



UNIVERSIDADE DE BRASÍLIA
FACULDADE DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

DANIELLE POMPEU NORONHA PONTES

**CLASSIFICAÇÃO DE CONTAS ENTRE BOTS E HUMANAS QUE DIVULGAM
ARTIGOS CIENTÍFICOS NO X: CONTRIBUIÇÕES PARA APRIMORAMENTO DOS
INDICADORES ALTMÉTRICOS**

BRASÍLIA - DF
2025

DANIELLE POMPEU NORONHA PONTES

**CLASSIFICAÇÃO DE CONTAS ENTRE *BOTS* E HUMANAS QUE DIVULGAM
ARTIGOS CIENTÍFICOS NO X: CONTRIBUIÇÕES PARA APRIMORAMENTO DOS
INDICADORES ALTMÉTRICOS**

Tese para defesa apresentada ao Programa de Pós-Graduação em Programa de Pós-Graduação em Ciência da Informação da Universidade de Brasília como requisito para a obtenção do título de Doutora em Ciência da Informação.

Orientador: Prof. Dr. João de Melo Maricato

BRASÍLIA - DF

2025

UNIVERSIDADE DE BRASÍLIA

PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

Ata Nº: 101

Aos treze dias do mês de outubro do ano de dois mil e vinte e cinco, instalou-se a banca examinadora de Tese de Doutorado da aluna Danielle Pompeu Noronha Pontes, matrícula 210002964. A banca examinadora foi composta pelos professores Dr. José Eduardo Santarem Segundo (Membro Titular externo, USP), Dra. Márcia Regina da Silva (Membro Titular externo, USP), Dr. Dalton Lopes Martins (Membro Titular Interno, UnB), Dr. Fernando Cesar Lima Leite (Suplente, UnB) e Dr. João de Melo Maricato (Orientador/Presidente). A discente apresentou o trabalho intitulado “CLASSIFICAÇÃO DE CONTAS ENTRE BOTS E HUMANAS QUE DIVULGAM ARTIGOS CIENTÍFICOS NO X: CONTRIBUIÇÕES PARA APRIMORAMENTO DOS INDICADORES ALTMÉTRICOS”.

Concluída a exposição, procedeu-se a arguição do(a) candidato(a), e após as considerações dos examinadores o resultado da avaliação do trabalho foi:

() Pela aprovação do trabalho;

(X) Pela aprovação do trabalho, com revisão de forma, indicando o prazo de até 30 (trinta) dias para apresentação definitiva do trabalho revisado;

() Pela reformulação do trabalho, indicando o prazo de (Nº DE MESES) para nova versão;

() Pela reprovação do trabalho, conforme as normas vigentes na Universidade de Brasília.

Conforme os Artigos 34, 39 e 40 da Resolução 0080/2021 - CEPE, o(a) candidato(a) não terá o título se não cumprir as exigências acima.

Dr. João de Melo Maricato
(Orientador/Presidente)

Dr. José Eduardo Santarem Segundo, USP
(Membro Titular externo)

Dra. Márcia Regina da Silva
(Membro Titular externo, USP)

Dr. Dalton Lopes Martins
(Membro Titular Interno, UnB)

Dr. Fernando Cesar Lima Leite
(Suplente, UnB)

Danielle Pompeu Noronha Pontes
(Doutorando)



Documento assinado eletronicamente por **Joao de Melo Maricato, Professor(a) de Magistério Superior da Faculdade de Ciência da Informação**, em 04/11/2025, às 11:00, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Jose Eduardo Santarem Segundo, Usuário Externo**, em 14/11/2025, às 09:27, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Dalton Lopes Martins, Professor(a) de Magistério Superior da Faculdade de Ciência da Informação**, em 14/11/2025, às 16:52, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Marcia Regina da Silva, Usuário Externo**, em 14/11/2025, às 17:17, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Danielle Pompeu Noronha Pontes, Usuário Externo**, em 25/11/2025, às 12:15, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



Documento assinado eletronicamente por **Elmira Luzia Melo Soares Simeao, Coordenador(a) da Pós-Graduação da Faculdade de Ciência da Informação**, em 03/12/2025, às 12:01, conforme horário oficial de Brasília, com fundamento na Instrução da Reitoria 0003/2016 da Universidade de Brasília.



A autenticidade deste documento pode ser conferida no site http://sei.unb.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **13350906** e o código CRC **1B96B0CC**.

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

PP814cc

POMPEU NORONHA PONTES, DANIELLE
CLASSIFICAÇÃO DE CONTAS ENTRE BOTS E HUMANAS QUE DIVULGAM
ARTIGOS CIENTÍFICOS NO X: CONTRIBUIÇÕES PARA APRIMORAMENTO
DOS INDICADORES ALTMÉTRICOS / DANIELLE POMPEU NORONHA
PONTES; orientador JOÃO DE MELO MARICATO. Brasília, 2025.
205 p.

Tese(Doutorado em Ciência da Informação) Universidade de
Brasília, 2025.

1. Altimetria. 2. Redes socias. 3. Detecção de bots. 4. X
- twitter. I. DE MELO MARICATO, JOÃO, orient. II. Título.

A Deus fonte de toda sabedoria e força.
Aos meus pais Ana, Batista que são fonte de inspiração. Minha
eterna gratidão pela base sólida que me deram a vida toda.
Ao Pontes, amor da minha vida, companheiro de sonhos e desafios,
obrigada por cada palavra, pensamento e oração.
Às minhas filhas amadas Sofia e Clarissa, a quem desejo ser fonte
de inspiração e de onde vem grande parte da minha força.
Para vocês, dedico não só esta tese, mas todos os meus dias.
Sem vocês, nada disso teria sentido.

AGRADECIMENTOS

A Deus, pela vida, pela saúde e pela sabedoria concedida em cada etapa desta caminhada acadêmica.

À minha família - meus pais, João Batista Filgueiras de Noronha e Ana Lúcia Pompeu Noronha, pelo amor incondicional, exemplo de ética e perseverança. Ao meu esposo Raimundo Pereira Pontes Filho, pelo apoio emocional, pela paciência, pelo incentivo diário e pelas orações. Às minhas filhas, Sofia e Clarissa de Noronha Pontes, minhas maiores incentivadoras, que celebram comigo cada conquista.

Ao meu orientador, Prof. Dr. João de Melo Maricato, pela orientação competente, por me dar a honra de ser sua orientanda, pelas inúmeras horas de orientação, por dividir comigo o seu saber, pelo rigor científico e pela generosidade intelectual que foram fundamentais para o desenvolvimento desta tese.

Aos membros da banca, Profa. Dra. Márcia Regina da Silva, Prof. Dr. Eduardo Santarém e Prof. Dr. Dalton Martins, pela leitura atenta e pelas valiosas contribuições, desde a qualificação, que enriqueceram significativamente este trabalho.

Aos amigos do programa de pós-graduação, especialmente Diana, Vanuza, Thais e Marcos, pelo companheirismo, pelo apoio mútuo e por toda palavra de incentivo ao longo desta jornada.

A Universidade do Estado do Amazonas (UEA) e aos professores do Núcleo de Computação da UEA que, generosamente, me permitiram os dois anos de licença para qualificação.

Aos amigos do Ocean, em especial ao Professor Rubelmar Neto que muito contribuiu para a validação do modelo apresentado; e aos demais professores e pesquisadores que, direta ou indiretamente, contribuíram para meu crescimento acadêmico.

A Universidade Federal do Amazonas (UFAM) e a Universidade de Brasília (UNB) que ofertaram o Dinter na pessoa da Profa. Dra. Célia Simonetti Barbalho que, desde muito antes, me incentivou a realizar o doutorado em Ciência da Informação.

Aos meus irmãos de sangue Raphael e Thiago Noronha, às minhas irmãs e aos meus irmãos em Cristo do “Mães que oram pelos Filhos” e da Equipe de Nossa Senhora (6F) por toda oração e incentivo.

Por fim, não poderia deixar de consagrar cada palavra escrita neste trabalho à Santíssima Trindade, Pai, Filho e Espírito Santo, pelas mãos generosas de Nossa Senhora. Que este trabalho, voltado ao estudo da desinformação na divulgação científica, possa ser um instrumento de reflexão e de transformação, reafirmando que a ciência só cumpre plenamente seu papel quando está comprometida com a verdade, a ética e o bem coletivo.

*“A língua dos sábios destila o conhecimento,
a boca dos insensatos fervilha de estupidez”.*

Provérbios 15,2

RESUMO

A disseminação automatizada de publicações científicas por meio de contas robotizadas em redes sociais representa um desafio à validade dos indicadores altmétricos. Esta pesquisa teve como objetivo desenvolver e validar métodos computacionais para a detecção de *bots* e a classificação automatizada de contas da plataforma X (antigo *Twitter*) com foco na divulgação científica. De natureza aplicada e abordagem qualiquantitativa, a pesquisa utilizou metodologia baseada no modelo *CRoss-Industry Standard Process for Data Mining* (CRISP-DM), combinando técnicas de aprendizado de máquina, curadoria de dados e análise estatística. Como principais resultados, foram desenvolvidos: (i) uma metodologia de inteligência artificial para classificação de contas, testando os algoritmos *k-nearest neighbors* (KNN), Árvore de decisão, *Random Forest* e *Extreme Gradient Boosting* (XGBoost), sendo este último o que apresentou melhor desempenho (acurácia de 94,29% e AUC de 0,9392); (ii) um *dataset* rotulado inédito, composto por 13.767 contas da rede X identificadas como humanas ou *bots*, construído por meio da integração de fontes reconhecidas da literatura e da validação manual rigorosa; e (iii) o protótipo SciBotScan, uma aplicação interativa desenvolvida em Python e hospedada em ambiente *web* (Streamlit), capaz de estimar a probabilidade de uma conta ser automatizada com base em 46 variáveis preditivas. Os resultados revelam que, embora os *bots* representem apenas 4% das contas analisadas, eles são responsáveis por 22% das postagens, com atividade média por conta significativamente superior à das contas humanas. Observou-se ainda que a presença e o impacto de *bots* variam entre áreas do conhecimento, com destaque para as divisões de Ciências Matemáticas, Computação e Ciências Físicas. A pesquisa conclui que a adoção de modelos especializados e bases de dados representativas é fundamental para aprimorar a confiabilidade dos indicadores altmétricos e mitigar distorções provocadas por automação, contribuindo para uma avaliação mais justa e transparente da produção científica em redes sociais.

Palavras-Chave: altmetria; redes sociais; inteligência artificial; detecção de *bots*; X - *twitter*.

ABSTRACT

The automated dissemination of scientific publications by fake accounts on social media poses a growing threat to the reliability of altmetric indicators. This research aimed to develop and validate computational methods for bot detection and automated classification of accounts on the X platform (formerly twitter) focused on scholarly communication. Characterized as applied research with a qualitative-quantitative approach, the study followed the *CRoss-Industry Standard Process for Data Mining* (CRISP-DM) methodological model, combining machine learning techniques, data curation, and statistical analysis. The main contributions of this thesis include: (i) an artificial intelligence methodology for account classification, evaluating the performance of k-nearest neighbors (KNN), Random Forest, and *Extreme Gradient Boosting* (XGBoost) algorithms the latter achieving the best results, with 94.29% accuracy and an AUC of 0.9392; (ii) a novel labeled dataset comprising 13,767 X accounts categorized as human or bot, built through the integration of recognized academic sources and rigorous manual verification; and (iii) the SciBotScan prototype, an interactive application developed in Python and deployed on a web platform (Streamlit), capable of estimating the likelihood that an account is automated based on 46 predictive features. The findings show that although bots represent only 4% of the analyzed accounts, they are responsible for 22% of the scientific posts, exhibiting a significantly higher posting frequency than human accounts. Additionally, the presence and influence of bots vary across fields of knowledge, with greater concentration in Mathematical Sciences, Computing, and Physical Sciences. The study concludes that the adoption of specialized models and representative datasets is essential to enhance the reliability of altmetric indicators and mitigate distortions caused by automation, thereby promoting more transparent and equitable evaluations of scientific impact on social media.

Keywords: altmetrics; social media; artificial intelligence; bot detection; X - twitter.

LISTA DE FIGURAS

Figura 1 - Ciclo de vida do processo de mineração de dados para classificação de contas X baseado na metodologia CRISP-DM.....	73
Figura 2 - Infográfico com as atividade da etapa de compreensão preparação de dados e modelagem.....	76
Figura 3 - Imagem da Tela do <i>Export Research Outputs</i> no Altmetric onde é realizado o download do arquivo.....	79
Figura 4 - Relatório de Classificação – XGBoost	130
Figura 5 - Interface do protótipo para classificação de contas (SciBotScan)	134
Figura 6 - Fluxograma de execução do protocolo de Revisão Sistemática de Literatura	190

LISTA DE GRÁFICOS

Gráfico 1 - <i>Summary plot</i> dos valores SHAP das <i>features</i> geradas	129
Gráfico 2 - Contas <i>Bots/Humanas</i> por Divisão.....	140
Gráfico 3 - <i>Posts</i> gerados por <i>bots/humanos</i> por Divisão	142
Gráfico 4 - Média de <i>posts</i> por artigos	154
Gráfico 5 - Média de <i>posts</i> por contas.....	155
Gráfico 6 - Comparação entre quantidade de contas, <i>posts</i> e artigos por divisão ..	156

LISTA DE QUADROS

Quadro 1 - Referências sobre uso de Web Scraping em pesquisas altmétricas	84
Quadro 2 - Conjuntos de dados utilizados como base para treinamento e avaliação da classificação de contas no X, com foco em perfis humanos, <i>bots</i> e categorias altmétricas	89
Quadro 3 - Quadro de síntese dos hiperparâmetros configurados no algoritmo XGBoost, apresentando os valores adotados e a respectiva descrição de sua função, no contexto da solução desenvolvida para a classificação de contas automatizadas no X.....	108
Quadro 4 - Indicadores de presença e engajamento em redes sociais por área do conhecimento	148
Quadro 5 - Média de menções por divisão de áreas de conhecimento, considerando contas, <i>posts</i> e artigos	152
Quadro 6 - Verificação das Hipóteses de Pesquisa: Fontes e Estratégias Analíticas	165
Quadro 7 - Resultados da busca bibliográfica (30/09 a 03/10/2024) sobre detecção de <i>bots</i> e métricas altmétricas nas bases consultadas	196
Quadro 8 - Resultados da aplicação dos critérios de exclusão ao protocolo de busca bibliográfica nas bases consultadas.....	197
Quadro 9 - Correspondência entre Passos Metodológicos, Bases de Dados e Algoritmos	202

LISTA DE TABELAS

Tabela 1 - Cáculo amostral de contas que postaram os artigos da amostra por divisão (área de conhemento) ForCode.....	91
Tabela 2 – Comparativo dos resultados com 05, 10,15, 20, 25, 30,35 ,40 45 e 46 variáveis preditoras (<i>features</i>).....	113
Tabela 3 - Desempenho dos modelos na identificação de <i>bots</i> na plataforma X, com métricas específicas para a classe <i>bot</i> e indicadores gerais de acurácia e robustez.....	123
Tabela 4 - Presença de Contas <i>bots</i> x Humanas	138
Tabela 5 - Quantidade de <i>posts</i> e contas <i>bot/humanos</i> por divisão.....	144
Tabela 6 - Quantidade de <i>posts</i> publicados por contas	158
Tabela 7 - Número de artigos publicados por contas	159
Tabela 8 - Quantidade de artigos por contas	161
Tabela 9 - Quantidade de <i>post</i> por contas	162
Tabela 10 - Quantidade de divisões por contas	163

LISTA DE ABREVIATURAS E SIGLAS

AAS	<i>Altmetric Attention Score</i>
ACM	<i>Association for Computing Machinery</i>
ACMDL	<i>ACM Digital Library</i>
AHCI	Índice de Citação de Artes e Humanidades
AM	Aprendizagem de Máquina
API	Interface de Programação de Aplicações
AUC	Curva característica de operação do receptor
BERT	Bidirectional Encoder Representations from Transformers
CNetS	<i>CompleXNetworks and Systems Research</i>
CNNs	Redes Neurais Convolucionais
INFOCOM	<i>Conference on Computer Communications</i>
CRISP-DM	<i>CRoss-Industry Standard Process for Data Mining</i>
DL	<i>Deep Learning</i> / Aprendizado profundo
DM	<i>Data Mining</i>
DOI	<i>Digital object identifier</i>
DT	<i>Decision Trees</i>
Flesch	Frequência de vocábulos científicos e índice médio de legibilidade
FN	Falso Negativo
ForCode	<i>Fields of Research</i>
FP	Falso Positivo
GANs	Redes Adversariais Generativas
GCN	Redes Convolucionais de Grafos
GNN / GNNs	Redes Neurais Profundas/ <i>Graph Neural Networks</i>
IA	Inteligência Artificial
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
INFOCOM	<i>Conference on Computer Communications</i>
ICDM	<i>International Conference on Data Mining</i>
IUNI	<i>Network Science Institute</i>
LLM	<i>Large Language Models</i>
KNN	<i>K Nearest Neighbor</i>

LR	<i>Logistic Regression</i>
LSTMs	Redes de Memória de Longo Prazo
ML	<i>Machine Learning</i>
NB	<i>Naïve Bayes</i>
NN	<i>Neural Networks</i>
OSoMe	Observatório de Mídias Sociais
PCA	Análise de Componentes Principais
PGD	Plano de Gestão de Dados
PLN	Processamento de Linguagem Natural
PRISMA	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</i>
RF	<i>Random Forest</i>
RNN	<i>Recurrent Neural Network</i>
SCIE	Índice de Citações de Ciência Expandido
SHAP	<i>Shapley Additive exPlanations</i>
SSCI	Índice de Citação de Ciências Sociais
SVM	<i>Support Vector Machine</i>
VN	<i>Verdadeiro Negativo</i>
VP	Verdadeiro Positivo
WOS	<i>Web of Science</i>
XGBoost	<i>Extreme Gradient Boosting</i>

SUMÁRIO

1	INTRODUÇÃO	20
1.1	JUSTIFICATIVA	25
1.2	OBJETIVOS	26
1.2.1	Objetivo Geral	26
1.2.2	Objetivos Específicos	26
1.3	HIPÓTESES DE PESQUISA.....	27
1.4	ESTRUTURA DO TRABALHO	28
2	REFERENCIAL TEÓRICO-CONCEITUAL	30
2.1	<i>BOTS EM MÍDIAS SOCIAIS E SUAS IMPLICAÇÕES PARA A AVALIAÇÃO CIENTÍFICA.....</i>	30
2.1.1	Definição de <i>bots</i> , ciborgues e contas automatizadas	31
2.1.2	Tipologias de contas <i>bots</i>	31
2.1.3	Impactos dos <i>bots</i> sobre a comunicação científica, engajamento artificial, e integridade das métricas.....	33
2.1.4	Questões éticas do uso da automação na divulgação científica	33
2.2	ALTMETRIA E REDES SOCIAIS: O CASO DO X.....	36
2.2.1	Fundamentos da altmetria	36
2.2.2	O papel do X como fonte de dados altmétricos	39
2.2.3	Engajamento de diferentes tipos de contas e a produção de indicadores altmétricos	40
2.2.4	Distorções causadas por <i>bots</i> nos scores altmétricos	42
2.3	TÉCNICAS DE DETECÇÃO DE <i>BOTS</i> : ABORDAGENS CLÁSSICAS E AVANÇADAS	44
2.3.1	Métodos tradicionais baseados em atributos de conta e comportamento	44
2.3.2	Aplicações recentes de Inteligência Artificial (IA) e aprendizado profundo na pesquisa científica.....	46
2.3.3	Literatura sobre detecção de <i>bots</i> em contextos científicos	48
2.3.4	Estudos experimentais da literatura sobre classificação de <i>bots</i> utilizando o XGBoost	53
2.3.5	Detecção de <i>bots</i> com o Botometer	55
2.4	MODELAGEM SUPERVISIONADA E ENGENHARIA DE FEATURES NA DETECÇÃO DE <i>BOTS</i> NO X	57

2.4.1	Principais algoritmos supervisionados: KNN, Árvore de Decisão, Random Forest e XGBoost.....	57
2.4.1.1	<i>K-Nearest Neighbors (KNN)</i>	57
2.4.1.2	<i>Árvore de decisão.....</i>	58
2.4.1.3	<i>Random Forest.....</i>	59
2.4.1.4	<i>Extreme Gradient Boosting (XGBoost).....</i>	60
2.4.2	Características/features utilizadas para treinamento de modelos de IA	61
3	METODOLOGIA E PROCEDIMENTOS METODOLÓGICOS	70
3.1	<i>METODOLOGIA.....</i>	70
3.2	<i>PROCEDIMENTOS METODOLÓGICOS</i>	72
3.3	<i>COMPREENSÃO DE NEGÓCIOS</i>	74
3.4	<i>COMPREENSÃO E PREPARAÇÃO DOS DADOS.....</i>	75
3.4.1	II.a. Definição e seleção de amostra de artigos	77
3.4.1.1	<i>II.a.1 Download dos artigos científicos de cada Grupo do ForCode no Altmetrics.....</i>	77
3.4.1.2	<i>II.a.2 Cálculo amostral de número de artigos e sua seleção randômica</i>	80
3.4.2	II.b Construção do dataset de treinamento	81
3.4.2.1	<i>II.b.1 Coleta dos posts de cada artigo da seleção randômica</i>	81
3.4.2.2	<i>II.b.2 Classificação preliminar com base nos Datasets da literatura</i>	85
3.4.2.3	<i>II.b.3 Cálculo amostral do número de contas e seleção randômica das contas por divisão do ForCode</i>	90
3.4.2.4	<i>II.b.4 Classificação manual do dataset de treinamento</i>	92
3.4.3	II.c Processo de desenvolvimento do modelo de inteligência artificial	96
3.4.3.1	<i>(II.c.1) Engenharia de features</i>	99
3.4.3.2	<i>(II.c.2) Modelagem do Modelo de Aprendizagem de Máquina para classificação de contas do X baseado no XGBOOST</i>	105
3.4.3.3	<i>(II.c.3) Processo de treinamento e otimização do modelo XGBOOSTS.....</i>	109
3.4.3.4	<i>(II.c.4) Classificação das contas</i>	115
4	RESULTADOS E DISCUSSÕES	118
4.1	<i>CONSTRUÇÃO DE UM DATASET ROTULADO E REPRESENTATIVO DE CONTAS DA PLATAFORMA X</i>	118

4.2	COMPARAÇÃO DO DESEMPENHO DE DIFERENTES ALGORITMOS DE APRENDIZADO DE MÁQUINA	120
4.3	MODELO DE INTELIGÊNCIA ARTIFICIAL OTIMIZADO, COM BASE NO ALGORITMO XGBOOST, PARA CLASSIFICAR CONTAS	125
4.3.1	Importância das <i>Features</i>	126
4.3.2	Avaliação do Modelo Gerado	130
4.4	PROVA DE CONCEITO DO MODELO COMPUTACIONAL E PROTÓTIPO PARA AUTOMAÇÃO DE CLASSIFICAÇÃO DE CONTAS DO X.....	133
4.4.1	Tecnologias e ferramentas utilizadas	135
4.4.2	Desempenho do modelo	136
4.4.3	Variáveis de entrada na interface do usuário (via Streamlit)	136
4.5	PADRÕES DE COMPORTAMENTO E ENGAJAMENTO DE CONTAS HUMANAS E BOTS	138
4.5.1	Padrões comportamentais, segmentados por área do conhecimento, entre contas do X, classificadas por tipo de perfil	139
4.5.2	Caracterização da amostra de estudo e sua atuação na plataforma X	146
4.5.2.1	<i>Caracterização da coleta primária de dados</i>	146
4.5.2.2	<i>Caracterização da amostra randômica de artigos</i>	147
4.5.2.3	<i>Caracterização dos posts coletados</i>	151
4.5.2.4	<i>Caracterização das contas que comentaram os artigos da seleção randômica</i>	157
4.5.2.5	<i>Caracterização do comportamento das contas</i>	157
4.5.2.6	<i>Caracterização da amostra randômica de contas</i>	160
4.6	HIPÓTESES DE PESQUISA: STATUS DE VERIFICAÇÃO, EVIDÊNCIAS E TIPO DE ANÁLISE	164
5	CONCLUSÕES	167
	REFERÊNCIAS	172
	APÊNDICE A - PROCEDIMENTOS METODOLÓGICOS ADOTADOS PARA A REVISÃO SISTEMÁTICA DE LITERATURA	190
	APÊNDICE B - PLANO DE GESTÃO DE DADOS (PGD)	198

1 INTRODUÇÃO

Existe uma diversidade crescente de “ecossistemas sociais” que sustentam o uso acadêmico das mídias sociais. Os cientistas estão aproveitando o poder desse tipo de mídia para acelerar o ritmo em que eles estão desenvolvendo e compartilhando conhecimento, tanto no âmbito da comunidade científica como no âmbito do público em geral. Neste contexto, as mídias sociais em geral e o X¹ (antigo *Twitter*) em particular influenciam o ciclo de vida da publicação científica (Darling *et al.*, 2013). Para estes autores, esta rede social traz benefícios ao aumentar as conexões às redes acadêmicas, traz desenvolvimento mais rápido de ideias e discussões entre os pares e amplia a disseminação e discussão do conhecimento científico dentro e fora da academia.

Como consequência deste movimento crescente, as comunidades científicas começaram a adotar ativamente plataformas de mídia social para identificar o impacto e as influências da literatura acadêmica em públicos diversos, havendo interesse em saber se elas são capazes de oferecer indicadores complementares aos indicadores bibliométricos, para avaliação científica (Piem; Hemminger, 2010). Desta forma, a crescente utilização de mídias sociais tem despertado, nas comunidades científicas, o interesse em rastrear o impacto acadêmico por meio dos seus novos indicadores (Araujo; Caran; Souza, 2016).

Avaliar o resultado da pesquisa e medir o seu impacto é um componente essencial na evolução da ciência. O sistema atual de avaliação científica privilegia os indicadores bibliométricos quantitativos, sobretudo por meio da contagem de citações. Esse modelo beneficia os países da América do Norte e Europa, porque as bases de dados tradicionalmente utilizadas como fontes para a avaliação científica sub-representam a produção científica de países em desenvolvimento. Em contrapartida, os indicadores advindos das mídias sociais (indicadores altmétricos) têm sido entendidos como uma forma complementar de monitoramento, que podem auxiliar na medição mais ampla do impacto dos trabalhos acadêmicos, de modo a beneficiar os países não centrais (Alperin, 2013).

A altmetria, um termo derivado do Latim “*alternativus*” (escolha entre duas opções) e “*metric*” (métrica em inglês), representa uma abordagem inovadora na

¹ Criada como *Twitter* em 2006, foi adquirida por Elon Musk em 2022 e, desde então, passou a se chamar X.

avaliação do impacto da produção científica. Diferentemente das métricas tradicionais, como o fator de impacto de revistas acadêmicas, que se concentram principalmente em citações em publicações especializadas, a altmetria visa capturar a disseminação e o alcance de uma pesquisa em plataformas on-line e em redes sociais.

Essa métrica complementar reflete a mudança na forma como o conhecimento científico é comunicado e consumido na era digital. Em vez de depender exclusivamente de indicadores baseados em citações acadêmicas, a altmetria leva em consideração uma variedade de fontes de dados on-line para avaliar impactos mais amplos de resultados de pesquisas acadêmicas. Entre os elementos considerados, estão menções em *blogs*, postagens em redes sociais (como X e Facebook), compartilhamentos em plataformas como *Mendeley* e *Figshare*, além de cobertura midiática.

Segundo Tunger, Clermont e Meier (2018), o desenvolvimento de indicadores altmétricos se iniciou a partir de questionamentos em relação aos indicadores bibliométricos tradicionais, os quais não refletiam adequadamente o significado científico e social do trabalho científico na era da Internet. A circulação de publicações científicas em sites de notícias, em *blogs* científicos e em redes sociais, aliada às emergentes ferramentas de rastreamento dessas publicações na web, contribuíram para o desenvolvimento desses indicadores. Uma vez que os indicadores altmétricos estão se consolidando como uma nova forma de mensuração da atenção social da ciência, estudos como o de Tunger, Clermont e Meier (2018) passaram a investigar até que ponto eles podem ser utilizados em avaliações científicas e em diferentes campos de aplicação.

A velocidade de comunicação provocada pela conectividade através da web social remodela o panorama da comunicação científica, fomentando processos comunicativos mais ágeis, impulsionando o uso de indicadores altmétricos oriundos de mídias e redes sociais. Nesses ambientes, o acesso à informação é mais democrático e desburocratizado em relação às bases de dados tradicionais. Neste sentido, Thelwall (2020) sugere que é necessário compreender se os indicadores altmétricos podem ajudar a avaliar os resultados acadêmicos, uma vez que eles estão amplamente disponíveis. Para Alhoori *et al.* (2014), a altmetria tem potencial para medir o impacto mais amplo das atividades de pesquisa.

Contudo, o ambiente das mídias e redes sociais, caracterizada pela Web 2.0, ao mesmo tempo que possibilitou trocas de informações de maneira mais ágil e

democrática, trouxe dilemas que não existiam anteriormente, como, por exemplo, o uso de robôs automatizados ou comumente chamados de *bots*. Os *bots* não são necessariamente ruins, pois também podem ser utilizados para divulgar informações verdadeiras e relevantes. Mesmo assim, dependendo do contexto de uso, eles podem ser considerados prejudiciais, pois podem interferir nos significados dos dados, nos indicadores e nos processos de tomada de decisão. Um exemplo disso é a produção de indicadores feita a partir dos rastros deixados por agentes humanos e não humanos em redes sociais. A existência de *bots* é uma das dificuldades de se utilizar indicadores de mídias sociais para o estudo e a avaliação dos impactos das produções científicas.

No atual cenário das mídias sociais e da inteligência artificial conversacional, os incentivos para o desenvolvimento de softwares, capazes de imitar interações humanas, têm se intensificado. *Bots* sociais, por exemplo, operam de forma imperceptível nos ecossistemas digitais, e, embora alguns tenham propósitos benignos, muitos são utilizados com intenções maliciosas, como persuadir, difamar ou enganar, frequentemente motivados por interesses políticos ou econômicos.

No domínio científico, essa problemática assume contornos ainda mais delicados. Plataformas como o X exercem influência significativa sobre o ciclo de vida da publicação científica, ao mesmo tempo em que podem favorecer a circulação de ideias e o engajamento entre pares, também apresentam riscos à integridade dos indicadores de impacto, especialmente quando manipuladas por *bots* com o objetivo de inflar artificialmente a visibilidade de determinados conteúdos (Darling *et al.*, 2013).

Os impactos da ciência (dentro da própria academia ou na sociedade em geral), as correlações entre citação e menção nas mídias sociais, as coberturas de menções, os tipos de comunidade de atenção etc. não terão o mesmo significado se forem produzidos com dados de artigos científicos que contenham *bots* e/ou agentes humanos. No caso do X (Didegah; Mejlgård; Sørensen 2018), os *bots* automatizados aumentam a incerteza sobre os padrões de divulgação científica nessa rede social e levantam dúvidas em relação a confiabilidade dos seus dados.

Os *bots* são algoritmos que produzem conteúdos e interagem com os usuários. Esses mecanismos são responsáveis por uma proporção significativa das atividades on-line (Ferrara *et al.*, 2016). Eles são agentes de software, que utilizam dados e informações sociais, com a finalidade de aumentar a sua capacidade de

relacionamento, interação, persuasão e influência sobre atores humanos. Isso leva à necessidade de atenção para o papel desses mecanismos nos processos e ciclos informacionais (Nunes, 2020). Os dados altmétricos também não estão imunes à atuação destes *bots*.

A compra de um *bot*² para promover o próprio produto acadêmico pode ser visto como inapropriado no contexto acadêmico, porém, provavelmente, é aceitável que sejam utilizados para promover, por exemplo, automaticamente a conscientização sobre importantes questões ambientais. Assim, a existência desse tipo de serviço e a prevalência de postagens feitas por *bots* levantam questões sobre o uso de postagens de mídia social para avaliar a relevância social da ciência (Crotty, 2014), bem como acerca dos limites aceitáveis dessa forma de divulgação e do marketing científico como um todo. Outro exemplo de divulgação de produtos acadêmicos nas redes e mídias sociais é o realizado pelas revistas científicas, que comumente promovem seus artigos nesses ambientes. Independentemente dessas questões, a possibilidade de identificação de *bots* pode facilitar análises por tipos de contas que se tem interesse, diminuindo, ou controlando melhor, os vieses.

Segundo Aljohani; Fayoumi; Hassan (2020), é essencial analisar a influência social destes *bots* nos dados altmétricos. Os autores consideram fundamental investigar como esses mecanismos disseminam conteúdo e difundem informações e a porcentagem do conteúdo acadêmico produzido por eles nas mídias sociais. Diversas são as redes e mídias sociais existentes na atualidade e, por consequência, existe grande número de indicadores altmétricos que podem ser produzidos por meio delas. Dentre elas, o X tem grande importância, pois é a mídia com maior cobertura de indicadores extraídos de agregadores como a plataforma altmetric.com. Um estudo de 2017, com mais de 1,1 milhão de publicações, demonstra a relevância dos dados do X para altmetria: a rede social responde por mais de 90% de todas as menções sociais registradas (Hassan *et al.*, 2017).

Um dos estudos que tiveram foco na investigação do impacto destes *bots* na altmetria foi publicado por Haustein *et al.* (2016). Esses autores buscaram identificar a influência dos *bots* em índices altmétricos. Ao analisarem uma amostra aleatória de 800 contas de usuários do X, constataram que 8% dos dados altmétricos foram completamente afetados por contas de *bot* e 5% parcialmente. O estudo mostrou que

² O termo refere-se a contas de mídia social automatizadas, executadas por um algoritmo projetado para fazer *posts* sem intervenção humana.

9% dos *tweets* para artigos do *arXiv*³ foram gerados por contas automatizadas. Descobriram, também, que as contas automatizadas que tuitam sobre temas acadêmicos se comportam de maneira diferente das contas gerais de *bots* do X. As descobertas dos autores revelaram que as contas automatizadas do X publicam, em média, 4,6 a 7,1 *tweets* por dia, enquanto contas não automatizadas produzem, em média, 2,2 *tweets* por dia. Além disso, os autores identificaram que os critérios de *tweeting*, utilizados pelos *bots*, são geralmente aleatórios e não qualitativos (eles podem, por exemplo, tuitar aleatoriamente frases predefinidas, como provérbios).

Aljohani; Fayoumi; Hassan (2020) estudaram o comportamento e a influência desses mecanismos de mídias sociais em dados altmétricos, aplicando algumas técnicas de análises baseadas em inteligência artificial no X. As análises revelaram que os *bots* influenciaram 87% dos *tweets* relacionados aos dados altmétricos, interferindo fortemente neste tipo de métrica. Os autores aplicaram uma técnica de rede convolucional de grafos para classificação de *bots* em um conjunto de dados altmétricos. Eles chegaram a uma precisão/acurácia de 71% e o F1-score de 0,67 (esse score refere-se à média harmônica entre a precisão e o *recall*, que está muito mais próxima dos menores valores do que uma média aritmética simples). Os autores esclarecem que não encontraram um conjunto de dados rotulados por humanos para melhorar seu modelo (baseado em inteligência artificial). Assim, concluíram que a diferença entre contas de *bots* e de humanos não são claras, sendo complexa a detecção até mesmo por humanos.

Hassan *et al.* (2017), utilizando-se de outra técnica, também realizou análises no contexto de dados da altmetria no X. A pesquisa foi baseada em palavras-chave no X com as contas mais ativas (que tuitaram pelo menos 1.000 vezes). Os autores relataram que entre 2.043 contas, 248 (12%) foram identificadas como automatizadas, enquanto 305 (15%) foram identificadas como contas de editores ou periódicos (também automatizadas). Assim, 1.795 (88%) das contas encontradas foram classificadas como sendo de humanos e 27% como *bots*.

³ Repositório temático de arquivos *preprints* de artigos científicos nos campos da Matemática, Física, Ciência Da Computação, Biologia Quantitativa E Estatística.

1.1 JUSTIFICATIVA

A detecção de *bots* e contas falsas em plataformas de redes sociais ainda é um dos principais desafios para os investigadores de segurança cibernética (Thuraisingham, 2020). Considerando a carência de estudos sobre a interferência e detecção de *bots* no contexto da Altmetria no X, a realização desta tese justifica-se pela necessidade citada em Aljohani; Fayoumi; Hassan, (2020) e Hassan *et al.* (2017), de aprimorar a precisão e a confiabilidade dos indicadores altmétricos e garantir uma avaliação mais precisa do impacto acadêmico nas redes sociais.

A relevância desta pesquisa reside na necessidade de enfrentar os desafios impostos pelos vieses e pelas fragilidades induzidas por *bots*, os quais representam obstáculos significativos para a utilização efetiva da altmetria na avaliação e no monitoramento da pesquisa científica. Ao abordar tal questão, esta investigação busca aprimorar a compreensão dos impactos dessas distorções e desenvolver estratégias para mitigá-las, garantindo, assim, uma avaliação mais precisa e abrangente do impacto da pesquisa científica na comunidade acadêmica e na sociedade em geral.

Em resumo, essa pesquisa pode contribuir para o avanço da avaliação científica e para uma melhor compreensão do papel das mídias sociais na comunicação acadêmica e na compreensão dos vieses e das fragilidades causadas pelos *bots* que dificultam o uso pleno da altmetria para a avaliação e o monitoramento da pesquisa científica.

Segundo Orabi *et al.* (2020), algumas das lacunas identificadas nos trabalhos desenvolvidos para detecção de *bots* em redes sociais consiste no fato de que raramente são usados métodos que não seja o aprendizado de máquina supervisionado, além disso, a maioria dos conjuntos de dados públicos não são precisos ou grandes o suficiente. Os autores reforçam que sistemas integrados e detecção em tempo real são necessários para munir os usuários legítimos de ferramentas que possam distinguir contas automatizadas de não automatizadas.

Conforme descritos pelos autores em Balestrucci *et al.* (2019) e Minnich *et al.* (2017), a geração de contas de *bot* muda continuamente ao longo do tempo e, adicionalmente, algumas das contas fornecidas já foram suspensas pelo X. Assim, é necessário um pré-processamento para melhorar a usabilidade desta grande coleção de conjuntos de dados, removendo os identificadores daquelas contas que já foram removidas pelo X. Este aspecto também é crítico, uma vez que vários detectores de

bots anteriores não foram atualizados com as novas tendências e com os recursos que as contas de *bots* podem ter atualmente, de modo que não são mais tão confiáveis como costumavam ser (Martin-Gutierrez *et al.*, 2021).

1.2 OBJETIVOS

Esta seção tem como finalidade apresentar os objetivos desta pesquisa, os quais estão organizados em objetivo geral e objetivos específicos, conforme descrito a seguir.

1.2.1 Objetivo Geral

Desenvolver uma abordagem computacional, baseada em inteligência artificial, para a detecção de *bots* e a classificação automatizada de contas que divulgam publicações científicas na plataforma X, com o objetivo de reduzir distorções nos indicadores altmétricos e fortalecer sua aplicabilidade na avaliação da produção científica.

1.2.2 Objetivos Específicos

Mais especificamente, tem-se como objetivo:

- a. Construir um *dataset* rotulado e representativo de contas da plataforma X, classificadas como humanas ou automatizadas, com base em validação manual e integração de fonte confiável;
- b. Comparar o desempenho de diferentes algoritmos de aprendizado de máquina para detecção de *bots* em dados altmétricos, avaliando métricas como acurácia, AUC e F1-score na tarefa de classificação de contas;
- c. Desenvolver um modelo de inteligência artificial otimizado, com base no algoritmo *Extreme Gradient Boosting* (Xgboost), para classificar contas segundo sua natureza (humana ou *bot*), utilizando um conjunto de variáveis preditoras (*features*) selecionadas;
- d. Propor uma prova de conceito do modelo computacional e protótipo para automação de classificação de contas do X, com atuação altmétrica;

e. Identificar padrões de comportamento e engajamento, diferenciando contas humanas e *bots* por área do conhecimento, por volume de postagens e por participação altmétrica.

1.3 HIPÓTESES DE PESQUISA

Diante da complexidade do fenômeno investigado — a detecção de contas automatizadas no ecossistema altmétrico da plataforma X —, a presente pesquisa se propõe a investigar a atuação de contas automatizadas no compartilhamento de publicações científicas na plataforma X, com foco na detecção automatizada dessas contas, *bots* e humanas, e no impacto que exercem sobre os indicadores altmétricos.

Com base nos objetivos delineados, foram formuladas hipóteses de pesquisa destinadas a orientar: a delimitação dos percursos investigativos, a construção do dataset, a definição metodológica e a interpretação dos resultados. Cada hipótese reflete uma expectativa teórica derivada da revisão da literatura e direciona aspectos específicos da análise.

Na sequência, são apresentadas as hipóteses que sustentam esta investigação, organizadas de acordo com os objetivos específicos do estudo.

H01 – XGBoost apresenta desempenho superior aos demais modelos: acredita-se que o algoritmo XGBoost, por suas características técnicas, obtenha os melhores resultados na classificação de contas.

H02 – Bots postam mais que humanos: pressupõe-se que contas *bots* são responsáveis por um volume desproporcionalmente elevado de postagens, em relação às contas humanas.

H03 – A atuação de bots varia por área do conhecimento: parte-se da premissa de que o comportamento e a incidência de *bots* não são uniformes, apresentando variações conforme a área disciplinar.

H04 – Bots atuam em um maior número de áreas temáticas do que humanos: supõe-se que *bots* tendem a operar em várias áreas do conhecimento simultaneamente, em contraste com humanos, que têm atuação mais focada.

As hipóteses formuladas neste estudo serão retomadas ao longo das seções seguintes: na metodologia, ao explicitar as estratégias de verificação adotadas; nos

resultados, com base nas evidências empíricas obtidas; e, por fim, na conclusão, quando são discutidas em função dos achados e das implicações da pesquisa.

1.4 ESTRUTURA DO TRABALHO

A presente tese está estruturada em cinco seções, além das referências e de dois apêndices, organizados de modo a garantir a coerência, o desenvolvimento lógico e articulado dos conteúdos abordados. A primeira seção apresenta a introdução do tema, expondo o contexto da pesquisa, a justificativa teórica e prática que a fundamenta, bem como a definição dos objetivos gerais e específicos. Finaliza-se com uma breve descrição da estrutura do trabalho, oferecendo ao leitor uma visão abrangente de seu desenvolvimento.

A seção 2 é dedicado à fundamentação teórico-conceitual, contemplando a revisão sistemática da literatura sobre *bots* em mídias sociais, com ênfase na plataforma X e nos seus impactos sobre os indicadores altmétricos, apresentando a síntese dos resultados encontrados e a identificação de lacunas na literatura.

Além disso, discute-se a caracterização de *bots* em mídias sociais, com destaque para suas diferentes tipologias, implicações éticas e o emprego de técnicas de aprendizagem de máquina para sua detecção, com especial atenção a algoritmos como *K-Nearest Neighbors* (KNN), Árvore de Decisão, Floresta Randômica, *Botometer* e *Extreme Gradient Boosting* (XGBoost). Por fim, são analisados os fundamentos da altmetria, a vulnerabilidade dos indicadores altmétricos à ação de contas automatizadas e o comportamento de usuários na disseminação de publicações científicas.

São discutidas definições e tipologias de contas automatizadas, implicações éticas, fundamentos da altmetria, distorções causadas por *bots* em métricas científicas, bem como as principais abordagens e algoritmos aplicados à detecção de *bots*, com destaque para o XGBoost.

A seção 3 descreve de maneira detalhada a metodologia desenvolvida e os procedimentos metodológicos adotados para operacionalização prática em uma aplicação real, descrevendo a execução de cada etapa definida na metodologia, tais como a aplicação das técnicas de engenharia de *features*, pré-processamento de dados e treinamento supervisionado do modelo baseado no algoritmo XGBoost,

culminando na construção de um protótipo de classificação automatizada. Essa seção inclui as etapas de compreensão do domínio de pesquisa, preparação e tratamento dos dados, delineamento amostral, construção do *dataset* de treinamento e desenvolvimento do modelo de inteligência artificial, incluindo o *download* de bases de classificação temáticas, a realização de cálculos amostrais, a seleção randômica de artigos e contas, a coletas de *posts*, a fusão dos *datasets* e a construção do conjunto de treinamento para a modelagem de inteligência artificial.

A seção 4 expõe os resultados e realiza a discussão crítica dos achados, organizados conforme os objetivos específicos traçados. São apresentadas as análises e discussões oriundas da comparação entre diferentes metodologias de classificação, da análise dos conjuntos de dados de treinamento, da construção de uma base confiável de perfis rotulados, do desenvolvimento. Além de realizar a avaliação de um modelo de inteligência artificial para a classificação automática de contas, a investigação da distribuição de perfis na plataforma X e a identificação de padrões comportamentais segmentados por áreas do conhecimento. Ademais, propõe-se, como prova de conceito, um modelo computacional e um protótipo de automação para a classificação de contas com atuação altmétrica.

Por fim, a seção 5 apresenta as conclusões da pesquisa, sintetizando as principais contribuições científicas e práticas alcançadas, externando as limitações do estudo e propondo direções para investigações futuras. A tese é complementada pela seção de referências bibliográficas, que reúne as fontes consultadas ao longo do trabalho, e por apêndices que ampliam a transparência e a reproduzibilidade da pesquisa.

O Apêndice A reúne o protocolo metodológico utilizado na condução da Revisão Sistemática da Literatura (RSL), bem como o detalhamento do processo de análise sistemática desenvolvido. O Apêndice B, por sua vez, apresenta o Plano de Gestão de Dados (PGD), no qual são descritas as estratégias adotadas para a organização, preservação, disseminação e reutilização dos dados gerados durante a investigação.

2 REFERENCIAL TEÓRICO-CONCEITUAL

Esta seção apresenta os fundamentos teóricos e conceituais que sustentam esta pesquisa sobre a detecção de *bots* na plataforma X. A proposta é construir um panorama que une perspectivas da Ciência da Computação e da Ciência da Informação, articulando dimensões técnicas e informacionais associadas ao uso de contas automatizadas em redes sociais e seus impactos sobre a altmetria.

A partir dessa perspectiva interdisciplinar, a seção se organiza para explorar a interface entre altmetria, detecção de *bots* e classificação de contas na plataforma X, foco central da investigação. São discutidos os fundamentos da altmetria e suas distinções em relação à bibliometria tradicional, destacando o papel do X como principal fonte de dados altmétricos.

Em seguida, aborda-se o impacto das contas automatizadas sobre os indicadores de atenção on-line, considerando desde a distorção de métricas até os desafios éticos relacionados ao uso de automação na comunicação científica.

Por fim, examinam-se as principais abordagens e técnicas de detecção de *bots*, com especial atenção para o uso de aprendizado de máquina e de métodos de inteligência artificial, oferecendo um panorama crítico da literatura e dos instrumentos aplicados à identificação de contas automatizadas no contexto da ciência em rede.

2.1 BOTS EM MÍDIAS SOCIAIS E SUAS IMPLICAÇÕES PARA A AVALIAÇÃO CIENTÍFICA

Para compreender a atuação de *bots* em mídias sociais e suas implicações na avaliação científica, é fundamental delinejar os conceitos-chave relacionados às contas automatizadas, suas tipologias e formas de operação. As subseções a seguir apresentam, de forma articulada, uma definição conceitual de *bots*, ciborgues e contas automatizadas, descrevem as classificações mais comuns encontradas na literatura, discutem os impactos dessas contas sobre o ecossistema da comunicação científica — incluindo a geração de engajamento artificial e a distorção de métricas altmétricas — e abordam as principais questões éticas associadas ao uso de automação na divulgação de resultados científicos. Esses elementos fornecem a base teórico-

conceptual necessária para a análise crítica dos efeitos dessas tecnologias no campo da altmetria.

2.1.1 Definição de *bots*, ciborgues e contas automatizadas

Os "*bots*" são programas de computador que desempenham tarefas automatizadas em redes sociais a uma velocidade superior à dos seres humanos. Uma definição mais precisa dos *bots* em redes sociais pode ser encontrada em Benkler *et al.* (2017), que os descrevem como "um *software* de computador que gera conteúdo automaticamente e interage com usuários de mídias sociais para replicar e possivelmente modificar seu comportamento".

Por sua vez, os ciborgues são contas que apresentam uma mistura de comportamento humano e de *bot* (ou seja, trata-se de uma conta do X que publica *posts* automatizados e criados por uma pessoa). Para Haustein *et al.* (2016), todos os três tipos de contas - humano, *bot* e ciborgue - podem ser iniciados por uma pessoa (por exemplo, cientista, estudante, professor, profissional, leigo interessado e assim por diante) ou por uma organização (por exemplo, sociedade científica, editora, periódico, universidade, grupos de pesquisa, organização financiadora, jornal e assim por diante).

2.1.2 Tipologias de contas *bots*

A popularidade de plataformas de mídia social como o X levou à proliferação de *bots* (contas automatizadas), criando oportunidades e desafios na disseminação de informações. Preocupados com a influência desses mecanismos, muitas pesquisas sobre a identificação de perfis de *bots* concentraram-se principalmente em *bots* maliciosos.

De acordo com a classificação de Benkler *et al.* (2017), existem quatro categorias principais de *bots* de mídia social: *spambots*, *bots* sociais, *bots sybil* e ciborgues. Oentaryo *et al.* (2016) classificaram muitos *bots* como benignos e, a partir deste achado, propuseram uma categorização mais ampla de *bots* com base em seus comportamentos. Outros trabalhos como Chu *et al.* (2012), Gorwa e Guilbeault (2020) e Aguilera *et al.* (2023) apresentam classificações de categorias de *bots* de mídia

social, diferenciando-os pelo seu comportamento, utilidade ou forma de operar. Em resumo, as principais classificações encontradas foram:

- **Bots Sybil:** identidades anônimas usadas para gerar um impacto significativo (Gorwa; Guilbeault, 2020; Benkler *et al.*, 2017).

- **Consumer Bots:** *bots* cujo objetivo principal é obter conteúdo de diversas fontes e fornecer serviços de atualização, como boletins meteorológicos, horóscopo ou notícias para consumo ou uso pessoal. Seu conteúdo geralmente é benigno.

- **Bots Híbridos ou Ciborgues:** esta classificação corresponde a contas de usuários que são gerenciadas por usuários humanos com a ajuda ou suporte de um *bot* ou vice-versa (Chu *et al.*, 2012a). Os *bots* ciborgues representam uma categoria híbrida, sendo meio humanos e meio *bots* e operam na interface entre pessoas e *bots* (Benkler *et al.*, 2017).

- **Crawlers e scrapers:** são *bots* que foram projetados para indexar e arquivar dados de páginas da web, por exemplo, para disponibilizar os dados dos sites através de mecanismos de busca.

- **Chatbots:** são programas de computador que conversam com pessoas em tempo real utilizando linguagem natural. Frequentemente usado por empresas para facilitar a interação com o cliente, como responder a perguntas frequentes;

- **Spambots:** esses *bots* são responsáveis por disseminar conteúdo com informações falsas ou conteúdo inofensivo ou irrelevante de forma repetitiva. São *softwares* ou sistemas de computador que foram comprometidos por *malware* e estão sob o controle de terceiros. São utilizados para enviar um grande número de mensagens, publicidade, comentários etc. (Oentaryo *et al.*, 2016). Os spambots incluem *bots* promotores, spambots de URL e falsos seguidores, que se dedicam a disseminar *links* prejudiciais, mensagens indesejadas e a se apropriar de tópicos populares nas redes sociais (Meshram *et al.*, 2021; Benkler *et al.*, 2017);

- **Bots sociais:** ativos nas redes sociais, estes programas criam conteúdos por si próprios e frequentemente passam por pessoas genuínas (Oentaryo *et al.*, 2016). Os *bots* sociais são contas de usuário controladas por algoritmos que imitam a atividade humana, mas o fazem a uma velocidade notavelmente superior, enquanto conseguem ocultar, com sucesso, sua verdadeira identidade (Ferrara, 2018; Benkler *et al.*, 2017).

- **Sockpuppets:** são pessoas genuínas que assumem identidades falsas, muitas vezes para promover uma determinada causa ou produto, como promoção de

negócios, livros favoráveis e críticas de filmes. A intenção desses *sockpuppets* é criar preconceitos de opinião em relação a uma entidade e geralmente são empregados para enganar on-line (Maity *et al.*, 2017).

2.1.3 Impactos dos *bots* sobre a comunicação científica, engajamento artificial, e integridade das métricas

Os *bots* existem em grande quantidade nas redes sociais (Gilani *et al.*, 2017). De acordo com Gorwa e Guilbeault (2020), os *bots* são responsáveis por uma parcela considerável da atividade *on-line*, são usados para manipular algoritmos e sistemas de recomendação, reprimir ou promover o discurso político e podem ser cruciais na disseminação de "notícias falsas" e "hiperpartidárias".

A literatura revela que *bots* são um fenômeno multifacetado na disseminação científica no X, com implicações tanto positivas, quando usado para ampliar a divulgação científica, quanto negativas, quando se trata do uso dos indicadores altmétricos no contexto da avaliação. Por um lado, eles podem amplificar o alcance de pesquisas e facilitar a descoberta de novos conteúdos. Por outro, distorcem métricas de impacto e podem ser manipulados para criar uma falsa percepção de influência. Observa-se que a presença de *bots* no X tem gerado controvérsias. Embora possam contribuir para ampliar a visibilidade de publicações científicas, esses agentes automatizados também têm o potencial de distorcer métricas baseadas no volume de postagens.

Os *bots* representam, aproximadamente, 8,5% dos usuários do X conforme divulgado pelo X (Subrahmanian *et al.*, 2016). Um estudo sobre *bots* sociais mostra que de todos os usuários ativos de língua inglesa no X, 9% a 15% exibem comportamentos semelhantes a *bots* (Varol *et al.*, 2017b). Um estudo de Chu *et al.* (2012) descobriu que 10,5% das contas do X são *bots*, com um adicional de 36,2% classificados como "ciborgues" (definidos como humanos assistidos por *bots* ou *bots* assistidos por humanos).

2.1.4 Questões éticas do uso da automação na divulgação científica

A utilização de *bots* na divulgação científica pode assumir funções tanto construtivas quanto problemáticas, dependendo do contexto e da intencionalidade de

seu uso. Yoshikawa, Kubo e Yamamoto (2021), no estudo “*Twitter integration of chemistry software tools*”, mostraram que *bots* especializados na promoção de ferramentas químicas no X ajudam a tornar a pesquisa mais acessível e a conectar pesquisadores de diferentes áreas. De maneira similar, Xu *et al.* (2018), no artigo “*Who, what, why? An exploration of JoVE scientific video publications in tweets*”, observaram que *bots* contribuem para a promoção de vídeos científicos, ampliando o alcance de conteúdos acadêmicos nas plataformas sociais.

Por outro lado, a automação pode ser utilizada de forma antiética, como discutido por Ciupercă e Stanciu (2022) no estudo *Social networks-based alt-metrics and the unethical use of AI*. Os autores analisam o uso mal-intencionado de *bots* para manipular indicadores altmétricos, promovendo artificialmente artigos com o objetivo de inflar métricas de engajamento e impacto percebido. O estudo destaca a urgência de estratégias robustas para detecção e mitigação de *bots*, a fim de preservar a integridade das métricas acadêmicas.

As implicações éticas desse tipo de prática são significativas: ao gerar interações artificiais, *bots* podem induzir uma falsa percepção de relevância científica, comprometendo os processos de avaliação e comunicação da ciência. Além disso, conforme alertam Ciupercă e Stanciu (2022), essas ações não apenas levantam dilemas éticos, mas também podem acarretar consequências legais, sobretudo em contextos marcados pela disseminação de desinformação.

A interação entre *bots* e usuários humanos, portanto, pode tanto contribuir para a circulação legítima de conhecimento quanto distorcer os processos de avaliação acadêmica. Nesse sentido, estudos como os de Qi, AlKulaib e Broniatowski (2018) e Xu *et al.* (2018) contribuíram ao caracterizar comportamentos semelhantes a *bots* no X, sugerindo que a identificação de padrões de comportamento e, portanto, de engajamento, pode ser tão importante quanto a análise do conteúdo postado. Esses estudos ampliaram a compreensão sobre o papel dos *bots*, apontando que eles não apenas influenciam as métricas de impacto, mas também moldam a natureza da discussão acadêmica.

A partir desse caráter ambíguo entre uso legítimo e uso manipulativo da automação, emergem questões éticas centrais relacionadas à integridade da comunicação científica e à confiabilidade dos indicadores altmétricos. Conforme salientam Rodrigo Costas e Ronaldo Araújo em Maricato *et al.*, (2025), o fenômeno da automação não deve ser visto unicamente como uma ameaça, mas como um

elemento constitutivo do ecossistema comunicacional contemporâneo, cujo impacto depende de critérios claros de transparência, rastreabilidade e contextualização. A dificuldade reside em distinguir as interações automatizadas que ampliam o alcance da ciência daquelas que distorcem a percepção de impacto, produzindo métricas infladas e pouco representativas do engajamento real.

Nesse sentido, a manipulação de indicadores por *bots* e outras formas de automação configura um problema de natureza ética e teórica de primeira ordem. Rodrigo Costas (Maricato *et al.*, 2025) observa que o uso indiscriminado da altmetria em moldes avaliativos, sem curadoria ou ponderação das fontes, reforça o risco de se transformar um indicador relacional em uma métrica performativa. Ronaldo Araújo (Maricato *et al.*, 2025) amplia essa reflexão ao advertir que o uso indescriminado de métricas digitais pode fomentar comportamentos oportunistas entre pesquisadores, resultando em práticas de autopromoção automatizada e na emergência do que denomina “pesquisadores caçadores de *likes*”, voltados à visibilidade superficial em detrimento da relevância científica.

As implicações éticas dessa lógica performativa são múltiplas. Em primeiro lugar, a automação pode gerar falsas percepções de relevância, comprometendo a integridade das avaliações institucionais e de políticas públicas baseadas em métricas digitais. Em segundo, o uso de *bots* com fins de persuasão, difamação ou desinformação, fenômeno já documentado em estudos sobre a circulação de conteúdo científico em contextos de controvérsia, como vacinas e mudanças climáticas, representa uma ameaça à credibilidade da ciência e ao debate público. Ainda, a desigualdade estrutural entre países e instituições agrava essas distorções, pois grandes centros de pesquisa, com mais recursos para investir em automação e marketing digital, tendem a ampliar artificialmente sua presença online, reforçando hierarquias já consolidadas no sistema global de comunicação científica.

No plano regulatório, essas tensões se estendem à necessidade de mecanismos de governança e auditabilidade dos dados altmétricos. A ausência de transparência nas plataformas comerciais que coletam e indexam menções científicas, tais como Altmetric.com e Crossref Event Data, impede a verificação independente das fontes e favorece a opacidade algorítmica. A dependência de poucos provedores, aliada à incerteza jurídica sobre a coleta automatizada de dados (*web scraping*), torna urgente o desenvolvimento de marcos regulatórios que conciliem inovação tecnológica, proteção de dados e ética da informação.

Por fim, a discussão ética sobre o uso de IA e automação na divulgação científica deve ser compreendida sob uma perspectiva mais ampla, que considere os efeitos sociotécnicos da visibilidade digital. A mediação algorítmica das plataformas cujos mecanismos de ranqueamento e recomendação definem o que é visto, compartilhado ou amplificado, faz com que o alcance de um conteúdo científico não reflita necessariamente seu mérito, mas a lógica de atenção e engajamento que estrutura o ambiente informacional. Assim, a questão ética fundamental não é simplesmente a presença dos *bots*, mas o modo como eles participam na construção de significados, valores e hierarquias de reconhecimento na ciência.

Em síntese, a automação deve ser reconhecida como um componente inevitável e multifacetado da comunicação científica contemporânea. No entanto, sua incorporação ética requer curadoria, transparência e regulação, de modo que o uso de *bots* e algoritmos sirva à circulação legítima do conhecimento, e não à amplificação artificial de prestígio. A ética da altmetria, como defendem Rodrigo Costas e Ronaldo Araújo em Maricato *et al.*, (2025), reside menos em rejeitar a automação do que em governá-la, transformando-a em instrumento de mediação responsável entre ciência, tecnologia e sociedade.

2.2 ALTMETRIA E REDES SOCIAIS: O CASO DO X

A compreensão da altmetria no contexto das redes sociais exige a análise de seus fundamentos teóricos, das plataformas que atuam como fontes de dados e das formas de engajamento que moldam os indicadores produzidos. As subseções seguintes discutem, inicialmente, os princípios conceituais da altmetria, destacando em seguida o papel do X como fonte privilegiada de dados. Posteriormente, examinam-se as dinâmicas de engajamento entre diferentes tipos de contas e seus efeitos sobre a geração de indicadores, bem como as distorções ocasionadas pela atuação de *bots* nos scores altmétricos.

2.2.1 Fundamentos da altmetria

Atualmente, as redes sociais são um dos principais meios de comunicação em todo o mundo, gerando diariamente uma enorme quantidade de informações através

da interação das pessoas que utilizam essas plataformas. Nessas interações, as pessoas compartilham todo tipo de informação sobre si mesmas ou sobre outros usuários e expressam suas opiniões e pontos de vista sobre qualquer tema de interesse público ou privado (Mitchell, 1974). Essa crescente utilização de ferramentas de mídias sociais também tem sido utilizada para divulgação científica, despertando, nas comunidades científicas, o interesse em rastrear o impacto acadêmico por meio de novos indicadores (Araujo; Caran; Souza, 2016).

Aliado a esses fatores, a velocidade de comunicação, proporcionada pela conectividade, através da web social, remodela o cenário da comunicação científica, promovendo processos comunicativos mais ágeis e democráticos. Diferentemente das bases de dados tradicionais, esses ambientes permitem um acesso à informação menos burocratizado, favorecendo a disseminação do conhecimento de forma mais ampla.

Nesse contexto, surgem os indicadores altmétricos como resposta às limitações dos indicadores bibliométricos tradicionais. Segundo Tunger, Clermont e Meier (2018), o desenvolvimento de indicadores altmétricos se iniciou a partir de questionamentos em relação aos indicadores bibliométricos tradicionais, os quais não refletiam adequadamente o significado científico e social da produção acadêmica na era da Internet. A circulação de publicações científicas em sites de notícias, em blogs científicos e em redes sociais, aliada às emergentes ferramentas de rastreamento dessas publicações na web, contribuíram para o desenvolvimento desses indicadores. Uma vez que os indicadores altmétricos estão se consolidando como uma nova forma de mensuração da atenção social da ciência, estudos como o de Tunger, Clermont e Meier (2018) passaram a investigar até que ponto eles podem ser utilizados em avaliações científicas e em diferentes campos de aplicação.

Medir o impacto dos resultados de uma pesquisa é um componente essencial na evolução da ciência. Mesmo no panorama atual, com a crescente utilização das mídias sociais, o sistema vigente de avaliação científica privilegia os indicadores quantitativos, sobretudo por meio da contagem de citações. Esse modelo não representa, de forma eficiente, a produção científica, pois beneficia mais os países da América do Norte e Europa, uma vez que as bases de dados, tradicionalmente utilizadas como fontes para a avaliação científica, sub-representam a produção acadêmica de países em desenvolvimento (Melo; Trinca; Maricato, 2021).

Em contrapartida, os indicadores advindos das mídias sociais, conhecidos como indicadores altmétricos, têm sido entendidos como uma forma complementar de monitoramento, que podem auxiliar na medição mais ampla do impacto dos trabalhos acadêmicos de modo a beneficiar os países não centrais (Alperin, 2013). Todavia, a interpretação dos indicadores altmétricos apresentam desafios. Segundo Thelwall (2020), esses indicadores estão amplamente disponíveis, mas ainda é necessário compreender se eles podem ajudar a avaliar os resultados acadêmicos.

A conscientização da importância desses indicadores é evidente, destacando-se sua importância no ecossistema de comunicação acadêmica. No entanto, a interpretação eficaz desses dados é desafiadora, dada a falta de controle de qualidade e o potencial para manipulação ou inclusão de dados irrelevantes (Haustein; Bowman; Costas, 2016). Apesar desses desafios, os indicadores altmétricos têm valor na avaliação formativa, oferecendo indicações preliminares de possível impacto futuro para artigos individuais (Hassan *et al.*, 2019).

Ao contrário das métricas tradicionais, baseadas em citações, os indicadores altmétricos - como a contagem de leitores Mendeley - podem ser vistos como sinais precoces de impacto acadêmico e educacional, apresentando evidências antecipadas de interesse ou impacto, portanto, antes que as citações formais ocorram. No entanto, a interpretação de outros indicadores altmétricos, como a contagem de postagens em redes sociais, ainda é incerta, uma vez que esses dados tendem a refletir mais o envolvimento da comunidade acadêmica nas atividades de divulgação científica do que, de fato, o interesse do público em geral (Costas; Zahedi; Wouters, 2015; Haustein *et al.*, 2014). Contrariando parcialmente essa perspectiva, estudos apontam que a área de investigação biomédica concentra a maior média de menções por artigo na plataforma X, sugerindo um elevado grau de engajamento, além de ampla repercussão junto ao público leigo. Tal padrão revela um interesse social significativo por temas relacionados à saúde e à medicina (Costas; Zahedi; Wouters, 2015).

Essa dimensão social dos indicadores altmétricos reforça um de seus principais atributos: a capacidade de captar formas ampliadas de impacto da ciência para além do ambiente acadêmico tradicional. Nesse sentido, um dos benefícios atribuídos aos indicadores altmétricos é a possibilidade de mensurar os impactos da ciência para um público mais amplo, como profissionais, estudantes, agentes governamentais e o público em geral (Adie, 2014). Como consequência, bibliotecas e revistas digitais têm usado altmetria para identificar e rastrear indicadores de interesse

em um determinado artigo, bem como seu impacto nas plataformas de interação ciência-sociedade (Thelwall *et al.*, 2013). Medir esse impacto ampliaria o modelo de recompensa científica ao estimular os pesquisadores a promoverem o engajamento com o público.

2.2.2 O papel do X como fonte de dados altmétricos

As redes sociais digitais desempenham diferentes funções a depender do contexto e do propósito de uso: enquanto o Facebook é amplamente utilizado para manter conexões sociais; o YouTube se destaca pelo foco em entretenimento audiovisual; e o LinkedIn é voltado predominantemente para interações em ambientes profissionais (Drahošová; Balco, 2017). No entanto, entre as plataformas mais relevantes no cenário digital, o Facebook e o X merecem destaque, sendo este último largamente utilizado em pesquisas científicas que envolvem análise de dados em redes sociais.

Nesse contexto, o X se sobressai como uma das principais plataformas de disseminação de informações científicas no ambiente digital, especialmente por sua estrutura orientada à circulação rápida de conteúdos e sua penetração entre diferentes públicos (Tsou *et al.*, 2015). A plataforma de *microblogging* X tem se destacado como uma das ferramentas de redes sociais mais frequentemente utilizadas para menções a documentos acadêmicos (Costas; Zahedi; Wouters, 2015; Haustein *et al.*, 2014). Segundo Haustein *et al.* (2016), alguns autores argumentaram que o X é potencialmente capaz de captar o impacto de documentos acadêmicos em um público não acadêmico e pode, assim, servir como uma medida do impacto mais amplo da investigação.

Segundo Fang, Costas e Wouters (2022), há um movimento crescente em direção à partilha de informações acadêmicas por meio do X. Para esses autores, o X tornou-se, indiscutivelmente, uma das fontes de dados altmétricos mais cruciais, com mais de um terço dos artigos científicos recentes, sendo mencionados em postagens nesta rede social. Esse destaque é reforçado por estudos, como o de Barcelos e Maricato (2017), que identificaram que o X foi responsável por mais de 93% do número de menções em mídias sociais em artigos de periódicos em cientometria. Além disso, Joubert; Costas (2019) ressaltam que a ampla utilização do *microblogging* em escala global permite a realização de comparações entre países.

A escolha do X para a coleta de dados nesta pesquisa é motivada por fatores como a estrutura objetiva das postagens, o acesso facilitado por meio de sua Interface de Programação de Aplicações (API) pública e, sobretudo, pela expressiva influência do X nos indicadores altmétricos (Batinca; Treleaven, 2015). Em razão dessas características, o X foi selecionado como principal objeto de análise na presente investigação.

2.2.3 Engajamento de diferentes tipos de contas e a produção de indicadores altmétricos

Diversos estudos utilizam dados da rede X para identificar as características dos usuários que interagem com conteúdo científico. Haustein (2019) e Tsou *et al.*, (2015) identificaram uma predominância de contas individuais sobre contas institucionais. Já Joubert e Costas (2019), observaram que o maior volume de interações com ciência parte de membros acadêmicos, em comparação com o público geral.

Aproximadamente 15% dos acadêmicos usam o X, motivados principalmente pela disseminação de informações, construção de redes de contato e atualização científica (Haustein, 2019). No entanto, esse uso tende a ser passivo ou esporádico, em função de receios relacionados à reputação acadêmica (Nicholas *et al.*, 2015) e à confiabilidade da plataforma (Hayawi *et al.*, 2023; Segado-Boj *et al.*, 2019). Ainda assim, Barcelos e Maricato (2017) identificaram que o X é a rede que apresenta maior engajamento dos autores com suas próprias pesquisas.

Complementando essas análises, Maricato e Manso (2022) identificaram que pesquisadores representam 11,8% dos perfis que compartilham artigos, sendo responsáveis por 10,9% das citações e 11,8% das menções. Quando somadas as categorias docentes, pesquisadores e estudantes, chega-se a 21,3% dos perfis, 22,5% dos artigos e 21,4% das menções. Haustein (2019) encontrou percentuais semelhantes (22%).

Além de usuários individuais, os periódicos científicos também integram a rede de atenção. Segundo Nishikawa-pacher (2023), 25,2% dos periódicos indexados na *Web of Science* (SCIE, SSCI, AHCI) possuem conta oficial no X. Em média, essas contas emitem 116 mensagens por ano e interagem com outros usuários a cada dois

ou três *posts*. No perfil dessas contas, é comum a menção ao processo de revisão por pares (600 periódicos) e ao impacto baseado em citações (263 periódicos).

Apesar disso, há uma heterogeneidade significativa no comportamento das revistas científicas no X. A existência de perfis fraudulentos, criados por editores predatórios, reforça a importância de que as revistas mantenham perfis oficiais para proteger sua identidade institucional.

A criação de *posts* sobre ciência no X representa a disseminação da ciência por parte das contas acadêmicas, ao mesmo tempo em que abre espaço para novas interações e formas de recepção (Costas; Rijcke; Marres, 2021).

Segundo Fang, Costas e Wouters (2022b), o engajamento com *posts* acadêmicos pode ser medido por meio de quatro comportamentos: curtir (44%), repostar (36%), citar (9%) e responder (7%). Curtidas e repostagens são formas básicas de engajamento, pois não adicionam conteúdo novo; já citações e respostas expressam interações mais informativas e conversacionais.

Na literatura altmétrica, o comportamento de repostar é amplamente analisado. Priem e Costello (2010) observaram que repostagens representavam apenas 19% dos *posts* acadêmicos em uma amostra inicial. No entanto, estudos mais recentes (Alperin; Gomez; Haustein, 2019; Didegah; Mejlgård; Sørensen, 2018; Haustein, 2019) apontam que repostagens chegam a representar até metade das interações com artigos, sendo centrais para a disseminação científica entre comunidades (Araújo, 2020; Hassan *et al.*, 2019).

Mohammadi *et al.* (2018) investigaram as motivações subjacentes a esses comportamentos e identificaram que o *like* serve majoritariamente para sinalizar interesse, enquanto repostagens são motivadas pelo desejo de divulgar. Fang *et al.* (2021) acrescentam uma crítica importante: quase metade dos links encurtados em *posts* acadêmicos não são clicados, o que limita o alcance efetivo da informação.

Maricato e Manso (2022) identificaram uma média de 1,8 menções por artigo. Pesquisadores alcançam média de 1,9, médicos 1,8 e biólogos 1,7. Em relação ao número de seguidores, destacam-se professores (6%), estudantes (6%) e pesquisadores (5,9%), que juntos representam 18% dos seguidores. Apesar da importância dos perfis acadêmicos, o conjunto de usuários não acadêmicos desempenha um papel ainda mais expressivo na circulação da ciência.

Contudo, é preciso considerar que nem todas as interações observadas nas redes sociais digitais provêm de usuários humanos. O comportamento de *bots*

também tem sido foco de atenção. Haustein *et al.* (2016) propuseram uma distinção entre contas humanas, ciborgues e *bots*. Enquanto humanos operam com seleção qualitativa dos conteúdos, *bots* replicam automaticamente postagens com base em fontes e algoritmos não seletivos. As contas automatizadas postam de 4,6 a 7,1 vezes ao dia, enquanto contas humanas postam em média 2,2 vezes ao dia. Muitas dessas postagens são aleatórias e sem curadoria de conteúdo, o que desafia a validade das métricas altmétricas baseadas puramente em volume de postagens ou menções.

Essa atuação automatizada representa um fator crítico na interpretação dos indicadores altmétricos, uma vez que pode gerar distorções significativas nos scores que pretendem refletir o impacto social da produção científica — aspecto que será detalhado na seção seguinte.

2.2.4 Distorções causadas por *bots* nos scores altmétricos

Os estudos altmétricos não estão imunes a atuação dos *bots*. Aljohani, Fayoumi e Hassan (2020) destacam a importância de analisar a influência social desses agentes nos dados altmétricos, considerando não apenas como disseminam conteúdo e difundem informações, mas também a proporção de conteúdo acadêmico produzido por eles nas mídias sociais. O X, nesse contexto, assume papel central, sendo a principal fonte de indicadores monitorados por agregadores como o *Altmetric.com*, que rastreia cerca de 91% das atividades sociais a partir dessa plataforma (Hassan *et al.*, 2017).

A ação de *bots* maliciosos, como *spambots*⁴, seguidores falsos e outros, podem distorcer os indicadores altmétricos, disseminando conteúdo indesejado e prejudicando a credibilidade das métricas. Embora existam diversos estudos voltados para detecção de *bots* sociais no X (Morais; Digiampietri, 2021), ainda são escassas as investigações que abordam, especificamente, os efeitos desses mecanismos sobre dados altmétricos.

Um dos estudos que tiveram foco na investigação do impacto destes *bots* na altmetria foi publicado por Haustein *et al.* (2016). Esses autores examinaram a influência dos *bots* na contagem de *posts* do X relacionados a publicações científicas. Este trabalho pioneiro lançou luz sobre a maneira como as contas automatizadas

⁴ São programas de computador projetados para auxiliar no envio de spam, são “robôs de spam”.

geram um volume significativo de interações, questionando a validade dos indicadores altmétricos. Os autores analisaram uma amostra aleatória de 800 contas do X e constataram que 8% das contagens altmétricas foram totalmente afetadas por contas de *bots* e 5% afetadas parcialmente. O estudo mostrou ainda que 9% dos *posts* para envios de artigos do arXiv foram gerados por contas automatizadas (*bots*). Diante desses resultados, os autores enfatizaram a necessidade de diferenciar os *bots* com base no engajamento, sugerindo que a mera contagem de *posts* poderia não refletir com precisão o impacto de uma publicação. Essa investigação inicial estabeleceu um fundamento importante para futuras pesquisas, que passariam a explorar mais profundamente a dinâmica entre *bots* e usuários humanos.

Estudos adicionais corroboram esses achados. Shuai, Pepe e Bollen (2012) relataram que 50% dos *posts* sobre os artigos do arXiv foram produzidos exclusivamente por *bots*. Já Haustein (2019), ao analisar 2.043 contas altamente ativas no X (com mais de 1.000 postagens), identificou que 248 foram identificadas como automatizadas, enquanto 305 como contas de editores ou periódicos.

Ellis *et al.* (2021), no estudo “*Recent trends in the use of social media in parasitology and the application of alternative metrics*”, examinaram o uso do X e de outras plataformas na parasitologia. Eles descobriram que, embora o X seja a principal rede usada para disseminação, há uma forte presença de *bots* promovendo o compartilhamento de conhecimento. Já Orduña-Malea; Font-Julián (2022), em “*Are patents linked on Twitter? A case study of Google patents*”, focaram na disseminação de patentes, concluindo que *bots* são responsáveis por mais de 50% das menções, sugerindo um impacto significativo na percepção de inovação.

Além disso, estudos recentes têm proposto abordagens inovadoras para mensurar o impacto acadêmico, mas também evidenciam limitações dos indicadores altmétricos diante da atuação de *bots*. Nishikawa-Pacher (2023), no estudo “*Measuring serendipity with altmetrics and randomness*”, introduziu um *bot* programado para recomendar publicações científicas de forma aleatória, com o objetivo de testar o conceito de serendipidade como novo indicador de impacto. Os resultados sugerem que o engajamento gerado por essas recomendações automatizadas pode ocorrer independentemente da relevância intrínseca do conteúdo, manifestando que fatores contextuais e artificiais podem influenciar os indicadores altmétricos. Essa evidência reforça a necessidade de cautela na

interpretação dessas métricas, uma vez que a atuação de *bots* pode distorcer a percepção do impacto real de uma publicação científica.

2.3 TÉCNICAS DE DETECÇÃO DE *BOTS*: ABORDAGENS CLÁSSICAS E AVANÇADAS

Diante da relevância crescente dos *bots* na dinâmica das redes sociais e de seu impacto direto sobre os indicadores altmétricos, torna-se fundamental compreender as abordagens existentes para a identificação desses agentes automatizados. A atuação dos *bots*, especialmente no X, tem despertado o interesse da comunidade científica, que busca desenvolver métodos eficazes para sua detecção e análise. Nesse contexto, a próxima seção apresenta estudos que exploram técnicas de detecção de *bots* na plataforma X no contexto geral e em especial no contexto da divulgação científica. Esse levantamento bibliográfico visa identificar as principais metodologias utilizadas, bem como os desafios e avanços no campo.

2.3.1 Métodos tradicionais baseados em atributos de conta e comportamento

Os procedimentos mais comuns de defesa contra *bots* consistem em detectar estas contas e deletá-las, seja no momento de criação da conta ou depois que se fundem nas mídias sociais. As abordagens predominantes concentram-se na análise de conteúdo das postagens e nas características dos perfis, com ênfase em atributos comportamentais e informacionais que permitam distinguir *bots* maliciosos de usuários humanos. Ainda assim, a literatura apresenta um conjunto diversificado de técnicas de detecção, refletindo a complexidade do fenômeno.

Aguilera *et al.* (2023), em estudo publicado na revista IEEE Access (v. 11), realizaram uma análise comparativa de dez algoritmos de aprendizado de máquina aplicados à detecção de *bots* na plataforma X. Os resultados denotaram que o algoritmo *Random Forest* (Floresta Randômica) obteve o melhor desempenho entre os modelos avaliados, atingindo uma taxa de precisão de 90%, o que reforça sua eficácia no contexto de classificação supervisionada em redes sociais.

Os autores também discutem trabalhos anteriores que corroboram a eficiência do *Random Forest* para essa finalidade. Alarifi; Alsaleh; Al-Salman (2016) desenvolveram um classificador supervisionado com três categorias — *bots*, híbridos

e usuários humanos — utilizando quatro algoritmos: *Random Forest*, *BayesNet*, *Support Vector Machine* (SVM) e redes neurais artificiais multicamadas. Dentre os modelos testados, o *Random Forest* e o *BayesNet* apresentaram os melhores desempenhos, ambos com acurácia de 91%.

No estudo de Barbon *et al.* (2018), os autores propuseram uma abordagem baseada na Transformada Wavelet Discreta para extrair padrões estilísticos das postagens, com o objetivo de classificar os autores como humanos, *bots* legítimos ou *bots* maliciosos. O modelo foi treinado com o algoritmo *Random Forest* utilizando dois conjuntos de dados reais do *Twitter* (Super Bowl XLVIII e Seattle Seahawks), alcançando uma acurácia média de 94,47%. Resultado que evidencia a robustez do *Random Forest* na identificação de padrões linguísticos automatizados.

Por fim, Chu *et al.* (2012) concentraram-se na distinção entre humanos, *bots* e ciborgues, considerando como uma das principais variáveis o padrão temporal das postagens. Os autores observaram que menores níveis de entropia nos intervalos de postagem estão associados a comportamentos automatizados, enquanto valores mais elevados significam maior probabilidade de atividade humana. Embora a identificação de humanos e *bots* tenha se mostrado mais precisa, a classificação de ciborgues apresentou maior taxa de falsos positivos devido ao seu comportamento híbrido. O modelo proposto alcançou acurácia de 96% na classificação binária entre *bots* e humanos.

No contexto nacional, Pontes e Maricato (2022) propuseram um *workflow* baseado em aprendizado supervisionado, com foco na aplicabilidade em ambientes altmétricos. Utilizando o algoritmo KNN, implementado na plataforma KNIME, os autores classificaram 52% das contas analisadas como *bots*. Apesar de menos sofisticados que modelos baseados em grafos, o método destacou-se por sua replicabilidade e viabilidade computacional.

Em um experimento posterior, Pontes e Maricato (2024a) compararam o desempenho de três algoritmos supervisionados — KNN, Árvore de Decisão (*Decision Tree*) e Floresta Randômica (*Random Forest*) — na detecção de contas automatizadas. O conjunto de dados incluiu variáveis como número de postagens, seguidores, seguidos, ano de criação da conta, média de postagens diárias e tempo total de atividade. Os resultados externaram que o *Random Forest* obteve o melhor desempenho na classificação por categoria de perfil (acurácia de 71,77%), seguido

pela Árvore de Decisão, que se destacou na classificação por tipo de perfil (acurácia de 64,78%). O algoritmo KNN apresentou desempenho inferior em ambas as tarefas.

Embora os resultados tenham sido satisfatórios, os modelos avaliados apresentaram limitações em termos de robustez classificatória, o que elucida a complexidade envolvida na distinção entre perfis automatizados e humanos com base apenas em variáveis descritivas. As análises acusaram que as variáveis mais relevantes para a predição foram o número de contas seguidas, número de seguidores, ano de criação da conta e volume total de postagens — sendo esta última menos relevante que o fator temporal representado pelo ano de criação.

Tais achados ressaltam a necessidade de abordagens mais abrangentes, que integrem variáveis temporais e históricas, além de aspectos relacionados ao conteúdo e à análise de sentimentos. Ainda assim, o modelo *Random Forest*, treinado com variáveis como ano de criação, número de postagens, contas seguidas e seguidores, apresentou desempenho razoável, destacando-se como o mais promissor entre os avaliados, embora com margem para aprimoramentos, especialmente no que se refere à redução da taxa de erro e ao aumento da precisão preditiva.

Adicionalmente, a ampla adoção do *Random Forest* em ferramentas comerciais de detecção de *bots*, como o *Botometer* e o *Bot Sentinel*, evidencia sua popularidade e eficácia. A preferência por esse algoritmo deve-se, em grande parte, à sua alta acurácia, eficiência computacional e robustez frente a conjuntos de dados heterogêneos — características fundamentais em ambientes dinâmicos e suscetíveis a manipulações automatizadas, como as redes sociais.

2.3.2 Aplicações recentes de Inteligência Artificial (IA) e aprendizado profundo na pesquisa científica

Nos últimos anos, a detecção de *bots* evoluiu significativamente com o uso de aprendizado profundo. Hayawi *et al.* (2023), na revisão sistemática “*Social media bot detection with deep learning methods*”, compararam técnicas tradicionais de aprendizado de máquina (ML) (*Machine Learning*) com abordagens de aprendizado profundo (DL) (*Deep Learning*) para detecção de *bots* em âmbito geral das redes sociais. Os autores concluíram que técnicas de DL, como redes neurais convolucionais (CNN) (*Convolutional Neural Network*) e redes neurais recorrentes

(RNN) (*Recurrent Neural Network*), são mais eficazes e eficientes na detecção de *bots* em redes sociais, superando modelos clássicos de ML em termos de precisão e escalabilidade.

Bhattacharyya e Kulkarni (2024), no estudo “*Machine Learning-Based Detection and Categorization of Malicious Accounts on Social Media*”, propuseram um modelo de aprendizado de máquina para classificar contas no X em diferentes categorias, como contas genuínas, *spambots* sociais, *spambots* tradicionais e seguidores falsos, baseados no modelo DistilBERT, obtendo aproximadamente 91% de acurácia na classificação de contas automatizadas. Isso sugere que, com abordagens mais sofisticadas, é possível detectar diferentes tipos de automação que afetam a validade dos indicadores altmétricos. Da mesma forma, Ye e Na (2024) propuseram o uso de redes neurais baseadas em grafos (GNN) (*Graph Neural Networks*) integradas ao *Bidirectional Encoder Representations from Transformers* (BERT), demonstrando alto desempenho na distinção entre contas humanas e *bots*.

Em Ye e Na (2024), os autores propuseram um sistema de classificação de usuários que mencionam publicações científicas, utilizando redes neurais profundas (GNNs) e modelos de classificação de texto BERT. A análise revelou que 71% das menções vieram de usuários do público geral, enquanto apenas 12,48% dos *posts* eram de acadêmicos. Isso reforça que o X atrai um público diversificado, além da academia, sugerindo que os indicadores altmétricos devem ser interpretadas com cautela, pois podem não refletir exclusivamente o impacto entre pesquisadores, demonstrando que compreender quem está envolvido na comunicação acadêmica no X é essencial para validar o score altmétrico dos artigos.

Riazi e Wang (2024) propuseram uma abordagem de inferência causal determinística para classificar padrões de difusão de informação em redes sociais como *broadcast* (centralizado) ou *viral* (distribuído). Embora não aplicada especificamente ao contexto da divulgação científica, sua análise avulta diferenças estruturais na atuação de *bots* e humanos: enquanto usuários humanos tendem a ocupar posições centrais na propagação de conteúdos, *bots* atuam predominantemente na etapa de amplificação, com menor influência individual. Essa distinção reforça a compreensão da desinformação como fenômeno estruturado, não apenas pela presença de *bots*, mas também pela forma como eles interagem com os demais agentes na rede.

Martín-Gutiérrez *et al.* (2021) reforçam esse ponto ao demonstrar como o uso de aprendizado profundo aprimora a identificação de interações automatizadas, enquanto Bhosale e Di Troia (2022) propõem o uso da Lei de Benford em conjunto, com métodos de aprendizado de máquina para identificar perfis automatizados, sugerindo novas abordagens para caracterização de comportamento anômalo.

Dentre os estudos aplicados a contextos científicos, destaca-se o trabalho de Aljohani, Fayoumi e Hassan (2020), que investigaram a detecção de *bots* na plataforma X, por meio de diferentes abordagens, com ênfase na utilização de **redes convolucionais de grafos (GCNs)**. Os autores identificaram três classes principais de técnicas recorrentes na literatura: métodos baseados na análise de informações de redes sociais, métodos baseados em inteligência artificial e métodos de aprendizado de máquina voltados à distinção entre contas humanas e automatizadas.

No estudo, essas abordagens foram aplicadas de forma combinada à análise de dados altmétricos, revelando que os *bots* estavam associados a **87% das postagens relacionadas à altmetria**, em uma interferência significativa nesse tipo de métrica. A aplicação do modelo baseado em GCNs resultou em **acurácia de 71% e F1-score de 0,67**, valor que representa a média harmônica entre precisão e recall, sendo mais sensível a desempenhos baixos do que uma média aritmética simples.

Os autores ressaltam, entretanto, a ausência de conjuntos de dados rotulados manualmente, o que limitou o aprimoramento do modelo exclusivamente baseado em inteligência artificial. Assim, observa-se que a detecção de *bots* em contextos científicos representa um desafio metodológico complexo, o que reforça a necessidade de uma literatura específica dedicada a essa temática, a ser explorada no tópico seguinte.

2.3.3 Literatura sobre detecção de *bots* em contextos científicos

O uso de redes sociais, como o X, na disseminação científica, tem recebido crescente atenção nos últimos anos, especialmente no âmbito da altmetria. A ascensão dos indicadores altmétricos como alternativa ou complemento às métricas tradicionais de impacto científico tem intensificado o debate sobre a confiabilidade dos dados provenientes de plataformas digitais. Nesse contexto, destaca-se uma preocupação emergente: o papel desempenhado por contas automatizadas (*bots*) e seu impacto na confiabilidade dos indicadores altmétricos. A presença de *bots*, ao

promover artificialmente a disseminação de publicações acadêmicas, representa um dos obstáculos à credibilidade desses indicadores.

Nesse contexto, estudos recentes têm investido no desenvolvimento de métodos automatizados de detecção, com destaque para a aplicação de técnicas de aprendizado de máquina, como uma via promissora para mitigar os vieses introduzidos por tais agentes. Esta seção apresenta uma síntese dos principais estudos dedicados à detecção de *bots* em contextos científicos, com base nas questões de pesquisa delineadas no protocolo de revisão de literatura (Apêndice A).

Para uso mais eficaz dos indicadores altmétricos, é essencial desenvolver métodos robustos de detecção de *bots* e considerar a complexidade das interações entre diferentes tipos de usuários nas redes sociais. Haustein *et al.* (2016) demonstram que *bots* são responsáveis por um número expressivo de menções a artigos científicos. No entanto, essas interações diferem das realizadas por usuários humanos, pois geralmente não envolvem engajamento genuíno com o conteúdo. Os autores sugerem, portanto, que se faça a distinção entre menções meramente “bibliográficas” e aquelas que refletem uma apropriação mais significativa da informação. Essa preocupação motivou uma série de estudos voltados à detecção de *bots*, à identificação de seus comportamentos característicos e à compreensão dos efeitos desses agentes nos indicadores altmétricos.

A literatura aponta que *bots* se diferenciam dos usuários humanos por apresentarem padrões de postagem intensivos e repetitivos, além de interações geralmente descontextualizadas. Estudos como o de Ye e Na (2024) apontam que, durante a pandemia de COVID-19, *bots* desempenharam papel relevante na disseminação de conteúdos científicos, registrando frequência de postagem significativamente superior à dos usuários humanos. Enquanto estes enriquecem os *posts* com comentários ou discussões, *bots* reproduzem conteúdos de forma automatizada e padronizada.

Nos estudos de Ye e Na (2024), “*Who is mentioning COVID-19 articles on twitter? Classifying twitter users in the context of scholarly communication*”, os autores analisaram como as contas automatizadas promoviam publicações científicas relacionadas à COVID-19, revelando que uma parcela significativa dos *posts* sobre o tema origina-se dos *bots*. Eles descobriram que *bots* contribuíam com 2,9% dos *posts*, atuando como amplificadores de informação e comportando-se de forma diferente de

outros *bots* não científicos. Esses resultados ressaltam a necessidade de diferenciar entre tipos de automação e suas motivações ao avaliar o impacto acadêmico on-line.

A principal crítica à presença de *bots* refere-se ao comprometimento da validade dos indicadores altmétricos. Conforme argumentam Haustein *et al.* (2016), o alto volume de postagens automatizadas pode inflar artificialmente a visibilidade de artigos científicos, sem refletir um impacto acadêmico ou social autêntico. Essa distorção compromete a confiança em métricas que deveriam representar formas legítimas de engajamento com a produção científica.

Para enfrentar esse desafio, diversos pesquisadores têm recorrido a ferramentas de visualização e análise de redes, como Gephi e NodeXL, além de modelos de aprendizado profundo, especialmente baseados em BERT e GNN.

Entretanto, um dos principais obstáculos identificados pela literatura reside na rápida evolução dos algoritmos utilizados pelos próprios *bots*, que dificulta a sua identificação. Ye e Na (2024) destacam que há variedade de tipos de *bots*, com diferentes finalidades e padrões comportamentais, tornando o processo de identificação ainda mais complexo.

Dentre as estratégias propostas para mitigar os efeitos da atuação dos *bots*, destacam-se o desenvolvimento de sistemas robustos de detecção, a implementação de verificações de identidade e o uso de indicadores complementares. Nishikawa-Pache (2023), por exemplo, propõe o uso de métricas de serendipidade altmétrica como alternativa para filtrar interações genuínas. A conscientização da comunidade científica também é fundamental para lidar com esse desafio.

Importante salientar que nem todos os *bots* afetam os indicadores altmétricos da mesma maneira. Enquanto *bots* de spam produzem conteúdos irrelevantes em grande escala, *bots* promocionais podem amplificar conteúdos científicos legítimos, ainda que distorçam a percepção do engajamento real (Orduña-Malea; Font-Julián, 2022b). Já os *bots* acadêmicos, como os estudados por Ye e Na (2024), podem desempenhar um papel funcional na disseminação de conhecimento, desde que suas ações sejam interpretadas com cautela.

As tendências futuras apontam para o aprimoramento contínuo das técnicas de aprendizado de máquina, com destaque para o uso de processamento de linguagem natural na diferenciação entre contas humanas e automatizadas (Bhattacharyya; Kulkarni, 2024). Além disso, cresce a importância da colaboração entre pesquisadores, plataformas de mídia social e instituições acadêmicas, no sentido de

estabelecer diretrizes éticas e técnicas para o monitoramento da comunicação científica mediada por *bots*.

O estudo de Arroyo-Machado, Herrera-Viedma e Torres-Salinas (2023) introduz uma perspectiva distinta ao aplicar a ferramenta *BotometerLite API*, baseada em aprendizado de máquina, para identificar *bots* na disseminação de artigos em Ciências Sociais. Os resultados expressaram uma presença relativamente baixa de *bots* (3,61% dos usuários), porém com variações significativas entre disciplinas. Esse achado sugere que o impacto de *bots* pode ser sensível ao campo científico analisado, reforçando a necessidade de abordagens metodológicas específicas por domínio.

Complementando essas abordagens, Haustein (2019) e Haustein *et al.* (2016) apontam limitações nas ferramentas estatísticas existentes, como o BOTORNOT, ao lidarem com *bots* acadêmicos, que se comportam de forma diferente de *bots* sociais comuns. Os autores sugerem a adoção de métricas baseadas em engajamento e diferenciam entre humanos, ciborgues e *bots* como categorias relevantes para o refinamento dos modelos de detecção. Essa crítica evidencia que a eficácia dos modelos depende da acurácia na captura dos padrões específicos de comportamento dos *bots* voltados à comunicação científica.

Em síntese, a comparação entre os estudos revela uma diversidade de métodos aplicados à detecção de *bots* em contextos altmétricos. De tal modo que desde modelos baseados em GCNs, passando por algoritmos supervisionados simples (KNN), até o uso de APIs especializadas e métodos manuais, notabiliza-se que não há uma abordagem única ideal. A escolha do método depende das características do corpus analisado, dos recursos computacionais disponíveis e da necessidade de interpretabilidade do modelo. Estudos futuros devem avançar na integração entre diferentes abordagens, buscando soluções mais robustas e sensíveis às especificidades do ecossistema altmétrico.

Com base nos artigos selecionados na revisão sistemática de literatura apresentada no Apêndice A, foram identificadas lacunas importantes na literatura que limitam o avanço do conhecimento sobre a influência de *bots* nos indicadores altmétricos. A seguir, são discutidos aspectos ainda pouco explorados ou que carecem de aprofundamento teórico e metodológico, complementados por referências adicionais que enriquecem a análise:

a. Falta de consenso sobre medidas de impacto: embora diversos estudos reconheçam que *bots* influenciam a contagem de *posts* e os indicadores altmétricos,

ainda não há consenso sobre como interpretar esses dados. Destaca-se a necessidade de desenvolver métricas mais robustas e confiáveis para avaliar o impacto real das publicações científicas (Konkiel, 2016; Saqr; López-Pernas; Apiola, 2023). Além disso, é fundamental explorar como métodos de *machine learning* e ferramentas analíticas podem ser aplicados para validar a autenticidade das interações (Bhattacharyya; Kulkarni, 2024).

b. Necessidade de taxonomias mais abrangentes: apesar de algumas tentativas de classificar usuários (como na pesquisa de Ye e Na (2024), ainda não há uma taxonomia universalmente aceita que capture a diversidade de comportamentos e tipos de contas (humanos vs. *bots*) nas redes sociais. A exploração de categorias mais detalhadas, como a apresentada por Bhattacharyya e Kulkarni (2024), que distingue entre *spambots* sociais, *spambots* tradicionais e seguidores falsos, poderia enriquecer a análise de redes sociais e proporcionar uma classificação mais precisa.

c. Sub-representação de determinantes sociais e comportamentais: A maioria dos estudos se concentra em análises quantitativas, como número de postagens e menções. São raras as investigações que abordam aspectos sociais, culturais e comportamentais que motivam os usuários a interagirem com conteúdos científicos. Xu *et al.* (2018), por exemplo, destacam a importância da comunicação científica em formatos alternativos, como vídeos, apontando para a necessidade de abordagens mais interpretativas.

d. Desenvolvimento de métodos de detecção de *bots*: embora haja avanços na detecção de *bots* usando aprendizado profundo, como discutido no artigo de Bhattacharyya e Kulkarni (2024), ainda há uma lacuna na aplicação prática dessas técnicas em ambientes de pesquisa e sua integração em plataformas de mídia social. Pesquisas futuras podem explorar a eficácia de métodos de detecção em tempo real e suas implicações para a integridade dos dados (Qi; AlKulaib; Broniatowski, 2018a; Bhosale; Di Troia, 2022).

e. Impacto de *bots* na disseminação de diferentes tipos de conteúdo: há uma escassez de estudos que diferenciam o impacto de *bots* na disseminação de diferentes tipos de conteúdo científico (artigos, vídeos, patentes etc.). Enquanto alguns artigos se concentram em artigos acadêmicos, outros discutem vídeos científicos (Xu *et al.*, 2018). No contexto geral, Riazi e Wang (2024) apontam a necessidade de explorar como esses diferentes formatos são afetados em termos de engajamento e alcance.

f. Conexões entre mídia social e avaliação acadêmica: muitos estudos abordam a relação entre indicadores altmétricos e métricas tradicionais, mas há uma necessidade crescente de examinar como essas métricas influenciam as decisões de avaliação acadêmica, como promoções e concessões de financiamento. Um exame mais profundo dessa interseção pode fundamentar políticas de avaliação e práticas acadêmicas (Haustein, 2019; Saqr; López-Pernas; Apiola, 2023).

g. Interseção de ética e IA nas mídias sociais: O uso antiético de inteligência artificial para manipulação de métricas altmétricas é uma preocupação crescente. Mais estudos são necessários para investigar como práticas éticas podem ser integradas ao uso de bots e IA em contextos científicos (Ng; Verhoeff; Steen, 2023).

h. Evolução das plataformas e suas dinâmicas: Com mudanças contínuas nas redes sociais, como o *rebranding* para “X” e alterações nas políticas de acesso a dados, torna-se essencial estudar como essas transformações afetam a comunicação científica e a atuação de *bots* (Alani, 2021; Bhattacharyya; Kulkarni, 2024, Riazi; Wang, 2024).

As lacunas identificadas evidenciam que, apesar dos avanços nas investigações sobre a comunicação científica em mídias sociais, o campo ainda precisa de aprofundamento em diversos aspectos. A elaboração de metodologias mais rigorosas, a classificação minuciosa dos perfis de usuários e a incorporação de dimensões sociais e éticas constituem etapas fundamentais para uma compreensão mais abrangente dos efeitos dos *bots* e dos indicadores altmétricos na disseminação da produção científica.

2.3.4 Estudos experimentais da literatura sobre classificação de *bots* utilizando o XGBoost

Diversos trabalhos têm explorado a aplicação do XGBoost, tanto de forma isolada quanto em conjunto com outras técnicas e estratégias para melhorar a acurácia da detecção de contas automatizadas, particularmente em contextos altmétricos e redes sociais. Observa-se um crescente interesse no uso do algoritmo XGBoost para a detecção de *bots* na plataforma X, dada a sua robustez em lidar com dados estruturados, sua capacidade de generalização e alto desempenho em tarefas de classificação supervisionada.

O estudo de Mohanty *et al.* (2022) destaca o uso de métodos de aprendizado de máquina do tipo ensemble, como Adaboost, *Gradient Boosting* e XGBoost, na construção de um framework de detecção de *bots*. Com base em um conjunto de dados contendo cerca de 2.700 usuários, o XGBoost alcançou uma acurácia de 89%, demonstrando desempenho comparável ao *Gradient Boosting*, que obteve 89,5%. A proposta reforça a viabilidade do XGBoost como componente central de modelos que visam identificar contas automatizadas com precisão elevada.

Owais, Shoaib e Waseem (2023), por sua vez, propõem uma abordagem híbrida, baseada na integração de múltiplas categorias de atributos — perfil, comportamento, conteúdo, rede e tempo — aplicadas ao conjunto de dados Twibot-22. O modelo baseado em XGBoost superou os classificadores de referência, atingindo métricas expressivas: acurácia de 0,898, F1-score de 0,875, precisão de 0,876 e recall de 0,899. Esses resultados reforçam o potencial do XGBoost em contextos de classificação complexos e multivariados, especialmente quando suportado por engenharia de atributos criteriosa.

No trabalho de Lingam *et al.* (2022), o foco recai sobre o desafio do desbalanceamento de dados, comum em tarefas de detecção de *bots*. Para enfrentá-lo, os autores propõem o GG-XGBoost, um modelo que integra XGBoost com redes adversariais generativas e unidades recorrentes (GRU-GAN), possibilitando a geração de dados sintéticos balanceados. O modelo obteve acurácia de 92,84%, superando modelos baseados apenas em *deep learning* (CNN, LSTM, Bi-LSTM), o que evidencia a eficácia da combinação entre técnicas generativas e aprendizado baseado em árvores.

Por fim, o estudo de Duddu e Bhavani (2023) propõe o modelo ACTHU_XGB, que utiliza o XGBoost como classificador, aplicado a dados codificados por meio da abordagem ACTHU (*Hashtags*, URLs, menções, conteúdo textual e tempo) combinada a uma variável de sentimento. A pesquisa emprega os *datasets* Twibot-20 e Cresci-17, propõe ainda a fusão de diferentes estratégias de codificação, incluindo abordagens baseadas em DNA e perfil de usuário. O uso do XGBoost neste modelo é apontado como central para atingir melhores desempenhos na classificação entre contas genuínas e automatizadas.

Em síntese, os estudos apontam que o XGBoost se consolidou como uma técnica de destaque na detecção de *bots* em redes sociais, seja por sua performance competitiva, sua flexibilidade na integração com outras técnicas (como GANs e

codificações específicas), ou por sua capacidade de lidar com dados complexos e desbalanceados.

2.3.5 Detecção de *bots* com o Botometer

O Botometer constitui um caso emblemático para refletir sobre as metodologias disponíveis para a detecção de *bots* em mídias sociais, especialmente no X. Embora a presente pesquisa não se proponha a utilizar diretamente essa ferramenta, sua ampla adoção na literatura científica e a sofisticação de seus métodos tornam-na uma referência importante para compreender as possibilidades e as limitações das abordagens automatizadas de identificação de contas não humanas. Assim, o exame de suas características técnicas, as aplicações empíricas e os impactos em diferentes campos do conhecimento contribuem para o mapeamento crítico das opções metodológicas atualmente disponíveis para o estudo da automação em ambientes digitais.

O Botometer é uma ferramenta desenvolvida pela Universidade de Indiana para detectar *bots* no X, com base em algoritmos de aprendizado de máquina. Treinado a partir de dados rotulados, ele calcula uma pontuação que expressa a probabilidade de uma conta ser automatizada. A ferramenta analisa mais de mil características, distribuídas em seis categorias, como perfil, rede, conteúdo e sentimento. Sua validação científica e aplicação em diversos contextos garantem destaque frente a outras ferramentas semelhantes.

A coleta de dados pelo Botometer é feita via API do X, acessando os 200 *posts* mais recentes da conta e informações da rede associada. Por sua natureza dinâmica, os resultados podem mudar com o tempo, especialmente em contas muito ativas ou suspensas. Por isso, é recomendável que a análise de *bots* ocorra logo após a coleta dos dados. A precisão do Botometer é alta, com AUC de 0,99 na versão mais recente, sobressaindo sua eficácia na distinção entre contas humanas e *bots*.

O Botometer tem sido amplamente usado em pesquisas. Um exemplo é o estudo de Keller e Klinger (2019), que investigaram o aumento de *bots* durante campanhas eleitorais na Alemanha. Eles observaram que a presença de *bots* subiu de 7,1% para 9,9%, com atuação pouco relacionada a debates políticos nacionais. Outro estudo, de Broniatowski *et al.* (2018), analisou a atuação de *bots* em conteúdos

sobre vacinas, revelando como *bots* e *trolls* contribuem para a desinformação em saúde pública.

No campo da saúde, Allem; Escobedo; Dharmapuri (2020) utilizaram o Botometer para distinguir *bots* em discussões sobre cannabis, usando classificadores de texto para identificar os principais temas tratados nas postagens. Já na área financeira, Fan; Talavera; Tran (2020) mostraram que a atividade de *bots* no X influencia diretamente a volatilidade e o volume de negociação de ações. Em outro estudo, Gorodnichenko; Pham; Talavera (2021) combinaram critérios comportamentais e o Botometer para detectar *bots*, confirmado a consistência entre abordagens.

Pesquisas sobre desinformação também destacaram o uso do Botometer. Grinberg *et al.* (2019) revelaram que uma pequena parcela de usuários concentra a maior parte da exposição e do compartilhamento de notícias falsas. Já o Pew Research Center constatou que 66% dos links para sites populares eram postados por *bots*, sendo que um número reduzido de contas automatizadas era responsável por uma parcela significativa dessa atividade, sem evidência clara de viés político (Wojcik; Weisel; Smith, 2018).

No contexto altmétrico, Arroyo-machado; Herrera-viedma; Torres-salinas (2023) utilizaram o Botometer para avaliar a presença de *bots* em menções a artigos científicos no X. A presença média foi de 3,61% dos usuários e 3,85% dos *posts*, com maior impacto em áreas como Criminologia e Matemática. Apesar disso, nas Ciências Sociais e Humanas, o impacto altmétrico dos *bots* foi considerado mínimo, reforçando que a influência varia significativamente entre disciplinas.

A trajetória da pesquisa sobre *bots* no X mostra um avanço metodológico e analítico relevante. Desde as primeiras abordagens descritivas até os modelos baseados em aprendizado de máquina, os estudos têm se ajustado às mudanças tecnológicas e ao comportamento on-line. O Botometer aparece como uma ferramenta central nesse processo, contribuindo para o entendimento do papel dos *bots* na comunicação científica, política e social, além de evidenciar a importância de análises críticas e éticas sobre o impacto da automação.

2.4 MODELAGEM SUPERVISIONADA E ENGENHARIA DE *FEATURES* NA DETECÇÃO DE *BOTS* NO X

A eficácia na detecção de *bots* depende tanto da escolha adequada dos algoritmos supervisionados quanto da definição criteriosa das características utilizadas no treinamento. As *features* correspondem a atributos extraídos das contas e de suas interações na plataforma, permitindo a identificação de padrões típicos de comportamento automatizado. As subseções seguintes apresentam os principais algoritmos empregados nesse processo, bem como as variáveis utilizadas, ressaltando sua importância para a construção de classificadores mais precisos e interpretáveis.

2.4.1 Principais algoritmos supervisionados: KNN, Árvore de Decisão, Random Forest e XGBoost

No âmbito da detecção de *bots* em redes sociais, diferentes algoritmos de aprendizado supervisionado têm sido aplicados com o objetivo de identificar padrões que distinguem contas humanas de contas automatizadas. Cada técnica apresenta especificidades em termos de funcionamento, capacidade de generalização e complexidade computacional, aspectos que influenciam diretamente sua eficácia na classificação. As subseções a seguir descrevem, de forma individualizada, os principais algoritmos empregados nesse contexto, iniciando pelo *K-Nearest Neighbors* (KNN).

2.4.1.1 *K-Nearest Neighbors* (KNN)

O algoritmo KNN é uma técnica de aprendizado de máquina que pode ser aplicada a problemas de classificação e regressão. Sua abordagem baseia-se na determinação da classe de um ponto de dados com base na maioria das classes dos *k* pontos mais próximos a ele, onde *k* é um parâmetro definido pelo usuário. Diferentemente de alguns outros algoritmos, o KNN não exige um treinamento explícito, uma vez que o modelo é simplesmente memorizado durante a fase de treinamento. Vale ressaltar que o KNN é sensível à escala dos recursos, frequentemente demandando normalização dos dados (Brink, 2016).

O Algoritmo KNN considera os objetos do conjunto de treinamento mais próximos do ponto de teste, em que K é um parâmetro do algoritmo. Quando o valor de K é maior que 1, para cada ponto de teste, são obtidos K vizinhos. Cada vizinho é associado a uma classe, sendo possível que diferentes vizinhos sejam agregados de modo a classificar o ponto de teste (Aguilera *et al.*, 2023). Posteriormente, as classes tomam valores em um conjunto discreto e cada vizinho é associado a uma classe. O objeto de teste é classificado na classe mais votada. Dizendo de maneira mais simplista, o algoritmo KNN classifica um novo dado (y) a partir da comparação com o conjunto de dados previamente rotulado (x), atribuindo a y a classe predominante entre seus vizinhos mais próximos.

2.4.1.2 Árvore de decisão

O algoritmo de árvore de decisão é uma técnica de aprendizado de máquina que modela decisões complexas por meio de uma estrutura hierárquica semelhante a uma árvore. Essa árvore é construída durante a fase de treinamento, quando o modelo busca identificar as características mais relevantes e as decisões que levam a uma classificação ou previsão desejada (Aguilera *et al.*, 2023).

Durante a construção da árvore de decisão, o algoritmo avalia diversas características do conjunto de dados para determinar qual delas oferece a melhor divisão dos dados em subconjuntos mais homogêneos. O critério utilizado para essa avaliação pode variar, sendo comumente o ganho de informação ou o índice de Gini. O processo é repetido recursivamente até que seja atingido um critério de parada, como um número máximo de nós ou a homogeneidade dos subconjuntos (Brink, 2016).

A árvore resultante é uma representação gráfica e lógica das decisões tomadas pelo modelo, onde cada nó interno representa uma decisão com base em uma característica, cada ramo representa o resultado dessa decisão, e as folhas representam classes ou valores previstos. Isso confere à árvore de decisão uma interpretabilidade única, permitindo que especialistas e usuários compreendam facilmente o raciocínio subjacente às previsões (Brink, 2016).

Durante a fase de inferência, o modelo utiliza a estrutura da árvore para guiar novas instâncias de dados através das decisões, levando à classificação ou previsão desejada. Além da interpretabilidade, a árvore de decisão é apreciada pela sua

capacidade de lidar com dados categóricos e numéricos, sendo um algoritmo versátil, aplicável a uma variedade de problemas, incluindo classificação e regressão (Brink, 2016).

Apesar das suas vantagens, árvores de decisão podem ser suscetíveis a *overfitting*⁵, especialmente em conjuntos de dados complexos. Estratégias como a poda da árvore e o uso de *ensemble methods*, como as Florestas Randômicas, podem ser empregadas para mitigar esses desafios e melhorar o desempenho geral do modelo.

As árvores de decisão são versáteis, sendo aplicáveis tanto a tarefas de classificação quanto de regressão. Esses modelos operam através da divisão do conjunto de dados com base em decisões "sim" ou "não" relacionadas a diferentes características, formando assim uma estrutura de árvore que orienta as decisões. O treinamento envolve a escolha das melhores características para dividir os dados em cada nó da árvore, com base em critérios como ganho de informação ou índice de Gini. Uma característica notável das árvores de decisão é a interpretabilidade, pois refletem uma sequência lógica de decisões comprehensíveis.

2.4.1.3 *Random Forest*

O algoritmo de Floresta Randômica é uma técnica de aprendizado de máquina baseada em ensemble, projetada para melhorar o desempenho e a generalização de modelos de árvores de decisão individuais. Essa abordagem é particularmente eficaz para mitigar o *overfitting* e aumentar a robustez do modelo (Brink, 2016).

A Floresta Randômica opera construindo várias árvores de decisão independentes durante a fase de treinamento. Cada árvore é treinada em uma amostra aleatória do conjunto de dados, com substituição (*bootstrap sampling*), e as decisões em cada nó da árvore são tomadas com base em um subconjunto aleatório de características. Essa aleatoriedade na seleção de dados e características contribui para a diversidade entre as árvores individuais, tornando a Floresta Randômica mais resistente a padrões específicos do conjunto de treinamento.

Segundo Brink (2016), Durante a fase de inferência, as previsões de cada árvore são combinadas para formar uma previsão final. Para problemas de

⁵ O termo refere-se ao ajuste excessivo. É um comportamento indesejável de aprendizado de máquina.

classificação, isso pode envolver uma votação majoritária, enquanto para problemas de regressão, a média das previsões individuais é frequentemente utilizada.

A principal vantagem da Floresta Randômica reside na redução da variância e na capacidade de lidar com sobreajuste, tornando-se uma escolha popular para uma variedade de tarefas de aprendizado de máquina. Além disso, a interpretabilidade do modelo é preservada, pois as decisões individuais de cada árvore ainda podem ser analisadas (Brink, 2016).

É importante notar que a Floresta Randômica pode ser ajustada através da parametrização, como o número de árvores na floresta, a profundidade máxima de cada árvore, e o número de características consideradas em cada divisão. Esse ajuste fino permite que o especialista em inteligência artificial adapte o modelo às características específicas do problema em questão, maximizando o desempenho e a eficácia da Floresta Randômica em diferentes cenários.

Em resumo, a floresta randômica (ou aleatória) é uma técnica de ensemble que utiliza múltiplas árvores de decisão para melhorar o desempenho e a generalização do modelo. Durante o treinamento, são criadas várias árvores de decisão independentes, e as previsões de cada árvore são combinadas para gerar uma previsão mais robusta. Cada árvore é treinada em uma amostra aleatória dos dados e/ou com características aleatórias, o que ajuda a reduzir a chance de *overfitting*. A floresta randômica proporciona variedade e redução de variância em comparação com uma única árvore de decisão, resultando em um modelo mais estável e menos propenso a *overfitting*. A escolha entre esses algoritmos depende das características específicas do problema em questão, do tipo de dados e dos requisitos desejados para o modelo.

2.4.1.4 *Extreme Gradient Boosting* (XGBoost)

O XGBoost é um dos algoritmos mais populares e poderosos para problemas de aprendizado supervisionado, especialmente para classificação e regressão. Ele é baseado em árvores de decisão, mas implementa técnicas avançadas para aumentar a eficiência, a precisão e a velocidade do treinamento. Abaixo está uma explicação detalhada do XGBoost:

XGBoost é um *ensemble model* que utiliza a técnica de *boosting* (Chen; Guestrin, 2016). O *boosting* consiste em combinar múltiplos modelos fracos (no caso

do XGBoost, árvores de decisão) para criar um modelo forte. Isso é feito treinando uma árvore por vez, de modo que cada nova árvore corrige os erros das anteriores. Esse modelo foi projetado para ser:

- **Eficiente**: rápido em grandes *datasets*.
- **Escalável**: suporta paralelismo e funciona bem em sistemas distribuídos.
- **Flexível**: permite ajuste fino de hiperparâmetros.

O XGBoost usa o método de *Gradient Boosting*, cujo modelo é inicializado com um valor preditivo básico, treina sequencialmente árvores de decisão, ajustando-se para minimizar o erro residual.

Uma vez apresentadas as abordagens de aprendizagem de máquina que são referenciadas em trabalhos acadêmicos para detecção de *bots* no contexto geral, torna-se pertinente ampliar a análise para além da identificação técnica, incorporando uma perspectiva mais integrada com o contexto altmétrico. A seção seguinte propõe essa reflexão, ao explorar a interseção entre altmetria, *bots* e a classificação de contas na plataforma X, examinando como a presença de perfis automatizados pode impactar a mensuração da atenção científica nas redes sociais e quais estratégias vêm sendo desenvolvidas para lidar com esse fenômeno no campo da comunicação científica.

2.4.2 Características/features utilizadas para treinamento de modelos de IA

A busca por entender quais são as características dos *bots* no X não é recente e tem se mostrado cada vez mais relevante na identificação destes agentes. Algumas pesquisas como a de Gilani *et al.* (2017) ainda estão interessadas em explorar o papel destes mecanismos automatizados no ecossistema social mais amplo e principalmente em como as suas características comportamentais diferem das dos humanos. Os autores levantam questões sobre como os *bots* operam, interagem e afetam a produção de conteúdo online, exploram quais são os comportamentos típicos dos humanos e dos *bots*.

Varol *et al.* (2017) mostram que a seleção de amigos e seguidores está correlacionada com a probabilidade de contas *bots*. Também destacam como os *bots* usam diferentes estratégias de repostagem e menção ao interagir com humanos ou outros *bots*. Os autores concluem caracterizando subclasses de comportamentos de

contas. Os *clusters* identificados por esta análise apontam, principalmente, para três tipos de *bots*:

- *Self Promoters*: Contas legítimas focadas em autopromoção, geralmente com padrões de comportamento homogêneos.
- *Spammers*: Contas altamente ativas com poucos seguidores, associadas à disseminação automatizada de conteúdo.
- Contas de Aplicativos Conectados: Contas que compartilham automaticamente conteúdo de outras plataformas via aplicativos vinculados, muitas vezes mantidas por usuários inativos.

Esses resultados enfatizam que o X hospeda uma variedade de usuários com comportamentos diversos, o que é verdade para contas humanas e *bots*. Em alguns casos, a fronteira que separa esses dois grupos não é nítida e uma conta pode apresentar características de ambos. Segundo os autores Varol *et al.* (2017), os seres humanos tendem a interagir com mais contas semelhantes a humanos do que contas semelhantes a *bots*, em média. Outro fato observado é que a reciprocidade dos laços de amizade é maior para os humanos.

O estudo realizado por Varol *et al.* (2017) apresenta uma estrutura de detecção de *bots* do X, que utiliza mais de mil características para avaliar em que medida uma conta do X apresenta semelhança com as características conhecidas dos *bots* sociais. Nesta pesquisa, os dados coletados usando a API do X são sintetizados em 1.150 recursos⁶ classificados em seis classes diferentes. De forma convergente, Morais e Digiampietri (2021) afirmam que um aspecto fundamental no processo de criação de mecanismos de detecção de *bots* reside na definição das características destes agentes, podendo estas serem agrupadas em **seis** grupos distintos:

1. Baseadas no **usuário**: número de amigos e seguidores, número de *posts* produzidos (por exemplo, número de *posts*), descrição do perfil e configurações;
2. Baseadas nos **amigos**: relações com outros usuários da rede, incluindo o envio de menções, respostas e compartilhamento de mensagens, ser mencionado ou ter suas mensagens compartilhadas/encaminhadas;
3. Baseadas na **rede**: diferentes tipos de interação na rede, considerando a frequência de interação e co-ocorrências de *hashtags*;

⁶ Desde julho de 2023 disponível no acesso pago da API do X

4. **Temporais**: consideram a atividade do usuário durante diferentes intervalos de tempo;
5. Baseadas no **Conteúdo**: o tipo de linguagem utilizada, comprimento das mensagens e a entropia do texto sendo publicado;
6. Baseadas em **Sentimento**: a atitude ou o humor de uma conversa ou mensagem.

A análise da literatura especializada sobre o tema permite a ampliação das classes e dos tipos de atributos empregados na detecção de *bots*. A seguir, são apresentadas todas as categorias identificadas, acompanhadas das respectivas referências bibliográficas que as mencionam:

- a. **Recursos baseados no usuário**: os recursos derivados dos metadados do usuário têm sido empregados na categorização de usuários (Mislove *et al.*, 2011; Ferrara, Emilio; *et al.*, 2016). Esses atributos englobam a contagem de amigos e seguidores, a quantidade de postagens geradas pelos usuários, bem como as informações de descrição e configurações do perfil. Metadados do perfil, como biografia, frequência de postagens, fotos de perfil e de capa e tipos de interações são examinados para identificar contas de *bots*, segundo Didegah (2018) e Pontes e Maricato (2023b). Fang, Costas e Wouters (2022) detalham características como o número de seguidores, o número de listas em que o usuário é mencionado, o número de usuários seguidos (amigos), o total de curtidas concedidas desde a criação da conta e o número de *posts* realizados. Bhattacharya *et al.* (2014) também consideram a contagem de favoritos, a contagem de seguidores e a idade da conta (idade do tweet) como variáveis relevantes. Balestrucci *et al.* (2019) sugerem que detalhes da conta, como "class-A *features*" e a análise comportamental baseada no "DNA *fingerprint*" (Cresci *et al.*, 2017) são importantes para identificar usuários. Crotty (2014) menciona a capacidade de *networking* de um autor, como o número de seguidores, como um fator que pode ser medido.
- b. **Recursos relacionados aos amigos**: o Xativamente fomenta a interconectividade, na qual os usuários estabelecem relações de seguidor-amigo (*followee*). A disseminação de conteúdo ocorre de forma interpessoal, através de postagens e repostagens que podem ser direcionados a usuários específicos por meio de menções. Quatro tipos de conexões são considerados: repostagens (quando a conta repostada), menções, ser repostado (quando a

conta é repostada) e ser mencionado. Para cada categoria, é possível extrair atributos abordando elementos como o uso da linguagem, horário local e níveis de popularidade, entre outros. É importante notar que, devido às restrições impostas pela API do X em 2023, as informações relativas a seguidores/followees não são disponibilizadas, o que dificulta a identificação de contas *bot*.

- c. **Recursos de rede:** Muito semelhante aos recursos relacionados aos amigos citados anteriormente, a estrutura da rede contém informações de importância fundamental para a caracterização de diversos modos de comunicação. O foco aqui consiste em explorar a estrutura e as interações do usuário dentro da rede social. Com efeito, a utilização de recursos de rede desempenha um papel importante em tarefas como a detecção de estratégias políticas (Ratkiewicz *et al.*, 2011). É possível identificar três tipos distintos de redes com base nos quatro tipos de conexões citados no item b.: repostagens, menções e redes de co-ocorrência de *hashtags*. Nas redes de repostagens e menções, os usuários constituem os nós, apresentando conexões direcionadas que refletem a direção da disseminação da informação, seja em direção ao usuário que efetua a repostagem ou ao mencionado. Por sua vez, as redes de co-ocorrência de *hashtags* estabelecem conexões não direcionadas entre os nós de *hashtag* sempre que duas *hashtags* coincidem em um *post*. Todas essas redes são ponderadas de acordo com a frequência de interações ou co-ocorrências. Para cada uma delas, são calculados conjuntos de atributos, abrangendo as distribuições de grau ponderado interno e externo, densidade e coeficiente de agrupamento. Importante ressaltar que tanto o grau superior quanto a força superior representam métricas de popularidade (Ratkiewicz *et al.*, 2011). Costas, Rijcke e Marres (2021) abordam "acoplamentos heterogêneos" e a aplicação de perspectivas de rede para analisar as interações entre ciência e sociedade por meio de métricas de mídias sociais. Eles também incluem "redes de dois modos" que podem envolver (re)postagens, postagens no Facebook e comentários on-line, todos ligados a objetos científicos e à análise de *hashtags* em redes sociais. Ferrara e Yang (2015) fazem referência à estrutura da comunidade como um aspecto relevante para a análise. Didegah (2018) considera os "tipos de interações (respostas ou comentários)" que um usuário teve em seus *posts* originais ou repostagens como um componente da análise

de rede. Além disso, Fang, Costas e Wouters (2022) incluem "menções a usuários" como uma *feature* de tweet, refletindo a dinâmica de interação na rede.

- d. Características temporais:** Pesquisas anteriores sugeriram que o padrão temporal da produção e do consumo de conteúdo tem o potencial de fornecer informações cruciais sobre campanhas on-line e seu desenvolvimento (Ghosh; Surachawala; Lerman, 2011; Ferrara *et al.*, 2016; Chavoshi; Hamooni; Mueen, 2016). Com o objetivo de capturar esse sinal, múltiplos atributos temporais relacionados à atividade do usuário podem ser avaliados, englobando taxas médias de produção de *posts* em diversos intervalos temporais, bem como distribuições de intervalos entre eventos. Neste caso, relacionam-se aos padrões de tempo e frequência da atividade do usuário. A idade do post, ou o tempo transcorrido desde sua publicação, é uma variável utilizada por Bhattacharya *et al.* (2014) para medir o engajamento. Em estudos de altmetria, Barcelos e Maricato (2017) observaram que as maiores médias de atenção on-line para artigos foram registradas entre 81 e 161 dias de exposição nas mídias sociais, embora o tempo de exposição não tenha sido um fator principal para o aumento da pontuação altmétrica. Araujo, Caran e Souza (2016) identificaram uma correlação forte e inversamente proporcional entre o tempo de publicação de um artigo e o número de menções altmétricas, apontando que artigos mais antigos tendem a receber menos menções atuais.
- e. Recursos de conteúdo e linguagem:** inúmeras pesquisas recentes têm sublinhado a significância dos recursos de conteúdo e linguagem na revelação das características das interações nas plataformas de mídias sociais (Mocanu *et al.*, 2013; Botta; Moat; Preis, 2015). Varol *et al.* (2017) recolhem estatísticas relacionadas ao tamanho e à entropia do texto presentes nas postagens no X e extraí recursos linguísticos por meio da aplicação da técnica de marcação de partes da fala, que identifica diversas categorias de componentes de linguagem natural. Neste caso, o foco está na análise do texto das postagens. Aguilera *et al.* (2023) consideram filtros de spam, palavras de baixo calão e erros de ortografia para avaliar a credibilidade do texto. Segundo Agarwal *et al.* (2011), as *features* de conteúdo podem ser representadas por contagens de características variadas, valores numéricos (como pontuações de dicionários de afetos) e valores booleanos, que predizem, por exemplo, a presença de

pontos de exclamação ou texto em letras maiúsculas; eles também distinguem *features* polares e não polarizadas, bem como aquelas que capturam estatísticas sobre partes-da-fala (POS). O Processamento de Linguagem Natural (PLN) é uma técnica fundamental para a análise textual, conforme Batrinca *et al.* (2015). Didegah (2018) distingue entre tweets originais e retweets e verifica se contêm texto adicional ou apenas o título/URL do artigo. Fang, Costas e Wouters (2022) incluem o uso de *hashtags* e menções a outros usuários como *features* de postagens. Além disso, Bhattacharya *et al.* (2014) utilizam *hashtags* e URLs como variáveis em seus modelos.

f. Características de sentimento: a análise de sentimento representa uma ferramenta de grande relevância para a descrição das emoções veiculadas em um fragmento de texto e, de modo mais abrangente, para a avaliação da atitude ou do estado de ânimo presentes em uma conversa completa. O sentimento extraído de discussões nas plataformas de mídias sociais tem sido empregado na previsão de eventos do mundo *off-line*, abrangendo, por exemplo, as flutuações no mercado financeiro (Bollen; Mao; Zeng, 2011), e tem sido reconhecido pelo seu impacto na disseminação de informações (Mitchell, Lewis *et al.*, 2013; Ferrara; Yang, 2015). Alguns trabalhos usam técnicas de extração de sentimento para gerar uma ampla gama de atributos relacionados ao sentimento, englobando pontuações de excitação, valência e dominância (Warriner; Kuperman; Brysbaert, 2013), pontuações de felicidade (Kloumann *et al.*, 2012), polarização e intensidade (Wilson; Wiebe; Hoffmann, 2005), análise de sentimentos e de comportamento de usuários em relação às publicações com o objetivo de identificar padrões de comportamento de *bots* (Pontes; Maricato, 2024b), bem como escores de *emoticons* (Agarwal *et al.*, 2011). Segundo Pontes e Maricato (2024), padrões de comportamento de *bots* e contas humanas na mídia X foram investigados por meio da análise de sentimentos, utilizando o ChatGPT 3.5 para classificar postagens como positivas, negativas ou neutras, demonstrando sua capacidade de compreender o contexto das mensagens e o resumo do artigo. Esse estudo revelou que, enquanto as contas humanas tendem a ter postagens predominantemente negativas (91%), os *bots* mostram uma variação percentual mais distribuída entre sentimentos positivos (8,3%), negativos (33,3%) e neutros (26,7%), o que foi interpretado como um comportamento

mais volátil ou contraditório em comparação com os humanos, que demonstram maior coerência. Apesar dos desafios da linguagem informal no X que podem levar a erros de ortografia e gramática, a análise de sentimento pode prever o impacto de artigos de pesquisa, com postagens positivas e neutras associadas a um impacto maior, conforme apontado por Barreto et al. (2023) e Hassan *et al.* (2023). Adicionalmente, Agarwal *et al.* (2011) propuseram a utilização de *features* como a soma das pontuações de polaridade de palavras e a distinção entre *features* polares e não polares para a análise de sentimento, enquanto Bhattacharya *et al.* (2014) incluíram o "sentimento positivo" como uma variável em seus modelos.

- g. Caracterização das interações do usuário:** se preocupa em caracterizar a conectividade social, o fluxo de informações e as propriedades compartilhadas dos usuários. É analisada a criação de laços sociais por contas com diferentes pontuações de *bot* e suas interações por meio de conteúdo compartilhado. É possível agrupar contas e investigar propriedades compartilhadas de usuários em cada cluster. As pontuações do *bot* são calculadas com base no modelo treinado a partir do conjunto de dados mesclado (Varol *et al.*, 2017).
- h. Conectividade social:** segundo Varol *et al.* (2017) os atributos referentes à conectividade social são: *bots* simples seguem os usuários aleatoriamente; as contas humanas geralmente seguem outras contas humanas; humanos são normalmente seguidos por outros humanos, mas também por *bots* sofisticados; *bots* tendem a seguir outros *bots* e são seguidos principalmente por *bots*; *bots* simples também podem atrair a atenção humana em casos em que humanos seguem *bots* benignos, como aqueles que compartilham notícias. A pontuação de reciprocidade de um usuário é definida como a fração de amigos que também são seguidores. Os autores observam que as contas humanas retribuem mais *bot* se correlaciona com menor reciprocidade.
- i. Fluxo de informações:** em termos de menções/repostagens no X, os *bots* simples tendem a repostar uns aos outros. *Bots* mais sofisticados repostam, mas não mencionam humanos, pois eles podem ser incapazes de se envolver em trocas significativas com humanos. Embora os humanos também repostem os *bots*, pois estes podem postar conteúdo interessante, eles não mencionam os *bots* diretamente no *post*.

j. **Outras / híbridas:** incluem abordagens mais complexas ou combinatórias. Chawla e Kapoor (2023) propõem uma estrutura híbrida para detecção de *bots* no X, combinando "digital DNA" com BERT. O conceito de "DNA digital" ou "social fingerprinting", desenvolvido por Cresci *et al.* (2016, 2017, 2018), é citado por Balestrucci *et al.* (2019), Broniatowski *et al.* (2018) e Chawla e Kapoor (2023) como um método que codifica informações comportamentais da atividade da conta de um usuário para detectar spambots e grupos de spambots. Além disso, Pontes e Maricato (2023a, 2023b) utilizaram o Botometer, um classificador de aprendizado de máquina supervisionado que distingue contas de *bots* e humanas com base em suas características.

Os autores Gilani *et al.* (2017) descobriram que os humanos geram muito mais conteúdo novo, enquanto os *bots* dependem mais de repostagens. Também observaram tendências menos intuitivas, como a propensão dos *bots* a postar no X mais URLs e a fazer upload de mídias mais volumosas (por exemplo, imagens). Analisando a interconectividade social entre *bots* e humanos para caracterizar como eles influenciam as inter-relações no X, os autores observaram que, embora as contribuições humanas sejam geralmente consideradas mais importantes através de métricas típicas (por exemplo, número de curtidas, repostagens), os *bots* ainda sustentam uma influência importante sobre a produção e propagação de conteúdo.

Segundo Gilani *et al.* (2017), duas linhas principais de pesquisa na caracterização de contas do X são relevantes: (i) análises sociais, demográficas e comportamentais de *bots* ou humanos; e (ii) o impacto dos *bots* em ambientes sociais. Em Freitas *et al.* (2015) são usadas três métricas para quantificar a presença de *bots* sociais: seguidores, pontuação de popularidade e interação baseada em mensagens (outros usuários favoritando, repostando, respondendo ou mencionando o *bot*). Eles descobriram que os *bots* nem sempre são captados com sucesso pelas defesas do X (apenas 38 dos 120 *bots* foram suspensos ao longo de 30 dias).

Em Wei e Nguyen (2023), os autores compararam o modelo proposto, chamado BOTLE, com modelos de linha de base que usam o mesmo conjunto de dados cres-2017⁷. Os autores observam que a maioria dos algoritmos/sistemas de última geração para detecção de *bots* do X exigem grande número de dados e atributos. Os sistemas estudados pelos autores requerem engenharia de atributos baseada de seis a mais

⁷ Disponível em: <http://mib.projects.iit.cnr.it/dataset.html>

de 1.000 recursos. A engenharia de atributos requer muito trabalho e é cara em termos de coleta e pré-processamento de dados. Em contrapartida, o modelo *Recurrent Neural Network* (RNN), proposto pelos autores Wei e Nguyen (2023), não depende de nenhuma engenharia de atributos e usa apenas o conteúdo textual dos *posts*.

Os resultados reportados pelos autores dos trabalhos apresentados ilustram desafios contemporâneos na detecção de *bots* como, por exemplo, a dependência de técnicas que exigem extensiva engenharia de atributos e grande volume de dados estruturados. Nesse sentido, observa-se uma tendência crescente na adoção de abordagens que buscam reduzir a dependência de extensa engenharia de atributos, destacando-se, nesse cenário, os modelos fundamentados exclusivamente na análise de conteúdo textual. Tal movimento evidencia a valorização de métodos mais generalizáveis e adaptativos, capazes de operar com dados menos estruturados e, portanto, mais alinhados à dinâmica heterogênea e em constante transformação das redes sociais.

3 METODOLOGIA E PROCEDIMENTOS METODOLÓGICOS

O presente capítulo descreve a metodologia e os procedimentos metodológicos adotados no desenvolvimento da pesquisa, contemplando desde a fundamentação metodológica até a execução prática das etapas necessárias para a construção e validação do modelo proposto. Inicialmente, são apresentadas a abordagem metodológica e a definição dos procedimentos adotados, seguidas pela compreensão do negócio e pela preparação dos dados utilizados. Na sequência, detalha-se o processo de definição e seleção da amostra de artigos, a construção do *dataset* de treinamento e as etapas de desenvolvimento do modelo de inteligência artificial, abrangendo a engenharia de *features*, a modelagem supervisionada com XGBoost, o processo de treinamento e otimização, bem como a classificação final das contas. Esse percurso metodológico busca garantir a consistência científica e a reproduzibilidade da pesquisa.

3.1 METODOLOGIA

A presente pesquisa caracteriza-se como sendo de **natureza aplicada**, uma vez que visa à geração de conhecimentos com utilidade prática, por meio do desenvolvimento de processos e soluções tecnológicas baseadas em inteligência artificial (Prodanov, 2012). O objetivo central é propor uma solução computacional para um problema específico e concreto que é a identificação das contas automatizadas no X, que interferem no cálculo do score altmétrico de publicações científicas.

Do ponto de vista da abordagem metodológica do problema, a presente pesquisa classifica-se como **qualiquantitativa**, uma vez que combina técnicas estatísticas e computacionais com análise interpretativa dos padrões de comportamento observados (Prodanov, 2012). A vertente quantitativa está presente na utilização de dados numéricos referentes às contas em redes sociais, os quais requerem técnicas estatísticas e de inteligência artificial para serem processados, classificados e analisados. Por sua vez, a dimensão qualitativa manifesta-se na interpretação dos comportamentos e padrões identificados, considerando aspectos contextuais e não mensuráveis diretamente, os quais são fundamentais para

compreender a atuação de contas automatizadas no ecossistema de disseminação científica.

Segundo Fonseca (2002), como as amostras geralmente são grandes e consideradas representativas da população, os resultados são tomados como se constituíssem um retrato real de toda a população alvo da pesquisa. A pesquisa quantitativa se centra na objetividade. Influenciada pelo positivismo, considera que a realidade só pode ser compreendida com base na análise de dados brutos recolhidos com auxílio de instrumentos padronizados e neutros.

Quanto aos objetivos, esta pesquisa é de cunho **exploratório e descritivo**, pois busca compreender e descrever um fenômeno em maior profundidade. É exploratória por buscar aprofundar o entendimento de um fenômeno ainda pouco sistematizado (Marconi; Lakatos, 2016). É exploratória, ao investigar a atuação de *bots* no contexto altmétrico, e descritiva por se dedicar à caracterização empírica dos perfis, comportamentos e padrões de engajamento das contas analisadas.

Segundo Marconi e Lakatos (2016), estudos exploratórios descritivos têm como finalidade descrever com precisão determinados fenômenos, ampliando o entendimento sobre suas múltiplas dimensões. Nesses tipos de investigação, é comum a utilização de abordagens quantitativas, qualitativas ou combinadas, acompanhadas da coleta de informações detalhadas, obtidas por meio da observação direta ou de outras técnicas de investigação. Nessas circunstâncias, valoriza-se a consistência e a relevância dos dados coletados.

Quanto aos procedimentos técnicos, ou seja, a forma de obtenção e de análise dos dados utilizados para responder aos objetivos da pesquisa, quando é necessário traçar um modelo conceitual e operativo, denominado de design (Prodanov, 2012), adota-se uma abordagem **empírico-computacional**. Essa classificação é mais adequada que a de pesquisa experimental em sentido estrito, uma vez que o estudo não realiza manipulação direta de variáveis em ambiente controlado, como tradicionalmente se exige nesse tipo de experimento (Lakatos; Marconi, 2016). Em vez disso, a pesquisa envolve a formulação e o teste de hipóteses por meio de modelagem computacional e avaliação sistemática de dados observacionais, obtidos a partir da plataforma X. Trata-se, portanto, de um estudo que se ancora em métodos empíricos — baseados em observações e medições reais — combinados com técnicas computacionais para simulação, análise preditiva e classificação automática,

em conformidade com práticas descritas na ciência de dados e na chamada ciência da ciência (Liu *et al.*, 2023).

No que diz respeito às técnicas e ferramentas utilizadas, esta pesquisa adotou a integração de técnicas de estatística, mineração de dados e algoritmos de aprendizagem de máquina para lidar com grandes volumes de informações extraídas da rede social. Considerando a complexidade envolvida nesse tipo de tarefa — que exige etapas sucessivas de entendimento do problema, tratamento e modelagem dos dados, validação dos resultados e análise crítica — tornou-se necessário adotar uma metodologia que estruturasse todo o ciclo de vida do processo analítico de forma sistemática e replicável.

Para isso, foi empregada a metodologia *CRoss-Industry Standard Process for Data Mining* (CRISP-DM), amplamente reconhecida por sua flexibilidade e aplicabilidade em diferentes contextos de mineração de dados. A adoção do modelo CRISP-DM se justifica pela evolução da área de *Data Mining* (DM) que, ao longo dos anos, evidenciou a necessidade de normas e padrões metodológicos capazes de orientar projetos com grande volume e diversidade de dados. Modelos tradicionais não contemplavam de maneira adequada os requisitos de projetos dessa natureza. Em resposta a essa lacuna, um consórcio de empresas fornecedoras e usuárias de soluções de data mining iniciou, ainda na década de 1990, o desenvolvimento do CRISP-DM — um modelo de processo não proprietário, neutro e de domínio público, que descreve as práticas mais comuns em projetos de mineração de dados (Shearer, 2000; Ramos *et al.*, 2020).

3.2 PROCEDIMENTOS METODOLÓGICOS

A escolha metodológica constitui um eixo fundamental para a validade e a robustez desta pesquisa, uma vez que define as estratégias adotadas para transformar os dados coletados em evidências capazes de responder às questões de investigação. Considerando a natureza do problema e a complexidade dos dados analisados, optou-se pela aplicação do modelo CRISP-DM, cuja relevância está consolidada tanto na literatura acadêmica quanto em práticas de mercado.

Nesta seção, são descritos os procedimentos metodológicos empregados, enfatizando a adaptação proposta por Pontes e Maricato (2022) e a forma como suas etapas foram operacionalizadas ao longo do estudo. Tal abordagem possibilitou

alinhar a estrutura teórico-metodológica da pesquisa a um processo sistemático e replicável, garantindo transparência e rigor científico.

Com base nessa adaptação, o modelo CRISP-DM foi aplicado nesta investigação em três etapas principais:

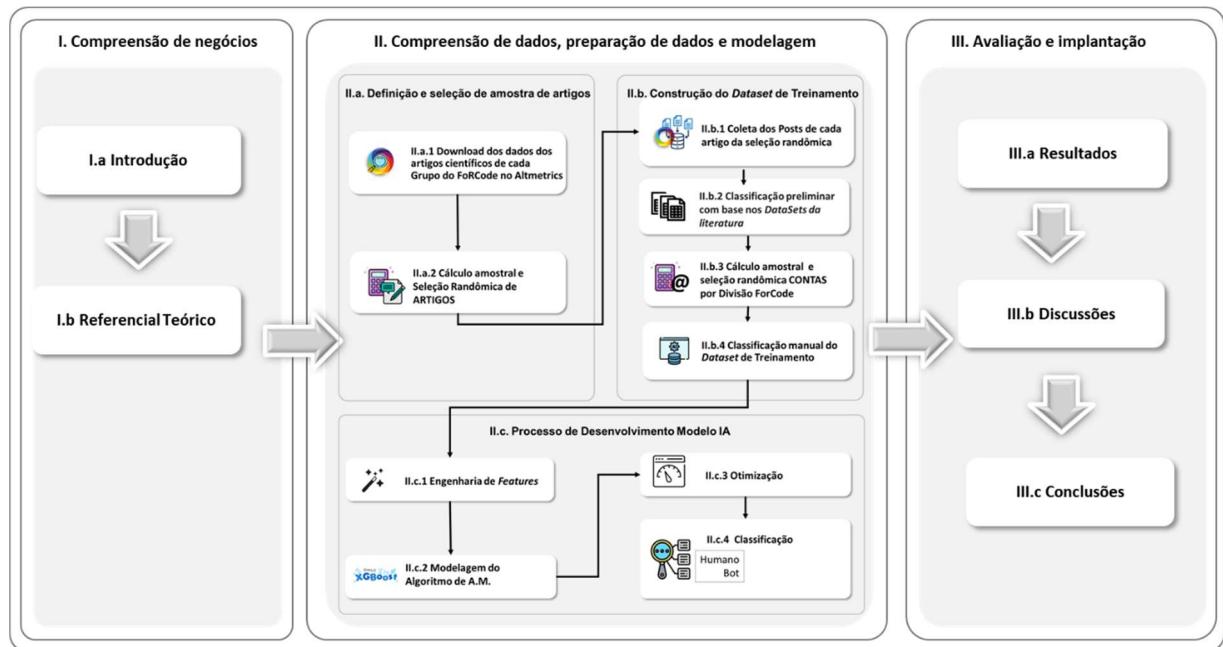
(1) a compreensão do negócio, representada pelas seções de introdução e referencial teórico da tese, que contextualizam o problema de pesquisa;

(2) a compreensão e preparação dos dados, bem como a modelagem, descritas nos procedimentos metodológicos adotados; e

(3) a avaliação e implantação dos resultados, abordadas nas seções de resultados, discussões e conclusões da tese.

A Figura 1 apresenta o processo metodológico para classificação de contas do X, especificando o ciclo de vida do processo de mineração de dados baseado na metodologia CRISP-DM.

Figura 1 - Ciclo de vida do processo de mineração de dados para classificação de contas X baseado na metodologia CRISP-DM



Fonte: Dados da Pesquisa (2025).

A seguir são detalhados os procedimentos metodológicos que compõem o ciclo de vida do processo adotado nesta pesquisa.

3.3 COMPREENSÃO DE NEGÓCIOS

Segundo Shearer (2000), a fase mais importante de um projeto de mineração de dados, a fase de entendimento de negócio, foca em entender o objetivo do projeto a partir da perspectiva do negócio, convertendo esse conhecimento em uma definição de problemas de mineração de dados e, em seguida, desenvolvendo um plano preliminar projetado para atingir os objetivos. Para entender quais dados e como devem ser analisados, é vital que os profissionais envolvidos entendam completamente os negócios para os quais estão encontrando uma solução. A fase de entendimento dos negócios envolve várias etapas importantes, incluindo a determinação dos objetivos de negócios, a avaliação da situação, a determinação das metas e a produção do plano do projeto.

Para a compreensão do problema, além de um estudo bibliográfico que constituiu o referencial teórico deste trabalho, alguns estudos preliminares foram desenvolvidos para testar técnicas e hipóteses com o objetivo de validar a metodologia proposta. Os dois principais estudos resultaram nos artigos de Pontes e Maricato (2022) e Pontes e Maricato (2023).

No primeiro trabalho Pontes e Maricato (2022) foi desenvolvido um *workflow* de classificação baseado em um *dataset* genérico⁸ e no algoritmo de classificação KNN. Neste, os autores observaram que o uso de um *dataset* genérico, ou seja, onde existem contas rotuladas entre *bots*, ciborgues e humanas em contextos variados, não é suficiente para identificar *bots* que atuam com dados de atenção altmétrica.

O segundo trabalho de Pontes e Maricato (2023) utilizou a ferramenta de classificação *Botometer*⁹. Neste, os resultados expressaram que o *Botometer* não classificou de forma eficiente contas automatizadas que claramente influenciaram o índice altmétrico dos artigos estudados. Em ambos os trabalhos, evidenciou-se indícios de limitação dos métodos generalistas quando aplicados ao contexto altmétrico, destacando a necessidade de abordagens específicas para investigar a atuação e o impacto dos *bots* na disseminação científica mediada por redes sociais.

⁸ Entende-se por "dataset genérico" o conjunto de dados composto por contas que não seguem uma temática específica, abrangendo publicações de diversos assuntos na plataforma X.

⁹ O Botometer é um sistema de aprendizado de máquina que avalia contas da plataforma X, atribuindo escores que indicam a probabilidade de serem operadas de forma automatizada.

3.4 COMPREENSÃO E PREPARAÇÃO DOS DADOS

A etapa de compreensão e preparação dos dados constitui um componente fundamental do processo baseado na metodologia CRISP, sendo responsável por garantir qualidade, consistência e relevância das informações que serão analisadas. Essa fase tem início com a coleta preliminar dos dados e se estende até a verificação de sua integridade e utilidade para os objetivos do estudo.

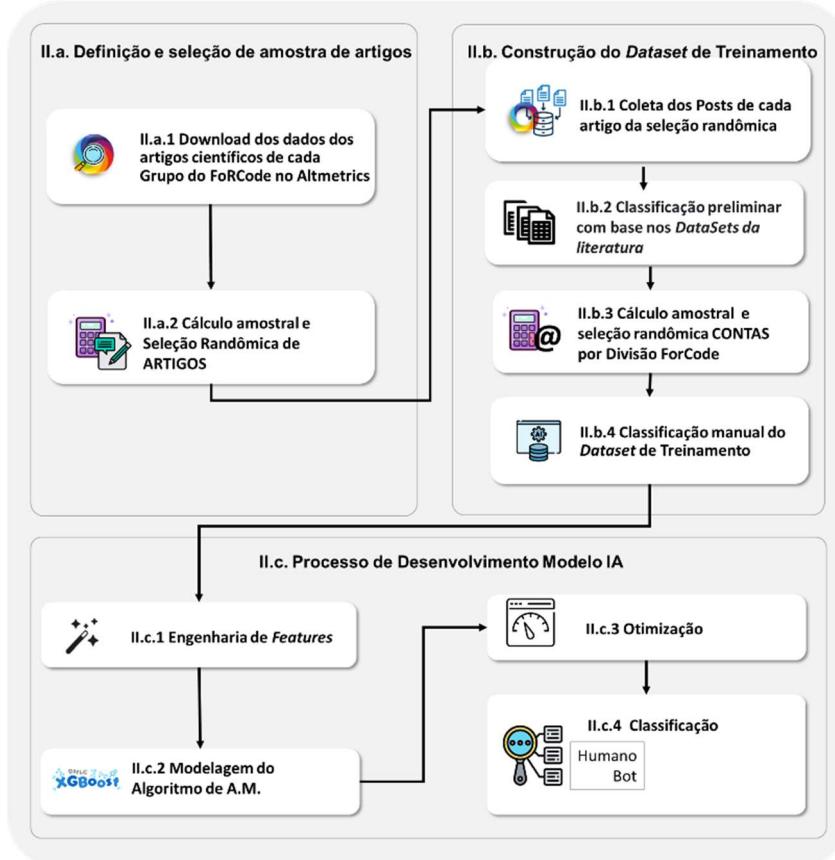
De acordo com Shearer (2000), a fase de entendimento dos dados envolve uma série de procedimentos destinados a permitir que o analista compreenda o conteúdo e a estrutura do conjunto de dados, identifique possíveis problemas relacionados à sua qualidade, descubra padrões e informações iniciais relevantes e identifique subconjuntos potencialmente interessantes. Essa compreensão preliminar é essencial para a formulação de hipóteses e para a definição das estratégias de análise subsequentes.

Shearer (2000) propõe que essa fase seja estruturada em quatro etapas principais:

- (i) Coleta de dados iniciais;
- (ii) Descrição dos dados;
- (iii) Exploração dos dados;
- (iv) Verificação da qualidade dos dados.

Considerando as particularidades deste estudo, as etapas foram ajustadas de modo a alinhar-se aos objetivos e ao contexto da pesquisa, conforme ilustrado no infográfico da Figura 2 que apresenta o fluxograma detalhado do processo metodológico de construção e aplicação de um modelo de IA voltado à classificação de contas do X que divulgam artigos científicos, com o objetivo de identificar se são humanos ou *bots*.

Figura 2 - Infográfico com as atividades da etapa de compreensão preparação de dados e modelagem



Fonte: Dados de Pesquisa (2025)

A imagem está dividida em três grandes seções, cada uma com suas respectivas etapas:

II.a. Definição e seleção de amostra de artigos:

- II.a.1 Download dos artigos científicos de cada Grupo do ForCode no Altmetrics;
- II.a.2 Cálculo amostral de número de artigos e Seleção Randômica de artigos.

II.b Construção do Dataset de Treinamento:

- II.b.1 Coleta dos Posts de cada artigo da seleção randômica;
- II.b.2 Classificação preliminar com base nos datasets da literatura;
- II.b.3 Cálculo amostral do número de contas e Seleção Randômica das CONTAS por Divisão do ForCode¹⁰;

¹⁰ Disponível em: <https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-research-classification-anzsrc/latest-release>

II.b.4 Classificação manual do *dataset* de Treinamento.

II.c Processo de Desenvolvimento do Modelo de IA

II.c.1 Engenharia de *features*;

II.c.2 Modelagem do Algoritmo de AM com o XGBOOST;

II.c.3 Otimização;

II.c.4 Classificação.

A seguir, cada uma dessas atividades será detalhada, com a finalidade de explicitar os procedimentos metodológicos adotados nesta etapa da pesquisa.

3.4.1 II.a. Definição e seleção de amostra de artigos

A etapa de definição e seleção da amostra de artigos constitui a base inicial para a construção do *dataset* utilizado neste estudo, uma vez que assegura a representatividade e a robustez estatística dos dados analisados. Nessa fase, foram estabelecidos critérios de recuperação e seleção de artigos científicos a partir da classificação ForCode no Altmetrics, seguidos de procedimentos de cálculo amostral e de seleção randômica. Tal processo visa garantir a diversidade temática dos registros e a validade metodológica necessária para a etapa subsequente de coleta de dados e análise.

3.4.1.1 II.a.1 Download dos artigos científicos de cada Grupo do ForCode no Altmetrics

A categorização dos artigos selecionados foi baseada na Classificação Padrão de Pesquisa da Austrália e Nova Zelândia (ANZSRC) (*Australian and New Zealand Standard Research Classification*) que define os Códigos de Área de Pesquisa (FoR) (*Field of Research*) como critério de classificação temática das publicações científicas, possibilitando a segmentação dos dados conforme as áreas do conhecimento, que são os mesmos utilizados pela plataforma Altmetric.com.

A ANZSRC foi desenvolvida com o propósito de apoiar a coleta, a análise e a disseminação de estatísticas relacionadas a atividades de pesquisa e de desenvolvimento experimental (P&D) na Austrália e na Nova Zelândia. Essa taxonomia oferece uma estrutura padronizada para a classificação das atividades científicas, favorecendo a organização dos dados dos artigos, permitindo a

comparação entre diferentes domínios de pesquisa. A adoção dessa classificação contribui para a precisão analítica e para a identificação de padrões e tendências no impacto social e científico das publicações dentro de distintas áreas do conhecimento.

A tabela utilizada nesta etapa corresponde ao arquivo em formato Excel denominado **DAD_00_Tabela_ForCode.xlsx**. Este arquivo contém a estrutura de classificação utilizada na codificação das áreas de conhecimento conforme o sistema ANZSRC 2020. A planilha é composta por informações relativas às 23 divisões principais de áreas de conhecimento, apresentadas por meio de colunas com seus respectivos códigos e descrições. Além disso, inclui os dados referentes aos grupos, que representam as subáreas de conhecimento, também identificados por seus respectivos códigos e descrições. Ao todo, são contemplados 212 grupos, permitindo uma categorização detalhada e padronizada dos campos de pesquisa.

De posse da estrutura de códigos das áreas de conhecimento, foi realizada a coleta dos dados dos artigos de cada área. Foram coletados os arquivos de *Research Output* e *Save Research* de cada grupo da tabela FoR com os dados dos artigos do grupo. Através da opção *Save Research* do Altmetric.com foi possível fazer o download de um arquivo que reúne os dados de todos os artigos de uma determinada pesquisa. Neste caso, o critério da pesquisa deve ser o número do grupo FoR e a restrição de que fossem selecionados apenas artigos científicos.

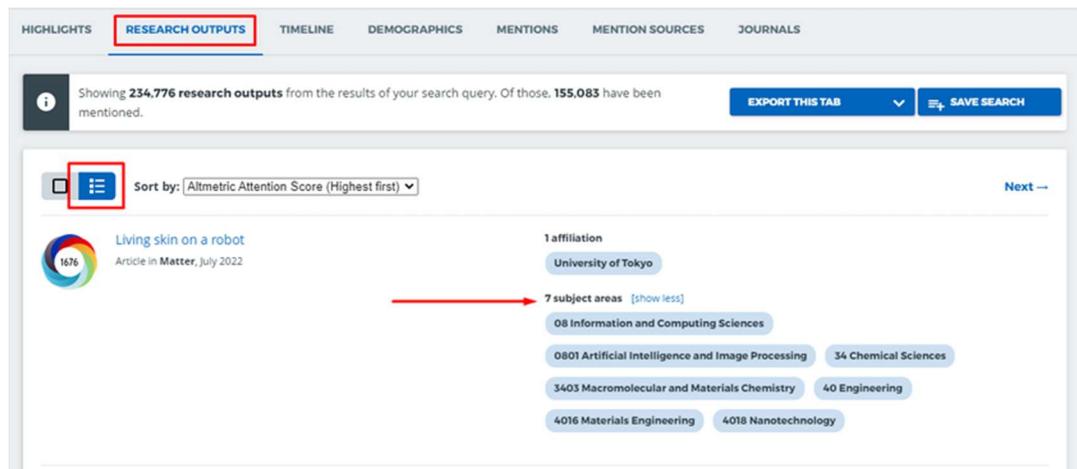
Os arquivos oriundos do *Research Output*, disponibilizado pelo Altmetric é composto pelos dados: *Altmetric Attention Score*, *Title*, *Journal/Collection Title*, *Journal ISSNs*, *Authors at my Institution*, *Departments*, *Output Type*, *OA Status*, *OA Type*, *Subjects (FoR)*, *Sustainable Development Goals*, *Affiliations (GRID)*, *Funder*, *Publication Date*, *DOI*, *ISBN*, *National Clinical Trial ID*, *URI*, *PubMed ID*, *PubMedCentral ID*, *Handle.net IDs*, *ADS Bibcode*, *arXiv ID*, *RePEc ID*, *SSRN*, *URN*, *News mentions*, *Blog mentions*, *Policy mentions*, *Patent mentions*, *X mentions*, *Peer review mentions*, *Weibo mentions*, *Facebook mentions*, *Wikipedia mentions*, *Google+ mentions*, *LinkedIn mentions*, *Reddit mentions*, *Pinterest mentions*, *F1000 mentions*, *Q&A mentions*, *Video mentions*, *Syllabi mentions*, *Number of Mendeley readers*, *Number of Dimensions citations*, *Details Page URL*, *Badge URL*, *Publisher Names*

A seleção dos grupos foi feita com base na disponibilização dos arquivos de artigos de cada grupo pelo download do Altmetric¹¹. A Figura 3 destaca a página da

¹¹ Disponível em: https://www.altmetric.com/explorer/highlights?field_of_research_codes%5B%5D=42

Plataforma *Altmetric* que apresenta a opção de fazer download dos artigos através da opção *export tab*¹².

Figura 3 - Imagem da Tela do *Export Research Outputs* no Altmetric no qual é realizado o download do arquivo



Fonte: Altmetric (2023).

Desta forma, a partir dos 212 grupos categorizados segundo o sistema FoR, foi possível realizar o download dos dados de 144 grupos na plataforma Altmetric.com. Durante o processo de coleta, constatou-se uma limitação técnica relacionada à instabilidade do sistema de origem, que impossibilitou a obtenção dos dados de 69 grupos. Embora tenham sido realizadas múltiplas tentativas de extração em diferentes dias e condições de acesso, o problema não foi superado, o que levou à exclusão desses casos do corpus analisado. Essa limitação decorre de fatores externos ao controle do pesquisador e deve ser considerada ao interpretar os resultados.

Os arquivos extraídos foram organizados em um diretório específico (denominado DRV_01_AltmetricResearcOutput)¹³ e salvos individualmente conforme o código identificador de cada grupo, seguindo o padrão de nomenclatura: “0-xxxx-Altmetric - Research Output.csv”, onde xxxx representa código o grupo coletado. No total, foram coletados dados referentes a 8.565.607 artigos, consolidados no arquivo “DAD_02_CalcloAmostraArtigos.xlsx”, os quais serviram de base para a próxima etapa da pesquisa.

¹² Disponível em: <https://help.altmetric.com/support/solutions/articles/6000240906-searching-by-subject-area>

¹³ No Apêndice B, é apresentada a estrutura de diretórios em que os datasets e scripts dos algoritmos foram armazenados.

3.4.1.2 II.a.2 Cálculo amostral de número de artigos e sua seleção randômica

A fim de garantir a representatividade da amostra e mitigar possíveis vieses na análise, adotou-se uma estratégia de amostra probabilística estratificada, por meio da seleção randômica de artigos científicos, tendo como critério de estratificação os 144 grupos previamente identificados.

Essa abordagem visou garantir a representatividade estatística da diversidade disciplinar contemplada no estudo, ao mesmo tempo que permitiu uma análise detalhada das dinâmicas de disseminação em escala reduzida e controlada.

Para cada grupo, foi calculado o tamanho amostral com base no número total de artigos com score altmétrico e menções no X, considerando um nível de confiança de **(99%) e da taxa e erro (5%)**. O arquivo **DAD_02_CalcularAmostraArtigos.xlsx**, com a consolidação da quantidade de artigos por grupo (colunas: CODFOR e TOTAL), foi a fonte de entrada de dados para o algoritmo implementado em Python (**ALG_01_CalcularAmostraArtigos.py**), desenvolvido com o objetivo de estimar, de forma automatizada, o tamanho ideal da amostra de artigos a ser analisada em cada grupo (subárea de conhecimento). Como resultado do processamento computacional, obteve-se um tamanho amostral total de 93.528 artigos, extraída de uma população de 8.565.607 artigos.

Após a definição do tamanho da amostra para cada grupo, os *Digital object identifier* (DOI) foram selecionados aleatoriamente por meio da execução de um algoritmo desenvolvido em Python, denominado **ALG_02_SelecionarRandomArtigos.py**, o qual automatizou o processo de seleção randômica dos artigos. Como resultado desse algoritmo, foi gerado o arquivo **DAD_03_ArtigosSelecionadosRandom.xlsx**, que reúne, de forma consolidada, o conjunto completo de artigos selecionados por grupo, juntamente com seus respectivos dados. Os dados referentes a esses DOIs foram extraídos dos arquivos obtidos na base *Research Output* e organizados em um único documento, que constitui a amostra final de artigos analisada neste estudo.

Destaca-se que este arquivo comprehende um total de 93.528 artigos científicos, correspondendo à amostragem integral a ser utilizada nas etapas subsequentes da análise. A escolha randômica teve como objetivo reduzir vieses sistemáticos associados à popularidade de temas ou à visibilidade de determinados periódicos, promovendo, assim, maior robustez na análise inferencial dos dados altmétricos.

3.4.2 II.b Construção do *dataset* de treinamento

A etapa de construção do *dataset* de treinamento constitui um processo central para a detecção de *bots*, uma vez que define a base empírica necessária ao desenvolvimento e validação do modelo de inteligência artificial. Nessa fase, são descritos os procedimentos de coleta de *posts* vinculados aos artigos previamente selecionados, a classificação preliminar fundamentada em *datasets* da literatura, bem como o cálculo amostral e a seleção randômica de contas segundo as divisões do ForCode. Por fim, contempla-se a classificação manual do conjunto de dados, assegurando maior robustez e confiabilidade ao corpus utilizado no treinamento do modelo.

3.4.2.1 II.b.1 Coleta dos *posts* de cada artigo da seleção randômica

A construção do *dataset* de treinamento teve início com a coleta de *posts* publicados na plataforma X que mencionavam artigos científicos previamente selecionados por amostragem probabilística estratificada. A amostra foi composta por 93.528 artigos, distribuídos entre 144 grupos de subáreas e 22 divisões do conhecimento, assegurando representatividade disciplinar. A partir dessa base, foram extraídos 758.288 *posts* que faziam referência direta aos artigos, identificados por meio de DOIs, títulos ou URLs indexadas. Essas publicações foram emitidas por 362.395 contas únicas, apresentando alto grau de capilaridade na circulação e no engajamento com a produção científica nas redes sociais.

Cada *post* e respectiva conta foram vinculados ao artigo que motivou sua publicação, estabelecendo uma rede relacional entre conteúdo científico e sua repercussão on-line. Essa estruturação viabiliza, nas etapas analíticas subsequentes, a exploração integrada entre características da produção acadêmica, padrões de disseminação digital e possíveis indícios de comportamento automatizado. A organização dos dados respeitou a estratificação original por divisão e grupo de área do conhecimento, o que permite comparações consistentes entre domínios científicos distintos.

Essa etapa constitui a base empírica do estudo, fornecendo os subsídios necessários para os módulos analíticos que visam compreender os fatores associados

à visibilidade altmétrica, bem como investigar a atuação de agentes automatizados (*bots*) na difusão da ciência em ambientes digitais.

A obtenção dos *posts* foi realizada a partir da definição da amostra de artigos, através de um algoritmo desenvolvido em Python para coletar, por meio da plataforma Altmetric, as postagens no X que mencionaram, comentaram ou compartilharam os artigos selecionados na seção 3.4.1. Devido às limitações de exportação das postagens no ambiente da plataforma *Altmetric*, as contas foram coletadas através de *web scraping*¹⁴. A *Altmetric* apresenta na web a visualização apenas dos últimos 10.000 *posts* de cada artigo. Foram coletados a identificação da conta, o texto do *post*, a data e hora de publicação e o número de seguidores da conta que postou.

Com base no artigo "*Web Scraping for Research Legal, Ethical, Institutional, and Scientific Considerations*" (Brown *et al.*, 2024), pode-se argumentar fortemente a favor do uso de *web scraping* como um método essencial e cada vez mais relevante para a coleta de dados em pesquisas científicas, especialmente no cenário atual de acesso restrito aos dados on-line. O artigo destaca que pesquisadores de diversas áreas frequentemente utilizam dados da internet para conduzir pesquisas, produzindo evidências relevantes sobre o comportamento humano. No entanto, a crescente importância do uso de grande quantidade de texto para o treinamento de inteligências artificiais generativa levou plataformas a restringirem significativamente o acesso a dados por meio de canais oficiais, como APIs (*Application Programming Interfaces*) (Brown *et al.*, 2024).

A crescente restrição ao acesso público, para fins acadêmicos, a dados de plataformas digitais tem impactado diretamente a viabilidade de pesquisas empíricas no campo. Casos emblemáticos, como a descontinuação do acesso gratuito à API do X para pesquisadores — substituído por um modelo de cobrança com valores proibitivos — e o anúncio do encerramento do *CrowdTangle* pela Meta, evidenciam essa tendência de redução progressiva na oferta de dados por vias institucionais e transparentes. Como consequência, muitos pesquisadores têm recorrido ao uso de técnicas de *web scraping* como alternativa metodológica para viabilizar a coleta de dados necessários ao desenvolvimento de suas investigações.

O *web scraping* é definido, em Brown *et al.*, 2024, como a coleta automatizada de dados disponíveis na internet, por meio da captura de informações destinadas à

¹⁴ Processo de mineração de dados que permite a extração de dados de sites da web de maneira automatizada e em larga escala.

exibição em páginas web ou aplicativos. Embora não seja uma prática recente — remontando aos primórdios da web — tem se consolidado como uma metodologia central de coleta de dados em diversas áreas, sendo empregada tanto para fins comerciais quanto para interesse público, como no jornalismo investigativo, no ativismo digital e em pesquisas acadêmicas. A ampla utilização da *web scraping* por diferentes setores é avultada pela expressiva quantidade de estudos baseados em conjuntos de dados obtidos por meio dessa técnica (Quadro 1), o que reforça sua relevância para as ciências sociais e para outras áreas do conhecimento.

Embora o ambiente regulatório em torno do *web scraping* permaneça incerto, Brown *et al.* (2024) argumentam que, dada a crescente necessidade de dados para pesquisa e a restrição de outras fontes, o *web scraping* se tornará ainda mais frequente. Para auxiliar os pesquisadores, o artigo propõe um quadro abrangente para o *web scraping* na pesquisa em ciências sociais, examinando as considerações legais, éticas, institucionais e científicas.

Brown *et al.* (2024) não apenas reconhecem a importância do *web scraping*, mas também fornecem orientações importantes para que os pesquisadores conduzam essa prática de forma científica, legítima e ética, minimizando riscos legais e maximizando o impacto de sua pesquisa. Eles abordam questões críticas que os pesquisadores devem considerar ao usar o *web scraping* como ferramenta de coleta de dados, oferecendo um *checklist* para ajudar a pensar nessas considerações.

Em suma, Brown *et al.* (2024) defendem o uso do *web scraping* como um método de coleta de dados valioso e, muitas vezes, necessário na pesquisa científica, especialmente em um contexto de crescente fechamento de APIs e outras fontes de dados formais. A capacidade de acessar e analisar dados da web de forma automatizada permite investigar fenômenos sociais e on-line que, de outra forma, seriam inacessíveis devido às restrições impostas pelas plataformas.

A análise dos estudos que empregaram *web scraping* em pesquisas altmétricas apresentada no Quadro 1 revela uma crescente preocupação com a abrangência, a qualidade e a volatilidade dos dados coletados por agregadores como Altmetric, PlumX e Crossref Event Data. O uso de técnicas automatizadas de raspagem de dados surge como uma estratégia metodológica fundamental para auditar fontes externas, complementar dados não disponíveis via API e examinar a persistência de menções em mídias sociais e blogs. Ortega (2020), por exemplo, utilizou scripts de raspagem para auditar mais de 50 mil links extraídos diretamente das páginas de

detalhes das plataformas, identificando problemas como links quebrados e inconsistências na cobertura. O Quadro 1 apresenta essa e outras referências que empregaram o *Web Scraping* para coleta de dados em estudos altmétricos.

Quadro 1 - Referências sobre uso de Web Scraping em pesquisas altmétricas

Referência	Uso de web scraping	Fontes coletadas
ORTEGA, J. L. Availability and audit of links in altmetric data providers: link checking of blogs and news in Altmetric.com, Crossref Event Data and PlumX. <i>Journal of Altmetrics</i> , v. 2, n. 1, p. 4, 2019. DOI: 10.29024/joa.14	Raspagem HTML das páginas /details/... de Altmetric.com e PlumX para extrair 51.000 URLs de blogs/notícias	Blogs e portais de notícias listados em Altmetric, PlumX e CED
ROBINSON-GARCÍA, N. et al. Nuevos datos, nuevas posibilidades: revelando el interior de Altmetric.com. <i>El profesional de la información</i> , v. 23, n. 4, p. 359-366, 2014. DOI: 10.3145/epi.2014.jul.03	Crawler varreu o site Altmetric.com para capturar JSON embutido em páginas badge além da API	Tweets, Mendeley, Facebook, Blogs, News, Reddit etc.
FANG, Z.; DUDEK, J.; COSTAS, R. Facing the volatility of tweets in altmetric research. <i>Journal of the Association for Information Science & Technology</i> , v. 73, n. 8, p. 1192-1195, 2022. DOI: 10.1002/asi.24624	Raspagem diária das páginas Web de 37.918 tweets para verificar exclusões e variações de menções	Páginas públicas de tweets mencionando DOIs de artigos COVID-19
ORTEGA, J. L. The coverage of blogs and news in the three major altmetric data providers. <i>Proc. 17th Int. Conf. on Scientometrics & Informetrics</i> , 2019.	Scrapers percorreram 90.000 páginas em Altmetric.com, PlumX e Crossref Event Data para comparar cobertura	Blogs/news em três provedores
STRIJKER, S. J.; SUGIMOTO, C. R. Stranger things: the vanishing of the Altmetric Attention score values in LIS publications. <i>Scientometrics</i> , v. 129, 2024. DOI: 10.1007/s11192-024-05011-5	Raspagem periódica das páginas /details/doi/... de Altmetric para monitorar oscilações do AAS ao longo de um ano	Páginas Altmetric "/details/doi/..."

Fonte: Dados de Pesquisa (2025).

Outros estudos, como o de Robinson-García *et al.* (2014), demonstram como o *web scraping* pode ser combinado à extração de dados em JSON embutidos em páginas da web para ampliar a visibilidade dos dados que não são acessíveis por meios oficiais. Essa abordagem permitiu investigar a composição granular das métricas, incluindo fontes como Facebook, Mendeley e blogs científicos. Já Fang, Dudek e Costas (2022) empregaram a raspagem de páginas públicas de *posts* para monitorar a volatilidade das menções altmétricas ao longo do tempo, evidenciando a fragilidade das métricas diante de deleções e alterações nas plataformas sociais.

De modo geral, o uso de *web scraping* nesses trabalhos mostra-se metodologicamente legítimo e cada vez mais necessário, especialmente diante das limitações impostas por APIs restritivas e pela natureza dinâmica dos dados sociais. Além de ampliar o escopo da coleta, essas técnicas permitem análise mais crítica da confiabilidade e estabilidade dos dados altmétricos, fornecendo subsídios valiosos para a avaliação da disseminação científica em ambientes digitais.

A coleta dos *posts* vinculados à amostra de artigos selecionados aleatoriamente foi realizada por meio da execução do script em Python **ALG_03_ColetarPosts.py**. O referido algoritmo foi programado para percorrer automaticamente os DOIs constantes na planilha **DAD_03_ArtigosSelecionadosRandom.xlsx**, acessando as páginas correspondentes a cada publicação e extraiendo os *posts* vinculados a elas na plataforma analisada. Como resultado dessa coleta automatizada, obtiveram-se 758.288 *posts*, os quais foram organizados e armazenados na planilha **DAD_04_PostsColetados**, com o objetivo de subsidiar as análises posteriores da pesquisa.

Os dados coletados possibilitaram a realização de análises, como o número de contas únicas de usuários da mídia social, a quantidade de postagens no X gerados por toda conta, e a ocorrência de postagens sobre múltiplos artigos por uma mesma conta. Essas análises visam identificar possíveis padrões de comportamento automatizado, permitindo a detecção de contas com indícios de atividade humana ou *bot*, ou seja, não orgânica.

3.4.2.2 II.b.2 Classificação preliminar com base nos Datasets da literatura

Na aprendizagem de máquina supervisionada, o *dataset* de treinamento é um ponto de extrema relevância. No estudo preliminar apresentado no artigo Pontes e Maricato (2022), utilizou-se, como base para treinamento, o *dataset* disponibilizado por Gutiérrez, (2020) no *Kaggle*¹⁵. Este conjunto de dados é composto por 37.437 registros de diferentes contas de usuários no X. Cada linha contém o ID do usuário e a variável de destino ou resposta. A variável de destino é denotada como *account_type* e tem valores únicos (*bot* ou humano). Este conjunto de dados tem

¹⁵ Disponível em: <https://www.kaggle.com/datasets/davidmartngutierrez/twitter-bots-accounts>

25.013 contas genéricas de usuários anotadas/categorizadas como contas humanas e 12.425 como *bots*. Já no estudo preliminar, apresentado no artigo Pontes e Maricato, (2023) foi utilizado o classificador *Botometer*, que também é genérico e segmenta os atores entre *bot* e humano.

Os estudos desenvolvidos por Pontes e Maricato (2022; 2023) evidenciam que o uso de bases de dados e de classificadores genéricos não tem se mostrado eficaz na identificação de agentes da rede de atenção altmétrica. Esses trabalhos demonstram a necessidade de construir bases de treinamento específicas, compostas por contas efetivamente vinculadas a essa rede, a fim de garantir maior acurácia na tarefa de classificação. Tal limitação parece decorrer da complexidade inerente à caracterização de *bots* e de outros agentes que atuam na disseminação de conteúdo científico nas redes sociais.

Para construir uma base de treinamento robusta e adequada, torna-se fundamental compreender os perfis dos diversos atores envolvidos na rede de atenção altmétrica. Nesse contexto, destaca-se o papel desempenhado por pesquisadores na divulgação do conhecimento científico por meio de plataformas como o X — um tema recorrente e central nos estudos sobre métricas científicas em redes sociais (altmetria). Diversas abordagens são propostas na literatura com o objetivo de identificar, caracterizar e mapear a presença e o comportamento de acadêmicos ativos em mídias sociais, contribuindo para o aprimoramento das análises altmétricas (Mongeon; Bowman; Costas, 2023a).

Existem alguns conjuntos de dados publicados para resolver o problema de identificação de *bots* a partir de uma perspectiva de classificação binária, como alguns já citados. Além disso, alguns desses conjuntos de dados já foram usados para treinar e avaliar o chamado serviço *Botometer* (Martini *et al.*, 2021). No entanto, conforme descritos pelos autores em Balestrucci *et al.* (2019) e Minnich *et al.* (2017), a geração de contas automatizadas (*bots*) é um processo em constante evolução, o que torna obsoletos muitos dos *datasets* disponíveis, que incluem perfis já suspensos pela plataforma X. Diante disso, torna-se necessário atualizar periodicamente os detectores de *bots*, incorporando novos padrões comportamentais e características emergentes. Além disso, é imprescindível realizar um pré-processamento dos dados para remover perfis inativos ou excluídos, a fim de preservar a qualidade e a usabilidade das bases. Apenas com essas atualizações, os detectores poderão manter níveis adequados de precisão e confiabilidade.

Devido às restrições impostas pela política de uso da plataforma X, os conjuntos de dados utilizados em pesquisas anteriores costumam disponibilizar apenas os identificadores das contas, sem atributos descritivos ou comportamentais relevantes, o que dificulta sua utilização em análises mais aprofundadas. A partir de julho de 2023, o cenário tornou-se ainda mais desafiador, uma vez que o acesso à API da plataforma passou a ser ainda mais restritivo, inviabilizando a extração em larga escala de dados.

Diante desse contexto, adotou-se uma estratégia metodológica alternativa e mais direcionada, com a identificação, por meio da plataforma Altmetric, das contas que mencionaram os artigos incluídos na amostra deste estudo. Na etapa seguinte, essas contas foram analisadas manualmente por meio da interface da própria plataforma X.

Com o intuito de mitigar as limitações impostas pelo acesso restrito aos dados, também foram incorporadas informações provenientes de outras pesquisas científicas, com o objetivo de realizar uma classificação preliminar das contas identificadas via Altmetric que compõem o *dataset* desenvolvido neste estudo. A seguir, apresentam-se as principais fontes utilizadas:

An open data set of scholars on X (Mongeon; Bowman; Costas, 2023a): Este artigo desenvolve um *dataset* de contas X acadêmicas através da relação de autores do *OpenAeX* com usuários do X identificados na plataforma *Crossref Event Data*¹⁶. O procedimento de correspondência é descrito e validado com dados ORCID. A abordagem dos autores combina quase 500.000 acadêmicos, através do ORCID, com suas contas no X com um nível de alta precisão e recall moderado. Este artigo apresenta o desenvolvimento de um *dataset* de contas acadêmicas na plataforma X, construído a partir da correspondência entre autores registrados no *OpenAeX* e usuários do X identificados por meio da plataforma *Crossref Event Data*. O procedimento de vinculação é detalhado e validado com o apoio de dados provenientes do ORCID. A metodologia adotada pelos autores permitiu relacionar aproximadamente 500.000 pesquisadores a suas respectivas contas na plataforma X, alcançando alta precisão e recall moderado no processo de correspondência.

O conjunto de dados resultante é descrito e disponibilizado publicamente, com o objetivo de apoiar pesquisas mais aprofundadas sobre as interações de acadêmicos

¹⁶ Crossref Event Data registra menções a pesquisas em mídias sociais, blogs, notícias e outros meios online. Os dados são abertos e acessíveis via API: <https://www.crossref.org/event-data>

na plataforma X. O *dataset* inclui, entre outras informações, uma coluna que indica os critérios utilizados com sucesso na identificação da correspondência e outra que assinala se a vinculação foi considerada válida após inspeção manual.

The X accounts of scientific journals (Nishikawa-pacher, 2023): Este artigo apresenta um *dataset* composto de um conjunto de dados de 3.485 contas do X pertencentes a uma amostra de 13.821 periódicos listados nos três principais índices da *Web of Science*: Índice de Citações de Ciência expandido (SCIE), o Índice de Citação de Ciências Sociais (SSCI) e o Índice de Citação de Artes e Humanidades (AHC). O método consistiu na identificação e na coleta sistemática de perfis institucionais no X, com análise de métricas como frequência de postagens, interações, menções a revisão por pares e fatores de impacto. A pesquisa revelou grande variabilidade entre áreas e editoras, destacando a importância de perfis oficiais diante da existência de contas enganosas. O *dataset* está disponível para análises cientométricas futuras, apresentado em uma planilha Excel, contendo o id do autor na base de dados *Openalex* e o id no X.

Characterization of the communities of attention interacting with scientific papers on X (Maricato; Manso, 2022): Este artigo tem como objetivo analisar as comunidades de atenção à ciência na plataforma X. Para isso, foram identificadas 2.437 contas de usuários que compartilharam publicações científicas, as quais foram integralmente submetidas a um processo de categorização manual. A análise considerou múltiplas dimensões dos perfis, incluindo: tipo de perfil (pessoal, institucional ou de personagem), categoria (acadêmica, profissional ou social), ocupação, área de atuação e vínculos geográficos e institucionais. As informações utilizadas para essa classificação foram extraídas diretamente das descrições públicas dos perfis na plataforma X.

O Quadro 2 apresenta uma síntese dos diferentes *datasets* de treinamento utilizados na classificação preliminar das contas do X, aplicados no contexto da pesquisa sobre detecção de *bots* e perfis humanos envolvidos na divulgação de artigos científicos. Cada linha do quadro representa um conjunto de dados utilizados para treinamento e/ou validação de modelos de aprendizado de máquina.

Quadro 2 - Conjuntos de dados utilizados como base para treinamento e avaliação da classificação de contas no X, com foco em perfis humanos, *bots* e categorias altmétricas

Nome do Dataset	Referência	Número de Contas	Características Principais
Dataset Gutiérrez	Gutiérrez (2020)	37.437	Contas genéricas de usuários do X; classificação binária (<i>bot</i> ou humano); disponibilizado no Kaggle.
Botometer Dataset	Pontes e Maricato (2023)	Não especificado	Utilização do serviço Botometer para classificação; identificação de limitações em contas altmétricas.
Scholars on X	Mongeon, Bowman e Costas (2023)	~500.000	Acadêmicos identificados no X; correspondência validada com dados da OpenAlex e ORCID; alta precisão.
Scientific Journals on X	Nishikawa-pacher (2023)	3.485	Contas do X associadas a periódicos indexados na Web of Science; dados organizados em planilha Excel.
Comunidades de Atenção Altmétrica	Maricato e Manso (2022)	2.437	Perfis do X classificados manualmente em categorias (pessoal, institucional e de caráter); análise de vínculos institucionais e áreas de atuação.

Fonte: Dados de Pesquisa (2025)

Os usos desses conjuntos de dados possibilitaram a criação de uma base de treinamento mais alinhada às características da rede altmétrica, incluindo a diversidade de agentes relevantes para a disseminação de publicações científicas. Assim, buscou-se aprimorar o desempenho dos algoritmos de aprendizagem de máquina aplicados à classificação de contas no X, considerando as especificidades do ecossistema altmétrico e mitigando potenciais vieses decorrentes do uso de bases genéricas.

A partir do processo de coleta dos 758.288 *posts*, foram identificadas 362.395 contas. Embora fosse possível proceder diretamente a seleção aleatória das contas, optou-se, nesta etapa, pela classificação preliminar das contas coletadas do Altmetric através dos três *datasets* (Mongeon; Bowman; Costas, 2023a; Nishikawa-pacher, 2023; Maricato; Manso, 2022) previamente obtidos.

Essa classificação preliminar foi operacionalizada por meio da execução do algoritmo **ALG_04_CruzarDatasets.py**, responsável pela comparação dos identificadores das contas.

Simultaneamente, realizou-se a identificação automática de perfis — compreendendo as etapas de (i) detecção de *bots* autodenominados, (ii) identificação de revistas e (iii) identificação de professores/pesquisadores — por meio do algoritmo

ALG_05_DetectarPersonas.py, o que resultou na geração da planilha intitulada ***DAD_07_PostsIdentificacaoPersonas.xlsx***.

Os dados consolidados foram posteriormente reorganizados e estruturados na tabela ***DAD_06_ContasPorGrandeArea.xlsx***, na qual são estabelecidas as relações entre as contas analisadas e suas respectivas grandes áreas do conhecimento.

3.4.2.3 II.b.3 Cálculo amostral do número de contas e seleção randômica das contas por divisão do ForCode¹⁷

De modo análogo ao procedimento adotado para a seleção dos artigos, foi realizada uma amostragem randômica das contas responsáveis por postar os artigos científicos incluídos na amostra. O objetivo principal dessa seleção aleatória foi minimizar a ocorrência de vieses na análise da atuação de *bots*, assegurando maior rigor estatístico à investigação.

Para a definição do tamanho da amostra de contas, considerou-se o agrupamento por Divisão (área de conhecimento), conforme classificação ForCode. Inicialmente, foram identificadas e quantificadas as contas vinculadas aos artigos da amostra, com base na Divisão correspondente de cada publicação. Essa etapa foi viabilizada a partir dos dados consolidados na planilha ***DAD_08_PostsCombinados.xlsx***.

A seguir, procedeu-se o cálculo do tamanho amostral necessário para cada Divisão, considerando um nível de confiança de 99% e uma margem de erro de 5%. O cálculo foi automatizado por meio do algoritmo ***ALG_06_CalcularAmostraContas.py***, que processou os dados registrados na planilha ***DAD_09_AmostraContas_Calculo.xlsx***.

O resultado final foi consolidado na planilha ***DAD_10_AmostraContas_Resultados.xlsx***, na qual são apresentados os tamanhos amostrais definidos para cada Divisão (estrato). O tamanho total da amostra foi de 13.986 contas distintas, distribuídas proporcionalmente entre as Divisões, conforme demonstrado na Tabela 1.

¹⁷ Disponível em: <https://www.abs.gov.au/statistics/classifications/australian-and-new-zealand-standard-research-classification-anzsrc/latest-release>

A Tabela 1 apresenta o cálculo amostral das contas que postaram os artigos incluídos na amostra da pesquisa, distribuídos por divisões disciplinares conforme a codificação *Fields of Research* (ForCode). Cada linha da tabela expressa a divisão (código e descrição), o número total de contas identificadas (NUMCONTAS), os parâmetros utilizados no cálculo amostral (margem de erro de 5% e nível de confiança de 99%) e o tamanho da amostra resultante para cada área.

Tabela 1 - Cálculo amostral de contas que postaram os artigos da amostra por divisão (área de conhecimento) ForCode.

DIVISÃO	DESC	NUM. CONTAS	TAM. AMOSTRA
30	AGRICULTURAL, VETERINARY AND FOOD SCIENCES	20.698	643
31	BIOLOGICAL SCIENCES	24.391	646
32	BIOMEDICAL AND CLINICAL SCIENCES	25.509	647
33	BUILT ENVIRONMENT AND DESIGN	9.752	621
34	CHEMICAL SCIENCES	10.830	625
35	COMMERCE, MANAGEMENT, TOURISM AND SERVICES	25.422	647
36	CREATIVE ARTS AND WRITING	15.765	637
37	EARTH SCIENCES	24.514	646
38	ECONOMICS	10.911	626
39	EDUCATION	14.968	635
40	ENGINEERING	23.396	645
41	ENVIRONMENTAL SCIENCES	27.121	648
42	HEALTH SCIENCES	38.972	652
43	HISTORY, HERITAGE AND ARCHAEOLOGY	11.853	628
44	HUMAN SOCIETY	48.254	655
46	INFORMATION AND COMPUTING SCIENCES	46.284	654
47	LANGUAGE, COMMUNICATION AND CULTURE	17.517	639
48	LAW AND LEGAL STUDIES	29.499	649
49	MATHEMATICAL SCIENCES	5.032	586
50	PHILOSOPHY AND RELIGIOUS STUDIES	19.540	642
51	PHYSICAL SCIENCES	7.790	611
52	PSYCHOLOGY	6.769	604
			13.986

Fonte: Dados de pesquisa (2025).

A distribuição evidencia a diversidade de áreas científicas contempladas na base analisada, assegurando representatividade estatística para as análises subsequentes sobre comportamento e engajamento das contas.

Com base no cálculo amostral previamente realizado e nos dados obtidos na etapa de coleta dos *posts*, procedeu-se à seleção aleatória de 13.986 contas, distribuídas de acordo com as Divisões estabelecidas pelo sistema de classificação ForCode. Essa seleção foi conduzida de forma automatizada por meio do algoritmo **ALG_07_SelecionarRandomContas.py**, garantindo a representatividade proporcional de cada estrato, conforme os tamanhos amostrais definidos. O processo utilizou como referência o *dataset* proveniente da etapa de classificação preliminar, no qual algumas contas já se encontravam classificadas, assegurando um critério de seleção objetivo e estatisticamente válido.

A execução do algoritmo teve como entrada os arquivos **DAD_10_AmostraContasResult.xlsx** e **DAD_06_ContasPorGrandeArea.xlsx** resultando na planilha **DAD_11_ContasSelecionadas.xlsx**, contendo a distribuição das contas selecionadas por grande área do conhecimento.

Na sequência, o algoritmo **ALG_08_CriarDatasetFinal.py** foi responsável por integrar os dados das contas selecionadas com os respectivos *posts*, originando a planilha **DAD_12_AmostraFinal_Contas.xlsx**. Ao término deste processo, obteve-se uma amostra composta por 13.767 contas únicas, que, em conjunto, realizaram 67.283 *posts*. Este conjunto de dados está consolidado no arquivo **DAD_13_AmostraFinal_Posts.xlsx**, o qual servirá de base para as análises subsequentes.

3.4.2.4 II.b.4 Classificação manual do dataset de treinamento

O *dataset* de treinamento foi constituído a partir do total de contas selecionadas com base no cálculo amostral das contas, conforme descrito na etapa 3.4.2.3. Esse conjunto de dados, representativo da população de contas analisadas, foi utilizado no treinamento supervisionado do modelo de aprendizado de máquina, garantindo a adequação estatística e a diversidade necessária para o desenvolvimento e a validação do classificador.

A base de dados de treinamento é um elemento essencial no processo de modelagem preditiva, pois contém exemplos rotulados que orientam o algoritmo na identificação e na generalização de padrões. Cada entrada nesse conjunto está associada a uma classificação previamente definida — como conta humana ou

automatizada (*bot*) — permitindo que o modelo aprenda a distinguir entre diferentes tipos de comportamento com base nas variáveis observadas.

Ao “aprender” com os dados do *dataset* de treinamento, o modelo ajusta seus parâmetros internos, buscando maximizar a acurácia das previsões futuras. Assim, a qualidade, a diversidade e o equilíbrio das classes presentes nesse conjunto influenciam diretamente a eficácia do algoritmo na tarefa de classificação.

Após a seleção randômica das contas, foram identificadas as contas que já tinham alguma classificação realizada previamente (seção 3.4.2.2) a partir dos *datasets* de Mongeon, Bowman e Costas (2023), Nishikawa-pacher (2023) e Maricato e Manso (2022). Embora o mapeamento dessas bases não permita, por si só, afirmar com certeza que uma conta é ou não automatizada, sua integração com outras análises de comportamento, tais como número de postagens e de artigos comentados, contribui significativamente para a acurácia da classificação final.

Desta forma, dentro da seleção randômica das contas, foram identificados:

- 66 *bots* AUTODETECT que, após avaliação manual, ficaram confirmados 57;
- 205 PERSONAS AUTO DETECT;
- 122 contas do *dataset* de Pontes e Maricato (2022);
- 96 contas do *dataset* de Nishikawa-Pache (2023);
- 1.428 contas do *dataset* de Mongeon, Bowman e Costas (2023).

Na sequência, foi realizada a contabilização do número de postagens por conta, considerando todo o universo de dados coletados — e não apenas os *posts* relacionados aos artigos da amostra, garantindo uma visão mais ampla do comportamento da conta. Foram analisados o número de artigos compartilhados por cada conta, bem como as suas respectivas divisões e grupos aos quais pertencem.

Partindo do pressuposto de que *bots* são sistemas programados para executar ações repetitivas de forma contínua, considera-se como premissa que contas automatizadas tendem a apresentar padrões regulares de atividade, sendo pouco provável que se limitem a uma única postagem. Essa premissa é corroborada por Varol *et al.* (2017b), cujo estudo apresenta um framework para detecção de *bots* no X, baseado na análise de mais de mil características, incluindo padrões de rede, frequência de postagens, interações e séries temporais de atividade. Os autores estimam que entre 9% e 15% das contas ativas na plataforma são automatizadas,

responsáveis por uma parcela significativa do conteúdo publicado — o que reforça a hipótese de que *bots* mantêm um comportamento sistemático e recorrente.

A partir dos pressupostos apresentados, foram definidos os seguintes critérios para a classificação automática das contas, com base em seus padrões de atividade na plataforma:

1. Contas que realizaram apenas uma postagem e cuja publicação ocorreu há mais de 150 dias foram classificadas como humanas;
2. Contas que postaram duas vezes, com intervalo superior a 150 dias entre as publicações, também foram consideradas humanas.

As contas que não se enquadram nesses critérios foram submetidas à análise manual.

Para essa etapa, toma-se como referência Pozzana e Ferrara (2020), os quais destacam que métricas de interação social são fundamentais para compreender a dinâmica comportamental dos usuários em mídias sociais. Tais métricas, como o número de postagem por conta, também podem ser aproveitadas para aprimorar estratégias de detecção de *bots*.

A análise manual foi conduzida com base nos seguintes elementos:

1. *Descrição e imagem do perfil (bio)*: verificação da existência e do conteúdo da biografia. Muitos *bots* operam de forma anônima, a falta de informações identificáveis nas biografias pode causar confusão sobre a autenticidade das interações na plataforma (Madahali; Tian, 2023; Qi; AlKulaib; Broniatowski, 2018a). *Bots* benignos normalmente incluem “*bot*” em seus nomes de usuário ou biografias, designando claramente sua natureza automatizada. Esses *bots* podem fornecer serviços úteis, como atualizações de notícias ou entretenimento e são projetados para ajudar em vez de enganar (Madahali; Tian, 2023; Veale; Cook, 2018).

2. *Frequência de postagens*: embora seja consenso que *bots* geralmente postam com maior frequência e em intervalos regulares, enquanto humanos tendem a apresentar maior variabilidade nos horários e frequência de postagem (Gilani *et al.*, 2019), essa distinção pode não ser aplicável diretamente à divulgação científica. Estudos preliminares revelam que muitos *bots* vinculados a revistas ou editoras publicam em intervalos determinados pela data de publicação dos artigos.

3. *Padrão de postagens*: *bots* tendem a manter padrões consistentes, ao passo que contas humanas apresentam maior variabilidade em seus comportamentos (Gilani *et al.*, 2019; Gilani *et al.* 2017).

4. *Número de contas seguidas (amigos)*: observação da quantidade de perfis seguidos como possível indicador de comportamento automatizado (Varol *et al.*, 2017b).

5. *Tom pessoal das postagens*: *bots* costumam compartilhar links, realizar repostagens ou divulgar conteúdo promocional com mais frequência, enquanto usuários humanos tendem a expressar opiniões pessoais, experiências e pensamentos originais. Essa diferença ressalta o caráter transacional das interações de *bots*, em contraste com a comunicação mais pessoal e diversa dos humanos (Gilani *et al.*, 2019).

6. *Conteúdo das postagens*: *bots* podem compartilhar links, repostar ou publicar conteúdo promocional com mais frequência, enquanto os humanos tendem a compartilhar opiniões pessoais, experiências e pensamentos originais (Gilani *et al.*, 2019).

7. *Interatividade*: humanos demonstram maior propensão ao engajamento social, respondendo, comentando ou interagindo com outras contas com maior frequência (Pozzana; Ferrara, 2020).

Com base na contabilização do número de postagens por conta, considerando todo o universo de dados coletados, foram analisados o número de artigos compartilhados por cada conta, bem como as respectivas divisões e grupos aos quais pertencem. Dentro dos critérios iniciais adotados para a classificação automática contabilizou-se:

1. Um total de 6.690 contas que realizaram apenas uma postagem e cuja publicação ocorreu há mais de 150 dias foram classificadas como humanas. 120 contas com 1 *post* e menos de 150 dias de postagem foram verificadas manualmente.

2. Um total de 2.094 contas que postaram duas vezes, com intervalo superior a 150 dias entre as publicações, também foram consideradas humanas, 99 contas foram verificadas manualmente, pois só tinham duas postagens, mas a diferença entre as datas de postagem era de menos de 150 dias.

Desta forma, 8.784 contas foram pré-classificadas a partir dos procedimentos descritos e 4.983 contas, que não se enquadram nesses critérios, foram submetidas à análise manual.

Entre as contas a serem classificadas manualmente, identificaram-se 291 contas que não existiam ou foram suspensas. Esse fenômeno é comum de acontecer no X devido à prática de suspensão de contas com comportamentos suspeitos. Essas

contas foram classificadas com base na observação do conteúdo dos *posts*. Os critérios de classificação manual seguiram o que foi exposto na seção “Construção e Classificação do *Dataset* de Treinamento”.

Como resultado da classificação do *dataset* de treinamento foram identificados: 822 contas *bots* que realizaram 20.542 *posts* e 12.945 contas humanas que realizaram 46.741 *posts*.

3.4.3 II.c Processo de desenvolvimento do modelo de inteligência artificial

A estrutura da solução proposta para a classificação de contas automatizadas na plataforma X foi operacionalizada por meio de um pipeline computacional dividido em quatro passos principais, implementados no *script* **ALG_09_PIPELINE_XGBOOST.PY**. Esse pipeline tem como objetivo garantir a padronização do pré-processamento, a extração sistemática de atributos, a modelagem supervisionada, a aplicação em dados inéditos e a geração automatizada de relatórios de desempenho.

O Passo 1 consiste na geração das variáveis preditoras (*features*) especificamente concebidas para representar aspectos comportamentais relevantes — como frequência e regularidade das postagens, padrões de interação, diversidade temática e temporalidade das publicações. A extração das *features* se deu a partir de duas bases de dados: uma base referente às contas (usuários) e outra aos respectivos *posts*. Nessa fase, foram extraídas 46 *features* que abrangem aspectos quantitativos, textuais, temporais, sentimentais, estatísticos e morfológicos, tanto dos conteúdos postados quanto dos perfis dos usuários. Essas variáveis incluem métricas como legibilidade textual, frequência de *emojis*, índice de entropia dos nomes de usuário, variação temporal das postagens, diversidade vocabular, entre outras que serão apresentadas a seguir.

A inclusão dessas variáveis visou ampliar a capacidade preditiva do modelo e refinar a distinção entre perfis humanos e automatizados, contribuindo para o aprimoramento da acurácia dos resultados. As *features* foram agregadas por conta e consolidadas em uma única base de dados estruturada, que serviu como insumo para a modelagem.

O Passo 2 compreende o treinamento do modelo de classificação baseado no algoritmo XGBoost, algoritmo selecionado na avaliação dos modelos descritos nos

resultados (seção 4.3). Inicialmente, os dados foram particionados em conjuntos de treino e teste com estratificação da variável-alvo (humano vs. *Bot*). Em seguida, foram realizadas a imputação de valores ausentes e a normalização das variáveis. O modelo foi ajustado com parâmetros otimizados para lidar com desbalanceamento entre classes (por meio da métrica *scale_pos_weight*) e avaliou sua performance por meio das métricas F1-score, AUC-ROC e Kappa. Com o intuito de aprimorar a interpretabilidade do modelo e identificar as variáveis mais relevantes, aplicou-se a técnica *Shapley Additive exPlanations* (SHAP), a qual possibilitou o ranqueamento das 46 *features* de acordo com sua contribuição individual para a predição, permitindo compreender o impacto de cada variável na tomada de decisão do classificador. O modelo proposto utiliza a totalidade das 46 variáveis preditoras geradas e o limiar de decisão foi calibrado com base na maximização do F1-score para a classe *bot*, de modo a otimizar a capacidade de detecção de contas automatizadas.

O Passo 3 do *pipeline* trata da geração das 46 *features* para um novo conjunto de contas ainda não rotuladas, neste caso o conjunto de mais de 350.000 contas extraídas da coleta dos *posts*, processo descrito na seção 3.4.2.1. O processo de extração e agregação de atributos segue a mesma lógica da etapa inicial, garantindo compatibilidade com o modelo previamente treinado. Após a geração das variáveis, os dados foram salvos em formato compatível para posterior inferência.

Por fim, o Passo 4 realiza a aplicação do modelo final nos dados das novas contas. Para isso, o pipeline carrega os artefatos salvos (modelo, *scaler*, imputador, *threshold* e lista das 46 *features*), normaliza os dados inéditos e realiza a predição das probabilidades de cada conta ser um *bot*. Com base no *threshold* ajustado, cada conta é então classificada como “humana” ou “*bot*”. Além disso, o script gera visualizações interativas, como histogramas de probabilidade e gráficos de distribuição de classes, e exporta um relatório final com as classificações realizadas.

Essa estrutura modular e sequencial do pipeline permite não apenas a automação completa do processo de classificação, mas também a rastreabilidade dos resultados, a análise explicativa das decisões do modelo e a reusabilidade da arquitetura em novos contextos e domínios de aplicação.

Ressalva-se que o modelo tem como base um conjunto de dados desbalanceado, no qual o número de contas humanas era significativamente maior do que o número de *bots*, fator determinante para escolha do modelo a ser adotado.

Na sequência, cada etapa do processo será apresentada e descrita em detalhes, de modo a evidenciar os procedimentos adotados e as decisões metodológicas envolvidas na construção do modelo.

Passo 1 – Construção das variáveis (*features*) a partir de dados brutos

A primeira etapa tem como foco transformar os dados brutos em uma base consolidada de atributos (chamados *features*) que possam ser utilizados por algoritmos de aprendizado de máquina. A geração das *features* foi operacionalizada através do script em python ***ALG_09_Pipeline_XGBoost.PY***. O script se estrutura em quatro grandes etapas e a geração de *features* corresponde ao PASSO um (1) que corresponde à criação e à estruturação dos dados, à engenharia de *features* e ao pré-processamento dos dados. Os dados brutos utilizados no *script* são oriundos de dois arquivos do Excel:

- **Tabela das Contas (df_contas): *DAD_15_DatasetClassificadoManual.xlsx*:** contém informações consolidadas sobre as contas que fazem parte do *dataset*. Foram considerados os seguintes atributos:
 - **ACCOUNT:** Identificador da conta.
 - **CLASSFINAL:** Classificação final da conta (ex.: *bot*, humano).
 - **NOME:** Nome do usuário associado à conta.
 - **QTDPOSTS:** Número total de postagens feitas pela conta.
 - **QTDDOIS:** Quantidade de artigos (DOIs) diferentes mencionados pela conta.
 - **QTDSUBAREA:** Quantidade de subáreas de conhecimento abordadas nos *posts* da conta.
 - **QTDGRANDEAREA:** Quantidade de grandes áreas de conhecimento abordadas nos *posts* da conta.
- **Tabela de Interações com os dados dos posts (df_posts): *DAD_16_TodosPostsTreinamento.xlsx*:** contém informações sobre o engajamento das contas em relação a postagens sobre artigos científicos no X, ou seja, os *posts* realizados por essas contas sobre os artigos da amostra, incluindo o conteúdo textual, a data e hora da publicação, o número de seguidores no momento da postagem e a área de conhecimento associada. Foram considerados os seguintes atributos:

- **DOI**: Identificador digital do artigo científico mencionado no *post*.
- **ACCOUNT**: Nome ou identificador da conta que fez a postagem.
- **QTDSEGUIDORES**: Número de seguidores da conta que fez o *post*.
- **CONTEUDOPOST**: Texto ou conteúdo do *post*.
- **DATATEMPO**: Data e hora em que o *post* foi publicado.
- **GRANDEAREA**: Área do conhecimento à qual o artigo citado no *post* pertence (ex.: Ciências Exatas, Ciências Humanas).
- **SUBAREA**: Subárea específica do conhecimento relacionada ao artigo citado (ex.: Inteligência Artificial dentro de Ciências Exatas).

Esses atributos da tabela original *df_posts* foram utilizados para capturar padrões de engajamento típicos de *bots*, como a frequência de postagem excessiva ou estrutura repetitiva de postagens.

Esses dados são integrados para extrair um conjunto robusto de características sobre o comportamento e o perfil das contas. Isso inclui métricas quantitativas, como média de seguidores e tamanho médio dos *posts*; métricas linguísticas, como número de palavras únicas, proporção de letras maiúsculas, uso de *emojis*, *hashtags*, *links* e pontuação; além de indicadores de sentimento, repetição de postagens, presença de termos científicos e até legibilidade textual. Também são construídas variáveis temporais (como frequência de postagem por dia útil ou fim de semana) e medidas sobre o nome e o identificador da conta (por exemplo, entropia do nome, presença da palavra “*bot*”, similaridade entre nome e *username* etc.).

Após esse rico processamento, os dados são agregados por conta, formando uma única linha (registro) para cada usuário, com dezenas de variáveis representando suas características. Este arquivo foi salvo como **DAD_14_DadosFeaturesTreinamento.xlsx** e passa a ser a base principal para treinamento do modelo.

3.4.3.1 (II.c.1) Engenharia de features

A detecção de *bots* no X depende da análise de uma gama diversificada de características, ou *features*, que capturam as propriedades das contas e seus padrões de comportamento. Essas *features* são essenciais para treinar modelos de aprendizado de máquina que distinguem usuários humanos de contas automatizadas.

A engenharia de *features* consistiu na criação e na seleção das *features* utilizadas durante o treinamento do modelo para diferenciar contas humanas e *bots* em redes sociais. Na oportunidade, foram extraídas e processadas diversas *features* quantitativas, textuais e temporais. Essas variáveis foram selecionadas para capturar padrões de comportamento, estilo de escrita e interações, permitindo a construção de um modelo de aprendizado de máquina eficiente.

A literatura externa que a identificação de *bots* baseia-se em atributos como frequência de postagens, estrutura da rede de conexões e padrões linguísticos. Aljohani *et al.* (2020) ressaltam a importância da análise do conteúdo textual e das redes de relacionamento, pois *bots* tendem a apresentar centralidade excessiva e vocabulário restrito. Hayawi *et al.* (2023) reforçam que a combinação de múltiplos atributos oferece maior precisão na identificação de contas não humanas.

As *features* foram escolhidas com base nos estudos revisados sobre detecção de *bots*, visando capturar diferenças quantitativas, linguísticas e comportamentais entre contas automatizadas e humanas. Foram definidas as *features* passíveis de extração a partir das tabelas-base.

As *features* foram extraídas a partir de duas fontes principais oriundas da estruturação dos dados:

1. df_contas → Informações gerais sobre as contas (incluindo a classificação manual: humano/bot).
2. df_posts → Dados sobre as postagens feitas pelas contas, incluindo texto, horário e interações.

Essas *features* foram organizadas e agrupadas em oito categorias distintas:

a. Features gerais das contas: identificação da conta, classificação final nome de exibição.

1. **ACCOUNT** (Identificador único da conta): Cada conta foi identificada por um código único para referência.
2. **CLASSFINAL** (Classificação final da conta: humano (0) ou *bot* (1)): Essa variável indica a classificação manual atribuída a cada conta, servindo como o rótulo para o modelo de aprendizado de máquina.
3. **NOME** (Nome ou usuário da conta): Identificação textual da conta (não utilizada no modelo, apenas para referência).

b. Features de atividade da conta: quantidade total de postagens, número de DOIs distintos mencionados, diversidade de subáreas e grandes áreas temáticas e média de seguidores por postagem.

4. **QTDPOSTS** (Número total de postagens feitas pela conta): representa a quantidade de postagens publicadas pela conta. *Bots* tendem a ter altas taxas de postagem em comparação com humanos (Balestrucci *et al.*, 2019),(Arroyo-Machado; Herrera-Viedma; Torres-Salinas, 2023).

5. **QTDDOIS** (Número de publicações que uma determinada conta postou: indica quantos DOIs (*Digital Object Identifier*), o que pode ser útil para avaliar a disseminação de artigos científicos por *bots*.

6. **QTDSUBAREA** (Número de subáreas diferentes em que a conta publica): captura a diversidade de temas abordados. *bots* podem se concentrar em um único tema ou disseminar conteúdo sobre várias áreas.

7. **QTDGRANDEAREA** (Número de grandes áreas científicas cobertas pelas postagens da conta): similar à variável anterior, mas focada em categorias amplas do conhecimento.

8. **media_seguidores** (Média de seguidores por postagem): *bots* podem apresentar padrões atípicos de seguidores, seja por contagem excessivamente baixa ou alta (Aguilera *et al.*, 2023).

c. Features temporais: horário médio das postagens, intervalo médio entre postagens, volume de postagens em dias úteis e finais de semana (Ferrara *et al.*, 2016) (Balestrucci *et al.*, 2019) (Adel Alipour; Orji; Zincir-Heywood, 2022).

9. **hora_media_post** (Horário médio das postagens da conta): verifica se há um padrão fixo de horário de postagem, algo que pode indicar automação.

10. **intervalo_medio_post** (Tempo médio entre postagens da conta, em minutos): *bots* frequentemente possuem intervalos curtos e regulares entre postagens, enquanto humanos tendem a ter variação maior.

11. **posts_dia_util** (Número de postagens feitas em dias úteis (segunda a sexta)): *bots* podem ter atividade constante, independentemente do dia da semana, enquanto humanos tendem a publicar mais em determinados dias.

12. **posts_fim_de_semana** (Número de postagens feitas no sábado e domingo): verifica se a atividade da conta muda no final de semana.

A análise dessas variáveis permitiu identificar características de automação, como o uso de textos repetitivos e o compartilhamento excessivo de links; horários de atividade das contas, permitindo identificar padrões previsíveis típicos de *bots*.

d. Features Textuais: comprimento médio dos *posts*, número de palavras e caracteres, diversidade de vocabulário, uso médio de *hashtags*, links, menções, sinais de pontuação e letras maiúsculas (Aguilera *et al.*, 2023) (Ferrara *et al.*, 2016).

13. **tamanho_medio_post** (Média do comprimento das postagens em caracteres): ajuda a identificar contas que publicam mensagens curtas e repetitivas (característica comum de *bots*) ou *posts* mais longos e variados.

14. **media_caracteres** (Média do número de caracteres por postagem): semelhante a *tamanho_medio_post*, mas calculada a partir de todas as postagens da conta.

15. **media_palavras** (Média do número de palavras por postagem): *bots* podem usar frases curtas e repetitivas, enquanto humanos escrevem textos mais variados.

16. **diversidade_vocabulario** (Média da diversidade do vocabulário da conta): calculada como a relação entre palavras únicas e total de palavras usadas em postagens. *Bots* geralmente apresentam vocabulário limitado.

17. **media_hashtags** (Média de *hashtags* usadas por postagem): *bots* podem utilizar muitas hashtags para impulsionar engajamento.

18. **media_links** (Média de links por postagem): *bots* frequentemente compartilham links externos para disseminar conteúdo automaticamente.

19. **media_mentions** (Média de menções a outras contas por postagem): *bots* podem mencionar muitos usuários para gerar engajamento artificial.

20. **media_exclamacoes** (Média de pontos de exclamação usados por postagem): o uso excessivo de exclamações pode indicar padrões automatizados de escrita emocional.

21. **media_interrogacoes** (Média de pontos de interrogação usados por postagem): captura padrões linguísticos comuns em postagens automáticas.

22. **taxa_maiusculas** (Proporção média de letras maiúsculas por postagem): *bots* podem utilizar muitas letras maiúsculas para chamar atenção.

23. **media_pontuacao_excessiva** (Média de uso de pontuação repetitiva como "!!!" ou "..."): a presença frequente de pontuação exagerada pode ser um indicativo de automação.

e. **Features Relacionadas a Emojis e Sentimento:** média de emojis por *post* e polaridade emocional média das postagens, a ser calculada por meio da biblioteca TextBlob. (Ferrara *et al.*, 2016) (Balestrucci *et al.*, 2019).

24. **media_emojis** (Média de emojis usados por postagem): *bots* podem ter padrões específicos de uso de emojis, como emoji único repetido ou ausência total de emojis.

25. **media_sentimento** (Média da polaridade do sentimento das postagens): analisada com o TextBlob, essa variável captura a tendência emocional das postagens. *Bots* podem ter polaridade extrema (muito positiva ou negativa) ou neutra (Ferrara *et al.*, 2016).

f. **Features de repetição e legibilidade:** proporção de postagens repetidas, frequência de vocábulos científicos e índice médio de legibilidade (Flesch), calculado com a biblioteca TextStat (Balestrucci *et al.*, 2019) (Aguilera *et al.*, 2023).

26. **taxa_posts_repetidos** (Proporção de postagens repetidas dentro da conta): *bots* frequentemente publicam o mesmo conteúdo várias vezes, enquanto humanos tendem a criar textos mais variados.

27. **media_palavras_cientificas** (Média de palavras associadas a temas científicos): calculada a partir da presença de palavras-chave como "pesquisa", "artigo" e "científico". Ajuda a identificar *bots* focados em disseminação científica.

28. **media_legibilidade** (Média do índice de legibilidade de Flesch das postagens): calculada com o TextStat, esse índice mede o quanto fácil um texto é de ser lido. *Bots* podem gerar textos muito simples ou extremamente complexos.

g. **Features baseadas na coluna NOME (nome do usuário exibido):** (nome do usuário exibido): características morfológicas e visuais do nome, como tamanho, número de palavras, proporção de letras maiúsculas, presença de números, emojis e nível de entropia (aleatoriedade) (Ferrara *et al.*, 2016; Balestrucci *et al.*, 2019). Para essa identificação foram utilizadas as funções nativas do Python emoji e math.

29. **tamanho_name** (Número total de caracteres no nome): pode indicar nomes curtos ou longos, comuns em usuários automatizados.

30. **quantidade_palavras_name** (Número de palavras no nome): *bots* tendem a ter menos palavras ou nomes sem estrutura comum.

31. **maioria_maiuscula_name** (Proporção de letras maiúsculas no nome): um valor alto pode indicar formatações não naturais.

32 **contém_números_name** (Indica (1/0) se o nome contém ao menos um número): números são menos comuns em nomes reais.

33. **entropia_name** (Medida da aleatoriedade dos caracteres do nome): alta entropia pode indicar strings geradas automaticamente.

34. **contém_emoji_name** (Indica (1/0) se o nome contém emoji): emojis são mais comuns em perfis informais ou promocionais.

h. Features baseadas na coluna ACCOUNT (nome de usuário / @handle): medidas estruturais e padrões de formatação do nome de usuário, como número de dígitos, proporção de caracteres especiais, entropia, presença da palavra "bot" e repetição de caracteres. Usa a biblioteca *re* para expressões regulares e bibliotecas padrão como *math* (Balestrucci *et al.*, 2019).

35. **tamanho_username** (Número total de caracteres no username): *bots* podem usar strings curtas ou muito longas.

36. **digitos_username** (Contagem de dígitos (0–9) no username): muitos dígitos podem indicar geração automatizada.

37. **letras_username** (Número de letras (a–z, A–Z) no username): pode ser usada para calcular proporções de letras.

38. **especial_username**: quantidade de caracteres especiais (que não são letras nem números).

39. **proporcao_digitos_username** (Proporção de dígitos em relação ao total de caracteres do username): Útil para medir “robotização”.

40. **tem_numeros_fim_username**: indica (1/0) se o username termina com números, o que é comum em nomes gerados automaticamente.

41. **contém_bot_username**: indica (1/0) se a palavra "bot" aparece no username. Indicador direto de automação.

42. **entropia_username** (Medida de entropia do username): nomes muito aleatórios tendem a ter alta entropia.

43. **tem_repeticoes_username** (Indica (1/0) se há caracteres repetidos consecutivos no username): padrões como "aa", "zz", etc., podem sugerir geração automática.

i. **Features de similaridade entre nome (NOME) e username (ACCOUNT):** avaliam a coerência entre o nome completo e o username por meio de métricas de similaridade de caracteres (Jaccard e Levenshtein) e verificação de abreviação. Utiliza as bibliotecas **difflib, collections**, e funções personalizadas de comparação (Ferrara 2017; Arroyo-Machado; Herrera-Viedma; Torres-Salinas, 2023).

44. **similaridade_jaccard**: similaridade entre os conjuntos de caracteres de nome e username. Mede sobreposição dos caracteres.

45. **similaridade_levenshtein** (Similaridade com base em edição (Levenshtein): avalia o quanto semelhante são nome e username em termos de inserções, deleções e substituições.

46. **username_é_abreviação** (Indica (1/0) se o username é uma abreviação plausível do nome). Reflete coerência entre os dois campos.

Após a etapa de processamento, os dados foram agregados por conta, resultando em um único registro por usuário, contendo 46 de *features* representativas de suas características comportamentais, textuais, temporais e de perfil. O resultado dessa agregação foi salvo no arquivo **DAD_14_DadosFeaturesTreinamento.xlsx**, que passou a constituir a base principal para o treinamento dos modelos de aprendizagem de máquina.

As *features* extraídas foram consolidadas em uma matriz final (df_final), estruturada para representar cada conta como uma instância única no conjunto de dados. Essa matriz serviu como entrada para os algoritmos de detecção de *bots*, possibilitando a identificação de padrões preditivos capazes de distinguir, de forma automatizada, perfis humanos de contas automatizadas.

3.4.3.2 (II.c.2) Modelagem do Modelo de Aprendizagem de Máquina para classificação de contas do X baseado no XGBOOST

Com a base de dados processada e o modelo definido, teve início a segunda etapa da pesquisa: o treinamento do classificador. Essa fase corresponde ao PASSO

dois (2), que integra o script em Python ***ALG_09_Pipeline_XGBoost.PY***, responsável por operacionalizar toda a pipeline de modelagem.

O processo de construção do classificador envolveu diversas etapas, incluindo o pré-processamento realizado previamente, a divisão dos dados em conjuntos de treino e teste (executada no ***ALG_09_Pipeline_XGBoost - PASSO 1***), o treinamento inicial do modelo, a aplicação de técnicas para lidar com o desbalanceamento entre classes — com ênfase na correta identificação da classe minoritária (*bots*) —, além da posterior otimização de hiperparâmetros.

Ao final do processo, o modelo foi utilizado para classificar as 361.000 contas coletadas, buscando maximizar sua capacidade preditiva e a distinção precisa entre contas humanas e automatizadas. A seguir, são detalhados os procedimentos executados, os ajustes aplicados e os resultados alcançados nessa etapa da pesquisa.

a. Pré-processamento e preparação dos dados

Inicialmente, foram carregados os dados processados (***ALG_09_Pipeline_XGBoost - PASSO2***), que continham diversas *features* extraídas das contas de redes sociais. Durante o pré-processamento, foram realizados os seguintes ajustes essenciais:

- Tratamento de valores ausentes:
 - Para colunas numéricas, os valores ausentes foram preenchidos com a média da coluna correspondente.
 - Para colunas categóricas, utilizamos a moda (valor mais frequente) ou uma *string* vazia.
- Conversão da variável-alvo (CLASSFINAL):
 - As classes foram convertidas para valores numéricos, em que "Humano" foi representado como 0 e "Bot" como 1.
- Separação entre *features* (X) e variável-alvo ¹⁸(y):

¹⁸ Atributo que representa o resultado a ser previsto pelo modelo; neste estudo, corresponde à indicação de uma conta como *bot* (classe positiva) ou *não bot* (classe negativa).

- Remoção das colunas irrelevantes como identificadores (ACCOUNT, NOME) para garantir que apenas as características relevantes fossem utilizadas no treinamento do modelo.

Após essa etapa, a distribuição das classes foi então analisada, revelando um desequilíbrio significativo, com uma predominância de contas humanas em relação às contas de *bots*. Esse fator foi determinante para as escolhas metodológicas subsequentes, uma vez que o desbalanceamento impacta diretamente o desempenho do modelo, podendo levá-lo a favorecer a classe majoritária. Com o objetivo de mitigar esse efeito, foram adotadas estratégias específicas de balanceamento, conforme detalhado ao longo deste trabalho.

Antes do treinamento, os dados passam por etapas de pré-processamento: valores ausentes são imputados (preenchidos com a média) e todas as variáveis são padronizadas (normalizadas) para evitar que discrepâncias de escala impactem o modelo.

b. Definição dos hiperparâmetros

Para que um modelo de inteligência artificial apresente bons resultados, é fundamental ajustar corretamente seus hiperparâmetros, que são configurações internas definidas antes do início do treinamento e que controlam como o algoritmo aprende. Diferentemente dos parâmetros comuns — que o próprio modelo ajusta automaticamente durante o processo de aprendizagem — os hiperparâmetros precisam ser definidos pelo pesquisador, podendo influenciar diretamente a precisão, a velocidade e a capacidade de generalização do modelo. No caso desta pesquisa, foi utilizado o algoritmo XGBoost, amplamente reconhecido por seu desempenho em tarefas de classificação. Ajustes criteriosos foram realizados para reduzir o sobreajuste (quando o modelo “decora” os dados de treino e perde a capacidade de prever novos casos) e melhorar a capacidade preditiva (habilidade de fazer previsões corretas em dados inéditos). O Quadro 3 apresenta, de forma organizada, os principais hiperparâmetros adotados, seus valores e uma breve descrição funcional, permitindo compreender como cada configuração contribuiu para o desenvolvimento do modelo e para lidar com o desafio do desbalanceamento entre classes presentes na base de dados. Assim, os hiperparâmetros foram ajustados para reduzir o sobreajuste e melhorar a capacidade preditiva do modelo.

O Quadro 3 apresenta os principais hiperparâmetros do algoritmo XGBoost utilizados no desenvolvimento do modelo de classificação proposto nesta pesquisa. Para melhor visualização, estão organizados em três colunas: o nome do parâmetro, o valor utilizado durante o treinamento do modelo e uma descrição funcional de cada parâmetro.

Quadro 3 - Quadro de síntese dos hiperparâmetros configurados no algoritmo XGBoost, apresentando os valores adotados e a respectiva descrição de sua função, no contexto da solução desenvolvida para a classificação de contas automatizadas no X

Parâmetro	Valor Utilizado	Descrição
Objective	binary:logistic	Define que o problema é de classificação binária com saída em probabilidade.
base_score	None	Pontuação inicial para todos os exemplos. Útil para bases desbalanceadas extremas.
Booster	None	Tipo de modelo base ('gbtree', 'gblinear', 'dart'). Padrão é 'gbtree'.
Call-backs	None	Lista de funções de callback durante o treino, como salvar checkpoints.
colsample_bylevel	None	Proporção de <i>features</i> por nível em cada árvore.
colsample_bynode	None	Proporção de <i>features</i> por nó em cada split.
colsample_bytree	0.7	Proporção de <i>features</i> amostradas para cada árvore.
Device	None	Define o uso de CPU ou GPU.
early_stopping_rounds	None	Número de rounds sem melhora antes de parar o treinamento (requer eval_set).
enable_categorical	False	Habilita o uso nativo de variáveis categóricas.
eval_metric	aucpr	Métrica de avaliação: AUC da curva de Precisão-Revocação.
feature_types	None	Tipos de <i>features</i> como 'float', 'categorical'.
Gamma	2	Ganho mínimo necessário para fazer um split. Ajuda a regularizar.
grow_policy	None	Política de crescimento da árvore ('depthwise', 'lossguide').
importance_type	None	Critério para cálculo da importância das <i>features</i> .
interaction_constraints	None	Restrições nas interações entre variáveis.
learning_rate	0.03	Taxa de aprendizado. Controla o impacto de cada árvore.
max_bin	None	Número máximo de bins usados no histograma para splits.
max_cat_threshold	None	Limiar máximo para categorias ao usar uma-hot.
max_cat_to_onehot	None	Número máximo de categorias convertidas para one-hot.
max_delta_step	None	Passo máximo de ajuste de peso por árvore.
max_depth	3	Profundidade máxima das árvores. Controla a complexidade.
max_leaves	None	Número máximo de folhas (usado com grow_policy='lossguide').
min_child_weight	5	Peso mínimo de amostras para uma folha ser criada.
Missing	nan	Valor considerado como ausente (NaN).
monotone_constraints	None	Impõe restrições de monotonicidade.

multi_strategy	None	Estratégia para classificação multiclasse.
n_estimators	200	Número de árvores (iterações).
n_jobs	None	Número de threads para paralelização.
num_parallel_tree	None	Número de árvores paralelas por rodada.
random_state	42	Semente de aleatoriedade para reproduzibilidade.
reg_alpha	None	Regularização L1 (lasso).
reg_lambda	None	Regularização L2 (ridge).
sampling_method	None	Método de amostragem ('uniform', 'gradient_based').
scale_pos_weight	15.75826086956 5218	Peso da classe positiva para lidar com desbalanceamento.
subsample	0.8	Proporção de amostras usadas por árvore.
tree_method	None	Algoritmo de construção de árvore ('auto', 'hist', 'gpu_hist').
validate_parameters	None	Valida os parâmetros ao iniciar o modelo.
Verbosity	None	Nível de logging (0=silencioso, 3=detalhado).
use_label_encoder	False	Evita uso do codificador interno de rótulo (obsoleto).

Fonte: Dados de Pesquisa

Para mitigar os efeitos do desbalanceamento entre as classes, foi utilizado o parâmetro *scale_pos_weight* do XGBoost, que ajusta a penalização atribuída à classe minoritária (*bots*) durante o treinamento. Conforme recomendado na documentação oficial do XGBoost (XGboost Developers, 2022), o valor sugerido para esse parâmetro é dado pela razão entre o número de instâncias da classe majoritária (negativa) e o número de instâncias da classe minoritária (positiva), conforme expressa na fórmula:

$$\text{Scale_pos_weight} = \text{sum}(n_negativas) / \text{sum}(n_positivas).^{19}$$

Esse ajuste tem como objetivo compensar a desproporção entre as classes, atribuindo maior peso às amostras da classe menos representada, o que reduz o viés do modelo em favor da classe dominante. A divisão dos dados foi realizada em conjuntos de treino (70%) e teste (30%), assegurando a preservação da proporção original das classes por meio da estratégia de estratificação (stratify=y).

3.4.3.3 (II.c.3) Processo de treinamento e otimização do modelo XGBOOSTS

O processo de treinamento e otimização do modelo XGBoost constitui uma etapa fundamental no desenvolvimento da solução proposta, uma vez que define os parâmetros e ajustes necessários para alcançar o melhor desempenho preditivo. Nessa fase, busca-se equilibrar a capacidade de generalização do algoritmo com a

¹⁹ Onde *n_negativos* é o número de amostras da classe majoritária e *n_positivos* da classe minoritária. Esse ajuste é especialmente útil em contextos de classificação binária com forte desbalanceamento.

minimização de erros de classificação, por meio da calibração de hiperparâmetros e da avaliação sistemática de métricas de desempenho. Assim, o procedimento não apenas garante maior robustez ao modelo, mas também contribui para a adequação da técnica ao contexto específico da detecção de contas automatizadas no X.

a. Divisão dos dados e avaliação inicial

Para a construção do modelo de classificação, foi realizada a divisão dos dados em dois subconjuntos: um destinado ao treinamento (70%) e outro à validação (30%). A divisão preservou a proporção original entre as classes da variável-alvo, por meio da técnica de estratificação (*stratify=y*), assegurando a representatividade da classe minoritária.

A avaliação inicial do modelo foi realizada utilizando os hiperparâmetros padrão do algoritmo XGBoost, com o objetivo de estabelecer uma linha de base para comparação com modelos otimizados. Os principais indicadores de desempenho a serem considerados incluem acurácia, recall e precisão da classe minoritária, permitindo avaliar se o modelo apresenta viés para a classe majoritária.

b. Estratégias para mitigação do desequilíbrio de classes

Dado o desequilíbrio entre as classes da variável-alvo, foi realizada a aplicação de estratégias específicas para aumentar a sensibilidade do modelo à classe minoritária, mantendo níveis aceitáveis de desempenho global. As técnicas a seguir foram consideradas:

- Ajuste do parâmetro *scale_pos_weight*: Calculou-se a razão entre o número de instâncias da classe majoritária e da classe minoritária, configurando o parâmetro *scale_pos_weight* com esse valor. Essa abordagem visa reduzir o viés do modelo, atribuindo maior peso à classe com menor representatividade, incentivando o algoritmo a prestar maior atenção a ela durante o processo de treinamento.
- Ajuste de hiperparâmetros do XGBoost: Modificaram-se os seguintes hiperparâmetros do modelo, com o objetivo de melhorar sua capacidade de generalização e evitar sobreajuste²⁰:

²⁰ *Sobreajuste (overfitting)* ocorre quando o modelo aprende padrões específicos demais do conjunto de treino, incluindo ruídos ou flutuações aleatórias, perdendo capacidade de generalização para novos dados. O XGBoost oferece diversos hiperparâmetros para mitigar esse problema, como *max_depth*,

- **learning_rate**: Redução para valores menores (ex.: 0.05), a fim de permitir um aprendizado mais progressivo e refinado;
- **max_depth**: Limitação da profundidade das árvores (ex.: 4), para controle da complexidade do modelo;
- **subsample** e **colsample_bytree**: Definição de valores intermediários (ex.: 0.8), promovendo regularização adicional por meio da amostragem parcial de observações e variáveis.
- Implementação de *Early Stopping*: Durante o treinamento, utilizou-se o recurso de parada antecipada (*early stopping*), com avaliação periódica do desempenho do modelo no conjunto de validação. Caso não sejam observadas melhorias após um número predefinido de iterações consecutivas (ex.: 10), o processo de treinamento deve ser interrompido automaticamente. Essa técnica utilizada contribui para a prevenção do sobreajuste, garantindo maior robustez do modelo.

c. Ajuste fino para equilíbrio entre *recall* e precisão

Após a aplicação das estratégias iniciais para mitigação do desequilíbrio entre classes, tornou-se necessário realizar um ajuste fino do modelo, de modo a otimizar simultaneamente sua capacidade de identificar corretamente instâncias da classe minoritária e minimizar o número de classificações incorretas. Essa etapa buscou encontrar um ponto de equilíbrio entre a sensibilidade do modelo à classe minoritária (*recall*) e a redução de falsos positivos (precisão), por meio de ajustes adicionais no processo de modelagem. Desta forma, foram adotadas as seguintes ações:

- Refinamento de hiperparâmetros: realização de novos testes com variações mais sutis nos parâmetros anteriormente definidos;
- Ajuste do limiar de decisão (*threshold*): avaliação do desempenho do modelo com diferentes valores de corte para a probabilidade predita (ex.: entre 0.4 e 0.5), a fim de identificar o ponto ótimo de balanceamento entre *recall* e precisão.

Durante o processo experimental, foram conduzidos diferentes testes com o objetivo de otimizar o desempenho do modelo de inteligência artificial baseado no algoritmo XGBoost. A principal meta foi maximizar a capacidade de detecção da

subsample, colsample_bytree e learning_rate, que atuam diretamente no controle da complexidade do modelo (XGBoost Developers, 2022).

classe minoritária (*bots*), enfrentando os desafios típicos de desbalanceamento de classes presentes em cenários altmétricos.

Inicialmente, o modelo foi avaliado utilizando o *threshold* padrão de 0,5. Nessa configuração, obteve-se um F1-score de 0,5359 e recall de 0,4534 para a classe *bot*, com acurácia global de 95,30%. Apesar da elevada acurácia, os resultados evidenciaram um viés na classificação, com desempenho significativamente superior na detecção de contas humanas.

Buscando aumentar a sensibilidade do modelo para contas automatizadas, foi realizado um segundo teste com *threshold* ajustado manualmente para 0,75. No entanto, os resultados permaneceram praticamente inalterados em relação ao limiar padrão, sugerindo que ajustes manuais não seriam suficientes para romper o viés de classificação.

A fim de superar essa limitação, foi aplicada a técnica de otimização do *threshold* com base na curva de precisão e revocação (*precision-recall curve*). Essa abordagem identificou como ponto ótimo o limiar de 0,82, valor que maximiza o F1-score da classe minoritária. Com essa calibragem, o modelo passou a apresentar um F1-score de 0,5479 e *recall* de 0,5789 para *bots* — representando um ganho real na sensibilidade da detecção. Embora tenha ocorrido uma ligeira queda na acurácia global (para 94,29%), o valor da área sob a curva ROC (ROC AUC) foi de 0,9392, o que demonstra excelente capacidade discriminativa. O índice Kappa de 0,5175 reforça a robustez do modelo, sugerindo concordância moderada entre predição e rótulo real.

Visando ainda aprimorar a performance do modelo, foram conduzidos testes com diferentes conjuntos de variáveis preditoras: 05, 10, 15, 20, 25, 30, 35, 40, 45 e 46 *features*. A estratégia adotada baseou-se na seleção de atributos com maior importância explicativa segundo os valores SHAP (*SHapley Additive exPlanations*), mantendo a configuração dos hiperparâmetros constante, bem como a divisão dos dados de treino e teste. A Tabela 2 apresenta o comparativo dos resultados de desempenho do modelo XGBoost treinado com diferentes conjuntos de variáveis preditoras: 05, 10, 15, 20, 25, 30, 35, 40, 45 e 46 *features*. São exibidas métricas fundamentais para a avaliação de modelos de classificação binária.

Tabela 2 – Comparativo dos resultados com 05,10, 15, 20, 25, 30, 35, 40, 45 e 46 variáveis preditoras (*features*)

Conjunto de <i>Features</i>	Threshold Ótimo	F1-score (<i>Bot</i>)	Recall (<i>Bot</i>)	Precision (<i>Bot</i>)	ROC AUC	Kappa	Accuracy
05 variáveis	0,82	0,4831	0,5789	0,4245	0,9269	0,4440	0,9259
10 variáveis	0,87	0,5249	0,4899	0,5654	0,9355	0,4970	0,9470
15 variáveis	0,84	0,5296	0,5425	0,5174	0,9382	0,4990	0,9424
20 variáveis	0,84	0,5267	0,5182	0,5356	0,9372	0,4972	0,9443
25 variáveis	0,87	0,5309	0,4696	0,6105	0,9382	0,5052	0,9504
30 variáveis	0,82	0,5399	0,5749	0,5090	0,9387	0,5088	0,9414
35 variáveis	0,79	0,5371	0,6154	0,4765	0,9390	0,5036	0,9366
40 variáveis	0,87	0,5397	0,4818	0,6134	0,9386	0,5141	0,9509
45 variáveis	0,82	0,5426	0,5668	0,5204	0,9396	0,5122	0,9429
46 variáveis	0,82	0,5479	0,5789	0,5200	0,9392	0,5175	0,9429

Fonte: Dados de Pesquisa (2025).

Nos experimentos com número reduzido de variáveis (5 a 15), verifica-se uma oscilação entre recall e precisão. O modelo com cinco variáveis alcançou recall elevado (0,5789), mas acompanhado de baixa precisão (0,4245), o que indica uma tendência a identificar a maior parte das contas automatizadas ao custo de aumentar a incidência de falsos positivos. À medida que o número de variáveis aumenta para 10 e 15, observa-se um equilíbrio mais favorável entre precisão e recall, aproximando-se de um ponto de maior estabilidade. Essa transição sugere que o acréscimo de variáveis relevantes amplia a capacidade do modelo em distinguir com mais segurança entre classes, reduzindo a sobreposição de padrões.

Nos cenários intermediários, especialmente entre **20 e 30 variáveis**, os resultados tornam-se mais consistentes. O conjunto de 25 variáveis destacou-se pela obtenção da maior precisão (0,6105), ainda que à custa de uma redução no recall (0,4696). Essa configuração mostra-se particularmente útil em situações em que a prioridade é minimizar a ocorrência de falsos positivos, privilegiando classificações mais conservadoras. Já o modelo com 30 variáveis, em contrapartida, apresentou recall elevado (0,5749) e precisão estável (0,5090), o que favorece aplicações voltadas à detecção ampliada de *bots*, mesmo que com algum aumento de falsos positivos.

Com a ampliação do número de variáveis para **35 a 45**, o desempenho do modelo atinge maior maturidade. O cenário de 35 variáveis, por exemplo, resultou no recall mais elevado de todos (0,6154), revelando-se particularmente eficaz em contextos em que é crítico não deixar *bots* passarem despercebidos. Entretanto, tal

desempenho foi acompanhado por uma redução na precisão (0,4765), o que implica em maior risco de sobreclassificação. Em contraste, o cenário de 40 variáveis apresentou o efeito inverso, com alta precisão (0,6134) e recall inferior (0,4818), reforçando a ideia de que a escolha do modelo deve considerar o *trade-off* entre essas métricas, em função dos objetivos analíticos.

A avaliação comparativa dos modelos com diferentes quantidades de variáveis preditoras evidencia a consistência do desempenho do XGBoost, independentemente da dimensionalidade do conjunto de *features*. Observa-se que o valor da métrica **ROC AUC** manteve-se elevado em todos os cenários ($\approx 0,93\text{--}0,94$), indicando robustez da capacidade discriminativa do classificador. De forma semelhante, a **acurácia global** apresentou variação pouco expressiva, situando-se entre 0,9259 e 0,9509, o que sugere estabilidade na classificação geral entre contas humanas e automatizadas. Contudo, ao aprofundar a análise nas métricas direcionadas à classe minoritária (*bots*), emergem diferenças relevantes que orientam a escolha do modelo mais adequado em função da estratégia de aplicação.

Por fim, o modelo com **46 variáveis** consolidou-se como a configuração mais equilibrada e robusta, apresentando os melhores valores combinados de F1-score (0,5479), recall (0,5789), precisão (0,5200) e coeficiente Kappa (0,5175). Esses resultados indicam não apenas a capacidade do modelo em capturar de maneira consistente as nuances da classe *bot*, mas também maior concordância além do acaso, o que é fundamental para análises em larga escala. Assim, a comparação demonstra que, embora cenários intermediários (25 ou 35 variáveis) possam ser úteis em aplicações específicas — respectivamente, quando se busca maior precisão ou maior sensibilidade —, a configuração com 46 variáveis oferece o desempenho mais robusto em termos gerais, sendo a mais indicada para aplicações em contextos de monitoramento altmétrico.

Esses resultados confirmam que, para o objetivo específico de identificar *bots* — uma tarefa de alta complexidade em contextos altmétricos —, a utilização de um conjunto mais amplo e cuidadosamente selecionado de variáveis tende a favorecer o equilíbrio entre sensibilidade e precisão.

A diferença observada entre os testes das variáveis reforça a importância da riqueza de atributos no mapeamento de padrões comportamentais sutis, característicos de *bots*, conforme discutido por Varol *et al.* (2017) e Ferrara *et al.* (2016). Além disso, a estabilidade dos resultados com os *thresholds* otimizados (0,82

e 0,84) demonstra que a calibragem do limiar de decisão é uma etapa necessária para maximizar o desempenho na classe minoritária, sem comprometer significativamente a acurácia geral.

Esses achados comprovam a relevância de estratégias de otimização que combinem ajuste do *threshold*, balanceamento das classes por meio do parâmetro *scale_pos_weight*, e seleção criteriosa de variáveis preditoras. Tais estratégias são determinantes para o sucesso da classificação automatizada em ambientes altmétricos, especialmente diante da crescente sofisticação dos comportamentos automatizados nas redes sociais (Varol *et al.*, 2017; Pozzana; Ferrara, 2020); Madahali; Tian, 2023).

Por fim, o modelo, o escalador, o imputador, o *threshold* ideal e a lista das 46 melhores *features* são salvos como arquivos.pkl, que permitem a reutilização futura. Um relatório em PDF é gerado automaticamente, contendo gráficos explicativos (matriz de confusão, importância das variáveis, desempenho do modelo) e um sumário textual das métricas, servindo como documento técnico e suporte à tese.

3.4.3.4 (II.c.4) *Classificação das contas*

Com o modelo de aprendizado de máquina devidamente treinado e validado, e com o *dataset* de treinamento consolidado, procedeu-se a aplicação do modelo sobre o conjunto de contas coletadas. Nessa etapa, as contas foram automaticamente classificadas de acordo com o tipo de usuário, distinguindo-se entre perfis humanos e automatizados (*bots*). Essa etapa é constituída pelos passos 3 e 4.

Passo 3 – Extração das 46 *features* para novas contas

Com o modelo treinado, a terceira etapa consiste em aplicar exatamente os mesmos processos de extração de variáveis a um novo conjunto de dados — agora composto por contas que não pertenciam ao *dataset* de teste/treinamento.

Os arquivos utilizados como entrada foram:

- **DAD_17_ContasIntegralizado.xlsx:** base de dados das novas contas.

- **DAD_18_PostsIntegralizado.xlsx**: base contendo os *posts* feitos por essas novas contas.

O mesmo conjunto de transformações e cálculos descrito no Passo 1 foi executado. Foram novamente extraídas características textuais, métricas temporais, estruturais e sintáticas, além de medidas linguísticas, de legibilidade e de similaridade entre nome e identificador da conta.

Essas variáveis são reunidas em uma base de dados chamada **DAD_19_DadosFeaturesNovasContas.xlsx**, pronta para ser submetida ao modelo.

Passo 4 – Classificação das novas contas

Na última etapa, o modelo treinado é utilizado para predizer a probabilidade de que cada uma das novas contas seja um *bot*. Para isso, o *script* carrega os arquivos salvos anteriormente:

- **MDL_ModeloXgboost.pkl**: o modelo de *machine learning* treinado.
- **MDL_Scaler.pkl** e **MDL_Imputer.pkl**: os objetos responsáveis por normalizar e preencher dados ausentes, da mesma forma que no treinamento.
- **MDL_ThresholdOtimo.pkl**: o melhor limiar de decisão já calculado.
- **MDL_Top46Features.pkl**: a lista das variáveis que devem ser utilizadas.

A base de novas contas é então carregada, e apenas as 46 variáveis esperadas são selecionadas. Os dados são padronizados e enviados ao modelo, que retorna uma probabilidade para cada conta. Se essa probabilidade ultrapassar o valor de *threshold* salvo, a conta é classificada como *bot*; caso contrário, como humana.

Os resultados finais — com probabilidade prevista, classe binária e rótulo textual (“Humano” ou “Bot”) — foram exportados para um arquivo Excel denominado **DAD_20_ClassificacaoNovasContas.xlsx**. As análises resultantes dessa classificação estão descritas na seção de Resultados e Discussões.

Este modelo, composto por quatro grandes etapas, possibilita a construção de uma solução completa e reproduzível para a identificação de *bots* em postagens acadêmicas no X, com base em atributos linguísticos, comportamentais e estruturais.

Todo o processo foi documentado, os parâmetros otimizados foram preservados e os resultados podem ser interpretados e auditados por meio dos relatórios e gráficos gerados. A estruturação cuidadosa dos dados e a explicação do funcionamento do modelo tornam essa abordagem cientificamente rigorosa e acessível.

4 RESULTADOS E DISCUSSÕES

O presente capítulo apresenta os resultados obtidos a partir da aplicação dos procedimentos metodológicos descritos anteriormente, estando norteado pelos objetivos traçados para esta investigação. Inicialmente, são expostos os resultados decorrentes da construção de um *dataset* rotulado e representativo de contas da plataforma X. Na sequência, discutem-se os desempenhos comparativos dos algoritmos de aprendizado de máquina aplicados à detecção de *bots*. Em seguida, detalham-se as observações oriundas da proposição de um modelo otimizado com base no algoritmo XGBoost, sua avaliação e a elaboração de um protótipo computacional. Por fim, são discutidos os padrões de comportamento e engajamento identificados entre contas humanas e automatizadas, além da análise das hipóteses de pesquisa, contemplando evidências empíricas e tipos de verificação aplicados.

4.1 CONSTRUÇÃO DE UM *DATASET* ROTULADO E REPRESENTATIVO DE CONTAS DA PLATAFORMA X

A adoção de critérios de classificação consistentes, associados a uma curadoria manual rigorosa, permitiu o desenvolvimento de uma base de dados que representa o comportamento dos atores envolvidos na disseminação científica em ambientes digitais. A aplicação da metodologia desenvolvida, resultou na construção de uma base de dados robusta e confiável para o treinamento supervisionado de modelos de aprendizado de máquina, contemplando a classificação de contas do X em perfis humanos e automatizados (*bots*). A base consolidada resultou na identificação de 822 contas de *bots*, responsáveis por 20.542 postagens, e 12.945 contas humanas, responsáveis por 46.741 postagens.

Para alcançar esse resultado, inicialmente foram integrados diferentes conjuntos de dados de referência: Mongeon, Bowman e Costas (2023), Nishikawa-pacher (2023) e Maricato e Manso (2022). Embora essas bases não permitam, por si sós, a identificação conclusiva de *bots*, sua combinação com análises de comportamento fortaleceu significativamente a acurácia da classificação.

A etapa preliminar de classificação das contas foi realizada por meio de critérios automatizados, baseando-se em padrões objetivos de atividade de postagem. Contas

com apenas uma publicação, realizada há mais de 150 dias, foram classificadas automaticamente como humanas, totalizando 6.690 perfis. De modo semelhante, contas com exatamente duas postagens e intervalo superior a 150 dias entre elas também foram consideradas como operadas por usuários humanos, abrangendo outras 2.094 contas.

As contas que não atenderam a esses critérios foram submetidas a uma análise manual, conduzida com base em parâmetros reconhecidos na literatura especializada em detecção de *bots*. Para essa etapa, foram adotados critérios qualitativos e quantitativos extraídos de estudos como os de Varol *et al.* (2017), Qi *et al.* (2018), Pozzana; Ferrara (2020) e Madahali; Tian (2023). Entre os aspectos avaliados destacam-se: a descrição do perfil, o padrão e a frequência de postagens, o número de amigos e seguidores, o conteúdo textual das publicações, o tom das mensagens e o grau de interatividade com outros usuários.

Durante o processo de revisão manual, foram analisadas individualmente 4.983 contas. Desse total, 291 perfis se encontravam inexistentes ou suspensos pela plataforma X. Nestes casos, a classificação foi realizada com base nos dados anteriores à suspensão, levando em consideração o histórico de publicações ainda disponível.

Além das contas classificadas manualmente ou via regras automáticas, a base de dados contou com integrações complementares, provenientes de diferentes fontes de referência. Foram incluídas 66 contas previamente identificadas como *bots* por sistemas automatizados (AUTODETECT), das quais 57 foram confirmadas como *bots* após verificação manual. Também foram integradas 205 contas associadas a pessoas humanas reais, bem como 122 perfis oriundos do *dataset* desenvolvido por Maricato e Manso (2022), 96 contas provenientes do estudo de Nishikawa-pacher (2023) e 1.428 contas extraídas do *dataset* de Mongeon, Bowman e Costas (2023).

Essa etapa de curadoria e integração de dados contribuiu para a formação de um conjunto robusto e validado de contas rotuladas, essencial para o treinamento supervisionado do modelo de inteligência artificial proposto nesta pesquisa.

Os resultados apresentados nesta seção oferecem suporte direto a dois conceitos previamente apresentados neste trabalho: a ideia de que modelos treinados com dados altimétricos específicos tendem a superar aqueles baseados em dados genéricos, e a premissa de que bases especializadas são mais eficazes na tarefa de classificação. A primeira suposição encontra apoio na constatação de que *datasets*

amplamente utilizados, como o de Gutiérrez (2020), não se mostraram suficientemente representativos no contexto da divulgação científica.

De forma complementar, observa-se que a curadoria e integração de conjuntos de dados voltados especificamente à altmetria — como as bases de Mongeon, Bowman e Costas (2023), Nishikawa-pacher (2023) e Maricato e Manso (2022), favoreceram a formação de um *dataset* mais coerente com o perfil das contas relevantes, aumentando a capacidade discriminativa dos modelos. Esses achados reforçam a importância da utilização de bases customizadas para o treinamento de classificadores no ecossistema altmétrico, mitigando os riscos associados ao uso crítico de bases genéricas.

O *dataset* resultante deste trabalho²¹ servirá como suporte essencial para o treinamento e a validação dos modelos de aprendizagem de máquina nas etapas subsequentes desta pesquisa.

4.2 COMPARAÇÃO DO DESEMPENHO DE DIFERENTES ALGORITMOS DE APRENDIZADO DE MÁQUINA

No desenvolvimento de soluções baseadas em aprendizagem de máquina, a escolha do modelo a ser adotado desempenha um papel central na eficácia da classificação. Com o objetivo de selecionar o algoritmo mais adequado para a detecção de *bots* no contexto altmétrico, foi conduzida uma trajetória investigativa composta por estudos sucessivos que analisaram, testaram e compararam diferentes abordagens de classificação supervisionada, partindo de métodos tradicionais até técnicas mais avançadas.

A primeira etapa dessa trajetória foi apresentada em Pontes e Maricato (2022), no artigo “*Detecção de bots que divulgam artigos científicos no Twitter: contribuições para aprimoramento dos indicadores altmétricos*”. Nesse estudo, propôs-se um workflow implementado na ferramenta KNIME para classificar contas como humanas ou *bots*, utilizando o algoritmo KNN. Foram analisadas 6.301 contas, das quais 6.094 puderam ser classificadas com base no modelo, resultando em uma distribuição de 2.900 contas (48%) classificadas como humanas e 3.194 (52%) como *bots*.

²¹ Disponível em: <https://zenodo.org/records/17050653>
(DATA/DAD_15_DatasetClassificadoManual.xlsx)

Em Pontes e Maricato (2023), ampliou-se a investigação com o uso da ferramenta *Botometer* para classificar 18.316 contas que divulgaram artigos científicos na plataforma X. O estudo identificou três categorias principais de usuários — *bots*, humanos e ciborgues — e revelou que aproximadamente 27% das contas eram automatizadas, destacando a influência significativa desses agentes na construção dos indicadores altimétricos. Esses achados reforçaram a necessidade de desenvolver mecanismos mais eficazes para filtrar interações não autênticas, a fim de preservar a integridade das métricas de atenção social.

Posteriormente, em Pontes e Maricato (2024), foi realizada uma análise comparativa entre três algoritmos supervisionados amplamente reconhecidos na literatura: KNN, Árvore de Decisão (*Decision Tree*) e Floresta Randômica (*Random Forest*). O objetivo foi verificar a acurácia na classificação de contas com base em características estruturais dos perfis. O *Random Forest* apresentou o melhor desempenho entre os três, com acurácia superior na classificação por categoria. No entanto, os resultados ainda não foram considerados plenamente satisfatórios, acusando a necessidade de testar abordagens mais robustas e sensíveis à complexidade do fenômeno investigado.

Nesse sentido, a pesquisa incorporou um novo eixo de análise voltado à linguagem natural e aos aspectos comportamentais dos usuários. Em Pontes e Maricato (2024) foi realizada uma análise de sentimentos com apoio do modelo ChatGPT, aplicada às postagens de contas com elevado número de interações em torno de uma publicação científica. O estudo revelou padrões distintos entre *bots* e humanos quanto ao conteúdo textual, vocabulário e expressividade emocional.

Essa linha de investigação foi aprofundada em Pontes e Maricato (2025), que comparou o desempenho dos modelos ChatGPT 3.5 e Gemini 1.5 Flash na análise de sentimentos de postagens sobre COVID-19 no X. Foram analisadas 3.493 postagens de 41 contas previamente classificadas como *bots* ou humanas. Os resultados refletiram que *bots* tendem a emitir mensagens mais polarizadas e negativas, enquanto humanos apresentaram maior neutralidade e consistência. Além disso, os *Large Language Models* (LLM) se mostraram úteis, não apenas na tarefa classificatória, mas principalmente como ferramentas para extração de *features* linguísticas, permitindo capturar aspectos semânticos e interacionais relevantes. A linguagem natural, nesse contexto, passa a representar uma dimensão fundamental

para distinguir contas automatizadas, pois reflete nuances da comunicação interpessoal que *bots* ainda têm dificuldade em reproduzir com autenticidade.

Paralelamente, a ausência de um conjunto de dados rotulados específicos para *bots*, voltados à divulgação científica, exigiu o uso de bases genéricas. Assim, os experimentos iniciais foram realizados com o *dataset* disponibilizado por Gutiérrez (2020) na plataforma *Kaggle*, composto por 37.437 contas do X rotuladas como “*bot*” ou “humano”.

Como desdobramento dessa trajetória e com base nas limitações observadas nos algoritmos previamente testados, optou-se, nesta tese de doutorado, por implementar um experimento sistemático comparativo entre os quatro algoritmos supervisionados: KNN, Árvore de Decisão, *Random Forest* e XGBoost. O objetivo foi testar o modelo mais eficaz para ser utilizado na construção de uma solução automatizada, capaz de classificar contas da plataforma X entre humanas e *bots*, com aplicabilidade direta na filtragem de menções altmétricas.

Para garantir isonomia nos experimentos e comparabilidade dos resultados, os modelos foram treinados e avaliados a partir de um pipeline unificado de pré-processamento, incluindo tratamento de dados ausentes, normalização, balanceamento de classes e particionamento estratificado. O XGBoost foi incorporado por sua reconhecida capacidade de lidar com grandes volumes de dados, alta complexidade e variáveis ruidosas, além de apresentar mecanismos integrados de regularização e controle de *overfitting* — características particularmente relevantes no contexto da detecção de *bots*.

A seguir, são apresentadas as análises comparativas realizadas entre os quatro algoritmos avaliados, evidenciando os critérios de avaliação utilizados, os ajustes de hiperparâmetros aplicados e os resultados alcançados com cada abordagem.

A aplicação comparativa dos algoritmos supervisionados KNN, Árvore de Decisão, *Random Forest* e XGBoost foi realizada por meio do script em python, com o objetivo de identificar qual deles é mais eficiente para detecção de contas automatizadas (*bots*) na plataforma X. O experimento utilizou um conjunto de dados contendo 46 variáveis preditoras relacionadas a atributos textuais, temporais, de perfil, legibilidade e comportamento das contas e publicações, com ênfase na capacidade dos modelos em reconhecer com precisão a classe minoritária (*bots*).

Todos os modelos foram treinados e avaliados com base em um conjunto de dados estratificado, dividido em 70% para treino e 30% para teste. Foi aplicado um

pipeline unificado de pré-processamento para garantir isonomia nos experimentos. Valores ausentes foram tratados com imputação pela média (`SimpleImputer(strategy="mean")`) e, quando necessário, os dados foram normalizados com StandardScaler. O foco principal da avaliação foi a classe *bot* (rotulada como 1), utilizando as métricas: *F1-score*, *Recall*, *Precision*, Acurácia, AUC (Área sob a curva ROC) e Coeficiente Kappa de Cohen. O limiar de decisão (*threshold*) foi ajustado com base na curva de precisão-revocação do modelo XGBoost, sendo definido em 0,82 por apresentar o melhor equilíbrio entre *Precision* e *Recall* para *bots*.

A Tabela 3 sintetiza o desempenho dos modelos testados na identificação de contas automatizadas na plataforma X. São apresentadas métricas específicas para a classe *bot* — *F1-score*, *recall* e *precision* —, bem como indicadores globais de robustez e qualidade preditiva, como a área sob a curva ROC (ROC_AUC), o coeficiente Kappa e a acurácia.

Tabela 3 - Desempenho dos modelos na identificação de *bots* na plataforma X, com métricas específicas para a classe *bot* e indicadores gerais de acurácia e robustez

Modelo	F1_bot	Recall_bot	Precision_bot	ROC_AUC	Kappa	Accuracy
KNN	0.024	0.012	0.600	0.750	0.021	94.0%
Árvore de Decisão	0.406	0.433	0.382	0.694	0.366	92.4%
Floresta Randômica	0.143	0.077	1.000	0.933	0.135	94.5%
XGBoost	0.548	0.579	0.520	0.939	0.517	94.3%

Fonte: Dados de pesquisa (2025)

O primeiro modelo testado foi o KNN, configurado com `n_neighbors=5`, métrica padrão e Euclidiana e pesos uniformes. Como o algoritmo é sensível à escala dos dados, foi aplicada normalização nos dados de entrada. Apesar do *Precision* aparentemente elevado (0,60), o modelo apresentou desempenho extremamente insatisfatório, com *Recall* de apenas 1,21% e F1-score de 0,024, como pode ser observado na Tabela 3. Isso revela que o modelo raramente identifica corretamente uma conta automatizada, mesmo com o *threshold* ajustado. O AUC de 0,749 sugere alguma capacidade de separação, mas que não se traduz em uma classificação

prática eficaz. Esses resultados demonstram que o KNN é inadequado para bases de dados desbalanceadas, como as utilizadas neste estudo.

Em seguida, foi aplicada a Árvore de Decisão, com os parâmetros padrão do *scikit-learn* e *random_state*=42 para reproduzibilidade. Este modelo, por natureza, não exige escalonamento dos dados. A árvore apresentou resultados mais equilibrados em relação ao KNN, com *F1-score* de 0,406 para a classe *bot*, *Recall* de 43,32% e *Precision* de 0,38. O valor do Coeficiente Kappa foi de 0,365, espelhando um desempenho significativamente acima do acaso. Apesar de inferior ao XGBoost, a Árvore de Decisão oferece como vantagem adicional à interpretabilidade dos resultados, podendo ser considerada uma opção em contextos em que a transparência do modelo é prioritária.

A Floresta Randômica, por sua vez, foi configurada com *n_estimators*=100, *class_weight=balanced* para lidar com o desbalanceamento, e *random_state*=42. Também sem necessidade de escalonamento, o modelo apresentou *Precision* perfeita (1.000) para a classe *bot*, mas com um *Recall* muito baixo (7,69%), resultando em um *F1-score* de apenas 0,143. A alta precisão (*Precision*) denota que os poucos *bots* que foram classificados como tal, realmente pertenciam à classe, mas o baixo índice apresentado no *Recall* evidencia que a maioria dos *bots* foi classificada erroneamente como humana. A AUC de 0,933 sugere que o modelo tem capacidade de distinguir as classes. O modelo apresentou acurácia global de 94,48% e Kappa de 0,135.

Por fim, foi aplicado o XGBoost, modelo que apresentou o melhor desempenho global. A configuração utilizada incluiu *max_depth*=3, *min_child_weight*=5, *gamma*=2, *subsample*=0,8, *colsample_bytree*=0,7, *learning_rate*=0,03 e *n_estimators*=200, com *scale_pos_weight* ajustado com base na razão entre classes (*bots* e humanos). O modelo foi treinado sobre os dados escalonados, utilizando todas as 46 variáveis previamente extraídas. O XGBoost obteve *F1-score* de 0,5479, *Recall* de 57,89% e *Precision* de 0,52 para *bots*. A AUC foi a mais alta entre os modelos (0,9392), bem como o Coeficiente Kappa (0,517), externando um equilíbrio ideal entre detecção da classe minoritária e controle de falsos positivos. Esses resultados confirmam a superioridade do XGBoost na tarefa de identificação de contas automatizadas neste cenário, tanto em termos quantitativos quanto em robustez estatística.

Conclui-se que o XGBoost foi o modelo com melhor desempenho entre os algoritmos avaliados, seguido pela Árvore de Decisão. A Floresta Randômica apresentou resultados promissores, embora dependa de ajustes específicos de limiar

para a adequada identificação da classe minoritária. Já o KNN demonstrou desempenho insatisfatório, sendo considerado inadequado para este contexto.

A adoção do XGBoost justifica-se, não apenas pelas limitações observadas nas abordagens anteriores, mas também por suas capacidades superiores na modelagem de padrões comportamentais complexos e sutis. Sua estrutura baseada em *gradient boosting* confere elevada sensibilidade a correlações não lineares entre variáveis, característica fundamental para a tarefa de distinguir perfis humanos de automatizados no contexto da disseminação científica em redes sociais.

A escolha de um modelo mais robusto contribuiu significativamente para o aprimoramento metodológico da presente pesquisa, permitindo o desenvolvimento de estratégias classificatórias mais eficazes e sensíveis às particularidades do ambiente altmétrico, com implicações relevantes para a avaliação da atenção social à produção científica.

Os resultados apresentados nesta seção reforçam que o algoritmo XGBoost apresenta desempenho superior na tarefa de classificação de contas automatizadas em comparação aos demais modelos testados. A aplicação de um pipeline padronizado e a análise sistemática das métricas revelaram que o XGBoost obteve os melhores índices de F1-score, Recall, AUC e Coeficiente Kappa, superando significativamente os algoritmos KNN, Árvore de Decisão e Floresta Randômica. A robustez estatística, aliada à capacidade do modelo de lidar com desbalanceamento e capturar padrões complexos em dados heterogêneos, justifica sua adoção como solução central nesta pesquisa. Dessa forma, a hipótese é integralmente sustentada pelos achados empíricos, reforçando a adequação do XGBoost no contexto da altmetria e da detecção de *bots*.

4.3 MODELO DE INTELIGÊNCIA ARTIFICIAL OTIMIZADO, COM BASE NO ALGORITMO XGBOOST, PARA CLASSIFICAR CONTAS

A metodologia desenvolvida para o modelo de inteligência artificial baseado no algoritmo XGBoost, detalhada na seção 4, foi aplicada conforme descrito no seção 5. Essa aplicação resultou em um classificador com desempenho expressivo na tarefa de distinção entre contas humanas e automatizadas (*bots*) na plataforma X. O modelo final atingiu uma área sob a curva ROC (ROC AUC) de 0,9392, o que revela excelente capacidade discriminativa entre as duas classes. Para a classe minoritária (*bots*), foi

identificada uma pontuação F1 de 0,5479 e um índice Kappa de 0,5175, valores que notabilizaram concordância substancial entre os rótulos reais e as previsões realizadas pelo classificador. O limiar de decisão ótimo (*threshold*) foi calibrado em 0,82, maximizando o desempenho na detecção de *bots* — foco principal da pesquisa.

Esses resultados reforçam a eficácia do modelo treinado, mesmo em um cenário de desbalanceamento, em que as contas humanas superam numericamente as automatizadas. A acurácia geral foi de 94,29%, com alta precisão e recall para a classe humana (f1-score de 0,9695), mas desempenho moderado para a classe de *bots*, refletindo os desafios inerentes à detecção automatizada de comportamentos artificiais.

Esses achados demonstram que a utilização de modelos baseados em XGBoost, aliados a estratégias de explicabilidade e otimização de limiar de decisão, constitui uma abordagem promissora para a classificação automatizada de contas do X, contribuindo diretamente para a mitigação de vieses em indicadores altmétricos e para a integridade da disseminação científica em redes sociais.

Quanto a otimização do modelo, a análise comparativa entre conjuntos com 05 a 46 *features* mostrou que uma redução excessiva de atributos compromete a performance, enquanto conjuntos mais amplos e selecionados aumentam a capacidade discriminativa. Esses achados reforçam a necessidade de modelos especializados e bases de dados representativas para a detecção eficaz de *bots* em ambientes científicos.

Essas descobertas corroboram a literatura recente, segundo a qual a presença de *bots* em redes sociais acadêmicas tende a ser subestimada quando métodos tradicionais de detecção são empregados. Demonstram, ainda, a efetividade da abordagem adotada em aumentar a sensibilidade do modelo para a detecção de *bots*, corroborando a necessidade de estratégias robustas de balanceamento de classes, conforme discutido por Varol *et al.* (2017), Madahali; Tian (2023) e Pozzana e Ferrara (2020).

4.3.1 Importância das *Features*

O Gráfico 1 apresenta um *summary plot* dos valores SHAP (*SHapley Additive exPlanations*), técnica de interpretação baseada na teoria dos jogos, utilizada para explicar os impactos das variáveis preditoras em modelos complexos de aprendizado

de máquina, como o XGBoost. Esse tipo de visualização demonstra, de forma integrada, como as características comportamentais e de conteúdo influenciam a predição de *bots*, permitindo compreender tanto a importância quanto o efeito individual de cada *feature* na saída do modelo — aspecto fundamental em domínios como a detecção de *bots* em mídias sociais, onde comportamentos automatizados frequentemente se disfarçam como humanos.

A visualização revela que a variável QTDPOSTS, representante da quantidade total de postagens sobre artigos científicos por conta, é a que exerce maior impacto preditivo sobre o modelo. Observa-se que valores altos dessa variável (coloridos em tons rosados) tendem a empurrar a predição para a classe positiva (*bot*), o que é coerente com a literatura especializada, pois *bots* são frequentemente programados para publicar com alta frequência. A contribuição positiva de valores altos dessa variável para a classificação como *bot* exprime que o volume de atividade é um forte indício de automação.

Na mesma direção, a variável **media_tempo_entre_posts**, que representa o intervalo médio de tempo entre postagens, apresenta impacto significativo inverso: valores baixos dessa variável (ou seja, postagens feitas em curtos intervalos) aumentam a probabilidade de a conta ser classificada como *bot*. Isso reforça a ideia de que a cadência de postagem acelerada, muitas vezes incompatível com padrões humanos, é uma característica-chave de contas automatizadas.

Outra variável de destaque é **media_exclamacoes**, que mede a frequência de uso de pontos de exclamação. O uso excessivo de pontuação enfática pode denotar um comportamento automatizado que visa chamar atenção ou simular entusiasmo humano, funcionando como um sinal heurístico de persuasão artificial. A variável **media_seguidores**, por sua vez, apresenta forte impacto negativo quando seus valores são baixos, sugerindo que contas com poucos seguidores tendem a ser classificadas como *bots*. Isso se deve ao fato de que contas automatizadas geralmente apresentam baixa atratividade social, sendo seguidas por poucas outras contas reais.

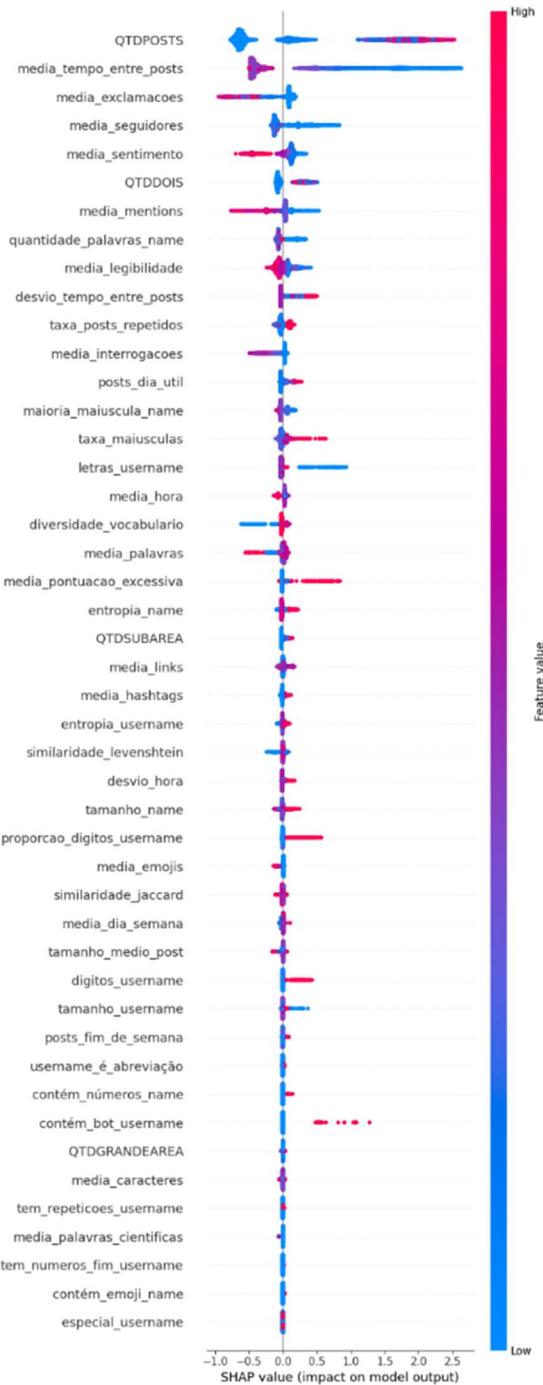
A variável **media_sentimento**, que reflete a polaridade média do grupo de *posts* tuitados por uma conta, também tem influência considerável. O modelo exprime que valores extremos dessa variável estão associados a comportamentos automatizados, o que pode se dever ao fato de que *bots* são utilizados para disseminar conteúdo emocionalmente carregado, muitas vezes polarizado, como forma de

engajamento ou manipulação discursiva. Já **QTDDOIS**, variável de contagem que representa o número de artigos que uma conta postou, parece representar uma métrica complementar de atividade, como número de repostagens ou respostas e se mostra fortemente relacionada à classificação como *bot*, apontando que padrões de interação específicos também são preditivos da automação.

Adicionalmente, variáveis como **media_mentions** e **quantidade_palavras_name** sugerem que a forma como *bots* interagem (marcando frequentemente outros usuários) e como se identificam (com nomes longos ou complexos) também são fatores discriminantes. A variável **media_legibilidade** apresenta forte impacto negativo: textos com baixa legibilidade, uma característica comum em mensagens geradas automaticamente, influenciam a predição de *bots*. Essa evidência reforça o papel das métricas linguísticas como discriminadores eficazes de autoria humana versus automatizada.

Outras variáveis como **desvio_tempo_entre_posts**, **taxa_posts_repetidos** e **media_interrogacoes** contribuem para o modelo, ao capturar nuances temporais e estilísticas do comportamento de postagem, evidenciando que padrões rítmicos e estilísticos artificiais são eficazes para separar *bots* de humanos. Embora o impacto dessas variáveis seja relativamente menor que o das variáveis mais destacadas, sua presença reforça a robustez do modelo ao considerar aspectos multifacetados da atividade social on-line.

Por fim, o conjunto de variáveis relacionadas à estrutura dos nomes e usernames, como **entropia_name**, **letras_username**, **proporcao_digitos_username**, **contém_bot_username** e **username_é_abreviação**, demonstra que o modelo também é sensível a padrões nominais que frequentemente denunciam a artificialidade de uma conta. Usernames com alta aleatoriedade, presença de números, abreviações ou termos como “bot” são comumente encontrados em contas automatizadas e funcionam como fortes indicativos de sua natureza não humana.

Gráfico 1 - *Summary plot* dos valores SHAP das *features* geradas

Fonte: Dados de Pesquisa - PDF_RelatorioModeloXgboost..pdf (2025).

Em síntese, o gráfico SHAP, apresentado no Gráfico 1, fornece uma interpretação transparente do processo decisório do modelo de classificação, evidenciando que variáveis associadas à frequência de postagens, padrões de tempo, características linguísticas e estrutura de nomes são os principais determinantes para a identificação de *bots* em redes sociais. A análise reforça a importância da

modelagem comportamental e linguística na construção de detectores automáticos de contas falsas e sobressai como a integração de diferentes domínios de atributos — atividade, textualidade e identidade — contribui para o desempenho preditivo do modelo.

4.3.2 Avaliação do Modelo Gerado

Os resultados obtidos com o modelo de aprendizado de máquina acusam um desempenho global robusto, realçado principalmente pela área sob a curva ROC (AUC ROC) de 0,9392, como pode ser observado na Figura 4. Esse valor sugere excelente capacidade discriminativa entre as classes "bot" e "humano", demonstrando que o modelo é altamente eficaz em atribuir escores de probabilidade diferenciados às instâncias de cada classe ao longo de todo o espectro de limiares. Essa característica é particularmente desejável em contextos de classificação binária com forte desbalanceamento, como é o caso da presente aplicação.

Figura 4 - Relatório de Classificação – XGBoost

```
==== RELATÓRIO DE CLASSIFICAÇÃO ====
Threshold ótimo (focado em F1 para Bots): 0.82
F1 (classe BOT): 0.5479
ROC AUC: 0.9392
Kappa: 0.5175

Classification Report:
precision      recall      f1-score      support
          0      0.9730      0.9660      0.9695      3884
          1      0.5200      0.5789      0.5479       247
accuracy                           0.9429      4131
macro avg       0.7465      0.7725      0.7587      4131
weighted avg    0.9459      0.9429      0.9443      4131
```

Fonte: Dados de Pesquisa - relatorio_modelo_xgboost_final 06052025.pdf (2025).

A avaliação do modelo XGBoost foi conduzida com foco na otimização do F1-score da classe minoritária (*BOT*), utilizando como estratégia a calibragem do limiar de decisão. O limiar ótimo identificado foi de 0,82 — superior ao ponto de corte padrão de 0,5 — o que representa uma escolha metodologicamente adequada diante da

assimetria entre as classes. Tal ajuste visa mitigar o viés do classificador em favor da classe majoritária, favorecendo a sensibilidade do modelo para a detecção de *bots*, mesmo que à custa de certa perda em precisão.

A adoção desse limiar resultou em um F1-score de 0,5479 para a classe *BOT*, refletindo um equilíbrio entre precisão (0,5200) e revocação (recall = 0,5789). Embora esse desempenho possa parecer moderado, é representativo de uma efetividade razoável, dada a proporção reduzida da classe de interesse no conjunto de dados (apenas 247 instâncias de um total de 4.131). Em cenários com classes desbalanceadas, o F1-score é amplamente reconhecido como uma das métricas mais adequadas, pois considera simultaneamente os erros do tipo I e II, oferecendo uma visão mais realista da capacidade de generalização do modelo.

A acurácia global obtida foi de 94,29%, um valor elevado, mas que requer interpretação cautelosa, dado que essa métrica pode ser inflacionada em contextos com predominância de uma classe. Por esse motivo, foi também calculado o índice Kappa de Cohen, cujo valor foi de 0,5175. Este coeficiente ajusta a acurácia, levando em consideração a classificação esperada pelo acaso. Neste caso, aponta para um nível moderado de concordância entre as previsões do modelo e os rótulos reais — o que é indicativo de um desempenho confiável, mesmo sob forte desbalanceamento.

A análise detalhada do relatório de classificação confirma essas observações. A classe 0 (HUMANO), amplamente majoritária, apresenta métricas quase perfeitas: precisão de 0,9730, revocação de 0,9660 e F1-score de 0,9695. Já a classe 1 (*BOT*), apesar do desempenho inferior, alcança resultados satisfatórios, considerando sua baixa representatividade. As médias macro e ponderada do F1-score (0,7587 e 0,9443, respectivamente) oferecem conclusões adicionais: enquanto a média ponderada reflete o domínio estatístico da classe humana no total de instâncias, a média macro equilibra a contribuição de ambas as classes, notabilizando que, mesmo diante do desbalanceamento, o modelo conseguiu preservar uma performance razoável na detecção de *bots*.

Em síntese, os resultados demonstram que o modelo XGBoost, ajustado por meio da calibragem do limiar de decisão, é efetivo para a tarefa de identificação de contas automatizadas em redes sociais. A combinação de elevada AUC ROC, concordância moderada, segundo o índice Kappa e desempenho relativamente satisfatório na classe minoritária, ressalta que o modelo está ajustado de forma apropriada às características complexas do problema. O uso de métricas

complementares à acurácia, como F1-score e AUC, revela-se essencial para avaliações criteriosas da performance em cenários marcados por assimetrias entre classes e consequências práticas relevantes da detecção correta.

Ao final do processo, foram salvos todos os artefatos do modelo, incluindo: o classificador treinado (MDL_ModeloXgboost.pkl), o objeto de escala (MDL_Scaler.pkl), o imputador (MDL_Imputer.pkl), o *threshold* ótimo (MDL_ThresholdOtimo.pkl) e a lista das 46 *features* coletadas (MDL_Top46Features.pkl). Esses componentes viabilizam a replicação do experimento e sua posterior aplicação em novos conjuntos de dados.

Os resultados e as discussões deste estudo são apresentados nesta seção, organizados conforme os objetivos específicos estabelecidos. Inicialmente, são apresentados os resultados da comparação de diferentes metodologias de aprendizagem de máquina para a classificação de contas da plataforma X, de acordo com o tipo de perfil (*bot* ou humano). Em seguida, são apresentados os resultados e as devidas análises dos efeitos da utilização de distintos conjuntos de dados de treinamento sobre o desempenho dos modelos de classificação.

Também são apresentados resultados obtidos com a aplicação da metodologia desenvolvida para construção de uma base de dados rotulada, contemplando a classificação das contas por tipo de perfil, a qual serviu de suporte para o desenvolvimento de um modelo de inteligência artificial.

O desenvolvimento do modelo de IA, dedicado à classificação automática de contas do X, proposto na metodologia (seção 4), é apresentado como resultado na seção 5, onde são detalhados os algoritmos e os dados utilizados em cada etapa do processo, garantindo, assim, a repetibilidade da pesquisa. Nesta ocasião, discute-se sobre as dificuldades e os avanços obtidos.

A análise quantitativa da distribuição de contas que compartilham artigos científicos permitiu identificar proporções relativas a perfis humanos e *bots*, contribuindo para uma melhor compreensão da dinâmica de disseminação científica na plataforma. Além disso, padrões comportamentais específicos foram identificados em diferentes áreas do conhecimento, considerando o tipo de perfil e a categoria de usuário.

Os resultados obtidos com a implementação do modelo final baseado em XGBoost e a análise de interpretação com SHAP permitem sustentar algumas das hipóteses formuladas nesta pesquisa. A hipótese H02 – *Bots* postam mais que

humanos, mesmo sendo minoria, segundo a qual *bots* tendem a publicar com maior frequência do que usuários humanos, ainda que representem uma minoria no conjunto analisado. O impacto da variável QTDPOSTS na classificação de *bots*, conforme evidenciado pelos valores SHAP, confirma essa tendência. Adicionalmente, foi constatado que contas humanas têm menor volume de postagens e foco temático mais específico, sugerindo que contas humanas apresentam menor volume de postagens e maior especificidade temática. Tal constatação foi corroborada pela análise das variáveis media_tempo_entre_posts, media_legitibilidade e media_sentimento. Assim, tais indicadores revelam que contas humanas tendem a operar com cadência mais espaçada e conteúdo mais legível, características opostas às observadas nos perfis automatizados.

Por fim, na próxima seção, é apresentado um modelo computacional e um protótipo funcional como prova de conceito, visando à automação da classificação de contas com influência altmétrica.

4.4 PROVA DE CONCEITO DO MODELO COMPUTACIONAL E PROTÓTIPO PARA AUTOMAÇÃO DE CLASSIFICAÇÃO DE CONTAS DO X

O *SciBotScan*²² é um protótipo computacional interativo, baseado em aprendizado de máquina, desenvolvido com o objetivo de estimar a probabilidade de uma conta da plataforma X ser automatizada (*bot*) ou humana, com foco específico na disseminação de artigos científicos. A proposta se insere no contexto da altmetria e visa fortalecer a confiabilidade dos indicadores de impacto alternativos ao identificar comportamentos automatizados na divulgação científica. A aplicação do protótipo está inserida no contexto da altmetria, sendo voltada à mitigação dos efeitos da automação na disseminação científica e na construção de indicadores de impacto alternativos.

O sistema foi construído a partir de um modelo de aprendizado de máquina supervisionado, utilizando o algoritmo XGBoost, selecionado por sua alta capacidade preditiva, robustez a dados desbalanceados e boa interpretabilidade.

O treinamento foi realizado sobre uma base rotulada composta por 13.767 contas, utilizando 46 *features* preditivas que abarcam quatro grandes categorias: textualidade das postagens, comportamento temporal, atividade da conta e estrutura

²² Disponível em: <https://scibotscanv3.streamlit.app/>

dos nomes de usuário. A classificação (*bot* ou *humano*) foi obtida a partir da integração de métodos automatizados, validações cruzadas com bases reconhecidas da literatura e inspeção manual. Ao todo, o *dataset* continha 822 contas de *bots* e 12.945 contas humanas, com mais de 67 mil postagens analisadas.

A Figura 5 apresenta a interface do protótipo desenvolvido, denominado *SciBotScan*, projetado para a classificação de contas que divulgam publicações científicas na plataforma X.

Figura 5 - Interface do protótipo para classificação de contas (SciBotScan)



Fonte: Dados de Pesquisa (2025).

A interface apresenta um campo para inserção de informações do perfil, como nome da conta, identificador, número de seguidores, quantidade de *posts* e de artigos compartilhados, além de dados temporais (datas e horários das postagens). Também é possível colar diretamente o conteúdo dos *posts* que deverão ser analisados. Após a execução do modelo, o sistema fornece o resultado da classificação, demonstrando

se a conta é humana ou automatizada, acompanhado da probabilidade estimada. A parte inferior da interface exibe um gráfico semicircular de interpretação imediata, que representa visualmente a probabilidade de automação atribuída à conta avaliada. Essa interface busca aliar clareza e praticidade, permitindo que pesquisadores e usuários interpretem os resultados de forma rápida e acessível.

4.4.1 Tecnologias e ferramentas utilizadas

O protótipo foi implementado em Python 3.12 e emprega um conjunto de bibliotecas robustas para machine learning, pré-processamento e visualização:

- **Modelagem e persistência de artefatos:**
 - XGBoost: construção do modelo classificatório;
 - *scikit-learn*: escalonamento de dados, imputação de valores faltantes;
 - joblib: salvamento e carregamento de modelos, escaladores e imputadores.
- **Pré-processamento e extração de *features*:**
 - pandas e numpy: manipulação tabular e vetorial;
 - textblob, textstat, emoji: análise textual e emocional;
 - re, unidecode, difflib: extração de padrões de nomes e similaridade.
- **Visualização:**
 - matplotlib: geração de gráficos personalizados (incluindo o gráfico do tipo gauge chart semicircular, que representa visualmente a probabilidade de automação).
- **Interface de usuário:**
 - Streamlit: construção da interface interativa, responsiva e leve, permitindo ao usuário testar a classificação de novas contas em tempo real;
 - Estilização personalizada em CSS foi aplicada para garantir uma apresentação visual compatível com ambientes acadêmicos.
- **Repositório e implantação:**
 - GitHub: utilizado como repositório do projeto e controle de versões;

- O protótipo foi configurado para execução em ambiente on-line, sendo possível seu deployment via Streamlit Cloud²³.

4.4.2 Desempenho do modelo

Os principais indicadores de desempenho do modelo treinado foram:

A análise de interpretabilidade baseada em SHAP *values*, que identificou variáveis mais influentes como: número total de postagens, frequência entre postagens, uso de pontuação emocional (exclamações), polaridade de sentimento e número de seguidores.

4.4.3 Variáveis de entrada na interface do usuário (via Streamlit)

O usuário do protótipo insere informações diretamente pela interface web, sendo essas posteriormente processadas para gerar as *features* necessárias. As variáveis de entrada são:

- Dados da conta (df_contas):

NOME: nome do perfil (exibido no X);

ACCOUNT: nome de usuário da conta;

QTDPOSTS: número total de postagens coletadas;

QTDSEGUIDORES: total de seguidores da conta;

QTDDOIS: número de DOIs compartilhados;

QTDSUBAREA: número de subáreas científicas associadas;

QTDGRANDEAREA: número de grandes áreas científicas associadas.

- Dados das postagens (df_posts):

CONTEUDOPOST: texto da postagem (até 10 postagens por conta);

DATATEMPO: data e horário individual de cada postagem, combinados a partir de campos separados fornecidos pelo usuário.

Essas entradas alimentam uma rotina de engenharia de atributos, que gera automaticamente variáveis como legibilidade, uso de hashtags, presença de termos

²³ Disponível em: <https://streamlit.io/cloud>

científicos, emoção textual (sentimento), padrão de horário das postagens, entre outros. Também são extraídas métricas complexas como entropia, similaridade de nomes e presença de emojis no nome do perfil e no nome de usuário.

O *SciBotScan* representa uma contribuição significativa para o monitoramento da integridade das métricas altmétricas, permitindo maior acurácia na identificação de padrões automatizados de disseminação científica e fomentando práticas mais transparentes na avaliação do impacto social da produção acadêmica.

A implementação do protótipo *SciBotScan* representou um importante desdobramento prático da presente pesquisa, ao materializar os conceitos teóricos e metodológicos discutidos ao longo do trabalho em uma ferramenta funcional e interativa. Sua construção não apenas viabilizou a aplicação real do modelo de detecção de *bots* na plataforma X, mas também reforçou a pertinência das escolhas realizadas ao longo do pipeline desenvolvido. Em especial, a adoção do algoritmo XGBoost como núcleo do sistema foi respaldada pelo desempenho alcançado nos experimentos, o que evidenciou sua robustez e justificou sua centralidade na arquitetura da ferramenta.

A configuração do sistema com base em um conjunto de dados altmétricos rotulados e representativos e a utilização de variáveis comportamentais e textuais — como frequência de postagens, número de DOIs, subáreas e áreas do conhecimento associadas às contas — tornou possível estruturar um arcabouço que poderá ser explorado em investigações futuras sobre padrões de atuação, engajamento e distribuição temática de *bots* e humanos na divulgação científica.

Dessa forma, o protótipo *SciBotScan*, além de uma aplicação técnica, também oferece um ponto de partida concreto para estudos subsequentes, ao aliar rigor analítico, inovação metodológica e aplicabilidade prática no campo da altmetria e da detecção de automatismos em redes sociais acadêmicas. O desenvolvimento do protótipo antecipa sua viabilidade técnica e utilidade prática.

Por fim, a arquitetura do *SciBotScan*, baseada em dados validados manualmente e oriundos de fontes especializadas, oferece evidências adicionais para as hipóteses da pesquisa, demonstrando como a curadoria rigorosa pode ampliar a representatividade e a eficácia de sistemas classificatórios aplicados à altmetria.

4.5 PADRÕES DE COMPORTAMENTO E ENGAJAMENTO DE CONTAS HUMANAS E BOTS

A análise quantitativa da distribuição de contas na plataforma X, a partir da classificação automatizada, realizada pelo modelo de inteligência artificial desenvolvido nesta pesquisa, revela um panorama relevante do comportamento e da participação de diferentes perfis na disseminação de artigos científicos. A base total é composta por 362.395 contas, responsáveis por 758.288 postagens relacionadas a publicações acadêmicas, como pode ser observado na Tabela 4.

Tabela 4 - Presença de Contas *bots* x Contas Humanas

Classe	Qtd account	Qtd Posts	% Contas	% Posts
Bot	13.433	169.219	4%	22%
Humano	348.962	589.069	96%	78%
Total Geral	362.395	758.288		

Fonte: Dados de Pesquisa (2025).

Dentre essas contas, 13.433 foram classificadas como *bots*, representando apenas 4% do total de perfis identificados. Contudo, esses agentes automatizados foram responsáveis por 169.219 postagens, o que corresponde a 22% de todos os conteúdos compartilhados. Por outro lado, as contas humanas, que compreendem 96% da amostra (348.962 contas), realizaram 589.069 postagens, ou seja, 78% do total.

Essa desproporcionalidade entre a representatividade numérica das contas e o volume de postagens evidencia um padrão de comportamento significativamente distinto entre os dois grupos. Os *bots*, ainda que em minoria, demonstram uma capacidade de publicação intensiva, sugerindo estratégias de automação voltadas à amplificação de determinados conteúdos. Em média, cada *bot* realizou cerca de 12,6 postagens, enquanto as contas humanas apresentaram uma média inferior, de aproximadamente 1,7 postagem por conta.

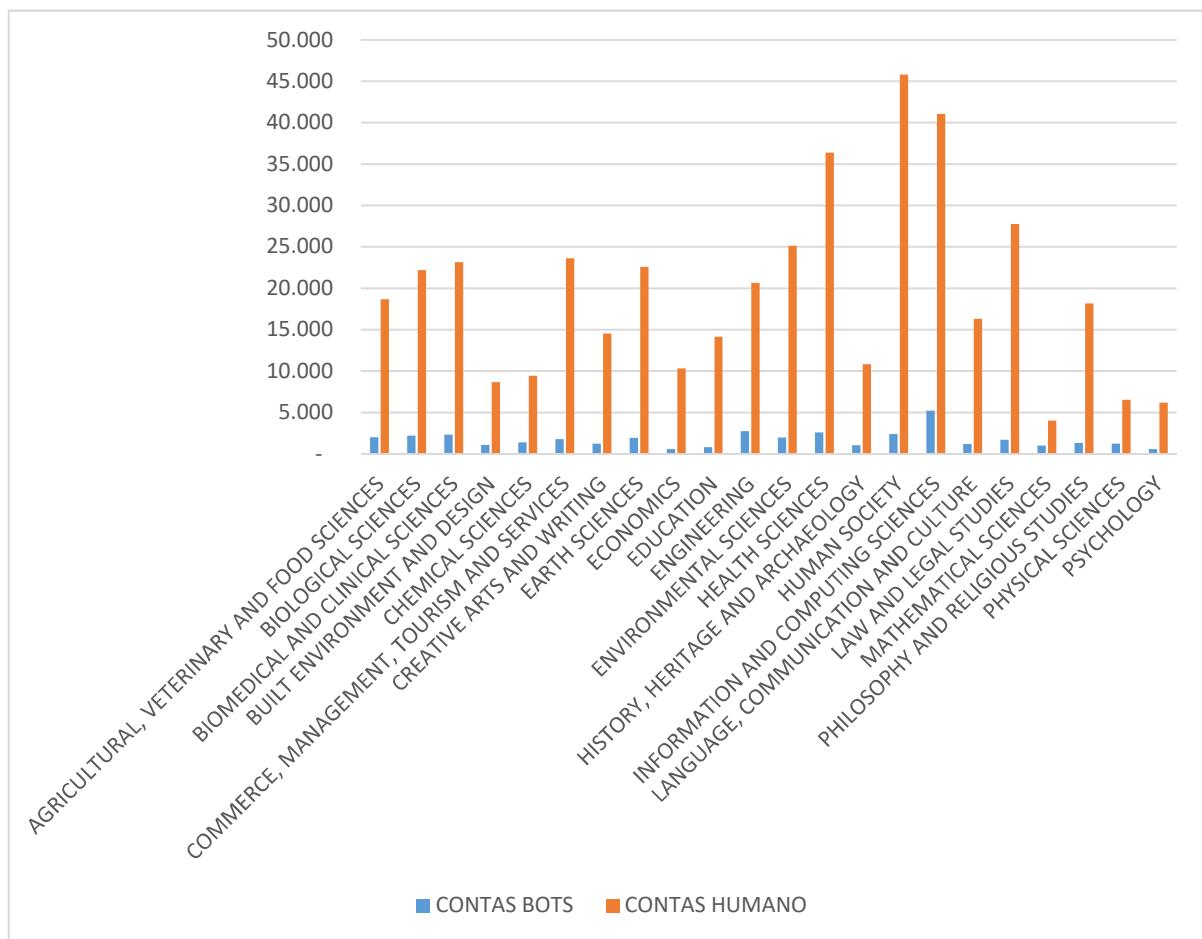
Esse achado reforça discussões anteriores sobre o papel dos *bots* na disseminação massiva e potencialmente artificial de conteúdos científicos, interferindo, direta ou indiretamente, nos indicadores altmétricos. A predominância humana em número de contas expressa um uso mais orgânico da plataforma, porém menos concentrado em volume de compartilhamentos por perfil. Já os *bots*, apesar

de numericamente menos expressivos, exercem um impacto desproporcional sobre a visibilidade de determinados artigos, o que pode distorcer a interpretação do alcance e da influência de publicações científicas quando apenas métricas de volume são consideradas.

Esses resultados sustentam a necessidade de um olhar crítico sobre as métricas baseadas em redes sociais, em especial na altmetria, e a urgência de mecanismos capazes de identificar e mitigar a influência de contas automatizadas na análise do impacto científico. Além disso, os dados corroboram a relevância da presente pesquisa ao desenvolver ferramentas capazes de distinguir perfis com comportamentos divergentes, contribuindo para um ambiente informacional mais transparente e confiável no ecossistema acadêmico digital.

4.5.1 Padrões comportamentais, segmentados por área do conhecimento, entre contas do X, classificadas por tipo de perfil

A análise da distribuição de contas de *bots* e contas humanas por divisão do conhecimento, apresentada no Gráfico 2, revela padrões significativos e discrepâncias que ajudam a compreender a atuação automatizada na disseminação científica em plataformas digitais. Com um total de 38.387 contas de *bots* frente a 426.151 contas humanas, os *bots* representam, em média, 8,3% do total de contas por divisão. Apesar de numericamente inferiores, o impacto dessas contas automatizadas varia consideravelmente entre as diferentes áreas.

Gráfico 2 - Contas *Bots*/Humanas por Divisão

Fonte: Dados de Pesquisa (2025).

Como pode ser observado no Gráfico 2, a divisão com maior número absoluto de contas de *bots* é *Information and Computing Sciences*, com 5.228 perfis automatizados, seguida por *Human Society* (2.410) e *Engineering* (2.729). Essas áreas se destacam tanto pela quantidade de contas, quanto por sua afinidade com tecnologias computacionais ou por sua intensa produção de conhecimento técnico. A presença ampliada de *bots* em *Information and Computing Sciences* pode sugerir um uso estratégico da automação por pesquisadores ou instituições para promover artigos em ambientes digitais de alta rotatividade informacional.

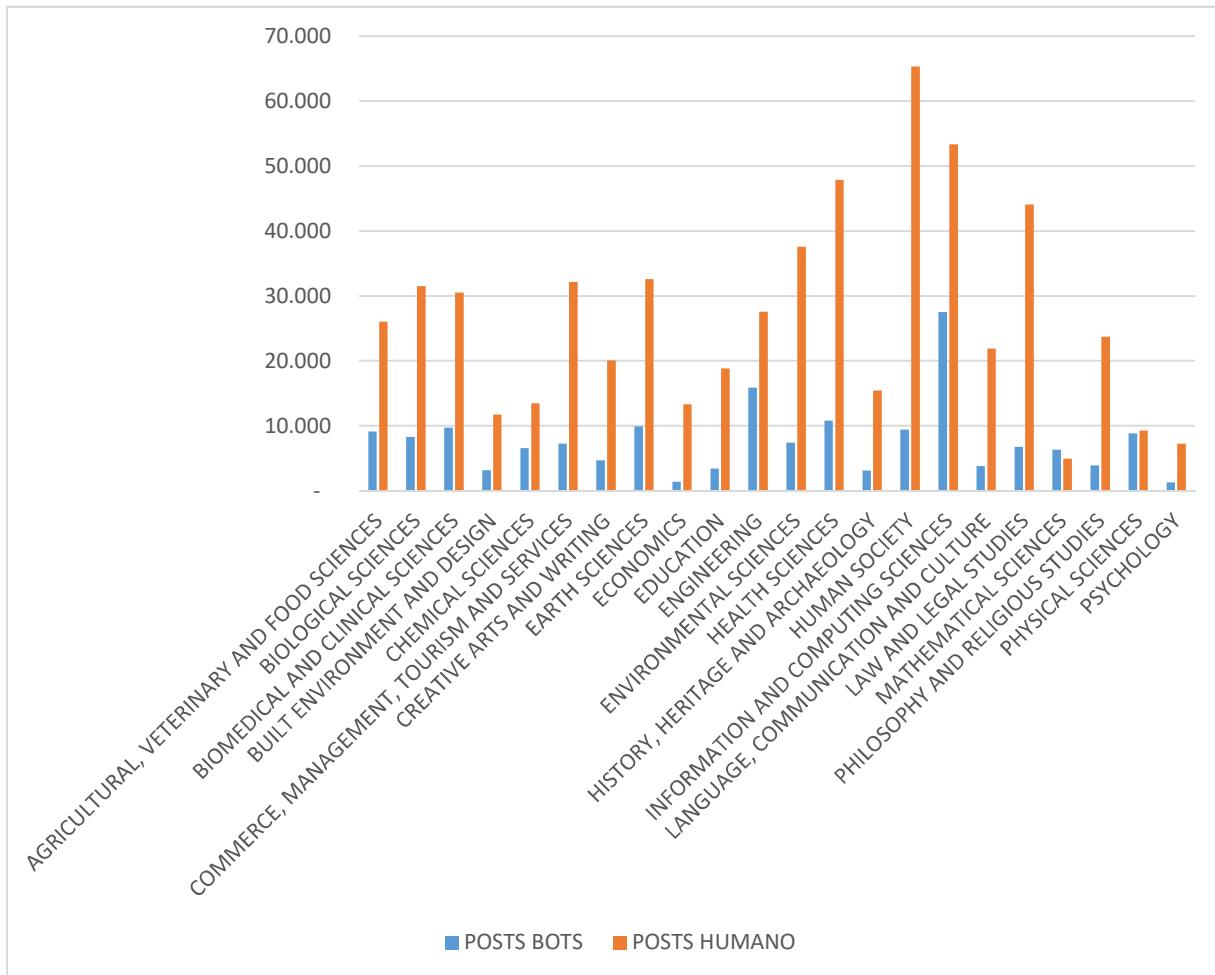
Em termos proporcionais, algumas divisões exibem índices particularmente elevados. *Mathematical Sciences* apresenta a maior concentração relativa de contas de *bots*, com 20,1% das contas sendo automatizadas (1.009 de 5.029 contas totais). Também se destacam *Psychology* (8,7%) e *Engineering* (11,7%). Já em áreas como *Education* (5,4%) e *Law and Legal Studies* (5,8%), a presença de *bots* é consideravelmente menor, sugerindo menor adesão à automação ou possível

predominância de práticas comunicacionais mais convencionais, pautadas em interações humanas diretas.

A dispersão desses valores pode ser quantificada por meio do desvio-padrão da proporção de *bots*, que é de aproximadamente 3,6 pontos percentuais. Isso sugere uma variação moderada, mas significativa, entre as áreas. As divisões com valores acima de um desvio-padrão da média (isto é, superiores a 11,9%) podem ser classificadas como *outliers positivos*, refletindo forte presença automatizada na composição das contas que divulgam conteúdo científico.

Portanto, a partir da análise estratificada por divisão, constata-se que a presença de *bots* na comunicação científica varia de forma relevante entre as áreas do conhecimento, tanto em termos absolutos quanto proporcionais. Esse cenário corrobora a necessidade de monitoramento contínuo da influência automatizada nas métricas altmétricas e sugere que a interpretação desses indicadores deve considerar o contexto tecnológico e comunicacional de cada área.

A análise quantitativa da distribuição de postagens realizadas por *bots* e humanos em diferentes áreas de conhecimento apresentada no Gráfico 3 revela disparidades significativas na participação automatizada na divulgação científica. Com base nos dados fornecidos, observa-se que, embora o volume absoluto de postagens humanas (588.454) seja substancialmente superior ao de *bots* (168.762), a distribuição relativa entre as divisões evidencia padrões que merecem investigação mais profunda.

Gráfico 3 - Posts gerados por *bots*/humanos por Divisão

Fonte: Dados de Pesquisa (2025).

Conforme demonstrado no Gráfico 3, do ponto de vista proporcional, algumas divisões apresentam uma presença de *bots* significativamente acima da média geral, que é de aproximadamente 22,28%. Destaca-se, por exemplo, a área de *Mathematical Sciences*, em que os *bots* são responsáveis por 56,2% das postagens, tornando-se a única divisão em que os *bots* superam numericamente os humanos. Situação semelhante, embora menos acentuada, ocorre em *Physical Sciences* (48,8%) e *Information and Computing Sciences* (34,0%).

Tais discrepâncias sugerem que áreas com forte presença computacional, matemática ou técnicas específicas de automação podem ser mais suscetíveis à atuação de contas automatizadas. Isso pode decorrer tanto de estratégias institucionais que automatizam a divulgação de artigos nessas áreas quanto da adoção de práticas padronizadas que facilitam a replicação de conteúdos por *bots*.

Por outro lado, divisões como *Human Society* (12,6%), *Law and Legal Studies* (13,3%) e *Health Sciences* (18,4%) apresentam percentuais de postagens feitas por *bots* consideravelmente abaixo da média, o que pode estar associado a uma comunicação científica mais centrada em perfis pessoais, acadêmicos ou institucionais geridos por humanos, além da menor padronização dos conteúdos divulgados, dificultando a automação.

A análise estatística por meio do coeficiente de variação (CV) para a proporção de *bots* por divisão revela uma alta dispersão nos dados ($CV \approx 48\%$), o que confirma a heterogeneidade no comportamento de automação entre as áreas. Essa variabilidade sugere que qualquer generalização do papel dos *bots* na disseminação científica deve considerar as especificidades disciplinares.

Além disso, a correlação de Pearson entre o total de postagens e a proporção de *bots* por área é moderadamente negativa (estimada em torno de -0,43), denotando que áreas com maior volume de postagens tendem a apresentar menor proporção de *bots*. Isso pode refletir um predomínio do engajamento humano em áreas com maior visibilidade pública ou institucional.

Em suma, os dados revelam que a influência dos *bots* na divulgação científica via redes sociais não é homogênea entre os campos do conhecimento. A preponderância de *bots* em divisões específicas pode distorcer os indicadores altmétricos, demandando mecanismos de correção ou ponderação que considerem a origem das postagens.

A correlação entre a distribuição das contas automatizadas (*bots*) e o volume de postagens por divisão do conhecimento permite um aprofundamento analítico que reforça a hipótese de que a influência dos *bots* na disseminação científica não está apenas na quantidade de contas, mas principalmente no volume e no padrão de atividade dessas contas.

Tabela 5 - Quantidade de *posts* e contas *bot/humanos* por divisão

DIVI- SÃO	DESCRÍCÃO	POSTS				CONTAS			
		BOTS		HUMANO		BOTS		HUMANO	
		QTD	%	QTD	%	QTD	%	QTD	%
30	AGRICULTURAL, VETERINARY AND FOOD SCIENCES	9.148	26,0%	26.023	74,0%	2.015	9,7%	18.683	90,3%
31	BIOLOGICAL SCIENCES	8.303	20,9%	31.519	79,1%	2.191	9,0%	22.190	91,0%
32	BIOMEDICAL AND CLINICAL SCIENCES	9.737	24,2%	30.540	75,8%	2.325	9,1%	23.176	90,9%
33	BUILT ENVIRONMENT AND DESIGN	3.172	21,3%	11.744	78,7%	1.085	11,1%	8.664	88,9%
34	CHEMICAL SCIENCES	6.572	32,8%	13.449	67,2%	1.390	12,8%	9.436	87,2%
35	COMMERCE, MANAGEMENT, TOURISM AND SERVICES	7.276	18,5%	32.145	81,5%	1.796	7,1%	23.620	92,9%
36	CREATIVE ARTS AND WRITING	4.683	18,9%	20.086	81,1%	1.234	7,8%	14.520	92,2%
37	EARTH SCIENCES	9.920	23,3%	32.582	76,7%	1.924	7,9%	22.585	92,1%
38	ECONOMICS	1.400	9,5%	13.312	90,5%	594	5,4%	10.314	94,6%
39	EDUCATION	3.423	15,4%	18.841	84,6%	814	5,4%	14.149	94,6%
40	ENGINEERING	15.901	36,6%	27.541	63,4%	2.729	11,7%	20.649	88,3%
41	ENVIRONMENTAL SCIENCES	7.415	16,5%	37.573	83,5%	1.967	7,3%	25.131	92,7%
42	HEALTH SCIENCES	10.834	18,5%	47.861	81,5%	2.590	6,6%	36.369	93,4%
43	HISTORY, HERITAGE AND ARCHAEOLOGY	3.113	16,8%	15.450	83,2%	1.030	8,7%	10.817	91,3%
44	HUMAN SOCIETY	9.436	12,6%	65.317	87,4%	2.410	5,0%	45.814	95,0%
46	INFORMATION AND COMPUTING SCIENCES	27.509	34,0%	53.337	66,0%	5.228	11,3%	41.039	88,7%
47	LANGUAGE, COMMUNICATION AND CULTURE	3.797	14,8%	21.873	85,2%	1.186	6,8%	16.314	93,2%
48	LAW AND LEGAL STUDIES	6.772	13,3%	44.043	86,7%	1.713	5,8%	27.760	94,2%
49	MATHEMATICAL SCIENCES	6.321	56,2%	4.930	43,8%	1.009	20,1%	4.020	79,9%
50	PHILOSOPHY AND RELIGIOUS STUDIES	3.890	14,1%	23.741	85,9%	1.324	6,8%	18.189	93,2%
51	PHYSICAL SCIENCES	8.855	48,8%	9.281	51,2%	1.244	16,0%	6.540	84,0%
52	PSYCHOLOGY	1.285	15,0%	7.266	85,0%	589	8,7%	6.172	91,3%
	TOTAIS	168.762	22%	588.454	78%	38.387	8,26%	426.151	92%

Fonte: Dados de Pesquisa (2025)

Ao considerar os dados previamente analisados — que revelam que *bots* representam, em média, 8,3% das contas — e os dados de postagens, com 168.762 publicações de *bots* (Tabela 5), o que equivale a 22,28% do total de postagens,

observa-se uma desproporcionalidade expressiva: apesar de numericamente minoritários, os *bots* são responsáveis por uma fração consideravelmente maior da atividade de disseminação. Essa diferença sugere que contas automatizadas são, em média, mais ativas que contas humanas, o que pode inflar métricas altmétricas em determinadas áreas (Ferrara, Emilio *et al.*, 2016).

Essa descoberta encontra paralelo no estudo de *Shao et al.* (2018), que analisou o papel dos *bots* sociais na disseminação de desinformação. Segundo os autores, embora esses *bots* representassem apenas 6% das contas analisadas, foram responsáveis por 31% de todos os tweets vinculados a conteúdos de baixa credibilidade e por 34% das postagens sobre artigos provenientes dessas fontes. Esses dados reforçam a hipótese de que os *bots* exercem um impacto desproporcional na circulação de informações, contribuindo para a amplificação de conteúdos potencialmente enganosos.

Tomando como exemplo a divisão *Mathematical Sciences*, que apresenta a maior proporção de contas de *bots* (20,1%), verifica-se também que 56,2% das postagens nessa área são feitas por *bots*. Esse achado reforça a hipótese de que há uma correlação direta entre a presença de contas automatizadas e o volume de postagens automatizadas. O mesmo padrão se repete em *Physical Sciences*, em que os *bots* compõem uma fração expressiva, tanto em número de contas quanto em postagens (48,8%).

Entretanto, algumas áreas apresentam uma dissociação entre o número de contas e a produção de postagens. Por exemplo, *Human Society* registra um alto número absoluto de *bots* (2.410), mas apenas 12,6% das postagens são feitas por essas contas. Isso pode apontar que, embora existam *bots* cadastrados nessas divisões, seu nível de atividade é inferior ao observado em áreas como *Mathematical Sciences* ou *Information and Computing Sciences*. De modo similar, *Psychology*, com 8,7% das contas sendo *bots*, apresenta apenas 15% de postagens automatizadas, abaixo da média global.

Já em *Information and Computing Sciences*, a convergência é clara: com a maior quantidade de contas de *bots* (5.228) e um dos maiores volumes absolutos de postagens automatizadas (27.509), essa divisão se destaca como um núcleo de atividade automatizada altamente expressiva, tanto em presença quanto em produtividade.

A análise estatística da proporção de *bots* por postagens (com desvio-padrão de aproximadamente 11,3 pontos percentuais e coeficiente de variação de 48%) reforça a heterogeneidade entre as divisões. Ao cruzar esse dado com o desvio-padrão das proporções de contas (3,6 p.p.), percebe-se que a variação no uso de automação é mais acentuada quando se considera mais a produção de conteúdo que apenas a existência de contas. Isso revela que, além da presença, a intensidade da atividade dos *bots* é um fator chave para compreender sua influência sobre os indicadores altmétricos.

Dessa forma, os dados apontam que a atuação automatizada não apenas varia entre áreas, mas manifesta-se de forma desproporcional em termos de produção, com implicações diretas na interpretação dos indicadores de impacto. A correlação entre número de contas e volume de postagens automatizadas, embora positiva, não é linear em todas as divisões, o que sugere que políticas institucionais, perfis de periódicos e estratégias de divulgação também podem influenciar esse fenômeno.

4.5.2 Caracterização da amostra de estudo e sua atuação na plataforma X

Esta seção dedica-se à caracterização detalhada da amostra utilizada na etapa empírica da pesquisa, com ênfase na atuação das contas analisadas na plataforma X.

4.5.2.1 Caracterização da coleta primária de dados

A fase inicial da coleta de dados deste experimento consistiu na análise de um amplo conjunto de dados bibliográficos, totalizando 17.915.482 artigos científicos, distribuídos em 144 áreas distintas do conhecimento, conforme a classificação adotada. Dentre o total de artigos coletados, 5.637.521 registros (31,46%) não apresentavam qualquer pontuação altmétrica, o que sugere ausência de menções em fontes monitoradas pela ferramenta Altmetric.com, ou eventualmente, limitações nos processos de indexação. Considerando-se exclusivamente as menções no X, observou-se que 9.800.586 artigos (54,72%) não foram mencionados nessa rede social, sendo, portanto, excluídos das análises subsequentes por não atenderem ao critério de presença na plataforma.

Após esta filtragem, o número de artigos com menções no X totalizou 8.114.896 publicações, o que representa 45,28% do total inicial. Esses registros compõem o núcleo da amostra analítica utilizada nas etapas posteriores do estudo, voltadas à investigação da dinâmica de disseminação científica por meio da mídia social.

Em síntese, a análise preliminar evidencia que mais da metade (55%) dos artigos científicos publicados nas áreas contempladas não receberam qualquer menção no X, o que levanta questões relevantes sobre a visibilidade social da produção científica, as desigualdades entre disciplinas e o papel de redes sociais específicas na circulação de conhecimento acadêmico.

4.5.2.2 Caracterização da amostra randômica de artigos

A análise por área temática FoR dos artigos que compõem a amostra permitiu identificar padrões relevantes no que diz respeito à visibilidade social da produção científica, medida por meio do *Altmetric Attention score* (AAS) e da distribuição das menções em redes sociais e outras fontes de atenção pública. Os dados sugerem que determinadas áreas do conhecimento alcançam níveis significativamente mais elevados de engajamento em plataformas digitais, enquanto outras demonstram visibilidade bastante reduzida. Tal cenário demonstra a heterogeneidade estrutural da altmetria entre disciplinas, exigindo uma análise diferenciada por campo de atuação científica.

O Quadro 4 apresenta um panorama detalhado da presença e do engajamento das diferentes áreas do conhecimento nas redes sociais, com base nos dados altmétricos obtidos para os artigos da amostra. As colunas incluem métricas como o AAS, número de postagens na plataforma X, vídeos, citações na Wikipedia, além de menções em outras redes como Facebook, Reddit, Pinterest e LinkedIn. Também são sugeridos os totais de leitores no Mendeley e de citações em bases como Dimensions, proporcionando uma visão abrangente da disseminação e do impacto social da produção científica por divisão temática. Esses dados subsidiam análises comparativas sobre o comportamento de disseminação e o alcance social dos artigos, permitindo identificar áreas com maior atividade automatizada ou engajamento espontâneo.

De acordo com o apresentado no Quadro 4, as divisões com os maiores volumes absolutos de menções na rede X incluem *Human Society* (81.104 menções),

Information and Computing Sciences (85.012) e *Health Sciences* (60.211). A elevada frequência de menções, combinada ao número igualmente expressivo de leitores no Mendeley e de citações na Dimensions, sugere um forte engajamento social e acadêmico nessas áreas. No entanto, a análise altimétrica exige cautela quanto ao perfil das interações: o volume, por si só, não assegura autenticidade. Considerando que a rede X é reconhecida pela alta presença de *bots*, especialmente em temas de ampla visibilidade pública e política, levanta-se a hipótese de que áreas como *Human Society* e *Health Sciences*, frequentemente debatidas em contextos sociais sensíveis, podem estar mais expostas à interferência automatizada — seja na amplificação artificial de conteúdo, seja na disseminação de determinadas retóricas.

Quadro 4 - Indicadores de presença e engajamento em redes sociais por área do conhecimento

Descrição Divisão	Média	Soma Menções												Soma Citações	Soma Leitores
		AAS	X	Video	F1000	Wikipedia	Pinterest	Reddit	Facebook	LinkedIn	News	Blogs	Policy	Patent	
30 - AGRICULTURAL, VETERINARY AND FOOD SCIENCES	9,27	37.165	95	22	473	-	104	1.205	1	3.668	744	681	522	164.009	291.180
31 - BIOLOGICAL SCIENCES	15,50	42.343	112	45	1.614	-	87	1.086	-	3.206	833	138	1.323	153.779	219.561
32 - BIOMEDICAL AND CLINICAL SCIENCES	13,54	41.356	498	40	365	-	124	1.738	-	4.972	668	334	1.172	154.056	245.859
33 - BUILT ENVIRONMENT AND DESIGN	7,46	15.582	16	1	187	1	27	221	-	1.031	181	242	114	66.829	172.065
34 - CHEMICAL SCIENCES	6,32	20.518	40	11	244	-	39	327	-	1.009	285	10	897	107.878	114.499
35 - COMMERCE, MANAGEMENT, TOURISM AND SERVICES	10,12	41.694	30	1	417	-	110	719	-	3.918	713	1.088	52	258.401	602.683
36 - CREATIVE ARTS AND WRITING	10,01	26.757	99	2	538	-	32	894	1	2.258	459	107	67	41.092	79.623
37 - EARTH SCIENCES	15,56	44.293	70	-	968	-	126	778	-	6.692	1.487	561	70	182.099	230.917
38 - ECONOMICS	14,55	15.696	10	-	203	-	44	170	-	646	299	2.010	66	74.976	73.700
39 - EDUCATION	10,81	24.227	18	-	129	-	12	969	1	1.554	308	312	12	70.052	172.420
40 - ENGINEERING	5,47	44.934	70	5	710	-	111	1.040	2	3.913	675	420	2.355	412.260	512.577
41 - ENVIRONMENTAL SCIENCES	16,99	47.866	39	33	440	-	78	1.034	1	4.725	885	812	71	153.162	292.807
42 - HEALTH SCIENCES	16,72	60.211	401	19	523	3	57	3.213	1	4.647	550	565	378	118.407	269.676

43 - HISTORY, HERITAGE AND ARCHAEOLOGY	21,04	20.770	36	6	2.153	1	73	707	-	3.405	625	38	13	31.879	50.279
44 - HUMAN SOCIETY	16,88	81.104	66	3	1.571	-	194	1.526	-	6.968	1.635	1.708	16	179.176	323.524
46 - INFORMATION AND COMPUTING SCIENCES	8,06	85.012	89	23	607	-	151	889	2	2.747	817	212	2.497	301.124	472.713
47 - LANGUAGE, COMMUNICATION AND CULTURE	9,85	28.279	35	2	630	-	47	652	-	1.518	381	131	14	61.531	111.038
48 - LAW AND LEGAL STUDIES	13,78	55.113	28	2	491	-	99	584	-	2.974	1.305	843	18	67.612	129.829
49 - MATHEMATICAL SCIENCES	3,22	11.562	9	1	170	-	16	139	-	215	122	67	160	75.122	46.957
50 - PHILOSOPHY AND RELIGIOUS STUDIES	8,96	29.870	28	1	761	-	64	738	-	1.273	414	70	9	41.973	76.904
51 - PHYSICAL SCIENCES	9,43	18.878	67	2	1.224	-	93	653	-	3.688	981	26	731	187.992	143.651
52 - PSYCHOLOGY	18,79	9.110	19	1	130	-	25	285	-	847	219	40	21	26.458	49.359
Total Geral	11,06	802.340	1.875	220	14.548	5	1.713	19.567	9	65.874	14.586	10.415	10.578	2.929.867	4.681.821

Fonte: dados de pesquisa (2025)

Por outro lado, divisões que apresentam escore médio elevado de atenção altmétrica — como *History, Heritage and Archaeology* (21,04), *Psychology* (18,79) e *Environmental Sciences* (16,99) — nem sempre são aquelas que concentram o maior volume bruto de menções. Esse dado sugere que a qualidade das menções (avaliada com base na origem, no engajamento e na influência do autor) pode ser mais relevante que a quantidade. Tal padrão está alinhado à lógica do algoritmo Altmetric, que atribui maior peso a postagens provenientes de autores influentes. Nessas divisões, a hipótese de atuação de *bots* parece menos evidente, uma vez que interações automatizadas dificilmente alcançam altos níveis de engajamento qualificado. Ainda assim, essa inferência deve ser feita com cautela, pois *bots* sofisticados são capazes de simular comportamentos humanos e interações convincentes.

A presença em múltiplas plataformas também constitui um indicador importante. Áreas como *Engineering, Biomedical and Clinical Sciences* e *Information and Computing Sciences* demonstram ampla distribuição de menções, incluindo registros no LinkedIn, Reddit, Wikipedia e Facebook. Destaca-se, nesse contexto, o elevado número de patentes associadas a *Information and Computing Sciences*

(2.497) e *Engineering* (2.355), o que sugere uma intersecção direta entre ciência e inovação tecnológica.

A análise das menções em fontes tradicionais de mídia, como *News* e *Blogs*, reforça a visibilidade pública de determinadas áreas. *Human Society*, *Environmental Sciences* e *Health Sciences* lideram em menções jornalísticas (6.968, 4.725 e 4.647, respectivamente). Embora tal volume possa refletir interesse legítimo, também é possível que derive de estratégias automatizadas de disseminação, sobretudo em áreas com forte apelo político ou emergências de interesse coletivo.

Por sua vez, divisões com baixo escore médio de AAS, como *Mathematical Sciences* (3,21) e *Engineering* (5,47), mesmo com altos índices de leitores no Mendeley e de citações acadêmicas, revelam uma desconexão entre impacto social e impacto científico. Essa lacuna pode ser explicada por perfis de comunicação menos voltados à mídia social, bem como por menor suscetibilidade dessas áreas à automação de interações. Tais características limitam sua presença em espaços de visibilidade pública, o que sugere que a altmetria não é igualmente eficaz como métrica transversal entre diferentes campos disciplinares.

No tocante à presença na Wikipedia, observam-se indícios de atenção orgânica em áreas como *History*, *Heritage and Archaeology* (2.153 menções) e *Human Society* (1.571). A Wikipedia, embora passível de edição automatizada, é geralmente menos suscetível a *bots* simples, dado seu modelo de moderação comunitária. Assim, a proeminência dessas áreas na plataforma pode refletir uma circulação mais espontânea e legitimada dos conteúdos científicos.

Com base nos dados analisados, é possível propor hipóteses exploratórias que podem orientar as investigações realizadas neste trabalho. Primeiramente, áreas com alta presença em redes sociais suscetíveis à atuação de *bots* — como X e Facebook —, especialmente aquelas com forte apelo público, podem estar mais expostas à interferência automatizada. Isso inclui divisões como *Human Society*, *Health Sciences*, *Environmental Sciences* e *Information and Computing Sciences*. Em segundo lugar, *Altmetric scores* medianos, combinados com altos volumes de menções, mas com baixa diversidade de fontes, podem apontar amplificação artificial da visibilidade, como observado em *Engineering*. Em contrapartida, áreas com score médio elevado e menor volume bruto de menções, como *Psychology*, tendem a apresentar interações mais orgânicas e qualificadas, menos propensas à automação.

A presença de *bots*, portanto, não deve ser considerada um fenômeno homogêneo, mas sim variável conforme o perfil temático da disciplina, seu grau de exposição pública e as características das plataformas de disseminação científica utilizadas, aspectos que devem ser considerados na construção de modelos de IA para detecção de *bots*. A análise dos dados revela, de forma mais ampla, que a visibilidade social da ciência é profundamente desigual entre as divisões disciplinares.

Embora o Altmetric represente uma ferramenta promissora para a avaliação do impacto social da pesquisa, sua interpretação exige contextualização, análise crítica e sensibilidade às especificidades de cada área. Tais elementos reforçam a importância de se adotar critérios diferenciados nas políticas públicas de avaliação científica, garantindo maior equidade e acurácia na mensuração da visibilidade da ciência na esfera social.

4.5.2.3 Caracterização dos posts coletados

A análise dos dados referentes à atuação das contas na disseminação de artigos científicos na plataforma X, obtidos a partir de uma amostra estratificada por grupos e organizada por divisões do conhecimento, revela padrões informacionais distintos que refletem tanto características disciplinares quanto dinâmicas próprias das redes sociais digitais. A amostra analisada contabilizou 758.288 *posts*, emitidos por 519.631 registros de contas, com referências a 93.529 artigos científicos distribuídos em 22 divisões. Embora o número total de contas não represente perfis únicos — já que uma mesma conta pode atuar em mais de uma divisão —, a distribuição dos dados permite inferências relevantes sobre o comportamento das instâncias envolvidas na comunicação científica em ambientes digitais. O Quadro 5 apresenta a distribuição dos dados que serão analisados.

De forma geral, a média de *posts* por artigo foi de 8, enquanto a média de *posts* por conta ficou em 1,46, prenunciando que, em termos agregados, a maioria das contas realizou no máximo uma ou duas postagens. No entanto, quando se examinam as divisões de maneira isolada, observa-se uma variação significativa nos padrões de divulgação. As áreas de *health sciences*, *psychology*, *environmental sciences* e *biological sciences* destacaram-se com as maiores médias de *posts* por artigo, situando-se entre 12 e 15 postagens por publicação científica. Essa alta taxa de

menções pode estar associada à relevância pública desses campos, à presença de comunidades engajadas e à capacidade de gerar repercussão midiática e social.

Quadro 5 - Média de menções por divisão de áreas de conhecimento, considerando contas, *posts* e artigos

DESCRÍÇÃO DIVISÃO	QTD CONTAS	QTD POSTS	QTD ARTIGOS	MÉDIA POSTS/ARTIGOS	MÉDIA POSTS/CONTAS
30 - AGRICULTURAL, VETERINARY AND FOOD SCIENCES	23.765	35.209	5.871	6,0	1,48
31 - BIOLOGICAL SCIENCES	26.982	39.874	3.298	12,1	1,48
32 - BIOMEDICAL AND CLINICAL SCIENCES	28.256	40.362	4.617	8,7	1,43
33 - BUILT ENVIRONMENT AND DESIGN	10.282	14.929	2.573	5,8	1,45
34 - CHEMICAL SCIENCES	12.241	20.054	3.297	6,1	1,64
35 - COMMERCE, MANAGEMENT, TOURISM AND SERVICES	28.451	39.464	5.839	6,8	1,39
36 - CREATIVE ARTS AND WRITING	16.767	24.794	3.592	6,9	1,48
37 - EARTH SCIENCES	28.767	42.533	5.232	8,1	1,48
38 - ECONOMICS	11.441	14.726	1.305	11,3	1,29
39 - EDUCATION	16.371	22.277	2.626	8,5	1,36
40 - ENGINEERING	28.330	43.504	11.670	3,7	1,54
41 - ENVIRONMENTAL SCIENCES	31.085	45.084	3.895	11,6	1,45
42 - HEALTH SCIENCES	41.846	58.771	3.947	14,9	1,40
43 - HISTORY, HERITAGE AND ARCHAEOLOGY	12.753	18.583	1.942	9,6	1,46
44 - HUMAN SOCIETY	54.421	74.878	6.558	11,4	1,38
46 - INFORMATION AND COMPUTING SCIENCES	53.505	80.963	8.430	9,6	1,51
47 - LANGUAGE, COMMUNICATION AND CULTURE	18.632	25.710	3.263	7,9	1,38
48 - LAW AND LEGAL STUDIES	33.394	50.881	4.467	11,4	1,52
49 - MATHEMATICAL SCIENCES	5.624	11.272	2.617	4,3	2,00
50 - PHILOSOPHY AND RELIGIOUS STUDIES	20.715	27.681	3.252	8,5	1,34
51 - PHYSICAL SCIENCES	9.234	18.173	4.578	4,0	1,97
52 - PSYCHOLOGY	6.769	8.566	660	13,0	1,27
Total Geral	519.631	758.288	93.529	8	1,46

Fonte: dados de pesquisa (2025).

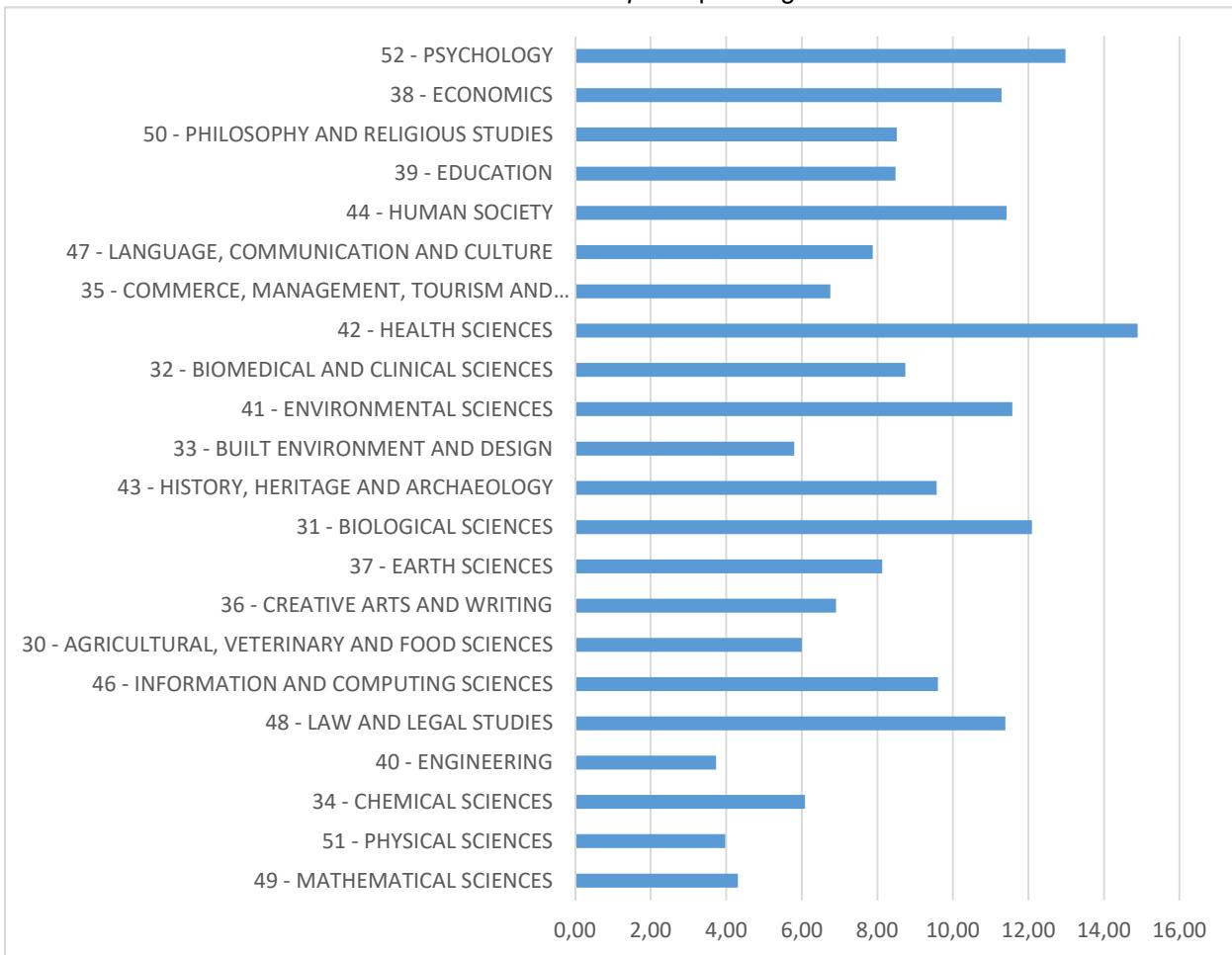
Em contraste, áreas como *engineering*, *physical sciences* e *mathematical sciences* apresentaram médias reduzidas, com aproximadamente 4 *posts* por artigo. Esse padrão pode refletir menor apelo midiático, especificidade técnica ou ainda a ausência de estratégias de divulgação amplificadas por parte das comunidades dessas disciplinas. É possível que, nessas áreas, a disseminação científica dependa

mais de agentes institucionais do que de uma rede ampla e distribuída de usuários interessados.

Em relação à média de *posts* por artigos, o Gráfico 4 apresenta a média de postagens por artigo científico, considerando o total de contas (humanas e automatizadas) que mencionaram publicações em cada uma das grandes áreas do conhecimento. Os dados revelam uma distribuição heterogênea do volume médio de disseminação por área, com destaque para *Health Sciences* (42), que lidera com uma média superior a 15 postagens por artigo, seguida por *Psychology* (52) e *Biological Sciences* (31), ambas com médias acima de 12.

Outras áreas com desempenho expressivo incluem *Law and Legal Studies* (48), *Human Society* (44) e *Information and Computing Sciences* (46), todas com valores entre 10 e 12 postagens por artigo. Em contraste, áreas como *Engineering* (40), *Chemical Sciences* (34), *Agricultural, Veterinary and Food Sciences* (30) e *Built Environment and Design* (33) registraram médias mais baixas, inferiores a 6 postagens por artigo. Esses resultados evidenciam uma discrepância nos níveis de visibilidade altmétrica entre as diferentes disciplinas, possivelmente associada ao interesse social, à frequência de publicação, à natureza do conteúdo e às estratégias de divulgação adotadas. A análise sustenta a hipótese de que a disseminação científica nas redes sociais varia significativamente por área do conhecimento (H03), contribuindo para a compreensão dos padrões de engajamento altmétrico no ecossistema digital.

Gráfico 4 - Média de posts por artigos



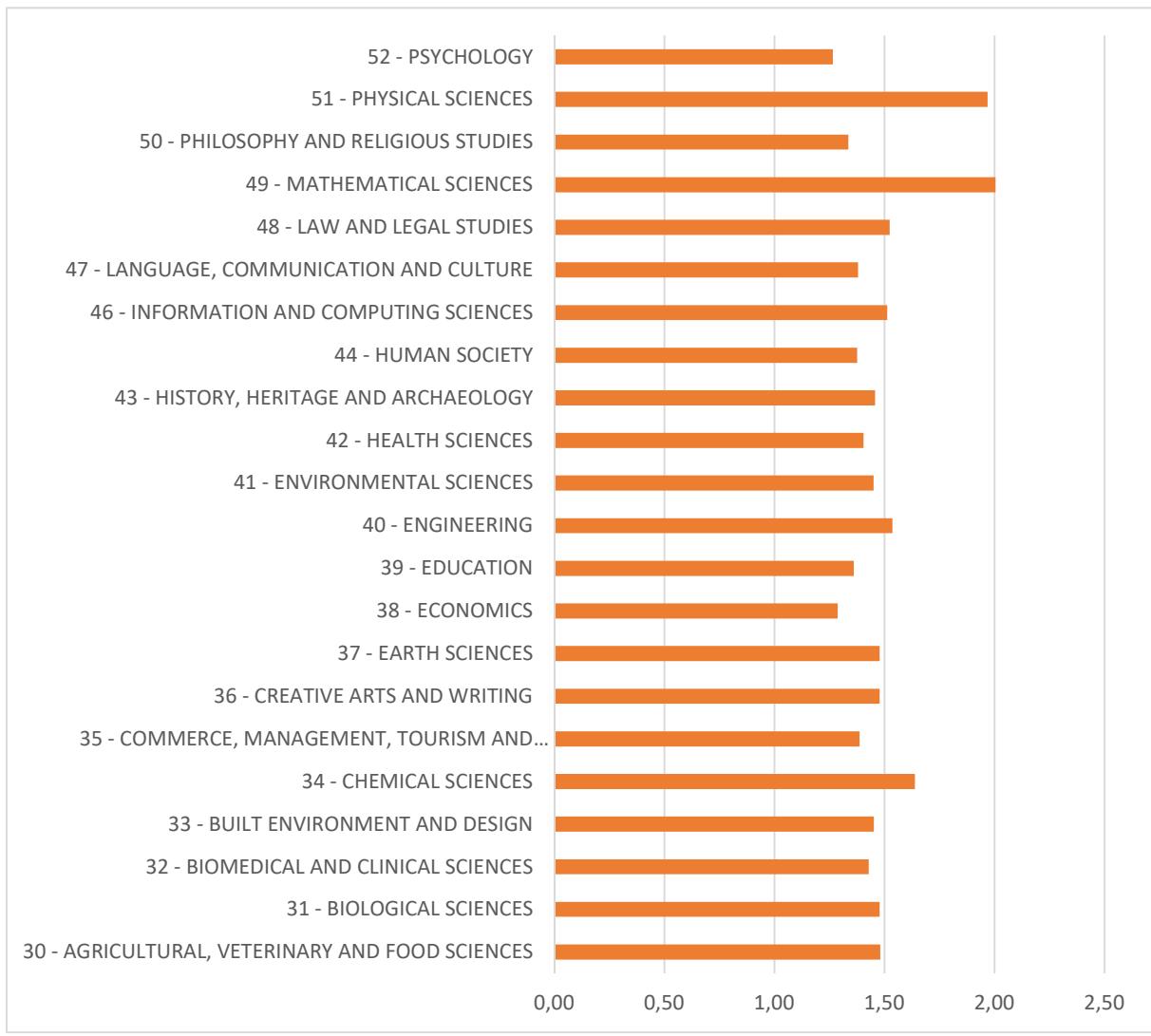
Fonte: dados de pesquisa (2025)

No que se refere à média de postagens por conta, o Gráfico 5 evidencia um aspecto relevante da dinâmica de disseminação por área do conhecimento. Observa-se que áreas como *Engineering*, *Physical Sciences* e *Mathematical sciences*, *chemical sciences*, apresentam médias significativamente superior em comparação a outras divisões, com valores próximos ou superiores a 1,6 postagens por conta. Isso manifesta que determinadas contas realizam múltiplas postagens, o que pode sinalizar a existência de perfis especializados na divulgação científica ou, alternativamente, levantar a hipótese da atuação de perfis automatizados ou semiautomatizados, como *bots*, especialmente em domínios de menor engajamento espontâneo.

A literatura recente tem apontado para a crescente presença de agentes automatizados em ambientes digitais de comunicação científica, sendo particularmente ativos em plataformas como o X, onde atuam replicando, amplificando ou reiterando informações de modo sistemático. Embora os dados analisados não

permitam afirmar com certeza a presença de *bots*, os padrões observados expressam a necessidade de investigações complementares com abordagens voltadas à detecção de comportamentos automatizados.

Gráfico 5 - Média de *posts* por contas

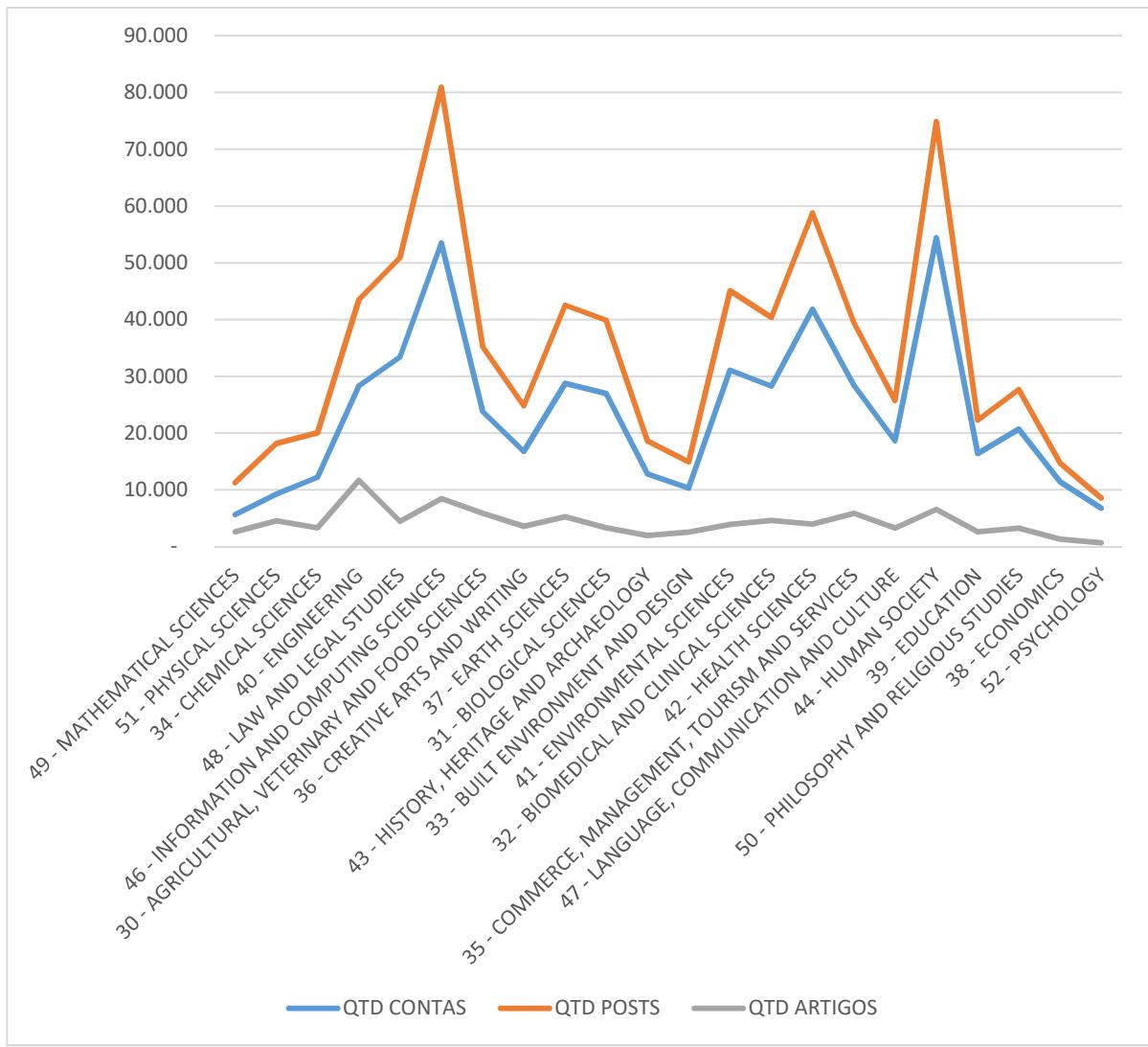


Fonte: Dados de pesquisa (2025).

Em contraposição, divisões como *Psychology*, *Education* e *Economics* apresentaram as menores médias de *posts* por conta, com valores entre 1,27 e 1,36, o que sugere um padrão de engajamento mais distribuído entre diversos perfis e possivelmente mais espontâneo. Nesses casos, a divulgação científica parece resultar de uma rede mais ampla de atores, cujas postagens ocorrem com menor frequência individual, mas maior diversidade de vozes.

O Gráfico 6 apresenta a comparação entre a quantidade de contas, *posts* e artigos distribuídos por divisão de área do conhecimento. Esse recurso visual permite observar a relação entre o volume de perfis identificados, o número de publicações realizadas e a produção científica correspondente, possibilitando identificar áreas com maior intensidade de engajamento e divulgação de artigos na plataforma X.

Gráfico 6 - Comparação entre quantidade de contas, *posts* e artigos por divisão



Fonte: Dados de pesquisa (2025).

As divisões de Sociedade *Human Society* e *Information and Computing Sciences* merecem atenção especial. Ambas combinam elevado número de artigos, grande volume de contas e altas taxas de postagens totais, o que revela uma estrutura de divulgação densa, com ampla participação de usuários e potencial engajamento institucional. No caso das Ciências da Computação, observa-se um comportamento

que pode conjugar engajamento orgânico com possíveis estratégias de disseminação sistemática, dada a familiaridade dos atores da área com práticas de automação e infraestrutura digital.

A partir desses resultados, constata-se que o ecossistema altmétrico é atravessado por lógicas disciplinares distintas e por dinâmicas próprias da circulação de informação nas redes. A interpretação desses dados exige, portanto, sensibilidade metodológica, evitando a adoção de métricas absolutas ou comparações diretas entre áreas sem a devida contextualização. As evidências apontam para a coexistência de estratégias orgânicas e estruturadas de divulgação, o que reforça a necessidade de investigações futuras sobre os agentes responsáveis por essa circulação, incluindo a identificação de perfis automatizados, a análise de temporalidade das postagens e a estrutura de redes de disseminação em torno dos artigos científicos.

4.5.2.4 Caracterização das contas que comentaram os artigos da seleção randômica

Analizando as interações sociais associadas a essa amostra, foi revelado um total de 362.395 contas únicas que realizaram ao menos uma menção aos artigos selecionados. No total, foram identificados 758.293 *posts* na plataforma X, demonstrando um volume expressivo de atividade em torno das publicações incluídas nesta etapa da investigação.

Esses dados secundários constituem a base para as análises avançadas de comportamento informacional e identificação de padrões de engajamento, incluindo, entre outros aspectos, a caracterização de contas (usuários), a frequência e distribuição temporal das postagens e a investigação sobre a presença de comportamentos automatizados (*bots*) na disseminação de literatura científica.

4.5.2.5 Caracterização do comportamento das contas

Com base nos dados da amostra estratificada, é possível realizar uma análise detalhada do comportamento das contas envolvidas na disseminação de artigos científicos na plataforma X. Esta análise fornece subsídios importantes para a compreensão da dinâmica da comunicação científica em redes sociais, revelando padrões de participação, recorrência e concentração de atividade entre os usuários.

A primeira dimensão observada refere-se à frequência de postagens por conta, conforme apresentado na Tabela 6. Os dados demonstram que 71,5% das contas realizaram apenas uma postagem, o que denota um padrão fortemente esporádico de participação. Quando se consideram as contas que realizaram até duas postagens, esse percentual se eleva para cerca de 86% do total, revelando que a esmagadora maioria das contas atua de forma pontual no compartilhamento de artigos científicos. Apenas 4,58% das contas realizaram mais de seis postagens, o que caracteriza um subconjunto restrito, mas potencialmente influente, de usuários mais ativos.

Tabela 6 - Quantidade de *posts* publicados por contas

QTD POSTS	QTDCONTAS	PERCENTUAL
1	259.213	71,53%
2	51.019	14,08%
3	20.502	5,66%
4	9.578	2,64%
5	5.492	1,52%
+ DE 6	16.591	4,58%

Fonte: Dados de pesquisa (2025).

Esse comportamento concentrado se repete quando se analisa o número de artigos distintos, mencionados por conta, como apresentado na Tabela 7. Aproximadamente 78% das contas publicaram apenas um artigo, e cerca de 90% mencionaram, no máximo, dois artigos diferentes. Em contraposição, apenas 10% das contas publicaram entre três e dois mil artigos, configurando um padrão de distribuição fortemente assimétrico. Dentro deste grupo, destaca-se um subconjunto ainda mais restrito: somente 16 contas publicaram sobre mais de 300 artigos, sendo responsáveis por um volume altamente expressivo de disseminação. Tal concentração de atividade em poucas contas levanta hipóteses relevantes sobre o perfil e os objetivos desses usuários. É possível que esse grupo inclua perfis institucionais, iniciativas de curadoria científica, ou ainda agentes automatizados, programados para amplificar conteúdos acadêmicos em larga escala.

Tabela 7 - Número de artigos publicados por contas

NÚMERO DE ARTIGOS	QTD CONTAS	PERCENTUAL
1	284.322	78,46%
2	41.413	11,43%
3	14.533	4,01%
4	6.818	1,88%
5	3.794	1,05%
6	2.436	0,67%
7	1.667	0,46%
8	1.202	0,33%
9	830	0,23%
10	661	0,18%
11-100	3.899	1,08%
101 a 500	118	0,03%
501 a 1789	6	0,00%

Fonte: Dados de pesquisa (2025).

A estrutura de distribuição observada tanto no número de postagens quanto na diversidade de artigos mencionados segue um padrão de distribuição de cauda longa, típico de sistemas sociotécnicos com alto grau de concentração informacional. Em tais sistemas, poucos atores concentram grande parte da atividade, enquanto a maioria participa de forma ocasional. Esse fenômeno tem implicações relevantes para a análise altimétrica e para a interpretação dos indicadores de atenção on-line à produção científica. Em particular, a presença de contas hiperativas, ainda que minoritária, pode distorcer a percepção de impacto social de determinados artigos ou áreas do conhecimento, sobretudo se tais contas forem operadas de forma automatizada ou por redes coordenadas de disseminação.

A predominância de contas que publicam apenas uma vez e sobre um único artigo também sugere que, para grande parte dos usuários, a menção a publicações científicas é episódica e possivelmente motivada por interesses pessoais, institucionais ou contextuais pontuais. Por outro lado, a existência de contas altamente ativas levanta questões sobre as estratégias de promoção científica e sobre a organicidade das interações observadas nas redes. A literatura recente aponta que perfis automatizados podem desempenhar papel central na propagação de conteúdos acadêmicos, influenciando algoritmos de recomendação e contribuindo para a viralização seletiva de determinadas publicações (Haustein *et al.*, 2016; Ferrara, 2020).

Em síntese, a análise dos dados demonstra que a comunicação científica em redes sociais é marcada por uma forte desigualdade participativa. De um lado, a grande maioria das contas atua de forma isolada, de outro uma minoria concentra grande parte da atividade de divulgação. Essa estrutura tem implicações significativas para a interpretação das métricas altmétricas, pois sugere que o impacto on-line de um artigo pode depender menos da mobilização espontânea de muitos usuários e mais da atuação intensiva de poucos agentes especializados. Tal constatação reforça a necessidade de abordagens analíticas que considerem o perfil das contas, a frequência de postagem e a diversidade de artigos compartilhados, bem como a possibilidade de atividade automatizada, como variáveis fundamentais para uma avaliação crítica e informada do impacto social da produção científica nas redes digitais.

4.5.2.6 Caracterização da amostra randômica de contas

A terceira etapa da coleta de dados desta pesquisa contemplou a construção de uma amostra probabilística estratificada por meio da seleção randômica de contas que compuseram o conjunto de treinamento destinado ao modelo de aprendizado de máquina. O critério adotado para estratificação baseou-se nas 22 divisões de conhecimento previamente identificadas no conjunto de dados original. Essa abordagem metodológica teve como propósito fundamental preservar a representatividade estatística das diferentes disciplinas abrangidas pelo estudo, além de possibilitar uma análise mais precisa das dinâmicas de disseminação científica nas redes sociais em um contexto controlado e reduzido.

A amostra estratificada, caracterizada por um volume significativo de informações, compreendeu um total de 67.283 postagens realizadas por 13.767 contas únicas, que por sua vez comentaram sobre 28.727 artigos científicos distintos. Essa diversidade numérica sugere a amplitude e a complexidade da interação entre usuários e produção científica divulgada nas redes sociais, proporcionando uma base robusta para a identificação de padrões comportamentais, inclusive na eventual atuação automatizada de *bots*.

Analizando o comportamento das contas selecionadas, apresentado na tabela 9, constatou-se que a maioria delas exibe uma participação limitada, concentrando suas atividades em um número reduzido de artigos científicos. Como pode ser

observado na Tabela 9, especificamente 59,68% das contas comentaram em apenas um artigo, enquanto 15,12% abordaram dois artigos distintos, e 7,28% trataram de três artigos. Observa-se ainda que aproximadamente 12% das contas tiveram participação mais intensa, interagindo com quatro a dez artigos diferentes. Por outro lado, é importante salientar que existe um pequeno grupo de contas com comportamento incomum, com cerca de 5,40% comentando entre 11 e 100 artigos e uma fração ainda menor, de apenas 0,33%, discutindo entre 101 e 1.789 artigos. Esta última categoria, embora pequena numericamente, pode exprimir comportamentos potencialmente automatizados, que frequentemente são característicos de contas identificadas como *bots*.

Observou-se que uma parcela expressiva do público – provavelmente composto por leitores ou pesquisadores individuais – manifesta-se em torno de ocorrências pontuais de interesse, ao passo que o estreitíssimo contingente de contas com atividade extensa pode refletir tanto perfis altamente especializados quanto *bots* programados para disseminar múltiplos conteúdos simultaneamente.

Tabela 8 - Quantidade de artigos por contas

QTD DOIS	QTD CONTAS	PERCENTUAL
1	8.216	59,68%
2	2.082	15,12%
3	1.002	7,28%
4 A 10	1.667	12,11%
11 a 100	743	5,40%
101 a 1789	45	0,33%

Fonte: Dados de pesquisa (2025).

Considerando especificamente o volume de postagens produzidas pelas contas, nota-se também uma concentração significativa das atividades em poucos *posts*, como demonstrado na Tabela 9. A maior parte das contas (54,67%) realizou apenas uma postagem, seguida por aquelas que produziram exatamente duas postagens, representando 15,93% do total. Outro grupo relevante, correspondente a 22,30% das contas, mostrou uma frequência intermediária, realizando entre três e dez postagens. Contudo, há novamente um subgrupo minoritário que exibe alta frequência de postagem, com 6,60% das contas realizando entre 11 e 100 postagens e apenas 0,51% ultrapassando 100 postagens, chegando até a 1.977 publicações. Esses casos

extremos de alta frequência, ainda que escassos, reforçam a hipótese da presença de *bots*, cujos padrões de comportamento costumam envolver repetidas interações em curtos períodos. Esses *bots*, ao multiplicarem repostagens e menções, têm potencial de inflar artificialmente métricas de atenção, comprometendo a interpretação da influência genuína de um trabalho científico.

Tabela 9 - Quantidade de *post* por contas

QTD POSTS	QTD CONTAS	PERCENTUAL
1	7.527	54,67%
2	2.193	15,93%
DE 3 A 10	3.070	22,30%
DE 11 A 100	907	6,60%
101 A 1977	70	0,51%

Fonte: Dados de pesquisa (2025).

Outro aspecto importante avaliado nesta etapa da análise refere-se à abrangência disciplinar das interações feitas pelas contas. De acordo com os valores organizados na Tabela 10, observou-se uma tendência clara de especialização temática entre as contas da amostra, com 63,30% comentando artigos relacionados a apenas uma divisão do conhecimento. Esse dado sugere uma segmentação forte na interação, em que a maioria das contas se restringe a interesses temáticos específicos. Essa concentração de temas por conta acusa que a maioria dos usuários mantém interesses concentrados, coerentes com o perfil de especialistas ou entusiastas humanos. Em contrapartida, 17,47% das contas apresentam interações em duas divisões distintas, enquanto 7,40% interagem com artigos de três divisões. Além disso, uma parcela reduzida das contas (1,45%) apresentou comportamento interdisciplinar acentuado, discutindo artigos de mais de nove divisões do conhecimento. Embora não se possa concluir diretamente, esse comportamento altamente diversificado também pode expressar a ação de *bots* projetados para disseminação ampla e generalizada de conteúdo científico, independentemente da especificidade temática.

Tabela 10 - Quantidade de divisões por contas

QTD DE DIVISÕES	QTD CONTAS	PERCENTUAL
1	8.715	63,30%
2	2.405	17,47%
3	1.019	7,40%
4	549	3,99%
5	340	2,47%
6	214	1,55%
7	128	0,93%
8	117	0,85%
9	80	0,58%
DE 10 A 22	200	1,45%

Fonte: Dados de pesquisa (2025).

Do ponto de vista do treinamento de modelos de *machine learning*, reconhecer esse perfil heterogêneo é fundamental. A preponderância de contas de baixa atividade fornece exemplos típicos de comportamento humano, úteis para calibrar classificadores. Entretanto, o conjunto de perfis de atividade intensa – tanto em termos de número de artigos quanto de postagens e amplitude disciplinar – deve ser cuidadosamente inspecionado. Caso sejam identificados como *bots*, a inclusão desses registros no *dataset* de treinamento pode enviesar o modelo, levando-o a interpretar erroneamente padrões automatizados como características de contas humanas.

No contexto do tópico 4.6, cinco hipóteses da pesquisa foram diretamente avaliadas, relacionadas aos padrões de comportamento e de engajamento de contas humanas e *bots*. A análise realizada revelou distinções significativas entre os dois grupos, especialmente no que diz respeito ao volume de postagens, à abrangência temática e à atuação por área do conhecimento. Constatou-se que, embora representem uma minoria numérica, as contas automatizadas foram responsáveis por um número desproporcionalmente elevado de postagens, o que corrobora a hipótese H02 – *Bots* postam mais que humanos, mesmo sendo minoria. Esse comportamento intensivo é indício de estratégias automatizadas voltadas à amplificação de conteúdos científicos, reforçando a necessidade de mecanismos que filtrem interações não autênticas na construção dos indicadores altmétricos.

Além disso, os dados evidenciaram que *bots* atuam de forma mais difusa entre áreas e subáreas científicas, com maior presença em múltiplas divisões temáticas,

enquanto os perfis humanos tendem a concentrar suas interações em domínios específicos de interesse. Tal achado confirma a hipótese H04 – *Bots* atuam em mais divisões temáticas do que humanos. A hipótese H03 – A atuação de *bots* varia por área do conhecimento também encontra respaldo nesta análise, uma vez que foi observada variação no padrão de atuação dos *bots* conforme a área do conhecimento, com predominância de perfis automatizados em grandes áreas como Ciências da Saúde e Ciências Biológicas, o que pode refletir estratégias específicas de disseminação científica nesses campos.

Por fim, a análise linguística e de engajamento revelou que os *bots* apresentam padrões expressivos e distintos em relação aos humanos, utilizando com maior frequência estratégias como repetição textual, polarização emocional e marcações recorrentes. Esses padrões não apenas diferenciam os comportamentos automatizados, como também fortalecem a eficácia dos classificadores ao capturar nuances de atuação não humana, reforçando a importância de métricas linguísticas e comportamentais na construção de modelos de detecção.

4.6 HIPÓTESES DE PESQUISA: STATUS DE VERIFICAÇÃO, EVIDÊNCIAS E TIPO DE ANÁLISE

As hipóteses formuladas orientaram o delineamento metodológico, a construção da base de dados e a interpretação dos resultados obtidos. Cada uma delas está associada a um ou mais objetivos específicos da pesquisa e foi verificada por meio de abordagens distintas, conforme o tipo de dado, a técnica utilizada e a natureza da inferência. Algumas hipóteses foram testadas por meio de métricas de desempenho quantitativas (como F1-score, AUC e Kappa), outras foram confirmadas a partir de análises empíricas baseadas nos dados obtidos, enquanto algumas foram examinadas qualitativamente, com base em interpretações geradas por técnicas de explicabilidade como o SHAP.

O Quadro 6 apresenta uma síntese das hipóteses de pesquisa formuladas. Para cada hipótese, são externados os objetivos específicos relacionados, o tipo de análise adotada (empírica, estatística, qualitativa, procedural ou exploratória), a fonte de evidência correspondente na tese (seções ou tabelas) e o status de confirmação com base nos resultados obtidos. Essa organização oferece uma visão integrada do percurso analítico da pesquisa, permitindo acompanhar a articulação

entre hipóteses, métodos e resultados, e assegurando a rastreabilidade e a coerência entre os elementos constitutivos do trabalho científico, bem como fornecendo transparência ao processo de investigação. A consolidação dessas informações em formato tabular contribui para a rastreabilidade dos resultados e reforça a coerência entre os objetivos da pesquisa, os métodos empregados e as evidências produzidas.

Quadro 6 - Verificação das Hipóteses de Pesquisa: Fontes e Estratégias Analíticas

Hipótese	Obj.	Tipo de Análise	Fonte/ Evidência na Tese	Status
H01 – XGBoost tem desempenho superior aos demais modelos	B	Estatística	Seção. 4.3 – Tabela 3 (F1, AUC, Kappa)	Comprovada
H02 – <i>Bots</i> postam mais que humanos, mesmo sendo minoria	E	Estatística descritiva	Seção. 4.5 – Tabela 5 e discussão	Comprovada
H03 – A atuação de <i>bots</i> varia por área do conhecimento	E	Estatística e segmentação	Seção. 4.5.1 – Gráficos 5 e 6; Tabela 4	Comprovada
H04 – <i>Bots</i> atuam em mais divisões temáticas do que humanos	E	Descritiva / Exploratório	Seção. 4.5.1.2 – Tabela 9	Parcialmente confirmada

Fonte: Dados da Pesquisa (2025).

As hipóteses formuladas nesta pesquisa foram verificadas por meio de estratégias variadas de análise, sempre alinhadas aos objetivos específicos. A seguir, descreve-se de forma integrada como cada hipótese foi avaliada com base nos dados obtidos e nas técnicas aplicadas.

A hipótese H01 previu que o algoritmo XGBoost teria desempenho superior aos demais testados. Essa verificação foi realizada por meio da aplicação de um pipeline unificado, utilizando os algoritmos KNN, Árvore de Decisão, Floresta Randômica e XGBoost, sendo este último o que apresentou melhores métricas preditivas.

Considerando o impacto dos *bots* nos indicadores altmétricos, a hipótese H02 sustentava que contas automatizadas publicariam mais frequentemente do que contas humanas, mesmo sendo numericamente minoritárias. Essa hipótese foi verificada por meio da análise da distribuição de postagens, ressaltando o comportamento dinâmico das contas automatizadas. A hipótese H03, por sua vez, afirmava que a atuação de *bots* variava conforme a área do conhecimento das publicações compartilhadas. Essa proposição foi confirmada com base na análise temática dos *posts*, que demonstrou

uma presença mais marcante de *bots* em determinadas disciplinas, como ciências exatas e computação.

Por fim, a hipótese H04 previa que *bots* atuariam em mais divisões temáticas do que contas humanas, sugerindo uma estratégia de disseminação automatizada mais ampla e inespecífica. Essa hipótese foi parcialmente confirmada, sendo observada uma maior dispersão temática entre os *bots*. De modo complementar, os dados analisados indicam padrões de atuação mais focados e com menor frequência entre os usuários humanos.

5 CONCLUSÕES

A presente pesquisa teve como objetivo central investigar e desenvolver métodos de detecção de *bots* na plataforma X com foco na disseminação de publicações científicas, contribuindo para o aprimoramento das métricas altmétricas. Para isso, foram aplicadas abordagens de aprendizado de máquina, estratégias de curadoria de dados e análises quantitativas e comportamentais, de modo a oferecer uma visão abrangente e aprofundada do fenômeno da automação na comunicação científica em redes sociais.

Os resultados obtidos revelam que modelos baseados em algoritmos de árvore, especialmente a Floresta Randômica e o XGBoost, demonstraram melhor desempenho na tarefa de classificação automática de contas, superando significativamente métodos como o KNN. O modelo final treinado com XGBoost atingiu alta acurácia (94,29%), elevada AUC (0,9392) e desempenho satisfatório na identificação de contas automatizadas, mesmo diante do desafio de desbalanceamento entre classes. A calibragem do *threshold* para maximização do F1-score da classe minoritária (*bots*) foi fundamental para aprimorar a sensibilidade do modelo.

Outro achado relevante da pesquisa foi a constatação da inadequação de bases genéricas de detecção de *bots* para o contexto específico da altmetria. A construção de uma base de dados rotulada, com perfis classificados manual e automaticamente a partir de fontes especializadas, mostrou-se essencial para o treinamento de modelos mais representativos do ecossistema científico digital. Essa base, composta por mais de 13 mil contas rotuladas, serviu de alicerce para o desenvolvimento de um classificador robusto, alinhado às especificidades do campo da altmetria.

A análise quantitativa revelou uma desproporcionalidade expressiva entre a quantidade de contas de *bots* (4%) e seu volume de atividade (22% das postagens), indicando que contas automatizadas, apesar de minoritárias, exercem impacto significativo sobre a visibilidade de conteúdos científicos. A média de postagens por conta de *bot* (12,6) foi substancialmente superior à média observada entre contas humanas (1,7), reforçando a hipótese de que *bots* operam com estratégias de amplificação automatizada, com potenciais efeitos distorcidos sobre os indicadores altmétricos.

Adicionalmente, verificou-se que a influência dos *bots* não é homogênea entre as áreas do conhecimento. Divisões como Mathematical Sciences, Information and Computing Sciences e Physical Sciences apresentaram proporções elevadas tanto em número de contas automatizadas quanto em volume de postagens geradas por *bots*. Em contrapartida, áreas como *Law and Legal Studies*, *Education* e *Health Sciences* exibiram baixa incidência de *bots* e maior predominância de perfis humanos. A análise estatística demonstrou alta variabilidade entre as divisões, com coeficiente de variação de 48% na proporção de postagens por *bots*, o que aponta para padrões disciplinares distintos quanto à adoção de estratégias automatizadas de divulgação.

Essas descobertas ressaltam a necessidade de abordagens específicas e monitoramento contínuo para a identificação de contas automatizadas no contexto científico. A presença significativa de *bots* em determinadas áreas pode comprometer a interpretação dos indicadores altmétricos, distorcendo avaliações de impacto e engajamento. A metodologia desenvolvida nesta pesquisa, baseada na construção de uma base rotulada, na aplicação de modelos avançados de classificação e na análise estratificada por divisão científica, oferece subsídios técnicos e conceituais para o aprimoramento das práticas de análise altmétrica.

A análise dos resultados permitiu também verificar a superioridade do modelo XGBoost em relação aos demais algoritmos testados, bem como a relevância do uso de SHAP para a explicabilidade do modelo. As evidências indicaram ainda que *bots*, apesar de numericamente minoritários, apresentam atividade significativamente superior à dos humanos e operam com maior diversidade temática (H04), ao passo que contas humanas demonstram menor volume de postagens e foco mais restrito por área do conhecimento.

Os resultados obtidos nesta tese evidenciam que as distorções causadas pela atuação de *bots* nos scores altmétricos exigem uma revisão crítica da interpretação do *Altmetric Attention Score* (AAS), sobretudo em áreas mais suscetíveis à automação, como *Information and Computing Sciences*, *Human Society*, *Engineering* e *Environmental Sciences*. Nas análises realizadas, observou-se que essas divisões concentram um número expressivo de contas automatizadas, alcançando 20,1% em *Mathematical Sciences* e mais de cinco mil perfis em *Information and Computing Sciences*, o que resulta em um volume inflado de menções e, consequentemente, em scores potencialmente superestimados. Esse grau de automação é particularmente relevante em plataformas como o X e o Facebook, nas quais mecanismos de

postagens automatizada amplificam artificialmente a visibilidade de determinados artigos ou periódicos (Maricato *et al.*, 2025).

Além do efeito quantitativo, constatou-se que áreas com *scores* medianos, mas alta concentração de menções repetitivas e baixa diversidade de fontes, podem refletir formas de engajamento não orgânico, distintas daquelas em que o alto desempenho altmétrico resulta de interações qualificadas e diversificadas. Dessa forma, o AAS, quando interpretado isoladamente, tende a capturar mais a intensidade do tráfego automatizado do que a qualidade da atenção social. A presença de *bots*, portanto, não é homogênea, variando conforme o perfil temático e o grau de exposição pública de cada disciplina, aspecto também destacado por Araújo em Maricato *et al.* (2025). Assim, propõe-se que o AAS seja interpretado de modo contextualizado, levando em consideração a probabilidade de interferência automatizada e as características comunicacionais das áreas, de forma a aprimorar a acurácia das avaliações de impacto social e institucional da pesquisa em consonância com o que é apresentado em Maricato *et al.* (2025).

Essa constatação implica uma mudança de paradigma onde o AAS deixa de ser compreendido como um indicador estritamente quantitativo e passar a ser concebido como um índice relacional e contextual, cuja validade depende da detecção e ponderação da influência automatizada de maneira coerente com a crítica de Rodrigo Costas em Maricato *et al.*, (2025) ao “excesso de mimetização da lógica de citação” e à necessidade de reformulação conceitual da altmetria como um espaço de *acoplamentos heterogêneos*.

Ferramentas como o SciBotScan, desenvolvida nesta tese, oferecem subsídios concretos para essa reinterpretação, ao integrar variáveis comportamentais tais como frequência de postagens, repetição textual, polaridade de sentimento, legibilidade e outras, capazes de gerar indicadores de organicidade. A combinação desses parâmetros com o AAS tradicional pode originar métricas ponderadas por autenticidade, distinguindo o engajamento genuíno da amplificação artificial e permitindo que instituições valorizem práticas legítimas de divulgação científica, sem confundi-las com estratégias automatizadas de autopromoção.

As implicações desses resultados estendem-se ao campo mais amplo da avaliação científica. A presença significativa de contas automatizadas em determinadas áreas, especialmente nas ciências computacionais, engenharias e ciências sociais digitais, pode comprometer a comparação interinstitucional e reduzir

a fidedignidade dos indicadores altmétricos, essa preocupação também é levantada por Araújo em Maricato *et al.*, (2025), ao afirmar que o uso apressado das métricas pode gerar “pesquisadores caçadores de likes” e comprometer a integridade avaliativa. Com ferramentas como o SciBotScan, é possível oferecer mecanismos de monitoramento e ponderação da automação, contribuindo para mitigar esse risco e fortalecer a transparência dos processos avaliativos. Sua aplicação pode subsidiar universidades, agências de fomento e periódicos científicos na formulação de políticas mais responsáveis de uso da altmetria, alinhadas às diretrizes internacionais de integridade e ciência aberta.

Por fim, os resultados desta tese indicam que as *features* utilizadas no modelo de detecção de *bots*, abrangendo aspectos temporais, linguísticos, relacionais e semânticos, constituem um caminho promissor para o aprimoramento dos indicadores altmétricos. Ao capturar padrões de cadência, legibilidade, repetição e vínculo com o objeto científico, essas variáveis permitem distinguir interações orgânicas de atividades automatizadas, revelando dimensões qualitativas do impacto informacional que escapam às métricas baseadas apenas em volume — convergindo com a defesa de Rodrigo Costas e Ronaldo Araújo em Maricato *et al.*, (2025) de que a altmetria deve priorizar abordagens qualitativas e interpretativas em detrimento de contagens brutais.

A incorporação dessas *features* em sistemas de mensuração altmétrica abre espaço para o desenvolvimento de indicadores mais sensíveis à organicidade das interações e ao valor comunicativo das trocas informacionais.

Embora o modelo apresentado represente um avanço na curadoria e filtragem de dados altmétricos, sua plena integração às práticas de avaliação científica requer estudos futuros. Entre as possibilidades de continuidade, destacam-se a criação de métricas derivadas que ponderem a probabilidade de automação e a qualidade do engajamento, bem como a adaptação das variáveis aos novos padrões de produção de conteúdo mediados por inteligência artificial. Tais aprimoramentos poderão contribuir para a formulação de indicadores mais transparentes, éticos e robustos, capazes de reconhecer a legitimidade dos *bots* enquanto instrumentos de divulgação científica, mas sem permitir que sua atuação distorça a compreensão do impacto real da comunicação acadêmica nas redes digitais.

Em termos epistemológicos, esta tese propõe um deslocamento na compreensão dos indicadores altmétricos, especialmente do Altmetric Attention Score

(AAS), que devem ser interpretados não mais como medidas puramente quantitativas, baseadas em contagens brutas, mas como construções informacionais que requerem leitura qualificada apartir de uma curadoria e contextualizada. A atenção científica nas redes sociais não é homogênea: resulta da interação de múltiplos agentes (humanos, institucionais e automatizados), e de diferentes formas de engajamento, que incluem desde a replicação automática até o comentário crítico e a mediação interpretativa. Sob essa perspectiva, o score deixa de ser um “fato dado” e passa a ser compreendido como um “fato construído”, cuja validade depende da explicitação dos critérios de coleta, ponderação e transparência quanto à origem das interações que o compõem.

Esse reposicionamento não implica rejeitar a automação, mas reconhecer sua presença e seus efeitos. Admite-se que *bots* podem exercer um papel legítimo na difusão científica, especialmente quando operam como extensões comunicacionais de instituições, periódicos ou repositórios. Contudo, sua atuação deve ser conceitualmente demarcada, de modo que sua influência seja visível, proporcional e justificada. O propósito, portanto, não é excluir os agentes automatizados, mas integrá-los de forma ética e interpretável, evitando que a performatividade dos números, aquilo que o indicador “faz” no mundo, se sobreponha ao seu significado, aquilo que ele efetivamente “representa” sobre o mundo. Assim, a organicidade do engajamento, entendida como expressão de mediação humana, ancoragem documental e diversidade de vozes, constitui o princípio regulador da interpretação do AAS e dos indicadores altmétricos em geral.

Conclui-se, portanto, que o uso de inteligência artificial aplicada à detecção de *bots* na altmetria é não apenas viável, mas necessário, especialmente em um cenário informacional crescente e mais automatizado. A pesquisa aqui apresentada contribui para a transparência na mensuração da influência científica em redes sociais e aponta caminhos para a mitigação de vieses algorítmicos e para a promoção de métricas mais confiáveis e representativas da produção acadêmica.

REFERÊNCIAS

- ADEL ALIPOUR, Sanaz; ORJI, Rita; ZINCIR-HEYWOOD, Nur. Security of Social Networks: Lessons Learned on Twitter Bot Analysis in the Literature. *In: ARES 2022: THE 17TH INTERNATIONAL CONFERENCE ON AVAILABILITY, RELIABILITY AND SECURITY*, 23 ago. 2022. **Proceedings of the 17th International Conference on Availability, Reliability and Security** [...]. Vienna Austria: ACM, 23 ago. 2022. p. 1–9. DOI 10.1145/3538969.3544450. Disponível em: <https://dl.acm.org/doi/10.1145/3538969.3544450>. Acesso em: 19 jun. 2025.
- ADIE, E. The grey literature from an altmetrics perspective—opportunity and challenges. **Research Trends**, [s. l.], v. 37, n. 37, p. 23–25, 2014. .
- AGARWAL, A.; XIE, B.; VOVSHA, I.; RAMBOW, O.; PASSONNEAU, R. Sentiment analysis of Twitter data. 2011. **Workshop on Languages in Social Media** [...]. [S. l.]: ACL, 2011. p. 30–38.
- AGUILERA, Ana; QUINTEROS, Pamela; DONGO, Irvin; CARDINALE, Judith. CrediBot: Applying Bot Detection for Credibility Analysis on Twitter. **IEEE Access**, [s. l.], v. 11, p. 108365–108385, 2023. DOI 10.1109/access.2023.3320687. Disponível em: <https://ieeexplore.ieee.org/document/10267924/>. Acesso em: 30 abr. 2025.
- ALANI, Mohammed M. Big data in cybersecurity: a survey of applications and future trends. **Journal of Reliable Intelligent Environments**, [s. l.], v. 7, n. 2, p. 85–114, jun. 2021. DOI 10.1007/s40860-020-00120-3. Disponível em: <https://link.springer.com/10.1007/s40860-020-00120-3>. Acesso em: 2 maio 2025.
- ALARIFI, Abdulrahman; ALSALEH, Mansour; AL-SALMAN, AbdulMalik. Twitter turing test: Identifying social machines. **Information Sciences**, [s. l.], v. 372, p. 332–346, dez. 2016. DOI 10.1016/j.ins.2016.08.036. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0020025516306077>. Acesso em: 21 ago. 2025.
- ALHOORI, Hamed; FURUTA, Richard; TABET, Myrna; SAMAKA, Mohammed; FOX, Edward A. Altmetrics for country-level research assessment. 8839., 2014. **The Emergence of Digital Libraries – Research and Practices. ICADL 2014. Lecture Notes in Computer Science** [...] [...]. Switzerland: Springer, Cham, 2014. v. 8839, p. 59–64. DOI https://doi.org/10.1007/978-3-319-12823-8_7.
- ALJOHANI, Naif Radi; FAYOUMI, Ayman; HASSAN, Saeed-Ul. Bot prediction on social networks of Twitter in altmetrics using deep graph convolutional networks. **Soft Computing**, [s. l.], v. 24, n. 15, p. 11109–11120, 1 ago. 2020. DOI 10.1007/s00500-020-04689-y. Disponível em: <https://doi.org/10.1007/s00500-020-04689-y>. Acesso em: 30 abr. 2025.

ALLEM, Jon Patrick; ESCOBEDO, Patricia; DHARMAPURI, Likhit. Cannabis surveillance with twitter data: Emerging topics and social bots. **American Journal of Public Health**, [s. l.], v. 110, n. 3, p. 357–362, 2020. DOI <https://doi.org/10.2105/AJPH.2019.305461>.

ALPERIN, Juan Pablo. Ask not what altmetrics can do for you, but what altmetrics can do for developing countries. **Bulletin of the American Society for Information Science and Technology**, [s. l.], v. 39, n. 4, p. 18–21, 2013. DOI <https://doi.org/10.1002/bult.2013.1720390407>.

ALPERIN, Juan Pablo; GOMEZ, Charles J.; HAUSTEIN, Stefanie. Identifying diffusion patterns of research articles on Twitter: A case study of online engagement with open access articles. **Public Understanding of Science**, [s. l.], v. 28, n. 1, p. 2–18, 2019. <https://doi.org/10.1177/0963662518761733>.

ARAÚJO, Carlos Alberto Ávila. A pós-verdade como desafio central para a ciência da informação contemporânea. **Em Questão**, [s. l.], v. 27, n. 1, p. 13–29, 2020. <https://doi.org/10.19132/1808-5245271.13-29>.

ARAUJO, Ronaldo Ferreira; CARAN, Gustavo Miranda; SOUZA, Iara Vidal Pereira de. Orientação temática e coeficiente de correlação para análise comparativa entre altmetrics e citações. **Em Questão**, [s. l.], v. 22, n. 3, p. 184, 2016. DOI <https://doi.org/10.19132/1808-5245223.184-200>.

ARROYO-MACHADO, Wenceslao; HERRERA-VIEDMA, Enrique; TORRES-SALINAS, Daniel. The Elon Musk paradox: Quantifying the Presence and Impact of Twitter Bots on Altmetrics with Focus in Social Sciences. 2023. **27th International Conference on Science, Technology and Innovation Indicators (STI 2023)** [...]. [S. l.]: International Conference on Science, Technology and Innovation Indicators, 2023. Disponível em: <https://dapp.orvium.io/deposits/644235015db3c5af25159230/view>. Acesso em: 30 abr. 2025.

ARROYO-MACHADO, Wenceslao; HERRERA-VIEDMA, Enrique; TORRES-SALINAS, Daniel. The Elon Musk paradox : Quantifying the Presence and Impact of Twitter Bots on Altmetrics with Focus in Social Sciences. 2023. **27th International Conference on Science, Technology and Innovation Indicators (STI 2023)** [...]. Leiden, The Netherlands: [s. n.], 2023. p. 1–7.

BALESTRUCCI, Alessandro; DE NICOLA, Rocco; INVERSO, Omar; TRUBIANI, Catia. Identification of credulous users on Twitter. In: SAC '19: THE 34TH ACM/SIGAPP SYMPOSIUM ON APPLIED COMPUTING, 8 abr. 2019. **Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing** [...]. Limassol Cyprus: ACM, 8 abr. 2019. p. 2096–2103. DOI 10.1145/3297280.3297486. Disponível em: <https://dl.acm.org/doi/10.1145/3297280.3297486>. Acesso em: 30 abr. 2025.

BALESTRUCCI, Alessandro; INVERSO, Omar; DE NICOLA, Rocco; TRUBIANI, Catia. Identification of credulous users on twitter. **Proceedings of the ACM Symposium on Applied Computing**, [s. l.], v. Part F1477, p. 2096–2103, 2019. DOI <https://doi.org/10.1145/3297280.3297486>.

BARBON, Sylvio; CAMPOS, Gabriel F. C.; TAVARES, Gabriel M.; IGAWA, Rodrigo A.; JR, Mario L. Proença; GUIDO, Rodrigo Capobianco. Detection of Human, Legitimate Bot, and Malicious Bot in Online Social Networks Based on Wavelets. **ACM Transactions on Multimedia Computing, Communications, and Applications**, [s. l.], v. 14, n. 1s, p. 1–17, 31 mar. 2018. DOI 10.1145/3183506. Disponível em: <https://dl.acm.org/doi/10.1145/3183506>. Acesso em: 21 ago. 2025.

BARCELOS, J.; MARICATO; J. M. Academic production mentions in social media: altmetric study of visibility and public engagement with Scientometric papers. 2017. **XVIII. ENANCIB - National Meeting of Research in Information Science [...]**. Marília: [s. n.], 2017.

BARCELOS, Janinne; MARICATO, João de Melo. Menções sobre a produção acadêmica nas mídias sociais : estudo altmétrico de visibilidade e engajamento público com artigos da Scientometrics. [s. l.], 2017. Disponível em: <http://repositorio.unb.br/handle/10482/31538>. Acesso em: 30 abr. 2025.

BATRINCA, Bogdan; TRELEAVEN, Philip C. Social media analytics: a survey of techniques, tools and platforms. **AI and Society**, [s. l.], v. 30, n. 1, p. 89–116, 2015. DOI <https://doi.org/10.1007/s00146-014-0549-4>.

BENKLER, Y; FARIS, Robert; HAL ROBERTS BRUCE ETLING; NIKKI BOURASSA; ZUCKERMAN, Ethan. **Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election**. Cambridge, v. 7641, 2017.

BHATTACHARYYA, Ajay; KULKARNI, Adita. Machine Learning-Based Detection and Categorization of Malicious Accounts on Social Media. [S. l.: s. n.], 2024. p. 328–337. DOI https://doi.org/10.1007/978-3-031-61281-7_23.

BHOSALE, Sanmesh; DI TROIA, Fabio. **Twitter Bots' Detection with Benford's Law and Machine Learning**. [S. l.]: Springer Nature Switzerland, 2022. v. 1683 CCIS. DOI https://doi.org/10.1007/978-3-031-24049-2_3

BOLLEN, Johan; MAO, Huina; ZENG, Xiaojun. Twitter mood predicts the stock market. **Journal of Computational Science**, [s. l.], v. 2, n. 1, p. 1–8, 2011. DOI <https://doi.org/10.1016/j.jocs.2010.12.007>

BOTTA, Federico; MOAT, Helen Susannah; PREIS, Tobias. Quantifying crowd size with mobile phone and Twitter data. **Royal Society Open Science**, [s. l.], v. 2, n. 5, p. 1–6, 2015. DOI <https://doi.org/10.1098/rsos.150162>

BRINK, Henrik. **Real-World Machine Learning**. New York: Manning Publications Co. LLC, 2016.

BRONIATOWSKI, David A.; JAMISON, Amelia M.; QI, Si Hua; ALKULAIB, Lulwah; CHEN, Tao; BENTON, Adrian; QUINN, Sandra C.; DREDZE, Mark. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. **American Journal of Public Health**, [s. l.], v. 108, n. 10, p. 1378–1384, 2018. DOI <https://doi.org/10.2105/AJPH.2018.304567>

BROWN, Megan A.; GRUEN, Andrew; MALDOFF, Gabe; MESSING, Solomon; SANDERSON, Zeve; ZIMMER, Michael. Web Scraping for Research: Legal, Ethical, Institutional, and Scientific Considerations. [s. l.], 19 dez. 2024. DOI 10.48550/arXiv.2410.23432. Disponível em: <http://arxiv.org/abs/2410.23432>. Acesso em: 2 maio 2025.

CHAVOSHI, Nikan; HAMOONI, Hossein; MUEEN, Abdullah. Identifying correlated bots in Twitter. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, [s. l.], v. 10047 LNCS, p. 14–21, 2016. DOI https://doi.org/10.1007/978-3-319-47874-6_2

CHEN, Tianqi; GUESTIN, Carlos. XGBoost: A Scalable Tree Boosting System. *In: KDD '16: THE 22ND ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING*, 13 ago. 2016. **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining** [...]. San Francisco California USA: ACM, 13 ago. 2016. p. 785–794. DOI 10.1145/2939672.2939785. Disponível em: <https://dl.acm.org/doi/10.1145/2939672.2939785>. Acesso em: 30 maio 2025.

CHU, Zi; GIANVECCHIO, Steven; WANG, Haining; JAJODIA, Sushil. Detecting automation of Twitter accounts: Are you a human, bot, or cyborg? **IEEE Transactions on Dependable and Secure Computing**, [s. l.], v. 9, n. 6, p. 811–824, 2012a. DOI <https://doi.org/10.1109/TDSC.2012.75>

CHU, Zi; GIANVECCHIO, Steven; WANG, Haining; JAJODIA, Sushil. Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg? **IEEE Transactions on Dependable and Secure Computing**, [s. l.], v. 9, n. 6, p. 811–824, nov. 2012b. DOI 10.1109/TDSC.2012.75. Disponível em: <https://ieeexplore.ieee.org/document/6280553>. Acesso em: 30 abr. 2025.

CIUPERĂ, Ella Magdalena; STANCIU, Alexandru. Social networks-based alt-metrics and the unethical use of AI. *In: INTERNATIONAL CONFERENCE ON VIRTUAL LEARNING - VIRTUAL LEARNING - VIRTUAL REALITY (17TH EDITION)*, 23 dez. 2022. **Proceedings of the International Conference on Virtual Learning - VIRTUAL LEARNING - VIRTUAL REALITY (17th edition)** [...]. [S. l.]: The National Institute for Research & Development in Informatics - ICI Bucharest (ICI Publishing House), 23 dez. 2022. p. 167–178. DOI 10.58503/icvl-v17y202214. Disponível em:

<https://icvl.eu/vol-17-2022/social-networks-based-alt-metrics-and-the-unethical-use-of-ai/>. Acesso em: 30 abr. 2025.

COSTAS, Rodrigo; DE RIJCKE, Sarah; MARRES, Noortje. “Heterogeneous couplings”: Operationalizing network perspectives to study science-society interactions through social media metrics. **Journal of the Association for Information Science and Technology**, [s. l.], v. 72, n. 5, p. 595–610, 2021. DOI <https://doi.org/10.1002/asi.24427>

COSTAS, Rodrigo; ZAHEDI, Zohreh; WOUTERS, Paul. Do “altmetrics” correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. **Journal of the Association for Information Science and Technology**, [s. l.], v. 66, n. 10, p. 2003–2019, 28 out. 2015. DOI <https://doi.org/10.1002/asi.23309>

CROTTY, David. Altmetrics: Finding Meaningful Needles in the Data Haystack. **Serials Review**, [s. l.], v. 40, n. 3, p. 141–146, 3 jul. 2014. DOI <https://doi.org/10.1080/00987913.2014.947839>

DARLING, Emily; SHIFFMAN, David; CÔTÉ, Isabelle; DREW, Joshua. The role of Twitter in the life cycle of a scientific publication. **Ideas in Ecology and Evolution**, [s. l.], v. 6, 2013. DOI 10.4033/iee.2013.6.6.f. Disponível em: <https://ojs.library.queensu.ca/index.php/IEE/article/view/4625>. Acesso em: 30 abr. 2025.

DE MELO MARICATO, João; DE CASTRO MANSO, Bruno Lara. Characterization of the communities of attention interacting with scientific papers on Twitter: altmetric analysis of a Brazilian University. **Scientometrics**, [s. l.], v. 127, n. 7, p. 3815–3835, 2022. DOI <https://doi.org/10.1007/s11192-022-04442-2>

DE MORAIS, Daniel Marques Gomes; DIGIAMPIETRI, Luciano Antonio. Methods and Challenges in Social Bots Detection: A Systematic Review. **ACM International Conference Proceeding Series**, [s. l.], , p. 21–28, 2021. DOI <https://doi.org/10.1145/3466933.3466973>

DETECTION OF SOCIAL BOTS IN TWITTER NETWORK. In: DUDDU, Mahesh Chandra; DURGA BHAVANI, S. **Algorithms for Intelligent Systems**. Singapore: Springer Nature Singapore, 2023. p. 655–668. DOI 10.1007/978-981-99-1435-7_53. Disponível em: https://link.springer.com/10.1007/978-981-99-1435-7_53. Acesso em: 1 maio 2025.

DIDEGAH, Fereshteh; MEJLGAARD, Niels; SØRENSEN, Mads P. Investigating the quality of interactions and public engagement around scientific papers on Twitter. **Journal of Informetrics**, [s. l.], v. 12, n. 3, p. 960–971, 2018. DOI <https://doi.org/10.1016/j.joi.2018.08.002>

DRAHOŠOVÁ, Martina; BALCO, Peter. The analysis of advantages and disadvantages of use of social media in European Union. **Procedia Computer Science**, [s. l.], v. 109, p. 1005–1009, 2017. DOI DOI <https://doi.org/10.1016/j.procs.2017.05.446>

ELLIS, John; ELLIS, Bethany; TYLER, Kevin; REICHEL, Michael P. Recent trends in the use of social media in parasitology and the application of alternative metrics. **Current Research in Parasitology and Vector-Borne Diseases**, [s. l.], v. 1, n. January, p. 100013, 2021. DOI <https://doi.org/10.1016/j.crvbd.2021.100013>.

FAN, Rui; TALAVERA, Oleksandr; TRAN, Vu. Social media bots and stock markets. **European Financial Management**, [s. l.], v. 26, n. 3, p. 753–777, 2020. DOI <https://doi.org/10.1111/eufm.12245>.

FANG, Zhichao; COSTAS, Rodrigo; TIAN, Wencan; WANG, Xianwen; WOUTERS, Paul. How is science clicked on Twitter? Click metrics for Bitly short links to scientific publications. **Journal of the Association for Information Science and Technology**, [s. l.], v. 72, n. 7, p. 918–932, jul. 2021. DOI 10.1002/asi.24458. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.24458>. Acesso em: 30 abr. 2025.

FANG, Zhichao; COSTAS, Rodrigo; WOUTERS, Paul. User engagement with scholarly tweets of scientific papers: a large-scale and cross-disciplinary analysis. **Scientometrics**, [s. l.], v. 127, n. 8, p. 4523–4546, 2022a. DOI <https://doi.org/10.1007/s11192-022-04468-6>.

FANG, Zhichao; COSTAS, Rodrigo; WOUTERS, Paul. User engagement with scholarly tweets of scientific papers: a large-scale and cross-disciplinary analysis. **Scientometrics**, [s. l.], v. 127, n. 8, p. 4523–4546, 2022b. DOI <https://doi.org/10.1007/s11192-022-04468-6>.

FANG, Zhichao; DUDEK, Jonathan; COSTAS, Rodrigo. Facing the volatility of tweets in altmetric research. **Journal of the Association for Information Science and Technology**, [s. l.], v. 73, n. 8, p. 1192–1195, ago. 2022. DOI 10.1002/asi.24624. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.24624>. Acesso em: 2 maio 2025.

FERRARA, Emilio. Measuring Social Spam and the Effect of Bots on Information Diffusion in Social Media. [s. l.], , p. 229–255, 2018. DOI https://doi.org/10.1007/978-3-319-77332-2_13

FERRARA, Emilio. Measuring social spam and the effect of bots on information diffusion in social media. [s. l.], 27 ago. 2017. DOI 10.1007/978-3-319-77332-2_13. Disponível em: <http://arxiv.org/abs/1708.08134>. Acesso em: 30 abr. 2025.

FERRARA, Emilio; VAROL, Onur; DAVIS, Clayton; MENCZER, Filippo;; FLAMMINI, Alessandro. The rise of social bots. **Communications of the ACM**, [s. l.], v. 59, n. 7, p. 96–104, 24 jun. 2016. <https://doi.org/10.1145/2818717>.

FERRARA, Emilio; VAROL, Onur; DAVIS, Clayton; MENCZER, Filippo; FLAMMINI, Alessandro. The rise of social bots. **Communications of the ACM**, [s. l.], v. 59, n. 7, p. 96–104, 24 jun. 2016. DOI 10.1145/2818717. Disponível em: <https://dl.acm.org/doi/10.1145/2818717>. Acesso em: 30 abr. 2025.

FERRARA, Emilio; YANG, Zeyao. Quantifying the effect of sentiment on information diffusion in social media. **PeerJ Computer Science**, [s. l.], v. 2015, n. 9, p. 1–10, 2015. DOI <https://doi.org/10.7717/peerj-cs.26>.

FREITAS, Carlos; BENEVENUTO, Fabricio; GHOSH, Saptarshi; VELOSO, Adriano. Reverse Engineering Socialbot Infiltration Strategies in Twitter. 25 ago. 2015. **Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015** [...]. New York, NY, USA: ACM, 25 ago. 2015. p. 25–32. DOI <https://doi.org/10.1145/2808797.2809292>.

GALVÃO, Taís Freire; TIGUMAN, Gustavo Magno Baldin; SARKIS-ONOFRE, Rafael; PAGE, Matthew J.; MCKENZIE, Joanne E.; BOSSUYT, Patrick M.; BOUTRON, Isabelle; HOFFMANN, Tammy C.; MULROW, Cynthia D.; SHAMSEER, Larissa; TETZLAFF, Jennifer M.; AKL, Elie A.; BRENNAN, Sue E.; CHOU, Roger; GLANVILLE, Julie; GRIMSHAW, Jeremy M.; HRÓBJARTSSON, Asbjørn; LALU, Manoj M.; LI, Tianjing; LODER, Elizabeth W.; MAYO-WILSON, Evan; MCDONALD, Steve; MCGUINNESS, Luke A.; STEWART, Lesley A.; THOMAS, James; TRICCO, Andrea C.; WELCH, Vivian A.; WHITING, Penny; MOHER, David. A declaração PRISMA 2020: Diretriz atualizada para relatar revisões sistemáticas. **Epidemiologia e Servicos de Saude**, [s. l.], v. 31, n. 2, p. 1–12, 2022. DOI <https://doi.org/10.5123/S1679-49742022000200033>.

GHOSH, Rumi; SURACHAWALA, Tawan; LERMAN, Kristina. **Entropy-based Classification of “Retweeting” Activity on Twitter**. [s. l.], v. 11, 2011.

GILANI, Zafar; FARAHBAKHSH, Reza; TYSON, Gareth; CROWCROFT, Jon. A Large-scale Behavioural Analysis of Bots and Humans on Twitter. **ACM Transactions on the Web**, [s. l.], v. 13, n. 1, p. 1–23, 28 fev. 2019. DOI 10.1145/3298789. Disponível em: <https://dl.acm.org/doi/10.1145/3298789>. Acesso em: 3 maio 2025.

GILANI, Zafar; FARAHBAKHSH, Reza; TYSON, Gareth; WANG, Liang; CROWCROFT, Jon. Of bots and humans (on twitter). **Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2017**, [s. l.], , p. 349–354, 2017. DOI <https://doi.org/10.1145/3110025.3110090>

GILANI, Zafar; KOCHMAR, Ekaterina; CROWCROFT, Jon. Classification of Twitter Accounts into Automated Agents and Human Users. In: ASONAM '17: ADVANCES IN SOCIAL NETWORKS ANALYSIS AND MINING 2017, 31 jul. 2017. **Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017** [...]. Sydney Australia: ACM, 31 jul. 2017. p. 489–496. DOI 10.1145/3110025.3110091. Disponível em: <https://dl.acm.org/doi/10.1145/3110025.3110091>. Acesso em: 30 abr. 2025.

GORODNICHENKO, Yuriy; PHAM, Tho; TALAVERA, Oleksandr. Social media, sentiment and public opinions: Evidence from #Brexit and #USElection. **European Economic Review**, [s. l.], v. 136, p. 103772, 2021. DOI <https://doi.org/10.1016/j.euroecorev.2021.103772>.

GORWA, Robert; GUILBEAULT, Douglas. Unpacking the Social Media Bot: A Typology to Guide Research and Policy. **Policy and Internet**, [s. l.], v. 12, n. 2, p. 225–248, 2020. DOI <https://doi.org/10.1002/poi3.184>.

GRINBERG, Nir; JOSEPH, Kenneth; FRIEDLAND, Lisa; SWIRE-THOMPSON, Briony; LAZER, David. Political science: Fake news on Twitter during the 2016 U.S. presidential election. **Science**, [s. l.], v. 363, n. 6425, p. 374–378, 2019. DOI <https://doi.org/10.1126/science.aau2706>.

GUTIÉRREZ, David Martín. Twitter Bots Accounts [Data set]. 2020. **Kaggle**. DOI <https://doi.org/10.34740/KAGGLE/DSV/1623389>. Disponível em: <https://www.kaggle.com/davidmartngutierrez/twitter-bots-accounts/metadata>. Acesso em: 16 abr. 2022.

HASSAN, Saeed UI; BOWMAN, Timothy D.; SHABBIR, Mudassir; AKHTAR, Aqsa; IMRAN, Mubashir; ALJOHANI, Naif Radi. Influential tweeters in relation to highly cited articles in altmetric big data. **Scientometrics**, [s. l.], v. 119, n. 1, p. 481–493, 2019. DOI <https://doi.org/10.1007/s11192-019-03044-9>

HASSAN, Saeed-UI; BOWMAN, Timothy D.; SHABBIR, Mudassir; AKHTAR, Aqsa; IMRAN, Mubashir; ALJOHANI, Naif Radi. Influential tweeters in relation to highly cited articles in altmetric big data. **Scientometrics**, [s. l.], v. 119, n. 1, p. 481–493, abr. 2019. DOI 10.1007/s11192-019-03044-9. Disponível em: <http://link.springer.com/10.1007/s11192-019-03044-9>. Acesso em: 30 abr. 2025.

HASSAN, Saeed-UI; IMRAN, Mubashir; GILLANI, Uzair; ALJOHANI, Naif Radi; BOWMAN, Timothy D.; DIDEGAH, Fereshteh. Measuring social media activity of scientific literature: an exhaustive comparison of scopus and novel altmetrics big data. **Scientometrics**, [s. l.], v. 113, n. 2, p. 1037–1057, nov. 2017. DOI 10.1007/s11192-017-2512-x. Disponível em: <http://link.springer.com/10.1007/s11192-017-2512-x>. Acesso em: 1 maio 2025.

HAUSTEIN, Stefanie. Scholarly twitter metrics. **Springer Handbooks**, [s. l.], , p. 729–760, 2019. DOI https://doi.org/10.1007/978-3-030-02511-3_28

HAUSTEIN, Stefanie; BOWMAN, Timothy D.; COSTAS, Rodrigo. Interpreting ‘Altmetrics’: Viewing Acts on Social Media through the Lens of Citation and Social Theories. In: SUGIMOTO, Cassidy R. (org.). **Theories of Informetrics and Scholarly Communication**. [S. l.]: De Gruyter, 2016. p. 372–406. DOI 10.1515/9783110308464-022. Disponível em: <https://www.degruyter.com/document/doi/10.1515/9783110308464-022/html>. Acesso em: 19 jun. 2025.

HAUSTEIN, Stefanie; BOWMAN, Timothy D.; HOLMBERG, Kim; TSOU, Andrew; SUGIMOTO, Cassidy R.; LARIVIÈRE, Vincent. Tweets as impact indicators: Examining the implications of automated “bot” accounts onTwitter. **Journal of the Association for Information Science and Technology**, [s. l.], v. 67, n. 1, p. 232–238, jan. 2016. DOI 10.1002/asi.23456. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.23456>. Acesso em: 30 abr. 2025.

HAUSTEIN, Stefanie; LARIVIÈRE, Vincent; THELWALL, Mike; AMYOT, Didier; PETERS, Isabella. Tweets vs. Mendeley readers: How do these two social media metrics differ? **IT - Information Technology**, [s. l.], v. 56, n. 5, p. 207–215, 2014. DOI <https://doi.org/10.1515/itit-2014-1048>

HAYAWI, Kadhim; SAHA, Susmita; MASUD, Mohammad Mehedy; MATHEW, Sujith Samuel; KAOSAR, Mohammed. Social media bot detection with deep learning methods: a systematic review. **Neural Computing and Applications**, [s. l.], v. 35, n. 12, p. 8903–8918, 2023. DOI <https://doi.org/10.1007/s00521-023-08352-z>

JOUBERT, Marina; COSTAS, Rodrigo. Getting to Know Science Tweeters: A Pilot Analysis of South African Twitter Users Tweeting about Research Articles. **Journal of Altmetrics**, [s. l.], v. 2, n. 1, p. 2, 2019a. DOI <https://doi.org/10.29024/joa.8>

JOUBERT, Marina; COSTAS, Rodrigo. Getting to Know Science Tweeters: A Pilot Analysis of South African Twitter Users Tweeting about Research Articles. **Journal of Altmetrics**, [s. l.], v. 2, n. 1, p. 2, 14 maio 2019b. DOI 10.29024/joa.8. Disponível em: <https://account.ll-j-ja.ubiquityjournal.website/index.php/ll-j-ja/article/view/8>. Acesso em: 30 abr. 2025.

KELLER, Tobias R.; KLINGER, Ulrike. Social Bots in Election Campaigns: Theoretical, Empirical, and Methodological Implications. **Political Communication**, [s. l.], v. 36, n. 1, p. 171–189, 2019. <https://doi.org/10.1080/10584609.2018.1526238>. KLOUMANN, Isabel M.; DANFORTH, Christopher M.; HARRIS, Kameron Decker; BLISS, Catherine A.; DODDS, Peter Sheridan. Positivity of the English language. **PLoS ONE**, [s. l.], v. 7, n. 1, p. 0–6, 2012. DOI <https://doi.org/10.1371/journal.pone.0029484>

KONKIEL, Stacy. Altmetrics: diversifying the understanding of influential scholarship. **Palgrave Communications**, [s. l.], v. 2, n. 1, p. 1–7, 23 ago. 2016. DOI 10.1057/palcomms.2016.57. Disponível em: <https://www.nature.com/articles/palcomms201657>. Acesso em: 30 abr. 2025.

LINGAM, Greeshma; YASASWINI, Bhavya; JAGADAMBA, P.V.S.Lakshmi; KOLLIBOYANA, Niharika. An Improved Bot Identification with Imbalanced Data using GG-XGBoost. *In: 2022 INTERNATIONAL CONFERENCE ON INTELLIGENT TECHNOLOGIES (CONIT)*, 24 jun. 2022. **2022 2nd International Conference on Intelligent Technologies (CONIT)** [...]. Hubli, India: IEEE, 24 jun. 2022. p. 1–6. DOI 10.1109/conit55038.2022.9848252. Disponível em: <https://ieeexplore.ieee.org/document/9848252/>. Acesso em: 1 maio 2025.

LIU, Lu; JONES, Benjamin F.; UZZI, Brian; WANG, Dashun. Data, measurement and empirical methods in the science of science. **Nature Human Behaviour**, [s. l.], v. 7, n. 7, p. 1046–1058, 1 jun. 2023. DOI 10.1038/s41562-023-01562-4. Disponível em: <https://www.nature.com/articles/s41562-023-01562-4>. Acesso em: 27 jul. 2025.

MADAHALI, Lale; TIAN, Jin. Bots: Genuine or Malicious. *In: 2023 IEEE 14TH ANNUAL UBIQUITOUS COMPUTING, ELECTRONICS & MOBILE COMMUNICATION CONFERENCE (UEMCON)*, 12 out. 2023. **2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)** [...]. New York, NY, USA: IEEE, 12 out. 2023. p. 0669–0673. DOI 10.1109/UEMCON59035.2023.10316169. Disponível em: <https://ieeexplore.ieee.org/document/10316169/>. Acesso em: 3 maio 2025.

MAITY, Suman Kalyan; CHAKRABORTY, Aishik; GOYAL, Pawan; MUKHERJEE, Animesh. Detection of sockpuppets in social media. **CSCW 2017 - Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing**, [s. l.], p. 243–246, 2017. DOI <https://doi.org/10.1145/3022198.3026360>.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. [S. l.]: Editora Atlas Ltda, 2016.

MARICATO, João de Melo; MANSO, Bruno. Characterization of the communities of attention interacting with scientific papers on Twitter: altmetric analysis of a Brazilian University. **Scientometrics**, [s. l.], v. 127, n. 7, p. 3815–3835, jul. 2022. DOI 10.1007/s11192-022-04442-2. Disponível em: <https://link.springer.com/10.1007/s11192-022-04442-2>. Acesso em: 30 abr. 2025.

MARICATO, João de Melo; SILVA, Márcia Regina da; MANSO, Bruno Lara de Castro; SILVA, Vanessa Suelen da. Passado, presente e futuro da altmetria na perspectiva de especialistas de diferentes regiões globais: entrevistas com Rodrigo Costas e Ronaldo Araújo. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, Brasil, v. 16, p. e-240602, 2025. DOI: [10.11606/issn.2178-2075.incid.2025.240602](https://doi.org/10.11606/issn.2178-2075.incid.2025.240602). Disponível em: <https://revistas.usp.br/incid/article/view/240602>. Acesso em: 31 out. 2025.

MARTIN-GUTIERREZ, David; HERNANDEZ-PENALOZA, Gustavo; HERNANDEZ, Alberto Belmonte; LOZANO-DIEZ, Alicia; ALVAREZ, Federico. A Deep Learning Approach for Robust Detection of Bots in Twitter Using Transformers. **IEEE Access**, [s. l.], v. 9, p. 54591–54601, 2021. DOI <https://doi.org/10.1109/ACCESS.2021.3068659>

MARTÍN-GUTIÉRREZ, David; HERNÁNDEZ-PEÑALOZA, Gustavo; HERNÁNDEZ, Alberto Belmonte; LOZANO-DIEZ, Alicia; ÁLVAREZ, Federico. A Deep Learning Approach for Robust Detection of Bots in Twitter Using Transformers. **IEEE Access**, [s. l.], v. 9, p. 54591–54601, 2021. DOI 10.1109/ACCESS.2021.3068659. Disponível em: <https://ieeexplore.ieee.org/document/9385071>. Acesso em: 30 abr. 2025.

MARTINI, Franziska; SAMULA, Paul; KELLER, Tobias R.; KLINGER, Ulrike. Bot, or not? Comparing three methods for detecting social bots in five political discourses. **Big Data and Society**, [s. l.], v. 8, n. 2, 2021. DOI <https://doi.org/10.1177/20539517211033566>.

MELO, João Henrick Neri De; TRINCA, Tatiane Pacanaro; MARICATO, João De Melo. Limites dos indicadores bibliométricos de bases de dados internacionais para avaliação da Pós-Graduação brasileira: a cobertura da Web of Science nas diferentes áreas do conhecimento. **Transinformação**, [s. l.], v. 33, p. e200071, 2021. DOI 10.1590/2318-0889202133e200071. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-37862021000100508&tlang=pt. Acesso em: 29 maio 2025.

MESHRAM, Er. Pranay; BHAMBULKAR, Rutika; POKALE, Puja; KHARBIKAR, Komal; AWACHAT, Anushree. Automatic Detection of Fake Profile Using Machine Learning on Instagram. **International Journal of Scientific Research in Science and Technology**, [s. l.], p. 117–127, 2021. DOI <https://doi.org/10.32628/ijsrst218330>

MINNICH, Amanda; CHAVOSHI, Nikan; KOUTRA, Danai; MUEEN, Abdullah. BotWalk: Efficient adaptive exploration of twitter bot networks. **Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2017**, [s. l.], , p. 467–474, 2017. DOI <https://doi.org/10.1145/3110025.3110163>

MISLOVE, Alan; LEHMANN, Sune; AHN, Yong Yeol; ONNELA, Jukka Pekka; ROSENQUIST, J. Niels. Understanding the Demographics of Twitter Users. **Proceedings of the 5th International AAAI Conference on Weblogs and Social Media, ICWSM 2011**, [s. l.], n. January, p. 554–557, 2011. DOI <https://doi.org/10.1609/icwsm.v5i1.14168>

MITCHELL, J. Clyde. Social Networks. **Annual Review of Anthropology**, [s. l.], v. 3, p. 279–299, 1974. Disponível em: <http://www.jstor.org/stable/2949292>. Acesso em: 30 abr. 2025.

MITCHELL, Lewis; FRANK, Morgan R.; HARRIS, Kameron Decker; DODDS, Peter Sheridan; DANFORTH, Christopher M. The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place. **PLoS ONE**, [s. l.], v. 8, n. 5, 2013. DOI <https://doi.org/10.1371/journal.pone.0064417>.

MOCANU, Delia; BARONCELLI, Andrea; PERRA, Nicola; GONÇALVES, Bruno; ZHANG, Qian; VESPIGNANI, Alessandro. The Twitter of Babel: Mapping World Languages through Microblogging Platforms. **PLoS ONE**, [s. l.], v. 8, n. 4, 2013. DOI <https://doi.org/10.1371/journal.pone.0061981>.

MOHAMMADI, Ehsan; THELWALL, Mike; KWASNÝ, Mary; HOLMES, Kristi L. Academic information on Twitter: A user survey. **PLoS ONE**, [s. l.], v. 13, n. 5, p. 1–18, 2018. DOI <https://doi.org/10.1371/journal.pone.0197265>.

MOHANTY, Sanjukta; DWIVEDY, Satya Prakash; ACHARYA, Arup Abhinna; MOHAPATRA, Suvakanta; SAHOO, Shivam Swastik; SAMAL, Sibadatta; SAMAL, Smritisrita. Enhancing the Detection of Social bots on Twitter using Ensemble machine Learning Technique. In: 2022 INTERNATIONAL CONFERENCE ON ADVANCEMENTS IN SMART, SECURE AND INTELLIGENT COMPUTING (ASSIC), 19 nov. 2022. **2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)** [...]. Bhubaneswar, India: IEEE, 19 nov. 2022. p. 1–6. DOI 10.1109/assic55218.2022.10088372. Disponível em: <https://ieeexplore.ieee.org/document/10088372/>. Acesso em: 1 maio 2025.

MONGEON, Philippe; BOWMAN, Timothy D.; COSTAS, Rodrigo. An open data set of scholars on Twitter. **Quantitative Science Studies**, [s. l.], v. 4, n. 2, p. 314–324, 2023a. DOI https://doi.org/10.1162/qss_a_00250.

MONGEON, Philippe; BOWMAN, Timothy D.; COSTAS, Rodrigo. An open data set of scholars on Twitter. **Quantitative Science Studies**, [s. l.], v. 4, n. 2, p. 314–324, 1 maio 2023b. DOI 10.1162/qss_a_00250. Disponível em: https://doi.org/10.1162/qss_a_00250. Acesso em: 30 abr. 2025.

NG, Jeremy Y.; VERHOEFF, Natasha; STEEN, Jeremy. What are the ways in which social media is used in the context of complementary and alternative medicine in the health and medical scholarly literature? a scoping review. **BMC Complementary Medicine and Therapies**, [s. l.], v. 23, n. 1, p. 32, 2 fev. 2023. DOI 10.1186/s12906-023-03856-6. Disponível em: <https://bmccomplementalternmed.biomedcentral.com/articles/10.1186/s12906-023-03856-6>. Acesso em: 2 maio 2025.

NISHIKAWA-PACHE, Andreas. Measuring serendipity with altmetrics and randomness. 2023. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/09610006221124338>. Acesso em: 30 abr. 2025.

NISHIKAWA-PACHER, Andreas. Measuring serendipity with altmetrics and randomness. **Journal of Librarianship and Information Science**, [s. l.], v. 55, n. 4, p. 1078–1087, 2023. DOI <https://doi.org/10.1177/09610006221124338>.

NISHIKAWA-PACHER, Andreas. The Twitter accounts of scientific journals: a dataset. **Insights: the UKSG Journal**, [s. l.], v. 36, 2023. DOI <https://doi.org/10.1629/uksg.593>.

NUNES, Amanda Maria de Almeida. **MÁQUINAS SOCIAIS E A DESINFORMAÇÃO EM REDE: O papel das entidades de software na formação de opinião na Internet**. 2020. 175 f. Universidade Federal de Pernambuco, Recife, 2020.

OENTARYO, Richard J.; MURDOPO, Arinto; PRASETYO, Philips K.; LIM, Ee Peng. On profiling bots in social media. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, [s. l.], v. 10046 LNCS, n. March 2018, p. 92–109, 2016. DOI https://doi.org/10.1007/978-3-319-47880-7_6.

ORABI, Mariam; MOUHEB, Djedjiga; AL AGHBARI, Zaher; KAMEL, Ibrahim. Detection of Bots in Social Media: A Systematic Review. **Information Processing and Management**, [s. l.], v. 57, n. 4, 2020. DOI <https://doi.org/10.1016/j.ipm.2020.102250>.

ORDUÑA-MALEA, Enrique; FONT-JULIÁN, Cristina I. Are patents linked on Twitter? A case study of Google patents. **Scientometrics**, [s. l.], v. 127, n. 11, p. 6339–6362, 2022. DOI <https://doi.org/10.1007/s11192-022-04519-y>

ORTEGA, José-Luis. Altmetrics data providers: A meta-analysis review of the coverage of metrics and publication. **El Profesional de la Información**, [s. l.], v. 29, n. 1, 30 jan. 2020. DOI 10.3145/epi.2020.ene.07. Disponível em: <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2020.en.e.07>. Acesso em: 2 maio 2025.

OWAIS, Muhammad; SHOAIB, Muhammad; WASEEM, Muhammad. Towards a Comprehensive Approach for Socialbot Detection on Twitter: Integrating Multiple Features. [s. l.], 5 abr. 2023. DOI 10.21203/rs.3.rs-2764404/v1. Disponível em: <https://www.researchsquare.com/article/rs-2764404/v1>. Acesso em: 1 maio 2025.

PONTES, Danielle; MARICATO, João. DETECÇÃO DE BOTS QUE DIVULGAM ARTIGOS CIENTÍFICOS NO TWITTER : CONTRIBUIÇÕES PRELIMINARES PARA APRIMORAMENTO DOS INDICADORES ALTMÉTRICOS. 436., 2022. **XXII ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO** [...]. PORTO ALEGRE RS: ANCIB, 2022. v. 436.

PONTES, Danielle Pompeu Noronha; MARICATO, João De Melo. Classificação de comunidades de atenção altmétrica no X avaliando a acurácia de algoritmos sob

diferentes variáveis. *In: 9º ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA - EBBC, 2024a. Anais do 9º Encontro Brasileiro de Bibliometria e Cientometria - EBBC [...].* [s. l.]: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2024. DOI 10.22477/ix.ebbc.323. Disponível em: <https://ebbc.inf.br/ojs/index.php/ebbc/article/view/323>. Acesso em: 23 jul. 2025.

PONTES, Danielle Pompeu Noronha; MARICATO, João de Melo. Classificação de contas do Twitter através do botometer. *In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, ISSN 2177-3688.*, 2023. **Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação**, [...]. Aracaju-SE: Brapci, 2023. v. ISSN 2177-3688, . Disponível em: <https://brapci.inf.br/v/258099>. Acesso em: 30 abr. 2025.

PONTES, Danielle Pompeu Noronha; MARICATO, João de Melo. DETECÇÃO DE BOTS QUE DIVULGAM ARTIGOS CIENTÍFICOS NO TWITTER: CONTRIBUIÇÕES PARA APRIMORAMENTO DOS INDICADORES ALTMÉTRICOS. [s. l.], 2022.

PONTES, Danielle Pompeu Noronha; MARICATO, João De Melo. Padrões de comportamento das contas bots e humanas na mídia x por meio da análise de sentimentos no chatgpt. *In: 9º ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA - EBBC, 2024b. Anais do 9º Encontro Brasileiro de Bibliometria e Cientometria - EBBC [...].* [s. l.]: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2024. DOI 10.22477/ix.ebbc.326. Disponível em: <https://ebbc.inf.br/ojs/index.php/ebbc/article/view/326>. Acesso em: 30 maio 2025.

PONTES, Danielle Pompeu Noronha; MARICATO, João De Melo. Usando análise de sentimentos para diferenciar bots e humanos na divulgação de publicações científicas sobre COVID-19 na mídia social X: um estudo com ChatGPT 3.5 e Gemini 1.5 Flash. **Biblios Journal of Librarianship and Information Science**, [s. l.], n. esp., p. e001, 3 jul. 2025. DOI 10.5195/biblios.2025.1297. Disponível em: <http://biblios.pitt.edu/ojs/biblios/article/view/1297>. Acesso em: 24 jul. 2025.

POZZANA, Iacopo; FERRARA, Emilio. Measuring Bot and Human Behavioral Dynamics. **Frontiers in Physics**, [s. l.], v. 8, p. 125, 22 abr. 2020. DOI 10.3389/fphy.2020.00125. Disponível em: <https://www.frontiersin.org/article/10.3389/fphy.2020.00125/full>. Acesso em: 3 maio 2025.

PRIEM, Jason; COSTELLO, Kaitlin Light. How and why scholars cite on Twitter. **Proceedings of the American Society for Information Science and Technology**, [s. l.], v. 47, n. 1, p. 1–4, nov. 2010. DOI 10.1002/meet.14504701201. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/meet.14504701201>. Acesso em: 30 abr. 2025.

PRIEM, Jason; HEMMINGER, Bradely H. Scientometrics 2.0: New metrics of scholarly impact on the social Web. **First Monday**, [s. l.], 2 jul. 2010. <https://doi.org/10.5210/fm.v15i7.2874>.

PRODANOV, Cleber Cristiano. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico.** [S. l.]: Universidade Feevale, 2012.

QI, SiHua; ALKULAIB, Lulwah; BRONIATOWSKI, David A. Detecting and Characterizing Bot-Like Behavior on Twitter. [S. l.: s. n.], 2018a. p. 228–232. DOI https://doi.org/10.1007/978-3-319-93372-6_26.

QI, SiHua; ALKULAIB, Lulwah; BRONIATOWSKI, David A. Detecting and Characterizing Bot-Like Behavior on Twitter. In: THOMSON, Robert; DANCY, Christopher; HYDER, Ayaz; BISGIN, Halil (org.). **Social, Cultural, and Behavioral Modeling.** Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018b. v. 10899, p. 228–232. DOI 10.1007/978-3-319-93372-6_26. Disponível em: http://link.springer.com/10.1007/978-3-319-93372-6_26. Acesso em: 3 maio 2025.

RATKIEWICZ, Jacob; CONOVER, Michael; MEISS, Mark; GONÇALVES, Bruno; PATIL, Snehal; FLAMMINI, Alessandro; MENCZER, Filippo. Truthy: mapping the spread of astroturf in microblog streams. In: WWW '11: 20TH INTERNATIONAL WORLD WIDE WEB CONFERENCE, 28 mar. 2011. **Proceedings of the 20th international conference companion on World wide web** [...]. Hyderabad India: ACM, 28 mar. 2011. p. 249–252. DOI 10.1145/1963192.1963301. Disponível em: <https://dl.acm.org/doi/10.1145/1963192.1963301>. Acesso em: 3 maio 2025.

RIAIZI, Amin; WANG, Yingfeng. Differentiating broadcast from viral: a causal inference approach for information diffusion analysis. **Applied Intelligence**, [s. l.], , p. 10374–10385, 2024a. DOI <https://doi.org/10.1007/s10489-024-05723-4>.

RIAIZI, Amin; WANG, Yingfeng. Differentiating broadcast from viral: a causal inference approach for information diffusion analysis. **Applied Intelligence**, [s. l.], v. 54, n. 21, p. 10374–10385, 14 ago. 2024b. DOI 10.1007/s10489-024-05723-4. Disponível em: <https://doi.org/10.1007/s10489-024-05723-4>. Acesso em: 30 abr. 2025.

ROBINSON-GARCÍA, Nicolás; TORRES-SALINAS, Daniel; ZAHEDI, Zohreh; COSTAS, Rodrigo. New data, new possibilities: exploring the insides of *Altmetric.com*. **El Profesional de la Información**, [s. l.], v. 23, n. 4, p. 359–366, 1 maio 2014. DOI 10.3145/epi.2014.jul.03. Disponível em: <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2014.jul.03>. Acesso em: 2 maio 2025.

SAQR, Mohammed; LÓPEZ-PERNAS, Sonsoles; APIOLA, Mikko. Capturing the Impact and the Chatter Around Computing Education Research Beyond Academia in Social Media, Patents, and Blogs. In: APIOLA, Mikko; LÓPEZ-PERNAS, Sonsoles; SAQR, Mohammed (org.). **Past, Present and Future of Computing Education**

Research. Cham: Springer International Publishing, 2023. p. 171–191. DOI 10.1007/978-3-031-25336-2_9. Disponível em: https://link.springer.com/10.1007/978-3-031-25336-2_9. Acesso em: 2 maio 2025.

SCHOLARLY TWITTER METRICS. **Cham:** Springer International Publishing, 2019. Disponível em: http://link.springer.com/10.1007/978-3-030-02511-3_28. Acesso em: 30 abr. 2025.

SEGADO-BOJ, Francisco; DÍAZ-CAMPO, Jesús; FERNÁNDEZ-GÓMEZ, Erika; CHAPARRO-DOMÍNGUEZ, María-Ángeles. Spanish academics and social networking sites: Use, non-use, and the perceived advantages and drawbacks of Facebook, Twitter, LinkedIn, ResearchGate, and Academia.edu. **First Monday**, [s. l.], 30 abr. 2019. DOI <https://doi.org/10.5210/fm.v24i5.7296>.

SHAO, Chengcheng; CIAMPAGLIA, Giovanni Luca; VAROL, Onur; YANG, Kai-Cheng; FLAMMINI, Alessandro; MENCZER, Filippo. The spread of low-credibility content by social bots. **Nature Communications**, [s. l.], v. 9, n. 1, p. 4787, 20 nov. 2018. DOI 10.1038/s41467-018-06930-7. Disponível em: <https://www.nature.com/articles/s41467-018-06930-7>. Acesso em: 15 maio 2025.

SHEARER, C. The CRISP-DM Model: The New Blueprint for Data Mining. **Journal of Data Warehousing**, [s. l.], v. 5, n. 4, p. 13–22, 2000. .
SHUAI, Xin; PEPE, Alberto; BOLLEN, Johan. How the Scientific Community Reacts to Newly Submitted Preprints: Article Downloads, Twitter Mentions, and Citations. **PLoS ONE**, [s. l.], v. 7, n. 11, p. 1–8, 2012.
DOI <https://doi.org/10.1371/journal.pone.0047523>

SUBRAHMANIAN, V. S.; AZARIA, Amos; DURST, Skylar; KAGAN, Vadim; GALSTYAN, Aram; LERMAN, Kristina; ZHU, Linhong; FERRARA, Emilio; FLAMMINI, Alessandro; MENCZER, Filippo. The DARPA Twitter Bot Challenge. **Computer**, [s. l.], v. 49, n. 6, p. 38–46, 2016. <https://doi.org/10.1109/MC.2016.183>.
THELWALL, Mike. The Pros and Cons of the Use of Altmetrics in Research Assessment. **Scholarly Assessment Reports**, [s. l.], v. 2, n. 1, p. 1–9, 2020. DOI <https://doi.org/10.29024/sar.10>

THELWALL, Mike; TSOU, Andrew; WEINGART, Scott; HOLMBERG, Kim; HAUSTEIN, Stefanie. Tweeting links to academic articles. **Cybermetrics**, [s. l.], v. 17, n. 1, p. 1–8, 2013.

THURAISINGHAM, Bhavani. The role of artificial intelligence and cyber security for social media. **Proceedings - 2020 IEEE 34th International Parallel and Distributed Processing Symposium Workshops, IPDPSW 2020**, [s. l.], , p. 1116–1118, 2020. DOI <https://doi.org/10.1109/IPDPSW50202.2020.00184>

TSOU, Andrew; BOWMAN, Tim; GHAZINEJAD, Ali; SUGIMOTO, Cassidy. Who Tweets about Science? *In:* 15TH INTERNATIONAL SOCIETY OF

SCIENTOMETRICS AND INFORMETRICS, 29 jul. 2015. **Proceedings of ISSI 2015** [...]. Istanbul, Turkey: Bogaziçi University Printhouse, 29 jul. 2015.

TUNGER, Dirk; CLERMONT, Marcel; MEIER, Andreas. Altmetrics: State of the Art and a Look into the Future. **IntechOpen**, [s. l.], v. 32, n. tourism, p. 123–134, 2018. DOI <http://dx.doi.org/10.5772/intechopen.76874>.

VAROL, Onur; FERRARA, Emilio; DAVIS, Clayton A.; MENCZER, Filippo; FLAMMINI, Alessandro. Online human-bot interactions: Detection, estimation, and characterization. **Proceedings of the 11th International Conference on Web and Social Media, ICWSM 2017**, [s. l.], , p. 280–289, 2017a. DOI <https://doi.org/10.1609/icwsm.v11i1.14871>.

VAROL, Onur; FERRARA, Emilio; DAVIS, Clayton; MENCZER, Filippo; FLAMMINI, Alessandro. Online Human-Bot Interactions: Detection, Estimation, and Characterization. **Proceedings of the International AAAI Conference on Web and Social Media**, [s. l.], v. 11, n. 1, p. 280–289, 3 maio 2017b. DOI 10.1609/icwsm.v11i1.14871. Disponível em: <https://ojs.aaai.org/index.php/ICWSM/article/view/14871>. Acesso em: 27 jul. 2025.

VEALE, Tony; COOK, Mike. **Twitterbots: Making Machines that Make Meaning**. [S. l.]: The MIT Press, 2018. DOI 10.7551/mitpress/10859.001.0001. Disponível em: <https://direct.mit.edu/books/book/4113/TwitterbotsMaking-Machines-that-Make-Meaning>. Acesso em: 3 maio 2025.

WARRINER, Amy Beth; KUPERMAN, Victor; BRYSBAAERT, Marc. Norms of valence, arousal, and dominance for 13,915 English lemmas. **Behavior Research Methods**, [s. l.], v. 45, n. 4, p. 1191–1207, 2013. <https://doi.org/10.3758/s13428-012-0314-x>. WEI, Feng; NGUYEN, Uyen Trang. Twitter Bot Detection Using Neural Networks and Linguistic Embeddings. **IEEE Open Journal of the Computer Society**, [s. l.], v. PP, p. 1–12, 2023. DOI <https://doi.org/10.1109/OJCS.2023.3302286>.

WILSON, Theresa; WIEBE, Janyce; HOFFMANN, Paul. Recognizing contextual polarity in phrase-level sentiment analysis. **HLT/EMNLP 2005 - Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference**, [s. l.], v. 7, n. 5, p. 347–354, 2005. DOI <https://doi.org/10.3115/1220575.1220619>.

WOJCIK, Stefan; WEISEL, Rachel; SMITH, Aaron. Bots in the Twittersphere. **Pew Research Center**, [s. l.], 9 abr. 2018. Disponível em: <https://www.pewresearch.org/internet/2018/04/09/bots-in-the-twittersphere/>. Acesso em: 13 set. 2024.

XGBOOST DEVELOPERS. **XGBoost Parameters**. 2022. **XGBoost 2.0.3 documentation**. Disponível em: <https://xgboost.readthedocs.io/en/stable/parameter.html#parameters-for-tree-booster>. Acesso em: 24 ago. 2025.

XU, Shenmeng; YU, Houqiang; HEMMINGER, Bradley M.; DONG, Xie. Who, what, why? An exploration of JoVE scientific video publications in tweets. **Scientometrics**, [s. l.], v. 117, n. 2, p. 845–856, nov. 2018. DOI 10.1007/s11192-018-2880-x. Disponível em: <http://link.springer.com/10.1007/s11192-018-2880-x>. Acesso em: 30 abr. 2025.

YE, Yingxin Estella; NA, Jin Cheon. Who is mentioning COVID-19 articles on twitter? Classifying twitter users in the context of scholarly communication. **Social Network Analysis and Mining**, [s. l.], v. 14, n. 1, 2024. doi <https://doi.org/10.1007/s13278-024-01236-7>.

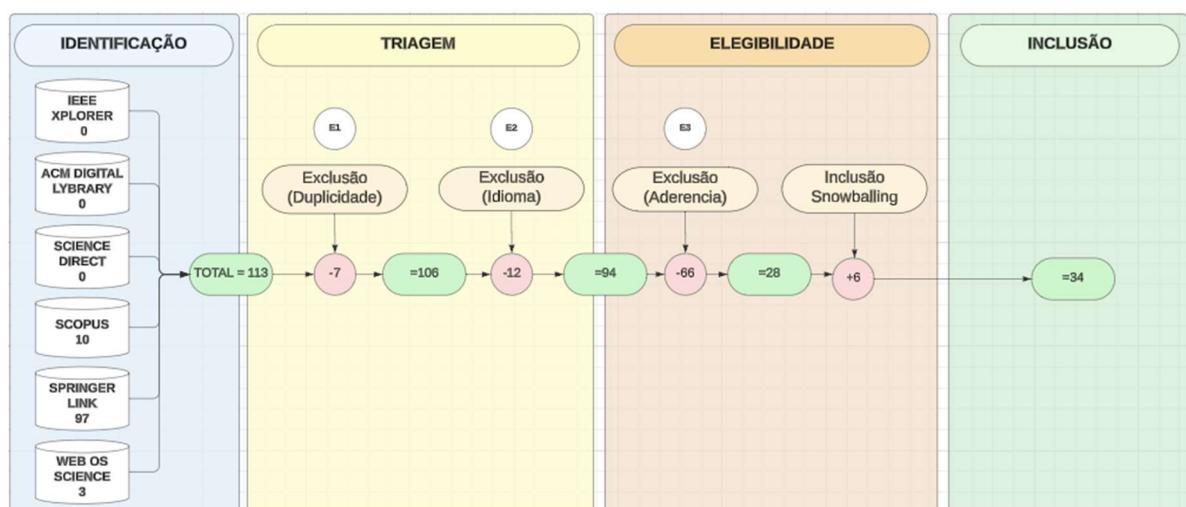
YOSHIKAWA, Naruki; KUBO, Ryuichi; YAMAMOTO, Kazuki Z. Twitter integration of chemistry software tools. **Journal of Cheminformatics**, [s. l.], v. 13, n. 1, p. 1–6, 2021. doi <https://doi.org/10.1186/s13321-021-00527-x>.

APÊNDICE A - PROCEDIMENTOS METODOLÓGICOS ADOTADOS PARA A REVISÃO SISTEMÁTICA DE LITERATURA

A adoção de protocolos bem estabelecidos é fundamental para a condução de revisões sistemáticas de literatura, garantindo rigor metodológico e transparência na coleta e análise de dados. O protocolo PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) é amplamente utilizado como referência para a condução de revisões e metanálises, estruturando o processo em etapas que incluem a identificação, triagem, avaliação de elegibilidade e inclusão dos estudos (Galvão *et al.*, 2022). Cada uma dessas etapas é cuidadosamente delineada para assegurar que as revisões sejam abrangentes e replicáveis, proporcionando um relato claro e objetivo dos métodos utilizados. Neste contexto, o presente protocolo segue a orientação do PRISMA, com adaptações específicas para a revisão de literatura sobre detecção de *bots* na altmetria, assegurando uma seleção criteriosa das fontes e uma análise profunda das evidências encontradas para responder às questões de pesquisa previamente definidas.

Respeitando as 4 etapas, o protocolo proposto é composto por sete atividades: (1) definição de protocolo, (2) formulação de questões de pesquisa, (3) construção de strings de busca, (4) seleção de bases de dados, (5) aplicação dos critérios de inclusão e exclusão, (6) seleção dos estudos e (7) síntese dos resultados. A Figura 6 apresenta o fluxograma detalhado de execução do protocolo.

Figura 6 - Fluxograma de execução do protocolo de Revisão Sistemática de Literatura



Fonte: Dados da pesquisa (2025).

O protocolo de pesquisa foi definido para orientar a revisão sistemática de literatura sobre a detecção de *bots* no Twitter que promovem artigos científicos e afetam os *scores* altmétricos. O protocolo abrange:

1. **As questões de pesquisa** derivadas do objetivo principal.
2. **A estratégia de busca** para localizar artigos relevantes nas principais bases de dados a partir da formulação da **String de Busca**.
3. **A seleção das bases de dados** de acordo com as especificidades do tema
4. **Os critérios de inclusão e exclusão** para selecionar os estudos a serem analisados.

A.1 Questões de pesquisa

A questão principal da pesquisa é: “*Quais são os métodos, ferramentas e desafios na detecção de bots que influenciam os scores altmétricos de artigos científicos no X?*”

A partir desta questão, foram derivadas as questões secundárias apresentadas a seguir.

RQ1: Quais são os métodos e algoritmos existentes para detectar *bots* no X?

RQ2: Quais características comportamentais distinguem *bots* de usuários humanos na divulgação de artigos científicos?

RQ3: Quais métricas e atributos específicos (e.g., frequência de postagens, rede de conexões, padrões de linguagem) são mais usados para identificar *bots* no contexto da disseminação de artigos?

RQ4: Como o uso de *bots* afeta a precisão e a validade das métricas altmétricas no contexto acadêmico?

RQ5: Quais ferramentas e *frameworks* têm sido utilizados para análise de *bots* no X?

RQ6: Quais são os principais desafios e limitações na detecção de *bots* que influenciam altmétricas?

RQ7: Quais são as implicações éticas e legais no uso de *bots* para aumentar artificialmente a disseminação de artigos?

RQ8: Quais abordagens foram propostas para mitigar a influência dos *bots* na mensuração de impacto de artigos científicos?

RQ9: Como diferentes tipos de *bots* (e.g., *bots* de spam, *bots* promocionais, *bots* acadêmicos) afetam, de maneira diferenciada, os *scores* altmétricos?

RQ10: Quais são as tendências futuras na detecção e na mitigação do impacto de *bots* na altmetria?

A.2. Estratégia de Busca: Formulação das Strings de Busca

A formulação das strings de busca segue uma estratégia baseada em palavras-chave para cobrir conceitos centrais relacionados a bots, X e altmétria. As palavras-chave foram organizadas em três categorias:

- **Categoria 1 (Detecção de Bots):** ("bot detection" OR "bot identification" OR "social bots" OR "Twitter bots")
- **Categoria 2 (Altmetrics):** ("altmetrics" OR "social media metrics" OR "article-level metrics" OR "research impact")
- **Categoria 3 (X e Divulgação de Artigos):** ("Twitter" OR "social media" OR "scholarly communication" OR "scientific dissemination")

Essas palavras-chave foram combinadas com operadores booleanos, resultando na seguinte string de busca:

(("bot detection") OR ("bot identification") OR ("social bots") OR ("Twitter bots"))
AND (("altmetrics") OR ("social media metrics") OR ("article-level metrics") OR ("research impact")) AND (("Twitter") OR ("social media") OR ("scholarly communication") OR ("scientific dissemination"))

A string de busca foi aplicada aos metadados ABSTRACT, TÍTULO E PALAVRAS-CHAVE

A.3. Seleção das Bases de Dados

A escolha das bases de dados para uma revisão de literatura é um passo fundamental para garantir que a pesquisa seja abrangente, de alta qualidade e representativa do estado da arte. A seleção deve levar em consideração o escopo dos artigos indexados, a relevância para a área de estudo e a credibilidade das fontes. Desta forma, as bases de dados escolhidas para a revisão abrangem periódicos especializados em Ciência da Computação, Ciência da Informação e estudos interdisciplinares entre comunicação científica e redes sociais. A seguir, serão

apresentadas as bases de dados selecionadas para a revisão sobre a detecção de *bots* no X e o impacto nas altmétricas:

ACM Digital Library, uma das principais fontes de pesquisa em computação e tecnologia da informação. É especialmente relevante para estudos de detecção de *bots* e redes sociais, pois possui uma vasta coleção de conferências e periódicos em aprendizado de máquina, inteligência artificial, mineração de dados e análise de redes sociais — temas essenciais para o desenvolvimento de algoritmos de detecção e análise de comportamento automatizado. Além disso, a ACM cobre tópicos interdisciplinares, como o impacto de novas tecnologias em diferentes áreas, incluindo comunicação acadêmica.

IEEE Xplore, uma referência mundial em pesquisas tecnológicas e de engenharia, com uma forte ênfase em sistemas de informação e inteligência artificial. Como a detecção de *bots* no X geralmente envolve a aplicação de técnicas avançadas de IA e aprendizado de máquina, a IEEE se torna relevante para identificar métodos de ponta e avanços no uso de redes neurais, algoritmos de processamento de linguagem natural (PLN) e detecção automatizada de padrões de comportamento. Ela também oferece acesso a conferências de alto impacto, como a *International Conference on Data Mining* (ICDM) e a *Conference on Computer Communications* (INFOCOM).

Science Direct, uma das maiores coleções de artigos acadêmicos revisados por pares, abrange áreas que vão desde ciências sociais até tecnologia, o que permite acessar estudos que discutem não apenas a detecção de *bots*, mas também as suas implicações sociais e éticas no uso de métricas alternativas (altnétricas) para a disseminação de conhecimento. É uma base importante para capturar estudos interdisciplinares que combinam ciência da informação, comunicação acadêmica e ética digital.

Scopus, uma das maiores bases de dados multidisciplinares e amplamente utilizada para revisões de literatura por oferecer uma cobertura extensa de periódicos de alta qualidade e conferências em diversas disciplinas. Ela fornece não apenas artigos, mas também análises biométricas, citações e métricas de impacto, permitindo a identificação de tendências de pesquisa e autores influentes. Scopus é particularmente importante para mapear o campo de estudo de forma ampla e identificar conexões entre disciplinas como ciência da computação, comunicação e ciências sociais.

Springer Link: base de dados que combina estudos teóricos e práticos em tecnologia, ciências sociais e ciências aplicadas. Seu valor está na ampla coleção de livros, capítulos e periódicos que exploram os métodos de detecção de *bots*, além das questões teóricas e práticas relacionadas ao uso de IA e suas implicações. É ideal para encontrar publicações mais extensas e detalhadas que abordam aspectos interdisciplinares e contribuem para a formação de uma base teórica sólida.

Web of Science: base multidisciplinar com foco em estudos de impacto e métricas de citação. Ela é frequentemente usada em análises bibliométricas e altmétricas, facilitando o estudo do impacto de artigos e da influência de *bots* nas métricas acadêmicas. A Web of Science é particularmente útil para capturar tendências de citações e análises altmétricas, permitindo uma análise profunda de como a disseminação artificial (por meio de *bots*) pode distorcer o impacto percebido de trabalhos científicos.

A escolha dessas bases se baseou em três critérios principais:

1. Cobertura Multidisciplinar: Para estudar detecção de *bots* no X, é necessário acessar artigos de várias áreas: ciência da computação, comunicação, ciência da informação, ética e ciências sociais. Por isso, bases como Scopus, Web of Science e Science Direct são essenciais para dar uma visão completa.

2. Qualidade e Rigor Acadêmico: Bases como ACM, IEEE e Springer são reconhecidas pelo rigor em suas publicações, garantindo que apenas estudos revisados por pares e de alta qualidade sejam incluídos na revisão.

3. Foco em Tecnologia e Ciências Sociais: A combinação de bases como ACM, IEEE e as coleções jurídicas e sociais, como Springer, permite mapear um panorama abrangente, cobrindo desde os aspectos técnicos (detecção de *bots*) até suas consequências sociais (impacto nas altmétricas).

As bases citadas, foram selecionadas por cobrirem tópicos de tecnologia, redes sociais e métricas acadêmicas, garantindo a inclusão de estudos interdisciplinares. Essa escolha estratégica garante que a revisão de literatura seja abrangente, interdisciplinar e cientificamente robusta, proporcionando uma análise detalhada dos métodos, dos desafios e dos impactos da detecção de *bots* no campo da comunicação acadêmica. A seguir serão apresentados os Critérios de Inclusão e Exclusão adotados neste protocolo.

A.4. Critérios de Inclusão e Exclusão

Os critérios definidos para seleção de artigos visam garantir que apenas estudos relevantes e de alta qualidade sejam considerados:

- **Critérios de Inclusão:**
 - IC1: Estudos que abordam a detecção de *bots* no X em contextos de divulgação científica.
 - IC2: Artigos que discutem como *bots* afetam a altmétria ou métricas de impacto de artigos acadêmicos.
 - IC3: Publicações revisadas por pares, com foco em IA, ciência da informação ou altmetria.
- **Critérios de Exclusão:**
 - EC1: Artigos EM DUPLICIDADE
 - EC1: Artigos QUE NÃO ESTAVAM EM INGLÊS
 - EC2: Estudos que tratam de *bots*, mas não no contexto de divulgação científica ou altmétricas.

A.5. Seleção dos Estudos

O processo de seleção dos estudos foi composto por quatro etapas:

- b. **Remoção de duplicatas** e artigos que atendem aos critérios de exclusão.
- c. **Leitura dos títulos, resumos e palavras-chave** para uma triagem inicial de relevância.
- d. **Avaliação do texto completo** para garantir que o estudo aborda os métodos e os desafios na detecção de *bots* que afetam altmétricas.

A.6. Extração e Análise dos Dados

As buscas foram realizadas no período de 30/09/2024 a 03/10/2024. A chave de busca foi aplicada nas bases de dados nos campos de Título, Palavras-chaves e abstract, filtrando por artigos.

Os resultados da busca utilizando a *string* de pesquisa proposta mostram uma variação significativa na recuperação de estudos entre diferentes bases de dados. A consulta foi estruturada para capturar publicações relacionadas à detecção de *bots*

em ambientes de redes sociais, com foco em métricas alternativas e disseminação científica.

Conforme pode ser observado no Quadro 7, a base de dados Springer Link apresentou o maior número de registros recuperados (97), sugerindo que esta fonte contém um volume considerável de estudos relevantes para o tópico. Esse resultado pode ser explicado pela natureza multidisciplinar e pelo enfoque em comunicação científica e impacto social, áreas fortemente representadas nos periódicos indexados nessa plataforma.

Quadro 7 - Resultados da busca bibliográfica (30/09 a 03/10/2024) sobre detecção de *bots* e métricas altmétricas nas bases consultadas

DATA DE BUSCA:	30/09/2024 a 03/10/2024							
CHAVE DE BUSCA:	(("bot detection") OR ("bot identification") OR ("social bots") OR ("Twitter bots")) AND ((("altmetrics") OR ("social media metrics") OR ("article-level metrics") OR ("research impact")) AND ((("Twitter") OR ("social media") OR ("scholarly communication") OR ("scientific dissemination"))))							
BASES	CM DL	IEEE XPLORER	SCIE NCE DIRECT	COPUS S	RINGER LINK	SP	OS	TOTAL
QTD DE ARTIGOS		0	0	0	1	97		113

Fonte: Dados da pesquisa (2025).

Em contrapartida, *Scopus* e *Web of Science* (WOS) recuperaram menos publicações, com 10 e 6 registros, respectivamente. Embora o número seja inferior ao da *Springer Link*, essas bases são reconhecidas pela qualidade e pelo rigor na seleção de documentos, o que indica que os artigos encontrados possuem alto grau de relevância e rigor metodológico.

Por outro lado, nenhuma publicação foi encontrada nas bases *ACMDL*, *IEEE Xplore* e *ScienceDirect* para a *string* de busca empregada. A ausência de resultados nestas bases pode ser atribuída ao fato de que esses repositórios têm maior foco em computação, engenharia e tecnologia, áreas nas quais o termo "altmetrics" e tópicos relacionados à disseminação científica ainda são menos abordados ou não integram o foco principal dos artigos.

Dessa forma, a análise preliminar revela que os estudos sobre detecção de *bots* com aplicação em altmetria e disseminação científica ainda estão em desenvolvimento e são predominantemente discutidos em fontes com uma orientação mais interdisciplinar, como a *Springer Link*.

Os dados extraídos de cada estudo selecionado incluirão:

- **Métodos de detecção de bots** empregados e sua eficácia.
- **Características comportamentais** usadas para distinguir *bots* de humanos.
- **Impacto dos bots nos scores altmétricos** e análise de casos específicos.
- **Ferramentas e frameworks** utilizados.
- **Desafios e oportunidades** para novas pesquisas na área.

O Quadro 8 apresenta os resultados da aplicação dos critérios de exclusão ao protocolo de busca bibliográfica nas bases de dados selecionadas. Após a recuperação inicial de 113 artigos, foram realizadas etapas de filtragem a fim de garantir a consistência e relevância do corpus analisado. O primeiro critério (EC1) eliminou trabalhos duplicados entre bases, totalizando sete exclusões. Em seguida, o critério de idioma (EC2) resultou na retirada de doze artigos não compatíveis com o escopo linguístico definido. Por fim, o critério de aderência temática (EC3) excluiu 66 publicações que, embora recuperadas pela estratégia de busca, não apresentavam alinhamento direto com os objetivos da pesquisa. Após esses procedimentos, permaneceu um total de 28 artigos, que constituem a base final para análise.

Quadro 8 - Resultados da aplicação dos critérios de exclusão ao protocolo de busca bibliográfica nas bases consultadas

BASES:	CMDL	XPLORE	IEE DIRECT	SCIENCE	COPUS	S ER LINK	SPRING	OS	TOTAL
QTD DE ARTIGOS:		0	0	0	1	97			13
CRITÉRIOS DE EXCLUSÃO									
EC1 (DUPLICIDADES)					2	4			
EC2 (IDIOMA)						12		2	
EC3 (ADERÊNCIA)						66		6	
						TOTAL:		8	

Fonte: Dados da pesquisa (2025).

APÊNDICE B - PLANO DE GESTÃO DE DADOS (PGD)

1. Descrição dos Dados

Tipos de dados: Tabelas Excel, arquivos CSV, scripts Python, bases de dados de artigos científicos, *posts* de redes sociais, *datasets* de contas classificadas.

Volume estimado: Aproximadamente 11.2 GB.

Formatos dos arquivos: XLSX, CSV, PY, PKL.

Origem dos dados: Altmetric.com, bases de dados acadêmicas, coleta própria em redes sociais.

2. Coleta e Documentação dos Dados

Métodos de coleta: Download direto de bases de dados (Altmetric.com), extração automatizada via scripts Python e coletas manuais.

Ferramentas e softwares utilizados: Python (bibliotecas pandas, numpy, requests), Excel, GitHub, Google Colab.

Padrões de metadados adotados: Campos estruturados conforme Field of Research (FoR Code - ANZSRC 2020).

3. Organização e Armazenamento

Estrutura de organização dos arquivos compactados(.zip):

- **/DATA.zip (3.155.077 kb)**
 - **/DRV_01_AltmetricResearcOutput**
 - DAD_00_Tabela_ForCode.xlsx
 - DAD_02_CalcularMostraArtigos.xlsx
 - DAD_03_ArtigosSelecionadosRandom.xlsx
 - DAD_04_PostsColetados.xlsx
 - DAD_05_PostsCruzamentoDatasets.xlsx
 - DAD_06_ContasPorGrandeArea.xlsx
 - DAD_07_PostsIdentificacaoPersonas.xlsx
 - DAD_08_PostsCombinados.xlsx
 - DAD_09_AmostraContas_Calcular.xlsx
 - DAD_10_AmostraContasResult.xlsx
 - DAD_11_ContasSelecionadas.xlsx
 - DAD_12_AmostraFinal_Contas.xlsx
 - DAD_13_AmostraFinal_Posts.xlsx
 - **/SCRIPTS.zip (11kb)**

- ALG_01_CalcularAmostraArtigos.py
- ALG_02_SelecionarRandomArtigos.py
- ALG_03_ColetarPosts.py
- ALG_04_CruzarDatasets.py
- ALG_05_DetectarPersonas.py
- ALG_06_CalclarAmostraContas.py
- ALG_07_SelecionarRandomContas.py
- ALG_08_CriarDatasetFinal.py
- **/MODELO.zip (328.927)**
 - ALG_09_Pipeline_XGBoost.py
 - MDL_Top46Features.pkl
 - MDL_ModeloXgboost.pk
 - MDL_Scaler.pkl
 - MDL_Imputer.pkl
 - MDL_ThresholdOtimo.pkl
 - DAD_14_DadosFeaturesTreinamento.xlsx
 - DAD_15_DatasetClassificadoManual.xlsx
 - DAD_16_TodosPostsTreinamento.xlsx
 - DAD_17_ContasIntegralizado.xlsx
 - DAD_18_PostsIntegralizado.xlsx
 - DAD_19_DadosFeaturesNovasContas.xlsx
 - DAD_20_ClassificacaoNovasContas.xlsx
 - PDF_RelatorioModeloXgboost.pdf

Procedimentos de controle de versão: Controle via ZENODO.

Local de armazenamento primário: <https://zenodo.org/records/17050653>.

4. Compartilhamento e Acesso

Estratégia de compartilhamento de dados:

- Dados derivados e scripts serão disponibilizados em <https://zenodo.org/records/17050653>.

Repositórios pretendidos para disponibilização: Zenodo.

Licença de uso dos dados: Creative Commons CC-BY 4.0 para scripts.

6. Responsabilidades

Responsável principal pela gestão dos dados: Danielle Pontes.

7. Recursos Necessários

Infraestrutura de armazenamento: ZENODO.

Ferramentas de gerenciamento de dados: Python, Excel.

8. Processo de Utilização dos Algoritmos

9.1 Download da Tabela Field of Research (FoRCode)

- **Arquivo:** DAD_00_Tabela_ForCode.xlsx

9.2 Download dos Artigos Científicos por Grupo FoRCode

- **Entrada:** IDs dos 212 grupos FoR
- **Saída:** ARTIGOS-FORCODE/[código]_Altmetric.csv

9.3 Cálculo Amostral de Número de Artigos

- **Algoritmo:** ALG_01_CalcularAmostraArtigos.py
- **Entrada:** DAD_02_CalcularAmostraArtigos.xlsx
- **Saída:** Tamanhos amostrais

9.4 Seleção Randômica de Artigos

- **Algoritmo:** ALG_02_SelecionarRandomArtigos.py
- **Entrada:** Amostras
- **Saída:** DAD_03_ArtigosSelecionadosRandom.xlsx

9.5 Coleta dos *Posts* de Cada Artigo da Seleção Randômica

- **Algoritmo:** ALG_03_ColetarPosts.py
- **Entrada:** DOIs
- **Saída:** DAD_04_PostsColetados

9.6 Fusão dos *Datasets*

- **Algoritmo:** ALG_04_CruzarDatasets.py
- **Entrada:** Posts coletados + bases externas
- **Saída:** DAD_05_PostsCruzamentoDatasets.xlsx

9.7 Cálculo Amostral de Número de Contas por Divisão FoRCode

- **Algoritmo:** ALG_06_CalcularAmostraContas.py
- **Entrada:** DAD_06_ContasPorGrandeArea.xlsx
- **Saída:** DAD_10_AmostraContas_Resultados.xlsx

9.8 Seleção Randômica de Contas por Divisão FoRCode

- **Algoritmos:** ALG_07_SelecionarRandomContas.py

e

ALG_08_CriarDatasetFinal.py

- **Entrada:** Contas por divisão
- **Saída:** DAD_13_AmostraFinal_Posts.xlsx

9.9 Construção e Classificação do *Dataset* de Treinamento

- **Algoritmos:** ALG_05_DetectarPersonas.py e cruzamento com bases externas
- **Saída:** *Dataset* de treinamento rotulado (humano/*bot*)

9.10 Desenvolvimento e Otimização do Modelo XGBoost

- **Algoritmo:** ALG_09_Pipeline_XGBoost.py
- **Entrada:** df_final contendo *features* extraídas
- **Saída:** Modelo treinado para classificação de contas

Por fim, o quadro 9 apresenta a sistematização dos passos metodológicos da pesquisa, relacionando cada etapa do processo com os respectivos arquivos de dados gerados e os algoritmos implementados. A estrutura evidencia a sequência lógica desde o download e organização inicial dos dados dos artigos científicos, passando pelo cálculo amostral, seleção randômica de artigos e contas, coleta e classificação preliminar dos *posts*, até a criação do *dataset* final. Além disso, contempla a fase de desenvolvimento do modelo preditivo (XGBoost), incluindo arquivos de treinamento, bases integralizadas, scripts de processamento e modelos finais. Dessa forma, o quadro funciona como um mapa de reproduzibilidade, permitindo rastreabilidade e transparência na condução da pesquisa.

Quadro 9 – Correspondência entre Passos Metodológicos, Bases de Dados e Algoritmos

Passos Metodológicos	Arquivos de Dados		Algoritmos
	Nome	Descrição	
II.a.1 Download dos artigos científicos (ForCode no Altmetric)	DAD_00_Tabela_ForCode.xlsx	Tabela com todos os códigos de grupos e divisões	—
	DRV_01_AltmetricResearcOutput.zip	Arquivo zipado contendo as 144 tabelas <i>research altmetric</i> output de cada grupo/divisão com os dados dos artigos	—
II.a.2 Cálculo amostral de artigos	DAD_02_CalculoAmostraArtigos.xlsx	Tabela com a relação dos 144 grupos e o respectivo resultado do cálculo do tamanho da amostra para cada grupo	ALG_01_CalcularAmostraArtigos.py
II.a.2 Seleção Randômica de artigos	DAD_03_ArtigosSelecionadosRandom.xlsx	Relação dos 93.529 artigos selecionados randomicamente	ALG_02_SelecionarRandomArtigos.py
II.b.1 Coleta dos posts dos artigos selecionados	DAD_04_PostsColetados.xlsx	Relação dos 758288 posts realizados sobre os 93529 artigos	ALG_03_ColetarPosts.py
II.b.2 Classificação preliminar com base em datasets da literatura	DAD_05_PostsCruzamentoDatasets.xlsx	Relação dos 758288 posts - com o resultado do cruzamento dos datasets e identificação das personas	ALG_04_CruzarDatasets.py

	DAD_06_ContasPorGrandeArea.xlsx	Relação das 464787 contas que postaram sobre os artigos da amostra com a sua divisão	
	DAD_07_PostsIdentificacaoPersonas.xlsx	<i>Posts</i> com identificação de personas	ALG_05_DetectarPersonas.py
II.b.3 Cálculo amostral do número de contas	DAD_08_PostsCombinados.xlsx	Relação dos 758288 posts com código do grupo do artigo que o post comenta	ALG_06_CalclarAmostraContas.py
	DAD_09_AmostraContas_Calculo.xlsx	Tabela com a quantidade de contas por divisão	
	DAD_10_AmostraContasResult.xlsx	Tabela com o cálculo do tamanho da amostra de cada divisão	
II.b.3 Seleção Randômica de contas	DAD_11_ContasSelecionadas.xlsx	Resultado da seleção randômica das contas	ALG_07_SelecionarRandomContas.py
II.b.3 Criação do dataset final de contas e posts	DAD_12_AmostraFinal_Contas.xlsx	Relação das contas da amostra final	ALG_08_CriarDatasetFinal.py
	DAD_13_AmostraFinal_Posts.xlsx	Relação de posts com coluna selecionados (onde 1 = selecao da amostra de artigos)	
II.c Processo de Desenvolvimento do Modelo (XGBoost)	DAD_14_DadosFeaturesTreinamento.xlsx	Dados das conta do usuário com dezenas de variáveis representando suas características(FEATURES).	ALG_09_Pipeline_XGBoost.py

	DAD_15_DatasetClassificadoManual.xlsx	DFCONTAS - <i>Dataset</i> de treinamento classificado manualmente. Informações consolidadas sobre as contas que fazem parte do <i>dataset</i> .	
	DAD_16_TodosPostsTreinamento.xlsx	DFPOSTS - Posts realizados pelas contas em DFCONTAS, sobre os artigos da amostra.	
	MDL_Top46Features.pkl	Modelo	
	MDL_ModeloXgboost.pk	Modelo	
	MDL_Scaler.pkl	Modelo	
	MDL_Imputer.pkl	Modelo	
	MDL_ThresholdOtimo.pkl	Modelo	
	PDF_RelatorioModeloXgboost.pdf	Pdf	
	DAD_17_ContasIntegralizado.xlsx	Base com dados das novas contas	
	DAD_18_PostsIntegralizado.xlsx	Base contendo os <i>posts</i> feitos por essas novas contas	
	DAD_19_DadosFeaturesNovasContas.xlsx	Características textuais, métricas temporais, estruturais e sintáticas, além de medidas linguísticas, de legibilidade e de similaridade entre nome e identificador da conta	

	DAD_20_ClassificacaoNovasContas.xlsx	Resultados finais — com probabilidade prevista, classe binária e rótulo textual (“Humano” ou “Bot”)	
--	--------------------------------------	---	--
