



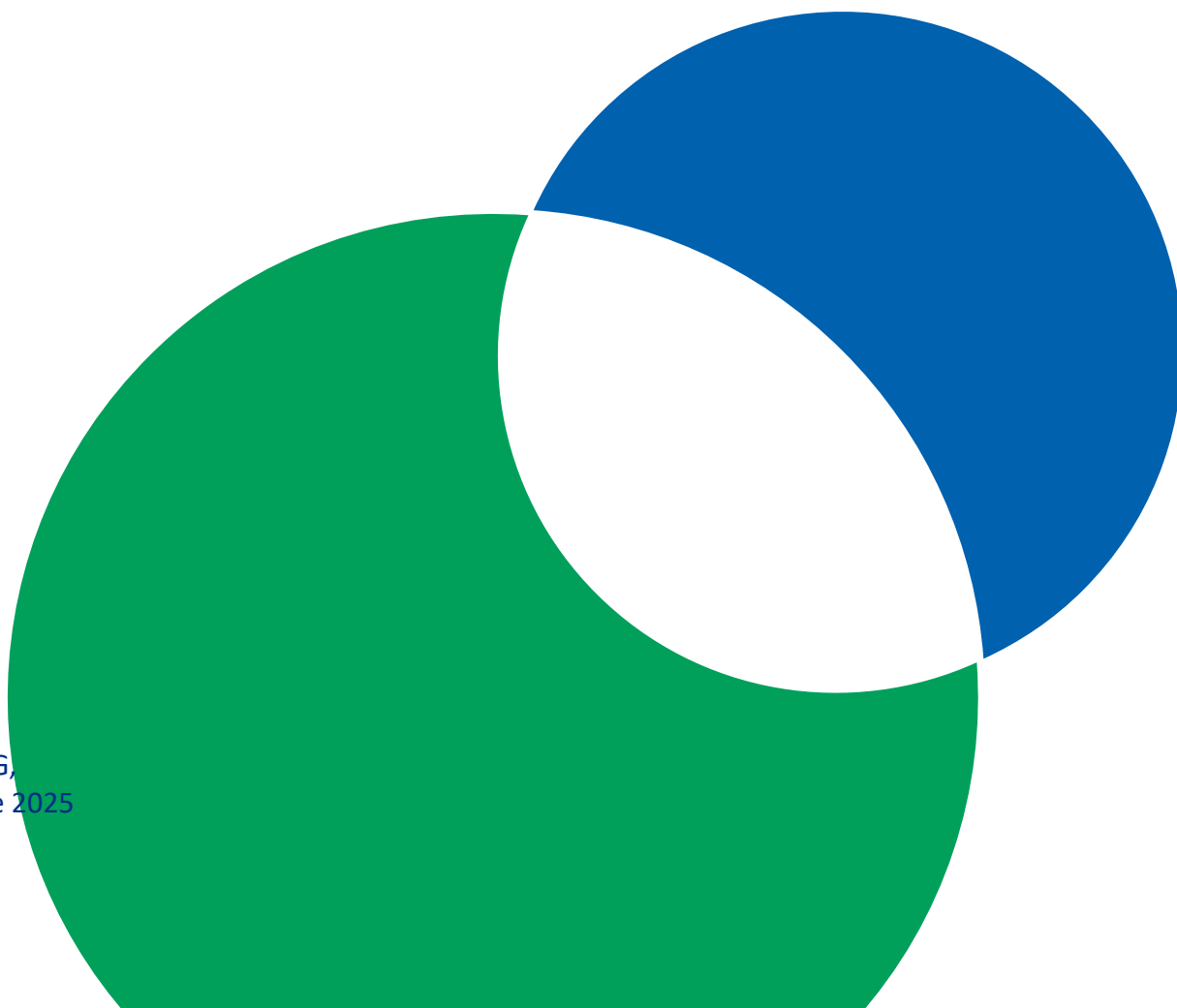
UNIFEI
UNIVERSIDADE FEDERAL DE ITAJUBÁ

UNIVERSIDADE FEDERAL DE ITAJUBÁ
Programa de Pós-Graduação em Desenvolvimento,
Tecnologias e Sociedade

***MACHINE LEARNING PARA ANÁLISE EXPLORATÓRIA E
EVIDENCIAÇÃO EMPÍRICA DA RELAÇÃO ENTRE SOLIDÃO, POLÍTICA E
MÍDIAS SOCIAIS NO BRASIL***

Jenaina de Fatima dos Santos

Itajubá- MG,
Outubro de 2025



UNIVERSIDADE FEDERAL DE ITAJUBÁ
INSTITUTO DE ENGENHARIA DE PRODUÇÃO E GESTÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM DESENVOLVIMENTO, TECNOLOGIAS E
SOCIEDADE

Jenaina de Fatima dos Santos

***MACHINE LEARNING* PARA ANÁLISE EXPLORATÓRIA E**
EVIDENCIAÇÃO EMPÍRICA DA RELAÇÃO ENTRE SOLIDÃO,
POLÍTICA E MÍDIAS SOCIAIS NO BRASIL

Dissertação apresentada ao Programa de Pós-Graduação em Desenvolvimento, Tecnologias e Sociedade da Universidade Federal de Itajubá, como parte dos requisitos para obtenção do título de Mestre.

Área de concentração: Desenvolvimento, Tecnologias e Sociedade

Linha de Pesquisa: Desenvolvimento e Tecnologia

Orientador: Prof. Dr. Jesús Antonio García Sánchez

Coorientador: Prof. Dr. Adilson Da Silva Mello

Itajubá - MG
Outubro de 2025

Copyright © 2025 – Jenaina de Fatima dos Santos, Programa de Pós-Graduação em Desenvolvimento, Tecnologias e Sociedade.

Esta dissertação é um trabalho original, escrito exclusivamente para este fim, e todos os autores cujos estudos e publicações contribuíram para o trabalho foram devidamente citados. A reprodução parcial é permitida com reconhecimento do autor e referência ao grau, ano letivo, instituição e data da defesa pública.

contato: jenainasantos60@gmail.com

Como citar: SANTOS, J. ***Machine learning para análise exploratória e evidênciação empírica da relação entre solidão, política e mídias sociais no Brasil*** Dissertação de mestrado. Itajubá, Programa de pós-graduação em Desenvolvimento, Tecnologias e Sociedade, Universidade Federal de Itajubá, 2025.

Ficha catalográfica elaborada pela Biblioteca Mauá
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

S237m Santos, Jenaina
MACHINE LEARNING PARA ANÁLISE EXPLORATÓRIA E
EVIDENCIAÇÃO EMPÍRICA DA RELAÇÃO ENTRE SOLIDÃO,
POLÍTICA E MÍDIAS SOCIAIS NO BRASIL / Jenaina
Santos; orientador Jesús Sánchez ; coorientador
Adilson Mello. -- Itajubá, 2025.
150 p.

Dissertação (Mestrado - Programa de Pós-Graduação
em Desenvolvimento, Tecnologias e Sociedade) --
Universidade Federal de Itajubá, 2025.

1. Solidão. 2. Política. 3. Mídias Sociais. 4.
Machine Learning . 5. Severidade da solidão. I.
Sánchez , Jesús, orient. II. Mello, Adilson,
coorient. III. Título.

FOLHA DA APROVAÇÃO

Jenaina de Fatima dos Santos

***MACHINE LEARNING* PARA ANÁLISE EXPLORATÓRIA E EVIDÊNCIAÇÃO EMPÍRICA DA RELAÇÃO ENTRE SOLIDÃO, POLÍTICA E MÍDIAS SOCIAIS NO BRASIL**

Dissertação apresentada ao programa de Pós-Graduação em Desenvolvimento, Tecnologias e Sociedade da UNIFEI, como parte dos requisitos para obtenção do título de mestre.

Área de concentração: Desenvolvimento, Tecnologias e Sociedade

Linha pesquisa: Desenvolvimento e Tecnologias

Aprovado em:

BANCA EXAMINADORA

Prof.^a Dra. Lina Maria Garcés Rodriguez
Avaliadora externa

Prof.^a Dra. Denise Pereira De Alcantara Ferraz
Avaliadora interna

Prof. Dr. Jesús Antonio Garcia Sanchez
Orientador

Itajubá - MG
Outubro de 2025

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL GENERATIVA

O(s) autor(es) autor(es) do manuscrito acima identificado, declara(m) que:

1. Uso de Ferramentas de Inteligência Artificial Generativa:

Em caso afirmativo, responda e/ou marque as caixas correspondentes referentes ao uso de Inteligência Artificial Generativa durante o desenvolvimento desta pesquisa:

Ferramenta(s): *Deepseek e ChatGPT*

Propósito(s):

- Revisão de linguagem:** melhoria da gramática, ortografia, clareza e estilo do texto.
- Geração de texto:** criação de rascunhos, sugestões ou partes específicas do conteúdo, devidamente revisadas pelos autores.
- Análise de dados:** auxílio para processamento e na interpretação inicial de conjuntos de dados.
- Visualização de dados ou imagens:** auxílio na geração ou refinamento de figuras, gráficos ou ilustrações.
- Sugestões metodológicas:** fornecimento de recomendações ou ideias para aprimorar abordagens de pesquisa.
- Outros (especificar):** auxílio para revisão, interpretação e correção das fórmulas matemáticas e gráficos utilizadas na etapa de processamento, devidamente revisada pelos autores.

2. Autoria e Supervisão Humana:

- A utilização de ferramentas de inteligência artificial foi exclusivamente para os propósitos acima declarados. Toda supervisão, validação e autoria intelectual do manuscrito são de responsabilidade dos autores humanos.

3. Conformidade com Princípios Éticos:

- O uso das ferramentas de IA seguiu práticas éticas e não incluiu atividades como fabricação de dados, plágio, manipulação de figuras ou qualquer prática que comprometa a integridade científica.

4. Transparência:

- Declaramos que as informações acima são completas e verdadeiras. Caso seja identificada qualquer inconsistência ou uso inadequado, aceitamos que o manuscrito seja submetido a investigação e, se necessário, rejeitado ou retratado.

Assinatura do(s) autor(es):

Assinatura do(s) orientador(es):

À minha família, especialmente aos pilares desta minha jornada na pesquisa e ciência: à minha mãe, Eloiza dos Santos, e ao meu sobrinho, Isaac dos Santos.

E a todos e todas que, em um mundo hiperconectado, experienciam a solidão social como um exílio adoecedor.

AGRADECIMENTOS

A Deus e a Nossa Senhora Aparecida, por me sustentarem com fé e coragem nesta jornada como pesquisadora e no sonho de me tornar cientista.

A mim, pela minha dedicação exclusiva, resiliência e persistência diante das complexidades de desenvolver uma pesquisa complexa, conduzindo-a da melhor forma possível durante estes anos.

À minha família, especialmente à minha mãe, Eloiza dos Santos; às minhas irmãs Patrícia, Maria Angélica, Elisângela, Josiane, Conceição, Elenize, Josélia, Juciara e Israely dos Santos; ao meu irmão, Isaias dos Santos; ao meu sobrinho, Isaac dos Santos; e às minhas sobrinhas Sthefanny e Jaqueline, pelo acolhimento e apoio incondicional em todos os momentos desta longa caminhada.

Aos meus amigos de longa data e aos que fiz no mestrado, em especial a Milena Cabral, por entender as minhas ausências durante este tempo. E ao Matheus Mendonça dos Reis, meu apoio emocional e melhor companhia, por me ajudar nos momentos mais difíceis, pelos trabalhos e momentos que compartilhamos.

Ao meu orientador, Prof. Dr. Jesús Sánchez, pela parceria, paciência, ensinamentos, contribuições e pelo incentivo constante ao desenvolvimento desta pesquisa.

Aos docentes que contribuíram significativamente para este trabalho: meu coorientador, Prof. Dr. Adilson Melo, por me guiar na construção teórica e na busca por caminhos mais assertivos para o tema. E ao Prof. Dr. Leandro de Jesus, por abrir espaço em sua disciplina e em seu laboratório para que eu aprendesse sobre a aplicação de modelos de IA.

À minha banca examinadora, Prof.^a Dra. Lina Rodriguez e Prof.^a Dra. Denise Ferraz, pelas valiosas contribuições que conduziram e enriqueceram esta pesquisa.

Aos participantes e colaboradores que contribuíram para a formação do banco de dados empírico, seja respondendo ou divulgando o formulário. Vocês foram peça fundamental na construção deste estudo.

Aos pesquisadores do Laboratório TIES (Pesquisa em Tecnologias Integradas e Engenharia Sustentável), em especial à Prof.^a Dra. Carolina Quintero Rodriguez, pelos conselhos e pelo espaço de estudo que também foram essenciais.

Aos meus colegas de pesquisa, com destaque para a turma de mestrado do DTECS em 2023 pelas ricas discussões e trocas que enriqueceram minha reflexão e fortaleceram minha pesquisa, em especial a Tamara Silva Rios, Ana Paula Ribeiro Fortes, Gabriel José Dias

Ferreira, Klercia Cristiane Costa Fonseca, João Pedro Lopes Dias Ribeiro e Joyce Oliveira Ribeiro. E às pesquisadoras que conheci nesse tempo e foram fundamentais para o meu amadurecimento na pesquisa: à Dra. Samanta Borges Pereira, pelo apoio e pelos conselhos que nortearam meu primeiro ano de jornada acadêmica; e a Helena de Oliveira Vítório, cujo trabalho inspirador na representação estudantil e no ativismo na pós-graduação reforçou a importância do engajamento coletivo. E a Isabel Pacheco pelas conversas, críticas e preocupações.

Às psicólogas Dra. Mariella Brighenti, Prof.^a Ma. Taciane Marques Castelo Branco e Prof.^a Ma. Kethulin Franco, que me incentivaram na pesquisa e no mestrado ainda durante minha graduação em Psicologia.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro indispensável à realização deste trabalho (Código de Financiamento 001).

A todos os docentes do DTECS que, ao longo desses anos, me ensinaram a tecer a interdisciplinaridade da minha pesquisa, cada qual com seu campo, seu método e sua complexidade.

Ao Programa de Pós-Graduação em Desenvolvimento, Tecnologias e Sociedade (DTECS) por contribuir com a minha formação como pesquisadora e futura cientista.

À Associação de Pós-Graduandos da UNIFEI (APG UNIFEI), em especial à gestão Vanguarda Científica, da qual fiz parte, pelo espaço de diálogo e pela mobilização constante em pautas políticas essenciais à nossa trajetória acadêmica.

Por fim, a Universidade Federal de Itajubá (UNIFEI), por acolher e sustentar estes espaços essenciais para pesquisa, ciência e minha trajetória como pesquisadora.

“Proteja-me do que eu quero.” - Jenny Holzer

RESUMO

A solidão na contemporaneidade emerge como um problema social complexo e crise global de saúde pública em diversos países. No Brasil, ainda há escassez de estudos sobre os impactos da solidão, seus fatores correlatos e as implicações demográficas e econômicas, em grande parte devido à literatura limitada e à padronização recente de instrumentos, como a Escala da solidão UCLA, para avaliação desse construto. Este estudo adota uma abordagem interdisciplinar, integrando métodos de IA, especificamente técnicas estatísticas computacionais, para análise de dados psicológicos, com o objetivo de explorar e evidenciar empiricamente as associações entre comportamentos relacionados a esse problema pouco investigado. Através de treinamento de algoritmos de *machine learning*, buscou-se analisar os padrões complexos da associação entre os níveis de solidão, comportamento político e uso de mídias sociais no contexto das eleições presidenciais brasileiras de 2022. Foram aplicados instrumentos de autorrelato de forma online, como a Escala Brasileira de Solidão UCLA e questionários adaptados, em uma amostra de 176 participantes, majoritariamente residentes do Sul de Minas Gerais. Os algoritmos foram implementados em *Python* no ambiente *VS Code*. Os resultados principais revelaram por meio do *K-means* a alta dimensionalidade dos dados limitou o agrupamento completo durante o pré-treinamento, mas revelou *clusters* que destacam a severidade da solidão por faixa etária, com maior prevalência em jovens adultos. O *KNN* não evidenciou associações significativas entre as variáveis analisadas, encontrando uma relação muito baixa apenas entre severidade da solidão e uso de mídias sociais. Já o *DBSCAN* conseguiu agrupar mais claramente os *clusters* de tal forma que foi possível encontrar relações entre algumas das severidades das três variáveis, especialmente quando tais severidades apresentavam valores baixos. Os resultados indicaram poucos casos de alta severidade para solidão e comportamento político na amostra. Não houve nenhum caso de severidade extrema associado ao uso de mídias sociais. Estes últimos resultados podem estar associados à limitação da amostra e dos dados coletados. Concluiu-se que este estudo contribui para o esclarecimento de hipóteses sobre as relações entre os construtos analisados, além de proporcionar uma metodologia que permita, eventualmente, avaliar a relação entre três fenômenos sociais emergentes. No contexto investigado, os dados revelaram associações muito fracas entre solidão, comportamento político e uso de mídias sociais. Esses resultados sugerem a necessidade de mais estudos, principalmente ajustes no banco de dados para melhor representatividade, abrangência e refinamento dos modelos implementados, considerando a complexidade dos dados e suas particularidades.

Palavras-chave: Solidão; Política; Mídias sociais; *Machine learning*; Análise exploratória dos dados; Severidade da solidão.

ABSTRACT

Loneliness in contemporary society has emerged as a complex social problem. It is increasingly recognized as a global public health crisis due to its significant impact on populations across various countries. In Brazil, there is still a scarcity of studies on the impacts of loneliness, its correlated factors, and its demographic and economic implications, largely due to limited literature and the recent standardization of assessment tools, such as the UCLA Loneliness Scale, for evaluating this construct. This study integrates artificial intelligence methods, specifically computational statistical techniques, for the analysis of psychological data. The aim was to provide empirical evidence for the associations between behaviors related to this under-investigated issue. By training machine learning algorithms, it was possible to analyze the complex patterns of association between levels of loneliness, political behavior, and social media use within the context of the 2022 Brazilian presidential elections. Online self-report instruments were administered to a sample of 176 participants, predominantly residents of Southern Minas Gerais. Algorithms were implemented in Python within the VS Code environment. Analysis with K-means indicated that the high dimensionality of the data prevented complete clustering during pre-training. However, the analysis revealed clusters highlighting the severity of loneliness by age group, with a higher prevalence among young adults. KNN did not show significant associations between the analyzed variables, finding only a very weak relationship between loneliness severity and social media use. In contrast, DBSCAN managed to cluster the data more clearly, making it possible to identify relationships between some of the severity levels of the three variables, especially when these severity levels were low. The results indicated few cases of high severity for loneliness and political behavior in the sample. No cases of extreme severity associated with social media use were found. These latter results may be associated with the limitations of the sample and the collected data. In conclusion, this study contributes to clarifying hypotheses about the relationships between the analyzed constructs and provides a methodology that may eventually allow for the assessment of the relationship between three emerging social phenomena. Within the investigated context, the data revealed very weak associations between loneliness, political behavior, and social media use. These findings suggest the need for further studies, primarily adjustments to the database for better representativeness and scope, as well as refinements to the implemented models, considering the complexity and particularities of the data.

Keywords: *Loneliness; Politics; Social medias; Machine learning; Exploratory data analysis; Loneliness severity.*

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1- Gráfico sobre a contextualização sobre a relação entre solidão, política e mídias sociais..... | 24 |
| Figura 2 - Gráfico do número de publicações por países sobre a relação entre solidão, política e mídias sociais..... | 24 |
| Figura 3- Mapa do número de ocorrência e força de ligação das variáveis..... | 26 |
| Figura 4- Gráfico da contextualização histórica de publicações sobre a relação entre solidão e política..... | 29 |
| Figura 5- Gráfico dos países que mais publicaram sobre a relação entre solidão e política..... | 30 |
| Figura 6- Mapa do número de ocorrência e força de ligação das variáveis..... | 31 |
| Figura 7- Gráfico do histórico de publicações sobre solidão e mídias sociais..... | 33 |
| Figura 8- Gráfico do número de estudos publicados sobre solidão e mídias sociais..... | 34 |
| Figura 9- Mapa do número de ocorrência e força de ligação entre solidão e mídias sociais..... | 35 |
| Figura 10- Gráfico de tendência temporal da relação entre política e mídias sociais no contexto global..... | 38 |
| Figura 11- Gráfico análise temporal-geográfica dos estudos que relacionam política e mídias sociais | 39 |
| Figura 12- Mapa do número de ocorrência e força de ligação entre as variáveis política e mídias sociais..... | 40 |
| Figura 13- Diferentes dimensões da solidão..... | 50 |
| Figura 14- Efeitos da solidão na cognição social..... | 51 |
| Figura 15- Características da desinformação e má informação | 57 |
| Figura 16- Nível de confiança na vacina posição..... | 58 |
| Figura 17- Neurônio artificial..... | 64 |
| Figura 18- Neurônio humano..... | 64 |
| Figura 19- Múltiplas camadas em uma rede neural biológica (córtex humano) | 64 |
| Figura 20- Neurônio artificial de Frank Rosenblatt..... | 64 |
| Figura 21- Mineração dos dados..... | 65 |
| Figura 22- Fluxograma das etapas gerais de aplicação do ML..... | 69 |
| Figura 23- Etapas do estudo sobre o panorama geral da literatura..... | 70 |
| Figura 24- Fluxograma das etapas da implementação do K-means..... | 85 |
| Figura 25- Fluxograma das etapas de implementação do KNN..... | 88 |

| | |
|---|-----|
| Figura 26- Fluxograma das etapas de implementação do DBSCAN..... | 90 |
| Figura 27- Gráfico do método do Cotovelo..... | 94 |
| Figura 28- Gráfico PCA do <i>K-means</i> | 94 |
| Figura 29- Gráfico sobre a importância das <i>features</i> no <i>K-means</i> | 95 |
| Figura 30- Gráfico de distribuição da severidade da solidão por cluster..... | 96 |
| Figura 31- Gráfico agrupamento da intensidade da solidão por faixa etária..... | 97 |
| Figura 32- Gráfico da importância das variáveis por severidade no <i>KNN</i> | 98 |
| Figura 33- Matriz de relação entre severidade da solidão e comportamento político..... | 99 |
| Figura 34- Matriz de relação entre severidade da solidão e uso de mídias..... | 100 |
| Figura 35- Matriz de relação entre comportamento político e uso de mídias..... | 101 |
| Figura 36- Gráfico 3D para a relação entre solidão, política e mídias sociais..... | 101 |
| Figura 37- Matriz de correlação entre as <i>features</i> do <i>KNN</i> | 105 |
| Figura 38- Relação entre as variáveis..... | 106 |
| Figura 39- Agrupamento dos participantes em clusters..... | 108 |

LISTA DE QUADROS

| | |
|--|----|
| Quadro 1- Hipóteses testáveis..... | 19 |
| Quadro 2- Artigo que explora a relação entre solidão, política e mídias sociais no Brasil..... | 27 |
| Quadro 3- Artigo que explora a relação entre solidão e política publicados no Brasil..... | 32 |
| Quadro 4- Artigos brasileiros que exploram a relação entre solidão e mídias sociais..... | 37 |
| Quadro 5- Artigos brasileiros que exploram a relação entre política e mídias sociais no Brasil..... | 42 |
| Quadro 6- Classificação dos modelos de agrupamento ML aplicados neste estudo..... | 66 |
| Quadro 7- organização das dimensões das questões sobre comportamento político..... | 81 |
| Quadro 8- organização das dimensões das questões sobre uso de mídias sociais..... | 81 |

LISTA DE TABELAS

| | |
|---|-----|
| Tabela 1- Variáveis de maior interesse em estudos sobre solidão, política e mídias sociais... | 26 |
| Tabela 2- Variáveis de maior interesse na busca geral 2..... | 32 |
| Tabela 3- Variáveis de maior interesse na busca 3..... | 36 |
| Tabela 4- Variáveis de maior interesse na busca 4..... | 41 |
| Tabela 5- Resultados da busca ampla do panorama geral da literatura..... | 72 |
| Tabela 6- Resultados das buscas delimitadas sobre o cenário brasileiro nos últimos cinco anos..... | 74 |
| Tabela 7- Variáveis sociodemográfica..... | 76 |
| Tabela 8- Variáveis socioeconômicas..... | 77 |
| Tabela 9- Perfil dos participantes que utilizaram mídias sociais..... | 78 |
| Tabela 10- Variável perfil político dos participantes..... | 78 |
| Tabela 11- Relação entre Severidade da Solidão e comportamento político | 104 |
| Tabela 12- Relação entre Severidade da Solidão e uso de mídias sociais..... | 106 |
| Tabela 13- Relação entre Severidade de comportamento político e Uso de Mídias sociais. | 107 |

SUMÁRIO

| | |
|---|-----|
| 1 INTRODUÇÃO | 17 |
| 1.1. Problema..... | 18 |
| 1.2. Hipóteses testáveis..... | 19 |
| 1.3. Objetivos..... | 19 |
| 1.4. Justificativa..... | 20 |
| 1.5. Estrutura do texto..... | 21 |
| 2 PANORAMA GERAL DA LITERATURA | 23 |
| 2.1. Contextualização sobre a relação entre solidão, política e mídias sociais..... | 23 |
| 2.2. Contextualização sobre a relação entre solidão e política..... | 29 |
| 2.3. Contextualização sobre a relação entre solidão e mídias sociais..... | 33 |
| 2.4. Contextualização sobre a relação entre política e mídias sociais..... | 38 |
| 3 FUNDAMENTAÇÃO TEÓRICA | 46 |
| 3.1. A natureza multidimensional da solidão: discussão histórica contextual..... | 46 |
| 3.2. A natureza da solidão na Psicologia e suas interfaces com a saúde mental..... | 48 |
| 3.3. Evidências sobre a relação entre solidão, política e mídias sociais..... | 51 |
| 3.4. Evidências sobre a relação entre solidão e política no Brasil..... | 53 |
| 3.5. Evidências sobre a relação entre solidão e mídias sociais no Brasil | 54 |
| 3.6. Evidências sobre a relação entre política e mídias sociais no Brasil..... | 55 |
| 4 METODOLOGIA | 62 |
| 4.1. Interdisciplinaridade do método..... | 62 |
| 4.2. Classificação do método de IA..... | 65 |
| 4.3. Materiais e método..... | 69 |
| 5 ANÁLISE E DISCUSSÃO DOS RESULTADOS | 93 |
| 5.1. Análise e discussão dos resultados da saída do <i>K-means</i> | 93 |
| 5.2. Análise e discussão resultados da saída do <i>DBSCAN</i> | 97 |
| 5.3. Análises estatística descritivas de suporte | 104 |
| 5.4. Discussão analítica dos resultados..... | 108 |
| 5.5. Limitações do estudo..... | 111 |
| 6 CONSIDERAÇÕES FINAIS | 112 |
| REFERÊNCIAS BIBLIOGRÁFICAS | 114 |
| APÊNDICE | 124 |
| ANEXO | 145 |

CAPÍTULO 1

1. INTRODUÇÃO

A solidão é um fenômeno historicamente recorrente, cujo sentido varia conforme o contexto social, grupos e a época. Há tempos assumiu um caráter ambíguo entre sua essencialidade ou causa de temor (Minois, 2019). Na contemporaneidade, embora vivenciada individualmente, reflete problemas coletivos, configurando-se em uma crise global (O'Rourke, 2024). Está sendo evidenciado como problema de saúde pública emergente, tendo seus riscos reconhecidos pela Organização Mundial da Saúde (OMS) (Braga *et al.*, 2022).

Em países centrais, como EUA, Reino Unido, Japão, a solidão tem sido reconhecida pelas suas implicações na saúde pública, sendo até mesmo entendida como uma “epidemia”. Isso tem gerado implicações democráticas nestes países, pois este fenômeno tem criado uma tendência de coesão social. Deste modo, buscam desenvolver políticas públicas para combatê-la (Hertz, 2021; Hawkley; Cacioppo, 2013).

Houve um aumento significativo no número de pessoas afetadas pela solidão nos últimos anos, tornando-a onerosa em diversos países (Barreto *et al.*, 2024). Os agravantes da solidão decorrem das remodelações da globalização, que promovem tendências ao individualismo e enfraquecem os laços comunitários (Hertz, 2021). Em um sistema de retroalimentação, um fator central apontado tanto como causa quanto agravante desse fenômeno é o modelo de interação mediado pelas novas tecnologias de comunicação e informação, como o uso crescente de mídias sociais (Hertz, 2021; Peterson *et al.*, 2025; Minois, 2019). Conforme Hertz (2021) e Han (2015), a arquitetura dessas plataformas, seguindo um viés neoliberal, fomenta uma "conexão social" paradoxalmente solitária, o que gera fragmentação política (Hertz (2021).

Trata-se de uma discrepância cognitiva entre a percepção dos níveis de interações e a deficiência nas relações sociais. Existem diferentes manifestações, dentre elas, as comportamentais antecedentes (como a falta de habilidades sociais) quanto as que acompanham essa experiência (como a falta de assertividade). Neste sentido, há uma preocupação com o estado crônico da solidão, e com os níveis considerados “normais”, pois ambos podem afetar a forma com o sujeito percebe suas conexões. (Prohaska *et al.*, 1981).

Essa experiência subjetiva pode ser vivenciada em qualquer fase da vida, ocorre sobretudo onde as situações sociais e a manutenção das relações são mais afetadas (Weeks,

1994; Berguno *et al.*, 2004). Contudo, na fase jovem-adulta, essa experiência também não é incomum. Apesar de ser marcada por maiores períodos de socialização e necessidades sociais (Prohaska *et al.*, 1981), essa fase se caracteriza pela maior transitoriedade das relações, o que conseqüentemente gera insatisfação com a realização social, aumentando a sensação de solidão nesse grupo (Weeks, 1994).

Nesse contexto, a solidão, ao adquirir relevância psicológica como causa de problemas interpessoais (Hertz, 2021; Barroso *et al.*, 2016), está sendo instrumentalizada por governos populistas e atores neoliberais, afetando grupos vulneráveis e contribuindo para segregação social (Hetz, 2021; Peterson *et al.*, 2025). Pesquisas recentes têm analisado a solidão a partir de seus indicadores de privação social, tanto objetiva quanto subjetiva, que afetam não apenas a saúde dos indivíduos, mas também seu engajamento sociopolítico. Um debate emergente na literatura investiga a possível associação entre altos níveis de solidão social e a adoção de posturas políticas extremistas (Langenkamp, 2025).

Essa associação entre solidão e política foi analisada em países como Alemanha, Áustria, Holanda, Croácia, Dinamarca, França, Hungria, Suécia e Suíça, evidenciando uma correlação significativa entre solidão e preferência por grupos políticos radicais e que causam segregação social (Langenkamp; Bienstock, 2022; Petherick *et al.*, 2023). Langenkamp (2025) busca comprovar empiricamente essa relação, examinando comportamentos como: o apoio eleitoral a partidos populistas de extrema-direita e processos de radicalização individual, com ênfase no crescimento do extremismo antifeminista (Langenkamp, 2025).

No Brasil ainda há poucos estudos que apontam a relação entre três fenômenos: solidão intensa, comportamento político engajado e uso de mídias sociais. Deste modo, através desta pesquisa, analisamos a associação entre estas variáveis, tendo como contexto empírico, o segundo turno das eleições para presidente do Brasil em 2022. Para investigar esta associação, propõe-se uma abordagem interdisciplinar, integrando uso de modelos de agrupamentos de *machine learning* e com técnicas da psicologia para análise de padrões ocultos e mensuráveis, buscando testar as hipóteses apresentadas neste estudo.

1.1. Problema de pesquisa

A questão central proposta é: Há uma relação entre os níveis de solidão, os padrões de comportamento político e o uso de mídias sociais entre eleitores durante o segundo turno das eleições presidenciais brasileiras de 2022?

1.2. Hipóteses testáveis:

Como apontando, a solidão também se configura-se como um problema social contemporâneo, resultante direto das remodelações políticas e tecnológicas das últimas décadas (Hertz, 2021). Esta hipótese articula três dimensões fundamentais que também são as variáveis desta pesquisa, sendo estas os níveis de severidade da solidão, comportamento político engajado e intensidade do uso das mídias sociais para interações políticas.

A dinâmica de interação entre esses elementos configura um sistema relacional complexo que reflete as disfunções estruturais da sociedade. Para permitir a testagem empírica através dos modelos *machine learning* (seção 4.2.1. no capítulo 4), essa hipótese geral foi desdobrada em hipóteses específicas, como mostra no Quadro 1, examinando a associação significativa entre as principais variáveis de forma bivariada (H1, H2 E H3), e posteriormente integradora (H4). Foi necessário, ainda, operacionalizar as hipóteses iniciais, originalmente genéricas, em uma escala de severidade, adotando o padrão de alta severidade como parâmetro para todas as variáveis analisadas.

Quadro 1- Hipóteses testáveis

| Hipótese | Análise considerada genérica | Testar associação por nível de severidade |
|----------|---|--|
| H1 | Associação entre solidão e comportamento político. | Há associação significativa entre solidão intensa e comportamento político engajado intenso. |
| H2 | Associação entre solidão e uso de mídias sociais. | Há associação significativa entre solidão intensa e maior uso de mídias sociais. |
| H3 | Associação entre comportamento político e uso de mídias sociais. | Há associação significativa entre alto uso de mídias sociais e comportamento político engajado intenso. |
| H4 | Associação entre solidão, comportamento político e uso de mídias sociais. | Há uma relação integradora entre solidão intensa, o engajamento político intenso e o alto uso de mídias sociais. |

Fonte: elaborado pelos autores.

1.3. Objetivos

1.3.1. Objetivo Geral

Evidenciar, através de modelos de inteligência artificial, a relação entre os níveis de solidão, comportamento político e mídias sociais no contexto das eleições presidenciais no Brasil em 2022.

1.3.2. Objetivo Específico

- Evidenciar de forma holística a conexão entre solidão, política e mídias sociais, considerando fatores históricos e contextuais relacionados;
- Caracterizar os principais indicadores da conexão entre essas variáveis no Brasil;

- Mapear os padrões de comportamento autorrelatados por eleitores acerca da intersecção entre solidão, política e mídias sociais no contexto do segundo turno das eleições de 2022 no Brasil;
- Analisar, de forma exploratória, as associações entre as variáveis, utilizando algoritmos de *ML* de agrupamentos para obter evidências empíricas da associação entre as variáveis.

1.4. Justificativa

A solidão tem sido abordada principalmente como um problema de saúde individual em estudos da psicologia, os quais se mostram insuficientes para uma compreensão ampla do fenômeno, uma vez que ele envolve condições sociais estruturais (Barreto et al., 2024). Estudos consideram a solidão não só como um problema individual, mas também político (Wilkinson, 2025) e de justiça social (Barreto et al., 2024). Segundo estes estudos, existem determinantes sociais como questões feministas (Wilkinson, 2025) e de minorias sociais (Barreto et al., 2024).

Pesquisas recentes já apontam impactos deste fenômeno, estabelecendo relações significativas entre solidão, privação social e engajamento político, investigando sua possível associação com posturas extremistas (Langenkamp, 2025). Evidências empíricas em contextos europeus, em países como Alemanha, Áustria e França, demonstram correlações consistentes entre níveis elevados de solidão e apoio a grupos políticos radicais (Peterson et al., 2025).

Enquanto alguns países avançam no entendimento e enfrentamento do problema, populações de países periféricos enfrentam impactos ainda mais severos, principalmente entre grupos minoritários (Hertz, 2021), mas ainda com pouco esclarecimento sobre esse fenômeno. No cenário brasileiro, a investigação empírica da solidão como fator causal de problemas sociais e políticos esbarra em significativas lacunas metodológicas (Barroso et al., 2016).

A recente validação da Escala Brasileira de Solidão UCLA (Barroso et al., 2016) representa um avanço crucial, embora tardio, para mensurar objetivamente esse fenômeno em nossa realidade sociocultural. Importante destacar que os estudos disponíveis se concentram em grupos etários específicos, permitindo a compreensão de fatores cruciais como cognição social e viés de desajustabilidade social (Júnior et al., 2023).

Apesar de muitos fatores sociodemográficos serem considerados para explicar o aumento da solidão, os *insights* sobre essa associação ainda são escassos. Desse modo, tais variáveis são frequentemente tratadas meramente como "fatores de risco" que aumentam a

vulnerabilidade à solidão. A frequência do controle dessas variáveis também é baixa, o que dificulta a análise dos mecanismos psicológicos subjacentes de interesse (Barreto et al., 2024).

Ainda assim, diante deste quadro, a aplicação de *Machine Learning (ML)* emerge como abordagem promissora para análise da associação entre diferentes tipos de variáveis, capaz de capturar padrões não lineares e interações complexas. Assim, associou-se a solidão a outras variáveis relacionadas a problemas de saúde mental (Abplanalp et al., 2024). Estes métodos superam as limitações das análises estatísticas tradicionais, particularmente relevantes para identificar correlações ocultas entre as muitas variáveis (Géron, 2021).

Portanto, este estudo justifica-se pela urgência em superar as limitações dos estudos nacionais existentes e fornecer evidências empíricas sobre a associação entre as variáveis através de modelos de *ML* (como *K-means*, *KNN* e *DBSCAN*) que analisam padrões não aparentes em abordagens convencionais subsidiando melhor compreensão sobre a realidade brasileira.

1.5. Estrutura da dissertação

Esta pesquisa está organizada da seguinte forma:

- **Capítulo 1-** introdução apresentada anteriormente.
- **Capítulo 2-** apresenta o panorama geral da temática na literatura através de uma abordagem quali quantitativa para o mapeamento crítico e holístico da literatura global sobre a conexão entre as variáveis solidão, política e mídias sociais; a identificação de insights e lacunas, com ênfase na escassez de estudos brasileiros que articulem as três variáveis nos últimos cinco anos.
- **Capítulo 3-** apresenta a fundamentação teórica a partir da natureza multidimensional da solidão; A solidão na psicologia e suas interfaces com a saúde mental; as evidências sobre a conexão entre solidão, política e mídias sociais; as evidências sobre a conexão entre solidão e política no Brasil; as evidências sobre a conexão entre solidão e mídias sociais no Brasil; as evidências sobre a conexão entre política e mídias no Brasil.
- **Capítulo 4-** apresenta a metodologia e a abordagem interdisciplinar seguida para construção do método proposto para análise de um problema complexo; aplicação de modelos *machine learning*; treinamento *K-means*, *KNN* e *DBSCAN* para análise do comportamento dos dados empíricos e limitações da pesquisa.
- **Capítulo 5-** apresentação da análise e discussão dos resultados através da testagem das hipóteses específicas.

- **Capítulo 6- considerações finais-** respostas às questões de pesquisa a partir dos principais avanços teóricos-metodológicos e sugestões para pesquisas futuras.
- **Referências bibliográficas**
- **Apêndices**
- **Anexos**

CAPÍTULO 2

2. PANORAMA GERAL DA LITERATURA

Neste capítulo, apresenta-se o panorama geral da temática por meio de uma revisão integrada da literatura para compreensão da temática e enquadramento do problema através da combinação de abordagens sistemática e bibliométricas, conforme o delineamento detalhado na seção 4.3.1.1. do capítulo 4.

Foram buscadas evidências de conexão entre as variáveis solidão, política e mídias sociais, na busca para atender aos dois primeiros objetivos específicos deste estudo. As buscas foram realizadas através de estudo de revisão sistemática em bases de dados eletrônicas, seguindo etapas cuja análise qualiquantitativa resultaram em 4 seções: seção 2.1.: contextualização da relação entre solidão, política e mídias sociais para análise (análise do panorama geral e do contexto brasileiro), seção 2.2.: contextualização sobre solidão e política (análise do panorama geral e do contexto brasileiro); seção 2.3.: contextualização sobre solidão e mídias sociais (análise do panorama geral e do contexto brasileiro); por fim seção 2.4.: contextualização sobre política e mídias sociais (análise do panorama e do contexto brasileiro).

2.1. Contextualização sobre a relação entre solidão, política e mídias sociais

A Figura 1 apresenta o histórico de publicações que exploram a relação entre solidão, política e mídias sociais, através da qual é possível observar que, nos últimos oito anos (2015 a 2023), o volume de publicações sobre o tema que integra as três variáveis tem sido relativamente baixo. Em 2022 houve um pequeno aumento, com a publicação de dois documentos que se manteve em 2023, indicando uma tendência crescente do número de publicações destes anos. Esta análise contextual histórica da temática na literatura, evidencia que mesmo a se tratando de um tema emergente no cenário global, a relação entre solidão, política e mídias sociais ainda tem recebido pouca atenção e o que indica muitas lacunas não identificadas na produção acadêmica mundial.

Neste cenário temporal, o pequeno aumento no volume de publicações em 2023, faz com que haja evidências sobre o interesse de pesquisa sobre a temática diante de eventos marcados pelas transformações tecnológicas (ascensão das redes sociais em 2015) e política

(conflitos políticos, principalmente no Brasil em 2015 a 2023) e suas afetações a saúde pública (as afetações da pandemia da COVID-19 em 2021 a 2023).

Figura 1- Contextualização da relação entre solidão, política e mídias sociais – número de publicações por ano

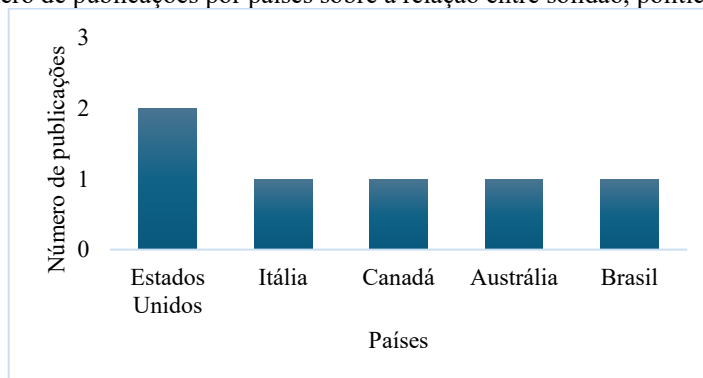


Fonte: elaborada pelos autores, 12 de outubro de 2024.

A relação entre estas variáveis tem sido pouco explorada no cenário mundial. Apenas cinco países publicaram estudos sobre a temática. A Figura 2 apresenta um gráfico com estes quatro países que publicaram estudos sobre a associação entre as variáveis solidão, política e mídias sociais no período de 2015 a 2023, sendo estes Estados Unidos, Itália, Canadá, Brasil e Austrália. Os Estados Unidos foi o país que mais publicou documento até o momento, apresentando apenas dois estudos. O que os torna polos centrais da produção acadêmica sobre esta temática para entendimento da relevância dos contextos sociais e políticos que evidenciam esta relação integradora entre estas variáveis investigadas.

O Brasil, assim como o Canadá e Austrália, embora estejam entre os cinco principais, demonstram um interesse ainda limitado na exploração da relação entre as três variáveis investigadas. Há uma necessidade de mais exploração para compreensão da temática no contexto destes países.

Figura 2- Número de publicações por países sobre a relação entre solidão, política e mídias sociais



Fonte: elaborada pelos autores, 12 de outubro de 2024.

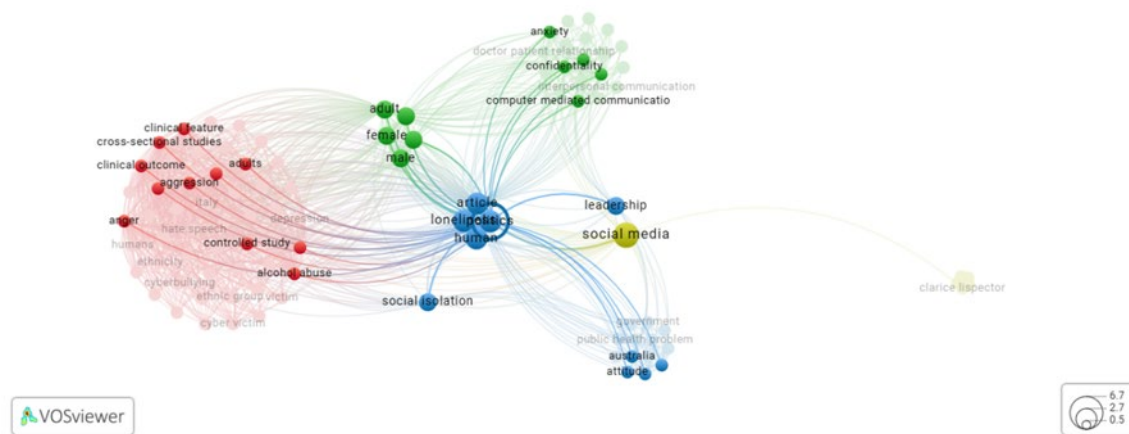
Em relação ao número de ocorrências das variáveis desta pesquisa, a Figura 3 evidencia, por meio da visualização da rede formada por *clusters* e seus agrupamentos, que 'mídia social' (identificada pelo *cluster* verde claro) é a variável mais utilizada ou de maior ocorrência nos estudos sobre o tema. Isso demonstra um maior interesse por essa variável nos últimos cinco anos. Ao analisar o mapa de relações entre os termos, o *cluster* que representa 'mídias sociais' é maior em comparação aos demais termos presentes no mapa. Seguindo o mesmo critério de análise, o *cluster* que representa a variável 'solidão' indica que esta é a segunda variável de maior ocorrência em estudos, seguida por 'política', que ocupa a terceira posição.

No que se refere à força de ligação, ou seja, a quantidade de vezes que um termo foi conectado a outros assuntos no mapa, 'mídias sociais' possui o maior número de ligações, seguida por 'solidão' e, posteriormente, 'política'.

Ainda conforme o mapa bibliométrico da Figura 3, é possível evidenciar os principais assuntos estudados que relacionam essas variáveis. O mapa é composto por 99 itens que são os termos organizados em 4 *clusters*. Dentre eles, se encontram os aspectos de relações interpessoais humanas, como 'humano', 'solidão' e 'política', estão no mesmo agrupamento (*cluster* azul). Esses termos estão diretamente ligados ao maior agrupamento do mapa (*clusters* vermelhos), que aborda temas relacionados à saúde, como 'abuso de álcool', 'agressão', 'raiva', 'resultados clínicos' e 'recursos clínicos'. Também aparecem variáveis ligadas ao tipo de estudo, como 'transversais' e 'controlados', além de questões demográficas, como 'adultos' e 'população LGBTQIA+', entre outros assuntos. Esta análise auxilia a identificar tendência de maior interesse de pesquisa nos últimos anos, indicando questões de saúde pública, métodos de estudos controlados e perfil demográfico dos grupos mais estudados em relação à temática.

Ainda na Figura 3, os *clusters* verde-escuro estão divididos em dois agrupamentos menores onde um aborda questões demográficas, como 'adulto', 'feminino' e 'masculino', enquanto o outro está ligado a temas de saúde mental, como 'ansiedade', e tecnologia, como 'confidencialidade' e 'comunicação mediada por computador'. Já os elementos azuis, também separados em agrupamentos menores, representam assuntos relacionados à saúde (como 'problema de saúde pública'), 'atitude' e 'governo'.

Figura 3- Mapa do número de ocorrência e força de ligação entre solidão, política e mídias sociais



Fonte: elaborado pelos autores.

Na Tabela 1, foram quantificados os resultados visualizados no mapa anterior, organizando em ordem decrescente as variáveis investigadas neste estudo de acordo com o total de ocorrências e a força de ligação com outros assuntos na literatura mundial que utilizam esses termos. Como já apresentado anteriormente, 'mídias sociais' é a variável mais explorada dentro da temática estudada nos últimos anos, registrando o maior total de ocorrências (4) e a maior força de ligação (113) em comparação com as demais variáveis. Já 'solidão' e 'política' apresentaram o mesmo total de ocorrências (3) e força de ligação (3). Esses dados evidenciam que 'mídias sociais' é a variável que tem recebido maior enfoque e despertado mais interesse nos estudos, em comparação com as demais variáveis analisadas.

Tabela 1- variáveis de maior interesse em estudos sobre solidão, política e mídias sociais

| Variáveis da pesquisa | Total de ocorrência em estudos | Total de força de ligação |
|-----------------------|--------------------------------|---------------------------|
| <i>social media</i> | 4 | 113 |
| <i>loneliness</i> | 3 | 108 |
| <i>politics</i> | 3 | 108 |

Fonte: elaborado pelos autores.

2.1.1. Estudos sobre a relação entre solidão, política e mídias sociais

Foram encontrados poucos artigos, sendo apenas um estudo sobre o contexto brasileiro, seguindo os critérios de busca apresentados na Seção 4.4.1.1. O único artigo encontrado, no Capítulo 4, foi analisado para compreender suas principais características, sendo estes as variáveis principais, covariáveis, lacunas e *insights* transcritos e organizados para aprofundamento.

O Quadro 2 apresenta trechos relevantes dos estudos realizados nos Estados Unidos, Canadá, Austrália e Itália, além do Brasil, sendo apontados os períodos e contexto de análise da associação entre as variáveis. Em 2015, a Austrália já demonstrava preocupação com a associação entre comportamento político e estado de saúde (Beste, 2015). Com o contexto da pandemia de COVID-19, cresceu significativamente o número de estudos que evidenciaram essa relação, utilizando dados de usuários do Twitter em países como Itália, EUA, Canadá e Brasil. Em 2021, o estudo de Josiowicz (2021) analisou *tweets* por meio de técnicas de inteligência artificial e mineração de texto, identificando que os tópicos "solidão" e "política" faziam referência a autores como Clarice Lispector e Jorge Luis Borges, refletindo mobilizações diante de transformações sociopolíticas e a busca por um sentido terapêutico durante a pandemia.

Em 2022, nos EUA, a solidão foi apontada como uma preocupação de saúde mental que se intensificou nas redes sociais a partir desse período (Zhang et al., 2022). No mesmo ano, no Canadá, Card e Hepburn (2022) destacaram a influência de crenças neoliberais no contexto pandêmico, reforçando comportamentos políticos e impactos na saúde mental, com a corrosão de vínculos sociais e o fomento a práticas individualistas.

Ainda em 2023, Mejova e colaboradores (2023) mostraram, por meio de *tweets*, que a manifestação da solidão varia conforme o gênero do usuário e evolui temporalmente. No mesmo ano, na Itália, Lo Moro *et al.* (2023) alertaram que as redes sociais se tornaram espaços de tensão e degradação social, onde a cyberagressão emergiu como uma preocupação principalmente entre adultos, afetando grupos politicamente vulneráveis e gerando consequências psicológicas significativas.

Quadro 2- Artigo que explora a relação entre solidão, política e mídias sociais na literatura

| Artigos | País | Principais variáveis | Covariáveis | Lacunas | Insights |
|--|-----------|---|---|---|---|
| 1. O político: um problema de saúde pública? John B Best -2015 | Austrália | Comportamentos políticos e indicadores de saúde | Ideologia/partido político, interesses econômicos, ciclos eleitorais, nível de desenvolvimento do país/região e contexto socioeconômico | Não opera ou faz mensuração do comportamento político como a saúde pública; ausência de dados de estudos longitudinais para provar a relação causal; não especifica de que maneira o político impacta a saúde da população e uso do termo político de forma generalizada. | Comportamento político associado aos riscos à saúde e apresenta implicações práticas. |

| | | | | | |
|--|---|---|---|---|---|
| 2. Humanidades digitais e leitura no <i>Twitter</i> : um placebo que cura em tempos de COVID-19 -Alejandra Judith Josiowicz -2021 | Brasil | - Leitura digital; <i>Twitter</i> ; humanidades digitais; Pandemia da COVID-19; | -Isolamento social; perfil dos usuários; Jorge Luís Borges; Clarice Lispector; países; tendências de <i>hashtags</i> e algoritmos do <i>Twitter</i> | -Viés de amostragem; impacto offline; outras plataformas de mídias sociais. | - Paradoxos do <i>Twitter</i> ; espaço de manifestação psicoemocional e mobilização política; novas formas de leitura digitalizada; humanidades digitais. |
| 3- A pandemia de COVID-19 e as preocupações com a saúde mental no <i>Twitter</i> nos Estados Unidos. -Senqi Zhang, Li Sun, Daiwei Zhang, Pin Li, Yue Liu, Ajay Anand, Zidian Xie e Dongmei Li -2022 | Estados Unidos | Pandemia da COVID-19 e preocupações com a saúde | Volume Total de tweets; variáveis demográficas; localização geográfica e eventos específicos | Viés da amostragem; validação psicológica; causalidade e contextualização dos dados. | Dinâmica temporal; evolução das preocupações a partir da pandemia da COVID-19, disparidade demográfica e inovação metodológica. |
| 4. O neoliberalismo está nos matando? Um estudo transversal sobre o impacto das crenças neoliberais na saúde e no bem-estar social em meio à pandemia de COVID-19. -Kiffer G. Card e Kirk J. Hepburn -2022 | Canadá | Crenças neoliberais; saúde mental, saúde física e bem-estar social | Características sociodemográficas; contexto da pandemia; apoio social e fatores de personalidade. | Viés de autorrelato, natureza transversal e limitação da generalização da amostra. | Associação significativa entre a adesão a crenças neoliberais e piores resultados em saúde mental, saúde física e bem-estar social durante a pandemia; interiorização do fracasso; corrosão do apoio social e evidencição quantitativa. |
| 5 Explorando a ciberagressão e suas consequências para a saúde mental em adultos: um estudo transversal nacional italiano -Giuseppina Lo Moro, Giacomo Scaioni,Manuela Martella,Alessio Pagani,Gianluca Colli,Fabrizio Bert andRoberta Siliquini -2023 | Itália | Cyberagressão e saúde mental e bem-estar psicossocial. | Características demográficas; padrões de Uso da internet; fatores Psicossociais e experiências Prévias. | Natureza transversal; viés da memória; limitação da generalização cultural e mecanismo de mediação. | Prevalência significativa de ciberagressão em adultos; impacto na saúde mental; identificação de grupos vulneráveis e as necessidades de intervenções. |
| 6-Gênero na expressão da solidão no <i>Twitter</i> durante os confinamentos da COVID-19 -Yelena Mejova | Não especificado-análise de tweets classificados como em inglês | Gênero e solidão | Contexto temporal; localização geográfica; padrões linguísticos e características da rede social | Identificação de gênero; viés da amostra; contexto cultural limitado e causalidade. | Diferença na manifestação da solidão entre gênero; padrões temporais de manifestação da solidão; diferença de |

| | | | | | |
|---|--|--|--|--|--|
| Yelena Mejova, Anya Hommadova Lu Anya Hommadova Lu -2023 | | | | | busca por apoio social entre gênero e inovação metodológica. |
|---|--|--|--|--|--|

Fonte: Elaborado pelos autores.

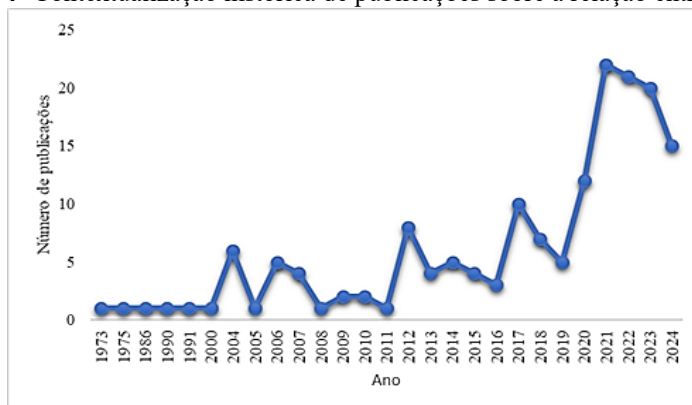
Devido à escassez de estudos, principalmente no Brasil (apenas 1 artigo), sobre a relação integradora entre as variáveis solidão, política e mídias sociais, buscou-se explorar mais amplamente a literatura para identificar outras formas de conexão entre elas. Para isso, foram realizadas combinações por pares dos termos 'solidão', 'política' e 'mídias sociais', seguindo os critérios no delineamento de estudo revisão, apresentado na seção 4.3.1.1. do Capítulo 4.

2.2. Contextualização sobre a relação entre solidão e política

A Figura 4 ilustra o histórico de estudos sobre a relação entre solidão e política, mostrando que o tema vem sendo explorado globalmente há mais de cinco décadas (de 1973 a 2024). Nesse período, observa-se um crescimento constante no número de estudos a partir de 2019, com um pico de publicações em 2021, ano marcado pela pandemia da COVID-19. No entanto, a partir de 2022 até 2024, houve uma pequena queda no volume de publicações, mas ainda são anos que se destacam nestas cinco décadas.

Através desta contextualização histórica é possível perceber como a temática evoluiu ao longo dos anos e quais os principais eventos globais neste período que aumentaram o interesse pela investigação da relação entre solidão e política, principalmente entre os anos 2021, 2022 e 2023, anos marcados pela pandemia e intensa polarização política em países como o Brasil.

Figura 4- Contextualização histórica de publicações sobre a relação entre solidão e política



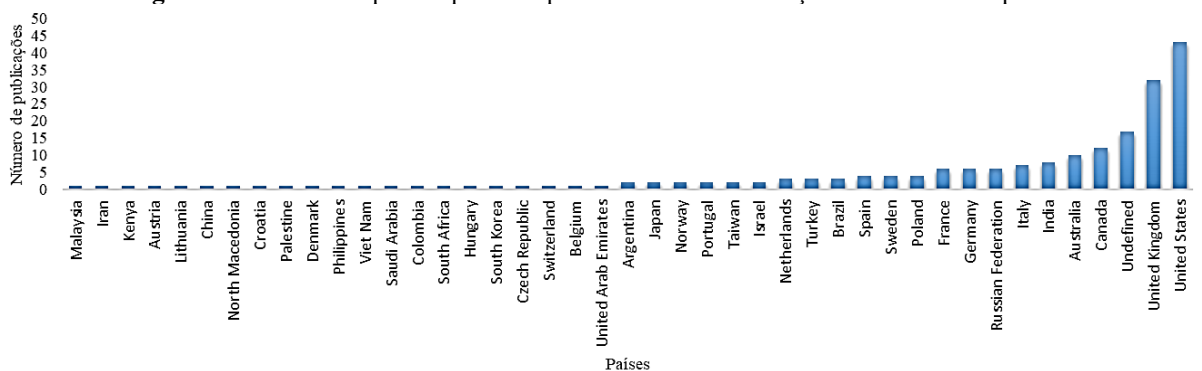
Fonte: elaborada pelos autores, 12 de outubro de 2024.

A relação entre solidão e política também é pouco explorada em escala global, com maior interesse e produção acadêmica concentrada em países do Norte Global, considerados

elites econômicas mundiais. O gráfico da Figura 5 evidencia que, no período de 1973 a 2024, os Estados Unidos têm maior interesse pela temática, liderando o volume de estudos publicados com 43 documentos, seguidos pelo Reino Unido, com 32 documentos. Entre os 15 países que mais publicaram sobre o assunto, o Brasil aparece em décima terceira posição, com apenas 3 documentos, demonstrando uma exploração ainda incipiente da temática.

Diante da contextualização global apresentada, pode-se evidenciar que os países centrais como EUA e Reino Unido tem investido e apresentado crescente interesse, sendo estes contextos importantes para serem analisados para compreensão da temática. Já o Brasil, ainda apresenta obscurantismo frente a temática, evidenciando uma grande lacuna no que tange a relação entre solidão e política que precisa ser explorada.

Figura 5- Gráfico dos países que mais publicaram sobre a relação entre solidão e política



Fonte: elaborada pelos autores, 12 de outubro de 2024.

A Figura 6 apresenta o mapa de visualização da rede formada por todas as palavras-chave presentes nos títulos, resumos e palavras-chave dos 164 documentos analisados. O mapa é composto por 1.000 itens ou termos, organizados em 19 *clusters* (agrupamentos) representados por cores diferentes. A ocorrência e a força de ligação entre esses termos podem ser visualizadas por meio da disposição dos *clusters*.

Em relação às variáveis estudadas, o mapa evidencia que 'solidão' é o termo mais utilizado nas pesquisas, apresentando a maior ocorrência, seguido por 'humano' e, posteriormente, 'política'. Esses três termos pertencem ao mesmo agrupamento, o *cluster 2*, que é o segundo maior grupo, com 87 termos representados por *clusters* na cor verde escuro. O *cluster* que representa 'solidão' tem um tamanho maior em comparação aos demais, com pouca diferença em relação ao termo 'humano', que é o maior do agrupamento. Isso demonstra que esses termos têm sido mais explorados e são de maior interesse na pesquisa sobre a temática, em comparação com outros presentes no mapa.

Tabela 2- Variáveis de maior interesse na busca geral 2

| Variáveis | Total de ocorrência | Total de força de ligação |
|---------------------|---------------------|---------------------------|
| <i>loneliness</i> | 64 | 1793 |
| <i>politics</i> | 54 | 1631 |
| <i>social media</i> | 4 | 104 |

Fonte: Elaborada pelos autores.

2.2.1. Estudos brasileiros que relacionam solidão e política

Para aprofundar a compreensão do estudo brasileiro e do contexto que envolve solidão e política, o Quadro 3 apresenta as características desse estudo, destacando tópicos considerados relevantes para a temática explorada. De modo geral, trata-se do mesmo artigo encontrado na busca anterior (busca 1), publicado em 2021. Aborda o contexto sociopolítico durante a pandemia de COVID-19, explorando a relação entre solidão, política e mídias sociais. Destaca o uso do *Twitter* como modo de lidar com as transformações sociais decorrentes da pandemia, que impactaram a saúde psicológica e a política na dimensão relacional.

Embora a solidão e seja uma variável presente neste estudo, não foi aprofundada e explorada diretamente. Citada como uma experiência subjetiva presente nos *tweets* dos usuários, a solidão nesse período somou-se ao isolamento, modulando as formas de relação com o *Twitter*.

Já a política foi um pouco mais explorada em comparação a solidão, devido ao contexto de crise de saúde pública com a pandemia da Covid-19 e as ações do governo de Jair Bolsonaro frente a este contexto. Foi analisada através de uma experiência discursiva manifestada pelos seus aspectos positivos e negativos pelo apoio ou oposição a este governo.

Ambas as variáveis tiveram impactos significativos na saúde das pessoas. Porém, o foco maior desse estudo foi nas práticas de leitura digital durante a pandemia COVID-19 através do *Twitter*, evidenciando esta plataforma de mídia social como variável principal.

Quadro 3- Artigo que explora a relação entre solidão e política publicados no Brasil

| Título, autores e ano | Principais variáveis | Covariáveis | Lacunas | Insights |
|---|---|--|--|---|
| 1. Humanidades digitais e leitura no <i>Twitter</i> : um placebo que cura em tempos de COVID-19 -Alejandra Judith Josiowicz -2021 | Leitura digital; <i>Twitter</i> ; humanidades digitais; Pandemia da COVID-19; | Isolamento social; perfil dos usuários; Jorge Luís Borges; Clarice Lispector; países; tendências de <i>hashtags</i> e algoritmos do <i>Twitter</i> | Viés de amostragem; impacto offline; outras plataformas de mídias sociais. | Paradoxos do <i>Twitter</i> ; espaço de manifestação psicoemocional e mobilização política; novas formas de leitura digitalizada; humanidades digitais. |

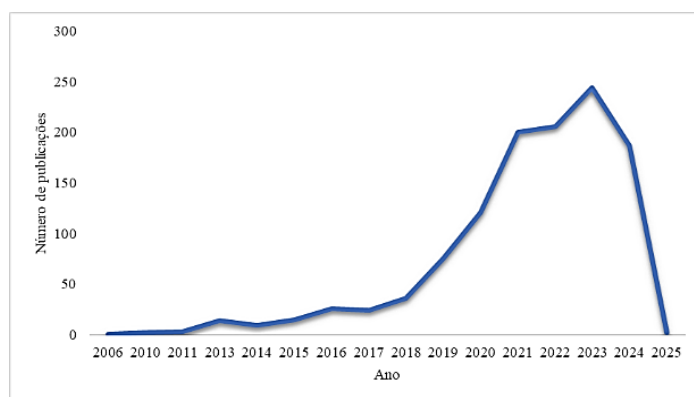
Fonte: elaborado pelos autores.

2.3. Contextualização sobre a relação entre solidão e mídias sociais

O histórico de publicações mostra que a relação entre solidão e mídias sociais vem sendo investigada há 18 anos, com estudos publicados entre 2006 e 2024, sendo tema de maior interesse para as áreas de estudos como medicina, psicologia e ciências sociais. Este histórico de publicações evidencia a evolução da temática ao longo dos anos.

Observa-se na Figura 7 um crescimento significativo no número de publicações a partir de 2021, atingindo o maior volume em 2023. Em 2024, ano ainda em curso, há uma queda no volume de publicações. Isso identifica um crescente interesse de investigação desta temática, se tornando uma preocupação a relação entre solidão e mídias sociais principalmente durante eventos que intensificaram o uso de tecnologias entre os anos 2021 e 2023, como por exemplo a pandemia da COVID-19.

Figura 7-Gráfico do histórico de publicações sobre solidão e mídias sociais

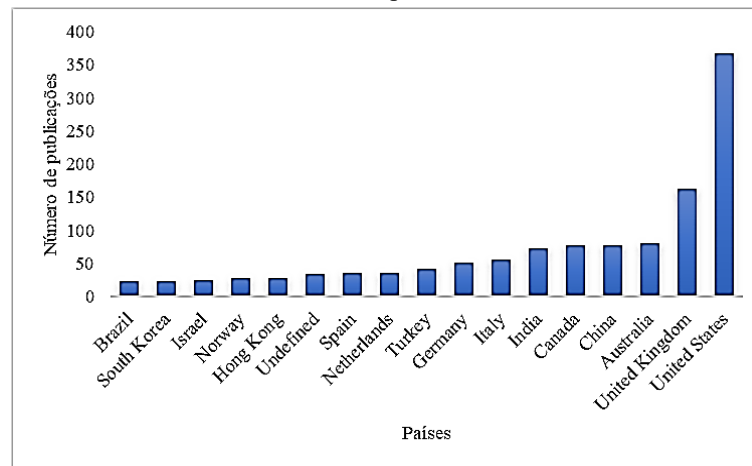


Fonte: elaborado pelos autores.

Nesse período entre 2006 e 2024, como mostra no gráfico da Figura 8, é possível analisar o volume de publicações dos países que mais publicaram estudos sobre a relação entre solidão e mídias sociais. Os Estados Unidos (EUA) lideram com o maior número de publicações, totalizando 358 documentos. Em segundo lugar, aparece o Reino Unido, com 163 estudos publicados. Já o Brasil demonstra uma exploração ainda incipiente do tema, com apenas 23 documentos publicados no mesmo período.

Deste modo, identifica-se ainda na Figura 8 que os países que lideram a produção acerca da temática são países centrais, podendo evidenciar o investimento e relevância para este contexto e culturas. O Brasil apresenta uma lacuna significativa na exploração sobre a relação entre solidão e mídias sociais, necessitando de uma exploração maior do contexto brasileiro.

Figura 8- Gráfico do número de estudos publicados sobre solidão e mídias sociais



Fonte: elaborada pelos autores.

A Figura 9 mostra o mapa bibliométrico de rede formado por todas as palavras-chave presentes nos títulos, resumos e palavras-chave dos 1.165 documentos analisados. Esta análise permite compreender a relação entre solidão e mídias sociais, sendo possível visualizar a conexão destas variáveis com outros temas que são prioritários para este estudo.

O mapa é composto por 1.000 itens ou termos, organizados em 9 *clusters*, que evidenciam a relação entre as variáveis e os temas mais explorados nos estudos. A variável 'mídia social', representada pelo cluster azul (sexto maior agrupamento de termos), destaca-se por seu tamanho maior em comparação aos demais *clusters*, indicando sua maior ocorrência nos estudos.

Além disso, em termos de conexões com outros assuntos, 'mídias sociais' é o segundo termo com maior força de relação, antecedido apenas pelo termo 'humano' (representado pelo cluster verde). Isso demonstra que, dentro das temáticas exploradas na literatura mundial relacionadas à solidão, 'mídias sociais' tem sido a variável mais estudada e conectada a diversos temas nos últimos anos.

Já o termo 'solidão', também representado pelo cluster azul claro, está no mesmo agrupamento que 'mídias sociais' e é o segundo termo de maior ocorrência. Ambas as variáveis estão relacionadas a uma ampla gama de assuntos nas pesquisas.

No maior agrupamento do mapa (299 termos), representado pelo cluster vermelho, destacam-se temas importantes para a compreensão do tema desta pesquisa, como saúde psicológica pública (medo, frustração, ligação social, psicologia social, empatia, interpessoal, entre outros); COVID-19 (infecção pelo Coronavírus, pandemia, isolamento social); e tecnologias de informação e comunicação (videoconferência, e-mail, realidade virtual).

Tabela 3- Variáveis de maior interesse na busca 3

| Variáveis | Total de ocorrência | Total de força de ligação |
|---------------------------|----------------------------|----------------------------------|
| <i>social media</i> | 830 | 16717 |
| <i>loneliness</i> | 768 | 16223 |
| <i>Health care policy</i> | 16 | 566 |
| <i>Health policy</i> | 9 | 303 |
| <i>Public policy</i> | 4 | 173 |
| <i>Policy</i> | 4 | 100 |
| <i>Politics</i> | 0 | 0 |

Fonte: elaborada pelos autores.

2.3.1. Estudos brasileiros que relacionam solidão e mídias sociais

O Quadro 4 apresenta os resultados da busca delimitada para o contexto brasileiro, destacando trechos relevantes para a compreensão da temática investigada, com base em artigos que exploram a relação entre solidão e mídias sociais no Brasil através de métodos empíricos ou de revisão sobre estudos empíricos. Foram incluídos 5 estudos para análise aprofundada, sendo estes organizados dos mais recentes (publicados em 2023) aos mais antigos (publicados em 2020, sendo dois 2 artigos publicados em 2021).

Quadro 4- Artigos brasileiros que exploram a relação entre solidão e mídias sociais

| Título, autores e ano | Variáveis principais | Covariáveis | Lacunas | Insights |
|--|---|---|---|---|
| 1. Solidão conectada: vivências de depressão em grupo fechado do Facebook -Adroaldo de Jesus Belens; Ligia Maria Rangel Santos e Jane Mary de Medeiros Guimarães -2023 | - Depressão; grupo fechado no <i>Facebook</i> ; e perfil dos participantes | - Estigma social; acesso a saúde mental; algoritmos do <i>Facebook</i> e eventos externos; | - Viés de seleção participantes; ética na pesquisa; longo prazo; comparação com outros espaços e dados demográficos limitados | - Ambivalência do apoio online; hierarquias invisíveis; dinâmicas do grupo de <i>Facebook</i> ; solidão e conexão digital linguagem específica e efeito da pandemia |
| 2. Solidão e desconexão social em tempos de pandemia entre brasileiros: evidências da iniciativa ELSI COVID-19 -Juliana Lustosa Torres, Luciana de Souza Braga, Bruno de Souza Moreiram, Camila Menezes Sabino Castro, Camila Teixeira Vaz e Amanda Cristina de Souza Andrade -2022 | - Solidão autorrelatada; desconexão social; fatores sociodemográficos (pessoas com 50 anos ou mais); | - Saúde mental e física; restrições pandêmicas; acesso à tecnologia; suporte social; e condições socioeconômicas | - Viés de amostragem; limitações temporais; mensuração da solidão e fatores culturais. | - Prevalência da solidão em pessoas com 50 anos ou mais; fatores protetivos do uso de mídias sociais; suporte familiar; impacto das políticas públicas, restrições locais e desigualdades estruturais |
| 3. Impacto de curto prazo das medidas de distanciamento social durante a pandemia de COVID-19 na função cognitiva e na percepção de saúde de idosos brasileiros: um estudo pré-pós -Pedro HTQ Almeida, Lilian D. Bernardo, Tatiana B. Pontes Jane A. Davis, Tainá MS Deodoro, Rafaela G. Ferreira Klyсна I. Souza e Joy C. MacDermid -2021 | -Função cognitiva; percepção de saúde; exposição ao distanciamento social e fatores sociodemográficos | - Acesso à tecnologia para contato sociais; condições de saúde prévias; suporte social e nível socioeconômico. | - Viés de amostragem; fatores de saúde; limitações temporais; instrumentos de medida e variáveis comportamentais | - Declínio cognitivo acelerado; efeito protetor da tecnologia; fatores de resiliência |
| 1. Humanidades digitais e leitura no <i>Twitter</i> : um placebo que cura em tempos de COVID-19 -Alejandra Judith Josiowicz -2021 | - Leitura digital; <i>Twitter</i> ; humanidades digitais; Pandemia da COVID-19; | -Isolamento social; perfil dos usuários; Jorge Luís Borges; Clarice Lispector; países; tendências de hashtags e algoritmos. | -Viés de amostragem; impacto offline; outras plataformas de mídias sociais. | - Paradoxos do <i>Twitter</i> ; espaço de manifestação psicoemocional e mobilização política; novas formas de leitura digitalizada; humanidades digitais. |
| 5. Estudantes timorenses em rede: laços sociais e conforto emocional diante da solidão da experiência migratória no Brasil e em Portugal - Silvia Garcia Nogueira -2020 | - Experiência migratória; laços sociais; conforto emocional e adaptação cultural | - Brasil e Portugal como países de acolhimento; acesso à tecnologia (mídias sociais) e experiências prévia | - Viés de amostragem; abordagem temporal; comparação limitada; variáveis econômicas e perspectiva interseccional | - Redes transnacionais como suporte; diferenças entre Brasil e Portugal; estratégias de enfrentamento; uso intensivo de redes sociais e papel das universidades |

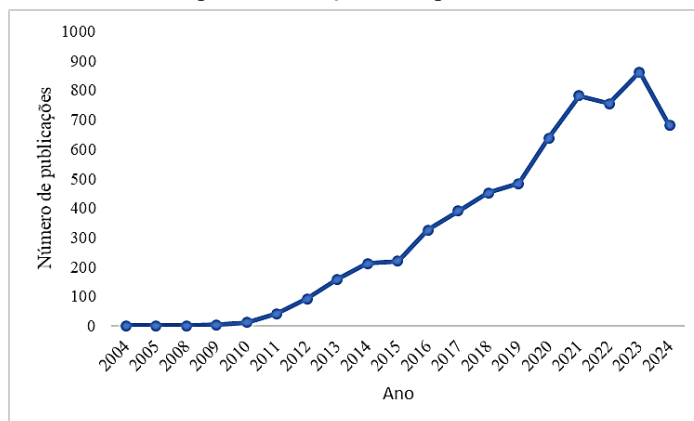
Fonte: elaborado pelos autores.

2.4. Contextualização sobre a relação entre política e mídias sociais

No que se refere sobre a relação entre política e mídias sociais, a Figura 10 ilustra a evolução temporal dessas publicações, que se estendem de 2004 a 2024 (em 20 anos), permitindo critérios para contextualização histórica acerca da temática. Observa-se um crescimento exponencial no número de estudos a partir de 2017, com um pico entre 2017 e 2020. Em 2021, houve uma redução no volume de publicações, seguida por uma retomada do crescimento em 2023, que atingiu seu ápice nesse ano. No entanto, em 2024, verifica-se uma nova queda, embora este ano ainda esteja em curso.

A tendência temporal neste gráfico aponta padrões de eventos globais que influenciam no interesse pelo tema, principalmente a partir de 2017, com o crescimento exponencial de publicações. A partir desta conjuntura, reflete-se a importância das mídias sociais como ferramentas de influência, mobilização e discussão política, principalmente em um contexto global marcado por diversas transformações digitais e polarizações. A partir do crescimento de publicação entre os anos 2017 e 2020, precedido de uma queda em 2021 e retomada dessas publicações em 2023, evidenciam-se os eventos que afetaram a política, tais como eleições controversas, crises políticas, aumento do uso das mídias sociais durante a pandemia da COVID-19.

Figura 10- Tendência temporal da relação entre política e mídias sociais no contexto global



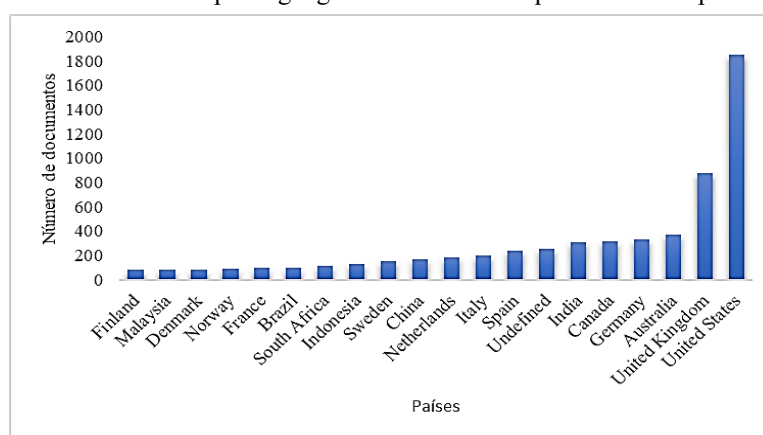
Fonte: elaborada pelos autores.

A Figura 11 apresenta um gráfico a análise temporal e geográfica da produção acadêmica acerca da temática, comparando diferentes regiões e contextos globais e seus engajamentos na produção de estudos acerca da temática explorada. Destacam-se países com maior produção de estudos sobre a relação entre as duas variáveis analisadas ao longo dos vinte anos.

Os Estados Unidos lideram com um volume significativo de publicações, totalizando 1.850 documentos. Em segundo lugar, aparece o Reino Unido, com 878 estudos publicados. O Brasil, por sua vez, demonstra menor engajamento na temática, ocupando a 14ª posição entre os dezenove países mais produtivos, com apenas 100 documentos publicados.

Essa baixa produção sugere uma evidência limitada sobre a relação entre política e mídias no contexto brasileiro. Deste modo, pode ser evidenciada liderança dos países centrais na produção de estudos acerca da temática, mas também a desigualdade nesta produção, apontando lacuna na quantidade de estudos brasileiros acerca do tema.

Figura 11-Gráfico análise temporal-geográfica dos estudos que relacionam política e mídias sociais



Fonte: elaborada pelos autores.

A Figura 12 exibe o mapa bibliométrico de rede formado por todas as palavras-chave presentes nos títulos, resumos e palavras-chave dos 6.105 documentos analisados. Este tipo de análise auxilia na compreensão do conhecimento existente acerca da temática, bem como identificação das tendências e focos de pesquisa a partir dos termos de maior ocorrência e suas conexões.

O mapa é composto por 1.000 itens ou termos, organizados em 6 clusters, que revelam a frequência de ocorrência e a força de conexão entre as variáveis 'política' e 'mídias sociais', além de outros termos explorados na literatura. O *cluster* verde, que representa 'mídias sociais', é o de maior ocorrência, evidenciado pelo tamanho do cluster em relação aos demais. Já o termo 'política', representado pelo *cluster* azul claro, ocupa a segunda posição em termos de ocorrência. Em relação às conexões dessas variáveis com outros temas, 'mídias sociais' apresenta o maior número de ligações, seguida por 'política'.

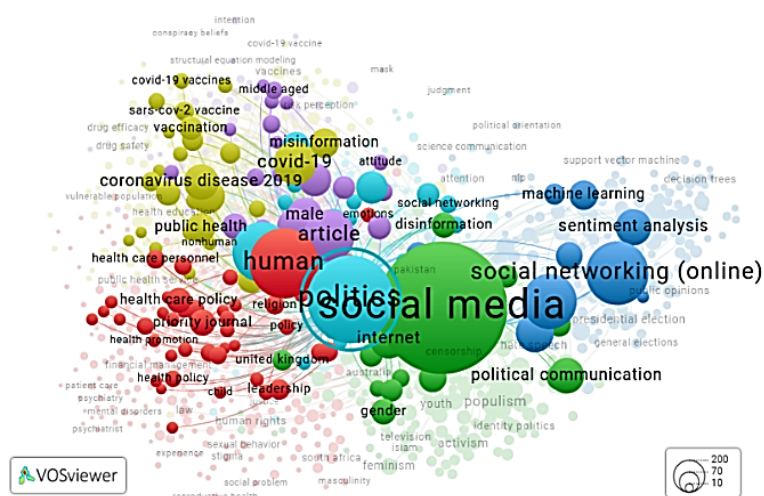
Dentre os termos mais associados às variáveis estudadas, destacam-se aqueles que compõem o maior agrupamento, representado pelos *clusters* vermelhos. Além do termo

'humano', que possui a maior ocorrência e força de ligação, estão presentes termos relacionados a dimensões como políticas públicas; políticas de saúde e cuidados com a saúde; promoção da saúde; religião; crianças; isolamento social; *'priority journal'*; e liderança. Outros termos associados a aspectos demográficos (adolescentes, pessoas pretas, crianças, entre outros); saúde (ansiedade, depressão, aborto, *bullying*, alcoolismo, psicologia, medicina, entre outros); sentimentos (raiva); e comportamento (agressão) não aparecem de forma evidente no mapa, mas integram o maior grupo, totalizando 315 itens.

O segundo maior agrupamento, representado pelo *cluster* verde, contém 186 itens e inclui termos relacionados a política, internet e gênero, que são mais visíveis no mapa. Também fazem parte desse grupo termos como mudança climática; afeto; *'dataset'*; ativismo político; ativismo digital; desinformação; política digital; democracia; eleições; campanhas eleitorais; feminismo; extremismo; partidos políticos; polarização política; Donald Trump, entre outros que, embora não estejam evidentes no mapa, foram considerados para a compreensão da temática explorada.

O terceiro maior agrupamento, com 159 itens, reúne termos relacionados às novas tecnologias de comunicação e informação, como redes sociais *on-line*; técnicas de inteligência artificial, como *'machine learning'* e análise de sentimentos. Outros termos que não estão evidentes no mapa, mas são relevantes para a compreensão da temática investigada, incluem *Twitter, Facebook*; inteligência artificial; aprendizado profundo; *fake news*; processamento de linguagem natural; eleições presidenciais; opinião pública; classificação de texto; e computação em ciências sociais, entre outros.

Figura 12-Mapa do número de ocorrência e força de ligação entre política e mídias sociais



Fonte: elaborado pelos autores.

A Tabela 4 apresenta os dados quantitativos extraídos a partir do mapa bibliométrico, destacando as temáticas que relacionam política e mídias sociais. Ao apresentar quantificação do total de ocorrência e conexões entre as variáveis estudadas, pode-se analisar o eixo central dos estudos a partir dos termos mais estudados e a relação entre eles dentro da temática. Como já evidenciado no mapa, o termo 'mídias sociais' possui a maior ocorrência, apresentando o total de 3.241 menções. Além disso, é o termo que apresenta o maior número de conexões com outros assuntos, totalizando 24.212 ligações nos últimos anos. Em segundo lugar, o termo 'política' aparece com 1.427 ocorrências e demonstra uma forte relação com outros temas, somando 17.873 ligações.

O termo 'humano', que não se destacava nas buscas anteriores, surge como o terceiro termo de maior ocorrência e força de ligação nos estudos analisados. Por outro lado, o termo 'solidão' não aparece como tema de pesquisa relevante na relação entre política e mídias sociais nos últimos anos.

Em resumo, na Tabela 4 observa-se os dados quantitativos que mostram que 'mídias sociais' tem sido a variável de maior interesse nos estudos, tanto em termos de ocorrência quanto de conexões com outros temas, seguida de perto por 'política', que também apresenta uma forte relevância na literatura analisada. O termo revela aspectos comportamentais e subjetivos a serem explorados neste contexto. Já o termo solidão dentro desta temática tem sido negligenciado, se tornando uma lacuna ao relacionar política e mídias sociais.

Tabela 4- Variáveis de maior interesse na busca 4

| Variáveis | Total de ocorrências | Total de força de ligação |
|---------------------|----------------------|---------------------------|
| <i>social media</i> | 3241 | 24212 |
| <i>politics</i> | 1427 | 17873 |
| <i>Loneliness</i> | 0 | 0 |

Fonte: elaborada pelos autores.

2.4.1. Estudos brasileiros que relacionam as variáveis política e mídias sociais

O Quadro 4 apresenta os resultados da busca delimitada para o contexto brasileiro, destacando trechos relevantes para a compreensão da temática investigada, com base em artigos que exploram a relação entre política e mídias sociais no Brasil através de métodos empíricos ou de revisão sobre estudos empíricos.

Ainda sobre o Quadro 4, os achados estão organizados em ordem cronológica decrescente, abrangendo publicações de 2020 a 2024. A análise revela uma quantidade significativa de estudos mais recentes que se concentram nos efeitos da polarização política no

Brasil, particularmente no contexto da pandemia de COVID-19 e sobre a gestão do governo Jair Bolsonaro, com destaque para o papel das plataformas de mídias sociais nesse processo.

Paralelamente, investigações sobre outros aspectos das mídias digitais ganharam relevância nos últimos anos, incluindo: a atuação de algoritmos no reforço de estereótipos racistas, os mecanismos de disseminação de desinformação, e as influências sobre o comportamento social e eleitoral. Essas evidências apontam para o crescente impacto das tecnologias de comunicação, como redes sociais e sistemas de *chatbots*, no cenário político nacional, especialmente no que diz respeito aos seus vieses algorítmicos e efeitos sociocomportamentais.

Quadro 5- Artigos brasileiros que exploram a relação entre política e mídias sociais no Brasil

| Título, autores e ano | Variáveis principais | Covariáveis | Lacunas | Insights |
|--|--|---|---|--|
| 1.Os que alcançam a bolha e o discurso incivilizado na esfera pública online polarizada -Jordânia K. Kobellarz, Milos; Brocic,Daniel Prata, Thiago H. Silva -2024 | Exploradores de bolhas, incivilidade do discurso, contexto político nacional e neutralidade ideológica | Políticos e não políticos; plataformas de mídias sociais (Twitter e Facebook) | Viés da plataforma, nuances da cultura, longevidade | Perfil político de pessoas neutras também possuem incivilidade, efeito no contexto político (temas que se tornam arena do antagonismo político). |
| 2. Investigando o 'efeito Bolsonaro' na propagação da pandemia de Covid-19: uma análise empírica de dados observacionais no Brasil -Mireille Razafindrakoto, François Roubaud,Marta Reis Castilho, Valéria Pero, João Saboia - 2024 | Efeito Bolsonaro, região e disseminação da COVID-19 | Fatores socioeconômicos políticos Locais; fases da pandemia; comportamento da População | Subnotificação de casos e mortes por COVID-19 no Brasil; dificuldade de isolar os “efeitos Bolsonaro” de outros fatores; viés político; falta de dados comportamentais diretos. | Efeito negacionista e relação entre polarização política e saúde e heterogeneidade regional. |
| 3.O impacto das teorias da conspiração e do conhecimento sobre vacinas na intenção de vacinação: um estudo longitudinal - Lucas Henriques Viscardi, Felipe Vilanova, Felipe Carvalho Novaes, Lessandra Michelin, Ângelo Brandelli Costa - 2024 | Teorias da conspiração; conhecimento sobre vacinas; intensão de vacinação e tempo. | Fatores demográficos; fatores psicológicos (confiança e aversão); fatores sociais (redes sociais, grupo de referência (família, amigos); contexto pandêmico | Viés do autorrelato; fatores não mensurados, viés das redes sociais (algoritmos), | Efeitos da teoria da conspiração; conhecimento como fator protetivo; heterogeneidade grupal |

| | | | | |
|--|---|--|--|--|
| <p>4. O preconceito das redes sociais: verdade e ansiedade nas políticas de engajamento - Paulo Beer, André Ferreira Bezerra, Beatriz Moreli Taldivo, Eduardo Pavinato Olimpio, Fabio Redkowiez Rodrigues Gomes, Jéssica Lopes Camargo, Lívia de Lucas Olivero -2024</p> | <p>Viés algorítmico; desinformação e informação; sofrimento psicológico; políticas de engajamento</p> | <p>Fatores demográficos; comportamento do usuário; (Tempo de uso) Contexto político e social (eleições, crises) que aumentam a polarização e Design de plataforma:</p> | <p>Viés de seleção das mídias digitais; dados limitados a saúde mental; falta de transparência algorítmica; efeitos culturais e métricas de engajamento.</p> | <p>Algoritmos priorizam estudos que geram reações fortes; exposição de usuários a visões cada vez mais extremas; saúde mental em risco por excesso de exposição conspiratórias (ansiedade e desconfiança social)</p> |
| <p>5. Evolução da democracia brasileira: revelando a dinâmica eleitoral em questões políticas, negatividade e aclamação - Isabella Gonçalves, Mathias-Felipe de Lima-Santos Vicente Fenoll e Yossi David -2024</p> | <p>Ciclo eleitoral; estratégia de comunicação política e influência da mídia; comportamento do eleitorado; polarização; saúde democrática</p> | <p>Contexto socioeconômico; Demografia; fatores institucionais e eventos Externos.</p> | <p>Dados limitados em eleições presidenciais; dificuldade em separar efeitos da mídia de fatores estruturais e viés temporal.</p> | <p>Aumento da negatividade pós-2016 e relação com polarização; temas políticos dominantes; mídia social (amplificação de discursos radicais versus mídia tradicional) e regionalismo.</p> |
| <p>6. Censura: uma reação à desinformação na World Wide Web - Alexandre Eustáquio Perpétuo Braga, Adilson Luiz Pinto, Enrique Muriel-Torrado, Moisés Lima Dutra -2024</p> | <p>Desinformação; Web Políticas de Censura; Contexto Político-Social,</p> | <p>-Legislação local, tipo de plataforma de redes sociais; perfil dos usuários e tecnologia utilizada.</p> | <p>-A desinformação assume diversas conotações</p> | <p>Educação dos usuários; desenvolvimento de competência informacional; contexto “infodêmico”, impacto na Liberdade de expressão; eficácia da censura e reação dos Usuários</p> |
| <p>7. Política e confiança na vacinação contra a COVID-19: um estudo transversal brasileiro - Marco Antonio Catussi Paschoalotto, Joana Cima, Eduardo Costa, Sara Valente de Almeida, Joana Gomes da Costa, João Vasco Santos, Claudia Souza Passador, João Luiz Passador & Pedro Pitar Barros -2024</p> | <p>Orientação política: posicionamento, exposição a informações e fatores institucionais.</p> | <p>Fatores sociodemográficos; contexto de saúde; religiosidade e efeito temporal.</p> | <p>-Viés de autorrelato, Pessoas com desconfiança prévia podem adotar certas posições políticas, não o contrário; amostragem: possível sub-representação de grupos vulneráveis (ex.: populações rurais); falta de análise qualitativa: não explora porquês por trás da desconfiança (ex.: medo de efeitos colaterais, teorias conspiratórias). Comparações internacionais</p> | <p>-Polarização política; efeito da mídia: Quem consome redes sociais (WhatsApp); desigualdades regionais: regiões com maior acesso a SUS; papel das instituições</p> |
| <p>8. Briga política ofusca defesa da vacinação: como o debate sobre vacinas</p> | <p>Enquadramento do debate sobre vacinas; engajamento com conteúdo antivacina e</p> | <p>Perfil dos usuários; Algoritmos do Twitter; Contexto</p> | <p>-Viés da plataforma, investigando apenas o Twitter; causalidade;</p> | <p>-Politização da saúde; declarações de Bolsonaro contra vacinas e</p> |

| | | | | |
|---|--|---|---|---|
| no Twitter brasileiro foi enquadrado por antivacinas durante o governo Bolsonaro - Marina Verjovsky, Mariana Porto Barreto, Isabella Carmo, Bruno Coutinho, Lilian Thomer, Sérgio Lifschitz, Claudia Jurberg, -2023 | contexto de aprovação das vacinas | mediático e fatores temporais. | limitação da análise qualitativa e viés temporal restringindo-se apenas ao governo Bolsonaro. | estratégias antivacina (falsos especialistas e “liberdade individual”) falta de protagonismo científico. |
| 9. O papel das plataformas de redes sociais para a legitimação discursiva: desvendando o neoliberalismo por trás do discurso sobre as universidades públicas - Daniel S. Lacerda e Rita de Cássia T. Santos -2023 | Legitimação discursiva do neoliberalismo; conteúdo das redes sociais (enquadramento temático; atores dominantes e estratégias discursivas) e plataformas analisadas: | Contexto político (corte orçamentários no ensino superior sob Bolsonaro); reformas neoliberais em curso; crises institucionais: greves universitárias e Trajetória midiática. | Viés de seleção; dificuldade de mensuração; causalidade complexa; amostra temporal e viés ideológico do pesquisador. | Algoritmo como ferramenta poderosa que pode manipular a interação e comunicação; importância de (re)imaginar criticamente a universidade; desvendar o neoliberalismo por trás da avaliação moral das universidades públicas |
| 10. Análise da ação política do Movimento de Reforma Sanitária Brasileira na pandemia de COVID-19: 2020-2021 - Jamilli Silva Santos e Carmem Fontes Teixeira -2023 | Reforma sanitária; medidas sanitárias; políticas públicas e SUS. | Cenário político partidário; crise econômica e orçamento; movimentos sociais e desinformação e negacionismo. | Efetividade concreta; divergências internas e comparação internacional e papel da mídia. | Estratégias de resistência; inovações com uso de redes sociais; impacto na opinião pública e lições para crises futuras, |
| 11. Flutuações de longo prazo na polarização no Brasil – Análise do compartilhamento de informações políticas entre 2011 e 2019 -Marcelo Alves dos Santos Junior -2023 | Nível de polarização política; temas polarizadores; mudanças na polarização ao longo de eventos críticos e diferenças entre grupos | Eventos políticos e crises institucionais; estratégias de comunicação digital; cobertura midiática e contexto socioeconômico | Se o aumento no compartilhamento reflete polarização real ou apenas ativismo digital; viés das plataformas e efeito de bolhas ideológicas | Padrões de viralização; eventos críticos como aceleradores; mudanças no comportamento digital e resiliência democrática |
| 12. Atingir a bolha pode não ser suficiente: o papel da mídia na polarização política online. -Jordânia K. Kobellarz, Miloš Bročić, Alexandre R. Graeml, Daniel Prata & Thiago H. Silva -2022 | Exposição a conteúdo político online; Polarização política; papel da mídia e engajamento em redes sociais. | Algoritmos de plataformas; identificação partidária; eventos políticos e desinformação | Dificuldade em determinar se a mídia causa polarização ou apenas reflete divisões; efeitos de longo prazo; plataformas não analisadas e viés de amostragem. | Exposição cruzada limitada; mídia tradicional como moderadora; grupos ideológicos podem reagir de forma diferente. |
| 13. É verdade! Eu vi no WhatsApp”: Redes Sociais, Covid-19 e Orientação Político-Ideológica no Brasil | Consumo e compartilhamento de informações sobre COVID-19 no WhatsApp; | Identificação partidária; nível de educação; perfil demográfico e contexto pandêmico | Viés de plataforma; amostragem limitada; efeitos comportamentais e dinâmicas de grupo. | Hiperpartidarização da desinformação; WhatsApp como "ecossistema fechado"; |

| | | | | |
|---|---|---|---|---|
| -Mozer de Miranda Ramos, Rodrigo de Oliveira Machado & Elder Cerqueira-Santos -2022 | Orientação político; Credibilidade atribuída às informações e engajamento em desinformação. | | | Efeito de homofilia ideológica; Papel dos "influenciadores de grupo |
| 14. Desinformação nas redes sociais durante a pandemia do novo coronavírus: um estudo de caso quali-quantitativo do Brasil. -Priscila Biancovilli, Lilla Makszin e Cláudia Jurberg -2022 | Tipos de desinformação; Plataformas analisadas; Engajamento; Perfil dos disseminadores | Contexto político; Cronologia da pandemia; Fatores socioeconômicos e algoritmos de redes sociais. | Amostra limitada; dinâmicas de grupos privados; Impacto comportamental; comparação internacional | Narrativas predominantes; teorias conspiratórias sobre origem do vírus; plataformas de mídias sociais (WhatsApp e Twitter); estratégias de viralização; Apropriação de imagens de autoridades e conexão política. |
| 1. Humanidades digitais e leitura no <i>Twitter</i> : um placebo que cura em tempos de COVID-19 -Alejandra Judith Josiowicz -2021 | Leitura digital; Twitter; humanidades digitais; Pandemia da COVID-19; | Isolamento social; perfil dos usuários; Jorge Luís Borges; Clarice Lispector; países; tendências de hashtags e algoritmos do <i>Twitter</i> | Viés de amostragem; impacto offline; outras plataformas de mídias sociais. | Paradoxos do <i>Twitter</i> ; espaço de manifestação psicoemocional e mobilização política; novas formas de leitura digitalizada; humanidades digitais. |
| 16. Qualidade da argumentação em tweets políticos: o que é e como medir -Cássio Faria da Silva; Amanda Pontes Rassi; Jackson Wilke da Cruz Souza; Renata Ramisch; Roger Alfredo de Marci Rodrigues Antunes; Helena de Medeiros Caseli -2021 | Qualidade da argumentação; características dos tweets; perfil dos autores; orientação ideológica; cargos políticos e engajamento. | Contexto político; Igoritmos do Twitter; educação e capital cultural; plataforma como espaço discursivo | Viés de plataforma; Subjetividade na avaliação; amostra temporal; efeito do anonimato; impacto real | Baixa qualidade predominante; uso frequente de simplificações binárias; diferenças ideológicas; formato vs. conteúdo; políticos vs. cidadãos. |
| 17. Narrativas entre ciência e política no ativismo cloroquina - André Luiz Machado das Neves; Breno de Oliveira Ferreira -2020 | Narrativas discursivas; Uso de evidências; Atores envolvidos; contexto da pandemia; plataformas de disseminação | Posicionamento político; Capital científico dos atores; Eventos críticos; Desigualdade de acesso à saúde | Viés de confirmação; Impacto comportamental; abordagem temporal; análise de audiência; comparação internacional | Relação ciência e política-política; politização da metodologia; padrões narrativos; assimetria discursiva e efeito plataforma |

Fonte: elaborado pelos autores.

O estudo evidenciou o contexto de polarização política no Brasil nos últimos anos, destacando a influência de comportamentos políticos extremista, predominantemente manifestados nas mídias sociais, sobre as eleições presidenciais de 2022. Tais dinâmicas emergem como objeto crucial de análise para sua associação com experiências de solidão no período, o que não foi apontado pelos estudos sobre o tema até o momento dessa pesquisa, principalmente ao que se refere a formação de bolhas políticas ideológicas.

CAPÍTULO 3

3. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, buscou-se interpretar e aprofundar em discussão sobre a solidão e os dados coletados à luz da literatura, situando-os em um contexto já existente no campo científico, especialmente no Brasil. Também, responde de forma mais aprofundada o segundo objetivo específico deste estudo, sendo inicialmente através de uma revisão bibliográfica em corpus pré-selecionados de obras fundamentais. Posteriormente, os resultados da análise qualitativa da revisão sistemática da literatura.

Como analisado no capítulo anterior, com base na revisão sistemática, existem poucas evidências da relação integrativa entre solidão, política e mídias sociais na literatura global. Ainda assim, os estudos buscam explorar essa relação e seus aspectos importantes, contribuindo para o apontamento dos principais *insights* e lacunas sobre a temática até o momento e permitindo aprofundamento sobre o cenário brasileiro.

Dividiu-se este capítulo em 6 seções que discutem o tema dentro de uma progressão cronológica e/ou temática sobre a conexão entre as variáveis estudadas: seção 3.1: visão geral sobre a natureza multidimensional da solidão; seção 3.2.: o conceito de solidão na psicologia; seção 3.3.: a relação integradora entre solidão, política e mídias sociais, segundo os estudos mais recentes sobre o cenário mundial; seção 3.4.: a relação entre solidão e política, segundo estudos brasileiros nos últimos anos; seção 3.5.: a relação entre solidão e mídias, segundo estudos brasileiros nos últimos anos; seção 3.6.: a relação entre política e mídias sociais, segundo estudos brasileiros nos últimos anos.

3.1. Discussão histórica contextual sobre a natureza multidimensional da solidão

A solidão é uma experiência subjetiva, mas também social (Barreto et.al 2024; Wilkinson, 2022), sendo constante ao longo da história (Minois, 2019). A espécie humana sempre foi condicionada a viver e a pensar socialmente. Tudo o que foi criado pelo ser humano formaliza essa conexão a grupos variados e a uma sociedade. Ao mesmo tempo, a solidão esteve presente em todas as épocas e culturas, assumindo diferentes significados conforme o contexto histórico e social, revelando sua natureza ambígua (Minois, 2019).

Como demonstra Minois (2019), a solidão sempre oscilou entre ser uma experiência subjetiva valorizada como elemento essencial para a introspecção e temida como fenômeno de desamparo, dependendo do período histórico, eventos sociais e como determinados atores

sociais os percebiam. Essa dualidade do significado da solidão está enraizada até em sua etimologia, sendo confundida com o isolamento em vários idiomas. Enquanto o termo originado do latim, *sollitudo* remete a um lugar deserto, hostil, incivilizado e desumanizado, e *solitarius* descreve o indivíduo isolado da civilização urbana. Essa dialética entre o coletivo e o individual é central para o entendimento deste fenômeno. O ser humano considerado como um "animal social", também na sociedade desenvolve a consciência de sua singularidade. Suas vivências sociais podem condicioná-lo ao refúgio voluntário ou ao exílio imposto (Minois, 2019).

No entanto, o fenômeno da solidão no mundo moderno passou a ser visto como problema social (Hertz, 2021) ou uma patologia devido a hiperconexão das massas (Minois, 2019). O reflexo de transformações da globalização como a urbanização e a digitalização ampliam a sensação de solidão do sujeito em espaços mesmo em meio a multidões (Hertz, 2021).

Este fenômeno já havia sido apontado pelas suas afetações na política no século XX. Arendt (2013) analisa a solidão como fenômeno político em regimes totalitários, marcada pela perda da capacidade de pensamento e ação política. Mesmo com contatos sociais, o indivíduo vivencia um abandono radical em sua subjetivação política, sendo como um não-pertencimento que ameaça a empatia e a pluralidade (Arendt, 2013). Essa condição, segundo ela, pode levar tanto à redescoberta do sentido da vida quanto à adesão a ideologias extremistas.

Em uma perspectiva mais atual e neuropolítica, Dias (2021) aborda a solidão, entendendo-a como sintoma biopsicossocial. Investiga como disfunções neuroquímicas e experiências precoces de exclusão moldam padrões de isolamento e experiências de solidão.

Nesse mesmo sentido sintomatológico de problemas sociais, dentre as abordagens sobre doenças neurais contemporâneas, Han (2015) introduz uma crítica tecnocultural ao capitalismo neoliberal, descrevendo a solidão moderna também como sintoma de patologias causadas pela hiperconexão digital que impactou as formas de interação social.

Hertz (2021) amplia a análise para uma perspectiva econômico-tecnológica, mostrando como o capitalismo neoliberal, ao priorizar lucro e urbanização em detrimento do bem-estar, transformou a solidão em um "mal do século XXI". Seus estudos destacam a sensação de sub-representação política, especialmente em grupos marginalizados, e sua relação com a polarização.

Ainda, sobre sintomas de doenças contemporâneas, a pandemia da COVID-19 evidenciou manifestações intensas desse fenômeno pré-existente (Hwang et al., 2020; Hertz, 2021).

Diante dessa diversidade de perspectivas sobre a natureza da solidão, este trabalho concentra-se nas causas estruturais que tornaram este fenômeno como um problema sociopolítico da contemporaneidade, isto é, a relação entre questões macro (transformações sociais) que afetam a percepção do sujeito sobre suas relações, analisando a interrelação de três fatores: solidão, política e mídias sociais.

Diante das limitações apresentadas para entendimento da natureza dinâmica da solidão como um individual e social, propõe-se discutir esta dinâmica complexa a partir de estudos comportamentais e psicológicos.

3.2. Solidão segundo a psicologia e suas interfaces com a saúde mental

Nesta seção apresenta a trajetória teórica, desafios de mensuração e interfaces com a saúde mental, principalmente de acordo autores da abordagem social e cognitiva comportamental.

Russell et al. (1980) introduziram a abordagem tridimensional da solidão, conceituando-a como uma experiência emocional complexa que surge da frustração de necessidades afetivas e sociais. Desenvolveram a Escala *UCLA* de Solidão, instrumento fundamental para mensuração. Weiss (1975) complementou essa visão ao diferenciar solidão emocional (falta de vínculos íntimos) e social (falta de rede de apoio), baseando-se em teorias do apego (Bowlby, 1969 a 1982 apud Weis 1975).

Neste sentido houve a consolidação do conceito multidimensional a partir de Rokach & Heather (1997) que propuseram que a solidão é uma experiência paradoxal entre o anseio por conexão e a incapacidade de mantê-la. Identificaram sete fatores socioculturais que fragilizam vínculos, dentre eles o isolamento, estresse, mobilidade geográfica.

Para análise dos avanços na compreensão clínica e social, de acordo Rokach (2013) refinou o modelo, distinguindo solidão (subjetiva) de isolamento (objetivo). Seu modelo de cinco fatores referentes ao sofrimento emocional e inadequação social, através dos quais se permitiu análises mais precisas. O que Russell *et al.* (2010) reforça com a necessidade de instrumentos específicos, como a *UCLA Scale*, para avaliar componentes distintos da solidão.

Já Cacioppo *et al.* (2009 a 2015) trouxeram a perspectiva evolucionista, interpretando a solidão como "dor social" adaptativa, com bases neurobiológicas e consequências para a saúde

mental. Assim, a solidão tem sido associada ao desenvolvimento de alguns transtornos mentais e intensificação de seus sintomas, tais como psicose, transtornos de personalidade (Hawkey; Cacioppo, 2013), depressão e anedonia social (Cacioppo *et al.*, 2009, 2015). Tal construto tem sido considerado importante para a saúde física; mortalidade; saúde mental e o funcionamento cognitivo (Hawkey; Cacioppo, 2013).

Em relação a estudos em grupos específicos e contextualizações nacionais sobre solidão, Barroso *et al.* (2016) destacaram a vulnerabilidade de idosos brasileiros, associando solidão a limitações físicas, falta de suporte social e fatores sociodemográficos como gênero feminino, baixa escolaridade. Uma das lacunas identificadas, segundo estes autores, foi a falta de estimativas nacionais devido à escassez de instrumentos adaptados ao contexto brasileiro até 2015. Também há limitações quanto ao entendimento de estimativas sobre a prevalência da solidão como problema crônico, possivelmente devido à falta de consenso conceitual sobre essa experiência que pode ocorrer em qualquer fase da vida (infância, adolescência, idade adulta ou velhice) (Júnior *et al.*, 2023)

Ainda sobre o contexto brasileiro, Júnior *et al.* (2023) focaram em adultos e idosos (acima dos 50 anos), relacionando solidão a indicadores de saúde mental negativa (depressão e distúrbios do sono) e comportamentais (cognição e desejabilidade social). Uma das limitações apontadas foram em relação aos estudos transversais não estabelecem causalidade, exigindo pesquisas longitudinais.

As tendências atuais e desafios apontadas por Abplanalp *et al.* (2024) ressaltou a complexidade de analisar múltiplas variáveis relacionadas a cognição social e depressão, defendendo o uso de modelos de *ML* para capturar padrões não lineares da solidão e isolamento social em problemas psicológicos que afetam negativamente o convívio social e na comunidade.

Neste sentido, a experiência de solidão pode diminuir ou extinguir o isolamento social, quando está ligada a motivação para conexão ou reconexão. Se a solidão se cronificar e não receber tratamento adequado, tem graves consequências para a cognição, emoção, comportamento e saúde humana (Hawkey; T Cacioppo, 2013).

Para Barreto *et al.* (2024), embora os avanços nos estudos psicológicos tenham ampliado a compreensão sobre a prevalência da solidão em certos grupos, contextos e associados variáveis demográficas, tais contribuições ainda não se mostraram suficientes para uma compreensão integral do fenômeno. Esses estudos tendem a priorizar mecanismos que operam em nível individual, como preditores, mediadores e desfechos, o que tem gerado evidências

principalmente sobre a saúde individual. Os autores ressaltam, portanto, a necessidade de investigar determinantes sociais e estruturais associados à solidão, de modo a avançar em explicações que ultrapassem a perspectiva individual.

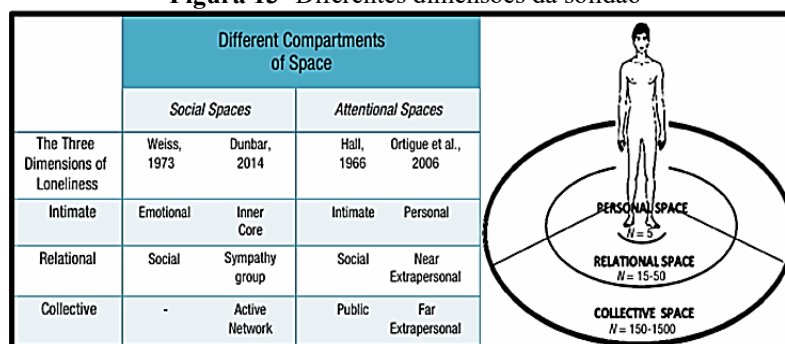
3.2.1. Solidão coletiva: instrumentalização para investigação empírica e tratamentos

A investigação sobre a solidão enquanto fenômeno psicológico tem evoluído significativamente nas últimas décadas. Cramer e Barry (1999) destacam os avanços na conceitualização e mensuração deste construto complexo, que apresenta múltiplos preditores e manifestações. Contudo, persistem divergências na comunidade científica quanto à sua natureza unidimensional ou multidimensional, resultando na ausência de um instrumento psicométrico universalmente aceito.

Estudos utilizando medidas de autoavaliação demonstram consistentemente a natureza multifatorial da solidão, revelando quatro componentes principais: solidão social, solidão emocional, afeto negativo e solidão familiar. Esta multidimensionalidade é corroborada por instrumentos amplamente validados, como a Escala UCLA de Solidão (Russell, 1996 apud Cacioppo *et al.*, 2015), que diferencia especificamente entre solidão social e emocional.

Como ilustrado na Figura 13, a solidão manifesta-se de forma diferenciada conforme o espaço relacional, processos atencionais e contexto populacional, podendo ser categorizada em três tipos distintos: solidão íntima, relacional e coletiva. Particularmente, a solidão coletiva, foco deste estudo, representa uma dimensão anteriormente negligenciada na literatura qualitativa (Cacioppo *et al.*, 2015), caracterizando-se pela desconexão percebida de identidades sociais valorizadas e a ruptura com redes de pertencimento.

Figura 13- Diferentes dimensões da solidão



Fonte: Cacioppo *et al.*, 2015.

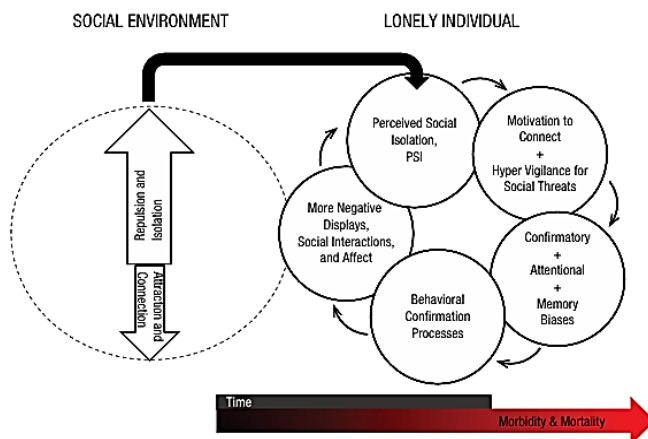
Ainda de acordo com Cacioppo *et al.*, (2015), do ponto de vista evolutivo, a solidão coletiva sugere o desenvolvimento de mecanismos motivacionais duais (para conexões íntimas

e grupais), com função adaptativa na construção identitária e manutenção de padrões cooperativos em contextos adversos.

No que concerne às intervenções terapêuticas, evidências apontam maior eficácia de abordagens grupais em comparação com individuais. Destacam-se quatro modalidades principais: ampliação de oportunidades de contato social, fortalecimento de apoio social, treino de competências sociais e modulação de cognições sociais adaptativas. Dentre estas, a abordagem cognitivo-comportamental (TCC) tem-se mostrado particularmente efetiva, atuando sobre pensamentos automáticos negativos, hipóteses relacionais defeituosas e vieses cognitivos (Cacioppo *et al.*, 2015).

O modelo neurocognitivo proposto por Cacioppo *et al.* (2015) - representado na Figura 14, estabelece que a solidão emerge como posição involuntária no continuum social (interações que têm continuidade e variações), desencadeando processos mal adaptativos como hipervigilância a ameaças sociais; vieses atencionais e de memória e padrões comportamentais de retraimento. Estes componentes constituem alvos primários para intervenções baseadas em evidências (Cacioppo *et al.*, 2015).

Figura 14-Os efeitos da solidão na cognição social. Modificado de J. T. Cacioppo e Hawkey (2009).



Fonte: Cacioppo *et al.*, 2015.

3.3.Evidências sobre a relação entre solidão, política e mídias sociais

Conforme os resultados da Seção 1 do Capítulo 2, poucos estudos brasileiros exploram diretamente a relação integradora entre solidão, política e mídias sociais, o que exigiu a inclusão de pesquisas internacionais de outros países para aumentar o escopo do estudo, tais como Itália, Austrália e EUA, ampliando a compreensão do tema.

No que tange aos comportamentos políticos e isolamento, Best (2015) investiga comportamentos políticos de risco em atores políticos fora das mídias sociais, destacando uma

tríade de fatores: insônia, tédio e isolamento. O estudo analisou habitantes de Canberra (Austrália), cujo isolamento geográfico e social influenciou sua participação política, muitas vezes reduzida ao consumo de informações digitais distorcidas. Best (2015) ressalta que a privação de sono, a falta de exercícios físicos e a repetição de ações frenéticas na política podem levar a ciclos viciosos, como a propagação de fofocas e a idealização errônea de líderes como figuras solitárias.

Card e Hepburn (2022) identificaram a pandemia de COVID-19 como um evento que amplificou problemas sociais que já vinham sendo agravados pelo neoliberalismo. Segundo estes autores, a relação entre neoliberalismo e saúde mental é complexa. O neoliberalismo é associado aos piores resultados em saúde mental coletiva por centralizar questões individuais. A solidão é uma experiência paradoxal nesse contexto, onde foram evidenciado menores níveis devido a autoeficácia. Trata-se de um mecanismo de defesa por meio do qual essa ideologia oferece uma sensação de controle e autossuficiência que protege contra o sentimento de isolamento. Essa proteção é individualista, mascarando o custo social mais amplo (Card; Hepburn,2022).

Na questão de saúde mental e política na pandemia, Zhang et al. (2022) examinaram o impacto do isolamento social durante a COVID-19 nos EUA, utilizando dados do *Twitter* (março de 2020 a janeiro de 2021). A análise revelou picos de discussões sobre solidão, política e respostas governamentais, com maior preocupação com saúde mental entre homens brancos, exceto na faixa etária de 18 a 29 anos. Apesar dos avanços na vacinação, as sequelas mentais do isolamento permaneceram como um agravante pós-pandêmico.

Lo Moro *et al.* (2023) focaram na ciberagressão na Itália, vinculando o aumento do uso de redes sociais (*Facebook*, *Twitter* e *WhatsApp*) a problemas de saúde mental em grupos minoritários (mulheres, LGBTQA+ e estrangeiros). Temas políticos e identitários geraram conflitos online, exacerbando sentimentos de raiva e tristeza entre vítimas e perpetradores.

No que se refere a expressão da solidão e diferenças de Gênero Mejova e Hommadova (2023) analisaram *tweets* com os termos "solidão" e "solitário" (2019–2021), destacando que: mulheres expressaram mais solidão no *Twitter*, especialmente em contextos como maternidade e cuidados, e receberam maior apoio social. Homens relataram menos essa sensação de solidão, e isso pode ter relação ao estigma associado à fragilidade masculina, embora tenham sofrido com o aumento da mortalidade pandêmica. A cultura individualista foi apontada como fator agravante da solidão entre homens jovens nos EUA.

Já no contexto brasileiro, analisa-se as humanidades digitais e seu engajamento nas questões de saúde e sociopolítica, Josiowicz (2021) investigou reações psicológicas e políticas no *Twitter* durante a pandemia no Brasil. Destacou-se o uso terapêutico da leitura digital, com citações frequentes a Clarice Lispector (autoajuda) e Jorge Luis Borges (crítica política). Técnicas de IA foram aplicadas para classificar os sentimentos e motivações por trás dessas postagens, revelando como a plataforma serviu tanto para escape emocional quanto para mobilização política.

3.4.Evidências sobre a relação entre solidão e política no Brasil

A relação entre solidão e política no contexto brasileiro, especialmente em conexão com o fenômeno da polarização atual, permanece pouco explorada na literatura. Essa lacuna é evidenciada pela escassez de estudos identificados ao longo dos últimos anos no Brasil e no mundo, conforme demonstram os dados estatísticos da revisão bibliográfica (ver Seção 2, Capítulo 2). A análise bibliométrica revela que “solidão” é o termo mais recorrente e frequentemente associado a outros conceitos, enquanto “política” aparece em posição secundária, sugerindo uma dissociação temática ainda pouco abordada de forma integrada.

Nestes estudos, aponta-se que a solidão emergiu como um fenômeno amplificado no Brasil pela pandemia de COVID-19, associado tanto a questões de saúde quanto a transformações sociais e econômicas. Dois fatores foram determinantes nesse processo: o isolamento social prolongado e o aumento do uso de plataformas digitais, como o *Twitter*, onde sentimentos de solidão e debates políticos se manifestaram de forma interligada (Josiowicz, 2021).

Nesse cenário, as desigualdades sociais e econômicas, especialmente entre o Norte e o Sul Global, foram intensificadas (Mordomo, 2020 apud Josiowicz, 2021). Paralelamente, observou-se um crescimento das leituras digitais com fins terapêuticos, tanto no Brasil quanto em outros países. Segundo Olave (2018 apud Josiowicz, 2022), essa prática serviu como mecanismo para processar emoções, lidar com experiências subjetivas e enfrentar o estresse do isolamento. Autores como Clarice Lispector (abordagem psicológica) e Jorge Luis Borges (dimensão sociopolítica) foram frequentemente citados pelos usuários do *Twitter*, revelando como a literatura foi instrumentalizada para construir significados sociais e confrontar a solidão (Williams, 1977 apud Josiowicz, 2021).

A pandemia, enquanto fenômeno sociocultural (Trindade Lima, 2020 apud Josiowicz, 2021), impactou tanto esferas individuais quanto coletivas. No plano subjetivo, a solidão foi

amplamente discutida como um estado afetivo marcante do período. No âmbito coletivo, destacou-se o cenário político brasileiro, especialmente as polarizações em torno do ex-presidente Jair Bolsonaro, cujas menções no *Twitter* geraram mobilizações massivas de apoio ou crítica (Josiowicz, 2021).

3.5.Evidências sobre a relação entre solidão e mídias sociais no Brasil

Esta seção discute os principais temas abordados nos estudos brasileiros publicados nos últimos cinco anos que exploram a relação direta entre solidão e mídias sociais, conforme os resultados apresentados na seção 3 do capítulo 2. A análise está organizada por enfoques temáticos e covariáveis relevantes, destacando as contribuições de cada estudo.

Ernesto et al. (2020) analisaram aspectos da solidão migratória e uso de mídias sociais. Investigaram como estudantes universitários timorenses que migraram para o Brasil e Portugal lidam com a solidão migratória, explorando o papel das mídias sociais nesse processo. A solidão migratória foi descrita como uma experiência ambivalente, percebida tanto de forma negativa (pelo distanciamento afetivo) quanto positiva (como parte do processo de adaptação). Os relatos destacaram a dificuldade de estabelecer novas conexões no país de acolhimento, especialmente nos primeiros anos, agravada por fatores como barreiras linguísticas (Morais, 2012, apud Ernesto *et al.*, 2020).

Quanto ao uso das mídias sociais, aplicativos como *WhatsApp*, *Skype* e *Facebook* foram essenciais para amenizar a solidão, facilitando a comunicação com familiares e mantendo vínculos afetivos. No entanto, os autores também apontaram riscos, como o uso excessivo e o vício em redes sociais devido à falta de limites no tempo de conexão. Essas plataformas foram avaliadas tanto como ferramentas de suporte emocional quanto como possíveis fontes de dependência (Ernesto *et al.*, 2020).

Josiowicz (2021) evidencia o *Twitter* como a plataforma de mídia social muito utilizada para leitura digital, principalmente durante o isolamento prolongado ocasionado pela pandemia da COVID-19. O aumento deste tipo de leitura, pelo *Twitter*, foi considerada pelos seus “efeitos terapêuticos” para lidar com a solidão e como espaço de mobilização política para manifestar oposição ao governo. Através de *twittes* dos usuários, principalmente no Brasil, onde estes fizeram citações de forma correta, mas também incorreta de obras dos autores Jorge Luis Borges e Clarice Lispector pode-se analisar o comportamento dessas chamadas “humanidades digitais”.

Almeida *et al.* (2021) analisaram a solidão junto ao Isolamento Social na Pandemia de COVID-19. Para estes autores, os efeitos o distanciamento social na saúde mental e física de idosos brasileiros (mais de 60 anos), com destaque para mulheres acima de 67 anos que moravam com o cônjuge e próximas aos filhos. Embora o estudo não tenha focado diretamente na solidão, os resultados indicaram piora na saúde mental durante o isolamento, sugerindo que a solidão e o uso de mídias sociais podem ser fatores relevantes para intervenções futuras.

Em contrapartida, Torres *et al.* (2020) investigaram especificamente a solidão pandêmica e pré-pandêmica em idosos (mais de 50 anos), utilizando uma escala de percepção de solidão associada a variáveis sociodemográficas e de saúde. Curiosamente, os resultados apontaram uma menor prevalência de solidão durante a pandemia em comparação ao período anterior. Os autores atribuíram esse fenômeno ao aumento das interações virtuais (via mídias sociais), que teriam mitigado os efeitos do isolamento físico.

Belens *et al.*, (2023), através de um estudo netnográfico evidenciaram a relação entre solidão, depressão e comunidades virtuais do *Facebook*. Exploraram aspectos da depressão associadas ao grupo fechado do *Facebook* voltado para pessoas com este transtorno psicológico. Os participantes descreveram-se como indivíduos que se sentiam "invisíveis" no cotidiano, mas encontravam no grupo um espaço de acolhimento e identificação. A solidão emergiu como uma subcategoria do sofrimento psíquico associado à depressão, sendo descrita como uma experiência "insuportável".

O estudo destacou o papel das mídias sociais na criação de espaços homogêneos de pertencimento, onde indivíduos marginalizados em suas interações *offline* podem reconstruir conexões sociais. O *Facebook*, nesse contexto, foi visto como um ambiente “seguro” para compartilhamento de vivências e suporte emocional (Belens *et al.*, 2023).

3.6. Evidências sobre a relação entre política e mídias sociais no Brasil

Esta seção apresenta as principais discussões acerca da relação entre política e mídias sociais explorada no contexto brasileiro nos últimos cinco anos. Como evidenciado por indicadores analisados na seção 4 no capítulo 2, a relação direta entre estas variáveis vem sendo mais investigada, onde pode ser apontado os principais fatores contextuais políticos e tecnológicos relacionados a problemática deste estudo.

Embora as mídias sociais sejam espaços propícios para engajamento, Kobellarz *et al.*, (2024) ressaltaram seus riscos políticos, caracterizando-as como uma "espada de dois gumes" que intensifica a polarização. Inicialmente vistas como ferramentas de renovação democrática,

hoje estão associadas a "bolhas de filtro" e "câmaras de eco". Deste modo, ao promover discursos incivilizados presentes em contextos mais polarizados, as plataformas de mídias sociais são espaços que têm contribuído para ascensão dos governos populistas alcançarem bolhas “neutras” (Kobellarz *et al.*, 2024).

Beer *et al.* (2024) destacam o papel das mídias sociais no cenário político brasileiro pós-manifestações de 2013, impulsionadas pelo mensalão. Movimentos anticorrupção e antipetismo ganharam força, com plataformas como *Facebook* sendo usadas para convocar protestos. Mídias alternativas, como a *Mídia Ninja*, ampliaram perspectivas ignoradas pela imprensa tradicional. O crescimento de pautas relacionadas a Jair Bolsonaro refletiu insatisfação com o petismo, pavimentando a ascensão da extrema-direita online (Beer *et al.*, 2024).

Entre 2010 e 2020, com a popularização de *smartphones* e *WhatsApp*, conteúdos curtos e de apelo emocional ampliaram o acesso à informação política (Beer *et al.*, 2024). Bolsonaro adotou estratégias semelhantes às de Trump, usando *microtargeting* e *profiling* para segmentar eleitores (Biancovilli *et al.*, 2021).

O clima de intensa polarização é persistente no país desde 2016, onde aconteceu o golpe da presidente Dilma Rousseff (Neves; Ferreira, 2021). No entanto, a polarização tornou-se evidente em 2018, com a disputa entre Bolsonaro (fim do PT) e Haddad (continuidade do PT). (Kobellarz *et al.*, 2024). A partir daí o cenário de polarização no Brasil se concentra de forma mais intensa e reduzida a dois espectros políticos: antipetismo (direita) versus antibolsonarismo (esquerda). O antipetismo traz como pautas principais, para engajar seus movimentos, o combate a “corrupção” e a defesa da “família” e dos “valores cristãos” (Neves; Ferreira, 2021).

Gonçalves *et al.* (2024) corroboram sobre o entendimento deste cenário, destacando que as mídias sociais foram palco de campanhas polarizadas que fortaleceram Bolsonaro em 2018, e contribuíram para sua derrota em 2022. O *Twitter* (X) emergiu como plataforma-chave para mobilização e análise política, embora também facilitasse campanhas negativas, especialmente com temas como a Lava Jato (Gonçalves *et al.*, 2024).

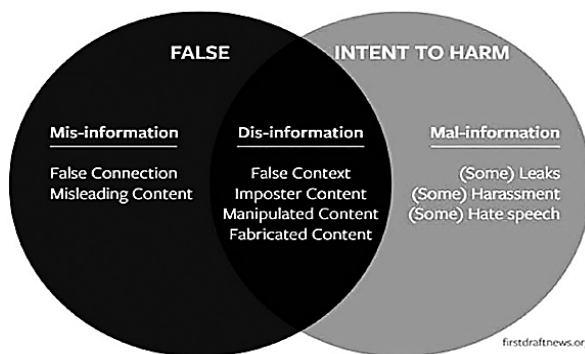
Durante a pandemia de COVID-19, o Brasil experienciou a relação de tensão entre as dimensões de saúde e economia (Ramos *et al.*, 2022), destacando-se pela má gestão da crise sanitária, influenciada pelo "efeito Bolsonaro" — postura negacionista que rejeitou evidências científicas e agravou a crise (Razafindrakoto *et al.*, 2024 e Biancovilli *et al.*, 2021). Viscardi *et al.* (2024) acrescentam que teorias da conspiração, como o movimento antivacina, foram instrumentalizadas por grupos conservadores, afetando a cobertura vacinal.

Ramos *et al.*, (2022), apontam que, nesse contexto da pandemia da COVID-19 foi potencializado o uso das mídias sociais a finalidade de alcançar e aproximar pessoas, sendo mais utilizadas (número de participantes e frequência): o *WhatsApp*, seguido do *Instagram* e *YouTube*. Contudo, práticas inadequadas ao enfrentamento da COVID-19 estão relacionadas a consumo e disseminação de notícias falsas pelo *WhatsApp*, sendo também esta plataforma, além do *Instagram* e *Facebook*, a mais utilizada para esta finalidade. Diferentes formas de se produzir conteúdo afetam o processo da população ao acesso à informação falsa. Falta de confiança na mídia social como meio de informação e mais confiança no governo são fatores ligados ao seu engajamento (Ramos *et al.*, 2022).

Braga *et al.* (2024) defendem a regulamentação das mídias sociais, que, segundo a jornalista Maria Ressa (Nobel 2010), ameaçam a democracia. A "infodemia" — desinformação disseminada deliberadamente — alimenta discursos polarizadores e manipulação (Wardle; Derakhshan, 2018 apud Braga *et al.*, 2024). A preocupação com a “infodemia” vem desde 2016, com a eleição presidenciais dos EUA, onde popularizou mais a desinformação na *internet*, utilizada como instrumento política para diminuir a crença da população em mídias tradicionais.

Ainda assim, segundo estudos, é complexo caracterizar notícias falsas, como ilustra a Figura 15, compreendendo características de conceitos que se diferenciam e se integram entre si: informações incorretas (*mis-information*), desinformação (*dis-information*) e má informação (*mal-information*) (Biancovilli *et al.*, 2021).

Figura15- Características da Desinformação, desinformação e má informação



Fonte: Wardle e Derakhsahan, (2018) adaptada por Biancovilli *et al.* (2021).

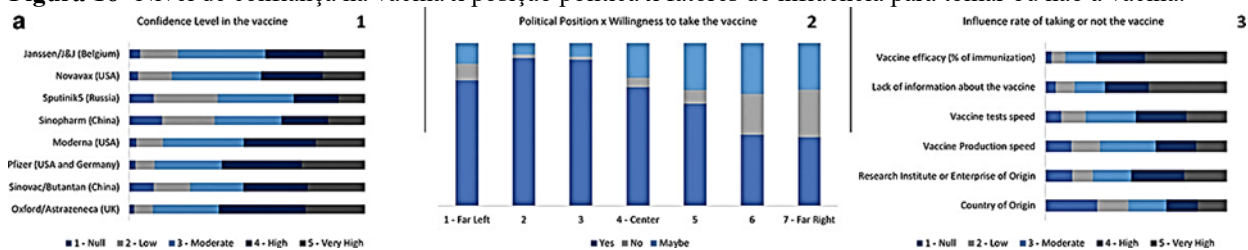
Destaca-se, nesse contexto, a questão da censura, que permeia as relações sociais há anos por meio de costumes, religião, poder político e até mesmo da lei. Atualmente, nas plataformas de mídias sociais, os usuários também podem ter suas publicações censuradas de

acordo com as políticas e termos de serviço dessas plataformas (Jornal do Comércio, 2021 apud Braga *et al.*, 2024).

Para analisar os desafios da censura e da desinformação nas mídias sociais, Braga *et al.* (2024) relacionam as áreas da Ciência da Informação e do Direito, discutindo a atuação do Judiciário brasileiro e os limites constitucionais. É essencial diferenciar liberdade de expressão, direito fundamental assegurado pela ONU em 1948, da censura, estabelecendo parâmetros legais para evitar que o discurso ilícito seja legitimado sob o pretexto de liberdade (Braga *et al.*, 2024). Os autores destacam a complexidade de identificar e combater a desinformação online, que envolve fatores subjetivos como o desejo de pertencimento, aceitação e imitação, reforçados pelas "bolhas de filtro", "redes de confiança" e "câmaras de eco". Além disso, ressaltam o uso de ferramentas computacionais, como mineração de texto, para análise de discursos (Bakir; McStay, 2017 apud Braga *et al.*, 2024).

Em outra perspectiva, Paschoalotto *et al.* (2024) investigaram a confiança na vacinação contra a COVID-19, analisando percepções, fatores influenciadores e o posicionamento político dos brasileiros. Por meio de uma pesquisa transversal online com questões sobre vacinação da COVID-19, aplicada em diversas mídias sociais, constatou-se que, além de questões como eficácia da vacina, desinformação e processo de produção, havia uma clara correlação entre posicionamento político e decisão de vacinar-se, como ilustrado na Figura 16. Indivíduos de centro-direita e extrema-direita mostraram maior relutância (com respostas predominantes de "Não" e "Talvez"), enquanto os de centro e centro-esquerda apresentaram maior adesão ("Sim"). Na extrema-esquerda, houve uma leve redução na intenção de vacinação, mas ainda com predominância de respostas positivas (Paschoalotto *et al.*, 2024).

Figura 16- Nível de confiança na vacina x posição política x fatores de influência para tomar ou não a vacina.



Fonte: Paschoalotto *et al.* (2024).

Verjovsky *et al.* (2023), aponta que a pandemia foi alvo de polarização político na comunidade *Twitter* junto discursos midiáticos populistas promovidos por antivacinas. Houve

reação dos influenciadores pró-vacina forma inarticulada que priorizaram críticas aos atores, atitudes e políticas antivacinação em detrimento da promoção das vacinas.

Lacerda e Santos (2023) investigaram o contexto empírico das universidades brasileiras e o viés neoliberal presente nos discursos sobre essas instituições nas mídias sociais. Segundo os autores, há uma disputa pela legitimação da universidade como espaço público e financiado pelo governo, intensificada pelo papel das mídias sociais. Essas plataformas permitem a manipulação do modelo cognitivo no processo comunicativo, pois, por meio de suas arquiteturas de informação, alteram a forma como os discursos de deslegitimação do ensino superior público são consumidos, reproduzidos e difundidos.

Durante sua campanha presidencial em 2018, Jair Bolsonaro propôs a privatização das universidades públicas. Esse posicionamento reflete a expansão do neoliberalismo e o avanço da mercantilização do ensino superior (Miller, 2010; Nurunnabi, 2018; Willmott, 1995). No entanto, a universidade deve ser defendida como um espaço autônomo que, além de atender a demandas de mercado, deve promover justiça social e valores democráticos (Chatelain-Ponroy et al., 2018; Clegg et al., 2011; Contu, 2018) (apud Lacerda; Santos, 2023).

Deste modo, Santos e Teixeira (2023), analisaram, durante a pandemia da COVID-19, a ação política no movimento da reforma sanitária brasileira, especialmente da Cebes e Abrasco entre os anos de 2020 a 2021. Este contexto foi marcado por um intenso retrocesso nas políticas públicas e de saúde em particular em que, além de exigir adaptação e criatividade das organizações da sociedade civil, configurou-se como uma necessidade de retomar o debate político sobre o papel do Estado e a responsabilidade social no campo da saúde. Deste modo, com as limitações imposta pelas medidas de distanciamento social, Cebes e a Abrasco utilizam novos meios de articulação política para garantir a amplificação do ativismo, sendo através da intensificação do uso meios de informação e comunicação virtual.

Santos Junior (2023) analisaram empiricamente através das mídias sociais a dieta informacional e o fluxo de informações políticas para entendimento sobre a segregação ideológica, sendo estes padrões isolados de citação de plataformas de mídias. Há um questionamento se os padrões se mantêm no longo prazo ou são condicionados por momentos de alta tensão, como eleições.

Kobellarz *et al.* (2022) analisam um fenômeno observado tanto no Brasil quanto no Canadá: a existência de domínios neutros que apresentam dados equilibrados para ambos os lados do espectro político. Esses domínios oferecem conteúdo imparcial, permitindo um entendimento comum sobre diversos temas. No entanto, os autores constataram que,

independentemente do assunto, os usuários tendem a interagir principalmente com conteúdo que reforçam suas convicções pré-existentes. O comportamento de compartilhamento (*retweets*) varia significativamente conforme o tema e a orientação política do domínio - sendo mais frequente em domínios polarizados.

Os mesmos autores destacam que, na tentativa de reduzir a polarização, plataformas e organizações têm buscado promover conteúdos neutros. Contudo, mesmo quando expostos a perspectivas contrastantes, os usuários demonstram preferência por selecionar informações que confirmem suas crenças anteriores, mantendo-se dentro de suas bolhas ideológicas (Kobellarz *et al.*, 2022).

Biancovilli *et al.*, (2021), apontam a necessidade de investigação sobre o quanto o poder de predição da orientação política (traços de conservadorismo e autoritarismo), afetam a confiança em mídias sociais e no engajamento em comportamento de saúde. Sendo considerado limitado, para os autores, o acesso a pessoas com perfil conservador, com tendência ao espectro da direita política.

No entanto, as plataformas de mídias sociais também podem ser espaço de manifestação sobre saúde mental e políticas. Josiowicz (2021) analisaram aspectos de saúde mental e comportamento sociopolíticos dos usuários de mídia social como *Twitter*, através do crescimento de leituras digitais durante o isolamento da pandemia da COVID-19. A leitura digital não foi somente um tipo de escape para enfrentamento de problemas psicoemocionais que o isolamento prolongado na pandemia gerou. Mas também como forma de mobilização e oposição política as ações do governo da extrema que não mitigou os impactos crise sanitária (Josiowicz, 2021).

A partir do entendimento de que argumentação é um dos aspectos da comunicação humana. O domínio político, de acordo com Menton *et al.* (2021), postagens do *Twitter* se caracterizam em seus aspectos textuais de argumentação pela frequência de apelo emocional negativo. Trata-se de uma plataforma reconhecido pela qualidade de credibilidade, clareza e organização. É considerado como argumento neste espaço qualquer expressão de opinião sem evidências por meio de *tweets* (Menton *et al.*, 2021).

Ainda sobre o contexto da pandemia da COVID-19 e políticas de saúde pública emerge a preocupação do campo da psicologia social e política para produção de subjetividades no enfrentamento de crise. Conforme Neves e Ferreira (2021), estão envolvidas as questões psicossociais relacionadas aos aspectos individuais e sociais da democracia.

No Brasil, a expansão do neoliberalismo e outras dinâmicas da extrema direita se deram principalmente através da redução de políticas de proteção social e bem-estar nos últimos anos. Estes autores apontam a influência do governo aos cuidados da saúde pública, dando mais fama e expectativas sobre o uso de medicamentos (hidroxicloroquina e cloroquina) para cura de doenças até então de causas desconhecidas e complexas como a COVID-19. Este cenário foi marco por disputas de verdades e de narrativas, de um modo que os ativistas de cloroquina buscavam desmoralizar os produtores de conhecimento (Neves; Ferreira,2021).

CAPÍTULO 4

4. METODOLOGIA

Neste capítulo, são apresentados os procedimentos metodológicos adotados para alcançar os objetivos deste estudo e responder suas questões de pesquisa. A exposição da metodologia foi organizada em quatro seções: a Seção 4.1 aborda a interdisciplinaridade do método; a Seção 4.2 caracteriza a IA como um campo interdisciplinar; a Seção 4.3 faz a classificação do método de IA utilizado; e a Seção 4.4 descreve os materiais e métodos. A metodologia utilizada tem sua justificativa na complexidade da problemática e da relação entre as variáveis envolvidas, havendo mais lacunas que *insights* na literatura. Os resultados gerados por estes procedimentos foram analisados e discutidos neste capítulo e no Capítulo 5.

4.1. Interdisciplinaridade do método de pesquisa

Este estudo adota uma abordagem interdisciplinar para sua construção metodológica, integrando a IA e psicologia cognitivo-comportamental aplicados para entendimento de três fenômenos sociais e se estão interrelacionados: solidão, comportamento político e uso de mídias sociais no eleitorado brasileiro, caracterizando um problema complexo envolvendo múltiplas variáveis.

Para enfrentar os desafios contemporâneos em sua complexidade integral, Morin (2005) aponta a necessidade de uma profunda reforma de pensamento, considerando urgente a conexão entre os campos do conhecimento para construir o que ele define como a verdadeira interdisciplinaridade. Trata-se de uma abordagem integradora através da qual se combinam diferentes campos, até mesmo os mais distintos como as ciências da materialidade e as ciências humanas a nível de indivíduo. Esta comunicação, que ultrapassa fronteiras disciplinares, é importante para resolução de problemas da realidade contemporânea que se caracteriza como complexa e híbrida (Raynaut, 2011).

Conforme Júnior e Neto (2011), essa convergência entre ciências exatas e humanas promove uma comunicação construtiva entre áreas distintas, permitindo a análise de dados empíricos e modelos teóricos, visando superar as fronteiras rígidas entre esses campos do conhecimento.

O método empregado organiza tanto a coleta de dados quanto a modelagem teórica, diferenciando os campos de observação para identificar, analisar e combinar as propriedades específicas de cada área, articulando ou confrontando suas características (Júnior; Neto, 2011).

Raynaut (2011) aponta que esta colaboração entre ciências da materialidade e ciências humanas é um desafio, mas muito almejada no desenvolvimento de pesquisa interdisciplinares

Morin (2005) aponta que no mundo atual o conhecimento se fragmenta em disciplinas incapazes de dialogar entre si, de modo que o saber compartimentalizado tem dificultado a compreensão a unidade dos fenômenos complexos e desafios sociais. A proposta interdisciplinaridade adotada tem criado apenas especialistas que reafirmam as fronteiras existentes do conhecimento. A construção do método científico, portanto, deve abandonar a ilusão do conhecimento fragmentado para construir conceitos migrantes que rescindam estas fronteiras, pensando a partir das emergências sistêmicas (Morin, 2005).

4.1.1. Inteligência Artificial: Um campo em diálogo interdisciplinar com a psicologia

A IA é um campo interdisciplinar que integra conhecimentos de neurociência, psicologia, computação e matemática para desenvolver sistemas capazes de aprendizagem (Russell; Norvig, 2022). Trata-se de um campo que vem florescendo na atualidade com a revolução dos dados (Russell; Norvig, 2022), trazendo inúmeros benefícios para resolução de problemas em diversos outros campos, mas também medos e receios quanto aos riscos (Mitchell, 2019) quanto ao processamento de dados através de algoritmos enviesados e pouco adaptado para a diversidade social (O'neil, 2021).

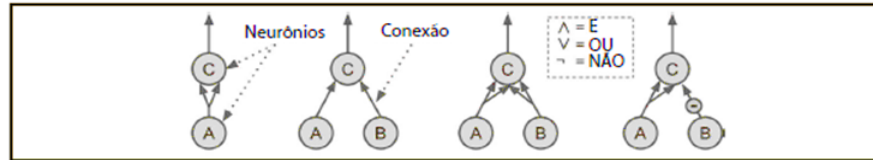
O uso de modelos de IA foi regulamentado no Brasil em 2024 (Senado Federal, 2024), sendo instituídas estratégias brasileiras de Inteligência Artificial (Portaria MCTI nº 4.979/2024), com o objetivo de fomentar a inovação e o desenvolvimento de soluções que garantam seu uso ético e alinhado ao progresso social (Ministério da Ciência, Tecnologia e Informação, 2021).

A IA baseia-se em dois pilares essenciais: os processos cognitivos humanos (inteligência) e sua implementação computacional (Russell; Norvig, 2022). Têm auxiliado na tomada de decisão frente a problemas cada vez mais complexos na atualidade, especialmente na área da saúde (Hamet; Tremblay, 2017). São utilizados algoritmos para execução de tarefas com a finalidade solucionar estes problemas, através de instruções (Nadikattu, 2016).

Algoritmos de IA aprendem com dados fornecidos (Nadikattu, 2016). Tratam-se de modelos matemáticos inspirados no funcionamento neural biológico, simulando o processamento de informações para resolução de problemas complexos (Russell; Norvig, 2022). O neurônio artificial, conceito pioneiro de McCulloch e Pitts (Figura 17), replica o

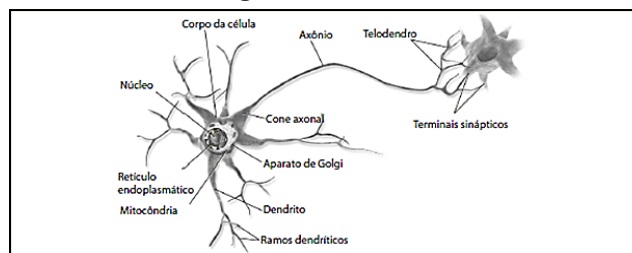
comportamento básico dos neurônios humanos (Figura 18) que se conectam em uma rede neural biológica processando informações (Figura 19).

Figura 17- neurônio artificial



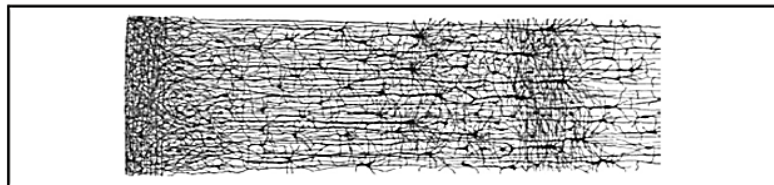
Fonte: Géron (2021).

Figura 18- Neurônio humano



Fonte: Géron (2021).

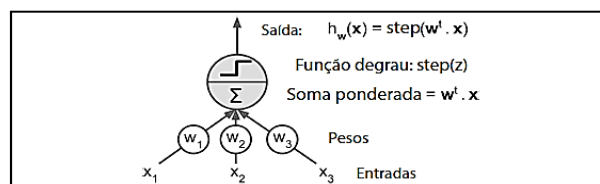
Figura19- Múltiplas camadas em uma rede neural biológica (córtex humano)



Fonte: Géron (2021).

A evolução dos modelos incorporou contribuições da psicologia: inicialmente pelo behaviorismo (relação estímulo-resposta) e posteriormente pela psicologia cognitiva, que compreende o cérebro como sistema de processamento de informações (Russell; Norvig, 2022). Arquiteturas mais simples, como o *Perceptron* de Rosenblatt (1957) (Figura 20), demonstraram como unidades lineares podem processar dados de entrada para gerar resultados, estabelecendo os fundamentos das modernas redes neurais artificiais (Géron, 2021).

Figura 20-Neurônio artificial de Frank Rosenblatt (1957)



Fonte: Géron (2021).

Atualmente a IA é considerada como um ramo da engenharia (Hamet; Tremblay, 2017), sendo integrada a vários setores da sociedade, principalmente na saúde, auxiliando na resolução de problemas cada vez mais complexos (Nadikattu, 2016).

4.2. Classificação do método de IA

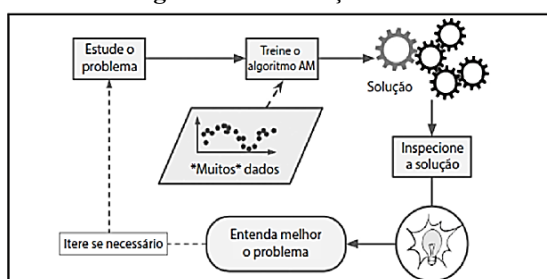
A IA possui métodos que auxiliam a lidar com a nova realidade dos dados. A partir disso, tornou-se possível tratar com problemas complexos que envolvem quantidade de dados produzidos pelas pessoas na atualidade. Um destes métodos ou subcampos da IA é o *ML*, implementado nesta pesquisa (Géron, 2021).

4.2.1. Método *Machine learning*

Trata-se de um método que permite que sistemas computacionais aprendam automaticamente a partir de dados, identificando padrões complexos e correlações não evidentes para resolver problemas (Géron, 2021). Se destaca pela capacidade de analisar grandes volumes de dados através de algoritmos matematicamente fundamentados, extraindo informações sem ou com pouca intervenção humana contínua (Heineman; Pollice; Selkow, 2009).

Assim, ainda de acordo Géron (2021), poderão ser reveladas correlações não esperadas e novas tendências que levarão a formas novas de compreender tais problemas. Isso significa que podem ser aplicadas em grandes e novas quantidades de dados com a finalidade de descobrir padrões não aparentes, através do que é denominado mineração de dados (Figura 21) (Géron, 2021).

Figura 21- Mineração dos dados



Fonte: Géron (2021).

Conforme Géron (2021), os algoritmos podem ser categorizados segundo três dimensões principais: quanto ao nível de supervisão (incluindo métodos supervisionados, não supervisionados, semi-supervisionados e por reforço), quanto ao fluxo de processamento (aprendizado online ou em lotes) e quanto à metodologia (baseada em instâncias ou em modelos preditivos).

Neste estudo, adota-se uma abordagem *ML* híbrida, utilizando aprendizado não supervisionado para exploração inicial dos dados e técnicas supervisionadas para análises preditivas mais refinadas, permitindo tanto a descoberta de padrões intrínsecos quanto a validação de hipóteses específicas (Mitchell, 1997; Géron, 2021). Esta combinação metodológica oferece uma compreensão abrangente dos dados, destacando o potencial do *ML* para aplicações práticas em pesquisa e tomada de decisão.

Neste estudo, para a análise dos dados subjetivos foram utilizados modelos de agrupamento de *ML*. O Quadro 5 mostra a classificação do treinamento quanto ao tipo de técnica de treinamento, algoritmo e finalidade de seu uso neste estudo. Em seguida, uma descrição mais detalhada de como funciona cada um deles. Todos os parâmetros dos algoritmos treinados foram definidos com base empírica para análise dos dados que tem como atributos analisados objetos abstratos que representam pessoas e casos reais, justificando destas técnicas estatísticas computacionais para observar o comportamento de relação entre eles.

Quadro 6- Classificação dos modelos de agrupamento *ML* aplicados neste estudo

| Treinamento | Algoritmo | Técnicas e objetivo | Finalidade na pesquisa |
|--------------------|---|---|--|
| Não supervisionado | <i>K-means</i> | Agrupamento dos dados não rotulados através de clusterização baseada em proximidade por similaridade com os centroides com o objetivo de explorar agrupamentos naturais em subconjunto inicial (n=150) | Utilizado como etapa exploratória iniciação para identificação de perfis comportamentais distintos com base em variáveis demográficas, socioeconômicas, políticas e de uso de mídias sociais. Serviu de base para selecionar variáveis relevantes a serem usadas nas etapas seguintes. |
| | <i>Density-Based Spatial Clustering of Applications with Noise</i> ou Agrupamento Espacial Baseado em Densidade de Aplicações com Ruído (<i>DBSCAN</i>) | Agrupamento de dados com base em pontos de densidade local através de clusterização, usando dois parâmetros de distância com o objetivo de detectar agrupamentos por densidade e identificar <i>outlier</i> . | Identificação de padrões de agrupamento baseados em densidade nas associações entre as principais variáveis, evidenciando os comportamentos compartilhados na formação de bolhas, atribuindo-lhes rótulos para aplicação do método supervisionado (KNN) Também foi utilizado para testagem das hipóteses |
| Supervisionado | <i>k-Nearest Neighbours</i> ou <i>K-vizinhos</i> mais próximos (<i>KNN</i>) utilizado. | - Predição da severidade das variáveis, atribuindo uma classe com base em padrões aprendidos em dados anteriores com o objetivo de classificar as variáveis por níveis de severidades com base em similaridade entre os perfis. | Previsão de categorias das variáveis solidão, política e mídias sociais com base nos agrupamentos de severidade identificados pelo <i>DBSCAN</i> e validação das hipóteses. |

Fonte: adaptado de Géron (2021); Mitchell (1997); Han et.al. (2011).

4.2.1.1. Treinamento não supervisionado com algoritmos de clusterização

No treinamento não supervisionado os dados disponibilizados para o treinamento dos algoritmos não são rotulados, isto é, não possuem classificação com base em resposta ou rótulos pré-definidos. Os algoritmos aprenderam de forma autônoma a encontrar conexões entre os dados para formação dos agrupamentos (Géron, 2021).

Uma das mais importantes tarefas executadas pelos algoritmos deste tipo é a clusterização ou agrupamento de atributos dos dados pela distância entre os pontos, com base na classificação destes por sua semelhança ou aproximação (Géron, 2021). Os *clusters* formados por dados abstratos que representam os sujeitos estudados devem apresentar alta homogeneidade interna, ou seja, alta similaridade entre os pontos de um mesmo grupo. Também, devem apresentar alta heterogeneidade externa, ou seja, a diferença significativa em relação aos demais *clusters* (Han et al. 2011). Dentre os algoritmos que executam esse tipo de tarefa se encontram o *K-means* (Géron, 2021) e o *DBSCAN* (Han et al., 2011).

O *K-means* foi utilizado na etapa exploratória inicial para identificação de perfis comportamentais distintos com base em variáveis demográficas, socioeconômicas, políticas e de uso de mídias sociais. Serviu de base para selecionar variáveis relevantes a serem usadas nas etapas seguintes, com a finalidade de visualização do comportamento de agrupamento natural entre os dados e *insights* para implantação de modelos mais eficazes para lidar com as características dos dados disponibilizados (Géron, 2021).

O *K-means* agrupa dados, buscando fazer a separação de amostras em n agrupamentos de variância semelhante. Dessa maneira, minimiza a inércia, que é calculada por meio da soma das distâncias quadradas de cada ponto, que deve resultar em um valor menor possível na etapa escolha do *centroíde* de seu *cluster*¹. Neste algoritmo os *clusters* precisam ser de formato convexos e isotrópicos, (Scikit-learn, 2025 cap.2). ²Busca-se instâncias centradas em torno de um ponto específico denominadas *centróides*, podendo ser visualizadas no diagrama de *Voronoi* que se trata da divisão do conjunto de dados entorno de pontos centrais como a saída dos dados gerada pelo algoritmo (Géron, 2021). Neste estudo, o *k-means* foi utilizado.

Já o *DBSCAN* foi utilizado para agrupamento com base na densidade de dados que estão próximos por similaridade. Diferentemente do *K-means*, esse algoritmo não requer a definição prévia do número de *clusters*. Trata-se de um modelo capaz de identificar grupos com formas

¹ Equação para soma dos quadrados dentro do *cluster*: $\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|^2)$, onde N amostras X em K aglomerados disjuntos C , cada um descrito pela média μ_j das amostras no *cluster*.

² Dados convexos e isotrópicos

arbitrárias, uma vez que sua entrada tem um parâmetro reduzido e maior capacidade de lidar com *outliers*. (Han *et al.*, 2011). Neste estudo, este algoritmo foi utilizado verificar os padrões de agrupamentos baseados em densidade, encontrados na associação entre as variáveis são confiáveis para evidenciar padrões comportamentais compartilhados. A partir dos clusters formados por este modelo, foram identificados os grupos formados com base na densidade dos dados de associação entre as variáveis por severidade e atribuindo-lhes rótulos aos dados para treinamento do KNN.

4.2.1.2. Treinamento supervisionado com o KNN para classificação dos grupos identificados

No aprendizado supervisionado são disponibilizadas alternativas desejadas (rótulos) aos algoritmos, atribuindo a eles dois tipos de tarefas típicas: classificação e predição. Na classificação é feita a previsão de classes conforme os atributos (variáveis) dos dados disponibilizados. Já na predição é utilizado a modelo regressão, principalmente quando se trata de valores (Géron, 2021). Dentre os algoritmos que executam este tipo de técnica, foi implementado inicialmente para análise da associação entre as variáveis neste estudo o KNN.

O algoritmo KNN é um modelo para reconhecimento de padrões baseado em instâncias correspondentes a pontos em um espaço n-dimensional. A definição dos vizinhos mais próximos de novos exemplos é estabelecida por meio da distância *euclidiana* que é calculado através da equação $\|x - \mu_i\|^2$.³(Mitchell, 1997). É uma abordagem considerada simples que pode aprender características representadas tanto por atributos qualitativos quanto os quantitativos (Han *et al.*, 2011; Batista, 2003).

Esse algoritmo analisa a proximidade entre atributos dos dados, realizando predições de classificação ou regressão com base nos K-vizinhos mais próximos de uma entrada, utilizando exemplos do conjunto de treinamento. Após a apresentação de um novo exemplo, um conjunto de casos semelhantes é recuperado e utilizado para sua classificação, tudo isso sem a necessidade de gerar regras explícitas (Han *et al.*, 2011; Batista, 2003).

Trata-se de um método (*Instance-based*) cujo processamento é por analogia ou comparações através da classificação de exemplos novos. Para contribuição de cada um dos K-*vizinhos* é dado um peso conforme a distância entre vizinhos e o candidato, sendo peso maior para os próximos e peso menor para os mais distantes (Batista,2003; Han *et al.*, 2011).

³ Distância Euclidiana representada matematicamente pelo cálculo: $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ ou $\|x - \mu_i\|^2$

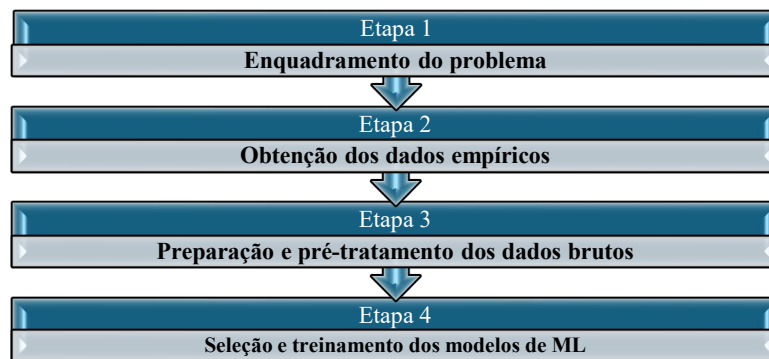
Este algoritmo teve como finalidade de exploração dos agrupamentos identificados através da previsão de categorias das variáveis por níveis de severidade com base na proximidade de sujeitos com características semelhantes. Também através dele, foram testadas as hipóteses deste estudo.

4.3. Materiais e métodos

Para aplicação das técnicas de *ML*, seguiu-se o método clássico encontrado na literatura de Géron (2021). Este método computacional foi adaptado para a pesquisa científica deste trabalho, sendo integradas a ele abordagens qualitativas e contextualizadas para compreensão de um problema que envolve variáveis complexas. Este método foi antecedido de uma revisão sistemática de literatura com análise estatística e bibliométrica para compreensão do tema.

A Figura 22 mostra o fluxograma das etapas gerais seguidas para o desenvolvimento dos modelos *ML* que envolvem: I) análise do panorama da problemática através da literatura; II) obtenção dos dados empíricos por meio de abordagens psicológicas; III) identificação e visualização dos dados brutos; IV) preparação dos dados brutos; V) seleção e treinamento dos modelos *ML*. Estas etapas serão detalhadas nas próximas seções.

Figura 22: Fluxograma das etapas gerais de aplicação do *ML*



Fonte: baseado em Gerón (2021).

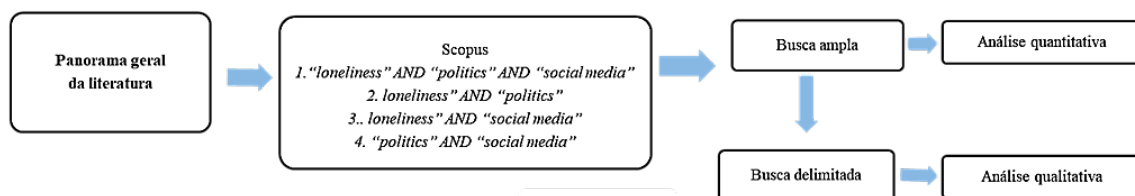
4.3.1. Análise do panorama geral da problemática através da literatura

Esta seção apresenta processo que forneceu conhecimento significativo para compreensão da temática e, posteriormente, contribuiu para o enquadramento do problema de pesquisa. Fez-se uma pesquisa de revisão sistemática de literatura para análise panorâmica que integrou a análise quantitativa e qualitativa, com base em Donato, H. e Donato, M. (2019). Justifica-se esta revisão como parte importante que antecede a pesquisa empírica, visando auxiliar na compreensão, validação e determinação das variáveis envolvidas na problemática investigada (Marconi; Lakatos, 2017).

Através desta revisão, foi traçado um panorama geral sobre os estudos que relacionam as três variáveis principais (solidão, política e mídias sociais), no qual fez-se o levantamento e análise de estudos que evidenciam a conexão entre as variáveis para percepções estruturais, abrangentes, estatísticas dos fatores relacionados e característica desta conexão em um contexto específico.

Como mostram as etapas do estudo do panorama geral da literatura (Figura 23), em virtude do tempo hábil para realização deste estudo, utilizou-se apenas a base de dados *Scopus* para a realização das buscas, devido à sua abrangência e confiabilidade como base de dados. E dada a complexidade da temática explorada, as buscas e a análise dos resultados foram conduzidas, seguindo em etapas distintas e uma abordagem quali-quantitativa.

Figura 23- Etapas do estudo sobre o panorama geral da literatura



Fonte: Elaborado pelos autores.

Na primeira etapa, buscou-se a relação entre as três variáveis investigadas (solidão, política e mídias sociais), utilizando a combinação dos termos-chave em inglês: *"loneliness" AND "politics" AND "social media"*. Para análise dos dados, utilizaram-se termos amplos nas buscas — solidão, política e mídias sociais — para maximizar a cobertura de estudos sobre sua relação integrada, dada a raridade de pesquisas que abordam a solidão como causa primária de problemas midiáticos e políticos. Adotou-se 'política relacional' (sem especificar tipos de relações institucionais ou comportamentais) e 'mídias sociais' (sem delimitar plataformas), garantindo abrangência na análise.

Adotou-se uma abordagem quali-quantitativa para análise dos dados, permitindo a identificação e classificação das variáveis mais estudadas, bem como das tendências e critérios da produção acadêmica sobre o tema.

Com base nos resultados iniciais, novas buscas foram realizadas, combinando os termos em pares: *"loneliness" AND "politics"*, *"loneliness" AND "social media"* e *"politics" AND "social media"*. Foram realizadas inicialmente de forma ampla (sem filtros) e posteriormente, foram delimitadas (com aplicação de filtros). Nas buscas delimitadas, seguiu-se os critérios de

inclusão e exclusão previamente estabelecidos, priorizando estudos atuais que trazem enfoques para o contexto brasileiro.

As buscas foram realizadas e datadas no dia 12 de outubro de 2024, sendo os resultados apresentados por meio de gráficos gerados a partir das estatísticas da própria base de dados, destacando os estudos encontrados. Em seguida, os mesmos dados exportados da base de dados foram analisados por meio do *software VOSviewer*, gerando mapas bibliométricos que exploram os termos-chave da pesquisa presentes nos títulos, resumos e palavras-chave dos estudos.

Estes dados quantitativos foram posteriormente analisados de forma descritiva. Já as buscas delimitadas para aprofundar a análise no cenário brasileiro nos últimos cinco anos. Incluíram a aplicação de filtros da própria *Scopus* para inclusão de estudos relacionados. A interpretação dos resultados desta busca foi feita através de uma análise qualitativa, alinhada às necessidades e especificidades deste estudo.

4.3.1.1. Análise dos resultados das buscas gerais

As buscas amplas têm como finalidade trazer evidências sobre o panorama geral (contexto global e histórico de publicações) da temática. Garantiu-se que os títulos, resumos e palavras-chave dos documentos incluíssem os termos-chave de forma parcial ou total. Os resultados dessa busca foram analisados descritivamente, apresentando a evolução global do tema, conforme exposto no capítulo 2. Para isso, não foram aplicados filtros quanto a países, anos, áreas de estudo, idiomas, tipos de documentos ou acesso, entre outros critérios.

Na Tabela 5, mostram-se os resultados das buscas amplas. Na primeira Busca Geral (BG1), utilizando a combinação dos termos '*loneliness*', '*politics*' e '*social media*', foram encontrados 6 documentos, dos quais 5 foram incluídos para análise, por estarem finalizados e com acesso aberto no momento da busca. Na segunda Busca Geral (BG2), ao relacionar os termos '*loneliness*' e '*politics*', foram encontrados e incluídos 164 documentos para análise. Na terceira Busca Geral (BG3), evidenciou-se a relação entre '*loneliness*' e '*social media*', que tem sido mais explorada em comparação às BG1 e BG2, com um total de 1.165 estudos publicados em âmbito mundial incluídos para análise. Por fim, na quarta Busca Geral (BG4), destacou-se a forte conexão entre '*social media*' e '*politics*', com 6.105 documentos publicados e incluídos para análise. Esse volume evidencia o crescente interesse acadêmico pelo tema.

Tabela 5- Resultados da busca geral do panorama geral da literatura

| Busca geral na base Scopus | Termos combinados | Resultados quantitativos |
|----------------------------|---|--------------------------|
| Busca geral 1 | "loneliness" and "politics" and "social medias" | 6 documentos |
| Busca geral 2 | "Loneliness" and "politics" | 164 documentos |
| Busca geral 3 | "loneliness" and "social medias" | 1165 documentos |
| Busca geral 4 | "politics" and "social medias" | 6105 documentos |

Fonte: elaborado pelos autores.

Os dados exportados da *Scopus* através das buscas amplas foram analisados de forma quantitativa, utilizando o *Excel* para a criação de gráficos e tabelas e o *Vosviewer* para análise bibliométrica.

Na análise bibliométrica, com o auxílio do *software VOSviewer*, seguiu-se a proposta de Zupic e Čater (2015), através da qual fez-se o mapeamento e a visualização dos indicadores relacionados às três variáveis do estudo e suas interações. Foram aplicadas medidas específicas para investigar padrões de ocorrência e conexões entre os termos presentes nos títulos, resumos e palavras-chave.

Os mapas bibliométricos gerados pelo *VOSviewer* foram fundamentais para identificar padrões de conexões conceituais entre as temáticas exploradas nos estudos. Nesses mapas, é possível analisar os termos ou variáveis mais pesquisadas (ocorrência) e classificar a relação entre eles na literatura (força de ligação) (Zupic; Čater,2015).

A ocorrência é representada pelo tamanho dos clusters (grupos de palavras-chave), onde clusters maiores indicam maior frequência de determinada palavra-chave nas pesquisas, e clusters menores indicam menor frequência. A força de ligação é representada por linhas coloridas que conectam as palavras-chave, indicando a intensidade da relação entre elas. Essa análise auxilia no entendimento dos indicadores sobre como o tema está sendo abordado na literatura, conforme apresentada no Capítulo 2.

4.3.1.2. Análise dos resultados das buscas delimitadas

Foi feita uma análise qualitativa com objetivo de aprofundar a compreensão dos estudos brasileiros sobre a relação entre solidão, política e mídias sociais, destacando as principais características das pesquisas mais recentes. Para tanto, realizaram-se buscas delimitadas mediante a aplicação de filtros específicos nas bases de dados, seguindo critérios de inclusão e exclusão predefinidos. Foram incluídos para análise: estudos brasileiros; artigos científicos; publicados nos últimos cinco anos; concluídos e com acesso aberto.

Organizou-se um quadro analítico contendo as principais características dos estudos que contribuem para entendimento da temática no Brasil, incluindo: Variáveis principais, covariáveis, lacunas e insights. A interpretação dos resultados considerou as necessidades específicas desta pesquisa, servindo como base para a fundamentação teórica apresentada no Capítulo 2. Como abordagem de análise destes estudos, fez-se a leitura preliminar de títulos e resumos. A seleção dos artigos foi a partir da análise se estes atendiam aos critérios específicos, tais como foco no contexto brasileiro recente (últimos 5 anos); que abordam a solidão como construto de saúde mental; cujos participantes são adultos (maiores de 18 anos); traz enfoque na política relacional e cenário eleitoral polarizado. Posteriormente, fez-se a leitura integral dos textos selecionados.

A Tabela 6 apresenta os resultados das buscas delimitadas, devido ao tempo hábil de realização da pesquisa e para análise mais aprofundada do contexto específico. Primeira busca, relacionando os termos "*loneliness*" AND "*politics*" AND "*social media*" resultou em apenas 1 artigo qualificado. A segunda busca relacionando os termos: "*loneliness*" AND "*politics*" resultou inicialmente em 3 documentos brasileiros. Porém, nesta busca foi excluído 2 artigos de revisão não aplicável ao contexto brasileiro de polarização política atual, sendo selecionados apenas 1 artigo para análise aprofundada. Na terceira Busca, relacionando os termos: "*loneliness*" AND "*social media*" resultou inicialmente em 23 documentos. No entanto, foram excluídos os artigos que não atendiam os critérios estabelecidos, tais como estudos não focados no Brasil (5); revisões sem enfoque brasileiro (5); duplicatas (1); diretrizes (1); artigos sem acesso completo (1); editoriais (1); cartas (1) e estudos com adolescentes (3). Nesta busca foram selecionados apenas 5 artigos para análise. Na quarta busca, relacionando os termos "*politics*" AND "*social media*", busca atualizada no dia 18 de dezembro de 2024, teve como resultados iniciais 39 artigos. Nesta etapa, foram excluídos os artigos que não atendiam os critérios estabelecidos, tais como os que não abordavam o cenário brasileiro (2) e não tinham foco na dimensão relacional ou saúde mental (4); artigos que não trazem enfoque direto ou nas eleições presidenciais polarizadas do Brasil nos últimos anos (12) artigos que trazem mais enfoque na ferramenta de IA, analisando informações textuais nas mídias sociais ou fora delas (4), sendo selecionados 17 artigos para análise qualitativa aprofundada.

Deste modo, foram selecionados de maneira geral, através das buscas realizadas, o total de 24 artigos, estudos brasileiros sobre a temática explorada, para análise qualitativa e aprofundamento.

Tabela 6- Resultados das buscas delimitadas sobre o cenário brasileiro nos últimos cinco anos

| Busca delimitada na base Scopus | Termos combinados | Resultados quantitativos |
|--|---|---------------------------------|
| Busca delimitada 1 | “loneliness” and “politics” and “social medias” | 1 artigo |
| Busca delimitada 2 | “Loneliness” and “politics” | 1 artigo |
| Busca delimitada 3 | “loneliness” and “social medias” | 5 artigos |
| Busca delimitada 4 | “politics” and “social medias” | 17 artigos |
| Total de artigos selecionados= 24 artigos | | |

Fonte: elaborado pelos autores.

Os resultados demonstram ainda baixo interesse acadêmico para exploração da conexão entre as variáveis de forma integrada e a relação direta ou mediadora entre solidão com as demais variáveis (política ou mídias). Porém, é crescente interesse acadêmico na intersecção entre política e mídias sociais, particularmente no contexto brasileiro recente. A análise qualitativa desses estudos permitiu identificar padrões e particularidades relevantes para a compreensão do fenômeno investigado, bem como as principais variáveis envolvidas, as lacunas e os *insights* desses estudos, conforme exposto nos Capítulos 2 e 3.

4.3.2. Obtenção do conjunto de dados empíricos

Nesta seção, apresenta-se o método utilizado para levantamento do conjunto de dados empíricos, sendo estes dados comportamentais, sociodemográficos e socioeconômicos para treinamento dos modelos de *ML*, buscando evidenciar a relação entre as variáveis.

Neste estudo, os dados são objetos abstratos que representam pessoas a serem agrupadas. Foram coletados dados sobre percepção dos níveis de solidão e comportamentais sobre o engajamento político e uso de mídias sociais através de um estudo empírico, exploratório de abordagem quali-quantitativa, o qual se deu por etapas (Marconi; Lakatos, 2017): i) definição da amostragem ii) elaboração e pré-teste dos instrumentos e iii) aplicação dos instrumentos.

Este projeto foi submetido e aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Itajubá, sob o Parecer nº 7.923.684 e CAAE: 76839323.3.0000.0356, em 24 de outubro de 2025. A pesquisa foi conduzida em estrita conformidade com as diretrizes e normas regulamentares de pesquisas envolvendo seres humanos, seguindo a Resolução nº 510/2016 do Conselho Nacional de Saúde (anexo A).

De acordo com Marconi e Lakatos (2017), justifica-se o delineamento da pesquisa empírica pela busca de respostas para a problemática estudada, visando levantar dados que auxiliem a comprovar hipóteses, descobrir novos fenômenos e estabelecer relações entre eles.

A finalidade desta pesquisa é coletar de dados diretamente com os sujeitos pesquisados, no local onde os fenômenos ocorrem.

Este tipo de investigação tem caráter exploratório-descritivo, buscando descrever as relações entre as variáveis, as características do fenômeno e da população. Através de uma abordagem que combina análises empíricas e teóricas, permite a descrição do fenômeno por meio de métodos qualitativos e quantitativos, conforme o acúmulo de informações obtidas (Marconi e Lakatos, 2017).

4.3.2.1. Identificação do conjunto de dados empíricos

Nesta seção, apresenta-se a amostra estudada, buscando representatividade de forma aleatória e pela probabilidade de integração para refletir o fenômeno estudado e seus vieses. O contexto empírico deste estudo é o segundo turno das eleições para presidente do Brasil em 2022, de modo que participaram desta pesquisa 176 pessoas, dos quais, inicialmente, foram identificados o seu perfil através das variáveis sociodemográficas, socioeconômicas, uso de mídias sociais e perfil político. Posteriormente, foram coletados dados sobre a percepção dos níveis de solidão, autopercepção do engajamento político *offline* ou locais e nas mídias sociais durante o segundo turno das eleições para presidente do Brasil em 2022.

Embora a proposta seja o treinamento de modelos para análise do eleitorado no Brasil, a pesquisa possui limitação de abrangência na amostra, apresentando alcance de respondentes apenas de uma microrregião, sendo esta a cidade de Itajubá e circunvizinhanças, o que constitui uma restrição metodológica que impede a generalização dos resultados. Deste modo, cabe enfatizar que os modelos não aprenderam sobre o eleitorado do Brasil, mas sobre padrão do eleitorado de uma microrregião específica. O *Dataset* desenvolvido com os dados brutos coletados sobre níveis de solidão, comportamento político e uso de mídias sociais se encontra disponível para consulta no Apêndice D. As características dos respondentes serão apresentadas nas seções a seguir.

4.3.2.1.1. Características sociodemográfica para a relação entre solidão, política e mídias sociais

Conforme mostra a Tabela 7, que mostra informações sobre a variável sociodemográfica, no que se refere às variáveis sociodemográficas, foi feita a identificação dos participantes quanto à faixa etária, gênero, etnia e região em que residem. A maior parte da amostra (82%) reside em Itajubá e em áreas urbanas. A faixa etária variou entre 18 e mais de

60 anos, sendo possível observar apenas pessoas adultas participaram deste estudo, prevalecendo a participação de pessoas entre 31 e 44 anos (33%). Quanto ao gênero, houve prevalência de participação do gênero feminino, sendo representado por 57,10% da amostra. O gênero masculino teve a segunda maior quantidade de participantes (42,90%) dos participantes. Já pessoas que se identificam com outros gêneros não tiveram representação. Houve participação de pessoas de várias etnias, porém havendo uma maior representação de pessoas que se autodeclararam brancas (76%) e posteriormente pardos (18,30%). Já em relação ao estado civil, 46,6% se identificam como solteiros.

Tabela 7- Variáveis sociodemográfica

| Variáveis | Número | Porcentagem |
|---------------------------|---------------|--------------------|
| Faixa etária | | |
| 18 anos | 4 | 2,3% |
| 19 a 24 anos | 6 | 3,4% |
| 25 a 30 anos | 4 | 2,3% |
| 31 a 44 anos | 58 | 33% |
| 45 a 54 anos | 28 | 15,9% |
| 55 a 60 anos | 13 | 7,4% |
| 60 anos | 2 | 1,1% |
| Mais de 60 anos | 10 | 5,7% |
| Gênero | | |
| Feminino | 101 | 57,4% |
| Masculino | 75 | 42,6% |
| Outros | 0 | 0% |
| Etnia | | |
| Branco | 133 | 75,6% |
| Indígena | 1 | 0,6% |
| Pardo | 32 | 18,2% |
| Preto | 10 | 5,7% |
| Estado civil | | |
| Solteiro | 82 | 46,6% |
| Casado (a) | 67 | 38,1% |
| Divorciado (a) | 17 | 9,7% |
| Outros | 10 | 5,7% |
| Natural de Itajubá | | |
| Sim | 120 | 68,2% |
| Não | 56 | 31,8% |
| Reside em Itajubá | | |
| Sim | 145 | 82,4% |
| Não | 31 | 17,6% |

Fonte: elaborada pelos autores.

4.3.2.1.2. Características socioeconômicas para a relação entre solidão, político e mídias sociais

A Tabela 8 apresenta informações sobre a variável socioeconômica, sendo identificado o perfil de escolaridade dos participantes e a renda individual. 24% da amostra é representada por pessoas que possuem pós-graduação, seguido de pessoas com ensino superior incompleto (21,10%). Em relação à renda financeira aproximada, houve bastante variação quanto ao valor

da renda de cada participante. A maior parte da amostra (21,6%) possui renda de até 2 salários-mínimos, seguidas de pessoas que possuem uma renda superior a 7 salários-mínimos (18,80%).

Tabela 8- variáveis socioeconômicas

| Variáveis | Número | Porcentagem |
|---|---------------|--------------------|
| Escolaridade | | |
| Ensino fundamental completo | 0 | 0 |
| Ensino Fundamental incompleto | 0 | 0 |
| Ensino médio completo | 15 | 8,5% |
| Ensino médio incompleto | 2 | 1,1% |
| Ensino superior completo | 30 | 17% |
| Ensino superior incompleto | 37 | 21% |
| Pós-graduação | 43 | 24,4% |
| Mestrado | 23 | 13,3% |
| Doutorado | 25 | 14,2% |
| Renda individual | | |
| Menos de 1 salário-mínimo (Menos de R\$ 1412) | 22 | 12,5% |
| 1 salário-mínimo (até R\$ 1.412) | 11 | 6,3% |
| De 1 a 2 salários-mínimos (até R\$ 2.824) | 38 | 21,6% |
| De 2 a 3 salários-mínimos (até R\$4.236) | 24 | 13,6% |
| De 3 a 4 salários-mínimos (até R\$ 5.548) | 22 | 12,5% |
| De 4 a 5 salários-mínimos (até R\$ 7.060) | 24 | 13,6% |
| De 5 a 6 salários-mínimos (até R\$ 8.472) | 12 | 6,8% |
| De 6 a 7 salários-mínimos (até R\$ até R\$ 9.884) | 8 | 4,5% |
| Superior a 7 salários-mínimos (mais de R\$ 9.884) | 33 | 18,8% |

Fonte: elaborada pelos autores.

4.3.2.1.3. Característica de uso de mídias sociais para a relação entre solidão, política e mídias sociais

Quanto às variáveis do uso de mídias sociais, foi identificado, como mostra a Tabela 9, o uso de algumas plataformas de mídias sociais como *Instagram*, *Facebook*, *Youtube*, *Tik-Tok* e *Telegram* para acompanhamento de assuntos sobre política a partir de 2022. 87,5% dos respondentes utilizam estas plataformas de mídias sociais nos últimos anos. Dentre estes, a maioria (65,3%) é usuário do Instagram seguido de Youtube para acompanhar informações sobre política. O *Telegram* foi a plataforma menos utilizada por eles nestes últimos anos para acompanhamento de informações sobre política.

Tabela 9- perfil dos participantes que utilizaram mídias sociais

| Variáveis | Número | Porcentagem |
|--|---------------|--------------------|
| Utiliza plataforma de mídias sociais | | |
| Sim | 170 | 96,6% |
| Não | 6 | 3,4% |
| Plataforma que utiliza para acompanhar informações sobre política | | |
| Facebook | 53 | 30,1% |
| Instagram | 115 | 65,3% |
| Linkedin | 28 | 15,9% |
| Tik Tok | 17 | 9,7% |
| Telegram | 9 | 5,1% |
| Twitter | 29 | 16,5% |
| Youtube | 91 | 51,7% |
| Outros | 24 | 13,6% |
| Não utilizo | 22 | 12,5% |

Fonte: elaborada pelos autores.

4.3.2.1.4. Características do perfil político para a relação entre solidão, política e mídias sociais

Por fim, a Tabela 10 mostra a variável perfil político dos participantes por aptidão ao voto, identificando os participantes que estiveram aptos e os que participaram no segundo turno das eleições para presidente do Brasil em 2022. 86,4 % dos participantes estavam aptos a votar nessas eleições. Também foram identificados os que votaram no segundo turno das eleições em 2022, sendo 84,1% dos participantes disseram ter participado do segundo turno das eleições para presidente do Brasil em 2022.

Tabela 10- variável perfil político dos participantes

| Variáveis | Número | Porcentagem |
|-----------------------|---------------|--------------------|
| Apto a votar | | |
| Sim | 152 | 86,4% |
| Não | 21 | 11,9% |
| Prefiro não responder | 3 | 11,9% |
| Votou | | |
| Sim | 148 | 84,1% |
| Não | 27 | 15,3% |
| Prefiro não responder | 1 | 0,6% |

Fonte: elaborada pelos autores.

4.3.2.2. Elaboração e pré-teste das ferramentas de coleta de dados

Para a coleta de dados, foram utilizados os seguintes instrumentos padronizados e adaptados para o contexto desta pesquisa: i) Questionário sociodemográfico e socioeconômico, com o objetivo de identificar o perfil dos participantes; ii) Escala Brasileira de Solidão (UCLA), para mapear os níveis de solidão; e iii) Questionário sobre autopercepção de comportamentos políticos engajados, para avaliar o engajamento político tanto nas relações sociais presenciais quanto nas mídias sociais durante o segundo turno das eleições presidenciais no Brasil em 2022.

O questionário sociodemográfico e socioeconômico (Apêndice B) foi composto por questões que abrangem características sociodemográficas, socioeconômicas, perfil eleitoral e uso de mídias sociais. As variáveis sociodemográficas investigadas incluíram idade, gênero, etnia, estado civil e região de residência no Brasil. Já as questões socioeconômicas abordaram o nível educacional e a renda dos participantes. Sobre o perfil eleitoral, foram incluídas perguntas relacionadas à participação nas eleições, como a aptidão para votar e a participação no segundo turno das eleições presidenciais brasileiras de 2022. No que diz respeito ao uso de mídias sociais, buscou-se identificar se os participantes eram usuários dessas plataformas e quais utilizavam com maior frequência para acompanhar e interagir sobre temas políticos. Esse questionário teve como finalidade estabelecer as variáveis relacionadas ao perfil sociodemográfico, socioeconômico, político e de uso de mídias sociais.

A Escala Brasileira de Solidão (UCLA) versão 1 (Anexo B), adaptada e validada para o Brasil por Barroso, Andrade, Midgett e Carvalho (2016), com base na escala original de Russel, Peplau e Cutrona (1980), foi utilizada para medir sentimentos associados à solidão (Hertz, 2021). Composta por 20 questões, a escala avalia sentimentos e ações relacionados à solidão, com respostas baseadas em uma escala *Likert* de quatro pontos, variando de 0 (nunca) a 4 (muito frequente). A pontuação máxima possível é de 60 pontos. O objetivo dessa escala foi mapear e investigar indicadores negativos de solidão, identificando a ausência de conexões sociais por meio da autopercepção dos participantes sobre seus sentimentos de isolamento (Barroso, 2016).

Para interpretar e analisar os resultados da escala da solidão, adotou-se, em código, a classificação proposta por Ferreira e Casemiro (2020), que categoriza os níveis de solidão com base na pontuação total: solidão mínima (1 a 22 pontos), solidão leve (23 a 35 pontos), solidão moderada (36 a 47 pontos) e solidão intensa (48 a 60 pontos). Neste estudo, a escala foi utilizada para mapear os níveis de solidão após o período eleitoral do segundo turno das eleições presidenciais brasileiras de 2022. Os dados obtidos serão posteriormente correlacionados com as demais variáveis investigadas.

Por fim, o questionário de autopercepção de comportamentos político engajado e uso de mídias sociais para interações políticas (Apêndice C) foi desenvolvido com base em instrumentos internacionais previamente validados conforme apresentado a seguir. Composto por 28 questões fechadas e afirmativas organizadas em itens, podendo ser respondidas em uma escala *Likert* de quatro pontos (0 a 4) que mede a frequência com que os participantes se percebem envolvidos em determinados comportamentos políticos.

Para a elaboração deste questionário, foram adaptados itens dos instrumentos internacionais para o desenvolvimento destas questões sobre uso da *internet* e comportamento social e político do *Inquérito Social Europeu (European Social Survey - ESS)*; questões sobre polarização política e uso de mídias sociais dos estudos de Wang et al. (2023), Wakefield e Wakefield (2023) e Yan et al. (2021); tópicos sobre temas que dividem a sociedade do *Teste Político de 8 Valores (IDRlabs, n.d.)*; e itens sobre posicionamento político do *Teste de Coordenadas Políticas (IDRlabs, n.d.)*, além de contribuições da revisão metodológica e dos objetivos específicos deste estudo. As questões sobre política visaram analisar padrões comportamentais de engajamento político, durante as eleições presidenciais no Brasil em 2022, por meio da autopercepção dos participantes sobre interesse político; participação em movimentos políticos, identificação com partidos políticos; afastamento ou a aproximação de indivíduos por afinidades ideológicas, tanto nas mídias sociais (uso de mídias sociais) quanto no contexto presencial.

O questionário passou por um pré-teste, aplicado de forma impressa e individualmente a um grupo de cinco voluntários, que forneceram *feedbacks* sobre a clareza e a pertinência das questões. As questões selecionadas para a pesquisa final foram aquelas que melhor capturaram comportamentos de aproximação e evitação de indivíduos com ideologias ou preferências partidárias contrastantes. Foi esclarecido que esta ferramenta de coleta de dados políticos não teve pretensão de determinar com exatidão a afiliação política ou a influência dela na cognição dos participantes.

Durante a implantação do algoritmo *KNN*, baseado nos ajustes de pontuação feitas com a variável *solidão* anteriormente, foram analisados os níveis de severidade do comportamento político engajado presencial (23 questões) e uso de mídias sociais para interações políticas (5 questões), ajustando os valores da pontuação obtidas nas respostas através da escala 0 a 4 das questões correspondentes.

Adotou-se, no código, o cálculo dos níveis de severidade do comportamento político engajado para verificar comportamento antagônicos. Na falta de referências, até o alcance da revisão realizada pelos autores, que consigam classificar a severidade do antagonismo político, os autores deste trabalho realizam uma proposta de classificação, sem validação, que possibilitem a verificação da relação entre as variáveis consideradas. A classificação utilizada realizou uma categorização em quatro grupos, sendo equitativamente distribuídos entre os valores mínimos e máximos possíveis das pontuações do questionário sobre as perguntas de engajamento político. As questões foram separadas segundo a dimensão mostrada no Quadro

7. As categorias consideradas foram mínimo, leve, moderado e intenso, como pode ser visto nesta seção.

Quadro 7- Questões sobre comportamento político engajado

| Nº da questão | Itens |
|---------------|---|
| 1 | Me informei (o) sobre as ações do candidato a presidente da república que pertencia ao partido/coligação que me identifico. |
| 2 | Trabalhei para um partido político ou movimento cívico |
| 3 | Usei um emblema autocolante de campanha/movimento |
| 4 | Participei de uma manifestação ou convocação feita por partidos políticos |
| 5 | Tenho uma posição no espectro político, onde 0 representa nulo e 4 mais à esquerda. |
| 6 | Tenho uma posição no espectro político, onde 0 representa nulo e 4 mais à direita. |
| 7 | Me interessei mais pela política |
| 8 | Me envolvi mais com assuntos ligados a política |
| 9 | Tive preferência por algum partido ou coligação político |
| 10 | Contatei um representante do poder local (prefeitos, deputados etc.) |
| 11 | Conversei sobre notícias políticas e assuntos públicos com minha família. |
| 12 | Conversei sobre notícias políticas e de assuntos públicos com amigos próximos ou conhecidos. |
| 13 | Conversei sobre notícias políticas e assuntos públicos com colegas de trabalho. |
| 14 | Conversei sobre notícias políticas e de assuntos públicos com vizinhos. |
| 15 | Conversei sobre notícias políticas e de assuntos públicos com desconhecidos. |
| 16 | Me relacionei mais com pessoas que têm o mesmo interesse partidário que eu. |
| 17 | Frequentei espaços que declararam apoio ao mesmo partido que pelo qual me interesse |
| 18 | Evitei pessoas que criticam o partido que me identifico. |
| 19 | Evitei pessoas que têm ideologias políticas se opõem a minha. |
| 20 | Rompí relacionamentos por causa do partido político das pessoas. |
| 21 | Quando alguém criticou o partido que apoio pareceu um insulto pessoal. |
| 22 | Quando alguém elogiou um partido que apoio, pareceu um elogio pessoal |
| 23 | Me preocupei com o que as pessoas pensam sobre o partido que apoio. |

Fonte: elaborado pelos autores.

Já no uso de mídias sociais (Quadro 8), também foi feita uma categorização, de forma análoga à realizada para os aspectos políticos. As questões foram organizadas em dimensões, que consideravam o uso de mídias, sendo classificadas como: uso baixo, moderado e alto.

Quadro 8- organização das dimensões das questões sobre uso de mídias sociais

| Nº da questão | Item | Tipo de comportamento analisado |
|---------------|--|--|
| 24 | Ver conteúdo político | Exposição passiva |
| 25 | Interagir com conteúdo semelhante | Engajamento ativo (com viés confirmatório) |
| 26 | Bloquear/dar <i>unfollow</i> em discordantes | Evitamento |
| 27 | Discutir/hostilizar partido | Confronto hostil |
| 28 | Excluir amigos por discordância | Ruptura de laços |

Fonte: elaborado pelos autores.

4.3.2.3. Aplicação das ferramentas de coleta de dados

A coleta de dados foi realizada por meio de um questionário eletrônico, desenvolvido na plataforma *Google Forms*. O formulário foi estruturado em três seções: Seção 1 contendo questões sobre identificação do perfil dos participantes; Seção 2 contendo todos os itens da

Escala Brasileira de Solidão (UCLA); e a Seção 3 com o questionário sobre autopercepção de comportamentos políticos.

A pesquisa foi conduzida no primeiro semestre de 2024, entre os meses de maio e julho, período escolhido por estar fora do calendário eleitoral das eleições municipais daquele ano. Essa decisão visou evitar possíveis interferências ou vieses decorrentes do clima político eleitoral.

A aplicação do instrumento foi através do *link* do formulário compartilhado, utilizando o método de amostragem bola-de-neve, através do qual os pesquisadores responsáveis divulgaram a pesquisa em suas redes de contato. O público foi convidado a participar voluntariamente, por meio da divulgação em plataformas de mídias sociais, como *WhatsApp*, *Instagram*, *LinkedIn* e *Facebook*. Além disso, a pesquisa foi divulgada em uma rádio local e pelo e-mail institucional da Universidade Federal de Itajubá (UNIFEI), na qual os pesquisadores estão vinculados. Em todas as divulgações, foram informados os objetivos do estudo, a data de início e término da pesquisa, bem como o local e a forma de aplicação dos instrumentos. Juntamente com o questionário, os participantes receberam o Termo de Consentimento Livre e Esclarecido (TCLE) (Apêndice A), no qual foram detalhadas as informações sobre a pesquisa, incluindo seus objetivos, benefícios, riscos mínimos, procedimentos de armazenamento e tratamento dos dados.

O estudo seguiu rigorosamente as diretrizes da Lei Geral de Proteção de Dados (LGPD), garantindo o sigilo e a proteção das informações coletadas. Foram adotadas medidas para prevenir riscos e garantir a segurança dos participantes, como a não divulgação de dados pessoais; a garantia de que os pesquisadores não teriam acesso aos endereços IP dos dispositivos utilizados e o armazenamento dos dados em dispositivos *offline* até o término da pesquisa.

Os pesquisadores mantiveram-se disponíveis para esclarecer dúvidas e oferecer suporte aos participantes, por meio de contato via *e-mail* e presencialmente na universidade, em horários previamente divulgados. Caso necessário, uma psicóloga integrante da pesquisa ficou à disposição para auxiliar os participantes que eventualmente apresentassem sentimentos de ansiedade, estresse, incômodo ou desconforto durante ou após a resposta ao questionário. Foi reforçado aos participantes que sua participação era voluntária e que poderiam desistir da pesquisa a qualquer momento, sem nenhum prejuízo. Além disso, foi esclarecido que o estudo não tinha como objetivo favorecer ou prejudicar qualquer partido político, mantendo-se neutro em relação às preferências políticas dos participantes.

4.3.3. Preparação inicial dos dados brutos

Para o treinamento dos algoritmos, os dados empíricos foram visualizados e preparados conforme as etapas gerais de Géron (2021). Estes dados são de casos reais obtidos através de instrumentos de autorrelatos, conforme visto na seção 4.3.2.3.

Para o desenvolvimento do modelo de aprendizado de máquina, foi estabelecido um ambiente isolado utilizando o *Visual Studio Code (VS Code)* como ferramenta principal de edição e execução de código. A implementação foi realizada em *Python* (versão 3.12.4), linguagem reconhecida por sua eficiência em aplicações científicas e de IA (McKinney, 2013), destacando-se por sua natureza dinâmica, interpretada e com capacidade de integração de sistemas para análise de dados, computação interativa e visualização. Desta forma, foram empregadas as seguintes bibliotecas *Python* especializadas: *NumPy* para operações numéricas avançadas; *Pandas* para manipulação e análise estruturada de dados; *Scikit-learn* para implementação dos algoritmos de aprendizado de máquina e *Matplotlib* e *Seaborn* para geração de visualizações e análises exploratórias.

No *VS Code* foram importadas as bibliotecas *Python* em etapas para análise da estrutura e análise inicial dos dados. O conjunto de dados, originalmente no formato *CSV* e coletado via *Google Forms*, foi carregado no ambiente de desenvolvimento. Uma análise preliminar através do método *head* revelou uma matriz composta por: dados para implementação do *K-means* com 150 linhas e 66 colunas e dados para implementação do *KNN* e do *DBSCAN* com 177 registros (176 participantes e 1 linha de cabeçalho); 66 atributos ou variáveis de tipos diversos (numéricos e textuais) e dados categorizados como binários (respostas dicotômicas) ou multivariados (escolhas múltiplas).

4.3.3.1. Pré-tratamento dos dados brutos

Nesta etapa os dados foram padronizados através da identificação e tratamento de valores nulos ou ausentes e as correções de *outliers* e remoção de dados irrelevantes. Também se fez a transformação de atributos textuais (gênero, etnia) em representações numéricas.

Foi feita a engenharia de *features* que consistiu no cálculo do índice de severidade da solidão para agregação das respostas da Escala de Solidão (20 itens) em pontuações únicas (escala 1-60) como descrito na Seção 4.3.2.1.. Em seguida fez-se simplificação de variáveis complexas (padrões de uso de redes sociais) e exclusão de colunas não contributivas para o modelo

Já na etapa de normalização, normalizou-se os valores numéricos e a divisão aleatória do *dataset* de 150 instâncias para treinamento para a implementação do *K-means* e 177 instâncias para implementação do *DBSCAN* e do *KNN*. Posteriormente, fez-se a preparação final dos dados por meio da categorização de variáveis em binárias (respostas sim ou não) e multivariadas (faixa etária, escolaridade e renda).

Este processo de preparação garantiu a qualidade dos dados para as etapas subsequentes de treinamento e validação dos modelos, cujos resultados serão detalhados no capítulo seguinte. A abordagem seguiu as práticas recomendadas por Géron (2021) para projetos de aprendizado de máquina, assegurando a robustez e confiabilidade dos resultados.

4.3.4. Seleção e treinamento dos modelos de ML

Nesta etapa, seguiu-se uma estratégia combinada de tipos de algoritmos com treinamentos diferentes, em duas fases: treinamento não-supervisionado e treinamento supervisionado. O treinamento não supervisionado permitiu uma compreensão mais abrangente da correlação entre as variáveis investigadas (incluindo solidão, políticas, uso de mídias, socioeconômicas, sociodemográficas e perfil político). Já o treinamento supervisionado buscou evidenciar associação entre solidão, política e mídias sociais através da predição de categorias dos agrupamentos identificados ou rótulos.

Nas seções deste capítulo serão descritas apenas o processamento dos dados de entrada, sendo a Seção 4.4.4.1. treinamento do *k-means*; seção 4.4.4.2. a treinamento do *DBSCAN* e Seção 4.4.4.3. treinamento do *KNN*. Todos os treinamentos seguem as etapas gerais que envolve pré-processamento dados; seleção de hiperparâmetros; treinamento e validação dos modelos (Géron, 2021). No entanto, cada algoritmo contém etapas específicas e pré-tratamento que foram apresentadas na etapa de implementação. Os dados de saída dos algoritmos serão apresentados e discutidos na análise e discussão dos resultados, Capítulo 5.

4.3.4.1. Implementação do *K-means* para análise exploratória inicial dos dados

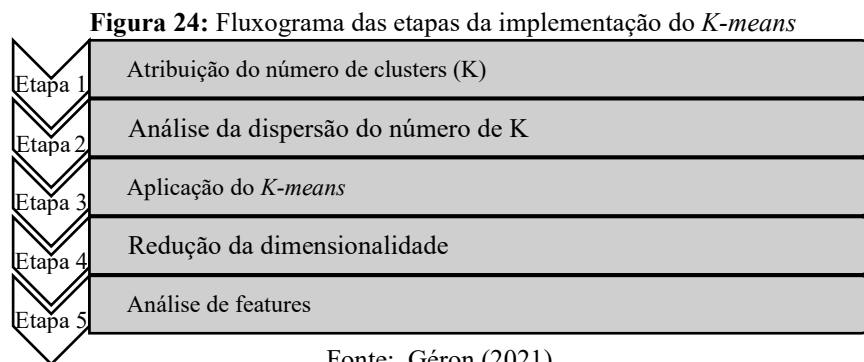
Inicialmente, realizou-se um pré-teste com aprendizado não supervisionado, aplicando o algoritmo *K-means* para uma avaliação exploratória inicial dos dados em quantidade menor, escolhendo de forma aleatória 150 instâncias (implementação do modelo apresentada no apêndice E). Além das instancias, o banco de dados possui no total 63 colunas, representando a variável sociodemográfica (6 colunas), socioeconômicas (2 colunas); perfil dos participantes

que utilizam médias (2 colunas); perfil político (2 colunas); solidão (20 colunas); comportamento político engajado (23 colunas) e uso de médias (5 colunas).

Essa abordagem teve como finalidade orientar o pré-processamento e preparar os dados para etapas posteriores de treinamento do modelo mais adequado para lidar com a variabilidade e correlação destes dados, auxiliando na seleção das variáveis mais importantes.

Por meio desse treinamento, foi possível identificar o comportamento dos dados, incluindo padrões e agrupamentos naturais, através da visualização preliminar da estrutura do conjunto de dados agrupados (*clusterizados*).

Após um pré-tratamento básico dos dados, que incluiu a eliminação de algumas variáveis consideradas de pouca importância com base em critérios estratégicos, aplicou-se o *K-means* ao conjunto pré-teste para análise do comportamento dos dados, conforme a Figura 24 que mostra o fluxograma das etapas de treinamento. Foram seguidas as seguintes etapas específicas na implementação: i) definição do hiperparâmetro ou escolha do número de *clusters*; ii) análise da dispersão do número de clusters; iii) treinamento do *K-means*; iv) análise da variância e redução da dimensionalidade e v) análise das variáveis mais importantes (*features*).



Fonte: Géron (2021).

Primeiramente, atribuiu-se e especificou-se ao algoritmo a quantidade de *clusters* (*K*) que deveriam ser identificados e associados a cada instância. Fez-se um agrupamento, selecionando estes *k* pontos como centros ou centróides. Atribuiu-se a cada ponto no conjunto de dados foi atribuída sua proximidade ao centróide. Assim minimiza-se a soma dos quadrados dentro do agrupamento (WCSS), sendo definido como (Hozumi *et al.*, 2021):

$$J = \sum_{j=1}^k \sum_{x_i \in C_j} \|x_i - \mu_j\|_2^2 \quad (1)$$

Onde k é o número de clusters, C_i é o conjunto de pontos pertencentes aos *cluster* i ; x é o ponto de dados μ_i centroide do *cluster* i (média dos pontos em C_i) e $\|x_i - \mu_j\|_2^2$ é distância euclidiana quadrada.

Utilizou-se o Método do Cotovelo, uma técnica gráfica para determinar e ajustar o número ideal de *clusters* avaliação, indicando se pode ser considerada ótima para implementação do modelo. O valor de K encontrado foi considerado ótimo, ou seja, a quantidade mais adequada para a realização da clusterização. A análise foi baseada nos intervalos de K , observando-se a formação de uma curvatura característica na linha do gráfico, em formato de cotovelo. O ponto de maior variação (onde se forma o cotovelo) foi considerado como indicador do valor ideal de K . Ao escolher $K=4$, significa que a escolha do número ideal de *clusters* foi realizada considerando que a inclinação da curva é quase constante entre 11 e 4 *clusters*, após isso, a dispersão aumenta muito mais rápido, mudando drasticamente a inclinação, como pode ser visto na seção 5.1 no capítulo 5.

Após a avaliação da escolha do número de clusters, iniciou-se o treinamento do modelo *K-means*, sendo acionado o *labels* (rótulos atribuídos ao conjunto de dados) do *dataframe* (conjunto de dados em movimento no programa) original e verificados os resultados onde ele encontrou os 4 *centróides* (0,1,2 e 3).

Foram analisadas todas as características dos levantados (sociodemográficos, socioeconômicos, perfil político, perfil de uso de mídias sociais, nível de solidão, comportamento político engajado e uso de mídias sociais para interação política) dos *clusters* que representa os 150 participantes quanto a proximidade de suas instâncias, utilizando o método *PCA* (*Principal Component Analysis*) para reduzir a quantidade de variáveis mantendo o máximo possível da informação presente nos dados, ao identificar e projetar os dados em um espaço 2D nas direções de maior variação (Hozumi et al., 2021).

Posteriormente, para visualização das variáveis consideradas de maior importância pelo modelo, utilizou-se o método *random forest* que classificou estas variáveis (*features*) pelo grau de importância atribuído para o modelo de forma hierárquica de modo a definir os grupos por *clusters*, servindo de apoio ao que foi demonstrado através do PCA (Marteletto, 2022). Por meio do *random forest*,

Por fim, foi analisada a variável de severidade da solidão e sua relação com outras variáveis na clusterização através da visualização gráfica de dados agrupados por *clusters*, sendo plotado o primeiro gráfico sobre a distribuição da severidade da solidão e o segundo

distribuição da faixa etária por *clusters* devido a maior influência apresentada na associação dentre as variáveis de perfil disponibilizadas.

4.3.4.2. Implementação do *DBSCAN* para agrupamento de dados densos e identificação dos grupos

Nesta etapa realizou-se o treinamento do *DBSCAN*, a partir do *insight* da implementação *K-means*, para uma nova clusterização entre os dados e compreensão global do comportamento destes dados que não se relacionaram a partir do formato e de ruídos não explícitos anteriormente (apêndice F). Para este modelo, foram selecionadas e associados as variáveis principais (solidão, comportamento político e uso de mídias sociais), de modo que a partir dos grupos formados serão atribuídos rótulos para aplicação de um modelo supervisionado posteriormente.

Na etapa de seleção das *features*, o *dataset* utilizado para o treinamento (176 participantes e 48 questões ou colunas), sendo selecionadas apenas as colunas relacionadas às variáveis solidão (20 colunas), comportamento político engajado (23 colunas) e uso de mídias sociais para interações políticas (5 colunas). Como já explicado, as colunas referentes a variáveis sociodemográficas, socioeconômicas, perfil eleitoral e perfil de uso de mídias sociais foram excluídos conforme o *insight* obtido por meio do *K-means*.

Fizeram-se ajustes iniciais para normalizar os dados do *dataset* antes do treinamento do *DBSCAN*, como o preenchimento de valores ausentes e a transformação de colunas binárias ('votou no 2º turno da eleição' e 'apto a votar no 2º turno da eleição') e multiclasse ('mídias sociais utilizadas'). Essas colunas foram codificadas, ou seja, ajustadas de acordo com as demais variáveis do *dataset*.

Também foram definidas as colunas correspondentes às variáveis comportamento político (26 colunas) e uso de mídias sociais (5 colunas) e solidão (20 colunas) para ajustar a escala de severidade, obtendo-se seus respectivos níveis. Na etapa de classificação dos dados, reduziram-se e classificaram-se os dados de autopercepção de comportamentos políticos e uso de mídias sociais, seguindo como referência o índice de ajuste aplicado à escala de solidão, com ajustes dos valores obtidos por pontuação, de acordo com a escala e a pontuação total do conjunto de itens das questões.

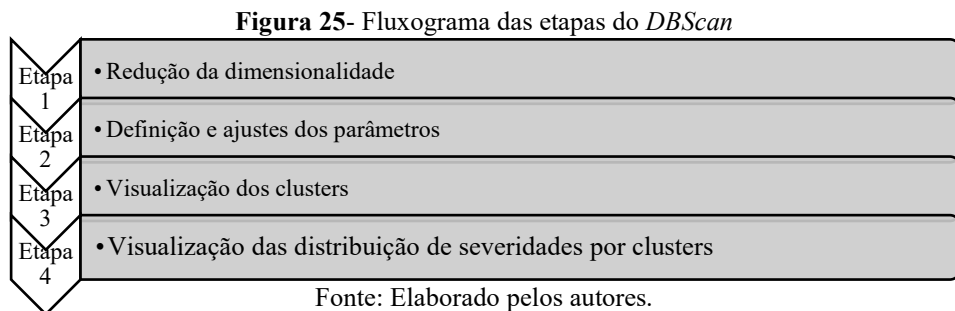
Para esta classificação dos comportamentos políticos e do uso de mídias sociais, conforme descrito na Seção 4.3.2.2., realizou-se uma distribuição uniforme, dividindo os dados

em quatro grupos de severidade, entre os valores mínimos possíveis (zero) e os máximos (92 para comportamento político e 20 para uso de mídias sociais) de cada conjunto de perguntas.

Efetuiu-se o agrupamento dos itens em dimensões e a padronização, em que a soma da pontuação de cada respondente resultou nas seguintes categorias: mínimo (0 a 23), leve (24 a 46), moderado (47 a 69) e intenso (70 a 92). Para determinar o nível de severidade do uso de mídias sociais de cada participante, organizaram-se os itens em uma escala de 0 a 4, resultando nas seguintes classificações: muito baixo (0 a 5), baixo (6 a 10), moderado (10 a 15) e alto (15 a 20). Posteriormente, normalizaram-se as colunas numéricas restantes, exceto as categóricas, para utilizar as severidades e convertê-las em categorias numéricas para *ML*.

Removeram-se as colunas originais das escalas para evitar redundância, e aplicou-se a normalização às colunas numéricas remanescentes. As severidades categóricas não foram normalizadas, a fim de preservar sua utilização e permitir a conversão em categorias numéricas para análise.

Após o pré-processamento, foi treinado o *DBSCAN*, seguindo as principais etapas do fluxograma da Figura 25.



Para implementação do *DBSCAN*, seguiram-se mais etapas de preparação dos dados como codificação ordinal para análise e visualização das *features* que correspondem às escalas de severidade da solidão, comportamento político e uso das mídias. Em seguida, fez-se a correlação entre as escalas plotadas através de uma matriz de correlação, que, análise como as três as variáveis, mostra-se a força de correlação entre elas. Nesta matriz de correlação, a análise quantitativa dessa correlação feita através do coeficiente de correlação entre as variáveis, verificando se a relação entre as variáveis é linear através da comparação se os valores aproximam ou afastam da média (de +1 a -1).

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2} \cdot \sqrt{\sum(y_i - \bar{y})^2}} \quad (2)$$

Onde r_{xy} coeficiente de correlação entre as variáveis x e y ; x_i é o valor da variável x para o indivíduo i ; \bar{x} é a média de todos os valores da variável; y_i é o valor da variável y para o indivíduo i e \bar{y} é a média de todos os valores da variável y .

Foram classificadas e distribuídas as variáveis por grau de semelhança, verificando se as intervenções no modelo, baseadas na quantidade de pontos classificados calculou-se a acurácia desta classificação para analisar a confiabilidade do modelo para esta análise.

O *DBSCAN* foi implementado com parâmetros ajustados para definição de como os *clusters* foram formados, por meio da definição da distância do raio da vizinhança ($eps=50$) que serão parte de um mesmo *clusters*. Foi definido também o número de mínimo de pontos ou de vizinhos formados ($min_samples=5$), visualizado *clusters DBSCAN* (sem elipses) e a distribuição das severidades dentro de cada *cluster*.

Em seguida, para verificar a distribuição das severidades de cada variável em cada cluster, foram realizadas gráficas com cada uma das severidades de cada variável. Utilizou-se o método *t-SNE* ajustado para visualização e avaliação por meio dos ajustes dos dados de alta dimensão. Este método foi utilizado para converter as similaridades em probabilidades conjuntas, minimizando a divergência *KL (Kullback-Leibler)* entre as distribuições nos espaços original e reduzido (*Scikit-learn*, 2007; Hozumi et al., 2021).

O *t-SNE* foi feito relacionando as variáveis, através de visualizações por severidade dos agrupamentos dos 176 dados, considerando unicamente duas dimensões no espaço reduzido *t-SNE1* e *t-SNE2*. Os resultados da saída deste treinamento foram apresentados e analisados na Seção 5.2. deste capítulo.

4.3.4.3. Implementação do *KNN* para classificação dos grupos

Nesta etapa, realizou-se o treinamento do *KNN* para explorar a associação entre as variáveis principais, com base em rótulos conhecidos (agrupamentos definidos) e na validação dos *insights* obtidos na fase não supervisionada com a implementação do *DBSCAN* (Apêndice G). Desta forma, foi possível priorizar as variáveis mais importantes e suas características a fim de classificar a associação entre elas por similaridade, explorar a relação entre severidades das variáveis, utilizando técnicas de classificação.

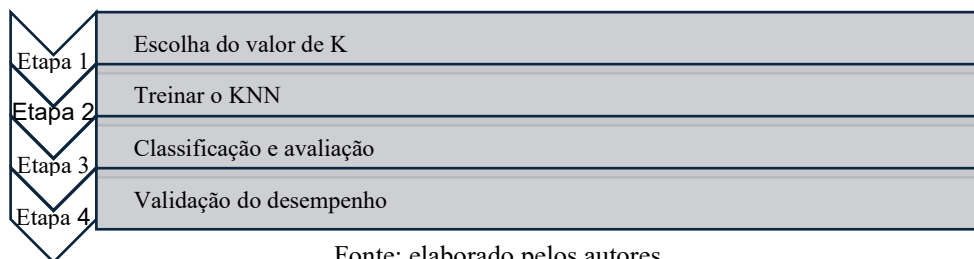
Como já apresentado na etapa de preparação dos dados Seção 4.4.3. deste capítulo, este treinamento seguiu com o mesmo *dataset* e as mesmas etapas iniciais gerais (pré-processamento) seguidas no *DBSCAN*, tais como seleção e codificação das principais *features* (variáveis) correspondentes as variáveis solidão (20 colunas), comportamento político (26

colunas) e uso de mídias sociais (5 colunas); classificação e ajustes dos valores destas variáveis; normalização das colunas numéricas restantes (exceto as categóricas) e remoção das colunas originais das escalas para evitar redundância. Posteriormente, para usar as severidades, foram convertidas para categorias numéricas para *ML*.

Além disso, criaram-se variáveis *dummies* (X e Y) para as severidades, em que X corresponde à severidade da solidão (mínima, leve, moderada e intensa), do comportamento político e do uso de mídias sociais (mínimo, leve, moderado e intenso) e uso de mídias (muito baixo, baixo, moderado e alto). A variável Y corresponde apenas à severidade da solidão. Os atributos relacionados a essas variáveis foram, posteriormente, subdivididos em conjuntos de treino e teste para aplicação do *KNN*, onde foram utilizados 20% dos dados para teste e 80% para treino (*teste_size=0.2*) para reprodutibilidade (*random_state=42*).

A Figura 26 mostra fluxograma das etapas específicas da implementação do *KNN* para análise da correlação entre as variáveis solidão, política e uso de mídias sociais: i) escolha do hiperparâmetro (valor de K vizinhos); ii) treinar o *KNN*; iii) predição ou classificação e avaliação dos dados e iv) validação do desempenho do modelo.

Figura 26- Fluxograma das etapas de implementação do *KNN*



Fonte: elaborado pelos autores.

Após a instanciação do modelo, definiu-se o hiperparâmetro do modelo de treinamento, atribuindo o número de vizinhos igual a 4. Em seguida, verificou-se a acurácia do modelo. Antes de implementação do *KNN*, realizou-se algumas análises iniciais de visualização e estatística de correlação para compreensão das saídas do algoritmo.

Analisou-se a importância e conseqüentemente a priorização das variáveis (*features*) para análise associativa, utilizando o método *random forest por meio do qual* plotou um gráfico de barras para hierarquização das variáveis de maior importância (Martelletto, 2022). Foram selecionadas as variáveis para redução da alta dimensionalidade dos dados e normalização das variáveis correspondentes ao comportamento político e uso das mídias sociais, com o objetivo de verificar quais delas impactam mais significativamente a acurácia do modelo e o cálculo de similaridade entre elas.

Para avaliar a associação significativa entre as variáveis mais relevantes do estudo, antecedendo o *KNN*, foi conduzida uma análise de contingência ou de interdependência entre as variáveis (comportamento resposta é influenciado por um estímulo que o antecede) por meio da matriz de relação bivariada, avaliando três associações principais: i) severidade da solidão versus comportamento político, ii) severidade da solidão versus uso de mídias sociais, e ii) comportamento político versus uso de mídias sociais.

Os resultados da análise de interdependência foram organizados em matrizes quadradas de contingência 4×4 , onde as linhas representam as categorias (níveis de severidade) de uma variável considerada independente e as colunas as categorias (níveis de severidade) da outra variável considerada dependente, com as células exibindo a contagem absoluta de indivíduos em cada combinação possível.

Para a avaliação das associações entre variáveis categóricas, foram aplicados o teste *qui-quadrado* (χ^2) e seu *p-valor*, com o objetivo de testar a hipótese de independência entre as variáveis. Utilizou-se este coeficiente *V* de *Cramer* (medida de força de associação entre duas variáveis) para mensurar a magnitude da associação, cujos valores variam de 0 (nenhuma associação) a 1 (associação perfeita), para maiores detalhes ver o Anexo C.

Posteriormente, como feito no *K-means*, no *KNN* também utilizou-se o *PCA* (Análise de Componentes Principais) para examinar onde há maior variação dos dados cuja finalidade neste modelo é de analisar a proximidade entre as *features*, bem como a consistência e separabilidade dos *clusters* formados anteriormente em um espaço *euclidiano*. O *PCA* (Hozumi et al., 2021).

Como saída, fez-se uma projeção dos dados em um espaço tridimensional, sendo plotado em um gráfico *3D* para visualização tridimensional da distribuição das variáveis em categorias, permitindo verificar se os pontos ou vizinhos de *K* estipulados para o treinamento do modelo pertenciam à mesma classe e identificar possíveis sobreposições. Por meio deste método, também foi possível verificar alguns *outliers*, analisar a relação entre os componentes principais e as variáveis originais, comparar classes reais versus preditas e avaliar a densidade dos dados.

Em seguida, semelhante ao que foi seguido no treinamento do *DBSCAN* (Equação 2 na Seção 4.3.4.2.), empregou-se, para avaliação dos resultados visualizados anteriormente através do *PCA*, uma matriz de correlação para uma análise quantitativa dessa correlação feita através do coeficiente de correlação entre as variáveis, verificando se a relação entre as variáveis é linear através da comparação se os valores aproximam ou afastam da média (de +1 a -1).

Foram classificadas e distribuídas as variáveis por grau de semelhança, verificando se as intervenções no modelo, baseadas na quantidade de pontos classificados calculou-se a acurácia desta classificação para analisar a confiabilidade do modelo para esta análise.

CAPÍTULO 5

5. ANÁLISE EXPLORATÓRIA E DISCUSSÃO DOS RESULTADOS

Neste capítulo, são apresentados os resultados das análises descritivas realizadas com os algoritmos *K-means* e *DBSCAN*, implementados para processar o *dataset* que contém informações sobre as variáveis solidão, comportamento político e uso de mídias sociais. Também, fez-se uma conexão com os achados da literatura através de uma discussão analítica aprofundada.

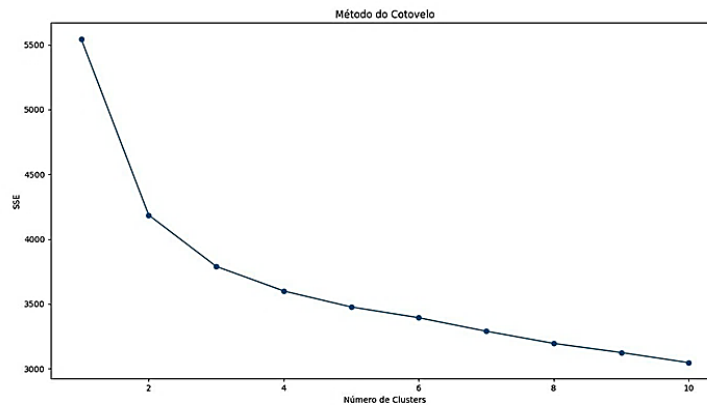
Neste capítulo, discutem-se ainda as contribuições desses resultados para as hipóteses específicas deste estudo, propondo possíveis aprimoramentos nos dados empíricos que permitam a aplicação de modelos de IA mais precisos, os quais possam oferecer melhores subsídios para a compreensão do problema em questão (Q3), respondendo o quarto objetivo específico.

Este capítulo foi dividido em 5 Seções: a Seção 5.1. análise e discussão dos resultados do *K-means* para testagem da hipótese de severidade da solidão em adultos; a Seção 5.2. análise da e discussão dos resultados do *DBSCAN* para testagem das hipóteses de associação entre as variáveis de forma integrada e pareadas; a Seção 5.3. análise da e discussão dos resultados do *KNN*. a Seção 5.4. discussão analítica dos resultados e a Seção 5.5. limitações da pesquisa.

5.1. Análise dos resultados de saída do *K-means* para análise exploratória inicial

Ao implementar o algoritmo não supervisionado, *K-means*, a Figura 27 mostra o gráfico do método do cotovelo onde apresentou dispersão *SSE* no eixo vertical e o número de *clusters* no eixo horizontal. A figura mostra que a modelo, com os dados considerados, não consegue encontrar um número de *clusters* adequado, pois a curva não atinge um patamar horizontal, apresentando uma mudança brusca no formato do cotovelo. Contudo, considerando que as análises realizadas com o *K-means* foram exploratórias, foi atribuído $K=4$ como o número de *clusters*, pois a partir deste ponto, a dispersão foi mantida baixa de modo que a curva apresentou uma inclinação constante. Isso indica ainda a necessidade de mais instâncias ou a redução do número de variáveis ou atributos.

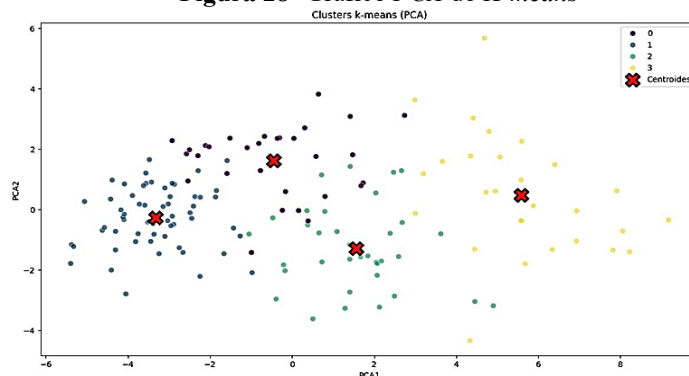
Figura 27- Método do Cotovelo



Fonte: elaborado pelos autores

Embora não estivesse entre os objetivos do trabalho, foi utilizado o método *K-means* para verificar se existia ou não uma influência das variáveis socioeconômicas e suas relações com os três grupos de variáveis principais deste projeto (solidão, política e mídias sociais). Na Figura 28, mostram-se os resultados do método *PCA*, com duas componentes principais, considerando todo o conjunto de variáveis presentes no *dataset* inicial. Observou-se que os *clusters* não foram detectados ou agrupados devidamente, pois apresentam alta dimensionalidade dos dados que corresponde a grande número de variáveis ou características desses dados, sendo necessário identificar as consideradas mais importantes para agrupamento.

Figura 28- Gráfico PCA do K-means



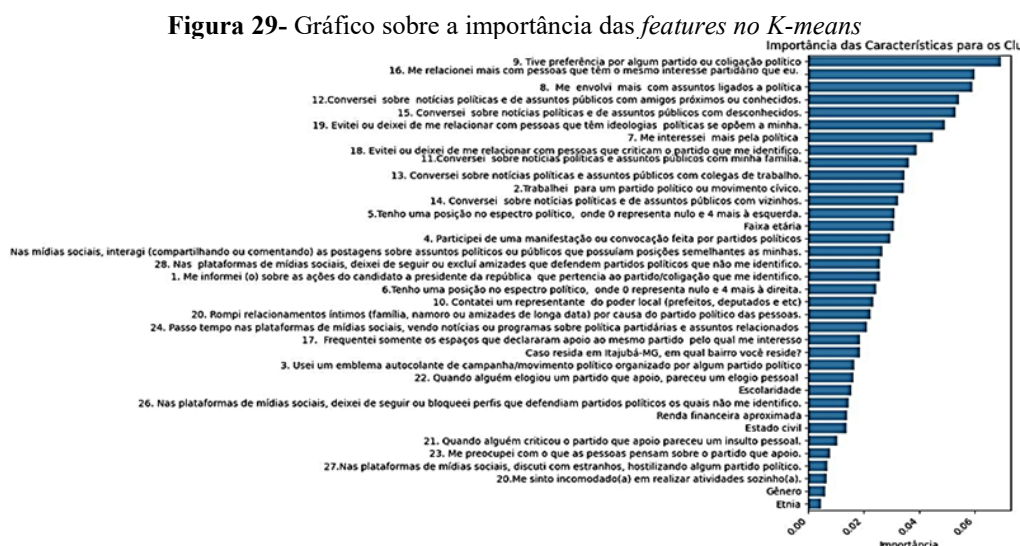
Fonte: elaborado pelos autores.

Na Figura 29, mostram-se os resultados obtidos através do *random forest*. É possível observar nesta figura que as variáveis relacionadas ao comportamento político, como a preferência por partido político, tiveram mais influência na formação dos *clusters*, quando comparadas às variáveis sociodemográficas, como gênero e etnia. Isso foi observado em outras

variáveis de perfil dos participantes como escolaridade, renda e cidade onde os participantes residem.

Em relação às variáveis sociodemográficas consideradas pelo modelo como de menor influência, pode-se justificar pelo viés da amostra, pois conforme os dados de identificação do perfil, na Seção 4.4.2.1., no Capítulo 4, houve prevalência do gênero feminino e etnia branca na amostra analisada. Já o perfil socioeconômico prevaleceu nas pessoas que possuem pós-graduação e renda de até 2 salários-mínimos. A literatura também não evidenciou empiricamente a influência destas variáveis na associação entre solidão, política e uso de mídias sociais. Para análise de grandes sistemas de produções culturais que possuem muitos dados empíricos, Josiowicz (2021) aponta que o campo humanidades digitais têm considerado mais os métodos computacionais em detrimento de variáveis sociodemográficas (como gênero e etnia) e questões geopolíticas.

Ainda, na Figura 29, também, nota-se que a falta de influência das variáveis do perfil como sociodemográfico (gênero, etnia e estado civil) e socioeconômicas (escolaridade e renda) está relacionada alta variabilidade destas características que operam em escalas distintas, dificultando se correlacionarem empiricamente, provocando o mal funcionamento do algoritmo. Aponta-se também um viés da seleção amostra, com falta de critérios bem definidos que tornou a amostra muito ampla para análise. Estes resultados indicam que estas características não precisam ser consideradas para implementação de outros algoritmos.



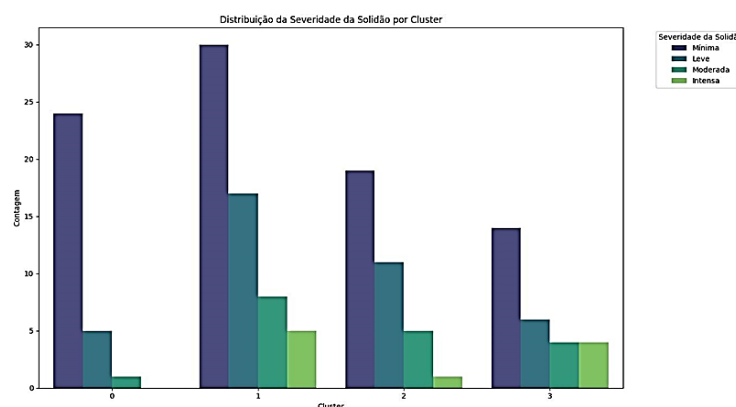
Fonte: elaborado pelos autores.

Foi possível observar a distribuição de variáveis por *clusters*, como visto nas Figuras 30 a 31. Considerando a análise exploratória inicial, desta etapa, foram escolhidas algumas das

variáveis (faixa etária e intensidade da solidão – escolhidas de forma aleatória) e verificar sua distribuição dentro de cada um dos *clusters*. Assim, na Figura 30, analisa-se a distribuição da intensidade da solidão por *cluster*, onde a vertical apresenta a pontuação das respostas e a horizontal os *clusters* que representam o agrupamento em *clusters* dos participantes da pesquisa. Neste gráfico, a solidão mínima está representada em azul escuro, a solidão leve em azul claro, a solidão moderada em verde escuro e a solidão intensa em verde claro.

Ao comparar os graus de severidade da solidão presente em cada *cluster*, onde o *cluster* 0 é o segundo agrupamento onde existem mais casos de solidão mínima e o único que não há casos de solidão intensa. O *cluster* 1 é o primeiro onde há mais casos de solidão mínima e demais casos de solidão intensa. O *cluster* 2 é o terceiro agrupamento com mais casos de solidão mínima, possuindo menos casos de solidão intensa. Por fim, o *cluster* 3 é o agrupamento onde há menos casos de solidão mínima e o segundo com mais casos de solidão intensa. Os resultados mostrados na figura, não mostram uma tendência clara da influência da severidade da solidão na distribuição dos *clusters*.

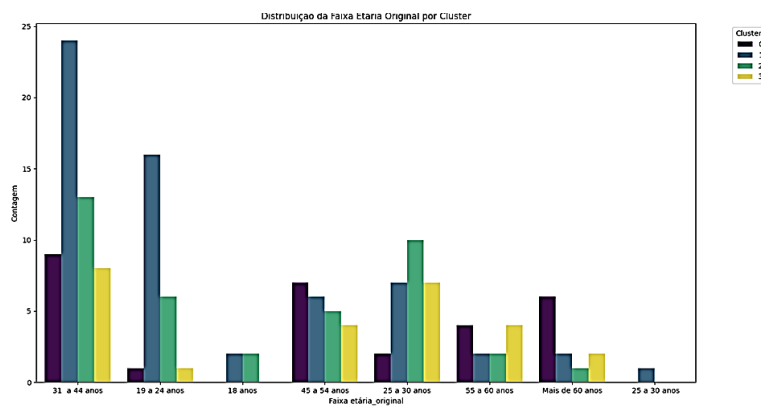
Figura 30- Gráfico de distribuição da severidade da solidão por cluster



Fonte: Elaborada pelos autores.

Da mesma forma, na Figura 31, é visualizada a distribuição da severidade da solidão através da clusterização por faixa etária. Em detalhe, observa-se a distribuição dos cinco *clusters*, sendo feito agrupamento conforme a intensidade de solidão apresentada por faixa etária dos participantes. Na faixa etária de 31 a 44 anos existem mais casos de solidão leve e solidão intensa. Esta faixa representa o maior grupo de participantes na amostra (33%). Ao revelar esta classificação de severidade alta apresentou mais casos neste grupo, estes resultados contribuem para compreensão da solidão como problema significativo que pode afetar indivíduos nesta faixa etária.

Figura 31- Gráfico agrupamento da intensidade da solidão por faixa etária



Fonte: Elaborada pelos autores.

Como verificado nos resultados anteriores, o algoritmo de clusterização não supervisionado foi considerado limitado para lidar com todas as variáveis nos dados disponibilizados, apresentando falhas para capturar a relação complexa entre elas. Observou-se que o *K-Means* não lidou adequadamente com a heterogeneidade e a complexidade dos dados, gerando clusters pouco definidos e com alta dispersão.

Estas variáveis possuem muitas características que demonstraram não se ajustar naturalmente a dimensionalidade do *K-means* para que este modelo pudesse agrupá-las. O que tornou preciso limitar a atribuição de importância, descartando colunas (*features*) do *dataset* onde há muitas variações de características que possuem pouca influência na associação entre as variáveis principais (solidão, política e uso de mídias sociais), segundo o modelo.

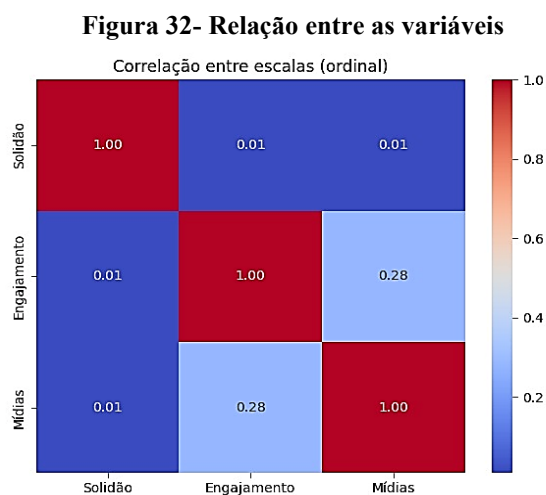
5.2. Análise dos resultados do *DBSCAN* para identificação dos grupos

A implementação do *K-means*, explorada na seção anterior, revelou a existência de padrões complexos que demandavam tanto uma delimitação mais focada das variáveis quanto métodos analíticos mais assertivos. Diante disso, procedeu-se à redução de variáveis, selecionando-se aquelas de maior relevância para os objetivos da pesquisa: solidão, comportamento político e uso de mídias sociais. Estas variáveis foram mantidas como foco da análise devido ao seu maior interesse teórico e aos indicativos preliminares de interação comportamental. Essa análise dos clusters obtidos com o *DBSCAN* fornece uma base empírica robusta para avaliar as hipóteses de pesquisa propostas, permitindo uma investigação que vai além das relações lineares e captura a complexidade das interações entre as variáveis.

Os resultados do *K-means* evidenciaram limitações na identificação de padrões claros entre as variáveis, atribuíveis à alta variabilidade comportamental dos dados. Assim, conforme previsto na metodologia, o *DBSCAN* foi selecionado para analisar o *dataset* final, utilizando-se

as três variáveis selecionadas em seus respectivos níveis de severidade, excluindo-se componentes socioeconômicos, sociodemográficos e de perfil específico de uso de mídias.

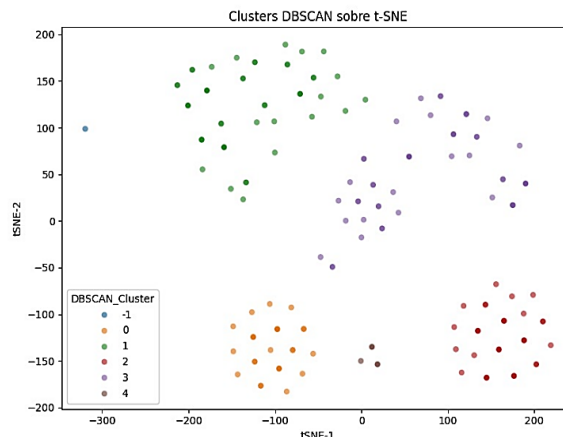
Na Figura 38, apresentam-se os resultados, da avaliação da relação entre as variáveis através da matriz de correlação entre as escalas das variáveis severidade da solidão, comportamento político de engajamento e uso de mídias. A matriz obtida evidencia uma relação muito baixa entre solidão e uso de mídias sociais; assim como também para a relação entre o engajamento político e a solidão, ambas com valores ao redor de 0,01, onde 0,00 representa o valor mínimo e 1 o valor máximo da relação. Contudo, a relação entre engajamento político e mídias apresenta um valor um pouco maior, 0,28.



Fonte: elaborado pelos autores.

Como explicado anteriormente, na Figura 33 mostra a saída do *DBSCAN* onde usam-se duas dimensões de *t-SNE* para poder visualizar de forma quantificável as severidades das três variáveis (solidão, comportamento político de engajamento e uso de mídias sociais). Após o uso do *DBSCAN*, os 176 participantes foram agrupados em 5 grupos (*clusters*). Os clusters formados foram: -1 em azul claro (este grupo representa os indivíduos que não conseguiram ser agrupados com o método); cluster 0 em laranja; *cluster 1* em verde; *cluster 2* em rosa; *cluster 3* em roxo e *cluster 4* em marrom.

Figura 33- Agrupamento dos participantes em *clusters*



Fonte: elaborada pelos autores.

Como pode ser visto através desses agrupamentos através do *DBSCAN*, a metodologia permitiu uma visualização por *clusters* que indica algumas relações entre as variáveis, mostrando-se adequada para testar associações bivariadas. Os resultados, contudo, não sustentam as relações diretas e intensas previstas. Os achados apontam para um cenário mais complexo de padrões de associação entre as variáveis, onde a tendência da amostra é possuir um perfil comportamental com solidão mínima e o baixo uso de mídias, sendo o engajamento político intenso padrão atípico e raro.

Foi evidenciado que os maiores agrupamentos (*cluster 1* e *cluster 3*) são mais heterogêneos e possuem padrões de comportamentos diversos entre os participantes, sendo para casos de severidade de solidão (mínima, leve, moderada e intensa), comportamento político (leve, mínimo e moderado) e uso de mídias sociais (baixo, moderado e muito baixo). E não foram encontrados casos de engajamento intenso ou uso de mídias altos nestes agrupamentos.

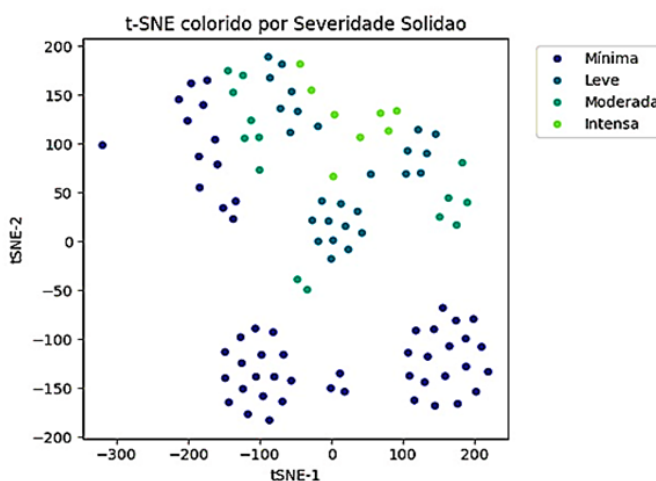
Também, nota-se que o comportamento do *cluster -1*, ou *outlier*, que muitas vezes denominados de ruído ou dados que não conseguiram ser classificados, possui padrão atípico na amostra ao apresentar severidade de solidão mínima, severidade intensa para comportamento político engajado e severidade moderada para uso de mídias.

Já os grupos homogêneos formados para severidade (*cluster 0*, *cluster 2* e *cluster 4*) possuem prevalência de severidade de solidão mínima e uso de mídias baixo, porém varia entre casos de severidade para comportamento político, sendo severidade predominantemente leve no *cluster 0*, predominantemente moderada no *cluster 4* e predominantemente mínima no *cluster 2*.

Isso evidencia que existem grupos de pessoas com solidão mínima com baixo uso de mídias sociais que possuem comportamento político leve, moderado ou mínimo. As Figuras 34, 35 e 36 detalham a distribuição das severidades de cada variável dentro dos *clusters*.

Ao utilizar o *t-SNE* para observação e avaliação dos agrupamentos realizado pelo *DBSCAN*, a Figura 34 representa a distribuição das severidades da solidão em cada um dos 5 *clusters* obtidos pelo método *DBSCAN* (inicialmente mostrados na Figura 34). Em azul escuro apresenta-se a severidade da solidão mínima; em azul claro apresenta-se a severidade da solidão leve; em verde escuro apresenta-se a severidade da solidão moderada e em verde claro apresenta-se a severidade da solidão intensa. Foi possível notar que o *cluster* 1 (maior agrupamento de participantes formado) possui casos variados de solidão (mínima, leve, moderada e intensa). O *cluster* 3 também possui vários casos de severidade de solidão, exceto de severidade mínima. O *cluster* 4 possui apenas casos de solidão mínima entre os participantes. O *cluster* -1 representa os casos que não conseguiram ser agrupados.

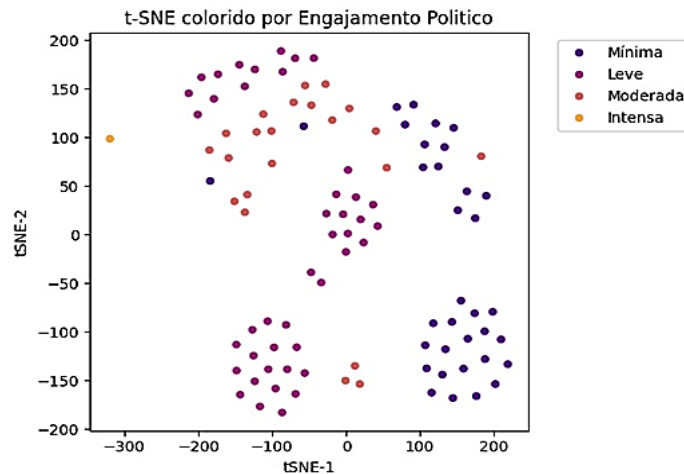
Figura 34- agrupamento dos participantes por severidade da solidão



Fonte: elaborada pelos autores.

Ao analisar a severidade de comportamento político de engajamento, a Figura 35 mostra que o *cluster* 1 apresenta casos variados de engajamento político com severidade leve, moderada e mínima, sem nenhum caso de severidade intensa entre os participantes agrupados neste cluster. O *cluster* 0, 2 e 4 apresentam todos os indivíduos com severidade de comportamento político mínimo. Os *cluster* 1 e 3 são os que apresentam indivíduos com severidade da solidão mínima e intensa. Um único participante (*cluster* -1) não agrupado possui severidade intensa para o comportamento político engajado. Já o *cluster* 0 possui casos predominantes de severidade leve e o *cluster* 2 de severidade mínima.

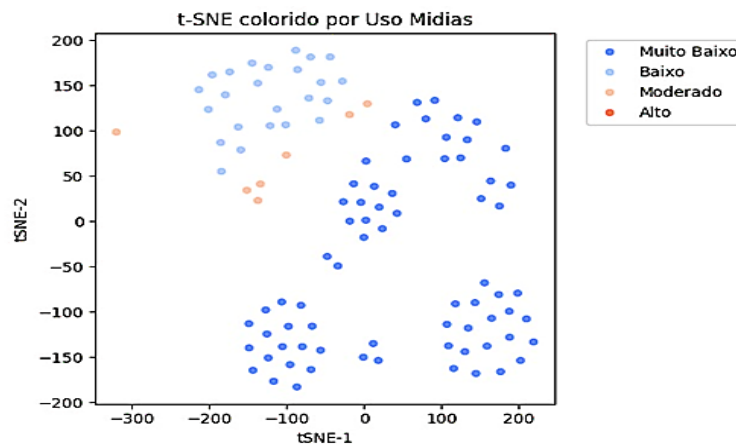
Figura 35- agrupamento dos participantes por severidade de comportamento político engajado



Fonte: elaborada pelos autores.

Ao analisar a severidade do uso de mídias sociais, a Figura 36 mostra que o *cluster 1* apresenta casos de severidade moderada e baixa entre os participantes agrupados. O *cluster -1* tem apenas um caso de severidade moderada para o uso de mídias sociais. Já os *cluster 0,2,3* e *4* possuem apenas casos de severidade muito baixa entre os participantes. Não foi encontrado nenhum caso de severidade alta entre os participantes.

Figura 36- agrupamento dos participantes por severidade do uso de mídias sociais



Fonte: elaborado pelos autores.

5.2.1. Testagem das hipóteses através dos resultados de saída DBSCAN

Os resultados de saída do *DBSCAN* não oferecem suporte para as hipóteses apresentadas nesta pesquisa (Quadro 1 no Capítulo 1), conforme observado anteriormente, nas Figuras 34, 35 e 36.

Embora a Hipótese 1 (H1) que previa uma relação positiva entre solidão intensa e engajamento político intenso não foi suportada, pois nenhum *cluster* combinou essas variáveis,

de acordo com o que foi observado nas Figuras 34 e 35. Os indivíduos com solidão intensa (*Clusters* 1 e 3) não se agruparam com base em alto engajamento político. O padrão identificado foi distinto, de modo que a solidão mínima emergiu como o perfil estável mais comum, associando-se principalmente a vários níveis de engajamento político (leve, mínimo e moderado). Contrariando a hipótese, os *outliers* (*Cluster* -1) revelaram um perfil oposto onde o engajamento político intenso coexiste com solidão mínima. Esse achado indica que o envolvimento político intenso, no contexto amostral estudado, pode ocorrer independentemente da experiência de solidão crônica significativa.

A hipótese 2 (H2), embora propusesse que a solidão intensa estaria associada a um alto uso de mídias sociais, os resultados do *DBSCAN* não sustentam esta associação. Ao contrário disso, como foi evidenciado na Figura 36, nenhum *cluster* foi caracterizado por alto uso de mídias sociais. Ao analisar as Figura 34 e Figura 36, observou-se que os sujeitos com solidão intensa se agrupam com base em um padrão mais baixo e moderado de uso de mídias sociais.

Já a Hipótese 3 (H3), embora previa uma associação positiva entre engajamento político intenso e alto uso de mídias sociais. O *DBSCAN* oferece suporte parcial, porém fraco, a esta hipótese. Isso é observado no *cluster* -1, nas Figuras 35 e 36 analisadas, sendo este o único que se aproxima do perfil previsto, apresentando a combinação de engajamento intenso com uso moderado (e não alto) de mídias. Este caso isolado sugere que a relação pode existir, mas é rara ou atípica na população estudada. Isso pode ser devido a uma evidência contrária e uma limitação amostral, na qual a ausência de um *cluster* dedicado a engajamento intenso e uso de mídias sociais alto.

A Figura 35 evidenciou a escassez geral de casos de engajamento intenso, presente apenas no *Cluster* -1, enquanto a Figura 36 mostrou a ausência de casos de alto uso de mídias sociais. Isto indica que, se essa relação existe, ela não é suficientemente representada ou *clusterizável* nos dados disponíveis. O padrão mais comum observado através desse modelo foi o engajamento político moderado e leve, presente nos *Clusters* 1, 3 e 4, que aparece majoritariamente associado ao uso muito baixo ou baixo de mídias. Isto mostra que níveis mais comuns de engajamento não dependem de um consumo elevado de plataformas.

Vale ressaltar que a luz da literatura, apresentada nas Seções de análise qualitativa no Capítulo 1 e discutida no Capítulo 2, também existe ausência de suporte empírico para essas hipóteses diretas. Conforme discutido, a investigação sobre essas relações no cenário brasileiro recente, não existem estudos quantitativos sobre a solidão como fenômeno sociopolítico no Brasil, que afeta jovens adultos até o momento. Estudos apontam a solidão associada a

depressão (Belens *et al.*, 2023), sendo de maior prevalência em idosos (Torres *et al.*, 2022) e manifestada nas mídias sociais mediante ao isolamento social prolongado (Josiowicz, 2021). No entanto, a solidão como manifestação da experiência migratória, foi uma questão social mais abrangente apontada (Nogueira, 2020).

Em relação ao uso de mídias como variável associada a esse fenômeno, ainda é incipiente, com estudos qualitativos frequentemente focados em plataformas específicas, *Twitter e Facebook*, e em dinâmicas particulares como mobilização e apoio emocional (Josiowicz, 2021; Belens *et al.* 2023; Torres *et al.*, 2023). O *Twitter* foi analisado como espaço dual, terapêutico através de leitura digital em tempos de recessão social (Josiowicz, 2021). O *Facebook*, de acordo com Belens *et al.* (2023), tem um papel na mediação de interações sociais mais prolongadas e na formação de comunidades temáticas. Através desta plataforma, criam-se redes de apoio que mitigam sentimentos de solidão associados ao transtorno. Torres *et al.* (2020) evidenciam o impacto positivo das interações mediadas por plataformas digitais na redução da solidão entre indivíduos acima de 50 anos.

A presente pesquisa buscou um enfoque quantitativo-comportamental mais amplo, abarcando múltiplas plataformas, com destaque para o uso frequente do *Instagram e YouTube*, majoritários na amostra (Tabela 9 na Seção 4.3.2.1.3.), mas esbarrou na própria realidade capturada, sendo um perfil majoritário de uso baixo de mídias para interação política no contexto analisado e a dissociação entre os estados intensos das variáveis. Isso sugere que as relações podem ser mais indiretas, mediadas ou contextualmente específicas do que as hipóteses bivariadas pressupunham para o cenário político investigado.

Existem marcadores de polarização política intensa no Brasil como o negacionismo, (Kobellarz *et al.*, 2024), principalmente o negacionismo científico e político (Razafindrakoto *et al.*, 2024), a resistência à vacinação contra COVID-19 (Viscardi *et al.*, 2024; Verjovsky *et al.*, 2024), o antipetismo (Gonçalves *et al.*, 2024), as percepções de censura em mídias sociais (Braga *et al.*, 2024), a desconfiança institucional, os debates sobre universidade pública (Lacerda; Santos, 2023), as discussões sobre políticas públicas e saúde (Santos; Teixeira, 2023), e o questionamento da ciência (Neves; Ferreira, 2020). Embora estes indicadores sejam evidentes de polarização no debate público apontados na literatura, Seção 3.4. no Capítulo 2, não foram indicadores capturados para manifestação como um perfil comportamental *clusterizável* nesta amostra. Isso aponta para a necessidade de modelos analíticos mais complexos que incorporem variáveis contextuais e mediadoras nesse sentido.

5.3. Análise descritivas de suporte sobre a associação entre as variáveis principais

Como as correlações lineares não capturaram relações robustas, a partir da análise de *clusters* por meio da implementação do *DBSCAN*, buscou-se evidências que investiguem se existem correlação entre as variáveis. Fez-se a contextualização estatística dos *clusters* identificados, apresentando análises descritivas das relações bivariadas entre as variáveis.

O algoritmo *KNN*, inicialmente selecionado para análise do *dataset* final com as mesmas três variáveis e pré-processamento do *DBSCAN* com intuito de classificação dos agrupamentos formados, foi descartado. Houve um erro de processamento, sendo a falta de rótulos no conjunto de dados após aplicação do *DBSCAN* que não conseguiu ser corrigido no tempo hábil dessa pesquisa. No entanto, o código de processamento do *KNN* se encontra completo no Apêndice G, considerando as correlações entre as variáveis realizadas no pré-processamento dos dados. Estas correlações entre as variáveis para clareza sobre a falta de evidências que sustentem as hipóteses apresentadas neste estudo foram apresentadas a seguir.

5.3.1. Hipótese 1: Associação entre solidão e comportamento político

A H1 propôs uma relação entre solidão intensa e o elevado comportamento político engajado. Em detalhe, nota-se que quando se analisa uma associação direta entre solidão intensa e comportamento político engajado intenso, a exploração realizada pelo modelo não evidenciou tal relação com significância estatística. Como mostra a Tabela 11, o teste estatístico aplicado para avaliar a significância dessa relação (*Qui-quadrado* = 6.80, *p-valor* = 0.6) não apresentou resultados significativos (considerando $p = 0,05$), indicando que não há correlação estatisticamente relevante entre os construtos nos dados disponíveis. Contudo, nota-se sim uma relação entre o comportamento político mínimo e leve com a severidade da solidão mínima, indicando valores elevados de correlação para eles.

Tabela 11 - Relação entre Severidade da Solidão e comportamento político

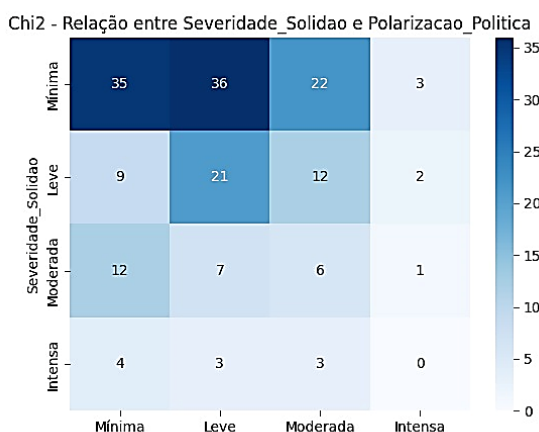
| Severidade da solidão | comportamento político | | | |
|-----------------------|------------------------|------|----------|---------|
| | Mínimo | Leve | Moderado | Intenso |
| Mínimo | 35 | 36 | 22 | 3 |
| Leve | 9 | 21 | 12 | 2 |
| Moderado | 12 | 7 | 6 | 1 |
| Intenso | 4 | 3 | 3 | 0 |

Fonte: elaborada pelos autores.

Na Figura 37, mostram-se os resultados por meio da matriz cruzada de relação entre severidade da solidão e comportamento político, sendo possível observar a contagem indivíduos em cada combinação das categorias. A maioria dos indivíduos concentra-se em

solidão mínima ou leve e comportamento político engajado mínimo ou leve. Observa-se também que nas categorias extremas (solidão intensa e comportamento político intenso) têm poucos casos (até 3 indivíduos). Nota-se que há uma possível tendência à medida que a solidão aumenta, de modo que o número de pessoas com comportamento político mínimo também diminui (de 35 indivíduos para 4). Contudo o comportamento intenso é raro em todos os níveis de solidão.

Figura 37- matriz de relação entre severidade da solidão e comportamento político



Fonte: elaborada pelos autores.

Os resultados sugerem que a amostra estudada não apresentou representatividade adequada para os extremos de solidão e comportamento político. Esse achado sugere que o perfil demográfico da amostra analisada, onde houve prevalência de respondentes jovens de 31 a 40 anos, gênero feminino, etnia branca, pós-graduados e com renda individual de dois salários-mínimos, não foi representativo para evidenciação do problema que associa níveis altos de solidão e comportamento político engajado com tendência radical. Isso se deve ao viés da amostra por ser um público mais jovem e com maior grau de escolaridade, como também pode ser inerente aos dados de autorrelato, considerando o intervalo temporal entre o período eleitoral (2022) e a coleta de dados fora deste período mais intenso (2023), que pode ter atenuado os efeitos do contexto nas percepções de solidão e do próprio comportamento político.

A natureza genérica das variáveis analisadas, também, pode ter influenciado, particularmente no que diz respeito à validação dos indicadores de comportamento político neste contexto que careceram de especificidade para capturar adequadamente a tendência investigada. Para abordar a variável política, torna-se importante caracterizar a severidade através de indicadores específicos de polarização político-partidária no Brasil em 2022.

5.3.2. Hipótese 2: Associação entre severidade da solidão e uso de mídias

A Tabela 12 mostra os resultados de saída na análise feita com apoio do *KNN*, segundo eles não foi evidenciada uma significância estatística sobre a relação entre solidão intensa e alto uso de mídias sociais, sendo apresentada uma discrepância entre os valores observados e esperados por meio do teste *Qui-quadrado* igual a 7.21 e *p-valor* igual a 0.3019.

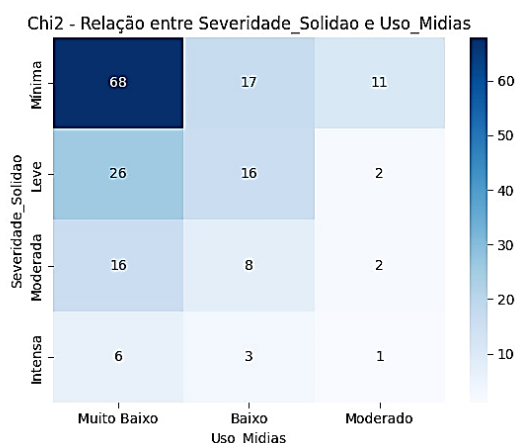
Contudo, novamente, quando analisada a relação entre a severidade da solidão mínima e o uso muito baixo de mídias sociais, nota-se que existe sim uma relação. Como visto na Figura 34, onde a matriz de confusão proporciona um elevado valor de correlação (68).

| Severidade da Solidão | Uso de Mídias | | |
|-----------------------|---------------|-------|----------|
| | Muito Baixo | Baixo | Moderado |
| Mínima | 68 | 17 | 11 |
| Leve | 26 | 16 | 2 |
| Moderada | 16 | 8 | 2 |
| Intensa | 6 | 3 | 1 |

Fonte: elaborada pelos autores.

A Figura 38 mostra também os resultados da utilização da matriz de relação cruzada para a distribuição da severidade da solidão em relação ao uso de mídias sociais. Dentre os padrões observados, a maior parte dos indivíduos se encontram na categoria muito baixo para uso de mídias, independentemente da solidão. A solidão mínima tem a maior frequência em todos os níveis de uso de mídias. O uso moderado é raro, especialmente em níveis mais altos de solidão, sendo apenas 1 caso em intensa.

Figura 38- Matriz de relação entre severidade da solidão e uso de mídias



Fonte: elaborada pelos autores.

Neste sentido, H2 deste estudo que associa valores intensos de solidão ao uso de mídias sociais, observa-se que o modelo apoiou parcialmente esta hipótese ao identificar relações de

similaridade entre essas variáveis. Contudo, essa ausência de significância estatística nesta associação ($p > 0.05$) indica que o uso isolado de mídias para relações política não classifica o sujeito como mais solitário.

Faz-se necessário também escolher e investigar uma plataforma específica mais utilizada para mobilização política, frequência do uso, o tipo de consumo e perfil dos usuários para interações políticas se acontece de forma passiva ou ativa nesta situação analisada. A literatura ainda apresenta muitas lacunas que possam associar o uso frequente destas plataformas de mídias sociais e altos níveis de severidade de solidão, principalmente em períodos como os das eleições presidenciais.

5.3.3. Hipótese 3: Associação entre comportamento político e mídias sociais

Este estudo propôs associação entre o comportamento político engajado e o uso de mídias sociais. Inicialmente, analisando a relação entre intenso comportamento político engajado e alto uso de mídias sociais (H3). Segundo os resultados observados na Tabela 13, estas variáveis, não apresentaram associação estatisticamente significativa nos dados analisados (em níveis intensos e elevados). E não há evidências suficientes para confirmar definitivamente a ausência de relação entre elas. O teste estatístico para H3 também revelou uma discrepância significativa entre os valores observados e esperados, com resultado de *Qui-quadrado* igual a 75.38 e $p\text{-valor} < 0.001$. Contudo, da mesma forma das relações anteriores, as variáveis Uso de mídias e Comportamento político apresentaram relação quando analisados os níveis mínimo e muito baixo, respectivamente.

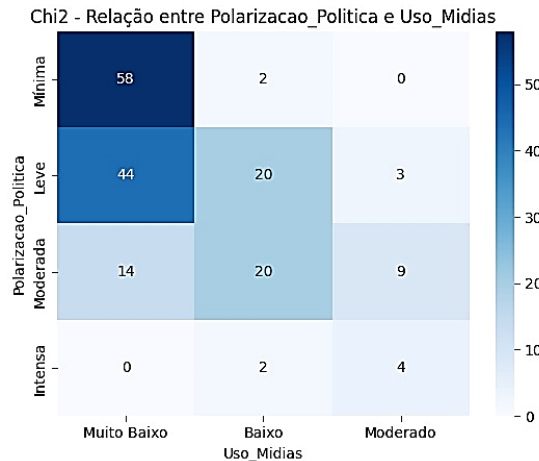
Tabela 13- Relação entre comportamento político e uso de Mídias

| Comportamento político | Uso de Mídias | | |
|------------------------|---------------|-------|----------|
| | Muito Baixo | Baixo | Moderado |
| Mínimo | 58 | 2 | 0 |
| Leve | 44 | 20 | 3 |
| Moderado | 14 | 20 | 9 |
| Intenso | 0 | 2 | 4 |

Fonte: elaborado pelos autores.

A Figura 39 mostra a matriz cruzada para a distribuição do comportamento político em relação ao uso de mídias sociais. Dentre os padrões observados foi possível analisar que o comportamento político "mínimo" está fortemente associado ao uso "muito baixo" de mídias (58 casos) e nenhum uso "moderado" (0 casos). O comportamento "intenso" ocorre apenas com uso "baixo" ou "moderado" (nunca "muito baixo"). Já o comportamento "moderado" tem maior equilíbrio entre uso "baixo" e "moderado".

Figura 39- Matriz de relação entre comportamento político e uso de mídias.



Fonte: elaborado pelos autores.

A análise descritiva de suporte das correlações bivariadas entre as variáveis solidão, comportamento político e uso de mídias sociais por níveis de severidade apresentaram-se fracas e não significativas. Compreende-se, portanto, que não oferecem suporte inicial às hipóteses de associação linear.

5.4. Discussão analítica dos resultados

A implementação do *DBSCAN* e a análise estatística de dados empíricos de uma microrregião brasileira evidenciaram a complexidade de padrões associativos entre as variáveis estudadas. Os dados subjetivos da amostra mostraram a necessidade de superar as limitações presentes na literatura predominante sobre solidão não só como fenômeno social, mas também psicológico no Brasil. Essas limitações das pesquisas sobre a temática são de ordem psicológica, conceitual, sociodemográfica e metodológica.

Os padrões identificados pelos algoritmos de *machine learning* corroboram o entendimento da solidão como fenômeno multidimensional e complexo (Cacioppo et al., 2015). Por meio deste método, considerado efetivo para análise da solidão (Abplanalp et al., 2024), as associações investigadas entre solidão, comportamento político e uso de mídias sociais revelaram a natureza não linear do fenômeno. Contudo, os modelos de clusterização utilizados demonstraram-se inadequados para capturar associações integradoras no contexto específico das eleições presidenciais brasileiras de 2022.

A ausência de clusters integradores, a tendência a associações bivariadas específicas e a exclusão de variáveis socioeconômicas e sociodemográficas configuram a complexidade de

compreender a solidão contemporânea para além de questões individuais, evidenciando determinantes sociais contextuais.

Este achado revela-se particularmente relevante ao contrastar com estudos realizados no Norte Global, onde a solidão tem sido tratada como risco à saúde pública, apontando para tendência de cronificação associada a questões sistêmicas (Hertz, 2021; Hawkey; Cacioppo, 2013). A preocupação com os riscos desse fenômeno refletido no Sul Global e em populações socialmente vulneráveis (Hertz, 2021) pode envolver variáveis que o presente estudo não conseguiu capturar, devido às limitações posteriormente apontadas.

Destaca-se que, no Brasil, as limitações conceituais e instrumentais precisam ser superadas. A compreensão da solidão como fenômeno psicológico ainda é predominante, com pouca variabilidade das variáveis demográficas envolvidas, apresentando lacunas significativas sobre seus aspectos sociais (Barroso et al., 2019; Junior et al., 2023).

No que se refere à associação entre solidão, comportamento político e uso de mídias sociais na amostra investigada, os resultados sugerem uma dinâmica distinta: a maioria dos participantes apresenta baixos níveis de solidão em todos os níveis de engajamento político. Este padrão desafia a generalização da narrativa de risco de cronificação do problema associado ao cenário político e ao uso intensivo de mídias sociais na polarização, apontando para a influência de mediações culturais e contextuais específicas. Sugere que as redes de apoio social no Brasil podem ser mais resilientes e que a formação de bolhas políticas digitais pode ocorrer por motivos não relacionados a níveis elevados de solidão, aspectos que merecem investigação aprofundada no Sul Global.

Para a compreensão abrangente da solidão, considerando os determinantes políticos e contextuais apresentados, enfatiza-se a necessidade de superar limitações instrumentais historicamente enfrentadas no país, como a validação tardia da Escala de Solidão UCLA (Barroso et al., 2016) e a escassez de estudos com amostras representativas abrangentes (Sandy Júnior et al., 2023). Essas limitações relacionam-se à divergência conceitual que permeia o campo de análise da solidão como fenômeno psicológico de alcance social (Minois, 2019; Hertz, 2021).

Os resultados obtidos reforçam a necessidade de desenvolver bases teórico-metodológicas sensíveis às particularidades contextuais brasileiras, capazes de capturar a interseção entre determinantes sociais, trajetórias individuais e mediações tecnológicas que conformam a experiência contemporânea da solidão. O período da pandemia de COVID-19 sob o governo Bolsonaro, reconhecido na literatura como contexto de manifestação e intensificação

desses fenômenos (Josiowicz, 2021; Torres et al., 2022), evidencia como a solidão se manifesta de forma singular conforme o contexto político e social (Minois, 2019), demandando abordagens analíticas igualmente contextualizadas para sua compreensão adequada.

5.5. Limitações da pesquisa

Embora este estudo forneça informações importantes sobre evidências de associação entre solidão, política e mídias sociais através de modelos de *ML*, esta investigação enfrentou uma série de limitações metodológicas, contextuais e conceituais que devem ser reconhecidas na interpretação dos resultados.

A principal delas está no fato de que a hipótese central de uma associação integradora entre solidão, engajamento político e uso de mídias sociais não pôde ser testada diretamente. A complexidade e a natureza não-linear dos dados revelaram que essa triangulação que é mais complexa do que uma relação causal simples, impedindo uma verificação empírica direta. Neste estudo não foram investigadas as variáveis mediadoras que poderiam explicar os mecanismos não percebidos claramente nas relações observadas.

O estudo enfrentou limitações para articular o conceito de solidão como fenômeno social dentro do referencial da terapia cognitivo-comportamental (TCC), devido ao tempo hábil da pesquisa e à escassez de pesquisas brasileiras que entrassem nessa discussão. Sugere-se integrar perspectivas sobre solidão social com abordagens psicológicas específicas. Esta lacuna na literatura nacional reflete um desafio mais amplo do campo que envolvem a integração entre dimensões sociais, estruturais e psicológicas.

Metodologicamente, o desenho transversal da pesquisa, com dados coletados em um único momento (2024), impossibilita entender trajetória e da solidão e dos comportamentos ao longo do tempo para caracterização do fenômeno, limitando a inferência de causalidade. Todos os dados basearam-se em autorrelato, podendo estarem enviesados pela memória sobre o evento, má recordação e sentimentos ou comportamentos menos intensos, mais elaborados e desejabilidade social, particularmente problemática para um questionário que remetia a comportamentos durante as eleições de 2022.

Na operacionalização dos algoritmos de *ML*, uma limitação significativa ocorreu com o *KNN* (*K-Nearest Neighbors*), onde não foi possível superar a falta de rótulos prévios para um treinamento supervisionado ideal, devido o tempo hábil do mestrado. Esta limitação foi contornada com o uso do *DBSCAN*, que identificou os grupos de forma não supervisionada. É crucial notar, porém, que os clusters identificados são descrições matemáticas da estrutura dos

dados; a interpretação causal e psicológica de porque esses perfis comportamentais existem permanece como uma inferência. Carece também de uma validação psicológica externa para os perfis gerados pelo modelo, o que significa que, embora estatisticamente robustos, os clusters não foram submetidos a uma avaliação qualitativa ou a critérios clínicos externos para confirmar sua coerência e significado psicológico no mundo real.

Do ponto de vista contextual, a pesquisa foi limitada por significativas lacunas na literatura científica. Observou-se uma escassez de dados robustos sobre a triangulação investigada, especialmente no cenário brasileiro. Há uma falta de consenso sobre o tema do extremismo político, com uma notável falta de dados e teorias sobre a manifestação de extremismo em indivíduos que se orientam para a esquerda do espectro político, em comparação com a relativa abundância de estudos focados no extremismo de direita. Este viés da literatura pode ter influenciado o marco teórico e a interpretação dos resultados.

O banco de dados desenvolvido também apresentou limitações, devido ao tamanho e o perfil da amostra (majoritariamente pós-graduadas) não permitem a generalização dos resultados para a população brasileira. Adicionalmente, identificou-se um viés temporal, uma vez que a pesquisa analisou fenômenos relacionados às eleições de 2022 com dados coletados em 2024. Quanto as ferramentas, a falta de contextualização e validação do questionário sobre comportamento político e as limitações dos instrumentos de medida de solidão disponíveis no país são fatores a serem considerados.

Por fim, as variáveis de perfil da amostra, bem como as sociodemográficas, socioeconômicas, qual mídia mais utilizada e perfil eleitoral foram excluídas, pois sua alta variabilidade introduziu ruído analítico. A identificação de relações não-lineares sinaliza a necessidade de futuras pesquisas que desenvolvam modelos específicos para testar essa complexa associação, preferencialmente com desenhos longitudinais, amostras representativas e delimitadas, a inclusão de mediadores psicológicos e, crucialmente, estudos qualitativos mistos para fornecer a necessária validação psicológica aos perfis estatísticos identificados, além de investigar de forma equilibrada as manifestações de engajamento político extremo em todo o espectro ideológico.

CAPÍTULO 6

6. CONSIDERAÇÕES FINAIS

A compreensão da solidão contemporânea demanda uma perspectiva sociopolítica, configurando-se como uma questão de saúde pública. O fenômeno, de natureza crônica em países do Norte Global, apresenta potencial de expansão no Sul Global, com crescente incidência em diferentes grupos etários e contextos. A literatura evidencia uma falta de consenso conceitual, resultando em definições ambíguas e fronteiras teóricas divergentes, o que torna premente a superação de abordagens centradas exclusivamente no indivíduo em favor da incorporação de determinantes sociais, políticos e de justiça social.

No contexto brasileiro, identificam-se lacunas significativas na investigação da relação entre solidão, política e mídias sociais. Embora existam estudos sobre a relação política-mídias sociais, influenciada pela polarização partidária e intensificada durante a pandemia de COVID-19, as associações entre solidão e política permanecem pouco exploradas. A relação entre solidão e mídias sociais, por sua vez, é um tema mais recorrente na literatura, mas ainda carece de evidências estatísticas robustas e de instrumentos validados. Ademais, faltam análises que considerem os diferentes níveis de severidade do problema com maior enfoque em variáveis sociodemográficas e socioeconômicas. A escassez de dados robustos e a predominância de evidências discursivas sobre estatísticas limitam a compreensão do fenômeno, agravada pela validação tardia de instrumentos como a Escala de Solidão *UCLA* no país.

Esta investigação, partindo da hipótese de associação linear entre solidão, engajamento político e uso de mídias sociais no Brasil de 2022, adotou uma abordagem metodológica híbrida e inovadora, integrando análise bibliográfica, psicometria e machine learning. Tal percurso contribuiu para uma compreensão multidimensional do problema, por meio do desenvolvimento de um conjunto de dados sobre as variáveis estudadas e da exploração de suas associações, revelando padrões complexos.

A aplicação de algoritmos de agrupamento (*K-means*, *DBSCAN* e *KNN*) demonstrou a superação de abordagens genéricas na investigação. Por meio deles, revelou-se a complexidade nas relações, refutando a hipótese inicial de uma associação linear entre as variáveis solidão, comportamento político e uso de mídias sociais. Identificou-se uma significativa heterogeneidade nos dados, com destaque para perfis contra-intuitivos, como indivíduos com alta politização e baixa solidão.

O modelo *DBSCAN* demonstrou maior eficácia em capturar a complexidade não linear do fenômeno, revelando a ausência de relação direta entre a intensidade do uso de mídias sociais e o engajamento político nos perfis majoritários. O perfil mais comum e estável identificado foi o de indivíduos com solidão mínima e baixo uso de mídias, independentemente do nível de engajamento político. Contudo, o achado mais significativo e contra-intuitivo foi a detecção de um grupo distinto com engajamento político intenso associado a solidão mínima. Esse perfil demonstra que a politização não é necessariamente motivada por sentimentos de solidão intensa, indicando a existência de outros fatores primários impulsionadores do engajamento político em ambientes digitais, os quais requerem investigação futura.

Não foi possível realizar uma análise efetiva por meio do *KNN*, devido à falta de atribuição de rótulos durante o processamento do conjunto de dados, conforme exigido no treinamento supervisionado. No entanto, sugere-se o uso de modelos supervisionados para a classificação dos agrupamentos formados, conforme proposto inicialmente.

Conclui-se pela natureza multifacetada dessas relações, nas quais a solidão não se apresenta como pré-condicionante do engajamento político. Os resultados não apenas respondem aos objetivos propostos, mas também delimitam com clareza o terreno empírico e metodológico a ser explorado em pesquisas futuras, realçando a complexidade de estudar fenômenos e a associação entre essas variáveis estudadas, as questões psicossociais envolvidas em um cenário digital e politicamente polarizado.

Recomendam-se futuras investigações com amostras representativas, incorporação de variáveis psicológicas mediadoras, validação de instrumentos para o contexto político brasileiro e aplicação de métodos mistos para interpretação dos perfis identificados, particularmente o grupo de alto engajamento político e baixa solidão. O estudo evidencia a inadequação de modelos lineares para fenômenos psicossociais complexos e a necessidade de abordagens interdisciplinares na análise da solidão em contextos do Sul Global.

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, P.; BERNARDO, L.; PONTES, T.; DAVIS, J.; DEODORO, T.; FERREIRA, R.; SOUZA, K.; MACDERMID, J.. Short-term impact of social distancing measures during the COVID-19 pandemic on cognitive function and health perception of Brazilian older adults: A pre-post study. **Journal of Applied Gerontology**, v. 40, n. 9, p. 934-942, 2021. Disponível em: https://journals.sagepub.com/doi/full/10.1177/07334648211015458?casa_token=OXkqicuRjeUAAAAA%3Aui2FcDsWmGyfr5rAJnScrfwxWssbfDh7AI5jP7aKCNJ_tNYQvPE0evQ1oRxGc5ON0CagzLF715WdrsM. Acesso em: 11 de jun. De 2025.

AL-SAGGAF, Y.; RAHMAN, M.; WIIL, U.. Using data mining to discover new patterns of social media and smartphone use and emotional states. **Social Network Analysis and Mining**, v. 14, n. 1, p. 90, 2024. Disponível em: <https://link.springer.com/article/10.1007/s13278-024-01243-8>. Acesso em: 20 de maio de 2024.

ABPLANALP, Samuel J.; GREEN, Michael; WYNN, Jonathan; EISENBERGER, Naomi; HORAN, William; LEE, Junghee ; MCCLEERY, Amanda; MIKLOWITZ, David; REDDY, L. Felice; REAVIS, Eric. Using machine learning to understand social isolation and loneliness in schizophrenia, bipolar disorder, and the community. **Schizophrenia**, v. 10, n. 1, p. 88, 2024. Disponível em: <https://www.nature.com/articles/s41537-024-00511-y#citeas>. Acesso em: 3 de jul. de 2025.

BARRETO, Manuela; DOYLE, David Matthew; QUALTER, Pamela. Changing the narrative: loneliness as a social justice issue. **Political Psychology**, v. 45, p. 157-181, 2024. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/pops.12965>. Acesso em: 24 de ago. de 2025.

BARROSO, Sabrina Martins et al. Evidências de validade da Escala Brasileira de Solidão UCLA. **Jornal Brasileiro de Psiquiatria**, v. 65, n. 1, p. 68-75, 2016. Disponível em: <https://www.scielo.br/j/jbpsiq/a/MjfQxLfddpKdQpKmZMB64DN/abstract/?lang=pt>. Acesso em: 3 de dez. de 2024.

BATISTA, Gustavo Enrique de Almeida Prado Alves. **Pré-processamento de dados em aprendizado de máquina supervisionado**. 2003. Tese (Doutorado) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2003. Disponível em: <https://www.teses.usp.br/teses/disponiveis/55/55134/tde-06102003-160219/publico/TeseDoutorado.pdf>. Acesso em: 10 de maio de 2025.

BEER, Paulo et al. The bias of social media: Truth and anxiety in engagement policies. **Revista Latinoamericana de Psicopatologia Fundamental**, v. 27, p. e231071, 2024. Disponível em: <https://www.scielo.br/j/rlpf/a/gGjXbJ8cbjPJVkStLgSj96r/abstract/?lang=en>. Acesso em: 10 de maio de 2025.

BELENS, Adroaldo; SANTOS, Ligia; GUIMARÃES, Jane. Connected loneliness: experiences of depression in a closed Facebook group. **Interface-Comunicação, Saúde, Educação**, v. 27, p. e220357, 2023. Disponível em: <https://www.scielo.br/j/icse/a/wp3phJLDdb9Lhk4LBGmTPhGF/abstract/?lang=en>. Acesso em: 4 de maio de 2025.

BERGUNO, George et al. Children's experience of loneliness at school and its relation to bullying and the quality of teacher interventions. **The qualitative report**, v. 9, n. 3, p. 483-499, 2004. Disponível em: https://scholar.google.com/scholar?hl=pt-BR&as_sdt=0%2C5&q=Children%E2%80%99s+experience+of+loneliness+at+school+and+its+relation+to+bullying+and+the+quality+of+teacher+interventions&btnG=. Acesso: 3 de abril de 2025.

BEST, JB. The politician: a public health problem? **Med J Aust**, v. 202, n. 8, p. 452-3, 4 maio 2015. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25929513/>. Acesso em: 3 de dez. de 2024.

BIANCOVILLI, Priscila; MAKSZIN, Lilla; JURBERG, Claudia. Misinformation on social networks during the novel coronavirus pandemic: a quali-quantitative case study of Brazil. **BMC Public Health**, v. 21, n. 1, p. 1200, 2021. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8220426/>. Acesso em: 5 de abril de 2025.

BRAGA, Alexandre Eustáquio Perpétuo et al. Censorship: A Reaction to Disinformation on the World Wide Web. **Investigación bibliotecológica**, v. 38, n. 98, p. 187-206, 2024. Disponível em: https://www.scielo.org.mx/scielo.php?pid=S0187-358X2024000100187&script=sci_arttext. Acesso em: 25 de maio de 2025.

BUTLER, J. El capitalismo tiene sus límites. In: V. V. A. A. Sopa de Wuhan. Pensamiento contemporáneo em tiempos de pandemias. Editora ASPO, 2020.

BROCK, A. From the Blackhand Side: Twitter as a cultural Conversation. **Journal of Broadcasting & Electronic Media**, v. 56, n. 4, p. 529-549, 2012. Disponível: <https://www.tandfonline.com/doi/abs/10.1080/08838151.2012.732147>. Acesso: 14 de jun. de 2025.

BROCK, A. Análise Crítica Tecnocultural do Discurso. In: SILVA, T. (org.). Comunidades, algoritmos e ativismos digitais: olhares afro-diaspóricos. São Paulo: LiteraRUA, 2020.

CACIOPPO, Stephanie et al. Loneliness: Clinical import and interventions. **Perspectives on psychological science**, v. 10, n. 2, p. 238-249, 2015. Disponível em: https://journals.sagepub.com/doi/10.1177/1745691615570616?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed. Acesso em: 20 de jun. de 2025.

DIAS, Joana Amaral. **O Cérebro da Política**. Leya, 2014.

DONATO, Helena; DONATO, Mariana. Etapas na condução de uma revisão sistemática. **Acta médica portuguesa**, v. 32, n. 3, p. 227-235, 2019. Disponível em: <https://www.actamedicaportuguesa.com/revista/index.php/amp/article/view/11923>. Acesso em: 10 de fev. de 2024.

CALLAWAY, E. et al. The push and pull of digital humanities: topic modeling the “what is digital humanities?” genre. **Digital Humanities Quarterly**. v. 14, n. 1, 2020. Disponível em:

<https://www.proquest.com/docview/2553557762?fromopenview=true&pq-origsite=gscholar&sourcetype=Scholarly%20Journals>. Acesso em: 13 de maio de 2025.

CACIOPPO, John T. et al. Loneliness within a nomological net: An evolutionary perspective. **Journal of research in personality**, v. 40, n. 6, p. 1054-1085, 2006. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S0092656606000055?casa_token=u95HRMWiJF4AAAAA:YNoiqDhGcd71ykEPO7X1chOqngOXsBU8cHP3DeYs5uWW1CCOeDpLEC6jmjowEqyyDU8Z1_f6uRQ. Acesso em: 10 de maio de 2025.

CARD, Kiffer G.; HEPBURN, Kirk J. Is neoliberalism killing us? A cross sectional study of the impact of neoliberal beliefs on health and social wellbeing in the midst of the COVID-19 pandemic. **International Journal of Social Determinants of Health and Health Services**, v. 53, n. 3, 2022. Disponível em: <https://journals-sagepub.com.ez38.periodicos.capes.gov.br/doi/10.1177/00207314221134040>. Acesso em: 2 de jul.de 2024.

CHEN, M. Keith; ROHLA, Ryne. The effect of partisanship and political advertising on closefamily ties. **Science**, v. 360, n. 6392, p. 1020-1024, 2018. Disponível em: <https://www.science.org/doi/abs/10.1126/science.aag1433>. Acesso em: 15 de março de 2023.

DAVIDOWITZ, SETH. **Todo mundo Mente: o que a internet e os dados dizem sobre quem realmente somos**. 1a Ed, Rio de Janeiro: Alta Books, 2018.

FARIA DA SILVA, Cássio et al. Quality of argumentation in political tweets: what is and how to measure it. **Revista de Estudos da Linguagem**, v. 29, n. 4, 2021. Disponível: https://openurl.ebsco.com/EPDB%3Aged%3A15%3A913151/detailv2?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Aged%3A151693238&crl=c&link_origin=scholar.google.com.br

GONÇALVES, Isabella et al. Evolution of Brazilian democracy: Unveiling election dynamics in political issues, negativity, and acclaim. **Politics and Governance**, v. 12, p. 1-16, 2024. Disponível em: <https://researchers.mq.edu.au/en/publications/evolution-of-brazilian-democracy-unveiling-election-dynamics-in-p>. Acesso em: 5 de agos. De 2025.

FARIA DA SILVA, Cássio et al. Quality of argumentation in political tweets: what is and how to measure it. **Revista de Estudos da Linguagem**, v. 29, n. 4, 2021. Disponível em: https://openurl.ebsco.com/EPDB%3Aged%3A15%3A913151/detailv2?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Aged%3A151693238&crl=c&link_origin=scholar.google.com.br. Acesso em: 3 de abril de 2025.

FERREIRA, Heloísa; CASEMIRO, Níldila. Solidão em idosos e fatores associados. **Revista Família, Ciclos de Vida e Saúde no Contexto Social**, v. 9, n. 1, p. 90-98, 2021. Disponível em: <https://www.redalyc.org/journal/4979/497968968009/html/>. Acesso em: 20 de jun. de 2024.

GÉRON, Aurélien. **Mãos à obra: aprendizado de máquina com Scikit-Learn. Keras & TensorFlow: Conceitos, ferramentas e técnicas para a construção de sistemas inteligentes**. [S. l.: sn], 2021.

GALHARDI, C.; Freire, N.; FAGUNDES, M.; MINAYO, M.; CUNHA, I. Fake News and Vaccine Hesitancy in the COVID-19 Pandemic in Brazil. **Ciência & Saúde Coletiva** v.1849, n.58. Disponível em: <https://doi.org/10.1590/1413-81232022275.24092021EN>. Acesso em: 20 de maio de 2024.

HAN, B. **Psicopolítica: o neoliberalismo e as novas técnicas de poder**. Áyiné, 2018.

HAN, Chan. **Infocracia: digitalização e a crise da democracia**. Petrópolis: Vozes, 2022.

HAN, Chan. **Sociedade do cansaço**. Petrópolis: Vozes, 2015.

HAN, Jiawei; KAMBER, Micheline; PEI, Jian. **Data Mining: Concepts and Techniques**. 3. ed. Burlington: Morgan Kaufmann, 2011. Disponível em: <https://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>. Acesso em: 24 de junho de 2025.

HAWKLEY, Louise C.; CACIOPPO, John T. Loneliness matters: A theoretical and empirical review of consequences and mechanisms. **Annals of behavioral medicine**, v. 40, n. 2, p. 218-227, 2010. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC3874845/>. Acesso em: 8 de agos. de 2025.

HEINEMAN, George T.; POLLICE, Gary; SELKOW, Stanley. **Algorithms in a Nutshell**. O'Reilly Media, Inc., 2009.

HWANG, Tzung-Jeng et al. Loneliness and social isolation during the COVID-19 pandemic. **International psychogeriatrics**, v. 32, n. 10, p. 1217-1220, 2020. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1041610224037773>. Acesso em: 2 de fev. de 2025.

HERTZ, N. **O século solitário: como restaurar a conexão humana em um mundo fragmentado**. 1a ed. Rio de Janeiro: Record, 2021.

HOZUMI, Yuta et al. UMAP-assisted K-means clustering of large-scale SARS-CoV-2 mutation datasets. **Computers in biology and medicine**, v. 131, p. 104264, 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0010482521000585>. Acesso em: 26 de agos. de 2025.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Portal do IBGE, 2022. Disponível em: <https://cidades.ibge.gov.br/brasil/mg/itajuba/panorama>. Acesso em: 18 de jul. de 2024.

JOSIOWICZ, Alejandra Judith. HUMANIDADES DIGITAIS E LEITURA NO TWITTER: “UM PLACEBO SANADOR EM TEMPOS DE COVID-19”. **Estudos Históricos** (Rio de Janeiro), v. 34, n. 73, p. 343-366, 2021. Disponível em: <https://www.scielo.br/j/eh/a/dhbtnyJQdTFmV8Fvpkwc98t>. Acesso em: 3 de maio de 2025.

KOBELLARZ, Jordan K. et al. Bubble reachers and uncivil discourse in polarized online public sphere. **Plos one**, v. 19, n. 6, p. e0304564, 2024. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0304564>. 10 de fev. de 2025.

KOBELLARZ, Jordan K. et al. Reaching the bubble may not be enough: news media role in online political polarization. **EPJ Data Science**, v. 11, n. 1, p. 47, 2022. Disponível em: https://epjds.epj.org/articles/epjdata/abs/2022/01/13688_2022_Article_357/13688_2022_Article_357.html. Acesso em: 15 de maio de 2025.

KOLAS, Janelle; VON MÜHLENEN, Adrian. Addicted to socialising and still lonely: A comparative, corpus-driven analysis of problematic social networking site use. **Journal of Behavioral Addictions**, v. 13, n. 1, p. 163-176, 2024. Disponível em: <https://akjournals.com/view/journals/2006/13/1/article-p163.xml>. Acesso em: 20 de maio de 2024.

LACERDA, Daniel S.; RITA DE CÁSSIA, T. Santos. The role of social network platforms for discursive legitimation: Unveiling neoliberalism behind the discourse on public universities. **M@ n@ gement**, v. 26, n. 4, p. 52-67, 2023. Disponível em: https://shs.cairn.info/article/MANA_264_0052?tab=texte-integral. Acesso em: 13 de abril de 2025.

LAMOTHE, Ryan. State-corporate capitalism, political polarization, and a culture of unquiet desperation: A pastoral analysis and responses. **Pastoral Psychology**, v. 61, p. 15-29, 2012.

LANGENKAMP, Alexander. Linking social deprivation and loneliness to right-extreme radicalization and extremist antifeminism. **Current Opinion in Behavioral Sciences**, v. 63, p. 101525, 2025. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2352154625000440?via%3Dihub#bib31>. Acesso em: 20 de maio de 2025.

LEE, M.; MARTIN, J. L. Coding, counting and cultural cartography. **American Journal of Cultural Sociology**, v.3, p. 1-33, 2015. <https://doi.org/10.1057/ajcs.2014.13>

LEUNG, C; HO, M.; BHARWANI, A.; COGO-MOREIRA, H.; WANG, Y.; CHOW, M. S; NI, M.. “Mental Disorders Following COVID-19 and Other Epidemics: A Systematic Review and Meta-Analysis. **Translational Psychiatry**; V. 12, n. 1, 2022. <https://doi.org/10.1038/s41398-022-01946-6>.

LO MORO, Giuseppina et al. Exploring cyberaggression and mental health consequences among adults: An Italian nationwide cross-sectional study. **International journal of environmental research and public health**, v. 20, n. 4. Disponível: <https://www.mdpi.com/1660-4601/20/4/3224>. Acesso em: 5 de maio de 2025.

PETERSON, Delaney et al. Loneliness is positively associated with populist radical right support. **Social Science & Medicine**, v. 366, 2025. Disponível em: <https://www.sciencedirect.com/science/article/pii/S027795362500005X?via%3Dihub#bib47>. Acesso em: 20 de jul. de 2025.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. Atlas, 2017.

MARTELETTO, Servio Reinaldo. Técnicas de seleção de atributos através do Random Forest: um estudo de caso para detecção de tendências em séries temporais financeiras. 2022. Dissertação (Mestrado em Modelagem de Sistemas Complexos) – Universidade de São Paulo, São Paulo, 2022. Disponível em: https://www.teses.usp.br/teses/disponiveis/100/100132/tdc-08032023-134543/publico/dissertacao_mestrado_SM_versao_corrigida.pdf. Acesso em: 3 de agos. de 2025.

MCKINNEY, Wes. **Python for data analysis**. O'Reilly Media, Inc., 2013.
MEJOVA, Yelena; HOMMADOVA LU, Anya. Gender in the disclosure of loneliness on Twitter during COVID-19 lockdowns. **Frontiers in Digital Health**, v. 5, 2023. Disponível em: <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2023.1297983/full>. Acesso em: 4 de maio de 2024.

MESSING, S.; WESTWOOD, S. J. Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. **Communication Research**, v.: 4, n:8, 1042-1063, 2014. Disponível em: https://journals.sagepub.com/doi/full/10.1177/0093650212466406?casa_token=Q6SVpz5ATBkAAAAA%3AzmQHihlb2flPj46F0cbUo8Gkf89LvDP93mCOwzOvDRSF3SegrFXynuDoTqUNaUQLgYxkfKIQSFZSLw. Acesso em: 12 de maio de 2023.

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES (MCTI). Inteligência Artificial. Brasília, 2021. Disponível em: <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/inteligenciaartificial#:~:text=Institu%C3%ADda%20pela%20Portaria%20MCTI%20n%C2%BA,v%C3%A1rias%20vertentes%2C%20que%20estimulem%20a>. Acesso em: 28 de maio de 2025.

MITCHELL, T. M. In **Machine Learning**. McGraw-Hill, 1997
MITCHELL, M.. **Artificial Intelligence: A Guide for Thinking Humans**. Penguin, 2019.
MINOIS, Georges. **A História da solidão e dos solitários**. Unesp: 2019.
MORIN, Edgar. **Ciência com Consciência**. Tradução de Maria D. Alexandre e Maria Alice Sampaio Dória. 8ª ed. Rio de Janeiro: Bertrand Brasil, 2005.

MORETTI, F. Patterns and interpretation. Pamphlet 15. Pamphlets of the Stanford Literary Lab., 2017.

MOTOKI, Fabio; PINHO NETO, Valdemar; RODRIGUES, Victor. More human than human: measuring ChatGPT political bias. **Public Choice**, v. 198, n. 1, p. 3-23, 2024. Disponível: <https://link.springer.com/article/10.1007/s11127-023-01097-2?fbclid=IwAR0Irs7jL2zsYHXSXPgO8v5C-uKEz-ccAQyt55LdVAWyi7Hrpe9gugRa7sw>
NADIKATTU, Rahul Reddy. The emerging role of artificial intelligence in modern society. *International Journal of Creative Research Thoughts*, 2016.
NEELY, S. R. Politically motivated avoidance in social networks: A study of facebook and the 2020 presidential election. **Social media and Society**, v.7, n.4, 2021. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/20563051211055438>. Acesso em: 12 de maio de 2023.

NEVES, André Luiz Machado das; FERREIRA, Breno de Oliveira. NARRATIVES BETWEEN SCIENCE AND POLITICS IN CHLOROQUINE ACTIVISM. *Psicologia & Sociedade*, v. 32, p. e020006, 2020. Disponível: <https://www.scielo.br/j/psoc/a/HKcwyWYWjB39fcFbbxFf44J/abstract/?lang=en>

NOGUEIRA, Silvia Garcia. Os estudantes timorenses em rede: laços sociais e conforto emocional frente à solidão da experiência migratória no Brasil e em Portugal. *Etnográfica. Revista do Centro em Rede de Investigação em Antropologia*, v. 24, n. 3), p. 577-601, 2020. Disponível: <https://journals.openedition.org/etnografica/7773>

O'NEIL. "Clinical Guidelines for the Use of Lifestyle-Based Mental Health Care in Major Depressive Disorder: World Federation of Societies for Biological Psychiatry (WFSBP) and Australasian Society of Lifestyle Medicine (ASLM) Taskforce". *The World Journal of Biological Psychiatry*: V.24, n.5, 2023. <https://doi.org/10.1080/15622975.2022.2112074>.

O'ROURKE, Hannah M. The global crisis of loneliness: a call for contextualised, mechanistic research. *The Lancet Healthy Longevity*, v. 5, n. 4, p. e241-e242, 2024. Disponível em: [https://www.thelancet.com/journals/lanhl/article/PIIS2666-7568\(24\)00030-8/fulltext](https://www.thelancet.com/journals/lanhl/article/PIIS2666-7568(24)00030-8/fulltext). Acesso em: 20 de maio de 2025.

PASCHOALOTTO, Marco Antonio Catussi et al. Politics and confidence toward the COVID-19 vaccination: A Brazilian cross-sectional study. *Human Vaccines & Immunotherapeutics*, v. 20, n. 1, p. 2318139, 2024. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/21645515.2024.2318139>

Pavel Hamet, Johanne Tremblay. Artificial intelligence in medicine. *Metabolism*. V. 69. Supplement. 2017. Pages S36-S40. ISSN 0026-0495, <https://doi.org/10.1016/j.metabol.2017.01.011>

PETERSON, Delaney et al. Loneliness is positively associated with populist radical right support. *Social Science & Medicine*, v. 366, p. 117676, 2025. Disponível em: <https://www.sciencedirect.com/science/article/pii/S027795362500005X?via%3Dihub#bib47>. Acesso em: 20 de jul. de 2025.

PITTMAN, Matthew; REICH, Brandon. Social media and loneliness: Why an Instagram picture may be worth more than a thousand Twitter words. *Computers in human behavior*, v. 62, p. 155-167, 2016. Disponível em: https://www.sciencedirect.com/science/article/pii/S0747563216302552?casa_token=aF1xqMmLookAAAAA:1x6vQK19GqoRVKU04QPieMEIXZQZIXwpV39epsakahQ61anTXOKitVi0mbu98Fywb1GvAzVanNg. Acesso em: 20 de maio de 2024.

PRECIADO, P. B. Aprendiendo del virus. In: V. V. A. A. Sopa de Wuhan. Pensamiento contemporáneo em tiempos de pandemias. Editora ASPO, 2020. p. 163-185.

PROHASKA, Thomas et al. Consensus statement: loneliness in older adults, the 21st century social determinant of health?. *BMJ open*, v. 10, n. 8, p. e034967, 2020. Disponível: <https://bmjopen.bmj.com/content/10/8/e034967.abstract>. Acesso em: 13 de abril de 2025.

RAMOS, Mozer de Miranda; MACHADO, Rodrigo de Oliveira; CERQUEIRA-SANTOS, Elder. "It's true! I saw it on WhatsApp": Social media, Covid-19, and political-ideological

orientation in Brazil. **Trends in Psychology**, v. 30, n. 3, p. 570-590, 2022. Disponível: <https://link.springer.com/article/10.1007/s43076-021-00129-4>

RAZAFINDRAKOTO, Mireille et al. Investigating the ‘Bolsonaro effect’ on the spread of the Covid-19 pandemic: an empirical analysis of observational data in Brazil. **Plos one**, v. 19, n. 4, p. e0288894, 2024. Disponível: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0288894>

RAYNAUT, Claude. Os desafios contemporâneos da produção do conhecimento: O apelo para interdisciplinaridade. Interthesis. 2014. Disponível em: <https://doi.org/10.5007/18071384.2014v11n1p1>

ROHLA, Ryne; JOHNSTON, R.; JONES, K., MANLEY, D. Spatial scale and the geographical polarization of the American electorate. **Political Geography**, v. 65, p. 117-122, 2018. Disponível em: https://www.sciencedirect.com/science/article/pii/S096262981730238X?casa_token=qlpO1bXCmNUAAAAA:RaMFLkOzzdT-IKVlczfDJ7MderXubTk3SJczT6HXeWaT1Vwdi7f76FgMfPLYmKy1b6VDMka92is. Acesso em: 14 de março de 2023.

RUSSEL DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. *JPers Assess.* 1996;66(1):20-40.

SANDY JÚNIOR, Paulo Afonso; BORIM, Flávia Silva Arbex; NERI, Anita Liberalesso. Solidão e sua associação com indicadores sociodemográficos e de saúde em adultos e idosos brasileiros: ELSI-Brasil. **Cadernos de saúde pública**, v. 39, 2023.

SANTOS, Jamilli Silva; TEIXEIRA, Carmen Fontes. Political action analysis of the Brazilian Health Care Reform Movement in the COVID-19 pandemic: 2020-2021. **Ciência & Saúde Coletiva**, v. 28, p. 1287-1296, 2023. Disponível em: <https://www.scielo.br/j/csc/a/B63CBTqG3rchQs4yCxvpB5J/?lang=pt>

SANTOS JUNIOR, Marcelo Alves dos. Long-term Fluctuations in Polarization in Brazil—Analysis of Political Information Sharing Between 2011 and 2019. *Dados*, v. 66, p. e20200076, 2022. Disponível em: <https://www.scielo.br/j/dados/a/R7CjTc36wwQ8JRqJXD3qTRt/abstract/?format=html&lang=en&stop=previous>

SENADO FEDERAL. **Senado aprova regulamentação da inteligência artificial; texto vai à Câmara.** Brasília, 10 dez. 2024. Disponível em: <https://www12.senado.leg.br/noticias/materias/2024/12/10/senado-aprova-regulamentacao-da-inteligencia-artificial-texto-vai-a-camara>. Acesso em: 28 de maio de 2025.

Teste Político dos 8 Valores. IDRLabs. Disponível em: <https://www.idrlabs.com/pt/teste-politico-dos-8-valores/teste.php>. Acesso em: 13 de set de 2023.

Teste de Coordenadas Políticas. IDRLabs. Disponível em: <https://www.idrlabs.com/pt/coordenadas-politicas/teste.php>. Acesso em: 13 de set de 2023.

TORRES, Juliana Lustosa et al. Loneliness and social disconnectedness in the time of pandemic period among Brazilians: evidence from the ELSI COVID-19 initiative. **Aging & Mental Health**, v. 26, n. 5, p. 898-904, 2022. Disponível: https://www.tandfonline.com/doi/full/10.1080/13607863.2021.1913479?casa_token=zfcdhgHktq0AAAAA%3AtNC5usXwLII89cNuoMHuOLyfEzkomwpoPHotBglogmfWD9gSoGBfYs5CzI36f7eFVV0y6T_hj1cv_MPtFQ.

TUBADJI, Annie. You'll never walk alone: Loneliness, religion, and politico-economic transformation. **Politics & Policy**, v. 51, n. 4, p. 661-695, 2023. Disponível: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/polp.12538?src=getftr>. Acesso em: 20 de maio de 2024.

UNDERWOOD, T. A genealogy of Distant Reading. **Digital Humanities Quarterly**, v. 11, n. 2, p. 1-12, 2017.

VAN DER VEEN, Olaf. Political Polarisation Compared: Creating a Comprehensive Index of Political Polarisation. URL: <https://www.google.com/url>, 2021.

VISCARDI, Lucas Henriques et al. The impact of conspiracy theories and vaccine knowledge on vaccination intention: a longitudinal study. **Ciência & Saúde Coletiva**, v. 29, p. e00352023, 2024. Disponível

WANG., Y., AHMED, S.; BEE, A. Selective avoidance as a cognitive response: Examining the political use of social media and surveillance anxiety in avoidance behaviours. **Behaviour and Information Technology**, .2023. Disponível em: https://www.tandfonline.com/doi/full/10.1080/0144929X.2023.2182609?casa_token=RXudh20kSdRUAAAAA%3A3q72SV6P3-3CqSp9srNhm-5UbDeQj9OeSwKYUCFTWjld7Znmgc8ZE329xGZuCxnnbxhCBle4--5drwyQ. Acesso em: 15 de março de 2023.

WEISS, Robert. Loneliness: The experience of emotional and social isolation. MIT press, 1975.

RUSSELL, Stuart; NORVIG, Peter; INTELLIGENCE, Artificial. A modern approach. Artificial Intelligence. Prentice-Hall, Englewood Cliffs, v. 25, n. 27, p. 79-80, 1995.

ROKACH, Ami; HEATHER, Brock. 1997. Loneliness: A multidimensional experience” **Psychology: A Journal of Human Behavior**: v. 34, n.1, 1997. <https://psycnet.apa.org/record/1997-38784-001>

ROKACH, Ami. **Loneliness Updated: Recent research on loneliness and how it affects our lives**. Routledge, 2013.

TORRES, J.; BRAGA, L.; MOREIRA, B.; CASTRO, S.; VAZ, C.; ANDRADE, A.; Andrade, F.; LIMA-COSTA, M.; CAIAFFA, W. “Loneliness and Social Disconnectedness in the Time of Pandemic Period among Brazilians: Evidence from the ELSI COVID-19 Initiative”. **Aging & Mental Health**, v.26, n. 5, 2022: 898–904. Disponível em: <https://doi.org/10.1080/13607863.2021.1913479>. Acesso em: 20 de maio de 2024.

VERJOVSKY, Marina et al. Political quarrel overshadows vaccination advocacy: How the vaccine debate on Brazilian Twitter was framed by anti-vaxxers during Bolsonaro administration. **Vaccine**, v. 41, n. 39, p. 5715-5721, 2023. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/37550146/>. Acesso em: 4 de agos. de 2025.

ZHANG, Senqi et al. The COVID-19 pandemic and mental health concerns on Twitter in the United States. **Health data science**, v. 2022, 2022. Disponível em: <https://spj-science.org.ez38.periodicos.capes.gov.br/doi/10.34133/2022/9758408>. Acesso em: 25 de abril de 2025.

ZUPIC, I.; ČATER, T.. “Bibliometric methods in management and organization”. **Organizational research methods**, v. 18, n. 3, 2015. <https://journals.sagepub.com/doi/10.1177/1094428114562629>. Acesso em: 3 de fev. de 2024.

WEEKS, DJ. A review of loneliness concepts, with particular reference to old age. *International Journal of Geriatric Psychiatry*. 1994;9:345–355.

WEISS, R. S. (Ed.). (1973). *Loneliness: The experience of emotional and social isolation*. Cambridge, MA: MIT Press.

WAKEFIELD, Robin L.; WAKEFIELD, Kirk. The antecedents and consequences of intergroup affective polarisation on social media. **Information Systems Journal**, v. 33, n. 3, p. 640-668, 2023. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1111/isj.12419>. Acesso em: 5 de abril de 2023.

Welcome to python.Org. Python.Org, 12 de dezembro de 2023, <https://www.python.org/>.
WEEKS, D. G., Michela, J. L., Peplau, L. A., & Bragg, M. E. The relation between loneliness and depression: A structural equation analysis. *Journal of Personality and Social Psychology*, 39, 1238–1244, 1980. doi:10.1037/h0077709. Disponível em: <https://psycnet.apa.org/record/1981-24754-001>. Acesso em: 5 de dez. de 2024.

WILKINSON, Eleanor. Loneliness is a feminist issue. *Feminist theory*, v. 23, n. 1, p. 23-38, 2022. Disponível em: <https://journals-sagepub-com.ez38.periodicos.capes.gov.br/doi/epub/10.1177/14647001211062739>. Acesso em: 24 de ago. de 2025.

YAN, BW; HSIA, RY; YEUNG, V. et al. Changes in mental health following the 2016 presidential election. **Journal of general internal medicine**, v. 36, p. 170-177, 2021. Disponível em: <https://link.springer.com/article/10.1007/s11606-020-06328-6>. Acesso em: 10 de março de 2023.

APÊNDICE

A- TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

Você está sendo convidada (o) a participar, desta pesquisa intitulada “**MACHINE LEARNING PARA ANÁLISE EXPLORATÓRIA E EVIDENCIAÇÃO EMPÍRICA DAS RELAÇÕES ENTRE SOLIDÃO, POLÍTICA E MÍDIAS SOCIAIS NO BRASIL**”, realizado pela pesquisadora **Jenaina de Fatima dos Santos**, aluna regular do Programa de Pós-Graduação em Desenvolvimento, Tecnologias e Sociedade da Universidade Federal de Itajubá, sob orientação do **Prof. Dr. Jesus Antonio Garcia Sanchez**.

Este estudo tem como objetivo analisar, através de modelos de inteligência artificial, as afetações da solidão no eleitorado brasileira mediante sua relação com a polarização política e uso das plataformas de mídias sociais. Contará com uma amostragem de no mínimo 100 pessoas do eleitorado de Itajubá, Minas Gerais, com idade entre 18 anos a 60 anos, aptos a votar e que façam uso de alguma mídia social. Tal estudo se faz importante, pois contribui para ampliar o entendimento do sentimento de solidão associado a problemas estruturais e sociais, como a desconexão política no contexto atual de polarização e do advento das mídias sociais. Visa em pensar formas de restabelecer as conexões humanas, refletindo sobre a importância de promover habilidades para convivência social e política. Vale esclarecer que a presente pesquisa não possui ligação ou posição partidária, tampouco vínculo com instituições ou com representações políticas partidárias. Também não pretende favorecer ou prejudicar algum partido. Leia com calma e atenção e tempo o presente termo.

PARTICIPAÇÃO NO ESTUDO

Sua participação neste estudo será de forma voluntária. Não acarretará nenhum valor econômico a ser recebido ou pago e você não precisará se deslocar. A pesquisa acontecerá de duas formas: 1) entrevista de forma presencial e individual, sendo disponibilizados em folhas impressas e 2) em ambiente virtual, através de um questionário eletrônico, no qual precisará de acesso à internet para respondê-lo.

Os pesquisadores solicitarão participação por meio de plataformas de mídias sociais e aplicarão os instrumentos Escala brasileira de Solidão UCLA e aos Questionários (demográfico e socioeconômico e de autopercepção do comportamento político, adaptados no *Google Forms*, sendo disponibilizado por meio *link*, precisando apenas ter acesso à internet para participar. Este *link* estará anexo em mensagem aos participantes, sendo um convite com algumas informações sobre a pesquisa. Por meio destes instrumentos serão analisados percepção sobre seus sentimentos de solidão e comportamentos políticos engajados e uso de plataformas de mídias sociais para interações políticas. Ao participar deste estudo será permitido que somente os pesquisadores responsáveis tenham acesso aos seus dados sobre sua saúde socioemocional e preferências políticas. Você tem liberdade de se recusar a participar e ainda se recusar a continuar participando em qualquer fase da pesquisa. Sempre que quiser poderá pedir mais informações sobre a pesquisa através do e-mail dos pesquisadores.

RISCOS

Os procedimentos adotados nesta pesquisa obedecem aos Critérios da Ética em Pesquisa com Seres Humanos, conforme Resolução no. 196/96 do Conselho Nacional de Saúde. Deste modo, a sua participação nesta pesquisa não traz complicações legais, oferecendo riscos mínimos a você. Haverá a possibilidade de sentir cansaço, desencadear reações antagônicas, desconforto e constrangimento ao responder as questões sensíveis que compõem o questionário.

Assim, para evitar e/ou reduzir os efeitos e as condições adversas, que possam causar dano ao participante da pesquisa, medidas de providências e cautelas serão empregadas pelos pesquisadores, tais como: os pesquisadores se colocaram à disposição para os devidos entendimentos e intervenção para amenizar tais sensações de desconforto. Haverá uma psicóloga na equipe de pesquisadores, que ficará à disposição para auxiliar caso os participantes sintam este desconforto. Após os participantes responderem a escala e questionário impresso ou eletrônico, os pesquisadores se comprometem a não divulgar qualquer informação pessoal dos participantes, apenas dados estatísticos, seguindo um tratamento de caráter sigiloso das informações, estando em conformidade com a Lei Geral de Proteção de Dados (LGPD). Após o fim da pesquisa, estes dados serão deletados de todos os dispositivos. Apenas a pesquisadora responsável e seu orientador poderão ter acesso a estes dados. Os dados que de alguma forma identifiquem os participantes coletados ficarão armazenados em dispositivos sem acesso à internet até o fim desta pesquisa, os quais não poderão ser compartilhados. Serão mantidos e disponibilizados para novos estudos apenas os dados brutos, isto é, os números e textos como resposta das questões de pesquisa em formato de tabela.

Também, com o intuito de propor uma redução, caso o entrevistado sinta algum tipo de desconforto ou risco, o mesmo poderá sair do estudo quando quiser, sem precisar se justificar e sem qualquer prejuízo a ele. No que se refere à participação, respondendo o questionário eletrônico, será assegurado pelos pesquisadores que suas informações não fornecidas, bem como o seu IP, não serão acessadas pelos pesquisadores. Caso desista de participar da pesquisa após responder ao questionário eletrônico, não será possível a retirada de suas respostas do banco de dados, uma vez que o questionário é anônimo, sendo necessário avisar os pesquisadores para não seguir com o uso destas informações.

BENEFÍCIOS

Ao participar desta pesquisa não há previsão de benefício direto individual participante ao, mas sim benefícios coletivos. Espera-se que este estudo traga informações que visem a promoção do bem-estar social no contexto político. E a longo prazo, espera-se que haja mais enfoque de novos estudos e criação de políticas públicas que prezam pelo diagnóstico e tratamento da solidão, amenizando seus riscos na atualidade.

SIGILO E PRIVACIDADE

Será respeitada sua privacidade, de modo que seu nome e qualquer outro dado que possa te identificar serão mantidos em sigilo durante todo o processo. Os pesquisadores se responsabilizam pela guarda e confidencialidade das informações, bem como o não compartilhamento e a não exposição dos dados pessoais na pesquisa, preservando assim o anonimato destes dados, durante todas as fases da pesquisa. Os dados obtidos não serão utilizados para outros fins que não seja o explícito nestes termos. Estes dados serão armazenados em dispositivo sem acesso a internet, até a finalização da pesquisa. Posteriormente, serão deletados. Serão divulgados somente dados estatísticos a partir das análises feitas. E devido ao fato do estudo ser realizado em ano eleitoral, a divulgação da pesquisa será feita após as eleições, sendo através de publicações em revistas científicas e participação em congressos científicos.

AUTONOMIA

Será garantida assistência a você de forma imediata, integral e gratuita, durante, após e/ou na interrupção da pesquisa. Assim como o livre acesso a todas as informações e esclarecimentos sobre o estudo e suas consequências durante a pesquisa, exceto se houver justificativa metodológica para tal (caso a informação venha a interferir nos métodos ou no

desfecho da pesquisa), apreciada e aprovada pelo Sistema CEP/CONEP. Você tem plena liberdade de se recusar a participar deste estudo ou retirar seu consentimento, em qualquer fase da pesquisa sem penalização alguma por parte dos pesquisadores ou da instituição.

RESSARCIMENTO E INDENIZAÇÃO

Os dados da pesquisa serão coletados presencialmente e todo material utilizado, bem como suas despesas e custos, será de inteira responsabilidade dos pesquisadores, o que não traz custos diretos ao participante. Caso o participante sofra danos diretos durante a pesquisa, haverá reparação material pelos pesquisadores envolvidos.

CONTATO

Você pode contactar os pesquisadores Jenaina de Fatima dos Santos e Jesus Antonio Garcia Sanchez pelos e-mails d2023100403@unifei.edu.br e jesus@unifei.edu.br ou ir até p Instituto de Engenharia de Produção e Gestão da Universidade Federal de Itajubá, onde está localizado o programa de Pós-graduação referido que estes pesquisadores estão vinculado, nas segundas-feiras e sextas-feiras, das 15 horas às 17 horas. Caso ache que a pesquisa não está sendo realizada de tal forma ou que está sendo prejudicado de alguma maneira, entre em contato com o Comitê de Ética em Pesquisa da Universidade Federal de Itajubá coordenado pelo Prof. Dr. Rossano Gimenes, situado na Av. BPS, 1303, Bairro Pinheirinho, Itajubá – MG, 37500 903, entrando em contato pelo telefone (35) 3629-1330 e (35) 99758-0874 ou pelo e-mail: cep@unifei.edu.br.

CONSENTIMENTO

Entendi todas as informações presentes neste Termo de Consentimento Livre e Esclarecido e tive a oportunidade de discutir e esclarecer sobre as informações relacionadas à pesquisa, ficando satisfeito (a) com as respostas. Entendo que receberei uma via impressa e datada deste documento e que outra via assinada e datada será arquivada pelos pesquisadores. Poderei também solicitar o acesso ao registro do consentimento sempre que necessário. Por fim, fui orientado (a) a respeito do que foi mencionado neste termo e compreendo a natureza e o objetivo do estudo, manifesto meu livre consentimento em participar, estando totalmente ciente de que não há nenhum valor econômico a receber ou a pagar por minha participação.

Observação: Não assine esse termo se ainda tiver dúvida a respeito

Itajubá- MG, _____ de _____ de _____.

Assinatura do (a) Participante da Pesquisa

Assinatura da Pesquisadora

B- QUESTIONÁRIO DE IDENTIFICAÇÃO DO PERFIL DOS PARTICIPANTES

A seguir serão apresentadas questões sociodemográficas, socioeconômicas, política e sobre uso de mídias sociais. Por favor, leia atentamente e assinale a que corresponde ao seu perfil.

E-mail: _____

Dados sociodemográficos

1. Faixa etária

18 anos 19 a 24 anos 25 a 30anos 31 a 44 anos
 45 a 54 anos 55 a 60 anos Mais de 60 anos

2. Gênero

Feminino Masculino outros _____

3. Estado civil

Solteiro (a) casada (o) . Divorciado (a) Outros

4. Etnia

Branco Indígena Pardo Preto

5. É natural de Itajubá, Minas Gerais?

sim não

6. Atualmente mora em Itajubá, Minas Gerais?

Sim Não

7. Caso resida em Itajubá - MG, em qual bairro você reside? Caso não resida em Itajubá responda não. _____

Dados socioeconômicos

8. Nível educacional

- Ensino fundamental completo
- Ensino Fundamental incompleto
- Ensino médio completo
- Ensino médio incompleto
- Ensino superior completo
- Ensino superior incompleto
- Pós-graduação
- Mestrado
- Doutorado
- Pós-doutorado

9. Renda familiar mensal

- Menos de 1 salário mínimo (Menos de R\$ 1412)
- 1 salário mínimo (até R\$ 1.412)
- De 1 a 2 salários mínimos (até R\$ 2.824)
- De 2 a 3 salários mínimos (até R\$4.236)
- De 3 a 4 salários mínimos (até R\$ 5.548)
- De 4 a 5 salários mínimos (até R\$ 7.060)
- De 5 a 6 salários mínimos (até R\$ 8.472)
- De 6 a 7 salários mínimos (até R\$ 9.884)
- Superior a 7 salários mínimos (mais de R\$ 9.884)

Dados sobre uso das mídias sociais

10. Você tem utilizado alguma plataforma de mídia social nos últimos anos (facebook, Instagram, Twitter, tik tok, Youtube...)?

Sim Não

11. Se sim, qual (quais) plataforma de mídia (s) social (ais) você tem utilizado para acompanhar assuntos políticos?

- Facebook
- Instagram
- LinkedIn
- Tik Tok
- Telegram
- Twitter
- Youtube
- Outra
- Não utilizo.

Dados eleitorais

12. Esteve apto a votar no segundo turno das eleições para presidente do Brasil em 2022?

- Sim
- Não

13. Votou no segundo turno das eleições para presidente do Brasil em 2022?

- Sim
- Não

**C- QUESTIONÁRIO SOBRE A AUTOPERCEPÇÃO DE COMPORTAMENTOS
POLÍTICOS E USO DE MÍDIAS SOCIAIS PARA INTERAÇÕES POLÍTICAS
DURANTE O SEGUNDO TURNO DAS ELEIÇÕES PARA PRESIDENTE DO
BRASIL EM 2022**

Instruções: Esta é uma ferramenta adaptada para análise do comportamento político engajado e uso de mídias para interações política. Abaixo você encontrará 28 questões afirmativas para serem respondidas referentes à sua autopercepção sobre seu comportamento político engajado e seu uso de mídias sociais para interações políticas **a partir do segundo turno das eleições para presidente do Brasil em 2022.**

Por favor, responda cada uma delas, indicando o nível de frequência em que você se percebeu tendo certas atitudes e afetos. Utilize a escala **Likert de 4 pontos, onde muito frequente= 4, frequentemente= 3, eventualmente= 2, raramente=1, nunca= 0.**

Ao final do questionário, você terá uma questão fechada sobre como participou desta pesquisa e uma aberta para descrever como se sentiu ao responder este questionário. É importante frisar que não há respostas certas ou erradas, apenas as que você se identifica no momento. Sua participação nesta pesquisa é muito importante para que possamos realizar o nosso estudo. Desde já agradecemos sua colaboração!

| | | | | | |
|---|---|---|---|---|---|
| 1. Me informei (o) sobre as ações do candidato a presidente da república que pertencia ao partido/coligação que me identifico. | 0 | 1 | 2 | 3 | 4 |
| 2. Trabalhei para um partido político ou movimento cívico | 0 | 1 | 2 | 3 | 4 |
| 3. Usei um emblema autocolante de campanha/movimento | 0 | 1 | 2 | 3 | 4 |
| 4. Participei de uma manifestação ou convocação feita por partidos políticos | 0 | 1 | 2 | 3 | 4 |
| 5. Tenho uma posição no espectro político, onde 0 representa nulo e 4 mais à esquerda. | 0 | 1 | 2 | 3 | 4 |
| 6. Tenho uma posição no espectro político, onde 0 representa nulo e 4 mais à direita. | 0 | 1 | 2 | 3 | 4 |
| 7. Me interessei mais pela política | 0 | 1 | 2 | 3 | 4 |
| 8. Me envolvi mais com assuntos ligados a política | 0 | 1 | 2 | 3 | 4 |
| 9. Tive preferência por algum partido ou coligação político | 0 | 1 | 2 | 3 | 4 |
| 10. Contatei um representante do poder local (prefeitos, deputados etc.) | 0 | 1 | 2 | 3 | 4 |
| 11. Conversei sobre notícias políticas e assuntos públicos com minha família. | 0 | 1 | 2 | 3 | 4 |
| 12. Conversei sobre notícias políticas e de assuntos públicos com amigos próximos ou conhecidos. | 0 | 1 | 2 | 3 | 4 |
| 13. Conversei sobre notícias políticas e assuntos públicos com colegas de trabalho. | 0 | 1 | 2 | 3 | 4 |
| 14. Conversei sobre notícias políticas e de assuntos públicos com vizinhos. | 0 | 1 | 2 | 3 | 4 |
| 15. Conversei sobre notícias políticas e de assuntos públicos com desconhecidos. | 0 | 1 | 2 | 3 | 4 |
| 16. Me relacionei mais com pessoas que têm o mesmo interesse partidário que eu. | 0 | 1 | 2 | 3 | 4 |
| 17. Frequentei espaços que declararam apoio ao mesmo partido que pelo qual me interessei | 0 | 1 | 2 | 3 | 4 |
| 18. Evitei pessoas que criticam o partido que me identifico. | 0 | 1 | 2 | 3 | 4 |
| 19. Evitei pessoas que têm ideologias políticas se opõem a minha. | 0 | 1 | 2 | 3 | 4 |
| 20. Rompi relacionamentos por causa do partido político das pessoas. | 0 | 1 | 2 | 3 | 4 |
| 21. Quando alguém criticou o partido que apoio pareceu um insulto pessoal. | 0 | 1 | 2 | 3 | 4 |
| 22. Quando alguém elogiou um partido que apoio, pareceu um elogio pessoal | 0 | 1 | 2 | 3 | 4 |
| 23. Me preocupei com o que as pessoas pensam sobre o partido que apoio. | 0 | 1 | 2 | 3 | 4 |
| 24. Passo tempo nas plataformas de mídias sociais, vendo notícias ou programas sobre política partidárias e assuntos relacionados | 0 | 1 | 2 | 3 | 4 |
| 25. Nas plataformas de mídias sociais, deixei de seguir ou bloqueei perfis que defendiam partidos políticos os quais não me identifico. | 0 | 1 | 2 | 3 | 4 |

| | | | | | |
|---|---|---|---|---|---|
| 26. Nas plataformas de mídias sociais, discuti com estranhos, hostilizando algum partido político. | 0 | 1 | 2 | 3 | 4 |
| 27. Nas plataformas de mídias sociais, deixei de seguir ou excluí amizades que defendem partidos políticos que não me identifico. | 0 | 1 | 2 | 3 | 4 |
| 28. Nas plataformas de mídias sociais, deixei de seguir ou excluí amizades que defendem partidos políticos que não me identifico. | 0 | 1 | 2 | 3 | 4 |

(2) Jowell and the Central Coordinating Team, European Social Survey 2006/2007: Technical Report, London: Centre for Comparative Social Surveys, City University (2007);(3)Wang, Y., Ahmed, S., & Bee, A. W. T..elective avoidance as a cognitive response: examining the political use of social media and surveillance anxiety in avoidance behaviours. *Behaviour & Information Technology*, 1-15, 2023.;(4)Wakefield, R. L., & Wakefield, K. . The antecedents and consequences of intergroup affective polarisation on social media. *Information Systems Journal*, 33(3), 640-668,2023 (5)Yan, BW; Hsia, RY; Yeung, V. et al. Mudanças na saúde mental após a eleição presidencial de 2016. *J GEN INTERN MED* 36, 170–177 (2021). <https://doi.org/10.1007/s11606-020-06328-6>

D- DATASET DA PESQUISA

SANTOS, Jenaina; SANCHEZ, Jesus (2025), *Dataset* sobre a associação entre solidão, comportamento político e mídias sociais no contexto do segundo turno das eleições presidenciais de 2022 em uma microrregião brasileira, Mendeley Data, V1, doi: 10.17632/983b46494s.1. Disponível em: <https://data.mendeley.com/datasets/983b46494s/1>

E- CÓDIGO DE IMPLEMENTAÇÃO DO K-MEANS EM PYTHON

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.cluster import KMeans
from sklearn.decomposition import PCA
from sklearn.ensemble import RandomForestClassifier
import matplotlib.pyplot as plt
import seaborn as sns

# Caminho para o arquivo Excel
file_path = r'E:\Jesus\Pesquisa\Jenaina\PesquisaRespostas.xlsx'

# Ler os dados da planilha
try:
    df = pd.read_excel(file_path)
except Exception as e:
    print(f"Erro ao ler o arquivo: {e}")
    raise

# Inicializar o LabelEncoder
label_encoder = LabelEncoder()

# Listas das colunas categóricas conforme suas especificações
binary_columns = [
    'Gênero',
    'É natural de Itajubá-MG?',
    'Mora em Itajubá-MG atualmente?',
    'Utiliza alguma plataforma de mídia social nos últimos anos?',
    'Votou no segundo turno das eleições para presidente do Brasil em
2022?'
]

multi_columns = [
    'Faixa etária',
    'Etnia',
    'Estado civil',
    'Escolaridade',
    'Renda financeira aproximada',
    'Qual mídia social você tem utilizado para acompanhar assuntos
políticos?',
    'No segundo turno das eleições para presidente do Brasil em 2022,
você esteve apto a votar?'
]

# Colunas que precisam de transformação numérica direta
direct_transform_columns = [
    'Faixa etária',
    'Gênero',
    'Etnia',
    'Estado civil',
    'É natural de Itajubá-MG?',
    'Mora em Itajubá-MG atualmente?',
    'Caso resida em Itajubá-MG, em qual bairro você reside?',
    'Escolaridade',
```

```

'Renda financeira aproximada',
'Utiliza alguma plataforma de mídia social nos últimos anos?',
'No segundo turno das eleições para presidente do Brasil em 2022,
você esteve apto a votar?',
'Votou no segundo turno das eleições para presidente do Brasil em
2022?'
]

# Substituir valores ausentes por uma string específica, por exemplo,
'missing'
df.fillna('missing', inplace=True)

# Transformar as colunas categóricas
category_mappings = {}
for col in direct_transform_columns:
    if col in df.columns:
        df[col] = df[col].astype(str)
        df[col + '_original'] = df[col] # Salvar a coluna original
        df[col] = label_encoder.fit_transform(df[col])
        category_mappings[col] = {index: label for index, label in
enumerate(label_encoder.classes_)}
    else:
        print(f"A coluna {col} não existe no DataFrame")

# Categorizar 'Qual mídia social você tem utilizado para acompanhar
assuntos políticos?'
if 'Qual mídia social você tem utilizado para acompanhar assuntos
políticos?' in df.columns:
    df['Uso_Mídias_Sociais'] = df['Qual mídia social você tem utilizado
para acompanhar assuntos políticos?'].apply(
        lambda x: 'Não utiliza mídias sociais' if x == 'Não utiliza
mídias sociais' else 'Utiliza mídias sociais para assuntos políticos'
    )
else:
    print(f"A coluna 'Qual mídia social você tem utilizado para
acompanhar assuntos políticos?' não existe no DataFrame")

# Verificar a nova coluna
print(df[['Qual mídia social você tem utilizado para acompanhar
assuntos políticos?', 'Uso_Mídias_Sociais']].head())

# Definir colunas relacionadas à solidão
loneliness_columns = df.columns[12:32] # Colunas 14 a 34

# Ajustar os valores das colunas de solidão
df[loneliness_columns] = df[loneliness_columns].replace({
    0: 0,
    1: 1,
    2: 2,
    3: 3,
    4: 3
})

# Calcular a severidade da solidão
df['Severidade_Solidão'] = df[loneliness_columns].sum(axis=1)
df['Severidade_Solidão'] = pd.cut(df['Severidade_Solidão'], bins=[-1,
22, 35, 47, 60], labels=['Mínima', 'Leve', 'Moderada', 'Intensa'])

```

```

# Excluir as colunas específicas mencionadas
columns_to_exclude = [
    'É natural de Itajubá-MG?',
    'Mora em Itajubá-MG atualmente?',
    'No segundo turno das eleições para presidente do Brasil em 2022,
    você esteve apto a votar?',
    'Votou no segundo turno das eleições para presidente do Brasil em
    2022?'
]

df.drop(columns=columns_to_exclude, inplace=True, errors='ignore')

# Excluir colunas de dummies criadas anteriormente
dummy_columns =
df.columns[df.columns.str.contains('|'.join(['Facebook', 'YouTube',
'Instagram', 'Twitter', 'LinkedIn', 'TikTok', 'WhatsApp', 'Reddit',
'Snapchat', 'Pinterest']))]
df.drop(columns=dummy_columns, inplace=True, errors='ignore')

# Verificar se todas as colunas a serem escaladas são numéricas
excluded_columns = [
    'ID',
    'Severidade_Solidão'
] + list(loneliness_columns)

columns_to_scale =
df.select_dtypes(include=[np.number]).columns.difference(excluded_columns)

# Normalizar os dados (exceto as colunas excluídas)
scaler = StandardScaler()
df[columns_to_scale] = scaler.fit_transform(df[columns_to_scale])

# Verificar os dados normalizados
print(df.head())

# Encontrar o número ideal de clusters usando o método do cotovelo
sse = []
for k in range(1, 11):
    kmeans = KMeans(n_clusters=k, random_state=42)
    kmeans.fit(df[columns_to_scale])
    sse.append(kmeans.inertia_)

# Plotar o gráfico do cotovelo
plt.figure(figsize=(10, 6))
plt.plot(range(1, 11), sse, marker='o')
plt.xlabel('Número de Clusters')
plt.ylabel('SSE')
plt.title('Método do Cotovelo')

# Definir o número ideal de clusters (ajuste com base no gráfico do
cotovelo)
num_clusters = 4

# Treinar o modelo k-means
kmeans = KMeans(n_clusters=num_clusters, random_state=42)

```

```

kmeans.fit(df[columns_to_scale])

# Adicionar as labels ao dataframe original
df['Cluster'] = kmeans.labels_

# Verificar os resultados
print(df.head())

# Analisar características dos clusters
cluster_analysis = df.groupby('Cluster')[columns_to_scale].mean()
print("Análise dos Clusters:\n", cluster_analysis)

# Plotar os clusters em um gráfico 2D usando PCA
pca = PCA(n_components=2)
pca_result = pca.fit_transform(df[columns_to_scale])
df['PCA1'] = pca_result[:, 0]
df['PCA2'] = pca_result[:, 1]

plt.figure(figsize=(10, 6))
sns.scatterplot(x='PCA1', y='PCA2', hue='Cluster', data=df,
palette='viridis', s=50, alpha=0.7, edgecolor='w')

# Plotar os centros dos clusters
centroids_pca = pca.transform(kmeans.cluster_centers_)
plt.scatter(centroids_pca[:, 0], centroids_pca[:, 1], s=300, c='red',
marker='X', edgecolor='black', linewidths=2, zorder=10,
label='Centroides')
plt.legend()
plt.title('Clusters k-means (PCA)')
plt.xlabel('PCA1')
plt.ylabel('PCA2')

# Importância das características para os clusters
rf = RandomForestClassifier(n_estimators=100, random_state=42)
rf.fit(df[columns_to_scale], df['Cluster'])
importances = rf.feature_importances_
features = columns_to_scale

feature_importances = pd.DataFrame({'Feature': features, 'Importance':
importances})
feature_importances = feature_importances.sort_values(by='Importance',
ascending=False)
print("Importância das Características:\n", feature_importances)

# Visualização da importância das características
plt.figure(figsize=(15, 10))
sns.barplot(x='Importance', y='Feature', data=feature_importances)
plt.title('Importância das Características para os Clusters')
plt.xlabel('Importância')
plt.ylabel('Características')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()

# Visualização usando gráficos de barras
plt.figure(figsize=(15, 8))
sns.countplot(x='Faixa etária', hue='Cluster', data=df,
palette='viridis')

```

```

plt.xlabel('Faixa etária')
plt.ylabel('Contagem')
plt.legend(title='Clusters', bbox_to_anchor=(1.05, 1), loc='upper
left')
plt.title('Distribuição de Faixa Etária por Cluster')
plt.tight_layout()

plt.figure(figsize=(15, 8))
sns.countplot(x='Escolaridade', hue='Cluster', data=df,
palette='viridis')
plt.xlabel('Escolaridade')
plt.ylabel('Contagem')
plt.legend(title='Clusters', bbox_to_anchor=(1.05, 1), loc='upper
left')
plt.title('Distribuição de Escolaridade por Cluster')
plt.tight_layout()

if 'Faixa etária_original' in df.columns:
    # A coluna existe, então pode plotar o gráfico
    plt.figure(figsize=(15, 8))
    sns.countplot(x='Faixa etária_original', hue='Cluster', data=df,
palette='viridis')
    plt.xlabel('Faixa etária_original')
    plt.ylabel('Contagem')
    plt.legend(title='Clusters', bbox_to_anchor=(1.05, 1), loc='upper
left')
    plt.title('Distribuição da Faixa Etária Original por Cluster')
    plt.tight_layout()
else:
    # A coluna não existe, então informe o erro
    print("Erro: A coluna 'Faixa etária_original' não existe no
dataframe.")

plt.figure(figsize=(12, 8))
sns.countplot(x='Cluster', hue='Severidade_Solidão', data=df,
palette='viridis')
plt.title('Distribuição da Severidade da Solidão por Cluster')
plt.xlabel('Cluster')
plt.ylabel('Contagem')
plt.legend(title='Severidade da Solidão', bbox_to_anchor=(1.05, 1),
loc='upper left')
plt.tight_layout()

# A coluna existe, então pode plotar o gráfico
# plt.figure(figsize=(15, 8))
# sns.countplot(x='8. Me envolvi mais com assuntos ligados a
política_original', hue='Cluster', data=df, palette='viridis')
# plt.xlabel('8. Me envolvi mais com assuntos ligados a
política_original')
# plt.ylabel('Contagem')
# plt.legend(title='Clusters', bbox_to_anchor=(1.05, 1), loc='upper
left')
# plt.title('Distribuição Envolvimento Assuntos Política Original por
Cluster')
# plt.tight_layout()

plt.show()

```

F- CÓDIGO DE IMPLEMENTAÇÃO DO DBSCAN EM PYTHON

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler, LabelEncoder,
OrdinalEncoder
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, classification_report,
confusion_matrix
from sklearn.ensemble import RandomForestClassifier
from sklearn.cluster import DBSCAN, KMeans
from sklearn.decomposition import PCA
from sklearn.manifold import TSNE
from scipy.stats import chi2_contingency
import matplotlib.pyplot as plt
import seaborn as sns

# Caminho para o arquivo Excel
file_path =
r'E:\Jesus\Pesquisa\Orientados\Jenaina\PesquisaRespostasA.xlsx'

# Ler os dados da planilha
df = pd.read_excel(file_path)
df.fillna('missing', inplace=True)

# Codificação
label_encoder = LabelEncoder()
categorical_columns = ['Votou no 2° turno', 'Apto a votar no 2° turno',
'Mídias sociais utilizadas']
for col in categorical_columns:
    if col in df.columns:
        df[col] = df[col].astype(str)
        df[col] = label_encoder.fit_transform(df[col])

# Seleção de colunas das escalas
loneliness_cols = df.columns[3:22]
polarization_cols = df.columns[22:48]
midias_cols = df.columns[48:52]

# Ajuste dos dados numéricos e criação de severidades
df[loneliness_cols] = df[loneliness_cols].apply(pd.to_numeric,
errors='coerce').fillna(0).clip(0, 3)
df['Severidade_Solidao'] = df[loneliness_cols].sum(axis=1)
df['Severidade_Solidao'] = pd.cut(df['Severidade_Solidao'], bins=[-1,
22, 35, 47, 60], labels=['Mínima', 'Leve', 'Moderada', 'Intensa'])

df[polarization_cols] = df[polarization_cols].apply(pd.to_numeric,
errors='coerce').fillna(0).clip(0, 3)
df['Polarizacao_Politica'] = df[polarization_cols].sum(axis=1)
df['Polarizacao_Politica'] = pd.cut(df['Polarizacao_Politica'], bins=[-1,
23, 46, 69, 92], labels=['Mínima', 'Leve', 'Moderada', 'Intensa'])

df[midias_cols] = df[midias_cols].apply(pd.to_numeric,
errors='coerce').fillna(0).clip(0, 4)
df['Uso_Midias'] = df[midias_cols].sum(axis=1)
df['Uso_Midias'] = pd.cut(df['Uso_Midias'], bins=[-1, 5, 10, 15, 20],
labels=['Muito Baixo', 'Baixo', 'Moderado', 'Alto'])
```

```

# Remover colunas redundantes
df.drop(columns=list(loneliness_cols) + list(polarization_cols) +
list(midias_cols), inplace=True)

# Normalização dos dados numéricos restantes
excluded_cols = ['Severidade_Solidao', 'Polarizacao_Politica',
'Uso_Midias']
numeric_cols =
df.select_dtypes(include=np.number).columns.difference(excluded_cols)
scaler = StandardScaler()
df[numeric_cols] = scaler.fit_transform(df[numeric_cols])

# Codificação ordinal para análise e visualização
for col in ['Severidade_Solidao', 'Polarizacao_Politica',
'Uso_Midias']:
    df[col] = df[col].astype('category')

encoder = OrdinalEncoder()
encoded = encoder.fit_transform(df[['Severidade_Solidao',
'Polarizacao_Politica', 'Uso_Midias']])

# Correlação entre escalas
plt.figure(figsize=(6, 5))
sns.heatmap(np.corrcoef(encoded.T), annot=True, xticklabels=['Solidão',
'Polarização', 'Mídias'], yticklabels=['Solidão', 'Polarização',
'Mídias'], cmap='coolwarm', fmt=".2f")
plt.title("Correlação entre escalas (ordinal)")
plt.tight_layout()
plt.show()

# PCA
pca = PCA(n_components=2)
pca_result = pca.fit_transform(encoded)
plt.figure(figsize=(8, 6))
plt.scatter(pca_result[:, 0], pca_result[:, 1],
c=df['Severidade_Solidao'].cat.codes, cmap='viridis', alpha=0.7)
plt.title("PCA - Projeção 2D das Escalas")
plt.xlabel("Componente Principal 1")
plt.ylabel("Componente Principal 2")
plt.colorbar(label='Severidade Solidão (códigos)')
plt.tight_layout()
plt.show()

# t-SNE ajustado
tsne = TSNE(n_components=2, perplexity=30, learning_rate=100,
n_iter=1000, random_state=42)
tsne_result = tsne.fit_transform(encoded)
df['tSNE-1'] = tsne_result[:, 0]
df['tSNE-2'] = tsne_result[:, 1]

# Visualizações coloridas por severidade
def plot_tsne_by_feature(feature, palette):
    g = sns.scatterplot(data=df, x='tSNE-1', y='tSNE-2', hue=feature,
palette=palette)
    g.set_title(f"t-SNE colorido por {feature.replace('_', ' ')}")
    plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')

```

```

plt.tight_layout()
plt.show()

plot_tsne_by_feature('Severidade_Solidao', 'viridis')
plot_tsne_by_feature('Polarizacao_Politica', 'plasma')
plot_tsne_by_feature('Uso_Midias', 'coolwarm')

# DBSCAN com parâmetros ajustados
dbscan = DBSCAN(eps=50, min_samples=5)
df['DBSCAN_Cluster'] = dbscan.fit_predict(tsne_result)

# Visualizar clusters DBSCAN (sem elipses)
plt.figure(figsize=(8, 6))
sns.scatterplot(x='tSNE-1', y='tSNE-2', hue='DBSCAN_Cluster', data=df,
palette='tab10', legend='full', alpha=0.6)
plt.title("Clusters DBSCAN sobre t-SNE")
plt.tight_layout()
plt.show()

# Visualizar distribuição das severidades dentro de cada cluster
for feature in ['Severidade_Solidao', 'Polarizacao_Politica',
'Uso_Midias']:
    print(f"\nDistribuição de '{feature}' por cluster DBSCAN:")
    crosstab = pd.crosstab(df['DBSCAN_Cluster'], df[feature],
normalize='index')
    print(crosstab.round(2))
    crosstab.plot(kind='bar', stacked=True, colormap='viridis',
figsize=(8, 5))
    plt.title(f'Distribuição de {feature} por Cluster DBSCAN')
    plt.xlabel('Cluster')
    plt.ylabel('Proporção')
    plt.legend(title=feature, bbox_to_anchor=(1.05, 1), loc='upper
left')
    plt.tight_layout()
    plt.show()

# KMeans (para comparação)
kmeans = KMeans(n_clusters=4, random_state=42)
kmeans_labels = kmeans.fit_predict(encoded)

plt.figure(figsize=(8, 6))
plt.scatter(pca_result[:, 0], pca_result[:, 1], c=kmeans_labels,
cmap='Accent', alpha=0.7)
plt.title("Clusters KMeans sobre PCA")
plt.xlabel("PCA 1")
plt.ylabel("PCA 2")
plt.tight_layout()
plt.show()

```

G- CÓDIGO DE IMPLEMENTAÇÃO DO KNN EM PYTHON

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler, LabelEncoder#,
OneHotEncoder
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split#, GridSearchCV,
cross_val_score, validation_curve
from sklearn.metrics import accuracy_score, classification_report,
confusion_matrix, f1_score
from sklearn.ensemble import RandomForestClassifier
#from sklearn.pipeline import Pipeline
#from sklearn.decomposition import PCA
from scipy.stats import chi2_contingency
#from imblearn.over_sampling import SMOTE # Para balanceamento de
classes
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import seaborn as sns

# Caminho para o arquivo Excel
#file_path =
r'E:\Jesus\Pesquisa\Orientados\Jenaina\PesquisaRespostasA.xlsx'

# Ler os dados da planilha
try:
    df = pd.read_excel('PesquisaRespostasA.xlsx')
except Exception as e:
    print(f"Erro ao ler o arquivo: {e}")
    raise

# Preencher valores ausentes
df.fillna('missing', inplace=True)

# Transformações de colunas categóricas (exemplo para colunas binárias
e multiclasse)
label_encoder = LabelEncoder()

# Exemplo para colunas que precisam ser codificadas (ajuste conforme
suas colunas)
categorical_columns = ['Votou no 2° turno', 'Apto a votar no 2° turno',
'Mídias sociais utilizadas']
for col in categorical_columns:
    if col in df.columns:
        df[col] = df[col].astype(str)
        df[col] = label_encoder.fit_transform(df[col])

# Definir colunas das escalas (ajuste o intervalo conforme o seu
arquivo)
# Exemplo baseado nos seus índices (ajuste conforme necessário):
loneliness_cols = df.columns[3:22] # solidão (20 colunas)
polarization_cols = df.columns[22:48] # polarização política (26
colunas)
midias_cols = df.columns[48:52] # mídias sociais (4 colunas)

# Ajustar valores para a escala de solidão
```

```

# Caso seus dados usem escala 0-4 e 4 seja o máximo, pode manter, ou
ajustar aqui:
df[loneliness_cols] = df[loneliness_cols].apply(pd.to_numeric,
errors='coerce').fillna(0).clip(0, 3)

# Soma das respostas para criar severidade (exemplo)
df['Severidade_Solidao'] = df[loneliness_cols].sum(axis=1)
df['Severidade_Solidao'] = pd.cut(df['Severidade_Solidao'],bins=[-1,
22, 35, 47, 60], labels=['Mínima', 'Leve', 'Moderada', 'Intensa'])

# Polarização política
df[polarization_cols] = df[polarization_cols].apply(pd.to_numeric,
errors='coerce').fillna(0).clip(0, 3)
df['Polarizacao_Politica'] = df[polarization_cols].sum(axis=1)
df['Polarizacao_Politica'] = pd.cut(df['Polarizacao_Politica'], bins=[-
1, 23, 46, 69, 92], labels=['Mínima', 'Leve', 'Moderada', 'Intensa'])

# Uso de mídias sociais
df[midias_cols] = df[midias_cols].apply(pd.to_numeric,
errors='coerce').fillna(0).clip(0, 4)
df['Uso_Midias'] = df[midias_cols].sum(axis=1)
df['Uso_Midias'] = pd.cut(df['Uso_Midias'], bins=[-1, 5, 10, 15, 20],
labels=['Muito Baixo', 'Baixo', 'Moderado', 'Alto'])

# Remover colunas originais das escalas para evitar redundância
df.drop(columns=list(loneliness_cols) + list(polarization_cols) +
list(midias_cols), inplace=True)

# Normalizar colunas numéricas restantes (exceto as severidades
categóricas)
excluded_cols = ['Severidade_Solidao', 'Polarizacao_Politica',
'Uso_Midias']
numeric_cols =
df.select_dtypes(include=np.number).columns.difference(excluded_cols)
scaler = StandardScaler()
df[numeric_cols] = scaler.fit_transform(df[numeric_cols])

# Para usar as severidades, converter para categorias numéricas para ML
for col in ['Severidade_Solidao', 'Polarizacao_Politica',
'Uso_Midias']:
    df[col] = df[col].astype('category')

# Criar variáveis dummy para severidades
X = pd.get_dummies(df[['Severidade_Solidao', 'Polarizacao_Politica',
'Uso_Midias']], drop_first=True)
y = df['Severidade_Solidao'].cat.codes
#y = X.copy() # Considerando multiclasse/multi-target

# Separar em treino/teste
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)

# Treinar KNN
knn = KNeighborsClassifier(n_neighbors=4)
knn.fit(X_train, y_train)
y_pred = knn.predict(X_test)

```

```

print("Acurácia do modelo KNN (severidades das escalas):",
accuracy_score(y_test, y_pred))
print(classification_report(y_test, y_pred))

# Treinar Random Forest
rf = RandomForestClassifier(n_estimators=100, random_state=42)
rf.fit(X_train, y_train)
y_pred_rf = rf.predict(X_test)

print("Acurácia do modelo Random Forest (severidades das escalas):",
accuracy_score(y_test, y_pred_rf))
print(classification_report(y_test, y_pred_rf))

# Importância das features (Random Forest)
importances = rf.feature_importances_
feature_importances = pd.DataFrame({'Feature': X.columns, 'Importance':
importances}).sort_values(by='Importance', ascending=False)

plt.figure(figsize=(10, 6))
sns.barplot(data=feature_importances, y='Feature', x='Importance',
palette='viridis')
plt.title('Importância das Features no Random Forest')
plt.xlabel('Importância')
plt.ylabel('Feature')
plt.tight_layout()
plt.show()

# Análise estatística de relação entre escalas
def analyze_relationship(scale1, scale2):
    crosstab = pd.crosstab(df[scale1], df[scale2])
    chi2, p, dof, expected = chi2_contingency(crosstab)
    print(f"\nRelação entre {scale1} e {scale2}")
    print("Tabela cruzada:\n", crosstab)
    print(f"Qui-quadrado: {chi2:.2f}, p-valor: {p:.4f}")
    sns.heatmap(crosstab, annot=True, cmap="Blues")
    plt.title(f"Chi2 - Relação entre {scale1} e {scale2}")
    plt.show()

analyze_relationship('Severidade_Solidao', 'Polarizacao_Politica')
analyze_relationship('Severidade_Solidao', 'Uso_Midias')
analyze_relationship('Polarizacao_Politica', 'Uso_Midias')

# Visualização 3D das categorias (pontos com jitter para melhor
visualização)
def visualize_3d():
    fig = plt.figure(figsize=(8, 8))
    ax = fig.add_subplot(111, projection='3d')

    x = df['Severidade_Solidao'].cat.codes + np.random.uniform(-0.2,
0.2, len(df))
    y = df['Polarizacao_Politica'].cat.codes + np.random.uniform(-0.2,
0.2, len(df))
    z = df['Uso_Midias'].cat.codes + np.random.uniform(-0.2, 0.2,
len(df))

    scatter = ax.scatter(x, y, z, c=df['Severidade_Solidao'].cat.codes,
cmap='viridis', alpha=0.7)

```

```

ax.set_xlabel('Severidade Solidão')
ax.set_ylabel('Polarização Política')
ax.set_zlabel('Uso de Mídias')
plt.title('Relação 3D: Solidão x Polarização x Mídias')
plt.colorbar(scatter, label='Severidade Solidão (códigos)')
plt.show()

visualize_3d()

# Função corrigida para treinar e avaliar KNN para prever uma
# severidade a partir das outras duas
def knn_relationship(target, features):
    print(f"\nTreinando KNN para prever '{target}' usando {features}
    como preditores.")

    # Preparar dados
    X = pd.get_dummies(df[features], drop_first=True)
    y = df[target].cat.codes # Target como numérico

    # Separar treino/teste
    X_train, X_test, y_train, y_test = train_test_split(X, y,
    test_size=0.2, random_state=42)

    # Treinar KNN
    knn = KNeighborsClassifier(n_neighbors=5)
    knn.fit(X_train, y_train)

    # Prever e avaliar
    y_pred = knn.predict(X_test)
    print(f"Acurácia KNN para {target}: {accuracy_score(y_test,
    y_pred):.4f}")
    print("Matriz de Confusão:\n", confusion_matrix(y_test, y_pred))
    print(classification_report(y_test, y_pred))

    # Plot matriz de confusão
    plt.figure(figsize=(6, 5))
    sns.heatmap(confusion_matrix(y_test, y_pred), annot=True, fmt='d',
    cmap='Blues')
    plt.title(f'KNN - Matriz de Confusão para {target}')
    plt.xlabel('Predito')
    plt.ylabel('Verdadeiro')
    plt.show()

from sklearn.preprocessing import OrdinalEncoder
encoder = OrdinalEncoder()
encoded = encoder.fit_transform(df[['Severidade_Solidao',
'Polarizacao_Politica', 'Uso_Midias']])
sns.heatmap(np.corrcoef(encoded.T), annot=True,
xticklabels=encoder.categories_, yticklabels=encoder.categories_)
plt.title("Correlação entre escalas (codificadas ordinalmente)")
plt.show()

# Rodar a função para prever Severidade_Solidao
knn_relationship('Severidade_Solidao', ['Polarizacao_Politica',
'Uso_Midias'])

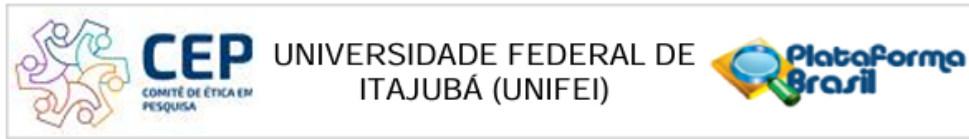
# Outras severidades

```

```
knn_relationship('Polarizacao_Politica', ['Severidade_Solidao',  
'Uso_Midias'])  
knn_relationship('Uso_Midias', ['Severidade_Solidao',  
'Polarizacao_Politica'])
```

ANEXOS

A- PARECER CONSUBISTANCIADO DO CEP DA UNIFEI



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: MACHINE LEARNING PARA ANÁLISE EXPLORATÓRIA E EVIDENCIAÇÃO EMPÍRICA DAS RELAÇÕES ENTRE SOLIDÃO, POLÍTICA E MÍDIAS SOCIAIS NO BRASIL

Pesquisador: JENAINA DE FATIMA DOS SANTOS

Área Temática:

Versão: 4

CAAE: 76839323.3.0000.0356

Instituição Proponente: UNIVERSIDADE FEDERAL DE ITAJUBA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 7.923.684

Apresentação do Projeto:

Trata-se de uma pesquisa que se orienta por uma abordagem interdisciplinar quali-quantitativa. Será realizada na cidade de Itajubá, Minas Gerais. Conterá com a participação de eleitores com idade entre 18 anos a 60 anos. Tem como objetivo analisar a relação entre as variáveis referentes à solidão, política e mídias sociais. Na análise dos resultados será empregado o método de Inteligência Artificial, machine learning. Os

resultados podem expor os grupos formados e os padrões comportamentais semelhantes de associação entre as variáveis (solidão, política e uso

de mídias sociais) no Brasil. O entendimento da solidão como problema social e estrutural pode ser evidenciado por estas associações feitas com estas variáveis no contexto atual.

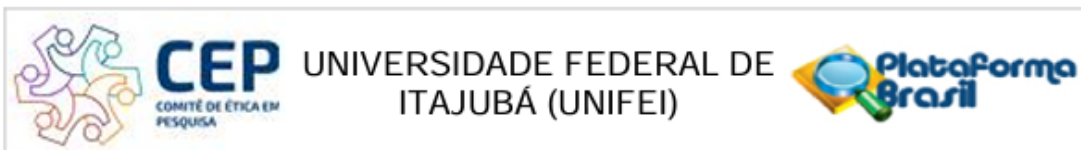
Objetivo da Pesquisa:

Evidenciar, através de modelos de inteligência artificial, a relação entre os níveis de solidão, comportamento político e mídias sociais no contexto das eleições presidenciais no Brasil em 2022.

Avaliação dos Riscos e Benefícios:

Os riscos e benefícios estão devidamente expostos e foram objeto de discussão na

Endereço: Av. BPS 1303 Bairro: Pinheirinho Prédio da Administração Central 4º Andar PRPPG Sala do
Bairro Pinheirinho **CEP:** 37.500-903
UF: MG **Município** ITAJUBA
Telefone (35)3629-1330 **E-** cep@unifei.edu.br



Continuação do Parecer: 7.923.684

primeira fase de tramitação. Como se trata de uma emenda, observou-se que não houve nenhuma alteração significativa neste contexto.

Comentários e Considerações sobre a Pesquisa:

Este parecer se debruça sobre uma emenda da pesquisa, já avaliada, mas que sofreu alterações em razão de recomendações da banca de qualificação de mestrado, que não produziram mudanças essenciais no corpo do estudo.

Considerações sobre os Termos de apresentação obrigatória:

Sem alterações em relação ao primeiro processo de análise, portanto, estão adequados.

Recomendações:

Nada a recomendar.

Conclusões ou Pendências e Lista de Inadequações:

Considerando que é uma emenda que contempla sugestões da banca de qualificação, as quais não modificaram fundamentalmente a estrutura do estudo e não houve geração de questões éticas adicionais, a emenda foi aceita e aprovada.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

| Tipo Documento | Arquivo | Postagem | Autor | Situação |
|---|---------------------------------------|------------------------|------------------------------|----------|
| Informações Básicas do Projeto | PB_INFORMAÇÕES_BÁSICAS_2630927_E2.pdf | 17/09/2025 17:34:23 | | Aceito |
| TCLE / Termos de Assentimento / Justificativa de Ausência | Carta_de_Emenda_.pdf | 17/09/2025 17:27:47 | JENAINA DE FATIMA DOS SANTOS | Aceito |
| Folha de Rosto | folhaDeRosto.pdf | 17/09/2025 17:15:18 | JENAINA DE FATIMA DOS SANTOS | Aceito |
| Solicitação Assinada pelo Pesquisador Responsável | Solicitacao_Emenda.pdf | 17/05/2024 23:23:09 | JENAINA DE FATIMA DOS SANTOS | Aceito |
| TCLE / Termos de Assentimento / Justificativa | Termo_Emenda_Modificado.pdf | 17/05/2024 22:31:32 | JENAINA DE FATIMA DOS SANTOS | Aceito |

Endereço: Av. BPS 1303 Bairro: Pinheirinho Prédio da Administração Central 4º Andar PRPPG Sala do
Bairro Pinheirinho **CEP:** 37.500-903
UF: MG **Município** ITAJUBA
Telefone (35)3629-1330 **E-** cep@unifei.edu.br

Continuação do Parecer: 7.923.684

| | | | | |
|--|-------------------------------|------------------------|------------------------------------|--------|
| de Ausência | Termo_Emenda_Modificado.pdf | 17/05/2024 22:31:32 | JENAINA DE FATIMA DOS SANTOS | Aceito |
| Projeto Detalhado / Brochura Investigador | Projeto_Emenda_Modificado.pdf | 17/05/2024 22:29:11 | JENAINA DE FATIMA DOS SANTOS | Aceito |
| Projeto Detalhado / Brochura Investigador | PROJETO_MODIFICADO.pdf | 20/03/2024 21:37:02 | JENAINA DE FATIMA DOS SANTOS | Aceito |
| Solicitação registrada pelo CEP | Solicitacao_cep.pdf | 20/03/2024 21:36:48 | JENAINA DE FATIMA DOS SANTOS | Aceito |
| Projeto Detalhado / Brochura Investigador | Projetode_pesquisa.pdf | 21/12/2023 17:59:53 | JENAINA DE FATIMA DOS SANTOS | Aceito |
| TCLE / Termos de Assentimento / Justificativa de Ausência | Termode_consentimento.pdf | 21/12/2023 17:49:31 | JENAINA DE FATIMA DOS SANTOS | Aceito |

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

ITAJUBA, 24 de Outubro de 2025

Assinado por:
Luiz Felipe Silva
(Coordenador(a))

Endereço: Av. BPS 1303 Bairro: Pinheirinho Prédio da Administração Central 4º Andar PRPPG Sala do
Bairro Pinheirinho **CEP:** 37.500-903
UF: MG **Município** ITAJUBA
Telefone (35)3629-1330 **E-** cep@unifei.edu.br

B- ESCALA BRASILEIRA DE SOLIDÃO UCLA (UCLA-BR)

Instruções: Abaixo encontram-se várias afirmativas sobre a forma como alguém pode se sentir. Por favor, leia atentamente todas elas e marque nas alternativas a frequência com que você se sente, utilizando a escala *likert* de 4 pontos (0 a 4):

Nunca= 0 Raramente=1 Algumas vezes= 2 frequentemente=3 muito frequente 4= muito frequente

| | | | | |
|---|---|---|---|---|
| 1.Eu me sinto infeliz por fazer tantas coisas sozinho(a). | 0 | 1 | 2 | 3 |
| 2.Eu não tolero ficar tão sozinho(a). | 0 | 1 | 2 | 3 |
| 3.Eu sinto que não tenho companhia. | 0 | 1 | 2 | 3 |
| 4.Eu sinto que ninguém me compreende. | 0 | 1 | 2 | 3 |
| 5.Eu fico esperando as pessoas me ligarem ou escreverem. | 0 | 1 | 2 | 3 |
| 6.Eu sinto que não tenho ninguém a quem eu possa recorrer. | 0 | 1 | 2 | 3 |
| 7. Eu não me sinto próximo a ninguém. | 0 | 1 | 2 | 3 |
| 8.Sinto que meus interesses e ideias não são compartilhados por aqueles que me rodeiam. | 0 | 1 | 2 | 3 |
| 9.Eu me sinto excluído(a). | 0 | 1 | 2 | 3 |
| 10.Eu me sinto completamente sozinho(a). | 0 | 1 | 2 | 3 |
| 11.Eu sou incapaz de me aproximar e de me comunicar com as pessoas ao meu redor. | 0 | 1 | 2 | 3 |
| 12.Eu sinto que minhas relações sociais são superficiais. | 0 | 1 | 2 | 3 |
| 13.Eu me sinto carente de companhia. | 0 | 1 | 2 | 3 |
| 14.Eu sinto que ninguém me conhece realmente bem. | 0 | 1 | 2 | 3 |
| 15.Eu me sinto isolado(a) das outras pessoas. | 0 | 1 | 2 | 3 |
| 16.Sou infeliz estando tão excluído(a). | 0 | 1 | 2 | 3 |
| 17.Para mim é difícil fazer amigos. | 0 | 1 | 2 | 3 |
| 18.Eu me sinto bloqueado(a) e excluído(a) por outras pessoas. | 0 | 1 | 2 | 3 |
| 19.Sinto que as pessoas estão ao meu redor, mas não estão comigo. | 0 | 1 | 2 | 3 |
| 20. Eu me sinto incomodado(a) em realizar atividades sozinho(a). | 0 | 1 | 2 | 3 |

Referências: (1) Barroso SM, Andrade VS, Midgett AH, Carvalho RGN. Evidências de validade da Escala Brasileira de Solidão UCLA. J Bras Psiquiatr, v.65, n.1, 2016.

C- GRANDEZAS ESTATISTICAS UTILIZADAS

Para a avaliação das associações entre variáveis categóricas, foram aplicados o teste *qui-quadrado* (χ^2) e seu *p-valor*, com o objetivo de testar a hipótese de independência entre as variáveis. Utilizou-se este coeficiente V de *Cramer* (medida de força de associação entre duas variáveis) para mensurar a magnitude da associação, cujos valores variam de 0 (nenhuma associação) a 1 (associação perfeita).

O χ^2 foi calculado em código através da equação:

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^n \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (\text{A_C1})$$

Onde O_{ij} é o valor da linha i e da coluna j ; E_{ij} são as frequências esperadas sob a hipótese de independência. Somam-se todas as células da tabela de contingência.

O *p-valor* foi calculado como a probabilidade de obter um valor do qui-quadrado igual ou superior ao observado, sob a hipótese nula de independência, utilizando a distribuição qui-quadrado com os graus de liberdade apropriados.

$$p - \text{valor} = P(\chi_v^2 \geq \chi_{obs}^2) \quad (\text{A_C2})$$

A confiabilidade dos resultados foi reforçada por meio da acurácia⁴ dos modelos de classificação. E das métricas de desempenho calculadas para cada classe, como precisão (*precision*), revocação (*recall*) e F1-score, definidas respectivamente pelas equações:

$$\text{Precision} = \frac{TP}{TP + FN} \quad (\text{A_C3})$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (\text{A_C4})$$

$$F_score = 2 \times \frac{\text{Precision} \times \text{recall}}{\text{Precision} + \text{recall}} \quad (\text{A_C5})$$

Onde TP corresponde aos verdadeiros positivos; FP corresponde aos falsos positivos e FN corresponde ao falso negativo.

Também foram computadas as médias macro (média simples entre as classes) e ponderada (considerando o número de amostras por classe), para fornecer uma avaliação mais

⁴ Acurácia é a métrica de avaliação do modelo através da qual se calcula a soma dos verdadeiros positivos e negativos (TP+TN) dividida pela soma do verdadeiros positivos e negativos e falsos positivos e falsos negativos (FP+ FN)

$$\text{Acuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

robusta do desempenho global do modelo, visualizações gráficas por meio de *heatmaps* complementaram a análise estatística, permitindo identificar padrões relevantes nas distribuições conjuntas das variáveis e oferecendo uma compreensão mais abrangente das inter-relações presentes nos dados.

As equações para calcular a média máxima e ponderada da precisão, da sensibilidade e do *F1-score*, onde n_i é o número de amostras de classes i , são:

$$\text{Macro precision} = \frac{1}{N} \sum_{i=1}^N \text{Precision}_i \quad (\text{A_C6})$$

$$\text{Macro recall} = \frac{1}{N} \sum_{i=1}^N \text{recall}_i \quad (\text{A_C7})$$

$$\text{Macro F1 - score} = \frac{1}{N} \sum_{i=1}^N \text{F1}_i \quad (\text{A_C8})$$

$$\text{Weighted Precision} = \frac{\sum_{i=1}^N (n_i \times \text{Precision}_i)}{\sum_{i=1}^N n_i} \quad (\text{A_C9})$$

$$\text{Weighted recall} = \frac{\sum_{i=1}^N (n_i \times \text{Recall}_i)}{\sum_{i=1}^N n_i} \quad (\text{A_C10})$$