

Renato Barreira

**Em Busca do Algoritmo Bolsonarista:
Aprendizado de Máquina e Aprendizagem
Profunda aplicados à Ciência Política**

Brasil

2024

Renato Barreira

**Em Busca do Algoritmo Bolsonaroista:
Aprendizado de Máquina e Aprendizagem Profunda
aplicados à Ciência Política**

Dissertação apresentada ao Programa de Pós-
Graduação em Ciência Política da UNIRIO,
para obtenção do grau de Mestre.

Universidade Federal do Estado do Rio de Janeiro - UNIRIO

Centro de Ciências Jurídicas e Políticas - CCJP

Programa de Pós-Graduação em Ciência Política - PPGCP

Orientador: Steven Dutt-Ross

Brasil

2024

Renato Barreira

Em Busca do Algoritmo Bolsonarista:
Aprendizado de Máquina e Aprendizagem Profunda aplicados à Ciência Política/
Renato Barreira. – Brasil, 2024–
76p. : il. (algumas color.) ; 30 cm.

Orientador: Steven Dutt-Ross

Dissertação (Mestrado) – Universidade Federal do Estado do Rio de Janeiro - UNIRIO
Centro de Ciências Jurídicas e Políticas - CCJP
Programa de Pós-Graduação em Ciência Política - PPGCP, 2024.

1. Comportamento Político. 2. Aprendizado de Máquina. 3. Aprendizagem Profunda. 4. Bolsonarismo.

Renato Barreira

**Em Busca do Algoritmo Bolsonarista:
Aprendizado de Máquina e Aprendizagem Profunda
aplicados à Ciência Política**

Dissertação apresentada ao Programa de Pós-
Graduação em Ciência Política da UNIRIO,
para obtenção do grau de Mestre.

Trabalho aprovado. Brasil, Fevereiro de 2025:

Steven Dutt-Ross
Orientador

Vinicius Pinheiro Israel
Convidado 1

Carlos Eduardo Mello
Convidado 2

Brasil
2024

Este trabalho é dedicado.

Agradecimentos

Os agradecimentos principais são direcionados a todos que participaram direta ou indiretamente desse trabalho. Seja ao corpo docente do PPGCP-UNIRIO, que acolheu e acreditou no projeto, seja ao PPGI-UNIRIO, que acolheu um estudante de humanas.

Os agradecimentos especiais para Téo e Lisa Barreira; Juliana Porto; M.Sc. Nathália Ribeiro - por ter acreditado e dado oportunidade às minhas ideias; M.Sc. Carolini Silva; Defensora Pública do Estado do Rio de Janeiro Dra. Luciana Fernandez; alguns da minha turma de mestrado; e Wagner Reis - por, surpreendentemente, ter sugerido o tema que deu o pontapé inicial desse trabalho.

“Admito, com sinceridade que, ao formular minhas propostas, fui guiado por juízos de valor e por algumas predileções de ordem pessoal. Mas espero que as propostas se tornem aceitáveis para os que apreciam não só o rigor lógico, mas também a ausência de dogmatismos; para os que se importam com as aplicações práticas, mas se interessam ainda mais pelas aventuras da ciência, pelas descobertas que uma após outra, nos acareiam com novas e inesperadas perguntas, obrigando-nos a tentar encontrar respostas novas e insuspeitadas” (Karl Popper, A Lógica Pesquisa Científica)

Resumo

O objetivo principal desse trabalho é utilizar métodos computacionais para prever o comportamento característico da militância “bolsonarista”. A partir do uso de aprendizado de máquina, especificamente métodos de agrupamento e regressões linear e logística, e aprendizagem profunda, especificamente com rede neural *transformers*, busca-se achar o modelo ótimo para prever o “bolsonarismo” nos eleitores. Os resultados demonstram que é possível utilizar métodos de aprendizado de máquina e aprendizagem profunda para prever comportamento político permitindo uma análise mais aprofundada destes. Além disso os resultados indicam o potencial transformador que esses métodos possuem na ciência política computacional.

Palavras-chave: Comportamento Político, Aprendizado de Máquina, Aprendizagem Profunda, Bolsonarismo.

Abstract

The main objective of this study is to use computational methods to predict the characteristic behavior of “Bolsonarist” militancy. By employing machine learning methods, specifically clustering and linear and logistic regressions, as well as deep learning, particularly with transformer neural networks, the aim is to find the optimal model to predict “Bolsonarism” among voters. The results demonstrate that it is possible to use machine learning and deep learning methods to predict political behavior, enabling a more in-depth analysis of these dynamics. Furthermore, the findings highlight the transformative potential these methods hold for computational political science.

Keywords: Political Behavior, Machine Learning, Deep Learning, Bolsonaroism.

Resumen

El objetivo principal de este trabajo es utilizar métodos computacionales para prever el comportamiento característico de la militancia “bolsonarista”. A través del uso de aprendizaje automático, específicamente métodos de agrupamiento y regresiones lineales y logísticas, y aprendizaje profundo, particularmente con redes neuronales *transformers*, se busca encontrar el modelo óptimo para predecir el “bolsonarismo” en los votantes. Los resultados demuestran que es posible utilizar métodos de aprendizaje automático y aprendizaje profundo para predecir el comportamiento político, permitiendo un análisis más detallado de estos fenómenos. Además, los resultados destacan el potencial transformador que estos métodos tienen en la ciencia política computacional.

Palabras clave: Comportamiento Político, Aprendizaje Automático, Aprendizaje Profundo, Bolsonarismo.

Lista de ilustrações

Figura 1 – Clusterização Fuzzy C-Means Bolsonarismo	35
Figura 2 – Clusterização Fuzzy C-Means Antipetismo	36
Figura 3 – ROC - Curve - Bolsonarismo e Antipetismo	44
Figura 4 – ROC - Curve - Antipetismo	44
Figura 5 – ROC - Curve - Bolsonarismo	45
Figura 6 – Gráfico de Palavras-Chave	55
Figura 7 – Gráfico de bigramas chave	56
Figura 8 – Análise de Sentimentos	57
Figura 9 – Gráfico “wordfish” de Posição Relativa	59

Lista de tabelas

Tabela 1	– Resultados do Modelo OLS com Ponto de Corte de 0.75	37
Tabela 2	– Coeficientes do Modelo INTERV_MIL_CRIME	38
Tabela 3	– Coeficientes do Modelo INTERV_MIL_CORRUPCAO	38
Tabela 4	– Coeficientes do Modelo INTERV_MIL_INST_POL	38
Tabela 5	– Resumo das Métricas das Matrizes de Confusão	41
Tabela 6	– Desempenho de Validação Cruzada de Diferentes Modelos e Configurações	64
Tabela 7	– Parâmetros e Fluxo de Trabalho para Inferência com Bert-Vi	65
Tabela 8	– Tabela Verdade das Predições do Modelo	65

Sumário

1	INTRODUÇÃO	23
2	NAVALHA DE OCKHAM DO BOLSONARISMO	29
2.1	Bolsonarismo	31
2.2	Metodologia	32
2.3	Resultados	34
2.4	Conclusão	45
3	REDES NEURAIS E POLARIZAÇÃO POLÍTICA - MODELO BERT-VI	47
3.1	Metodologia	50
3.1.1	Banco de Dados	50
3.1.2	Análise Exploratória	52
3.2	BERT-Vi	60
3.3	Resultados	63
3.4	Conclusão	66
4	CONCLUSÕES E CONSIDERAÇÕES FINAIS	69
	REFERÊNCIAS	71

1 Introdução

A hipótese principal desse trabalho é que é possível utilizar aprendizado de máquina para descobrir e prever um determinado comportamento político. No caso estudo, o que será definido, nas próximas seções, como “bolsonarismo”.

Isso ocorre no contexto dos resultados das eleições presidenciais de 2018 e 2022 no Brasil, que evidenciaram uma profunda cisão ideológica entre os eleitores, demonstrando não apenas um alinhamento polarizado — centrado em líderes carismáticos — mas também uma firmeza de posicionamentos que se mantém ao longo do tempo. Esse fenômeno se insere em um contexto mais amplo de mudanças sociais e tecnológicas, em que as redes digitais atuam como palco privilegiado para a difusão de informações (ou desinformações) e para o acirramento de debates políticos. A rápida expansão do acesso à internet e o uso cada vez mais intenso de redes sociais transformaram a forma como os eleitores discutem política, formam opiniões e se mobilizam, reforçando dinâmicas de “bolhas ideológicas” e, consequentemente, acirrando conflitos e divisões.

Um dos trabalhos que evidenciam essa polarização é o artigo “Adesão ao Bolsonarismo: Dilemas, Contradições e Retórica em Argumentos de seus Militantes”. (2024). Nesse trabalho os autores analisam os discursos de 15 militantes pró-Bolsonaro para entender como eles justificam sua adesão ao movimento. Utilizando a psicologia social discursiva, os autores identificaram quatro principais argumentos: a defesa da liberdade individual, a moralidade cristã, a transparência e simplicidade de Jair Bolsonaro e o desejo de romper com o sistema político tradicional. Esses argumentos, entretanto, são marcados por contradições, como o apoio a medidas autoritárias em nome da liberdade, a defesa de valores cristãos em contraste com práticas éticas questionáveis e a promoção de rupturas com o sistema que ameaçam princípios democráticos.

O estudo também aponta os dilemas ideológicos presentes nos discursos, como as tensões entre tolerância e preconceito, liberdade e ordem, conservadorismo e mudança. Além disso, destaca as estratégias retóricas usadas para sustentar os argumentos como factuais, apresentando os militantes como racionais e preocupados com a verdade. A pesquisa contribui para entender as complexidades do “bolsonarismo”, revelando como seus apoiadores articulam valores aparentemente contraditórios para justificar sua adesão ao movimento.

Contudo, é importante destacar o trabalho “Um Brasil Dividido? Reflexões sobre a Polarização Política e Social no Brasil Contemporâneo” (2024). Nele, embora a autora analise a polarização política no Brasil atual, destacando a divisão entre os polos identificados com Lula e Bolsonaro, a análise sugere que o fenômeno Bolsonaro é um

sintoma que é um aprofundamento, e não a origem, da polarização contemporânea. Ela chega a essa conclusão a partir da investigação de cinco processos socio-históricos desde a fundação da Nova República que influenciam essa polarização: reformas no sistema político institucional, mudanças no contexto econômico nacional, transformações estruturais na sociedade brasileira, evolução da mídia tradicional e novas tecnologias de informação, e alterações no contexto político internacional.

Em meio a essa conjuntura, a necessidade de métodos e análises sobre o comportamento do eleitorado brasileiro se mostra urgente. A polarização não se manifesta apenas em termos de adesão a partidos ou figuras públicas, mas também em como certos grupos interpretam valores democráticos, defendem pautas específicas e reagem a narrativas de ódio ou de exclusão. Dessa forma, a análise de grandes volumes de dados políticos — como pesquisas eleitorais, interações em redes sociais e indicadores socioeconômicos — surge como ferramenta essencial para investigar padrões de comportamento, motivações, perfis de voto e, sobretudo, para compreender como e por que determinados grupos se mantêm firmes em suas posições políticas.

Nesse sentido, o atual cenário político brasileiro oferece uma oportunidade singular para o desenvolvimento de estudos que combinem ciência política, análise de dados e técnicas avançadas de aprendizado de máquina. A complexidade das interações sociais e políticas na era digital exige abordagens que permitam capturar não apenas as preferências declaradas dos eleitores, mas também a dinâmica de como essas preferências se formam, evoluem e se cristalizam ao longo do tempo. Por meio de metodologias interdisciplinares, torna-se possível identificar fatores que contribuem para a consolidação de narrativas polarizadoras, mapear mecanismos de influência e prever tendências futuras — contribuindo, assim, para um debate mais qualificado sobre o fortalecimento da democracia e a superação de divisões tão profundas na sociedade.

Sletge *et al.* observam que os cientistas políticos têm adotado o aprendizado de máquina como método empírico, com um aumento significativo no uso nos últimos 10 anos. Os autores identificaram que a aplicação do aprendizado de máquina é mais comum nas áreas de comunicação política e estudos de conflito e paz, sendo os métodos mais utilizados a modelagem de tópicos (*Topic Modeling*), a máquina de vetores de suporte (SVM - *Supported Vector Machines*) e a floresta aleatória (*Random Forest*). Além disso, os autores esboçam uma crítica no sentido de que a transparência no relato da otimização dos modelos de aprendizado de máquina por meio do ajuste de hiperparâmetros poderia ser aprimorada, e que os pesquisadores deveriam realizar suas próprias comparações ao selecionar o modelo mais adequado (SLEGTE; DROOGENBROECK; GINIS, 2024).

Portanto, esse trabalho tem como objetivo secundário explorar a multidisciplinaridade entre Ciência de Dados e Ciência Política de maneira a entender como o desenvolvimento de sistemas de informação impacta a qualidade e a abrangência das soluções

criadas em análise de comportamento político nesta disciplina.

Para contextualização, esse trabalho toma para si os seguintes conceitos de aprendizado de máquina. O aprendizado supervisionado (SML) é o paradigma mais prevalente, no qual o objetivo é treinar um modelo a partir de um conjunto de entradas e rótulos, usado para fazer previsões, e onde o desempenho é medido em dados fora da amostra. Aqui, é importante destacar a distinção entre predição no sentido formal e previsão. Enquanto a predição estima o resultado com base em valores de características fora da amostra, a previsão tenta prever eventos futuros (JORDAN M. I., 2015). Problemas de classificação tratam de dados discretos, enquanto problemas de regressão tratam de dados contínuos.

O aprendizado de máquina não supervisionado (UML) tem como objetivo desvendar a estrutura subjacente em dados não rotulados (JORDAN M. I., 2015): uma representação dos dados de entrada que seja mais compreensível do que os próprios dados. Nesse contexto, as principais categorias são as técnicas de redução de dimensionalidade e de agrupamento (*clustering*). Há ainda o aprendizado semi-supervisionado (Semi-SML) que pode ser considerado uma combinação de aprendizado supervisionado e não supervisionado, onde dados não rotulados ou informações auxiliares são utilizados para complementar dados rotulados em um contexto de aprendizado supervisionado.

Embora o aprendizado de máquina seja eficaz para identificar associações entre variáveis, autores como Breiman (BREIMAN, 2001) destacam que esses modelos não permitem derivar relações causais de forma direta. Isso levanta a questão de uma possível tensão entre esses métodos e a preferência da ciência política por inferências causais baseadas em análises empíricas. Contudo, esse trabalho se posiciona e toma para si as tendências recentes na causalidade moderna no aprendizado de máquina. De modo que busca-se - além da análise de variáveis - o objetivo de descoberta causal e estimativa de efeito causal no estudo do comportamento político (NOGUEIRA AR., 2022).

Nesse contexto, a Teoria da Representação (*Representation Theory*) fornece um arcabouço para modelar a complexidade das interações sociais e políticas na era digital. Ela se concentra em como informações podem ser estruturadas e processadas para capturar e representar propriedades fundamentais de sistemas complexos, como comportamentos políticos, dentro de um contexto matemático ou algorítmico.

Como discutido por Plagnol(PLAGNOL, 2023), enfatiza-se a análise das representações e suas propriedades lógicas. Isso é útil na análise dos mecanismos de representação política, essencial para compreender como diferentes ideologias, como o “bolsonarismo”, são representadas e interpretadas nos dados.

Portanto, a Teoria da Representação permite entender como padrões complexos de dados — como os encontrados em estudos eleitorais — podem ser mapeados para estruturas matemáticas mais simples e manipuláveis, sem perder as características essenciais

do sistema original. Um exemplo disso é o trabalho de Goppa(GOPPA, 1995). O autor explora a aplicação de teoria da representação em grupos simétricos e sua incorporação com a teoria da informação, o que pode ter implicações para entender dinâmicas de grupos no contexto da análise do comportamento político.

Teoria da Representação é utilizada também no artigo “*Making sense to decreasing citizen eParticipation through a social representation lens*” de Pozzebon, Cunha e Coelho(POZZEBON; CUNHA; COELHO, 2016). Nele, os autores utilizam a lente das representações sociais para compreender a diminuição da “eParticipação” dos cidadãos. O estudo investiga como diferentes grupos interpretam e reagem a “eParticipação”, revelando que as representações sociais influenciam significativamente as motivações e comportamentos dos cidadãos em relação à participação eletrônica. Essa perspectiva de análise é relevante para estudos de teorias de representação em sistemas informacionais, já que busca explicar como percepções subjetivas e coletivas moldam a interação dos cidadãos com plataformas de “eGovernança”.

Portanto, o objetivo geral desse trabalho é explorar a multidisciplinaridade entre Ciência de Dados e Ciência Política de maneira a entender como o desenvolvimento de sistemas de informação impacta a qualidade e a abrangência das soluções criadas em análise de comportamento político nesta disciplina.

Para isso o trabalho realiza dois experimentos. No primeiro capítulo “Navalha de Ockham do Bolsonarismo”, onde se busca identificar e analisar, a partir do banco de dados ESEB-2022, os grupos “bolsonarista” e “antipetista” entre os eleitores brasileiros. Para isso, utiliza-se o algoritmo de clusterização *Fuzzy C-Means* (FCM), que atribui graus de pertencimento aos indivíduos conforme suas respostas sobre preferência política e ideologia. Em seguida, aplicam-se modelos de regressão (linear e *Fuzzy Regression Discontinuity*) para avaliar o efeito do “antipetismo” sobre o “bolsonarismo”, focando em eleitores de alto grau de afinidade com Bolsonaro. Além disso, são analisadas variáveis relacionadas ao apoio a intervenções militares. Para verificar como “bolsonarismo” e “antipetismo” se associam ao apoio a um golpe militar, desenvolvem-se três modelos de regressão logística, cada um voltado para um tipo específico de “intervenção” (crimes, corrupção e problemas institucionais). Por fim, o efeito combinado de “bolsonarismo” e “antipetismo” também é explorado por meio de um termo de interação no modelo de regressão.

No segundo capítulo, “Redes Neuras e Polarização Política”, trata da aplicação de redes neurais (NNs), sobretudo modelos de aprendizagem profunda (*deep learning*) como BERT e XLM-RoBERTa, para analisar a polarização política durante a pandemia de 2020. Primeiramente, é feita uma análise exploratória usando os pacotes quanteda (BENOIT et al., 2018) e (JOCKERS, 2015) do R (R Core Team, 2019). A partir dessa análise exploratória, desenvolve-se uma rede neural *transformers* chamada BERT-Vi: um modelo de classificação para entrevistas de profundidade, focando especialmente na

avaliação do governo Bolsonaro (2019–2022). A principal justificativa é que métodos tradicionais têm limitações ao lidar com dados não estruturados e padrões complexos típicos do comportamento político.

2 Navalha de Ockham do Bolsonarismo

Conforme exposto, os resultados das eleições presidenciais de 2018 e 2022 no Brasil revelaram um eleitorado polarizado e rígido. Esse cenário oferece uma oportunidade singular para desvendar padrões e tendências que moldam essa polarização já que a crescente complexidade das interações sociais e políticas na era digital exige abordagens inovadoras para compreender e prever comportamentos políticos.

Utilizar aprendizado de máquina para analisar e compreender o comportamento político dos eleitores brasileiros, buscando capturar relações de necessidade e suficiência em padrões e resultados específicos. Nesse sentido, a Teoria da Representação é fundamental para entender como diferentes tipos de informações e dados podem ser transformados e interpretados dentro de um sistema. A análise de comportamentos políticos complexos por meio de ciência de dados se beneficia dessa teoria ao possibilitar a representação adequada de dados políticos e sociais, permitindo a identificação e análise de padrões.

Este capítulo visa suprir uma lacuna crítica na intersecção entre Ciência Política e Ciência de Dados, oferecendo uma abordagem inovadora para compreender o comportamento político por meio de algoritmos de aprendizado de máquina. A pesquisa não apenas avançará o desenvolvimento teórico e empírico da Ciência Política, mas também aprimorará o rigor metodológico ao integrar técnicas avançadas de análise de dados. Ao explorar grandes volumes de dados políticos, este estudo tem o potencial de fornecer novos *insights* e ferramentas para entender as dinâmicas políticas contemporâneas.

Desta maneira, busca-se a “navalha de ockham do bolsonarismo”. Isto é, Navalha de Ockham, também conhecida como princípio da parcimônia ou princípio da economia, constitui um importante recurso heurístico oriundo da escolástica, atribuído a Guilherme de Ockham (1288–1347). De acordo com esse princípio, entre múltiplas explicações igualmente adequadas para descrever um fenômeno, deve-se escolher aquela que apresente o menor número de variáveis e hipóteses correlacionadas. Na prática, “cortam-se” as suposições desnecessárias a fim de manter apenas o essencial para explicar o fato observado.

Enquanto parte integrante do método científico e da filosofia da ciência, a Navalha de Ockham não representa o único critério para avaliar teorias. Entretanto, apresenta vantagens notáveis, pois hipóteses mais simples facilitam a verificação empírica e a replicação dos resultados. A aplicação moderna desse ideal de simplicidade pode ser observada no âmbito matemático e no aprendizado de máquina.

Na área de aprendizado de máquina, a Navalha de Ockham — ou princípio da parcimônia — é frequentemente associada à adoção de modelos mais simples e menos propensos a sobreajuste (*overfitting*) (DHERIN; MUNN; BARRETT, 2021; MINGARD et

al., 2023; ANTAL et al., 2024; STERKENBURG, 2024). Em termos práticos, a ideia de “cortar” supostos desnecessários converge para a busca de modelos com número reduzido de parâmetros ou variáveis, sem perder a capacidade de explicar os dados. Dessa forma, reduz-se a complexidade ao mesmo tempo em que se preserva a capacidade preditiva, tornando os algoritmos mais interpretáveis e menos suscetíveis a ruídos ou variações não essenciais nos dados.

Na teoria de representação de sistemas informacionais, a Navalha de Ockham — ou princípio da parcimônia — exerce papel fundamental ao orientar a concepção de modelos e estruturas que sejam, simultaneamente, eficientes e suficientes. Em outras palavras, busca-se representar o conjunto de informações de modo a captar suas características essenciais, sem incorrer em redundâncias desnecessárias ou complexidades que dificultem a manutenção e escalabilidade do sistema.

A aplicação de técnicas de aprendizado de máquina, guiadas por estruturas representacionais, pode preencher lacunas teóricas ao identificar e explorar padrões complexos. Por exemplo, os movimentos de “antipetismo” e “bolsonarismo” podem ser representados como subespaços ou componentes de uma estrutura maior, capturando relações independentes e interdependentes entre eles.

Nesse contexto, o objetivo específico desse experimento é entender o que torna uma pessoa “bolsonarista”. Ou seja, a partir do banco de dados [ESEB-2022 ESTUDO ELEITORAL BRASILEIRO](https://www.cesop.unicamp.br/por/banco_de_dados/v/4680)¹, será feito um estudo exploratório do uso de aprendizado estatístico aplicados à análise de comportamento político a fim de verificar se há conexões explícitas entre os marcos teóricos do cânone de Ciência Política sobre “bolsonarismo” e os dados disponíveis sobre o eleitor brasileiro.

Ao testar o modelo canônico de “bolsonarismo”, a teoria de representação pode ajudar a formalizar e verificar as relações entre características dos eleitores e comportamentos preditos pelos modelos, fornecendo uma estrutura sólida para validar inferências.

Com base no referido banco de dados, este trabalho empregará o método *Fuzzy C-Means (FCM)* para, além de redimensionar os dados, identificar o agrupamento associado ao “bolsonarismo”, atribuindo a ele graus de pertencimento. A partir da variável *proxy* gerada, serão analisados os efeitos do “antipetismo” (também derivado pelo FCM) e das variáveis relacionadas ao apoio a um golpe militar sobre o grau de pertencimento ao “bolsonarismo”. Com isso objetiva-se explorar a pesquisa empírica em Ciência Política além dos métodos lineares e de covariâncias clássicos em prol de uma abordagem mais multidisciplinar e contemporânea.

É exatamente nesse contexto que pesquisas interdisciplinares em sistemas de informação apresentam-se como fundamentais para promover descobertas e desenvolver

¹ https://www.cesop.unicamp.br/por/banco_de_dados/v/4680

modelos orientados por dados para compreender o comportamento desse eleitor. Desta maneira é possível reunir abordagens, técnicas e ferramentas de diversas disciplinas, como Mineração de Dados, Aprendizado de Máquina, Estatística, Visualização, Aprendizagem Profunda, entre outras. Quando integradas ao conhecimento especializado de cada área da literatura canônica de Ciência Política, essas ferramentas abrem oportunidades para a criação de novas metodologias, modelagens e aplicações em pesquisas inovadoras no campo da busca por algoritmos de comportamento político.

2.1 Bolsonarismo

O principal referencial teórico a ser testado na análise de dados é o cânone de Ciência Política sobre “bolsonarismo”. O estudo de Oswaldo E. do Amaral ([AMARAL, 2020](#)) investiga os determinantes dos votos para Jair Bolsonaro nas eleições presidenciais de 2018 no Brasil. Essas eleições surpreenderam analistas ao quebrar a lógica competitiva entre o PT e o PSDB que prevalecia desde 1994.

Para essa análise, Amaral utilizou

- **Análise de Regressão Logística Multinomial:** utilizada para analisar os votos no primeiro turno da eleição presidencial de 2018. Comparou eleitores que votaram em Ciro Gomes, Jair Bolsonaro e Fernando Haddad, com este último grupo servindo como categoria de referência;
- **Análise de Regressão Logística Binomial:** Aplicada para analisar os votos no segundo turno. Comparou eleitores que votaram em Jair Bolsonaro com os que votaram em Fernando Haddad, novamente com Haddad servindo como categoria de referência;
- **Testes Multivariados:** Foram realizados três testes diferentes para compreender melhor o impacto das variáveis selecionadas. Primeiro Teste: Incluiu apenas as variáveis de avaliação retrospectiva. Segundo Teste: Incluiu apenas as variáveis de identificação política. Terceiro Teste (Modelo Geral): Incluiu tanto as variáveis de avaliação retrospectiva quanto as de identificação política, além das variáveis sociodemográficas e contextuais.

Os resultados mostraram que os modelos de identificação política apresentaram resultados mais robustos do que os modelos de avaliação retrospectiva, indicando que as variáveis de identificação política tiveram um impacto mais forte na explicação dos votos na eleição presidencial de 2018.

Entre as variáveis, vale destacar o “antipetismo”. O aumento significativo do “antipetismo”, tornando-se um fator crucial na vitória de Bolsonaro. Resultando num impacto de que os eleitores que rejeitavam o PT tinham dez vezes mais chances de votar

em Bolsonaro. Além disso, ser cristão pentecostal aumentou a probabilidade de votar em Bolsonaro, destacando seu alinhamento com líderes pentecostais em questões morais (AMARAL, 2020, p.10).

Já no trabalho de Renno (RENNÓ, 2020), os resultados indicam que o apoio a Bolsonaro foi influenciado por uma combinação de ressentimento contra o PT, uma perspectiva de reação cultural, visões estritas sobre lei e ordem, liberalismo econômico e rejeição de políticas sociais. A análise distingue entre eleitores consistentemente apoiadores de Bolsonaro e outros eleitores, destacando as variações na intensidade do apoio ao longo da campanha.

Entre os fatores apontados por Renno, se destacam a) Retrocesso Cultural: o apoio a Bolsonaro foi em parte um reflexo de uma reação contra as mudanças progressistas de valores sociais, como direitos LGBTQ+ e políticas de inclusão social; b) Lei e Ordem: os eleitores de Bolsonaro mostraram um forte apoio a políticas severas de combate ao crime, como a redução da maioria penal e a pena de morte; c) Economia: os eleitores de Bolsonaro tendem a apoiar o liberalismo econômico, favorecendo privatizações e reduzindo a intervenção estatal na economia; d) Rejeição ao PT: um fator significativo foi o “antipetismo”, ou seja, o ressentimento contra o PT, que foi exacerbado por escândalos de corrupção e a percepção de incompetência administrativa; e) Políticas Sociais: houve uma oposição a políticas sociais específicas, especialmente as ações afirmativas e as cotas raciais.

Diferente de Renno e Amaral, este trabalho utilizará uma variável *proxy* redimensionada (explicada na próxima seção) como variável alvo. Além disso, diferentemente de uma análise baseada em efeitos das variáveis, busca-se analisar qual a variável, entre “antipetismo” e “apoio a intervenção militar” tem melhor capacidade preditiva para essa variável alvo.

2.2 Metodologia

A partir da base de dados ESEB-2022 ESTUDO ELEITORAL BRASILEIRO, usa-se o *Fuzzy C-Means* para clusterização a fim de separar a classe “Bolsonarista” e “Antipetista” das demais classes, atribuindo um grau de pertencimento a essas duas classes para a análise posterior. A classe “bolsonarista” será obtida a partir do redimensionamento das variáveis discretas: gosta do Bolsonaro² (0 não gosta e 10 gosta) e ideologia³ (0 esquerda e 10 direita). Portanto, quanto mais de direita e mais gostar do Bolsonaro, para fins dessa análise, é mais “bolsonarista”. Desta maneira busca-se não só redimensionar o dado, mas testar também se há convergência dos entrevistados nesse ponto.

Da mesma maneira, a classe “antipetista” será obtida a partir do redimensionamento

² Referência no questionário: Q17-2 https://www.cesop.unicamp.br/por/banco_de_dados/v/4680

³ Referência no questionário: Q18

das variáveis discretas: gosta do Lula⁴ (0 não gosta e 10 gosta) e gosta do Partido dos Trabalhadores (PT)⁵ (0 não gosta e 10 gosta). Para, mais uma vez, não só redimensionar os dados, mas para testar a convergência dos entrevistados em torno dos conceitos chaves para essa análise.

O *Fuzzy C-Means* (FCM) é um algoritmo de agrupamento (*clustering*) baseado em lógica fuzzy, proposto inicialmente por Dunn em 1973 (DUNN, 1973) e aprimorado por Bezdek em 1981 (BEZDEK, 1981). Ele realiza uma partição suave dos dados, onde cada elemento pode pertencer a múltiplos grupos com diferentes graus de pertencimento.

O algoritmo atribui níveis de associação a cada ponto de dados com base na sua distância para os centros dos clusters, facilitando soluções de agrupamento flexíveis. Uma melhoria notável é o algoritmo *Membership Scaling Fuzzy C-Means* (MSFCM), que acelera a convergência escalando os graus de associação com base na proximidade do centro do *cluster*, mantendo alta qualidade de agrupamento (ZHOU et al., 2021).

Depois desse redimensionamento, serão feitas uma regressão linear simples (RL) e uma regressão descontinuada (*Fuzzy Regression Discontinuity* - FRD) para analisar os efeitos do “antipetismo” sobre o “bolsonarismo”. A RL mostrará o efeito da primeira sobre a segunda e a FRD mostrará o efeito após o corte de tratamento. O corte é necessário porquanto o objeto dessa análise são os eleitores com alto grau de pertencimento ao grupamento “bolsonarista”.

A *Fuzzy Regression Discontinuity* (FRD) é uma variação do método de *Regression Discontinuity* (RD), uma abordagem de avaliação de impacto utilizada em estudos econômicos para identificar efeitos causais. No contexto do RD tradicional, a atribuição a um tratamento é determinada por uma regra de corte clara em uma variável de continuidade. No entanto, no FRD, a probabilidade de receber o tratamento não é perfeitamente determinística na regra de corte, resultando em uma descontinuidade na probabilidade de tratamento em vez de uma descontinuidade no tratamento em si (HAHN; TODD; KLAUW, 2001; IMBENS; LEMIEUX, 2008; LEE; LEMIEUX, 2010).

Os conceitos principais envolvem a descontinuidade na probabilidade de tratamento. No FRD, a atribuição ao tratamento é parcialmente determinada pela regra de corte, criando uma descontinuidade na probabilidade de tratamento na fronteira dessa regra. Além disso, a variável indicadora da posição em relação ao ponto de corte atua como uma variável instrumental para o tratamento, permitindo a estimativa do efeito causal mesmo quando a adesão ao tratamento é imperfeita. Assim como no RD clássico, a identificação dos efeitos causais no FRD baseia-se na comparação de unidades próximas ao ponto de corte, tomando para si que essas unidades são comparáveis, exceto pela descontinuidade na probabilidade de receber o tratamento.

⁴ Referência no questionário: Q17-5

⁵ Referência no questionário: Q16-5

O uso de *Fuzzy C-Means*, FRD e Regressão Logística vai ao encontro da Teoria da Representação porquanto ela reflete uma abordagem para traduzir dados comportamentais em representações matemáticas que podem ser exploradas de forma mais eficiente. De maneira a oferecer uma base teórica para entender como esses métodos preservam (ou transformam) as propriedades fundamentais dos dados.

Em seguida são analisadas as variáveis relacionadas ao apoio de golpe militar (“intervenção militar”). Analisa-se como o apoio a essas “intervensões” está relacionada com as variáveis de grau de pertencimento ao agrupamento “bolsonarista” e ao agrupamento “antipetista”. Para enriquecimento da análise, utiliza-se o “antipetismo” como variável alvo também. Três modelos de regressão logística para avaliar a influência das referidas variáveis e sua interação sobre a probabilidade de ser a favor de uma intervenção militar em diferentes contextos: crimes⁶, corrupção⁷ e problemas institucionais⁸.

Cada modelo de regressão logística foi ajustado a uma única variável-alvo binária (apoia ou não a “intervenção militar”) no lugar de modelo de regressão logística multinomial. Assim busca-se analisar cada variável isoladamente.

Para testar o efeito combinado do “bolsonarismo” e “antipetismo” em um modelo de regressão logística, foi incluído um termo de interação no modelo. O termo de interação permitirá analisar se a combinação dessas duas variáveis tem um efeito adicional nas probabilidades de ser a favor de uma intervenção militar além das variáveis isoladamente. Ao interpretar os coeficientes do modelo com interação, além dos coeficientes individuais de “bolsonarismo” e “antipetismo”, foi adicionado um coeficiente para o termo de interação. Esse coeficiente indicará o efeito adicional quando ambas as variáveis estão presentes juntas. Os resultados finais serão discutidos na próxima seção.

2.3 Resultados

Para a clusterização da classe “bolsonarismo”, foram utilizadas as variáveis discretas: gosta do Bolsonaro (0 não gosta e 10 gosta) e ideologia (0 esquerda e 10 direita). O algoritmo do *Fuzzy C-Means* foi configurado para identificar 3 agrupamentos (*clusters*) a fim de melhor representar a polarização política atual. Como é possível ver na Figura 1, os três agrupamentos foram identificados e a variável alvo desse trabalho é o grau de pertencimento ao agrupamento 2. O tamanho é proporcional ao grau de pertencimento a este e é calculado a partir do centroide. Contudo, é importante destacar um achado pertinente nessa análise: a existência de uma “direita não bolsonarista” (agrupamento 1) mais numerosa que a “direita bolsonarista” alvo. O agrupamento 3 é um grupo difuso com todos que não pertencem

⁶ Ref. no questionário: Q30a

⁷ Ref. no questionário: Q30b

⁸ Ref. no questionário: Q30c

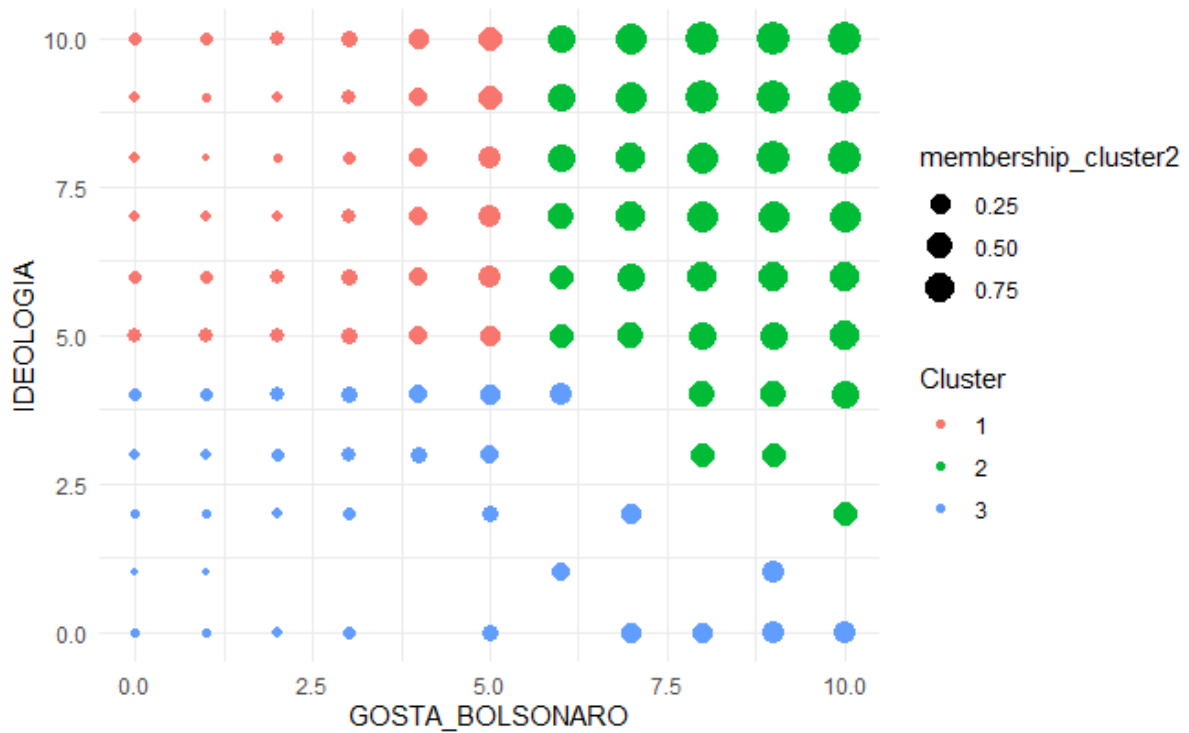


Figura 1 – Clusterização Fuzzy C-Means Bolsonarismo

Fonte: Elaboração própria.

nem a “direita não bolsonarista” e nem a “direita bolsonarista”, por essa característica, a análise desse agrupamento exigiria um artigo próprio, o que foge do escopo desse trabalho.

A teoria da representação fornece a base para estruturar fenômenos complexos, como a afiliação ideológica, em um espaço matemático. No caso do *Fuzzy C-Means*, a capacidade de atribuir graus de pertencimento a diferentes agrupamentos reflete a ideia central da representação: traduzir ambiguidades e transições graduais, típicas do comportamento humano, em um sistema formal. Essa abordagem nuançada permite representar o “bolsonarismo” e o “antipetismo” como não mutuamente exclusivos, capturando a sobreposição e a interação entre esses fenômenos.

Para o agrupamento do “antipetismo” (*cluster 1*), foram utilizadas as variáveis discretas: gosta do Lula (0 não gosta e 10 gosta) e gosta do PT (0 não gosta e 10 gosta). Como é possível ver na Figura 2, o agrupamento 1 é considerado o “antipetismo” e o tamanho é proporcional ao grau de pertencimento a esse agrupamento. Conforme esperado pela literatura canônica de comportamento político, aqui temos uma relação linear bem estabelecida e clara.

O método Fuzzy C-Means foi usado não só como uma redução de dimensionalidade, mas também como método para testar a proposição teórica atribuindo um grau de pertencimento aos conceitos de “bolsonarismo” e “antipetismo”.

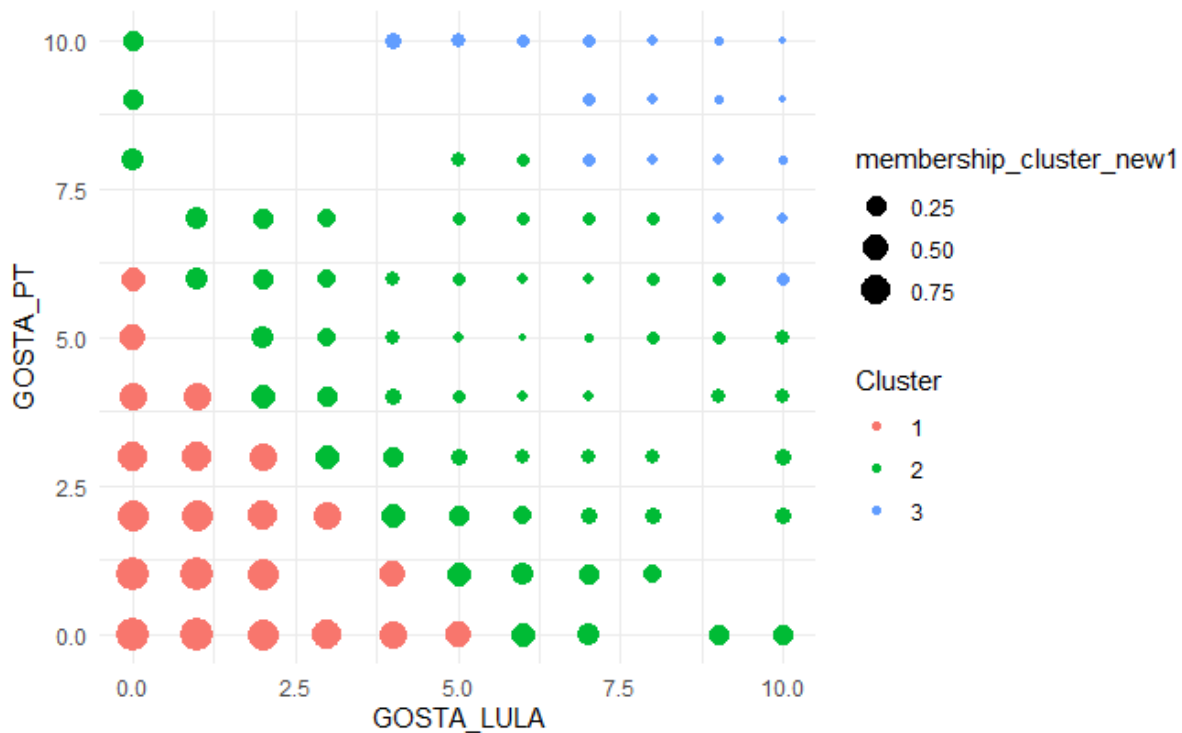


Figura 2 – Clusterização Fuzzy C-Means Antipetismo

Fonte: Elaboração própria.

Como a análise busca entender o efeito do “antipetismo” não só em todos os graus de “bolsonarismo”, mas para entender como especificamente o “antipetismo” (em sua alta pertinência) afeta o “bolsonarismo” (em sua alta pertinência) a variável de tratamento foi definida como “antipetismo > 0.75”, garantindo que a análise é feita naqueles que são realmente “antipetistas”. Portanto, além da regressão linear, foi feita também uma regressão descontinuada (FRD) para verificar se de fato o alto grau de “antipetismo” está associado ao alto grau de “bolsonarismo”.

Conforme a tabela 1: a) intercepto (Constante): 0.1382. Este coeficiente é altamente significativo com um p-valor menor que 0.0001. Isso significa que o valor esperado de “bolsonarismo” é 0.1382 quando as variáveis “antipetismo” e tratamento são zero; b) “antipetismo”: 0.7587. Este coeficiente também é altamente significativo, com um p-valor menor que 0.0001. Cada aumento de uma unidade em “antipetismo” está associado a um aumento de 0.7587 unidades em bolsonarismo, mantendo constante a variável tratamento; c) “treatment”: -0.1100. Este coeficiente é marginalmente significativo, com um p-valor de 0.072. Isso indica que há uma leve redução de 0.1100 unidades em “bolsonarismo” para indivíduos logo acima do ponto de corte (“antipetismo > 0.75”) comparado com aqueles logo abaixo.

A *Fuzzy Regression Discontinuity* utiliza conceitos de representação para lidar com cenários onde a atribuição de tratamento não é clara. Ao explorar causalidades em situações

Tabela 1 – Resultados do Modelo OLS com Ponto de Corte de 0.75

	Bolsonarismo
Intercepto	0.1382*** (0.011)
Antipetismo	0.7587*** (0.065)
Tratamento (Antipetismo > 0.75)	-0.1100 (0.061)
Observações	1,504
R ²	0.506

Fonte: Elaboração própria

complexas, essa técnica encapsula relações que não podem ser claramente definidas em fronteiras rígidas, mas que ainda podem ser representadas matematicamente. Isso reflete um aspecto central da teoria da representação: estruturar e analisar relações onde o rigor absoluto não é aplicável.

A regressão linear simples demonstra uma relação positiva entre “antipetismo” e “bolsonarismo evidente”. O coeficiente positivo e altamente significativo de “antipetismo” indica uma forte relação positiva de maneira que o aumento no “antipetismo” está associado ao aumento no “bolsonarismo”.

Contudo, a descontinuidade marginalmente significativa no ponto de corte sugere que a variável tratamento tem um coeficiente negativo e marginalmente significativo. Isso implica que, ao cruzar o ponto de corte de 0.75 para “antipetismo”, há uma leve redução no “bolsonarismo”. Este efeito é menor do que o observado com o ponto de corte de 0.5.

O impacto da descontinuidade no “bolsonarismo” é reduzido, sugerindo que a influência direta do “antipetismo” sobre o “bolsonarismo” é mais uniforme em níveis mais altos de “antipetismo”.

Em conclusão, os resultados desta análise indicam uma relação positiva e significativa entre “antipetismo” e “bolsonarismo”. Indivíduos com maior “antipetismo” tendem a exibir maiores níveis de “bolsonarismo”. No entanto, a descontinuidade específica observada em torno do ponto de corte de 0.75 é menos pronunciada, sugerindo que a variação em “bolsonarismo” é menos sensível a mudanças em “antipetismo” em níveis mais altos. Esse resultado é consistente com o agrupamento de dados na Figura 1, que indicou a existência de uma “direita não-bolsonarista” mais numerosa que a “direita bolsonarista” fazendo com que a relação após o corte (“antipetismo” > 0.75) seja negativa.

Portanto, uma hipótese a ser levantada, que será testada a seguir, é que, embora o “antipetismo” seja necessário para o “bolsonarismo”, ele não é suficiente. Sugerindo que, para que um “antipetista” seja “bolsonarista”, é necessário analisar outras variáveis, no caso, as relacionadas ao conservadorismo político.

Para isso, serão analisadas as variáveis relacionadas ao apoio de golpe militar (“intervenção militar”). Analisa-se se o apoio a essas “intervensões” está relacionada com as variáveis de grau de pertencimento aos agrupamentos “bolsonarista” e “antipetista”. Nessa parte, o “antipetismo” é utilizado também como variável alvo.

Cada modelo de regressão logística foi ajustado a uma única variável-alvo binária (apoia ou não a “intervenção militar”) para cada uma das três questões referentes a intervenção militar: a) em caso de corrupção, b) em caso de crime, c) em caso de instabilidade política; no lugar de modelo de regressão logística multinomial. Assim busca-se analisar cada variável isoladamente. Ao interpretar os coeficientes do modelo com interação, além dos coeficientes individuais de “bolsonarismo” e “antipetismo”, esse coeficiente indicará o efeito adicional quando ambas as variáveis estão presentes juntas.

Tabela 2 – Coeficientes do Modelo INTERV_MIL_CRIME

	Coeficientes
BOLSONARISMO	0.6706568
ANTIPETISMO	0.3791844
BOLSONARISMO:ANTIPETISMO	0.7961020
Intercepto	-0.9520012

Fonte: Elaboração própria

Tabela 3 – Coeficientes do Modelo INTERV_MIL_CORRUPCAO

	Coeficientes
BOLSONARISMO	1.6041268
ANTIPETISMO	0.6106723
BOLSONARISMO:ANTIPETISMO	0.4997436
Intercepto	-0.5861708

Fonte: Elaboração própria

Tabela 4 – Coeficientes do Modelo INTERV_MIL_INST_POL

	Coeficientes
BOLSONARISMO	1.6788182
ANTIPETISMO	0.2756693
BOLSONARISMO:ANTIPETISMO	0.1453656
Intercepto	-1.001

Fonte: Elaboração própria

A regressão logística com termos de interação modela como múltiplos fatores se combinam para formar padrões eleitorais. Aqui, a teoria da representação é referenciada para traduzir interações complexas entre o “antipetismo”, o “bolsonarismo” e o apoio a intervenções militares em termos matemáticos que capturam a essência dessas interações sem simplificá-las excessivamente.

O coeficiente *log odds* (ou logaritmo das chances) é definido como o logaritmo natural das chances de um evento ocorrer. As chances (*odds*) são definidas como a razão entre a probabilidade de um evento ocorrer (p) e a probabilidade de ele não ocorrer ($1 - p$):

$$Odds = \frac{p}{1 - p}$$

O *log odds* é então dado por:

$$LogOdds = \ln \left(\frac{p}{1 - p} \right)$$

Os logaritmos das chances (*log odds*) são amplamente utilizados no contexto da regressão logística devido à sua capacidade de estabelecer uma relação linear entre os preditores e os logaritmos das chances da variável dependente. Essa abordagem é adequada para a modelagem de resultados binários, uma vez que transforma as probabilidades, originalmente não lineares, em uma escala linear modelada.

Além disso, essa transformação possui outra característica: ela assegura que as probabilidades previstas pelo modelo permaneçam dentro do intervalo $[0,1]$. Isso é fundamental, pois probabilidades fora desse intervalo não são interpretáveis no contexto de eventos binários. Dessa forma, o uso dos logaritmos das chances não apenas facilita a interpretação dos coeficientes estimados, mas também preserva a coerência estatística do modelo (AGRESTI, 2013; HOSMER; LEMESHOW; STURDIVANT, 2013).

No modelo que avalia o apoio à intervenção militar em casos de criminalidade muito alta, o “bolsonarismo” apresentou um coeficiente logístico de 0,67, correspondente a uma razão de chances (*odds ratio*) de aproximadamente 1,96 ($\exp(0,67) \approx 1,96$). Isso indica que, para cada ponto adicional na escala de “bolsonarismo”, as chances de um indivíduo apoiar a intervenção aumentam em cerca de 96%, mantendo constante o nível de “antipetismo”. Este, por sua vez, teve um coeficiente de 0,37 ($OR \approx 1,46$), sugerindo um aumento de 46% nas chances de apoio à intervenção para cada unidade adicional.

O termo de interação entre “bolsonarismo” e “antipetismo” foi positivo (0,79), com uma *odds ratio* de aproximadamente 2,22 ($\exp(0,79)$). Esse efeito conjunto é mais do que aditivo, sugerindo que indivíduos que combinam altos níveis nas duas variáveis têm uma probabilidade significativamente maior de apoiar a intervenção. O intercepto do modelo foi de $-0,95$, indicando uma razão de chances de aproximadamente 0,39 quando

todas as variáveis independentes estão em zero, o que representa uma baixa propensão à intervenção nesse ponto de referência.

No modelo voltado à intervenção militar em casos de corrupção, o “bolsonarismo” manteve um efeito ainda mais expressivo, com coeficiente de 1,60 ($OR \approx 4,97$), sugerindo que as chances de apoio aumentam quase cinco vezes por ponto adicional na escala. O “antipetismo” também teve efeito positivo, com coeficiente de 0,61 ($OR \approx 1,84$), enquanto a interação entre as variáveis foi de 0,49 ($OR \approx 1,65$). Nesse caso, embora o efeito combinado ainda seja relevante, ele é menos pronunciado do que no cenário da criminalidade. O intercepto foi de $-0,58$ ($OR \approx 0,56$), indicando uma probabilidade base moderadamente baixa de apoio.

No modelo sobre intervenção militar em situações de instabilidade política, o coeficiente de “bolsonarismo” foi de 1,67 ($OR \approx 5,36$), o mais alto entre os modelos analisados, sugerindo forte relação entre essa variável e o apoio à intervenção. O “antipetismo” teve um coeficiente mais modesto (0,27; $OR \approx 1,32$), e a interação entre ambas foi de 0,14 ($OR \approx 1,16$), indicando um efeito conjunto relativamente pequeno. O intercepto, de $-1,00$ ($OR \approx 0,37$), novamente aponta para uma propensão basal baixa de apoio à intervenção.

Em todos os modelos, o “bolsonarismo” apresentou impacto substancial e consistente na probabilidade de apoio a intervenções militares, sendo especialmente influente nos contextos de corrupção e instabilidade institucional. O “antipetismo”, embora também associado positivamente ao apoio, exerce efeito de menor magnitude. Os termos de interação revelam que o alinhamento simultâneo com essas duas orientações políticas intensifica o apoio à intervenção, sobretudo no contexto da criminalidade, sugerindo que o autoritarismo popular pode emergir com mais força quando essas atitudes políticas se somam.

Esses resultados sugerem que o apoio a intervenções militares está fortemente associado ao “bolsonarismo”, com um efeito adicional significativo quando o “antipetismo” também está presente. Assim pode-se distanciar o “bolsonarista” do “antipetista” médios, no que tange questões como o apoio a essas intervenções. Esses achados vão ao encontro da literatura canônica de ciência política como exposto na seção anterior.

A Tabela 5 apresenta métricas clássicas utilizadas para avaliar o desempenho de modelos de classificação. Estas métricas permitem verificar, de maneira quantitativa, a qualidade das predições estatística dos resultados. Descrevem-se brevemente cada métrica, discutem-se os valores obtidos e fornecem-se referências acadêmicas que embasam tais escolhas de avaliação.

A Acurácia (Accuracy) é a proporção de predições corretas em relação ao total de instâncias classificadas. No conjunto “Geral” e “Bols.” (“Bolsonarismo”), a acurácia atingiu 0,68 (68%), enquanto no conjunto “Antipt.” obteve-se 0,63 (63%). Esses valores indicam que, de modo geral, o modelo classifica corretamente aproximadamente dois terços das

Tabela 5 – Resumo das Métricas das Matrizes de Confusão

Métrica	Geral	Bols.	Antipt.
Acurácia	0.68	0.68	0.63
IC 95%	(0.61, 0.75)	(0.61, 0.75)	(0.55, 0.70)
P-Valor [Acc > NIR]	0.00	0.00	0.03
Kappa	0.34	0.34	0.24
Sensibilidade	0.82	0.80	0.71
Especificidade	0.51	0.53	0.53
Vlr. Pred. +	0.68	0.68	0.65
Vlr. Pred. -	0.70	0.68	0.59
Acc. Balanceada	0.66	0.67	0.62

Fonte: Elaboração própria.

instâncias. Embora úteis, valores de acurácia sozinhos podem ser enganosos em cenários com classes desbalanceadas (FAWCETT, 2006).

O Intervalo de Confiança de 95% (IC 95%) fornece uma faixa de valores que, com 95% de confiança, contém a verdadeira acurácia populacional. Por exemplo, para a acurácia de 0,68 no conjunto “Geral” e “Bols.” (“Bolsonarismo”), o IC 95% foi [0,61; 0,75]. Isso significa que, se repetíssemos o experimento muitas vezes, 95% dos intervalos obtidos em cada repetição conteriam a verdadeira acurácia do modelo dentro desse intervalo. Esse parâmetro é fundamental para avaliar a estabilidade estatística dos resultados (ALTMAN; BLAND, 1994).

O p-valor apresentado corresponde a um teste estatístico que verifica se a acurácia do classificador excede a taxa de acerto de um classificador nulo ou aleatório, muitas vezes referida como *No Information Rate* (NIR). Quando o p-valor é muito baixo (por exemplo, $< 0,05$), rejeita-se a hipótese nula de que o modelo não supera o “acerto ao acaso”.

Observa-se que, para o conjunto “Geral” e “Bols.”, o p-valor foi 0,00, enquanto para “Antipt.” foi 0,03. Ambos indicam evidência estatística de que o modelo supera o acerto puramente aleatório, embora o conjunto “Antipt.” apresente uma significância menor ($p = 0,03$) em comparação aos demais ($p = 0,00$).

O coeficiente Kappa (“Cohen’s Kappa”) mede o grau de concordância entre as predições do modelo e os rótulos reais, levando em conta o acaso. Nos resultados apresentados, Kappa variou entre 0,24 (“Antipt.”) e 0,34 (“Geral” e “Bols.”). Em termos gerais, um Kappa em torno de 0,20–0,40 é considerado “justo” (*fair*) ou “razoável”; isso indica que, embora haja concordância além do acaso, ainda existe margem considerável para melhoria no desempenho do classificador.

Sensibilidade (*Sensitivity* ou *Recall*): é a fração de verdadeiros positivos corretamente identificados pelo modelo. Quanto maior a sensibilidade, menor a probabilidade

de que o modelo deixe de detectar exemplos positivos. Especificidade (*Specificity*): é a fração de verdadeiros negativos corretamente identificados. Quanto maior a especificidade, menor a probabilidade de que o modelo atribua de forma incorreta um rótulo positivo a um exemplo negativo (POWERS, 2011). Nos resultados, as sensibilidades variam de 0,71 (“Antipt.”) a 0,82 (“Geral”), sugerindo que o modelo tem boa capacidade de identificar positivamente a maioria dos casos positivos. As especificidades, por outro lado, variam de 0,51 a 0,53, indicando uma performance mais modesta em reconhecer corretamente os casos negativos.

Valor Preditivo Positivo (*Positive Predictive Value* – PPV): indica a probabilidade de uma classificação positiva estar correta. Valor Preditivo Negativo (*Negative Predictive Value* – NPV): indica a probabilidade de uma classificação negativa estar correta. No presente estudo, o PPV situou-se entre 0,65 (“Antipt.”) e 0,68 (“Geral” e “Bols.”), enquanto o NPV variou de 0,69 (“Antipt.”) a 0,72 (“Bols.”). Esses valores mostram que, de modo geral, o modelo tende a classificar positivamente de forma correta em cerca de 65–68% das vezes em que prevê positivo, ao passo que cerca de 69–72% das predições negativas são efetivamente negativas (POWERS, 2011).

Por último, a acurácia balanceada é a média entre Sensibilidade e Especificidade. Esse indicador é particularmente relevante quando se lida com desequilíbrio de classes, pois mitiga o efeito de haver uma classe com número muito superior de amostras. Os valores variaram entre 0,62 (“Antipt.”) e 0,67 (“Bols.”), sugerindo que, embora a sensibilidade seja relativamente alta, a menor especificidade puxa a acurácia média para valores próximos de 0,65 (BRODERSEN et al., 2010).

Observa-se que os modelos “Geral” e “Bols.” (“Bolsonarismo”) apresentam acurácias e Kappas semelhantes, com sensibilidade elevada ($\geq 0,80$) e especificidade por volta de 0,50. Já o modelo “Antipt.” possui acurácia um pouco menor (0,63) e Kappa de 0,24, indicando menor concordância com os rótulos verdadeiros. Ainda assim, os testes estatísticos (p-valores) mostram que as acurácias excedem significativamente o acerto aleatório, o que valida o uso prático desses modelos dependendo do contexto.

A sensibilidade relativamente alta indica que o modelo identifica corretamente a maioria dos casos positivos, o que pode ser desejável em cenários onde se busca minimizar falsos negativos (p. ex., detecção de fraudes ou doenças). Entretanto, a especificidade mediana alerta para a possibilidade de muitos falsos positivos, o que pode ter consequências adversas dependendo do domínio de aplicação (BISHOP, 2006).

Ao comparar os modelos de regressão logística com “bolsonarismo” e “antipetismo” individualmente como variáveis explicativas, é possível observar algumas diferenças de desempenho. O modelo com “bolsonarismo” demonstrou uma acurácia ligeiramente superior (68,21%) em comparação com o modelo que utiliza “antipetismo” (63,01%). Além disso, o modelo com “bolsonarismo” apresentou uma sensibilidade melhor (80,21%) em relação ao

outro modelo (70,83%), indicando maior eficácia na identificação de ser contra (classe 0). No entanto, ambos os modelos apresentaram a mesma especificidade (53,25%), indicando uma eficácia equivalente na identificação da classe a favor (classe 1).

No que tange a precisão do modelo, o “bolsonarismo” foi ligeiramente superior (68,14%) em comparação com o modelo com “antipetismo” (65,38%). O coeficiente kappa do primeiro modelo foi maior (0,34), indicando uma concordância moderada, enquanto o segundo modelo teve um kappa menor (0,24), indicando uma concordância justa.

Na síntese do necessário, o modelo que utiliza “bolsonarismo” como variável explicativa individualmente apresentou uma performance geral superior em comparação com o modelo que utiliza “antipetismo”. No entanto, ambos os modelos demonstraram uma precisão moderada e apresentam potencial para melhorias. A otimização desses modelos, especialmente na especificidade e na precisão da classe “a favor” (classe 1), poderia resultar em desempenhos mais robustos e eficazes. Conforme esse trabalho vem desenvolvendo, isso ocorre porquanto há “antipetistas” que não são “bolsonaristas” e, por essa razão, não compartilham dos mesmos traços que a literatura canônica sobre “bolsonarismo” estabelece.

Como a regressão logística avalia variáveis dependentes binárias, métrica AUC (Área Sob a Curva) ROC (*Receiver Operating Characteristic*) é amplamente utilizada para avaliar a performance desses modelos binários.

AUC (Área Sob a Curva) ROC é uma métrica utilizada para avaliar a performance de um modelo de classificação binária. AUC-ROC, onde ROC significa *Receiver Operating Characteristic* (Característica de Operação do Receptor), é uma forma gráfica de representar a relação entre a taxa de verdadeiros positivos (True Positive Rate - TPR) e a taxa de falsos positivos (*False Positive Rate* - FPR) em diferentes pontos de corte do modelo (ZHOU; OBUCHOWSKI; MCCLISH, 2011; HANLEY; MCNEIL, 1982; FAWCETT, 2006).

O modelo que considera apenas o “bolsonarismo” (AUC = 0,72) apresenta a melhor capacidade discriminativa entre os três modelos testados. O modelo que considera apenas o “antipetismo” (AUC = 0,63) demonstra a menor capacidade discriminativa, embora ainda forneça alguma informação preditiva. Já o modelo que incorpora a interação entre “bolsonarismo” e “antipetismo” (AUC = 0,70) também apresenta uma boa capacidade discriminativa, mas não supera o desempenho do primeiro modelo com “bolsonarismo” isoladamente.

Esses resultados sugerem que “bolsonarismo” é uma variável preditiva mais robusta do que “antipetismo” na previsão do apoio à intervenção militar. A interação entre as duas variáveis acrescenta algum valor preditivo, mas não o suficiente para superar o poder preditivo do “bolsonarismo” por si só.

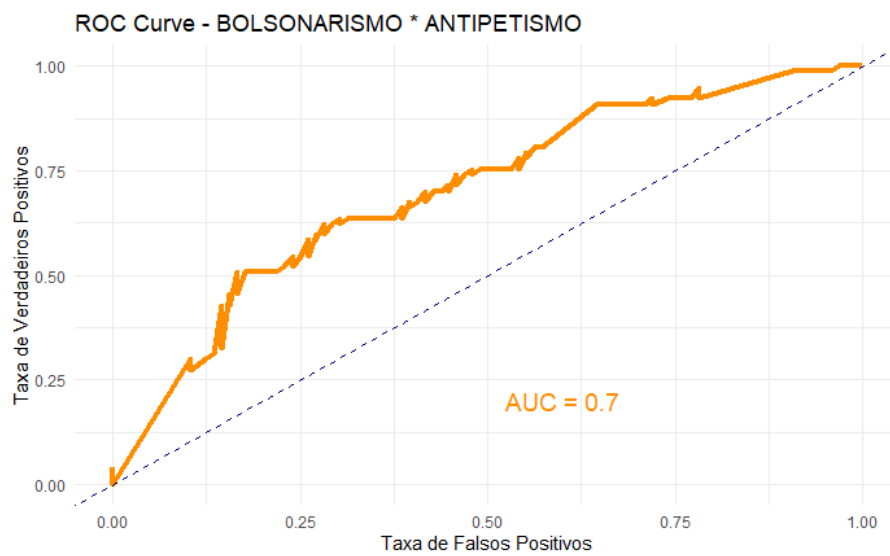


Figura 3 – ROC - Curve - Bolsonarismo e Antipetismo

Fonte: Elaboração própria

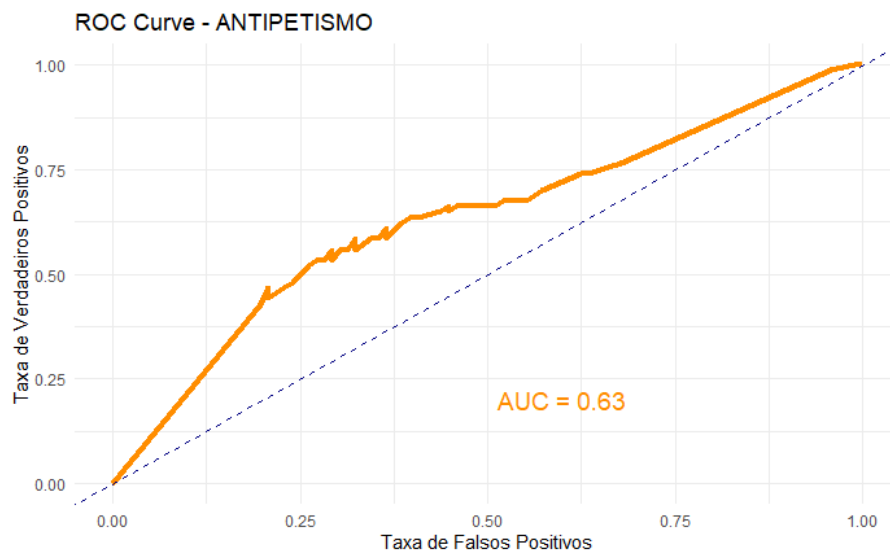


Figura 4 – ROC - Curve - Antipetismo

Fonte: Elaboração própria

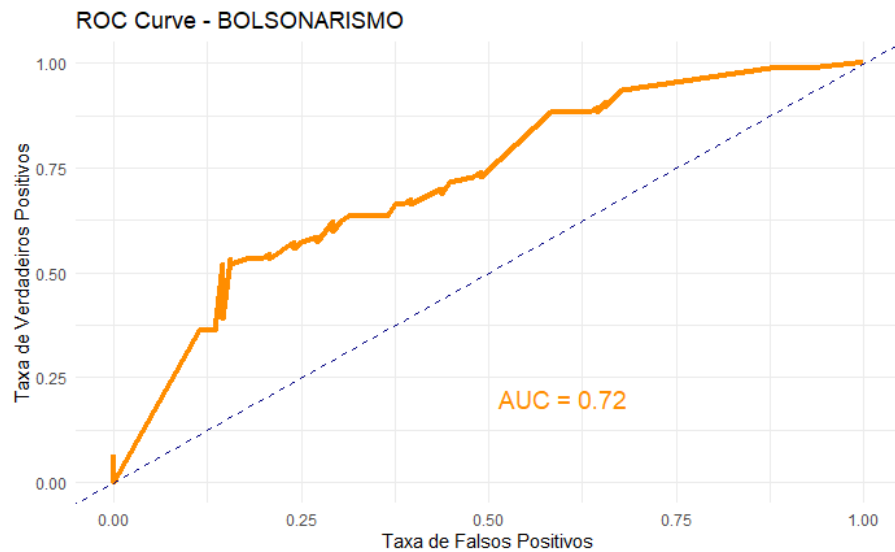


Figura 5 – ROC - Curve - Bolsonarismo

Fonte: Elaboração própria

2.4 Conclusão

Através do uso de *Fuzzy C-Means* e *Fuzzy Regression Discontinuity*, este estudo demonstrou a eficácia dessas ferramentas em identificar e analisar padrões complexos dentro do comportamento eleitoral, como o “bolsonarismo” e o “antipetismo”. A capacidade do *Fuzzy C-Means* de atribuir graus de pertencimento a diferentes agrupamentos permitiu uma análise mais nuançada das variáveis envolvidas, enquanto a *Fuzzy Regression Discontinuity* ofereceu uma maneira robusta de explorar relações causais em cenários onde a atribuição de tratamento não é perfeitamente clara.

Os resultados indicam que, embora o “antipetismo” seja um fator significativo na determinação do “bolsonarismo”, ele não é suficiente por si só. Outros fatores, como o conservadorismo político, desempenham um papel crucial. Além disso, a análise de regressão logística com termos de interação revelou a complexidade das interações entre “bolsonarismo” e “antipetismo”, especialmente no contexto de apoio a intervenções militares.

A identificação de elementos adicionais, como o conservadorismo político, como fatores determinantes no comportamento eleitoral destaca a importância de expandir o espaço representacional para acomodar novas dimensões explicativas. Esse processo reflete a capacidade da teoria da representação de integrar novas variáveis e relações em uma estrutura analítica coesa.

A teoria da representação oferece o alicerce conceitual para que ferramentas de aprendizado de máquina avancem na análise do comportamento eleitoral, permitindo uma tradução rigorosa, porém flexível, de dinâmicas sociais complexas em estruturas formais. Isso enriquece o entendimento de fenômenos como o “bolsonarismo” e o “antipetismo”,

proporcionando insights mais profundos e aplicáveis à ciência política e à análise de dados.

Entro os pontos deixados em aberto nesse trabalho para que possa ser explorado futuramente, pode-se citar explorar ainda mais as características dessa “direita não bolsonarista”. As eleições regionais de 2024 quebraram o paradigma das polarizadas eleições de 2022. Diferente destas, os eleitores orbitaram partidos de direita, centro direita e centro, indo de encontro com a dicotomia Partido dos Trabalhadores (PT) e Partido Liberal (PL) que ocorreu na última eleição presidencial. Esse resultado vai ao encontro desse trabalho porquanto na primeira análise demonstrou-se a existência de uma “direita não bolsonarista” mais numerosa e expressiva que a “direita bolsonarista”.

Este trabalho reforça a importância de métodos avançados e flexíveis, como os de aprendizado estatístico, para capturar a realidade multifacetada da política contemporânea. Ao integrar essas técnicas, pesquisadores podem obter *insights* mais precisos e aprofundados, contribuindo para a evolução da Ciência Política. No contexto das transformações metodológicas, os sistemas informacionais oferecem um caminho promissor para superar as limitações dos métodos tradicionais, proporcionando uma ferramenta poderosa para a análise e a previsão de comportamentos políticos complexos.

Adicionalmente, a lógica de pensamento científico proposta por Karl Popper enfatiza a importância da falseabilidade nas hipóteses científicas. Segundo Popper, uma hipótese científica deve ser passível de ser testada e potencialmente refutada para ser considerada válida. A aplicação sistemas computacionais na ciência política segue essa premissa ao permitir a construção de modelos que podem ser rigorosamente testados e desafiados, proporcionando um avanço no rigor científico das análises políticas (POPPER, 2002).

Portanto, a continuidade na exploração e desenvolvimento dessas abordagens é não apenas desejável, mas necessária para avançar a compreensão dos fenômenos políticos. Com isso, espera-se que futuras pesquisas possam refinar e expandir o uso dessas técnicas, ampliando ainda mais seu impacto na análise política e nas tomadas de decisão informadas.

A crescente complexidade das interações sociais e políticas na era digital exige abordagens inovadoras para entender e prever comportamentos políticos. Sistema de Informações, com suas técnicas de análise de grandes volumes de dados, oferecem uma oportunidade única para desvendar padrões e tendências que, de outra forma, permaneceriam ocultos.

3 Redes Neurais e Polarização Política - Modelo BERT-Vi

Este capítulo explora o uso de redes neurais (NNs) para analisar a polarização política nas eleições presidenciais de 2018 e 2022 no Brasil. Métodos tradicionais apresentam dificuldades ao lidar com dados não estruturados e padrões complexos inerentes ao comportamento político. Ao aplicar modelos avançados de *deep learning*, incluindo BERT e XLM-RoBERTa, este trabalho classifica posições políticas a partir de dados de entrevistas e pesquisas a fim de demonstrar como os princípios da teoria da representação aprimoram a interpretabilidade e o poder preditivo. Esta pesquisa fornece uma estrutura escalável e robusta para a compreensão do comportamento político, contribuindo para a ciência política computacional e para a análise de democracias polarizadas em todo o mundo.

As eleições presidenciais brasileiras de 2018 e 2022 evidenciaram um eleitorado polarizado e rígido, onde divisões ideológicas e partidárias se consolidaram de forma notória não só entre família e amigos, como também na grande mídia e nas instituições democráticas. Esse cenário apresenta uma oportunidade única para explorar padrões e tendências que contribuem para essa polarização, permitindo um entendimento mais profundo dos mecanismos sociais e políticos em jogo. Essa crescente complexidade das interações sociais e políticas - alavancada pela era digital - exige abordagens atualizadas para compreender e prever comportamentos políticos.

Nesse contexto, a aplicação de redes neurais (*Neural Networks*, NNs) e técnicas de aprendizado de máquina (*Machine Learning*, ML) na ciência política tem ganhado destaque ao longo da última década (SLEGTE; DROOGENBROECK; GINIS, 2024). Redes neurais, especialmente os modelos de aprendizagem profunda (*deep learning*), têm demonstrado grande potencial em diversas áreas de pesquisa da ciência política porque elas permitem lidar com dados não estruturados e ao identificar padrões, ampliando as possibilidades de análise.

As NNs têm ampla aplicação em tarefas de processamento de linguagem natural (*Natural Language Processing*, NLP) podendo ser facilmente associadas a diversos contextos relacionados à ciência política. Essas tarefas incluem: a) interpretar emoções, opiniões e tendências em postagens de plataformas como Twitter, Facebook e outras redes sociais de maneira a mapear a polarização política, monitorar a reação pública a eventos e até prever mudanças no comportamento eleitoral; b) categorização de documentos políticos, como manifestos eleitorais e programas partidários, identificando os temas principais, as prioridades políticas e as orientações ideológicas subjacentes com maior precisão; c) processar e analisar grandes volumes de texto para identificar padrões latentes que revelam

ideologias políticas, narrativas dominantes e estratégias discursivas. Esses tipos de análise são utilizados para compreender as dinâmicas das campanhas políticas e a evolução das ideologias ao longo do tempo (CHATSIOU; MIKHAYLOV, 2020).

Além do uso em NLP, as NNs têm sido utilizadas para prever resultados eleitorais e analisar o comportamento dos eleitores. Como elas são capazes de lidar com relações não-lineares, frequentemente presentes em fenômenos sociais, isso torna essas ferramentas ideias para análises preditivas. Por exemplo, redes neurais podem integrar e analisar grandes volumes de dados, como histórico de votos, dados socioeconômicos, pesquisas de opinião e tendências nas redes sociais, para prever com precisão o desempenho de candidatos e partidos em eleições. As NNs têm se mostrado capazes de detectar padrões complexos nas decisões dos eleitores, incluindo a influência de campanhas publicitárias, *fake news* e interações em redes sociais. Esses modelos ajudam a entender como os eleitores respondem a eventos específicos e como seus comportamentos se moldam ao longo do tempo (KHAZE; MASDARI; HOJJATKHAH, 2013; ZHOU; MAKSE, 2019; PINEDA et al., 2022).

Nesse contexto, esse trabalho se posiciona no sentido de que a aplicação de NNs na análise do comportamento político oferece uma ferramenta para desvendar padrões e relações complexas em dados eleitorais. Dessa maneira, como a polarização política e a rigidez ideológica têm marcado os ciclos eleitorais recentes, os modelos NNs apresentam uma oportunidade para superar os métodos tradicionais de análise ¹.

Assim, o objetivo geral desse trabalho é entender como a Aprendizagem Profunda, uma subárea da Inteligência Artificial, é utilizada para detectar posicionamentos políticos em um contexto de entrevistas de profundidade aplicadas em grupos focais, especificamente a avaliação do Governo Bolsonaro (2019-2022). Os objetivos específicos desse trabalho são: a) fazer uma análise exploratória das respostas dadas pelos entrevistados com o intuito de encontrar padrões e convergências que sustentem a construção de um modelo específico para essa análise; b) modelar uma NN capaz de identificar padrões consistentes nos dados, de modo a inferir relações causais ou correlacionais que expliquem determinados comportamentos ou resultados políticos.

Para chegar a esses objetivos, será construído um modelo de rede neural de classificação que sustente que as respostas convergem de maneira preditiva. Isso implica no desenvolvimento de uma arquitetura capaz de categorizar dados de entrada de forma consistente com padrões previamente identificados. Esse modelo de NN será projetado para capturar relações complexas e não lineares entre as respostas, garantindo que as previsões geradas reflitam a convergência dos dados observados em direção a resultados rotulados.

¹ Entre os métodos tradicionais, pode-se mencionar (VEIGA; GONDIM, 2001)

Antes de abordar a metodologia de trabalho, é importante destacar que esse trabalho se posiciona e vai ao encontro da Teoria da Representação (ARMENTA; JODOIN, 2021). A Teoria da Representação, empregada para estudar como objetos matemáticos podem ser representados em espaços vetoriais, encontra aplicabilidade significativa em redes neurais, especialmente em suas camadas de transformação e aprendizado de características (*features*).

No domínio político, essa teoria permite que as representações dos modelos reflitam melhor as relações e estruturas subjacentes nos dados, tais como: a) aplicação de representações permite que o modelo capture as interações entre dimensões ideológicas e sociais, como progressismo e conservadorismo, ou entre alinhamentos partidários; b) utilizar princípios de simetria ajuda a garantir que as transformações respeitem as propriedades dos dados, como simetrias regionais ou invariância a traduções temporais em dados eleitorais longitudinais.

No caso desse trabalho, incorporar representações teóricas aos *embeddings* de texto (por exemplo, usando BERT) alinha as palavras e frases com as dimensões que serão observadas. Representações fundamentadas teoricamente tornam os modelos mais interpretáveis, como, por exemplo, o uso de *embeddings* supervisionados alinhados com categorias políticas específicas pode facilitar a interpretação das saídas do modelo em termos de conceitos políticos. De maneira a construir modelos interpretáveis que aumentem a confiança dos analistas políticos ao aplicá-los em seus estudos.

Entre os trabalhos que referenciam a Teoria da Representação em NNs, pode-se citar *The Representation Theory of Neural Networks* 2021, neste trabalho, os autores demonstram que NNs podem ser representadas por meio da teoria matemática das representações de *quivers*. Os autores provam que uma rede neural é uma representação de um *quiver* com funções de ativação, introduzindo o conceito de *network quiver*. Essa interpretação algébrica fornece inferências sobre como as redes neurais processam e armazenam dados, mapeando-os para um espaço geométrico chamado espaço de módulos, que é definido pelo grafo orientado subjacente à rede.

Pode-se citar também o *Representation Learning: A Review and New Perspectives* 2013, neste artigo, os autores revisam trabalhos em aprendizado de características não supervisionado e aprendizagem profunda, abordando avanços em modelos probabilísticos, *autoencoders*, aprendizado de variedades (*manifold learning*) e redes profundas. Além disso, os autores discutem como aprender representações dos dados pode facilitar a extração de informações úteis ao construir classificadores ou outros preditores, enfatizando a importância do aprendizado de representações na aprendizagem de máquina.

Portanto, esse trabalho se encaixa no âmbito da Teoria da Representação ao abordar como diferentes formas de dados políticos, como transcrições de entrevistas e respostas a pesquisas podem ser sistematicamente transformadas em representações estruturadas

que permitem análise. A Teoria da Representação oferece o fundamento matemático e conceitual para compreender essas transformações, especialmente em redes neurais, onde dados brutos são codificados em espaços vetoriais de alta dimensão para capturar relações e padrões latentes.

Ao utilizar técnicas como *word embeddings*, extração de características e arquiteturas grupo-equivariantes, a abordagem desse trabalho assegura que informações críticas sejam preservadas e interpretáveis. Essa coerência não apenas aprimora a capacidade preditiva do modelo desenvolvido, mas também destaca a busca pelo rigor teórico da metodologia, posicionando-a em um discurso mais amplo sobre como os dados são abstraídos, processados e analisados em aprendizado de máquina e ciência política computacional. Sob essa perspectiva, busca-se contribuir para a exploração de como a representação impacta a compreensão e modelagem de comportamentos políticos complexos.

3.1 Metodologia

3.1.1 Banco de Dados

A pesquisa² foi realizada entre os dias 8 e 20 de abril de 2021, na cidade do Rio de Janeiro. Durante o período, foram conduzidas oito entrevistas de profundidade em grupos focais utilizando a plataforma virtual gratuita *Google Meet*. Cada sessão teve aproximadamente duas horas de duração, com uma média de oito participantes por grupo, totalizando 62 pessoas. A seleção dos entrevistados abrangeu as quatro zonas da cidade: Central, Norte (ZN), Sul (ZS) e Oeste (ZO). Os grupos foram compostos de forma mista, incluindo homens e mulheres, com idades entre 25 e 55 anos, pertencentes à classe C em termos de rendimento mensal. (ESTRELLA; BORBA, 2024).

Esse método baseia-se na ideia de que os participantes são reciprocamente influenciados pelas dinâmicas de poder que emergem ao longo das discussões, o que contribui para aprofundar as conclusões sobre os consensos e divergências relacionados aos temas abordados. Além disso, considera-se que esses pequenos grupos frequentemente refletem, em suas conversas, os discursos predominantes em suas interações macrosociais. No entanto, é importante levar em conta fatores que podem interferir no processo, como a simultaneidade entre pensar e ouvir, o receio de desaprovação social e o efeito de persuasão. Apesar dessas limitações, o moderador pode empregar estratégias específicas para mitigar tais desafios (VEIGA; GONDIM, 2001; ESTRELLA; BORBA, 2024).

A partir da coleta dessas imagens, foram usados aplicativos³ de *video-to-text* para

² Esse trabalho agradece aos responsáveis pelo trabalho com os grupos focais utilizado por terem cedido as imagens de vídeo dos encontros para análise desse artigo: Prof. Dr. Felipe Borba, Profa. Dra. Luciana Veiga, M.Sc. Thais Estrella de Alencar.

³ Nota do Autor: para o leitor curioso, os aplicativos usados foram <https://transkriptor.com/> e <https://happyscribe.com>.

transcrever as entrevistas em sua totalidade. Em seguida, foram selecionadas somente as respostas referentes à pergunta de avaliação do Governo de Jair Messias Bolsonaro. Embora os grupos focais tenham avaliado outros assuntos, para esse trabalho, focou-se somente na avaliação geral e econômica do referido governo.

A partir da separação desse texto, o banco de dados foi criado a partir de entradas de até 360 caracteres, contando uma afirmação e a justificativa dessa afirmação - quando se aplica. Como o objetivo é classificar opiniões contra o Governo Bolsonaro, essas entradas foram rotuladas com o valor 1; todas as outras entradas: opiniões mistas, positivas ou divagações foram classificadas como 0. Essa separação foi feita porquanto o objetivo do modelo é identificar aqueles que se posicionam de maneira objetiva e enfática contra o Governo Bolsonaro.

A abordagem metodológica para classificar opiniões sobre o Governo Bolsonaro se justifica porquanto rotular entradas como opostas (1) ou não opostas (0) ao Governo Bolsonaro corresponde aos métodos de classificação binária usados na análise de sentimento político. Por exemplo, o estudo *“Sentiment is Not Stance: Target-Aware Opinion Classification for Political Text Analysis”* enfatiza a importância de distinguir entre sentimento e postura em textos políticos, defendendo uma classificação para capturar com precisão as opiniões políticas (BESTVATER; MONROE, 2023).

Ao escolher limitar as entradas a 360 caracteres, incluindo declarações e justificativas, buscou-se consistência com metodologias que focam em expressões de opinião. A pesquisa sobre análise de sentimento da comunicação política frequentemente envolve a análise de segmentos curtos de texto para capturar efetivamente a essência das opiniões políticas. Entre essas metodologias pode-se citar o estudo *“Sentiment analysis of political communication: combining a dictionary approach with crowdcoding”* onde são discutidos métodos para analisar textos políticos, destacando a relevância do comprimento do texto e do conteúdo na classificação de sentimento (HASELMAYER; JENNY, 2017).

Empregar técnicas de aprendizado supervisionado, onde entradas são rotuladas manualmente para treinar modelos, é uma prática padrão na classificação de texto. Ao se concentrar em entradas que se opõem explicitamente ao Governo Bolsonaro busca-se garantir clareza na classificação, semelhante a metodologias que visam posições políticas específicas. Esse trabalho vai ao encontro do artigo *“Stance detection: a practical guide to classifying political belief in texts”* onde é fornecida a orientação sobre como detectar posições políticas explícitas em textos, apoiando o foco em expressões claras de opinião em sua abordagem (BURNHAM, 2024).

Treinar redes neurais requer um conjunto de dados suficientemente grande para garantir desempenho e capacidade de generalização (ZHANG et al., 2024; YU; WANG; SHI, 2023). Embora o conjunto inicial contenha 600 entradas equilibradas, expandir esse conjunto de dados é essencial para treinar o modelo de rede neural de maneira a performar

para a tarefa designada.

Para ampliar o conjunto de dados trabalhado, foram usadas duas de técnicas de aumento de dados (*data augmentation*) - que envolvem a criação de versões modificadas dos dados existentes para aumentar artificialmente o tamanho do conjunto. A primeira técnica foi substituição de sinônimos (*synonym replacement*) e a segunda técnica foi retradução (*back-translation*) usando idiomas não latinos⁴, especificamente o coreano, japonês e mandarim. Ambos métodos são estabelecidos no processamento de linguagem natural (PLN) para aumentar a diversidade dos dados e melhorar a capacidade de generalização além de reduzir o risco de *overfitting*. (IWANA; UCHIDA, 2021).

No que tange a substituição de Sinônimos (*Synonym Replacement*): Essa técnica consiste em substituir palavras de uma sentença por seus sinônimos, gerando novas frases que mantêm o significado original. Essa variação lexical ajuda os modelos a generalizarem melhor, ao expô-los a diferentes formas de expressar o mesmo conceito.

Já a retradução (*back-translation*) envolve traduzir uma frase do idioma original para um idioma-alvo e, em seguida, traduzi-la de volta para o idioma original. Esse processo introduz variações linguísticas naturais, criando versões parafraseadas do texto original. O uso de idiomas não latinos, como coreano, japonês e mandarim, pode enriquecer ainda mais o conjunto de dados, adicionando estruturas sintáticas e expressões diversificadas. A utilidade da retradução na ampliação de dados tem sido demonstrada em diversos estudos, como em *Data Augmentation Techniques for Machine Translation of Code-Switched Texts: A Comparative Study* 2023, que evidencia sua eficácia em melhorar a robustez dos modelos.

Assim, esse trabalho usará três bancos de dados para a modelagem, o banco de dados raiz com 600 entradas para análise exploratória e teste da capacidade preditiva. O segundo banco de dados apenas como substituição de sinônimos com aproximadamente 1500 entradas e o último com quase 6000 entradas para modelagem. Ambas as classes devidamente balanceadas em todos os banco de dados.

Ao integrar essas estratégias de ampliação de dados, esse trabalho busca aumentar a variabilidade e a riqueza do seu conjunto de treinamento, o que é fundamental para desenvolver modelos de redes neurais robustos e generalizáveis em tarefas de processamento de linguagem natural (PLN).

3.1.2 Análise Exploratória

Antes de desenvolver um modelo de aprendizagem profunda, é fundamental realizar uma análise exploratória dos textos para compreender a estrutura, os padrões e as caracte-

⁴ Foram utilizados idiomas não latinos para pleitear garantir uma maior variedade léxica

rísticas do conjunto de dados. Essa etapa preliminar não apenas oferece uma inferência sobre os dados, mas também demonstra uma possível convergência ou não do modelo.

No contexto deste estudo, serão usados os pacotes *quanteda* e *syuzhet* no R (R Core Team, 2021). O pacote *quanteda* (BENOIT et al., 2018) oferece ferramentas para análise quantitativa de texto, possibilitando tarefas como tokenização, análise de frequência e criação de matrizes de características de documentos. Com esse pacote, busca-se examinar frequências de palavras, *n-grams* e diversidade lexical no conjunto de dados, revelando temas e identificando possíveis vieses ou desequilíbrios para o posterior modelo que será treinado.

Complementando essa abordagem, o pacote *syuzhet* (JOCKERS, 2015) facilita a análise de sentimentos e a detecção de emoções, permitindo analisar o tom emocional do texto e categorizá-lo em estados afetivos. Essa análise é útil para compreender como o sentimento varia em diferentes partes do conjunto de dados, o que pode ter implicações em tarefas posteriores, como classificação ou predição.

Ao combinar essas ferramentas, essa análise exploratória visa criar uma base para o desenvolvimento de um modelo de aprendizagem profunda. Essa etapa destacará padrões que justifiquem que o modelo seja treinado.

A análise exploratória de texto desempenha um papel fundamental ao conectar dados com o arcabouço da teoria da representação nos Sistemas de Informação (SI). A teoria da representação em SI foca em como a informação é codificada, estruturada e transformada para facilitar a interpretação e a tomada de decisões. Essa perspectiva está alinhada com os objetivos da análise exploratória, cujo propósito é revelar padrões, relações e estruturas que possam informar o desenvolvimento de modelos preditivos ou outros quadros analíticos (RECKER et al., 2019; BURTON-JONES et al., 2017; HART; GREGOR, 2007).

No contexto deste estudo, o uso dos pacotes *quanteda* e *syuzhet* para análise exploratória apoia diretamente os princípios da Teoria da Representação por meio de codificação dos dados em representações estruturadas. O pacote *quanteda* permite a transformação de dados textuais em formas estruturadas, como matrizes de características de documentos (*document-feature matrices*) e *n-grams*. Essas representações capturam propriedades lexicais e sintáticas do texto, oferecendo uma base para a análise de relações e padrões, de modo a preservar o significado contido nos dados originais.

Além disso, por meio do pacote *syuzhet*, a análise de sentimentos e emoções adiciona uma camada de representação semântica aos dados. Os tons emocionais e os estados afetivos, extraídos como características estruturadas, representam elementos presentes nos dados. Isso atende às necessidades representacionais dos Sistemas de Informação (SI), ao capturar tanto informações explícitas quanto latentes.

Portanto, a fase exploratória garante que os dados textuais sejam não apenas representados, mas também alinhados aos requisitos dos modelos de aprendizagem profunda nas etapas posteriores. Ao identificar características-chave, essa análise contribui para a construção de representações que são tanto interpretáveis quanto preditivas, atendendo aos dois objetivos principais da teoria da representação nos Sistemas de Informação (SI): precisão e significância.

Primeiramente, os dados textuais do banco de dados raiz são transformados em formas tokenizadas, com remoção de pontuação, símbolos, números, URLs e *stopwords*. Além disso, um conjunto adicional de palavras de alta frequência e contextualmente irrelevantes (como “é”, “acho”, “então”) é excluído para refinar o conjunto de dados⁵.

A primeira análise proposta é palavras-chave (*keyness*). A análise de *keyness* é uma técnica na linguística de *corpus*, permitindo que pesquisadores identifiquem palavras ou frases que ocorrem com frequência estatisticamente significativa em um *corpus* específico em comparação com um *corpus* de referência. Esse método revela ênfases temáticas e padrões ideológicos nos dados textuais (GABRIELATOS, 2018).

As figuras 6 e 7 oferecem uma análise comparativa dos dados textuais extraídos dos *corpora* “contra” (entrevistados que se declararam explicitamente críticos ao governo Bolsonaro) e “naocontra” (entrevistados à favor, opiniões ambíguas e neutros), com foco na frequência lexical e na divergência temática. Essas análises revelam as diferenças na estrutura do discurso, nas escolhas lexicais e nas prioridades temáticas entre perspectivas opostas sobre o Governo Bolsonaro.

No que tange a diferenciação lexical, no corpus “contra”, palavras como “postura”, “presidente” e “problema” predominam, refletindo um foco na crítica às qualidades de liderança e questões relacionadas à governança. Por outro lado, no corpus “naocontra”, termos como “culpa”, “Bolsonaro” e “PT” (Partido dos Trabalhadores) são mais frequentes, indicando uma narrativa que pode deslocar a responsabilidade para entidades políticas opostas ou defender o Governo Bolsonaro.

É possível destacar também - no que tange padrões temáticos - que o corpus “contra” enfatiza temas relacionados a falhas na governança (“problema” e “criminoso”), à distribuição de vacinas (“vacinas”, “Butantan”) e ao impacto social. Já o corpus “naocontra” é caracterizado por termos ideológicos e avaliativos mais amplos, como “governo”, “culpa” e “STF” (Supremo Tribunal Federal), sugerindo uma moldura discursiva que aborda questões sistêmicas em vez de focar na responsabilidade individual.

Essa divergência lexical destaca diferentes estratégias discursivas. O corpus “contra” busca deslegitimar o Governo Bolsonaro por meio de críticas específicas e diretas. Já o corpus “naocontra” tende a redirecionar a crítica ou contextualizar ações dentro de um

⁵ Todos os detalhes do pré-processamento pode ser acessado na página do github do autor [BERT-VI](#)

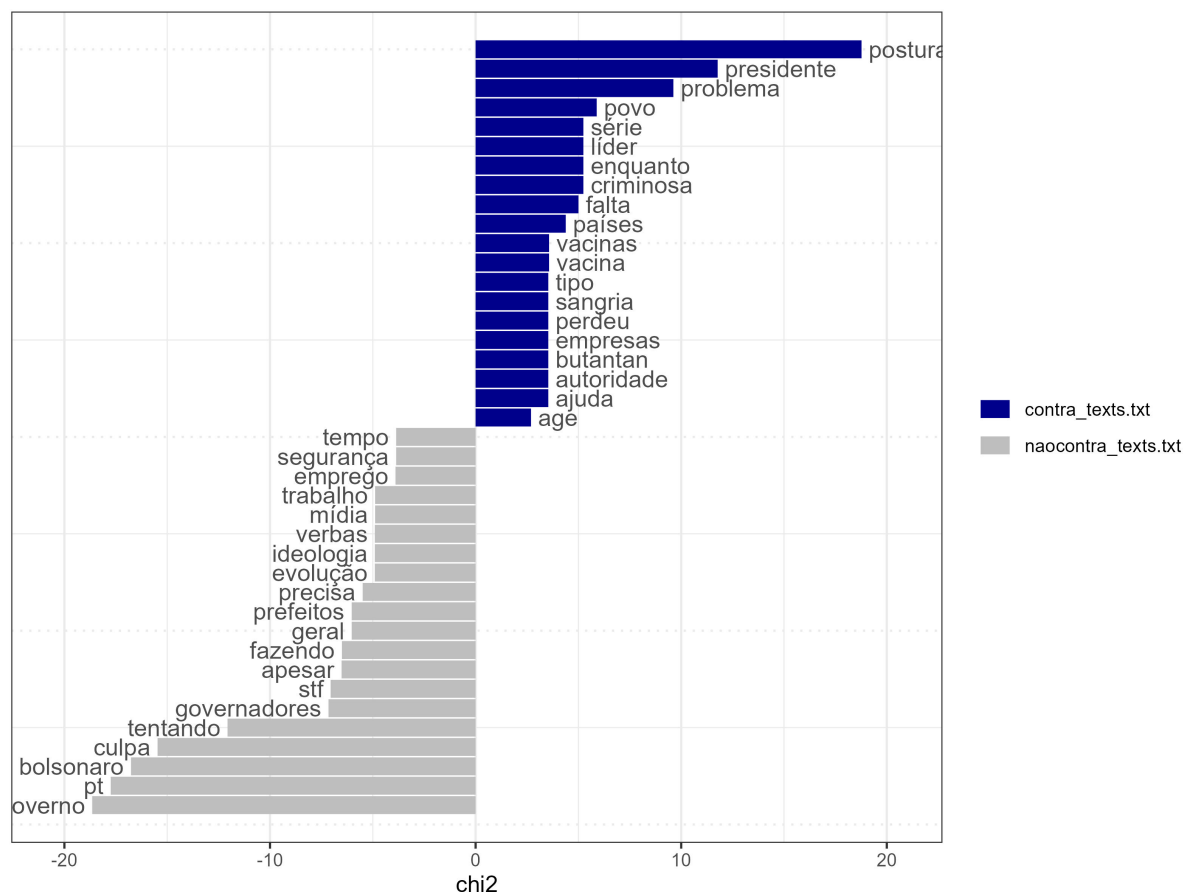


Figura 6 – Gráfico de Palavras-Chave

Fonte: Elaboração própria.

panorama político mais amplo, afastando a ênfase de uma responsabilidade individual. Esses padrões evidenciam não apenas diferenças na escolha de palavras, mas também nas prioridades discursivas e nas abordagens narrativas dos dois grupos analisados.

Na figura 7, ao examinar os bigramas para compreender as estruturas temáticas, pode-se destacar termos como “burrice_negacionismo”, “postura_presidente” e “grandes_empresas” no *corpus* “contra”. Destaca maneira, o foco em percepções de incompetência, má gestão durante a pandemia e favorecimento empresarial em políticas governamentais ficam mais latentes. Outros bigramas, como “estelionato_eleitoral”, sugerem problematizações e acusações política eleitoreiras.

Já no *corpus* “naocontra”, frases como “governo_ruim” e “votei_Bolsonaro” indicam discussões sobre a qualidade da governança e decisões pessoais de voto, isso ocorre porquanto tanto apoiadores declarados como críticos ao governo estão agrupados juntos. Nas frases “ideologia_bolsonaro” e “linha_conservadora” apontam para alinhamento ideológico entre o entrevistado e o Governo.

Uma das características da polarização discursiva desses textos é que o *corpus*

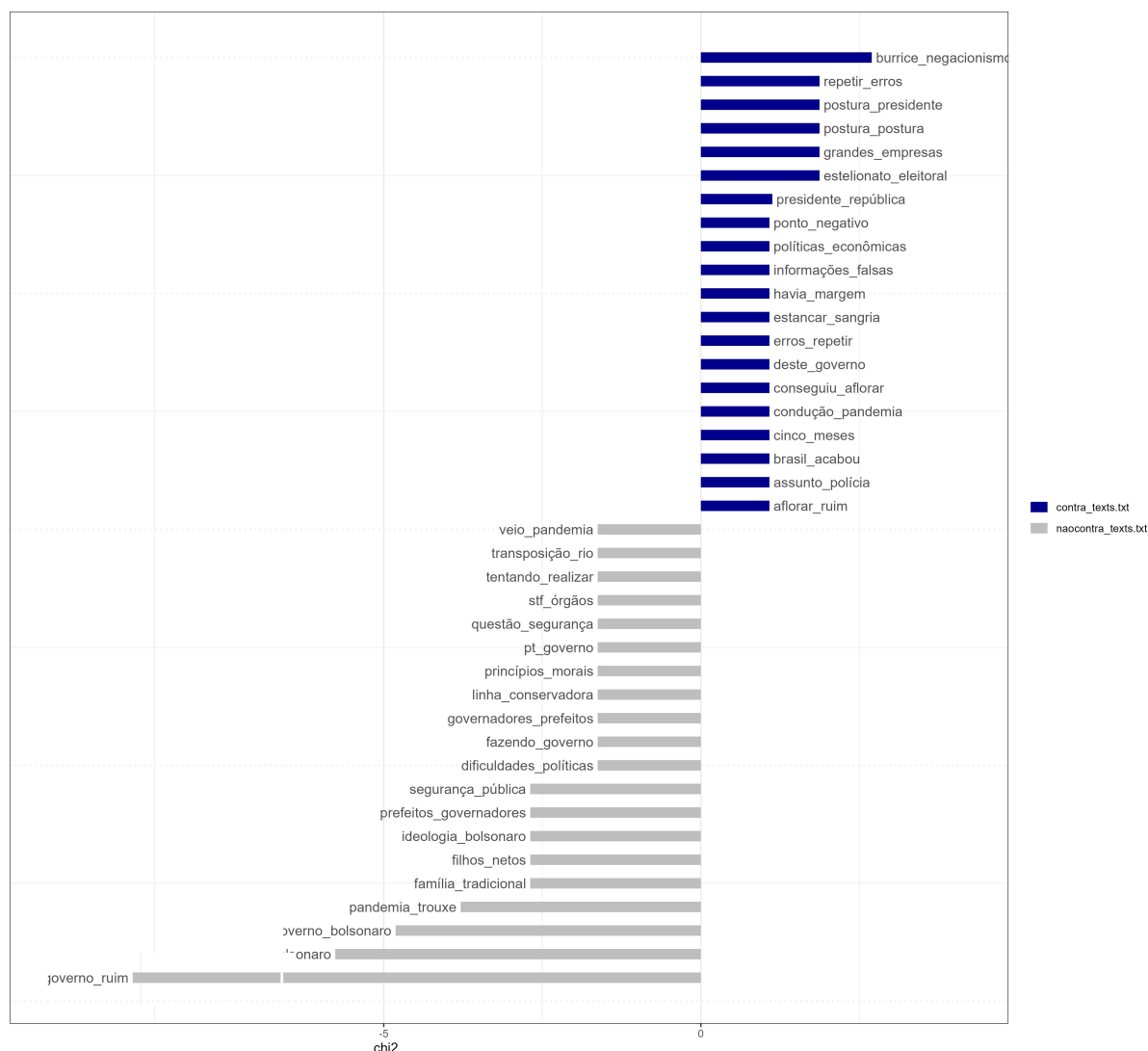


Figura 7 – Gráfico de bigramas chave

Fonte: Elaboração própria.

“contra” foca em críticas intensas sobre falhas de governança e liderança, com o objetivo de enfatizar a responsabilização e o corpus “naocontra” adota narrativas de defesa ideológica e desafios sistêmicos, que aparentam mitigar críticas diretas ao Governo Bolsonaro.

Os resultados destacam a natureza polarizada do discurso, com ambos os grupos utilizando estratégias lexicais e temáticas distintas para sustentar suas narrativas. Esses achados estão alinhados com padrões mais amplos de polarização política no Brasil, ressaltando como a linguagem reflete e reforça divisões ideológicas (SOUSA; FILHO; LARRAIN, 2024; ESCOSTEGUY-MEDRONHO, 2024; BüLOW; ABERS, 2024).

Ao identificar essas diferenças discursivas, esse trabalho contribui para uma compreensão mais detalhada da comunicação política, facilitando a construção de modelos que possam prever alinhamentos ideológicos ou sentimentos com base em dados textuais,

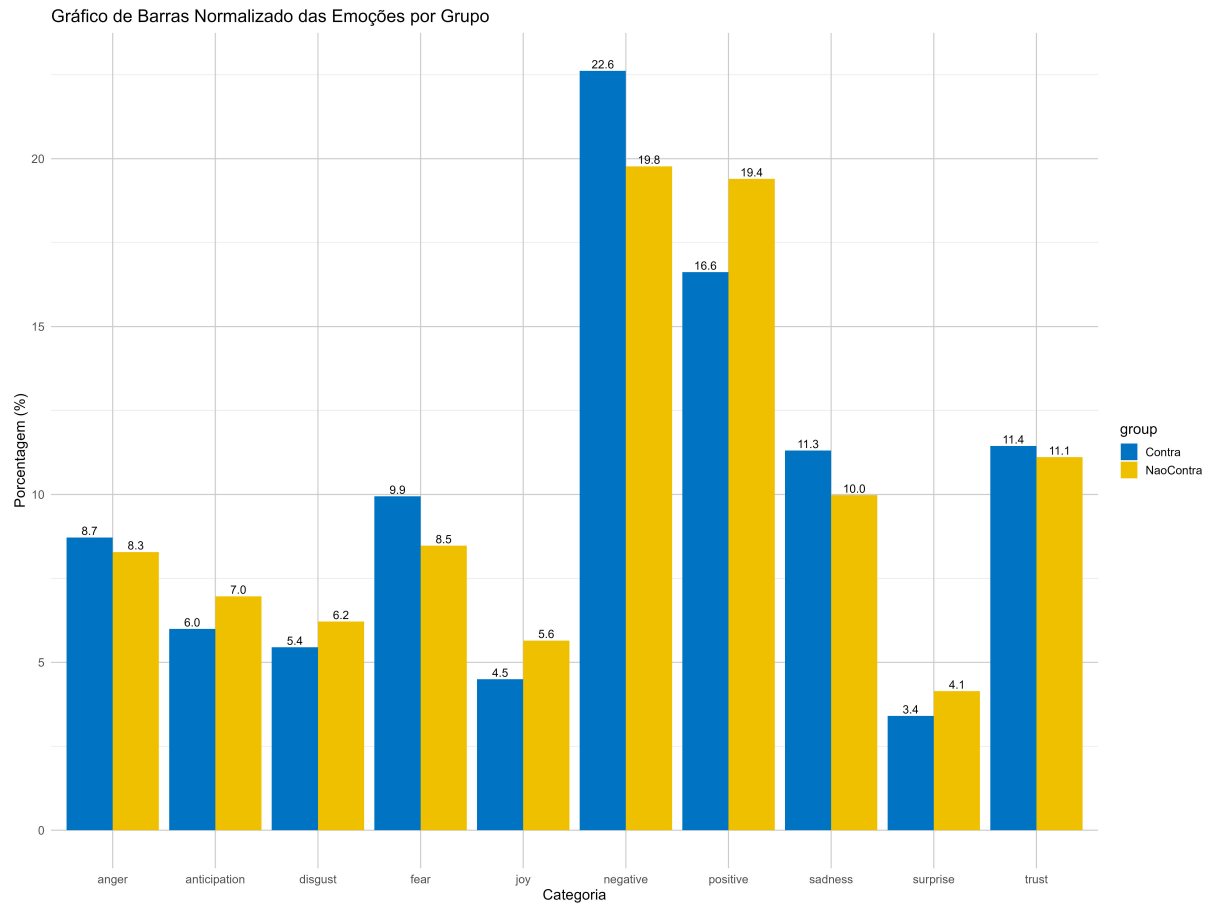


Figura 8 – Análise de Sentimentos

Fonte: Elaboração própria.

que servirá para etapas subsequentes de modelagem, como classificação de sentimentos, com o objetivo de capturar as estruturas presentes no discurso polarizado.

A análise de sentimentos tornou-se uma ferramenta central na pesquisa de comportamento político, oferecendo inferências sobre opinião pública, resultados eleitorais e discussões sobre políticas públicas (BESTVATER; MONROE, 2023; RITA; ANTÔNIO; AFONSO, 2023). A Figura 8 apresenta a distribuição normalizada das emoções entre os dois grupos, “contra” e “naocontra”, obtida utilizando o pacote *syuzhet* no R. O objetivo dessa análise é compreender o tom emocional presente nos textos, especialmente no contexto da polarização política em torno da figura de Bolsonaro.

O grupo “contra” apresenta níveis ligeiramente mais altos de raiva (8,7%) em comparação ao grupo “naocontra” (8,3%). Isso está alinhado com a postura oposicionista do grupo “contra”, onde a raiva provavelmente é direcionada a falhas percebidas ou insatisfação com a liderança. De forma similar, os textos do grupo “contra” mostram uma porcentagem marginalmente maior de nojo (7,0%) em relação ao grupo “naocontra” (5,4%), destacando o caráter crítico e orientado para a rejeição desse discurso. O medo

é ligeiramente mais prevalente no grupo “contra” (9,9%) do que no “naocontra” (8,5%). A tristeza é mais proeminente nos textos do grupo “contra” (11,3%) do que no grupo “naocontra” (10%), possivelmente indicando decepção em relação aos resultados do Governo. Essas diferenças nas emoções negativas refletem as nuances discursivas de cada grupo e reforçam a importância do tom emocional no contexto da polarização política.

O grupo “naocontra” apresenta mais expressões de alegria (5,6%) em comparação ao grupo “contra” (4,5%). Isso indica que narrativas de apoio tendem a incluir um enquadramento emocional mais positivo, possivelmente refletindo satisfação com a liderança ou otimismo em relação ao futuro. Esse padrão de emoção positiva ressalta como os grupos estruturam seus discursos de maneiras contrastantes, utilizando sentimentos positivos para reforçar suas narrativas e legitimar suas perspectivas.

No que tange a polaridade geral, o grupo “contra” apresenta uma porcentagem maior de sentimento negativo geral (22,6%) em comparação ao grupo “naocontra” (19,8%). Isso reforça a postura crítica do grupo “contra” em relação à liderança de Bolsonaro. Por outro lado, o grupo “naocontra” exibe uma porcentagem maior de sentimento positivo (19,4%) do que o grupo “contra” (16,6%), refletindo uma narrativa mais favorável e de apoio à liderança de Bolsonaro.

Os resultados destacam contextos distintos entre os grupos “contra” e “naocontra”, refletindo suas diferenças ideológicas e discursivas. O grupo “contra” é caracterizado por um tom mais negativo e emocionalmente carregado, alinhado à sua postura crítica. Em contraste, o grupo “naocontra” apresenta um perfil emocional relativamente equilibrado, com maior prevalência de emoções positivas, indicando uma postura de apoio ou defesa.

Essa divergência emocional ressalta o papel dos sentimentos na formação do discurso político para análises preditivas subsequentes. Ao incorporar esses padrões emocionais, futuras análises podem aprofundar a compreensão da polarização ideológica e de seus fundamentos emocionais, contribuindo para a construção de modelos mais sofisticados do comportamento político e do discurso.

Por último, o modelo de escalonamento *Wordfish*(SLAPIN; PROKSCH, 2008) é uma ferramenta utilizada em metodologias de análise de texto como dados (*text-as-data*), permitindo posicionar documentos em uma escala ideológica. Esse método estima as posições ideológicas relativas dos documentos com base nas frequências de palavras, tornando-o particularmente útil para a análise de discursos políticos.

Nesta análise na figura 9, o modelo *Wordfish* foi aplicado aos dois *corpora*: “contra” e “naocontra”, que representam posições contrastantes em relação à liderança de Bolsonaro. Os valores estimados de *theta* no eixo x representam as posições ideológicas relativas inferidas pelos padrões de uso das palavras, conforme exposto na figura.

A análise de escalonamento do *Wordfish* está perfeitamente alinhada com as infe-

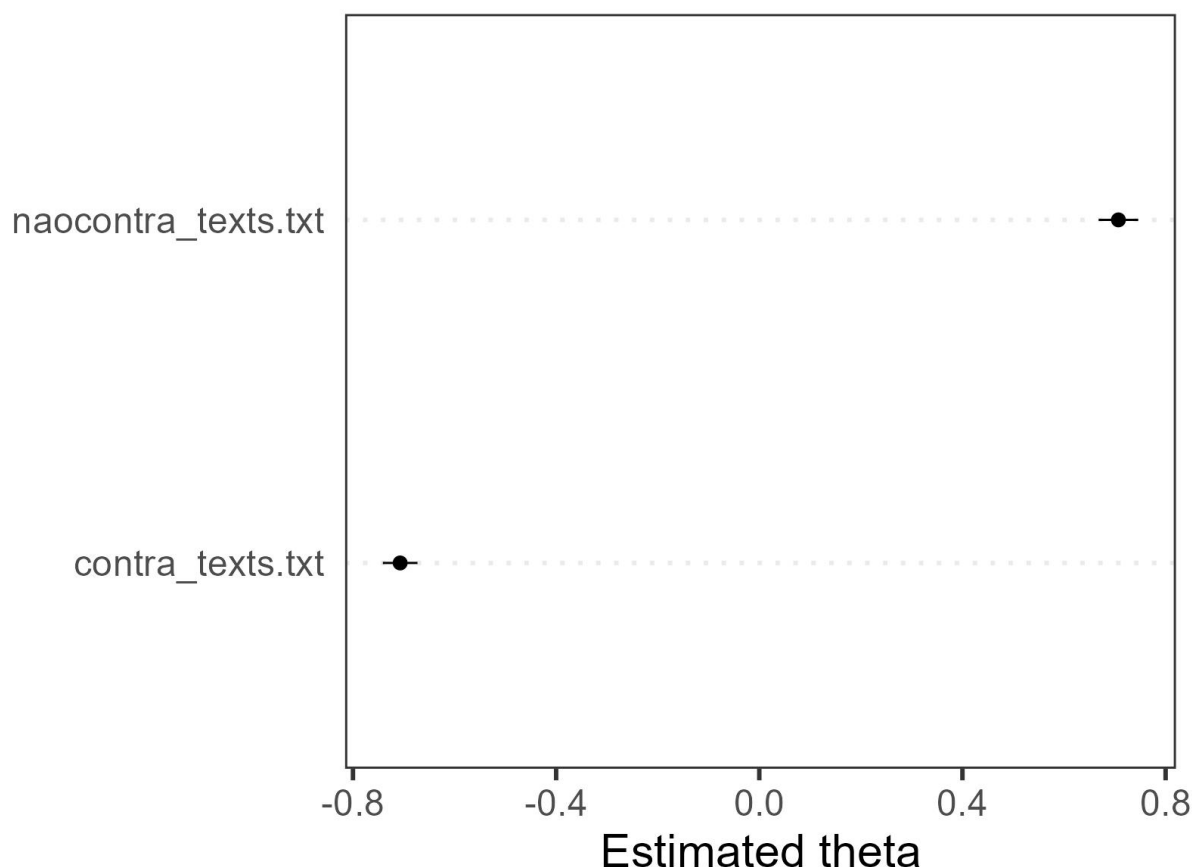


Figura 9 – Gráfico “wordfish” de Posição Relativa

Fonte: Elaboração própria.

rências obtidas nas análises lexicais, de sentimentos e temáticas anteriores, proporcionando uma compreensão mais ampla da polarização ideológica presente no conjunto de dados. Em conjunto, essa análise reforça a interpretação das estruturas linguísticas e emocionais nos *corpora* “contra” e “naocontra” anteriores.

Como destacado anteriormente, na análise de palavras-chave (*keyness*). O *corpus* “contra” consistentemente enfatizou termos como “postura”, “problema” e “criminoso”, refletindo um foco em críticas e questões de governança. Por outro lado, o *corpus* “naocontra” priorizou termos como “culpa”, “governo” e “Bolsonaro”, sugerindo estratégias de defesa ou reestruturação narrativa. Esses padrões lexicais fornecem a base para a capacidade do modelo *Wordfish* de separar ideologicamente os dois *corpora*. O uso frequente de termos carregados ideologicamente garante que o escalonamento reflita a polarização textual visto na figura.

Além disso, os resultados do modelo *Wordfish* estão alinhados com os achados da análise de sentimentos realizada com o pacote *syuzhet*. Os textos do grupo “contra” apresentaram níveis mais altos de emoções negativas, como raiva, medo e tristeza,

consistentes com sua posição crítica no extremo negativo da escala. Por outro lado, os textos do grupo “naocontra” exibiram mais emoções positivas, como alegria e confiança, correlacionando-se com sua postura ideológica de apoio. A análise de sentimento fornece uma camada adicional de validação para os resultados do escalonamento, demonstrando que sentimento e ideologia estão interligados nos dados textuais. Isso reforça a ideia de que as emoções desempenham um papel crucial na construção e expressão das posições ideológicas.

Quando integradas, as análises utilizadas por esse trabalho oferecem uma visão da polarização ideológica no discurso configurando padrões de comportamento político específicos de seus nichos. Os textos do grupo “contra” representam uma narrativa crítica e emocionalmente carregada, fundamentada em falhas de governança e insatisfação, enquanto os textos do grupo “naocontra” refletem uma perspectiva defensiva e alinhada ideologicamente, marcada por confiança e justificativas.

Essa abordagem integrada demonstra como métodos quantitativos de análise de texto podem se complementar para revelar dinâmicas ideológicas e emocionais mais sutis. A convergência desses achados fornece evidências robustas da polarização no discurso político, destacando a inter-relação entre linguagem, sentimentos e ideologia.

Essa convergência de resultados não apenas reforça a validade das análises exploratórias, mas também destaca a viabilidade de desenvolver um modelo de aprendizagem profunda. O *framework* de aprendizagem profunda permite desenvolver um modelo preditivo onde os padrões sutis podem ser identificados, como marcadores lexicais, tons emocionais e posicionamentos ideológicos, permitindo previsões robustas de comportamentos políticos e classificações de novos textos.

Com base nos resultados obtidos dessa análise integrada, o modelo de aprendizagem profunda terá uma fundação teórica e empiricamente fundamentada. Essa abordagem garante que o modelo seja metodologicamente sólido e capaz de aproveitar as características linguísticas e ideológicas presentes nos *corpora* para desenvolver um modelo preditivo. Além disso, essa estratégia contribuirá para o avanço no estudo do comportamento político, oferecendo uma ferramenta para compreender, analisar e prever dinâmicas complexas e não lineares.

3.2 BERT-Vi

A crescente disponibilidade de dados em estudos políticos não só criou uma janela de oportunidade para o desenvolvimento de modelos de aprendizado de máquina e aprendizagem profunda, mas também surge como uma necessidade para analisar e classificar a complexidade destes. Nesta seção, serão descritos as etapas de pré-processamento do conjunto de dados, a arquitetura do modelo, o desenvolvimento, treinamento e avaliação

de uma arquitetura de rede neural projetada para classificar dados textuais dos banco de dados exposto. Para esse modelo, será utilizada uma abordagem de aprendizado por transferência com o modelo XLM-RoBERTa, ajustado em um conjunto de dados ampliado, aproveitando os avanços recentes em arquiteturas baseadas em *transformers*.

As redes neurais são modelos computacionais inspirados nas estruturas neurais biológicas do cérebro humano. Elas consistem em camadas de nós interconectados (ou “neurônios”) que aprendem a reconhecer padrões complexos nos dados (LECUN; BENGIO; HINTON, 2015; SCHMIDHUBER, 2015; GOODFELLOW; BENGIO; COURVILLE, 2016).

Uma rede neural geralmente é composta por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Cada neurônio possui parâmetros associados (pesos e *biases*) que ajustam a maneira como os sinais de entrada são processados. Na etapa de passagem direta (*forward pass*), os dados percorrem a rede, camada por camada, gerando previsões na saída. Após calcular uma previsão, a perda (ou erro) é determinada em relação aos rótulos reais, em seguida essa perda é então propagada de volta pela rede, permitindo que os pesos sejam ajustados para minimizar erros futuros (retropropagação ou *backpropagation*) (RUMELHART; HINTON; WILLIAMS, 1986).

Nos neurônios, as funções de ativação não lineares (e.g., ReLU, Sigmoid, Tanh) são um componente crucial das redes neurais, permitindo que elas aprendam e representem funções que mapeiam entradas para saídas. Sem essas ativações não lineares, uma rede composta apenas de transformações lineares (multiplicações de matrizes) estaria limitada a aprender mapeamentos lineares, independentemente do número de camadas.

Outra característica importante são os algoritmos de otimização (SGD, Adam, AdamW) que ajustam os pesos da rede com base nos gradientes calculados, com o objetivo de reduzir gradualmente a função de perda (loss).

Ao longo da última década, as redes neurais se destacaram em tarefas como classificação de imagens, reconhecimento de fala e processamento de linguagem natural (NLP). No campo do NLP, os grandes modelos de linguagem pré-treinados tornaram-se o padrão predominante (HOWARD; RUDER, 2018; DEVLIN et al., 2019; RAFFEL et al., 2020).

Tradicionalmente, a construção de um modelo de NLP exigia uma grande quantidade de dados rotulados. No entanto, o aprendizado por transferência transformou radicalmente esse paradigma. O aprendizado por transferência (PAN; YANG, 2010) consiste em utilizar um modelo previamente treinado em um grande conjunto de dados (geralmente com tarefas auto-supervisionadas) e, em seguida, ajustá-lo (*fine-tuning*) em um conjunto de dados menor e específico para a tarefa.

Modelos como BERT (DEVLIN et al., 2019), RoBERTa (LIU et al., 2019), GPT (RADFORD et al., 2018), XLM (LAMPLE; CONNEAU, 2019) e XLM-RoBERTa (CON-

NEAU et al., 2020) são exemplos de grandes modelos de linguagem treinados em corpora de texto em diversas línguas. Esses modelos aprendem *embeddings* contextuais de palavras, capturando características semânticas e sintáticas que se generalizam entre diferentes tarefas.

O conjunto de dados utilizado para o treinamento e avaliação foi derivado de um *corpus* pré-processado de amostras textuais anotadas quanto ao alinhamento político (“contra” ou “naocontra”). Conforme descrito na seção banco de dados, foi aplicada ampliação de dados (*data augmentation*) para aumentar a diversidade e o tamanho do conjunto de dados. Foram utilizados dois conjuntos, um com aproximadamente 1500 entradas e o outro com 6000 entradas balanceadas.

Para preparar os dados para entrada no modelo, foi realizada a tokenização utilizando o tokenizer pré-treinado do Bertimbau(SOUZA; NOGUEIRA; LOTUFO, 2020) e do XLM-RoBERTa, conforme detalhado por Conneau et al. 2019. A tokenização garante que o modelo consiga lidar com textos multilíngues, preservando informações essenciais para a tarefa de classificação proposta.

Foram duas arquiteturas testadas: XLM-RoBERTa e Bertimbau. Ambos modelos transformers, sendo, o primeiro, multilíngue e, o segundo, português brasileiro. Esses modelos seguem princípios de aprendizado por transferência, permitindo a adaptação de um modelo pré-treinado para tarefas específicas (HOWARD; RUDER, 2018), como será exposto a seguir.

As amostras de texto foram tokenizadas utilizando o *AutoTokenizer* da biblioteca *Transformers* da HuggingFace(WOLF et al., 2020), projetado para processar entradas compatíveis com os modelos utilizados. Esse processo converte o texto em tokens, enquanto realiza o truncamento ou preenchimento (*padding*) das sequências para um comprimento máximo fixo, garantindo um processamento em lotes (*batch processing*).

Para o ajuste fino, uma camada de *dropout* com uma taxa de 0,3 é utilizada para mitigar o sobreajuste (*overfitting*) (SRIVASTAVA et al., 2014) e uma camada linear mapeia os *embeddings* para as duas classes-alvo. Além disso, a maior parte das camadas dos modelos permanecem congeladas. Apenas as últimas 4 camadas são descongeladas, permitindo ajustes específicos para a tarefa, enquanto retêm o conhecimento linguístico de uso geral inerentes aos modelos.

No que tange o treinamento e otimização, a função de perda *cross-entropy* também foi utilizada para otimizar o modelo para classificação binária, garantindo um equilíbrio entre penalizar previsões incorretas e promover as corretas. O otimizador AdamW (LOSHCHILOV; HUTTER, 2019) foi empregado para ajustar os pesos do modelo, incorporando *weight decay* para melhorar a generalização. Um agendador de taxa de aprendizado linear com etapas de warm-up foi usado para garantir a convergência durante o treinamento.

Outra característica importante da rede neural desenvolvida foi a técnica de *early stopping* implementada para interromper o treinamento quando o desempenho na validação estabilizou, reduzindo o risco de *overfitting* e otimizando o tempo de treinamento.

Foi realizada uma validação cruzada estratificada com 5 *folds* para avaliar a robustez do modelo. A estratificação garantiu a representação proporcional de ambas as classes nos subconjuntos de treinamento e validação, minimizando vieses (KOHAVI, 1995). Cada *fold* foi composto por: a) treinamento: 80% dos dados; b) validação: 10% dos dados; c) teste: 10% dos dados. Isso assegurou uma avaliação abrangente em todas as amostras.

Para as métricas de avaliação, foi utilizado o F1-Score (Ponderado) porque ele captura o equilíbrio entre precisão (*precision*) e revocação (*recall*) (LIPTON; ELKAN; NARAYANASWAMY, 2014). Foi utilizada também a acurácia (*Accuracy*) porquanto ela mede a proporção de previsões corretas em relação ao total de amostras, oferecendo uma métrica direta do desempenho geral do modelo. Essa abordagem garante uma avaliação do modelo considerando a variação entre diferentes subconjuntos de dados.

Essa seção demonstrou a arquitetura da rede neural desenvolvida por esse trabalho (chamada de BERT-Vi⁶) para classificar textos politicamente polarizados utilizando o modelo XLM-RoBERTa e o BERTIMBAU. Ao integrar aprendizado por transferência, validação cruzada e estratégias otimização, o modelo alcançou alto desempenho, mantendo a capacidade de generalização. Os resultados, expostos na próxima sessão, validam o uso de arquiteturas baseadas em *transformers* para tarefas de classificação de texto complexas no campo da ciência social computacional.

3.3 Resultados

Os dados da validação cruzada dos modelos estão resumidos na Tabela 6. Como esperado, cada configuração apresenta características distintas em termos de acurácia, F1-score e capacidade de generalização.

O modelo BERTimbau, treinado com um conjunto de dados de 1500 entradas, utilizando um tamanho de lote de 8 e uma taxa de aprendizado de 3e-5, alcançou um F1-score médio de 0,8774 ($\pm 0,0214$) e uma acurácia média de 0,8777 ($\pm 0,0211$). Esses resultados indicam um desempenho robusto e consistente, com baixa variabilidade entre os *folds* , demonstrando forte capacidade de generalização, especialmente em conjuntos de dados menores.

Em comparação, o modelo RoBERTa, treinado sob a mesma configuração de 1500 entradas e tamanho de lote de 8, apresentou um F1-score médio de 0,8334 ($\pm 0,0167$) e uma acurácia média de 0,8336 ($\pm 0,0165$). Embora o desempenho tenha sido consistente, foi

⁶ Anexos a esse trabalho estão no [Github](#) e no [Hugging Face](#) do autor

ligeiramente inferior ao BERTimbau, sugerindo que ajustes nos hiperparâmetros poderiam melhorar os resultados.

Tabela 6 – Desempenho de Validação Cruzada de Diferentes Modelos e Configurações

Modelo	Banco de Dados	Lote	F1-Score ($\pm$ DP)	Acurácia ($\pm$ DP)
BERTimbau	1500	8	0.8774 ($\pm$ 0.0214)	0.8777 ($\pm$ 0.0211)
Roberta	1500	8	0.8334 ($\pm$ 0.0167)	0.8336 ($\pm$ 0.0165)
Roberta	1500	16	0.7822 ($\pm$ 0.0185)	0.7827 ($\pm$ 0.0188)
Roberta	6000	8	0.9137 ($\pm$ 0.0125)	0.9138 ($\pm$ 0.0125)
Roberta	6000	16	0.9067 ($\pm$ 0.0080)	0.9068 ($\pm$ 0.0080)

Fonte: Elaboração própria.

Ao aumentar o tamanho do lote para 16, o RoBERTa, ainda com 1500 entradas, apresentou um desempenho mais modesto, com um F1-score médio de 0,7822 ($\pm$ 0,0185) e uma acurácia média de 0,7827 ($\pm$ 0,0188). Apesar de mostrar consistência entre os *folds*, a queda nas métricas gerais indica que o aumento no tamanho do lote pode não ser ideal para conjuntos menores.

Por outro lado, quando o RoBERTa foi treinado com um conjunto de dados ampliado para 6000 entradas e utilizando um tamanho de lote de 8, os resultados foram significativamente melhores, com um F1-score médio de 0,9137 ($\pm$ 0,0125) e uma acurácia média de 0,9138 ($\pm$ 0,0125). Esse desempenho excepcional destaca a influência positiva de conjuntos de dados maiores na generalização e na capacidade de classificação do modelo. Mesmo com o tamanho do lote aumentado para 16, o RoBERTa treinado com 6000 entradas apresentou resultados excelentes, alcançando um F1-score médio de 0,9067 ($\pm$ 0,0080) e uma acurácia média de 0,9068 ($\pm$ 0,0080), com variabilidade mínima entre os folds. Apesar de um pequeno declínio em relação ao lote menor, a configuração demonstrou ser altamente eficiente.

Esses resultados indicam que conjuntos de dados maiores contribuem significativamente para a melhora do desempenho, especialmente em modelos como o RoBERTa, que se beneficiam de entradas amplas e diversificadas. O BERTimbau, por outro lado, teve uma performance proporcionalmente melhor considerando o tamanho do conjunto de dados se destacando em cenários com conjuntos de dados menores, demonstrando ser uma opção melhor em contextos onde a quantidade de dados disponíveis é limitada. Ajustes no tamanho do lote e em outros hiperparâmetros podem ainda otimizar o desempenho, dependendo do tamanho do conjunto de dados e da arquitetura do modelo.

A aplicação do modelo pré-treinado BERT-Vi foi conduzida para prever rótulos no conjunto de dados base contendo as 600 entradas originais antes do *data augmentation*. Esse modelo, previamente ajustado (*fine-tuned*) em um conjunto de dados específico, foi utilizado sem qualquer ampliação ou modificação adicional nesta fase, com foco exclusivo

na inferência.

O conjunto utilizado para a predição contém 600 amostras textuais originais equilibradas entre duas classes: 300 amostras para a classe 0 (“contra”) e 300 amostras para a classe 1 (“naocontra”). Cada entrada de texto foi tokenizada utilizando o mesmo *tokenizer* empregado durante o treinamento com os mesmo parâmetros. Os dados processados foram processados em uma classe de conjunto de dados personalizada, projetada para gerar entradas tokenizadas em um formato compatível com o *PyTorch* (PASZKE et al., 2019).

O conjunto de dados foi então inserido em um *DataLoader*, permitindo o processamento em lotes (batch size) de 16, priorizando a eficiência e otimização do hardware durante a fase de inferência.

Tabela 7 – Parâmetros e Fluxo de Trabalho para Inferência com Bert-Vi

Parâmetro/Etapa	Valor/Descrição
Nome do Modelo	xlm-roberta-large
Tamanho do Conjunto de Dados	600 entradas
Comprimento Máximo da Sequência	360
Tamanho do Lote	16
Caminho do Checkpoint Pré-Treinado	best_model_fold4.pth
Tokenizador	AutoTokenizer
Camadas Descongeladas	Últimas 4 camadas
Formato de Saída da Predição	CSV com rótulos previstos
Dispositivo	GPU (se disponível)

Fonte: Elaboração própria.

Ao aplicar o modelo pré-treinado BERT-Vi, as predições resultaram em 295 amostras para a classe 0 e 305 para a classe 1. Esse leve desequilíbrio nas predições (aproximadamente 49,17% para a classe 0 e 50,83% para a classe 1) demonstra que o modelo foi capaz de generalizar bem, embora com uma ligeira inclinação para a classe 1.

As predições estão alinhadas de forma próxima à distribuição real do conjunto de dados, indicando que o modelo BERT-Vi capturou os padrões aprendidos durante as fases de pré-treinamento e ajuste fino. No entanto, os casos de classificações incorretas devem ser analisados para identificar padrões onde o modelo possa ter dificuldade em diferenciar entre as classes “contra” e “naocontra”.

Tabela 8 – Tabela Verdade das Predições do Modelo

Rótulo Real	Predito: Classe 0 (“naocontra”)	Predito: Classe 1 (“contra”)
Classe 0 (“naocontra”)	295 (VN)	5 (FP)
Classe 1 (“contra”)	5 (FN)	295 (VP)

Fonte: Elaboração própria.

A aplicação do modelo BERT-Vi ao banco de dados original ajuda a demonstrar sua eficácia em generalizar e prever dados não vistos de maneira a gerar confiança no desempenho do modelo na tarefa que ele foi treinado pra fazer. Contudo, é fundamental realizar análises adicionais das classificações incorretas e melhorar ainda mais a precisão das predições.

3.4 Conclusão

Esse trabalho, utilizando técnicas computacionais de análise quantitativa de texto e NNs, tomou para si o desafio de enfrentar e compreender os nuances do comportamento político polarizado, refletindo a crescente relevância dessas metodologias na pesquisa em ciência política. O uso do modelo de aprendizagem profunda, particularmente arquiteturas baseadas em *transformers*, demonstrou potencial no processamento de dados não estruturados e na identificação de padrões, ampliando as possibilidades de análise de textos de comportamento político.

Para isso foram implementados e avaliados modelos baseados em *transformers* BERTimbau e XLM-RoBERTa com diferentes versões dos conjuntos de dados, configurações e tarefas. Os principais resultados incluem a) modelos treinados em conjuntos menores (1500 amostras) demonstraram desempenho moderado, com o BERTimbau alcançando um F1-score médio de 0,8774 ($\pm 0,0214$); b) conjuntos maiores (6000 amostras) melhoraram significativamente o desempenho dos modelos, com o XLM-RoBERTa obtendo um F1-score médio de 0,9137 ($\pm 0,0125$) em configurações específicas; c) a variabilidade entre os *folds* permaneceu baixa, indicando generalização consistente nos diferentes subconjuntos de dados.

O modelo desenvolvido BERT-Vi foi aplicado ao conjunto base de 600 amostras de texto, resultando em predições equilibradas (295 para a classe “contra” e 305 para a classe “não contra”). Demonstrando a capacidade do modelo de predição de dados e sua utilidade em aplicações práticas.

Baseando-se nos princípios da Teoria da Representação, o estudo utilizou *embeddings* e técnicas de extração de características para estruturar dados políticos. Essa abordagem alinhou os *embeddings* textuais com dimensões ideológicas e temáticas, aumentando a interpretabilidade e a capacidade preditiva.

Esse trabalho contribui tanto para a ciência política quanto para os sistemas de informação, combinando rigor teórico com inovação metodológica. A integração da Teoria da Representação ao processo de modelagem destaca sua relevância para estruturar dados políticos em redes neurais, preservando dimensões ideológicas e sociais, tornando os modelos tanto preditivos quanto interpretáveis. Do ponto de vista da ciência política, a pesquisa demonstra como técnicas computacionais avançadas podem ser aplicadas para

compreender e modelar comportamentos complexos. A capacidade de prever resultados eleitorais ou mapear a polarização política usando dados não estruturados representa um avanço metodológico significativo em relação às abordagens tradicionais.

Embora esse estudo ressalte o potencial das redes neurais na captura e análise de fenômenos políticos, algumas limitações, como a dependência de conjuntos de dados balanceados e a necessidade de ajustes extensivos de hiperparâmetros, apontam caminhos para futuras pesquisas. Técnicas de ampliação de dados, abordagem de desbalanceamentos de classes e exploração de conjuntos de dados multilíngues poderiam ampliar ainda mais a aplicabilidade desses métodos.

Por fim, destaca-se a importância de abordagens interdisciplinares, combinando aprendizado de máquina, ciência política e sistemas de informação para enfrentar os desafios em um cenário político polarizado e em evolução. Ao avançar o conjunto de ferramentas metodológicas disponíveis para os pesquisadores, este estudo busca estabelecer as bases para análises sobre os mecanismos que moldam o comportamento e a tomada de decisão política.

Para trabalhos futuros e questões em aberto: i) Interpretação de emoções, opiniões e tendências em plataformas de mídia social para mapear a polarização política, monitorar reações públicas e prever mudanças no comportamento eleitoral; b) Categorização de documentos políticos, como manifestos partidários, para identificar prioridades temáticas e fundamentos ideológicos; c) análise de grandes volumes de texto para detectar padrões latentes que revelem ideologias políticas, narrativas dominantes e mudanças estratégicas no discurso.

Portanto, redes neurais empregadas para prever resultados eleitorais e analisar o comportamento dos eleitores, sua capacidade de lidar com relações não lineares em fenômenos sociais as torna ferramentas eficazes para análises preditivas. Por exemplo, essas redes podem integrar diversos conjuntos de dados, como histórico de votos, variáveis socioeconômicas, pesquisas de opinião e tendências de mídia social, para prever com precisão o desempenho eleitoral. Isso permite compreender como os eleitores respondem a eventos e como seus comportamentos evoluem ao longo do tempo.

4 Conclusões e Considerações Finais

Neste trabalho - “Em Busca do Algoritmo Bolsonarista: Aprendizado de Máquina e aprendizagem profunda aplicados à Ciência Política” -, analisou-se a militância do “Bolsonarismo” utilizando *clustering* de *machine learning*, regressões lineares e logísticas, além de aprendizagem profunda com redes baseadas em *transformers*. Essa abordagem possibilita uma compreensão detalhada do comportamento político e destaca o potencial transformador desses métodos na ciência política computacional.

No primeiro capítulo foi demonstrado que o “antipetismo” é um movimento autônomo e independente do “Bolsonarismo”. Isso fica claro com o resultado das eleições regionais de 2024. As eleições regionais de 2024 quebraram o paradigma das polarizadas eleições de 2022. Diferente destas, os eleitores orbitaram partidos de direita, centro direita e centro, indo de encontro com a dicotomia Partido dos Trabalhadores (PT) e Partido Liberal (PL) que ocorreu na última eleição presidencial.

No Rio de Janeiro, Eduardo Paes (Partido Social Democrático - PSD) conseguiu unir a maior parte dos eleitores ganhando no primeiro turno contra o candidato Alexandre Ramagem (PL). Até mesmo a candidatura à esquerda de Tarcísio Motta (Partido Socialismo e Liberdade - PSOL) ficou inexpressiva frente a candidatura ideologicamente neutra de Eduardo Paes.

Em São Paulo, no primeiro turno, o eleitor paulista se pulverizou quase homogeneamente entre três candidatos: Ricardo Nunes (Movimento Democrático Brasileiro - MDB), Guilherme Boulos (PSOL), e Pablo Marçal (Partido Renovador Trabalhista Brasileiro - PRTB). O resultado mostra um eleitor à direita pulverizado entre uma direita (supostamente) moderada representada pelo candidato do MDB e outra direita extremista com características populistas e antissistêmica representada pelo candidato do PRTB.

Semelhante ao caso paulista, mas sem um candidato à esquerda, o eleitor de BH ficou pulverizado entre quatro candidaturas à direita e centro direita: Bruno Engler (PL), Fuad Noman (PSD), Mauro Tramonte (Republicanos) e Gabriel Azevedo (MDB). Os candidatos do PL e do Republicanos se alinham com uma direita mais demarcada ideologicamente num espectro extremista, enquanto os candidatos do PSD e MDB, disputavam os eleitores mais moderados tentando apelar também aos eleitores de esquerda.

Portanto, como demonstrado no capítulo “Navalha de Ockham do Bolsonarismo”, a “direita não-bolsonarista” é autônoma e independente da “direita bolsonarista”, e essa desvinculação - seja oportuna ou permanente - ficou latente nas eleições de 2024.

No capítulo “Redes Neurais e Polarização Política - Modelo BERT-Vi”, ficou

demonstrada, a partir da análise de entrevistas de profundidades em grupos focais, a natureza polarizada do discurso, com ambos os grupos utilizando estratégias lexicais e temáticas distintas para sustentar suas narrativas. Com essa dissociação foi possível treinar um modelo de rede neural - BERT-Vi - para prever comportamento político a favor ou não do governo Bolsonaro.

A capacidade de usar NLPs e LLMs encontram terreno fértil na Ciência Política pois com eles é possível analisar grandes volumes de dados e encontrar nuances ainda mais sutis na diferenças de comportamentos políticos. Assim, analisar entrevistas de profundidade se torna uma tarefa menos custosa, podendo aumentar ainda mais a quantidade de dados de entrada, expandindo a Ciência Política para além das análises de dados tabulares e da análise qualitativa de textos.

Na síntese do necessário, os resultados obtidos não apenas confirmaram a eficácia dessas abordagens em tarefas de classificação, mas também contribuíram para o avanço metodológico no campo da ciência política computacional. Além de oferecer uma base sólida para futuras investigações, este trabalho reforça a relevância do uso de ferramentas tecnológicas no entendimento de dinâmicas políticas contemporâneas, propondo novas possibilidades para a análise de dados em cenários marcados por complexidade social e ideológica. Espera-se que os achados apresentados inspirem estudos futuros que aprofundem a interseção entre inteligência artificial e ciências sociais, promovendo uma compreensão mais rica e abrangente dos desafios políticos do século XXI. Assim, esta pesquisa não apenas contribui para a literatura acadêmica, mas também propõe soluções práticas e inovadoras para desafios contemporâneos, alinhando-se ao esforço contínuo de integrar ciência, tecnologia e impacto social.

Referências

- AGRESTI, A. *Categorical Data Analysis*. 3rd. ed. Wiley, 2013. ISBN 978-0470463635. Disponível em: <<https://onlinelibrary.wiley.com/book/10.1002/9781118712931>>. Citado na página 39.
- ALTMAN, D. G.; BLAND, J. M. Diagnostic tests 1: sensitivity and specificity. *BMJ*, v. 308, n. 6943, p. 1552, 1994. Citado na página 41.
- AMARAL, O. E. d. The victory of jair bolsonaro according to the brazilian electoral study of 2018. *Brazilian Political Science Review*, v. 14, n. 1, p. e0004, 2020. Disponível em: <<http://doi.org/10.1590/1981-3821202000010004>>. Citado 2 vezes nas páginas 31 e 32.
- ANTAL, B. B. et al. Achieving occam’s razor: Deep learning for optimal model reduction. *PLOS Computational Biology*, Version 2, July 18 2024. Disponível em: <<https://doi.org/10.1371/journal.pcbi.1012283>>. Citado 2 vezes nas páginas 29 e 30.
- ARMENTA, M.; JODOIN, P.-M. The representation theory of neural networks. *Mathematics*, v. 9, n. 24, 2021. ISSN 2227-7390. Disponível em: <<https://www.mdpi.com/2227-7390/9/24/3216>>. Citado na página 49.
- BENGIO, Y.; COURVILLE, A.; VINCENT, P. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 35, n. 8, p. 1798–1828, 2013. Citado na página 49.
- BENOIT, K. et al. quanteda: An r package for the quantitative analysis of textual data. *Journal of Open Source Software*, v. 3, n. 30, p. 774, 2018. Disponível em: <<https://quanteda.io>>. Citado 2 vezes nas páginas 26 e 53.
- BESTVATER, S. E.; MONROE, B. L. Sentiment is not stance: Target-aware opinion classification for political text analysis. *Political Analysis*, v. 31, n. 2, p. 235–256, 2023. Citado 2 vezes nas páginas 51 e 57.
- BEZDEK, J. C. *Pattern Recognition with Fuzzy Objective Function Algorithms*. [S.l.]: Plenum Press, New York, 1981. Citado na página 33.
- BISHOP, C. M. *Pattern Recognition and Machine Learning*. [S.l.]: Springer, 2006. Citado na página 42.
- BREIMAN, L. Statistical modeling: The two cultures. statistical science. *Statist. Sci.* 16(3): 199–231, 2001. Citado na página 25.
- BRODERSEN, K. H. et al. The balanced accuracy and its posterior distribution. In: *2010 20th International Conference on Pattern Recognition*. [S.l.]: IEEE, 2010. p. 3121–3124. Citado na página 42.
- BURNHAM, M. Stance detection: a practical guide to classifying political beliefs in text. *Political Science Research and Methods*, p. 1–18, 2024. Citado na página 51.

BURTON-JONES, A. et al. Assessing representation theory with a framework for pursuing success and failure. *MIS Quarterly: Management Information Systems*, University of Minnesota Press (Project Muse), v. 41, n. 4, p. 1307–1333, dez. 2017. ISSN 0276-7783. Citado na página 53.

BüLOW, M. von; ABERS, R. N. Denialism and populism: Two sides of a coin in jair bolsonaro's brazil. *Government and Opposition*, v. 59, n. 4, p. 998–1016, 2024. Citado na página 56.

CHATSIOU, K.; MIKHAYLOV, S. J. Deep learning for political science. In: CURINI, L.; FRANZESE, R. J. (Ed.). *The SAGE Handbook of Research Methods in Political Science and International Relations*. Thousand Oaks, CA: SAGE Publications, 2020. Citado na página 48.

CONNEAU, A. et al. Unsupervised cross-lingual representation learning at scale. *CoRR*, abs/1911.02116, 2019. Disponível em: <<http://arxiv.org/abs/1911.02116>>. Citado na página 62.

CONNEAU, A. et al. Unsupervised cross-lingual representation learning at scale. *arXiv preprint arXiv:1911.02116*, 2020. Disponível em: <<https://arxiv.org/abs/1911.02116>>. Citado na página 62.

DEVLIN, J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*. [S.l.: s.n.], 2019. p. 4171–4186. Citado na página 61.

DHERIN, B.; MUNN, M.; BARRETT, D. G. T. The geometric occam's razor implicit in deep learning. *CoRR*, abs/2111.15090, 2021. Disponível em: <<https://arxiv.org/abs/2111.15090>>. Citado 2 vezes nas páginas 29 e 30.

DUNN, J. C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, Taylor & Francis, v. 3, n. 3, p. 32–57, 1973. Citado na página 33.

ESCOSTEGUY-MEDRONHO, A. Um brasil dividido? reflexões sobre a polarizaÇão política e social no brasil contemporâneo. *Lua Nova: Revista de Cultura e Política*, CEDEC, n. 122, p. e122041ae, 2024. ISSN 0102-6445. Disponível em: <<https://doi.org/10.1590/0102-001041ae/122>>. Citado 2 vezes nas páginas 23 e 56.

ESTRELLA, T.; BORBA, F. Legitimidade democrática: análise qualitativa da percepção do eleitorado no governo bolsonaro. *Revista Brasileira de Ciência Política*, Universidade de Brasília. Instituto de Ciência Política, v. 43, p. e268564, 2024. ISSN 0103-3352. Disponível em: <<https://doi.org/10.1590/0103-3352.2024.43.268564>>. Citado na página 50.

FAWCETT, T. An introduction to roc analysis. *Pattern Recognition Letters*, Elsevier, v. 27, n. 8, p. 861–874, 2006. Citado 2 vezes nas páginas 41 e 43.

GABRIELATOS, C. Keynes analysis: Nature, metrics and techniques. In: TAYLOR, C.; MARCHI, A. (Ed.). *Corpus Approaches To Discourse: A Critical Review*. Oxford: Routledge, 2018. cap. 11. Citado na página 54.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. MIT Press, 2016. Disponível em: <<https://www.deeplearningbook.org/>>. Citado na página 61.

GOPPA, V. D. Group representations and algebraic information theory. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, v. 59, n. 6, p. 25–50, 1995. Citado na página 26.

HAHN, J.; TODD, P.; KLAAUW, W. Van der. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, JSTOR, v. 69, n. 1, p. 201–209, 2001. Citado na página 33.

HAMED, I.; HABASH, N.; VU, N. T. *Data Augmentation Techniques for Machine Translation of Code-Switched Texts: A Comparative Study*. 2023. Disponível em: <<https://arxiv.org/abs/2310.15262>>. Citado na página 52.

HANLEY, J. A.; MCNEIL, B. J. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*, Radiological Society of North America, v. 143, n. 1, p. 29–36, 1982. Citado na página 43.

HART, D.; GREGOR, S. (Ed.). *Information Systems Foundations: Theory, Representation and Reality*. Canberra, Australia: ANU E Press, 2007. Disponível em: <<https://press.anu.edu.au/publications/series/information-systems-foundations/information-systems-foundations-theory>>. Citado na página 53.

HASELMAYER, M.; JENNY, M. Sentiment analysis of political communication: combining a dictionary approach with crowdcoding. *Quality & Quantity*, v. 51, p. 2623–2646, November 2017. Published: 21 September 2016. Disponível em: <<https://doi.org/10.1007/s11135-016-0412-4>>. Citado na página 51.

HOSMER, D. W.; LEMESHOW, S.; STURDIVANT, R. X. *Applied Logistic Regression*. 3rd. ed. Wiley, 2013. ISBN 978-0470582473. Disponível em: <<https://onlinelibrary.wiley.com/book/10.1002/9781118548387>>. Citado na página 39.

HOWARD, J.; RUDER, S. Universal language model fine-tuning for text classification. In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL)*. [S.l.: s.n.], 2018. p. 328–339. Citado 2 vezes nas páginas 61 e 62.

IMBENS, G.; LEMIEUX, T. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, Elsevier, v. 142, n. 2, p. 615–635, 2008. Citado na página 33.

IWANA, B. K.; UCHIDA, S. An empirical survey of data augmentation for time series classification with neural networks. *PLOS ONE*, Public Library of Science, v. 16, n. 7, p. 1–32, 07 2021. Disponível em: <<https://doi.org/10.1371/journal.pone.0254841>>. Citado na página 52.

JOCKERS, M. L. *Syuzhet: Extract Sentiment and Plot Arcs from Text*. [S.l.], 2015. Disponível em: <<https://github.com/mjockers/syuzhet>>. Citado 2 vezes nas páginas 26 e 53.

JORDAN M. I., . M. T. M. Machine learning: Trends, perspectives, and prospects. science. *Political Studies Review*, 2015. Citado na página 25.

KHAZE, S. R.; MASDARI, M.; HOJJATKHAH, S. Application of artificial neural networks in estimating participation in elections. *CoRR*, abs/1309.2183, 2013. Disponível em: <<http://arxiv.org/abs/1309.2183>>. Citado na página 48.

KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: MORGAN KAUFMANN PUBLISHERS INC. *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*. [S.l.], 1995. p. 1137–1143. Citado na página 63.

LAMPLE, G.; CONNEAU, A. Cross-lingual language model pretraining. In: *Advances in Neural Information Processing Systems (NeurIPS)*. [s.n.], 2019. v. 32. Disponível em: <<https://arxiv.org/abs/1901.07291>>. Citado na página 61.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, 2015. Citado na página 61.

LEE, D. S.; LEMIEUX, T. Regression discontinuity designs in economics. *Journal of Economic Literature*, American Economic Association, v. 48, n. 2, p. 281–355, 2010. Citado na página 33.

LIPTON, Z. C.; ELKAN, C.; NARAYANASWAMY, B. Optimal thresholding of classifiers to maximize f1 measure. In: ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. [S.l.], 2014. p. 1742–1753. Citado na página 63.

LIU, Y. et al. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019. Disponível em: <<https://arxiv.org/abs/1907.11692>>. Citado na página 61.

LOSHCHILOV, I.; HUTTER, F. Decoupled weight decay regularization. In: *International Conference on Learning Representations*. [s.n.], 2019. Disponível em: <<https://arxiv.org/abs/1711.05101>>. Citado na página 62.

MINGARD, C. et al. *Do deep neural networks have an inbuilt Occam's razor?* 2023. Disponível em: <<https://arxiv.org/abs/2304.06670>>. Citado 2 vezes nas páginas 29 e 30.

NOGUEIRA AR., P. A. R. S. e. a. Methods and tools for causal discovery and causal inference. *WIREs Data Mining and Knowledge Discovery* 12: e1449, 2022. Citado na página 25.

PAN, S. J.; YANG, Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, v. 22, n. 10, p. 1345–1359, 2010. Citado na página 61.

PASZKE, A. et al. Pytorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems*. [S.l.]: Curran Associates, Inc., 2019. v. 32, p. 8026–8037. Citado na página 65.

PINEDA, A. M. et al. *Machine learning for the prediction of voter model opinions through complex network structures*. 2022. Disponível em: <<https://arxiv.org/abs/2211.07420>>. Citado na página 48.

PLAGNOL, A. *Logic and theory of representation*. 2023. Disponível em: <<https://arxiv.org/abs/2307.10629>>. Citado na página 25.

- POPPER, K. *A Lógica da Pesquisa Científica*. 12. ed. São Paulo: Cultrix, 2002. Citado na página 46.
- POWERS, D. M. Evaluation: From precision, recall and f-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, v. 2, n. 1, p. 37–63, 2011. Citado na página 42.
- POZZEBON, M.; CUNHA, M. A.; COELHO, T. R. Making sense to decreasing citizen eparticipation through a social representation lens. *Information and Organization*, v. 26, n. 3, p. 84–99, 2016. ISSN 1471-7727. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1471772716301221>>. Citado na página 26.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2019. Disponível em: <<https://www.R-project.org/>>. Citado na página 26.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>. Citado na página 53.
- RADFORD, A. et al. Improving language understanding by generative pre-training. *OpenAI*, 2018. Disponível em: <https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf>. Citado na página 61.
- RAFFEL, C. et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, v. 21, n. 140, p. 1–67, 2020. Citado na página 61.
- RECKER, J. et al. Information systems as representations: A review of the theory and evidence. *Journal of the Association for Information Systems*, v. 20, n. 6, 2019. Disponível em: <<https://aisel.aisnet.org/jais/vol20/iss6/5>>. Citado na página 53.
- RENNÓ, L. R. The bolsonaro voter: Issue positions and vote choice in the 2018 brazilian presidential elections. *Latin American Politics and Society*, Cambridge University Press, v. 62, n. 4, p. 1–23, 2020. Disponível em: <https://www.researchgate.net/publication/342978868_The_Bolsonaro_Voter_Issue_Positions_and_Vote_Choice_in_the_2018_Brazilian_Presidential_Elections>. Citado na página 32.
- RITA, P.; ANTÔNIO, N.; AFONSO, A. Social media discourse and voting decisions influence: sentiment analysis in tweets during an electoral period. *Social Network Analysis and Mining*, Springer, v. 13, p. 46, 2023. Disponível em: <<https://doi.org/10.1007/s13278-023-01048-1>>. Citado na página 57.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, v. 323, n. 6088, p. 533–536, 1986. Disponível em: <<https://doi.org/10.1038/323533a0>>. Citado na página 61.
- SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural Networks*, v. 61, p. 85–117, 2015. Citado na página 61.
- SLAPIN, J. B.; PROKSCH, S.-O. A scaling model for estimating time-series party positions from texts. *American Journal of Political Science*, Wiley, v. 52, n. 3, p. 705–722, 2008. Disponível em: <<https://doi.org/10.1111/j.1540-5907.2008.00338.x>>. Citado na página 58.

- SLEGTE, J. de; DROOGENBROECK, F. V.; GINIS, V. The use of machine learning methods in political science: An in-depth literature review. *International Political Science Review*, 2024. Online First. Disponível em: <<https://doi.org/10.1177/14789299241265084>>. Citado 2 vezes nas páginas 24 e 47.
- SOUSA, R. S. d.; FILHO, P. d. O.; LARRAIN, L. F. C. R. Adesão ao bolsonarismo: Dilemas, contradições e retórica em argumentos de seus militantes. *Psicologia: Ciência e Profissão*, Conselho Federal de Psicologia, v. 44, p. e270308, 2024. ISSN 1414-9893. Disponível em: <<https://doi.org/10.1590/1982-3703003270308>>. Citado 2 vezes nas páginas 23 e 56.
- SOUZA, F.; NOGUEIRA, R.; LOTUFO, R. Bertimbau: Pretrained bert models for brazilian portuguese. *arXiv preprint arXiv:2009.11942*, 2020. Disponível em: <<https://arxiv.org/abs/2009.11942>>. Citado na página 62.
- SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, v. 15, n. 56, p. 1929–1958, 2014. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>. Citado na página 62.
- STERKENBURG, T. F. Statistical learning theory and occam’s razor: The core argument. *Minds and Machines*, v. 35, n. 1, p. 3, November 27 2024. ISSN 1572-8641. Disponível em: <<https://doi.org/10.1007/s11023-024-09703-y>>. Citado 2 vezes nas páginas 29 e 30.
- VEIGA, L. F.; GONDIM, S. M. G. A utilização de métodos qualitativos na ciência política e no marketing político. *Opinião Pública*, UNICAMP, Campinas, v. 7, n. 1, p. 1–15, 2001. Citado 2 vezes nas páginas 48 e 50.
- WOLF, T. et al. Transformers: State-of-the-art natural language processing. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics, 2020. p. 38–45. Disponível em: <<https://www.aclweb.org/anthology/2020.emnlp-demos.6>>. Citado na página 62.
- YU, H.; WANG, Q. F.; SHI, J. Y. Data augmentation generated by generative adversarial network for small sample datasets clustering. *Neural Processing Letters*, Springer, v. 55, p. 8365–8384, 2023. Disponível em: <<https://doi.org/10.1007/s11063-023-11315-z>>. Citado na página 51.
- ZHANG, J. et al. A unified search framework for data augmentation and neural architecture on small-scale image data sets. *IEEE Transactions on Cognitive and Developmental Systems*, v. 16, n. 2, p. 501–510, 2024. Citado na página 51.
- ZHOU, S. et al. A new membership scaling fuzzy c-means clustering algorithm. *IEEE Transactions on Fuzzy Systems*, IEEE, v. 29, n. 9, p. 3003441, 2021. Disponível em: <<https://dx.doi.org/10.1109/tfuzz.2020.3003441>>. Citado na página 33.
- ZHOU, X. H.; OBUCHOWSKI, N. A.; MCCLISH, D. K. *Statistical Methods in Diagnostic Medicine*. 2nd. ed. Hoboken, NJ: Wiley, 2011. ISBN 978-0-471-72738-0. Citado na página 43.
- ZHOU, Z.; MAKSE, H. A. Artificial intelligence for elections: the case of 2019 argentina primary and presidential election. *CoRR*, abs/1910.11227, 2019. Disponível em: <<http://arxiv.org/abs/1910.11227>>. Citado na página 48.