ANNALS OF DATA ANALYSIS

Populism in Words: Artificial Intelligence and Presidential Speech Analysis in Brazil (2003–2022)

Isabela do Canto[1] * - https://orcid.org/0009-0002-0720-3014
Antonio Nipo [2] - https://orcid.org/0009-0001-1618-9499

[1] Political Science Department, Universidade Federal de Pernambuco, Recife, Brazil
Email: isabela.canto@ufpe.br

[2] Electrical Engineering Department, Escola Politécnica de Pernambuco, Recife, Brazil
Email: afbn@poli.br

RESEARCH ARTICLE

Annals of Data Analysis
https://periodicos.ufpe.br/revistas/ADA/
ISSN Online: xxxx-xxxx

Submitted on: 12.12.2025.
Accepted on: 03.08.2025.
Published on: 03.08.2025.

Abstract

Text analysis and classification are a central area of Natural Language Processing (NLP), facing significant challenges when dealing with the complexities of human language, such as ambiguity, irony, and contextual nuances. These difficulties justify the traditional use of holistic qualitative approaches in political analysis. Intending to develop more efficient and replicable techniques, this work uses populism as a case study, as it is a complex political phenomenon of growing relevance in contemporary democracies. The research presents a methodological approach for identifying discursive populism in Brazil between 2003 and 2022, applying a set of NLP and machine learning techniques. We investigate the effectiveness of different models, including classical approaches (SVM, Random Forest), deep learning (DeBERTa), and generative artificial intelligence (Gemini), to classify presidential speeches based on the "people vs. elite" dichotomy. The results demonstrate that while traditional and deep learning models have limited performance due to data complexity and imbalance, the Gemini model stands out as a superior screening tool compared to the dictionary-based approach. The analysis of Gemini's "classification errors" not only reveals the insufficiency of minimalist definitions of populism but also illuminates the nuances of populist rhetoric in Brazil, such as Lula's conciliatory populism and Bolsonaro's implicit performance. We conclude that the most robust approach to political discourse analysis lies in the symbiosis between the scale and efficiency of computational tools and the depth and rigor of human qualitative analysis, reinforcing the importance of hybrid methods in data science applied to the social sciences.

Keywords

Populism, Natural Language Processing (NLP), Large Language Models (LLM), Presidential Rhetoric, Text Classification.

1. Introduction

Text analysis in political science, especially regarding political speeches, constitutes a terrain of high complexity and, simultaneously, enormous analytical potential. The ambiguous, fluid, and sometimes contradictory nature of language hinders the application of purely

quantitative methods and, historically, has led to the predominance of qualitative approaches, which are more sensitive to context and discursive nuances. However, such methods present significant limitations regarding replicability, scale of analysis, and comparability between different cases. In this sense, the development of hybrid methodologies, which combine the precision of computational techniques with human interpretive depth, represents a strategic opportunity for the renewal of studies on political discourse.

Populism emerges as a privileged object in this debate, both for its empirical relevance in contemporary democracies and for its resistance to stable and consensual definitions. Defined here in minimalist terms - the opposition between a “virtuous people” and a “corrupt elite” - the concept has been criticized for its excessive elasticity, which allows it to encompass heterogeneous and even contradictory phenomena. This characteristic creates a fundamental gap: how to delimit, measure, and empirically compare a phenomenon expressed in such diverse discursive registers? Literature has resorted to dictionary methods and manual analysis to handle this problem, but both strategies have limitations: dictionaries tend to reduce discourse to superficial term counts, and their creation is subject to criticism regarding potential bias, while qualitative analysis lacks scale and faces replication difficulties.

It is at this point that advancements in Natural Language Processing (NLP) and machine learning techniques become particularly promising. By allowing the automatic classification of large textual corpora and the identification of latent patterns, these methods offer greater scope and empirical consistency. However, they also face obstacles: data imbalance, the difficulty of capturing irony, context, or discursive performance, and the tendency to reproduce theoretical or technical biases embedded in the models. Thus, rather than replacing qualitative analysis, the use of computational tools should be understood as a complementary resource capable of expanding, but not exhausting, the interpretation of the populist phenomenon.

This work contributes to this effort by proposing and testing a methodology for identifying populism in Brazilian presidential speeches between 2003 and 2022. The period covers governments with contrasting rhetorical styles - Lula, Dilma Rousseff, Michel Temer, and Jair Bolsonaro - offering a fertile empirical field to evaluate the performance of different techniques. The study compares traditional models (such as SVM and Random Forest), deep learning (DeBERTa), and generative artificial intelligence (Gemini). By exploring both classification hits and "errors," the research demonstrates how the limits of minimalist definitions of populism can be tested against empirical evidence, revealing, for example, Lula's conciliatory rhetoric or Bolsonaro's implicit performance. The central objective, therefore, is not only to improve the automated detection of populism but also to contribute to the conceptual and methodological refinement of the field, indicating how hybrid methods can overcome persistent gaps. By articulating scale, rigor, and critical interpretation, this research seeks to demonstrate that true robustness in political discourse analysis stems from the symbiosis between Data Science and Political Science, rather than the primacy of one over the other.

The adopted methodological approach is hybrid, combining:

1. Computational techniques (supervised and unsupervised classifiers) for large-scale analysis;
2. Qualitative validation to ensure coherence with theoretical concepts;

3. Critical analysis of the limits of traditional definitions of populism against the patterns identified in the data.

2. Theoretical Framework

Discourse analysis, especially when focused on complex phenomena like populism, has historically been supported by qualitative methods that allow for the detailed interpretation of speeches, narratives, and rhetorical strategies used by populist leaders and movements. These approaches, often called "holistic classification," involve the careful reading of political texts, the identification of linguistic patterns, and the contextualization of messages within specific socio-historical scenarios. Hawkins (2009) and Tamaki and Fuks (2020) emphasize that this methodology is essential for understanding the nuances of populist rhetoric, especially because populism frequently resorts to metaphors, ironies, and emotional mobilization strategies that are not easily captured by pure quantitative techniques. Furthermore, qualitative analysis allows for situating speeches within their cultural and political contexts, offering a richer and more contextualized understanding of populist dynamics.

One of the main benefits of qualitative analysis is its methodological flexibility, which enables adaptation to different theories and analytical perspectives. In studies on populism, this means researchers can apply different theoretical approaches, such as Mudde's ideational populism (2017) or Lynch and Casimiro's authoritarian populism (2022), adjusting their analyses according to the specific characteristics of the studied textual corpus. Additionally, qualitative analysis is fundamental for identifying discursive variations among different populist leaders, revealing how the same phenomenon can take distinct forms depending on the national context, political culture, or predominant social demands. These differences require careful analysis to capture the rhetorical intentions and specific legitimation strategies of each movement.

A widely used approach is the dictionary method, which consists of creating lists of words representing interest groups. In the case of this research, one list represents the "people" and another represents the "elite". Words are chosen by researchers based on theoretical and ideological criteria. The text is then analyzed quantitatively, and sentences or documents are classified based on the presence and frequency of these predefined terms. The main advantage of this technique lies in its simplicity of application and the clarity of the process. However, by reducing the complexity of such intricate ideas, it becomes inflexible and unable to capture other contextual ways of referring to them. This often results in misclassifications and difficulty in measuring abstract concepts like populism itself. This conceptual rigidity, which ignores rhetorical subtleties, creates a need for more sophisticated approaches capable of capturing ideological nuances emerging from contextual relationships between terms rather than their mere isolated presence. It is also criticized for potential sample bias and the choice of words that may leave out sentences containing the subjective idea of populism without strictly following classic keywords.

The transition to more sophisticated computational methods inaugurates a context-sensitive approach, overcoming this rigidity. An example of this evolution is found in Dai (2019), who utilizes the word embedding technique to classify populism contextually. In this approach, words cease to be mere accounting units and are transformed into vectors positioned in a multidimensional space, where proximity reflects not only semantic similarity but also ideological affinity inferred from their relationships within the corpus. Large Language Models (LLMs) like Gemini operate under this logic of contextual semantic interpretation through

word embedding. Instead of simply checking for the occurrence of a keyword, as in dictionary logic, it evaluates the sentence's content in its entirety, seeking to identify the presence of the antagonistic "people vs. elite" duality, even when expressed implicitly or with a vocabulary not predicted by the theoretical dictionary. This capability allows it to overcome limitations in capturing irony or metaphors, as well as vulnerability to researcher selection bias in word list construction. In practice, this means the LLM avoids excluding genuinely populist sentences that would be discarded by a dictionary filter, which would lead to the loss of data relevant to the analysis.

This development does not diminish the importance of Qualitative Analysis. However, it faces criticism regarding its limitations in handling large volumes of data, as its manual nature makes application to large databases - such as collections of presidential speeches, interviews, opinion articles, or social media posts - impractical. This limits the capacity for longitudinal studies that seek to identify discursive evolution patterns over time or comparative analyses between countries or leaders. Furthermore, the subjectivity of coders is a point of discussion, as different analysts may interpret the same text in different ways, affecting the replicability of results. Given these challenges, political science has sought to complement traditional qualitative analysis with quantitative and computational methods, creating a hybrid approach capable of balancing interpretive depth with scale and replicability.

The intersection of these two perspectives allows researchers to explore both general patterns and contextual nuances of populist discourse, ensuring a more robust and precise analysis. Machine learning-based methods and Natural Language Processing (NLP) allow for the systematic identification of discursive patterns in large volumes of textual data, facilitating the detection of populist elements in presidential speeches. For populism analysis specifically, data depends on whether the research focus is on the people - in a discussion more oriented toward the field of Public Opinion - or the elite. Respectively, social media content and party manfestos/speeches are used. Sim (2013) highlights that using automated methods also helps reduce the subjectivity typical of qualitative analysis, increasing the consistency of results and allowing for the reproduction of studies by other researchers.

However, despite these advantages, computational tools also present significant limitations, especially regarding the interpretation of complex contexts and the understanding of rhetorical nuances. Machine learning models frequently fail to detect irony, sarcasm, and implicit cultural references, which can lead to imprecise classifications. One of the main quantitative methods of textual analysis includes supervised machine learning. In this approach, the model is trained with a previously classified dataset. This learning involves several stages of text preprocessing (tokenization, lemmatization, stopword removal) for model application. Many techniques and models fall into this category, also known as Natural Language Processing (NLP). This is the case for word embedding, neural networks, and Fine-tuning Language Models. Neural network models such as BERT and RoBERTa have the capacity to understand syntax and semantics because they were pre-trained on vast textual datasets. This model can further be specialized for tasks through the fine-tuning process, one of the most advanced approaches today. It consists of subjecting the already trained language model to additional training using a more restricted dataset focused on the desired task - in this case, ideological classification. A notable example is the work of Bonikowski et al. (2022), in which the RoBERTa model is used to measure the presence of populism, nationalism, and authoritarianism in U.S. presidential campaign speeches.

These approaches still need to be verified and validated based on qualitative analyses. This practice helps mitigate the individual limitations of each method while strengthening the reliability and interpretability of the data. The integration of qualitative and quantitative analyses is, therefore, a recommended methodological strategy for complex studies, such as those investigating populism and democratic erosion. Besides identifying populist elements, this methodology allows for evaluating how authoritarian rhetoric evolves over time, reflecting changes in the political and social context and ensuring that identified patterns align with the theoretical characteristics of populism. The way validation is performed varies considerably among researchers, ranging from comparison with human tabulation to expert surveys (Bonikowski et al., 2022; Di Cocco and Monechi, 2021; Schwarzbözl and Fatke, 2017), out-of-sample predictions (Dai, 2019), databases like the Global Populism Database (Raveau et al., 2024), or machine learning classification model evaluation metrics - which, during the construction of learning and testing banks, underwent human analysis (van der Veen et al., 2024).

The research is based on the ideational definition of populism, which understands the phenomenon as a set of ideas postulating the moral opposition between a "virtuous and homogeneous people" and a "corrupt elite". For this approach, populism is not a full ideology but a "central idea" that can be attached to different ideologies, such as left-wing populism (focused on criticizing economic elites) and right-wing populism (focused on defending conservative values and attacking institutions). The dichotomous operationalization of this definition is the starting point for applying computational techniques.

3. Methodology

3.1. Data Acquisition and Preprocessing

The analysis corpus was built from two sources: Ricci's (2021) database of official Brazilian presidential speeches for the periods of Lula (2003-2010), Dilma Rousseff (2011-2016), and Michel Temer (2016-2018), and a new collection of Jair Bolsonaro's speeches (2019-2022). Ricci's original database uses a dictionary technique to pre-filter potentially populist sentences before manual classification. To mitigate the bias of this filter, a new statistically representative sample of the total corpus, without the dictionary filter, was created for this study.

The study utilizes a hybrid approach, combining large-scale quantitative analysis with manual qualitative validation. To ensure a robust analysis and mitigate the bias present in previous dictionary-filtered datasets, this research employs a statistically representative sample of the unfiltered total corpus.

The NLP pipeline involved standard linguistic cleaning - including lemmatization, tokenization, and stopword removal - alongside vectorization methods such as TF-IDF and Word Embeddings (Word2Vec and BERT). For deep learning architectures (CNN, LSTM, and DeBERTa), preprocessing was adjusted to preserve sequential information and contextual richness. To address the severe class imbalance between populist and non-populist sentences, the Synthetic Minority Over-sampling Technique (SMOTE) and internal model rebalancing were applied.

3.2. Applied Models

In this research, distinct classification models were employed, ranging from traditional machine learning techniques to more recent approaches based on Deep Learning and models without supervised learning but equipped with advanced linguistic processing. Within machine learning, traditional models used were Random Forest and Support Vector Machine (SVC). Random Forest, an ensemble model based on multiple decision trees, is widely recognized for its robustness against noisy data and its ability to reduce overfitting. Support Vector Machine proved useful for high-dimensionality problems, as is common in textual data, with its ability to find the optimal separating hyperplane between classes.

For deep learning models, widely used architectures in NLP tasks were chosen: CNN (Convolutional Neural Network) and LSTM (Long Short-Term Memory). CNN has shown great efficiency in capturing local patterns and semantic structures in short texts. In turn, LSTM specializes in sequence processing, being capable of capturing long-term dependencies between words and sentences in discourse, a crucial skill for understanding the ideological and rhetorical context present in political messages. Validation involved a direct comparison between model classifications and those made by human specialists. To evaluate the models and identify error patterns, besides traditional metrics - such as accuracy, precision, sensitivity, and F1-score - sentences from a representative sample in each approach were analyzed qualitatively.

3.5.1. Support Vector Machine (SVM)

SVM is a probabilistic model. However, it uses the idea that there is an optimal hyperplane—through support vectors, which determine the spatial position of points (words) closest to the hyperplane - that understands the logic of sentence classification, handling semantically complex data better and remaining resistant to outliers. Like Naive Bayes, it applies best to well-defined classes; linguistic ambiguity and the possibility of using the same word in both populist and non-populist sentences affect the result, but the use of sample rebalancing techniques justifies its application in the study given its excellent generalization capacity. However, performance can be negatively affected when data is not linearly separable or is imbalanced, which is the case for an abstract ideological classification like populism.

3.5.2. Random Forest

This technique builds a forest through the creation of independent decision trees formed by different random samples of the set of sentences, combining results to understand the researched pattern. It is widely used due to the reduction of overfitting provided by multiple decision trees, and it is excellent for complex, imbalanced data with high robustness against noise and variability.

3.5.3. CNN (Convolutional Neural Network)

Its operation is based on applying convolutional filters that scan the text in expressions (n-grams), capturing local patterns such as the co-occurrence of ideological terms (e.g., "corrupt elite", "good people"). Afterward, significant terms are selected, reducing dimensionality and highlighting relevant patterns. This structure is ideal for detecting key expressions characteristic of populist speeches. While efficient for local patterns and fixed syntactic structures, it has less contextual memory. Thus, it is less effective at analyzing long-term relationships and may lose semantic nuances in broader discursive strategies.

3.5.4. LSTM (Long Short-Term Memory)

LSTM is a variation of recurrent neural networks (RNNs) developed to overcome the long-term memory loss that limits traditional RNNs. Its architecture consists of memory cells and three gates: input, forget, and output. These gates regulate information flow, allowing the network to retain or discard data based on relevance across the textual sequence. This model can maintain contextual information over time, capturing long-term dependencies and semantic nuances in extensive speeches. Its main limitation for Humanities applications is the need for large volumes of classified data to reach satisfactory performance.

3.5.6. DeBERTa (Decoding-enhanced BERT with disentangled attention)

In addition to classic models and CNN/LSTM architectures, this research incorporated DeBERTa, an advanced language model based on the Transformer architecture. DeBERTa represents a significant evolution over models like BERT by introducing a disentangled attention mechanism. This technique separates the word vector representation into two components: content (meaning) and relative position within the sentence. This separation is particularly advantageous for political discourse analysis. In a context-dependent concept like populism, word relationships and positions are as important as the words themselves. By treating content and position independently, DeBERTa gains a deeper understanding of syntactic and semantic nuances.

3.6. Evaluation Metrics

Validation is essential to ensure results are methodologically reliable for political classification. In this project, manual classification was taken as the gold standard. Differences between categorizations were analyzed qualitatively. Metrics used were: Accuracy[1], Precision[2], Sensitivity (Recall)[3], and F1-score[4].

For this study, the F1-score was chosen as the primary metric as it provides a balance between concerns regarding false positives and false negatives in a political context.

[1] Measures the proportion of correct classifications out of the total. In imbalanced datasets, it can be misleading.
[2] The proportion of correct positive (populist) classifications among those predicted as positive. It minimizes false positives.
[3]Perceives how many, among all populist sentences, were correctly classified, even at the cost of false positives. It is useful for identifying all populist instances.
[4] The harmonic balance between precision and sensitivity, fundamental for imbalanced databases. It provides a balanced view of performance.

3.7. Qualitative Validation

Beyond quantitative metrics, a qualitative evaluation stage was incorporated. The objective was to validate results and deepen the understanding of error patterns. A statistically representative sub-sample of classifications was selected. This sample was then subjected to manual interpretive analysis conducted "blindly" - without the analyst knowing the model's classification. This ensured impartiality and provided a richer comparative analysis.

4. Results

4.1. Quantitative Analysis

The comparison between Random Forest, SVM, DeBERTa, and Gemini reveals distinct performance profiles. In all cases, accuracy proved misleading due to data imbalance. With the exception of Gemini, all presented accuracies above 98%, which do not reflect their real capacity to identify populist discourse. Performance can be grouped into three categories:

Classic Models (Random Forest and SVM): Both presented unsatisfactory and similar performance. With low F1-Scores (around 0.25), they failed in both precision and sensitivity. This demonstrates they lack the capacity to learn subtle patterns in the minority class.

Deep Learning Model (DeBERTa): With an F1-Score of 0.48 - nearly double the classic models - DeBERTa was the most balanced approach. Its performance was balanced between precision (51.22%) and sensitivity (45.65%).

Generative AI (Gemini): Gemini presented a unique and extreme profile. It maximized sensitivity to a nearly perfect level (98.18%) at the cost of extremely low precision (4.34%). It functioned as a nearly exhaustive screening tool, ensuring virtually no populist sentence was left behind.

Table 1 - Metrics per Model

Model	Accuracy	Precision	Sensitivity	F1-Score
Random Forest	98.89%	2.500	2.222	2.353
SVM	98.58%	2.121	3.111	2.523
DeBERTa	99.04%	5.122	4.565	4.828
Gemini	79.91%	434	9.818	831

4.2. Qualitative Analysis

The quantitative results reveal a significant performance gap between models, yet the qualitative depth of the presidential speeches from 2003 to 2022 explains why automated classification remains a challenge. The following sections detail the rhetorical nuances that defined each administration.

4.2. The Accuracy Paradox and Model Limitations

While models like Random Forest, SVM, and DeBERTa showed accuracies above 98%, this metric is misleading due to the severe class imbalance - where non-populist sentences constitute the overwhelming majority of the corpus. DeBERTa proved to be the most balanced autonomous classifier, with an F1-score (0.483) nearly double that of traditional models, indicating a superior ability to capture contextual nuances. In contrast, the Gemini model functioned as a high-sensitivity screening tool (recall of 0.982), ensuring that virtually no potentially populist sentence was missed for human review.

4.3.1. Lula (2003–2010): Conciliatory vs. Polarizing Rhetoric

The primary challenge in classifying Luiz Inácio Lula da Silva's discourse was the high frequency of "false positives". While his speeches frequently utilized terms such as "the people," "the poor," and "the rich," they were framed within a project of social inclusion and institutional conciliation rather than existential confrontation.

- Intention: Lula's rhetoric focused on expanding rights and consumption for marginalized sectors without delegitimizing his opponents as "enemies of the people".
- International Stance: In the international arena, he framed the inequality between rich and poor nations as a call for solidarity and negotiation, reinforcing a pragmatic rather than populist logic.

4.3.2. Dilma Rousseff (2011–2016): From Technocracy to Reactive Populism

Dilma Rousseff's early years were characterized by a technocratic and institutionalist communication style with low emotional intensity. However, a significant shift occurred during the 2015-2016 institutional crisis.

- The Inflexion: Facing impeachment, Rousseff adopted a "reactive populism," framing herself as a representative of the popular vote betrayed by a "conspiratorial political elite".
- Mechanism: This was an episodic use of populism as a survival strategy rather than a systematic characteristic of her administration.

4.3.3. Michel Temer (2016–2018): The Anti-Populist Pole

Michel Temer represents the complete absence of populist rhetoric. Both automated models and manual validation confirmed zero populist instances in his speeches.

- Technocratic Identity: Temer explicitly rejected the search for popularity, positioning himself as a "technical manager" of economic rationality and institutional transition.
- Establishment Representation: His lack of a direct electoral mandate and his reliance on partisan mediation made a "voice of the people" narrative symbolically non-viable.

4.3.4. Jair Bolsonaro (2019–2022): Subtle and Adaptive Populism

Jair Bolsonaro presents the most complex case, where populism often transcends the minimalist "people vs. elite" formula. His rhetoric frequently generated "false positives" because the antagonism was often implicit or built through a curated persona.

- Implicit Antagonism: Bolsonaro constructed himself as the only virtuous leader ("A person who believes in God... loyal to his people"), causing the listener to subvert the "elite" as any institution (media, judiciary) that opposed him.
- Institutional Erosion: During the COVID-19 pandemic, his populism turned anti-systemic, deturpating constitutional phrases like "all power emanates from the people" to justify personal power over democratic checks and balances.

5. Conclusion

The main methodological contribution of this study is the demonstration of Generative AI's superiority as a screening tool for political discourse analysis compared to dictionary-based approaches. The "errors" in Gemini's classification provided empirical evidence of the insufficiency of minimalist definitions. The tool captured nuances like Lula's conciliatory false positives and Bolsonaro's implicit populist tactics. This suggests that contemporary populism in Brazil operates through subtle and implicit mechanisms. Ultimately, research robustness is achieved through a hybrid approach where machines optimize data screening and human analysis provides the necessary interpretive depth.

The research highlights a significant paradigm shift in political science, demonstrating that the future of discourse analysis lies in a hybrid symbiosis between computational scale and human interpretative depth. By identifying the "Accuracy Paradox" - where models show high accuracy due to data imbalance while failing to detect subtle ideological cues - the study proves that traditional metrics like accuracy are insufficient for analyzing complex phenomena like populism. Instead, generative AI like Gemini emerges as a strategically valuable "high-recall" screening tool that acts as a "fine-mesh net," ensuring that nearly all potentially populist sentences are captured for subsequent qualitative analysis. This allows researchers to move beyond the limitations of dictionary-based filters, which are often rigid and prone to researcher bias, and focus human effort on the most semantically rich and ambiguous data points.

Furthermore, the study's exploration of classification "errors" serves as a powerful diagnostic tool for refining political theory, revealing that minimalist definitions of populism often fail to capture the contemporary reality of the phenomenon. The discovery of "Lula's Conciliatory Populism" and "Bolsonaro's Implicit Performance" suggests that populist rhetoric in Brazil has evolved into more sophisticated, adaptive, and subtle forms that do not always rely on the explicit "povo vs. elite" formula. By showing how digital-era leaders use persona-building and institutional delegitimation to mobilize followers, this research provides a methodological roadmap for monitoring authoritarian trends and strengthening democratic resilience. It reinforces the necessity of interdisciplinary approaches, where data science optimizes data processing while political science provides the critical context needed to understand the risks these discourses pose to institutional stability.

Acknowledgements

The completion of this research was made possible through the support and contribution of several individuals and institutions. Recognition is due to the Universidade Federal de Pernambuco (UFPE) and its Political Science Department, which provided the intellectual environment and institutional support necessary for the development of this work. Gratitude is also extended to the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for providing the graduate scholarship that allowed for full dedication to this project. Significant acknowledgment is owed to the authors of the Ricci (2021) dataset, whose foundational work on presidential populism in Brazil served as the primary data source for this study.

Conflict of Interest Declaration

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report.

References

1. Bonikowski, B., Luo, Y., & Stuhler, O. (2022). Politics as usual? Measuring populism, nationalism, and authoritarianism in U.S. presidential campaigns (1952–2020) with neural language models. Sociological Methods & Research.
2. Dai, Y. (2019). Measuring populism in context: A supervised approach with word embedding models.
3. Di Cocco, J., & Monechi, B. (2021). How populist are parties? Measuring degrees of populism in party manifestos using supervised machine learning. Political Analysis.
4. Finchelstein, F. (2019). Do fascismo ao populismo na história. Almedina.
5. Foa, R. S., & Mounk, Y. (2016). The danger of deconsolidation: The democratic disconnect. Journal of Democracy, 27(3), 5–17.
6. Hawkins, K. A., & Kaltwasser, C. R. (2017). The ideational approach to populism. Latin American Research Review, 52(4), 513–528.
7. Hawkins, K. A. (2009). Is Bolivia becoming a new Venezuela? Survey evidence from the 2008 recall referendum. [This citation is used in your theoretical framework regarding "holistic classification"].
8. King, G., Keohane, R. O., & Verba, S. (1994). Designing social inquiry: Scientific inference in qualitative research. Princeton University Press.
9. Lynch, C., & Cassimiro, P. H. (2022). O populismo reacionário. Editora Contracorrente.
10. Moffit, B. (2015). How to perform crisis: A model for understanding the key role of crisis in contemporary populism. Government and Opposition, 50(2), 189–217.
11. Mudde, C. (2017). The populist radical right: A fourth wave of deconsolidation? American Political Science Review, 111(4), 699–717.

12. Nunes, F., & Traumann, T. (2023). Biografia do abismo: Como a polarização divide famílias, desafia empresas e compromete o futuro do Brasil. HarperCollins Brasil.
13. Raveau, M. P., Fuentes-Bravo, C., Ángel-Fernández, M., Coyoumdjian, J. P., & del Solar, M. J. (2024). “It's not the what but (also) the how”: Characterizing left-wing populism in political texts. Frontiers in Political Science.
14. Ricci, M. (2021). Populismo presidencial no Brasil (1989-2018): Um estudo de caso sobre os governos Collor, FHC, Lula e Dilma. Revista Brasileira de Ciências Sociais, 36(107).
15. Schwarzbözl, T., & Fatke, M. (2017). Measuring populism in social media data.
16. Sim, Y., Acree, B. D., Gross, J. H., & Smith, N. A. (2013). Measuring ideological proportions in political speeches. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (pp. 91–101).
17. Tamaki, E. R., & Fuks, M. (2020). Populism in Brazil’s 2018 general elections: An analysis of Bolsonaro’s campaign speeches. Lua Nova: Revista de Cultura e Política, 103–127.
18. Urbinati, N. (2019). Me the people: How populism transforms democracy. Harvard University Press.
19. van der Veen, O., Dzebo, S., Littvay, L., Hawkins, K., & Dar, O. (2024). Classifying populist language in American presidential and governor speeches using automatic text analysis.Cooper, B. S. (2018). *Interactive Planning and Sensing for Aircraft in Uncertain Environments with Spatiotemporally Evolving Threats* (Doctoral dissertation, Department of Aerospace Engineering, The Pennsylvania State University).