



Before the Votes Are Cast: Electoral Forecasting in
OLPR
Predicting the 2026 Brazilian Chamber of Deputies
Election using Machine Learning algorithms

Akira Pinto-Medeiros
University of Pittsburgh

Bruno Wilhelm Speck
USP

ReDem Working Paper Series
Vol. 3
Nº 6
Outubro, 2026

Before the Votes Are Cast: Electoral Forecasting in OLPR

Predicting the 2026 Brazilian Chamber of Deputies Election using Machine Learning algorithms

Akira Pinto-Medeiros, Bruno Wilhelm Speck

Resumo

Introdução: A previsão eleitoral em sistemas proporcionais de lista aberta é um desafio metodológico significativo. Com 513 cadeiras distribuídas em 27 distritos eleitorais, mais de trinta partidos registrados e milhares de candidatos competindo simultaneamente, a eleição para a Câmara dos Deputados brasileira é especialmente difícil de prever. Ao contrário das corridas majoritárias, não existem pesquisas de opinião pública cobrindo sistematicamente as eleições legislativas brasileiras. Este trabalho utiliza técnicas de Machine Learning para prever o resultado das eleições de 2026 para a Câmara. **Métodos:** Usamos modelos de aprendizado de máquina treinados em quatro ciclos eleitorais anteriores. Empregamos três algoritmos de classificação --- Regressão Logística, Gradient Boosting e uma Rede Neural Multilayer Perceptron (MLP) --- treinados em um conjunto de candidatos competitivos ($N = 9.946$). Os modelos utilizam como preditores a receita de campanha, a incumbência, o gênero, o nível educacional e a família ideológica do partido. **Resultados:** Os três modelos alcançam forte desempenho fora da amostra ($AUC > 0,90$) e suas previsões são amplamente consistentes entre si. No nível partidário, o modelo MLP prevê o União Brasil como o maior vencedor (+20 cadeiras), enquanto o PL deve perder 16 cadeiras apesar de continuar sendo o maior partido individual. **Discussão:** Os resultados confirmam que receita de campanha, incumbência, gênero e escolaridade são preditores robustos e consistentes do sucesso eleitoral na Câmara dos Deputados ao longo de quatro ciclos eleitorais. Os resultados são preliminares e serão avaliados em relação aos resultados oficiais do TSE após 4 de outubro de 2026.

Palavras-chave

Eleições brasileiras, Câmara dos Deputados, aprendizado de máquina, previsão eleitoral, financiamento de campanha, incumbência, 2026.

Akira Pinto-Medeiros

University of Pittsburgh

akira.pinto@pitt.edu

<https://orcid.org/0000-0003-2763-8824>

Bruno Wilhelm Speck

USP

bruno.speck@gmail.com

<https://orcid.org/0000-0001-5085-6825>

Abstract

Introduction: Electoral forecasting in open-list proportional representation systems is a significant methodological challenge. With 513 seats distributed across 27 electoral districts, over thirty registered parties, and thousands of candidates competing simultaneously, the Brazilian Chamber of Deputies election is especially difficult to predict. Unlike majoritarian races, there are no public opinion polls systematically covering Brazilian legislative elections. This paper uses machine-learning techniques to predict the outcome of the 2026 Chamber of Deputies election. **Methods:** We use machine-learning models trained on four previous electoral cycles. We employ three classification algorithms --- Logistic Regression, Gradient Boosting, and a Multilayer Perceptron Neural Network (MLP) --- trained on a dataset of competitive candidates ($N = 9,946$). The models use campaign revenue, incumbency status, gender, educational level, and party ideological family as predictors. **Results:** All three models achieve strong out-of-sample performance ($AUC > 0.90$) and their predictions are broadly consistent with one another. At the party level, the MLP model predicts União Brasil as the biggest winner (+20 seats), while PL is predicted to lose 16 seats despite remaining the largest individual party. **Discussion:** The results confirm that campaign revenue, incumbency, gender, and educational attainment are robust and consistent predictors of electoral success in the Chamber of Deputies across four electoral cycles. The results are preliminary and will be evaluated against the official TSE results after October 4, 2026.

Keywords

Brazilian elections, Chamber of Deputies, machine learning, electoral prediction, campaign finance, incumbency, 2026.

Note: *This is a working paper circulated for discussion prior to the October 4, 2026 election. The predictions herein are statistical estimates based on historical patterns and do not constitute endorsements of any candidate or party. An earlier version of this paper was presented at the 2026 MPSA Annual Conference.*

1 Introduction

Electoral prediction is hard. In single-member district systems like the United States or the United Kingdom, predicting outcomes is relatively tractable: a district elects one representative, party labels serve as powerful heuristics, and decades of polling infrastructure have been built around this simple structure (Jacobson 1978). Brazil is the opposite of this scenario in almost every meaningful way.

Brazil elects its 513 federal deputies through an open-list proportional representation (OLPR) system across 27 multi-member electoral districts — one per state and the Federal District. District magnitudes range from 8 seats (the smallest states) to 70 seats (São Paulo), with a national average above 19. Under this system, voters cast ballots for individual candidates rather than party lists, and seats are allocated proportionally to parties based on total votes, with seats within each party going to its individually highest-vote-getters (Nicolau 2007). The implications of this institutional design for predictability are severe.

First, Brazil operates under extreme multipartism. Following successive party fragmentations and mergers, the 2018 election saw over 21 parties win at least one seat in the Chamber of Deputies, giving Brazil one of the highest effective numbers of parliamentary parties in the world. This fragmentation is not merely a curiosity: it means that no single party or bloc dominates any district, that coalition politics govern the assembly of winning legislative majorities, and that candidate entry is essentially unrestricted. In 2022, over 9,400 candidates competed for 513 seats nationally.

Second, under open-list proportional representation, co-partisans compete against each other as much as against candidates from rival parties (Carey and Shugart 1995; Ames 1995). A candidate's strategic problem is not simply to defeat the opposing party — it is to accumulate more personal votes than fellow party members while still contributing to the party's total vote share. Ames (2001) famously described Brazilian legislative electoral strategy as driven by geographic concentration of personal vote networks, with deputies cultivating localized bases of support within their districts. The personal vote logic of OLPR means that individual candidate characteristics — resources, visibility, incumbency, personal network — matter more than in closed-list or single-member systems (Carey and Shugart 1995). Recent studies, though, have proposed the idea that intra-party competition is less frequently observed than previously expected (Cheibub et al. 2023; Cheibub and Sin 2020).

Third, the combination of high district magnitude and a large number of competing parties means that the effective electoral threshold varies across different electoral districts. In some states, where the population is large, the electoral threshold (*"quociente eleitoral"*) reaches 300,000 votes), while in smaller states with fewer voters, the electoral threshold can be of only 35,000 votes. This creates a situation where thousands of candidates register, most of whom have no realistic prospect of election, but whose very presence reshapes the vote distribution in ways that are difficult to model. For many, this produces a uniquely volatile and unpredictable electoral environment, even by the standards of other multiparty democracies in Latin America (Mainwaring et al. 2018).

These features combine to make Brazilian federal elections exceptionally difficult to forecast. Unlike U.S. House races, where incumbency rates above 90% and partisan voting patterns allow for relatively confident predictions from structural models (Jacobson 2009), Brazilian deputies face a more uncertain reelection environment. Ames (2001) estimated that incumbency reelection rates in Brazil hover around 70%, substantially lower than in the United States, and that

the personal vote basis of electoral success means that candidate-level characteristics drive outcomes in ways that aggregate models struggle to capture. In Brazil, there is no equivalent of Nate Silver's FiveThirtyEight (<https://data.fivethirtyeight.com/>) for the Brazilian Chamber of Deputies — no systematic, publicly available forecasting infrastructure that combines polling, structural models, and district-level expertise.

This paper attempts to fill that gap. While most existing work focuses on explaining past outcomes, we take a different approach to predict electoral outcomes. Using machine-learning models trained on four electoral cycles (2010, 2014, 2018, and 2022), we predict which candidates are most likely to win seats in the 2026 Chamber of Deputies election (October 4, 2026). Following (Cheibub and Sin 2020), we restrict our training dataset to *competitive candidates*— those who obtained at least 1% of the district's electoral quota (*quociente eleitoral*) in their district — in order to train the models on the genuinely contested portion of the race and reduce the noise introduced by non-viable candidacies. We employ three classifiers: Logistic Regression with year and state fixed effects, Gradient Boosting, and a Multilayer Perceptron Neural Network, and we generate predictions for all 513 seats across all 27 districts. We compare predictions across three algorithms providing a solid set of predictions for the election that is coming.

Beyond generating predictions, the paper speaks to a broader substantive question: what attributes determine electoral success in an OLPR system with high district magnitudes and extreme multipartism? Are those elections as messy and hard to follow as the institutional design suggest? Our results show that campaign revenue, incumbency status, gender, educational attainment, and ideological family are consistent and strong predictors of electoral success among competitive candidates. These findings contribute to the growing literature on political inequality and candidate competition in Brazil (Speck 2016; Janusz et al. 2022; Janusz and Campos 2021; Sacchet et al. 2025) and highlight the role of structural advantages — financial, institutional, and demographic — in shaping who gets elected to one of the world's most fragmented legislatures.

The paper proceeds as follows. Section 2 describes the data. Section 3 presents the methods. Section 4 reports the logistic regression results and model performance. Section 5 presents the 2026 predictions. Section 6 discusses limitations, and Section 7 concludes.

2 Data

2.1 Training Data: 2010–2022 Elections

The training dataset covers four electoral cycles: 2010, 2014, 2018, and 2022. It is assembled from three primary sources: (i) candidate and result data from the *Tribunal Superior Eleitoral* (TSE), Brazil's highest electoral authority; (ii) campaign finance records from the TSE; and (iii) additional incumbency data compiled by the authors. The training dataset is organized at the candidate-level and contains 9,946 observations after restricting to competitive candidates.

[Table 1](#) shows the number of candidates and elected deputies per cycle, both for the full universe and for the competitive subset used in training.

Table 1: Candidates, Competitive Candidates, and Elected Deputies by Election Year

Year	All Candidates	Competitive Candidates	Elected
2010	4,888	2,164	513
2014	5,870	2,246	513
2018	7,629	3,055	513
2022	9,455	3,719	513
Total	27,842	9,946	2,052

2.2 2026 Prediction Data

For the 2026 prediction, we use data on all registered candidates compiled from the TSE. Incumbency status for 2026 candidates is assigned by cross-referencing registered candidates with the official list of deputies currently serving in the 57th Legislature (2023–2027), obtained via the *Câmara dos Deputados* open data API (dadosabertos.camara.leg.br). This approach captures incumbents regardless of party changes between 2022 and 2026, using name and state-level matching rather than candidate number matching. Campaign revenue for 2026 is drawn from TSE partial disclosure filings (until September 16th 2026); candidates with no reported revenue are coded as having zero campaign resources.

2.3 Key Variables

The following variables are used as features in all models:

- **Total campaign revenue** (in millions of BRL): the primary predictor, drawn from TSE finance records. For pre-2015 elections, this includes corporate donations; for 2018, 2022, and 2026, this reflects post-ban public funding.
- **Incumbency status**: binary indicator (1 = current deputy running for re-election; 0 = challenger).
- **Gender**: binary dummy (1 = male; 0 = female).
- **Educational level**: binary indicator for college attendance (`edu_college`).
- **Ideological family**: categorical variable classifying parties as Left, Right, Center, or Other (inconsistent classification among scholars), following the coding scheme of (Bolognesi et al. 2022), (Tarouco and Madeira 2015), and (Zucco and Power 2024).
- **District magnitude**: number of seats available in each state-level district.
- **Election year** and **state**: included as categorical fixed effects.

3 Methods

3.1 Defining Competitive Candidates

A central methodological decision in this paper is the restriction of the analysis to somehow *competitive* candidates. In Brazilian federal elections, a large majority of registered candidates receive very few votes and have a negligible probability of election (Cheibub and Sin 2020).

Including these candidates in the training data introduces substantial noise and inflates model accuracy artificially (a model that simply predicts “not elected” for everyone achieves a high level of accuracy).

We define competitive candidates as those who obtained at least 1% of the *quociente eleitoral* (electoral quota) in their district. The electoral quota is defined as:

$$QE_d = \frac{\text{Total valid votes in district}_d}{\text{Magnitude of district}_d}$$

Candidates with $\frac{\text{votes}}{QE_d} \geq 0.01$ are classified as competitive. This filter retains approximately 38–45% of all registered candidates, depending on the election year, and captures essentially all candidates who have any realistic path to election ([Table 1](#) shows the number of competitive candidates in each election).

3.2 Logistic Regression with Fixed Effects

The baseline model is a logistic regression with year and district fixed effects. The probability of election is modeled as:

$$Pr(Y_c = 1) = \frac{e^{x_c\beta}}{1 + e^{x_c\beta}}$$

which is equivalent to modeling the log-odds of election as a linear function of candidate characteristics:

$$\log\left(\frac{Pr(Y_c = 1)}{1 - Pr(Y_c = 1)}\right) = \alpha_0 + \beta_1 \text{Revenue}_c + \beta_2 \text{Ideology}_c + \beta_3 \text{Gender}_c + \beta_4 \text{College}_c + \beta_5 \text{Incumbent}_c + \gamma_t + \delta_d + \varepsilon_c$$

where $Y_c = 1$ if candidate c was elected and $Y_c = 0$ otherwise, γ_t are election year fixed effects, and δ_d are electoral district fixed effects.

3.3 Machine-Learning Classifiers

In addition to logistic regression, we train six machine-learning classifiers: a Gaussian Naive Bayes, a KNN Classifier (K-Nearest Neighbors), a Decision Tree Classifier, a Random Forest Classifier, a Gradient Boosting Classifier, and a Multilayer Perceptron (MLP) Neural Network. All models are estimated using scikit-learn (Python), with hyperparameter selection via 10-fold cross-validated grid search optimizing for AUC.

Preprocessing pipeline. Numeric features (revenue, gender, partyideology, incumbency) are standardized using `StandardScaler`. Categorical features (year, college) are one-hot encoded.

3.4 Generating 2026 Predictions

For the 2026 election, we used the two top-performer Machine Learning Classifiers (MLPClassifier, and GradientBoosting), in addition to a Logistic Regression classifier, to generate predicted probabilities of election for each candidate. To respect the institutional constraint that exactly M_d candidates are elected in district d (where M_d is the district's magnitude), we do not use a fixed 0.5 probability threshold. Instead, within each state, we rank candidates by their predicted probability and assign the top M_d candidates as predicted winners. This ensures

that the total number of predicted elected candidates is exactly 513, and that the predictions respect district-level seat allocations.

4 Results: What Predicts Electoral Success?

4.1 Logistic Regression Results on Training data

[Table 2](#) presents odds ratios from the logistic regression model with the state and year fixed effects. These results were obtained from the training dataset (2010, 2014, 2018, and 2022 elections) and they shed the light over the statistical importance of each predictor in estimating the electoral success across 2010, 2014, 2018, and 2022 elections.

Table 2: Odds Ratios from Logistic Regression

Variable	Odds Ratio	CI Low	CI High	p-value	
Total Revenue (Millions)	4.657	4.221	5.138	0.000	***
Incumbent	5.445	4.647	6.381	0.000	***
Gender (Male = 1)	1.466	1.217	1.767	0.000	***
Ideology: Left	2.021	1.426	2.863	0.000	***
Ideology: Right	1.586	1.130	2.227	0.008	**
Ideology: Other	1.193	0.789	1.805	0.402	
Education (College = 1)	1.620	1.348	1.948	0.000	***

Year fixed effects

District fixed effects

N = 9,946.

McFadden Pseudo- R^2 = 0.368.

Significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.10$

The results reveal several strong and consistent predictors of electoral success among competitive candidates:

Campaign revenue is a powerful predictor. Each additional million BRL in campaign revenue quadruples the odds of election (OR = 4.65). This finding is consistent with the extensive literature on political finance in Brazil (Speck 2016; Marcelino 2010; Cervi et al. 2015; Peixoto 2010, 2012; Jairo Junckes et al. 2019).

Incumbency is the strongest predictor, with incumbents having odds of election nearly 5.5 times higher than non-incumbents (OR = 5.44). This large advantage likely reflects name recognition, access to institutional resources, and stronger party support. Results speak to studies that highlighted the importance of incumbency in the Brazilian legislative elections (Perissinotto and Fernandes Veiga 2014; Cervi et al. 2015).

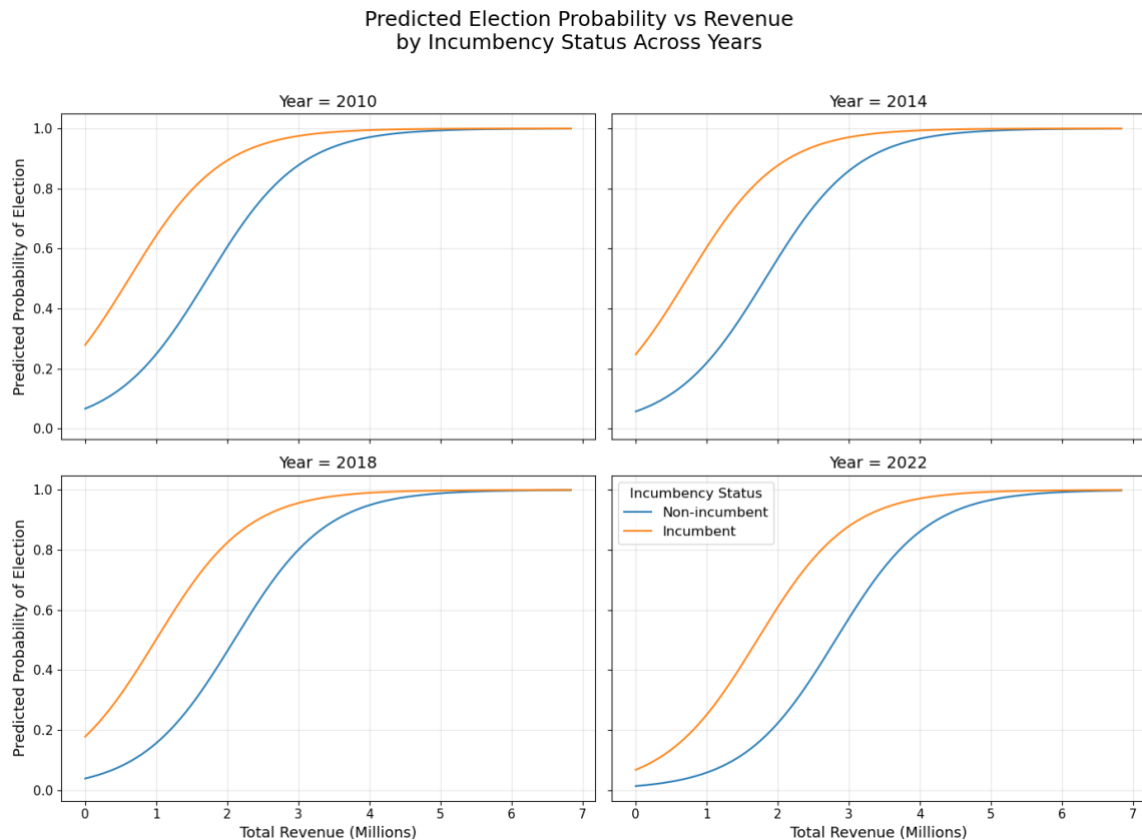
Gender remains a significant predictor, with male candidates having 46% higher odds of election than female candidates, holding other factors constant. This persists even among competitive candidates and reinforces findings in the literature on gender inequality in Brazilian electoral politics (Janusz et al. 2022; Sacchet et al. 2025).

Educational attainment: candidates who attended college have 62% higher odds of election compared to those who never attended college (OR = 1.620). This pattern has been discussed by scholars such as (Perissinotto and Fernandes Veiga 2014) who described the imbalanced distribution of college degree holders among the elected candidates in older elections (1998, 2002, and 2006).

Ideology: both left-wing (OR = 2.021) and right-wing (OR = 1.586) candidates have higher odds of election than centrist or unclassified candidates. The higher odds ratio seen in left-wing parties might be a factor of more strict selection process and candidate selections in these parties. Across all years in the training data set, the number of candidates running in left-wing parties (2,761) was more than 60% smaller than the number of candidates running in right-wing parties (7,178). This may reflect the more centralized organizational structure of left-wing parties, where national party elites exert greater influence over local and regional decisions such as candidate selection (Sousa Braga et al. 2009; Floriano Ribeiro 2013).

[Figure 1](#) below, displays the predicted probability of election as a function of campaign revenue, separately for incumbents and non-incumbents. Three patterns stand out. First, incumbency confers a substantial electoral advantage across all years: at any given level of campaign revenue, incumbents consistently show higher predicted probabilities of election than non-incumbents. Second, campaign revenue matters enormously for both groups — the predicted probability rises steeply with revenue and approaches 1.0 for well-funded candidates regardless of incumbency status. Third, the incumbency advantage appears to narrow over time. In 2010 and 2014, even incumbents with zero reported revenue had predicted probabilities above 0.25. This is likely a result of the corporate donations ban when the Supreme Court declared corporate donations to candidates as unconstitutional. By 2022, the intercept for incumbents drops closer to 0.07, suggesting that the post-ban public funding regime reduced the baseline advantage of holding office.

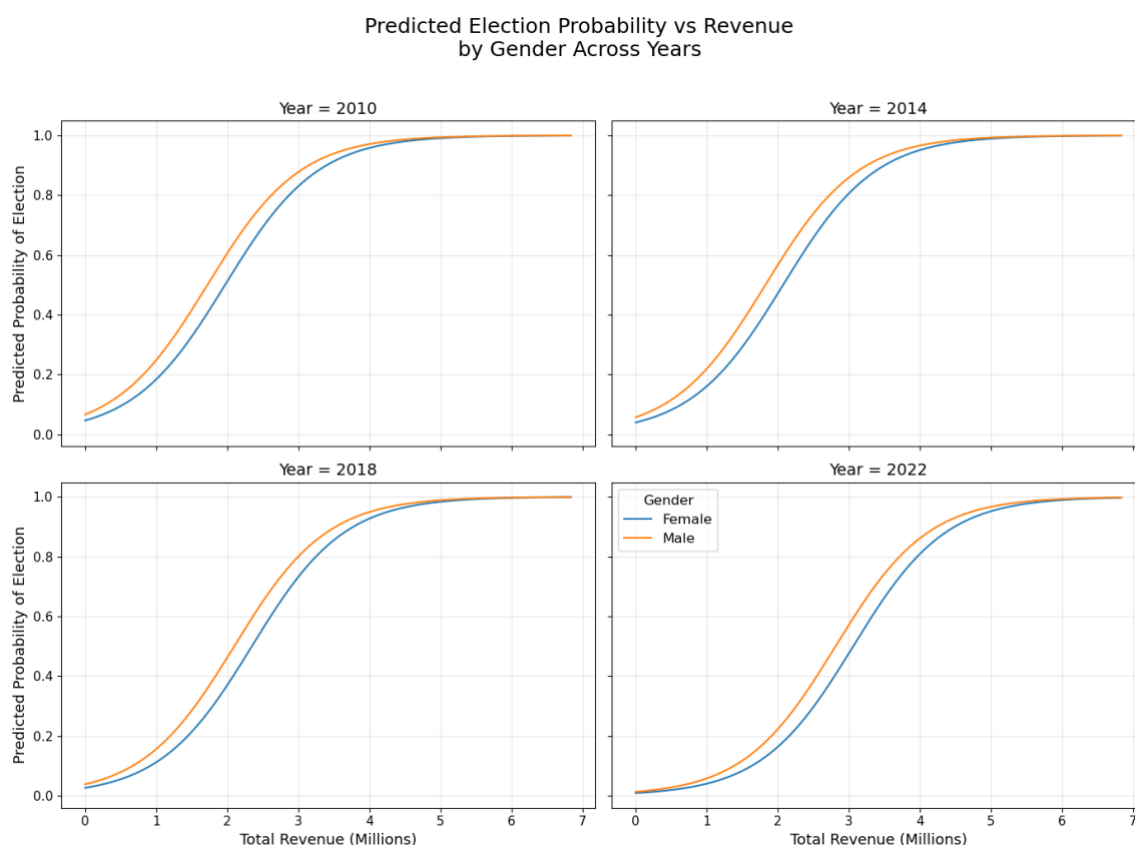
Figure 1: Predicted probability of election by revenue, and incumbency status



Other variables held at their modal or mean values. Source: TSE. Elaborated by the authors.

[Figure 2](#) displays the predicted probability of election as a function of campaign revenue, separately for male and female candidates across all four electoral cycles. The gender gap is consistent and persistent across all years: at any given level of campaign revenue, male candidates show a higher predicted probability of election than their female counterparts. This means that female candidates need to raise *more* money than male candidates to achieve the same probability of winning. Notably, the gender gap does not disappear after the 2015 corporate donation ban. These findings are consistent with recent work showing that female candidates remain systematically underfunded relative to male candidates even under the public funding regime established after 2015 (Janusz et al. 2022; Sacchet et al. 2025).

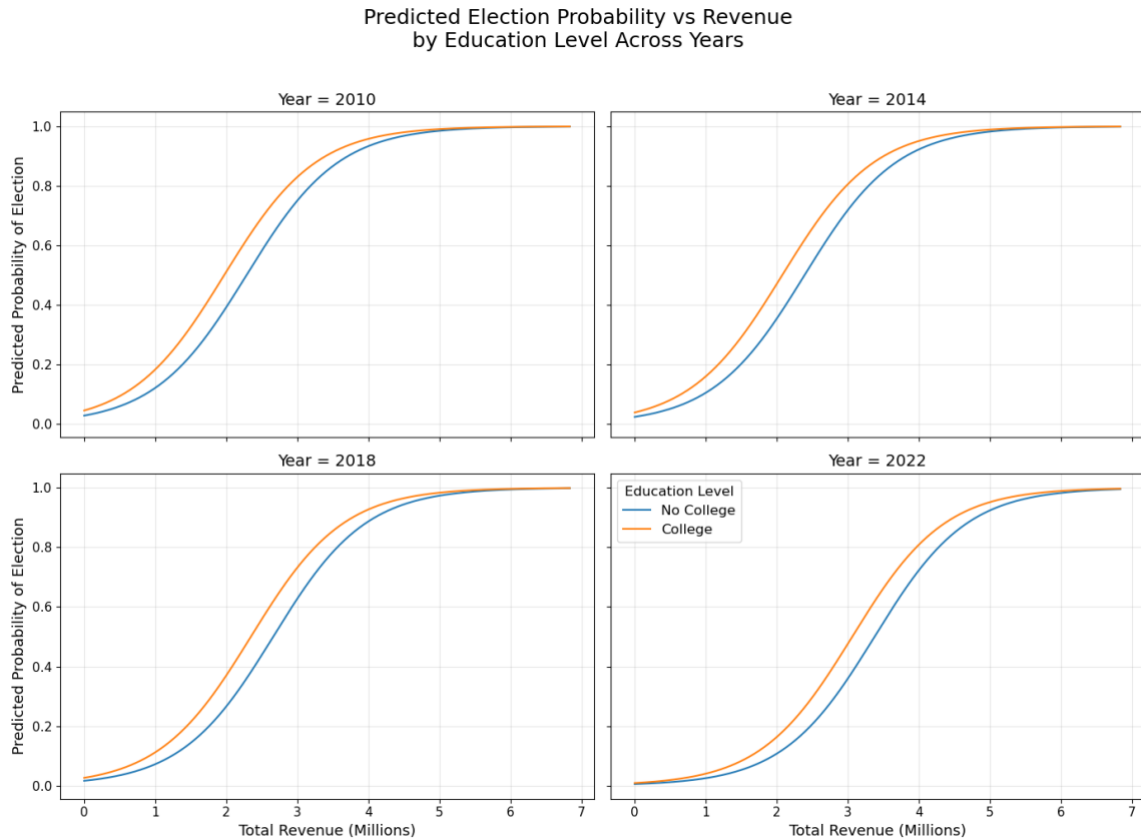
Figure 2: Predicted probability of election by revenue and gender



All other variables held at their baseline values (No College, non-incumbent, modal state and ideological family)

[Figure 3](#) displays the predicted probability of election as a function of campaign revenue, separately for college-educated and non-college candidates across all four electoral cycles. The college premium is consistent and persistent: at any given level of campaign revenue, candidates with a college degree show a higher predicted probability of election than those without one. This advantage is visible across all four years and mirrors the pattern observed for gender — non-college candidates need to raise more money than college-educated candidates to achieve the same predicted probability of winning. This finding is consistent with the literature suggesting that party elites systematically allocate more resources to college-educated candidates, believing they possess the skills necessary to run a competitive campaign (Janusz et al. 2022), which in turn reinforces the structural advantages of higher education in electoral competition.

Figure 3: Predicted probability of election by revenue and educational level



All other variables held at their baseline values (female candidates, non-incumbent, modal state and ideological family). Source: TSE. Elaborated by the authors.

Taken together, Figures 1, 2, and 3 reveal a consistent pattern across all four electoral cycles: campaign revenue is the dominant predictor of electoral success, but its effect is systematically conditioned by structural candidate characteristics. Incumbents, male candidates, and college-educated candidates all enjoy a persistent electoral advantage over their counterparts at equivalent levels of campaign spending — they achieve higher predicted probabilities of election with the same amount of money.

4.2 Model Performance

Table 3 reports performance metrics for seven classifiers evaluated using a 10-fold stratified cross-validation procedure. In this approach, the dataset is randomly partitioned into 10 equal subsets (folds), each containing approximately 894–895 candidates with a consistent share of elected candidates (approximately 17.4–17.5% per fold). The model is then trained on 9 of the 10 folds and evaluated on the remaining fold, repeating this process 10 times so that every observation is used for evaluation exactly once. Performance metrics — accuracy, precision, recall, F1 score, and AUC — are averaged across all 10 folds, producing a more reliable and generalizable assessment of model performance than a single train-test split. This procedure reduces the risk that results are driven by a particular random partition of the data, and ensures that the reported metrics reflect how well each model is expected to perform on genuinely unseen observations — such as the 2026 candidates for whom no outcome is yet known.

Table 3: Model Performance on Test Set (Competitive Candidates, 2010–2022)

Model	Accuracy	Precision	Recall	F1	AUC
Logistic Regression	0.878	0.710	0.506	0.591	0.896
Gaussian Naïve Bayes	0.882	0.678	0.619	0.647	0.859
K-Nearest Neighbors	0.885	0.705	0.586	0.640	0.910
Decision Tree	0.883	0.681	0.622	0.650	0.906
Random Forest	0.883	0.726	0.529	0.612	0.913
Gradient Boosting	0.882	0.694	0.586	0.635	0.919
MLP Neural Network	0.885	0.695	0.611	0.650	0.918

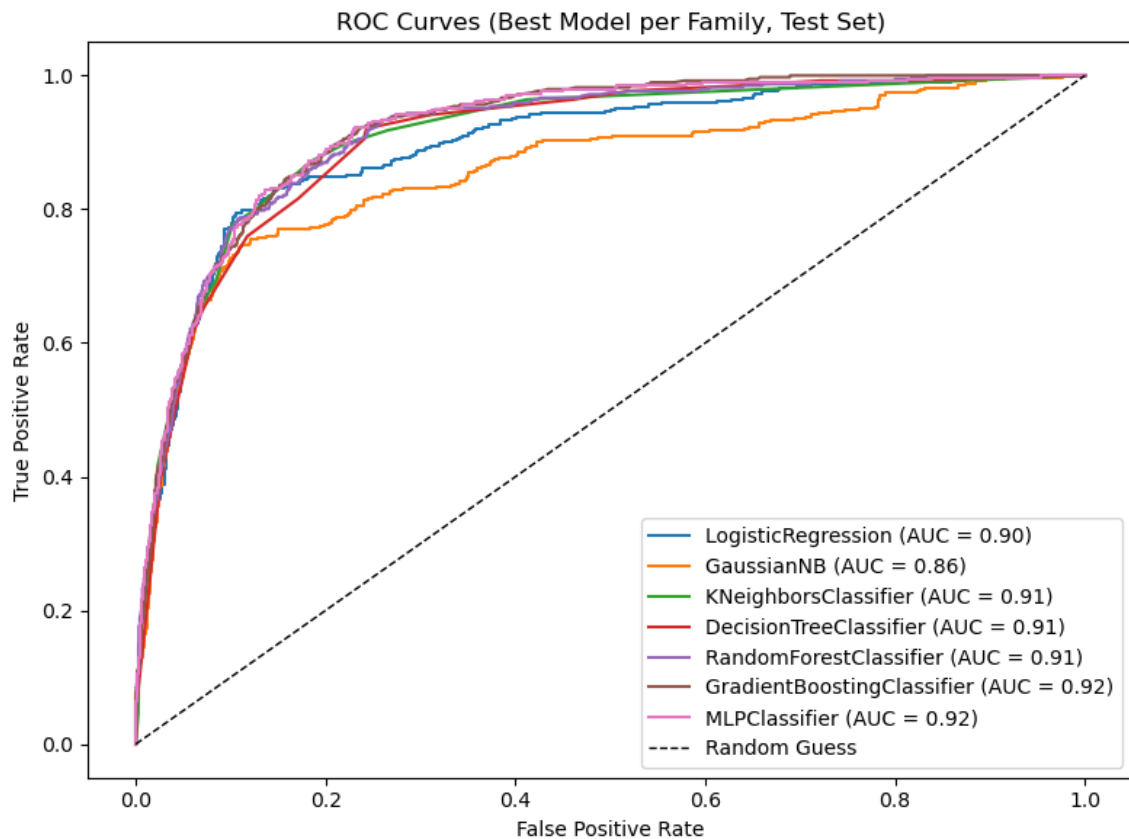
Best values in bold.

The MLP and Gradient Boosting classifiers achieve the strongest overall performance, though results are broadly comparable across all seven algorithms. To interpret these metrics, it is useful to briefly describe what each one captures. *Accuracy* measures the share of candidates correctly classified as elected or not elected; at 0.885, the MLP correctly classifies nearly 9 out of 10 competitive candidates. *Precision* measures, among all candidates the model predicts as elected, what share actually wins a seat — a high precision model is conservative and rarely makes false predictions of victory. *Recall* measures, among all candidates who actually win a seat, what share the model correctly identifies — a high recall model misses few actual winners. The *F1* score is the harmonic mean of precision and recall, penalizing models that perform well on one but poorly on the other; the MLP achieves the highest F1 (0.650), reflecting the best overall balance between these two competing objectives. Finally, the *Area Under the ROC Curve* (AUC) measures the probability that the model assigns a higher predicted probability to a randomly chosen winner than to a randomly chosen loser — an AUC of 1.0 indicates perfect ranking, while 0.5 indicates no better than random guessing. Gradient Boosting achieves the highest AUC (0.919), making it the strongest model for ranking candidates by their likelihood of election. The MLP Classifier, however, leads on accuracy (0.885) and F1 score (0.650), and trails Gradient Boosting by only 0.001 AUC points — a negligible difference in practice. Given this, we generate 2026 predictions using both models, alongside Logistic Regression as a more interpretable baseline. Presenting results from multiple algorithms allows us to assess the robustness of the predictions: candidates consistently identified as likely winners across all three models represent the most confident forecasts, while disagreements between models highlight genuinely uncertain races. All models comfortably outperform random classification (AUC = 0.5).

Compared to models trained on the full (non-competitive) candidate sample, performance on the competitive subset is lower — accuracy drops from approximately 0.94 to 0.88, and AUC from 0.96 to 0.92. This is expected: restricting to competitive candidates removes the “easy” cases (candidates with near-zero votes who are obviously not going to win), making the classification task genuinely harder and more informative. A model that achieves 0.88 accuracy on competitive candidates is doing something substantively meaningful; a model that achieves 0.94 on the full sample is largely just identifying the large pool of non-viable candidacies.

[Figure 4](#) shows the ROC curves for all models evaluated on the test set.

Figure 4: ROC curves for all classifiers (test set, competitive candidates)



Source: elaborated by the authors.

5 2026 Predictions

5.1 Overall Summary

All three models — Logistic Regression, Gradient Boosting, and MLP — generate predictions for 513 seats across 27 electoral districts. [Table 4](#) summarizes the degree of agreement across models.

Table 4: Model Agreement on 2026 Predictions

Agreement level	N candidates
Predicted elected by all 3 models	318
Predicted elected by exactly 2 models	165
Predicted elected by exactly 1 model	255
Total predicted elected (any model)	513

Candidates predicted elected by all three models represent the most confident predictions — these are individuals for whom the historical patterns in the training data converge on a clear prediction of victory.

The three algorithms used for prediction differ substantially in their underlying logic, which is precisely why using all three together strengthens the robustness of the forecasts. *Logistic Regression* is the most interpretable of the three: it models the log-odds of election as a linear

function of candidate characteristics, producing coefficients that can be directly read as the contribution of each variable to the probability of winning. Its main limitation is that it assumes a linear relationship between predictors and the outcome on the log-odds scale, and cannot capture complex interactions between variables without them being explicitly specified by the researcher. In the context of Brazilian elections, where the relationship between money and votes is non-linear and context-dependent, this is a meaningful constraint.

Gradient Boosting belongs to the family of ensemble methods: it builds a sequence of decision trees, where each tree is trained to correct the errors of the previous one, iteratively improving predictive accuracy. This approach is particularly well-suited to capturing non-linear relationships and complex interactions between variables — for instance, the possibility that the effect of campaign revenue on election probability differs depending on incumbency status or district magnitude. Its main disadvantage is interpretability: unlike logistic regression, the internal logic of a gradient boosting model is difficult to summarize in a simple equation or table of coefficients. It is also more computationally demanding and prone to overfitting if hyperparameters are not carefully tuned, which is why we rely on cross-validated grid search to select the optimal configuration.

The Multilayer Perceptron (MLP) Neural Network is a feedforward artificial neural network that learns representations of the input data through multiple layers of interconnected nodes. Like Gradient Boosting, it can capture highly non-linear patterns in the data that simpler models miss. The MLP used here has a single hidden layer of 50 neurons with a hyperbolic tangent activation function, trained using backpropagation. Its key advantage is flexibility — it makes very few assumptions about the functional form of the relationship between predictors and outcome. Its disadvantages mirror those of Gradient Boosting: it is a black-box model whose predictions are difficult to interpret substantively, and it requires careful regularization (controlled here via the alpha penalty parameter) to avoid overfitting on the training data.

In practice, the three models agree on the vast majority of predicted winners, which is reassuring: the core signal in the data — that well-funded incumbents from established party blocs are most likely to win — is strong enough to be detected by algorithms as different as a linear model and a neural network. Disagreements between models tend to cluster around candidates in the middle of the predicted probability distribution, where the race is genuinely uncertain and small differences in model assumptions produce different classifications. These borderline cases are precisely where the three-model consensus approach adds the most value: rather than committing to a single prediction, we flag them as uncertain and let the reader weigh the evidence accordingly.

5.2 State-Level Illustration: Paraná and Rio Grande do Sul

To illustrate the predictions in concrete terms, [Table 5](#) and [Table 6](#) show the predicted elected candidates for Paraná (PR) and Rio Grande do Sul (RS), respectively, across all three models. Both states are mid-sized districts (PR: 30 seats; RS: 31 seats) with competitive multi-party races. Predictions are generated by three machine-learning models: Multilayer Perceptron (MLP), Gradient Boosting (GB), and Logistic Regression (LR). Within each table, candidates are ranked by the **Avg.** column, which reports the average predicted probability of election across the three models — this ranking determines who is assigned a predicted seat in each district, with the top M_d candidates (where M_d is the district magnitude) classified as predicted winners. The **Ag.** column indicates how many of the three models independently predict the candidate as elected: candidates with **Ag. = 3** are predicted winners by all three algorithms and represent our most confident forecasts, while those with **Ag. = 1** or **Ag. = 2** are the genuinely uncertain races where model disagreement signals

electoral volatility. The complete predictions for all 27 states are available in the supplementary material.

Table 5: Predicted Elected Candidates — Paraná (PR), 2026 (30 seats)

Rank	Candidate	Party	MLP	GB	LR	Avg.	Ag.
1	Carol Dartora	PT	0.988	0.856	0.933	0.926	3
2	Diego Garcia	UNIÃO	0.956	0.874	0.945	0.925	3
3	Luisa Canziani	UNIÃO	0.961	0.846	0.921	0.909	3
4	Beto Preto	PSD	0.944	0.828	0.895	0.889	3
5	Luciano Ducci	PSB	0.947	0.770	0.915	0.877	3
6	Leandre	PSD	0.961	0.735	0.920	0.872	3
7	Cristina Silvestri	PP	0.954	0.859	0.701	0.838	3
8	Felipe Francischini	PODE	0.896	0.799	0.741	0.812	3
9	Delegado Matheus Laiola	UNIÃO	0.953	0.440	0.930	0.774	2
10	Karime Fayad	MDB	0.947	0.730	0.599	0.759	3
11	Roberto Requião	PDT	0.879	0.604	0.779	0.754	3
12	Paulo Litro	UNIÃO	0.791	0.716	0.738	0.748	3
13	Luiz Nishimori	PSD	0.887	0.614	0.713	0.738	3
14	Marco Brasil	PSD	0.842	0.769	0.562	0.724	3
15	Dilceu Sperafico	PP	0.855	0.687	0.610	0.717	3
16	Aliel Machado	PV	0.956	0.199	0.936	0.697	2
17	Welter	PT	0.781	0.789	0.405	0.658	2
18	Tadeu Veneri	PT	0.782	0.765	0.407	0.651	2
19	Aline Sleutjes	PL	0.941	0.485	0.525	0.650	2
20	Carlise Kwiatkowski	PL	0.941	0.485	0.524	0.650	2
21	Pedro Lupion	REPUBLICANOS	0.763	0.615	0.531	0.636	2
22	Zeca Dirceu	PT	0.804	0.584	0.496	0.628	3
23	Beto Richa	PSDB	0.867	0.449	0.548	0.621	2
24	Sergio Souza	MDB	0.742	0.688	0.404	0.611	2
25	Rosangela Moro	PL	0.944	0.270	0.559	0.591	2
26	Stephanes	PL	0.756	0.521	0.477	0.585	1
27	Tião Medeiros	PP	0.781	0.241	0.675	0.566	1
28	Naiany Hruschka	PP	0.924	0.383	0.368	0.558	2
29	Luciano Alves	PODE	0.755	0.558	0.278	0.530	1
30	Marlei Fernandes	PT	0.834	0.546	0.151	0.510	2

Notes: MLP = Multilayer Perceptron; GB = Gradient Boosting; LR = Logistic Regression. Avg. = average predicted probability across the three models. Ag. = number of models (out of 3) predicting election. Ranked by Avg. Candidates predicted by only 2 models (ranks 9, 16, 17, 18, 19, 24, 25, 27, 28, 29) and by only 1 model (rank 30) represent the most uncertain seats in the district.

Table 6: Predicted Elected Candidates — Rio Grande do Sul (RS), 2026 (31 seats)

Rank	Candidate	Party	MLP	GB	LR	Avg.	Ag.
1	Denise Pessôa	PT	0.988	0.913	0.930	0.944	3
2	Covatti Filho	PP	0.958	0.889	0.953	0.933	3
3	Bibo Nunes	PL	0.959	0.877	0.962	0.933	3
4	Marcelo Moraes	PL	0.959	0.877	0.962	0.933	3
5	Carlos Gomes	REPUBLICANOS	0.953	0.879	0.931	0.921	3
6	Maria Do Rosario	PT	0.988	0.786	0.921	0.898	3
7	Giovani Cherini	PL	0.959	0.762	0.963	0.895	3
8	Afonso Motta	PDT	0.931	0.877	0.844	0.884	3
9	Afonso Hamm	PP	0.946	0.781	0.901	0.876	3
10	Bohn Gass	PT	0.925	0.871	0.816	0.871	3
11	Marcon	PT	0.922	0.864	0.805	0.864	3
12	Fernanda Melchionna	PSOL	0.988	0.702	0.846	0.845	3
13	Luciano Azevedo	PSD	0.929	0.754	0.840	0.841	3
14	Heitor Schuch	PSD	0.927	0.754	0.834	0.838	3
15	Danrlei	PSD	0.926	0.748	0.833	0.836	3
16	Lucas Redecker	PSD	0.936	0.690	0.864	0.830	3
17	Busato	UNIÃO	0.933	0.698	0.854	0.828	3
18	Any Ortiz	PP	0.960	0.580	0.908	0.816	3
19	Daiana Santos	PCdoB	0.987	0.791	0.669	0.816	3
20	Alceu Moreira	MDB	0.946	0.594	0.902	0.814	3
21	Alexandre Lindenmeyer	PT	0.929	0.673	0.835	0.812	3
22	Osmar Terra	PL	0.818	0.594	0.844	0.752	2
23	Fran Bayer	REPUBLICANOS	0.935	0.546	0.721	0.734	2
24	Maurício Marcon	PL	0.824	0.716	0.486	0.675	3
25	Reginete Bispo	PT	0.914	0.743	0.289	0.649	2
26	Daniel Trzeciak	PSDB	0.866	0.505	0.539	0.637	2
27	Pompeo De Mattos	PDT	0.797	0.613	0.469	0.626	2
28	Pedro Westphalen	PP	0.778	0.440	0.655	0.624	1
29	Laura Sito	PT	0.925	0.551	0.328	0.601	2
30	Onyx Lorenzoni	PP	0.781	0.241	0.675	0.566	1
31	Enfermeira Fabia Richter	PSD	0.922	0.401	0.356	0.560	2

Notes: MLP = Multilayer Perceptron; GB = Gradient Boosting; LR = Logistic Regression. Avg. = average predicted probability across the three models. Ag. = number of models (out of 3) predicting election. Ranked by Avg. Candidates predicted by only 2 models (ranks 23, 25, 27, 30) and by only 1 model (ranks 28, 31) represent the most uncertain seats in the district.

5.3 Full Predictions

The complete list of predicted elected candidates across all 27 states and all three models is publicly available. The file contains, for each of the 513 predicted seats: candidate name, party, state, predicted probability from each of the three models and a consensus indicator showing how many models agree on the prediction.¹

5.4 Parties outcome predictions

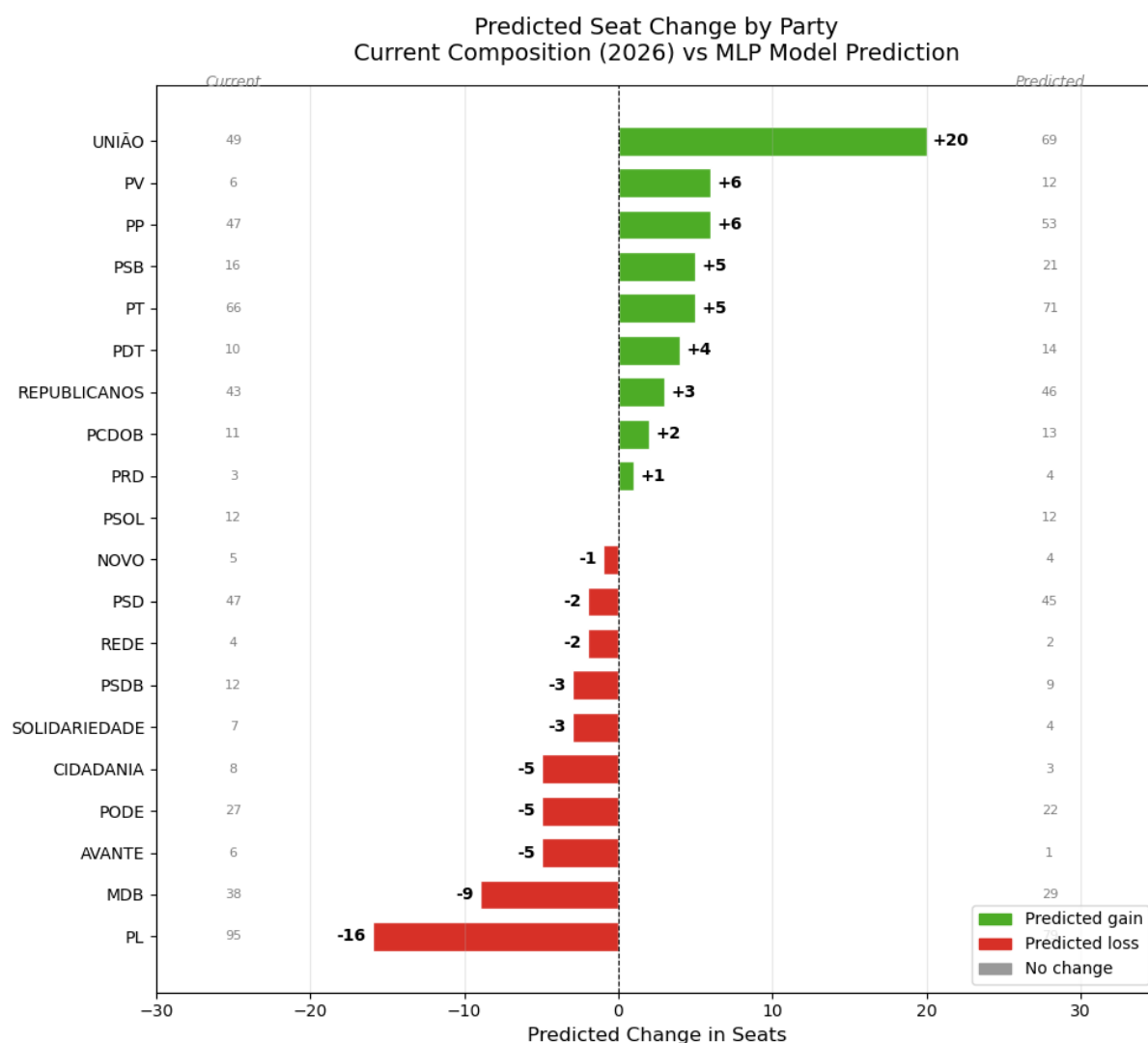
[Figure 5](#) displays the predicted change in seat share for each party, comparing the current composition of the Chamber (as of mid-2026, following the *janela partidária* which closed on April 3, 2026) with the MLP model's predictions for the 2026 election. The most striking prediction is the substantial gain forecast for **União Brasil** (+20 seats, from 49 to 69), **PP** (+9) and **Republicanos** (+6) are also predicted to gain seats. On the left, **PT** (+5), **PSB** (+5) and **PV** (+6), and **PDT** (+4) show modest predicted gains.

On the losing side, the model predicts **PL** as the biggest loser (-16 seats, from 95 to 79), which would still leave it as the largest individual party but with a meaningfully reduced delegation. **MDB** (-9) and **AVANTE** (-5) are also predicted to lose substantially, **PSOL** is the only party predicted to hold its seats exactly.

These predictions should be interpreted with caution for two reasons. First, the model was trained on candidate-level features and does not explicitly model party-level dynamics such as coalition strategies or the effects of the *federações*. Second, the comparison baseline reflects the current composition *after* the 2026 party switching window, meaning some of the predicted “losses” for parties like PL may partly reflect deputies who already switched parties and are now running under a different label — an effect our model captures through incumbency status but not through party affiliation directly.

¹ All the predictions are available here: [Download Link](#)

Figure 5: Predicted Seat Change by Party (national aggregate)

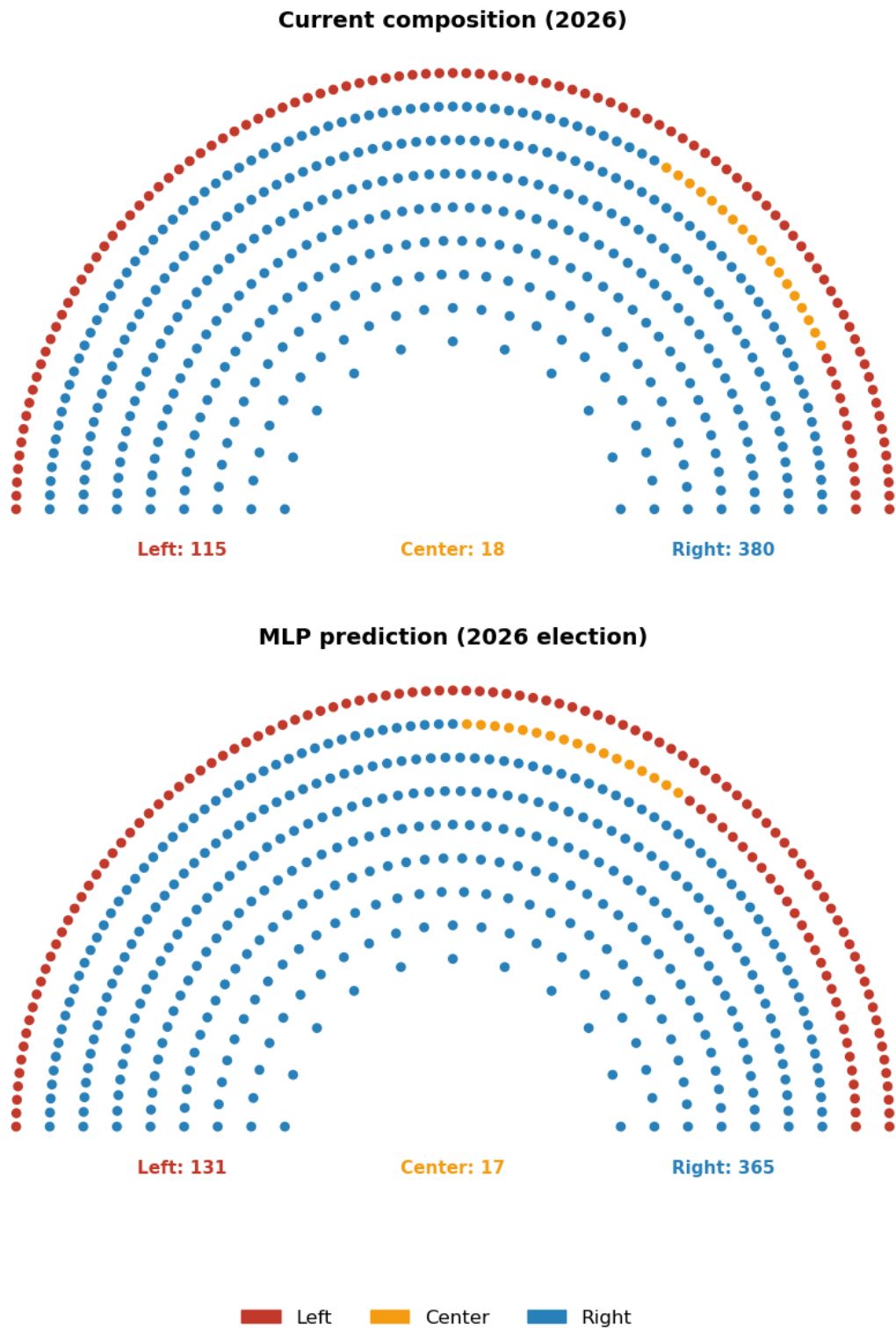


Predictions from the MLPClassifier. Source: elaborated by the authors.

Finally, considering the parties' ideological categorization made by (Bolognesi et al. 2022), we present the possible variation of ideological blocs based on the MLPClassifier predictions (Figure 6). The two hemicycles displayed reveal a Chamber that is already overwhelmingly right-leaning, and the MLP model predicts minimal structural change in this regard. In the current composition, the right bloc holds 380 of 513 seats (74.1%), while the left holds 115 (22.4%) and the center just 18 (3.5%). The predicted composition for 2026 shows a marginal shift: the right bloc is forecast to lose 15 seats (from 380 to 365), while the center (PV, CIDADANIA, REDE) loses 1 seats (from 18 to 17) and the left gains 16 (from 115 to 131). These changes are small enough to suggest that the ideological balance of the Chamber is unlikely to shift meaningfully regardless of which candidates win the contested seats. The Brazilian Chamber of Deputies, under any plausible outcome consistent with our model, will remain a right-dominant lower house.

Figure 6: Predicted Seat Change by ideological bloc (national aggregate)

Brazilian Chamber of Deputies — Seat Distribution by Ideological Bloc
Party categorization source: Bolognesi et al. (2023). Elaborated by the authors.



Predictions from the MLPClassifier. Source: elaborated by the authors.

6 Limitations and Caveats

Several limitations should be kept in mind when interpreting these predictions:

Training data does not include 2026. Models are trained exclusively on 2010–2022 data. If the 2026 electoral environment differs substantially from past cycles — due to new party configurations, the impact of social media mobilization, or macroeconomic shocks — model predictions may be less accurate.

Campaign revenue data is partial. The 2026 campaign finance data used here reflects TSE partial disclosures available as of late September 2026. Final reported revenues may differ, and candidates who report revenue late in the cycle are coded as having zero resources, which may underestimate their competitiveness.

Conservative by design. By restricting predictions to competitive candidates, we are implicitly assuming that the eventual winners will come from this pool. In practice, a small number of seats may be won by candidates who did not meet the 1% quota threshold in past elections. The predictions should therefore be interpreted as describing the most likely winners among competitive candidacies, not as an exhaustive prediction of all 513 seats.

7 Conclusion

This paper demonstrates that machine-learning models trained on historical electoral data can generate informative predictions for Brazilian Chamber of Deputies elections. Using data from four electoral cycles (2010–2022) and restricting the analysis to competitive candidates, we find that campaign revenue, incumbency status, gender, educational attainment, and ideological family are consistent and strong predictors of electoral success. The models achieve AUC values above 0.90 (Table 3) on held-out data, suggesting strong generalizability, which should be sustained at the 2026 election.

Among the models tested, the MLP Neural Network and Gradient Boosting Classifier perform best overall, while Logistic Regression remains competitive and offers the advantage of interpretable coefficients. All three models are used to generate 2026 predictions, with seats awarded to the top-ranked candidates within each district.

The results will be evaluated against official TSE results following the October 4 election, providing an out-of-sample test of the models' predictive validity in a real future election.

References

- Ames, Barry. 1995. "Electoral Strategy Under Open-List Proportional Representation." *American Journal of Political Science* 39 (2): 406–33. <https://doi.org/10.2307/2111619>.
- Ames, Barry. 2001. *The Deadlock of Democracy in Brazil*. The University of Michigan Press.
- Bolognesi, Bruno, Ednaldo Ribeiro, and Adriano Codato. 2022. "Uma Nova Classificação Ideológica Dos Partidos Políticos Brasileiros." *Dados* 66 (September): e20210164. <https://doi.org/10.1590/dados.2023.66.2.303>.
- Carey, John M, and Matthew Soberg Shugart. 1995. "Incentives to Cultivate a Personal Vote: A Rank Ordering of Electoral Formulas." *Electoral Studies* 14 (4): 417–39. [https://doi.org/10.1016/0261-3794\(94\)00035-2](https://doi.org/10.1016/0261-3794(94)00035-2).

- Cervi, Emerson Urizzi, Adriano Codato, Renato Perissinotto, and Luiz Domingos Costa. 2015. "Dinheiro, Profissão e Partido: A Vitória Na Eleição Para Deputado Federal No Brasil Em 2010." *Revista Sociedade e Estado* 30 (1): 17. <https://doi.org/10.1590/S0102-69922015000100011>.
- Cheibub, José Antônio, Andrea Junqueira, and Thiago de Miranda Queiroz Moreira. 2023. *Preventing a War of All Against All: Geographic Sorting in in OLPR Electoral Systems*.
- Cheibub, José Antonio, and Gisela Sin. 2020. "Preference Vote and Intra-Party Competition in Open List PR Systems." *Journal of Theoretical Politics* 32 (1): 70–95. <https://doi.org/10.1177/0951629819893024>.
- Floriano Ribeiro, Pedro. 2013. "Organização e Poder Nos Partidos Brasileiros." *Revista Brasileira de Ciência Política* 10 (April): 225–65. <https://doi.org/http://dx.doi.org/10.1590/S0103-33522013000100007>.
- Jacobson, Gary C. 1978. "The Effects of Campaign Spending in Congressional Elections." *The American Political Science Review* 72 (2): 469–91. <https://doi.org/10.2307/1954105>.
- Jacobson, Gary C. 2009. "Measuring Campaign Spending Effects in US House Elections." In *Capturing Campaign Effects*, 4th edition. The University of Michigan Press.
- Jairo Junckes, Ivan, Rodrigo Rossi Horochovski, Neilor Fermino Camargo, Edson Armando Silva, and Alides Baptista Chimin Junior. 2019. "Poder e Democracia: Uma Análise Da Rede de Financiamento Eleitoral Em 2014 No Brasil." *Revista Brasileira de Ciências Sociais* 34 (100): 1–22. <https://doi.org/http://dx.doi.org/10.1590/3410006/2019>.
- Janusz, Andrew, Sofi-Nicole Barreiro, and Erika Cintron. 2022. "Political Parties and Campaign Resource Allocation: Gender Gaps in Brazilian Elections." *Party Politics* 28 (5): 854–64. <https://doi.org/10.1177/13540688211018424>.
- Janusz, Andrew, and Luiz Augusto Campos. 2021. "Candidate Advertisements and Afro-Brazilian Political Marginalization." *Latin American Research Review* 56 (4): 761–78. <https://doi.org/10.25222/larr.888>.
- Mainwaring, Scott, Timothy J. Power, and Fernando Bizarro. 2018. "The Uneven Institutionalization of a Party System: Brazil." In *Party Systems in Latin America: Institutionalization, Decay, and Collapse*, 1st ed., edited by Scott Mainwaring. Cambridge University Press.
- Marcelino, Daniel. 2010. "Sobre Dinheiro e Eleições: Um Estudo Dos Gastos de Campanha Para o Congresso Nacional Em 2002 e 2006." Dissertação de {Mestrado}, Universidade de Brasília (UnB). <https://repositorio.unb.br/handle/10482/6669>.
- Nicolau, Jairo. 2007. "The Open-List Electoral System in Brazil." *Dados* 3 (SE): 0–0. http://socialsciences.scielo.org/scielo.php?script=sci_abstract&pid=S0011-52582007000100003&lng=en&nrm=iso&tlng=en.
- Peixoto, Vitor de Moraes. 2010. "Eleições e Financiamento de Campanhas No Brasil." Tese de {Doutorado}, Instituto Universitário de Pesquisas do Rio de Janeiro - IUPERJ.
- Peixoto, Vitor de Moraes. 2012. "Impacto Dos Gastos de Campanhas Nas Eleições Legislativas de 2010: Uma Análise Quantitativa." *8º Encontro Nacional Da ABCP* (Gramado, RS, Brasil), 1–32.
- Perissinotto, Renato, and Fernanda Fernandes Veiga. 2014. "Profissionalização Política, Processo Seletivo e Recursos Partidários: Uma Análise Da Percepção Dos Candidatos Do PT, PMDB, PSDB e DEM Nas Eleições Para Deputado Federal de 2010." *Opinião Pública* 20 (1): 49–66. <https://doi.org/10.1590/S0104-62762014000100003>.
- Sacchet, Teresa, Hannah Maruci Aflalo, Marcus Vinícius C. Alves, and Vanilda Souza Chaves. 2025. "Não é Só Sobre Dinheiro: Diferentes Dimensões Do Financiamento Eleitoral Em Perspectiva de Gênero e Raça No Brasil." *Revista Brasileira de Ciência Política* 44: e291774. <https://doi.org/https://doi.org/10.1590/0103-3352.2025.44.291774pt>.

- Sousa Braga, Maria do Socorro, Luciana Fernandes Veiga, and Angel Miríade. 2009. "Recrutamento e Perfil Dos Candidatos e Dos Eleitos à Câmara Dos Deputados Nas Eleições de 2006." *Revista Brasileira de Ciências Sociais* 24 (70): 124–42. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-69092009000200008&lng=en&nrm=iso&tlng=pt.
- Speck, Bruno Wilhelm. 2016. "Game over: Duas Décadas de Financiamento de Campanhas Com Doações de Empresas No Brasil." *Revista de Estudos Brasileños* 3 (4): 125–35. <https://reb.universia.net/article/view/1834/game-over-veinte-anos-financiacion-campanas-empresas-privadas-brasil>.
- Tarouco, Gabriela da Silva, and Rafael Machado Madeira. 2015. "Os Partidos Brasileiros Segundo Seus Estudiosos: Análise de Um Expert Survey." *Civitas - Revista de Ciências Sociais* 15 (March): e24–39. <https://doi.org/10.15448/1984-7289.2015.1.18077>.
- Zucco, Cesar, and Timothy J. Power. 2024. "The Ideology of Brazilian Parties and Presidents: A Research Note on Coalitional Presidentialism Under Stress." *Latin American Politics and Society* 66 (1): 178–88. <https://doi.org/10.1017/lap.2023.24>.

Autoria

Akira Pinto-Medeiros

É doutorando em Ciência Política pela Universidade de Pittsburgh (Pensilvânia, EUA), mestre em Ciência Política pela Universidade de São Paulo (USP), e especialista em Legislativo e Democracia no Brasil pela Escola do Parlamento da Câmara Municipal de São Paulo.

Bruno Wilhelm Speck

Bruno Wilhelm Speck é professor do Departamento de Ciência Política da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo (FFLCH-USP).

Normas para colaboração



○ **ReDem Working paper do Instituto Nacional de Ciência e Tecnologia (INCT) Representação e Legitimidade Democrática (ReDem)**

aceita somente notas de pesquisa originais. Elas devem apresentar resultados substantivos de pesquisas empíricas a partir da análise de dados ainda não publicados.

Os relatos de pesquisas devem conter até 5 mil palavras. A decisão sobre a sua publicação cabe à Editora a partir da avaliação de dois pareceristas.

O resumo dessas notas de pesquisa deve ser redigido no formato IMRAD (introdução, materiais e métodos, resultados e discussão e deve destacar as conclusões). O título da nota de pesquisa deve conter até 100 caracteres com espaços. Cada nota de rodapé deve conter no máximo 400 caracteres com espaços. As referências bibliográficas utilizadas serão apresentadas ao final da nota de pesquisa, listadas em ordem alfabética obedecendo ao padrão APA 2019 (7ª edição).

As notas de pesquisa devem ser enviadas por e-mail ao seguinte endereço coordenacao@redem.tec.br aos cuidados da Editora.



Copyright© 2026
Instituto Nacional de Ciência e Tecnologia
Representação e Legitimidade Democrática (INCT-ReDem)

**ReDem Working paper. Instituto Nacional de Ciência e
Tecnologia Representação e Legitimidade Democrática**

Editora: Maiane Bittencourt (UFPR; INCT ReDem)

Conselho editorial: Adriano Codato (UFPR), Bruno Bolognesi (UFPR), Carolina de Paula (UERJ), Ednaldo Ribeiro (UEM), Graziella Testa (FGV), Jaqueline Zulini (FGV), Karolina Roeder (UNINTER), Luciana Veiga (UNIRIO/FGV), Maria Alejandra Nicolás (UNILA), Maria Cecilia Eduardo (UFPR), Nara Pavão (UFPE), Renato Perissinotto (UFPR)

Financiamento:

CNPq. Processo n. 406649/2022-7
Fundação Araucária INC2024191000001

**Instituto Nacional de Ciência e Tecnologia
Representação e Legitimidade Democrática**

Universidade Federal do Paraná
Departamento de Ciência Política
Rua General Carneiro, 460, sala 515
Curitiba - PR - BRASIL CEP 80060-150
E-mail: inct.redem@gmail.com ■ URL: <https://redem.tec.br/>

Direitos e permissões

O texto e os dados nesta publicação podem ser reproduzidos desde que a fonte seja citada.

O *Working Paper ReDem* dissemina resultados de trabalhos em andamento para incentivar a troca de ideias. Os documentos são assinados pelos(as) autores(as) e devem ser citados de acordo com a referência abaixo.

Os resultados, interpretações e conclusões expressos no working paper são de responsabilidade das(os) autoras(es) e não refletem necessariamente as visões do INCT ReDem.

Os **ReDem Working papers** estão disponíveis online [<https://redem.tec.br/>]. Assinaturas podem ser feitas por e-mail coordenacao@redem.tec.br

Como citar

Pinto-Medeiros, A., & Speck, B. W. (2026). Before the Votes Are Cast: Electoral Forecasting in OLPR. Predicting the 2026 Brazilian Chamber of Deputies Election using Machine Learning algorithms. *ReDem Working Paper*, INCT ReDem, 3(6), outubro.

Rights and Permissions

The text and data in this publication may be reproduced as long as the source is cited.

The ReDem Working Paper disseminates the findings of work in progress to encourage the exchange of ideas. The papers are signed by the authors and should be cited accordingly.

The findings, interpretations, and conclusions expressed are the responsibility of the authors and do not necessarily reflect the views of the INCT ReDem.

Working papers are available online [<https://redem.tec.br/>] and subscriptions can be requested by email to coordenacao@redem.tec.br

How to cite

Pinto-Medeiros, A., & Speck, B. W. (2026). Before the Votes Are Cast: Electoral Forecasting in OLPR. Predicting the 2026 Brazilian Chamber of Deputies Election using Machine Learning algorithms. *ReDem Working Paper*, INCT ReDem, 3(6), October.