Journal of Production, Operations and Systems Management

e-ISSN: 1984-2430

Open Access

Responsible Editor: Professor Hermes Moretti Ribeiro da Silva

Received: 04/04/23 Approved: 15/09/23

PREDICTING SATISFACTION WITH DEMOCRACY IN BRAZIL CONSIDERING DATA FORM AN OPINION SURVEY

PREDIÇÃO DA SATISFAÇÃO EM RELAÇÃO À DEMOCRACIA NO BRASIL CONSIDERANDO DADOS DE PESQUISA DE OPINIÃO

Douglas Martins de Souza Rosa[1]
Bruno Samways dos Santos[2]
Rafael Henrique Palma Lima[3]

ABSTRACT

Purpose – This article compared machine learning algorithms in the context of satisfaction with democracy in Brazil. The models were trained with data from the Latinobarómetro survey, a private non-profit institution.
Theoretical foundation – The Support Vector Classifier (SVC), Random Forest (RF), and Artificial Neural Networks (ANN) classification techniques were described, followed by evaluation metrics, such as accuracy, precision, recall, f1-score and the area under a receiver operating characteristic (auc-roc).
Methodology – The data set was cleaned and the questionnaire was reduced to variables related to the local democracy index (IDL). Then, attribute transformations were performed, hyperparameters were analyzed and subsets with different class balances were created to evaluate the performance of the classifiers in different scenarios. Also, the attributes that most contributed to satisfaction with democracy were analyzed.
Results – The best classifier was RF for the class of those dissatisfied with democracy, however, the ANN and SVC techniques obtained better results in the class of satisfied individuals. Evaluating the most important attributes for satisfaction with democracy, it was identified that they are related to the country's economic situation and political and governmental issues.
Research implications – The models created were mainly able to identify people dissatisfied with democracy. The most important variables in this context were economy performance, government, political positioning, and democracy. This indicates directions for future studies and enables the development of strategies to change the perception of dissatisfied individuals.
Originality/Value – Exploring data on democracy from the Latinobarómetro using machine learning techniques in the context of Brazil.
Keywords - Machine learning; Democracy; Classification; Satisfaction.

RESUMO

Objetivo – O presente artigo comparou algoritmos de aprendizagem de máquina no contexto da satisfação quanto à democracia no Brasil. Os modelos foram treinados com dados da pesquisa de opinião do Latinobarómetro, uma instituição privada sem fins lucrativos.
Fundamentação teórica – Foram descritas as técnicas de classificação *Support Vector Classifier* (SVC), *Random Forest* e Redes Neurais Artificiais (RNA), seguido pelas métricas de avaliação, como a acurácia, precisão, *recall* e *f1-score*.
Metodologia – Realizou-se a limpeza do conjunto de dados e reduziu-se o questionário para variáveis relacionadas com o índice de democracia local (IDL). Depois, foram feitas as transformações de atributos e criados subconjuntos com diferentes balanceamentos de classe para avaliar o desempenho dos classificadores em diferentes cenários. Também, analisou-se os atributos que mais contribuíram para a satisfação em relação à democracia.
Resultados – O melhor classificador foi o *Random Forest*, com resultados superiores aos outros métodos aplicados, principalmente para a classe específica de insatisfeitos com a democracia. Avaliando-se os atributos mais importantes para a satisfação quanto à democracia, foram identificados que eles estão relacionados com a situação econômica do país e assuntos de cunho político e governamental.
Implicações da pesquisa – Os modelos criados foram capazes de identificar, principalmente, as pessoas insatisfeitas com a democracia e as variáveis mais importantes neste contexto, oportunizando pesquisas sobre a relação entre economia, governo, posicionamento político e democracia, enquanto elaboram-se de possíveis estratégias para mudança de percepção dos insatisfeitos.
Originalidade/valor – Explorar os dados sobre democracia do Latinobarómetro a partir de técnicas de aprendizagem de máquina no contexto do Brasil.
Palavras-chave: Aprendizagem de máquina; Democracia; Classificação; Satisfação.

DOI: 10.15675/gepros.2965

[1] Federal Technological University of Paraná, Londrina Campus, Av. dos Pioneiros, 3131 - Jardim Morumbi, Londrina - PR, 86036-370, Brazil. Email: douglas0395@gmail.com. Orcid: https://orcid.org/0009-0007-3401-4215

[2] Email: brunosantos@utfpr.edu.br Orcid: https://orcid.org/0000-0001-7919-1724

[3] Email: rafaelhlima@utfpr.edu.br Orcid: https://orcid.org/0000-0002-9098-3025

1. INTRODUCTION

Knowledge Discovery in Databases (KDD) is the process of discovering non-trivial and potentially valuable patterns in a dataset. Raschka (2015), Géron (2019), and Tan et al. (2019) mention that various sectors of the economy can be impacted by studies involving the phases of KDD, such as marketing, education, finance, industry, and others, as many companies and institutions generate large amounts of data that can be explored. Some sectors analyze and interpret this data inefficiently, often manually, resulting in a slower exploration and interpretation process, thus delaying the time to obtain results.

With the volume and complexity of data increasing at a significant pace, developers are constantly challenged to create more efficient data manipulation programs (Yida Tao et al., 2021). Due to the growth in the amount of generated data, the use of technology for storing and interpreting it becomes essential, requiring the utilization of methodologies and algorithms for the data collection, transformation, exploration, data mining, and interpretation process.

Several algorithms have been developed based on statistical learning, machine learning (ML), and artificial intelligence (AI) to support and enable data mining (Plotnikova et al., 2020). One of the crucial phases of the KDD process, data mining, can be defined as the application of algorithms to extract patterns from databases, patterns that may be unknown or not yet identified and might be discovered after much time and manual effort (Bhojani & Bhatt, 2016). It is worth noting that this pattern extraction and relationship identification can be carried out on datasets considered large (Morabito, 2016).

One of the fields where data mining can be applied is in the area of sentiment analysis, also known as opinion mining. According to Bing Liu (2012), after the year 2000, this became a highly active research field, widely used in almost all sectors due to the large volume of opinion data generated for processing, making the use of mining techniques in this context quite appealing.

The specialized literature includes some related works that aim to apply ML techniques to study social phenomena using government databases or opinion surveys. For example, Nascimento, Barone, and Castro (2019) applied the Random Forest (RF), Support Vector Machines (SVM), and Artificial Neural Networks (ANNs) techniques to study data from the World Values Survey, which aims to assess the beliefs and values of the studied populations. The authors examined the degree of political participation of the population in advanced democracies compared to other countries. In another study, Kang et al. (2021) used the Decision Tree (DT) technique to investigate respondents' perceptions of the quality of federal health agencies in the United States, enabling the identification of key factors influencing satisfaction, with a focus on the fairness of procedures used and employee commitment.

Another example is the work conducted by Broderstadt (2023), who used the RF technique to study the main predictors of satisfaction with democracy in Europe using data from the European Social Survey. The authors found that satisfaction with the economy, fairness in government procedures, and the speed of addressing population issues were the primary predictors of satisfaction with democracy.

Questionnaires concerning the measurement of variables related to democracy, in general, are linked to economics and political sciences, with researchers interested in examining the determining factors and consequences of democratic transitions in national-level data, needing to choose an indicator that genuinely reflects the level of democracy in different countries (Gründler & Krieger, 2021). Applications of ML techniques have also been found in the field of political stability (Tsyganov, 2021) and election prediction in Latin America (Brito & Adeodato, 2023). However, there are no research studies related to the application of ML techniques for classifying citizens' satisfaction with democracy based on an opinion survey.

The objective of this study is to apply data mining techniques, more specifically machine learning classification algorithms, to study satisfaction with democracy in Brazil and gain insights into the factors that most influence this process. Data provided by Latinobarómetro, a nonprofit organization based in Chile that conducts public opinion surveys in 18 Latin American countries with over 20,000 respondents, were used. To select features and streamline the questionnaire, the Local Democracy Index (LDI) (Silva & Mizuca, 2021) was used as a reference.

The database provided by Latinobarómetro has been used in various academic research studies. For instance, Pecorari and Cuesta (2023) applied machine learning techniques to study the trust of respondents in the governments of Caribbean countries. Berry and Rodriguez (2010) used statistical techniques applied to the 2005 Latinobarómetro survey to investigate the primary factors of dissatisfaction with democracy. Marek (2021) applied statistical techniques to Latinobarómetro data from 2002 to 2010 to explore the correlation between the population's perception of police and national government corruption. Overall, there is a wide diversity in the use of this research data in academic literature (Power & Jamison, 2005; Ateca-Amestoy et al., 2014; Barredo Ibáñez, 2018; Saravia & Marroquín, 2023), but there are few studies that focus on the application of machine learning techniques and studies exclusively data from Brazil.

Following this introductory section, the remainder of the article has been divided into five additional sections. The second section elaborates on the methods used in this research and the evaluation metrics. The third section describes the Latinobarómetro dataset and the sequence of steps undertaken in the research development. The fourth section presents the results obtained from the techniques and the most important variables for the RF model, along with a brief analysis of the results. The fifth section discusses the most relevant variables based on the literature, followed by the contributions by the classification models and the limitations of this study. Finally, the sixth section succinctly concludes the article by summarizing what was accomplished and discovered in this research, pointing to potential future research directions.

2. LITERATURE REVIEW

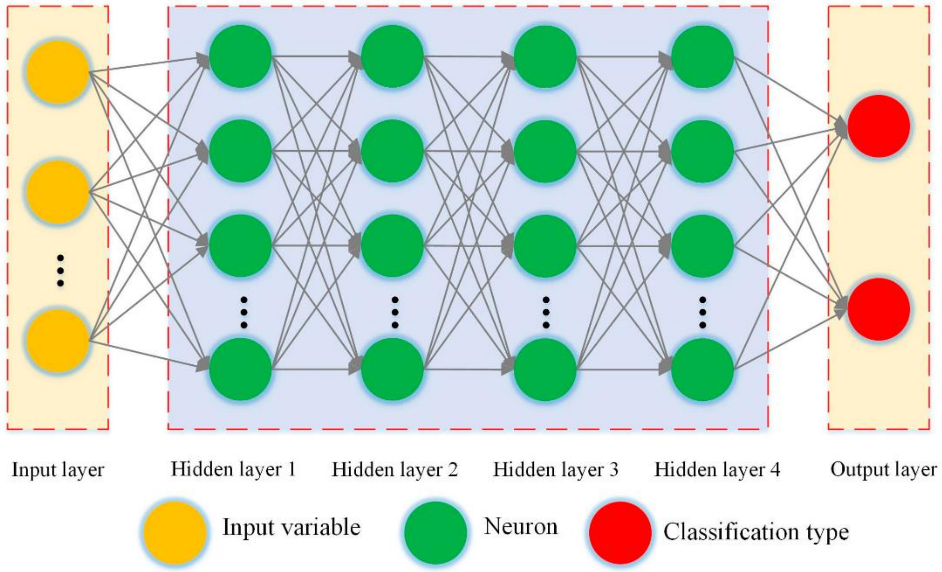
This section describes the classification techniques used and the metrics used to assess the performance of the models.

2.1 ML classification techniques

Artificial Neural Networks (ANNs) are so named because they emulate the functioning process of biological neurons in the brain to solve complex problems (Raschka, 2015). Artificial neurons exhibit a communication pattern similar to their biological counterparts, receiving information from other neurons, processing it, and transmitting it to other neurons through synaptic bridges (Aggarwal, 2018). Figure 1 depicts a typical network known as the Multilayer Perceptron.

Figure 1

Multilayer Perceptron ANN with four hidden layers



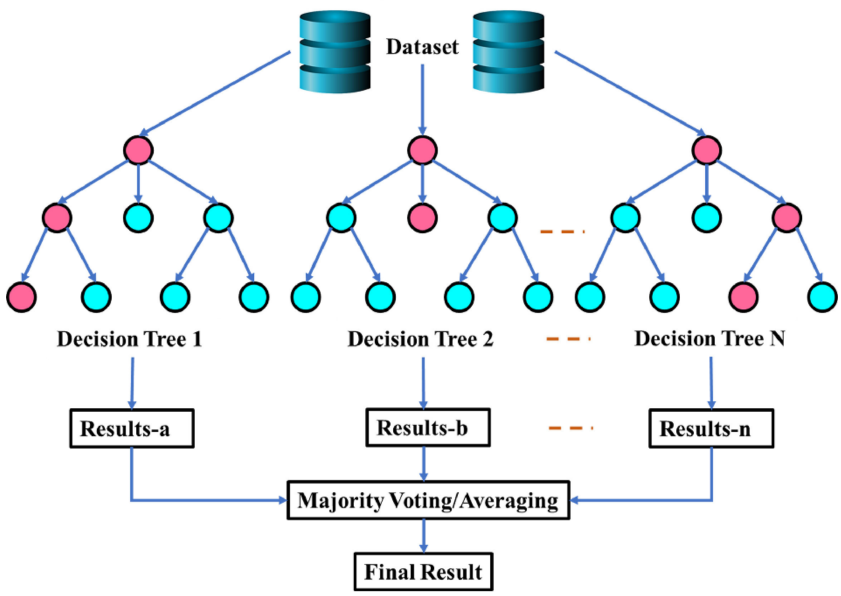
Source: Jiang et al. (2018).

The number of neurons in the input layer is determined by the quantity of available input attributes, while the number of neurons in the hidden layer and the number of hidden layers are configurable parameters of the model (Aggarwal, 2018, Géron, 2019).

Introduced by Breiman (2001), the Random Forest (RF) can be considered an ensemble model, meaning it employs multiple classifiers deemed "weak learners" and combines them to form a "strong learner" classifier. This method utilizes several Decision Trees (DT) structured in parallel, rendering it robust for joint prediction of the trees, often yielding superior results compared to a single decision tree (Speiser et al., 2019). Figure 2 illustrates an example of RF.

Figure 2

Random Forest for classification or regression

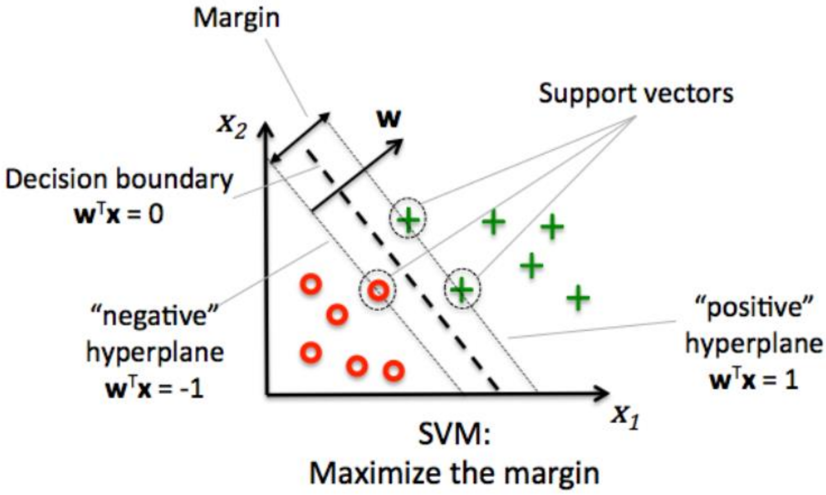


Source: Liu et al. (2021).

Finally, Support Vector Machines (SVM), or more specifically the Support Vector Classifier (SVC) used for classification tasks, aims to optimize (maximize) a margin during algorithm training. This margin is the distance between the decision boundary (hyperplane) and the training samples that are closest to this hyperplane, known as support vectors (Raschka, 2015). Figure 3 illustrates class separation with the creation of the margin and its hyperplanes.

Figure 3

Class separation with the creation of the margin and its hyperplanes.



Source: Raschka (2015).

2.2 Model evaluation metrics

The classification techniques applied in this paper were compared based on their precision, accuracy, recall, and F1-score. These metrics are derived from the confusion matrix (Witten & Frank, 2005), as shown in Table 1.

Table 1

Description of the confusion matrix

		Actual class	
		+	-
Predicted class	+	*True Positive (TP)*	*False Positive (FP)*
	-	*False Negative (FN)*	*True Negative (TN)*

Source: Valero-Carreras et al. (2023).

Equations 1 to 4 are derived from the confusion matrix and determine how accuracy, recall, precision, and f1-score are measured in this paper.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F1\text{-}score = 2 \times \frac{precision \times recall}{precision + recall} \tag{4}$$

Equation (1) demonstrates how well the model performs in classifying true instances, both negative and positive. Equation (2) evaluates performance from the perspective of the true positive class, indicating the level of accuracy in classifying positive instances among those that were indeed positive. Equation (3) assesses the level of correct classifications that the model identified as positive out of all those it classified as such. Equation (4) presents the harmonic mean between precision and recall, as these two metrics can exhibit conflicts. Lastly, the area under the receiver operating characteristic curve (AUC-ROC) was also employed as a metric, as it defines the discriminative power between two classes based on probability.

3. METHODOLOGY

This work involves the application of Machine Learning algorithms for classifying respondents into two categories: satisfied or dissatisfied with democracy. Thus, this study utilizes a supervised learning approach, as the dataset contains information about input variables and their corresponding output classes, enabling algorithms to learn from these samples as examples (Müller & Guido, 2017).

The steps of this research were inspired by similar studies that relied on questionnaire data to train classification algorithms and provide insights for addressing the original problem. For instance, the study conducted by Lee & Pak (2023) utilized data from a survey conducted in South Korea to assess suicidal ideation among adults using RF, SVM, and XGBoost techniques. The research stages were as follows:

1. Data collection and description, aimed at comprehending the available variables and data types.
2. Identification of the output variable, which will be used to determine the class of each respondent.
3. Application of techniques to address class imbalance, which becomes necessary when one of the classes is underrepresented in the dataset.
4. Selection of predictor variables to be considered in the training and testing stages of the ML models.
5. Training and testing of the ML algorithms chosen for the study.
6. Evaluation of the results obtained by the models.

The following subsections describe the main stages of this research, detailing the dataset used, as well as the procedures for data preparation, dataset balancing, training, and result evaluation.

3.1 Description of the dataset

The dataset was obtained from the Latinobarómetro website, specifically for the year 2020, which was the most recent dataset available when this work was carried out. Using the Latinobarómetro dataset, the questionnaire applied was examined, and each question was mapped according to its similarity to the areas of the Local Democracy Index (LDI) as detailed in Table 2. The LDI was developed by the *Atuação Institute* and encompasses the following dimensions and attributes.

Table 2

Dimensions and items evaluated by the Local Democracy Index (LDI)

Electoral Process	Political Participation	Democratic Culture
Democratic Choice	Wide Sense	Cognitive Dimension
Integrity	Strict Sense	Community Life
Inclusion		Norms and Values

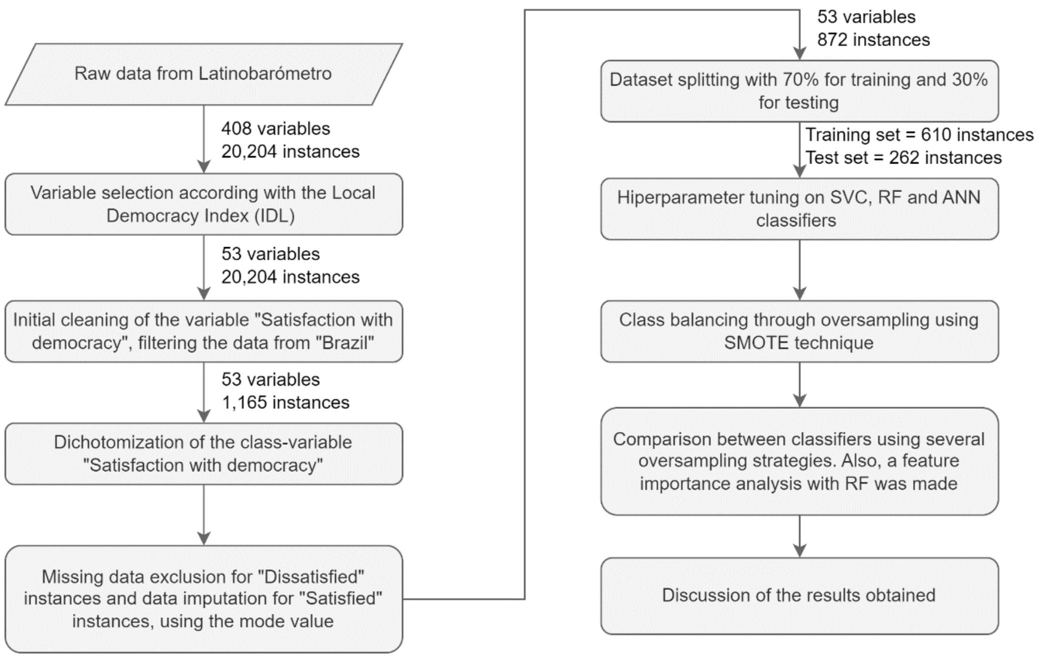
Civil Rights and Liberties	Functioning of Local Government
Civil Liberties	Checks and Balances
Freedom of Speech	Transparency
Economic Freedom	Control
Access to Justice	Responsiveness
Fair Treatment	Public Safety

3.2 Data preparation and training procedures

The data extracted from the Latinobarómetro dataset was preprocessed, transformed, evaluated, and subsequently discussed in the sequence illustrated in Figure 4.

Figure 4

Steps for preparing the dataset and training the models



The first stage of the study involved selecting questions from the Latinobarómetro database that aligned with the LDI metric. Out of a total of 408 questions, it was possible to reduce them to 53 questions (attributes). The Latinobarómetro institution, which conducted the survey, performs preliminary data cleaning and provides a codebook along with the survey information and the data collection instrument.

Even with this initial cleaning, a second stage was necessary, which involved removing data related to non-effective responses such as “don't know”, “no response”, “not applicable”, “not asked”, and “I don't know”, which remained in the dataset concerning the output variable

"Satisfaction with democracy". During this stage, only data relevant to the context of Brazil were considered, thus reducing the original dataset of 20,204 rows (instances) to 1,165 rows, in order to narrow the study to the specific context of Brazil.

In the third stage, the output variable, i.e., the level of satisfaction with democracy, was dichotomized, resulting in the following classes:

- Class 1 ("Dissatisfied"): included individuals who responded as "Not very satisfied" and "Not satisfied".
- Class 2 ("Satisfied"): in this case, respondents who answered the questionnaire as "Very satisfied" and "Quite satisfied" were considered.

In the fourth stage, all instances related to the "Dissatisfied" class with input attributes marked as "don't know", "no response", "not applicable", "not asked", and "I don't know" were excluded, reducing the count from 908 to 615 instances. As for those "Satisfied" with democracy, in order to maintain a balanced dataset, these non-effective responses were replaced by the mode of the data for their respective variables, retaining all 257 instances. For instance, in the case of the variable "Trust in the Judiciary", if an instance from the "Satisfied" class had responded with an input attribute like "I don't know", that response was replaced with the mode value for "Trust in the Judiciary". If it belonged to the "Dissatisfied" class, the entire instance was excluded.

In the fifth stage, Python 3 was used within the Jupyter Notebook environment, an interactive web-based computing platform. To implement the ML techniques, libraries such as Pandas, NumPy, Scikit-learn, and Imbalanced-learn were utilized. Upon importing the initial file in CSV (comma-separated values) format, the resulting data matrix consisted of 872 rows and 53 columns, meaning 872 instances and 53 attributes (or variables).

These data were then split into training and test sets, with 70% and 30% of the data, respectively. The training set was initially used for hyperparameter analysis across the three ML techniques. The parameters configured for each of the techniques align with those presented in Table 3. It is worth noting that the primary metric of interest was accuracy, with a k-fold cross-validation, where $k = 10$. The highlighted values are the ones selected through validation.

Table 3

Hyperparameters tested and selected

Technique	Hyperparameters: [values]	Description
SVC	*kernel*: [**'linear'**, 'rbf', 'sigmoid']	*Kernel* function
	C: [**0.5**, 1.0, 3]	Regularization parameter
RF	*n_estimators*: [50, 100, **500**, 1000]	Number of decision trees in the forest
	min_samples_split: [2, **5**, 10]	Minimum number of samples for a node to Branch
ANN	*hidden_layer_sizes*: [(10,), (20,), **(30,)**]	Number of neurons in the hidden layer
	max_iter: [200, 500, **1000**]	Maximum number of iterations in the *Backpropagation* algorithm

Table 3 displays the best parameters for each model; however, it is worth noting that for each of them, there was a maximum difference of 2 percentage points (p.p.) between the investigated configurations.

Following this stage, the training set was balanced due to the disparity in the number of instances for the output variable, using the Smote technique (Synthetic Minority Oversampling Technique), as proposed by the authors (Chawla et al., 2002). The Smote technique has the advantage of being easily combined with other techniques (Chawla, 2006).

The SMOTE approach is widely recognized in the literature and considered one of the most reliable techniques for class balancing (Waqar et al., 2021). It is extensively used for balancing binary class datasets (Blagus & Lusa, 2013), as is the case in this research. In Python 3, the *Imblearn* library was employed with different sampling strategies. It is important to note that this library interprets the sampling value (e.g., 0.5) as the proportion of instances the minority class should have in relation to the majority class, not with respect to the entire dataset. For example, in a scenario where the majority class has 1,000 instances, the minority class will have 500 instances when the sampling strategy is set to 0.5. If the value is 1.0, both classes will have the same number of instances due to an increase in examples of the minority class.

4. RESULTS AND DISCUSSION

The metrics described in Section 2 were applied to the assess the performance of the three classification algorithms used in this paper: SVC, RF, and ANN. Initially, a SMOTE sampling strategy with a value of 0.5 was used, resulting in an increase in instances related to the "Satisfied" class, which is the minority class in the original dataset. In this context, the metrics for overall classification and for each of the classes were analyzed (see Table 4).

Table 4

Results with sampling strategy of 0.5

Class	SVC			RF			ANN		
	Precision	Recall	*f1*	Precision	Recall	*f1*	Precision	Recall	*f1*
Dissatisfied	81%	86%	84%	78%	94%	85%	82%	87%	84%
Satisfied	60%	51%	55%	70%	34%	46%	62%	53%	57%
Accuracy	76%			77%			77%		
AUC-ROC	77%			80%			78%		

In Table 4, it is observed that the accuracies were very close among the three techniques analyzed, with only a 1 percentage point (p.p.) difference. For the "Dissatisfied" class, SVC and ANN achieved results above 80% for precision, recall, and f1-score. However, RF stood out with a 94% recall and an 85% f1-score, while ANN achieved a precision of 82%. Examining the minority class ("Satisfied"), it was rare to find results above 60%, with only RF and ANN achieving precision values of 70% and 62%, respectively. However, these values remained low when compared to the metrics for the "Dissatisfied" class. This raised questions about the effectiveness of a 50% balance for the minority class, leading to further analysis of variations in the balancing method in this article. Regarding the AUC-ROC metric, it was observed that RF also performed the best with an 80% score.

Through the method known as "feature importances" in the RF technique, it was determined which attributes had the most influence on the output classes, selecting those with an importance score above 0.04. The results of this analysis are displayed in Table 5.

It is important to emphasize that the "feature importance" method is based on the Gini impurity index. Initially, the Gini index is minimized when both probabilities are close to zero, and a total decrease in the Gini index is calculated after each node split in the tree. After this process is carried out for each tree in the forest, the average of the trees is computed, and the higher the average impurity decrease, the more important the input attribute becomes (Saarela & Jauhiainen, 2021).

Table 5

Features with importance above 0.4

Rank	Importance	Feature
1	0.1123	Satisfaction with the economic situation (overall)
2	0.0582	Confidence in the national government
3	0.0484	Age
4	0.0440	Self appointment in the left-right spectrum
5	0.0434	Future economic situation of the country

Considering the attributes selected by the method (Table 5), it can be observed that the attributes linked to the country's economic situation, both current and future, were the first and fifth most important attributes, respectively. These two are highlighted in gray in the table. Meanwhile, the cells with a blue background are related to political and governmental matters, encompassing both structural and ideological issues. Lastly, age is the only demographic attribute that emerged as one of the most relevant for the model.

After identifying the attributes that made the most significant contributions to the RF model, a new round of algorithm training was conducted, considering only the attributes from Table 4 and setting the "*sampling_strategy*" parameter to 0.5. The results obtained are presented in Table 6.

Table 6

Results using the most relevant features and sampling_strategy of 0.5

Class	SVC			RF			ANN		
	Precision	Recall	*f1*	Precision	Recall	*f1*	Precision	Recall	*f1*
Dissatisfied	77%	90%	83%	79%	90%	84%	80%	90%	84%
Satisfied	60%	36%	45%	62%	41%	49%	63%	43%	52%
Accuracy	74%			76%			76%		
AUC-ROC	77%			74%			75%		

The difference between the results obtained with the new models trained using only the most important attributes (Table 6), compared to the results from the original set of attributes (Table 4), is summarized in Table 7, expressed in percentage points (p.p.).

Table 7

Performance difference (in p.p) when comparing the second model against the first model

Class	SVC			RF			ANN		
	Precision	Recall	*f1*	Precision	Recall	*f1*	Precision	Recall	*f1*
Dissatisfied	-4	+4	-1	+1	-4	-1	-2	+3	0
Satisfied	0	-15	-10	-8	+7	+3	+1	-10	-5
Accuracy	-2			-1			-1		
AUC-ROC	0			-6			-3		

A slight improvement was observed for instances in the "Satisfied" class with the RF method, but a deterioration occurred in the cases of SVC and ANN. The recall still remains significantly lower when compared to the results obtained for the specific “Dissatisfied” class. This means that there seems to be no indication that a high number of input attributes is harming the performance of the models created, particularly for a better understanding of the “Satisfied” class.

In pursuit of improved performance and algorithm learning, other scenarios for “*sampling_strategy*” were attempted, first using only the original dataset (identified as Smote = 0), and then ranging from 0.5 to 1, with increments of 0.1. Table 8 presents the precision metric, which aims to measure, for a specific class, how accurately the model managed to correctly predict the instances.

Table 8

Precision for each value tested for sampling strategy

SMOTE	**Precision “Dissatisfied”**			**Precision “Satisfied”**		
0 - 1	**SVC**	**RF**	**ANN**	**SVC**	**RF**	**ANN**
0.0	80%	75%	82%	61%	69%	60%
0.5	81%	78%	82%	60%	70%	62%
0.6	84%	78%	82%	59%	70%	57%
0.7	85%	80%	82%	55%	71%	55%
0.8	85%	81%	82%	57%	69%	58%
0.9	85%	81%	82%	54%	67%	54%
1.0	85%	81%	84%	52%	63%	62%

It can be observed in Table 8 that the precisions of SVC and RF benefit from balancing through SMOTE, while ANN shows a rather specific improvement for the “Dissatisfied” class, but a decline for the opposite class. This highlights the impact of the growth of one class at the expense of the other.

Table 9 represents the variation in recall, which aims to measure, for a specific class, how well the algorithm is at correctly identifying a given output class in relation to the true classes, examining the proportion of samples from the class of interest that were classified correctly for the “Satisfied” and “Dissatisfied” groups.

Table 9

Recall for each sampling_strategy value tested

SMOTE	***Recall* “Dissatisfied”**			***Recall* “Satisfied”**		
0 - 1	**SVC**	**RF**	**ANN**	**SVC**	**RF**	**ANN**
0.0	88%	94%	85%	46%	24%	54%
0.5	86%	94%	87%	51%	34%	53%
0.6	83%	94%	83%	62%	37%	57%
0.7	78%	93%	81%	67%	42%	58%
0.8	79%	91%	83%	67%	46%	55%
0.9	77%	90%	80%	66%	49%	57%
1.0	74%	88%	84%	68%	51%	62%

Table 9 shows that there is a decline in recall for all three techniques concerning the “Dissatisfied” class, but there appears to be a noteworthy improvement for the “Satisfied” class. This demonstrates that one class is favored at the expense of the other, much like in precision.

Table 10 represents the variation in the f1-score, defined as the harmonic mean of precision and recall, for the “Satisfied” and “Dissatisfied” groups.

Table 10

F1-scores for each sampling_strategy value tested

SMOTE	*f1-score* “Dissatisfied”			*f1-score* “Satisfied”		
0 - 1	**SVC**	**RF**	**ANN**	**SVC**	**RF**	**ANN**
0.0	84%	84%	84%	53%	35%	57%
0.5	84%	85%	84%	55%	46%	57%
0.6	83%	85%	83%	61%	48%	57%
0.7	81%	86%	82%	61%	53%	56%
0.8	82%	86%	83%	61%	55%	56%
0.9	81%	85%	81%	59%	56%	55%
1.0	79%	84%	84%	59%	57%	62%

Table 10 summarizes the results for the f1-score, identifying that there is practically no systemic change in this metric, which is a consequence of the variations found in the previous precision and recall. The only technique that benefited from the increasing sample sizes was RF, with an increase of up to 22 p.p. for the “Satisfied” class, which is the minority in the original dataset.

Finally, Table 11 presents the variation in accuracy and AUC-ROC for the “Satisfied” and “Dissatisfied” groups based on different sampling strategy values.

Table 11

Accuracy and AUC-ROC for each sampling strategy value tested

SMOTE	Accuracy			AUC-ROC		
0 - 1	**SVC**	**RF**	**ANN**	**SVC**	**RF**	**ANN**
0.0	76%	75%	76%	77%	79%	77%
0.5	76%	77%	77%	77%	80%	78%
0.6	77%	77%	75%	77%	79%	74%
0.7	75%	78%	74%	78%	80%	75%
0.8	76%	78%	75%	77%	79%	77%
0.9	74%	78%	73%	78%	79%	74%
1.0	73%	77%	78%	78%	79%	77%

From the tables showing the variation of the SMOTE sampling strategy, including the analysis in Table 11, it is clear that there are some overall gains for the SVC and RF techniques in terms of accuracy, precision, and f1-score, while AUC-ROC remains relatively stable for these two techniques. In contrast, ANN does not show improvement when attempting to balance the minority class.

Based on the results in Table 5, it was possible to identify some of the factors that had the most influence on the RF algorithm's outcome regarding respondents' satisfaction with democracy. In this case, it was found that economic factors carry significant weight in the decision, followed by trust in the government, age, and political stance, which aligns with the research by Resende and Epitácio (2014). Their study suggests that a country's economic situation influences individual perceptions of satisfaction with democracy, along with public policies and their effective implementation. Additionally, it was noted that satisfaction with the democratic system is linked to factors related to the country's economic trajectory, such as the level of per capita income (Resende & Epitácio, 2014).

The techniques compared in this study are similar to those used by other authors such as Nascimento, Barone, and Castro (2019) and Broderstadt (2023), which allowed for the identification that the RF and SVM techniques are suitable for studying satisfaction with democracy using the Latinobarómetro dataset. Furthermore, varying the "sampling strategy" parameter of the SMOTE algorithm enables an improvement in the identification of individuals satisfied with democracy, which is relevant when the dataset is unbalanced with a low number of satisfied respondents, as they constitute the minority class in this study.

The performance on the training set was also evaluated to check for potential overfitting in the models. Implementing the three machine learning models for the training set classification revealed accuracies above 85% for SVC and RF, and above 90% for ANN. This indicates the possibility of overfitting in the training phase, making generalization more challenging. To address this issue, several approaches were adopted in this study: (i) Hyperparameter analysis; (ii) Attribute reduction; (iii) Class balancing. Another way to investigate this situation is by increasing the sample size, which may be possible for future studies using data from Latinobarómetro surveys considering previous years.

The results of this study are consistent with other related works. For instance, Broderstadt's study (2023) based on European data also identified that economic performance is one of the primary predictors of satisfaction with democracy. Therefore, this serves as an indication of voter expectations in democratic countries, where the population anticipates that democracy will facilitate economic prosperity.

Similar results were indicated by Sima and Huang (2023), who argue that economic prosperity can be better for democratic countries than autocratic ones, but only if there is already an adequate economic framework for democracy, placing these countries in what is termed "Strong Democracy." However, in an indicator containing information on income, education, natural resource dependence, and inequality, approximately 40% of democratization cases after 1960 are classified as "Weak Democracy," and these nations have a lower potential for government transparency, higher political corruption, and social instability (Sima & Huang, 2023). In this context, it is understood that the variables of greater importance for the RF model in the fields of politics and economics appear to be associated with satisfaction with democracy.

As a practical implication of this study, these models could focus on dissatisfied voters to identify the attributes of greatest importance for their dissatisfaction with democracy. A similar strategy was reported by Albright (2016) regarding the presidential elections in the United States of America for Donald Trump's first term, where data were used to identify voters who could be persuaded to change their vote in his favor. Therefore, dissatisfied voters in Brazil could be more properly targeted by electoral campaigns and candidates.

In addition to their use in electoral campaigns, this research highlights some of the most important variables in the context of democracy, which can be employed to enhance policies and development plans in the realms of civil liberties, particularly those related to economic freedom. It also pertains to political participation in a broader sense, creating opportunities to engage in political parties and municipal councils. Additionally, it underscores the importance of monitoring government activities, which should aim to increase transparency and policies across all three branches of government, thereby enhancing public trust.

A limitation of this study is that the models did not achieve satisfactory results for the "satisfied" class, even with the use of the SMOTE technique. It is possible that the size of the dataset may not be sufficient, and there may be conflicting responses among different individuals, which hinders classifier training and may result in potential overfitting. Thus, it is understood that a strategy for future work would be to consolidate datasets from different years to attempt to expand the training set, seeking to improve the learning of classifiers, especially in relation to the “satisfied” class.

5. CONCLUSION

The use of ML techniques on opinion survey datasets can help validate assumptions, generate a model that can predict new opinions, and uncover the factors that influence these opinions the most. In this regard, this study employed ML techniques to assess the satisfaction of respondents with democracy and to identify the factors that lead to satisfaction and dissatisfaction with democracy in Brazil.

This research identified that variables related to economic satisfaction, trust in the government, and age have a more significant influence on satisfaction with democracy than the respondent's political position in the right-left spectrum, based on the attribute values generated by the RF algorithm.

The classification models yielded promising results for the specific case of the “Dissatisfied” class. However, the “Satisfied” class obtained inferior results that need a better understanding by the machine learning models, especially with datasets that contain more examples of this class for the training phase. When compared to the “Dissatisfied” class, in one of the best models trained by ANN, for example, the precision was 62%, recall was 62%, and the f1-score was 62%, with an accuracy of 77% for the model, considering a SMOTE of 1.0, highlighting how the dataset's class imbalance influences the output.

Future research can use datasets from previous years to increase its size and potentially enhance the classifier performance, especially when classifying individuals who are “Satisfied” with democracy. It is also suggested that future work focuses on other Latin American countries to compare the results with the findings of this research, which only used data from Brazil.

REFERENCES

Aggarwal, C. C. (2018). Neural Networks and Deep Learning. In *Neural Networks and Deep Learning*. Springer International Publishing. https://doi.org/10.1007/978-3-319-94463-0

Albright, J. (2016, November 26). *How Trump’s campaign used the new data-industrial complex to win the election*. Https://Blogs.Lse.Ac.Uk/Usappblog/2016/11/26/How-Trumps-Campaign-Used-the-New-Data-Industrial-Complex-to-Win-the-Election/.

Ateca-Amestoy, V., Aguilar, A. C., & Moro-Egido, A. I. (2014). Social Interactions and Life Satisfaction: Evidence from Latin America. *Journal of Happiness Studies*, *15*(3), 527–554. https://doi.org/10.1007/s10902-013-9434-y

Barredo Ibáñez, D. (2018). Religious Commitment, Subjective Income, and Satisfaction towards the Functioning of Democracy in Latin America. A Mediation Analysis Model Based on Latinobarómetro. *Religions*, *9*(6), 198. https://doi.org/10.3390/rel9060198

Berry, B. J. L., & Rodriguez, O. S. T. (2010). Dissatisfaction with Democracy: Evidence from the Latinobarómetro 2005. *Journal of Politics in Latin America*, *2*(3), 129–142. https://doi.org/10.1177/1866802X1000200305

Bhojani, S., & Bhatt, N. (2016). Data Mining Techniques and Trends – A Review. *Global Journal for Research Analysis*, *5*(5), 252–254.

Blagus, R., & Lusa, L. (2013). SMOTE for high-dimensional class-imbalanced data. *BMC*

Bioinformatics, *14*(1), 106. https://doi.org/10.1186/1471-2105-14-106

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1007/9781441993267_5

Brito, K., & Adeodato, P. J. L. (2023). Machine learning for predicting elections in Latin America based on social media engagement and polls. *Government Information Quarterly*, *40*(1), 101782. https://doi.org/10.1016/j.giq.2022.101782

Broderstadt, T. S. (2023) An Empirical Evaluation of Explanations for Political System Support. Political Research Quarterly, 76 (3). https://doi.org/10.1177/10659129231156388

Chawla, N. V. (2006). Data Mining for Imbalanced Datasets: An Overview. In *Data Mining and Knowledge Discovery Handbook* (pp. 853–867). Springer-Verlag. https://doi.org/10.1007/0-387-25465-x_40

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Géron, A. (2019). *Hands-on: Machine Learning with Scikit-Learn, Keras & Tensorflow* (2nd ed.). O'Reilly Media.

Gründler, K., & Krieger, T. (2021). Using Machine Learning for measuring democracy: A practitioners guide and a new updated dataset for 186 countries from 1919 to 2019. *European Journal of Political Economy*, *70*, 102047. https://doi.org/https://doi.org/10.1016/j.ejpoleco.2021.102047

Jiang, W., He, G., Long, T., Ni, Y., Liu, H., Peng, Y., Lv, K., & Wang, G. (2018). Multilayer perceptron neural network for surface water extraction in landsat 8 OLI satellite images. *Remote Sensing*, *10*(5). https://doi.org/10.3390/rs10050755

Kang, I., Kim, N., Loh, W., & Bichelmeyer, B. (2021) A Machine-Learning Classification Tree Model of Perceived Organizational Performance in U.S. Federal Government Health Agencies. Sustainability, 13(18). https://doi.org/10.3390/su131810329

Lee, J., & Pak, T. (2022) Machine learning prediction of suicidal ideation, planning, and attempt among Korean adults: A population-based study. SSM - Population Health, 19.

Liu, Y., Esan, O. C., Pan, Z., & An, L. (2021). Machine learning for advanced energy materials. *Energy and AI*, *3*, 100049. https://doi.org/10.1016/j.egyai.2021.100049

Marek, J. (2021). The Public Perception of Police Corruption in Venezuela and its Effect on National Government. *Revista Sul-Americana de Ciência Política*, *1*(3), 1–21.

Morabito, V. (2016). *The Future of Digital Business Innovation*. Springer International Publishing. https://doi.org/10.1007/978-3-319-26874-3

Müller, A. C., & Guido, S. (2017) Introduction to Machine Learning with Python: a guide for data scientists, O'Reilly Media.

Nascimento, F.M., Barone, D., & Castro, H. C. (2019) Social Activism Analysis: An Application of Machine Learning in the World Values Survey. Proceedings of the MLDM, v. 1, pp. 28-38.

Pecorari, N., & Cuesta, J. (2023). Citizen Participation and Political Trust in Latin America and the Caribbean: A Machine Learning Approach. *Policy Research Wrking Papers*, 10335.

Plotnikova, V., Dumas, M., & Milani, F. (2020). Adaptations of data mining methodologies: a systematic literature review. *PeerJ Computer Science*, *6*, e267. https://doi.org/10.7717/peerj-cs.267

Power, T. J., & Jamison, G. D. (2005). Desconfiança política na América Latina. *Opinião Pública*, *11*(1), 64–93. https://doi.org/10.1590/S0104-62762005000100003

Raschka, S. (2015). *Python Machine Learning* (1st ed.). Packt Publishing Ltd.

Resende, R. C., & Epitácio, S. D. S. F. (2014). Desenvolvimento econômico e satisfação com a democracia: uma análise da América Latina. *Ciências Sociais Unisinos*, *50*(2). https://doi.org/10.4013/csu.2014.50.2.03

Saarela, M., & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. *SN Applied Sciences*, *3*(2), 272. https://doi.org/10.1007/s42452-021-04148-9

Saravia, A., & Marroquín, A. (2023). Political ideology mismatch as a determinant of the intention to migrate: evidence from Latin America. *International Journal of Social Economics*, *50*(1), 97–110. https://doi.org/10.1108/IJSE-01-2022-0028

Silva, D. R. de M., & Mizuca, H. D. de. (2021). *Índice de democracia local: estudos a partir da experiência de São Paulo* (1st ed.). Instituto Atuação.

Sima, D., & Huang, F. (2023). Is democracy good for growth? — Development at political transition time matters. *European Journal of Political Economy*, 102355. https://doi.org/10.1016/j.ejpoleco.2022.102355

Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, *134*, 93–101. https://doi.org/10.1016/j.eswa.2019.05.028

Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *Introduction to Data Mining* (2nd ed.). Pearson Prentice Hall.

Tsyganov, V. (2021). Supervised Machine Learning of Citizens and Political Stability. *IFAC-PapersOnLine*, *54*(13), 611–616. https://doi.org/10.1016/j.ifacol.2021.10.517

Valero-Carreras, D., Alcaraz, J., & Landete, M. (2023). Comparing two SVM models through different metrics based on the confusion matrix. *Computers & Operations Research*, *152*, 106131. https://doi.org/10.1016/j.cor.2022.106131

Waqar, M., Dawood, H., Dawood, H., Majeed, N., Banjar, A., & Alharbey, R. (2021). An

Efficient SMOTE-Based Deep Learning Model for Heart Attack Prediction. *Scientific Programming*, *2021*, 1–12. https://doi.org/10.1155/2021/6621622

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2nd ed.). Elsevier Ltd.

Yida Tao, Shan Tang, Yepang Liu, Zhiwu Xu, & Shengchao Qin. (2021). Speeding Up Data Manipulation Tasks with Alternative Implementations: An Exploratory Study. *ACM Transactions on Software Engineering and Methodology*, *30*(4), 1–28. https://doi.org/https://doi.org/10.1145/3456873