Analysis of economic satisfaction using machine learning models and explainable artificial intelligence

Análise da satisfação com a economia a partir de modelos de aprendizado de máquina e inteligência artificial explicativa

Luiz Fernando Menegazzo Ferreyra
https://orcid.org/0009-0008-9219-6560
Estudante de Engenharia de Produção. Universidade Tecnológica Federal do Paraná – Campus Londrina (UTFPR) – Brasil. luizferreyra@alunos.utfpr.edu.br

Yasser Bulaty Tauil
https://orcid.org/0009-0001-4804-2416
Estudante de Engenharia de Produção. Universidade Tecnológica Federal do Paraná – Campus Londrina (UTFPR) – Brasil. yasser@alunos.utfpr.edu.br

Helton Messias Adigneri
https://orcid.org/0000-0002-2652-6508
Bacharel em Engenharia de Produção. Universidade Estadual de Maringá – Campus Maringá (UEM) – Brasil. pg405633@uem.br

Bruno Samways dos Santos
https://orcid.org/0000-0001-7919-1724
Doutor em Engenharia de Produção e Sistemas. Universidade Tecnológica Federal do Paraná – Campus Londrina (UTFPR) – Brasil. brunosantos@utfpr.edu.br

Rafael Lima
https://orcid.org/0000-0002-9098-3025
Doutor em Engenharia de Produção. Universidade Tecnológica Federal do Paraná – Campus Londrina (UTFPR) – Brasil. rafaelhlima@utfpr.edu.br

ABSTRACT

The economic satisfaction of a nation can reflect citizens' perceptions of their government's performance, and machine learning models can help uncover non-trivial information from such data. In this context, this article aimed to analyze the satisfaction of Latin American citizens with their country's economy. To achieve this, six traditional classifier algorithms and four ensemble models were used, with a final application of an explainable method (SHapley Additive exPlanations, SHAP) to analyze the key factors contributing to economic satisfaction. The models were trained and tested on a dataset comprising data from the 2020 and 2023 Latinobarómetro surveys, totaling 27,600 instances in the final set. As a result, it was found that the Random Forest was the best individual model, while the stacking ensemble achieved the best performance in classifying between "satisfied" and "dissatisfied" citizens. The SHAP method revealed that "satisfaction with democracy" and "perception of the country's progress" are the main factors influencing economic satisfaction. This study offers insights for public managers on how to improve their citizens' economic satisfaction.
Keywords: Economic satisfaction. Machine learning. Explainable Artificial Intelligence. Latin America.

RESUMO

A satisfação econômica de uma nação pode refletir a percepção dos cidadãos sobre o desempenho de seus respectivos governos, e modelos de aprendizado de máquina podem auxiliar na descoberta de informações não triviais contidas em dados desta natureza. Neste sentido, o objetivo deste artigo foi analisar a satisfação dos cidadãos latino-americanos sobre a economia do seu país. Para isso, foram utilizados seis algoritmos classificadores tradicionais e mais quatro modelos *ensemble*, com a aplicação final de um método explicativo (*SHapley Additive exPlanations*, SHAP), analisando os principais fatores que contribuem para a satisfação econômica. Os modelos foram treinados e testados em um conjunto de dados composto pelos anos de 2020 e 2023 da pesquisa do Latinobarómetro, totalizando 27.600 instâncias no conjunto final. Como resultado, verificou-se que a Floresta Aleatória foi o melhor modelo individual, enquanto o *stacking ensemble* obteve o melhor desempenho para a classificação entre "satisfeitos" e "insatisfeitos". O método SHAP mostrou que a "satisfação com a democracia" e a "percepção sobre o progresso do país" sãos os principais fatores que influenciam na satisfação econômica. Este trabalho oferece caminhos nos quais gestores públicos podem atuar para a melhoria da satisfação econômica de seus cidadãos.

Palavras-chave: Satisfação econômica. Aprendizado de máquina. Inteligência Artificial Explicativa. América Latina.

Recebido em 31/08/2024. Aprovado em 04/11/2024. Avaliado pelo sistema *double blind peer review*. Publicado conforme normas da APA.
https://doi.org/10.22279/navus.v15.2006

1 INTRODUCTION

The advancement in understanding satisfaction involves recognizing the "object of satisfaction," as it is asserted that the satisfaction a person feels with life as a whole is distinct from specific satisfaction with work, marriage, or housing (Veenhoven, 1996). For instance, in the domain of consumption, satisfaction can be summarized as the attainment of expected quality in the purchase of products or the procurement of services (Martínez-Navalón *et al.*, 2021). Consumer satisfaction is addressed in various fields of knowledge, and regarding total quality, economic globalization, and strategic management, the concept permeates the entire organization (Bortolotti *et al.*, 2012).

The analysis of satisfaction has been researched since 1989, beginning with a study in Sweden to measure consumer satisfaction using the Swedish Customer Satisfaction Barometer (SCSB) (Bortolotti *et al.*, 2012). Data on how to measure and understand consumer and citizen satisfaction is highly relevant for strategic business and governmental planning, as this metric can guide which factors or variables are crucial in changing satisfaction, thereby improving people's well-being and institutional efficiency.

Regarding the general population of a nation, surveys on satisfaction with services within a country are key in evaluating a government, revealing judgments about the quality of services offered to citizens thus far. However, general satisfaction is multifactorial and subjective, making this task more complex (Van Ryzin, 2004). Therefore, machine learning (ML) methods can be useful in the objective and quantitative analysis of satisfaction data.

ML algorithms are being widely utilized in various fields to analyze satisfaction, including healthcare, products and services, economics, and education, among others. For example, authors such as Abdelkader *et al.* (2022), Chamorro-Atalaya *et al.* (2022), Langan and Harris (2024), and Liang and Jia (2023) applied ML in the context of education and teaching. In the field of products and services, Zaghloul *et al.* (2024) applied ML to evaluate e-commerce products, Li *et al.* (2024) and Noviantoro and Huang (2022) studied satisfaction in airline companies, and Joolfoo *et al.* (2022) in the telecommunications sector. Also, Polce *et al.* (2021), Zhang *et al.* (2021), Sabarmathi and Chinnaiyan (2019), and Kowalski (2017) employed ML in satisfaction analysis within the healthcare area.

In Latin America, data from the Latinobarómetro survey evaluates public opinion in 18 countries concerning democracy, economy, and society. With this publicly available dataset, recent studies by Pecorari and Cuesta (2023) applied ML techniques to analyze citizen participation and political trust, Rosa *et al.* (2023) classified democracy in Brazil, and Tauil *et al.* (2024) focused on economic classification, also applying classification algorithms.

In this context, the present article seeks to analyze the application of machine learning techniques to classify data concerning economic satisfaction in Latin American countries, including ensemble classifier models (bagging, boosting, voting, stacking) and the SHAP technique (SHapley Additive exPlanations) to evaluate predictor variables. In addition to the use of algorithms, the study also compares different years of questionnaire application in Latin America, aiming to identify potential variations in economic satisfaction across the continent.

Following this introductory section, the article comprises four main sections. Section 2 outlines concepts related to data mining, machine learning, ensemble classifiers, and interpretive models. The third section details the research sequence, including the treatment of the datasets used and the analysis of base algorithm hyperparameters. The fourth section presents the results obtained by the classifiers, offering comparisons and interpretations of predictor variables through the SHAP model. Finally, the conclusion and suggestions for future research are provided in Section 5.

2 DATA MINING AND MACHINE LEARNING

According to Yadav *et al*. (2022), data mining is essentially the process of discovering interesting patterns, models, and other types of knowledge in large datasets (Han *et al*., 2022). Mining is part of the Knowledge Discovery in Databases (KDD) process, which initially requires the selection, cleaning, and transformation of data, before applying a mining task. On the other hand, ML models necessitate the application of an algorithm to learn from the data, with the main types being supervised learning, unsupervised learning, and reinforcement learning.

Kalita (2022) notes that the datasets used for supervised learning are labeled, meaning each example in the dataset has an associated "outcome" or "summary" value that depends on the details of the example. This is added to the attributes (or features) used to describe the details of the example.

Žižka *et al*. (2019) explain that unsupervised learning does not rely on a "teacher"; learners must learn on their own, and the available training samples do not have their appropriate class labels. As a result, it is not directly possible to reveal what is relevant to each class. Moreover, it is not known which (or how many) classes exist for a specific case. Thus, it is said that the algorithms seek to "naturally" find patterns among the instances based on the available attributes (or variables).

In reinforcement learning, an agent learns to perform a task within an environment. The reinforcement learning agent has a repertoire or set of basic actions it can execute, and at any given time, it is assumed to be "residing" in a set of states. When the agent reaches the final state, the environment, a teacher, or the agent itself provides a reward. Thus, most actions are not rewarded, but rewards are given infrequently or "rarely" (Kalita, 2022). For this article, the classification task was utilized, as a predefined label ("economic satisfaction") was used as a reference for training the machine learning algorithm.

2.1 Traditional machine learning models

This article applied several ML algorithms for data classification: Decision Trees, Random Forest, XGBoost, Naïve Bayes, Support Vector Machines (SVM), Logistic Regression, and a combination of these methods (also known as “ensembles”) using strategies such as voting, stacking, bagging, and boosting.

A classifier based on Decision Trees is structured as a tree-like algorithm similar to a flowchart, where each internal node (non-leaf node) represents a test on an attribute, each branch represents the outcome of the test, and each leaf node (or terminal node) contains a class label. The highest node in the tree is the root node. The process of learning decision trees is performed using class-labeled training tuples (Han *et al*., 2022).

Žižka *et al*. (2019) state that Random Forest employs simultaneous voting by multiple expert algorithms during training, with the outcome determined by the majority of votes (or by averaging for regression tasks). It randomly selects attributes for each split node in each sub-tree, in addition to randomly selecting subsets of training samples using the bagging technique (Breiman, 2001).

According to Zou *et al*. (2022), XGBoost is based on gradient-boosted decision trees. It begins by creating several weak learners, primarily regression trees, to train these learners. After training, a weighted combination is performed to obtain the final regression model. During construction, new learners are added based on the residual error from the last iteration of the weak learner.

Inspired by Bayes' theorem and the calculation of conditional probabilities, the method estimates the label of a new record based on probability distributions previously calculated using labeled data (Da Silva *et al*., 2023). It receives labeled training data denoted by training and label, and produces a structured output labeled to receive test data (Brunton & Kutz, 2019).

According to Han *et al*. (2022), Support Vector Machines (SVMs) are a method for classifying linear and nonlinear data. A nonlinear mapping is applied to transform the original training data into a higher-dimensional space, where the algorithm seeks the optimal linear separating hyperplane (i.e., a "decision boundary" that separates tuples from one class from another). With an appropriate nonlinear mapping to a sufficiently high-dimensional space, the data from two classes can always be separated by a hyperplane. Thus, the SVM finds this hyperplane using support vectors, which are the "essential" training tuples, and margins (defined by the support vectors).

Yadav *et al*. (2022) explain that Logistic Regression (LR) is used to predict the probability of a target or dependent variable that is dichotomous in nature, meaning there are only two possible classes (either 0 or 1). The method performs mathematical modeling to predict the probability of an event occurring, based on the analysis of the relationship between the available variables (Ariza & Santos, 2023).

2.2 Ensemble methods

Han *et al*. (2022) mention that an ensemble learning model combines a series of base classifiers (learning models) to create a composite and enhanced classification model. This method returns a class prediction based on the votes of the base classifiers. There are different types of ensemble classifiers, including bagging, boosting, voting, and stacking.

For the bagging strategy, the term "bagging" stands for "bootstrap aggregating", where each training set is a sample with replacement, and the aggregated classifier counts the votes and assigns the class with the majority of votes to a new instance (Jafarzadeh *et al*., 2021). This model can also be applied to predict continuous values by calculating the average value of each prediction for a given test tuple.

A boosting classifier is designed to produce a prediction rule by combining flexible classifiers in sequence, generating a more powerful classifier based on the adjusted weights of previous classifiers' performance (Naem *et al*., 2018). The first classifier is trained with the training instances, and those incorrectly classified have a higher probability of being selected for the second classifier, continuing until a stopping criterion is met (Kadkhodaei *et al*., 2020).

Commonly used, the voting model is a process in which multiple learning techniques are applied, or the same technique is used multiple times to create the base classifiers, where each of these bases is trained with distinct data. This process makes classification predictions, where the highest vote or score assigned to a prediction is accepted (Géron, 2019; Tauil *et al*., 2024).

The ensemble stacking learning method consists of two phases: base classifier and meta-classifier (Nipa *et al*., 2024). At the base classifier level, the training set is used to train models and make predictions. In contrast, in the meta-classifier, the metadata is used for training, while the output of the base classifier is mapped to the actual classification label (Jiang *et al*., 2019).

2.3 Explainable model

Dandolo *et al*. (2023) cite that ML models have limitations in explaining their internal functioning, often referred to as "black box" models. Consequently, there is a lack of understanding regarding which information the algorithm utilized to comprehend the relationship between input and output variables. To overcome these limitations, the field of Explainable Artificial Intelligence (XAI) has emerged as a type of AI that allows ML models to provide explanations focused on "why" the system reached a particular decision, exploring its logical paradigms (Vishwarupe *et al*., 2022). In this context, SHapley Additive exPlanations (SHAP) can reveal relevant information about the relative influence of input variables on the analyzed classes (Zheng *et al*., 2023).

This model generates SHAP values that indicate the contribution of each attribute in a specific sample, and the predictive model returns a projected output for each separate sample (Amin *et al*., 2023). It leverages game theory to explore the reasons behind the formation of the machine learning model in a particular way, thereby providing a better understanding of the model (Lan *et al*., 2024).

3. MATERIALS AND METHODS

The data used in this study were obtained from surveys conducted in 2020 and 2023 by the Latinobarómetro Corporation, with both datasets undergoing preprocessing and splitting into training and testing sets. Variables were empirically selected based on their relevance to the problem, ensuring their mutual presence in each dataset. Consequently, 19 attributes were selected from each year's dataset, with minimal differences between the years, such as accentuation of specific names and classifications, which were subsequently unified during preprocessing. The most significant challenge identified was that one variable related to the respondent's country of origin lacked information for one of the 18 countries present in the 2020 data. Therefore, it was necessary to exclude this country to maintain consistency in the results of future analyses, thus aligning the variables passed to the algorithms.

The selected output variable was the respondent's assessment of their "general satisfaction with the economy" in their country, with all six different response options present. To transform the problem into a binary classification, the classes were grouped as 0 or 1 according to their correspondence, with 0 representing the "dissatisfied" group and 1 representing the "satisfied" group. The transformation of the classes is presented in Table 1.

Table 1
Treatment of responses for the class variable

Classes from the original output	New output
Very satisfied	Satisfied (1)
Somewhat satisfied	
Somewhat dissatisfied	Insatisfied (0)
Very dissatisfied	
Don’t know	No cases (excluded)
No response	

To handle missing data, no imputation methods were used, therefore, instances with incomplete data were removed from both datasets. This strategy was chosen to maintain greater reliability in the model training phase, while still retaining a substantial amount of information even after excluding the missing data. Additionally, attributes related to gender, country, race, and religion were binarized, and the data were subsequently standardized and normalized. The data processing and cleaning resulted in a final dataset consisting of 27,600 instances and 57 columns, with 14,032 instances for training and 13,568 for testing.

Subsequently, during the algorithm application, the preprocessed 2020 dataset was first used as the training data, while the 2023 instances were used to test the models' effectiveness. This approach allowed for the assessment of compatibility between the datasets from different years, ensuring that evaluating both years with the algorithm would not affect the results, as the questions asked of the respondents remained the same over a short period. After this step, the class balancing for the training set reached 11,760 "dissatisfied" (83.81%) and 2,272 "satisfied" (16.19%).

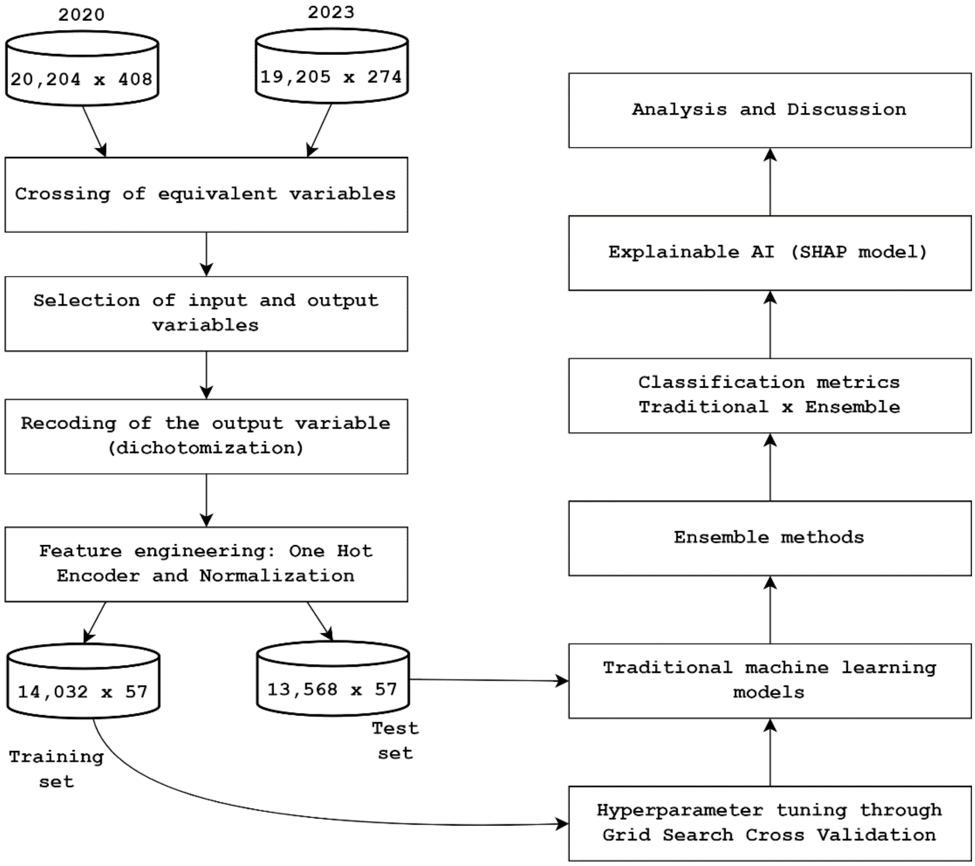
For the classification task, the methods Random Forest (RF), Logistic Regression (REG), Bernoulli Naïve Bayes (BNB), Support Vector Machine (SVM), XGBoost (XGB), and Neural Networks (Neural) were used. To enhance these classifiers, the Grid Search method was employed for hyperparameter tuning within a 5-fold cross-validation and using accuracy as the reference metric. Table 2 summarizes the best hyperparameters after tuning.

Table 2
The best parameters for each model after the Grid Search

Algorithm	**Chosen hyperparameters**
RF	(n_estimators=100, max_depth=20, min_samples_split=10, min_samples_leaf=1, max_features='sqrt')
REG	(penalty='l2', C=0.001, solver='liblinear')
BNB	(alpha=1.0, binarize=1.0, fit_prior=True)
SVM	(C=1, kernel='rbf', gamma='scale', max_iter=1000, probability=True)
XGB	(colsample_bytree=1.0, gamma=0.2, learning_rate=0.1, max_depth=3, n_estimators=100)
Neural	(hidden_layer_sizes=(128,), activations='sigmoid', optimizer='rmsprop')

After training each traditional ML algorithm, ensemble methods were adopted among the classifiers using the dataset that demonstrated the best accuracy. This efficiency was evaluated based on the metrics of accuracy, precision, recall, and f1-score. Figure 1 presents the overall flowchart of the entire data processing and model evaluation, implemented in Python programming language and its libraries.

Figure 1
Research workflow



Finally, after evaluating the performance of the classifiers and ensemble methods with the aid of graphs, the SHAP library was applied to the algorithms that demonstrated the highest quality. For this step, the Kernel Explainer function from the SHAP library, with the algorithm being trained on the 2020 dataset and tested on the 2023 dataset, following the same method as the classifiers.

In the end, this technique provides a means to understand the decision-making method of the classifier, elucidating the key factors of the highest-performing black-box model. This enabled the testing and formulation of hypotheses regarding the categories that most significantly influence public satisfaction or dissatisfaction with the functioning of a country’s economy, cross-referencing these results with articles found in the literature.

4 RESULTS AND DISCUSSION

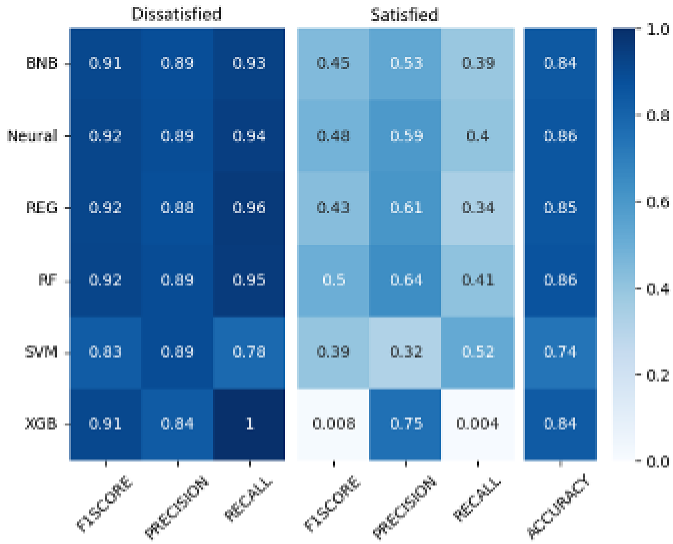
The results obtained after applying the described methods were divided based on the different types of criteria analyzed: (1) classifier algorithm and (2) ensemble methods. Additionally, preference was given to presenting only the results after applying the Grid Search method with all models already tuned to the hyperparameters that maximized accuracy. Visualizations were provided for better interpretations.

Initially, using the prepared dataset as described in the previous section, all resulting instances were employed in the classifier algorithms. To effectively demonstrate the results achieved, all algorithms were tested with the same input data, with the accuracy output revealing the best results, which are analyzed in this section.

4.1 Performance of the classifiers

Firstly, Figure 2 illustrates the performance of the classifiers without any ensemble method applied, with the algorithm following the parameters and inputs previously described. Thus, the heatmap information reflects the metrics used to measure the accuracy of each algorithm, with darker colors representing better results, or closer to the value of 1, and lighter colors indicating poorer classifiers, with metrics closer to a null value.

Figure 2
Classification results for individual algorithms



There was a general difficulty among the algorithms in evaluating instances representing people “satisfied” with the economy. All metrics for the "satisfied" block showed lower values compared to the other block. This result can be explained by the imbalance in the training dataset, where instances of dissatisfied people were more prevalent than their counterparts. The Random Forest classifier had the best result in general, with a high f1-score for the “dissatisfied” class, and 0.5 f1-score for “satisfied”.

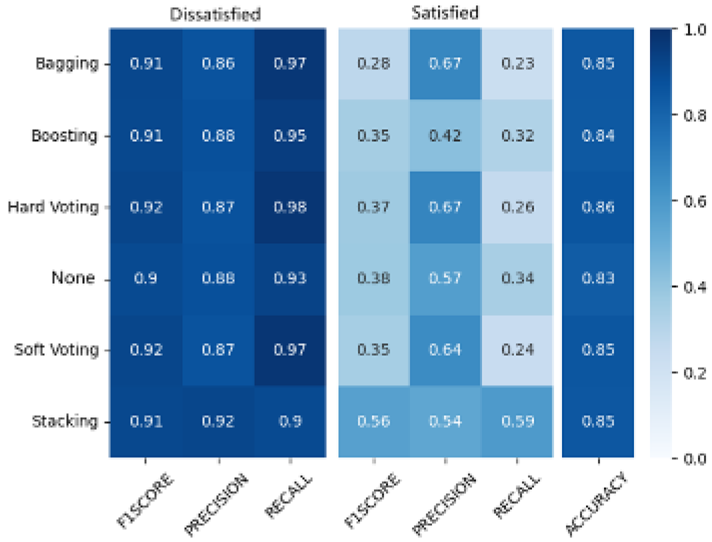
Despite this, it is notable that among the instances, Logistic Regression presented the best evaluation metrics for satisfied people, being considered, for this research, the best classifier among the others. This result stems from its better balance between instances, represented by its high precision rate for the satisfied class while maintaining a higher f1-score and recall compared to the other classifiers.

On the other hand, the XGBoost method obtained the worst results, clearly affected by the imbalance in the input instances, resulting in nearly null f1-score and recall figures for the minority class.

4.2 Performance ensemble methods

From all the classifiers previously used, while maintaining the same data input, the methods bagging, boosting, stacking, and voting (both in their hard and soft forms for this last) were applied in search of an improvement in overall accuracy, especially in the minority target variable. This type of model has considerable potential for forming an algorithm with greater effectiveness in achieving better results (Ogutu *et al.*, 2022; Sagi & Rokach, 2018). Therefore, the average performance of the models is represented in Figure 3.

Figure 3
Results for the ensemble algorithms



Evaluating the average results of the ensemble algorithms in Figure 3, it is observed that the models continued to perform better for the more favored class. For the "Satisfied" class, the method that best managed to balance the predictions was Stacking, with an f1-score of 56%. Regarding accuracies, although the stacking method did not have the highest value for the "dissatisfied" class, it demonstrated the best balance and is therefore classified by this research as the method with the best overall results. It is important to note that the results for "None" are related to the simple average of the individual classifiers as shown in Figure 2.

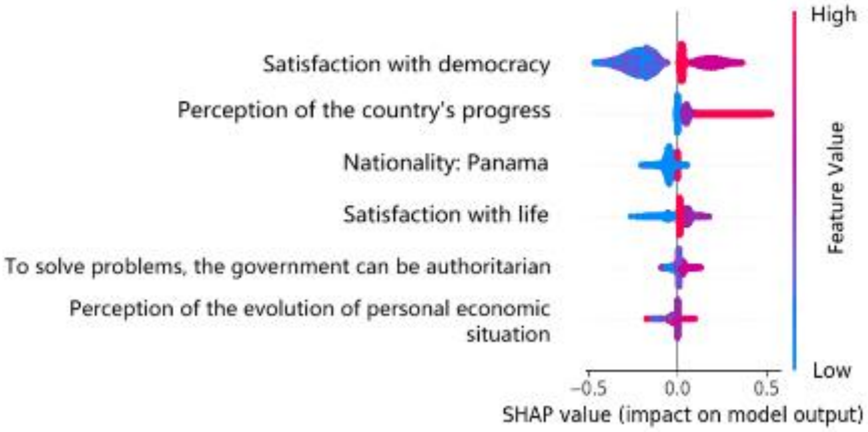
4.3 Feature analysis through explainable model

Even after identifying a classifier algorithm with the best performance in the study, the decision-making process indicating which factors most influenced the determination of whether an individual was satisfied or dissatisfied with the economic situation remained unclear. To clarify the prediction method used by the algorithm, the SHAP method was applied, as it is specifically designed to better visualize the decision-making process of black-box models like some of those used in this research. Given that, the stacking ensemble achieved the best results, so SHAP was applied solely to this method to investigate the decision factors.

The application of SHAP, following the parameters and procedures outlined in the previous section, resulted in the graph presented in Figure 4. This figure displays the names of the features that most significantly influenced the stacking method's decision-making process, along with the values classified as influential for determining satisfaction or dissatisfaction. Since not all instances exerted a strong influence on the decision-making process of this model, many of them resulted in SHAP graphs without a defined SHAP value trend according to the variable's value. These were therefore excluded from Figure 4, which includes only the six most important variables.

Figure 4

SHAP results obtained from the staking ensemble



First, is noteworthy that much of the evaluation regarding economic performance is related to the perception of other relevant aspects of human life. According to the SHAP analysis, "satisfaction with democracy" is the most significant factor, with individuals more content with their country's democracy tending to also assess the economy more positively. This finding aligns with the historical context of the continent, where maintaining democracy in a nation is highly correlated with its economic situation (Bozzetto & Amador, 2022).

Following democracy, the assessment of the country's progress stands out, where individuals who perceive their country as more prosperous tend to evaluate the economy more favorably. Generally, the evaluation of this progress is multifactorial, with studies confirming that indicators such as innovation (Zhylinska *et al.*, 2019), research and development (Khan, 2015), and education (Bah, 2023) are directly related to economic growth, which is closely linked to a country's progress.

The third most important characteristic according to the SHAP analysis is a specific result tied to just one nation. The graph in Figure 4 shows that Panamanian respondents are more likely to positively evaluate their country's economy, indicating that Panamanian citizens were more satisfied with their economy than those in other Latin American countries. Although there are no direct studies on this relationship, the Inter-American Development Bank (IDB, 2008) has reported on the satisfaction ranking of Panamanian citizens regarding life, which is also an important variable for predicting economic satisfaction.

The fourth and sixth factors, respectively, are satisfaction with life and perception of personal financial improvement, highlighting a close relationship between personal life quality factors and national economic performance. A related study by Cahill *et al.* (2015) found that a worsening economic situation, leading to increased perceived risk of unemployment, causes greater job dissatisfaction, impacting personal financial progression. Satisfaction with life in other regions of the world is positively related to several factors that influence the economy, such as income and wealth (D'Ambrosio *et al.*, 2009), economic freedom (Graafland & Compen, 2012), and Gross Domestic Product (VeČerník & Mysíková, 2015).

The fifth characteristic identified by SHAP as important for the stacking model is the acceptance of authoritarian initiatives by the state if they solve societal problems. Individuals more favorable to this type of governance demonstrated a greater likelihood of being economically satisfied. Although this finding contrasts with the most significant factor (democracy), it was not considered as confident as the other alternatives. The graph shows a large overlap of positive and negative values for both classes of the interest variable. Nevertheless, it suggests that the perception that an authoritarian regime can solve societal problems remains strong in some Latin

American countries due to the historical instability of democratic regimes in the region.

As seen from their absence in both models, other characteristics of the respondents, such as gender, age, religion, and race, did not significantly influence economic satisfaction according to the SHAP investigation. These absences may indicate that, despite the clear cultural differences among individuals from various Latin American countries, economic satisfaction is defined by universal factors that transcend community barriers. This conclusion is relevant as it suggests that populations understand progress similarly, and sovereign states should follow a similar path to better serve their citizens with a more advanced economy.

Finally, it was found that the stacking ensemble model achieved the best results among the models analyzed, especially when compared to the prediction made by individual classifiers. This conclusion was based not only on the overall accuracy, which was lower than the other models but rather on the better balance between the classes of the variable of interest, thereby providing results more aligned with reality in assessing individuals satisfied and dissatisfied with their country's economic situation. This factor is crucial because, with only high accuracy, a model might favor instances of economic dissatisfaction simply because they are the majority of recorded cases on the continent.

Thus, the stacking model effectively combined classifiers with low accuracy for the minority variable, creating a new algorithm that better identified the actual satisfaction of individuals. On the other hand, the other ensemble models did not achieve satisfactory results, possibly due to data-related issues and the difficulty of predicting economic satisfaction when the proportion of outcomes for the variable of interest indicates a significant rarity of such individuals in the surveyed population.

5 CONCLUSION

Based on the results achieved by the interpretation model, this study identified several key points relevant to the study of public perception of the economy. The strong association found between the evaluation of democracy and economic satisfaction is well-supported by existing literature, confirming the study’s success in reinforcing these findings, particularly in Latin America.

Furthermore, a significant distinction was observed between traditional classifiers and ensemble methods. Among the individual classifiers, Random Forest showed superior performance, providing high accuracy for the “satisfied” class with fewer samples, along with better balance among precision, and recall metrics. This result is particularly relevant for future research involving classifiers for similar variables.

Regarding ensemble methods, the stacking model achieved the best results, surpassing the Random Forest classifier in general. The observed outcome might be specific to the dataset and variables analyzed, or due to characteristics of the chosen classifiers. Further studies on the efficiency of these algorithms are recommended to clarify these aspects. It is important to include data from past years, seeking to have a higher training set, especially for the “satisfied” class.

In conclusion, this work met its objectives by comparing classifiers and addressing issues related to public satisfaction with the economy in Latin America. Its contributions are valuable for academics, professionals, policymakers, and others interested in public economic perception studies, providing a substantial resource for the field of computational intelligence.

6 ACKNOWLEDGMENTS

The authors wish to thank the Universidade Tecnológica Federal do Paraná for the Scientific Initiation scholarship awarded to the second author of this work, through Edital 02/2023 PROPPG – PIBIC, Cycle 2023-2024.

REFERENCES

Abdelkader, H. E., Gad, A. G., Abohany, A. A., & Sorour, S. E. (2022). **An Efficient Data Mining Technique for Assessing Satisfaction** Level With Online Learning for Higher Education Students During the COVID-19. IEEE Access, 10, 6286–6303. https://doi.org/10.1109/ACCESS.2022.3143035

Amin, M. N., Ahmad, W., Khan, K., Nazar, S., Arab, A. M. A., & Deifalla, A. F. (2023). **Evaluating the relevance of eggshell and glass powder for cement-based materials using machine learning and SHapley Additive exPlanations (SHAP) analysis**. Case Studies in Construction Materials, 19, e02278. https://doi.org/10.1016/j.cscm.2023.e02278

Ariza, V. M. P., & Santos, B. S. dos. (2023). **Classificação da percepção de servidores públicos federais em relação a atos de corrupção utilizando algoritmos de aprendizado de máquina**. Brazilian Journal of Production Engineering, 9(4), 166–178. https://doi.org/10.47456/bjpe.v9i4.42073

Bah, I. A. (2023). **The relationship between education and economic growth**: A cross-country analysis. Research, Society and Development, 12(5), e19312540522. https://doi.org/10.33448/rsd-v12i5.40522

Bortolotti, S. L. V., Moreira Junior, F. de J., Bornia, A. C., Sousa Júnior, A. F. de, & Andrade, D. F. de. (2012). **Consumer satisfaction and item response theory**: creating a measurement scale. Gestão & Produção, 19(2), 287–302. https://doi.org/10.1590/S0104-530X2012000200005

Bozzetto, M., & Amador, R. (2022). **Experiências Institucionais e Confiança na Democracia**: Correlações entre as Instabilidades na América Latina. Congreso Latinoamericano y Caribeño de Ciencias Sociales. “Democracia, Justicia e Igualdad,” 282–306.

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

Brunton, S. L., & Kutz, J. N. (2019). **Data-Driven Science and Engineering**. Cambridge University Press. https://doi.org/10.1017/9781108380690

Cahill, K. E., McNamara, T. K., Pitt-Catsouphes, M., & Valcour, M. (2015). **Linking shifts in the national economy with changes in job satisfaction, employee engagement and work–life balance**. Journal of Behavioral and Experimental Economics, 56, 40–54. https://doi.org/10.1016/j.socec.2015.03.002

Chamorro-Atalaya, O., Arce-Santillan, D., Arévalo-Tuesta, J. A., Rodas Camacho, L., Dávila-Laguna, R. F., Alejos-Ipanaque, R., & Moreno-Chinchay, L. R. (2022). **Supervised learning using support vector machine applied to sentiment analysis of teacher performance satisfaction**. Indonesian Journal of Electrical Engineering and Computer Science, 28(1), 516. https://doi.org/10.11591/ijeecs.v28.i1.pp516-524

Da Silva, A. O., Raminelli, D. G. de T. L., Dos Santos, B. S., & Lima, R. H. P. (2023). **Classificação das percepções de stakeholders sobre o futuro do Brasil utilizando aprendizado de máquina**. AtoZ: Novas Práticas Em Informação e Conhecimento, 12, 1. https://doi.org/10.5380/atoz.v12i0.84075

D’Ambrosio, C., Frick, J. R., & Jäntti, M. (2009). **Satisfaction with Life and Economic Well-Being**: Evidence from Germany. Schmollers Jahrbuch, 129(2), 283–295. https://doi.org/10.3790/schm.129.2.283

Dandolo, D., Masiero, C., Carletti, M., Dalle Pezze, D., & Susto, G. A.

(2023). **AcME—Accelerated model-agnostic explanations:** Fast whitening of the machine-learning black box. Expert Systems with Applications, 214, 119115. https://doi.org/10.1016/j.eswa.2022.119115

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

Graafland, J., & Compen, B. (2012). **Economic Freedom and Life Satisfaction:** A Cross Country Analysis. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2057751

Han, J., Pei, J., & Tong, H. (2022). **Data Mining:** Concepts and Techniques(4th ed., Vol. 1). Morgan Kaufmann.

Inter-American Development Bank. (2008). **Idb | faster economic growth hurts life satisfaction in latin america and the caribbean.** https://www.iadb.org/en/news/faster-economic-growth-hurts lifesatisfactionlatin-america-and-caribbean

Jafarzadeh, H., Mahdianpari, M., Gill, E., Mohammadimanesh, F., & Homayouni, S. (2021). **Bagging and Boosting Ensemble Classifiers for Classification of Multispectral, Hyperspectral** and PolSAR Data: Comparative Evaluation. Remote Sensing, 13(21), 4405. https://doi.org/10.3390/rs13214405

Jiang, W., Chen, Z., Xiang, Y., Shao, D., Ma, L., & Zhang, J. (2019). SSEM: **A Novel Self-Adaptive Stacking Ensemble Model for Classification.** IEEE Access, 7, 120337–120349. https://doi.org/10.1109/ACCESS.2019.2933262

Joolfoo, K. M. B. A., Jugurnauth, R. A., & Joolfoo, M. B. A. (2022). **Application of Machine Learning in Predicting Customer Satisfaction of Telecom Service Providers.** 2022 4th International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM), 1–10. https://doi.org/10.1109/ELECOM54934.2022.9965212

Kadkhodaei, H. R., Moghadam, A. M. E., & Dehghan, M. (2020). **HBoost:** heterogeneous ensemble classifier based on the Boosting method and entropy measurement. Expert Systems with Applications, 157, 113482. https://doi.org/10.1016/j.eswa.2020.113482

Kalita, J. (2022). **Machine Learning.** Chapman and Hall/CRC. https://doi.org/10.1201/9781003002611

Khan, J. (2015). **The Role of Research and Development in Economic Growth.** MPRA.

Kowalski, R. (2017). **Patients' written reviews as a resource for public healthcare management in England.** Procedia Computer Science, 113, 545–550. https://doi.org/https://doi.org/10.1016/j.procs.2017.08.275

Lan, H., Wang, S., & Zhang, W. (2024). **Predicting types of human-related maritime accidents with explanations using selective ensemble learning and SHAP method.** Heliyon, 10(9), e30046. https://doi.org/10.1016/j.heliyon.2024.e30046

Langan, A. M., & Harris, W. E. (2024). **Metrics of student dissatisfaction and disagreement:** longitudinal explorations of a national survey instrument. Higher Education, 87(2), 249–269. https://doi.org/10.1007/s10734-023-01004-0

Li, Q., Jing, R., & Xihua Zhu. (2024). **Determinants of travel satisfaction**

for commercial airlines: A data mining approach. Engineering Applications of Artificial Intelligence, 133, 108597. https://doi.org/10.1016/j.engappai.2024.108597

Liang, W., & Jia, C. (2023). **Application of improved neighbor propagation algorithm in international communication and cooperation to promote internationalization of higher education.** Computer Applications in Engineering Education, 31(3), 696–709. https://doi.org/10.1002/cae.22578

Martínez-Navalón, J.-G., Gelashvili, V., & Gómez-Ortega, A. (2021). **Evaluation of User Satisfaction and Trust of Review Platforms:** Analysis of the Impact of Privacy and E-WOM in the Case of TripAdvisor. Frontiers in Psychology, 12. https://doi.org/10.3389/fpsyg.2021.750527

Naem, A. A., Ghali, N. I., & Saleh, A. A. (2018). **Antlion optimization and boosting classifier for spam email detection.** Future Computing and Informatics Journal, 3(2), 436–442. https://doi.org/10.1016/j.fcij.2018.11.006

Nipa, N., Riyad, M. H., Satu, S., Walliullah, Howlader, K. C., & Moni, M. A. (2024). **Clinically adaptable machine learning model to identify early appreciable features of diabetes.** Intelligent Medicine, 4(1), 22–32. https://doi.org/10.1016/j.imed.2023.01.003

Noviantoro, T., & Huang, J.-P. (2022). **Investigating airline passenger satisfaction:** Data mining method. Research in Transportation Business & Management, 43, 100726. https://doi.org/10.1016/j.rtbm.2021.100726

Ogutu, R. V. A., Rimiru, R. M., & Otieno, C. (2022). **Target Sentiment Analysis Ensemble for Product Review Classification.** Journal of Information Technology Research, 15(1), 1–13. https://doi.org/10.4018/JITR.299382

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.

Polce, E. M., Kunze, K. N., Fu, M. C., Garrigues, G. E., Forsythe, B., Nicholson, G. P., Cole, B. J., & Verma, N. N. (2021). **Development of supervised machine learning algorithms for prediction of satisfaction at 2 years following total shoulder arthroplasty.** Journal of Shoulder and Elbow Surgery. https://doi.org/10.1016/j.jse.2020.09.007

Rosa, D. M. de S., Dos Santos, B. S., & Lima, R. H. P. (2023). **Predicting satisfaction with democracy in Brazil considering data form an opinion survey.** Revista Gestão Da Produção Operações e Sistemas, 18. https://doi.org/10.15675/gepros.2965

Sabarmathi, G., & Chinnaiyan, R. (2019). **Reliable Machine Learning Approach to Predict Patient Satisfaction for Optimal Decision Making and Quality Health Care.** 2019 International Conference on Communication and Electronics Systems (ICCES), 1489–1493. https://doi.org/10.1109/ICCES45898.2019.9002593

Sagi, O., & Rokach, L. (2018). **Ensemble learning:** A survey. WIREs Data Mining and Knowledge Discovery, 8(4). https://doi.org/10.1002/widm.1249

Tauil, Y. B., Santos, B. S. dos, & Lima, R. H. P. (2024). **Machine learning techniques in classifying satisfaction with the economy of Latin American Citizens.** Observatorio de La Economía Latinoamericana, 22(5), e4912. https://doi.org/10.55905/oelv22n5-199

Van Ryzin, G. G. (2004). **The Measurement of Overall Citizen Satisfaction.**

Public Performance & Management Review, 27(3), 9–28. http://www.jstor.org/stable/3381143

VeČerník, J., & Mysíková, M. (2015). **GDP and life satisfaction in European countries – focus on transition.** Post-Communist Economies, 27(2), 170–187. https://doi.org/10.1080/14631377.2015.1026687

Veenhoven, R. (1996). **Developments in satisfaction-research.** Social Indicators Research, 37(1), 1–46. https://doi.org/10.1007/BF00300268

Vishwarupe, V., Joshi, P. M., Mathias, N., Maheshwari, S., Mhaisalkar, S., & Pawar, V. (2022). **Explainable AI and Interpretable Machine Learning**: A Case Study in Perspective. Procedia Computer Science, 204, 869–876. https://doi.org/10.1016/j.procs.2022.08.105

Yadav, V., Dubey, A. K., Singh, H. P., Dubey, G., & Suryani, E. (2022). **Process Mining Techniques for Pattern Recognition.** CRC Press. https://doi.org/10.1201/9781003169550

Zaghloul, M., Barakat, S., & Rezk, A. (2024). **Predicting E-commerce customer satisfaction**: Traditional machine learning vs. deep learning approaches. Journal of Retailing and Consumer Services, 79, 103865. https://doi.org/10.1016/j.jretconser.2024.103865

Zhang, S., Chen, J. Y., Pang, H. N., Lo, N. N., Yeo, S. J., & Liow, M. H. L. (2021). **Development and internal validation of machine learning algorithms to predict patient satisfaction after total hip arthroplasty.** Arthroplasty, 3(1), 33. https://doi.org/10.1186/s42836-021-00087-3

Zheng, X., Xie, Y., Yang, X., Amin, M. N., Nazar, S., Khan, S. A., Althoey, F., & Deifalla, A. F. (2023). **A data-driven approach to predict the compressive strength of alkali-activated materials and correlation of influencing parameters using SHapley Additive exPlanations (SHAP) analysis.** Journal of Materials Research and Technology, 25, 4074–4093. https://doi.org/10.1016/j.jmrt.2023.06.207

Zhylinska, O., Chernyak, O., & Bazhenova, O. (2019). **The role of innovations in driving economic growth**: case of advanced economies. Globalization and Business, 4(7), 11–15. https://doi.org/10.35945/gb.2019.07.001

Žižka, J., Dařena, F., & Svoboda, A. (2019). **Text mining with machine learning**: Principles and Techniques (1st ed.). CRCPress.

Zou, Z., Wang, L., Chen, J., Long, T., Wu, Q., & Zhou, M. (2022). **Research on peanut variety classification based on hyperspectral image.** Food Science and Technology, 42. https://doi.org/10.1590/fst.18522