Anales del Instituto de Actuarios Españoles

4ª Época, Número 30, Año 2024/1-20
DL M-3160-1961 – ISSN 0534-3232 – eISSN 2531-2308

UNDERSTANDING CONFIDENCE IN BANKS: THE ROLE OF PERSONAL CHARACTERISTICS AND ARTIFICIAL INTELLIGENCE

ENTENDIENDO LA CONFIANZA EN LOS BANCOS: EL PAPEL DE LAS CARACTERÍSTICAS PERSONALES Y LA INTELIGENCIA ARTIFICIAL

Raúl Gómez-Martínez
Universidad Rey Juan Carlos. Móstoles, Spain.
ORCID: https://orcid.org/0000-0003-3575-7970
raul.gomez.martinez@urjc.es

Benito Pérez-González
Universidad Internacional de La Rioja (UNIR). Logroño, Spain.
ORCID: https://orcid.org/0000-0002-6191-2885
benito-perez@unir.net

María Luisa Medrano-García
Universidad Rey Juan Carlos. Móstoles, Spain.
ORCID: https://orcid.org/0000-0003-1844-1034
marialuisa.medrano@urjc.es

Jose Torres-Pruñonosa
Universidad Internacional de La Rioja (UNIR). Logroño, Spain.
ORCID: https://orcid.org/0000-0003-2276-1655
jose.torresprunonosa@unir.net
(Autor para correspondencia/Corresponding author)

Fecha de recepción/Date of reception: 2 March 2024
Fecha de aceptación/Date of acceptance: 26 April 2024
Fecha de publicación/Date of publicartion: 31 July 2024

ABSTRACT

Confidence in banks and financial institutions is a cornerstone of financial stability and economic prosperity. This study investigates the relationship between personal characteristics and confidence in banks, recognizing the pivotal role of trust in shaping individuals' perceptions of financial institutions. Through a mixed-methods approach combining survey techniques and artificial intelligence modelling, we analyse data collected from a representative sample of the university community. Our findings highlight the significant influence of demographic factors such as age, gender and education level on confidence in banks. Moreover, we validate our hypothesis using metrics such as ROC Area and PRC Area, indicating the predictive power of personal characteristics in determining confidence in banks. The sensitivity analysis further elucidates the relative importance of different predictors in shaping confidence levels. The implications of our research extend to policymakers, financial institutions and researchers, offering insights for tailored interventions, customer engagement strategies, and future investigations. By deepening our understanding of the drivers of confidence in banks, this study contributes to the enhancement of financial stability and consumer trust in the banking

DOI 10.26360/2024_03

sector.

Keywords: financial institutions, artificial intelligence, survey analysis, banking

RESUMEN

La confianza en los bancos y las instituciones financieras es una piedra angular de la estabilidad financiera y la prosperidad económica. Este estudio investiga la relación entre las características personales y la confianza en los bancos, reconociendo el papel fundamental de la confianza en la formación de las percepciones de los individuos sobre las instituciones financieras. Mediante un enfoque de métodos mixtos que combina técnicas de encuesta y modelización de inteligencia artificial, analizamos datos recogidos de una muestra representativa de la comunidad universitaria. Nuestros resultados ponen de relieve la influencia significativa de factores demográficos como la edad, el sexo y el nivel educativo en la confianza en los bancos. Además, validamos nuestra hipótesis utilizando métricas como el área ROC y el área PRC, que indican el poder predictivo de las características personales a la hora de determinar la confianza en los bancos. El análisis de sensibilidad aclara aún más la importancia relativa de los distintos predictores en la configuración de los niveles de confianza. Las implicaciones de nuestra investigación se extienden a los responsables políticos, las instituciones financieras y los investigadores, ofreciendo ideas para intervenciones a medida, estrategias de captación de clientes y futuras investigaciones. Al profundizar en el conocimiento de los factores que impulsan la confianza en los bancos, este estudio contribuye a mejorar la estabilidad financiera y la confianza de los consumidores en el sector bancario.

Palabras clave: instituciones financieras, inteligencia artificial, análisis de inspección, banco

1. INTRODUCTION

Confidence in financial institutions, particularly banks, plays a pivotal role in the stability and prosperity of economies (Hammond & Opoku, 2023). Understanding the factors that influence citizens' confidence in banks is crucial for policymakers, financial institutions and researchers alike. In this study, we employ a mixed methodology approach integrating survey techniques and artificial intelligence modelling to investigate the relationship between personal characteristics and confidence in banks (Henrique et al., 2019).

The study begins with the administration of a structured survey designed to gather information on participants' personal characteristics and their confidence in banks. Emphasizing representativeness and diversity, the survey is distributed among the university community via social networks, ensuring confidentiality, privacy, and informed consent.

Subsequently, the collected data undergo descriptive and inferential statistical analysis, exploring associations between variables of interest and identifying possible subgroups or patterns within the sample. To deepen our understanding, we use a mixed methodology, incorporating supervised machine learning techniques, particularly Random Forest (Goulet Coulombe, P., 2024; Huang et al., 2023; Yego et al., 2023; Andrade & Valencia, 2021), to develop and validate an AI model predicting confidence in banks.

The validation of our hypothesis hinges on the performance metrics of the AI model, specifically the ROC Area (Rice & Harris, 2005; Ekström et al., 2023, Bouallègue, & Richardson, 2022) and PRC Area (Lei et al., 2022; El Fouki et al., 2019). ROC (Receiver Operating Characteristic) area measures the trade-off between true positive rate and false positive rate, whereas PRC (Precision-Recall Curve) area measures precision-recall trade-off. Theses metrics measure the model's ability to distinguish between positive and negative classes and evaluate the precision-recall trade-off, respectively. Through a sensitivity analysis, we assess the impact of including or removing variables on the model's predictive power.

Our findings underscore the significant influence of personal characteristics such as age, gender, education level and perception of current issues on confidence in banks. While certain predictors may wield stronger influence, a nuanced understanding requires taking into account various demographic and attitudinal factors.

Implications of our research extend to policymakers, financial institutions and researchers, informing tailored interventions, customer engagement strategies and future avenues of inquiry. By delving into longitudinal studies, cross-cultural comparisons, qualitative research and the impact of interventions, future research can deepen our understanding of confidence in banks and contribute to the development of more effective policies and strategies for fostering financial stability and consumer trust.

The structure of the paper is as follows. Both the objectives and the literature review will be shown in section 2. Section 3 will explain the methodology used. The results will be written in section 4, whereas they will be discussed in section 5. Finally, conclusions, limitations and futures lines of research will be explained in section 6.

2. THEORETICAL FRAMEWORK

The aim of this article is to explore the connection between human traits and bank confidence, acknowledging the critical role that trust plays in influencing people's opinions about financial institutions.

Confidence in banks and financial firms is crucial for several reasons. First of all, because it has to do with the stability of the financial system (Xu, 2020). Confidence ensures the stability of the financial system. When individuals and businesses trust banks and financial firms to safeguard their money and investments, they are more likely to participate in financial activities such as saving, investing and borrowing. This confidence helps to maintain liquidity and smooth functioning of financial markets. Second, economic growth is involved. Confidence in banks and financial institutions fosters economic growth. When people have trust in the banking system, they are more willing to invest in businesses, buy homes and make other long-term financial commitments (Nasreen et al., 2020). This investment stimulates economic activity, creates jobs and drives overall economic prosperity (Mikhaylov et al, 2023). The third reason is risk management. Confidence encourages effective risk management. Banks and financial firms play a critical role in managing and distributing risk in the economy (Mangala & Sony, 2020). When stakeholders have faith in these institutions' ability to assess and manage risks, they are more likely to engage in financial transactions, which contributes to a healthier allocation of capital and resources. Fourth, consumer protection has to be taken into consideration. Confidence promotes consumer protection. Strong confidence in banks and financial firms often correlates with robust regulatory frameworks and risk management practices. These institutions are incentivized to maintain high standards of transparency, accountability and customer service to preserve public trust (Van der Cruijsen et al., 2023). Fifth, confidence in banks and financial firms is essential for global financial stability (Kusi et al, 2023). In an interconnected global economy, the trustworthiness of financial institutions transcends national borders. Confidence in major financial centres and institutions can have ripple effects across the entire global financial system. Finally, confidence in banks and financial firms contributes to social stability (De Leon, 2020). Financial instability or the collapse of major financial institutions can have severe social consequences, including unemployment, poverty and social unrest. Maintaining confidence helps mitigate these risks and fosters social cohesion (Karkošková, 2023). All in all, confidence in banks and financial firms is vital for the 1) stability of the financial system, 2) economic growth, 3) effective risk management, 4) consumer protection, 5) global financial stability and 6) social stability. Building and maintaining this confidence require strong regulatory oversight, prudent risk management practices, transparency, and accountability from financial institutions.

Measuring confidence in banks involves assessing the level of trust and belief that various stakeholders have in the financial institution. Several methods and indicators can be used to gauge confidence in banks. First, conducting surveys among bank customers to gather feedback on their satisfaction, trust, and overall confidence in the bank (Chauhan et al., 2023). Questions can cover aspects such as customer service, reliability, and perceived financial stability. Second, market perception is an item to be taken into account. Monitoring the perception of the bank in the financial markets, including the stock market. Stock prices and credit ratings can provide insights into how investors and creditors perceive the bank's financial health and management (Josephson & Shapiro, 2020). Third, regulatory assessments are involved. Regulatory bodies often assess the financial health and compliance of banks. Regular examinations and stress tests conducted by regulatory authorities can offer an objective measure of a bank's stability and adherence to industry standards (Vanclay, 2020). Fourth, deposit and withdrawal trends are necessary to be taken into account. Analysing patterns in deposit and withdrawal activities can indicate the level of confidence among depositors. Sudden, large withdrawals may suggest a lack of confidence, while steady or increasing deposits may signify trust (Andros et al., 2020). Fifth, publicly Available Financial Reports Examining the bank's financial reports, including balance sheets, income statements and other financial disclosures, can provide insights into its financial stability. Transparency and clear communication in financial reporting can enhance confidence (Isidro et al., 2020). Sixth, monitoring Credit Default Swap (CDS) Spreads associated with a bank can indicate market participants' perception of the institution's creditworthiness. Wider spreads may suggest increased perceived risk (Andres et al., 2021). Seventh, analysing media coverage and social media sentiment can provide a qualitative measure of public opinion. Positive or negative sentiments expressed by the public and media can influence confidence levels (Kuchciak & Wiktorowicz, 2021). Eight, the level of government support or intervention during financial crises can impact confidence (Laeven & Valencia; Lee et al., 2020). Government actions to stabilize or rescue a bank may influence perceptions of its stability. Ninth, peer Comparisons needs to be taken into account. Comparing a bank's performance and reputation with its peers in the industry can offer insights into its relative standing. If a bank consistently outperforms or underperforms its peers, it may affect confidence (Laeven & Valencia, 2020). Last but not least, employee satisfaction and turnover rates are good indicators to be used. The morale and turnover rates of bank employees can be indicative of the internal health of the institution. High employee satisfaction may contribute to a positive external perception (Bhardwaj et al, 2021).

The combination of multiple indicators and methods provides a more comprehensive understanding of confidence in banks. It is essential to consider both quantitative and qualitative factors to form a well-rounded assessment.

Using artificial intelligence (AI) to analyse survey data and identify the best predictors for banks' confidence can be a powerful approach (Zhang & Lu, 2021). Here is a general outline of how you could proceed:

1. Data Collection: Gather survey data from bank customers, focusing on questions related to their satisfaction, trust and confidence in the bank. Ensure the survey covers various aspects such as customer service, reliability, financial stability and overall experience.

2. Data Preparation: Clean and preprocess the survey data to handle missing values, outliers and inconsistencies. Convert categorical variables into numerical representations if needed and ensure the data is in a format suitable for analysis.

3. Feature Selection: Use AI techniques, such as feature selection algorithms, to identify the most relevant predictors for banks' confidence. These algorithms can automatically select the subset of survey questions/features that have the strongest predictive power (Bigné, 2020).

4. Model Training: Train a machine learning model (Witten & Frank, 2005; Russell, 2018; Marqués, 2020) using the selected features and the survey responses as the target variable (banks' confidence). Choose an appropriate machine learning algorithm for the task, such as regression, classification, or ensemble methods.

5. Model Evaluation: Evaluate the trained model's performance using validation techniques such as cross-validation or holdout validation (Yadav & Shukla, 2016). Measure metrics like accuracy, precision, recall, or F1-score to assess how well the model predicts banks' confidence.

6. Interpretation: Interpret the results to understand which survey questions/features have the most significant impact on banks' confidence. Identify key insights and patterns that emerge from the analysis.

7. Deployment: Deploy the trained model to predict banks' confidence using new survey data. Integrate the model into a workflow where it can provide real-time predictions or insights to stakeholders.

8. Monitoring and Iteration: Continuously monitor the model's performance and update it as needed with new data. Iterate on the model to improve its accuracy and relevance over time.

By leveraging AI techniques to analyse survey data, you can identify the factors that most strongly influence banks' confidence, helping financial institutions better understand their customers' perceptions and prioritize areas for improvement.

Taking this into account, the study hypothesis is:

H_0: *The personal characteristics of individuals are capable of predicting confidence in banks.*

2. METHODOLOGY

In this study, a quantitative methodology was used based on data collection through a survey specifically designed to investigate citizens' perception of banks confidence. The survey was administered to a representative sample of the university community with the aim of gathering information about the respondents' personal characteristics and their confidence in institutions and political parties.

The survey was distributed using social networks, with an emphasis on ensuring the representativeness and diversity of the sample. Measures were included to ensure the confidentiality and privacy of participants, as well as to obtain their informed consent to participate in the study.

The collected data were analysed using descriptive and inferential statistical techniques (Martín & Caballero, 2020). Associations between variables of interest were assessed, and additional analyses were conducted to explore possible subgroups or patterns within the sample.

In this study, a mixed methodology was employed that integrates survey techniques and artificial intelligence (AI) modelling to investigate the relationship between personal characteristics and confidence in institutions and political parties. The research was conducted in two main phases: data collection through a survey and the development and validation of an AI model to predict confidence in these entities.

In the first phase, a structured survey was designed consisting of two main parts: the first part included questions related to participants' personal characteristics such as age, gender, educational level, political affiliation, among others, which would serve as predictor variables in the AI model (Angelov et al.; Zang and Lu, 2021). The second part of the survey consisted of questions aimed at measuring participants' confidence in banks. These target questions were defined as the variables to be predicted in the AI model.

In the second phase, the collected dataset was used to develop and validate an AI model to predict confidence in banks. Supervised machine learning techniques of Random Forest were employed to train the model using predictor variables obtained from the survey (Goulet Coulombe, P., 2024; Huang et al., 2023; Yego et al., 2023). Random Forest is a popular and powerful option in machine learning for various scenarios, and it could be an appropriate choice for your study for several reasons:

- High accuracy: Random Forest tends to provide accurate results across a wide range of datasets, even in datasets with noise or irrelevant features. This makes it suitable when seeking a reliable model to predict confidence in institutions and political parties.

- Robustness to overfitting: Random Forest tends to be less prone to overfitting than other machine learning algorithms, meaning it is able to generalize well to unseen data. This is important to ensure that your model can make accurate predictions in real-world situations.
- Effective handling of categorical and numerical data: Random Forest can effectively handle both categorical and numerical variables in datasets, making it versatile for a variety of problems, including those involving surveys with questions of different types.
- Scalability: Random Forest can scale well to large datasets, meaning it can handle surveys with a large number of observations and features without compromising model performance.
- Feature interpretation: Random Forest provides a measure of feature importance, which can help understand which variables have the greatest impact on confidence in institutions and political parties in your study.

In summary, Random Forest is a solid choice for modelling the relationship between personal characteristics and political confidence in your study due to its accuracy, robustness, ability to handle different types of data, and ease of interpretation. Nevertheless, it is always important to evaluate various algorithms and techniques to find the best solution for your specific problem.

The model was trained using a portion of the data and validated using the remaining portion using 10-fold cross-validation techniques. 10-fold cross-validation is a commonly used technique in machine learning to evaluate the performance of a predictive model. In this technique, the dataset is randomly divided into 10 equal parts, or "folds," of approximately equal size.

The 10-fold cross-validation process is carried out in several stages:

1. Division of the dataset: the dataset is divided into 10 equal parts, called "folds."
2. Iterations of training and evaluation: the model is trained and evaluated 10 times. In each iteration, a different fold is selected as the test set, and the other 9 folds are used as the training set.
3. Performance evaluation: after each iteration, a performance metric (such as accuracy, sensitivity, specificity, etc.) is calculated using the test set. These metrics are averaged at the end of the 10 iterations to obtain an estimate of the model's performance.
4. Obtaining average performance metrics: at the end of the cross-validation process, an estimate of the model's performance is obtained by averaging the metrics calculated in each iteration.

10-fold cross-validation is useful because it provides a more robust evaluation of the model's performance than simply dividing the dataset into a training set and a test set. By performing multiple iterations of training and evaluation on different subsets of data, a more reliable estimate of the model's expected performance in practice is obtained.

This technique is particularly useful when a moderate-sized dataset is available and a more precise evaluation of the model's performance is desired without wasting valuable data on a fixed test set.

For model validation, the accuracy and generalization of the model were assessed using appropriate performance metrics such as accuracy, sensitivity and specificity. Evaluating a machine learning model using multiple metrics such as area under the ROC curve (ROC AUC), and area under the precision-recall curve (PRC AUC) provides a more comprehensive understanding of the model's performance and its capabilities in different aspects. Evaluating a model using multiple metrics provides a more complete and balanced assessment of its performance, helping to better understand its strengths and weaknesses in different aspects of the classification problem.

The metrics used to evaluate the model's accuracy are, on the one hand, the area under the ROC curve (Receiver Operating Characteristic) is a metric used to evaluate the performance of a binary classification model. The ROC curve is a graphical representation that shows the relationship between the true positive rate (TPR) and the false positive rate (FPR) at different classification thresholds. On the other hand, the area under the precision-recall curve (PRC) is a metric used to evaluate the performance of a binary classification model. The PRC curve is a graphical representation that shows the relationship between precision and recall (sensitivity) at different classification thresholds.

The interpretation of these metrics (both ROC and PRC area) may vary depending on the context and the specific problem, but here are some general guidelines:

- Equal to 0: If the metric is zero, it means that the model has no ability to distinguish between positive and negative classes. Essentially, the model is predicting at random.
- Between 0 and 0.5: This range indicates poor model performance. It means that precision is lower than the true positive rate, which is worse than a random approach.
- Between 0.5 and 0.7: In this range, the model has limited predictive ability but may offer some value. However, precision and true positive rate are still considered less than ideal.
- Between 0.7 and 0.9: This range indicates good model performance. The model is able to provide good precision as the true positive rate increases.
- Between 0.9 and 1: An area under the PRC curve above 0.9 is considered excellent. It indicates that the model has high precision even at high true positive rates, which is highly desirable in many applications.

Therefore, we will validate H_0 if the metrics ROC area and PRC area are greater than 0.5, although we will qualify as efficient models those that show metrics above 0.7.

Furthermore, we will conduct a sensitivity analysis by including and removing variables to identify which predictors are most effective in predicting confidence in banks.

As far as the software is concerned, WEKA, or Waikato Environment for Knowledge Analysis (Attwal and Dhiman, 2020), is a popular suite of machine learning software written in Java, developed at the University of Waikato in New Zealand, chosen for this study. It provides a comprehensive set of tools for data preprocessing, classification, regression, clustering, association rules mining and visualization. WEKA is widely used for research, education and application development in the field of machine learning and data mining due to its user-friendly interface, extensive collection of algorithms and open-source nature.

4. DATA

The survey has been distributed through an online form accessible by means of a link. The form consists of the questions included in Annex 1, whereas demographic information is asked, (such as age, gender, material status, education level) along with financial literacy, whether they considered if they are up to date with current affairs, their preference channel to be informed, the first newspaper to be consulted in the event of searching for economic and political information, the football team they support and the last match result, its perception of the weather in the day were the survey was carried out, and how they felt that day, using a 4-point Likert Scale variable. Finally, how must trust is placed in banks was asked also using a 4-point Likert Scale.

The data collection process involved the dissemination of an online survey across various social media platforms and through internal communication channels within the university community. Respondents were encouraged to participate voluntarily, and their responses remained anonymous to ensure confidentiality.

Within the survey, several variables were measured to capture a comprehensive understanding of participants' demographics, attitudes, and behaviours. These variables included age, gender, marital status, level of education completed, and financial literacy level. Additionally, participants were asked about their awareness of current affairs, preferred channels for information consumption, and their inclination towards seeking economic-political information from various sources.

Furthermore, the survey assessed respondents' affiliations with sports teams, the outcomes of their team's latest matches, current weather conditions, and their current mood. Notably, the primary variable of interest, the level of trust in banks, was measured using a Likert scale ranging from 1 to 4, where respondents indicated their degree of confidence in banking institutions.

While the survey methodology facilitated data collection from a diverse range of individuals within the university community, it is essential to acknowledge several limitations. One potential limitation is the inherent bias introduced by the voluntary nature of participation, as individuals who chose to respond may hold different perspectives compared to those who did not participate. Additionally, the distribution method of the survey via social media and university channels may have resulted in a sample that is not entirely representative of the broader population. These limitations should be considered when interpreting the findings and generalizing the results to the larger population.

5. RESULTS

The values obtained from the 60 surveys collected throughout the month of November, 2023 are explained both in this paragraph for qualitative variables and in Table 1 for quantitative ones. As far as age is concerned, 38% were younger than 24 years old, 47% were between 25 and 44 years old, 13% were between 45 and 64 years old and 2% was higher than 65 years 2. When it comes to **Gender**, 53% were men, 45% were women and 2% declared themselves as "Others". In regards to **Marriage**, 42% were single, 33% were living along with their mate, 22% were single and 3% declared to be in other situations. In terms of the **Mood** variables, 37% were Real Madrid's supporters, 18% were supporters of Atlético de Madrid, 12% of Futbol Club Barcelona, 8% of other teams and 25% did not like football; 55% of the sample answered that their favourite team won the last game, 2% that they draw, 4% that they lost and 27% that they did not follow any team; when it comes to today's weather perception, 35% perceived it was sunny, 17% cloudy, 15% rainy and 2% a very much indeed terrible day; finally, concerning the question related to how they felt, the results were a 4-scale Likert and are included in Table 1. When it comes to **Studies** variables, on the one hand, 30% were finished a bachelor; 27% took a Master, 25% finished their High School, 17% took a PhD and 2% finished only primary school; and on the other hand, the results of their perception in regards to financial literally is shown in Table 1. As far as **News** variables are concerned, 95% of the sample though that they were aware of news, whereas only 5% were not; as to the first source of information they would search in regards to economic and political affairs, there is a great variety of response: 23% responded "El País", 18% "El Mundo", 5% "ABC", 5% "El Confidencial", 2% "20 minutos", 2% "La Razón", 2% "La Vanguardia", 2% "La Voz de Galicia", 2% "OK Diario" and 40% other options; finally, when it comes to the channel, 88% preferred the digital channels (Internet), 8% radio or television and 3% written channels (newspapers).

Table 1: Descriptive results of quantitative variables.

	Financial literacy	Today's feeling	Trust on banks
Mean	3,00	3,00	2,35
Variance	0,89	0,89	0,75

Using machine learning techniques (Annex 2), the metrics described above, ROC Area and PRC Area have been obtained (Table 2). The table indicates whether a variable has been considered for model training. A "1" signifies that the variable has been included in the training process, while the absence of a "1" indicates that the variable has not been utilized for training. Additionally, the TP rate (True Positive Rate) is shown, indicating the model's ability to correctly identify positive cases and the FP rate (False Positive Rate), indicating the proportion of negative cases that the model incorrectly classifies as positive.

Table 2: 10 folds cross validation

Attributes	**All**	**Without Gender**	**Without mood**	**Without marriage**	**Without studies**	**Without news**
What is your age?	1	1	1	1	1	1
Gender	1		1	1	1	1
Marriage	1	1	1		1	1
Completed level of education	1	1	1	1		1
Your financial literacy is	1	1	1	1		1
Do you consider yourself up to date with current affairs?	1	1	1	1	1	
To inform yourself, what channel do you prefer?	1	1	1	1	1	
If you had to search for information on an economic-political topic, would you first consult...	1	1	1	1	1	
What team do you support?	1	1		1	1	1
Has your team won its last match?	1	1		1	1	1
What is the weather like today?	1	1		1	1	1
How do you feel today?	1	1		1	1	1
How much trust do you place in banks?	1	1	1	1	1	1
TP Rate	0.48	0.50	0.47	0.50	0.50	0.45
FP Rate	0.30	0.31	0.29	0.31	0.32	0.35
Precision	0.49	0.49	0.46	0.48	0.48	0.46
Recall	0.48	0.50	0.47	0.50	0.50	0.45
F-Measure	0.48	0.48	0.46	0.48	0.48	0.44
MCC	0.20	0.20	0.18	0.21	0.19	0.12
ROC Area	0.67	0.67	0.67	0.66	0.60	0.67
PRC Area	0.55	0.57	0.53	0.54	0.50	0.58

The results obtained indicate the performance of the model under different scenarios where specific predictors are removed. The discussion of the results obtained will be considered in the following paragraphs.

ROC Area (see Figure 2). This metric measures the model's ability to distinguish between positive and negative classes across different thresholds. A higher ROC Area indicates better overall performance. In this case, we see that removing certain predictors does not significantly impact the ROC Area, as the values remain relatively consistent around 0.67. However, removing predictors related to the level of education or perception of current affairs slightly decreases the ROC Area, suggesting these predictors might contribute somewhat to the model's predictive ability.

PRC Area (see Figure 3). This metric evaluates the precision-recall trade-off, particularly important when dealing with imbalanced datasets. A higher PRC Area indicates better precision and recall performance. Similar to the ROC Area, the PRC Area remains relatively stable across most scenarios, hovering around 0.55 to 0.58. Nonetheless, removing predictors related to mood or education level leads to a slight decrease in the PRC Area, suggesting these predictors might have some impact on the model's precision and recall.

Figure 1: ROC Area

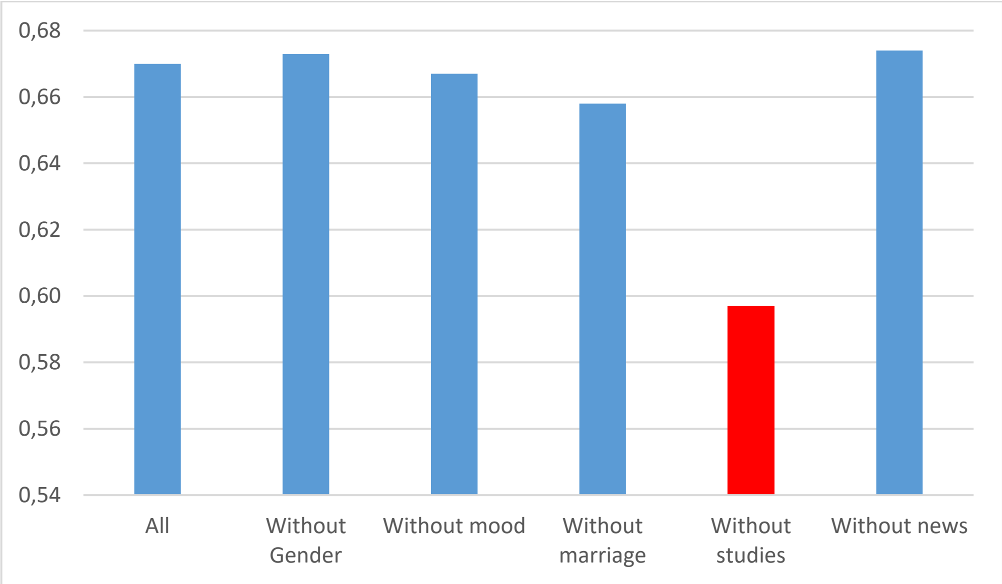
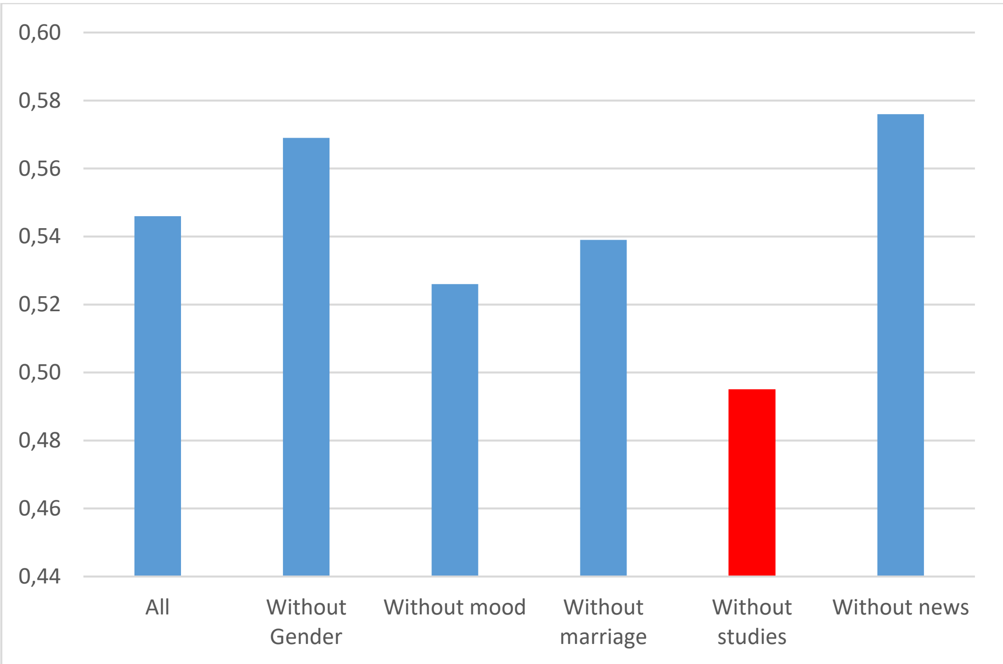


Figure 2: PRC Area



Overall, these results suggest that while certain predictors may have a minor influence on the model's performance, the overall predictive ability remains relatively consistent across different scenarios. All the validations reflect ROC and PRC area higher than 0.5, therefore, the models have predictable capacity and the hypothesis H_0 is validated.

6. CONCLUSIONS

In this paper, we have explored the relationship between personal characteristics and confidence in banks, recognizing the critical importance of confidence in financial institutions for the stability and prosperity of economies (De Leon; Xu, 2020; Ozili, 2020). Through a mixed methodology approach combining survey techniques and artificial intelligence modelling (Angelov et al.; Zang and Lu, 2021), we have gained valuable insights into the factors influencing confidence in banks.

Our findings suggest that personal characteristics such as age, gender, education level and perception of current issues play significant roles in shaping individuals' confidence in banks. While certain predictors may have a stronger influence than others, it is clear that a multifaceted understanding of confidence in banks requires consideration of various demographic and attitudinal factors.

The validation of our hypothesis, supported by metrics such as ROC Area and PRC Area, underscores the relevance of personal characteristics in predicting confidence in banks. Furthermore, the sensitivity analysis conducted by including and removing variables provides additional nuance to our understanding of the predictive power of different factors.

The implications of our findings extend to policymakers, financial institutions and researchers alike. Recognizing the importance of personal characteristics in shaping confidence in banks, policymakers can tailor regulatory frameworks and interventions to address specific demographic segments or attitudinal patterns. By fostering transparency, accessibility, and trust-building measures, policymakers can enhance overall confidence in the financial system.

Financial institutions can leverage the insights from this study to refine their customer engagement strategies and risk management practices. By understanding the demographic profiles and attitudes of their customers, banks can tailor their services and communications to better meet their needs and expectations, thereby, strengthening customer relationships and loyalty (Karkošková, 2023).

Researchers can build upon this study by delving deeper into the complex interplay between personal characteristics, confidence in banks and broader socio-economic dynamics. Future research could explore longitudinal data to assess how confidence in banks evolves over time and in response to external events such as financial crises or regulatory reforms. Additionally, qualitative research methods such as interviews or focus groups could provide deeper insights into the underlying motivations and perceptions driving confidence in banks.

Building on the foundation laid by this study, several avenues for future investigations emerge:

1. Longitudinal Studies: Conduct longitudinal studies to track changes in confidence in banks over time and identify key drivers of these changes.
2. Cross-Cultural Comparisons: Explore cross-cultural variations in confidence in banks to better understand how cultural factors influence attitudes towards financial institutions.
3. Qualitative Research: Employ qualitative research methods to uncover the nuanced motivations and perceptions underlying confidence in banks among different demographic groups.
4. Impact of Interventions: Evaluate the impact of regulatory interventions, financial literacy programs, and customer engagement initiatives on confidence in banks.
5. Machine Learning Techniques: Further refine machine learning models to enhance their predictive accuracy and robustness in predicting confidence in banks.

By addressing these areas of inquiry, future research can deepen our understanding of confidence in banks and contribute to the development of more effective policies and strategies for fostering financial stability and consumer trust.

7. ACKNOWLEDGMENTS

This paper has been financed by PP-2023-10 Research Project of Universidad Internacional de la Rioja.

8. REFERENCES

Andrade, J. L., & Valencia, J. L. (2021, December). Modeling lapse rates using machine learning: a comparison between survival forests and cox proportional hazards techniques. *Anales del Instituto de Actuarios Españoles, 27, 161-183.* https://doi.org/10.26360/2021_7

Andres, C., Betzer, A., Doumet, M. (2021): Measuring changes in credit risk: The case of CDS event studies, *Global Finance* Journal, *49,* 100647. https://doi.org/10.1016/j.gfj.2021.100647

Andros, S., Akimova, L., & Butkevich, O. (2020). Innovations in management of banks deposit portfolio: structure of customer deposit. *Marketing & Management of Innovations, (2).* https://doi.org/10.21272/mmi.2020.2-15

Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: an analytical review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *11(5), e1424.* https://doi.org/10.1002/widm.1424

Attwal, K. P. S., & Dhiman, A. S. (2020). Exploring data mining tool-Weka and using Weka to build and evaluate predictive models. *Advances and Applications in Mathematical Sciences*, *19(6), 451-469*

Bhardwaj, A., Mishra, S., & Jain, T. K. (2021). An analysis to understanding the job satisfaction of employees in banking industry. *Materials Today: Proceedings, 37, 170-174.* https://doi.org/10.1016/j.matpr.2020.04.783

Bigné, E. (2020). Teaching in business administration: a customized process driven by technological innovations. *Journal of Management and Business Education, 3(1), 4–15.* https://doi.org/10.35564/jmbe.2020.0002

Bouallègue, Z.B., & Richardson, D.S. (2022). On the ROC Area of Ensemble Forecasts for Rare Events, 27(5), 787–796. https://doi.org/10.1175/WAF-D-21-0195.1

Chauhan, J. S., Kaushal, D., & Pais, R. (2023). Factors Determining the Bank Selection Criteria of Consumers: A Banking Customers Survey. *European Economic Letters (EEL)*, *13*(*1*)*, 286-291.* https://doi.org/10.52783/eel.v13i1.169

De Leon, M. V., & Tubay, J. B. (2020). Ethical leadership of supervisors and internal social capital in a financial institution. *International Journal of Economics and Business Administration.* https://doi.org/10.35808/ijeba/462

Ekström, J., Åkerrén Ögren, J., & Sjöblom, T. (2023). Exact Probability Distribution for the ROC Area under Curve. *Cancers 15, 1788.* https://doi.org/10.3390/cancers15061788

El Fouki, M., Aknin, N., & El Kadiri, K. E. (2019). Towards an improved classification model based on deep Learning and nearest rules strategy. *Campus Virtuales, 8(1), 111-119*. http://www.uajournals.com/campusvirtuales/journal/14/9.pdf

Goulet Coulombe, Ph. (2024). The macroeconomy as a random forest. *Journal of Applied Econometrics (early view)*. https://doi.org/10.1002/jae.3030

Hammond, P., & Opoku, M.O. (2023). The mediating effect of going concern and corporate reporting in the relationship between corporate governance and investor confidence in financial institutions. *Heliyon 9(10), e20447*. https://doi.org/10.1016/j.heliyon.2023.e20447

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications, 124, 226-251*. https://doi.org/10.1016/j.eswa.2019.01.012

Huang, F., Xiong, H., Chen, S., Lv, Z., Huang, J., Chang, Z., & Catani, F. (2023). Slope stability prediction based on a long short-term memory neural network: comparisons with convolutional neural networks, support vector machines and random forest models. *International Journal of Coal Science & Technology 10 (18)*. https://doi.org/10.1007/s40789-023-00579-4

Isidro, H., Nanda, D., & Wysocki, P. D. (2020). On the relation between financial reporting quality and country attributes: Research challenges and opportunities. *The Accounting Review*, *95(3), 279-314.* https://dx.doi.org/10.2139/ssrn.3427028

Josephson, J., & Shapiro, J. (2020). Credit ratings and structured finance. *Journal of Financial Intermediation, 41, 100816.* https://doi.org/10.1016/j.jfi.2019.03.003

Karkošková, S. (2023). Data governance model to enhance data quality in financial institutions. *Information Systems Management*, *40(1), 90-110.* https://doi.org/10.1080/10580530.2022.2042628

Kuchciak, I., & Wiktorowicz, J. (2021). Empowering financial education by banks—social media as a modern channel. *Journal of Risk and Financial Management, 14(3), 118.* https://doi.org/10.3390/jrfm14030118

Kusi, B. A., Forson, J. A., Adu-Darko, E., & Agbloyor, E. (2023). Global financial crisis, international capital requirement and bank financial stability: an international evidence. *Journal of Financial Regulation and Compliance, 31(2), 237-258.* https://doi.org/10.1108/JFRC-04-2022-0057

Laeven, L., & Valencia, F. (2020). Systemic banking crises database II. *IMF Economic Review, 68, 307-361.* https://doi.org/10.1057/s41308-020-00107-3

Lee, S. J., Posenau, K. E., & Stebunovs, V. (2020). The anatomy of financial vulnerabilities and banking crises. *Journal of Banking & Finance, 112, 105334.* https://doi.org/10.1016/j.jbankfin.2018.04.013

Lei, X., Mohamad, U.H., Sarlan, A., Shutaywi, M., Daradkeh, Y. I., & Mohammed, H.O. (2022). Development of an intelligent information system for financial analysis depend on supervised machine learning algorithms. *Information Processing & Management, 59(5)*. https://doi.org/10.1016/j.ipm.2022.103036

Mangala, D., & Soni, L. (2023). A systematic literature review on frauds in banking sector. *Journal of Financial Crime, 30(1), 285-301.* https://doi.org/10.1108/JFC-12-2021-0263

Marqués, F. (2020). Machine Learning. Técnicas de Análisis Supervisado: Modelos Predictivos Multiecuacional: Independently Published

Martín, E., & Caballero, R. (2020). Las bases de Big Data. España: Los Libros de La Catarata

Mikhaylov, A., Dinçer, H., & Yüksel, S. (2023). Analysis of financial development and open innovation oriented fintech potential for emerging economies using an integrated decision-making approach of MF-X-DMA and golden cut bipolar q-ROFSs. *Financial Innovation*, *9(1), 1-34.* https://doi.org/10.1186/s40854-022-00399-6

Nasreen, S., Mahalik, M. K., Shahbaz, M., & Abbas, Q. (2020). How do financial globalization, institutions and economic growth impact financial sector development in European countries? *Research in International Business and Finance*, *54, 101247.* https://doi.org/10.1016/j.ribaf.2020.101247

Ozili, P. K. (2020). Theories of financial inclusion. In *Uncertainty and challenges in contemporary economic behaviour,* 89-115. Emerald Publishing Limited. https://doi.org/10.1108/978-1-80043-095-220201008

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior, 29(5), 615–620*. https://doi.org/10.1007/s10979-005-6832-7

Russell, R. (2018). Machine learning: guía paso a paso para implementar algoritmos de machine learning con Python: CreateSpace Independent Publishing Platform

Van der Cruijsen, C., de Haan, J., & Roerink, R. (2023). Trust in financial institutions: A survey. *Journal of economic surveys*, *37(4), 1214-1254.* https://doi.org/10.1111/joes.12468

Vanclay, F. (2020). Reflections on Social Impact Assessment in the 21st century. *Impact Assessment and Project Appraisal*, *38*(2)*, 126-131.* https://doi.org/10.1080/14615517.2019.1685807

Witten, I. H., & Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques, Second Edition. Netherland: Elsevier Science

Xu, X. (2020). Trust and financial inclusion: A cross-country study. *Finance Research Letters*, *35*, 101310. https://doi.org/10.1016/j.frl.2019.101310

Yadav, S., & Shukla, S. (2016, February). Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In *2016 IEEE 6th International conference on advanced computing (IACC) (pp. 78-83)*. IEEE. https://doi.org/10.1109/IACC.2016.25

Yego, N.K.K. , Nkurunziza, J., & Kasozi, J. (2023). Predicting health insurance uptake in Kenya using Random Forest: An analysis of socio-economic and demographic factors. *Plos One*. https://doi.org/10.1371/journal.pone.0294166

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. *Journal of Industrial Information Integration*, *23*, 100224. https://doi.org/10.1016/j.jii.2021.100224

9. ANNEX

9.1. Annex 1: Survey questions

- Predictors or personal characteristics
  - What is your age?
    - Under 24 years
    - 25 to 44 years old
    - 45 to 64 years old
    - Over 65 years old
  - Are you...
    - Male
    - Female
    - Other
  - Your marital status is...
    - Single
    - Married
    - Divorced
    - In a relationship
    - Other
  - Completed level of education
    - Primary
    - Secondary Education (ESO)
    - High School - Vocational Training
    - University
    - Postgraduate
    - Doctorate
  - Your financial literacy is (1-4)
  - Do you consider yourself up to date with current affairs?
    - Yes
    - No
  - To inform yourself, what channel do you prefer?
    - Digital media (Internet)
    - Written media (Newspaper)
    - Radio or television
  - If you had to search for information on an economic-political topic, would you first consult...
    - El Pais
    - El Mundo
    - 20 Minutos
    - El Español
    - ABC
    - El Confidencial
    - La Razón
    - Ok Diario
    - El Periódico
    - La Vanguardia
    - El Correo
    - La Voz de Galicia
    - Others
  - What team do you support?
    - Real Madrid
    - FC Barcelona
    - Atlético de Madrid
    - Other
    - I'm not interested in football
  - Has your team won its last match?
    - Yes
    - No
    - Draw
    - I don't follow any team
  - What is the weather like today?
    - Sunny and radiant

        • Cloudy
        • Raining
        • It's a terrible day
    - How do you feel today? (1-4)
• Confidence in financial institutions
    - How much trust do you place in banks? (1-4)

9.2. Annex 2:

```
=== Run information ===

Scheme:       weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:     Confianza en los agentes económico-políticos. (respuestas) - Respuestas de formulario 1-
weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.Remove-R13-
weka.filters.unsupervised.attribute.Remove-R13-15,17-23-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last
Instances:    60
Attributes:   13
              ¿Qué edad tienes?
              Eres...
              Estas...
              Nivel de estudios completado
              Tu cultura financiera es
              ¿Consideras que estás al tanto de la actualidad?
              Para informarte, qué canal prefieres
              Si tuvieses que buscar información sobre algún tema económico-político primero consultarías en
              ¿De qué equipo eres?
              ¿Ha ganado tu equipo su último partido?
              ¿Hoy hace un día?
              Hoy estás:
              ¿Qué confianza depositas en los bancos?
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 0.07 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          29               48.3333 %
Incorrectly Classified Instances        31               51.6667 %
Kappa statistic                          0.2024
Mean absolute error                      0.2912
Root mean squared error                  0.3898
Relative absolute error                 87.0619 %
Root relative squared error             95.5962 %
Total Number of Instances               60

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,273    0,163    0,273      0,273    0,273      0,109    0,727     0,297     1
                 0,630    0,485    0,515      0,630    0,567      0,145    0,645     0,620     2
                 0,368    0,146    0,538      0,368    0,438      0,251    0,647     0,560     3
                 0,667    0,018    0,667      0,667    0,667      0,649    0,836     0,699     4
Weighted Avg.    0,483    0,295    0,486      0,483    0,477      0,197    0,670     0,546

=== Confusion Matrix ===

  a  b  c  d   <-- classified as
  3  7  1  0 |  a = 1
  5 17  5  0 |  b = 2
  2  9  7  1 |  c = 3
  1  0  0  2 |  d = 4

*******************

=== Run information ===

Scheme:       weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:     Confianza en los agentes económico-políticos. (respuestas) - Respuestas de formulario 1-
weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.Remove-R13-
weka.filters.unsupervised.attribute.Remove-R13-15,17-23-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-R2
Instances:    60
Attributes:   12
              ¿Qué edad tienes?
              Estas...
```

```
              Nivel de estudios completado
              Tu cultura financiera es
              ¿Consideras que estás al tanto de la actualidad?
              Para informarte, qué canal prefieres
              Si tuvieses que buscar información sobre algún tema económico-político primero consultarías en
              ¿De qué equipo eres?
              ¿Ha ganado tu equipo su último partido?
              ¿Hoy hace un día?
              Hoy estás:
              ¿Qué confianza depositas en los bancos?
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          30               50      %
Incorrectly Classified Instances        30               50      %
Kappa statistic                          0.2119
Mean absolute error                      0.2904
Root mean squared error                  0.39
Relative absolute error                 86.8338 %
Root relative squared error             95.6321 %
Total Number of Instances               60

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,182    0,102    0,286      0,182    0,222      0,096    0,725     0,360     1
                 0,667    0,515    0,514      0,667    0,581      0,153    0,630     0,643     2
                 0,421    0,171    0,533      0,421    0,471      0,269    0,678     0,564     3
                 0,667    0,018    0,667      0,667    0,667      0,649    0,848     0,701     4
Weighted Avg.    0,500    0,305    0,486      0,500    0,484      0,204    0,673     0,569

=== Confusion Matrix ===

  a  b  c  d   <-- classified as
  2  8  1  0 |  a = 1
  3 18  6  0 |  b = 2
  1  9  8  1 |  c = 3
  1  0  0  2 |  d = 4

************************

=== Run information ===

Scheme:       weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:     Confianza en los agentes económico-políticos. (respuestas) - Respuestas de formulario 1-
weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.Remove-R13-
weka.filters.unsupervised.attribute.Remove-R13-15,17-23-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-R9-12
Instances:    60
Attributes:   9
              ¿Qué edad tienes?
              Eres...
              Estas...
              Nivel de estudios completado
              Tu cultura financiera es
              ¿Consideras que estás al tanto de la actualidad?
              Para informarte, qué canal prefieres
              Si tuvieses que buscar información sobre algún tema económico-político primero consultarías en
              ¿Qué confianza depositas en los bancos?
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 0.03 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          28               46.6667 %
Incorrectly Classified Instances        32               53.3333 %
Kappa statistic                          0.1823
Mean absolute error                      0.2841
```

```
Root mean squared error                  0.4018
Relative absolute error                 84.9516 %
Root relative squared error             98.5357 %
Total Number of Instances               60

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,364    0,163    0,333      0,364    0,348      0,194    0,737     0,396     1
                 0,630    0,455    0,531      0,630    0,576      0,175    0,618     0,592     2
                 0,316    0,171    0,462      0,316    0,375      0,164    0,666     0,522     3
                 0,333    0,035    0,333      0,333    0,333      0,298    0,868     0,431     4
Weighted Avg.    0,467    0,290    0,463      0,467    0,459      0,181    0,667     0,526

=== Confusion Matrix ===

  a  b  c  d   <-- classified as
  4  6  1  0 |  a = 1
  5 17  5  0 |  b = 2
  2  9  6  2 |  c = 3
  1  0  1  1 |  d = 4

*******************************

=== Run information ===

Scheme:       weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:     Confianza en los agentes económico-políticos. (respuestas) - Respuestas de formulario 1-
weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.Remove-R13-
weka.filters.unsupervised.attribute.Remove-R13-15,17-23-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-R3
Instances:    60
Attributes:   12
              ¿Qué edad tienes?
              Eres...
              Nivel de estudios completado
              Tu cultura financiera es
              ¿Consideras que estás al tanto de la actualidad?
              Para informarte, qué canal prefieres
              Si tuvieses que buscar información sobre algún tema económico-político primero consultarías en
              ¿De qué equipo eres?
              ¿Ha ganado tu equipo su último partido?
              ¿Hoy hace un día?
              Hoy estás:
              ¿Qué confianza depositas en los bancos?
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          30               50      %
Incorrectly Classified Instances        30               50      %
Kappa statistic                          0.2091
Mean absolute error                      0.2912
Root mean squared error                  0.392
Relative absolute error                 87.0459 %
Root relative squared error             96.117  %
Total Number of Instances               60

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,182    0,102    0,286      0,182    0,222      0,096    0,748     0,334     1
                 0,704    0,515    0,528      0,704    0,603      0,191    0,635     0,614     2
                 0,368    0,171    0,500      0,368    0,424      0,217    0,614     0,526     3
                 0,667    0,018    0,667      0,667    0,667      0,649    0,813     0,695     4
Weighted Avg.    0,500    0,305    0,482      0,500    0,480      0,205    0,658     0,539

=== Confusion Matrix ===

  a  b  c  d   <-- classified as
  2  7  2  0 |  a = 1
  3 19  5  0 |  b = 2
  1 10  7  1 |  c = 3
  1  0  0  2 |  d = 4

**********************************

=== Run information ===
```

```
Scheme:       weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:     Confianza en los agentes económico-políticos. (respuestas) - Respuestas de formulario 1-
weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.Remove-R13-
weka.filters.unsupervised.attribute.Remove-R13-15,17-23-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-R4-5
Instances:    60
Attributes:   11
              ¿Qué edad tienes?
              Eres...
              Estas...
              ¿Consideras que estás al tanto de la actualidad?
              Para informarte, qué canal prefieres
              Si tuvieses que buscar información sobre algún tema económico-político primero consultarías en
              ¿De qué equipo eres?
              ¿Ha ganado tu equipo su último partido?
              ¿Hoy hace un día?
              Hoy estás:
              ¿Qué confianza depositas en los bancos?
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          30               50      %
Incorrectly Classified Instances        30               50      %
Kappa statistic                          0.1979
Mean absolute error                      0.3039
Root mean squared error                  0.4036
Relative absolute error                 90.8675 %
Root relative squared error             98.9734 %
Total Number of Instances               60

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,091    0,082    0,200      0,091    0,125      0,013    0,610     0,247     1
                 0,667    0,545    0,500      0,667    0,571      0,123    0,574     0,555     2
                 0,474    0,195    0,529      0,474    0,500      0,288    0,585     0,520     3
                 0,667    0,000    1,000      0,667    0,800      0,809    0,836     0,699     4
Weighted Avg.    0,500    0,322    0,479      0,500    0,478      0,189    0,597     0,495

=== Confusion Matrix ===

  a  b  c  d   <-- classified as
  1  8  2  0 |  a = 1
  3 18  6  0 |  b = 2
  1  9  9  0 |  c = 3
  0  1  0  2 |  d = 4

********************************

=== Run information ===

Scheme:       weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation:     Confianza en los agentes económico-políticos. (respuestas) - Respuestas de formulario 1-
weka.filters.unsupervised.attribute.Remove-R1-weka.filters.unsupervised.attribute.Remove-R13-
weka.filters.unsupervised.attribute.Remove-R13-15,17-23-weka.filters.unsupervised.attribute.NumericToNominal-
Rfirst-last-weka.filters.unsupervised.attribute.Remove-R6-8
Instances:    60
Attributes:   10
              ¿Qué edad tienes?
              Eres...
              Estas...
              Nivel de estudios completado
              Tu cultura financiera es
              ¿De qué equipo eres?
              ¿Ha ganado tu equipo su último partido?
              ¿Hoy hace un día?
              Hoy estás:
              ¿Qué confianza depositas en los bancos?
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

RandomForest

Bagging with 100 iterations and base learner

weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities
```

```
Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          27               45      %
Incorrectly Classified Instances        33               55      %
Kappa statistic                          0.1301
Mean absolute error                      0.2834
Root mean squared error                  0.3906
Relative absolute error                 84.7135 %
Root relative squared error             95.7947 %
Total Number of Instances               60

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0,182    0,041    0,500      0,182    0,267      0,219    0,740     0,367     1
                 0,556    0,545    0,455      0,556    0,500      0,010    0,631     0,658     2
                 0,421    0,293    0,400      0,421    0,410      0,127    0,670     0,560     3
                 0,667    0,018    0,667      0,667    0,667      0,649    0,854     0,702     4
Weighted Avg.    0,450    0,346    0,456      0,450    0,437      0,117    0,674     0,576

=== Confusion Matrix ===

  a  b  c  d   <-- classified as
  2  8  1  0 |  a = 1
  1 15 11  0 |  b = 2
  1  9  8  1 |  c = 3
  0  1  0  2 |  d = 4

/*/*/*/*/*/*/*/*/*/*/*/*/


                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
Weighted Avg.    0,483    0,295    0,486      0,483    0,477      0,197    0,670     0,546
Weighted Avg.    0,500    0,305    0,486      0,500    0,484      0,204    0,673     0,569
Weighted Avg.    0,467    0,290    0,463      0,467    0,459      0,181    0,667     0,526
Weighted Avg.    0,500    0,305    0,482      0,500    0,480      0,205    0,658     0,539
Weighted Avg.    0,500    0,322    0,479      0,500    0,478      0,189    0,597     0,495
Weighted Avg.    0,450    0,346    0,456      0,450    0,437      0,117    0,674     0,576
```