

Harnessing Machine Learning for Predictive Insights in Mobile Banking: A Malaysian Perspective

*Note: Sub-titles are not captured in Xplore and should not be used

Hui Shan Lee*

*Faculty of Accountancy and
Management*

Universiti Tunku Abdul Rahman

Kajang, Malaysia

hslee@utar.edu.my

*Corresponding author

Kee Seng Kuang

*Lee Kong Chian Faculty of
Engineering and Science*

Universiti Tunku Abdul Rahman

Kajang, Malaysia

kuangks@utar.edu.my

Bik Kai Sia

*Faculty of Accountancy and
Management*

Universiti Tunku Abdul Rahman

Kajang, Malaysia

siabk@utar.edu.my

Kok Chin Khor

*Lee Kong Chian Faculty of
Engineering and Science*

Universiti Tunku Abdul Rahman

Kajang, Malaysia

kckhor@utar.edu.my

Abstract— Access to financial services, particularly mobile banking, helps mitigate poverty and strengthen the economic growth and resilience of a developing country like Malaysia. Although such access is important, and the Malaysian government has made notable progress in the mobile banking service, its widespread adoption is still lacking. In this study, we investigated and identified the important factors affecting mobile banking adoption in Malaysia. We utilized the J48 decision tree algorithm from WEKA to explain the demographic attributes that predict the adoption of mobile banking from a dataset provided by the Global Findex database. The results show that education level is the most influential factor, followed by age and employment status. Understanding these three factors can help shape the development of policies so that Malaysia can improve mobile banking adoption and build a more inclusive financial system. Applying the J48 decision tree algorithm is innovative in this context, as it offers a transparent and visual representation of decision-making through data splits, unlike traditional statistical models that can be harder to interpret due to complicated formulas.

Keywords—J48 decision tree, data mining, machine learning, digital banking, financial inclusion, education

I. INTRODUCTION

Mobile banking has become a significant drive in changing the global financial landscape as it has the potential for inclusion and promoting economic development. It is common nowadays for individuals to own at least one online banking account because this enables them to perform transactions, which is also a crucial step toward broader financial inclusion [1]. Such accounts act as gateways to various financial services and making global access to these accounts is the priority of the World Bank Group (WBG) [2]. Strategizing financial inclusion globally is the priority of WBG because the strategy aligns with seven of the 17 Sustainable Development Goals [3], [4]. These seven SDGs are, namely, eradicating poverty (SDG 1), ending hunger and promoting sustainable agriculture (SDG 2), promoting health and well-being (SDG 3), achieving gender equality and economic empowerment for women (SDG 5), promoting economic growth and jobs (SDG 8), supporting

industry, innovation, and infrastructure (SDG 9), and reducing inequality (SDG 10). Taking the previous COVID-19 outbreak as an example, the outbreak has highlighted the urgent need for expanding digital financial inclusion, especially to financially excluded and underserved populations. Such needs can be catered to by leveraging cost-effective digital solutions to offer formal financial services.

According to the Global Findex Database 2021 survey, 76% of the global adult population have a bank or mobile account, compared to only 51% in 2011 [5]. Fig 1 shows that account ownership in developing economies increased from 63% to 71% between 2017 and 2021, owing to services such as mobile money. Even though there is progress in financial inclusion, many adults in developing countries have no access to formal financial accounts, causing significant economic and social costs. Such exclusion limits them financially, such as saving, accessing credit, and participating in the formal economy, thus increasing their vulnerability to financial shocks and perpetuating poverty. Expanding financial access shall empower individuals in economic growth, leading to asset building, reducing transaction costs, and fostering participation in the global digital economy.

The percentage of account ownership by region and demographic factors is shown in Fig 2. The World Bank identifies the following factors as the key barriers to financial inclusion: education, employment, and gender inequality. Despite a global rise in financial inclusion, substantial gaps can still be seen in Malaysia, especially regarding demographic factors. The demographic factors influencing financial inclusion in Malaysia may be specific and differ from those of other developing countries analyzed by the World Bank. Understanding these disparities through ML models shall help Malaysia achieve financial inclusion better.

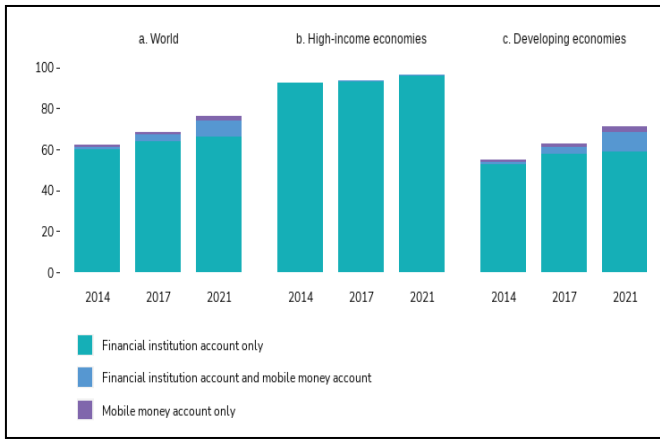


Fig. 1. Adults with an account (%) 2014-2021. Source: Extracted from [5]

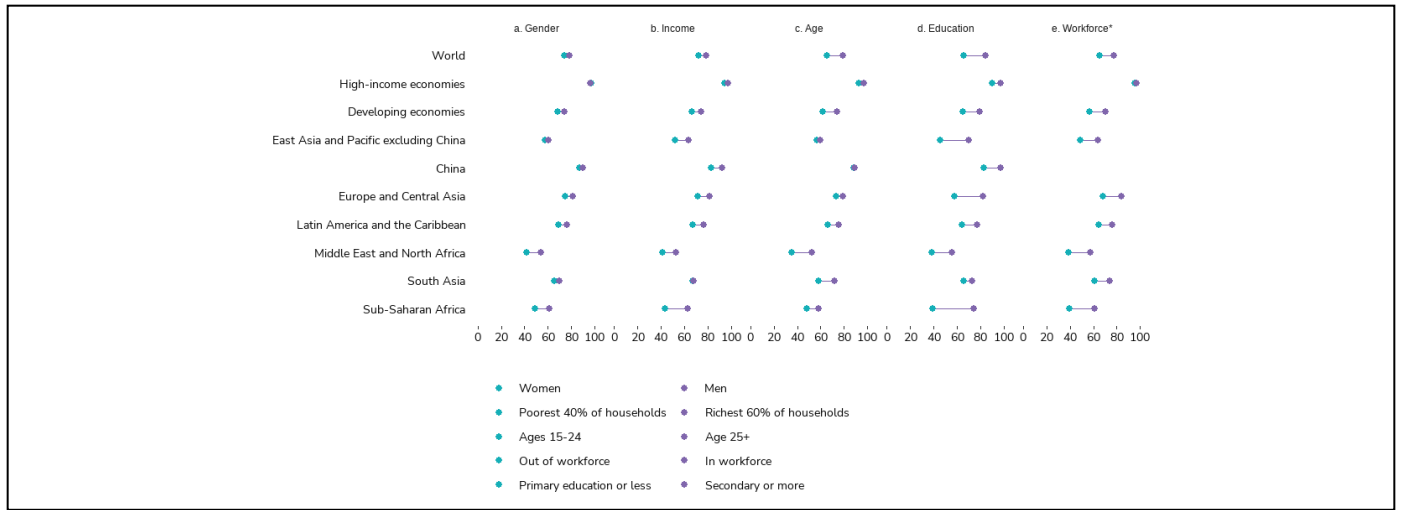


Fig. 2. Account ownership by regions and demographic factors in 2021. Source: Extracted from [5]

In Malaysia, providing financial services using mobile banking platforms can help communities to reduce poverty and foster economic strength. However, despite the apparent benefits, poor digital skills and low financial literacy limit the adoption of digital banking in the country [6]. Inevitably, the attempt to fully adopt mobile banking is still far from complete, even though the Malaysian government has laid a solid foundation for the banking infrastructure.

With the demographic factors, we use the predictive capabilities of J48, a decision tree ML algorithm, to provide insights into mobile banking adoption in Malaysia. We emphasize the demographic-specific interventions for a broader adoption of Malaysia's financial inclusion goals. Our study reveals that the strongest predictor of mobile banking adoption in Malaysia is education level. The likeliness of using mobile banking services is higher among those with higher education levels. Age and employment status are also identified as the key factors in the adoption of mobile banking services. Those young and employed individuals are likelier to the services.

Our findings shall contribute to improving financial inclusion in Malaysia through mobile banking. Since education is a critical factor, efforts to enhance digital literacy and financial

education should be poured to advance mobile banking usage in the country. We offer an alternative approach to the traditional statistical approaches commonly used in financial inclusion research. The alternative approach was conducted by leveraging the J48 algorithm to provide predictive insights into the demographic factors that drive mobile banking adoption in Malaysia.

II. LITERATURE REVIEW

A. Theoretical Development

There are three theories that support the demographic attributes to explain mobile banking adoption, namely, the Technology Acceptance Model (TAM) [7], the Unified Theory of Acceptance and Use of Technology (UTAUT) [8], and the Financial Literacy Theory [9]. TAM contends that perceived usefulness and ease of use are the major technology adoption

factors. This model can assist in explaining how people's perceptions of mobile banking's benefits and how usability influence their decision to use these services, especially regarding education and digital literacy. On the other hand, UTAUT broadens TAM by including aspects such as social impact, conducive environments, and performance expectations. This theory can provide a thorough knowledge of how demographic factors such as age and work status interact with social and organizational factors to influence mobile banking uptake. Furthermore, Financial Literacy Theory emphasizes the importance of an individual's knowledge and grasp of financial concepts in making sound financial decisions. It can explain how education improves people's capacity to understand and effectively use mobile banking services.

B. Empirical Studies

Traditionally, researchers in the research field relied on financial ratios and econometric models to assess performance. It is common now for finance researchers to build ML models from their data. The models discover data patterns and allow the researcher to make decisions. Many studies have now aligned with the current trend of applying ML to the financial sector.

The study by [10] predicts the operational performance and efficiency of selected banking sectors in India by using ML

models, i.e., decision trees, random forests, etc. Other studies, i.e., [11] and [12], showed the advantages of utilizing ML for better predictions, risk assessments, and operational efficiencies.

ML models with good predictive capabilities are crucial for the financial industry to understand customer behavior and optimize their marketing efforts to cater to the increasing competition and the need for personalized customer engagement. For example, the studies by [13] and [14] predict customer behavior in the financial industry. With the approaches proposed by the two studies, decisions made by financial institutions can be more data-driven.

The existing literature strongly focuses on applying ML models in various aspects of the banking sector. Nevertheless, a gap exists in utilizing ML in mobile banking adoption in emerging markets like Malaysia. Limited attention has been given to how ML techniques can help with the following. Firstly, predicting the demographic factors that drive mobile banking uptake. Secondly, exploring the potential of ML algorithms to optimize customer experience, personalize mobile banking services, and enhance user engagement in the context of Malaysia's diverse market.

As Malaysia has a digital landscape that is now rapidly evolving, the insights gained by this study can thus influence future banking strategies in the country. With ML, this study provides an option to close the gap for shaping the future of mobile banking in Malaysia by predicting mobile banking adoption and offering valuable insights into demographic factors.

III. DATA AND METHODOLOGY

A. Background of the research data

We used the data from the 2021 World Bank Global Findex survey. The dataset is yielded by the Global Findex database that was launched by the World Bank in 2011. This database has a comprehensive source of information on adult savings, borrowing, payments, and risk management.

The dataset we used includes responses from over 150,000 individuals across 144 countries based on nationally representative surveys; it provides updated insights into using formal and informal financial services and access to them. In Malaysia, data was collected from 1,000 participants between December 17, 2021, and January 28, 2022, through landline and mobile phone interviews conducted in English, Malay, or Chinese. 419 (41.9%) of these participants are mobile bank users (label Yes), and the remaining 581 (58.1%) are non-mobile bank users (label No). We briefly describe the seven attributes of the dataset in Table 1. The data type of the attributes is categorical.

TABLE I. DATASET DESCRIPTION

No	Attribute Name	Possible Values
1	Gender	F / M
2	Education	Primary/ Secondary/ Tertiary
3	Income	Highest/ Middle/ Lowest
4	Employment	Employed/ Unemployed

5	Urbanicity	Urban/ Rural
6	Age	<20 / 20-39 / 40-59 / >60
7	Mobile banking user	Yes / No

B. Classification algorithms

The J48 algorithm from the WEKA software was used on this dataset to develop a predictive model that classifies mobile banking users or non-mobile banking users. When a model is built, a flowchart-like tree structure shall be developed, which can be used to predict the banking users. By analyzing its structure, we can identify the most significant attributes influencing the banking users. The J48 decision tree algorithm was chosen due to its strong statistical significance and interpretative capabilities.

IV. FINDINGS

The flowchart-like tree structure illustrated in Fig. 3 was the model representation of the J48 algorithm. It comprises one root node, several intermediate lower-level nodes, and leaves or decision nodes. The structure's ease of interpretation makes the data easy to understand and evaluate. Interpreting the structure from the top root node to the bottom decision nodes shall generate rules that help the model decide whether or not a user is a mobile banking user. The algorithm pruned the constructed tree to enhance generalization. The final tree consists of 24 nodes, with 16 decision nodes representing "Yes" and "No".

Table II shows the performance evaluation of the J48 model. Since the algorithm was employed on this relatively balanced dataset, accuracy is thus important for evaluating the model's overall performance. The average accuracy of the model is 70.1%, demonstrating the model's capability in classifying banking users. The model's classification performance is considered satisfactory because it exceeds the random chance rate of over 50%. Additionally, it is comparable to the rates found in the reviewed literature, further affirming its acceptability [15].

TABLE II. PERFORMANCE EVALUATION OF THE J48 MODEL USING 10-FOLD CROSS-VALIDATION

Correctly Classified Instances	701	70.1%
Incorrectly Classified Instances	299	29.9%
Kappa Statistic	0.3853	
Root Mean Squared Error	0.3652	
Root Mean Squared Error	0.4362	
Relative Absolute Error	75.0062%	
Root Relative Squared Error	88.4047%	
Total Number of Instances	1000	

Note: the j48 is optimized with parameter value M=5, where M is the minimum number of instances per leaf

We presented our performance evaluation on the J48 model using the confusion matrix shown in Table III. The model accurately identified 268 instances as mobile banking users (Yes) and 433 as non-mobile banking users (No). On the other hand, 299 (151+148) instances were classified incorrectly, with

a slightly higher misclassification rate for non-mobile banking users compared to mobile banking users.

TABLE III. THE CONFUSION MATRIX OF THE BUILT MODEL

Predicted values	Actual values	
	Yes	No
Yes	268	151
No	148	433

Researchers commonly use accuracy (or True Positive rate) to evaluate a model's performance. However, accuracy does not

relevant attribute for splitting the training set and lead to decision nodes. Fig. 3 shows the flowchart-like tree structure of the J48 model we built. The first split is the "education" attribute, indicating that the education level highly affects the adoption of mobile banking. This split is crucial because it determines the most influential attribute in separating the data into classes, determining whether or not an individual is a mobile banking user. Higher-educated individuals may be more tech-savvy and understand the advantages of mobile banking; thus, they are more likely to use the technology. Conversely, less educated individuals may be less familiar with online banking technologies and are concerned about security and usability; thus, the adoption rate of online banking is low.

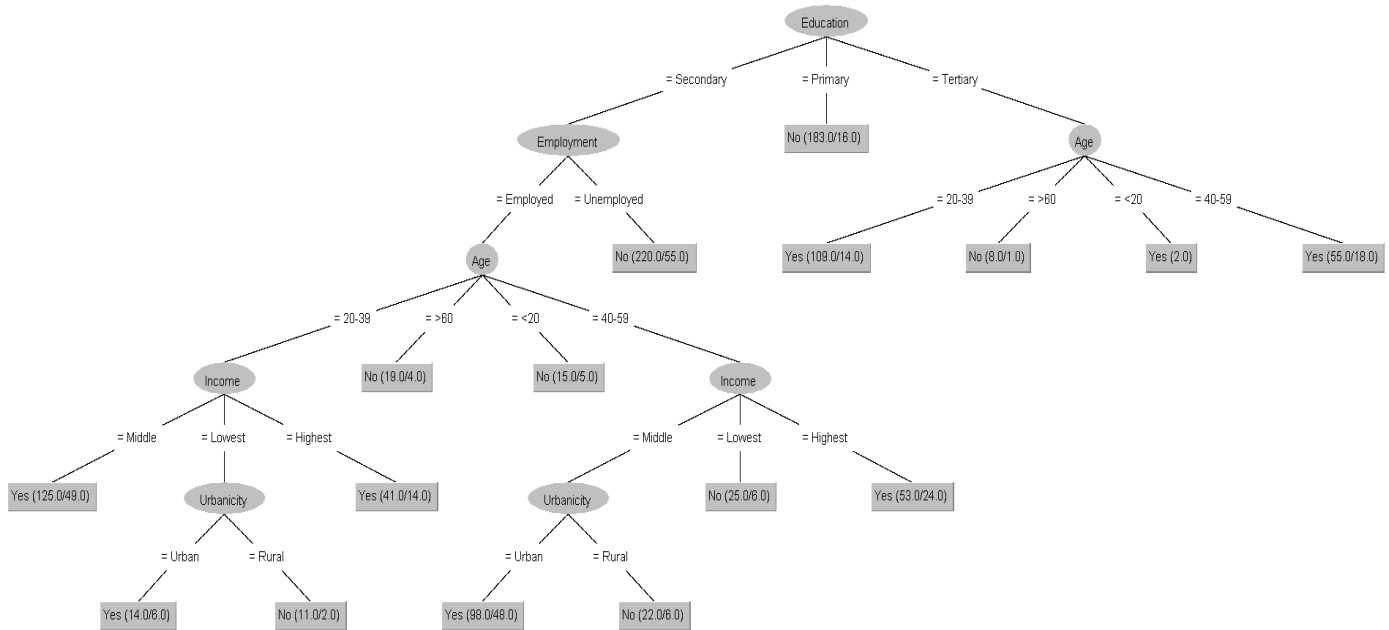


Fig. 3. J48 Decision Tree Algorithms in predicting demographics attributes on mobile banking users

give us a comprehensive understanding of the model's performance. Accuracy provides an overall classification performance evaluation of a model but does not reveal a model's performance on each class label (like Yes and No in our study). To examine the model performance for each class, we used the True Positive rate (TP rate) (Table IV). The TP rate for class mobile banking users (Yes) is 0.640, while the TP rate is 0.745 for non-mobile banking users. Both accuracy and TP rates show that the J48 model has a satisfactory classification performance in overall and for each class.

TABLE IV. CLASSIFIER OUTPUT

TP Rate (or Accuracy)	0.701
TP Rate (Yes)	0.640
TP Rate (No)	0.745
Precision	0.701

The J48 algorithm used in this study reveals the impact of the dataset attributes on mobile banking adoption. The splitting criteria used in this decision tree algorithm can identify the most

The splitting is followed by the "employment" and "age" attributes. "Employment" has been identified by the algorithm as a key attribute for further differentiating mobile and non-mobile banking users based on their education levels. Employed individuals have steady incomes, which correlates with the use of mobile banking. Due to their work nature, they are likely exposed to technology at work; thus, it is more comfortable or convenient for them to use mobile banking services. Unemployed individuals or informal employees may be inclined to traditional banking services and adopt mobile banking services less.

Splitting the training set using the "age" attribute by the algorithm also shows differences in mobile banking usage across age groups. Younger individuals below age 60 are more comfortable with mobile banking technology as the technology can provide convenience, speed, and integration with other online services. On the other hand, older individuals above 60 years old may be reluctant to use mobile banking services. They are less tech-savvy and concerned with the security of the

services. Usually, older individuals prefer traditional banking methods.

V. CONCLUSIONS

A key driver to global financing transformation is mobile banking. Mobile banking offers significant potential for financial inclusion and economic growth. Owning a mobile banking account is important so that individuals can store funds and conduct transactions. In addition, these accounts also serve as gateways to various financial services locally and globally. Through the Universal Financial Access strategy, WBG states that owning these accounts to achieve global access aligns with several SDGs of the United Nations.

Despite advancements in mobile banking services in developing countries, including Malaysia, many adults still lack access to the services. Mobile banking can enhance financial inclusion in Malaysia, but low digital skills and financial literacy hinder it. This study aims to use a machine learning algorithm, the J48 decision tree, to identify the demographic factors that influence the adoption of mobile banking. The results highlight that education level, age, and employment status are the key demographic attributes influencing mobile banking adoption. Education emerged as the strongest out of these three demographic attributes. It implies that higher education levels increase the likelihood of mobile banking usage. Our study offers several theoretical and practical implications, which shall be discussed as follows.

A. Theoretical implications

According to TAM, perceived ease of use and usefulness are important in promoting mobile banking adoption. Education shapes these perceptions as educated individuals commonly comprehend and use mobile banking services well. Higher-educated individuals usually find mobile banking more intuitive and beneficial; thus, it leads to higher adoption rates for this individual group. The theoretical understanding that education is an important factor in mobile banking adoption aligns well with our study's results. Our study also reinforces the Financial Literacy Theory. The theory suggests that individuals with greater financial literacy are more likely to make informed decisions about financial products. Educated individuals have a better understanding of the benefits of mobile banking, thus leading to a higher adoption rate.

Our findings regard employment and age as key factors to mobile banking adoption, which can be theoretically tied to UTAUT. Formally employed individuals may face greater performance expectancy (i.e., mobile banking helps them manage their finances efficiently) and facilitating conditions (access to smartphones, digital literacy, etc.), which increase their likelihood of adoption. On the other hand, age highlights differences in effort expectancy. Younger users may find mobile banking easy and convenient to use. It is opposed to older users who perceive more effort and complexity, thus discouraging adoption.

B. Practical implications

Banks can adjust their marketing strategies accordingly to improve mobile banking adoption. Conducting awareness campaigns about the benefits and security of mobile banking

could be helpful to those with lower education levels. Adding focuses like security, ease of use, and personal support in mobile banking to the campaigns can also help older individuals.

Providing feedback by banks to make mobile banking more accessible and user-friendly helps individuals with different education attainments. Particularly for younger individuals, banks can also design features that suit them, such as seamless integration with other financial tools, instant notifications, and mobile-first services.

Different marketing strategies can be set for individuals with different employment statuses. Marketing messages regarding the convenience of mobile banking can be sent to employed individuals. On the other hand, strategies emphasizing ease of use and low-cost mobile banking can also be formed for unemployed or informally employed individuals.

Legal stability significantly impacts the growth of financial services [16]. Thus, the government is pivotal in creating robust legal frameworks to safeguard consumer rights, transactions, and regulations in mobile banking. Addressing these areas helps the government to reduce barriers in education, employment, and age to banking adoption. Banks can then develop targeted marketing strategies and design services for their customers, knowing there is support for a legal and secure mobile banking environment.

In conclusion, mobile banking adoption is influenced by education, employment, and age. Understanding these factors allows banks to better tailor marketing strategies, product offerings, and customer support to reach different demographic groups. So, when promoting mobile banking services, banks must consider demographic realities in addition to technological features.

REFERENCES

- [1] S. Yap, H. S. Lee, and P. X. Liew, "The Roles of Insurance and Banking Services on Financial Inclusion," *SAGE Open*, vol. 14, no. 2, pp. 1–19, 2024, doi: 10.1177/21582440241252268.
- [2] The World Bank, "Financial Inclusion Overview," 2021. <https://www.worldbank.org/en/topic/financialinclusion/overview> (accessed Sep. 18, 2024).
- [3] S. Yap, H. S. Lee, and P. X. Liew, "The role of financial inclusion in achieving finance-related sustainable development goals (SDGs): A cross-country analysis," *Econ. Res. Istraživanja*, vol. 36, no. 3, p., 2023, doi: 10.1080/1331677X.2023.2212028.
- [4] The World Bank, "UFA2020 Overview: Universal Financial Access by 2020," 2018. <https://www.worldbank.org/en/topic/financialinclusion/brief/achieving-universal-financial-access-by-2020> (accessed Sep. 28, 2024).
- [5] The Global Findex Database, "The Global Findex Database 2021," 2023. <https://www.worldbank.org/en/publication/globalfindex> (accessed Sep. 28, 2024).
- [6] Bank Negara Malaysia, "Annual Report 2023," 2024. doi: 10.3934/energy.2024013.
- [7] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Q. Manag. Inf. Syst.*, vol. 13, no. 3, pp. 319–339, 1989, doi: 10.2307/249008.
- [8] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User Acceptance of Information Technology: Toward a Unified View," vol. 27, no. 3, pp. 425–478, 2003.

- [9] A. Lusardi and O. S. Mitchell, "Financial Literacy and Retirement Planning: New Evidence from the Rand American Life Panel," *SSRN Electron. J.*, 2011, doi: 10.2139/ssrn.1095869.
- [10] A. Joshi, S. Sharma, N. V. M. Rao, and A. K. Vaish, "Usage of Machine Learning Algorithm Models to Predict Operational Efficiency Performance of Selected Banking Sectors of India," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 12, no. 6, pp. 105–114, 2022, doi: 10.46338/ijetae0622_14.
- [11] G. Samara *et al.*, "Securing cryptocurrency transactions: Innovations in malware detection using machine learning," *Int. J. Data Netw. Sci.*, vol. 8, no. 4, pp. 2055–2066, 2024, doi: 10.5267/j.ijdns.2024.7.003.
- [12] Y. Liu, "View of Discussion on the Enterprise Financial Risk Management Framework Based on AI Fintech," *Decis. Mak. Appl. Manag. Eng.*, vol. 7, no. 1, pp. 254–269, 2024.
- [13] W. Saengthongrattachot, A. Na-Udom, and J. Rungrattanaubol, "A Comparison of Machine Learning Techniques for Classification in Bank Marketing Data," *Thai J. Math.*, vol. 2022, no. Special Issue, pp. 157–168, 2022.
- [14] F. Safarkhani and S. Moro, "Improving the accuracy of predicting bank depositor's behavior using a decision tree," *Appl. Sci.*, vol. 11, no. 19, 2021, doi: 10.3390/app11199016.
- [15] G. Miari, Z. Almeshaima, E. Hanafy, A. Muftah, and A. Zaki, "A Predictive Approach to Mobile-only Banking Adoption in Bahrain: Evidence from Decision Trees," *2023 4th Int. Conf. Data Anal. Bus. Ind. ICDABI 2023*, pp. 295–300, 2023, doi: 10.1109/ICDABI60145.2023.10629233.
- [16] H. S. Lee, F. F. Cheng, S. C. Chong, and B. K. Sia, "Influence of macroeconomics factors and legal stability to the insurance growth in the ASEAN-5 countries," *J. Ekon. Malaysia*, vol. 52, no. 2, pp. 219–229, 2018, doi: 10.17576/JEM-2018-5202-18.