Assessing the Importance of ICT on High-Educated Poverty in East Java Using Random Forest Methods

Dias Satria[1] | *University of Brawijaya, Malang, Indonesia*
Radityo Putro Handrito | *University of Brawijaya, Malang, Indonesia*
Girindra Mega Paksi | *University of Brawijaya, Malang, Indonesia*
Patricia Juan Aurellia | *University of Brawijaya, Malang, Indonesia*

Received 3.8.2023 (revision received 30.10.2023), **Accepted** (reviewed) 28.11.2023, **Published** 15.3.2024

Abstract

Poverty is one of the main issues in the economic development of a country, including Indonesia. This was exacerbated when the Covid-19 pandemic had an impact on economic activities, so there are educated people who are still below the poverty line. Therefore, strategies in tackling the problem of poverty need to be developed in line with the times that have entered the era of digitalization. The research aimed to examine the importance of Information and Communication Technologies (ICTs) on high-educated poverty in East Java. Data obtained from the National Survey in Indonesia is analyzed with the random forest algorithm as a classification method in machine learning. The analysis shows that social media, laptop ownership, age, higher education, and internet access are the five most important variables for high-educated poverty classification in East Java. Based on this, a number of recommendations can be made to policy makers regarding the effect of digitalization on high-educated poverty in East Java.

Keywords	***DOI***	***JEL code***
Education, digitalization, machine learning, poverty, random forest	*https://doi.org/10.54694/stat.2023.38*	*I24, I32, O12*

INTRODUCTION

The Covid-19 pandemic has exacerbated global poverty levels. Lockdowns, business closures and travel restrictions have affected the informal sector and the small and medium enterprise sector, and millions of workers in these sectors have lost their jobs. To mitigate the pandemic‘s impact on global poverty,

[1] University of Brawijaya, Jl. Veteran No 10–11, Ketawanggede, Kec. Lowokwaru, Kota Malang, Jawa Timur 65145, Indonesia. Corresponding author: e-mail: dias.satria@ub.ac.id, phone: (+62)81333828319. ORCID: <https://orcid.org/0000-0002-4068-6807>.

many countries and international organizations have increased aid and financial support, such as direct cash transfers, food subsidies, and other social assistance programs. However, these efforts are far from enough, and more support is needed to strengthen social security systems in the future and ensure that people most vulnerable to poverty are protected.

Poverty is one of the main issues affecting human development (Nili, 2019). Indonesia is one of the countries with 9.78% poverty or around 27 million people living below the poverty line (Finch and Finch, 2020). Poverty in Indonesia is high when compared to countries in Southeast Asia, such as Malaysia and Thailand, which are only 0.4% and 6.1%. The poverty rate in Indonesia is still high when compared to international standards. According to World Bank standards, a person is categorized as living in poverty if income is less than US$1.90 per day (Diop and Asongu, 2021). In this case, Indonesia still has a lot of homework to do to overcome the problem of poverty. Indonesia has 38 provinces with various characteristics. One of the provinces in Indonesia that has a considerable economic potential is East Java. As the second largest province in Indonesia after Papua, East Java has quite diverse economic sectors, such as industry, agriculture, fisheries, and tourism. Despite its huge economic potential, East Java is also one of the provinces with a significant poverty rate in Indonesia. Therefore, a study of poverty in East Java can provide an overview of the factors that cause high poverty rates.

Poverty rates must be reduced for several important reasons, including the human right to have an equal access to resources and economic opportunities. A high poverty rate indicates that a large number of individuals do not have sufficient access to the resources necessary to satisfy basic needs such as food, clothing and housing. In addition, poverty can reduce people‘s quality of life and social well-being. High poverty rates can also hinder economic progress, as poor individuals have fewer resources and less capacity to actively participate in the economy (Abduvaliev and Bustillo, 2019). Similarly, in terms of peace, a high level of poverty can lead to social tensions and exacerbate social conflicts, thus threatening the stability and peace of society.

In recent years, the impact of technology on poverty has gained an increased attention and interest. Technology has the potential to improve economic opportunities and access to essential services, yet there are also concerns about the unequal distribution of technology and its negative effects (Chernova et al., 2019). The relationship between technology and poverty is complex and multi-faceted, and the full impact of technology on poverty is still not well understood. This highlights the need for more comprehensive and in-depth research to better understand the complex relationship between technology and poverty, and to develop effective solutions to address poverty and promote economic growth.

Based on Central Bureau of Statistics publication data for 2020, there are 19.5% of the poor population in East Java who have completed 12 years of compulsory education. In other words, there are still people with high school, vocational, and even university education who are still below the poverty line, called high-educated poverty. Over the past few years, the proportion of poor people with secondary or higher education in East Java has also shown an increasing number. There are several potential impacts if there are people with high education but classified as poor, including: poor individuals may struggle to secure jobs that match their qualifications, or even access to business opportunities or capital. This can hinder their ability to improve their standard of living and fully leverage their skills. Poor individuals may also lack equal access to social resources such as connections, information, or family and friend support. This can impact their mental health and overall well-being. Therefore, currently many factors that can affect poverty are increasingly complex. Not only about education, but also matters related to skills and the times, namely digitalization.

The purpose of this study is to address the research gap by exploring the relationship of technology on high-educated poverty in East Java. The research question for this study is as follows: Does technology, the use of social media, e-commerce, and digital banking, have an impact on high-educated poverty in East Java? To address this research question, a literature review was conducted to examine existing research on the relationship between technology and poverty and we focused on high-educated poverty in East Java. The literature review revealed that there have been studies examining the potential benefits of technology in alleviating poverty, such as improving access to education and financial services. However, the full impact of technology on poverty is still not well understood, and more comprehensive and in-depth research is needed to better understand the complex relationship between technology and poverty.

To analyze the relationship between technology and poverty, we will be using the machine learning method called Random Forest (RF). RF is a popular machine learning algorithm that can handle large datasets and complex relationships between variables. This method will allow us to examine the relationship between various factors, including technology use and access, socio-economic status, and poverty. RF is a robust and efficient method that can handle both continuous and categorical variables, and it can also handle missing data and outliers (Ao et al., 2019). The results generated by RF will provide insights into the factors that contribute to poverty and the importance of technology on the model of high-educational poverty in East Java. With the use of Susenas data and the RF method, this study aims to fill the research gap in understanding the complex relationship between technology and poverty and to inform policy decisions to address poverty and promote economic growth in East Java.

The policy implications of this study are important, as the findings can inform policy-making decisions related to technology and poverty. For instance, the results of this study can provide valuable insights into the need for government programs and initiatives to address the unequal distribution of technology and ensure that its benefits are distributed equitably. Additionally, the results of this study can inform the development of policies and initiatives aimed at improving digital literacy and technology access for those in poverty, to ensure that they are able to fully benefit from the opportunities that technology can provide.

Based on the backgrounds above, this study was conducted to exploring the importance of technology on high-educated poverty in East Java using machine learning method such as RF. This study is organized into five sections. Introduction explains the Covid-19 impact on Indonesia, especially high-educated poverty in East Java, and how the digitalization related to skills. Section 1 provides theoretical support and past studies related to high-educated poverty and digitalization in Indonesia and worldwide. Section 2 provides information about research analysis and tools to know the importance of digitalization on high-educated poverty in East Java using RF. Section 3 discusses the results of the analysis and explains the types of digitalization that most importance to the model of high-educated poverty in East Java. The last section presents the concluding remarks.

1 LITERATURE REVIEW

In terms of poverty, Hofmarcher (2021) found an important role of the education sector in poverty alleviation programs in Europe. Meo et al. (2020) gave a fresh idea about the non-linear relationship between unemployment, governance, and poverty in Pakistan. Long-term unemployment will create too much immoral crime in countries such as frustration, homelessness, family tension, loss of confidence, social isolation, self-esteem, and poverty (Siddiqa, 2021). In Indonesia, many studies have investigated the determinants and nature of poverty as it relates to poverty reduction, unemployment, and national economic growth. For example, Erlando et al. (2020) found a relationship between economic growth, inequality and poverty. Investment has a direct effect on poverty, while economic growth has no direct effect on poverty. Poverty in Indonesia is also caused by many other factors, including population growth, investment, education, health, market structure, and government regulation (Jacobus et al., 2019).

In this era, we know about digitalization. Digitalization is the process of transforming traditional processes or activities into digital form through the use of digital technologies such as computers, the internet, and other electronic devices. In digitalization, data that was previously stored in physical form, such as paper or books, is recorded and stored in digital form. Digitalization enables faster, more efficient, and easily accessible access to information. Digitalization has impacted many aspects of human life such as communication, education, health, entertainment, and business. Digital technology can enable new business models in the form of digital services, digital platforms, digital tools or infrastructure (McQuire, 2021), digital artifacts (Sebastiani, 2021), or service innovations enabled through the Internet (Vakulenko et al., 2019).

E-commerce is a form of digitalization because it enables the digitization of business transactions, which previously could only be conducted physically. The rapid advancement of information and communication technologies, combined with globalization, has resulted in a significant impact on economic life through electronic commerce. E-commerce has become accessible and convenient, with applications and progress seen in all economic sectors. Its availability on the internet and instantaneous nature are the main reasons behind these developments. E-commerce‘s features have made it widely used in economic life, with its usage becoming increasingly diverse as more people use desktop computers and mobile phones (Guven, 2020). With mobile applications, transactions can be made at any time and place. Additionally, e-commerce also allows for the digitization of promotion, where businesses can promote their products or services through digital platforms such as social media.

Social media is an online platform that enables individuals or groups to interact, communicate, and share information over the internet (Lovari and Valentini, 2020). Before the existence of social media, social interactions were more physical and limited by geographical distance. However, with the advent of social media, social interaction can occur online, where people can interact with others who are located all around the world without being hindered by geographical distance. Social media also enable the digitalization of content, where information and content such as images, videos, or text can be uploaded and accessed online. Thus, social media allow for the digitalization of social interaction and content, which is one aspect of digitalization in general.

Digital banking refers to banking services that are provided through digital channels such as mobile applications, internet banking, and other online platforms. Digital banking has become increasingly popular in recent years due to advancements in technology and the growing trend of digitalization. Digital banking is a type of digitalization that has a strong connection to business. Banks around the globe are shifting towards digital banking services to offer their customers convenience and speed. These services involve opening accounts, transferring funds, paying bills, and more (Wewege et al., 2020). One of the main advantages of digital banking is its ease and convenience. Customers can access banking services at any time and from any location using electronic devices such as smartphones, tablets, or laptops. Furthermore, digital banking provides lower transaction costs than traditional banking services. By using digital banking, businesses can accelerate transaction processes, save time and cut costs. As a result, digital banking is increasingly becoming the go-to solution for businesses seeking to boost efficiency and productivity in their operations.

The literature for this study provides a thorough examination of the research done on the connection between ICTs access to finance, and poverty. It also offers an overview of the research on how new technologies can promote digital financial services. The growing body of literature suggests that ICTs play a significant role in promoting economic and social improvement. Research also indicates that improved access to finance can help reduce poverty and enhance household welfare. According to Kendall et al. (2010), developed countries have greater access to finance compared to developing countries. ICTs can provide access to financial services, information, and e-banking to the poor, especially in remote areas. Isukul and Tantua (2021), and Mishra and Singh Bisht (2013) found that mobile technology was an effective tool for financial inclusion in remote regions. Chatterjee (2020), and Pradhan et al. (2021) confirmed the positive impact of ICTs on economic growth through financial inclusion. Bhavnani et al. (2008) demonstrated the increased social and economic benefits of mobile phones in rural areas and predicted the poverty alleviation effects of mobile phones. Chib et al. (2015) reviewed the ICTs literature and progress in measuring its impact and concluded that prior to 2002, researchers focused more on macro-level ICTs linkages, while there was less attention given to its role in poverty alleviation at the micro level.

Research on poverty has been conducted with various methods of analysis. Faharuddin and Endrawati (2022) classified poverty in Indonesia with statistical learning with logistic regression analysis. But, such a method has high requirements on the sample size of the data, so that modeling of poverty in a region continues to grow. Alsharkawi et al. (2021) conducted research on multidimensional poverty in Jordan with machine learning with the LightGBM algorithm and Bagged Decision Tree. Yao et al. (2023) conducted research to estimate poverty reduction in China spatially by utilizing deep learning Long Short Term Memory (LSTM) analysis, a neural network technique. Kaur and Kaur (2020) conducted a study comparing several classification methods in machine learning based on their accuracy. The results showed that Random Forest is better than Logistic Regression, Decision Tree, and K-Nearest Neighbors. Based on previous research, we employed the RF algorithm as a machine learning method to explore the factors associated with educated poverty in East Java. It is important to note that the use of RF aims to analyze relationships and associations between variables, and it does not establish causal relationships.

The literature review provides a comprehensive understanding of the relationship between ICTs, access to finance, and poverty. The studies reviewed suggest that ICTs play a significant role in promoting economic and social improvement and improved access to finance can help reduce poverty and enhance household welfare. The literature highlights the positive impact of ICTs in providing financial services, information, and e-banking to the poor. The review further highlights the shift in focus from macro-level ICTs linkages to its role in poverty alleviation at the micro level in recent years. Overall, the literature supports the idea that ICTs can have a positive impact on poverty alleviation through improved access to finance.

2 DATA AND METHODOLOGY

2.1 Data sources and description

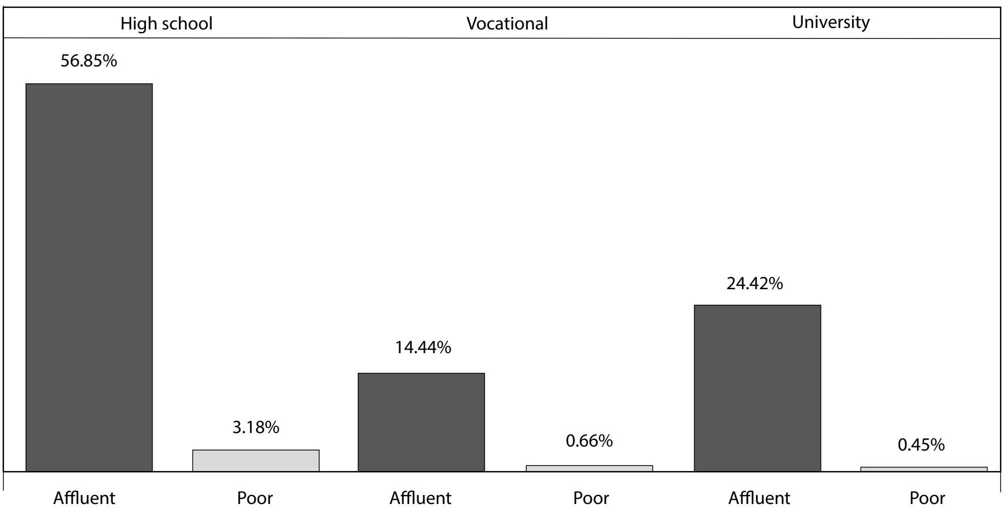
This study focuses to examine the impact of ICTs (e-commerce, laptop usage, handphone usage, internet access, social media, and digital banking) on poverty, particularly high-educated poverty, in East Java. In addition to ICTs as the independent variable, this study also involves variables that contain basic information on each respondent that can affect poverty, including age, marital status, gender, and education level. The internet access question in this study allows one to access the Internet anywhere and through personal or company-owned devices. This is fine because this variable provides information on whether a person uses the internet in their daily life. Furthermore, the ICT indicator in the form of laptop or cellphone ownership specifically requires personal ownership not belonging to the company where one works. This variable provides information that technology in the form of laptops and cellphones is needed by someone in their daily life to communicate not only at work. In order to do this, the study utilizes national survey data from the Susenas survey in 2020. The data used was obtained from a survey with community respondents in East Java with the provision of having received at least a high school education with the total of 30 719 respondents of the age of 18–60 years which is the working age in Indonesia.

East Java is one of the provinces in Indonesia with the second highest population density in Indonesia after West Java. This high population density is certainly vulnerable to economic problems, including poverty. East Java is also one of the provinces that has high economic growth potential in Indonesia (Sengaji et al., 2019). Therefore, understanding the factors that affect poverty in East Java and how to reduce it can help in designing more effective government policies and programs in reducing poverty in the province. In addition, a study of poverty in East Java can also provide greater insight into the problem of poverty in Indonesia in general, and can serve as a basis for comparing poverty rates in other provinces in Indonesia.

Poverty is a condition in which a person or family does not have the ability to satisfy minimum basic needs, such as food, clothing, housing, education, and health services. BPS measures poverty with reference to the Poverty Line or Minimum Poverty Line (MPL), which is the threshold value of income needed to fulfill these basic needs. The poverty data obtained is categorized into 2 categories, namely 0 for non-poor where a person is above the poverty line and 1 for poor where a person is below the poverty line.

The focus of this study is high-educated poverty in East java. High-educated poverty is a term that refers to the condition in which an individual or a group of people who have achieved higher education still live in poverty. High-educated poverty can occur due to several factors, such as difficulty in finding a job that matches their educational qualifications, low salaries and wages, lack of access to resources and capital, and lack of skills and experience in managing finances and businesses. High-educated poverty can have a significant impact on a person's life, such as difficulty in meeting basic needs, difficulty in obtaining adequate healthcare services, and limited access to economic and social opportunities. This can lead to a decline in the quality of life and well-being of an individual. To address high-educated poverty, efforts are needed from various stakeholders, such as the government, educational institutions, and the private sector. This research focuses on educated people in East Java who have studied high school, vocational school, and higher education with the distribution of data shown in Figure 1.

Figure 1 Information criteria and entropy of the various LCA models



Source: Own construction

Figure 1 provides information that most affluent people have a senior high school education. However, high school graduates also show the highest poverty rate compared to other graduates. When calculated at each level of education, the highest percentage of poverty occurs among senior high school graduates, which is 3.18%. Meanwhile, the poverty rate for vocational graduates is 0.66% and university graduates show the lowest poverty rate, at 0.45%. This documents that even though a person has taken a high level of education or in other words has been educated, there is still the possibility of poverty caused by other factors. The occurrence of poverty in educated communities can be caused by the use of digital technology that continues to develop in this day and age. This digital technology is the focus of this study to determine its effect on high-educated poverty in East Java.

This research also highlights the characteristics of high-educated poverty in East Java. The variables to be observed in this study are shown in Table 1.

Table 1 The variables

Variable	Category
Age	18–60 years
Gender	0: Male
	1: Female
Marital status	0: Not marriage
	1: Marriage
Internet access	0: No
	1: Yes
Mobile phone usage	0: No
	1: Yes
Social media usage	0: No
	1: Yes
Digital banking	0: No
	1: Yes
Laptop usage	0: No
	1: Yes
Education	0: University
	1: Vocational
	2: High school
Poverty status	0: Affluent
	1: Poor

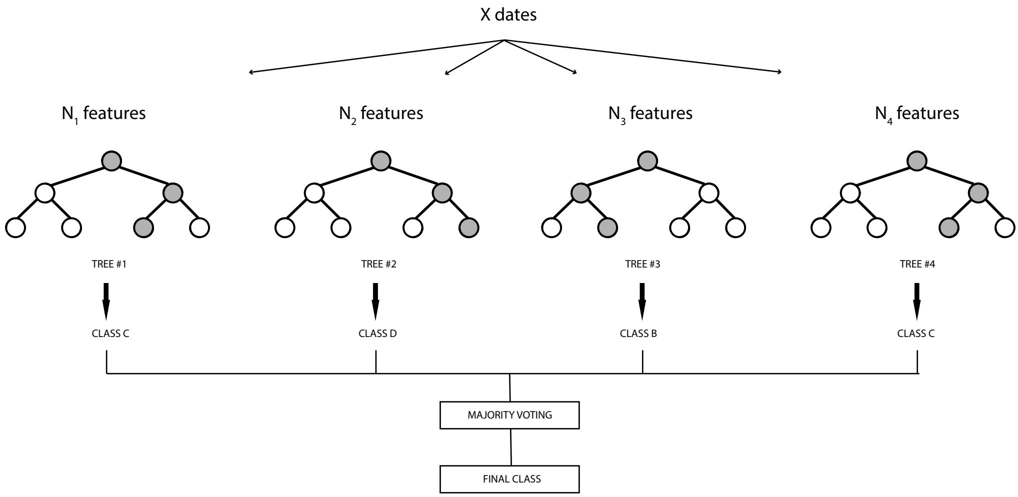
Source: Own construction

2.2 Random forest for machine learning agorithm

The machine learning technique used in this study is Random Forest (RF) analysis. RF is a type of machine learning algorithm that uses decision trees to make predictions. This method is particularly useful for this study because it can handle complex interactions between variables and handle large amounts of data. RF is one of the classification methods introduced by Breiman in 2001 as a development of the Classification and Regression Tree (CART) method. It is used to improve classification accuracy. The RF algorithm has demonstrated significant success as a classification and regression method for various general purposes. As a machine learning algorithm, it can handle large-scale problems involving big data and can be readily customized for various ad-hoc learning tasks (Mirzaei et al., 2021; Yangyudongnanxin, 2021).

The RF classification method consists of a combination of independent CART classification trees through a randomization process to form a tree on the sample and factor data. Therefore, this process will produce different classification trees. From a set of decision trees, it is expected to obtain a small correlation between trees to reduce prediction errors (Breiman, 2001). The underlying approach involves constructing multiple decision trees by randomly selecting training samples, and the output category is determined through the principle of majority voting, as demonstrated in Figure 2. In this context, the predictions for unknown samples are established based on the majority voting principle. In this study, random forest analysis will be used to examine the importance of technology to high-educated poverty modeling.

Figure 2 Poverty status on education



Source: Boateng et al. (2020)

The RF algorithm also provides a measure of feature importance. Feature importance is a measure that helps you understand which features (variables or attributes) are most influential in making predictions. Feature importance can provide insights into the relative significance of different input variables in a predictive model. It is important to note that feature importance does not establish causation, but it does help to identify which features have the most impact on the model's performance. Feature importance serves as a cornerstone for optimizing predictive models, aiding in the selection of critical features, and providing a deeper understanding of the interplay between variables and the target outcome. This information can be used to guide further research and policy interventions. This can be used to understand the importance of technology (social media, e-commerce, and digital banking) on poverty, as well as the role of education in alleviating poverty.

The use of the Susenas survey and RF analysis allows for a robust examination of the importance of technology on poverty in East Java. The survey data provides a rich and diverse data source, while the machine learning technique provides a powerful tool for analyzing complex relationships between variables. By combining these two methods, this study aims to provide a comprehensive and in-depth analysis of the impact of technology on poverty in East Java, with a focus on high-educated poverty.

The Susenas (national survey) data is collected and prepared for analysis. The data should be cleaned and checked for missing or incorrect values. The relevant variables that are believed to impact poverty in East Java, including education levels, use of technology (social media, e-commerce, and digital banking), and economic status, are selected. The data is then split into two parts, one for training and one for testing. The training data is used to build the RF model, while the testing data is used to validate the model. The random forest algorithm is applied to the training data. This involves creating a large number of decision trees, each of which is trained on a different subset of the data. The model then uses these decision trees to make predictions about the target variable (poverty status) based on the input variables.

The RF model is evaluated using the testing data. This involves comparing the model's predictions to the actual values and calculating various performance metrics, such as accuracy, precision, recall, and F1 score (Jiang et al., 2021; Matloob et al., 2021). For model evaluation and in Formulas (1), (2), (3), and (4), we expressed these evaluation metrics as follows:

$$Accuracy\left(Acc\right)=\frac{TP+TN}{TP+TN+FP+FN}\,, \tag{1}$$

$$Recall\left(Re\right)=\frac{TP}{TP+TN}\,, \tag{2}$$

$$Precision\left(Pre\right)=\frac{TP}{TP+FP}\,, \tag{3}$$

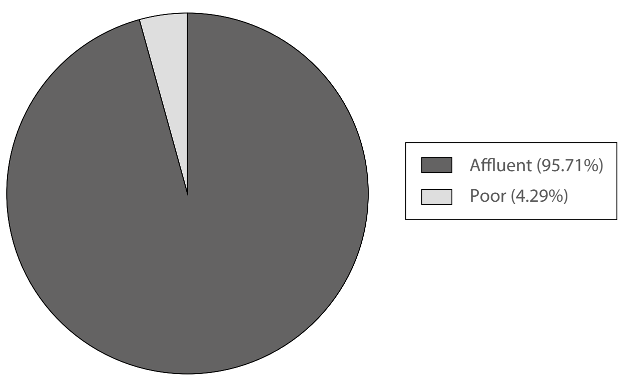
$$F1-Score=2\frac{\left(Pre\right)\left(Re\right)}{Pre+Re}\,. \tag{4}$$

In calculating accuracy for various types of models or performance evaluation methods, four terms are commonly used: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP occurs when the model or system correctly identifies a positive event as positive. TN occurs when a negative event is correctly identified as negative. On the other hand, FP happens when the model or system wrongly identifies a negative event as positive, while FN occurs when a positive event is wrongly identified as negative (Shah et al., 2020). In the context of classification, TP and TN refer to the number of true positive and true negative predictions, respectively, while FP and FN refer to the number of false positive and false negative predictions. Although accuracy can be calculated using TP, TN, FP, and FN, other performance metrics such as precision, recall, or F1-score may also be necessary to provide a more comprehensive evaluation of the model‘s performance If the performance of the model is not satisfactory, various parameters of the RF algorithm can be adjusted to improve its performance. This may involve changing the number of trees, the size of the subsets used for training, or the type of impurity measure used.

3 RESULTS AND DISCUSSION

The examination aimed to provide a preliminary understanding of the distribution of the data, which would aid in further analysis and interpretation of the results. The graphical representation provided valuable insights into the nature of the data, allowing for the identification of any potential issues or trends in the data. Through this examination, the study aimed to lay the foundation for a more in-depth and comprehensive analysis of the data in the future. The study aims to investigate the importance of ICTs on poverty in educated individuals in East Java. One of the key findings from the study is that the majority of the educated population in East Java is classified as affluent, with poverty rates at 4.29% as shown in Figure 3.

Figure 3 Percentage of poverty status in East Java



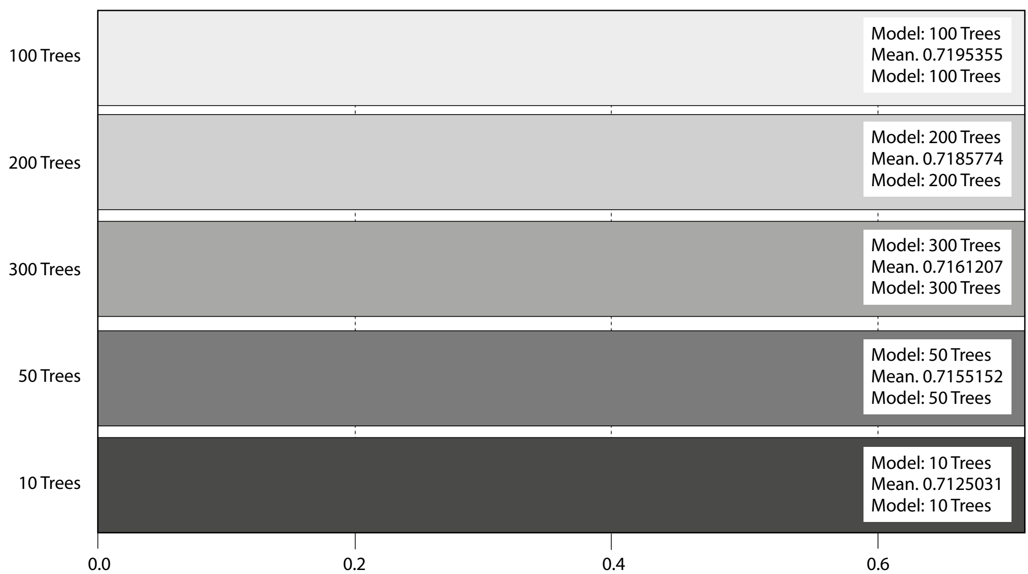
Source: Own construction

This result highlights the challenge of data imbalance in machine learning problems, as the majority class is more frequently represented in the data. Imbalance data can negatively impact the quality of the results obtained from machine learning models. The models tend to predict the majority class more often, as they have more examples to learn from. This can lead to biased results and poor performance in predicting minority or rare classes in the data. To address this issue, the study utilizes Synthetic Minority Over-Sampling Technique (SMOTE), a popular balanced data sampling technique. SMOTE creates synthetic data by modifying the unbalanced dataset and generating a balanced dataset from the unbalanced dataset. This helps to overcome the imbalance and increase the representation of the minority class in the data (Karthik and Krishnan, 2021).

The study highlights the importance of considering data imbalance and utilizing appropriate techniques such as SMOTE to overcome the challenge in machine learning problems. The results from the RF model provide a valuable insight into the importance of ICTs on poverty in educated individuals in East Java and demonstrate the effectiveness of the balanced data sampling technique in overcoming data imbalance. These findings can be used to inform future research and guide the development of more effective strategies to address poverty through the use of ICTs.

The RF model was optimized using a 10-fold cross-validation test mode with 1 000 iterations. As a result, the model is constructed using 10 trees on the grounds that a simple model can have the same accuracy as 300 trees (Figure 4). Based on the law of parsimony, we used the simplest model with 10 trees.

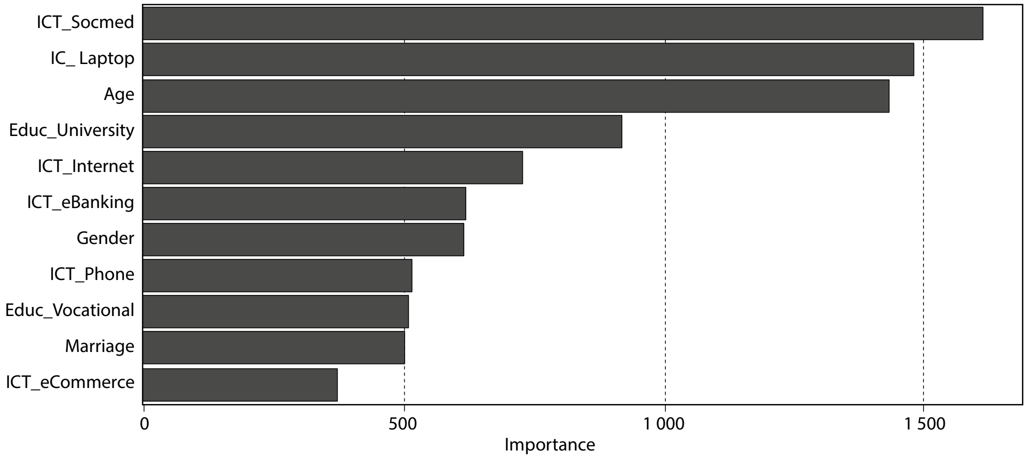
Figure 4 Selecting number of trees



Source: Own construction

Identifying the important variables in this experiment is crucial for determining if there are other factors that affect poverty classification aside from income. This is important in guiding us towards building multidimensional poverty indicators for poverty classification. Conversely, the RF algorithm is capable of generating a list of important variables by using the average merit calculation. Figure 5 presents the ranking of important variables from this experiment. It is apparent that out of the 11 variables, social media literacy, laptop ownership, age, higher education, and internet access are the five most important variables for high-educated poverty classification in East Java.

Figure 5 The variable importance



Source: Own construction

The accuracy of the RF model is evaluated using the testing data. This involves comparing the model's predictions to the actual values and calculating various performance metrics, such as accuracy, precision, recall, and F1 score are shown in Table 2.

Table 2 Evaluation model

Metric	Value
Accuracy	0.713
Precision	0.972
Recall	0.719
F1 score	0.826

Source: Own construction

Accuracy with a value of around 71.3% gives an overall picture that the model is able to correctly predict around 71.3% of individuals in the high-educated poverty group. However, a deeper focus on the precision metric indicates that the model is very efficient in identifying individuals who are actually in high-educated poverty. With a precision of about 97.2%, about 97.2% of the model's positive predictions are true positives, reducing the risk of misclassifying individuals who are not actually experiencing poverty as being in high-educated poverty. Similarly, the recall metric (Recall or Sensitivity) which has a value of around 71.9% highlights the model's ability to identify individuals who are actually in high-educated poverty. This reinforces the belief that the model is able to capture about 71.9% of all high-educated poverty cases in the data, creating a strong basis for more appropriate countermeasures. Furthermore, the F1 Score value of around 0.826 reflects a good balance between the model's ability to classify outcomes with a high degree of precision and its ability to identify actual cases of high-educated poverty. In this context, the F1 Score illustrates that the model is successful in providing overall good results in modeling high-educated poverty, considering both precision and recall.

Based on Figure 4, the results from the RF model suggest that using social media, having laptop, using internet and digital banking are among the top variables ICTs that have importance to classification high-educated poverty in East Java. The direct applications of ICTs, such as providing information about markets by Nwafor et al. (2020), opportunities, and employment, as well as improving skills and education, health care by Tortorella et al. (2022), and delivery of government services, are critical in addressing the welfare of the poor. These technologies also have the potential to empower the poor by improving communication between them and the government, giving them new avenues to voice their grievances. It should be noted that while access to ICTs is important, it is not the only a variable that matters in high educated poverty. Age, gender, education level, marital status, and mobile phone ownership also play a role in classification of poverty. Age, gender, education level, marital status, and posession of mobile phone also play a role in this model.

The presence of social media in an individual‘s life is also seen as a critical factor of high-educated poverty classification. Social media provides people with access to information and communication that can improve their economic and social well-being. For instance, individuals can use social media to connect with potential employers, access job opportunities, market their goods and services, and access education and training. Oh and Syn (2015) indicated the advantages of social media. Social media can also serve as a platform for building networks, seeking support, and exchanging ideas and information. All these benefits can contribute to an improvement in the economic status and overall well-being of individuals. However, excessive and uncontrolled use of social media can disrupt productivity and lead to unhealthy dependence (Achmad et al., 2022). Therefore, it is important for individuals to manage their social media usage and find the right balance between work and social life.

The importance of having a laptop in high-educated poverty is shown in the RF results. This is because someone who is in the poverty category will usually find it difficult to buy a laptop. Of course, this will hamper one‘s work skills and creativity considering that this is the digital era. Laptops provide access to information and opportunities that can improve economic and social well-being. This includes access to online markets, education and employment opportunities, as well as communication platforms and social media. With laptops, individuals can engage in digital commerce and grow their businesses, access financial services, and receive and transmit information. This enables them to improve their financial literacy and overall standard of living.

Age is also a critical factor in high-educated poverty classification. Generally, as individuals grow older and gain experience, they are able to acquire more skills, knowledge, and assets. These resources can help individuals to participate more effectively in the economy, thus improving their financial stability and reducing poverty. This is consistent with research by Baker et al. (2018) that the older individuals may have a greater sense of financial literacy, which can enable them to access financial services and make better financial decisions.

The use of the Internet is also the importance variable in the model of high-educated poverty in East Java. The Internet provides individuals with access to a wealth of information, as well as new and innovative ways of participating in the economy. For instance, individuals can use the Internet to access online markets, education, and job opportunities. The Internet can also be used to connect with others and build networks, which can be critical in securing support, resources, and information. All of these benefits can contribute to an improvement in the financial stability and overall well-being of individuals.

Having a mobile phone is also seen as an important factor in high-educated poverty modeling. Mobile phones provide individuals with access to information and communication that can improve their economic and social well-being. For instance, individuals can use mobile phones to access online markets, education, and job opportunities. Mobile phones can also be used for financial services, such as mobile banking, which can greatly enhance financial inclusion and improve financial stability. Furthermore, mobile phones can provide individuals with a sense of security and connect them with others, which can be critical in times of need and uncertainty.

CONCLUSION

The results of the RF model revealed that the top five most important variables in high-educated poverty in East Java are owning a laptop, having a social media account, age, using the internet, and having a mobile phone. The findings show that ICT use is the most important variable for classifying high-educated poverty in East Java. The ownership of a laptop and having a social media account, both ranking high in importance, signal the growing relevance of digital tools and online presence in today's interconnected world. These findings may hint at the potential opportunities and challenges faced by highly educated individuals, as these two variables could relate to job search, remote work, or online networking for economic advancement.

Technology, in the form of social media presence, laptop ownership and Internet access, is identified in this study as a factor that is important to modeling the high-educated poverty. However, we understand that poverty cannot be simplified into a direct result of not having technology. Poverty is a complex and often multidimensional phenomenon, influenced by various factors such as income, education, employment, and access to resources. Therefore, this study found that technology was one of the factors that contribute to overcoming poverty, but not the only cause. In-depth understanding needs to be done on how technology can act as a tool of reducing high-educated poverty, while still understanding the complexity of the poverty phenomenon and its impact on technology access.

In order to reduce the level of high-educated poverty in East Java, a number of concrete recommendations can be made to the government. First, it is necessary to launch an ICTs training program that focuses on educated people who still experience poverty. This includes training in computer use, applications, and internet access. In addition, initiatives should provide free or affordable internet access in areas where the majority of the population is out of reach thereof. This program should cover both rural and less developed urban areas in East Java.

Second, the government needs to allocate resources for digital infrastructure development in remote and isolated areas. This includes the development of strong internet networks and stable electricity. Collaboration with internet and electricity service providers, as well as Village-owned enterprises, can ensure reliable access to digital technology. Furthermore, digital-based economic empowerment programs should be a priority. This involves training in app creation, e-commerce, digital marketing and online business management. Financial support and mentoring for local digital entrepreneurs should be strengthened. Fourth, the government should pass strict regulations related to data security and user privacy in the use of digital technology. An independent oversight body should be established to deal with data and privacy breaches. Finally, regular monitoring and evaluation of these programs should be conducted. Collaboration with the private sector, financial institutions, and non-governmental organizations that have competencies in digital technology and poverty alleviation can support the implementation of these recommendations. With these measures, the government is expected to make a significant contribution in addressing high-educated poverty in East Java through better and effective utilization of the potential of digitalization.

There are several limitations that need to be acknowledged when interpreting the results of this study. Firstly, the study was conducted solely in East Java, Indonesia, which may limit its generalizability to other regions or countries. Secondly, the study was based on cross-sectional data, meaning that causality between the independent and dependent variables cannot be established. Finally, self-reported data, such as the information collected in this study, may be prone to measurement errors such as recall bias or social desirability bias. As a result, these limitations need to be considered when interpreting the findings of this study.

To further advance the understanding of the relationship between ICT and poverty among educated individuals in East Java, future research could be enhanced to address the limitations of this study. One potential approach is to include other socio-economic variables that may have an impact on poverty, such as household income and employment status, which could provide a more comprehensive understanding of the relationship between ICT and poverty. In addition, analyzing the impact of ICTs on poverty at the community level can reveal potential benefits and challenges in implementing ICT interventions in poor areas. Finally, conducting longitudinal studies can help to better understand changes in poverty levels over time, and the long-term impact of ICT interventions on poverty alleviation. By exploring these recommendations, future research can provide a more in-depth and comprehensive understanding of the role of ICT in alleviating poverty among educated individuals in East Java.

References

ABDUVALIEV, M., BUSTILLO, R. (2019). Impact of remittances on economic growth and poverty reduction amongst CIS countries [online]. *Post-Communist Economies*, 1–22. <https://doi.org/10.1080/14631377.2019.1678094>.

ACHMAD, W., SUDRAJAT, A., FAIZA, S., OLLIANTI, R. N. (2022). The Influence of Social Media on Teenagers' Lifestyles: Behavioral Analysis Among Adolescents in Bandung. *Journal on Education*, 5(3): 10356–10363.

ALSHARKAWI, A., AL-FETYANI, M., DAWAS, M., SAADEH, H., ALYAMAN, M. (2021). Poverty Classification Using Machine Learning: The Case of Jordan [online]. *Sustainability* (Switzerland), 13(3): 1–16. <https://doi.org/10.3390/su13031412>.

AO, Y., LI, H., ZHU, L., ALI, S., YANG, Z. (2019). The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling [online]. *Journal of Petroleum Science and Engineering*, 174: 776–789. <https://doi.org/10.1016/j.petrol.2018.11.067>.

BAKER, H. K., KUMAR, S., GOYAL, N., GAUR, V. (2018). How financial literacy and demographic variables relate to behavioral biases [online]. *Emmerald Insight*, 45(1): 124–146. <https://doi.org/10.1108/MF-01-2018-0003.

BHAVNANI, A., CHIU, R. W.-W., JANAKIRA, S., SILARZSKY, P. (2008). Ict Policy Division Global Information and Communications Department (Gict) the Role of Mobile Phones in Sustainable Rural Poverty Reduction. In: *Global Information and Communications Department*.

BOATENG, E. Y., OTOO, J., ABAYE, D. A. (2020). Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review [online]. *Journal of Data Analysis and Information Processing*, 8(4): 341–357. <https://doi.org/10.4236/jdaip.2020.84020>.

BREIMAN, L. (2001). Random Forests [online]. *Machine Learning*, 45: 5–32. <https://doi.org/10.1023/A:1010933404324>.

CHATTERJEE, A. (2020). Financial inclusion, information and communication technology diffusion, and economic growth: A panel data analysis [online]. *Information Technology for Development*, 26(3): 607–635. <https://doi.org/10.1080/02681102.2020.1734770>.

CHERNOVA, V. Y., ZOBOV, A. M., STAROSTIN, V. S. (2019). Dynamics of digital inequality in Russia [online]. *Media Watch*, 10(2): 225–234. <https://doi.org/10.15655/mw/2019/v10i2/49622>.

CHIB, A., MAY, J., BARRANTES, R. (2015). Impact of information society research in the global south [online]. In: *Impact of Information Society Research in the Global South*. <https://doi.org/10.1007/978-981-287-381-1>.

DIOP, S., ASONGU, S. A. (2021). The Covid-19 Pandemic and the New Poor in Africa: The Straw That Broke the Camel's Back [online]. In: *Forum for Social Economics*, Routledge. <https://doi.org/10.1080/07360932.2021.1884583>.

ERLANDO, A., RIYANTO, F. D., MASAKAZU, S. (2020). Financial inclusion, economic growth, and poverty alleviation: Evidence from eastern Indonesia [online]. *Heliyon*, 6(10): e05235. <https://doi.org/10.1016/j.heliyon.2020.e05235>.

FAHARUDDIN, F., ENDRAWATI, D. (2022). Determinants of working poverty in Indonesia [online]. *Journal of Economics and Development*, 24(3): 230–246. <https://doi.org/10.1108/jed-09-2021-0151>.

FINCH, W. H., FINCH, M. E. H. (2020). Poverty and Covid-19: Rates of Incidence and Deaths in the United States During the First 10 Weeks of the Pandemic [online]. *Frontiers in Sociology*, 5(June): 1–10. <https://doi.org/10.3389/fsoc.2020.00047>.

GUVEN, H. (2020). Industry 4.0 and Marketing 4.0 [online]. In: *Perspective of Digitalization and E-Commerce, Agile Business Leadership Methods for Industry 4.0*, 25–46. <https://doi.org/10.1108/978-1-80043-380-920201003>.

HOFMARCHER, T. (2021). The effect of education on poverty: A European perspective [online]. *Economics of Education Review*, 83(April): 102124. <https://doi.org/10.1016/j.econedurev.2021.102124>.

ISUKUL, A., TANTUA, B. (2021). Financial Inclusion in Developing Countries: Applying Financial Technology as a Panacea [online]. *South Asian Journal of Social Studies and Economics*, January: 42–60. <https://doi.org/10.9734/sajsse/2021/v9i230237>.

JACOBUS, E. H., KINDANGEN, P., WALEWANGKO, E. N. (2019). Analisis Faktor-Faktor Yang Mempengaruhi Kemiskinan Rumah Tangga Di Sulawesi Utara [online]. *Jurnal Pembangunan Ekonomi Dan Keuangan Daerah*, 19(7): 86–103. <https://doi.org/10.35794/jpekd.19900.19.7.2018>.

JIANG, T., LI, J. P., HAQ, A. U., SABOOR, A., ALI, A. (2021). A Novel Stacking Approach for Accurate Detection of Fake News [online]. *IEEE Access*, 9: 22626–22639. <https://doi.org/10.1109/ACCESS.2021.3056079>.

KARTHIK, M. G., KRISHNAN, M. B. M. (2021). Hybrid random forest and synthetic minority over sampling technique for detecting internet of things attacks [online]. *Journal of Ambient Intelligence and Humanized Computing*, 0123456789. <https://doi.org/10.1007/s12652-021-03082-3>.

KAUR, I., KAUR, J. (2020). Customer churn analysis and prediction in banking industry using machine learning [online]. *PDGC 2020–2020 6th International Conference on Parallel, Distributed and Grid Computing*, 434–437. <https://doi.org/10.1109/PDGC50313.2020.9315761>.

KENDALL, J., MYLENKO, N., PONCE, A. (2010). Measuring Financial Access around the World. *Policy Research Working Paper 5253 Abstract*, March: 1–21.

LOVARI, A., VALENTINI, C. (2020). Public Sector Communication and Social Media [online]. In: *The Handbook of Public Sector Communication*, 315–328. <https://doi.org/10.1002/9781119263203.ch21>.

MATLOOB, F., GHAZAL, T. M., TALEB, N., AFTAB, S., AHMAD, M., KHAN, M. A., ABBAS, S., SOOMRO, T. R. (2021). Software defect prediction using ensemble learning: A systematic literature review [online]. *IEEE Access*, 9: 98754–98771. <https://doi.org/10.1109/ACCESS.2021.3095559>.

MCQUIRE, S. (2021). Urban Digital Infrastructure, Smart Cityism, and Communication: Research Challenges for Urban E-Planning [online]. *International Journal of E-Planning Research*, 10(3): 1–10. <https://doi.org/10.4018/IJEPR.20210701.oa1>.

MEO, M. S., KUMAR, B., CHUGHTAI, S., KHAN, V. J., DOST, M. K. BIN, NISAR, Q. A. (2020). Impact of Unemployment and Governance on Poverty in Pakistan: A Fresh Insight from Non-linear ARDL Co-integration Approach [online]. *Global Business Review*. <https://doi.org/10.1177/0972150920920440>.

MIRZAEI, S., VAFAKHAH, M., PRADHAN, B., ALAVI, S. J. (2021). Flood susceptibility assessment using extreme gradient boosting (EGB), Iran [online]. *Earth Science Informatics*, 14(1): 51–67. <https://doi.org/10.1007/s12145-020-00530-0>.

MISHRA, V., SINGH BISHT, S. (2013). Mobile banking in a developing economy: A customer-centric model for policy formulation [online]. *Telecommunications Policy*, 37(6–7): 503–514. <https://doi.org/10.1016/j.telpol.2012.10.004>.

NILI, S. (2019). Global poverty, global sacrifices, and natural resource reforms [online]. *International Theory*, 11(1): 48–80. <https://doi.org/10.1017/S1752971918000209>.

NWAFOR, C. U., OGUNDEJI, A. A., VAN DER WESTHUIZEN, C. (2020). Adoption of ICT-based information sources and market participation among smallholder livestock farmers in South Africa [online]. *Agriculture* (Switzerland), 10(2). <https://doi.org/10.3390/agriculture10020044>.

OH, S., SYN, S. Y. (2015). Motivations for Sharing Information and Social Support in Social Media: A Comparative Analysis of Facebook, Twitter, Delicious, YouTube, and Flickr [online]. *Journal of the American Society for Information Science and Technology*, 66: 2045–2060. <https://doi.org/10.1002/asi>.

PRADHAN, R. P., ARVIN, M. B., NAIR, M. S., HALL, J. H., BENNETT, S. E. (2021). Sustainable economic development in India: The dynamics between financial inclusion, ICT development, and economic growth [online]. *Technological Forecasting and Social Change*, 169(May 2020): 120758. <https://doi.org/10.1016/j.techfore.2021.120758>.

SEBASTIANI, A. (2021). Digital artifacts and landscapes. Experimenting with placemaking at the impero project [online]. *Heritage*, 4(1): 281–303. <https://doi.org/10.3390/heritage4010018>.

SENGAJI, M., SASONGKO, S., SAKTI, R. (2019). Determinant of Regional Economic Growth in East Java, Indonesia [online]. *Jurnal Ekonomi Dan Studi Pembangunan*, 11(2): 104–115. <https://doi.org/10.17977/um002v11i22019p104>.

SHAH, K., PATEL, H., SANGHVI, D., SHAH, M. (2020). A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification [online]. *Augmented Human Research*, 5(1). <https://doi.org/10.1007/s41133-020-00032-0>.

SIDDIQA, A. (2021). Determinants of Unemployment in Selected Developing Countries: A Panel Data Analysis [online]. *Journal of Economic Impact*, 3(1): 19–26. <https://doi.org/10.52223/jei3012103>.

TORTORELLA, G. L., FOGLIATTO, F. S., SAURIN, T. A., TONETTO, L. M., MCFARLANE, D. (2022). Contributions of Healthcare 4.0 digital applications to the resilience of healthcare organizations during the Covid-19 outbreak [online]. *Technovation*, 111(September 2021): 102379. <https://doi.org/10.1016/j.technovation.2021.102379>.

VAKULENKO, Y., SHAMS, P., HELLSTRÖM, D., HJORT, K. (2019). Service innovation in e-commerce last mile delivery: Mapping the e-customer journey [online]. *Journal of Business Research*, 101(June 2018): 461–468. <https://doi.org/10.1016/j.jbusres.2019.01.016>.

WEWEGE, L., LEE, J., THOMSETT, M. C. (2020). Disruptions and Digital Banking Trends. *Journal of Applied Finance & Banking*, 10(6): 15–56.

YANGYUDONGNANXIN, G. (2021). Financial Credit Risk Control Strategy Based on Weighted Random Forest Algorithm [online]. *Scientific Programming*. <https://doi.org/10.1155/2021/6276155>.

YAO, Y., ZHOU, J., SUN, Z., GUAN, Q., GUO, Z., XU, Y., ZHANG, J., HONG, Y., CAI, Y., WANG, R. (2023). Estimating China's poverty reduction efficiency by integrating multi-source geospatial data and deep learning techniques [online]. *Geo-Spatial Information Science*, 1–17. <https://doi.org/10.1080/10095020.2023.2165975>.