Journal of Islamic Contemporary Accounting and Business
JICAB Vol. 3 No. 1, March 2025
E-ISSN: 3021-7105

Financing-Limit Prediction Classifier in Islamic Bank Using Tree-Based Algorithms

Mutya Qurratu'ayuni Mustafa[1], Muhammad Riza Iqbal Latief[2], Dewi Febriani[3]
[1,3,]Tazkia University, [2]CEP-CCIT Faculty of Engineering Universitas Indonesia
Email: 1902.mutya.027@student.tazkia.ac.id[1], riza@eng.ui.ac.id[2], dewifebriani@tazkia.ac.id[3]

A B S T R A C T

Islamic bank is one of financial institutions that has been proved to be the catalyst to end extreme poverty in the world. However, amid the massive development of industry 5.0, the research about technology adaptation in Islamic bank is still considered rare. The aim of this study is to develop a technology that will help Islamic bank in making their financing decision more efficient. By using the current outstanding financing data in an Islamic bank, this study proposes a machine learning algorithm that could predict a financing-limit based on customer classification. The tree-based learning algorithms used to build the algorithm has shown impressive results. The results shows that the basic algorithm which is Decision Tree gives 86% prediction accuracy. The algorithm is then improved by using Random Forest algorithm. The Random Forest algorithm gives 91% prediction accuracy which significantly improve the base learning algorithm. Future research in this area is needed as the need to implement sophisticated technology is prominent to make Islamic banking more accessible across the globe.

A R T I C L E I N F O

Article History:
Submitted/Received 16 Feb 2025
First Revised 26 Feb 2025
Accepted 1 March 2025
First Available online 12 March 2025
Publication Date 12 March 2025

Keyword:
Islamic Bank, Financing-Limit Prediction, Machine learning, Decision Tree, Random Forest

A B S T R A K

Bank syariah merupakan lembaga keuangan yang telah terbukti menjadi katalis untuk mengakhiri kemiskinan ekstrim di dunia. Namun, di tengah masifnya perkembangan industri 5.0, penelitian tentang adaptasi teknologi di bank syariah masih terbilang langka. Penelitian ini bertujuan untuk membantu bank syariah dalam membuat keputusan pembiayaan yang lebih efisien dengan teknologi. Dengan menggunakan data pembiayaan aktif di bank syariah, penelitian ini mengusulkan algoritma pembelajaran mesin yang dapat memprediksi batas pembiayaan berdasarkan klasifikasi pelanggan. Algoritma yang digunakan untuk pembelajaran adalah algoritma berbasis pohon. Hasil menunjukkan bahwa algoritma dasar yaitu Decision Tree memberikan akurasi prediksi sebesar 86%. Algoritma tersebut kemudian diperbaiki dengan menggunakan algoritma Random Forest yang memiliki akurasi prediksi 91%. Penelitian selanjutnya di bidang ini sangat diperlukan demi kemajuan industri perbankan syariah di dunia.

Kata Kunci:
Bank Syariah, Batas Pembiayaan, Pembelajaran Mesin, Pohon Keputusan.

INTRODUCTION

Financial institutions play a crucial role in Society 5.0 by providing capital to businesses across all sectors, impacting entrepreneurs, corporations, and households alike. Banks, in particular, contribute by financing business projects, household mortgages, and essential needs (Campa, 2022), while also supporting industries aligned with the Sustainable Development Goals (SDGs), making them key catalysts for achieving these goals. To address the evolving needs of their customers, banks have advanced their technology, as they face risks from all sectors of the economy. This requires continuous evaluation of their risk management processes and capital adequacy (Campa, 2022).

Islamic banks offer financing to both individuals and businesses, distinguishing themselves from traditional banks by using an Islamic financing system based on profit and loss sharing, rather than interest-based loans. This model has proven effective for global development and supports shared prosperity and poverty reduction (World Bank, 2015). Given that Islamic banks act as investors sharing both profits and losses, assessing financing applications requires careful evaluation. Machine learning (ML), a branch of Artificial Intelligence (AI), has been increasingly adopted by banks to improve risk assessments. ML has been successfully used to predict credit scores and customer creditworthiness (Ala, 2022; Aniceto et al., 2020; Dwiandriani & Mauritsius, 2021; Jutasompakorn et al., 2022; Khatir & Bee, 2022; Motwani et al., 2018; Orlova, 2021; Shih et al., 2022).

However, determining the appropriate amount of credit to extend is another challenge for credit risk teams. Business financing requires more thorough analysis due to the larger amounts involved compared to individual financing. Higher amounts come with higher risks, and the financing application process can be time-consuming. For instance, a financing officers (AO) in an Islamic bank noted that processing one business financing application typically takes one month (Hakim, 2023). To expand their market, banks need to shorten this analysis process to handle more applications. To the best of the author's knowledge, there has been no research focused on leveraging ML to address this challenge.

This study proposes the development of a system that helps Islamic banks assess business financing applications more efficiently. By analyzing existing financing data, the system will predict the maximum financing amount a business can potentially receive, enabling financing officers to make faster, more informed decisions. This study uses Decision Tree and Random Forest to develop an efficient predictive model for classifying customers based on their financing eligibility. This study contributes to the existing literature by offering a novel machine learning application in the Islamic banking context, focusing-limit prediction. The findings are expected to help Islamic banks improve their decision-making processes, reduce manual errors, and enhance overal operational efficiency.

LITERATURE REVIEW

Financing-Limit in Islamic Bank

A financing facility is the facility given by a bank to its potential customers to finance their businesses with a certain amount of finance within a year. The amount of financing-limit for a certain customer is the same as the amount of financing facility he gets from the bank. This means that a company can apply for multiple financing within a year to the bank, as many as it does not exceed their financing limit. This financing facility is applicable for customers who already have partnerships with the bank. The general process in applying financing-facility in Islamic bank can be seen as follows:

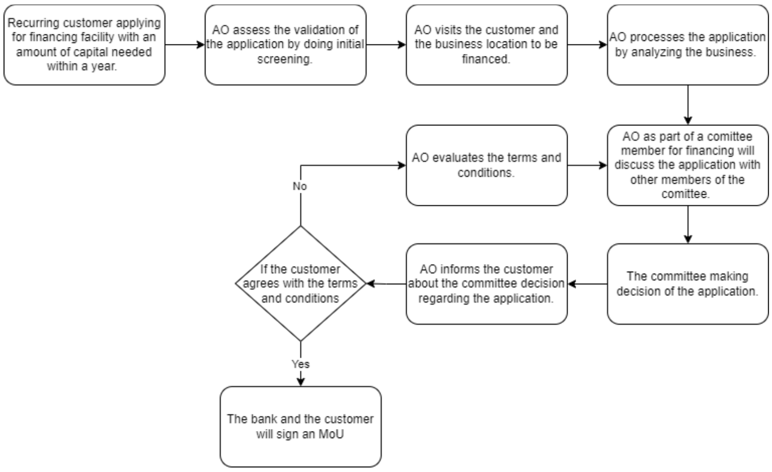

Figure 1: Financing application process in Islamic bank

Machine Learning

Machine learning is the process of a device to be able to perform certain tasks that have not been able to be performed before by humans or doing the human tasks in a better way because of changes produced by the learning process. The way machines can learn is through data. Machine learning studies are interested in the development of computer algorithms to transform data into intelligent action (Lantz, 2015). This makes machine learning an intersection between three fields of study which are domain knowledge that provides availability of data, statistical methods, and computing power.

Tree-Based Algorithms

Decision Tree

Decision tree is part of a non-parametric supervised learning algorithm that implements a tree structure to make predictions based on the relationships among the features given (Lantz, 2015). In decision trees, the decision is made from a wider scope and narrows it down until the final decision is made to predict a class value. The nodes represent features (attributes), branch represents decision (rule), and leaf nodes represent outcomes.

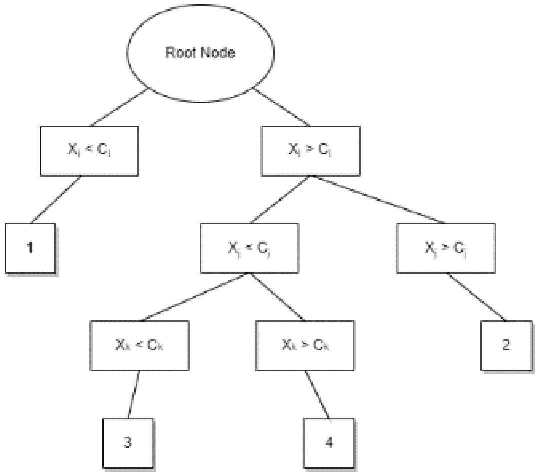

Figure 2: Decision Tree algorithm illustration

Random Forest

Random forest is part of supervised learning algorithm that can be used for both classification and regression problem (Mbaabu, 2020). This algorithm is built upon a decision tree as its base algorithm, and at the same time diminishing the limitations that exist in decision tree algorithm (Xu et al., 2021). Random forest minimizes overfitting in the model and increasing its precision compared to decision tree method. Random forest algorithm consists of many decision trees. The data is obtained by implementing a bagging method to decision tree as its base learner. The sample that obtained from bootstrapping process is called “Forest” as it depicts the existence of many “trees” which comes from the collection of decision trees. Next, the prediction will be based on the aggregate decision from all decision trees.

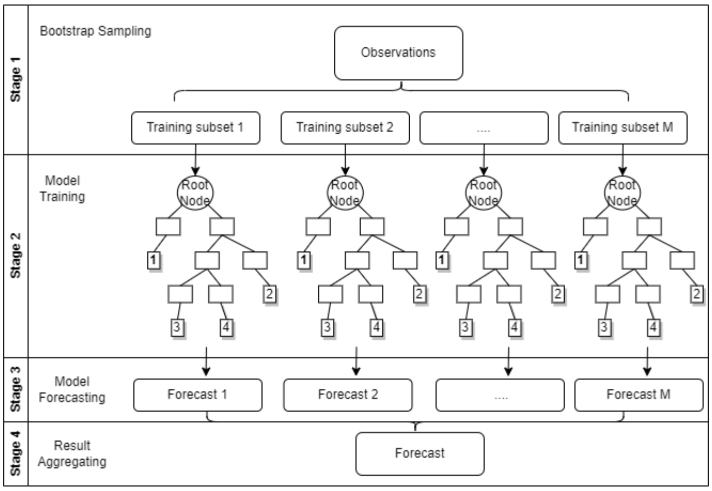

Figure 3: Random Forest algorithm illustration

Related Research in Financing-Limit and Machine Learning Applications in Finance

Many studies present literature contributions in many aspects of financing in financial institutions. (Agier & Szafarz, 2013; Cozarenco & Szafarz, 2020; Hodula et al., 2023) had observed the implication of loan ceiling and its impact on the lending behavior in society. It can be seen that although the loan ceiling has given a betterment in lending behavior as it can be seen in the research by (Hodula et al., 2023), it is still prompt to bias as can be seen in the research of (Agier & Szafarz, 2013; Cozarenco & Szafarz, 2020) especially in terms of gender equality and equality in prosocial lending. Other researchers had tried to find a way to solve problems in the financial industry using a technological approach. Some are intended to mitigate the risks that potentially exist in financing (Laopodis, 1999; Norden & Weber, 2010; Zhang et al., 2022). Subsequently, others are intended to create better financing terms for both customers and financial institutions (Kandil, 2009; Oyebode & Orji, 2020). Some studies are also discussing about the potential use of technology in Islamic finance for example in credit scoring (Muryanto et al., 2022; Muslimin & Mauritsius, 2022).

Loan ceiling or financing ceiling practice has existed in both conventional and Islamic financial institutions and the bias that currently exists, may exists in both worlds too. However, with the different scheme that Islamic finance have, to the best of author knowledge, there has not been any research that tackle the bias problem in Islamic financial institutions perspective. This study attempts to address the problem by incorporating technology to predict customer eligibility for financing amount based on their current financing conditions in Islamic bank using machine learning algorithms.

Journal of Islamic Contemporary Accounting and Business
JICAB Vol. 3 No. 1, March 2025
E-ISSN: 3021-7105

RESEARCH METHODOLOGY

This study implements a machine learning framework called CRISP- DM. CRISP-DM stands for Cross Industry Standard Process for Data Mining which is a process that designs the base methodology for a data science project (Hotz, 2023). According to KDnuggets polls, datascience-pm.com polls, and researched Google search volumes, CRISP-DM is the most popular framework used for datascience projects in 2020. This study analyzes the data with various scenarios. Throughout the process, different algorithms and parameters tuning might be applied to produce higher (acceptable) accuracy results. The experiment scenarios can be seen as follows:

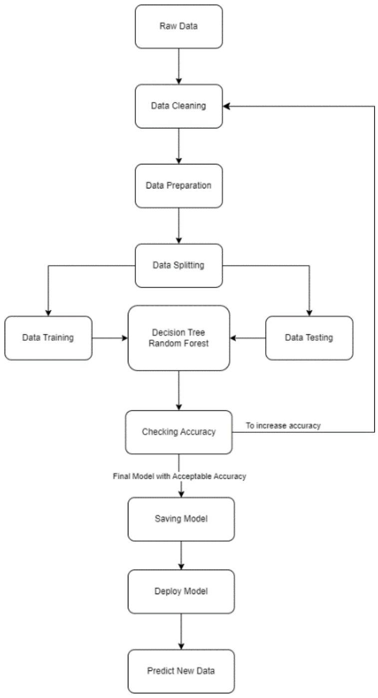

Figure 4: CRISP-DM framework illustration

Analysis Steps

1. **Business Understanding**

Financing is the most important part of Islamic banking industry (IB). It is the part where IBs are making profit from. Productive financing is way trickier compared to consumer financing because the bigger risks it imposes on the company. Most of the decision making in the application process is handled by humans, hence, it will take a very long time to process one productive financing application. While consumer productive application decision can take days, productive financing application decision could take months before it gets approved. This is due to the deeper analysis process of financing for business. The financing facility is the best scenario in which IBs can have certainty in having customers throughout the business year.

Because financing facilities are only given to current customers of IBs, there is a need for a mechanism to detect whether the current customers of IBs are worth giving the financing facility. In this study, the sample of the current outstanding productive-financing data in a syariah bank is analyzed to predict each customer financing-limit category.

2. **Data Understanding**

The dataset that was used in this study contains 42,583 data with 130 columns related to the current outstanding financing in an Islamic bank. 958 of them are the current outstanding productive financing while the rest consists of consumer financing and pawnshop. From the 130 columns provided, 18 columns are chosen based on their relations in the decision making to give financing facility. After the data has been imported, unnecessary and redundant variables are deleted. In this case, jenis_margin and uraian variables are deleted. The data shape after cleaning resulted in 516 rows and 19 columns for further preparation. The 19 columns that will be used are as follows:

Table 1: Variables

No	Variables	Explanation
1	Financing bank	The amount of outstanding productive financing
2	Tenure	The term for payment for each financing
3	Collateral value	The amount of collateral that goes with the financing
4	Economy sector	Economic sector category of a business
5	Type of business owner	The type of the busines owner
6	Type of funding usage	The type of funding usage
7	Type of business	The type of the business whict gets financed
8	Down payament	The frond-payament that were given as part of the join investment
9	The type of customer	The type of customer
10	The type of product	The type of financing product used
11	Segement category	The type of segment the business is related
12	Business code	The type of business based on code
13	Profit xbrl	The amount of the business current profit
14	Plafonds facilities	The amount of financing facility (if any) given to a business that have MoU with the bank
15	Ratio_rbh	The ratio of profit sharing that bank already received from customer
16	Equivalent_rate_contract	The margin rate for each financing

Based on an interview with a financing risk manager in an Islamic Bank in Indonesia, the most important features to notice in analysing productive financing application are business revenue, type of receivables, equivalent rate contract, and collateral values (Sudiaman, 2023). This study hypothesized that all the variables above, will determine the amount of financing-limit (plafond_fasilitas) given to a certain customer. However, the amount of financing-limit may vary for each customer because the needs of capital needed for every business may differ. To address

Journal of Islamic Contemporary Accounting and Business
JICAB Vol. 3 No. 1, March 2025
E-ISSN: 3021-7105

that problem, this study proposes a solution by creating categorical range for financing-limit. The categories that will be implemented is based on the interview with the Account Officer (AO) of productive financing. According to (Hakim, 2023), he divides customer against the financing facility as follows:

Table 2: Customer segmentation in productive financing

Types of business	Amount of financing facility given
Small-medium enterprise	500 millions -1 billions
Medium enterprise	1 billions - 5 billions
Big enterprise	5 billions - 10 billions
Financing syndicate of large enterprise	10 billions - 30 billions or more

3. Data Preparation

a. Selecting Data

There are many types of productive financing products offered by the Islamic bank. However, generally it is divided into two categories, contractual based financing, and non-contractual based financing. For this study, only non- contractual based productive financing is used for the analysis.

b. Cleaning Data

This process is conducted to filter any missing data, irrelevant inputs, and unnecessary data. In this process, the data contains inconsistency is deleted. Next, the data is also being filtered so that only productive financing using *musharaka* and *mudaraba* will be used in the analysis. Lastly, the data containing missing values is deleted. These steps are being taken to ensure the completeness and robustness of data.

c. Constructing Data

The data is also reconstructed because the variable eir_contract and equivalen_rate_contract contains the same information. For efficiency of data, the column eir_contract and equivalen_rate_contract are merged in the column equivalen_rate_contract. Next, the eir_contract column is deleted.

d. Transforming Data

In this phase, the dataset was reformed so that it is suitable for a classification analysis task. The transformation process was conducted. All categorical data are transformed into numerical-categorical data so that it is suitable for the machine to learn from it. For analysis process, each category of financing limit will be reformatted as follows:

Table 3: Customer segmentation in productive financing reformatted

Types of customer	Amount of financing facility given
Bronze customer	500 millions -1 billions
Silver customer	1 billions - 5 billions
Gold customer	5 billions - 10 billions
Platinum customer	10 billions - 30 billions or more

e. Modelling

The modelling process is the process where the dataset is trained with different kinds of algorithms. The goal of this process is to find the best algorithm that best predicts the customer financing-limit category. For each algorithm, there will be parameters passed to process the data. The algorithms experiment for this data can be seen as follows:

Table 4: Algorithms Experiments

Model	**Parameter**	**Value**
Decision tree clasifier	Max_depth	2
	Max_depth	best
Random forest classifier	default	default
	Max_depth	11
	n_estimator	221

RESULT AND DISCUSSION

Evaluation on Predicting Using Decision Tree Classifier

Evaluation Results Decision Tree Classifier with set parameters

a. Confusion Matrix

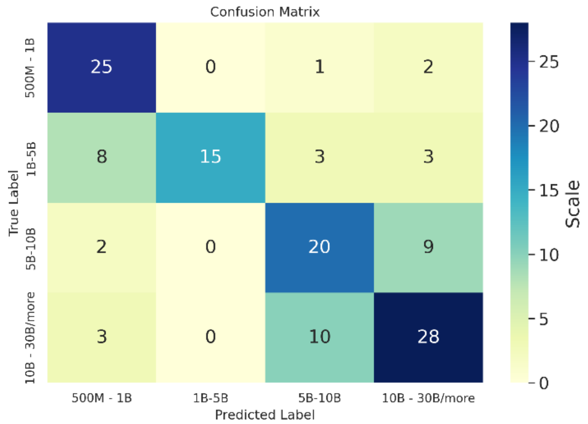

Figure 5: Confusion Matrix-Decision Tree 1

Based on the confusion matrix above, the model can predict 25 instances correctly from the total of 38 bronze customers which can be given financing from 500 million to 1 billion rupiah. It mislabeled 8 instances as silver customer, 2 instances as gold customer, and 8 instances for platinum customer. In short, the model is underperformed when it must classify bronze customers. The high number of mislabeled instances as platinum customers could potentially result in a big loss for the bank.

Secondly, the model can predict all instances correctly from silver customers which can be given financing from 1 billion to 5 billion rupiah. Thirdly, the model could predict 20 instances correctly from the total of 34 gold customers which can be given financing up to 10 billion rupiah. It mislabeled 1 instance as bronze customer, 3 instances as silver customer, and 10 instances as platinum customer. Based on the result, the model performed best in predicting silver customers as it classifies all the instances correctly. However, it performed poorly when classifying silver customers. The high number of misclassified silver customers as platinum customers confront the bank with higher risk of financing.

Journal of Islamic Contemporary Accounting and Business
JICAB Vol. 3 No. 1, March 2025
E-ISSN: 3021-7105

Lastly, the model successfully classifies 28 instances from the total of 42 platinum customers which can be given financing up to 30 billion rupiah or more in special cases. It mislabeled 2 instances as bronze customer, 3 instances as silver customer, and 9 instances for gold customer. The model performed well in classifying platinum customers. Further analysis of the confusion matrix will be explained in the classification report.

b. Classification Report

Classification Report

	precision	recall	f1-score	support
1	0.66	0.89	0.76	28
2	1.00	0.52	0.68	29
3	0.59	0.65	0.62	31
4	0.67	0.68	0.67	41
accuracy			0.68	129
macro avg	0.73	0.68	0.68	129
weighted avg	0.72	0.68	0.68	129

Figure 6: Classification Report - Decision Tree 1

Classification report is used to evaluate each model's performance. Based on the recall, for decision tree classifier algorithm, the model can predict 66% of bronze customers correctly. For silver customers, the model can predict 100% correctly. Meanwhile, the model can predict only 59% correctly for gold customers. Lastly, it can predict 67% of platinum customers correctly.

For precision, for all instances that the model predicts is bronze customer, the model got correct in 89% of it. for all instances that the model predicts is silver customer, the model got correct only 52% of it. for all instances that the model predicts is gold customer, the model got correct only 65% of it. And for all instances that the model predicts is platinum customer, the model got correct only 68% of it. For overall, the f-1 score is still far from 1. Hence, it can be said that decision tree classifier algorithm is not suitable to model the customer financing-limit prediction.

c. Features Importance

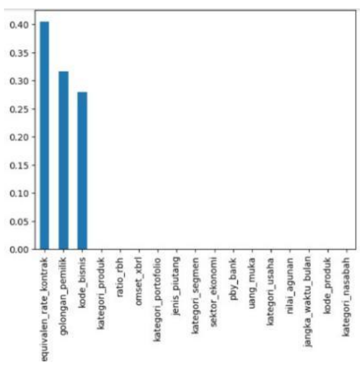

Figure 7: Features Importance using Decision Tree

Features importance is a visualization to see the reasoning behind tree algorithm. Based on the evaluation using feature importance, the decision tree classifier that is used in this study only takes three parameters to determine the category of their customers. The highest determinant for the tree is *equivalent_rate_kontrak* (equivalent_rate_contract) variable with the probability of

Journal of Islamic Contemporary Accounting and Business
JICAB Vol. 3 No. 1, March 2025
E-ISSN: 3021-7105

0.42, followed by *golongan_pemilik* (owner category) with the probability of 0.32 and *kode_bisnis* (business_code) with the probability of 0.26.

d. Decision Tree Visualization

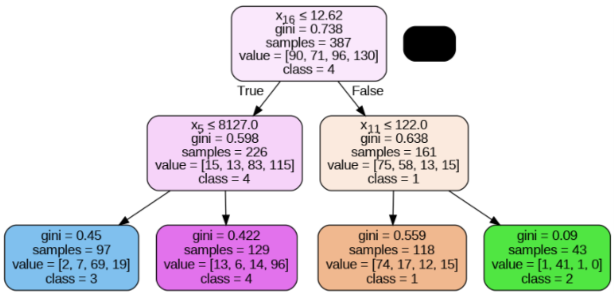

Figure 8: Decision Tree 1 Visualization

To see how the algorithm makes the decision based on the variables importance it chose, the tree diagram is plotted as follows. X16 shows how equivalent rate contract becomes the root node of the tree. If the equivalent rate contract is less than or equal to 12.62, the model will categorize the customer as platinum customer, otherwise, it would categorize it as bronze customer. Next, if the equivalent rate is less than 12.62 and its owner category is less than or equal 8127 then it will be categorized as gold customer, otherwise, it will be still categorized as platinum customer. On the other hand, if the equivalent rate contract is higher than 12.62 and the business category is less than 122, then it will categorize the customer as bronze customer. However, if the business category is higher than 122, then it will categorize the customer as a silver customer. Based on this observation, the model cannot optimize the variables and tends to generalize the data. Hence, it got the overall accuracy of 68% and this proves that it is not suitable for this study's use case.

Evaluation Results on Tuned Decision Tree Classifier (Max Depth = best)

a. Confusion Matrix

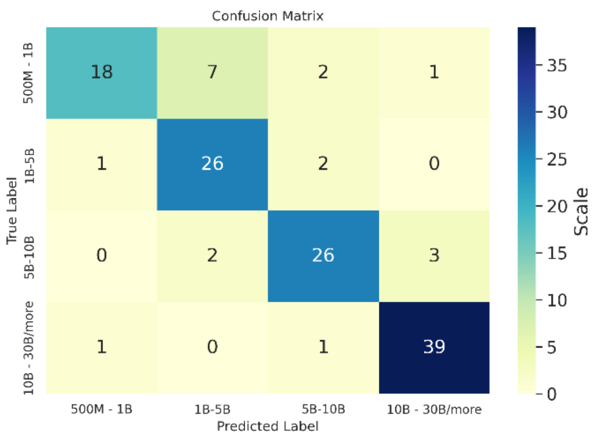

Figure 9: Confusion Matrix-Decision Tree 2

Based on the confusion matrix above, the model can predict 18 instances correctly from the total of 20 bronze customers which can be given financing from 500 million to 1 billion rupiah.

It mislabeled 1 instance as silver customer, and 1 instance for platinum customer. In short, the tune model resulted in better prediction of bronze customers.

Secondly, the model can predict 26 instances correctly from 35 silver customers which can be given financing from 1 billion to 5 billion rupiah. It mislabeled 7 instances as bronze customer, and 2 instances as gold customer. This shows that the model performs fairly well in predicting silver customers. Thirdly, the model could predict 26 instances correctly from the total of 31 gold customers which can be given financing up to 10 billion rupiah. It mislabeled 2 instances as bronze customer, 2 instances as silver customer, and 1 instance as platinum customer. Based on the result, the model could predict gold customers better than the previous model. It reduces the misclassified platinum customers from 10 to only 1 misclassified instance.

Lastly, the model successfully classifies 39 instances from the total of 43 platinum customers which can be given financing up to 30 billion rupiah or more in special cases. It mislabeled 1 instance as bronze customer, and 3 instances for gold customer. The model performed better in classifying platinum customers compared to the previous model. The further of analysis the confusion matrix will be explained in the classification report.

b. Classification Report

Classification Report

	precision	recall	f1-score	support
1	0.81	0.79	0.80	28
2	0.82	0.97	0.89	29
3	0.85	0.74	0.79	31
4	0.93	0.93	0.93	41
accuracy			0.86	129
macro avg	0.85	0.85	0.85	129
weighted avg	0.86	0.86	0.86	129

Figure 10: Classification Report-Decision Tree 2

Based on the classification report above, the accuracy has increased by 18%. Based on the precision, for bronze customers, the tuned model can predict 9% higher than before. For silver customers, the new model has lower recall than the base model by 30%. Meanwhile, the improved model can predict gold customers and platinum customers higher by 20% and 25% compared to the base model respectively. In short, the tuning could improve the precision of the model by roughly 10 to 30 percent for all classes.

For the recall, the new model also showed improvements in most of the classes. For bronze customers the recall has improved by 13%. However, there is a slight decrease in recall for silver customers by 0.3%. For gold and platinum customers the recall has increased by 15% and 26% respectively. This shows that the tuning could improve the recall by roughly 10 to 20 percent.
These improvements are manifested in the higher point of accuracy and the overall f1 score of 0.86. Hence, the new model is a good model for this study use case.

Journal of Islamic Contemporary Accounting and Business
JICAB Vol. 3 No. 1, March 2025
E-ISSN: 3021-7105

c. Features Importance

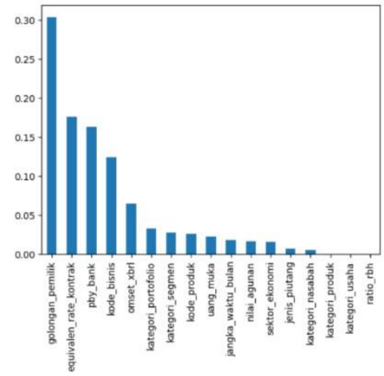

Figure 11: Features Importance on Decision Tree 2 Model

The tuned model shows a more robust understanding of the data. From the graph above, there are 14 variables considered in making the decision. The structure of the top three variables is also changed. The variable with higher probability of consideration is owner's category by 0.30. This is followed by the equivalent rate contract variable by roughly 0.17 and outstanding financing by roughly 0.16.

d. Decision Tree Visualization

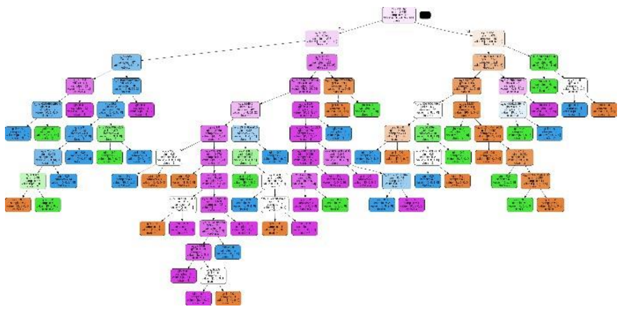

Figure 12: Decision Tree 2 Visualization

The visualization above is the manifestation of the features importance from the tuned model. The more machines learn from various variables, the deeper tree it will be to make decision. It can be concluded that while the maximum depth of the tree is low, the model tends to generalize the data and has a poor performance. On the other hand, the deeper a tree gets shows higher performance in classifying the data into its correct classes.

Evaluation of Prediction Using Random Forest Classifier

Evaluation Results on Random Forest Classifier with Default Parameters

a. Confusion Matrix

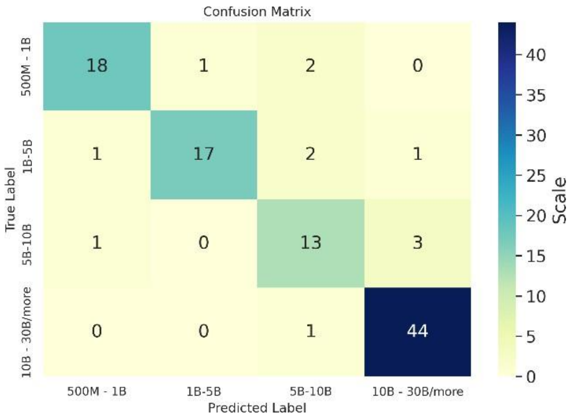

Figure 13: Confusion Matrix-Random Forest 1

Based on the confusion matrix above, the model can predict 18 instances correctly from the total of 20 bronze customers which can be given financing from 500 million to 1 billion rupiah. It mislabeled 1 instance as silver customer, and 1 instance for gold customer. This model is better at predicting bronze customers because it could reduce the misclassification as platinum customer to 0. Secondly, the model can predict 17 instances correctly from 18 silver customers which can be given financing from 1 billion to 5 billion rupiah. It mislabeled only 1 instance as bronze customer. This shows that the model performs very well in predicting silver customers.

Thirdly, the model could predict 13 instances correctly from the total of 18 gold customers which can be given financing up to 10 billion rupiah. It mislabeled 2 instances as bronze customer, 2 instances as silver customer, and 1 instance as platinum customer. This model could maintain the fair result of the tuned decision tree by keeping only 1 misclassified instance as platinum. Lastly, the model successfully classifies 44 instances from the total of 43 platinum customers which can be given financing up to 30 billion rupiah or more in special cases. It mislabeled 1 instance as silver customer, and 3 instances for gold customer. The model performed better in classifying platinum customers compared to the previous model. The further of analysis the confusion matrix will be explained in the classification report.

b. Classification Report

	precision	recall	f1-score	support
1	0.96	0.87	0.91	30
2	0.86	0.90	0.88	20
3	0.90	0.90	0.90	20
4	0.89	0.94	0.91	34
accuracy			0.90	104
macro avg	0.90	0.90	0.90	104
weighted avg	0.91	0.90	0.90	104

Figure 14: Classification Report - Random Forest 1

For random forest classification with default parameters, based on the recall, for regression tree classifier algorithm that trained with the best model, the model can predict 87% of bronze

Journal of Islamic Contemporary Accounting and Business
JICAB Vol. 3 No. 1, March 2025
E-ISSN: 3021-7105

customers correctly. For silver customers, the model can predict 90% correctly. Meanwhile, the model can predict 90% correctly of gold customers. Lastly, it can predict 94% of platinum customers correctly.

For precision, for all instances that the model predicts is bronze customer, the model got correct in 96% of it. for all instances that the model predicts is silver customer, the model got correct only 86% of it. for all instances that the model predicts is gold customer, the model got correct only 90% of it. And for all instances that the model predicts is platinum customer, the model got correct only 89% of it. For overall, random forest regression shows f1-score 0.90. Hence, it can be said that random forest classifier is suitable to predict customers financing limit category.

Evaluation Results on Tuned Random Forest Classifier

a. Confusion Matrix After Tuning

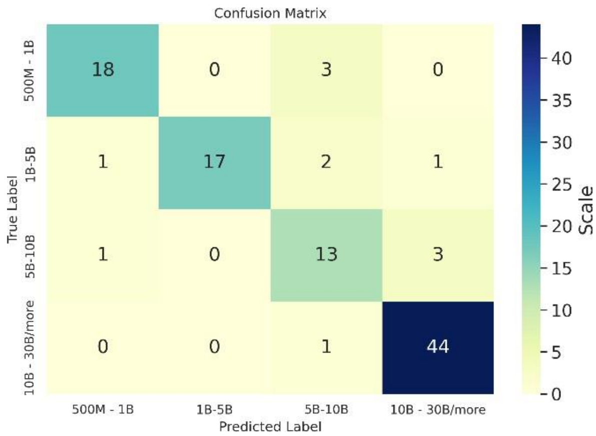

Figure 15: Confusion Matrix-Random Forest 2

Based on the confusion matrix for the tuned random forest, the ability of the model to predict bronze customers is the same as the base model. Secondly, the model can predict all instances correctly from silver customers. This is an improvement as the model could predict all instances correctly. Thirdly, the model could predict 13 instances correctly from the total of 19 gold customers which can be given financing up to 10 billion rupiah. It mislabeled 3 instances as bronze customer, 2 instances as silver customer, and 1 instance as platinum customer. This model could maintain the fair result of the tuned decision tree by keeping only 1 misclassified instance as platinum. Lastly, the tuned model maintains the performance of the base model in predicting platinum customers. The model performed better in classifying platinum customers compared to the previous model. Further analysis of the confusion matrix will be explained in the classification report.

b. Classification Report after tuning

	precision	recall	f1-score	support
1	0.96	0.87	0.91	30
2	0.86	0.90	0.88	20
3	0.90	0.95	0.93	20
4	0.91	0.94	0.93	34
accuracy			0.91	104
macro avg	0.91	0.91	0.91	104
weighted avg	0.92	0.91	0.91	104

Figure 16: Classification Report-Random Forest 2

From the tuned random forest, based on the recall, for regression tree classifier algorithm, the model can predict 87% of bronze customers correctly. For silver customers, the model can predict 90% correctly. Meanwhile, the model can predict 95% correctly of gold customers. Lastly, it can predict 94% of platinum customers correctly.

For precision, for all instances that the model predicts is bronze customer, the model got correct in 96% of it. for all instances that the model predicts is silver customer, the model got correct only 86% of it. for all instances that the model predicts is gold customer, the model got correct only 90% of it. And for all instances that the model predicts are platinum customers, the model got correct only 91% of it. For overall, random forest regression shows f1-score 0.91. After tuning the model, random forest classifier can predict the platinum customer better, and it shows higher accuracy by 1% compared to the random forest classifier model without tuning.

c. Features Importance

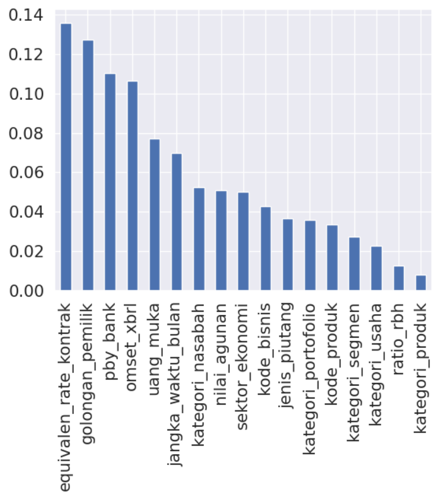

Figure 17: Features Importance-Random Forest 2

Based on the final model of random forest classifier, the variable importance of the model is extracted. From the graph above, it can be concluded that random forest classifier utilizes all the existing variables. By utilizing more variables, random forest classifier could create a robust model that results in higher accuracy. This study proves that by utilizing more variables, the accuracy of the model can be improved. According to the model, the three highest predictors for the model is equivalent rate contract with the probability of 0.13. Next is the owner's category with the probability roughly 0.12, and bank current financing with the probability of roughly 0.11.

Based on the four experiments of the trial, the summary of the model development in this study can be seen in the table below:

Table 5: Experiments Summary

Model	Parameter	Value	Accuracy	Precision	Recall	F-1 Score
Decision tree clasifier	Max_depth	2	68%	0.73	0.68	0.68
	Max_depth	best	86%	0.85	0.85	0.85
Random forest classifier	default	default	90%	0.90	0.90	0.90
	Max_depth	11	91%	0.91	0.91	0.91
	n_estimator	221				

CONCLUSION

This study has implemented various types of trees algorithm that best predicts customer financing-limit category. The base algorithm for this study case is a decision tree with the accuracy of 68% and performed well after tuning with 86% accuracy. The model is developed further by using random forest classification that resulted in 90% accuracy. To optimize the result, hyperparameter tuning is conducted using random search cross validation method. This optimization increases the accuracy to 91%. Hence, this study suggests that random forest classifier is the algorithm that is suitable for this use case.

This study also shows the impact of features importance to algorithms performance. The lack of identifying features of importance resulted in low accuracy as is shown in the decision tree model result. Meanwhile, the ability of random forest classifier to identify the features of importance in each variable resulted in a model with higher accuracy. In short, this study proves that the ability of a model to identify features importance determines the success rate of the model itself.

In a nutshell, customer financial-limit category classifier is beneficial for AO to detect business prospects in their current customers. Through the establishment of this technology, it is hoped to boost the efficiency of the marketing team in Islamic. This will make Islamic bank relevant to society 5.0 and becoming the catalyst to end extreme poverty around the world by promoting a robust and fair economics system. More implementations of artificial intelligence in various aspects of Islamic bank are encouraged to create more Islamic-compliance artificial intelligence in the future. Future research might use advanced algorithms such as support vector machine (SVM), deep learning, and K-nearest neighbor (KNN) to create a better model for this use case.

REFERENCES

Agier, I., & Szafarz, A. (2013). Microfinance and Gender: Is There a Glass Ceiling on Loan Size? *World Development*, *42*(1), 165–181. https://doi.org/10.1016/j.worlddev.2012.06.016

Ala, M. (2022). A deep learning model for behavioural credit scoring in banks. *Neural Computing and Applications*, *34*(8), 5839–5866. https://doi.org/10.1007/s00521-021-06695-z

Aniceto, M. C., Barboza, F., & Kimura, H. (2020). Machine learning predictivity applied to consumer creditworthiness. *Future Business Journal*, *6*(1), 1–14. https://doi.org/10.1186/s43093-020-00041-w

Campa, J. M. (2022). *The Role of Law in Transitioning to a Sustainable Economy*. https://doi.org/10.2139/ssrn.2816914

Cozarenco, A., & Szafarz, A. (2020). The regulation of prosocial lending: Are loan ceilings effective? *Journal of Banking and Finance*, *121*, 105979. https://doi.org/10.1016/j.jbankfin.2020.105979

Dwiandriani, F., & Mauritsius, T. (2021). Credit Scoring Using Machine Learning : A Case Study. *ICIC International*, *15*(12), 1329–1336. https://doi.org/10.24507/icicel.15.12.1329

Hakim, H. R. (2023). *Interview: How Financing Works in Islamic Bank?*

Hodula, M., Melecký, M., Pfeifer, L., & Szabo, M. (2023). Cooling the mortgage loan market: The effect of borrower-based limits on new mortgage lending. *Journal of International Money and Finance*, *132*. https://doi.org/10.1016/j.jimonfin.2023.102808

Hotz, N. (2023). *What is CRISP DM?* https://www.datascience-pm.com/crisp-dm-2/

Jutasompakorn, P., Perdana, A., & Balachandran, V. (2022). *Enhancing decision making with machine learning : The case of aurora crowdlending platform*. *0*(0), 1–9. https://doi.org/10.1177/20438869211060847

Kandil, M. (2009). On the relation between financial flows and the trade balance in developing countries. *Journal of International Trade and Economic Development*, *18*(3), 373–393. https://doi.org/10.1080/09638190902986520

Khatir, A. A. H. A., & Bee, M. (2022). Machine Learning Models and Data-Balancing Techniques for Credit Scoring: What Is the Best Combination? *Risks*, *10*(9). https://doi.org/10.3390/risks10090169

Lantz, B. (2015). *Machine Learning with R 2nd Edition* (2nd Editio). Packt Publishing.

Laopodis, N. T. (1999). Optimal prediction rule: An application to debt reschedulings. *Applied Economics*, *31*(1), 17–26. https://doi.org/10.1080/000368499324525

Mbaabu, O. (2020). *Introduction to Random Forest in Machine Learning*. https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/

Motwani, A., Chaurasiya, P., & Bajaj, G. (2018). Predicting Credit Worthiness of Bank Customer with Machine Learning over Cloud. *International Journal of Computer Sciences and Engineering*, *6*(7), 1471–1477. https://doi.org/10.26438/ijcse/v6i7.14711477

Muryanto, Y. T., Kharisma, D. B., & Ciptorukmi Nugraheni, A. S. (2022). Prospects and challenges of Islamic fintech in Indonesia: a legal viewpoint. *International Journal of Law and Management*, *64*(2), 239–252. https://doi.org/10.1108/IJLMA-07-2021-0162

Muslimin, F., & Mauritsius, T. (2022). Implementation of Machine Learning Technology for Consumer Credit Scoring in Banking Industry: Study Case of Pt Bank Bni Syariah. *Journal of Theoretical and Applied Information Technology*, *100*(7), 1628–1642.

Norden, L., & Weber, M. (2010). Credit line usage, checking account activity, and default risk of bank borrowers. *Review of Financial Studies*, *23*(10), 3665–3699. https://doi.org/10.1093/rfs/hhq061

Orlova, E. V. (2021). *Methodology and Models for Individuals ' Creditworthiness Management Using Digital Footprint Data and Machine Learning Methods*.

Oyebode, O., & Orji, R. (2020). A hybrid recommender system for product sales in a banking environment. *Journal of Banking and Financial Technology*, *4*(1), 15–25. https://doi.org/10.1007/s42786-019-00014-w

Shih, D., Wu, T., Shih, P., Lu, N., & Shih, M. (2022). A Framework of Global Credit-Scoring Modeling Using Outlier Detection and Machine Learning in a P2P Lending Platform. *Mathematics*, *10*(2282), 1–13. https://doi.org/https://doi.org/10.3390/ math10132282

Sudiaman, R. (2023). *Interview: Productive Financing Risk Assessment*.

World Bank. (2015). *Islamic Finance*. https://www.worldbank.org/en/topic/financialsector/brief/islamic-finance

Xu, J., Lu, Z., & Xie, Y. (2021). Loan default prediction of Chinese P2P market : a machine learning methodology. *Scientific Reports*, 1–19. https://doi.org/10.1038/s41598-021-98361-6

Zhang, W., Yan, S., Li, J., Tian, X., & Yoshida, T. (2022). Credit risk prediction of SMEs in supply chain finance by fusing demographic and behavioral data. *Transportation Research Part E: Logistics and Transportation Review*, *158*(December 2021), 102611. https://doi.org/10.1016/j.tre.2022.102611