International Journal of Educational Review, Law And Social Sciences (IJERLAS Journal) | ISSN (e): 2808-487X
Volumes 5 No. 4 (2025)

CREDIT MANAGEMENT ANALYSIS TO PREDICT EARLY NON-PERFORMING LOANS AT PT. BANK NEGARA INDONESIA (PERSERO) TBK

Hardiansyah[1]; Nazaruddin[2]; Muhammad Anggia Muchtar[3];
Magister Manajemen, Universitas Sumatera Utara, Medan[1,2,3]
Corresponding Email : hardiansyahardi90@gmail.com[1]; nazarmtd60@gmail.com[2]; anggi.muchtar@usu.ac.id[3];

Received : 21 May 2025
Revised : 30 May 2025
Accepted : 17 June 2025
Published : 27 July 2025
DOI : https://doi.org/10.54443/ijerlas.v5i4.3615
Link Publish : https://radjapublika.com/index.php/IJERLAS

Abstract
The purpose of this study is to improve credit quality in designing a customer loan default prediction system so that it can reduce the possibility of huge losses in banking in order to produce a model that has high accuracy and recall rate at PT. Bank Negara Indonesia (Persero) Tbk. The entire population in this study was sampled as many as 467 customers. Using the random forest method, the results obtained are: 1) descriptive analysis shows that the variables of customer credit history, payment ratio, age, collateral value, and affiliate balances influence the occurrence of bad debt, 2) the prediction model identifies patterns of customers at risk of bad debt, 3) the application of the credit model appropriately can reduce the level of bad debt.

Keywords: ***Bad Debt, Credit Management, Credit Risk.***

INTRODUCTION

Banking Financial Institutions play a role as credit distributors to the public in carrying out their business in the banking industry. Every bank must be fully aware of the various business risks it faces. Bank Negara Indonesia (BNI) is a conventional bank that provides banking services with savings products, loans, and other banking services. BNI has several savings and loan products (credit), one of which is BNI Griya, a type of loan product specifically for customers who need funds for Home Ownership Credit (KPR), building renovations, and for multi-purpose needs. A bank is said to have a high NPL if the number of non-performing loans exceeds the amount of loans granted to debtors. A high NPL will increase costs, both for providing reserves for productive assets and other expenses. In other words, the higher a bank's NPL, the more it will disrupt its performance. A high NPL indicates poor business operations at a bank, leading to various problems such as inability to pay third parties, reduced capital, and bad debts. This also leads to a decline in profits because, rather than setting aside reserves based on credit collectibility, the bank must accept a reduction in its revenue sources (Dwihandayani, 2017).In designing a loan default prediction system that can reduce the likelihood of bad debts with high accuracy and recall rates, PT Bank Negara Indonesia (Persero) Tbk contributes to improving the accuracy of credit analysis and strengthening its risk management strategy.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Credit Management

Firdaus (2009) stated that credit management is the management of credit carried out by banks including planning, organizing, implementing, and supervising in such a way that the credit runs well in accordance with the agreement between the bank and the debtor. Credit management is the main key for national banking to survive in tight competition, and will provide the expected income or profit. The steps taken by banking in Indonesia to achieve healthy credit conditions are: a) Credit Planning; b) Credit granting process and credit administration; c) Credit granting analysis; d) Types of interest rates; and e) Credit supervision.

Credit Risk

The factors causing bad debt are internal and external factors. Some internal factors are: a) Inaccurate analysis, so it is impossible to predict what will happen during the credit period. For example, credit is not given according to needs, so that customers are unable to pay installments that exceed their capabilities; b) Collusion between bank officials who handle credit and customers, so that the bank decides on credit that should not be given. For example, the bank over-transacts the value of collateral; c) Limited knowledge of bank officials regarding the type of debtor's business, so they cannot conduct analysis properly and accurately; d) Excessive interference from related parties, for example, commissioners, bank directors so that officers are not independent in deciding on credit; and e) Weaknesses in providing guidance and monitoring of debtor credit.

Non-Performing Loan (NPL)

Non-Performing Loan(NPL) is a credit that has matured installments but is not paid on time by the debtor so that there are arrears (Putri, 2016). There are three types of NPL credit (Ismail, 2013), namely a) Sub-standard including the category of problematic receivables with payments of <90-180 days. If the debtor-bank relationship deteriorates, creditors will not be able to trust the customer's financial history; b) Suspicious loans are loans whose interest or principal payments are delayed from 180-270 days on the terms that the debtor is required to do; and c) Bad credit is an incident where the bank experiences a loan loss because the debtor does not pay the loan for more than 270 days. So, non-performing loans are arrears from debtors on loans that have matured.

Credit Variable (Data Input)

The variables used in this study cover various aspects related to credit and customer profiles, which serve as early indicators in detecting potential bad debts: a) Kol Date (Collectibility Date); b) Open Account (Account Opening Date); c) Max Credit (Credit Ceiling); d) Debit Balance (Loan Balance); e) Installment Amount; f) Affiliate Balance; g) Workplace; and h) Address.

Random Forest Algorithm

This algorithm combines tree predictors or decision trees, where each tree relies on a random vector value sampled freely and evenly across all trees in the forest. Numerous studies have shown that Random Forests have good predictive performance in regression and clarification in various fields, such as finance, remote sensing, and genetic and biomedical analysis.

Data Mining

*Data mining*is a set of procedures used to extract additional value from information in a database that was previously unknown manually. This process utilizes extraction and mining of important patterns from existing data to obtain models or functions that describe and differentiate data concepts or labels (Ismanto & Novalia 2021).

Machine Learning

*Machine Learning*Machine learning is a modern scientific methodology that can perform automated procedures to generate predictions about a phenomenon by observing previous events, namely by searching for patterns in a given data set. Currently, machine learning has become a commonly used method to solve tasks or problems in everyday life that require the process of extracting large data sets. Broadly speaking, there are two types of machine learning: supervised learning and unsupervised learning.

RESULTS AND DISCUSSION

Data Understanding

This study used private data obtained by Bank BNI, which comprises customer data on credit at the bank. The collected dataset consists of 10 attributes and has been cleaned of missing values, as shown in the table below, representing the variables used in the study.

Table 1.1 Dataset

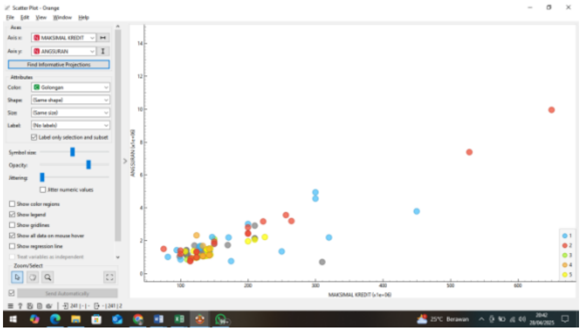
No	Name	Type	Role	Values
1	Group	Categorical	Target	1, 2, 3, 4, 5
2	ID Number	Numeric	Features	
3	Max Kredit	Numeric	Features	
4	Installment Value	Numeric	Features	
5	Debit Balance	Numeric	Features	
6	Affiliate Balance	Numeric	Features	
7	Income	Categorical	Features	Fixed Income, Non Fixed Income
8	The product	Categorical	Features	BNI GRIYA SUBSIDIARY,

No	Name	Type	Role	Values
				COMMERCIAL
9	Workplace	Text	Meta	
10	Address	Text	Meta	

Source: Research data processing (2025)

Scatter Plot Visualization

A type of two-dimensional graphic visualization to display points based on two or more variables, where each point on the scatter plot represents one observation or data.



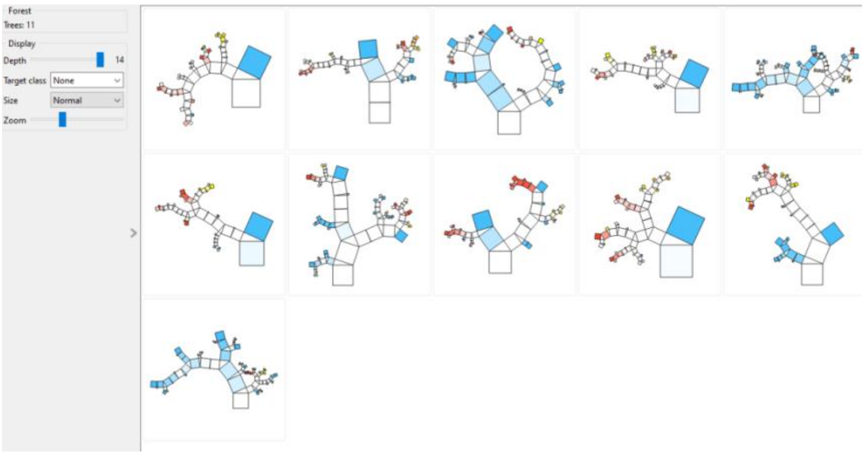
Source: Research data processing (2025)

Figure 1.1 Scatter Plot Visualization

The image above shows a correlation between credit smoothness and non-performing loans by category using the Orange data mining tool. The blue dots, representing borrowers with a smooth status (class 1), dominate areas with small loan ceilings and low total bills. Meanwhile, the yellow dots, representing borrowers with a substandard status (class 5), dominate areas with large loan ceilings and high total bills. This descriptive analysis concludes that credit smoothness can be influenced by loan ceilings and total bills. Furthermore, the scatterplot results also indicate that some data points are too far apart from the other data groups. Some of these points need to be removed for the machine learning algorithm to work optimally.

Pythagorean Random Forest

*Random Forest*is an ensemble learning algorithm consisting of many decision trees used for classification or regression. It is a very powerful and popular method in machine learning.



Source: Research data processing (2025

Figure 1.2 Pythagorean Random Forest

Figure 1.2 shows a visualization of a Pythagorean forest with 11 trees, illustrating the working mechanism of the random forest algorithm in determining predictions. The Pythagorean concept helps handle uncertainty or ambiguity in input data, resulting in more accurate predictions and greater tolerance to noise.

Evaluation

The predictive performance of the random forest algorithm can be measured using various metrics, including AUC (area under the curve), CA (accuracy), F1 Score, Precision, Recall, and MCC. The results of these various metrics in the table below show quite good predictive performance, approaching a value of 1.

Table 1.2 Random Forest Model Results

	AUC	CA	F1	Prec	Recall	MCC
Random Forest Model	0.963	0.865	0.853	0.845	0.865	0.672

Source: Research data processing (2025)

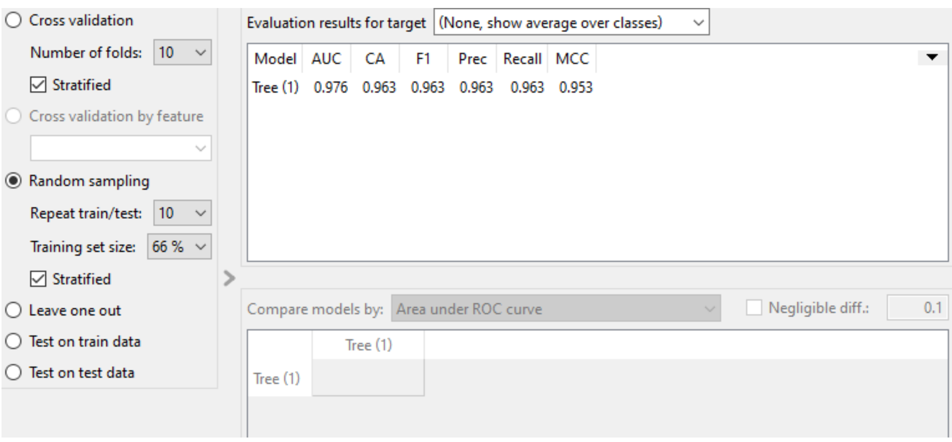
Table 1.2 illustrates the results of the Random Forest model performance evaluation for this study. This model achieved excellent performance with an AUC value of 0.963, indicating the model's ability to distinguish classifications with high precision. Accuracy (CA) of 0.865 reflects the model's overall accuracy. The F1 Score of 0.853 and Precision of 0.845 demonstrate a good balance between precision and recall for classifying positive cases. Recall of 0.865 demonstrates the model's ability to recognize a large proportion of positive cases. The Matthews Correlation Coefficient (MCC) of 0.672 indicates a strong correlation between model predictions and actual labels, confirming the reliability of the Random Forest model in this analysis. Therefore, it can be seen from the confusion matrix results table to see the prediction results of the random forest algorithm to show accurate results with a total of 467 data.

Table 1.3 Confusion Matrix Results

Actual	***Predicted***						**∑**
	Group	**1**	**2**	**3**	**4**	**5**	
	1	339	8	0	2	1	**350**
	2	6	43	0	3	2	**54**
	3	1	3	0	2	3	**9**
	4	2	6	0	8	6	**22**
	5	2	2	1	3	24	**32**
	∑	**350**	**62**	**1**	**18**	**36**	**467**

Source: Research data processing (2025)

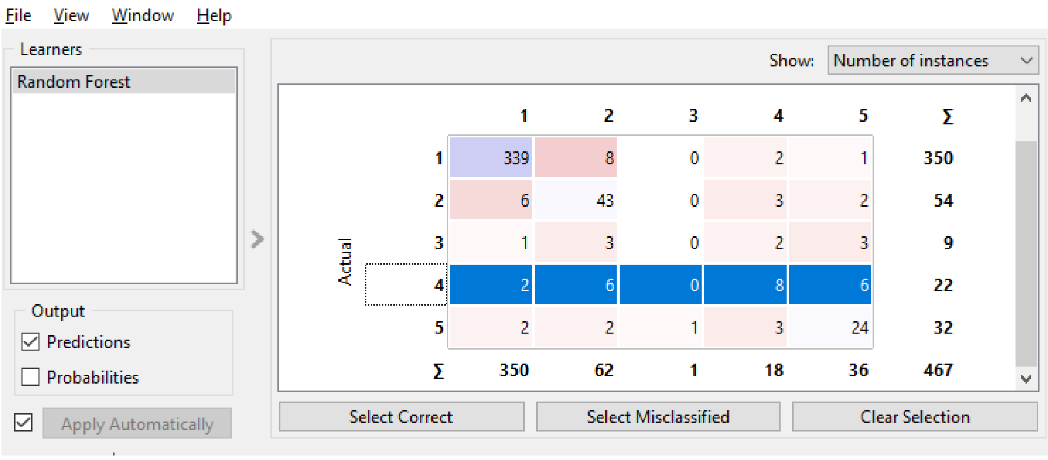
In the class 1 credit category, there were 338 correct predictions with a total data of 354. In the class 2 category there were 41 correct data out of a total prediction of 61 data. For the group 3 category, the correct prediction is in accordance with the total data. And for group 4 there are 4 correct out of a total of 13 data. In group 5 there are 21 correct data with a total of 39 data. The figure below illustrates the overall prediction process using the Random Forest algorithm. The research process begins with the initial step of data cleaning to handle missing

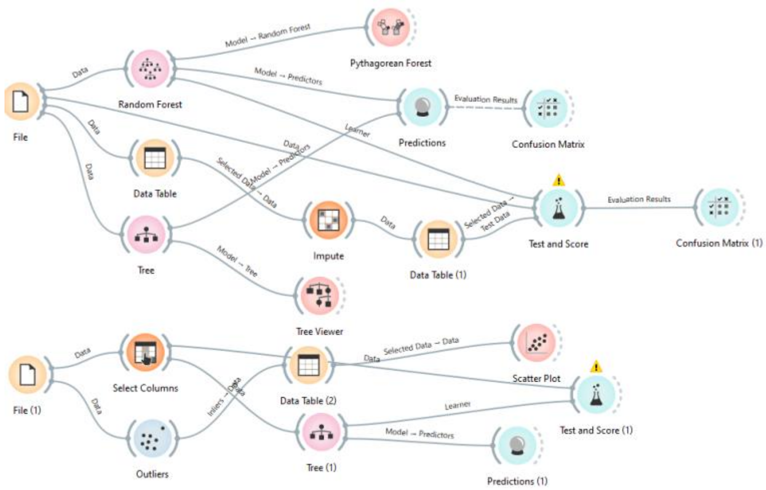


Source: Research data processing (2025)

values and reduce outliers that could potentially affect prediction accuracy. Next, the data is prepared for further analysis.

Figure 1.3 Random Forest Model Data Processing





Source: Research data processing (2025)

Figure 1.4 Confusion Matrix

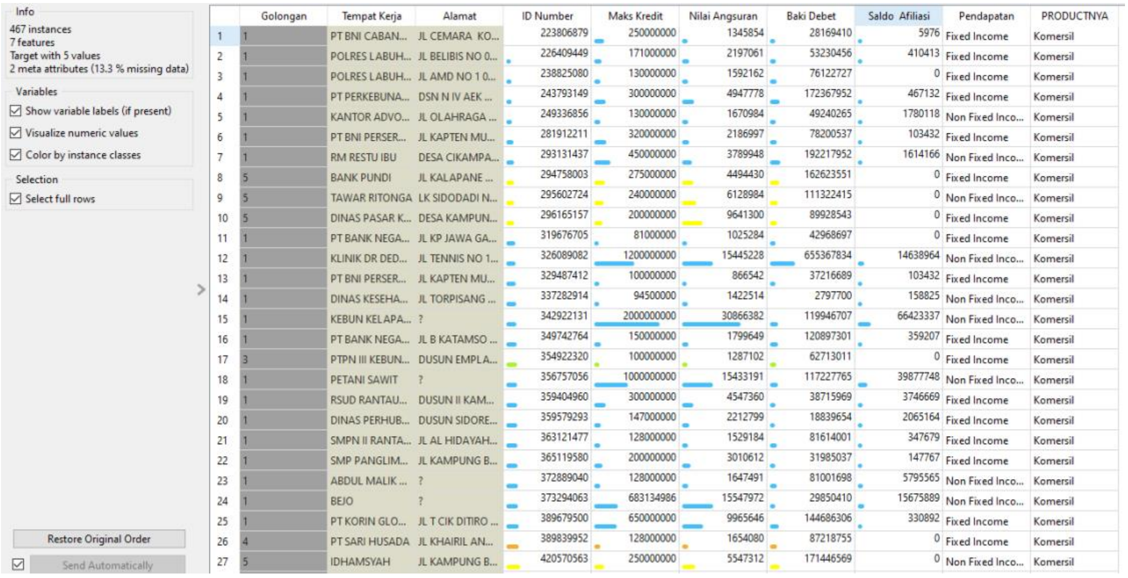
Source: Research data processing (2025)

Figure 1.5 Random Forest Algorithm Prediction Flow Using Orange

Next, the Random Forest algorithm is used in a predetermined configuration. This configuration includes the number of trees and the maximum depth. The resulting data from this model is trained and used to train the model to find important patterns in the data set. The test data will be used to evaluate the model's performance, generating reliable predictions to understand and anticipate relevant phenomena in credit data. Therefore, this research is expected to make a significant contribution to more proactive credit risk management, which relies on more accurate and efficient prediction of non-performing loans in the microfinance context. This is due to the number of customers with fixed income, which is 277 customers and 189 customers with non-fixed income. 377 customers have BNI Griya Subsidi Menpe products, and 90 customers have commercial products. Class, seen from the group, the number of groups consists of 1 to 5 groups, where group I numbers 350, for group II there are 54 customers, for group III there are 9 customers, for group IV there are 22 customers, and finally for group V there are 32 customers.

For the product, customers are divided into 2 categories of criteria, namely Commercial and BNI GRIYA SUBSIDI MENPE with a total of 87 customers for each product, and 381 customers for Griya Subsidi Menpe.

Source: Orange data mining data processing (2025)



Info
467 instances
7 features
Target with 5 values
2 meta attributes (13.3 % missing data)
Variables
Show variable labels (if present)
Visualize numeric values
Color by instance classes
Selection
Select full rows
Restore Original Order
Send Automatically

	Golongan	Tempat Kerja	Alamat	ID Number	Maks Kredit	Nilai Angsuran	Baki Debet	Saldo Afiliasi	Pendapatan	PRODUCTNYA
1	1	PT BNI CABAN...	JL CEMARA KO...	223806879	250000000	1345854	28169410	5976	Fixed Income	Komersil
2	1	POLRES LABUH...	JL BELIBIS NO 0...	226409449	171000000	2197061	53230456	410413	Fixed Income	Komersil
3	1	POLRES LABUH...	JL AMD NO 1 0...	238825080	130000000	1592162	76122727	0	Fixed Income	Komersil
4	1	PT PERKEBUNA...	DSN N IV AEK ...	243793149	300000000	4947778	172367952	467132	Fixed Income	Komersil
5	1	KANTOR ADVO...	JL OLAHRAGA ...	249336856	130000000	1670984	49240265	1780118	Non Fixed Inco...	Komersil
6	1	PT BNI PERSER...	JL KAPTEN MU...	281912211	320000000	2186997	78200537	103432	Fixed Income	Komersil
7	1	RM RESTU IBU	DESA CIKAMPA...	293131437	450000000	3789948	192217952	1614166	Non Fixed Inco...	Komersil
8	5	BANK PUNDI	JL KALAPANE ...	294758003	275000000	4494430	162623551	0	Fixed Income	Komersil
9	5	TAWAR RITONGA	LK SIDODADI N...	295602724	240000000	6128984	111322415	0	Non Fixed Inco...	Komersil
10	5	DINAS PASAR K...	DESA KAMPUN...	296165157	200000000	9641300	89928543	0	Fixed Income	Komersil
11	1	PT BANK NEGA...	JL KP JAWA GA...	319676705	81000000	1025284	42968697	0	Fixed Income	Komersil
12	1	KLINIK DR DED...	JL TENNIS NO 1...	326089082	1200000000	15445228	655367834	14638964	Non Fixed Inco...	Komersil
13	1	PT BNI PERSER...	JL KAPTEN MU...	329487412	100000000	866542	37216689	103432	Fixed Income	Komersil
14	1	DINAS KESEHA...	JL TORPISANG ...	337282914	94500000	1422514	2797700	158825	Non Fixed Inco...	Komersil
15	1	KEBUN KELAPA...	?	342922131	2000000000	30866382	119946707	66423337	Non Fixed Inco...	Komersil
16	1	PT BANK NEGA...	JL B KATAMSO ...	349742764	150000000	1799649	120897301	359207	Fixed Income	Komersil
17	3	PTPN III KEBUN...	DUSUN EMPLA...	354922320	100000000	1287102	62713011	0	Fixed Income	Komersil
18	1	PETANI SAWIT	?	356757056	1000000000	15433191	117227765	39877748	Non Fixed Inco...	Komersil
19	1	RSUD RANTAU...	DUSUN II KAM...	359404960	300000000	4547360	38715969	3746669	Fixed Income	Komersil
20	1	DINAS PERHUB...	DUSUN SIDORE...	359579293	147000000	2212799	18839654	2065164	Fixed Income	Komersil
21	1	SMPN II RANTA...	JL AL HIDAYAH...	363121477	128000000	1529184	81614001	347679	Fixed Income	Komersil
22	1	SMP PANGLIM...	JL KAMPUNG B...	365119580	200000000	3010612	31985037	147767	Fixed Income	Komersil
23	1	ABDUL MALIK ...	?	372889040	128000000	1647491	81001698	5795565	Non Fixed Inco...	Komersil
24	1	BEJO	?	373294063	683134986	15547972	29850410	15675889	Non Fixed Inco...	Komersil
25	1	PT KORIN GLO...	JL T CIK DITIRO ...	389679500	650000000	9965646	144686306	330892	Fixed Income	Komersil
26	4	PT SARI HUSADA	JL KHAIRIL AN...	389839952	128000000	1654080	87218755	0	Fixed Income	Komersil
27	5	IDHAMSYAH	JL KAMPUNG B...	420570563	250000000	5547312	171446569	0	Non Fixed Inco...	Komersil

Figure 1.6 Table Data

CONCLUSION

Based on the analysis of the research results that have been conducted, it can be concluded that this study uses a random forest algorithm to predict the high risk of bad debt in banks. This algorithm achieves a 96% accuracy rate on a 90:10 dataset. Descriptive analysis shows the relationship between various variables such as bank credit risk, loan interest rates, loan interest rates, and loan interest rates with the potential for high credit default. This can be used to develop credit risk mitigation strategies. This model predicts banks with high credit risk, where 96% of BNI Griya Rantau Prapat Branches fall into this risk category. This study recommends that banks focus on reducing credit risk by evaluating risk factors and implementing selective credit restructuring.

REFERENCES

Dwihandayani, D. (2017). Analisis Kinerja Non Performing Loan (Npl) Perbankan Di Indonesia Dan Faktor-Faktor Yang Mempengaruhi NPL. Jurnal Ilmiah Ekonomi Bisnis, 22(3), 228985.

Firdaus, R., & Ariyanti, M. (2009). *Manajemen perkreditan bank umum teori, masalah, kebijakan dan aplikasinya lengkap dengan analisis kredit*. ALFABETA.

Ismanto, E., & Novalia, M. (2021). Komparasi Kinerja Algoritma C4.5, Random Forest, dan Gradient Boosting untuk Klasifikasi Komoditas. *Techno.Com*, *20*(3), 400–410. https://doi.org/10.33633/tc.v20i3.4576

Pahlevi, O.-, Amrin, A.-, & Handrianto, Y.-. (2023). Implementasi Algoritma Klasifikasi Random Forest Untuk Penilaian Kelayakan Kredit. *Jurnal Infortech*, *5*(1), 71–76. https://doi.org/10.31294/infortech.v5i1.15829

Putri, Y. M. W., & Akmalia, A. (2016). Pengaruh CAR, NPL, ROA dan LDR Terhadap Penyaluran Kredit Pada Perbankan. Journal Balance, XIII(2), 82– 93, ISSN Print: 1693-9352, e-ISSN: 2614-820X.

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*, 102611. https://doi.org/10.1016/j.tre.2022.102611.