Classification of Customer Loans Using Hybrid Data Mining

Eka Praja Wiyata Mandala[1], Eva Rianti[2], Sarjon Defit[3]

[1,2,3]Department of Computer Science, Universitas Putra Indonesia YPTK Padang, Indonesia

[1]ekaprajawm@upiyptk.ac.id, [2]evarianti@upiyptk.ac.id, [3]sarjond@yahoo.co.uk

Abstract - At this time, loans are one of the products offered by banks to their customers. BPR is an abbreviation of Bank Perkreditan Rakyat. BPR is one of the banks that provide loans to their customers. The problem that occurs is that the number of loans given to customers is often not on target and does not meet the criteria. We propose a hybrid data mining method which consists of two phases, first, we will cluster the eligibility of customers to be given a loan using the k-means algorithm, second, we will classify the loan amount using data from the clustering of eligible customers using k-nearest neighbors. As a result of this study, we were able to cluster 25 customers into 2 clusters, 10 customers into the "Not Feasible" cluster, 15 customers into the "Feasible" cluster. Then we also succeeded in classifying customers who applied for new loans with occupation is Entrepreneur, salary is ≥ IDR 5000000, loan guarantees Proof of Vehicle Owner, account balance is < IDR 5000000 and family members is ≥ 4. And the results, classified as Loans with a small amount. We obtained the level of validity of the data testing of each input variable to the target variable reached 97.57%.

Keywords: clustering, classification, customer loans, hybrid data mining

I. INTRODUCTION

Loans are one of the main components of a bank. Banks try hard to offer loans to their customers because the main income of the bank comes from the profits obtained from the distribution of these loans. Bank will approve the loan submitted by the customer after verifying and validating a set of predetermined criteria [1]. Banks handle various types of loans that will be given to their customers, such as loans for personal purposes, loans for housing, loans for business purposes, and others. After the customer applies for a loan, the bank will validate the customer's eligibility to obtain a loan or not [2]. Loans are the main income priority for banks, but often loans will harm the bank if the customer is not smooth in paying, there will be bad loans [3].

Banks have important requirements in defending themselves so that they always exist and always become the trust of customers. The important requirement is to always check the background of the customer who will apply for a loan to reduce the risk that will harm the bank [4]. Many customers apply for loans to banks, but sometimes banks have limited funds and sometimes banks only provide loans to limited people so that customers have to seek loans from parties outside the bank [5].

In this paper, we find a problem in PT. Bank Perkreditan Rakyat (BPR) Bukit Cati Pematang Panjang in terms of providing loans to customers. BPRs accept loan applications from customers every month, but the funds available at BPRs are not enough for all loan applications, so customers who have strong connections can get loans. Another problem is that the amount of loans given sometimes does not match the criteria that have been set, so it is often not on target. Customers who meet the criteria get a small loan, while customers who do not meet the criteria can get a large loan.

To solve this problem, we propose a hybrid data mining method. Hybrid data mining is a combination of various selection features and classification algorithms to support the decision-making process [6]. In the hybrid data mining method, an important influencing factor is the process of generating various classifications to create a common model. There are many different ways to generate different and varied classifiers [7]. We propose a two-phase process, the first phase, clustering the eligibility of customers in applying for loans, then the second phase, eligible customers will be classified based on the loan amount.

In the first phase, we use a clustering approach using the k-means algorithm. Clustering is used to group customers into Feasible and Not Feasible clusters. In this phase, customers who do not meet the criteria can be screened so that of all customers who apply for loans, not all of them will be approved. K-means is one of the techniques used for unsupervised clustering. Labels are not owned by the data sample to classify the data into different classes. Data instances will be clustered according to the closest distance to the centroid. In the k-means algorithm, k indicates the number of clusters and each instance in the cluster has similarities [8]. K-means can optimize the clustering results gradually and redistribute the target dataset to each centroid constantly. K-means has advantages in simplicity, speed, and objectivity [9]. K-Means algorithm can be used to group data based on item data, for example the best-selling items, items that are selling well and items that are not selling well [10].

In the second phase, we use a classification approach using the k-nearest neighbor (KNN) algorithm. Classification is used to obtain the loan amount to be given to customers. Classification is a task in data mining to predict the future by testing the past with distance calculations [11]. Classification is a process to find a model or function that describes and distinguishes a class or concept from the data [12]. Classification is the process of assigning items to a predetermined set. The classification process is also referred to as pattern recognition [13]. KNN is the simplest classification algorithm [14]. KNN is often used as a method to solve complex classifier problems. KNN is included in the non-parametric learning algorithm. This means making no assumptions about the distribution of the data. Non-parametric algorithms such as KNN make decisions based on the entire training data set [15]. KNN performs classification with learning data in multidimensional space. This space is divided into sections that represent the learning data criteria [16].

By combining two techniques in data mining, clustering and classification, we hope to be able to determine the eligibility of customers to be given loans and the amount of loans that will be given to eligible customers. The aim of this study is to assist BPRs in making decisions about loan applications. Making the right decisions will help BPRs in approving loans that meet the criteria so that the loans provided can be right on target.

Previous study on classification for loan applications has been carried out using the rough set method. The aim of this study is to conduct an analysis to determine the effectiveness of the method based on rough set theory in binary classification of loan applications. This study adds a number of sub-scenarios, so that this study produces a total of 242 different classification results. These results indicate that there is a significant increase in the classification of rules [17]. Another study on the classification for the bank loan sector has also been conducted using KNN. The purpose of this study is to obtain a performance analysis of the KNN method used in loan applications. Performance analysis is obtained using metrics such as Jaccard index, F1-score and LogLoss. The results of this study obtained that the accuracy of KNN performance using Jaccard is 0.703704, using F1-score is 0.686067 and using Log Loss is N/A [18]. In another study, KNN was used to predict loan eligibility for cooperative customers. The purpose of this study is to predict the feasibility of lending to make the right decisions and to determine the results of the evaluation, accuracy, and validation of the k-NN algorithm in loan eligibility. The results of this study indicate that the accuracy of KNN before validation was 87.78% with AUC of 0.95%. After validation, the validation split decreased by 2% to 85.71% with an AUC of 0.836% [19].

All of the above studies use only one method, not using a combination of methods for classification in loan applications. Studies using a combination of methods have also been carried out, namely combining the Support Vector Machine (SVM) and K-Nearest Neighbor (KNN). The purpose of this study is to predict loan applications using SVM and KNN. The results of this study indicate that by using SVM and KNN, it is possible to predict the probability of customers getting a loan [20].

II. METHOD

We propose a combination of methods in data mining called hybrid data mining by combining clustering and classification techniques. In the first phase, we use clustering to cluster a customer's eligibility for a loan. For the clustering, we use the k-means algorithm. In the second phase, we use classification to determine the amount of loan that will be given to customers who are eligible for loans. For classification, we use the k-nearest neighbor algorithm. We propose a research methodology as shown in Fig. 1.

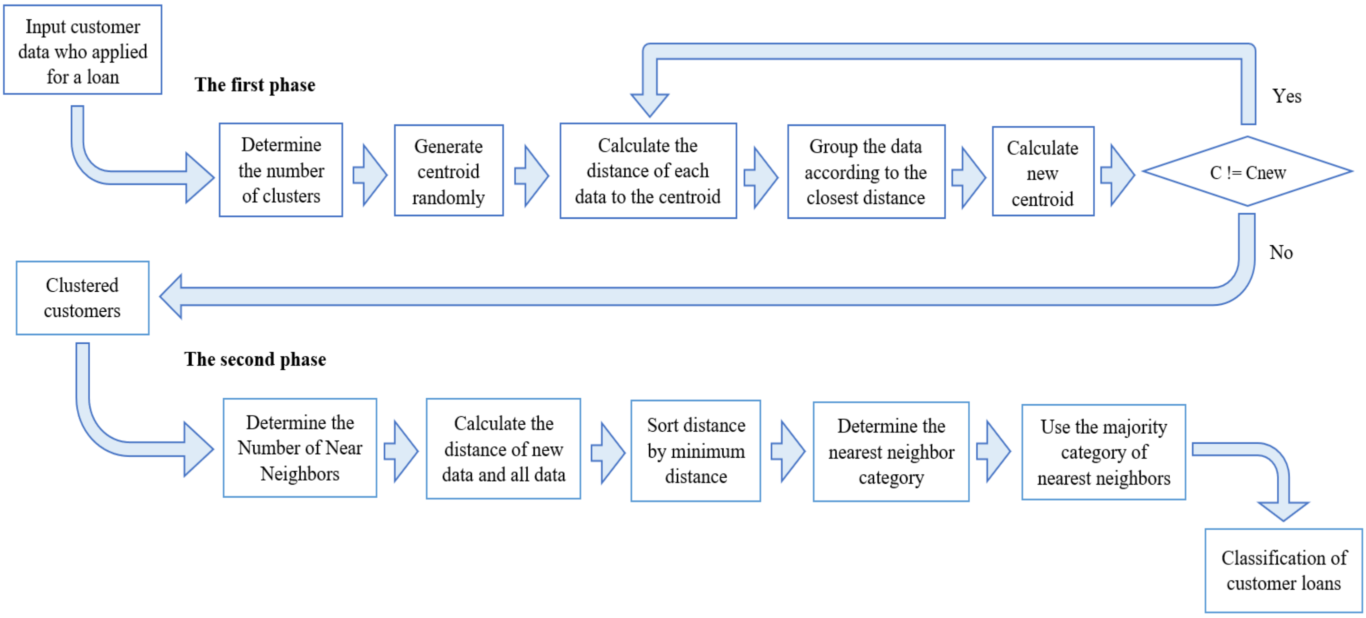

Fig. 1 Research steps

Fig. 1 illustrates the research methodology we used in this paper. This paper begins with data from customers who apply for loans as input data. Customer data is obtained from BPR. The data will be processed in two phases to obtain accurate classification results.

In the first phase, the process begins by determining the number of clusters to determine how many groups will be created, in this paper we group them into 2 clusters, "Feasible" and "Not Feasible". Then, we generate centroids randomly according to the number of clusters that have been created, so we generate 2 centroids. Then, we calculate the distance between each data to the centroid by using the correlation formula between two objects, Euclidean Distance (1).

$$D_{11} = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2} \qquad (1)$$

Next, we group the data based on the shortest distance from the data to each centroid. Then, we calculate the new centroid from the results of grouping the data that has been done previously. If the new centroid and the old centroid are not the same, then repeat the calculation of the distance of each data to each new centroid. If there is no change in cluster members, then the cluster data of customers who are eligible for loans can be obtained.

In the second phase, the data of customers who have been clustered who are entitled to receive loans are used as input data, meaning that customer data with the "Feasible" cluster is used as input. In this phase, the data classification process of customers who apply for new loans is carried out. The classification process starts from determining the number of nearest neighbors, where we determine the number of nearest neighbors is 3 because we have 2 decisions, "Loans with large amounts" and "Loans with a small amount". Then, we calculate the distance of the new data with all the data using the Euclidean Distance formula. Then, we sort the data based on the shortest distance between the new data and the input data, where the results will indicate which new data will be closer to the decision. Then, we determine the category of the nearest neighbor, and use the majority category of the nearest neighbor, so that the customer loan classification will be obtained based on the loan amount.

In this paper, we use customer data at PT. Bank Perkreditan Rakyat (BPR) Bukit Cati Pematang Panjang applied for a loan as training data. We obtained data for 25 customers who applied for loans. We will process this customer data to obtain clusters of customers who are eligible for loans and the classification of loan amounts to eligible customers. We use 5 predictor variables, including occupation, salary, loan guarantees, account balances and family members. The customer data we obtained can be seen in Table I.

TABLE I
CUSTOMER'S DATA AS A TRAINING DATA

Cust	Occupation	Salary	Loan Guarantees	Account Balance	Family Members
1	Entrepreneur	≥ IDR 2000000	Certificate	≥ IDR 5000000	< 4
2	Government Employees	≥ IDR 2000000	Proof of Vehicle Owner	≥ IDR 5000000	< 4
3	Freelancer	< IDR 2000000	Electronics	< IDR 5000000	< 4
4	Entrepreneur	≥ IDR 2000000	Electronics	< IDR 5000000	< 4
5	Government Employees	≥ IDR 2000000	Proof of Vehicle Owner	≥ IDR 10000000	< 4
6	Entrepreneur	< IDR 2000000	Electronics	≥ IDR 5000000	≥ 4
7	Entrepreneur	> IDR 2000000	Certificate	≥ IDR 10000000	≥ 4
8	Government Employees	≥ IDR 2000000	Position Decree	≥ IDR 5000000	< 4
9	Entrepreneur	≥ IDR 2000000	Certificate	≥ IDR 10000000	≥ 4
10	Government Employees	≥ IDR 2000000	Certificate	≥ IDR 10000000	< 4
11	Entrepreneur	≥ IDR 2000000	Certificate	≥ IDR 5000000	≥ 4
12	Government Employees	≥ IDR 5000000	Position Decree	≥ IDR 5000000	≥ 7
13	Entrepreneur	≥ IDR 5000000	Certificate	≥ IDR 5000000	≥ 4
14	Entrepreneur	≥ IDR 2000000	Certificate	≥ IDR 5000000	< 4
15	Government Employees	≥ IDR 2000000	Position Decree	≥ IDR 5000000	≥ 4
16	Entrepreneur	< IDR 2000000	Electronics	< IDR 5000000	≥ 4
17	Entrepreneur	≥ IDR 2000000	Certificate	< IDR 5000000	< 4
18	Freelancer	< IDR 2000000	Electronics	< IDR 5000000	≥ 4
19	Entrepreneur	≥ IDR 2000000	Certificate	≥ IDR 5000000	≥ 4
20	Entrepreneur	≥ IDR 2000000	Certificate	< IDR 5000000	≥ 4
21	Freelancer	< IDR 2000000	Electronics	< IDR 5000000	≥ 7
22	Entrepreneur	< IDR 2000000	Electronics	< IDR 5000000	≥ 7
23	Freelancer	< IDR 2000000	Proof of Vehicle Owner	< IDR 5000000	≥ 7
24	Entrepreneur	< IDR 2000000	Electronics	≥ IDR 5000000	≥ 7
25	Freelancer	< IDR 2000000	Proof of Vehicle Owner	≥ IDR 5000000	≥ 7

From Table I, we use 25 customers as training data. However, this data must be converted into numeric form so that it can be processed using the algorithm that we will use. To convert to numeric form, we assign weights to all attribute of the predictor variables. The weights are set with a range of 0-1. The weights we set can be seen in Table II.

III. RESULT AND DISCUSSION

We change the customer data in Table I to numeric by entering the weights that have been set in Table II. The process of converting text data into numeric data aims to facilitate the calculation of the distance between data and each centroid. Customer data that has been converted to numeric can be seen in Table III.

TABLE III
WEIGHT OF ALL ATTRIBUTE

Predictor Variables	Attribute of Predictor	Weight
Occupation	Government Employees	1,0
	Entrepreneur	0,7
	Freelancer	0.4
Salary	≥ IDR 5000000	1,0
	≥ IDR 2000000	0,7
	< IDR 2000000	0.4
Loan Guarantees	Position Decree	1,0
	Certificate	0,8
	Proof of Vehicle Owner	0,6
	Electronics	0,4
Account Balance	≥ IDR 10000000	1,0
	≥ IDR 5000000	0,7
	< IDR 5000000	0.4
Family Members	< 4	1,0
	≥ 4	0,7
	≥ 7	0.4
Decision of Loan Amount	Loans with large amounts	1,0
	Loans with a small amount	0,7

TABLE III
TRAINING DATA IN NUMERIC

Cust.	Occupation	Salary	Loan Guarantees	Account Balance	Family Members
1	0,7	0,7	0,8	0,7	1
2	1	0,7	0,6	0,7	1
3	0,4	0,4	0,4	0,4	1
4	0,7	0,7	0,4	0,4	1
5	1	0,7	0,6	1	1
6	0,7	0,4	0,4	0,7	0,7
7	0,7	0,7	0,8	1	0,7
8	1	0,7	1	0,7	1
9	0,7	0,7	0,8	1	0,7
10	1	0,7	0,8	1	1
11	0,7	0,7	0,8	0,7	0,7
12	1	1	1	0,7	0,4
13	0,7	1	0,8	0,7	0,7
14	0,7	0,7	0,8	0,7	1
15	1	0,7	1	0,7	0,7
16	0,7	0,4	0,4	0,4	0,7
17	0,7	0,7	0,8	0,4	1
18	0,4	0,4	0,4	0,4	0,7
19	0,7	0,7	0,8	0,7	0,7
20	0,7	0,7	0,8	0,4	0,7
21	0,4	0,4	0,4	0,4	0,4
22	0,7	0,4	0,4	0,4	0,4
23	0,4	0,4	0,6	0,4	0,4
24	0,7	0,4	0,4	0,7	0,4
25	0,4	0,4	0,6	0,7	0,4

In the first phase, we will process the numerical data above using the k-means algorithm to obtain clusters of customers who are eligible for loans. The k-means algorithm begins by determining the number of clusters. We set 2 clusters to group the above 25 customers. The clusters we defined are "Feasible" and "Not Feasible". Then randomly generate centroids and calculate the distance of each data to each centroid using the Euclidean Distance formula. Then group each data based on the shortest distance to the centroid to get the members of each cluster. Then calculate the new centroid, if the new centroid is the same as the old centroid, then the process is complete, but if the new centroid is not the same as the old centroid, then the distance of each data must be recalculated to the new centroid, then grouped again, until there are no more member changes from each cluster.

We carried out the testing process using Weka version 3.8.3. After we do the calculations and testing, we have to go through 8 iterations to get the members of each cluster. The results of the clustering of tests using Weka, from 25 training data 10 customers (40% of data) became members of "cluster0" and 15 customers (60% of data) became members of "cluster1". Where cluster0 refers to the "Not Feasible" cluster and cluster1 refers to the "Feasible" cluster. The cluster results can be visualized in the form of a scatter plot as shown in Fig. 2

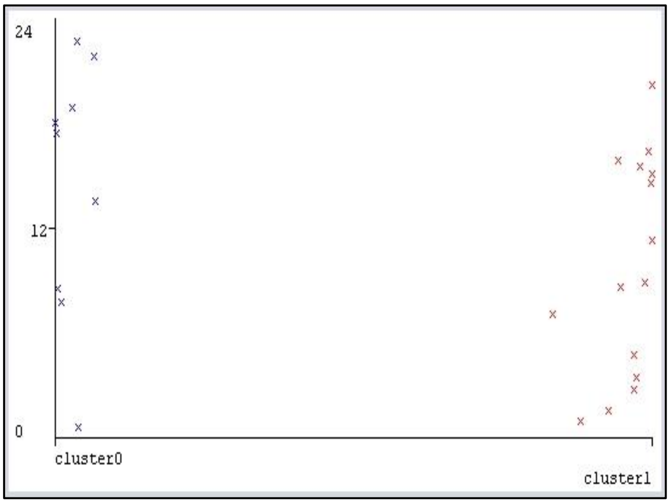

Fig. 2 Scatter plot of clustering results

In the second phase, the data from the "Feasible" cluster will be used as training data to determine the loan amount to be given to new customers as test data. In this phase, we use the k-nearest neighbor (KNN) algorithm. KNN begins by determining the number of nearest neighbors, in this case, we determine the number of nearest neighbors is 3, because the loan amount that we will classify is 2, "Loans with large amounts" and "Loans with a small amount". Then we will calculate the distance between the test data and all the training data. Then we will sort the distance by the closest distance.

Then we will determine the nearest neighbor category and we will use the majority category of the nearest neighbor. The training data that will be used in KNN can be seen in Table IV.

From Table IV, we can see that there are 2 classifications, "Loans with large amounts" given to 7 customers and "Loans with a small amount" given to 8 customers. In KNN, the classification process requires new customer data as test data to determine the amount of loan that will be given to the new customer. New customer data as test data can be seen in Table V.

Based on the test data in Table V, we tested KNN using Weka 3.8.3. We enter training data along with test data to get the results of the classification of the number of loans given to customers. From the test results, we find that the 26th customer is classified as "Loans with a small amount" with a prediction margin of -0.882 as shown in Fig. 3

To measure the size of the validity of using construct validity. The criteria for conducting the test are to calculate the product moment correlation (r) and then testing with the t-test. The calculation results are obtained by comparing the count value with the value in the table as in (2) [21].

$$rxy = \frac{\sum xy}{\sqrt{(\sum x^2)}\sqrt{(\sum y^2)}} \tag{2}$$

$$rxy = \frac{2.38}{2.439262184}$$

$$= 0.975704873 \ x\ 100\%$$

$$= 97{,}57\ \%$$

From the results of the measurement of the validity of the input variable and the target variable, the validity level reached 97.57%. This shows that the level of validity of the data from each variable has a high validity. The attributes of all input variables match the attributes contained in the target variable in the test data.

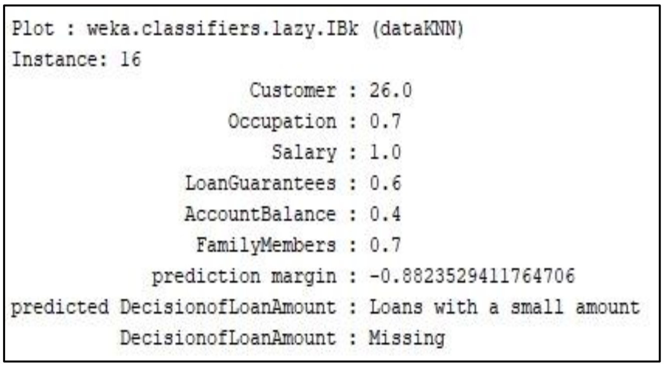

Fig. 3 Classification result from testing using Weka

TABLE IV
TRAINING DATA FROM “FEASIBLE” CLUSTER

Cust	Occupation	Salary	Loan Guarantees	Account Balance	Family Members	Decision of Loan Amount
1	0,7	0,7	0,8	0,7	1	Loans with large amounts
2	1	0,7	0,6	0,7	1	Loans with large amounts
5	1	0,7	0,6	1	1	Loans with large amounts
7	0,7	0,7	0,8	1	0,7	Loans with large amounts
8	1	0,7	1	0,7	1	Loans with a small amount
9	0,7	0,7	0,8	1	0,7	Loans with large amounts
10	1	0,7	0,8	1	1	Loans with large amounts
11	0,7	0,7	0,8	0,7	0,7	Loans with a small amount
12	1	1	1	0,7	0,4	Loans with a small amount
13	0,7	1	0,8	0,7	0,7	Loans with a small amount
14	0,7	0,7	0,8	0,7	1	Loans with large amounts
15	1	0,7	1	0,7	0,7	Loans with a small amount
17	0,7	0,7	0,8	0,4	1	Loans with a small amount
19	0,7	0,7	0,8	0,7	0,7	Loans with a small amount
20	0,7	0,7	0,8	0,4	0,7	Loans with a small amount

TABLE V
TESTING DATA

Cust	Occupation	Salary	Loan Guarantees	Account Balance	Family Members
26	0,7	1	0,6	0,4	0,7

IV. CONCLUSION

In this paper, we have discussed the problems that occur in BPR regarding customer loans. We use a hybrid data mining approach to help BPRs overcome these problems. Our process consists of two phases. In the first phase, we use the k-means algorithm to cluster the eligibility of customers who will be given loans. Of the 25 customers, we obtained 15 customers who are eligible for loans. In the second phase, we use the k-nearest neighbor algorithm to classify the loan amount. From 15 customers as training data, we provide 1 test data for testing. With the following customer data used as test data, occupation is Entrepreneur, salary is ≥ IDR 5000000, loan guarantees is Proof of Vehicle Owner, account balance is < IDR 5000000 and family members is ≥ 4, it can be concluded that the customer is classified as Loans with a small amount. The use of the hybrid method between k-means and k-nearest neighbors has a significant impact in determining which customers are eligible to receive loans and the amount of loans granted to these customers. From the results of validity testing, we obtained the level of validity of the data testing of each input variable to the target variable reached 97.57%. The limitation of this paper is that our research is only limited to classification, in the future we will try to predict the risk of loans submitted by customers.

REFERENCES

[1] A. Dagar, "A Comparative Study on Loan Eligibility," *Int. J. Sci. Res. Eng. Trends*, vol. 7, no. 3, pp. 1646–1649, 2021.

[2] M. A. Sheikh, A. K. Goel, and T. Kumar, "An Approach for Prediction of Loan Approval using Machine Learning Algorithm," *Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020*, no. Icesc, pp. 490–494, 2020.

[3] J. Sanjaya, E. Renata, V. E. Budiman, F. Anderson, and M. Ayub, "Prediksi Kelalaian Pinjaman Bank Menggunakan Random Forest dan Adaptive Boosting," *J. Tek. Inform. dan Sist. Inf.*, vol. 6, no. 1, pp. 50–60, 2020.

[4] K. Gupta, B. Chakrabarti, A. A. Ansari, S. S. Rautaray, and M. Pandey, "Loanification - Loan Approval Classification using Machine Learning Algorithms," *SSRN Electron. J.*, pp. 1–4, 2021.

[5] A. Kulothungan, Neha, Himanshu, and K. Gupta, "Loan Forecast by Using Machine Learning," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 7, pp. 894–900, 2021.

[6] J. Nalić, G. Martinović, and D. Žagar, "New hybrid data mining model for credit scoring based on feature selection algorithm and ensemble classifiers," *Adv. Eng. Informatics*, vol. 45, no. March, p. 101130, 2020.

[7] M. Abedini, F. Ahmadzadeh, and R. Noorossana, "Customer credit scoring using a hybrid data mining approach," *Kybernetes*, vol. 45, no. 10, pp. 1576–1588, 2016.

[8] M. S. Rao, C. Sekhar, and D. Bhattacharyya, "Comparative Analysis of Machine Learning Models on Loan Risk Analysis," in *Machine Intelligence and Soft Computing*, 2020, pp. 81–90.

[9] Z. Zhu and N. Liu, "Early Warning of Financial Risk Based on K-Means Clustering Algorithm," *Complexity*, vol. 2021, 2021.

[10] F. Indriyani and E. Irfiani, "Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means," *JUITA J. Inform.*, vol. 7, no. November, pp. 109–113, 2019.

[11] J. Sinaga and B. Sinaga, "Data Mining Classification Of Filing Credit Customers Without Collateral With K-Nearest Neighbor Algorithm (Case study: PT. BPR Diori Double)," *J. Comput. Networks, Archit. High Perform. Comput.*, vol. 2, no. 2, pp. 204–210, 2020.

[12] I. Riswanto and R. H. Laluma, "Klasifikasi Kelayakan Pinjaman Pada Koperasi Karyawan Menggunakan Metode Naïve Bayes Classifier Berbasis Web," *Infotronik J. Teknol. Inf. dan Elektron.*, vol. 5, no. 1, pp. 11–16, 2020.

[13] F. Soleymani, H. Masnavi, and S. Shateyi, "Classifying a lending portfolio of loans with dynamic updates via a machine learning technique," *Mathematics*, vol. 9, no. 1, pp. 1–15, 2021.

[14] H. Brawijaya, S. Samudi, and S. Widodo, "Komparasi Algoritma K-Nearest Neighbor dan Naiive Bayes Pada Pengobatan Penyakit Kutil Menggunakan Cryotheraphy," *JUITA J. Inform.*, vol. 7, no. 2, p. 93, 2019.

[15] M. A. Mukid, T. Widiharih, A. Rusgiyono, and A. Prahutama, "Credit scoring analysis using weighted k nearest neighbor," *J. Phys. Conf. Ser.*, vol. 1025, no. 1, 2018.

[16] N. S. H. Pratama, D. T. Afandi, Mulyawan, Iin, and N. D. Nuris, "Menurunkan Presentase Kredit Macet Nasabah Dengan Menggunakan Algoritma K-Nearest Neighbor," *Inf. Syst. Educ. Prof.*, vol. 5, no. 2, pp. 131–140, 2021.

[17] J. Becker, A. Radomska-Zalas, and P. Ziemba, "Rough set theory in the classification of loan applications," *Procedia Comput. Sci.*, vol. 176, pp. 3235–3244, 2020.

[18] K. Hemachandran, P. M. George, R. V. Rodriguez, R. M. Kulkarni, and S. Roy, "Performance analysis of k-nearest neighbor classification algorithms for bank loan sectors," *Adv. Parallel Comput.*, vol. 38, pp. 9–13, 2021.

[19] Roviani, D. Supriadi, and I. D. Iskandar, "Prediction of Cooperative Loan Feasibility Using the K- Nearest Neighbor Algorithm," *J. PILAR Nusa Mandiri*, vol. 17, no. 1, pp. 39–46, 2021.

[20] K. Kaarthik, G. Dharanidharan, R. B. Navalarasu, and G.

Sabarinathan, “Machine Learning Based Loan Prediction System Using Svm and Knn Algorithms,” *Turkish J. Physiother. Rehabil.*, vol. 32, no. 2, pp. 3214–3219, 2021.

[21] S. T. Safitri, D. M. Kusumawardani, C. Wiguna, D. Supriyadi, and I. Yulita, “MEASUREMENT OF VALIDITY AND RELIABILITY OF CUSTOMER SATISFACTION QUESTIONER in E-BOARDING APPICATIONS,” *J. Pilar Nusa Mandiri*, vol. 16, no. 1, pp. 1–6, 2020.