2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE) took place 20-21 November 2019 in Yogyakarta, Indonesia.

IEEE catalog number:	CFP19G48-ART
ISBN:	978-1-7281-5118-2

Organizing Committee

General Chair

Kusrini Kusrini (AMIKOM Yogyakarta University, Indonesia)

Publication Chair

Sumarni Adi (Universitas AMIKOM Yogyakarta, Indonesia)
Akhmad Dahlan (Universitas Amikom Yogyakarta, Indonesia)

Secretary

Yoga Pristyanto (Universitas Amikom Yogyakarta, Indonesia)
Nila Puspitasari (Universitas AMIKOM Yogyakarta, Indonesia)

Treasurer

Sumarni Adi (Universitas AMIKOM Yogyakarta, Indonesia)

Chair of Technical Program Committee

Emha Taufiq Luthfi (Universitas AMIKOM Yogyakarta, Indonesia)
Lesnanto Multa Putranto (UGM, Indonesia)
Arief Setyanto (Universitas AMIKOM Yogyakarta, Indonesia)
Andi Sunyoto (Universitas AMIKOM Yogyakarta, Indonesia)
Ferry Wahyu Wibowo (Universitas Amikom Yogyakarta, Indonesia)

Technical Program Committee

Intan Ermahani A. Jalil	Universiti Teknikal Malaysia Melaka	Malaysia
Mohd Helmy Abd Wahab	Universiti Tun Hussein Onn Malaysia	Malaysia
MS. Hendriyawan Achmad	Yogyakarta University of Technology	Indonesia
Sumarni Adi	Universitas AMIKOM Yogyakarta	Indonesia
I Adiyasa	Gadjah Mada University	Indonesia
I Made Agastya	Universitas Amikom Yogyakarta	Indonesia
Mohd Khairul Ikhwan Ahmad	Universiti Tun Hussein Onn Malaysia	Malaysia
Ali Qusay Al-Faris	Assumption College	USA
Shajith Ali	SSN College of Engineering, Chennai	India
Gede Angga Pradiptha	Institut Teknologi dan Bisnis STIKOM Bali	Indonesia
Igi Ardiyanto	Universitas Gadjah Mada	Indonesia
Dhani Ariatmanto	Universitas Amikom Yogyakarta	Indonesia
Agus Aribowo	Universitas Pembangunan Nasional Veteran Yogyakarta	Indonesia
Andria Arisal	Indonesian Institute of Sciences	Indonesia
Arti Arya	PESIT-Bangalore South Campus	India
Ika Astuti	Universitas Amikom Yogyakarta	Indonesia
Hasyim Asyari	Universitas Muhammadiyah Surakarta	Indonesia
Azizul Azizan	Universiti Teknologi Malaysia (UTM)	Malaysia
Aslina Baharum	Universiti Malaysia Sabah	Malaysia
Aashish Bardekar	Sipna College of Engineering and Technology, Amravati	India
Agus Bejo	Universitas Gadjah Mada	Indonesia
Adha Imam Cahyadi	Universitas Gadjah Mada	Indonesia
Alessandro Carrega	CNIT	Italy
Maria Chiara Caschera	CNR	Italy
Selem Charfi	HD Technology	France
Wichian Chutimaskul	King Mongkut's University of Technology Thonburi	Thailand
Domenico Ciuonzo	University of Naples Federico II, IT	Italy
Akhmad Dahlan	Universitas Amikom Yogyakarta	Indonesia

Ni Ketut Dewi Ari Jayanti	Institute of Technology and Business STIKOM Bali	Indonesia
Andi Wahju Rahardjo Emanuel	Universitas Atma Jaya Yogyakarta	Indonesia
Ahmad Fajar	Bina Nusantara University	Indonesia
Muhammad Faris	Universitas Gadjah Mada	Indonesia
Hanif Fatta	Universitas AMIKOM Yogyakarta	Indonesia
Anggit Ferdita Nugraha	Universitas AMIKOM Yogyakarta	Indonesia
Alireza Ghasempour	ICT Faculty	USA
Arifiyanto Hadinegoro	Amikom Universty Yogyakarta	Indonesia
Byeong-jun Han	Soongsil University	Korea (South)
Muh Hanafi	Universitas Amikom Yogyakarta	Indonesia
Seng Hansun	Universitas Multimedia Nusantara	Indonesia
Hartatik Hartatik	STMIK AMIKOM Yogyakarta	Indonesia
Mardhiya Hayaty	Universitas AMIKOM Yogyakarta	Indonesia
Purwono Hendradi	Universitas Muhammadiyah Magelang	Indonesia
Roberto Carlos Herrera Lara	National Polytechnic School	Ecuador
Indriana Hidayah	Universitas Gadjah Mada	Indonesia
Tonny Hidayat	Universitas AMIKOM Yogyakarta	Indonesia
Nor Hikmah	Universitas AMIKOM	Indonesia
Danial Hooshyar	Korea University	Korea (South)
Hozairi Hozairi	Islamic University of Madura	Indonesia
Imelda Imelda	Universitas Budi Luhur	Indonesia
Rini Indrayani	Universitas Amikom Yogyakarta	Indonesia
Nurulisma Ismail	Universiti Malaysia Perlis	Malaysia
Iswandi Iswandi	Gadjah Mada University	Indonesia
Ramkumar Jaganathan	VLB Janakiammal College of Arts and Science	India
Arihant Jain	Jaipur Engineering College & Research Centre	India
Biao Jiang	The City University of New York	USA
Dimitrios Kallergis	University of West Attica	Greece
Sandy Kosasi	STMIK Pontianak	Indonesia

Krisnawati Krisnawati	University of AMIKOM Yogyakarta	Indonesia
Ryan Kristianto	Universitas Amikom Yogyakarta	Indonesia
Sumit Kushwaha	KNIT Sultanpur	India
Kusnawi Kusnawi	AMIKOM University	Indonesia
Kusrini Kusrini	AMIKOM Yogyakarta University	Indonesia
Pavel Loskot	Swansea University	United Kingdom (Great Britain)
Emha Taufiq Luthfi	Universitas AMIKOM Yogyakarta	Indonesia
Mahdin Mahboob	Stony Brook University	USA
Robert Marco	STMIK Amikom Yogyakarta	Indonesia
Prita Dewi Mariyam	Universitas Indonesia	Indonesia
Ahlihi Masruro	Universitas AMIKOM Yogyakarta	Indonesia
Ratheesh Kumar Meleppat	University of California Davis	USA
Ahmed Mobashsher	The University of Queensland	Australia
Seyed Sahand Mohammadi Ziabari	Vrije University of Amsterdam	The Netherlands
Amrit Mukherjee	Jiangsu University	China
I Wayan Mustika	Universitas Gadjah Mada	Indonesia
Syibrah Naim	Universiti Sains Malaysia	Malaysia
Warsun Najib	Universitas Gadjah Mada	Indonesia
Asro Nasiri	University of Amikom Yogyakarta	Indonesia
Shah Nazir	University of Peshawar	Pakistan
Ruzelita Ngadiran	Universiti Malaysia Perlis	Malaysia
Bhanu Nugraha	Universitas Amikom Yogyakarta	Indonesia
Hanung Nugroho	Universitas Gadjah Mada	Indonesia
Prapto Nugroho	Universitas Gadjah Mada	Indonesia
Nitish Ojha	DIT University, Dehradun	India
Ilker Ali Ozkan	Selcuk University	Turkey
Oktavia Permata	Institut Teknologi Telkom Surabaya	Indonesia
Anugerah Persada	Universitas Gadjah Mada	Indonesia
Kiran Sree Pokkuluri	Shri Vishnu Engineering College for Women	India

N. Prabaharan	SASTRA Deemed University	India
Gede Pramudya Ananta	Universiti Teknikal Malaysia Melaka	Malaysia
Andri Pranolo	Universitas Ahmad Dahlan	Indonesia
Prihandoko Prihandoko	University of Gunadarma	Indonesia
Yoga Pristyanto	Universitas Amikom Yogyakarta	Indonesia
Reza Pulungan	Universitas Gadjah Mada	Indonesia
Nila Puspitasari	Universitas AMIKOM Yogyakarta	Indonesia
Guntur Putra	University of New South Wales	Australia
Lesnanto Multa Putranto	UGM	Indonesia
Ali Rafiei	University of Technology Sydney	Australia
Suwanto Raharjo	Informatics of The Institut Sains & Teknologi AKPRIND Yogyakarta	Indonesia
Hemant Kumar Rath	Tata Consultancy Services	India
Rizky Rizky	Universitas AMIKOM Yogyakarta	Indonesia
Rumini Rumini	Universitas AMIKOM Yogyakarta	Indonesia
G. p. Sajeev	Amrita Vishwa Vidyapeetham	India
Andreas Sandiwan	Gadjah Mada University	Indonesia
Banu Santoso	Universitas AMIKOM Yogyakarta	Indonesia
Leo Santoso	Petra Christian University	Indonesia
Sayantam Sarkar	Vijaya Vittala Institute of Technology	India
Theopilus Bayu Sasongko	Universitas AMIKOM Yogyakarta	Indonesia
Mithileysh Sathiyanarayanan	City, University of London	United Kingdom (Great Britain)
Vaibhav Saundarmal	Marathwada Institute of Technology, Aurangabad	India
Enny Sela	Universitas Teknologi Yogyakarta	Indonesia
Erni Seniwati	Universitas AMIKOM Yogyakarta	Indonesia
Anindita Septiarini	Univeristas Mulawarman	Indonesia
Amel Serrat	USTO MB	Algeria
Bayu Setiaji	Universitas AMIKOM Yogyakarta	Indonesia
Noor Akhmad Setiawan	Universitas Gadjah Mada	Indonesia
Emy Setyaningsih	Institute of Science & Technology AKPRIND	Indonesia

Arief Setyanto	Universitas AMIKOM Yogyakarta	Indonesia
Iwan Setyawan	Satya Wacana Christian University	Indonesia
Andik Setyono	Universitas Dian Nuswantoro	Indonesia
Aditi Sharma	Quantum University, Roorkee, Uttarakhand	India
Vesh Raj Sharma Banjade	Intel Corporation	USA
Abdul Samad Shibghatullah	UCSI University	Malaysia
Monika Sikri	Cisco Systems India Pvt Ltd.	India
Amando Singun	Higher College of Technology, Muscat	Oman
Achmad Solichin	Universitas Budi Luhur	Indonesia
Iickho Song	Korea Advanced Institute of Science and Technology	Korea (South)
Yi-Jen Su	Shu-Te University	Taiwan
Suhirman Suhirman	Universitas Teknologi Yogyakarta	Indonesia
Andi Sunyoto	Universitas AMIKOM Yogyakarta	Indonesia
Sutarman Sutarman	Faculty Of Information Technology and Business, University Technoloy of Yogyakarta	Indonesia
Melwin Syafrizal	Universitas AMIKOM Yogyakarta	Indonesia
Srinivasulu Tadisetty	Kakatiya University College of Engineering and Technology	India
Chakib Taybi	Mohammed First University	Morocco
Ivanna Timotius	Satya Wacana Christian University	Indonesia
Evi Triandini	Institut Teknologi dan Bisnis STIKOM Bali	Indonesia
Hastari Utama	Universitas Amikom Yogyakarta	Indonesia
Ema Utami	Universitas Amikom Yogyakarta	Indonesia
Dario Vieira	EFREI	France
Anik Vitianingsih	Universitas Dr Soetomo	Indonesia
Mochammad Wahyudi	Universitas Gadjah Mada	Indonesia
Sri Ngudi Wahyuni	Universitas AMIKOM Yogyakarta	Indonesia
Fikri Waskito	Universitas Gadjah Mada	Indonesia
Leong Wen Chek	University of Malaya	Malaysia
Sunu Wibirama	Universitas Gadjah Mada	Indonesia
Adi Wibowo	Diponegoro University	Indonesia

Ferry Wahyu Wibowo	Universitas Amikom Yogyakarta	Indonesia
Sigit Wibowo	Gadjah Mada University	Indonesia
Teguh Wibowo	Gadjah Mada University	Indonesia
Ainul Yaqin	Universitas Amikom Yogyakarta	Indonesia
Thaweesak Yingthawornsuk	King Mongkut's University of Technology Thonburi	Thailand
Uky Yudatama	Universitas Indonesia	Indonesia
Go Yun Il	Heriot-Watt University Malaysia	Malaysia
Muhammad Yusuf	University of Trunojoyo, Madura	Indonesia
Sri Zuliana	UIN Sunan Kalijaga	Indonesia
Nur Zareen Zulkarnain	Universiti Teknikal Malaysia Melaka	Malaysia

Additional Reviewers

Intan Ermahani A. Jalil	Universiti Teknikal Malaysia Melaka	Malaysia
Mohd Helmy Abd Wahab	Universiti Tun Hussein Onn Malaysia	Malaysia
MS. Hendriyawan Achmad	Yogyakarta University of Technology	Indonesia
I Wayan Adiyasa	Gadjah Mada University	Indonesia
Ali Qusay Al-Faris	Assumption College	USA
Shajith Ali	SSN College of Engineering, Chennai	India
Gede Angga Pradiptha	Institut Teknologi dan Bisnis STIKOM Bali	Indonesia
Dhani Ariatmanto	Universitas Amikom Yogyakarta	Indonesia
Agus Sasmito Aribowo	Universitas Pembangunan Nasional Veteran Yogyakarta	Indonesia
Andria Arisal	Indonesian Institute of Sciences	Indonesia
Arti Arya	PESIT-Bangalore South Campus	India
Aslina Baharum	Universiti Malaysia Sabah	Malaysia
Maria Chiara Caschera	CNR	Italy
Wichian Chutimaskul	King Mongkut's University of Technology Thonburi	Thailand
Domenico Ciuonzo	University of Naples Federico II, IT	Italy
Ni Ketut Dewi Ari Jayanti	Institute of Technology and Business STIKOM Bali	Indonesia
Andi Wahju Rahardjo Emanuel	Universitas Atma Jaya Yogyakarta	Indonesia
Ahmad Nurul Fajar	Bina Nusantara University	Indonesia
Anggit Ferdita Nugraha	Universitas AMIKOM Yogyakarta	Indonesia
Alireza Ghasempour	ICT Faculty	USA
Seng Hansun	Universitas Multimedia Nusantara	Indonesia
Hartatik Hartatik	STMIK AMIKOM Yogyakarta	Indonesia
Mardhiya Hayaty	Universitas AMIKOM Yogyakarta	Indonesia
Roberto Carlos Herrera Lara	National Polytechnic School	Ecuador
Tonny Hidayat	Universitas AMIKOM Yogyakarta	Indonesia
Danial Hooshyar	Korea University	Korea (South)
Hozairi Hozairi	Islamic University of Madura	Indonesia
Rini Indrayani	Universitas Amikom Yogyakarta	Indonesia
Nurulisma Ismail	Universiti Malaysia Perlis	Malaysia
Ramkumar Jaganathan	VLB Janakiammal College of Arts and Science	India
Biao Jiang	The City University of New York	USA
Sandy Kosasi	STMIK Pontianak	Indonesia
Krisnawati Krisnawati	University of AMIKOM Yogyakarta	Indonesia
Ryan Putranda Kristianto	Universitas Amikom Yogyakarta	Indonesia
Sumit Kushwaha	KNIT Sultanpur	India
Pavel Loskot	Swansea University	United Kingdom (Great Britain)
Prita Dewi Mariyam	Universitas Indonesia	Indonesia
Ratheesh Kumar Meleppat	University of California Davis	USA
Amrit Mukherjee	Jiangsu University	China
Syibrah Naim	Universiti Sains Malaysia	Malaysia
Asro Nasiri	University of Amikom Yogyakarta	Indonesia
Shah Nazir	University of Peshawar	Pakistan
Ruzelita Ngadiran	Universiti Malaysia Perlis	Malaysia
Hanung Adi Nugroho	Universitas Gadjah Mada	Indonesia
Prapto Nugroho	Universitas Gadjah Mada	Indonesia
Ilker Ali Ozkan	Selcuk University	Turkey
Oktavia Ayu Permata	Institut Teknologi Telkom Surabaya	Indonesia
Gede Pramudya Ananta	Universiti Teknikal Malaysia Melaka	Malaysia
Prihandoko Prihandoko	University of Gunadarma	Indonesia
Guntur Dharma Putra	University of New South Wales	Australia
Lesnanto Multa Putranto	UGM	Indonesia
Suwanto Raharjo	Informatics of The Institut Sains & Teknologi AKPRIND Yogyakarta	Indonesia
Bambang Riyono	Gadjah Mada University	Indonesia
Leo Santoso	Petra Christian University	Indonesia
Vaibhav Dudhaji Saundarmal	Marathwada Institute of Technology, Aurangabad	India
Enny Sela	Universitas Teknologi Yogyakarta	Indonesia
Ririn Septrisulviani	Universitas Gadjah Mada	Indonesia
Amel Serrat	USTO MB	Algeria

Lilis Susanti Setianingsih	Universitas Gadjah Mada	Indonesia
Noor Akhmad Setiawan	Universitas Gadjah Mada	Indonesia
Iwan Setyawan	Satya Wacana Christian University	Indonesia
Abdul Samad Shibghatullah	UCSI University	Malaysia
Achmad Solichin	Universitas Budi Luhur	Indonesia
Iickho Song	Korea Advanced Institute of Science and Technology	Korea (South)
Andi Sunyoto	Universitas AMIKOM Yogyakarta	Indonesia
Melwin Syafrizal	Universitas AMIKOM Yogyakarta	Indonesia
Srinivasulu Tadisetty	Kakatiya University College of Engineering and Technology	India
Chakib Taybi	Mohammed First University	Morocco
Ivanna Timotius	Satya Wacana Christian University	Indonesia
Evi Triandini	Institut Teknologi dan Bisnis STIKOM Bali	Indonesia
Anik Vega Vitianingsih	Universitas Dr Soetomo	Indonesia
Leong Wen Chek	University of Malaya	Malaysia
Adi Wibowo	Diponegoro University	Indonesia
Ainul Yaqin	Universitas Amikom Yogyakarta	Indonesia
Thaweesak Yingthawornsuk	King Mongkut's University of Technology Thonburi	Thailand
Uky Yudatama	Universitas Indonesia	Indonesia
Sri Utami Zuliana	UIN Sunan Kalijaga	Indonesia

2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)

Parallel Session 1-A

Parallel Session 1-B

Parallel Session 1-C

Parallel Session 1-D

Parallel Session 1-E

Parallel Session 2-A

Parallel Session 2-B

Parallel Session 2-C

Parallel Session 2-D

Parallel Session 2-E

Parallel Session 3-A

Parallel Session 3-B

Parallel Session 3-C

Parallel Session 3-D

Parallel Session 3-E

Comparison of Classification Methods using Historical Loan Application Data

Yohanes R. Laberto Kelen
laberto.kelen@gmail.com
Magister Informatika, Universitas Atma Jaya Yogyakarta[1]
Indonesia 55281

Andi Wahju Rahardjo Emanuel
andi.emanuel@uajy.ac.id
Magister Informatika,Universitas Atma Jaya Yogyakarta
Indonesia 55281

Abstract—Every year, the number of cooperatives in the province of East Nusa Tenggara continues to grow. Cooperatives are present with the aim of helping the community on the financial side. The cooperative offers the principle of saving and providing low-interest loans to its members. But there are times when lending is subjective. This condition is a major factor in the occurrence of errors in providing credit that leads to congestion (non-performing loans). This study focuses on the comparison of five classification methods using historical loan application data for a Multipurpose Cooperative in East Nusa Tenggara. The 5 methods are Naïve Bayes, K-Nearest Neighbor (KNN), Support Vector Machine (SVM) Random Forest, and C4.5. In the test results, it turns out that the C4.5 Method has better accuracy and a smaller error rate.

Keywords—big data, classification, cooperative, Historical Loan Application.

I. INTRODUCTION

Based on article 4 in Indonesian Law Number 17 of 2012 concerning cooperatives, cooperatives have a goal to prosper members in particular and society in general [1]. The cooperative in the island province of East Nusa Tenggara is expected to be the leading spearhead in the economy in the community.

The development of cooperatives in East Nusa Tenggara continues to increase. The number of cooperatives in this province in 2018 was 4,317 units. There was a significant increase in the number compared to 2014, which amounted to 2,818 active cooperative units. Cooperatives help the community of East Nusa Tenggara because it is considered very easy to reach and not difficult for the community to get funds. The interest offered by the cooperative is so small that the community is interested in joining the Cooperative.

But at a Multipurpose Cooperative in East Nusa Tenggara, researchers found cases of Non-Performing Loan (NPL). NPL is the failure of customers to repay loans. It turns out that this case often occurs in cooperatives, as seen from the percentage of NPL from 2014-2017, which reached 76%. It turned out that the consideration of lending, which was declared feasible by the creditor was sometimes subjective. Loans are given to members because there are elements of kinship or acquaintances without looking at the loan criteria themselves. This is also a major factor in the cause of non-performing loans.

In this study, five classification methods will be compared using historical loan application data. The loan application data consists of 6 criteria, namely the loan amount, collateral amount, loan period, collateral status, type of interest, and marital status. The contribution of the proposed research is finding the best method of accuracy, so it can be used as a reference for further research, such as to develop a knowledge base and decision support system. So that going forward, the process of granting credit to customers can pay attention to these available criteria.

II. RELATED WORKS

The following are some previous studies on the classification methods that will be discussed in this paper. Naive Bayes Classifier is a popular algorithm that can be used to predict singers just by lyrics. Using available datasets, this method successfully predicts with very good results [2]. Naïve Bayes ability is used to classify SMS messages for Disaster Response. The test results get an accuracy of up to 89% [3]. Using the K-Nearest Neighbor classification approach, grape leaf disease is classified effectively. Although leaf images and identification of leaf disease are not easy tasks using images of agronomic plants [4].

Support Vector Machine (SVM) is the new machine learning algorithm used in pattern recognition problems including texture classification [5]. SVM minimizes structural risk, which excludes traditional empirical risk minimization, which is commonly used in conventional neural networks [5][6]. SVM is used to classify Bangla Language scripts with accuracy reaching an average of 94.78% [7]. Decision trees are algorithms that are very fast to be trained and evaluated. This method is also very easy to interpret and can be visualized [8]. Random forest and C4.5 are methods in the decision tree algorithm.

Random forest-based techniques can be used more useful for unsupervised learning [9]. C4.5 algorithm is a classification algorithm with a decision tree technique. C4.5 is the development of the ID3 algorithm [10]. This method is often used because accuracy tends to be high and has been used by many researchers for classification purposes [11].

III. METHODOLOGY

A. Machine Learning Classifier

In this paper, the author examines five different classification methods. Some of these algorithms will be explained in this section:

(1) Naïve Bayes is a classification supervised learning algorithm [12]. This is a popular algorithm based on conditional probability theorems to determine new feature vector classes.

(2) K-Nearest Neighbor (KNN) classification algorithm that is safe and efficient can hide class labels and data access patterns. Viewed from the length of the classification time, the results of the performance analysis of this method achieved 17 times better performance than the existing scheme [13].

(3) Support Vector Machine (SVM) is SVM is a well-known classification technique based on theory statistical learning [14]. SVM consists of a two classes classification model; a model defined in the feature space of the largest linear classifier [15];

(4) Random forest is an algorithm which based techniques can be used more useful for unsupervised learning [9]. RF is very efficient with low computational costs, few parameter settings, and good performance [16].

(5) C4.5 algorithm is a classification algorithm with a decision tree technique [17]. C4.5 is the development of the ID3 algorithm [10]. This method is often used because accuracy tends to be high and has been used by many researchers to classify things [11].

B. Data source

The authors have conducted research at a Multipurpose Cooperative in East Nusa Tenggara. They have allowed the author to use their historical loan application data for 2016-2017. The number of data samples used is 2500 loan application data. The author uses a ratio of 80:20 for the test scenario. So, 2000 data is used for training data and the rest is used as test data.

This historical loan application data has been edited so that attributes are formed. Table 1 shows 6 attribute variables. The six attributes include loan amount (A1), collateral amount (A2), loan period (A3), collateral status (A4), loan interest (A5), and marital status (A6).

TABLE I. DISTRIBUTION OF ATTRIBUTES

Code	Attribute Name
A1	Loan amount
A2	Collateral amount
A3	Loan period
A4	Collateral status
A5	Loan interest
A6	Marital status

C. Discretization

Discretization converts continuous data values into categorical data. In this paper, discretion is carried out on class attributes, namely the loan amount (A1), collateral amount (A2) and load period (A3) attributes.

TABLE II. LOAN AMOUNT CLASSIFICATION

Class	Intervals (IDR)	Class Name
1	< 2.000.000	Very Low (VL)
2	2.000.000 – 4.000.000	Low (L)
3	4.000.000 – 6.000.000	Moderate (M)
4	6.000.000 – 8.000.000	High (H)
5	> 8.000.000	Very High (VH)

This classification is intended to label data. In table 2, it can be seen that there are 5 classes for the loan amount attribute (A1). The class is very low (VL), low (L), moderate (M), high (H) and very high (VH).

TABLE III. COLLATERAL AMOUNT CLASSIFICATION

Class	Intervals (IDR)	Class Name
1	< 2.000.000	Low (L)
2	2.000.000 – 4.000.000	Moderate (M)
3	> 4.000.000	High (H)

Class for loan value (A2) consists of low (L), moderate (M), and high (H) can be seen in table 3. The amount of collateral is one of the conditions in making a loan. collateral is usually seen from the number of members' deposits in the cooperative. But it can also be in the form of proof of ownership of valuables. in this study, collateral is used in the form of cooperative deposits.

TABLE IV. LOAN PERIOD CLASSIFICATION

Class	Intervals	Class Name
1	6 months	Very slow (VS)
2	10 months	Slow (S)
3	12 months	Moderate (M)
4	18 months	Fast (F)
5	24 months	Very fast (VF)

Class for loan period (A5) can be seen in table 4. the five classes are very slow (VS), slow (S), moderate (M), fast (F), and very fast (VF). The loan period is the time needed by the cooperative members to repay the loan to the cooperative.

All attribute elements have several categories which can be seen in table 5.

TABLE V. DISTRIBUTION OF CLASS

Attribute Code	Class Name (Class Code)
A1	Very Low (VL), Low (L), Moderate (M), High (H), Very High (VH)
A2	Low (L), Moderate (M), High (H)
A3	Very slow (VS) Slow (S), Moderate (M), Fast (F), Very fast (VF)
A4	True (T), False (F)
A5	Down interest (D), Flat interest (F)
A6	Married (M), Single (S)

After the steps above, the next step is to determine the attribute results (A8). Authors aiming at classification will classify two classes, *Lancar* and *Macet*. *Lancar* class declared for cooperative members who will successfully repay the loans; and *Macet* class if the loan will fails to pay off or non-performing loan (NPL). Testing the data will be explained in the next sub-chapter.

D. Evaluation

The confusion matrix is used to find the accuracy of classification model in testing several methods in this study [10] [11]. The confusion matrix is a matrix formed by several training data and test data.

The confusion matrix size n x n shows the results of prediction and classification, where n is the number of predetermined class attributes. The four table fields contain the number of documents classified as True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

TABLE VI. CONFUSION MATRIX

		Actual	
		YES	***NO***
Predicted	***YES***	TP	FP
	NO	FN	TN

TP is the amount of positive data that is predicted right. TN is the amount of positive data that is predicted wrong. FP is the amount of negative data that is predicted right. TN is the amount of negative data that is predicted wrong.

Accuracy is the percentage of the results of the comparison of cases that were correctly identified by the number of all cases. The accuracy of the classifier can be found by formula 1.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision, the positive predictive value between samples taken. The following formula calculates precision:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

APER: (Apparent Error Rate) or the so-called error rate is an evaluation measure that is used to see opportunities for classification errors produced by a classification function. The smaller the APER value, the better the classification results. The calculation of the APER can be seen in equation 3.

$$APER = \frac{FP + FN}{TP + TN + FP + FN} \quad (3)$$

IV. EXPERIMENTAL RESULTS

The authors tested 500 test data on 2000 training data with WEKA data mining software that produced a pretty good percentage. The classification results are shown in table 7.

TABLE VII. CLASSIFICATION TEST

Methods	Correctly classified	Incorrectly classified	%
NB	491	9	98.2
KNN	495	5	99
SVM	493	7	98.6
RF	495	5	99
C4.5	496	4	99.2

C4.5 classifier algorithm gives the highest classification in terms of correctly classified instances [18]. Table 8 shows the results of the classification of the random forest method and KNN obtained a value of 99% and the highest classification obtained by method C4.5 with a value of 99.2%.

TABLE VIII. CONFUSION MATRIX FROM TEST RESULT

Methods	TP	TN	FP	FN
NB	328	0	9	163
KNN	334	2	3	161
SVM	332	2	5	161
RF	334	2	3	161
C4.5	335	2	2	161

Table 8 shows a confusion matrix classified by WEKA data mining software with class results (A8). Table 9 shows the comparison results. Each method tested has a value of True Positive and False Negative are greater than True Negative and False Positive. This proved that the success rate of classification correctly is very high according to table 7.

TABLE IX. COMPARISON OF RESULT

Methods	Accuracy	Precision	APER
NB	98.2	100	1.80
KNN	99	99.4	1.00
SVM	98.6	99.4	1.40
RF	99	99.4	1.00
C4.5	99.2	99.41	0.80

The results of the testing experiment show that the C4.5 method has the smallest APER value with 0.80. C4.5 method also has the biggest accuracy with 99.2%. Naïve Bayes has the biggest precision with 100%.

The classification method C4.5 has a good result than the other four classification methods with a dataset for applying the credit to a cooperative in East Nusa Tenggara. The C4.5 method on WEKA Data Mining Software turns out to produce a decision tree. The decision tree can be converted into programming logic.

TABLE X. KNOWLEDGE BASE

If A1 = Very Low then *Lancar*
If A1 = Low then *Lancar*
If A1 = Moderate then *Lancar*
If A1 = High then
    If A6 = Married then *Macet*
    If A6 = Single then *Lancar*
If A1 = Very High then
    If A6 = Married then *Macet*
    If A6 = Single then
        If A2 = Low then
            If A5 = Flat Interest then *Macet*
            If A5 = Down Interest then *Lancar*
        If A2 = High then *Macet*
        If A2 = Moderate then
            If A5 = Flat Interest then
                If A4 = True then *Macet*
                If A4 = False then *Lancar*
            If A5 = Down Interest then *Macet*

The logic in table 10 can be implemented into a knowledge base for a decision support system.

V. Conclusion and Discussion

Each loan submitted by the members of the cooperative must be well considered by the creditor. Credit criteria are expected to be an important concern. Avoid accepting loan proposals based on subjective factors. This paper shows the results of the highest testing accuracy obtained by method C4.5 with a percentage of 99.41%. Naïve Bayes has the best precision of the five tested methods.

The results are obtained from the calculation of the confusion matrix. The confusion matrix is generated from WEKA data mining software. Accuracy can be increased by adding training data, attributes. Besides that, you can also use algorithm optimization. The results of this paper are not only comparisons between the five classification methods. But there is a decision tree based on the C4.5 algorithm. This decision tree is expected to be a reference for creditors to determine loan proposals that fit the criteria.

References

[1] "UNDANG-UNDANG REPUBLIK INDONESIA NOMOR 17 TAHUN 2012 TENTANG PERKOPERASIAN," pp. 1–56, 2012.

[2] D. Buži and J. Dobša, "Lyrics Classification using Naive Bayes," *2018 41st Int. Conv. Inf. Commun. Technol. Electron. Microelectron.*, pp. 1011–1015, 2018.

[3] A. J. Ordoñez, R. E. J. Paje, and R. N. Naz, "SMS Classification Method for Disaster Response using Naïve Bayes Algorithm," pp. 233–236, 2018.

[4] I. Conference and C. Systems, "An Individual Grape Leaf Disease Identification Using Leaf Skeletons and KNN Classification," 2017.

[5] G. S. Raghtate, "Comparison of Classification Methods with Second Order Statistical Analysis and Wavelet Transform For Texture Image Classification," 2015.

[6] S. Begum, R. Sarkar, and A. Adaboost, "Cancer Classification from Gene Expression Based Microarray Data Using SVM Ensemble," pp. 13–16, 2015.

[7] M. K. Shukla, A. Rana, and H. Banka, "Classification of the Bangla Script Document using SVM," pp. 31–34, 2016.

[8] S. H. Ahn *et al.*, "Comparison of Clustering Methods for Obesity Classification," *2018 IEEE Int. Conf. Ind. Eng. Eng. Manag.*, pp. 1821–1825, 2018.

[9] S. Hussain, S. Muhammad, and M. Majid, "Segmentation of glioma tumors in brain using deep convolutional neural network," in *Neurocomputing*, 2018, vol. 282, pp. 248–261.

[10] C. Algorithm and J. A. Suyatno, "Rainfall Forecasting in Bandung Regency using," *2018 6th Int. Conf. Inf. Commun. Technol.*, vol. 0, no. c, pp. 324–328, 2018.

[11] A. G. Pertiwi, T. Widyaningtyas, and U. Pujianto, "Classification of Province Based on Dropout Rate," pp. 410–413, 2017.

[12] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, "A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms," in *Mobile Information Systems*, 2018, vol. 2018.

[13] H. Kim and H. Kim, "A Privacy-Presering k NN Classification Algorithm using Yao ' s garbled circuit on Cloud Computing," 2017.

[14] H. Elaidi, Y. Elhaddar, Z. Benabbou, and H. Abbar, "An Idea of a Clustering Algorithm using Support Vector Machines Based on Binary Decision Tree," in *2018 International Conference on Intelligent Systems and Computer Vision, ISCV 2018*, 2018, vol. 2018-May, no. 5, pp. 1–5.

[15] G. Cheng and X. Tong, "Fuzzy Clustering Multiple Kernel Support Vector Machine," in *2018 International Conference on Wavelet Analysis and Pattern Recognition*, 2018, vol. 2018-July, pp. 7–12.

[16] H. Zhong, X. Song, and L. Yang, "Vessel Classification from Space-based AIS Data Using Random Forest," in *2019 5th International Conference on Big Data and Information Analytics (BigDIA)*, 2019, pp. 9–12.

[17] M. E. Cintra, M. C. Monard, and H. A. Camargo, "A Fuzzy Decision Tree Algorithm Based on C4.5," *Mathware & Soft. Computing Magazine*, vol. 20, pp. 56–62.

[18] H. Erol, "Classification Performances Of Data Mining Clustering Algorithms For Remotely Sensed Multispectral Image Data," pp. 1–4, 2018.

the 4th icitisee 2019

International Conference on Information Technology, Information Systems, and Electrical Engineering

UNIVERSITAS AMIKOM YOGYAKARTA

IEEE

CERTIFICATE

This certificate is presented to

Mr. Andi Wahju Rahardjo Emanuel

as

for outstanding contribution in the 4th 2019 International Conference on Information Technology, Information System, and Electrical Engineering held in Yogyakarta, Indonesia on 20 - 21 November 2019

Conference No. **#48480**

ID number 1129597

ICITISEE

Dr. Kusrini, M.Kom.
General Chair

IEEE INDONESIA SECTION

Department of Electrical and Information Engineering Universitas Gadjah Mada

UNIVERSITAS AMIKOM PURWOKERTO