

# PENGEMBANGAN MODEL PERSETUJUAN KREDIT NASABAH BANK DENGAN ALGORITMA KLASIFIKASI NAÏVE BAYES, DECISION TREE, DAN ARTIFICIAL NEURAL NETWORK

Puji Rahmawati<sup>1</sup>, Aisyah Larasati<sup>1\*</sup>, Marsono<sup>2</sup>

<sup>1</sup>Program Studi Teknik Industri, Fakultas Teknik, Universitas Negeri Malang

<sup>2</sup>Program Studi Pendidikan Teknik Mesin, Fakultas Teknik, Universitas Negeri Malang  
Jl. Semarang No.5, Malang, Indonesia 65145

(Received: March 15, 2021/ Accepted: December 2, 2021)

## Abstrak

Kredit merupakan salah satu solusi bagi masyarakat untuk mendapatkan pinjaman dari bank dengan waktu pembayaran yang dilakukan secara berkala sesuai dengan ketentuan yang telah disetujui baik oleh bank maupun masyarakat itu sendiri sebagai calon nasabah. Berdasarkan hal tersebut, penelitian ini bertujuan untuk melakukan klasifikasi data profil nasabah yang memiliki peluang untuk mengajukan kredit pinjaman atau tidak menggunakan data mining dengan bantuan tiga algoritma yaitu Naïve Bayes, Decision Tree, Artificial Neural Network (ANN), dan melihat kriteria profil nasabah yang dihasilkan pada masing-masing hasil prediksi sebagai acuan bagi bank dalam penawaran kredit kepada nasabah. Evaluasi hasil ketiga algoritma yang digunakan dilakukan dengan melihat nilai gain ratio, confusion matrix, dan kurva Receiver Operating Characteristics (ROC) dengan melihat nilai Area Under Curve (AUC) yang dihasilkan. Hasil yang didapatkan menunjukkan bahwa algoritma ANN menghasilkan nilai akurasi tertinggi dengan akurasi sebesar 99.61% dan AUC = 0.983. Disusul dengan decision tree menghasilkan nilai akurasi sebesar 99.36% dan AUC = 0.999. Terakhir, algoritma naïve bayes dengan nilai akurasi = 90.79% dan AUC sebesar 0.935. Karakteristik nasabah yang memiliki peluang untuk mengajukan kredit jika income rata-rata pada rentang 101-116 atau > 116 USD setiap bulannya, age = 23 - 43 tahun, family >4, memiliki CD Account, dan CC Average = 3 - 9 USD. Terdapat tujuh faktor importance yang mempengaruhi hasil tiga algoritma tersebut yaitu income, mortgage, family, CD account, CC average, education, dan age.

**Kata kunci:** kredit; data mining; naïve bayes; decision tree; artificial neural network

## Abstract

*[The Development of Credit Approval Model using Naïve Bayes, Decision Tree, and Artificial Neural Network Algorithm]* Credit is one of the solutions for the public to get a loan from a bank with payment that is made periodically in accordance with the terms agreed by both the bank and the community itself as a potential customer. Based on this, this study aims to classify customer profile data who have the opportunity to apply for a loan or not use data mining with the help of three algorithms, namely Naïve Bayes, Decision Tree, Artificial Neural Network (ANN), and see the customer profile criteria produced. in each of the prediction results as a reference for the bank in offering credit to customers. Evaluation of the results of the three algorithms used is done by looking at the value of the gain ratio, confusion matrix, and Receiver Operating Characteristics (ROC) curve by looking at the resulting Area Under Curve (AUC) value. The results obtained show that the ANN algorithm produces the highest accuracy value with an accuracy of 99.61% and AUC = 0.983. Followed by the decision tree, the accuracy value is 99.36% and AUC = 0.999. Lastly, the naïve Bayes algorithm with an accuracy value of = 90.79% and an AUC of 0.935. Characteristics of customers who have the opportunity to apply for credit if the average income is in the range 101-116 or > 116 USD each month, age = 23 - 43 years, family > 4, has a CD Account, and CC Average = 3 - 9 USD. There are seven factors of importance that affect the results of the three algorithms, namely income, mortgage, family, CD account, CC average, education, and age.

**Keywords:** credit; data mining; naïve bayes; decision tree; artificial neural network

---

\*Penulis Korespondensi.

E-mail: aisyah.larasati.ft@um.ac.id

## 1. Pendahuluan

Kredit merupakan salah satu jalan pintas yang terkadang diambil oleh masyarakat untuk memenuhi kebutuhannya. Oleh sebab itu, sangat penting untuk dapat melakukan identifikasi atau evaluasiajuan kredit pinjaman (Hadianto, dkk. 2019) serta pengawasan nasabah yang telah mengajukan kredit pinjaman melalui data historis yang ada, agar bank dapat memberikan keputusan yang tepat dan dapat terhindar dari risiko kredit yang sering terjadi saat ini (Ivandari, 2015; Rani, 2015; Pandie, 2018).

Menurut Undang-Undang Republik Indonesia Pasal 1 Nomor 10 Tahun 1998, kredit merupakan tagihan atau penyediaan uang yang dapat dinyatakan sama dengan itu, antara bank dengan pihak lain, dengan mengharuskan pihak lain tersebut untuk membayar tagihannya dengan jangka waktu yang telah ditentukan. Pinjaman kredit yang diberikan oleh bank sering kali mendapatkan kendala karena tidak semua nasabah setiap bulannya akan membayar bunga sesuai dengan waktu yang telah ditentukan. Hal tersebut oleh bank sering disebut dengan kredit macet. Banyak dari calon nasabah melakukan kecurangan dalam proses pengajuan, sehingga tingkat kredit macet ini kian hari kian meningkat (Iriadi & Nuraeni, 2016; Nuraeni, 2017).

Berdasarkan hal tersebut, *data mining* dipilih untuk dapat membantu memberikan penyelesaian masalah tersebut. *Data Mining* adalah proses yang dilakukan untuk mendapatkan informasi tersembunyi dari sekumpulan data, baik itu berupa pola (*pattern*) ataupun hubungan dengan tujuan untuk melakukan klasifikasi, *clustering*, ataupun visualisasi data (Fiska, 2017). Pemilihan teknik *data mining* didasarkan pada kemampuannya yang dapat mengidentifikasi pola dalam data yang belum diketahui sebelumnya secara otomatis, khususnya pada data dengan jumlah yang besar. Sehingga pemrosesan menjadi lebih mudah dengan waktu yang lebih sedikit dibandingkan dengan teknik manual. Terdapat beberapa penelitian sejenis dengan melakukan identifikasi nasabah bank seperti menggunakan bantuan algoritma *decision tree* atau C4.5 (Ivandari, 2015; Syafnur, 2017; Sari, 2015; Honesqi, 2017; Rani, 2015; Iriadi & Nuraeni 2016). Namun, dari penelitian terdahulu belum terdapat penggunaan kombinasi tiga algoritma serta karakteristik atau profil nasabah pada masing-masing luaran kelas yang dihasilkan.

Penelitian ini bertujuan untuk melakukan klasifikasi dengan identifikasi kelompok nasabah yang memiliki peluang untuk mengajukan atau tidak mengajukan kredit pinjaman dari data historis profil nasabah yang dimiliki oleh bank. Klasifikasi digunakan dengan bantuan tiga algoritma *data mining* yaitu *Naïve Bayes*, *Decision Tree*, dan *Artificial Neural Network* (ANN). Klasifikasi adalah proses untuk menemukan pola dalam data yang kemudian akan digunakan untuk menentukan kelompok atau anggota data lain yang memiliki pola yang sama dengan data tersebut. ANN terdiri dari *neuron* buatan dengan tiga lapisan umum yaitu *input layer*, *hidden layer*, dan *output layer* (Larasati, 2019). Sedangkan *decision tree* salah satu algoritma yang paling sering digunakan, serta dapat

memproses data dengan tipe numerik atau kategorik (Larasati, 2019).

Hasil yang didapatkan akan di evaluasi dengan melihat *gain ratio*, *confusion matrix*, ataupun nilai dari *ROC curve* yang didapatkan. *Gain ratio* merupakan salah satu nilai dalam algoritma *decision tree* sebagai penentu simpul yang akan dibangun menjadi sebuah pohon (*tree*). Atribut yang memiliki nilai *gain* yang tertinggi akan dipilih menjadi simpul akar atau teratas (Hidayah, 2019). Kemudian, dalam *confusion matrix*, terdapat empat bagian yang digunakan untuk melihat performansi hasil dari proses klasifikasi yang memberikan nilai presisi, akurasi dan nilai *recall*, antara lain (Nuraeni, 2017), yaitu *True Negative* (TN), *False Negative* (FN), *True Positive* (TP), dan *False Positive* (FP). Terakhir, kurva ROC merupakan salah satu kurva yang sering digunakan untuk menilai hasil klasifikasi atau prediksi dengan sumbu horizontal menggambarkan nilai negatif (*false positive*), sedangkan sumbu vertikal menggambarkan nilai positif (*true positive*) (Hadianto, dkk. 2019). Di dalam ROC, terdapat *Area Under Curve* (AUC) yang menunjukkan luas area di bawah kurva, semakin tinggi (mendekati 1) nilai luasnya maka akan semakin bagus hasil klasifikasi yang didapatkan.

## 2. Metode

Penelitian ini menggunakan data sekunder *Tehra Bank* yang didapatkan melalui *machine learning repository* mengenai profil nasabah bank dengan total 5000 data dengan 14 variabel, yaitu *ID customer*, *ZIP code*, *age*, *experience*, *income*, *family*, *CCAverage*, *Education*, *mortgage*, *securities account*, *CD account*, *online*, *credit card*, dan *personal loan* (Y). Proses eksperimen digunakan pada parameter masing-masing algoritma untuk melihat hubungan atau pengaruh perlakuan yang diberikan terhadap variabel penelitian tersebut (Iriadi & Nuraeni, 2016). Pada tahapan ini digunakan *pre-experimental design* yaitu dengan memberikan perlakuan (*treatment*) terhadap variabel, dan hasil yang didapatkan diukur, sehingga dapat diambil kesimpulan dari nilai yang dihasilkan.

Pada tahapan *pre-processing*, dua variabel yaitu *ID Customer* dan *ZIP Code* dikeluarkan karena dianggap tidak memberikan informasi signifikan yang berhubungan dengan tujuan penelitian. Kemudian, data dibersihkan dengan menghilangkan *missing value*, dan *redundant* pada data. Data di proses menggunakan bantuan *rapidminer software 9.6.000*. Digunakan operator-operator pada *rapidminer* untuk membantu dalam proses analisis, diantaranya operator *select attributes* yang digunakan untuk memilih variabel yang akan digunakan. Operator *split data* untuk menentukan rasio partisi antara data *training* dan data *testing*. Kemudian, operator *set role* digunakan untuk menentukan variabel terikat (*label*), dan dihubungkan dengan operator *cross validation* untuk melakukan proses *training* dan *testing*. Pada sub-bagian *training* menggunakan operator algoritma yang akan digunakan baik itu *naïve bayes*, *decision tree*, atau *neural network*. Lalu, pada sub-bagian *testing* digunakan operator *apply model* dan *performance* untuk mengukur tingkat akurasi model tersebut. Setelah itu, model dijalankan

**Tabel 1.** Parameter Algoritma *Neural Network*

Parameter	Nilai	Deskripsi
<i>Training cycles</i>	1000	Digunakan untuk mengetahui jumlah pengulangan <i>training</i> yang harus dilakukan untuk menghasilkan <i>error</i> terkecil dengan nilai mulai dari 1 sampai tak hingga (Badrul, 2016)
<i>Learning rate</i>	0.1	Parameter yang digunakan untuk menentukan bobot dari neuron dengan nilai berupa bilangan positif kurang dari 1
<i>Momentum</i>	0.28	Parameter untuk dapat mempercepat <i>learning rate</i> , meningkatkan <i>convergence</i> dengan nilai mulai dari 0 sampai 1 (Badrul, 2016)

**Tabel 2.** *Setting* Parameter Awal

Parameter	Nilai
<i>Training Cycle</i>	100 sampai 1000
<i>Learning rate</i>	0.1 sampai 1
<i>Momentum</i>	0.1 sampai 1
<i>Hidden Layer</i>	12 sampai 36 <i>node</i>

**Tabel 3.** Parameter Model *Decision Tree*

Parameter	Nilai
<i>Criterion</i>	<i>Gain Ratio</i>
<i>Maximal depth</i>	10
<i>Confidence</i>	0.2
<i>Minimal gain</i>	0.05
<i>Minimal Leaf size</i>	2
<i>Minimal size for split</i>	3
<i>Number of pre-pruning</i>	4
<i>Split ratio</i>	0.6 : 0.4

**Tabel 4.** Parameter Model *Naive Bayes*

Parameter	Nilai
<i>Laplace Correction</i>	Digunakan untuk menghindari terjadinya nilai 0 pada nilai probabilitas atribut yang digunakan yang dapat mempengaruhi hasil yang didapatkan
Partisi <i>split data</i>	60% ( <i>training</i> ) : 40% ( <i>testing</i> ) 70% ( <i>training</i> ) : 30% ( <i>testing</i> ) 80% ( <i>training</i> ) : 20% ( <i>testing</i> )

**Tabel 5.** *Split Data Naive Bayes*

Percobaan	<i>Training</i>	<i>Testing</i>	Akurasi (%)	Presisi (%)	<i>Recall</i> (%)	AUC
A	0.8	0.2	90.47	50.38	62.68	0.935
B	0.7	0.3	90.57	50.86	62.48	0.935
C	0.6	0.4	90.79	51.79	63.48	0.935

untuk mendapatkan nilai performansi hasil model yang telah digunakan. Akurasi hasil terlihat pada *performance vector*. Grafik hasil *training* dan *testing* data sebelumnya terlihat pada kurva *Receiver Operating Characteristic* (ROC) dengan melihat nilai *Area Under Curve* (AUC). Akurasi hasil juga terlihat pada *confusion matrix* yang dihasilkan yaitu nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN).

Parameter yang diubah sebagai pemberian perlakuan (*treatment*) pada algoritma *neural network* seperti terlihat pada **Tabel 1**. *Setting* parameter yang digunakan pada penelitian ini seperti pada **Tabel 2**. Tahapan proses untuk algoritma *naive bayes* dan *decision tree* hampir sama dengan *neural network*, parameter dan *setting* yang digunakan pada masing-masing algoritma berbeda. Pada algoritma *naive bayes* pertama dilakukan eksperimen pada perbandingan *split data training* dan *testing* yaitu dengan menggunakan tiga model perbandingan: 80% : 20%, 70% : 30%, dan

60% : 40%. Perbandingan yang menghasilkan nilai akurasi tertinggi akan dipilih untuk diaplikasikan pada keseluruhan proses. Parameter model *Naive Bayes* dapat dilihat pada **Tabel 4**. Sedangkan parameter yang digunakan pada algoritma *decision tree* dapat dilihat pada **Tabel 3**. Masing-masing parameter pada tiap algoritma akan disesuaikan berdasarkan hasil luaran yang didapatkan.

### 3. Hasil dan Pembahasan

Berikut hasil analisis yang diperoleh dari algoritma *Naive Bayes*, *Decision Tree*, dan *Neural Network*.

#### A. Algoritma *Naive Bayes*

Pada **Tabel 5**, dilakukan eksperimen dengan menggunakan tiga kombinasi *split data*, yaitu A dengan rasio *training* dan *testing* sebesar 0.8 : 0.2, B dengan rasio 0.7 : 0.3, dan C dengan rasio 0.6 : 0.4. Dapat dilihat bahwa C dengan rasio 0.6 (60%) : 0.4 (40%) menghasilkan nilai akurasi tertinggi, sehingga rasio

**Tabel 6. Performance Training Naïve Bayes**

	<i>True Tidak</i>	<i>True Ya</i>	<i>Class Precision (%)</i>
<b>Pred. Tidak</b>	22862	950	96.01
<b>Pred. Ya</b>	1537	1651	51.79
<b>Class Recall (%)</b>	93.70	63.48	
<b>Akurasi (%)</b>	90.79		

**Tabel 7. Performance Testing Naïve Bayes**

	<i>True Tidak</i>	<i>True Ya</i>	<i>Class Precision (%)</i>
<b>Pred. Tidak</b>	2541	106	96.00
<b>Pred. Ya</b>	170	183	51.84
<b>Class Recall (%)</b>	93.73	63.32	
<b>Akurasi (%)</b>	90.80		

**Tabel 8. Precision Naïve Bayes**

	<i>True Tidak</i>	<i>True Ya</i>	<i>Class Precision (%)</i>
<b>Pred. Tidak</b>	22862	950	96.01
<b>Pred. Ya</b>	1537	1651	51.79
<b>Class Recall (%)</b>	93.70	63.48	
<b>Akurasi (%)</b>	51.79		

**Tabel 9. Recall Naïve Bayes**

	<i>True Tidak</i>	<i>True Ya</i>	<i>Class Precision (%)</i>
<b>Pred. Tidak</b>	22862	950	96.01
<b>Pred. Ya</b>	1537	1651	51.79
<b>Class Recall (%)</b>	93.70	63.48	
<b>Recall (%)</b>	63.48 ( <i>Positive Class: Ya</i> )		

**Tabel 10. Perbandingan Hasil Akurasi pada Operator Split Ratio**

<i>Split Ratio</i>	<i>Akurasi (%)</i>	<i>RMSE</i>	<i>AUC</i>
0.6 : 0.4	99.22	0.077	0.997
0.7 : 0.3	99.14	0.084	0.997
0.8 : 0.2	98.38	0.113	0.989

tersebut yang diaplikasikan dalam model klasifikasi untuk memproses data. Pada umumnya, rasio yang digunakan oleh peneliti lain adalah 70% : 30%, tetapi hal tersebut tergantung kembali data yang digunakan. Rasio 60% : 40% dapat menghasilkan nilai akurasi tertinggi karena jumlah data *testing* yang digunakan dapat lebih merepresentasikan kondisi data yang digunakan dengan nilai 40%.

a. Performansi

Hasil performansi *training* dan *testing* setelah dilakukan pemrosesan pada data, dapat dilihat pada **Tabel 6**. Nilai akurasi yang dihasilkan sebesar 90.79%, dimana nilai tersebut memperlihatkan kedekatan antara nilai aktual dan prediksi yang dihasilkan, dimana semakin tinggi nilai performansi yang didapatkan, maka semakin bagus model tersebut diterapkan pada data. Begitupun pada **Tabel 7** yang menunjukkan hasil performansi *testing* yang dihasilkan.

b. Presisi dan *Recall*

Hasil presisi dan *recall* setelah dilakukan pemrosesan data ditunjukkan pada **Tabel 8** dan **9**. Pada **Tabel 8**, nilai presisi sebesar 51.79% menunjukkan nilai dimana data yang berhasil diprediksi “Ya” dengan benar dibandingkan dengan jumlah data yang diprediksi “Ya” tetapi

pada aktual terklasifikasi tidak melakukan pengajuan kredit. Sedangkan pada **Tabel 9**, nilai *recall* yang dihasilkan sebesar 63.48%, dimana hal tersebut menunjukkan bahwa data yang diprediksi “Ya” dengan benar dibandingkan dengan jumlah data yang terklasifikasi “Ya” namun dalam aktual nasabah tersebut tidak melakukan pengajuan kredit.

Terdapat *sample distribution* yang menampilkan distribusi pada kelas data yaitu “Ya” atau “Tidak”, masing-masing kelas memiliki 8 distribusi dengan kelas “Tidak” bernilai 0.904, dan “Ya” bernilai 0.096. Sedangkan nilai AUC yang dihasilkan sebesar 0.926. Semakin tinggi nilai AUC yang didapatkan, maka semakin bagus pula model tersebut diterapkan pada data yang digunakan.

**B. Algoritma Decision Tree**

Sebelumnya dilakukan percobaan dengan membandingkan menggunakan nilai *default* pada parameter *decision tree* untuk melihat pengaruh perbedaan nilai yang digunakan pada parameter *split ratio* untuk melihat perbandingan atau rasio yang menghasilkan nilai tertinggi untuk nantinya digunakan pada proses eksperimen selanjutnya. Nilai *default* pada parameter *decision tree* dapat dilihat pada **Tabel 10**.

**Tabel 11.** Hasil Parameter *Decision Tree*

No	A	B	C	D	E	F (%)	G
1	0.1	0.01	2	2	2	99.13	0.993
2	0.1	0.02	3	3	3	98.72	0.969
3	0.1	0.03	4	4	4	98.86	0.994
4	0.1	0.04	5	5	5	98.64	0.993
5	0.1	0.05	6	6	6	98.36	0.985
6	0.2	0.01	3	4	5	99.28	0.998
7	0.2	0.02	4	5	6	99.03	0.996
8	0.2	0.03	5	6	2	99.17	0.994
9	0.2	0.04	6	2	3	98.51	0.954
<b>10</b>	<b>0.2</b>	<b>0.05</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>99.36</b>	<b>0.999</b>
11	0.3	0.01	4	6	3	98.51	0.950
12	0.3	0.02	5	2	4	98.67	0.993
13	0.3	0.03	6	3	5	98.37	0.985
14	0.3	0.04	2	4	6	99.23	0.997
15	0.3	0.05	3	5	2	98.61	0.959
16	0.4	0.01	5	3	6	98.62	0.993
17	0.4	0.02	6	4	2	97.98	0.939
18	0.4	0.03	2	5	3	99.13	0.996
19	0.4	0.04	3	6	4	99.09	0.997
20	0.4	0.05	4	2	5	98.84	0.994
21	0.5	0.01	6	5	4	98.32	0.985
22	0.5	0.02	2	6	5	99.10	0.996
23	0.5	0.03	3	2	6	99.12	0.997
24	0.5	0.04	4	3	2	98.48	0.950
25	0.5	0.05	5	4	3	98.36	0.949

Keterangan:

A : *Confidence*

C : *Leaf Size*

E : *Number preprunning*

G : *AUC (Area Under Curve)*

B : *Minimal Gain*

D : *Size for Split*

F : *Akurasi*

**Tabel 12.** *Performance Training Decision Tree*

	<i>True Tidak</i>	<i>True Ya</i>	<i>Class Precision (%)</i>
<b>Pred. Tidak</b>	24382	156	99.36
<b>Pred. Ya</b>	17	2445	99.31
<b>Class Recall (%)</b>	99.93	94.00	
<b>Akurasi (%)</b>	99.36		

Dapat dilihat bahwa percobaan dengan rasio sebesar 0.6 untuk *training* dan 0.4 *testing* menghasilkan nilai akurasi tertinggi sebesar 99.22%. Selain itu, RMSE yang dihasilkan memberikan nilai terendah diantara ketiga percobaan yang dilakukan. Kemudian, nilai AUC sebesar 0.997 menunjukkan bahwa rasio tersebut menghasilkan model klasifikasi yang sangat baik yaitu dalam rentang 0.9 – 1.0. Nilai RMSE menunjukkan besaran nilai *error* yang didapatkan dari hasil prediksi. Semakin kecil nilai RMSE yang dihasilkan, maka semakin baik atau akurat hasil prediksi yang didapatkan dibandingkan dengan nilai aktualnya. Untuk itu, pada penentuan parameter lainnya, akan digunakan nilai *split ratio* yaitu sebesar 0.6. Kemudian, perlakuan (*treatment*) dilakukan pada lima parameter yaitu *minimal gain*, *leaf size*, *size for split*, *number preprunning*, dan *confidence*. Hasil yang didapatkan dapat dilihat pada **Tabel 11**.

*Minimal gain* merupakan nilai bobot pada setiap *node* yang dimiliki. *Node* akan melakukan percabangan jika nilai *gain* yang dimiliki lebih besar dari nilai yang ditetapkan pada *setting* parameter. Kedua, *Minimal leaf*

*size*, yaitu jumlah contoh (*examples*) yang ada pada subsetnya. Maka setiap *tree* memiliki *leaf* dengan jumlah *leaf size* yang telah ditentukan oleh *setting* parameter. Ketiga, *Minimal Size for Split*, yaitu ketika *node* akan melakukan *splitting* ketika *size* yang dimiliki lebih dari atau sama dengan *minimal size for split* yang telah di tentukan pada *setting* parameter. Terakhir, *number of preprunning*, yaitu ketika *splitting* pada *tree* dibatasi untuk menghindari terjadinya *splitting* pada *node* yang tidak memberikan nilai tambah pada proses pengambilan keputusan menggunakan *tree* tersebut.

Pada **Tabel 11** dilakukan 25 kombinasi percobaan yang dihasilkan dengan bantuan *software* Minitab 19 dengan 5 level dan 5 faktor.

a. Performansi

Hasil performansi *training* dan *testing* setelah dilakukan pemrosesan pada data dapat dilihat pada **Tabel 12**. Nilai performansi *training* dan *testing* pada **Tabel 12** dan **Tabel 13**, dengan masing-masing sebesar 99.36% dan 97.90%, menunjukkan nilai kedekatan antara nilai yang diprediksi dan nilai aktual yang dihasilkan.

**Tabel 13. Performance Testing Decision Tree**

	True Tidak	True Ya	Class Precision (%)
Pred. Tidak	2692	44	98.39
Pred. Ya	19	245	92.80
Class Recall (%)	99.30	84.78	
Akurasi (%)	97.90		

**Tabel 14. Precision Testing Decision Tree**

	True Tidak	True Ya	Class Precision (%)
Pred. Tidak	24382	156	99.36
Pred. Ya	17	2445	99.31
Class Recall (%)	99.93	94.00	
Akurasi (%)	99.32		

**Tabel 15. Recall Decision Tree**

	True Tidak	True Ya	Class Precision (%)
Pred. Tidak	24382	156	99.36
Pred. Ya	17	2445	99.31
Class Recall (%)	99.93	94.00	
Recall (%)	94.00		



**Gambar 1. Hasil Decision Tree**

Nilai tersebut menunjukkan bahwa model *decision tree* yang digunakan dapat dengan baik melakukan pemrosesan pada data yang digunakan. Semakin tinggi persentase nilai performansi yang dihasilkan, maka semakin bagus model tersebut untuk digunakan.

b. Presisi dan Recall

Hasil presisi dan *recall* yang didapatkan setelah dilakukan pemrosesan pada data dapat dilihat pada **Tabel 14** dan **15**. Pada **Tabel 14** dan **15**, dimana masing-masing menunjukkan nilai presisi dan *recall*. Nilai presisi merepresentasikan bahwa data telah memberikan informasi dengan tepat sesuai yang diinginkan, serta hasil yang dihasilkan relevan dengan yang ditunjukkan oleh sistem.

Sedangkan untuk nilai AUC yang didapatkan sebesar 0.999, dimana hasil tersebut termasuk dalam klasifikasi yang sangat baik. Nilai tersebut menyatakan bahwa model yang digunakan sangat baik untuk proses klasifikasi.

**Gambar 1** menampilkan hasil luaran *tree* yang didapatkan.

**C. Algoritma Neural Network**

Penentuan parameter awal dilakukan dengan menggunakan operator *optimize parameter* yang ditampilkan pada **Tabel 16**, dan nilai terbaik dihasilkan pada iterasi ke-10 dengan akurasi sebesar 99.28 %. Pada **Tabel 16** dapat dilihat bahwa dilakukan 10 kali kombinasi percobaan dengan nilai mengacu pada nilai parameter *training cycle* dengan kelipatan 100 yang dimulai dari nilai 100, 200, ....., 1000. Sedangkan

**Tabel 16.** Penentuan Nilai Paramater

No	A	B	C	D (%)	E (%)	F
1	100	0.1	0.64 = 0.6	98.83	98.18	0.978
2	200	0.46 = 0.5	0.82 = 0.8	98.85	98.00	0.976
3	300	0.1	0.46 = 0.5	99.07	97.92	0.978
4	400	0.28 = 0.3	0.1	99.11	97.90	0.978
5	500	0.64 = 0.6	0.28 = 0.3	99.18	97.96	0.979
6	600	0.28 = 0.3	0.1	99.20	98.00	0.979
7	700	0.28 = 0.3	0.1	99.21	98.06	0.979
8	800	0.28 = 0.3	0.1	99.20	97.96	0.980
9	900	1	0.1	99.24	98.02	0.976
10	1000	0.1	0.28 = 0.3	99.28	98.00	0.980

Keterangan:

A : *Training Cycle*

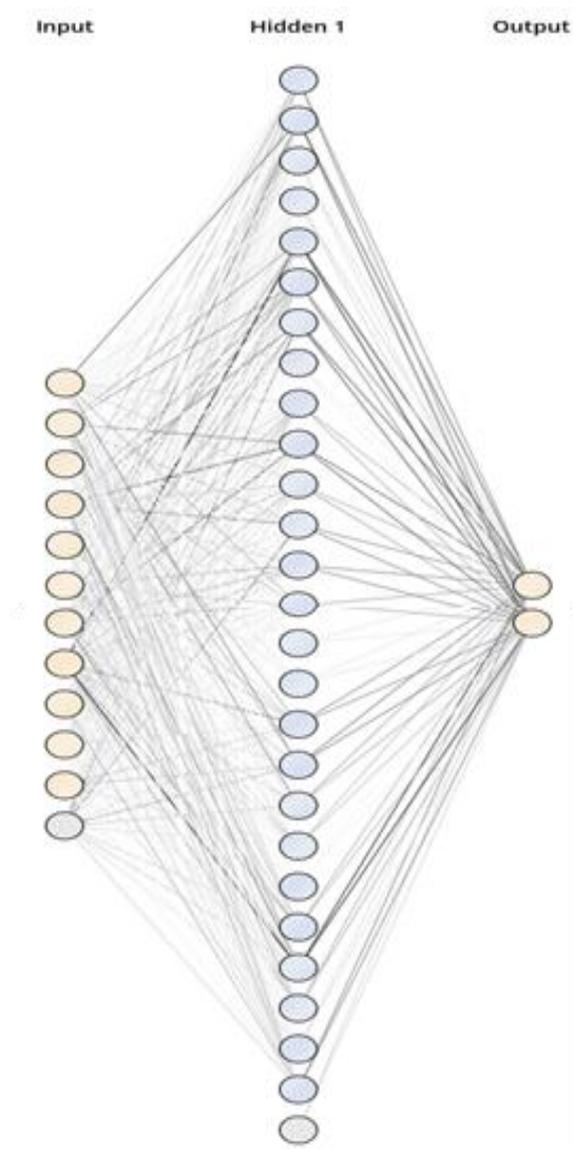
B : *Learning rate*

C : *Momentum*

D : *Performance training*

E : *Performance Testing*

F : *AUC (Area Under Curve)*



**Gambar 2.** Hasil ANN

parameter lain dihasilkan dari nilai kombiansi yang didapatkan dengan bantuan operator *optimize parameter*. Operator *optimize parameter* untuk menentukan nilai akurasi tertinggi ditentukan dengan melakukan 36 iterasi pada tiga parameter yaitu *training cycles*, *learning rate*, dan *momentum*. **Gambar 2**

menunjukkan struktur luaran yang dihasilkan dari model *neural network*.

Pada struktur ANN yang dibentuk, terdapat satu *hidden layer* (lapisan tersembunyi). Jumlah tersebut dipilih karena pada umumnya, dalam penelitian jumlah yang *hidden layer* yang digunakan adalah satu *layer* (Larose & Larose, 2014), dan disamping itu belum ada

**Tabel 17.** Penentuan Jumlah *Hidden Layer Model*

No.	Hidden Layer Size	Perform. Training (%)	Perform. Testing (%)	AUC
1	12	99.50	97.86	0.976
2	14	99.56	97.92	0.983
3	16	99.56	98.22	0.983
4	18	99.57	98.12	0.982
5	20	99.60	98.30	0.982
6	22	99.56	98.14	0.981
7	24	99.59	98.04	0.983
8	26	99.61	98.26	0.983
9	28	99.59	98.14	0.984
10	30	99.60	98.18	0.983
11	32	99.58	98.26	0.982
12	34	99.59	98.20	0.982
13	36	99.58	98.14	0.983

**Tabel 18.** Performance Training Artificial Neural Network

	True Tidak	True Ya	Class Precision (%)
Pred. Tidak	40666	169	99.59
Pred. Ya	5	4160	99.88
Class Recall (%)	99.99	96.10	
Akurasi (%)	99.61 +/- 0.06 (micro average: 99.61%)		

**Tabel 19.** Performance Testing Artificial Neural Network

	True Tidak	True Ya	Class Precision (%)
Pred. Tidak	4492	60	98.68
Pred. Ya	27	421	93.97
Class Recall (%)	99.40	87.53	
Akurasi (%)	98.26 +/- 0.39 (micro average: 98.26%)		

**Tabel 20.** Presisi Artificial Neural Network

	True Tidak	True Ya	Class Precision (%)
Pred. Tidak	40666	60	98.68
Pred. Ya	5	421	93.97
Class Recall (%)	99.99	87.53	
Akurasi (%)	99.88 +/- 0.17 (micro average: 99.88%) (positive class: 2)		

penelitian yang menetapkan jumlah *hidden layer* lebih dari satu. Sedangkan untuk jumlah *node* atau *hidden layer size* yang berada dalam *layer* tersebut berjumlah 36, dimana maksimal *node* yang digunakan adalah 3 kali jumlah variabel *input* nya. Oleh karena itu, jumlah minimum *node* yang digunakan berkelipatan 2 yaitu 12 (minimal) – 14 -16, ..., - 36 *node* (maksimal).

Penelitian ini menggunakan 12 variabel, dimana jumlah tersebut menjadi jumlah *node* yang digunakan pada *input layer* (lapisan masukan). Jumlah keseluruhan untuk masing-masing *layer* sesuai dengan hasil performansi terbaik pada **Tabel 17** adalah *input layer* = 12 *node*, *hidden layer* = 26, dan *output layer* = 2.

#### a. Performansi

Hasil performansi yang didapatkan setelah diproses dapat dilihat pada **Tabel 18**. Nilai performansi *training* dan *testing* pada **Tabel 18** dan **19** menunjukkan kedekatan hasil data pada nilai aktual dan nilai prediksi yang didapatkan. Selain itu, hal tersebut menunjukkan bahwa

model yang digunakan dapat dengan baik memberikan informasi yang sesuai dengan yang diinginkan, serta hasil yang didapatkan relevan sesuai dengan yang ditujukan pada sistem. Selain itu, dapat dilihat bahwa nilai *recall* yang dihasilkan lebih rendah dari presisi yang dihasilkan. Hal tersebut mengindikasikan bahwa *false positive* dihindari terjadi daripada *true positive*, sehingga pada penelitian ini akan lebih baik jika nasabah yang melakukan pengajuan kredit tetapi saat diprediksi nasabah tersebut tidak mengajukan, daripada nasabah tersebut tidak mengajukan namun dalam hasil prediksi menunjukkan bahwa nasabah tersebut akan mengajukan kredit. Oleh karena itu, penambahan biaya operasional sangat dihindari, dan lebih baik jika bank kehilangan nasabah potensial.

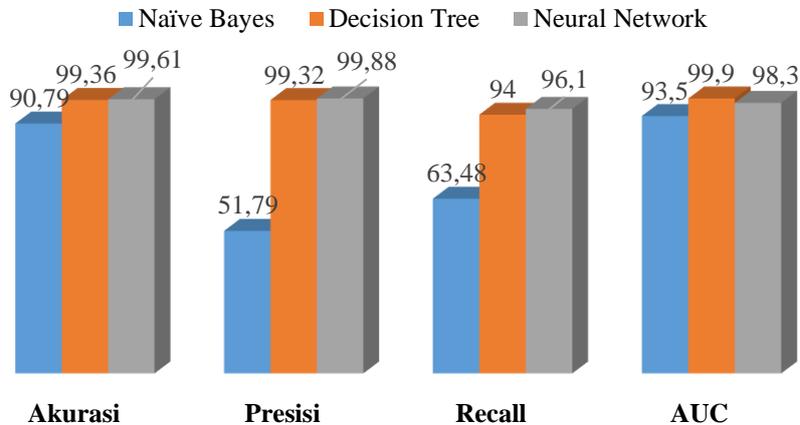
#### b. Presisi dan Recall

Hasil presisi dan *recall* yang didapatkan setelah diproses. Kemudian, pada **Tabel 20** dan **Tabel 21** dihasilkan nilai presisi dan *recall*,

**Tabel 21.** Recall Artificial Neural Network

	True Tidak	True Ya	Class Precision (%)
Pred. Tidak	40666	169	99.59
Pred. Ya	5	4160	99.88
Class Recall (%)	99.99	96.10	
Recall (%)	96.10 +/- 0.49 (micro average: 96,10%) (positive class: 2)		

**Komparasi Hasil Tiga Algoritma**



**Gambar 3.** Komparasi Hasil Tiga Algoritma

dimana masing-masing merepresentasikan bahwa data yang diberikan telah sesuai dengan informasi yang diminta. Berikut komparasi hasil tiga algoritma yang telah dihasilkan dapat dilihat pada **Gambar 3**. Pada gambar tersebut menunjukkan algoritma ANN menghasilkan nilai tertinggi dibandingkan dua algoritma yang lain, dilihat dari histogram yang dihasilkan.

Penggunaan operator *optimize parameter* pada beberapa algoritma bertujuan untuk dapat meningkatkan nilai persentase performansi model (Tantithamthayorn, 2018). Pada ANN sebagai algoritma dengan hasil performansi terbaik, *training cycle* menghasilkan nilai akurasi yang semakin rendah jika nilainya semakin tinggi, dimana *training cycle* ini digunakan untuk mengurangi *error* pada data, sehingga semakin banyak perulangan yang dilakukan *error* yang dihasilkan akan semakin menurun (Badrul, 2016). Pada parameter *learning rate*, jika nilainya semakin besar proses yang dilakukan akan berjalan lebih cepat namun akan terjadinya osilasi pada bobot dan akan menurunkan nilai akurasi yang didapatkan (Badrul, 2016). Sebaliknya, jika nilai *learning rate* yang diberikan semakin rendah, akurasi yang didapat akan semakin tinggi. Kemudian, untuk parameter *momentum* digunakan untuk dapat mempersingkat waktu proses pembelajaran (*learning*) dan mengurangi osilasi bobot (Badrul, 2016). Biasanya nilai *learning rate* akan berbanding terbalik dengan nilai *momentum* untuk bisa mendapatkan nilai akurasi terbaik.

Pada algoritma *decision tree* terdapat beberapa perubahan nilai akurasi dari perlakuan (*treatment*) yang diberikan. Pertama, saat *confidence* berada pada nilai paling minimum yaitu 0.1, *minimal gain* meningkat, sedangkan tiga parameter lain bernilai setara, semakin

tinggi nilai pada ketiga parameter tersebut, maka semakin rendah akurasi yang didapatkan. Kemudian, jika nilai parameter *confidence* naik, maka nilai akurasi cenderung akan semakin menurun. Cenderung saat *confidence* semakin besar, maka *node* yang memiliki *error* tinggi dapat ditoleransi untuk digunakan dalam proses membangun *tree*. Selain itu, jika nilai pada parameter *number of prepruning* meningkat, dengan nilai *minimal leaf size* yang lebih kecil, cenderung nilai akurasi yang dihasilkan yang semakin tinggi.

*Prunning* digunakan untuk memangkas jumlah cabang yang tidak memberikan informasi terhadap target variabel, sehingga nilai akurasi akan semakin meningkat (Ahlemeyer-Stubbe & Coleman, 2014). Keempat, jika nilai pada parameter *leaf size* dinaikkan, maka akurasi yang dihasilkan akan sebaliknya yaitu mengalami penurunan. Semakin besar nilai yang diberikan pada parameter *leaf size*, maka akan semakin besar pula jumlah *leaf* yang dibentuk pada *tree*. *Leaf size* yang terlalu besar akan menghambat dalam proses pembentukan *tree* yang dilakukan.

Terakhir, algoritma *naïve bayes* menghasilkan nilai akurasi yang cukup tinggi pula yaitu sebesar 90.79%, presisi = 51.79%, *recall* = 63.48%, dan AUC = 0.935. Hasil tersebut menunjukkan bahwa *naïve bayes* telah bekerja sangat baik dalam mengklasifikasikan profil nasabah bank, walaupun nilainya masih di bawah dua algoritma yang lain yaitu ANN dan *decision tree*. Secara keseluruhan terdapat tujuh faktor *importance* yang mempengaruhi hasil penelitian yaitu *mortgage*, *income*, *family*, *CD account*, *CC average*, *education*, dan *age*. Pertama, *mortgage*, semakin tinggi nilai *mortgage* yang dimiliki maka akan semakin tinggi kecenderungan nasabah untuk mengajukan kredit dan taat dalam pembayaran angsuran setiap bulannya untuk menghindari penyitaan

yang akan dilakukan oleh bank jika seluruh kredit yang diberikan tidak dibayarkan (Supriyadi & Kartikasari, 2015). Pada kondisi saat ini, kebutuhan akan *mortgage* semakin tinggi dengan berbagai persyaratan yang dimiliki, sehingga tidak menutup kemungkinan bahwa tingkat ajuan kredit juga akan berbanding lurus dengan hal tersebut.

Kedua, *income*, semakin tinggi pendapatan seseorang maka semakin tinggi pula tingkat konsumsi yang akan dibutuhkan, dan hal tersebut memiliki pengaruh yang signifikan terhadap kecenderungan nasabah untuk mengajukan kredit (Aris & Syechalad, 2018; Riyanto, 2011; Pranata & Widarno, 2014; Sandria, dkk. 2016). Seiring dengan waktu, kebutuhan manusia akan barang-barang kebutuhan sehari-hari baik itu barang primer, sekunder, ataupun tersier akan terus meningkat. Tetapi peningkatan tersebut tidak berbanding lurus dengan tingkat pendapatan yang dihasilkan. Untuk dapat memenuhi seluruh kebutuhan yang dimiliki, maka salah satu cara yang dapat dilakukan adalah dengan mengajukan kredit pinjaman. Sehingga, saat ini banyak sekali masyarakat yang mengajukan kredit untuk usaha, ataupun kebutuhan lainnya.

Ketiga, *family* (keluarga), semakin banyak jumlah tanggungan yang dimiliki maka pengeluaran nasabah tersebut akan semakin tinggi (Marantika & Sampurno, 2013). Pendapatan rumah tangga belum dapat menutupi jumlah pengeluaran kebutuhannya, dan tidak semua anggota keluarga dapat menghasilkan pendapatan. Untuk itu, pengajuan kredit merupakan salah satu jalan yang dapat ditempuh untuk dapat memenuhi kebutuhan sehari-hari yang dimiliki.

Keempat, *CD account (Certificate of Deposit Account)*, dimana jika nasabah yang memiliki memiliki *CD account* cenderung akan mengajukan kredit pinjaman, dan sebaliknya, jika nasabah tersebut tidak memiliki, maka cenderung tidak akan mengajukan kredit. Sejalan dengan penelitian Tipa & Purba (2018) yang mengatakan bahwa *CD account* yang merupakan salah satu jaminan memiliki pengaruh signifikan terhadap keputusan pengajuan kredit. Kelima, *CC average (Credit Card Average)*, dimana nasabah yang memiliki *CC average* yang besar memiliki kecenderungan untuk mengajukan kredit, dan jika nilainya rendah maka nasabah akan cenderung tidak akan mengajukan kredit. Penggunaan kartu kredit (*credit card*) merupakan hal yang umum terjadi pada saat ini. Namun, dari penggunaan tersebut ternyata masih banyak yang belum dapat memenuhi tagihannya. Untuk itu, tidak jarang jika nasabah pengguna kartu kredit, juga mengajukan kredit pinjaman di bank untuk memenuhi tagihan yang mereka miliki.

Keenam, *education* (pendidikan). Nasabah yang berada pada kelompok *undergraduate* cenderung untuk tidak mengajukan kredit karena kebutuhan yang diperlukan masih rendah, sehingga pemenuhan dengan dana tambahan tidak begitu dibutuhkan. Namun lainnya halnya dengan kelompok nasabah *graduate* dan *professional*, cenderung akan mengajukan kredit karena didukung dengan pengalaman serta pengetahuan terhadap berbagai risiko dan prosedur dalam pengajuan kredit (Rahman & Widyarti, 2017).

Hal tersebut sejalan pula dengan penelitian Suprpto & Puryandi (2020), menyatakan bahwa tingkat pendidikan memiliki pengaruh signifikan terhadap pengajuan atau pengambilan kredit. Setelah lulus, umumnya kebutuhan akan barang-barang akan semakin meningkat, sejalan dengan adanya kegiatan baru yang harus dilakukan, dimana kegiatan tersebut juga membutuhkan dana tambahan. Sehingga, kredit merupakan salah satu pengajuan yang dapat dilakukan untuk dapat membantu dalam memenuhi kebutuhan tersebut.

Terakhir, *age* (umur). Nasabah yang berada pada umur antara 23 – 43 tahun, akan cenderung mengajukan kredit, sedangkan yang berada pada umur antara 28 – 59 tahun cenderung tidak akan mengajukan pinjaman. Nasabah pada kelompok usia relatif muda menyadari jika usia yang dimiliki termasuk usia produktif dan dapat dijadikan sebagai salah satu jaminan dalam pengajuan kredit, karena tingkat peluang pengembaliannya masih tinggi (Pratiwi, 2015). Jika umur yang dimiliki nasabah lebih dari angka tersebut, maka dapat dinyatakan nasabah tersebut tidak dalam usia produktif dan kecenderungannya adalah tingkat produktifitas yang dimiliki akan semakin menurun, sehingga berdampak pula pada penghasilan yang didapatkan. Sehingga, ketujuh faktor yang dihasilkan, dapat dijadikan sebagai pertimbangan bagi bank untuk memberikan kredit dan penawaran kredit pula.

#### 4. Kesimpulan

Hasil luaran yang didapatkan dari ketiga algoritma yang digunakan yaitu *naïve bayes*, *decision tree*, dan *artificial neural network* menunjukkan bahwa ANN memberikan luaran terbaik dengan akurasi sebesar 99.61%. Disusul oleh *decision tree* dengan akurasi sebesar 99.36%, dan *naïve bayes* sebesar 90.79%. ANN sangat baik dalam melakukan klasifikasi data profil nasabah, namun *naïve bayes* dan *decision tree* tidak kalah baik dibuktikan dengan hasil luaran performansi yang dihasilkan dengan nilai rata-rata di atas 90%. Nasabah yang berpeluang mengajukan kredit memiliki kriteria, yaitu  $CCAvg \leq 9$ ,  $mortgage > 565$ ,  $income$  antara 108–190 *dollar* setiap bulannya,  $CC average = 3 - 9$  USD,  $mortgage > 565$  USD,  $income$  antara 101–116 atau lebih dari 116 USD setiap bulannya, memiliki *CD Account*, *age* (umur) diantara 23 - 43 tahun, dan *family* (keluarga) berjumlah lebih dari  $\geq 3$  anggota,  $education = undergraduate$  dan *graduate*.

Sedangkan nasabah yang berpeluang tidak mengajukan kredit memiliki kriteria, yaitu  $CC average = 2 - 9$  USD,  $mortgage \leq 565$ ,  $income$  antara 101 -114 per bulan,  $age = 28 - 59$  tahun, tidak memiliki *CD Account*,  $family = 3 - 4$  anggota,  $education = graduate$  dan *professional*. Terdapat 6 faktor penting yang mempengaruhi luaran yang dihasilkan ketiga model algoritma tersebut yaitu *income*, *CC Average*, *CD Account*, *mortgage*, *family*, dan *education*.

## 5. Daftar Pustaka

- Ahlemeyer-Stubbe, A., & Coleman, S. (2014). A Practical Guide to Data Mining for Business And Industry. *John Wiley & Sons*.
- Aris, M., & Syechalad, M. N. (2018). Pengaruh Pola Konsumsi Terhadap Permintaan Kredit Konsumtif di Banda Aceh. *Jurnal Ilmiah Mahasiswa Ekonomi Pembangunan*, 3(1), 100-109. Dari <http://www.jim.unsyiah.ac.id/EKP/article/view/6886>.
- Badrul, M. (2016). Optimasi Neural Network Dengan Algoritma Genetika Untuk Prediksi Hasil Pemilukada. *Bina Insani ICT Journal*, 3(1), 229-242.
- Fiska, R. R. (2017). Penerapan Teknik Data Mining dengan Metode Support Vector Machine (SVM) untuk Memprediksi Siswa yang Berpeluang Drop Out (Studi Kasus di SMKN 1 Sutera). *SATIN-Sains dan Teknologi Informasi*, 3(1), 15-23. DOI: 10.33372/stn.v3i1.200.
- Hadianto, N., Novitasari, H. B., & Rahmawati, A. (2019). Klasifikasi Peminjaman Nasabah Bank Menggunakan Metode *Neural Network*. *Pilar Nusa Mandiri: Journal of Computing and Information System*, 15(2), 163-170. DOI: 10.33480/pilar.v15i2.658.
- Hidayah, N. (2019). Sistem Klasifikasi Penerima Beras Miskin Menggunakan Algoritma *Decision Tree C4.5*. Tesis. *Yogyakarta: Fakultas Teknologi Informasi dan Elektro Universitas Teknologi Yogyakarta*.
- Honesqi, H. D. (2017). Klasifikasi *Data Mining* Untuk Menentukan Tingkat Persetujuan Kartu Kredit. *Jurnal Teknoif ITP*, 5(2), 57-62. DOI: 10.21063/jtif.2017.v5.2.57-62.
- Iriadi, N., & Nuraeni, N. (2016). Kajian Penerapan Metode Klasifikasi *Data Mining* Algoritma C4.5 Untuk Prediksi Kelayakan Kredit Pada Bank Mayapada Jakarta. *Jurnal Teknik Komputer*, 2(1), 132-137. Dari <http://ejournal.bsi.ac.id/ejournal/index.php/jtk/article/view/371>.
- Ivandari. (2015). Model Pohon Keputusan untuk Klasifikasi Persetujuan Kredit. *Jurnal Litbang Kota Pekalongan*, 9, 124-136.
- Larasati, A., Farhan, M., Rahmawati, P., Azzahra, N., Hajji, A. M., & Handayani, A. N. (2019, November). Designing Classification Models of Patron Visits to an Academic Library using Decision Tree. In *2019 1st International Conference on Engineering and Management in Industrial System (ICOEMIS 2019)*, 139-145.
- Larasati, A., Hajji, A. M., & Dwiastuti, A. (2019, May). The Relationship Between Data Skewness and Accuracy of Artificial Neural Network Predictive Model. In *IOP Conference Series: Materials Science and Engineering*, 523(1), 012070. *IOP Publishing*. DOI: 10.1088/1757-899X/523/1/012070.
- Larose, D. T., & Larose, C. D. (2014). Discovering Knowledge in Data: An Introduction to Data Mining (Vol. 4). *John Wiley & Sons*.
- Marantika, C. R. & Sampurno, R. D. (2013). Analisis Faktor-Faktor Yang Mempengaruhi Kelancaran Pengembalian Kredit Usaha Rakyat (KUR) Mikro. *Diponegoro Journal of Management*, 2(2), 1-14. Dari <http://ejournal-s1.undip.ac.id/index.php/dbr>.
- Nuraeni, N. (2017). Penentuan Kelayakan Kredit dengan Algoritma Naive Bayes Classifier: Studi Kasus Bank Mayapada Mitra Usaha Cabang PGC. *Jurnal Teknik Komputer AMIK BSI*, 3(1), 9-15. Dari <https://ejournal.bsi.ac.id/ejournal/index.php/jtk/article/view/1337>.
- Pandie, E.S.Y. (2018). Implementasi Algoritma *Data Mining Naive Bayes* Pada Koperasi. *Jurnal Icon*, 6(1), 15-20. Dari <http://ejournal.undana.ac.id/jicon/article/view/350>.
- Pemerintah Indonesia. (1998). Undang-Undang Republik Indonesia Nomor 10 Tahun 1998 tentang Perubahan atas Undang-Undang Nomor 7 Tahun 1992 tentang Perbankan. Lembaran RI Tahun 1998 No.182. *BPHN* (online), (<http://www.bphn.go.id>), diakses 7 Nopember 2020.
- Pranata, G. & Widarno, B. (2014). Pengaruh Tingkat Suku Bunga, Pendapatan, Status Pekerjaan, Jangka Waktu Kredit dan Tingkat Pendidikan Terhadap Jumlah Pengambilan Kredit. *Jurnal Akuntansi dan Sistem Teknologi Informasi*, 10(2), 151-165.
- Pratiwi, I. A. (2015). Determinasi Permintaan Kredit Usaha Rakyat (KUR) PT Bank Rakyat Indonesia (Studi Pada Usaha Mikro Kecil dan Menengah Kabupaten Jombang). Skripsi. *Malang: FEB Universitas Brawijaya*.
- Rahman, M. I. & Widyarti, E. T. (2017). Analisis Pengaruh Tingkat Pendapatan, Tingkat Pendidikan, Suku Bunga, Penetrasi Demografis dan Geografis Perbankan Terhadap Total Kredit UMKM (Studi Kasus Kota dan Kabupaten di Jawa Tengah Tahun 2011-2015). *Diponegoro Journal of Management*, 6(2), 1-14. Dari <http://ejournal-s1.undip.ac.id/index.php/management>.
- Rani, N. L. (2015). Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit. *Jurnal KomTekInfo Fak. Ilmu Komputer*, 2(2), 33-38. DOI: 10.33330/jurteksi.v4i1.30.
- Riyanto, N. E. (2011). Pengaruh Tingkat Pendidikan Dan Pendapatan Terhadap Keputusan Mengambil Kredit (Studi Kasus Pada Anggota KPRI Bhakti Wuluhan Kabupaten Jember Tahun 2010). Skripsi. *Jember: FKIP UNEJ*.
- Sandria, D, Adnan, N. & Yuliana, S. (2016). Analisis Faktor Yang Mempengaruhi Permintaan Kredit Pemilikan Rumah (KPR) di Kota Palembang: Kasus Nasabah. *Jurnal Economic*

- Pembangunan*, 14(2), 54-58. Dari <https://ejournal.unsri.ac.id/index.php/jep/index>.
- Sari, I. U. (2015). Analisis Dengan Metode Klasifikasi Menggunakan *Decision Tree* Untuk Prediksi Penentuan Resiko Kredit Pada Bank Bukopin BATAM. *CBIS Journal*, 3(1), 67-78. Dari <http://ejournal.upbatam.ac.id/index.php/cbis/article/view/387>.
- Suprpto, E. & Puryandani, S. (2020). Analisis Pengaruh Kualitas Layanan, Suku Bunga, Pendapatan, Dan Pendidikan Terhadap Keputusan Nasabah Mengambil Kredit Produktif Di Bank Jateng Capem Margasari. *Magisma Jurnal Ilmiah Ekonomi dan Bisnis*, 8(1), 73-79. DOI: 10.35829/magisma.v1i1.71.
- Supriyadi, A. & Kartikasari, D. (2015). Hubungan Analisis Kredit Terhadap Kualitas Pembiayaan KPR pada PT XYZ Cabang Batam. *Jurnal Akuntansi Ekonomi dan Manajemen Bisnis*, 3(2), 128-134
- Syafnur, A. (2017). Analisis Dengan Metode Klasifikasi Menggunakan *Decision Tree* untuk Memprediksi Penentuan Resiko Kredit Bank. *Jurnal Teknologi dan Sistem Informasi*, 4(1), 101-106.
- Tantithamthavorn, C., Mcintosh, A., Hassan, A. E., Member, S., & Matsumoto, K. (2018). The Impact of Automated Parameter Optimization on Defect Prediction Models. *IEEE Transaction on Software Engineering*, 5589 (1), 32. DOI: 10.1109/TSE.2018.2794977.
- Tipa, H. & Purba, M. A. (2018). Analisis Faktor-Faktor Yang Mempengaruhi Pengambilan Kredit oleh Pelaku Usaha Properti di Kota Batam. *Jurnal Akuntansi Bareleng*, 3(1), 49-56. DOI: <https://doi.org/10.33884/jab.v3i1.1612>.