
Pemodelan Sistem Prediksi Kelayakan Pengajuan Kredit Kepemilikan Rumah Dengan Metode C4.5 Dan Naive Bayes

Lingga Desyanita¹, Arief Wibowo²

¹Fakultas Teknologi Informasi, Magister Ilmu Komputer, Universitas Budi Luhur
Jalan Cileduk Raya Petukangan Utara Jakarta Selatan DKI Jakarta, e-mail:
linggadesyaniitaa@gmail.com

²Fakultas Teknologi Informasi, Magister Ilmu Komputer, Universitas Budi Luhur
Jalan Cileduk Raya Petukangan Utara Jakarta Selatan DKI Jakarta, e-mail:
arief.wibowo@budiluhur.ac.id

ARTICLE INFO

Article history: Received 24 Oktober 2020
Received in revised form 2 November 2020
Accepted 10 November 2020 Available online
10 Desember 2020

ABSTRACT

A house for every human being is the main and most important need compared to others needs in general. A financial institution is an institution engaged in the financial sector where its customers are people from various walks of life with various behaviors. Lending is a business activity that carries a high risk and affects the business continuity of a banking company. The problem that is often faced in providing home loans is determining the decision to extend credit to prospective customers, while another problem is that not all home loan payments by customers can run well or commonly known as bad credit. One of the causes of bad credit is an assessment error in making credit decisions. Data mining is a process used to analyze cases in order to find the best performance of an algorithm being tested. One way to get information or patterns from a large data set is to use techniques in data mining. There are many classification methods that can be used to produce precise accuracy values. In this study, two classification algorithm methods are used in classifying the home crediting dataset, namely the C4.5 decision tree algorithm and the Naïve Bayes algorithm. The comparison of the two algorithms produces an accuracy value for the Naïve Bayes algorithm of 36.36% and the Decision Tree C4.5 algorithm has an accuracy rate of 59.54%.

Keywords: Data Mining, Classification, Naïve Bayes, C4.5, application for credit

1. Pendahuluan

Rumah merupakan salah satu kebutuhan paling penting dan utama bagi setiap manusia. Badan usaha yang berfungsi untuk menghimpun dana dari masyarakat dalam bentuk simpanan dan menyalurkannya kepada masyarakat berupa kredit adalah Bank. Kegiatan usaha memiliki tingkat resiko tinggi dan berpengaruh terhadap keberlangsungan usaha suatu perusahaan perbankan disebut dengan pemberian kredit.

Suatu keadaan dimana seorang nasabah sudah tidak mampu membayar sebagian hingga seluruh kewajibannya kepada bank sesuai dengan yang telah dijanjikan ataupun disepakati sejak awal disebut kredit macet. Adapun faktor penyebab terjadinya kredit macet antara lain terjadinya

kesalahan dalam penilaian dalam keputusan kredit, terlambatnya pembayaran angsuran dari waktu yang telah ditentukan, bertambahnya angsuran yang dimiliki setiap nasabah dan juga penggunaan overlimit.

Berdasarkan data tahun 2015 sampai dengan 2020. Berdasarkan gambar 1.2 menunjukkan bahwa kenaikan kredit macet terjadi setiap tahunnya, kenaikan npl tertinggi terjadi di tahun 2017 sebesar 4.68% dari seluruh Kredit Pemilikan Rumah pada tahun 2017 menyandang status kredit macet sebagai faktor utama. Lemahnya pengawasan dalam proses pemberian kredit dan tahap pelaksanaannya dapat meningkatkan presentasi terjadinya kredit macet. Oleh karena itu, harus dilakukan pengurangan tingkat resiko dengan meningkatkan kehati-hatian dan memperhatikan berbagai faktor yang bisa menyebabkan terjadinya kredit macet dalam pemberian pinjaman kepada nasabahnya.

Berdasarkan penelitian yang telah ada sebelumnya bahwa metode klasifikasi dapat memberikan prediksi dalam penentuan calon nasabah yang akan mengajukan kredit sehingga dapat meminimalisir terjadinya kredit macet setelah pengajuan kredit disetujui. Maka pada penelitian ini, penulis akan mencoba untuk mengimplementasikan data persetujuan kredit di bank XYZ dengan menggunakan metode klasifikasi yang hasilnya diharapkan dapat membantu pihak bank dan para pengambil keputusan untuk menentukan kriteria nasabah yang layak untuk mengajukan kredit.

2. Metodologi Penelitian

Dalam penelitian ini, metode yang digunakan adalah penggabungan dari Tahapan *Knowledge Discovery in Database (KDD)* dan Model Prototype. Untuk mengekstrak pola atau model dari data dengan menggunakan suatu algoritma yang spesifik menggunakan metode *KDD*. Sedangkan, untuk proses yang melibatkan hubungan kerja antara perancang dan pengguna menggunakan model prototype.

Pressman (2001) menyatakan bahwa seringkali seorang pelanggan mendefinisikan serangkaian sasaran umum bagi perangkat lunak, tetapi tidak mengidentifikasi kebutuhan *input*, pemrosesan, ataupun *output* detail.

Metode digunakan dalam perancangan perangkat lunak yang akan dibangun dimana pengguna hanya memberikan beberapa kebutuhan umum perangkat lunak tanpa detail *input*, proses dan detail output yang tidak memperhatikan efisiensi algoritma yang digunakan dan tingkat adaptasi terhadap sistem informasi. Figure 2.1 adalah Metode Penelitian pada penelitian ini :

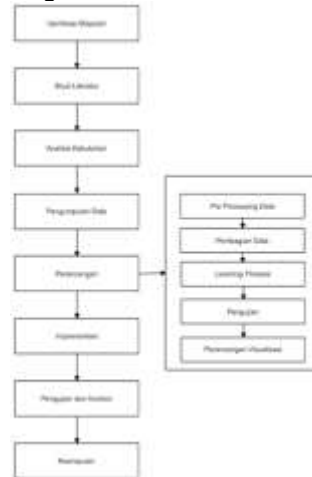


Figure 2.1 Metodologi Penelitian

1. Identifikasi Masalah

Identifikasi dilakukan dengan menganalisis permasalahan yang ada untuk memprediksi pengajuan kredit yang akan diputus.

2. Studi Literatur

Studi dan penelitian mengenai klasifikasi yang telah dilakukan oleh beberapa peneliti yang nantinya akan digunakan sebagai referensi teori yang relevan dengan permasalahan yang terjadi.

3. Analisis Kebutuhan

Analisis kebutuhan merupakan tahap untuk melakukan identifikasi hal-hal yang dibutuhkan dan dilakukan oleh sistem.

4. Pengumpulan Data

Pengumpulan data langsung dari bank xyz dalam bentuk format .xlsx. Data yang terkumpul adalah data pengajuan kredit kepemilikan rumah (KPR) di Jabodetabek pada bulan Januari 2016 hingga Desember 2019.

5. Perancangan

Dilakukan *preprocessing* data baik secara manual maupun menggunakan *Microsoft Excel*. Proses *preprocessing* dilakukan dengan tiga tahap yaitu data *selection*, *cleansing* dan *transformation*. Setelah dilakukan *preprocessing*, selanjutnya melakukan pembagian untuk data menjadi 80:20 dimana 80% merupakan data latih dan 20% adalah data uji.

6. Pengujian dan Analisis

Pengujian dan analisis dilakukan dari hasil implementasi yang sudah dilakukan dengan *confusion matrix* dan kurva *ROC* dimana hasil dari klasifikasi tersebut akan dianalisa dari kelas masuk yaitu *false positives* dan *false negatives* yang didasarkan oleh rule. Untuk pengujian sistem menggunakan *blackbox validation* sebagai dasar *testing* pengujian yang merupakan syarat dari fungsionalitas sistem untuk mengetahui kualitas *dashboard* yang dihasilkan.

7. Kesimpulan

Kesimpulan dan saran untuk penelitian selanjutnya agar dapat terus dikembangkan jauh lebih baik lagi.

3. Analisis dan Hasil

Pembahasan dan analisis dari hasil penelitian terhadap hal-hal yang berkaitan dengan penerapan dari langkah-langkah penelitian yang sudah dirancang pada penelitian ini untuk dapat menjelaskan dan menguraikan dengan jelas perancangan, pengolahan dan hasil implementasi model dengan melakukan pengujian-pengujian serta analisis dari suatu pengujian dan model yang telah dilakukan.

3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data untuk pemilihan prediksi kelayakan pengajuan kredit, di dapat dari sebuah perusahaan perbankan di Indonesia adapun *dataset* yang akan digunakan adalah data dari bulan Januari 2016 sampai dengan Desember 2019 memiliki 16 *atribut*, terdiri dari 15 *predictor* dan 1 hasil. Atribut yang dijadikan *label class* adalah atribut status pengajuan. Atribut yang menjadi parameter terlihat pada tabel 3.1.

Table 3.1 Atribut Data Master

No	Atribut	Pengertian
1	Limit Pinjaman	Limit kredit yang diajukan
2	Nomor Pengajuan	Nomor urut pengajuan nasabah
3	Jangka Waktu	Jangka waktu yang diajukan
4	Tipe Produk	Jenis kredit yang diajukan
5	Tipe Agunan	Tipe agunan
6	Harga Agunan	Harga beli agunan yang akan diajukan
7	Status Agunan	Kondisi rumah yang akan dibeli
8	Umur	Umur nasabah
9	Pendapatan	Jumlah pendapatan perbulan nasabah
10	Jenis Pekerjaan	Jenis pekerjaan nasabah
11	Nomor Persetujuan	Nomor urut persetujuan nasabah
12	Wilayah	Wilayah Pengajuan
13	Tgl lahir	Tanggal lahir nasabah
14	App_date	Tanggal pengajuan kredit

No	Atribut	Pengertian
15	Kpts_date	Tanggal persetujuan kredit
16	Status Pengajuan	Status Pengajuan yang didapat oleh nasabah

3.2. Preprocessing Data

Data kredit yang diperoleh adalah 230.650 *record* pengajuan kredit periode Januari 2016 – Desember 2019 pada Bank XYZ. Data pengajuan pada tahun 2016 sebanyak 78.789 *record*, tahun 2017 sebanyak 55.981 *record*, tahun 2018 sebanyak 42.990 *record*, dan tahun 2019 sebanyak 52.890 *record*. Untuk mendapatkan data yang berkualitas peneliti melakukan teknik *cleaning*, *integration*, *selection* dan *transformation* di dalam tahap *preprocessing*.

a. Data Cleaning dan Data Integration

Data *Cleaning* dalam penelitian ini adalah sebuah proses melakukan *replace missing value* dan menghilangkan data-data yang tidak lengkap. Pada data pengajuan kredit terdapat *missing value* sebanyak 103.793 *record* dengan melakukan *replace missing value* secara manual pada Ms. Excel dengan mengubah isi *field* masing-masing *record* dari (*blank*) menjadi *null*. Merapikan tanggal lahir, tanggal pengajuan dan tanggal keputusan kredit yang tidak sesuai dengan format tanggal.

Data *Integration* dalam penelitian ini adalah penghubungan antara beberapa tabel untuk mengetahui *sub-category* item dengan berbagai jenis nama produk yang ada dan jenis agunan yang tersedia.

Data yang digunakan adalah data pengajuan kredit selama empat tahun yaitu pada tahun 2016 sampai dengan tahun 2019 dengan 15 *predictor* atribut dan 1 atribut label class yang terbagi menjadi 2 kelas yaitu kelas debitur yang diterima dan kelas debitur yang ditolak.

Data yang diperoleh dikumpulkan sebagai data *training* dan data *testing* dimana pembagian atas *dataset* tersebut sebanyak 80% dan 20%. Terdiri dari 4 tabel, dimana masing-masing tabel terdapat 1 *field* yang digunakan untuk menghubungkan antar tabel yaitu Nomor_Pengajuan.

b. Data Selection, Data Reduction dan Data Transformation

Data *Selection* dalam penelitian ini adalah sebuah proses perubahan setiap nama atribut yang memiliki penulisan elemen dipisahkan dengan satu karakter garis bawah (*_*), tidak memiliki spasi dan juga pemilihan atribut yang didasarkan pendapat officer supporting bank xyz dimana atribut-atribut yang terpilih merupakan atribut-atribut dalam prediksi kelayakan pengajuan kredit yang berpotensi terhadap macet tidaknya suatu pinjaman nasabah dikemudian hari.

Data *Reduction* dalam penelitian ini adalah sebuah proses menghilangkan data yang tidak digunakan terkait kebutuhan penelitian sesuai dengan diskusi dengan officer supporting bank XYZ. Dimana hasil dari diskusi yang dilakukan antara lain :

1. Peneliti menghapus sebanyak 103.793 data *noisy* yang memiliki nilai *null*, dimana data tersebut dianggap tidak dapat digunakan karena tidak memiliki nilai yang pasti.
2. Menghapus enam *atribut* yaitu *atribut* nomor pengajuan, nomor keputusan, wilayah, tgl lahir, App_date, Kpts_date.
3. Menghapus data wilayah selain wilayah Jabodetabek sebanyak 227.789 *record*, dikarenakan wilayah digunakan oleh peneliti hanya wilayah Jabodetabek. Kemudian setelah proses dilakukan terdapat pengurangan data. Sehingga jumlah data yang digunakan menjadi 2.861.

Pada tahapan *transformation* data dilakukan sebuah proses pengambilan data dari *database* dan melakukan proses perubahan data ke bentuk yang dibutuhkan untuk melakukan tahapan analisis. Pada tahap ini dilakukan penentuan parameter yang akan digunakan yang terdiri dari 10 *atribut*, dimana 9 merupakan *atribut predictor* dan 1 *atribut* hasil. Adapun *atribut* yang menjadi parameter dapat dilihat pada tabel 3.3.

Table 3.2 Atribut Data Master

No	Atribut	Pengertian	Kategori
1	Limit Pinjaman	Limit Kredit yang diajukan	a. Rendah

No	Atribut	Pengertian	Kategori
			b. Menengah c. Tinggi
2	Jangka Waktu	Jangka Waktu yang diajukan	a. Pendek b. Menengah c. Panjang
3	Tipe Produk	Jenis Kredit yang diajukan	a. New Booking b. Take Over
4	Tipe Agunan	Tipe Agunan	a. Rumah Tinggal b. Apartment
5	Harga Agunan	Harga Beli Agunan yang akan diajukan	a. Rendah b. Menengah c. Tinggi
6	Status Agunan	Kondisi Rumah yang akan dibeli	a. Siap Huni b. Belum Siap Huni
7	Umur	Umur Nasabah	a. <21thn b. > 21thn
8	Pendapatan	Jumlah pendapatan perbulan nasabah	a. Rendah b. Menengah c. Tinggi
9	Jenis Pekerjaan	Jenis Pekerjaan Nasabah	a. Pegawai b. Wiraswasta
10	Status Pengajuan	Status Pengajuan yang didapat oleh nasabah	a. Terima b. Tolak

c. Data *Training* dan *Testing*

Data yang tersedia sebanyak 2.861 *record* yang siap digunakan untuk data latih dan data uji dengan menggunakan Ms. Excel, pembagian data latih dan data uji dibagi secara manual dan random dalam pemilihan data yang akan digunakan untuk data latih maupun data uji sehingga jumlah perbandingan kedua data tersebut adalah 80:20. Penulis melakukan data uji untuk pembagian data latih dan data uji dengan perbandingan 80 : 20 seperti pada tabel 3.9.

Table 3.9 Data *Training* dan *Testing*

Pembagian	Persentase	Total
<i>Data Latih</i>	80%	2,300
<i>Data Uji</i>	20%	561
Total	100%	2,861

Berdasarkan tabel 3.10 dapat diketahui bahwa persentase data menjadi 2 kelompok yaitu 80% untuk data latih dan 20% untuk data Uji. Pembagian *data latih* dan *uji* dengan skala 80 : 20 berdasarkan jumlah nasabah dalam kategori terima dan tolak dari atribut status pengajuan. Terdapat 1,307 nasabah dalam kategori diterima dan 1,554 nasabah yang masuk dalam kategori tolak. Setelah mengetahui jumlah nasabah dalam status terima dan tolak, selanjutnya dilakukan pembagian data 80% dari jumlah nasabah yang berstatus terima dan tolak untuk dijadikan *data latih* dan untuk *data uji* diambil sebanyak 20% yang merupakan sisa dari jumlah nasabah yang berstatus terima dan tolak. Adapun pembagian tersebut dapat di lihat pada table 3.10.

Table 3.10 Pembagian Data *Latih* dan *Uji*

	Data Uji	Data Latih	Grand Total
Terima	222	1,085	1,307
Tolak	339	1,215	1,554
Grand Total	561	2,300	

Setelah melakukan pembagian *data Latih* dan *uji* sehingga data siap untuk dilatih dan di ujikan menggunakan model dengan algoritma Naïve Bayes dan C4.5.

3.2. Hasil Pembuatan Model

Pada tahap ini akan dijelaskan mengenai prediksi pengajuan kredit nasabah menggunakan algoritma Naïve Bayes dan Decision Tree C4.5.

a. Penyusunan Model *Naïve Bayes*

Perhitungan manual dengan algoritma *Naïve Bayes* menggunakan *sample dataset* yang ada setelah dilakukan pemilihan atribut. Adapun langkah-langkah untuk membuat algoritma *Naïve Bayes*, antara lain :

1. Data sample yang digunakan dalam perhitungan ini sejumlah 60 data.
2. Hitung jumlah probabilitas, jika data yang tersedia merupakan data numerik maka :
 - a. Hitung nilai mean dan standar deviasi dari setiap parameter data numerik
 - b. Hitung nilai probabilitik dengan menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut

Pada penelitian ini data sample yang digunakan bukan merupakan data numerik sehingga dapat menggunakan perhitungan jumlah probabilitasnya saja.
3. Hasil nilai dalam tabel mean, standar deviasi dan probabilitas.

b. Penyusunan Model *Decision Tree C4.5*

Perhitungan manual dengan menggunakan Algoritma C4.5 menggunakan *sample dataset* yang ada setelah dilakukan pemilihan atribut. Adapun langkah-langkah untuk membuat algoritma C4.5, antara lain :

1. Siapkan *data latih*. *Data latih* yang digunakan dalam contoh perhitungan ini adalah 60 *sample data*.
2. Hitung nilai *info* untuk mengetahui nilai *split* dari data yang berupa angka.
3. Kemudian hitung nilai *entropy* dari setiap atribut.
4. Setelah itu akan didapatkan nilai *gain*. Nilai *gain* yang paling besar dari setiap atribut yang akan nantinya digunakan sebagai percabangan pertama.
5. Untuk mendapatkan percabangan selanjutnya dilakukan perhitungan sama dengan langkah sebelumnya.
6. Jika ada atribut yang memiliki nilai = 0, artinya atribut tersebut sudah dapat dikatakan NO (ditolak pada kasus ini).

3.2. Hasil Komparasi

Multi class confusion matrix digunakan pada penelitian ini untuk menguji kinerja dari dua metode yang diterapkan yaitu algoritma *Naïve Bayes* dan algoritma *Decision Tree C4.5* yang akan menghasilkan nilai *accuracy*, *precision* dan *recall*. Untuk melihat *table confusion matrix* untuk metode *Naïve Bayes* dapat dilihat pada table 3.11 dan untuk metode *Decision Tree C4.5* dapat dilihat pada table 3.12.

Table 3.11 Hasil Confusion Matrix untuk metode *Naïve Baye*

Confusion Matrix		
n = 561	Positive (Actual)	Negative (Actual)
Positive (Predicted)	103	238
Negative (Predicted)	119	101

Table 3.12 Hasil Confusion Matrix untuk metode *Decision Tree C4.5*

Confusion Matrix		
n = 561	Positive (Actual)	Negative (Actual)
Positive (Predicted)	132	137
Negative (Predicted)	90	202

a. *Accuracy*

Accuracy menggambarkan nilai akurat model yang dapat mengklasifikasikan dengan benar. *Accuracy* merupakan rasio prediksi benar dengan keseluruhan data. Untuk menghitung akurasi menggunakan metode *Naïve Bayes* didapat *true positive*, maka dapat dihitung sebagai berikut :

$$\begin{aligned}
 & Accuracy_{Naïve Bayes} \\
 &= \frac{(TP + TN)}{(TP + FP + FN + TN)} \times 100 \\
 &= \frac{(103 + 101)}{(103 + 238 + 119 + 101)} \times 100 \\
 &= 36.36\%
 \end{aligned}$$

$$\begin{aligned}
 & Accuracy_{Naïve Bayes} \\
 &= \frac{(TP + TN)}{(TP + FP + FN + TN)} \times 100 \\
 &= \frac{(132 + 202)}{(132 + 137 + 90 + 202)} \times 100 \\
 &= 59.54\%
 \end{aligned}$$

b. *Precision*

Precision menggambarkan tingkat keakuratan antara data yang diminta dengan hasil prediksi yang diberikan oleh model. *Precision* merupakan proses perbandingan antara rasio prediksi benar *positive* dengan keseluruhan hasil yang diprediksi *positive*. Sehingga dari semua kelas *positive* yang telah diprediksi dengan benar, berapa banyak data yang benar *positive*, maka dapat dihitung sebagai berikut :

$$\begin{aligned}
 Precision_{Naive\ Bayes} &= \frac{TP}{(TP + FP)} \times 100 \\
 &= \frac{103}{(103 + 238)} \times 100 \\
 &= 30.21\%
 \end{aligned}
 \qquad
 \begin{aligned}
 Precision_{C4.5} &= \frac{TP}{(TP + FP)} \times 100 \\
 &= \frac{132}{(132 + 137)} \times 100 \\
 &= 49.07\%
 \end{aligned}$$

c. *Recall*

Recall menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. *Recall* merupakan rasio prediksi benar *positive* dibandingkan dengan keseluruhan data yang benar *positive*, dapat dihitung sebagai berikut :

$$\begin{aligned}
 Recall_{Naive\ Bayes} &= \frac{TP}{(TP + FN)} \times 100 \\
 &= \frac{103}{(103 + 119)} \times 100 \\
 &= 46.40\%
 \end{aligned}
 \qquad
 \begin{aligned}
 Recall_{C4.5} &= \frac{TP}{(TP + FN)} \times 100 \\
 &= \frac{132}{(132 + 90)} \times 100 \\
 &= 59.46\%
 \end{aligned}$$

d. Kurva ROC

Kurva ROC (*Receiver Operating Characteristic*) menggambarkan dari hasil nilai yang telah didapatkan pada perhitungan dengan *confusion matrix* yaitu antara *False Positive Rate* dengan *True Positive Rate*, maka dapat dihitung sebagai berikut :

$$\begin{aligned}
 FPR_{Naive\ Bayes} &= \frac{False\ Positive}{(False\ Positive + True\ Negative)} \\
 &= \frac{119}{(119 + 101)} = 0.541
 \end{aligned}
 \qquad
 \begin{aligned}
 FPR_{C4.5} &= \frac{False\ Positive}{(False\ Positive + True\ Negative)} \\
 &= \frac{90}{(90 + 202)} = 0.308
 \end{aligned}$$

Dari hasil pengujian menggunakan *confussion matrix* yang mengacu pada table 3.11 dan table 3.12, maka dapat disimpulkan prediksi pengajuan kredit menggunakan metode algoritma C4.5 lebih baik dibanding menggunakan metode algoritma *Naïve Bayes*, dikarenakan C4.5 mendapatkan *Accuracy*, *Precision* dan *Recall* lebih besar dibandingkan dengan *Naïve Bayes*. Tabel hasil perbandingan dari klasifikasi *Naïve Bayes* dan C4.5 dapat dilihat pada tabel 3.13.

Table 3.13 Hasil Perbandingan Performa dari klasifikasi Naïve Bayes dan Decision Tree C4.5

	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>AUC</i>
<i>Naïve Bayes</i>	36.36%	30.21%	46.40%	0.633
Decision Tree C4.5	59.54%	49.07%	59.46%	0.630

Dari Table 3.13 Nilai *AUC* untuk algoritma decision tree C4.5 yaitu 0.630 lebih rendah dibandingkan *Naïve Bayes* sebesar 0.633, dari kedua algoritma yang digunakan dimana menghasilkan Nilai *AUC* <0.700 maka kedua algoritma tersebut masuk dalam kategori *poor classification*. Jika dilihat dari sisi *accuracy*, untuk algoritma decision tree C4.5 memiliki nilai *accuracy* lebih tinggi dibandingkan *Naïve Bayes* yaitu sebesar 59.54%. sehingga dapat disimpulkan bahwa algoritma decision tree C4.5 lebih baik dibandingkan algoritma *Naïve Bayes*.

3.2. Penerapan Aplikasi Data Mining

Untuk menghasilkan suatu prototipe yang dapat memprediksi pengajuan kredit setiap nasabah agar tidak menimbulkan kredit macet dikemudian hari. Sehingga rule yang dihasilkan adalah pembuatan prototipe aplikasi berbasis website dengan menggunakan metode algoritma *Naïve Bayes* dan algoritma decision tree C4.5 agar dapat mempermudah pihak bank dalam memutuskan suatu pengajuan kredit yang diajukan oleh nasabah.

Adapun alur dari *Activity Diagram* yang diimplementasikan kedalam prototipe prediksi pengajuan kredit kepemilikan rumah di bank XYZ terdapat pada Figure 3.1:

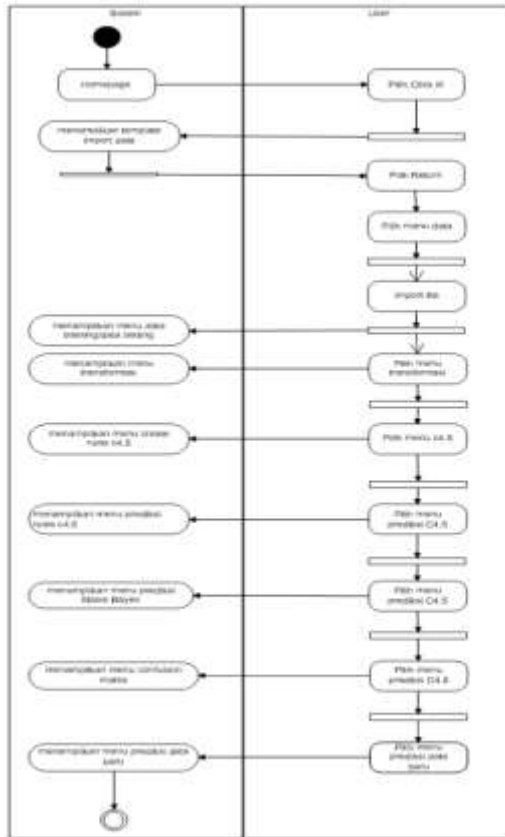


Figure 3.1 Activity Diagram

Prototipe yang digunakan dalam penelitian ini berbasis website dengan bahasa pemrograman PHP dan *database* MySQL. Adapun tampilan dari screen homepage hingga screen prediksi data baru dapat dilihat pada figure 3.2 sampai dengan gambar 4.19 :

a. Tampilan menu homepage

Pada halaman menu homepage. User dapat melihat logo dan juga dapat memilih click it, untuk mengetahui bagaimana template dari form excel yang tersedia seperti pada figure 3.2.



Figure 3.2 Tampilan Menu Homepage

b. Tampilan menu template import data

Pada halaman ini tersusun format excel untuk dapat di import ke dalam sistem ini seperti pada figure 3.3.



Figure 3.3 Tampilan Menu Template import data

- c. Tampilan menu import data
 Pada figure 3.4 merupakan merupakan proses import data dari excel yang belum di transformasi dan untuk figure 3.5 adalah hasil data yang telah di import.



Figure 3.4 Tampilan Menu Import Data



Figure 3.5 Tampilan Hasil Import Data

- d. Tampilan menu Transformasi Data
 Pada figure 3.6 merupakan hasil proses dari menu transformasi data *training* dan data *testing*.



Figure 3.6 Tampilan Menu Transformasi Data

- e. Tampilan menu C4.5 dan Prediksi C4.5
 Pada figure 3.7 menampilkan menu proses dan hasil dari *create rule decision tree* dengan C4.5. sedangkan pada figure 3.8 menampilkan hasil prediksi data *testing* maupun *training*.



Figure 3.7 Menu C4.5



Figure 3.8 Tampilan Menu Prediksi Data dengan C4.5

- f. Tampilan menu naïve bayes
 Pada figure 3.9 menampilkan hasil prediksi data *training* (data latih) dan data *testing* (data uji) dengan menggunakan algoritma Naïve Bayes.



Figure 3.9 Tampilan Menu Prediksi dengan Naïve Bayes

- g. C.matrix
 Pada figure 3.10 sampai dengan gambar 4.12 menampilkan hasil dari *Confusion Matrix*.



Figure 3.10 Tampilan Menu *Confusion Matrix*



Figure 3.11 Hasil *Confusion Matrix Naïve Bayes*



Figure 3.12 Hasil Confusion Matrix Decision Tree C4.5

- h. Prediksi data baru
 Pada figure 3.13 dan figure 3.14 merupakan proses untuk memprediksi data baru dengan menggunakan C4.5.



Figure 3.14 Hasil List Data Prediksi Data Baru



Figure 3.13 Tampilan Menu Prediksi Data Baru

3.2. Pengujian *Prototype* Perangkat Lunak

Pengujian *prototype* ini menggunakan metode *blackbox*. Tujuannya ialah untuk mengetahui apakah *prototype* aplikasi yang dibangun sesuai dengan fungsional yang diharapkan. Nantinya aplikasi akan diberikan beberapa jenis kondisi (*input*) dan beberapa jenis keluaran (*output*) dibandingkan dengan hasil yang diharapkan.

- a. Pengujian *Import* data

Table 3.15 Pengujian *Import* Data

Input	Yang diharapkan	Kesimpulan
File yang akan diupload	File dalam bentuk .xls	Diterima
Klik <i>insert</i>	File berhasil di <i>import</i> oleh sistem	Diterima

Klik Data	Training	Sistem akan menampilkan data <i>training</i> yang terdapat pada <i>database</i>	Diterima
Klik Data	Testing	Sistem akan menampilkan data <i>testing</i> yang terdapat pada <i>database</i>	Diterima

Klik Delete Data	Sistem akan menampilkan pilihan menu <i>delete</i> data	Diterima
------------------	---	----------

b. Pengujian Menu Lainnya

Table 3.16 Pengujian Menu Lainnya

Input	Yang diharapkan	Kesimpulan
Menu Transformasi	Sistem akan menampilkan pilihan data yang akan di transformasi	Diterima

Pilihan Data Transformasi	Sistem akan mentransformasikan data sesuai dengan kondisi yang ada	Diterima
Klik Create pada menu C4.5	Sistem akan melakukan proses <i>create rule</i> pada data yang telah dipilih	Diterima
Klik menu C. Matrix	Sistem akan menampilkan tampilan Tabel <i>Confusion Matrix</i>	Diterima

Setelah dilakukan pengujian *blackbox* didapatkan hasil bahwa aplikasi yang dibangun telah berjalan dengan baik sesuai dengan yang diharapkan.

4. Kesimpulan

4.1. Kesimpulan

Berdasarkan penelitian yang sudah dilakukan untuk memprediksi kelayakan pengajuan kredit kepemilikan rumah pada bank XYZ dengan membandingkan dua metode algoritma yaitu algoritma Naïve Bayes dan C4.5, maka kesimpulan yang dapat ditarik antara lain :

1. Penggunaan algoritma *Decision Tree C4.5* lebih baik daripada algoritma *Naïve Bayes* untuk memprediksi kelayakan pengajuan kredit kepemilikan rumah sesuai dengan kriteria yang diinginkan pihak Bank XYZ.
2. Hasil pengujian dua metode klasifikasi yaitu *Naïve Bayes* dan *Decision Tree C4.5* didapat bahwa algoritma *Decision Tree C4.5* tingkat akurasi yang dimiliki sebesar 59.54% untuk memprediksi kelayakan pengajuan kredit lebih baik dibandingkan dengan algoritma *Naïve Bayes* sebesar 36.36%.

4.2. Saran

Penelitian ini masih bisa dikembangkan lagi dengan atribut atau studi kasus yang berbeda sehingga dapat memperoleh hasil yang jauh lebih baik dan perlu ditambahkan beberapa perbandingan dari metode klasifikasi lainnya agar dapat meningkatkan nilai akurasi dan performa untuk menghasilkan hasil prediksi yang jauh lebih baik lagi.

Daftar Pustaka

- [1] M. dan S. Kuncoro, "Manajemen Perbankan Teori dan Aplikasi," Edisi Kedu., Yogyakarta: BPFE, 2011.
- [2] M. H. Rifqo and A. Wijaya, "Implementasi Algoritma Naive Bayes Dalam Penentuan Pemberian Kredit," *Pseudocode*, 2017, doi: 10.33369/pseudocode.4.2.120-128.
- [3] Bank Indonesia, "Undang-Undang Nomor 10 Tahun 1998 Tentang Perbankan," *Undang. Nomor 10 Tahun 1998 tentang Perbank.*, 1998, doi: 10.1007/s13398-014-0173-7.2.
- [4] J. Han and M. Kamber, *Data Mining Concept and Tehniques*. 2006.
- [5] F. Gorunescu, "Data mining: Concepts, models and techniques," *Intell. Syst. Ref. Libr.*, 2011, doi: 10.1007/978-3-642-19721-5.
- [6] S. Budi, *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. 2007.
- [7] E. Kusriani dan Taufik Luthfi, "Algoritma Data Mining," Yogyakarta: Andi, 2009, pp. 1–212.
- [8] J. Han, M. Kamber, and J. Pei, *Data Transformation by Normalization*. 2011.
- [9] D. A. Kurniawan and Y. I. Kurniawan, "Aplikasi Prediksi Kelayakan Calon Anggota Kredit Menggunakan Algoritma Naïve Bayes," *J. Teknol. dan Manaj. Inform.*, 2018, doi: 10.26905/jtmi.v4i1.1831.

- [10] S. Wahyuningsih and D. R. Utari, "Perbandingan Metode K-Nearest Neighbor , Naive Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit," *Konf. Nas. Sist. Inf. 2018*, 2018.
- [11] L. Farokhah and R. D. Indahsari, "Implementasi Decision Tree C4.5 Dalam Penentuan Pinjaman Uang Di Koperasi XYZ di Banjarmasin," *Klik - Kumpul. J. Ilmu Komput.*, 2019, doi: 10.20527/klik.v6i3.268.
- [12] A. U. Zailani and N. L. Hanun, "Penerapan Algoritma Klasifikasi Random Forest Untuk Penentuan Kelayakan Pemberian Kredit Di Koperasi Mitra Sejahtera," *Infotech J. Technol. Inf.*, 2020, doi: 10.37365/it.v6i1.61.
- [13] B. Ramadhani, "Klasifikasi Metode Naive Bayes Untuk Kelancaran Pembayaran Kredit Leasing Sepeda Motor," *Technol. J. Ilm.*, 2017, doi: 10.31602/tji.v8i3.1131.