

BAB III

METODOLOGI PENELITIAN

3.1. Waktu Penelitian

Waktu penelitian ini dimulai pada bulan Mei 2022 sampai dengan bulan September 2021.

3.2. Tempat Penelitian

Tempat penelitian ini berlokasi di Koperasi Gentiaras Pringsewu Lampung dengan Kepala bagian Koperasi.

3.3. Metode Pengumpulan Data

Dalam metode pengumpulan data ini mempunyai peranan yang penting untuk mendapatkan suatu informasi dari penelitian yang dilakukan. Data yang relevan dengan pokok pembahasan adalah indikator keberhasilan penelitian. Pengumpulan data harus dilakukan dengan cara yang sangat tepat. Dalam metode pengumpulan data ini, penulis menggunakan beberapa metode yaitu :

3.3.1 Wawancara

Metode wawancara merupakan metode pengumpulan data dengan cara mengajukan beberapa pertanyaan langsung kepada narasumber. Wawancara dilakukan untuk mengetahui secara langsung kepada kepala koperasi Gentiaras Pringsewu Lampung. Dalam kegiatan ini diajukan pertanyaan lisan dalam usaha untuk mendapatkan dan melengkapi data yang akan diperoleh. Wawancara dilakukan di Koperasi Gentiaras Pringsewu Lampung.

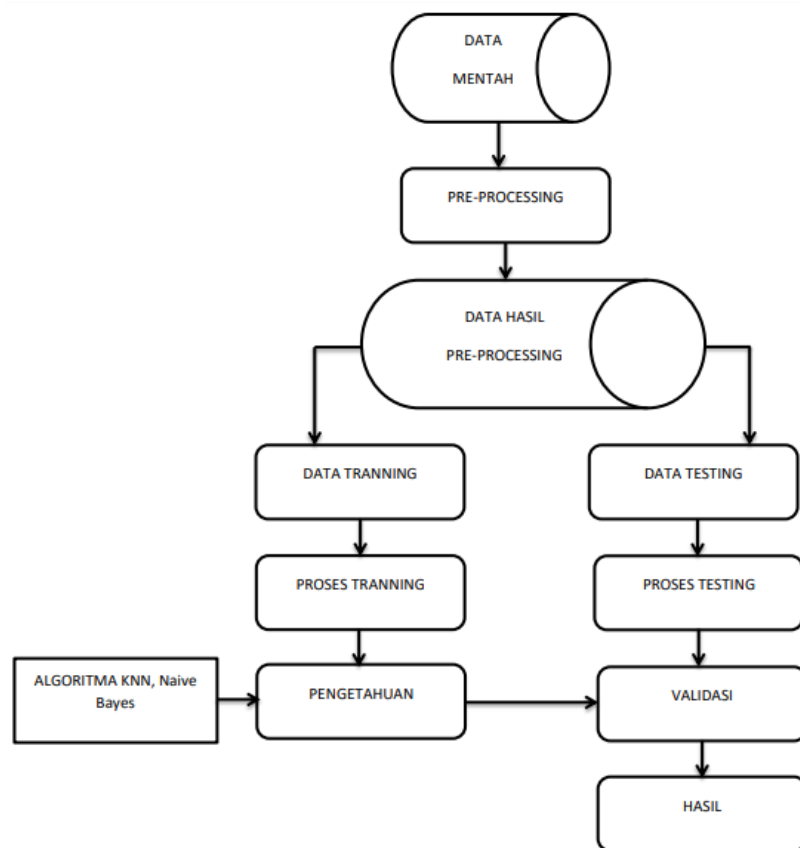
3.3.2 Studi Pustaka

Metode studi pustaka ini dilakukan dengan cara mempelajari jurnal, buku-buku atau studi literatur terdahulu yang berkaitan dengan kredit pinjaman, data mining, klasifikasi dan algoritma data mining. Setelah data dikumpulkan maka dilakukan pembelajaran mengenai data yang akan digunakan. Pembelajaran data diharapkan dapat membiasakan peneliti dengan data-data yang sudah dikumpulkan

serta dapat menemukan wawasan awal mengenai informasi apa saja yang bisa didapatkan didalamnya.

3.4. Tahapan Penelitian

Dalam penelitian ini dilakukan tahap-tahap untuk menyelesaikan penelitian ini agar sesuai dengan tujuan. Tahapan penelitian seperti Gambar 3.1



Gambar 3.1 Tahapan Penelitian

Berdasarkan gambar tahapan penelitian diatas dapat dijelaskan sebagai berikut:

1. Data mentah direduksi terlebih dahulu dan dimasukkan ke dalam Microsoft Excel untuk menjadi data training, total data yang dikumpulkan 334 data.
2. Dengan menggunakan metode pre-processing data diolah untuk menghilangkan duplikat dan missing, yang sudah tersedia pada software RapidMiner Studio.
3. Selanjutnya data testing akan di uji dengan data latih yang telah diperoleh.
4. Selanjutnya data diuji dengan menggunakan algoritma KNN, Naive Bayes dan Logistic Reggresion.

5. Data yang telah diuji kemudian dibandingkan mana hasil performa algoritma yang terbaik.

3.5. Metode Analisis Data

Teknik analisa data menggunakan data kuantitatif berupa angka atau numerik dan nominal. Pada penelitian ini, analisa data dilakukan melalui data pada koperasi Gentiaras Pringsewu Lampung kemudian data diolah dan di uji dalam pengujian pada algoritma Naive Bayes, KNN dan Logistic Regression Tahapan metode analisis data adalah sebagai berikut :

3.5.1 Fase Pemahaman Data

Data yang digunakan adalah data sekunder, yang didapat dari koperasi Gentiaras Pringsewu Lampung di dalam data tersebut dapat diketahui lancar dan tidak lancar pemberian pinjamna yang terdiri dari 8 atribut predictor dan 1 atribut target. Atribut-atribut yang menjadi parameter terlihat pada tabel 3.1

Tabel 3.1 Atribut Data

Atribut	No anggota
	Nama
	Nilai Pinjaman
	Jangka waktu
	Sisa saldo pinjaman
	Jumlah tanggungan
	Status
	Pendidikan Terakhir
	Predikat
	Jaminan

3.5.2 Tahapan Data Persiapan (Data Preparation)

Pada tahap ini menjelaskan tentang tahap awal data mining. Data yang telahdidapatkan akan diolah ke format yang dibutuhkan, pengelompokkan dan penentuan attribut, variabel. Dalam melakukan pengolahan data awal, akan di

training atau di hitung dan dilakukan beberapa tahapan agar didapatkan data yang bisa digunakan untuk tahap selanjutnya.

1. Data Selection

Tahap data selection atau pemilihan data merupakan tahap pemilihan atribut dari data yang akan dianalisis, karena tidak semua data yang terdapat dalam data mentah akan digunakan. Sehingga didapat beberapa atribut yang akan digunakan. Hasil tabel data selection dapat dilihat pada Tabel 3.2

Tabel 3.2 Data Selection

Atribut	Indicator	Detail Penggunaan
No anggota	X	Tidak digunakan
Nama	V	Digunakan
Nilai Pinjaman	V	Digunakan
Jangka waktu	V	Digunakan
Jumlah tanggungan	V	Digunakan
Status	V	Digunakan
Pendidikan Terakhir	V	Digunakan
Predikat	V	Digunakan
Jaminan	V	Digunakan

Tabel diatas menerangkan atribut yang akan digunakan dan tidak digunakan dalam penelitian ini. Indikator “V” adalah menandakan bahwa variabel tersebut digunakan, sedangkan untuk indikator “X” adalah menandakan bahwa atribut tersebut dieliminasi atau tidak digunakan pada tahap penentuan kriteria. Pengeliminasian beberapa atribut tersebut karena tidak mempengaruhi hasil dari proses penilaian.

2. Pre-Processing (data cleaning)

Tahap pre-processing merupakan tahap dilakukannya pembersihan terhadap data missing value yaitu data yang tidak konsisten atau kosong, dan juga dilakukannya pemilihan variabel terhadap data yang ingin

digunakan pada proses data mining. Klasifikasi kelayakan kredit pinjaman koperasi Gentiaras Pringsewu Lampung akan menjadi keputusan lancar atau tidak lancar.

3. Data Integration dan Transformation

Untuk meningkatkan akurasi dan efisiensi algoritma. Data dalam penelitian ini dikategorikan atau diklasifikasi. Setelah mendapatkan data selanjutnya adalah mengklasifikasi atribut-atribut karena berpengaruh terhadap hasil prediksi. Klasifikasi data ini dilakukan dengan bantuan tool Ms.Excel. Data kosumen yang telah diklasifikasi diambil 7 atribut. Berikut atribut- atributnya adalah :

- a. Nama : merupakan atribut yang berisikan nama konsumen
- b. Nilai Pinjaman : merupakan nilai pinjaman dari konsumen terhadap koperasi Gentiaras Lampung.
- c. Jangka Waktu : merupakan durasi atau batas waktu pinjaman yang harus dikembalikan oleh konsumen terhadap koperasi.
- d. Sisa saldo pinjaman : merupakan sisa jumlah pinjaman konsumen atas pembayaran yang telah dilakukan secara kredit.
- e. Jumlah tanggungan : merupakan jumlah tanggungan keluarga dari konsumen.
- f. Pendidikan terakhir : merupakan latar belakang pendidikan terakhir dari konsumen koperasi.
- g. Predikat : merupakan atribut yang berisikan keterangan hasil dari semua pengecekan terhadap persyaratan konsumen yaitu “lancar dan tidak lancar”
- h. Jaminan : Merupakan dokumen yang harus dilampirkan yang menjadi jaminan pada saat pengajuan kredit.

4. Pembagian Dataset

Sumber data yang dihasilkan berupa dataset dengan jumlah 334 record data yang berisi data nasabah. Kemudian data tersebut akan dijadikan landasan awal proses mining data sehingga didapatkan suatu output signifikansi dan analisa hasil yang menyatakan layak tidaknya dalam

penentuan pemberian kredit. Set data dibagi kedalam dua bagian yakni ada data testing yang merupakan data yang digunakan untuk mengukur sejauh mana classifier berhasil melakukan klasifikasi dengan benar. Data training (training set) merupakan data yang digunakan oleh algoritma klasifikasi (KNN, Naïve Bayes dan Logistic Regression) untuk membentuk sebuah model classifier. Model ini merupakan representasi pengetahuan yang akan digunakan untuk prediksi kelas data baru yang belum pernah ada.

Tabel set data ini merupakan suatu data murni yang belum diolah sedemikian rupa menjadi data training. Jika kita lihat secara seksama pada tabel data set ini maka akan didapati beberapa atribut yang tidak digunakan dan record yang duplikat.

Tidak hanya dibagi kedalam data training saja yang digunakan sebagai landasan pemodelan algoritma Naive Bayes, KNN dan Logistic Regression namun dari sejumlah besar dataset yang ada diperlukan pembagian data uji (data testing) yang digunakan untuk menguji sejauh mana classifier berhasil melakukan klasifikasi dengan benar.

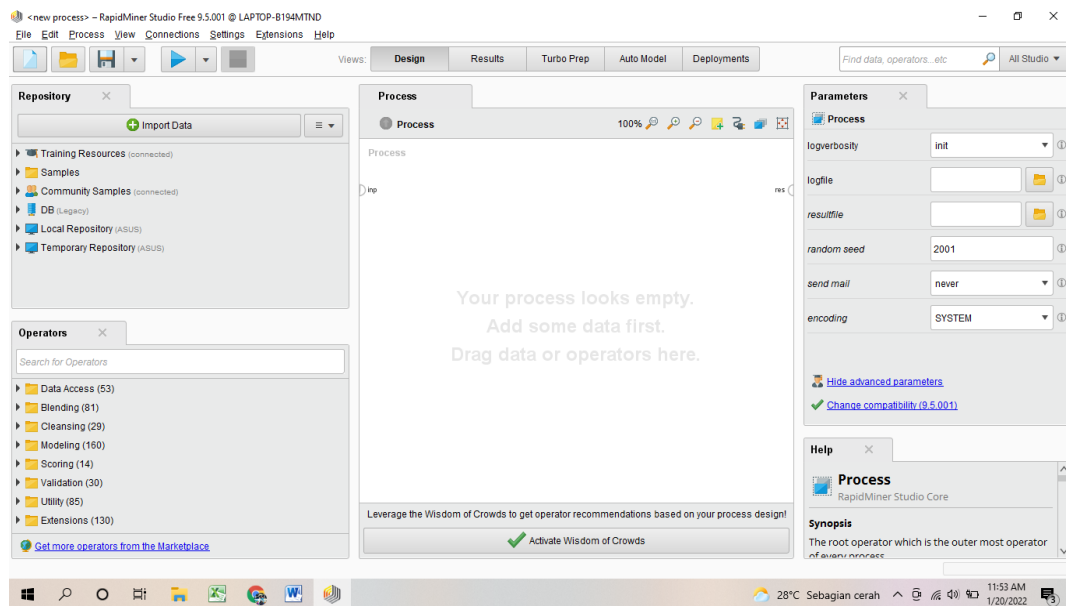
5. Pemodelan

Pada tahap ini dilakukan pemilihan teknik pemodelan yang tepat. Pada penelitian ini, menggunakan model atau algoritma Naïve Bayes, KNN dan Logistic Regression dalam RapidMiner yang merupakan solusi untuk melakukan analisis terhadap data mining, klasifikasi dan analisis prediksi. RapidMiner merupakan perangkat lunak yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang diintegrasikan pada produknya sendiri. Setelah melewati tahap persiapan pengolahan Data, sehingga pada penelitian ini digunakan dataset sebanyak 334 record.

3.6. Implementasi dan Pengujian Metode

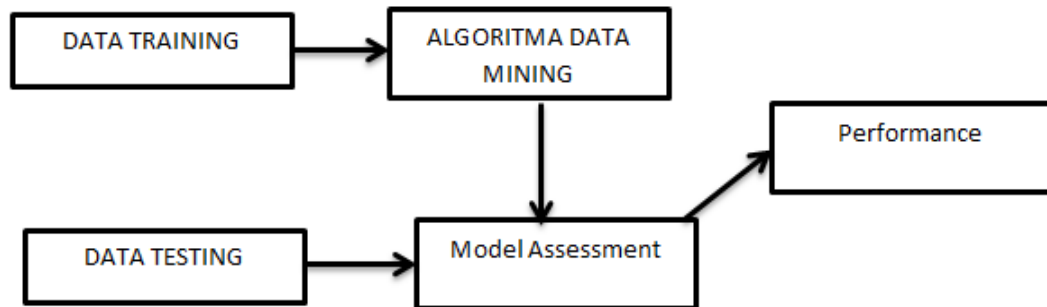
Implementasi pengujian klasifikasi dan keakurasian pada penelitian ini menggunakan software RapidMiner. RapidMiner merupakan pemrograman lunak yang bekerja dalam pengolahan data. RapidMiner merupakan bahasa pemrograman yang mempunyai cakupan kemampuan yang luas dengan

menggunakan prinsip dan algoritma data mining. Disamping itu RapidMiner dapat mengekstrakan pola-pola dari data akhir yang sangat besar dengan mengkombinasikan metode statistika, kecerdasan buatan dan database. RapidMiner dikhususkan untuk penggunaan data mining, model yang disediakan cukup lengkap diantaranya model Naive Bayes, KNN dan Logistic Regression lain-lainnya. Berikut adalah tampilan awal tools Rapidminer yang dapat dilihat pada gambar 3.2



Gambar 3.2 Rapidminer

Tahapan-tahapan pengujian klasifikasi dan akurasi menggunakan Rapidminer dapat dilihat pada Gambar 3.3 :



Gambar 3.3 Tahapan Pengujian Klasifikasi

3.7. Instrument Penelitian

Proses klasifikasi kredit pinjaman koperasi untuk mengetahui lancar atau tidak lancar nasabah dalam melakukan pembayaran. Adapun instrument yang dibutuhkan dalam melakukan penelitian ini dilakukan berdasarkan permasalahan yang telah diuraikan pada bab sebelumnya. Adapun instrumen yang dibutuhkan dalam melakukan penelitian ini yaitu :

1. Bahan

Bahan penelitian ini yang dibutuhkan adalah data kredit pinjaman di koperasi Gentiaras Pringsewu Lampung, yang akan digunakan untuk mencari perhitungan akurasi dan mengklasifikasikan nasabah dalam pemberian pinjaman.

2. Peralatan

Tujuan dari peralatan adalah untuk mengetahui kebutuhan sistem agar mempermudah perancangan. Peralatan ini meliputi kebutuhan perangkat keras dan perangkat lunak. Kebutuhan-kebutuhan dibawah ini merupakan kebutuhan minimal dari sistem:

a. Kebutuhan Perangkat Lunak (Software)

Perangkat lunak yang dibutuhkan adalah sebagai berikut :

➤ Sistem Operasi Windows 10

Untuk mendukung penelitian ini, maka penulis menggunakan sistem operasi windows 10 dengan versi 64 bit.

➤ Microsoft excel 2017

Sebagai media penulisan datasheet.

➤ RapidMiner

Tools yang akan digunakan untuk melihat hasil klasifikasi dan akurasi dari algoritma yang digunakan terhadap datasheet yang sedang diteliti.

b. Kebutuhan Perangkat Keras (Hardware)

Selain kebutuhan perangkat lunak (Software), juga dibutuhkan perangkat keras (Hardware) yang digunakan pada penelitian ini, laptop dengan spesifikasi :

➤ Processor Intel Core i3-6006U, 2.0 GHz. X441U

➤ RAM 4 GB

➤ 1 TB HDD