

BAB III

METODOLOGI PENELITIAN

3.1 KERANGKA KERJA PENELITIAN

Untuk membantu dalam penyusunan penelitian ini, maka perlu adanya kerangka kerja (*framework*) yang jelas tahapan-tahapannya. Kerangka kerja ini merupakan langkah-langkah yang akan dilakukan dalam penyelesaian masalah yang akan dibahas. Adapun kerangka kerja penelitian yang akan digunakan adalah sebagai berikut :



Gambar 3. 1 Kerangka Kerja Penelitian

Pada Gambar 3.1, maka dapat diuraikan masing-masing tahapan dalam penelitian adalah sebagai berikut :

1. Perumusan Masalah

Masalah yang dirumuskan dalam penelitian ini adalah bagaimana penerapan *Algoritma Naïve Bayes* dalam memprediksi tingkat persetujuan kartu kredit.

2. Penentuan Tujuan

Tujuan dari penelitian ini adalah untuk mengetahui seberapa akurat tingkat akurasi *Algoritma Naïve Bayes* dalam memprediksi tingkat persetujuan kartu kredit.

3. Mempelajari Literatur

Mempelajari literatur-literatur yang dapat mencapai tujuan penelitian, literatur-literatur bersumber dari buku-buku perpustakaan Universitas Dinamika Bangsa Jambi, jurnal dan jaringan internet. Literatur-literatur yang digunakan nanti dilampirkan dalam daftar pustaka.

4. Pengumpulan Data dan Informasi

Dalam pengumpulan data, penulis mendapatkan *dataset online* yang terdapat pada sebuah *website* yang bernama "www.kaggle.com" di internet.

5. Proses Data Mining

Data Mining adalah proses pengekstrasian *knowledge* yang tersimpan dalam dataset bervolume besar. Untuk mendapatkan *knowledge* dalam dataset.

6. Pengujian

Pada tahap ini dilakukan pengujian dari hasil yang didapat dari tahap sebelumnya sebagai pedoman untuk mendapatkan hasil tingkat persetujuan kartu kredit.

7. Pembuatan Laporan

Pada tahap ini dilakukan pembuatan laporan, membuat hasil akhir dari suatu kegiatan penelitian berdasarkan data dan fakta yang telah diamati pada saat meneliti.

3.2 KAJIAN DATASET

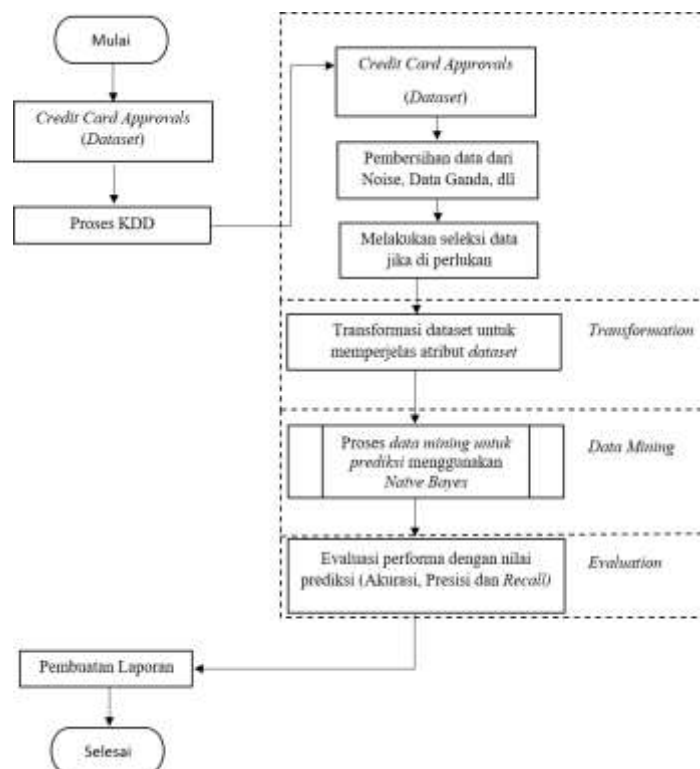
Metode pengumpulan data yang digunakan dalam penelitian adalah mencari dan mengambil *dataset online* dari sebuah website yang bernama “*www.kaggle.com*” (<https://www.kaggle.com/datasets/samueltcortinhas/credit-card-approval-clean-data>) untuk dijadikan pengujian dalam penelitian ini. Pengumpulan data dilakukan untuk memperoleh hasil dan informasi yang dibutuhkan dalam hal mencapai tujuan penelitian.

Penulis mengambil dataset dari website Kaggle yang berjudul “*Credit Card Approvals (Clean Data)*” yang berjumlah 690 data dengan 16 atribut yaitu *Gender, Age, Debt, Married, BankCustomer, Industry, Ethnicity, YearsEmployed, PriorDefault, Employed, CreditScore, DriversLicense, Citizen, ZipCode, Income, Approved*. Termasuk atribut “*Approved*” sebagai label (*class*). Dalam penelitian [42], menggunakan 2/3 untuk *data training* dan 1/3 untuk *data testing*. Secara umum, proporsi *data training* dan *data testing* sebesar 60:40,

70:30. atau 80:20, bahkan 90:10 jika dataset relatif besar. Kemudian penulis akan membagi data tersebut untuk *data training* sebanyak 60%, 70% dan 80%, selanjutnya *data testing* sebanyak 20%, 30% dan 40%.

3.3 ALUR DATA MINING PADA PENELITIAN

Pada tahap ini, akan dilihat apa saja tahapan dan proses penelitian yang akan dilakukan oleh penulis, adapun alur penelitian yang akan digunakan dalam penelitian adalah sebagai berikut:



Gambar 3. 2 Alur Data Mining

Berdasarkan alur penelitian pada Gambar 3. 2, maka dapat diuraikan masing-masing alur dalam penelitian adalah sebagai berikut:

1. *Credit Card Approvals Dataset*

Pada tahap ini penulis mengumpulkan data persetujuan kartu kredit yang diperlukan untuk penelitian. Kumpulan *dataset* yang digunakan bersumber dari website *kaggle* yang berjudul “*Credit Card Approvals (Clean Data)*” yang berisi sebanyak 690 data.

2. Proses KDD

Pada tahapan ini penulis menggunakan proses *KDD (Knowledge Discovery in Database)* yang dapat dijelaskan sebagai berikut:

a. *Data Selection dan Preprocessing*

Pemilihan data berupa sekumpulan data (*dataset*) yang perlu dilakukan sebelum tahap penggalian informasi dalam *data mining*. Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *Knowledge Discovery in Database (KDD)*. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang *noise*, dan memperbaiki kesalahan pada data, seperti kesalahan dalam pengetikan. Setelah memperoleh data bersih, lakukan pemilihan data jika diperlukan, bisa menggunakan *forward selection* atau *backward selection*.

b. *Transformation*

Melakukan perubahan isi daripada nilai atribut yang digunakan dan menghapus atribut yang tidak digunakan dalam proses *data mining*

nanti, Selanjutnya memperjelas nama dari atribut yang akan digunakan.

c. *Data Mining*

Data Mining adalah pencarian pola atau informasi yang menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Pemilihan metode atau *algoritma* yang tepat sangat bergantung pada tujuan *KDD*. Pada tahap ini penulis menggunakan 2 jenis *tools*, yaitu *WEKA* dan *RapidMiner*. Data yang digunakan penulis pada pengujian ini akan dilakukan split data untuk *data training* dan *data testing* sebanyak 60%:40%, 70%:30% dan 80%:20%. Metode yang digunakan dalam penelitian ini adalah *Algoritma Naïve Bayes* untuk memprediksi tingkat persetujuan kartu kredit.

d. *Evaluation*

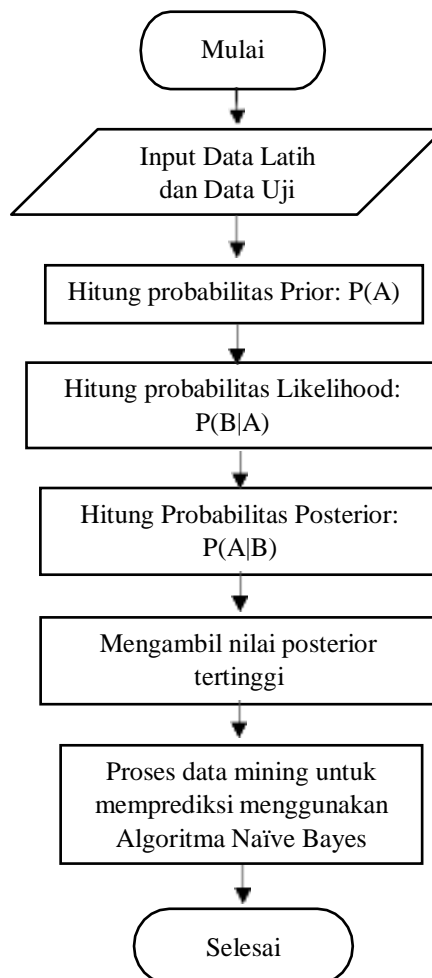
Pola informasi yang dihasilkan dari proses *data mining* harus disajikan dalam format yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini memeriksa apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau *hipotesis* yang ada.

3. Pembuatan Laporan

Pada tahap ini penulis membuat laporan dari hasil keseluruhan proses penelitian dan pengujian data yang telah dilakukan oleh penulis.

3.4 ANALISA TEKNIK KLASIFIKASI

Pada bagian ini data dan informasi yang diperoleh dan diproses dengan menggunakan metode *Algoritma Naïve Bayes* untuk mendapatkan hasil yang sesuai. Berikut tahapan klasifikasi dengan menggunakan *Algoritma Naïve Bayes*.



Gambar 3. 3 Desain Sistem Klasifikasi Dengan *Algoritma Naïve Bayes*

Pada proses ini akan dibuat suatu perancangan dari sistem klasifikasi kartu kredit melalui penerapan *Naïve Bayes* dengan menggunakan *flowchart* pada gambar 3. 3 sebagai penjelasan kebutuhan fungsional.

Tahapan-tahapan dari metode *Naïve Bayes* yang dipakai pada penelitian ini terdiri dari 3 proses perhitungan utama. Yang pertama adalah menghitung probabilitas *Prior*, kemudian menghitung probabilitas *Likelihood*, lalu dilanjutkan dengan menghitung probabilitas *Posterior*. Setelah itu dilakukan pengambilan nilai *Posterior* tertinggi, kemudian dilakukan proses *data mining* untuk memprediksi tingkat persetujuan kartu kredit dengan menggunakan *algoritma Naïve Bayes*.

3.5 ALAT PENELITIAN

Penulis mengolah data dan bahan penelitian dengan menggunakan beberapa alat dan perangkat, antara lain:

1. *Hardware* sebagai berikut:
 - a. Komputer, dengan processor Ryzen 7 2700X (3.70 GHz), RAM 16GB, SSD 1 TB, VGA RTX 2060
 - b. Monitor Xiaomi 34 inci
 - c. Mouse Press Play
 - d. Keyboard HyperX
2. *Software* dengan keterangan sebagai berikut:
 - a. OS Windows 10 (64 bit)
 - b. *Microsoft Excel & Word 2016*
 - c. *WEKA*
 - d. *RapidMiner Studio Professional 7.1.001*.