

BAB II

TINJAUAN PUSTAKA

2.1 Teori

Penulis akan melakukan beberapa pembahasan mengenai teori yang akan berkaitan dengan penelitian ini, dimana nanti akan bermanfaat dalam membantu proses pelaksanaan penelitian serta membantu dalam memahami lebih jauh hal yang nantinya akan berkaitan dengan penelitian yang dilakukan. Teori yang akan dibahas terdiri dari teori umum dan teori khusus.

2.1.1 Teori Umum

Berikut adalah penjelasan mengenai teori umum yang berkaitan dengan penelitian yang dilakukan.

2.1.1.1 *Knowledge Discovery In Database (KDD)*

Knowledge Discovery In Database atau yang biasanya disebut KDD adalah sebuah proses untuk melakukan ekstraksi atau penggalian terhadap Informasi pengetahuan baru yang bernilai dari sekumpulan besar data yang tersimpan di dalam basis data, gudang data ataupun sarana penyimpanan data informasi lainnya (Pratiwi, Yunita, & Andrianto, 2018). Menurut pendapat Nurachim (2019) *Knowledge Discovery In Database (KDD)* yaitu suatu yang proses terstruktur yang bertujuan untuk melakukan indentifikasi terhadap pola dalam suatu data yang masif

dan rumit, dengan pola data yang bersifat valid, terkini serta bermanfaat dan bisa dengan mudah dimengerti.

Berdasarkan tanggapan yang telah dinyatakan diatas dapat disimpulkan bahwa KDD merupakan sebuah teknik yang diterapkan untuk memperoleh pengetahuan dan pola dari sekumpulan data yang besar dan kompleks yang tersimpan dalam sebuah *database* dimana hasil pola data yang didapatkan harus valid, baru dan bermanfaat dalam membantu proses pengambilan keputusan.

Tahapan proses KDD secara keseluruhan salah satunya adalah *data mining*, KDD dan *data mining* seringkali dipersamakan dalam proses penggalian data informasi dan pola tersembunyi dalam suatu basis data yang masif, Pada dasarnya KDD dan *data mining* tersebut mempunyai kaitan dengan satu sama lain tetapi konsepnya tidaklah sama.

Menurut Mardi (2016). berikut adalah proses tahapan dalam KDD secara umum yaitu :

1. *Data Selection*

Adalah sebuah proses dimana dilakukanya pemilihan atau seleksi data dalam sekumpulan data operasional, dimana proses seleksi data harus dilaksanakan sebelum proses *mining* terhadap data informasi dalam KDD dijalankan.

2. *Cleaning / Pre-processing*

Proses tahapan yang harus dilakukan setelah seleksi data yaitu melakukan *cleaning* terhadap data informasi yang akan menjadi prioritas utama dalam KDD. Tahapan pembersihan biasanya terdiri dari atas kegiatan seperti menghapus duplikasi data, melakukan pemeriksaan inkonsistensi data, mengubah kesalahan data dan memperbanyak data yang telah ada dengan dukungan data lainnya yang berkaitan dengan KDD.

3. *Transformation*

Tahapan transformasi adalah tahapan dimana data yang dipilih mengalami proses transformasi yang disebut *coding*, supaya data sesuai dengan proses tahapan dalam *data mining*.

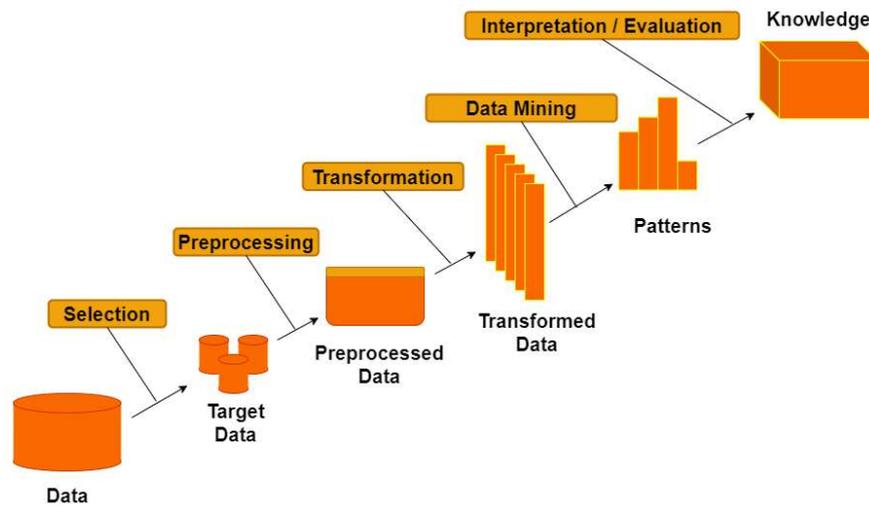
4. *Data Mining*

Data mining merupakan sebuah proses tahapan dimana dilakukannya kegiatan pencarian terhadap *pattern*, *rules* dan *information* dalam data yang telah dipilih dengan memanfaatkan metode, teknik dan algoritma yang terdapat dalam data mining sesuai dengan tujuan yang ingin dicapai dalam KDD.

5. *Interpretation / Evaluation*

Pada tahap akhir ini adalah tahap dimana dilakukannya proses penyajian hasil dari pengujian proses *data mining* yang berupa sebuah pola informasi yang disajikan dalam bentuk yang gampang dipahami oleh pihak berkepentingan.

Berikut dibawah ini adalah ilustrasi keluruhan tahapan proses dari *Knowledge Discovery in Database (KDD)*.



Gambar 2.1 Tahapan Proses dalam *Knowledge Discovery In Database (KDD)*

Sumber : (Penulis, 2021)

2.1.1.2 *Data Mining*

Data mining merupakan suatu tahapan proses yang mempunyai tujuan untuk menemukan relasi baru yang mempunyai arti dengan cara memisahkan sekumpulan besar data dalam ukuran masif yang tersimpan dalam *database* menggunakan metode inferensi pola, metode statistik dan matematis. Pada *data mining* dalam prosesnya akan dilakukan ekstraksi informasi yang bernilai dengan melakukan analisa terhadap data informasi untuk menemukan *patterns* atau relasi keterkaitan tertentu dari data-data yang berukuran masif (Iriadi & Nuraeni, 2016). Menurut Hadi (2018) *data mining* yaitu proses yang memanfaatkan teknologi

machine learning dalam menganalisis serta melakukan ekstraksi pengetahuan secara otomatis, selain itu *data mining* juga digunakan untuk melakukan proses pencarian terhadap kemungkinan atau pola yang diharapkan dari sebuah *database* yang berukuran besar untuk mendukung proses pengambilan keputusan pada waktu mendatang. Selain pendapat diatas Elisa (2017) juga pernah menjelaskan bahwa *data mining* itu adalah sebuah proses yang terdiri dari kegiatan seperti penggalian dan penambangan pengetahuan dari sekumpulan data yang berjumlah besar pada sebuah basis data, dimana tujuan dilakukannya penggalian data adalah untuk memperoleh sebuah wawasan baru yang tersembunyi dalam sebuah *database*.

Berdasarkan pandangan yang telah diungkapkan diatas dapat ditarik kesimpulan bahwa *data mining* itu merupakan sebuah teknik ataupun metode yang digunakan dalam mengumpulkan informasi yang mempunyai nilai dan manfaat yang masih belum diketahui sebelumnya dalam sekumpulan informasi yang jumlahnya sangat besar, dimana kumpulan informasi tersebut masih belum dimanfaatkan secara optimal. Dimana dalam proses pengumpulan informasi tersebut juga memanfaatkan bantuan dari beberapa teknologi seperti *machine learning* dan lainnya dalam mengumpulkan informasi, hasil informasi yang diperoleh bisa berupa aturan *pattern* ataupun relasi antar variabel tertentu.

2.1.1.3 Pengelompokan Data Mining

Menurut Zaman (2016) *data mining* bisa dikategorikan dalam beberapa kelompok berdasarkan tugas masing-masing yaitu sebagai berikut :

1. Deskripsi

Deskripsi adalah metode yang digunakan untuk memberikan penjelasan atau gambaran tentang pola dan tren yang ada pada data yang disimpan.

2. Estimasi

Estimasi sangat mirip dengan klasifikasi, kecuali bahwa variabel target untuk estimasi adalah angka dan bukan kategori atau model yang dibentuk menggunakan kumpulan data lengkap yang dilengkapi dengan nilai variabel sasaran sebagai nilai prediksi.

3. Prediksi

Prediksi melakukan prakiraan terhadap nilai yang masih belum diketahui dan memprediksi nilai masa depan.

4. Klasifikasi

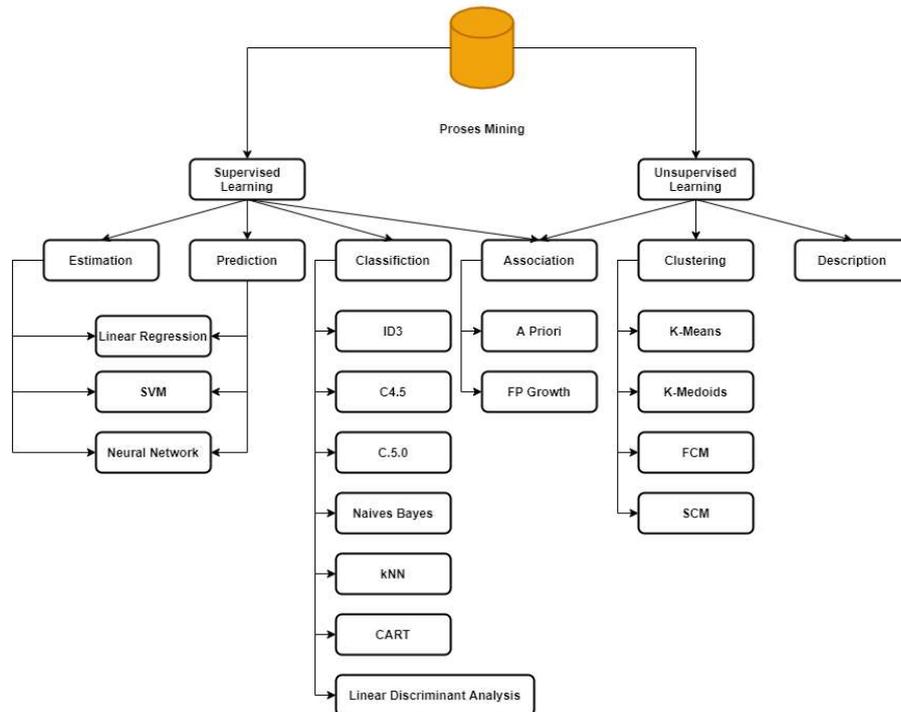
Klasifikasi mempunyai *item* variabel kategori seperti dalam pengelompokan pendapatan yang dapat dibagi menjadi tiga kelompok item yaitu tinggi, sedang, dan rendah.

5. Pengklasteran

Clustering adalah proses yang terdiri dari aktivitas-aktivitas seperti pengelompokan *record*, *monitoring* serta melakukan pembentukan *class object* yang mempunyai kemiripan.

6. Asosiasi

Asosiasi berkewajiban untuk menemukan atribut yang terjadi atau muncul pada waktu yang sama. Dalam industri bisnis lebih sering dikenal sebagai analisis keranjang belanja.



Gambar 2.2 Pengelompokan Teknik *Data Mining*

Sumber : (Penulis, 2021)

2.1.1.4 Klasifikasi

Klasifikasi adalah sebuah teknik yang digunakan dalam *data mining* untuk menyajikan data dalam bentuk kategori yang telah ditentukan atau yang disebut *class*, dimana pada teknik klasifikasi terdapat sebuah proses pencarian terhadap model ataupun fungsi yang berguna dalam mengilustrasikan dan mengelompokkan *class* dari suatu data supaya bisa dimanfaatkan untuk memprediksi *class* dari objek

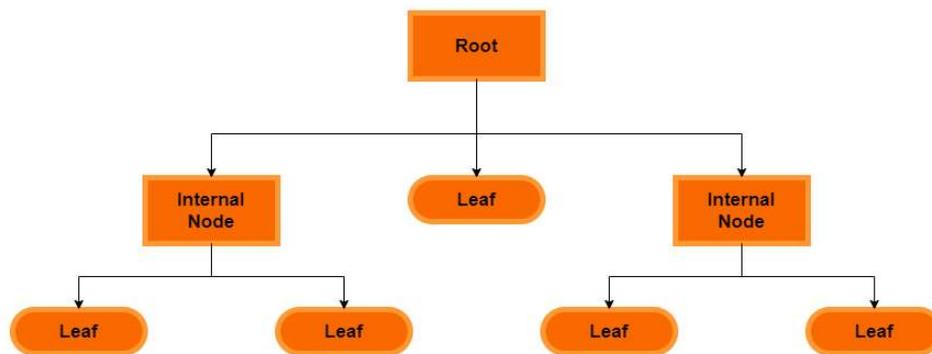
yang *class* labelnya masih belum diketahui (Rafiska, Defit, & Nurcahyo, 2018). Sedangkan menurut Romli & Zy (2020) klasifikasi merupakan jenis teknik pengolahan data yang digunakan untuk memisahkan sebuah objek menjadi sejumlah kelas yang berbeda sesuai dengan jumlah kelas yang diharapkan. Selain itu, teknik ini juga digunakan untuk menemukan pola yang dapat memisahkan kelas data satu sama yang lain untuk menentukan objek yang termasuk dalam kategori tertentu, yaitu membandingkannya dengan ciri-ciri variabel kelompok yang telah ditentukan.

2.1.1.5 Pohon Keputusan

Pohon keputusan atau yang dikenal dengan *Decision Tree* merupakan sebuah susunan pohon yang dimana terdiri dari *internal node* (*node* bukan daun) yang masing-masing mewakili sebuah variabel atribut, dan masing-masing cabang adalah perwakilan dari keadaan variabel yang merupakan proses dari hasil pembagian uji coba, sedangkan *leaf node* (*node* daun) berperan dalam mewakili kategori *class* tertentu, dan pada tingkatan node paling tinggi dalam sebuah pohon keputusan yaitu *root node* (*node* akar) yang merupakan variabel atribut yang mempunyai pengaruh paling besar pada suatu *class* (Salim & Giap, 2017).

Hal paling utama yang harus diperhatikan dalam pembangunan pohon keputusan yaitu pada proses pemisahan kriteria dimana termasuk juga kriteria dalam membuat sebuah cabang serta indikator yang akan digunakan untuk menghentikan pencabangan, *decision tree* dibangun dengan memakai kumpulan dari *data training* yang dijadikan sebagai data pengkajian (Hendrian, 2018).

Secara umum pohon keputusan memperoleh sebuah hasil dengan dilakukannya proses pencarian dari tingkat tertinggi sampai yang paling rendah dan pada saat melakukan proses klasifikasi terhadap data yang tidak dikenal maka akan dilakukan pengujian terhadap nilai atribut dengan memeriksa jalan dari *node root* hingga *node leaf* lalu dilakukan prediksi terhadap *class* yang dimiliki oleh suatu data baru tertentu (Salim & Giap, 2017).



Gambar 2.3 Model Pohon Keputusan

Sumber : (Penulis, 2018)

2.1.1.6 Algoritma C4.5

Algoritma C4.5 adalah salah satu model yang digunakan dalam metode klasifikasi *data mining*, dimana model ini digunakan untuk membentuk sebuah pohon keputusan serta melakukan proses klasifikasi data baik untuk yang menggunakan angka maupun kategori. Hasil yang didapatkan setelah dilakukannya tahapan proses klasifikasi yaitu berbentuk aturan pola yang bisa dipakai untuk

melanjutkan tahapan proses prediksi menggunakan hasil nilai prediksi yang didapatkan dari *record* data yang lama hingga yang terkini.

Model algoritma C4.5 pada dasarnya merupakan versi terbaru dari hasil pengembangan versi lamanya yaitu algoritma ID3 (*Iterative Dichotomiser 3*), yang dimana sehingga membuat algoritma C4.5 mempunyai kemampuan untuk mengatasi masalah seperti ketiadaan data dan data yang bersifat pecahan. Menurut Rahman (2015) Algoritma C4.5 adalah salah satu metode yang dapat digunakan dalam membangun sebuah pohon keputusan sesuai dengan data yang telah disiapkan. Dari penjelasan di atas bisa disimpulkan bahwa algoritma C4.5 pada dasarnya adalah model yang digunakan untuk membantu menyusun struktur pohon keputusan sesuai dengan pemilihan atribut dengan tingkat prioritas paling tinggi atau nilai dari *gain* tertinggi sesuai dengan nilai *entropy* atribut yang dijadikan sebagai proses atribut klasifikasi.

Berikut adalah penjelasan tentang tahapan proses umum dalam membangun sebuah struktur dari pohon keputusan yaitu :

1. Tahapan awal yang perlu dilakukan yaitu melakukan pememilihan atribut atau deskripsi data yang akan dijadikan sebagai akar dari pohon keputusan.
2. Setelah itu dilanjutkan dengan membuat struktur cabang dari pohon keputusan dengan nilai.

3. Tahapan langkah ketiga yaitu lakukan proses pembagian atau pemberian kasus permasalahan pada struktur cabang yang dimiliki oleh pohon keputusan.
4. Dan pada langkah terakhir yang perlu dilakukan adalah mengulangi proses dari langkah pertama sampai akhir bagi setiap cabang hingga kasus permasalahan pada cabang mempunyai kelas atau kategori yang sama.

Ketika langkah pertama dilakukan dalam tahapan pembangunan sebuah pohon keputusan yaitu ketika melakukan seleksi atribut atau deskripsi data dapat dilakukan dengan mencari *gain values* dengan hasil yang paling tinggi dari semua nilai yang dimiliki oleh atribut yang ada. Berikut adalah rumus yang digunakan untuk melakukan proses pencarian terhadap nilai *gain* tertinggi yaitu :

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Rumus 2.1 Perhitungan *Gain*

Berikut adalah deskripsi simbol dari rumus :

S : himpunan kasus

A : atribut atau deskripsi data

n : jumlah partisi yang ada pada atribut atau deskripsi data bagian (A)

$|S_i|$: total kasus yang terdapat pada partisi yang ke- i

$|S|$: total kasus yang terdapat dalam S

Sementara untuk rumus yang digunakan untuk mencari *entropy values* bisa dilihat dibawah.

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i$$

Rumus 2.2 Perhitungan *Entropy*

Berikut adalah deskripsi simbol dari rumus :

S : himpunan kasus

A : atribut atau deskripsi data

n : total partisi yang ada pada atribut atau deskripsi data bagian (S)

p_i : proporsi dari S_i kepada S

2.1.2 Teori Khusus

Berikut adalah penjelasan tentang beberapa teori khusus yang mempunyai hubungan dengan penelitian yang dilakukan.

2.1.2.1 Kredit

Kredit adalah sebuah proses dimana pihak kreditor memberikan penyediaan dana kepada pihak debitor sesuai dengan kesepakatan yang sudah disetujui

bersama, dengan syarat pihak debitor harus dapat melunasi semua utangnya bersama dengan bunga kreditnya dalam jangka waktu yang telah ditentukan sesuai hasil kesepakatan yang telah disepakati sebelumnya (Khasanah, 2019). Sedangkan menurut Rai & Purnawati (2017) kredit merupakan hak untuk mendapatkan pembayaran atau kewajiban untuk melakukan pembayaran pada waktu tertentu yang telah ditetapkan dimasa mendatang karena atas penyerahan barang atau jasa yang dilakukan pada saat ini.

Berdasarkan penjelasan di atas dapat disimpulkan bahwa kredit adalah salah satu bentuk cara yang dapat digunakan untuk memperoleh dukungan finansial melalui dana yang dipinjamkan dari pihak kreditur yang dimana nanti proses pembayarannya akan dilakukan secara angsuran pada tiap periode tertentu berdasarkan perjanjian yang telah dibuat sebelumnya.

2.1.2.2 Kelayakan Kredit

Kelayakan kredit merupakan suatu proses penilaian terhadap kemampuan pemohon kredit untuk melakukan kewajibannya dalam membayar uang yang dipinjamkan dimana kelayakan kredit biasanya akan ditentukan oleh seorang analis kredit (Dyah, Nur, & Heny, 2019). Dari penjelasan diatas dapat disimpulkan bahwa kelayakan kredit adalah kegiatan dimana dilakukannya analisis untuk memperkirakan atau melakukan pengukuran terhadap kemampuan peminjam dalam melakukan pembayaran atas dana yang telah dipinjam serta dengan bunga pinjamannya.

2.1.2.3 Prinsip 5C 7P dan 3R

Prinsip 5C 7P dan 3R adalah sebuah sistem penilaian ataupun pengukuran terhadap kelayakan kredit seorang calon debitur yang biasanya digunakan dalam dunia perbankan seperti bank ataupun lembaga peminjaman lainnya, berikut dibawah adalah penjelasan lebih dalam mengenai isi dari prinsip-prinsip diatas.

1. Prinsip dari 5C terdiri dari :

a. *Character*

Adalah penilaian terhadap sifat ataupun watak seorang calon debitur apakah layak untuk diberikan kredit ataupun tidak, biasanya hal yang dinilai adalah tentang latar belakang dari pekerjaannya maupun kehidupan pribadi seperti gaya hidup dan kondisi keluarga.

b. *Capacity*

Adalah penilaian terhadap kemampuan calon debitur dalam membayar kredit yang telah dipinjamkan, biasanya hal yang dinilai adalah terkait tentang pendapatan yang dihasilkan oleh calon debitur, pengalaman yang dimiliki dalam mengelola bisnisnya, yang sehingga dapat diketahui kemampuan dalam mengembalikan kredit yang dipinjamkan.

c. *Capital*

Adalah penilaian terhadap pemanfaatan atas modal yang dimiliki apakah sudah efektif atau belum, yang dimana informasi tersebut dapat dilihat dari laporan keuangan perusahaan neraca dan laporan

laba rugi yang nantinya akan dilakukan proses pengukuran seperti dari segi likuiditas, solvabilitas, rentabilitas dan lainnya.

d. *Condition*

Adalah penilaian terhadap kondisi dari ekonomi, politik dan sosial untuk sekarang dan juga untuk masa depan apakah bidang usaha yang akan diinvestasi benar-benar mempunyai prospek yang memuaskan, sehingga masalah akan kredit untuk muncul menjadi rendah.

e. *Collateral*

Adalah penilaian terhadap jaminan yang dititipkan oleh calon debitur baik yang bersifat fisik maupun nonfisik, dan untuk nilai dari jaminanya harus lebih tinggi dari pinjaman yang diberikan.

2. Prinsip dari 7P terdiri dari :

a. *Personality*

Adalah penilaian terhadap personalitas ataupun kepribadian dan kesehatan mental dari seorang calon debitur pada masa lalu maupun sekarang.

b. *Party*

Adalah penilaian terhadap calon debitur dengan mengklasifikasikannya ke dalam golongan ataupun kategori-kategori tertentu, calon debitur yang tergolong dalam kategori tertentu akan mendapatkan pelayan dan fasilitas yang berbeda-beda dari bank.

c. *Purpose*

Adalah penilaian yang dilakukan untuk mengetahui tujuan ataupun alasan calon debitur melakukan pengajuan kredit.

d. *Prospect*

Adalah penilaian yang dilakukan untuk mengetahui prospek dari usaha ataupun bisnis dari calon debitur apakah akan menguntungkan atau tidak pada masa yang akan datang.

e. *Payment*

Adalah penilaian yang dilakukan untuk mengetahui bagaimana cara calon debitur untuk mengembalikan dana yang telah dipinjamkan atau dari mana saja sumber dana yang akan digunakan melunaskan kredit yang dipinjam.

f. *Profitability*

Adalah penilaian terhadap kemampuan calon debitur dalam menghasilkan keuntungan atau kekayaan ekonomi dalam periode waktu tertentu.

g. *Protection*

Adalah penilaian yang dilakukan untuk mengetahui bagaimana cara calon debitur dapat menjamin keamanan dan perlindungan atas dana atau kredit yang dipinjamkan. biasanya perlindungan yang ditawarkan dapat berupa jaminan barang, asuransi maupun saksi.

3. Prinsip dari 3R terdiri dari :

a. *Returns*

Adalah penilaian terhadap hasil ataupun keuntungan yang akan dicapai oleh usaha atau bisnis dari calon debitur setelah kredit diberikan, apakah nanti hasil yang dicapai dapat membayar kembali pinjaman yang telah dipinjamkan.

b. *Repayment*

Adalah penilaian yang dilakukan untuk mengetahui kemampuan calon debitur untuk melunaskan pinjaman kreditnya pada jangka waktu yang telah ditentukan.

c. *Risk Bearing Ability*

Adalah penilaian yang dilakukan untuk mengetahui kemampuan perusahaan calon debitur dalam menghadapi atau menangani resiko yang mungkin akan terjadi.

2.1.2.4 Aplikasi WEKA

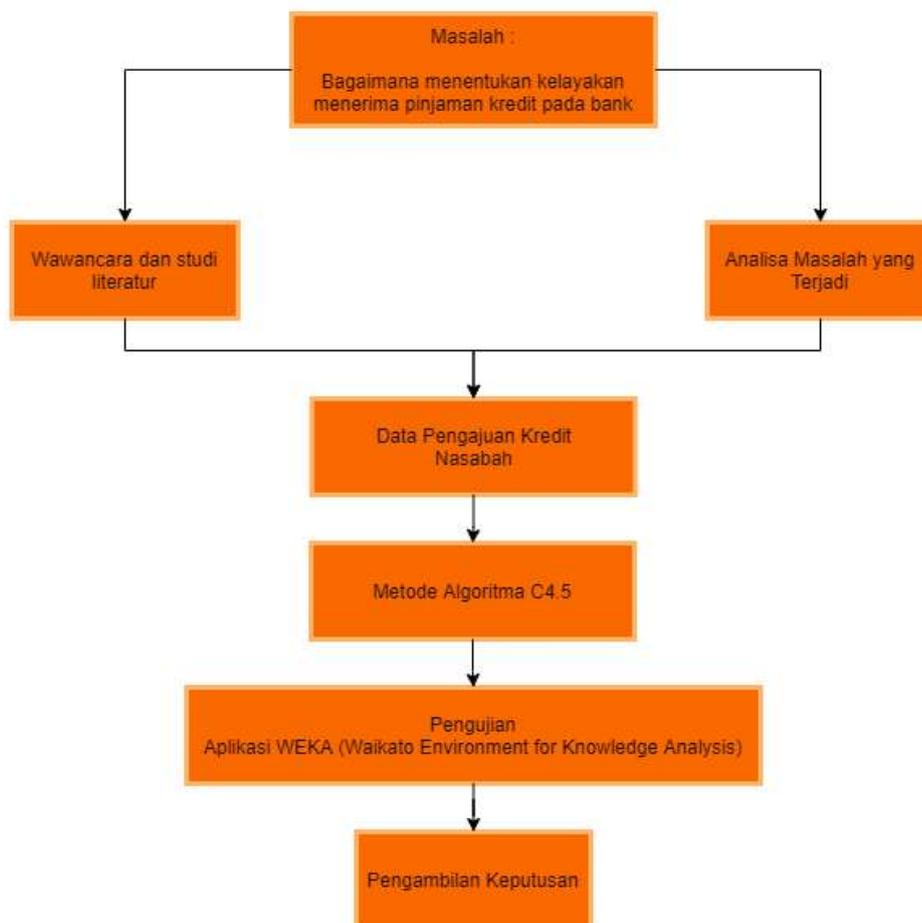
Waikato Environment for Knowledge Analysis atau yang biasanya lebih dikenal sebagai WEKA merupakan sebuah paket *tools machine learning* yang sangat mudah untuk digunakan. Program WEKA dirancang oleh Universitas Waikato, New Zealand untuk digunakan dalam kegiatan penelitian dan pendidikan. WEKA mempunyai sekumpulan algoritma dan fungsi visualisasi yang dapat digunakan dalam proses analisis data serta pemodelan prediktif, dan salah satu algoritma yang terdapat pada WEKA adalah pembelajaran mesin atau lebih dikenal dengan *Machine Learning* yang dimana membuatnya dapat digunakan untuk menyelesaikan masalah data mining. Perangkat *software* ini dibuat dengan menggunakan bahasa pemrograman berorientasi objek yaitu Java sehingga program

ini dapat dijalankan pada semua *platform* dan WEKA juga telah menerapkan semua teknik untuk klasifikasi dan regresi yang dimana yaitu seperti pohon keputusan, *Support Vector Machines* (SVM), *logistic* dan *linier*, *multi layers perceptrons* dan metode *nearest neighbour* (Sartika & Sensuse, 2017). Program WEKA juga mampu untuk melakukan pengindentifikasian terhadap informasi yang tersembunyi dalam *database* dan *file system* degan memilih melalui *user interface* yang cepat dan mudah untuk dioperasikan (Ainurrohmah, 2021).

2.2 Kerangka Pemikiran

Penulis telah membuat sebuah kerangka pemikiran sesuai dengan penelitian yang akan dilakukan, tujuan dibuatnya kerangka pemikiran ini adalah untuk memperjelas langkah – langkah penyelesaian masalah yang akan diambil dan dimplementasikan dalam menyelesaikan masalah penelitian.

Berikut dibawah adalah ilustrasi dari kerangka pemikiran yang akan digunakan.



Gambar 2.4 Kerangka Pemikiran

Sumber : (Penulis, 2021)

2.3 Hipotesis Penelitian

Berikut dibawah ini adalah uraian tentang hipotesis atau dugaan sementara dari hasil penelitian yang dimana masih diperlukan proses pengujian lebih lanjut tentang kebenarannya yaitu :

1. *Data Mining* yang menggunakan model algoritma C4.5 dapat digunakan dalam menganalisis kelayakan menerima pinjaman kredit pada bank.
2. Hasil yang diperoleh dari penelitian dapat meningkatkan kualitas hasil analisa dalam menentukan kelayakan menerima pinjaman kredit pada bank.

2.4 Penelitian Terdahulu

Berikut dibawah ini adalah beberapa penelitian yang pernah dilakukan oleh peneliti sebelumnya, dimana penelitian tersebut akan dijadikan sebagai bahan referensi dan dasar dalam melakukan penelitian karena penelitian tersebut mempunyai kaitan langsung dengan topik penelitian yang nanti akan dilakukan yaitu sebagai berikut :

1. Penelitian yang dilakukan oleh (Dyah et al., 2019) yang meneliti tentang “Penggunaan pohon keputusan dengan algoritma C4.5 dalam menganalisis kelayakan kredit nasabah” menjelaskan bahwa dari hasil analisis dan pembahasan menunjukkan bahwa penggunaan algoritma C4.5 sangat efektif dalam menentukan kelayakan kredit nasabah karena didapatkannya sebuah pola penentuan.
2. Penelitian yang dilakukan oleh (Khasanah, 2017) yang meneliti tentang “Penentuan kelayakan kredit dengan model algoritma C4.5” menjelaskan bahwa penggunaan model algoritma C4.5 dalam menentukan kelayakan kredit calon nasabah lancar dan bermasalah menunjukkan nilai akurasi

yang cukup bagus yaitu 88.52% dengan 270 jumlah *data training* milik nasabah.

3. Penelitian yang dilakukan oleh (Setiawan, 2020) yang meneliti tentang “Implementasi algoritma C4.5 dalam memprediksi kelayakan diberikanya kredit nasabah” menjelaskan bahwa nilai akurasi dari penelitian perlu ditingkatkan, karena nilai akurasi dapat dipengaruhi oleh banyak hal diantaranya yaitu jumlah data trainingnya serta detail dari isi *data training*.
4. Penelitian yang dilakukan oleh (Santoso & Sekardiana, 2019) yang meneliti tentang “Pengaplikasian model algoritma C4.5 dalam membantu proses penentuan hak pemberian kredit” menjelaskan bahwa algoritma C4.5 terbukti akurat dalam digunakan dalam penentuan permohonan kelayakan kredit karena didapatkan nilai akurasi *data training* sebesar 95% dengan error rate yang sebesar 6.6%
5. Pada penelitian yang dilakukan oleh (Hadi, 2018) yang berjudul tentang “Penerapan Metode Algoritma C4.5 dalam Menganalisa Pangajuan Kredit pada Koperasi Jasa Keuangan (KJKS) Syariah Kelurahan Limau Manis Selatan” menyimpulkan bahwa aplikasi *data mining* yang dibuat berdasarkan algoritma C4.5 mampu untuk mengurangi resiko terjadinya kredit macet, dengan dilakukan perhitungan dapat menghasilkan *decision tree* yang berisi *rules* atau aturan yang dapat digunakan dalam penentuan pengajuan kredit calon debitur.

6. Pada penelitian yang dilakukan oleh (Pratama, Kurniawati, Larbona, & Haryanti, 2019) yang meneliti tentang “Prediksi masalah kredit macet dengan melakukan klasifikasi nasabah menggunakan algoritma C4.5” menyimpulkan bahwa perhitungan nilai *entropy* dan *gain* dari algoritma C4.5 pada masing-masing atribut dan sub atribut bisa dengan baik mengidentifikasi kelayakan kredit, selain itu karakteristik data yang diklasifikasikan menggunakan algoritma C4.5 juga sangat jelas dan mudah dipahami baik struktur pohon keputusan yang dihasilkan maupun *rule if-then*.
7. Pada penelitian yang dilakukan oleh (Hartini & Kurahman, 2020) yang berjudul tentang “Sistem Pakar Pemilihan Calon Debitur Kredit Motor Dengan Algoritma C4.5 Pada PT. Federal International Finance” menyimpulkan bahwa dengan perpaduan antara sistem pakar dan data mining bisa digunakan untuk mendukung proses pemecahan masalah diberbagai bidang, yang dimana salah satunya yaitu melakukan klasifikasi terhadap calon debitur yang ingin melakukan pinjaman kredit.
8. Pada penelitian yang dilakukan oleh (Setianingrum, Hindayanti, Cahya, & Purnia, 2021) yang meneliti tentang “Perbandingan antara metode algoritma C4.5 dan metodo algoritma K-NN dalam menganalisa kredit macet” menjelaskan bahwa antara kedua metode tersebut berdasarkan hasil perbandingan yang dilakukan menunjukkan bahwa metode algoritma C4.5 lebih akurat dari pada KNN dalam hal menentukan kredit macet.

9. Pada penelitian yang dilakukan oleh (Aji & Sunyoto, 2021) berdasarkan hasil penelitian yang dilakukan menunjukkan bahwa dengan di implementasikannya algoritma C4.5 dapat membantu dalam menganalisis performa dari siswa, karena dengan algoritma C4.5 dapat dengan mudah untuk membentuk sebuah model klasifikasi.
10. Pada penelitian yang dilakukan oleh (Amin, Indwiarti, & Sibaroni, 2015) yang berjudul tentang “*Implementation of Decision Tree Using C4.5 Algorithm in Decision Making of Loan Application by Debtor*” berdasarkan hasil dari penelitian yang dilakukan menunjukkan bahwa jumlah *data training set* yang besar tidak selalu akan memberikan hasil akurasi dan presisi yang tinggi karena itu juga tergantung pada kualitas dari *data training set* yang digunakan.
11. Pada penelitian yang dilakukan oleh (Awaludin, Gerhana, Darmalaksana, & Rahman, 2021) yang berjudul tentang “*Comparison of Decision Tree C4.5 Algorithm with K-Nearest Neighbor (KNN) Algorithm in Hadith Classification*” berdasarkan hasil dari penelitian yang dilakukan menunjukkan bahwa hasil akurasi rata-rata yang dihasilkan dari penggunaan algoritma C4.5 sebanyak lima kali percobaan lebih tinggi dari metode KNN, dengan tingkat akurasi sebesar 70.53% untuk C4.5 dan sebesar 66.36% untuk KNN. Kemudian juga dikatakan bahwa nilai akurasi juga dapat dipengaruhi oleh kesalahan pada proses penyediaan label kelas di tahapan persiapan data yang pada akhirnya akan mempengaruhi hasil prediksi yang dihasilkan.

12. Pada penelitian yang dilakukan oleh (Siswanto et al., 2019) yang berjudul tentang “*The Feasibility of Credit Using C4.5 Algorithm Based on Particle Swarm Optimization Prediction*” berdasarkan hasil penelitian yang dilakukan didapatkan tingkat akurasi yang tinggi yaitu sebesar 90.99% dengan menggunakan algoritma C4.5 dalam memprediksi kelayakan kredit, dan dengan bantuan dari PSO telah meningkatkan tingkat akurasi menjadi sedikit lebih tinggi dari sebelumnya yaitu sebesar 91.18% walaupun peningkatannya tidak bisa dikatakan sangat besar tetapi tetap sangat berpengaruh terhadap kualitas dari prediksi yang dilakukan.