

**PENERAPAN DATA MINING UNTUK MEMPREDIKSI  
KELAYAKAN KREDIT NASABAH PADA PT. BPR  
INDOBARU FINANSIA MENGGUNAKAN  
ALGORITMA C4.5**

**SKRIPSI**



**Oleh:**

**Melva Marsyah Adella Tuhumena  
181510087**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
2022**

**PENERAPAN DATA MINING UNTUK MEMPREDIKSI  
KELAYAKAN KREDIT NASABAH PADA PT. BPR  
INDOBARU FINANSIA MENGGUNAKAN  
ALGORITMA C4.5**

**SKRIPSI**

**Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana**



**Oleh**

**Melva Marsyah Adella Tuhumena  
181510087**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNIK DAN KOMPUTER  
UNIVERSITAS PUTERA BATAM  
2022**

## SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini saya:

Nama : Melva Marsyah Adella Tuhumena  
NPM : 181510087  
Fakultas : Teknik dan Komputer  
Program Studi : Sistem Informasi

Menyatakan bahwa “Skripsi” yang saya buat dengan judul:

### **Penerapan Data Mining Untuk Memprediksi Kelayakan Kredit Nasabah Pada PT. BPR Indobaru Finansia Menggunakan Algoritma C4.5**

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, didalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun

Batam, 08 Agustus 2022



**Melva Marsyah Adella Tuhumena**  
**181510087**

**PENERAPAN DATA MINING UNTUK MEMPREDIKSI  
KELAYAKAN KREDIT NASABAH PADA PT. BPR  
INDOBARU FINANSIA MENGGUNAKAN  
ALGORITMA C4.5**

**SKRIPSI**

**Untuk memenuhi salah satu syarat  
memperoleh gelar Sarjana**

**Oleh  
Melva Marsyah Adella Tuhumena  
181510087**

**Telah disetujui oleh Pembimbing pada tanggal  
seperti tertera dibawah ini**

**Batam, 08 Agustus 2022**



**Sasa Ani Arnomo, S.Kom., M.SI., Ph.D.  
Pembimbing**

## ABSTARK

Bank adalah lembaga yang didedikasikan untuk layanan keuangan. Salah satu mitra pemerintah yang dituntut untuk mensukseskan pembangunan, melalui jasa-jasa layanan kredit yang diperlukan bagi masyarakat dalam melakukan kegiatan usaha. Kredit adalah uang yang diberikan oleh bank berdasarkan perjanjian pinjaman. Penentuan kelayakan kredit seringkali tidak akurat, sehingga menghasilkan skor kredit yang rendah karena kesalahan marketing saat melakukan analisis. Analisis kredit dilakukan dengan menggunakan 5C, yang meliputi *character, capacity, capital, collateral, dan condition of economy*. Namun, metode ini menimbulkan masalah dengan proses yang lama dikarenakan banyaknya calon nasabah. Oleh karena itu, pihak marketing harus terlebih dahulu mengevaluasi dan menentukan keputusan diawal sebelum memberikan kredit. Dalam penelitian ini akan dilakukan analisis data pemohon kredit pinjaman menggunakan klasifikasi data mining yakni algoritma C4.5 dengan menggunakan beberapa atribut diantaranya pekerjaan, karakter, pendapatan, plafond pinjaman, dan jaminan yang akan dihitung dengan perhitungan manual dan perhitungan software weka untuk mempermudah uji kelayakan pemberian kredit pinjaman dalam mendapatkan hasil yang akurat dan cepat. Dari 115 jumlah sampel kasus yang terdiri dari 74 kasus layak dan 41 kasus tidak layak data didapatkan dari PT BPR Indobaru Finansia. Dari perhitungan metode algoritma C4.5 didapatkan hasil akhir dari pohon keputusan faktor utama dalam memprediksi pemberian kelayakan kredit yang menjadi acuan utama dalam memberikan kredit adalah karakter, pendapatan dan jaminan dan menghasilkan 7 rule dengan jumlah class layak sebanyak 4 rule dan jumlah class yang tidak layak sebanyak 3 rule serta melakukan perhitungan confusion matriks untuk menghitung tingkat akurasi sebesar 98% sehingga dapat disimpulkan bahwa penelitian ini dapat membantu dan mendukung dalam mengambil keputusan layak dan tidak layak dalam pemberian kredit pinjaman cepat dan efektif.

Kata kunci: Penerapan, Data Mining, Algoritma C4.5, kredit pinjaman, Prediksi kelayakan kredit

## **ABSTRACT**

*A bank is an institution dedicated to financial services. One of the government partners who is required to succeed in development, through credit services needed for the community in carrying out business activities. Credit is money provided by a bank under a loan agreement. Creditworthiness determination is often inaccurate, resulting in a low credit score due to marketing errors when conducting analysis. Credit analysis is carried out using the 5C, which includes character, capacity, capital, collateral, dan condition of economy. However, this method causes problems with the long process due to the large number of potential customers, therefore, the marketing party must first evaluate and determine the decision at the beginning before giving credit. In this study, data analysis of loan credit applicants will be carried out using the classification of data mining, namely the C4.5 algorithm using several attributes including work, character, income, loan ceiling, and guarantees which will be calculated by manual calculations and weka software calculations to facilitate the service test of lending in obtaining accurate and fast results. Of the 115 total sample cases consisting of 74 decent cases and 41 unfit cases, data was obtained from PT BPR Indobaru Finansia. From the calculation of the C4.5 algorithm method, the final result of the decision tree of the main factors in predicting the provision of creditworthiness which is the main reference in providing credit is character, income and guarantees and produces 7 rules with the number of decent classes as many as 4 rules and the number of classes that are not as many as 3 rules and perform matrix confusion calculations to calculate an accuracy rate of 98% so that it can be concluded that this research can help and support in making feasible and unfit decisions in providing fast and effective loan credit*

*Keywords: Application, Data Mining, C4.5 Algorithm, credit loans, Creditworthiness prediction*

## KATA PENGANTAR

Puji syukur penulis ucapkan tuhan YME yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Sistem Informasi Universitas Putera Batam.

Masih banyak terdapat kekurangan dari skripsi ini suatu kritik dan saran akan senantiasa penulis terima dengan senang hati. Dengan segala keterbatasan, penulis menyadari skripsi ini terwujud dari beberapa dukungan, bimbingan, nasehat tanpa bantuan dan bantuan dari beberapa pihak. Maka, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI. selaku Rektor Universitas Putera Batam.
2. Bapak Muhammad Rasid Ridho, S.Kom., M.SI. selaku Ketua Program Studi Sistem Informasi Universitas Putera Batam.
3. Bapak. Sasa Ani Arnomo, S.Kom., M.SI., Ph.D. selaku dosen pembimbing skripsi pada program studi sistem informasi dan dosen pembimbing akademik yang telah memberi support kepada penulis untuk menyelesaikan skripsi ini.
4. Dosen dan Staff Universitas Putera Batam.
5. Kedua orang tua penulis yang selalu mendoakan dan memberikan motivasi. kepada penulis agar penelitian ini selesai tepat waktu.
6. Teman-teman seperjuangan yang bersedia membagi ilmunya dan sharing pendapat dalam pengerjaan skripsi ini.
7. Semua pihak yang telah bersedia meluangkan waktu, tenaga dan pikirannya dalam memberikan data dan informasi selama penulis membuat skripsi yang tidak dapat penulis sebutkan satu-persatu

Semoga tuhan membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam, 08 Agustus 2022

Melva Marsyah Adella Tuhumena

## DAFTAR ISI

Halaman

<b>HALAMAN SAMPUL</b> .....	i
<b>HALAMAN JUDUL</b> .....	ii
<b>SURAT PERNYATAAN</b> .....	iii
<b>HALAMAN PENGESAHAN</b> .....	iv
<b>ABSTARK</b> .....	v
<b>ABSTRACK</b> .....	vi
<b>KATA PENGANTAR</b> .....	vii
<b>DAFTAR ISI</b> .....	viii
<b>DAFTAR GAMBAR</b> .....	xi
<b>DAFTAR TABEL</b> .....	xii
<b>DAFTAR RUMUS</b> .....	xiii
<b>BAB I PENDAHULUAN</b> .....	1
1.1. Latar Belakang .....	1
1.2. Identifikasi Masalah.....	3
1.3. Batasan Masalah .....	4
1.4. Rumusan Masalah.....	4
1.5. Tujuan Penelitian .....	5
1.6. Manfaat Penelitian .....	5
1.6.1. Manfaat Teoritis.....	5
1.6.2. Manfaat Praktis .....	6
<b>BAB II TINJAUAN PUSTAKA</b> .....	7
2.1. Tinjauan Teori Umum.....	7
2.1.1. Knowledge Discovery in Database (KDD).....	7
2.2. Tinjauan Teori Khusus.....	10
2.2.1. Algoritma C4.5 .....	10
2.2.1.1. Decision Tree .....	11
2.2.2. Kredit .....	14
2.2.2.1. Perjanjian Kredit .....	14
2.2.2.2. Jaminan Kredit .....	15
2.2.2.3. Prinsip Penilaian Kredit .....	15



2.2.2.4.	Prosedur Pemberian Kredit .....	17
2.2.3.	Bank .....	18
2.2.4.	Waikato Environment for Knowledge Analysis (WEKA) .....	19
2.2.4.1.	Confusion Matrix .....	19
2.3.	Kerangka Pemikiran.....	21
2.4.	Penelitian Terdahulu.....	22
<b>BAB III METODE PENELITIAN .....</b>		<b>25</b>
3.1.	Design Penelitian .....	25
3.2.	Objek Penelitian.....	27
3.2.1.	Lokasi Penelitian.....	27
3.2.2.	Jadwal Penelitian .....	27
3.2.2.1.	Populasi.....	28
3.2.3.	Populasi dan Sampel.....	28
3.2.3.1.	Sampel.....	28
3.2.4.	Variabel Penelitian.....	28
3.3.	Teknik Pengumpulan Data.....	29
3.4.	Model Penelitian .....	30
<b>BAB IV PEMBAHASAN.....</b>		<b>32</b>
4.1.	Hasil Penelitian .....	32
4.2.	Pembahasan.....	32
4.2.1.	Melakukan Pra-Proses Data Weka.....	33
4.2.2.	Analisis Data.....	33
4.2.3.	Seleksi Data .....	33
4.2.4.	Transformasi Data.....	33
4.2.5.	Pohon Keputusan .....	35
4.2.6.	Proses Data Menggunakan Aplikasi WEKA .....	54
4.2.7.	Proses pengujian aplikasi weka .....	59
4.2.7.1.	Hasil Pengujian .....	61
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>62</b>
5.1.	Kesimpulan .....	62
5.2.	Saran .....	63
<b>DAFTAR PUSTAKA .....</b>		<b>65</b>
<b>LAMPIRAN 1 PENDUKUNG PENELITIAN .....</b>		<b>67</b>
<b>LAMPIRAN 2 DAFTAR RIWAYAT HIDUP .....</b>		<b>75</b>

**LAMPIRAN 3 SURAT KETERANGAN IZIN PENELITIAN..... 77**

## DAFTAR GAMBAR

	Halaman
<b>Gambar 2.1</b> Kerangka Pemikiran .....	21
<b>Gambar 3.1</b> Desain penelitian .....	25
<b>Gambar 4.1</b> Pohon perhitungan node .....	42
<b>Gambar 4.2</b> hasil pohon keputusan node kedua .....	48
<b>Gambar 4.3</b> Pohon keputusan node ketiga .....	53
<b>Gambar 4.4</b> Data ms. Excel .....	54
<b>Gambar 4. 5</b> Data pada notepad.....	55
<b>Gambar 4.6</b> Tampilan awal aplikasi weka.....	55
<b>Gambar 4.7</b> Tampilan memilih file .....	56
<b>Gambar 4. 8</b> Tampilan proses Pengujian data atribut.....	56
<b>Gambar 4. 9</b> Tampilan memilih metode algoritma C4.5.....	57
<b>Gambar 4. 10</b> Tampilan Proses Visualize Tree .....	58
<b>Gambar 4.11</b> Hasil pohon keputusan dari aplikasi weka.....	58
<b>Gambar 4.12</b> Tampilan confusion matrix .....	60

## DAFTAR TABEL

	Halaman
<b>Tabel 2.1</b> Confusion matrix .....	20
<b>Tabel 2.1</b> Penelitian Terdahulu .....	22
<b>Tabel 3.1</b> Jadwal penelitian .....	28
<b>Tabel 3.2</b> Variabel Penelitian .....	29
<b>Tabel 4.1</b> Klasifikasi atribut Pekerjaan .....	33
<b>Tabel 4.2</b> Klasifikasi atribut Karakter .....	34
<b>Tabel 4.3</b> Klasifikasi atribut Plafond pinjaman.....	34
<b>Tabel 4.4</b> Klasifikasi atribut Pendapatan.....	35
<b>Tabel 4.5</b> Klasifikasi atribut Jaminan.....	35
<b>Tabel 4.6</b> Hasil perhitungan node pertama .....	41
<b>Tabel 4.7</b> Hasil perhitungan node kedua .....	47
<b>Tabel 4.8</b> Hasil perhitungan node ketiga.....	52
<b>Tabel 4.9</b> Pengujian Confusion Matrix .....	60

## DAFTAR RUMUS

	Halaman
<b>Rumus 2.1</b> Perhitungan Gain .....	13
<b>Rumus 2.2</b> Perhitungan Entropy .....	13
<b>Rumus 2.3</b> Confusion matrix .....	20

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Kebijakan ekonomi yang dibuat oleh pemerintah pada bidang perbankan merupakan salah satu bidang yang mendapat perhatian pemerintah. Dalam hal ini dikarenakan masyarakat mendapat permodalan usaha lebih banyak melalui bank. Bank merupakan suatu lembaga yang bergerak dalam bidang jasa keuangan. Salah satu rekan pemerintah yang dituntut untuk mensukseskan pembangunan, melalui jasa pemberian layanan kredit yang sangat dibutuhkan masyarakat dalam menjalankan kegiatan usaha.

Kredit adalah uang yang diberikan oleh bank sebagai hasil dari perjanjian pinjaman. Pemberi pinjaman membayar bunga dan harus membayar kembali uang tersebut setelah jangka waktu tertentu. Sebelum memberikan pinjaman, bank harus menilai kemampuan nasabah untuk membayar kembali pinjamannya. Hal ini dilakukan untuk mengurangi risiko kerugian akibat nasabah tidak mampu membayar utangnya. Penentuan kelayakan kredit seringkali tidak akurat, sehingga menghasilkan skor kredit yang rendah karena kesalahan marketing saat melakukan analisis. Analisis kredit dilakukan dengan menggunakan 5C, yang meliputi *character, capacity, capital, collateral, dan condition of economy*. Namun, metode ini menimbulkan masalah dengan proses yang lama dikarenakan banyaknya calon nasabah. Oleh karena itu, pihak marketing harus terlebih dahulu mengevaluasi dan menentukan keputusan diawal sebelum memberikan kredit. Masalah ini dapat diselesaikan dengan menggunakan metode data mining. Data mining adalah studi

sistematis untuk menemukan pola dalam kumpulan data yang besar. Model ini dibangun dengan melatih algoritma data mining yang digunakan untuk menentukan apakah seorang nasabah memenuhi syarat untuk mendapatkan pinjaman atau tidak.

PT. Bank Perkreditan Rakyat Indobaru Finansial berlokasi di Jl. Laksamana Bintan, kompleks Tanah Mas Blok. C No.4-5 Kelurahan Teluk Tering, Kecamatan Batam Kota, Kota Batam. berdiri sejak tanggal 30 Maret 2009 Bank konvensional yang sektor usaha difokuskan pada penyaluran kredit kepada usaha kecil dan menengah (UMKM), termasuk investasi modal kerja, investasi, konsumsi dan jalur kredit pembiayaan. Penentuan kelayakan pemberian kredit yang dilakukan oleh PT. BPR Indobaru Finansial dengan menggunakan prinsip 5C memiliki hasil belum akurat dan proses cukup lama terlihat dengan masih adanya kredit macet untuk itu PT. BPR Indobaru Finansial mungkin dapat melakukan analisis kredit dengan menggunakan C4.5.

Algoritma C4.5 merupakan salah satu solusi yang dapat digunakan untuk menyelesaikan permasalahan operasional yang berkaitan dengan credit rating. Keluaran dari algoritma C4.5 adalah sebuah pohon keputusan terstruktur yang dapat digunakan untuk mentransformasikan dataset menjadi sebuah pohon keputusan dengan menggunakan aturan keputusan. Menentukan model pohon keputusan untuk meningkatkan akurasi dalam menganalisa kelayakan kredit yang diajukan calon debitur. Dalam melakukan berbagai fungsi untuk mengurangi risiko macet. Karakteristik yang digunakan adalah: pekerjaan, karakter, plafond pinjaman, pendapatan dan jaminan. Tujuan dari penelitian ini adalah untuk mengetahui keputusan layak dan tidak layaknya calon nasabah untuk di berikan kredit.

Permasalahan dari definisi di atas adalah masih adanya proses yang lama dan kurang akurat dalam menentukan kualitas pemberian kredit. Penggunaan algoritma C4.5 seharusnya meningkatkan akurasi pemeriksaan kredit untuk menentukan kesesuaian.

Menggunakan algoritma ini akan mempercepat proses peramalan yang paling efisien dalam menentukan pemberian kredit. Ini sangat penting karena algoritme dapat mengirim pesan lebih tepat dan efisien pada BPR Indobaru Finansial. Maka BPR Indobaru Finansial dijadikan bahan penelitian penulis untuk melakukan penelitian dengan judul: **“PENERAPAN DATA MINING UNTUK MEMPREDIKSI KELAYAKAN KREDIT NASABAH MENGGUNAKAN ALGORITMAN C4.5”**.

## **1.2. Identifikasi Masalah**

Berdasarkan uraian latar belakang diatas maka dapat di identifikasikan beberapa permasalahan yang terjadi:

1. Penentuan kelayakan kredit masih belum akurat dikarenakan pihak marketing masih belum tepat dalam menganalisis calon nasabah
2. Dengan adanya jenis data debitur yang banyak sehingga membutuhkan waktu yang Panjang untuk menentukan kelayakan kredit nasabah BPR Indobaru Finansial.
3. Menggunakan analisis kredit menggunakan prinsip 5C membutuhkan waktu proses yang lama dan hasil yang kurang akurat terbukti dengan masih banyak kredit macet.



### **1.3. Batasan Masalah**

Untuk memperjelas pembahasan penelitian ini, penulis memberikan kesimpulan dan permasalahan dalam penelitian ini:

1. Data yang digunakan pada penelitian ini menggunakan Data Nasabah Pinjaman Kredit Mikro Pinjaman Multi Guna pada BPR Indobaru Finansial.
2. Penentuan kelayakan kredit nasabah menggunakan sistem informasi yang dilakukan dengan menggunakan metode dari data mining dan proses klasifikasi decision tree serta perhitungan algoritma dari C4.5 dengan pengujian dari data mining dilakukan menggunakan software atau perangkat lunak WEKA 3.9.2.
3. Hasil dari penelitian yang dibuat adalah dokumen yang berisi penjelasan temuan hasil dari analisa mengenai sistem informasi setelah melalui proses analisis dan evaluasi dan rekomendasi terkait dengan hasil dari analisa mengenai sistem informasi yang didapat.

### **1.4. Rumusan Masalah**

Berdasarkan identifikasi masalah diatas maka permasalahan yang akan di bahas antara lain yaitu:

1. Bagaimana memanfaatkan data mining untuk membantu pihak marketing dan bank dalam memutuskan nasabah yang layak diberikan kredit agar mengurangi kemacetan kredit dengan Metode algoritma C4.5 ?
2. Bagaimana menerapkan algoritma C4.5 dalam menentukan kelayakan kredit pada nasabah BPR Indobaru Finansial ?

3. Bagaimana cara memudahkan BPR untuk menentukan kelayakan kredit dengan memanfaatkan data mining dengan pengujian dari data mining dilakukan menggunakan software atau perangkat lunak WEKA 3.9.2.

### **1.5. Tujuan Penelitian**

Tujuan yang ingin dicapai dalam penelitian ini dapat diuraikan antara lain sebagai berikut yaitu:

1. Mengetahui cara memanfaatkan data mining untuk membantu pihak bank dalam memutuskan nasabah yang layak diberikan kredit agar mengurangi kemacetan kredit dengan Metode algoritma C4.5.
2. Mengetahui cara menerapkan algoritma C4.5 dalam menentukan kelayakan kredit pada nasabah BPR Indobaru Finansial.
3. Mengetahui cara memudahkan BPR untuk menentukan kelayakan kredit dengan memanfaatkan data mining dengan pengujian dari data mining dilakukan menggunakan software atau perangkat lunak WEKA 3.9.6

### **1.6. Manfaat Penelitian**

Penelitian ini diharapkan berguna bagi pembaca. Manfaat dari penelitian yang dilakukan ini dibagi menjadi dua bagian yaitu sebagai berikut:

#### **1.6.1. Manfaat Teoritis**

Manfaat teoritis yang diharapkan dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Menetapkan konsep berkenaan algoritma C4.5 dalam memprediksi Kelayakan Kredit di BPR Indobaru.

2. Hal ini terkait dengan proses pengumpulan data menggunakan metode algoritme C4.5 sebagai *suplemen* untuk penelitian selanjutnya..
3. Pengetahuan tentang *data mining* dapat diperluas, terutama berkenaan dengan algoritma C4.5 pada data yang dianggap tidak berguna atau diabaikan oleh pemilik data.

### **1.6.2. Manfaat Praktis**

Selain manfaat utama, manfaat praktis yang diharapkan dari penelitian ini, antara lain:

1. Dapat difungsikan sebagai tolak ukur untuk memprediksi Kelayakan pemeberian kredit yang tepat dan akurat di BPR Indobaru Batam sehingga dapat mengurangi kredit yang macet.
2. Memberikan saran kepada BPR Indobaru dalam memprediksi Kelayakan pemberian kredit.
4. Memperoleh ilmu baru tentang teknik data mining khususnya metode algoritma c45 dalam memprediksi kelayakan pemberian kredit yang tepat supaya pihak BPR Indobaru dapat meningkatkan kredit analisis yang baik dalam pemberian kredit.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Tinjauan Teori Umum**

##### **2.1.1. Knowledge Discovery in Database (KDD)**

*Knowledge Discovery in Databases (KDD)* adalah teknik yang mengumpulkan data historis untuk menemukan pola, pola, atau hubungan dalam database besar. KDD Ini didefinisikan sebagai penghapusan informasi rahasia, terenkripsi dan anonim dari pemrosesan data. Proses penyimpanan data meliputi hasil proses pengumpulan data (proses menemukan pola dan struktur data) dan penemuan pengetahuan serta menerjemahkan hasilnya dengan cara yang mudah dipahami (Sularno & Anggraini, 2017).

Proses KDD menurut (Mardi, 2017) dapat didefinisikan, yaitu:

1. Data Selection

Data harus dipilih dari kumpulan data aktif sebelum fase penemuan pengetahuan (KDD) dapat dimulai di database. Data yang Anda gunakan untuk proses pengumpulan data *ter-save* kedalam folder terpisah dari database produksi.

2. Pre-processing / Cleaning

Sebelum memulai proses pengumpulan data, harus menjalankan proses pembersihan pada data yang merupakan KDD dalam *database*. Secara khusus, proses pembersihan meliputi penghapusan data duplikat, pemeriksaan data yang tidak konsisten, dan koreksi kesalahan data seperti

kesalahan ejaan. Sebuah sistem pengayaan juga dikembangkan. Artinya, proses “menambah” data yang ada dengan data atau informasi penting lainnya yang diperlukan untuk penemuan pengetahuan KDD dalam *database*, seperti data eksternal atau informasi penting lainnya.

### 3. Transformation

Sangat cocok untuk proses pengumpulan data karena coding adalah proses mengubah data yang dipilih. Proses pengkodean Knowledge Discovery (KDD) Ini adalah proses alami dalam database berdasarkan jenis atau pola informasi yang dimasukkan ke dalam database.

### 4. Data Mining

Menurut (Rumahorbo & Arnomo, 2020) Data mining ialah tahapan menemukan informasi penting yang terdapat dikumpulan database yang besar. Data mining adalah proses menemukan pola atau informasi menarik pada data terpilih dengan menggunakan metode atau teknik tertentu. Teknik, metode, atau algoritma penambangan data sangat berbeda (Septiani, 2017). Memilih metode atau algoritma yang tepat tergantung pada tujuan dan metode deteksi sensitivitas KDD pada data. Data mining, sering juga disebut *knowledge discovery in database* (KDD). Data mining adalah untuk mengekstrasikan atau “menambang” pengetahuan dari kumpulan banyak data., adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menentukan pola keteraturan, pola hubungan dalam set data berukuran besar (Septiani, 2017).

Menurut Bramer dalam jurnal (Mardi, 2017) Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD). Maka dari itu, data mining dapat digunakan untuk mengekstrak informasi dan pengetahuan penting yang tersembunyi dari data set yang besar. Data mining dapat dibagi menjadi beberapa kelompok seperti dibawah ini (Maulana & Fajrin, 2018):

1. Deskripsi

Deskripsi adalah cara untuk menjelaskan pola dan pola dalam data yang disimpan.

2. Estimasi

Penilaian seperti kategoris, tetapi perubahan objektif dalam penilaian lebih bersifat kuantitatif daripada kategoris. Model dibangun menggunakan seluruh dataset, memberikan nilai variabel target sebagai nilai prediksi.

3. Prediksi

Prediksi adalah mempertimbangkan nilai yang tidak diketahui dan memprediksi nilai masa depan.

4. Klasifikasi

Klasifikasi menggunakan variabel yang berbeda, misalnya, distribusi pendapatan dapat dibagi menjadi tiga kategori: pendapatan besar, pendapatan menengah dan pendapatan kecil.

5. Pengklasteran

Pengklasteran untuk mencatat, mengamati atau melakukan pengamatan terhadap sifat-sifat hal-hal yang mempunyai persamaan.

6. Asosiasi

Tugas asosiasi adalah menemukan sumber daya untuk bekerja sama.

7. Interpretation / Evaluation

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

Tahap ini merupakan bagian dari proses *Knowledge Discovery in Database* (KDD) yang disebut interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

## 2.2. Tinjauan Teori Khusus

### 2.2.1. Algoritma C4.5

Menurut Sularno dalam jurnal (Elisa & Harman, 2019) Algoritma C4.5 adalah salah satu metode untuk membuat decision tree berdasarkan training data yang telah disediakan. Algoritma C4.5 adalah algoritma yang paling umum digunakan untuk merencanakan pohon keputusan karena keunggulannya yang besar dibandingkan algoritma lainnya. Selain itu, keuntungan dari algoritma C4.5 adalah mudah untuk dijelaskan, memiliki akurasi yang dapat diterima, dapat

menangani fitur khusus dengan baik, dan dapat membuat pohon keputusan yang dapat menangani fitur dan angka yang berbeda (Muhammad Rizal, 2019).

Algoritma juga dapat berguna untuk membangun pohon keputusan (*decision tree*) menurut (Sulaeman & Rilmansyah, 2021) sebagai berikut :

1. Memilih atribut sebagai node akar
2. Membuat cabang untuk tiap-tiap nilai.
3. Membagi kasus dalam cabang.
4. Mengulangi proses untuk setiap cabang sampai semua kasus memiliki kelas yang sama.

#### **2.2.1.1. Decision Tree**

*Decision tree* merupakan metode klasifikasi dan prediksi yang mengubah data menjadi pohon keputusan dan aturan-aturan keputusan. Sebuah model keputusan terdiri dari seperangkat aturan yang membagi besar, populasi heterogen menjadi populasi yang lebih kecil yang memiliki jumlah yang lebih besar dari variabel target. Data dalam pohon keputusan biasanya disajikan dalam format tabel dalam bentuk atribut dan struktur data. Komponen yang dibuat oleh deklarasi properti berfungsi sebagai dasar untuk membangun pohon. (Wahyuni & Ginting, 2018). Dalam tahapannya, pohon keputusan melewati proses yang dimulai dengan semua potongan data, dengan nilai yang berbeda dan pengetahuan yang lengkap, dan berakhir dengan sebuah keputusan.

Selanjutnya, demikian penjelasan dari kelebihan dan kekurangan yang dimiliki oleh *decision tree*, yaitu:

1. Kelebihan dari *decision tree*



- 1) Saat menggunakan pohon keputusan, jika cakupan datanya kompleks, itu menjadi pohon yang sederhana untuk digunakan sebagai dasar pengambilan keputusan.
- 2) Keuntungan menggunakan pohon keputusan adalah Anda dapat menghapus perhitungan yang tidak perlu sehingga ketika Anda memulai proses analisis, hanya gunakan kriteria yang memenuhi kebutuhan pengujian.
- 3) Ketika pohon keputusan digunakan dalam proses pemilihan integritas atau akurasi data, hasil pemilihan integritas atau akurasi data berbeda dengan ketika - mengorganisir data menurut kriteria lain.
- 4) *Decision tree* bertindak untuk hindari masalah lain dengan menggunakan sejumlah kecil kelompok atau artikel tanpa kehilangan kualitas keputusan akhir.

## 2. Kekurangan dari *decision tree*

- 1) Saat menggunakan pohon keputusan, jika jumlah hasil keputusan yang akan diperoleh dan kriteria yang digunakan besar, maka waktu yang dibutuhkan untuk membuat keputusan bertambah dan diperlukan sistem pendukung memori yang besar.
- 2) Sebuah pohon keputusan berisi ringkasan jumlah kesalahan yang terjadi di setiap bagian dari proses di pohon keputusan.
- 3) Sulit untuk membuat pohon keputusan dengan biaya tinggi.

Saat menggunakan pohon keputusan, hasil atau kualitas keputusan sangat bergantung pada model atau struktur pohon keputusan (Harryanto & Hansun,

2017). Untuk memilih atribut sebagai node akar, didasarkan pada nilai Gain tertinggi dari atribut-atribut yang ada. Untuk menghitung Gain digunakan rumus seperti tertera dalam persamaan berikut:

$$\text{Gain}(S,A) = \text{Entropy}(S) + \sum_{i=1}^n * \text{Entropy}(S_i)$$

Keterangan:

S : Himpunan Kasus

A : Atribut

n : Jumlah Partisi Atribut **Rumus 2.1** Perhitungan Gain

|S<sub>i</sub>| : Jumlah Kasus pada partisi ke-i

|S| : Jumlah Kasus dalam S

Sebelum menghitung nilai Gain, yang perlu dilakukan yaitu mencari nilai Entropy. Entropy digunakan untuk menentukan seberapa informatif sebuah masukan atribut untuk menghasilkan keluaran atribut. Rumus dasar dari Entropy tersebut adalah sebagai berikut:

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

Keterangan:

S : Himpunan Kasus

A : Fitur **Rumus 2.2** Perhitungan Entropy

n : Jumlah partisi S

p<sub>i</sub> : Proporsi dari S<sub>i</sub> terhadap S

### **2.2.2. Kredit**

Menurut asal mulanya kata kredit berasal dari kata credere yang artinya adalah kepercayaan, maksudnya adalah apabila seseorang memperoleh kredit maka berarti mereka memperoleh kepercayaan. Sedangkan bagi si pemberi kredit artinya memberikan kepercayaan kepada seseorang bahwa uang yang dipinjamkan pasti kenbali. Pengertian kredit menurut Undang-Undang Perbankan nomor 10 tahun 1998 adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan dengan pihak lain yang mewajibkan pihak meminjam melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga. Sedangkan pengertian pembiayaan adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan antara bank dengan pihak lain yang mewajibkan pihak yang dibiayai untuk mengembalikan uang atau tagihan tersebut setelah jangka waktu tertentu dengan imbalan atau bagi hasil (Putra et al., n.d.).

#### **2.2.2.1. Perjanjian Kredit**

Perjanjian kredit bank berbeda dengan perjanjian pinjaman menurut hukum perdata. Perjanjian kredit ini tidak tunduk pada ketentuan Bab 13 Pasal III KUHP perdata. Dengan kata lain, perjanjian pinjaman bank adalah perjanjian anonim, karena secara khusus diatur oleh hukum perdata atau hukum perbankan. Dasar hukum adalah kontrak atau kesepakatan antara bank dan peminjam berdasarkan prinsip kebebasan berkontrak.

#### 2.2.2.2. Jaminan Kredit

Jaminan pinjaman didefinisikan sebagai transfer properti atau pernyataan hak untuk membayar kembali pinjaman. Persetujuan pinjaman diperoleh melalui pertimbangan yang cermat dari perilaku peminjam, stresor, investasi dan prospek bisnis.

#### 2.2.2.3. Prinsip Penilaian Kredit

Sebagian besar standar penilaian kredit berlaku. Ini adalah analisis 5-C; Analisis 7-P dan bukti konsep. Analisis 5C adalah sebagai berikut.:

1. *Character*, dalam hal ini watak atau watak seseorang, merupakan agama masa depan. Tujuannya adalah untuk meyakinkan pihak bank bahwa watak atau watak peminjam benar-benar dapat dipercaya.
2. *Capacity (capability)*, untuk melihat kemampuan calon nasabah dalam membayar kredit dihubungkan dengan kemampuan mengelola bisnis serta kemampuan mencari laba.
3. *Capital*, dimana untuk mengetahui sumber-sumber pembiayaan yang dimiliki nasabah terhadap usaha yang akan dibiayai oleh Bank.
4. *Collateral*, merupakan jaminan yang diberikan calon nasabah baik yang bersifat fisik maupun non fisik. Jaminan hendaknya melebihi jumlah kredit yang diberikan.
5. *Condition*, dalam menilai kredit hendaknya dinilai kondisi ekonomi sekarang dan untuk dimasa yang akan datang sesuai sektor masing-masing.

Analisis 7 P adalah sebagai berikut:

1. Personality

Yaitu menilai nasabah dari segi kepribadiannya atau tingkah lakunya sehari-hari maupun masa lalunya. Personality juga mencakup sikap, emosi, tingkah laku, dan tindakan nasabah dalam menghadapi suatu masalah.

2. Party

Yaitu mengklasifikasikan nasabah ke dalam klasifikasi tertentu atau golongan-golongan tertentu berdasarkan modal, loyalitas, serta karakternya sehingga nasabah dapat digolongkan ke golongan tertentu dan akan mendapatkan fasilitas kredit yang berbeda pula dari bank.

3. Purpose

Yaitu untuk mengetahui tujuan nasabah dalam mengambil kredit termasuk jenis kredit yang diinginkan nasabah

4. Prospect

Yaitu untuk menilai usaha nasabah di masa yang akan datang apakah menguntungkan atau tidak, atau dengan kata lain mempunyai prospek atau sebaliknya.

5. Payment

Merupakan ukuran bagaimana cara nasabah mengembalikan kredit yang telah diambil atau dari sumber mana saja dana untuk pengembalian kredit yang diperolehnya.

6. Profitability

Untuk menganalisis bagaimana kemampuan nasabah dalam mencari laba.

7. Protection Tujuannya adalah bagaimana menjaga kredit yang dikucurkan oleh bank, tetapi melalui suatu perlindungan. Perlindungan dapat berupa jaminan barang atau orang atau jaminan asuransi. (Pato et al., 2013)

#### **2.2.2.4. Prosedur Pemberian Kredit**

Menurut (Yasman & Afriyeni, 2019) Dalam industri perbankan, sistem pembayarannya sama dengan bank, dengan sedikit perbedaan. Perbedaannya hanya dapat dikaitkan dengan faktor-faktor yang ditetapkan oleh bank dan tujuan pinjaman. Pada PT. Bank Perkreditan Rakyat Indobaru Finansial Proses pemberian kredit pada nasabah adalah sebagai berikut:

1. Nasabah mengajukan permohonan kredit lengkap dokumen-dokumen pendukung (perizinan, laporan, keuangan, data identitas, akta perusahaan, data penjualan, data agunan, dll).
2. Permohonan tersebut diproses setelah data-data yang diperlukan lengkap. Proses dimulai dengan melakukan kunjungan ke lokasi usaha, wawancara, dan penilaian agunan dan meminta data informasi debitur secara online ke Bank Indonesia (*BI Checking*). Istilah BI Checking/SID yang kini telah beralih dari Bank Indonesia ke Otoritas Jasa Keuangan (OJK) dengan nama Sistem Layanan Informasi Keuangan (SLIK). SLIK berisikan informasi pencarian, data pokok debitur, kolom pemilik atau pengurus (badan usaha), fasilitas beserta ringkasan kredit dan garansi yang diberikan, serta kolom kredit atau pembiayaan. SLIK merupakan sistem informasi yang dikelola OJK untuk mendukung pengelolaan dan

penyelenggaraan layanan informasi sektor keuangan. SLIK dapat digunakan di bidang keuangan untuk mempercepat proses investasi, mengatur pinjaman atau manajemen risiko keuangan, memantau kualitas peminjam, mengelola sumber daya koresponden SLIK, untuk memperkuat dan mengembangkan kemitraan antara jurnal SLIK dan lainnya.

3. Analisis kredit kemudian melakukan analisis/penilaian kredit 5C berdasarkan informasi yang diterima dan hasil survei lapangan dan wawancara. Temuan ini tercermin dalam laporan kredit, juga dikenal sebagai peringkat kredit atau rekomendasi kredit.
4. Jika setelah pemeriksaan pemeriksa kredit menentukan kemungkinan terjadinya kasus, pemeriksaan kredit direkomendasikan kepada manajer departemen atau pimpinan perusahaan (perusahaan yang menyetujui pinjaman atas nama yang memerlukan persetujuan).
5. Setelah disetujui, akan diperoleh Surat Keterangan Kredit (SP2K) sesuai dengan cara yang diminta oleh nasabah.

### **2.2.3. Bank**

Dari asal usul bank, berasal dari bahasa Italia banca, yang berarti bainis, sebuah kata yang sering digunakan oleh bankir Italia ketika mereka menawarkan pinjaman di pasar. Oleh karena itu, menurut Pasal 10 Negara Republik Indonesia tanggal 10 November 1998, pada tahun 1998, bank dapat menghimpun uang dari masyarakat dalam bentuk uang tunai dan menyalurkannya dalam bentuk pinjaman dan/atau jenis lainnya. Satu. Memperbaiki kondisi kehidupan masyarakat. (Ritha et al., 2016).

#### **2.2.4. Waikato Environment for Knowledge Analysis (WEKA)**

Menurut Effendy dalam jurnal (Arnomo, 2021) Aplikasi WEKA adalah sistem yang digunakan untuk melakukan pekerjaan dalam perbandingan terhadap algoritma dari mesin pembelajaran atau machine learning dalam penggunaan teknik data mining. WEKA adalah perangkat lunak pembelajaran mesin populer yang ditulis dalam bahasa pemrograman *Java*. WEKA dikembangkan oleh Universitas Waikato, Selandia Baru. WEKA berisi kumpulan algoritma dan alat visualisasi untuk analisis dan prediksi data. Algoritma pembelajaran mesin WEKA dapat digunakan untuk memecahkan masalah di bidang data mining. WEKA menerapkan semua proses pembelajaran untuk pemrosesan dan regresi. Artinya, pohon keputusan, mesin vektor pendukung (SVM), logistik dan linier, perceptron multi-layer, dan aplikasi WEKA di dekatnya memungkinkan pengguna untuk menemukan data dari database dan file melalui seleksi dan bidang visual. Sistem kemudahan penggunaan (Seminar & Matematika, 2021).

WEKA dapat digunakan untuk menyelesaikan masalah-masalah data mining di dunia nyata terutama klasifikasi yang menggunakan machine learning sebagai pendekatannya. WEKA mengandung tools pre-processing data, klasifikasi, regresi, clustering, aturan asosiasi, dan visualisasi. Pengujian dengan menggunakan WEKA lebih fleksibel dan efektif karena pohon keputusan yang terbentuk hampir sama dengan perhitungan secara manual (Yulia & Azwanti, 2018).

##### **2.2.4.1. Confusion Matrix**

Menurut Pramita dalam jurnal (Pratiwi et al., 2020) *Confusion Matrix* adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada



konsep data mining. Evaluasi dengan menggunakan metode confusion matrix menghasilkan nilai akurasi, presisi dan recall. Menurut Han dalam jurnal (Ginting et al., 2020) Akurasi dalam klasifikasi dalam Data Mining merupakan persentase ketepatan pada record data yang telah diklasifikasikan secara benar dan dilakukan pengujian pada hasil klasifikasi Presisi atau confidence merupakan proporsi pada kasus yang diprediksi positif, dimana data yang sebenarnya juga positif. Menurut Abdullah dalam jurnal (Ginting et al., 2020) Recall atau sensitivity adalah proporsi kasus positif yang diprediksi dengan benar.

Pengukuran akurasi dilakukan dengan metode pengujian confusion matrix dapat dilihat pada table berikut:

*Tabel 2. 1 Confusion matrix*

Correct Classification	Classification	
	Positif	Negatif
Positif	TP	TN
Negatif	FP	FN

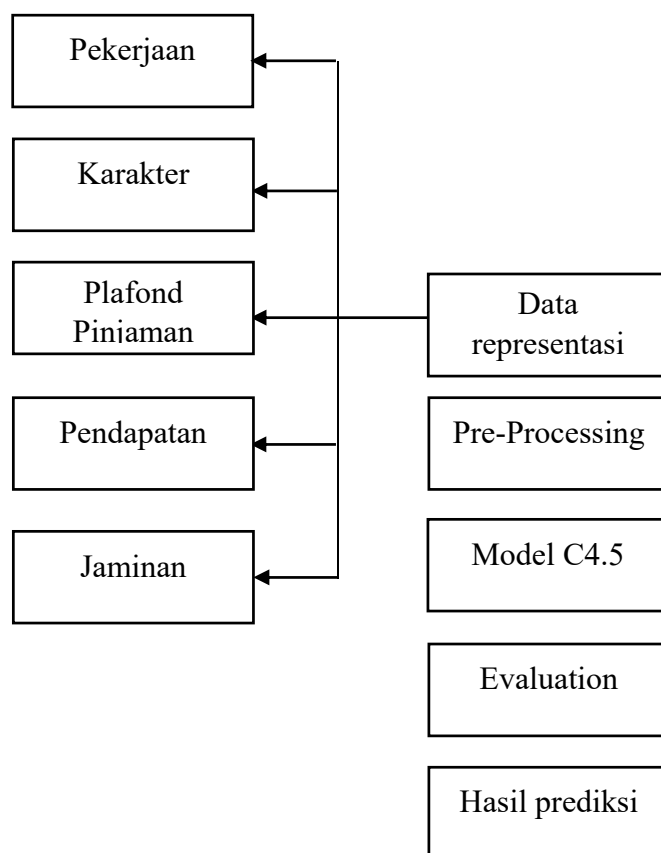
$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FN} \times 100\%$$

$$\text{Precision} = \frac{TP}{TP+FN} \times 100\%$$

**Rumus 2.3** confusion matrix

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\%$$

### 2.3. Kerangka Pemikiran



**Gambar 2.1** Kerangka Pemikiran

Pada tahapan representasi Data yang dikumpulkan dari PT BPR Indobaru Bulan juli 2021 sampai Febuari 2022 sebanyak 115 record dan memiliki atribut Pekerjaan, Karakter, Plafond pinjaman, Pendapatan, dan Jaminan. dari semua atribut, merupakan nilai kategorikal dan bukan nilai angka tertera penjelasan pada tabel 3.2 variabel penelitian. Akan dilakukan beberapa penyeleksian untuk menghasilkan data yang dibutuhkan. Melakukan teknik preprocessing yang pertama melakukan data cleaning bekerja untuk membersihkan nilai yang kosong, karna adanya data yang tidak konsisten atau mungkin tupel yang kosong (missing values dan noisy) Dari validasi data menghasilkan 115 data calon nasabah yang

dapat dihitung. Data reduction data yang digunakan untuk data training mungkin terlalu besar, hanya beberapa data yang diperlukan sehingga data yang tidak diperlukan akan dihapus. pada penelitian ini metode yang digunakan algoritma C4.5, Evaluasi dilakukan dengan cara mengamati hasil klasifikasi dari penerapan C4.5. Pengukuran tingkat akurasi dilakukan dengan menggunakan evaluasi model confusion matrix Dengan demikian dapat diketahui berapa tinggi akurasi dari algoritma tersebut (Nadiyah, n.d, 2020).

#### 2.4. Penelitian Terdahulu

Penelitian sebelumnya menjadi acuan dalam penelitian. Para peneliti juga mengidentifikasi beberapa temuan sebelumnya tentang kepercayaan, termasuk:

*Tabel 2. 2 Penelitian Terdahulu*

No.	Penelitian	Judul	Hasil Penelitian
1	Rika Nofitri, Juna Eska 2018	Implementasi Data Mining Klasifikasi C4.5 Dalam Menentukan Kelayakan Pengambilan Kredit	bentuk pohon keputusan dalam memberikan rekomendasi terhadap pengambilan keputusan untuk kelayakan pengambilan kredit adalah dilihat kriteria pertama yang menjadi root yaitu jaminan. Kriteria jaminan mempunyai sub root kedua yaitu watak. Kemudian subroot dari watak yaitu surat izin usaha. Selanjutnya sub root yang terakhir yaitu jenis kredit
2	Aldi Zein Pratama, Laela Kurniawati <sup>1</sup> , Simson Larbona <sup>1</sup> , Tuti Haryanti	Algoritma C4.5 Untuk Klasifikasi Nasabah Dalam Memprediksi Kredit Macet	Algoritma C4.5, yang menghitung entropi dan nilai untuk setiap fitur dalam sekumpulan fitur, bekerja dengan baik. Algoritma C4.5 juga

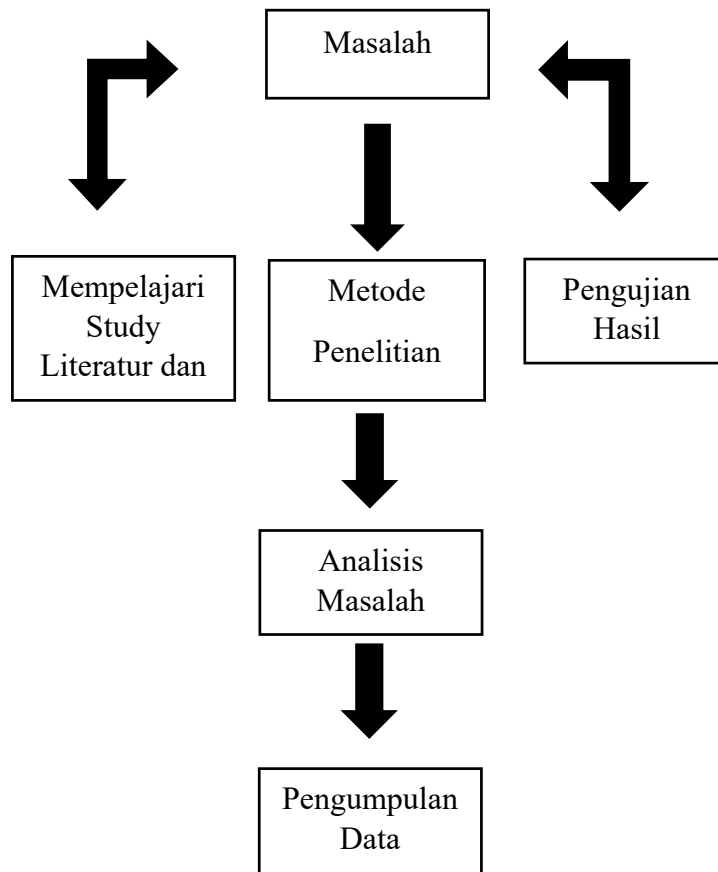
	2019		merupakan salah satu algoritma dari sistem pengolahan data yang sangat berguna untuk pengolahan data karena karakteristik dari klasifikasi data. Tetapi jenis pohon keputusan dan jenis aturan menjadi sangat jelas dan memungkinkan pengguna untuk dengan mudah mengekstrak informasi tentang data yang diminta.
3	Khotibul Umam, Diah Puspitasari, Acamad Nurhadi 2020	Penerapan Algoritma C4.5 Untuk Prediksi Loyalitas Nasabah PT Erdika Elit Jakarta	Berdasarkan pohon keputusan yang dibangun, karakteristik yang paling berpengaruh terhadap retensi pelanggan adalah pendidikan sebagai akar dari pohon keputusan dan memiliki nilai tertinggi di antara fitur lainnya dengan nilai 1,545292721. Di sisi lain, sifat pelanggan selalu hadir pada harga akhir dan nilai keuntungan sebesar 0,623919119, sehingga tidak berpengaruh signifikan terhadap retensi pelanggan.
4	Teguh Budi Santoso, Dela Sekardiana 2019	Penerapan Algoritma C4.5 Untuk Penentuan Kelayakan Pemberian Kredit (Studi Kasus : Koperia - Koperasi Warga Komplek Gandaria)	penggunaan data mining dengan algoritma C4.5 terbukti akurat dalam penentuan pengajuan kelayakan kredit. Hal ini dibuktikan dengan hasil evaluasi penelitian bahwa algoritma

			C4.5 mendapatkan nilai akurasi data training serta data testing menggunakan algoritma C4.5, kedua proses tersebut rata-rata 93%, dan error rate 6.6%
5	Arno Prayogo Nawary, Kurniati 2021	Penerapan Data Mining Dalam Memprediksi Kelancaran Kredit Nasabah Menggunakan Algoritma C4.5 (Studi Kasus Pada Pt. Astra International (Auto 2000 Plaju))	Penerapan Data Mining Dalam Memprediksi Kelancaran Kredit Nasabah Menggunakan Algoritma C4.5 (Studi Kasus Pada Pt. Astra International (Auto 2000 Plaju))

**BAB III**  
**METODE PENELITIAN**

**3.1. Design Penelitian**

Dalam hal ini peneliti memaparkan hasil dari desain penelitian yang telah dibuat menggunakan cara yang terstruktur agar dapat diambil sebagai acuan dalam melakukan penelitian dan dapat membantu peneliti pada saat proses penelitian, desain penelitian ini juga berperan sebagai pedoman bagi peneliti.



*Gambar 3. 1 Desain penelitian*

1. Mendeskripsikan masalah  
Penulis mencoba menjelaskan masalah terkait kredit yang diidentifikasi dalam versi algoritma yang dimulai dengan C4.5.
2. Analisa Masalah  
Langkah ini merupakan langkah selanjutnya setelah masalah diidentifikasi. Penulis mencoba untuk mengeksplorasi isu-isu yang berbeda.
3. Mempelajari dan memahami literatur  
Penulis berusaha mencari dan memahami berbagai sumber teori yang mendukung tinjauan pustaka agar dapat digunakan sebagai pedoman bagi penulis untuk membimbing penelitiannya.
4. Proses pengumpulan data  
Selama ini, peneliti pergi ke lembaga penelitian untuk melakukan wawancara dan bertanya jawab untuk mendapatkan informasi yang relevan sebanyak mungkin.
5. Analisa menentukan kelayakan kredit menggunakan C4.5  
Diawali dengan proses membangun pohon keputusan menggunakan algoritma model C4.5 untuk membuat seperangkat aturan yang digunakan untuk menentukan nilai kredit.
6. Implementasi C4.5  
Penulis menentukan integritas data pada akar pohon keputusan, kemudian menghitung nilai nilai absolut dari data tersebut, yang digunakan dalam proses konstruksi pohon.

## 7. Pengujian Hasil

Pada langkah terakhir ini, setelah menghitung nilai maksimum, peneliti melakukan eksperimen dan membangun pohon keputusan menggunakan WEKA versi 3.9.6

### 3.2. Objek Penelitian

Pada bagian topik penelitian, penulis menjelaskan area penelitian dan rencana proyek penelitian, sebagai berikut:

#### 3.2.1. Lokasi Penelitian

Penelitian ini berlokasi di PT. BPR Indobaru Finansia berlokasi di Jl. Laksamana Bintan, kompleks Tanah Mas Blok. C No.4-5 Kelurahan Teluk Tering, Kecamatan Batam Kota, Kota Batam.

#### 3.2.2. Jadwal Penelitian

Di bawah ini adalah jadwal penelitian penulis dalam melakukan penelitian terkait dengan penentuan kelayakan pelanggan PT. Bank Perkreditan Rakyat Indobaru Finansia.

*Tabel 3. 1 Jadwal Penelitian*

Kegiatan	Maret 2022				April 2022				Mei 2022				Juni 2022				Juli 2022			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Perencanaan Sistem	■	■	■	■																
Analisis sistem					■	■	■	■												
Desain									■	■	■	■								
Pengkodean													■	■	■	■				





**Tabel 3. 2 Variabel Penelitian**

No	Atribut	Penjelasan
1	Pekerjaan	Karyawan dan Usaha (BPR Indobaru Finansia pinjaman Mikro memiliki 2 fasilitas pinjaman untuk karyawan dan usaha)
2	Karakter	Baik, Cukup, Tidak Baik .( karakter dapat dinilai , dengan melakukan cek pada BI checking atau dengan menanyakan kolega dan orang disekitar calon debitur, sehingga pegawai dapat menilai karakter dari calon debitur)
3	Plafond Pinjaman	Pinjaman mikro multiguna memiliki limit pinjaman Rp50.000.000, Pinjaman Rendah (Rp5.000,000,- sampai Rp15.000.000) Sedang (Rp16.000,000 sampai Rp 30.000.000) Tinggi (Rp31.000.000 sampai Rp 50.000.000)
4	Pendapatan	Besar (Penghasilan > 3 kali dari angsuran), Sedang (Penghasilan 2 kali angsuran ), Kecil (penghasilan < 1kali angsuran)
5	Jaminan	Tinggi ( > 3 kali jumlah hutang), Cukup ( 2 kali Jumlah hutang ) Rendah (1 kali jumlah hutang)

### 3.3. Teknik Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi yang diperlukan untuk mencapai tujuan penelitian, atau untuk menentukan informasi

yang berkaitan dengan objek penelitian atau situasi di lokasi penelitian. Proses pengumpulan datanya adalah sebagai berikut:

1. Tinjauan pustaka adalah tinjauan pustaka dalam bidang studi tertentu. Survei adalah ringkasan singkat penelitian, diskusi, dan temuan tentang suatu topik, biasanya disusun secara kronologis atau tematis. Dalam hal ini, peneliti mempelajari berbagai artikel yang berkaitan dengan masalah yang sedang dipecahkan.
2. Dalam penelitian menggunakan Teknik observasi Metode pengumpulan data yang digunakan untuk mengamati langsung, melihat dan mengambil suatu data di PT. Bank Perkreditan Rakyat Indobaru Finansia
3. Metode makalah penelitian digunakan untuk menganalisis sumber tekstual masalah data di lokasi penelitian atau untuk mengekstrak dari kumpulan data yang ada sesuai dengan kebutuhan penelitian yang dilakukan. Dalam hal ini, peneliti mengamati dan mempelajari data-data calon nasabah di PT BPR Indobaru Finansia
4. Melakukan Teknik wawancara Untuk memahami aturan yang perlu ditetapkan dalam mengidentifikasi data calon nasabah, Penelitian ini mewawancarain pihak marketing mikro.

#### **3.4. Model Penelitian**

Penulis melakukan pencarian tipe menggunakan model data mining dan algorithm pada C4.5. Awalnya, data kinerja pelanggan diklasifikasikan sesuai dengan 115 kriteria, dan kemudian dihitung menggunakan algoritma C4.5 saat membangun struktur pohon keputusan untuk mendapatkan entropy dan nilai

maksimum, hal ini dilakukan beberapa kali hingga bilangan keputusan nol sebelum menerima model sebagai aturan kredit pertama pada nasabah PT BPR Indobaru Finansia.