

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian terkait

Penelitian ini sebelumnya dilakukan studi kepustakaan dari penelitian terdahulu sebagai dasar atau acuan untuk menyelesaikan tugas akhir. Dari studi kepustakaan tersebut, penulis menemukan beberapa penelitian yang mendukung untuk mengangkat topik yang berkaitan dengan penelitian penulis. Oleh karena itu, akan dilakukan langkah kajian terhadap beberapa penelitian mengenai algoritma yang digunakan serta yang akan diangkat oleh penulis dalam tugas akhir ini.

Tabel 2.1 *State-of-the-Art*

NO	Judul Penelitian	Metode	Hasil Penelitian
1.	Prediksi Kredit Macet Melalui Perilaku Nasabah Pada Koperasi Simpan Pinjam Menggunakan Metode Klasifikasi C4.5 [3]	Algoritma Decision Tree C4.5	Menggunakan data nasabah berjumlah 1312, Memiliki akurasi sebesar 92 %. Dapat digunakan untuk mendeteksi perilaku nasabah yang akan diberikan kredit.
2	A New Credit Scoring Method Based on Rough Sets and Decision Tree [4]	Algoritma Decision Tree C4.5 dengan metode dasar Rought Sets - Penggunaan rought set, untuk	Menggunakan data German set dan Australia Set memperoleh akurasi sebesar 73 % dan 86.

		meningkatkan performa dan akurasi dengan mengurangi atribut yang redundan.	penggunaan Rought sets dapat menaikkan akurasi.
3	Credit Scoring Menggunakan Metode Support Machine Vector Teknik Seleksi Atribut Berbasis Chi Squared Stastistic Dan Particle Swarm Opimazation [5]	Support Machine Vector Berbasis Chi Squared Stastistik Dan Particle Swarm Optimazation	Credit scoring menggunakan data set German dengan 20 <i>predictor variables weighted chi squared statistic</i> , dengan 19 atribut lalu terpilih 14 atribut untuk PSO proses, SVM Berbasis Chi Squared Stastistik Dan Particle Swarm Optimazation menghasilkan akurasi 77, 80%
4	Comparative Study Of Data Mining Model For Credit Card Aplication Scoring In Bank [6]	Naïve bayes Classifier dan ID3 classifier	Menggunakan data set 1000 record dan memiliki 2 kelas label (<i>Approved Or Rejected</i>), mengkomparasi antara Naive bayes dan ID3 menghasilkan akurasi sebesar 82 %

			dan 76%.
--	--	--	----------

2.2 Tinjauan Pustaka

2.2.1 Bank

Bank sebagai lembaga keuangan yang berfungsi sebagai perantara keuangan (*financial intermediary*) yaitu menghimpun dana dari masyarakat dalam bentuk simpanan dan menyalurkannya kepada masyarakat yang membutuhkan dalam bentuk kredit. Menurut UU RI No. 21 tahun 2008 tentang Perbankan Syariah yang dimaksud dengan bank adalah “Bank adalah badan usaha yang menghimpun dana dari masyarakat dalam bentuk Simpanan dan menyalurkannya kepada masyarakat dalam bentuk kredit dan/atau bentuk lainnya dalam rangka meningkatkan taraf hidup rakyat” [2].

2.2.2 Kredit

Pengertian kredit telah diatur dalam undang-undang yang berlaku di Indonesia yaitu : Undang-Undang No.14 tahun 1967 bab 1 , 2 yang menyatakan :” Kredit adalah penyedia uang atau yang disamakan dengan itu berdasarkan persetujuan pinjam meminjam antara bank dengan lain pihak dalam hal mana pihak peminjam berkewajiban melunasi hutang setelah jangka waktu tertentu dengan jumlah bunga yang telah ditentukan”. Selanjutnya pengertian tersebut disempurnakan lagi dalam Undang-Undang No.7 tentang Perbankan sebagaimana telah diubah dengan Undang - Undang No10 tahun 1998, yang mendefinisikan kredit sebagai berikut: ” Kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu berdasarkan persetujuan atau kesepakatan panjam meminjam antara bank dengan pihak lain yang

mewajibkan pihak peminjam untuk melunasi hutangnya setelah jangka waktu tertentu dengan jumlah bunga” [2] .

2.2.3 Prinsip - Prinsip Perkreditan

Prinsip perkreditan disebut juga sebagai konsep 5C, Pada dasarnya konsep 5C ini akan dapat memberikan informasi mengenai tekad baik dan kemampuan membayar nasabah untuk melunasi kembali pinjaman beserta bunganya. Prinsip 5C tersebut antara lain adalah [2] :

1. *Character*

Menilai karakter untuk mengetahui apakah calon debitur merupakan orang yang jujur dan punya etikat baik untuk melakukan kewajiban yang harus dilakukan oleh debitur.

2. *Capacity*

Menilai kemapuan debitur untuk melakukan pelunasan kewajibannya dari usaha yang dilakukan atau usaha yang akan dibiayai oleh kredit bank.

3. *Capital*

Menilai besar atau kecilnya modal seorang calon debitur, serta menilai modal akan didistribusikan.

4. *Collateral*

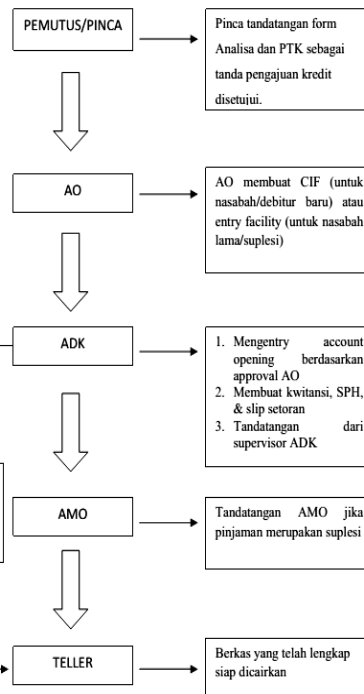
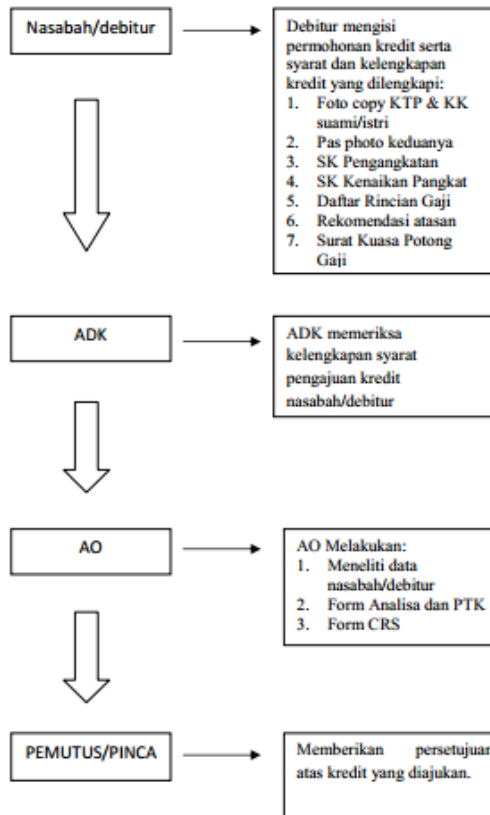
Menilai jaminan fisik harta benda yang akan digunakan sebagai penjamin apabila seorang calon debitur tidak bisa memenuhi kewajiban, misal terjadi sesuatu seperti kecelakaan dan hal lain yang mengakibatkan peminjam tidak dapat melakukan pelunasan.

5. *Condition of Economy*

Melihat bagaimana situasi, kondisi ekonomi dan sektor usaha calon debitur atau peminjam . Agar mengurangi resiko yang muncul yang disebabkan kondisi ekonomi.

2.2.4 Prosedur Pemberian Kredit

Prosedur secara umum pemberian kredit serta penilaian di dunia perbankan adalah sebagai berikut yang dapat dilihat pada gambar 2.1 [2] :



Keterangan:

AD (administrasi officer)

AO(Account Officer)

ADK (administrasi Kredit)

AMO (supervisor)

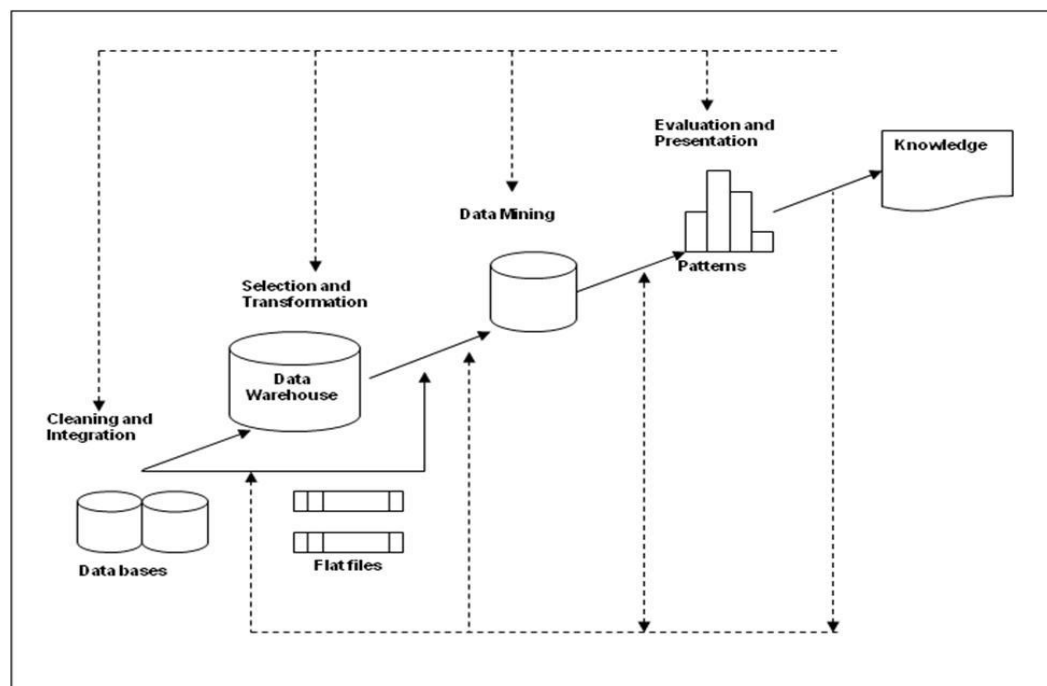
Jika nasabah/debitur baru maka berkas-berkas dibawa ke Teller untuk dicairkan

Gambar 2.1 Alur pemberian kredit bank

Untuk mengetahui layak atau tidak layaknya kredit yang akan diberikan kepada calon debitur, maka diperlukan suatu penilaian kelayakan pemberian kredit dengan menggunakan prinsip 5C.

2.2.5 Tahapan Data Mining

Data Mining dan Knowledge Discovery Database untuk mencari informasi dari suatu basis data yang menghasilkan suatu informasi yang tersembunyi. Proses tahapan KDD akan dijelaskan sebagai berikut: [7]



Gambar 2 . 2 Tahapan Data Mining KDD

1) Seleksi data / Data Selection

Seleksi atau pemilihan untuk mendapatkan data baru dari sekumpulan data. Dilakukan sebelum tahapan penggalian informasi. Kemudian data baru yang berasal dari hasil seleksi akan digunakan dalam proses data mining.

2) **Pembersihan Data**

Tahapan pembersihan data yang berguna untuk menghilangkan noise, data yang tidak relevan, *typhografi* (kesalahan dalam pengetikan), membuang atribut – atribut yang tidak relevan dengan hipotesa. Pembersihan ini akan mempengaruhi performa data mining.

3) **Transformasi**

Proses pengubahan data ke bentuk yang sesuai dengan teknik data mining yang digunakan, karena beberapa teknik data mining membutuhkan format data yang khusus.

4) **Data Mining**

Menemukan pola atau informasi yang berhasil digali, dengan mengimplementasikan teknik, metode dan algoritma data mining yang bervariasi.

5) **Interpretasi/Evaluasi**

Pola maupun informasi yang didapatkan dari proses data mining kemudian dievaluasi dan di presentasikan kedalam bentuk yang mudah untuk dimengerti. Pada Tahapan ini pola tersebut akan diperiksa yang bertentangan dengan hipotesa yang sebelumnya telah ada.

2.2.6 Tugas Dalam Data Mining

Pengelompokan data mining berdasarkan tugasnya [8]

1. Deskripsi

Untuk menggambarkan pola dan kecenderungan yang ada dalam data, misalnya untuk mengetahui banyak sedikitnya dukungan dalam pemilihan gubernur. Deskripsi pola kecenderungan akan dapat memberikan penjelasan untuk suatu pola dan kecenderungan.

2. Estimasi

Estimasi memiliki atribut target yang lebih kearah numerik dari pada katagori, Estimasi serupa dengan klasifikasi perbedaanya adalah variable targetnya yang berupa numeric. Misalnya mengukur estimasi tekanan darah pada suatu rumah sakit berdasarkan jenis kelamin, umur , berat badan dan level sodium darah. Hubungan antara tekanan darah dan nilai variable prediksi pada proses pembelajaran akan menghasilkan estimasi.

3. Prediksi

Prediksi dinilai pada masa mendatang, misalnya

- Prediksi curah hujan dalam tiga bulan yang akan datang.
- Prediksi kelulusan mahasiswa.

4. Klasifikasi

Klasifikasi terdapat variable target dan variable katagori, variable katagori merupakan penentu variable target. Misalnya adalah

- Memperkirakan apakah pemmberian kredit akan terjadi kredit lancar atau kredit macet
- Memperkirakan apakah debitur layak atau tidak layak untuk diberikan kredit

5. Pengklusteran

Pengelompokan record, mengamati kelas objek – objek untuk mengetahui kemiripan. Kumpulan record yang memiliki kemiripan antara satu dengan yang lain dan record yang tidak memiliki kemiripan dengan record – record yang lain, Pengklusteran berbeda dengan klasifikasi karena pengklusteran tidak memiliki variable target dan juga tidak melakukan klasifikasi, memprediksi atau mengestimasi nilai dari suatu variable target. Pengklusteran membagi semua data menjadi kelompok yang memliki kesamaan atau homogen, kesamaan dalam record pada suatu kelompok memiliki nilai maksimal. Sedangkan kemiripan record dalam kelompok lain memiliki nilai minimal.

6. Asosiasi

Asosiasi adalah untuk menemukan atribut yang muncul pada satu waktu. Dalam perdagangan disebut analisa keranjang pasar, contoh penggunaan asosiasi dalam dunia bisnis adalah:

- Untuk mengetahui jumlah pelanggan operator yang diharapkan memberikan respon positif pada penawaran *upgrade* layanan yang diberikan.
- Mengetahui pola beli konsumen , dengan mengetahui pembelian barang yang dilakukan bersamaan dan yang tidak dibeli secara bersamaan.

2.2.7 Metode Klasifikasi

suatu pekerjaan dalam menilai objek data kemudian dimasukkan kedalam kelas tertentu dari sekelompok kelas yang tersedia disebut klasifikasi. Pekerjaan utama yang dilakukan dalam mengklasifikasikan suatu objek antara lain, pembangunan model sebagai *prototype* yang akan disimpan dalam bentuk memori dan penggunaan model ini berfungsi untuk mengenalkan atau mengklasifikasikan suatu objek data lain untuk mengetahui dikelas mana objek data tersebut didalam model yang sudah disimpan. [9]

empat komponen komponen dalam klasifikasi dilakukan berdasarkan [10]:

A. Kelas

Variabel dependen berupa kategorikal yang mempresentasikan “label” yang ada dalam obyek. Contohnya: Kelayakan pemberian kredit (Layak atau Tidak Layak)

B. Variabel Predictor

Variabel Independen direpresentasikan sebagai karakteristik data (atribut). Contohnya:, umur, Karakter, Penghasilan, Jumlah Tanggungan ,Nilai Anggunan.

C. Dataset Training

Suatu data set berisi nilai dari kedua komponen diatas yang digunakan untuk menentukan kelas yang cocok berdasarkan predictor.

D. Testing dataset

Memuat data baru untuk diklasifikasikan oleh model yang telah dibuat kemudian dievaluasi akurasi klasifikasinya.

2.2.8 Algoritma Decision Tree C4.5

Algoritma C4.5 merupakan algoritma yang berguna dalam membentuk pohon keputusan. Metode prediksi dan klasifikasi yang digunakan untuk eksplorasi data , menemukan hubungan tersembunyi diantara variabel target dengan beberapa calon variabel input disebut dengan pohon keputusan. Pohon keputusan sangat bagus untuk langkah awal pada proses permodelan bahkan ketika menjadi model akhir dari beberapa teknik, karena pohon keputusan merupakan perpaduan antara permodelan dan eksplorasi data. Hal pertama yang harus dilakukan untuk membangun pohon keputusan adalah memilih atribut sebagai akar, kemudian dari akar tersebut dibuat cabang untuk tiap nilai. Selanjutnya kasus dibagi dalam cabang. Berikutnya proses diulangi pada setiap cabang hingga seluruh cabang memiliki kelas yang sama [8].

Pemilihan atribut dengan akar berdasarkan jumlah gain yang memiliki nilai terbesar dari atribut yang tersedia. Gain dapat dihitung dengan rumus berikut ini [8].

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

Keterangan:

S : Himpunan kasus

A : Atribut

N : Jumlah partisi atribut A

| Si | : Jumlah kasus pada partisi ke-i

| S | : Jumlah kasus dalam S

Dari perhitungan didapatkan nilai gain dari atribut dengan nilai terbesar.. Gain merupakan *attribute selection measure* berguna untuk memilih *test attribute* pada pohon. Atribut yang memiliki *information gain* terbesar terpilih sebagai *test Attribute* dalam *node* tertentu.

Untuk menghitung nilai entropi dengan menggunakan persamaan berikut :

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

Keterangan :

S : Himpunan kasus

A : Atribut

N : Jumlah partisi S

Pi : Proporsi dari Si terhadap S

Pada algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

1. Memilih atribut untuk menjadi akar
2. Membuat nilai untuk tiap – tiap cabang

3. Membagi kasus dalam seluruh cabang.

Kemudian mengulangi proses pada semua cabang hingga seluruh kasus pada cabang mempunyai kelas yang sama [8].

2.2.9 Confusion Matrix

Confusion matrix merupakan salah satu metode untuk menilai performa klasifikasi yang didasarkan pada objek benar dan salah. Serta memberi keputusan yang didapatkan dalam training dan testing. Informasi di dalam *Confusion matrix* memiliki sifat aktual dan diprediksi didalam sistem klasifikasi [8].

Tabel 2.2 Tabel *Confusion Matrix*

<i>Classification</i>	<i>Predicted Class</i>		
		<i>Class = Yes</i>	<i>Class = No</i>
<i>Observed Class</i>	<i>Class = Yes</i>	<i>A</i> (<i>true positif – tp</i>)	<i>B</i> (<i>false negative – fn</i>)
	<i>Class = No</i>	<i>C</i> (<i>false positif – fp</i>)	<i>D</i> (<i>true negative – tn</i>)

Keterangan:

- True Positive* (tp) = proporsi positif dalam data set yang diklasifikasikan positif
- True Negative* (tn) = proporsi negative dalam data set yang diklasifikasikan negative
- False Positive* (fp) = proporsi negatif dalam data set yang diklasifikasikan positif
- False Negative* (fn) = proporsi negative dalam data set yang diklasifikasikan negative.

Berikut adalah persamaan model *confusion matrix*:

- a. Nilai akurasi merupakan proporsi jumlah prediksi benar dihitung dengan persamaan berikut ini:

$$acc = \frac{tp + tn}{tp + tn + fp + fn}$$

- b. *Recall* adalah membandingkan proporsi tupelo positif dihitung dengan persamaan berikut ini:

$$Sensitivity = \frac{tp}{tp + fn}$$

- c. *Specificity* adalah memabndikan proporsi tupelo negative dihitung dengan persamaan berikut :

$$Specificity = \frac{tn}{tn + fp}$$

- d. *Precision* merupakan jumlah proporsi kasus yang memiliki diagnosis hasil positif dihitung dengan persamaan berikut:

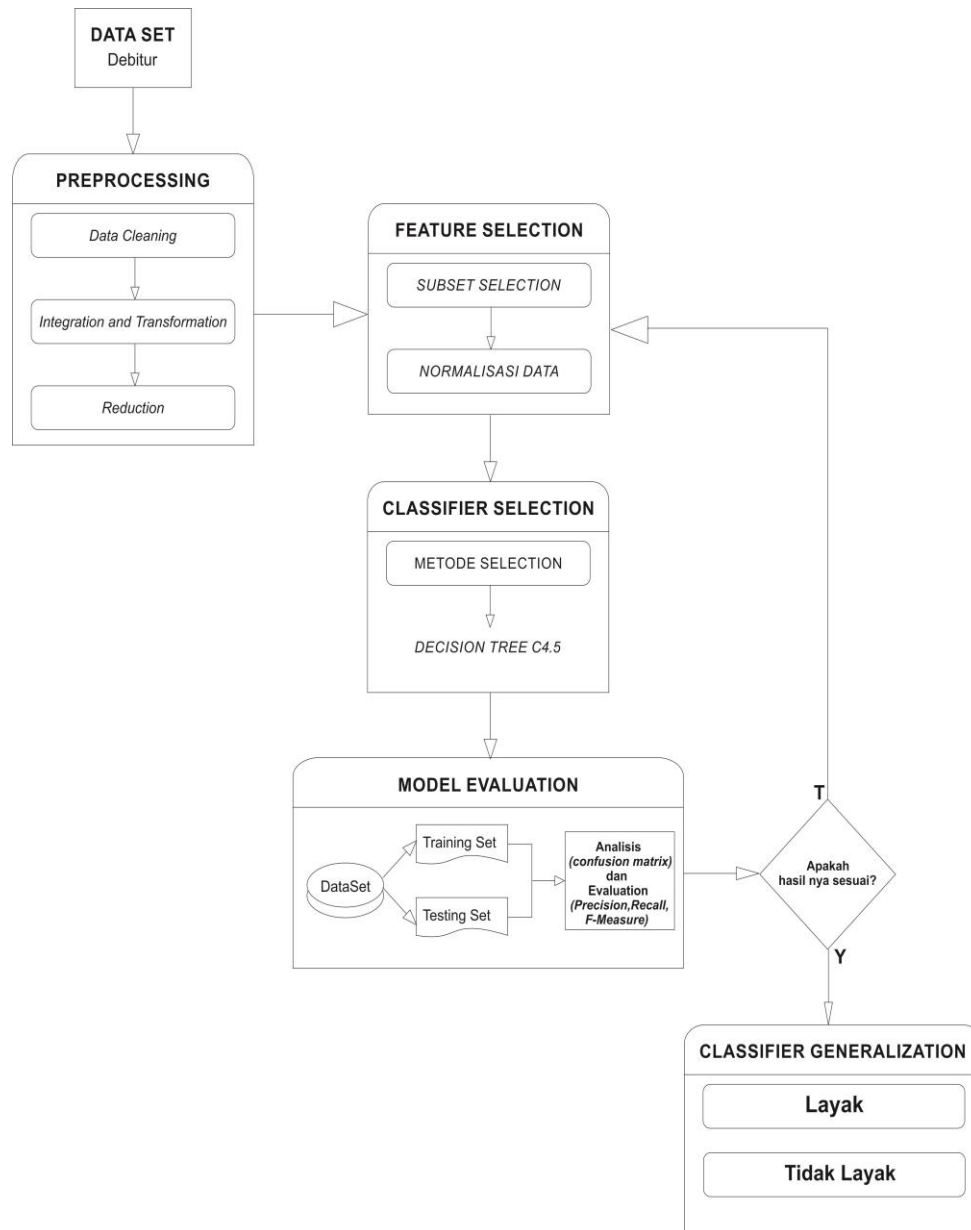
$$PPV = \frac{tp}{tp + fp}$$

- e. *Negative predictive value* merupakan jumlah proporsi kasus yang memiliki diagnosis hasil positif dihitung dengan persamaan berikut :

$$NPV = \frac{tn}{tn + fn}$$

2.3 Kerangka Pemikiran

Pada penelitian tugas akhir ini dibuat kerangka penelitian bertujuan agar penelitian dapat dilakukan secara bertahap serta konsisten dan merupakan garis besar dari langkah – langkah penelitian yang dilakukan dalam penelitian.



Gambar 2. 3 Kerangka Pemikirann