

PENGEMBANGAN MODEL *CREDIT SCORING* DI PERUSAHAAN X

SKRIPSI

Diajukan untuk memenuhi salah satu syarat guna mencapai gelar
Sarjana dalam bidang ilmu Teknik Industri

Disusun oleh:

Nama : Christophorus Adrian

NPM : 2015610143



**PROGRAM STUDI SARJANA TEKNIK INDUSTRI
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG
2019**

**FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS KATOLIK PARAHYANGAN
BANDUNG**



Nama : Christophorus Adrian
NPM : 2015610143
Jurusan : Teknik Industri
Judul Skripsi : PENGEMBANGAN MODEL *CREDIT SCORING* DI
PERUSAHAAN X

TANDA PERSETUJUAN SKRIPSI

Bandung, 12 Agustus 2019

Ketua Program Studi Sarjana Teknik Industri

(Romy Loice, S.T., M.T)

Pembimbing

(Fransiscus Rian Pratikto, S.T., M.T., MIE.)



Jurusan Teknik Industri
Fakultas Teknologi Industri
Universitas Katolik Parahyangan

Pernyataan Tidak Mencontek atau Melakukan Tindakan Plagiat

Saya, yang bertanda tangan di bawah ini,

Nama : Christophorus Adrian

NPM : 2015610143

dengan ini menyatakan bahwa skripsi dengan judul:

“PENGEMBANGAN MODEL *CREDIT SCORING* DI PERUSAHAAN X”

adalah hasil pekerjaan saya dan seluruh ide, pendapat atau materi dari sumber lain telah dikutip dengan cara penulisan referensi yang sesuai.

Pernyataan ini saya buat dengan sebenar-benarnya dan jika pernyataan ini tidak sesuai dengan kenyataan, maka saya bersedia menanggung sanksi yang akan dikenakan kepada saya.

Bandung,

Christophorus Adrian
2015610143

ABSTRAK

Perusahaan X merupakan perusahaan yang bergerak pada bidang perbankan. Salah satu jasa yang diberikan oleh perusahaan X adalah pemberian kredit kepada pengguna jasa perusahaan X. Pemberian kredit melalui kartu kredit merupakan salah satu sumber utama dari pendapatan perusahaan. Pada data yang diberikan oleh perusahaan X terdapat kredit yang macet sebesar 22 miliar rupiah atau setara dengan 16% dari total tagihan kredit yang diberikan kepada 17707 pengguna kartu kredit. Kredit dikatakan macet ketika pengguna jasa tidak dapat melakukan pembayaran dengan lunas sesuai dengan ketentuan waktu yang telah ditetapkan. Kredit dikatakan lancar ketika pengguna jasa dapat melunasi tagihan sesuai pada waktu yang telah ditetapkan.

Berdasarkan penelitian yang telah dilakukan dalam pengambilan keputusan apakah kredit akan macet atau tidak terdapat beberapa variabel yang perlu diperhatikan, variabel-variabel tersebut antara lain kode cabang, jumlah kartu, total pemakaian tunai, total pemakaian, outstanding, rasio pembayaran, persentase overlimit, rasio pembayaran tiga bulan, pemakaian enam bulan per limit, rasio pembayaran enam bulan, sisa tagihan per jumlah kartu, sisa tagihan per limit, total pemakaian per limit dan jumlah tahun sejak pembukaan kredit. Dengan menggunakan variabel-variabel tersebut sebagai *input* dibuat model *decision tree* dan *neural network*. *Decision tree* yang akan dirancang menggunakan algoritma C5.0 dan CHAID, kedua algoritma tersebut akan menghasilkan jumlah cabang yang berbeda karena menggunakan cara kerja yang berbeda. *Neural network* akan menggunakan tipe MLP, dimana MLP merupakan tipe *neural network* yang sering digunakan untuk klasifikasi prediksi.

Berdasarkan ketiga model yang terbentuk perusahaan sebaiknya menerapkan metode *decision tree* C5.0 karena metode *decision tree* merupakan metode paling mudah diimplementasi. Algoritma yang dipilih adalah C5.0 karena memiliki tingkat keakurasian yang lebih tinggi dibandingkan dengan algoritma CHAID. Walau *neural network* memiliki tingkat keakurasian yang cukup tinggi metode tersebut sulit diimplementasi.

ABSTRACT

Company X is a company engaged in the banking sector. One of the services provided by company X is the provision of credit cards. Giving credit through credit cards is one of the main sources of company revenue. In the data provided by company X, there are bad loans of 22 billion rupiah, equivalent to 16% of the total credit bills given to 17707 credit card user. Credit is called bad credit when a service user cannot make a payment in full in accordance with the stipulated time. Credit is called current credit when service user can pay off bills according to the stipulated time.

Base on the research that has been done, the decision on whether credit will bad credit or current credit are affected by several variables. These variables include kode cabang, jumlah kartu, total pemakaian tunai, total pemakaian, outstanding, rasio pembayaran, persentase overlimit, rasio pembayaran tiga bulan, pemakaian enam bulan per limit, rasio pembayaran enam bulan, sisa tagihan per jumlah kartu, sisa tagihan per limit, total pemakaian per limit and jumlah tahun sejak pembukaan kredit. Decision tree and neural network model can be created using these variables as inputs. The decision tree will be designed using C5.0 and CHAID algorithms, the two algorithms will produce different number of branches because they differ in how they work. Neural network will use MLP type, since MLP is a type that is often used for prediction classification.

Based on the three models formed the company should use the C5.0 decision tree model because the decision tree method is the easiest method to implement. The algorithm chosen is C5.0 because it has a higher level of accuracy compared to the CHAID algorithm Although the neural network has a high level of accuracy, the method is difficult to implement.

KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa atas segala berkat dan karunianya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Penerapan Model *Credit Scoring* Di Perusahaan X". Pembuatan skripsi ini diajukan sebagai salah satu syarat untuk mencapai gelar sarjana pada bidang ilmu Teknik Industri Universitas Katolik Parahyangan Bandung.

Pada proses pembuatan skripsi ini, penulis mendapat banyak sekali bantuan dan dukungan dari berbagai pihak berupa komentar, saran, dan motivasi yang sangat membantu sehingga skripsi ini dapat terselesaikan dengan baik. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada pihak yang telah mendukung proses pembuatan skripsi ini, terutama kepada :

1. Bapak Fransiscus Rian Praktiko, S.T., M.T., M.S. selaku dosen pembimbing yang telah meluangkan waktunya untuk membimbing, memberikan saran, memberikan arahan, mengevaluasi, dan memberi semangat kepada penulis sehingga penulis dapat menyelesaikan skripsi dengan baik.
2. Bapak Romy Loice, S.T., M.T. selaku Kepala Program Studi Teknik Industri dan Koordinator Skripsi yang telah memberikan arahan dan gambaran mengenai skripsi melalui proses *briefing*.
3. Ibu Catharina Badra Nawangpalupi S.T., M.Eng.Sc., MTD., Ph.D. dan Ibu Paulina Kus Ariningsih, S.T., M.Sc selaku dosen penguji sidang skripsi yang telah memberikan saran dan masukan kepada penulis dalam proses pembuatan skripsi ini.
4. Orang tua yang selalu memberikan dukungan, semangat, dan doa kepada penulis selama menempuh perkuliahan di Teknik Industri Universitas Katolik Parahyangan dan selama proses pembuatan skripsi.
5. Pihak-pihak lainnya yang tidak dapat disebutkan satu per satu yang turut memberikan bantuan dan dukungan kepada penulis selama proses pembuatan skripsi.

Penulis berharap dengan terselesaikannya skripsi ini, Perusahaan X dapat melakukan pengambilan keputusan dalam menilai kredit macet pengguna

jasa kartu kredit dengan sistematis. Selain itu, penulis juga berharap agar skripsi ini dapat berguna bagi pihak peneliti lain atau pihak lain yang berkepentingan. Akhir kata, semoga Tuhan Yang Maha Esa membalas semua kebaikan yang telah diberikan oleh setiap pihak.

Bandung, Juli 2019

Penulis

DAFTAR ISI

ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR	ix
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	
I.1 Latar Belakang Masalah.....	I-1
I.2 Identifikasi Masalah.....	I-3
I.3 Pembatasan Masalah dan Asumsi Penelitian.....	I-5
I.4 Tujuan Penelitian	I-6
I.5 Manfaat Penelitian	I-6
I.6 Metodologi Penelitian	I-6
I.7 Sistematika Penulisan	I-9
BAB II TINJAUAN PUSTAKA	
II.1 Kredit	II-1
II.2 <i>Data Mining</i>	II-1
II.2.1 Klasifikasi	II-2
II.3 CRISP-DM	II-3
II.4 Analisis Variabel dan Skala Pengukuran Variabel	II-4
II.5 Uji Asosiasi	II-5
II.6 <i>Decision Tree</i>	II-5
II.6.1 C5.0	II-6
II.6.2 <i>Chi-squared Automatic Interaction Detector (CHAID)</i>	II-7
II.7 <i>Neural Network</i>	II-7
BAB III PENGUMPULAN DAN PENGOLAHAN DATA	
III.1 Pemahaman Penelitian	III-1
III.2 Pemahaman Data	III-2
III.3 Persiapan Data	III-4

III.3.1 Uji Asosiasi	III-4
III.4 Pemodelan	III-14
III.4.1 Pemodelan <i>Decision Tree</i>	III-14
III.4.1.1 Pemodelan <i>Decision Tree</i> C5.0	III-14
III.4.1.2 Pemodelan <i>Decision Tree</i> CHAID.....	III-15
III.4.2 Pemodelan <i>Neural Network</i>	III-16
III.5 Perbandingan Model	III-21
III.6 Usulan Penerapan	III-23
BAB IV ANALISIS	
IV.1 Analisis Variabel dan Skala Pengukuran Variabel.....	IV-1
IV.2 Analisis Persiapan Data.....	IV-2
IV.3 Analisis Pemodelan <i>Decision Tree</i> C5.0	IV-3
IV.4 Analisis Pemodelan <i>Decision Tree</i> CHAID.....	IV-4
IV.5 Analisis Pemodelan <i>Neural Network</i>	IV-4
IV.6 Analisis Pemilihan Model	IV-5
IV.7 Analisis Penerapan	IV-5
BAB V KESIMPULAN DAN SARAN	
V.1 Kesimpulan.....	V-1
V.2 Saran.....	V-2
DAFTAR PUSTAKA	
LAMPIRAN	
RIWAYAT HIDUP PENULIS	

DAFTAR TABEL

Tabel III.1 Tabel Skala Pengukuran Variabel	III-3
Tabel III.2 Contoh Hasil Asosiasi Variabel Interval-Interval	III-4
Tabel III.3 Rekapitulasi Asosiasi Variabel Interval-Interval	III-5
Tabel III.4 Hasil Uji Asosiasi Variabel Ordinal-Ordinal.....	III-9
Tabel III.5 Tabel Contoh Hasil Tabulasi Silang Variabel Nominal-Ordinal.....	III-10
Tabel III.6 Tabel Hasil <i>Chi Square</i> Variabel Nominal-Ordinal.....	III-10
Tabel III.7Tabel Hasil Asosiasi Variabel Nominal-Ordinal.....	III-10
Tabel III.8 Rekapitulasi ASosiasi Variabel Nominal-Ordinal	III-10
Tabel III.9 Tabel Hasil Asosiasi Variabel Nominal-Interval.....	III-11
Tabel III.10 Rekapitulasi Asosiasi Variabel Nominal-Interval	III-11
Tabel III.11 Tabel Hasil Asosiasi Variabel Ordinal-Interval	III-12
Tabel III.12 Rekapitulasi Asosiasi Ordinal-Interval.....	III-12
Tabel III.13 Rekapitulasi Variabel Terpilih	III-13
Tabel III.14 <i>Predictor Importance Decision Tree</i> C5.0	III-15
Tabel III.15 <i>Predictor Importance Decision Tree</i> CHAID.....	III-16
Tabel III.16 Rekapitulasi Bobot Kombinasi	III-17
Tabel III.17 <i>Predictor Importance Neural Network</i>	III-21
Tabel III.18 Perbandingan Tingkat Keakurasian Antar Model	III-21
Tabel III.19 Perbandingan <i>Predictor Importance</i> Antar Metode	III-22
Tabel III.20 Perbandingan <i>Output</i> Masing-masing Model Dengan Data Awal. III-22	
Tabel III.21 Contoh Usulan Penerapan <i>Decision Tree</i> pada Microsoft Excel ..	III-23

DAFTAR GAMBAR

Gambar I.1 Metodologi Penelitian.....	I-8
Gambar II.1 Enam Fase CRISP-DM	II-3
Gambar II.2 Model Sederhana <i>Neural Networks</i>	II-8
Gambar III.1 Hasil Analisis Model <i>Decision Tree</i> C5.0.....	III-14
Gambar III.2 Hasil Analisis <i>Decision Tree</i> CHAID.....	III-15
Gambar III.3 Ilustrasi Jaringan Hasil Pemodelan	III-16
Gambar III.4 Hasil Analisis <i>Neural Network</i>	III-20

DAFTAR LAMPIRAN

Lampiran A Uji Asosiasi Variabel Nominal-Ordinal

Lampiran B Uji Asosiasi Variabel Nominal-Interval

Lampiran C Uji Asosiasi Variabel Ordinal-Interval

Lampiran D Model *Decision Tree* C5.0

Lampiran E Model *Decision Tree* CHAID

Lampiran F Model *Neural Network*

BAB I

PENDAHULUAN

Pada bab pendahuluan akan dibahas mengenai latar belakang permasalahan, identifikasi permasalahan, penentuan batasan dan asumsi penelitian, tujuan dari penelitian, manfaat penelitian, metodologi penelitian yang dilakukan, serta sistematika penulisan dari penelitian yang dilakukan.

I.1 Latar Belakang Masalah

Menurut Otoritas Jasa Keuangan (2016) kredit merupakan fasilitas keuangan yang memungkinkan seseorang atau badan usaha untuk membeli produk dan membayarnya kembali dalam jangka waktu yang ditentukan dengan dikenakan biaya. Pembayaran yang dilakukan konsumen tidak selalu sesuai dengan jangka waktu yang ditentukan sehingga dapat merugikan perusahaan. Untuk meminimasi kemungkinan terjadinya pembayaran yang terlambat, perusahaan melakukan seleksi terhadap calon konsumen yang mengajukan kredit. Seleksi tersebut dinamakan *credit scoring*.

Credit scoring membantu perusahaan dalam membuat keputusan untuk memberikan kredit kepada konsumen yang mengajukan kredit. *Credit scoring* sendiri merupakan hasil dari sebuah algoritma menggunakan data historis yang dimiliki perusahaan. Data historis yang dapat digunakan bisa berupa besarnya total pemakaian, tagihan yang tidak terbayar dan lain-lain. Bila perusahaan memiliki *credit scoring* yang baik maka perusahaan dapat meminimasi calon konsumen yang akan mengalami kredit macet, karena bila sebuah bank memiliki tingkat kredit macet yang tinggi maka keberlangsungan dari bank tersebut akan terancam.

Perusahaan X merupakan perusahaan yang bergerak di bidang perbankan. Salah satu jasa yang diberikan oleh perusahaan X adalah pemberian kredit. Perusahaan X dalam memberikan kredit kartu kredit macet yang terjadi sebesar IDR 23.763.692107 atau dengan kata lain kredit yang macet sebesar 16%. Sedangkan perusahaan lain memiliki tingkat kemacetan pada kartu kredit sebesar 2,8% pada tahun 2018. Perbedaan yang cukup besar antara

perusahaan X dengan perusahaan lain menunjukkan bahwa perusahaan perlu melakukan perbaikan dalam penilaian kredit pengguna jasa.

Menurut Turner, Casem dan Nazemets (1993) teknik industri merupakan ilmu yang didasarkan oleh beberapa ilmu, antara lain matematika, fisika, ilmu sosial, dan manajemen. Ilmu-ilmu tersebut digunakan untuk merancang, memperbaiki, mengembangkan, dan mengaplikasikan sistem yang terintegrasi, di mana sistem yang terintegrasi mencakup manusia, *material*, informasi, peralatan, dan energi. Informasi menjadi salah satu elemen penting yang dipelajari oleh teknik industri, di mana teknik industri menggunakan informasi untuk memecahkan masalah. Salah satu tujuan dari mempelajari informasi adalah pengembangan dari suatu sistem. Pengembangan sistem yang ada dapat dilakukan dengan penggalian pengetahuan dari informasi-informasi yang dimiliki.

Menurut Larose dan Larose (2014) *data mining* merupakan proses pencarian pola dan tren dari sebuah kelompok data dengan jumlah besar. Pengolahan data dilakukan mengubah data menjadi informasi yang berguna bagi pengguna data. *Data mining* memiliki fungsi klasifikasi, di mana fungsi klasifikasi digunakan untuk mengklasifikasikan variabel dengan mempertimbangkan beberapa variabel (Larose dan Larose, 2014). Fungsi klasifikasi dapat membantu pengguna data untuk mengetahui resiko dari sebuah kejadian berdasarkan beberapa variabel jika jumlah data terlalu banyak atau jika jumlah variabel yang terlalu banyak. Terdapat beberapa algoritma yang dapat digunakan untuk klasifikasi serta perlu dilakukan penyesuaian antara algoritma yang digunakan dengan kasus yang terjadi sehingga didapatkan hasil yang dapat diimplementasikan.

Klasifikasi yang berguna untuk mengetahui resiko dari sebuah kejadian akan membantu perusahaan dalam mengetahui kemungkinan atau kecenderungan konsumen. Salah satu contoh kegunaan klasifikasi adalah mengetahui kecenderungan konsumen yang hendak menggunakan jasa peminjaman kredit akan mengalami pembayaran kredit macet atau lancar berdasarkan data historis yang ada. Implementasi untuk perusahaan dapat membantu perusahaan dalam mengambil keputusan dalam waktu yang singkat dan akurat. Objek dari penelitian yang dilakukan merupakan salah satu perusahaan yang memberikan jasa peminjaman kredit. Perusahaan yang memberikan peminjaman kredit tersebut masih mengalami banyaknya

pembayaran kredit macet sehingga dibutuhkan perbaikan. Perbaikan yang akan dilakukan menggunakan klasifikasi bertujuan untuk membantu perusahaan dalam melakukan pengembangan sistem terutama dalam membantu perusahaan mengambil keputusan apakah konsumen akan mengalami kredit macet atau lancar. Proses klasifikasi menggunakan *Decision Tree* algoritma C5.0 dan *Chi-square interaction Detector* (CHAID) serta *neural network* yang berbeda akan menghasilkan nilai performansi yang berbeda. Hasil performansi tersebut dapat dibandingkan dan berdasarkan perbandingan antara hasil performansi dari masing-masing algoritma dapat diketahui metode apa yang sesuai untuk diaplikasikan terhadap kasus tersebut.

Pengembangan sebuah sistem dapat dilakukan dengan memperbaiki atau melakukan pembaharuan dari sistem yang ada. Penelitian ini dilakukan untuk memperbaiki proses pengambilan keputusan terhadap konsumen yang mengajukan kredit. Proses yang akan diperbaiki bertujuan untuk membantu perusahaan dalam mengambil keputusan dalam *credit scoring* sehingga permasalahan kredit macet dapat diminimasi.

I.2 Identifikasi Masalah

Perusahaan X merupakan perusahaan yang bergerak di bidang perbankan. Salah satu layanan yang diberikan oleh perusahaan X adalah memberikan kredit kepada masyarakat. Dalam pemberian kredit kepada masyarakat perusahaan X dituntut untuk mampu mengelola pinjaman konsumennya dengan cermat. Salah satu cara Perusahaan dalam mengelola pinjaman konsumennya dengan melakukan *credit scoring* terlebih dahulu. *Credit scoring* merupakan model yang dapat membantu perusahaan dalam memprediksi calon konsumen akan mengalami pembayaran kredit macet atau tidak. Meski telah dilakukan *credit scoring* menurut salah satu kepala cabang dari perusahaan X masih terdapat kasus *fraud* di mana konsumen tidak dapat membayar kredit yang dipinjam. *Credit scoring* yang dilakukan merupakan hasil rapat pada bagian kredit dan hasil pencarian informasi mengenai konsumsi pengguna jasa seperti banyaknya pengeluaran, tagihan, dan lain-lain.

Data yang digunakan untuk melakukan *credit scoring* diambil berdasarkan data peminjaman kredit yang dilakukan oleh konsumen sebelumnya. Data-data yang digunakan mencakup jumlah kartu aktif yang dimiliki pelanggan,

total saldo pemakaian kartu kredit, jumlah maksimum limit kartu kredit yang dapat digunakan, dan lain-lain. Data-data historis yang ada tersebut dipelajari oleh perusahaan kemudian berdasarkan pengetahuan yang didapatkan dibuat serangkaian penilaian-penilaian untuk dijadikan acuan penentu apakah calon konsumen akan mengalami kredit macet atau tidak.

Jumlah data yang tidak sedikit dan bervariasi menyebabkan sulit untuk dilakukan penentuan peluang apakah konsumen pengguna jasa kartu kredit akan mengalami kemacetan dalam pembayaran atau tidak. Menurut Larose dan Larose (2014) jumlah data yang cukup banyak dan bervariasi tersebut cocok untuk diolah menggunakan *data mining*. Fungsi *data mining* yang akan digunakan adalah klasifikasi, di mana fungsi tersebut dapat digunakan untuk menggolongkan berdasarkan data yang ada serta dapat memprediksi data baru masuk ke dalam kelompok yang mana (Larose dan Larose, 2014). Klasifikasi yang dilakukan memiliki variabel target apakah kredit dari konsumen macet atau tidak. Berdasarkan data yang dikumpulkan dari data historis yang dilakukan perusahaan terdapat kemungkinan data tersebut dapat membantu perusahaan dalam memprediksi konsumen akan mengalami kredit macet atau lancar.

Dalam pengklasifikasian yang akan dilakukan akan digunakan beberapa metode. Metode yang digunakan antara lain *decision trees* dan *neural network*. Menurut Wendler dan Grottrup (2016) *decision trees* merupakan salah satu metode klasifikasi berdasarkan aturan atau *rule-based*. Klasifikasi berdasarkan aturan dilakukan dengan menginspeksi data mentah untuk aturan dan struktur umum yang ada pada target (Wendler dan Grottrup, 2016). *Decision trees* sebagai klasifikasi berdasarkan aturan memiliki kelebihan untuk lebih mudah dimengerti dan diinterpretasi tanpa menggunakan pengetahuan statistika (Wendler dan Grottrup, 2016).

Algoritma yang digunakan untuk metode *decision tree* adalah C5.0 dan *Chi-squared Automatic Interaction Detector* (CHAID). Menurut Wendler dan Grottrup (2016) *Decision tree* merupakan metode klasifikasi yang menghasilkan aturan dan struktur yang dapat ditemukan dengan menganalisa variabel target, berdasarkan aturan-aturan yang ditemukan tersebut dijadikan dasar dalam pengambilan keputusan. *Decision tree* dipilih karena dapat menghasilkan prediksi yang cepat serta hasil yang mudah diinterpretasi (Wendler dan Grottrup, 2016). C5.0 dipilih karena algoritma tersebut dapat menghasilkan *error* yang cukup kecil

serta algoritma C5.0 menghasilkan pohon keputusan yang cenderung kecil (Wendler dan Grottrup, 2016). CHAID dipilih karena CHAID akan terus membentuk cabang hingga tidak dapat dibentuk lagi cabang selain itu CHAID dapat menggunakan berbagai macam data, seperti nominal, ordinal dan interval yang memungkinkan pohon keputusan yang terbentuk memperhitungkan seluruh variabel yang ada (Wendler dan Grottrup, 2016).

Menurut Wendler dan Grottrup (2016) metode *neural network* merupakan salah satu metode klasifikasi terbaik karena tingkat kompleksitas dan fleksibilitas yang tinggi. Kompleksitas dan fleksibilitas yang tinggi tersebut menyebabkan *neural network* dapat diaplikasikan terhadap berbagai macam permasalahan, selain itu *neural network* memiliki sistem pembelajaran yang dapat mengatasi hubungan kompleks dari variabel serta *neural network* cocok untuk mengolah data yang memiliki jumlah besar (Wendler dan Grottrup, 2016). Keunggulan *neural network* karena tingkat kompleksitas dan fleksibilitas yang tinggi serta data yang digunakan dalam jumlah besar dan cenderung memiliki hubungan yang kompleks antar variabelnya sehingga metode ini dipilih untuk membentuk model prediksi *credit scoring*. *Neural network* yang akan digunakan adalah *multilayer perceptrons* (MLP). MLP merupakan tipe *neural network* yang sangat cocok untuk klasifikasi prediksi dimana pada penelitian ini merupakan penelitian klasifikasi prediksi terhadap kredit pengguna jasa.

Berdasarkan identifikasi masalah yang telah dipaparkan, dibuat rumusan masalah seperti berikut:

1. Bagaimana perbandingan performansi model *decision tree* dan *neural network* dalam memprediksi resiko kredit pada perusahaan X?
2. Model prediksi mana yang sebaiknya digunakan oleh perusahaan X?

I.3 Pembatasan Masalah dan Asumsi Penelitian

Pembatasan masalah yang digunakan dan dibuat dalam penelitian bertujuan agar penelitian yang dilakukan tidak keluar dari konteks penelitian dan dapat menjawab tujuan diadakannya penelitian. Batasan masalah dalam penelitian ini adalah:

1. Penelitian dilakukan dengan metode *data mining*.
2. Pemodelan dilakukan menggunakan *Clementine*.
3. Penelitian hanya dilakukan hingga tahap evaluasi model.

Selain batasan masalah penelitian ini memiliki asumsi yang digunakan. Asumsi yang digunakan dalam penelitian ini adalah data yang digunakan dapat merepresentasikan keadaan saat ini.

I.4 Tujuan Penelitian

Tujuan penelitian merupakan sesuatu yang ingin dicapai dari penelitian dan sesuai dengan identifikasi masalah serta rumusan masalah yang telah dirumuskan. Adapun tujuan yang ingin dicapai dari penelitian sebagai berikut:

1. Mengetahui perbedaan performansi akurasi dari *decision tree* dan *neural network*.
2. Mengetahui model yang cocok dalam memprediksi resiko kredit pada perusahaan X

I.5 Manfaat Penelitian

Penelitian yang dilakukan diharapkan dapat memberikan beberapa manfaat bagi peneliti dan pembaca. Berikut adalah hal-hal yang diharapkan dapat menjadi manfaat dari penelitian ini.

1. Perusahaan mendapatkan solusi dari permasalahan yang ada, berupa model klasifikasi konsumen kartu kredit
2. Pembaca dapat menambah ilmu mengenai *data mining* khususnya fungsi klasifikasi.
3. Memberikan refrensi bagi penelitian yang akan dilakukan selanjutnya.

I.6 Metodologi Penelitian

Pada bab ini akan dibahas mengenai metodologi penelitian yang akan dilakukan. Tahapan-tahapan yang dilakukan dalam penelitian ini dapat dilihat pada gambar I.1 dan penjelasan dari tahapan-tahapan yang dilakukan adalah sebagai berikut.

1. Identifikasi dan Perumusan Masalah
Setelah dilakukan penelitian pendahuluan dilakukan identifikasi masalah terhadap objek penelitian. Berdasarkan identifikasi masalah dilakukan perumusan masalah dari objek penelitian.
2. Penentuan Batasan dan Asumsi Masalah

Dalam Tahapan ini ditentukan batasan dan asumsi dengan masalah yang diteliti. Batasan dan asumsi ini ditentukan agar penelitian tidak keluar dari konteks permasalahan.

3. Studi Literatur

Setelah ditentukan tujuan penelitian dilakukan studi literature terkait yang berhubungan dengan permasalahan. Studi literature dilakukan terhadap refrensi buku, jurnal, dan artikel.

4. Pemahaman Tujuan Bisnis

Pada tahap ini dilakukan pemahaman terhadap tujuan serta kebutuhan bisnis. Berdasarkan pemahaman tersebut ditentukan permasalahan *data mining* dan tujuan *data mining*.

5. Pemahaman Data

Pemahaman data dilakukan agar dapat menemukan informasi tambahan yang ada pada data. Tahapan dilakuan dengan pengumpulan data terlebih dahulu kemudian dilakukan eksplorasi data untuk menemukan wawasan serta informasi dalam data.

6. Persiapan Data

Persiapan data dilakukan agar data yang ada dapat diolah. Persiapan data dilakukan untuk pemodelan agar dapat dilakukan pemodelan dengan baik. Persiapan dilakukan dengan melakukan uji asosiasi dan penentuan variabel. Sehingga dapat dilakukan pemodelan berdasarkan data yang telah disiapkan

7. Perancangan Model

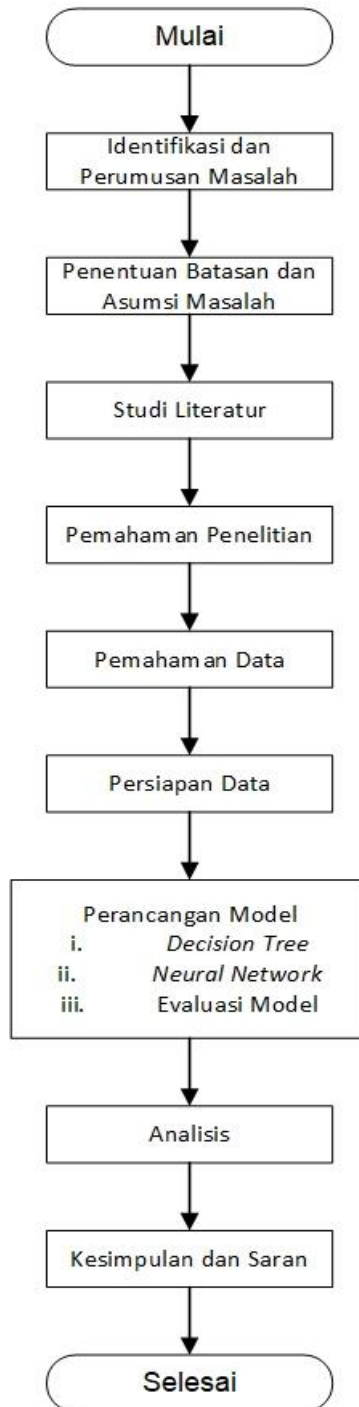
Setelah dilakukan persiapan data, data kemudian diolah. Pengolahan data yang dilakukan adalah pembentukan model klasifikasi dengan menggunakan *decision tree* dengan algoritma C5.0 dan CHAID serta *neural network*. Model dibentuk untuk meramalkan apakah konsumen akan mengalami kredit macet atau tidak. Selain pembentukan model dilakukan juga evaluasi terhadap model yang dibentuk untuk melihat model yang cocok untuk digunakan.

8. Analisis

Setelah diolah hasil penelitian akan dilakukan analisis. Analisis akan berisi mengenai persiapan data, pemodelan data, dan hasil evaluasi dari model.

9. Kesimpulan dan Saran

Setelah penelitian dilakukan akan dilakukan penarikan kesimpulan dari hasil pengolahan data serta pemberian usulan untuk menjawab tujuan penelitian.



Gambar I.1 Metodologi Penelitian

I.7 Sistematika Penulisan

Dalam penelitian ini sistematika penulisan dibuat bertujuan untuk memudahkan penyampaian dari penelitian dengan membagi penelitian menjadi 5 bab yang terdiri atas pendahuluan, tinjauan pustaka, pengumpulan dan pengolahan data, analisis, serta kesimpulan dan saran. Berikut akan dijelaskan bagian sistematika penulisan yang lebih rinci.

BAB I PENDAHULUAN

Bab I merupakan bagian awal yang menjelaskan mengenai latar belakang permasalahan yang terjadi, dilanjutkan dengan identifikasi permasalahan yang memunculkan rumusan masalah. Pada bab ini juga ditentukan batasan serta asumsi penelitian, tujuan dari penelitian, manfaat penelitian. Metodologi penelitian yang dilakukan, serta sistematika penulisan dari penelitian yang dilakukan.

BAB II TINJAUAN PUSTAKA

Bab II menjelaskan mengenai teori-teori yang digunakan dalam penelitian. Teori yang akan dipaparkan adalah *data mining*, CRISP-DM, skala pengukuran, uji korelasi, klasifikasi, *decision tree*, dan *neural network*

BAB III PENGUMPULAN DAN PENGOLAHAN DATA

Bab ini akan berisi proses pengumpulan data, pemahaman data dan pemilihan variabel. Pemilihan variabel dilakukan dengan melakukan uji korelasi. Kemudian dilakukan pemodelan dari *decision tree* dan *neural network*. Setelah itu dilakukan perbandingan model yang telah dibentuk.

BAB IV ANALISIS

Bab ini akan berisi analisis dari penelitian yang telah dilakukan. Analisis akan mencakup pemodelan *decision tree*, pemodelan *neural network*, dan perbandingan model yang dihasilkan

BAB V KESIMPULAN DAN SARAN

Bab ini akan berisi mengenai kesimpulan dari penelitian yang telah dilakukan. Kesimpulan yang diambil merupakan jawaban dari tujuan penelitian yang telah dilakukan, selain itu terdapat saran dari penelitian yang telah dilakukan.