

**PERBANDINGAN ALGORITMA NAIVE BAYES DAN
DECISION TREE DALAM MEMPREDIKSI APPROVAL
PEMBERIAN KREDIT STUDI KASUS: PT BATAVIA
PROSPERINDO FINANCE TBK. CABANG PRINGSEWU**

TESIS



Disusun Oleh:

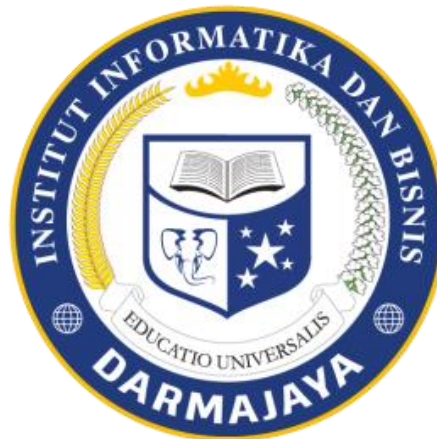
Rendi Irawan
1921211022

**PROGRAM STUDI MAGISTER TEKNIK INFORMATIKA
PROGRAM PASCASARJANA
INSTITUT INFOTMATIKA & BISNIS DARMAJAYA
BANDAR LAMPUNG
2021**

**PERBANDINGAN ALGORITMA NAIVE BAYES DAN DECISION TREE
DALAM MEMPREDIKSI APPROVAL PEMBERIAN KREDIT STUDI
KASUS: PT BATAVIA PROSPERINDO FINANCE TBK.
CABANG PRINGSEWU**

TESIS

**Diajukan Sebagai Salah Satu Syarat untuk Mencapai Gelar
MAGISTER TEKNIK INFORMATIKA
Pada Program Studi Magister Teknik Informatika
IIB Darmajaya Bandar Lampung**



Disusun Oleh :

**Rendi Irawan
1921211022**

**MAGISTER TEKNIK INFORMATIKA
PROGRAM PASCASARJANA
INSTITUT INFORMATIKA DAN BISNIS DARMAJAYA
BANDAR LAMPUNG
2021**



PERNYATAAN

Saya yang bertanda tangan dibawah ini menyatakan, bahwa skripsi yang saya ajukan ini adalah hasil karya saya sendiri, tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesajanaan di suatu perguruan tinggi atau karya yang pernah ditulis atau diterbitkan orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka. Karya ini adalah milik saya dan pertanggungjawaban sepenuhnya berada di pundak saya.

Bandar Lampung, Februari 2021

Rendi Irawan
1921211022

HALAMAN PERSETUJUAN

Judul Tesis	Perbandingan Algoritma Naive Bayes Dan Decision Tree Dalam Memprediksi Approval Pemberian Kredit Studi Kasus: PT Batavia Prosperindo Finance Tbk. Cabang Pringsewu
Nama Mahasiswa	Rendi Irawan
NPM	1921211022
Program Studi	Magister Teknik Informatika

Disetujui oleh :

Pembimbing

Ketua Program Studi



Dr.Ir.Kurnia Muludi, M.S.Sc
NIP. 19640616 198902 1 001

Dr.Sutedi, S.Kom., M.T.I
NIK. 00600303

HALAMAN PENGESAHAN

Telah diuji dan dipertahankan didepan Tim Penguji Tesis
Program Pascasarjana IIB Darmajaya dan dinyatakan diterima
untuk memenuhi syarat guna memperoleh gelar
Magister Teknik Informatika

Mengesahkan,

- | | |
|---|--------------|
| 1. Tim Penguji | Tanda Tangan |
| Ketua Penguji | |
| Joko Triloka, S.Kom., MT, Ph.D | (.....) |
| NIK. 00610303 | |
| Penguji I | |
| Dr. Chairani, S.Kom., M.Eng | (.....) |
| NIK. 01190305 | |
| 2. Dekan Fakultas Ilmu Komputer | |
| Program Studi Magister Teknik Informatika | |

Zaidir Jamal, ST., M.Eng
NIK.00590203

Tanggal Lulus Ujian Tesis : Februari 2021

DAFTAR RIWAYAT HIDUP

1. Identitas

a. Nama	Rendi Irawan
b. NPM	1921211022
c. Tempat/Tanggal Lahir	Kutawaringin, 12 Januari 1993
d. Agama	Islam
e. Alamat	Jl.Seroja Gg.Apel Lk V Rt/Rw.005/- Kel. Pringsewu Barat Kec. Pringsewu Kab.Pringsewu -35373
f. Suku	Jawa
g. Kewarganegaraan	Indonesia
h. E-mail	ren.irawan33@gmail.com
	HP : 0822 8145 3169

2. Riwayat Pendidikan

a. Sekolah Dasar	: SDN 10 Bandung Baru - Pringsewu
b. Sekolah Menengah Pertama	: SMPN 1 Adiluwih - Pringsewu
c. Sekolah Menengah Atas	: SMK Karya Bhakti - Pringsewu
d. Sarjana	: STMIK Pringsewu

Dengan ini saya menyatakan bahwa semua keterangan yang saya sampaikan di atas adalah benar.

Yang menyatakan,

Bandar Lampung, 26 Februari 2021

Rendi Irawan

1921211022

PERSEMBAHAN

Segala yang kuraih adalah kehendak Tuhan YME dan bukti kasih sayang dari orang-orang yang menyayangiku, dengan mengucapkan syukur kepadaNya dan atas segala limpahan nikmatNya kepadaku dan segala ketulusan dan kerendahan hati kupersembahkan sebuah karya kecil hasil jerih payah perjuanganku ini untuk:

1. Orang tuaku,Istri dan Calon Bauh Hatiku yang sangat aku cintai Bpk. Sakiman,Ibu Pains dan Gandis Ghardya Demusfita telah berkorban serta sabar dengan segenap jiwa dan raga, mencurahkan segala cinta serta senantiasa berdoa demi keberhasilanku
2. Dosen pembimbingku Bpk. Dr.Ir.Kurnia Muludi, M.S.Sc yang telah meluangkan waktunya untuk menuntun penulis serta memberikan koreksi dan saran.
3. Almamaterku tercinta Perguruan Tinggi IIB Darmajaya Bandar Lampung.

MOTTO

“Biarkan Orang Lain Menilai Kita Dengan Penilaiannya Masing-Masing
Tetaplah Selalu Berbuat Baik Menurut Kita Itu Baik.
Sebaik Baik Nya Pemberi Nilai Yaitu Allah Swt. ”

(Penulis)

ABSTRAK**PERBANDINGAN ALGORITMA NAIVE BAYES DAN
DECISION TREE DALAM MEMPREDIKSI APPROVAL
PEMBERIAN KREDIT STUDI KASUS:
PT BATAVIA PROSPERINDO FINANCE TBK.
CABANG PRINGSEWU**

Oleh

Rendi Irawan

ren.irawan33@gmail.com

Pertumbuhan ekonomi suatu bangsa sangat dipengaruhi oleh perkembangan dinamis dan kontribusi nyata dari dunia *finance* yang juga merupakan lembaga keuangan dimana kegiatan operasionalnya memberikan dana pinjaman dengan prinsip kepercayaan masyarakat. PT Batavia Prosperindo Finance Tbk merupakan perusahaan yang bergerak dibidang pembiayaan dana pinjaman dimana saat ini faktor yang besar pengaruhnya adalah kurang akuratnya penilaian terhadap kemampuan *debitur* didalam melunasi pinjamannya sehingga sering menyebabkan kredit yang bermasalah. Perusahaan mengalami kesulitan dalam menentukan pemberian pinjaman kredit karena belum adanya sistem untuk pengambilan keputusan dalam proses *Approval* data nasabah dalam memprediksi kelayakan pemberian kredit. Untuk itu akan dikembangkan sebuah penerapan *data mining* yang berfungsi untuk menentukan kriteria nasabah. Teknik *data mining* yang diterapkan adalah klasifikasi, sedangkan metode klasifikasi yang digunakan adalah *Decision Tree* (pohon keputusan) dan *Naïve Bayes*. Terdapat tiga parameter uji yang digunakan sebagai evaluasi sistem yaitu *accuracy*, *precision* dan *recall*. Data akan diuji dengan menggunakan *k-folds cross validation* (k=10). Dari hasil perbandingan tersebut didapat hasil dari akurasi algoritma *Decision Tree* lebih unggul dibanding dengan algoritma *Naïve Bayes*. Hasil yang didapat adalah algoritma *Decision Tree* memiliki tingkat akurasi sebesar 85,22 dan algoritma *Naïve Bayes* memiliki tingkat akurasi sebesar 73,67%. Dan kurva ROC *decision tree* tingkat diagnosa adalah *excellent classification* sedangkan *Naïve Bayes* adalah *good classification*.

Kata Kunci: Kredit, *Data Mining*, *Decision Tree*, *Naïve Bayes*

ABSTRACT**PERBANDINGAN ALGORITMA NAIVE BAYES DAN
DECISION TREE DALAM MEMPREDIKSI APPROVAL
PEMBERIAN KREDIT STUDI KASUS:
PT BATAVIA PROSPERINDO FINANCE TBK.
CABANG PRINGSEWU**

By

Rendi Irawan

ren.irawan33@gmail.com

The economic growth of a nation is strongly influenced by dynamic developments and real contributions from the world of finance, which is also a financial institution where its operations provide loan funds based on the principle of public trust. PT Batavia Prosperindo Finance Tbk is a company engaged in the financing of loan funds where currently the major influencing factor is the inaccurate assessment of the debtor's ability to pay off his loan, which often causes credit problems. The company has difficulty in determining credit lending because there is no system for decision making in the process of approving customer data in predicting credit worthiness. For this reason, an application of data mining will be developed which functions to determine customer criteria. The data mining technique applied is classification, while the classification method used is the Decision Tree (decision tree) and Naïve Bayes. There are three test parameters used for system evaluation, namely accuracy, precision and recall. Data will be tested using k-folds cross validation ($k = 10$). From the results of these comparisons, the accuracy of the Decision Tree algorithm is superior to the Naïve Bayes algorithm. The results obtained are that the Decision Tree algorithm has an accuracy rate of 85.22 and the Naïve Bayes algorithm has an accuracy rate of 73.67%. And the ROC curve of the decision tree for the diagnostic level is excellent classification while Naïve Bayes is a good classification.

Keywords: Credit, Data Mining, Decision Tree, Naïve Bayes

KATA PENGANTAR

Segala puji syukur saya panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat dan kasih-Nya sehingga penulis dapat menyelesaikan penyusunan tesis ini sebagai salah satu syarat untuk menyelesaikan program Pendidikan Magister Teknik Informatika di Program Pascasarjana.

Penulis menyadari bahwa dalam penyusunan tesis ini dapat diselesaikan berkat dukungan dan bantuan semua pihak. Oleh karena itu dalam kesempatan ini penulis menyampaikan terima kasih kepada semua pihak yang telah memberikan kontribusi dalam penyelesaian tesis ini, secara khusus ucapan terima kasih penulis sampaikan kepada yang terhormat :

1. Bapak DR.Andi Desfiandi.,S.E.,M.A. Selaku Ketua Yayasan Institut Informatika dan Bisnis Darmajaya Bandar Lampung.
2. Bapak Ir.Firmansyah.,YA.,MBA.,M.Sc. Selaku Rektor Institut Informatika dan Bisnis Darmajaya Bandar Lampung.
3. Bapak Dr.R.Z.Abdul Aziz., S.T., M.T. Selaku Wakil Rektor I Bidang Akademik Institut Informatika dan Bisnis Darmajaya Bandar Lampung.
4. Bapak Zaidir Jamal, S.T., M.Eng. Selaku Dekan Fakultas Ilmu Komputer Institut Informatika dan Bisnis Darmajaya Bandar Lampung.
5. Bapak Dr.Sutedi, S.Kom., M.T.I. Selaku Ketua Program Studi Magister Teknik Informatika Institut Informatika dan Bisnis Darmajaya Bandar Lampung.
6. Bapak Dr.Ir.Kurnia Muludi, M.S.Sc. Selaku Dosen Pembimbing yang telah membantu membimbing dan mengarahkan serta memberikan petunjuk sehingga tesis ini dapat diselesaikan dengan baik.
7. Para dosen, staf dan karyawan Informatika Institut Informatika dan Bisnis Darmajaya Bandar Lampung yang telah memberi bantuan baik langsung maupun tidak langsung selama saya menjadi mahasiswa.

8. Orangtua,Istri dan keluargaku yang tiada hentinya memotivasi dan terus memberikan semangat, pengertian, dan kesabaran serta doanya.
9. Sahabat dan Teman-teman yang selalu memberikan dukungan dan semangat.
10. Semua Pihak yang telah memberikan bantuan dan petunjuk sehingga saya dapat lebih mudah dalam menyusun skripsi ini.
11. Almamaterku tercinta.

Demikian banyaknya bantuan berbagai pihak kepada penulis, tentunya tidak menutup kemungkinan bahwa hasil dari laporan ini masih ada kekurangan dan masih jauh dari taraf sempurna. Oleh karena itu, kritik dan saran guna perbaikan di masa depan adalah mutlak sangat penulis perlukan. Semoga laporan tesis ini bermanfaat bagi setiap pembacanya.

Bandar Lampung, 26 Februari 2021

Penulis,

Rendi Irawan

DAFTAR ISI

Halaman

HALAMAN <i>COVER</i> DALAM	
LEMBAR PERNYATAAN	
LEMBAR PENGESAHAN TESIS	
ABSTRAK	i
<i>ABSTRACT</i>	ii
KATA PENGANTAR	iii
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Masalah Penelitian	3
1.2.1 Identifikasi Masalah	4
1.2.3 Rumusan Masalah	4
1.3 Tujuan dan Manfaat Penelitian	5
1.3.1 Tujuan Penelitian	5
1.3.2 Manfaat Penelitian	5
1.4 Tata Urut Penulisan	6
1.5 Daftar Pengertian	7
BAB II LANDASAN TEORI DAN KERANGKA KONSEP	8
2.1 Tinjauan Pustaka	8
2.1.1 Pengertian Kredit	8
2.1.2 Data Mining	9
2.1.3 Klasifikasi	15
2.1.4 <i>K-Fold Cross Validation</i>	15
2.1.5 Algoritma <i>Naïve Bayes</i>	16
2.1.6 Algoritma <i>Decision Tree</i>	17
2.1.7 Pemilihan Variabel	19
2.1.8 Pengujian Akurasi dan Validasi Metode Klasifikasi <i>Data Mining</i>	19
2.1.9 Pengertian Nasabah	21
2.2 Tinjauan Studi	22
2.3 Tinjauan Obyek Penelitian	24
2.4 Kerangka Konsep/Pola Pikir Pemecahan Masalah	24
2.5 Hipotesis	26
BAB III METODOLOGI DAN RANCANGAN/DESAIN PENELITIAN	27
3.1 Metode Penelitian	27
3.2 Metode Pemilihan Sampel	29

3.3 Tahap-Tahap Penelitian	30
3.3.1 Metode Pengumpulan Data	30
3.3.2 <i>Dataset</i> Penelitian	31
3.4 Metode-metode Pendekatan Penyelesaian Permasalahan	33
3.4.1 Analisis Sistem Berjalan	33
3.4.2 Langkah-langkah Penelitian.....	33
3.4.3 Desain Global Sistem Baru	37
3.5 Instrumentasi	39
3.6 Teknik Analisis, Desain, dan Pengujian.	39
3.6.1 Teknik Analisis	39
3.6.2 Desain Proses Klasifikasi pada Prototipe.....	40
3.6.3 Teknik Pengujian	40
BAB IV PEMBAHASAN HASIL PENELITIAN.....	42
4.1 Akuisisi Data.....	42
4.2 Seleksi Atribut	43
4.3 Pendefinisian Atribut Kelengkapan	44
4.3.1 Pengujian Klasifikasi Dengan Algoritma <i>Decision Tree</i>	48
4.3.2 Pengujian Klasifikasi Dengan Algoritma Naïve Bayes	61
4.3.3 <i>Confusion Matrix</i> , <i>F-measure</i> dan Kurva ROC <i>Decision Tree</i>	63
4.3.4 <i>Confusion Matrix</i> , <i>F-measure</i> dan Kurva ROC <i>Naïve Bayes</i>	65
4.4 Evaluasi dan Validasi.....	66
4.4.1 <i>Confusion Matrix</i> , <i>F-measure</i> dan Kurva ROC <i>Decision Tree</i>	66
4.4.2 <i>Confusion Matrix</i> , <i>F-measure</i> dan Kurva ROC <i>Naïve Bayes</i>	68
4.5 Hasil Komparasi.....	69
4.6 Penerapan Algoritma Terpilih.....	70
4.7 Pengujian <i>Prototype</i> Perangkat Lunak (<i>Software</i>).....	76
4.7.1 Pengujian <i>Prototype</i>	76
4.7.2 Lingkungan Pengujian	78
4.7.3 Hasil Pengujian Kualitas Perangkat Lunak (<i>Software</i>).....	78
4.8 Implikasi Penelitian	80
4.8.1 Aspek Sistem.....	80
4.8.2 Aspek Penelitian Lanjut	80
BAB V PENUTUP.....	81
5.1 Simpulan	81
5.2 Saran	82

DAFTAR PUSTAKA
LAMPIRAN-LAMPIRAN
RIWAYAT HIDUP SINGKAT

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Prinsip Data Mining	9
Gambar 2.2 Proses Data Mining	13
Gambar 2.3 Kerangka Pemikiran.....	25
Gambar 3.1 Proses CRISP-DM.....	28
Gambar 3.2 Langkah-langkah Penelitian.....	34
Gambar 3.3 Bagan Alir Algoritma <i>Naïve Bayes</i>	37
Gambar 3.4 Bagan Alir Algoritma <i>Decision Tree</i>	38
Gambar 4.1 Proses Akuisisi Data.....	42
Gambar 4.2 Hasil Pohon Keputusan Menggunakan RapidMiner.....	58
Gambar 4.3 Nilai AUC dengan <i>Decision Tree</i>	64
Gambar 4.4 Nilai AUC dengan <i>Naïve Bayes</i>	66
Gambar 4.5 Nilai Perbandingan AUC	70
Gambar 4.6 <i>Activity Diagram Login</i>	71
Gambar 4.7 <i>Use Case Login</i>	72
Gambar 4.8 <i>Class Diagram</i>	72
Gambar 4.9 <i>Flowcart Prototype</i>	73
Gambar 4.10 <i>Form Login</i>	74
Gambar 4.11 <i>Form User</i>	74
Gambar 4.12 <i>Form Upload</i>	75
Gambar 4.13 Halaman <i>Dashboard</i>	75
Gambar 4.14 <i>Detail Nasabah</i>	76

DAFTAR TABEL

	Halaman
Tabel 2.1 Model <i>Confusion Matrix</i>	20
Tabel 2.2 Ringkasan Tinjauan Studi	22
Tabel 3.1 <i>Dataset</i> Penelitian	32
Tabel 3.2 Atribut Data Master	35
Tabel 4.1 Atribut Data Penelitian.....	44
Tabel 4.2 Sample Data Training	46
Tabel 4.3 Hasil Perhitungan Entropi.....	57
Tabel 4.4 Perhitungan <i>Probabilitas Prior</i>	61
Tabel 4.5 Probabilitas <i>Posterior</i> Data.....	62
Tabel 4.6 <i>Confesion Matrik</i> Dengan <i>Decision Tree</i>	63
Tabel 4.7 <i>Confesion Matrik</i> Dengan <i>Naïve Bayes</i>	65
Tabel 4.8 Evaluasi <i>Confesion Matrik</i> Dengan <i>Decision Tree</i>	67
Tabel 4.9 Evaluasi <i>Confesion Matrik</i> Dengan <i>Naïve Bayes</i>	68
Tabel 4.10 Perbandingan Akurasi dan AUC.....	69
Tabel 4.11 Pengujian Kotak Hitam.....	76
Tabel 4.12 Pengujian File <i>Upload</i>	77
Tabel 4.13 Pengujian <i>Dashboard</i>	77
Tabel 4.14 Komponen <i>Software Quality Assurance</i>	79
Tabel 4.15 Skor Matrik	79

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Daftar Pertanyaan Wawancara Dngan Pakar	86
Lampiran 2. Kuisisioner terhadap Prototipe	88
Lampiran 3. Dataset Nasabah.....	89
Lampiran 4. Hasil Testing Data Prediksi Nasabah Disetujui.....	94
Lampiran 5. Hasil Testing Data Prediksi Nasabah Ditolak	95
Lampiran 6. <i>Source Code</i>	96
Lampiran 7. Daftar Riwayat Hidup	102

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pertumbuhan ekonomi suatu bangsa sangat dipengaruhi oleh perkembangan dinamis dan kontribusi nyata dari dunia *Finance* yang juga merupakan lembaga keuangan bekerja berdasarkan kepercayaan masyarakat dimana kegiatan operasionalnya adalah memberi pinjaman dana kepada masyarakat. Karenanya *Finance* juga berperan serta dalam meningkatkan kesejahteraan masyarakat dengan cara memberikan jasa-jasa pinjaman modal atau bentuk – bentuk lainnya dalam rangka meningkatkan taraf hidup rakyat banyak[1].

Dalam peraturan Presiden No 9 Tahun 2009 pasal 1, bahwa Lembaga Pembiayaan adalah badan usaha yang melakukan kegiatan pembiayaan dalam bentuk penyediaan dana modal atau barang. Perusahaan Pembiayaan adalah badan usaha yang khusus didirikan untuk melakukan Sewa Guna Usaha, Anjak Piutang, Pembiayaan Konsumen[2]. Berdasarkan UU tersebut, segala bentuk kredit yang dilakukan harus berdasarkan pada persetujuan pinjam meminjam, dimana akan ada suatu analisa yang dilakukan untuk menentukan sebuah pengambilan keputusan

PT Batavia Prosperindo Finance Tbk merupakan perusahaan yang bergerak di bidang pembiayaan dana pinjaman dimana saat ini banyak sekali mengalami fluktuasi ekonomi karena banyaknya nasabah tertunggak/macet yang menyebabkan perusahaan harus lebih tepat dalam memberikan pinjaman pada nasabah. Kendala yang dihadapi yaitu perusahaan mengalami kesulitan dalam menentukan pemberian pinjaman kredit karena belum adanya sistem untuk pengambilan keputusan dalam proses *Approval* data nasabah dalam memprediksi kelayakan pemberian kredit. Pemberian pinjaman dana dalam bentuk kredit kepada nasabah berpotensi tidak kembalinya dana yang telah disalurkan. Sistem

yang tepat merupakan hal yang sangat diperlukan pada saat pemberian kredit kepada nasabah /*debitur*. Sistem ini merupakan bagian dari pengendalian intern yang digunakan oleh finance sebagai pedoman dalam pemberian kredit.

Nasabah PT Batavia Prosperindo Finance Tbk seringkali mengajukan pinjaman ulang atau Top Up penambahan plafon kredit yang sedang berjalan ataupun yang sudah lunas dalam masa pinjaman. Dalam hal ini untuk memprediksi proses Approval kredit sebagai acuan di lihat dari histori paymen yang sudah berjalan. Dalam penelitian ini akan dilakukan percobaan penambangan data dari tahun 2018 -2020 data yang diambil sebanyak 300 *record*, untuk memprediksi apakah nasabah yang mengajukan permohonan kredit dapat di setuju atau tidak oleh pihak finance mengacu pada aspek 5C menurut [3] yaitu meliputi ; *character* (sifat), *capacity* (kemampuan), *capital* (simpanan/modal), *colleteral* (jaminan) dan *condition of economy* (kondisi ekonomi) dengan menggunakan klasifikasi *data mining*.

Dari uraian tersebut, dapat disimpulkan bahwa dalam memberikan kredit PT Batavia Prosperindo Finance Tbk harus mampu meningkatkan sistem pemberian kredit dan berusaha untuk meminimalkan resiko kredit yang bermasalah. Berdasarkan latar belakang tersebut penulis tertarik untuk mengangkat judul "**Perbandingan Algoritma Naive Bayes Dan Decision Tree Dalam Memprediksi Approval Pemberian Kredit Studi Kasus : PT Batavia Prosperindo Finance Tbk. Cabang Pringsewu**"

Pemilihan teknik[4] *data mining* yang tepat merupakan kunci keakuratan hasil dari prediksi suatu kasus yang dianalisa kegiatan yang meliputi pengumpulan dan pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data yang berukuran besar, Keluaran yang dihasilkan oleh klasifikasi data mining dapat digunakan untuk memperbaiki pengambilan keputusan proses (Approval) dibagian analis dalam pemberian kredit. Penelitian ini akan membandingkan dua metode algoritma tersebut yang bertujuan untuk mengetahui tingkat akurasi tertinggi yang akan di jadikan acuan untuk proses Approval kredit pada PT Batavia Prosperindo Finance Tbk untuk menekan

tingkat resiko kemacetan pembayaran nasabah, yaitu dengan metode *Decision tree* dengan *Naïve bayes*.

Decision tree adalah metode klasifikasi yang mudah untuk diinterpretasi oleh manusia *Decision tree* merupakan salah satu metode yang dapat diterapkan pada data kredit karena Algoritma tersebut bekerja dari atas ke bawah, mencari pada setiap tahap atribut untuk membaginya ke dalam bagian terbaik class tersebut dan memproses secara rekursif submasalah yang dihasilkan dari pembagian tersebut yang dapat mengubah data seperti pekerjaan, umur, penghasilan, dan kolektibilitas menjadi pohon keputusan dan aturan-aturan keputusan yang kemudian dapat digunakan untuk klasifikasi nasabah kredit berdasarkan status kelancaran.

Naive bayes merupakan salah satu algoritma pengklasifikasian dengan metode probabilitas dan statistik sederhana yang berdasar pada teorema keputusan bayes merupakan pendekatan statistik yang fundamental dalam pengenalan pola. Naive bayes didasarkan pada asumsi penyederhaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output yang berbeda dengan asumsi independensi yang kuat. Hasil yang diperoleh dari perhitungan naïve bayes dapat mengklasifikasikan nasabah kredit berdasarkan status kelancaran[5]. Algoritma Naive Bayes dapat diartikan sebagai sebuah metode yang tidak memiliki aturan, Naive Bayes menggunakan cabang matematika yang dikenal dengan teori probabilitas untuk mencari peluang terbesar dari kemungkinan klasifikasi dengan cara melihat frekuensi tiap klasifikasi pada data training untuk pengambilan suatu keputusan.

1.2 Masalah Penelitian

1.2.1 Identifikasi Masalah

Dari beberapa uraian yang dikemukakan pada latar belakang, maka dapat diidentifikasi masalah-masalah sebagai berikut:

1. Terjadinya kredit beresiko tinggi/macet setelah disetujui karena tidak adanya sistem yang dapat memprediksi resiko pemberian kredit , diluar

“BI *checking*” sebagai tahap awal seleksi untuk setiap nasabah lama yang mengajukan kredit kembali.

2. Belum tersedianya *tools* evaluasi pengajuan kredit yang mampu memprediksi data-data nasabah dalam melakukan peminjaman kembali dengan lebih efektif dan efisien dengan menerapkan metode prediksi.
3. Belum adanya sebuah metode klasifikasi prediksi dalam menentukan pola data-data nasabah yang akan mengajukan kembali pinjaman untuk dilakukan evaluasi berdasarkan riwayat peminjaman sebelumnya.

1.2.2 Batasan Masalah

Adapun batasan masalah tersebut yaitu sebagai berikut:

1. Algoritma yang digunakan adalah *Naive Bayes* dan *Decision Tree* dengan perhitungan menggunakan bantuan *tools Rapidminer*
2. Data set yang digunakan adalah nasabah yang sudah memiliki kontrak di PT Batavia Prosperindo Finance Tbk dan ingin mengajukan pinjaman kembali.
3. Sistem yang dibuat berupa *prototype* aplikasi berbasis *web*.
4. Objek yang diteliti adalah dokumen-dokumen pengajuan kredit nasabah kepada *finance* dalam bentuk *softcopy* nasabah selama periode 2018-2020, dengan jumlah data yang digunakan sebanyak 300 *record*.

1.2.3 Rumusan Masalah

Berdasarkan latar belakang, identifikasi dan batasan masalah, maka permasalahan penelitian dapat dirumuskan sebagai berikut:

1. Apakah metode klasifikasi *data mining* dapat digunakan untuk memprediksi keputusan pemberian kredit di PT Batavia Prosperindo Finance Tbk?
2. Manakah diantara metode *Naive Bayes* dan *Decision Tree* yang lebih akurat dalam memprediksi pemberian kredit kepada nasabah.

3. Bagaimana hasil akurasi dari algoritma terpilih jika diterapkan ke sebuah aplikasi *prototype* dalam memprediksi data nasabah dengan menggunakan metode klasifikasi data *mining*?

1.3 Tujuan dan Manfaat penelitian

1.3.1 Tujuan Penelitian

Penelitian tesis ini bertujuan untuk merekomendasikan *prototype* metode klasifikasi *data mining* untuk prediksi pemberian kredit kepada nasabah berdasarkan metode klasifikasi *data mining* yaitu *algoritma Naive Bayes* dan *Decision Tree* yang selanjutnya algoritma terbaik dapat dipilih untuk penentuan prediksi evaluasi nasabah yang melakukan pengajuan pinjaman.

1.3.2 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Manfaat Teoritis:
 - a. Meningkatkan akurasi prediksi dengan metode klasifikasi *data mining* untuk memprediksi pemberian kredit kepada nasabah.
 - b. Menghasilkan rekomendasi algoritma yang akurasinya lebih baik dalam memprediksi pemberian kredit dibandingkan metode prediksi sebelumnya.
2. Manfaat Praktis
 - a. Perusahaan dapat memprediksi dalam hal pemberian kredit kepada nasabahnya dengan lebih baik dan meningkatkan keakurasian aktual dalam pengambilan keputusan pemberian kredit.
 - b. Dengan meningkatkan keakurasian *forecast* dengan aktual pemberian kredit maka dapat mempermudah bagian *Credit Sales Officer* dan *Agen Sales Management* dalam memberikan keputusan dalam hal pengajuan kredit dari nasabah sehingga dapat meminimalisir kredit yang bermasalah.

1.4 Tata Urut Penulisan

Dalam penelitian tesis ini, sistematika penulisan yang disusun adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi tentang penjelasan latar belakang, masalah penelitian, tujuan dan manfaat penelitian, tata urut penulisan naskah dan daftar pengertian.

BAB II LANDASAN TEORI DAN KONSEP KERANGKA

Bab ini berisi penjabaran tentang tinjauan pustaka, tinjauan studi, tinjauan objek penelitian, kerangka konsep dalam penelitian.

BAB III METODOLOGI DAN DESAIN PENELITIAN

Bab ini berisi penjelasan tentang metode penelitian, pemilihan sampel, pengumpulan data, instrumentasi, teknik analisis, perancangan, pengujian data, langkah-langkah penelitian, dan jadwal penelitian.

BAB IV PENUTUP

Bab ini berisi tentang simpulan dan saran penelitian yang sedang dilakukan.

1.5 Daftar Pengertian

Dalam butir 1.5 ini akan dijelaskan mengenai definisi atau pengertian kata-kata/istilah yang sering digunakan untuk menjelaskan teknologi yang digunakan, yaitu:

<i>Data Mining</i>	Perangkat lunak yang digunakan untuk menemukan pola tersembunyi, tren, maupun aturan-aturan yang terdapat dalam basis data berukuran besar dan menghasilkan aturan-aturan yang digunakan untuk memperkirakan perilaku di masa mendatang[4].
Prediksi	Sama dengan ramalan atau perkiraan dapat juga dikatakan hasil dari kegiatan memprediksi atau meramal atau

	memperkirakan.
<i>Data training</i>	Kumpulan data yang akan digunakan untuk pembentukan pola algoritma.
<i>Data testing</i>	data yang digunakan untuk menguji pola algoritma yang telah terbentuk.
Akurasi	Tingkat kedekatan pengukuran kuantitas terhadap nilai yang sebenarnya.
Variabel	Suatu besaran yang nilainya selalu berubah-ubah (tidak tetap) yang digunakan untuk objek suatu penelitian atau yang menjadi fokus di dalam penelitian.
Nasabah	Pihak yang menggunakan jasa finance, termasuk pihak yang tidak memiliki rekening namun memanfaatkan jasa finance untuk melakukan transaksi keuangan.
<i>Naïve Bayes</i>	Sebuah Metode untuk pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes.
<i>Decision Tree</i>	Sebuah struktur pohon, dimana setiap node pohon merepresentasikan atribut yang telah di uji.

BAB II

LANDASAN TEORI DAN KERANGKA KONSEP

2.1 Tinjauan Pustaka

2.1.1 Pengertian Kredit

Istilah kredit berasal dari bahasa Latin *credo*, yang berarti ‘saya percaya’, yang merupakan kombinasi dari bahasa Sansekerta *cred* yang artinya ‘kepercayaan’ dan bahasa Latin *do* yang artinya ‘saya tempatkan’. Kredit yang diberikan oleh finance atas dasar kepercayaan sehingga pemberian kredit merupakan pemberian kepercayaan finance kepada nasabahnya[5]. Menurut[6] undang-undang no.10 tahun 1998 tentang perubahan undang-undang no.7 tahun 1992 undang-undang tentang pembiayaan kendaraan, dalam bab 1 ketentuan umum pasal 1, yang dimaksud dengan kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam-meminjam antara finance dan pihak lain yang mewajibkan pihak peminjam untuk melunasi hutangnya setelah jangka waktu tertentu dengan pemberian bunga.

Peranan finance sebagai lembaga keuangan berperan besar dalam pemberian kredit kepada masyarakat tepatnya kepada nasabahnya, dimana pemberian kredit merupakan kegiatan utamanya. Sebagai sebuah lembaga keuangan, aset terbesar yang dimiliki oleh finance adalah aset financial dan aset utama dari finance adalah kredit yang disalurkan kepada debitur dalam hal ini adalah nasabahnya.

Ada beberapa kriteria dalam penilaian kredit kepada nasabah yang harus dilakukan atau prinsip utama yang berkaitan dengan kondisi calon nasabah. Prinsip itu dikenal dengan istilah 5C [3] yaitu :

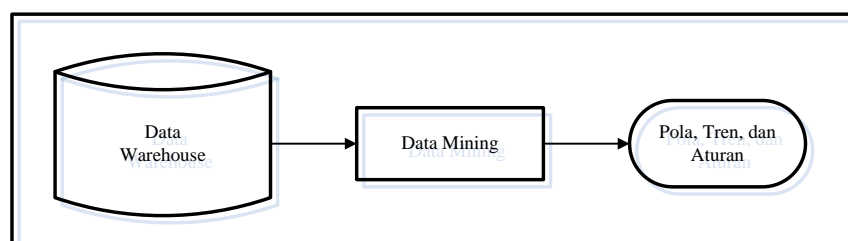
1. *Character*

Penilaian terhadap karakter atau kepribadian calon peminjam untuk memperkirakan kemungkinan bahwa peminjam dapat memenuhi kewajibannya.

2. *Capacity*
Kemampuan peminjam (*debitur*) untuk melakukan pembayaran. Kemampuan ini di ukur dari catatan prestasi peminjam di masa lalu.
3. *Capital*
Penilaian terhadap kemampuan modal yang dimiliki oleh calon peminjam, di ukur dengan posisi usaha atau jumlah penghasilan.
4. *Collateral*
Jaminan yang dimiliki calon peminjam. Penilaian ini untuk lebih meyakinkan bahwa jika suatu resiko kegagalan pembayaran terjadi, maka jaminan dapat dipakai sebagai pengganti kewajiban.
5. *Conditions*
Pihak *finance* akan melihat jenis usaha yang dilakukan oleh calon peminjam, hal ini dilakukan karena kondisi eksternal memiliki pengaruh yang kuat atau cukup besar dalam proses berjalannya usaha dari calon peminjam dalam jangka panjang.

2.1.2 *Data Mining*

Data mining atau penambangan data adalah perangkat lunak yang digunakan untuk menemukan pola tersembunyi, tren, maupun aturan-aturan yang terdapat dalam basis berukuran besar dan menghasilkan aturan-aturan yang digunakan untuk memperkirakan perilaku di masa mendatang[4].



Gambar 2.1 Prinsip *Data Mining*[7]

Data mining terdiri dari *extract*, *transform*, dan memuat data transaksi ke sistem *data warehouse*, disimpan dan mengelola data dalam sistem database

multidimensi, menyediakan akses data untuk analisis bisnis dan profesional teknologi informasi, menganalisis data oleh aplikasi perangkat lunak, menyajikan data dengan format yang berguna, seperti grafik atau tabel *Data mining* sering dikatakan berurusan dengan "penemuan pengetahuan" dalam basis data. Hal yang menarik, *data mining* menjadi perangkat yang membantu para pemakai untuk menemukan jawaban atas pertanyaan-pertanyaan yang tidak pernah mereka pikirkan sebelumnya.

Istilah[4] *data mining* memiliki hakikat sebagai disiplin ilmu yang tujuan utamanya adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki. Ada beberapa peran utama dalam *data mining* antara lain:

1. *Description*

Cara yang digunakan untuk menggambarkan sekumpulan data secara ringkas. Banyak cara yang digunakan dalam memberikan gambaran secara ringkas bagi sekumpulan data yang besar jumlahnya dan banyak macamnya, yaitu Deskripsi Grafis, Deskripsi Lokasi dan Deskripsi keragaman.

2. *Estimation*

Algoritma estimasi yang biasa digunakan adalah: *Linear Regression*, *Neural Network*, *Support Vector Machine*. Algoritma estimasi mirip dengan algoritma klasifikasi, tapi variabel target adalah berupa bilangan numerik dan bukan kategorikal (nominal). Model dibangun dari data dengan *record* yang lengkap, yang menyediakan nilai dari variabel sebagai prediktor, kemudian estimasi nilai dari variabel target ditentukan berdasarkan nilai dari variabel prediktor. Penentuan kebijakan atau suatu nilai pada proses yang akan dilakukan. Estimasi dapat dilakukan dari data-data lama yang akan diolah.

3. *Prediction*

Algoritma prediksi sama dengan algoritma estimasi dimana label/target/class bertipe numerik, bedanya adalah data yang digunakan merupakan data rentetan waktu (*data time series*). Sifat prediksi bisa

menghasilkan *class* berdasarkan berbagai atribut yang kita sediakan. Penentuan hasil dari proses yang sedang berlangsung. Data-data yang digunakan untuk prediksi berasal dari data yang ada saat proses sedang berlangsung. Istilah prediksi kadang digunakan juga untuk klasifikasi, tidak hanya untuk prediksi *time series*, karena sifatnya yang bisa menghasilkan *class* berdasarkan berbagai atribut yang kita sediakan.

4. *Classification*

Algoritma yang menggunakan data dengan *target/class/label* berupa nilai kategorikal (nominal). Pengelompokan data-data yang ada menjadi dalam kelompok yang sudah ditentukan nama kelompoknya. Metode yang cocok untuk klasifikasi, yakni: *Naïve Bayes*, *K-Nearest Neighbor*, *C4.5*, *ID3*, *CART*, *Linear Discriminant Analysis*, dan yang lainnya. Contoh, apabila *target/class/label* adalah pendapatan, maka bisa digunakan nilai nominal (kategorikal) pendapatan besar, menengah, kecil.

5. *Clustering*

Klastering adalah pengelompokan data, hasil observasi dan kasus ke dalam *class* yang mirip. Suatu klaster (*cluster*) adalah koleksi data yang mirip antara satu dengan yang lain, dan memiliki perbedaan bila dibandingkan dengan data dari klaster lain. Metode yang cocok untuk klastering, yakni: *K-Means*, *K-Medoids*, *Self-Organizing Map (SOM)*, *Fuzzy C-Means*, dan yang lainnya. Perbedaan utama algoritma klastering dengan klasifikasi adalah klastering tidak memiliki *target/class/label*, jadi termasuk *unsupervised learning*. Klastering sering digunakan sebagai tahap awal dalam proses *data mining*, dengan hasil klaster yang terbentuk akan menjadi input dari algoritma berikutnya yang digunakan.

6. *Association*

Algoritma *association rule* (aturan asosiasi) adalah algoritma yang menemukan atribut yang “jalan bersamaan”. Dalam dunia bisnis, sering disebut dengan *affinity analysis* atau *market basket analysis*. Algoritma *association rules* berangkat dari pola “*If antecedent, then consequent*,” bersamaan dengan pengukuran *support (coverage)* dan *confidence*

(*accuracy*) yang terasosiasi dalam aturan. Algoritma *association rule* diantaranya adalah: *Apriori algorithm*, *FP-Growth algorithm*, *GRI algorithm*.

Berikut beberapa alasan *data mining* menjadi penting saat ini:

1. Data yang tersedia sangat besar:

Selama *decade* terakhir harga *hardware* terutama harga *hardisk* telah menurun drastis. Bersama dengan ini, perusahaan telah menyimpan data yang sangat besar melalui banyak aplikasi. Dengan semua data ini untuk mengeksplorasi, perusahaan menginginkan dapat menemukan pola dan informasi yang tersembunyi sehingga dapat membantu mengarahkan strategi bisnis perusahaan menjadi lebih baik.

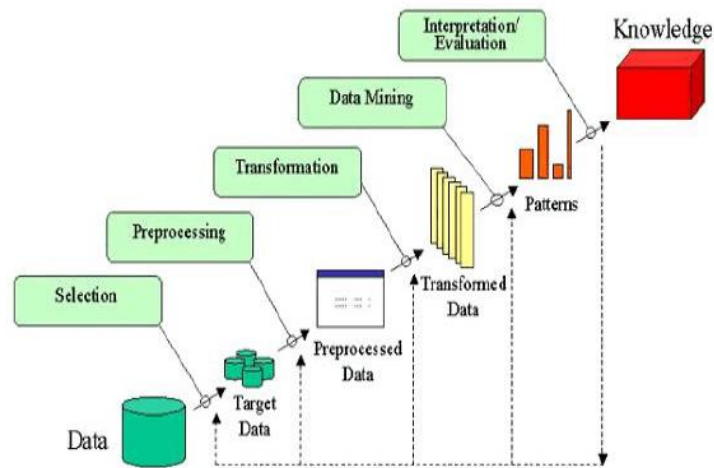
2. Meningkatkan kompetisi:

Kompetisi yang tinggi sebagai dampak dari modernisasi pasar dan distribusi seperti *internet* dan telekomunikasi. Perusahaan-perusahaan di seluruh dunia menghadapi persaingan dan kunci untuk keberhasilan bisnis adalah kemampuan untuk mempertahankan pelanggan dan memperoleh pelanggan baru. *Data mining* adalah teknologi yang memungkinkan perusahaan untuk menganalisis faktor-faktor yang mempengaruhi masalah ini.

3. Teknologi yang telah siap:

Teknologi *data mining* yang sebelumnya hanya ada dalam wilayah akademik, tapi sekarang banyak dari teknologi tersebut telah matang dan siap diterapkan dalam perusahaan. Algoritma yang lebih akurat, efektif dan dapat menangani data yang semakin banyak dan rumit. Selain itu pemrograman aplikasi antar muka *data mining* telah distandarisasi yang akan memungkinkan para pengembang untuk membangun aplikasi *data mining* akan lebih baik.

Sebagai suatu rangkaian proses, *data mining* dapat dibagi menjadi beberapa tahap proses yang diilustrasikan pada gambar II-2.



Gambar 2.2 Proses *Data Mining*[8]

Pada gambar 2.2 digambarkan proses *data mining* adalah data dipilih, dibersihkan, dan dilakukan *preprocessing* mengikuti pedoman dan *knowledge* dari ahli domain yang menangkap dan mengintegrasikan *data internal* dan *eksternal* ke dalam tinjauan organisasi secara menyeluruh. Penggunaan algoritma *data mining* dilakukan pada langkah ini untuk menggali data yang terintegrasi untuk memudahkan identifikasi informasi bernilai. Keluaran dari *data mining* dievaluasi untuk melihat apakah *knowledge domain* ditemukan dalam bentuk rule yang telah diekstraksi dari jaringan.

Adapun proses dari KDD(*Knowledge Discovery in Database*) sebagai berikut [8]:

1. *Data Selelection*

Pemilihan data dari sekumpulan data operasional perlu dilakukan tahap penggalan informasi sebelum KDD di mulai.

2. *Preporcessing*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* dengan tujuan untuk membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses *enrichment*,

yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

3. *Transformation*

Proses *coding* pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam *database*

4. *Data Mining*

Proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu

5. *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 KDD yang disebut dengan *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya atau tidak.

Evaluasi adalah kunci ketika membuat aplikasi berbasis *data mining*. Ada berbagai macam cara melakukan evaluasi. Evaluasi ternyata tidak sederhana yang kita bayangkan. Jika kita memiliki data yang kita gunakan dalam proses pelatihan, maka tidak serta merta kita menjadikan data tersebut sebagai indikator keberhasilan aplikasi yang kita buat. Oleh karena itu, kita membutuhkan metode tertentu guna memprediksi performa berdasarkan eksperimen untuk berbagai macam data selain *data training* tersebut.

Pada umumnya data yang cukup banyak dapat dimanfaatkan untuk pengujian. Hanya saja masalah yang kerap dijumpai adalah datanya. Oleh karena itu, kita harus memastikan data yang akan kita gunakan baik untuk pelatihan maupun untuk pengujian merupakan data yang berkualitas.

2.1.3 Klasifikasi

Klasifikasi dalam *data mining* merupakan metode pembelajaran data untuk memprediksi nilai dari sekelompok atribut. Algoritma klasifikasi akan menghasilkan sekumpulan aturan yang disebut *rule* yang akan digunakan sebagai *indicator* untuk dapat memprediksi kelas dari data yang ingin diprediksi[1]. Klasifikasi digunakan dalam banyak sekali bidang, dan secara teori algoritma klasifikasi sama seperti otak manusia.

Otak manusia mampu mengolah data yang sudah ada sebagai pengalaman dalam bertindak. Dalam *data mining* ada beberapa algoritma klasifikasi yang banyak digunakan dalam masyarakat atau dalam penelitian secara luas diantaranya adalah *Decision / classification trees*, *Bayesian classifier / Naïve Bayes classifiers*, *Neural networks*, Analisa Statistik, Algoritma Genetika, *Rough sets*, *K-Nearest Neighbor*, metode *Rule Based*, *Memory Based reasoning*, dan *Support Vector Machine (SVM)*.

Tujuan dari algoritma klasifikasi adalah untuk menemukan relasi antara beberapa variable yang tergolong dalam kelas yang sama. Relasi tersebut akan digambarkan dengan aturan-aturan agar dapat memprediksi kelas dari data yang *attribute* nya sudah diketahui.

Penilaian algoritma klasifikasi biasanya dilihat dari akurasi model. Akurasi model merupakan ketepatan model dalam memprediksi kelas data. Selain akurasi kecepatan pembentukan model, kemampuan algoritma dalam mengatasi data yang tidak relevan atau bahkan data yang tidak lengkap, serta kemampuan algoritma ketika diterapkan pada data jumlah besar maupun kecil.

2.1.4 K-Fold Cross Validation

Dalam penelitian ini,[4] metode yang digunakan untuk menguji pola klasifikasi adalah dengan metode *k-fold cross validation*. Dalam *k-fold cross validation* data dibagi menjadi *k* bagian, D_1, D_2, \dots, D_k , dan masing-masing D memiliki jumlah data yang sama. Kemudian akan dilakukan proses perulangan sebanyak *k*, dimana dalam setiap perulangan ke-*i*, D_1 akan dijadikan *data testing*, dan sisanya akan digunakan sebagai *data training*.

Pengujian dengan $k=5$ atau $k=10$ dapat digunakan untuk memperkirakan tingkat kesalahan yang terjadi, sebab data training pada setiap *fold* cukup berbeda dengan *data training* yang asli. Secara keseluruhan, 5 atau 10 *fold cross validation* sama-sama direkomendasikan dan disepakati bersama. Menghitung nilai akurasi dapat dilakukan dengan menggunakan persamaan:

$$\text{Akurasi} = \frac{\text{Jumlah klasifikasi benar}}{\text{Jumlah data uji}} \times 100 \% \quad (2.1)$$

2.1.5 Algoritma Naïve Bayes

Algoritma *Naïve Bayes* merupakan metode yang dapat digunakan untuk mengklasifikasikan sekumpulan data. Algoritma ini memanfaatkan metode probabilitas dan Statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya [5][4].

Naïve Bayes merupakan *machine learning* yang menggunakan perhitungan probabilitas yang menggunakan konsep pendekatan *Bayesian*. Kata *Naïve*, yang terkesan merendahkan, berasal dari asumsi *independensi* pengaruh nilai suatu atribut dari probabilitas pada kelas yang diberikan terhadap nilai atribut lainnya. Penggunaan teorema Bayes pada algoritma *Naïve Bayes* yaitu dengan mengkombinasikan prior probability dan probabilitas bersyarat dalam sebuah rumus yang bisa digunakan untuk menghitung probabilitas tiap klasifikasi yang mungkin.

Rumus *Naïve Bayes* nya adalah :

$$P(H|X) = \frac{P(H)P(X|H)}{P(X)} \quad (2.2)$$

Keterangan :

X = data dengan kelas yang belum diketahui

H = hipotesis data X, merupakan suatu kelas yang spesifik

$P(H|X)$ = probabilitas hipotesis H berdasar kondisi X (posteriori probability)

$P(H)$ = probabilitas hipotesis H (posteriori probability)

$P(X|H)$ = probabilitas X berdasar kondisi H

$P(X)$ = probabilitas dari X

atau

$$\text{Posterior Probability} = \frac{\text{Prior Probability} \times \text{likelihood}}{\text{evidance}} \quad (2.3)$$

2.1.6 Algoritma *Decision Tree*

Decision Tree adalah sebuah struktur pohon, dimana setiap node pohon merepresentasikan atribut yang telah di uji, setiap cabang merupakan suatu pembagian hasil uji, dan node daun (*leaf*) merepresentasikan kelompok kelas tertentu. *Level node* teratas dari sebuah *Decision Tree* adalah node akar (*root*) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada kelas tertentu[4].

Pohon keputusan sendiri merupakan pendekatan “*divide and conquer*” dalam mempelajari masalah dari sekumpulan data independen yang digambarkan dalam bagan pohon. Pohon keputusan juga merupakan sekumpulan pertanyaan yang tersusun secara sistematis, dimana setiap pertanyaan yang ada menentukan percabangan berdasarkan nilai atribut dan berhenti pada daun dari pohon yang merupakan prediksi dari kelas variabel.

Semua metode yang menghasilkan pohon keputusan bekerja mulai dari akar yang paling atas, dengan menguji data yang ada dan dipecah kedalam setiap cabang dari pohon yang lebih kecil. Keputusan dipilih berdasarkan kesesuaian antara data yang diuji dengan aturan dari cabang pohon keputusan yang ada hingga mencapai daun sehingga nilai kelas yang terdapat pada daun diberikan pada tupel data yang diperiksa tersebut.

Ada beberapa proses yang harus diperhatikan dalam pembentukan struktur pohon, yaitu :

- a) Pilih *root* berdasarkan gain ratio terbesar
- b) Pilih internal *root* / cabang *root* berdasar *gain ratio* terbesar setelah menghapus atribut yang telah terpilih sebagai *root*.
- c) Ulangi sampai semua atribut terhitung nilai *gain rationya*.

Dalam pembuatan pohon keputusan, setiap algoritma menerapkan ukuran pemilihan atribut yang berbeda-beda. Ukuran pemilihan atribut merupakan ukuran yang digunakan dalam menentukan kriteria yang terbaik untuk mengelompokkan *tuple*. Ukuran pemilihan atribut ini juga disebut sebagai *splitting rules* karena menentukan bagaimana data akan dipisahkan kesetiap cabang. C4.5 yang merupakan pengembangan dari ID3 menggunakan *Information gain* untuk ukuran pemilihan atribut.

Information gain diciptakan oleh Claude Shannon dengan mempelajari nilai informasi dari data, dan menggunakan nilai tersebut sebagai acuan dalam menentukan atribut yang akan digunakan dalam menyusun pohon keputusan. Atribut yang dipilih akan menghasilkan partisi dengan data yang lebih seragam, dan dapat menghasilkan pohon keputusan yang sesederhana mungkin dengan perulangan yang sedikit. Berikut persamaan data dalam *tuple* D.

$$Info(D) = \sum_{i=1}^n -p_i \log_2(p_i) \quad (2.4)$$

Dimana p_i merupakan probabilitas tuple dalam D yang menjadi kelas C_i dengan asumsi $|C_{i,D}|/|D|$. $Info(D)$ atau disebut juga *entropy* dari D merupakan rata rata informasi yang diperlukan untuk identifikasi tuple dalam D.

Jika Nilai A adalah nilai diskrit maka data D akan dipisahkan sejumlah nilai data A sehingga nilai setiap cabang akan murni dan sejenis. Setelah percabangan pertama, jumlah percabangan yang mungkin terjadi diukur dengan persamaan.

$$Info_A(D) = \sum_{j=1}^v \frac{|D_j|}{|D|} \times Info(D_j) \quad (2.5)$$

Dimana $\frac{|D_j|}{|D|}$ merupakan bobot dari partisi j. $Info_A(D)$ merupakan informasi yang diperlukan untuk mengklasifikasikan tuple dari D pada partisi A. Semakin kecil hasil persamaan ini, semakin baik pula partisi yang dihasilkan. Nilai dari sebuah atribut menentukan penting tidaknya atribut tersebut dalam penyusunan pohon keputusan. Semakin besar nilai atribut berarti semakin baik pula klasifikasi yang akan terbentuk jika cabang dipecah menurut atribut tersebut, untuk melihat nilai dari atribut A digunakan persamaan.

$$Gain(A) = Info(D) - Info_A(D) \quad (2.6)$$

2.1.7 Pemilihan Variabel

Pemilihan variabel yang juga disebut sebagai pemilihan atribut, digunakan pada *dataset* untuk menemukan pola yang penting dalam *data mining*. Pemilihan variabel digunakan untuk pengurangan dimensi pada *dataset*. Pemilihan variabel digunakan untuk melakukan eliminasi variabel yang tidak *relevan* dan *redundan*, yang dapat menyebabkan kebingungan dalam penggunaan variable.

Pemilihan variabel dapat mengurangi dimensi data, hal ini memungkinkan lebih efektif dalam operasi agar lebih cepat dari beberapa algoritma data mining. Dengan adanya pemilihan variabel membuat algoritma data mining lebih cepat dan lebih efektif.

Penggunaan pemilihan variabel pada *dataset* yang menggunakan variabel bebas dapat meningkatkan performa model. Pemilihan variabel juga merupakan proses yang cukup memakan biaya, dan juga berten.tangan dalam asumsi awal, bahwa semua informasi diperlukan untuk mencapai akurasi yang maksimal.

Metode yang dapat digunakan untuk pemilihan variabel antara lain *Backward Elimination*, *Forward Selection*, *Genetic Algorithm*, dan yang lainnya. Metode-metode tersebut digunakan dalam penelitian *data mining* agar dapat menghasilkan variabel yang relevan dalam penelitian.

Pemilihan variabel dengan filter model ini lebih murah dalam komputasi karena tidak melibatkan induksi algoritma dalam prosesnya.

2.1.8 Pengujian Akurasi dan Validasi Metode Klasifikasi Data Mining

Untuk menguji model, pada penelitian ini, digunakan metode *Confusion Matrix*, dan kurva *ROC (Receiver Operating Characteristic)*.

1. *Confusion matrix*

Metode ini menggunakan tabel matriks seperti pada Tabel 2.1, jika data set hanya terdiri dari dua kelas, kelas yang satu dianggap sebagai positif dan yang lainnya *negative*

Tabel 2.1 Model *Confusion Matrix*

<i>Correct classification</i>	<i>Classified as</i>	
	+	-
+	<i>True Positives</i>	<i>False Negatives</i>
-	<i>False Positives</i>	<i>True Negatives</i>

True positives adalah jumlah *record* positif yang diklasifikasikan sebagai positif, *false positives* adalah jumlah *record* negatif yang diklasifikasikan sebagai positif, *false negatives* adalah jumlah *record* positif yang diklasifikasikan sebagai negatif, *true negatives* adalah jumlah *record* negatif yang diklasifikasikan sebagai negative, kemudian masukkan data uji. Setelah data uji dimasukkan ke dalam *confusion matrix*, hitung nilai-nilai yang telah dimasukkan tersebut untuk dihitung jumlah *sensitivity (recall)*, *specificity*, *precision* dan *accuracy*. *Sensitivity* digunakan untuk membandingkan jumlah TP terhadap jumlah *record* yang positif sedangkan *specificity* adalah perbandingan jumlah TN terhadap jumlah *record* yang negatif. Untuk menghitung digunakan persamaan di bawah ini:

$$\text{Sensitivity} = \frac{TP}{P} \quad (2.7)$$

$$\text{Specificity} = \frac{TN}{P} \quad (2.8)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2.9)$$

$$\text{Accuracy} = \text{sensitivity} \frac{P}{P+N} + \text{specificity} \frac{N}{P+N}$$

Keterangan:

TP = jumlah *true positives*

TN = jumlah *true negatives*

P = jumlah *record* positif

N = jumlah *tupel* negatif

FP = jumlah *false positives*

2. Kurva ROC (*Receiver Operating Characteristic*)

Kurva ROC atau *Receiver Operating Characteristic* menunjukkan akurasi dan membandingkan klasifikasi secara visual. ROC mengekspresikan *confusion matrix*. ROC adalah grafik dua dimensi dengan *false positives* sebagai garis horisontal dan *true positives* sebagai garis vertikal[9][10].

2.1.9 Pengertian Nasabah

Nasabah adalah pihak yang menggunakan jasa bank atau lembaga keuangan bukan bank. Penghimpunan dana dan pemberian kredit merupakan pelayanan jasa perbankan yang utama dari semua kegiatan lembaga keuangan. Berdasarkan Pasal 1 angka (16) UU Perbankan diintroduksikan rumusan nasabah yaitu nasabah adalah pihak yang menggunakan jasa bank.

Rumusan tersebut kemudian diperinci pada butir berikutnya, yaitu sebagai berikut:

- a. Nasabah Penyimpan adalah nasabah yang menempatkan dananya di bank dalam bentuk simpanan berdasarkan perjanjian bank dengan nasabah yang bersangkutan.
- b. Nasabah Debitur adalah nasabah yang memperoleh fasilitas kredit atau pembiayaan berdasarkan prinsip syariah atau yang dipersamakan dengan itu berdasarkan perjanjian bank dengan nasabah yang bersangkutan.

2.2 Tinjauan Studi

Berikut adalah ringkasan dari beberapa penelitian sebelumnya yang terkait dengan klasifikasi *data mining*.

Tabel 2.2 Ringkasan Tinjauan Studi

No	Judul	Metode	Tujuan	Variabel/atribut	Hasil
1.	“Algoritma Naïve Bayes, Decision Tree, dan SVM untuk Klasifikasi Persetujuan Pembiayaan Nasabah Koperasi Syariah”	Pengujian menggunakan metode Algoritma Decision tree, Naïve Bayes dan SVM	menerapkan data mining untuk dapat membantu melakukan klasifikasi persetujuan pembiayaan nasabah	Status, tanggungan, usia, pendidikan, pekerjaan, pendapatan, status rumah, jaminan, jumlah pinjaman, lama pinjaman	Algoritma yang telah dilakukan diuji menggunakan tools Rapid Miner. Diperoleh hasil akurasi sebesar tertinggi 94% pada algoritma SVM.
2.	“Perbandingan Metode K-Nearest Neighbor, Naïve Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit”[1]	Algoritma K-NN, Naïve Bayes dan Decision Tree	Data ini mengacu dan dapat digunakan dalam memprediksi kelayakan pemberian kredit untuk calon debitur	Jenis kelamin, umur, Pendidikan terakhir, status pernikahan, pekerjaan, penghasilan, jumlah pinjaman, jenis pinjaman, jangka waktu	Decision Tree (J-48) dengan akurasi sebesar 92,21%, algoritma K-Nearest Neighbor memiliki tingkat akurasi sebesar 81,82% dan Naïve Bayes memiliki tingkat akurasi sebesar 81,83%.

Tabel 2.2 (lanjutan)

No	Judul	Metode	Tujuan	Variabel/atribut	Hasil
3.	“Aplikasi metode K-Nearest Neighbor dan analisis Diskriminan untuk analisis resiko kredit pada koperasi simpan pinjam di kopinkra sumber rejeki”[12]	Algoritma K-NN dan Analisis Diskriminan Metode pengujian : <i>K-Folds Cross Validation</i>	Dapat memprediksi potensi kredit macet dan menentukan kelancaran simpan pinjam di kopinkra sumber ejeti	penghasilan, pinjaman, jangka waktu dan jaminan	Ketepatan prediksi resiko dengan metode KNN 84,33% pada nilai K adalah 7 dan Metode Diskriminan sebesar 76,5%.
4.	“analisis klasifikasi pada nasabah kredit koperasi menggunakan decision tree c4.5 dan naïve bayes”[13]	Metode pengujian menggunakan decision tree c4.5 dan naïve bayes	untuk membentuk model decision tree C4.5 dan naïve bayes untuk klasifikasi nasabah kredit berdasarkan nilai kolektibilitasnya	Status nasabah, Umur, Desa, Kecamatan, Status Marital, Nilai Pinjaman, Jumlah angsuran	Decision tree C4.5 memiliki kecenderungan tingkat akurasi yang lebih tinggi yaitu sebesar 71,91%, 68,03%, dan 66,84%, sedangkan naïve bayes memiliki akurasi sebesar 67,01%, 64,66%, dan 65,82%.
5.	“Komparasi Algoritma Klasifikasi Data Mining untuk Evaluasi Pemberian Kredit”[14]	Algoritma C4.5 dan Naïve Bayes	Analisa kredit nasabah sebelum diabil keputusan pemberian kredit	Nama, Jenis Kelamin, Umur, Pinjaman, Tenor, Angsuran, Plafon Teoritis, Tunggakan Pokok, Tunggakan bunga	Tingkat akurasi yang lebih baik adalah menggunakan algoritma C4.5 yaitu nilai akurasinya 88,90% sedangkan dengan Naïve Bayes yaitu 80,00%

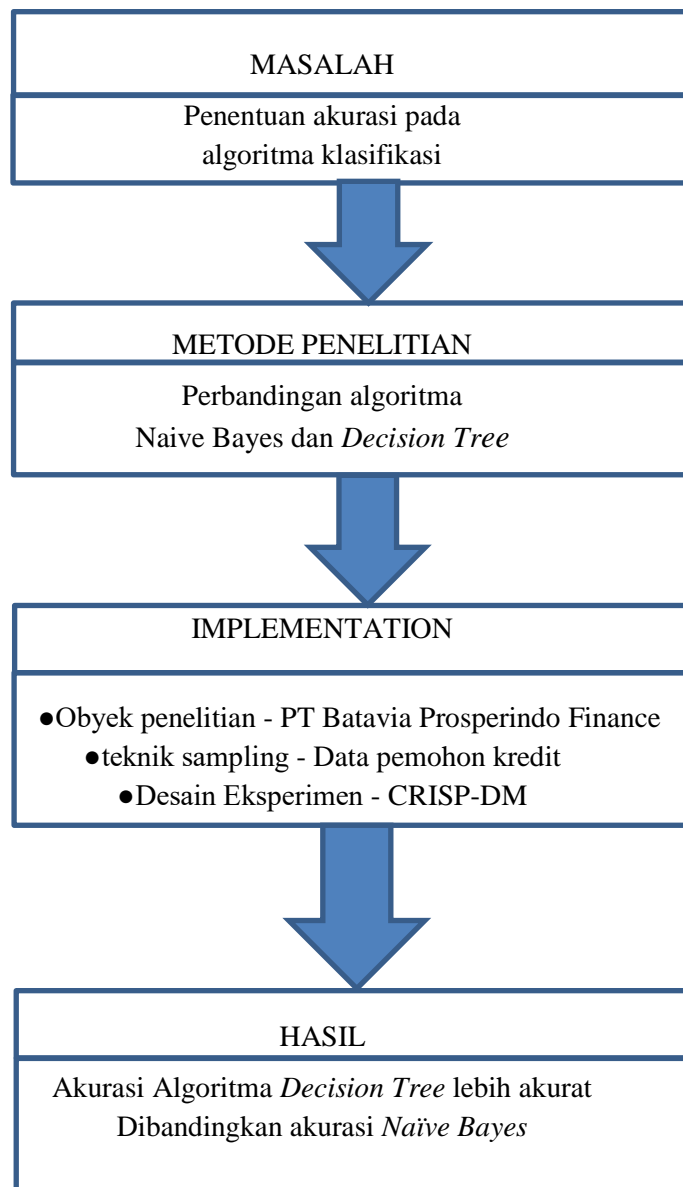
Pada tabel 2.2. Ringkasan tinjauan studi, hampir semua penelitian menggunakan metode klasifikasi *data mining*. Dalam hal ini penulis mencoba membandingkan dua metode klasifikasi *data mining* yaitu *Naïve Bayes* dan *Decision Tree* untuk melihat akurasi yang tepat dari penghitungan ke dua metode tersebut penulis menggunakan pengujian *k-folds cross validation* dengan $k=10$ agar dapat mengetahui tingkat akurasinya terhadap prediksi pemberian kredit pada finance. Nantinya hasil dari pengujian tersebut akan dipilih salah satu metode dan kemudian akan dimodifikasi modelnya sehingga diharapkan dapat menaikkan nilai pengujian baik akurasi, precision, recall dan lain-lain. Hasil pemilihan metode dan modifikasi model akan diterapkan dalam pengembangan prototype prediksi pemberian kredit untuk calon nasabah.

2.3 Tinjauan Obyek Penelitian

PT Batavia Prosperindo Finance Tbk adalah suatu perusahaan *finance* yang bergerak di bidang pembiayaan untuk kendaraan roda empat, terutama kendaraan bekas jenis Penumpang/Pribadi (Passenger) dan Niaga (Commercial) yang memiliki banyak penghargaan di dunia usaha dan bisnis perbankan, yang memiliki banyak anak cabang yang tersebar di seluruh Indonesia dan dikenal sebagai salah satu pembiayaan yang memiliki pelayanan terbaik kepada para debitur. Dengan penawaran dari suku bunga rendah dan proses pengajuan kredit tidak sulit menjadikan perusahaan ini menjadi tolak ukur bagi seluruh finance di Indonesia. Adapun visi dan misi dari perusahaan ini adalah pelopor dalam memberikan solusi finansial yang inovatif dan menjadikan hidup lebih bernilai untuk kebutuhan kendaraan dan permodalan dalam dunia usaha serta bisnis.

2.4 Kerangka Konsep / Pola Pikir Pemecahan

Penelitian ini dilakukan dengan judul “**Perbandingan Algoritma *Naive Bayes* Dan *Decision Tree* Dalam Memprediksi Approval Pemberian Kredit Studi Kasus : PT Batavia Prosperindo Finance Tbk. Cabang Pringsewu**”. Maka dapat dibangun kerangka konsep penelitian tahapan sebagai berikut:



Gambar 2.3 Kerangka Pemikiran

Penjelasan untuk kerangka Pemikiran pada gambar 2.3 yaitu:

a. Masalah / *Problems*

Problems atau permasalahan yang ada pada penelitian ini adalah penentuan akurasi pada algoritma Klasifikasi.

b. Metode Penelitian / *Approach*

Approach atau metode penelitian yang dipakai pada kasus ini adalah dengan perbandingan *Naive Bayes* dan *Decision Tree*.

c. *Implementation*

Penelitian ini menggunakan data Nasabah dalam pengajuan kredit pada PT Batavia Prosperindo Finance , dan urutan desain eksperimen menggunakan CRISP-DM (*Cross Standard Industry Process for Data Mining*)

d. Hasil / *Result*

Result atau hasil penelitian yang didapat adalah akurasi dari Algoritma *Decision Tree* lebih tepat akurasinya dibanding dengan metode *Naïve Bayes* untuk penentuan Prediksi kelayakan pemberian Kredit

2.5 Hipotesis

Dari ke dua klasifikasi data mining yang diuji yaitu *Naïve Bayes* dan *Decision Tree* merupakan metode yang dapat digunakan dalam menentukan prediksi pemberian kredit kepada nasabah.

1. Dengan menggunakan metode klasifikasi *Naïve Bayes* dan *Decision tree* diduga dapat digunakan untuk prediksi penentuan nasabah dalam melakukan pengajuan kembali pinjaman dengan evaluasi data nasabah sebelumnya.
2. Penggunaan kedua algoritma yaitu *Naïve Bayes* dan *Decision tree* salah satunya akan menghasilkan nilai yang lebih tinggi, nilai yang lebih tinggi yang nantinya akan dijadikan sebagai acuan dalam menentukan evaluasi nasabah melakukan pinjaman kembali.
3. Nilai akurasi dari *Decision Tree* diduga memiliki nilai yang lebih tinggi dibandingkan *Naïve Bayes* .

BAB III

METODOLOGI DAN DESAIN PENELITIAN

3.1 Metode Penelitian

Untuk mengetahui bagaimana seharusnya penelitian dilakukan yaitu dengan menggunakan metode penelitian. Metode penelitian adalah suatu cara atau prosedur untuk mencari, memperoleh, mengumpulkan dan mencatat data yang digunakan dalam menyusun laporan penelitian[1].

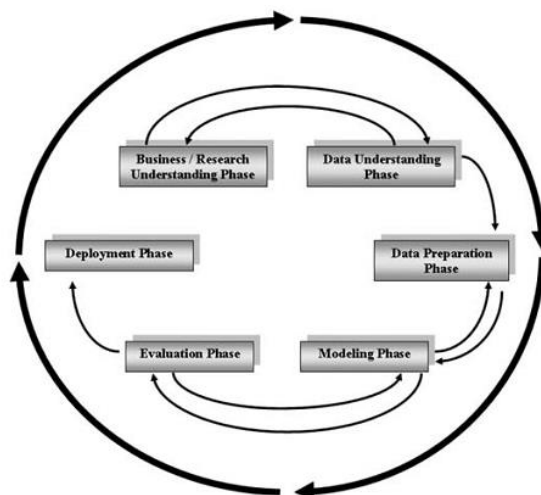
Penulis menggunakan jenis penelitian yang dilihat dari tujuannya adalah penelitian eksperimen. Penelitian eksperimen adalah penelitian yang menguji hipotesis mengenai hubungan sebab akibat[14].

Pendekatan pada penelitian ini adalah eksperimen dengan cara menguji metode melalui suatu prototype system yang keefektifannya akan diuji dengan menggunakan kelengkapan (*recall*) dan ketepatan (*precision*).

Berdasarkan data dan informasi yang diolah pada penelitian ini, penelitian ini tergolong didalam jenis penelitian kualitatif. Penelitian kualitatif adalah penelitian yang hipotesisnya tidak dapat diuji dengan teknik-teknik statistik[15] .

Untuk melakukan pengujian metode menggunakan K-Fold Crossvalidation (k=10) yang akan menampilkan nilai *akurasi*, *precision*, *recall*, *compareROC* pada masing-masing metode yang dibandingkan. Hasil dari pengujian tersebut akan menunjukkan algoritma yang tepat hal ini dilihat dari nilai akurasi untuk diterapkan pada prototype rancangan aplikasi prediksi pemberian kredit sebagai dasar rekomendasi dalam pengambilan keputusan pengajuan kredit dari nasabah.

Data mining adalah sebuah proses sehingga dalam melakukan prosesnya harus sesuai prosedur yaitu proses CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yaitu sebagai keseluruhan proses, *reprocessing* data, pembentukan model, model evaluasi, dan akhirnya penyebaran model[14].



Gambar 3.1 Proses CRISP-DM[16]

Dalam gambar 3.1 dijelaskan bahwa metode penelitian eksperimen, digunakan model proses CRISP-DM (*Cross Standard Industry Process for Data Mining*) yang terdiri dari 6 tahapan[16]:

1. *Business Understanding*
2. *Data Understanding*
3. *Data Preparation*
4. *Modelling*
5. Evaluasi
6. *Deployment*

Penelitian ini menggunakan metode studi pustaka sebagai tahap awal dengan mempelajari landasan teori mengenai *data mining* menggunakan algoritma klasifikasi yaitu *Naïve Bayes* dan *Decision Tree* pada beberapa literatur dan referensi lainnya. Referensi mencakup data-data dari jurnal, *internet*, *e-book*, dan dokumen lainnya yang berkaitan dengan penelitian ini.

Dan akan menggunakan *dataset* yang dijadikan *data training* maupun *data testing* sebanyak 300. Dari 300 data tersebut akan dibagi menjadi dua bagian, 240 digunakan untuk *data training* dan 60 digunakan untuk *data testing*. Dengan menggunakan *dataset* sebanyak 300 dan atribut yang terdiri dari 13 atribut, akan dilakukan beberapa penyeleksian untuk menghasilkan data yang dibutuhkan, tahapannya yaitu:

1. *Data Cleaning*

Untuk membersihkan nilai yang kosong, sebagai contoh *atribut tujuan_pinjaman*.

2. *Data Integration*.

Berfungsi menyatukan tempat penyimpanan yang berbeda kedalam satu data. Dalam kasus ini hanya ada satu tempat penyimpanan data yaitu *alamat_nasabah*.

3. *Data Reduction*

Jumlah atribut yang digunakan mungkin terlalu banyak, dari 15 atribut yang digunakan hanya 13 atribut yang diperlukan, dan atribut yang tidak diperlukan akan dihapus.

Data training akan digunakan untuk pembentukan pola algoritma klasifikasi data mining dari metode *Naïve Bayes* dan *Decision Tree*. *Data testing* digunakan untuk menguji pola algoritma yang telah dibentuk. Proses penentuan di terima atau di tolak pengajuan kreditnya seorang calon debitur oleh pihak Batavia dengan mengacu pada 5C (*character, capacity, capital, colleteral, condition of economy*) dari calon *debitur*.

Dalam pengembangannya akan dibuat sebuah aplikasi dengan PHP. Hasil penelitian ini adalah Perbandingan metode klasifikasi *data mining* untuk analisis prediksi pemberian kredit dengan metode *Naïve Bayes* dan *Decision Tree* agar dapat meminimalkan kasus kredit macet.

Penulis menggunakan jenis penelitian yang dilihat dari tujuannya adalah penelitian eksperimen. Penelitian eksperimen adalah penelitian yang menguji hipotesis mengenai hubungan sebab akibat[1].

Pendekatan pada penelitian ini adalah eksperimen dengan cara menguji metode melalui suatu *prototype* sistem yang keefektifannya akan diuji dengan menggunakan kelengkapan akurasi(*accuration*), (*recall*), dan ketepatan (*precision*).

3.2 Metode Pemilihan Sampel

Pemilihan sampel populasi berkaitan dengan pengajuan kredit dari nasabah. Penarikan sampel merupakan proses pilihan sejumlah elemen dari

populasi. *Probability sampling* adalah teknik pengambilan sampel yang memberikan peluang yang sama bagi setiap unsur atau anggota populasi untuk dipilih menjadi sampel[1].

Metode penarikan sampel yang dilakukan dalam penelitian ini adalah *purposive* (disengaja). Populasi adalah wilayah *generalisasi* yang terdiri atas objek/subjek yang mempunyai *kuantitas* dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya [1]. Populasi kasus yang digunakan dalam penelitian ini adalah data nasabah yang pernah mengajukan pinjaman di PT Batavia Prosperindo Finance dalam kurun waktu 2018-2020 sebanyak 300 nasabah.

3.3 Tahap Tahap Penelitian

3.3.1 Metode Pengumpulan Data

Dalam penelitian ini untuk mendapatkan data yang diharapkan maka peneliti mencari, mempelajari, serta mendalami berbagai literatur baik jurnal, buku, ataupun referensi-referensi lainnya yang berhubungan dengan topik penelitian ini.

1. Metode observasi

Pengamatan langsung dilakukan di PT Batavia Prosperindo Finance dengan Marketing dan *Service Management* untuk mengumpulkan data yang berhubungan dengan prediksi pengajuan kredit calon nasabah dengan cara mengamati dan mencatat secara sistematis masalah-masalah yang diselidiki dan meneliti secara langsung terhadap objek yang akan diteliti.

2. Wawancara

Tanya jawab langsung dilakukan kepada pihak-pihak yang berkompeten atau berkepentingan dalam menentukan metode yang cocok untuk prediksi pemberian kredit. Dari hasil wawancara ini diharapkan dapat menambah kelengkapan data yang diperoleh dari hasil pengamatan.

3. Studi Pustaka

Suatu bentuk riset yang menggunakan proses pencarian data dengan cara mencari, membaca buku dan mengolah isi dari beberapa referensi buku yang dapat dijadikan tujuan dalam pencarian data. Data yang diperoleh dari studi pustaka inilah yang disebut dengan data sekunder, tujuan dari data sekunder ini adalah sebagai landasan teori untuk menganalisa pemecahan masalah di dalam tesis ini.

4. Kuesioner

Angket atau kuesioner adalah teknik pengumpulan data melalui formulir-formulir yang berisi pertanyaan-pertanyaan yang diajukan secara tertulis pada seseorang atau sekumpulan orang untuk mendapatkan jawaban atau tanggapan dan informasi yang diperlukan oleh peneliti [17]. Metode ini digunakan untuk mendapatkan berbagai informasi tambahan yang mendukung dalam penelitian. Hasil dari kuesioner akan digunakan untuk penilaian terhadap *prototype* (sistem baru yang dibuat) setelah dilakukan prediksi klasifikasi kedua metode dengan memberikan daftar pertanyaan sesuai dengan uji kuesioner *Software Quality Assurance*

3.3.2 Dataset Penelitian

Berikut ini adalah tabel yang menunjukkan dataset penelitian nasabah yang melakukan pinjaman kredit di PT Batavia Prosperindo Finance Tbk. Terdapat 387 data nasabah seperti ditunjukkan pada tabel 3.1

No.	no_Kontrak	nama	dob	gender
1	074372190071	TASRIP	04/04/1977	L
2	074372190072	EKO SANTOSO	09/09/1979	P
3	074372190074	YENI EFTIANINGSIH	16/02/1978	L
4	074372190075	RESTU NURUL FATIMAH	29/03/1975	L
5	074372190076	AGUS SURAHMAN	28/12/1975	L
6	074372190079	JAENURI	06/08/1979	L
7	074372190080	ISRIYADI	10/01/1975	L
8	074372190082	OPTO SUBAGIO	18/02/1978	L
9	074372190084	SUNGATNO	17/04/1979	P
10	074372190086	DEDE MAHENDRA *(Nasabah GP)	13/08/1979	L
11	074372190087	HENDRA RUDHIANATA	30/10/1977	L
12	074372190088	ISLAH *(Nasabah GP)	01/01/1977	L
13	074372190090	YENIAR EFENDI	18/09/1977	L
14	074372190091	DARMAWAN	14/04/1979	L
15	074372190092	NOVRIADI	18/08/1978	L
16	074372190096	RIZKI INDRA SETIAWAN *(Nasabah GP)	17/11/1977	L
17	074372190099	MUKHLIS HS *(Nasabah GP)	27/12/1975	L
18	074372190101	RINI PUJI LESTARI	23/12/1977	P
19	074372190103	TUNUT	10/03/1979	L
20	074372190104	NANANG PURWANTO *(Nasabah GP)	01/07/1978	P
21	074372190105	DEDI ARPANI	09/05/1978	L
22	074372190111	MIRHAN	14/05/1978	L
23	074372190112	NOPRIANDI	17/08/1975	L
24	074372190113	ZULYADI MUNZIRI	17/08/1976	L
25	074372190118	RATNA ARIANTI	17/08/1977	P
...
387	074372190153	MISDI *(Nasabah GP)	17/08/1997	L

Tabel 3.1 *Dataset* Penelitian

3.4 Metode-metode Pendekatan Penyelesaian Permasalahan

3.4.1 Analisis Sistem Berjalan

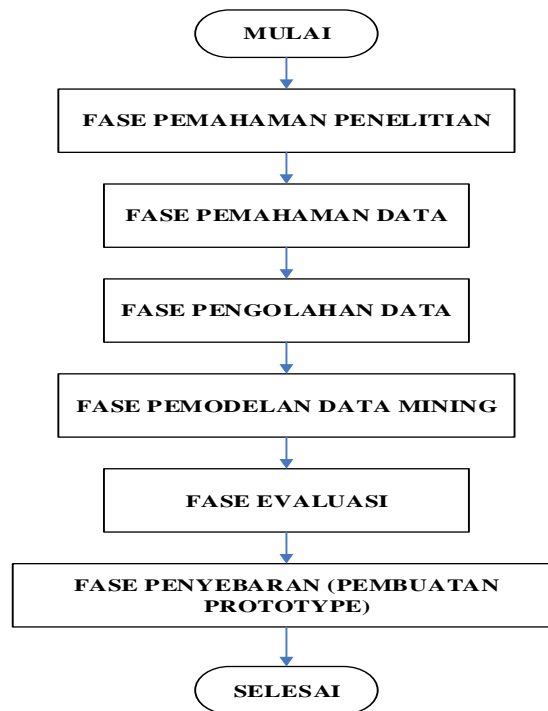
Menurut hasil wawancara yang dilakukan dengan bagian *General Marketing* dan Area *Sales Service Management* dari PT Batavia Prosperindo Finance Tbk., Cabang Pringsewu masih menggunakan cara yang manual dari data-data di Microsoft Excel.

Proses sistem berjalan kegiatan pengajuan kredit masih kurang efisien khususnya pada pengambilan keputusan seorang pimpinan dalam menentukan nasabah bisa disetujui kembali pengajuan kreditnya. Sebagian besar penentuan kriteria hanya diambil kriteria yang bersifat umum saja misal, jenis jaminan, nilai jaminan, dan penghasiln. Padahal masih banyak kriteria-kriteria yang bisa dijadikan sebagai penunjang pengajuan kredit nasabah disetujui. Data dan informasi yang terkait dengan persetujuan nasabah mengajukan pinjaman umumnya bersifat banyak dan kegiatan manual tidak memungkinkan untuk mengolah system secara cepat, tepat dan akurat. Karena itu diperlukan suatu algoritma terbaik untuk penentuan disetujui atau tidaknya pengajuan pinjaman oleh nasabah yang bisa menghasilkan nilai keefektifan dan tepat sasaran.

3.4.2 Langkah-langkah Penelitian

Setiap tahun dilakukan penilaian oleh bagian pengambil keputusan yang mencakup dua hal penting, yakni *review* terhadap riwayat pinjaman nasabah dan rencana pengajuan kembali pinjaman nasabah dengan menggunakan data-data yang sudah ada dan riwayat peminjaman nasabah sebelumnya. Dalam hal ini manajemen sales mendiskusikan dengan *supervisor*/atasanlangsung-nya. Dalam peraturan perusahaan dapat ditentukan bahwa yang berwenang dalam membuat keputusan adalah agen peminjaman dan agen sales manager. Penelitian ini dilaksanakan berdasarkan alur penelitian seperti yang ditunjukkan pada Gambar 3.2.

Langkah-langkah yang digunakan dalam penelitian ini menggunakan model CRISP-DM (*Cross Standard Industries Process for Data Mining*), dalam metode ini terdapat 6 tahapan[14][16]:



Gambar 3.2 Langkah-langkah Penelitian

Keterangan:

1. *Business/Research Understanding Phase*

Permasalahan yang ada di PT Batavia Prosperindo Finance Tbk adalah belum adanya sistem yang digunakan dalam melakukan analisis prediksi pemberian kredit kepada calon nasabah. Berangkat dari permasalahan ini, pada fase pemahaman bisnis, penulis mengumpulkan data yang terkait dengan penghitungan prediksi, langkah ini dilakukan dengan cara melakukan wawancara dan observasi di objek penelitian setempat, wawancara dilakukan dengan bagian *Marketing* dan *Service Management* yang berwenang menentukan kebijakan pemberian kredit kepada nasabah di PT Batavia Prosperindo Tbk. Data yang dikumpulkan dengan beberapa atribut yaitu: no_kontrak, nama, accnumber, accname, tgl_lahir, gender, alamat, kota, no_hp, pendidikan, status, jumlah_tanggungan, status_tempat_tinggal, pekerjaan, penghasilan_per_tahun, tgl_transaksi, tgl_bayar_akhir, tunggakan, merk_asset, jenis_asset, nama_asset, tahun_asset, no_rangka, no_mesin, no_polisi, bpkb, warna, nilai_jaminan, jumlah_pinjamanmax, tujuan_pinjaman, cmo.

Data yang dikumpulkan dengan beberapa *field* dengan jumlah atribut sebanyak 31 yaitu :

Tabel 3.2 Atribut *Data Master*

No	Atribut	Keterangan
1	no_kontrak	nomor kontrak data nasabah
2	nama	Nama nasabah
3	accnumber	No rekening nasabah
4	accname	Nama pemegang rekening
5	Tgl_lahir (<i>dob</i>)	tanggal, bulan dan tahun lahir
6	gender	Jenis kelamin
7	Alamat	Alamat nasabah
8	Kota	Kategori wilayah berdasarkan wilayah/area
9	No_hp	Mobile phone
10	pendidikan	Pendidikan terakhir
11	Status	Status pernikahan
12	Jumlah_tanggungan	Jumlah tanggungan keluarga
13	Status_tempat_tinggal	Kepemilikan rumah
14	pekerjaan	Jenis pekerjaan
15	Penghasilan_per_tahun	Penghasilan per tahun
16	Tgl_transaksi	Tanggal transaksi terakhir melakukan pinjaman
17	Tgl_bayar_akhir	Tanggal pelunasan terakhir
18	tunggakan	Jumlah tunggakan
19	Merk_asset	Merek asset/jaminan
20	Jenis_asset	Jenis asset/jaminan
21	Nama_asset	nama asset/jaminan
22	Tahun_asset	Tahun pembuatan asset/jaminan
23	No_rangka	Nomer Rangka kendaraan
24	No_mesin	Nomer mesin kendaraan
25	No_polisi	Nomer polisi kendaraan
26	bpkb	Nomer BPKB kendaraan
27	warna	Warna kendaraan
28	Nilai_jaminan	Jumlah nilai taksiran jaminan
29	Jumlah_pinjamanmax	Jumlah pinjaman maksimal yang diinginkan
30	Tujuan_pinjaman	Tujuan pinjaman
31	cmo	Nama sales penjamin

2. *Data Understanding Phase* (Fase Pemahaman Data)

Pada tahap ini mulai dilakukan analisis data nasabah dimana data yang diambil hanya data pada tahun 2018-2020. Pada fase pemahaman ini dilakukan wawancara dengan staf marketing dan *service* managemen untuk mendalami data dan penentuan metode prediksi yang sedang berjalan. Pada tahap ini mulai dilakukan menganalisis data *master* yang direlasikan dengan data sekunder. Data yang diambil akan dikumpulkan sebagai data *training* dan data *testing* dengan pembagian *dataset* sebesar 80% dan 20% dari 300 *record*. Pada fase pemahaman ini peneliti melakukan wawancara dengan staf *database* untuk mendalami data dan penentuan metode prediksi pengajuan kredit yang sedang berjalan. Adapun *field* dari data *training* yaitu: *gender*, *age*, *pendidikan_terakhir*, *status_perkawinan*, *jumlah_tanggungan*, *status_tempat_tinggal*, *pekerjaan*, *penghasilan_pertahun*, *nilai_jaminan*, *tujuan_pinjaman*, *jumlah_pinjaman*, *hasil*.

3. *Data Preparation Phase* (Fase Pengolahan Data)

Dari fase pemahaman data akan menjadi modal peneliti untuk masuk ke fase pengolahan data, tahap ini meliputi semua kegiatan untuk membangun *dataset* akhir data yang akan diproses pada tahap pemodelan (*modeling*).

4. *Modeling Phase* (Fase Pemodelan)

Pada tahap ini membandingkan, memilih dan menerapkan teknik pemodelan klasifikasi *data mining* yaitu *Naïve Bayes* dan *Decision Tree*. Data yang diperoleh dari fase pengolahan data akan digunakan pada proses ini.

5. *Evaluation Phase* (Fase Evaluasi)

Pada tahap ini akan dilakukan evaluasi terhadap keefektifan dari perbandingan klasifikasi data mining, sebelum digunakan dan menentukan apakah model dapat mencapai tujuan yang diterapkan pada fase *business understanding*. Ketepatan atau ketelitian menjadi kriteria performance suatu metode peramalan. Ketepatan atau ketelitian tersebut dapat

dinyatakan sebagai kesalahan dalam peramalan. Metode yang digunakan dalam fase ini adalah *k-fold cross validation*.

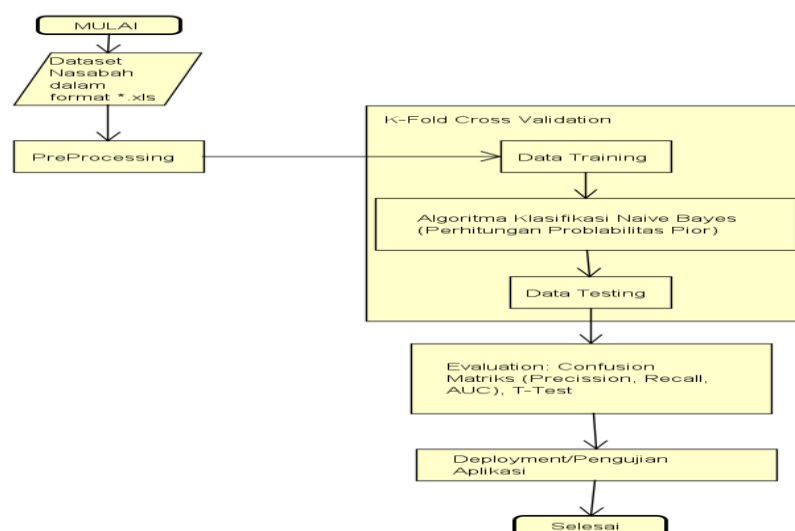
6. *Deployment Phase* (Fase Penyebaran)

Tahap *deployment* dilakukan dengan pembuatan prototype klasifikasi *data mining* dengan menggunakan model yang terbentuk pada proses sebelumnya, pemodelan *prototype* menggunakan UML dengan *use case diagram* dan *activity diagram*, Untuk pengujian aplikasi akan digunakan metode pengujian kotak hitam (*Black Box Testing*). Hasil dari fase ini diharapkan sebuah aplikasi *data mining* yang siap digunakan dan bermanfaat untuk PT Batavia Prosperindo Finance Tbk.

3.4.3 Desain global sistem baru

3.4.3.1 Algoritma Naïve Bayes

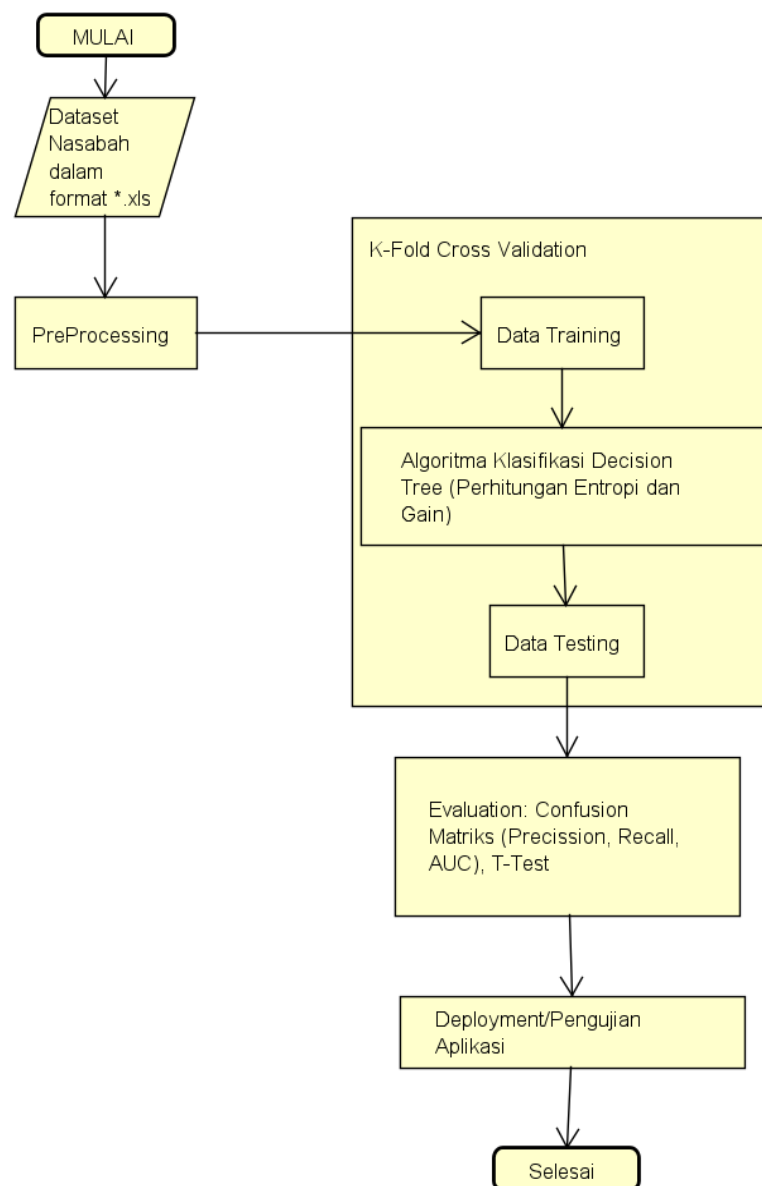
Teknik data mining untuk klasifikasi pengajuan pinjaman dengan Naïve Bayes dapat digunakan untuk membantu pengambilan keputusan pemberian kredit dan menganalisa kemampuan nasabah dalam membayar kredit. Algoritma Naïve Bayes digunakan untuk mencari nilai probabilitas prior masing-masing kriteria. Bagan alir perhitungan algoritma Naïve Bayes seperti ditunjukkan pada gambar di bawah ini.



Gambar 3.3 Bagan Alir Algoritma *Naïve Bayes*

3.4.3.2 Algoritma *Decision Tree*

Algoritma *Decision Tree* digunakan untuk mencari nilai *entropi* per-atribut dan menghitung nilai *gain* per-atribut dimana kriteria dengan nilai *gain* tertinggi menjadi akar dari *tree*(pohon). Bagan alir algoritma *decision tree* seperti ditunjukkan pada gambar di bawah ini:



Gambar 3.4 Bagan Alir Algoritma *Decision tree*

3.5 Instrumentasi

Instrumentasi yang akan digunakan dalam penelitian ini adalah sebagai berikut.

1. Instrumen berupa data nasabah PT Batavia Prosperindo Finance Tbk yang akan digunakan untuk pengajuan kredit dalam bentuk *softcopy* karena kendala kerahasiaan Perusahaan.
2. Instrumen wawancara berupa pertanyaan terstruktur untuk mendapat hasil informasi mengenai prosedur analisis pengolahan data.
3. Instrumen studi pustaka yaitu mengumpulkan referensi yang berkaitan dengan data mining klasifikasi baik yang terkait dengan prediksi maupun studi komparasi.
4. Instrumen perangkat keras (*hardware*) berupa laptop DELL-S2D0UO4 dengan spesifikasi sebagai berikut.
 - a. Processor : Intel(R)Core(TM) i5-4310U CPU@ 2.00GHz (4CPUs),~2.6GHz.
 - b. Harddisk : 500 GB
 - c. RAM : 4096 MB
 - d. Sistem Operasi : Windows 10 Pro 64-bit (10.0,Build 18362)
5. Instrumen perangkat lunak (*software*) yaitu
 - a. Rapid Miner 9 : *tools* yang membantu dalam pengujian
 - b. XAMPP : server *localhost* saat mendesign sistem
 - c. Sublime Text : aplikasi untuk pemrograman PHP
 - d. Google Chrome : *browser* untuk menampilkan program

3.6 Teknik Analisis, Desain, dan Pengujian

3.6.1 Teknik Analisis

Teknik analisa deskriptif dilakukan untuk menganalisa data yang akan dilakukan terhadap hasil pengumpulan data dengan studi pustaka, wawancara dan observasi untuk mendapatkan spesifikasi kebutuhan sistem yang akan dikembangkan.

Teknik analisis yang akan dilakukan menggunakan metode-metode klasifikasi *data mining* antara lain: *Naive Bayes* dan *Decision Tree*. Metode-metode tersebut digunakan untuk mengolah data pengajuan kredit dari nasabah yang akan mengajukan pinjaman kembali, guna menghasilkan prediksi yang akurat dalam analisis pemberian kredit.

Data yang didapat akan dibagi menjadi dua *set* yaitu: data *training* dan data *testing*. Hasil dari masing-masing metode dengan data latih akan dibandingkan hasil pengujiannya dengan menggunakan *k-fold cross validation* dengan $k=10$ untuk mendapatkan hasil berupa nilai akurasi, *precision*, *recall*, *ROC's curve*.

3.6.2 Desain Proses Klasifikasi pada *Prototype*

Proses yang akan dirancang dalam prototipe sistem antara lain :

1. *Import data excel*

Import data dilakukan untuk memasukkan data yang akan diprediksi ke dalam prototipe yang akan dirancang. Format data berupa .csv atau xls.

2. *Preprocessing*

Data yang di-*import* akan dicek keseluruhan isinya oleh prototipe untuk mengetahui kelayakan dalam pemrosesan. Proses pengecekan dapat berupa pengecekan *missing value*, perbedaan format data, dan lain-lain.

3. Proses prediksi

Setelah data bersih maka akan dilakukan prediksi terhadap data uji menggunakan metode dengan nilai akurasi terbaik yang telah melalui tahapan komparasi dalam proses analisis pada penelitian. Hasil prediksi berupa akurasi prediksi penentuan kredit nasabah. Dalam prototipe dinyatakan dalam notasi Diterima dan Ditolak.

3.6.3 Teknik Pengujian

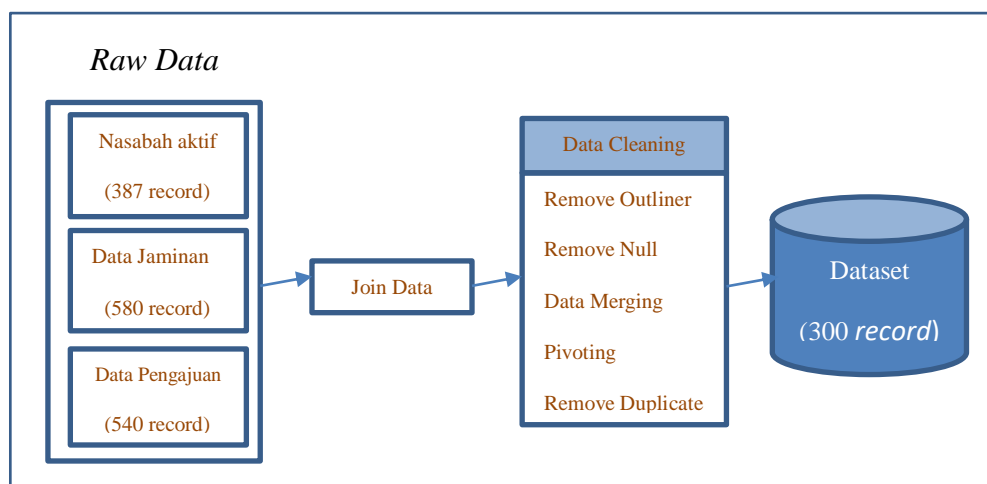
Teknik pengujian terhadap metode yang akan dilakukan menggunakan *k-folds cross validation* dengan $k=10$. [1] Metode ini membagi data latih secara acak menjadi 10 bagian dengan jumlah yang hampir sama pada masing-masing kelompok. Pada setiap perulangan dalam proses *training*, maka 1 bagian data digunakan sebagai data uji dan 9 bagian data lainnya sebagai data latih. Proses *training* dilakukan sebanyak 10 kali data *testing*. Hasil pengujian akan didapatkan dengan menghitung rata-rata nilai-nilai statistik pengujian pada keseluruhan perulangan.

BAB IV

PEMBAHASAN HASIL PENELITIAN

4.1 Akuisisi Data

Akuisisi data merupakan sistem yang digunakan untuk mengambil, mengumpulkan dan menyiapkan data yang sedang berjalan, kemudian data tersebut diolah lebih lanjut dalam komputer untuk keperluan tertentu [2]. Sistem akuisisi data dapat didefinisikan sebagai suatu sistem yang berfungsi untuk mengambil, mengumpulkan dan menyiapkan data, hingga memprosesnya untuk menghasilkan data yang dikehendaki. Jenis serta metode yang di pilih pada umumnya bertujuan untuk menyederhanakan setiap langkah yang dilaksanakan pada keseluruhan proses. Suatu sistem akuisisi data pada umumnya dibentuk sedemikian rupa sehingga sistem tersebut berfungsi untuk mengambil, mengumpulkan dan menyimpan data dalam bentuk yang siap untuk diproses lebih lanjut.



Gambar 4.1 Proses Akuisisi Data

Gambar 4.1 menyajikan proses akuisisi data penelitian sehingga diperoleh dataset penelitian. Beberapa data sumber yang ditarik dari pusat penyimpanan data yaitu: (1) data nasabah aktif, (2) data jaminan nasabah, (3) data pengajuan. Seluruh data sumber selanjutnya digabungkan menjadi satu dataset dan dilakukan proses pembersihan data. Proses pembersihan data diperlukan untuk memastikan

dataset yang diperlukan sudah sesuai. Berikut tahapan proses data *cleaning*:

1. *Remove outlier*, untuk menghilangkan data yang memiliki nilai yang salah dan diluar jangkauan nilai yang wajar.
2. *Remove null*, untuk menghilangkan record atau field yang kosong
3. *Data merging*, untuk menggabungkan beberapa record dengan data yang memiliki arti yang sama.
4. *Pivoting*, untuk mengubah dimensi data nasabah dari baris ke kolom.
5. *Remove duplicate*, untuk menghilangkan data ganda yang memiliki nilai dan arti yang sama.

Proses akuisisi data menghasilkan *dataset* penelitian sebanyak 300 *record* dimana terdapat 239 disetujui dan 61 ditolak. Pemilihan data dalam penelitian ini dituangkan dalam pembuatan model sistem. Model sistem yang akan dibuat memiliki batasan-batasan sebagai berikut:

1. Perancangan yang dibuat dengan menggunakan penalaran *data* dengan menggunakan metode *Naïve Bayes* dan *Decision Tree*.
2. Pembuatan aturan dalam basis pengetahuan dibantu oleh Kepala HRD, Kepala Marketing Manager dan Kepala GA di PT Batavia Prosperindo Finance Tbk.

4.2 Seleksi Atribut

Pada tahap awal penelitian, seleksi terhadap atribut evaluasi penentuan pengajuan pinjaman kredit pada Tabel 4.1 berikut dilakukan untuk mengurangi atribut awal. Pengurangan atribut dilakukan berdasarkan nilai *information gain* dari masing-masing atribut terhadap *entropi* kelas targetnya. Semakin tinggi nilai *information gain* dapat dikatakan bahwa atribut tersebut semakin penting atau semakin mempengaruhi keputusan kelas targetnya.

Tabel 4.1 Atribut Data Penelitian Berdasarkan *Information Gain*

No.	Atribut	Tipe Atribut	Nilai Atribut
1.	<i>Gender</i>	Binomial	Pria Wanita
2.	<i>Age</i>	Binomial	<=40 tahun >40 tahun
3.	Pendidikan_terakhir	Polinomial	Magister Sarjana Diploma SMA/SMK SMP SD
4.	Status	Polinomial	Belum Menikah Menikah Janda/Duda
5.	Jumlah_tanggungan	Binomial	<=2 >2
6.	Status_tempat_tinggal	Polinomial	Sewa Milik Sendiri Milik Keluarga Rumah Dinas
7.	Pekerjaan	Polinomial	PNS Wiraswasta TNI/Polri Guru/Dosen Petani Lain-lain
8.	Penghasilan_per_tahun	Binomial	<=45 juta >45 juta
9.	Nilai_jaminan	Binomial	<100 juta >100 juta
10.	Tujuan_Pinjaman	Binomial	Modal Usaha Kendaraan
11.	Pinjaman_max	Binomial	<=60 juta >60 juta
12.	Hasil <i>Survey</i>	Binomial	Disetujui(<i>Approval</i>) Ditolak (<i>Rejected</i>)

4.3 Pendefinisian Atribut Kelengkapan

Penentuan atribut-atribut yang digunakan dilakukan setelah melakukan wawancara dengan bagian terkait mengenai atribut apa saja yang paling penting digunakan untuk evaluasi nasabah yang melakukan pengajuan pinjaman. Hasilnya ditetapkan 11 buah atribut data pengajuan yaitu: *gender*, *age*, *pendidikan_terakhir*, *status*, *jumlah_tanggungan*, *status_tempat_tinggal*, *pekerjaan*, *penghasilan_perthn*, *nilai_jaminan*, *tujuan_pinjaman*, dan *pinjaman_max*. Atribut-atribut ini menjadi *input* pada perancangan *prototype*, sementara untuk *output*-nya adalah Disetujui atau Ditolak.

Hal yang pertama dilakukan dalam penelitian ini adalah pemakaian dua algoritma yaitu *Naïve Bayes* dan *Decision Tree* untuk membentuk sebuah model. Model yang dihasilkan akan digunakan untuk mengetahui pola kelayakan pemberian kredit kepada calon nasabah yang mengajukan kredit dengan melihat data-data atribut yang telah ditentukan sebelumnya.

Pada saat penelitian ini dilakukan proses validasi untuk menemukan, dan mengkonversi data agar dapat digunakan dalam metode algoritma *data mining* dan memperoleh akurasi serta *performance* yang baik. Dalam *dataset* yang akan digunakan ini, validasi data yang digunakan adalah dengan menghapus data yang tidak lengkap atau kosong yang tidak memiliki nilai (*null*). Setelah itu dilakukan seleksi atribut untuk memilih atribut mana saja yang dibutuhkan dari dataset yang digunakan dalam proses menganalisis kelayakan pemberian kredit kepada calon nasabah. Berikut diberikan beberapa data training seperti tabel 4.2

Tabel 4.2 Sampel Data Training

No.	No Kontrak	Nama	Age	Gender	Pend.ter akhir	Status	Jumlah Tangg.Kel.	Kepemilikan Rumah	Pekerjaan	Penghasilan/tahun	Nilai Jaminan	Tujuan Pinjaman	Pinjaman max	HASIL
1	074372180004	TASRIP	KP	L	SMA	Menikah	K	Milik Sendiri	Wiraswasta	K	K	Kendaraan	B	Disetujui
2	074372180005	EKO SANTOSO	P	L	SMA	Menikah	K	Milik Sendiri	Wiraswasta	K	K	Modal Usaha	K	Disetujui
3	074372180010	YENI EFTIANINGSIH	KP	P	SMA	Menikah	K	Milik Sendiri	Wiraswasta	K	B	Kendaraan	K	Disetujui
4	074372180011	RESTU NURUL FATIMAH	KP	P	S1	Menikah	K	Milik Keluarga	Guru/Dosen	B	B	Kendaraan	B	Disetujui
5	074372180012	AGUS SURAHMAN	P	L	SMA	Menikah	B	Milik Sendiri	PNS	K	K	Kendaraan	B	Disetujui
6	074372180016	JAENURI	KP	L	SMA	Menikah	K	Milik Sendiri	Wiraswasta	B	B	Kendaraan	B	Disetujui
7	074372190002	ISRIYADI	P	L	SMA	Menikah	K	Milik Sendiri	Wiraswasta	K	K	Kendaraan	K	Disetujui
8	074372190004	CIPTO SUBAGIO	P	L	SMA	Menikah	K	Milik Sendiri	Wiraswasta	B	B	Kendaraan	B	Disetujui
9	074372190006	SUNGATNO	P	L	SMP	Menikah	K	Milik Sendiri	Petani	K	B	Kendaraan	B	Disetujui
10	074372190007	DEDE MAHENDRA (Nasabah GP)	KP	L	SMA	Menikah	K	Milik Keluarga	Petani	B	K	Kendaraan	B	Disetujui
11	074372190008	HENDRA RUDHIANATA	KP	L	SMA	Menikah	B	Milik Sendiri	Petani	K	K	Modal Usaha	K	Disetujui
12	074372190009	ISLAH *(Nasabah GP)	KP	L	SMA	Menikah	K	Milik Sendiri	Wiraswasta	B	B	Kendaraan	B	Disetujui
13	074372190012	YENIAR EFENDI	KP	L	S2	Menikah	B	Milik Keluarga	PNS	K	K	Modal Usaha	B	Ditolak
14	074372190013	DARMAWAN	KP	L	SMA	Menikah	B	Milik sendiri	Wiraswasta	K	K	Kendaraan	B	Ditolak
15	074372190014	NOVRIADI	KP	L	SMA	Menikah	B	Milik Sendiri	Wiraswasta	K	K	Modal Usaha	K	Ditolak
16	074372190016	RIZKI INDRA SETIAWAN (Nasabah GP)	KP	L	SMA	Menikah	B	Milik Keluarga	Wiraswasta	B	B	Kendaraan	B	Ditolak
17	074372190017	MUKHLISH *(Nasabah GP)	KP	L	SMP	Duda	B	Milik Sendiri	Wiraswasta	B	K	Modal Usaha	K	Ditolak
18	074372190018	RINI PUJI LESTARI	KP	P	SMP	Menikah	B	Milik Sendiri	Lain-lain	K	B	Kendaraan	B	Ditolak
19	074372190019	TUNUT	P	L	SMA	Belum Menikah	K	Milik Keluarga	Petani	B	B	Modal Usaha	B	Ditolak
20	074372190021	NANANG PURWANTO (Nasabah GP)	P	L	SMA	Menikah	K	Rumah Dinas	TNI/POLRI	B	K	Kendaraan	K	Ditolak
21	074372190025	DEDI ARPANI	P	L	SMA	Belum Menikah	K	Sewa	Wiraswasta	B	B	Kendaraan	B	Disetujui
22	074372190031	MIRHAN	P	L	SMA	Menikah	B	Sewa	Wiraswasta	K	K	Kendaraan	B	Disetujui
23	074372190032	NOPRIANDI	P	L	SMA	Menikah	K	Milik Keluarga	Wiraswasta	K	K	Kendaraan	B	Disetujui
24	074372190034	ZULYADI MUNZIRI	P	L	SMA	Menikah	B	Milik Sendiri	Wiraswasta	K	K	Kendaraan	B	Disetujui
25	074372190035	SYAHRUDDIN AL JAWI	P	L	SMA	Menikah	K	Milik Sendiri	Petani	K	K	Kendaraan	K	Disetujui
26	074372190036	MIMBAR WIBOWO (Nasabah GP)	P	L	SMA	Menikah	K	Milik Sendiri	Petani	B	B	Kendaraan	B	Disetujui
27	074372190037	FIRDAUS SE	P	L	SMA	Menikah	B	Milik Sendiri	Petani	B	B	Kendaraan	B	Disetujui
28	074372190038	CATUR DWIYATNO	P	L	SMA	Menikah	K	Milik Sendiri	PNS	B	B	Kendaraan	B	Disetujui
29	074372190041	HOTIB	KP	L	SMA	Menikah	K	Milik Sendiri	Petani	K	K	Kendaraan	K	Disetujui

Tabel 4.2 (lanjutan)

No.	No Kontrak	Nama	Age	Gender	Pend.ter akhir	Status	Jumlah Tangg.Kel.	Kepemilikan Rumah	Pekerjaan	Penghasilan/tahun	Nilai Jaminan	Tujuan Pinjaman	Pinjaman max	HASIL
30	074372190042	ELIA NOVITA WATI *(Nasabah GP)	KP	P	SMA	Menikah	K	Milik Sendiri	Petani	B	K	Kendaraan	B	Ditolak
31	074372190043	ADYANTO *(Nasabah GP)	P	L	SMA	Menikah	K	Milik Sendiri	Petani	K	K	Kendaraan	K	Disetujui
32	074372190046	TONI	P	L	S1	Menikah	K	Milik Sendiri	Wiraswasta	K	K	Kendaraan	K	Disetujui
33	074372190047	IMAM SOPINGI	P	L	S1	Menikah	K	Milik Sendiri	PNS	B	B	Kendaraan	B	Disetujui
34	074372190049	PONDI	P	L	SMA	Menikah	B	Milik Sendiri	Petani	B	B	Kendaraan	B	Disetujui
35	074372190050	ANTON PRAYOGO	P	L	SMA	Belum Menikah	K	Milik Sendiri	Petani	B	B	Kendaraan	B	Disetujui
36	074372190051	YULI ASMARA	P	P	S1	Belum Menikah	K	Sewa	Wiraswasta	K	K	Kendaraan	K	Disetujui
37	074372190053	MALA SARI	P	P	DIPLOMA	Belum Menikah	K	Milik Sendiri	PNS	K	K	Kendaraan	B	Disetujui
38	074372190054	JUNAIDI	P	L	SMA	Belum Menikah	K	Milik Keluarga	Wiraswasta	K	B	Kendaraan	B	Disetujui
39	074372190056	BAMBANG GUNAWAN	P	L	SMP	Menikah	B	Milik Keluarga	Wiraswasta	K	K	Kendaraan	K	Disetujui
40	074372190058	AMRAN SE	P	L	SMA	Menikah	K	Milik Keluarga	Wiraswasta	B	B	Modal Usaha	B	Disetujui
41	074372190059	SURA A TMA JA	P	L	SMA	Menikah	K	Milik Keluarga	Wiraswasta	B	B	Modal Usaha	B	Disetujui
42	074372190063	FITRI LESTARI	P	P	SMA	Menikah	K	Milik Keluarga	Wiraswasta	B	B	Kendaraan	B	Disetujui
43	074372190065	ROSIHUN *(Nasabah GP)	P	L	S1	Menikah	K	Milik Sendiri	Lain-lain	K	K	Modal Usaha	K	Disetujui
44	074372190066	KODAR AMINUDIN	KP	L	DIPLOMA	Menikah	K	Milik Sendiri	Lain-lain	K	K	Modal Usaha	K	Disetujui
45	074372190067	MASKUR	KP	L	S1	Menikah	B	Milik Sendiri	Wiraswasta	K	K	Modal Usaha	B	Disetujui
46	074372190071	SYAHRIL ANWAR	P	L	S1	Menikah	K	Milik Sendiri	Petani	K	K	Kendaraan	K	Disetujui
47	074372190072	IKA RACHIMAH	P	P	S1	Menikah	B	Milik Sendiri	Wiraswasta	B	B	Kendaraan	B	Ditolak
48	074372190074	ZIKRANI	P	L	SMA	Menikah	B	Milik Sendiri	Wiraswasta	B	B	Kendaraan	B	Disetujui
49	074372190075	NURHADI	P	L	SMA	Menikah	B	Milik Sendiri	Wiraswasta	B	K	Modal Usaha	B	Disetujui
50	074372190076	DEDI MARGONO	P	L	SMA	Belum Menikah	K	Milik Sendiri	Wiraswasta	B	B	Kendaraan	B	Ditolak
51	074372190079	KUSDA RMAJI	P	L	SMA	Belum Menikah	K	Milik Sendiri	Petani	K	K	Kendaraan	K	Ditolak
52	074372190080	BURHANUDIN RAIS	P	L	SMA	Belum Menikah	K	Milik Sendiri	Wiraswasta	K	K	Kendaraan	K	Ditolak
53	074372190082	HARDI YANTO *(Nasabah GP)	P	L	DIPLOMA	Belum Menikah	K	Milik Sendiri	Petani	B	K	Kendaraan	B	Ditolak
54	074372190084	ENDRIYANI	P	P	DIPLOMA	Belum Menikah	B	Sewa	Wiraswasta	B	B	Modal Usaha	B	Ditolak
55	074372190086	EDI PRASTYO *(Nasabah GP)	P	L	S1	Belum Menikah	K	Sewa	Petani	B	K	Modal Usaha	K	Ditolak
56	074372190087	JAMRONI	KP	L	SMP	Duda	K	Milik Sendiri	Wiraswasta	B	B	Kendaraan	B	Ditolak
57	074372190088	SUSANTO *(Nasabah GP)	KP	L	SMP	Menikah	K	Milik Sendiri	Petani	K	B	Kendaraan	B	Ditolak
58	074372190090	HARDIONO	KP	L	SD	Duda	K	Milik Sendiri	Petani	B	K	Kendaraan	B	Ditolak
59	074372190091	ARI JOHARI	P	L	SMA	Menikah	B	Milik Sendiri	Petani	B	K	Kendaraan	B	Ditolak
60	074372190092	IBNU SAYAF *(Nasabah GP)	P	L	S1	Menikah	K	Milik Sendiri	PNS	K	K	Kendaraan	K	Ditolak
...
300	074372200111	MUHAMMAD RAZIK	P	L	S1	Menikah	K	Milik Sendiri	Wiraswasta	B	B	Kendaraan	B	Disetujui

4.3.1 Pengujian Klasifikasi dengan Algoritma *Decision Tree*

Pada penelitian ini akan dilakukan perhitungan manual pada *data sample*. Perhitungan dilakukan pada 30 data sampel dari 300 *data sample* berdasarkan data yang terdapat pada tabel 4.2.

Berdasarkan tabel 4.2 dibuat tabel keputusan untuk menentukan kelayakan pemberian kredit kepada calon nasabah dengan melihat *atribut dataset* yang telah ditentukan. Langkah pertama menghitung total *entropi* dari jumlah kasus seluruhnya, jumlah keputusan “Disetujui” maupun “Ditolak”, dengan persamaan:

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

$$Entropy\ Total = \left(-\frac{21}{30} \log_2\left(\frac{21}{30}\right)\right) + \left(-\frac{9}{30} \log_2\left(\frac{9}{30}\right)\right) = 0.88129$$

Sehingga dari perhitungan diatas didapat total *entropi* = 0.88129

Kemudian menghitung total *entropi* dan *gain* dari semua kasus yang terbagi berdasarkan atribut “*age*”, “*gender*”, “*pend_terakhir*”, “*status*”, “*juml_tangg_kel*”, “*status_tempat_tinggal*”, “*pekerjaan*”, “*penghasilan_per_tahun*”, “*nilai_jaminan*”, “*tujuan_pinjaman*”, dan “*pinjaman_max*”.

1. Atribut *age*

Untuk atribut *age* pada tabel 4.2 terdapat 30 kasus dengan 15 nilai kategori ‘<=40’; dan 15 nilai kategori ‘>40’, dengan perincian untuk nilai kategori ‘<=40’ 13 ‘disetujui’ dan 2 ‘ditolak’, untuk nilai kategori ‘>40’ 8 ‘disetujui’ dan 7 ‘ditolak’, sehingga nilai *entropi age* dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy\ (<=40) = \left(-\frac{13}{15} \log_2\left(\frac{13}{15}\right)\right) + \left(-\frac{2}{15} \log_2\left(\frac{2}{15}\right)\right) = 0.56536$$

$$Entropy\ (>40) = \left(-\frac{8}{15} \log_2\left(\frac{8}{15}\right)\right) + \left(-\frac{7}{15} \log_2\left(\frac{7}{15}\right)\right) = 0.9967$$

Kemudian menghitung nilai *gain age* sebagai berikut:

$$Gain_{(age)} = Entropy\ Total - \left(\frac{15}{30} \times Entropy(" \leq 40")\right) + \frac{15}{30} \times Entropy(" > 40")$$

$$Gain_{(age)} = 0.88129 - \left(\frac{15}{30} \times (0.56536)\right) + \frac{15}{30} \times (0.9967)$$

$$Gain_{(age)} = 0.10026$$

2. Atribut *gender*

Untuk atribut *gender* pada tabel 4.2 terdapat 30 kasus dengan 26 nilai ‘laki-laki’; dan 4 nilai ‘perempuan’, dengan perincian untuk nilai ‘laki-laki’ 19 ‘disetujui’ dan 7 ‘ditolak’, untuk nilai ‘perempuan’ 2 ‘disetujui’ dan 2 ‘ditolak’, sehingga nilai *entropi gender* dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy\ ('laki-laki') = \left(-\frac{19}{26} \log_2 \left(\frac{19}{26}\right)\right) + \left(-\frac{7}{26} \log_2 \left(\frac{7}{26}\right)\right) = 0.84041$$

$$Entropy\ ('perempuan') = \left(-\frac{2}{4} \log_2 \left(\frac{2}{4}\right)\right) + \left(-\frac{2}{4} \log_2 \left(\frac{2}{4}\right)\right) = 1$$

Kemudian menghitung nilai *gain jenis gender* sebagai berikut:

$$Gain_{(gender)} = Entropy\ Total - \left(\frac{26}{30} \times Entropy\ ('laki.laki')\right) + \frac{4}{30} \times Entropy\ ('perempuan')$$

$$Gain_{(gender)} = 0.88129 - \left(\frac{26}{30} \times (0.84041)\right) + \frac{4}{30} \times (1)$$

$$Gain_{(gender)} = 0.01965$$

3. Atribut *pend_terakhir*

Untuk atribut *pend_terakhir* pada tabel 4.2 terdapat 30 kasus dengan 3 nilai 'SMP'; 25 nilai 'SMA'; 1 nilai 'S1'; dan 1 nilai 'S2', dengan perincian untuk nilai 'SMP' 1 'disetujui' 2 'ditolak', untuk nilai 'SMA' 19 'disetujui' 6 'ditolak', untuk nilai 'S1' 1 'disetujui' 0 'ditolak', untuk nilai 'S2' 0 'disetujui' 1 'ditolak' sehingga nilai entropi *pend_terakhir* dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy('SMP') = \left(-\frac{1}{3}\log_2\left(\frac{1}{3}\right)\right) + \left(-\frac{2}{3}\log_2\left(\frac{2}{3}\right)\right) = 0.91825$$

$$Entropy('SMA') = \left(-\frac{19}{25}\log_2\left(\frac{19}{25}\right)\right) + \left(-\frac{6}{25}\log_2\left(\frac{6}{25}\right)\right) = 0.79503$$

$$Entropy('S1') = \left(-\frac{1}{1}\log_2\left(\frac{1}{1}\right)\right) + \left(-\frac{0}{1}\log_2\left(\frac{0}{1}\right)\right) = 0$$

$$Entropy('S2') = \left(-\frac{0}{1}\log_2\left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1}\log_2\left(\frac{1}{1}\right)\right) = 0$$

Kemudian menghitung nilai *gain* *pend_terakhir* sebagai berikut:

$$\begin{aligned} Gain_{(pend_terakhir)} = Entropy\ Total - & \left(\frac{3}{30} \times Entropy("SMP")\right) + \\ & \left(\frac{25}{30} \times Entropy("SMA")\right) + \frac{1}{30} \times Entropy("S1") \\ & + \frac{1}{30} \times Entropy("S2") \end{aligned}$$

$$\begin{aligned} Gain_{(pend_terakhir)} = 0.88129 - & \left(\frac{3}{30} \times (0.91825)\right) + \frac{25}{30} \times (0.79503) + \\ & \left(\frac{1}{30} \times (0) + \left(\frac{1}{30} \times (0)\right)\right) \end{aligned}$$

$$Gain_{(pend_terakhir)} = 0.12696$$

4. Atribut status

Untuk atribut status pada tabel 4.2 terdapat 30 kasus dengan 2 nilai 'belum_menikah', 27 nilai 'menikah', dan 1 nilai 'duda' dengan perincian untuk nilai 'belum_menikah' 1 'disetujui' dan 1 'ditolak', untuk nilai 'menikah' 20 'disetujui' dan 7 'ditolak', untuk nilai 'duda' 0 'disetujui' dan 1 'ditolak', sehingga nilai *entropi* status dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy ('belum_menikah') = \left(-\frac{1}{2} \log_2 \left(\frac{1}{2}\right)\right) + \left(-\frac{1}{2} \log_2 \left(\frac{1}{2}\right)\right) = 1$$

$$Entropy ('menikah') = \left(-\frac{20}{27} \log_2 \left(\frac{20}{27}\right)\right) + \left(-\frac{7}{27} \log_2 \left(\frac{7}{27}\right)\right) = 0.82557$$

$$Entropy ('duda') = \left(-\frac{0}{1} \log_2 \left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} \log_2 \left(\frac{1}{1}\right)\right) = 0$$

Kemudian menghitung nilai *gain* status sebagai berikut:

$$Gain_{(status)} = Entropy Total - \left(\frac{2}{30} \times Entropy ('belum_menikah') + \frac{27}{30} \times Entropy ('menikah') + \frac{1}{30} \times Entropy ('duda')\right)$$

$$Gain_{(status)} = 0.88129 - \left(\frac{2}{30} \times (1) + \frac{27}{30} \times (0.82557) + \frac{1}{30} \times (0)\right)$$

$$Gain_{(status)} = 0.07157$$

5. Atribut juml_tangg_kel

Untuk atribut juml_tangg_kel pada tabel IV-2 terdapat 30 kasus dengan 19 nilai kategori '<=2'; dan 11 nilai kategori '>2', dengan perincian untuk nilai kategori '<=2' 16 'disetujui' dan 3 'ditolak', untuk nilai kategori '>2' 5 'disetujui' dan 6 'ditolak', sehingga nilai entropi juml_tangg_kel dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy ("<=2") = \left(-\frac{16}{19} \log_2 \left(\frac{16}{19}\right)\right) + \left(-\frac{3}{19} \log_2 \left(\frac{3}{19}\right)\right) = 0.62922$$

$$Entropy(">2") = \left(-\frac{5}{11} \log_2 \left(\frac{5}{11}\right)\right) + \left(-\frac{6}{11} \log_2 \left(\frac{6}{11}\right)\right) = 0.99402$$

Kemudian menghitung nilai *gain* *juml_tangg_kel* sebagai berikut:

$$Gain_{(juml_tangg_kel)} = Entropy\ Total - \left(\frac{19}{30} \times Entropy("<= 2")\right) +$$

$$\frac{11}{30} \times Entropy("> 2")$$

$$Gain_{(juml_tangg_kel)} = 0.88129 - \left(\frac{19}{30} \times (0.62922)\right) + \frac{11}{30} \times (0.99402)$$

$$Gain_{(juml_tangg_kel)} = 0.11831$$

6. Atribut *status_tempat_tinggal*

Untuk atribut *status_tempat_tinggal* pada tabel IV-2 terdapat 30 kasus dengan 2 nilai 'sewa'; 21 nilai 'milik_sendiri', 6 nilai 'milik_keluarga', dan 1 nilai 'dinas', dengan perincian untuk nilai kategori 'sewa' semuanya "disetujui", untuk nilai kategori 'milik_sendiri' 16 'disetujui' dan 5 'ditolak', untuk nilai kategori 'milik_keluarga' 3 'disetujui' dan 3 'ditolak', dan untuk nilai kategori 'dinas', 0 'disetujui' dan 1 'ditolak', sehingga nilai entropi *status_tempat_tinggal* dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy('sewa') = \left(-\frac{2}{2} \log_2 \left(\frac{2}{2}\right)\right) + \left(-\frac{0}{2} \log_2 \left(\frac{0}{2}\right)\right) = 0$$

$$Entropy('milik_sendiri') = \left(-\frac{16}{21} \log_2 \left(\frac{16}{21}\right)\right) + \left(-\frac{5}{21} \log_2 \left(\frac{5}{21}\right)\right) = 0.79185$$

$$Entropy('milik_keluarga') = \left(-\frac{3}{6} \log_2 \left(\frac{3}{6}\right)\right) + \left(-\frac{3}{6} \log_2 \left(\frac{3}{6}\right)\right) = 1$$

$$Entropy('dinas') = \left(-\frac{0}{1} \log_2 \left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} \log_2 \left(\frac{1}{1}\right)\right) = 0$$

Kemudian menghitung nilai *gain* *status_tempat_tinggal* sebagai berikut:

$$Gain_{(status_temp_tinggal)} = Entropy\ Total - \left(\frac{2}{30} \times Entropy('sewa')\right) +$$

$$\begin{aligned} & \frac{21}{30} x \text{Entropy}('milik_sendiri') + \\ & \frac{6}{30} x \text{Entropy}('milik_keluarga') \\ & + \frac{1}{30} x \text{Entropy}('dinas')) \end{aligned}$$

$$\begin{aligned} \text{Gain}_{(\text{status_temt_tinggal})} &= 0.88129 - \left(\frac{2}{30} x (0) + \frac{21}{30} x (0.79185) + \right. \\ & \left. \left(\frac{6}{30} x (1) + \frac{1}{30} x (0) \right) \right) \end{aligned}$$

$$\text{Gain}_{(\text{status_temt_tinggal})} = 0.12699$$

7. Atribut pekerjaan

Untuk atribut pekerjaan pada tabel 4.2 terdapat 30 kasus dengan 15 nilai kategori ‘wiraswasta’; dan 3 nilai kategori ‘PNS’; 9 nilai kategori ‘petani’; 1 nilai kategori ‘TNI/POLRI’, 1 nilai kategori ‘guru/dosen’, dan 1 nilai kategori ‘Lain-lain’, dengan perincian untuk nilai kategori ‘wiraswasta’, 11 “disetujui” dan 4 “ditolak”; untuk nilai kategori ‘PNS’ 2 “disetujui” dan 1 “ditolak”; untuk nilai kategori ‘petani’, 7 “disetujui” dan 2 “ditolak”; untuk nilai kategori ‘TNI/POLRI’ semuanya “ditolak”, untuk nilai kategori ‘guru/dosen’ semuanya “diterima”, untuk nilai kategori ‘lain-lain’ semuanya “ditolak”, sehingga nilai entropi pekerjaan dapat dihitung untuk setiap kasus sebagai berikut:

$$\text{Entropy}('wiraswasta') = \left(-\frac{11}{15} \log_2 \left(\frac{11}{15} \right) \right) + \left(-\frac{4}{15} \log_2 \left(\frac{4}{15} \right) \right) = 0.83668$$

$$\text{Entropy}('PNS') = \left(-\frac{2}{3} \log_2 \left(\frac{2}{3} \right) \right) + \left(-\frac{1}{3} \log_2 \left(\frac{1}{3} \right) \right) = 0.91825$$

$$\text{Entropy}('petani') = \left(-\frac{7}{9} \log_2 \left(\frac{7}{9} \right) \right) + \left(-\frac{2}{9} \log_2 \left(\frac{2}{9} \right) \right) = 0.76412$$

$$\text{Entropy}('tni_polri') = \left(-\frac{0}{1} \log_2 \left(\frac{0}{1} \right) \right) + \left(-\frac{1}{1} \log_2 \left(\frac{1}{1} \right) \right) = 0$$

$$Entropy('guru') = \left(-\frac{1}{1} \log_2 \left(\frac{1}{1}\right)\right) + \left(-\frac{0}{1} \log_2 \left(\frac{0}{1}\right)\right) = 0$$

$$Entropy('lain_lain') = \left(-\frac{0}{1} \log_2 \left(\frac{0}{1}\right)\right) + \left(-\frac{1}{1} \log_2 \left(\frac{1}{1}\right)\right) = 0$$

Kemudian menghitung nilai *gain* pekerjaan sebagai berikut:

$$\begin{aligned} Gain_{(pekerjaan)} = Entropy Total - & \left(\frac{15}{30} \times Entropy('wiraswasta') + \right. \\ & \frac{3}{30} \times Entropy('pns') + \frac{9}{30} \times Entropy('petani') + \\ & \frac{1}{30} \times Entropy('tni_polri') + \frac{1}{30} \times Entropy('guru') \\ & \left. + \frac{1}{20} \times Entropy('lain_lain')) \right) \end{aligned}$$

$$\begin{aligned} Gain_{(pekerjaan)} = 0.88129 - & \left(\frac{15}{30} \times (0.83668) + \frac{3}{30} \times (0.91825) + \right. \\ & \left(\frac{9}{30} \times (0.76412) + \left(\frac{1}{30} \times (0) + \left(\frac{1}{30} \times (0) \right. \right. \right. \\ & \left. \left. \left. + \left(\frac{1}{30} \times (0)\right)\right)\right) \right) \end{aligned}$$

$$Gain_{(pekerjaan)} = 0.11418$$

8. Atribut penghasilan_per_tahun

Untuk atribut penghasilan_per_tahun pada tabel IV-2 terdapat 30 kasus dengan 16 nilai kategori '<=45 juta'; dan 14 nilai kategori '>45 juta', dengan perincian untuk nilai kategori '<=45 juta' 12 'disetujui' dan 4 'ditolak', untuk nilai kategori '>45 juta' 9 'disetujui' dan 5 'ditolak', sehingga nilai *entropi* penghasilan_per_tahun dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy("<=45") = \left(-\frac{12}{16} \log_2 \left(\frac{12}{16}\right)\right) + \left(-\frac{4}{16} \log_2 \left(\frac{4}{16}\right)\right) = 0.81128$$

$$Entropy(">45") = \left(-\frac{9}{14} \log_2 \left(\frac{9}{14}\right)\right) + \left(-\frac{5}{14} \log_2 \left(\frac{5}{14}\right)\right) = 0.94037$$

Kemudian menghitung nilai *gain* penghasilan_per_tahun sebagai berikut:

$$Gain_{(\text{penghasilan_per_tahun})} = Entropy\ Total - \left(\frac{16}{30} \times Entropy(" \leq 45")\right) + \frac{14}{30} \times Entropy(" > 45")$$

$$Gain_{(\text{penghasilan_per_tahun})} = 0.88129 - \left(\frac{16}{30} \times (0.81128)\right) + \frac{14}{30} \times (0.94037)$$

$$Gain_{\text{penghasilan_per_tahun}} = 0.00977$$

9. Atribut nilai_jaminan

Untuk atribut nilai_jaminan pada tabel 4.2 terdapat 30 kasus dengan 17 nilai kategori '<=100 juta'; dan 13 nilai kategori '>100 juta', dengan perincian untuk nilai kategori '<=100' 11 'disetujui' dan 6 'ditolak', untuk nilai kategori '>100' 10 'disetujui' dan 3 'ditolak', sehingga nilai *entropi* nilai_jaminan dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy (" \leq 100") = \left(-\frac{11}{17} \log_2 \left(\frac{11}{17}\right)\right) + \left(-\frac{6}{17} \log_2 \left(\frac{6}{17}\right)\right) = 0.9367$$

$$Entropy (" > 100") = \left(-\frac{10}{13} \log_2 \left(\frac{10}{13}\right)\right) + \left(-\frac{3}{13} \log_2 \left(\frac{3}{13}\right)\right) = 0.7794$$

Kemudian menghitung nilai *gain* nilai_jaminan sebagai berikut:

$$Gain_{(\text{nilai_jaminan})} = Entropy\ Total - \left(\frac{17}{30} \times Entropy(" \leq 100")\right) + \frac{13}{30} \times Entropy(" > 100")$$

$$Gain_{(\text{nilai_jaminan})} = 0.88129 - \left(\frac{17}{30} \times (0.9367)\right) + \frac{13}{30} \times (0.7794)$$

$$Gain_{(\text{nilai_jaminan})} = 0.01276$$

10. Atribut tujuan_pinjaman

Untuk atribut tujuan_pinjaman pada tabel 4.2 terdapat 30 kasus dengan 6 nilai kategori 'modal_usaha'; 24 nilai kategori 'kendaraan', dengan perincian untuk nilai kategori 'modal_usaha, 2 "disetujui" dan 4 "ditolak"; untuk nilai kategori 'kendaraan', 19 'disetujui' dan 5 "ditolak, sehingga nilai *entropi* tujuan_pinjaman dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy ('modal_usaha') = \left(-\frac{2}{6}\log_2\left(\frac{2}{6}\right)\right) + \left(-\frac{4}{6}\log_2\left(\frac{4}{6}\right)\right) = 0.91825$$

$$Entropy ('kendaraan') = \left(-\frac{19}{24}\log_2\left(\frac{19}{24}\right)\right) + \left(-\frac{5}{24}\log_2\left(\frac{5}{24}\right)\right) = 0.73831$$

Kemudian menghitung nilai gain tujuan_pinjaman sebagai berikut:

$$Gain_{(tujuan_pinjaman)} = Entropy\ Total - \left(\frac{6}{30} \times Entropy('modal_usaha') + \frac{24}{30} \times Entropy('kendaraan')\right)$$

$$Gain_{(tujuan_pinjaman)} = 0.88129 - \left(\frac{6}{30} \times (0.91825) + \frac{24}{30} \times (0.73831)\right)$$

$$Gain_{(tujuan_pinjaman)} = 0.10699$$

11. Atribut pinjaman_max

Untuk atribut pinjaman_max pada tabel 4.2 terdapat 30 kasus dengan 9 nilai kategori '<=60 juta'; dan 21 nilai kategori '>60 juta', dengan perincian untuk nilai kategori '<=60 juta' 6 'disetujui' dan 3 'ditolak', untuk nilai kategori '>60 juta' 15 'disetujui' dan 6 'ditolak', sehingga nilai *entropi* pinjaman_max dapat dihitung untuk setiap kasus sebagai berikut:

$$Entropy ("≤60") = \left(-\frac{6}{9}\log_2\left(\frac{6}{9}\right)\right) + \left(-\frac{3}{9}\log_2\left(\frac{3}{9}\right)\right) = 0.91825$$

$$Entropy (">60") = \left(-\frac{15}{21}\log_2\left(\frac{15}{21}\right)\right) + \left(-\frac{6}{21}\log_2\left(\frac{6}{21}\right)\right) = 0.86311$$

Kemudian menghitung nilai *gain* pinjaman_max sebagai berikut:

$$Gain_{(\text{pinjaman_max})} = Entropy\ Total - \left(\frac{9}{30} \times Entropy("<= 60")\right) + \frac{21}{30} \times Entropy("> 60")$$

$$Gain_{(\text{pinjaman_max})} = 0.88129 - \left(\frac{9}{30} \times (0.91825)\right) + \frac{21}{30} \times (0.86311)$$

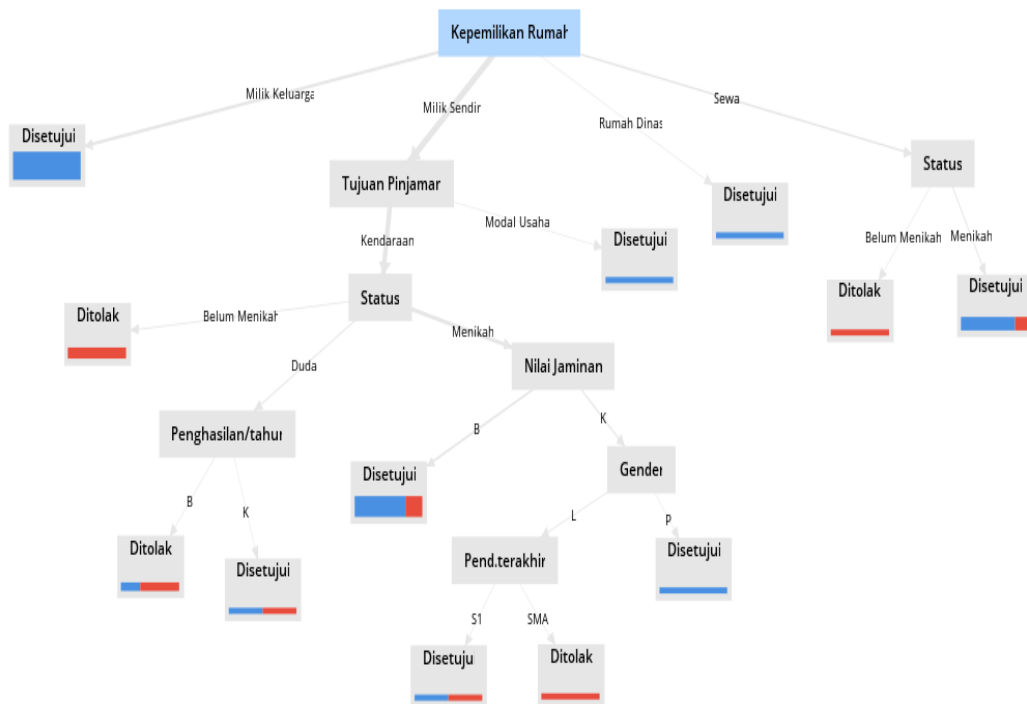
$$Gain_{\text{pinjaman_max}} = 0.00161$$

Dari hasil perhitungan total *entropi* dan *gain* setiap atribut diatas dapat dibuat dalam bentuk tabel 4.3 berikut ini:

Tabel 4.3 Hasil perhitungan *entropi* dan *gain* dari setiap atribut sampel

Atribut	Range	Jumlah kasus	Disetujui	Ditolak	Entropy	Gain
Total		30	21	9	0.88129	
Age						0.10026
	<=40	15	13	2	0.56536	
	>40	15	8	7	0.99672	
gender						0.01965
	L	26	19	7	0.84041	
	P	4	2	2	1	
Pend_terakhir						0.12696
	S2	1	0	1	0	
	S1	1	1	0	0	
	SMA	25	19	6	0.79503	
	SMP	3	1	2	0.91825	
status						0.07157
	belum menikah	2	1	1	1	
	menikah	27	20	7	0.82557	
	duda	1	0	1	0	
juml_tangg_kel						0.11831
	<=2	19	16	3	0.62922	
	>2	11	5	6	0.99402	
status_tempat_tinggal						0.12699
	sewa	2	2	0	0	
	milik sendiri	21	16	5	0.79185	
	milik keluarga	6	3	3	1	
	dinas	1	0	1	0	
pekerjaan						0.11418
	wiraswasta	15	11	4	0.83668	
	pns	3	2	1	0.91825	
	petani	9	7	2	0.76412	
	tni/polri	1	0	1	0	
	guru/dosen	1	1	0	0	
	lain-lain	1	0	1	0	
penghasilan_per_tahun						0.00977
	<=45	16	12	4	0.81128	
	>45	14	9	5	0.94037	
nilai_jaminan						0.01276
	<=100 juta	17	11	6	0.93671	
	>100 juta	13	10	3	0.7794	
tujuan_pinjaman						0.10699
	modal usaha	6	2	4	0.91825	
	kendaraan	24	19	5	0.73831	
pinjaman_max						0.00161
	<=60 juta	9.00	6	3	0.91825	
	>60 juta	21.00	15	6	0.86311	

Dari tabel 4.3 diperoleh bahwa atribut dengan nilai *gain* tertinggi adalah *status_tempat tinggal* dengan nilai 0,12699, maka *status_tempat tinggal* menjadi *node* akar. Atribut *status_tempat tinggal* terdiri dari 4 kategori nilai yaitu “Milik Keluarga”, “Milik Sendiri”, “Rumah Dinas”, dan “Sewa”, oleh karena itu diagram pohon yang terbentuk memiliki empat cabang. Diagram pohon keputusan setelah dicabangkan dengan *node status_tempat tinggal* digambarkan pada gambar 4.2



Gambar 4.2 Hasil Pohon Keputusan Menggunakan RapidMiner

Berdasarkan gambar 4.2 diatas, semua kasus sudah masuk pada kelasnya, sehingga pohon keputusan diatas merupakan pohon keputusan terakhir yang terbentuk dan diperoleh juga *rule-rule* untuk klasifikasi nasabah yang Diterima(A) dan Ditolak(R) dibawah ini yaitu:

IF *status_tinggal* == 'Milik Sendiri' THEN

 IF *tujuan_pinjam* == 'Kendaraan' THEN

 IF *marital* == 'Duda' THEN

 IF *penghasilan_tahun* == 'K' THEN

```

        pred_hasil = "Disetujui(A)"
    ELSEIF penghasilan_tahun == 'B' THEN
        pred_hasil = "Ditolak(R)"
    ENDIF
ENDIF
ENDIF
ENDIF
ELSEIF status_tinggal== 'Milik Sendiri' THEN
    IF tujuan_pinjam== 'Kendaraan' THEN
        IF marital == 'Menikah' THEN
            IF nilai_jaminan == 'K' THEN
                IF gender == 'L' THEN
                    IF pendidikan == 'S1' THEN
                        pred_hasil = "Disetujui(A)"
                    ELSEIF pendidikan == 'SMA' THEN
                        pred_hasil = "Ditolak(R)"
                    ENDIF
                ENDIF
            ENDIF
        ENDIF
    ENDIF
ENDIF
ENDIF
ENDIF
ELSEIF status_tinggal== 'Milik Sendiri' THEN
    IF tujuan_pinjam== 'Kendaraan' THEN
        IF marital == 'Menikah' THEN
            IF nilai_jaminan == 'K' THEN
                IF gender == 'P' THEN
                    pred_hasil = "Disetujui(A)"
                ELSEIF nilai_jaminan == 'B' THEN
                    pred_hasil = "Disetujui(A)"
                ENDIF
            ENDIF
        ENDIF
    ENDIF
ENDIF
ENDIF

```

```

        ENDIF
    ENDIF
ENDIF
ELSE IF status_tinggal== 'Milik Sendiri' THEN
    IF tujuan_pinjam== 'Kendaraan' THEN
        IF marital == 'Menikah' THEN
            IF nilai_jaminan == 'B' THEN
                pred_hasil = "Disetujui(A)"
            ENDIF
        ENDIF
    ENDIF
ENDIF
ENDIF
ENDIF
ELSE IF status_tinggal== 'Milik Sendiri' THEN
    IF tujuan_pinjam== 'Modal Usaha' THEN
        pred_hasil = "Disetujui(A)"
    ELSE IF status_tinggal== 'Milik Sendiri' THEN
        IF tujuan_pinjam== 'Kendaraan' THEN
            IF marital == 'Belum Menikah' THEN
                pred_hasil = "Ditolak(R)"
            ENDIF
        ENDIF
    ENDIF
ENDIF
ELSE IF status_tinggal== 'Rumah Dinas' THEN
    pred_hasil = "Disetujui(A)"
ELSE IF status_tinggal== 'Sewa' THEN
    IF marital == 'Menikah' THEN
        pred_hasil = "Disetujui(A)"
    ELSEIF marital == 'Belum Menikah' THEN
        pred_hasil = "Ditolak(R)"
    ELSE IF status_tinggal== 'Milik Keluarga' THEN
        pred_hasil = "Disetujui(A)" ENDIF

```

4.3.2 Pengujian Klasifikasi dengan Algoritma *Naive Bayes*

Penggunaan algoritma *Naive Bayes* menggunakan *data training* pada Tabel 4.2 dimulai dengan melakukan perhitungan *probabilitas prior* untuk mengetahui nilai yang disetujui dan ditolak untuk semua *record*. Pada *data training* jumlah *record* sebanyak 30, dimana nasabah yang disetujui sebanyak 21 (dua puluh satu) dan yang ditolak sebanyak 9 (sembilan). Berikut hasil perhitungan *prior probability* terlihat pada tabel dibawah.

Tabel 4.4 Perhitungan *Probabilitas Prior*

Atribut	Range	Jumlah kasus	Disetujui	Ditolak	p(x c1)	
					Disetujui	Ditolak
Total		30	21	9	0.7	0.3
Age						
	<=40	15	13	2	0.43333	0.06667
	>40	15	8	7	0.26667	0.23333
gender						
	L	26	19	7	0.63333	0.23333
	P	4	2	2	0.06667	0.06667
Pend_terakhir						
	S2	1	0	1	0.00000	0.03333
	S1	1	1	0	0.03333	0.00000
	SMA	25	19	6	0.63333	0.20000
	SMP	3	1	2	0.03333	0.06667
status						
	belum menikah	2	1	1	0.03333	0.03333
	menikah	27	20	7	0.66667	0.23333
	duda	1	0	1	0.00000	0.03333
juml_tangg_kel						
	<=2	19	16	3	0.53333	0.10000
	>2	11	5	6	0.16667	0.20000
status_tempat_tinggal						
	sewa	2	2	0	0.06667	0.00000
	milik sendiri	21	16	5	0.53333	0.16667
	milik keluarga	6	3	3	0.10000	0.10000
	dinas	1	0	1	0.00000	0.03333
pekerjaan						
	wiraswasta	15	11	4	0.36667	0.13333
	pns	3	2	1	0.06667	0.03333
	petani	9	7	2	0.23333	0.06667
	tni/polri	1	0	1	0.00000	0.03333
	guru/dosen	1	1	0	0.03333	0.00000
	lain-lain	1	0	1	0.00000	0.03333
penghasilan_per_tahun						
	<=45	16	12	4	0.40000	0.13333
	>45	14	9	5	0.30000	0.16667
nilai_jaminan						
	<=100 juta	17	11	6	0.36667	0.20000
	>100 juta	13	10	3	0.33333	0.10000
tujuan_pinjaman						
	modal usaha	6	2	4	0.06667	0.13333
	kendaraan	24	19	5	0.63333	0.16667
pinjaman_max						
	<=60 juta	9.00	6	3	0.20000	0.10000
	>60 juta	21.00	15	6	0.50000	0.20000

Untuk menentukan kasus baru termasuk disetujui atau ditolak, dilakukan perhitungan *probabilitas posterior* berdasarkan *probabilitas prior* yang telah dihitung sebelumnya pada Tabel 4.4. Perhitungan *probabilitas posterior* untuk menentukan *data testing* termasuk klasifikasi mana yang diterima atau ditolak terdapat disajikan pada Tabel 4.5. Misalkan diambil sebuah *data testing* X dengan nilai seperti pada Tabel 4.5 kolom dua, untuk menentukan kelas mana, dilakukan perhitungan *probabilitas posterior* yang hasilnya terdapat pada Tabel 4.5 kolom tiga dan empat.

Tabel 4.5 *Probabilitas Posterior Data*

Data X	Atribut	Nilai	p(X Ci)	
			DISETUJUI	DITOLAK
Age		32	0.43333	0.06667
Gender		L	0.63333	0.23333
Pend. Terakhir		SMA	0.63333	0.20000
Status		Menikah	0.66667	0.23333
Jml Tangg Kel.		1	0.53333	0.10000
Status Tempat Tinggal		Milik Sendiri	0.53333	0.16667
Pekerjaan		Wiraswasta	0.36667	0.13333
Penghasilan_per_tahun		36,295,000	0.40000	0.13333
Nilai Jaminan		80,000,000	0.36667	0.20000
Tujuan Pinjaman		Kendaraan	0.63333	0.16667
Pinjaman_max		59,100,000	0.20000	0.10000

Dari tabel diatas terdapat beberapa langkah untuk menghitung, yaitu:

- a. $P(X|C_i)$ = P(X| remark = Disetujui)

$$= 0.43333 \times 0.63333 \times 0.63333 \times 0.66667 \times 0.53333$$

$$\times 0.53333 \times 0.36667 \times 0.40000 \times 0.36667 \times$$

$$0.63333 \times 0.20000$$

$$= 0,000224$$
- b. $P(X|C_i)$ = P(X| remark = Ditolak)

$$= 0.06667 \times 0.23333 \times 0.20000 \times 0.23333 \times 0.10000$$

$$\times 0.16667 \times 0.13333 \times 0.13333 \times 0.20000 \times$$

$$0.16667 \times 0.10000$$

$$= 0,0000000007169$$

- c. $P(X|C_i) P(C_i) = P(X | \text{remark} = \text{Disetujui}) P(\text{remark} = \text{Disetujui})$
 $= 0,7 \times 0,000224$
 $= \mathbf{0.000157}$
- d. $P(X|C_i) P(C_i) = P(X | \text{remark} = \text{Ditolak}) P(\text{remark} = \text{Ditolak})$
 $= 0,3 \times 0,0000000007169$
 $= 0,000000000215$

Dari hasil perhitungan di atas, didapat nilai $P(X|C_i)$ dan $P(X|C_i) P(C_i)$ lebih besar untuk *remark* = Disetujui sehingga dapat disimpulkan bahwa *data testing* tersebut termasuk klasifikasi “**Disetujui**”.

4.3.3 Confusion Matrix, F-measure dan Kurva ROC Decision Tree

Pengujian dilakukan dengan *confusion matrix* yang terdiri dari *accuracy*, *precision* dan *recall* dilakukan pada dataset sebanyak 300 data yang diolah dengan menggunakan *Decision Tree*. Pengujian *confusion matrix* untuk dataset yang diolah menggunakan *Decision Tree* dapat dilihat pada tabel dibawah ini.

Tabel 4.6 *Confusion Matrix* dengan *Decision Tree*

Dataset			
	A	R	Precision
Pred A	235	38	86.08%
Pred R	4	23	85.19%
Recall	98.33%	37.70%	

Nilai akurasi dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned} \text{akurasi} &= \frac{(TN + TP)}{(TN + FN + FP + TP)} \\ &= \frac{(23 + 235)}{(23 + 4 + 38 + 235)} \\ &= \frac{258}{300} = 0,8600 \end{aligned}$$

$$\mathbf{\text{akurasi} = 86,00\%}$$

Nilai *precision* dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned} \text{precision} &= \frac{TP}{(TP + FP)} \\ &= \frac{235}{(235 + 38)} \end{aligned}$$

$$= \frac{235}{273} = 0,8608$$

$$\mathbf{precision = 86,08\%}$$

Nilai *recall* dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned} recall &= \frac{TP}{(TP + FN)} \\ &= \frac{235}{(235 + 4)} \\ &= \frac{235}{239} = 0,9833 \end{aligned}$$

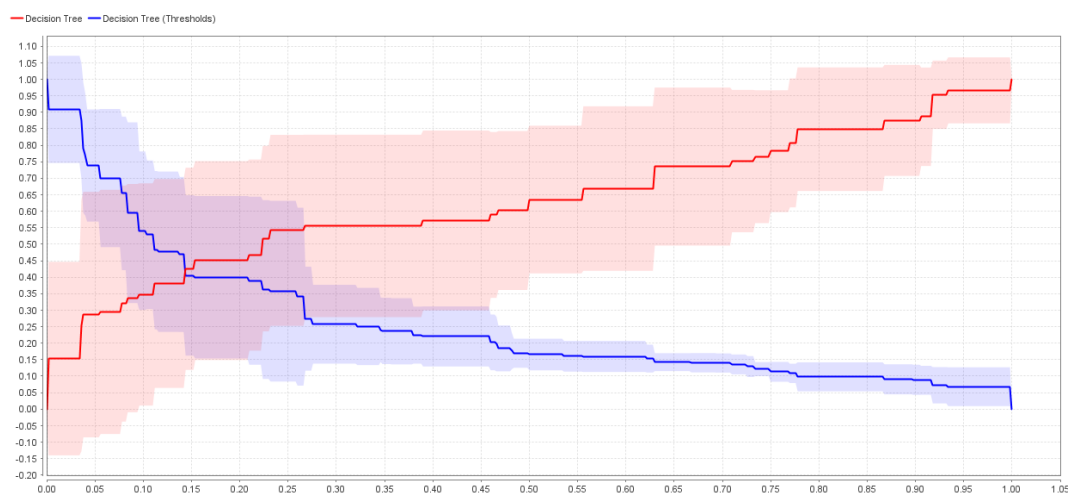
$$\mathbf{recall = 98,33\%}$$

Nilai *F-measure* dapat dihitung dengan persamaan berikut:

$$\begin{aligned} f_measure &= \frac{2 \times recall \times precision}{(recall + precision)} \\ &= \frac{1,6928}{1,8441} = 0,9179 \end{aligned}$$

$$\mathbf{f_measure = 91,79\%}$$

Kurva ROC untuk *Decision Tree* sebagai berikut:



Gambar 4.3 Nilai AUC dengan *Decision Tree*

4.3.4 Confusion Matrix, F-measure dan Kurva ROC Algoritma Naïve Bayes

Pengujian *confusion matrix* untuk *dataset* yang diolah menggunakan algoritma *Naïve Bayes* dapat dilihat pada tabel dibawah ini.

Tabel 4.7 *Confusion Matrix* dengan algoritma *Naïve Bayes*

dataset			
	A	R	Precision
Pred A	234	53	81.53%
Pred R	5	8	61.54%
Recall	97.91%	13.11%	

Nilai akurasi dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned}
 \text{akurasi} &= \frac{(\text{TN} + \text{TP})}{(\text{TN} + \text{FN} + \text{FP} + \text{TP})} \\
 &= \frac{(8 + 234)}{(8 + 5 + 53 + 234)} \\
 &= \frac{242}{300} = 0,8067
 \end{aligned}$$

$$\text{akurasi} = \mathbf{80,67\%}$$

Nilai *precision* dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned}
 \text{precision} &= \frac{\text{TP}}{(\text{TP} + \text{FP})} \\
 &= \frac{234}{(234 + 53)} \\
 &= \frac{234}{287} = 0,81533
 \end{aligned}$$

$$\text{precision} = \mathbf{81,53\%}$$

Nilai *recall* dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned}
 \text{recall} &= \frac{\text{TP}}{(\text{TP} + \text{FN})} \\
 &= \frac{234}{(234 + 5)}
 \end{aligned}$$

$$= \frac{234}{239} = 0,97907$$

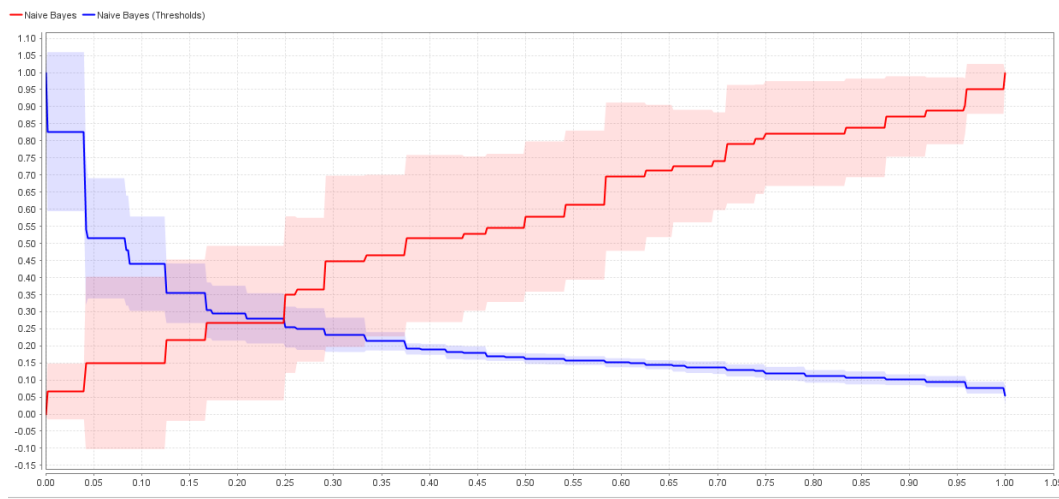
$$\mathbf{recall = 97,91\%}$$

Nilai *F-measure* dapat dihitung dengan persamaan berikut:

$$\begin{aligned} f_measure &= \frac{2 \times recall \times precision}{(recall + precision)} \\ &= \frac{2 \times (0,97907) \times (0,81533)}{(0,97907 + 0,81533)} \\ &= \frac{1,59653}{1,79440} = 0,88973 \end{aligned}$$

$$\mathbf{f_measure = 88,97\%}$$

Kurva ROC untuk algoritma *Naïve Bayes* sebagai berikut:



Gambar 4.4 Nilai AUC dengan Algoritma *Naïve Bayes*

4.4 Evaluasi dan Validasi

4.4.1 Evaluasi *Confusion Matrix*, *F-measure* dan Kurva ROC *Decision Tree*

Pengujian dilakukan dengan *confusion matrix* yang terdiri dari akurasi (*accuracy*), *precision* dan *recall* dilakukan pada 45 *data testing* yang diolah dengan menggunakan *Decision Tree*. Pengujian *confusion matrix* untuk *data testing* yang diolah menggunakan *Decision Tree* dapat dilihat pada tabel dibawah ini.

Tabel 4.8 *Confusion Matrix* dengan *Decision Tree*

Dataset			
	A	R	Precision
Pred A	35	6	85.37%
Pred R	1	3	75.00%
Recall	97.22%	33.33%	

Nilai akurasi dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned}
 \text{akurasi} &= \frac{(\text{TN} + \text{TP})}{(\text{TN} + \text{FN} + \text{FP} + \text{TP})} \\
 &= \frac{(3 + 35)}{(3 + 1 + 6 + 35)} \\
 &= \frac{38}{45} = 0,8444
 \end{aligned}$$

$$\text{akurasi} = \mathbf{84,44\%}$$

Nilai *precision* dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned}
 \text{precision} &= \frac{\text{TP}}{(\text{TP} + \text{FP})} \\
 &= \frac{35}{(35 + 6)} \\
 &= \frac{35}{41} = 0,8537
 \end{aligned}$$

$$\text{precision} = \mathbf{85.37\%}$$

Nilai *recall* dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned}
 \text{recall} &= \frac{\text{TP}}{(\text{TP} + \text{FN})} \\
 &= \frac{35}{(35 + 1)} \\
 &= \frac{35}{36} = 0,9722
 \end{aligned}$$

$$\text{recall} = \mathbf{97.22\%}$$

Nilai *F-measure* dapat dihitung dengan persamaan berikut:

$$\begin{aligned}
 f_measure &= \frac{2 \times recall \times precision}{(recall + precision)} \\
 &= \frac{2 \times (0,9722) \times (0,8537)}{(0,9722 + 0,8537)} \\
 &= \frac{1,6599}{1,8259} = 0,9091
 \end{aligned}$$

$$f_measure = 90.91\%$$

4.4.2 Evaluasi Confusion Matrix, F-measure dan Kurva ROC Naïve Bayes

Pengujian *confusion matrix* untuk *dataset* yang diolah menggunakan algoritma *Naive Bayes* dengan 45 *data testing* dapat dilihat pada tabel dibawah ini.

Tabel 4.9 *Confusion Matrix* dengan algoritma *Naive Bayes*

dataset			
	A	R	Precision
Pred A	30	9	76.92%
Pred R	6	0	0.00%
Recall	83.33%	0.00%	

Nilai akurasi dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned}
 \text{akurasi} &= \frac{(TN + TP)}{(TN + FN + FP + TP)} \\
 &= \frac{(0 + 30)}{(0 + 6 + 9 + 30)} \\
 &= \frac{30}{45} = 0,6667
 \end{aligned}$$

$$\text{akurasi} = 66.67\%$$

Nilai *precision* dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned}
 \text{precision} &= \frac{TP}{(TP + FP)} \\
 &= \frac{30}{(30 + 9)}
 \end{aligned}$$

$$= \frac{30}{39} = 0.7692$$

$$\mathbf{precision = 76.92\%}$$

Nilai *recall* dari *confusion matrix* tersebut adalah sebagai berikut:

$$\begin{aligned} recall &= \frac{TP}{(TP + FN)} \\ &= \frac{30}{(30 + 6)} \\ &= \frac{30}{36} = 0.8333 \end{aligned}$$

$$\mathbf{recall = 83.33\%}$$

Nilai *F-measure* dapat dihitung dengan persamaan berikut:

$$\begin{aligned} f_measure &= \frac{2 \times recall \times precision}{(recall + precision)} \\ &= \frac{2 \times (0,8333) \times (0,7692)}{(0,8333 + 0,7692)} \\ &= \frac{1,2819}{1,6025} = 0,7999 \end{aligned}$$

$$\mathbf{f_measure = 79,99\%}$$

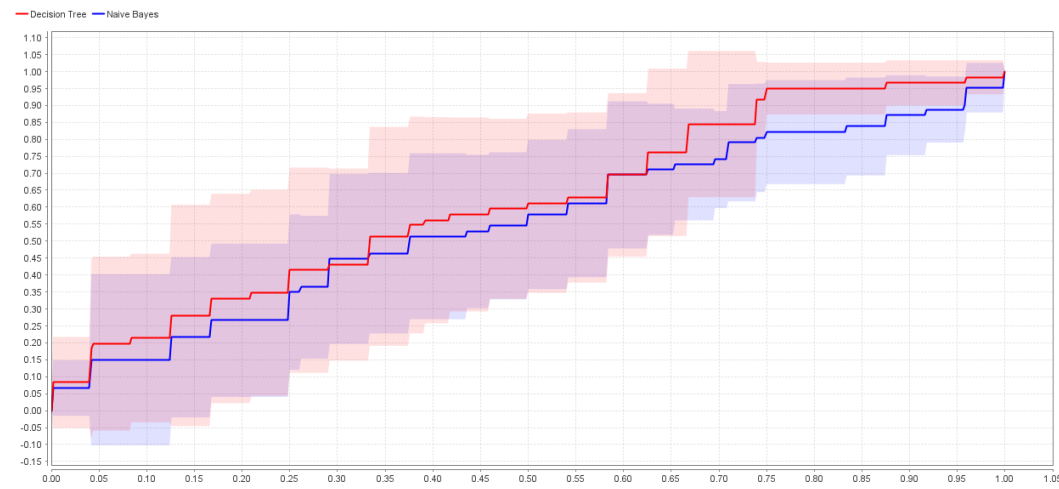
4.5 Hasil Komparasi

Model dengan algoritma *Decision Tree* dan *Naïve Bayes* untuk prediksi kelayakan pemberian kredit kepada nasabah yang diuji tingkat akurasi menghasilkan perbandingan nilai akurasi (*accuracy*) dan nilai AUC. Dengan dataset nasabah sebagai data uji. Algoritma *Decision Tree* mendapatkan akurasi yang paling tinggi yaitu sebesar 85,22%. Sedangkan algoritma *Naïve Bayes* mendapatkan akurasi sebesar 73,67%. Seperti yang ditunjukkan pada tabel 4.10.

Tabel 4.10 Perbandingan Akurasi dan AUC Metode *Decision Tree* dan *Naïve Bayes*

Metode Data Mining	Akurasi		AUC		Perbandingan Nilai Akurasi	Perbandingan Nilai AUC
	Training	Testing	Training	Testing		
<i>Decision Tree</i>	86,00%	84,44%	93,80%	97,40%	85,22%	95,60%
<i>Naïve Bayes</i>	80,67%	66,67%	91,20%	87,90%	73,67%	89,55%

Kurva ROC untuk perbandingan dua metode algoritma *Decision Tree* dan *Naïve Bayes* sebagai berikut:



Gambar 4.5 Nilai Perbandingan AUC

Hasil perhitungan divisualisasikan dengan kurva ROC (*Receiver Operating Characteristic*) atau AUC (*Area Under Curve*). ROC memiliki tingkat nilai diagnosa yaitu [18]:

- Akurasi bernilai $0.90 - 1.00 = \textit{excellent classification}$
- Akurasi bernilai $0.80 - 0.90 = \textit{good classification}$
- Akurasi bernilai $0.70 - 0.80 = \textit{fair classification}$
- Akurasi bernilai $0.60 - 0.70 = \textit{poor classification}$
- Akurasi bernilai $0.50 - 0.60 = \textit{failure}$

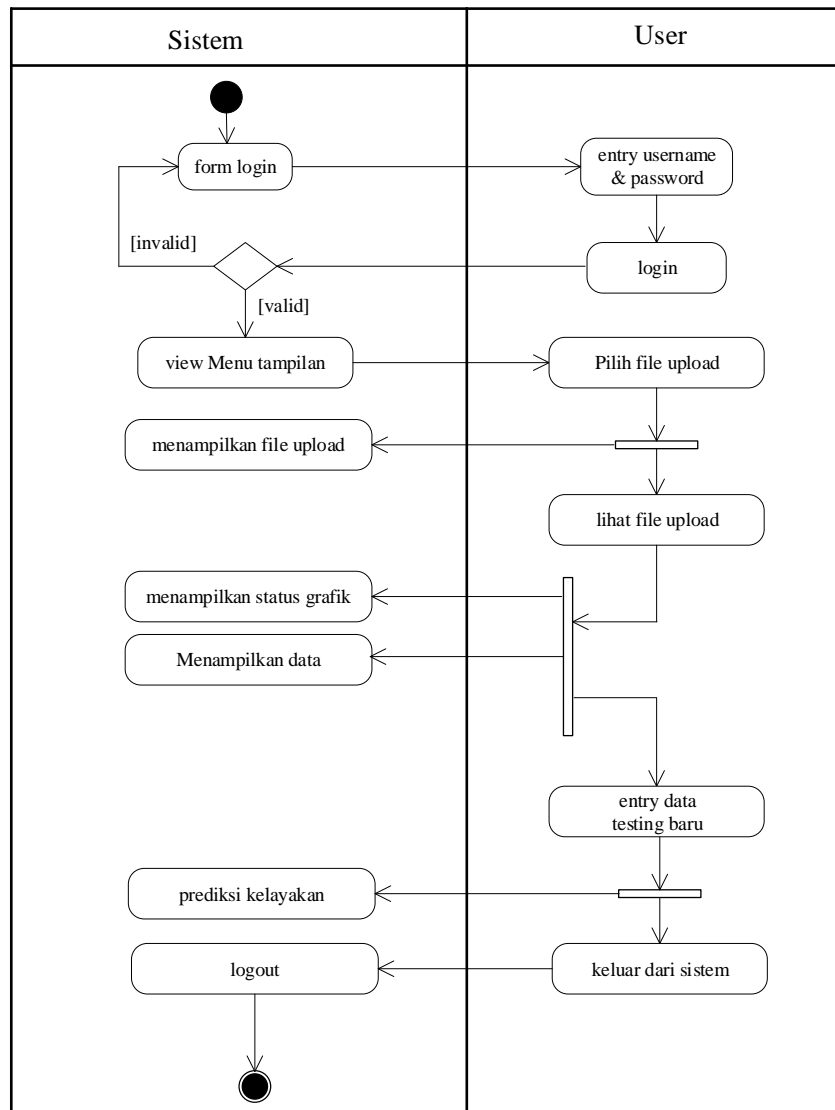
Hasil yang didapat dari pengolahan ROC menggunakan data training untuk algoritma *decision tree* sebesar 0.9560 dengan tingkat diagnosa *excellent classification*, sedangkan algoritma *Naïve Bayes* sebesar 0.8955 dengan tingkat diagnosa *good classification*.

4.6 Penerapan Algoritma Terpilih

Dari hasil evaluasi dan validasi diatas dapat diketahui bahwa metode *Decision Tree* memiliki tingkat akurasi dan *perfoma* yang baik, sehingga *rule* yang dihasilkan oleh metode *Decision Tree* dapat dijadikan sebagai *rule* untuk

pembuatan *prototype* yang dapat memudahkan dalam memprediksi pemberian pinjaman kembali kepada nasabah.

Berikut adalah alur dari *Activity Diagram*, *Use Case*, *Class Diagram* dan *Flow Chart* yang diimplementasikan kedalam *prototype* Prediksi Kelayakan Pemberian Kredit kepada nasabah.

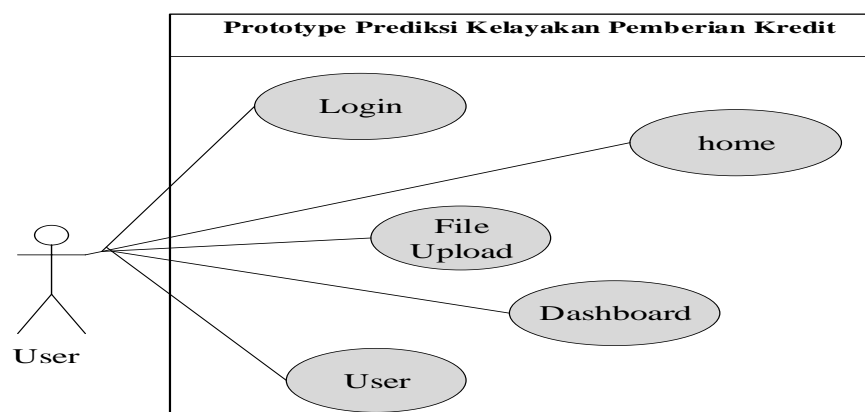


Gambar 4.6 Activity Diagram Login

User membuka aplikasi dengan memasukkan *login* pada sistem dan akan menampilkan *menu* tampilan dengan menampilkan pilihan untuk *upload file* yang dipilih *user* pada saat yang sama sistem juga dapat menampilkan data dari *file* yang di *upload* yang dipilih *user*, selanjutnya sistem akan *import upload file* dan

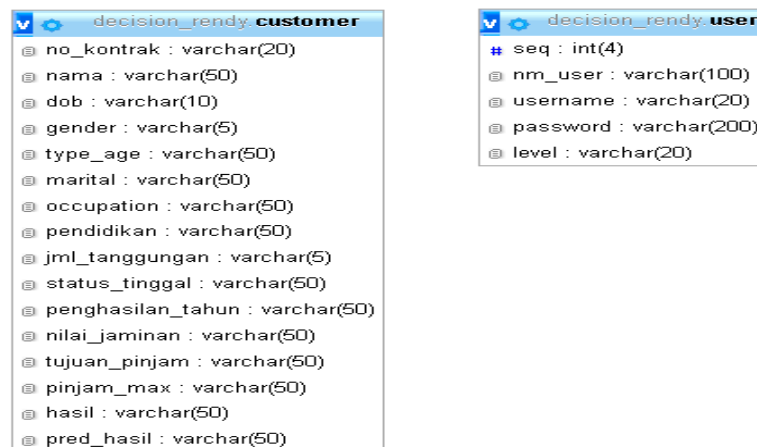
memberikan *summary* dari data *upload*. *User* akan meng-*upload* lagi pilihan *file* data *testing* dan menampilkan grafik prediksi dari data yang dipilih. Setelah selesai *user* akan keluar (*logout*) dari aplikasi. Dibawah ini adalah *use case* dari *menu login* dan *class diagram*

a. *Use Case Diagram*

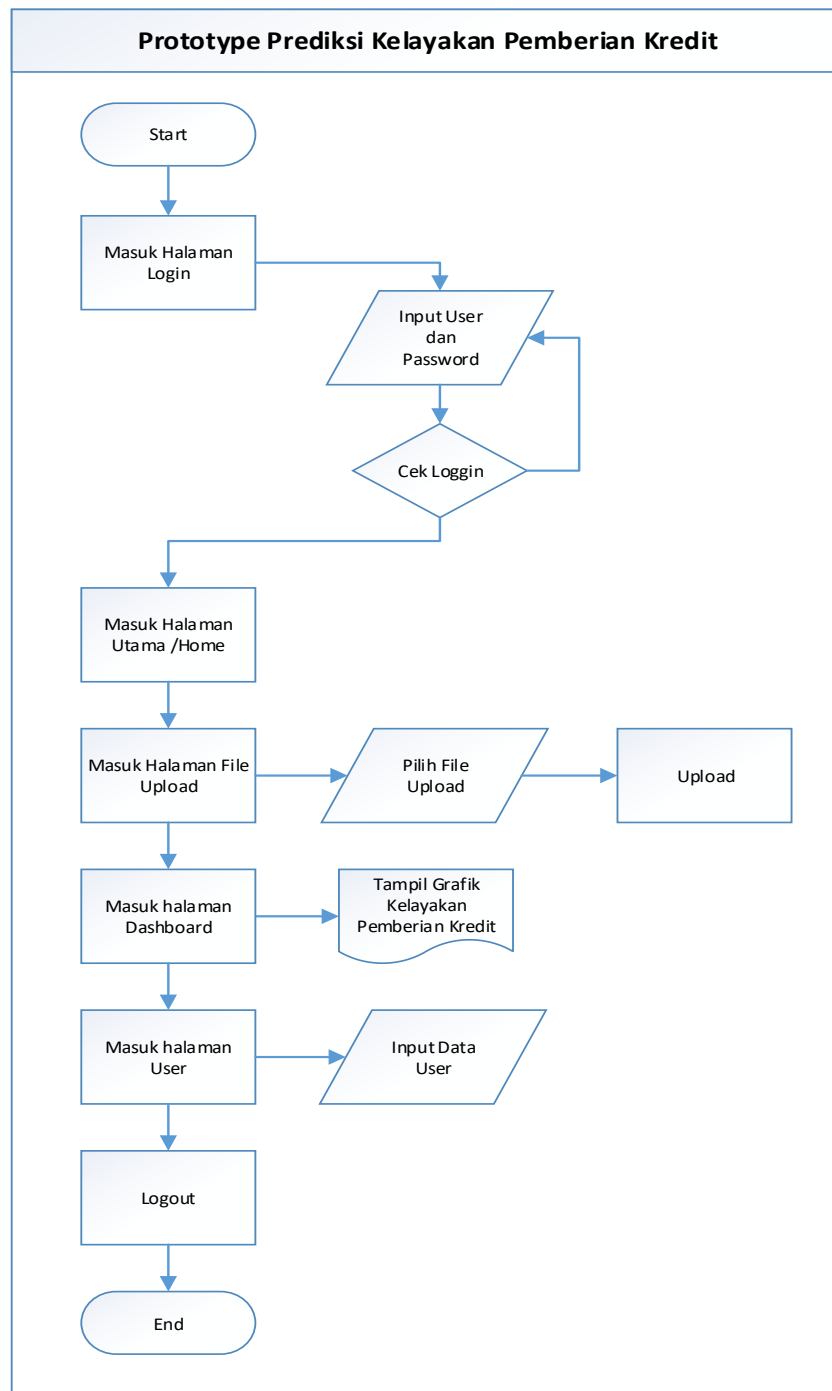


Gambar 4.7 *Use Case Login*

b. *Class diagram*



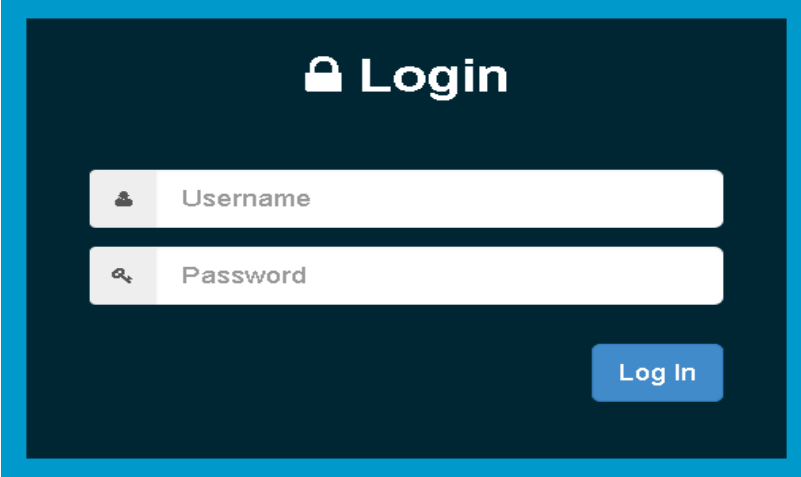
Gambar 4.8 *Class Diagram*



Gambar 4.9 *Flow Chart Prototype* Prediksi Kelayakan Pemberian Kredit

Prototype yang digunakan dalam penelitian ini dibuat berbasis *web* dengan menggunakan PHP dan *database* menggunakan *MySQL*. Tampilan untuk *Form* utama *Graphical User Interface (GUI) prototype* prediksi kelayakan pemberian kredit dapat dilihat pada gambar dibawah.

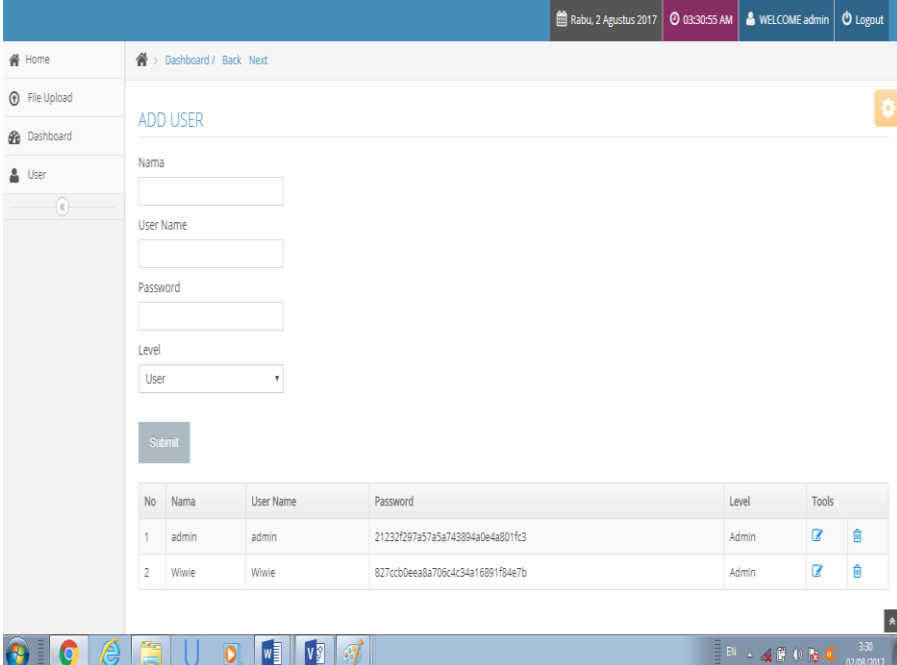
1. Tampilan Login



The image shows a login form with a dark blue background and a light blue border. At the top center, there is a white padlock icon followed by the word "Login" in white. Below this, there are two white input fields. The first field is labeled "Username" and has a small person icon on the left. The second field is labeled "Password" and has a small magnifying glass icon on the left. At the bottom right of the form, there is a blue button with the text "Log In" in white.

Gambar 4.10 *Form Login*

2. Tampilan Halaman Utama



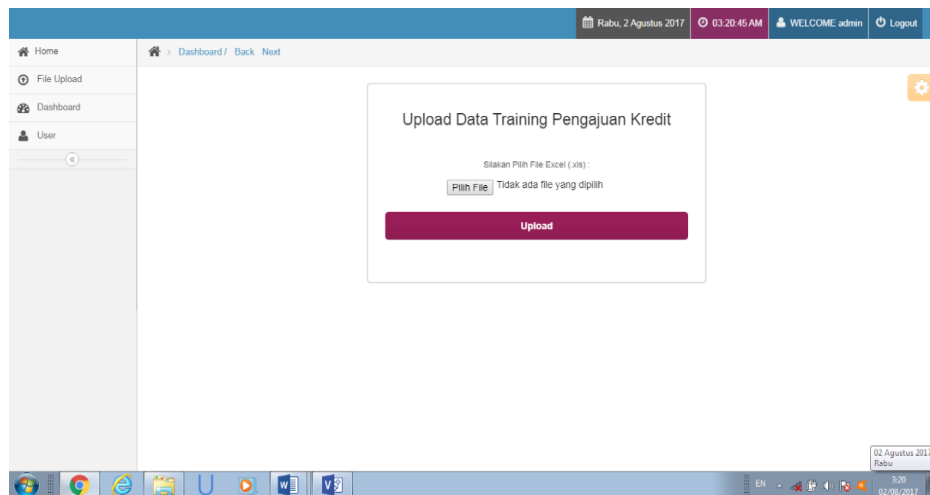
The image shows a dashboard page with a blue header. The header contains the date "Rabu, 2 Agustus 2017", the time "03:30:55 AM", the user name "WELCOME admin", and a "Logout" button. The main content area is divided into a left sidebar and a main panel. The sidebar has a "User" menu item. The main panel has a title "ADD USER" and a form with fields for "Nama", "User Name", "Password", and "Level" (a dropdown menu set to "User"). Below the form is a "Submit" button. At the bottom of the main panel, there is a table with the following data:

No	Nama	User Name	Password	Level	Tools
1	admin	admin	21232f297a57a5a743894a0e4a801fc3	Admin	Edit Delete
2	Wiwie	Wiwie	827ccb0eea8a706c4c34a16891f84e7b	Admin	Edit Delete

Gambar 4.11 *Form User*

Pihak admin dapat menambahkan *user* pada menu *user*.

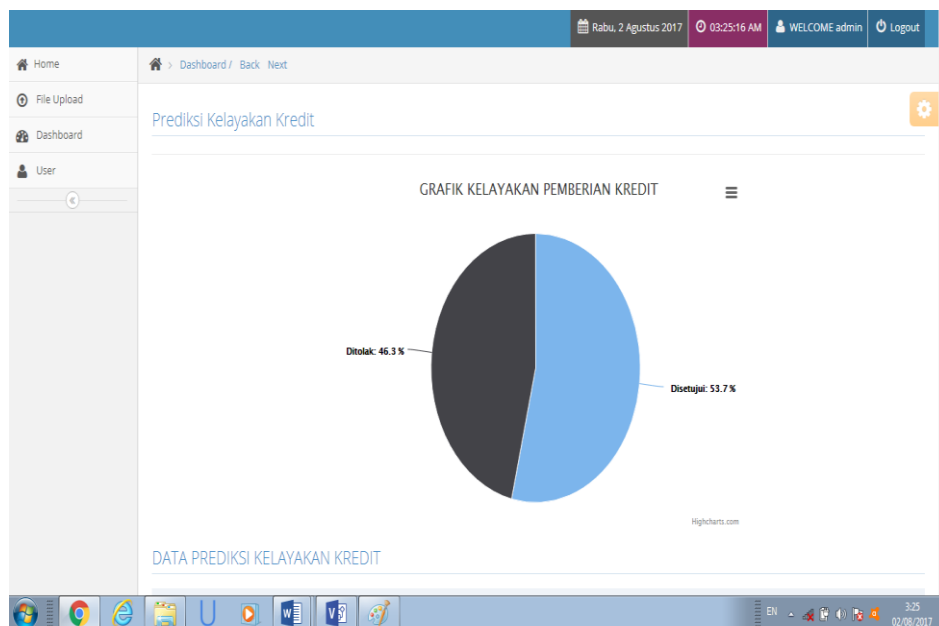
3. Tampilan Halaman *Upload*



Gambar 4.12 Form Upload Data Training Pengajuan Kredit

Pada menu file *upload*, user dapat meng-klik pilih *file* untuk memilih *file* data yang akan di uji. Selanjutnya akan ditampilkan hasil import data dari *file* yang dipilih.

4. Tampilan *Dashboard*



Gambar 4.13 Halaman *Dashboard*

Pada menu *Dashboard*, user dapat melihat grafik dari *file* yang telah dipilih, dengan mengklik *Dashboard*. Menu ini menampilkan juga detail dari hasil prediksi kelayakan kredit.

5. Tampilan *Detail Form*

DATA DETAIL KREDITUR Disetujui

Excel

No	Usia	Gender	Pendidikan Terakhir	Status	Jml Tanggungan	Status Tmpt Tinggal	Pekerjaan	Jml Pinjaman Yg Ditinginkan	Tujuan Pinjaman	Income/Year	Hasil	Pred. Hasil
1	52	M	S2	Menikah	1	Milik Sendiri	Wiraswasta	5500	Personal Loan	11250	Disetujui	Disetujui
2	47	M	S1	Menikah	2	Milik Keluarga	Guru	250	CarLoan	725	Disetujui	Disetujui
3	48	F	S1	Menikah	4	Milik Keluarga	Pegawai Swasta	500	Housing Loan	750	Disetujui	Disetujui
4	45	F	S2	Menikah	3	Milik Sendiri	Wiraswasta	2700	Housing Loan	6500	Disetujui	Disetujui
5	49	M	S2	Belum Menikah	1	Milik Keluarga	Wiraswasta	800	Personal Loan	1500	Disetujui	Disetujui
6	40	F	S2	Menikah	2	Milik Sendiri	Wiraswasta	600	Housing Loan	875	Disetujui	Disetujui
7	52	M	S1	Menikah	1	Milik Keluarga	Pegawai Swasta	500	Personal Loan	675	Disetujui	Disetujui
8	50	M	S1	Menikah	2	Milik Sendiri	Pegawai Negeri	250	CarLoan	250	Disetujui	Disetujui
9	47	M	S3	Menikah	3	Milik Sendiri	Dosen	300	Personal	250	Disetujui	Disetujui

Gambar 4.14 Detail Nasabah

Detail dari hasil prediksi akan ditampilkan secara rinci dan dapat dilihat dalam bentuk Microsoft Excel.

4.7 Pengujian *Prototype Perangkat Lunak (Software)*

4.7.1 Pengujian *Prototype*

Pengujian dilakukan dengan tujuan untuk mengetahui apakah aplikasi dibangun sesuai dengan fungsional yang di harapkan.

Tabel 4.11 Pengujian Kotak Hitam

Kelas Uji	Butir Uji	Jenis Pengujian
<i>File Upload</i>	<i>Pilih File</i>	<i>Black Box</i>
	<i>Upload File</i>	<i>Black Box</i>
<i>Dashboard</i>	<i>Lihat Grafik</i>	<i>Black Box</i>

A. Pengujian *File Upload*

Pada pengujian *file upload* terbagi atas dua bagian yaitu, pilih *file* dan *upload file*.

Tabel 4.12 Pengujian *File Upload*

Kasus dan Hasil Uji (Data sesuai)			
Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
<i>File</i> yang akan di <i>upload</i>	File dalam bentuk <i>ext.xls</i>	Dapat melakukan pilih <i>file upload</i>	Diterima
Klik <i>Upload</i>	File berhasil diimport oleh sistem	Dapat melakukan <i>upload data</i>	Diterima
Kasus dan Hasil Uji (Data salah)			
Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
<i>Extension</i> berbeda, tidak sesuai dengan sistem	Data tidak tersimpan	Data tidak tersimpan dan menampilkan kesalahan, sesuai yang diharapkan	Diterima

B. Pengujian *Dashboard*

Pada pengujian *Dashboard user* hanya meng-klik *Dashboard*.

Tabel 4.13 Pengujian *Dashboard*

Kasus dan Hasil Uji (Data sesuai)			
Data Masukan	Yang Diharapkan	Pengamatan	Kesimpulan
<i>File</i> yang telah di <i>upload</i>	Berhasil menampilkan Grafik persentase	Dapat menampilkan grafik presentase	Diterima
Klik Tombol Detail	Dapat menampilkan detail <i>customer</i> yang Disetujui ataupun Ditolak	Dapat menampilkan sesuai yang diharapkan	Diterima

Berdasarkan pengujian *Black Box* yang sudah dilakukan, menjelaskan bahwa aplikasi yang dibangun sudah berjalan dengan baik dan sesuai dengan yang diharapkan.

4.7.2 Lingkungan Pengujian

Lingkungan pengujian memberikan gambaran tentang spesifikasi *hardware* dan *software* yang digunakan oleh pengguna dalam proses pengujian sistem, baik pengujian validasi maupun kualitas. Spesifikasi tersebut diperoleh dalam proses observasi berdasarkan aspek sistem. Berikut ringkasan singkat spesifikasi dari perangkat keras dan perangkat lunak yang digunakan oleh pengguna untuk proses pengujian,

1. *Hardware*,
 - a. Processor AMD Dual Core E1-6010 1.35GHz
 - b. Harddisk 500 GB
 - c. RAM / Memory 2 GB
2. *Software*,
 - a. OS Win XP/7/8
 - b. XAMPP
 - c. PHP
 - d. MySQL
 - e. Web Browser (Mozilla Firefox/Google Chrome)

Perangkat lunak yang digunakan untuk membangun aplikasi *prototype* aplikasi prediksi kelayakan pemberian kredit adalah menggunakan *Software open source* yaitu PHP, XAMPP dan MySQL untuk *database*. Ditinjau dari segi kehandalan *software* tersebut sudah teruji dan sudah banyak digunakan untuk membangun sistem tanpa terkendala masalah lisensi karena bersifat *open source*.

4.7.3 Hasil Pengujian Kualitas Perangkat Lunak (*Software*)

Untuk dapat mengetahui kesesuaian yang diharapkan pada *prototype* sistem informasi prediksi kelayakan pemberian kredit maka metode yang digunakan untuk pengukuran perangkat lunak tersebut adalah *Software Quality*

Assurance (SQA). Adapun komponen yang menjadi tolak ukur yang digunakan pada pengujian ini dapat dilihat pada tabel 4.14.

Tabel 4.14 Komponen *Software Quality Assurance*

No	Metrik	Deskripsi
1	<i>Auditability</i>	Kesesuaian terhadap standar
2	<i>Accuracy</i>	Ketepatan perhitungan
3	<i>Completeness</i>	Kelengkapan kebutuhan
4	<i>Error Tolerance</i>	Toleransi terhadap kesalahan
5	<i>Execution Efficiency</i>	Performa <i>run-time</i>
6	<i>Operability</i>	Kemudahan untuk dioperasikan
7	<i>Simplicity</i>	Kemudahan untuk dipahami
8	<i>Training</i>	Kemudahan pembelajaran

Ada delapan komponen dalam *Software Quality Assurance* (SQA) yang digunakan. Dari delapan komponen tersebut akan dibuat delapan pertanyaan untuk angket yang akan disebarakan kepada lima orang *user* yang diambil secara acak.

Tabel 4.15 Skor Metrik

User	SKOR METRIK								SKOR
	1	2	3	4	5	6	7	8	
1	100	80	80	80	100	80	80	80	85
2	80	60	80	60	80	80	80	80	75
3	100	80	80	80	100	100	100	80	90
4	80	80	80	80	100	100	80	100	88
5	80	80	60	80	80	80	80	80	78
RATA - RATA									83

Untuk menghitung nilai rata-rata dari keseluruhan kuisisioner adalah sebagai berikut:

$$\text{Rata - rata} = \frac{85 + 75 + 90 + 88 + 78}{5} = 83$$

Skor rata-rata yang dihasilkan adalah 83% merupakan hasil skor yang cukup tinggi, sehingga dapat disimpulkan bahwa kualitas perangkat lunak “*prototype* prediksi kelayakan pemberian kredit kepada nasabah” ini dapat

dinyatakan baik untuk dilanjutkan implementasinya sebagai sebuah sistem Prediksi Kelayakan Pemberian Kredit dengan fitur yang lebih lengkap (*Full Version*).

4.8 Implikasi Penelitian

Implikasi penelitian ini berkisar pada aspek sistem dan aspek penelitian lanjut. Aspek sistem terkait dengan teknis operasional, desain *hardware* dan *software* yang diperlukan. Sedangkan aspek penelitian lanjut berkaitan dengan penelitian lanjutan yang diperlukan untuk meningkatkan kualitas penelitian sebelumnya.

4.8.1 Aspek Sistem

Untuk menerapkan hasil penelitian ini dibutuhkan dukungan sistem yang baik, sehingga pihak yang berkepentingan dapat menggunakan hasil penelitian ini untuk memprediksi kelayakan pemberian kredit untuk nasabah. Oleh karena itu dibutuhkan sarana dan prasarana yang memadai yang terdiri dari *hardware*, *software* (sistem operasi dan aplikasi yang dibuat pada tahap *deployment*), dan infrastruktur lainnya guna memberikan hasil yang terbaik, spesifikasi *hardware* dan *software* yang dapat digunakan dalam penelitian ini memiliki spesifikasi minimal seperti pada butir 4.5.2 Lingkungan Pengujian Sistem.

4.8.2 Aspek Penelitian Lanjut

Penulis menyadari keterbatasan dalam penelitian saat ini, diharapkan pada penelitian selanjutnya yang sejenis dapat dipertimbangkan hal-hal berikut:

1. Mengkombinasikan lebih banyak metode dalam Analisa data dan penyelesaian masalah, sehingga didapat sebuah sistem yang lebih efektif dan efisien dalam pengolahan ataupun penyajian informasi.
2. Pengelolaan waktu penelitian agar dapat lebih dimaksimalkan, mengingat pendeknya waktu yang tersedia.
3. Peran dari responden sangatlah penting dalam mendukung penelitian ini terutama responden yang berkaitan langsung dengan objek penelitian.

4. Penelitian ini dapat dikembangkan dengan algoritma klasifikasi yang lain yang terdapat dalam *data mining*, seperti algoritma *Neural Network*, *K-Means* atau *SVM (Support Vector Machine)*.

BAB V

PENUTUP

5.1 Simpulan

Dari pengukuran kinerja dengan melakukan komparasi dua algoritma yang telah dilakukan berdasarkan jumlah data maka dapat disimpulkan bahwa algoritma *Decision Tree* memiliki kemampuan dalam pengambilan keputusan untuk menentukan nasabah potensial yang disetujui dan ditolak untuk pengajuan pinjaman yang ditawarkan perusahaan. Kedua algoritma *Decision Tree* dan *Naïve Bayes* dapat digunakan dalam menentukan nasabah yang potensial di PT Batavia Prosperindo Tbk. Cabang Pringsewu. Kedua algoritma ini dikomparasi kemudian diuji akurasi. Tingkat akurasi tertinggi lah yang digunakan dalam menentukan nasabah berpotensi yang ingin kembali mengajukan pinjaman. Adapun hasil penelitian dari percobaan yang telah dilakukan dapat disimpulkan bahwa:

1. Algoritma *Decision Tree* memiliki tingkat akurasi tertinggi yaitu 85,22% sedangkan *Naïve Bayes* sebesar 73,67% maka selisih diantaranya sebesar 11,55%. Model algoritma *Decision Tree* memiliki AUC sebesar 0.9560 dan *Naïve Bayes* 0.8955, dari nilai AUC tersebut algoritma *Decision Tree* termasuk dalam kategori *excellent classification* dan *Naïve Bayes* *good classification* maka Algoritma *Decision Tree* bisa diimplementasikan dalam menentukan calon nasabah berpotensi.
2. *Rule* yg dihasilkan algoritma *Decision Tree* diterapkan dalam *prototype* prediksi nasabah yang ingin mengajukan pinjaman kembali dengan hasil akurasi pengujian verifikasi *prototype* sebesar 83,00%. Berdasarkan akurasi yg dihasilkan *prototype* menunjukkan bahwa metode dan *prototype* yg diterapkan sudah baik dalam memprediksi pengajuan pinjaman oleh nasabah lama.

Dari hasil penelitian ini diharapkan algoritma terpilih yaitu algoritma *Decision Tree* dalam memprediksi keputusan nasabah untuk mengajukan pinjaman lebih tepat dan cepat, sehingga membantu pencapaian kinerja perusahaan.

5.2 Saran

Agar penelitian ini bisa ditingkatkan, berikut adalah saran-saran yang diusulkan:

1. Penelitian ini diharapkan dapat digunakan pihak Perusahaan sebagai bahan pertimbangan dalam menentukan dan menganalisa data calon nasabah sehingga dapat meningkatkan akurasi ketika melakukan klasifikasi nasabah yang potensial dalam mengajukan kembali pinjaman pada PT Batavia Prosperindo Tbk.
2. Penelitian ini dapat dikembangkan dengan metode optimasi lainnya seperti Ant Colony Optimization (ACO), Genetic Algorithm (GA), dan lainnya.
3. Penelitian ini dapat dikembangkan dengan metode klasifikasi data mining lainnya seperti Support Vector Machine, KNN dan lainnya untuk melakukan perbandingan

Implementasi calon nasabah yang telah diklasifikasi pada PT Batavia Prosperindo Tbk. Cabang Pringsewu memberikan dampak positif terhadap perusahaan. Oleh karena itu data-data yang banyak dapat diolah kemudian dicarikan polanya sehingga menemukan pengetahuan baru agar data-data yang banyak tidak terbuang sia-sia. Sehingga data-data tersebut dapat dibuat berubah menjadi sebuah pengetahuan baru melalui data mining.

Aplikasi data mining dapat lebih dikembangkan dengan menyesuaikan kasus yang ada serta penyesuaian dengan bahasa pemrograman yang dipilih harus lebih akurat dengan sistematika proses algoritma data mining yang dipilih. Perangkat keras dan perangkat lunak pendukung pun harus disesuaikan seiring berkembangnya teknologi, serta kaderisasi *operator* dan pembelajaran harus dilaksanakan agar eksistensi sistem data mining khususnya dalam menentukan nasabah asuransi yang potensial dapat terus berjalan dan mampu memberikan solusi pemecahan masalah dalam mengolah data yang lebih baik berdasarkan akurasi.

DAFTAR PUSTAKA

- [1] S. Wahyuningsih and D. R. Utari, “Perbandingan Metode K-Nearest Neighbor , Naive Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit,” *Konf. Nas. Sist. Inf. 2018*, pp. 619–623, 2018, [Online]. Available:
<http://jurnal.atmaluhur.ac.id/index.php/knsi2018/article/view/424/349>.
- [2] P. Studi and B. Kuliah, “Distribusi II Distribusi II,” pp. 145–148, 2006.
- [3] A. Samhudi, F. Ekonomi, U. Islam, K. Muhammad, and A. Al, “Penerapan Prinsip 5C Terhadap Pengambilan Keputusan Kredit Pada PT . BRI Unit Handil,” *J. Komun. Bisnis Dan Manaj.*, vol. 6, no. 1, 2019.
- [4] I. Menarianti, “Klasifikasi data mining dalam menentukan pemberian kredit bagi nasabah koperasi,” vol. 1, no. 1, 2015.
- [5] D. Puspitasari, S. S. Al Khautsar, and W. P. Mustika, “Algoritma Naïve Bayes Untuk Memprediksi Kredit Macet Pada Koperasi Simpan Pinjam,” *J. Inform. Upgris*, vol. 4, no. 2, 2019, doi: 10.26877/jiu.v4i2.2919.
- [6] P. Effendi, “Kedudukan Para Pihak dalam Perjanjian Satandar Perbankan Ditinjau dari Kitab Undang-Undang Hukum Tentang Perlindungan Konsumen,” *J. Pro Huk.*, vol. IV, no. 2, pp. 87–92, 2015.
- [7] Arifin, “Data Warehouse , Data Mart , OLAP , dan Data Mining,” *Sist. Inf. - Udinus 1*, pp. 1–35, 2017.
- [8] E. P. W. Mandala and D. E. Putri, “Prediksi Jumlah Pemberian Kredit Kepada Nasabah Di Bank Perkreditan Rakyat Dengan Algoritma C 4.5,” *Komtekinfo*, vol. 5, no. 1, pp. 70–80, 2018.
- [9] M. F. Fibrianda and A. Bhawiyuga, “Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, pp. 3112–3123, 2018.
- [10] A. Ashari, “Performance Comparison between Naïve Bayes , Decision Tree and k-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool,” vol. 4, no. 11, pp. 33–39, 2013.

- [11] N. N. P. Galih. Surono, “Journal of technology information,” *Http://Jurnal.Kampuswiduri.Ac.Id/*, vol. 5, no. 1, pp. 25–30, 2020, [Online]. Available: <http://jurnal.kampuswiduri.ac.id/index.php/infoteh/article/view/79/67>.
- [12] Muhammad Sholeh, “Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2014 Yogyakarta, 15 November 2014 ISSN: 1979-911X,” *Snast*, no. November, pp. 211–216, 2014.
- [13] A. N. Kholifah and N. Insani, “Analisis Klasifikasi Pada Nasabah Kredit Koperasi X Menggunakan Decision Tree C4. 5 Dan Naive Bayes,” *J. Pendidik. Mat. dan Sains*, vol. 5, no. 6, pp. 1–8, 2016.
- [14] S. Masripah, “Komparasi Algoritma Klasifikasi Data Mining untuk Evaluasi Pemberian Kredit,” vol. 3, no. 1, pp. 187–193, 2016.
- [15] D. sudrajat and I. Moha, “Ragam Penelitian Kualitatif.” 2019, doi: 10.31227/osf.io/jaxbf.
- [16] D. T. Larose and C. D. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition*, vol. 9780470908. 2014.
- [17] Mardalis, *Metode Penelitian Suatu Pendekatan Proposal*, Jakarta: Bumi Aksara, 2008.
- [18] Gorunescu, F. (2011). *Data Mining Concept Model and Techniques*. Berlin: Springer. ISBN 978-3-642-19720-8

LAMPIRAN-LAMPIRAN

Lampiran 1: Daftar Pertanyaan Wawancara dengan Pakar

Daftar Pertanyaan dengan Pakar Penelitian Evaluasi Pengajuan Kredit Pinjaman oleh Nasabah pada PT Batavia Prosperindo Finance Tbk

Nama : Bambang Widianto
 Jabatan : Supervisor & BA Manger
 Tanggal : 15 Desember 2020

Untuk pertanyaan di bawah ini silakan dijawab dengan singkat dan jelas

Kebutuhan Fungsional :

1. Bagaimana sistem pengajuan kredit pinjaman yang digunakan saat ini?
 Jawab: menggunakan Data - Data yang dimasukkan ke Excel
2. Bagaimana proses evaluasi pengajuan pinjaman dilakukan?
 Jawab: melihat skor kredit dari data-data nasabah yang pernah mengajukan pinjaman.
3. Apa saja atribut sebagai penentu pengajuan pinjaman kredit?
 Jawab: pendidikan, status perkawinan, jumlah tanggungan, kepemilikan rumah, pekerjaan, penghasilan per bulan, nilai jaminan
4. Apa saja kendala yang dihadapi pada proses evaluasi pengajuan kredit?
 Jawab: Tidak ada tools evaluasi pengajuan pinjaman yang cepat dan akurat
5. Apa yang diharapkan apabila dilakukan penelitian evaluasi pengajuan pinjaman kredit terhadap nasabah yang ingin mengajukan kredit kembali?
 Jawab: Ada tools evaluasi pengajuan pinjaman yang terkomputerisasi yg bisa mengolah data secara cepat dan akurat
6. Dari atribut-atribut yang ditentukan dimasukkan dalam himpunan input, range, dan domain

No.	Atribut	Range	Domain
1.	Gender	Binomial	- L - P
2.	Age	20-80 tahun	≤ 40 > 40
3.	Pendidikan terakhir	S1 - S2	- SD - Diploma - SMP - S1 - SMA - S2
4.	Status perkawinan	-	Belum Menikah Menikah janda/puda
5.	Jumlah tanggungan	0 - 6	≤ 2 > 2
6.	Status tempat tinggal	-	Sewa - Diras Milik sendiri Keluarga
7.	Pekerjaan	-	PNS - Guru/Dok Wiraswasta - Petani TNI/POLRI - lain-lain
8.	Penghasilan per tahun	-	≤ 45 juta > 45 juta.
9.	Nilai jaminan	-	≤ 100 juta > 100 juta.
10.	Tujuan pinjaman	-	Modal Usaha Kendaraan
11.	Pinjaman maksimal yang bisa diproses	-	≤ 60 juta > 60 juta.

Lampiran 2. Kuesioner Terhadap Prototipe

Kuesioner Prototype Evaluasi Pengajuan Pinjaman Kredit Menggunakan Decision Tree dan Naïve Bayes pada PT Batavia Prosperindo Finance Tbk

a. Informasi Umum

Nama : Ferikortita
 Jabatan : Brand Manager
 Pendidikan : Diploma
 Jenis Kelamin : Pria Wanita
 Email : ferikortita17@gmail.com
 Tanggal : 15 Februari 2021

Menindaklanjuti hasil penelitian ini, peneliti ingin melakukan pengujian dan evaluasi aplikasi *prototype* tersebut. *Software Quality Assurance* (SQA) ini dilakukan untuk menilai kemampuan perangkat lunak yang berhubungan dengan variable independen yaitu: *auditability, accuracy, completeness, error tolerance, execution efficiency, operability, simplicity, dan training*

Petunjuk Pengisian

b. Informasi Kepentingan dan Kepuasan

Beri nilai dari 0 - 100 pada jawaban nilai yang tersedia pada kolom kepentingan dan kepuasan menurut penilaian pendapat anda.

Tabel 1.1 Komponen *Software Quality Assurance*

No	Metrik	Deskripsi	Nilai
1	<i>Auditability</i>	Kesesuaian terhadap standar	85
2	<i>Accuracy</i>	Ketepatan perhitungan	85
3	<i>Completeness</i>	Kelengkapan kebutuhan	85
4	<i>Error Tolerance</i>	Toleransi terhadap kesalahan	82
5	<i>Execution Efficiency</i>	Performa <i>run-time</i>	80
6	<i>Operability</i>	Kemudahan untuk dioperasikan	85
7	<i>Simplicity</i>	Kemudahan untuk dipahami	85
8	<i>Training</i>	Kemudahan pembelajaran	82

No.	No Kontrak	Nama	Age	Gender	Pend.terakhir	Status	lah Tangg.	pemilikan Ruma	Pekerjaan	nghasilan/ta	Nilai Jaminan	ujuan Pinjam	njaminan ma	HASIL
141	074372190235	AFIT ABDIN	30	L	SMA	Belum Menikah	0	Milik Sendiri	Petani	113.220.000	129.000.000	Kendaraan	102.000.000	Disetujui
142	074372190236	WINDRANI	34	P	SMA	Belum Menikah	0	Milik Sendiri	Petani	62.825.000	85.000.000	Kendaraan	68.000.000	Disetujui
143	074372190237	ROSIWATI	35	P	SMA	Menikah	1	Milik Sendiri	Petani	46.735.000	140.000.000	Modal Usaha	72.800.000	Disetujui
144	074372190240	Y EKO GUNAWANS (Nasabah GP)	53	L	SMA	Menikah	5	Milik Sendiri	Petani	78.600.000	91.500.000	Kendaraan	72.972.000	Disetujui
145	074372190243	DIARAHMAN	52	L	SMP	Menikah	2	Milik Sendiri	Petani	82.750.000	115.000.000	Kendaraan	91.125.000	Disetujui
146	074372190244	SUBIYANTI	34	P	SMA	Menikah	2	Milik Sendiri	Petani	62.230.000	112.000.000	Kendaraan	88.150.000	Disetujui
147	074372190245	SYARIF RAWAN	25	L	SMA	Menikah	3	Milik Sendiri	Petani	131.100.000	145.000.000	Modal Usaha	115.000.000	Disetujui
148	074372190246	DEVY SETAWAN	26	L	SMA	Menikah	2	Milik Sendiri	Petani	45.344.000	70.000.000	Modal Usaha	48.000.000	Disetujui
149	074372190247	HASAN BASRI	32	L	S1	Menikah	2	Milik Sendiri	PNS	13.752.000	145.000.000	Kendaraan	74.500.000	Disetujui
150	074372190249	BAYU SETAWAN	34	L	SMP	Menikah	3	Milik Sendiri	Wiraswasta	50.544.000	70.000.000	Kendaraan	53.500.000	Disetujui
151	074372190250	FATHURRAZI	24	L	SMP	Menikah	2	Milik Sendiri	Wiraswasta	57.070.000	105.000.000	Kendaraan	61.130.000	Disetujui
152	074372190251	WINOTO	46	L	SMP	Menikah	2	Milik Sendiri	Wiraswasta	69.160.000	90.000.000	Kendaraan	72.000.000	Disetujui
153	074372190252	IMAM SUBAHR	46	L	SMA	Menikah	2	Milik Sendiri	Wiraswasta	42.250.000	57.000.000	Kendaraan	41.850.000	Disetujui
154	074372190254	WILDAN HISTORI	34	L	SMP	Menikah	3	Milik Sendiri	Wiraswasta	59.930.000	90.000.000	Kendaraan	62.370.000	Disetujui
155	074372190255	JUBAIDI	33	L	SMA	Menikah	1	Milik Sendiri	Wiraswasta	80.640.000	147.000.000	Kendaraan	115.675.000	Disetujui
156	074372190257	IMAM BUCHORI	30	L	SMA	Menikah	1	Milik Sendiri	Wiraswasta	49.790.000	70.000.000	Kendaraan	52.425.000	Disetujui
157	074372190258	SYARIPUDIN (Nasabah GP)	45	L	SMA	Menikah	0	Milik Sendiri	Wiraswasta	77.850.000	93.000.000	Kendaraan	71.425.000	Disetujui
158	074372190259	NARA OKTARIO	56	L	SMA	Menikah	2	Milik Sendiri	Wiraswasta	50.700.000	70.000.000	Kendaraan	54.000.000	Disetujui
159	074372190261	HENDI ARI	54	L	SMA	Menikah	1	Milik Sendiri	Wiraswasta	48.400.000	70.000.000	Modal Usaha	52.300.000	Disetujui
160	074372190262	SIDIK ARIFIN	46	L	SMA	Menikah	1	Milik Sendiri	Wiraswasta	47.710.000	65.000.000	Modal Usaha	50.225.000	Disetujui
161	074372190263	TOPK HDAYAT (Nasabah GP)	46	L	SMA	Menikah	2	Milik Sendiri	Wiraswasta	88.100.000	115.000.000	Kendaraan	88.500.000	Disetujui
162	074372190264	DEDI SUTRISNO	47	L	SMA	Menikah	2	Milik Sendiri	Wiraswasta	42.250.000	55.000.000	Kendaraan	41.180.000	Disetujui
163	074372190265	SUKMIN	47	L	SMA	Menikah	2	Milik Sendiri	Petani	76.830.000	105.000.000	Modal Usaha	82.300.000	Ditolak
164	074372190266	BAMBANG RAWAN	48	L	SMA	Menikah	3	Milik Sendiri	Petani	43.940.000	60.000.000	Kendaraan	45.225.000	Disetujui
165	074372190267	BUDI YANTO	50	L	SMA	Menikah	1	Milik Sendiri	Petani	64.740.000	90.000.000	Kendaraan	67.375.000	Disetujui
166	074372190268	AHMAD KRISPANTORO	30	L	SMA	Menikah	1	Milik Sendiri	Petani	59.800.000	90.000.000	Modal Usaha	61.550.000	Disetujui
167	074372190269	NUR ANI	50	P	SMA	Menikah	0	Milik Keluarga	Petani	9.538.000	104.000.000	Kendaraan	52.500.000	Ditolak
168	074372190270	ERNA SARI	33	P	SMA	Menikah	2	Milik Keluarga	Wiraswasta	69.420.000	93.000.000	Kendaraan	74.340.000	Ditolak
169	074372190271	SUAJITOKO	32	L	S1	Menikah	3	Milik Keluarga	PNS	31.710.000	60.000.000	Kendaraan	44.000.000	Disetujui
170	074372190274	GUNADI RIZAL	43	L	S1	Menikah	1	Milik Keluarga	PNS	37.875.000	90.000.000	Kendaraan	51.130.000	Disetujui
171	074372190275	MUKTI KURNAWAN	42	L	S1	Menikah	1	Milik Sendiri	Wiraswasta	65.610.000	120.000.000	Kendaraan	66.905.000	Disetujui
172	074372190277	MAHRUB	42	L	S1	Menikah	0	Milik Sendiri	Wiraswasta	77.350.000	102.500.000	Kendaraan	81.900.000	Disetujui
173	074372190279	BUDIONO	44	L	S1	Menikah	2	Milik Sendiri	Wiraswasta	50.220.000	66.000.000	Kendaraan	55.215.000	Disetujui
174	074372190280	ION CHANDRA	45	L	S1	Menikah	2	Milik Sendiri	Wiraswasta	46.650.000	78.000.000	Kendaraan	61.670.000	Disetujui
175	074372190282	SUBROTO	51	L	S1	Menikah	3	Rumah Dinas	TNI/POLRI	30.750.000	49.000.000	Kendaraan	38.725.000	Disetujui
176	074372190283	DEDI HARYANTO	50	L	S1	Menikah	2	Milik Keluarga	Guru/Dosen	95.004.000	101.500.000	Kendaraan	81.200.000	Ditolak
177	074372190284	JAMATUN	50	P	SMA	Menikah	5	Rumah Dinas	TNI/POLRI	94.365.000	122.000.000	Modal Usaha	97.330.000	Ditolak
178	074372190285	FERA ANSEKA	34	P	SMA	Menikah	0	Milik Sendiri	Wiraswasta	61.830.000	80.000.000	Kendaraan	61.960.000	Ditolak
179	074372190286	HI DEDDY S. SOERGWA	34	L	DIPLOMA	Menikah	1	Milik Sendiri	Wiraswasta	67.310.000	130.000.000	Kendaraan	63.515.000	Disetujui
180	074372190287	NGADINEM	33	P	DIPLOMA	Menikah	3	Milik Sendiri	Wiraswasta	39.906.000	90.000.000	Modal Usaha	40.000.000	Disetujui
181	074372190288	RESTU AGUNG PRAYUGO	32	L	DIPLOMA	Menikah	4	Milik Sendiri	Wiraswasta	100.305.000	128.000.000	Kendaraan	102.275.000	Disetujui
182	074372190289	ERLICA	42	P	SMA	Menikah	4	Milik Sendiri	Wiraswasta	143.640.000	190.000.000	Kendaraan	144.000.000	Disetujui
183	074372190292	SUPLI	45	L	SMA	Menikah	2	Milik Sendiri	Wiraswasta	41.250.000	85.000.000	Kendaraan	55.215.000	Disetujui
184	074372190293	YANI KURNAWATI	46	P	SMP	JANDA	2	Milik Sendiri	Wiraswasta	95.580.000	127.500.000	Kendaraan	98.575.000	Disetujui
185	074372190295	EDI PURWANTO (Nasabah GP)	47	L	SD	Menikah	2	Milik Sendiri	Petani	66.000.000	105.000.000	Kendaraan	60.550.000	Disetujui
186	074372190296	DAVID ANDRIANTO (Nasabah GP)	34	L	SMA	Menikah	2	Sewa	Petani	55.800.000	95.000.000	Kendaraan	61.475.000	Ditolak
187	074372190297	EKA RUSDANTO	25	L	DIPLOMA	Menikah	2	Sewa	Petani	34.200.000	60.000.000	Kendaraan	44.850.000	Disetujui
188	074372190299	AHMAD HDAYAT	26	L	DIPLOMA	Menikah	3	Milik Keluarga	Wiraswasta	49.275.000	67.000.000	Kendaraan	50.225.000	Disetujui
189	074372190300	ANDRI SUSANTO	32	L	DIPLOMA	Menikah	2	Milik Keluarga	Wiraswasta	80.028.000	102.000.000	Kendaraan	81.600.000	Disetujui
190	074372190301	GATNO	34	L	DIPLOMA	Menikah	5	Milik Keluarga	Wiraswasta	47.550.000	100.000.000	Modal Usaha	62.350.000	Disetujui
191	074372190302	BUCHORI BUKSR	24	L	DIPLOMA	Belum Menikah	0	Milik Keluarga	Wiraswasta	74.655.000	110.000.000	Kendaraan	77.000.000	Disetujui
192	074372190305	ENDRO WIDYANTORO	46	L	SMA	Menikah	1	Milik Sendiri	Petani	68.175.000	98.000.000	Modal Usaha	68.340.000	Disetujui
193	074372190307	SUMARNO	46	L	SMA	Menikah	3	Milik Sendiri	Petani	36.400.000	130.000.000	Kendaraan	45.675.000	Disetujui
194	074372190308	SAPRUDIN	34	L	SMA	Menikah	4	Milik Sendiri	Wiraswasta	51.996.000	65.000.000	Kendaraan	49.700.000	Disetujui
195	074372190310	DODI KURNAWAN	50	L	SMA	Menikah	3	Rumah Dinas	Guru/Dosen	84.500.000	116.500.000	Kendaraan	92.925.000	Disetujui
196	074372190312	SARYONO	52	L	SMA	Menikah	1	Milik Keluarga	Wiraswasta	135.360.000	140.000.000	Kendaraan	112.000.000	Disetujui
197	074372190314	ARISKA JULIANSYAH	47	L	SMA	Menikah	1	Milik Sendiri	Wiraswasta	40.000.000	65.000.000	Kendaraan	50.200.000	Disetujui
198	074372190315	ENI WINARTI (Nasabah GP)	45	P	SMA	Menikah	0	Milik Sendiri	Wiraswasta	73.000.000	102.000.000	Kendaraan	73.600.000	Disetujui
199	074372190316	DEDY SANTOHA	50	L	SMA	Menikah	2	Milik Sendiri	Wiraswasta	37.520.000	65.000.000	Kendaraan	45.560.000	Disetujui
200	074372190319	RFAN (Nasabah GP)	47	L	SMA	Menikah	3	Milik Sendiri	Wiraswasta	119.200.000	130.000.000	Kendaraan	103.740.000	Disetujui
201	074372190320	ANDIKA PRASETYA	49	L	SMA	Menikah	1	Rumah Dinas	TNI/POLRI	69.020.000	80.000.000	Kendaraan	62.475.000	Disetujui
202	074372190321	EVA APRILANTI A. MICKEL	51	P	SMA	Menikah	1	Milik Keluarga	Wiraswasta	93.772.000	115.500.000	Kendaraan	92.200.000	Disetujui
203	074372190322	FEWIN HAIRONI	48	L	SMA	Menikah	4	Milik Keluarga	Petani	36.720.000	59.000.000	Kendaraan	44.225.000	Disetujui
204	074372190323	SOFYAN EFFENDI	47	L	SMA	Menikah	2	Milik Keluarga	Petani	113.680.000	140.000.000	Kendaraan	111.760.000	Disetujui
205	074372190324	ROHDI	40	L	SMA	Menikah	2	Milik Keluarga	Petani	76.860.000	93.000.000	Kendaraan	74.300.000	Disetujui
206	074372190325	AVITINA KURNIA WIATI	50	P	SMA	Menikah	3	Milik Keluarga	Petani	39.680.000	85.000.000	Kendaraan	50.210.000	Disetujui
207	074372190327	RIZON ANTON (Nasabah GP)	55	L	SMA	Menikah	2	Rumah Dinas	TNI/POLRI	97.960.000	110.000.000	Kendaraan	88.000.000	Ditolak
208	074372190328	MIRWAN EFFENDI	35	L	S1	Menikah	0	Sewa	Buruh	84.980.000	100.000.000	Modal Usaha	79.875.000	Disetujui
209	074372190331	AGUNG SAPITRO	26	L	S1	Menikah	0	Sewa	Wiraswasta	56.100.000	64.000.000	Modal Usaha	46.000.000	Disetujui
210	074372190332	BUDI HARYADI	35	L	S2	Belum Menikah	0	Sewa	Buruh	114.800.000	146.000.000	Kendaraan	113.525.000	Disetujui

Lampiran 4. Hasil Testing Data Prediksi Disetujui

PREDIKSI CUSTOMER DISETUJUI																
LIST CUSTOMER																
NO	NO KONTRAK	NAMA	DOB	GENDER	AGE	MARITAL	PEKERJAAN	PENDIDIKAN	JML TANGGUNGAN	STATUS TMPT TINGGAL	PENGHASILAN/TAHUN	NILAI JAMINAN	TUJUAN PINJAM	PINJAMAN MAX	HASIL	PREDIKSI HASIL
1	74372190071	SYAHRIL ANWAR	04/04/1977	L	P	Menikah	Petani	S1	K	Milik Sendiri	K	K	Kendaraan	K	Disetujui	Disetujui(A)
2	74372190072	IKA RACHIMAH	09/09/1979	P	P	Menikah	Wiraswasta	S1	B	Milik Sendiri	B	B	Kendaraan	B	Ditolak	Disetujui(A)
3	74372190074	ZIKRANI	16/02/1978	L	P	Menikah	Wiraswasta	SMA	B	Milik Sendiri	B	B	Kendaraan	B	Disetujui	Disetujui(A)
4	74372190075	NURHADI	29/03/1975	L	P	Menikah	Wiraswasta	SMA	B	Milik Sendiri	B	K	Modal Usaha	B	Disetujui	Disetujui(A)
5	74372190088	SUSANTO *(Nasabah GP)	01/01/1977	L	KP	Menikah	Petani	SMP	K	Milik Sendiri	K	B	Kendaraan	B	Ditolak	Disetujui(A)
6	74372190092	IBNU SAYAF *(Nasabah GP)	18/08/1978	L	P	Menikah	PNS	S1	K	Milik Sendiri	K	K	Kendaraan	K	Ditolak	Disetujui(A)
7	74372190096	FURZAN	17/11/1977	L	P	Menikah	PNS	SMA	K	Milik Sendiri	B	B	Kendaraan	B	Disetujui	Disetujui(A)
8	74372190099	ZUNANDA SUPIYADI	27/12/1975	L	P	Menikah	Guru/Dosen	SMA	B	Milik Sendiri	K	B	Kendaraan	B	Disetujui	Disetujui(A)
9	74372190101	DWI ASTUTI *(Nasabah GP)	23/12/1977	P	P	Menikah	Petani	SMA	K	Milik Sendiri	K	K	Kendaraan	K	Disetujui	Disetujui(A)
10	74372190103	JAJUN HOLIK *(Nasabah GP)	10/03/1979	L	KP	Menikah	Wiraswasta	S1	K	Rumah Dinas	K	K	Kendaraan	K	Disetujui	Disetujui(A)
11	74372190104	AYUNA DEWI	01/07/1978	P	P	Menikah	Wiraswasta	S1	K	Sewa	B	B	Kendaraan	B	Disetujui	Disetujui(A)
12	74372190105	ADI SUNTORO *(Nasabah GP)	09/05/1978	L	P	Menikah	Wiraswasta	DIPLOMA	K	Sewa	B	B	Kendaraan	B	Disetujui	Disetujui(A)
13	74372190111	NANANG WISMA	14/05/1978	L	P	Menikah	Wiraswasta	DIPLOMA	B	Sewa	B	B	Kendaraan	B	Disetujui	Disetujui(A)
14	74372190112	BUDIIONO	17/08/1975	L	P	Belum Menikah	Lain-lain	SMA	B	Milik Keluarga	K	K	Kendaraan	K	Disetujui	Disetujui(A)
15	74372190113	UMARUDIANTO	17/08/1976	L	P	Belum Menikah	Lain-lain	SMA	B	Milik Keluarga	K	B	Kendaraan	B	Disetujui	Disetujui(A)
16	74372190118	RATNA ARIANTI	17/08/1977	P	P	Belum Menikah	Wiraswasta	SMA	B	Milik Keluarga	K	K	Modal Usaha	K	Disetujui	Disetujui(A)
17	74372190119	RIVANTORO MANULANG	17/08/1978	L	P	Belum Menikah	Wiraswasta	SMA	B	Milik Keluarga	K	K	Modal Usaha	K	Disetujui	Disetujui(A)
18	74372190122	SURYADI	17/08/1979	L	P	Belum Menikah	Lain-lain	SMA	B	Milik Keluarga	K	K	Modal Usaha	K	Disetujui	Disetujui(A)
19	74372190123	HENIWATI	17/08/1980	P	P	Belum Menikah	Lain-lain	S1	B	Milik Keluarga	B	K	Modal Usaha	B	Disetujui	Disetujui(A)
20	74372190124	AHANI *(Nasabah GP)	17/08/1981	L	P	Belum Menikah	Lain-lain	SMA	K	Milik Keluarga	B	K	Modal Usaha	B	Disetujui	Disetujui(A)
21	74372190125	BADRUN	17/08/1982	L	P	Belum Menikah	Petani	SMA	K	Milik Keluarga	K	K	Modal Usaha	K	Disetujui	Disetujui(A)
22	74372190126	PUSPA WARDI	17/08/1983	P	P	Menikah	Petani	SMA	K	Milik Keluarga	B	B	Kendaraan	B	Disetujui	Disetujui(A)
23	74372190127	SUPARNO	17/08/1984	L	KP	Duda	Petani	SMA	K	Milik Keluarga	B	B	Kendaraan	B	Disetujui	Disetujui(A)
24	74372190129	BAYU ISMAWAN	17/08/1985	L	KP	Duda	PNS	S2	K	Milik Sendiri	K	K	Kendaraan	K	Disetujui	Disetujui(A)
25	74372190131	SUSILOWATI	17/08/1986	P	P	Menikah	Petani	SMA	B	Milik Sendiri	B	B	Kendaraan	B	Disetujui	Disetujui(A)
26	74372190132	TRI WINARTI	17/08/1987	P	P	Menikah	Petani	SMA	K	Milik Sendiri	K	K	Kendaraan	K	Disetujui	Disetujui(A)
27	74372190133	MEI RAHMAWATI SE	17/08/1988	P	P	Menikah	Petani	SMA	K	Milik Sendiri	B	B	Kendaraan	B	Disetujui	Disetujui(A)
28	74372190137	RUDI TANZIRO	17/08/1989	L	P	Menikah	Petani	SD	K	Sewa	B	K	Kendaraan	B	Disetujui	Disetujui(A)
29	74372190138	SARNUBI	17/08/1990	L	P	Menikah	Petani	SMP	K	Sewa	B	B	Kendaraan	B	Ditolak	Disetujui(A)
30	74372190139	SURYADI	17/08/1991	L	P	Menikah	Petani	SMP	K	Milik Keluarga	K	K	Kendaraan	K	Disetujui	Disetujui(A)
31	74372190141	JAMAWATI	17/08/1992	P	KP	Menikah	TNI/POLRI	SMP	K	Rumah Dinas	B	K	Kendaraan	B	Disetujui	Disetujui(A)
32	74372190142	YUNI FITRIASARI *(Nasabah GP)	17/08/1993	P	KP	Menikah	Guru/Dosen	SMA	K	Milik Sendiri	B	K	Modal Usaha	K	Disetujui	Disetujui(A)
33	74372190145	HADRIYANDI	17/08/1994	L	KP	Duda	Wiraswasta	SMP	K	Milik Sendiri	K	K	Kendaraan	K	Ditolak	Disetujui(A)
34	74372190147	OCKE ADITTA SAPUTRA	17/08/1996	L	KP	Menikah	Wiraswasta	SMA	B	Milik Sendiri	B	B	Kendaraan	B	Disetujui	Disetujui(A)

Lampiran 5. Hasil Testing Data Prediksi Ditolak

PREDIKSI CUSTOMER																
LIST CUSTOMER																
NO	NO KONTRAK	NAMA	DOB	GENDER	AGE	MARITAL	PEKERJAAN	PENDIDIKAN	JML TANGGUNGAN	STATUS TMPT TINGGAL	PENGHASILAN/TAHUN	NILAI JAMINAN	TUJUAN PINJAM	PINJAMAN MAX	HASIL	PREDIKSI HASIL
1	74372190076	DEDI MARGONO	28/12/1975	L	P	Belum Menikah	Wiraswasta	SMA	K	Milik Sendiri	B	B	Kendaraan	B	Ditolak	Ditolak(R)
2	74372190079	KUSDARMAJI	06/08/1979	L	P	Belum Menikah	Petani	SMA	K	Milik Sendiri	K	K	Kendaraan	K	Ditolak	Ditolak(R)
3	74372190080	BURHANUDIN RAIS	10/01/1975	L	P	Belum Menikah	Wiraswasta	SMA	K	Milik Sendiri	K	K	Kendaraan	K	Ditolak	Ditolak(R)
4	74372190082	HARDI YANTO *(Nasabah GP)	18/02/1978	L	P	Belum Menikah	Petani	DIPLOMA	K	Milik Sendiri	B	K	Kendaraan	B	Ditolak	Ditolak(R)
5	74372190084	ENDRIYANI	17/04/1979	P	P	Belum Menikah	Wiraswasta	DIPLOMA	B	Sewa	B	B	Modal Usaha	B	Ditolak	Ditolak(R)
6	74372190086	EDI PRASTYO *(Nasabah GP)	13/08/1979	L	P	Belum Menikah	Petani	S1	K	Sewa	B	K	Modal Usaha	K	Ditolak	Ditolak(R)
7	74372190087	JAMRONI	30/10/1977	L	KP	Duda	Wiraswasta	SMP	K	Milik Sendiri	B	B	Kendaraan	B	Ditolak	Ditolak(R)
8	74372190090	HARDIONO	18/09/1977	L	KP	Duda	Petani	SD	K	Milik Sendiri	B	K	Kendaraan	B	Ditolak	Ditolak(R)
9	74372190091	ARI JOHARI	14/04/1979	L	P	Menikah	Petani	SMA	B	Milik Sendiri	B	K	Kendaraan	B	Ditolak	Ditolak(R)
10	74372190146	NELI ANTON	17/08/1995	L	KP	Duda	Wiraswasta	SMA	K	Milik Sendiri	B	K	Kendaraan	B	Disetujui	Ditolak(R)
11	74372190153	MISDI *(Nasabah GP)	17/08/1997	L	P	Menikah	Wiraswasta	SMA	K	Milik Sendiri	K	K	Kendaraan	K	Ditolak	Ditolak(R)

Lampiran 6. Source Code

File_upload.php

```

<?php
//membaca file excel yang diupload
$data = new Spreadsheet_Excel_Reader($_FILES['userfile']['tmp_name']);
//membaca jumlah baris dari data excel
$baris = $data->rowcount($sheet_index=0);
//nilai awal counter jumlah data yang sukses dan yang gagal diimport
$sukses = 0;
$gagal = 0;
//import data excel dari baris kedua, karena baris pertama adalah nama kolom
for ($i=2; $i<=$baris; $i++) {
    $no_kontrak = trim($data->val($i,1));
    $nama = trim($data->val($i,2));
    $dob = trim($data->val($i,3));
    $gender = trim($data->val($i,4));
    $type_age = trim($data->val($i,5));
    $marital = trim($data->val($i,6));
    $occupation = trim($data->val($i,7));
    $pendidikan = trim($data->val($i,8));
    $jml_tanggungan = trim($data->val($i,9));
    $status_tinggal = trim($data->val($i,10));
    $penghasilan_tahun = trim($data->val($i,11));
    $nilai_jaminan = trim($data->val($i,12));
    $tujuan_pinjam = trim($data->val($i,13));
    $pinjam_max = trim($data->val($i,14));
    $hasil = trim($data->val($i,15));
    $pred_hasil = trim($data->val($i,16));
    IF ($status_tinggal== 'Milik Sendiri') {
    IF ($tujuan_pinjam== 'Kendaraan') {
        IF ($marital == 'Duda') {
            IF($penghasilan_tahun == 'K') {
                $pred_hasil = "Disetujui(A)";
            }
            ELSEIF ($penghasilan_tahun == 'B') {
                $pred_hasil = "Ditolak(R)";}}}}

```

```

IF ($status_tinggal== 'Milik Sendiri') {
IF ($tujuan_pinjam== 'Kendaraan') {
    IF ($marital == 'Menikah') {
        IF($nilai_jaminan == 'K') {
            IF($gender == 'L') {
                IF($pendidikan == 'S1') {
                    $pred_hasil = "Disetujui(A)";
                }
                ELSEIF ($pendidikan == 'SMA') {
                    $pred_hasil = "Ditolak(R)";
                }
            }
        }
    }
}
}

IF ($status_tinggal== 'Milik Sendiri') {
IF ($tujuan_pinjam== 'Kendaraan') {
    IF ($marital == 'Menikah') {
        IF($nilai_jaminan == 'K') {
            IF($gender == 'P') {
                $pred_hasil = "Disetujui(A)";
            }
            ELSEIF ($nilai_jaminan == 'B') {
                $pred_hasil = "Disetujui(A)";
            }
        }
    }
}

IF ($status_tinggal== 'Milik Sendiri') {
IF ($tujuan_pinjam== 'Kendaraan') {
    IF ($marital == 'Menikah') {
        IF($nilai_jaminan == 'B') {
            $pred_hasil = "Disetujui(A)";
        }
    }
}

IF ($status_tinggal== 'Milik Sendiri') {
IF ($tujuan_pinjam== 'Modal Usaha') {
    $pred_hasil = "Disetujui(A)";
}

IF ($status_tinggal== 'Milik Sendiri') {
IF ($tujuan_pinjam== 'Kendaraan') {
    IF ($marital == 'Belum Menikah') {
        $pred_hasil = "Ditolak(R)";
    }
}

IF ($status_tinggal== 'Rumah Dinas') {

```

```

        $pred_hasil = "Disetujui(A)";
    }
IF ($status_tinggal== 'Sewa') {
    IF ($marital == 'Menikah') {
        $pred_hasil = "Disetujui(A)";
    }
    ELSEIF ($marital == 'Belum Menikah') {
        $pred_hasil = "Ditolak(R)";
    }
}
IF ($status_tinggal== 'Milik Keluarga') {
    $pred_hasil = "Disetujui(A)";
}

    $query          =          "INSERT          INTO          CUSTOMER
(no_kontrak,nama,dob,gender,type_age,marital,occupation,pendidikan,jml_tanggu
ngan,status_tinggal,penghasilan_tahun,nilai_jaminan,tujuan_pinjam,pinjam_max,t
hasil,pred_hasil)          values
('$no_kontrak','$nama','$dob','$gender','$type_age','$marital','$occupation','$pendi
dikan','$jml_tanggungannya','$status_tinggal','$penghasilan_tahun','$nilai_jaminan','$t
ujuan_pinjam','$pinjam_max','$thasil','$pred_hasil)";
$hasil = mysql_query($query, $koneksi);
//menambah counter jika berhasil atau gagal
if($hasil) $sukses++;
else $gagal++;
}
if ($upload) {
    //tampilkan report hasil import
    echo "<p>Jumlah data sukses diimport : ".$sukses."<br>";
    echo "Jumlah data gagal diimport : ".$gagal."<p>";
}
?>
<!doctype html>
<html lang="en">
<head>
    <meta charset="UTF-8" />
    <title>Upload File Nasabah</title>
<style type="text/css">
a{ text-decoration: none; color: #333}

```

```
h1{ font-size: 1.9em; margin: 10px 0}
p{ margin: 8px 0}
*{
    margin: 0;
    padding: 0;
    box-sizing: border-box;
    -webkit-box-sizing: border-box;
    -moz-box-sizing: border-box;
    -webkit-font-smoothing: antialiased;
    -moz-font-smoothing: antialiased;
    -o-font-smoothing: antialiased;
    font-smoothing: antialiased;
    text-rendering: optimizeLegibility;
}
body{
    font: 12px Arial,Tahoma,Helvetica,FreeSans,sans-serif;
    text-transform: inherit;
    color: #333;
    background: #e7edee;
    width: 100%;
    line-height: 18px;
}
.wrap{
    width: 500px;
    margin: 15px auto;
    padding: 20px 25px;
    background: white;
    border: 2px solid #DBDBDB;
    -webkit-border-radius: 5px;
    -moz-border-radius: 5px;
    border-radius: 5px;
    overflow: hidden;
    text-align: center;
}
.status{
    /*display: none;*/
    padding: 8px 35px 8px 14px;
```

```
margin: 20px 0;
text-shadow: 0 1px 0 rgba(255, 255, 255, 0.5);
color: #468847;
background-color: #dff0d8;
border-color: #d6e9c6;
-webkit-border-radius: 4px;
-moz-border-radius: 4px;
border-radius: 4px;
}
input[type="submit"] {
  cursor:pointer;
  width:100%;
  border:none;
  background:#991D57;
  background-image:linear-gradient(bottom, #8C1C50 0%, #991D57 52%);
  background-image:-moz-linear-gradient(bottom, #8C1C50 0%, #991D57 52%);
  background-image:-webkit-linear-gradient(bottom, #8C1C50 0%, #991D57 52%);
  color:#FFF;
  font-weight: bold;
  margin: 20px 0;
  padding: 10px;
  border-radius:5px;
}
input[type="submit"]:hover {
  background-image:linear-gradient(bottom, #9C215A 0%, #A82767 52%);
  background-image:-moz-linear-gradient(bottom, #9C215A 0%, #A82767 52%);
  background-image:-webkit-linear-gradient(bottom, #9C215A 0%, #A82767 52%);
  -webkit-transition:background 0.3s ease-in-out;
  -moz-transition:background 0.3s ease-in-out;
  transition:background-color 0.3s ease-in-out;
}
input[type="submit"]:active {
  box-shadow:inset 0 1px 3px rgba(0,0,0,0.5);
}
</style>
</head>
<body>
```

```
<div class="wrap">
  <h1>Upload Data Nasabah</h1>
  <br />
  <p>Silakan Pilih File Excel (.xls) :</p>
  <!-- Multiple file upload html form-->
  <form      method="post"      enctype="multipart/form-data"      action="<?php
    $_SERVER['PHP_SELF']; ?>">
    <input name="userfile" type="file" accept="application/msexcel" >
    <input name="upload" type="submit" value="Upload">
  </form>
</div>
</body>
</html>
```