

**SISTEM REKOMENDASI KELAYAKAN KREDIT
MENGUNAKAN METODE RANDOM FOREST PADA BRI
KANTOR CABANG PELAIHARI**

PROPOSAL SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

DISUSUN OLEH:

MUHAMMAD IRHAMNA PUTRA H76215022

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2019

PERSETUJUAN PEMBIMBING

JUDUL : SISTEM REKOMENDASI KELAYAKAN KREDIT
MENGGUNAKAN METODE RANDOM FOREST PADA BRI
KANTOR CABANG PELAIHARI
NAMA : MUHAMMAD IRHAMNA PUTRA
NIM : H76215022

Mahasiswa tersebut telah melakukan proses bimbingan dan dinyatakan layak
untuk mengikuti Sidang Proposal skripsi

Surabaya, 02 Mei 2019.

Menyetujui,

Dosen Pembimbing I

Dosen Pembimbing II

(Ahmad Yusuf, M. Kom)
NIP. 199001202014031003

(Nita Yalina, S.Kom., M.MT)
NIP. 198702082014032003

LEMBAR PENGESAHAN

JUDUL : SISTEM REKOMENDASI KELAYAKAN KREDIT
MENGUNAKAN METODE RANDOM FOREST PADA BRI
KANTOR CABANG PELAIHARI

NAMA : MUHAMMAD IRHAMNA PUTRA

NIM : H76215022

Proposal skripsi tersebut telah dipresentasikan pada Sidang Proposal Skripsi di
depan Dosen Penguji pada tanggal 02 Mei 2019.

Menyetujui,

Dosen Pembimbing I

Dosen Pembimbing II

(Ahmad Yusuf, M. Kom)
NIP. 199001202014031003

(Nita Yalina, S.Kom., M.MT)
NIP. 198702082014032003

Dosen Penguji I

Dosen Penguji II

(Mujib Ridwan, S.Kom., M.T)
NIP. 198604272014031004

(Nurissaidah Ulinnuha, M. Kom)
NIP. 199011022014032004

Mengetahui,

Ketua program Studi

Ketua jurusan

(Muhammad Andik Izzuddin, MT)
NIP. 198403072014031001

(Mujib Ridwan, S.Kom., M.T)
NIP. 198604272014031004

DAFTAR ISI

PERSETUJUAN PEMBIMBING.....	i
LEMBAR PENGESAHAN	ii
DAFTAR ISI.....	iii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah	3
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	4
1.6. Sistematika Penulisan Skripsi	5
BAB II TINJAUAN PUSTAKA	6
2.1. Tinjauan Penelitian Terdahulu	6
2.2. Dasar Teori.....	8
2.2.1. Sistem Pendukung Keputusan (SPK).....	8
2.2.2. Klasifikasi	9
2.2.3. Algoritma Random Forest.....	9
2.2.4. <i>Confusion Matrix</i> dan <i>Accuracy</i>	13
2.2.5. Analisa Kredit	14
2.2.6. Database	15
2.3. Integrasi Keilmuan.....	15
BAB III METODOLOGI PENELITIAN	18
3.1. Desain Penelitian	18
3.1.1. Perumusan Masalah	19
3.1.2. Studi Pustaka.....	19
3.1.3. Pengumpulan Data	19
3.1.4. Pengolahan Data	21
3.1.5. Pengembangan Sistem	22
3.1.6. Pengujian Sistem.....	22
DAFTAR PUSTAKA	24

BAB I

PENDAHULUAN

1.1. Latar Belakang

Bank merupakan salah satu lembaga yang memberikan pelayanan dalam jasa intermediasi keuangan. Menurut UU No 10 Tahun 1998 tanggal 10 November 1998 tentang perbankan, jasa atau usaha perbankan meliputi tiga kegiatan, yaitu menghimpun dana, menyalurkan dana dan memberikan jasa bank lainnya. Dalam kegiatan menyalurkan dana lebih tepatnya dalam pemberian kredit, ada sebuah istilah resiko kredit yang wajib dikelola oleh bank manapun. Resiko kredit ini adalah kegagalan atau ketidakmampuan dari debitur untuk memenuhi kewajibannya dalam hal pembayaran sesuai dengan ketentuan yang telah disepakati. Pihak bank tentu memerlukan ketelitian yang sangat tinggi dalam memilih calon debitur untuk mengurangi resiko kredit tersebut. Dalam resiko kredit, terdapat sebuah istilah kredit macet yang mana merupakan masalah utama dalam resiko kredit yang menjadi satu masalah yang kompleks. Kurang tepatnya melakukan penilaian awal pada calon debitur merupakan penyebab dari masalah ini.

Bank Rakyat Indonesia atau kerap dikenal sebagai Bank BRI tercatat sebagai bank yang paling besar meraup untung pada tahun 2017, berfokus pada penyaluran kredit ke usaha mikro, kecil dan menengah (UMKM) berhasil mencetak laba sebesar Rp. 13,4 triliun atau naik sekitar 10,4% dibanding periode tahun sebelumnya. Adapun pertumbuhan laba ini utamanya terdorong oleh tingginya penyaluran kredit. Total penyaluran kredit BRI sepanjang tahun 2017 tercatat sebesar Rp 687,9 triliun atau naik 11,8%. Terkait dengan rasio kredit bermasalah atau *Non Performing Loan* (NPL), pada tahun 2017 dengan pertumbuhan kredit yang signifikan tersebut, BRI berhasil menjaga rasio NPL gross nya sebesar 2,16% atau turun dibandingkan dengan NPL gross pada tahun 2016 sebesar 2,22%. Namun peningkatan tersebut merupakan statistik penurunan NPL secara garis besar, faktanya dalam lapangan ada beberapa cabang yang masih memiliki peningkatan

NPL misalnya Bank BRI Kantor Cabang Pelaihari. Dalam wawancara secara pribadi, disebutkan bahwa NPL pada kantor cabang tersebut mengalami peningkatan yang nilainya tidak boleh disebutkan. Hal ini dikarenakan pertumbuhan ekonomi di masyarakat yang mempengaruhi terhadap nilai NPL ini. Maka dari itu, untuk menekan peningkatan nilai NPL ini diharuskan untuk mengelola kelayakan kredit secara lebih mendetail lagi untuk calon nasabah atau debitur baru.

Kelayakan kredit merupakan kriteria penentu terhadap layak atau tidaknya calon debitur dalam hal diberikan kredit agar tidak menimbulkan kemacetan kredit yang berimbas pada rasio kredit bermasalah atau *Non Performing Loan* (NPL). Kelayakan kredit tersebut dapat diukur dengan skor kredit yang merupakan angka yang diberikan kepada suatu individu yang menjelaskan seberapa besar kemungkinan individu tersebut dalam memenuhi kewajiban finansialnya. Hal ini dapat di lihat sebagai probabilitas standar untuk mengukur calon debitur. Pemberi pinjaman, terutama institusi bank atau Lembaga lainnya memperhitungkan kemungkinan standar ini untuk menetapkan calon debitur apakah pinjaman ini nantinya akan masuk ke dalam kategori pinjaman baik (*Good Loan* atau pinjaman yang memiliki kemungkinan untuk membayar kembali pinjaman) atau pinjaman buruk (*Bad Loan* atau pinjaman yang tidak memiliki kemungkinan untuk membayar kembali pinjaman) berdasarkan probabilitas standar tersebut, sehingga individu tersebut dapat diklasifikasikan ke kategori yang telah ada. Atas dasar hal ini, dibutuhkanlah sebuah sistem pendukung keputusan untuk menilai calon debitur yang layak diberikan kredit.

Sistem Pendukung Keputusan adalah sistem berbasis komputer yang memiliki tujuan untuk membantu dalam pengambilan keputusan dengan memanfaatkan data dan model tertentu untuk memecahkan berbagai persoalan yang tidak terstruktur. Secara umum sistem pendukung keputusan memiliki peran untuk membantu pengambilan keputusan secara efektif. Sistem pendukung keputusan menggunakan data, memberikan antarmuka pengguna yang mudah, dan dapat menggabungkan pemikiran pengambil keputusan.

Metode Random Forest merupakan metode yang cukup efektif yang mampu memberikan rekomendasi terbaik dibandingkan dengan metode – metode *machine learning* lainnya. *Random Forest* adalah salah satu klasifikasi *data mining* yang merupakan metode pohon gabungan yang berasal dari pengembangan metode *Classification and Regression Tree* (CART), yaitu dengan menerapkan metode *bootstrap aggregating* (Bagging) dan *random feature selection* (Breiman 2001). Metode *random forest* ini telah diaplikasikan pada berbagai permasalahan dalam situasi penelitian yang berkecimpung pada kesehatan, bisnis, Pendidikan dan lainnya.

Sehubungan dengan pesatnya perkembangan teknologi di era ini, diperlukan sebuah sistem pendukung keputusan dengan cara menganalisa data rekapan calon debitur yang telah atau tidak diberikan pinjaman yang ada pada tahun sebelumnya. Pada tahap selanjutnya sistem akan menghitung kelayakan calon debitur baru berdasarkan data rekapan tersebut. Dengan harapan sistem ini kedepannya akan membantu pihak bank dalam mengurangi resiko kredit.

1.2. Perumusan Masalah

Dari latar belakang diatas dapat ditarik rumusan masalah sebagai berikut.

1. Bagaimana mengembangkan aplikasi penentuan kelayakan kredit?
2. Bagaimana evaluasi metode *random forest*?

1.3. Batasan Masalah

Adapun Batasan – Batasan masalah dalam penelitian ini adalah:

1. Data penelitian yang digunakan didapatkan dari Kantor BRI Cabang Pelayari, Kalimantan Selatan.
2. Data yang digunakan merupakan data pinjaman berdasarkan pinjaman yang berlangsung pada tahun 2015 – 2019.
3. Sistem pendukung keputusan ini akan dibangun berbasis *web*.

1.4. Tujuan Penelitian

1. Dapat mengembangkan aplikasi penentuan kelayakan kredit dengan menggunakan metode Random Forest.

2. Dapat membuktikan bahwa metode Random Forest mampu mengurangi resiko kredit pada Bank.

1.5. Manfaat Penelitian

Hasil dari penelitian yang diharapkan dapat memberikan manfaat bagi semua pihak yang terlibat dalam penelitian ini, diantaranya:

1. Bagi Peneliti
 - a. Sebagai media dalam menerapkan ilmu pengetahuan yang telah diperoleh selama melaksanakan kuliah.
 - b. Sebagai salah satu syarat dalam memperoleh gelar Sarjana Komputer Program Studi Sistem Informasi di Universitas Islam Negeri Sunan Ampel Surabaya.
2. Bagi Pembaca
 - a. Sebagai referensi ilmu pengetahuan dalam pengembangan aplikasi perhitungan dan metode Random Forest.
 - b. Membuka wawasan dalam pengembangan sebuah sistem dengan menggunakan metode Random Forest.
3. Bagi Bank
 - a. Dengan adanya aplikasi penentuan kelayakan kredit, mampu mengurangi resiko kredit dalam menentukan calon debitur baru.
 - b. Memberikan penilaian terhadap calon debitur baru yang berhak menerima kredit dari pihak bank sesuai dengan keadaan sebenarnya.
 - c. Digunakan sebagai dasar pengambilan keputusan yang terstruktur untuk mendukung proses utang piutang pada Bank BRI Cabang Pelayari.
 - d. Dapat digunakan untuk penentuan kelayakan pemberian kredit pada debitur.
 - e. Menjadi jembatan antara Calon Debitur Baru dan Pihak Bank dalam proses pengajuan kredit.

1.6. Sistematika Penulisan Skripsi

Adapun sistematika penulisan skripsi ini terdiri atas beberapa bagian, yakni sebagai berikut:

1. Bagian awal terdiri dari halaman sampul.
2. Bagian isi terdiri dari :

- a. BAB I PENDAHULUAN

Pendahuluan menjelaskan latar belakang permasalahan yang diangkat dalam penelitian, perumusan masalah, tujuan penelitian, dan manfaat yang diperoleh dari penelitian yang dilakukan.

- b. BAB II TINJAUAN PUSTAKA

Bab ini berisi tentang penjelasan penelitian sebelumnya yang relevan dengan penelitian ini, dan penjelasan singkat tentang teori – teori yang terkait dengan penelitian ini seperti konsep sistem informasi, metode yang terkait dengan penelitian ini seperti konsep sistem informasi, metode pengembangan sistem informasi, Bahasa pemrograman, metode random forest dan sebagainya.

- c. BAB III METODOLOGI PENELITIAN

Bab ini berisikan alur sistematika penelitian yang terdiri dari tahap identifikasi kebutuhan, tahap perencanaan dan tahap pengembangan prototipe dan tahap implementasi.

- d. BAB IV HASIL DAN PEMBAHASAN

Bab ini berisikan hasil dan pembahasan bagaimana data diolah berdasarkan algoritma *Random Forest* dan bagaimana suatu sistem dikembangkan.

- e. BAB V PENUTUP

Bab ini berisikan kesimpulan dari analisa masalah serta saran – saran dari penulis, sehingga apa yang menjadi tujuan dari penelitian ini dapat terwujud.

BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Penelitian Terdahulu

Dalam memberikan pemahaman mengenai keterkaitan penelitian terdahulu dengan penelitian yang akan dilakukan, dapat dilihat lebih lanjut dalam Tabel 2.1:

Tabel 2.1 Daftar Penelitian Terdahulu

No	Topik	Metode	Hasil
1.	KLASIFIKASI PEMBIAYAAN WARUNG MIKRO MENGGUNAKAN METODE <i>RANDOM FOREST</i> DENGAN TEKNIK <i>SAMPLING</i> KELAS <i>IMBALANCED</i> (Studi Kasus: Data Nasabah Pembiayaan Warung Mikro Bank Syariah Mandiri KC Jambi) (Widiastuti, 2018)	<i>Random Forest</i>	Metode <i>Random Forest</i> digunakan pada <i>Imbalanced Data</i> sehingga ketika di <i>training</i> akurasi mencapai 82.54%, namun nilai <i>sensitivity</i> bernilai 0, sehingga akurasi dianggap bias. Hasil klasifikasi didapatkan dengan nilai estimasi error OOB sebesar 25.44%
2.	SISTEM PENDUKUNG KEPUTUSAN KELAYAKAN PEMBERIAN KREDIT PADA BANK TABUNGAN NEGARA (BTN) MENGGUNAKAN ALGORITMA C4.5 (Octabriyantiningtyas, 2016)	<i>Decision Tree</i> – C4.5	Sistem dievaluasi dan dikategorikan baik, sehingga dapat disimpulkan bahwa sistem bersifat <i>user friendly</i> . Berdasarkan hasil penelitian, tingkat akurasinya rule – rule dalam sisitem ini melebihi 50%. Sehingga dapat digunakan sebagai dasar pengembangan sistem pendukung keputusan untuk Analisa pemberian kredit.
3.	SISTEM PENDUKUNG KEPUTUSAN PEMBERIAN KREDIT KOPERASI SERBA USAHA BERKAH TIRAM JAYA 4MENGGUNAKAN METODE <i>ANALYTIC</i>	• <i>Analytic Hierarchy Process</i> (AHP)	Sistem dapat membantu pihak koperasi untuk menentukan penerima kredit dengan mengambil nilai hasil akhir nasabah dengan prioritas tertinggi.

	<i>HIERARCHY PROCESS</i> (AHP) (Nurdiyanto and Minarto, 2017)		
4.	SISTEM PENDUKUNG KEPUTUSAN KREDIT USAHA RAKYAT PT. BANK RAKYAT INDONESIA UNIT KALIANGGRIK MAGELANG (Nugroho, Kusrini and Arief, 2015)	<i>K-Nearest Neighbors</i> (KNN)	Sistem dapat digunakan untuk menentukan baik, buruk, normal calon penerima kredit KUR. Penerapan algoritma KNN diperoleh tingkat <i>error rate</i> sebesar 6,98% dan 93,023% akurat.
5.	Klasifikasi Nasabah Thera Bank Membeli <i>Personal Loan</i> Menggunakan Metode Klasifikasi Dalam <i>Machine Learning</i> (Putri, 2018)	<ul style="list-style-type: none"> • <i>Random Forest</i> • <i>Naives-Bayes Classifier</i> • <i>K-Nearest Neighbors</i> • <i>Decision Tree</i> • <i>Random Forests</i> 	Dengan membandingkan beberapa metode klasifikasi dalam <i>machine learning</i> , didapatkan bahwa metode terbaik adalah <i>random forest</i> dengan <i>feature selection</i> dan <i>hyperparameter tuning</i> .
6.	Kategorisasi Teks pada Hadits Sahih Al-Bukhari menggunakan <i>Random Forest</i> (Afianto <i>et al.</i> , 2017)	<i>Random Forests</i>	Kategorisasi dokumen hadits sahih Al-Bukhari dengan menggunakan metode <i>Random Forests</i> dengan mekanisme <i>Preprocessing</i> berupa <i>Stemming</i> , <i>Case Folding</i> dan <i>Filtration</i> menghasilkan nilai <i>F1-Score</i> sebesar 90%.
7.	PREDIKSI LAMA STUDI MAHASISWA DENGAN METODE RANDOM FOREST (Budi Adnyana, 2016)	<i>Random Forests</i>	Hasil pengujian menunjukkan nilai <i>accuracy</i> sebesar 83.54%.
8.	An up-to-date comparison of state-of-the-art classification algorithms (Zhang <i>et al.</i> , 2017)	11 Algoritma	Dengan membandingkan 11 algoritma populer dalam machine learning, didapatkan 5 algoritma yang telah di

			urutkan berdasarkan peringkat tinggi akurasi, yakni: <ol style="list-style-type: none"> 1. GBDT 2. <i>Random Forests</i> 3. SVM 4. C4.5
--	--	--	---

Dari beberapa matrik penelitian diatas dapat dikatakan bahwa saat ini banyak penelitian terhadap analisis hingga klasifikasi terhadap kelayakan kredit, sistem pendukung keputusan serta penggunaan *machine learning* untuk keperluan klasifikasi hingga *data mining* dari berbagai macam akademisi.

Adapun penggunaan algoritma *random forest* merupakan salah satu algoritma yang populer digunakan dalam hal *classification* dan *clustering*. Berdasarkan penelitian sebelumnya, *random forest* memiliki tingkat akurasi yang tinggi. Sehingga, peneliti tertarik untuk menggunakan *random forests* dalam klasifikasi terhadap analisa kelayakan kredit pada bank.

2.2. Dasar Teori

Dasar teori merupakan teori yang relevan yang digunakan untuk menjelaskan tentang variabel yang akan diteliti.

2.2.1. Sistem Pendukung Keputusan (SPK)

Sistem pendukung keputusan adalah sistem yang dibangun untuk mendukung pengambilan keputusan dalam suatu organisasi perusahaan yang dirancang untuk pengembangan efektifitas dan produktifitas para manajer dengan bantuan teknologi komputer (Manurung, 2010)

Sebuah sistem keputusan merupakan sebuah model dari sistem yang terdiri atas keputusan yang bersifat tertutup atau terbuka. Sistem keputusan tertutup berarti keputusan dipisah dari masukan yang tidak diketahui lingkungannya, sedangkan sistem keputusan terbuka memandang keputusan sebagian berada dalam suatu lingkungan yang rumit dan sebagian tak diketahui. Keputusan dipengaruhi oleh lingkungan dan pada gilirannya proses keputusan kemudian mempengaruhi lingkungan.

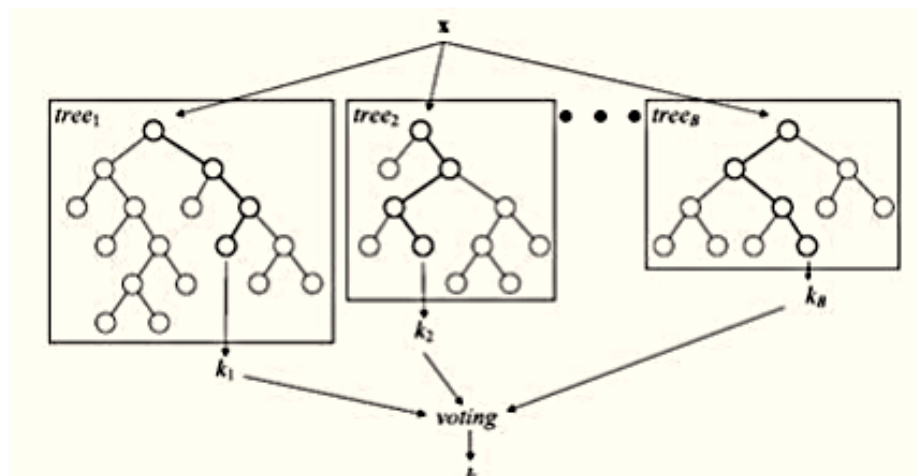
Dalam keputusan model tertutup, komputer bertindak sebagai sebuah alat penghitung untuk bisa menghitung hasil optimum. Dalam keputusan model terbuka, komputer bertindak sebagai pembantu bagi pengambilan keputusan dalam menghitung, menyimpan, mencari kembali, menganalisis data dan sebagainya. Perancangan tersebut memungkinkan manusia pengambil keputusan mengalokasikan tugas bagi dirinya atau pada komputer.

2.2.2. Klasifikasi

Klasifikasi merupakan proses untuk menempatkan suatu objek ke dalam suatu kategori / kelas yang telah terdefiniskan sebelumnya berdasarkan model tertentu. Data mining merupakan penjelasan tentang masa lalu dan prediksi masa depan berdasarkan analisa pada sekelompok data. Secara umum, proses klasifikasi dimulai dengan diberikannya sejumlah data yang menjadi acuan untuk membuat aturan klasifikasi data. Data ini biasanya disebut dengan *training sets*. Dari *training sets* tersebut, kemudian dapat dibuat suatu model untuk mengklasifikasikan data. Model tersebut kemudian digunakan sebagai acuan untuk mengklasifikasikan data – data yang belum diketahui kelasnya yang biasa disebut dengan *datasets set*. (Nugroho, Kusrini and Arief, 2015).

2.2.3. Algoritma Random Forest

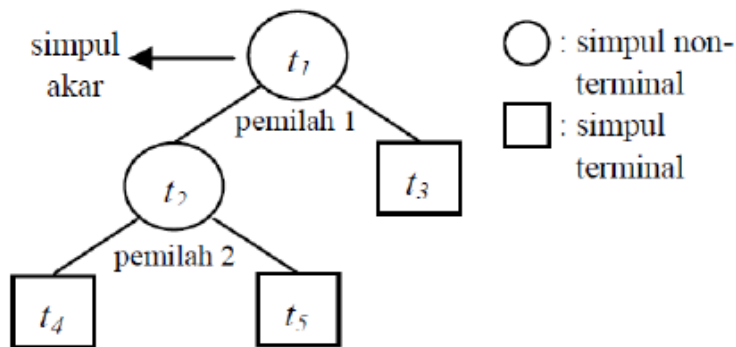
Random Forest diusulkan oleh Tin Kam Ho pada tahun 1995 dengan penelitiannya yang berjudul *Random Decision Forest* (Ho, 1995), yang kemudian dikembangkan Leo Breimann dan kemudian di patenkan pada tahun 2001 (Breiman, 2001). *Random Forest* merupakan salah satu metode *ensemble* yang bertujuan untuk meningkatkan akurasi suatu klasifikasi data dari sebuah pemilah tunggal yang tidak stabil melalui kombinasi banyak pemilah dari suatu metode yang sama dengan proses *voting* untuk memperoleh prediksi klasifikasi akhir (van Wezel and Potharst, 2007)



Gambar 2.1: Contoh struktur *Random Forest* (Budi Adnyana, 2016)

Random Forest adalah pengembangan lebih lanjut dari metode *Classification and Regression Tree* (CART) yang menerapkan metode *Bootstrap Aggregating* (*Bagging*) dan *Random Feature Selection*. Yaitu dengan melakukan modifikasi terhadap *bagging* yang membangun susunan pohon yang tidak berkorelasi lalu menghitung rata – ratanya. Dalam beberapa kasus, performa dari *Random Forest* mirip dengan *Boosting* namun lebih sederhana dan mudah untuk di *train* dan di *tune*. Sehingga, *Random Forests* menjadi salah satu metode klasifikasi populer dan di implementasikan di berbagai bidang (Hastie, Tibshirani and Friedman, 2001). CART sendiri merupakan metode eksplorasi data yang didasari pada teknik pohon keputusan. Pohon klasifikasi dihasilkan saat peubah respon berupa data kategorik, sedangkan pohon regresi dihasilkan saat peubah respons berupa data numerik. Untuk membangun pohon klasifikasi CART, meliputi tiga hal berikut yakni (Budi Adnyana, 2016):

1. Pemilihan pemilah (*Split*)
2. Penentuan simpul terminal
3. Penandaan label kelas.



Gambar 2.2: Struktur Pohon pada Metode CART (Budi Adnyana, 2016)

Dalam *Random Forest*, banyak pohon yang ditumbuhkan sehingga terbentuklah sebuah hutan (*forest*), kemudian dilakukan analisis pada kumpulan pohon tersebut. Pada gugus data yang terdiri atas n amatan dan p peubah penjelas, *Random Forest* dilakukan dengan cara:

1. Melakukan penarikan contoh acak berukuran n dengan pemulihan pada gugus data dimana tahapan ini merupakan tahapan *bootstrap*.
2. Dengan menggunakan contoh *bootstrap*. Pohon dibangun sampai mencapai ukuran maksimu (tanpa pemangkasan). Pada setiap simpul, pemilihan pemilah dilakukan dengan memilih m peubah penjelas secara acak, dimana $m \ll p$, lalu pemilah terbaik dipilih berdasarkan m peubah penjelas tersebut dimana tahapan ini disebut dengan tahapan *random feature selection*.
3. Ulangi langkah 1 dan 2 sebanyak k kali, sehingga terbentuklah sebuah hutan yang terdiri atas k pohon.

Untuk melakukan *splitting*, jika didalam *decision tree* menggunakan perhitungan terhadap *entropy* dan *gain*. Maka didalam metode *random forest* menggunakan *Gini index*. Yang mana sebelumnya juga dilakukan *feature selection*, yakni perhitungan untuk menentukan *feature* mana yang menjadi aturan dari *node* pohon keputusan. Untuk perhitungan *Gini index* terhadap *classification*, dilakukan perhitungan sebagai berikut (Breiman and Cutler, 2005):

$$Gini = N_L \sum_{k=1, \dots, K} p_{kL}(1 - p_{kL}) + N_R \sum_{k=1, \dots, K} p_{kR}(1 - p_{kR}) \quad (2.1)$$

Dimana: p_{kL} merupakan proporsi dari kelas k yang ada di node kiri dan

p_{kR} merupakan proporsi dari kelas k yang ada di node kanan.

Untuk mendapatkan hasil yang optimal, metode *random forest* harus menentukan m jumlah variabel predictor yang diambil secara acak dan k pohon yang akan dibentuk. Nilai k yang disarankan untuk digunakan pada metode *bagging* yang telah cobakan adalah $k = 50$ telah memberikan hasil memuaskan untuk klasifikasi (Breiman, 2001). sedangkan $k \geq 100$ cenderung menghasilkan tingkat misklasifikasi rendah (Sutton, 2004).

Respons suatu amatan diprediksi dengan menggabungkan (*Aggregating*) hasil prediksi k pohon. Pada masalah klasifikasi, dilakukan berdasarkan *Majority Vote* (suara terbanyak). Untuk menduga adanya kesalahan dalam klasifikasi *random forest*, dapat dilakukan dengan perolehan *error* OOB. Data OOB (*Out of bag*) adalah data yang tidak termuat dalam contoh *bootstrap*. Adapun perolehan *error* OOB dapat dilakukan dengan cara sebagai berikut (Budi Adnyana, 2016):

1. Melakukan prediksi terhadap setiap data OOB pada pohon yang bersesuaian. Data OOB (*Out of bag*) adalah data yang tidak termuat dalam contoh *bootstrap*.
2. Secara rata – rata, setiap amatan gugus data asli akan menjadi data OOB sekitar 36% dari banyak pohon. Untuk itu, pada langkah 1, masing – masing amatan gugus data asli mengalami sebanyak sekitar sepertiga kali dari banyak pohon. Jika a adalah sebuah amatan dari gugus data asli, maka hasil prediksi *random forest* terhadap a adalah gabungan dari hasil prediksi setiap kali a menjadi data OOB.
3. Error OOB dihitung dari proporsi misklasifikasi hasil prediksi *random forest* dari seluruh amatan gugus asli.

Disarankan agar mengamati *error* OOB saat k kecil, lalu memilih m yang menghasilkan *error* OOB terkecil. Jika *random forest* dilakukan dengan menghasilkan *variable importance*, disarankan untuk menggunakan banyak pohon, misalnya 1000 pohon atau lebih. Jika peubah penjelas yang dianalisis sangat

banyak, nilai tersebut dapat lebih besar agar *variable importance* yang dihasilkan semakin stabil (Breiman and Cutler, 2005).

Dengan diterapkannya metode *random forest* pada permasalahan untuk rekomendasi kelayakan kredit, diharapkan dapat membantu pihak Bank dalam meninjau calon pemohon kredit untuk menekan dan mengurangi adanya *bad loan* di Bank BRI kantor cabang pelaihari untuk periode berikutnya.

2.2.4. Confusion Matrix dan Accuracy

Confusion matrix adalah suatu alat yang berfungsi untuk melakukan pengukuran ketika menganalisis sebuah *classifier*. Apakah *classifier* tersebut baik dalam mengenali *tuple* dari kelas yang berbeda. Nilai dari *True-Positive* dan *True-Negative* memberikan informasi ketika *classifier* melakukan klasifikasi dan memiliki data yang bernilai benar. Sedangkan *False-Positive* dan *False-Negative* memberikan informasi ketika *classifier* memiliki kesalahan dalam melakukan klasifikasi data (Han, Kamber and Pei, 2011).

Adapun bentuk *confusion matrix* serta penjelasannya dapat dilihat pada gambar 2.3:

		Predicted class		
		yes	no	Total
Actual class	yes	TP	FN	P
	no	FP	TN	N
Total		P'	N'	P + N

Gambar 2.3: *Confusion Matrix* menampilkan total *positive* dan *negative tuple* (Han, Kamber and Pei, 2011).

1. TP (*True Positive*), merupakan jumlah data dengan nilai sebenarnya positif dan nilai prediksi positif.
2. FP (*False Positive*), merupakan jumlah data dengan nilai sebenarnya negatif dan nilai prediksi positif.
3. FN (*False Negative*), merupakan jumlah data dengan nilai sebenarnya positif dan nilai prediksi negatif.
4. TN (*True Positive*), merupakan jumlah data dengan nilai sebenarnya negatif dan nilai prediksi negatif.

Confusion Matrix dapat digunakan untuk mengukur sebuah evaluasi, misalnya mengukur *accuracy*. *Accuracy* yang dimaksudkan adalah presentase keakuratan dari set tes *tuple* yang diklasifikasikan dengan benar oleh *classifier*. Adapun rumusnya sebagai berikut:

$$accuracy = \frac{TP + TN}{P + N} \quad (2.2)$$

Perhitungan untuk *accuracy* juga dapat digunakan dengan menambahkan pengukuran *sensitivity* dan *specificity*. Dimana *sensitivity* juga di anggap sebagai *True-Positive* (proporsi dari *tuple* positif yang telah diidentifikasi dengan benar) dan *specificity* dianggap sebagai *True-Negative* (proporsi dari *tuple* negatif yang telah diidentifikasi dengan benar). Rumus dari *sensitivity*, *specificity* dan *accuracy* dengan menambahkan dua pengukuran tersebut adalah sebagai berikut:

$$sensitivity = \frac{TP}{P} \quad (2.3)$$

$$specificity = \frac{TN}{N} \quad (2.4)$$

$$accuracy = sensitivity \frac{P}{(P + N)} + specificity \frac{N}{(P + N)} \quad (2.5)$$

2.2.5. Analisa Kredit

Dalam upaya memperkecil risiko dalam memberikan pinjaman (kredit), bank terlebih dahulu mempertimbangkan beberapa hal yang terkait dengan itidak baik (*willingness to pay*) dan kemampuan membayar (*ability to pay*) calon nasabah untuk melunasi kembali pinjaman beserta bunganya. Hal – hal tersebut tertuang dalam istilah “*The Five C of Credit Analysis*”, sebagai berikut (Pandie, 2012):

1. *Character* atau aspek karakter, memuat data tentang penilaian terhadap karakter calon debitur dilihat dari berbagai sifat misalnya watak, kemauan, kejujuran dan pengalaman hutang masa lalu.
2. *Capacity* atau aspek kemampuan, memuat data tentang penilaian terhadap kemampuan berusaha, kemampuan pasaran, kemampuan membayar kembali hutangnya masa lalu dan hubungan dengan rekan usahanya.

3. *Capital* atau aspek permodalan, memuat data tentang penilaian terhadap data keuangan calon debitur yang meliputi harta lancar, harta tetap, hutang dan sebagainya.
4. *Conditions of Economy* atau aspek kondisi ekonomi, memuat data tentang penilaian kondisi usaha, kondisi rumah tangga, kondisi usaha yang berkaitan dengan kondisi ekonomi pada umumnya, dan sebagainya.
5. *Collateral* atau aspek jaminan, memuat data tentang agunan yang akan disediakan kebendaannya, keberadaannya, kondisi jaminannya, nilai jualnya, penilaian terhadap kelayakannya, dan sebagainya.

2.2.6. Database

Database secara umum merupakan koleksi dari data – data yang terorganisasi sedemikian rupa sehingga dapat dengan mudah disimpan dan dimanipulasi (diperbarui, dicari, diolah dengan perhitungan tertentu, serta dihapus). Data merupakan fakta – fakta tentang segala sesuatu yang dapat direkam dan disimpan pada media komputer. Pada saat ini, data tidak hanya mengandung text saja, melainkan objek seperti dokumen, gambar hingga suara (Qomary, 2018).

2.3. Integrasi Keilmuan

Kredit merupakan salah satu jasa yang disediakan oleh bank. Kredit merupakan istilah lain dari hutang. Bedanya, pengajuan kredit kepada bank dapat dilakukan dengan berbagai cara, ada yang mengajukan kredit dengan melakukan akad terlebih dahulu, mengajukan kredit dengan mengajukan jaminan (menggadaikan). Dan yang jelas, hutang piutang harus dicatat dan nantinya harus di bayarkan kembali seperti perjanjian awal.

Sebagaimana yang tertulis dalam Al-Qur'an surah Al-Baqarah ayat 282 yang memiliki arti :

Hai orang-orang yang beriman, apabila kamu bermu'amalah tidak secara tunai untuk waktu yang ditentukan, hendaklah kamu menuliskannya. Dan hendaklah seorang penulis di antara kamu menuliskannya dengan benar. Dan janganlah penulis enggan menuliskannya sebagaimana Allah mengajarkannya, meka hendaklah ia menulis, dan hendaklah orang yang berhutang itu

mengimlakkan (apa yang akan ditulis itu), dan hendaklah ia bertakwa kepada Allah Tuhannya, dan janganlah ia mengurangi sedikitpun daripada hutangnya. Jika yang berhutang itu orang yang lemah akalnya atau lemah (keadaannya) atau dia sendiri tidak mampu mengimlakkan, maka hendaklah walinya mengimlakkan dengan jujur. Dan persaksikanlah dengan dua orang saksi dari orang-orang lelaki (di antaramu). Jika tak ada dua orang lelaki, maka (boleh) seorang lelaki dan dua orang perempuan dari saksi-saksi yang kamu ridhai, supaya jika seorang lupa maka yang seorang mengingatkannya. Janganlah saksi-saksi itu enggan (memberi keterangan) apabila mereka dipanggil; dan janganlah kamu jemu menulis hutang itu, baik kecil maupun besar sampai batas waktu membayarnya. Yang demikian itu, lebih adil di sisi Allah dan lebih menguatkan persaksian dan lebih dekat kepada tidak (menimbulkan) keraguanmu. (Tulislah mu'amalahmu itu), kecuali jika mu'amalah itu perdagangan tunai yang kamu jalankan di antara kamu, maka tidak ada dosa bagi kamu, (jika) kamu tidak menulisnya. Dan persaksikanlah apabila kamu berjual beli; dan janganlah penulis dan saksi saling sulit menyulitkan. Jika kamu lakukan (yang demikian), maka sesungguhnya hal itu adalah suatu kefasikan pada dirimu. Dan bertakwalah kepada Allah; Allah mengajarmu; dan Allah Maha Mengetahui segala sesuatu.(Qs. Al-Baqarah : 282)

Dari ayat dan artinya di atas dijelaskan bahwa sebagai manusia dan ingin melakukan utang piutang dengan seseorang, hendaknya dituliskan segala utang piutangnya baik yang kecil maupun yang besar. Tak lupa menuliskan pula jumlah, dan tempo pembayarannya agar menghindari ketidak adilan dan keraguan. Selanjutnya pula dijelaskan hendaknya adanya seorang juru tulis pula yang menuliskan perihal utang piutang tersebut dengan tidak merusak sedikitpun dari perjanjian dan jumlah utang yang telah dikatakannya. Juru tulis tersebut adalah orang adil yang tidak memihak sebelah pihak. Hendaknya pemberi utang dan penerima utang mengutarakan maksudnya satu sama lain agar ditulis oleh juru tulis dan tidak mengurangi sedikitpun hak orang lain demi kepentingan pribadi (Al-Maraghi, 1986).

Adapun maksud dari tafsir ayat di atas apabila di integrasikan dengan penelitian ini adalah pencatatan utang yang dapat disamakan dengan *database*.

Database pada sistem rekomendasi kelayakan kredit ini di asumsikan mencatat data dari para debitur yang mengajukan pinjaman yang nantinya akan diproses terlebih dahulu oleh sistem lalu disetujui oleh pihak atasan. Data debitur akan tersimpan terlebih dahulu didalam *database* sehingga dapat di simpulkan bahwa penyimpanan data calon debitur pada *database* tersebut merupakan salah satu upaya pencatatan utang piutang yang telah di jelaskan pada ayat diatas.

Adapula ayat yang membahas tentang segera membereskan hutang. Yang dimaksudkan adalah apabila telah diikat perjanjian hutang untuk jangka waktu tertentu, maka wajiblah janji itu ditepati dan pihak yang berhutang perlu membereskan hutangnya menurut perjanjian tersebut (Ya'qub, 1992). Menepati janji merupakan hal yang wajib dan setiap orang sangat dipertanggung jawabkan akan janji-janjinya. Sebagaimana firman Allah yang tertulis dalam Al-Qur'an surah Al-Isra ayat 34 yang memiliki arti:

Dan janganlah kamu mendekati harta anak yatim, kecuali dengan cara yang lebih baik (bermanfaat) sampai ia dewasa dan penuhilah janji; sesungguhnya janji itu pasti diminta pertanggungan jawabnya.(QS Al Isra : 34)

Ayat tersebut apabila berfokus pada potongan ayat terakhir, disebutkan bahwa manusia wajib memenuhi janji karena janji nantinya akan diminta pertanggung jawabannya. Dan apabila di integrasikan ke-ilmuannya dengan penelitian ini, penulis mencoba untuk membangun sebuah sistem pendukung keputusan yang dapat memberikan keputusan apakah seseorang layak dan mampu mengemban tanggung jawab apabila diberikan kredit atau utang berdasarkan kemampuan jasmani dan rohani seseorang tersebut.

Diasumsikan bahwa dengan adanya penelitian ini nantinya dapat dilakukan pencatatan utang piutang oleh seseorang dalam hal ini pencatatan data calon debitur yang mengajukan pinjaman kedalam *database*. Serta menemukan prediksi untuk keputusan terhadap sebuah pinjaman dari calon debitur yang bertanggung jawab.

BAB III

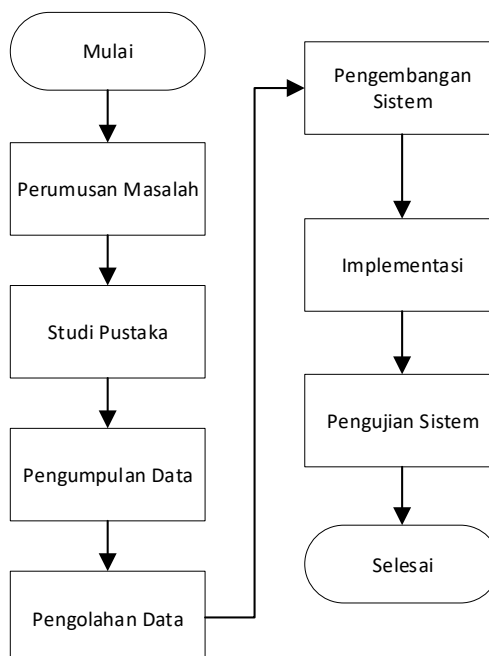
METODOLOGI PENELITIAN

Sistem rekomendasi kelayakan kredit ini adalah alat bantu untuk memudahkan pihak bank dalam peninjauan calon nasabah atau debitur yang sedang melakukan pengajuan, apakah nantinya calon nasabah atau debitur tersebut dapat dikategorikan menjadi *good loan* atau *bad loan*.

Sistem rekomendasi kelayakan ini berisi beberapa parameter yang telah ditentukan yang sebelumnya dianalisis terlebih dahulu dalam analisis kredit 5C. sehingga bertujuan untuk meminimalisir pihak bank mendapati *bad loan* dan meningkatkan *good loan*.

3.1. Desain Penelitian

Adapun desain penelitian dideskripsikan menggunakan alur *flowchart*. Hal ini dilakukan oleh peneliti bertujuan agar mempermudah alur penelitian, berikut alur dari penelitian ini:



Gambar 3.1: Alur Penelitian

Berikut adalah pembahasan dari proses alur metodologi yang telah tergambarakan diatas:

3.1.1. Perumusan Masalah

Pada tahap ini, penulis merumuskan masalah yang telah menjadi latar belakang untuk membuat penelitian ini. Masalah yang di angkat pada penelitian ini seperti halnya tertera dalam latar belakang yakni mengenai rekomendasi kelayakan kredit. Rekomendasi kelayakan kredit merupakan salah satu upaya dari pihak bank untuk menghitung resiko kredit dari calon nasabah atau debitur baru. Dengan melakukan metode klasifikasi data mining dari data kredit yang sudah ada, sehingga setelah melakukan perhitungan pada data pengajuan yang baru didapatkan sebuah hasil apakah nanti pengajuan tersebut akan masuk *good loan* atau *bad loan*. Yang nantinya keputusan lebih lanjut masih berada di tangan pihak yang terkait dalam hal ini manajer.

3.1.2. Studi Pustaka

Adapun dalam tahapan ini penulis melakukan studi pustaka yakni melakukan pendekatan lebih lanjut terhadap klasifikasi dalam data mining. Didapatkan hasil untuk melakukan klasifikasi dalam data mining ada beberapa metode yang populer diantara nya *KNN*, *Random Forest*, *Logistic Regression*, *Decision Tree*, dll. Setelah melakukan perbandingan terhadap studi pustaka lebih lanjut, didapatkan bahwa *Random Forest* merupakan metode yang populer dan memiliki akurasi prediksi yang paling tinggi dibandingkan metode klasifikasi lainnya dalam data mining. Sehingga, dalam penelitian ini penulis menggunakan metode *Random Forest* untuk proses klasifikasi.

3.1.3. Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini dengan menggunakan metode – metode yang diharapkan dapat mengumpulkan data secara terperinci dan jelas, metode yang digunakan adalah:

1. Melakukan wawancara kepada pihak Bank BRI Kantor Cabang Pelaihari, dalam hal ini melakukan dengan Asisten Manajer untuk mengetahui permasalahan apa saja yang harus diperhatikan dalam proses penentuan

kelayakan kredit pada calon debitur, memperoleh deskripsi kriteria untuk penentuan kelayakan kredit, memperoleh data mengenai kategori calon debitur.

2. Melakukan studi literatur lebih lanjut dalam memahami metode *Random Forest* dalam proses penentuan kelayakan kredit calon debitur baru berdasarkan kriteria yang ada.
3. Melakukan penghitungan terhadap data sekunder. Data sekunder yang dimaksud adalah data yang telah didapatkan dari *database* Bank BRI Kantor Cabang Pelaihari.

Adapun data sekunder yang dapat di ambil dari Bank BRI Kantor Cabang Pelaihari memiliki atribut dan di deskripsikan sebagai berikut:

Tabel 3.1 Deskripsi Atribut Data

No	Nama Data	Jenis Data	Deskripsi
1.	Usia	Kuantitatif	Digunakan untuk klasifikasi dalam perihal segmentasi usia. Contoh: '>60', '<60', dan '≤50'
2.	Alamat	Kualitatif	Digunakan untuk klasifikasi dalam perihal segmentasi daerah asal dari debitur. Contoh: 'Bati – Bati', 'Sungai Kiram', 'Pelaihari'
3.	Jenis Pekerjaan	Kualitatif	Digunakan untuk klasifikasi dalam perihal segmentasi jenis pekerjaan debitur. Contoh: 'Swasta', 'Wiraswasta'
4.	Penghasilan Perbulan	Kuantitatif	Digunakan untuk klasifikasi dalam perihal penghasilan dari debitur. Contoh: '>15.000.000', '≤30.000.000'
5.	Status Pernikahan	Kualitatif	Digunakan untuk klasifikasi dalam perihal segmentasi status pernikahan dari debitur. Contoh: 'Menikah', 'Belum Menikah'
6.	Jumlah Tanggungan	Kuantitatif	Digunakan untuk klasifikasi dalam perihal berapa jumlah kepala yang ditanggung oleh debitur. Contoh: '4', '3', '2'
7.	Status Checking BI	Kualitatif	Digunakan untuk klasifikasi dalam perihal status BI Checking debitur.

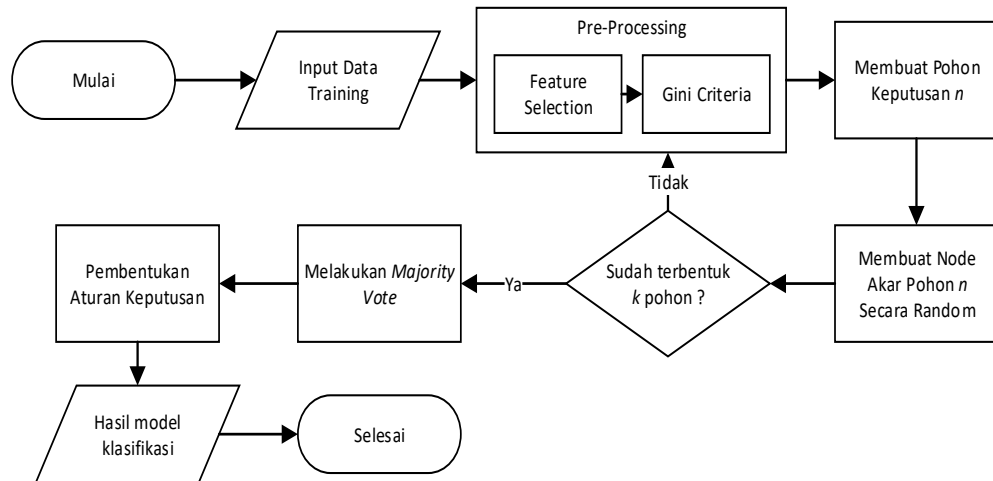
			Contoh: 'Normal', 'Macet', 'Lancar'
8.	Status Sebagai Nasabah	Kualitatif	Digunakan untuk klasifikasi dalam perihal status apakah debitur sudah menjadi nasabah BRI atau belum. Contoh: 'Aktif', 'Tidak Aktif'
9.	Tujuan Kredit / Jenis Kredit / Sektor Kredit	Kuantitatif	Digunakan untuk klasifikasi dalam perihal tujuan dari kredit atau pada sektor apa kredit tersebut diajukan oleh debitur. Contoh: 'Tambang', 'Usaha', 'KUR'
10.	Jumlah Kredit Yang Di Ajukan	Kuantitatif	Digunakan untuk klasifikasi dalam perihal jumlah kredit yang akan di ajukan debitur. Contoh: '>100.000.000', '≤100.000.000'
11.	Jangka Waktu Kredit	Kuantitatif	Digunakan untuk klasifikasi dalam perihal jangka waktu atau tenor yang akan diajukan debitur. Contoh: '12 bulan', '36 bulan', '48 bulan'
12.	Jaminan	Kualitatif	Digunakan untuk klasifikasi dalam perihal bentuk jaminan yang di ajukan oleh debitur. Contoh: 'Sertifikat Rumah', 'Sertifikat Tanah'
13.	Status Pinjaman	Kualitatif	Digunakan untuk klasifikasi dalam perihal status terkini dari pinjaman debitur. Contoh: 'Lancar', 'Macet'

3.1.4. Pengolahan Data

Dengan data yang telah terkumpul lalu diproses melalui tahap pengolahan data. Adapun langkah – langkah dalam pengolahan data yakni sebagai berikut:

1. Hasil studi literatur akan digunakan sebagai acuan dalam proses perhitungan metode *Random Forest* dalam membuat keputusan penentuan kelayakan kredit kepada calon debitur.
2. Melakukan analisa dari data yang diperoleh melalui wawancara mengenai proses bisnis dan permasalahan dalam penentuan kelayakan kredit kepada calon debitur.
3. Melakukan analisa dari data yang diperoleh melalui wawancara mengenai deskripsi kriteria yang digunakan dalam penentuan kelayakan kredit kepada calon debitur.

4. Melakukan perhitungan dan analisa terhadap *history* data nasabah yang telah didapat dari bank dengan metode *Random Forest*. Yang digambarkan dalam alur sebagai berikut:



Gambar 3.2: Flowchart Pembuatan Pohon pada Metode Random Forests dalam tahap *training*

3.1.5. Pengembangan Sistem

Pengembangan sistem untuk penentuan kelayakan kredit ini menggunakan teknik atau metode *waterfall*. Adapun terdapat lima proses tahapan yang harus dilalui oleh metode *waterfall* (Safitri and Supriyadi, 2015):

1. Analisis Kebutuhan
2. Desain Sistem
3. Implementasi
4. *Testing*
5. Perawatan

3.1.6. Pengujian Sistem

Untuk pengujian sistem, sistem akan diuji dengan melakuka perbandingan hasil dari sistem yang telah dibuat dengan data *real* yang telah didapat. Lalu dengan menggunakan metode pengujian *confusion matrix* dan perhitungan *accuracy*.

3.2. Tempat dan Waktu Penelitian

Penelitian ini dilaksanakan dari April 2018 hingga Juli 2018. Lokasi yang digunakan dalam penelitian ini adalah Bank BRI Kantor Cabang Pelaihari.

Adapun alur waktu dalam penelitian ini dijabarkan dalam tabel *timeline* seperti berikut:

No	Kegiatan	April				Mei				Juni				Juli			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Perumusan Masalah																
2	Studi Pustaka																
3	Pengumpulan Data																
4	Pengolahan Data																
5	Pengembangan Sistem																
6	Implementasi																
7	Pengujian Sistem																
8	Kesimpulan dan Pelaporan																

DAFTAR PUSTAKA

- Afianto, M. F. *et al.* (2017) 'Kategorisasi Teks pada Hadits Sahih Al-Bukhari menggunakan Random Forest', 4(3), pp. 4874–4881.
- Al-Maraghi, A. M. (1986) *Tarjamah Tafsir Al-Maraghi*.
- Breiman, L. (2001) 'Random forests', pp. 1–33.
- Breiman, L. and Cutler, A. (2005) *Random Forests*. Available at: <https://www.stat.berkeley.edu/users/breiman/RandomForests/> (Accessed: 15 April 2019).
- Budi Adnyana, I. M. (2016) 'Prediksi Lama Studi Mahasiswa Dengan Metode Random Forest (Studi Kasus: Stikom Bali)', *CSRID (Computer Science Research and Its Development Journal)*, 8(3), pp. 201–208. doi: 10.22303/csrid.8.3.2016.201-208.
- Han, J., Kamber, M. and Pei, J. (2011) *Data Mining Concept and Techniques, Data Mining: Concepts and Techniques*. doi: 10.1016/B978-0-12-381479-1.00001-0.
- Hastie, T., Tibshirani, R. and Friedman, J. (2001) 'The Elements of Statistical Learning The Elements of Statistical Learning'. Available at: <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>.
- Ho, T. K. (1995) 'Random Decision Forest', 47, pp. 4–5.
- Manurung, P. (2010) *Sistem Pendukung Keputusan Seleksi Penerima Beasiswa Dengan Metode AHP dan TOPSIS*. Universitas Sumatera Utara.
- Nugroho, A., Kusriani, K. and Arief, M. R. (2015) 'Sistem Pendukung Keputusan Kredit Usaha Rakyat PT. Bank Rakyat Indonesia Unit Kiangkrik Magelang', *Creative Information Technology Journal*, 2(1), p. 1. doi: 10.24076/citec.2014v2i1.33.
- Nurdiyanto, H. and Minarto, S. Y. (2017) 'Pada Koperasi Serba Usaha Berkah Tiram Jaya Menggunakan Metode Analytic Hierarchy Process (Ahp)'.
'
- Octabriyantiningtyas, D. (2016) *Sistem Pendukung Keputusan Kelayakan Pemberian Kredit Pada Bank Tabungan Negara (BTN) Menggunakan Algoritma C4.5*. Universitas Airlangga.
- Pandie, E. S. Y. (2012) *Sistem Informasi Pengambilan Keputusan Pengajuan Kredit Dengan Algoritma K-Nearest Neighbour (Studi Kasus: Koperasi Simpan Pinjam)*. Universitas Diponegoro.
- Putri, C. B. (2018) *Klasifikasi Nasabah Thera Bank Membeli Personal Loan Menggunakan Metode Klasifikasi Dalam Machine Learning Pendahuluan Metodologi Penelitian*. Institut Teknologi Sepuluh Nopember.

- Qomary, L. N. (2018) *Rancang Bangun Sistem Pendukung Keputusan Pemilihan Program Latihan Fitnes Metode Case Based Reasoning*.
- Safitri, S. T. and Supriyadi, D. (2015) 'Rancang Bangun Sistem Informasi Praktek Kerja Lapangan Berbasis Web dengan Metode Waterfall', pp. 3–8.
- Sutton, C. D. (2004) 'Classification and Regression Trees, Bagging, and Boosting', *Handbook of Statistics*, 24(04), pp. 303–329. doi: 10.1016/S0169-7161(04)24011-1.
- van Wezel, M. and Potharst, R. (2007) 'Improved customer choice predictions using ensemble methods', *European Journal of Operational Research*, 181(1), pp. 436–452. doi: 10.1016/j.ejor.2006.05.029.
- Widiastuti, J. (2018) *Klasifikasi Pembiayaan Warung Mikro Menggunakan Metode Random Forest Dengan Teknik Sampling Kelas Imbalanced*. Universitas Islam Indonesia.
- Ya'qub, H. (1992) *Kode Etik Dagang Menurut Islam (Pola Pembinaan Hidup dalam Berekonomi)*.
- Zhang, C. *et al.* (2017) 'An up-to-date comparison of state-of-the-art classification algorithms', *Expert Systems with Applications*. Elsevier Ltd, 82, pp. 128–150. doi: 10.1016/j.eswa.2017.04.003.