

BAB III

METODE PENELITIAN

3.1. Bahan Penelitian

Bahan yang digunakan dalam penelitian ini adalah data yang diperoleh dari salah satu perusahaan *multifinance* cabang Karawang, yaitu data kredit motor baru, motor bekas, elektronik dan pinjaman dana tahun 2020 yang berisi nama, alamat, jenis pengajuan kredit, penghasilan perbulan, dan atribut lain yang mendukung untuk prediksi kelayakan pemberian kredit berupa *softfile* dengan format *spreadsheet*.

3.2. Peralatan Penelitian

Penelitian ini dilakukan dengan menggunakan peralatan berupa perangkat keras dan perangkat lunak.

1. Perangkat keras yang digunakan dalam penelitian ini adalah:
 - a. Laptop
 - b. *Processor* N2920 (2M Cache, 2.0 GHz)
 - c. 4 GB RAM DDR3
 - d. *Hardisk Drive* 500 GB.
2. Perangkat lunak yang digunakan pada laptop antara lain:
 - a. Perangkat lunak Sistem operasi laptop, sebagai sistem yang digunakan dalam melakukan penelitian.
 - b. Perangkat lunak *Document* 2016, digunakan untuk penyusunan laporan.
 - c. Perangkat lunak *Spreadsheet* 2016, digunakan untuk memilih data yang akan digunakan menjadi dataset.
 - d. *Flatform* RapidMiner Studio, digunakan untuk mencocokkan data yang telah dihitung.
 - e. Python versi 3.6.8, digunakan untuk menulis kode program.

3.3. Waktu Penelitian dan Tempat Penelitian

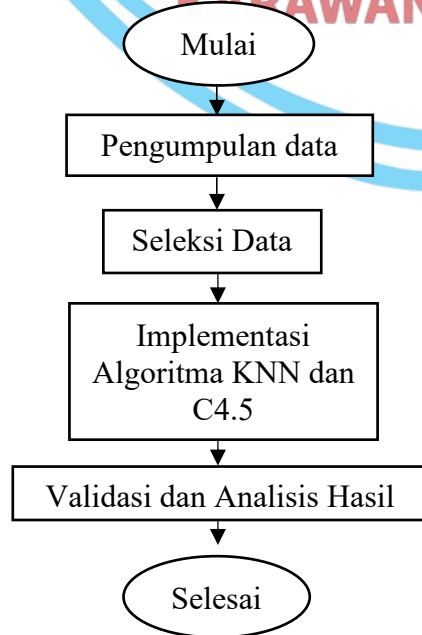
Lokasi penelitian dilakukan di salah satu perusahaan *multifinance* cabang Karawang sejak bulan Januari 2021. Adapun waktu penelitian terdapat pada tabel di bawah ini :

Tabel 3. 1 Waktu Penelitian

Kegiatan	2021																							
	Januari				Februari				Maret				April				Mei				Juni			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Analisis Kebutuhan																								
Pengumpulan data																								
Seleksi data																								
Pengelolaan data																								
Pengujian Hasil																								
Pembuatan Laporan																								

3.4. Prosedur Penelitian

Dari permasalahan yang dihadapi perusahaan yaitu kredit macet. Untuk itu dibutuhkan komputasi yang dapat membantu meminimalisasir risiko kredit macet. Maka alur penelitian yang akan dilakukan adalah seperti pada gambar berikut :



Gambar 3. 1 Alur Penelitian

1 Pengumpulan Data

Pada tahap ini proses yang dilakukan adalah wawancara di perusahaan *multifinance* cabang Karawang, bertujuan melakukan pengumpulan data untuk bahan penelitian. Data yang dikumpulkan berupa *softfile* dengan format spreadsheet yaitu data nasabah dari jenis data kredit motor baru, motor bekas, barang elektronik dan pinjaman dana pada tahun 2020.

2 Seleksi Data

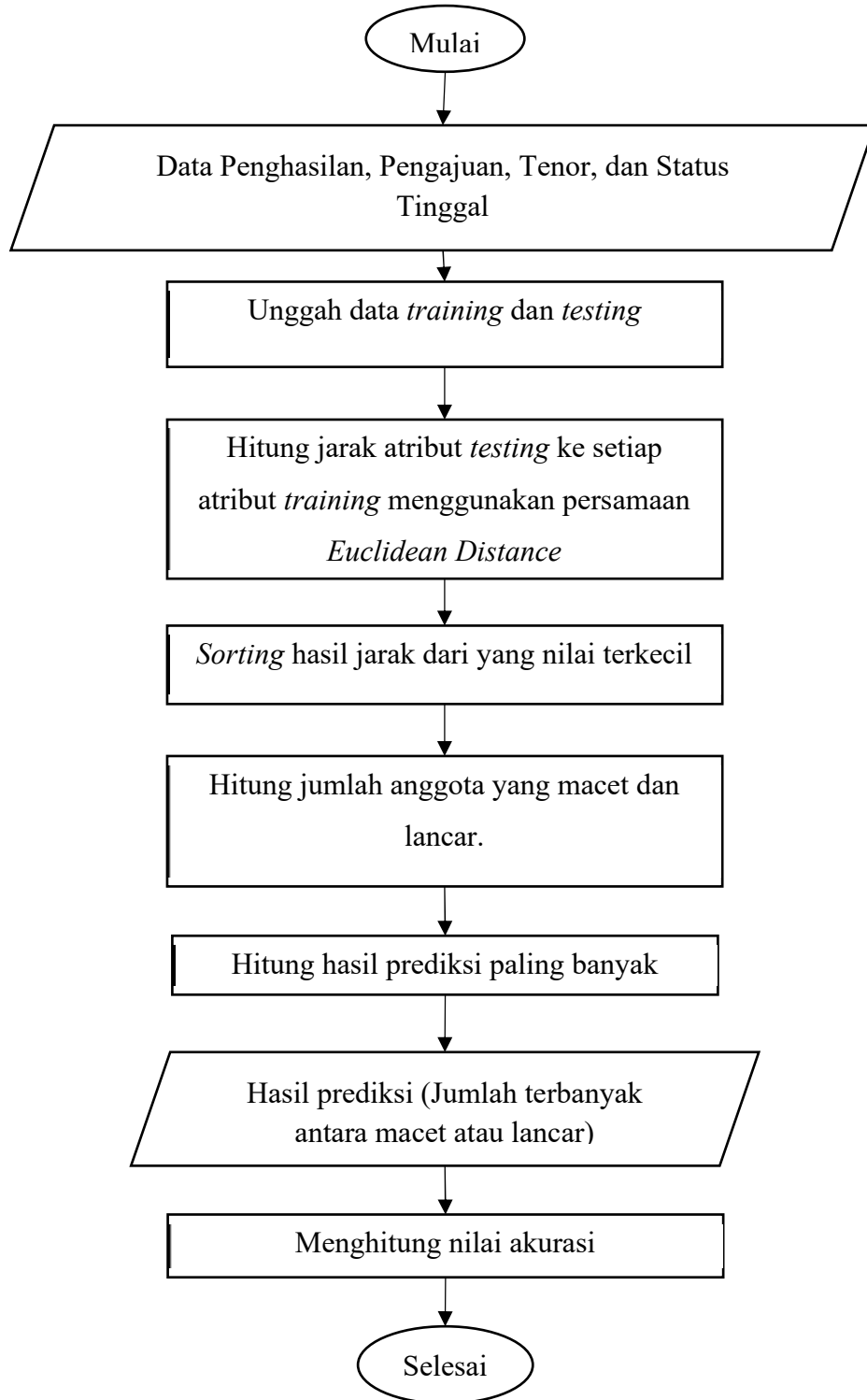
Proses seleksi data yaitu membersihkan data dari atribut – atribut yang tidak diperlukan, data yang tidak diperlukan yaitu atribut kode kantor, nomor kontrak dan jumlah angsuran. Setelah tahap penyeleksian dilakukan, maka akan menghasilkan dataset yang akan segera diolah. Hasil seleksi data adalah data nama customer, alamat, pekerjaan, penghasilan dan *object*/jenis kredit. Contoh dari hasil seleksi data bisa dilihat pada tabel berikut :

Tabel 3. 2 Contoh seleksi data penelitian

Nama Cust	Alamat	Pekerjaan	Penghasilan	Obj
ADE RIDWAN	BAKAN LIO RT 029 RW 012. Kel Karyasari Kec Rengasdengklok Kabupaten Karawang	JASA	3890030	NMC
ARIP ROHMAN	DUSUN BAYUR I RT 13 RW 04 Kel Payungsari Kec Pedes Kab Karawang	BURUH/PRT	1322353	MPF
ONIN	KP SAYURAN RT 05 RW 03 Kel kertasari Kab. Bekasi	PEDAGANG	10505428	NMC
SAAN	DUSUN KRAJAN RT 02 RW 01 Kel Tambaksumur, Tirtajaya, Karawang	PETANI	7815815	NMC

3 Implementasi Algoritma KNN dan C.45

Tahapan perhitungan algoritma KNN dapat digambarkan dengan *Flowchart* seperti pada Gambar 3.2.



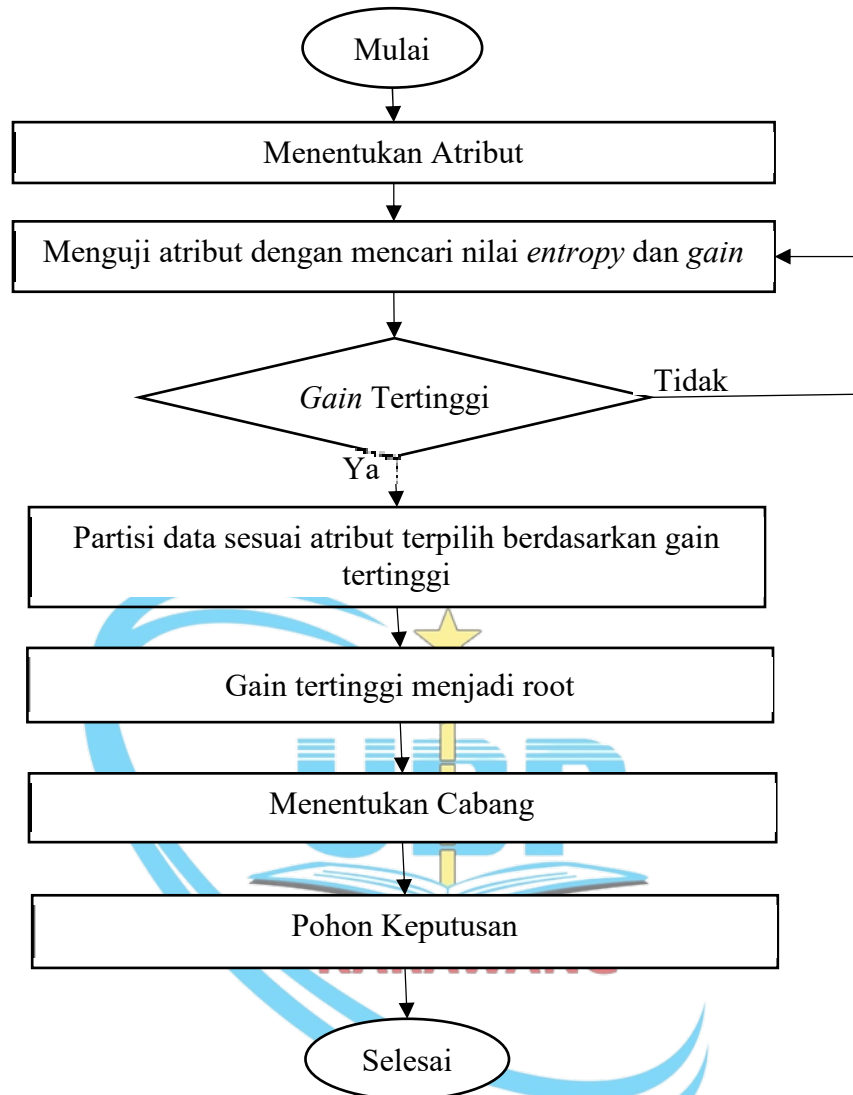
Gambar 3. 2 Flowchart KNN

Pada proses ini yaitu mengimplementasikan algoritma KNN untuk perhitungan menggunakan *platform* RapidMiner dan Sistem yang dibangun menggunakan program Python. Selain mengimplementasikan pada *platform*, pada tahap ini juga akan dilakukan perhitungan secara manual dengan menggunakan rumus *Euclidean Distance*.

Gambar 3.2 menjelaskan bahwa dalam proses perhitungan algoritma KNN yang pertama dilakukan adalah *input* data berdasarkan atribut yang telah ditentukan (Penghasilan, Pengajuan, Tenor, dan Status Tinggal). Kemudian menentukan data *testing* dan *training* yang mana data *training* diambil dari datasets yang digunakan untuk perhitungan dalam algoritma sedangkan data *testing* didapatkan dari hasil perhitungan, lalu dihitung menggunakan persamaan *Euclidean Distance*. Selanjutnya hasil dari perhitungan di urutkan dari nilai terendah, dan untuk menentukan prediksinya hitung jumlah yang paling banyak antara lancar atau macet berdasarkan jumlah tetangga terdekat yang telah ditentukan. Untuk tahap akhir adalah menghitung nilai akurasi, agar bisa mengetahui nilai persentase dari prediksi menggunakan algoritma KNN.



Pada proses implementasi algoritma C4.5 juga sama seperti KNN, yaitu perhitungan menggunakan *platform* RapidMiner dan bahasa pemrograman python. Alur proses dari algoritma C4.5 dimulai dari menentukan atribut yang digunakan. Kemudian melakukan pengujian atribut dengan cara mencari *gain* tertinggi, berdasarkan perhitungan *entropy* dari masing masing atribut. Apabila ditemukan *gain* tertinggi, maka *gain* tersebut akan menjadi *root* awal. Selanjutnya dilakukan penentuan cabang dengan cara yang sama, yaitu dengan melihat dari *gain* tertinggi dari hasil setiap partisi. Kemudian pembuatan pohon keputusan, dan pada pohon keputusan tersebut sudah diketahui hasil dari prediksi macet dan lancar. perhitungan manual dari algoritma C4.5 digambarkan menggunakan *Flowchart* seperti pada Gambar 3.3 berikut.



Gambar 3. 3 Flowchart C4.5

4 Validasi dan Analisis Hasil

Proses validasi ini diambil dari hasil akurasi yang telah didapatkan dari algoritma KNN dan C4.5. Kemudian algoritma yang mempunyai nilai akurasi yang tinggi bisa dijadikan patokan oleh perusahaan untuk menentukan kelayakan kredit. Untuk mendapatkan nilai akurasi dalam memprediksi kredit macet, digunakan persamaan sebagai berikut:

$$Akurasi = \frac{Jumlah\ Keputusan\ yang\ Sesuai}{Jumlah\ Data\ yang\ Diuji} \times 100\% \quad (4)$$