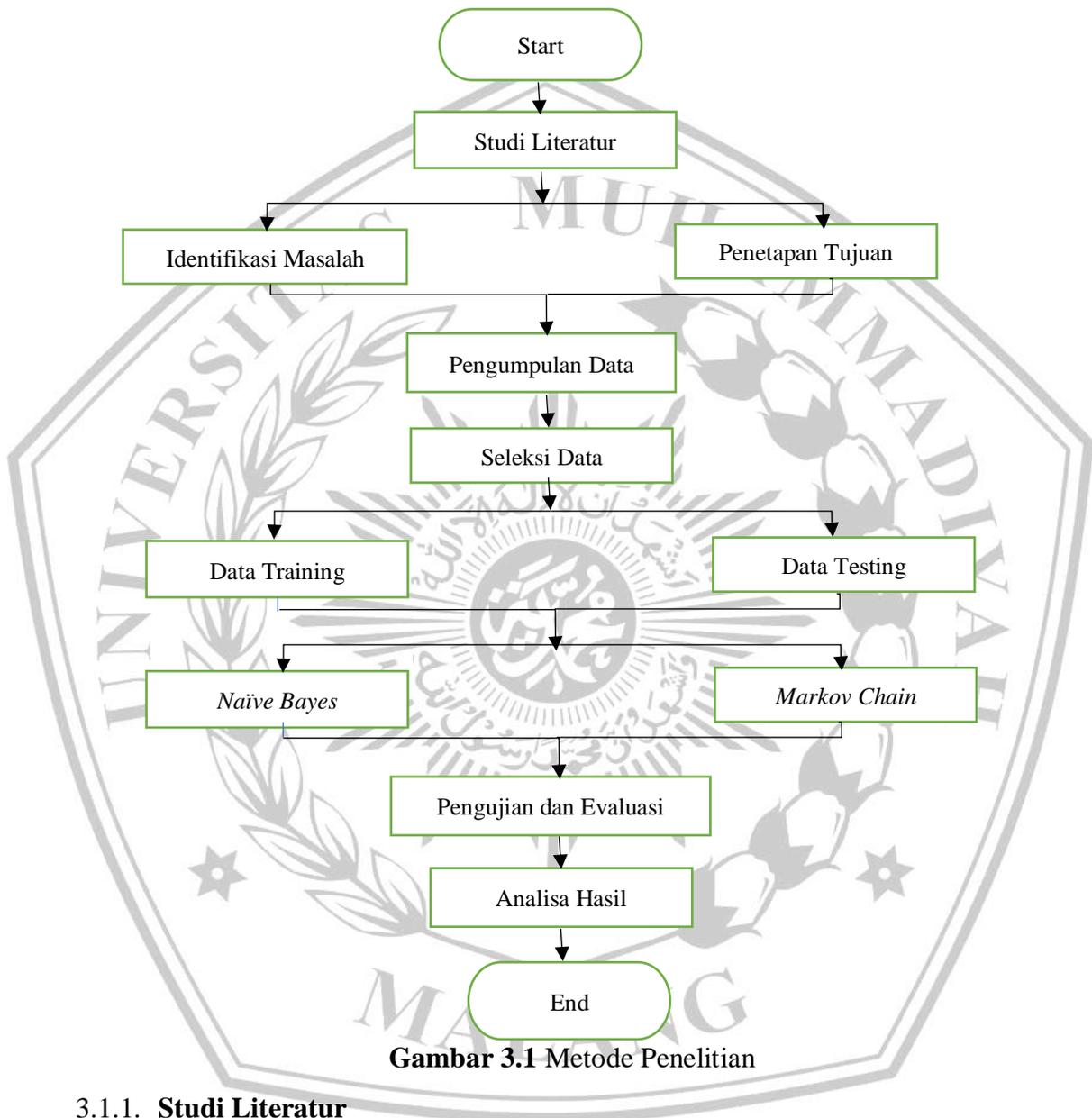


BAB III

METODOLOGI PENELITIAN

3.1. Tahapan Penelitian

Uraian tahapan penelitian yang dilakukan pada penelitian ini antara lain:



3.1.1. Studi Literatur

Setelah merencanakan penelitian, langkah awal yang diambil adalah melakukan studi literatur dengan mengumpulkan beberapa referensi yang berkaitan dengan kelulusan mahasiswa, algoritma Naive Bayes, Markov Chains, serta topik-topik lain yang terkait dengan penelitian yang akan dilakukan.

3.1.2. Identifikasi Masalah

Identifikasi masalah melibatkan pengenalan terhadap masalah yang ada. Dalam penelitian ini, masalah yang diangkat adalah untuk menilai dan membandingkan efektivitas algoritma Naive Bayes dan Markov Chains dalam memprediksi kelulusan mahasiswa jurusan Teknik Elektro di Universitas Muhammadiyah Malang.

3.1.3. Penetapan Tujuan

Tujuan dari penelitian ini adalah untuk mengevaluasi serta membandingkan efektivitas dua metode analisis, yakni algoritma Naive Bayes dan Markov Chains, dalam memprediksi kelulusan mahasiswa pada program studi Teknik Elektro di Universitas Muhammadiyah Malang. Penelitian ini akan menyelidiki dan menganalisis seberapa baik kedua algoritma ini dapat menghasilkan prediksi yang akurat berdasarkan data yang tersedia dari mahasiswa-mahasiswa yang telah lulus atau sedang menyelesaikan studi mereka di universitas tersebut.

3.1.4. Pengumpulan Data

Dalam penelitian ini, teknik pengumpulan data mencakup dua pendekatan:

- a. Studi kepustakaan: Dilakukan untuk menganalisis isu utama dalam penelitian, yaitu kelulusan mahasiswa program Teknik Elektro di Universitas Muhammadiyah Malang.
- b. Studi dokumentasi: Digunakan untuk memperoleh data dan informasi dari arsip atau dokumen, khususnya terkait dengan catatan alumni Universitas Muhammadiyah Malang.

3.1.5. Seleksi Data

Setelah data terkumpul, langkah selanjutnya adalah melakukan seleksi variabel yang relevan dengan fokus pada penelitian ini. Proses ini melibatkan analisis korelasi antara setiap variabel dan label yang menjadi pusat perhatian dalam penelitian mengenai kelulusan mahasiswa program studi Teknik Elektro di Universitas Muhammadiyah Malang.

Tabel 3.1 Variabel Data

No	Variabel Data
1	Nama
2	NIM
3	Tahun Ajaran
4	Tahun Lulus
5	Lama Studi
6	Status
7	IPK

3.1.6. Data Training dan Data Testing

Setelah melakukan proses pembersihan data, selanjutnya data dipilah menjadi dua bagian utama, yaitu data training dan data testing, dengan menggunakan proporsi yang bervariasi. Proporsi data training yang digunakan adalah 60%, 70%, 80%, dan 90% dari total data yang tersedia. Misalnya, untuk data training sebesar 60%, digunakan 60% dari keseluruhan data sebagai bagian untuk proses pelatihan model, sementara 40% sisanya dipisahkan untuk pengujian model. Proporsi ini sama berlaku untuk data training sebesar 70%, 80%, dan 90%.

Dalam tahap penelitian ini, peneliti bertujuan untuk mengevaluasi dan membandingkan hasil akurasi dari berbagai pembagian data training tersebut, dengan harapan dapat mengidentifikasi proporsi mana yang memberikan hasil prediksi yang paling optimal.

3.1.7. *Naïve Bayes dan Markov Chains*

Setelah data dikelompokkan, langkah berikutnya adalah mengolahnya menggunakan algoritma Naive Bayes. Proses dimulai dengan menulis skrip algoritma Naive Bayes di MATLAB. Setelah skrip selesai disusun, algoritma akan mengambil data training, menghitung jumlah dan probabilitasnya. Langkah-langkahnya mencakup menghitung jumlah kelas atau label terlebih dahulu, menemukan probabilitasnya, mengalikan nilai probabilitas variabel, dan membandingkan hasil dari setiap label yang ada. Label dengan probabilitas tertinggi akan dianggap sebagai prediksi yang paling mungkin terjadi.

Selanjutnya, data juga diolah dan dibandingkan dengan algoritma Markov Chains. Langkah pertama adalah membuat skrip Matlab untuk algoritma Markov Chains. Setelah skrip selesai dibuat, dilakukan running dengan menggunakan data yang ada. Setelah algoritma selesai bekerja, hasil yang dihasilkan oleh Matlab diamati dan dibandingkan.

3.1.8. **Pengujian dan Evaluasi**

Setelah implementasi algoritma Naive Bayes dan Markov Chains selesai, langkah selanjutnya adalah melakukan pengujian dan evaluasi. Tujuannya adalah untuk melihat seberapa banyak hasil prediksi yang benar dan tidak benar.

3.1.9. **Analisa Hasil**

Setelah melihat hasil dari penelitian ini, akan dievaluasi apakah Naive Bayes dan Markov Chains efektif dalam memprediksi kelulusan mahasiswa jurusan Teknik Elektro di Universitas Muhammadiyah Malang atau tidak.

3.2. **Jenis dan Sumber Data**

Data yang digunakan dalam penelitian ini berasal dari sumber data sekunder yang diperoleh dari Universitas Muhammadiyah Malang. Data ini mencakup informasi tentang alumni dan mahasiswa baru dari jurusan Teknik Elektro Universitas Muhammadiyah Malang untuk tahun masuk periode 2015-2019, yang termasuk dalam program Sarjana (S1). Penelitian ini memanfaatkan data

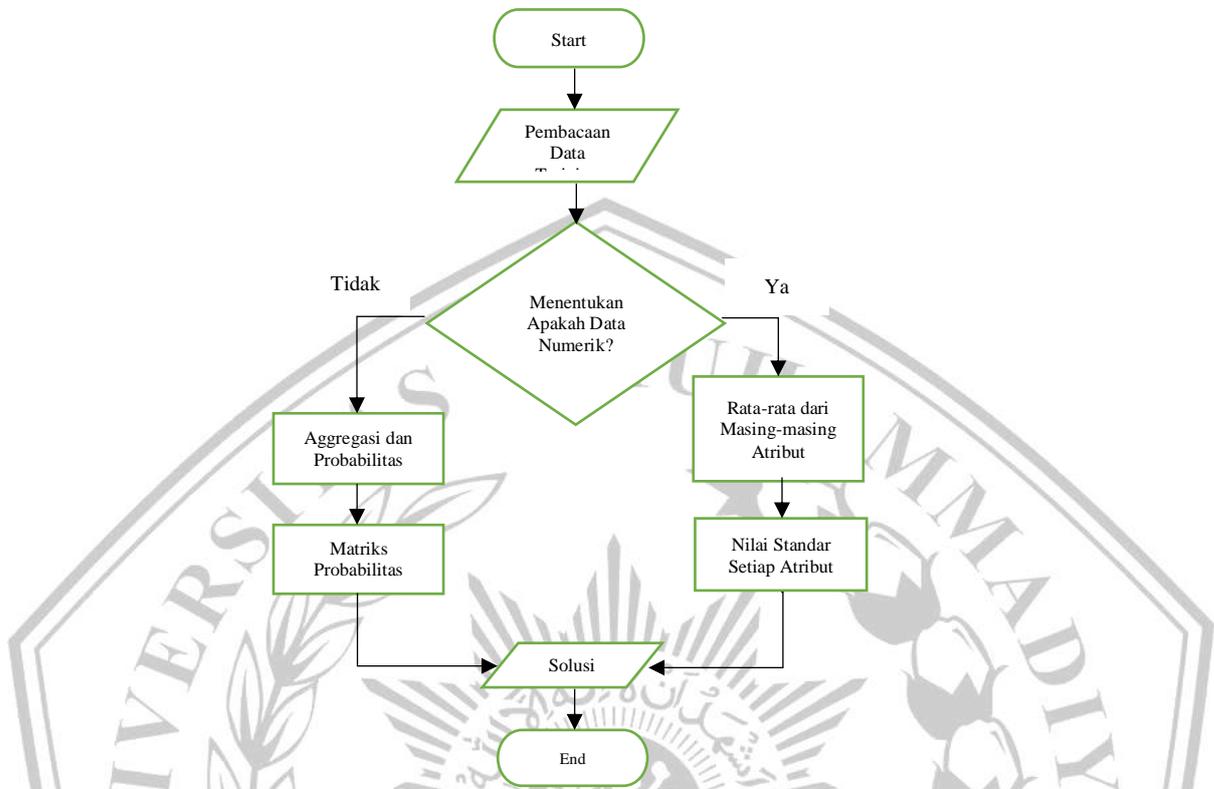
tersebut sebagai basis untuk analisis prediksi kelulusan mahasiswa serta implementasi algoritma seperti Naive Bayes dan Markov Chains.

3.3. Algoritma *Naive Bayes*

Berikut adalah tahapan analisis algoritma Naive Bayes yang lebih diuraikan:

1. Langkah pertama adalah memasukkan data training ke dalam algoritma.
2. Selanjutnya, dilakukan pengecekan terhadap jenis data training untuk menentukan langkah selanjutnya:
 - Jika data tersebut berupa data numerik, maka dilakukan perhitungan rata-rata (mean) dan deviasi standar (standar deviasi) dari setiap parameter untuk menggambarkan distribusi data.
 - Jika data tidak berupa numerik, langkah selanjutnya adalah menghitung probabilitas untuk setiap kategori atau nilai dalam data tersebut. Hal ini dilakukan dengan menghitung jumlah data yang sesuai dari kategori yang sama dan membaginya dengan jumlah total data dalam kategori tersebut, kemudian tabel probabilitas dibuat berdasarkan hasil perhitungan ini.
3. Setelah langkah-langkah di atas selesai, akan diperoleh nilai tabel mean, standar deviasi, dan probabilitas yang digunakan sebagai dasar untuk proses selanjutnya dalam analisis menggunakan algoritma Naive Bayes.

Diagram alur (flowchart) tahapan analisis *Naive Bayes* dapat dilihat di bawah ini untuk lebih memahami prosesnya.



Gambar 3.2 Flowchart *Naive Bayes*

3.4. Algoritma *Markov Chains*

Untuk analisis data menggunakan *Markov Chains* dapat dilakukan langkah-langkah sebagai berikut:

- a. Mengambil data
 - b. Menentukan nilai peluang kelulusan mahasiswa
 - c. Membuat matriks peluang kelulusan mahasiswa
 - d. Membuat matriks peluang emisi B
 - e. Menentukan nilai vektor peluang awal
 - f. Menghitung peluang observasi dengan menggunakan algoritma *Forward Backward*
 - g. Menghitung prediksi parameter *Markov Chain*
- Penarikan Kesimpulan