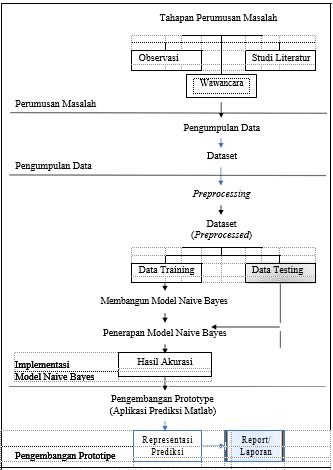
**BAB III**

**METODELOGI PENELITIAN**

* 1. **Diagram Alur Metodologi Penelitian *Naive Bayes***

Pada diagram alur metodologi penelitian ini, penulis menjelaskan alur proses penelitian yang dilakukan. Penelitian ini akan diberikan ruang lingkup/pembahasan agar pembahasan yang dilakukan dapat terstruktur dan terorganisir dengan baik serta dapat mengefektifkan penelitian yang dilakukan karena proses awal hingga akhir telah ditentukan. Alur metodologi yang digunakan oleh penulis adalah algoritma *Naive Bayes Classifier*. Secara bertahap kerangka kerja penelitian ini dapat dijelaskan pada gambar berikut:

Gambar 3.1: Alur Penelitian *Naive Bayes*

* + 1. **Tahapan-Tahapan Dalam Penelitian *Naive Bayes***
       1. **Perumusan Masalah**

Dalam merumuskan masalah, terlebih dahulu dilakukan beberapa tahapan sebagai berikut ini :

1. Observasi

Observasi dilakukan yaitu dengam melihat atau menemukan temuan sebuah kondisi dan kasus pada kampus Politeknik Pos Indonesia. Dari hasil observasi diketahui bahwa terdapat temuan yang dapat dijadikan sebagai bahan penyelesain masalah yang dapat terjadi pada sistem pendidikan Politeknik Pos Indonesia yaitu pada kasus penilaian dan kelulusan mahasiswa.

1. Wawancara

Wawancara dilakukan dengan proses tanya jawab dengan dosen beserta pegawai akademik dan kemahasiswaan di Politeknik Pos Indonesia untuk mengetahui lebih jelasnya masalah apa yang terjadi di Politeknik Pos Indonesia mencakup rana pendidikan tersebut dan penanganan apa yang bias dijadikan solusi terbaik untuk menyelesaikan masalah atau kasus penilaian dan kelulusan mahasiswa.

1. Studi Literatur

Studi Literatur sebagai dasar penulis untuk mendapatkan referensi yang mengenai teori-teori yang berkaitan dengan masalah-masalah yang bersumber dari buku literatur. Berdasarkan hasil studi pustaka yang telah dilakukan penulis, ditemukan solusi pemecahan masalah tentang penilaian dan kelulusan mahasiswa dengan menerapkan prediksi menggunakan algoritma *Naive Bayes Classifier*.

* + 1. **Pengumpulan Data**

Dalam pengumpulan data penulis dalam proses awal sebelum pengolahan data harus memiliki data untuk melakukan proses prediksi kelulusan mahasiswa, maka penulis melakukan pengumpulan data sebagai berikut :

* + - 1. **Pengumpulan Data**

Dalam proses pengumpulan data yang dilakukan pertama kali ialah menerima data yang akan digunakan sesuai dengan kasus atau masalah yang diselesaikan. Data tersebut berasal dari pihak Akademik Kemahasiswaan. Adapun data yang diterima berupa data history akademik mahasiswa D4 Jurusan Teknik Informatika selama 3 tahun terakhir.

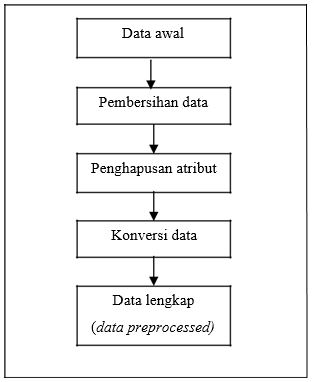
* + - 1. **Dataset**

Pada proses pengumpulan data, data yang diterima berupa dataset yaitu kumpulan history akademik mahasiswa D4 Teknik Informatika. Terdapat 500 data yang akan dikelola. Adapun cakupan data yang digunakan ialah berdasarkan semester 1 sampai dengan semester 6. Data yang diambil hanya berdasarkan semester 1 – semester 6 dikarenakan pada rentan waktu semester 7 – semester 8 akan difungsikan solusi hasil dari prediksi kelulusan mahasiswa yang telah didapatkan sehingga dapat meminimalisir terjadinya penurunan prestasi akademik dan tingkat kelulusan mahasiswa D4 Teknik Informatika Politeknik Pos Indonesia.

* + 1. **Implementasi Metode Naive Bayes**

Dari kajian pustaka penelitian, penerapan metode *Naive* *Bayes Classifier* berpotensi menentukan probabilitas dari prediksi kelulusan mahasiswa D4 Teknik Informatika guna meminimalisir penurunan prestasi akademik mahasiswa dengan memberikan penanganan awal terhadap potensi ketidaklulusan mahasiswa.

* + - 1. ***Preprocessing***

Sebelum tahap implementasi diterapkan, tahap *preprocessing* terlebih dahulu dilakukan. Jumlah data awal yang dapat diperoleh dari pengumpulan data yaitu sebanyak 500 data, tetapi tidak semua data digunakan dan tidak semua atribut dipergunakan karena data tersebut harus melalui tahap pengolahan awal data atau disebut dengan *preparation data*.

Gambar 3.2. *Preprocessing*

Adapun data yang diterima berupa data history akademik mahasiswa D4 Jurusan Teknik Informatika selama 3 tahun terakhir.

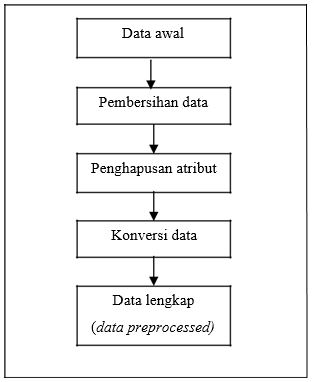
* + - 1. **Dataset**

Pada proses pengumpulan data, data yang diterima berupa dataset yaitu kumpulan history akademik mahasiswa D4 Teknik Informatika. Terdapat 500 data yang akan dikelola. Adapun cakupan data yang digunakan ialah berdasarkan semester 1 sampai dengan semester 6. Data yang diambil hanya berdasarkan semester 1 – semester 6 dikarenakan pada rentan waktu semester 7 – semester 8 akan difungsikan solusi hasil dari prediksi kelulusan mahasiswa yang telah didapatkan sehingga dapat meminimalisir terjadinya penurunan prestasi akademik dan tingkat kelulusan mahasiswa D4 Teknik Informatika Politeknik Pos Indonesia.

* + 1. **Implementasi Metode Naive Bayes**

Dari kajian pustaka penelitian, penerapan metode *Naive* *Bayes Classifier* berpotensi menentukan probabilitas dari prediksi kelulusan mahasiswa D4 Teknik Informatika guna meminimalisir penurunan prestasi akademik mahasiswa dengan memberikan penanganan awal terhadap potensi ketidaklulusan mahasiswa.

* + - 1. ***Preprocessing***

Sebelum tahap implementasi diterapkan, tahap *preprocessing* terlebih dahulu dilakukan. Jumlah data awal yang dapat diperoleh dari pengumpulan data yaitu sebanyak 500 data, tetapi tidak semua data digunakan dan tidak semua atribut dipergunakan karena data tersebut harus melalui tahap pengolahan awal data atau disebut dengan *preparation data*.

Gambar 3.2. *Preprocessing*

Untuk mendapatkan data yang memiliki kualitas bagus, beberapa tahap harus dilakukan, berikut adalah tahapannya:

* + - * 1. ***Data Cleaning***

Data yang tidak berkualitas akan menghasilkan data mining yang tidak berkualitas. Keputusan yang berkualitas harus didasarkan pada data yang berkualitas (data ganda atau data yang telah hilang dapat menyebabkan ketidak benaran). Ekstaksi data, pembersihan, dan transformasi data merupakan tugas utama dalam *data warehouse.* Tugas- tugas *data cleaning,* yaitu:

1. Mengisi nilai yang telah hilang pada data yang tidak lengkap (*missing value*).
2. Mengidentifikasi atau menghilangkan *outliers* dan memperhalus *data noise.*
3. Memperbaiki ketidak konsistenan data.
4. Memecahkan redudansi yang disebabkan oleh integrasi data.
   * + - 1. ***Data Integration* dan *Transformation***

Dapat menigkatkan akurasi dan efisiensi algoritma. Algoritma *Naive Bayes* *Classifier* mampu melakukan proses data yang bernilai nominal, ordinal, ataukontinu. Sehingga nilai dari tiap-tiap attribut yang terdapat pada dataset tidak perlu ditrasnformasikan.

1. Mengintegrasikan berbagai database atau file- file.
2. Transformasi data (*data transformation*).
3. Normalisasi dan agregasi.
   * + - 1. ***Data Reduction***

Proses yang dilakukan dengan mereduksi suatu dataset yaitu dengan mengurangi jumlah atribut dan *record* yang ada sehingga lebih sedikit tetapi masih tetap informatif.

1. Mendapatkan hasil representasi pada volume data yang sebelumnya sudah terkurangi namun tetap menghasilkan analisis yang sama.

Diskritisasi data adalah bagian dari proses reduksi data, namun pada bagian ini ini adalah penting untuk data-data yang numerik.

* + - 1. **Dataset (*preprocessed*)**

Dataset (*preprocessed*) adalah dataset yang sudah melalui *preprocessing* diantaranya sudah melalui proses *cleaning, integration* dan *transformation*, dan *reduction*, sehingga dataset tersebut sudah siap digunakan untuk data traning dandata testing.

* + - * 1. ***Data Training***

Pada proses pemisahan data, data *training* biasanya digunakan sebagai pola data pembentukan model data mining. Adapun banyaknya data *training* yang digunakan ialah sebesar 450 data.

* + - * 1. ***Data Testing***

Pada proses pemisahan data, data *testing* biasanya digunakan sebagai pengujian dalam menetapkan klasifikasi atau prediksi. Adapun banyaknya data *testing* yang digunakan ialah sebesar 50 data.

* + - * 1. **Membangun Model *Naive Bayes Classifier***

Berdasarkan pada *data training* yang sebelumnya sudah melalui *prepocessing* sehingga dapat digunakan untuk membangun model *Naive Bayes* *Classifier* yaitu menyelesaikan masalah klasifikasi *class* atau atribut dalam data.

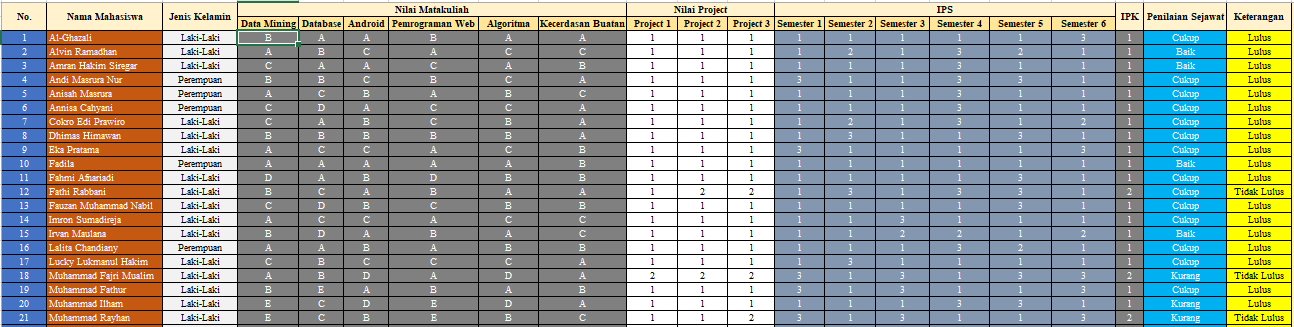
* + - * 1. **Penerapan Model *Naive Bayes Classifier***

Penerapan model *Naive Bayes Classifier* dilakukan berdasarkan model yang telah dibangun sebelumnya yang digunakan untuk menentukan atribut atau *class* dari suatu data baru yang atribut atau *class* nya belum diketahui sebelumnya.

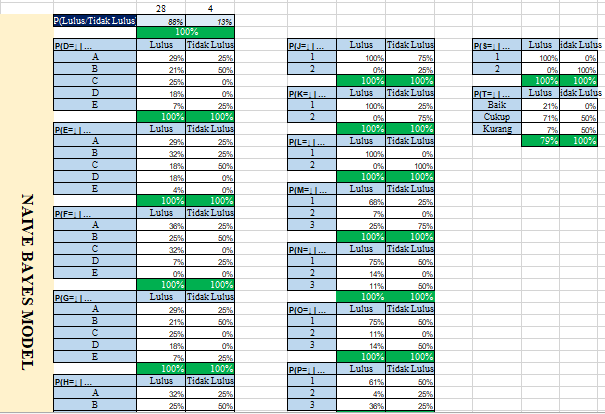
* + - * 1. **Hasil Akurasi**

Hasil akurasi merupakan hasil nilai akurasi yang didapat dari penerapan algoritma *Naive Bayes Classifier*. Untuk menghitung akurasi dari pola yang didapat,pengujian *Confusion Matrix* dan *Receiver Operating Characteristics* (ROC) merupakan metode akurasi dalam konsep data mining. Perhitungan akurasi dapat dibantu dengan penerapan rumus perhitungan *Naive Bayes* pada *Ms.Excel* ataupun *Rapid Miner.*

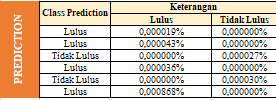
Adapun hasil penerapan rumus Naive Bayes pada Ms.Excel menciptakan akurasi yang sangat baik dengan contoh dataset yang dapat dilihat sebagai berikut:

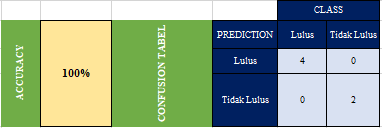


Gambar 3.3 Contoh Dataset Perhitungan *Naive Bayes*



Gambar 3.4 Perhitungan Model *Naive Bayes*



Gambar 3.5 Hasil Prediksi

Gambar 3.6 Hasil Akurasi dan Confusion Tabel

* + 1. **Pengembangan Prototipe**

Pada pengembangan prototipe terdapat 2 (dua) tahapan yaitu penerjemah pola dan *development prototype.* Penerjemah pola pada proses KDD merupakan prosesuntuk mengetahui apakah hasil pola dari penerapan metode yang digunakan sesuai atau bertentangan dengan fakta yang ada. Pada pengembangan prototipe dalam penelitian ini dilakukan melalui tahapan antara lain, sebagai berikut:

* + - 1. **Development Prototype (Aplikasi Prediksi *berbasis web*)**

Pengembangan prototipe digunakan untuk merepresentasikan dalam penelitian ini, yaitu dengan mengimplementasikan aplikasi *web*. Hasil yang didapat dari implementasi tersebut dapat digunakan oleh pihak kampus untuk melakukan penanganan lebih dini atau menciptakan solusi terhadap prediksi kelulusan mahasiswa sehingga meningkatkan *history* prestasi akademik mahasiswa terkait.

* + - 1. **Representasi Prediksi**

Penerapan algoritma *Naive Bayes Classifier* pada aplikasi *web* yang hasilnya penggambaran dari prediksi kelulusan mahasiswa D4 Teknik Informatika dan akan memberikan solusi dan penanganan terhadap hasil prediksi tersebut.

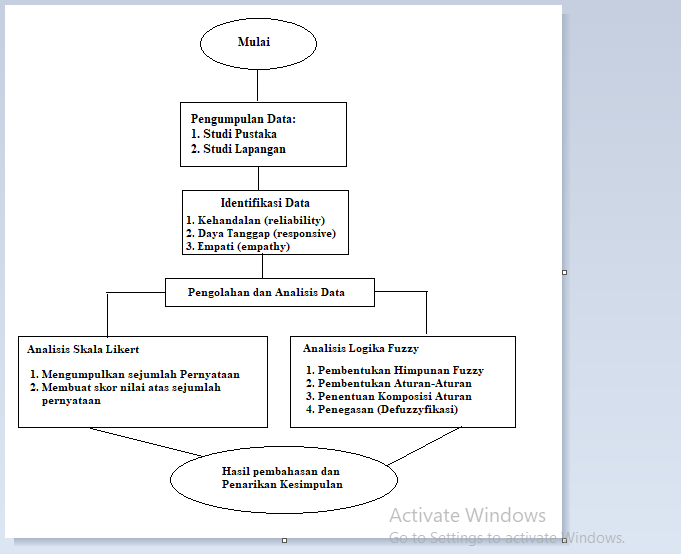
* + - 1. ***Report* atau Laporan**

*Report* atau laporan adalah *view* atau tampilan hasil dari prediksi berupaklasifikasi mahasiswa yang akan lulus dan tidak lulus berdasarkan *history* nilai dan kriteria yang digunakan dalam perhitungan prediksi tersebut.

* 1. **Diagram Alur Metodologi Penelitian *Fuzzy Logic***

Pada diagram alur metodologi penelitian ini, penulis menjelaskan alur proses penelitian yang dilakukan. Penelitian ini akan diberikan ruang lingkup/pembahasan agar pembahasan yang dilakukan dapat terstruktur dan terorganisir dengan baik serta dapat mengefektifkan penelitian yang dilakukan karena proses awal hingga akhir telah ditentukan. Alur metodologi yang digunakan oleh penulis adalah algoritma *Fuzzy Logic*.

*Fuzzy Logic* ini akan digunakan dan dikombinasikan dengan metode *Naive Bayes* untuk menentukan prediksi kelulusan mahasiswa. Hasil dari perhitungan *Fuzzy Logic* akan dipergunakan sebagai penilaian dimana menjadi salah satu kriteria dalam perhitungan prediksi kelulusan mahasiswa. Secara bertahap kerangka kerja penelitian ini dapat dijelaskan pada gambar berikut :



Gambar 3.6 Alur Penelitian *Fuzzy Logic*

* + 1. **Pengumpulan Data**

Penelitian ini menggunakan sumber, yaitu:

1. Studi Pustaka

Pengumpulan data dengan cara membaca dan mengkaji buku-buku serta jurnal nasional maupun internasional secara teoritis yang berkaitan dengan metode yang dibahas.

1. Studi Lapangan (menyebarkan Kuesioner)

Dengan menyebarkan daftar pertanyaan kepada 17 responden, yaitu kepada para mahasiswa D4 Teknik Informatika yang berada dalam kelas yang sama.

* + 1. **Identifikasi Data**

Identifikasi data dilakukan untuk menentukan vareabel dan semesta pembicaraan yang diperlukan dalam melakukan perhitungan dan analisis masalah.

1. Data-data kuesioner yang diambil mencakup aspek-aspek:
2. Kehandalan (*Reliability*)
3. Daya tanggap (*Responsiveness*)
4. Empati (*Emphaty*)
5. Ketiga aspek/dimensi tersebut diatas dinilai dengan menggunakan kategori pengukuran:
6. Kategori 1 : Sangat Kurang Baik (SKB)
7. Kategori 2 : Kurang Baik (KB)
8. Kategori 3 : Cukup Baik (CB)
9. Kategori 4 : Baik (B)
10. Kategori 5 : Sangat Baik (SB)
11. Contoh Pertanyaan Kuesioner :

Tabel 3.1 Pertanyaan Kuesioner

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Pernyataan** | **1** | **2** | **3** | **4** | **5** |
| Teman anda mencatat dan merangkum materi yang diberikan baik ketika didalam kelas maupun setelah kelas selesai. |  |  |  |  |  |
| Teman anda melakukan diskusi materi perkuliahan dengan anda diluar jam kelas. |  |  |  |  |  |
| Teman anda aktif dalam mengajak anda untuk mengerjakan tugas bersama. |  |  |  |  |  |
| Tugas teman anda selesai dengan baik tanpa mengerjakannya bersama dengan teman yang lain (mengerjakan sendiri). |  |  |  |  |  |
| Teman anda pandai dalam presentasi ataupun menjawab pertanyaan |  |  |  |  |  |
| Teman anda berani dalam melakukan presentasi |  |  |  |  |  |
| Teman anda mampu membuat keputusan dengan baik dan cepat |  |  |  |  |  |
| Teman anda tidak mudah putus asa / menyerah |  |  |  |  |  |
| Teman anda selama diskusi kelompok mengajukan ide yang dapat membantu tugas yang dikerjakan |  |  |  |  |  |
| Teman anda mendengarkan pendapat teman yang lainnya ketika berdiskusi |  |  |  |  |  |
| Teman anda handal dalam mengorganisasi ide-ide dalam diskusi kelompok ataupun tugas perindividu |  |  |  |  |  |
| Teman anda membantu anda ketika anda dalam kesusahan |  |  |  |  |  |

* + 1. **Metode Analisis Data**

Dalam pengelohan data yang telah dikemukakan, digunakan dua macam metode analisis, yaitu:

* + - 1. **Analisis Skala Likert**

Menurut Freedy Rangkuti (2002:38), adapun langkah-langkah dalam mengerjakannya ialah sebagai berikut :

* 1. Mengumpulkan sejumlah pertanyaan-pertanyaan yang berkaitan dengan masalah yang akan diteliti. Responden diharuskan memilih salah satu dari sejumlah kategori jawaban yang tersedia, kemudian masing-masing jawaban diberi nilai tertentu.
  2. Membuat nilai total untuk setiap orang dengan menjumlahkan nilai untuk jawaban.
  3. Menilai kekompakan antar pertanyaan.
     + 1. **Analisis logika Fuzzy**

Pengelolahan data dapat dilakukan dengan bantuan *software* matlab .8.0 ataupun dengan perhitungan manual. Dengan menggunakan fasilitas yang disediakan pada toolbox fuzzy pada *software* yang digunakan, dapat dilakukan langkah langkah sebagai berikut:

1. Pembentukan himpunan fuzzy

Dalam tahap ini dilakukan pembentukan himpunan fuzzy berdasarkan hasil analisis data parameter, yaitu dengan menetapkan nilai linguistic dan numeris.

1. Pembentukan aturan aturan *(Rule Base)*

Proses ini berfungsi untuk untuk mencari suatu nilai fuzzy output dari fuzzy input. Prosesnya adalah sebagai berikut: suatu nilai fuzzy input yang berasal dari proses fuzzifikasi kemudian dimasukkan kedalam sebuah aturan yang telah dibuat untuk dijadikan sebuah fuzzy*.*

1. Penentuan komposisi aturan *(Fuzzyfikasi)*

Penentuan komposisi aturan merupankan proses perubahan suatu nilai *crips* ke dalam variabel fuzzy yang berupa variabel linguistic yangnantinya akan dikelompokkan menjadi himpunan fuzzy

1. Penegasan *(defuzzyfikasi)*

Proses defuzzyfikasi adalah suatu himpunan fuzzy yang diperoleh dari dari komposisi aturan aturan fuzzy, sedangkan output yang dihasilkann merupakan suatu bilangan pada domain himpunan fuzzy tersebut.

* + 1. **Penarikan Kesimpulan Penelitian *Fuzzy Logic***

Penarikan kesimpulan merupakan tahap akhir dari penelitian, dimana dilakukan pendokumentasian riset secara keseluruhan. Hasil akhir dari penelitian menggunakan metode *Fuzzy Logic* akan digunakan sebagai salah satu kriteria untuk menghasilkan prediksi kelulusan mahasiswa yang dikombinasikan dengan metode algoritma *Naive Bayes.*