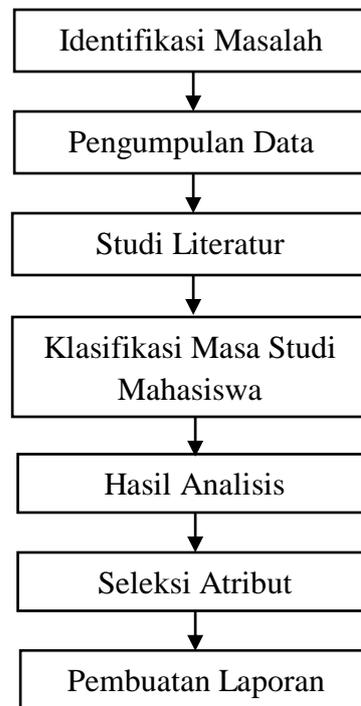


## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 KERANGKA KERJA PENELITIAN**

Kerangka kerja penelitian merupakan tahapan-tahapan yang dilakukan selama mengerjakan penelitian. Kerangka kerja penelitian dibuat agar mempermudah pencapaian hasil penelitian, dapat menyelesaikan penelitian tepat waktu dan penelitian dapat berjalan sesuai dengan yang diharapkan. Adapun kerangka kerja penelitian yang digunakan dapat dilihat pada gambar 3.1.



**Gambar 3.1 Kerangka Kerja Penelitian**

Berdasarkan kerangka kerja yang telah digambarkan diatas, maka dapat diuraikan pembahasan dari masing-masing tahapan penelitian yaitu:

1. Identifikasi Masalah

Pada tahap ini penulis mengidentifikasi hal apa saja yang menjadi permasalahan *data mining* mengenai lama masa studi pada data mahasiswa di STIKOM DB, sehingga penulis dapat menentukan rencana kerja serta menentukan data apa saja yang akan dibutuhkan dalam penelitian ini.

2. Pengumpulan Data

Pada tahap ini penulis melakukan pengumpulan data dan informasi yang dibutuhkan dalam penelitian. Pengumpulan data ini dilakukan dengan beberapa metode yaitu:

- a. Pengamatan (*Observasi*)

Metode ini dilakukan dengan pengamatan secara langsung pada objek yang akan diteliti. Dalam hal ini penulis mengamati bagaimana pengelolaan pada data-data mahasiswa, sehingga penulis mendapatkan *softcopy* data-data mahasiswa tahun 2012 hingga 2016. Data mahasiswa ini diberikan oleh kepala BAAK STIKOM DB. Adapun atribut data mahasiswa yang diperoleh yaitu nama, nim, jenis kelamin, sekolah asal, jurusan sekolah asal, IPK, seluruh nilai matakuliah yang dikontrak, tanggal lulus, tahun lulus, dan angkatan.

- b. Wawancara

Pada metode ini penulis melakukan wawancara singkat secara langsung

dengan kepala BAAK untuk memperoleh data dan informasi akurat mengenai data mahasiswa yang penulis butuhkan.

c. Penelitian Kepustakaan (*Library Research*)

Untuk menunjang penelitian yang akan dilakukan, penulis melakukan *library research* yakni dengan mencari data-data dari buku maupun jurnal penelitian sejenis yang berhubungan dengan metode yang penulis gunakan dalam penelitian ini yaitu metode *Naïve Bayes*.

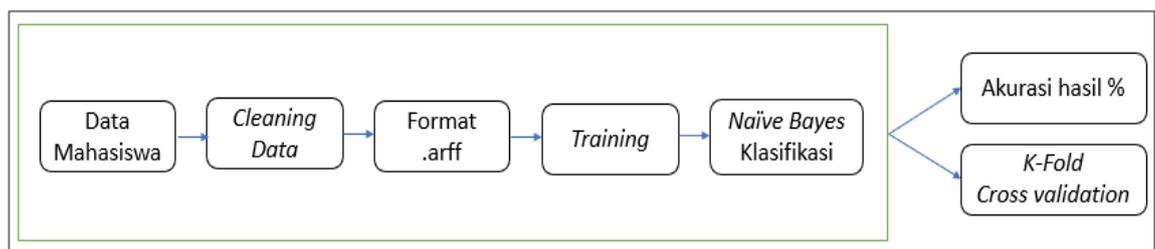
3. Studi Literatur

Pada tahap ini penulis melakukan pencarian terhadap landasan-landasan teori yang diperoleh dari berbagai sumber seperti buku, jurnal ilmiah dan juga referensi lainnya untuk melengkapi penelitian baik mengenai konsep dan teori sehingga memiliki acuan yang baik dan relevan.

4. Klasifikasi Lama Masa Studi Mahasiswa

pengklasifikasian data mahasiswa dilakukan dengan beberapa tahapan.

Adapun alur proses pengklasifikasian dapat dilihat pada gambar 3.2.



**Gambar 3.2 Alur proses klasifikasi dan analisis *naïve bayes***

Berikut ini penjelasan mengenai tahapan-tahapan pada alur proses pengklasifikasian:

#### A. Data Mahasiswa

Data Mahasiswa merupakan data-data yang dimiliki oleh mahasiswa yang telah lulus tahun 2012 dan 2013. Dari sekian banyak atribut data mahasiswa, atribut data yang digunakan yaitu nama, nim, jenis kelamin, sekolah asal, jurusan sekolah asal, IPK, angkatan (tahun masuk), tanggal lulus, tahun lulus.

#### B. *Cleaning Data*

*Cleaning data* merupakan *preproses* dimana pada tahap ini dilakukan pembersihan atau pemisahan data dari *noise* data dan data yang tidak konsisten seperti mahasiswa yang memiliki atribut tidak lengkap. Selain itu ada beberapa hal yang penulis lakukan dalam proses *cleaning data* yaitu sebagai berikut:

- a. Mengklasifikasikan IPK yang berupa angka menjadi 3 kategori yaitu  $>3,5$ =Tinggi,  $\leq 3,5$ =Sedang,  $\leq 2,6$  = Rendah.
- b. Mengklasifikasikan kelas lulus menjadi 3 kategori dihitung dari tahun masuk, tahun lulus dan tanggal lulus yaitu lulus 7 semester menjadi kategori CEPAT. lulus 8 semester menjadi kategori TEPAT dan lulus di atas 8 semester masuk kategori LAMBAT.
- c. Mengklasifikasikan sekolah asal menjadi tiga kategori yaitu SMA, MAN dan SMK.

Kemudian hasil dari *cleaning data* ini disusun dalam bentuk format arff pada tools Notepad++.

### C. Format arff

WEKA mendukung beberapa format file dalam input-nya diantaranya adalah format arff. *Attribute-Relation File Format* (ARFF) merupakan tipe *file text* yang berisi berbagai *instance* data yang berhubungan dengan suatu set atribut data. kemudian data-data ini disusun dengan menggunakan tanda koma (,) sebagai pemisah data per-atributnya.

### D. Training

Pada tahap ini dilakukan proses *training* oleh algoritma klasifikasi naïve bayes untuk membentuk sebuah model *classifier* pada data mahasiswa yang sudah ada. Model ini merupakan representasi pengetahuan yang akan digunakan untuk prediksi kelas data baru yang belum pernah ada.

### E. Naïve Bayes Klasifikasi

Pada tahapan ini, dilakukan proses pengklasifikasian data menggunakan tools WEKA terhadap data mahasiswa yang sudah *ditesting* menggunakan model *classifier* yang sudah dibentuk.

### F. Akurasi Hasil

Akurasi Hasil menunjukkan kedekatan hasil pengukuran dengan nilai sesungguhnya.

### G. K-fold Cross Validation

*Cross Validation* merupakan salah satu teknik untuk menilai/memvalidasi keakuratan sebuah model yang dibangun

berdasarkan dataset tertentu. Salah satu metode *cross-validation* yang populer adalah *K-Fold Cross Validation*.

#### 5. Hasil Analisis

Hasil dari analisis akan dinilai keakuratan data model yang sudah di dapatkan dengan perhitungan *naïve bayes* menggunakan teknik *Cross Validation*. Salah satu metode *cross-validation* yang populer dan akan digunakan adalah *K-Fold Cross Validation*. Dalam teknik ini *dataset* dibagi menjadi sejumlah K-buah partisi secara acak. Kemudian dilakukan sejumlah K-kali eksperimen, dimana masing-masing eksperimen menggunakan data partisi ke-K sebagai data *testing* dan memanfaatkan sisa partisi lainnya sebagai data *training*. Dimana hasil dengan persentasi terbesar yang akan di ambil. Selain itu, bentuk visualisasi dari setiap atribut dapat dilihat pada tools WEKA.

#### 6. Seleksi Atribut

Seleksi atribut ini dilakukan untuk mengetahui atribut apa saja yang memengaruhi hasil klasifikasi. Seleksi atribut ini menggunakan algoritma yang umum digunakan yaitu algoritma Classification Attribute Evaluation (ClassificationAttributeEval) pada tools WEKA.

#### 7. Penulisan Laporan

Setelah semua tahapan penelitian dilakukan, maka akan dibuat laporan sebagai dokumentasi penelitian agar dapat dimanfaatkan pada waktu yang akan datang. Baik oleh peneliti sendiri maupun peneliti lainnya.

### 3.2 ALAT BANTU PENELITIAN

Dalam melakukan penelitian, penulis juga memerlukan alat bantu. Alat bantu yang digunakan dalam melakukan analisis serta pengujian untuk mendapatkan hasil akurasi agar sesuai dengan hasil yang diinginkan. Adapun alat bantu yang digunakan yaitu:

#### 1. Perangkat Keras (*Hardware*)

Perangkat keras yang penulis gunakan dalam penelitian ini adalah sebagai berikut:

- a. Satu unit laptop Lenovo Ideapad™ 110 dengan spesifikasi Processor AMD A8-7410 APU with AMD Radeon R5 Graphics 2.20Ghz, RAM 4GB / Harddisk 1TB, Monitor 15.6 inch.
- b. Satu unit Printer Cannon Pixma MX497
- c. Satu unit Mouse Logitech M171

#### 2. Perangkat Lunak (*Software*)

Perangkat lunak yang penulis gunakan dalam penelitian ini yaitu:

- a. Sistem Operasi Windows 10 64-bit
- b. Microsoft Office 2016
- c. Tools WEKA
- d. Notepad++