

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1. Tempat dan waktu Penelitian**

Penelitian ini dilaksanakan di Universitas Muhammadiyah Yogyakarta di ruang Biro Sistem Informasi, gedung AR. Fachruddin B. Adapun waktu penelitian ini dilaksanakan dari bulan Agustus 2016 - Februari 2017.

#### **3.2. Alat dan Bahan**

Untuk mendukung proses penelitian ini penulis menggunakan beberapa alat untuk mendukung berjalannya perancangan dan implementasi data.

##### **3.2.1 Alat Penelitian**

Alat yang digunakan untuk melakukan penelitian ini berupa perangkat keras dan lunak.

##### **a. Perangkat Keras**

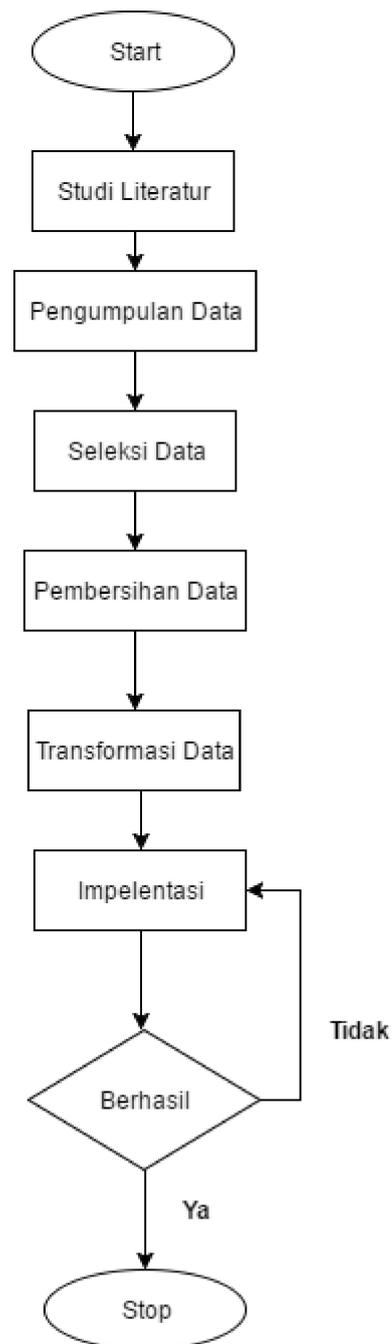
1. *Personal Computer* atau laptop.
2. *Intel i3-2310 2.10 GHZ with Nvidia Geforce 540 M Cuda 2 GB*
3. *6 GB Random Acces Memory (RAM)*

##### **b. Perangkat Lunak**

1. *Operating system windows 8.1 64 bit*
2. *RapidMiner*
3. *Sql server 2014 management studio*

### 3.3. Alur Penelitian

Dalam melakukan penelitian ini, penulis menerapkan alur penelitian yang dapat dilihat pada gambar 3.1



**Gambar 3. 1** *Flowchart* alur penelitian

### 3.4. Studi Literatur

Langkah pertama yang dilakukan dalam penelitian ini adalah studi literatur. Studi Literatur dilakukan untuk mengumpulkan informasi yang diperlukan dalam penelitian ini. Informasi tersebut diperoleh dengan mempelajari dan membaca literatur - literatur yang ada hubungannya dengan permasalahan yang menjadi

### 3.5 Pengumpulan Data

Tahapan selanjutnya adalah pengumpulan data. Tahap ini merupakan tahapan yang penting bagi penulis karena dapat mempengaruhi proses implementasi dan kesimpulan apabila pengumpulan data tidak dilakukan dengan benar. Data yang digunakan dalam penelitian diperoleh dari Biro Sistem Informasi (BSI) Universitas Muhammadiyah Yogyakarta.

### 3.6 Seleksi Data

Data yang ada di *database* tidak semuanya dipakai untuk penelitian ini, oleh karena hanya data yang sesuai untuk dianalisis dan kemudian di *filter* di *database*. Sebagai contoh faktor kecenderungan orang membeli dua barang sekaligus dalam kasus *market basket analysis*, tidak perlu mengambil nama pelanggan cukup dengan id pelanggan untuk menemukan pola-pola pembelian produk.

### 3.7 Pembersihan Data

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten. Pada umumnya data yang diperoleh, baik dari *database* Universitas maupun hasil eksperimen, mempunyai isi yang tidak sempurna seperti data yang hilang, data yang tidak valid dan data yang duplikat. Selain itu ada juga atribut-

atribut data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga juga akan mempengaruhi kinerja dari teknik *data mining* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

### **3.8 Transformasi Data**

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining* karena beberapa metode *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Dalam penelitian ini data akan diubah ke dalam format *csv* melalui *excel*. Transformasi data dapat memudahkan proses *mining* di *software Rapidminer*.

### **3.9 Implementasi**

Merupakan suatu proses utama saat metode ini dilakukan untuk menemukan pengetahuan berharga atau sebuah informasi dari data. Teknik yang digunakan pada proses *mining* ini adalah prediksi dengan menggunakan algoritma *Naïve Bayes* dan *software* yang digunakan dalam proses *mining* ini yaitu *RapidMiner*.

### **4.0. Analisis Data**

Data yang didapatkan dari proses tahapan-tahapan *data mining* kemudian diolah kembali menggunakan *Rapidminer* untuk mendapatkan hasil analisis yang lebih akurat. Atribut yang diuji adalah nilai IPK semester I-IV mahasiswa angkatan 2010 yang telah dinyatakan lulus, jenis kelamin, jurusan sekolah dan asal pulau sebagai data *training*. Kemudian atribut yang digunakan untuk data *testing* adalah nilai IPK semester I-IV mahasiswa angkatan 2014 yang belum lulus, jenis kelamin, asal pulau, jurusan sekolah dan angkatan 2011 yang telah lulus yang

diambil secara acak. Hasil dari analisis prediksi menggunakan *Rapid Miner* dengan algoritma *Naive bayes* adalah dalam bentuk persentase.