

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Data merupakan hal yang *crucial* pada saat ini. Baik bagi perusahaan, pemerintahan, instansi, maupun masing-masing individu yang ada. Setiap harinya, jumlah data yang ada semakin bertambah dan semakin menumpuk. Dengan perkembangan internet pada saat ini dan kemudahan mengaksesnya, jumlah data maupun tumpukan data akan semakin bertambah. Timbunan data-data tersebut dapat menjadi timbunan yang tidak berguna jika tidak ada yang dapat memanfaatkannya dengan baik. Pemrosesan data-data secara manual sangat sulit untuk dilakukan melihat banyaknya data yang ada saat ini. Salah satu data yang menyediakan informasi yang bermanfaat adalah data di dalam bidang pendidikan. Terdapat cukup banyak data dalam bidang tersebut yang setiap semesternya dapat bertambah. Sayangnya, pengelolaan data dalam bidang tersebut hanya sekedar melihat secara sekilas saja. Dengan jumlah data yang memadahi dan cara yang tepat dapat diperoleh sebuah informasi yang mungkin sangat bermanfaat untuk membenahi atau memajukan bidang pendidikan.

*Data mining* merupakan proses untuk mencari sebuah pola dalam sekumpulan data yang besar sehingga diperoleh sebuah pengetahuan yang bermanfaat (Han, et al. 2012:8). Penerapan data mining sendiri dapat dilakukan di dalam berbagai bidang yang ada sekarang. Terdapat dua kategori di dalam pencarian pola dalam *data mining* yaitu *descriptive*, misalnya *clustering* dan *predictive*, misalnya *classification* (Han, et al. 2012: 15). Data yang digunakan akan mempengaruhi *error rate* dari teknik yang digunakan untuk mengelola data tersebut. Implementasi data mining dalam bidang pendidikan khususnya untuk tingkat universitas akan sangat membantu di dalam memperoleh suatu pola atau *pattern* data baik dari data mahasiswa maupun dari data dosen. Dengan studi kasus terhadap data mahasiswa pada beberapa fakultas di Universitas Kristen Maranatha dilakukanlah penerapan *data mining* untuk mendapatkan pola atau *pattern* yang dapat bermanfaat bagi universitas maupun fakultas yang ada.

Tugas Akhir ini merupakan pengembangan ide dari hibah bersaing (Ayub & Kristanti, 2013). Dalam Tugas Akhir ini akan dibuat dan diimplementasikan sebuah aplikasi untuk *clustering* dan *classification* yang membantu analisis data-data tersebut. Dilakukan juga analisa terhadap pola atau *pattern* yang terbentuk dari hasil *classification* dan *clustering* yang dilakukan dalam Tugas Akhir. Sehingga diperoleh hasil yang bermanfaat bagi universitas untuk meningkatkan kualitas pendidikan mahasiswa yang ada.

### **1.2. Rumusan Masalah**

Rumusan masalah yang muncul dari latar belakang tersebut adalah:

1. Pola atau *pattern* seperti apa yang dihasilkan dari *classification* yang mempengaruhi tingkat *drop-out*, lama masa studi, dan tingkat IPK mahasiswa?
2. Di dalam melakukan *clustering*, metode apa yang disarankan untuk melakukan *clustering* terhadap *dataset* mahasiswa Universitas Kristen Maranatha?

### **1.3. Tujuan**

Tujuan dilakukannya penelitian dan pembuatan aplikasi adalah:

1. Memperoleh pola atau *pattern* yang dihasilkan dari *classification* yang mempengaruhi tingkat *drop-out*, lama masa studi, dan tingkat IPK mahasiswa.
2. Mendapatkan metode yang disarankan untuk melakukan *clustering* terhadap *dataset* mahasiswa Universitas Kristen Maranatha.

#### 1.4. Batasan Masalah

Beberapa batasan masalah yang diberikan adalah:

1. Data yang digunakan adalah data mahasiswa beberapa fakultas yaitu fakultas Ekonomi dan fakultas Teknologi Informasi yang sudah berupa *schema*.
2. Data Mahasiswa yang digunakan dipisahkan menjadi dua, yaitu data Mahasiswa Aktif tahun pertama (semester 1 & 2) dan data Mahasiswa Lulusan.
3. *Schema* yang digunakan dalam Tugas Akhir ini berasal dari hibah bersaing (Ayub & Kristanti, 2013) yang diimplementasikan sebagai data *mart* dalam Kerja Praktek (Natalia, 2013).
4. Metode *classification* yang digunakan adalah *NBTree* sedangkan metode *clustering* yang digunakan adalah *DBScan* dan *XMeans*.
5. Aplikasi yang dibuat menggunakan *library* yang telah disediakan oleh WEKA dan dikhususkan untuk mengolah data mahasiswa di Universitas Kristen Maranatha.
6. Dalam *classification*, hal yang dianalisis meliputi data mahasiswa aktif untuk mencari *attribute* yang mempengaruhi *drop-out*, data mahasiswa lulus untuk mencari *attribute* yang mempengaruhi IPK dan lama studi.
7. Dalam *clustering*, dilakukan pembentukan *cluster* dengan metode yang ada yaitu *X-Means* dan *DBScan*, dilakukan perbandingan untuk mencari metode yang sesuai terhadap *dataset* yang digunakan.

#### 1.5. Sistematika Penulisan

Sistematika penulisan untuk tugas akhir ini adalah :

##### BAB I PENDAHULUAN

Bab ini menjelaskan latar belakang, rumusan masalah, tujuan, batasan masalah dan sistematika pembahasan.

##### BAB II LANDASAN TEORI

Bab ini menjelaskan teori-teori yang berkaitan dengan pembuatan sistem dan mendukung pembuatan sistem atau aplikasi yang ada.

### BAB III ANALISIS DAN DISAIN

Bab ini menjelaskan bagaimana analisis keadaan, kebutuhan dari aplikasi, perancangan aplikasi, UML, dan gambaran arsitektur dari aplikasi yang dibuat.

### BAB IV PENGEMBANGAN PERANGKAT LUNAK

Bab ini digunakan untuk menjelaskan mengenai implementasi dari teknik serta perancangan aplikasi dan hasil yang didapat dari percobaan yang telah dilakukan dalam Tugas Akhir.

### BAB V TESTING DAN EVALUASI SISTEM

Bab ini menjelaskan tentang pengujian dari teknik atau aplikasi yang telah dibuat beserta *testing* terhadap hasil percobaan yang telah dilakukan dalam Tugas Akhir.

### BAB VI KESIMPULAN DAN SARAN

Bab ini menjelaskan mengenai kesimpulan yang didapatkan dari hasil implementasi atau penelitian.