

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan. Data mining mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan.

Hasil dari aplikasi data mining tersebut di evaluasi untuk menemukan suatu informasi/pengetahuan baru yang menarik dan bernilai bagi perusahaan, dan kemudian di visualisasikan agar mempermudah bagi user memilih informasi-informasi yang mempunyai arti bagi pendukung keputusan.

Salah satu proses dalam data mining adalah klasifikasi, pada klasifikasi diberikan sejumlah *record* yang dinamakan *training set*, yang terdiri dari beberapa atribut, salah satu atribut menunjukkan kelas untuk *record*. Tujuan dari klasifikasi adalah untuk menemukan model dari *training set* yang membedakan *record* kedalam kategori atau kelas yang sesuai, model tersebut kemudian digunakan untuk mengklasifikasikan *record* yang kelasnya belum diketahui sebelumnya. Salah satu metode yang digunakan di dalam klasifikasi adalah pengklasifikasian menggunakan *naive bayes* (probabilitas bayesian).

Dalam menjalankan perannya di dunia perkuliahan, para mahasiswa pada umumnya dihadapkan pada pemikiran tentang seberapa besar pencapaian yang telah mereka raih selama menjalani proses perkuliahan, apa saja yang telah mereka dapatkan dalam perkuliahan, dan berawal dari pemikiran-pemikiran tersebut, mahasiswa cenderung mencari cara maupun alasan agar mereka dapat lebih maju dan terdorong untuk dapat mencapai prestasi yang maksimal. Banyak hal yang dapat dilakukan para mahasiswa untuk bisa mendapatkan sesuatu yang lebih dari sekedar apa yang bisa mereka dapatkan di perkuliahannya. Di antaranya adalah dengan mengikuti kegiatan-kegiatan di luar agenda kurikulum kampus yang dinilai dapat menggali potensi dan menumbuhkan semangat atau

motivasi untuk mencapai suatu penguasaan, penalaran, maupun kemampuan yang lebih baik di bidangnya.

Pada saat memasuki universitas, para mahasiswa juga cenderung memiliki tujuan-tujuan yang hendak dicapai, ada yang memiliki target agar mendapat nilai baik dan lulus dengan nilai baik pula. Sebagian mahasiswa lainnya ada juga yang memiliki target agar kuliahnya lancar dan lulus dengan *cum-laude*. Sekelompok mahasiswa lainnya, mencita-citakan memperoleh pekerjaan berharap dapat menempati posisi strategis dan mendapatkan gaji yang memadai. Pata aktivis kampus ada yang ingin terjun ke dunia politik dan memiliki keinginan menduduki jabatan - jabatan tertentu, sehingga kesuksesan tersebut akan berimbas naiknya pamor mereka di mata masyarakat, dan lain sebagainya.

### **1.2. Rumusan Masalah**

Permasalahan yang terjadi dapat dirumuskan sebagai berikut :

- a. Belum diperoleh sebuah informasi yang berkaitan dengan kasifikasi mahasiswa pandai sampai dengan sangat kurang pandai.
- b. Belum diterapkannya sebuah teknik data mining dengan algoritma *Naive Bayes Clasifier* untuk klasifikasi mahasiswa pandai.

### **1.3. Tujuan Penelitian**

Tujuan dari dilakukannya penelitian ini adalah :

- a. Menemukan suatu informasi mengenai data mahasiswa pandai dan tidak pandai.
- b. Menerapkan dan melakukan analisis klasifikasi dengan *Naive Bayes Classifier (NBC)*.

### **1.4. Batasan Masalah**

- a. Data yang digunakan adalah data mahasiswa semester 1 sampai dengan 8, Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta angkatan 2011.
- b. Menerapkan teknik klasifikasi dengan algoritma *Naive Bayes* untuk

menemukan data mahasiswa pandai sampai sangat kurang pandai pada mahasiswa Jurusan Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta Angkatan 2011.

### **1.5. Output dan Manfaat**

Dalam penelitian ini akan menghasilkan keluaran berupa klasifikasi data mahasiswa pandai sampai sangat kurang pandai berdasarkan dari data IPK mahasiswa semester 8, Jurusan Teknik Informatika, sehingga dapat memberikan masukan kepada mahasiswa dan orang tua yang belum mengerti tentang cara penilaian Universitas.

Adapun manfaat yang di dapat dalam penelitian ini adalah :

- a. Manfaat untuk IPTEK yaitu sebagai bahan analisis untuk membantu mengambil keputusan proses pengolahan data dalam jumlah yang besar dan klasifikasi data menjadi 4 kelas serta penerapan teknik data mining.
- b. Manfaat untuk *user* adalah untuk mempermudah dalam pengelompokkan nilai yang berlaku di lingkungan Universitas.

### **1.6. Sistematika Penulisan**

Skripsi ini ditulis dan diuraikan dalam lima (5) bagian besar, yaitu BAB I sampai dengan BAB V. Adapun isi dari bab - bab tersebut adalah sebagai berikut :

#### **BAB I PENDAHULUAN**

Bab ini memberikan gambaran secara garis besar mengenai isi skripsi sehingga pembaca dapat memahaminya dengan mudah. Isi dari bab ini adalah latar belakang, rumusan masalah, tujuan penelitian, batasan masalah, *output* dan manfaat, serta sistematika penulisan yang digunakan untuk menyusun laporan ini.

#### **BAB II TINJAUAN PUSTAKA**

Bab ini berisi tentang penjelasan mengenai latar belakang mahasiswa, data mining, teknik klasifikasi data mining, klasifikasi, teorema *Naive Bayes*.

#### **BAB III BAHAN DAN METODE**

Bab ini berisi tentang penjelasan mengenai bahan - bahan seperti perangkat keras (*hardware*) dan perangkat lunak (*software*) serta metodologi yang

digunakan dalam penelitian ini.

#### BAB IV PEMBAHASAN DAN HASIL

Bab ini berisi tentang penjelasan mengenai hasil uji coba yang telah dilakukan dalam penelitian ini, serta pembahasan atau analisa dari hasil uji coba tersebut.

#### BAB V PENUTUP

Bab ini berisi tentang penjelasan mengenai kesimpulan dan saran yang dapat membangun serta meningkatkan maksud dan tujuan dari penelitian ini menuju arah yang lebih baik.

#### DAFTAR PUSTAKA

#### RIWAYAT HIDUP

#### LAMPIRAN

