

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Jumlah data mahasiswa baru yang diterima di UPN "Veteran" Yogyakarta dalam jumlah besar, tetapi mahasiswa yang lulus tepat waktu jauh lebih sedikit dibandingkan mahasiswa yang diterima. Sehingga terjadi penumpukan data mahasiswa dalam jumlah tinggi disetiap periode kelulusan.

Berdasarkan tumpukan data akademik dari data mahasiswa baru dan data kelulusan, yang mempunyai kesamaan pada atribut-atributnya apabila diolah dan dianalisa menggunakan Teknik klasifikasi data mining dengan menggunakan metode Algoritma *K-Nearest Neighbor* akan ditemukan pola serta pengetahuan dan informasi baru yang bermanfaat.

*K-Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru (*testing data*) dengan kasus lama (*training sample*), yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur/atribut yang ada.

Dengan menggunakan fitur dan atribut yang ada pada database akademik yaitu Nilai Rata-rata Ebtanas Murni(NEM/UN), Nilai Rata-rata STTB, Propinsi asal SMA, Status SMA(Negeri/Swasta), Jurusan SMA(IPA/IPS) dan Jenis Kelamin kemudian direlasikan dengan data kelulusan mahasiswa, maka seorang mahasiswa baru dapat diprediksi masa studinya dengan menggunakan database kelulusan mahasiswa. Diharapkan hasil Prediksi mahasiswa baru di awal ini dapat membantu dan meningkatkan kualitas serta memotivasi mahasiswa baru nantinya dalam menempuh perkuliahan di UPN "Veteran" Yogyakarta.

Untuk merealisasikan hal tersebut maka perlu dibuat suatu aplikasi untuk memprediksi lama studi mahasiswa dengan teknik data mining menggunakan metode K-Nearest Neighbor yang berbasis web sehingga mudah diakses di lingkungan UPN "Veteran" Yogyakarta.

Hasil informasi dari proses *Data mining* ini dapat digunakan untuk mengambil keputusan yang nantinya dapat berguna dan membantu manajemen UPN "Veteran" Yogyakarta dalam mempertimbangkan dan mengambil keputusan lebih lanjut tentang salah satu faktor yang mempengaruhi masa studi

mahasiswa baru berdasarkan Nilai dari atribut-atribut yang ada dalam database.

## **1.2 Rumusan Masalah**

Permasalahan yang dibahas dalam skripsi ini adalah bagaimana membuat aplikasi berbasis web yang menghasilkan suatu nilai *output* yang dapat digunakan untuk prediksi masa studi berdasarkan korelasi antara tingkat/data kelulusan dengan data utama akademik mahasiswa dengan menerapkan teknik *data mining* menggunakan metode Algoritma *K-Nearest Neighbor*.

## **1.3 Ruang Lingkup**

Ruang lingkup dalam perancang aplikasi ini dibatasi pada:

1. Sumber data yang digunakan adalah :
  - a. Data Utama Akademik Mahasiswa(Data Primer) yaitu data awal mahasiswa baru UPN "Veteran" Yogyakarta dari tahun 2005 s/d 2009 sejumlah  $\pm$  1000 data.
  - b. Data Kelulusan dari tahun 2005 s/d 2009.
2. Pengolahan data menggunakan Teknik *classification data mining* dengan menggunakan metode Algoritma *K-Nearest Neighbor*.

3. Menganalisa korelasi antara data akademik mahasiswa yaitu variabel Nilai Rata-rata Ebtanas Murni(NEM/UN), Nilai Rata-rata STTB, Propinsi asal SMA, Status SMA(Negeri/Swasta), Jurusan SMA(IPA/IPS) dan Jenis Kelamin, dengan data kelulusan mahasiswa menggunakan aplikasi berbasis web.
4. Aplikasi dibangun berbasis web dengan menggunakan *tools* Dreamweaver CS6 dengan Pemrograman PHP dan MySQL.

#### **1.4 Tujuan**

Tujuan penulisan skripsi ini adalah Menghasilkan suatu aplikasi berbasis web untuk memprediksi masa studi mahasiswa baru dengan Teknik data *mining* menggunakan Metode Algoritma *K-Nearest Neighbor*.