

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Salah satu fakta yang terjadi di lembaga pendidikan STIKOM Dinamika Bangsa Jambi adalah ledakan pertumbuhan data mahasiswa. Ini terjadi setiap tahun dan setiap penerimaan mahasiswa baru. Data menentukan bahwa jumlah mahasiswa baru lebih banyak dibandingkan dengan mahasiswa yang lulus. Waktu studi mahasiswa merupakan salah satu faktor yang menentukan kualitas dari suatu lembaga pendidikan tinggi. Lembaga pendidikan tinggi perlu melakukan analisa terhadap hasil belajar mahasiswa sehingga dapat memprediksi mahasiswa yang mungkin bermasalah. Hasil prediksi dapat digunakan sebagai peringatan dini sehingga dapat dicari solusi untuk membantu mahasiswa lulus tepat waktu. Namun pada saat ini, tumpukan data tersebut belum dimanfaatkan secara maksimal. Padahal tumpukan data tersebut dapat menjadi sebuah informasi yang bermanfaat dengan menggunakan suatu teknik yaitu teknik *data mining*.

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. tujuan utama *data mining* adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki, Mujib Ridwan et al. (2013 : 60).

Metode klasifikasi algoritma k-nearest neighbor merupakan salah satu metode pengklasifikasian data yang memiliki konsistensi yang kuat, dengan cara mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama berdasarkan pencocokan bobot (Kusrini dan Lutfhi, 2009). Algoritma ini lebih efektif dalam melakukan training data yang besar dan dapat menghasilkan data yang lebih akurat. Konsep penelitian dengan algoritma K-nearest neighbor telah banyak dilakukan oleh peneliti-peneliti sebelumnya diantaranya yaitu, k-nearest neighbor merupakan teknik yang sangat sederhana, efisien dan efektif dalam bidang pengenalan pola, kategori teks, pengolahan objek dan lain-lain, karena kesederhanaan pengolahannya dan mampu melakukan training data dalam jumlah yang besar (Bathia, 2010). Algoritma K-NN dapat mengurangi waktu pemrosesan dan memberikan akurasi yang baik dalam sampel pengujian (Khamar, 2013). K-nearest neighbor mengklasifikasikan kasus berdasarkan kesamaan data dalam sampel pengujian (Katarina, 2013).

Hal inilah yang melatar belakangi penulis untuk melakukan penelitian guna memberi solusi terhadap masalah yang terjadi dengan mengangkat judul **“PENERAPAN DATA MINING UNTUK MEMPREDIKSI MASA STUDI MAHASISWA TEKNIK INFORMATIKA STIKOM DB JAMBI MENGGUNAKAN METODE K-NEAREST NEIGHBOR”** yang nantinya dapat mengklasifikasikan lama masa studi / kategori kelulusan mahasiswa yang telah menempuh kuliah dengan menggunakan dasar data nilai IP dari semester sebelumnya.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang diatas, maka rumusan masalah pada penelitian ini adalah :

1. Bagaimana mengklasifikasikan dan menganalisis prediksi lama masa studi mahasiswa Teknik Informatika STIKOM DB JAMBI menggunakan metode *K Nearest-Neighbor* ?
2. Bagaimana mengevaluasi hasil dari prediksi kelulusan menggunakan algoritma *K Nearest-Neighbor* ?

1.3 BATASAN MASALAH

Agar penelitian ini dapat berjalan dengan baik, terarah, dan tidak keluar dari topik pembahasan maka penulis menetapkan batasan-batasan masalah yaitu :

1. Data yang diambil adalah data mahasiswa STIKOM Dinamika Bangsa Jambi angkatan tahun 2016 (berjumlah 165 data) dan mahasiswa angkatan 2013 yang sudah lulus sebelumnya (berjumlah 153 data).
2. Variabel yang digunakan dalam penelitian ini adalah nilai Mata Kuliah dan Indek Prestasi Persemester.
3. Pengujian hasil analisis menggunakan *tools* WEKA.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut:

1. Mengklasifikasikan dan menganalisis data-data mahasiswa teknik informatika dengan metode *K-Nearest Neighbor* sehingga dapat memprediksi lama masa studinya.
2. Mengevaluasi hasil dari prediksi kelulusan menggunakan algoritma *K-Nearest-Neighbor*.

1.4.2 Manfaat Penelitian

Adapun manfaat yang dapat diperoleh dalam penelitian ini adalah sebagai berikut :

1. Dapat mengetahui prediksi lama masa studi masing-masing mahasiswa.
2. Dapat menggunakan hasil analisis untuk menentukan strategi dalam meningkatkan kualitas dan kuantitas kelulusan mahasiswa pada tahun-tahun berikutnya.
3. Penulis dapat menambah ilmu dan wawasan baru mengenai pengklasifikasian dan analisis prediksi lama masa studi mahasiswa dapat digunakan sebagai acuan untuk melakukan penelitian berikutnya.

1.5 SISTEMATIKA PENULISAN

Adapun sistematika penulisan tugas akhir ini, penulis menguraikan dalam beberapa bab yaitu:

BAB I : PENDAHULUAN

Pada bab ini akan diuraikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan dan manfaat penulisan, dan sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab landasan teori ini membahas tentang teori-teori dan pendapat para ahli yang berhubungan dengan permasalahan yang dianalisis. Teori-teori yang digunakan antara lain mengenai data mining, klasifikasi, prediksi, lama masa studi, *K-Nearest Neighbor*, dan WEKA.

BAB III : METODOLOGI PENELITIAN

Pada bab ini menjelaskan tentang kerangka kerja penelitian, metode pengumpulan data, metode klasifikasi, serta alat bantu yang digunakan pada penelitian ini.

BAB IV : ANALISIS

Pada bab ini dilakukan perhitungan analisis menggunakan metode *K-Nearest Neighbor* terhadap data-data mahasiswa yang tersedia.

BAB V : HASIL ANALISIS DAN VISUALISASI

Pada bab ini akan ditampilkan hasil dari analisis dan bentuk visualisasi analisis dari tools WEKA yang digunakan.

BAB VI : PENUTUP

Pada bab ini merupakan penutup dari penelitian ilmiah ini yang berisi kesimpulan dari pembahasan bab-bab sebelumnya dan juga saran-saran yang berguna bagi pihak-pihak yang berkaitan dengan penelitian ilmiah ini.