

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data akademik yang ada pada suatu perguruan tinggi akan semakin bertambah seiring dengan berlangsungnya proses kegiatan akademik. Tidak stabilnya rasio tingkat kelulusan mahasiswa program studi Sistem Informasi pada Universitas Sari Mutiara Indonesia menjadi tugas yang berat bagi program studi. Hal ini menciptakan kondisi yang membuat adanya suatu penumpukan data. Perguruan tinggi dituntut untuk menyelenggarakan pendidikan yang berkualitas bagi mahasiswa sehingga menghasilkan sumber daya manusia yang memiliki daya saing. Dalam perguruan tinggi mahasiswa merupakan aset yang sangat penting bagi institusi pendidikan oleh karena itu perlu diperhatikan tingkat kelulusan mahasiswa tepat pada waktunya [1]. Selain dari pada itu penelitian ini dilakukan karena masih belum adanya penelitian yang melakukan tentang prediksi kelulusan di Universitas Sari Mutiara Indonesia.

Data mining merupakan proses penggalian informasi dan berguna dari set data besar yang melibatkan konsep interdisipliner yang relatif baru yang melibatkan analisis data dan penemuan pengetahuan dari database dan menggunakan pendekatan multi-sisi yang mencakup analisis statistik, visualisasi data, penemuan pengetahuan, pengenalan pola dan manajemen basis data [2]. Algoritma yang digunakan untuk melakukan fungsi klasifikasi adalah algoritma *k*-NN. Algoritma *k*-Nearest Neighbors adalah algoritma *supervised learning* dimana hasil dari *instance* yang baru

diklasifikasikan berdasarkan mayoritas dari kategori k -tetangga atau jarak terdekat [3]. Algoritma *K-Nearest Neighbor* (k -NN) merupakan sebuah metode untuk melakukan klasifikasi terhadap obyek baru berdasarkan (k) tetangga terdekatnya [4]. Tujuan dari algoritma ini adalah untuk mengklasifikasikan obyek baru berdasarkan atribut dan sample-sample dari training data. Penelitian sebelumnya oleh Mustakim prediksi dilakukan terhadap mahasiswa program studi sistem informasi angkatan 2014/2015 sebagai data testing dengan jumlah 50 data, serta berdasarkan dari data angkatan 2012/2013 sebagai data training dengan jumlah 165 data yang menghasilkan pengujian akurasi sebesar 82% [5]. Penelitian sebelumnya oleh Muhammad Sukri memprediksi kelulusan tepat waktu bagi mahasiswa baru. Dari hasil pengujian dengan menerapkan algoritma *k-Nearest Neighbors* dan menggunakan data sampel alumni tahun wisuda 2004 s.d 2010 untuk kasus lama dan data alumni tahun wisuda 2011. Pada metode *k-Nearest Neighbors* ini memiliki tingkat akurasi yang cukup besar yaitu sebesar 83,36% [6]. Tingkat akurasi ini dapat diperbaiki dengan menambah jumlah data training.

Oleh karena itu dengan adanya sebuah prediksi kelulusan mahasiswa dengan teknik klasifikasi menggunakan *k-Nearest Neighbors* diharapkan agar dapat digunakan oleh pihak program studi untuk mencari solusi dan memberi perhatian khusus terhadap mahasiswa yang di prediksi tidak lulus tepat waktu (terlambat) sehingga mahasiswa tersebut diharapkan dapat memperbaiki indeks prestasinya tiap semester agar dapat lulus tepat waktu.

1.2 Rumusan Masalah

Adapun rumusan masalah berdasarkan uraian latar belakang di atas adalah menentukan cara pengklasifikasian untuk memprediksi kelulusan mahasiswa program studi sistem informasi pada Universitas Sari Mutiara Indonesia dengan menggunakan metode *k-Nearest Neighbors*, dan mengetahui nilai akurasi dari *k* terdekat.

1.3 Batasan Masalah

Agar pembahasan dalam penelitian ini dapat lebih jelas dan terarah maka penulis memberikan batasan terhadap permasalahan yang akan di teliti yang berfokus pada :

1. Hanya membahas tentang prediksi kelulusan mahasiswa sistem informasi berdasarkan sample data angkatan tahun 2014-2017.
2. Jumlah data training dan data testing yang digunakan dengan rasio 80 : 20
3. Fitur data yang digunakan dalam penelitian ini terdiri dari Nama, Nomor Induk Mahasiswa (NIM), Indeks Prestasi (IP) semester I-VI dan Status kelulusan.

1.4 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dari penelitian ini adalah untuk mendapatkan prediksi kelulusan mahasiswa berdasarkan algoritma *k-Nearest Neighbors* pada data mahasiswa di program studi sistem informasi Universitas Sari Mutiara Indonesia angkatan 2014-2017.

1.5 Manfaat Penelitian

Adapun manfaat penelitian yang dilakukan yaitu :

1. Mempermudah program studi untuk menginformasikan mahasiswa agar dapat lulus tepat waktu
2. Mengetahui jumlah mahasiswa yang lulus tepat waktu dan tidak tepat waktu pada program studi sistem informasi
3. Mengetahui nilai akurasi tertinggi dengan k terbaik.

1.6 Sistematika Penulisan

Sistematika penyusunan ini dibuat untuk memudahkan para pembaca dalam memahami isi dari penelitian ini. Sistematika penyusunan tersebut adalah sebagai berikut :

BAB I : PENDAHULUAN

Pendahuluan berisi mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, sistematika penulisan serta rencana kegiatan.

BAB II : LANDASAN TEORI

Dalam bab ini terdiri atas penelitian terdahulu dan dasar teori yang bersumber pada jurnal, menyajikan pembahasan teori yang meliputi Data Mining, Klasifikasi, dan k -Nearest Neighbors.

BAB III : METODOLOGI PENELITIAN

Pada bab ini berisi tentang analisis dan perancangan prediksi kelulusan mahasiswa dengan metode k -Nearest Neighbors.

BAB IV : HASIL DAN PEMBAHASAN

Pada bab ini berisi tentang hasil pengujian dari model klasifikasi prediksi kelulusan mahasiswa dengan metode *k-Nearest Neighbor*.

BAB V : KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dan saran dari penelitian yang sudah dilakukan.

