

BAB 1

PENDAHULUAN

1.1. LATAR BELAKANG

Kelulusan merupakan tujuan akhir bagi setiap orang dalam rangkaian proses menyelesaikan studi untuk memperoleh nilai, penghargaan, gelar ataupun yang lainnya. Bagi mahasiswa kelulusan merupakan suatu capaian yang luar biasa terutama bila dapat menyelesaikan studinya tepat waktu. Dalam perguruan tinggi mahasiswa merupakan asset yang sangat penting bagi institusi pendidikan oleh karena itu perlu diperhatikan tingkat kelulusan mahasiswa tepat pada waktunya (Saputra & Primadasa, 2018). Umumnya mahasiswa mampu menyelesaikan studi mereka secara tepat waktu, ada juga yang terlambat menyelesaikan studi (lebih dari 8 semester), namun tidak sedikit pula mahasiswa yang mampu menyelesaikan studinya lebih cepat selama 7 semester. Banyaknya mahasiswa yang masa studinya tepat waktu akan berdampak baik terhadap penilaian akreditasi program studi.

Pada Program Studi Teknik Informatika Universitas Muhammadiyah Gresik sendiri terdapat banyak mahasiswa yang lulus tidak tepat waktu (terlambat) dan mahasiswa yang *Drop Out* (DO) dikarenakan melebihi batas masa studi. Hal tersebut merupakan suatu masalah penting bagi Program Studi, dikarenakan jumlah total dari mahasiswa yang mendaftar tidak sebanding dengan jumlah total mahasiswa yang lulus (menyelesaikan studi) yang nantinya akan berdampak pada penilaian Akreditasi. Sehingga sangat penting bagi program studi mengetahui klasifikasi waktu penyelesaian studi mahasiswa menggunakan data mining. Informasi tersebut dapat menjadi dasar dalam membuat rencana strategis dalam rangkaian pembelajaran supaya meningkatkan jumlah mahasiswa dengan masa studi lulus tepat waktu.

Klasifikasi merupakan suatu proses untuk menemukan model atau fungsi untuk menggambarkan kelas dari suatu data. Proses yang digunakan untuk mendeskripsikan data yang penting serta dapat meramalkan kecenderungan data pada masa depan. Dalam klasifikasi terdapat beberapa algoritme antara lain, *K-Nearest Neighbor*, *Naïve Bayes*, *Neural Network*, *Decision Tree*, dan *Support*

Vector Machine (SVM). Algoritme *K-Nearest Neighbor* merupakan salah satu teknik klasifikasi data yang kuat, dengan cara mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama berdasarkan pencocokan bobot. Algoritme *K-Nearest Neighbor* merupakan satu metode yang termasuk kedalam *supervised learning*, dimana hasil dari klasifikasi ditentukan oleh kelas yang paling banyak muncul (mayoritas) (Saputra & Primadasa, 2018). Salah satu kelebihan dari metode KNN adalah tangguh terhadap *training data* dan efektif apabila *training data*-nya dalam jumlah yang besar (Amalia, 2018). Pada penelitian sebelumnya yang dilakukan oleh Aninda Zulaifa A. dan Yogie Indra K. (2019) yang berjudul “Aplikasi Klasifikasi Penerima Kartu Indonesia Sehat Menggunakan Algoritme *K-Nearest Neighbor*” mendapat kesimpulan, bahwa penerapan Algoritme *K-Nearest Neighbor* dirasa sesuai dilihat dari nilai pengujian data testing sebanyak 12 kali percobaan. Penelitian lainnya yang dilakukan oleh Yulia Rizki Amalia (2018) yang berjudul “Penerapan Data Mining Untuk Prediksi Penjualan Terlaris Menggunakan Metode *K-Nearest Neighbor*” mendapatkan kesimpulan, bahwa dengan menerapkan model prediksi menggunakan metode *K-Nearest Neighbor* mampu memberikan nilai akurasi klasifikasi yang tinggi. Pandu Yuli Santoso dan Dewi Kusumaningsih (2018) dengan judul penelitian “Algoritme *K-Nearest Neighbor* Untuk Memprediksi Kelulusan Ujian Nasional Berbasis Desktop Pada SMAN 12 Tangerang” mendapatkan kesimpulan, nilai akurasi rata-rata sistem yang didapat setelah dilakukan 5 kali testing dalam melakukan prediksi sebesar 88,5% dengan nilai tertinggi pada $K = 7$. Penelitian lainnya yang dilakukan oleh Eka Wahyu Sholeha, Selviana Yunita, dkk (2022) dengan judul penelitian “Analisis Sentimen Pada Agen Perjalanan Online Menggunakan Naïve Bayes dan *K-Nearest Neighbor*” mendapatkan kesimpulan, bahwa didapatkan akurasi tertinggi ketika seluruh data menggunakan huruf kecil untuk kedua Algoritme dengan akurasi 52,35%. Penelitian lainnya yang dilakukan oleh Raemon Syaljumari, dkk (2021) dengan judul penelitian “Akurasi Klasifikasi Pengguna Hotspot WiFi Dengan Menggunakan Metode *K-Nearest Neighbor*” mendapatkan kesimpulan, bahwa berdasarkan tabel presentase terbesar adalah $K = 3$ dengan presentase akurasi 95% dan *error* sebesar 5%. Penelitian lainnya yang dilakukan oleh Tias Mugi Rahayu,

Besse Arnawisuda Ningsi, Isnurani, dan Irvana Arofah (2021) dengan judul penelitian “Klasifikasi Ketepatan Waktu Kelulusan Mahasiswa Dengan Metode *Naïve Bayes*” mendapat kesimpulan, dari total data yang berjumlah 2478 data mahasiswa, diketahui 61,9% lulus tepat waktu berjumlah 697 mahasiswa dan 38,1% lulus tidak tepat waktu berjumlah 429 mahasiswa. Dari perhitungan menggunakan Algoritme *Naïve Bayes* bahwa 225 data testing yang diuji menunjukkan 156 data terklasifikasi secara benar. Hasil klasifikasi ketepatan kelulusan mahasiswa menggunakan Algoritme *Naïve Bayes* diperoleh tingkat akurasi sebesar 69,33%.

Dalam penelitian ini akan dilakukan klasifikasi waktu penyelesaian studi mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik. Secara dominasi waktu penyelesaian studi, bagaimana klasifikasi waktu penyelesaian studi dengan waktu 7 semester (lebih cepat), 8 semester (tepat waktu), dan lebih dari 8 semester (terlambat), sehingga klasifikasi waktu penyelesaian studi akan menggunakan 3 kelas tersebut. Adapun parameter yang digunakan dalam penelitian ini adalah Indeks Prestasi Semester (IPS) yang terdiri dari IPS 1, IPS 2, IPS 3, IPS 4, IPS 5, dan IPS 6, fitur Indeks Prestasi Kumulatif (IPK), Maka Kuliah Pilihan (MKP), Kuliah Kerja Nyata (KKN), dan Tugas Khusus (KP). Metode yang digunakan dalam melakukan penelitian ini adalah metode *K-Nearest Neighbor* (KNN) dan *Naïve Bayes*.

1.2. RUMUSAN MASALAH

Berdasarkan uraian dari latar belakang diatas maka rumusan masalah pada penelitian ini adalah sebagai berikut :

1. Bagaimana implementasi Algoritme *K-Nearest Neighbor* untuk klasifikasi kelulusan mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik ?
2. Bagaimana implementasi Algoritme *Naïve Bayes* dalam melakukan klasifikasi kelulusan mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik ?

3. Bagaimana perbandingan hasil evaluasi Algoritme *K-Nearest Neighbor* dan Algoritme *Naïve Bayes* dalam melakukan klasifikasi kelulusan mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik ?
4. Bagaimana implementasi rancang bangun sistem menggunakan metode *Waterfall* dalam melakukan klasifikasi kelulusan mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik dengan Algoritme *K-Nearest Neighbor* dan *Naïve Bayes* ?

1.3. TUJUAN PENELITIAN

Tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk mengetahui klasifikasi kelulusan mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik menggunakan Algoritme *K-Nearest Neighbor* dan *Naïve Bayes*.
2. Untuk mengetahui nilai akurasi antara Algoritme *K-Nearest Neighbor* dan Algoritme *Naïve Bayes* dalam melakukan klasifikasi kelulusan mahasiswa Program Studi Teknik Informatika Universitas Muhammadiyah Gresik.

1.4. BATASAN MASALAH

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Data training yang digunakan berasal dari 4 gelombang data yudisium terbaru Program Studi Teknik Informatika Universitas Muhammadiyah Gresik. (29 Januari 2021, 30 Juli 2021, 28 Januari 2022, dan 26 Juli 2022)
2. Data testing yang digunakan berasal dari 2 gelombang data yudisium Program Studi Teknik Informatika Universitas Muhammadiyah Gresik. (18 Januari 2020 dan 3 Agustus 2020)
3. Variabel yang digunakan pada penelitian ini adalah Indeks Prestasi Semester (IPS) yang terdiri dari IPS 1, IPS 2, IPS 3, IPS 4, IPS 5, dan

IPS 6, fitur Indeks Prestasi Kumulatif (IPK), Kuliah Kerja Nyata (KKN), Mata Kuliah Pilihan (MKP), dan Tugas Khusus (KP).

4. Metode yang digunakan adalah *K-Nearest Neighbor* dan *Naïve Bayes*.

1.5. MANFAAT PENELITIAN

Penelitian ini diharapkan dapat digunakan sebagai berikut :

1. Dapat memberikan prediksi klasifikasi kelulusan mahasiswa aktif Jurusan Teknik Informatika Universitas Muhammadiyah Gresik.
2. Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Gresik dapat memberikan masukan kepada mahasiswa aktif, terutama beberapa mahasiswa yang mengulang mata kuliah tertentu yang nantinya akan memperlama masa studi mereka.

1.6. METODOLOGI PENELITIAN

Tahapan dalam penelitian dan perancangan sistem yang akan dibangun sebagai berikut :

1. Studi Literatur

Proses studi literatur dimulai dengan mencari permasalahan yang akan menjadi topik penelitian. Dan dilanjutkan dengan mencari sumber referensi yang sesuai dengan topik penelitian melalui buku dan jurnal. Referensi tersebut guna menentukan variabel-variabel dan metode penyelesaian masalah yang akan digunakan.

2. Perancangan

Proses pengumpulan data ini dilakukan dengan melakukan peminjaman akun SIAKAD kepada Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Gresik guna melakukan pengambilan data yang akan dipakai dalam penelitian.

3. Analisis Data

Proses analisis dari data dilakukan untuk mengetahui keakuratan variabel terhadap topik penelitian dengan metode penyelesaian masalah yang digunakan.

4. Perancangan Sistem

Proses perancangan sistem ini meliputi perancangan basis data yang akan diterapkan melalui perancangan diagram alir, perancangan desain aktivitas sistem, dan perancangan desain tampilan sistem.

5. Implementasi dan Pengujian

Proses implementasi merupakan tahapan realisasi dari perancangan sistem yang sudah dilakukan ke dalam program. Dan dilanjutkan dengan proses pengujian program terkait fungsi-fungsi yang ada.

6. Penulisan Laporan

Proses penulisan laporan dilakukan dengan mendokumentasikan setiap perubahan dan hasil dari penelitian. Sehingga dapat digunakan sebagai referensi untuk pengembangan lebih lanjut.

1.7. SISTEMATIKA PENULISAN

Dalam penulisan laporan skripsi ini disajikan ke dalam beberapa bab, dan masing-masing bab dipaparkan dalam beberapa sub bab. Sistematika penulisan laporan skripsi ini bertujuan untuk memudahkan dalam memahami maksud dan tujuan dalam penelitian. Hal tersebut diantaranya sebagai berikut :

BAB I : PENDAHULUAN

Pada bab ini membahas mengenai latar belakang masalah, perumusan masalah, tujuan, batasan masalah, manfaat, metodologi penelitian, dan sistematika penulisan.

BAB II : TINJAUAN PUSTAKA

Pada bab ini menjelaskan dan menguraikan teori-teori yang digunakan sebagai pedoman penelitian.

BAB III : ANALISIS DAN PERANCANGAN SISTEM

Pada bab ini dijelaskan tentang perancangan yang akan dilakukan secara rinci beserta metode yang akan dilakukan.

BAB IV : IMPLEMENTASI DAN HASIL PENGUJIAN

Pada bab ini terdapat penjelasan *source code*, hasil pengujian sistem, serta analisa hasil pengujian.

BAB V : KESIMPULAN DAN SARAN

Pada bab ini akan dipaparkan kesimpulan dan saran dari hasil penelitian ini.

DAFTAR PUSTAKA

Pada bab ini terdapat semua referensi-referensi yang telah didapat dari buku, jurnal, internet, dan lain sebagainya.

