

1. Pendahuluan

1.1 Latar Belakang

Jumlah perguruan tinggi di Indonesia pada tahun 2018 sebanyak 4.690 yang terdiri dari 1.064 akademi, 278 politeknik, 2.531 sekolah tinggi, 215 institut, 583 universitas, dan 19 akademi komunitas [1]. Oleh karena itu, pada tahun 2017 Kemenristekdikti (Kementerian Riset, Teknologi, dan Pendidikan Tinggi) membuat lima klaster perguruan tinggi non politeknik di Indonesia berdasarkan empat aspek dan indikator. Salah satu aspek yang penting adalah aspek kelembagaan (28%) dengan indikator akreditasi institusi BAN PT (Badan Akreditasi Nasional Perguruan Tinggi), akreditasi program studi BAN PT, akreditasi internasional, dan jumlah mahasiswa asing [2].

Salah satu elemen penilaian yang mempengaruhi akreditasi BAN PT adalah profil lulusan (ketepatan waktu penyelesaian studi dan proporsi mahasiswa yang menyelesaikan studi dalam batas masa studi). Terdapat dua poin profil lulusan yaitu persentase kelulusan tepat waktu dan persentase mahasiswa yang DO atau mengundurkan diri [3]. Oleh karena itu, kondisi dan status mahasiswa perlu dianalisa dengan seksama sehingga dapat diberikan penanganan yang tepat.

Tingkat kelulusan tepat waktu mahasiswa dapat diprediksi dengan menggunakan *data mining*. *Data Mining* adalah pemecahan masalah dengan cara menganalisis data dan menemukan pola dalam data yang prosesnya bersifat otomatis atau semiotomatis [4]. Metode yang dapat digunakan dalam *Data Mining* salah satunya adalah klasifikasi data. Klasifikasi data adalah proses pembelajaran dari pemodelan yang telah terbentuk dan proses pengelompokan untuk memprediksi label kelas terhadap data yang ada [5].

Algoritma KNN (K-Nearest Neighbor) adalah algoritma klasifikasi data berdasarkan data pembelajaran yang diambil dari k tetangga terdekatnya. Algoritma ini dapat digunakan untuk melakukan prediksi dan *error rate* yang dihasilkan lebih baik dari kemungkinan minimum yang dihasilkan oleh algoritma Naïve Bayes. Dalam permasalahan apapun, *error rate* yang dihasilkan algoritma KNN tidak pernah lebih buruk dari dua kali *error rate* yang dihasilkan oleh algoritma Naïve Bayes [6].

1.2 Topik dan Batasannya

Terdapat beberapa perumusan masalah antara lain adalah cara mengimplementasikan algoritma KNN dalam prediksi kelulusan tepat waktu mahasiswa S1 Teknik Informatika Universitas Telkom, mengetahui periode yang tepat untuk melakukan klasifikasi, dan mengetahui faktor yang mempengaruhi kelulusan tepat waktu mahasiswa S1 Teknik Informatika Universitas Telkom. Batasan yang digunakan dalam melakukan penelitian berupa pengambilan studi kasus prediksi kelulusan tepat waktu mahasiswa S1 Teknik Informatika Universitas Telkom berdasarkan data akademik dan cakupan data yang digunakan dalam klasifikasi adalah data akademik mahasiswa S1 Teknik Informatika Universitas Telkom angkatan 2008-2011.

1.3 Tujuan

Tujuan dari tugas akhir ini adalah dapat mengimplementasikan algoritma KNN dalam prediksi kelulusan tepat waktu mahasiswa S1 Teknik Informatika Universitas Telkom, membangun model klasifikasi secara *times series* menggunakan algoritma KNN, dan mengetahui faktor yang mempengaruhi kelulusan tepat waktu mahasiswa S1 Teknik Informatika Universitas Telkom.

1.4 Organisasi Tulisan

Jurnal TA ini terdiri dari empat bagian setelah pendahuluan yaitu studi terkait, sistem yang dibangun, evaluasi, dan kesimpulan. Pada bagian studi pustaka terdapat teori dasar yang digunakan dalam prediksi kelulusan tepat waktu mahasiswa antara lain beberapa jurnal yang menggunakan algoritma KNN atau *K-Fold Cross Validation*, penjelasan *knowledge discovery*, *data mining*, klasifikasi, algoritma *K-Nearest Neighbor*, dan metode *K-Fold Cross Validation*. Bagian sistem yang dibangun menjelaskan tentang tahapan sistem yang dibangun yaitu pengolahan data akademik mahasiswa S1 Teknik Informatika Universitas Telkom angkatan 2008-2011, pembagian data akhir pemrosesan, klasifikasi menggunakan algoritma KNN, dan pengujian menggunakan *cross validation*. Bagian evaluasi menjelaskan mengenai hasil pengujian yang berupa akurasi dan analisis secara manual. Bagian kesimpulan berisi kesimpulan dari hasil pengujian dan analisis serta saran untuk penelitian lebih lanjut.