

1. Pendahuluan

Latar Belakang

Akreditasi adalah bentuk evaluasi atau penilaian mutu dan kelayakan institusi perguruan tinggi atau program studi yang dilakukan oleh organisasi atau Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT) [1]. Perguruan tinggi yang memiliki akreditasi bagus akan menjadi pilihan bagi calon mahasiswa. Salah satu penilaian akreditasi Badan Akreditasi Nasional Perguruan Tinggi adalah persentase kelulusan tepat waktu untuk setiap program [1]. Oleh karena itu, tingkat kelulusan menjadi hal yang penting terhadap penilaian akreditasi.

Persentase kelulusan tepat waktu pada perguruan tinggi dapat diprediksi dengan *learning analytics*. *Learning analytics* merupakan cara untuk mengolah data yang didapatkan dari pembelajaran[2]. *Learning analytics* mulai digunakan pada dunia pendidikan tinggi untuk meningkatkan kualitas pendidikan. Sehingga dapat memberikan manfaat yaitu peningkatan kelulusan, pengembangan kurikulum, peningkatan kinerja dosen, waktu penerimaan kerja setelah lulus, dan meningkatkan riset-riset di bidang pendidikan. Salah satu metode yang sering digunakan pada *learning analytics* adalah klasifikasi.

Penelitian prediksi kelulusan mahasiswa tepat waktu sebelumnya pernah dilakukan dengan menggunakan *k-nearest neighbor* [3]. Penelitian ini mengkombinasikan algoritma *naïve bayes* dan *k-nearest neighbor*. *Naïve bayes* digunakan karena memiliki kelebihan pada perhitungannya cepat dan sederhana [4]. *Naïve bayes* juga memiliki kekurangan satu probabilitas atribut tidak bisa mengukur besar tingkat keakuratan sebuah prediksi [4]. Algoritma *k-nearest neighbor* digunakan karena memiliki kelebihan dapat memproses data yang besar namun memiliki kelemahan membutuhkan waktu komputasi cukup tinggi [5]. Untuk mengatasi kelemahan dan mengoptimalkan kelebihan kedua algoritma tersebut, maka penelitian ini akan mengkombinasikan algoritma *naïve bayes* dan *k-nearest neighbor*. Metode kombinasi dilakukan dengan mencari atribut yang paling berpengaruh dengan *naïve bayes*. Lalu algoritma *k-nearest neighbor* hanya memperhitungkan atribut yang dipilih untuk dimasukkan ke model klasifikasi yang dibangun.

Prediksi dilakukan secara temporal, yakni dilakukan per tahun. Ini bertujuan untuk dapat memprediksi secara dini mahasiswa yang lulus tepat waktu atau tidak. Sehingga dapat melakukan tindak pencegahan terhadap mahasiswa yang beresiko lulus tidak tepat waktu ataupun *drop out*. Dataset yang digunakan untuk prediksi adalah data nilai dan mahasiswa S1 Teknik Informatika Universitas Telkom angkatan 2008-2011. Data nilai mahasiswa merupakan salah satu faktor terpenting dalam perkuliahan dan bersifat temporal, sehingga digunakan sebagai atribut pada penelitian ini. Tujuan dari penelitian ini adalah untuk mendapatkan hasil prediksi kelulusan mahasiswa secara temporal dengan algoritma *naïve bayes* dan *k-nearest neighbor*, mengetahui atribut mata kuliah yang mempengaruhi kelulusan mahasiswa dengan algoritma *naïve bayes*, dan menganalisis hasil prediksi.