

Abstrak

Kelulusan tepat waktu bagi seorang mahasiswa adalah hasil yang baik dan menjadi poin penting untuk akreditasi perguruan tinggi. Karena hal tersebut, memerlukan strategi yang baik. Persentase kelulusan tepat waktu pada perguruan tinggi dapat diprediksi dengan *data mining* dan *machine learning*. Tujuan dari penelitian ini adalah memberikan pencegahan dini terhadap mahasiswa yang beresiko lulus tidak tepat waktu, ada beberapa metode klasifikasi yang dapat digunakan untuk memprediksi. Riset ini mengkombinasikan algoritma *naïve bayes* dengan *principal component analysis* (PCA). PCA digunakan untuk menghasilkan data yang memiliki struktur yang lebih sederhana sehingga mencakup informasi dari data aslinya lebih baik, sehingga dapat memberikan hasil yang lebih optimal. Dataset yang akan digunakan adalah berupa *time series data* yang berupa data historis akademik mahasiswa Informatika Telkom University periode 2008 – 2011. Hasil yang didapatkan dengan algoritma kombinasi PCA dan *naïve bayes* adalah algoritma PCA menghasilkan beberapa *principal component* yang memberikan hasil yang optimal untuk digunakan oleh *naïve bayes*, dari proses tersebut dapat menghasilkan pertambahan akurasi rata – rata 6,04% lebih baik dibandingkan algoritma *naïve bayes* tanpa kombinasi.

Kata kunci: kelulusan, klasifikasi, machine learning, mahasiswa, *naïve bayes*, *principal component analysis*.