

KLASIFIKASI KETEPATAN LAMA STUDI MAHASISWA DENGAN ALGORITMA RANDOM FOREST DAN GRADIENT BOOSTING (Studi Kasus Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jakarta)

MUHAMMAD LABIB MU'TASHIM

ABSTRAK

Universitas umumnya pada setiap tahun menerima mahasiswa baru dan memiliki kuota yang berbeda pada setiap jurusannya, namun hal ini tidak diimbangi dengan jumlah mahasiswa yang lulus, sehingga mengakibatkan jumlah mahasiswa meningkat setiap tahunnya, begitu juga dengan Fakultas Ilmu Komputer (FIK) Universitas Pembangunan Nasional (UPN) Veteran Jakarta. Melimpahnya data akademik di FIK UPN Veteran Jakarta bisa diolah sesuai yang dibutuhkan dan berguna mencari informasi penting demi pengembangan fakultas menjadi lebih baik. Maka dari itu dilakukan penelitian untuk menganalisis mahasiswa yang lulus tepat waktu maupun tidak tepat waktu dengan data mining.

Penelitian ini menggunakan metode Random Forest dan Gradient Boosting untuk mengetahui tingkat akurasi dan menentukan mana model klasifikasi yang terbaik pada ketepatan lulus mahasiswa. Kedua algoritma ini memiliki akurasi yang sangat baik pada kebutuhan data yang sangat besar. Analisis menggunakan data mahasiswa S1 FIK UPN Veteran Jakarta angkatan 2015 - 2017. Hasil uji coba sampel pada 590 data, algoritma random forest 10 k-fold mendapatkan akurasi 82,64% dan pada gradient boosting 3 k-fold mendapatkan akurasi 79,66%. Hasil penelitian ini digunakan sebagai salah satu dasar pengambilan keputusan untuk menentukan kebijakan oleh pihak fakultas.

Kata kunci : Klasifikasi, Random Forest, Gradient Boosting, Kelulusan

CLASSIFICATION OF STUDENT STUDY DURATION USING RANDOM FOREST AND GRADIENT BOOSTING ALGORITHMS (Case Study of the Faculty of Computer Science, Pembangunan Nasional Veteran University Jakarta)

MUHAMMAD LABIB MU'TASHIM

ABSTRACT

Universities generally admit new students every year and have different quotas for each department, but this is not matched by the number of students graduating, resulting in the number of students increasing every year, as well as the Faculty of Computer Science (FIK) at the National Development University (UPN).) Jakarta Veterans. The abundance of academic data at FIK UPN Veteran Jakarta can be processed according to what is needed and is useful for finding important information for better faculty development. Therefore, research was conducted to analyze students who graduated on time or not on time with data mining.

This study uses the Random Forest and Gradient Boosting methods to determine the level of accuracy and determine which classification model is the best for the accuracy of student graduation. Both of these algorithms have very good accuracy on very large data requirements. The analysis uses data from S1 FIK UPN Veteran Jakarta students class 2015 - 2017. The results of the sample trial on 590 data, the 10 k-fold random forest algorithm obtains 82.64% accuracy and the 3 k-fold gradient boosting obtains 79.66% accuracy. The results of this study are used as a basis for decision making for determining policies by the faculty.

Keywords : *Classification, Random Forest, Gradient Boosting, Graduation*