

## ABSTRAK

Sistem akademik dirancang setiap instansi pendidikan yang mencakup seluruh proses pendidikan yang berlaku. Kualitas mahasiswa yang baik dipengaruhi berbagai faktor, dengan salah satu inti utamanya adalah sistem akademik yang tersedia. Diketahui dari penelitian sebelumnya bahwa kualitas seorang pelajar yang bisa disebut performa akademik, dapat diketahui melalui data historis proses admisi pelajar tersebut. Salah satu proses admisi yang sebelumnya berlaku di perguruan tinggi negeri Indonesia adalah Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN). Tujuan dari penelitian ini adalah mengolah data SNMPTN, yang digabungkan dengan data nilai Indeks Prestasi Kumulatif (IPK), sehingga dapat diproses menggunakan model pembelajaran mesin. Data yang digunakan diproses menggunakan metodologi pembelajar mesin untuk menghasilkan suatu model. Model tersebut dievaluasi yang kemudian dilakukan proses perbandingan dalam rangka mencari algoritma yang menghasilkan nilai evaluasi tertinggi. Model yang dibuat bertujuan untuk memprediksi performa akademik seseorang dengan mengklasifikasikannya kedalam 4 kelas label. Klasifikasi label ditentukan berdasarkan rentang nilai IPK. Algoritma yang digunakan untuk membuat model adalah algoritma *Supervised Learning* Klasifikasi, diantaranya *Decision Tree* (DT), *Random Forest* (RF), *Support Vector Machine* (SVM), dan *Extreme Gradient Boosting* (XGB). Penelitian dilakukan dalam 3 skema berdasarkan persentase antara data latih dengan data uji. Hasil yang diperoleh menunjukkan DT mengeluarkan nilai akurasi dan *precision* tertinggi, dengan nilai akurasi 0,79 dan nilai *precision* 0,56. Kemudian XGB menghasilkan nilai *recall* dan *f1-score* tertinggi, dengan nilai *recall* 0,35 dan nilai *f1-score* 0,36. Dengan meninjau perbandingan skala nilai *true positive* dan *true negative* yang dihasilkan setiap model pada tabel *confusion matrix* tergolong signifikan, maka dapat dipilih model dengan *f1-score* tertinggi sebagai model terbaik, yakni model dengan algoritma Extreme Gradient Boosting pada skema data latih-uji 70%-30%.

**Kata Kunci:** Admisi, Klasifikasi, Pembelajaran Mesin, Performa Akademik, *Supervised Learning*

## ***ABSTRACT***

The academic system designed every educational institution that covers the entire educational process applies. A good student's quality is influenced by a variety of factors, with one of the main core being the available academic system. It is known from previous research that the quality of a student, which can be called academic performance, can be determined through the historical data of the student's admission process. One of the previous admission processes in Indonesian state colleges was the National Selection of State Colleges (SNMPTN). The purpose of this research is to process the SNMPTN data, which is combined with the data value of the cumulative Index of Performance (IPK), so that it can be processed using a machine learning model. Data used is processed by using machine learning methodology to produce a model. The model is evaluated and then performed a comparison process in order to find the algorithm that produces the highest evaluation value. The model is designed to predict a person's academic performance by classifying it into four classes of labels. The classification of the label is determined on the basis of the range of IPK values. The algorithms used to create the model are the Supervised Learning Classification Algorithms, including Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Extreme Gradient Boosting. (XGB). The research was carried out in 3 schemes based on the percentage between the training data and the test data. The results obtained showed DT yielded the highest accuracy and precision values, with a precision value of 0.79 and a precisions value of 0.56. Then XGB yields the maximum recall and f1-score value, with the recall value of 0,35 and the f1 score of 0,36. By reviewing the comparison of the true positive and true negative values scales that each model generates on the matrix confusion table belongs significantly, then the model with the highest f1-score can be chosen as the best model, that is, the models with the Extreme Gradient Boosting algorithm on the training data scheme 70%-30%.

**Keywords:** Academic Performance, Admission, Classification, Machine-Learning, Supervised Learning