

ABSTRACT

The problem of on-time graduation of students at Telkom University, majoring in Information Systems, showed a decrease in the percentage of graduation until 2018. The inability of students to complete their studies within the specified time can have an impact on the institution's reputation, increased tuition fees and living costs, delays in professional careers, and overcapacity of campus facilities. To overcome this problem, a graduation prediction system is needed that is fast, accurate, and effective. This research applies data mining techniques by Naïve Bayes classification algorithm to predict on-time graduation. The model without SMOTE shows good performance with an accuracy of 86.38%, precision of 95.76%, and recall of 87.78%, while the accuracy varied between 76.1% to 91.5% using K-fold CV with 5 fold experiments. Meanwhile, the model with the application of SMOTE resulted in an accuracy of 83.93%, precision of 94.48%, and recall of 78.08%, with accuracy varying between 65.9% to 87.8%. Although SMOTE helped in some folds, the model without SMOTE showed a more stable and better performance in predicting graduation on time. In conclusion, the Naïve Bayes model without SMOTE is more recommended to be used in predicting the on-time graduation of students.

Keywords : Data Mining, Classification, Naïve Bayes, Crisp-DM