

Abstract

Work readiness for college graduates is an essential and significant thing to get a job immediately after graduation. But what happens is that many graduates are unemployed after graduation or do not get jobs that match the majors they have studied for more than four years. Therefore, by using a people analytics approach, this study aims to predict the work readiness of Telkom University students and find out what factors affect student work-readiness after graduation. The model built is a multi-classes classification model. This model uses Chi-square Test calculation for feature selection, Multinomial Logistic Regression and Random Forest as a classification method, and confusion matrix as an evaluation method. Multinomial Logistic Regression is used because several studies use this algorithm for categorical data, while Random Forest is used to compare which model produces better accuracy. This study conducted several test scenarios, which obtained the best model by performing hyperparameter tuning and handling unbalanced data with SMOTE-ENN. Handling imbalanced data with SMOTE-ENN is used to improve accuracy scores and predict classes well, especially for minority class. The best accuracy of the Multinomial Logistic Regression method is 53.9%, and Random Forest is 48.5%.

Keywords: People Analytics; Work Readiness; Students Performance; Multinomial Logistic Regression; Random Forest