Abstract

Ensuring graduates' smooth transition into the job market is crucial as competition rises with increasing graduate numbers. This research addresses predicting employability, focusing on Telkom University students' initial job income. Using a dataset of 6089 Telkom University 2022 alumni, split 80:20 for training and testing, the study utilized Support Vector Machine (SVM) for data analysis due to the limitations of traditional linear regression in handling potential non-linearity in the data. Feature manipulation techniques like Principal Component Analysis, Spearman's rank correlation, and the Chi-square test of independence were applied, followed by SMOTE-ENN to tackle data imbalance. The SVM model, with Randomized Search hyperparameter tuning and analyzed through Permutation Feature Importance, identified key employability factors. The enhanced SVM model, utilizing SMOTE-ENN, Spearman's rank correlation for feature selection, and Randomized Search, achieved precision, recall, f1-score, and accuracy of approximately 0.70, 0.73, 0.71, and 0.73, respectively. Competency features such as ethics, English skills, IT skills, and knowledge emerged as the most influential factors.

Keywords: feature manipulation, imbalanced dataset, tracer study analysis, student employability, SVM