

ABSTRACT

Internship programmes are an important part of preparing vocational students for the world of work. However, the internship selection process at the Faculty of Applied Sciences, Telkom University, still faces challenges in terms of lengthy timeframes and low acceptance rates. This study aims to develop a job recommendation and resume data extraction system based on Natural Language Processing (NLP), Machine Learning (ML), and Regular Expressions (Regex) to support the acceleration and improvement of the accuracy of the internship participant selection process. The system was built using the Random Forest algorithm for classification and Regex for extracting important information from resumes. Evaluation was conducted using the Accuracy, Precision, Recall, and F1-score metrics, as shown in the learning curve, as well as testing the accuracy of resume extraction through test cases. The test results showed that the system was able to achieve an Accuracy value of 100% and was able to extract resume data perfectly compared to the manual method. Therefore, this system is effective in assisting the Cooperation and Internship Service Unit in the selection process of internship participants and providing job recommendations that align with students' profiles.

Keywords: Resume data extraction, Job recommendation system, Natural Language Processing, Machine Learning, Regular Expression