

## **ABSTRACT**

### **A COMPARISON OF DECISION TREE, RANDOM FOREST, AND K-NEAREST NEIGHBOR ALGORITHMS IN PREDICTING EMPLOYEE TURNOVER**

**By:**

**AULIA MAHARANI**

E-mail: auliamaharani201101@gmail.com

Employee turnover represents a significant challenge for companies, as it can affect operational stability and efficiency. This study compared three machine learning algorithms—Decision Tree, Random Forest, and K-Nearest Neighbor (KNN)—to predict employee turnover using a dataset consisting of 4,653 employee records. The dataset included variables such as education level, year of joining, work location, salary level, age, gender, experience in the current field, and history of unemployment. The research methodology involved data collection, data exploration, preprocessing, model implementation, and classification model evaluation. The results concluded that the Random Forest algorithm achieved the highest accuracy at 86%, followed by Decision Tree at 84%, and KNN at 79%. Therefore, Random Forest was recommended as the most effective model for predicting employee turnover due to its more stable and accurate performance. This study contributes to helping companies better understand the factors influencing employees' decisions to stay or leave, thus serving as a foundation for developing more effective employee retention strategies.

**Keywords:** Employee Turnover, Machine Learning, Comparison



## **ABSTRACT**

### **A COMPARISON OF DECISION TREE, RANDOM FOREST, AND K-NEAREST NEIGHBOR ALGORITHMS IN PREDICTING EMPLOYEE TURNOVER**

**By:**

**AULIA MAHARANI**

E-mail: auliamaharani201101@gmail.com

Employee turnover represents a significant challenge for companies, as it can affect operational stability and efficiency. This study compared three machine learning algorithms—Decision Tree, Random Forest, and K-Nearest Neighbor (KNN)—to predict employee turnover using a dataset consisting of 4,653 employee records. The dataset included variables such as education level, year of joining, work location, salary level, age, gender, experience in the current field, and history of unemployment. The research methodology involved data collection, data exploration, preprocessing, model implementation, and classification model evaluation. The results concluded that the Random Forest algorithm achieved the highest accuracy at 86%, followed by Decision Tree at 84%, and KNN at 79%. Therefore, Random Forest was recommended as the most effective model for predicting employee turnover due to its more stable and accurate performance. This study contributes to helping companies better understand the factors influencing employees' decisions to stay or leave, thus serving as a foundation for developing more effective employee retention strategies.

**Keywords:** Employee Turnover, Machine Learning, Comparison

