Abstract:

The goal of this project was to analyze the employee attrition dataset of a company and develop a model to predict employee attrition. The dataset consisted of 1470 records and 35 variables. Exploratory data analysis was performed to understand the data and identify any patterns or correlations. Various visualization techniques were used to get insights from the data. Machine learning algorithms like Logistic Regression and Deep Learning were used to build the prediction model.

Logistic Regression algorithm gave an accuracy of 80% with a confusion matrix of 144, 0, 29, and 4. Deep Learning algorithm gave an accuracy of 78%. The results showed that the dataset had some imbalance between the target classes. The data also had some missing values and outliers, which were handled using appropriate techniques.

The analysis showed that there was a high correlation between employee attrition and some of the variables like Job Satisfaction, Environment Satisfaction, Work Life Balance, Age, and Distance From Home. The analysis also revealed that employees with higher salaries and job levels tend to have lower attrition rates.

The project concluded that employee attrition is a complex issue influenced by several factors. Machine learning models can be used to predict employee attrition with reasonable accuracy. The insights gained from this analysis can be used by organizations to identify the factors that influence employee attrition and take necessary actions to reduce it.

Researcher: sabune radhika

Qualification: Master's in Statistics

Thanks...