Forecasting Employee's Retention to Companies Using Deep Learning Models

By student:

Yaqeen Alsmadi

Supervisor

Dr .Huthaifa Riyad Aljawazneh

Al-Zaytoonah University of Jordan, 2024

Abstract

A crucial indicator of an organization's ability to reduce turnover among workers who quit or are fired within a given period of time is employee retention. Most firms place a high priority on retention since it directly affects overall success and corporate performance.

This thesis undertook the task of predicting employee retention using two distinct approaches with varying complexity levels: Deep Learning and Machine Learning algorithms based on dataset taken from an International Company that works on IT sector. In the realm of Deep Learning, the study employed deep neural networks (DNN) configured with four hidden layers, as well as two long short-term memory networks (LSTMs) using configurations of three and six units. On the Machine Learning level, the algorithms included Support Vector Machine (SVM) and a neural network with a single hidden layer. Due to the dataset's imbalance, five balancing algorithms were implemented: SMOTE, SMOTE-ENN, SMOTE-Tomek, SVM-SMOTE, and Borderline-SMOTE.

The efficiency of these algorithms was assessed using six key classification metrics: Accuracy, Recall, Specificity, Precision, Type I Error, and Type II Error. With respect to Machine Learning, the findings revealed that the LSTM model with six layers, using SMOTE as an advanced balancing technique, consistently outperformed other approaches across all metrics— Accuracy, Recall, Specificity, Precision, Type I Error, and Type II Error. This superior performance was achieved by integrating effective data balancing techniques, specifically SMOTE-Tomek and Borderline-SMOTE, which proved highly effective in enhancing predictive Accuracy. In summary, by employing sophisticated techniques like SMOTE-Tomek and Borderline-SMOTE across a spectrum of Deep Learning and Machine Learning algorithms, organizations can better address the complexities of employee turnover and optimize their workforce management strategies for sustained success. Below are several recommendations derived from the thesis findings: 1) Utilizing SMOTE as the primary method for data balancing, with two distinct types: SMOTE-Tomek and Borderline-SOMTE. 2) Utilizing LSTM with six layers yielded the optimal performance outcomes based on the six defined parameters.

Keywords: SMOTE, Deep Learning, Machine Learning, Employee Retention.