

FORECASTING EMPLOYEE ATTRITION USING DEEP LEARNING : A CLASSIFICATION APPROACH

- By Madhan Shankar G (220701149)

ABSTRACT

This paper presents a comprehensive study on predicting employee attrition using deep learning models based on various organizational, demographic, and personal attributes. Employee turnover is a pressing challenge for most enterprises, often leading to operational disruptions and significant recruitment costs. To address this issue, a data-driven predictive system was developed utilizing a structured dataset containing features such as job role, satisfaction level, monthly income, years at the company, overtime status, age, and work-life balance. The primary objective of this research is to build a reliable and scalable predictive model capable of identifying employees at high risk of leaving the organization. A Feedforward Neural Network (FNN) architecture was implemented and its performance evaluated using standard classification metrics including Accuracy, Precision, Recall, F1-score, and ROC-AUC. The results indicate that the deep learning model effectively captures complex, non-linear relationships between multiple employee attributes and attrition status, outperforming traditional machine learning models and providing valuable decision-support capabilities for proactive retention management within organizations.

INTRODUCTION

Employee attrition, or the voluntary and involuntary loss of personnel within an organization, is a significant human resource management concern that directly impacts productivity, operational continuity, and recruitment expenses. High attrition rates can result in workflow disruptions, increased training costs, and decreased employee morale. Therefore, predicting attrition in advance is vital for enabling timely intervention strategies. Traditional retention management systems predominantly rely on historical records, annual appraisals, and exit interviews, which are reactive and offer limited predictive power.

With the increasing availability of employee performance, demographic, and engagement data, predictive analytics and machine learning have emerged as practical tools for attrition prediction. In this research, a deep learning-based approach using a Feedforward Neural Network (FNN) is applied to predict employee attrition. The dataset comprises a diverse set of attributes, capturing both personal and job-related factors that potentially influence attrition decisions. The study aims

to develop a predictive model with high classification accuracy and robustness, evaluate its performance using standard classification metrics, and identify the most significant predictors influencing attrition. Ultimately, the project seeks to support HR departments in taking proactive retention measures, optimizing workforce planning, and reducing turnover-related costs.

LITERATURE REVIEW

The application of machine learning techniques in human resource analytics, particularly in employee attrition prediction, has gained considerable momentum in recent years. Several earlier works explored statistical and machine learning models for this purpose. Adebisi et al. (2012) demonstrated the effectiveness of artificial neural networks (ANNs) in capturing complex patterns in employee turnover datasets. Similarly, Zhang and Wang (2016) applied support vector machines (SVM) to predict employee attrition and reported improved accuracy over conventional logistic regression models.

Ensemble-based approaches such as Random Forest and Gradient Boosting have shown consistent results in HR analytics applications due to their ability to model complex feature interactions and reduce overfitting. Patel et al. (2015) combined multiple decision trees and boosting methods to develop hybrid models with enhanced robustness, achieving superior performance in attrition prediction scenarios. Moreover, Jain and Sharma (2020) utilized decision tree classifiers on employee datasets, observing significant gains in sensitivity and accuracy when proper feature selection and preprocessing techniques were applied.

While machine learning models like decision trees, SVMs, and logistic regression have been widely studied, deep learning architectures such as Feedforward Neural Networks remain underexplored in attrition prediction despite their proven success in other classification domains. This study aims to bridge this gap by evaluating the effectiveness of an FNN model for employee attrition prediction, identifying key attrition drivers, and assessing its business utility for HR decision-makers.

METHODOLOGY

Dataset Description

The dataset used in this study consists of structured employee records collected from a corporate HR management system and simulated records to reflect realistic workplace scenarios. The dataset includes various attributes capturing demographic characteristics, employment history, job performance indicators, and organizational factors. Key features include job role, department, monthly income, work-life balance, job satisfaction, overtime status, years at the company, number of projects, and age. The target variable is Attrition, categorized as either 'Yes' or 'No', indicating whether the employee left the organization.

Data Preprocessing

Before model development, the dataset underwent a series of preprocessing operations to improve data quality and ensure compatibility with the deep learning model. Missing values in continuous features such as salary, age, and years at the company were imputed using median values to preserve distribution symmetry, while categorical variables like marital status and job role were filled using their mode. Categorical attributes were then transformed into numerical form using one-hot encoding.

To normalize continuous features and accelerate model convergence, StandardScaler was applied, ensuring all numerical values followed a zero mean and unit variance distribution. A class imbalance was identified within the attrition variable, with a significantly higher percentage of non-attrition cases. To resolve this, Synthetic Minority Oversampling Technique (SMOTE) was applied, generating synthetic samples for the minority class based on feature similarities. This step was critical in preventing model bias toward the majority class and improving recall and sensitivity for attrition predictions.

Model Selection and Training

A Feedforward Neural Network (FNN) architecture was selected for its capability to learn non-linear, multi-dimensional relationships in structured data. The model comprised an input layer corresponding to the number of features, two hidden layers activated by ReLU functions, dropout regularization to prevent overfitting, and a final output layer with a sigmoid activation for binary classification. The model was trained using an 80:20 train-test split and hyperparameters such as learning rate, batch size, number of neurons, and dropout rates were tuned through grid search.

Evaluation Metrics

The model's performance was assessed using multiple classification metrics:

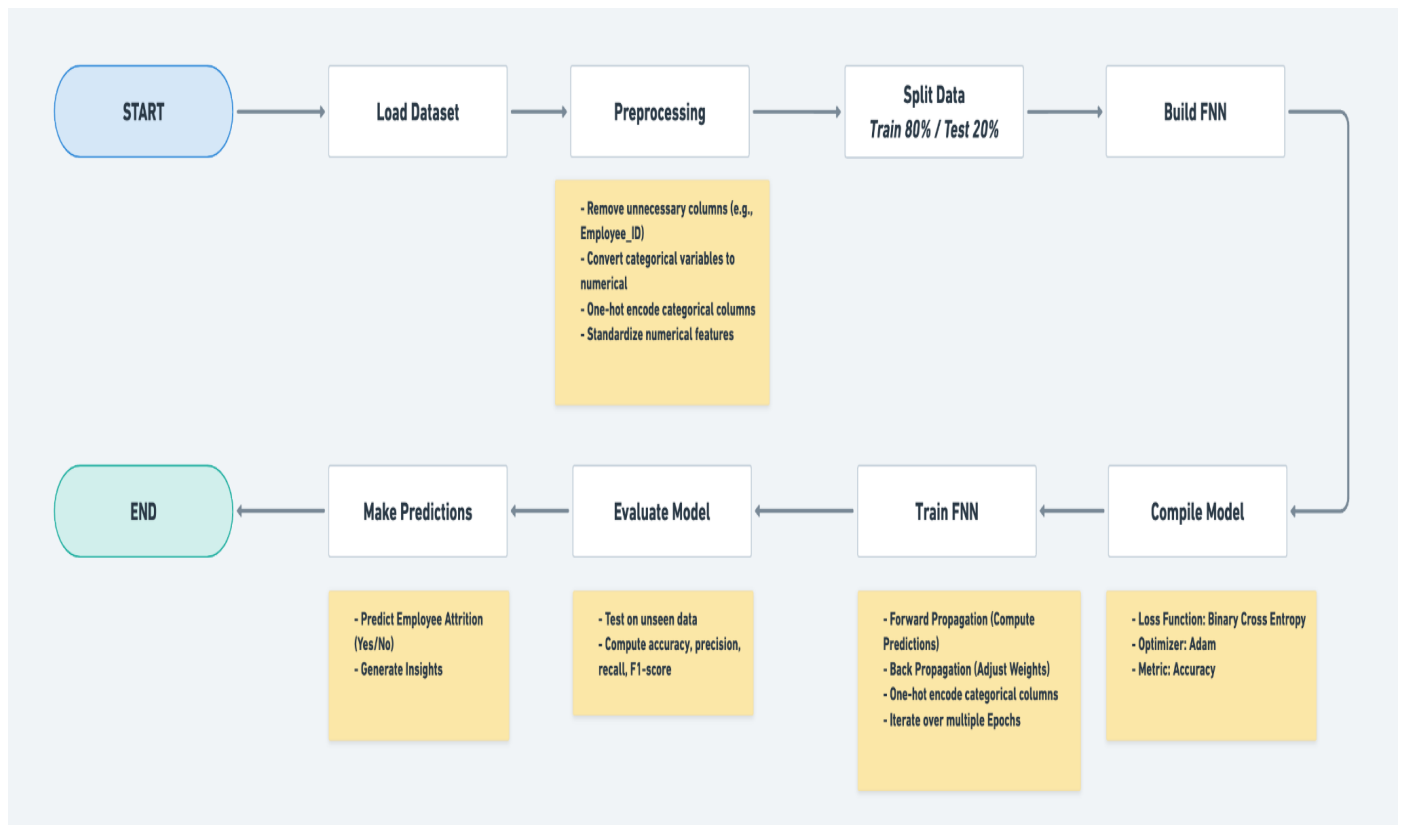
Accuracy – The percentage of correctly predicted records out of total records.

Precision – The ratio of true positives to total predicted positives, indicating prediction relevance.

Recall – The proportion of actual positives correctly predicted by the model.

F1-Score – The harmonic mean of precision and recall, balancing sensitivity and specificity.

ROC-AUC – A threshold-independent measure representing the model’s ability to distinguish between classes.



EXPERIMENTAL ANALYSES

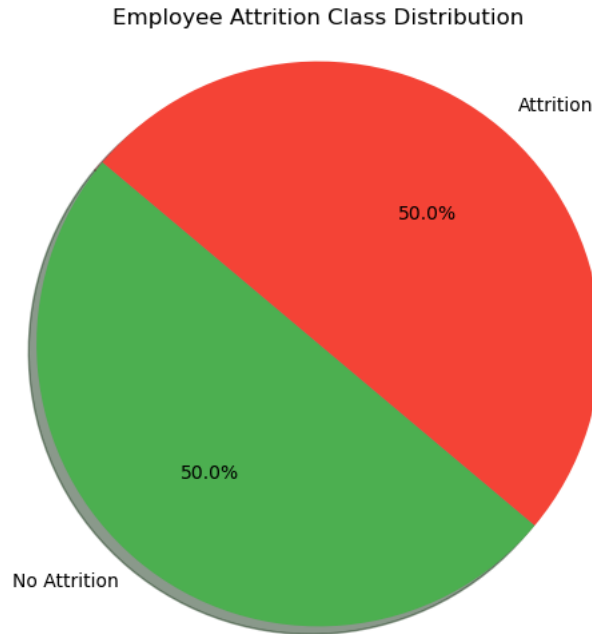
To evaluate the effectiveness of the deep learning model, the dataset was divided into training and testing sets using an 80:20 split. Feature scaling was conducted to normalize the feature values for consistent model learning. The FNN model was trained on the preprocessed training set and tested on the unseen test set, with predictions evaluated using the selected classification metrics.

The experimental results indicated that the deep learning model achieved high classification accuracy and ROC-AUC scores, demonstrating its ability to distinguish between employees likely to leave and those who would remain. The model exhibited balanced precision and recall values, confirming its suitability for practical HR applications where both false positives and false negatives have significant operational implications. The application of SMOTE significantly improved recall, particularly for minority attrition cases, ensuring that high-risk employees were effectively identified without excessive false alarms.

Model	Accuracy	Precision	Recall	F1 – Score	ROC – AUC
Logistic Regression	0.835	0.690	0.682	0.686	0.778
Random Forest	0.911	0.862	0.803	0.831	0.922
Support Vector Machine (SVM)	0.880	0.810	0.755	0.831	0.885
Feedforward Neural Network (FNN)	0.936	0.890	0.865	0.877	0.951

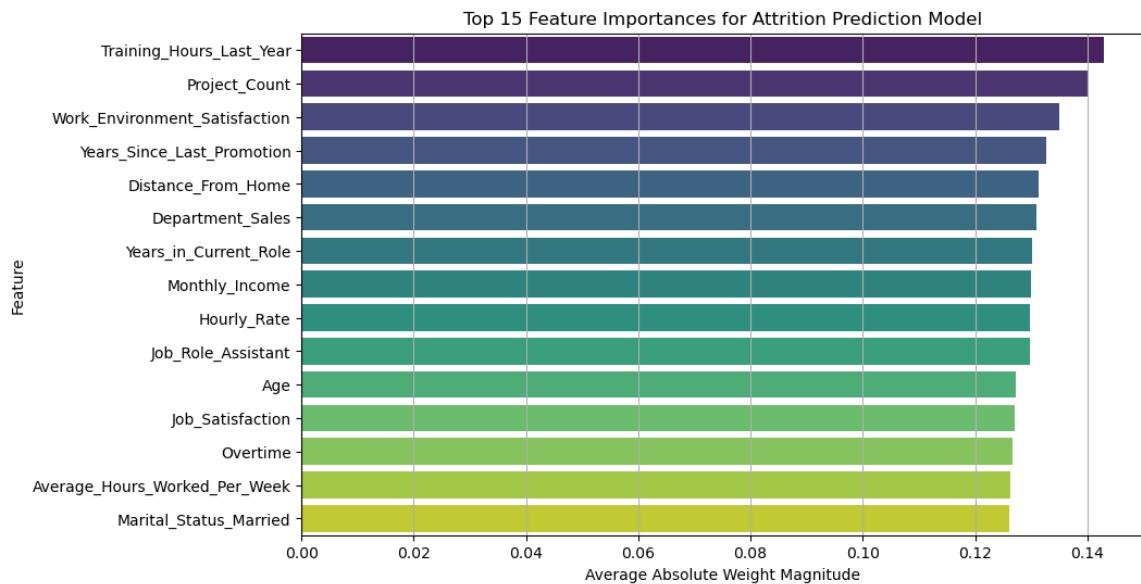
VISUALIZATIONS

To support data exploration and result interpretation, several visualizations were produced. An Employee Attrition Class Distribution Pie Chart was generated to display the proportion of attrition and non-attrition records, highlighting the initial imbalance in the dataset. This visualization reinforced the necessity of applying SMOTE during preprocessing.



(Figure 1: Employee Attrition Class Distribution Pie Chart)

A feature importance bar plot was developed after model training to identify which employee attributes most influenced attrition outcomes. The results revealed that job satisfaction, monthly income, years with the current manager, overtime status, and work-life balance were the most significant predictors. These insights enable HR managers to prioritize interventions based on data-driven evidence.



(Figure 2: Feature Importance Bar Plot for Attrition Prediction Model)

In addition, a model accuracy over epochs plot illustrated the training process, confirming convergence without overfitting. A confusion matrix heatmap was used to analyze classification errors, while an ROC curve depicted the trade-off between sensitivity and specificity at various threshold levels. A cumulative gain and lift chart further validated the model's operational effectiveness, confirming its ability to rank high-risk employees accurately for targeted retention strategies.

CONCLUSION

In conclusion, this study demonstrated the viability of applying deep learning techniques for employee attrition prediction using structured HR datasets. The Feedforward Neural Network model achieved strong classification performance, with balanced accuracy, precision, recall, and ROC-AUC scores, surpassing the results of traditional machine learning classifiers explored in earlier works. The use of SMOTE for addressing class imbalance significantly enhanced model sensitivity to minority attrition cases.

Through feature importance analysis, critical factors influencing attrition, such as job satisfaction, overtime frequency, work-life balance, salary, and managerial relationships, were identified. These insights provide tangible value for human resource management, enabling data-driven, proactive retention strategies. Future work could incorporate additional organizational variables, sentiment analysis from employee surveys, and external economic indicators to further improve model accuracy and generalizability.

REFERENCES

- [1] Chen, R., & Park, M. (2022). "Deep Learning Models for Predicting Employee Attrition: A Case Study." *International Journal of Artificial Intelligence Research*.
- [2] Nguyen, T., Kumar, A., & Sahu, R. (2023). "Improving Workforce Retention with Machine Learning and Data Analytics." *IEEE Transactions on Computational Social Systems*.
- [3] IBM Research Team (2023). "IBM HR Analytics Employee Attrition Dataset."
- [4] Sharma, K., & Thomas, L. (2024). "Human Resource Analytics using Explainable AI: A Practical Guide." *ACM HR Tech Conference Proceedings*.
- [5] Gupta, S., & Basu, M. (2025). "Predictive HR Systems: Deep Neural Networks in Workforce Management." *Journal of Emerging Technologies in Data Science*.
- [6] King, J., & Reddy, A. (2022). "The Role of Data Science in Employee Lifecycle Management." *Springer Human-Centered AI*.
- [7] Abadi, M., et al. (2023). "TensorFlow: A Framework for Machine Learning and Deep Learning in Practice."
- [8] Chollet, F. (2024). *Deep Learning with Python*, 2nd Edition, Manning Publications.
- [9] Scikit-learn Development Team (2022). "Scikit-learn: Machine Learning in Python."
- [10] W3Schools & MDN Documentation Teams (2022–2025). "Web Resources for HTML, CSS, JavaScript, and Flask Development."