

# **International Journal of Research Publication and Reviews**

Journal homepage: www.ijrpr.com ISSN 2582-7421

# **Employee Attrition Prediction Using Machine Learning**

Selvabharathi R<sup>[1]</sup>, Pravin P<sup>[2]</sup>, Chandru S<sup>[3]</sup>, Madhubalan S<sup>[4]</sup>, Nijanthan N<sup>[5]</sup>

[1,2,3,4] B.E. Computer Science and Engineering,

<sup>[5]</sup> Assistant Professor, M. Tech (IT),

Department of Computer Science and Engineering, Pavai College of Technology, Paavai Institutions, Paavai Nagar, NH-44, Pachal -637 018. Namakkal Dist., Tamil Nadu, India.

#### ABSTRACT :

Employee attrition is a critical concern for organizations as it directly impacts productivity, morale, and the overall cost of recruitment and training. Understanding and predicting which employees are at risk of leaving can help companies implement proactive retention strategies. The project begins with a comprehensive analysis of the IBM HR Analytics Employee Attrition & Performance dataset, which contains various employee-related features such as age, job role, monthly income, job satisfaction, years at the company, and overtime frequency. Data preprocessing steps, including handling missing values, feature selection, and normalization, were performed to prepare the dataset for training. Several machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), AdaBoost, Naive Bayes, Linear Discriminant Analysis, K-Nearest Neighbor, and Multilayer Perceptron (MLP), were implemented and compared. The final model provides not only accurate predictions but also interpretable insights into the reasons behind employee departures. This project demonstrates how machine learning can serve as a powerful tool in human resource management by transforming traditional reactive retention strategies into proactive decision-making processes.

Keywords :SVM, Machine learning, Naïve bayes, multilayer perceptron, Random forest, K-Nearest Neighbor, Adaboost.

### 1. Introduction

Employee attrition is a significant challenge for organizations, as it leads to increased recruitment costs, the loss of experienced personnel, and a decline in overall productivity. Predicting attrition enables companies to take proactive measures aimed at improving employee retention, boosting job satisfaction, and minimizing operational disruptions. Machine learning techniques offer an effective approach to analyzing employee data, uncovering patterns, and accurately forecasting attrition risks. Various supervised learning algorithms, including Logistic Regression, Decision Trees, Random Forest, Support Vector Machines (SVM), and Neural Networks, are commonly employed for this purpose. The primary objective of this study is to evaluate and compare the performance of these machine learning models to identify the most effective one for predicting employee turnover, thereby enabling organizations to make informed, data-driven decisions.

This was our very short introduction for the application of Employee attrition prediction, below will have the detailed explanations.

## 2. Objective

- To identify key factors influencing employee attrition by analyzing historical employee data.
- To build and compare multiple machine learning models.
- To evaluate the performance of the models using appropriate metrics.
- To provide actionable insights to HR teams to help them implement retention strategies.
- To develop a user-friendly system where HR professionals can input employee data.

## 3. Literature Review

- Employee Turnover Prediction with Machine Learning. A Reliable Approach.
- Predicting Employee Attrition Using Machine Learning Techniques.
- Prediction of Employee Attrition Using Machine Learning and Ensemble Methods.
- Prediction of Employee Turnover in Organizations Using Machine Learning Algorithms.
- Analyzing Employee Attrition Using Decision Tree Algorithms.

## 4. Existing System

- Traditional statistical methods and basic machine learning models are commonly used for predicting employee attrition.
- Logistic Regression and Decision Trees are widely applied but may not capture complex patterns in employee data.
- Many existing systems do not explore a diverse range of algorithms, limiting prediction accuracy.
- The lack of advanced techniques reduces the effectiveness of identifying at-risk employees.
- A more comprehensive machine learning approach is needed to improve the accuracy of attrition prediction.

#### 5. Proposed System

This project core features include:

- The system utilizes multiple supervised machine learning models to predict employee attrition with higher accuracyGitHub integrated codebase like repository fetching is possible now.
- Employee data, including factors like age, tenure, job role, and performance scores, is collected, cleaned, and preprocessed for analysis.
- Various machine learning models such as Logistic Regression, Naive Bayes, Decision Trees, Random Forest, AdaBoost, SVM, LDA, KNN, and MLP are trained and evaluated.
- The models are compared based on key performance metrics like accuracy, precision, recall, and computational efficiency to determine the most effective model.
- The final selected model provides HR teams with actionable insights to identify at-risk employees and implement strategies to reduce attrition.

## 6. Module Description

- Data Collection Module: This module gathers employee data from existing HR databases or external datasets. It includes attributes like age, department, salary, experience, job satisfaction, and performance metrics.
- Data Preprocessing Module: Cleans and prepares the data by handling missing values, encoding categorical data, and normalizing numerical features to ensure compatibility with machine learning algorithms.
- Model Building Module: Implements various machine learning models such as Logistic Regression, Decision Tree, Random Forest, SVM, and others to train on the employee data for predicting attrition.
- Evaluation Module: Assesses the performance of each model using evaluation metrics such as accuracy, precision, recall, and F1-score to select the most effective model.
- Prediction & Visualization Module: Uses the best-performing model to predict employee attrition and displays insights through charts and dashboards for HR decision-making.

## 7. Architecture Design

The system architecture design was the master piece of the project because this drives the entire success pattern and for right problem-solution answer for the nature of the project .This architecture was designed using user interface layer, application layer(backend), machine learning layer and database layer.

In User Interface Layer there is HR/Admin inputs employee data or uploads datasets (CSV/Excel) and View prediction results, graphs, and report In Application Layer (Backend) there is Handles input validation and business logic, Connects UI to ML models, Provides endpoints for predictions and visualizations.

In Machine Learning Laye there is Data Preprocessing for Cleans and transforms data, Model Training for Trains ML models like Logistic Regression, Random Forest, SVM and Model Evaluation for Compares models based on accuracy and other metrics and Prediction Engine for Uses the selected model to predict employee attrition.

In Database Layer there is Stores employee data, model results, and logs and Optionally maintains user information (login/roles).

### 9. Summary

This project aimed to predict employee attrition using machine learning on HR analytics data. Techniques like PCA and SMOTE were applied to handle imbalance and improve accuracy. Among various models, MLP with PCA and SMOTE gave the best results. The system shows strong potential to aid HR decisions effectively.

#### Acknowledgements

The Faculty Members, Authors, Co-Student and Co-Producer are the biggest contributor for the project completion. I now want to thank them with full of gratitude and full respect, without their contribution I didn't able to climb this height and finally my family backed a lot, my special thanks to them also.

#### **REFERENCES:**

- 1. P Aseel Qutub et al. "Prediction of Employee Attrition Using Machine Learning and Ensemble Methods". In: International Journal of Machine Learning and Computing 2021
- 2. M. V. Jose and T. Chakkalakkal. "Deep Learning-Based Prediction of Employee Turnover". In: International Journal of Advanced Computer Science and Applications, 2021
- 3. K. R. See and H. H. Soon" A Comparative Study of Supervised Machine Learning Algorithms for Employee Attrition Prediction Journal of Information and Communication Technology, In: Journal of Information and Communication Technology, 2020
- 4. A. Tzimas, G. Giannakopoulos, and C. Kyriazis. "Analyzing Employee Attrition Using Explainable AI Techniques". In: Expert Systems with Applications, 2020
- 5. Ahmed M. Alghamdi. "Predicting Employee Attrition Using Ensemble Learning Techniques". In: Procedia Computer Science, 2019