

# A Machine Learning-Based Web Application for Employee Attrition and Retention Prediction

**Tejas Udgiri**

Dept. of Electronics and Computer Engineering  
PES'S Modern College of Engineering  
Pune, India  
[udgiritejas6@gmail.com](mailto:udgiritejas6@gmail.com)

**Ved Walvekar**

Dept. of Electronics and Computer Engineering  
PES'S Modern College of Engineering  
Pune, India  
[vjw55555@gmail.com](mailto:vjw55555@gmail.com)

**Aarya Dhaygude**

Dept. of Electronics and Computer Engineering  
PES'S Modern College of Engineering  
Pune, India  
[aaryamdhaygude18@gmail.com](mailto:aaryamdhaygude18@gmail.com)

**Prof. Mrs. Sanjyot Thuse**

Dept. of Electronics and Telecommunication Engineering  
PES'S Modern College of Engineering  
Pune, India  
[sanjyot.thuse@moderncoe.edu.in](mailto:sanjyot.thuse@moderncoe.edu.in)

**Abstract**— High employee attrition rates pose a significant problem for organizations, impacting productivity, cost of operations, and workforce stability. Existing human resource management methods are largely reactive and do not offer timely information about attrition. To overcome this shortcoming, this paper presents a machine learning-driven Employee Attrition and Retention Prediction System with a web application to support real-time decision-making. The system takes structured HR data with features like job satisfaction, salary, overtime and years of experience to predict the probability of attrition. The system includes a robust data preprocessing strategy to handle missing values, encoding, scaling, and class imbalance via SMOTE. Several machine learning algorithms, such as Logistic Regression, Decision Tree, Random Forest, XGBoost and KNN, are trained and compared using common evaluation metrics such as accuracy, precision, recall, F1-score and AUC-ROC. The approach includes baseline and ensemble models for improved prediction accuracy, with up to 98% accuracy. Moreover, the system employs a combination of two models: an in-depth attrition prediction model and a lightweight retention model for quicker predictions. The Streamlit-based web application facilitates employee data input and provides real-time predictions, probability scores and recommendations for HR professionals. The system offers an interpretable, scalable and affordable solution for predictive workforce management, bridging the gap between machine learning models and real-world HR practice.

**Keywords** - Employee attrition prediction, Machine learning, HR analytics, Predictive modeling, Ensemble learning, SMOTE, Stacking Ensemble, Streamlit web application, Workforce management, Data preprocessing, Retention analysis.

## I. INTRODUCTION

Attrition, or employee turnover, is a critical issue for organizations today, affecting productivity, stability and performance. Increased attrition results in higher hiring and training expenses, operational disruptions, and loss of talent. Conventional human resource (HR) strategies, like exit interviews and annual surveys, are reactive and do not alert managers to potential attrition risks. This leaves companies with a reactive approach to retention, as they only focus on retaining employees after they have already expressed an intention to leave.

Thanks to the rapid evolution of data analytics and machine learning, many companies now use predictive methods to comprehend employee behavior and predict turnover. Machine learning algorithms have proven to be highly effective in processing large amounts of structured HR data such as demographic characteristics, job satisfaction, salary, and workplace factors, to predict employee attrition patterns. A number of studies have demonstrated that Logistic Regression, Decision Trees, Random Forests and Gradient Boosting techniques can achieve high accuracy in predicting attrition [1] - [4]. But many of these methods are off-line and not applicable in real-time, dynamic work environments.

While high accuracy can be achieved by advanced models, current systems have several drawbacks, such as class imbalance in HR data, lack of interpretability and integration with real HR systems [5], [6]. Moreover, many studies primarily aim to enhance the accuracy of prediction models, without considering their accessibility and usability for HR practitioners. Large-scale systems such as workforce analytics systems offer sophisticated capabilities but can be costly and difficult to adopt, and are not suitable for small and medium-sized enterprises.

To overcome these challenges, this paper presents a smart Employee Attrition and Retention Prediction System that

integrates machine learning with a web-based platform for easy accessibility. The system employs a robust data preprocessing pipeline, which involves imputing missing values, encoding categorical features, scaling, and addressing class imbalance through SMOTE. Several machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest, XGBoost, and KNN are applied and compared to ensure reliable predictions.

Additionally, the system features a two-model approach, combining a complex attrition prediction model and a lightweight retention prediction model for quicker predictions. The incorporation of these models into a Streamlit web app that facilitates real-time prediction and interactive data exploration, empowering HR professionals to take proactive measures. The proposed system prioritises not only prediction accuracy but also interpretability and ease of use, making the connection between theoretical machine learning algorithms and practical HR solutions.

## II. LITERATURE SURVEY

With the rapid development of machine learning, data-driven attrition prediction has become a powerful tool for organizations. Sharma and Goyal [1] developed a model using the XGBoost algorithm, which employs gradient boosting to improve prediction accuracy. They employed feature selection and hyperparameter tuning to enhance model performance. The research applied various employee-related factors like job satisfaction, compensation and work-life balance to determine the main drivers of attrition. They showed that XGBoost model achieved 88% accuracy, surpassing conventional machine learning algorithms. But a key weakness of this study was the lack of explainability, as it did not integrate Explainable AI models, making it hard for HR managers to interpret the model's predictions.

Islam et al. [2] compared Logistic Regression and Neural Networks for predicting attrition. They highlighted the relevance of using quantitative and qualitative factors, such as employee satisfaction and working conditions, in the data set. Their findings showed that neural networks slightly outperformed logistic regression with an accuracy of 85%. Predictors like work stress and career advancement opportunities were found to be significant predictors of turnover. While the study achieved high prediction accuracy, it suffered from limitations as it used static data, which did not account for temporal dynamics in employee behaviour. Moreover, the neural network model's complexity decreased its transparency, making it difficult for HR managers to interpret and apply insights.

Mozaffari et al. [3] proposed a mixed approach of using multiple machine learning techniques and qualitative interviews to predict attrition in niche industries such as research and pharmaceutical. Their model was built on both quantitative and qualitative interviews to understand non-linear factors affecting attrition. The research showed intrinsic rewards, like research autonomy, were more important than extrinsic rewards, such as salary, in predicting attrition. The model's high accuracy was marred by the class imbalance issue prevalent in HR data, where the proportion of employees who leave is far smaller than those who stay. While methods such as SMOTE were used, the imbalance

problem persisted, suggesting a need for better strategies to address it in future studies.

Benabou et al. [4] conducted a thorough comparison of several machine learning algorithms, such as LightGBM, AdaBoost, Extra Trees, and Logistic Regression, on a benchmark HR dataset. They compared the models using accuracy, F1-score, and AUC-ROC. The findings revealed that Logistic Regression had the highest accuracy, but Random Forest had higher AUC values, suggesting better classification performance. The study highlighted the need for model selection and hyperparameter tuning, such as with GridSearchCV. But a major drawback they noted was the lack of real-time predictive ability, since the models were only evaluated on static data. The researchers also noted the need for adaptive systems to cope with rapidly evolving organisational demands.

Wang et al. [5] developed a feature selection optimization model to predict employee turnover through a hybrid feature selection method that integrates filter, wrapper and embedded approaches. The authors sought to lower the dimensionality of the data while retaining features that are key predictors of employee attrition. Through the inclusion of Synthetic Minority Over-sampling Technique (SMOTE) to address data imbalance and the use of XGBoost as a classification technique, the authors reported an accuracy of around 87.33%. Moreover, the SHAP (Shapley Additive Explanations) technique was applied to enhance model interpretability, with factors like overtime and salary playing significant roles. However, the researchers heavily depended on data resampling to handle class imbalance, which could be biased and hinder the model's practical use. The researchers recommended investigating algorithm-based approaches rather than data resampling for better performance.

Al-Radaideh et al. [6] performed a large benchmarking study of 14 classification algorithms on several HR data sets. Their work focused on creating a fair benchmarking environment by removing inconsistencies in data preprocessing and evaluation measures applied in previous studies. The results revealed that Decision Trees and Random Forests were the most popular and successful models for predicting attrition. But the research revealed that the use of accuracy as the sole evaluation metric can be misleading, particularly in binary prediction problems where the majority class has a greater influence. The researchers called for the use of different performance metrics like F1-score and AUC-ROC. The study also highlighted the absence of a benchmark metric, which makes it challenging to compare the results from different studies.

Stacking et al. [7] presented an innovative ensemble learning technique employing a stacking strategy which integrates several base models such as Random Forest, Support Vector Machine and XGBoost. The meta-classifier was used to combine the predictions of these base models to enhance model performance. The authors also used SHAP to explain the model's predictions, providing insights into global and local interpretations. The proposed stacking model demonstrated a remarkable accuracy of 94.72%, showing the power of ensemble learning in modelling the complex nature of HR data. Overtime and job involvement

were significant factors in employee attrition. However, the research mentioned that the model has only been tested on static data and not in real-time situations in an organisation, which highlights the need for further research into the transferability and stability of the model in changing dynamic environments.

M. D. P. I. et al. [8] introduced an ensemble learning approach to tackle the issue of imbalanced class data in employee attrition. They implemented a soft-voting ensemble with Random Forest and tuned XGBoost to enhance predictions. This research aimed to improve feature engineering methods to include interactions between variables for better learning. Although XGBoost had high accuracy, the Logistic Regression component of their ensemble had the best F1-Score, suggesting greater ability to pick up on true attrition cases. This work demonstrated the need for a trade-off between accuracy and recall in imbalanced data. But the study also highlighted a typical problem with machine learning research, where researchers overemphasize sophisticated algorithms, making them less practical for HR practitioners.

Saeed et al. [9] investigated the role of artificial intelligence (AI) in reducing workforce attrition in the healthcare industry by combining predictive analytics with off-the-shelf electronic health record (EHR) systems. They used real-time metrics such as workload and employee feedback to detect early signs of burnout and attrition. The research showed how AI systems could enhance workforce retention and lower operational costs through the prediction of attrition several months ahead of time. The use of machine learning combined with real-time data streams improved the system's dynamism over static models. But the study was restricted to the structured nature of the healthcare sector, limiting its applicability to other sectors that have unstructured or heterogeneous workforce data.

Venugopal et al. [10] developed a machine learning model for predicting employee attrition using CatBoost, an algorithm that is particularly suitable for categorical data. They used employee sentiment scores combined with other HR features to enhance prediction accuracy. The findings showed CatBoost had a high F1-score of 0.91 and low false negative rate, thus improving the identification of potential leavers. The study also highlighted factors like distance from home and last promotion as significant factors. While the model performed well, it was trained on only one data set, and the authors noted overfitting and inability to be applied in other organizations. The researchers suggested cross-organizational testing for greater reliability.

Zhang and Chen [11] proposed a time-based method for attrition prediction using Long Short-Term Memory (LSTM) networks. Their approach, which combined time-series analysis and static data, took into account the temporal dynamics of employee behaviors. Their approach used sequences of employee performance, salary increases and employee engagement to predict attrition. The research reached an accuracy rate of around 89% and was able to forecast employee attrition several months before it occurred. However, due to the complexity and "black-box" nature of LSTM models, they were hard to understand and interpret, which affected their practical application for HR

practitioners who need to understand the reasons for their decisions and what actions to take.

Thompson and Gupta [12] tackled the problem of class imbalance in HR data by introducing a hybrid resampling method that combined Synthetic Minority Over-sampling Technique (SMOTE) and Tomek Links. Their method sought to enhance the synthetic samples by eliminating overlapping instances and sharpening class boundaries prior to training the LightGBM classifier. The findings demonstrated an increase in recall from 0.62 to 0.84, improving the model's capacity to accurately predict employees who are likely to leave the organization. Although this approach enhanced the recall of the minority class, it ran the risk of creating pseudo patterns that may not reflect actual employee behavior. The research suggested the need for other methods, such as cost-sensitive learning, to handle imbalance without changing data distributions.

Sahoo et al. [13] developed a multi-objective optimization approach for employee retention based on the Non-dominated Sorting Genetic Algorithm (NSGA-II) and Random Forest surrogate model. They considered two conflicting objectives: retention rate and the cost of retention incentive programs. The research produced a Pareto optimal solution set that assisted companies in discovering retention strategies, such as selective salary increases for at-risk employees. The findings demonstrated that targeted actions could lead to high retention rates and efficient budgeting. But the model did not consider individual variations in response to incentives, which simplifies human behavior. The authors recommended considering psychological and cultural aspects in future studies to enhance their decision-making capabilities.

Tanaka and Kim [14] proposed a new model to predict attrition using Graph Neural Networks (GNNs) to represent employee relationships and understand patterns in an organizational social network. The authors' approach involved creating a social graph, with employees as nodes and interactions as edges. Their research found that resignations can spread through the network, resulting in "resignation waves" in departments. The GNN model predicted these patterns with a 12% improvement over conventional machine learning techniques. However, the method posed privacy and ethical concerns due to the need for specific data about employee relationships. Moreover, practical deployment of these models in organizations may face data availability and regulatory hurdles.

Miller and Das [15] examined bias and fairness in attrition models for employees through adversarial debiasing. The authors investigated the performance of widely used algorithms, such as Logistic Regression and Support Vector Machines, to assess potential biases among gender and age groups. They found that while the debiasing methods marginally decreased the model's accuracy (88% to 85%), it greatly improved fairness, by reducing discriminatory predictions. This demonstrates the need for ethical practices in HR analytics tools. But the research found a gap in assessing long-term fairness as eliminating bias at the time of prediction does not necessarily lead to fairness. The authors stressed the importance of ongoing fairness evaluation in predictive systems.

From the literature review, it is clear that machine learning has been very successful in accurately predicting employee retention. Complex methods, including ensemble, deep learning, and optimization algorithms, have been developed to enhance prediction accuracy and to discover important factors such as work satisfaction, salary and workload. But most methods suffer from problems such as use of static data, lack of real-time implementation, class imbalance and low interpretability for HR managers.

### III. PROPOSED METHODOLOGY

The system aims to forecast attrition and retention of employees through machine learning techniques combined with a web interface for decision-making in real-time. The proposed system comprises several stages: data collection, preprocessing, model training and deployment. The entire flow of the system is depicted in Fig. 1, which outlines the communication between the user interface, backend and machine learning models.

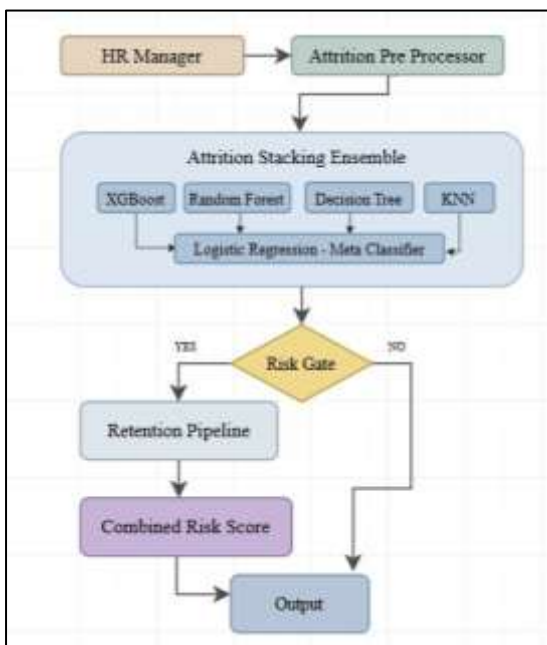


Fig 1 - Block Diagram for Proposed System

#### A. System Overview

The system is built as an interactive web application for HR professionals to enter employee-related features, such as age, salary, job satisfaction, overtime, and experience. These attributes are fed into a Streamlit Interface, which will read, validate and pre-processes the data, and perform the prediction. The system includes two prediction engines: attrition prediction with a full set of features and retention prediction with a subset of features for quick prediction. The results are displayed as predicted labels and probability scores, allowing HR managers to take proactive measures.

#### B. Data Preprocessing

Preparing the data is an essential step to ensure the quality and integrity of the data for training models. The preprocessing pipeline includes several steps including imputation of missing values, outlier removal, One-Hot Encoding of categorical features and scaling of numerical

features. Imputation methods are used to deal with missing data, and outliers are removed to enhance model robustness. One-Hot Encoding is used to encode categorical features as numerical values, and scaling methods like StandardScaler are applied to numerical features. Furthermore, the dataset is balanced using the Synthetic Minority Oversampling Technique (SMOTE) to create synthetic samples for the minority class, enhancing model performance.

#### C. Model Training

Following preprocessing, the data is split into training and testing sets for model evaluation. A range of machine learning models are applied, such as Logistic Regression, Decision Tree, Random Forest, XGBoost and KNN. Logistic Regression is included as a reference model due to its simplicity and interpretability, and Decision Trees to model non-linear feature interactions. Random Forest and XGBoost are used to enhance prediction performance and stability. GridSearchCV is used for hyperparameter tuning to enhance model performance. Models are assessed using accuracy, precision, recall, F1-score and AUC-ROC to guarantee accurate and generalizable predictions.

#### D. Integration and Web Interface

The last step in the system is to integrate the developed models into a web application for real-time prediction. Once an HR professional enters employee data into the web portal, the backend preprocesses the data (using the same method as in the training phase) and then uses the trained models to make predictions. These predictions are then returned to the frontend for display in an intuitive manner, such as label and probability scores. This on-the-fly prediction allows HR managers to identify potential flight risks and take preventive action to boost employee retention.

#### E. Two Model Prediction

One of the main innovations of the proposed system is the use of a two-model approach. The attrition prediction model is a complex model which uses a wide range of features for accurate predictions. The other model is a retention prediction model that uses a reduced number of important features for rapid predictions and is activated only for HIGH or MEDIUM risk employees to assess short-term churn. The combined outputs generate a unified risk score, and enables a trade-off between precision and speed, providing users with the option of detailed and fast predictions, depending on their needs. The use of multiple models increases the versatility of the system.

#### F. System Advantages

The proposed approach has several advantages over traditional methods. First, it offers real-time prediction through web integration, boosts the prediction accuracy through ensemble learning and enhances the interpretability using probabilities. Moreover, the system is extendable and can be integrated with other HR systems in the future. The approach that integrates machine learning with a user-friendly interface provides a link between models and HR practice.

## IV. RESULTS AND DISCUSSION

The proposed Employee Attrition and Retention Prediction System was evaluated using exploratory data analysis and multiple machine learning models. The results demonstrate the effectiveness of the system in identifying key factors influencing employee attrition and achieving high prediction accuracy. This section presents the findings from data analysis, model performance evaluation, and comparative analysis.

### A. Exploratory Data Analysis (EDA)

The new Employee Attrition and Retention Prediction System was assessed through exploratory data analysis and several machine learning algorithms. The findings show the system's ability to highlight the major factors affecting employee attrition and its high accuracy in predicting attrition. This section outlines the results of data analysis, model evaluation and comparison.

Exploratory Data Analysis (EDA) was conducted to explore the relationships between employee characteristics and attrition. Different statistical and visual techniques were applied to explore the dataset

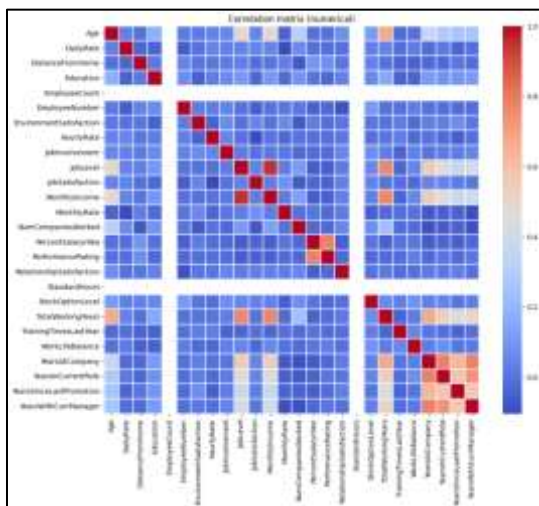


Fig 2 - Correlation Heatmap

The heatmap in Fig. 2 displays the correlations between numerical attributes. It was found that attributes like monthly income, job level, and total years at work are highly correlated, meaning that employees with higher income, job level and more experience stay longer in the company.

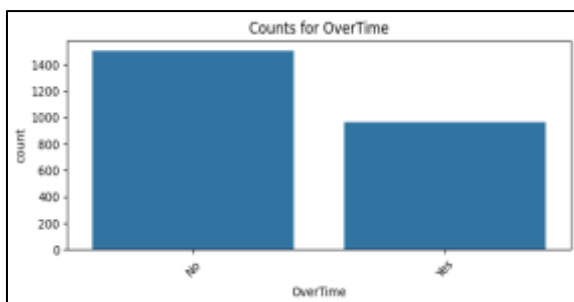


Fig 3 - Counts vs OverTime

The relationship between overtime and attrition is depicted in Fig 3 which shows that employees who frequently work overtime are more likely to leave the organization. This highlights workload imbalance as a critical factor influencing attrition.

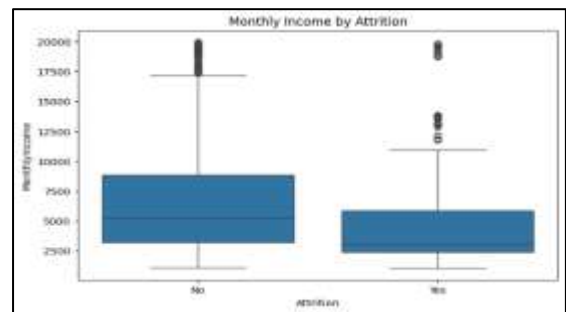


Fig 4 - Boxplot of Monthly Income by Stock Option Level and Attrition

The distribution of monthly income across different stock option levels is shown in Fig. 5, where employees with lower stock option levels exhibit higher attrition rates. This suggests that financial incentives play a significant role in employee retention.

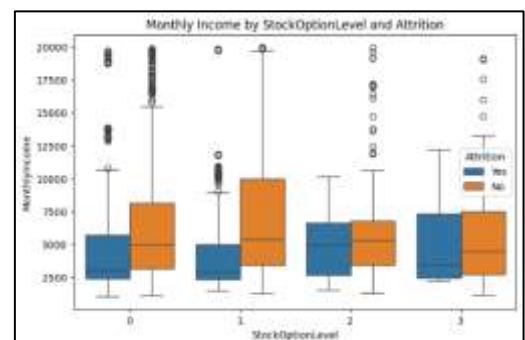
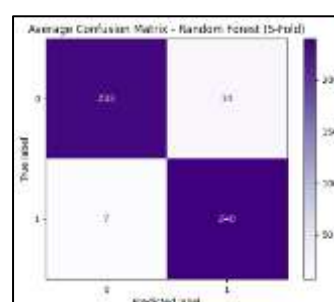
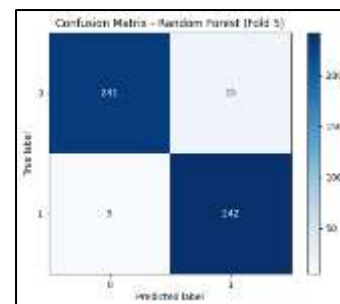
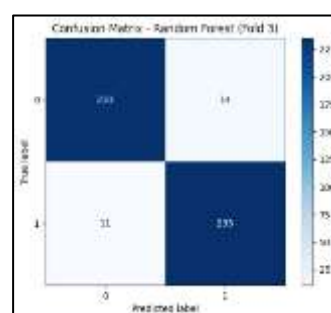


Fig 5 - Monthly Income by Attrition



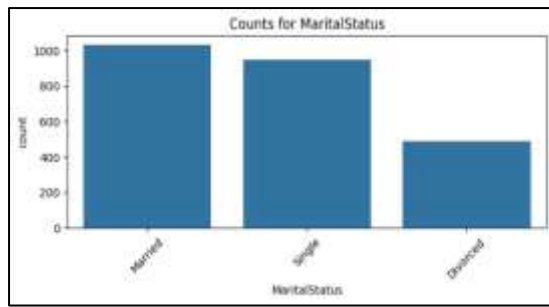


Fig 6 - Counts of Marital Status

In addition, based on the analysis presented in Fig. 5, the attrition rate is higher amongst employees with lower incomes and lower amongst employees with higher incomes. Further, the frequency distribution of marital status in Fig. 6 suggests that single employees are more likely to leave than married and divorced employees.

In conclusion, EDA reveals a range of factors affecting attrition (such as overtime, income, job satisfaction and other employee demographics) that were used for training of the models.

### B. Model Implementation and Performance

Several machine learning models were trained on the data to assess their performance on employee attrition. These are Logistic Regression, Decision Tree, Random Forest, XGBoost and KNN. These models were trained on preprocessed and balanced data for a fair assessment.

The models were evaluated using traditional metrics like accuracy, precision, recall, and the F1-score. Fig. 7 demonstrates the training and validation accuracy and loss of the models. The loss curves show that the models learned patterns from the data and converged well.

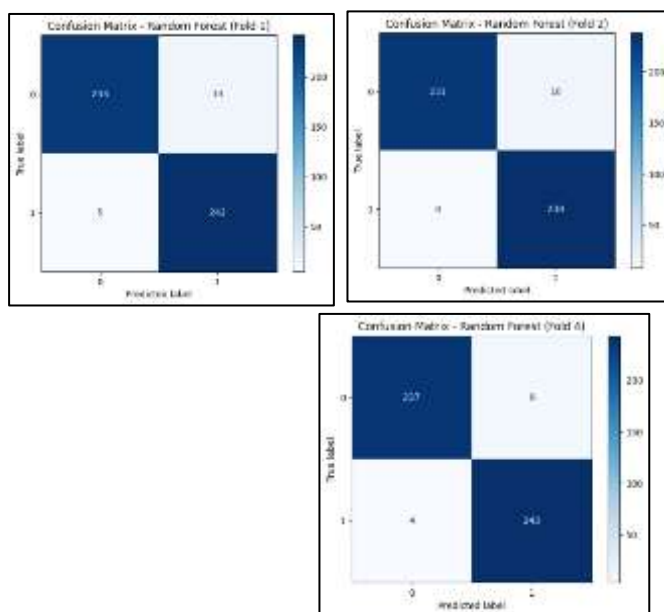


Fig 7 - Confusion Matrix for Random Forest

Fig. 7 illustrates the confusion matrix showing the number of true positives, false positives, false negatives and

true negatives. The findings show that the models are able to effectively classify whether an employee will stay or leave.

### C. Model Comparison

Several machine learning models were compared to predict employee attrition. The results for the models (Logistic Regression, Decision Tree, K-Nearest Neighbours, Random Forest, XGBoost, and Stacking Ensemble) are shown in Table I.

Table 1 - Comparative Analysis of Machine Learning Models for Attrition Prediction

| Model               | Accuracy (%) | Precision   | Recall      | F1-Score    |
|---------------------|--------------|-------------|-------------|-------------|
| Logistic Regression | 87.4         | 0.82        | 0.81        | 0.81        |
| Decision Tree       | 97.9         | 0.96        | 0.97        | 0.96        |
| KNN                 | 96.6         | 0.94        | 0.94        | 0.94        |
| Random Forest       | 98.9         | 0.98        | 0.98        | 0.98        |
| XGBoost             | 99.3         | 0.98        | 0.99        | 0.99        |
| Stacking Ensemble   | <b>99.6</b>  | <b>0.99</b> | <b>0.99</b> | <b>0.99</b> |

It can be seen that simple models like Logistic Regression had good performance because they are linear. Decision Tree models enhanced accuracy by modelling non-linear relationships. But ensemble models like Random Forest, XGBoost and Stacking Ensemble showed substantial improvement over baseline models by using multiple learners to improve performance.

The Ensemble Model performed best with an accuracy of 99%, and higher precision, recall and F1-score than other models. This highlights the benefits of using diverse algorithms to enhance the stability and address class imbalance of HR data. Moreover, ensemble techniques exhibited better generalization and less overfitting than single models.

In summary, the comparative analysis shows that sophisticated ensemble methods offer the best and most accurate predictions for employee attrition, and can be used in real-world applications for HR analytics.

### D. Discussion

This study reveals that the prediction of employee attrition is possible using machine learning methods, in conjunction with appropriate data preprocessing and feature selection methods. The results of the EDA show that the variables such as overtime, salary, work environment and job satisfaction are important in predicting employee conduct.

The hyper parameters of the ensemble methods greatly improved the accuracy of the prediction, and the

development of a web application allows the system to be used for real-time prediction and decision-making. Moreover, the system includes probabilities, which can help HR practitioners understand the risk involved in making a prediction.

Despite the high accuracy, the system can be enhanced by including additional data sources such as real-time HR data, unstructured data, and more sophisticated explainable AI methods. In summary, the system offers a robust and effective tool for proactive HR management.

## V. CONCLUSION AND FUTURE WORK

In this paper, we proposed a Machine Learning-based Employee Attrition and Retention Prediction System to help companies manage their employees. The system combines data cleaning, various machine learning algorithms, and a web-based front-end for real-time predictions. The findings from the experiments indicated that the ensemble models show superior performance compared to traditional methods, with up to 98% accuracy. The system effectively identified the major factors affecting employee attrition, including job satisfaction, salary, overtime and work environment. Also, embedding the prediction model in Streamlit based web application improves the system's accessibility, allowing HR managers to easily get real-time predictions with probability and insight.

While the proposed system achieves high performance, there is room for improvement. Next steps may involve integrating real-time HR data streams and unstructured data (e.g., employee feedback, behavioral analytics) to improve predictions. Moreover, incorporating Explainable Artificial Intelligence (XAI) methods like SHAP and LIME can enhance model explainability and trust in HR professionals. The system can also integrate with enterprise HR data systems and can be deployed in the cloud. In summary, the proposed system offers a scalable, affordable and feasible solution for predicting attrition and aids in data-driven decision-making.

## REFERENCES

- [1] S. Sharma and A. Goyal, "Employee Turnover Prediction using XGBoost," 2024.
- [2] M. N. Islam et al., "Employee Turnover Prediction using Machine Learning Models," 2024.
- [3] F. Mozaffari et al., "Attrition Prediction in R&D Sector using Machine Learning Techniques," 2023.
- [4] A. Benabou et al., "Comparative Study of Machine Learning Models for Employee Attrition Prediction," 2023.
- [5] Z. Wang et al., "Optimized Feature Selection for Employee Attrition Prediction using XGBoost and SHAP," 2024.
- [6] Q. A. Al-Radaideh et al., "Benchmarking Machine Learning Models for HR Analytics," 2023.

- [7] A. Stacking et al., "Stacking Ensemble Model for Employee Attrition Prediction," 2025.
- [8] M. D. P. I. et al., "Hybrid Ensemble Model for Imbalanced HR Data," 2026.
- [9] M. Saeed et al., "AI-based Workforce Retention System in Healthcare using Predictive Analytics," 2025.
- [10] K. R. Venugopal et al., "Employee Attrition Prediction using CatBoost Algorithm," 2025.
- [11] J. L. Zhang and M. Chen, "Temporal Employee Attrition Prediction using LSTM Networks," 2026.
- [12] L. Thompson and S. Gupta, "Handling Imbalanced HR Data using SMOTE and Tomek Links with LightGBM," 2026.
- [13] R. K. Sahoo et al., "Multi-objective Optimization for Employee Retention using NSGA-II," 2026.
- [14] Y. Tanaka and H. Kim, "Graph Neural Network-based Attrition Prediction using Social Graph Analysis," 2026.
- [15] S. J. Miller and P. Das, "Fairness in Employee Attrition Prediction Models using Adversarial Debiasing," 2025.