ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

Article Info

Date Received: 15/03/2026
Date Revised: 05/04/2026
Available Online: 27/04/2026

A Workforce Harmony Predictor System for Proactive Employee Churn Identification and Strategic Retention Intervention

1. N. Bhaskar, 2. N. Pavan Sai, 3. N. Yagna Harshitha, 4. N. D. S. N. V. G. Varalakshmi, 5. N. Anusha

Author Affiliations

1,2,3,4,5 B. Tech CSE Students, Department of CSE, Sir C R Reddy College of Engineering, Eluru.

DOI: 10.64264/ijisea/0732

ABSTARCT

Employee attrition has become a serious concern for many organizations, as it directly affects productivity, costs, and overall stability. Most traditional HR approaches depend on methods like exit interviews and surveys, which only provide insights after employees decide to leave. Because of this, organizations often struggle to take preventive actions. In this paper, we present a Workforce Harmony Predictor System, a machine learning-based solution designed to identify employees who are likely to leave the organization. The system uses important features such as satisfaction level, evaluation score, workload, tenure, and salary to make predictions. A Random Forest model is used, as it provides good accuracy while remaining easy to interpret. The system is implemented as a web application using Streamlit, allowing HR managers to perform both single and batch predictions. The model achieved an accuracy of 87% with a strong ROC-AUC score, indicating reliable performance. Overall, the proposed system helps organizations move from reactive to proactive decision-making and supports better employee retention strategies.

Key words: Employee Churn, Machine Learning, HR Analytics, Random Forest, Predictive Modeling, Attrition Prediction, Data Mining

1. INTRODUCTION

In recent years, digital technologies and data-driven approaches have started playing a major role in how organizations operate. One of the key challenges that companies continue to face is employee attrition, which directly affects productivity and increases operational costs [1], [2].

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

Most organizations still rely on traditional HR practices such as exit interviews and employee surveys to understand why employees leave. While these methods provide useful feedback, they are reactive in nature and do not help in predicting future attrition [3], [4]. As a result, companies often fail to take timely actions to retain employees. With the advancement of machine learning, it has become possible to analyze employee data and identify patterns related to churn. These models can learn from historical data and predict which employees are at risk of leaving, allowing organizations to take proactive steps [5], [6].

Several studies have shown that factors such as job satisfaction, workload, salary, and career growth opportunities significantly influence employee retention [7], [8]. By analyzing these factors, predictive systems can provide valuable insights for HR decision-making.

In this paper, we propose a Workforce Harmony Predictor System that uses machine learning techniques to predict employee churn. The system is designed to be simple, efficient, and practical for real-world use, helping organizations improve retention strategies and reduce attrition.

2. LITERATURE REVIEW

Employee churn prediction has been an important area of study in human resource analytics and data mining. Dolatabadi et al. (2017) proposed a churn prediction model using neural networks to identify patterns in employee behavior. Their work focused on improving prediction accuracy by leveraging data mining techniques. The study demonstrated that neural network-based approaches can effectively capture complex relationships in employee data. However, the model required significant computational resources, which could limit its applicability in smaller organizations.

Weeramantry et al. (2017) developed a web-based employee turnover analysis tool designed to assist HR departments in large-scale organizations. Their system provided visualization-based insights into employee attrition trends, helping decision-makers understand workforce patterns. While the tool was useful for analysis and reporting, it lacked predictive capabilities, making it less effective for proactive decision-making.

Alam Syah et al. (2018) conducted a comparative study of various machine learning models for employee churn prediction. Their research highlighted that ensemble methods, particularly Random Forest, achieved better performance compared to other algorithms. The study emphasized the importance of selecting appropriate models to improve prediction accuracy, although it mainly focused on model comparison rather than real-world implementation.

Carraher (2011) explored the relationship between employee satisfaction, benefits, and turnover behavior using statistical approaches. The study revealed that salary and job satisfaction are major factors influencing employee retention. However, the approach was limited to descriptive analysis and did not provide predictive insights.

Yiğit et al. (2017) applied data mining techniques to predict employee churn by analyzing historical employee data. Their study showed that machine learning models can identify patterns associated with attrition. However, the accuracy of the model depended heavily on data quality and preprocessing techniques.

Cheripelli et al. (2020) evaluated different machine learning models for employee churn prediction and found that advanced algorithms can significantly improve prediction performance. Similarly, Sisodia et al. (2017) compared multiple machine learning approaches and concluded that ensemble models provide better reliability and accuracy. Jain et al. (2021) further extended this work by applying machine learning techniques across multiple industries, demonstrating the effectiveness of predictive models in different domains.

Ghosh et al. (2013) studied employee behavior and identified key factors that influence an employee's decision to stay in an organization. Their research emphasized the role of organizational support, career growth, and job satisfaction in reducing employee turnover.

Overall, these studies highlight the growing importance of machine learning in employee churn prediction. While significant progress has been made in improving accuracy, many approaches still lack real-time implementation and practical usability, creating a need for more integrated and user-friendly systems.

A summary of key research works related to employee churn prediction is presented in Table 1.

Table 1: Literature Survey

S.No	Author	Year	Methodology	Result
1	Dolata Badi	2017	Neural Networks	High accuracy
2	Weeramantry	2017	Web-based Tool	Visualization only
3	Alam Syah	2018	ML Comparison	RF best
4	Carraher	2011	Statistical	Salary impact
5	Yigit	2017	Data Mining	Moderate accuracy
6	Cheri Pelli	2020	ML Models	Improved accuracy
7	Sisodia	2017	Comparative ML	Better Prediction
8	Jain	2021	ML Techniques	High performance
9	Ghosh	2013	Behavioral Study	Retention Factors

From Table 1, it can be observed that machine learning-based approaches provide better accuracy compared to traditional statistical methods. Among various techniques, ensemble methods such as Random Forest have shown improved performance in handling complex datasets. However, many existing systems still lack real-time implementation and user-friendly interfaces, which highlight the need for the proposed system.

3. EXISTING SYSTEM

Most existing systems used in organizations for managing employee churn are based on traditional and reactive approaches. These systems primarily focus on methods such as exit interviews, employee surveys, and manual HR data analysis. While these approaches provide useful insights into why employees leave, they do not help in predicting future attrition.

In many organizations, employee data is analyzed using spreadsheets and static reports, which makes the process time-consuming and prone to errors. These systems mainly provide descriptive information such as turnover rates and employee statistics, rather than offering predictive insights. As a result, organizations are unable to identify employees who are at risk of leaving in advance.

Another limitation of current systems is the lack of real-time analytics. Employee data is often analyzed periodically, which delays decision-making and reduces the effectiveness of retention strategies. Additionally, these systems do not integrate modern machine learning techniques, which limits their ability to handle large datasets and complex patterns.

Furthermore, existing HR systems lack interactive and user-friendly interfaces that can assist decision-makers in understanding employee behavior effectively. Without proper visualization and predictive capabilities, HR professionals find it difficult to take proactive measures.

These limitations highlight the need for a more advanced and data-driven system that can provide accurate predictions and real-time insights. The proposed system addresses these challenges by integrating machine learning techniques with an interactive platform, enabling organizations to identify at-risk employees and take timely actions to improve retention.

4. PROPOSED SYSTEM

The proposed system introduces a data-driven approach for predicting employee churn using machine learning techniques. Unlike traditional HR systems that rely on reactive methods, the proposed solution focuses on early identification of employees who are at risk of leaving the organization. This helps HR managers take timely actions and improve retention strategies.

The system is designed to be scalable, efficient, and user-friendly. It integrates data preprocessing, machine learning, and visualization into a single platform. The application allows HR users to input employee data either manually or through batch uploads and receive predictions in real time.

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

4.1 System Architecture

The overall architecture of the proposed system follows a layered approach, ensuring modularity and easy maintenance. The architecture mainly consists of data processing, machine learning, and application layers.

4.1.1 **Data Layer:** This layer handles employee data collected from CSV files or databases. It includes features such as satisfaction level, evaluation score, number of projects, salary, and tenure.

4.1.2 **Processing Layer:** In this stage, data preprocessing is performed. Missing values are handled, categorical variables are encoded, and numerical features are normalized to improve model performance.

4.1.3 **Machine Learning Layer:** The processed data is passed to the trained Random Forest model, which predicts whether an employee is likely to leave or stay. The model also provides a probability score.

4.1.4 **Application Layer:** A Streamlit based interface is used to interact with the system. Users can perform single or batch predictions and view results instantly.

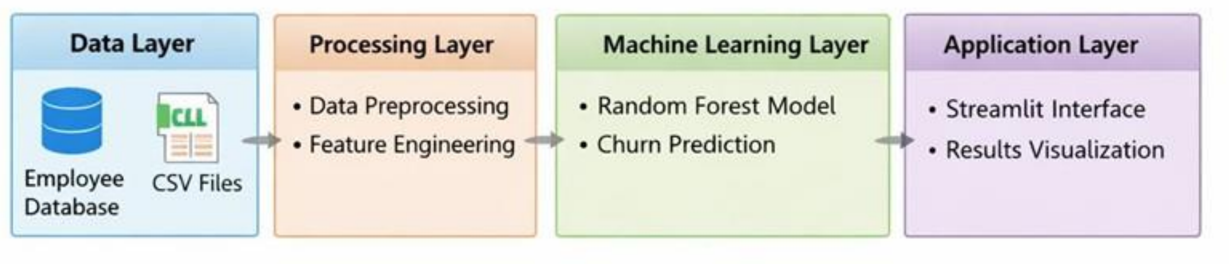

Figure 1: System Architecture of the Proposed System

4.2 System Modules

The proposed system is divided into several functional modules to ensure smooth operation and better organization.

Data Collection Module: Responsible for gathering employee data from various sources such as CSV files.
Data Preprocessing Module: Cleans and transforms data by handling missing values, encoding categorical features, and scaling numerical values.
Machine Learning Module: Implements the Random Forest algorithm for training and prediction.
Prediction Module: Generates churn predictions for both individual employees and multiple records.
Visualization Module: Provides graphical insights such as distributions, correlation heatmaps, and churn trends.

4.3 Dataset Description

The dataset used in this system contains employee-related attributes that influence churn behavior.
These include:
4.3.1 Satisfaction level
4.3.2 Last evaluation score
4.3.3 Number of projects
4.3.4 Average monthly working hours
4.3.5 Time spent in the company
4.3.6 Salary level
4.3.7 Promotion history
These features help the model learn patterns associated with employee attrition.

4.4 Data Preprocessing

Data preprocessing is an important step to ensure the quality and reliability of the model.
The dataset is first cleaned by handling missing values and removing inconsistencies. Categorical variables such as department and salary are converted into numerical form using encoding techniques. Numerical features are then normalized using scaling methods to improve model performance.

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

The processed data is divided into training and testing sets, allowing the model to learn and evaluate its performance effectively.

4.5 Algorithm Used

The proposed system uses the Random Forest algorithm for employee churn prediction. Random Forest is an ensemble learning method that combines multiple decision trees to improve accuracy and reduce overfitting.

The algorithm works by training several decision trees on different subsets of data and combining their outputs to produce a final prediction. It also provides feature importance, which helps identify the key factors influencing employee attrition.
The choice of Random Forest is based on its ability to handle both numerical and categorical data and its strong performance in classification tasks.

4.6 System Workflow

The workflow of the proposed system is designed to provide a smooth and efficient prediction process.

1. The user inputs employee data manually or uploads a dataset.
2. The system preprocesses the data by cleaning and transforming it.
3. The processed data is passed to the machine learning model.
4. The model generates predictions along with probability scores.
5. The results are displayed through the user interface with visual insights.

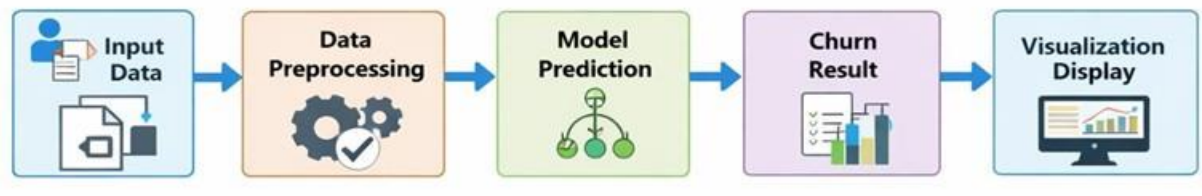

Figure 2: System Workflow of the Proposed System

5. RESULTS AND ANALYSIS

The proposed employee churn prediction system was evaluated based on both model performance and system usability. The results demonstrate that the system can provide accurate predictions along with meaningful visual insights, making it useful for real-world HR applications**.**

5.1 System Interface

The system provides a user-friendly interface that allows HR managers to easily navigate through different functionalities such as data upload, analysis, and prediction. The home page presents an overview of the system along with navigation options.

Figure 3: UI for Home Page

The interface is designed to be simple and intuitive, ensuring that even non-technical users can operate the system without difficulty**.**

5.2 Data Analysis and Visualization

The system includes an Exploratory Data Analysis (EDA) module that helps users understand the dataset before performing predictions. Users can upload employee data and view basic information such as dataset size, column names, and missing values.

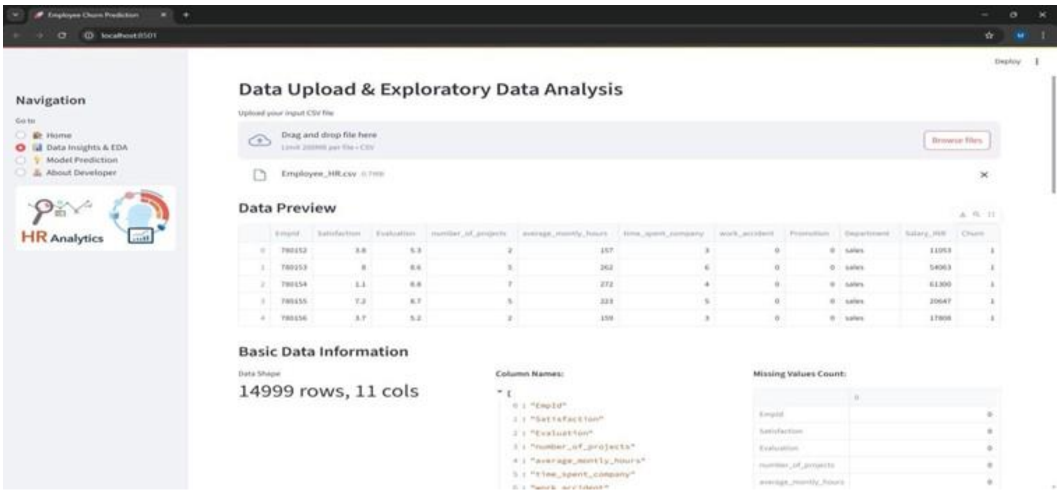

Figure 4: UI for Data Upload & EDA

To better understand employee attrition patterns, the system provides a visualization of churn distribution.

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

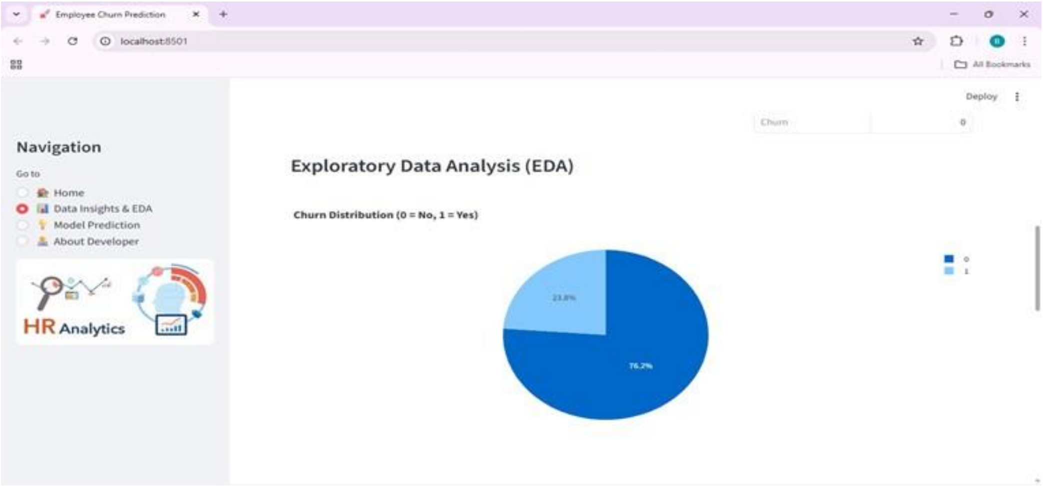

Figure 5: UI for EDA, Churn Distribution

From this visualization, it can be observed that most employees remain in the organization, while a smaller percentage leave. This helps in identifying whether the organization is facing high attrition.
The system also allows visualization of feature distributions such as satisfaction level and working hours, which helps in identifying patterns and outliers.

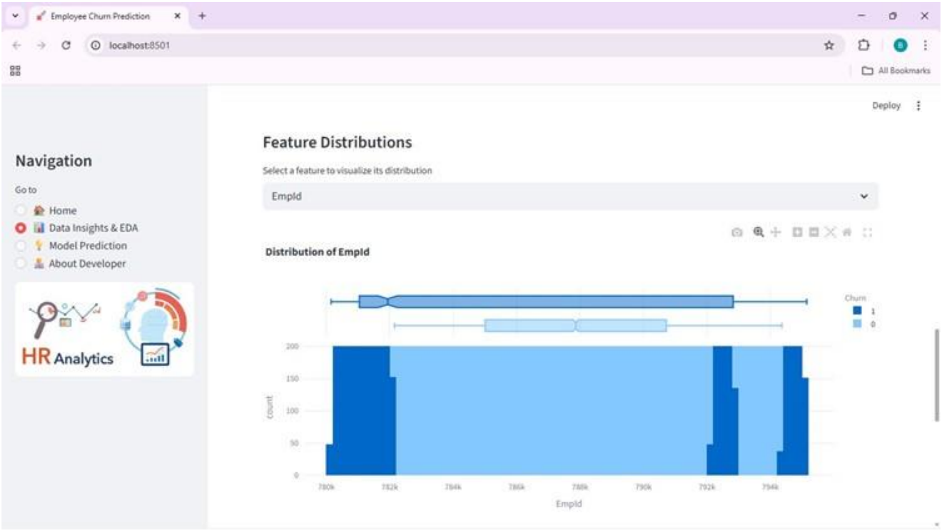

Figure 6: UI for Feature Distribution

In addition, a correlation heatmap is generated to understand the relationship between different features.

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

Figure 7: UI for Correlation Heatmap of Numerical Features

The heatmap shows that satisfaction level has a strong negative correlation with churn, indicating that lower satisfaction leads to higher attrition.

5.3 Prediction Results

The system supports both single employee prediction and batch prediction.

In single prediction mode, the user manually enters employee details, and the system predicts whether the employee is likely to leave along with a probability score.

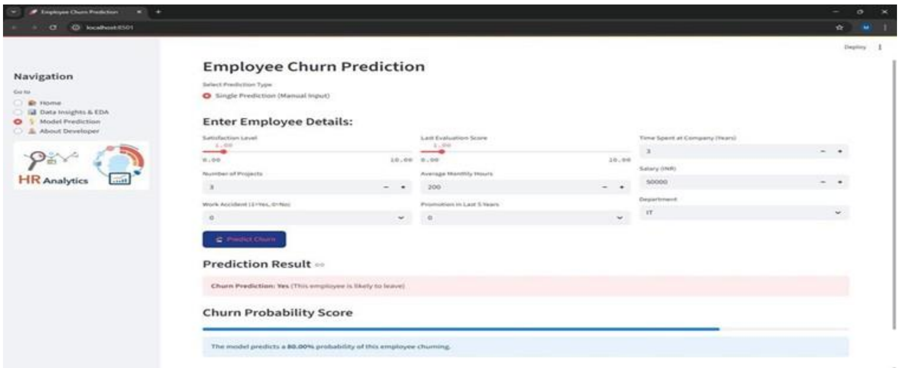

Figure 8: UI for Employee Churn Prediction (Single Prediction)

For large datasets, the system provides batch prediction functionality, where users can upload a CSV file containing multiple employee records.

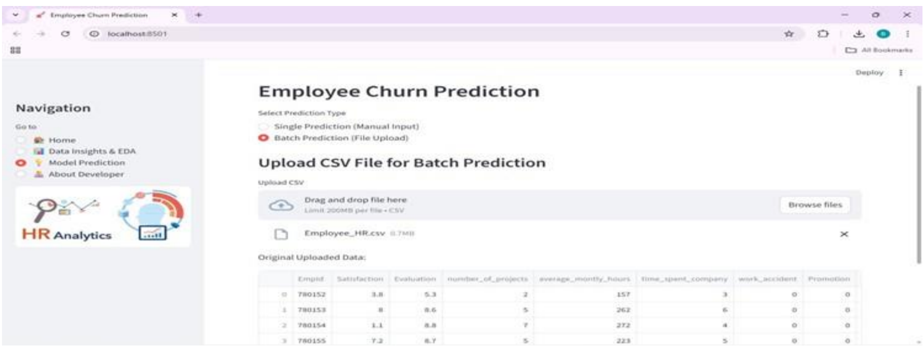

Figure 9: UI for CSV Upload (Batch Prediction)

ISSN: 2582 - 6379
IJISEA Publications
International Journal for Interdisciplinary Sciences and Engineering Applications
IJISEA - An International Peer- Reviewed Journal
2026, Volume 7 Issue 2
www.ijisea.org

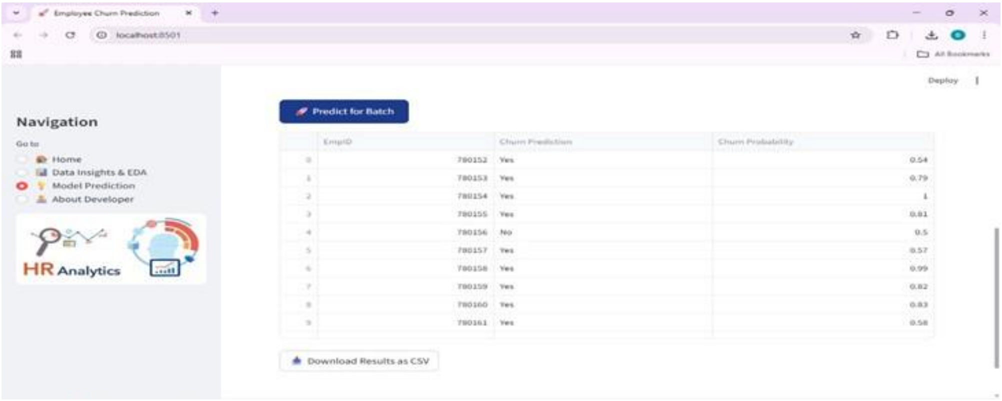

Figure 10: UI for Results with Download Option (Batch Prediction)

The results are displayed in a tabular format, including employee ID, prediction result, and churn probability. Users can also download the results for further analysis.

5.4 Model Performance

The performance of the model was evaluated using standard metrics such as accuracy, precision, recall, and F1-score. The model achieved an accuracy of 87%, indicating strong predictive performance.
The performance metrics of the proposed model are summarized in Table 4.1.

Table 2: Model Performance Metrics

Metric	Value
Accuracy	87%
Precision	84%
Recall	81%
F1-Score	82%
ROC-AUC	0.90

From Table 2, it can be observed that the model performs well across all evaluation metrics. The high ROC-AUC score indicates strong capability in distinguishing between employees who are likely to leave and those who are not.
These metrics confirm that the model performs well and can be reliably used for churn prediction.

5.5 Key Observations

From the analysis, several important insights were identified:
5.5.1 Employees with low satisfaction levels are more likely to leave.
5.5.2 High workload and long working hours contribute to attrition.
5.5.3 Salary and promotion opportunities significantly influence employee retention.
These insights can help HR managers design better retention strategies.

5.6 System Performance

The system also performs efficiently in real-time scenarios. Predictions are generated quickly, and batch processing allows multiple employee records to be analyzed within a short time.
The application remains responsive even when handling large datasets, making it suitable for practical deployment in organizations.

6. DISCUSSIONS

The results obtained from the proposed system clearly indicate that machine learning techniques can play an important role in solving real-world HR challenges. By using predictive models, organizations can shift from traditional reactive approaches to more proactive decision-making, which helps in identifying employees who are at risk of leaving at an early stage.

The system not only provides accurate predictions but also offers useful insights through data visualization, making it easier for HR professionals to understand employee behavior. This improves the overall decision-making process and supports the development of effective retention strategies.

However, the performance of the system largely depends on the quality of the data used. If the dataset contains missing values, inconsistencies, or incorrect information, the predictions may not be reliable. Therefore, proper data collection, preprocessing, and validation are essential to ensure accurate results.

In addition, while the model performs well for the given dataset, its effectiveness may vary across different organizations depending on their workforce structure and data availability. Despite these limitations, the proposed system provides a practical and efficient solution for employee churn prediction and can be further improved with more advanced techniques.

7. CONCLUSION

In this paper, a machine learning-based system for predicting employee churn was presented. The system integrates data preprocessing, predictive modeling, and a user-friendly interface to provide meaningful insights into employee behavior.

The results demonstrate that the proposed model achieves good performance and can effectively identify employees who are likely to leave the organization. By providing early predictions, the system enables organizations to take proactive measures, thereby reducing attrition and improving workforce stability.

Overall, the proposed system offers a simple, efficient, and scalable solution for employee churn prediction. It highlights the importance of using data-driven approaches in HR management and shows how machine learning can be applied to improve organizational decision-making.

8. FUTURE WORK

The proposed system can be further improved by incorporating more advanced machine learning techniques such as XGBoost, LightGBM, and deep learning models, which may enhance prediction accuracy and handle complex patterns more effectively.

Another important improvement would be the integration of the system with real-time HR management systems, allowing continuous monitoring of employee data and enabling instant predictions. This would help organizations respond more quickly to potential attrition risks.

In addition, deploying the system on cloud platforms can improve scalability and accessibility, making it suitable for large organizations with a high volume of employee data.

The system can also be enhanced by incorporating explainable AI techniques, which would help HR professionals understand the reasoning behind predictions and increase trust in the model. Furthermore, future work can focus on recommending personalized retention strategies based on employee risk levels, such as training programs, promotions, or workload adjustments.

9. REFERENCES

[1] Dolatabadi, S.H.; Keynia, F. Designing of customer and employee churn prediction model based on data mining method and neural predictor. In Proceedings of the 2nd International Conference on Computer and Communication Systems (ICCCS), Kraków, Poland, 11–14 July 2017; pp. 74–77.

[2] Weeramantry, T.T.; Thilakumara, C.N.; Wijesiri, K.N.A.C.; Fernando, N.I.; Thelijjagoda, S.; Gamage, A. ARROW: A web-based employee turnover analysis tool for effective human resource management in large-scale organizations. In Proceedings of the National Information Technology Conference (NITC), Colombo, Sri Lanka, 14– 15 September 2017; pp. 136–140.

[3] Seth Unga, S.; Perera, I. Impact of Performance Rewards on Employee Turnover in Sri Lankan IT Industry. In Proceedings of the Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka, 30 May–1 June 2018; pp. 114–119.

[4] Wei, G.U.O.; Tai, L.I. An empirical study on organizational commitment and turnover of it industry. In Proceedings of the International Conference on E-Business and E- Government, Guangzhou, China, 7–9 May 2010; pp. 904–906.
[5] David, S.; Kaushik, S.; Verma, H.; Sharma, S. Attrition in “IT” Sector. Int. J. Core Eng. Manag. IJCEM 2015, 2, 74–92.
[6] Mura, L.; Zsigmond, T.; Machová, R. The effects of emotional intelligence and ethics of SME employees on knowledge sharing in Central-European countries. Oeconomia Copernic. 2021, 12, 907–934. [CrossRef]
[7] Szeiner, Z.; Kovács, Á.; Zsigmond, T.; Mura, L.; Sanders, E.; Poor, J. An empirical study of consulting in a transitional economy in the Central European region during COVID-19. J. East. Eur. Cent. Asian Res. (JEECAR) 2022, 9, 471–485. [CrossRef]
[8] Alam Syah, A.; Salma, N. A Comparative Study of Employee Churn Prediction Model. In Proceedings of the 4th International Conference on Science and Technology (ICST), Yogyakarta, Indonesia, 7–8 August 2018; pp. 1–4.
[9] Carraher, S.M. Turnover prediction using attitudes towards benefits, pay, and pay satisfaction among employees and entrepreneurs in Estonia, Latvia, and Lithuania. Balt. J. Manag. 2011, 6, 25–52. [CrossRef]
[10] Yi ˘git, ˙I.O.; Shourabizadeh, H. An approach for predicting employee churn by using data mining. In Proceedings of the International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, 16–17 September 2017; pp. 1–4.
[11] Cheripelli, R.; Ajitha, P.V. Evaluation of Machine Learning Models for Employee Churn Prediction. Solid State Technol. 2020, 63, 2482–2487.
[12] Sisodia, D.S.; Vishwakarma, S.; Pujahari, A. Evaluation of machine learning models for employee churn prediction. In Proceedings of the International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 23–24 November 2017; pp. 1016– 1020.
[13] Jain, H.; Yadav, G.; Manoov, R. Churn Prediction and Retention in Banking, Telecom and IT Sectors Using Machine Learning Techniques. In Advances in Machine Learning and Computational Intelligence, 1st ed.; Springer: Singapore, 2021; Volume 1, pp. 137–156.
[14] Yeom, S.; Giacomelli, I.; Fredrikson, M.; Jha, S. Privacy risk in machine learning: Analyzing the connection to overfitting. In Proceedings of the 31st Computer Security Foundations Symposium (CSF), Oxford, UK, 9–12 July 2018; pp. 268–282.
[15] Ghosh, P.; Satyawadi, R.; Joshi, J.P.; Shadman, M. Who stays with you? Factors predict employees’ intention to stay. Int. J. Organ. Anal. 2013, 21, 288–312.