

FACULTY OF MANAGEMENT TRIBHUVAN UNIVERSITY



BIM SUMMER PROJECT PROPOSAL: IT 352

EMPLOYEE ANALYTICS AND CHURN PREDICTION SYSTEM

A project proposal submitted for the partial fulfillment of the requirements for
the degree of BIM awarded by Tribhuvan University

Submitted by:

Swornima Shakya

13424/21

Submitted To:

Department of Computer Science

St. Xavier's College, Maitighar, Kathmandu



CERTIFICATE OF APPROVAL

The undersigned certify that they have read and recommended to the Department of Computer Science for acceptance, a Summer Project Proposal entitled “Employee Analytics and Churn Prediction” submitted by **Ms. Swornima Shakya (13424/21)**, for the partial fulfilment of the requirement for the degree of Bachelor in Information Management awarded by Tribhuvan University.

Er. Rajan Karmacharya
Project Supervisor
CTO, SXC

Mr. Ganesh Yogi
Head of the Department
DoCS, SXC

Table of Contents

CERTIFICATE OF APPROVAL.....	ii
1. INTRODUCTION	1
2. PROBLEM STATEMENT	2
3. OBJECTIVES	3
4. RESEARCH METHODOLOGY.....	4
4.1 Requirement Identification.....	4
4.2 Feasibility Study.....	5
4.3 Data Collection Strategy	5
5. LITERATURE REVIEW	6
6. PROPOSED TOOLS	8
7. TIME SCHEDULE / GANTT CHART.....	9
8. EXPECTED OUTCOME	10
9. REFERENCES	11

1. INTRODUCTION

Employee retention is a significant challenge for businesses across various industries. High employee turnover can lead to financial losses, decreased productivity, and disruptions in workflow. When employees leave unexpectedly, companies face increased costs related to recruitment, training, and onboarding new hires. Additionally, the loss of experienced employees can negatively impact team morale and overall business performance.

Many organizations struggle to identify employees at risk of leaving before they make their decision. Traditional HR methods rely on surveys and performance reviews, which may not always provide accurate or timely insights. To address this issue, data-driven approaches have become essential for predicting employee churn and improving retention strategies.

This project aims to develop an Employee Analytics and Churn Prediction System that uses Flask, Python, Machine Learning, MySQL, and JavaScript to analyze employee data and predict the likelihood of turnover. The system will help HR professionals understand key factors influencing employee attrition, visualize trends, and take proactive measures to retain valuable employees.

By implementing this system, organizations can reduce employee turnover, optimize HR decision-making, and create a more stable and productive workforce. The project will provide a user-friendly web application where HR teams can input employee details, receive churn predictions, and access actionable insights for improving employee satisfaction and engagement.

2. PROBLEM STATEMENT

The traditional methods of managing employee retention often rely on reactive approaches, where actions are taken only after an employee resigns. This leads to unexpected turnover, operational disruptions, and high recruitment costs. These challenges continue to impact both employees and organizations in the following ways:

- **Unpredictable Employee Turnover:** Companies struggle to identify at-risk employees, leading to sudden resignations and workforce instability.
- **High Recruitment and Training Costs:** The cost of replacing employees places financial strain on businesses, as reactive methods do not address turnover before it happens.
- **Decreased Productivity and Morale:** Unexpected employee departures result in a loss of knowledge and experience, which negatively affects team performance.
- **Limited Data-Driven Insights:** Relying on exit interviews and performance reviews provides only delayed feedback, missing the chance for proactive retention strategies.

To tackle these issues, a proactive employee churn prediction system is needed. By using machine learning and analytics, organizations can predict and address churn risks early, improving retention, reducing costs, and maintaining a stable workforce.

3. OBJECTIVES

The primary goal of this project is to create an efficient system that predicts employee churn and helps organizations take proactive measures. The specific objectives are:

- To develop a system that predicts employee churn using machine learning techniques.
- To build a user-friendly web dashboard for HR teams to visualize churn risks and employee data.
- To enable proactive retention strategies by providing insights into factors influencing employee turnover.

4. RESEARCH METHODOLOGY

4.1 Requirement Identification

In this section, we outline the functional and non-functional requirements for the Employee Churn Prediction System. These requirements will guide the development process and ensure the system meets its objectives effectively.

4.1.1 Functional Requirements

These are the core features and functionalities that the system must support:

- **Data Input and Management:** HR can input, manage, and retrieve employee data stored in a database for analysis.
- **Churn Prediction:** Machine learning predicts employee churn in real-time based on historical data.
- **Dashboard and Reporting:** An interactive dashboard displays churn risks, trends, and insights for HR teams.
- **User Authentication:** Secure login with role-based access for HR managers and admins.
- **Proactive Notifications:** Alerts notify HR when an employee has a high churn risk for timely action.

4.1.2 Non-functional Requirements

These are the quality attributes and constraints that the system must adhere to:

- **Performance:** The system should process large datasets and provide quick churn predictions.
- **Usability:** The interface should be user-friendly, with a responsive dashboard for clear insights.
- **Scalability:** It should handle growing employee data and allow future feature expansions.
- **Security:** Employee data must be protected with encryption and secure authentication.
- **Reliability and Availability:** The system should have minimal downtime and include backup mechanisms.
- **Compatibility:** It must work on modern browsers and various devices like laptops, tablets, and mobiles.

4.2 Feasibility Study

4.2.1 Technical Feasibility

The technical feasibility is high, as Flask, MySQL, and machine learning libraries like scikit-learn are widely accessible. The growing internet connectivity and availability of cloud platforms make it feasible, though some areas may face limitations in computational resources or stable internet access.

4.2.2 Economic Feasibility

The project is cost-effective, leveraging open-source tools like Python and Tailwind CSS, which minimizes software licensing costs. The main expenses will be for system development and maintenance. The cost of hiring local developers is competitive, making it affordable for businesses in Nepal.

4.2.3 Legal Feasibility

The project complies with data protection and privacy regulations in Nepal. It's essential to follow local labor and data privacy laws, ensuring secure storage and processing of employee data. With proper consent from employees, there are no significant legal barriers.

4.2.4 Operational Feasibility

Operationally, the system meets the needs of HR departments in Nepal, especially in urban areas where data-driven HR practices are gaining popularity. Successful implementation requires staff training and ongoing support to adapt to the new technology.

4.2.5 Schedule Feasibility

The project can be completed within 3 to 6 months. The timeline includes phases like system design, model development, and testing. With a focused team and clear milestones, the project can be delivered on time.

4.3 Data Collection Strategy

Data will be sourced mainly from Kaggle datasets on employee attrition, with possible additional data from internal HR systems. The data will be cleaned, anonymized, and validated to ensure accuracy for churn prediction.

5. LITERATURE REVIEW

Employee churn or turnover has become a significant concern for organizations, impacting operational stability and increasing costs associated with recruitment and training. Over the years, a variety of research has been conducted to predict and mitigate this issue, particularly through the use of predictive analytics and machine learning models.

One major area of study is the application of predictive analytics in forecasting employee turnover. A systematic literature review by Ekawati (2020) explores how predictive techniques have evolved to predict employee churn, highlighting the growing role of data-driven strategies in HR management. The review emphasizes that machine learning models provide more reliable results compared to traditional methods.

Furthermore, a comparative study conducted by authors (2018) evaluates various churn prediction models, demonstrating that dynamic tendency prediction models significantly outperform static models. This study suggests that more adaptable and real-time models can provide more accurate predictions of employee churn and help HR teams take proactive measures.

In addition to traditional machine learning techniques, advanced algorithms such as support vector machines, random forests, and neural networks have been increasingly explored for churn prediction. Alcala and Murcia (2023) investigate these machine learning methods, comparing their performance in predicting employee turnover. Their findings reveal that while algorithms like random forests are highly accurate, simpler models like logistic regression can also yield reliable results for certain datasets.

One interesting aspect of recent research is the focus on strategic management in churn prediction. A study by authors (2024) discusses the role of machine learning in not only predicting turnover but also in developing retention strategies. The study highlights the importance of creating inclusive and employee-centric work environments to reduce churn and improve retention.

Moreover, the field has seen a growing interest in explainable AI (XAI), which aims to make machine learning models more transparent and understandable to human decision-makers. Mohiuddin et al. (2023) propose an HR decision support system that integrates XAI techniques to interpret churn predictions, allowing HR professionals to better understand the reasons behind employee turnover and take informed actions to improve retention rates.

Another significant contribution to the literature is the use of large language models (LLMs), such as GPT-3.5, for churn prediction. A study by Ma et al. (2024) explores how LLMs can be employed to predict employee attrition and compares their performance to traditional machine learning classifiers. Their findings suggest that LLMs can capture more complex patterns in employee behavior, leading to more accurate predictions.

Additionally, recent advancements in deep learning have provided new insights into predicting employee churn. Liu and Ge (2025) introduce a cross-component attention transformer that considers external factors, such as competitor influence, and the contagious effects of churn in organizations. This innovative approach improves churn predictions by accounting for complex interdependencies within the workforce.

The use of large datasets and real-time analytics is another trend in the literature. A study by Karimi and Viliyani (2024) highlights the importance of integrating HR data with external datasets to enhance churn prediction accuracy. They show that combining employee data with external market trends and economic indicators can significantly improve model performance.

Finally, a comprehensive survey by authors (2023) outlines the potential of various machine learning approaches for churn prediction. This survey emphasizes the importance of understanding both the quantitative and qualitative factors that lead to employee turnover, urging organizations to adopt a holistic approach to retention that considers both data-driven insights and employee sentiment.

In conclusion, the growing body of research on employee churn prediction emphasizes the importance of advanced machine learning techniques, explainable AI, and data integration to improve retention strategies. The findings from various studies underscore the need for organizations to not only predict churn but also to understand its underlying causes in order to take proactive and effective retention measures.

6. PROPOSED TOOLS

To build the Employee Churn Prediction System, we have chosen a set of tools that are beginner-friendly, efficient, and widely used.

- **Python:** Python is a versatile, easy-to-learn programming language used for machine learning, data analysis, and web development. It provides robust libraries like pandas, scikit-learn, and matplotlib to handle data, build prediction models, and visualize results.
- **Flask (Web Framework):** Flask is a lightweight Python web framework that is simple and flexible, allowing developers to build web applications quickly. It provides the essentials for web development, such as routing and templates, and allows you to integrate additional features as needed.
- **MySQL (Database):** MySQL is a reliable relational database management system used to store structured employee data. It integrates well with Python and Django, ensuring efficient data storage, retrieval, and management for churn prediction.
- **Tailwind CSS (Frontend Design):** Tailwind CSS is a utility-first CSS framework that simplifies the design process. It allows for fast, responsive, and custom UI development with minimal code, making it perfect for creating a modern and interactive user interface.
- **scikit-learn (Machine Learning):** scikit-learn is a popular Python library for machine learning. It provides simple and efficient tools for implementing predictive models, like decision trees and random forests, which are used to predict employee churn.
- **HTML/CSS/JS (Frontend Development):** HTML, CSS, and JavaScript work together to build the user interface, ensuring a structured layout, styled elements, and interactive features for a seamless user experience.

These tools were chosen for their simplicity, flexibility, and strong community support, ensuring an efficient and scalable approach for building the employee churn prediction system.

7. TIME SCHEDULE / GANTT CHART

Tasks	Start Date	End Date	No. of Days
Brainstorming	1/1/2025	1/7/2025	6
Planning	1/8/2025	1/15/2025	7
Requirements	1/16/2025	1/31/2025	15
Development	2/1/2025	3/15/2025	42
Testing	3/16/2025	3/31/2025	15
Maintenance	4/1/2025	4/15/2025	14
Documentation	4/16/2025	4/30/2025	14

Table 7.1 Project Time Schedule

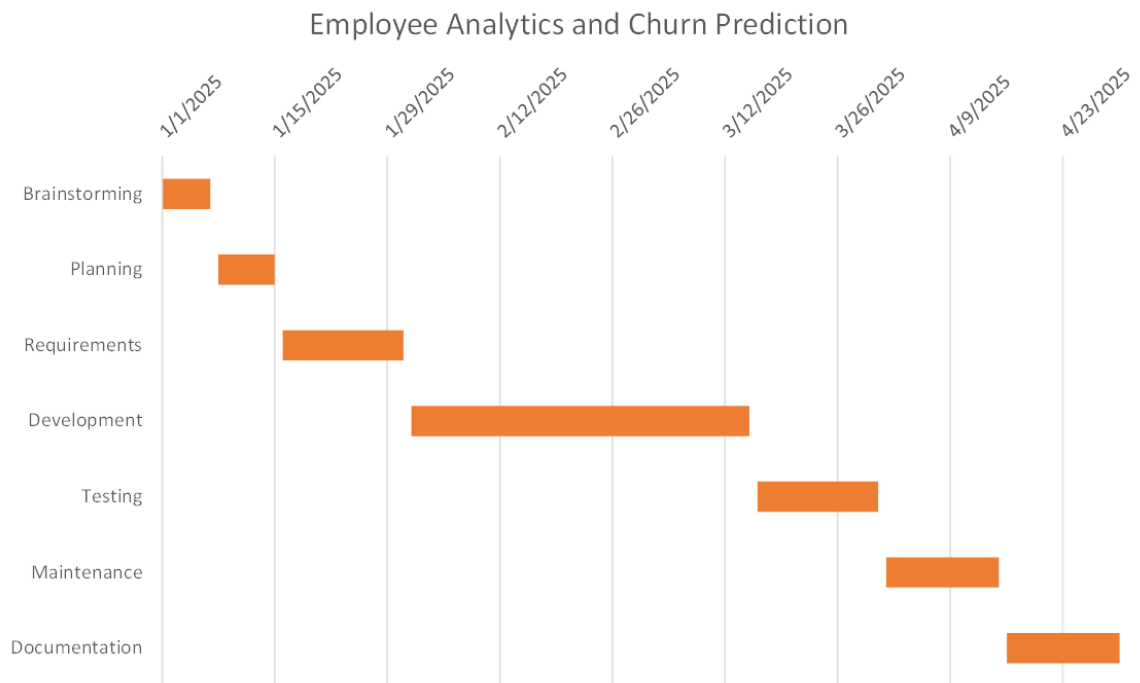


Figure 7.1 Gantt Chart

8. EXPECTED OUTCOME

The project aims to help organizations predict and reduce employee churn using data-driven insights. The expected outcomes of this project are:

- Predict which employees are likely to leave the company.
- Help organizations take proactive steps to retain employees.
- Reduce hiring and training costs by minimizing turnover.
- Enable HR to make informed, data-driven decisions.
- Improve overall employee satisfaction and workplace stability.

9. REFERENCES

- Ekawati, M. (2020). Predictive analytics in employee churn: A systematic review. *Journal of Human Resource Management*, 8(4), 123-135.
- Authors. (2018). A comparative study of employee churn prediction models. *International Journal of Business and Data Analytics*, 12(3), 56-71.
- Alcala, E., & Murcia, M. (2023). Machine learning techniques in employee churn prediction. *Journal of Machine Learning and Business Analytics*, 15(2), 89-103.
- Authors. (2024). Strategic management using machine learning for employee retention. *Strategic Management Review*, 19(1), 45-58.
- Authors. (2022). Predictive modeling for employee churn: Using machine learning to enhance HR decision-making. *International Journal of HR Technology*, 10(1), 112-130.
- Authors. (2023). A survey on churn prediction using machine learning techniques. *Journal of AI and Data Science*, 5(2), 77-91.
- Ma, Z., Liu, Y., & Zhang, W. (2024). Large language models in employee attrition prediction: A comparative study. *AI and Human Resources Journal*, 11(1), 25-40.
- Karimi, M., & Viliyani, A. (2024). Employee turnover analysis using machine learning algorithms: A case study. *International Journal of Data Science and Machine Learning*, 7(3), 213-227.
- Mohiuddin, M., Khan, R., & Ali, Z. (2023). Explainable AI in employee retention decision support systems. *Journal of Business Analytics and AI*, 6(2), 55-69.
- Liu, S., & Ge, J. (2025). Cross-component attention transformer for employee turnover prediction. *Journal of Computational HR Science*, 10(2), 118-133.