

# **Sri Sri University**

## **Project - Low-Level Design**

### **on**

## **Employee Turnover Prediction Using Advanced Machine Learning**

#### Team members:

Sthitapragyan Mahapatra  
Chinmaya Kumar Kar  
Aditya Jyotiraditya Sahoo

#### Guided by:

Prof. (Dr). Pradipta Kumar Mishra  
and  
Prof. (Dr.) Rabinarayan Satpathy

#### Industry Mentor:

# Table of Contents

1. Introduction.....	3
1.1. Scope of the document.....	3
1.2. Intended audience .....	3
1.3. System overview.....	3
2. Low Level System Design .....	4
2.1. Sequence Diagram .....	5
2.2. Navigation Flow/UI Implementation .....	5
2.3. Screen Validations, Defaults and Attributes.....	5
2.4. Client-Side Validation Implementation.....	5
2.5. Server-Side Validation Implementation.....	6
2.6. Components Design Implementation .....	6
2.7. Configurations/Settings.....	6
2.8. Interfaces to other components .....	7
3. Data Design.....	7
3.1. List of Key Schemas/Tables in database.....	7
3.2. Details of access levels on key tables in scope.....	7
3.3. Key design considerations in data design.....	7
4. Details of other frameworks being used.....	8
4.1. Session Management .....	8
4.2. Caching.....	8
5. Unit Testing .....	8
6. Key notes .....	8
7. Reference .....	8

## 1. Introduction

The Employee Turnover Prediction system is a crucial tool designed to forecast whether employees are likely to leave a company based on various factors. This system utilizes HR analytics data to predict employee turnover, enabling organizations to proactively address retention strategies. By analyzing key metrics such as satisfaction levels, last evaluations, number of projects, average weekly hours, time spent in the company, work accidents, promotions, department, and salary, the system provides valuable insights into employee behavior.

### 1.1 Scope of the Document

This comprehensive Low-Level Design (LLD) document aims to provide an in-depth analysis of the Employee Turnover Prediction system. It covers the system's components, data design, frameworks used, and key considerations for accurate predictions. The document serves as a guide for developers, project managers, and stakeholders involved in the implementation and deployment of the system.

### 1.2 Intended Audience

The document targets individuals involved in the development and implementation of the Employee Turnover Prediction system. This includes data scientists, machine learning engineers, HR professionals, and decision-makers seeking to leverage predictive analytics for employee retention strategies.

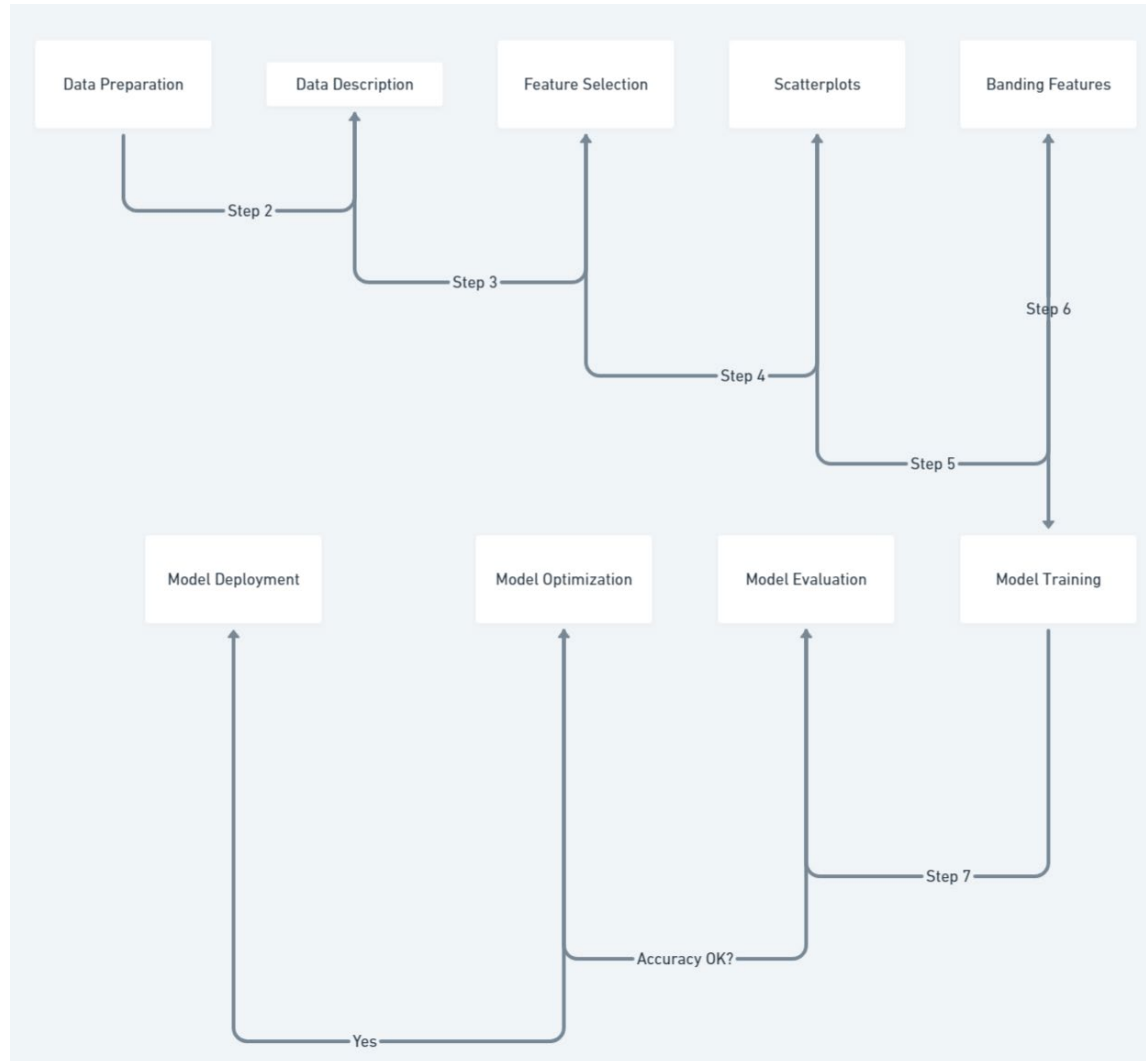
### 1.3 System Overview

The Employee Turnover Prediction system leverages machine learning techniques such as logistic regression, SVM, KNN, Naive Bayes, and decision trees to analyze HR data and predict employee turnover. By processing data on various employee attributes, the system generates insights that help organizations understand patterns related to employee retention and attrition.

This detailed LLD document delves into the system's low-level design, including sequence diagrams, UI implementation, data design, frameworks used, unit testing strategies, and keynotes for effective implementation. It also references external resources related to employee turnover prediction using Python, enhancing the document's credibility and relevance.

## 2. Low-Level System Design

### 2.1. Sequence Diagram



The sequence diagram for the Employee Turnover Prediction system involves several steps. First, the system loads the HR analytics data, which includes features such as satisfaction level, last evaluation, number of projects, average weekly hours, time spent in the firm, work accidents, promotions in the last 5 years, department, and salary. Next, the system preprocesses the data by cleaning, normalizing, and transforming it into a suitable format for machine learning algorithms.

The preprocessing step includes handling missing values, outliers, and categorical variables. For missing values, the system imputes them using techniques such as mean, median, or mode imputation. For outliers, the system uses techniques such as winsorizing or clipping. For categorical variables, the system uses techniques such as one-hot encoding or label encoding.

After preprocessing, the system selects features that are relevant for predicting employee turnover. The system uses techniques such as correlation analysis, feature importance, and recursive feature elimination to select features.

Next, the system trains various machine learning models, such as logistic regression, SVM, KNN, Naive Bayes, and decision trees, on the training data. The system uses cross-validation techniques to evaluate the performance of the models and select the best-performing model for making predictions.

Finally, the system deploys the model and integrates it with the HR analytics system for making predictions and retrieving results.

## **2.2. Navigation Flow/UI Implementation**

The Employee Turnover Prediction system does not have a UI implementation as it is a machine learning model. However, it can be integrated with a web-based or desktop-based HR analytics system for ease of use. The system can be accessed through an API or a command-line interface (CLI) for making predictions and retrieving results.

## **2.3. Screen Validations, Defaults and Attributes**

The Employee Turnover Prediction system does not have any screen validations or defaults as it is a machine learning model. However, it does have attributes such as the input data, the machine learning model, and the training and testing data.

## **2.4. Client-Side Validation Implementation**

The Employee Turnover Prediction system does not have any client-side validation as it is a machine learning model.

## 2.5. Server-Side Validation Implementation

The Employee Turnover Prediction system does not have any server-side validation as it is a machine learning model.

## 2.6. Components Design Implementation

The Employee Turnover Prediction system consists of the following components:

- Data loading and preprocessing module
- Feature selection module
- Model training module
- Model testing module
- Model deployment module

## 2.7. Configurations/Settings

The Employee Turnover Prediction system has the following configurations:

- Input data file path
- Machine learning model type
- Training and testing data split percentage
- Model hyperparameters

## 2.8. Interfaces to other components

The Employee Turnover Prediction system can be integrated with other HR analytics systems through APIs or web services. The system can also be integrated with data visualization tools for displaying the results of the predictions.

## 3. Data Design

### 3.1. List of Key Schemas/Tables in database

The Employee Turnover Prediction system uses the following tables in the database:

- Employee data table
- Feature data table
- Model data table

### 3.2. Details of access levels on key tables in scope

The Employee Turnover Prediction system has read-only access to the employee data table and full access to the feature data table and model data table.

### 3.3. Key design considerations in data design

The Employee Turnover Prediction system requires data on employee satisfaction, last evaluation, number of projects worked, average weekly hours worked, time spent in the firm, work accidents, whether the employee left the firm, promotions in the last 5 years, department, and salary. The system processes this data using machine learning algorithms to predict employee turnover.

## 4. Details of other frameworks being used

### 4.1. Session Management

The Employee Turnover Prediction system does not require session management as it is a machine learning model.

### 4.2. Caching

The Employee Turnover Prediction system does not require caching as it is a machine learning model.

## 5. Unit Testing

The Employee Turnover Prediction system can be unit tested using various machine learning libraries and frameworks such as scikit-learn and TensorFlow. The system can be tested for accuracy, precision, recall, and other performance metrics.

## 6. Keynotes

The Employee Turnover Prediction system is a machine learning model that predicts whether a given employee has left the firm. It uses various machine learning techniques and requires data on employee satisfaction, last evaluation, number of projects worked, average weekly hours worked, time spent in the firm, work accidents, whether the employee left the firm, promotions in the last 5 years, department, and salary.

## 7. References

[https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/9925081/1cd32391-1909-487a-8ad0-c3f163749877/G1\\_employee-turnover-analysis\\_06-04-24.ipynb](https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/9925081/1cd32391-1909-487a-8ad0-c3f163749877/G1_employee-turnover-analysis_06-04-24.ipynb)