Low Level Document (LLD)

Employee Attrition Prediction

Version number: 1.0

Last date of revision: 5 June 2023

**DECLARATION**

We declare that this written submission represents our ideas. it is our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources.

We also declare that we have adhered to all principles of academic honesty

and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission.

We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when

needed.

**Revision History**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Reviewer** | **Approver** | **Comments** |
| 0.1 | 01-05-2023 | Praveen Kumar | Thokchom Mobisana |  | Draft version |
| 0.2 | 02-05-2023 | Praveen Kumar | Chirag Sharma |  | Suggested some selections like key notes, screen validations and attributes to be added |
| 0.3 | 03-05-2023 | Thokchom Mobisana | Praveen Kumar |  | Suggested document format related comments like correction of version, adding one sections for open issues etc |
| 0.4 | 04-05-2023 | Thokchom Mobisana | Praveen Kumar |  | Suggested some changes like correct sequence diagram, changes in data design sections etc |
| 1.0 | 05-05-2023 | Praveen Kumar | Thokchom Mobisana |  | Baseline version |

Table of Contents

1. Introduction

1.1 Scope of the document

1.2 Intended Audience

1.3 System Overview

1. Project Briefing
2. Problem Statement
3. Problem Solution
4. Objective of Project
5. Scope of Project
6. Block Diagram
7. Requirements Gathering
8. Analysis
9. System Architecture

10.1 Data Preprocessing

10.2 Feature Selection

10.3 Model Development

10.4 Model Evaluation

10.5 Deployment

11.Screenshoots

**1. Introduction**

1.1 Scope of the Document:

This low-level document provides a detailed description of the design and implementation of the employee attrition prediction project using the random forest classifier algorithm.

1.2 Intended Audience:

This document is intended for developers, data scientists, and stakeholders involved in the project.

1.3 System Overview:

The system aims to predict employee attrition by analyzing various features related to job satisfaction, distance from work, compensation, and performance, among others.

**2. Project Briefing**

The objective of the project is to develop an employee attrition prediction system using machine learning techniques. The system aims to assist companies in predicting and identifying employees who are likely to leave the organization. By analyzing various factors such as job satisfaction, distance from work, compensation, and performance, the system will provide insights that can help reduce attrition rates and improve employee retention.

Employee attrition is a significant concern for companies, as high attrition rates can lead to productivity losses, increased recruitment and training costs, and a negative impact on company reputation. By accurately predicting which employees are at risk of leaving, companies can proactively implement retention strategies, offer incentives, or address any underlying issues to improve employee satisfaction and increase retention rates.

The project will utilize the Random Forest Classifier algorithm, a powerful and widely used machine learning algorithm for classification tasks. The algorithm will be trained on historical employee data, which will serve as the basis for predicting attrition probabilities. The system will involve various stages, including data preprocessing, feature selection, model training, model evaluation, and result visualization.

The system will provide an intuitive and user-friendly interface for visualizing the prediction results. This will enable companies to easily interpret and utilize the predictions to make informed decisions regarding employee retention. The system will also be designed to be scalable and adaptable, allowing for easy integration with existing HR systems or databases.

Overall, the employee attrition prediction system aims to assist companies in reducing attrition rates and improving employee satisfaction and retention. By accurately identifying employees at risk of leaving, companies can take proactive measures to address issues, improve employee engagement, and foster a positive work environment.

**3. Problem Statement**

The problem is to predict the attrition of an employee for the company based on the given dataset. The high attrition rate may indicate issues with the company's culture, pay scale, or other factors affecting employee satisfaction and retention. The goal is to develop a machine learning model that can accurately predict employee attrition, assisting the company in reducing its attrition rate and improving employee retention

**4. Problem Solution**

The solution involves using the random forest classifier algorithm to predict employee attrition. Random forest is a powerful ensemble learning method that combines multiple decision trees to make predictions. It is well-suited for classification problems, such as determining whether an employee will leave or stay with the company.

**5. Objective of the Project**

The objectives of the project are as follows:

Develop a machine learning model for predicting employee attrition.

Provide insights into the factors influencing attrition.

Assist the company in reducing attrition rates and improving employee retention.

**6. Scope of the Project**

The scope of the project includes:

Analyzing the provided dataset and performing data preprocessing tasks.

Selecting relevant features for the prediction model.

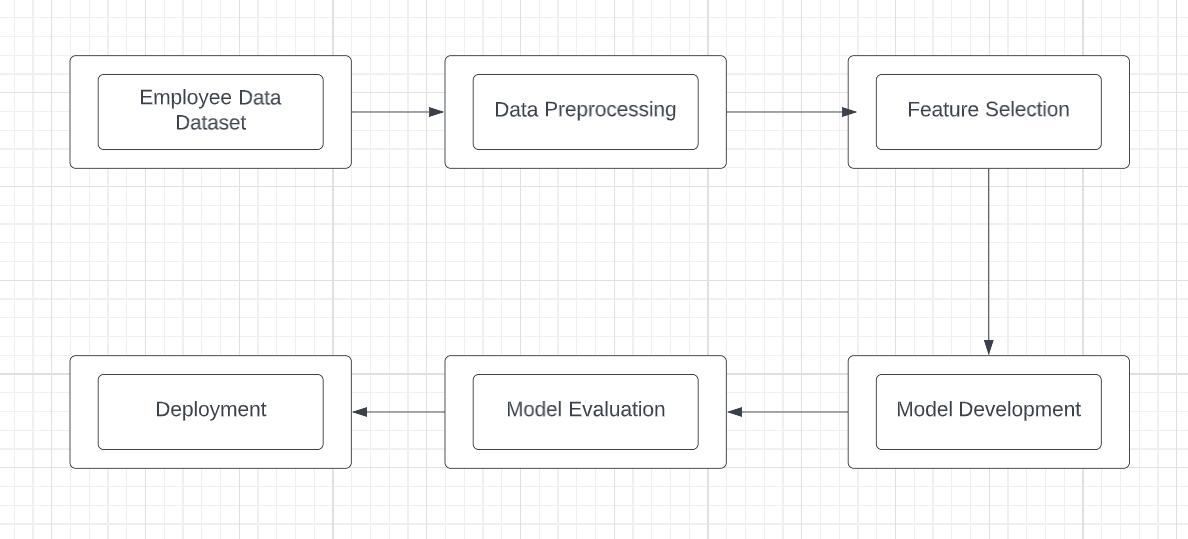
Implementing the random forest classifier algorithm for training and prediction.

Evaluating the model's performance and fine-tuning the parameters.

Deploying the trained model for making predictions on new, unseen data.

**7.Block Diagram**

The block diagram illustrates the flow of the employee attrition prediction system. Here's a brief description of each component:



**8. Requirements Gathering**

The requirements gathering process involved:

Gathering the dataset containing employee-related information.

Understanding the business goals and objectives related to employee attrition.

Identifying the desired features and metrics for evaluating the model's performance.

**9. Analysis**

The analysis phase included:

Exploratory data analysis of the dataset.

Data cleaning and preprocessing to handle missing values, outliers, and inconsistencies.

Feature engineering to extract meaningful information from the available features.

**10. System Architecture**

The system architecture comprises

Employee Data Dataset:This is the dataset containing employee-related information, including features such as job satisfaction, distance from work, compensation, and performance.

10.1Data Preprocessing:

This component is responsible for performing data cleaning, handling missing values, outlier detection and treatment, and data normalization or scaling.

10.2 Feature Selection:

This component involves selecting the most relevant features from the preprocessed data to be used in the prediction model. Feature selection techniques such as correlation analysis, feature importance ranking, or dimensionality reduction methods can be applied.

10.3 Model Development:

This component focuses on building the prediction model using the random forest classifier algorithm. The model is trained on the selected features from the dataset.

10.4 Model Evaluation:

This component assesses the performance of the trained model using appropriate evaluation metrics such as accuracy, precision, recall, and F1 score. Cross-validation or holdout validation techniques can be used to evaluate the model's generalization ability.

10.5 Deployment:The final trained model is deployed to make predictions on new, unseen employee data. It can be integrated into a larger system or used independently to provide predictions on employee attrition.

**11.Screenshots**

