Machine Learning Project Report: Employee Salary Prediction

1. Project Overview

This project involves predicting employee salaries based on various attributes using a Machine Learning regression model. The dataset used includes features such as department, experience, education, and other employee-specific details. The objective is to build a model that accurately estimates salaries using Python's scikit-learn library.

2. Tools and Technologies

- Python
- Pandas, Matplotlib for data analysis and visualization
- Scikit-learn for model building and evaluation
- Linear Regression model for prediction
- StandardScaler and LabelEncoder for preprocessing

3. Project Workflow

1. Loaded and explored the dataset using pandas.
2. Cleaned the data by removing null values and encoding categorical variables.
3. Scaled features using StandardScaler.
4. Split data into training and testing sets (80/20).
5. Trained a Linear Regression model.
6. Evaluated performance using R^2 and Mean Squared Error.
7. Visualized results using a scatter plot of actual vs predicted salaries.

4. Model Evaluation

The Linear Regression model was evaluated using Mean Squared Error (MSE) and R^2 score. A scatter plot comparing actual vs. predicted salary values was generated to visually assess prediction accuracy.