Salary Prediction using Ensemble Learning

1. Project Overview

This project focuses on predicting salaries using ensemble learning techniques. The objective is to build a robust machine learning model that can estimate salaries based on features such as experience, education, job title, and location.

2. Dataset

The dataset used includes various attributes such as job title, years of experience, education level, and location. These features are used to train the model to predict the target variable - salary.

3. Technologies Used

- Python
- Pandas
- NumPy
- Scikit-learn
- Matplotlib / Seaborn
- Jupyter Notebook

4. Ensemble Methods Used

- Random Forest
- Gradient Boosting (XGBoost / LightGBM)
- Voting Regressor (Combination of multiple models)

5. Model Evaluation

The models are evaluated using metrics such as Mean Squared Error (MSE), Root Mean Squared Error

(RMSE), and R^2 Score to determine prediction accuracy.

6. Results

The ensemble models showed improved performance over individual models, with Random Forest and Gradient Boosting delivering the most accurate results.

7. How to Run the Project

1. Clone the repository

2. Install the dependencies using `pip install -r requirements.txt`

3. Run the notebook or script

4. View the predictions and evaluation results

8. Future Work

- Incorporate more features (e.g., company size, industry)

- Use deep learning techniques

- Deploy the model as a web app

9. Contributors

- Your Name Here

10. License