Initial Project Report

Project Title: Predicting Employee Salary Based on Experience and Education Level

We intend to study the application of regression analysis to predict employee salaries using a dataset from Kaggle. The dataset encompasses various attributes, such as years of experience, education level, job role, and industry, which are hypothesized to influence an employee's salary. By applying multiple linear regression, Lasso regression, and possibly other regression techniques, Our goal is to develop a predictive model that accurately forecasts salary based on these factors. The project will involve data preparation, exploratory data analysis, feature engineering, model building and validation, and interpretation of the results. The dataset for this study is sourced from Kaggle's "Employee Salary Dataset," which provides a rich compilation of employee salary data across different sectors and job roles​.

Dataset Link: <https://www.kaggle.com/datasets/anninasimon/employee-salary-dataset>

Research Papers:

1. G. Wang, "Employee Salaries Analysis and Prediction with Machine Learning," 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Guangzhou, China, 2022, pp. 373-378, doi: 10.1109/MLISE57402.2022.00081. keywords:

Link: <https://ieeexplore.ieee.org/document/9943146>

The paper by Guanqi Wang proposes a method for predicting employee salaries using Machine learning, Linear regression and Polynomial regression. It involves collecting a dataset of employee salaries along with various features such as education level, years of experience, job role, and industry. The study applies machine learning algorithms, including regression models, decision trees, to analyze patterns and predict salaries based on the features. The effectiveness of the models would be evaluated using metrics such as mean squared error or R-squared values.

1. “Employee Salary Prediction”, 2022 by Tiasa Mukherjee, MS. B. Satyasaivani

Link: <https://www.ijariit.com/manuscript/employees-salary-prediction/>

The paper by Tiasa Mukherjee and MS. B. Satyasaivani proposes a method for predicting employee salaries using machine learning techniques. It focuses on identifying key attributes that influence salary and employing a machine learning algorithm, specifically supervised learning and linear regression, to predict the salaries based on these attributes. The approach is rooted in the broader context of machine learning's capability to make accurate predictions by learning from data.