This project aims to predict whether a person's income is greater than $50,000 or less than or equal to $50,000 per year using supervised machine learning algorithms.

The dataset used is the **UCI Adult Income Dataset**, originally from the 1994 U.S. Census database. It contains over 32,000 records with attributes such as age, education, occupation, hours per week, and more. The target variable is the income level (<=50K or >50K).

The data was preprocessed by handling missing values, encoding categorical variables, removing outliers, and applying feature scaling. A total of **9 different models** were trained and tested, including Logistic Regression, k-Nearest Neighbors (kNN), SVM, Naive Bayes, Decision Tree, Random Forest, MLP Classifier, Gradient Boosting, and XGBoost.

Performance evaluation was based on accuracy, precision, recall, and F1-score. After tuning the models using GridSearchCV, **Gradient Boosting and XGBoost** were found to perform best, achieving an accuracy of **84.4%**.

The entire project was implemented in **Google Colab** using Python libraries like Pandas, Scikit-learn, and XGBoost.