

Project Report

Adult Census Income Prediction

Document Version: 1.0

Date: 25-Sept-2023

Author: Abhishek Pandit

Table of Contents

1. Executive Summary
2. Introduction
 - Project Overview
 - Objectives
 - Scope
3. Methodology
 - Data Collection
 - Data Preprocessing
 - Model Selection and Training
 - Application Design
4. Implementation
 - Data Ingestion
 - Data Transformation
 - Model Training
 - Prediction Pipeline
 - User Interface
5. Results
 - Model Evaluation
 - User Testing
 - Performance Metrics
6. Discussion
 - Challenges Faced
 - Lessons Learned
 - Future Enhancements
7. Conclusion
8. References

1. Executive Summary

The Machine Learning Income Prediction App is a web-based application designed to predict a person's income based on various input features. This report provides a detailed overview of the project, including its objectives, methodology, implementation details, results, and future considerations.

2. Introduction

Project Overview

The project aims to develop a machine learning-powered income prediction application that assists users in estimating their income based on demographic and employment-related factors.

Objectives

- To create an accurate machine learning model for income prediction.
- To design an intuitive user interface for data input and result display.
- To provide users with a user-friendly web application for income estimation.

Scope

The project includes data collection, preprocessing, model training, application development, and testing. The application allows users to input data such as age, education, and hours worked per week, and it returns income predictions.

3. Methodology

Data Collection

- Data was collected from the "adult.csv" dataset, containing demographic and employment-related information.

Data Preprocessing

- Data was preprocessed to handle missing values and prepare it for model training.

Model Selection and Training

- Various machine learning models were evaluated using grid search and cross-validation.
- The best-performing model was selected and trained on the preprocessed data.

Application Design

- The application was designed using Streamlit for the user interface.
- User input data is preprocessed, and predictions are made using the selected model.

Adult Census Income Prediction

4. Implementation

Data Ingestion

- Raw data was loaded and preprocessed, resulting in training and test datasets.

Data Transformation

- Input features were transformed and preprocessed using pipelines for feature scaling and encoding.

Model Training

- Machine learning models were trained on the transformed training data, and the best model was selected.

Prediction Pipeline

- User input data is preprocessed using the preprocessor, and predictions are made using the selected model.

User Interface

- The Streamlit app provides a web-based interface for users to input data and view predictions.

5. Results

Model Evaluation

- The selected model's performance was evaluated using accuracy and other relevant metrics.

User Testing

- User testing was conducted to ensure the app's usability and accuracy.

Performance Metrics

- The app's performance metrics, including prediction accuracy, were measured and documented.

6. Discussion

Challenges Faced

- Challenges encountered during data preprocessing and model selection.
- Overcoming issues related to data quality and quantity.

Lessons Learned

- Key takeaways from the project, including best practices in machine learning and application development.

Future Enhancements

- Future improvements and features, such as user authentication and enhanced data input options.

7. Conclusion

The Machine Learning Income Prediction App successfully achieves its objectives by providing users with an accurate income prediction tool. The project demonstrates the integration of machine learning into web application development.

8. References

- List of data sources, libraries, and resources used in the project.

The following dependencies are required to run the app:

- Python
- pandas
- scikit-learn
- Streamlit
- Numpy

Software and account Requirement.

1. [Github Account](#)
2. [VS Code IDE](#)
3. [GIT cli](#)
4. [GIT Documentation](#)