Low Level Design

Adult census Income Prediction

Written By	Abhishek Pandit
Document Version	0.1
Last Revised Date	25 – Sep -2023

Table of Contents

1. Introduction

This Low-Level Design (LLD) document provides detailed information about the design and implementation of the Machine Learning Income Prediction App. It outlines the class structure , data structures, functions and methods

2. Design Overview

Architecture:

The application follows a modular architecture with the following components:

- Data Ingestion
- Data Transformation
- Model Training
- Prediction Pipeline
- Streamlit User Interface

Dependencies:

- Python
- Pandas
- Numpy
- Ipykernel
- seaborn
- scikit-learn
- Streamlit

3. Data Structures

Input Data:

- DataFrame for user input

Preprocessing Objects:

- Preprocessor object for data transformation

Model Objects:

- Trained machine learning model

4. Functions and Methods

- Data Ingestion (data_ingestion.py):
  - load_data(): Loads data from the dataset.
  - preprocess_data(): Handles data preprocessing.

- Data Transformation (data_transformation.py):
  - transform_data(): Applies data transformation and preprocessing.
  - create_preprocessor(): Creates a preprocessor object.

- Model Training (model_trainer.py):
  - train_model(): Trains machine learning models.
  - select_best_model(): Selects the best-performing model.

- Prediction Pipeline (prediction_pipeline.py):
  - preprocess_input(): Preprocesses user input data.
  - make_prediction(): Makes predictions using the selected model.

5. Conclusion

This Low-Level Design document provides an in-depth understanding of the architecture, data flow, class structure, sequence of operations, user interface, data structures, functions, methods, and exception handling strategies of the Machine Learning Income Prediction App. It serves as a valuable reference for developers and stakeholders involved in the project.