Low Level Design (LLD)

Low Level Design (LLD)

[**Adult\_income\_Prediction**](https://github.com/mohiteyashprogrammer/Adult_income_Prediction/tree/main)

Revision Number – 1.0

Last Date of Revision – 11-04-2023

**YASH MOHITE**

1

Low Level Design (LLD)

2

Low Level Design (LLD)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Contents** | | | | | | | | | | | | | | | |  |  |  |  |
| [**Abstract**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.gjdgxs) | |  | | | | | | | | | | | | | | [4](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.gjdgxs) | | | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| [**INTRODUCTION**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.30j0zll) | | | | |  | | | | | | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.30j0zll) | | | |
|  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | [**Why this LLD documentation?**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1fob9te) | | | | | | | | | | | | | | | [5](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1fob9te) | | |  |
|  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| [**1 Architecture**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3znysh7) | | | | | | | | | | | | | | | | [**3**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3znysh7) | | | |
|  |  | |  |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |
| [**2 Architecture Design**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3dy6vkm) | | | | | | | |  | | | | | | | | [**6**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3dy6vkm) | | | |
|  |  | | |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |
|  | [**2.1 Data Gathering from Main Source**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1t3h5sf) | | | | | | | | | | | | | | | [**6**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1t3h5sf) | | |  |
|  |  | | |  |  |  |  | |  |  |  |  | |  |  |  |  |  |  |
|  | [**2.2 Tools Used**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.17dp8vu) | | | |  | | | | | | | | | | |  | | [**6**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.17dp8vu) | |
|  |  | | |  |  |  |  | |  |  |  |  | |  |  |  |  | |  |
|  | [**2.3 Data Description**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3rdcrjn) | | | | | | | | | | | | | | | [**7**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3rdcrjn) | | |  |
|  |  | | |  |  |  |  | |  |  |  |  | |  |  |  |  | |  |
|  | [**2.4 Import Data into Database**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.26in1rg) | | | | | | | | | | | | |  | | [**7**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.26in1rg) | | |  |
|  |  | | |  |  |  |  | | |  |  |  | |  |  |  |  | |  |
|  | [**2.5 Export Data from Database**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.lnxbz9) | | | | | | | | | | | | | |  | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.lnxbz9) | | |  |
|  |  | | |  |  |  |  | | |  |  |  | | |  |  |  | |  |
|  | [**2.6 Data Pre-Processing**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.35nkun2) | | | | | | | | | |  | | | |  | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.35nkun2) | | |  |
|  |  | | |  |  |  |  | | |  |  |  | | | |  |  | |  |
|  | [**2.7 Modelling**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1ksv4uv) | | | | | | | | | |  | | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1ksv4uv) | | | |
|  |  | | |  |  |  |  | | |  | |  | | | |  |  | |  |
|  | [**2.8 UI Integration**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.44sinio) | | | | |  | | | | | | | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.44sinio) | | | |
|  |  | | | |  |  |  | | |  | |  | | | |  |  | |  |
|  | [**2.9 Data From User**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.2jxsxqh) | | | | | |  | | | | | | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.2jxsxqh) | | | |
|  |  | | | |  | |  | | |  | |  | | | |  |  | |  |
|  | [**2.10 Data Validation**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.z337ya) | | | | | | | | |  | | | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.z337ya) | | |  |
|  |  | | | |  | |  | | |  | |  | | | |  |  | |  |
|  | [**2.11 Rendering the Results**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3j2qqm3) | | | | | | | | | | |  | | | | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.3j2qqm3) | | |  |
|  |  | | | |  | |  | | | | |  | | | |  |  | |  |
|  | [**3 Deployment**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1y810tw) | | | |  | | | | | | |  | | | |  | [**8**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.1y810tw) | |  |
|  | [**3.1 Unit Test Cases**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.4i7ojhp) | | | | | |  | | | | | | | | | [**9**](https://docs.google.com/document/d/1CawpaB-mO_T9ew2gcRJPxTYNEbUemWHc/edit#heading=h.4i7ojhp) | | | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

3

Low Level Design (LLD)

**Abstract**

The objective of this project is to develop a predictive model that can accurately estimate the income level of adults based on various demographic, educational, and occupational features. The model should analyze a given individual's characteristics and provide a prediction of whether their income exceeds a certain threshold, such as $50,000 per year.

The problem involves leveraging historical data on adult income along with associated attributes to train the predictive model. The trained model will then be used to make income predictions for new individuals based on their input features.

The goal is to create an accurate and robust income prediction model that can assist in making informed decisions, such as resource allocation, policy planning, and financial assessments. The model should take into account factors such as education, occupation, work experience, demographics, and any other relevant attributes that may contribute to an individual's income level.

The success of the project will be evaluated based on the model's predictive performance, measured by appropriate evaluation metrics such as accuracy, precision, recall, or area under the receiver operating characteristic curve (AUC-ROC). The model should aim to achieve high accuracy and minimize false predictions.

The developed model should be applicable to diverse populations and exhibit fairness, ensuring that predictions are not biased or discriminatory based on factors such as gender, race, or ethnicity. Ethical considerations should be taken into account throughout the development process to ensure responsible and unbiased predictions.

Overall, the project aims to provide a reliable and interpretable income prediction model that can assist in understanding income disparities, identifying potential opportunities for improvement, and informing decision-making processes related to income distribution and economic policies.

4

Low Level Design (LLD)

**1 Introduction**

**1.1 Why this Low-Level Design Document?**

The main purpose of this LLD documentation is to feature the required details of the project and supply the outline of the machine learning model and also the written code. This additionally provides the careful description on however the complete project has been designed end-to-end.

**1.2 Architecture**



**2. Architecture Design**

This project is to make associate interface for the user to grasp their approximate flight price ticket worth, additionally to the present, in would like of obtaining the important time project expertise we have a tendency to square measure mercantilism the gathered information into our own information then begin the project from the scratch.

5

Low Level Design (LLD)

**2.1. Data Gathering**

The data for the current project is being gathered from Kaggle dataset, the link to the data is: <https://www.kaggle.com/nikhilmittal/flight-fare-prediction-mh>

**2.2. Tool Used**

• Python 3.7 is employed because the programming language and frame works like numpy, pandas, sklearn and alternative modules for building the model.

* PyCharm is employed as IDE.
* For visualizations seaborn and components of matplotlib are getting used
* For information assortment prophetess info is getting used version

management.

• Netlify is employed for deployment

**2.3 Data Description**

Predicting adult income is a complex task that involves analyzing various factors related to an individual's demographics, education, occupation, and other relevant variables. While I can provide you with some general information, please note that predicting income accurately for specific individuals can be challenging due to the wide range of factors involved and the inherent uncertainty of future outcomes.

6

Low Level Design (LLD)

**2.4 Import Data into Database**

* Created associate api for the transfer of the info into the Cassandra info, steps performed are:
* Connection is created with the info.
* Created a info with name flightfare.
* Cqlsh command is written for making the info table with needed parameters.
* And finally, a cqlsh command is written for uploading the knowledgeset into data table by bulk insertion.

**2.5 Export Data into Database**

In the above created api, the download url is also being created, which downloads the data into a csv

file format.

**2.6 Data Preprocessing**

Steps performed in pre-processing are:

* First the info sorts square measure being checked and located solely the value column is of sort number.
* Checked for null values as there square measure few null values, those rows square measure born.
* Converted all the desired column into the date time format.
* Performed one-hot cryptography for the desired columns.
* Scaling is performed for needed information.
* And, the info is prepared for passing to the machine learning formula

7

Low Level Design (LLD)

**2.7 Modelling**

The pre-processed information is then envisioned and everywhere the specified insights are being drawn. Though from the drawn insights, the info is at random unfold however still modelling is performed with completely different machine learning algorithms to form positive we tend to cowl all the chances and eventually, for sure random forest regression performed well and any hyperparameter calibration is finished to extend the model’s accuracy.

**2.8 UI Integration**

Both CSS and HTML files are being created and are being integrated with the created machine learning model. All the required files are then integrated to the app.py file and tested locally.Note CSS and HTML is not done by me.

**2.3 Data from User**

The data from the user is retrieved from the created HTML web page.

**2.4 Data Validation**

The data provided by the user is then being processed by app.py file and validated. The validated data is then sent for the prediction.

**2.11 Rendering Result**

The data sent for the prediction is then rendered to the web page.

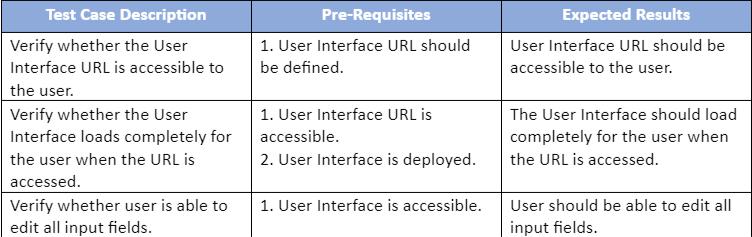
**3. Deployment**

The tested model is then deployed to Netlify. So, users can access the project from any internet devices.

8

Low Level Design (LLD)

**3.1 Unit Test**



9