Low Level Design

Adult Census Income Predictor

Written By	Shubham Rawat
Document Version	0.1
Last Revised Date	22 – Aug-2022

Document Control

Change Record:

Version	Date	Author	Comments
0.1	19 – Aug - 2022	Shubham Rawat	Introduction & Architecture defined, Architecture & Architecture Description appended and updated, Unit Test Cases defined and appended

Reviews:

Version	Date	Reviewer	Comments

Approval Status:

Version	Review Date	Reviewed By	Approved By	Comments

Contents

1. Introduction

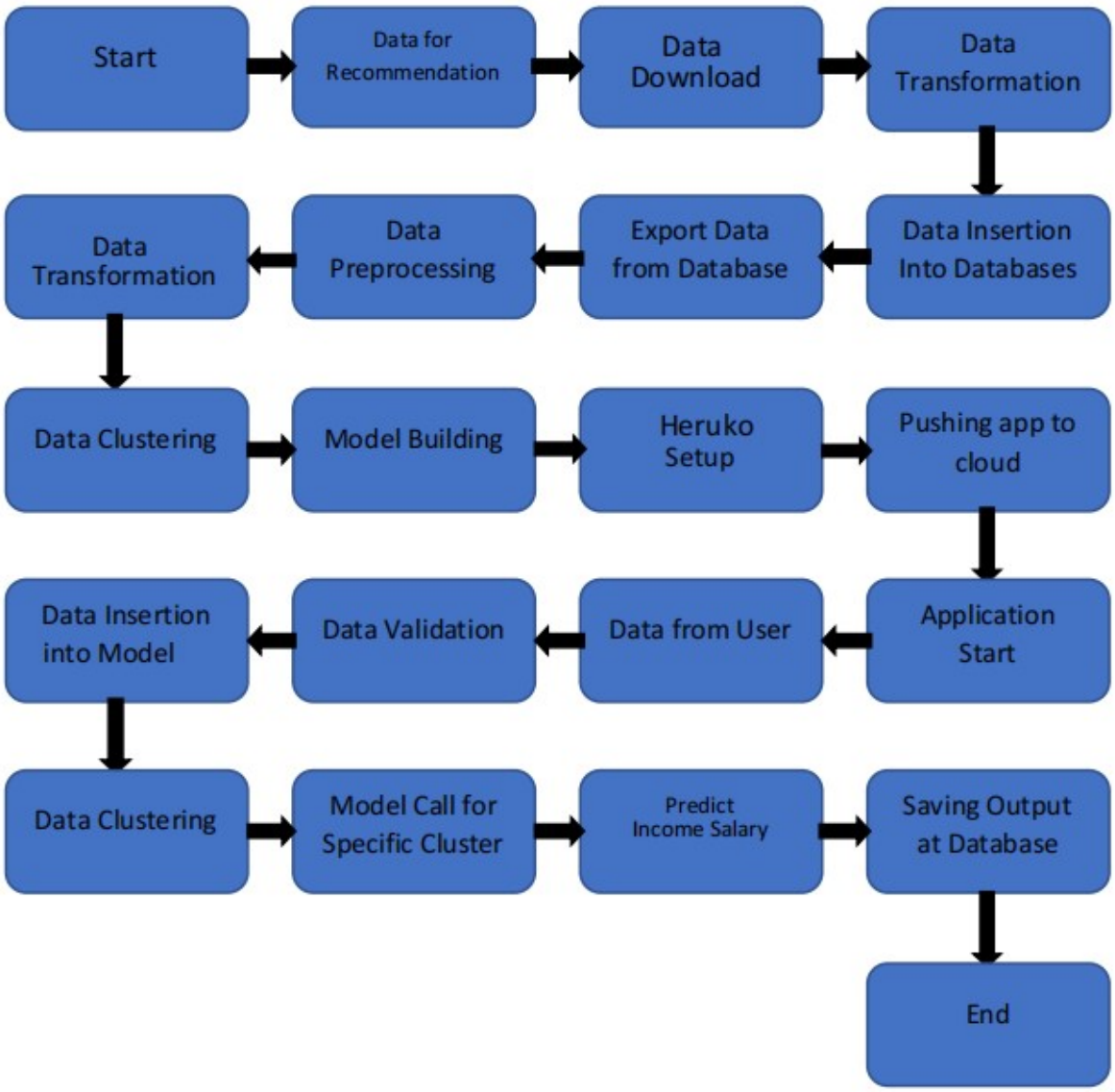
1.1. What is Low-Level design document?

The goal of LLD or a low-level design document (LLDD) is to give the internal logical design of the actual program code for Food Recommendation System. LLD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document.

1.2. Scope

Low-level design (LLD) is a component-level design process that follows a step-by step refinement process. This process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work

2. Architecture



3. Architecture Description

3.1. Data Description

This Data was extracted from the 1994 Census bureau database by Ronny Kohavi and Barry Becker (Data Mining and Visualization, Silicon Graphics). we Have more then thirty thousand data row and each row have fourteen columns such as age, race, Gender, Education Status, Marital Status, Job Title, Country of Origin etc.

3.2. Data Ingestion

Data is store in Git Hub so we receive git-hub link.
Data is Downloaded into local system and Processed further

3.3. Data Transformation

In the Transformation Process, we will convert our original data set which is in JSON format to CSV Format. And will merge it with the Scrapped data set.

3.4. Data Preprocessing

Data Pre-processing steps we could use are Null value handling, filling those null value with suitable values, Data Encoding and Data Standard Scales after scaling the data then we save that Prepocessing model as a pickle for further use.

Data leakage is prevented by applying fit_transform and saving data into NP.nyz format.

3.5. Export Data from Database

Data Export from Database .

3.6. Model Building

After clusters are created, we will find the best model for each cluster. For each cluster, algorithms will be passed with the best parameters derived from Grid-Search. We will calculate the MAE scores for models and select the model with the best score. Similarly, the models will be selected for each cluster.

3.7. Data from User

Here we will collect financial data from user such as age, education, gender etc.

3.8.Data Validation

Given data is validated on the bases of there name data type schema Column name and shape.
These Information is already decided with negotiation with Clients. According to agreement and conditions we Validate the data .

3.8. Recipe Recommendation & Saving Output in Database

After calling model Recipe/Output will be recommended, this output will be display

3.9. Deployment

We will be deploying the model to Heruko.
This is a work flow diagram for the Adult Census Income..