



High-LEVEL DESIGN

Fraud Transaction Detection



Document Version Control

Date Issued	Version	Description	Author
09.08.2022	V1.0	Initial LLD- V1.0	Shubham chavhan

Contents

Document Version Control	1
1.0 Introduction	3
1.1 What is Low-Level Design Document?	3
1.2 Scope	3
2.0 Architecture	4
3.0 Architecture Description	5
3.1 Data Description	5
3.2 Exploratory Data Analysis	5
3.3 Data Pre-processing	5
3.4 Model Building	5
3.5 Data Validation.....	6
3.6 Deployment.....	6
4.0 Unit Test Cases	7

1.0 Introduction

1.1 What is Low-Level Design Document?

The goal of LLD or Low-Level design document (LLDD) is to give the internal logical design of the actual program code. Low-Level design is created based on the High-Level design. 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.

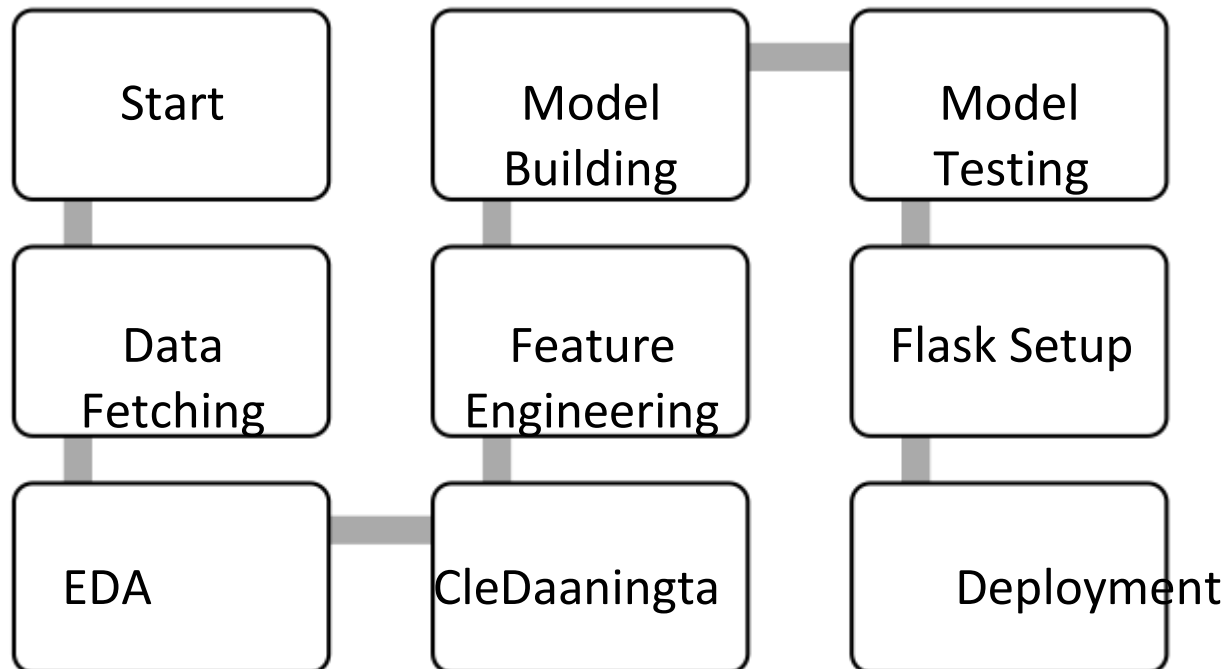
1.2 Scope

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.

1.3 Tool used:

Numpy, Pandas, Matplotlib, Sklearn, Seaborn

2.0 Architecture



3.0 Architecture Description

3.1 Data Description

The data is given by INeuron

3.2 Exploratory Data Analysis

Exploring the data by visualizing the distribution of values in some columns of the dataset, and the relationships between expenses and other columns. Visualizing the distribution of age, BMI (body mass index). Also checking the region wise have any differences in the expenses.

3.3 Data Pre-processing

If data is not suited to take place directly for the regression. Then, cleaning of dataset becomes important for using the data under various regression algorithms.

3.4 Model Building

After data pre-processing is done, we will split the dataset into training set and validation set. Then we will use training set for building the best model. The model will be trained on several algorithms. We will calculate RMSE and r^2 score for each model and select the model with the best score.