Low Level Design (LLD)

Store Sales Prediction

|  |  |
| --- | --- |
| Written By | Neet , Madhav , Pavitra |
| Document Version | Initial LLD -V1.0 |
| Last Revised Date | 03-04-2022 |

**Document Control**

**Change Record:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Comments** |
| **0.1** | **03-04-2022** | **Madhav**  **Pavitra**  **Madhav** | **Architecture data creation.** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Contents

1.Introduction.....................................................................................................................5

1.1 What is Low-Level design document........................................................................5

1.2 Scope..........................................................................................................................5

2. Architecture...................................................................................................................6

3. Architecture Description...............................................................................................7

3.1 Data Description.....................................................................................................7

3.2 Data Gathering ......................................................................................................7

3.3 Data Pre-Processing ..............................................................................................8

3.4 Feature Engineering...............................................................................................8

3.5 Parameter Tuning...................................................................................................8

3.6Deployment.............................................................................................................8

4. Unit Test Cases .............................................................................................................9

Abstract

In this era of modernisation and e-learning, everything is based onto internet which generate a huge amount of data, which can be used by humans in various ways to create more advancement in this humongous world of data. This whole process is completed with various machine learning algorithms to solve the purpose and get desired data, but particularly in this case we are trying to predict the sales of store using different machine learning techniques and trying to determine the best algorithm suited to our particular problem statement.

1. Introduction

* 1. What is a Low-Level design document?

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. Post-build, each component is specified in detail or in other words it can be describe as the goal of LLD 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 as well as It is used to design internals of the individual modules identified during HLD i.e., data structures and algorithms of the modules are designed and documented.

1.2. Scope

LLD helps in providing the interface for all classes that include public and private methods, including parameters, return. It also helps in describe and justify your choice of data structures as well as the major alternatives that you considered and why you made the choice.

Design documents are created to coordinate efforts of a large team, give them a stable reference point, and describe all parts of the software and how they will operate. It assures that the product is built to meet the needs and is on par with what was agreed upon prior to the inception of the product/software.

2. Architecture



3. Architecture Description

* 1. Data Description

[BigMart Sales Data](https://www.kaggle.com/brijbhushannanda1979/bigmart-sales-data) is the biggest publicly available recipe dataset. We have the train (8523) and test (5681) data set, and the train data set has both input and output variable(s). We need to predict the sales for the test data set.

|  |  |  |
| --- | --- | --- |
| Name | Data Type | Measurement |
| Item\_Identifier | String | Unique product ID |
| Item\_Weight | Float | Weight of product |
| Item\_Fat\_Content | String | Whether the product is low fat or not |
| Item\_Visibility | Float | The % of a total display area of all products in a store allocated to the particular product |
| Item\_Type | String | The category to which the product belongs |
| Item\_MRP | Float | Maximum Retail Price (list price) of the product |
| Outlet\_Identifier | String | Unique store ID |
| Outlet\_Establishment\_Year | Integer | The year in which the store was established |
| Outlet\_Size | String | The size of the store in terms of ground area covered |
| Outlet\_Location\_Type | String | The type of city in which the store is located |
| Outlet\_Type | String | Whether the outlet is just a grocery store or some sort of supermarket |
| Item\_Outlet\_Sales | Float | Sales of the product in the particular store. This is the outcome variable to be predicted. |

* 1. Data Gathering

Data source: [**https://www.kaggle.com/brijbhushannanda1979/bigmart-sales-data**](https://www.kaggle.com/brijbhushannanda1979/bigmart-sales-data)

Train and Test data are stored in .csv format.

* 1. Pre-processing

In data pre-processing all the processes required before sending the data for model building are performed. Like, here the ‘Item Visibility’ attributes are having some values equal to 0, which is not appropriate because if an item is present in the market, then its visibility can be 0. So, it has been replaced with the average value of the item visibility of the respective ‘Item Identifier’ category. New attributes were added named ‘’Outlet years”, where the given establishment year is subtracted from the current year. A new “Item Type” attribute was added which just takes the first two characters of the Item Identifier which indicates the types of the items. Then mapping of “Fat content” is done based on ‘Low’, ‘Reg’ and ‘Non-edible’.

* 1. Feature Engineering

A "feature" in the context of predictive modelling is just another name for *a*predictor variable*.*Feature engineering is the general term for creating and manipulating predictors so that a good predictive model can be created. After pre-processing it was found that some of the attributes are not important to the item sales for the particular outlet. So those attributes are removed. Even one hot encoding is also performed to convert the categorical features into numerical features.

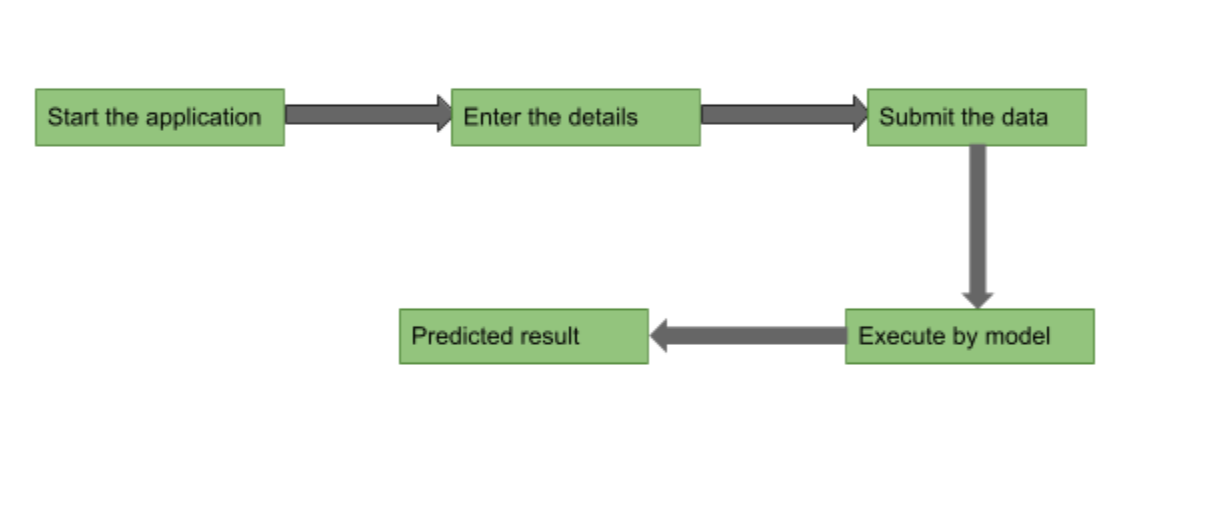
* 1. Parameter Tuning

Parameter tuning is crucial as they control the overall behaviour of a machine learning model. Every machine learning model will have different parameters that can be set before the learning process begins. Parameters are tuned using Randomized searchCV. The algorithm which we used is Random Forest Regressor.

* 1. Deployment

App link- Will appear in YouTube.

App demo link - https://youtu.be/1Gt8gyJR3TA



3. Unit Test Cases

|  |  |  |
| --- | --- | --- |
| Test Case Description | Pre-Requisite | Expected Result |
| Verify whether the Application URL is  accessible to the user | 1. Application URL  should be defined | Application URL should be  accessible to the user |
| Verify whether the Application loads  completely for the user when the URL  is accessed | 1. Application URL  is accessible  2. Application is  deployed | The Application should load  completely for the user when the  URL is accessed |
| Verify whether user is able to see input  fields | 1. Application is  accessible | User should be able to see input  fields |
| Verify whether user gets Submit  button to submit the inputs | 1. Application is  accessible | User should get Submit button to  submit the inputs |
| Verify whether the predicted  results are in accordance to the  selections user made | 1. Application is  accessible | The predicted results should  be in accordance to the selections  user made |