High Level Design (HLD)

High Level Design (HLD)

PETROL PRICE FORECASTING

Revision Number – 1.3

Last Date of Revision – 20/04/2022

Naveen Gupta

(Document Version Control

HIGH LEVEL DESIGN (HLD)

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Version | Description | Author |
| 17-04-2022 | 1.0 | Abstract, | Naveen |
|  |  | Introduction |  |
|  |  | Problem |  |
|  |  | Statement |  |
| 18-04-2022 | 1.1 | Design Flow | Naveen |
| 19-04-2022 | 1.2 | Performance | Naveen |
|  |  | Evaluation |  |

HLD)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Contents** | | | | | | | | | | | | | | | | | | | | | |
| [**Abstract**](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) | | | | | | | | | |  | | | | | | | | | | | |
|  |  |  |  | |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | [**Why this HLD documentation?**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1fob9te) | | | | | | | | | | | | | | | | | | | | |
|  |  |  |  | |  | |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |
| [**1 Description**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2et92p0) | | | | | | | |  | | | | | | | | | | | | | |
|  |  | |  |  |  | |  |  |  | |  |  |  |  |  |  |  |  |  |  |  |
|  | [**1.1**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.tyjcwt) | |  |  | [**Problem Perspective**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.tyjcwt) | | | | | | | | | | | | | | | | |
|  |  | |  | |  | |  | |  | |  |  |  |  |  |  |  |  |  |  |  |
|  | [**1.2 Problem Statement**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3dy6vkm) | | | | | | | | | | | | | | |  | | | | | |
|  |  | | | | | |  | |  | |  |  |  |  |  |  |  | |  |  |  |
|  | [**1.3 Proposed Solution**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1t3h5sf) | | | | | | | | | | | | | | |  | | | | | |
|  |  | | | | | |  | |  | |  |  |  |  |  |  |  | |  |  |  |
|  | [**1.4 Solution Improvements**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.4d34og8) | | | | | | | | | | | | | | | | | | | | |
|  |  | | | | | |  | |  | |  |  |  |  |  |  |  | |  |  |  |
|  | [**1.5 Technical Requirements**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2s8eyo1) | | | | | | | | | | | | | | | | | | |  | |
|  | [**1.6 Data Requirements**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.17dp8vu) | | | | | | | | | | | | | | |  | | | | | |
|  |  | | | | | |  | |  | |  |  |  |  |  |  |  | | | |  |
|  | [**1.7 Tools Used**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3rdcrjn) | | | | | | | | | |  | | | | |  | | | | | |
|  |  | |  |  | | |  | |  | |  |  |  |  |  | |  | | | |  |
|  | [**1.8**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.26in1rg) | |  | [**Constraints**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.26in1rg) | | | | | | | |  | | | | | | | | | |
|  |  | | | | | |  | |  | | |  |  |  |  | |  | | | |  |
|  | [**1.9 Assumptions**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.lnxbz9) | | | | | | | | | | | | | | | | | | | | |
|  | |  | | | | |  | |  | | |  |  |  |  | |  | | | |  |
| [**2 Design Flow**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.35nkun2) | | | | | | | | |  | | | | | | | | | | | | |
|  |  | | | | | |  | |  | | |  |  | |  | |  | | | |  |
|  | [**2.1 Modelling Process**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1ksv4uv) | | | | | | | | | | | | | | | | | | | | |
|  |  | | | | | |  | | | | |  |  | |  | |  | | | |  |
|  | [**2.2 Deployment Process**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.44sinio) | | | | | | | | | | | | | | | | | | | | |
|  | |  | |  | | |  | | | | |  |  | | | |  | | | |  |
| **2.3** | | | | **Logging** | | | | | | | | | | | | | | | | | |
| **2.2** | | | | **Error Handling** | | | | | | | | | | | | | | | | | |
| [**3 Performance Evaluation**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.4i7ojhp) | | | | | | | | | | | | | | | | |  | | | | |
|  | | |  |  | | |  | | | | |  |  | | | |  | | | |  |
|  | [**3.1**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2xcytpi) | |  | [**Reusability**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2xcytpi) | | | | | | | |  | | | | |  | | | | |
|  |  | |  |  | | |  | | | | |  |  | | | |  | | | |  |
|  | [**3.2**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1ci93xb) | |  | [**Application Compatibility**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.1ci93xb) | | | | | | | | | | | | | | | | | |
|  | [**3.3**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3whwml4) | |  | [**Resource Utilization**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.3whwml4) | | | | | | | | | | | | |  | | | |  |
|  |  | | | | | |  | | | | | |  | | | |  | | | | |
|  | [**3.2 Deployment**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.2bn6wsx) | | | | | | | | | | | | | | | |  | | | | |
|  |  | | | | | |  | | | | | |  | | | | | | | | |
| [**Conclusion**](https://docs.google.com/document/d/1ZLlZEVdMuxYl3o0PziQjZcUbNmAcz-PR/edit#heading=h.qsh70q) | | | | | | | | | | | | | | | | | | | | | |

**Abstract**

The liberalization of the petroleum sector in Morocco has a significant effect for petroleum product distributors. Since the beginning of December 2015, fuel prices are freely determined. This event presents many constraints affecting the balance of the sector plus the competition between its economic players. The lack of accompanying measures by the State makes this vital reform for public finances that stop subsidizing the price of gasoline vulnerable. With the halt of the competitive manufacturing's activity, Morocco's only refinery, distributors must, for their part, build up large stocks. As all fuel products are imported, we will be interested in the evolution by making forecasts of the price of fuels in the Moroccan market. In order to achieve their objectives, the oil companies must rely on precise forecasts. In this context, our paper aims mainly to study the time series of diesel and gasoline in order to provide precise forecasts to the company and to respect the permissible error margin of 3%. To this end, we worked with the FBPROPHET method. We found that the FBPROPHET method gives forecasts of the price of gasoline near the margin to be met for the first quarter of the current year with an average error margin of 2,855%. In addition, the assumption that the residuals are a Gaussian white noise has always been verified.

**1 Introduction**

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

The main purpose of this HLD 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 Description**

**Problem Perspective**

The Petrol Price Forecasting may be a machine learning model that helps Business to forecast the sale of the store’s products and helps the users to manage the sale of the products.

**1.3 Problem Statement**

The ONGCF is a organisation dedicated to the exploration and production of oil and

natural gas. Price information is supplied on a weekly basis. It seeks to forecast crude

oil prices for the following 16 months, from January 1, 2019 to April 1, 2020.

The main goal is to predict the forecast the prices based upon the best model as per

your choice.

**1.4. Project Solution**

Project requires the desired input of user from the created interface and method all the provided information to satisfy the wants of the machine learning model and at last show the expected output .

**1.5 Answer enhancements**

We will even predict that how much the sale of a product on yearly , quarterly ,monthly, daily or hourly basis that will more help for the business to getting insights of a sale forecasting in advance.

**1.6 Technical needs**

There are not any hardware needs needed for victimization this application, the user should have AN interactive device that has access to the web and should have the fundamental understanding of providing the input. And for the backend half the server should run all the package that's needed for the process and provided information to show the results.

**1.7 Information needs**

The info demand is totally supported the matter statement. and also, the information set is accessible on the Kaggle within the type of standout sheet(.xlsx), because the main theme of the project is to induce the expertise of real time issues, we have a tendency to once more mercantilism {the information into the prophetess data base and commerce it into csv format.

**1.8Tools Used**

* Python 3.9 is employed because the programming language and frame works like fbprophet , numpy , pandas, sklearn and alternative modules for building the model.
* Visual Studio Code is employed as IDE.
* For visualizations seaborn , plotly and components of matplotlib are getting used.
* For information assortment prophetess info is getting used.
* Front end development is completed victimization HTML/CSS.
* Flask is employed for each information and backend readying.
* GitHub is employed for version management.
* Heroku is employed for deployment.
* SQL SERVER IS USED FOR DATABASE.
* Power Bi is used for creating a report.

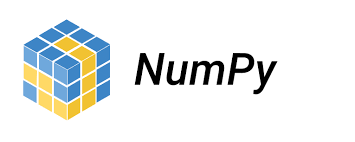
**1.9 Constraints**

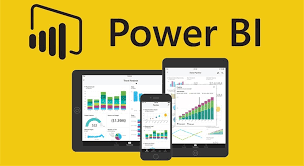
The Back Order Prediction answer should be user friendly, as automatic as attainable and also the user should not be needed to understand any of the operating.

**1.10 Assumptions**

The most objective of the project is to implement the utilization cases as for the new dataset that user provides through the programme. Machine learning model is employed for process the on top of computer file. It's additionally assumed that each one aspects of this project have the flexibility to figure along within the approach as the designer is expecting.





**2.1 and 2.2 Design Flow and Deployment Process**

PYTHON

SQL SERVER

Data (CSV)

EDA

DATA PREPROCESSING

IMPORT PYTHON LIBRARIES & READ DATA

FEATURE SELECTION

FITTING MODEL WITH FBPROPHET

FEATURE ENGINEERING

CREATING A WEB PAGE FOR DEPLOYMENT

SAVE MODEL IN PICKLE FILE

HYPER PARAMETER TUNNING

EXPORT DATA BACK TO SQL SERVER

DEPLOYMENT ON LOCAL HOST USING FLASK AND VS CODE

DEPLOYMENT ON HEROKU

CREATING POWER BI REPORT

INPUT VALUE & PREDICT FINAL RESULT

ARCHITEC

**2.3 Logging**

Each step is being logged within the system that runs internally, that shows the date time and therefore the processed that has been performed, work is completed in several layers as information, DEBUG, ERROR, WARNINGS. this provides US the perceive of the logged info.

**2.4 Error Handling**

Once ever a slip is occurred, the reason are logged in its several log file, in order that the developer will rectify the error.

ARCHITECTURE

**3 Performance analysis**

**3.1 Reusability**

Elements of the code written is accustomed different applications and therefore the rest is changed and be reused.

**3.2 Application Compatibility**

The various parts for this project are exploitation python as associate interface between them. Every element can have its own tasks to perform, and it's the work of the python to make sure correct transfer of data.

**3.3 Resource Utilization**

Once any task is performed, it'll doubtless; use all the process power offered till that performs is finished.

**3.4 Deployment**

The model is being deployed on Heroku.

**Conclusion**

The Petrol Price Forecasting will forecast the worth supported the trained knowledge set within the rule. Therefore, the user will recognize the approximate result for his or her product.