Architecture

Architecture

PETROL PRICE FORECASTING

Revision Number – 1.3

Last Date of Revision – 20/04/2022

Naveen Gupta

1

Architecture

Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Version | Description | Author |
| 18-04-2022 | 1.0 | Abstract, | Naveen |
|  |  | Introduction |  |
| 19-04-2022 | 1.1 | Architecture | Naveen |
| 20-04-2022 | 1.2 | Architecture | Naveen |
|  |  | Design |  |

2

ARCHITECTURE

**Contents**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [**Abstract**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.gjdgxs) | | | | | | | | | | | | | | | [4](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.gjdgxs) | | |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| [**Introduction**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.30j0zll) | | |  | | | | | | | | | | | | [4](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.30j0zll) | | | |
|  |  | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | [**Why this Architecture Design documentation?**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.1fob9te) | | | | | | | | | | | | | |  | | [4](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.1fob9te) |  |
| [**1 Architecture**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.3znysh7) | | | |  | | | | | | | | | | | [4](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.3znysh7) | | | |
|  | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| [**2 Architecture design**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.3dy6vkm) | | | | | | | |  | | | | | | | [5](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.3dy6vkm) | | |  |
|  |  | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | [**2.1 Data gathering from main source**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.1t3h5sf) | | | | | | | | | | | | | | [5](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.1t3h5sf) | | |  |
|  |  | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | [**2.2 Data description**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.4d34og8) | | | | | | | | | | | | |  | [5](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.4d34og8) | | |  |
|  |  | | | |  |  |  |  |  |  |  |  |  | |  |  |  |  |
|  | [**2.3 Upload data into Cassandra**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.2s8eyo1) | | | | | | | | | | | |  | | [5](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.2s8eyo1) | | | |
|  |  | | | |  |  |  |  | |  |  |  |  | |  |  |  |  |
|  | [**2.4 Export data from database**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.17dp8vu) | | | | | | | | | | |  |  | | [5](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.17dp8vu) | | |  |
|  |  | | | |  |  |  |  | |  |  |  | | |  |  |  |  |
|  | [**2.5 Data pre-processing**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.3rdcrjn) | | | | | | | | | | |  | | |  | [5](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.3rdcrjn) | |  |
|  |  | | | |  |  |  |  | |  |  | | | |  | |  |  |
|  | [**2.6 Modelling**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.26in1rg) | | | |  | | | | |  | | | | |  | | [6](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.26in1rg) | |
|  |  | | | |  |  |  |  | | |  | | | |  | |  |  |
|  | [**2.7 UI integration**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.lnxbz9) | | | | |  | | | | | | | | | [6](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.lnxbz9) | | |  |
|  |  | | | | |  |  |  | | |  | | | |  | |  |  |
|  | [**2.8 Data from user**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.35nkun2) | | | | | | | | | | | | | | [6](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.35nkun2) | | |  |
|  |  | | | | |  |  |  | | |  | | | |  | |  |  |
|  | [**2.9 Data validation**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.1ksv4uv) | | | | | | |  | | | | | | | [6](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.1ksv4uv) | | |  |
|  |  | | | | |  | |  | | |  | | | |  | |  |  |
|  | [**2.10 Rendering the results**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.44sinio) | | | | | | | | | |  | | | | [6](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.44sinio) | | |  |
|  | [**2.11 Deployment**](https://docs.google.com/document/d/14Wck6wDTC8bscpCxNjrSRu01-UC_tJ5Y/edit#heading=h.2jxsxqh) | | | | | | | | | |  | | | |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

3

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.

Introduction

Why this Architecture Design documentation?

The main objective of the Architecture design documentation is to provide the internal logic understanding of the Petrol Price Forecasting code. The Architecture design documentation is designed in such a way that the programmer can directly code after reading each module description in the documentation.

1 Architecture

ARCHITECTURE

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

ARCHITECTURE

2 Architecture design

This project is to create an interface for the user to know their approximate Petrol Price Forecasting, in addition to this, in need of getting the real time project experience we are importing the gathered data into our own database and then start the project from the scratch.

2.1 Data gathering from main source

The data for the current project is being gathered from Kaggle dataset, the link to the data is:

<https://www.kaggle.com/c/petrol-price-forecasting/data?select=test_data.csv>

2.2 Data description

There are about 1000 record of petrol price information with date and Petrol price column.

2.3 Upload data into SQL SERVER

Created an api for the upload of the data into the SQL SERVER database, steps performed are:

* Connection is made with the database.
* Created a database with name ineuron.
* Create command is written for creating the data table with required parameters.
* And finally, a insert command is written for uploading the dataset into sale table by bulk insertion.

2.4 Export data from database

In the above created api, the download url is also being created, which downloads the data into a csv file format and upload the csvs in sql server and then through pyodbc library connection is made with python and the data is extracting in python notebook to perform next steps.

2.5 Data pre-processing

Steps performed in pre-processing are:

* First the data types are being checked and found only the sale column is of type integer.
* Checked for null values as there are few null values, those rows are dropped.
* Converted all the required column into the date time format.
* Scaling is performed for required data.
* And, the data is ready for passing to the machine learning algorithm.

2.6 Modelling

The pre-processed data is then visualized and all the required insights are being drawn. Although from the drawn insights, the data is randomly spread but still modelling is performed with different machine learning algorithms to make sure we cover all the possibilities. And finally, as expected FBPROPHET performed well and further hyperparameter tuning is done to increase the model’s accuracy.

2.7 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.

2.8 Data from user

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

2.9 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.

2.10 Rendering the results

The data sent for the forecasting is then rendered to the web page.And the Power Bi report made through SQL Server to show the results and difference between actual and forecasting price.

2.11 Deployment

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