### CRUDE OIL PRICE PREDICTION

Bonafide record of work done by

**NAVEEN KARTHI P** (1902126)

**MAHENDRA PRAKASH R (**1902158)

**MUTHU RATHIS K (**1902124)

**RITHICK M** (1902155)

**Professional Readiness for Innovation, Employability, and Entrepreneurship**

**GUIDE: SHANTHI, SELVAGANESH**

**BACHELOR OF ENGINEERING**

**BRANCH: ELECTRONICS AND COMMUNICATION ENGINEERING**

****

**NOVEMBER 2022**

**DEPARTMENT OF** **ELECTRONICS AND COMMUNICATION ENGINEERING**

**SRI RAMAKRISHNA ENGINEERING COLLEGE**

**(Autonomous Institution) COIMBATORE – 641022**

Table of Contents

[**Professional Readiness for Innovation, Employability, and Entrepreneurship** 0](#_bookmark0)

[**PSG COLLEGE OF TECHNOLOGY** 0](#_bookmark1)

[INTRODUCTION 2](#_bookmark2)

* 1. [Project Overview 2](#_bookmark3)
  2. [Purpose 2](#_bookmark4)

[This project helps the People working in the investment of crude oil needs earlier crude oil](#_bookmark5) [price prediction system, which can help them to find the right time to buy crude oil so that they](#_bookmark5) [can increase profit from the purchase and reduce any substantial loss 2](#_bookmark5)

[LITERATURE SURVEY 3](#_bookmark6)

* 1. [Existing Problem 3](#_bookmark7)
  2. [Problem Definition 3](#_bookmark8)
  3. [References 3](#_bookmark9)

[IDEATION AND PROPOSED SOLUTION 4](#_bookmark10)

* 1. [Empathy Map 4](#_bookmark11)
  2. [Ideation and Brainstorming 5](#_bookmark12)
  3. [Proposed Solution 6](#_bookmark13)
  4. [Problem Solution Fit 7](#_bookmark14)

[REQUIREMENT ANALYSIS 8](#_bookmark15)

* 1. [Functional Requirements 8](#_bookmark16)
  2. [Non-Functional Requirements 8](#_bookmark17)

[PROJECT DESIGN 9](#_bookmark18)

* 1. [Dataflow Diagram 9](#_bookmark19)
  2. [Technical Architecture 9](#_bookmark20)
     1. [Component and Technologies 10](#_bookmark21)
     2. [Application Characteristics 10](#_bookmark22)
  3. [User Stories 11](#_bookmark23)

[PROJECT PLANNING AND SCHEDULING 13](#_bookmark24)

* 1. [Sprint Planning & Estimation 13](#_bookmark25)
  2. [Sprint Delivery Schedule 14](#_bookmark26)
  3. [Reports for JIRA 14](#_bookmark27)

[CODING AND SOLUTION 15](#_bookmark28)

[7.1 Feature 15](#_bookmark29)

[TESTING 17](#_bookmark30)

[8.1 Test Cases 17](#_bookmark31)

[RESULTS 18](#_bookmark32)

[9.1 Performance Metrics 18](#_bookmark33)

[PROS AND CONS 19](#_bookmark34)

* 1. [Pros 19](#_bookmark35)
  2. [Cons 19](#_bookmark36)

[CONCLUSION 20](#_bookmark37)

[FUTURE WORKS 21](#_bookmark38)

[APPENDIX 21](#_bookmark39)

* 1. [Source Code 21](#_bookmark40)
  2. [GitHub & Project Demo Link 23](#_bookmark41)

# CHAPTER 1

* 1. **Project Overview**

# INTRODUCTION

Oil demand is inelastic, therefore the rise in price is good news for producers because they will see an increase in their revenue. Oil importers, however, will experience increased costs of purchasing oil. Because oil is the largest traded commodity, the effects are quite significant. A rising oil price can even shift economic/political power from oil importers to oil exporters. The crude oil price movements are subject to diverse influencing factors.

This Guided Project mainly focuses on applying Neural Networks to predict the Crude Oil Price. This decision helps us to buy crude oil at the proper time. Time series analysis is the best option for this kind of prediction because we are using the previous history of crude oil prices to predict future crude oil. So, we would be implementing RNN (Recurrent Neural Network) with LSTM (Long Short-Term Memory) to achieve the task.

## Purpose

#### This project helps the People working in the investment of crude oil needs earlier crude oil price prediction system, which can help them to find the right time to buy crude oil so that they can increase profit from the purchase and reduce any substantial loss

# CHAPTER 2

# LITERATURE SURVEY

## Existing Problem

Crude oil is one of the most significant sources of energy available today. With over a third of all energy consumed worldwide, it continues to be the most popular fuel. In the current environment, when technology is taking over our lives and efforts are being made to reduce the need for human labor, the Artificial Neural Network Technique has emerged as one of the most valuable techniques for data prediction. This paper offers a method for predicting oil prices that uses an artificial neural network (Sigmoid Function with the Learning Algorithm). Complex and non-linear interactions between input and output can be modeled using ANNs. The ability of ANN to generalize allows it to infer relationships even in the absence of data or input after learning from the inputs. A trustworthy method for creating predictions, ANN also learns from hidden relationships in the data without imposing any fixed relationships on the data. Many economists and analysts forecast the price of crude oil using data transformation and regression techniques like autoregressive moving average (ARMA) models and vector autoregressive (VAR) models, each time using a different input value. They then plot the graph with their forecasted prices while considering the main economic factors.

## Problem Definition

The business people who invest or works in the crude oil field and petrol bunk owners requires a way to predict the crude oil price for the next day so that they can take some major investment decisions which can lead to business profit or reduce loss.

Persons working in the investment of crude oil needs a earlier crude oil price prediction system, which can help them to find the right time to buy crude oil so that they can increase profit from the purchase and reduce any substantial loss.

To develop a system that predicts the crude oil prices using LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit) and to display the results using python-flask app.

## References

* + 1. “A comparison between AdaBoost-LSTM and AdaBoost-GRU for improving forecast prediction”. Ganiyu Adewale Busari, Dong Hoon Lim. (2021)
    2. “Crude Oil Price Prediction Using LSTM Networks”. Varun Gupta, Ankit Pandey. (2018)
    3. “Crude Oil Price Prediction using Artificial Neural Network”. Nalini Gupta, Shobhit Nigam
    4. “Crude Oil Price Forecasting based on Support Vector Machines”. Wen Xie, Lean Yu, Shanying Xu, and Shouyang Wang.
    5. LSTM. https://blog.mlreview.com/understanding-lstm-and-its-diagrams-37e2f46f1714 6. GRU. https://medium.com/geekculture/understanding-basic-architecture-of-lstm-gru- diagramm atically-6365befc64d

# CHAPTER 3

# IDEATION AND PROPOSED SOLUTION

## Empathy Map

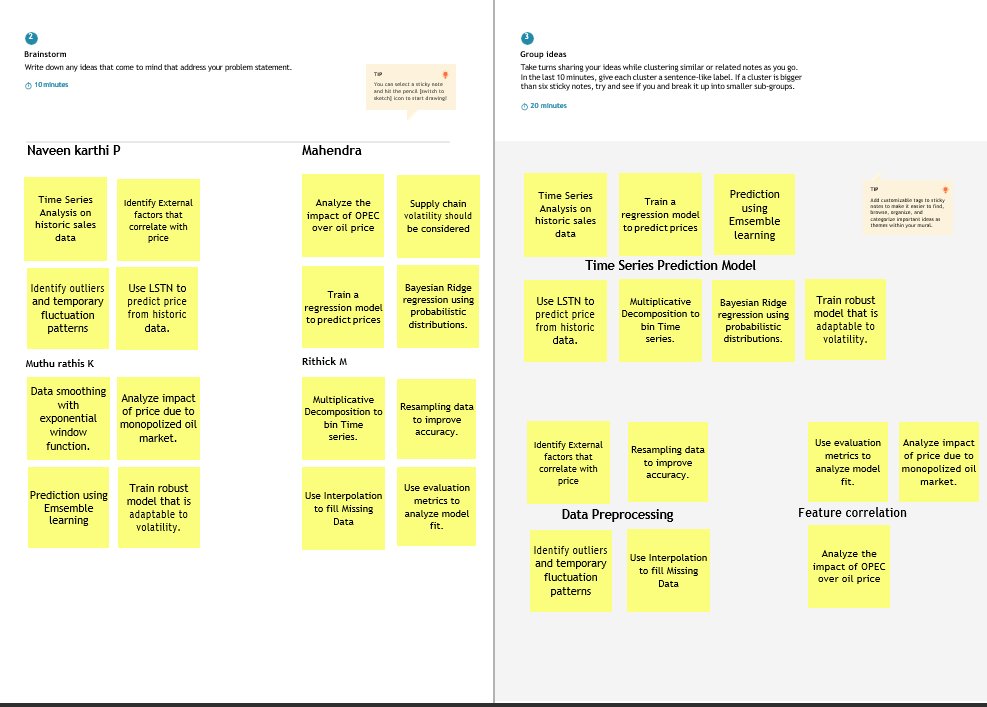
The primary purpose of the empathy map is to bridge the understanding of the user and developer. Figure 3.1 represents the empathy map for the Crude oil Price Prediction System.

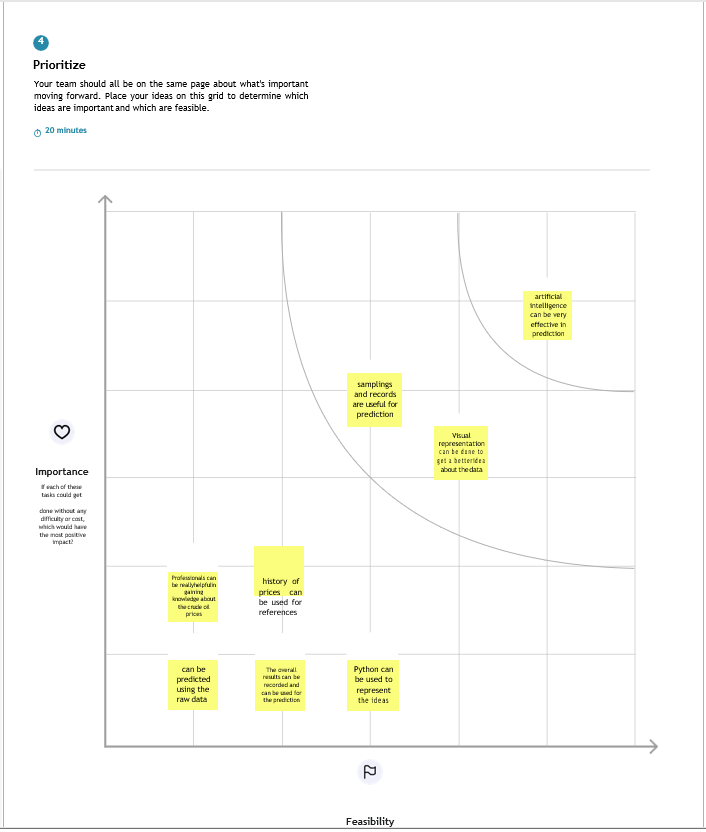


**Figure 3.1 – Empathy Map**

## Ideation and Brainstorming

This is often the most exciting stage in a project, because during Ideation and brainstorming, the aim is to generate a large quantity of ideas that the team can then filter and cut down into the best, most practical, or most innovative ones to inspire new and better design solutions and products. Figure 3.2 shows the stages of ideation and brainstorming for the Crude oil Price Prediction System.





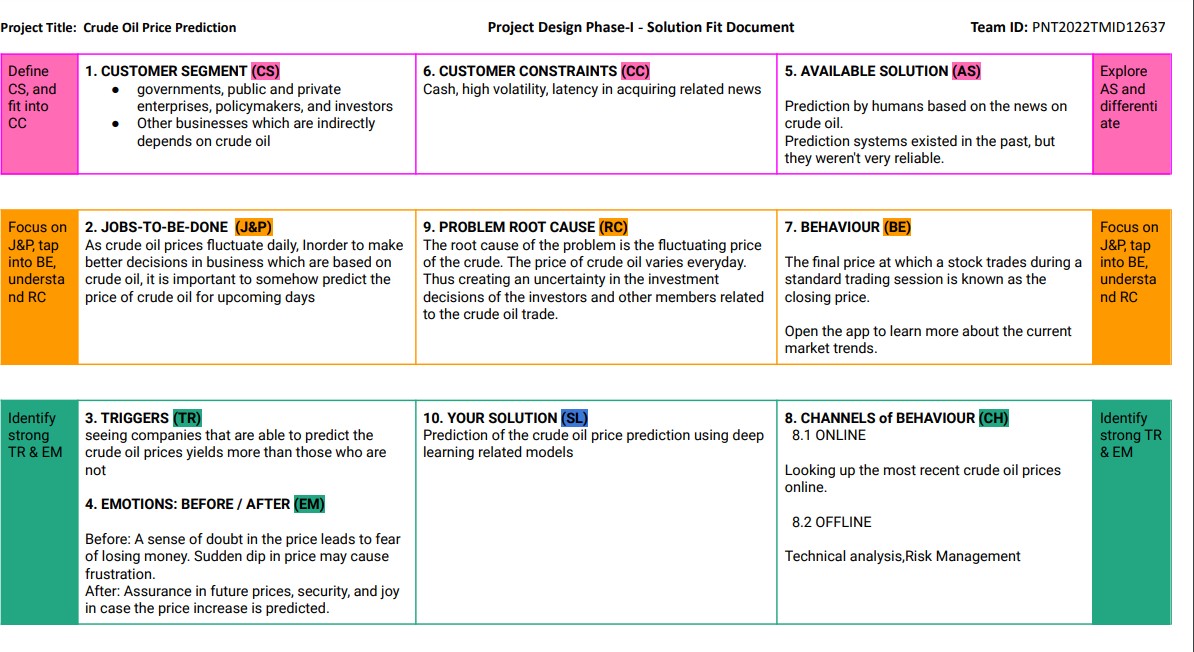
**Figure 3.2 – Ideation & Brainstorming**

## Proposed Solution

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Parameter** | **Description** |
| **1** | **Problem Statement (Problem to be solved)** | People working in the investment of crude oil needs an earlier crude oil price prediction system, which can help them to find the right time to buy crude oil so that they can  increase profit from the purchase and reduce any substantial loss. |
| **2** | **Idea / Solution description** | This solution uses, deep learning models for the prediction of crude oil price. The deep learning models that are used for time series prediction are used in this solution. Two such models are LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Units). They can be used in combination to make the prediction. |
| **3** | **Novelty / Uniqueness** | The solution tries to combine the two best time series prediction models and predict the oil price with least error. |
| **4** | **Social Impact / Customer Satisfaction** | This application will help people working in the area of investment to take better decision regarding investing on the crude oil. This will help them to predict days to buy and sell crude oil, as the prices are often variable. The customer satisfaction depends on the accuracy of the solution. The better the closeness of the predicted prices of the crude oil  to the original prices, more useful this solution becomes. |
| **5** | **Business Model (Revenue Model)** | The revenue model can be implemented as pay per month  use model. The user can pay for the service for a month. Or it can also be provided by a yearly subscription |
| **6** | **Scalability of the Solution** | The solution frontend pages are built modularly. More pages can be added as we require. In backend flask can orchestrate the different functions, so in the future new  functions and related pages can be added to the solution. |

## Problem Solution Fit

The problem solution fit is the solution one has found to address the problem of the customer. Figure 3.4 depicts the solution fit for the Crude oil Price Prediction System.



**Figure 3.4 – Problem Solution Fit**

# CHAPTER 4

# REQUIREMENT ANALYSIS

## Functional Requirements

Table 4.1 are the functional requirements of the proposed solution.

|  |  |  |
| --- | --- | --- |
| **FR**  **No.** | **Functional Requirement (Epic)** | **Sub Requirement (Story / Sub-Task)** |
| FR-1 | **Graph** | Showing Graph by obtaining the data from the Excel sheet. |
| FR-2 | **News** | Information of all oil prices will be updated by the admin |
| FR-3 | **Database** | Information of the crude oil price will be updated stored in excel sheet |

**Table 4.1 – Functional Requirements**

* 1. **Non-Functional Requirements**

|  |  |  |
| --- | --- | --- |
| **FR No.** | **Non-Functional Requirement** | **Description** |
| NFR-1 | **Usability** | It can use by wide variety of client as it is very simple to learn and not complex to proceed. |
| NFR-2 | **Security** | The information will be hashed so that it will be very secure to use. |
| NFR-3 | **Reliability** | It will be reliable that it can update with very time period so that the accuracy will be good. |
| NFR-4 | **Performance** | It will be performed fast and secure even at the lower bandwidth. |
| NFR-5 | **Availability** | Prediction will be available for every user. |
| NFR-6 | **Scalability** | we are going to use data in excel so it will be easily scalable. |

**Table 4.2 – Non-Functional Requirements**

# CHAPTER 5

* 1. **Dataflow Diagram**

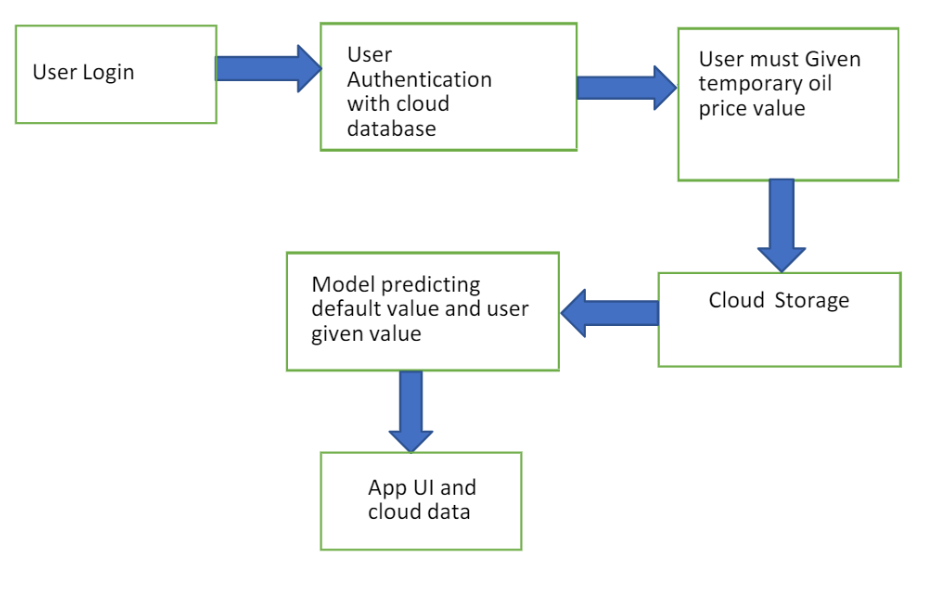
# PROJECT DESIGN

A Data Flow Diagram (DFD) is a traditional visual representation of how information flows within a system. A neat and clear DFD can thus depict the right amount of the system requirements graphically. It not only shows how data enters and leaves the system, but also what changes the information and where the data is stored. Figure 5.1 represents the DFD for the given project.

**Figure 5.1 – Dataflow Diagram**

## 

## Technical Architecture



**Figure 5.2 Technical Architecture**

### Component and Technologies

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Component** | **Description** | **Technology** |
| 1. | User Interface | How user interacts with application e.g. Web UI, Mobile App, Chatbot etc. | HTML, CSS, JavaScript  / Flask |
| 2. | Loading data | Converting the csv file to python object | Python |
| 3. | Pre-Processing  of data | Pre-Processing and normalizing the data  to get accurate results | Python |
| 4. | Cloud Database | Database Service on Cloud | IBM DB2, IBM Cloudant  etc. |
| 5. | File Storage | File storage requirements | IBM Block Storage or Other Storage Service or Local Filesystem |
| 6. | Machine Learning Model | Long short-term memory (LSTM) is an artificial neural network. Unlike standard feedforward neural networks, LSTM has feedback connections  GRU Gated recurrent units is like a long short-term memory (LSTM) with a forget gate, but has fewer parameters than  LSTM, as it lacks an output gate. | Object Recognition Model, etc |
| 7. | Infrastructure (Server / Cloud) | Application Deployment on Local System / Cloud Local Server Configuration: 2.5Ghz processor, 8GB RAM  Cloud Server Configuration: 4 GB GPU | Local, Cloud Foundry, Kubernetes, etc. |

**Table 5.2.1 – Components and Technologies**

### Application Characteristics

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Characteristics** | **Description** | **Technology** |
| 1. | Open-Source Frameworks | List the open-source frameworks used | Flask |
| 2. | Scalable Architecture | Justify the scalability of architecture (3 –  tier, Micro-services) | Cloud Foundry, IBM Cloudant |

|  |  |  |  |
| --- | --- | --- | --- |
| 3. | Availability | Justify the availability of application (e.g., use of load balancers, distributed  servers etc.) | Cloud Foundry |
| 4. | Performance | Design consideration for the performance of the application (number of requests per sec, use of Cache, use of  CDN’s) etc. | Cloud Foundry |

**Table 5.2.2 – Application Characteristics**

## User Stories

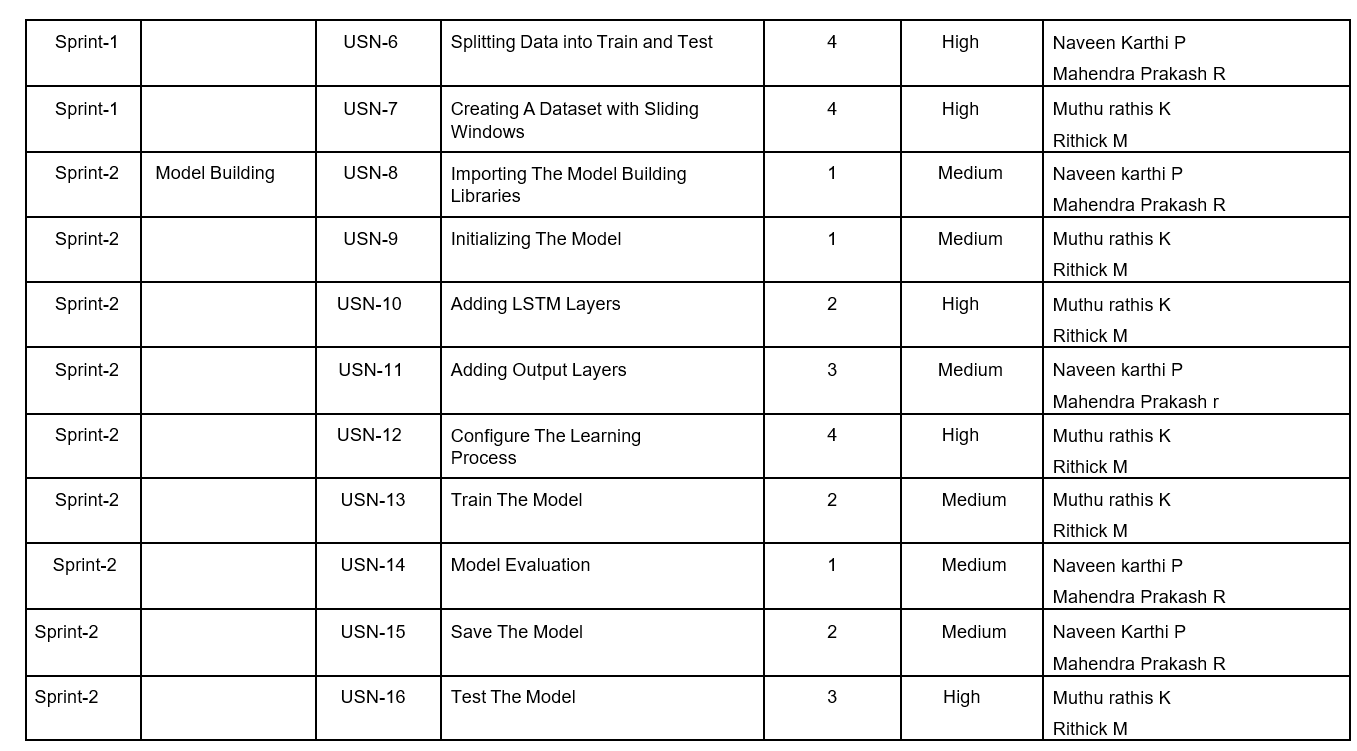
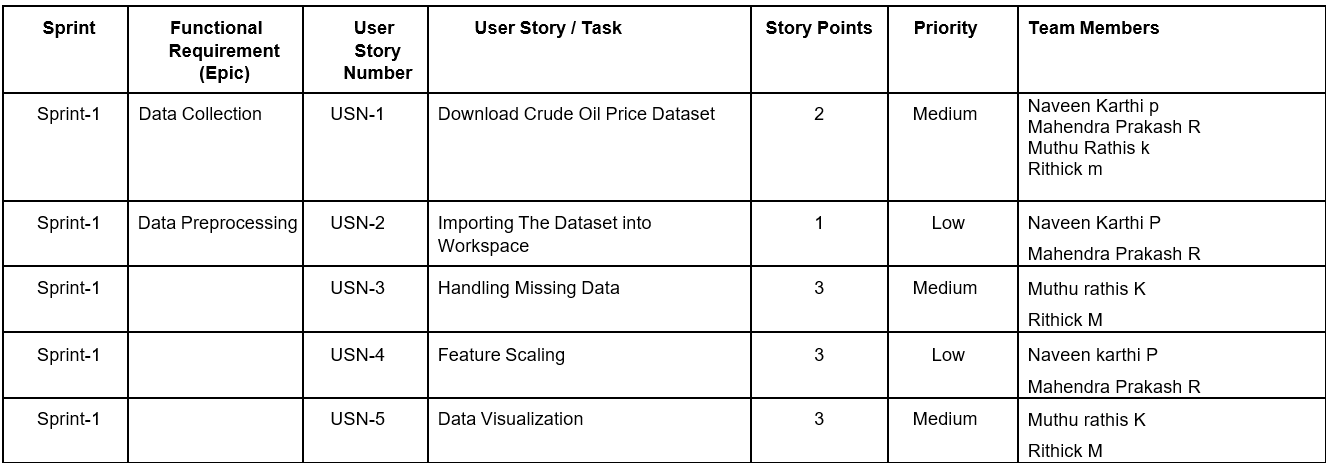
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **User Type** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** | **Acceptance criteria** | **Priority** | **Release** |
| Customer (Mobile user) | Application | USN-1 | You can download the crude oil price by opening the Google Play Store app directly as a user. | I can access own decisions. | High | Sprint-1 |
|  | Available Products | USN-2 | Users of the application may instantly update the energy and oil prices while using it  because there are so many different products in the crude oil price app. | I can receive the data once click then confirm | High | Sprint-1 |
|  | Additional Features | USN-3 | Users can read the most recent news and see oil price charts.  Major Energy Quotes User View  The user may use many colour schemes. | I can view then read the price prediction. | High | Sprint-2 |
|  | [Expectations](https://developer.ibm.com/patterns/visualize-unstructured-text/) | [USN-4](https://developer.ibm.com/patterns/visualize-unstructured-text/) | User Can Convert Currency And Exchange Rates | I can expect | Medium | Sprint-1 |
|  | [Login](https://developer.ibm.com/patterns/visualize-unstructured-text/) | [USN-5](https://developer.ibm.com/patterns/visualize-unstructured-text/) | Log in as a user without using your email address, username, or password. |  | High | Sprint-1 |
|  |  |  |  |  |  |  |
| [Customer (Web](https://developer.ibm.com/patterns/visualize-unstructured-text/) [user)](https://developer.ibm.com/patterns/visualize-unstructured-text/) |  |  | I can see the price of crude oil as a consumer. | I can view the price directly | High | Sprint-1 |
| [Customer Care](https://developer.ibm.com/patterns/visualize-unstructured-text/) [Executive](https://developer.ibm.com/patterns/visualize-unstructured-text/) |  |  | I am the user and I executive the pricing  history. | I can accept the terms | medium | Sprint-1 |
| [Administrator](https://developer.ibm.com/patterns/visualize-unstructured-text/) |  |  | As a manager, it anticipates the results. | Show the result | High | Sprint-1 |

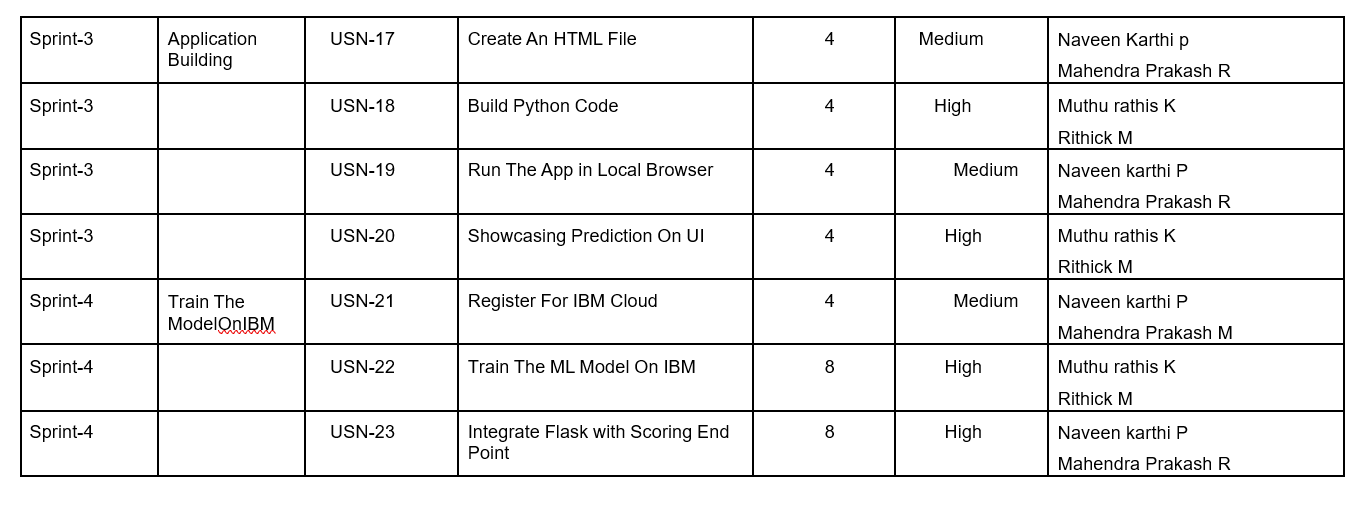
**Table 5.3 – User Stories**

# CHAPTER 6

# PROJECT PLANNING AND SCHEDULING

## Sprint Planning & Estimation





**Table 6.1 – Sprint Planning**

## Sprint Delivery Schedule

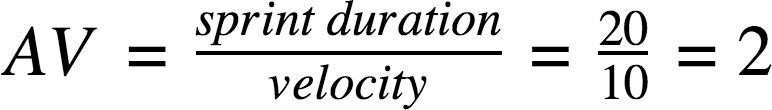
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Total Story Points** | **Duration** | **Sprint Start Date** | **Sprint End Date (Planned)** | **Story Points Completed (as on Planned End Date)** | **Sprint Release Date (Actual)** |
| Sprint-1 | 20 | 6 Days | 24 Oct 2022 | 29 Oct 2022 | 20 | 29 Oct 2022 |
| Sprint-2 | 20 | 6 Days | 31 Oct 2022 | 05 Nov 2022 | 20 | 05 Nov 2022 |
| Sprint-3 | 20 | 6 Days | 07 Nov 2022 | 12 Nov 2022 | 20 | 12 Nov 2022 |
| Sprint-4 | 20 | 6 Days | 14 Nov 2022 | 19 Nov 2022 | 20 | 19 Nov 2022 |

**Table 6.2 – Sprint Delivery Schedule**

## Reports for JIRA

**Velocity:**

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let’s calculate the team’s average velocity (AV) per iteration unit (story points per day)



**Burndown Chart:** A burndown chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.

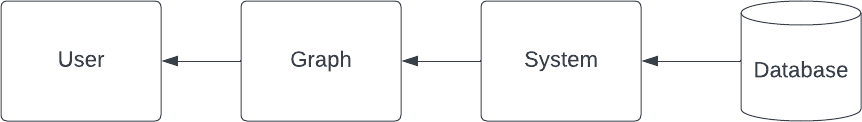


# CHAPTER 7 CODING AND SOLUTION

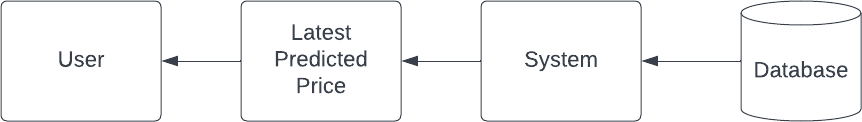
## 7.1 Feature

|  |  |  |
| --- | --- | --- |
| **FR**  **No.** | **Feature** | **Description** |
| FR-1 | Crude oil Price Graph | Showing the price of crude oil for respective dates in a graph with dates in x axis and crude oil prices on y axis |
| FR-2 | Current Price Prediction | Showing the last predicted price of the crude oil |
| FR-3 | Prediction based on user provided values | When user provided with the three days prices of the crude oil the application will give predicted price for the next day |

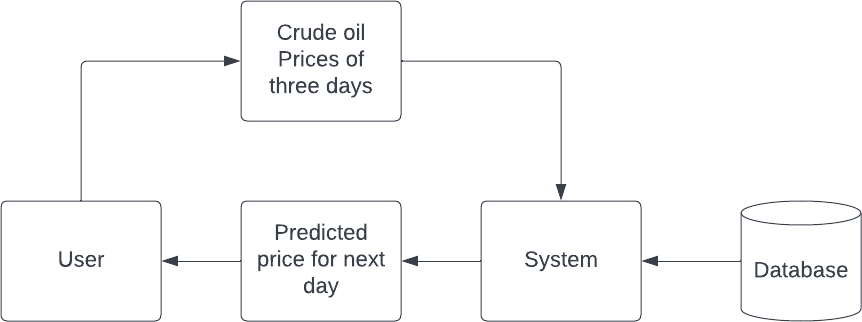
**Table 7.1 – Description for Feature**



**Figure 7.1 – Dataflow Diagram for Feature 1**

****

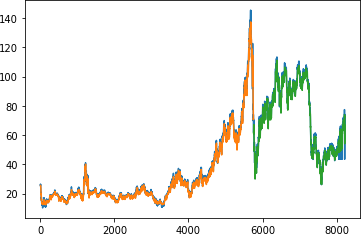
**Figure 7.1 – Dataflow Diagram for Feature 2**

****

**Figure 7.1 – Dataflow Diagram for Feature 3**

**8.1 Test Cases**

# TESTING

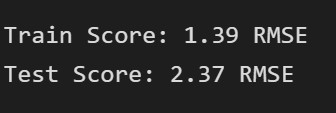
The test cases are window of closing prices, where the window size is 3. The test cases are sent to the model and the prediction is compared with the original closing price. The loss metric is used to analyze the performance of the model. Figure 8.1 shows the result after the testing. The blue line in the bottom shows the true closing prices. The orange lines denote the prediction using the training data. The green line denotes the prediction based on testing data.

**Figure 8.1 – Test Cases Run**

**9.1 Performance Metrics**

# RESULTS

The performance metric used to measure the model is RMSE (Root Mean Square Error). RMSE is measured on both the prediction on training data and the testing data. Lower the RMSE score, better is the accuracy of the model. The results of the RMSE are given in the figure 9.1. From the figure 9.1 it can be observed that the RMSE are quite low. This indicates the model is working better and the predictions are quite accurate.



**Figure 9.1 – Performance Metrics**

* 1. **Pros**

# PROS AND CONS

* + - The application we have created is user friendly
    - This application is flexible as user can choose the way they need to predict the price
    - User can either give crude oil prices of any three continues dates or can get the latest predicted price of crude oil

## Cons

As the data used in this project is not up to date the prices the model predicted will not be applicable for using in real world crude oil price prediction

# CHAPTER 11

# CONCLUSION

The prediction system works using the model that is built by combination of LSTM and GRU. The RMSE score for both the training and testing data is quite low. This shows that the accuracy of the model is good. A website is served using flask framework, which helps to enable the users to interact with the model. It helps the user to see the current predicted price the crude oil. And it helps to do prediction for manually entered crude oil closing price values.

# CHAPTER 12

# FUTURE WORKS

The model currently cannot update the prices to the current data automatically. Web automation can be enables to let the system update its database to current prices. And the model can be retrained on the updated data.

# CHAPTER 13

# APPENDIX

## Source Code

App.py

|  |  |
| --- | --- |
|  |  |
|  | from flask import Flask,render\_template,request,redirect |
|  | import joblib |
|  | from keras.models import load\_model |
|  |  |
|  | app = Flask(\_\_name\_\_) |
|  |  |
|  | @app.route('/',methods=["GET"]) |
|  | def index(): |
|  | return render\_template('index.html') |
|  |  |
|  | @app.route('/predict',methods=["POST","GET"]) |
|  | def predict(): |
|  | if request.method == "POST": |
|  | string = request.form['val'] |
|  | if(string ==""): |
|  | return render\_template('predict.html') |
|  | string = string.split(',') |
|  | x\_input = [eval(i) for i in string] |
|  | sc = joblib.load("scaler.save") |
|  | x\_input = sc.fit\_transform(np.array(x\_input).reshape(-1,1)) |
|  | x\_input = np.array(x\_input).reshape(1,-1) |
|  | x\_input = x\_input.reshape(1,-1) |
|  | x\_input = x\_input.reshape((1,10,1)) |
|  | model = load\_model('model.h5') |
|  | output = model.predict(x\_input) |
|  | val = sc.inverse\_transform(output) |
|  | return render\_template('predict.html' , prediction = round(val[0][0],2)) |
|  | if request.method == "GET": |
|  | return render\_template('predict.html') |
|  |  |
|  | if \_\_name\_\_=="\_\_main\_\_": |
|  | model = load\_model('model.h5') |
|  | app.run(debug=True) |

## GitHub & Project Demo Link

GitHub link: https://github.com/IBM-EPBL/IBM-Project-22841-1659859096

Project Demo Link: [https://drive.google.com/file/d/1JPQF0KSuGBJHQxoCcNTXRmldGj7fL0xm/view?usp=sharing](https://drive.google.com/drive/folders/1nHufpbcAsPpI1zr5xqFod9Qf2brMkKq4?usp=share_link)