# CRUDE OIL PRICE PREDICTION USING ARTIFICIAL INTELLIGENCE

**NAALAIYA TIRAN PROJECT BASED LEARNING ON PROFESSIONAL**

**READINESS FOR INNOVATION, EMPLOYABILITY**

**AND**

**ENTREPRENEURSHIP**

**A PROJECTREPORT**

| | |
|---|---|
| **DHYALAN C** | **611819106009** |
| **DINESH KUMAR S** | **611819106010** |
| **SATHISH G** | **611819106046** |
| **JAYAKUMAR S** | **611819106013** |

**TEAM ID**: PNT2022TMID40932

**FACULTY MENTORS NAME**: PRAKASAM L

**INDUSRTY MENTORS NAME**: SHANTHI

**EVALUATOR NAME:** SHANMUGAM S

## P.S.V. COLLEGE OF ENGINEERING AND TECHNOLOGY

**(An ISO 9001:2015 Certified Institution)**
**(Accredited by NAAC with ‘A’ Grade)**

**KRISHNAGIRI-635108**

**NOVEMBER 2022**

# ANNA UNIVERSITY: CHENNAI 600 025

## BONAFIDE CERTIFICATE

**Certified that this project report “CRUDE OIL PRICE PREDICTION USING ARTIFICIAL INTELLIGENCE” is the bonafide work of “DHYALAN C (611819106009), DINESH KUMAR (611819106010), SATHISH G (611819106046) and JAYAKUMAR (611819106013)” who carried out the project work under my supervision.**

| | |
|---|---|
| **SIGNATURE** | **SIGNATURE** |
| **Prof. C. THAMILARASI,M.E.,** | **Prof. L.PRAKASAM.,M.E.,** |
| **Head of the Department** | **Faculty Mentor,** |
| Dept of ECE, | Dept of ECE, |
| P.S.V.College of | P.S.V.College of |
| Engineering & Technology, | Engineering & Technology, |
| Krishnagiri-635 108. | Krishnagiri-635 108. |

Submitted for the **Naalaiya Tiran Project based learning on professional readiness for Innovation, Employability and Entrepreneurship** held on ………………….……………….at P.S.V College of Engineering and Technology, Krishnagiri.

| | |
|---|---|
| **INTERNAL EXAMINAR** | **EXTERNAL EXAMINAR** |

# TABLE OF CONTENTS

# ABSTRACT

Crude oil is the world's most leading fuel. The main advantages of crude oil are it has high density, it is easily available. Oil is used in almost all the industries. Oil is a Constant Power Source. The main aim of this project is to find the different models that efficiently fit the data points and predict the price of fuel with the help of machine learning models. This project works on comparing the different supervised learning models and brings a conclusion based on the efficiency. We have used three supervised learning models namely Random Forest Regression, Linear Regression and Decision Tree Regression to know which gives best in terms of accuracy and performance. These algorithms give a numeric value as output. So we can compare the output of these models with the actual models. Now-a-days the oil price has been increasing in leaps and bounds due to certain reasons like inflation worldwide. Hence these are derived or extracted from petroleum. To predict the values of petroleum like petroleum and Diesel within the future, we've decided to use the Machine Learning algorithms. We use performance metrics to find the performance of the supervised learning models based on their errored value. In this way we can compare different algorithms and find the best one for our problem statement.

*Keywords: Python, Algorithm, Price,Deep learning, Data, Prediction*

# LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1

# INTRODUCTION

## 1.1 Project Overview

Owing to the fact that crude oil provides around one-third of the world's energy needs, crude oil is important to the global economy. Additionally, changes in oil prices have a big impact on both countries' economies that export and buy oil. Forecasting the oil price accurately would assist policymakers in enacting the right legislation and selecting the best energy sources. However, because there are numerous factors that affect oil prices, forecasting researchers have found it difficult to estimate the price of crude oil. Economic growth, conflicts, wars, and breaking news all have a significant impact on oil price fluctuations additionally the basic market elements like supply, demand, and inventory. For instance, oil producers were paying buyers to take the commodity off their hands because they were concerned that storage space might be depleted in May 2020. On April 20, 2020, the price of WTI oil even became negative for the first time ever. Another recent example is the higher association between changes in crude oil prices and the severity of the COVID-19 epidemic. Since the majority of this information is found in unprocessed texts, characterizing and modelling these nonlinear and non quantitative factors is difficult.

## 1.2 Purpose

The three primary factors that impact the price of oil are:

- **Supply and demand**

The idea of supply and demand is rather simple.Price should rise as demand(or supply) rises or falls. Price should decrease when supply grows or as demand declines. Actually, the oil futures market is where the price of oil as we know it is set. A legally binding agreement known as an oil futures contract offers one the right to buy oil by the barrel at a specified price on a specified date in the future. In a futures contract, each party handles carrying out their portion of the deal before the deadline.

- **Cost of production**

Cost of production refers to the total cost incurred by a business to produce a specific quantity of a product or offer a service. Production costs may include things such as lab our, raw materials, or consumable supplies. In other words, the cost of production is defined as the expenditures incurred to obtain the factors of production such as lab our, land, and capital, that are needed in the production process of a product.

- **Market sentiment**

Sentiment is the other important factor that impacts oil prices. The simple expectation that oils demand would rise sharply at some point in the future can cause speculators and hedgers to buy up oil futures contracts, driving up oil prices now.

There used to be a recognizable seasonal swing in oil prices. As oil dealers anticipated a large demand for driving over the summer vacation, they increased in the spring. Prices fell in the fall and winter once the demand peaked.

Geopolitical instability and civil upheaval also have a significant impact on global supply and prices.

There are several reasons why oil prices are more unpredictable now, but five are the most significant.

- **The RussianInvasion of Ukraine**

Russia is the third-largest producer of liquid fuels and petroleum, so when the country invaded Ukraine in late February2022, it had immediate impact on Brent crude oil futures prices.10 As the conflict continued, the prices of crude oil settled in out on an upward trajectory, reaching nearly $130/b in early March, and staying well above $100/bin to April.

- **US Oil Supply**

The coronavirus pandemic and natural events are still affecting oil demand and supply.The U.S. experienced a drop in production followingHurricane Ida in September as the storm shut at least nine refineries.

The EIA estimates that U.S. crude oil production will average 12.01 million b/din 2022 and 12.95 million b/d in 2023.11

- **Diminished OPEC Output**

Oil price increases also reflect supply limitations by the Organization of the Petroleum Exporting Countries (OPEC) and OPEC partner countries. In 2020, OPEC cut oil production due to decreased demand during the pandemic. It gradually increased oil output through 2021 and into 2022. Supply chain disruptions in late 2021affected global trade a swell.

At its most recent meeting in December2021, OPEC stated it would continue to gradually adjust oil production upward by 0.4 million barrels per day (mb/d) in January 2022.

- **Natural Gas**

Countries in Asia have relied on coal to generate power, but recent shortages have turned them to natural gas. Higher temperatures in parts of Asia and Europe have led to high demand for natural gas to generate power.

COVID-19 has hampered Europe's natural gas production, and a colder-than- expected heating season in early 2021 reduced supplies further.

As a result, natural gas prices soared in 2021 and are expected to remain high in 2022 and affected countries have turned to gas-to-oil switching to reduce power generation costs.

- **Global InventoryDraw**

As a reduction in oil production continues globally, countries are forced to draw from their stored reserves (not including the strategic petroleum reserves). This steady draw of oil is contributing to the increase in prices because inventories are decreasing.

Models incorporating economic parameters such as supply, and demand and their determinants are known as structural models. Even though structural models are found to be the most logical ways of modelling the prices of industrial products, the price of crude oil is affected by many other factors. One of these factors is that the price of crude oil is determined in the futures market which enables the purchase of a predefined amount of oil at a particular price in the future. Additionally, only 1 percent of the crude oil traded in futures contracts results in the actual purchase of a physical commodity; its chief purpose is to make money out of price fluctuations in crude oil. Hence the price of crude oil behaves more like a financial asset and therefore is more representative of the expectations of traders rather than just predictions based on economic theories of supply and demand.

There are other categories of models which are non-structural and consider time variation of crude oil prices, known as time series models. It is difficult to obtain reliable data to formulate a structural model, while time series data for crude oil prices is easily available and hence it is easier to build a time series model. We focus on time series modelling of crude oil prices in this article.

In time series models, it is assumed that the current price of crude oil reflects the effects of all influencing factors, and that price forecasting can be done based on the behavior of past crude oil prices. The main assumption in such models is that the past behavioral oil prices can explain future prices. Although time series models can capture trends or any cyclical patterns in the data, there are limitations to the forecasting capability of these models when trend reversals are observed in the data, or the repeating pattern captured in the model is not followed in future prices.Different trends in a time series can be classified as increasing, decreasing and periodic patterns. Time series models are quite useful and forecast reasonably well when the data follows any of these types of trends.

We can easily observe the downtrends, uptrends and repeating patterns in crude oil prices within specific years. Crude oil monthly price data is obtained from the USEnergy Information Administration (EIA) website.1 Different subsets of crude oil price data are formed to demonstrate the utility of time series modelling and its limitations in some scenarios.

## Time Series Modelling Techniques

Several methods are proposed in the literature to build time series models.They include auto regressive integrated moving average (ARIMA), generalized auto regressive conditional heteroscedastic (GARCH), Holt-Winters, autoregressive neural networks, and support vector regression.2 Various hybrid models are also suggested such as combination of ARIMA and neural networks with support vector regression, genetic algorithms and wavelets.3-7 Discussion of various methodologies applied for crude oil price modelling can be found in review articles available in the literature.8,7 We have used ARIMA and auto regressive neural networks for modelling oil prices, as these techniques cover both linear and non-linear types of modelling. A short description of these methods is given below.

## ARIMA

ARIMA is the most widely used and well-known technique for time series analysis,developed by Box and Jenkins. In an ARIMA model, future values are predicted as a linear combination of previous oil prices and the associated errors. This model consists of three parts: the AR (auto regressive) component is a linear combination of past observations; MA (moving average)is a linear combination of laggederror terms; and I (integrated) replace the original series with differenced series.

## Auto regressive NeuralNetwork

An autoregressive neural network (ANN) is a non-linear model in which future prices are expressed as a non-linear function of lagged prices in the series, in contrast to linear modelling in ARIMA. Additionally, neural network-based models can learn and capture patterns in data sets without the need to specify the exact model form. Multilayer perceptron (MLP) is the most widely used ANN in forecasting problems. Typically, the model is composed of input layer, hidden layer and output layer. The connecting nodes in these layers are called neurons. Input to the neurons is mapped using transfer functions and the weighted average of output from all the nodes is sent to next layer. There are variousparameters that need to be specified for an ANN model: number of hidden layers, number of neurons in each layer, type of transfer function,and number of lags. The selection of appropriate networkparameters is crucial to the fitting and forecast accuracy of an ANN model. We have used the nnetar functionin R to build a neural network model.

## Benefits of predicting crude oil prices:

1. Some Sectors Thrive It probably counts as obvious that there are sectors that thrive when oil prices march upward. High prices for oil fuel the same sort of process as in any other sector; suppliers look for ways to provide more of the product and take advantage of those higher prices. For energy, then, that means opportunities for companies involvedin exploration (seismicsurvey, for instance), drilling, production and servicing.
2. New Technologies Become Viable Inexpensive oil is problematic for companies and industries looking to supplant oil. While most people can agree that there are vague and nebulous costs associated with accessing and utilizing oil (pollution, for starters), the United States has been reticent to translate those costs into higher energy taxes. What's more, it is not clear that higher taxes on fossil fuels in Europe and much of Asia really do anything to mitigate environmentaldamage beyond reducing consumption. All in all, then, when oil prices are low it is very hard for cleanerenergy technologies to compete effectively on price.

3. Changes in Behaviour For those who believe that burning oil (and other hydrocarbons) is generally a bad thing, higher prices that lead to lower use has to be counted as a benefit. When people are faced with higher prices and no obvious substitutes, they will consumeless assuming that their demand is relatively elastic.

4. Alternatives Come to the Fore If increased exploration and production is a normal by-product of higher oil prices, so too is substitution. When NaziGermany faced oil shortages in World War II, methods of producing oil, diesel and gasoline substitutes from vegetable oils, animal fats and coal

were thoroughly explored. Likewise, the oil crisis of the 1970s gave the developmentof ethanol in Brazil a majorboost.

# CHAPTER 2

# LITERATURE SURVEY

## 2.1 EXISTING PROBLEMS:

The existing problem can be broadly classified into the following

i. Predictive Analysis

ii. Determining the Crude Oil Price

iii. Neural Network for Predictive Analytics

iv. RNNLSTM Network

### A. Predictive Analytics

Predictive analytics is a cutting-edge field of study that employs statistical models and other scientific methods to assess hazy future opportunities to producing actual forecasts and verifying the accuracy of these forecasts in the real world [2]. The predictive analytics model can provide meaningful insights by extracting knowledge from data and use statistical or machine learning methods to assist with the analytical task.

### B. Determining the Crude Oil Price

Various significant elements, including a supply and demand curve, the present financial market, the commodities market, speculative factor, and geopolitical factor, may have an impact on fluctuations in crude oil prices, according to Miao et al. [3]. Each of these variables has a number of determining factors (sub-variables) that impact the priceof the commodity.

According to an article published on the Caltex website [4], the fuel (such as petrol) prices change is closely related to the cost of crude oil–and it has a long-term effect on the fluctuation of the commodityprice. Additionally, the cost of crude oil alonehas contributed to nearly 50 percent of the retailpetroleum price [4].

## C. Neural Network for Predictive Analytics

The neural networkcontains a set of neurons(or perceptron’s) which acts as processing units [5], interlinked, and may residewithin an extensivenetwork.
The most basic form of the neuralnetwork consists of an input layer, one hidden layer, and an output layer [6], as visualized in Figure 1. The number of hidden layers mayvary based on the complexity of computation.

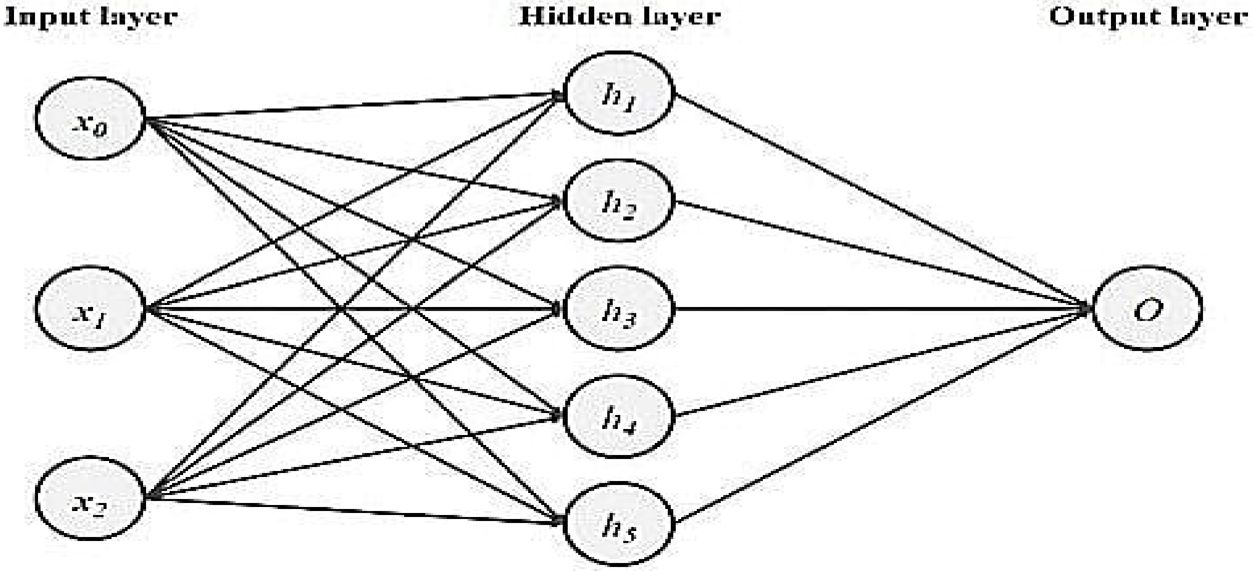


Figure 1 - A neuralnetwork

## D. RNN-LSTM Network

Traditional neural network techniques function well for applications requiring prediction, but they cannot store memories. On the other hand, the Recurrent Neural Network (RNN) is a section of a neural network that has been converted into a loop, providingit the ability to retain knowledge from its previousstate.

Hochreiter & Schmidhuber [7] have introduced the concept of Long-Short Term Memory (LSTM), which has proven its accuracy across various domains [7]. LSTM is a type of Recurrent Neural Network (RNN) that can learn long-termdependencies and is useful for a sequence-to-sequence prediction–such as prediction of upcoming crude oil pricesusing time-series data.

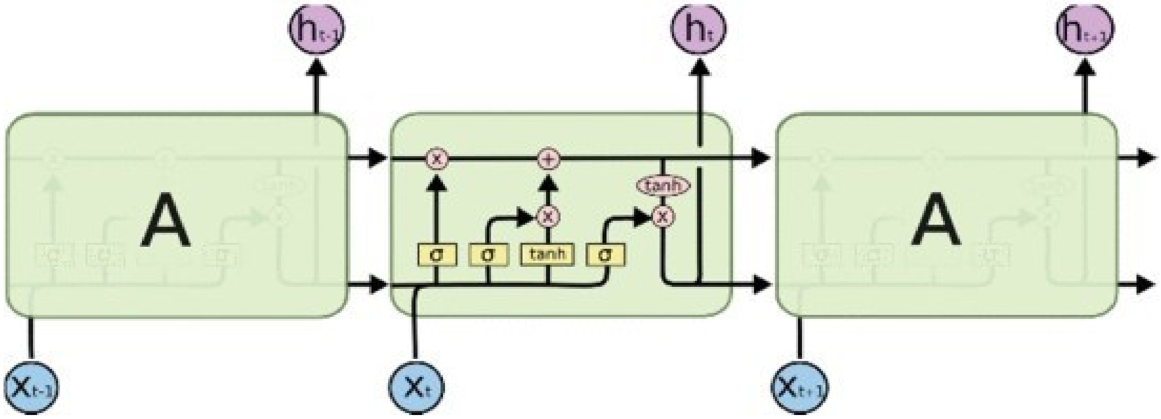


Figure 2 - The RNN-LSTMarchitecture

In our project “Crude Oil Price Prediction”, we proposed a solution which usesthe RNN LSTM method to solve the existing problem. Time series analysis algorithm is used to combine all the advantages of the above methods and to remove some of the disadvantages discussed in the above methods. Time series analysis is a specificway of analyzing a sequence of data points collected over an interval of time. In time series analysis, analysts record data points at consistent intervals over a set period of timerather than just recording the data points intermittently or randomly. This model is also trained using the Long Short Term Memory method in the Recurrent Neural Networkalgorithm which would have a greater efficiency.

## 2.2 References

1. N. Aziz, M. H. A. Abdullah and A. N. Zaidi, "Predictive Analytics for CrudeOil Price Using RNN-LSTM Neural Network," 2020 International Conference on Computational Intelligence (ICCI),2020, pp. 173-178,doi: 10.1109/ICCI51257.2020.9247665.

2. G. Shmueli and O. R. Koppius, "Predictive analytics in information systems research", *MIS Q*, pp. 553-572, 2011.

3. H.Miao, S. Ramchander, T. Wang and D. Yang, "Influential factors in crude oilpriceforecasting", *Energy Econ*, vol. 68, pp. 77-88, 2017.

4. "Determining Fuel Prices", Mar 2020, [online] Available:https://www.caltex.com/my/motorists/tips-resources/determining-fuel- prices.html.

5. M.H. A. Abdullah, "Comparative Analysisof Spatio/Spectro-Temporal

DataModelling Techniques", *Data Engineering And Information Security*, 2017.

6. M. H. A. Abdullah, M. Othman, S. Kasim and S. A. Mohamed, "Evolving Spiking NeuralNetworks Methods for Classification Problem : A Case Study in Flood Events Risk Assessment Evolving spiking neural networks methodsfor classification problem: a case study in flood events risk assessment", *Indones.J. Electr. Eng. Comput. Sci*,vol. 16, no. 1, pp. 222-229, 2019.

7. S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Comput*,vol. 9, no. 8, pp.

## 2.3 Problem Statement Definition

The price of crude oil has a significant impact on the environment globally,and its forecasts are particularly helpful to governments and industry. Crude oil is the most widely used fuel in the world. The ongoing application of statistics and econometric methods for crude oil, including AI Price forecasting could show reductions in the accuracy of the prediction.

In order to predict future crude oil using historical data on crude oil, RNN (Recurrent Neural Network) is utilised with long short-term memory. The effectiveness of the cost is calculated using the denote squared error. Using the pricing information in the crude oil materials, the proposed model'sperformance is assessed.

Since changes in the price of crude oil have a significant impact on national economies around the world, price forecasting can help reduce the risks brought on by oil pricevolatility.

Governments, public and private businesses, legislators, and investorsall placea high valueon price estimates.

**The project“Crude Oil Price Prediction”, has the following uniqueness and novelty:**

1. Thismodel is used to forecastfuture pricing and to manage oil use.
2. This price directly influences many different items and goods, and its variationshave an impacton the capital markets.
3. Important eventsalso have an impact on oil prices,in addition to economicfactors.

**The project “Crude Oil Price Prediction”, has the following business model:**

4. Itcan assist those who are making decisionsabout whether to buy or sell crudeoil, whether they are businesses, private investors, or individuals.
5. The benchmark model for predicting crude oil pricesuses RNN and LSTMmodels.

**The scalability of the solutionof this projectare:**

6. The dimensionsof the data are reducedusing the PCA, MDS, and LLE methods.
7. Enhance the RNN and LSTM models'accuracy.

# CHAPTER 3

# IDEATION AND PROPOSED SOLUTION

## 3.1 Empathy Map Canvas

An empathy map canvas is a more in-depth version of the original empathy map, which helps identify and describe the user’s needs and pain points. And this is valuable information for improving the user experience.

Teams rely on user insights to map out what is important to their target audience, what influences them, and how they present themselves. This information is then used to create personas that help teams visualize users and empathize with themas individuals, rather than just as a vague marketing demographic or account number.

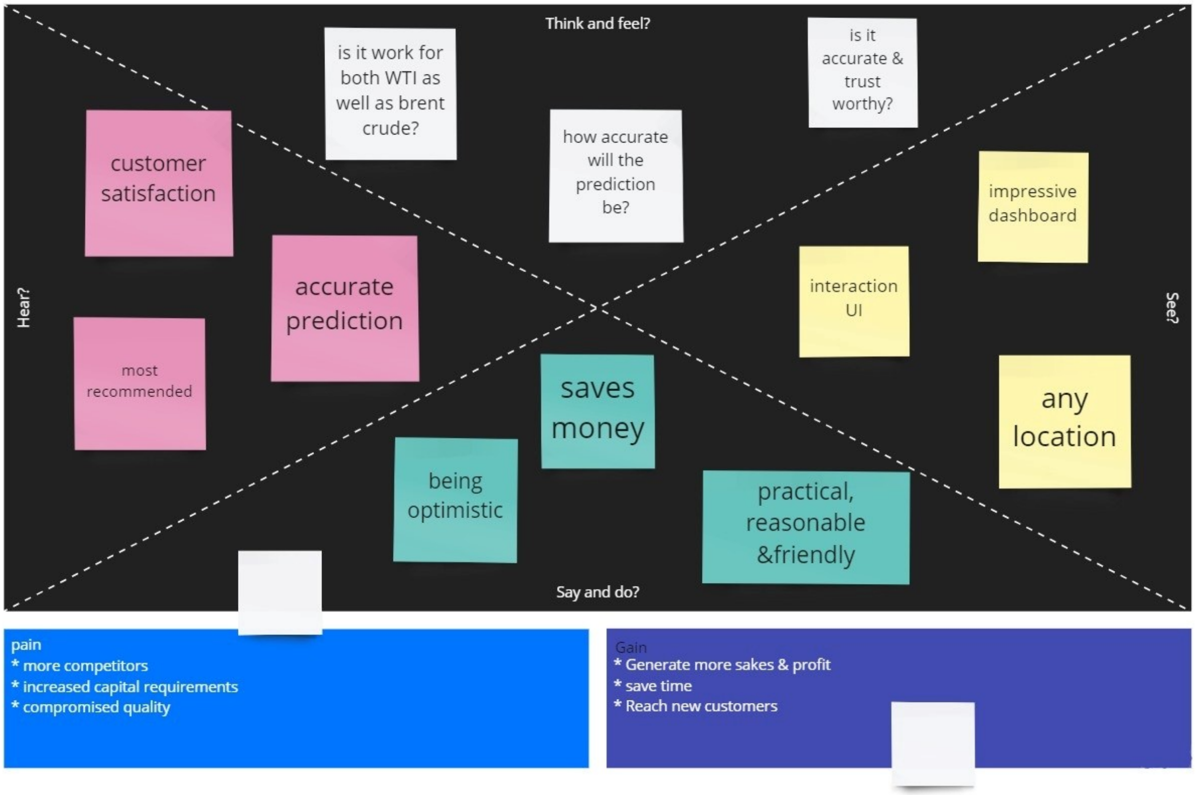
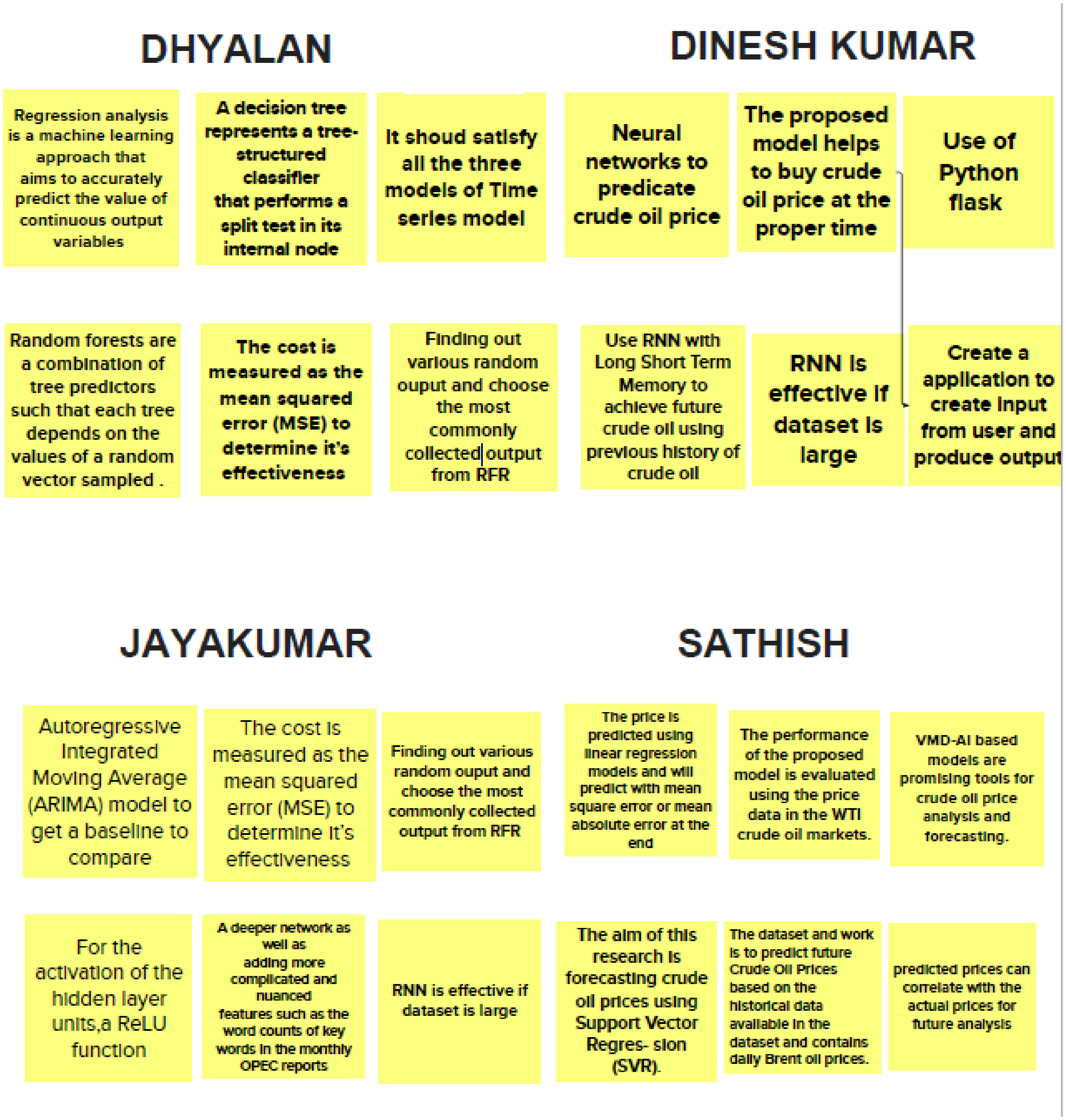


Figure 3 – Empathy Map Canvas

## 3.2 Ideation & Brainstorming

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich number of creative solutions.

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you are not sitting in the same room.



DHYALAN

Regression analysis is a machine learning approach that aims to accurately predict the value of continuous output variables

A decision tree represents a tree-structured classifier that performs a split test in its internal node

It shoud satisfy all the three models of Time series model

Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled .

The cost is measured as the mean squared error (MSE) to determine it's effectiveness

Finding out various random ouput and choose the most commonly collected output from RFR

DINESH KUMAR

Neural networks to predicate crude oil price

The proposed model helps to buy crude oil price at the proper time

Use of Python flask

Use RNN with Long Short Term Memory to achieve future crude oil using previous history of crude oil

RNN is effective if dataset is large

Create a application to create input from user and produce output

JAYAKUMAR

Autoregressive Integrated Moving Average (ARIMA) model to get a baseline to compare

The cost is measured as the mean squared error (MSE) to determine it's effectiveness

Finding out various random ouput and choose the most commonly collected output from RFR

For the activation of the hidden layer units,a ReLU function

A deeper network as well as adding more complicated and nuanced features such as the word counts of key words in the monthly OPEC reports

RNN is effective if dataset is large

SATHISH

The price is predicted using linear regression models and will predict with mean square error or mean absolute error at the end

The performance of the proposed model is evaluated using the price data in the WTI crude oil markets.

VMD-AI based models are promising tools for crude oil price analysis and forecasting.

The aim of this research is forecasting crude oil prices using Support Vector Regres- sion (SVR).

The dataset and work is to predict future Crude Oil Prices based on the historical data available in the dataset and contains daily Brent oil prices.

predicted prices can correlate with the actual prices for future analysis

Figure 4 - Brainstorm

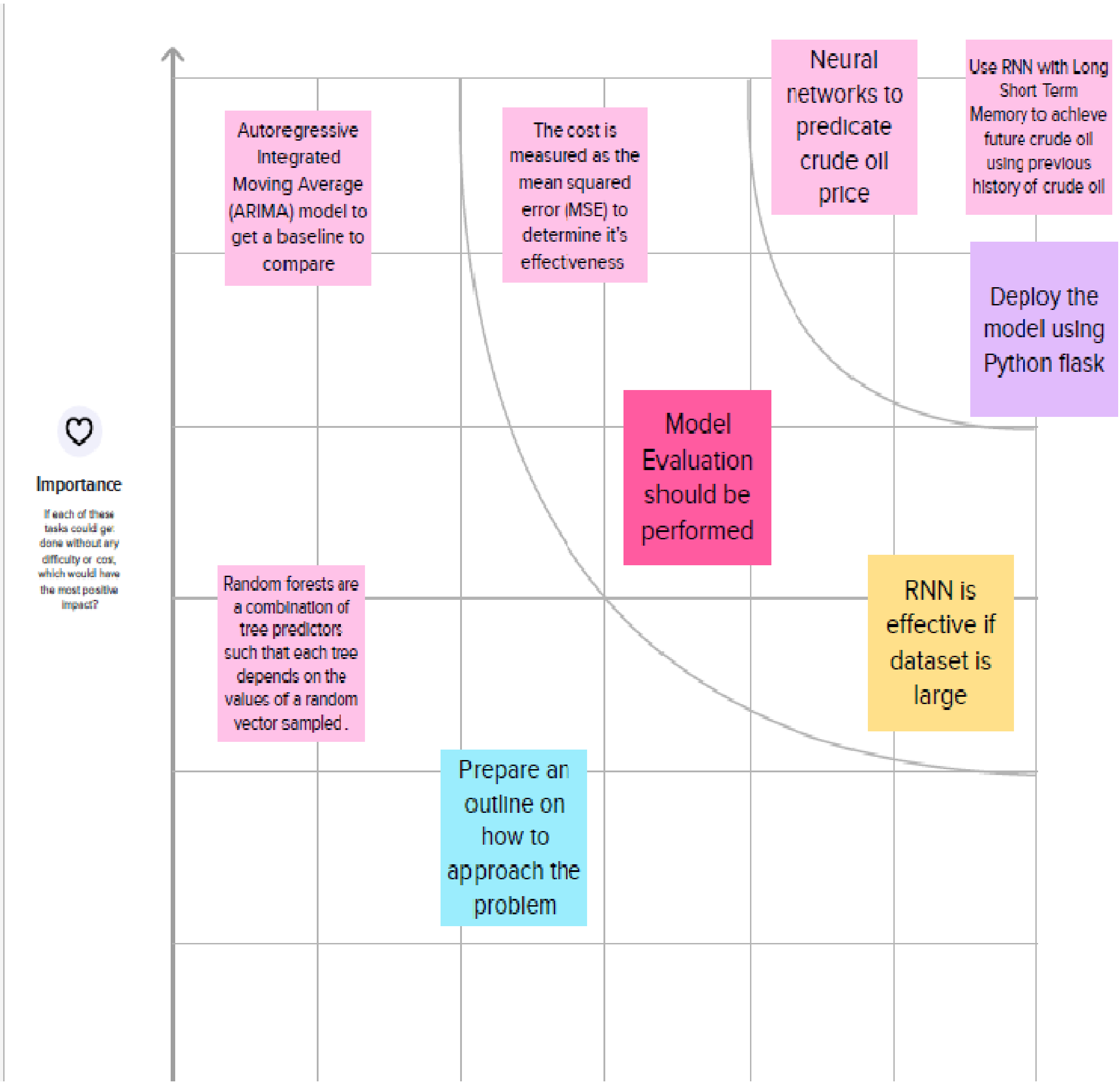

Importance
If each of these tasks could get done without any difficulty or cost, which would have the most positive impact?
Autoregressive Integrated Moving Average (ARIMA) model to get a baseline to compare
The cost is measured as the mean squared error (MSE) to determine it's effectiveness
Neural networks to predicate crude oil price
Use RNN with Long Short Term Memory to achieve future crude oil using previous history of crude oil
Deploy the model using Python flask
Model Evaluation should be performed
Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled .
RNN is effective if dataset is large
Prepare an outline on how to approach the problem

Figure 5 – Group Ideas & Prioritisation

### 3.3 Proposed Solution

The price of crude oil has a significant impact on the environment globally, and its forecasts are particularly helpful to governments and industry.

Crude oil is the most widely used fuel in the world. The ongoing application of statistics and econometric methods for crude oil, including AI Price forecasting could show reductions in the accuracyof the prediction.

In order to predict future crude oil using historical data on crude oil, RNN (Recurrent Neural Network) is utilised with long short-term memory. The effectiveness of the cost is calculated using the denote squared error. Using the pricing information in the crude oil materials, the proposed model'sperformance is assessed.

Since changes in the price of crude oil have a significant impact on national economies around the world, price forecasting can help reduce the risks brought on byoil price volatility.

Governments, public and private businesses, legislators, and investors all place a high valueon price estimates.

**The project "Crude Oil Price Prediction", has the following uniqueness and novelty:**

i. Thismodel is used to forecastfuture pricing and to manage oil use.

ii. This price directly influences many different items and goods, and its variationshave an impacton the capital markets.

iii. Important eventsalso have an impact on oil prices,in addition to economicfactors.

**The project “Crude Oil Price Prediction”, has the following business model:**

iv. Itcan assist those who are making decisionsabout whether to buy or sell crudeoil, whether they are businesses, private investors, or individuals.

v. The benchmark model for predicting crude oil pricesuses RNN and LSTMmodels.

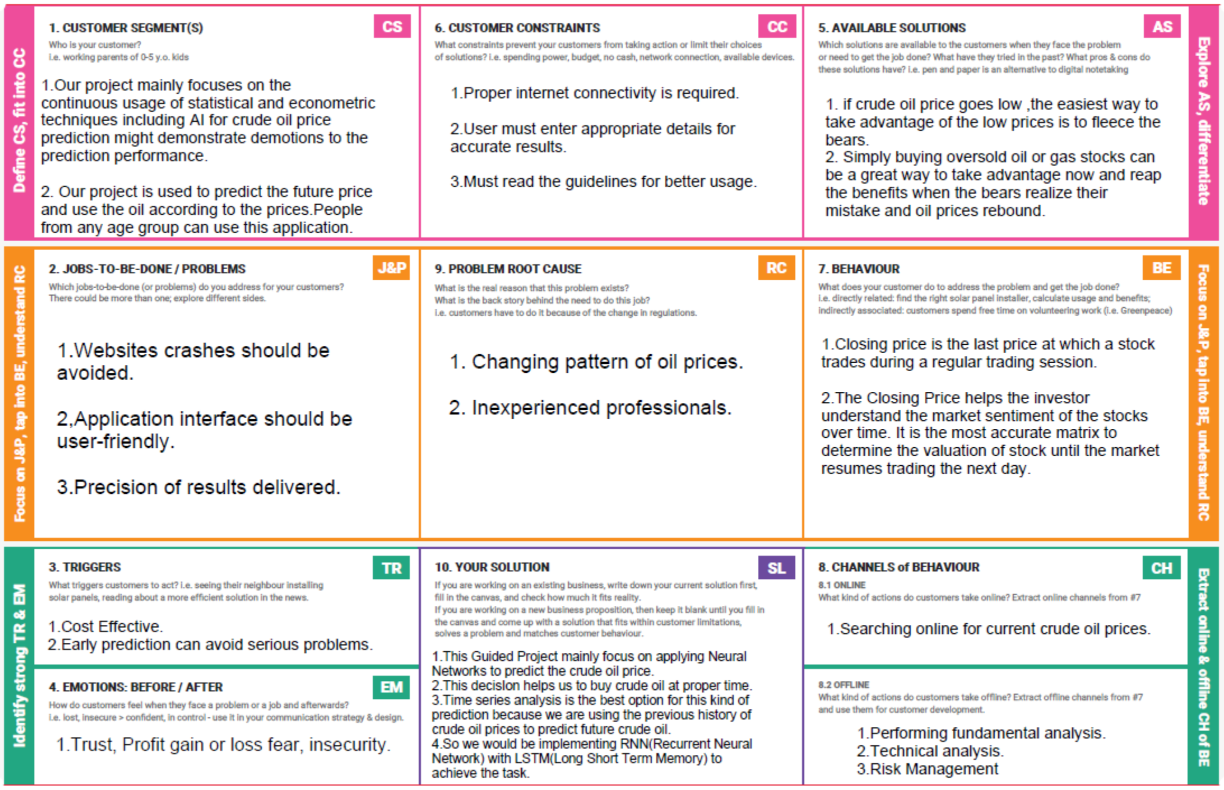
**The scalability of the solutionof this projectare:**

vi. Thedimensions of the data are reduced using the PCA, MDS, and LLE methods.

vii. Enhance the RNN and LSTM models'accuracy.

## 3.4 Problem Solution fit

Problem-Solution fit canvas is not just a mappingtool, but an actionable translation template, where you turn problemsinto solution and communication strategy, taking into account customer behaviour to increase your chances of solutionadoption. It gives you insightsinto how your idea couldfit the reality.


Define CS, fit into CC
CS
CC
AS
Explore AS, differentiate
Focus on J&P, tap into BE, understand RC
J&P
RC
BE
Focus on J&P, tap into BE, understand RC
Identify strong TR & EM
TR
EM
SL
CH
Extract online & offline CH of BE

**1. CUSTOMER SEGMENT(S)**

Who is your customer?
i.e. working parents of 0-5 y.o. kids

1.Our project mainly focuses on the continuous usage of statistical and econometric techniques including AI for crude oil price prediction might demonstrate demotions to the prediction performance.

2. Our project is used to predict the future price and use the oil according to the prices.People from any age group can use this application.

**6. CUSTOMER CONSTRAINTS**

What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devices.

1.Proper internet connectivity is required.

2.User must enter appropriate details for accurate results.

3.Must read the guidelines for better usage.

**5. AVAILABLE SOLUTIONS**

Which solutions are available to the customers when they face the problem or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking

1. if crude oil price goes low ,the easiest way to take advantage of the low prices is to fleece the bears.
2. Simply buying oversold oil or gas stocks can be a great way to take advantage now and reap the benefits when the bears realize their mistake and oil prices rebound.

**2. JOBS-TO-BE-DONE / PROBLEMS**

Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.

1.Websites crashes should be avoided.

2,Application interface should be user-friendly.

3.Precision of results delivered.

**9. PROBLEM ROOT CAUSE**

What is the real reason that this problem exists?
What is the back story behind the need to do this job?
i.e. customers have to do it because of the change in regulations.

1. Changing pattern of oil prices.

2. Inexperienced professionals.

**7. BEHAVIOUR**

What does your customer do to address the problem and get the job done?
i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace)

1.Closing price is the last price at which a stock trades during a regular trading session.

2.The Closing Price helps the investor understand the market sentiment of the stocks over time. It is the most accurate matrix to determine the valuation of stock until the market resumes trading the next day.

**3. TRIGGERS**

What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news.

1.Cost Effective.
2.Early prediction can avoid serious problems.

**4. EMOTIONS: BEFORE / AFTER**

How do customers feel when they face a problem or a job and afterwards?
i.e. lost, insecure > confident, in control - use it in your communication strategy & design.

1.Trust, Profit gain or loss fear, insecurity.

**10. YOUR SOLUTION**

If you are working on an existing business, write down your current solution first, fill in the canvas, and check how much it fits reality.
If you are working on a new business proposition, then keep it blank until you fill in the canvas and come up with a solution that fits within customer limitations, solves a problem and matches customer behaviour.

1.This Guided Project mainly focus on applying Neural Networks to predict the crude oil price.
2.This decision helps us to buy crude oil at proper time.
3.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.
4.So we would be implementing RNN(Recurrent Neural Network) with LSTM(Long Short Term Memory) to achieve the task.

**8. CHANNELS of BEHAVIOUR**

8.1 ONLINE
What kind of actions do customers take online? Extract online channels from #7

1.Searching online for current crude oil prices.

8.2 OFFLINE
What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development.

1.Performing fundamental analysis.
2.Technical analysis.
3.Risk Management

Figure 7 – Problem SolutionFit

# CHAPTER 4

# REQUIREMENT ALANLYSIS

## 4.1 FunctionalRequirement

Following are the functional requirements of the proposedsolution.

| FR No. | Functional Requirement (Epic) | Sub Requirement (Story / Sub-Task) |
|---|---|---|
| FR-1 | User Registration | 1. Registration through Form<br>2. Registration through Gmail<br>3. Registration through LinkedIn |
| FR-2 | User Confirmation | 1. Confirmation via Email<br>2. Confirmation via OTP |
| FR-3 | User Login | 1. Login through username and password<br>2. Login through Gmail<br>3. Login through LinkedIn |
| FR-4 | Primary specifics | 1. Sync oil priceevery second<br>2. Show Up and Down graph in real timein accordance withthe oil price |

| | | |
|---|---|---|
| FR-5 | Additional Requirement | 1. Read latest news<br>2. View price charts<br>3. Review futures on selected quotation<br>4. Analyse historical pricetrends<br>5. Check exchange rates andcommodities futures |
| FR-6 | System Responsibility | 1. Allowing the userto select a date<br>2. Track the precious results<br>3. The pricing newsshould be updated |

Table 1 – Functional Requirements

## 4.2 Non-Functional Requirements

Following are the non-functional requirements of the proposedsolution.

| **FR No.** | **Non-Functional Requirement** | **Description** |
|---|---|---|
| NFR-1 | **Usability** | 1. To utilise a system easily and accelerate routine operations, itmust havea logical userinterface.<br>2. Anyone who registers on the portal can utilise the system. |

| | | |
|---|---|---|
| NFR-2 | **Security** | The following is a list of some of the factors that have been foundto prevent malicious or unintentional access, usage, modification, destruction, or disclosure of the software:<br>1. Maintain particular log or historical data sets.<br>2. Apply specific cryptographymethods.<br>3. Limit the numberof devices that can access the website forpredicting theprice.<br>4. Verify the integrity of the data. |
| NFR-3 | **Reliability** | 1. At the time of entry, all user variable data will be committed tothe database.<br>2. By using the available backup procedures and techniques, datacorruption is avoided. |

| | | |
|---|---|---|
| NFR-4 | **Performance** | 1. The system mustallow for the simultaneous use of many usersatall times.<br>2. The accuracy of the price should be at the maximum. |
| NFR-5 | **Availability** | 1. The system should always be accessible, allowing for simple useraccess.<br>2. A replacement page will be displayed in the event that hardware or data base failure increases, and data should be obtained to restore the system. |
| NFR-6 | **Scalability** | 1. Identifies the maximum workloads at which the system will stilloperate well.<br>2. Focus on themeasurement of the system's response time undervarious load levels. |

Table 2 – Non-Functional Requirements

# CHAPTER 5

# PROJECT DESIGN

## 5.1 Data Flow Diagram

A Data Flow Diagram(DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system,what changes the information, and where data is stored.

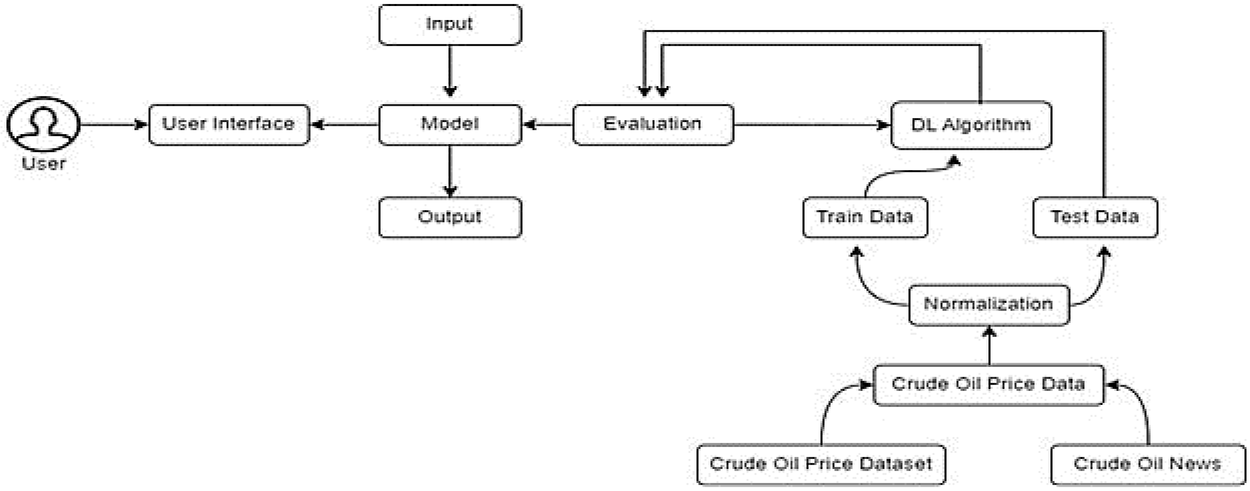


Figure 8 – Data Flow Diagram

## 5.2 Solution & Technical Architecture

Solution architecture is a complex process – with many sub-processes – that bridges the gap betweenbusiness problems and technology solutions. Its goals are to:

i. Find the best tech solutionto solve existingbusiness problems.

ii. Describe the structure, characteristics, behavior, and otheraspects of thesoftware to project stakeholders.

iii. Define features, development phases, and solutionrequirements.

iv. Provide specifications according to which the solution is defined, managed,anddelivered.

**Solution Architecture Diagram:**

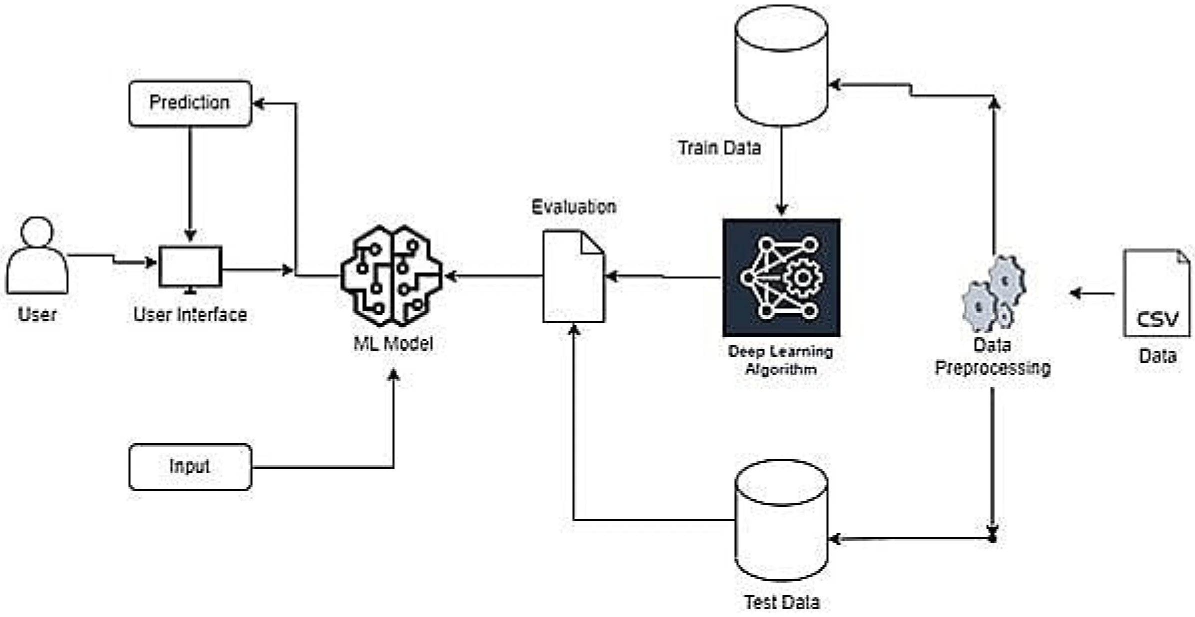


Figure 9 - Architecture Diagram

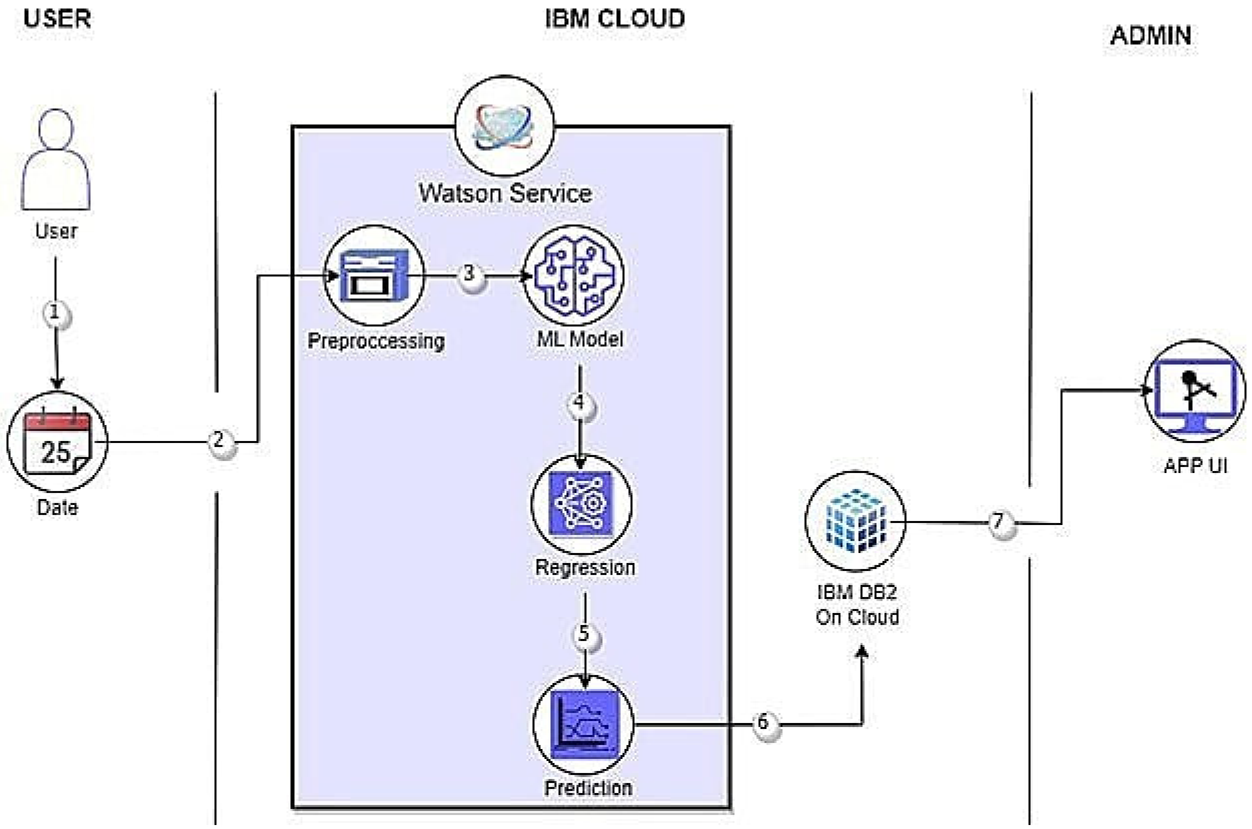


Figure 10 – Technical Architecture Diagram

## Components & Technologies

| S.No | Component | Description | Technology |
|---|---|---|---|
| 1. | User Interface | Through a webUI, the user can engagewith theapplication. | HTML, CSS, JavaScript / AngularJs / ReactJs etc. |

|  |  |  |  |
|---|---|---|---|
| 2. | Application Logic-1 | It has manyin built libraries which helpsinmachine learning | Python |
| 3. | Application Logic-2 | It helps to build machine learning model | IBM Watson Jupyter Notebook service |
| 4. | Application Logic-3 | It is fastand accurate | IBM Watson Assistant |
| 5. | Database | MySQL is used to store the user information andwarehouse the crude oil price | MySQL |
| 6. | Cloud Database | IBM Db2 is reliable and scalable | IBM DB2 |
| 7. | File Storage | Maintain files easily | Local Filesystem |
| 8. | External API-2 | Aadhar and customer KYC verification takesalittle amountof time | Aadhar API, etc. |
| 9. | Machine Learning Model | To recognize the patterns and trends | Sequential, Dense & LSTM Model |

| | | | |
|---|---|---|---|
| 10. | Infrastructure(Server / Cloud) | Application Deployment on Local System / Cloud<br>Local Server Configuration:<br>Cloud Server Configuration | Local System and IBM Watson |

Table 3 – Components & Technologies

## Application Characteristics

| S.No | Characteristics | Description | Technology |
|---|---|---|---|
| 1. | Open-Source Frameworks | Tensor flow – Implements model building andtraining.<br>Flask – Can handle multiple userrequest simultaneously.<br>Scikit learn– Contains model for classification, regression, clustering. | Tensor flow, Flask, Scikit learn. |
| 2. | Security Implementations | SHA-256 doesn't have any knownvulnerabilities | SHA-256. |
| 3. | Scalable Architecture | MySQL can storehuge amount of dataand it | MySQL |

| | | | |
|---|---|---|---|
| | | iseasily scalable. | |
| 4. | Availability | This application can be accessed from anywhereeasily and it is easily scalable. | IBM Watson Cloud. |
| 5. | Performance | Flask can handlemultiple user request simultaneously. | Flask |

Table 4 – Application Characteristics

## 5.3 User Stories

| **User Type** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** | **Acceptance criteria** | **Priority** | **Release** |
|---|---|---|---|---|---|---|
| Customer (Mobile user) | Registration | USN-1 | As a user,I can register | I can access my account | High | Sprint-1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | for the application by entering my email, password, and confirming my password. | / dashboard | | |
| | | USN-2 | As a user,I will receive confirmation email once I have registered for the | I can receive confirmation email & click confirm | High | Sprint-1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | application | | | |
| | | USN-3 | As a user,I can register for the application through Facebook | I can register & access the dashboard with Facebook Login | Low | Sprint-2 |
| | | USN-4 | As a user,I can register | I can register | Medium | Sprint-1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | for the application through Gmail | through already existing mail account. | | |
| | Login | USN-5 | As a user,I can log into the application by entering email & password | After registration, I can log in via only email & password. | High | Sprint-1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Dashboard | USN-6 | Display the oil price, line graph/ bar graph real time. | I can expect the prediction in various formats. | Low | Sprint-3 |
| Customer (Web user) | Login | USN-7 | As the user, I can login by using Gmail or Facebook account or LinkedIn or by registering. | Existing users can easily login. | High | Sprint-2 |
| Customer Care Executive | Support | USN-8 | The Customer care service will provide solutions for any FAQ and also provide | I can solve the problems raised. | High | Sprint-3 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | Chat-Bot. | | | |
| Administrator | Access Control | USN-9 | Admin can control the access of users. | Access permission for Users. | High | Sprint-4 |
| | Database | USN-10 | Admin can store the details of users. | Stores User details. | Medium | Sprint-4 |
| | News | USN-11 | Admin will give the recent news of Oil Prices. | Provide the recent oil prices. | Medium | Sprint-4 |
| | Notification | USN-12 | Admin will | Notification | High | Sprint- |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | notify when the oil prices changes. | by Gmail. | | 4 |

5 – User Stories

# CHAPTER 6

# PROJECT PLANNING & SCHEDULING

## 6.1 Sprint Planning & Estimation:

| S.No | Milestone | Activities | Team Members |
|---|---|---|---|
| 1. | Data Collection | Create Train and TestFolders | DHYALAN C,<br>DINESH KUMAR S,<br>JAYAKUMAR S,<br>SATHISH G |

| | | | |
|---|---|---|---|
| 2. | Data Pre-processing | Import the libraries, Handle the missing data, | DHYALAN C,<br>JAYAKUMAR S,<br>SATHISH G |
| 3. | | Feature scaling, Data Visualization, Splitting the data intotest, train,and split. | DHYALAN C,<br>DINESH KUMAR S,<br>SATHISH G |
| 4. | Model Building | Import the required model building libraries | DHYALAN C,<br>SATHISH G |
| 5. | Model Building | Initialize themodel | DHYALAN C,<br>JAYAKUMAR S, |
| 6. | Model Building | Add the LSTM Layers | DHYALAN C,<br>DINESH KUMAR S, |
| 7. | Model Building | Add the output layers | DHYALAN C,<br>SATHISH G |
| 8. | Model Building | Configure the learningprocess | DHYALAN C,<br>JAYAKUMAR S, |
| 9. | Model Building | Train the model | DHYALAN C,<br>DINESH KUMAR S, |
| 10. | Model Building | Model evaluation | DHYALAN C,<br>SATHISH G |
| 11. | Model Building | Fit and save the model | DHYALAN C,<br>JAYAKUMAR S, |

| | | | |
|---|---|---|---|
| | | | |
| 12 | Test theModel | Import the packages andload thesaved models | DHYALAN C,<br>DINESH KUMAR S, |
| | | | |
| 13. | | Load the test image, pre-process it, and predict | DHYAL AN C,<br>JAYAKUMAR S, |
| 14. | Application Building | Build a flask application | DHYALAN C,<br>DINESH KUMAR S, |
| 15. | | Build the HTML page | DHYALAN C,<br>DINESH KUMAR S, |
| 16. | | Output | DHYAL AN C,<br>JAYAKUMAR S, |

| | | | |
|---|---|---|---|
| 17. | Train The Model onIBM | Register for IBM Cloud | DHYALAN C,<br>DINESH KUMAR S, |
| 18. | | Train Image Classification Model | DHYALAN C,<br>DINESH KUMAR S, |
| 19. | Ideation Phase | Prepare Empathymap | DHYALAN C,<br>JAYAKUMAR S, |
| 20. | | Literature Surveyon the selected project & Information Gathering | DHYALAN C,<br>DINESH KUMAR S, |
| 21. | | Ideation | DHYALAN C,<br>DINESH KUMAR S, |

| | | | |
|---|---|---|---|
| 22. | Project Design PhaseI | Proposed Solution | DHYALAN C,<br>JAYAKUMAR S, |
| 23. | | Proposed Solution Fit | DHYALAN C,<br>JAYAKUMAR S, |
| 24. | | Solution Architecture | DHYALAN C,<br>DINESH KUMAR S, |
| 25. | Project Design PhaseII | Customer Journey | DHYALAN C,<br>JAYAKUMAR S, |
| 26. | | Functional Requirements | DHYALAN C,<br>SATHISH G |
| 27. | | Data Flow Diagram | DHYALAN C,<br>JAYAKUMAR S, |
| 28. | | Technology Architecture | DHYALAN C,<br>SATHISH G |
| 29. | Project Planning Phase | Prepare Milestone and ActivityList | DHYALAN C,<br>SATHISH G |

| | | | |
|---|---|---|---|
| 30. | Project Development Phase | Delivery of Sprint 1 | DHYALAN C,<br>JAYAKUMAR S, |
| 31. | | Delivery of Sprint 2 | DHYALAN C,<br>SATHISH G |
| 32. | | Delivery of Sprint 3 | DHYALAN C,<br>SATHISH G |
| 33. | | Delivery of Sprint 4 | DHYALAN C,<br>JAYAKUMAR S, |

Table 6 – Sprint Plan

## 6.2 Sprint Delivery Schedule

| **Title** | **Description** | **Date** |
|---|---|---|
| Literature Survey & Information Gathering | Literature survey on the selected project & gathering information byreferring the, technical papers, research publications etc. | 28 September 2022 |

|  |  |  |
|---|---|---|
| Prepare Empathy Map | Prepare Empathy Map Canvas to capture theuser Pains &Gains, Prepare listof problem statements | 11 September 2022 |
| Ideation | List the by organizing the brainstorming session and prioritize the top 3 ideas basedon the feasibility & importance. | 18 September 2022 |
| Proposed Solution | Prepare the proposed solution document, which includes the novelty,feasibility of idea, business model, socialimpact, scalability of solution, etc. | 23 October 2022 |
| Problem Solution Fit | Prepare problem - solution fit document. | 15 October 2022 |
| Solution Architecture | Prepare solution architecture document. | 10 October 2022 |
| Customer Journey | Prepare the customer journey maps to understand the user interactions &experiences with the application (entryto exit). | 09 October 2022 |
| Functional Requirement | Prepare the functional requirement document. | 20 October 2022 |

| | | |
|---|---|---|
| Data Flow Diagrams | Draw the dataflow diagrams and submit forreview. | 23 October 2022 |
| Technology Architecture | Prepare the technology architecture diagram. | 26 October 2022 |
| Prepare Milestone & Activity List | Prepare the milestones and activity list of the project. | 27 October 2022 |
| Sprint Schedule | Prepare spring plan | 27 October 2022 |
| Delivery of Sprint-1 | Develop & submitthe developed code. | 29 October 2022 |
| Delivery of Sprint-2 | Develop & submitthe developed code. | 06 November 2022 |
| Delivery of Sprint-3 | Develop & submitthe developed code. | 11 November 2022 |
| Delivery of Sprint-4 | Develop & submitthe developed code. | 15 November 2022 |

Table 7 – SprintPlan 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 | 06 Nov 2022 |
| Sprint-3 | 20 | 6 Days | 07 Nov 2022 | 12 Nov 2022 | 20 | 11 Nov 2022 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Sprint-4 | 20 | 6 Days | 12 Nov 2022 | 16 Nov 2022 | 20 | 16 Nov 2022 |

8 – Sprint DeliverySchedule

## 6.3 Reports From 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)

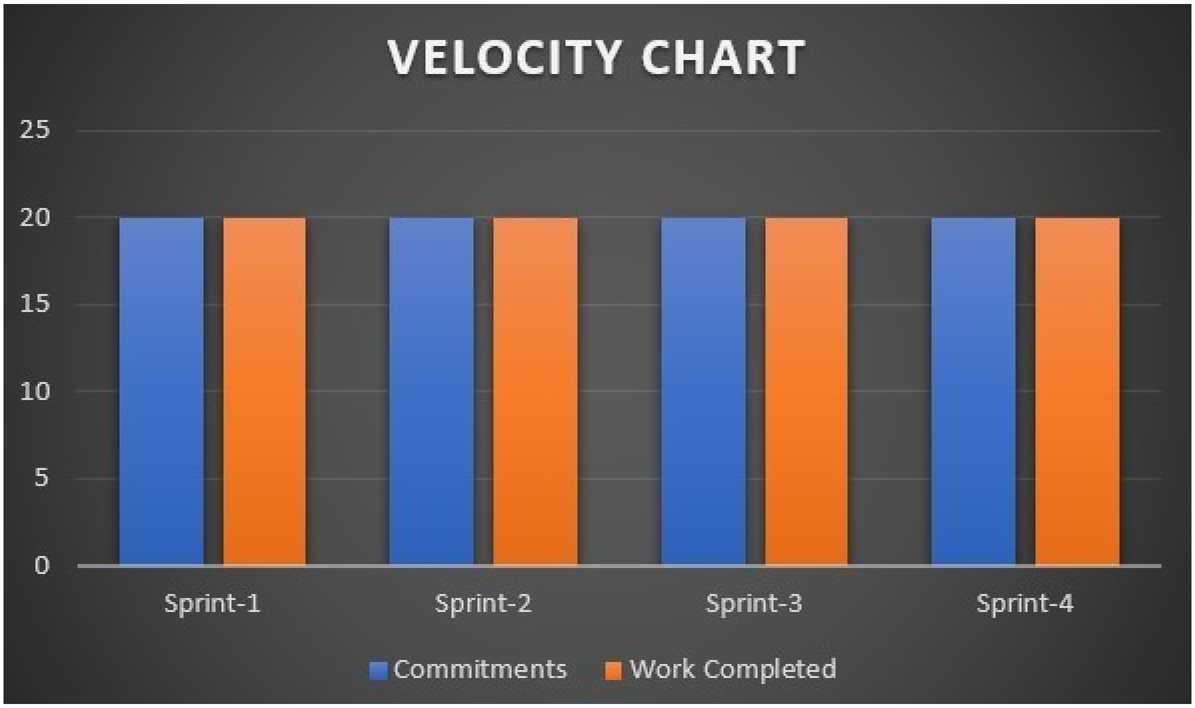


Figure 11 – Velocity Chart

### Burndown Chart:

A burn down 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.

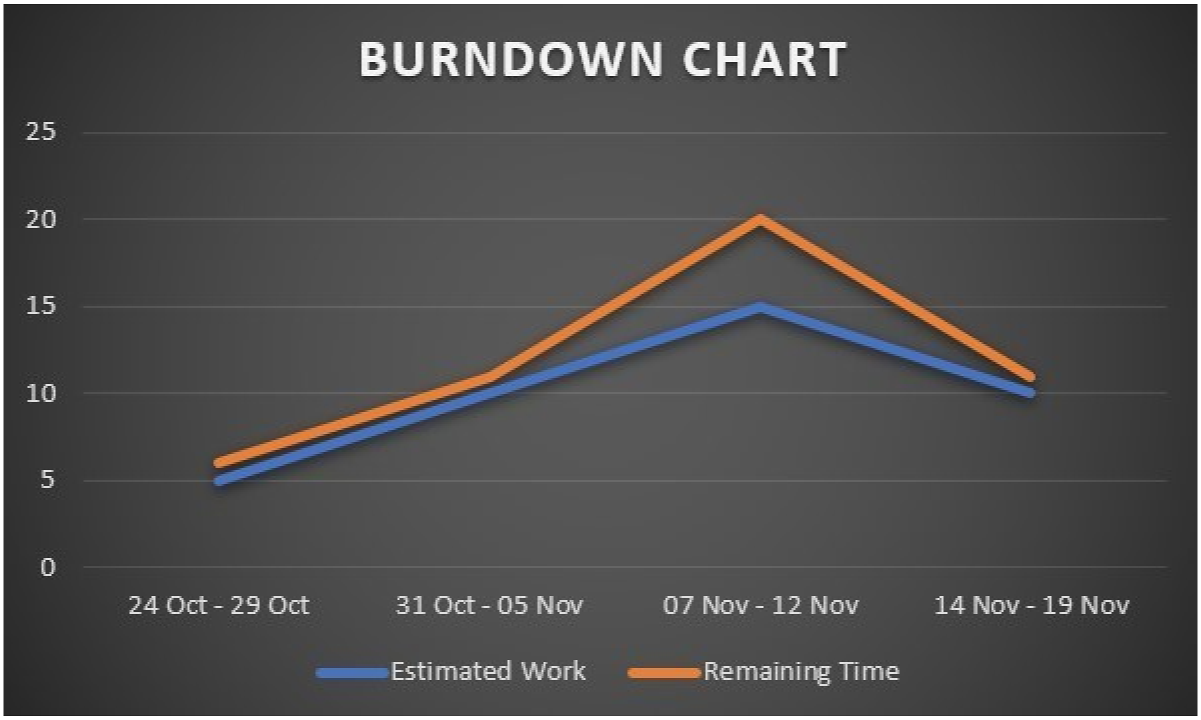


Figure 12 – Burndown Chart

# CHAPTER 7

# CODING & SOLUTIONING

## 7.1 InteractiveUI

The area where interactions between people and machines take place is known as a user interface (UI) in the subject of industrial design known as human-computer interaction. This interaction's purpose is to enable efficient machine operation and control from the human end, while the machine also feeds information back to the operators to support their decision-making. The general objective of user interface design is to provide an interface that makes it simple, effective, and pleasurable (user- friendly) to operatea machine in a way that yields the desiredoutcome (i.e., maximumusability). This typicallymeans that the machine reducesundesirable outputs to the user while simultaneously requiring the operator to input as little as possibleto produce the desired output.

We have included a user interface in our project to make it easier for users to forecast the price of crude oil in the future. Users simply need to visit the website to access the interface and can click a button to forecast the price. Once the button has been clicked, the user will be taken to another website where they can enter the price of crude oil for 10 days. In that case, the user should click Predict. The user can then view the price ofcrude oil after ten days.

## 7.2 Cloud Integration

The on-demand availability of computer system resources, in particular data storage (cloud storage) and processing power, without direct active supervision by the user, is known as cloud computing. Functions in large clouds are frequently dispersed over severalsites, each of which is a data centre. Cloud computing often uses a "pay as you go" model, which can help reduce capital expenses but may also result inunanticipated running expensesfor users. Cloud computing dependson resource sharing to accomplish coherence.

Our projectis cloud-integrated, allowingit to run anywhere and be accessible at any time. Anytime the user desires, they will be able to forecast the price of crude oil. Through the IBM Cloud, this is accomplished. On the IBM Watson Studio, whichmakes use of the Watson Machine Learning Platform, we developed and trained the model. We generated a deployment space and ran the code using the API key to deploy the model. The Flask app, which is used to link to the backend and frontend,was then finally integrated.

# CHAPTER 8

# TESTING

## 8.1 Test Cases

The following test scenarios were tested successfully.

**Test Scenarios**

i. Verify the UI elements on the home page

ii. Verify whether the user can navigate to the prediction page

iii. Verify the UI elements in the prediction page

iv. Verify user is able to enter a valuein the text box.

v. Verify user is able to enter numbersin the text box

vi. Verify model can handle no inputs

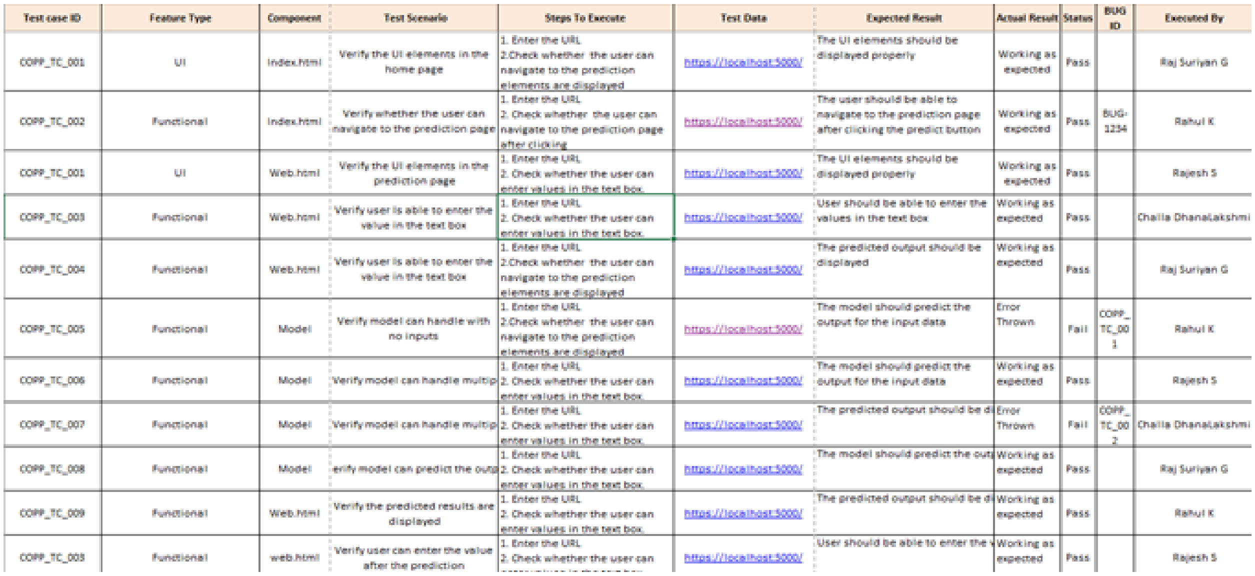
vii. Verify model can handle multipleinput

viii. Verify model can handle unsupported input

ix. Verify model can predict the output

x. Verify the predicted results are displayed

xi. Verify user can enter the value after the prediction



| Test case ID | Feature Type | Component | Test Scenario | Steps To Execute | Test Data | Expected Result | Actual Result | Status | BUG ID | Executed By |
|---|---|---|---|---|---|---|---|---|---|---|
| COPP_TC_001 | UI | Index.html | Verify the UI elements in the home page | 1. Enter the URL<br>2.Check whether the user can navigate to the prediction elements are displayed | https://localhost:5000/ | The UI elements should be displayed properly | Working as expected | Pass | | Raj Suriyan G |
| COPP_TC_002 | Functional | Index.html | Verify whether the user can navigate to the prediction page | 1. Enter the URL<br>2. Check whether the user can navigate to the prediction page after clicking | https://localhost:5000/ | The user should be able to navigate to the prediction page after clicking the predict button | Working as expected | Pass | BUG-1234 | Rahul K |
| COPP_TC_001 | UI | Web.html | Verify the UI elements in the prediction page | 1. Enter the URL<br>2. Check whether the user can enter values in the text box. | https://localhost:5000/ | The UI elements should be displayed properly | Working as expected | Pass | | Rajesh S |
| COPP_TC_003 | Functional | Web.html | Verify user is able to enter the value in the text box | 1. Enter the URL<br>2. Check whether the user can enter values in the text box. | https://localhost:5000/ | User should be able to enter the values in the text box | Working as expected | Pass | | Challa Dhanalakshmi |
| COPP_TC_004 | Functional | Web.html | Verify user is able to enter the value in the text box | 1. Enter the URL<br>2.Check whether the user can navigate to the prediction elements are displayed | https://localhost:5000/ | The predicted output should be displayed | Working as expected | Pass | | Raj Suriyan G |
| COPP_TC_005 | Functional | Model | Verify model can handle with no inputs | 1. Enter the URL<br>2.Check whether the user can navigate to the prediction elements are displayed | https://localhost:5000/ | The model should predict the output for the input data | Error Thrown | Fail | COPP_TC_001 | Rahul K |
| COPP_TC_006 | Functional | Model | Verify model can handle multip | 1. Enter the URL<br>2. Check whether the user can enter values in the text box. | https://localhost:5000/ | The model should predict the output for the input data | Working as expected | Pass | | Rajesh S |
| COPP_TC_007 | Functional | Model | Verify model can handle multip | 1. Enter the URL<br>2. Check whether the user can enter values in the text box. | https://localhost:5000/ | The predicted output should be di | Error Thrown | Fail | COPP_TC_002 | Challa Dhanalakshmi |
| COPP_TC_008 | Functional | Model | erify model can predict the outp | 1. Enter the URL<br>2. Check whether the user can enter values in the text box. | https://localhost:5000/ | The model should predict the out | Working as expected | Pass | | Raj Suriyan G |
| COPP_TC_009 | Functional | Web.html | Verify the predicted results are displayed | 1. Enter the URL<br>2. Check whether the user can enter values in the text box. | https://localhost:5000/ | The predicted output should be di | Working as expected | Pass | | Rahul K |
| COPP_TC_003 | Functional | web.html | Verify user can enter the value after the prediction | 1. Enter the URL<br>2. Check whether the user can | https://localhost:5000/ | User should be able to enter the v | Working as expected | Pass | | Rajesh S |

Figure 13 – Test Cases

## 8.2 User Acceptance Testing:

### Defect Analysis

| Resolution | Severity1 | Severity2 | Severity3 | Severity4 | Subtotal |
|---|---|---|---|---|---|
| By Design | 1 | 0 | 0 | 1 | 0 |
| Duplicate | 0 | 0 | 0 | 0 | 0 |
| External | 0 | 0 | 2 | 0 | 2 |
| Fixed | 4 | 1 | 0 | 1 | 6 |
| Not Reproduced | 0 | 0 | 0 | 0 | 0 |
| Skipped | 1 | 0 | 0 | 0 | 1 |
| Won't Fix | 1 | 0 | 1 | 1 | 3 |
| Totals | 7 | 1 | 3 | 3 | 1 |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | 2 |

Table 9 – Defect Analysis

**Test Case Analysis**

| Section | Total Cases | Not Tested | Fail | Pass |
|---|---|---|---|---|
| Print Engine | 10 | 0 | 2 | 8 |
| Client Application | 5 | 0 | 0 | 5 |
| Security | 1 | 0 | 0 | 1 |
| Outsource Shipping | 3 | 0 | 0 | 3 |
| Exception Reporting | 2 | 0 | 2 | 0 |
| Final ReportOutput | 4 | 0 | 0 | 4 |

Table 10 – Test Case Analysis

# CHAPTER 9

# RESULTS

## 9.1 Performance Metrics:

We attempted to forecast the output of the crude oil by entering various input variables in order to assess the accuracy and performance of this project. These are theinput values.

[0.44172960165852215, 0.48111950244335855, 0.49726047682511476, 0.4679401747371539, 0.4729749740855915, 0.47119798608026064, 0.47341922108692425, 0.4649785280616022, 0.4703835332444839, 0.47149415074781587]

The anticipated outcomeafter providing the input valuesis 0.46976325.

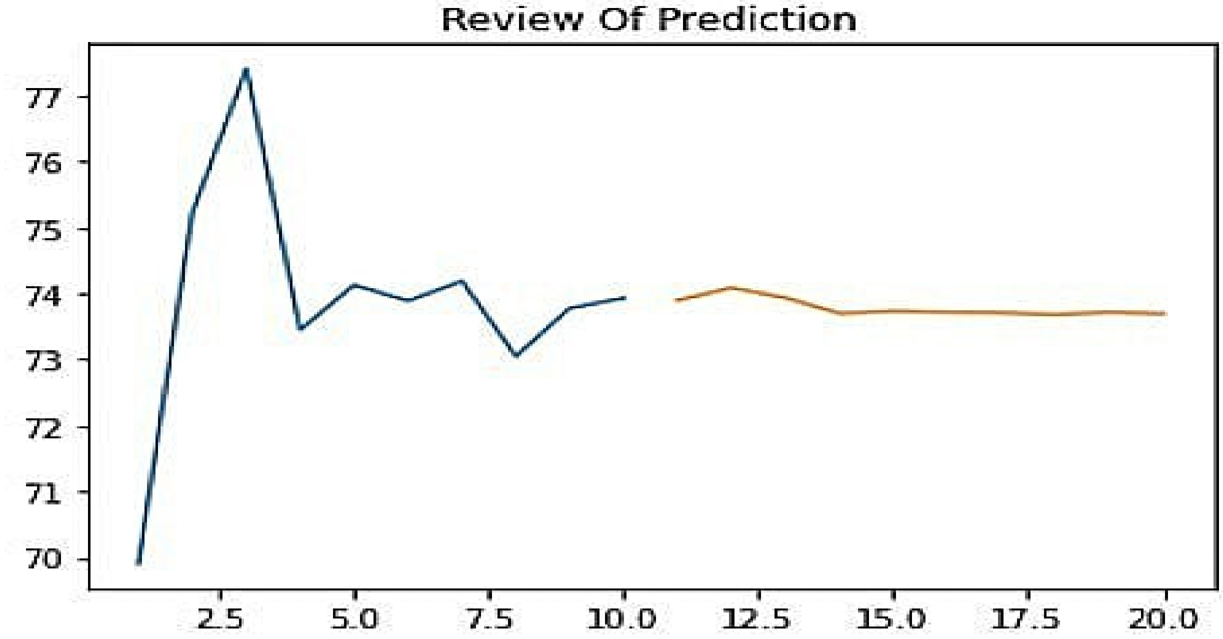
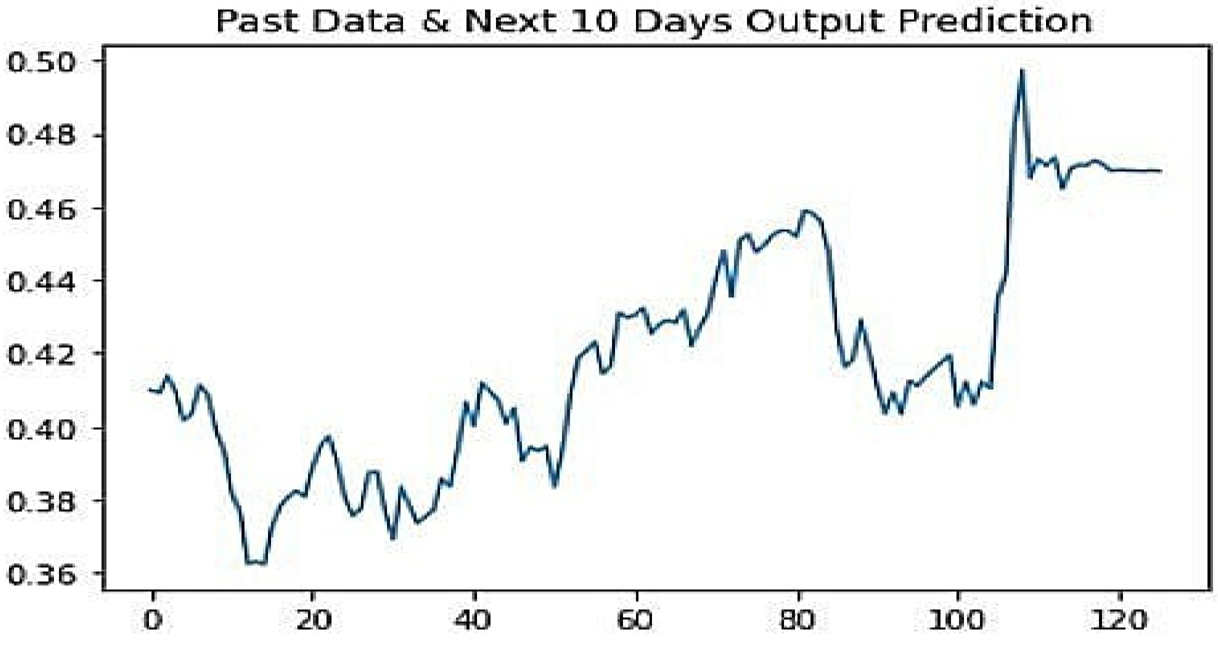


Figure 14 – Review of Prediction

Figure 13 givesa review of prediction how the systemhas predicted the future price

based on the given input values.



It can be seen that the graph was drawn using the provideddata and a projection forthe next 10 days. There was a little discrepancy between the output and the real pricing.

The developed system shows a clear prediction of the future prices which has very less deviations from the true prices by using LSTM in tensorflow and keras in python. There is always a thin line between the overfitting of the model and its best performance. This project helps a lot to learn about the developed model and the algorithm and using this model as a base, a much more complicated model can be easily developed. The facet of more prediction algorithms for crude oil can concoct with the help ofthis system.

This system concludes that the machine learning model LSTM (Long Short- Term Method) predictsthe future price of crude oil by bordering the actual price of the crude oil price.

# CHAPTER 10

# ADVANTAGES & DISADVANTAGES

## Advantages

1. HighAccuracy
2. Removes the investment bias
3. Develop the habit of complete analysis
4. Minimize our losses
5. Allows smart way of making money

### High Accuracy:

The model which we predicted had a high accuracy of above 90 per cent in all aspects.The other advantages of predicting the price of crude oil are discussed below.

### Removes the investment bias:

The Indian stock market offers a variety of chances for traders and investors, but it is also helpful to be aware of the market environment before taking a position ina particular stock. Take the weather prediction as an example to help you comprehend this; being aware of the weather forecast for the coming week enables you to make appropriate plans. The situation with stock market investments is comparable. Let'slook at a few of the major benefits connectedwith stock marketprediction now tohelp you grasp.

**Develop the habitof complete analysis:**

Investors don't always conduct a thorough research of the stock before learning how to anticipate the stock market and putting what they have learned into practise.They only start to establish the habit of comprehensive analysis before making any investing decisionsafter they learn how to apply formulaeand procedures to forecast

stock market movements. Once or initially, making a successful stock market prediction gives investors the confidence to form the habit of conducting a thorough analysis each time. Here, "complete analysis" refers to both the fundamental and the technical analysis of the stocks because the combination of these two forecasting methods results in predictions that are more precise.

**Minimise our losses:**

Another benefit of stock market prediction is that it significantly reduces your losses or restricts them. Investors sometimes make the error of not doing their studies thoroughly before learninghow to anticipate, which resultsin them frequently employing the incorrect prediction strategies. As a result, many put their money into the stocks based solely on intuition or merely wild estimates in the hopes that the prices will rise, and they will profit. They lose most of the time because it doesn't happen.They can reduce their losses by correctlyimplementing and using the appropriate forecast strategies. The converse of this is also true, and given the information provided, you can make wise selections.

**Allows smart way of makingmoney:**

Making steadily increasing profits through the use of your trading expertise and knowledge is the smart method to make money. The most desired and ideal approach to make money in the stock market is to become a day trader and make money every day, unless of course a person has long-term aspirations. But in order to do that, you must be aware of the various difficulties and difficulties that come with intraday trading, as well as how to deal with them. That can only occur when you understand how to forecast the stock market using a variety of tools and tactics and how to maximiseintraday trading, enablingyourself to consistently make money.

## Disadvantages

6. Forecasts are never 100 percent accurate

7. Itcan be time-consuming and resource-intensive

**Forecasts are never 100% accurate:**

Let’s face it: it’s hard to predict the future. Even if you have a great process in place and forecasting experts on your payroll, your forecasts will never be spot on.Some products and markets simply have a high level of volatility. And in general, there is just an endlessnumber of factorsthat influence demand.

**It can be time-consuming and resource-intensive:**

Forecasting involves numerous data gathering, data organizing, and coordination. Companiestypically employ a team of demand planners who are responsible for coming up with the forecast. But in order to do this well, demand planners need substantial input from the sales and marketing teams. In addition, it’s not uncommon for processesto be manual and labour-intensive, thus taking up a lot of time. Fortunately, if you have the right technology in place, this is much less of an issue.

□

# CHAPTER 11

□

□

# CONCLUSION

In today's world and in such a dynamic atmosphere where everyone wants to know what will happen in the future, artificial intelligence and deep learning are the foundation for upgradingtechnology. The path to future prediction has been established by several facilities. It previously hard to predictthe prices of cryptocurrencies since they change randomly, but machine learninghas made it feasible.

By integrating LSTM in TensorFlow and keras in Python, the constructed model demonstrates a clear prediction of the future prices with very little variance from the genuine prices. Between the model being overfitted and performing at its optimum, there is always a fine line. With a few minor adjustments, the model may be applied to different time series data. With the knowledge gained from this research, a far more complex model may be created with relative ease utilising the generated model and algorithm as a foundation. With the aid of this model, more prediction algorithms for bitcoin may be developed.

This projectcomes to the conclusion that the LSTM (Long Short-Term Method) machine learning algorithm predicts the future price of crude oil by edging the currentprice of the oil with high accuracy.

# CHAPTER 12

## FUTURE SCOPE

The Long Short-Term Method (LSTM) machine learning algorithm is shown to have a high degree of accuracy in predicting the future price of crude oil by edging thecurrent price of the oil.

In the future, it will be possible to estimate crude oil prices by taking into account additional variables that influence the price, such as tweets, national news, natural disasters, the cost of forecasting, conflict, demand, and floods. By doing this, the model'sprecision and accuracy would both be enhanced.

The dataset will be obtained from Kaggle, a sizable platform that is frequently used for data mining and doing analysis. The model would similarly be created using these elements. If this is carried out, the accuracy of forecasting the price of crude oil will exceed 98 percent.

# CHAPTER 13

# APPENDIX

## Source Code

### Building the model:

IMPORTING LIBRARIES

```
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import datetime
from pylab import rcParams
import matplotlib.pyplot as plt
import warnings
import itertools
import statsmodels.api as sm
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import LSTM
from keras.layers import Dropout
from sklearn.metrics import mean_squared_error
from keras.callbacks import ReduceLROnPlateau, EarlyStopping,
ModelCheckpoint
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
import seaborn as sns
sns.set_context("paper", font_scale=1.3)
sns.set_style('white')
```

```
import math
from sklearn.preprocessing import MinMaxScaler

# Input data files are available in the "../input/" directory.
# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the input directory
warnings.filterwarnings("ignore")
plt.style.use('fivethirtyeight')
import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

## IMPORTING DATA

```
dateparse = lambda x: pd.datetime.strptime(x, '%b %d, %Y')
#Read csv file
from google.colab import files
uploaded = files.upload()
import io
df = pd.read_excel(io.BytesIO(uploaded['Crude Oil Prices Daily.xlsx']))
df.head()
df[:10]

#Sort dataset by column Date
df = df.sort_values('Date')
df = df.groupby('Date')['Closing Value'].sum().reset_index()
df.set_index('Date', inplace=True)
df=df.loc[datetime.date(year=2000,month=1,day=1):]
df.head()
```

## DATA PRE-PROCESSING

```
def DfInfo(df_initial):
    # gives some infos on columns types and numer of null values
    tab_info = pd.DataFrame(df_initial.dtypes).T.rename(index={0: 'column type'})
    tab_info =
tab_info.append(pd.DataFrame(df_initial.isnull().sum()).T.rename(index={0: 'null
values (nb)'}))
    tab_info = tab_info.append(pd.DataFrame(df_initial.isnull().sum() /
df_initial.shape[0] * 100).T.
                               rename(index={0: 'null values (%)'}))
    return tab_info
DfInfo(df)
df.index
y = df['Closing Value'].resample('MS').mean()
y.plot(figsize=(15, 6))
plt.show()
rcParams['figure.figsize'] = 18, 8
decomposition = sm.tsa.seasonal_decompose(y, model='additive')
fig = decomposition.plot()
plt.show()
sc = MinMaxScaler(feature_range = (0, 1))
df = sc.fit_transform(df)
```

## TRAINING AND TESTING

```
train_size = int(len(df) * 0.70)
test_size = len(df) - train_size
train, test = df[0:train_size, :], df[train_size:len(df), :]
def create_data_set(_data_set, _look_back=1):
    data_x, data_y = [], []
    for i in range(len(_data_set) - _look_back - 1):
        a = _data_set[i:(i + _look_back), 0]
        data_x.append(a)
        data_y.append(_data_set[i + _look_back, 0])
    return np.array(data_x), np.array(data_y)
look_back =90
X_train,Y_train,X_test,Ytest = [],[],[],[]
X_train,Y_train=create_data_set(train,look_back)
X_train = np.reshape(X_train, (X_train.shape[0], X_train.shape[1], 1))
X_test,Y_test=create_data_set(test,look_back)
X_test = np.reshape(X_test, (X_test.shape[0], X_test.shape[1], 1))
```

## LSTM LAYER

```
regressor = Sequential()

regressor.add(LSTM(units = 60, return_sequences = True, input_shape =
(X_train.shape[1], 1)))
regressor.add(Dropout(0.1))

regressor.add(LSTM(units = 60, return_sequences = True))
regressor.add(Dropout(0.1))

regressor.add(LSTM(units = 60))
regressor.add(Dropout(0.1))

regressor.add(Dense(units = 1))



regressor.compile(optimizer = 'adam', loss = 'mean_squared_error')
reduce_lr = ReduceLROnPlateau(monitor='val_loss',patience=5)
history =regressor.fit(X_train, Y_train, epochs = 20, batch_size =
15,validation_data=(X_test, Y_test), callbacks=[reduce_lr],shuffle=False)
```

## MODEL TRAINING

```
train_predict = regressor.predict(X_train)
test_predict = regressor.predict(X_test)
train_predict = sc.inverse_transform(train_predict)
Y_train = sc.inverse_transform([Y_train])
test_predict = sc.inverse_transform(test_predict)
Y_test = sc.inverse_transform([Y_test])
```

PREDICTION

```
print('Train Mean Absolute Error:', mean_absolute_error(Y_train[0],
train_predict[:,0]))
print('Train Root Mean Squared Error:',np.sqrt(mean_squared_error(Y_train[0],
train_predict[:,0])))
print('Test Mean Absolute Error:', mean_absolute_error(Y_test[0],
test_predict[:,0]))
print('Test Root Mean Squared Error:',np.sqrt(mean_squared_error(Y_test[0],
test_predict[:,0])))
plt.figure(figsize=(8,4))
plt.plot(history.history['loss'], label='Train Loss')
plt.plot(history.history['val_loss'], label='Test Loss')
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epochs')
plt.legend(loc='upper right')
plt.show();
aa=[x for x in range(180)]
plt.figure(figsize=(8,4))
plt.plot(aa, Y_test[0][:180], marker='.', label="actual")
plt.plot(aa, test_predict[:,0][:180], 'r', label="prediction")
plt.tight_layout()
sns.despine(top=True)
```

```
plt.subplots_adjust(left=0.07)
plt.ylabel('Price', size=15)
plt.xlabel('Time step', size=15)
plt.legend(fontsize=15)
plt.show();
```

**Deploying on IBM Cloud:**

```
    get_ipython().system('pip install ibm_watson_machine_learning')

    from ibm_watson_machine_learning importAPIClient
    wml_credentials = {
       "url": "https://us-south.ml.cloud.ibm.com",
       "apikey": "uVEty-CB4dYcccQ_Jq9V-atVXmL1dByE_wiDm95lcyTQ"
    }

    client = APIClient(wml_credentials)

    def guid_from_space_name(client,
       NewSpace):space =
       client.spaces.get_details()
       return(next(item for item in space['resources'] if item['entity']["name"] ==
    NewSpace)['metadata']['id'])

    space_uid = guid_from_space_name(client,
    'NewSpace')print("Space UID = " + space_uid)

    client.set.default_space(space_uid)

client.software_specifications.list()
```

```
software_spec_id =
client.software_specifications.get_id_by_name('tensorflow_r
t22.1-py3.9') print(software_spec_id)

model.save('crude.h5')

get_ipython().system('tar -zcvf crude-oil.tgz Crude.h5')

software_space_uid =
client.software_specifications.get_uid_by_name('tensorflow_r
t22.1-py3.9') software_space_uid

model_details = client.repository.store_model(model='crude.tgz',meta_props={
client.repository.ModelMetaNames.NAME:"crude_oil_model",
client.repository.ModelMetaNames.TYPE:"tensorflow_2.7",
client.repository.ModelMetaNames.SOFTWARE_SPEC_UID:software_spec_id
}
                                    )
model_id =
client.repository.get_model_uid(model_details)
model_id

client.repository.download(model_id,'crude_oil_model.tar.gb')
```

## INTEGRATE FLASK WITH SCORING END POINT

**App.py**

```
from flask
importFlask,render_template,request,redirectimport
pandas as pd
import numpyas np
from flask importFlask, render_template, Response,
requestimport pickle
from sklearn.preprocessing
importLabelEncoderimport requests

# NOTE: you must manuallyset API_KEY below using information retrieved
fromyour IBM Cloudaccount.
API_KEY = "uVEty-CB4dYcccQ_Jq9V-atVXmL1dByE_wiDm95lcyTQ"
token_response = requests.post('https://iam.cloud.ibm.com/identity/token',
data={"apikey":API_KEY, "grant_type": 'urn:ibm:params:oauth:grant-
type:apikey'})mltoken= token_response.json()["access_token"]
header = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + mltoken}



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']
        string = string.split(',')
        temp_input = [eval(i) for i in string]

        x_input = np.zeros(shape=(1, 10))
        x_input.shape

        lst_ou
        tput =
        []
        n_ste
        ps =
        10
        i=0
        while
        (i<10
        ):
            if(len(temp_input)>10):
                x_input =
                np.array(temp_input[1:])
                x_input= x_input.reshape(1,-1)
                x_input =
                x_input.reshape((1,n_steps, 1)) yhat
                = model.predict(x_input, verbose=
                0)temp_input.extend(yhat[0].tolist())
                temp_input = temp_input[1:]

                lst_output.extend(yhat.to
                list())i=i+1

            else:
```

```
                x_input = x_input.reshape((1,
                n_steps,1)) yhat =
                model.predict(x_input, verbose= 0)
                temp_input.extend(yhat[0].tolist())
                lst_output.extend(yhat.tolist())
                i=i+1



          # NOTE: manuallydefine and pass the array(s)of values to be scored in
thenext line
          payload_scoring = {"input_data": [{   "values": [[x_input]]    }]}


          response_scoring = requests.post('https://us-
south.ml.cloud.ibm.com/ml/v4/deployments/7f67cbed-6222-
413b-9901-b2a72807ac82/predictions?version=2022-10-30',
json=payload_scoring, headers={'Authorization': 'Bearer' +
mltoken})
          predictions = response_scoring.json()
          print(response_scoring.json())


          val = lst_output[9]
          return render_template('web.html' , prediction = val)



    if request.method=="GET":
        return render_template('web.html')


if__name__=="__main__":
    model = load_model('C:/Users/rkara/IBM/Sprint -
    4/Crude_oil.tar.gz')app.run(debug=True)
```

## INDEX.HTML

```
<!DOCTYPE html>
<html>
<head>
<title>PRICE PREDICTION</title>
</head>
<body>
<p>&nbsp;</p>

<p>&nbsp;</p>
<style>
body {
  background-image: url('CR.webp');
  background-repeat: no-repeat;
    background-attachment: fixed;
  background-size: cover;
}
{
      min-height: 70%;
      }
      body,form {
      padding: 0;
      margin: 0;
      outline: none;
      font-family: Roboto, Arial, sans-serif;
      font-size: 14px;
      color: #FFFFFF;
      line-height: 22px;
     }
    .myDiv {
  border: 5px outset  #000000;
  background-color: #2F4F4F;
  text-align: center;
  font-family: Roboto, Arial, sans-serif;
      font-size: 14px;
      color: #FFFFFF;
}

</style>

<h1 style="text-align:center"><span style="color:#ffffff"><strong><span style="font-family:Arial,Helvetica,sans-serif"><span style="font-size:36px">CRUDE OIL PRICE PREDICTION&nbsp;</span></span></strong></span></h1>
<p>&nbsp;</p>
<p>&nbsp;</p>
<p>&nbsp;</p>
<form style="text-align:center">
  <label for="ENTER PRICE">ENTER PRICE:</label>
  <input type="date" id="PRICE" name="PRICE">
  <input type="submit">
</form>

<p>&nbsp;</p>
<p>&nbsp;</p>

<div class="myDiv">
  <h2>PREDICTED PRICE:</h2>
  <p>{______________}</P><br>
  </p>
</div>

</body>
</html>
```

## INDEX.HTML

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Home page</title>
    <link rel="stylesheet" href="style.css">
</head>
<body>

    <div class="main">
        <div class="navbar">
            <div class="icon">
                <h2 class="logo">CRUDE OIL</h2>
            </div>

            <div class="menu">
                <ul>
                    <li><a href="#">HOME</a></li>
                    <li><a href="#">ABOUT</a></li>
                    <li><a href="#">SERVICE</a></li>
                    <li><a href="#">CONTACT</a></li>
                </ul>
            </div>

            <div class="search">
```

```
                <input class="srch" type="search" name="" placeholder="Type To text">
                <a href="#"> <button class="btn">Search</button></a>
            </div>

        </div>
        <div class="content">
            <h1>Crude Oil<br><span>Price Prediction</span><br></h1>
            <p class="par"> Crude oil means a mixture of hydrocarbons that exists in
liquid phase in<br>
                natural underground reservoirs and remains liquid <br>at atmospheric
pressure
                after passing through <br>surface separating facilities.</p>

                <button class="cn"><a href="register.html">JOIN US</a></button>

                <div class="form">
                    <h2>Login Here</h2>
                    <input type="email" name="email" placeholder="Enter Email Here">
                    <input type="password" name="" placeholder="Enter Password Here">
                    <button class="btnn"><a href="#">Login</a></button>

                    <p class="link">Don't have an account<br>
                    <a href="#">Sign up </a> here</a></p>
                    <p class="liw">Log in with</p>

                    <div class="icons">
                        <a href="#"><ion-icon name="logo-facebook"></ion-icon></a>
                        <a href="#"><ion-icon name="logo-google"></ion-icon></a>
                    </div>

                </div>
                    </div>
                </div>
        </div>
    </div>
    <script src="https://unpkg.com/ionicons@5.4.0/dist/ionicons.js"></script>
</body>
```

```
</html>
```

## INDEX.HTML

```
<!DOCTYPE html>
<html>
    <head>
        <title>Registration Form</title>
        <link rel="stylesheet"
        href="register.css" type="text/css">
    </head>
    <body>
        <div class="main">
            <div class="register">
                <h2>Register Here</h2>
                <form id="register" method="post">
                    <label>First Name : </label>
                    <br>
                    <input type="text" name="fname"
                    id="name" placeholder="Enter Your First Name">
                    <br><br>
                    <label>Last Name : </label>
                    <br>
                    <input type="text" name="lname"
                    id="name" placeholder="Enter Your last Name">
                    <br><br>
                    <label>Your Age : </label>
                    <br>
                    <input type="number" name="age"
                    id="name" placeholder="How Old Are You">
                    <br><br>
                    <label>Email : </label>
                    <br>
                    <input type="email" name="email"
                    id="name" placeholder="Enter Your Valid Email">
                    <br><br>
                    <label>Gender : </label>
```

```
          <br>
          &nbsp; &nbsp; &nbsp;
          <input type="radio" name="gender"
          id="male">
          &nbsp;
          <span id="male">Male</span>
          &nbsp; &nbsp; &nbsp;
          <input type="radio" name="gender"
          id="female">
          &nbsp;
          <span id="female">Female</span>
          <br><br>
          <input type="submit" value="Submit"
          name="submit" id="submit">
        </form>
      </div>
    </div>
  </body>
</html>
```

### GITHUB & DEMO LINK

## GITHUB:

https://github.com/IBM-EPBL/IBM-Project-44943-1660727551

## DEMO LINK:

https://drive.google.com/file/d/1D_OyMGs_6zznSUMI3cHW6lwyhpecugOp/view?usp=share_link