CRUDE OIL PRICE

PREDICITION

A PROJECT REPORT

Submitted by

ANU PRATHA. K (950019106006)
MERCY GNANA PRIYA. R (950019106025)
ANJANA. N (950019106005)
MADHUMITHA. M (950019106021)

TEAM ID :PNT2022TMID4960

In partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING IN

ELECTRONICS AND COMMUNICATION

ENGINEERING

ANNA UNIVERSITY

PROGRESS THROUGH KNOWLEDGE

S.NO	TABLE OF CONTENT

1.INTRODUCTION

1.1.PROJECT OVERVIEW

Prediction of crude oil prices has been a wide topic for ages. People use their intuition and lot of techniques to guess the prices of crude oil. It takes a lot of knowledge about the crude oil to accurately predict it. Predicting the crude oil price is very significant in various economic, political and industrial areas, both for crude oil importer and exporter countries. Since the crude oil is important strategic resource around the globe; it has become the crucial commodity for the world's economy. Thus, prediction of prices of crude oil has always been considered as a very exciting and challenging task which drew the curiosity of professionals, researchers and organizations all over the world. Moreover, crude oil volatility has a critical impact on macroeconomic parameters such as such as inflation, unemployment, exchange rate, economic growth of countries whose economy rely heavily on crude oil export or import. Thus, crude oil price prediction can help governments of countries of the world in economic policymaking and make quick and operative economic decisions to hedge against probable risk in these economic parameters. Therefore, forecasting of crude oil prices is quite useful and is also the objective of this paper. In this, we have used LSTM based recurrent neural networks for the purpose of crude oil price prediction. Recurrent neural networks (RNN) have been proved to be one of the most powerful models for processing time-series based sequential data. LSTM is one of the most successful RNN architectures. LSTM introduces the memory cell, a unit of computation that replaces traditional artificial neurons in the hidden layer of the network. With these memory cells, networks are able to effectively associate memories and input remote in time, hence suit to grasp the structure of data dynamically over time with high prediction capacity. Recurrent neural network are a type of Neural Network where the output from previous step are fed as input to the current step. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases like when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to remember the previous words. Thus RNN came into existence, which solved this issue with the help of a Hidden Layer. The main and most important feature of RNN is Hidden state, which remembers some information about a sequence.

LSTMs have chain like structure with the repeating module having a different structure. There are four neural network layers which are interacting to each other in a special way. The key to LSTMs is the cell state, which is the horizontal line running through the top of the diagram. The cell state runs straight down the entire chain, with only some minor linear interactions. The information flows along it unchanged. The LSTM does have the ability to remove or add information to the cell state, carefully regulated by structures called gates. Gates are a way to optionally let information through. They are composed out of a sigmoid neural

network
layer and a point-wise multiplication operation. An LSTM has three of these gates, to protect and control the cell state.

1.2.PURPOSE:

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. Our work mainly focuses on applying Recurrent Neural Networks to predict the Crude Oil Price. This decision helps common people 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 price of the crude oil. So we would be implementing RNN (Recurrent Neural Network) with LSTM (Long Short Term Memory) to achieve the task. We will be experimenting with different types of models with varying number of epochs, look backs and other tuning methods.

2.LITERATURE SURVEY

2.1.EXISTING METHOD:

Crude oil forecasting is an important topic in financial and economic studies. Many studies have been performed to forecast the prices of crude oil.

- The first research about forecasting oil market is proposed by Amano (1987). The author used a small-scale econometric model for oil market prediction.
- Huntington (1994) utilized a sophisticated econometric model for predicting oil price in the 1980s.
- In another work, Gulen (1998) applied counteraction analysis to predict the WTI crude oil price.
- Barone-adesi et al. (1998) suggested a semiparametric approach based on the filtered historical simulation technique to forecast oil price.
- Kulkarni and Haidar [8] presented a model for forecasting crude oil spot price direction in the short- term, up to three days ahead based on multilayered feed-forward neural

network. They tested the relation between crude oil future prices and spot price. They found the evidence that future prices of crude oil contain new information about oil spot price detection.

- Hamdi and Aloui [9] performed a literature survey of numerous studies done on forecasting of crude oil price using artificial neural networks (ANN) until 2014. From the survey, they concluded that crude oil market is the most volatile commodity and forecasting oil price using nonlinear models such as ANN is the most suitable choice.
- Abdullah and Zeng [10] proposed Hierarchical Conceptual (HC) and Artificial Neural Networks- Quantitative (ANN-Q) model based on machine learning and computational intelligence techniques to predict the monthly WTI crude oil price for every barrel in USD. The results obtained from their study validated the effectiveness of data selection process by the proposed model which successfully extracts a comprehensive list of key factors that cause the crude oil price market to be volatile.
- Chen et al. [11] proposed a crude oil price forecasting model based on the deep learning model. They were able to analyze and model the crude oil price movement using the proposed deep learning model. They used the proposed model to capture the unknown complex non linear characteristics ofthe crude oil price movements. They evaluated the performance of the proposed model using the price data in the WTI crude oil markets.
- Huntington (1994) utilized a sophisticated econometric model for predicting oil price in the 1980s.
- In another work, Gulen (1998) applied co-integration analysis to predict the WTI crude oil price.
- Morana (2001) employed a semi-parametric approach investigated by Barone-adesi et al. (1998) to short-term forecast of Brent crude oil price.
- In another work, Tang and Hammoudeh (2002) utilized nonlinear regression to predict OPEC basket price.

Many Scientist and researcher have come across unique and variation of model for discovering and exploring and forecasting crude oil prices.

2.2.REFERENCES:

[1] Mohammad Reza Mahdiani and Ehsan Khamehchi, “A modified neural network model for predicting the crude oil price”, Intellectual Economics, vol. 10, no. 2, pp. 71-77, Aug. 2016.

[2] Manel Hamdi and Chaker Aloui, "Forecasting Crude Oil Price Using Artificial Neural Networks: A Literature Survey," Economics Bulletin, AccessEcon, vol. 35, no. 2, pp. 1339-1359, 2015.

[3] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Comput.,

vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[4] Hamdi and Aloui, "Machine learning approach for crude oil price prediction with Artificial Neural Networks-Quantitative (ANN-Q) model," The 2010International JointConference on Neural Networks (IJCNN), Barcelona, pp. 1-8, 2010.

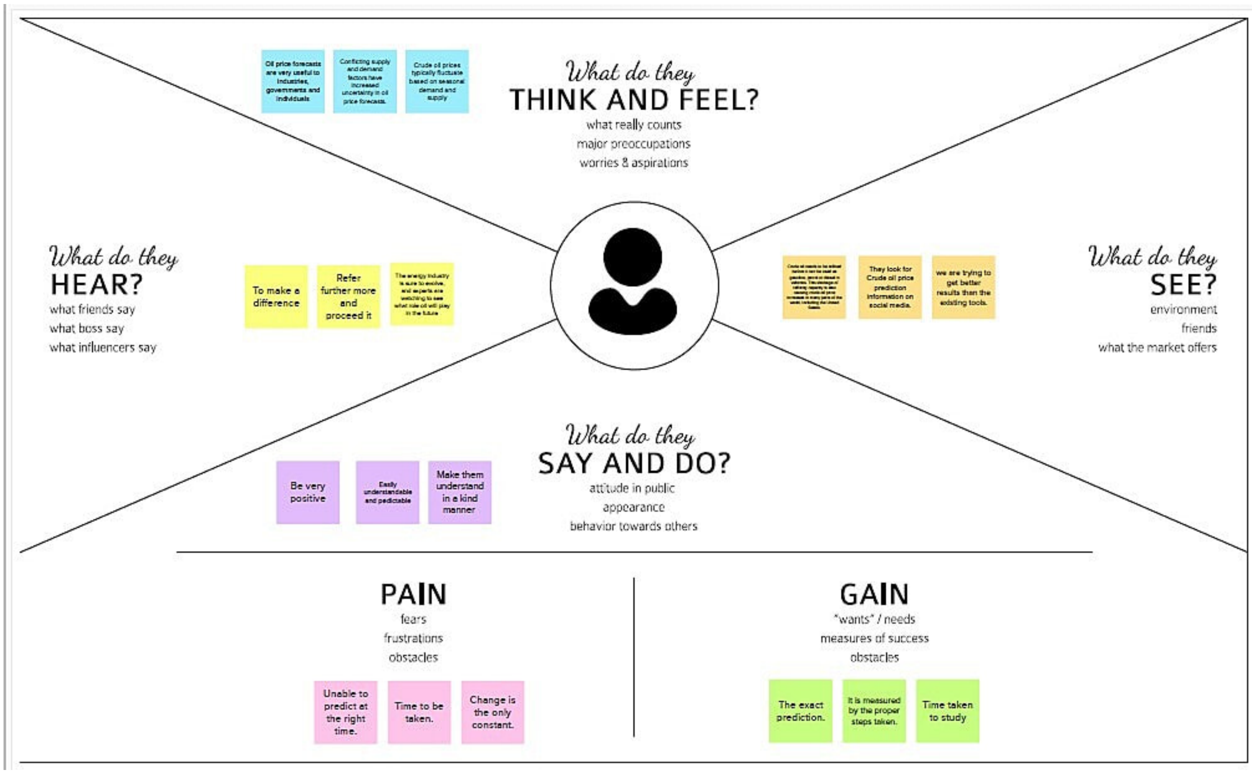
2.3.PROBLEM STATEMENT DEFINITION

Nowadays, the increased oil prices worldwide are having a great impact on all economic activities. Over the years there has been a fluctuation in petroleum prices, and a close consideration of the demand and supply side effects that sparked these price changes shows there is high probability that these changes will continue in the outlook period and beyond. Crude oil is amongst the most important resources in today's world. The evaporative nature of crude oil, its price prediction becomes extremely difficult and it is hard to precise with the same. Prediction of future crude oil price is considered a significant challenge due to the extremely complex and dynamic nature of the market and stakeholders perception.

Who does the problem affect?	Consumers, stakeholders
What are the boundaries of the problem?	Continuously capturing the unstable pattern of the crude oil prices.
What is the issue?	Crude-oil prices have always been volatile affecting the performance of the economy.
When does the issue occurs?	It occurs when the crude oil is influenced by many factors such as supply-and-demand gap, labour costs, amount of remaining resources, as well as stakeholders' perception.
Where is the issue occurring?	As India consumes around 2.2 million barrels of oil per day, the production is only about 0.8 million barrels per day . Therefore, 70% of its total oil consumption has to be imported. There occurs the issue.
Why is it important that we fix the problem?	As India ranks among the top 10 largest oil-consuming countries in the world. It is definitely mandatory to fix this problem.

3.IDEATION & PROPOSED SOLUTION

3.1.EMPATHY MAP CANVAS



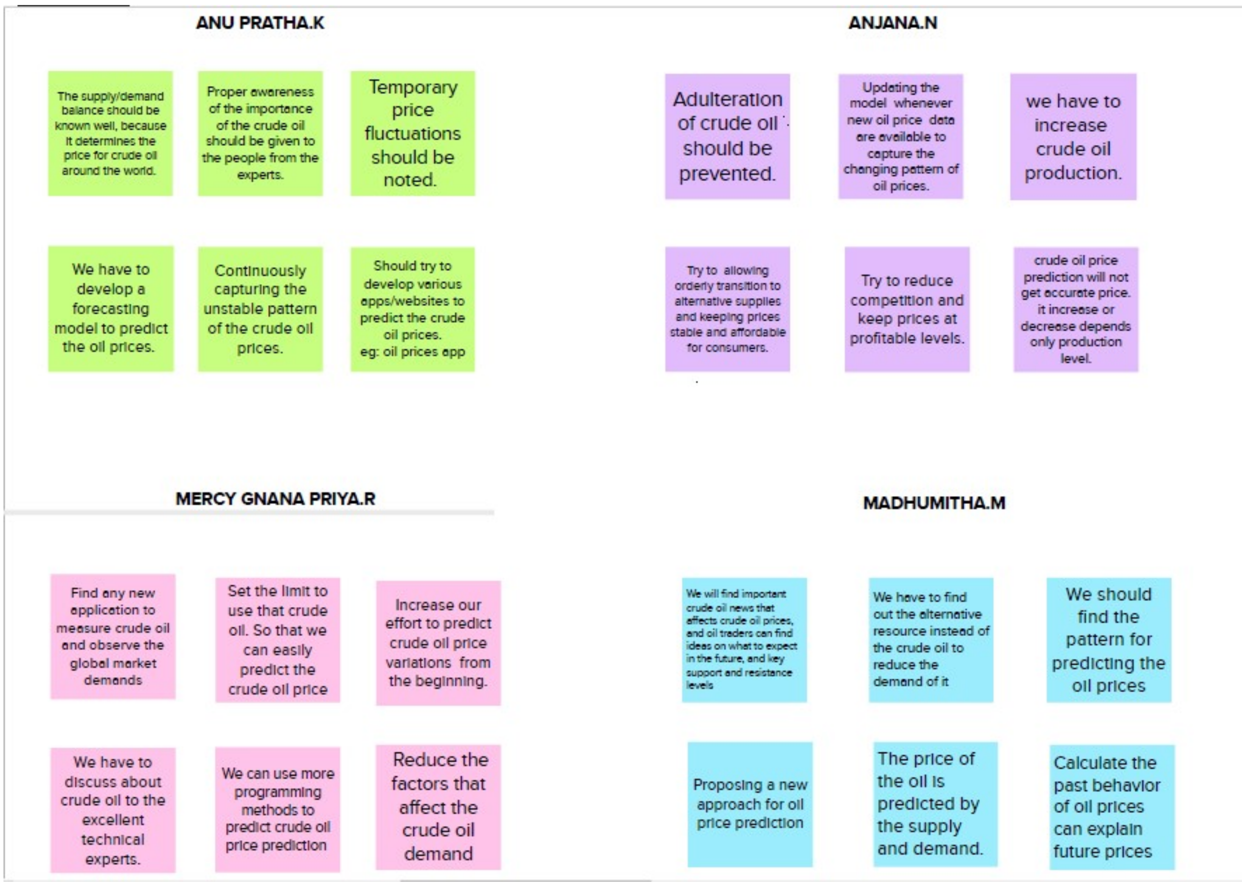
3.2.IDEATION & BRAINSTORMING

Problem Statements:

PROBLEM

How might we create a website to predict crude oil price prediction and to focus on exporters in exporting countries and generate revenue by selling our application.

Brainstorm:



ANU PRATHA.K
The supply/demand balance should be known well, because it determines the price for crude oil around the world.
Proper awareness of the importance of the crude oil should be given to the people from the experts.
Temporary price fluctuations should be noted.
We have to develop a forecasting model to predict the oil prices.
Continuously capturing the unstable pattern of the crude oil prices.
Should try to develop various apps/websites to predict the crude oil prices. eg: oil prices app
ANJANA.N
Adulteration of crude oil should be prevented.
Updating the model whenever new oil price data are available to capture the changing pattern of oil prices.
we have to increase crude oil production.
Try to allowing orderly transition to alternative supplies and keeping prices stable and affordable for consumers.
Try to reduce competition and keep prices at profitable levels.
crude oil price prediction will not get accurate price. it increase or decrease depends only production level.
MERCY GNANA PRIYA.R
Find any new application to measure crude oil and observe the global market demands
Set the limit to use that crude oil. So that we can easily predict the crude oil price
Increase our effort to predict crude oil price variations from the beginning.
We have to discuss about crude oil to the excellent technical experts.
We can use more programming methods to predict crude oil price prediction
Reduce the factors that affect the crude oil demand
MADHUMITHA.M
We will find important crude oil news that affects crude oil prices, and oil traders can find ideas on what to expect in the future, and key support and resistance levels
We have to find out the alternative resource instead of the crude oil to reduce the demand of it
We should find the pattern for predicting the oil prices
Proposing a new approach for oil price prediction
The price of the oil is predicted by the supply and demand.
Calculate the past behavior of oil prices can explain future prices

Group ideas:

The supply/demand balance should be known well, because it determines the price for crude oil around the world.

The price of the oil is predicted by the supply and demand.

Reduce the factors that affect the crude oil demand

Try to reduce competition and keep prices at profitable levels.

Increase our effort to predict crude oil price variations from the beginning.

We should find the pattern for predicting the oil prices

Should try to develop various apps/websites to predict the crude oil prices.
eg: oil prices app

Find any new application to measure crude oil and observe the global market demands

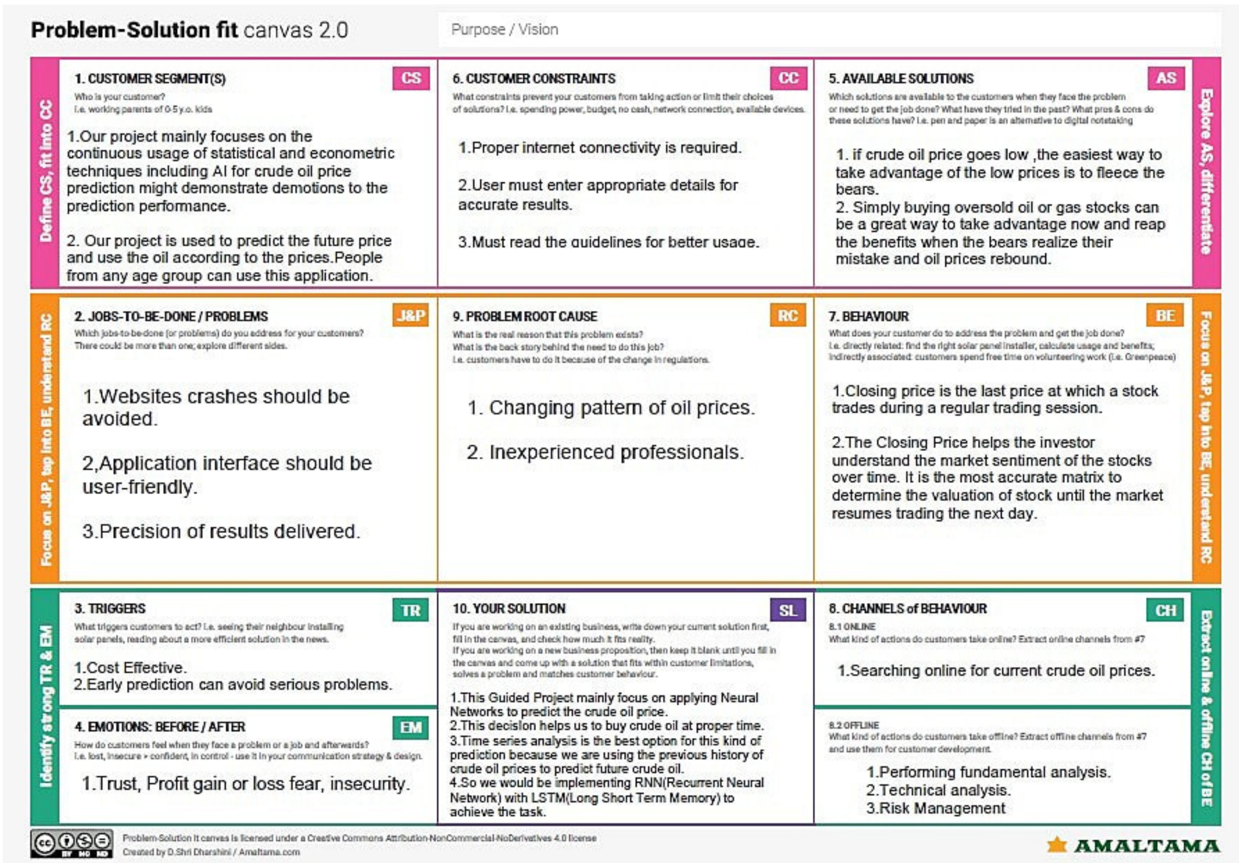
Priortize:

Set the limit to use that crude oil. So that we can easily predict the crude oil price
Adulteration of crude oil should be prevented.
We have to develop a forecasting model to predict the oil prices.
Find any new application to measure crude oil and observe the global market demands
Continuously capturing the unstable pattern of the crude oil prices.
Calculate the past behavior of oil prices can explain future prices
We should find the pattern for predicting the oil prices
We can use more programming methods to predict crude oil price prediction
Updating the model whenever new oil price data are available to capture the changing pattern of oil prices.
Proper awareness of the importance of the crude oil should be given to the people from the experts.
Temporary price fluctuations should be noted.
Try to allowing orderly transition to alternative supplies and keeping prices stable and affordable for consumers.
We will find important crude oil news that affects crude oil prices, and oil traders can find ideas on what to expect in the future, and key support and resistance levels
We have to discuss about crude oil to the excellent technical experts.
Importance
If each of these tasks could get done without any difficulty or cost, which would have the most positive impact?
we have to increase crude oil production.
We have to find out the alternative resource instead of the crude oil to reduce the demand of it
crude oil price prediction will not get accurate price. It increase or decrease depends only production level.
Reduce the factors that affect the crude oil demand
Proposing a new approach for oil price prediction
The supply/demand balance should be known well, because it determines the price for crude oil around the world.
Feasibility

3.3.PROPOSED SOLUTION

S.NO.	PARAMETER	DESCRIPTION
1.	Problem Statement (Problem to be solved)	The crude oil price has a huge impact on the world's economy. From the past few years, crude oil price fluctuates more than any other commodities prices. As the crude oil price depends on several external factors and there is high volatility predicting crude oil prices is very challenging.
2.	Idea/Solution Description	Continuously capturing the unstable pattern of the crude oil prices using new application.
3.	Novelty/Uniqueness	It is easy for the optimization algorithms to optimize the parameters to get the output. This helps the proposed model to illustrate lowest errors and better forecasting accuracy when compared to other models.
4.	Social Impact/Customer Satisfaction	● Crude oil prices are heavily influenced by non-market forces, including the Organization of the Petroleum Exporting Countries (OPEC) , which effectively acts as a multinational oil cartel. ● The reason why movements in oil price often surprise analysts is because there are hundreds of variables, each of them moving simultaneously in unpredictable ways.
5.	Business Model	Crude oil price prediction helps for the supply of crude oil is determined by the ability of oil companies to extract reserves from the ground and distribute them around the world.
6.	Scalability of the Solution	The volatility is continuously monitored for the price prediction of crude oil.

3.4.PROPOSED SOLUTION FIT



Problem-Solution fit canvas 2.0

Purpose / Vision

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) CS

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 CC

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 AS

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.

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

2. JOBS-TO-BE-DONE / PROBLEMS J&P

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 RC

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 BE

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.

Focus on J&P, tap into BE, understand RC

Identify strong TR & EM

3. TRIGGERS TR

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 EM

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 SL

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 CH

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

Extract online & offline CH of BE

AMALTAMA

4.REQUIREMENT ANALYSIS

4.1.FUNCTIONAL REQUIREMENTS

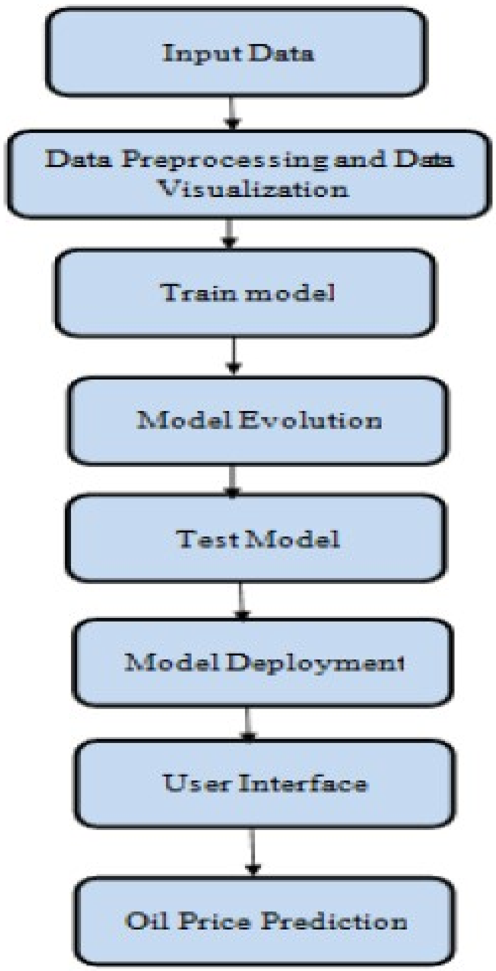
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form
FR-2	User Confirmation	Confirmation via SMS.
FR-3	Fetching input data	Give the model the input data.
FR-4	Generating Results	Prediction of Oil Prices.

4.2.NON-FUNCTIONAL REQUIREMENTS

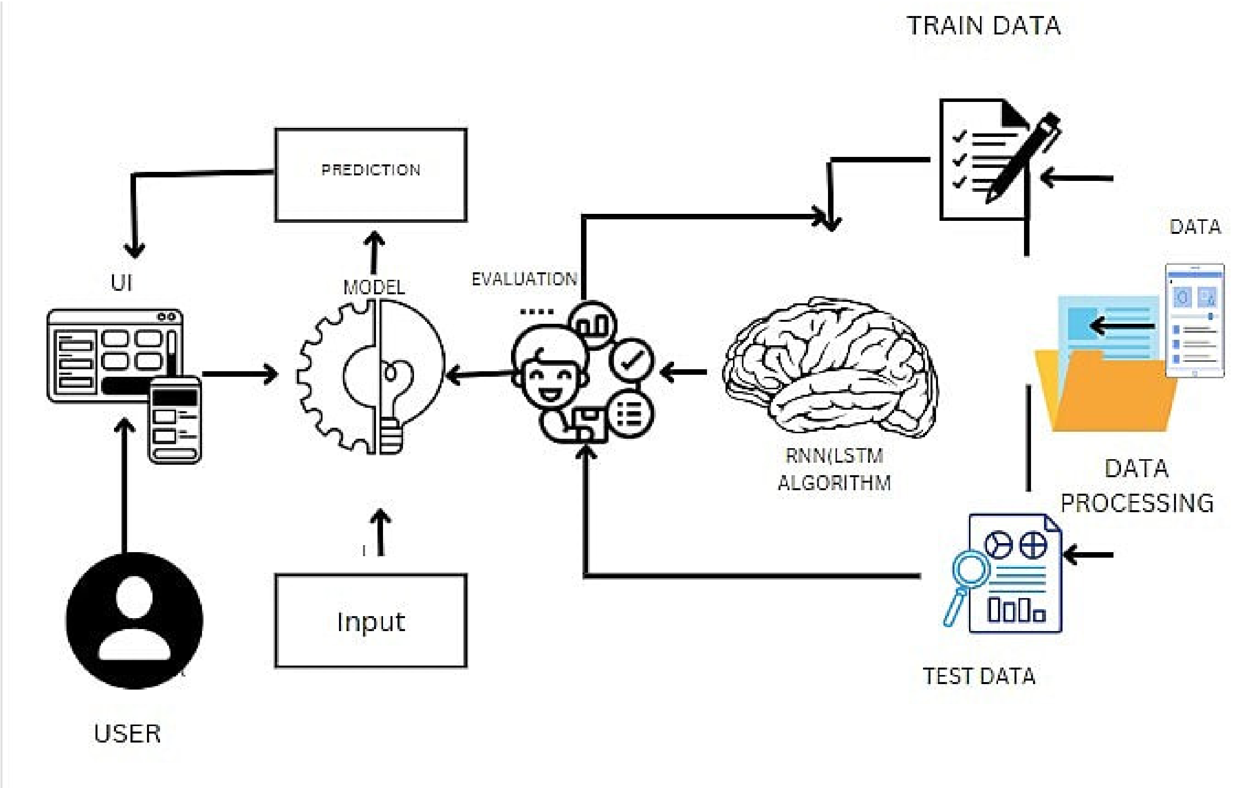
FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	People with no proper awareness about the regular change in price of crude oil can get to know, User interfaces are easy to use.
NFR-2	**Security**	Access permission for the local system must be given by the system's data administrator.
NFR-3	**Reliability**	The system will show the results regularly. Because there is very little variance from the prediction, the testing is highly dependable.
NFR-4	**Performance**	The front page will not exceed more than 2 seconds. Using LSTM networks gives highly performance.
NFR-5	**Availability**	The system is provided with the past ten days data.
NFR-6	**Scalability**	RNN (LSTM) network model works efficiently for large number of users.

5.PROJECT DESIGN

5.1.DATA FLOW DIAGRAMS



5.2.SOLUTION AND TECHNICAL ARCHIETECTURE



5.3.USER STORIES

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Front page	USN-1	As an user, I can view the front page of the website where the description about the application is available.	2	High	Anjana N and Madhumitha M
Sprint-2	Entering the past data	USN-2	As an user, I can enter the past ten days data.	2	Low	Anu Pratha K and Mercy Gnana Priya R
Sprint-3	Data Augmentation	USN-3	As a developer,the ANN dataset must be augmented.	1	High	Anjana N and Madhumitha M
Sprint-3	Classification of ANN model	USN-4	As a developer, the model must be classified.	3	High	Anu Pratha K and Mercy Gnana Priya R
Sprint-3	Compiling the model	USN-5	As a developer, the model must be compiled.	3	High	Anu Pratha K and Mercy Gnana Priya R
Sprint-4	Predicting the price of crude oil	USN-3	As an user, I can get the predicted value of crude oil price.	2	High	Anjana N,Madhumitha M,Anu Pratha K and Mercy Gnana Priya R

6.PROJECT PLANNING & SCHEDULING

6.1.SPRINT PLANNING & ESTIMATION

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Front page	USN-1	As an user, I can view the front page of the website where the description about the application is available.	2	High	Anjana N and Madhumitha M
Sprint-2	Entering the past data	USN-2	As an user, I can enter the past ten days data.	2	Low	Anu Pratha K and Mercy Gnana Priya R
Sprint-3	Data Augmentation	USN-3	As a developer,the ANN dataset must be augmented.	1	High	Anjana N and Madhumitha M
Sprint-3	Classification of ANN model	USN-4	As a developer, the model must be classified.	3	High	Anu Pratha K and Mercy Gnana Priya R
Sprint-3	Compiling the model	USN-5	As a developer, the model must be compiled.	3	High	Anu Pratha K and Mercy Gnana Priya R
Sprint-4	Predicting the price of crude oil	USN-3	As an user, I can get the predicted value of crude oil price.	2	High	Anjana N,Madhumitha M,Anu Pratha K and Mercy Gnana Priya R

MILESTONE & ACTIVITY LIST

Activity Number	Activity Name	Detailed Activity Description	Task Assigned	Status
1.1	Access Resources	Accesstheresources(courses)inproject dashboard.	All Members	COMPLETED
1.2	Rocket chat registration	Join the mentoring channel via platform& rocket-chat mobile app.	All Members	COMPLETED
1.3	Access workspace	Access the guided project workspace.	All Members	COMPLETED
1.4	IBM Cloud registration	Register on IBM Academic Initiative &Apply Feature code for IBM Cloud Credits.	All Members	COMPLETED
1.5	Project Repository Creation	Create GitHub account & collaboratewith Project Repository in project workspace.	All Members	COMPLETED
1.6	Environment Setup	Set-up the Laptop / Computers basedon the pre-requisites for each technology track.	All Members	COMPLETED
2.1	Literature survey	Literature survey on the selected project& Information Gathering.	All Members	COMPLETED
2.2	Technology Training	Attend the technology trainings as perthe training Calendar.	All Members	COMPLETED
2.3	Empathy Map	Prepare EmpathyMapCanvasto capture the user Pains &Gains, Preparelistofproblemstatements	All Members	COMPLETED
2.4	Technology Training	Attend the technology trainings as perthe training Calendar.	All Members	COMPLETED

2.5	Brainstorming	List the ideas (at least 4 per each team member) by organizing the brainstorm session and prioritize the ideas	All Members	COMPLETED
2.6	Technology Training	Attend the technology trainings as perthe training Calendar.	All Members	COMPLETED
3.1	Proposed Solution Document	Prepare theproposedsolution document, which includes the novelty, feasibility of idea, business model, social impact, scalability of solution, etc.	All Members	COMPLETED
3.2	Technology Training	Attend the technology trainings as perthe training Calendar.	All Members	COMPLETED
3.3	Problem - Solution fit & SolutionArchitecture	Prepare problem - solution fit document& Solution Architecture.	All Members	COMPLETED
3.4	Technology Training	Attend the technology trainings as per the training Calendar.	All Members	COMPLETED
4.1	Customer Journey Map	Prepare the customer journey maps to understand the user interactions & experiences with the application (entry to exit).	All Members	COMPLETED
4.2	Technology Training	Attend the technology trainings as perthe training Calendar.	All Members	COMPLETED
4.3	Functional Requirements& Data Flow Diagrams	Prepare theFunctionalRequirement Document &DataFlow Diagrams.	All Members	COMPLETED
4.4	Technology Architecture	Prepare Technology Architecture of the solution.	All Members	COMPLETED

4.5	Technology Training	Attend the technology trainings as per the training Calendar.	All Members	COMPLETED
5.1	Milestone&Activity List	Prepare Milestone & Activity List.	All Members	COMPLETED
5.2	Sprint Delivery Plan	Prepare Sprint Delivery Plan.	All Members	IN PROGRESS

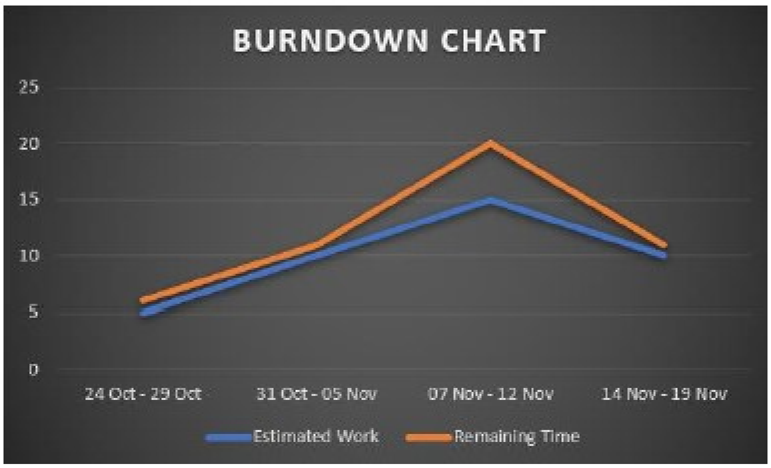
6	Data Collection	Collect datasets from different open sources likekaggle.com, data.gov, UCI machine learningrepository, etc.	All Members	COMPLETED
7.1	Data Pre-processing	Importing the Data Generator Library	All Members	COMPLETED
8.1	Model Building	Importing the model building libraries.	All Members	COMPLETED
8.2	Model Building	Training the Model	All Members	COMPLETED
8.3	Model Building	Save the model	All Members	COMPLETED
8.4	Model Building	Predictions	All Members	COMPLETED
9.1	Train LSTM Model on IBM	Register for IBM Cloud	All Members	COMPLETED
9.2	Deploy LSTM Model on IBM	Deploy the LSTM Model	All Members	COMPLETED

6.2.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	12	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	12	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	12	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	12	19 Nov 2022

BURNCHART



VELOCITY CHART

6.3.REPORTS FROM JIRA

7.CODING & SOLUTIONING

7.1.FEATURE 1

- Import the model building Libraries
- Initializing the model
- Adding LSTM Layers
- Adding Output Layer
- Configure the Learning Process
- Training the model
- Model Evaluation
- Save the Model
- Test the Model

LSTM

Long Short Term Memory , it is an advancement of RNN (Recurrent Neural Network) . To overcome the problems like vanishing gradient and exploiding gradient we go with LSTM. Unlike from a ANN , RNN can save the output on memory. So, there is a memory unit in LSTM. LSTM, it has default activation function(sigmoid and tanh). In LSTM method we use the previous weight, previous output, current output to obtain the next predicted output. LSTM don't invest on stock market.

- **Adding LSTM Layers**

For the LSTM layer, units is the number of LSTM neurons in the layer. 50 neurons will give the model high dimensionality, enough to capture the upwards and downward trends. Return sequences is True as we need to add another LSTM layer after the current one. Input shape corresponds to the number of time stamps and the number of indicators.

Adding Output Layers

The dense layer is a deeply connected neural network layer. It is the most common and frequently used layer. Finally, add the output layer. The output dimension is 1 since we are predicting 1 price each time. Understanding the model is a very important phase to properly use it for training and prediction purposes. Keras provides a simple method, summary to get the full information about the model and its layers.

Train The Model

For training the model. RNN weights are updated with a batch size . Try more batches and epochs if the loss of the model is not converging.
Epochs: an integer and number of epochs we want to train our model for.
Validation data can be either: an inputs and targets list, a generator. An inputs, targets, and sample_weights list which can be used to evaluate the loss and metrics for any model after any epoch has ended.

Model Evaluation

Finally, we need to check to see how well our model is performing on the test data.

Regression Evaluation Metrics:

1. Mean Squared Error (MSE):

MSE or Mean Squared Error is one of the most preferred metrics for regression problems. It is simply the average of the squared difference between the target value and the value predicted by the regression model. As it squares the differences, it penalizes even a small error which leads to over-estimation of how bad the model is. It is preferred more than other metrics because it is differentiable and hence can be optimized better.

2. Root Mean Square Error (RMSE):

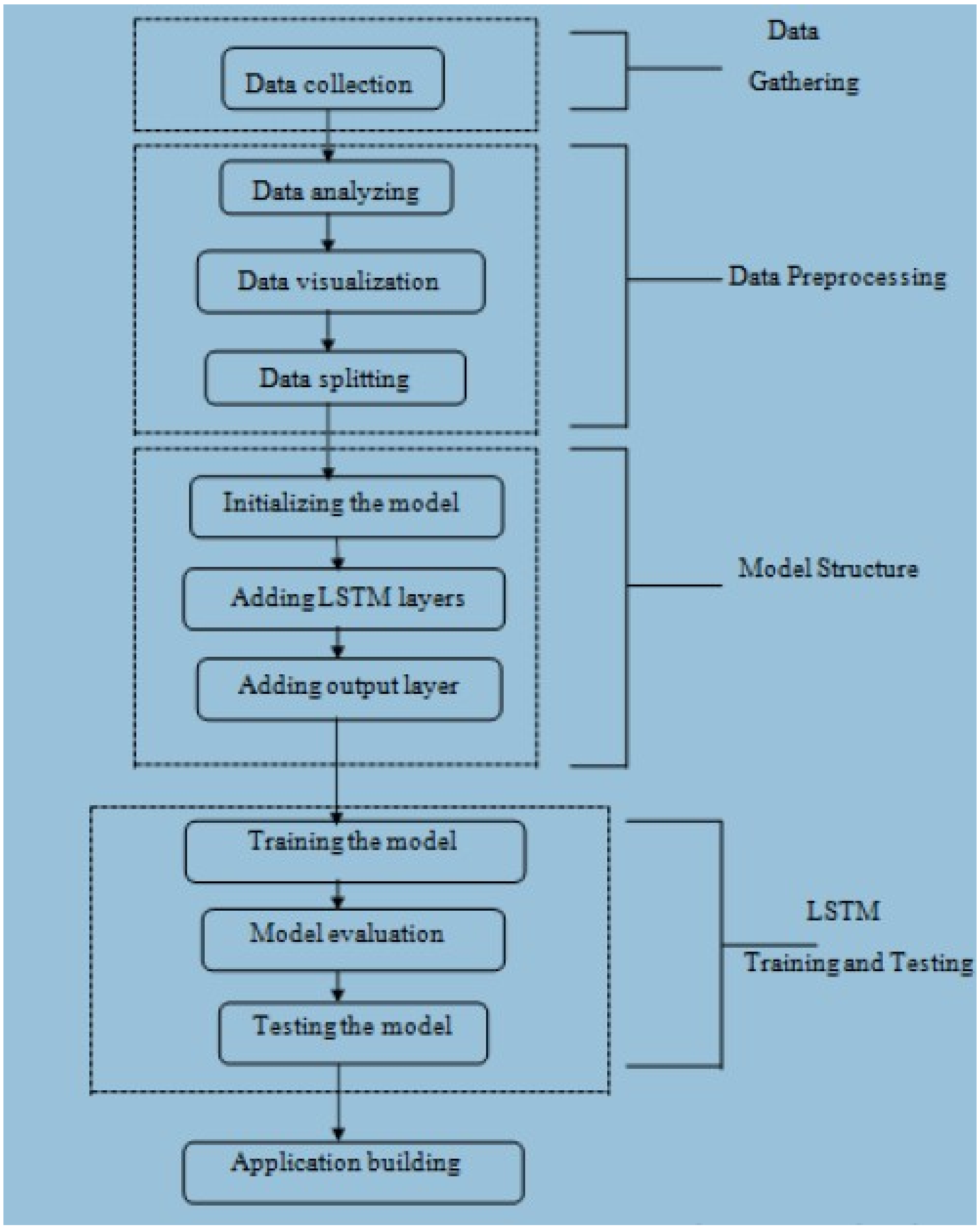
RMSE is the square root of the averaged squared difference between the target value and the value predicted by the model. It is preferred more in some cases because the errors are first squared before averaging which poses a high penalty on large errors. This implies that RMSE is useful when large errors are undesired.

Test the Model

Finally, we can generate predictions using the model for both the train and test to visualize the model. We must shift the predictions so that they align on the x-axis with the original dataset. Once prepared, the data is plotted, showing the original dataset in blue, the predictions for the training dataset in green, the predictions on the unseen test dataset in orange.

Prediction for the next 10 Days

Now let us predict the price of crude oil for the next 10 days. As the length of the test data is 2876, we are taking the previous 10 days input i.e., from index 2866 -2876 to predict 2867th day output. For predicting the next 10 days crude oil prices we consider n steps=10 We create the input for prediction, index starting from the date 10 days before the first date in the test dataset. Then, reshape the inputs to have only 1 column and predict using model predict predefined function.



Methodology Of proposed System

7.2.FEATURE 2(CODE)

Data_Preprocessing

Importing the libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

Importing the dataset

```
Data= pd.read_excel(r'C:\Users\user\Crude Oil\Crude Oil Prices Daily.xlsx')
```

Handling missing data

```
Data.isnull().any()

Data.isnull().sum()

Data.dropna(axis=0,inplace=True)

Data.isnull().sum()

Data_oil=Data.reset_index()['Closing Value']

Data_oil
```

Feature scaling

```
from sklearn.preprocessing import MinMaxScaler
scaler=MinMaxScaler(feature_range=(0,1))
Data_oil=scaler.fit_transform(np.array(Data_oil).reshape(-1,1))
```

Data visualization

```
plt.plot(Data_oil)
```

Splitting data into train and test

```
training_size=int(len(Data_oil)*0.65)
test_size=len(Data_oil)-training_size
train_Data,test_Data=Data_oil[0:training_size,:],Data_oil[training_size:len(Data_oil),:1]

training_size,test_size

train_Data.shape
```

Creating a dataset with sliding windows

```
#convert an array of values into a dataset matrix
def create_dataset(dataset, time_step=1):
    dataX, dataY = [], []
    for i in range(len(dataset)-time_step-1):

        a = dataset[i:(i+time_step), 0]
        dataX.append(a)
        dataY.append(dataset[i + time_step, 0])
    return np.array(dataX), np.array(dataY)

time_step = 10
X_train, y_train = create_dataset(train_Data, time_step)
X_test, ytest = create_dataset(test_Data, time_step)

print(X_train.shape), print(y_train.shape)

print(X_test.shape), print(ytest.shape)

X_train

X_train =X_train.reshape(X_train.shape[0],X_train.shape[1] , 1)
X_test = X_test.reshape(X_test.shape[0],X_test.shape[1] ,1)
```

Model_Building

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense
from tensorflow.keras.layers import LSTM

model=Sequential()

model.add(LSTM(50,return_sequences=True,input_shape=(10,1)))
model.add(LSTM(50,return_sequences=True))
model.add(LSTM(50))

model.add(Dense(1))

model.summary()

model.compile(loss='mean_squared_error',optimizer='adam')

model.fit(x_train,y_train,validation_data=(x_test,y_test),epochs=3,batch_size=64,verbose
=1)

##Transformback to original form
train_predict=scaler.inverse_transform(train_data)
test_predict=scaler.inverse_transform(test_data)

### Calculate RMSE performance metrics
import math
from sklearn.metrics import mean_squared_error
math.sqrt(mean_squared_error(train_data,train_predict))

from tensorflow.keras.models import load_model

model.save("crude_oil.h5")
!tar -zcvf crude-oil.tgz crude_oil.h5

### Plotting
look_back=10
```

```
trainpredictPlot = np.empty_like(Data_oil)
trainpredictPlot[:, :]= np.nan
trainpredictPlot[look_back:len(train_predict)+look_back, :] = train_predict

# shift test predictions for plotting
testPredictplot = np.empty_like(Data_oil)
testPredictplot[:,: ] = np.nan
testPredictplot[look_back:len(test_predict)+look_back, :] = test_predict

# plot baseline and predictions
plt.plot(scaler.inverse_transform(Data_oil))
plt.show()

len(test_data)

x_input=test_data[2866:].reshape(1,-1)
x_input.shape

temp_input=list(x_input)
temp_input=temp_input[0].tolist()

temp_input

lst_output=[]
n_steps=10
i=0
while(i<10):
   if(len(temp_input)>10):
#print(temp_input)
      x_input=np.array(temp_input[1:])
      print("{} day input {}".format(i,x_input))
      x_input=x_input.reshape(1,-1)
      x_input = x_input.reshape((1, n_steps, 1)) #print(x_input)
      yhat = model.predict(x_input, verbose=0)
      print("{} day output {}".format(i,yhat))
      temp_input.extend(yhat[0].tolist())
      temp_input=temp_input[1:] #print(temp_input)
```

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

day_new=np.arange(1,11)
day_pred=np.arange(11,21)
len(Data_oil)
plt.plot(day_new, scaler.inverse_transform(Data_oil[8206:]))
plt.plot(day_pred, scaler.inverse_transform(lst_output))
[]

df3=Data_oil.tolist()
df3.extend(lst_output)
plt.plot(df3[8100:])

df3=scaler.inverse_transform(df3).tolist()

plt.plot(scaler.inverse_transform(Data_oil))

!pip install watson-machine-learning-client

!pip install ibm_watson_machine_learning

from ibm_watson_machine_learning import APIClient
wml_credentials = {
    "url":"https://us-south.ml.cloud.ibm.com",
    "apikey":"M55SXB-FjuyroTNKBo1vl9lc1buIITQ7lWEPX54uggH5"
}

client = APIClient(wml_credentials)
```

```
client

client.spaces.get_details()

client.spaces.list()

space_uid = "6f47baab-1727-4d8d-aed0-9a4fb3dfe481"
space_uid

client.set.default_space(space_uid )

client.software_specifications.list()

software_space_uid =
client.software_specifications.get_uid_by_name("tensorflow_rt22.1-py3.9")
software_space_uid

model_details = client.repository.store_model(model="crude-oil.tgz", meta_props={
    client.repository.ModelMetaNames.NAME:"Crude Oil Price Prediction",
    client.repository.ModelMetaNames.TYPE:"tensorflow_2.7",
    client.repository.ModelMetaNames.SOFTWARE_SPEC_UID:software_space_uid
})

model_details

model_id = client.repository.get_model_id(model_details)
model_id
```

PYTHON CODE

app.py

```
import numpy as np
from flask import Flask, render_template,request

from tensorflow.keras.models import load_model


app = Flask(__name__)

model = load_model('crude_oil.h5',)




@app.route('/about')
def home1() :
   return render_template("index.html")


@app.route('/predict')
def home2() :
   return render_template("prediction.html")

@app.route('/')
def home() :
   return render_template("index.html")


@app.route('/login', methods=['POST'])

def login() :
   x_input=str(request.form['year'])
   x_input=x_input.split(',')
```

```
    print(x_input)
    for i in range(0, len(x_input)):
        x_input[i] = float(x_input[i])
    print(x_input)
    x_input=np.array(x_input).reshape(1, -1)
    temp_input=list(x_input)
    temp_input=temp_input[0].tolist()
    lst_output=[]
    n_steps = 10
    i = 0
    while(i<1):
        if(len(temp_input)>10):
            x_input=np.array(temp_input[1:])
            print("{} day input {}".format(i,x_input))
            x_input=x_input.reshape(1,-1)
            x_input=x_input.reshape((1, n_steps, 1))
            yhat = model.predict(x_input, verbose=0)
            print("{} day output {}".format(i,yhat))
            temp_input.extend(yhat[0].tolist())
            temp_input=temp_input[1:]
            lst_output.extend(yhat.tolist())
            i = i+1
        else:
            x_input = x_input.reshape((1, n_steps,1))
            yhat = model.predict(x_input, verbose=0)
            print(yhat[0])
            temp_input.extend(yhat[0].tolist())
            print(len(temp_input))
 lst_output.extend(yhat.tolist())
            i = i+1
    print(lst_output)
    return render_template("prediction.html",showcase = 'The next day predicted value
is:'+str(lst_output))


if __name__ == '__main__' :
    app.run(debug = True,port=5000)
```

8.TESTING

8.1.Test Cases:

8.2.User Acceptance Testing:

1.Purpose of Document

The purpose of this documentis to briefly explain the test coverageand open issues of the [ProductName] projectat the time of the release to User Acceptance Testing (UAT).

2.Defect Analysis

This reportshows the numberof resolved or closedbugs at each severity level, and howthey were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	6	1	2	1	10
Duplicate	1	0	0	0	1
External	1	0	0	1	2
Fixed	10	5	3	12	30
Not Reproduced	0	0	0	0	0
Skipped	0	0	0	1	1
Won't Fix	0	0	2	1	3
Totals	18	6	7	16	50

Test Case Analysis

This report shows the number of test cases that have passed, failed,and untested.

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	5	0	0	5
Client Application	30	0	0	30
Security	4	0	0	4
Outsource Shipping	12	0	0	12
Exception Reporting	5	0	0	5
Final Report Output	6	0	0	6

9.RESULTS

9.1.PERFORMANCSE METRICS:

S.No.	Parameter	Values
1.	Model Summary	Determining effective and efficient approach in predicting highly complex and volatile price like crude oil is a critical and challenging task in an economy of a nation. Most of the prediction techniques are designed focusing on statistical and econometrics point of view which has been helpful in numerous scenarios, however prediction using powerful AI tool like the LSTM of the DL is very rare. In this paper, we proposed a new crude oil price prediction technique based on complex network analysis and LSTM. In order to evaluate the effectiveness and robustness of the technique, we conducted the experiment on ten (10) different prices of crude oil across the world used by other researchers. From the experiment conducted we can conclude that, during the training process, the selection of batch size and number of LSTM layers has a great influence on the objective function value, fitting effect, and running time. The appropriate batch size and number of LSTM layers can effectively improve the model. Compared with the traditional and classic econometric prediction method, the model selects more datasets over a longer period of time as training samples. The LSTM prediction model has higher precision and wider application scenarios. The LSTM model can clearly predict the trend of crude oil price in the next time period.
2.	Accuracy	Training Accuracy -90% Validation Accuracy –above 90%

10.ADVANTAGES & DISADVANTAGES

ADVANTAGES:

- Crude oil price fluctuations have a far reaching impact on global economies and thus price forecasting can assist in minimising the risks associated with volatility in oil prices.

- The results is quite accurate with the accuracy up to 90%.

- Using LSTM, time series forecasting models can predict future values based on previous, sequential data. This provides greater accuracy for demand forecasters which results in better decision making for the business

DISADVANTAGES:

- LSTMs take longer to train.
- LSTMs require more memory to train.
- LSTMs are easy to overfit.
- Dropout is much harder to implement in LSTMs.
- LSTMs are sensitive to different random weight initializations.

11.CONCLUSION

This document report has clears that an LSTM network is better than other traditional neural network for forecasting prices as it aims in using back propagation model. Traditional neural network such as CNN on the other hand predicts the next outgoing but doesn’t necessarily save the previous data or connection which is based on feed-forwarding, in the sense the previous data is not necessary to predict the future data. LSTM focuses on storing the previous data and prediction which is rather encouraging and more approximate. The outcome derived are relatively encouraging. The results show that large look ups do not necessarily improve the accuracy of the predictions of crude oil prices. Hence it can be concluded, the model with single LSTM model is definitely the most accurate.

12.FUTURE SCOPE

In the coming future, fundamental indicators and market trends have been planned to be incorporated into a model which will help the proposed model perform more efficiently.

13.APPENDIX

SOURCE CODE:

Our project source code link:http://localhost:8888/tree/Crude%20Oil

Our Github link -https://github.com/IBM-EPBL/IBM-Project-46545-1660749455.git

DEMO VIDEO:

Demo video link - https://youtu.be/YkNBVUbAXOw

In the coming future, fundamental indicators and market trends have been planned to be incorporated into a model which will help the proposed model perform more efficiently.