CRUDE OIL PRICE PREDICTION

HX 8001-

Professional Readiness For Innovation, Employability and Entrepreneurship

IBM-Project-15664-165902682

TEAM ID : PNT2022TMID08321

Submitted by

MADALA SATISH (810419104057)
ALAVALA PRAKASH REDDY (810419104005)
BODDU BALAJI (810419104016)
GOTTAPU SAI JAGANNADHAM NAIDU (810419104026)

in partial fulfillment for the award of the degree

of

BATCHELOR OF ENGINNERING

IN

COMPUTER SCIENCE AND ENGINEERING

DHANALAKSHMI SRINIVASAN ENGINEERING COLLEGE

(AUTONOMOUS)

PERAMBALUR-621212

TABLE OF CONTENTS

CRUDE OIL PRICE PREDICTION

ABSTRACT

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. As the most traded commodity, oil, the repercussions are fairly substantial. Rising oil prices may even cause oil exporters to gain economic and political clout at the expense of oil importers. It is difficult to forecast the movements of the crude oil price because of its high degree of vacillation, unforeseen irregularity occurrences, and the intricate linkages involved between the market's elements. The price of crude oil is affected by a variety of variables. The major goal of this project is to employ neural networks to predict the price of crude oil. A solution to predict is offered using a machine learning and computational intelligence technique that combines historical quantitative data with qualitative data from expert opinion and news. This decision enables us to purchase crude oil at the appropriate moment. The greatest solution for this type of prediction is time series analysis since we are utilising past data on crude oil prices to forecast future crude oil prices. Therefore, implementing XGBOOST algorithm for predict the price value.

1. INTRODUCTION

1.1 PROJECT OVERVIEW

The international crude oil price has experienced the most significant volatility since the 2018 financial crisis. The oil market has taken on new features that affect the development of the global economy, national strategic security, and investor sentiment significantly. As the most important strategic resource around the globe, crude oil is the “key” commodity for the world’s economy. Therefore, forecasting crude oil price has always been considered as a very challenging task which drew the interest of researchers, practitioners and institutions. The price of oil is essentially determined by its supply demand. Therefore, it has been sometimes called “black gold” or “life blood”. Most countries heavily relied on imported crude oil in order to meet their energy needs. Oil exporting countries attempt to use oil as a weapon to perpetuate and wield political and economic power. Changes in world crude oil prices are becoming an increasing source of concern for government’s economic and organizational decisions. Knowing that every economic sector in the world is dependent on crude oil; hence any increase or decrease in the price of crude oil has a ripple effect on the global economy

1.2 PURPOSE

Numerous factors influence the price of crude oil. This project's main objective is to use neural networks to forecast crude oil prices. We are now able to buy crude oil at the right time thanks to this choice. Time series analysis is the best method for this kind of prediction because we are using historical data on crude oil prices to anticipate future crude oil prices. XGBOOST algorithm is being implemented as a result to estimate price value.

2. LITERATURE REVIEW

2.1 EXISTING PROBLEM

2.1.1 TITLE: Crude oil time series prediction model based on LSTM network with chaotic Henry gas solubility optimization, 2021

AUTHOR NAME: Aytac Altan

Estimating the price of crude oil, which is seen as an important resource for economic development and stability in the world, is a topic of great interest by policy makers and market participants. However, the chaotic and nonlinear characteristics of crude oil time series (COTS) make it difficult to estimate crude oil prices with high accuracy. To overcome these challenges, a new crude oil price prediction model is proposed in this study, which includes the long short-term memory (LSTM), technical indicators such as trend, volatility and momentum, and the chaotic Henry gas solubility optimization (CHGSO) technique. In the proposed model, features based on trend, momentum and volatility technical indicators are utilized. The features are obtained by using the trend indicators such as exponential moving average (EMA), simple moving average (SMA) and Kaufman's adaptive moving average (KAMA), the momentum indicators such as commodity channel index (CCI), rate of change (ROC) and relative strength index (RSI), and the volatility indicators such as average true range (ATR), volatility ratio (VR) and highest high-lowest low (HHLL). These indicators are obtained separately for the West Texas Intermediate (WTI) and Brent COTS. Especially, including the volatility indicator in the model is important in terms of the robustness of the proposed model. The results show that the proposed prediction model copes with the chaoticity and nonlinear dynamics of both WTI and Brent COTS.

2.1.2 TITLE: Prediction Model for the Viscosity of Heavy Oil Diluted with Light Oil Using Machine Learning Techniques, 2022

AUTHOR NAME: Xiaodong Gao

Due to the presence of asphaltene, the flow assurance of high viscosity crude oil becomes more challenging and costly to produce in wellbores and pipelines. One of the most effective ways to reduce viscosity is to blend heavy oil with light oil. However, the viscosity measurement of diluted heavy crude is either time-consuming or inaccurate. This work aims to develop a more accurate viscosity model of diluted heavy crude based on machine learning techniques. A multilayer neural network is used to predict the viscosity of heavy oil diluted with lighter oil. The input data used in the training include temperature, light oil viscosity, heavy oil viscosity, and dilution ratio. In this modeling process, 156 datasets were retrieved from the available iterature of various heavy-oil fields in China. Part of the data (80%) is used to train the developed models using Adam optimizer algorithms, while the other part of the data (20%) is used to predict the viscosity of heavy oil diluted with lighter. The performance and accuracy of the machine learning models were tested and compared with the existing viscosity models. It was found that the new model can predict the viscosity of diluted heavy oil with higher accuracy, and it performs better than other models. The absolute average relative error is 10.44%, the standard deviation of the relative error is 8.45%, and the coefficient of determination is R2 = 0.95. The viscosity predicted by the neural network outperformed existing correlations by the statistical analysis used for the datasets available in the literature. Therefore, the method proposed in this paper can better estimate the viscosity of diluted heavy crude oil and has important promotion value.

2.1.3 TITLE: An Explainable Machine Learning Framework for Forecasting Crude Oil Price during the COVID-19 Pandemic, 2022

AUTHOR NAME: Xinran Gao

Financial institutions, investors, central banks and relevant corporations need an efficient and reliable forecasting approach for determining the future of crude oil price in an effort to reach optimal decisions under market volatility. With oil being one of the most vital commodities in the world nowadays, fluctuations in the price of crude oil can have a substantial impact on global economic stability and development. As a globally priced commodity, it is not the current supply of crude oil or current economic growth that decides the price of crude oil, but rather the market’s expectations of future supply and demand trends, which largely determine the direction of crude oil price fluctuations. This paper presents an innovative research framework for precisely predicting crude oil price movements and interpreting the predictions. First, it compares six advanced machine learning (ML) models, including two state-of-the-art methods: extreme gradient boosting (XGB) and the light gradient boosting machine (LGBM). Second, it selects novel data, including user search big data, digital currencies and data on the COVID-19 epidemic. The empirical results suggest that LGBM outperforms other alternative ML models. Finally, it proposes an interpretable framework for facilitating decision making to interpret the prediction results of complex ML models and for verifying the importance of various features affecting crude oil price. The results of this paper provide practical guidance for participants in the crude oil market.

2.1.4 TITLE: Analysis and forecasting of crude oil price based on the variable selection-LSTM integrated model, 20212

AUTHOR NAME: Quanying Lu1

Since 2014, the international crude oil price has experienced the most significant volatility since the 2018 financial crisis. The oil market has taken on new features that affect the development of the global economy, national strategic security, and investor sentiment significantly. Especially as the primary alternative energy resources, the US tight oil production has been significant macroeconomic effects on the oil price. In recent years, the crude oil market has entered a new period of development and the core influence factors of crude oil have also been a change. Thus, we develop a new research framework for core influence factors selection and forecasting. Firstly, this paper assesses and selects core influence factors with the elastic-net regularized generalized linear Model (GLMNET), spike-slab lasso method, and Bayesian model average (BMA). Secondly, the new machine learning method long short-term Memory Network (LSTM) is developed for crude oil price forecasting. Then six different forecasting techniques, random walk (RW), autoregressive integrated moving average models (ARMA), elman neural Networks (ENN), ELM Neural Networks (EL), walvet neural networks (WNN) and generalized regression neural network Models (GRNN) were used to forecast the price. Finally, we compare and analyze the different results with root mean squared error (RMSE), mean absolute percentage error (MAPE), directional symmetry (DS). Our empirical results show that the variable selection-LSTM method outperforms the benchmark methods in both level and directional forecasting accuracy.

2.1.5 TITLE: Application of machine learning and artificial intelligence in oil and gas industry, 2021

AUTHOR NAME: Anirbid Sircar

Oil and gas industries are facing several challenges and issues in data processing and handling. Large amount of data bank is generated with various techniques and processes. The proper technical analysis of this database is to be carried out to improve performance of oil and gas industries. This paper provides a comprehensive state-of-art review in the field of machine learning and artificial intelligence to solve oil and gas industry problems. It also narrates the various types of machine learning and artificial intelligence techniques which can be used for data processing and interpretation in different sectors of upstream oil and gas industries. The achievements and developments promise the benefits of machine learning and artificial intelligence techniques towards large data storage capabilities and high efficiency of numerical calculations. In this paper a summary of various researchers work on machine learning and artificial intelligence applications and limitations is showcased for upstream and sectors of oil and gas industry. The existence of this extensive intelligent system could really eliminate the risk factor and cost of maintenance. The development and progress using this emerging technologies have become smart and makes the judgement procedure easy and straightforward. The study is useful to access intelligence of different machine learning methods to declare its application for distinct task in oil and gas sector.

2.2 REFERENCES

[1]. Aytac Altan, Crude oil time series prediction model based on LSTM network with chaotic Henry gas solubility optimization, 2021

[2]. Xiaodong Gao, Prediction Model for the Viscosity of Heavy Oil Diluted with Light Oil Using Machine Learning Techniques, 2022

[3]. Xinran Gao, An Explainable Machine Learning Framework for Forecasting Crude Oil Price during the COVID-19 Pandemic, 2022

[4]. Quanying Lu1, Analysis and forecasting of crude oil price based on the variable selection-LSTM integrated model, 20212

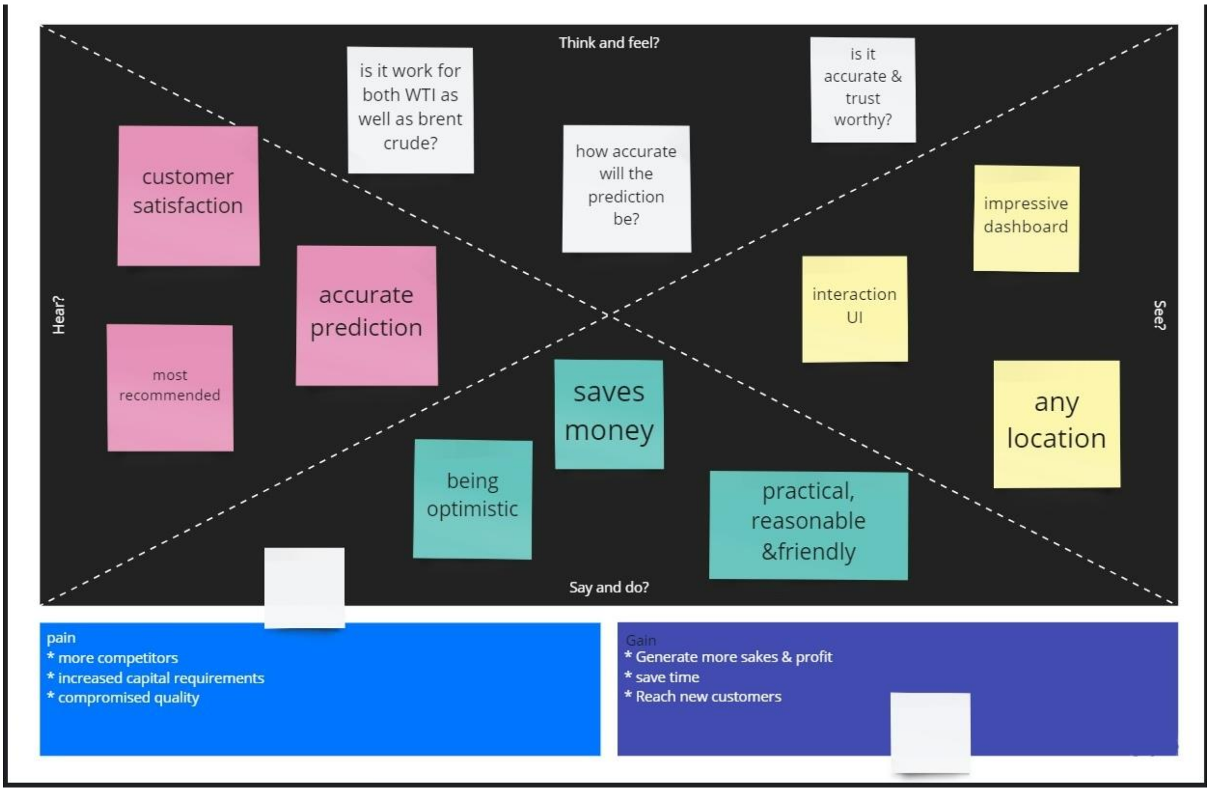
[5]. Anirbid Sircar, Application of machine learning and artificial intelligence in oil and gas industry, 2021

2.3 PROBLEM STATEMENT DEFINITION

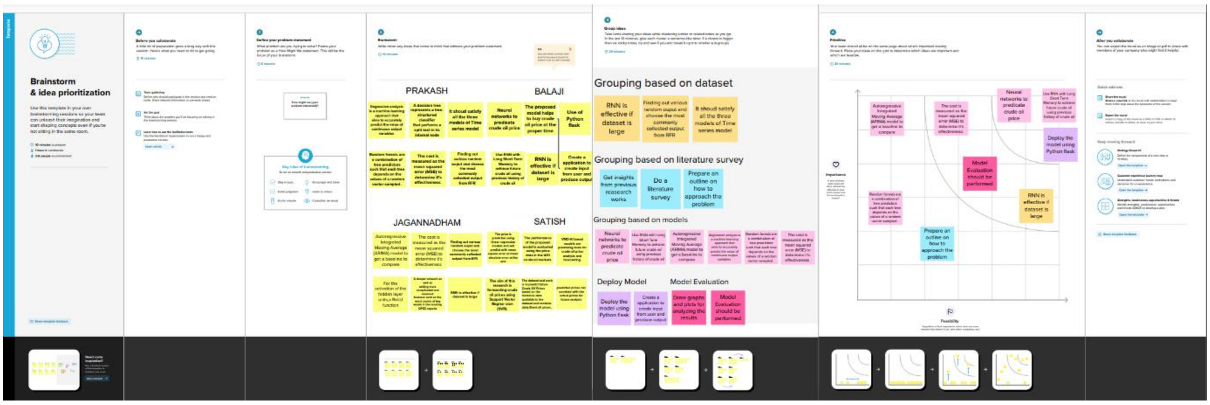
In existing system, the oil market's internal and external environments are evolving, and there are now a wide variety of complicated impacting elements. It is getting harder to identify practical elements and forecast oil prices as the variables driving global oil prices become increasingly complicated. Numerous previous types of research on the forecasting of crude oil prices demonstrate that the forecasting outcomes depend on the frequency and interval choice of the modelling sample data. Since crude oil's financial characteristics are steadily improving, the volatility of crude oil prices will inevitably have an impact on the profits of oil firms and the actions of investors. Therefore, it is crucial to accurately capture the current trend in global oil prices as well as conduct a systematic examination of the features of complicated international oil markets. It is challenging to identify the elements that most strongly influence the price of oil.

3. IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION & BRAINSTORMING



3.3 PROPOSED SOLUTION

Crude oil is the primary supply material for fuels and chemicals. As the main energy supply for industrialization, it plays a very vital role in the economic and industrial development of countries all over the world, and it is one of the most valuable natural resources affecting economic development. Predicting crude oil prices in a new and effective method is one problem that academics and practitioners are very concerned about all the time. To better analyze the changing trend of the crude oil market, it is necessary to determine the main factors affecting the price, determine the impact of each factor on price, and establish a forecasting model finally. The implementation is on predicting the crude oil prices for using machine learning Algorithm and plotting the graph based on prediction. Use XGBOOST algorithm which is feasible to some extend for the prediction of the crude oil prices. The Prediction of crude oil rates based on the previous datasets on the data and prices as the feature list are inputs and target list are predicted values. It can provide reference and theoretical support for the formulation of national energy security strategy and enterprise avoidance of market risks.

3.4 PROBLEM SOLUTION FIT

There is no organised means to gather future data. The system will forecast crude oil prices by entering historical data using a machine learning model and pre-trained dataset. Users should be able to interact with the suggested system both offline and online.

4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

Framework Creation

In this module, design the framework for Predicting crude oil price to get answers without any human assistance. Admin can train keywords with answers for future processing.

Dataset

In this module, oil price datasets should be uploaded as CSV files. Additionally, the information is kept in a database for later use. The dataset contains information on the crude oil price detail such as date, month, and year and oil price. These numbers are preserved as integer values and acquired from the Kaggle website.

Pre-processing

Modifying the raw data to make it appropriate for research on machine learning models is a vital stage in the process. To understand the dataset, we must use charts to plot the data, such as the histogram. As a result of the vast number of results for used cars, it is established that the dataset contains a significant number of outliers. Automobiles with the newest model years and low mileage (measured in kilometres travelled) command a higher price. However, other factors, such as accident history and physical condition, did not justify this premium sell.

Price Prediction

In this module, user can give the inputs in the user defined paramaters and get the current price details. Using machine learning model and trained dataset, the system can predict the accurate price details of the crude oil.

4.2 NON FUNCTIONAL REQUIREMENTS

Usability

The system shall allow the users to access the system with pc using web application. The system uses a web application as an interface. The system is user friendly which makes the system easy

Availability

The system is available 100% for the user and is used 24 hrs a day and 365 days a year. The system shall be operational 24 hours a day and 7 days a week.

Scalability

Scalability is the measure of a system's ability to increase or decrease in performance and cost in response to changes in application and system processing demands.

Security

A security requirement is a statement of needed security functionality that ensures one of many different security properties of software is being satisfied.

Performance

The information is refreshed depending upon whether some updates have occurred or not in the application. The system shall respond to the member in not less than two seconds from the time of the request submittal. The system shall be allowed to take more time when doing large processing jobs. Responses to view information shall take no longer than 5 seconds to appear on the screen.

Reliability

The system has to be 100% reliable due to the importance of data and the damages that can be caused by incorrect or incomplete data. The system will run 7 days a week. 24 hours a day.

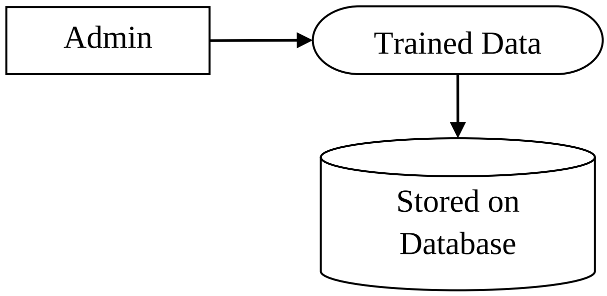
5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAMS

A two-dimensional diagram explains how data is processed and transferred in a system. The graphical depiction identifies each source of data and how it interacts with other data sources to reach a common output. Individuals seeking to draft a data flow diagram must identify external inputs and outputs, determine how the inputs and outputs relate to each other, and explain with graphics how these connections relate and what they result in. This type of diagram helps business development and design teams visualize how data is processed and identify or improve certain aspects.

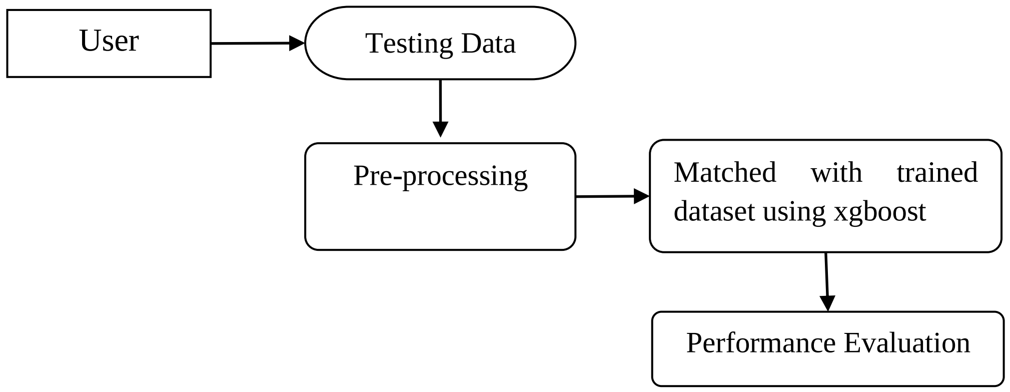
LEVEL 0

The Level 0 DFD shows how the system is divided into 'sub-systems' (processes), each of which deals with one or more of the data flows to or from an external agent, and which together provide all of the functionality of the system as a whole. It also identifies internal data stores that must be present in order for the system to do its job, and shows the flow of data between the various parts of the system.



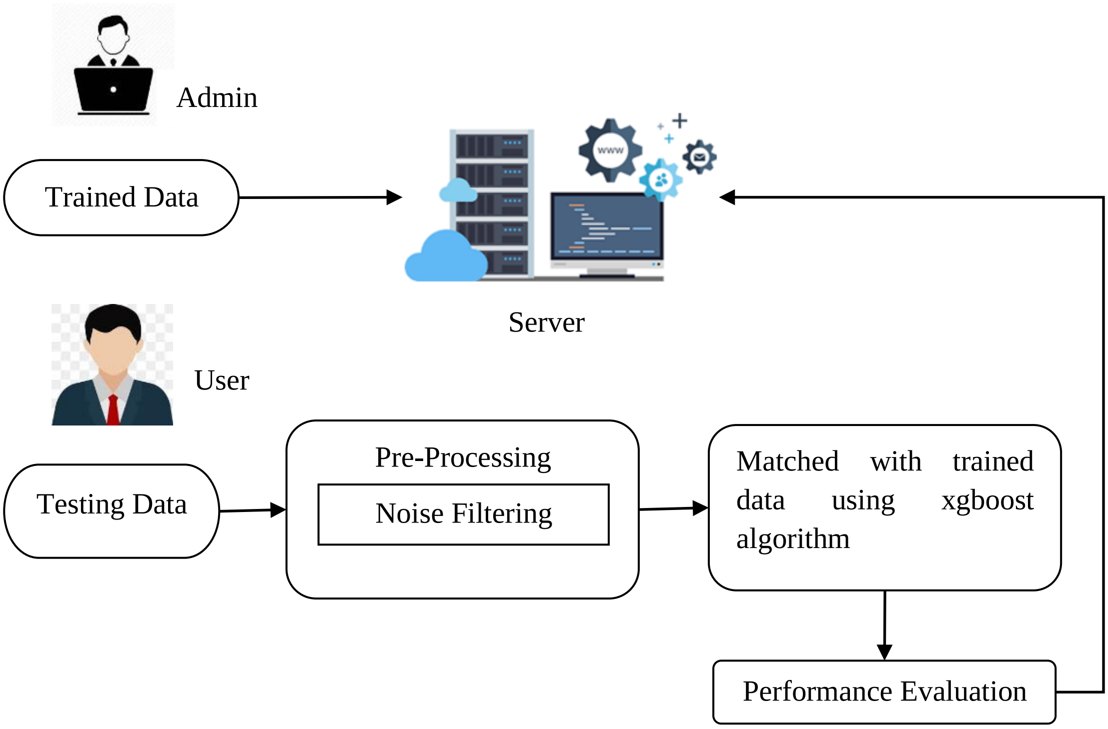
LEVEL 1

The next stage is to create the Level 1 Data Flow Diagram. This highlights the main functions carried out by the system. As a rule, to describe the system was using between two and seven functions - two being a simple system and seven being a complicated system. This enables us to keep the model manageable on screen or paper.



5.2 SOLUTION & TECHNICAL ARCHITECTURE

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. System architecture can comprise system components, the externally visible properties of those components, the relationships (e.g. the behavior) between them. It can provide a plan from which products can be procured, and systems developed, that will work together to implement the overall system. There have been efforts to formalize languages to describe system architecture, collectively these are called architecture description languages (ADLs).



5.3 USER STORIES

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	High	A.PRAKASH REDDY
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	High	M.SATISH
Sprint-1	Login	USN-3	As a user, I can log into the application by entering email & password.	High	G.SAI JAGANNADHAM NAIDU
Sprint-2	Input Necessary Details	USN-4	As a user, I can give Input Details to Predict Likeliness of crude oil	High	B.BALAJI
Sprint-2	Data Pre-processing	USN-5	Transform raw data into suitable format for prediction.	High	M.SATISH
Sprint-3	Prediction of Crude Oil Price	USN-6	As a user, I can predict Crude oil using machine learning model.	High	A.PRAKASH REDDY
Sprint-3		USN-7	As a user, I can get accurate prediction of crude oil	Medium	B.BALAJI
Sprint-4	Review	USN-8	As a user, I can give feedback of the application.	High	G.SAI JAGANNDHAM NAIDU

6. PROJECT PLANNING & SCHEDULING

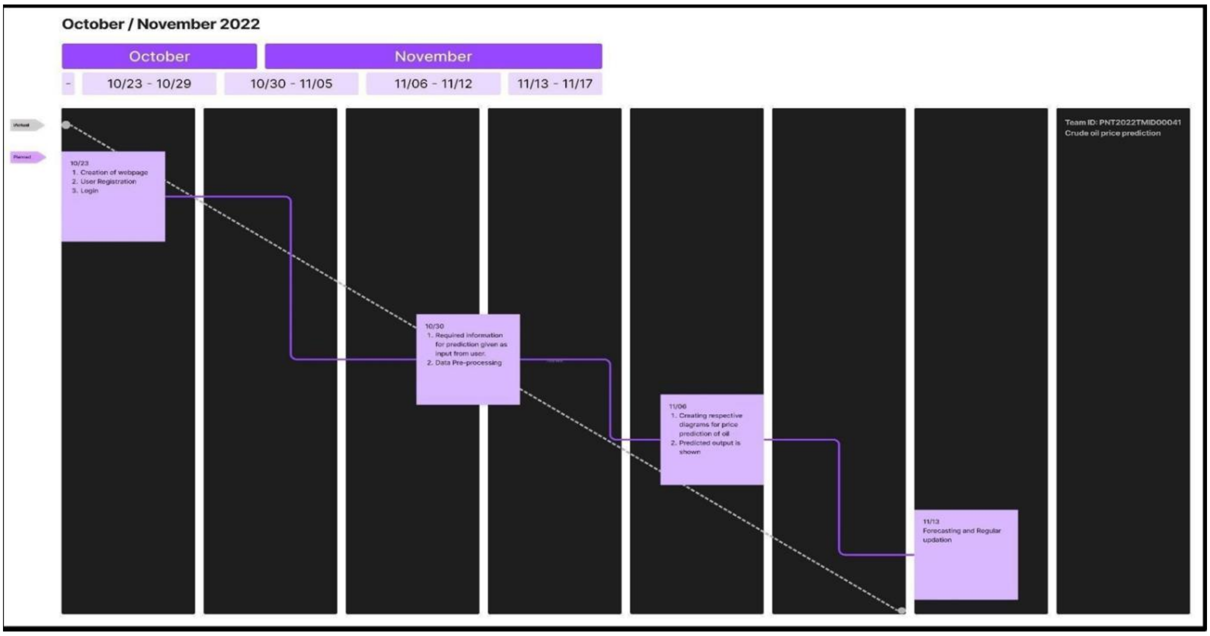
6.1 SPRINT PLANNING & ESTIMATION

Sprint	Duration	Sprint Start Date	Sprint End Date (Planned)	Sprint Release Date (Actual)
Sprint-1	6 Days	24 Oct 2022	29 Oct 2022	29 Oct 2022
Sprint-2	6 Days	31 Oct 2022	05 Nov 2022	
Sprint-3	6 Days	07 Nov 2022	12 Nov 2022	
Sprint-4	6 Days	14 Nov 2022	19 Nov 2022	

6.2 SPRINT DELIVERY SCHEDULE

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	High	A.PRAKASH REDDY
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	High	M.SATISH
Sprint-1	Login	USN-3	As a user, I can log into the application by entering email & password.	High	G.SAI JAGANNADHAM NAIDU
Sprint-2	Input Necessary Details	USN-4	As a user, I can give Input Details to Predict Likeliness of crude oil	High	B.BALAJI
Sprint-2	Data Pre-processing	USN-5	Transform raw data into suitable format for prediction.	High	M.SATISH
Sprint-3	Prediction of Crude Oil Price	USN-6	As a user, I can predict Crude oil using machine learning model.	High	A.PRAKASH REDDY
Sprint-3		USN-7	As a user, I can get accurate prediction of crude oil	Medium	B.BALAJI
Sprint-4	Review	USN-8	As a user, I can give feedback of the application.	High	G.SAI JAGANNADHAM NAIDU

6.3 REPORTS FROM JIRA



7. CODING & SOLUTIONING

7.1 FEATURE 1

Home.html(code):

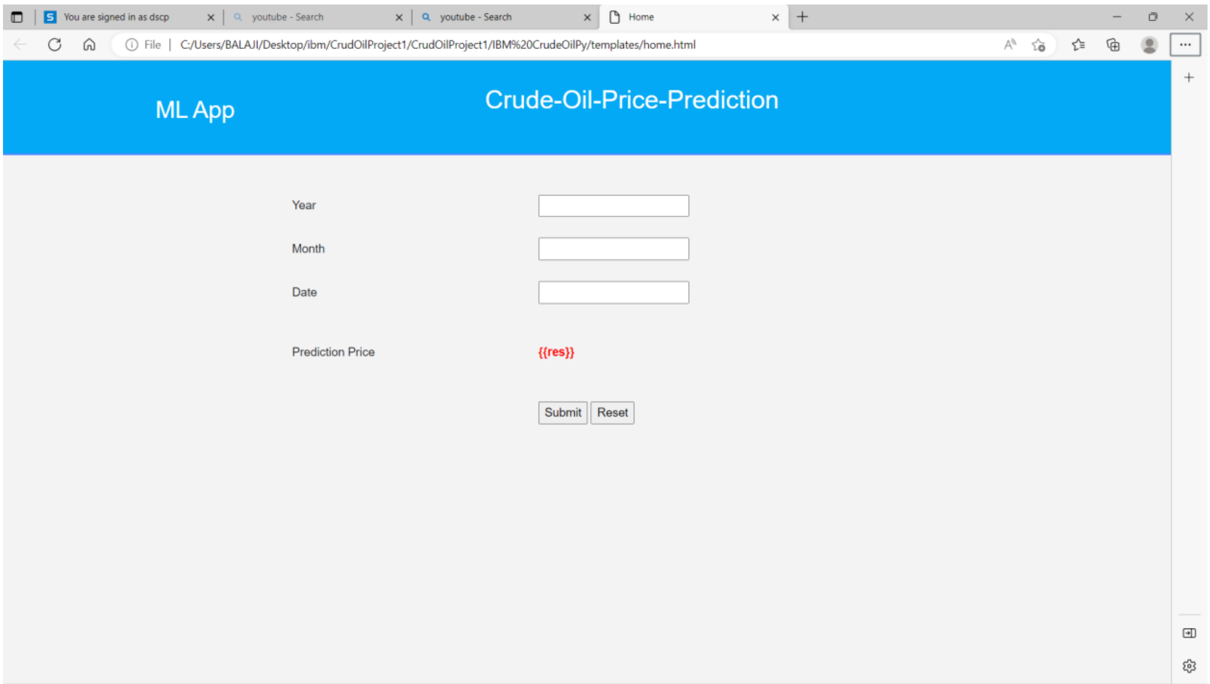
```
<!DOCTYPE html>
<html>
<head>
  <title>Home</title>
  <!-- <link rel="stylesheet" type="text/css" href="../static/css/styles.css">
-->
  <link rel="stylesheet" type="text/css" href="{{ url_for('static',
filename='css/styles.css') }}">
  <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.css
" integrity="sha384-
ggOyR0iXCbMQv3Xipma34MD+dH/1fQ784/j6cY/iJTQUOhcWr7x9JvoRxT2MZw1T"
crossorigin="anonymous">
    <style>
        body{
  font:15px/1.5 Arial, Helvetica,sans-serif;
  padding: 0px;
  background-color:#f4f3f3;
}

.container{
  width:100%;
  margin: auto;
  overflow: hidden;
}

header{
  background:#03A9F4;#35434a;
  border-bottom:#448AFF 3px solid;
  height:120px;
  width:100%;
  padding-top:30px;

}

.main-header{
      text-align:center;
      background-color: blue;
      height:100px;
      width:100%;
      margin:0px;
    }
```

```
#brandname{
  float:left;
  font-size:30px;
  color: #fff;
  margin: 10px;
}

header h2{
  text-align:center;
  color:#fff;

}



.btn-info {background-color: #2196F3;
  height:40px;
  width:100px;} /* Blue */
.btn-info:hover {background: #0b7dda;}


.resultss{
  border-radius: 15px 50px;
    background: #345fe4;
    padding: 20px;
    width: 200px;
    height: 150px;
}

    .style1 {
  color: #FF0000;
  font-weight: bold;
}
    </style>
</head>
<body>

  <header>
    <div class="container">
    <div id="brandname">
      ML App
    </div>
    <h2>Crude-Oil-Price-Prediction  </h2>

  </div>
  </header>

  <div class="ml-container">
```

```html
    <form id="form1" name="form1" method="post" action="/result">
  <table width="50%"  border="0" align="center">

        <tr>
          <td width="42%" > </td>
          <td width="27%"> </td>
        </tr>
        <tr>
          <td > </td>
          <td> </td>
        </tr>
        <tr>
          <td >Year</td>
          <td><input name="t1" type="number" id="t1"  required /></td>
          <td width="31%"></td>
        </tr>
        <tr>
          <td > </td>
          <td> </td>
          <td></td>
        </tr>
        <tr>
          <td >Month<br></td>
          <td><input name="t2" type="number" id="t2"  required /></td>
          <td> </td>
        </tr>

        <tr>
          <td ><br></td>
          <td> </td>
          <td> </td>
        </tr>
        <tr>
          <td >Date</td>
          <td><input name="t3" type="number" id="t3"  required /></td>
          <td> </td>
        </tr>
        <tr>
          <td > </td>
          <td> </td>
          <td></td>
        </tr>
        <tr>
          <td > </td>
          <td> </td>

          <td></td>
        </tr>
```

```html
          <tr>
            <td >Prediction Price </td>
            <td><span class="style1">{{res}}</span></td>
            <td></td>
          </tr>


          <tr>
            <td > </td>
            <td> </td>
            <td> </td>
          </tr>
          <tr>
            <td > </td>
            <td> </td>
            <td> </td>
          </tr>

          <tr>
            <td > </td>
            <td><input name="btn" type="submit" id="btn" value="Submit" />
                <input type="reset" name="Submit2" value="Reset" /></td>
            <td> </td>
          </tr>
        </table>
    <p> </p>
    <!-- <input type="text" name="comment"/> --><br/>
    </form>

  </div>




</body>
</html>
```

Home.html(output):



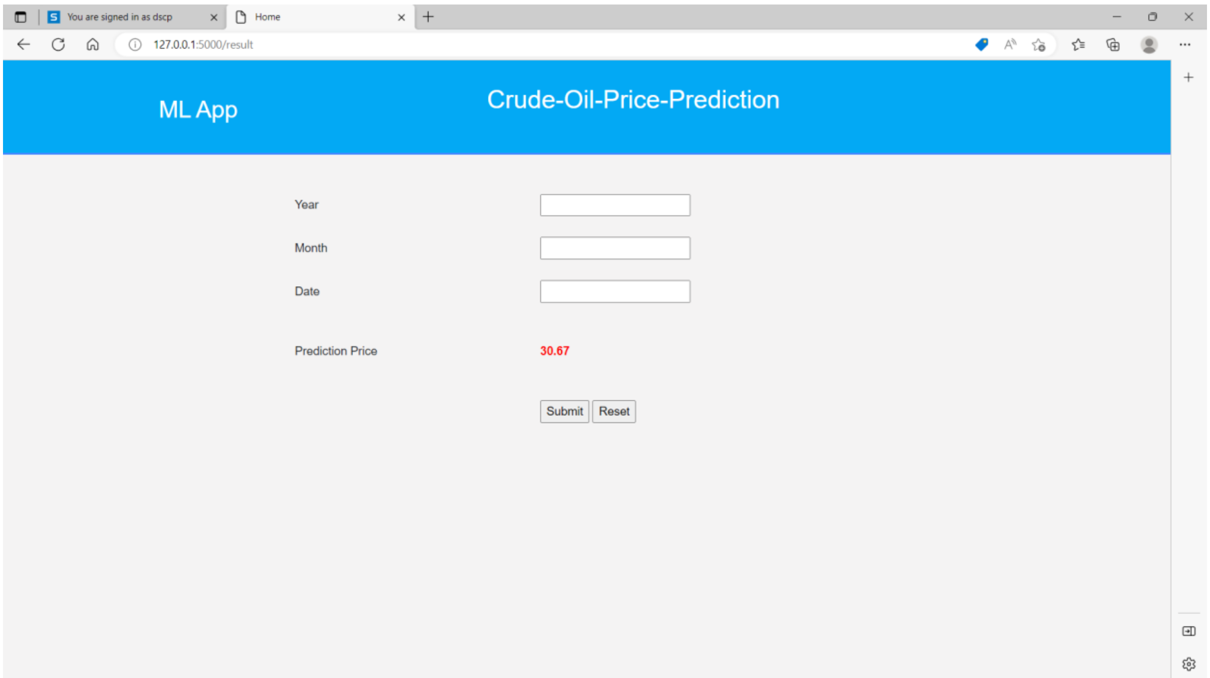
7.2 FEATURE 2

Result.html(code):

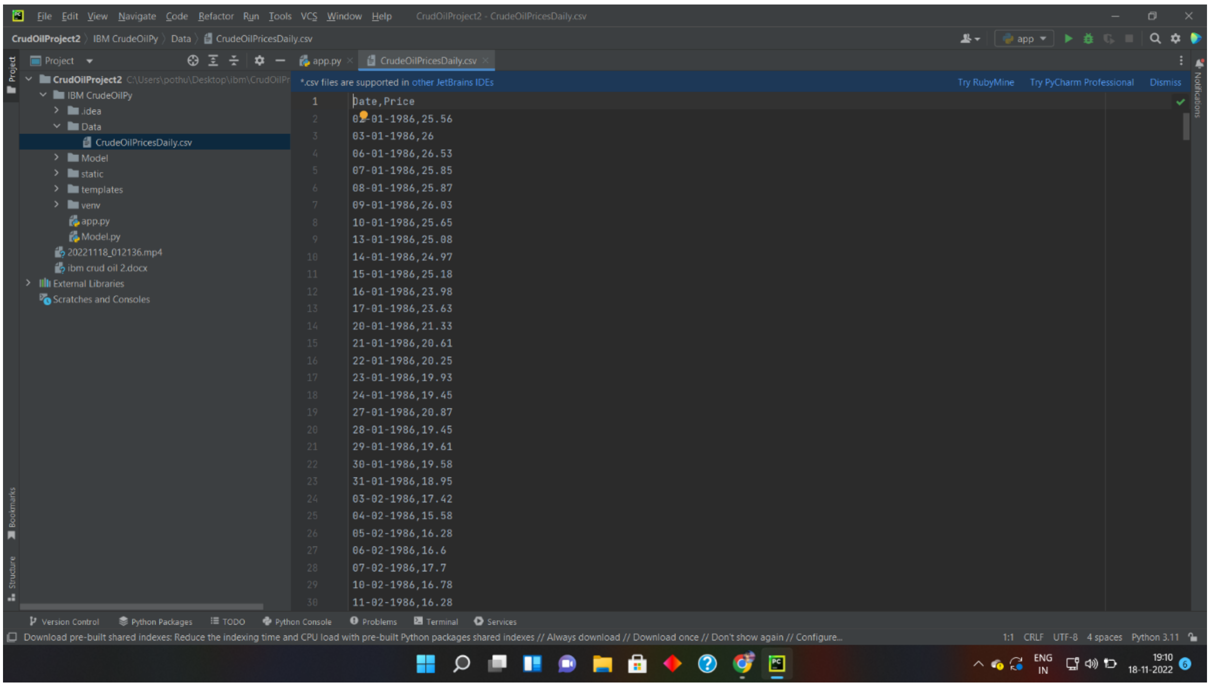
```
<!DOCTYPE html>
<html>
<head>
    <title></title>
    <link rel="stylesheet" type="text/css" href="{{ url_for('static',
filename='css/styles.css') }}">
    <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.css
" integrity="sha384-
ggOyR0iXCbMQv3Xipma34MD+dH/1fQ784/j6cY/iJTQUOhcWr7x9JvoRxT2MZw1T"
crossorigin="anonymous">
</head>
<body>

    <header>
        <div class="container">
        <div id="brandname">
            ML App
        </div>
        <h2>Twitter Sentiment Analysis</h2>
```

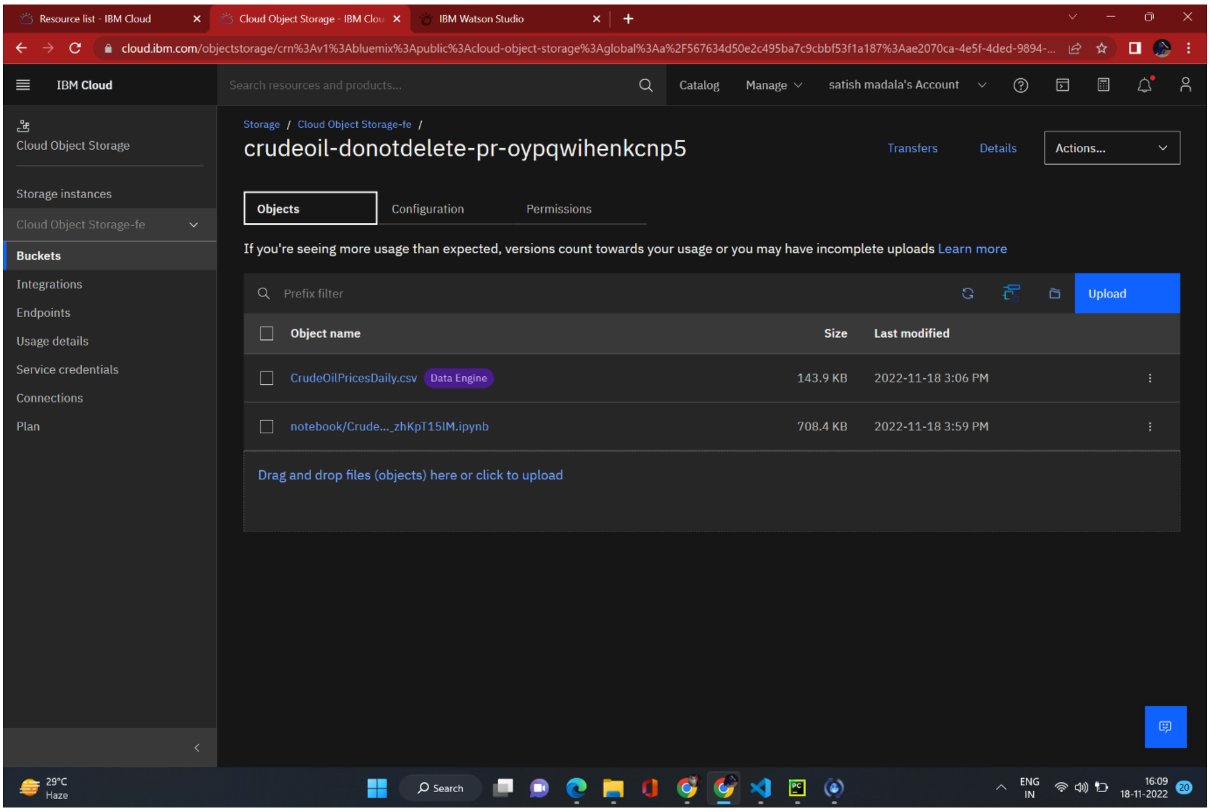
```
    </div>
    </header>


    <div class="results">



    {% if prediction == 1%}
    <h2 style="color:red;">Negative</h2>
    {% elif prediction == 0%}
    <h2 style="color:blue;">Postive</h2>
    {% endif %}

    </div>

</body>
</html>
```

Result.html(output):



7.3 DATABASE SCHEMA





8. TESTING

8.1 TEST CASES

A test case has components that describe input, action and an expected response, in order to determine if a feature of an application is working correctly. A test case is a set of instructions on "HOW" to validate a particular test objective/target, which when followed will tell us if the expected behavior of the system is satisfied or not.
Characteristics of a good test case:

- Accurate: Exacts the purpose.
- Economical: No unnecessary steps or words.
- Traceable: Capable of being traced to requirements.
- Repeatable: Can be used to perform the test over and over.
- Reusable: Can be reused if necessary.

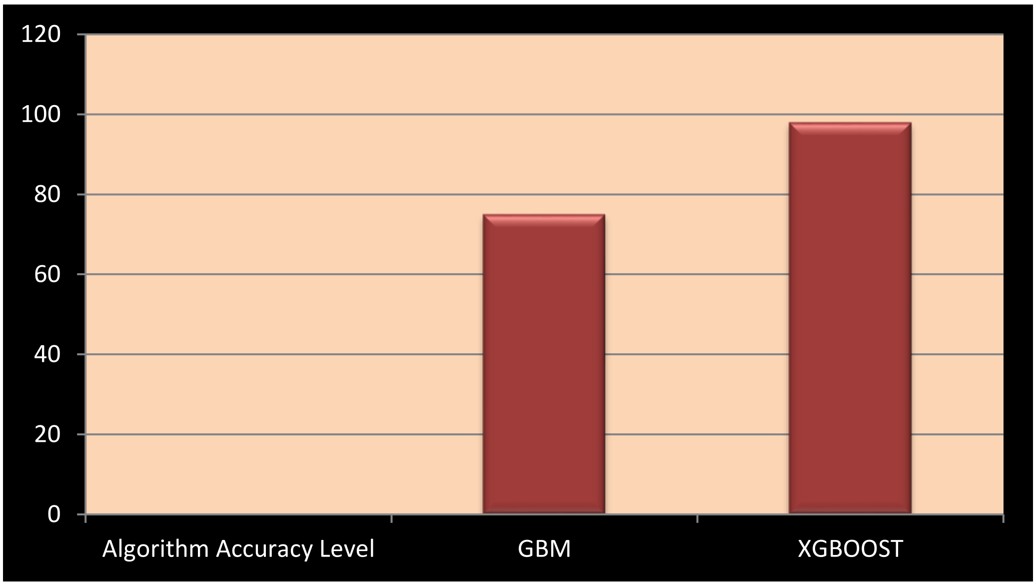
S.NO	Scenario	Input	Output	Status
1	User login	User name and password	Login	Login success.
2	Price Prediction	Give input in the user defined parameters	Predicting the crude oil prices by getting the previous prices	Details are stored in a database.

8.2 USER ACCEPTANCE TESTING

This sort of testing is carried out by users, clients, or other authorised bodies to identify the requirements and operational procedures of an application or piece of software. The most crucial stage of testing is acceptance testing since it determines whether or not the customer will accept the application or programme. It could entail the application's U.I., performance, usability, and usefulness. It is also referred to as end-user testing, operational acceptance testing, and user acceptance testing (UAT).

9. RESULTS

9.1 PERFORMANCE METRICS



10. ADVANTAGES & DISADVANTAGES

ADVANTAGE

- Give accurate result
- Easy to access and get the price
- Effective with large datasets

DISADVANTAGE

- Hard to find oil price
- Inefficient in accuracy
- Poor Customer support

11. CONCLUSION

Predicting Crude Oil prices is a very challenging problem due to the high volatility of oil prices. In this paper, we developed a new oil price prediction approach using ideas and tools from stream learning, a machine learning paradigm for analysis and inference of continuous flow of non-stationary data. Our stream learning model will be updated whenever new oil price data are available, and provided to model, so the model continuously evolves over time, and can capture the changing pattern of oil prices. In addition, updating the model requires only a small constant time per new data example, the experiment results show that our stream learning model outperformed four other popular oil price prediction models over a variety of forecast time horizons. This process is used to Predict the oil Prices. The prediction model predicts continuous valued functions.

12. FUTURE SCOPE

Future research may extend our work by considering a richer set of market variables, such as political or commercial factors and phases of economic instability, which are often determinants of crude oil price. Moreover, another direction for future research is the application of the proposed model to forecast the price of other commodities. Moreover, it is a worthwhile direction to explore the consideration of one or more computational cost factors when comparing different forecasting models. Therefore, calculations based on operational research methods might be a good direction.

13. APPENDIX

SOURCE CODE

Home.html:

```html
<!DOCTYPE html>
<html>
<head>
  <title>Home</title>
  <!-- <link rel="stylesheet" type="text/css" href="../static/css/styles.css">
-->
  <link rel="stylesheet" type="text/css" href="{{ url_for('static',
filename='css/styles.css') }}">
  <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.css
" integrity="sha384-
ggOyR0iXCbMQv3Xipma34MD+dH/1fQ784/j6cY/iJTQUOhcWr7x9JvoRxT2MZw1T"
crossorigin="anonymous">
    <style>
        body{
  font:15px/1.5 Arial, Helvetica,sans-serif;
  padding: 0px;
  background-color:#f4f3f3;
}

.container{
  width:100%;
  margin: auto;
  overflow: hidden;
}

header{
  background:#03A9F4;#35434a;
  border-bottom:#448AFF 3px solid;
  height:120px;
  width:100%;
  padding-top:30px;

}

.main-header{
      text-align:center;
      background-color: blue;
      height:100px;
      width:100%;
      margin:0px;
```

```
    }
#brandname{
  float:left;
  font-size:30px;
  color: #fff;
  margin: 10px;
}

header h2{
  text-align:center;
  color:#fff;

}



.btn-info {background-color: #2196F3;
  height:40px;
  width:100px;} /* Blue */
.btn-info:hover {background: #0b7dda;}


.resultss{
  border-radius: 15px 50px;
    background: #345fe4;
    padding: 20px;
    width: 200px;
    height: 150px;
}

    .style1 {
  color: #FF0000;
  font-weight: bold;
}
    </style>
</head>
<body>

  <header>
    <div class="container">
    <div id="brandname">
      ML App
    </div>
    <h2>Crude-Oil-Price-Prediction  </h2>

  </div>
  </header>

  <div class="ml-container">
```

```
  <form id="form1" name="form1" method="post" action="/result">
<table width="50%"  border="0" align="center">

      <tr>
        <td width="42%" > </td>
        <td width="27%"> </td>
      </tr>
      <tr>
        <td > </td>
        <td> </td>
      </tr>
      <tr>
        <td >Year</td>
        <td><input name="t1" type="number" id="t1"  required /></td>
        <td width="31%"></td>
      </tr>
      <tr>
        <td > </td>
        <td> </td>
        <td></td>
      </tr>
      <tr>
        <td >Month<br></td>
        <td><input name="t2" type="number" id="t2"  required /></td>
        <td> </td>
      </tr>

      <tr>
        <td ><br></td>
        <td> </td>
        <td> </td>
      </tr>
      <tr>
        <td >Date</td>
        <td><input name="t3" type="number" id="t3"  required /></td>
        <td> </td>
      </tr>
      <tr>
        <td > </td>
        <td> </td>
        <td></td>
      </tr>
      <tr>
        <td > </td>
        <td> </td>

        <td></td>
```

```html
            </tr>
            <tr>
              <td >Prediction Price </td>
              <td><span class="style1">{{res}}</span></td>
              <td></td>
            </tr>


            <tr>
              <td > </td>
              <td> </td>
              <td> </td>
            </tr>
            <tr>
              <td > </td>
              <td> </td>
              <td> </td>
            </tr>

            <tr>
              <td > </td>
              <td><input name="btn" type="submit" id="btn" value="Submit" />
                  <input type="reset" name="Submit2" value="Reset" /></td>
              <td> </td>
            </tr>
          </table>
    <p> </p>
    <!-- <input type="text" name="comment"/> --><br/>
    </form>

  </div>




</body>
</html>
```

Result.html:

```
<!DOCTYPE html>
<html>
<head>
    <title></title>
    <link rel="stylesheet" type="text/css" href="{{ url_for('static',
filename='css/styles.css') }}">
    <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.css
" integrity="sha384-
ggOyR0iXCbMQv3Xipma34MD+dH/1fQ784/j6cY/iJTQUOhcWr7x9JvoRxT2MZw1T"
crossorigin="anonymous">
</head>
<body>

    <header>
        <div class="container">
        <div id="brandname">
            ML App
        </div>
        <h2>Twitter Sentiment Analysis</h2>

    </div>
    </header>


    <div class="results">



    {% if prediction == 1%}
    <h2 style="color:red;">Negative</h2>
    {% elif prediction == 0%}
    <h2 style="color:blue;">Postive</h2>
    {% endif %}

    </div>

</body>
</html>
```

App.py

```
from flask import Flask, render_template, flash, request, session,send_file
from flask import render_template, redirect, url_for, request

import sys

import pickle


import numpy as np

app = Flask(__name__)
app.config['DEBUG']
app.config['SECRET_KEY'] = '7d441f27d441f27567d441f2b6176a'

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



@app.route("/result", methods=['GET', 'POST'])
def result():
    if request.method == 'POST':

        year = request.form['t1']
        month = request.form['t2']
        date = request.form['t3']








        filename = 'Model/prediction-rfc-model.pkl'
        classifier = pickle.load(open(filename, 'rb'))

        data = np.array([[year,month,date ]])

        my_prediction = classifier.predict(data)

        print(my_prediction)



        print(my_prediction[0])
```

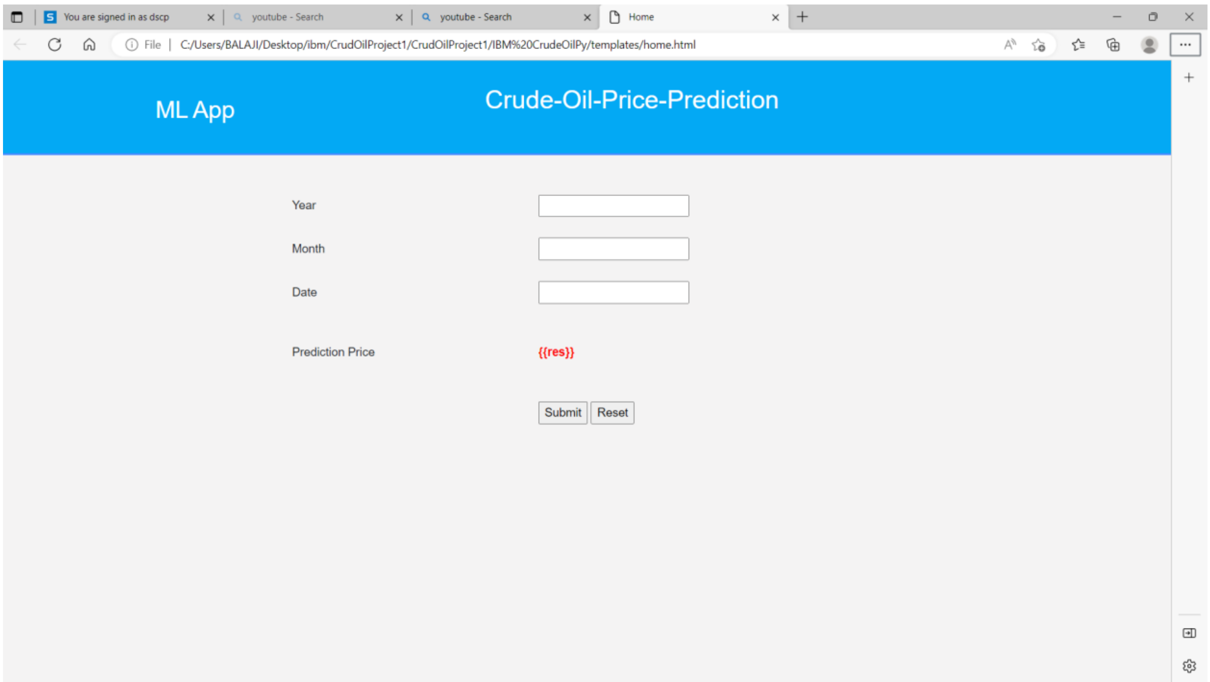
```python
        da = ("%.2f" % round(my_prediction[0], 2))

        return render_template('home.html',res=da)








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

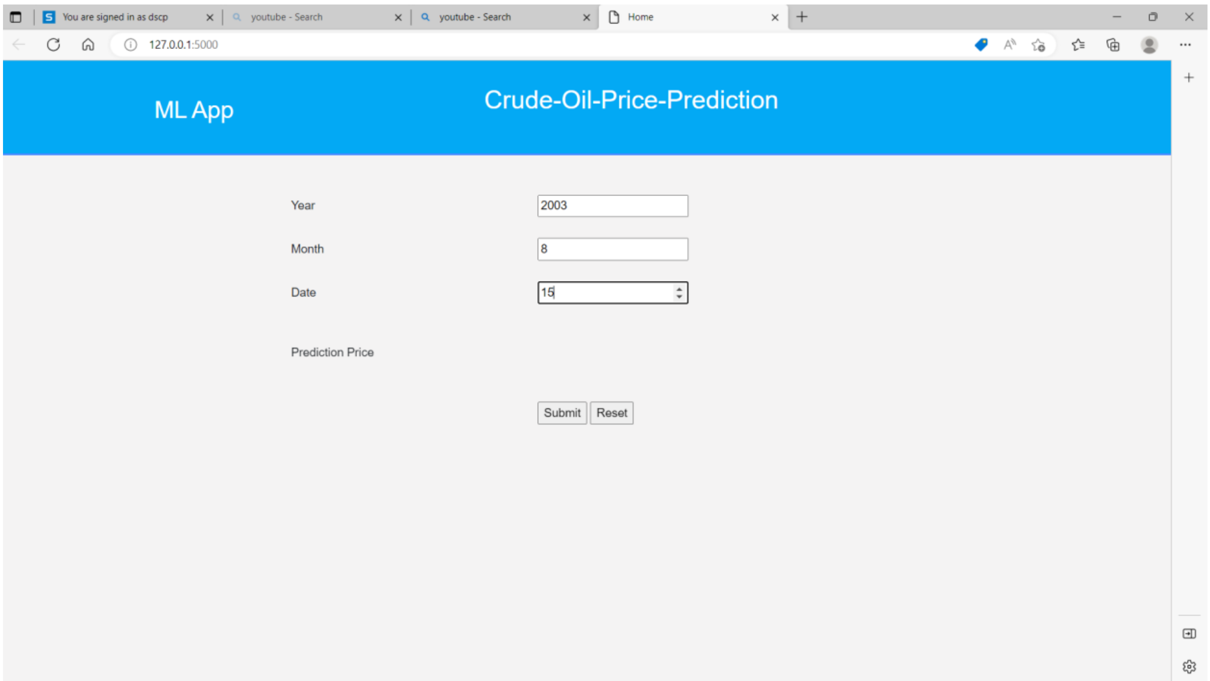
model.py

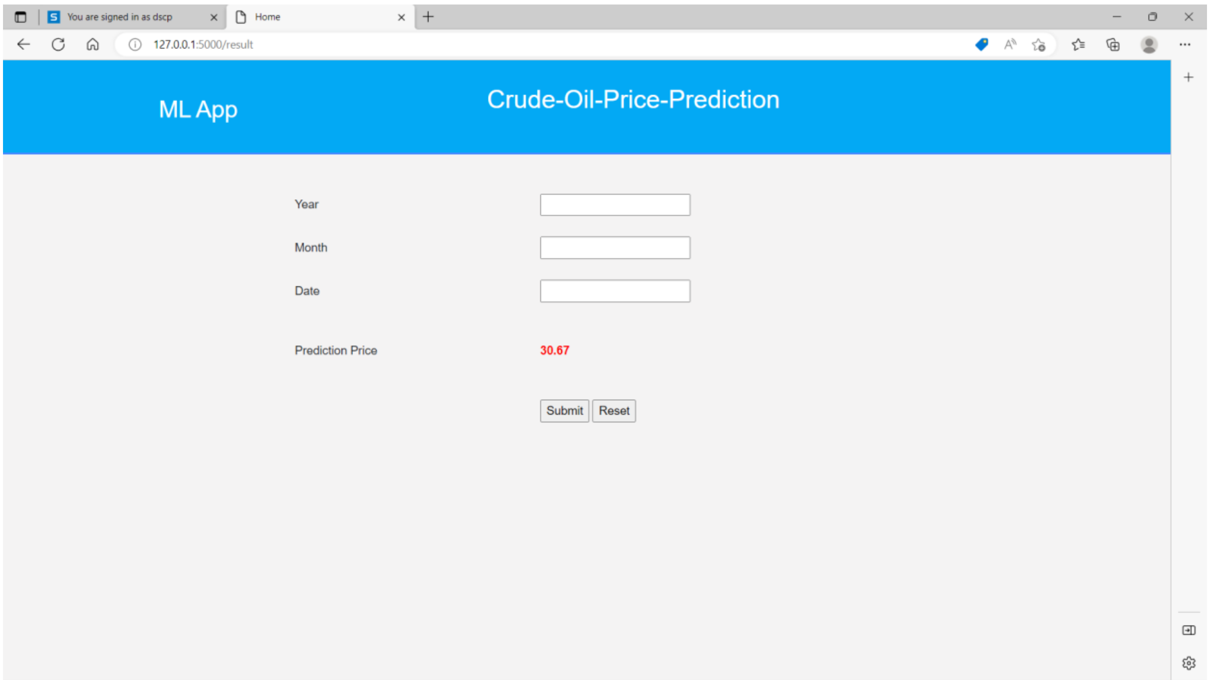
```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import pickle
import  warnings

warnings.filterwarnings("ignore", category=DeprecationWarning)

data=pd.read_csv('./Data/CrudeOilPricesDaily.csv')
print(data)
print(data.head())

data['year'] = pd.DatetimeIndex(data['Date']).year
data['month'] = pd.DatetimeIndex(data['Date']).month
data['day'] = pd.DatetimeIndex(data['Date']).day
data.drop('Date',axis=1,inplace=True)

data['Price'].fillna(data['Price'].median(),inplace=True)

X=data.drop('Price',axis=1)
Y=data['Price']


from sklearn.model_selection import train_test_split
```

```
X_train,X_test,y_train,y_test=train_test_split(X,Y,test_size=0.2,random_state=
42)
X_train.shape,X_test.shape,y_train.shape,y_test.shape




from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
from sklearn.metrics import plot_confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn.ensemble import RandomForestRegressor
regressor = RandomForestRegressor (n_estimators =1000, max_depth = 10,
random_state = 34)


regressor.fit (X_train, np.ravel(y_train, order = 'C'))




# Creating a pickle file for the classifier
filename = 'Model/prediction-rfc-model.pkl'
pickle.dump(regressor, open(filename, 'wb'))




filename = 'Model/prediction-rfc-model.pkl'
pickle.dump(regressor, open(filename, 'wb'))

filename = 'Model/prediction-rfc-model.pkl'
classifier = pickle.load(open(filename, 'rb'))


data = np.array([[1986,3,17]])
my_prediction = classifier.predict(data)
warnings.filterwarnings("ignore", category=DeprecationWarning)
print(my_prediction[0])
```

Output:


You are signed in as dscp
youtube - Search
youtube - Search
Home
File | C:/Users/BALAJI/Desktop/ibm/CrudOilProject1/CrudOilProject1/IBM%20CrudeOilPy/templates/home.html
ML App
Crude-Oil-Price-Prediction
Year
Month
Date
Prediction Price
{{res}}
Submit
Reset


You are signed in as dscp
youtube - Search
youtube - Search
Home
127.0.0.1:5000
ML App
Crude-Oil-Price-Prediction
Year
2003
Month
8
Date
15
Prediction Price
Submit
Reset



GITHUB & PROJECT DEMO LINK

Git-hub link : https://github.com/IBM-EPBL/IBM-Project-15664-1659602682

Demo video link: https://youtu.be/RTTNWBCtW30