NATURAL GAS PRICE PREDICTION

Using Random Forest Regression.

Developed by: K S Meghana, Spandana J, Meghana M, Sanjana G P, Lanchana Gudami

Smart Bridge-Remote Summer Internship Program

1.INTRODUCTION

Natural gas constitutes one of the most actively traded energy commodity with a significant impact on many financial activities of the world. The accurate natural gas price prediction and the direction of price changes are considered essential since these forecasts are utilized in energy sustainability planning, commodity trading and decision making. In this project, a new deep learning prediction model is proposed for forecasting natural gas price. Additionally, a significant advantage of the proposed model is its abilities to predict the price of natural gas on the following day (regression) and also to predict if the price on the next day will increase, decrease or stay stable (classification) with respect to today's price.

Being able to forecast natural gas price benefits various stakeholders and has become a very valuable tool for all market participants in competitive natural gas markets. Machine learning algorithms have gradually become popular tools for natural gas price forecasting. In this, we investigate data-driven predictive models for natural gas price forecasting based on common machine learning tools. Natural gas energy causes less pollution to the environment than other kinds of energy resource, it has received much more recognition recently. Natural gas, which is one of the most important energy resources, is going to play an expanded role in the future of global energy due to its significant environmental benefits. Forecasting natural gas prices is a powerful and essential tool which has become more important for different stakeholders in the natural gas market, allowing them to make better decisions for managing the potential risk, reducing the gap between the demand and supply, and optimizing the usage of resources based on accurate predictions. Accurate natural gas price forecasting not only provides an important guide for effective implementation of energy policy and planning, but also is extremely significant in economic planning, energy investment, and environmental conservation.

1.1OVERVIEW

Natural gas varies with season. In addition, natural gas supply, demand, storage, and imports are important indicators related to natural gas price. There are plenty of methods for analyzing and forecasting natural gas prices and machine learning is increasingly used. Machine learning algorithms can learn from historical relationships and trends in the data and make data-driven predictions or decisions.Here a new model for predicting price for natural gas by using Machine Learning concepts. Here some algorithms have been used to build the proposed model: Random Forest Regression, Linear Regression, Decision Tree, Multilinear Regression. By using the algorithm, a Flask model has been implemented and tested. The results have been discussed and a full comparison between algorithms was conducted. Random forest Regression was selected as best algorithm based on accuracy.

1.2 PURPOSE

The aim of this project is to discuss data-driven models for natural gas price forecasting, and make use of pandas, matplotlib, & seaborn libraries from python to extract the libraries for machine learning for the price prediction of natural gas. Specifically, we will introduce the fundamental frameworks, basic concepts, and major technical features of each machine learning approach; then we analyze and compare their practical applications in natural gas price prediction and investigate future developments with potential research significance of these approaches in the forecasting of energy prices field.

2. LITERATURE SURVEY

Data mining is the process of analyzing data from different perspectives and extracting useful knowledge from it. It is the core of knowledge discovery process. Natural gas price prediction is playing very important role in investments. However, there is no guarantee that the price prediction using historical data will be 100% accurate due to the uncertainty in the future. Different data mining techniques include classification, clustering, association rule mining, prediction and sequential patterns, neural networks, regression etc. To predict the natural gas prices we can use historical data. This method focuses on an analysis such as daily opening, high, low, and closing prices. In addition,

other features may be considered and used in the technical analysis for increasing an accuracy in the prediction. Gas price can fluctuate depending on political and economic conditions. Thus, investors have used fundamental and technical analysis simultaneously for the price prediction. The present study has presented the current state of the literature in the field of natural gas forecasting we perform a discussion regarding the numerical performance of our proposed model for forecasting natural gas price and movement as well as the main findings.

2.1EXISTING PROBLEM

Machine learning algorithms have gradually become popular tools for natural gas price forecasting. It has always been a difficult task to predict the exact daily price ofthe stock market index. Many factors such as political events, general economic conditions, and traders expectations may have an influence onthe spot price index.The problem of forecasting natural gas price and movement can be considered to belong on chaotic time series problems. This means that accurate and reliable predictions are almost impossible since these problems are close to random walk processes, while the identification of possible existing patters and their proper distinguishment among a large pool of noisy instances, seems to be a significantly challenging task. Though we all are aware of the fact that, gas consumption is traditionally high now a days, which leads to rising market prices. The problem is that, in contrast to popular belief, there is remarkable deviation in price almost throughout the year.In order to target the said problem, we use machine learning algorithm to examine if we can predict the future price by adding different parameters which may have external direct or indirect effect on natural gas price.

2.2PROPOSED SOLUTION

In this the main aim is to predict the Natural Gas Price. As the Crud oil Prices are Increased randomly the gas price also gets Increased. The aim of this project is to discuss data-driven models for natural gas price forecasting, which utilize machine learning approaches and Logistic Regression etc. Here, we have applied several input variables to predict the price direction. However, the prediction performance may be improved further by selecting effective input indicators. In addition, we can include a few other technical indicators that may affect the prediction performance. The models in the experiment have solved the problem as a regression. Regression model can be combined with Classification model predicting rise and fall to improve the directional accuracy for price

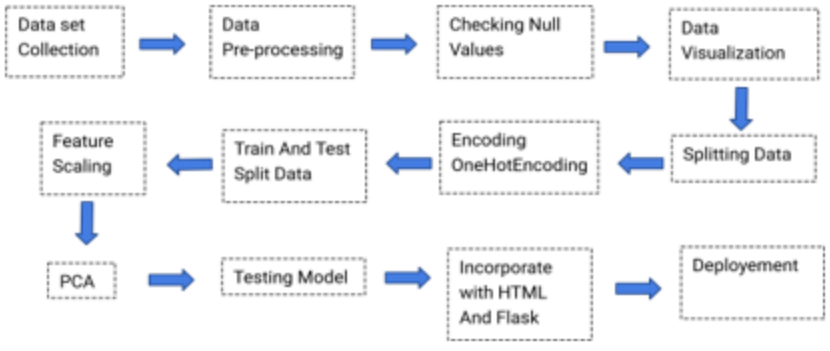
predictions. Moreover, the same model need to be tested with large amount of data for prediction accuracy because non-parametric modeling approach requires a lot more training data to estimate the mapping function. Regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of classes(classification) or mean prediction(regression) of the individual trees. Random forest method, involves training each decision tree on a different data sample where sampling is done with replacement. And also we have created an UI using the Flask for the price prediction, this UI will allows us to predict the price. We use machine learning algorithm to examine and we can predict the future price by adding different parameters which may have external direct or indirect effect on natural gas price. We will try to develop a forecasting model to predict the spot price of natural gas based on historical prices.

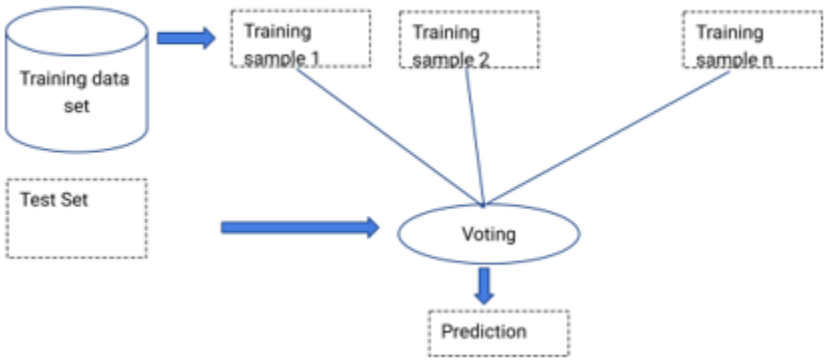
3.THEORETICAL ANALYSIS

While selecting the algorithm that gives an accurate prediction we gone through lot of algorithms which gives the results abruptly accurate and from them we selected only one algorithm for the prediction problem that is Random Forest Regression, it assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature that is how the prediction work great with the Random Forest Regression Algorithm.

At first we got like lot of worst accuracies because we tried lot of algorithms for the best accurate algorithm, finally after all of that we tried the best suitable algorithm which gives the prediction of 98% in Random Forest Regression and we developed it to use as a real time prediction problem for the price prediction.

3.1BLOCK DIAGRAM





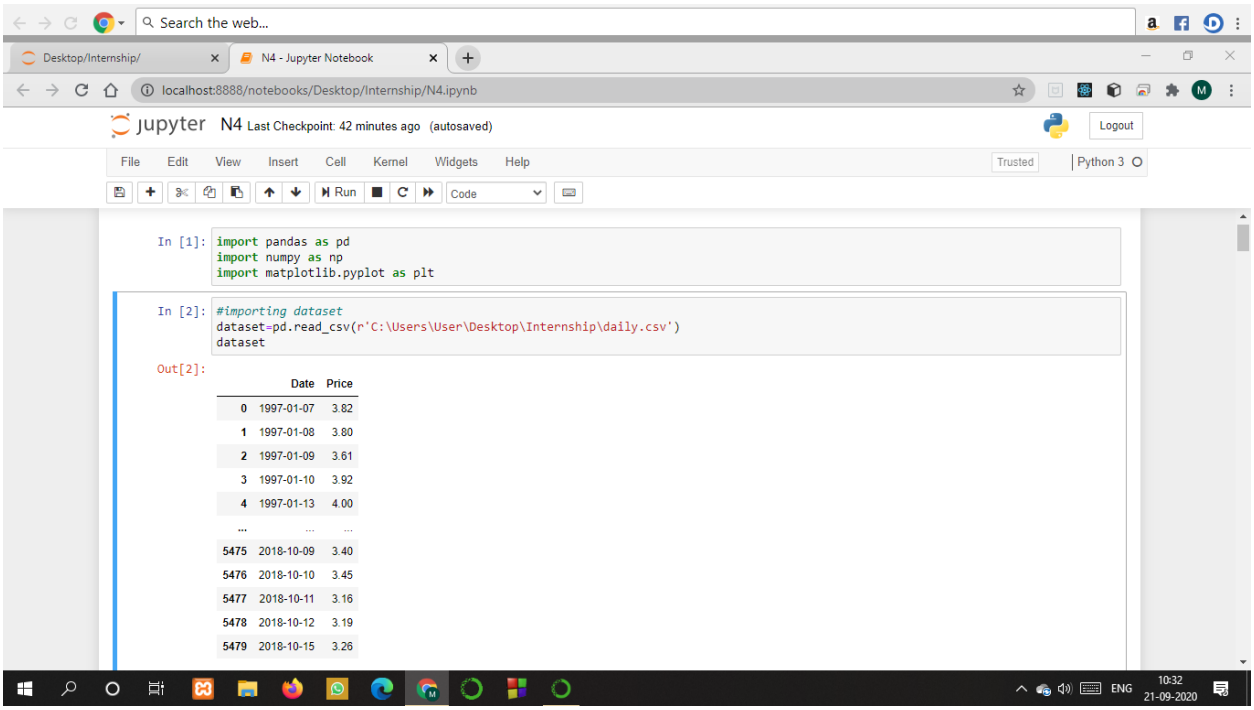
3.2 SOFTWARE DESIGNING

1. Jupyter Notebook Environment
2. Spyder IDE
3. Machine Learning Algorithms
4. Python (pandas, numpy, matplotlib, seaborn, sklearn)
5. HTML
6. Flask

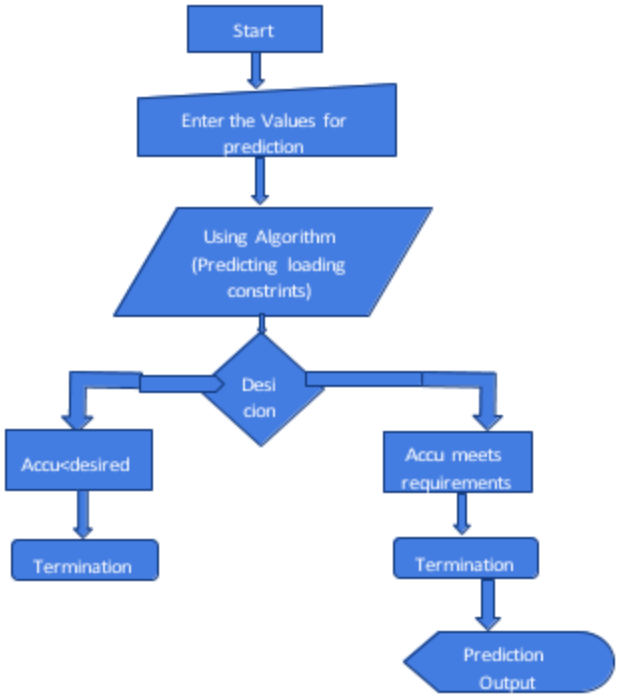
We developed this natural gas price prediction model by using the Python language which is a interpreted and high level programming language and using the Machine Learning algorithms for coding we used the Jupyter Notebook environment of the Anaconda distributions and the Spyder, it is an integrated scientific programming in the python language. For creating an user interface for the prediction we used the Flask. It is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions, and a scripting language to create a webpage is HTML by creating the templates to use in the functions of the Flask and HTML.

4. EXPERIMENTAL INVESTIGATION

We have included a few technical indicators that may affect the prediction performance. The models in the experiment have solved the problem as a regression. The presented algorithms highlighted that although models constitute a wide and efficient choice for addressing price detection also provide a significant boost in increasing the forecasting performance. The empirical results demonstrate the prediction methods have decent performance in forecasting natural gas price



5. FLOWCHART

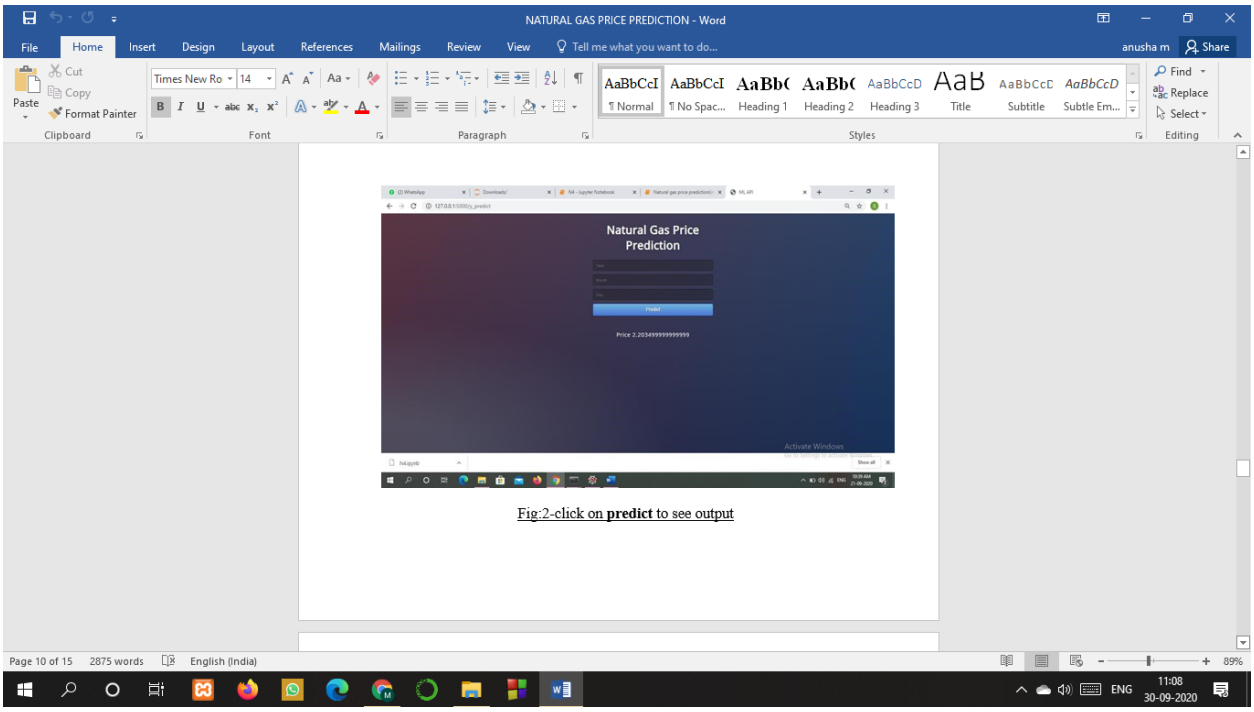


6. RESULT

In this paper, the Random Forest Regression algorithm is used to predict its performance, and compared with another three machine learning methods namely the decision tree, simple linear regression, multi linear regression. The obtained results are displayed in Table below. The results show that, the performance of decision tree and Random forest regression have comparable performance, but random forest regression still performs the best, with an accuracy of 98%, higher than the decision tree with an accuracy of 97% which is pretty good.

ALGORITHM	ACCURACY
1.Decision tree	97%
2.Random forest regression	98 %

Login page



7. ADVANTAGES AND DISADVANTAGES

Advantages:

1. **Natural Gas is Environmentally clean:** Among many burning fossil fuels available on planet earth, natural gas is the most cleanest .It is considered clean because unlike other gases, natural gas emits very less byproducts into the atmosphere as pollutants. This keeps the air we breath cleaner.
2. **It is economical :** Natural gas us not as expensive as other burning fuels.
3. **It is convenient :** If people use natural gas for cooking or electric purpose, It can be directed connected to the consumers house with the help of pipelines .This eliminates the long process of collecting, processing, packing and delivering the end product to the customers.
4. **Natural Gas Is safe to use :** One of the biggest fear we have when using LPG cylinders or any other type of fuel is the possibility of leakage. Gases

have the power of burning the things and people around and can be very dangerous. However, natural gas is lighter than air.in case of a leakage, it dissipates quickly into the air avoiding fire. Isn't that super safe?

5. **It is available abundantly:** Currently, experts state that amount of natural gas available is more. than crude oil or other such products.

Disadvantages:

6. **Limited qualities :** The problem in India is that we don't have vast reserves of natural gas which means that most of natural gas that is consumed by us has to be brought from other countries. Such constant purchase can turn into a rather expensive proposition over time.
7. **Natural gas is highly combustible :** Though natural gas is lighter than air. One cannot deny the fact that is highly combustible. As natural gas is odourless, It is difficult to select a leakage as well.
8. **Natural Gas is a Non-Reneweable source of energy :** Keeping aside the positive of using natural gas, Once should not forgot that it is a non-renewable source of energy. Experts state that natural gas will be depleted in the future and we will have to import it from other nations.
9. **Natural Gas Emits Carbon Dioxide :** one of the biggest disadvantage of natural gas is that it emits carbon dioxide which is bad for our atmosphere .constant introduction of Carbon dioxide into our atmosphere will lead to climate change and also global warming.
10. **Long processing process :** As natural gas has other components that has to be removed before using residential or commercial purpose, It take a lot and manpower to process it.

8. APPLICATION

1. As our logistic regression analysis and regression models with broader set of inputs can effectively predict the price of a natural gas.
2. Our models could easily be incorporated into phone application or website. And using our models would be convenient and cost-effective.
3. The main contribution of this work is the development of a new

forecasting model for the short-term prediction of natural gas price.

4. Additionally, a significant advantage of the proposed model is its abilities to predict the price of natural gas on the following day (regression) and also to predict if the price on the next day will increase, decrease or stay stable (classification) with respect to today's price.
5. Accurate natural gas price forecasting provides an important guide for significant in economic planning, energy investment, and environmental conservation.
6. Though we all are aware of the fact that, gas consumption is traditionally high which leads to rising market prices. The problem is that, in contrast to popular belief, there is remarkable deviation in price almost throughout the year.So here we use the Flask for the price prediction, this will allows us to predict the price. We use machine learning algorithm to examine if we can predict the future price by adding different parameters which may have external direct or indirect effect on natural gas price.

9. CONCLUSION

The present model has presented the current state in the field of natural gas forecasting. The empirical results demonstrate the prediction methods have decent performance in forecasting natural gas price. It has always been a difficult task to predict the exact daily price of the natural gas. But our model would be convenient and can predict the price of next 12 months also it is cost-effective. The presented algorithms highlighted that although models constitute a wide and efficient choice for addressing price detection also provide a significant boost in increasing the forecasting performance. The main contribution of this work is the development of a new forecasting model for the short-term prediction of natural gas price.

10. FUTURE SCOPE

In future the Random Forest Regression can be applied on data sets to investigate its

accuracy of a natural gas. A rigorous analysis of other machine learning algorithms other than these six can also be done in future to investigate the power of machine learning algorithms for the price prediction of natural gas. In further study, we will try to conduct experiments on larger data sets or try to tune the model so as to achieve the state -of-art performance of the model and a great UI support system making it complete web application model. The model for supervised learning can be iterated so that they can achieved the higher accuracy. Overall due to less number of samples, it was difficult to build the models with higher accuracy. Hence as more data comes in, these models can be reiterated. Also, GUI can be made more comprehensive and more flexibility can be provided to these model for the end user to tweak the parameters visualize the results.

11. BIBLIOGRAPHY

1. Natural gas price links to oil market frustrate regulators efforts to develop competition". New York Times. 29 October 2008.
2. Roben Farzad (19 April 2012). "High Oil Prices Cut the Cost of Natural Gas". Businessweek.com. Retrieved 15 May 2015.
3. "Archived copy". Archived from the original on 1 April 2017. Retrieved 13 May 2013.
4. Disavino, Scott; Krishnan, Barani (25 September 2014). "Henry Hub, king of U.S.
   natural gas trade, losing crown to Marcellus". Reuters. Retrieved 21 October 2014.
5. Natural Gas – Exports". The World Factbook. Central Intelligence Agency. Retrieved 11 June 2015.
6. "Background". Naturalgas.org. Archived from the original on 9 July 2014. Retrieved 14 July 2012.
7. "Electricity from Natural Gas". Archived from the original on 6 June 2014. Retrieved 10 November 2013.

11. APPENDIX

INDEX.HTML:

```
<!DOCTYPE html>
<html >
<!--From https://codepen.io/frytyler/pen/EGdtg-->
<head>
  <meta charset="UTF-8">
  <title>ML API</title>
        <link href='https://fonts.googleapis.com/css?family=Pacifico'rel='stylesheet' type='text/css'>
<link href='https://fonts.googleapis.com/css?family=Arimo' rel='stylesheet' type='text/css'>
<link href='https://fonts.googleapis.com/css?family=Hind:300'rel='stylesheet' type='text/css'>
<link href='https://fonts.googleapis.com/css?family=Open+Sans+Condensed:300' rel='stylesheet' type='text/css'>
<link rel="stylesheet" href="{{ url_for('static', filename='css/style.css') }}">

<style>
.login{
top:
20%; }
</style>
</head>

<body>
 <div class="login">
 <h1>Natural Gas Price Prediction</h1>

    <!-- Main Input For Receiving Query to our ML -->
   <form action="{{ url_for('y_predict')}}"method="post">

         <input type="number" name="year" placeholder="Year" required="required" />
      <input type="number" name="month" placeholder="Month" required="required" />
```

```
      <input type="number" name="day" placeholder="Day" required="required" />

                        <button type="submit" class="btn btn-primary btn-block btn-large">Predict</button>

    </form>

   <br>
   <br>
   {{ prediction_text }}

 </div>


</body>

</html>
```

APP.PY:

```
from flask import Flask, request,
render_template from joblib import load app
= Flask(__name__) model= load('sp1.save')
trans=load('sh1')

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

@app.route('/y_predict',methods=['POST'])
def y_predict():
    '''
    For rendering results on HTML GUI
```

```
    '''
    x_test = [[x for x in
request.form.values()]]     print(x_test)
prediction = model.predict(x_test)
print(prediction)     output=prediction[0]


    return render_template('index.html', prediction_text='Price {}'.format(output))


if __name__ == "__main__":
    app.run(debug=True)
```