Natural gas price prediction system using IBM Watson Machine learning Service

INDEX


Natural Gas
$2.200
$2.100
$2.000
$1.900
$1.800
$1.700
OILPRICE.com
28. Jul
29. Jul
30. Jul
31. Jul
3. Aug

INTRODUCTION

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.As indicated in the abovementioned existing studies that exploited machine learning tools for natural gas price prediction, ANN is widely used machine learning methods in forecasting natural gas prices.

PROBLEM STATEMENT

- The aim of this project is to build data-driven machine learning models for natural gas price prediction.
- Data preprocessing based on the dataset, training a regression model for prediction.
- Creation of API through Flask for a user friendly interface.

SOLUTION

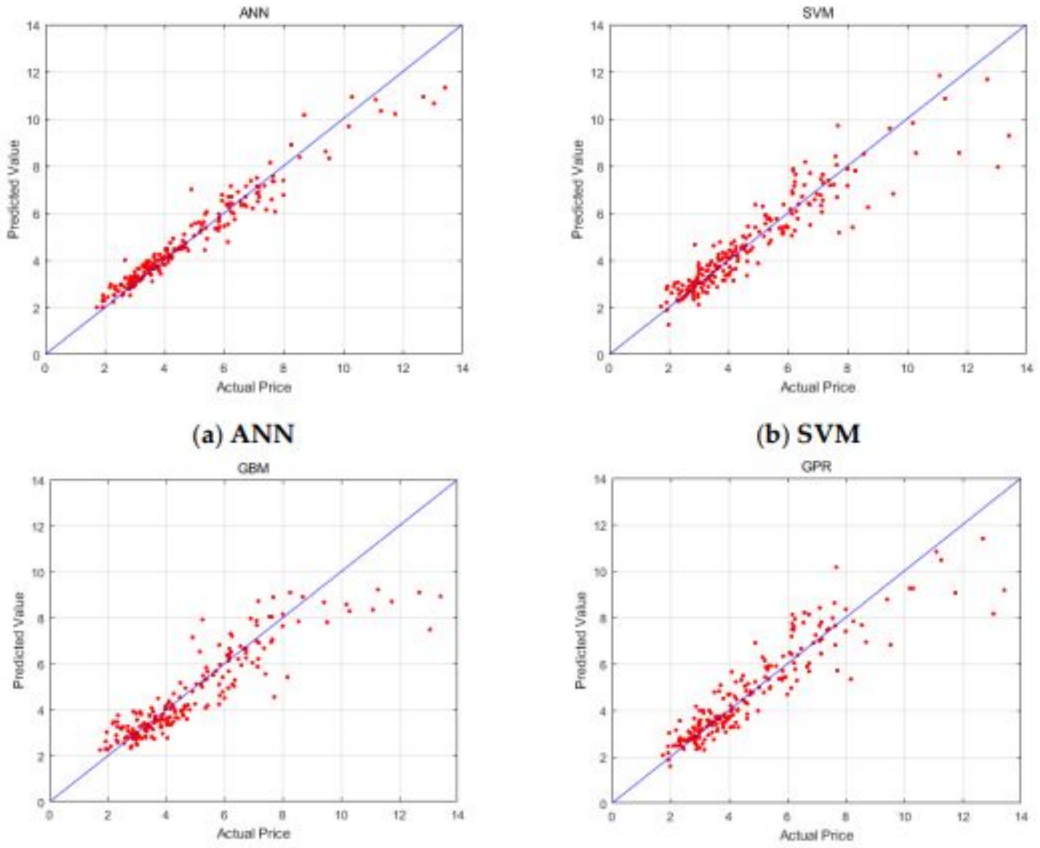
Natural gas has been proposed as a solution to increase the security of energy supply and reduce environmental pollution around the world. In this paper, we investigate data-driven predictive models for natural gas price forecasting based on common machine learning tools, i.e., artificial neural networks (ANN),support vector machines (SVM), gradient boosting machines (GBM), and Gaussian process regression (GPR).

Literature Survey

The poor track record of energy price forecasting models has encouraged analysts to turn to other sources of information about future energy prices, including most prominently, energy futures markets. Energy futures markets are ‘hubs’ that price and market natural gas. Walls (1995), examining several years of spot prices finds that, in general, gas futures are unbiased predictors of future spot prices whereas Herbet (1993) finds bias in natural gas futures prices where futures prices are greater than realized spot prices. Chinn et al (2005) finds futures prices to be unbiased predictors of future spot prices, with the exception of those in the natural gas market at the three month horizon and they slightly outperform time series models. This study builds upon the existing literature by investigating the accuracy of forecast methods up to the 24 month horizon.

EXPERIMENTAL INVESTIGATIONS

Based on the before-mentioned machine learning methods, prepared data, forecasting performance evaluation criteria, model validation technique, and selected model parameters, the empirical study is carried out. Observing data of four criteria can easily find that the forecasting performance of ANN and SVM is better than that of GBM and GPR. In particular, ANN is obviously superior to other methods while GBM has the worst behaviour. Overall, the performance ranking is ANN, SVM, GPR, and GBM from strong to weak. intuitively compares the prediction suitcases of four machine learning methods, which contains 214 observed values from January 2001 to October 2018. ANN delays two values, i.e., from March 2001. It can be obtained from Figure 4 that overall, all the methods have a strong prediction ability, since their predicted natural gas spot prices are close to the actual prices and approximately depict the characteristics of the Henry Hub natural gas spot price time series. From the view of the whole tendency, ANN outperforms others, in particular, for the prediction of abnormal values at the beginning of 2009 and the second half of 2010. SVM and GPR are inferior to ANN, but SVM excels the other three methods in terms of price prediction in the middle of 2008. GBM behaves the worst prediction ability among the methods, especially for outliers.



(a) ANN (b) SVM

(c) GBM (d) GPR

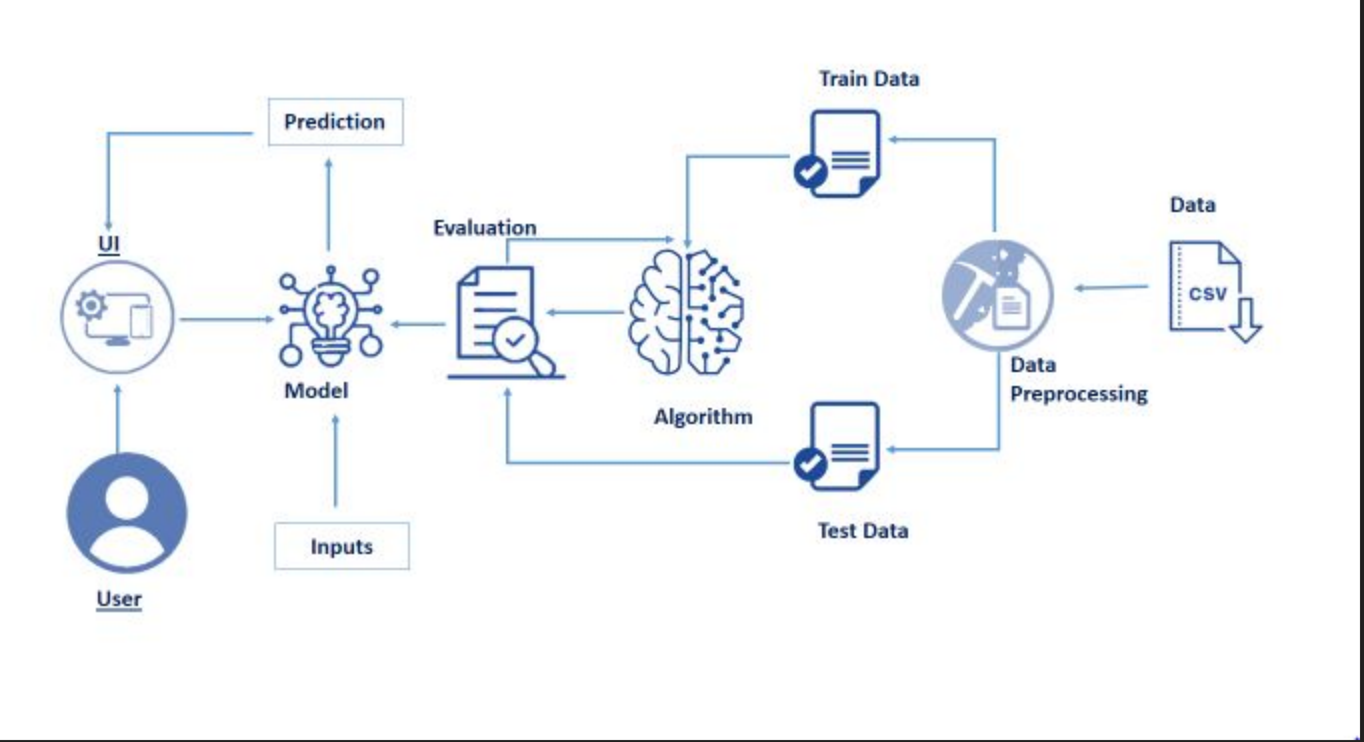
Figure 5. Distribution suitcases between predictive values and actual values.

SOFTWARE SPECIFICATIONS

- Jupyter Notebook- open source web application that you to create and share documents that contain live code, equations, visualizations and narrative text.

- Spyder- web API using the Flask framework

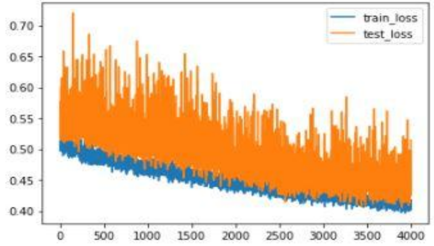
FLOWCHART


Train Data
Prediction
Data
UI
Evaluation
CSV
Model
Data
Preprocessing
Algorithm
Inputs
Test Data
User

OUTPUT

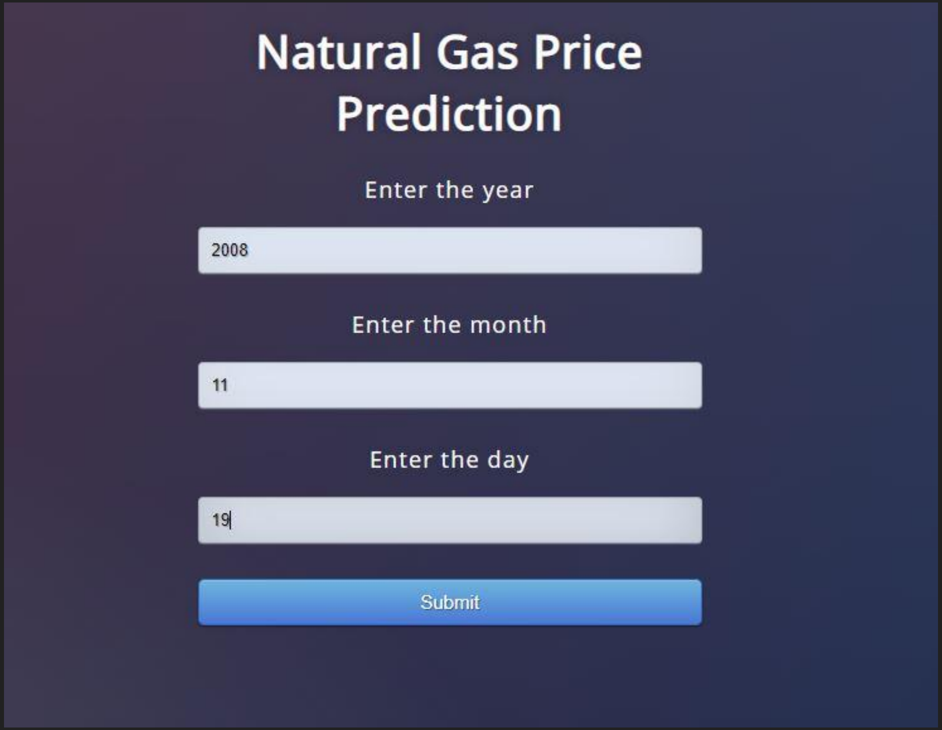
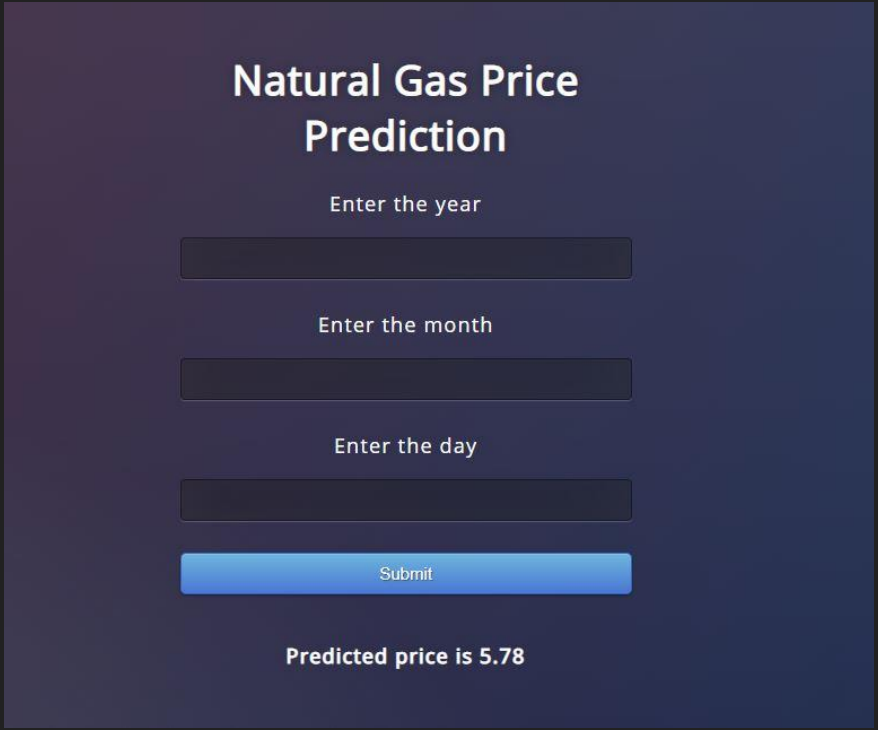
```
In [46]: from sklearn.metrics import r2_score
         accuracy=r2_score(y_test,y_pred)
         accuracy

Out[46]: 0.9029401956732639
```

```
In [54]: import matplotlib.pyplot as plt
         plt.plot(history.history['loss'],label="train_loss")
         plt.plot(history.history['val_loss'],label='test_loss')
         plt.legend()
         plt.show()
```



Verifying the model

```
In [ ]:  yp=reg.predict(sc.transform([[2008,7,30]]))

In [49]: yp

Out[49]: array([[8.949772]], dtype=float32)

In [50]: yp1=reg.predict(sc.transform([[2011,7,19]]))

In [51]: yp1

Out[51]: array([[3.9736688]], dtype=float32)

In [63]: reg.save("price.h5")

In [64]: import joblib
         joblib.dump(sc,"scaler")

Out[64]: ['scaler']
```


Natural Gas Price Prediction
Enter the year
2008
Enter the month
11
Enter the day
19
Submit

Natural Gas Price Prediction
Enter the year
Enter the month
Enter the day
Submit
Predicted price is 5.78

CONCLUSION

The aim of this study is to investigate natural gas price forecasting based on four machine learning methods (ANN, SVM, GBM, and GPR). Monthly Henry Hub natural gas spot price data from January 2001 to October 2018 (there are 215 observations) were used in four prediction methods. Nine variables were investigated as inputs, which are NGSP, WTI, HO, NGRR, HDD, CDD, NGMP, NGTC, NGUSV, and NGI. The method cross-validation was used in model training. Four forecasting performance evaluation criteria including R^2, MSE, RMSE, and MAPE are employed in prediction methods. Finally, the empirical results demonstrate that four prediction methods have decent performance in forecasting natural gas price. Overall, ANN and SVM have better forecasting performance than GBM and GPR. In particular, ANN obviously outperforms the other methods while GBM is the worst. This study could be improved by more thorough research, e.g., by comparing more different aspects of the prediction performance such as computation efficiency and using more diverse machine learning methods such as random forest.

FUTURE SCOPE

For future work, we will evaluate the effect of emerging machine learning algorithms, such as deep learning and reinforcement learning, on energy price and correlation prediction, the next step is to determine how these results can inform researchers and policy makers who utilize natural gas prices to develop federal energy policy. The analysis presented in this article is only possible because the futures market began to produce forward prices longer than a 12 month horizon in the mid-1990s. Allowing researchers to investigate this overlap with forecast models such as ANN to determine which approach, those based on economic models or on the market, is more accurate and unbiased.

10. BIBLIOGRAPHY

1. International Energy Agency (IEA). Key World Energy Statistics 2018. Available online: https://webstore.iea. org/key-world-energy-statistics-2018 (accessed on 30 April 2019).

2. Afgan, N.H.; Pilavachi, P.A.; Carvalho, M.G. Multi-criteria evaluation of natural gas resources. Energy Policy 2007, 35, 704–713. [CrossRef]

3. Abrishami, H.; Varahrami, V. Different methods for gas price forecasting. Cuad. Econ. 2011, 34, 137–144. [CrossRef] Energies 2019, 12, 1680 15 of 17

4. Busse, S.; Helmholz, P.; Weinmann, M. Forecasting day ahead spot price movements of natural gas—An analysis of potential influence factors on basis of a NARX neural network. In Proceedings of the Tagungsband der Multikonferenz Wirtschaftsinformatik (MKWI), Braunschweig, Germany; Available online: https://publikationsserver.tu-braunschweig.de/servlets/MCRFileNodeServlet/dbbs_derivate_00027726/Beitrag299.pdf (accessed on 30 April 2019).

5. Azadeh, A.; Sheikhalishahi, M.; Shahmiri, S. A Hybrid Neuro-Fuzzy Approach for Improvement of Natural Gas Price Forecasting in Vague and Noisy Environments: Domestic and Industrial Sectors. I