Project Report

Natural Gas Price Prediction

Done By: Dannapaneni Tripura

Index

1 Introduction

1.1 Overview

Natural gas is a non-renewable hydrocarbon used as a source of energy for heating, cooking, and electricity generation. It is also used as a fuel for vehicles and as a chemical feedstock in the manufacture of plastics and other commercially important organic chemicals.

The task here is to analyse the 'gasPrice' dataset and build data-driven machine learning models for natural gas price forecasting.

1.2 Purpose

It helps in managing the potential risk, reducing the gap between the demand and supply, and optimizing the usage of resources based on accurate predictions.

2 Literature Survey

2.1 Existing problem

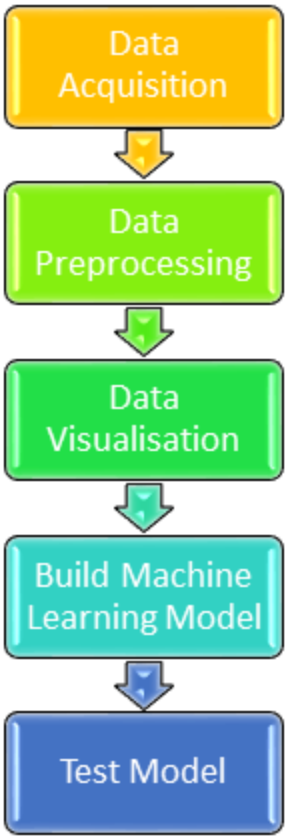
There exists many approches for predicting natural gas price. ANN and SVM are widely used machine learning models in forecasting natural gas price.

2.2 Proposed Solution

In this project the machine learning model which I used is Decison tree algorithm.

3 Theoretical Analysis

3.1 Block Diagram



3.2 Hardware / Software Requirements

- Windows OS 64 bit
- Eclipse
- Java Version 10.0
- Weka gui

4 Experimental Investigations

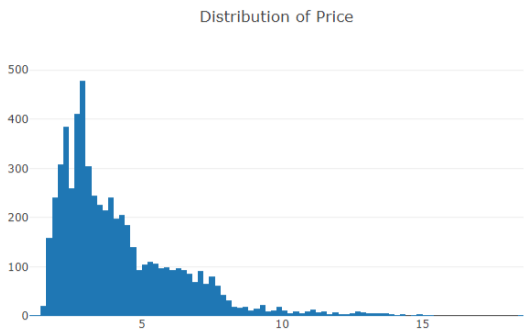
The dataset was a time series data with only 2 columns in it. They are:
- Date
- Price

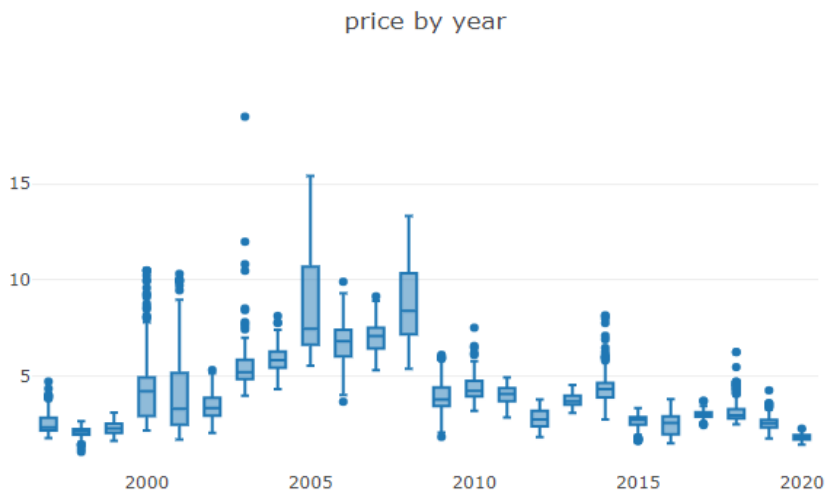
Size of dataset:
-Instances: 5938
-Rows: 2

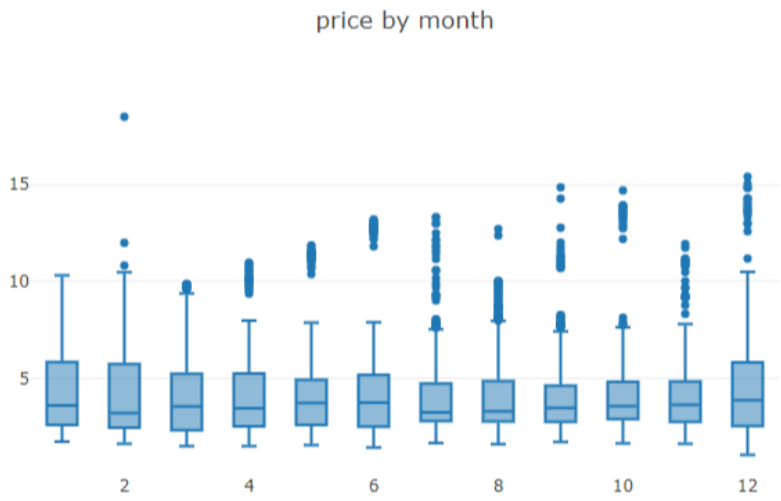
The dataset was then modified by adding the year, month and day columns to it and removing the date column.

When analysing the dataset it was found that there were missing values in 'Price' column present. So, using Weka the preprocessing was done.

Eclipse was used to construct different plots(such as histogram, box plot and bar plot) to analyse the data.

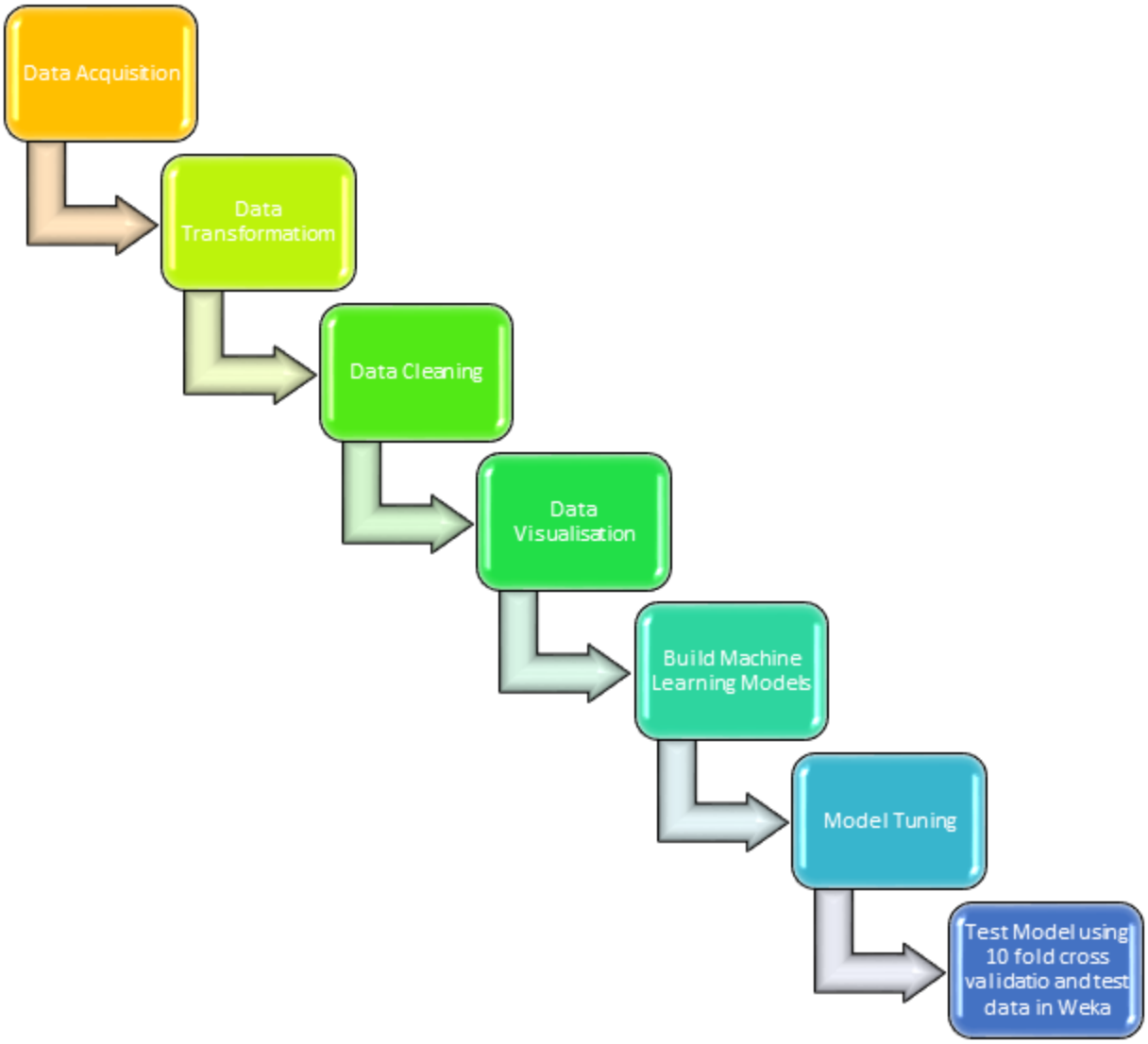






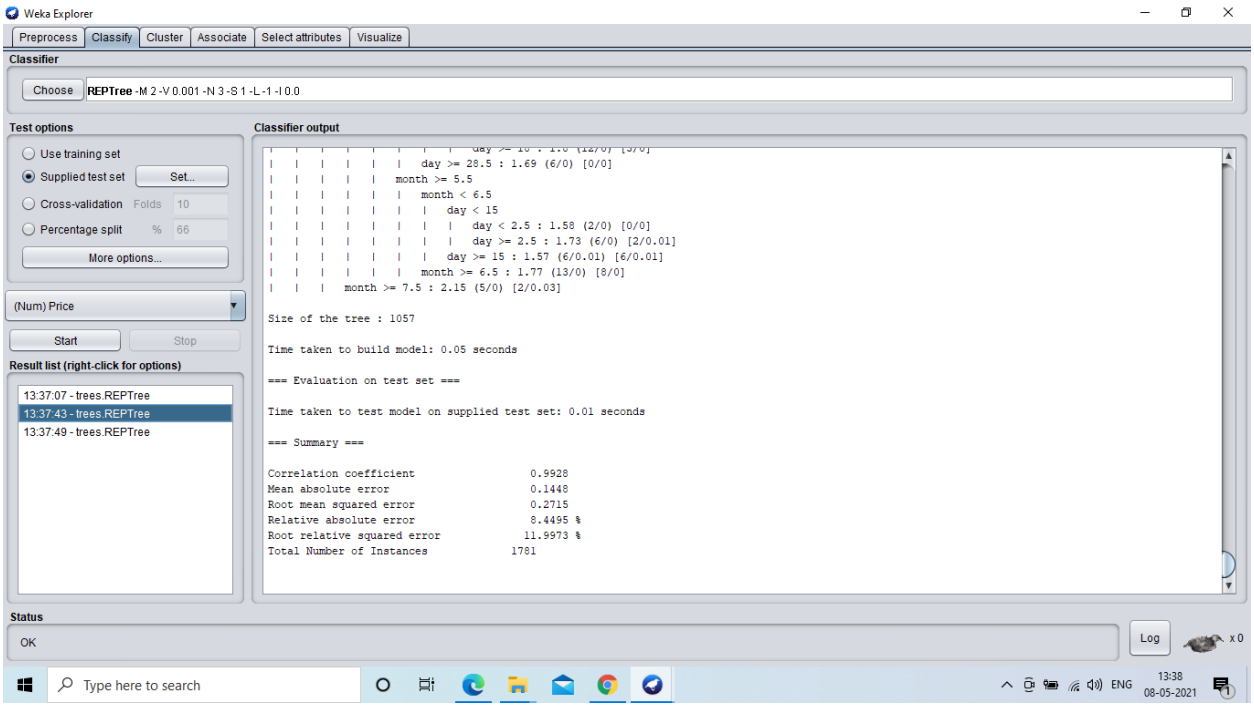
From the above two box plots it is observed that increase in price in 2005 reached a peak and that in a year the starting and ending months see the highest price of gas.

5 Flow Chart


Data Acquisition
Data Transformatiom
Data Cleaning
Data Visualisation
Build Machine Learning Models
Model Tuning
Test Model using 10 fold cross validatio and test data in Weka

6 Result

A test score of 0.99 and mean absolute error of 0.14 was obtained using the decision tree(i.e; REPTree) algorithm in Weka when the test set(which was randomly created) was supplied and thus successfully creating a model for Natural gas price prediction.



7 Advantages and Disadvantages

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.

The disadvantage is that new data should be taken and analysed continuously for more accurate prediction. That is it is harder to predict for long term.

8 Applications

- Economic planning
- Energy investment
- Environmental conservation

9 Conclusion

Natural gas will peak and fall drastically and abruptly. But, it helps 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.

10 Future Scope

Implementation of UI for this problem and thus make it possible for layman to understand the different trends that are found when analysing data.

11 Project link

https://github.com/smartinternz02/SPS-11180-Natural-gas-price-prediction