PETROL PRICE FORECASTING

CONTENTS

OBJECTIVE

The objective of this project is to perform the price trend forecast of petrol for time series.

ABSTRACT

Time series data abound in many realistic domains. The proper study and analysis of time series data help to make important decisions. Study of such data is very useful in many applications where there are trendy changes with time or specific seasonality as in electricity demand, cloud workload, weather and sales, cost of business products, etc. By understanding the nature of the time series and the objective of analysis, we have used different approaches to learn and extract meaningful information that can satisfy the business needs. The present paper covers and compares various forecasting algorithmic approaches and explores their limitations and usefulness for different types of time series data in different domains.

INTRODUCTION

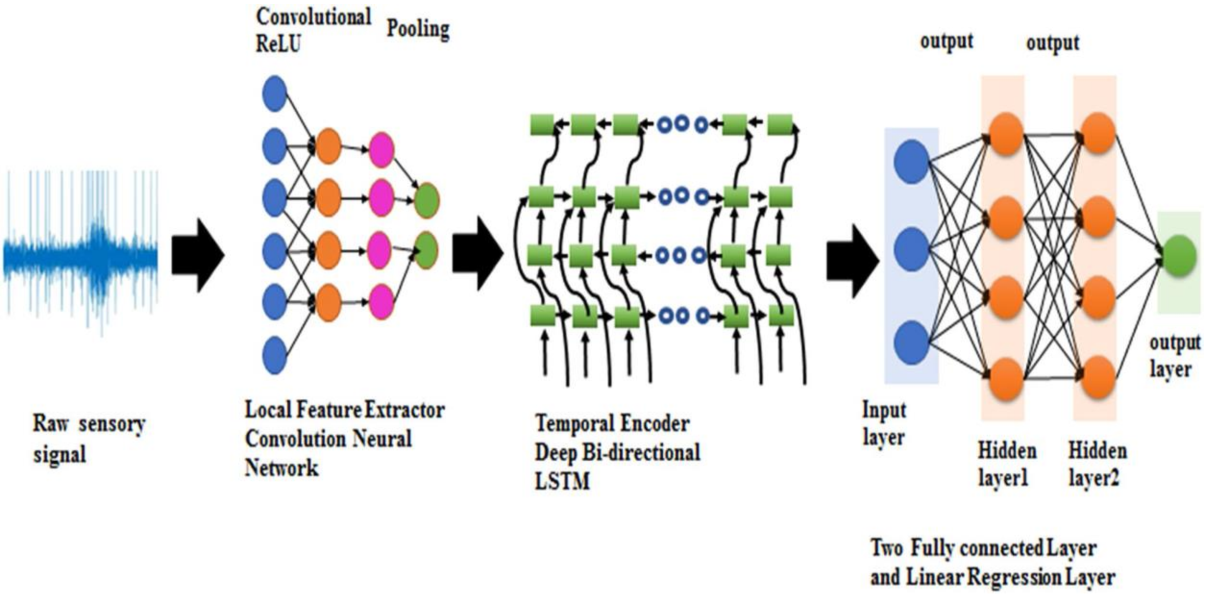
The concept of time series analysis has come a long way right since its inception. Plenty of research has been done in time series analysis to accomplish multiple objectives. Even in today’s machine learning and deep learning era, time series forecasting plays a crucial role to make important business decisions. Since the time series data are in real time, it can be used in organizations such as power management, stock market, health care, business, marketing, weather forecasting and many more. Time series prediction is essentially a part of temporal data mining and statistics. It is the process of careful collection and rigorous study of data that has been collected over a continuous period and development of a proper model that describes the inherent trend of the series. Time series forecasting is a process to predict the future with the help of history data.

Time Series Analysis

Time series data basically consist of systematic and non-systematic components. A systematic component has a consistency and recurrence that can be described by the model. A non-systematic component is a random phenomenon that cannot be directly modelled. The systematic component can be thought of a combination of level, trend, seasonality and the non-systematic component as a random variable or noise. Therefore, a decomposition technique extracts the trend and seasonal factors from time series data that helps to build the forecasting model. The more extensive decomposition includes local change effects, long-run effects, seasonal effects, etc. The removal of seasonal effects from data makes clear visibility of trend. The decomposition includes ways to find level, trend, seasonality, and noise from the data. The level is considered as mean pattern of the series while trend is harmonic pattern of the series. Seasonality is a very interesting characteristic, it is a short-term cyclic change over a long time series data, and finally some uncertain variation represents “noise” in time series data. All these four components may exist in additive or multiplicative form in the series data. However, the ultimate objective is to extract various patterns and an analysis can be made for depicting forecasts or future predictions which in turn helps in the growth of the business. But in this survey, prediction concept and the algorithms are the main foci. The kind of data that is considered here is called temporal data as it changes frequently with time and is represented using time stamps.

Time Series Forecasting Techniques

RNN (LSTM) models are one of the best models for extraction patterns of the input features and used to span over a long sequence. They can model problems seamlessly with multiple input variables and this ads as a great benefit in time series forecasting, where simple classic regression methods can be difficult to adapt to multiple input and multivariate forecasting problems. In simple words, LSTM provides a lot of flexibility over other methods. LSTMs are a form of recurrent neural networks (RNNs). They have the ability to retain information for long time with the help of their inner cells which can carry information unchanged. The network has control of the cell state where it can edit, add or delete information using gates as shown below. Mathematically, the main aim of LSTM is to find the conditional probability $p(y1,\ldots, yN \mid x1,\ldots, xN)$ where $(x1,\ldots, xN)$ are the input sequences and $(y1,\ldots, yN)$ are the corresponding output $\Phi(B\ (2)\ s)\_(B)(xt - \_) = \_(Bs)\_(B)wt$ sequences in fixed-dimensional representation. In case of LSTM, the model is learning one part of the time series prediction and wants to apply it to another part which is not able to find local changes in time series. Hence, we can conclude that LSTMs are not supposedly good for short period forecasting where learning on linearity's and their "stationarity" is less brittle.



PROBLEM STATEMENT:

The ONGCF is an organisation dedicated to the exploration and production of oil and natural gas. Price information is supplied on a weekly basis. It seeks to forecast crude oil prices for the following 16 months, from January 1, 2019 to April 1, 2020.

The main goal is to predict the forecast the prices based upon the best model as per your choice.

DATA REQUIREMENT:

In this model we going to predict the future trend of the petrol price, we need the data of history of petrol price with time.

TOOL USED:

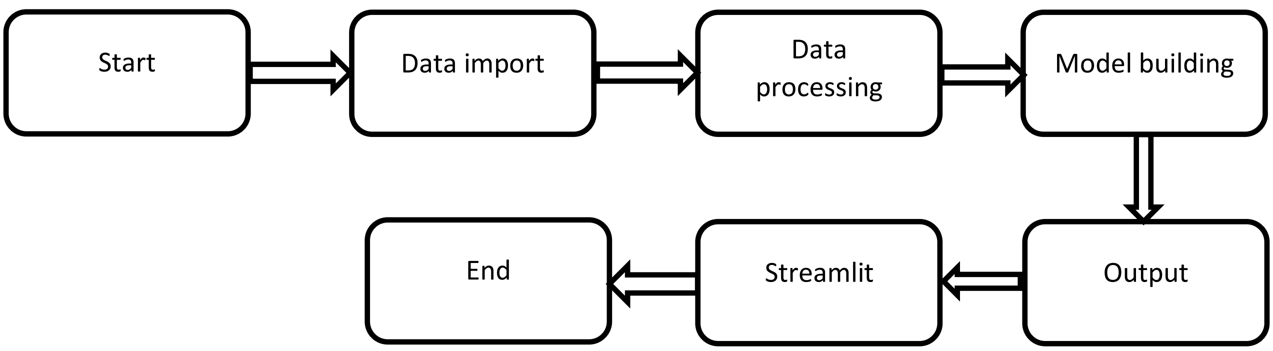
Python programming language and frameworks such as NumPy, Pandas, Scikit-Learn, Tenser Flow and Keras are used to build the whole model.

DESIGN DETAILS:

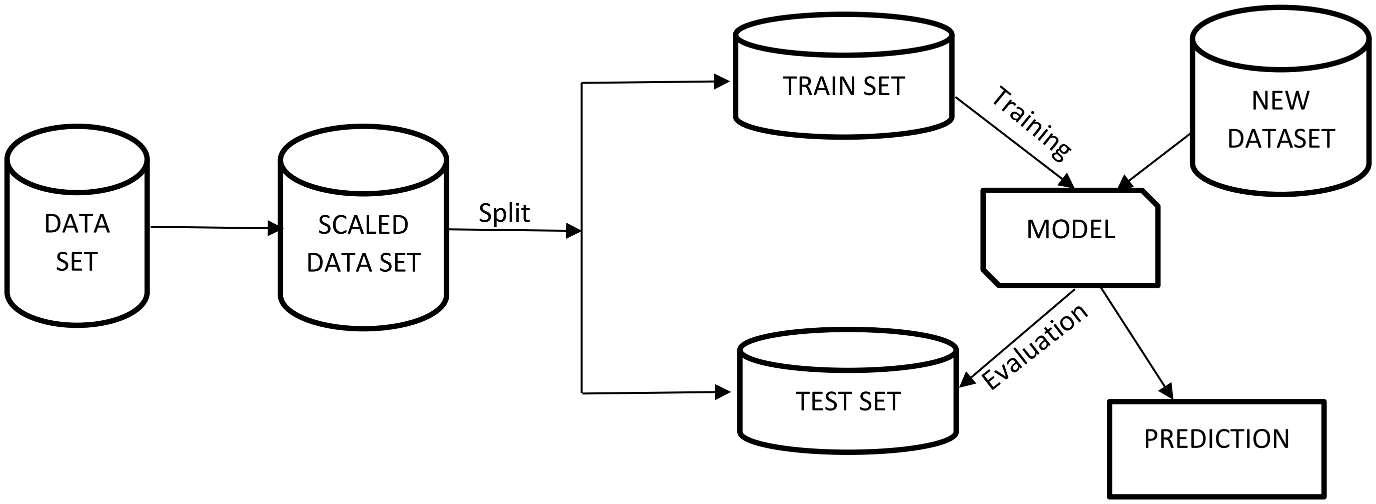
PROCESS FLOW:

For identifying the trend of future price of the petrol, we will use a deep learning base model .below is the process flow diagram is as shown below.

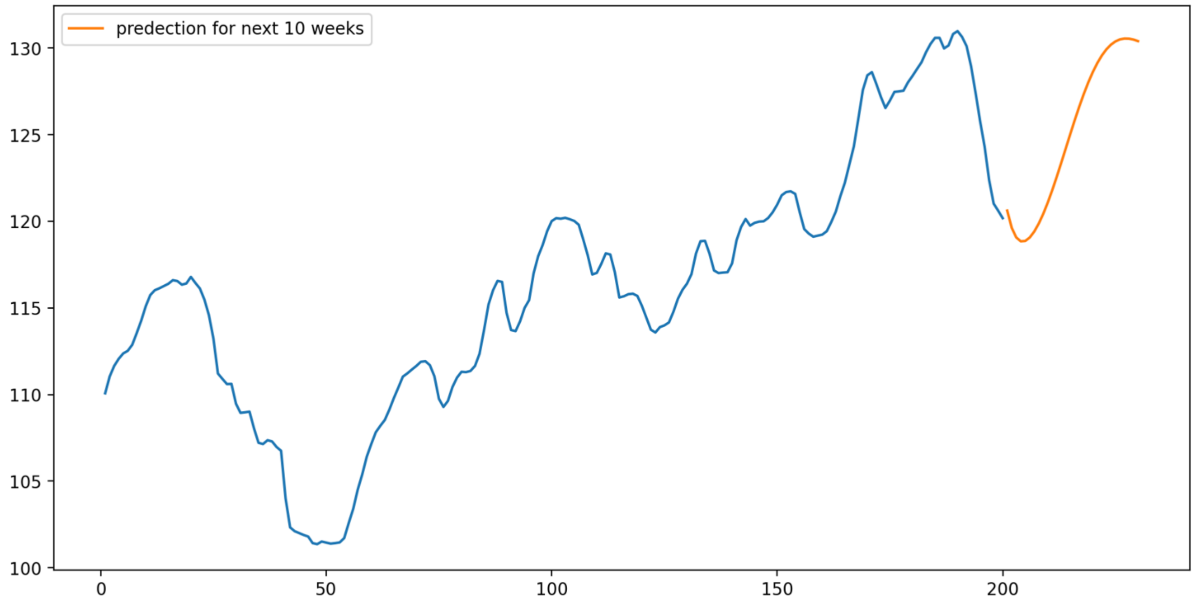
PROCESS METHODOLOGY:



MODEL TRAINING AND EVALUATION:



CONCLUSION:



The prediction says that the price trend of petrol will increase for next 10 weeks.