

Bachelor

Introduction

Accurate forecasts of natural gas prices are of great importance for a variety of stakeholders, including energy companies, consumers, financial markets, policymakers, and international traders.

For energy companies, natural gas is a key input for electricity production, heating, and industrial processes. Accurate price forecasts enable energy companies to manage production costs, optimize supply chain logistics, and make informed decisions about investments in infrastructure.

For consumers, natural gas is a critical energy source for heating, cooking, and hot water. Accurate price forecasts enable consumers to budget for their energy expenses and make informed decisions about their energy usage.

In financial markets, natural gas prices have a significant impact on commodity futures markets, stock markets, and currency markets. Accurate price forecasts can help investors make informed decisions and manage risk, reducing price volatility and ensuring economic stability.

Natural gas is also increasingly traded on global markets, and accurate price forecasts are critical for importers and exporters to make informed decisions about supply contracts and investments in infrastructure.

Finally, environmental policy plays a key role in energy production and consumption. Accurate price forecasts can inform policy decisions about energy production and consumption, and help guide efforts to reduce greenhouse gas emissions.

In summary, accurate forecasts of natural gas prices are crucial for ensuring energy security, promoting economic stability, managing production costs, reducing price volatility, and supporting efforts to address climate change.

Methodology

Time Series

A time series is a sequence of data points collected over time, where the time intervals between each data point are equally spaced. These data points can be anything that can be measured or observed, such as stock prices, temperature readings, or website traffic.

Time series data can be analyzed to identify patterns or trends over time, such as seasonal fluctuations, long-term trends, or short-term changes. This type of analysis can be useful for predicting future values, identifying anomalies or outliers, or understanding the underlying factors that contribute to the data.

Time series data is commonly used in a variety of fields, including finance, economics, meteorology, engineering, and many others. There are many statistical techniques and machine learning algorithms that can be applied to analyze and model time series data, such as ARIMA models, exponential smoothing, and neural networks.

Time series data is important because it allows us to analyze and understand how a variable changes over time. This can be useful in many different contexts, such as predicting future values, identifying patterns and trends, detecting anomalies or outliers, and understanding the underlying factors that contribute to the data.

Machine learning techniques can be particularly useful for analyzing time series data because they can automatically learn patterns and relationships in the data that may be difficult or impossible to detect manually. Some examples of machine learning techniques that can be applied to time series data include:

Time series forecasting: This involves using historical data to predict future values of a time series. Popular techniques for time series forecasting include ARIMA models, exponential smoothing, and machine learning algorithms such as neural networks.

Anomaly detection: This involves identifying unusual or unexpected patterns in a time series that may indicate a problem or opportunity. Machine learning techniques such as clustering, support vector machines (SVMs), and decision trees can be used for anomaly detection.

Classification: This involves assigning a label or category to each data point in a time series. For example, a time series of electrocardiogram (ECG) readings could be classified as normal or abnormal. Machine learning algorithms such as k-nearest neighbors (KNN), decision trees, and support vector machines (SVMs) can be used for classification.

Pattern recognition: This involves identifying recurring patterns or trends in a time series that may be useful for predicting future values or understanding underlying factors. Machine learning techniques such as recurrent neural networks (RNNs), convolutional neural networks (CNNs), and Hidden Markov Models (HMMs) can be used for pattern recognition.

Overall, machine learning can be a powerful tool for analyzing time series data, and can help us gain insights and make predictions that would be difficult or impossible to achieve with traditional statistical methods.

Autoregressive Processes

Autoregressive Processes, or AR processes for short, are a class of time series models that are commonly used for time series forecasting. The basic idea behind an AR process is that the value of a time series at a given time point depends on its past values. In other words, the model uses a linear combination of the past values of the series to predict the next value.

Mathematically, an AR process of order p , denoted $AR(p)$, is defined as:

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + e_t$$

where X_t is the value of the time series at time t , c is a constant, $\phi_1, \phi_2, \dots, \phi_p$ are the autoregressive coefficients, e_t is a random error term with mean 0 and variance σ^2 , and p is the order of the process.

To make a prediction for the next value of the series, we use the past p values of the series to calculate the predicted value:

$$\hat{X}_{t+1} = c + \phi_1 X_t + \phi_2 X_{t-1} + \dots + \phi_p X_{t-p+1}$$

The coefficients $\phi_1, \phi_2, \dots, \phi_p$ are typically estimated using a technique called maximum likelihood estimation, which involves finding the values of the coefficients that maximize the likelihood of observing the observed values of the series.

One important property of AR processes is stationarity, which means that the mean and variance of the series are constant over time. Stationarity is important because it allows us to make predictions that are accurate over time, and also simplifies the modeling process. To ensure stationarity, the roots of the AR characteristic polynomial must lie outside the unit circle, which can be checked using the roots of unity test.

Overall, AR processes are a simple but powerful class of time series models that can be used for time series forecasting and prediction. They are widely used in practice, and many more advanced time series models build on the basic framework of AR processes.

Experiments

Source of Data