The study of time series forecasting encompasses a range of methodologies, each with unique attributes suited to different kinds of data and forecasting needs. Our project delves into this realm, examining the effectiveness of both traditional statistical models and advanced neural network techniques in predicting future data points in a time series. The project is inspired by various scholarly works that have applied these methods across diverse domains, from financial market predictions to weather forecasting.

**SARIMA and SARIMAX Models:**

**(1) SARIMA Model:**

**What It Is:**

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is a prominent statistical method for time series forecasting, particularly effective in data with seasonal patterns. It extends the ARIMA model by incorporating seasonality.

**Components:**

* AR (p): Autoregression component, where 'p' represents the number of lag observations.
* I (d): Integration or differencing component, where 'd' is the number of differencing required to make the series stationary.
* MA (q): Moving average component, with 'q' indicating the size of the moving average window.
* Seasonal Elements: Comprises seasonal AR, differencing, and MA elements, denoted as (P, D, Q)s, where 's' indicates the seasonality period.

(1 - \sum\_{i=1}^{p} \phi\_i B^i)(1 - \sum\_{i=1}^{P} \Phi\_i B^{si})(1 - B)^d(1 - B^s)^D y\_t = (1 + \sum\_{i=1}^{q} \theta\_i B^i)(1 + \sum\_{i=1}^{Q} \Theta\_i B^{si})\varepsilon\_t

**SARIMA Model Hyperparameters:**

p (Autoregressive Order):

* Determines the number of lagged terms of the series in the model.
* Higher values can capture longer historical dependencies but might lead to overfitting.

d (Differencing Order):

* The number of times the data needs to be differenced to achieve stationarity.
* Excessive difference can lead to loss of information and over-differencing.

q (Moving Average Order):

* Specifies the size of the average moving window.
* It captures the relationship between an observation and a residual error from a moving average model applied to lagged observations.

P, D, Q (Seasonal Components):

* Seasonal counterparts to p, d, q.
* P, Q are similar to p, q but for the seasonal part of the model.
* D is the order of seasonal differencing. It removes seasonal patterns to make the series stationary.

s (Seasonal Period):

* The length of the seasonal cycle.
* For instance, s=12 for monthly data with annual seasonality.

**Uses of SARIMA Model:**

* Seasonal Time Series Forecasting: Effectively predicts data with clear seasonal patterns, like monthly sales or yearly climate variations.
* Economic Data Analysis: Used in analyzing and forecasting economic indicators such as GDP, inflation rates, and employment statistics.
* Inventory Management: Helps in predicting future inventory requirements based on historical demand patterns.
* Energy Consumption Forecasting: Useful in anticipating future energy demands, particularly with seasonal fluctuations.

**Strengths and Limitations of SARIMA Model:**

* **Strengths:**
* Excels in forecasting when data exhibits a clear seasonal pattern.
* Robust in handling stationary time series after differencing.
* Interpretability of model components (AR, MA, Seasonality).
* **Limitations:**
* Requires the time series to be stationary or made stationary through differencing.
* Not suitable for handling high-frequency data or very long series due to computational constraints.
* Cannot incorporate external variables influencing the series.

**(2) SARIMAX Model**

**What It Is:**

SARIMAX stands for Seasonal ARIMA with eXogenous variables. It builds upon SARIMA by including exogenous (external) variables, making it more versatile for complex datasets where additional factors influence the time series.

**SARIMAX Model Hyperparameters:**

Similar to SARIMA, with the addition of parameters for the exogenous variables.

Exogenous Variables (X):

* External variables or predictors that are believed to influence the forecasted variable.
* Their selection is crucial as irrelevant or poorly measured exogenous variables can degrade model performance.

**Uses of SARIMAX Model:**

* Incorporating External Factors: Ideal for scenarios where external variables (like economic policies or marketing efforts) influence the time series.
* Enhanced Economic Forecasting: More accurate economic forecasting by considering external economic indicators or policies.
* Environmental Modeling: Useful in ecological studies, factoring in external environmental factors for more accurate predictions.
* Market Analysis: Helps in understanding and predicting market trends by considering external influences like regulatory changes or competitors' actions.

**Strengths and Limitations of SARIMAX Model:**

* **Strengths:**
* Inherits all the advantages of SARIMA, with the added ability to include exogenous variables.
* More versatile in real-world scenarios where external factors impact the time series.
* **Limitations:**
* Complexity increases with the addition of exogenous variables, requiring careful selection and understanding of these variables.
* Like SARIMA, struggles with non-stationary high-frequency data.

**LSTM Models:**

**(3) LSTM-Univariate Model:**

**What It Is:**

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network designed to handle the limitations of traditional RNNs, particularly in learning long-term dependencies. In univariate LSTMs, the model predicts future values based on a single time series.

**Components:**

Cells with Gates: Each cell in LSTM has an input gate, output gate, and a forget gate, which collectively decide what information to retain or discard.

**Core equations of LSTM cells:**

Forget Gate: f\_t = \sigma(W\_f \cdot [h\_{t-1}, x\_t] + b\_f)

Input Gate: i\_t = \sigma(W\_i \cdot [h\_{t-1}, x\_t] + b\_i)

Cell State Update: \tilde{C}\_t = \tanh(W\_C \cdot [h\_{t-1}, x\_t] + b\_C)

Final Cell State: C\_t = f\_t \ast C\_{t-1} + i\_t \ast \tilde{C}\_t

Output Gate: o\_t = \sigma(W\_o \cdot [h\_{t-1}, x\_t] + b\_o)

Final Output: h\_t = o\_t \ast \tanh(C\_t)

**LSTM Model Hyperparameters:**

Number of Layers:

* Determines the depth of the LSTM network.
* More layers can model more complex patterns but increase the risk of overfitting and computational cost.

Number of Neurons per Layer:

* Controls the capacity of the network at each layer.
* More neurons allow the model to learn more complex representations but can lead to overfitting.

Learning Rate:

* The step size at each iteration while moving toward a minimum of a loss function.
* Too high a learning rate can cause the model to converge too quickly to a suboptimal solution, while too low a rate can make the training process very slow.

Epochs:

* The number of times the learning algorithm will work through the entire training dataset.
* More epochs can lead to better learning, up to a point where it starts overfitting.

Batch Size:

* The number of training examples utilized in one iteration.
* A smaller batch size often provides a regularization effect and lower generalization error.

Activation Functions:

* Functions like sigmoid, tanh, and ReLU that are used to determine the output of neural network nodes.
* They introduce non-linear properties to the network.

Dropout Rate:

* A regularization technique where randomly selected neurons are ignored during training to prevent overfitting.

Optimizer:

* Algorithms like Adam, SGD, RMSprop used to change the attributes of the neural network such as weights and learning rate to reduce losses.

**Uses of LSTM-Univariate Model:**

* Stock Price Prediction: Predicting future stock prices based on historical price data.
* Univariate Time Series Forecasting: Effective in scenarios with single-variable time series data, like temperature trends or individual product sales over time.
* Pattern Recognition: Identifying patterns or anomalies in univariate time series data, like detecting unusual spikes in network traffic.
* Signal Processing: Used in filtering, smoothing, and interpreting signal data over time.

**Strengths and Limitations of LSTM-Univariate Model:**

* **Strengths:**
* Capable of capturing long-term dependencies in data.
* Effective in handling non-linear patterns in time series.
* Suitable for large datasets and high-frequency data.
* **Limitations:**
* Requires substantial computational resources for training.
* May overfit on smaller datasets without proper regularization.
* Less interpretable than traditional statistical models.

**(4) LSTM-Multivariate Model:**

**What It Is:**

Multivariate LSTM models extend the univariate approach to incorporate multiple input features, making them suitable for complex scenarios where the target variable is influenced by several factors.

**LSTM Model Hyperparameters:**

Similar to the univariate model, with additional considerations for the interactions between different input features.

**Uses of LSTM-Multivariate Model:**

* Complex Time Series Forecasting: Handling and predicting scenarios where multiple variables interact, such as economic indicators.
* Sensor Data Analysis: Analyzing and forecasting based on data from multiple sensors, like in IoT applications or environmental monitoring.
* Financial Market Analysis: Incorporating various market indicators to predict stock or commodity prices.
* Predictive Maintenance: Anticipating equipment failures in industrial settings by analyzing multiple sensor readings.

**Strengths and Limitations of LSTM-Multivariate Model:**

* **Strengths:**
* Handles multiple input variables effectively, capturing complex interactions in the data.
* Retains the advantages of univariate LSTMs in learning long-term dependencies.
* **Limitations:**
* High computational complexity and data requirements.
* The complexity of the model can lead to challenges in interpretation and overfitting.

**Conclusion:**

* Model Selection Criteria: The choice between SARIMA/SARIMAX and LSTM models depends on the nature of the data, computational resources, and the specific requirements of the forecasting task.
* Use SARIMA/SARIMAX when dealing with seasonal data and when interpretability is crucial, especially if the impact of known external variables needs to be accounted for (SARIMAX).
* Opt for LSTM (univariate or multivariate) when dealing with large datasets or data with complex, non-linear patterns and long-term dependencies.

**Complementary Use:**

In practice, a combination of these models can be used, where traditional models like SARIMA/SARIMAX offer a strong baseline and LSTMs provide advanced modeling capabilities for more complex scenarios.

Ultimately, the effectiveness of a model in time series forecasting is contingent upon its alignment with the data characteristics, the forecasting horizon, and the specific use case requirements.