Forecasting Weekly Demand for a Soft Drink Product A Time Series Modeling Challenge

Imadul Islam Chowdhury
Business Insights and Analytics
Longo Faculty of Business
Humber Polytechnic
Toronto, Canada
N01588334@humber.ca

***Abstract*—predicting the demand remains one of the most challenging tasks in the production domain due to the inherent volatility and complexity of understanding the user needs. This paper examines the application of deep learning techniques to predict the demand, with a focus on models capable of capturing the non-linear patterns and long-term dependencies in production time series data. Traditional machine learning approaches like SARIMA has demonstrated varying degrees of effectiveness in forecasting short-term fluctuations. However, these models often struggle with the chaotic nature of the demand. In contrast, deep learning models, specifically Long Short-Term Memory (LSTM) networks, offer significant advantages by addressing the challenges of long-term dependencies and temporal data patterns. The paper evaluates multiple machine learning models, including LSTM, Prophet and SARIMA highlighting their ability in terms of predictive accuracy and robustness. However, through a series of experiments and comparisons, the study demonstrates that SARIMA model is particularly effective for predicting the demand as our dataset is a short term. The paper concludes by discussing the potential of the models and suggesting future research directions to optimize model architectures, enhance computational efficiency, and integrate novel data sources for more accurate production predictions.**

***Index Terms*—Demand Prediction, LSTM, SARIMA, Prophet.**

I. Introduction

Recently, A soft drink manufacturing company that distributes its main product nationwide has experienced fluctuations in the demand for this product, resulting in occurrences of both stock shortages and surplus in their inventory. These differences have led the company to a considerable financial implication, primarily due to increased logistics costs and product spoilage.

Considering this situation, the company has granted access to their historical weekly demand data for the product in question. The primary objective is to analyze the time series characteristics of this data and to construct a forecasting model that will accurately predict the demand for the next two weeks. The insights gained from this forecast will be instrumental in shaping future production strategies, enabling the company to better align its supply with consumer demand and mitigate the risks associated with inventory mismanagement.

To come up with a solution this essay looks at several forecasting models, lists their performances and recommendations for future production planning.

II. Methodology

a) Dataset Description

The dataset consists of three distinct columns: DATE, TIME, and DEMAND. The DATE column is formatted to represent weeks in the standard format of YYYY-MM-DD, which allows for easy identification of specific weeks within a calendar year. The TIME column denotes the hour of the transaction, utilizing a 24-hour clock format to provide precise timing. Lastly, the DEMAND column captures the total quantity of units sold during the specified week, offering valuable insights into sales performance and consumer behavior over time.

b) Data Preparation & Exploration

Python code has been to exploration and transformation for data. As par our findings, there are 1323 number of rows, and no null values was found.

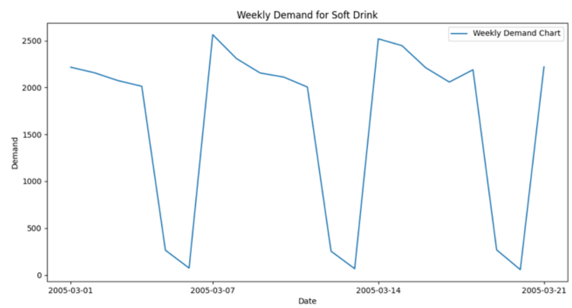

Fig 1: weekly demand

The time series graph presents a detailed depiction of fluctuations in weekly demand over a specified period. Upon careful examination, it becomes evident that the plot does not indicate any clear trends or seasonal patterns that could suggest a consistent increase or decrease in demand. Instead, the data appears to exhibit a level of randomness, with variations that do not conform to predictable cycles or trends, highlighting the complexity of consumer behavior during the observed timeframe.

c) Stationarity

I performed the Augmented Dickey-Fuller test to check the stationarity of data and found ADF: -0.90, p-value: 0.78 and Critical Values: {'1%':-4.01), '5%':-3.10), '10%':-2.69)} which demonstrates the time series is not stationary. Furthermore, I checked the rolling mean and standard deviation.

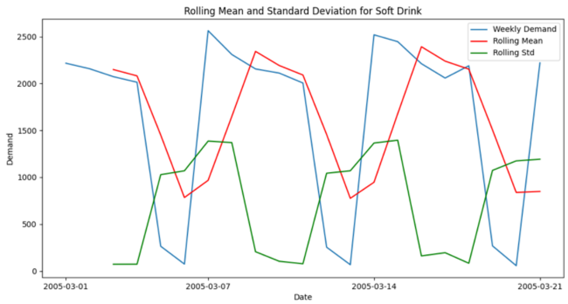

Fig 2: rolling mean and standard deviation

The graphical representation showcasing the rolling mean and standard deviation provides a comprehensive view of the fluctuations in weekly demand. This analysis highlights the inherent variations that occur over time, reflecting the dynamic nature of consumer behavior. Nevertheless, it is important to note that the data does not reveal any clear trends or seasonal patterns, suggesting that the demand remains relatively unpredictable and lacks consistent cyclical behavior.

d) Decomposition

An essential first step in comprehending the fundamental structure of time series data is time series decomposition. By breaking the series down into three separate parts trend, seasonality, and residual (noise) were better able to understand fluctuations and create precise forecasting models.

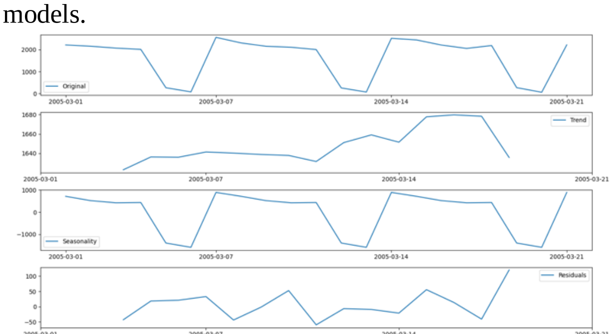

Fig 3: Additive Decomposition

The seasonal decomposition plot shows fluctuations in weekly demand. The seasonality component indicates a weekly pattern in the data.

III. FORECASTING MODELS

a. LSTM

Long Short-Term Memory (LSTM) networks enhance traditional Recurrent Neural Networks (RNN) by effectively handling long-term dependencies. LSTM utilizes cell states regulated by three gates: the input gate, output gate, and forget gate. The forget gate decides which information to discard, while the input gate allows new data into the cell state, enabling LSTMs to retain information from the distant past, a feature that RNNs do not possess [1].

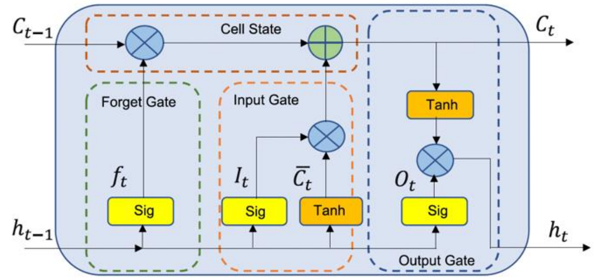

Fig 4: LSTM Architecture

The model uses ReLU activation to prevent vanishing gradients and improve training efficiency. LSTM is employed for its ability to handle long-term dependencies, predicting next two weeks demand.

b. SARIMA

ARIMA is a statistical method that integrates three key components: autoregression (AR), differencing (I), and moving average (MA). The AR part forecasts future values based on past data, the I part stabilizes the dataset through differencing to minimize seasonality, and the MA part enhances accuracy by considering previous errors [2].

The ARIMA model has a limitation in that it does not account for seasonality in time-series data with repeated cycles. To address this, the seasonal ARIMA (SARIMA) model is introduced, which incorporates the seasonal component in the ARIMA model and is capable of handling seasonality in time-series data [3].

To define the values of the (p, d, q) and (P, D, Q)s parameters for SARIMA, first need to identify d which is the number of times to difference the non-seasonal part of the series, and on the other side D is the number of times to variance of the seasonal part of the series. This is naturally done by using a stationarity test like the Augmented Dickey-Fuller (ADF) test, which tells the user whether a series is stationary or non-stationary. If the series is non-stationary, apply differencing until it becomes stationary. Next, identify the p and q using the ACF and PACF plots while observing where (and how) the autocorrelations and partial autocorrelations are cut off. Also use these plots to identify the seasonal components (P and Q). Finally, as for our dataset it has daily observations with weekly seasonality, I set s to 7 to capture the weekly cycles.

c. PROPHET

A Prophet is a forecasting tool introduced by Facebook. It is designed to analyse time series data with the additional capability of including non-linear trends with weekly and yearly seasonality and holiday effects. It can even tackle missing data easily. Prophet’s input is always a data frame with two columns: ds and y. However, also allows to augment other independent features that can contribute to forecasting [4].

IV. RESULTS

As par the model performances, the SARIMA model beats both LSTM and Prophet, achieving the lowest MAE (121.55), RMSE (165.86), and MAPE (1.07%), thus providing the most accurate forecasts.

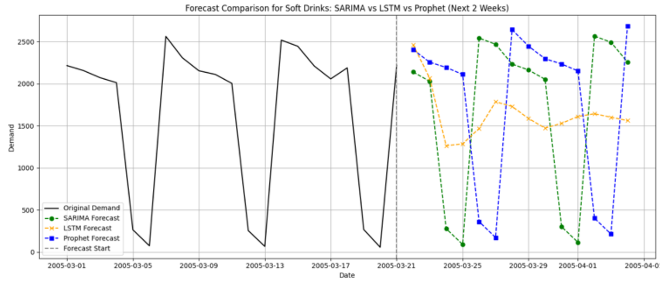

Fig 5: Forecast

LSTM performs moderately with higher error rates (MAE: 692.00, RMSE: 785.08, MAPE: 3.23%) but still outperforms Prophet, which has the highest error rates (MAE: 1251.02, RMSE: 1517.39, MAPE: 5.26%), making it the least reliable model.

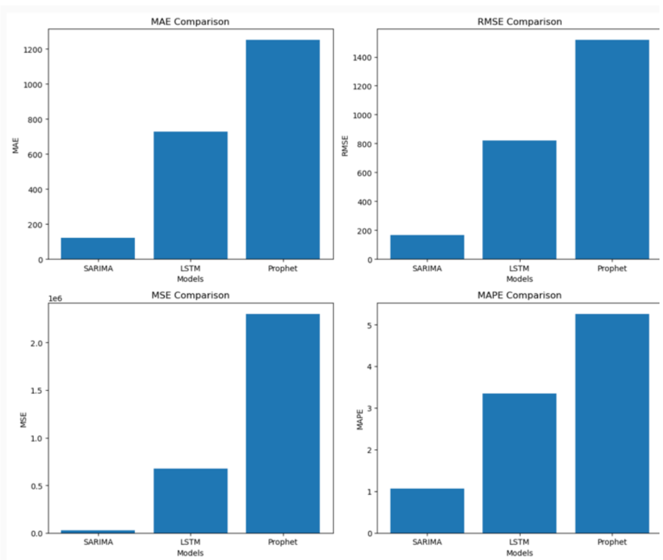

Fig 6: Model Comparison

Therefore, SARIMA is the most appropriate model for forecasting in this context.

V. RECOMMENDATION

With an average error of only 1.07% (MAPE), the SARIMA model shows strong predictive confidence, and the business should consider boosting production next two weeks. However, it is crucial to monitor external factors like market trends and unexpected disruptions, as they may impact forecast accuracy.

REFERENCES

[1] "Stock Market Prediction Using Deep Learning LSTM Model" by Shruti Goswami, IEEE International Conference on Smart Generation Computing, Communication and Networking, 2021. https://ieeexplore.ieee.org/document/9645837

[2] "Stock Price Prediction Using LSTM and ARIMA" by Apoorva Mahadik, Devyani Vaghela, and Amrapali Mhaisgawali, Second International Conference on Electronics and Sustainable Communication Systems (ICESC), IEEE Xplore, 2021. https://ieeexplore.ieee.org/document/9532655

[3] "Sales Forecasting: A Comparison of Traditional and Modern Times-Series Forecasting Models on Sales Data with Seasonality," by Y. Ali and S. Nakti, 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2023, pp. 159-163. https://ieeexplore.ieee.org/document/10112431

[4] "Time Series based Short term Load Forecasting using Prophet for Distribution system," by S. S. Kindalkar, A. R. Itagi, M. Kappali and S. Karajgi 2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Bangalore, India, 2022, pp. 1-6. https://ieeexplore.ieee.org/document/10084254