Order Demand Forecasting Through Customer Behavior and Seasonal Pattern

Prabahar Godwin James
Computer Science and Engineering,
Sri Sai Ram Institute of Technology,
Chennai, India.
Email : prabahar.cse@sairamtap.edu.in

Karthick M
Computer Science and Engineering,
Sri Sai Ram Institute of Technology,
Chennai, India.
Email : sit19cs000@sairamtap.edu.in

Syed Abuthahir A
Computer Science and Engineering,
Sri Sai Ram Institute of Technology,
Chennai, India.
Email : sit19cs150@sairamtap.edu.in

Tarun H
Computer Science and Engineering,
Sri Sai Ram Institute of Technology,
Chennai, India.
Email: sit19cs044@sairamtap.edu.in

***Abstract*— The Bullwhip Effect is a supply chain phenomenon wherein even slight changes in consumer demand at the store level will have an impact on all other supply chain participants. This frequently results in production and inventory imbalances as well as other issues that can harm the supply chain's effectiveness and profitability. In the supply chain, there are other factors that can affect the bullwhip effect, including Lack of communication may cause misunderstandings and incorrect readings of demand signals, which may result in overproduction, underproduction, or other inefficiencies. If the retailer's estimates were incorrect, it would have a significant negative impact on having too much inventory, raising the expense of maintaining it and lowering its profit of the firm. In order to solve the aforementioned issues and boost the company's profitability, an efficient forecasting technique must be used. This study suggests a deep learning model that, using data from the previous seven years—2011 to 2017—predicts product demand by examining a variety of variables. To anticipate the product demand sparked at the retailer level by the end customer, it specifically uses RNN-LSTM Model.**

Keywords—Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Deep Learning, Demand Forecasting, Neural Network.

I. INTRODUCTION

The technique of predicting future demand for a good or service is known as demand forecasting. For businesses to have the correct amount of inventory, manpower, and resources to meet consumer demand and optimize profitability, accurate demand forecasting is essential. This kind of information can assist merchants in determining the quantity of inventory needed to satisfy customers. If the quantity of inventory needed is known in advance, we can also optimize the manpower needed for maintenance, thus lowering the cost of maintenance.

The model is designed to forecast the ideal amount of stock needed for a specific product in a specific warehouse on a specific day. With historical data, which includes a retail store's past sales history, this will function. The information includes the product code, the warehouse where the item must be dispatched from, the category to which the item belongs, the date the demand was first noticed, and lastly the quantity of demand for the item.

II. RELATED WORKS

Hyojeoung Kim et al. [1] [2023] compared the CatBoost machine learning and CNN deep learning model and presented it as a single model CNN-CatBoost hybrid model prediction method that gives better performance. They also noticed that the accuracy changed when adding wind speed and precipitation to the hybrid model. Hyojeoung Kim [2023] proposed a solution for predicting solar radiation which will resolve the issues in solar energy due to climate change.

Marco Ratusny et al. [2] developed a framework, where they utilize data enrichment via synthetical training samples, Integrating synthetically generated data into the training phase allowed them to strengthen the inclusion of rare pattern variants that were identified during initial analysis. Actual customer data is used to benchmark the performance of the framework and it shows that the baseline CNN approach outperforms all available state-of-the-art benchmark models.

Uppala Meena Sirisha et al. [3w] studied the statistical methods- Autoregressive Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA) models, as well as the deep learning method, Long Short-Term Memory (LSTM) Neural Network model. The models were fitted and used to predict profit on test data, resulting in accuracies of approximately 93.84% (ARIMA), 94.378% (SARIMA), and 97.01% (LSTM). Forecasts for the next 5 years were made, and the results show that the LSTM method outperforms both statistical models in creating the best model.

Nithin Soundar S J et al. [4] proposed using a CNN- LSTM model to forecast retail demand. Equipped with the Swish Activation Function it works better than the traditional ReLU (Rectified Linear Unit). Data from 10 stores each consisting of 50 items are taken as input. The experiment results suggest using CNN- LSTM Model as it hasconsiderably lower RMSE (Root Mean-Squared Error).

Akshay Krishna et al. [5] have implemented normal regression techniques and as well as boosting techniques in our approach and have found that the boosting algorithms have better results than the regular regression algorithms. They observed that the AdaBoost algorithm has the highest RMSE value of 1350.72 and the algorithm with the least RMSE value is GradientBoost having 1088.64. It is concluded that without proper hyperparameter tuning the AdaBoost algorithm won't be able to perform as expected and the performance deteriorates.

This study done by Yeu-Shiang Huang, Chia-Hsien Ho, and Chih-Chiang Fang [6] considers a two-echelon supply chain with seasonal consumer demand, in which the impacts of the degree of information sharing on the supplier’s profits are investigated. Since the variance in the supplier’s inventory would marginally decrease as the degree of information sharing increases, the benefits gained by the supplier due to information sharing are thus a convex function. This can be used to obtain the optimal degree of information sharing with the aim of maximizing profits,the seasonal demand is described by a SARMA time series model. The results of sensitivity analyses show that the correlations of demand for successive periods and estimation errors would both have great effects on the benefits gained by information sharing.

III. METHODOLOGY

Fig 1: explains the process used to forecast product demand. It also demonstrates the data cleaning procedure, and finally, we obtain the dataset with the undesirable cells eliminated.

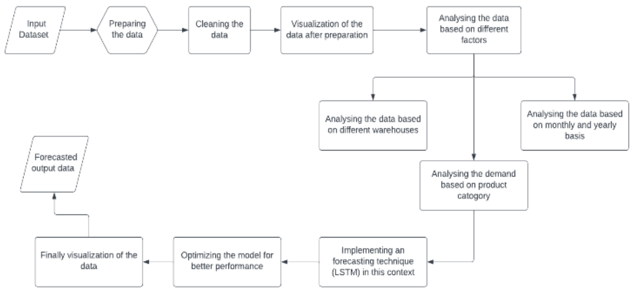


Fig 1: Block Diagram

The above block diagram shows the processing of the dataset from cleaning to final output this model specifically analyzes the dataset based on different factors which include warehouse-based analysis, analysis based on product demand, and analysis based on a monthly and yearly basis which would be assumed to have better accuracy in forecasting the demand for the product and meanwhile it can also be able to reduce the forecasts error which leads to the main problem of supply chain management that is bullwhip effect.

Fig 2: The graph below illustrates the warehouse-based analysis by designating each warehouse with a distinct hue. It depicts the demand induced by the end-user at the specific warehouse that is responsible for delivering the products within the region

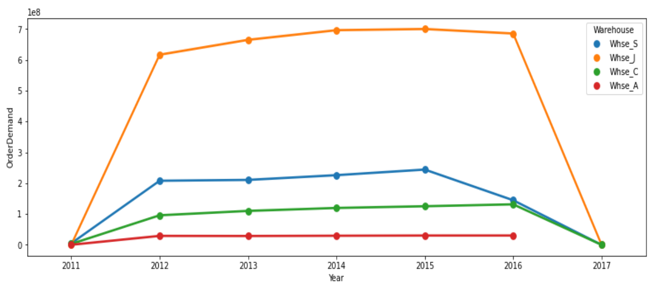


Fig 2: products demand at the specific warehouse

Fig 3: The graph below shows the monthly demand for the items from January to December from the years 2011 to 2017 as stated by the product code and categorized by the product category.

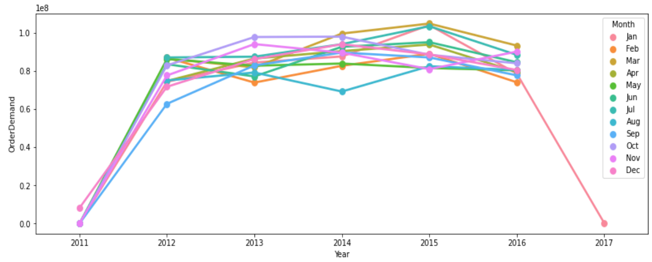


Fig 3: products demand on the monthly basis

The time-series algorithms should be translated into supervised machine-learning problems before being modeled using the method, i.e., a series of lists should be changed into combinations of input and output values. While a supervised machine learning issue comprises input and output values that a model can operate on, a time series can be imagined as columns of ordered values.

The data is then divided into train and test segments, with the train being 60% of the original data and the test segment comprising 40%.

A. Recurrent Neural Network

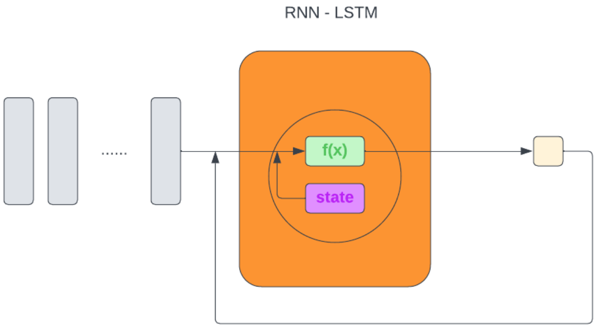
Recurrent neural networks (RNNs) are a type of artificial neural network designed for sequential data processing tasks. Unlike traditional feedforward neural networks, RNNs have loops in their network architecture that allow information to persist and be passed from one-time step to the next. This makes them particularly useful for tasks such as speech recognition, natural language processing, and time series analysis. The key advantage of RNNs is their ability to handle variable-length sequences of input data, which makes them well-suited for modeling real-world data that often have complex temporal dynamics. However, they can suffer from the problem of vanishing gradients, which can make training them more difficult than other types of neural networks. To overcome this, several variants of RNNs have been developed, such as long short-term memory (LSTM) networks and gated recurrent units (GRUs), which have proven to be very effective in many applications.

B. Long Short-Term-Memory Network Model

Recurrent neural networks (RNNs) of the Long Short-Term Memory (LSTM) variety are created to address the issue of vanishing gradients during training. This is done by capturing long-term dependencies in sequential data by selectively remembering or forgetting information from earlier time steps utilizing memory cells. An LSTM network's architecture comprises gates that regulate the information flow into and out of the memory cells. To improve the network's performance on a particular job, these gates can be changed during training.

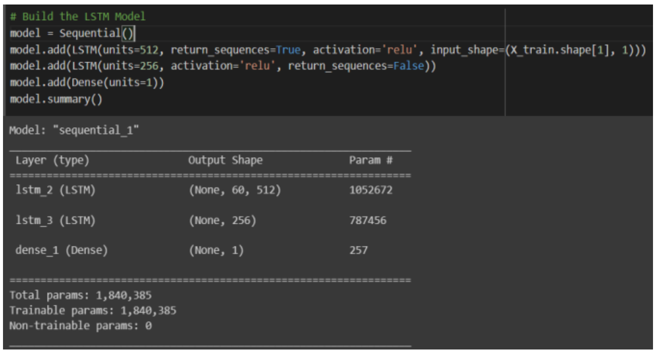
C. RNN-LSTM model

Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) cells are a popular type of neural network architecture that can handle sequential data by maintaining an internal state. LSTMs are particularly useful for solving problems with long-term dependencies, such as natural language processing and speech recognition. They are capable of learning long-term dependencies in the data by selectively retaining or discarding information at each time step based on a set of learned rules.



In contrast to ANN, the RNN model works by saving the output of one step and feeding it into the next stage of the input sequence. This procedure is repeated until the input layer's entire sequence has been processed.

Fig 4: It provides specific details regarding the model's executive summary.



```
# Build the LSTM Model
model = Sequential()
model.add(LSTM(units=512, return_sequences=True, activation='relu', input_shape=(X_train.shape[1], 1)))
model.add(LSTM(units=256, activation='relu', return_sequences=False))
model.add(Dense(units=1))
model.summary()

Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm_2 (LSTM)               (None, 60, 512)           1052672

 lstm_3 (LSTM)               (None, 256)               787456

 dense_1 (Dense)             (None, 1)                 257

=================================================================
Total params: 1,840,385
Trainable params: 1,840,385
Non-trainable params: 0
_________________________________________________________________
```

Fig 4: Model Summary

The mean squared error loss function was used to calculate the loss after fitting the model for 20 epochs.

D. *Activation Function*

One of the key mechanisms in neural networks that determine whether a neuron should be engaged or not is the activation function. Several well-liked activation functions exist. This paper uses the activation function named ReLU.

1) ReLU Activation Function:

A non-linear activation function is ReLU. Neurons are only deactivated when the output is zero. ReLU has an advantage over the other activation mechanisms in this situation.

Fig 5: The graph below represents the basic ReLU activation function.

f(y)=max (0, y)

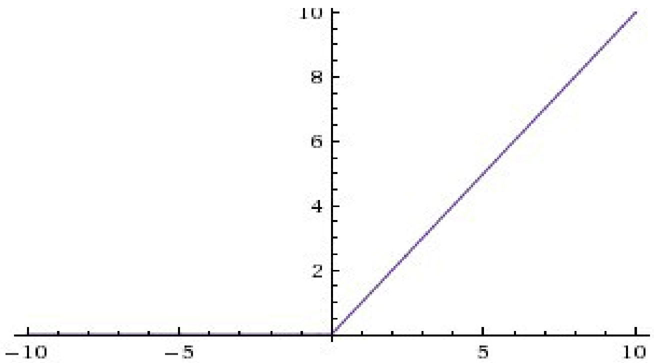


Fig 5 : ReLU Function

IV. TRAINING METHODS

We compared the performance of ETS, LSTM, and XGBoost models on a given dataset using MSE as the loss function and MAE as the performance metric. XGBoost, as a gradient boosting tree method, showed the best performance among the three models, while LSTM performed moderately well but required a larger dataset for training.: As shown in Table the XGBoost model outperformed the other two models, achieving the lowest MAE of 0.059. The LSTM model performed moderately well, achieving an MAE of 0.094, but took much longer to train due to the limited dataset. This suggests that the LSTM model requires a larger dataset for training, which was not available in this study. On the other hand, ETS did not perform well on this dataset, achieving an MAE of 0.149. This could be due to the fact that the ETS model was not able to capture the order effectively.

The sample dataset contains information about the attributes used in the data set for predicting the future demand for the product and that can be later used by the manufacturing firm to optimize the workflow inside their organization eventually resulting in eliminating the money spent on unwanted things and increase the profit of the organization.

Fig 6: It shows the basic attributes used in the dataset to predict the future demand of the product categorized by product code and product category.

1	Product_Code	Warehouse	Product_Category	Date	Order_Demand
2	Product_0993	Whse_J	Category_028	27-07-2012	100
3	Product_0979	Whse_J	Category_028	19-01-2012	500
4	Product_0979	Whse_J	Category_028	03-02-2012	500
5	Product_0979	Whse_J	Category_028	09-02-2012	500
6	Product_0979	Whse_J	Category_028	02-03-2012	500
7	Product_0979	Whse_J	Category_028	19-04-2012	500
8	Product_0979	Whse_J	Category_028	05-06-2012	500
9	Product_0979	Whse_J	Category_028	27-06-2012	500
10	Product_0979	Whse_J	Category_028	23-07-2012	500
11	Product_0979	Whse_J	Category_028	29-08-2012	500
12	Product_0979	Whse_J	Category_028	29-08-2012	500
13	Product_0979	Whse_J	Category_028	18-09-2012	500
14	Product_0979	Whse_J	Category_028	11-10-2012	500
15	Product_0979	Whse_J	Category_028	01-11-2012	500
16	Product_0979	Whse_J	Category_028	29-11-2012	500
17	Product_0979	Whse_J	Category_028	26-12-2012	500
18	Product_1159	Whse_J	Category_006	06-01-2012	50000
19	Product_1159	Whse_J	Category_006	18-01-2012	100000
20	Product_1159	Whse_J	Category_006	02-02-2012	50000
21	Product_1159	Whse_J	Category_006	22-02-2012	50000
22	Product_1159	Whse_J	Category_006	02-03-2012	50000
23	Product_1159	Whse_J	Category_006	09-03-2012	50000
24	Product_1159	Whse_J	Category_006	23-03-2012	50000
25	Product_1159	Whse_J	Category_006	06-04-2012	50000
26	Product_1159	Whse_J	Category_006	16-04-2012	50000
27	Product_1159	Whse_J	Category_006	07-05-2012	50000

Fig 6: Sample dataset

TABLE I OBSERVED MEAN SQUARED ERROR AND MEAN ABSOLUTE ERROR VALUE COMPARISON FOR ETS, XG-BOOST, AND LSTM MODEL

Model Name	**MSE (Mean Squared Error)**	**MAE (Mean Absolute Error)**
ETS	1.3192	0.149
XG-Boost	0.8219	0.059

LSTM	0.0387	0.1579

Fig 7: It shows the loss value for the given dataset using 20 Epochs and a batch size of 32 and the plot is drawn against the Loss and the corresponding epochs.

Fig 8: It represents the graph that clearly describes about the performance metrics of the given dataset and the model is plotted against the mean absolute error and its corresponding epoch.

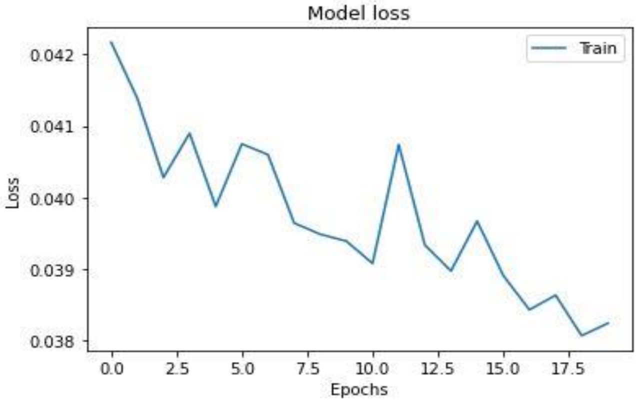


Fig 7: Model Loss for 20 Epochs using RNN-LSTM

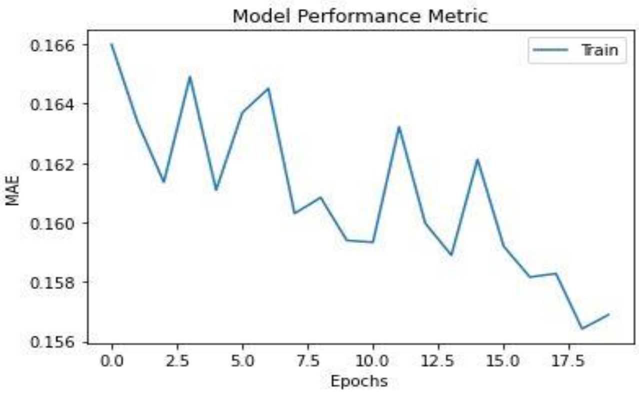


Fig 8: Performance metrics for 20 epochs

V. CONCLUSION

In this paper, we proposed an LSTM-based model for demand forecasting that attempts to resolve the bullwhip effect to an extent and in minimizing the forecasting error.The process of reducing the bullwhip effect in supply chain management is a challenging task as it depends on multiple factors such as lack of information sharing across borders, inaccuracy in the forecasted information, etc., In this paper, we have focused on a solution that will resolve one such factor of the occurrence of the bullwhip effect i.e, prediction of demand for the product. Our approach leverages the sequential nature of demand data and captures the complex relationships between historical and current demand, thereby providing a more reliable forecast. The proposed model has great potential for real-world application, particularly in industries where accurate demand forecasting is critical to maintaining efficient and profitable operations. The proposition aims to result in good satisfaction of the customer and eventually, it will also increase the profitability of the organization. our work contributes to the growing body of research aimed at improving demand forecasting methods and reducing the impact of supply chain disruptions.

REFERENCES

[1] H. Kim, S. Park, H. -J. Park, H. -G. Son and S. Kim, "Solar Radiation Forecasting Based on the Hybrid CNN-CatBoost Model," in IEEE Access, vol. 11, pp. 13492-13500, 2023, doi: 10.1109/ACCESS.2023.3243252.

[2] M. Ratusny, M. Schiffer and H. Ehm, "Customer Order Behavior Classification Via Convolutional Neural Networks in the Semiconductor Industry," in IEEE Transactions on Semiconductor Manufacturing, vol. 35, no. 3, pp. 470-477, Aug. 2022, doi: 10.1109/TSM.2022.3187100.

[3] U. M. Sirisha, M. C. Belavagi and G. Attigeri, "Profit Prediction Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison," in IEEE Access, vol. 10, pp. 124715-124727, 2022, doi: 10.1109/ACCESS.2022.3224938.

[4] S. S. J. Nithin, T. Rajasekar, S. Jayanthy, K. Karthik and R. R. Rithick, "Retail Demand Forecasting using CNN-LSTM Model," 2022 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2022, pp. 1751-1756, doi: 10.1109/ICEARS53579.2022.9752283.

[5] A. Krishna, A. V, A. Aich and C. Hegde, "Sales-forecasting of Retail Stores using Machine Learning Techniques," 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), Bengaluru, India, 2018, pp. 160-166, doi: 10.1109/CSITSS.2018.8768765.

[6] Y. -S. Huang, C. -H. Ho and C. -C. Fang, "Information Sharing in the Supply Chains of Products With Seasonal Demand," in IEEE Transactions on Engineering Management, vol. 64, no. 1, pp. 57-69, Feb. 2017, doi: 10.1109/TEM.2016.2623327.