Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

AI Driven Supply Chain Optimization for Deep Demand Forecasting in Grocery Retail

Vakkala Venusai[1], Gurram Sujankumar[2], Vemula Kumarswamy[3], Kota Lakshmi Narahari Reddy[4], Geetha D Devanagavi[5]

[1,2,3,4,5] Department of Computer Science & Engineering, Reva University, Bangalore, Karnataka, India

Abstract
Demand forecasting plays an important role for making the retail supply chains more sustainable and efficient, importantly in the grocery sector. Groceries are the things that does not last for the long time and the buying habits of people also changes a lot, so predicting or forecasting what they will buy is tricky. Recent studies show that traditional methods used for forecasting often does not work well when handling with the unpredictable and the complex sales trends [1]. It uses the advanced machine learning, deep learning techniques and the model is tested using actual Walmart sales data that includes the sales history as well as information about promotions, holidays, and economic conditions, which helped in making the results more accurate. Different forecasting methods are tested including the basic time series approaches and a machine learning model called XGBoost [2]. To make the model much better, a deep learning model was created which uses the LSTM networks along with Self Attention mechanisms, inspired by the recent research on the attention-based predictions [3]. The results show that these attention enhanced LSTM is much better for making the accurate predictions than the usual statistical methods. This model helps in reduction of the stock wastage due to expiry and maintaining the required stock to avoid the loss of sales. It allows for making the smarter decisions based on the data. This leads to the most efficient operations and less waste in the retail sector [5].

Keywords: Demand Forecasting, Supply Chain Optimization, LSTM, Self Attention, XGBoost, Retail Analytics.

1. Introduction

Demand forecasting is the important thing in managing the supply chain retail sector. In grocery stores predicting what quantity of each product the customer will buy is difficult because these items can spoil, the customer buying habits may change often and the factors like holidays, sales and the economy play's a big role. Changes in the prices, special offers and overall economic conditions all affect the how much people buy. Low forecast accuracy either could result in too much stock which wastes the money or not enough stock which hurt the customer and the sales for the store will be reduced. Research shows that poor forecast will affect the ability of the supply chain to adapt and make it difficult to run the business smoothly [6], [7].

In the retail industry conventional techniques like statistical time series models and moving averages are frequently used. But these methods works best only when the patterns are stable and follow a linear line which is not always the situation in today grocery stores that faces a lot of ups and downs [1], [9].

Business had started collecting the large amounts of the data from sales and the marketing campaigns because the retail sector expanded faster. This kind of data is very complicated and needs the better models which can handle the non-linear trends and the long-term patterns over the time [8]. Thanks to the advances in AI where the businesses are now using the machine learning and the deep learning techniques for demand forecasting. Models such as XGBoost have been particularly best when working with structured retail datasets as they can capture important patterns in historical sales information [2]. Also the LSTM networks are best for the time series forecasting because they remember the important patterns over long period of time [4]. New research has added that the attention mechanisms help's the models to focus more on the past important events and improving the prediction accuracy [3], [10]. Based on the advancements this paper introduces an AI based demand forecasting system that combines both the machine learning and deep learning techniques. It uses the sales history, promotions and holidays data to improve the weekly demand forecasting. The objective of this approach is to support the smarter inventory planning and reducing the wastage of product which improves the efficiency of the supply chain [5]. The remaining of this paper is structured as follows: Section 2 provides a comprehensive literature survey with a focus on approaches using the Artificial Intelligence for demand forecasting and supply chain optimization in the grocery retail sector. Section 3 provides the explanation about methodology and the techniques used in the study. Section 4 provides the comparison of the models using the performance metrics and graphical representations. Finally Section 5 gives the conclusion of the paper by summarizing the key findings and the future scope in AI based supply chain optimization.

2. Literature Survey

Some research shows that how artificial intelligence is used for predicting demand forecasting and supply chains in the grocery and retail industry. A deep learning model using by Bayesian methods, used to increase forecasting accuracy during unstable time like the COVID-19 [1]. Machine learning and deep learning models used to increase logistics and distribution networks in e-commerce platforms [2]. Hybrid models using both LSTM and CNN with attention mechanisms used for better forecasting accuracy in grocery sales [3]. Attention-based LSTM models are integrated with optimized algorithms have been proposed for dynamic pricing and replenishment in fresh grocery supermarkets [4]. Multi-agent deep reinforcement learning has been used to increase the demand forecasting and inventory management in retail supply chains [5]. AI-based models used to increase supply chains strength and risk management [6]. Machine learning models used for improving supply chain agility and sustainability by optimizing logistics and inventory management [7]. A review of deep learning models shows models such as CNN, LSTM, and GRU that it works well for predicting logistics [8]. Comparing of two different models shows that LSTM-based models that provides more accurate predictions than compared to regression models for food demand forecasting [9]. Studies shows that attention-based LSTM encoder-decoder models provides more accurate predictions for time series [10]. In summary, these studies show the importance of AI-based models for improving better accuracy for forecasting and efficiency in grocery retail supply chains.

3. Methodology

To address the problem of the demand forecasting in grocery retail, we choose a structured Artificial Intelligence based approach merging both Machine Learning and Deep Learning techniques. The methodology contains multiple steps that are data preprocessing, feature engineering, model development and performance evaluation. This framework is designed to forecast weekly sales of the store and individual product level using historical and external influencing factors.

3.1 Data Preprocessing

This project makes use of the Walmart store sales dataset, which provides weekly sales figures along with detailed information about each store, including promotional markdowns, holiday flags and relevant economic factors.

At first the dataset is moved to the location where the environment is set and all required libraries are installed. The preprocessing steps consists handling missing values or null values in promotional and economic features, removing duplicate records, concerting date variables into features such as year, month and week sorting data chronologically to preserve time-series order.

To ensure the consistency in the training of the model, numerical features are normalized using scaling techniques and categorical features such as holidays are encoded into numerical format. The dataset is then divided into two training and testing while ensuring that future records are not used to predict past values and avoiding any data leakage.

Figure 1 Feature Engineering of Dataset

	id	item_id	dept_id	cat_id	store_id	state_id	d	sales	date	wm_yr_wk	...	event_type_2	snap_CA	snap_TX	snap_WI	sell_
0	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_1	12	2011-01-29	11101	...	NaN	0	0	0	
1	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_2	4	2011-01-30	11101	...	NaN	0	0	0	
2	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_3	6	2011-01-31	11101	...	NaN	0	0	0	
3	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_4	0	2011-02-01	11101	...	NaN	1	1	0	
4	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_5	3	2011-02-02	11101	...	NaN	1	0	1	
5	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_6	10	2011-02-03	11101	...	NaN	1	1	1	
6	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_7	13	2011-02-04	11101	...	NaN	1	0	0	
7	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_8	1	2011-02-05	11102	...	NaN	1	1	1	
8	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_9	6	2011-02-06	11102	...	NaN	1	1	1	
9	FOODS_1_046_TX_2_validation	FOODS_1_046	FOODS_1	FOODS	TX_2	TX	d_10	2	2011-02-07	11102	...	NaN	1	1	0	

In Figure 1, It shows how feature engineering was carried on Walmart Retail dataset. The Original dataset is cleaned and it handles missing values or null values and removing duplicate records to ensure the dataset was accurate and consistent. The attributes such as week, month and year are captured from date fields to identify seasonality. Lag features and rolling statistical measures are then generated to represent past demand behavior. The processed features form the data is used by both machine learning and deep learning forecasting models.

3.2 Model Architecture

To evaluate the forecasting model performance two models are implemented.

3.2.1 XGBoost Model

XGBoost, a gradient boosting-based regression algorithm, is used as a machine learning baseline model. It captures non-linear relationships and the trends between the sales and variables which are essential. The both models are trained using the structured features that are generated during preprocessing.

3.2.2 LSTM Model

A Long Short-Term Memory (LSTM) network is implemented to model sequential sales data that is also known as the time series data. LSTM networks are capable of learning long-term dependencies in time-series data by maintaining internal memory states which is useful for the iterations of the whole process. Before passing the data into the network the data is reorganized into sequential format so that the model can understand the time-based patterns more effectively.

To improve both model's performance a self-attention layer is combined with the LSTM network layers. This attention mechanism that we used helps the model to identify past periods or patterns which are more important and giving higher weight to suitable weeks like holidays or promotional events. As a result the model produces better accurate forecast predictions.

3.3 Training and Evaluation

The data was split according to time. The 80 percentage of the dataset is used for training and the remaining 20 percent is used for the testing during the process. This process of training and testing is important because sales forecasting depends on sequence of the LSTM, so the future weeks cannot be used while training the model and improve the accuracy of the model.

We experimented with two models: XGBoost and LSTM. For XGBoost, different types of input features were tested, including previous week sales, rolling averages, promotional activity, and holiday indicators and many more features that are required for the forecast. After trying multiple and different parameter settings, the configuration that produced the lowest error is selected and therefore the error in the forecasting is reduced significantly.

Preparing the data for training of LSTM required a different approach. First the values are normalized and then arranged into the sequences using the sliding window format. This allowed the network to learn how demand changes from week to week. Training was done using the Adam optimizer and Early stopping was introduced to prevent the model from overfitting to the training data.

Finally the performance of both models was evaluated using MAE, RMSE, and MAPE. Lower values of this metrics indicate the better forecasting performance.

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Figure 2 Metrics of an XGBoost Model

```
# Explicit numeric feature list (VERY IMPORTANT)
feature_cols = [
    "lag_1", "lag_7", "lag_14",
    "rmean_7", "rmean_14",
    "sell_price",
    "snap_CA", "snap_TX", "snap_WI",
    "wday", "month"
]

# Use ONLY numeric features
X = train_df[feature_cols]
y = train_df["sales"]

# Time-based split
X_train, X_val, y_train, y_val = train_test_split(
    X, y, test_size=0.2, shuffle=False
)

# XGBoost baseline
model
    (function) learning_rate: Any
    learning_rate=0.05,
    max_depth=6,
    subsample=0.8,
    colsample_bytree=0.8,
    objective="reg:squarederror",
    random_state=42
)

# Train
model.fit(X_train, y_train)

# Predict
y_pred = model.predict(X_val)


# Metrics
rmse = np.sqrt(mean_squared_error(y_val, y_pred))
mae = mean_absolute_error(y_val, y_pred)


print(f"RMSE: {rmse:.2f}")
print(f"MAE   : {mae:.2f}")
[55] ✓ 0.1s
... RMSE: 12.26
    MAE   : 8.28
```

In Figure 2, It shows the performance result metrics from an XGBoost model. The model use sales lag, promotional flags and seasonal factors. The RMSE and MAE values help in understand the how demand is changing from the real demand. Even the model is good at handling non-linear patterns the value of error show it is not effective like deep learning techniques when it is time-based sequences.

Figure 3 Performance metrics of an LSTM Model

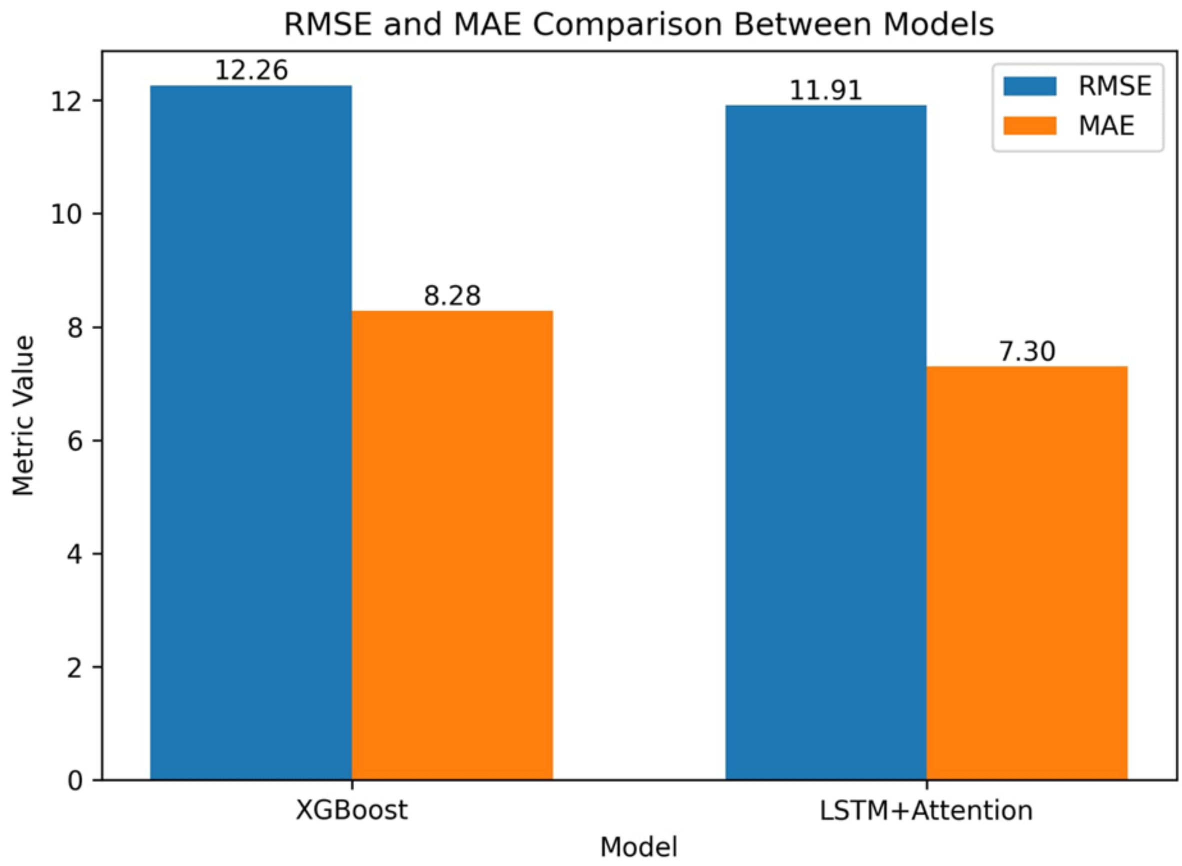
```
[57] ✓ 8m 31.0s                                                                    Python
... Model: "model"
    ________________________________________________________________________________________________
     Layer (type)                Output Shape         Param #     Connected to
    ================================================================================================
     input_1 (InputLayer)        [(None, 28, 11)]     0           []

     lstm (LSTM)                 (None, 28, 64)       19456       ['input_1[0][0]']

     dropout (Dropout)           (None, 28, 64)       0           ['lstm[0][0]']

     multi_head_attention (MultiHea  (None, 28, 64)   16640       ['dropout[0][0]',
     dAttention)                                                   'dropout[0][0]']

     tf.__operators__.add (TFOpLamb  (None, 28, 64)   0           ['dropout[0][0]',
     da)                                                           'multi_head_attention[0][0]']

     layer_normalization (LayerNorm  (None, 28, 64)   128         ['tf.__operators__.add[0][0]']
     alization)

     global_average_pooling1d (Glob  (None, 64)       0           ['layer_normalization[0][0]']
     alAveragePooling1D)

     dense (Dense)               (None, 64)           4160        ['global_average_pooling1d[0][0]'
                                                                  ]

    ...

    ===== FINAL LSTM + ATTENTION PERFORMANCE =====
    RMSE  : 11.91
    MAE   : 7.30
    Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

In Figure 3, It shows the performance metrics of an LSTM Attention model. The LSTM model manages the sales more better when compared with the XGBoost because it understand the long-term patterns.

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

The values of metrics are lower it means the forecasting will be more accurate. This indicate recurrent neural networks are better mainly when handling seasonal and historical trends.

Figure 4 Comparison of RMSE and MAE values between XGBoost and LSTM-Attention Models



In Figure 4, The graph compares the metric results of two models i.e XGBoost and the LSTM. The x axis displays the model's name and y axis display the error metric values. The LSTM Attention model gives the smaller error values i.e RMSE is 11.91 and MAE is 7.30. The lower error values indicate the model forecast more better this shows that recurrent neural networks are better than traditional approaches mainly when handling seasonal and historical trends.

4. Results and Discussion

The performances of XGBoost and LSTM models was evaluated using MAE and RMSE metrics on the test data set. The results that show LSTM model has achieved the lower prediction error compared to the XGBoost model evaluation metrics. This shows that deep learning models are well suited for predicting. LSTM model results are better because it learns long term sequential data and understands the data and how demand changes over time. It can also handle sudden changes in demand during holidays, and promotions and see the connection between one week and the next. On the other hand, XGBoost depends on manually added past value, and it cannot understand time sequences data, so it is difficult to predict when demand changes unexpectedly.

The forecasting plots results shows that LSTM predictions model follows the actual demand curve more closely, and XGBoost predictions model shows higher deviation during peak demand time. So, the deep learning-based LSTM models are more stable for predicting demand forecasting changes in real-world retail environments.

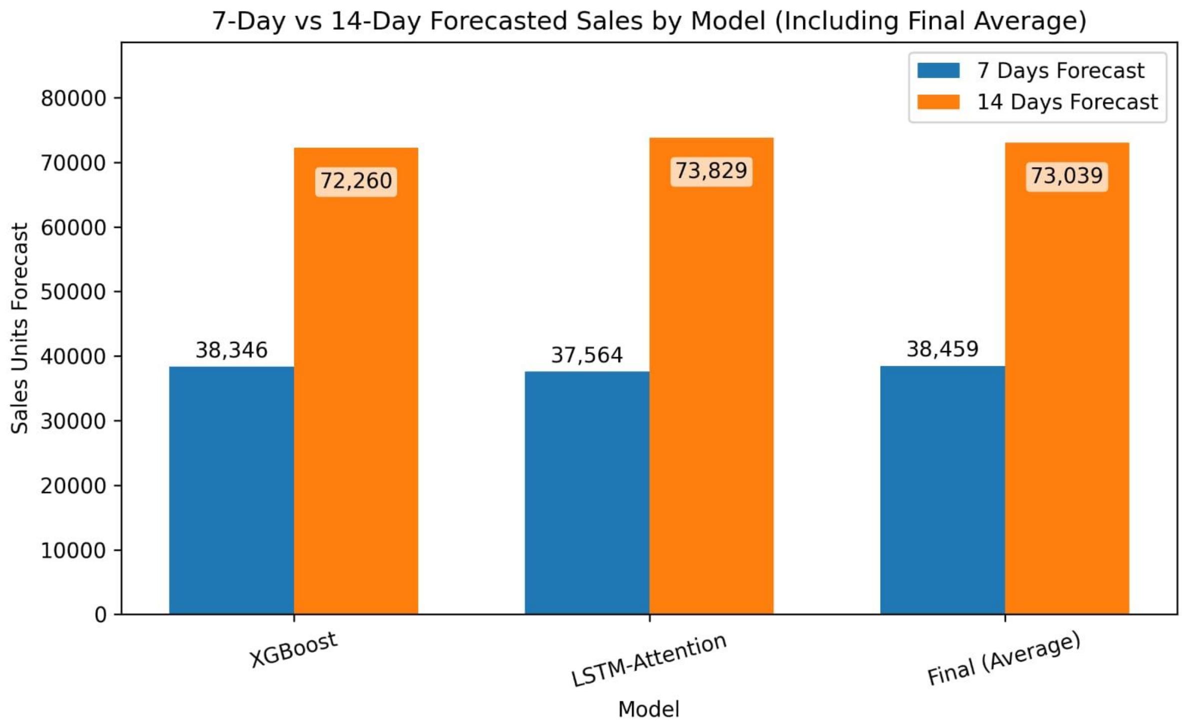
Figure 5 Model forecasting for 7 days

In Figure 5, it compares the 7-day demand forecasts from both XGBoost and the LSTM-Attention model. The LSTM-Attention model predictions follow the actual demand curve more closely, especially when demand goes high and low, and while the XGBoost shows larger errors. This shows that benefit of sequence learning for short-term forecasting.

Figure 6 Model forecasting for 14 days

In Figure 6, it compares the 14-day demand forecasts from both XGBoost and the LSTM-Attention model. The XGBoost model results shows more errors in forecasting over longer time period because it lacks temporal memory, and the LSTM-Attention model are more stable for predicting demand forecasting. The LSTM-attention mechanism helps the model to focus more on important historical data such as promotional week and holidays, and while improving long-term accuracy.

Figure 7 Comparison of 7 day and 14 day forecasted sales units for XGBoost, LSTM Attention and Final averaged prediction



In Figure 7, the graph gives the comparison units from the XGBoost model, LSTM model and final average forecast for both 7 days and 14 days. The x axis displays the models name and Final and y axis display the forecasted sales units. For 7 days forecast XGBoost predicted 38,346 units, LSTM Attention predicted 37,564 units, and final average forecast is 38,459 units which is a middle point between two models. For 14 days forecast combining both predictions gives more stable and reliable forecasts leading to consistent results at different time frames. This shows that deep learning models are good for predicting the demand in grocery retail sector.

5. Conclusion

This paper shows that machine learning and deep learning, how it works for demand forecasting in real-world retail environments. By analyzing real-world sales data and by using different feature engineering, multiple forecasting models were developed and evaluated. It shows that the LSTM model with strong self-attention mechanism, and it provides more accurate data and reliable predictions. The proposed system supports for better inventory planning, and reduces stockouts and wastage of products, and enables data-driven decision-making. Overall, this paper shows that forecasting systems in improving operational efficiency and sustainability.

References

1. Nader Al-Theeb, Hazem Smadi, Naser Al-Qaydeh, "Bayesian-Optimized Ensemble Deep Learning Models for Demand Forecasting in Volatile Situations: A Case Study of Grocery Demand During Covid-19 Outbreaks", Journal of Industrial Engineering and Management, 2025, 18 (1), 193–213.https://doi.org/10.3926/jiem.6571
2. Krishna Kumaar Ragothaman, "Smart Distribution in E-Commerce: Harnessing Machine Learning and Deep Learning Approaches for Improved Logistics", International Journal of Computational and Experimental Science and Engineering, 2025, 11 (1), 1286–1300. https://doi.org/10.22399/ijcesen.1157
3. Md Shariful Alam Bhuiyan, Sheikh Razia Sultana, Sadia Mahjabin Sanchary, Jannatul Maua, Md Nazmul Hossain Mir, Kamruddin Nur, "Hybrid LSTM-CNN with Attention Mechanism for Accurate and Scalable Grocery Sales Forecasting", 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), February 2025, 13–15. https://doi.org/10.1109/ECCE64574.2025.11012951
4. Xianchen Liu, Tianhui Zhang, Xinyu Zhang, Lingmin Hou, Zhen Guo, Yuanhao Tian, Yang Liu, "Integrating Attention-Enhanced LSTM and Particle Swarm Optimization for Dynamic Pricing and Replenishment Strategies in Fresh Food Supermarkets", 2025.
5. Yongbin Yang, Mengdie Wang, Jiyuan Wang, Pan Li, Mengjie Zhou, "Multi-Agent Deep Reinforcement Learning for Integrated Demand Forecasting and Inventory Optimization in Sensor-Enabled Retail Supply Chains", Sensors, 2025, 25 (8), 2428. https://doi.org/10.3390/s25082428
6. Jennifer Jones, "The Role of AI in Enhancing Supply Chain Resilience: Insights from Industry Experts", Preprints.org, April 2025. https://doi.org/10.20944/preprints202504.0741.v1
7. Vikram Pasupuleti, Bharadwaj Thuraka, Chandra Shikhi Kodete, Saiteja Malisetty, "Enhancing Supply Chain Agility and Sustainability Through Machine Learning: Optimization Techniques for Logistics and Inventory Management", Logistics, 2024, 8 (3), 73. https://doi.org/10.3390/logistics8030073
8. Ivan Lizandro Pampa Pampa, Lia Betsabe Rau Rosales, Giancarlo Sánchez Atúncar, "Machine Learning and Deep Learning for Demand Forecasting in Logistics: A Systematic Literature Review", 2024.
9. Sandeep Kumar Panda, Sachi Nandan Mohanty, "Time Series Forecasting and Modelling of Food Demand Supply Chain Based on Regressors Analysis", IEEE Access, 2023. https://doi.org/10.1109/ACCESS.2023.3266275
10. Xianyum Wen, Weibang Li, "Time Series Prediction Based on LSTM-Attention-LSTM Model", IEEE Access, 2023. https://doi.org/10.1109/ACCESS.2023.3276628