



IOT with Blockchain Enabled Machine Learning Model for Sales Prediction using Warehouse Management Data

¹Prasanna Kumar Reddy Gurijala, ²Sohit Reddy Kalluru, ³Pavana Kumari Gavva ⁴Lohith Reddy Kalluru

¹Solution Architect, ²Solution Architect, ³Senior Analyst ⁴ Cloud Developer

Abstract : Sales forecasting permits companies to design their manufacturing outputs that donate to enhance firms' inventory management through cost reduction. On the other hand, not every company has a similar ability to keep all essential data in the long term. So, time series with a trivial distance are general within productions, and difficulties occur owing to the small time series does not completely take sales' behavior. Supply Chain Management (SCM) plays a significant part in organizing and managing business processes, enlarging the organization's operational efficacy. Factors like customer satisfaction, product success, and an organization's growth are based on the effective implementation of SCM. SCM is essential to improve the basis and substructure in societies which in turn upsurges the economic development and living standard of society also. In this article, we present a novel Blockchain Enabled Machine Learning Model for Sales Prediction (BCE-MLMSP) model in Warehouse Management Data. The proposed BCE-MLMSP model exploits BC technology for secured data communication in sales prediction. Initially, the BCE-MLMSP technique takes place data normalization using Z-score normalization is performed. Next, the XGBoost algorithm is designed for the prediction process. Finally, an optimal parameter tuning of the XGBoost method is carried out employing the whale optimization algorithm (WOA). The experimental evaluation of the BCE-MLMSP technique occurs and the outcomes are examined under various aspects. The simulation study inferred the supremacy of the BCE-MLMSP method over current state of art approaches.

IndexTerms - Internet of Things, Machine Learning, Blockchain, Whale Optimization Algorithm, Sales Prediction

I.INTRODUCTION

The Internet of Things (IoT) refers to the connection of tools with limited proficiencies of the Internet. The major aspect of IoT is to simplify the distribution of the resources of the confined devices with other objects [1]. Due to the ubiquity and independence of the IoT environment, devices are more susceptible to attacks. Besides, through such ironic communication, the IoT will achieve a tipping point in that most of the created data on the Internet will originate in billions of devices, which are also limited resources to effectively implement data privacy and complex security strategies [2]. Then, the solution combines dispersed ledger techniques, like blockchain (BC), to IoT devices and the usage of smart bonds to execute operations depending on pre-defined instructions [3]. BCs have attracted important attention in past years due to their new features namely immutability, auditability, security, anonymity, and decentralization. In IoT, BC presents an unchallengeable inspection path of sensor annotations by keeping sensor data as BC connections [4].

Sales prediction has been a very important field to focus on. An optimal and efficient way of predicting has become necessary for all the vendors to withstand the marketing organization's effectiveness [5]. Physical infestation of this task can lead to strong errors resulting in poor administration of the organization, and most significantly should have been time-consuming, which is approximately not desired in this advanced world [6]. The main role of the global economy lies in the business areas that are predictable to generate suitable product numbers for the fulfillment of the complete requirements. It is thus considered significant that the company became capable of succeeding in this objective by using a prediction system [7]. This prediction process includes examining data from different sources like consumer behavior, market trends, and additional factors [8]. These studies will also aid companies in managing their financial resources efficiently. The forecasting system is applied for numerous objectives, including forecasting the upcoming demand for the service or products, and predicting how many of the product should be sold in a certain period. Machine learning (ML) is the domain where technologies increase the capacity to outperform humans in particular tasks [9]. They are utilized to do some specialized tasks logically and achieve superior results for the development of the current society [10].

In this article, we propose a novel Blockchain Enabled Machine Learning Model for Sales Prediction (BCE-MLMSP) model in Warehouse Management Data. The proposed BCE-MLMSP model exploits BC technology for secured data communication in sales prediction. Initially, the BCE-MLMSP technique takes place data normalization using Z-score normalization is performed. Next, the XGBoost algorithm is designed for the prediction process. Finally, an optimal parameter tuning of the XGBoost method is

carried out by utilizing the whale optimization algorithm (WOA). The experimental evaluation of the BCE-MLMSP method occurs and the outcomes are examined under various aspects. The simulation study inferred the supremacy of the BCE-MLMSP method over current state of art approaches.

2. Literature Review

Yadav and Singh [11] concentrates on optimizing the digital purchasing cost for the important selection of suppliers, which retains security, transparency, traceability, and whole data on the blockchain(BC). The probability sampling technique is utilized to create the data for emerging the ML-based method. ML gathers the value stated by the miners for emerging the authenticity variables for the selection of suppliers. Afterward, this real authenticity variable is used to express the MINLP method for digital purchasing issues. Nayak et al. [12] introduce how smart warehousing, energized IIoT, impacts real controlling, predictive analytics, and inventory management in perishable food supply chains. In addition, it attempts to discover the greater significance of IIoT-enabled smart warehouses on industry sustainability. Smart warehousing energized IIoT converts perishable food supply chains, encouraging sustainability through reduced food wastages, improved energy efficacy, and reliable source usage, changing the industry's future. Hu et al. [13] introduce a smart VSC management system, which gives decision support for VSC management in the COVID-19 disease. The method integrates BC, IoT, and ML, which efficiently tackles the 3 problems in VSC. The precision of BC assures entrust between investors. The real vaccine monitoring condition by the IoT assures the quality of the vaccine. ML predicts the demand for vaccines and performs SA on vaccine reviews to assist the company in enhancing the quality of vaccines.

In [14], a BC-based e-commerce analytics method is advanced to improve the digital supply chain. First, early operating decision support will be attained as BC technology, afterwards, the ML system will assist the logistics industry to handle data effectively and to predict active e-commerce demand. Afterwards, the presented method enables LSPs to adaptably reallocate the correct amount of sources in actual time to handle by hourly changing order advents in distribution centers. In addition, the presented method allows logistics experts to predict the trade's performance similar to e-commerce. Shah et al. [15] examine BC technology and ML will be utilized in this research to examine drug protection and confirm early delivery of medicines. The usage of BC technology introduces an SCM method based on a permitted BC network and role-based authorization. By utilizing data from the BC network-based ML methods will be utilized to forecast interruptions in shipments and later assist the supply chain managers in taking precautionary measures and storing additional supply chain expenditures. Sheriff and Aravindhar [16] propose a new method, which integrates BC technology for agri-food traceability with a DBN-based secure BC Management called DBCM for profit optimization. The goal of the Data-Based Decision Making with Hierarchical Ensemble Decision Support (DBCM with HEDS) technique, which was presented, is to reach secure data-driven decisions regarding the warehousing and manufacturing of agri-food products.

3. Proposed Methodology

In this study, we have developed an efficient BCE-MLMSP algorithm in WSN. The presented BCE-MLMSP technique involves BC technology, data normalization, XGBoost algorithm, and WOA-based hyperparameter tuning. Fig. 1 presents the complete process of the BCE-MLMSP method.

3.1. Blockchain Technology

The BCE-MLMSP technique exploits BC technology for secured data communication in sales prediction. BC is generally a dispersed, encrypted database, which can be presented to make it easier to follow operations and handle the resources in a commercial network [17]. Use something like a BC network, for design, completely all valuables might be sold and recorded, justifying business threats and reducing expenses for all applicants involved. It is a decentralized distributed ledger technique, which makes the archives of some digital strength unalterable and observable without the inference of an intermediary facilitator. It is usually a legally trailblazing platform, which creates much excitement attributed to its capacity to decrease corruption and vulnerabilities on a global scale. After effectively integrating into several industries, BC has been presented in various sectors along with the supply chain. For an evolving invention, BC within the supply chain is even a new idea that has sparked huge attention from specialists around the world. It can convert the data, which is slow in the supply chain function into a great tactical resource that could grant companies main benefits opposing their rivals. Using BC, companies can collect and distribute valued data, which can surely assist in predicting demand and using logistics resources for the success of organizational goals and strategic operations.

During the BC network, supply chain transactions are either started by authorized supply chain units or dynamically forced by smart agreements responding to pre-defined conditions. To reduce potential security problems, supply networks normally use accepted BCs that permit related data about supply chain transactions and events to be validated only by trustworthy and authorized supply chain members with certain qualifications. Particular user groups like end-users, are only authorized to access data on the BC network. After being anonymized and evaluated with a hash algorithm, BC data can be saved in time-recorded blocks, which are connected in sequential order. The blocks are applied to work together for the creation of a shared record that might be used such that the supply chain technology is traced such that the inventory provenance and asset could be verified. Based on BC networks that offer a unified.

3.2. Data Normalization

At first, the BCE-MLMSP method uses data normalization utilizing Z-score normalization. Z-score normalization, also recognized as standardization, converts sales prediction data by changing every data point into the number of standard deviations absent from the mean [18]. This model confirms that the information has a standard deviation of one and a mean of zero, making it very simple to analyze and compare the data. By using z-score normalization, variations and outliers in sales predictions are diminished, foremost to more consistent and accurate outcomes. This method is most convenient in refining the performance of ML methods that are employed for sales forecasting.

3.3. Prediction using XGBoost

Next, the XGBoost algorithm is designed for the prediction process. This algorithm enhances regularization, and loss function, and extensively enhances accuracy and efficiency [19]. The base learners are decision tree-based techniques.

For an assumed dataset $T = \{(x_i, y_i)\}$, the model's loss function is $l(y_i, \hat{y}_i)$, the regularization term $\Omega(f_k)$ and the predicted value \hat{y}_i is $\sum_{k=1}^K f_k(x_i)$. The objective function (Obj) calculation is given below:

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^m \Omega(f_k) \quad (1)$$

Here, n and m denote the number of samples and trees, respectively. Where $\Omega(f) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2$

The Eq. (2) experiences a first and second-order Taylor expansion.

$$Obj^{(t)} \approx \sum_{i=1}^n \left[l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] x_i + \Omega(f_t) \quad (2)$$

Whereas, t signifies the tth tree, g_i and h_i represents the first-and second order gradient number, respectively.

After eliminating the term constant, Eq. (3) was changed as below:

$$Obj^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] x_i + \Omega(f_t) \quad (3)$$

Express the data of jth leaf node as a group of nodes:

$$I_j = \{q(x_i) = j\} \quad (4)$$

Edit Eq. (5) and unite the first and secondary term co-efficient:

$$Obj^{(t)} = \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right] + \gamma T \quad (5)$$

Eq. (6) is changed into a calculation in terms of the weight vector ω_j . The optimal solution Obj_b and the equivalent weight vector ω^* were computed by Eq. (7), and an outcome is formulated below:

$$Obj_b = \frac{1}{2} \sum_{j=1}^T \frac{\sum_{i \in I_j} (g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \quad (6)$$

$$\omega^* = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (7)$$

3.4. WOA-based Parameter Tuning

Finally, an optimal parameter tuning of the XGBoost technique is executed by utilizing the WOA. This technique introduced by Mirjalili in 2016, is a novel method stimulated by the predatory performance of humpback whales [20]. The WOA initially comprises 3 phases namely bubble-net attacking, surrounding the prey, and seeking for the victim arbitrarily. The complete procedure is described below: Fig. 3 represents the flowchart of WOA.

Neighboring the prey: Primarily, the precise location of the target is unidentified. The WOA describes the location of the best reasonable individual in the present whale group as the assumed prey position. Afterward, the residual members of the whale cluster fine-tune their locations based on the recognized place of the optimum candidate individual. This upgrading model is expressed below:

$$\{D = |C \cdot X^*(j) - X(j)|\} X(j+1) = X^*(j) - A \cdot D \quad (8)$$

Whereas, X represents the vector location, X^* stands for the vector location of optimum outcome attained, A and C signify the coefficient vectors, || indicates the absolute rate, refers to the element wise multiplication, and j defines the present iteration.

It is essential to demonstrate that once a higher solution is exposed, the variable X can be upgraded in all the iterations. This formulation for A and C is written as:

$$\{A = 2ar - a \quad C = 2r\} \quad (9)$$

whereas a is linearly reduced from two to zero under the iteration courses, r illustrates the random vector from zero and one.

Bubble-net attack: In this phase, the computation of the prey's locations and the distance in the whale has been accompanied for the primary time. Employing these data, the whale defines its most probable next movement.

$$\{D' = |X^*(j) - X(j)| \quad X(j+1) = D' \cdot e^{bl_1} \cdot \cos(2\pi \cdot l_1) + X^* \cdot (j) \quad (10)$$

In which, b signifies the continuous that determines the logarithmic spiral shape and l_1 determines the randomly generated number among $[-1, 1]$.

Exploration for prey. In practical situations, members of whale groups can be involved in random searches for prey locations comparable to their places, boosting the method's capability for global search. This performance is determined by leading the searching agent to differ from the reference whale once the state of $|A| > 1$ is met.

$$\{D = |C \cdot X_{rand} - X| \quad X(j+1) = X_{rand} - A \cdot D \quad (11)$$

Whereas, X_{rand} demonstrates the random position vector elected in the existing population.

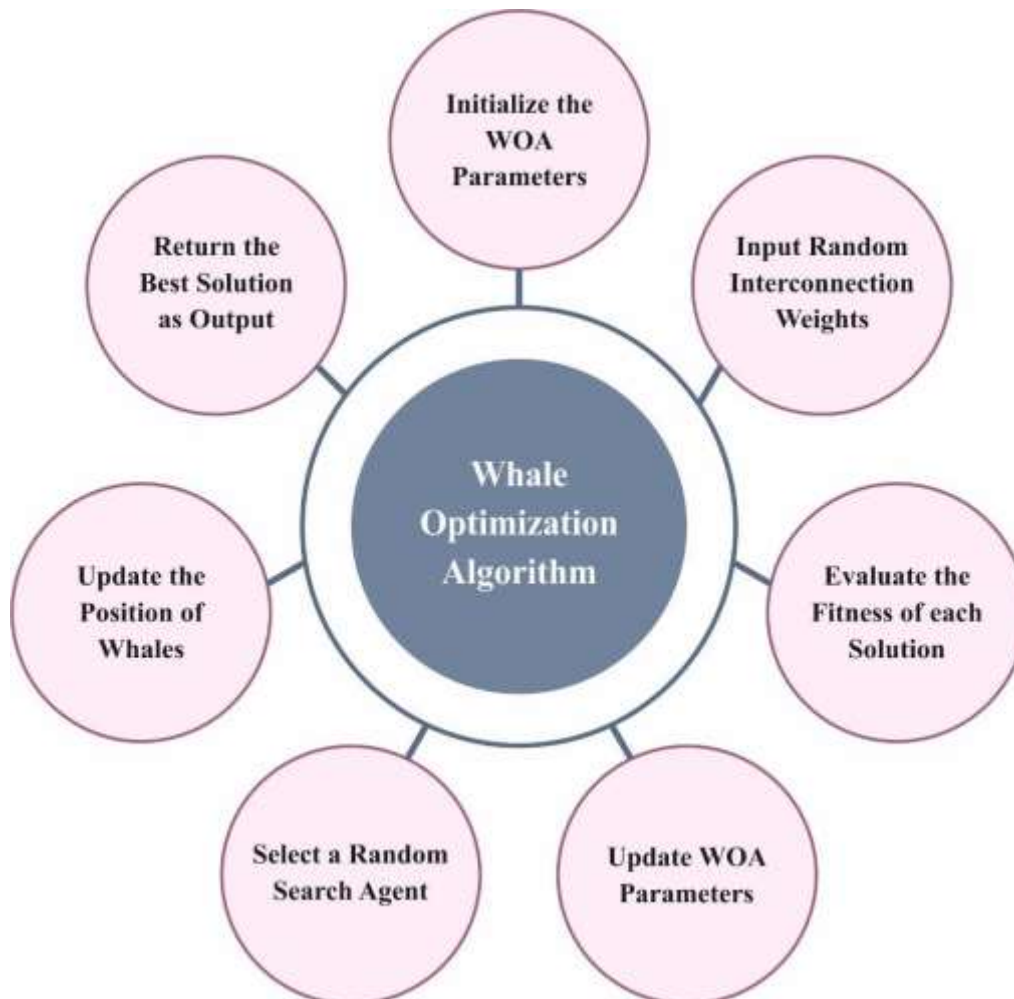


Fig. 3. Flowchart of WOA

Step1: Significance the data and accomplish normalization.

Step2: Create the structure of ENN. Determine variables for output and input layers.

Step3: Initialization of the thresholds and weights of NN.

Step4: Present the WOA to enhance the network. Always iterate utilizing the WOA to fine-tune the thresholds and ENN weights. Repeating this procedure still, the thresholds and weights come across the identified tolerance range. End the loop moment and continue to the following step.

Step5: Upgrade the thresholds and weights of networks depending on the values given by the method.

Step6: Output the training set and predictive rates.

In this study, the WOA is used to identify the hyperparameter convoluted in the XGBoost technique. The MSE examines the objective function and is determined as follows.

$$MSE = \frac{1}{T} \sum_{j=1}^L \sum_{i=1}^M (y_j^i - d_j^i)^2 \quad (12)$$

Where M and L signify the resultant value of layer and data respectively, y_j^i and d_j^i indicates the attained and appropriate magnitudes for unit j from the resultant layer of a network in time t respectively.

4. Performance Analysis

This article studies the performance investigation of the BCE-MLMSP model using benchmark dataset [21]. It can be selected eight features such as product ID, Product Name, Beginning Inventory, Inventory Received, Inventory Sold, Ending Inventory, Predicted Sales, and Difference.

Fig. 4 illustrates the correlation matrix made by the BCE-MLMSP method. The results identify that the BCE-MLMSP technique has an efficient prediction of all class labels accurately.

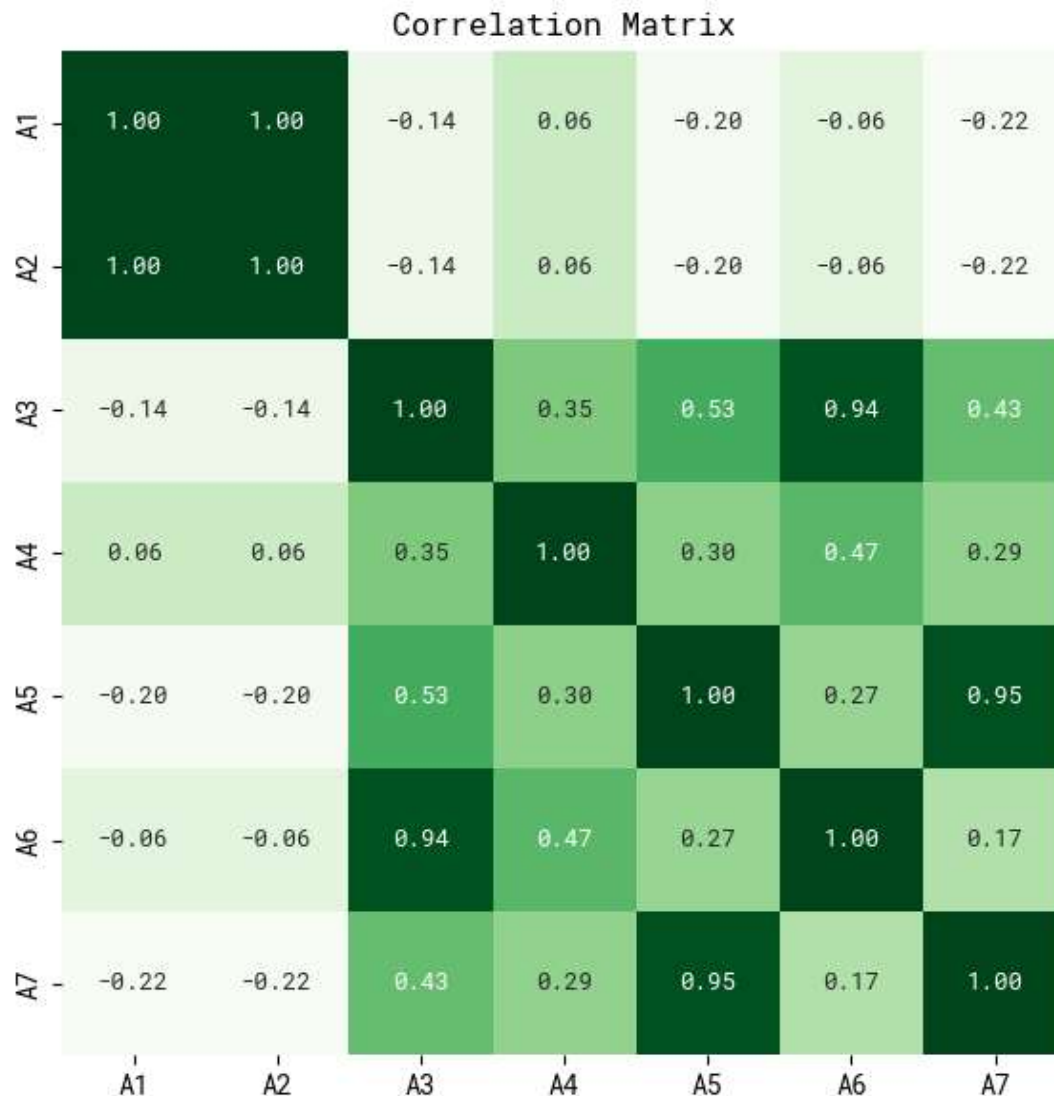


Fig. 4. Correlation Matrix of BCE-MLMSP technique

Fig. 5 distributes a warehouse outcomes study on the actual vs predicted BCE-MLMSP model under different epoch counts. Moreover, it is well-known that the variance between the actual and predicted values can be computed at the least.

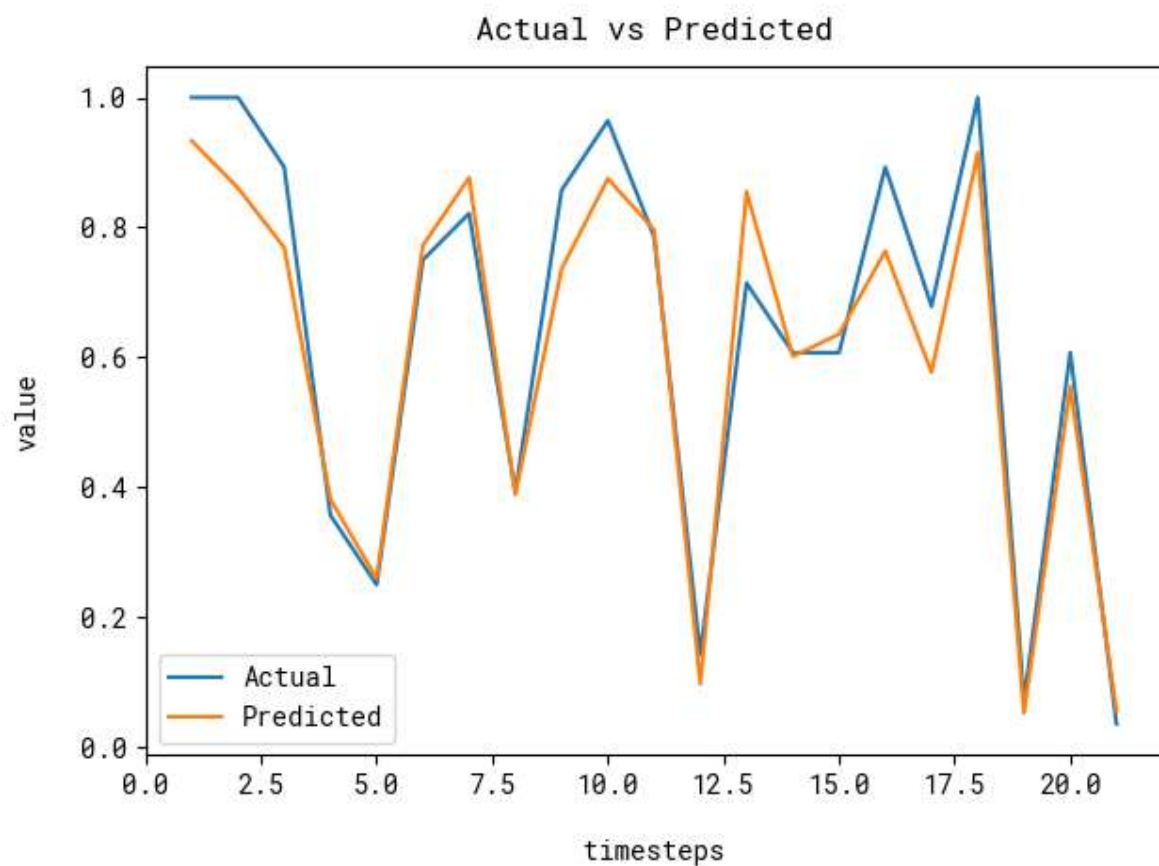


Fig. 5. Warehouse Results Analysis on Actual vs Predicted

Table 1 and Fig. 6 represent the classifier outcomes of the BCE-MLMSP technique with distinct metrics. The results inferred that the BCE-MLMSP technique obtained an MSE of 0.006. Likewise, the BCE-MLMSP technique has obtained an MAE of 0.062. Eventually, the BCE-MLMSP technique obtained a MAPE of 0.134.

Table 1 Classifier outcomes of BCE-MLMSP technique with distinct metrics

Metrics	Values
MSE	0.006
MAE	0.062
MAPE	0.134

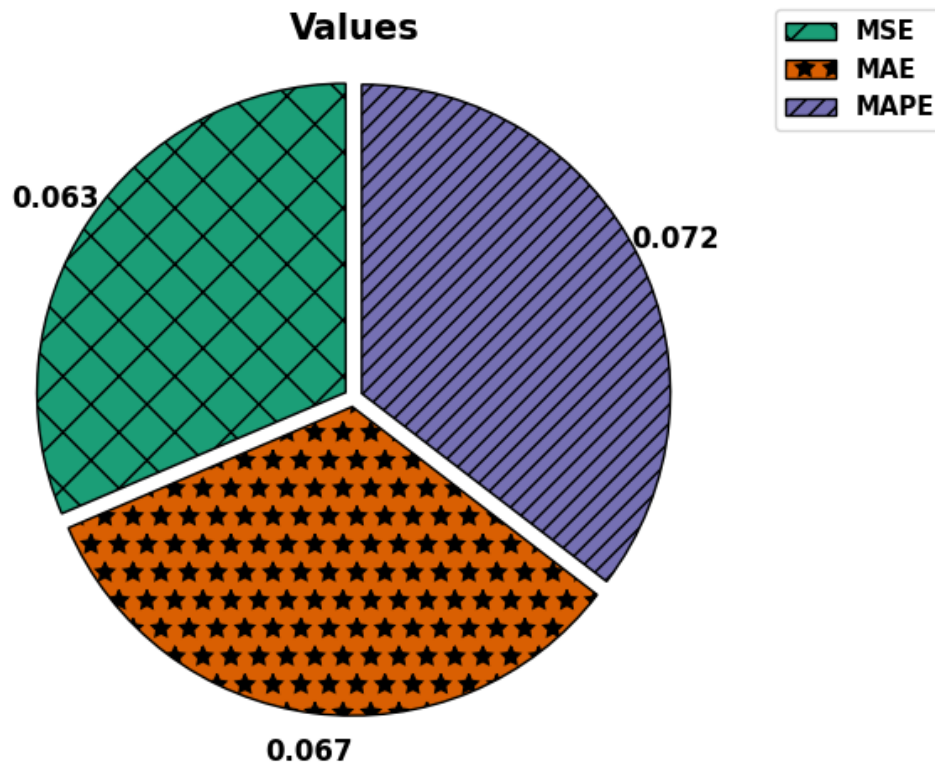


Fig. 6. Average of BCE-MLMSP technique with distinct metrics

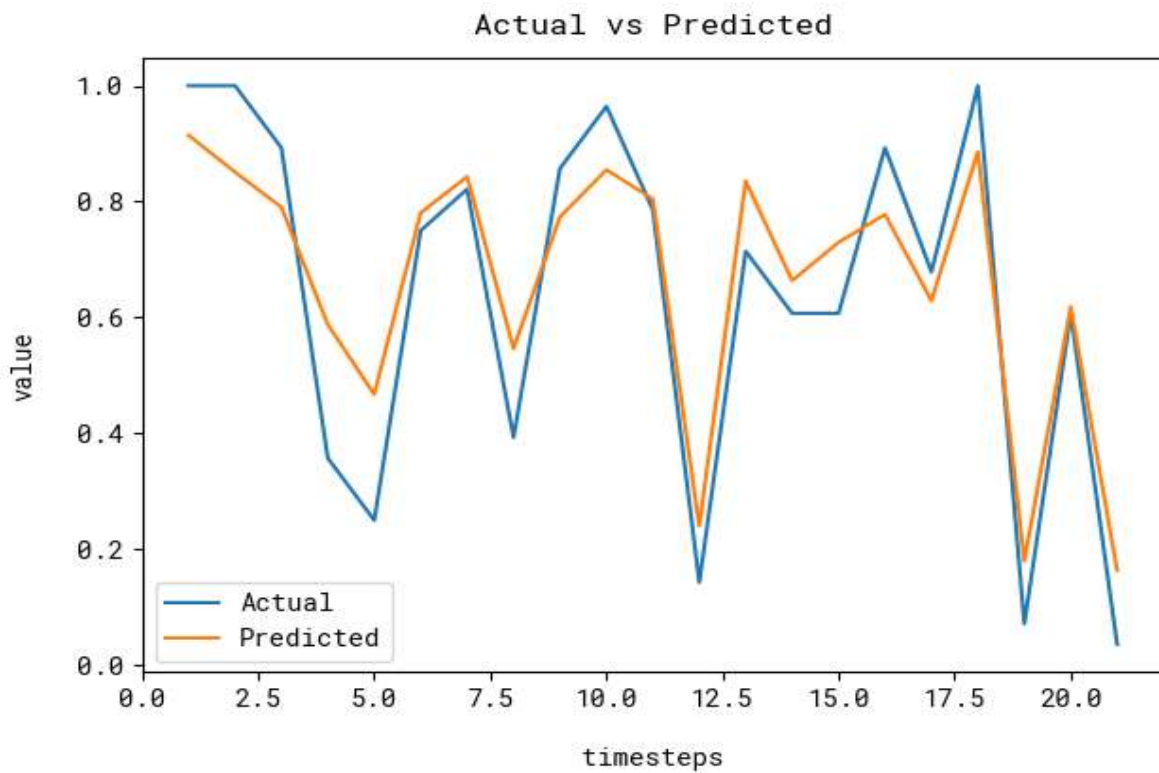


Fig. 7. Warehouse Results Analysis on Actual vs Predicted Epoch 25 of BCE-MLMSP technique

Fig. 7 displays the predicted warehouse result for the actual vs predicted BCE-MLMSP approach on the epoch count 25. The figure illustrates that the BCE-MLMSP methodology correctly predicted the warehouse result. It is also observed that the predicted values by the BCE-MLMSP model are adjacent to the actual values.

Fig. 8 illustrates the predicted warehouse result for the actual vs predicted BCE-MLMSP method on the epoch count 50. The figure points out that the BCE-MLMSP model accurately predicted the warehouse result. It is also noted that the predicted values by the BCE-MLMSP method are nearer to the actual values.

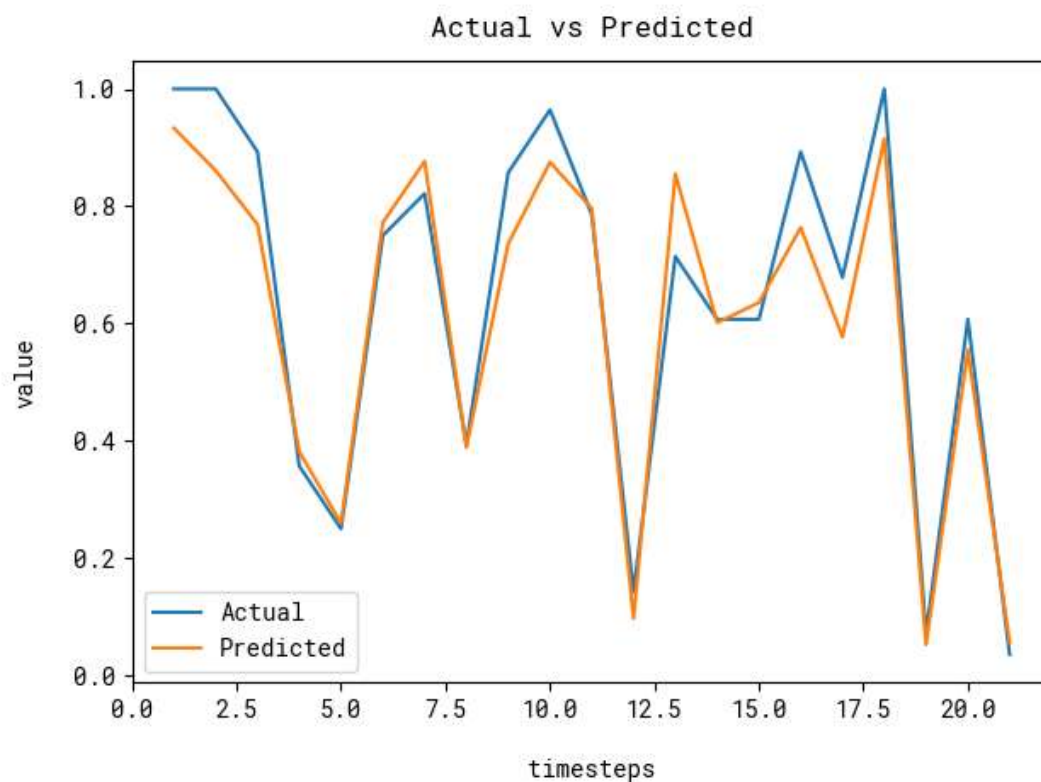


Fig. 8. Warehouse Results Analysis on Actual vs Predicted Epoch 50 of BCE-MLMSP technique

To demonstrate the proficiency of the BCE-MLMSP method, a detailed comparison study is made in Table 2 and Fig. 9. The experimental values inferred that the BCE-MLMSP approach has better performances in terms of MSE. The RsNet-50 model has exhibited poor performance with a higher value MSE of 0.026. At the same time, the ADAM model has attained a slightly lesser MSE of 0.021. Meanwhile, the DT and linear regression models have demonstrated closer values MSE of 0.016 and 0.011, respectively. Nevertheless, the BCE-MLMSP technique results in improved performance with a lower value of MSE of 0.006. Therefore, the BCE-MLMSP technique can be applied for enhanced sales prediction.

Table 2 MSE outcome of BCE-MLMSP technique with existing approaches

Method	MSE
BCE-MLMSP	0.006
Linear Regression	0.011
Decision Tree	0.016
ADAM Algorithm	0.021
RsNet-50 Model	0.026

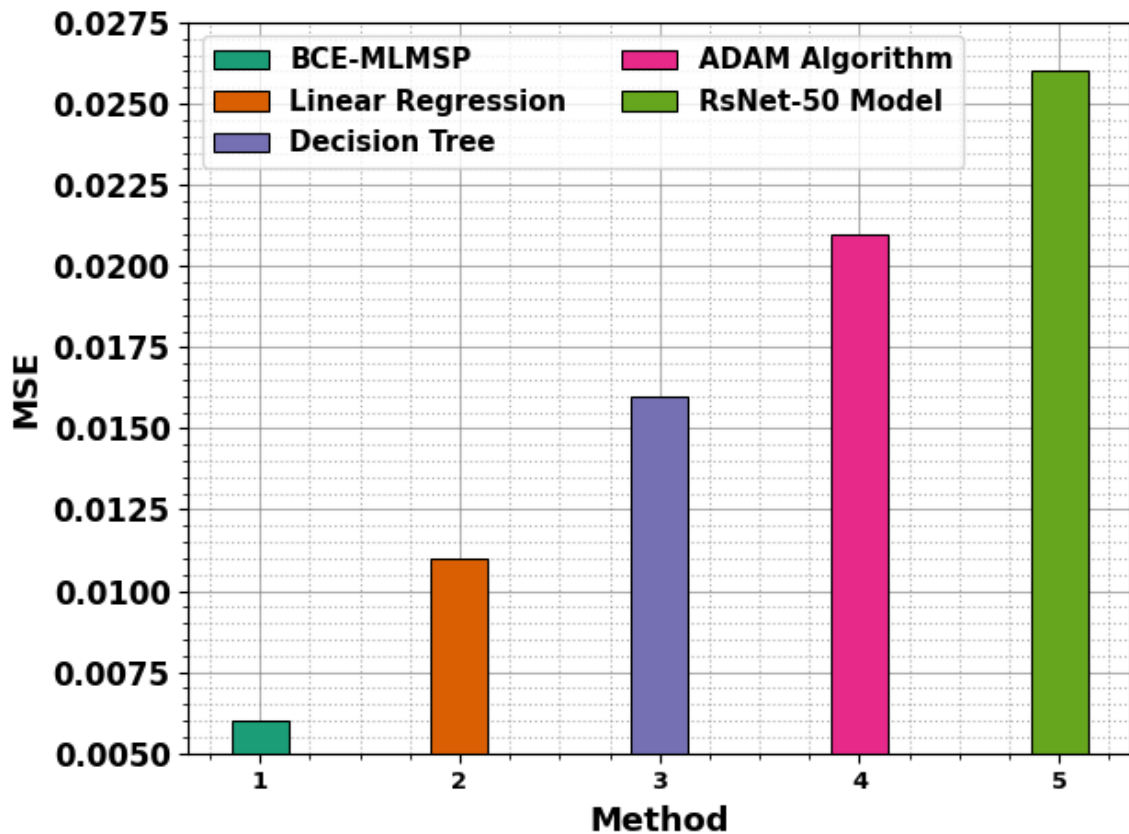


Fig. 9. MSE outcome of BCE-MLMSP technique with existing approaches

5. Conclusion

In this study, we present a new BCE-MLMSP model in Warehouse Management Data. The proposed BCE-MLMSP model used BC technology for secured data communication in the sales prediction. At first, the BCE-MLMSP technique takes place data normalization using Z-score normalization is performed. Next, the XGBoost method is designed for the prediction process. Lastly, an optimal parameter tuning of the XGBoost technique is carried out using the WOA. The experimental evaluation of the BCE-MLMSP method takes place and the results are examined under various aspects. The simulation study inferred the supremacy of the BCE-MLMSP method over current state of art methods.

References

- [1] Tsoumakas, G., 2019. A survey of machine learning techniques for food sales prediction. *Artificial Intelligence Review*, 52(1), pp.441-447.
- [2] Ensafi, Y., Amin, S.H., Zhang, G. and Shah, B., 2022. Time-series forecasting of seasonal items sales using machine learning—A comparative analysis. *International Journal of Information Management Data Insights*, 2(1), p.100058.
- [3] Kohli, S., Godwin, G.T. and Urolagin, S., 2020. Sales prediction using linear and KNN regression. In *Advances in Machine Learning and Computational Intelligence: Proceedings of ICMLCI 2019* (pp. 321-329). Singapore: Springer Singapore.
- [4] Pavlyshenko, B.M., 2019. Machine-learning models for sales time series forecasting. *Data*, 4(1), p.15.
- [5] Kilimci, Z.H., Akyuz, A.O., Uysal, M., Akyokus, S., Uysal, M.O., Atak Bulbul, B. and Ekemis, M.A., 2019. An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain. *Complexity*, 2019(1), p.9067367.
- [6] Weng, T., Liu, W. and Xiao, J., 2020. Supply chain sales forecasting based on lightGBM and LSTM combination model. *Industrial Management & Data Systems*, 120(2), pp.265-279.
- [7] Ji, S., Wang, X., Zhao, W. and Guo, D., 2019. An Application of a Three-Stage XGBoost-Based Model to Sales Forecasting of a Cross-Border E-Commerce Enterprise. *Mathematical Problems in Engineering*, 2019(1), p.8503252.
- [8] Bandara, K., Shi, P., Bergmeir, C., Hewamalage, H., Tran, Q. and Seaman, B., 2019. Sales demand forecast in e-commerce using a long short-term memory neural network methodology. In *Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part III 26* (pp. 462-474). Springer International Publishing.
- [9] Loureiro, A.L., Miguéis, V.L. and Da Silva, L.F., 2018. Exploring the use of deep neural networks for sales forecasting in fashion retail. *Decision Support Systems*, 114, pp.81-93.
- [10] Boone, T., Ganeshan, R., Jain, A. and Sanders, N.R., 2019. Forecasting sales in the supply chain: Consumer analytics in the big data era. *International journal of forecasting*, 35(1), pp.170-180.
- [11] Yadav, S. and Singh, S.P., 2024. Machine learning-based mathematical model for drugs and equipment resilient supply chain using blockchain. *Annals of Operations Research*, pp.1-75.
- [12] Nayak, A., Patnaik, A., Satpathy, I., Tripathy, S. and Patnaik, B.C.M., Role of Industrial Internet of Things (IIoT) Enabled Smart Warehousing for Sustainable Perishable Food Supply Chains. In *Digital Technology Enabled Circular Economy* (pp. 127-144). CRC Press.

- [13] Hu, H., Xu, J., Liu, M. and Lim, M.K., 2023. Vaccine supply chain management: An intelligent system utilizing blockchain, IoT and machine learning. *Journal of business research*, 156, p.113480.
- [14] Ho, G.T., Tang, Y.M., Lam, H.Y. and Tang, V., 2023, February. A blockchain-based decision support system for E-commerce order prediction. In *2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)* (pp. 041-045). IEEE.
- [15] Shah, Y., Verma, Y., Sharma, U., Sampat, A. and Kulkarni, V., 2023, January. Supply chain for safe & timely distribution of medicines using blockchain & machine learning. In *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1123-1129). IEEE.
- [16] Sheriff, I.M.M. and Aravindhar, D.J., 2024, January. Integrated Blockchain-Based Agri-Food Traceability and Deep Learning for Profit-Optimized Supply Chain Management in Agri-Food Supply Chains. In *2024 Fourth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)* (pp. 1-6). IEEE.
- [17] Hasan, N., Ferdouse, N., Khan, T., Adiba, A.F. and Shams, K.A.I., 2022. Implementing Blockchain in the supply chain of FMCG companies for better demand forecasting and logistics operation. *Supply Chain Insider*| ISSN: 2617-7420 (Print), 2617-7420 (Online), 8(1).
- [18] Henderi, H., Wahyuningsih, T. and Rahwanto, E., 2021. Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer. *International Journal of Informatics and Information Systems*, 4(1), pp.13-20.
- [19] Liu, M., Feng, Y., Yang, S. and Su, H., 2024. Dam Deformation Prediction Considering the Seasonal Fluctuations Using Ensemble Learning Algorithm. *Buildings*, 14(7), p.2163.
- [20] Xu, E. and Zhao, R., 2024, May. Research on Displacement Sensorless Control for Bearingless Synchronous Reluctance Motor Based on the Whale Optimization Algorithm–Elman Neural Network. In *Actuators* (Vol. 13, No. 5, p. 192). MDPI
- [21] <https://www.kaggle.com/datasets/lastman0800/warehouse-management-using-ai>

