E-COMMERCE DEMAND AND SUPPLY FORECASTING BY LONG SHORT-TERM MEMORY WITH RECUREENT NEURAL NETWORK

Monika Saini[1], Vinti Dhaka[2]

Department of Mathematics

Om Sterling Global University, Hisar,

Haryana, India

Abstract: Supply Chains are the network of facilities that not only includes retailers, distributors, transporters, manufacturers but also the customers. Therefore, it is vital to understand the important consumption and wishes of the purchasers as they are the prime nodal of each supply chain as they push various entities to supply and distribute. The availability chain facilities have now learned the importance of collaboration and coordination to fulfil the real demand. The entities also work cohesively to lower down the total cost of the supply chain. However, in the absence of such collaborations; a mismatch between the important and ideal world of supply chain networks occurs. To overcome these gaps this paper consists of application of machine learning techniques in supply chain management. It consists of cases of supply chain management such as demand forecasting, supply forecasting, text analytics, price panning and more to enhance their processes, reduce costs and risk, and increase revenue. It gives us a précis about all the important aspects of economy and how to understand and use them wisely.

Keywords: Supply Chain Network, Supply Management, Machine learning

1. INTRODUCTION

With the continuous development of the economic society, more and more enterprises gradually convert from the original business to cooperative pattern, which will greatly change the structure of modern industrial organization [6,23,34]. The traditional supply chain is based on the manufacturers, sellers, and consumers, each of which is an independent system and builds a relationship of interest through the circulation of products [9,17,21]. However, more and more cooperation has emerged in this link, such as commodity supply before payment and commodity selling before cost paying off, which is based on mutual trust. In addition, the exchanges and cooperation between various enterprises become more frequent with the continuous improvement of Internet information technology means. In the face of the impact from the Internet industry, disadvantages of the traditional supply chain system become more obvious, such as untimely communication of information and the uncoordinated supply and demand. As a result, the overall efficiency of the enterprise cannot be improved, because each link is restricted by the upper and lower companies, which is the main contradiction for modern enterprises [9]. With development of machine and deep learning technologies, the supply chain network is optimized continuously, more and more enterprises become the links in the supply chain. In short, it is of great practical significance to establish a supply chain network model satisfying the actual situation for enterprises to assume the smallest risks and obtain the greatest benefits in a market-balanced state, and realize the sustainable development. Most of the researches on supply chain network focus on simplifying practical problems and strengthening the connection between data and financial information among various enterprises [2,5,10,30]. Among them, Quddu et al. [33] revealed the impact of municipal solid waste utilization on the performance of the biofuel supply chain network using the combined sample average approximation algorithm. Mishra et al. [27] solved the design problems of the closed-loop supply chain using the genetic algorithm (GA) and optimization algorithms, and verified the feasibility of the model and the applicability of the developed solution method. Fu [14] constructed the enterprise supply chain model through the Cplex enhancing constraint method and social impact coefficient and verified it in wine companies, and the results proved that this method is effective and feasible. Madani et al. [28] developed a new supply chain model with the hybrid heuristic algorithm, and proved its application in practical applicability. Therefore, enterprises face how to use efficient algorithms to meet the stochastic needs of consumers for commodities, which will change with the change of commodity price, logistics, quality, and time [25]. Therefore, whether an enterprise can change and make corresponding decisions according to demand is directly related to the sustainable development of the enterprise under market competition. Using various algorithms and technologies to construct a decision model has become a hot research topic in this field.

1.1 Classic and Modern Supply chain Network

Design of the supply chain network is the process of constructing a supply chain model so that it can utilize the available resources, time, and geographic location better, thus bringing commodities to the market quickly. Based on correct and effective management, the overall function of the supply chain is greater than the sum of the parts of each chain, and the goal of profit maximization of the enterprises can be achieved. Fig. 1 shows the differences and connections between the traditional and the modern supply chain network. It illustrates that the traditional supply chain network is based on the manufacture – retailer – demand market, including all specific links of plans, raw materials procurement, manufacturing, transportation, and sales, so it is a line of service network with the purpose of guaranteeing the balance between supply and demand. The modern supply chain networks make full use of the advantages of network nerves to ensure that all links are interchangeable, interconnected, and affected mutually. Such supply chain logistics has not yet been applied, so it requires more technical and algorithmic supports. The original supply chain management model is improved and optimized based on the modern model in this study, involving three aspects. One is the selection of the supply chain network predication model with the purpose of improving the accuracy of commodity demand predication and reducing related risks. The other is the optimization of supply chain network algorithms with the main purpose optimizing the entire structure to improve operating efficiency. Another is the security of the supply chain network with main purpose of strengthening cooperation between various nodes to improve the efficiency. The models proposed later in this study are based on this, construction of modern supply chain logistics network is achieved through various algorithm and predication methods.

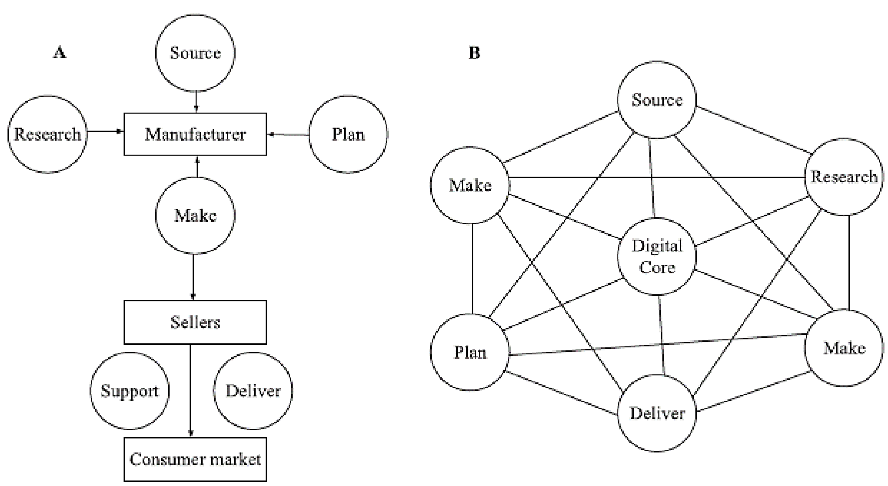

Fig.1: Conversion of supply chain network of the industrial organization [16]

1.2 Challenges in Supply Chain Management

The supply chain frequently changes. Hence, there evolves a requirement for meeting the new demands within maintaining a smooth sailing flow. Overall, supply chain management can potentially face several challenges such as [29]:

- Fluctuation in demand
- Inadequate inventory planning

- Backlogs of orders
- Uncertainties in logistics
- Communication gaps within the availability chain
- Shortages in supply

1.3 GAP IN PREVIOUS WORK

- Fluctuation in demand.
- Inadequate inventory planning.
- Backlogs of orders.
- Uncertainties in logistics.
- Communication gaps within the supply chain.
- Shortages in supply.

2. LITERATURE REVIEW

With the advancement of e-business and therefore the existence of technology to become more agile and dynamic; the firms are restrictive to have long-term relationships. Another reason could be not knowing the real demand of the purchasers and producing more in anticipation of the demand. The normal way of knowing and forecasting the demand of their own entity is now overtaken by new techniques of Machine Learning Approaches.

Kosasih, et.al [23] pose the supply chain visibility problem as a link prediction problem from the field of Machine Learning (ML) and propose the use of an automated method to detect potential links that are unknown to the buyer with Graph Neural Networks (GNN). Using a real automotive network as a test case, the researchers show that our method performs better than existing algorithms. Additionally, we use Integrated Gradient to improve the explainability of our approach by highlighting input features that influence GNN's decisions. Tugay, et.al [37] propose a new approach for demand prediction on an e-commerce web site. The proposed model differs from earlier models in several ways. The business model used in the e-commerce web site, for which the model is implemented, includes many sellers that sell the same product at the same time at different prices where the company operates a market place model. The demand prediction for such a model should consider the price of the same product sold by competing sellers along the features of these sellers. In this study we first applied different regression algorithms for specific set of products of one department of a company that is one of the most popular online e-commerce companies in Turkey. Then the researchers used stacked generalization or also known as stacking ensemble learning to predict demand. Finally, all the approaches are evaluated on a real-world data set obtained from the e-commerce company. The experimental results show that some of the machine learning methods do produce almost as good results as the stacked generalization method. Gao, et.al [16] propose a model combining with the network neural commodity demand predication method and the particle swarm optimization (PSO) algorithm to comprehensively evaluate the predication effect and algorithm performance by using the supply chain data of the enterprises, coming up with an optimal model. Results of the study show that: on national warehouses and regional warehouses, the difference between the predicted value and the actual value of autoregressive integrated (AR) mixture density networks (MDN) (AR-MDN) is 15%, the average outlier is between 450 and 150, the score of roots mean square error (RMSE) and mean absolute percentage error (MAPE) is 117.342 and 2.334, respectively. Feizabadi, et.al [12] developed a hybrid demand forecasting methods grounded on machine learning i.e., ARIMAX and Neural Network. Both time series and explanatory factors are feed into the developed method. The method was applied and evaluated in the context of functional product and a steel manufacturer. The statistically significant supply chain performance improvement differences were found across traditional and ML-based demand forecasting methods. The implications for the theory and practice are also presented. Aamer, et.al [1] focused on comprehensively overviewing machine learning applications in demand forecasting and underlying its potential role in improving the supply chain efficiency. A total of 1870 papers were retrieved from Scopus and Web of Science databases based on our string query related to machine learning. A reduced total of 79 papers focusing on demand forecasting were comprehensively reviewed and used for the analysis in this study. The result showed that neural networks, artificial neural networks, support vector regression, and support vector machine were among the most widely used algorithms in demand forecasting with 27%, 22%, 18%, and 10%, respectively. This accounted for 77% of the total reviewed articles. Most of the machine learning application (65%) was applied in the industry sector, and a limited number of articles (5%) discussed the agriculture sector. Baryannis, et.al [6] first propose a supply chain risk prediction framework using data-driven AI techniques and relying on the synergy between AI and supply chain experts. The researchers then explore the trade-off between prediction performance and interpretability by implementing and applying the framework on the case of predicting delivery delays in a real-world multi-tier manufacturing supply chain. Experiment results show that prioritizing interpretability over performance may require a level of compromise, especially with regard to average precision scores. Ni, et.al [2019] carried out to present the latest research trends in the discipline by analyzing the publications between 1998/01/01 and 2018/12/31 in five major databases. The quantitative analysis of 123 shortlisted articles showed that ML applications in SCM were still in a developmental stage since there were not enough high-yielding authors to form a strong group force in the research of ML applications in SCM and their publications were still at a low level; even though 10 ML algorithms were found to be frequently used in SCM, the use of these algorithms were unevenly distributed across the SCM activities most frequently reported in the articles of the literature. The aim of this study is to provide a comprehensive view of ML applications in SCM, working as a reference for future research directions for SCM researchers and application insight for SCM practitioners. Bousqaoui, et.al [5] examines multiple Machine Learning algorithms, explores their applications in the various supply chain processes, and presents a long short-term memory model for predicting the daily demand in a Moroccan supermarket. machine Learning or the ability of a machine to learn automatically has found applications in various fields. It has proven to be a valuable tool for aiding decision makers and improving the productivity of enterprise processes, due to its ability to learn and find interesting patterns in the data. Thereby, it is possible to improve supply chains processes by using Machine Learning which generates in general better forecasts than the traditional approaches.

Table 1. Literature Analysis on machine learning methods

Ref	Year	Aim	Techniques	Dataset	Gaps	Findings
[23]	[2022]	Machine learning approach for predicting hidden links in supply chain with graph neural networks	Machine Learning (ML) and Graph Neural Networks (GNN)	Real automotive dataset	Supply chain visibility gaps	Researchers have highlighted the importance of gaining visibility into procurement interdependencies between suppliers to develop more informed business contingency plans.
[37]	[2022]	Demand prediction using machine learning methods	ML and stacking ensemble learning	Real world data set obtained from the e-commerce company	As the data is not statistically significant between the proposed model and random forest, the proposed method can be used to forecast demand due to its accuracy with less data.	Experiments have shown that our approach predicts demand at least as good as single classifiers do, even better using much less training data (only %20 of the dataset).
[16]	[2022]	The analysis of commodity demand predication in supply chain network based on particle swarm optimization algorithm	Models: ARIMA and MLP-LSTM models Algorithms: PSO, ABC algorithm.	—	The data prediction in this study is homogeneous without consideration of special circumstances such as holidays; second, there is still a lack of verification and improvement in	The supply chain network model constructed in this study can provide enterprises with a good commodity demand predication method and improve the ability to respond to risks in the supply chain

					different regions and different time periods for the models	
[1]	[2020]	Data analytics in the supply chain management	ML methods	Digital databases	Most of the machine learning applications were found in the industry sector, and limited machine learning applications were found in the agriculture sector. This calls for more research needed in the agricultural area to improve data analytics' efficiency.	Based on the analysis, we concluded that machine learning algorithms could provide better accuracy and less computational cost for demand forecasting than traditional forecasting models.
[5]	[2019]	Machine learning applications in supply chains: Long short-term memory for demand forecasting	LSTM an ANFIS Model	Keras and TensorFlow Libraries	—	ML algorithms explores their applications in the various supply chain processes, and presents a long short-term memory model for predicting the daily demand in a Moroccan supermarket.
[6]	[2019]	Predicting supply chain risks using machine learning	SVM and decision tree models	SCRM within a real-world multi-tier aerospace manufacturing supply chain, with partners in Europe and Asia	—	Explore the case of SCRM within a real-world multi-tier aerospace manufacturing supply chain, with partners in Europe and Asia.

3. METHODLOGY

The selection of the LSTM network for forecasting in retail is based on several reasons. Some of them are as follows. LSTM networks recently have shown promising results in time-series forecasting tasks [13]. LSTM networks are capable of working well on linear and non-linear time-series [8]. Therefore, the decomposition of time-series into linear and non-linear components is not required. The data contains several demand patterns generated from online and offline sales. Thus, it could be conjectured that LSTM networks will handle the linear/non-linear demand variations well, eliminating the need for different methods for different demand series.

3.1 LSTM memory cell structure

LSTM networks belong to the class of recurrent neural networks (RNNs). RNNs have the property of information persistence - i.e., retaining the state variables across time steps [15], thus making sequential learning over time steps feasible. The architecture of an RNN is presented in Figure 3.1, where □□ is the input, St is the hidden state of the cell and □□ is the output of the RNN cell at time t. RNN can only handle short- term dependencies because it suffers from a vanishing gradient problem. The LSTM networks, on the other hand, have the capability to learn long-term dependencies [18].

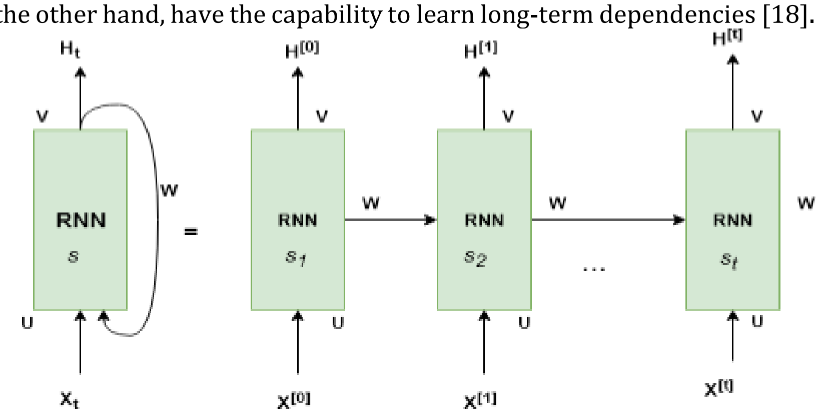

Fig.2: Recurrent Neural Network

The architecture of the LSTM network has three types of layers: 1) an input layer with a number of neurons equal to the number of input variables, 2) single or multiple hidden layers, and 3) an output layer with a number of neurons equal to the number of output variables. The hidden layers of LSTM networks consist of a memory cell. LSTM networks are superior to standard RNN due to the presence of this memory cell, which helps to retain information across time steps as this was not possible in earlier neural networks. The structure of the memory cell has three types of gates: 1) a forget gate (□□), 2) an input gate (□□), and 3) an output gate (□□). In memory cell (Figure 2), at each time step *t*, the input consists of an element from the input sequence (□□) and the output from the previous step ($h_{□-1}$). At cell state *t*: a) the forget gate takes these inputs and decide upon which information will be removed from memory, b) the input gate decides which information shall be added to memory (at cell state *t*), and c) the output gate decides the output of the memory block.

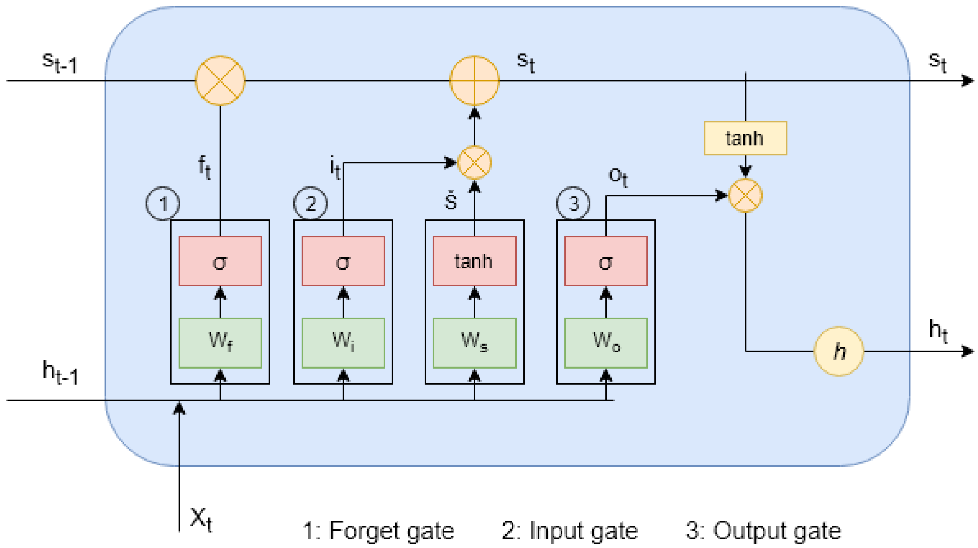

Fig.3: LSTM input data and memory cell architecture

LSTM networks network processes information through a sequence of four steps: In the first step, a sigmoid (a non-linear activation function) layer called the forget gate layer, which takes x_t and h_{t-1} as inputs and b_f as bias, computes the vector of activation values, f_t, for each of the values in cell state s_{t-1} within a normalized range between 0 (completely get rid- off) to 1 (completely keep). Then the activation value vector is calculated as follows [19]:

$$f_t = \sigma\,(W_{f,x}X_t + W_{f,h}h_{t-1} + b_f) \qquad (1)$$

In the second step, it is decided which information will be added to the memory cell state s_t. This step has two parts: first, candidate values $\tilde{s}_t$ are calculated. In the second step, an activation layer called the *input gate layer*, calculated as follows [19]:

$$\tilde{s}_t = tanh\,(W_{\tilde{s},x}X_t + W_{\tilde{s},h}h_{t-1} + b_{\tilde{s}t}) \qquad (2)$$

$$i_t = \sigma\,(W_{i,x}X_t + W_{i,h}h_{t-1} + b_i) \qquad (3)$$

In the third step, the cell state using new information was updated. Hadamard product was used in this step:

$$s_t = f_t . s_{t-1} + i_t . \tilde{s}_t \qquad (4)$$

In the last step, h_t, the output of the memory cell was calculated as follows [19]:

$$o_t = \sigma\,(W_{o,x}X_t + W_{o,h}h_{t-1} + b_o) \qquad (5)$$

$$h_t = o_t * \tanh(s_t) \qquad (6)$$

As shown in Figure 3.2, the input variables (demand and lagged variables) are inserted into the LSTM networks' input gate. LSTM network process the inputs step by step using Equations (1)-(6), and after completion of the process, it generates the final output sequence. The output vector is then the forecast of the network. LSTM networks are trained in multiple iterations known as epochs. During these iterations, bias and weights change to minimize the objective function across the training sets. For the task, the mean absolute error (MAE) is used as loss functions, and advanced hyper-parameter optimization through grid search is used to decide the final parameters of the prediction model [4]. Further details of hyper-parameter optimization are explained in Section 3.5.1 during empirical analysis.

Further, LSTM networks can provide a complete solution to model temporal and regression effects of the demand data. Moreover, errors are expected to minimize the proposed method [13]. However, the model was restricted to point forecasts; however, using bootstrap sampling, the prediction interval can be generated [3].

4. EXPERIMENTAL SETUP

4.1 Dataset

The dataset was taken from a multi-channel retailer selling packaged products through one online platform, and ten offline stores spread across a metro city. The data consist of weekly sales data for a number of products and span over 34 months of the time period. After performing initial data cleaning, data for ten products were obtained, each available for one online platform and ten offline stores, which leads to a total of 110 series. Along with available sales, date, price, item id, shop id, item category, item name variables, the calendar variables like a week, month, year, and a quarter are obtained from the date variable. The calendar variables are especially useful for regression methods. As the retailer requires demand forecasts for operational (short-term) decisions with a planning horizon of one week and one month, the last one-month was kept as a test dataset for out-of-sample weekly, and monthly evaluations utilized the rest of the data for training and validation.

4.2 Hardware/Software Platforms

The analysis was performed entirely in two freeware and open-source platforms: R statistical language and Python. Keras [8] was used (built on the top of TensorFlow, Theano, and CNTK, for implementing deep learning methods in Python. The python libraries hyperopt [4] and hyperas with Keras were used for the implementation of the grid search algorithm to find optimal settings of hyperparameters in LSTM networks.

The popular *randomforest* package [26] in R was used for fitting the RF model. For the benchmarking models, the auto.arima function and xreg parameter were used from the *forecast* package [20] for ARIMAX; furthermore, the *neuratnet* (Günther & Fritsch, 2010) package was used for neural networks, and the *caret* [24] package for training and evaluating the models. All the modeling and analysis are performed in a Core-i7 processor with 12 GBs of DDR4 RAM and 4GB of Graphic Processing Units (GPUs). Notably, LSTM networks are trained on GPUs.

4.2.1 Metrics

As in a typical forecasting competition, the performance of the proposed forecasting method and benchmarking methods is initially compared to three important characteristics. These are:

- *Bias* – to check the forecasting method for its tendency of over-forecasting or under- forecasting of actual values.
- *Accuracy* – to check how closely forecasted values conforms to the actual values.
- *Uncertainty* – to check the average deviation of the forecast from the mean forecast.

The metrics used to measure the characteristics mentioned above are as follows: mean error, mean absolute error, and mean squared error for bias, accuracy, and uncertainties, respectively [31]. An improvement to these metrics is made. The performance of a bunch of different methods is compared so that the relative measures will be more intuitive to implement. Also, the relative metrics are easy to interpret, scale, and allows to summarize the results from different series concisely. Therefore, relative mean error, relative mean absolute error, and relative mean squared errors are used to compare the performance of these methods. These relative errors are calculated by dividing the sum or mean of errors from the evaluated method by that of a benchmark method. The naïve method is used as the standard benchmark method to find the relative error for all other forecasting methods.

Furthermore, Pesaran and Timmermann [32] (PT) test is used to test the null hypotheses that forecast and actual values are independently distributed. The PT test will be applied to test the hypothesis that the predictive accuracy of the forecasting method is significant or not. Finally, to compare forecasting methods more comprehensively, the Diebold and Mariano (DM) [11] test is used. This statistical test is used to compare the forecast accuracy of any two forecasting methods. Through this test, it can be established whether the difference between accuracy from methods is statistically significant or not. This test tests the alternative hypotheses that the forecasts from method i have better accuracy than forecasts from method j, given $i, j \in$ a set of methods and $i \neq j$.

As the relative errors will be applied to multiple data series, therefore, relative errors across forecasts of various data series are aggregated to get a single metric to compare the forecasts. Therefore, average relative forecast errors are used. In this way, finally, five metrics are used to benchmark the performance of all forecasting methods. These are: average relative mean error (ARME), average relative mean absolute error (ARMAE), average relative mean squared error (ARMSE), the Pesaran and Timmermann [32] test, and the Diebold and Mariano [11] test.

4.3 Experimental Results

In this section, the forecasting performance of the model is presented, compared, and discussed. The forecasts for one-week and four weeks ahead were produced and benchmarked against random forests (RF), neural network (NN), ARIMAX, and multiple linear regression (MLR).

4.3.1 Results for the Online Channel: Online Shop

The forecasting methods are first applied to the data from the online store. The error metrics for the forecast are calculated, and the average of one-week ahead forecasting errors from all series for each forecasting method is presented in Table 2. These empirical results show that the forecasts from the hybrid time-series methods, i.e., deep learning and ARIMAx method, are more accurate, less biassed, and have less variance. These results for short-term forecasting horizons show that these methods are outperforming the benchmark methods on all three performance characteristics (with the respective metrics been: ARME, ARMAE, ARMSE).

Table 2. Online channel; one-week ahead: ARME, ARMAE, and ARMSE

	RF	NN	ARIMAX	MLR	LSTM
ARME	0.8854	0.9852	0.8735	0.8715	*0.6521*
ARMAE	0.9479	1.0415	0.8924	0.9812	*0.7154*
ARMSE	1.2461	1.4417	0.9680	0.9904	*0.7721*

Table 3 presents the average relative forecasting errors for the ten products for the forecasting horizon of 4 weeks (approximately one-month ahead). The results show that the most accurate is the LSTM networks based on average errors. The empirical results are the same across all three forecasting performance metrics. It can be learned that time-series methods (or sequence modeling methods) such as ARIMAX and LSTM are better performing in forecasting the pattern than the machine learning methods. The machine learning method can forecast some peak demand, which may be due to information available in explanatory variables available in data.

Table 3. Online channel; one-month (4 weeks) ahead: ARME, ARMAE, and ARMSE

	RF	NN	ARIMAX	MLR	LSTM
ARME	0.8481	0.9278	0.8437	0.6492	*0.6174*
ARMAE	0.9159	0.9856	0.8748	0.9222	*0.6895*
ARMSE	1.0365	1.2894	0.9381	0.9545	*0.7417*

Moreover, statistical significance tests were performed to empirically estimate the difference between the accuracy (forecasting errors) of the best method and the benchmarks above with the DM and PT tests (Table 4). The p-values cannot be aggregated across different products. Thus, the results are presented for weekly predictions for Product 1 in Table 3. The LSTM networks method outperformed all the other methods, and the differences in the forecasting performance are statistically significant at a 95% confidence level.

Table 4. Online channel; Statistical significance tests for Product #1 (weekly forecast). Panel A: DM test; Panel B: PT test.

A: DM Test						B: PT Test	
i *j* =	LSTM	ARIMAX	MLR	NN	RF	Method	Result
LSTM	-	0.0000	0.0000	0.0000	0.0000	RF	0.0000
ARIMAX		-	0.0001	0.0000	0.0210	NN	0.0000
MLR			-	0.0751	0.9849	ARIMAX	0.0000
NN				-	1.0000	MLR	0.0000
RF					-	LSTM	0.0000

Panel A in Table 4, the DM test, shows the p-values for the null hypothesis that the paired methods have equal performance. All individual hypotheses are rejected at 95% significance level over, and therefore, the deep learning method forecasts are superior to all benchmark methods: ARIMAX, NN, MLR, and RF. In panel A and panel B, it can be seen that the p-values for the PT test for the null hypothesis that actual data and forecast values are independently distributed. So, it can be assumed with high confidence that prediction and response from the forecasting method are not independently distributed, and thus, the selected forecasting methods have high prediction performance.

CONCLUSION

Deep learning forecasting method is used for demand forecasting. To benchmark the performance of the proposed forecasting method, a set of popular and widely used competitive methods was used, including Naïve, MLR, NN, ARIMAX, and RF. Three forecasting error metrics are employed ARME, ARMAE, and ARMSE as proxies for the bias, accuracy, and variance of the evaluated forecasts. Furthermore, the DM and PT statistical significance tests are used to attest to the empirical findings. The empirical evaluations over a real-world dataset of 110 time-series were performed and respective cues of information over two different channels: an online channel for ten products and offline channels, including the same ten products sold over ten different physical stores. All the analyses indicate that LSTM networks outperform all benchmarks in this study. The extensive empirical evidence presented here advocates the case for the potential of the use of deep learning methods for demand forecasting context, as well as calling for further research, testing, and development of these methods.

References

1. Aamer, A., Eka Yani, L., & Alan Priyatna, I. Data analytics in the supply chain management: Review of machine learning applications in demand forecasting. Operations and Supply Chain Management: An International Journal, 14(1), p. 1-13, 2020.
2. Abbasi, B., Babaei, T., Hosseinifard, Z., Smith-Miles, K., & Dehghani, M. Predicting solutions of large-scale optimization problems via machine learning: A case study in blood supply chain management. Computers & Operations Research, 119(2), p.104941, 2020.
3. Antipov, E. A., & Pokryshevskaya, E. B. Mass appraisal of residential apartments: An application of Random Forest for valuation and a CART-based approach for model diagnostics. Expert systems with applications, 39(2), p.1772-1778, 2012.
4. Bergstra, J., Yamins, D., & Cox, D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In International conference on machine learning, 17(3)p. 115-123, PMLR, 2013.
5. Bousqaoui, H., Achchab, S., & Tikito, K. Machine learning applications in supply chains: Long short-term memory for demand forecasting. In International Conference of Cloud Computing Technologies and Applications, 14(6)p. 301-317, Springer, Cham, 2019.
6. Baryannis, G., Dani, S., & Antoniou, G. Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. Future Generation Computer Systems, 16(4)101, p.993-1004, 2019.
7. Chong, E., Han, C., & Park, F. C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Systems with Applications,18(3) 83, p.187-205, 2017.
8. De Vass, T., Shee, H., & Miah, S. J. The effect of "Internet of Things" on supply chain integration and performance: An organisational capability perspective. Australasian Journal of Information Systems, 19(4)22, 2018.
9. Dash, R., McMurtrey, M., Rebman, C., & Kar, U. K. Application of artificial intelligence in automation of supply chain management. Journal of Strategic Innovation and Sustainability, 14(3), p.43-53, 2019.
10. Diebold, F.X.; Mariano, R. Comparing predictive accuracy. J. Bus. Econ. Stat.13, P.253–265, 1995.
11. Feizabadi, J. Machine learning demand forecasting and supply chain performance. International Journal of Logistics Research and Applications, 25(2), p.119-142, 2022.
12. Fischer, T., Krauss, C., & Treichel, A. Machine learning for time series forecasting-a simulation study (No. 02/2018). FAU Discussion Papers in Economics, 2018.
13. Fu, Y., & Zhu, J. Big production enterprise supply chain endogenous risk management based on blockchain. IEEE access, 7, p.15310-15319, 2019.
14. Graves, A., & Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural networks, 18(5-6), p.602-610, 2005.
15. Gao, Q., Xu, H., & Li, A. The analysis of commodity demand predication in supply chain network based on particle swarm optimization algorithm. Journal of Computational and Applied Mathematics, 19(3),400, p.113760, 2022.
16. Gawankar, S. A., Gunasekaran, A., & Kamble, S. A study on investments in the big data-driven supply chain, performance measures and organisational performance in Indian retail 4.0 context. International Journal of Production Research, 58(5), p.1574-1593, 2020.
17. Hochreiter, S., & Schmidhuber, J. Long short-term memory. Neural computation, 9(8), p.1735-1780, 1977.
18. Introduction to LSTM Units in RNN. Online Available at: https://www.pluralsight.com/guides/introduction-to-lstm-units-in-rnn
19. Hyndman, Rob J., Earo Wang, and Nikolay Laptev. Large-scale unusual time series detection. In 2015 IEEE international conference on data mining workshop (ICDMW),32(5), p. 1616-1619. IEEE, 2015.
20. Kochan, C. G., Nowicki, D. R., Sauser, B., & Randall, W. S. Impact of cloud-based information sharing on hospital supply chain performance: A system dynamics framework. *International Journal of Production Economics*,64(2) *195*, p.168-185, 2018.

21. Kone, E. R., & Karwan, M. H. Combining a new data classification technique and regression analysis to predict the Cost-To-Serve new customers. Computers & Industrial Engineering, 61(1), p.184-197, 2011.
22. Kosasih, E. E., & Brintrup, A. A machine learning approach for predicting hidden links in supply chain with graph neural networks. International Journal of Production Research, 60(17), p.5380-5393, 2022.
23. Kuhn, M. Building predictive models in R using the caret package. Journal of statistical software, 28, p.1-26,2008.
24. Li, Y., Wan, Z., & Liu, J. Bi-level programming approach to optimal strategy for vendor-managed inventory problems under random demand. The ANZIAM Journal, 59(2), p.247-270, 2019.
25. Liaw, A., & Wiener, M. Classification and regression by random Forest. R news, 2(3), p.18-22, 2002.
26. Mishra, J. L., Hopkinson, P. G., & Tidridge, G. Value Creation from Circular Economy led Closed Loop Supply Chains: A Case Study of Fast-Moving Consumer Goods, 2017.
27. Madani Saatchi, H., Arshadi Khamseh, A., & Tavakkoli-Moghaddam, R. Solving a new bi-objective model for relief logistics in a humanitarian supply chain by bi-objective meta-heuristic algorithms. Scientia Iranica, 28(5), p.2948-2971, 2021.
28. Makkar, S., Devi, G., & Solanki, V. K. Applications of machine learning techniques in supply chain optimization. In International Conference on Intelligent Computing and Communication Technologies,45(4) p. 861-869, Springer, Singapore, 2019.
29. Ni, D., Xiao, Z., & Lim, M. K. (2020). A systematic review of the research trends of machine learning in supply chain management. International Journal of Machine Learning and Cybernetics, 11(7), p.1463-1482, 2020.
30. Petropoulos, F., Kourentzes, N., & Nikolopoulos, K. Another look at estimators for intermittent demand. International Journal of Production Economics, 181, p.154-161, 2016.
31. Pesaran, M. H., & Timmermann, A. A simple nonparametric test of predictive performance. Journal of Business & Economic Statistics, 10(4), p.461-465, 1992.
32. Quddus, M. A., Chowdhury, S., Marufuzzaman, M., Yu, F., & Bian, L. A two-stage chance-constrained stochastic programming model for a bio-fuel supply chain network. International Journal of Production Economics, 195, p.27-44, 2018.
33. Schwarzmüller, T., Brosi, P., Duman, D., & Welpe, I. M. How does the digital transformation affect organizations? Key themes of change in work design and leadership. *Management Revue*, *29*(2), p.114-138, 2018.
34. Singh, A., Dwivedi, A., & Dubey, S. Rethink supply chain management: A machine learning perspective, 2022.
35. Supply chain management & Machine Learning. Online Available at: https://vitalflux.com/supply-chain-management-ai-machine-learning/
36. Tugay, R., & Oguducu, S. G. Demand prediction using machine learning methods and stacked generalization, 2020. arXiv preprint arXiv:2009.09756.