A Novel Ensemble Learning Approach for Intelligent Logistics Demand Management

Boyang Li[1*], *Yuhang Yang*[2], *Ziyu Zhao*[2], *Xin Ni*[2], *Diyang Zhang*[2]

[1] *Graduate School of Technology Management, Kyung Hee University, Republic of Korea*
[2] *Haide College, Ocean University of China, China*
liboyangline@yeah.net, yyh2469@stu.ouc.edu.cn, zhaoziyu6242@stu.ouc.edu.cn, nx571@stu.ouc.edu.cn, zhangdiyang@stu.ouc.edu.cn

Abstract

Logistics demand forecasting plays a crucial role in regulating logistics management activities, developing production plans, seeking maximum economic returns, and building smart logistics. Current studies have focused on forecasting logistics demand using various statistical algorithms and machine learning models. However, it is difficult for a single learner to forecast logistics demand time series with complex nonlinear fluctuation patterns. Therefore, a novel ensemble learning approach (Deep Logistics Demand Forecasting, DeepLDF) is introduced in this work to forecast logistics demand. DeepLDF consists of two different base learners, which are the Multi-scale Time-delay Convolution Model (MSTDCM) and the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. MSTDCM and SARIMA are combined for ensemble learning through a novel weight assignment approach. Based on six Singapore logistics demand data sets, DeepLDF is compared with nine different baselines. The experimental results show that DeepLDF performs well in fitting local extreme values and forecasting volatility. Overall, DeepLDF can forecast logistics demand well.

Keywords: Logistics demand, Nonlinear fluctuation patterns, Ensemble learning, Base learner

1 Introduction

In the context of economic globalization and regional economic integration, the logistics industry is gradually developing and growing [1]. After years of updating and iteration, the modern logistics patterns [2] have surpassed the traditional logistics patterns in terms of logistics support facilities, logistics informatization degree, and logistics specialized service level [3]. Especially, modern logistics has been able to realize automation, intelligence, and green logistics management [4]. Although some countries and regions are currently experiencing an imbalance between economic growth and logistics industry development, this imbalance is being gradually eliminated through emerging approaches such as automated allocation of market resources, supply chain optimization, and cross-border logistics cooperation [5]. Data released by the State Post Bureau of China show that the volume of express delivery business in China has rapidly increased from the initial 1.5 billion pieces to 20.67 billion pieces, with a Compound Annual Growth Rate (CAGR) of 43.9% on average, and has shown a trend of continuous rapid growth [6].

Logistics demand forecasting [7] is a way to forecast future logistics change trends based on historical logistics demand data, which usually involves the knowledge of statistics, operations research, and computer science. It is an important part of modern logistics. For production enterprises, accurate forecasting of logistics demand can assist enterprise managers in making reasonable and effective production plans, regulating their production capacity, and achieving a balance between production and demand [8]. For the transportation industry, accurate logistics demand forecasting can help freight drivers reduce the empty rate, save transportation time, and reduce transportation costs [9]. For consumers, logistics demand forecasting can increase the speed of cargo transportation and improve shopping satisfaction [10]. For the market, accurate logistics demand forecasting can meet market demand, cope with the uncertainty of the ordering cycle, mitigate the impact on the production sector due to delays and shortages of supplies, and stabilize prices, and manufacturing costs [11].

In this work, a novel ensemble learning approach (DeepLDF) is introduced to forecast logistics demand data with complex nonlinear fluctuation patterns. The main contributions of this work are as follows.

(1) It is difficult for a single learner to accurately forecast logistics demand data with complex nonlinear patterns. Therefore, a novel ensemble learning approach (DeepLDF) is introduced to model complex logistics demand time-series data.

(2) A novel approach for fusing the forecasting results of base learners is proposed. The final forecasting result is determined by the forecasting accuracy of each base learner. The final contribution weights of the base learners are proportional to the forecasting accuracy.

*Corresponding Author: Boyang Li; E-mail: liboyangline@yeah.net
DOI: https://doi.org/10.70003/160792642024072504002

2 Materials and Methods

2.1 Multi-scale Convolution

Multi-scale convolution [12] can process input features at multiple scales in parallel. For a feature map $X \in \mathbb{R}^{L\times W\times H}$, we can perform i convolutional transformations using the convolutional kernels of size $kernel_i$, which in turn yields i partial feature maps ($[X_0, X_1, \ldots, X_{i-1}]$). The feature maps for different parts represent deep features at different scales. The computation process of the feature map X_i of the ith part can be represented by Equation (1), and the multi-scale feature F can be represented by Equation (2). Where $kernel_i$ denotes the size of the ith convolution kernel. $Conv$ denotes the convolution process. $Concat$ denotes the feature map stitching process.

$$X_i = Conv(kernel_i \times kernel_i)(X). \tag{1}$$

$$F = Concat([X_0, X_1, \ldots, X_{i-1}]). \tag{2}$$

2.2 Designing the Forecasting Model

The logistics demand forecasting problem [13] is forecasting the logistics data y_t at the next moment (t) based on the historical logistics demand data ($[y_{t-n}, \ldots, y_{t-2}, y_{t-1}]$, where n denotes the sample number of historical data), as shown in Equation (3). Where $Framework$ denotes the modeling method, the input is $[y_{t-n}, \ldots, y_{t-2}, y_{t-1}]$, and the output is y_t.

$$y_t = Framework([y_{t-n}, \ldots, y_{t-2}, y_{t-1}]). \tag{3}$$

Overview of the DeepLDF

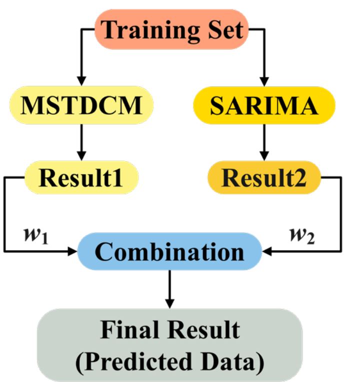

Figure 1. Framework of DeepLDF

As shown in Figure 1, two powerful base learners (MSTDCM and SARIMA) are introduced for ensemble learning to accurately forecast logistics demand. MSTDCM can mine the temporal dependence features of logistics demand time-series data at different scales and integrate them to fit the volatility and overall trend of the data. SARIMA can model the seasonality of logistics demand data, which in turn assists MSTDCM in improving the ability to fit local extreme values within each band. Besides, a novel approach for fusing the forecasting results of base learners is proposed to enhance the actual accuracy of ensemble learning, as shown in Equation (4). The final forecasting result is determined by the forecasting error of each base learner. The final contribution weights of the base learners are inversely proportional to the forecasting errors.

$$Final\ Result = \sum\nolimits_{i=1}^{2} Result(i) \frac{exp(RMSE(i))}{\sum\nolimits_{i=1}^{2} exp(RMSE(i))}. \tag{4}$$

$Final\ Result$ denotes the final forecasting result. $Result(i)$ denotes the forecasting result of the ith base learner. $exp(RMSE(i))/\sum_{i=1}^{2} exp(RMSE(i))$ denotes the fusion weight w_i of the ith base learner. exp denotes the natural constant. $exp(num)$ is equals to e^{num}. $RMSE(i)$ denotes the RMSE of the ith base learner.

MSTDCM

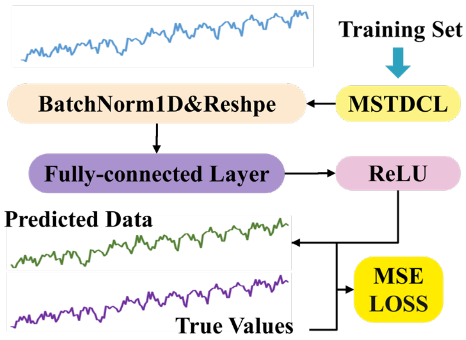

Figure 2. Structure of MSTDCM

MSTDCM is used to mine the multi-scale temporal correlation features of logistics demand data, as shown in Figure 2. Firstly, the training data are extracted by the Multi-scale Time-delay Convolution Layer (MSTDCL) with multi-scale deep features. Then, these deep features are mapped to the nonlinear sample space by the fully-connected layer and the ReLU activation function. Finally, the forecasting error of MSTDCM is accurately computed by the Mean Square Error (MSE) loss function.

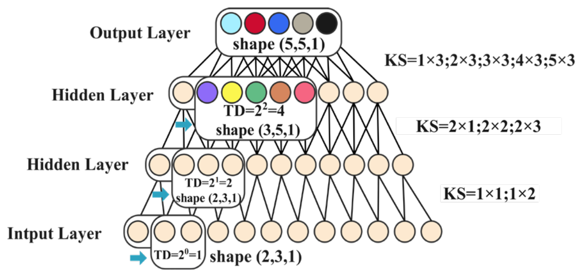

Figure 3. Working principle of the Multi-scale Time-delay Convolution Layer (MSTDCL)

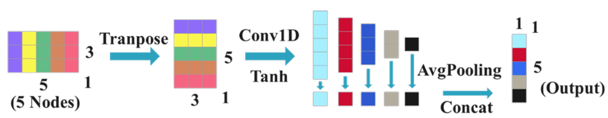

Figure 4. Example of multi-scale convolution operation

The specific computation process of MSTDCL is shown in Figure 3. On the one hand, when extracting temporal correlation features, we cannot know which feature mapping method is the best. Therefore, in the convolution process, we consider using Convolution Kernels (KSs) of different scales. During the convolution process, the width of the convolution kernel is kept constant and the length is varied from 1, incrementing by 1 each time until it equals the maximum length of the feature map. This allows considering multiple feature maps and saves computational costs. After obtaining the multi-scale feature maps, we process them by Average Pooling and obtain the result (Output) by the concatenation (Concat) operation, as shown in Figure 4. On the other hand, we map the features using a time sliding window [14]. We specify the Time Delay (TD) values within the time sliding window to vary with the convex function. This captures long-term historical information and extends the perceptual field of the convolution process [15]. We set the value of TD to vary according to 2 and 4. In this way, the forecasting approach does not lose local information because the oversized perceptual field [16-17].

3 Experiments

3.1 Data Sets and Baselines

Figure 5. Singapore (the study area)

We conduct experiments using six real-world Singapore logistics demand data sets (the study area is shown in Figure 5). We consider all historical data from 2005 to 2016 as training data and regard 2017 data as forecasting data. We set nine different baselines to validate the performance of DeepLDF because they include statistic methods, traditional machine learning models and deep neural networks.

ARIMA [18]: ARIMA is usually used to forecast future values in a stable time series.

SARIMA: Compared with ARIMA, SARIMA adds the technique of mining the seasonal features of the time series. Therefore, SARIMA is suitable for processing time-series data with significant seasonal features.

Ridge Regression (RR) [19]: Ridge regression is a complement to the least-squares regression. Ridge regression achieves higher computational accuracy by losing unbiasedness in exchange for high numerical stability.

Support Vector Regression (SVR) [20]: SVR is a Support Vector Machine (SVM) scheme for handling regression problems. SVR is a regression model that is mainly used to fit numerical values and is generally applied to scenarios with sparse features and a small number of features.

Recurrent Neural Network (RNN) [21]: RNN is very effective in dealing with time-series data. It can mine temporal relevance information as well as contextual semantic information in time-series data.

Long Short-term Memory (LSTM) [22]: The principle of LSTM is to retain long-term memory and forget unimportant information. LSTM is suitable for processing events with very long intervals and delays in time series.

Gated Recurrent United (GRU) [23]: Compared with LSTM, GRU is easier to be trained.

Temporal Convolutional Network (TCN) [24]: The architecture of TCN is not only more accurate than typical RNN, GRU, and LSTM but also simpler and clearer. Therefore, TCN is considered a suitable solution for applying Deep Neural Networks (DNNs) to time series.

MSTDCM: MSTDCM is a deep temporal feature extraction model based on multi-scale convolution and time-delay convolution operations.

3.2 Experimental Settings and Evaluation Metrics

In the experiment, we use the following computer configuration. Central Processing Unit (CPU): 11th Gen Intel (R) Core (TM) i9-11900 @ 2.50GHz. Graph Processing Unit (GPU): NVIDIA GeForce RTX 3060 (Memory: 12GB). Random Access Memory (RAM): 16GB x 4 (64GB in total).

During the training process, the data in 2017 is used for evaluation and the rest of the data is used for training. The learning rate is 0.001, the gradient optimizer is Adam optimizer, and the loss function is MSE Loss. (Experimentally, the performance of DeepLDF is optimal when the learning rate is 0.001. The experimental results are all results when the learning rate is equal to 0.001.)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i^{fore} - y_i\right|. \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i^{fore} - y_i\right)^2}. \tag{6}$$

In this experiment, we choose MAE and RMSE as the evaluation metrics for the experimental results. The methods for calculating MAE and RMSE are shown in Equation (5) and Equation (6), respectively. y_i represents the true value. y_i^{fore} represents the predicted value. n represents the total number of samples.

3.3 Logistics Demand Data Processing

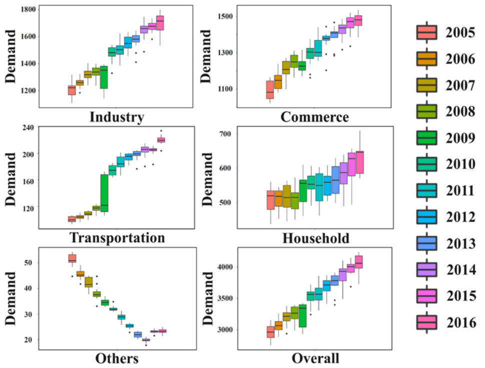

Figure 6. Logistics demand changes in six different logistics demand data sets

As shown in Figure 6, there are 6, 10, 5, 0, 9, and 5 outliers in the data sets Industry, Commerce, Transportation, Household, Others, and Overall, respectively. The horizontal axis represents the year and the vertical axis represents the logistics demand volume. The upper and lower ends of the solid part of each box indicate the upper quartile and lower quartile, respectively. The uppermost and lowermost parts of the box extensions indicate the maximum and minimum values, respectively. The horizontal line inside the box indicates the median. The points outside the box are outliers. These outliers may be caused by various errors, such as statistical errors or extreme data. These outliers may have a significant adverse impact on the data analysis and modeling process. Therefore, during data processing, all outliers are mapped between the lower and upper bounds of the box by a novel data deflation method that is scientifically valid in the study by Xu et al. [15, 25]. For these outliers, we cannot discard them because the weight of their influence on the logistics demand is important. Besides, this ensures that 100% (greater than or equal to 95%) of the important training information is not lost. The expression of the data deflation method is shown in Equation (7). V denotes the outlier. min and max denote the minimum and maximum values in the logistics demand data, respectively. lim_{max} and lim_{min} denote the upper and lower limits, respectively. V^{new} denotes the new value after mapping.

$$V^{new} = \begin{cases} \dfrac{(\lim_{max} - \lim_{min})(V - \lim_{max} + \lim_{min} - \min)}{\max - \lim_{max} + \lim_{min} - \min} + \lim_{min} \quad s.t.\ V \in (\lim_{max}, \max] \\ \dfrac{(\lim_{max} - \lim_{min})(V - \min)}{\max - \lim_{max} + \lim_{min} - \min} + \lim_{min} \quad s.t.\ V \in (\min, \lim_{min}]. \end{cases} \tag{7}$$

4 Results

4.1 Quantitative Analysis

The evaluation results of all forecasting models are shown in Table 1. As a traditional time-series data forecasting model, ARIMA can only forecast stable time-series data. When dealing with logistics demand data with complex nonlinear fluctuation patterns, ARIMA reaches its modeling bottleneck. SARIMA can handle logistics demand data with cyclical and seasonal features. Compared with ARIMA, SARIMA reduces MAE and RMSE by approximately 37.38% and 42.11%, respectively, on average. RR is a machine learning approach for modeling regression problems. However, RR inherently can only capture linear relationships and cannot model nonlinear features. Therefore, the forecasting results of RR are not good. SVR can map nonlinear time-series data to a high-dimensional space through high-dimensional mapping and make the mapped time-series data with linear features. Therefore, the forecasting performance of SVR is improved to a certain extent compared with RR. Compared with RR, SVR reduces MAE and RMSE by about 10.47% and 12.02% on average.

Table 1. Evaluation metrics for baselines and DeepLDF (The Best) on 6 data sets

	Overall		Industry		Commerce		Transportation		Household		Others	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
ARIMA	91.32	94.80	40.02	48.08	38.35	51.46	5.08	7.71	12.54	16.12	1.12	1.66
RR	88.65	89.72	38.86	42.71	34.19	46.84	5.06	7.62	12.88	16.94	1.44	1.82
SVR	80.44	81.03	34.77	36.26	31.18	41.29	4.18	5.99	10.51	14.87	1.06	1.53
RNN	81.32	82.68	35.14	37.06	30.62	40.48	4.63	6.54	11.33	15.09	1.13	1.66
GRU	74.78	76.27	32.51	33.29	26.56	34.90	3.98	5.65	9.66	12.30	1.03	1.49
LSTM	75.19	77.80	31.19	32.88	25.47	33.25	3.93	5.65	9.78	12.45	1.05	1.52
TCN	60.34	61.59	26.83	29.74	22.38	26.66	3.26	5.10	9.10	10.94	0.99	1.47
SARIMA	60.12	60.44	24.38	26.76	20.01	23.40	3.67	5.49	8.88	9.79	0.94	1.39
MSTDCM	42.94	43.12	13.98	15.73	16.14	17.22	2.18	2.84	6.57	6.80	0.73	0.96
The Best	**30.62**	**31.10**	**10.65**	**11.88**	**12.38**	**13.47**	**1.50**	**1.92**	**3.44**	**3.76**	**0.51**	**0.73**

RNN, GRU, and LSTM are classical forecasting models. Compared with RNN and GRU, LSTM can capture a longer period of historical information. Therefore, in this work, LSTM outperforms RNN and GRU. The residual structure in TCN overcomes the problem that traditional recurrent neural networks do not support parallel computation and slow training speed compared with the gate structure. Therefore, in this work, TCN has better generalization ability than RNN, GRU, and LSTM.

MSTDCM is a deep temporal feature mining model based on multi-scale convolution and time-delay convolution. We cannot know which feature mapping approach is the best. Therefore, we use different convolution kernels in MSTDCM to model deep temporal dependence features of logistics demand data. The period of historical data is large. Therefore, we use a time sliding window for feature mapping in MSTDCM. The MSTDCM is designed to calculate the temporal dependence. Therefore, almost all the historical information can be retained. The MSTDCL consists entirely of convolutional networks. The actual forecasting error of MSTDCM is minimal compared with other baselines.

Compared with baselines, DeepLDF has the best performance. On the one hand, SARIMA is better than MSTDCM in expressing the seasonal trends of logistics demand time-series data. SARIMA can compensate for the deficiency of the powerful MSTDCM in mining seasonal features. On the other hand, we use a novel approach to fuse the forecasting results of SARIMA and MSTDCM. For the base learner with high forecasting accuracy, we give it a large weight. Compared with the forecasting results of baselines, DeepLDF decreases MAE and RMSE by approximately 59.76% and 62.23% on average, respectively.

4.2 Qualitative Analysis

As we can see in Figure 7 and Figure 8, ARIMA can identify the overall trend of the logistics demand data. However, ARIMA cannot accurately fit the local extreme values within each band. Compared with ARIMA, SARIMA has better forecasting results for seasonal variation patterns, which is reflected in SARIMA's ability to accurately fit the local extreme values within each band. Figure 9 and Figure 10 show that the predicted values of RR are almost always large. Compared with RR, the forecasting accuracy of SVR is significantly improved. This is because of the ability of SVR to tap the nonlinear trend of the logistics demand time-series data using high-dimensional mapping. However, SVR is only able to fit the approximate trends of logistics demand data within each band to a certain extent. Therefore, SVR cannot forecast accurately.

From Figure 11, Figure 12, Figure 13, and Figure 14, we can find that the forecasting results of RNN are all on the low side. The fitting ability of RNN gradually decreases as the forecasting time increases. This is because the RNN cannot remember long-term historical data. Compared with RNN, the accuracy of both GRU and LSTM is improved to some extent. Figure 15 and Figure 16 show that both TCN and MSTDCM can forecast logistics demand with some degree of accuracy. Compared with TCN, MSTDCM has a better ability to control the extreme values of the data. Besides, compared with MSTDCM, the DeepLDF method is the best, as shown in Figure 17 and Figure 18.

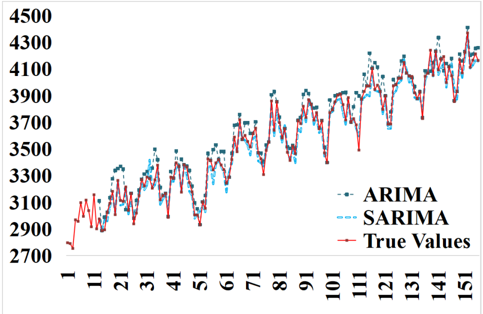

Figure 7. Forecasting results of ARIMA and SARIMA from 2006 to 2017

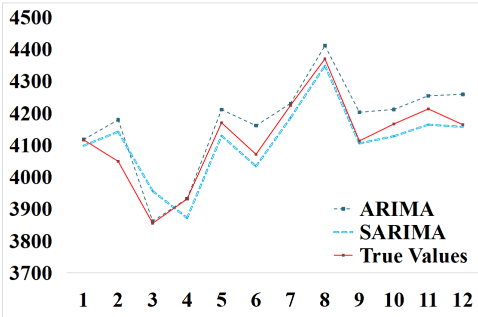

Figure 8. Forecasting results of ARIMA and SARIMA in 2017

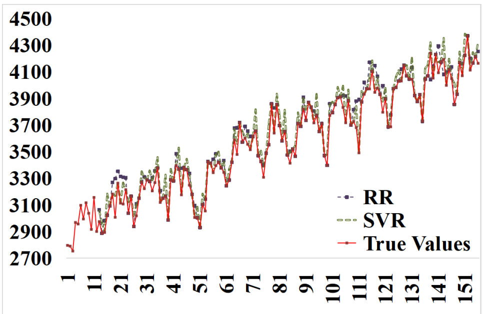

Figure 9. Forecasting results of RR and SVR from 2006 to 2017

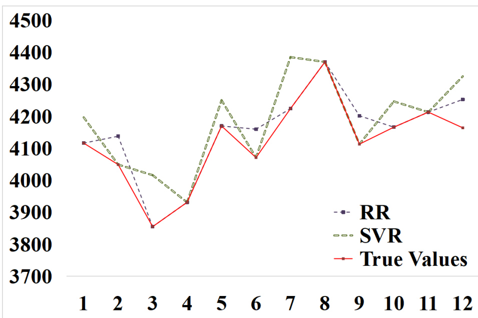

Figure 10. Forecasting results of RR and SVR in 2017

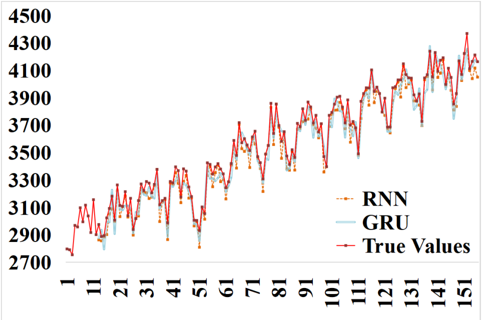

Figure 11. Forecasting results of RNN and GRU from 2006 to 2017

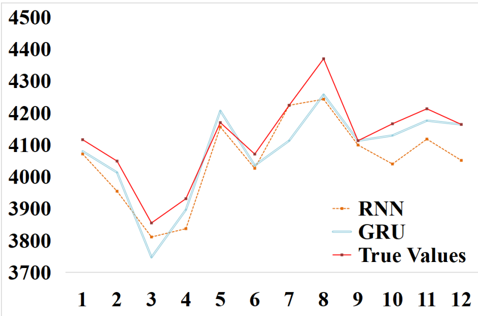

Figure 12. Forecasting results of RNN and GRU in 2017

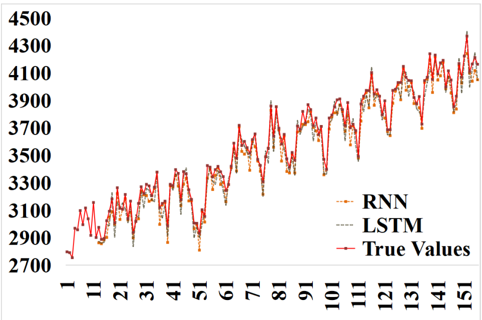

Figure 13. Forecasting results of RNN and LSTM from 2006 to 2017

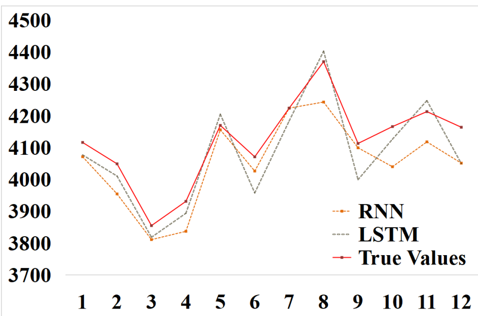

Figure 14. Forecasting results of RNN and LSTM in 2017

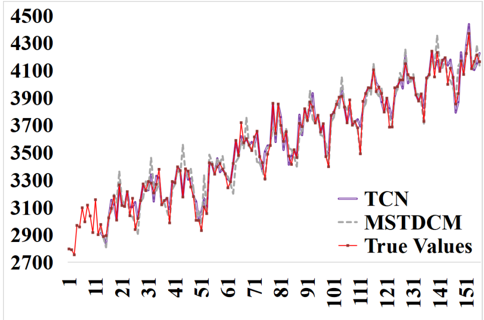

Figure 15. Forecasting results of logistics demand data from 2006 to 2017

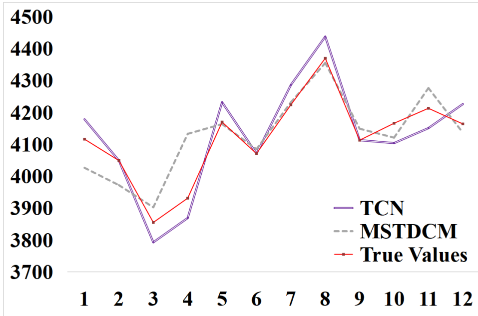

Figure 16. Forecasting results of logistics demand data in 2017

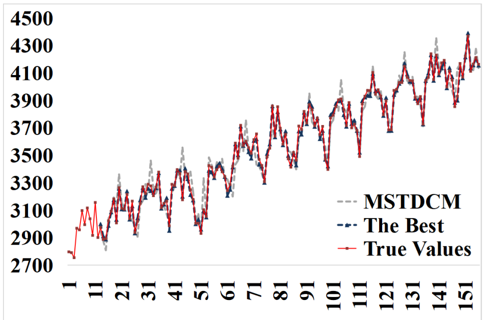

Figure 17. Forecasting results of logistics demand data from 2006 to 2017

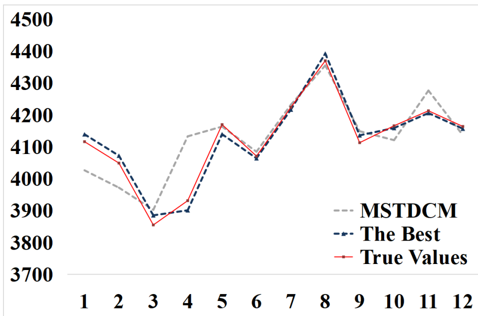

Figure 18. Forecasting results of logistics demand data in 2017

Table 2. Evaluation metrics for 3 types of MSTDCM on 6 data sets

	Overall		Industry		Commerce		Transportation		Household		Others	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
MSTDCM(3)	48.84	50.07	16.33	18.73	20.19	20.80	2.84	3.36	7.29	7.51	0.93	1.42
MSTDCM(5)	45.12	45.98	14.62	16.32	18.06	18.97	2.32	3.00	6.82	6.98	0.85	1.00
MSTDCM(4)	**42.94**	**43.12**	**13.98**	**15.73**	**16.14**	**17.22**	**2.18**	**2.84**	**6.57**	**6.80**	**0.73**	**0.96**

4.3 Ablation Study

Table 2 shows the actual forecasting results of the three different MSTDCMs in this work. Among them, MSTDCM(3) denotes the MSTDCM with one input layer, one hidden layer, and one output layer (three layers in total). In this case, the time delay (TD) values are set to 30 (from the input layer to the hidden layer) and 31 (from the hidden layer to the output layer). MSTDCM(4) denotes the MSTDCM with one input layer, two hidden layers, and one output layer (four layers in total). MSTDCM(4) is the MSTDCM we finally choose and use. MSTDCM(5) denotes the MSTDCM with one input layer, three hidden layers, and one output layer (five layers in total). In this case, the time delay (TD) values are set to 1 (from the input layer to the first hidden layer), 2 (from the first hidden layer to the second hidden layer), 3 (from the second hidden layer to the third hidden layer) and 4 (from the third hidden layer to the output layer). Compared with MSTDCM(3) and MSTDCM(5), MSTDCM(4) has an average reduction of approximately 10.36% and 10.69% in MAE and RMSE, respectively. Therefore, MSTDCM(4) with two hidden layers has the best results.

5 Discussion

There are a large number of classical methods in statistics and operations research. Yan et al. [26] used the inverse of variance weighted assignment method to combine the gray model and exponential smoothing model, and then built a port logistics demand forecasting model. They reduced the randomness of the original data by combining various forecasting models. Besides, this effectively improves the accuracy of the forecasting model. Zhang et al. [27] used the SARIMA-Markov to forecast the total coal transportation of the Da-Qin Railway in the future period. SARIMA is highly capable of tapping seasonal and cyclical variation patterns. However, SARIMA-Markov does not consider the many factors that affect coal transportation. Therefore, SARIMA-Markov is not strong in fitting local extreme values. Tonchiangsai et al. [28] used an improved ARIMA to forecast cable demand. However, as the size of cable demand data increases, the accuracy of ARIMA forecasting becomes less and less accurate. It is difficult for ARIMA to model accurately when dealing with time-series data with nonlinear patterns.

To achieve intelligent decision-making in logistics resource scheduling, many studies have started to explore the potential of deep learning in logistics demand forecasting work. Lou et al. [29] demonstrated the feasibility of deep learning models in logistics demand forecasting by Back Propagation (BP) neural networks. Therefore, they established a BP neural network based on a particle swarm optimization algorithm to achieve short-term forecasting of logistics demand. Leng et al. proposed some effective feature adaptive selection and weighting methods with high discriminative power [30], high accuracy and low correlation [31].

Compared with the current work, our work makes some improvements. DeepLDF consists of two base learners [32] (MSTDCM and SARIMA). Besides, a novel approach for fusing the forecasting results of the base learners is introduced in DeepLDF. The results of contrast and ablation experiments show that the four-layer MSTDCM structure designed based on logistics demand data can accurately fit the overall trend of logistics demand time-series data. SARIMA can compensate for the deficiency of MSTDCM in fitting the local extreme values within each band. Besides, the fusion method of forecasting results based on forecasting errors can improve the practical generalization ability of DeepLDF. Compared with MSTDCM and SARIMA, DeepLDF decreased the MAE and RMSE by approximately 41.09% and 41.22%.

6 Conclusion

To accurately forecast logistics demand with complex nonlinear fluctuation patterns, a novel ensemble learning-based DeepLDF model is introduced. DeepLDF includes two base learners, which are MSTDCM and SARIMA. Experimental results show that MSTDCM can mine the temporal correlation features of logistics demand time series by multi-scale convolution. Besides, MSTDCM can compute long-term historical information by time-delay convolution and time sliding window. SARIMA can compensate for the deficiency of MSTDCM in fitting local extreme values by modeling the seasonal variation pattern of logistics demand data. The fusion method of forecasting results based on forecasting errors can effectively improve the accuracy of DeepLDF.

Although DeepLDF has high forecasting accuracy, the time overhead of DeepLDF designed based on the ensemble learning idea may be large. In this work, we do not discuss the balance between forecasting accuracy and speed. Compared with the proposed MSTDCM, although the forecasting accuracy of DeepLDF is higher, the forecasting time may be longer. In some scenarios with high requirements for real-time performance, maybe MSTDCM is a better choice.

References

[1] D. Zhang, F. Zhang, Y. Liang, An evolutionary model of the international logistics network based on the Belt and Road perspective, *Physica A*, Vol. 572, Article No. 125867, June, 2021.

[2] X. Xu, Y. He, Blockchain application in modern logistics information sharing: A review and case study analysis, *Production Planning & Control*, pp. 1-15, April, 2022. https://doi.org/10.1080/09537287.2022.2058997

[3] M. Zielske, T. Held, Application of agile methods in traditional logistics companies and logistics startups: Results from a German Delphi Study, *Journal of Systems and Software*, Vol. 177, Article No. 110950, July, 2021.

[4] J. C. Ma, L. Wang, Z. Q. Jiang, W. Yan, W. X. Zhou, City logistics networks based on online freight orders in China, *Physica A*, Vol. 583, Article No. 126333, December, 2021.

[5] E. A. Lezhnina, Y. E. Balykina, Cooperation between sea ports and carriers in the logistics chain, *Journal of Marine Science and Engineering*, Vol. 9, No. 7, Article No. 774, July, 2021.

[6] J. S. Pan, Z. Fu, C. C. Hu, P. W. Tsai, S. C. Chu, Rafflesia Optimization Algorithm Applied in the Logistics Distribution Centers Location Problem, *Journal of Internet Technology*, Vol. 23, No. 7, pp. 1541-1555, December, 2022.

[7] S. S. Abosuliman, A. O. Almagrabi, Computer vision assisted human computer interaction for logistics management using deep learning, *Computers & Electrical Engineering*, Vol. 96, Article No. 107555, December, 2021.

[8] Y. Zhao, Z. Li, T. Wang, Internet-of-thing based real-time electrical market monitoring system design, *Computers & Electrical Engineering*, Vol. 100, Article No. 107872, May, 2022.

[9] K. Tsolaki, T. Vafeiadis, A. Nizamis, D. Ioannidis, D. Tzovaras, Utilizing machine learning on freight transportation and logistics applications: a review, *ICT Express*, pp. 1-12, February, 2022. https://doi.org/10.1016/j.icte.2022.02.001

[10] M. Zeng, R. Liu, M. Gao, Y. Jiang, Demand forecasting for rural e-commerce logistics: a gray prediction model based on weakening buffer operator, *Mobile Information Systems*, Vol. 2022, Article No. 3395757, July, 2022.

[11] X. Cao, J. Zhang, Preference learning and demand forecast, *Marketing Science*, Vol. 40, No. 1, pp. 62-79, January-February, 2021.

[12] S. Tang, Y. Zhu, S. Yuan, An improved convolutional neural network with an adaptable learning rate towards multi-signal fault diagnosis of hydraulic piston pump, *Advanced Engineering Informatics*, Vol. 50, Article No. 101406, October, 2021.

[13] P. He, S. Ma, W. Li, Efficient Barrage Video Recommendation Algorithm Based on Convolutional and Recursive Neural Network, *Journal of Internet Technology*, Vol. 22, No. 6, 1241-1251, November, 2021

[14] J. Wu, H. Xu, R. Sun, P. Zhuang, Road Boundary-Enhanced Automatic Background Filtering for Roadside LiDAR Sensors, *IEEE Intelligent Transportation Systems Magazine*, Vol. 14, No. 4, pp. 60-72, July-August, 2022.

[15] Z. Xu, Z. Lv, J. Li, A. Shi, A Novel Approach for Predicting Water Demand with Complex Patterns Based on Ensemble Learning, *Water Resources Management*, Vol. 36, No. 11, pp. 4293-4312, September, 2022.

[16] Z. Lv, J. Li, C. Dong, H. Li, Z. Xu, Deep learning in the COVID-19 epidemic: A deep model for urban traffic revitalization index, *Data & Knowledge Engineering*, Vol. 135, Article No. 101912, September, 2021.

[17] A. K. Dubey, A. Kumar, V. García-Díaz, A. K. Sharma, K. Kanhaiya, Study and analysis of SARIMA and LSTM in forecasting time series data, *Sustainable Energy Technologies and Assessments*, Vol. 47, Article No. 101474, October, 2021.

[18] A. L. Schaffer, T. A. Dobbins, S. A. Pearson, Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions, *BMC Medical Research Methodology*, Vol. 21, No. 1, pp. 1-12, March, 2021.

[19] A. R. Zarei, M. R. Mahmoudi, Evaluation and comparison of the effectiveness rate of the various meteorological parameters on UNEP aridity index using backward multiple ridge regression, *Water Resources Management*, Vol. 35, No. 1, pp. 159-177, January, 2021.

[20] X. Chang, Z. Feng, J. Wu, H. Sun, G. Wang, X. Bao, Understanding and Predicting the Short-Term Passenger Flow of Station-Free Shared Bikes: A Spatiotemporal Deep Learning Approach, *IEEE Intelligent Transportation Systems Magazine*, Vol. 14, No. 4, pp. 73-85, July-August, 2022.

[21] S. I. Popoola, B. Adebisi, M. Hammoudeh, H. Gacanin, G. Gui, Stacked recurrent neural network for botnet detection in smart homes, *Computers & Electrical Engineering*, Vol. 92, Article No. 107039, June, 2021.

[22] Y. Zhai, L. Yang, J. Yang, L. He, Z. Li, BadDGA: Backdoor Attack on LSTM-Based Domain Generation Algorithm Detector, *Electronics*, Vol. 12, No. 3, Article No. 736, February, 2023.

[23] X. Han, F. Chen, J. Ban, Music Emotion Recognition Based on a Neural Network with an Inception-GRU Residual Structure, *Electronics*, Vol. 12, No. 4, Article No. 978, February, 2023.

[24] Z. Lv, J. Li, H. Li, Z. Xu, Y. Wang, Blind travel prediction based on obstacle avoidance in indoor scene, *Wireless Communications & Mobile Computing*, Vol. 2021, Aricle No. 5536386, June, 2021.

[25] Z. Xu, Z. Lv, J. Li, H. Sun, Z. Sheng, A Novel Perspective on Travel Demand Prediction Considering Natural Environmental and Socioeconomic Factors, *IEEE Intelligent Transportation Systems Magazine*, Vol. 15, No. 1, pp. 136-159, January-February, 2023.

[26] P. Yan, L. Zhang, Z. Feng, J. Zhang, Research on logistics demand forecast of port based on combined

model, *Journal of Physics: Conference Series*, Vol. 1168, No. 3, pp. 1-6, March, 2019.
[27] C. Zhang, S. Liu, Forecast and Analysis of Coal Traffic in Daqin Railway Based on the SARIMA-Markov Model, *Journal of Advanced Transportation*, Vol. 2020, Article No. 1276305, December, 2020.
[28] K. Tonchiangsai, G. Boonsothonsatit, Electrical Cable Demand Prediction Using ARIMA, *Proceedings of the 10th International Conference on Industrial Technology and Management (ICITM)*, Cambridge, UK, 2021, pp. 111-114.
[29] X. Lou, M. Qin, Y. Zhao, C. Zhang, Research on Logistics Demand Forecasting Model of Truck Broker Based on PSO-BPNN, *International Core Journal of Engineering*, Vol. 4, No. 12, pp. 66-74, December, 2018.
[30] L. Leng, M. Li, C. Kim, X. Bi, Dual-source discrimination power analysis for multi-instance contactless palmprint recognition, *Multimedia Tools and Applications*, Vol. 76, No. 1, pp. 333-354, January, 2017.
[31] Z. Xu, Z. Lv, B. Chu, J. Li, A Fast Spatial-temporal Information Compression Algorithm for Online Real-time Forecasting of Traffic Flow with Complex Nonlinear Patterns, *Chaos, Solitons & Fractals*, Vol. 182, Article No. 114852, May, 2024.
[32] S. Ochella, M. Shafiee, F. Dinmohammadi, Artificial intelligence in prognostics and health management of engineering systems, *Engineering Applications of Artificial Intelligence*, Vol. 108, Article No. 104552, February, 2022.

Biographies

Boyang Li received her B.S. degree from Jilin University, China, in 2023. She is currently pursuing the master degree at Graduate School of Technology Management, Kyung Hee University, Republic of Korea. Her recent research interests are artificial intelligence and education technologies.

Yuhang Yang enrolled at Ocean University of China in 2020 to study Mathematics and Applied Mathematics. His recent research interest is intelligent logistics demand management.

Ziyu Zhao enrolled at Ocean University of China in 2020 to study Mathematics and Applied Mathematics. His recent interests are time series analysis and data statistics.

Xin Ni is currently pursuing the bachelor's degree with the Ocean University of China of Mathematics and Applied Mathematics. His recent research interest is big data mining.

Diyang Zhang enrolled at Ocean University of China in 2020 to study Mathematics and Applied Mathematics. His current research interest is time-series analysis.