

Cargo volume prediction of logistics sorting center based on GCN-BiLSTM

Huihui Guo^{*,1,a}

¹Communication University of China Beijing, China

^a2021213053005@cuc.edu.cn

Abstract. Taking into account the transport network and average cargo volume of each sorting center, a directed weighted graph is constructed in this paper. Next, the GCN model is used to extract the spatial characteristics of the transport connection information of the sorting center and input it into the BiLSTM network. The BiLSTM network uses the two-way information flow to learn the temporal characteristics, and then uses the GCN-BiLSTM model combined with the spatio-temporal characteristics of the sorting center to predict the daily cargo volume in the next 30 days. The integrated learning model based on ARIMA and BiLSTM is then used to predict the next 30 days of hourly cargo volume, and adjust and optimize. The results show that GCN-BiLSTM model and BiLSTM model improve the prediction performance.

Keywords: Ensemble learning; GCN model; BiLSTM model;

1. Introduction

The rapid development of the Internet has fundamentally changed the internal structure and business operation mode of the e-commerce logistics supply network. E-commerce logistics network is composed of multiple links in order fulfillment. As the middle link of the network, the sorting center maintains many important aspects of the logistics process such as transportation and warehousing. The cargo volume forecast of the sorting center, including the daily cargo volume forecast and hourly cargo volume forecast, can not only reflect the change law of logistics demand to a certain extent, but also further obtain the future change trend of logistics resource allocation through the forecast of freight volume. It can be said that the accuracy of the flow prediction affects the subsequent operation arrangement of all related logistics work.

It is worth noting that the transportation route of the network will also affect the cargo volume forecast of the sorting center. When the line relationship is adjusted and updated, the cargo distribution of each sorting center should also be updated in time to adapt to the new line relationship. At the same time, the personnel scheduling forecast caused by the change of cargo volume is also an urgent problem to be solved. Reasonable arrangement of regular and temporary workers according to the result of cargo volume prediction can improve the efficiency of cargo volume sorting and reduce personnel costs.

In 2007, Ju Songdong et al. put forward the theory of logistics network [1]. With the passage of time, a group mainly led by Yan Lijun et al. studied the logistics network of cities [2]. As for the cognition of logistics network, many famous scholars in China have conducted researches in various aspects: Xu Jie et al provided the theoretical research method of logistics network [3], pointing out that the study of logistics network is conducive to the further development of logistics enterprises. Chen Guorong et al. proposed a logistics network modeling method named NGM [4], and Huang Jicong published a method of applying mathematical modeling to logistics networks [5], which introduced logistics network models and established corresponding model algorithms to further promote the development of logistics management systems. Recently, Du Xiaohui et al. 's research on resource allocation of logistics network under random environment has made the logistics network model more effective [6].

In the multiple links of logistics network, the existence of sorting center is undoubtedly the core of the whole network, and the research on sorting center has become the academic focus. At present, there are many academic studies on the problems of cargo flow prediction and human resource

allocation. There is a great correlation between the operation rules of logistics network and the management efficiency of distribution center and the operation cost. A large number of scholars have studied the logistics prediction and the distribution of manpower: Hou Rui et al. MLP neural network model was proposed [7], Zhang Jin established prediction theory to study logistics demand [8], Geng Yong et al proposed neural network to analyze and forecast logistics demand [9], Zhang Peng further studied the logistics prediction method of neural network and proposed a three-layer neural network learning algorithm [10]. Cao Ping et al proposed GA-grey neural network [11] Li Minjie et al proposed RBF neural network model to forecast logistics demand [12], Feng Zhongmiao established GRNN neural network [13], and a team composed of wang Jiuhe et al established PSO-BP neural network to study logistics services [14]. Nie Yiwen et al. provided a new method for the study of logistics prediction through PCA-GA-BP neural network [15]. In the same year, a group led by Qiao Na et al. established GM(1,N) model [16]. Xu Guojun et al. studied the distribution of human capital [17], and Lu Jiangdong also expressed his own views on the management of human resources [18].

2. Correlation methodology

2.1 Establishment of GCN-BiLSTM model

2.1.1 GCN model:

In this paper, taking the sorting center Network as an example, we establish a Graph Convolutional Network (GCN) model, which is used to analyze the connection relationships of the transportation lines of the sorting center and determine the adjacency relationships between the sorting centers. In addition to the information of the nodes, we also consider the direction and weight information of the edges.

Based on Annex III and Annex IV, we correspond the new cargo volume data of 57 sorting centers. Based on these data, we construct a graph structure, in which each sorting center corresponds to a node in the graph, while the transportation line corresponds to the edge in the graph, and a directed weighted graph is obtained, in which the node represents the sorting center, and the edge represents the transportation line between the sorting center. The weight of the edge represents the cargo information of the transport line.

Next, we use GCN to process the graph structure data in order to extract the feature information and relationships between nodes. The core idea of GCN is to extract information from adjacency matrix and node feature matrix through multi-layer graph convolution operation. The principle of GCN is to process a set of graph data containing nodes, the features of these nodes form a dimensional feature matrix X , and the relations between each node form a dimensional adjacency matrix A . Taking sum as input. The specific formula is shown as formula (1)

$$H^{(l+1)} = \sigma(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} H^{(l)} W^{(l)}) \quad (1)$$

In practice, two ChebNet convolution operations are performed on GCN to fully extract the feature information between nodes. In the first ChebNet convolution operation, we mainly focus on extracting first-order neighbor features of nodes. The first-order neighbor feature refers to the feature of the sorting center directly connected to the target node, that is, the node adjacent to the target node. In the second ChebNet convolution operation, the second-order neighbor characteristics of the nodes are further considered. The second-order neighbor feature refers to the feature of the node connected to the target node through an edge, that is, the neighbor of the target node's neighbor. Through two ChebNet convolution operations, we can make full use of the first and second order neighbor features of the target node, so as to obtain richer node feature information and provide more accurate input data for subsequent tasks.

Finally, according to the node characteristics and connection relationships learned by the GCN model, as well as the edge weight information, we can evaluate the importance and load of different transport lines, so that the sorting center network can be analyzed and predicted.

2.1.2 BiLSTM model:

LSTM (Long Short-Term Memory network) solves the problem of long-term dependence in traditional recurrent neural networks and makes them perform better in processing time series data [2]. However, since the transmission of the cell state in LSTM is one-way from front to back, only the information of the past moment can be used, and the information of the future moment cannot be directly obtained. To solve this problem, BiLSTM (Bidirectional LSTM, Bidirectional Long short-term memory Network) was proposed. BiLSTM takes advantage of the flow of information in two directions, forward and backward, enabling it to consider both past and future information. Through recursion and feedback, BiLSTM can learn the characteristics of future volumes while using past information. In BiLSTM, input sequences are processed through LSTM layers in both directions at the same time, and their outputs are then spliced or merged to obtain a complete bidirectional representation of information. In this way, the model can not only capture the time before and after the past, but also predict the future, thus improving the accuracy and robustness of the prediction. BiLSTM has been widely used in various time series prediction tasks and has achieved remarkable results. The BiLSTM network unit is the same as the LSTM network unit as described in 5.2. Figure 1 shows the BiLSTM network structure. The BiLSTM network is a bidirectional structure that allows inputs to flow in both directions to preserve future and past information. Therefore, the BiLSTM-based model can better mine the correlation features of time series data.

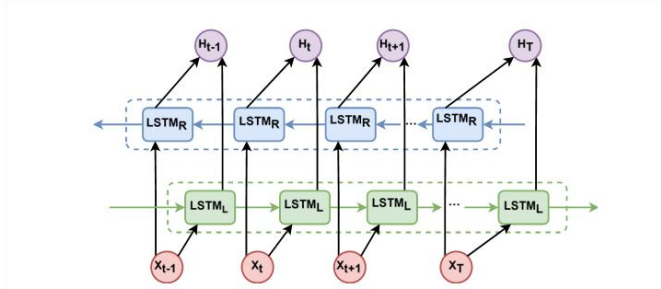


Figure 1: BiLSTM network structure

Assume that \vec{h}_t is the hidden layer state of the forward LSTM network at a certain time, and its calculation formula is shown in equation (2).

$$\vec{h}_t = \text{LSTM}(\vec{x}_t, \vec{h}_{t-j}) \quad (2)$$

Here x_t is the input at this time t and \vec{h}_{t-1} is the hidden layer state $t-1$ of the forward LSTM network at some point. \vec{h}_t is the hidden layer state of the LSTM network at t certain time, and its calculation formula is shown in equation (3):

$$\vec{h}_t = \text{LSTM}(\vec{x}_t, \vec{h}_{t-1}) \quad (3)$$

Here x_t is the input t at this time, and \vec{h}_{t-1} is the hidden layer state $t-1$ of the later LSTM network at a certain time. The output of the BiLSTM network is the combination of two hidden layer states, \vec{h}_t and \overleftarrow{h}_t , to form the entire hidden state network.

2.1.3 GCN-BiLSTM model:

First, the original cargo volume time series data of 57 sorting centers were pre-processed, such as data cleaning, missing value processing and standardization, and then the data set was divided into training set, verification set and test set. Then, through the two-layer GCN network, the spatial feature extraction of the transport connection information of the sorting center is realized, and the information is input into the BiLSTM network. The BiLSTM network uses two-way information flow to learn the timing features in the sequence on the time series and obtains the output values by optimizing the dropout and processing the six fully connected layers. In the training stage, the mean square error loss is calculated between the output value and the actual cargo volume, and the model is continuously optimized by the Adam optimizer. In the test phase, the predicted value of goods volume in the next 30 days by GCN-BiLSTM model is directly obtained without loss calculation and model optimization. Dropout is a regularization technique used to reduce overfitting in neural networks. During training, Dropout randomly sets the output of a subset of neurons to zero (that is, discard), updating their weights to zero with a certain probability. This prevents the neural network from becoming overly dependent on specific input features, thereby improving the generalization ability of the model. In this study, root mean square error (RMSE) and mean absolute error (MAE) were used as evaluation indexes to measure the model performance and analyze the prediction accuracy. The RMSE and MAE values can be viewed as absolute errors between the true and predicted values.

3. Experimental analysis

Compared with traditional LSTM, BiLSTM has significant advantages in that its bidirectional information flow mechanism enables the model to consider both past and future cargo information at the same time, so as to capture long-term dependencies in sequence data more comprehensively and improve the forecasting ability of the model. We use an integrated learning model based on GCN and Bidirectional Long Short-term Memory Network (BiLSTM) to predict the amount of goods per hour over the next 30 days. When analyzing the Annex 2 data, we found that the data is clearly cyclical, with distinct peaks and troughs occurring at different points in time each day. As shown in Figure 5-10, the method 5.3 still lacks in predicting the hourly volume of goods in the next 30 days. The high sequence of special festivals will affect the daily volume of goods in the following days, resulting in generally large predicted value of daily volume of goods. Therefore, the integrated learning model based on GCN and BiLSTM is used for improvement. We determine the weight of the two models by minimizing the root mean square error (RMSE) of each sorting center to obtain the final prediction result.

Based on the processed data, we collated it to obtain a data set of the daily volume of 57 sorting points in the past 4 months and the transportation route network of sorting centers. The data set is divided into training set, verification set and test set according to the ratio of 8:1:1, and the model RMSE and MAE are calculated. By using the GCN-BiLSTM model to train the data set, the cargo volume forecast of 57 sorting points in the next 30 days is obtained.

Table 1. Experimental results of cargo volume prediction for the next 30 days at SC33 sorting point

Date	RMSE	MAE
2023/12/1	3.534	4.414
2023/12/2	3.459	4.358
2023/12/3	3.745	4.574
2023/12/4	3.862	4.682
2023/12/5	4.273	5.799
2023/12/6	4.479	4.171
2023/12/7	4.635	3.854
2023/12/8	4.956	4.963

.....
2023/12/23	4.149	6.534
2023/12/24	5.455	6.459
2023/12/25	5.624	5.745
2023/12/26	6.368	6.862
2023/12/27	7.023	7.273
2023/12/28	8.671	7.479
2023/12/29	7.283	8.635
2023/12/30	9.579	6.956

According to the experimental results of the cargo volume prediction in the next 30 days at the SC33 sorting point of a sorting center, the GCN-BiLSTM model has better prediction performance in the prediction of different sorting centers. Both the RMSE error and MAE error of these models appear to be on the rise as the number of forecast days increases. On the forecast day, the GCN-BiLSTM model outperformed the LSTM method in terms of RMSE. The GCN-BiLSTM model, combined with spatial factors and bidirectional time flow, performs better than the LSTM model which only considers a single factor.

Visualize the prediction of SC43 sorting points, as shown in the figure 2 below:

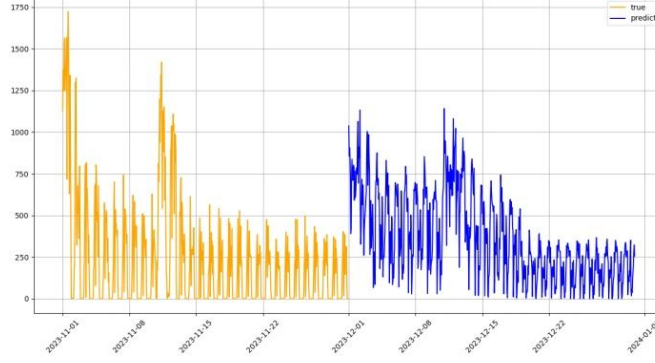


Figure 2: SC43 sorting point volume forecast for the next 30 days

According to the observation of the forecast trend chart of the cargo volume in the past and the next 30 days at the SC43 sorting point, the forecast shows that the cargo volume of "Double 12" is reduced compared with that of "Double 11", while the cargo volume of natural day is relatively stable, similar to the trend of the real cargo volume in November, showing a better forecast result. In addition, RMSE is 10.32, MAE is 11.84, low error values show that the prediction model fits the observed values well, further confirming the accuracy of the prediction.

By comparing the evaluation indicators of multiple sorting centers, we found that the integrated learning model based on ARIMA and BiLSTM is better than the integrated learning model based on LSTM and ARIMA in terms of cargo volume prediction, reflecting the advantages of BiLSTM model. Meanwhile, the attention mechanism is used to improve the model's attention to the data in a specific period of time. The GCN-BiLSTM model predicts the cargo volume in the next 30 days after the change of the transportation line network of the sorting center, and the methods such as restricting and improving the cargo volume in the next 30 days and hours also improve the accuracy of the prediction accuracy.

4. Conclusion

In this paper, a GCN-BiLSTM model is proposed, which has many advantages. Firstly, it extracts the spatial features of the transport connection relationship between the sorting centers through the GCN network, and learns the sequence features of the time series data through the BiLSTM network, which enables the model to effectively capture the relationship and structural features between the sorting centers and improve the modeling ability of the time series data.

Secondly, regularization using Dropout technology effectively reduces the overfitting risk of the model, enhances the generalization ability of the model, and makes the model more robust. In addition, the GCN-BiLSTM model has high flexibility and can be adjusted and improved according to different problems, which is suitable for various time series prediction tasks.

However, the GCN-BiLSTM model also has some shortcomings. First of all, the computational complexity of the model is high, involving the computation of multi-layer GCN network and BiLSTM network, so a lot of computational resources need to be consumed in the training and reasoning process. Secondly, it is difficult to adjust the hyperparameters of the model, which involves the selection and adjustment of multiple hyperparameters, which requires a lot of time and energy. Nevertheless, the GCN-BiLSTM model still has wide application prospects in the field of time series prediction, and its structure can be further extended and optimized to improve the performance and generalization ability of the model. At the same time, GCN-BiLSTM model can be widely used in stock forecasting, weather forecasting and other fields. The model performance can be improved by combining different neural network structures and data types.

References

- [1] JU Songdong, Xu Jie. Logistics network theory and its research significance and methods [J]. China Circulation Economy, 2007(08):10-13.
- [2] Yan Lijun, Yang Zhongzhen, Liu Chong, et al. Research on distribution and scale optimization of intermediate nodes in urban logistics Network [J]. Journal of Dalian University of Technology, 2007(03):414-418.
- [3] JU Songdong, Xu Jie. Logistics network theory and its research significance and methods [J]. China Circulation Economy, 2007(08):10-13.
- [4] Chen Guorong, Yan Ping, Peng Jun, et al. A logistics network modeling method based on growth [J]. Journal of South China University of Technology (Natural Science Edition), 2008(05):24-29+37.
- [5] HUANG J C. Logistics network model analysis based on Mathematical modeling [J]. Journal of Heilongjiang Ecological Engineering Vocational College, 2016, 29(03):33-35.
- [6] Du Xiaohui, Meng Fanfeng. Logistics network traffic distribution based on random environments research [J]. Water transport in China, 2021 (8) : 97-99..
- [7] Hou Rui, Zhang Bixi. Regional logistics demand forecasting Method based on MLP neural network and its application [J]. Systems Engineering Theory & Practice, 2005(12):43-47.
- [8] Zhang Jin. Research on L-OD Prediction Theory and Modern logistics planning method [D]. Southwest Jiaotong University, 2005.
- [9] Geng Yong, JU Songdong, Chen Yana. Logistics demand Analysis and prediction based on BP neural network [J]. Logistics Technology, 2007(07):35-37+73.
- [10] Zhang P. Research on prediction method based on neural network and its application in logistics system [D]. Wuhan University of Technology, 2002.
- [11] Cao Ping, Chen Fuji. GA - regional logistics demand forecasting of grey neural network [J]. Journal of Beijing institute of technology (social science edition), 2012, 14 (01) : 66-70.
- [12] Li Minjie, Wang Jian. Research on cold chain logistics demand prediction of aquatic products based on RBF neural network [J]. China Agricultural Resources and Regional Planning, 2020, 41(06):100-109.
- [13] Feng Shemiao. Research on Social logistics cost prediction based on GRNN neural Network [J]. Journal of Wuhan University of Technology (Transportation Science and Engineering Edition), 2010, 34(06):1209-1212.
- [14] Wang Jiuhe, Liu Huan, GAO Hui. Research on Cold chain logistics Service provider Selection based on rough PSO-BP neural network [J]. Industrial Engineering (in Chinese), 201, 24(02):10-18.
- [15] Nie Yiwen, Wang Wei. Forecasting of logistics demand in Anhui Province based on PCA-GA-BP neural network model [J]. Journal of Fuyang Normal University (Natural Science Edition), 2023, 40(04):96-102.

- [16] Jonna, Fang Qing. Research on demand prediction of agricultural products logistics in Hubei Province -- Based on GM(1,N) model [J]. Logistics Engineering and Management, 2019,45(10):66-69+74.
- [17] Xu Guojun, Xia Hong. On the distribution mechanism of human capital and individual income [J]. Journal of Ocean University of Qingdao (Social Science Edition),1999(01):51-58.
- [18] Lu Jiangdong. Problems and countermeasures of enterprise human resource management [J]. China Commerce and Trade,2014(28):53-54