



# Deep Data Analysis-Based Agricultural Products Management for Smart Public Healthcare

Wenjing Yan<sup>1</sup>, Zesheng Zhang<sup>1</sup>, Qingchuan Zhang<sup>1</sup>, Ganggang Zhang<sup>2</sup>, Qiaozhi Hua<sup>3\*</sup> and Qiao Li<sup>4</sup>

<sup>1</sup> National Engineering Laboratory for Agri-Product Quality Traceability, Beijing Technology and Business University, Beijing, China, <sup>2</sup> Digital Campus Construction Center, Capital Normal University, Beijing, China, <sup>3</sup> Computer School, Hubei University of Arts and Science, Xiangyang, China, <sup>4</sup> Chongqing Key Laboratory of Intelligent Perception and Blockchain Technology, Chongqing Technology and Business University, Chongqing, China

## OPEN ACCESS

### Edited by:

Ali Kashif Bashir,  
Manchester Metropolitan University,  
United Kingdom

### Reviewed by:

Durai Raj Vincent P. M.,  
VIT University, India  
Heng Wang,  
Henan Agricultural University, China  
Shizheng Zhang,  
Zhengzhou University of Light  
Industry, China  
Weizheng Wang,  
Cloud County Community College,  
United States

### \*Correspondence:

Qiaozhi Hua  
11722@hbuas.edu.cn

### Specialty section:

This article was submitted to  
Digital Public Health,  
a section of the journal  
Frontiers in Public Health

**Received:** 01 January 2022

**Accepted:** 21 February 2022

**Published:** 07 April 2022

### Citation:

Yan W, Zhang Z, Zhang Q, Zhang G,  
Hua Q and Li Q (2022) Deep Data  
Analysis-Based Agricultural Products  
Management for Smart Public  
Healthcare.  
Front. Public Health 10:847252.  
doi: 10.3389/fpubh.2022.847252

Agriculture is an indispensable public healthcare industry for human beings at any time and smart management of it is of great significance. Since substantial technical advance relies on long-term efforts and continuous progress, reasonably scheduling the distribution of agricultural products acts as a key aspect of smart public healthcare. The most intuitive factor affecting the distribution of agricultural products is its dynamic price. Forecasting price fluctuations in advance can optimize the distribution of agricultural products and pave the way to smart public healthcare. Most researchers study the prices of various agricultural products separately, without considering the interaction of different agricultural products in the time dimension. This study introduces a typical deep learning model named graph neural network (GNN) for this purpose and proposes deep data analysis-based agricultural products management for smart public healthcare (named GNN-APM for short). The highlight of GNN-APM is to take latent correlations among multiple types of agricultural products into consideration when modeling evolving rules of price sequences. A case study is set up with the use of real-world data of the agricultural products market. Simulative results reveal that the designed GNN-APM functions well.

**Keywords:** graph neural network, agricultural products, public healthcare, deep data analysis, smart management

## 1. INTRODUCTION

Since ancient times, agriculture has been a life industry for human survival, which is closely related to the most basic life guarantee of human beings. At present, food shortage is one of the most important problems faced by many regions in the world (1). This phenomenon is generally reflected in two aspects (2). For one thing, there is still room for improvement in current agricultural technology, which makes grain yield fail to meet the expectations (3). For another, due to the lack of scientific management and scheduling strategy, the production of agricultural products is not reasonable (4). The exploration of advanced agricultural technology has lasted for at least a 100 years, and some technological breakthroughs have been made in some key fields (5, 6). However, this process exerts an imperceptible influence as we all know, which needs to be accumulated over a long period of time to make progress (7). The use of advanced computing technology to manage the agricultural products market can improve the distribution and dispatch efficiency of global agricultural products to a certain extent (8), and then alleviate the problem of food shortage (9). The key to improving management efficiency is to forecast the market conditions of several major types

of agricultural products (10). To realize such a goal, data-driven methods are the most intuitive ways (11, 12).

Rationalizing the distribution of agricultural products is an important aspect of smart public healthcare. Predicting price fluctuations in advance can optimize the layout of agricultural products and pave the way for smart public healthcare. Many scholars have studied the agricultural products market in recent years (13–32). For example, Fan et al. (14) analyzed the value-added mechanism of agricultural products circulation value chain and put forward three optimization methods of agricultural products organization mode. These existing researches are mainly realized through the research methods of social sciences. Their analyses focus on the mechanism of social development and evolution but have to deal with large-scale computational tasks in a manual way. Without the assistance of intelligent computing, these methods are always faced with certain limitations. There are also some scholars who use intelligent computing methods to solve this problem. They regard the forecasting problem of the future agricultural products market as a time series forecasting problem based on historical data. However, most of them just treat different agricultural products as independent categories and then build time series prediction models separately. In this way, potential relationships between categories are ignored, so that the precision of the modeling process is reduced.

In order to solve the above challenges, a graph neural network (GNN) can be used to model this time series prediction problem (33). Different from the traditional model that deals with grid-structured data, GNN deals with the data of topological structure. GNN is based on deep learning (34) and is widely used in various fields due to its good performance and interpretability (35). It can deeply perceive the relationship between entities in the process of modeling by graph-structured data. Vertexes and edges connecting vertexes together constitute graph structure. The vertexes are object entities, and the edges are the specific relationship between the entities. Specifically, several common agricultural products can be regarded as entities, and the correlation between them can be regarded as edges, which together constitute a kind of graph network. By introducing a neural computing structure, a GNN model for time series prediction can be constructed. Therefore, this study designs a graph neural network-based smart management for the agricultural products market (GNN-APM). The main highlights of this article can be summed up as follows:

- The internal complexity of agricultural products systems is investigated for further management.
- The GNN is employed to construct a time-series price prediction method for agricultural products systems.
- A case study is carried out to evaluate the performance of the proposed method on real-world scenes.

## 2. SYSTEM MODEL

The left part of **Figure 1** illustrates the learning and training process inside GNN-APM. First, the feature space of different types of agricultural products is coded into vectorized

representations. Then a prediction model is obtained by learning the data samples. As a necessity in our life, agricultural products have a strong correlation among various categories, and these correlations are very meaningful and worth exploring further. The complex interrelation between different categories affects their demand to a great extent. The introduction of the GNN model deeply excavates the correlation between different categories of agricultural products and builds a graph structure with agricultural products as the entity. Through analysis, the market conditions of agricultural products can be predicted and further intelligent management of public health can be realized.

Generalized to the problem scenario in this study, there are several types of agricultural products whose market conditions need to be forecasted. The learning algorithm of GNN-APM is shown in **Algorithm 1**. Types of agricultural products are viewed as the set of nodes, and their internal relations are regarded as the set of edges. Market conditions of agricultural products refer to the average market price in this research and will be updated temporally. Each time that the market condition is updated, is defined as a timestamp  $t$  which ranges from 1 to  $M$ . During each timestamp, market conditions of all the agricultural types are denoted as  $V_i^{(t)}$ , where  $i$  is the index number of agricultural products types. Inputting market conditions data of  $M$  timestamps, the main goal is to predict unknown market conditions data of following timestamps. Obviously, internal relations among nodes are likely to influence the tendency of market conditions. Thus, a relationship-aware sequential forecasting problem is formulated, and the GNN model is adopted to deal with such a problem.

---

### Algorithm 1 | The learning algorithm of GNN-APM.

---

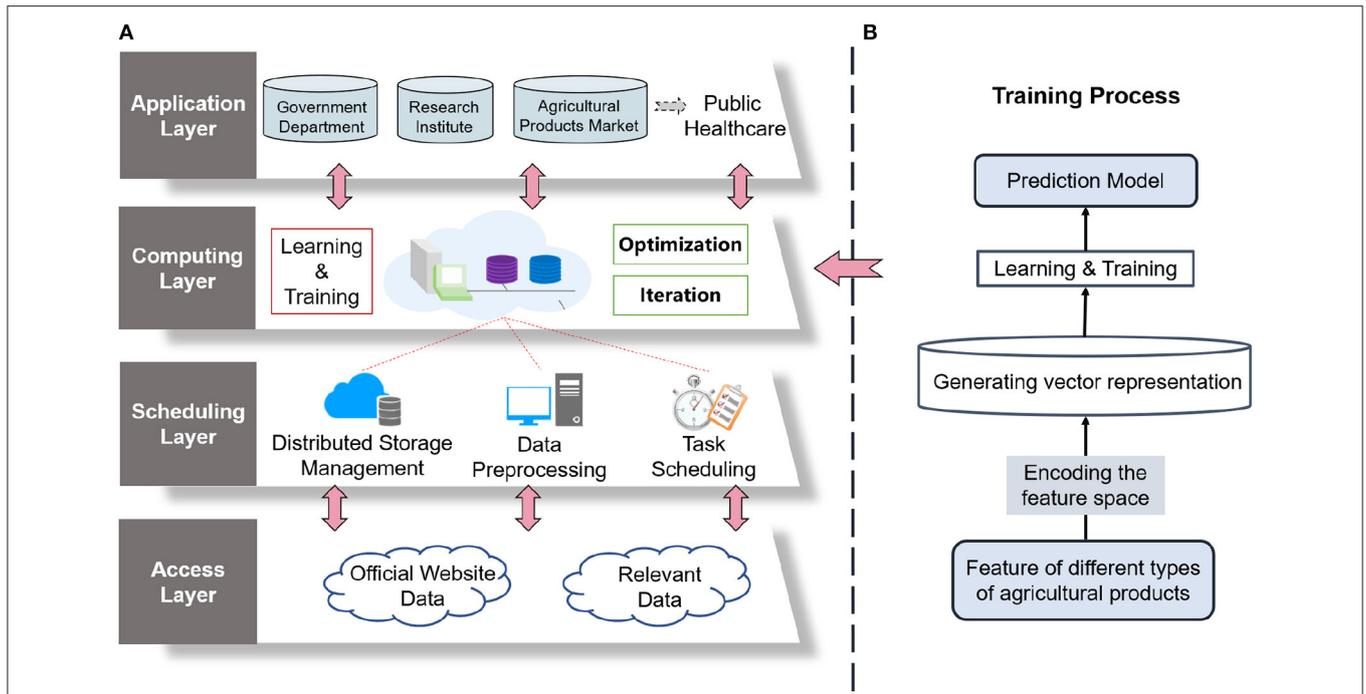
**Input:** The market condition dataset:  $D$ ; node set:  $V$ ; number of agricultural products:  $N$ ; total timestamp:  $M$ ; learning rate:  $l$ ; parameter set:  $\Theta$ ; penalty parameter:  $p$ ;

- 1: initial  $iter = 1$
- 2: **repeat**
- 3:   **for**  $t \in [1, M]$  **do**
- 4:     **for**  $i \in [1, N]$  **do**
- 5:       compute  $V_i^{(t)} = f(\theta(; ; id(V_i); \Theta)$ ;
- 6:       compute  $loss = \sum_{i=1}^N \sum_{t=1}^M [\|V_i^{(t)} - \hat{V}_i^{(t)}\| + p * \|\Theta\|_F^2]$ ;
- 7:       compute the gradient of  $\Theta$  according to  $loss$ ;
- 8:       update model parameters  $\Theta$  according to their gradients and learning rate  $l$ ;
- 9:     **end for**
- 10:   **end for**
- 11:    $iter = iter + 1$ ;
- 12: **until** convergence

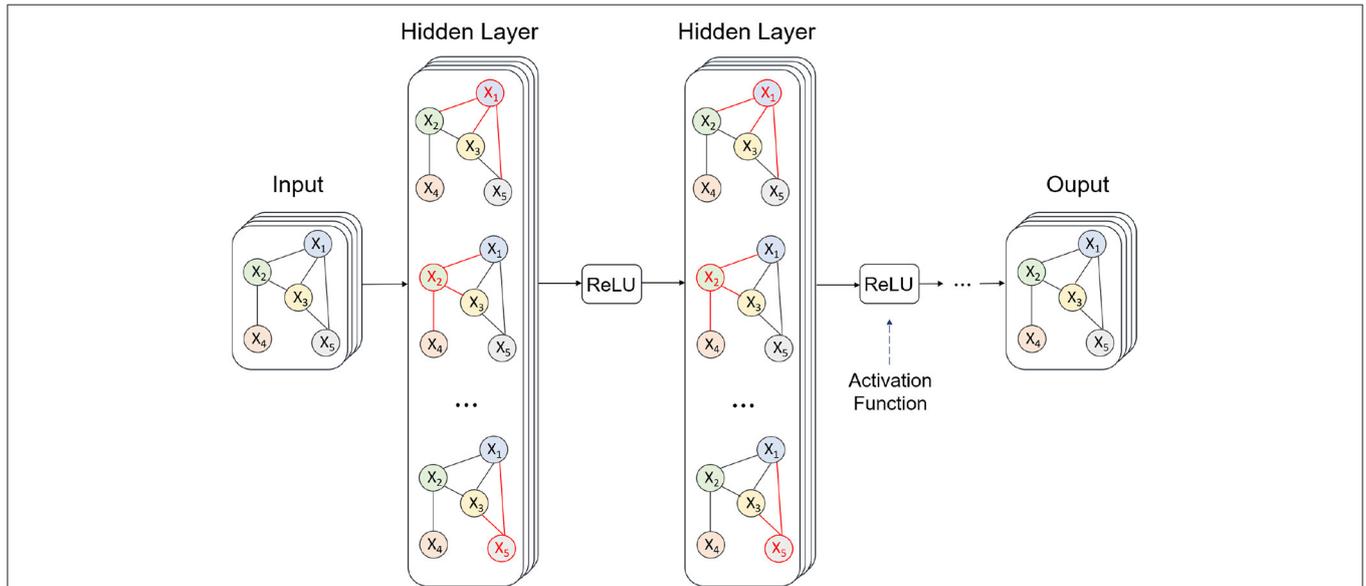
---

## 3. METHODOLOGY

Graph convolution network (GCN), a typical GNN model, is utilized here to model correlated sample space. The GCN extends the convolution operation to non-Euclidean data with graph



**FIGURE 1 |** System model and overview of the designed graph neural network-based smart management for agricultural products market (GNN-APM). **(A)** architecture and **(B)** training process in computing layer.



**FIGURE 2 |** Workflow of the graph convolutional network employed for prediction.

structure. It is a deep learning method for graph structured data. Graph data can naturally represent data structures in real life, such as traffic networks, communication networks, and social networks. In other words, it is the way to represent this kind of data format. Unlike image and text data, graph data has different local structures for each node. This is because the nodes in the

graph represent the different entities in the network, and the edges that connect the nodes represent the relationships between the entities.

As is shown in **Figure 2**, taking graph structure as input, GCN obtains new node representation through graph convolution operation on neighbor nodes of each node in the graph. Then,

all nodes are pooled to obtain the representation of the whole graph. In particular, an undirected graph with nodes is defined as  $G(V, E)$ , where  $V$  is the number of nodes and  $E$  is the edge between two nodes. Enumerating  $i$  from 1 to  $N$ ,  $v_i$  constitutes the node set  $V$ . Let  $j$  denote the index number of nodes different from node  $i$ , edge  $e_{ij}$  constitutes all the edges between pairs of nodes. Additionally, all the edge states inside graph  $G$  are able to make up an adjacency matrix  $A$ .

There are two kinds of GCN methods: spectral method and spatial method. Spectral CNN is the first method to construct a convolutional neural network on the graph. This method uses the convolution theorem on the graph to define graph convolution from the spectral domain. Specifically, it uses the convolution theorem to define the graph convolution operator in each layer. Under the guidance of the loss function, it learns the convolution kernel by gradient backpropagation and builds neural networks by stacking multiple layers. The spectral method is more general in most time-series prediction problems and is selected as the main technique of this study.

The spectral GCN method derives from the Fourier transform (FT) theory which can transform signals of the time domain into signals of the frequency domain. After such transformation, complicated convolution operations of the time domain can be approximated as multiplication operations of the frequency domain. The FT and inverse Fourier transform (IFT) are defined as follows:

$$\hat{x} = FT(x) = U^T x \tag{1}$$

$$x = IFT(x) = U \hat{x} \tag{2}$$

where  $U$  is the eigenvector that approximates the FT to matrix computation. Thus, the graph convolution operation is defined as follows:

$$x \otimes_G g = U \left( U^T x \odot U^T g \right) \tag{3}$$

where  $x$  is the input,  $\otimes_G$  is the graph convolution operator,  $g$  is the core, and  $\odot$  is the harmand product operator. Introducing Laplacian eigenvector as a basis function, the input signal can be expanded as:

$$x \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(\gamma) \end{bmatrix} = (u_1, u_2, \dots, u_\gamma) \begin{bmatrix} \hat{x}(1) \\ \hat{x}(2) \\ \vdots \\ \hat{x}(\gamma) \end{bmatrix} \tag{4}$$

Expanding  $g$  with matrix forms and then substituting Equation (4) into Equation (3), the following formula can be deduced:

$$x \otimes_G g = U \begin{bmatrix} \hat{g}(\lambda_1) & & \\ & \ddots & \\ & & \hat{g}(\lambda_\gamma) \end{bmatrix} \tag{5}$$

Letting  $Chev(\lambda_1)$  denote first-order Chebyshev polynomials of  $\lambda_1$ , the convolution operator  $g$  can be approximated as:

$$g_\theta = U \begin{bmatrix} Chev(\lambda_1) & & \\ & \ddots & \\ & & Chev(\lambda_\gamma) \end{bmatrix} \tag{6}$$

Among, the  $Chev(\lambda_\gamma)$  can be represented as:

$$Chev(\lambda_\gamma) = \xi_0 \Gamma_0(\lambda_\gamma) + \xi_1 \Gamma_1(\lambda_\gamma) \tag{7}$$

Hence, Equation (5) can be rewritten as the following formula:

$$x \otimes_G g = \left[ \alpha_0 - \alpha_1 \left( E^{-\frac{1}{2}} \cdot A \cdot E^{-\frac{1}{2}} \right) \right] x \tag{8}$$

where  $E$  is the degree matrix, and  $\alpha_0$  and  $\alpha_1$  are parameters to be learned. For simplicity, it is supposed to set  $\alpha_0 = \alpha_1 = -\theta$ . Therefore, the above equation can be rewritten as:

$$x \otimes_G g = \left[ \theta \left( E^{-\frac{1}{2}} \cdot A \cdot E^{-\frac{1}{2}} + D_\eta \right) \right] x \tag{9}$$

In order to facilitate searching for optimum, renormalization operation is conducted on the above formula:

$$E^{-\frac{1}{2}} \cdot A \cdot E^{-\frac{1}{2}} + D_\eta \approx E^{-\frac{1}{2}} \cdot \tilde{A} \cdot E^{-\frac{1}{2}} \tag{10}$$

$$\tilde{A} \approx A + D_\eta \tag{11}$$

$$\tilde{E}_{ij} \approx \sum_i \tilde{A}_{ij} \tag{12}$$

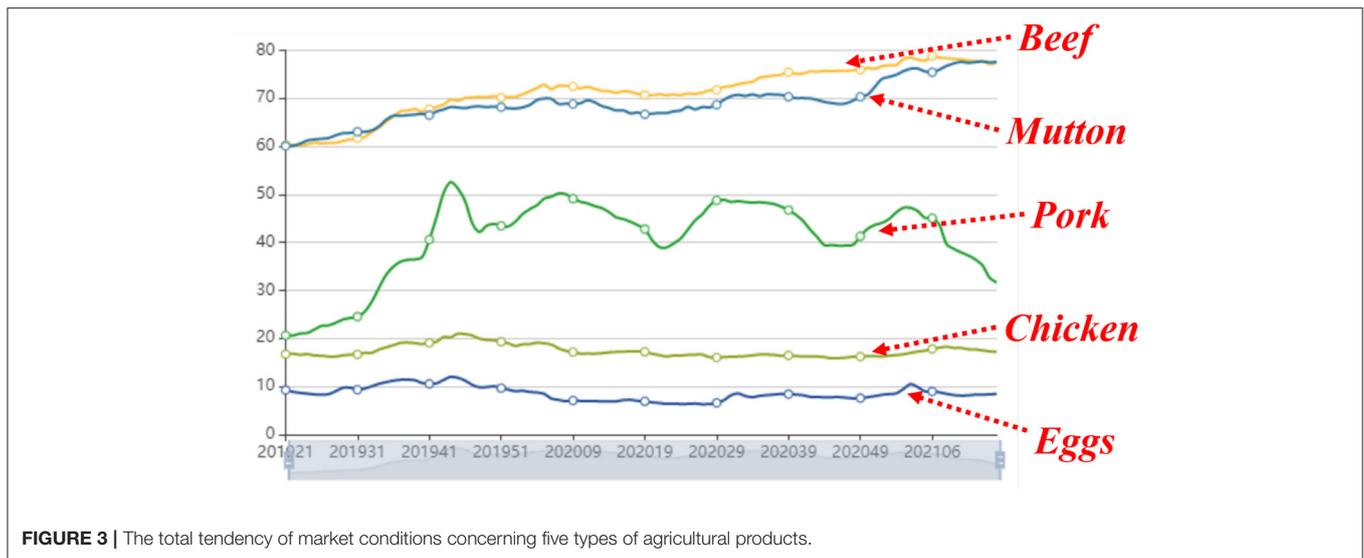
where  $\tilde{E}_{ij}$  is the degree matrix of the  $i$ -th node, and  $\sum_i \tilde{A}_{ij}$  is the number of edges between the  $i$ -th node and other nodes. The final expression of graph convolution operation can be represented as:

$$x \otimes_G g = \theta \left( E^{-\frac{1}{2}} \cdot \tilde{A} \cdot E^{-\frac{1}{2}} \right) x \tag{13}$$

At the  $t$ -th timestamp, the main input is related to the outputs of the previous several timestamps and the relation status among nodes. For the  $i$ -th node, its information state at the  $t$ -th timestamp can be represented as the following formula:

$$S_i^{(t)} = \beta \cdot O_i^{(t)} \cdot W_{S1} + (1 - \beta) \cdot a_i \cdot W_{S2} \tag{14}$$

where  $O_i^{(t)}$  is a vector that records output values at previous several timestamps,  $a_i$  is a vector that records relation status between the  $i$ -th node and other nodes,  $\beta$  is a tuning parameter that adjusts the weight of two parts in Equation (14), and  $W_{S1}$  and  $W_{S2}$  are parameters to be learned. It is widely known that the internal of GCN is the information propagation process, as GCN emphasizes modeling of various dynamic or static relations. Accordingly, the representative vectors for node status can be also



propagated to the following timestamps, which can be expressed as the following formula:

$$S_i^{(t+1)} = \sigma_1 \left[ \tilde{E}^{-\frac{1}{2}} \cdot S_i^{(t)} \cdot \tilde{E}^{-\frac{1}{2}} \cdot W_{S3} + b_{S1} \right] \quad (15)$$

where  $W_{S3}$  and  $b_{S1}$  are parameters, and  $\sigma_1(\cdot)$  is the Reluctant Unit activation function represented as follows:

$$\sigma_1(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (16)$$

Hence, the prediction result at the  $t$ -th timestamp can be calculated as the following formula:

$$V_i^{(t)} = \sigma_1 \left[ S_i^{(t)} \cdot W_{S4} + b_{S2} \right] \quad (17)$$

where  $W_{S4}$  and  $b_{S2}$  are parameters. As for training, the following optimization objective can be formulated to search for the optimal parameters:

$$\min \sum_{i=1}^N \sum_{t=1}^M \left[ \|V^{(t)} - \hat{V}^{(t)}\| + p \cdot \|\Theta\|_F^2 \right] \quad (18)$$

where  $V^{(t)}$  is the predicted result at the  $t$ -th timestamp,  $\hat{V}^{(t)}$  is the real result at the  $t$ -th timestamp,  $\Theta$  is the parameter set,  $p$  is the penalty parameter, and  $\|\cdot\|_F^2$  is the Frobenius norm. Finally, the Adam optimization algorithm can be utilized to search optimal solution for Equation 18.

### 4. SETTING OF CASE STUDY

To evaluate the GNN-APM designed in this study, real-world data is used here to set a simulative analysis situation. The real-world data was crawled from the official website of the Ministry

of agriculture of China <sup>1</sup>, including the market data of several key agricultural products from April 2019 to March 2021. The data demonstrates the average market price of agricultural products wholesale and contains five types of agricultural products. Market condition data for these agricultural products are updated once a week, and there are totally 96 weeks of data concerning the five types of agricultural products. Of all the five types of agricultural products, there are 10 kinds of node combinations, indicating 10 kinds of edges among these nodes. In other words, the adjacency matrix in this situation is a matrix with five columns and five lines, representing relation status between every two combinations of nodes.

During each round of simulations, these 10 groups of relations are randomly generated according to a Gaussian distribution whose mean is set to 0.5 and variance is set to 0.05. As for the ratio between training data and testing data, it is majorly set to 7:3 and 6:4. To quantify the error between prediction results and real results, two typical metrics are selected. They are mean absolute error (MAE) and root mean squared error (RMSE). In addition, two general prediction models are employed as baseline methods for comparison. Different from the designed GNN-APM in this study, the two baseline methods never take internal correlations among nodes into consideration. The two methods are the long short-term memory (LSTM) model and the multi-layer perceptron (MLP) model.

### 5. RESULTS AND ANALYSIS

In this study, the whole simulative experiments are composed of three parts. First, the fluctuation tendency of the involved five agricultural product types is visualized using a curve diagram. Second, the prediction efficiency of the GNN-APM on five objects is compared with two baseline methods. Third, the

<sup>1</sup><http://www.moa.gov.cn/>

**TABLE 1** | Mean absolute error (MAE) results when proportion of training data ranges from 50 to 70% and learning rate ranges from 0.001 to 0.002.

Setting	Learning rate: 0.001			Learning rate: 0.002		
	50%	60%	70%	50%	60%	70%
MLP	1.189	1.047	1.016	1.232	1.175	1.143
LSTM	1.116	1.052	1.044	1.139	1.091	1.107
GNN-APM	<b>0.964</b>	<b>0.927</b>	<b>0.875</b>	<b>0.989</b>	<b>0.903</b>	<b>1.016</b>

The bold values correspond to results of the proposed GNN-APM method.

**TABLE 2** | Root mean squared error (RMSE) results when the proportion of training data ranges from 50 to 70% and the learning rate ranges from 0.001 to 0.002.

Setting	Learning rate: 0.001			Learning rate: 0.002		
	50%	60%	70%	50%	60%	70%
MLP	1.664	1.458	1.373	1.761	1.663	1.589
LSTM	1.512	1.408	1.387	1.553	1.476	1.481
GNN-APM	<b>1.289</b>	<b>1.245</b>	<b>1.177</b>	<b>1.343</b>	<b>1.184</b>	<b>1.353</b>

The bold values correspond to results of the proposed GNN-APM method.

robustness of the GNN-APM is tested by changing several parameter combinations.

## 5.1. Data Pre-processing

**Figure 3** visualizes the total tendency of market conditions for five types of agricultural products. It can be observed from the figure that five curves from the bottom to the top correspond to eggs, chicken, pork, mutton, and beef. During a long period about nearly 2 years, eggs and chicken remain relatively stable, mutton and beef show an ascending tendency, and the pork fluctuates frequently. These five types of agricultural products possess their own fluctuation tendencies and satisfy the assumption of diversity. And it can be also seen that the fluctuation tendency of pork has some effect on the other four types of agricultural products. Thus, the assumption that correlations exist among these types of agricultural products is reasonable.

## 5.2. Performance Assessment

**Tables 1, 2**, respectively give MAE results and RMSE results of these experimental methods when the proportion of training data ranges from 50 to 70% and the learning rate ranges from 0.001 to 0.002. Each of them has five lines and seven rows. The first two lines list the experimental setting, and the other lines present the experimental results of three methods. The first row lists three experimental methods, the second to the fourth rows present results under a learning rate of 0.001, the fifth to the seventh rows present results under a learning rate of 0.002. It can be observed from the two tables that MAE results and RMSE results of GNN-APM are below two other baseline methods, regardless of the proportion of training data and learning rate. This demonstrates that the performance of the GNN-APM is better than baseline methods.

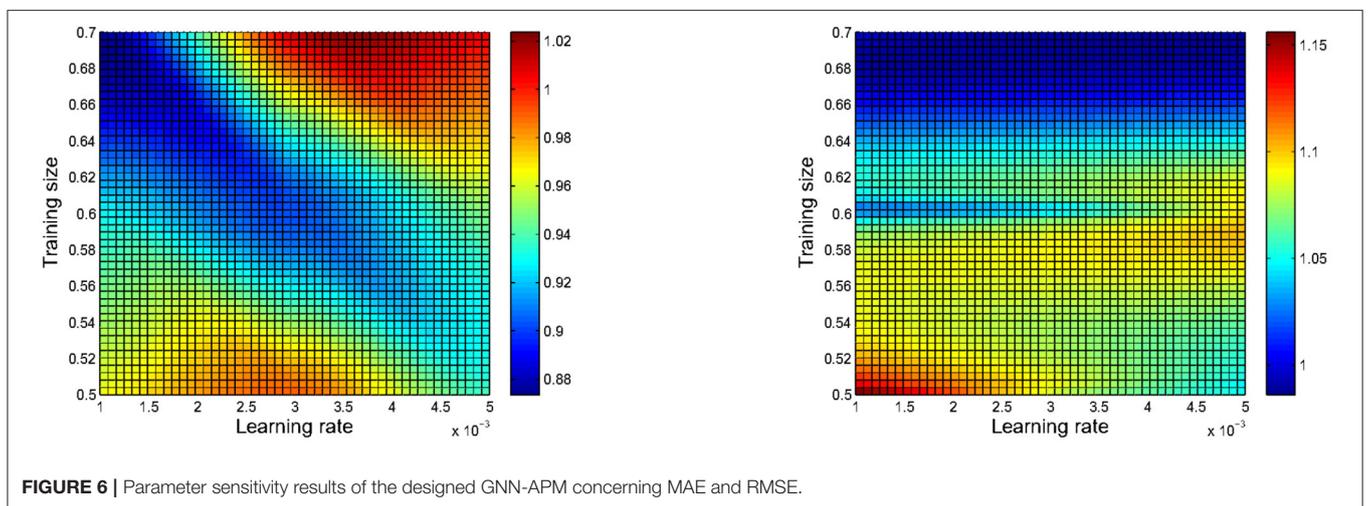
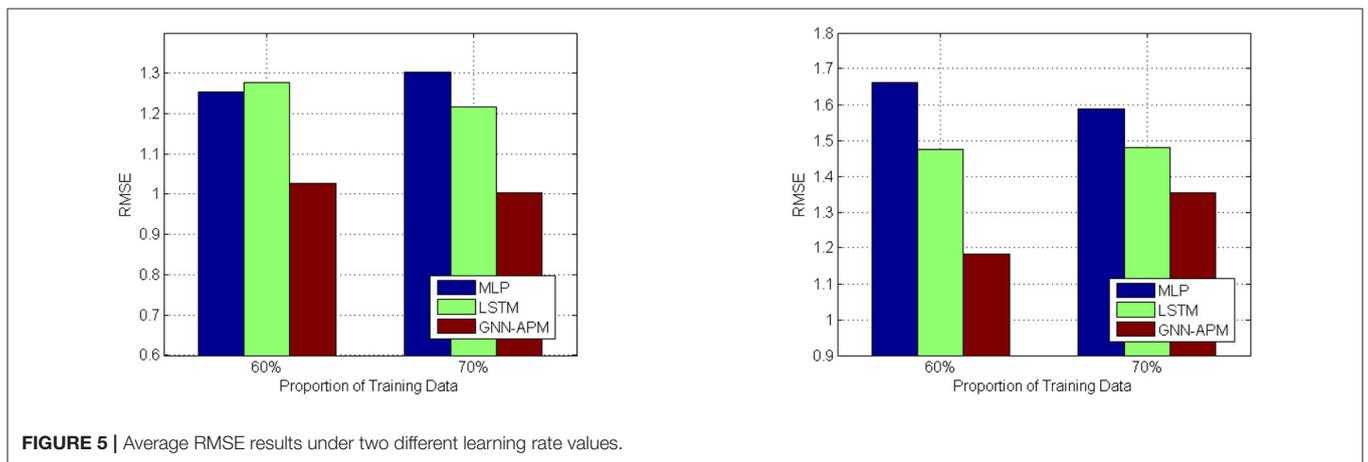
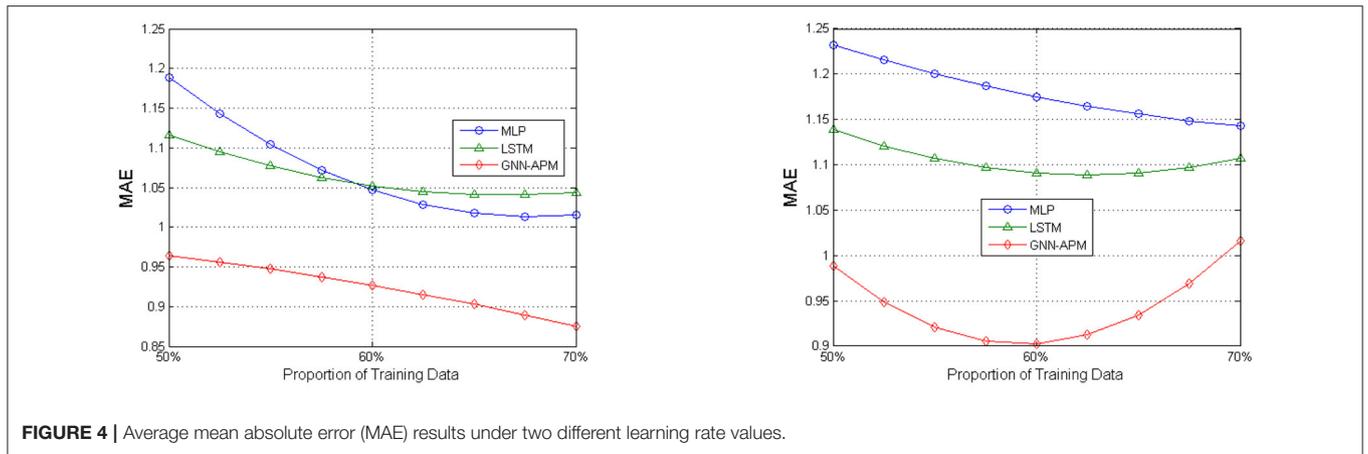
**Figures 4, 5** illustrates prediction efficiency with respect to using two metrics: MAE and RMSE. As there are totally five types

of agricultural products involved, the MAE results and RMSE results are obtained as the mean value of prediction results on the five types. This figure has two subfigures, corresponding to MAE results and RMSE results. Among them, **Figure 4** is the curve diagram and **Figure 5** is the bar diagram. For the former, the X-axis demonstrates three kinds of training sizes and the Y-axis demonstrates values of MAE results. For the latter, only the two most typical training sizes are utilized for evaluation. Thus, it has two clusters of bars, corresponding to RMSE results under two training sizes. It is clearly observed that the GNN-APM is always endowed with better prediction efficiency compared with baseline methods. To sum up, this group of simulative experiments well demonstrates the good performance of the designed GNN-APM.

In order to visualize the tendency of MAE results and RMSE results under different experimental settings, some of the results are illustrated with the use of curve diagrams or bar diagrams. **Figure 4** illustrates the MAE results of three methods under two different learning rate values: 0.001 and 0.002. It has two subfigures that correspond to results about two learning rate values. In each subfigure, the X-axis denotes the proportion of training data ranging from 50 to 70%, and the Y-axis denotes values of MAE results. **Figure 5** illustrates RMSE results of three methods when the training data size is set to 60 and 70%. This is because it can be seen from previous experiments that result under the two training data sizes are relatively better. It has two subfigures that correspond to two learning rate values. In each subfigure, the X-axis denotes two training sizes, and the Y-axis denotes values of RMSE results. It can be observed from these figures that values of GNN-APM are obviously below the other two methods and that values show descending tendency when the proportion of training data increases. Such results demonstrate the improvement process of methods with being trained more sufficiently. These figures show a better performance tendency of GNN-APM compared with two other baseline methods.

## 5.3. Parameter Sensitivity

Besides, it is also expected to explore parameter sensitivity of the GNN-APM, and relevant simulative results are illustrated in **Figure 6**. During this group of experiments, the GNN-APM is not compared with baseline methods and just performance of itself is explored. **Figure 6** has two subfigures, corresponding to sensitivity results using two different metrics: MAE and RMSE. Inside each subfigure, the X-axis denotes the change of learning rate, and the Y-axis denotes the change of training size. In the middle square area, the color depth indicates the different values of evaluation metrics. As the two subfigures are heatmaps, the color depth degree inside figures is able to indicate values of metrics. Each subfigure includes a squared area, gentle color change inside it indicates that the performance of GNN-APM fluctuates not heavily. It can be objectively found that color fluctuation in both two subfigures seems quite gentle, revealing that the GNN-APM is not susceptible to parameter change. In other words, the GNN-APM is always able to remain stable, no matter how the key parameters change. This group of simulative results well prove that the GNN-APM possesses proper robustness.



## 6. CONCLUSION

Agriculture has been viewed as the most fundamental industry since ancient times. Nowadays, E-commerce is an important

sales channel of agricultural products. To better manage and schedule the supply of agricultural products, dynamic price prediction for agricultural products in the E-commerce market is of great significance. To overcome the shortcomings of

existing research studies, this article proposes a deep learning-based price prediction model for agricultural products in the E-commerce market. In particular, the most typical GCN is utilized to establish a time-series prediction model for the dynamic price of agricultural products. In addition, the whole simulative experiments are composed of three parts. First, the fluctuation tendency of the involved five agricultural product types is visualized using a curve diagram. Second, the prediction efficiency of the GNN-APM on five objects is compared with two baseline methods. Third, the robustness of the GNN-APM is tested by changing several parameter combinations.

Nowadays, data mining and data management for many industries are gradually approaching the application of the Internet of Things (IoT), yielding such as mobile IoT (36, 37), financial IoT, medical IoT (38), cloud-assisted IoT (39), vehicular IoT (40), and industrial IoT (41, 42). As is known to all, the IoT is a kind of effective tool or platform to integrate multi-domain data and schedule business flows. To realize more effective scheduling management of the agricultural product market, designing an integrated microservice IoT platform that is embedded with robust artificial intelligence algorithms (43), is in urgent demand to deal with many disturbing issues in various industries. Thus, for future outlook, the authors plan to deeply investigate optimal scheduling and management schemes for the agricultural product market with the use of novel IoT-related technologies.

## REFERENCES

- Xiong H, Jin C, Alazab M, Yeh KH, Wang H, Gadekallu TRR, et al. On the design of blockchain-based ECDSA with fault-tolerant batch verification protocol for blockchain-enabled IoMT. *IEEE J Biomed Health Inform.* (2021). doi: 10.1109/JBHI.2021.3112693
- Yu K, Tan L, Yang C, Choo KKR, Bashir AK, Rodrigues JJPC, et al. A blockchain-based shamir's threshold cryptography scheme for data protection in industrial internet of things settings. *IEEE Internet Things J.* (2021). doi: 10.1109/JIOT.2021.3125190
- Wang W, Qiu C, Yin Z, Srivastava G, Gadekallu TR, Alsolami F, et al. Blockchain and PUF-based lightweight authentication protocol for wireless medical sensor networks. *IEEE Internet Things J.* (2021). doi: 10.1109/JIOT.2021.3117762
- Yu K, Tan L, Shang X, Huang J, Srivastava G, Chatterjee P. Efficient and privacy-preserving medical research support platform against COVID-19: a blockchain-based approach. *IEEE Consumer Electron Mag.* (2021) 10:111–20. doi: 10.1109/MCE.2020.3035520
- Zeng W, Guo Z, Yu K, Shen Y, Bashir AK, Al-Otaibi YD. Data-driven management for fuzzy sewage treatment processes using hybrid neural computing. *Neural Comput Appl.* (2021). doi: 10.1007/s00521-020-05655-3
- Wang W, Fida MH, Lian Z, Yin Z, Pham QV, Gadekallu TR, et al. Secure-Enhanced federated learning for ai-empowered electric vehicle energy prediction. *IEEE Consumer Electron Mag.* (2021). doi: 10.1109/MCE.2021.3116917
- Guo Z, Wang H. A deep graph neural network-based mechanism for social recommendations. *IEEE Trans Ind Inform.* (2021) 17:2776–83. doi: 10.1109/TII.2020.2986316
- Guo Z, Yu K, Li Y, Srivastava G, Lin JCW. Deep learning-embedded social internet of things for ambiguity-aware social recommendations. *IEEE Trans Netw Sci Eng.* (2021). doi: 10.1109/TNSE.2021.3049262
- Deepa N, Durai R, Senthil KN, Srinivasan K, Chang CY, Bashir AK. An efficient ensemble VTOPES multi-criteria decision-making model for sustainable sugarcane farms. *Sustainability.* (2019) 11:4288. doi: 10.3390/su11164288
- Li H, Chen X, Guo Z, Xu J, Shen Y, Gao X. Data-driven peer-to-peer blockchain framework for water consumption management. *Peer-to-Peer Netw Appl.* (2021) 14:2887–900. doi: 10.1007/s12083-021-01121-6
- Meng D, Xiao Y, Guo Z, Jolfaei A, Qin L, Lu X, et al. A data-driven intelligent planning model for UAVs routing networks in mobile internet of things. *Comput Commun.* (2021) 179:231–41. doi: 10.1016/j.comcom.2021.08.014
- Elavarasan D, Vincent DR, Sharma V, Zomaya AY, Srinivasan K. Forecasting yield by integrating agrarian factors and machine learning models: a survey. *Comput Electron Agric.* (2018) 155:24. doi: 10.1016/j.compag.2018.10.024
- Ahmmedi P, Rahimian M, Movahed RG. Theory of planned behavior to predict consumer behavior in using products irrigated with purified wastewater in Iran consumer. *J Cleaner Product.* (2021) 296:126359. doi: 10.1016/j.jclepro.2021.126359
- Fan P, Wang Y, Xu N. Value added mechanism and organisational model optimisation of agricultural products circulation value chain from the perspective of game theory. *Acta Agric Scand B Soil Plant Sci.* (2021) 71:215–23. doi: 10.1080/09064710.2021.1879927
- Assa H, Sharifi H, Lyons A. An examination of the role of price insurance products in stimulating investment in agriculture supply chains for sustained productivity. *Eur J Oper Res.* (2021) 288:918–34. doi: 10.1016/j.ejor.2020.06.030
- Ray P. Agricultural supply chain risk management under price and demand uncertainty. *Int J Syst Dyn Appl.* (2021) 10:17–32. doi: 10.4018/IJSDA.2021040102
- Yang B, Chang M, Wang K, Zhang R. The calendar effect of price-reduction auction of online agricultural products. *Int J Serv Technol Manag.* (2021) 27:229–47. doi: 10.1504/IJSTM.2021.115167
- Zhao H. Futures price prediction of agricultural products based on machine learning. *Neural Comput Appl.* (2021) 33:837–50. doi: 10.1007/s00521-020-05250-6

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: <http://www.moa.gov.cn/>.

## AUTHOR CONTRIBUTIONS

WY contributed significantly to theoretical analysis and manuscript preparation. ZZ performed the experiments and handled funding details. QZ contributed to the conception of the study and model formulation. GZ performed the data analyses and data visualization. QH provided many promising insights during the revision process. QL helped perform the analysis with constructive discussions. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported in part by Beijing Natural Science Foundation (No. 4202014), Natural Science Foundation of China (61873027), in part by Humanity and Social Science Youth Foundation of Ministry of Education of China (No. 20YJCZH229), in part by the Open Project Program of National Engineering Laboratory for Agri-Product Quality Traceability (AQT-2020-YB8), and in part by Hubei Natural Science Foundation (No. 2021CFB156).

19. Zhang D, Chen S, Liwen L, Xia Q. Forecasting agricultural commodity prices using model selection framework with time series features and forecast horizons. *IEEE Access*. (2020) 8:28197–209. doi: 10.1109/ACCESS.2020.2971591
20. Huang J, Chen Y, He S. The evolutionary analysis of agricultural production transaction under the price subsidy policy. *Int J Inf Syst Supply Chain Manag*. (2020) 13:73–97. doi: 10.4018/IJISSCM.2020010104
21. Yadava TG, Jayanna HS. Improvements in spoken query system to access the agricultural commodity prices and weather information in Kannada language/dialects. *J Intell Syst*. (2020) 29:664–87. doi: 10.1515/jisys-2018-0120
22. Zheng G, Zhang H, Han J, Zhuang C, Xi L. The research on agricultural product price forecasting service based on combination model. In: *13th IEEE International Conference on Cloud Computing, CLOUD 2020 Virtual Event, 18–24 October 2020*. Beijing: IEEE (2020). p. 4–9.
23. Guo H, Woodruff A, Yadav A. Improving lives of indebted farmers using deep learning: predicting agricultural produce prices using convolutional neural networks. In: *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020 The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020 The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020 New York, NY, USA, February 7–12, 2020 AAAI Press*. New York, NY: AAAI Press (2020). p. 13294–9. doi: 10.1609/aaai.v34i08.7039
24. Chen L, Li Z. Research on the price analysis and prediction method of agricultural products based on logistics information. *Clust Comput*. (2019) 22:14951–57. doi: 10.1007/s10586-018-2462-y
25. Alizamir S, Irvani F, Mamani H. An analysis of price vs. Revenue protection: government subsidies in the agriculture industry. *Manag Sci*. (2019) 65:32–49. doi: 10.1287/mnsc.2017.2927
26. Madaan L, Sharma A, Khandelwal P, Goel S, Singla P, Seth A. Price forecasting & anomaly detection for agricultural commodities in India. In: *Proceedings of the Conference on Computing & Sustainable Societies, COMPASS 2019 Accra, Ghana, July 3–5 2019*. Accra: ACM (2019). p. 52–64.
27. Silva J, Gaitán-Angulo M, Borré JR, Ayarza LPL, Lezama OBP, del Carmen Martínez Galán Z, et al. An early warning method for agricultural products price spike based on artificial neural networks prediction. In: *Computational Science - ICCS 2019 - 19th International Conference, Faro, Portugal, June 12–14, 2019 Proceedings, Part II. Vol. 11537 of Lecture Notes in Computer Science*. Faro: Springer (2019). p. 622–32.
28. Deng Y, Zhu Y, Zhang Q, Cheng X. Hadoop + spark platform based on big data system design of agricultural product price analysis and prediction by holtwinters. In: *Data Science - 5th International Conference of Pioneering Computer Scientists, Engineers and Educators, ICPCSEE 2019 Guilin, China, September 20–23, 2019 Proceedings, Part I. Vol. 1058 of Communications in Computer and Information Science*. Springer (2019). p. 103–23.
29. Dharavath R, Khosla E. Seasonal ARIMA to forecast fruits and vegetable agricultural prices. In: *IEEE International Symposium on Smart Electronic Systems, iSES 2019 (Formerly iNiS), Rourkela, India, December 16–18, 2019*. Rourkela: IEEE (2019). p. 47–52.
30. Chen Q, Lin X, Zhong Y, Xie Z. Price prediction of agricultural products based on wavelet analysis-LSTM. In: *2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking, ISPA/BDCLOUD/SocialCom/SustainCom*. Xiamen: IEEE (2019). p. 984–90.
31. Xiong T, Li C, Bao Y. Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: evidence from the vegetable market in China. *Neurocomputing*. (2018) 275:2831–44. doi: 10.1016/j.neucom.2017.11.053
32. Kurumatani K. Time series prediction of agricultural products price based on time alignment of recurrent neural networks. In: Wani MA, Kantardzic MM, Mouchaweh MS, Gama J, Lughofer E, editors. *17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018 Orlando, FL, USA, December 17–20*. Orlando, FL: IEEE (2018). p. 81–8.
33. Guo Z, Yu K, Jolfaei A, Bashir AK, Almagrabi AO, Kumar N. A fuzzy detection system for rumors through explainable adaptive learning. *IEEE Trans Fuzzy Syst*. (2021) 29:3650–64. doi: 10.1109/TFUZZ.2021.3052109
34. Yu K, Tan L, Lin L, Chen X, Yi Z, Sato T. Deep-Learning-Empowered breast cancer auxiliary diagnosis for 5GB remote e-health. *IEEE Wirel Commun*. (2021) 28:54–61. doi: 10.1109/MWC.001.2000374
35. Guo Z, Shen Y, Bashir AK, Yu K, wei Lin JC. Graph embedding-based intelligent industrial decision for complex sewage treatment processes. *Int J Intel Syst*. (2021). doi: 10.1002/int.22540
36. Zhang Q, Yu K, Guo Z, Garg S, Rodrigues JJPC, Hassan MM, et al. Graph neural networks-driven traffic forecasting for connected internet of vehicles. *IEEE Trans Netw Sci Eng*. (2021). doi: 10.1109/TNSE.2021.3126830
37. Yu K, Guo Z, Shen Y, Wang W, Lin JCW, Sato T. Secure artificial intelligence of things for implicit group recommendations. *IEEE Internet Things J*. (2021) 9:2698–707. doi: 10.1109/JIOT.2021.3079574
38. Sun Y, Liu J, Yu K, Alazab M, Lin K. PMRSS: privacy-preserving medical record searching scheme for intelligent diagnosis in IoT healthcare. *IEEE Trans Ind Inform*. (2021) 18:1981–90. doi: 10.1109/TII.2021.3070544
39. Wang T, Quan Y, Shen XS, Gadekallu TR, Wang W, Dev K. A privacy-enhanced retrieval technology for the cloud-assisted internet of things. *IEEE Trans Ind Inform*. (2021). doi: 10.1109/TII.2021.3103547
40. Feng C, Yu K, Aloqaily M, Alazab M, Lv Z, Mumtaz S. Attribute-based encryption with parallel outsourced decryption for edge intelligent IoT. *IEEE Trans Veh Technol*. (2020) 69:13784–95. doi: 10.1109/TVT.2020.3027568
41. Li H, Yu K, Liu B, Feng C, Qin Z, Srivastava G. An efficient ciphertext-policy weighted attribute-based encryption for the internet of health things. *IEEE J Biomed Health Inform*. (2021). doi: 10.1109/JBHI.2021.3075995
42. Wang W, Xu H, Alazab M, Gadekallu TR, Han Z, Su C. Blockchain-Based reliable and efficient certificateless signature for IIoT devices. *IEEE Trans Ind Inform*. (2021). doi: 10.1109/TII.2021.3084753
43. Tan L, Yu K, Shi N, Yang C, Wei W, Lu H. Towards secure and privacy-preserving data sharing for COVID-19 medical records: a blockchain-empowered approach. *IEEE Trans Netw Sci Eng*. (2021) 9:271–81. doi: 10.1109/TNSE.2021.3101842

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Yan, Zhang, Zhang, Zhang, Hua and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.