

Article

# A Heterogeneous Graph Enhanced LSTM Network for Hog Price Prediction Using Online Discussion

Kai Ye , Yangheran Piao , Kun Zhao and Xiaohui Cui \* 

Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, Wuhan 430072, China; kai\_ye@whu.edu.cn (K.Y.); CDra\_90n@whu.edu.cn (Y.P.); zhaokun@whu.edu.cn (K.Z.)

\* Correspondence: xcui@whu.edu.cn

**Abstract:** Forecasting the prices of hogs has always been a popular field of research. Such information has played an essential role in decision-making for farmers, consumers, corporations, and governments. It is hard to predict hog prices because too many factors can influence them. Some of the factors are easy to quantify, but some are not. Capturing the characteristics behind the price data is also tricky considering their non-linear and non-stationary nature. To address these difficulties, we propose Heterogeneous Graph-enhanced LSTM (HGLTSM), which is a method that predicts weekly hog price. In this paper, we first extract the historical prices of necessary agricultural products in recent years. Then, we utilize discussions from the online professional community to build heterogeneous graphs. These graphs have rich information of both discussions and the engaged users. Finally, we construct HGLTSM to make the prediction. The experimental results demonstrate that forum discussions are beneficial to hog price prediction. Moreover, our method exhibits a better performance than existing methods.



check for updates

**Citation:** Ye, K.; Piao, Y.; Zhao, K.; Cui, X. A Heterogeneous Graph Enhanced LSTM Network for Hog Price Prediction Using Online Discussion. *Agriculture* **2021**, *11*, 359. <https://doi.org/10.3390/agriculture11040359>

Academic Editor: Gioacchino Pappalardo

Received: 23 March 2021  
Accepted: 12 April 2021  
Published: 16 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** agricultural price; price forecast; hog price; heterogeneous graph

## 1. Introduction

Livestock is widely known as an important part of agriculture. According to the Food and Agriculture Organization (Food and Agriculture Organization of the United Nations (Available online: <http://www.fao.org/>, accessed on 20 January 2021), pork production plays an important role in meat production. Consequently, the production and consumption of agricultural products like pork affect many countries' economies and livelihoods around the world. Given the close connection between pork and people's lives, stable pork prices are important for economic and social stability. The prices of pork and hog not only influence the global agriculture market, but also government policies [1,2], water industry [3], food markets [4], oil prices [5] and other industries [6]. An accurate prediction of hog prices will provide favorable conditions for farmers, consumers, the government, and other participants. Government officials and other regulators can better understand the market and make policies accordingly. Consumers and farmers can make business adjustments to maximize their interest. Therefore, it is of great significance to capture the characteristics of hog prices and make accurate predictions.

In previous research, efforts have been made to predict future prices based on various historical factors, such as historical prices, climate change, seasonal factors, agricultural calamities, and other economic effects. However, many of these factors, such as capital operation, policy, and disease, are difficult to quantify to make a prediction, making it hard to choose the influencing factors.

To address this issue, we explore the influence of forum discussions on hog price prediction. As forum discussions contain people's analysis and reflect their expectations towards this topic, we assume that they include and interpret many factors, like the influence of consumer preferences, political events and other factors that are difficult to

quantify. In fact, including textual information such as news articles for classification or prediction problems is not rare, especially for stock price prediction [7] and many NLP tasks. It is very likely that the forum discussions can enhance hog price prediction. Besides this, hog prices follow a non-linear and non-stationary time-series. For this time-series prediction task, researchers seek statistical methods and later machine learning methods. To extend this research line of applying deep neural networks to extract the necessary features, we further construct heterogeneous graphs to capture the representations of online discussions, enhancing hog price prediction.

In this study, we explore the influence of forum discussions on hog price prediction and propose a method that predicts the weekly hog prices. We extract historical prices of hog, maize and bean, as well as forum discussions for hog price prediction. It has been proved that bean and maize prices can largely influence hog price [8]. More importantly, the historical prices of hog and maize are easy to quantify and acquire. After obtaining representations of forum discussions and price series, our HGLSTM will combine price features and discussion information to forecast hog prices.

Our contributions are summarized as follows:

- As far as we are concerned, this is the first study to make use of discussion information, acquired from the online professional pig community, for hog price prediction, and prove it to be effective;
- Due to our limited time and effort, we find no other research to deeply integrate discussion information and prices series based on heterogeneous graph for hog price forecast;
- We propose a heterogeneous graph-enhanced LSTM network (HGLSTM) and conduct extensive experiments to prove its effectiveness. Our experiments show that it outperforms state-of-the-art models.

This paper is organized in the following manner: In Section 2, we introduce some important related works. In Section 3, we discuss how our model is constructed and give necessary explanations. In Section 4, we illustrate the experimental design and results. We also give a brief analysis of the results in this section. Finally, in Section 5, we present our conclusions and insight into this task's possible future direction.

## 2. Related Work

In this section, we introduce some critical studies related to the price forecasting of agricultural commodities. Price forecasting is often regarded as a time-series prediction problem. Thus, traditional statistical and deep learning methods have been commonly used for this. Significant studies of natural language processing and deep neural network concerning our method will also be discussed.

### 2.1. Price Forecasting Using Statistical Methods

Regression methods, like the autoregressive integrated moving average (ARIMA), generalized ARIMA, and seasonal ARIMA, are often used to solve this type of task. They are usually classified as traditional statistical methods. The ARIMA model is exploited by researchers [9–12] for agricultural price prediction. When studying cocoa bean price forecasting, Assis and Remali [13] tried to figure out the best method in various time-series prediction models. Their experiments showed that the generalized ARIMA model achieved the best performance. Adanacioglu and Yercan [14] applied seasonal ARIMA to tomato price forecasting in Turkey. In an attempt to solve corn prices forecasting, Gu et al. [15] proposed a multivariate linear regression model. They tried to model the effect of supply and demand, but their model's performance is still not very desirable due to drastic changes in the corn market. BV and Dakshayini [16] tried to predict the prices and demand of tomatoes. Their study compared the performance of Holt Winter's model and other benchmark models, such as simple (multiple) linear regression. Their experiments presented huge variations between targets and predictions. They also concluded that seasonality was an influencing factor because Holt Winter's model, which considers seasonality, achieved the best performance.

Statistical methods show good performance on linear price series, but their performance drops drastically when faced with non-linear and non-stationary price series.

## 2.2. Price Forecasting Using Machine Learning Methods

Thanks to the rapid development of machine learning and deep learning algorithms, many researchers have developed new approaches to solve time-series forecasting problems. These new methods can extract hidden features from price series and, as a result, show a much better performance than traditional statistical models.

A back-propagation neural network proposed by Minghua et al. [17] is applied to the price forecasting of agricultural products. They conducted extensive experiments and found their proposed artificial neural network's superiority against a statistical method. Other researchers, such as Nasira and Hemegeetha [18], also exploit back-propagation neural network (BPNN) to predict tomatoes' prices. Trying to predict the non-linear garlic price series, Wang et al. [19] proposed a hybrid ARIMA support vector machine (SVM) model. Experimental results showed that this very model surpassed the performance of both single ARIMA and SVM. Hemegeetha and Nasira [20] proposed a radial basis function neural network (RBF) to predict tomato prices. Their model achieved better accuracy than the BPNN model. Using a chaotic neural network, Li et al. [21] found it to be a superior algorithm for weekly egg price prediction than ARIMA.

Many researchers have also made an effort to combine multiple models into a hybrid model. Luo et al. [22] propose three models and a hybrid model to forecast *Lentinus edodes* mushroom prices of Beijing. Their integrated model combines BPNN, RBF neural network, and genetic-algorithm-based neural network to achieve the best performance. Zhang et al. [23] proposes a quantile regression-based RBF (QR-RBF) neural network model to predict soybean prices in China. In the process of model optimization, they apply a gradient descent with GA to improve performance. Their experimental results align with previous studies [24,25].

Other researchers seek to preprocess the price series before feeding them into the model. Xiong, Li and Bao [26] first use the STL-based method to decompose the price series to predict cabbage, hot pepper, cucumber, kidney bean, and tomato prices. They consider the seasonal characteristics of vegetables and preprocess the time-series price data based on these characteristics. Their experiments prove the effectiveness of their method. To forecast vegetable prices, Li and Zheng [27] proposes a model that integrates an H-P filter and a neural network. Their study's main contribution is that they decompose trend and cyclical components in the price series and recombine prediction values using the H-P filter. Another study [28] aims to forecast five monthly crop prices in the Korean market. They propose the STL-LSTM model, which eliminates high seasonality in vegetable prices. Their model performance has improved a lot by doing so. Following this research line, Liu et al. [29] propose a model that divides hog price series into the trend and cyclical components. They use the most similar sub-series search method to predict them and recombine these components. Finally, with the help of support vector regression, they successfully forecast the hog prices.

Researchers also exploit other information to help forecast the price series. Yoo et al. [30] makes use of climate factors and production information along with trends and seasonality of price data for prediction. They aim to forecast the prices of Korean cabbage and achieve good results. Chen et al. [31] aims to predict cabbage prices in the Chinese market. They propose a wavelet analysis-based LSTM model. The wavelet method that removes noise from the price series, therefore, helps improve model performance.

As we can see, most researchers using deep neural networks include LSTM in their model, which is not surprising because LSTM has shown superiority in dealing with series data. With the development of the attention mechanism of Bahdanau et al. [32], researchers began to apply it to their model. The attention mechanism can assign weights for different input vectors, thus calculating each vector's importance value. There are many variants of attention, suggesting that the structure is very flexible and can be combined with many

existing models. Consequently, it has applications in various fields, such as classification, recommendation, regression, and price prediction.

Qin et al. [33] proposes a dual-stage attention-based recurrent neural network for stock price forecast. Feature attention and temporal attention are used in their model. Attention structure also helps to explain the correlations between input vectors and outputs. Ran et al. [34] addresses travel time prediction by an attention-based LSTM. The attention structure assigns different weights to different features, thus improving model efficiency. Li et al. [35] proposed evolutionary attention-based LSTM model explains the correlations between local features in time steps. Aiming to solve financial time series prediction, Zhang et al. [36] designs attention-based LSTM and addresses a long-term dependence issue.

We summarize the above literature review in the following table (Table 1):

**Table 1.** Summary of the literature review.

Model	Purpose	Input
ARIMA [13]	Cocoa bean	Price
Seasonal ARIMA [14]	Tomato	Price
Multivariate linear regression [15]	Corn	Price, production
BPNN [18]	Tomato	Price
ARIMA-SVM [19]	Garlic	Price
RBF [20]	Tomato	Price
Hybrid model of BPNN, RBF and GA [22]	Mushroom	Price
QR-RBF [23]	Soybean	Price, import/Output, consumer index, money supply
STL-LSTM [28]	Crop	Price, climate, trading volumes
Similar Sub-Series Search and SVM [29]	Hog	Price
Wavelet analysis based LSTM [31]	Cabbage	Price
Dual-stage attention based RNN [33]	Stock price	Price
Attention-based LSTM [34]	Travel time	time
STL-ATTLSTM [37]	Vegetable prices	Price, weather, market trading volumes

### 2.3. LSTM

Long short-term memory (LSTM) is a type of recurrent neural network (RNN). It is proposed by Hochreiter and Schmidhuber [38] to solve long-term dependency and gradient vanishing problems. An LSTM cell usually consists of an input gate, an output gate, a forget gate and a cell state. The structure of a LSTM cell is shown below (Figure 1):

As shown in Figure 1, for each element in the input sequence,  $h_t$  (the hidden state at time  $t$ ) is computed via the following functions

$$\begin{aligned}
 i_t &= \sigma(W^i x_t + U^i h_{t-1} + b^i) \\
 f_t &= \sigma(W^f x_t + U^f h_{t-1} + b^f) \\
 \tilde{C}_t &= \tanh(W^{\tilde{C}} x_t + U^{\tilde{C}} h_{t-1} + b^{\tilde{C}}) \\
 o_t &= \sigma(W^o x_t + U^o h_{t-1} + b^o) \\
 c_t &= \tilde{C}_t \cdot i_t + c_{t-1} \cdot f_t \\
 h_t &= o_t \cdot \tanh c_t
 \end{aligned} \tag{1}$$

where  $h_t$  is the hidden state at time  $t$ ,  $c_t$  is the cell state at time  $t$ ,  $x_t$  is the input at time  $t$ ,  $h_{t-1}$  is the hidden state at time  $t - 1$  or the initial hidden state, and  $i_t, f_t, \tilde{C}_t, o_t$  are the input, forget, cell, and output gates, respectively.  $\sigma$  is the sigmoid function and  $\cdot$  denotes element-wise matrix multiplication.

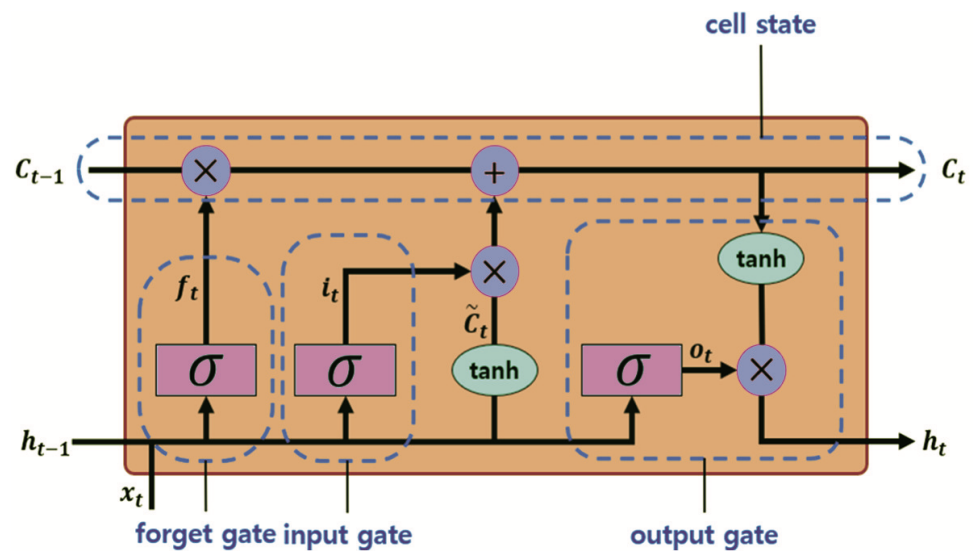


Figure 1. LSTM cell.

LSTM networks are well-suited to classifying, processing, and making predictions based on time series data. A lot of research [28,34,36,39,40] into agricultural price prediction have demonstrated the effectiveness of LSTM in dealing with prices series. Therefore, in this paper, we decide to follow this line of research by exploiting LSTM network to process price series.

### 3. Materials and Methods

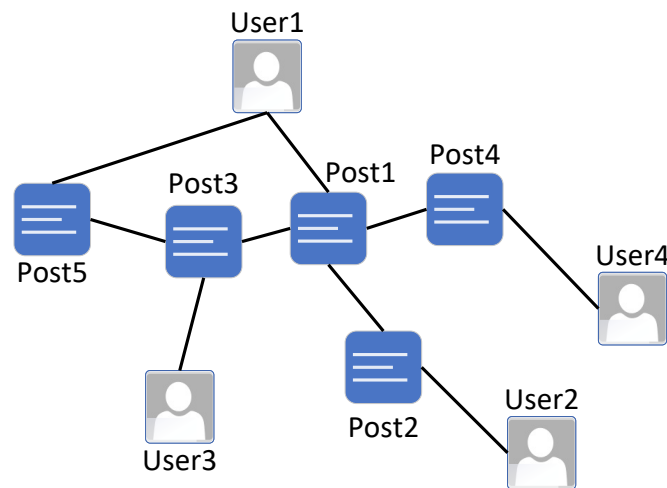
#### 3.1. Problem Statement

Let  $\mathcal{P} = \{p_1, p_2, \dots, p_{|\mathcal{P}|}\}$  be a thread consisting of a set of posts,  $\mathcal{U} = \{u_1, u_2, \dots, u_{|\mathcal{U}|}\}$  be a group of users participating in this thread uploading at least one post, where  $|\mathcal{P}|$  denotes the number of all posts involved and  $|\mathcal{U}|$  denotes the number of all users involved in this thread.

To make the best use of the discussion network and capture the user-enhanced semantic features, we construct the heterogeneous graph  $G = (V, E)$ , where  $V$  denotes the node set and  $E$  denotes the edge set.  $A \in \{0, 1\}^{|V| \times |V|}$  is the adjacency matrix of graph  $G$ . An example of this heterogeneous graph is shown in Figure 2. Considering the graph’s heterogeneity, there are two types of nodes: users node  $V_u$  and posts node  $V_p$ . Therefore, there are two types of edge: post-user edges  $E_{pu}$  and post-post edges  $E_{pp}$ . The connections between users are not considered in this study because, in a discussion thread, users’ connections are rare, thus contributing little to our goal. Moreover, we treat  $G$  as an undirected graph.

For historical prices, let  $\mathcal{X} = (x_1, x_2, \dots, x_{|\mathcal{X}|})$  be the processed weekly prices, where each  $x_i = (q_1, q_2, q_3)$ , and  $q_1, q_2, q_3$  denotes hog, maize, and bean price, respectively.

We regard this price prediction task as a binary classification problem.  $c \in \{0, 1\}$  denotes the label, where  $c = 1$  means hog price will increase next week and  $c = 0$  represents other situations. So our goal is to train a model  $f(\cdot)$  to predict the label of given input (forum discussions and historical prices).

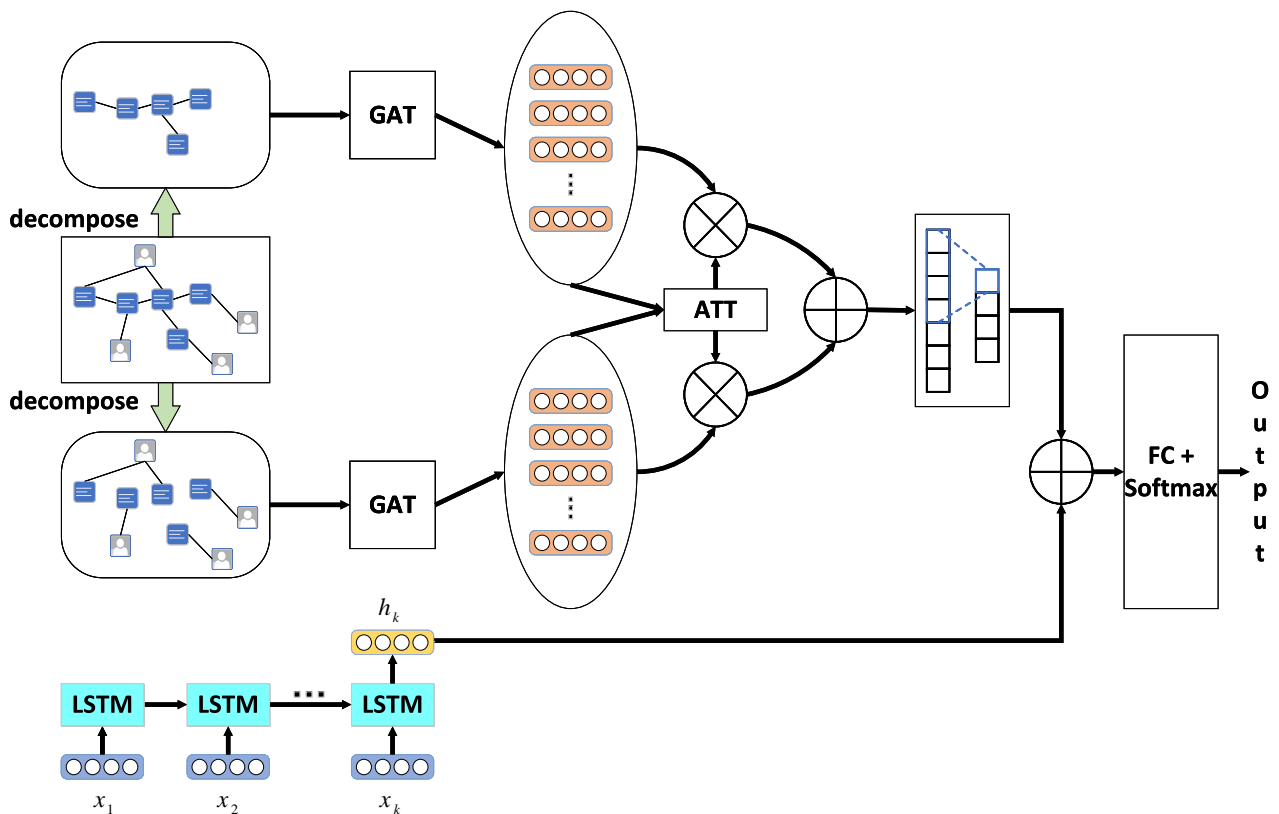


**Figure 2.** This is an example of the discussion network we make use of in this paper. Four users and five posts are engaged in this discussion network.

3.2. Overall Structure

In this paper, we propose a forecasting method for hog prices. The overall structure of our proposed method is shown in Figure 3. It will be explained in detail later. For clarity, all steps are presented below:

1. Necessary pre-processing of historical price data and discussion text;
2. Acquire hidden representation of price series via an LSTM network;
3. Construct a heterogeneous graph based on forum discussion network to capture semantic and network features.
4. Integrate the features extracted from the above process and make the prediction.



**Figure 3.** Overall structure of our proposed model.



### 3.3. Pre-Processing of Data

After acquiring raw data from the Internet, we have to do some data cleaning and pre-processing before feeding them into our model. For forum discussions, We first remove stop words and irregular words or expressions. Then, we use the nltk [41] package for tokenization and transform words into vectors with GloVe [42]. For price data, we replace the price’s absolute value with the change of price relative to the previous week. As our price data is weekly, we choose the thread with the most comments every week to make graphs used in later steps. We assume that the more comments a thread contains, the more information we can extract from the discussion, thus helping the price prediction.

### 3.4. Acquiring Hidden Representation of Price Series

Let  $\mathcal{S} = (x_1, x_2, \dots, x_k)$ , where  $k \in [1, |\mathcal{X}|]$ ,  $x_i$  is defined in Section 3.1. As shown in Figure 3, we feed  $\mathcal{S}$  into a one-layer LSTM network

$$(h_1, h_2, \dots, h_k) = LSTM(x_1, x_2, \dots, x_k) \tag{2}$$

and we use the representation of the last hidden state  $h_k$  as the feature of historical prices.

### 3.5. Constructing Heterogeneous Graph Based on Discussion Network

There are two types of relations in our constructed graph. To obtain a global representation combining semantics, propagation, and user information, we decompose the heterogeneous graph into a post-post subgraph and a post-user subgraph based on meta-path post-post and post-user. After decomposition, only one type of relationship is considered for each subgraph. This process is shown in Figure 3.

Then, we feed the subgraphs into GAT [43]. GATs have shown great capacity in capturing the graph structures. Therefore, we choose GAT in our work. We will describe the details here.

The propagation step from the  $l$ -th layer to the  $(l + 1)$ -th layer of GAT is

$$h_i^{(l+1)} = \sigma \left( \sum_{j \in \mathcal{N}(i) \cup \{i\}} \alpha_{ij}^{(l)} W^{(l)} h_j^{(l)} \right) \tag{3}$$

where  $h_i^{(l)} \in \mathbb{R}^d$  is the representation of node  $v_i$  in the  $l$ -th layer.  $W^{(l)}$  is a trainable weight matrix,  $\sigma$  is the ReLU activation function.  $\mathcal{N}(i)$  is the set of one-hop neighbors of node  $v_i$ ,  $v_i$  itself is also included in the set. And the attention coefficients  $\alpha_{ij}^{(l)}$  are computed as

$$\alpha_{ij}^{(l)} = \frac{\exp \left( \text{LeakyReLU} \left( \vec{a}^{(l)T} \left( W^{(l)} h_i^{(l)} \parallel W^{(l)} h_j^{(l)} \right) \right) \right)}{\sum_{k \in \mathcal{N}(i) \cup \{i\}} \exp \left( \text{LeakyReLU} \left( \vec{a}^{(l)T} \left( W^{(l)} h_i^{(l)} \parallel W^{(l)} h_k^{(l)} \right) \right) \right)} \tag{4}$$

$H^{(0)} \in \mathbb{R}^{|V| \times d}$  is the node embedding matrix. To extract the structure information of subgraphs, We reserve the matrix of activations in the  $l$ -th layer  $H^{(l)} \in \mathbb{R}^{|V| \times d}$  for later use.

Now that we have the node embedding matrix of post-post subgraph  $X_{pp}$  and that of post-user subgraph  $X_{pu}$ , after feeding them into GAT, we can obtain node representations (output)  $X'_{pp}$  and  $X'_{pu}$ , respectively.

The decomposed subgraphs contain different information. The post-post subgraph contains the semantic information of text contents and propagation features, while the post-user subgraph primarily contains user features and relations between the user and its post. To acquire a global and complete representation of heterogeneous graphs, we design an attention mechanism to fuse the information in different subgraphs together.

For this part, we have  $X'_{pp}$  and  $X'_{pu}$  as input, we need to calculate the weights of each subgraph  $\beta_{pp}$  and  $\beta_{pu}$

$$(\beta_{pp}, \beta_{pu}) = \text{attention}(X'_{pp}, X'_{pu}) \tag{5}$$

To learn the weights  $\beta_{pp}$  and  $\beta_{pu}$ , we first transform the representation of nodes in subgraphs into higher-level features by applying a linear transformation. Then, we compute an attention score for each node by doing dot product operations between the transformed node representations and a learnable weight vector  $\mathbf{a}$ . Next, we average the attention scores of all nodes in the subgraph and use it as the subgraph score. The score of the subgraph is computed as follows

$$e = \frac{1}{|X'|} \sum_{x_i \in X'} \mathbf{a}^T \cdot \tanh(Wx_i) \tag{6}$$

where  $e$  is the score of subgraph,  $W$  is the learnable weight matrix.  $W$  together with attention vector  $\mathbf{a}$  are shared by all subgraphs.

After above steps, we normalize the attention scores  $e$  ( $e_{pp}$  or  $e_{pu}$ ) using softmax function

$$\beta = \frac{\exp(e)}{\sum_j \exp(e_j)} \tag{7}$$

where  $e_j \in \{e_{pp}, e_{pu}\}$  and  $\beta$  ( $\beta_{pp}$  or  $\beta_{pu}$ ) denotes the weight of subgraphs.

Finally, with the learned weight  $\beta_{pp}$  and  $\beta_{pu}$ , we fuse the node representations in subgraphs to form a global representation of the heterogeneous graph

$$x_w = \frac{1}{|X'_{pp}|} \sum_{x_i \in X'_{pp}} \beta_{pp} x_i + \frac{1}{|X'_{pu}|} \sum_{x_j \in X'_{pu}} \beta_{pu} x_j \tag{8}$$

$$x_H = \text{Pooling}(x_w) \tag{9}$$

$x_w$  contains rich global relation information of the discussion network. Therefore, after a necessary pooling layer, we attain the discussion network's representation  $x_H$  for price prediction in a later section.

### 3.6. Intergrating Features and Making Prediction

As in Figure 3, after extracting  $x_H$  from the discussion network and  $h_k$  from price series, we finally combine those information and make the prediction, which is formulated as follows

$$P_k = [h_k; x_H] \tag{10}$$

where  $P_k$  is the concatenation of  $x_H$  and  $h_k$ . We then feed  $P_k$  into a simple feed-forward neural network with softmax function

$$\hat{y}_k = \text{softmax}(W^T P_k + b) \tag{11}$$

where  $W$  and  $b$  are learned parameters.  $\hat{y}_k$  is the predicted probability distributions.

Finally, in order to train the parameters, the cross-entropy loss is used as the model's objective function. The loss  $\mathcal{L}$  is computed as

$$\mathcal{L} = \sum_{i=1}^N \sum_{c \in \{0,1\}} -y_i \log \hat{y}_i \tag{12}$$

where  $y_i$  is [1,0] or [0,1].  $N$  is the number of training data and  $c$  indicates the class label.

## 4. Experiments and Results

### 4.1. Description of Data

In this study, we first collect historical prices of the hog, maize, and bean from 2013 to 2020. All these historical prices are available from <http://www.wind.com.cn/>, accessed on



7 February 2021. Then, we extract discussions from an online professional pig community (<https://bbs.zhue.com.cn/>, accessed on 24 February 2021). Figure 4 shows the historical price data we collected, and Figure 5 is an example of such discussion. As we can see, the discussion contains people’s analysis and reflects their expectations. As a result, such a discussion already includes many other factors, in Figure 5, it contains the influence of supply and demand.

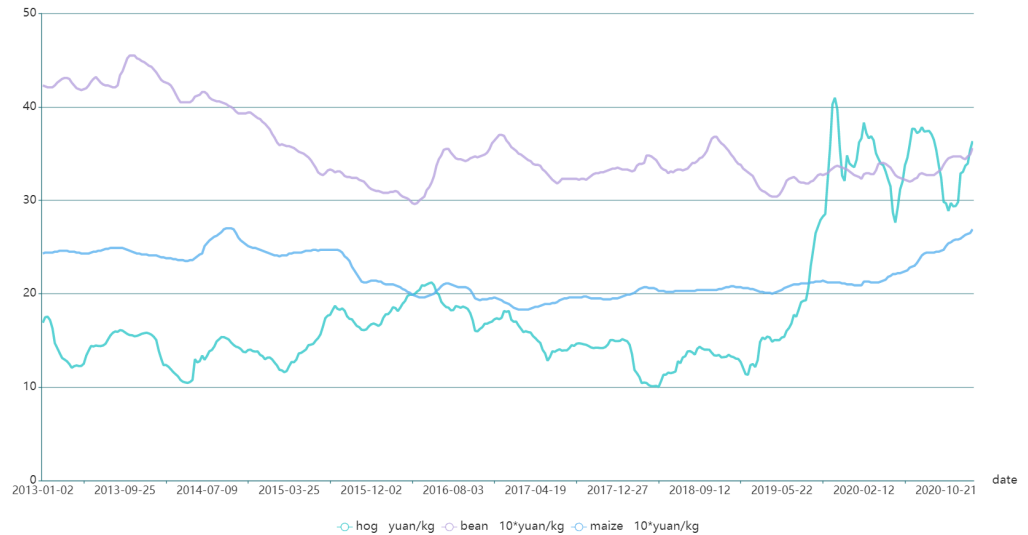


Figure 4. Weekly prices of the three agricultural products from 2013 to 2020.

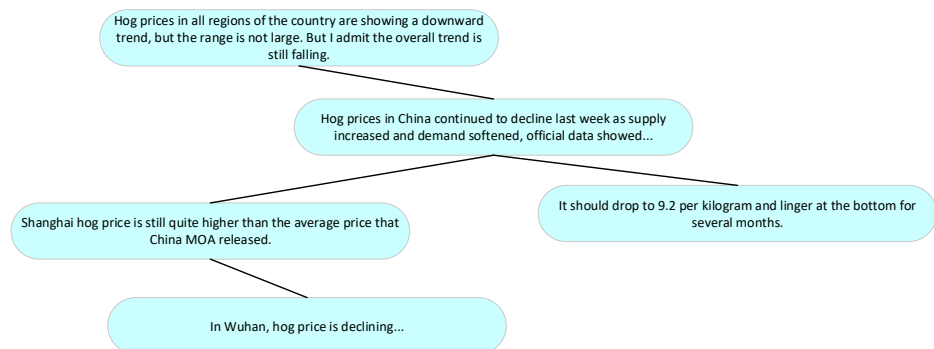


Figure 5. An example of the discussion content.

#### 4.2. Experimental Setup

When we make the dataset, each input price series  $(x_1, x_2, \dots, x_k), k \in [1, |\mathcal{X}|]$  has a corresponding discussion network and a label  $c$ .  $c = 1$  means hog price will increase next week and vice versa.

We implement our model and other compared models using PyTorch [44] and PyTorch-geometric [45]. Our experiments have been conducted on Tesla P100-16GB. We use the cross-entropy loss function and the Adagrad optimizer to train our model and set the learning rate as  $5 \times 10^{-3}$ .

#### 4.3. Evaluation Metrics

Considering that we transform the price prediction task into a binary classification problem, we use four popular performance indices to evaluate the models. These evaluation metrics are Accuracy, F1-score, Precision, and Recall. They are calculated as follows

$$\begin{aligned}
 Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\
 Precision &= \frac{TP}{TP + FP} \\
 Recall &= \frac{TP}{TP + FN} \\
 F1 &= \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}}
 \end{aligned}
 \tag{13}$$

where *TP* denotes true positive, *TN* denotes true negative, *FP* denotes false positive and *FN* denotes false negative.

#### 4.4. Competing Models

We have conducted extensive experiments to compare our proposed method’s performance with several popular methods for classification problems, single LSTM, multilayer perceptron (MLP), and STL-ATTTLSTM. We briefly discuss these competing methods here.

- Single LSTM: Proposed by Hochreiter and Schmidhuber [38], LSTM networks have shown superiority in processing time-series data. Therefore, LSTM networks are usually exploited when dealing with time series classification problems. In this study, we build a one-layer LSTM network for comparison;
- MLP: As a class of feedforward artificial neural network, multilayer perceptron usually consists of an input layer, an output layer, and several hidden layers. Researchers often make use of MLPs to solve regression problems. Since classification is a particular case of regression, MLPs also make good classifiers;
- STL-ATTTLSTM: Proposed by Yin et al. [37], STL-Attention-based LSTM is a state-of-the-art method to forecast the price of agricultural products. In their original paper, STL-ATTTLSTM makes use of several types of information to forecast monthly vegetable prices, such as vegetable prices, weather information, and market trading volumes [37]. According to their paper, the STL algorithm decomposes the price series into three parts: trend, seasonality, and remainder components. Then, they feed the remainder components into an LSTM network with an attention layer by removing the trend and seasonality components. Their experiments have shown promising results;
- BERTLSTM [46]: As BERT [47] has shown a great capacity to capture semantic information from text, Ko and Chang [46] exploited BERT to extract better representations of news article. After feeding the stock prices into LSTM module, they integrate price features and news features. Inspired by their study, we select BERTLSTM as one of the competing models;
- GCNLSTM [48]: GCN [49] is a popular model to extract hidden representation on graph structure data. Li et al. [48] proposed GCNLSTM for traffic flow prediction. They employ GCN to mine the spatial relationships of traffic flow. Then, they use LSTM module to extract temporal features. Finally, they design a structure to make the final prediction.

Table 2 shows the performance of our proposed method and all competing methods on the dataset.

**Table 2.** Experimental results on our dataset.

Method	Accuracy	Precision	Recall	F1
MLP	0.552	0.591	0.433	0.501
Single LSTM	0.741	0.742	0.767	0.754
STL-ATTTLSTM	0.792	0.763	0.799	0.781
BERTLSTM	0.809	0.783	0.809	0.796
GCNLSTM	0.814	0.812	0.804	0.808
proposed HGLSTM	0.832	0.838	0.812	0.825

#### 4.5. Discussion of Results

##### 4.5.1. Performance Comparison

As is shown in Table 2, deep neural networks achieve much better results than MLP. This phenomenon is very reasonable because deep neural networks have much more powerful learning abilities and can extract better representations. In contrast, typical MLP architectures are not deep, and they do not have many hidden layers, resulting in their relatively poor performance. Our experiments again prove the effectiveness of deep neural networks in classification.

LSTMs are well known for their effectiveness in dealing with series data. Our experiments also demonstrate this. As in Table 2, Single LSTM, STL-ATTLSTM, and our HGLSTM all contain LSTM networks, accounting for their better performance than MLP. Yin et al. [37] integrates the STL method and attention mechanism into LSTM. Therefore, their STL-ATTLSTM outperforms Single LSTM.

According to Table 2, our proposed Heterogeneous Graph-enhanced LSTM (HGLSTM) outperforms every competing method in terms of all metrics, indicating the effectiveness of our model. When we compare our HGLSTM with Single LSTM, the main difference is that HGLSTM includes the information of online discussions. This alone proves that discussion networks contain helpful information for hog price prediction.

It is also worth noting that, although we do not deal with Seasonality or Trends like STL-ATTLSTM, our HGLSTM still outperforms STL-ATTLSTM. According to their paper, the STL algorithm decomposes the price series into trend, seasonality, and remainder components before feeding the remainder components into LSTM. Their experiments proved the effectiveness of their model. Although we do not decompose the price series using the STL algorithm, our model still outperforms STL-ATTLSTM. This is mainly due to the introduction of discussion information, showing the success of including such information.

##### 4.5.2. Importance of Constructing Heterogeneous Graph

Now that we have proved the effectiveness of including discussion information, we still have various ways of capturing that representation. Thus, we further perform experiments to show that constructing the heterogeneous graph is the most effective way. We carefully choose two competing methods, GCNLSTM and BERTLSTM, which have been described in Section 4.4.

As Figure 2 shows, HGLSTM unquestionably outperforms both GCNLSTM and BERTLSTM. Here, we analyze the reasons for this. For BERTLSTM, it neglects both propagation structure and user information. Such information is indispensable for classification. GCNLSTM has a similar shortcoming. Its ability to model graph network enables it to exploit the propagation structure; however, it only allows one type of node and does not consider user information. Thus, GCNLSTM treats every node equally. This is definitely not good because, in a real discussion network, the credit of different users is not the same. High-credit users should influence the prediction far more than low-credit users. Adding user nodes into the graph, our proposed HGLSTM has two types of nodes, solving this problem.

## 5. Conclusions and Future Work

In this paper, we propose a heterogeneous graph-enhanced LSTM network for hog price prediction. We assume the online discussions can enhance hog price prediction and prove this through our experiments. To make the best use of discussions and user information, we resort to constructing the heterogeneous graphs. Our experiments demonstrate the effectiveness of incorporating online discussions and constructing heterogeneous graphs.

In the future, we plan to investigate how to combine discussion information and price series representations more efficiently and effectively.

**Author Contributions:** Conceptualization, K.Y., X.C., Y.P. and K.Z.; methodology, K.Y.; software, K.Y. and K.Z.; investigation, K.Y. and Y.P.; resources, X.C.; writing—original draft preparation, K.Y.; writing—review and editing, K.Y., X.C., Y.P. and K.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Key Research and Development Program of China No.2018YFC1604000, Fundamental Research Funds for the Central Universities No.2042017gf0035.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All the data used in this paper is available online and we have given the specific information about dataset in previous sections.

**Acknowledgments:** We would like to thank the support of National Key Research and Development Program of China No. 2018YFC1604000, Fundamental Research Funds for the Central Universities No. 2042017gf0035 as well as all the reviewers for their enlightening comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Kim, H.N.; Choi, I.C. The Economic Impact of Government Policy on Market Prices of Low-Fat Pork in South Korea: A Quasi-Experimental Hedonic Price Approach. *Sustainability* **2018**, *10*, 892. [\[CrossRef\]](#)
- Li, J.; Liu, W.; Song, Z. Sustainability of the Adjustment Schemes in China's Grain Price Support Policy: An Empirical Analysis Based on the Partial Equilibrium Model of Wheat. *Sustainability* **2020**, *12*, 6447. [\[CrossRef\]](#)
- Vandone, D.; Peri, M.; Baldi, L.; Tanda, A. The impact of energy and agriculture prices on the stock performance of the water industry. *Water Resour. Econ.* **2018**, *23*, 14–27. [\[CrossRef\]](#)
- Erokhin, V. Factors influencing food markets in developing countries: An approach to assess sustainability of the food supply in Russia. *Sustainability* **2017**, *9*, 1313. [\[CrossRef\]](#)
- Vu, T.N.; Ho, C.M.; Nguyen, T.C.; Vo, D.H. The Determinants of Risk Transmission between Oil and Agricultural Prices: An IPVAR Approach. *Agriculture* **2020**, *10*, 120. [\[CrossRef\]](#)
- Tomal, M.; Gumieniak, A. Agricultural Land Price Convergence: Evidence from Polish Provinces. *Agriculture* **2020**, *10*, 183. [\[CrossRef\]](#)
- Schumaker, R.P.; Chen, H. Textual analysis of stock market prediction using breaking financial news: The AZFin text system. *ACM Trans. Inf. Syst. (TOIS)* **2009**, *27*, 1–19. [\[CrossRef\]](#)
- Liu, Q. Price relations among hog, corn, and soybean meal futures. *J. Futur. Mark. Futur. Opt. Other Deriv. Prod.* **2005**, *25*, 491–514. [\[CrossRef\]](#)
- Darekar, A.; Reddy, A.A. Cotton price forecasting in major producing states. *Econ. Aff.* **2017**, *62*, 373–378. [\[CrossRef\]](#)
- Jadhav, V.; Reddy, C.B.V.; Gaddi, G. Application of ARIMA model for forecasting agricultural prices. *J. Agric. Sci. Technol.* **2017**, *19*, 981–992.
- Pardhi, R.; Singh, R.; Paul, R.K. Price Forecasting of Mango in Lucknow Market of Uttar Pradesh. *Int. J. Agric. Environ. Biotechnol.* **2018**, *11*, 357–363.
- Li, W.; Ding, W.; Sadasivam, R.; Cui, X.; Chen, P. His-GAN: A histogram-based GAN model to improve data generation quality. *Neural Netw.* **2019**, *119*, 31–45. [\[CrossRef\]](#)
- Assis, K.; Amran, A.; Remali, Y. Forecasting cocoa bean prices using univariate time series models. *Res. World* **2010**, *1*, 71.
- Adanacioglu, H.; Yercan, M. An analysis of tomato prices at wholesale level in Turkey: an application of SARIMA model. *Custos Gronegocio Line* **2012**, *8*, 52–75.
- Gu, Y.; Yoo, S.; Park, C.; Kim, Y.; Park, S.; Kim, J.; Lim, J. BLITE-SVR: New forecasting model for late blight on potato using support-vector regression. *Comput. Electron. Agric.* **2016**, *130*, 169–176. [\[CrossRef\]](#)
- BV, B.P.; Dakshayini, M. Performance analysis of the regression and time series predictive models using parallel implementation for agricultural data. *Procedia Comput. Sci.* **2018**, *132*, 198–207.
- Minghua, W.; Qiaolin, Z.; Zhijian, Y.; Jingui, Z. Prediction model of agricultural product's price based on the improved BP neural network. In Proceedings of the 2012 7th International Conference on Computer Science & Education (ICCSE), Melbourne, VIC, Australia, 14–17 July 2012; pp. 613–617.
- Nasira, G.; Hemageetha, N. Vegetable price prediction using data mining classification technique. In Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012), Salem, India, 21–23 March 2012; pp. 99–102. [\[CrossRef\]](#)
- Wang, B.; Liu, P.; Chao, Z.; Junmei, W.; Chen, W.; Cao, N.; O'Leary, G.M.; Wen, F. Research on hybrid model of garlic short-term price forecasting based on big data. *CMC Comput. Mater. Continua* **2018**, *57*, 283–296. [\[CrossRef\]](#)

20. Hemaetha, N.; Nasira, G.M. Radial basis function model for vegetable price prediction. In Proceedings of the 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, Salem, India, 21–22 February 2013; pp. 424–428. [[CrossRef](#)]
21. Li, Z.M.; Cui, L.G.; Xu, S.W.; Weng, L.Y.; Dong, X.X.; Li, G.Q.; Yu, H.P. Prediction model of weekly retail price for eggs based on chaotic neural network. *J. Integr. Agric.* **2013**, *12*, 2292–2299. [[CrossRef](#)]
22. Luo, C.; Wei, Q.; Zhou, L.; Zhang, J.; Sun, S. Prediction of vegetable price based on Neural Network and Genetic Algorithm. In Proceedings of the International Conference on Computer and Computing Technologies in Agriculture, Nanchang, China, 22–25 October 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 672–681.
23. Zhang, D.; Zang, G.; Li, J.; Ma, K.; Liu, H. Prediction of soybean price in China using QR-RBF neural network model. *Comput. Electron. Agric.* **2018**, *154*, 10–17. [[CrossRef](#)]
24. Asgari, S.; Sahari, M.A.; Barzegar, M. Practical modeling and optimization of ultrasound-assisted bleaching of olive oil using hybrid artificial neural network-genetic algorithm technique. *Comput. Electron. Agric.* **2017**, *140*, 422–432. [[CrossRef](#)]
25. Ma, C.; Shi, X.; Zhu, W.; Li, W.; Cui, X.; Gui, H. An Approach to Time Series Classification Using Binary Distribution Tree. In Proceedings of the 2019 15th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN), Shenzhen, China, 11–13 December 2019; pp. 399–404.
26. 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*, 2831–2844. [[CrossRef](#)]
27. Li, Y.; Li, C.; Zheng, M. A hybrid neural network and HP filter model for short-term vegetable price forecasting. *Math. Probl. Eng.* **2014**, *2014*.
28. Jin, D.; Yin, H.; Gu, Y.; Yoo, S.J. Forecasting of Vegetable Prices using STL-LSTM Method. In Proceedings of the 2019 6th International Conference on Systems and Informatics (ICSAI), Shanghai, China, 2–4 November 2019; pp. 866–871.
29. Liu, Y.; Duan, Q.; Wang, D.; Zhang, Z.; Liu, C. Prediction for hog prices based on similar sub-series search and support vector regression. *Comput. Electron. Agric.* **2019**, *157*, 581–588. [[CrossRef](#)]
30. Yoo, D. Developing vegetable price forecasting model with climate factors. *Korean J. Agric. Econ.* **2016**, *57*, 1–24.
31. Chen, Q.; Lin, X.; Zhong, Y.; Xie, Z. Price Prediction of Agricultural Products Based on Wavelet Analysis-LSTM. In Proceedings of the 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, China, 16–18 December 2019; pp. 984–990.
32. Bahdanau, D.; Cho, K.H.; Bengio, Y. Neural machine translation by jointly learning to align and translate. In Proceedings of the 3rd International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.
33. Qin, Y.; Song, D.; Cheng, H.; Cheng, W.; Jiang, G.; Cottrell, G.W. A dual-stage attention-based recurrent neural network for time series prediction. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 2627–2633.
34. Ran, X.; Shan, Z.; Fang, Y.; Lin, C. An LSTM-based method with attention mechanism for travel time prediction. *Sensors* **2019**, *19*, 861. [[CrossRef](#)]
35. Li, Y.; Zhu, Z.; Kong, D.; Han, H.; Zhao, Y. EA-LSTM: Evolutionary attention-based LSTM for time series prediction. *Knowl. Based Syst.* **2019**, *181*, 104785. [[CrossRef](#)]
36. Zhang, X.; Liang, X.; Zhiyuli, A.; Zhang, S.; Xu, R.; Wu, B. AT-LSTM: An attention-based LSTM model for financial time series prediction. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2019; Volume 569, p. 052037.
37. Yin, H.; Jin, D.; Gu, Y.H.; Park, C.J.; Han, S.K.; Yoo, S.J. STL-ATLSTM: Vegetable Price Forecasting Using STL and Attention Mechanism-Based LSTM. *Agriculture* **2020**, *10*, 612. [[CrossRef](#)]
38. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
39. Li, W.; Fan, L.; Wang, Z.; Ma, C.; Cui, X. Tackling mode collapse in multi-generator GANs with orthogonal vectors. *Pattern Recognit.* **2021**, *110*, 107646. [[CrossRef](#)]
40. Li, W.; Liang, Z.; Ma, P.; Wang, R.; Cui, X.; Chen, P. Hausdorff GAN: Improving GAN Generation Quality With Hausdorff Metric. *IEEE Trans. Cybern.* **2021**. [[CrossRef](#)] [[PubMed](#)]
41. Bird, S.; Klein, E.; Loper, E. Natural language processing with Python: analyzing text with the natural language toolkit. In Proceedings of the 5th International Conference on Learning Representations, Toulon, France, 24–26 April 2017.
42. Pennington, J.; Socher, R.; Manning, C.D. GloVe: Global Vectors for Word Representation. In Proceedings of the Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
43. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph Attention Networks. *arXiv* **2017**, arXiv:1710.10903.
44. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems 32*; Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R., Eds.; Curran Associates, Inc.: New York, NY, USA, 2019; pp. 8024–8035.
45. Fey, M.; Lenssen, J.E. Fast Graph Representation Learning with PyTorch Geometric. *arXiv* **2019**, arXiv:1903.02428.
46. Ko, C.R.; Chang, H.T. LSTM-based sentiment analysis for stock price forecast. *PeerJ Comput. Sci.* **2021**, *7*, e408. [[CrossRef](#)]
47. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2018**, arXiv:1810.04805.

- 
48. Li, Z.; Xiong, G.; Chen, Y.; Lv, Y.; Hu, B.; Zhu, F.; Wang, F. A Hybrid Deep Learning Approach with GCN and LSTM for Traffic Flow Prediction\*. In Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 27–30 October 2019; pp. 1929–1933. [[CrossRef](#)]
  49. Kipf, T.N.; Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. In Proceedings of the 5th International Conference on Learning Representations, Toulon, France, 24–26 April 2017.