

## Research Article

# Procurement Volume Prediction of Cross-Border E-Commerce Platform Based on BP-NN

Yaping Jiang 

*School of Business, Shanghai Normal University Tianhua College, Jiading District, Shanghai, China*

Correspondence should be addressed to Yaping Jiang; [jyp2186@sthu.edu.cn](mailto:jyp2186@sthu.edu.cn)

Received 9 June 2022; Revised 2 August 2022; Accepted 29 August 2022; Published 16 September 2022

Academic Editor: Chia-Huei Wu

Copyright © 2022 Yaping Jiang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

For logistics management, predictive analysis has never been more crucial than it is now, thanks to the vast number of deal data generated every second by electronic commerce. In an effort to enhance customer service and supply control, e-commerce businesses are progressively utilizing machine learning technologies to enhance projections. The back-propagation neural network (BP-NN) model is used to develop a C-A-BP forecasting model that considers commodity sales characteristics and the data series' trend. In order to predict each cluster, a C-BP-NN model is first created, adding sales information as influencing elements into the C-BP-NN model. An A-BP-NN model is used in combination with the ARIMA that is employed for the linear component. These two forecasting models are merged to provide the final results. By comparing the results of the ARIMA, BP-NN, A-BP-NN, and C-BP-NN using the information provided by Jollychic's cross-border platform, the A-C-BPNN was shown to be the best.

## 1. Introduction

Customer satisfaction in the supply chain industry is enhanced when goods are prestocked at local warehouses in numerous marketplaces throughout the world, reducing logistics time. Because of the globalized production and sales regions, it takes longer for cross-border e-commerce companies to make arrangements from the acquisition of commodities, shipping, and customs quality control to arrive at the final destination [1]. Data from e-commerce commodity sales form the foundation for logistics management, and hence algorithms and big data analysis methods are frequently used to forecast sales. This data is critical to cross-border e-commerce companies' worldwide supply chain plans [2].

The sophistication of the cross-border e-commerce sector affects sales predictions in a variety of ways, in addition to the volume and variety of transaction data. As a result, e-commerce businesses still face a problem when trying to increase the precision and accuracy of their sales forecasts [3].

ARIMA [4], Procurement Volume Prediction [5], and BP-NN [6] are some of the terms used to describe this research. The following is a brief summary of the main points. Supply chain management necessitates [7] that items be stored in preparation in local warehouses in numerous marketplaces throughout the world, which might significantly cut logistics time. Although the manufacturing and sales sectors of e-commerce items are globalized, it takes cross-border e-commerce firms longer to prepare for the purchase of supplies, transportation, and customs quality inspections, to name a few. Data from e-commerce commodity sales form the foundation for supply chain management, and hence algorithms and big data analysis technologies are frequently used to forecast sales. This data is critical to cross-border e-commerce companies' worldwide supply chain plans [8].

It is not just transaction data that affects sales projections in the cross-border e-commerce business, but many other elements, as well. As a result, e-commerce businesses still face a problem when trying to increase the efficiency of their sales forecasts.

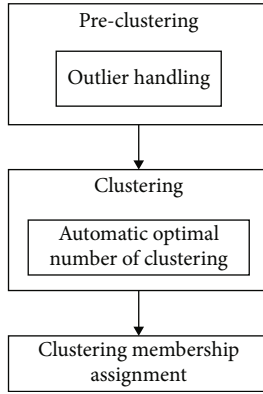


FIGURE 1: The two-step clustering method's main technology and procedures.

Much research has been done in the field of sales forecasting. Most of these research approaches to predicting sales use either time series (TS) models [9, 10] or machine learning (ML) algorithms [11].

In order to extrapolate emerging tendencies from existing data collected, TSs vary from moving averages [12] to the ARIMA groupings. Although useful for sales forecasting, TSs are limited by the expectation of linear development and cannot compensate for other effects like pricing and marketing changes. There are several studies in which univariate forecasting approaches have been used as a benchmark model [13, 14].

MLs have had a vital role in the field of predicting. State-of-the-art prediction models such as ANN [15], RBF [16], metaheuristic-based NNs (MHNN) [17], deep convolutional NN (DCNN) [18, 19], long short-term memory network [20, 21], SVR [22], and ELM [19] have had a significant impact on the current MLs [23].

MLs and TSs have been compared in various current prediction models. ANN outperforms ARIMA in forecasting demand, according to a study cited in [3]. Reference [24] evaluated classic statistical approaches, such as ARIMA and multivariate regression, and found that ANNs performed better than the more statistical approaches. Reference [25] compared SVR to ARIMA, Holt-Winters, and moving average for use in sales prediction with promotion influences.

TS-based MLs have also been used for sales forecasting. Referencing [19], it was shown that combining ARIMA and ANN to simulate both linear and nonlinear aspects of the dataset was more efficient. There was an ARIMA prediction model developed in [20] that was used to train and fit the BP-NN using its residuals. LSTM ensemble prediction was demonstrated in [21] using a unique approach that efficiently aggregates several prediction outputs from a series of independent LSTM networks. As more and more e-commerce data becomes available, several algorithms have been developed to better manage unpredictable sales trends and account for a variety of other aspects. Reference [22] presented a unique method for automatic learning useful features from structured information using CNNs. Training an LSTM model with marketing consumer demands and cross-series information was the goal of reference [23]. ELM was frequently used in predicting, which is more

important. Using ELM and distribution optimization, a new data-driven strategy for predicting user behavior was developed in [25]. In [26], a deep learning architecture was used to improve ELM's ability to predict wind speed.

Even though there are a variety of prediction models, the choice of approach is influenced by the features of different items. There is evidence to support the idea that product qualities can influence both search and sales since customers are more interested in the intrinsic properties of items when making a purchase decision. Clustering algorithms have been used to better represent the features of commodities in sales forecasting. When it comes to improving the outcomes of NNs, both fuzzy artificial NNs and aggregating approaches have been applied. Reference [26] used hierarchy self-organizing maps and principal component analysis to create the SVR to address the sales forecasting challenge. In reference [27], an ELM and composite linkage approach were used to develop a sales prediction model. Reference [28] developed an SVR-based clustering sales forecasting system. For forecasting computer retail sales, reference [29] created a clustering-based forecasting model that combines clustering and ML methodologies.

Three-phase BP-NN-based prediction models are created to focus on the two factors (sales attributes and data series trend) indicated above in this study, according to the literature review.

It is necessary to first incorporate numerous sales-influencing elements using the two-step clustering technique, which is an upgraded version of BIRCH. Each cluster is then modeled using the BP-NN approach, which has been proven to be a successful predictor in numerous data analysis contests, such as Kaggle and many recently published papers.

For the complex sections of sample data, we present an A-BP-NN approach that integrates the characteristics of ARIMA and BP-NN methods, respectively, to improve trend prediction performance. It is possible to build a final mixture strategy that takes into account the various factors that influence sales of products and the time series pattern. The C-A-BP-NN model is shown here.

To sum up, The BPNN and ARIMA have been shown in this paper to be speedy and efficient throughout the initial phase of global search, addressing the aforementioned drawbacks. Based on their advantages and limitations, this paper presents a hybrid computational technique that combines ARIMA with BPNN weight optimization. This process is known as ARIMA and BPNN computation. The ARIMA's global search capability and the BPNN's range search capability are both utilized by the mixed computation technique.

The main contributions of the paper are as follows:

- (i) We present a unique GA method, dubbed the BPNN and ARIMA, in which a novel ARIMA agent acts as an inventive operator, enhancing BPNN and ARIMA's exploitation and exploration capabilities in the issue search space
- (ii) We compare the proposed ARIMA performance to well-known and innovative metaheuristic algorithms using well-known benchmark issues

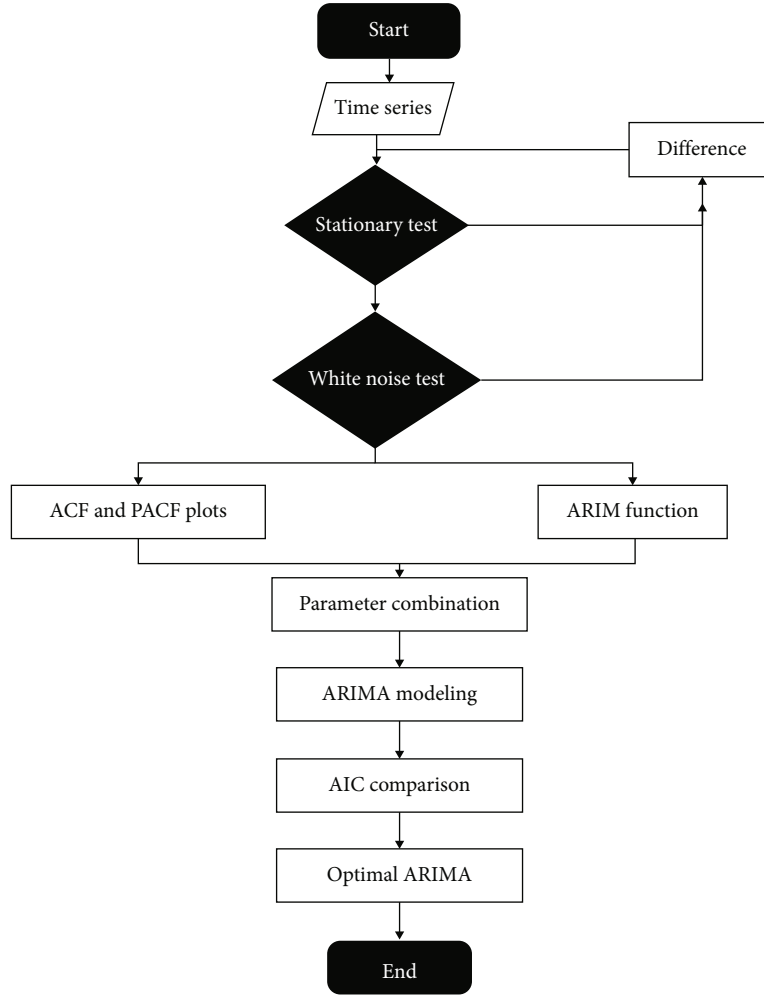


FIGURE 2: ARIMA model's parameter estimation techniques.

TABLE 1: The parameters' description of the BP-NN.

Parameters	Values
Learning rate	0.05
Number of input nodes	Number of features ( $N$ )
Number of hidden nodes	$2 \times N + 1$
Number of output nodes	1
Momentum	0.8
Number of iterations	1000

- (iii) We design a BPNN's training technique and demonstrate a significant increase in the performance of the proposed forecasting model
- (iv) For the first time, we forecast Procurement Volume Prediction of Cross-Border E-Commerce Platform
- (v) We compare our suggested method to state-of-the-art ANN to measure investment funds' forecasting accuracy

The remaining five parts of the document are as follows: Two-step clustering technique, ARIMA parameter estima-

tion method, and BP-NN model are only a few of the models and methods briefly discussed in Section 2. The BP-NN-based model is provided in Section 3 to anticipate sales attributes and trends of time series in a three-phase manner. This forecasting model's accuracy is demonstrated numerically in Section 4. Section 5 summarizes the findings and makes a suggestion for further study.

## 2. Methodologies

*2.1. Selecting the Appropriate Features.* The volume of data in the e-commerce industry has grown exponentially since the introduction of web technology. Big data is characterized by a range of formats, including Twitter, text, online, music, movie, click-stream, and log files. Techniques such as filter-based or wrapper feature extraction have been explored to minimize the dimension of the data to remove the most unnecessary and superfluous information from diverse sources. When predicting the accuracy of feature subsets in wrapper feature selection, effective training utilizes mathematical resampling (like cross-validation) to predict feature subset accuracy [30]. Different algorithms can use filter-based feature extraction to represent the identical dataset.

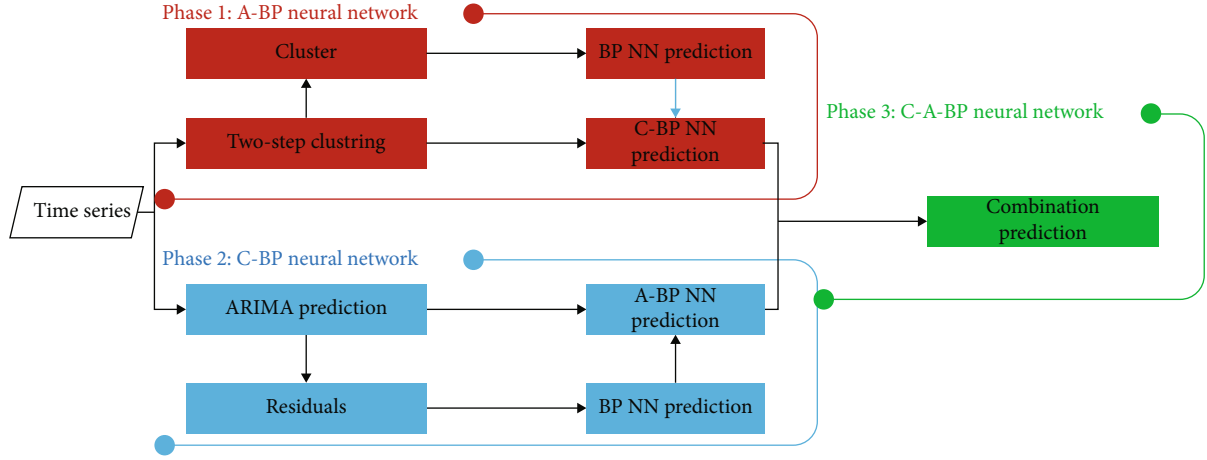


FIGURE 3: The suggested three-phase approach.

TABLE 2: Source data series description.

Data series	Data date + other fields
Goods information	Goods id; season; stock-keeping unit id; level; brand id
Customer behavior	Goods click; favorites click; cart click
Goods sales	Stock keeping unit sales; original shop price; goods price
Goods promote price	Goods promotion price; goods price
Goods id and stock-keeping unit id's relation	Stock keeping unit id; goods id
Marketing	Plan; marketing
Temperature	Temperature mean
Holidays	Holiday

TABLE 3: Clustering series description.

Fields	Meaning
Goods code	Goods code
Data date	Date
The sale of stock keeping unit	The sale of stock keeping unit
The code of stock keeping unit	The code of stock keeping unit
Goods price	Selling price
Goods click	# clicks on goods
Goods season	Seasonal features of goods
Original shop price	Tag price
Cart click	# clicks on purchasing carts
Favorites click	# clicks on favorites
Plan	The code of activity rhythm
Marketing	The code of activity type
Sale unique visitor	# unique visitors
Promotion	The code of promotion
The mean temperature	The mean temperature
Holiday	Vacation

In this research, the prediction and clustering techniques use wrapper feature selection to remove uninteresting qualities from multidimensional data based on SD, Pearson correlation coefficient, CV, and feature significance scores, with the following information.

Data dispersion is measured by SD, which may be computed as follows:

$$\delta = \sqrt{\frac{1}{M} \times \sum_{j=1}^M (y_i - \eta)^2}. \quad (1)$$

The coefficient of variation (CV) is a statistical measure of the variability of actual values in the data.

$$CV = \frac{\delta}{\eta}. \quad (2)$$

In statistical terms, PCC is a way of measuring the linear correlation between two variables.

$$PCC = \frac{1}{m-1} \sum_{j=1}^m \left( \frac{Y_i - \bar{Y}}{\delta_y} \right) \times \left( \frac{X_i - \bar{X}}{\delta_x} \right), \quad (3)$$

where  $((Y_i - \bar{Y})/\delta_y)$ ,  $\delta_y$ , and  $\bar{Y}$  show the standard deviation, standard score, and the mean of  $Y_i$ , respectively. Each feature is given a score based on how important it is to the model's enhanced decision trees. The more weight is given to a certain trait in decision trees, the more important it becomes. The importance of a single decision tree is derived by calculating the performance metric by each splitting

TABLE 4: The training, validation, and test sets' description.

Date set	Training 1	Training 2	Validation	Test
Samples	Clustering series	SKU code = 95144	SKU code = 95144	SKU code = 95144
# weeks	59	59	8	4
Start date	Apr.1 2017	Apr.1 2017	Oct.2 2018	Dec.5 2018
End date	Dec.29 2017	Dec.29 2017	Dec.29 2018	Jan.3 2019
First day	1	1	267	337
Last day	337	237	266	371

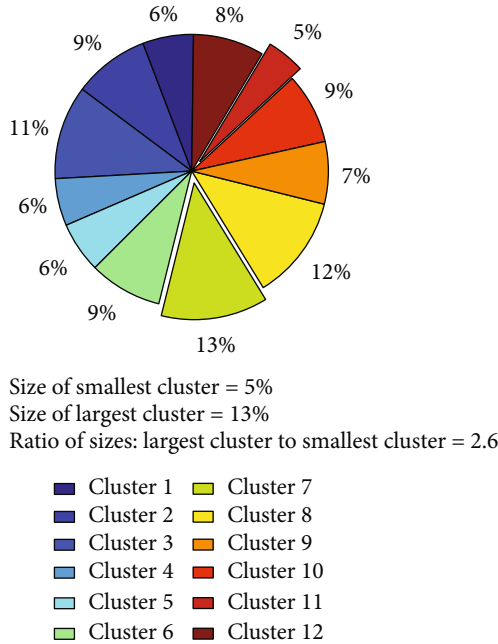


FIGURE 4: Cluster size of the two-step clustering system's clustering series.

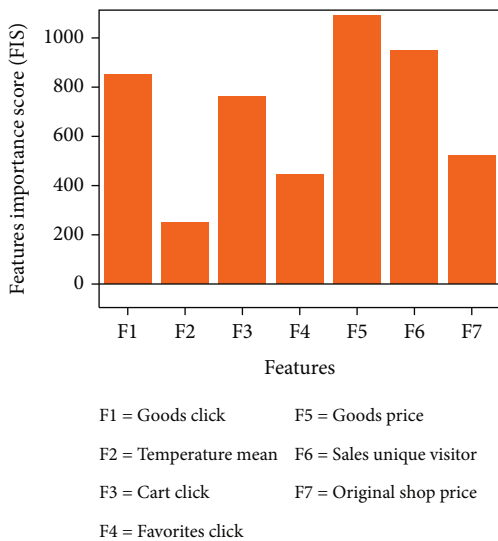


FIGURE 5: shows the C-BP-NN's component significance score.

point's adjusted set of measurements. An error function, such as Gini Index, may be used as a performance assessment, but it is not required. The relevance of each attribute is then aggregated throughout all of the decision trees.

2.2. Two-Phase Clustering Algorithm. When a dataset is broken down into many different subsets that are nearly identical, the procedure of clustering is used. Methods including segmentation, clustering, density-based categorization, grid-based clustering, and model-based clustering are frequently employed.

Clustering techniques are selected based mostly on the size and type of data that was gathered. When working with numerical or categorical data, standard techniques for clustering can be used. Reference [31] presented the BIRCH, a hierarchical technique that is particularly well-suited to big datasets with continuous characteristics. However, in this study, the two-step clustering approach in SPSS Modeler is recommended because of the huge and heterogeneous data. BIRCH's two-step clustering technique uses the scroll distance as a metric to calculate distances between categorical data as well as between continuous data. Preclustering monitoring is similar to BIRCH's two-step clustering procedure in that it keeps the focused parts of acquired information in the form of descriptive analysis. The dense zones are then clustered using a hierarchical clustering technique. Both BIRCH and the two-step clustering method are capable of handling mixed characteristics, but the two-step clustering technique has a novel technique for allocating cluster identification to noisy data.

Since it is a hierarchy approach, it is much more effective at coping with noise and inconsistencies than partitioning approaches. This approach is superior to others in that it automatically calculates the ideal number of clusters. The two-step clustering approach depicted in Figure 1 may be used to cluster items in e-commerce interaction datasets of large and heterogeneous transaction datasets.

2.2.1. Preclustering. Using the BIRCH algorithm's clustering feature (CF) tree growth, it is possible to execute outlier treatment while simultaneously reading each record in a data collection. The datasets in dense regions are then used to generate a CF tree, which yields subclusters.

2.2.2. Clustering. Based on hierarchical agglomerative clustering methods, the clusters are created by combining the

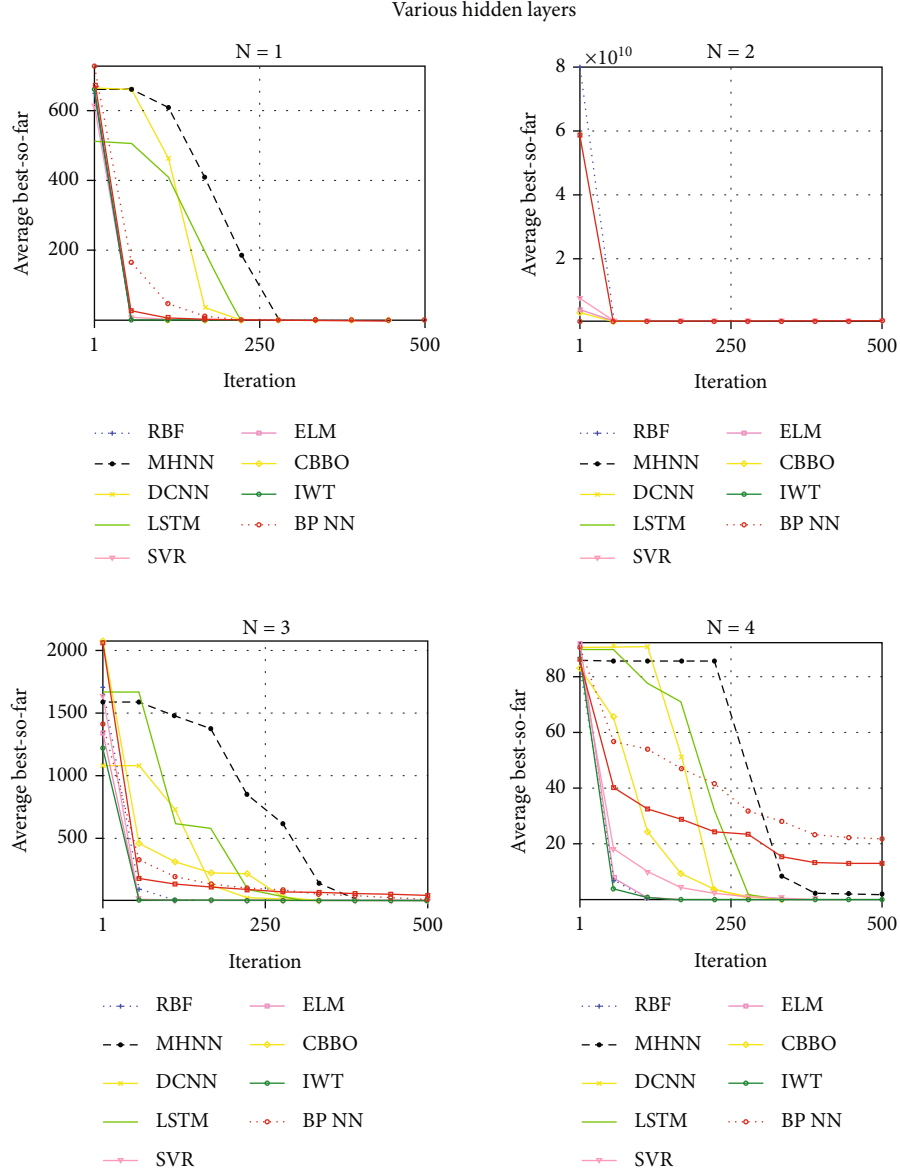


FIGURE 6: Convergence curve based on the various number of layers.

subclusters piece by piece and determining the ideal number based on the Bayesian information criteria.

**2.2.3. The Assignment of Cluster Membership.** It is determined which data records belong in which clusters by determining how far apart data records are from each other in terms of log-likelihood.

**2.2.4. Results' Validation.** In order to evaluate the performance of the clustering findings, the coefficient  $\varphi$  is used. Better clustering results can be achieved with a value of 1 greater than zero:

$$\varphi = \frac{\alpha_2 - \alpha_1}{\max(\alpha_2, \alpha_1)}, \quad (4)$$

where  $\alpha_1$  is the average distance among the sampling and

its groupings and  $\alpha_2$  is the average distance among the sampling and its separate cluster.

**2.3. The Optimization of the ARIMA model's Parameters.** Moving averages and autoregressive models are combined to form the ARIMA model. Reference [31] explains how to use the Box-Jenkins technique in the time series concept to build an ARIMA (p, d, and q) model. Since its parameters are often established by plotting ACF and PACF data, the ARIMA is susceptible to judgment variation when it comes to defining parameters. As an alternative, the function called ARIM () in the R package is often utilized to autonomously create an ideal ARIMA framework, which composes of the shortcoming of ARIMA during parameter judgment.

The ARIM () function is coupled with the findings of ACF and PACF plots in order to increase ARIMA's fitting

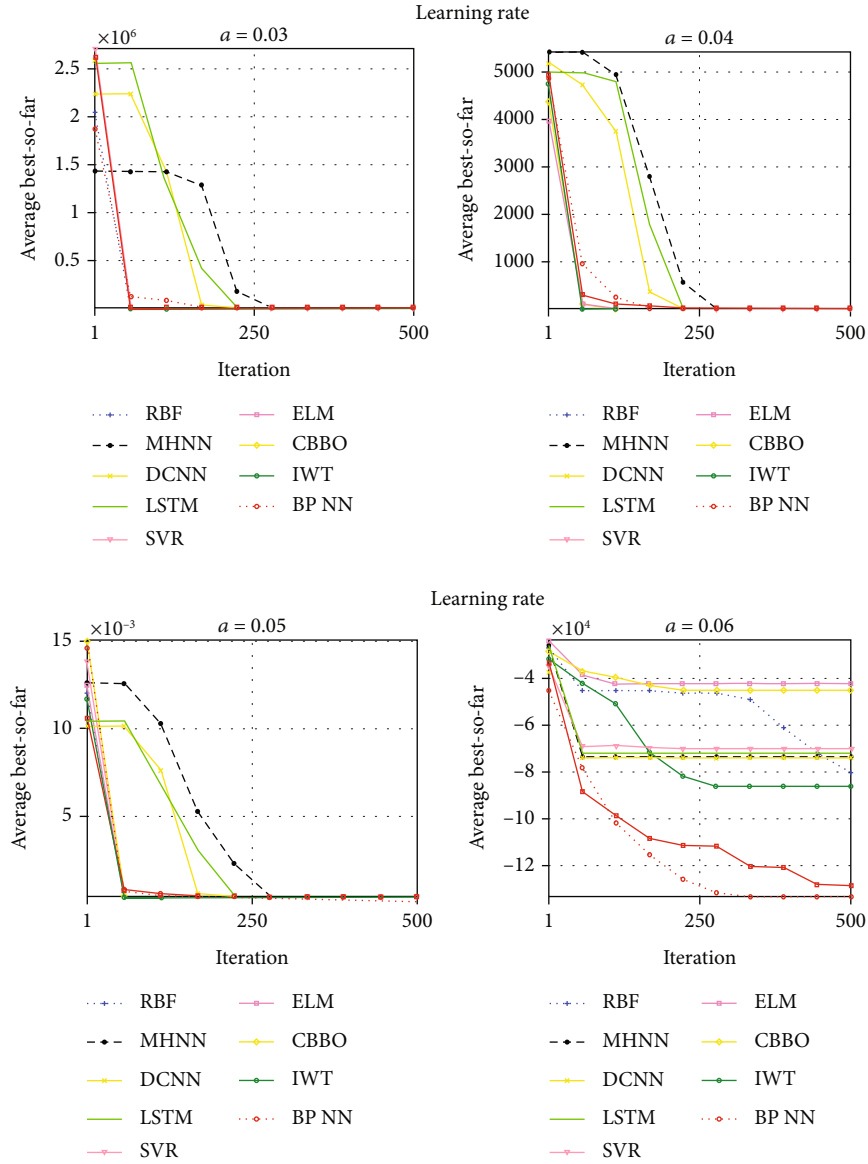


FIGURE 7: Convergence curve based on the various momentums.

performance, and this combined technique of parameter evaluation is now proposed. Figure 2 shows the steps, which are discussed below. ADF and Box–Pierce tests are used to verify the static and white noise prior to developing ARIMA. Time series can employ ARIMA if both stationary time series and white noise criteria pass. R application’s ARIM () function may be used to determine some parameter combinations, while ACF and PACF plots can be used to find other parameter combinations. Calculate the AIC scores for each scenario by simulating the ARIMA using a variety of hyperparameters. Finally, Find the best ARIMA parameter combinations with the smallest AIC possible.

**2.4. BP-NN Algorithm.** The “Back Propagation NN” proposed by Ian Goodfellow et al. Reference [27] is known as the BP-NN. In this study, rather than covering the impor-

tant core theory of the BP-NN, we focus on the algorithm’s methods.

**2.4.1. The Selection of Feature.** Data cleaning, feature extraction, and extraction of features based on feature relevance scores are the particular processes in the BP-NN feature selection process.

**2.4.2. The Training of the Model.** Default parameters are used to train the model depending on the characteristics that have been specified.

**2.4.3. The Optimization of Parameters.** In order to reduce the difference between expected and actual values, parameter optimization is used. According to Table 1, the method uses three different types of parameters.

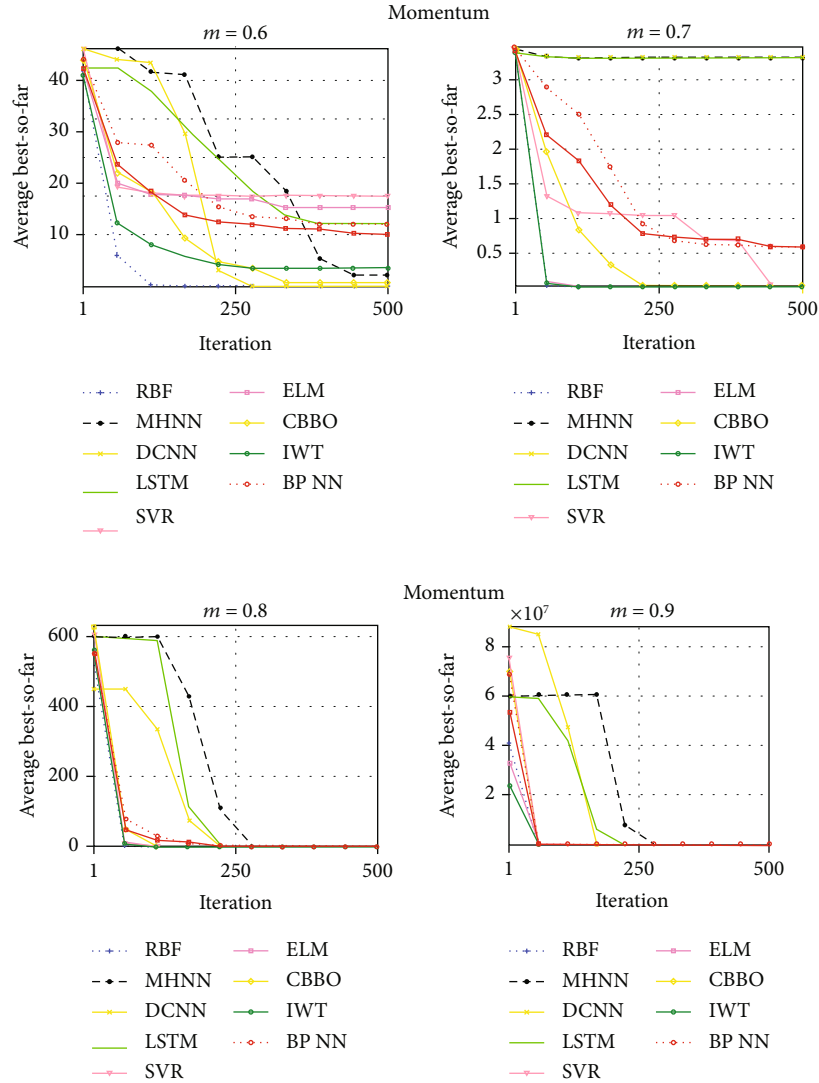


FIGURE 8: Convergence curve based on the various learning rates.

### 3. The Three-Phase Prediction Model

Research in this study proposes a three-phase C-A-BP-NN prediction system based on both the sales attributes and the data's tendency.

New NN models are developed for the first phase of forecasting in Phase 1, which combine distinct clustering characteristics into forecasts. BP-NNs are used to model each of the generated clusters, after which the two-step clustering approach is used to divide the commodities into distinct groups depending on characteristics.

Two NN models are described in Phase 2: the ARIMA-BP-NN architecture, which uses ARIMA's linear fitting capability and BP's nonlinear mapping ability to forecast trends in time series. The linear component of the data series is predicted using ARIMA, the nonlinear part is revised using ARIMA residuals, and a BP-NN is established using the rolling prediction approach.

The C-A-BP-NN is created by merging the C-BP-NN with the A-BP-NN at this stage. A-BP-NN and C-A-BP-

NN findings are given a weighted average in the C-A-BP-NN in order to reduce some errors of squares. This reflects the reliability and reliability of sales characteristics and data series. As seen in Figure 3, the proposed three-phase strategy is shown, and the details are provided below.

### 4. Experiment Results and Discussion

#### 4.1. Data Description

(i) An example of the C-A-BP-NN model's predicting performance is shown using the following datasets

(ii) Source data series

Eight data series are included in the source data series, as indicated in Table 2, between the dates of January 3rd, 2017, and December 29th, 2018.

(iii) Clustering series





FIGURE 9: The results of various algorithms.

TABLE 5: Statistical metrics for the C-A-BP-NN, C-BP-NN, A-BP-NN, and ARIMA.

Method	$R^2$	RMSE	MAE	RRMSE	MRE	MAPE
ARIMA	0.44	0.055	0.048	0.023	0.019	2.533
C-A-BP-NN	0.79	0.008	0.011	0.008	0.037	0.622
C-BP-NN	0.72	0.033	0.017	0.019	0.011	1.111
A-BP-NN	0.73	0.031	0.019	0.018	0.012	1.210

Ten continuous variables and six categorical features are derived from the reconstruction of the source data series in the collection of clustering characteristics. Table 3 depicts the clustering series feature descriptions.

**4.2. Experiments.** The following consistent experimental settings are developed to verify the theoretical model's performance in accordance with performance assessment indices.

**4.2.1. Uniform Dataset.** Training, validation, and test datasets are created, as shown in Table 4, to suit the requirements of different models. Here is a breakdown of how the data application works: A sample period of 381 days is covered by the clustering series. Training set 1, which includes data from the first 347 days of the clustering series, serves as the basis for two-step clustering models in the C-BP-NN model. After a two-step clustering process, the obtained data is utilized to build BP-NN models. The C-BP-NN model is tested using the remaining 34 days of data. As part of the validation process, two-step clustering is performed to partition and deploy the test set. (1) The ARIMA predicting is performed using the second training dataset, which contains data from the first to the 277th day, and the validation data is used to estimate ARIMA multicollinearity, which is then

utilized to build the A-BP-NN network. By using a specified set of data, it is then evaluated. C-BP and A-BP-NN designs were fitted with quick and effective weights using the latest 34 data samples.

**4.3. Uniform Evaluation Indexes.** Validating forecasting models have traditionally relied on a variety of performance criteria. In order to differentiate the best forecasting model, standard assessment metrics are used. A model's accuracy improves as its size decreases.

$$R^2 = 1 - \frac{\text{sum squared regression (SSR)}}{\text{sum of squares total (SST)}},$$

$$\text{MAE} = \left(\frac{1}{m}\right) \sum_{i=1}^m |Av_i - Pv'_i|,$$

$$\text{MAPE} = \frac{1}{m} \sum_{i=1}^m \left| \frac{Av_i - Pv'_i}{Av_i} \right| \times 100\%,$$

$$\text{RMSE} = \sqrt{\left(\frac{1}{m}\right) \sum_{i=1}^m (Av_i - Pv'_i)^2},$$

$$\text{RRMSE} = \sqrt{\left(\frac{1}{m}\right) \sum_{i=1}^m \left(\frac{Av_i - Pv'_i}{Av_i}\right)^2},$$

$$\text{MRE} = \left(\frac{1}{m}\right) \sum_{i=1}^m \frac{|Av_i - Pv'_i|}{|Av_i|}.$$
(5)

#### 4.4. Experiments of C-A-BP-NN Model

**4.4.1. C-BP-NN Model.** Training set 1 is the first dataset that is clustered using the two-step clustering technique.

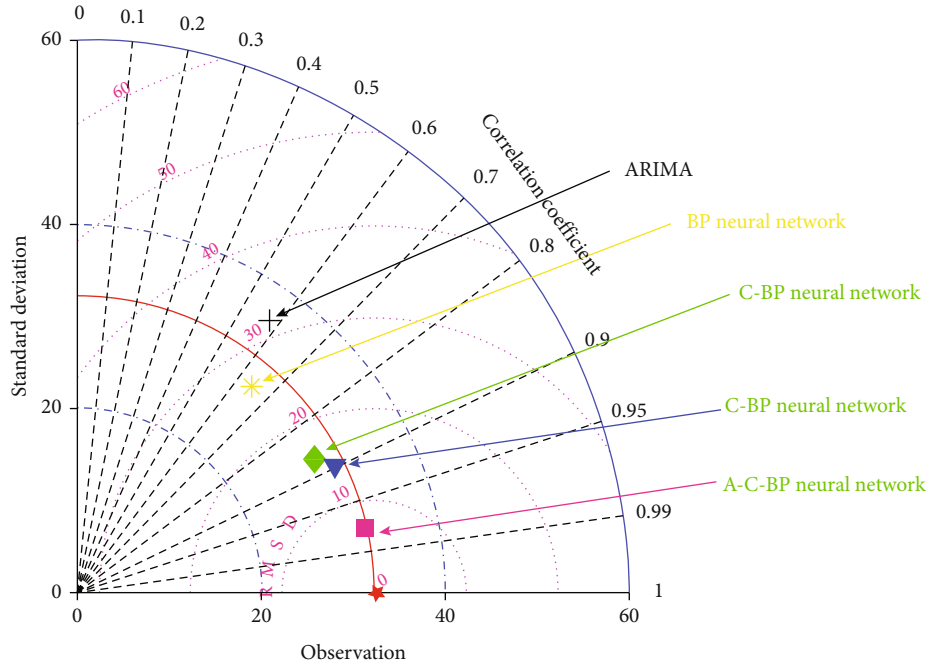


FIGURE 10: Tylor diagram.

Continuous characteristics are standardized; outliers are handled at a noise level of 25%; log-likelihood proximity is used to assess distance; BIC is used to cluster the data. Figure 4 illustrates the 2.64 ratios of sizes for each cluster, and the proportion is just right. As a result, the quality of the clusters is satisfactory.

(2) The C-BP-NN models are built in this step. The initial features are chosen from each of the twelve clusters according to the relevance of the features. C-BP-NN models for each cluster are then built using the chosen attributes of each cluster and the SKU marketing in Table 3 as the input and output variations, respectively

As an illustration of the modeling methods for the BP-NN, consider one of the 12 clusters.

Table 3 lists the characteristics, and Figure 5 shows the seven that have been narrowed down from that list. Among the most important elements are the following: F1, F3, F5, F6, and F7. Although F2 and F4 make less of an impact on the forecast, they are still important factors.

The cluster is pretrained by using the default input and output, which are the 11 cluster characteristics in Step 1 and the associated SKU sales in Table 3.

(3) Optimization of parameter values: When optimizing BP-NNs, the most important step is determining which input and output parameters are optimal. Parameter optimization, on the other hand, has less of an effect on the algorithm's accuracy. To improve the BP-NN, just the fundamental parameters, such as the number of layers, momentum, and the learning rate, are changed. The model can be balanced since increasing the number of layers will make the model more complicated and more likely to overfit, while raising the values of momentum and learning rate will make the model less conservative. Figures 6–8 demonstrate the average variations in the convergence curve based on the number of layers, momentum, and

learning rate, respectively, of the BP-NN and other benchmark prediction models

*4.5. Models for Comparison.* As a comparison between our suggested models and other models, the following methods are selected: ARIMA is a popular time series model for estimating future sales. By establishing the specified characteristics and the associated input and output, the BP-NN model is formed and optimized. In the NN of the C-BP system, to predict sales based on the two-step clustering model's results, the BP-NN considers commodity sales characteristics. In the A-BP-NN, the ARIMA residuals are revised using the A-BP-NN. When it comes to time series modeling, BP-NN is employed for the nonlinear component and ARIMA for the linear part. In the C-A-BP-NN, the C-BP and A-BP-NNs are combined in this model to maximize their respective strengths and minimize their respective weaknesses. The simulation results are shown in Figure 9 and Table 5. Also, the Tylor diagram for various utilized methods is shown in Figure 10.

## 5. Conclusions

C-A-BP-NN, a BP-NN prediction technique that takes into account the sales features and patterns of the sample data, was recommended in this study. First, the C-BP-NN is introduced, which combines the clustering and BP-NNs in order to reflect the sales characteristics of commodities in the forecasting process. Based on certain criteria, the prediction may be made using the two-step clustering approach, which divides data into numerous groups. It is then time to build unique clusters' C-PB-NNs with the help of a BP-NN model. The proposed A-BP-NN utilizes ARIMA's strengths in forecasting the trend of data series while also addressing the

nonlinear component of the data series, which ARIMA's drawbacks cannot. By comparing AICs for various parameters, we can find the best ARIMA model, and then we can use it to forecast the linear component of the data series. The ARIMA residuals are used as input and output for the rolling prediction of nonlinear data series by the trained BP-NN. By combining the anticipated residuals from the BP-NN with ARIMA forecast values, we arrive at the final A-BP-NN results.

To summarize, the C-A-BP-NN is constructed by combining the strengths of the C-BP and A-BP-NNs and weighting them appropriately. This means that a concatenation of the findings of the two models is used to arrive at the final prediction values. It is necessary to make comparisons with other existing models in order to determine the efficacy of the C-A-BP-NN. The C-A-BP-NN's predicting ability is evaluated using four popular evaluation indices, including ME, RRMSE, RMSE, and MAE. According to the results of this experiment, the C-A-BP-NN surpasses all other methods, demonstrating that this model may be used as a guide when choosing forecasting models for an e-commerce firm.

As an example of future research areas, instead of using only one model for all commodities, the e-commerce corporation should use various models for each commodity. The two possible expansions have been proposed for further study. On the one hand, because there could be no model with minimal values for all assessment indicators, selecting the ideal model becomes challenging. In order to tackle this challenge, a thorough assessment index of predicting ability will be developed. When it comes to sales forecasting, various significant aspects like inventory costs, order lead times, delivery times, and transportation times are really taken into account in an effort to improve inventory management.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

## References

- [1] V. Irina, "Blockchain technology in the international supply chains," *Communication*, vol. 1, no. 1, pp. 16–25, 2020.
- [2] A. Ellinger, H. Shin, W. M. Northington, F. G. Adams, D. Hofman, and K. O'Marah, "The influence of supply chain management competency on customer satisfaction and shareholder value," *Supply Chain Management: An International Journal*, vol. 17, no. 3, pp. 249–262, 2012.
- [3] S. Ji, X. Wang, W. Zhao, and D. Guo, "An application of a three-stage XGBoost-based model to sales forecasting of a cross-border e-commerce enterprise," *Mathematical Problems in Engineering*, vol. 2019, Article ID 8503252, 15 pages, 2019.
- [4] K. Gilbert, "An ARIMA supply chain model," *Management Science*, vol. 51, no. 2, pp. 305–310, 2005.
- [5] P. L. Carter, J. R. Carter, R. M. Monczka, T. H. Slight, and A. J. Swan, "The future of purchasing and supply: a ten-year forecast 1," *Journal of Supply Chain Management*, vol. 36, no. 1, pp. 14–26, 2000.
- [6] X. Cai, Y. Qian, Q. Bai, and W. Liu, "Exploration on the financing risks of enterprise supply chain using back propagation neural network," *Journal of Computational and Applied Mathematics*, vol. 367, article 112457, 2020.
- [7] M. A. Camilleri, "The rationale for responsible supply chain management and stakeholder engagement," *Journal of Global Responsibility*, vol. 8, no. 1, pp. 111–126, 2017.
- [8] X. Zhao, "A study on the applications of big data in cross-border E-commerce," in *2018 IEEE 15th International Conference on e-Business Engineering (ICEBE)*, pp. 280–284, Xi'an, China, 2018.
- [9] A. Mahmoodzadeh, M. Mohammadi, K. M. Gharrib Noori et al., "Presenting the best prediction model of water inflow into drill and blast tunnels among several machine learning techniques," *Automation in Construction*, vol. 127, article 103719, 2021.
- [10] A. Mahmoodzadeh, M. Mohammadi, H. Hashim Ibrahim et al., "Machine learning forecasting models of disc cutters life of tunnel boring machine," *Automation in Construction*, vol. 128, article 103779, 2021.
- [11] M. R. Mosavi and M. Khishe, "The use of radial basis function networks based on leader mass gravitational search algorithm for sonar dataset classification," *Journal of Electronic and Cyber Defence*, vol. 4, no. 2, pp. 39–52, 2020.
- [12] M. R. Mosavi, M. Kaveh, and M. Khishe, "Sonar data set classification using MLP NN trained by nonlinear migration rates BBO," in *The fourth Iranian conference on engineering electromagnetic (ICEEM 2016)*, pp. 1–5, Nowshar, Iran, 2016.
- [13] A. Saffari, S. H. Zahiri, M. Khishe, and S. M. Mosavi, "Design of a fuzzy model of control parameters of chimp algorithm optimization for automatic sonar targets recognition," *IJMT*, vol. 9, no. 1, pp. 1–44, 2020, [http://ijmt.iranjournals.ir/article\\_241126.html](http://ijmt.iranjournals.ir/article_241126.html).
- [14] M. Khishe, M. R. Mosavi, and A. Moridi, "Chaotic fractal walk trainer for sonar data set classification using multi-layer perceptron NN and its hardware implementation," *Applied Acoustics*, vol. 137, pp. 121–139, 2018.
- [15] M. R. Mosavi, M. Kaveh, M. Khishe, and M. Aghababaie, "Design and implementation a sonar data set classifier using multi-layer perceptron NN trained by elephant herding optimization," *IJMT*, vol. 5, no. 1, pp. 1–12, 2018.
- [16] Y. Wang, L. P. Yuan, M. Khishe, A. Moridi, and F. Mohammadzade, "Training RBF NN using sine-cosine algorithm for sonar target classification," *Archives of Acoustics*, vol. 45, 2020.
- [17] M. Kaveh, M. Khishe, and M. R. Mosavi, "Design and implementation of a neighborhood search biogeography-based optimization trainer for classifying sonar dataset using multi-layer perceptron neural network," *Analog Integrated Circuits and Signal Processing*, vol. 100, no. 2, pp. 405–428, 2019.
- [18] C. Wu, M. Khishe, M. Mohammadi, S. H. T. Karim, and T. A. Rashid, "Evolving deep convolutional neural network by hybrid sine-cosine and extreme learning machine for real-time COVID19 diagnosis from X-ray images," *Soft Computing*, pp. 1–20, 2021.
- [19] T. Hu, M. Khishe, M. Mohammadi, G.-R. Parvizi, S. H. T. Karim, and T. A. Rashid, "Real-time COVID-19 diagnosis

- from X-Ray images using deep CNN and extreme learning machines stabilized by chimp optimization algorithm,” *Bio-medical Signal Processing and Control*, vol. 68, article 102764, 2021.
- [20] M. Khishe and M. R. Mosavi, “Chimp optimization algorithm,” *Expert Systems with Applications*, vol. 149, article 113338, 2020.
- [21] M. Khishe, M. Nezhadshahbodaghi, M. R. Mosavi, and D. Martín, “A weighted chimp optimization algorithm,” *IEEE Access*, vol. 9, pp. 158508–158539, 2021.
- [22] W. Kaidi, M. Khishe, and M. Mohammadi, “Dynamic levy flight chimp optimization,” *Knowledge-Based Systems*, vol. 235, article 107625, 2021.
- [23] M. Khishe and A. Safari, “Classification of sonar targets using an MLP NN trained by dragonfly algorithm,” *Wireless Personal Communications*, vol. 108, no. 4, pp. 2241–2260, 2019.
- [24] I. Alon, M. Qi, and R. J. Sadowski, “Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods,” *Journal of Retailing and Consumer Services*, vol. 8, no. 3, pp. 147–156, 2001.
- [25] G. Di Pillo, V. Latorre, S. Lucidi, and E. Procacci, “An application of support vector machines to sales forecasting under promotions,” *4OR*, vol. 14, no. 3, pp. 309–325, 2016.
- [26] X. Luo, J. Sun, L. Wang et al., “Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy,” *IEEE Trans. Ind. Informatics*, vol. 14, no. 11, pp. 4963–4971, 2018.
- [27] C.-J. Lu and L.-J. Kao, “A clustering-based sales forecasting scheme by using extreme learning machine and ensembling linkage methods with applications to computer server,” *Engineering Applications of Artificial Intelligence*, vol. 55, pp. 231–238, 2016.
- [28] W. Dai, Y.-Y. Chuang, and C.-J. Lu, “A clustering-based sales forecasting scheme using support vector regression for computer server,” *Procedia Manufacturing*, vol. 2, pp. 82–86, 2015.
- [29] I.-F. Chen and C.-J. Lu, “Sales forecasting by combining clustering and machine-learning techniques for computer retailing,” *Neural Computing and Applications*, vol. 28, no. 9, pp. 2633–2647, 2017.
- [30] A. Freij, K. Walid, and M. Mustafa, “Deep learning model for digital sales increasing and forecasting: towards smart E-commerce,” *Journal of Cybersecurity and Information Management*, vol. 8, no. 1, pp. 26–34, 2021.
- [31] T. Zhang, R. Ramakrishnan, and M. Livny, “BIRCH: a new data clustering algorithm and its applications,” *Data Mining and Knowledge Discovery*, vol. 1, no. 2, pp. 141–182, 1997.