Hybrid Functional Link Neural Networks for Soybean Price Forecast

S. Dhanalakshmi[1](✉), S. Rajakumar[2], A. S. Anakath[3], R. Kannadasan[4], and S. Ambika[3]

[1] Department of Master of Computer Applications, Meenakshi Ramasawamy Engineering College, Ariyalur, Tamilnadu, India
sadhanamresearch@gmail.com

[2] Department of Computer Science and Mathematics, University College of Engineering, Ariyalur, Tamilnadu, India

[3] Department of Information Technology, E.G.S. Pillay Engineering College, Nagapattinam, Tamilnadu, India

[4] SCOPE, VIT University, Vellore, Tamilnadu, India
kannadasan.r@vit.ac.in

Abstract. Drastic change in crop prices is observed due to climatic changes, natural calamities and lack of quantity of a specific commodity. Crop price prediction plays key role in effective farm management. Farmers are not able to predict these crop prices and facing massive loss. These aspects pressure us to use advanced technology and develop accurate, reliable and efficient crop price prediction system. Crop price prediction also helps agriculture based industries and policymakers. There are many price-sensitive crops like tomatoes, onions, potatoes, Soybean and other food grains, which need prior price prediction so that farmers can take wise decisions on which crop to cultivate. Functional link neural network is chosen to develop Basic network for Soybeans price prediction. Optimization algorithms like whale optimization, particle swarm optimization and Harris Hawks Optimization are used to calculate appropriate biases and weights. Dataset is taken from daily reports issued by Chicago Mercantile Exchange (CME).Most efficient hybrid FLNN with associated Expansion function, activation function and learning scheme for predicting crop price could be found out through our study.

Keywords: Functional Link Neural Networks · Whale Optimization Algorithm · Particle Swarm Optimisation Algorithm · Harris Hawks Optimization Algorithm

1 Introduction

Due to climatic changes, natural calamities and also when there is less quantity of a particular commodity, crop prices vary in a larger scale, then its prices increase. These things hinder the use of advanced technology in agriculture. The estimation of crop prices plays a central role in successful farm management. Farmers cannot foresee the prices of these crops and face massive losses. For future, forecasting the price that a given

S. Tamane et al. (Eds.): ICAMIDA 2022, ACSR 105, pp. 569–581, 2023.
https://doi.org/10.2991/978-94-6463-136-4_48

crop would produce is extremely valuable when deciding which crop types to promote and grow. Several variables are used to predict future prices for a given crop which are just not limited to historical pricing but also Area planted, Area Harvested, production and crop health indicators. These and other potentially predictive variables are gathered from various sources of data which includes Temperature and Production. Functional link neural networks will be built using various Expansion functions like Chebyshev, legendre, laguarre, power and trigonometric polynomials using various activation functions like Rectified Linear Unit and Exponential Linear Unit. Optimization algorithms like whale optimization, particle swarm optimisation and Harris Hawks Optimization algorithms will be used to calculate appropriate biases and weights. The dataset is taken from reliable source which support dynamic change and accuracy.

2 Literature Survey

In [1] PSO-BP neural network is used to forecast the price of Green pepper in Danzhou. To optimize the initial weights and BP (Back propagation) threshold's PSO (Particle swarm optimization) algorithm is used. This model overcomes the local minima problem and over-fitting problem faced by traditional BP neural networks. Training error is reduced and effectively increased predicting precision compared with traditional BP neural networks. Time cost is more for BPNN. Particle swarm optimization algorithm (PSO) has very low convergence rate while iterating, Modern optimization algorithms can be used.

In [2] Auto regressive Integrated Moving Average (ARIMA) model is used for forecasting medium rice price. Two ARIMA models ARIMA (2, 0, 2) and ARIMA (1, 1, 2) are compared. ARIMA (1, 1, 2) is proved to give more accurate results than ARIMA (2, 0, 2). Limitation of ARIMA model is that it is accurate only for stationary time series and it requires minimum number of inputs to get accurate results. ARIMA is not robust method and time has to be updated at every iteration, there are many models for time series forecast so in future robust forecasting model could be developed.

In [3] Market price of tomato is predicted using Back-propagation neural network (BPNN) and Radial basis function neural network (RBF), Optimization Algorithms like Levenberg-Marquardt algorithm and gradient decent algorithm are compared and Levenberg-Marquardt algorithm was chosen. To measure neural network accuracy Mean square error is used. Accuracy of BPNN is 77.42% whereas accuracy of RBF is 85.55%. It is concluded that RBF neural network is more accurate and efficient when compared to Back-propagation neural net- work. Limitations of BPNN and RBFNN are time co stand Large MAPE value respectively. Optimisation techniques can be used to increase accuracy of model.

In [4] Prices of cabbage, bokchoy, watermelon, and cauliflower are predicted using nearest neighbour (NN), inverse distance weighting (IDW), Kriging method with Partial Least Square (KPLS), and artificial neural network (ANN). For calculating the accuracy mean absolute percentage error (MAPE) is used. Kringing method is proved to give more accuracy.Limitation of this paper is that feature selection is not effective. Temporal and Spatial features like climatic conditions, planting area location, and past trading price could be considered. In addition, Algorithms like particle swarm optimization, ant colony system could be used to select features effectively.

In [5] Grey Prediction Method GM (1,1) issued to predict prices of various agricultural products and is compared with Radial basis function neural network (RBF). To calculate accuracy mean absolute percentage error (MAPE) is used. For stable data it is proved that GM (1,1) is better RBF and for data which is not stable RBF is better. Limitation of GM (1,1) is that its MAPE value is very large for data which is not stable. Alternatives for GM (1,1) should be developed as Grey Prediction Method is unstable for large data.

In [6] chilli price is predicted using K-Nearest Neighbour (KNN) algorithm and Adaptive Synthetic (ADASYN) algorithm was used to overcome the issue of imbalanced classes. Using KNN algorithm without ADASYN produced high accuracy but low F1score, whereas using classification with ADASYN produced 100% accuracy and F1 score. Limitation of KNN algorithm is that it is not suitable for large datasets and ADASYN doesn't identify instances which are noisy. Modern classification algorithm could be used, as KNN is not efficient for large datasets.

In [7] Hybrid model is proposed where seasonal-trend decomposition procedures based on loss (STL) and extreme learning machines (ELMs) are combined for predicting short-, medium-, and long-term vegetable prices. STL-ELM model is compared with Seasonal Auto regressive Integrated Moving Average (SARIMA), Time delay neural network (TDNN), Support vector regression (SVR), ELM, and SARIMA-Kalman filter. Symmetric mean absolute percentage error (SMAPE) and MASE are the criteria for calculating accuracy. It is proved that STL-ELM is the best model and SARIMA is the worst model. Limitation of this paper is that it is restricted to point forecasting, interval forecasting is more preferred when compared to point forecasting. As future work proposed model could be applied using interval forecasting.

In [8] Radial Basis Function (RBF) Neural Network and Back Propagation (BP) Neural network are used to predict the price of ten agriculture products of china and both are compared. Mean Absolute Errors (MAE), Mean Absolute Percentage Errors (MAPE), Mean Square Error (MSE) are used to calculate the accuracy. It is proved that BP neural network has more accuracy when compared to RBF neural networks and time cost is more for BPNN. Limitations of BPNN and RBFNN are time cost and Large MAPE value respectively. New methods are to be proposed to reduce time cost of BP Neural network and MAPE on the RBF Neural network.

In [9] Artificial neural network method is applied to predict chilli price in Indonesia. Artificial neural network method is compared with Holt-Winter model Mean Absolute Percentage Errors (MAPE) issued to calculate the accuracy. It is proved that accuracy of Holt- winter model is greater than ANN method. Limitation of ANN is that its accuracy is less because of various fluctuations in data pattern of variables. NA As future work ANN model could be combined with any other forecasting method.

In [10] Auto regressive integrated moving average (ARIMA) is used as Feature selection method along with support vector regression (SVR), artificial neural net-works (ANNs) and multivariate adaptive regression splines (MARS). Root mean square error (RMSE), Mean absolute percentage error (MAPE), and mean absolute error (MAE) are used to calculate accuracy. Accuracy of Proposed model is compared with MARS, ARIMA, SVR; ANN It is proved that proposed model is more accurate. NA Integrated models can combined with time delay neural networks, extreme learning machines, and artificial immune systems.

In [11] this research uses the RNA to develop an early warning system for facing the increase in agricultural products, considering macro and micro economic variables and factors related to the seasons of the year. Can predict prices with a good accuracy and the use of RNA enhance the percentage of right prediction. Only monitors spikes and does not give any specific price prediction value. This paper only focuses on variables and factors effecting seasons of the year. All the other attributes which effect price variation are not considered.

In [12] Agricultural product price is predicted using K-neighbourhood, random forests, support vector machines and BP neural networks. Accuracy is compared using MAPE values and it is proved that BP neural networks model gives more accurate results. To overcome issues of computational complexity and over fitting in BPNN GA algorithm is used to optimize BP neural networks. Limitation of GA algorithm is that fitness function should be chosen with care. There is no standard data collection method when it comes to amount of crop produced. Amount sold or amount present in storage or proper storage.

In [13] BP neural network model is used to predict agriculture price prediction and MIV (Mean Impact Value) method is used to find strong features and then Variable Learning Rate Algorithm and Momentum Back Propagation are used for BP network design training. Find strong features NA Focuses on calculation methodology and not on precise dynamic data changes. Maximum attributes are not taken into consideration.

In [14] Usage of ARIMA model to predict garlic prices. Using the powerful data analysis function of R language, forecast the monthly average price of garlic in the first half of 2018 in Shandong province is measured. Suggests some changes in the forecast methods and strengthening market vision only applicable for garlic in Shandong province. This paper only focuses on soybean price pertaining to China and attributes relevant to it. If any other crop is used the factors considered may be susceptible to changes.

In [15] Using quantile regression models to describe the distribution of the soy-bean price range; and also using RBF neural networks to approximate the non-linear component of the soybean price. Almost all factors taken into consideration when calculating crop price. Only applies for soybeans. Here the same regression model is applied but on soybeans can be extended to other regions data as well.

In [16] a model is created based on 3 endogenous modules called upstream farms, downstream farms, and groundwater and 2 exogenous models called surface water and crop prices. This can be used on any type of crop. Just a theoretical model with no experimental validation is given for the model. Other factors such as climate change and amount of production spent is not included in the research and also better learning techniques can be employed using hybrid neural models.

In [17] Analysis of data mining techniques which are applied to agriculture and their applications to agriculture related areas is described. Gives comparison between different algorithms used. Nothing new or novel is proposed. It is just a review. Was a review only pertaining to efficiency but never was there more extensive research on better data collection and maintenance in terms of agricultural production and monitoring.

3 Methodology

3.1 Functional Link Neural Network

Functional link neural network is one of the higher order neural networks. Although they have linear nature, non-linear input-output relationships can be captured by FLNN's if acceptable polynomial inputs are provided. FLNN doesn't contain any hidden layers, so the learning algorithm is less complicated. FLNN's are suitable for time series data predictions. Output of node is computed using following equation.

$$p = f(b + W.Xt)$$

where p is the predicted output, b is the bias, W is the weight vector, X is the expanded input vector f and is the activation function.

3.2 Whale optimization Algorithm

Whale Optimisation Algorithm [18] is an extension to meta-heuristic algorithms developed by Seyed ali Mir jalili and Andrew Lewis which is based on the hunting habit of hump back whales known as bubble net hunting strategy. Hump back whales use two mechanisms for recognising the location of target prey and then attacking them. They encircle the target prey and then form bubble nets.

Encircling Prey

Whale optimisation algorithm defines the current candidate solution as prey. After defining best search agent, other search agents will keep on updating their positions towards already defined best search agent. This is represented by following mathematical equations:

$X(t + 1) = X * (t) - A.D$

$D = |c.x * (t) - X(t)|$

here t denotes the present iteration, X(t) and $X * (t)$ represent the position vector and the position vector of the optimum solution respectively.

$A = 2a.r$ a $C = 2.r$

here r [0, 1] is a randomly produced variable and "a" is a variable decreasing from 2 to 0. A vector typically takes part in the exploration so as to cover the entire search space.

Bubble-Net Attacking Technique

This is the exploitation phase and it has two approaches shrinking encircling mechanism and spiral updating of position. In shrinking encircling mechanism value of "a" is decreased from 2 to 0 and values of "A" lies between -1 and 1 thus reducing the circle made by whale for catching target prey. In spiral updation of position mechanism distance between whale and prey is calculated and spiral equation is created which is represented mathematically as.

$X(t + 1) = D.eb1. \mathrm{Cos}(2\pi 1) + X * (t)$

where $D = X * (t)$ X(t) is the distance between whale and prey, b is a constant and t is random number in $[-1, 1]$.

Humpback whales may follow either shrinking encircling mechanism or spiral updation of position mechanism with 50% probability which is represented mathematically as.

$X(t + 1) = X * (t) - A.D$ when $p < 0.5$
$= D.eb1. \cos(2\pi 1) + X * (t)$when $p \geq 0.5$

3.3 Harris Hawks Optimization Algorithm

HHO comes under swarm intelligence algorithms. It has been introduced recently and is inspired by nature. It was proposed by Heid ari et al. in 2019 inspired from the natural behaviour of Harris hawks. The chasing style of Harris hawks in order to trap the prey is very distinctive. HHO comes under population-based algorithm because in it many hawks (each one representing a candidate solution) cooperate with each other in order to follow the prey location (the fittest candidate solution) by using different chasing styles. The algorithm has good performance in handing many unconstrained and constrained problems Exploration and exploitation is balanced by this algorithm by six updating phases. It has two phases of exploration and four phases of exploitation.

Exploration phases: In this phase, two strategies are used to update the search agents. Between two updating methods one is selected with the probability of 50%. Based on randomly selected agent or based on the optimum or best solution obtained so far agents are updated. Transformation from exploration to exploitation: HHO model has an additional parameter which is escaping energy (E). It decreases with time for the transformation from exploration to exploitation. This idea is inspired from escaping rabbit as its energy decreases with time. Exploitation phases: HHO uses many mechanisms that are time-varying with a greedy scheme to model the exploitation phase. Depending on the energy (E) and the escaping behaviours of the rabbit one among four updating mechanisms is chosen. Soft besiege, hard besiege, soft besiege with progressive rapid dives, and hard besiege with progressive rapid dives are the employed phases.

3.4 Particle Swarm Optimization Algorithm

In computational science, particles warm optimization (PSO) is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. It solves a problem by having a population of candidate solutions, here dubbed particles, and moving these particles around in the search space according to simple mathematical formula over the particle's position and velocity. Each particle's movement is influenced by its local best known position, but is also guided toward the best known positions in the search-space, which are updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions.

$Vk + 1 = \omega V k + c1r1Pk - Xk + c2r2Pk - Xk$

where ω is inertia weight, $d = 1, 2, D$ and $j = 1, 2$ n, denotes current iterations, Vid is the particle velocity, c1 and c2 are learning factors, r1 and r2 are random number between0and1.Pi denotes individual extreme value and Pg denotes populations global extreme value. To prevent the particle's blind search, the position is generally limited in a certain range [−Xmax, Xmax], and the velocity is limited in a certain range − [Vmax, Vmax].

4 Prediction Process

4.1 Collection of Data

Dataset is taken from kaggle website; data for the dataset is taken from daily reports issued by Chicago Mercantile Exchange (CME). It contains data of Soy-bean prices and factors that affect soybean prices from 1962 to 2018.

4.2 Data Pre-processing

The dataset which we use for prediction contains enormous amounts of raw data around 21,800 samples which needs to be pre-processed. The null values are managed and duplicate records are removed. To observe the relationship between attributes we found the correlation matrix and highly correlated attributes are considered.

4.3 Optimizing FLNN Using Whale optimization, Harris Hawks Optimization, Particle Swarm Optimization

Functional link neural network is chosen to develop Basic network for prediction. Optimization techniques are to solve the issue of local minima. Optimization algorithms like whale optimization, Particle swarm optimization and Harris Hawks Optimization algorithms are used to calculate appropriate biases and weights.

5 FLNN—WOA

Inputs are fed to FL Neural Network and Price and output price is generated.

1. Predicted price is compared with actual price and MSE, MAE are calculated.
2. Whale optimization algorithm is used to minimize value of MSE.
3. Training is stopped when the lowest value of MSE is reached.
4. Input to the whale optimization algorithm is the objective function. First positions are initialized; evaluation is done by calculating fitness. Initially global best is initialized. If current best is better than global best, it is updated as global best. Finally [1–18] Global best is returned (Fig. 1, Fig. 2 and Fig. 3).

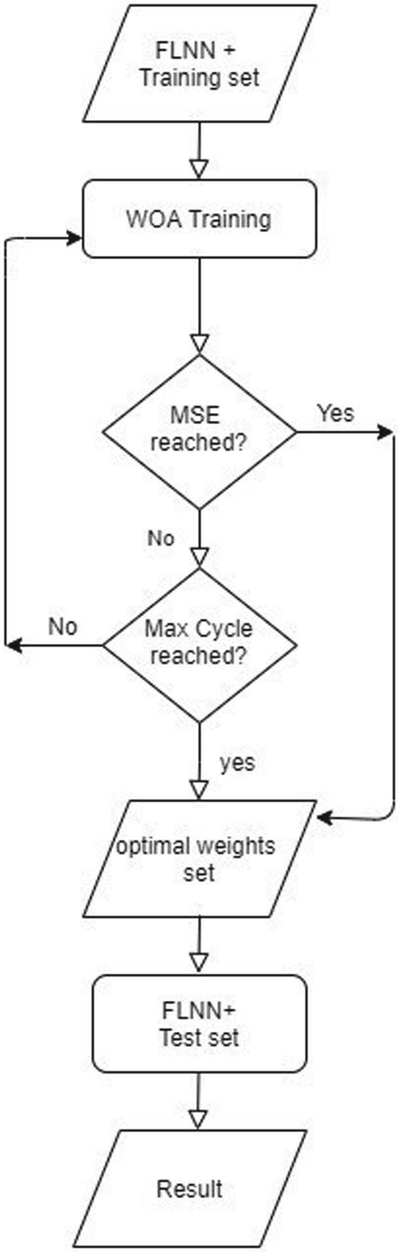

Fig. 1. Training scheme for FLNN and WOA

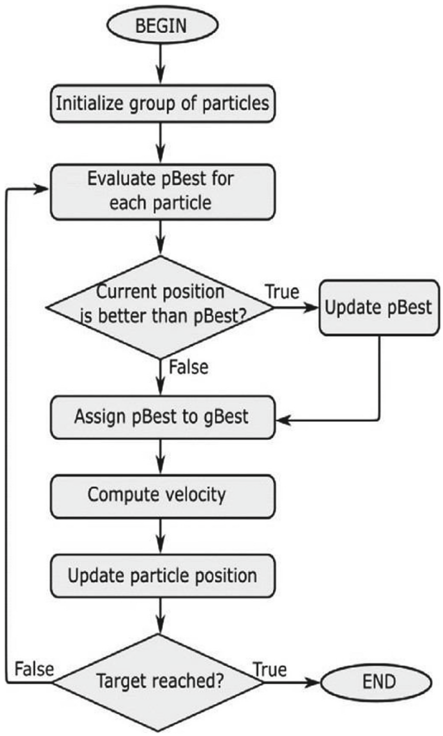

Fig. 2. Training scheme for FLNN and PSO

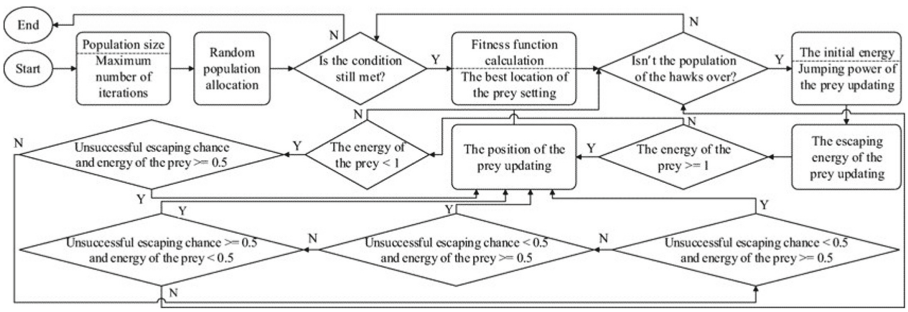

Fig. 3. Training scheme or FLNN and HHO

6 FLNN—PSO

1. Inputs are fed to FL Neural Network and Price and output price is generated.
2. Predicted price is compared with actual price and MSE, MAE are calculated.
3. Particle swarm optimization algorithm is used to minimize value of MSE.
4. Training is stopped when the lowest value of MSE is reached.

Table 1. Expansion function influence on FL-WOANN

Expansion function	MSE	MAE
Chebyshev	0.0057	0.0500
Legendre	0.0020	0.0297
Laguerre	0.0176	0.0938
Power	0.0105	0.0794
Trigonometric	0.0401	0.1617

7 FLNN—HHO

1. Inputs are fed to FL Neural Network and Price and output price is generated.
2. Predicted price is compared with actual price and MSE, MAE are calculated.
3. Harris hawks optimization algorithm is used to minimize value of MSE.
4. Training is stopped when the lowest value of MSE is reached.

8 Testing the Model

The Mean Squared Error (MSE), Mean Absolute Error (MAE) were computed to evaluate performance of FLNN — WOA, FLNN — HHO and FLNN — PSO. The two measures are defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y_i}\right)^2$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} = \frac{\sum_{i=1}^{n}|e_i|}{n}$$

Here n is the total number of test dataset and e_t is the difference between actual and predicted price of the t^{th} input.

9 Comparing Efficiencies

Three different optimization algorithms are applied for training a function all in k neural network, whale optimization algorithm (WOA), Particles warm optimization (PSO) and Harris Hawks optimization (HHO). MSE(Mean Square error) AND MAE(Mean absolute error) are computed. We tested expansion functions including Laguerre, Chebyshev, Trigonometric, Power and Legendre. Analysis of FLNN-WOANN is represented in Table 1 and Legend reproduces least MSE and MAE values 0.0020 and 0.0297 respectively after 50 epochs. Analysis of FLNN-PSONN is represented in Table 2 and power function gave least MSE and MAE values 0.0002 and 0.0095 respectively after 50 epochs. Analysis of FLNN-HHONN is represented in Table 3 and Legendre gave least MSE and MAE values 0.4757 and 0.6640 respectively after 50 epochs.

Table 2. Expansion function influence on FL-PSONN

Expansion function	MSE	MAE
Chebyshev	0.0004	0.0138
Legendre	0.0006	0.0138
Laguerre	0.0006	0.0183
Power	0.0002	0.0095
Trigonometric	0.0029	0.0363

Table 3. Expansion function influence on FL-HHONN

Expansion function	MSE	MAE
Chebyshev	0.5286	0.7058
Legendre	0.4757	0.6640
Laguerre	0.5209	0.6993
Power	0.5094	0.6907
Trigonometric	0.5287	0.7058

10 Conclusion

On comparing all the hybrid models developed it can be concluded that Function all in k neural network with power function as expansion function and Particles warm optimization algorithm as learning algorithm is found to give optimal results with least MSE and MAE values when compared to Function all in k neural network with Harris Hawks optimization Algorithm and whale optimization algorithm.

As future work, we are considering studying several other optimization algorithms with Functional link neural network for Price prediction.

References

1. Yu, Y. Zhou, H., Fu, J.: Research on agricultural product price forecasting model based on improved BP neural network. J Ambient Intell Human Comput (2018)
2. S, H., Nhita, F., Saepudin, D., Aditsania, A.: Chili Commodity Price Forecasting in Bandung Regency using the Adaptive Synthetic Sampling (ADASYN) and K-Nearest Neighbor (KNN) Algorithms. In:2019 International Conference on Information and Communications Technology (ICOIACT). pp. 434–438(2019)
3. Wang, B., Liu, P., Zhang, C., Wang, J., Peng, L.: Prediction of Garlic Price Based on ARIMA Model. Lecture Notes in Computer Science 11063(2018)
4. Hemageetha, N., Nasira, G.M.: Radial basis function model for vegetable price prediction. In:2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering. pp. 424–428(2013)

5. Yuehjen, E., Shao, J.T., Dai: Integrated Feature Selection of ARIMA with Computational Intelligence Approaches for Food Crop Price Prediction. Complexity 2018(2018)
6. Minghua, W., Qiaolin, Z., Zhijian, Y., Jingui, Z.: Prediction model of agricultural product's price based on the improved BP neural network. In: " 2012 7th International Conference on Computer Science & Education (ICCSE). pp. 613–617(2012)
7. Lu, Y., Yuping, L., Weihong, L., Qidao, S., Yanqun, L., Xiaoli, Q.: Vegetable Price Prediction Based on PSO-BP Neural Network. 8th International Conference on Intelligent Computation Technology and Automation (ICICTA) pp. 1093–1096(2015)
8. Zhang, D., Zang, G., Li, J., Ma, K., Liu, H.: Prediction of soybean price in China using QR-RBF neural network model. Computers and Electronics in Agriculture 154, 168–1699 (2018)
9. Ohyver, M., Pudjihastuti, H.: Arima model for forecasting the price of medium quality rice to anticipate price fluctuations. Procedia Computer Science 135, 701–711(2018)
10. &'Kaur, M.., Gulati, Heena, Kundra, H.: Data Mining in Agriculture on Crop Price Prediction: Techniques and Applications. International Journal of Computer Applications 99 (2014)
11. Anggraeni, W.: Agricultural Strategic Commodity Price Forecasting Using Artificial Neural Network. In: Research of Information Technology and Intelligent Systems (ISRITI). pp. 347–352(2018)
12. Peng, Y., Hsu, C., Huang, P.: An investigation of special approaches for crop price forecasting in different Taiwan markets. In: 2015 Conference on Technologies and Applications of Artificial Intelligence (TAAI). pp. 176–179(2015)
13. Zong, J., Zhu, Q.: Apply Grey Prediction in the Agriculture Production Price .In:2012 Fourth International Conference on Multimedia Information Networking and Security.pp.396–399 (2012)
14. Saeed, P., Langarudi, C.G., Silva: Crop Price Volatility and its Influence on Crop Patterns. In: Proceedings of the 2017 International Conference of the Computational Social Science Society of the Americas. vol. 4, pp. 1–10. Association for Computing Machinery(2017)
15. 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. Neuro computing pp. 2831–2844(2018)
16. Zong, J., Zhu, Q.: Price forecasting for agricultural products based on BP and RBF Neural Network. Computer Science and Automation Engineering pp. 607–610(2012)
17. Silva, J.: An Early Warning Method for Basic Commodities Price Based on Artificial Neural Networks. In: H., L., H., T., Z., W.(eds.)Advances in Neural Networks-ISNN2019.ISNN2019. vol. 11554. Springer(2019)
18. Mirjalili, S., Lewis, A.: The Whale Optimization Algorithm. Advances in Engineering Software 95, 51–67(2016)