RESEARCH ARTICLE

Macromolecular Symposia

www.ms-journal.de

Modeling Consumer Price Index: A Machine Learning Approach

Pradeepta Kumar Sarangi, Ashok Kumar Sahoo, and Sachin Sinha*

The change in price of a group of goods and services is reflected in terms of consumer price index (CPI), making it one of the most important economic indicators. This is also the mostly used measure of inflation. Forecasted CPI values help the Government to take corrective measures to control the economic conditions of the country. This paper implements and examines two machine learning models such as artificial neural network (ANN) and ANN model optimized with particle swarm optimization (PSO) known as ANN-PSO to assess the accuracy in predictability of CPI. The data set for four groups such as food and beverages, housing, clothing, and footwear used for the calculation of all India CPI has been taken from the official website of the Government of India. The mean absolute percentage error (MAPE) has been used as the validator for model accuracy. The MAPE calculated for all experiments are less than 10% which indicates that the ANN-PSO models used are highly accurate for prediction of CPI of India.

1. Introduction

Consumer price index (CPI) is a measure of change of selected goods and services over a period of time. Initially, it was used to provide a basis for increase in compensation of the workers so that the rising cost of living could be compensated to some extent. Later on, it came to be recognized as one of the biggest indicators for the inflation in the economy. CPI is not only used as the microeconomic indicator for inflation but is also used by various government bodies for monitoring price stability, deflators in national accounts, and many other areas.

All India CPI is released by the National Statistics Office, Ministry of Statistics and Program Implementation (MSPI) on the base 2012 = 100. The office subsequently also releases the consumer food price index (CFPI), both of which are considered to be the most significant indicators of inflationary tendencies in the market.

Forecasting inflation, both in the short and medium term, has always been a cause of concern for the economic policy makers and for the timely implementation of the monetary policies. Economic forecasting uses historical data points that have been released in the previous years to predict the future trend. More sophisticated is the method used for forecasting, better are the estimates and more effective is the policy formulation. The scope of machine learning tools has been ever increasing especially for forecasting purposes. It has been used in a variety of areas like stock index trading decision,[1] electrical load forecasting,[2,3] forecasting CPI,[4] Indian export predictions,[5] trend analysis for Indian automobile industry,[6] Forex trend analysis,[7] and many other areas.[8,9]

For forecasting CPI, a number of methods like vector error correction model (VCEM) and dynamic factor model (DFM), ARIMA, additive ARIMA, time series analysis, artificial neural network (ANN), AI, and Machine Learning tools have been used. This paper makes an attempt to make use of ANN and ANN-PSO models for forecasting CPI.

2. Review of Literature

A number of studies have been conducted and methods have been tested to come up with the most accurate predictive model for CPI. In ref., [10] the researchers have used VECM and DFM to forecast CPI inflation. According to the authors, the performance of VECM model was rated slightly better than the DFM model but the same predictive accuracy was associated with both the models through the Diebold–Mariano test.

Box–Jenkins ARIMA technique was used to model and forecast CPI in Belgium using time series data from 1960 to 2017. ARIMA (0, 2, 1) model was used for predicting CPI and it was the stable and acceptable model with optimal outcome.[11]

In another study,[12] authors report that monthly predictions of CPI are better with ARIMA (12, 1, 12) whereas in another work,[13] the additive ARIMA*ARIMAS model was used to analyze the trend of the CPI and to forecast the CPI for the coming months. It was found that the Holt-Winters model exhibited minimum error. Hybrid ANN-PSO models were used for forecasting the results and these were compared with the ARIMA, hybrid ANN-GA, ANN-BP, and linear models. It was found that the

P. K. Sarangi
Chitkara University Institute of Engineering and Technology
Chitkara University
Punjab 140401, India

A. K. Sahoo
Graphic Era Hill University
Dehradun 248002, India
E-mail: ashoksahoo2000@yahoo.com

S. Sinha
Christ (Deemed to be University)
Delhi 201003, India

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/masy.202100349

DOI: 10.1002/masy.202100349

Table 1. Interpretation of MAPE results (forecasting accuracy).

MAPE-value	Accuracy of the model
<10%	Very high
>10% and ≤20%	Good
>20% and ≤50%	Reasonable
>50%	Inaccurate

Source: Kasemset et al.[28]

hybrid ANN models and ANN-PSO models are the best fit models based on RMSE and mean absolute percentage error (MAPE) using time series data.[14] The ANN-PSOCoG model has also been used to create a network and to predict the closing price of stocks with accuracy.[15]

In order to identify personalized recommendations for users on the basis of their preferences and behavior, the Rating Filtering method, Movie Categories Based Filtering–Singular Value Decomposition (MCBF-SVD) algorithm was used and it was established that MCBF-SVD is highly effective in reducing errors of rating prediction models.[16] The prediction accuracy of deep convolutional neural networks was found to be more than 90% accurate,[17] in ref.[18] generative convolutional network was used and it was identified that the model having components: artificial feature generator, artificial label generator, and traditional DCNN model for prediction of movie ratings outperformed other baseline models. ANN classifiers were used to predict financial solvency of the companies and it was found that the accuracy provided by ANN was 88% and sensitivity rate was 96%.[19]

Researchers have undertaken a number of studies to assess the working of ANN-PSO models in the prediction of various economic and financial indices. A study was undertaken to predict the future price of Bitcoins using PSO, and the results were quite accurate[20] and flexible neural tree (FNT) was used for predicting the stock market values.[21] Multilayer perceptron (MLP) and long short-term memory networks were used to predict stock market indices of Bombay Stock Exchange (BSE)[22] while the Firefly algorithm was used to develop a new stock market prediction method for BSE Sensex, NSE Sensex, S&P 500, and FTSE.[23]

In order to understand the application of PSO to predict the prices in the stock market, its various modifications were taken into consideration.[24] To highlight buy order, sell order, or order for holding the stocks in the New York Stock Exchange (NYSE) and in Stock Exchange of Thailand (SET), PSO algorithm has been used to select the optimally weighted signals.[25] PSO has also been used to forecast the stock indices of S&P 500 and DJIA for the short and long term. Adaptive linear combiner (ALC) was used in the model, in which PSO modified its weights. It was compared with the MLP-based model and PSO-based model was found to be better.[26] In general, it can be concluded that PSO-based models fare much better as compared to other models.

3. Objective

This paper aims to design and implement neural network-based machine learning models optimized with PSO and to compare the results with simple ANN models in predicting the CPI of India.

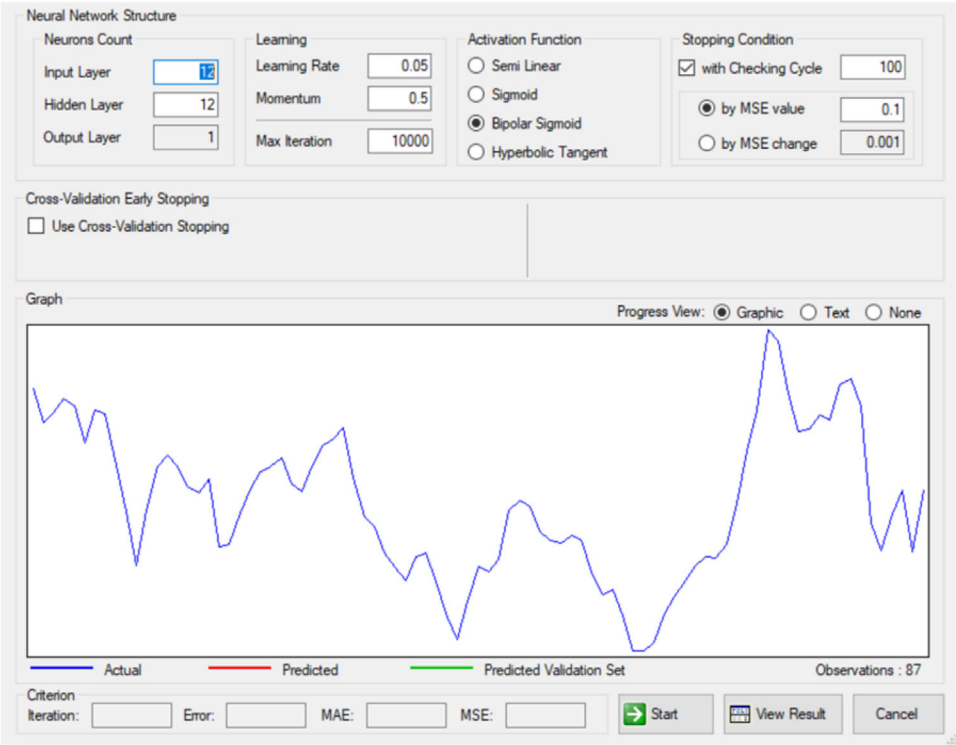

Figure 1. Training data set and parameters selection (food and beverages).

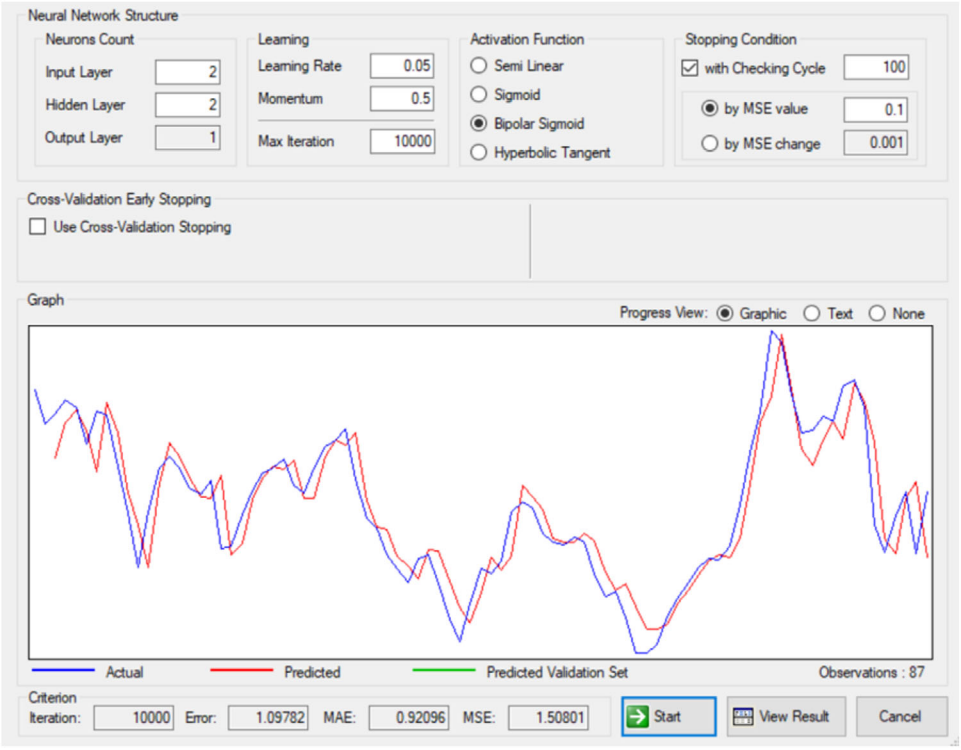

Figure 2. Training results.

Model Summary

	Value
Variable	bev
Included Observation	85 (After Adjusting Endpoints)
Network Archiecture	
Input Layer Neurons	2
Hidden Layer Neurons	2
Output Layer Neurons	1
Activation Function	BipolarSigmoidFunction
Back Propagation Learning	
Learning Rate	0.05
Momentum	0.5
Criteria	
Error	1.097820
MSE	1.508010
MAE	0.920956

Figure 3. Model summary.

4. Data Source

For the purpose of designing the model, secondary data (monthly CPI data for the period January 2013 to May 2021 with base year 2012) has been taken from MSPI.[27]

5. Implementation Design

The implementation is divided into two phases. In phase-1, simple neural network models have been implemented for various data sets of four categories namely, food and beverages, clothing and footwear, housing, and general index (all groups). In phase-2, a PSO model has been implemented using the same data set. Finally, results from both the implementations have been compared and analyzed.

6. Validation of the Models

According to Kasemset et al.,[28] the accuracy of a model can be validated in terms of the MAPE values which can be calculated by the formula in Equation 1.

$$\text{MAPE} = \frac{1}{n}\sum_{i}^{n} = 1\frac{|O_i - P_i|}{O_i} \times 100 \quad (1)$$

where P_i is the predicted value and O_i is the observed (actual) value.

The interpretation of model accuracy based on MAPE values are as in **Table 1**.

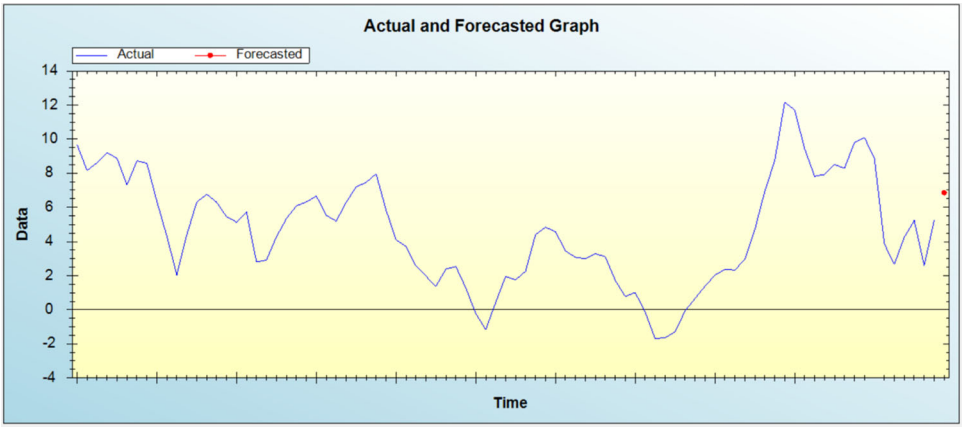

Figure 4. Actual and forecasted graph.

Table 2. Summary of experimental results.

Data set	MSE training	MAE training	MSE forecasted	MAPE forecasted	Accuracy of the model
Food and beverages	1.50	0.92	0.39	19.29	Good
Clothing and footwear	0.09	0.19	0.07	4.5	Very high
Housing	0.09	0.22	0.06	6.9	Very high
General index (all groups)	0.09	0.24	0.23	9.32	Very high

7. Phase-1: Implementation of ANN

In this phase, a simple neural network model has been implemented. The network is first trained with the training set and then it is tested for the test patterns and the error is calculated.

The monthly data for the period January 2014 to May 2021 has been used as training patterns and the value for the June 2021 is calculated through the trained network. The implementation details are given in **Figures 1–4**.

The Figure 1 shows the neural network analysis form where we have to set the architecture and parameters for the training of the network. The graph represents the actual values of the series.

The pattern was tested and trained with different architectures to identify the most suitable architecture for the input data. One such sample of the training of the network is given in Figure 2.

Figure 2 displays the training of the network with the architecture 2-2-1. The model summary and the forecasted value are given in Figures 3 and 4, respectively.

Similar experiments have been done for other architectures such as 3-3-1, 12-12-1, etc. and finally the architecture producing the lowest error for the test pattern is considered for analysis. The same methodology has been adopted for all other three data sets and outputs are given below (**Table 2**).

8. Phase-II: Implementation of Hybrid ANN-PSO Models

Traditional optimization methods often use a convergent stepwise technique to conduct a local search, which may result in

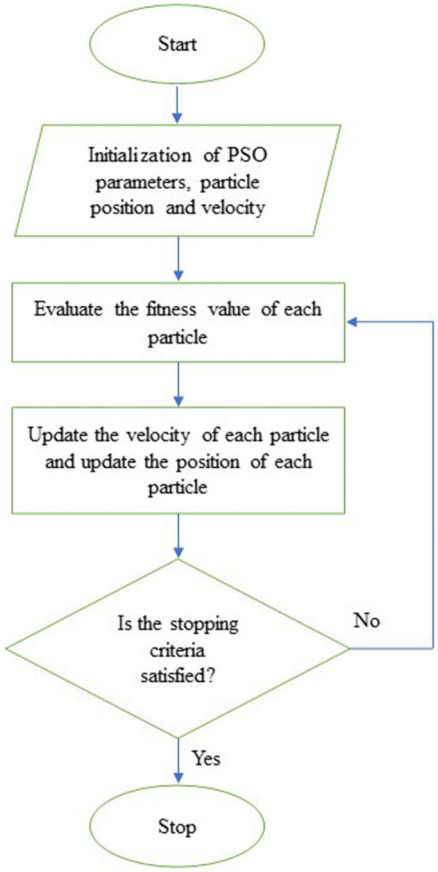

Figure 5. Flowchart for basic PSO algorithm. *Source*: author's diagram.

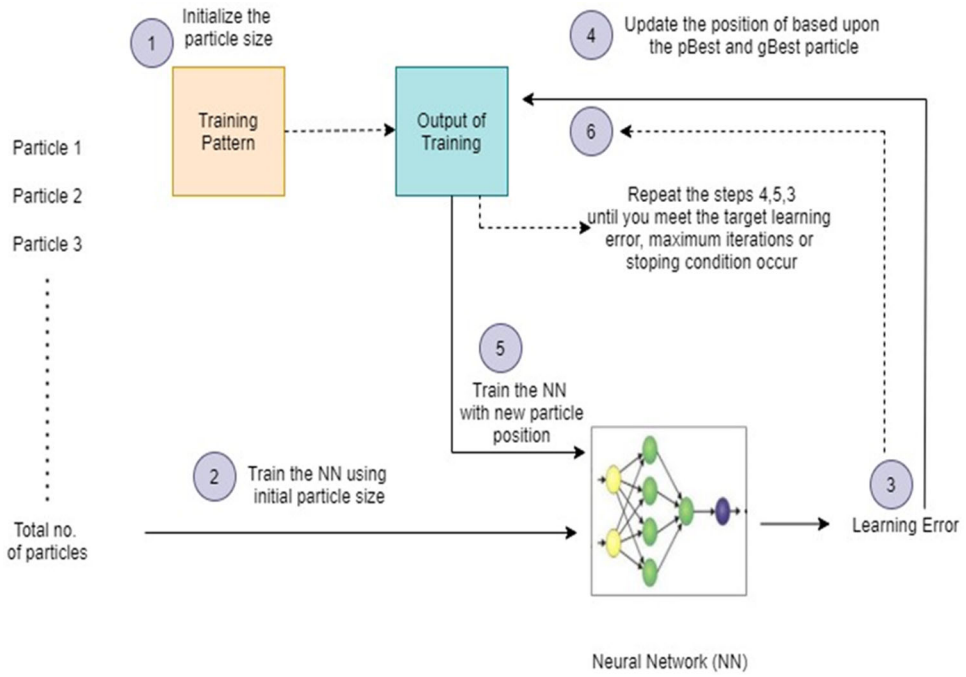

Figure 6. The training of ANN model using PSO algorithm. *Source*: ref.[29]

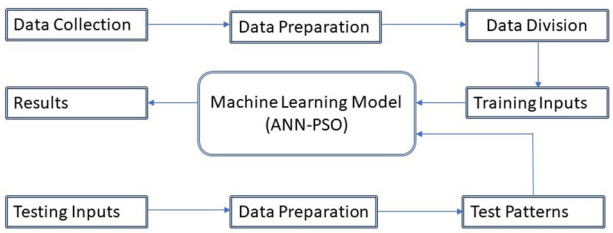

Figure 7. Block diagram: implementation plan for ANN-PSO model.

a local optimum solution. A global optimization procedure is necessary if the complicated function to be improved does not have convexity features that effectively fulfil that the local extreme point is a global optimum. PSO or particle swarm optimization is an advanced computational approach based on the flocking of birds, that is, swarming. PSO may be seen as an extension and improvement of the working concept of genetic algorithm without the use of typical evolution operators such as crossover and mutation. As a result, particle swarm methods are utilized to find the best overall performance. The flowchart for PSO is given in **Figure 5**.

The implementation design of ANN-PSO model is given in **Figure 6**.

The detailed implementation plan is given in **Figure 7**.

The experimental results and accuracy of every model is summarized below (**Table 3**).

9. Comparison of Results and Discussion

This paper orients around a neural network-based machine learning model optimized with PSO. The model has been designed and implemented, and results have been compared with the simple ANN models in predicting the CPI of India. The accuracy of the proposed models has been verified using MAPE. From the MAPE values, it is observed that the proposed ANN models are highly accurate in forecasting the CPI values except the data set of food and beverages where model accuracy is found to be good. It is also found that the ANN implementations require a lot of expertise since no defined rule is there to decide the perfect architecture. It requires a lot of trial experiments to find the best combination of architecture and parameter for a particular data set. When ANN was hybridized with PSO then the forecasting error decreased significantly indicating the high accuracy of the models. From the experimental values it is clearly observed that the ANN and ANN-PSO models are very much effective to be used in forecasting the CPI. However, care has to be taken in selecting the best architecture and parameter combination.

Table 3. Experimental results and model accuracy.

	ANN architecture	Population size	MAPE	Accuracy of the model
Food and beverages	3-2-1	10	4.35	Highly accurate
		20	4.89	Highly accurate
		30	6.34	Highly accurate
	3-3-1	10	6.54	Highly accurate
		20	4.03	Highly accurate
		30	7.13	Highly accurate
Clothing and footwear	3-3-1	10	2.39	Highly accurate
		20	2.99	Highly accurate
		30	3.17	Highly accurate
	6-3-1	10	3.98	Highly accurate
		20	2.81	Highly accurate
		30	2.01	Highly accurate
Housing	3-2-1	10	4.12	Highly accurate
		20	4.99	Highly accurate
		30	3.12	Highly accurate
	3-3-1	10	3.19	Highly accurate
		20	4.01	Highly accurate
		30	3.95	Highly accurate
General index (all groups)	3-3-1	10	5.19	Highly accurate
		20	5.99	Highly accurate
		30	5.78	Highly accurate
	6-3-1	10	3.78	Highly accurate
		20	3.09	Highly accurate
		30	2.12	Highly accurate

Source: Authors' computation.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

artificial neural network, consumer price index, machine learning, mean absolute percentage error, particle swarm optimization

Received: September 7, 2021

[1] W. C. Chiang, D. Enke, T. Wu, R. Wang, *Expert Syst. Appl.* **2016**, *59*, 195.
[2] A. K. Gupta, P. K. Sarangi, *J. Eng. Appl. Sci.* **2012**, *7*, 1017.
[3] N. Singh, R. Singh, *J. Eng. Appl. Sci.* **2009**, *4*, 88.
[4] C. Singla, A. K. Sahoo, *J. Technol. Manag. Growing Econ.* **2019**, *10*, 43.
[5] S. Singh, P. K. Sarangi, *Apeejay J. Manag. Sci. Technol.* **2014**, *2*, 29.
[6] M. Pant, S. Bano, P. K. Sarangi, *Apeejay J. Manag. Sci. Technol.* **2014**, *1*, 28.
[7] P. K. Sarangi, M. Chawla, P. Ghosh, S. Singh, P. K. Singh, *Mater. Today: Proc.* **2021**.
[8] P. K. Sarangi, P. Sarangi, *IUP J. Inf. Technol.* **2010**, *6*, 34.
[9] P. K. Sarangi, P. Sarangi, *IUP J. Comput. Sci.* **2010**, *4*, 15.
[10] R. Bhattacharya, M. Kapoor, Working Papers 20/323, National Institute of Public Finance and Policy, **2020**.
[11] T. Nyoni, MPRA Paper 92414, University Library of Munich, Germany, **2019**.
[12] W. Dongdong, *2010 WASE International Conference on Information Engineering*, **2010**, p. 307, https://doi.org/10.1109/ICIE.2010.79.
[13] V. Shinkarenko, A. Hostryk, LarysaShynkarenko, L. Dolinskyi, *Model. Manag. Emerg. Econ.* **2020**, *107*, 10002.
[14] A. Anand, L. Suganthi, *Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications*, **2020**. https://doi.org/10.4018/978-1-7998-0414-7.ch048.
[15] H. A. L. R. RazanJamous, M. El-Darieby, *Sci. Program.* **2020**, *2021*.
[16] X. Zhou, P. Lu, Z. Zheng, D. Tolliver, A. Keramati, *Reliab. Eng. Syst. Saf.* **2020**, *200*, 106931.
[17] G. Fiorillo, H. Nassif, *Struct. Infrastruct. Eng.: Maint. Manage. Life-Cycle Des. Perform.* **2020**, *16*, 1669.
[18] X. Ning, L. Yac, X. Wang, B. Benatallah, M. Dong, S. Zhang, *Pattern Recognit. Lett.* **2020**, *132*, 12.
[19] M. Abdullah, *J. Asian Bus. Econ. Stud.* **2021**, *303* (in press).
[20] N. I. Indera, I. M. Yassin, A. Zabidi, Z. I. Rizman, *J. Fund. Appl. Sci.* **2018**, *9*, 791.
[21] Y. H. Chen, B. Yang, A. Abraham, *Neurocomputing* **2007**, *70*, 697.
[22] R. Arjun, *Int. J. Res. Sci. Innov. (IJRSI)* **2018**, *5*.
[23] S. R. Das, D. Mishra, M. Rout, *Expert Syst. Appl.: X* **2019**, *4*.
[24] R. A. Jamous, E. E.l. Seidy, A. A. Tharwat, B. I. Bayoum, *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **2015**, *6*.
[25] C. Worasucheep, S. Nuannimnoi, R. Khamvichit, P. Attagonwantana, in *Proc. of the 2017 14th International Conference on Electrical Engineering/Electronics, Computer*, Telecommunications and Information Technology (ECTI-CON), Phuket, Thailand **2017**, pp. 497–500
[26] R. Maojhi, G. Panda, G. Sahoo, A. Panda, A. Choubey, in *Proc. of the 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, Hong Kong, China **2008**, pp. 1276-1282.
[27] The Ministry of Statistics and Programme Implementation, http://www.mospi.nic.in/ (downloaded on: June 26, 2021).
[28] C. Kasemset, N. Sae-Haew, A. Sopadang, *CMU J. Nat. Sci.* **2014**, *13*, 391.
[29] M. Q. Raza, M. Nadarajah, D. H. Hung, Z. Baharudin, *Sustain. Cities Soc.* **2017**, *31*, 264.