Fuzzy Time Series Forecasting Approach using LSTM Model

Radha Mohan Pattanayak[1], M.V.Sangameswar[1], Deepika Vodnala[2], Himansu Das[3]

[1] Godavari Institute of Engineering and Technology,
Department of Computer Science and Engineering,
India

[2] Vignana Bharathi Institute of Technology,
Department of Information Technology,
India

[3] KIIT Deemed to be University,
Department of Computer Engineering,
India

{radhamohan.pattanayak, sangamrjy, deepuvodnala19, das.himansu2007}@gmail.com

Abstract: In the present scenario, fuzzy time series forecasting (FTSF) is an interesting concept by the researchers to approach the uncertainty in the dataset. In the current study, we proposed a fuzzy long short term memory (FLSTM) model to forecast a wide range of time series (TS) dataset with less computational complexity. The present research mainly focuses on two issues such as (1) in order to obtain the number of intervals (NOIs) of the universe of discourse (UOD) the trend based discretization (TBD) approach is applied, and (2) the subscript of the fuzzy set associated with the crisp observation is considered to establish the fuzzy logical relationships (FLRs) for the proposed FLSTM model. To demonstrate the forecasting ability of the FLSTM model, six TS datasets with three profound FTSF models are considered in this paper. The empirical result analysis revealed that, in all measured the proposed model outperformed and showed better result than its alternatives. The outcome of the different FTSF models on different measures proves the outperformance of the FLSTM model than its competitors.

Keywords. Long short term memory (LSTM), fuzzy time series forecasting (FTSF), fuzzy logical relationships (FLRs), length of interval (LOI), number of Interval (NOI), time series (TS), fuzzy set theory (FST).

1 Introduction

Since from the last two decades, FTS method has been considered as a popular topic in research because of handling the ambiguity and incomplete information more efficiently in decision making (DM) system.

At first, Zadeh [1] implemented fuzzy set theory to challenge the linguistic terms in the fuzzy set.

Later, by adopting the fuzzy set concept, Song and Chissom [2, 3] first developed FTSF model and solved most of the forecasting issues in the DM problem.

Later, Chen [4] modified the FTS model by using simple arithmetic operation instead of complex matrix operations. During the 90's, many researchers used simplex structure for forecasting and due to the simplex structure, the first order FTS models have been facing more difficulties to forecast complex data.

By considering the above difficulties, Chen et al. [5] proposed high order FTS to forecast the student enrolment records. After that, many researchers [5–9] have continued their research in this domain to increase the accuracy of fuzzy forecasting model.

Recent papers [10–24] provide detailed literature on different steps of FTSF.

However, in the proposed FLSTM model we mainly concentrated on two major concepts such as: (1) to determine the NOIs of the UOD and to make partition the UOD into different intervals of equal proportion, the TBD approach [10] is applied, and (2) the subscript of the fuzzy set associated with the crisp observation is considered to establish the FLRs and later it is modeled using LSTM.

In the last few years, the interval determination methods faced difficulties to determine the LOI, and which is not clear to date.

In 2001, Huarng [7] proposed average and distribution based procedure to partition the UOD into equal proportion of length.

Later Panigrahi and Behera [17] proposed modified average based method, to obtain the intervals of the UOD with equal proportion.

Pattanayak et al. [20] proposed an entropy based neutrosophic model where an adaptive based partition method is considered to partition the time series into unequal length of intervals. The adaptive partition method improves the forecasting efficiency of the model by removing the outliers in the time series more efficiently.

Huarng and Yu [25] obtained an unequal proportion of length of the UOD by employing ratio based method.

Many researchers [26–29] also considered evolutionary algorithms (EAs) to optimize the LOI. But defining the initial parameter is an important factor in the efficiency of EA methods.

Motivated by this, in the present study we considered the TBD approach [10] to attain the NOIs of TS data and later it splits the UOD into equal proportion of length.

Coming to the concept of FLR, many researchers have followed different methods to model the FLRs for different FTSF model. In most of the research, the authors have been using either the highest membership or all membership values of the element to establish the FLRs. Pattanayak et al. [10] proposed a probabilistic intuitionistic model and considered a combination of mean of membership values and the data element to establish the FLRs. Researchers [8, 16, 30] considered all membership values of each interval and used it to model the FLR by using different neural network models.

Pattanayak et al. [11] considered both data and all membership values to establish the FLRs. Although by considering all membership values saves the loss of information in forecasting but it increases the size of the input pattern into high, which reduces the performance of the FTSF model. Observing from this, we have considered the subscript value of the fuzzy set related to the crisp observation to establish the FLRs of the TS data.

The rest of this research are formed as follows. Section 2, explains the preliminaries including different definitions of FTS, the working principle of LSTM. Section 3, demonstrates the execution process of proposed FLSTM model. Section 4, explains the experimental analysis. Section 5, explains the feature of the FLSTM model and highlights the future extension of the work.

2 Background Study

2.1 Definition of FTS

A conventional TS, $y(t)$ is called as a FTS if the values presented in TS are fuzzy values. Suppose $f(t)$ is the collection of different fuzzy values for different intervals of the TS, thereupon $S(t)$ is called a FTS prescribed on $y(t)$.

Definition 1 [11]:

Suppose $\{u_1, u_2 \cdots u_n\}$, are the elements of a universal set $\mathcal{U}$ for which a fuzzy set S can be defined as shown in Eq. (1):

$$f_i = \frac{f_{Si}(u_1)}{u_1} + \frac{f_{Si}(u_2)}{u_2} + \cdots\cdots + \frac{f_{Si}(u_n)}{u_n}, \quad (1)$$

where f_{Si} represents as the membership function of the fuzzy set f.

Definition 2 [11]:

Let in FTS, the instant $f(t-1)$ is used to define the value of its next instant $f(t)$, thereupon the above relation the FLR of the TS data is defined as $f(t-1) \rightarrow f(t)$ and it is termed as first order FTS.

Definition 3 [11]:

Let in FTS, the instants $f(t-1), f(t-2), \cdots f(t-n)$ are used to define the value of the instant $f(t)$, thereupon the above relation the FLR of the TS data is defined as $f(t-n), \cdots f(t-2), f(t-1) \rightarrow f(t)$ it is termed as high order FTS.

2.2 Concept of LSTM

In 1977, Hochreiter and Schmidhuber [31] developed a deep learning technique as long short term memory (LSTM) is another version of the recurrent neural network model. Unlike the traditional neural network model where the input data flows in a forward direction only, in LSTM due to a feedback connection in the architecture the data can be processed both in backward as well as the forward direction. The architecture of LSTM comprises three distinct layers namely the input layer, the hidden layer, and the output layer. The neurons in the hidden layer are connected self recurrently memory cells and expect from this the hidden layer consists of a special memory block to regulate the information transmission.

The memory unit is the collection of three gates such as the input gate, forget gate, and output gate. The different gates in the memory block are used to control the flow of information considering either by adding information or by removing information in the memory cell. The input gate controls the flow of activated values into the cell. The output gate controls the flow of computed output values of a cell, which transmits the output to other neurons.

Similarly, the forget state is responsible to discard the information from the memory cell which are less importance and no longer used in the network. In the present study the LSTM technique is considered to model the FLRs for the FLSTM model.

3 Methodology

Algorithm 1 explains the execution process of the proposed FLSTM model. The first few steps of the algorithm explains the different procedures used to obtain the RTV data of the TS data. The NOIs of the TS data are determined using TBD approach [10].

The subscript of the fuzzy set associated with the crisp observation is considered to establish the FLRs for the FLSTM model and later it is modeled by LSTM model. After that, the forecasted data are de-normalized and then the de-normalized values are defuzzified to get the actual forecasted value of the TS data.

Algorithm 1. The proposed FLSTM model

Input :	The TS data as $y = [y_1, y_2, \cdots\cdots, y_n]^{\mathrm{T}}$, length of training set (ℓ_{tr}), length of validation set (ℓ_v), and length of test set (ℓ_t).
Output:	Forecasted value.

1: Compute ℓ_0 and ℓ_1 for the TS data

$$\ell_1 = \frac{\sum_{i=1}^{n}\left(t_i - \frac{\sum_{i=1}^{n} t_i}{n}\right) * \left(y_i - \frac{\sum_{i=1}^{n} y_i}{n}\right)}{\sum_{i=1}^{n}\left(t_i - \frac{\sum_{i=1}^{n} t_i}{n}\right)^2}$$

$$\ell_0 = \frac{1}{n}\left(\sum_{i=1}^{n} y_i - \ell_1 \sum_{i=1}^{n} t_i\right)$$

2: Compute the trend value of each observation in the TS data.

$for\ t = 1\ to\ n (size\ of\ TS)$
$\quad u_t = \ell_0 + \ell_1 t$
$end\ of\ for$

3: Calculate the RTV values i.e. V of each observation in the TS data

$for\ t = 1\ to\ n\ (Size\ of\ TS)$
$\quad V_t = \left(y_t / u_t\right) * 100$
$end\ of\ for$

4: By taking into account to the median value of the RTV data, it is partitioned into two subgroups P and Q.

5: Compute the absolute first difference P_d and Q_d in both subgroups P and Q.

6: Obtain the average of the P_d and Q_d as P_v and Q_v for both the subsets P and Q respectively.

7: Obtain the deciding factor ϑ

$$\vartheta = 10^{floor(e)\left(\frac{P_v}{Q_v}\right)}$$

8: Obtain the NOIs i.e. k as follows

$$k = \left(\frac{(min(V) - \vartheta) - (max(V) + \vartheta)}{\vartheta}\right) + \ell_t \times n$$

9: Determine the UOD as: $\mathcal{U} = (min(y) - \delta, \; max(y) + \delta)$. Where, δ is the standard deviation.

10: Obtain the NOIs k of the TS.

11: Partition the UOD
$\mathcal{U} = \{ue_1, ue_2, ue_3, \cdots, ue_w,\}$ into k equal length of intervals and Obtain the mid-point m_i $(i = 1,2,\cdots, w)$ of each intervals.

//Fuzzify the TS data

12: $for \; r = 1 \; to \; n \; (length \; of \; TS)$
$\quad if \; y_r \; belongs \; to \; the \; partition \; ue_i$
$\quad\quad f_r(t) = i$
$\quad end \; of \; if$
$end \; of \; for$

//Establish the FLR by using LSTM

13: There after the order of the FLSTM model is obtained employing autocorrelation function.

14: Apply Min-Max normalization to obtain normalized in sample TS data $f'(t)$.

15: Transform the $f'(t)$, into $n - r$ patterns using sliding window protocol.

16: Divide the TS data into train, validation, and test set.

17: Obtain normalized forecasting result $\widehat{f}'(t)$ using the optimized parameters of LSTM.

18: De-normalize the forecasting result i.e. $\hat{f}(t)$

19: Obtain the forecasted value by defuzzify $\hat{f}(t)$ by using middle point fuzzy interval and obtain the actual forecasted value $\hat{y}$.

4 Experimental Setup and Results

The present study considered the Taiwan capitalization weighted stock index (TAIEX) TS data from the year 2005 to 2010 (as shown in Table 1), and three profound forecasting models as Aladag et al. [21], Aladag [23], and Bas et al.

[14] to investigate the forecasting accuracy of the proposed FLSTM model. All the models are implemented using MATLAB R2018b.

In all compared models the UOD of the TS data is calculated as $[min - d1, \; max + d2]$, where the UOD of the TS data is computed as $[min(y) - d1, \; max(y) + d2]$, where the function $max(y)$ and $min(y)$ shows the higher and lower value of the TS data and ten percent of $min(y)$ and $max(y)$ value are considered as $d1$ and $d2$ respectively.

In the proposed model the NOI of the TS data is obtained using TBD approach [10] and the same NOI value is considered in all compared models for partitioning the UOD into equal length of intervals. ACF function is employed to find the order of the TS data.

The subscript of the fuzzy set associated with the crisp observation is considered to establish the FLRs for the FLSTM model and later it is modeled by using LSTM. Each pattern in the experiment breaks into in-sample and out of sample pattern and the values are represented in Table 1. To find the efficiency of the FLSTM model, two error measures such as root mean square error (RMSE) and symmetric mean absolute percentage error (SMAPE) (as shown in Eq. (2-3)) are considered:

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2}, \qquad (2)$$

$$SMAPE = \frac{1}{n}\sum_{j=1}^{n}\frac{|y_j - \hat{y}_j|}{(|y_j| + |\hat{y}_j|)/2}, \qquad (3)$$

where $\hat{y}_j$, and y_j are predicted and actual values respectively.

In the experimental analysis, all four models are executed fifty number of times by conceding RMSE and SMAPE measure.

Table 2 presents the mean value of fifty executions resulted from different FTSF models considering RMSE measure.

Table 3 presents the mean value of fifty executions resulted from different FTSF models considering SMAPE measure. The analysis result from both Table 2 and Table 3 noticed that, the
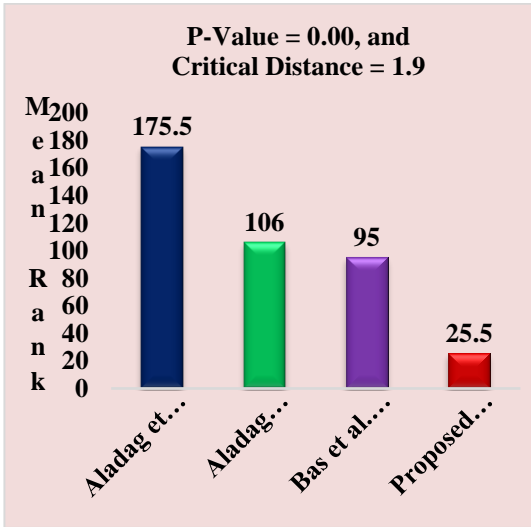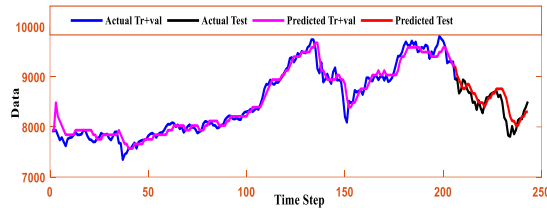
Fig. 1. The mean rank of different FTSF model

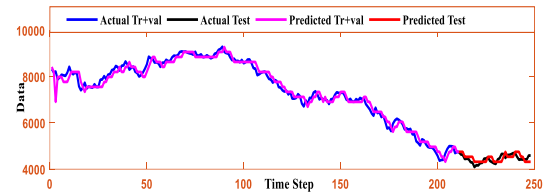**Fig. 2.** Actual and Forecasted value of TAIEX2007 TS data
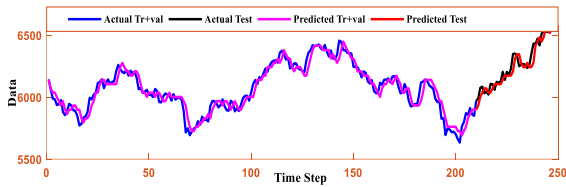

Fig. 3. Actual and Forecasted value of TAIEX2008 TS data

Fig. 4. Actual and Forecasted value of TAIEX2005 TS data
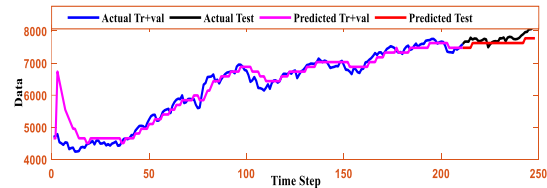

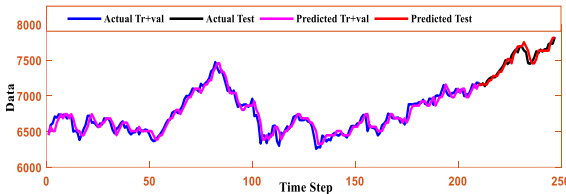**Fig. 5.** Actual and Forecasted value of TAIEX2009 TS data

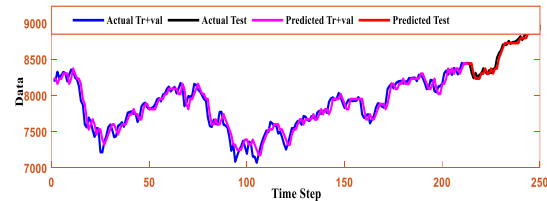**Fig. 6.** Actual and Forecasted value of TAIEX2006 TS data

Fig. 7. Actual and Forecasted value of TAIEX20010 TS data

proposed FLSTM model outperforms in all six TS datasets using RMSE as well as SMAPE measure.

The above analysis results concludes the outperformance of the FLSTM model than its competitor FTSF models. Later, to make a robust comparison the SMAPE result of all datasets together are collected and employed a Nemenyi [32] hypothesis test is performed with confidence level of 95%.

The obtained result from the hypothesis [32] are shown in Figure 1. From Fig 1 it clearly shows that, the proposed FLSTM model has the smaller mean rank 25.5 among all FTSF models.

The above analysis proves the proposed FLSTM model is more robust in nature than its competitor models. In order to show the intimacy among the TS data and the forecasted value resulted on the TAIEX TS data for the year 2005 to 2010 are depicted as shown in Figure (2-7).

Table 1. Description of different TS data

Year wise TAIEX TS Data	Total pattern	In-Sample data (85% of the TS data)		Out of sample data (15% of the TS data)	NOIs of the TS data
		Train (70% of the TS data)	Validation (15% of the TS data)	Test (15% of the TS data)	
2005	247	173	37	37	39
2006	247	173	37	37	39
2007	243	170	36	37	41
2008	248	174	37	37	39
2009	248	174	37	37	42
2010	250	175	38	37	39

Table 2. Mean RMSE result of different FTSF model (Best value is in Bold)

Year wise TAIEX TS data	Aladag et al. [21]	Aladag [23]	Bas et al. [14]	Proposed FLSTM
	Mean	Mean	Mean	Mean
2005	6243.39	154.84	110.02	**58.69**
2006	7475.06	426.60	429.46	**65.93**
2007	8428.90	375.23	393.05	**203.67**
2008	4440.41	2517	2517	**169.26**
2009	7751.75	995.09	566.11	**139.36**
2010	8584.86	509.92	462.33	**55.15**

Table 3. Mean SMAPE result of different FTSF model (Best value is in Bold)

Year wise TAIEX TS Data	Aladag et al. [21]	Aladag [23]	Bas et al. [14]	Proposed FLSTM
	Mean	Mean	Mean	Mean
2005	199.62	1.96	1.45	**0.74**
2006	198.73	5.53	5.56	**0.67**
2007	199.10	3.55	3.81	**1.88**
2008	199.10	44.11	44.11	**3.26**
2009	198.36	13.57	7.39	**1.52**
2010	198.87	5.73	5.19	**0.50**

5   Conclusion

The present research introduced a FTSF model using LSTM to forecast a wide range of TS datasets. To confirm the forecasting efficiency of the proposed FLSTM model, three profound FTSF models and the TAIEX TS data for six consecutive years from 2005 to 2010 are considered. The comparative result based on

Table 2 and Table 3, proves the outperformance of the FLSTM model on each dataset than the competitors by employing both the RMSE and SMAPE measure. Later, by considering all datasets together, a Nemenyi [32] hypothesis is conducted to test the statistical superiority of the FLSTM model. The outcome of the Nemenyi [32] hypothesis test from Fig 1 represents the statistical supremacy of the FLSTM model with lowest mean rank 25.5 than the competitors.

In the future, to increase the accuracy of the FTSF model one can use (1) any optimization techniques to obtain the membership value, (2) instead of using conventional FST, one can use either hesitant FST or neutrosophic FST can be employed to establish the fuzzification, and (3) instead of using ACF function one can use any significant input variable selection procedure to define the order of the TS data.

References

1. **Zadeh, L.A. (1965).** Fuzzy logic. Information and Control, Vol. 8, pp. 338–353.

2. **Song, Q., Chissom, B.S. (1993).** Forecasting enrollments with fuzzy time series - Part I. Fuzzy Sets and Systems, Vol. 54, pp. 1–9.

3. **Song, Q., Chissorn, B.S. (1994).** Forecasting enrollments with fuzzy time series-part II. Fuzzy Sets and Systems, Vol. 62, pp. 1–8.

4. **Chen, S.M. (1996).** Forecasting enrollments based on fuzzy time series. Fuzzy Sets and Systems, Vol. 81, pp. 311–319.

5. **Chen, S.M., Chung, N.Y. (2006).** Forecasting enrollments using high-order fuzzy time series and genetic algorithms. International Journal of Intelligent Systems, Vol. 21, pp. 485–501.

6. **Huarng, K., Yu, T.H.K. (2006).** The application of neural networks to forecast fuzzy time series. Physica A: Statistical Mechanics and its Applications, Vol. 363, pp. 481–491.

7. **Young, M. (1989).** The Technical Writer's Handbook. Mill Valley, CA: University Science.

8. **Yu, T.H.K., Huarng, K.H. (2010).** A neural network-based fuzzy time series model to improve forecasting. Expert Systems with Applications, Vol. 37, pp. 3366–3372.

9. **Teoh, H.J., Cheng, C.H., Chu, H.H., Chen, J.S. (2008).** Fuzzy time series model based on probabilistic approach and rough set rule induction for empirical research in stock markets. Data and Knowledge Engineering, Vol. 67, pp. 103–117.

10. **Pattanayak, R.M., Behera, H.S., Panigrahi, S. (2021).** A novel probabilistic intuitionistic fuzzy set based model for high order fuzzy time series forecasting. Engineering Applications of Artificial Intelligence, Vol. 99, pp. 104136.

11. **Pattanayak, R.M., Panigrahi, S., Behera, H.S. (2020).** High-Order Fuzzy Time Series Forecasting by Using Membership Values Along with Data and Support Vector Machine. Arabian Journal for Science and Engineering. Vol. 45, pp. 10311–10325.

12. **Pattanayak, R.M., Behera, H.S. (2018).** Higher Order Neural Network and Its Applications: A Comprehensive Survey. Advances in Intelligent Systems and Computing, Vol. 710, pp. 695–709.

13. **Pattanayak, R.M., Behera, H.S., Panigrahi, S. (2020).** A Novel Hybrid Differential Evolution-PSNN for Fuzzy Time Series Forecasting. Computational Intelligence in Data Mining, pp. 675–687.

14. **Bas, E., Grosan, C., Egrioglu, E., Yolcu, U. (2018).** High order fuzzy time series method based on pi-sigma neural network. Engineering Applications of Artificial Intelligence, 72, 350–356.

15. **Egrioglu, E., Aladag, C.H., Yolcu, U., Uslu, V.R., Basaran, M.A. (2010).** Finding an optimal interval length in high order fuzzy time series. Expert Systems with Applications, Vol. 37, pp. 5052–5055.

16. **Cagcag-Yolcu, O., Lam, H.K. (2017).** A combined robust fuzzy time series method for prediction of time series. Neurocomputing, Vol. 247, pp. 87–101.

17. **Panigrahi, S., Behera, H.S. (2020).** A study on leading machine learning techniques for high order fuzzy time series forecasting. Engineering Applications of Artificial Intelligence, Vol. 87, pp. 103245.

18. **Pattanayak, R.M., Behera, H.S., Rath, R.K. (2020).** A Higher Order Neuro-Fuzzy Time Series Forecasting Model Based on Un-equal Length of Interval. International Conference on Application of Robotics in Industry using Advanced Mechanisms, pp. 34–45.

19. **Pattanayak, R.M., Behera, H.S., Panigrahi, S. (2020).** A Multi-step-Ahead Fuzzy Time Series Forecasting by Using Hybrid Chemical Reaction Optimization with Pi-Sigma Higher-Order Neural Network. Computational Intelligence in Pattern Recognition, pp. 1029–1041.

20. **Pattanayak, R.M., Behera, H.S., Panigrahi, S. (2021).** A non-probabilistic neutrosophic entropy-based method for high-order fuzzy time-series

forecasting. Arabian Journal for Science and Engineering.

21. **Aladag, C.H., Basaran, M.A., Egrioglu, E., Yolcu, U., Uslu, V.R. (2009).** Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations. Expert Systems with Applications, Vol. 36, pp. 4228–4231.

22. **Das, H., Naik, B., Behera, H.S. (2020).** A hybrid neuro-fuzzy and feature reduction model for classification. Advances in Fuzzy Systems.

23. **Aladag, C.H. (2013).** Using multiplicative neuron model to establish fuzzy logic relationships. Expert Systems with Applications, Vol. 40, pp. 850–853.

24. **Das, H., Naik, B., Behera, H.S., Jaiswal, S., Mahato, P., Rout, M. (2020).** Biomedical data analysis using neuro-fuzzy model with post-feature reduction. Journal of King Saud University-Computer and Information Sciences, pp. 1–11.

25. **Huarng, K., Yu, T.H.K. (2006).** Ratio-based lengths of intervals to improve fuzzy time series forecasting. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 36, pp. 328–340.

26. **Kuo, I.H., Horng, S.J., Kao, T.W., Lin, T.L., Lee, C.L., Pan, Y. (2009).** An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. Expert Systems with Applications, Vol. 36, pp. 6108–6117.

27. **Kuo, I.H., Horng, S.J., Chen, Y.H., Run, R.S., Kao, T.W., Chen, R.J., Lai, J.L., Lin, T.L. (2020).** Forecasting TAIFEX based on fuzzy time series and particle swarm optimization. Expert Systems with Applications, Vol. 37, pp. 1494–1502.

28. **Eğrioğlu, E. (2012).** A new time-invariant fuzzy time series forecasting method based on genetic algorithm. Advances in Fuzzy Systems.

29. **Cai, Q., Zhang, D., Zheng, W., Leung, S.C.H. (2015).** A new fuzzy time series forecasting model combined with ant colony optimization and auto-regression. Knowledge-Based Systems, Vol. 74, pp. 61–68.

30. **Yolcu, U., Aladag, C.H., Egrioglu, E., Uslu, V.R. (2013).** Time-series forecasting with a novel fuzzy time-series approach: an example for Istanbul stock market. Journal of Statistical Computation and Simulation, Vol. 83, pp. 599–612.

31. **Hochreiter, S., Schmidhuber, J. (1997).** Long short-term memory. Neural Computation, Vol. 9, pp. 1735–1780.

32. **Demšar, J. (2006).** Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research, Vol. 7, pp. 1–30.