International Journal of Decision Inelligence
Vol 1, Issue 2, Spring 2023 , 45-53

A Hybrid Type-2 Fuzzy-LSTM Model (HT2FLSTM) for Prediction of Environmental Temporal Patterns

Aref Safari[a,*], Rahil Hosseini[a]

Department of Computer Engineering, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran

Received 04 March 2023; Accepted 14 April 2023

Abstract:

Computational intelligence methods, such as fuzzy logic and deep neural networks, are robust models to solve real-world problems. In many dynamic and complex problems, statistical attributes frequently change over the time. Recurrent neural networks (RNN) are suitable to model dynamic high-dimensional and non-linear state-space systems. Nevertheless, the RNN is incapable of modelling long-term dependencies in temporal data, and its learning using gradient descent is a complex and difficult task. The main objective of this research is extant a Hybrid Type-2 Fuzzy-LSTM Model (HT2FLSTM) for cope with uncertain and complex temporal time-series such as air pollution prediction. In this work, the proposed model aimed at overcoming the drawbacks of the existing methods and presenting a reliable, efficient and robust model for prediction of prediction of environmental temporal patterns through type-2 fuzzy systems and LSTM models. On the other hand, precise and reliable prediction can help reasonable strategies and assist the specialists in planning the best policies to model an event in uncertain circumstances.

***Keywords**:* Deep Learning; Type-2 Fuzzy Logic; LSTM Network; Time-Series Prediction.

1-Introduction

The mathematical issue of learning long-term dependencies in RNNs is that gradients propagated over many stages have a tendency to vanish or explode [1]. One of available solutions to deal with this issue is to utilize a model that operates at several time scales. Deep learning approaches are major extensions of machine learning. The deep learning methods enables the modeling of complex relationships and concepts of the learning procedure using multiple levels of layers and their representation. On the other side, the Long-Short-Term-Memory (LSTM) networks are based on the idea of generates the paths through time that have derivatives that neither vanish nor explode. The distribution of air pollutants, such as particle matters, involves a complex process depends on several factors during a time-series. Handling such problems requires intelligent models to address the high-order uncertainty in the air pollution issue's characteristics. This research aims at modeling uncertainty sources associated with non-stationary time series features in real-world applications such as air quality prediction.

The idea of introduce a self-loop to produce paths where the gradient can flow for extended times is an important influence of the original LSTM model [1]. By makes the weight of this self-loop gated, the time scale of integration can be changed dynamically. Nevertheless, one issue needs to be addressed in learning long-term problems; the uncertainty in temporal data and non-stationary problems [2]. The lack of inference and interpretability is that gap we aimed to handle with powerful tools in cope with high-order uncertainty; type-2 fuzzy logic systems. Besides, the fuzzy logic systems suffer from the lack of learning and adaptation mechanisms for their membership function parameters and inference rules. The fuzzy inference rules make feature extraction more human interpretable under uncertainty. Utilize the hybrid intelligent models brings the advantages of both fuzzy logic and deep learning approaches in model the uncertainty and extract useful high-level features from ambiguous data. Therefore, this study's proposed model is a win-win game for LSTM and type-2 fuzzy logic systems.

The HT2FLSTM model benefits from computational intelligence models' potential for solve the long-term learning dependencies problems in an uncertain time-series prediction problem.

[*] Corresponding Author. Email: safari.aref@gmail.com

1-1- Literature Review

An RNN has cycles in its graph that let it to keep information about historical inputs for an amount of time. The recurrent network transforms an input sequence into an output sequence while flexibly takes into account contextual information. Another architecture that makes it easy to remember information over long periods is called Long-Term Short-Term Memory (LSTM). The idea is that a network's activations are its short-term memory and its weights are its long-term memory. The LSTM architecture wants the short-term memory to last for a long time. It's composed of memory cells that have controllers say when to store or forget information.

Many types of research [3]-[4] have been reported for model uncertainty based on fuzzy logic. During the recent decade, applications of type-2 fuzzy logic systems have been grown in handles the high-order uncertainty, especially for dynamic and non-stationary problems in control and time-series prediction [5]-[6]. Various intelligent models, such as neuro-fuzzy and LSTM, have been applied to model and predict the time series data in uncertain environments [7]-[11]. In addition to the general benefits of RNNs for time series prediction, the LSTM network can also automatically learn the data's temporal dependencies [12]. An LSTM cell's benefit compared to a regular recurrent unit is its cell memory [13]-[14]. The cell vector of an LSTM has the capability of the information earlier stored memory and the part of its new information [15]. These capabilities can be used in the prediction of temporal problems. LSTM networks have an excellent capability for handle the complex problems. Many LSTM networks are introduced as deep learning approaches to resolve many issues, such as model the multi-layered inter-relationships between two-time series [16]-[17].

According to Table1, various hybrid fuzzy models applied to predict the time series data [18]-[24] have been reviewed. It has been proven that fuzzy systems, especially hybrid fuzzy models, have a great capability for solving complex and uncertain temporal problems, where the model estimates and predicts the similarity between two time series in uncertain conditions. [25]-[32].

Table 1:
Related fuzzy models for time-series predictions

Method	Limitations
FCM Time Series [18]	No optimum parameter
Fuzzy logic [19]	No data reduction
Fuzzy Markov [20]	Complexity of model
Fuzzy time series [23]	High-order uncertainty is not modeled
Fuzzy Neural [25]	Weights are fixed
Fuzzy-PSO [26]	Limited to mid-range series
Gustafson-Kessel Fuzzy Clustering [27]	Insufficient results for long time-series.
PSO-fuzzy time series [28]	Not reliable for long time-series
Fuzzy-NN time series clustering [29]	Time complexity is high
Evolutionary Type-2 Fuzzy Rule Learning [30]	Sensetive to initiate patameters
Type-2 Fuzzy DTW Time Series [31]	Not Reported
DT2FTW [32]	Not Reported

2-Theoretical Research Background

This section presents a brief overview of the RNN and the LSTM networks in A. It follows a review of interval type-2 fuzzy sets (IT2FS) concepts and their mathematic definitions in B.

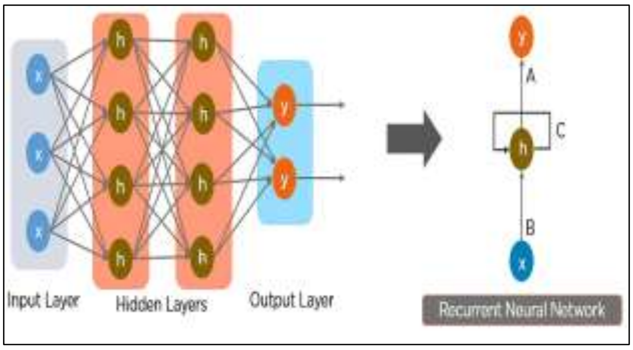

Fig.1. The architecture of a Simple Recurrent Neural Network

2-1-Recurrent Neural Networks and LSTM

RNNs can represent time series, which are presented utilizes data sequences. One of the most important RNN deep learning models is LSTM. The LSTM network benefits several gated units to avoid RNN traditional issues; gradient vanishing. LSTM has been widely applied in many areas, especially in time series prediction. The LSTMs have been developed to address classic RNNs' limitations by enhance the network structure's gradient vanishing. The use of a cell state attains this c_t, which stores long-term information as follows [13]:

Input gate:

$$i_t = \sigma(W_{i1}z_{t-1} + W_{i2}h_t + W_{i3}x_t + W_{i4}s + b_i) \quad (1)$$

Output gate:

$$o_t = \sigma(W_{o1}z_{t-1} + W_{o2}h_t + W_{o3}x_t + W_{o4}s + b_o) \quad (2)$$

Forget Gate:

$$f_t = \sigma(W_{f1}z_{t-1} + W_{f2}h_t + W_{f3}x_t + W_{f4}s + b_f) \quad (3)$$

where z_{t-1} is the hidden state of the LSTM at time $t-1$ and W is the weight matrices. As well as, t index is the time step, x is the input, and h is the output variables at time t, and σ is the sigmoid activation function. The gates adjust the states and hidden cell of the LSTM based on the following equations [13]:

Hidden state:

$$z_t = o_t \odot \mathrm{Tanh}(c_t) \quad (4)$$

Cell state:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{c1}z_{t-1} + \cdots + W_{c4}s + b_c) \quad (5)$$

where $\odot$ is the element-wise product, and tanh as the activation function.

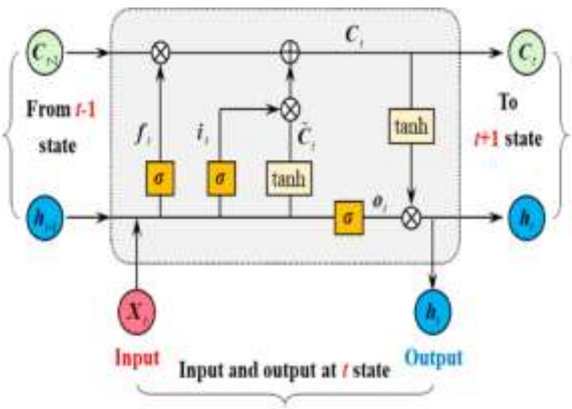

Fig.2. The architecture of the LSTM cell [14]

2.2 An Overview of Interval Type-2 Fuzzy Sets

An IT2FS $\tilde{A}$, is characterized through a type-2 MF $\mu_{\tilde{A}(x,u)}$ where $x \in X$ and $u \in J_x \subseteq [0,1]$ as follows [20]:

$$\tilde{A} = \left\{\left((x,u), \mu_{\tilde{A}(x,u)}\right) | \forall x \in X, \forall \in J_x \subseteq [0,1]\right\} \quad (6)$$

where, $0 \leq \mu_{\tilde{A}(x,u)} \leq 1$, X is the domain of fuzzy set and J_x is the domain of the secondary MF at x. $\tilde{A}$ is as [20]:

$$\tilde{A} = \frac{\int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x,u)}{x, uJ_x} \subseteq [0,1] \quad (7)$$

where $\iint$ represents union over all admissible x and u and,

$$\tilde{A} = \frac{\int_{x \in X} \int_{u \in J_x} 1}{x, u} = \frac{\int_{x \in X} \left[\frac{\int_{u \in J_x} 1}{u}\right]}{x} \quad (8)$$

where x is the primary variable, J_x , an interval in [0,1], is the primary MF of x, u is the secondary variable, and $\int_{u \in J_x}$ is the secondary MF at x. Uncertainty about $\tilde{A}$ is addressed by the union of all of the primary memberships, called the footprint of uncertainty (FOU) of $\tilde{A}$, i.e., $[FOU(\tilde{A})]$ as follows: [21]

$$FOU(\tilde{A}) = \bigcup_{x \in X} J_x \quad (9)$$

The membership grade of each element of an IT2FS is identified by an interval of $[\bar{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{A}}(x)]$. In the IT2FLS, the UMF and LMF can better represent input variables' uncertainties than type-1 fuzzy sets. Similarly, the FOU in the IT2FLS provides more degrees of freedom when desings a fuzzy system [22]. Membership function (MF) of a type-2 fuzzy set (T2FS) of a given element is itself a type- 1 fuzzy set (T1FS) [20]. An interval type-2 fuzzy system (IT2FS) architecture contains four components; fuzzifier, fuzzy rule base, inference engine, type-reducer, and defuzzifier [23].

3-The Proposed Model

This section presents the details of the proposed deep HT2FLSTM architecture. In section A, we present the general structure of the HT2FLSTM model. Then, in section B, the mathematical model of the HT2FLSTM is described, in sections C and D include detailed description of inference layers and the mathematical details of fuzzy layers. In Section E, the cell structure of the LSTM (HT2FLSTM Cell) has been proposed.

3-1-The Architecture of the proposed HT2FLSTM

The HT2FLSTM is fuzzified by the IT2F sets. The framework of the proposed HT2FLSTM includes five layers; the input layer, encoder layer, hidden layers, the decoder layer, and the output layer. The details of each layer are illustrated in Figure. 3.

Input Layer: The first layer is the input layer of HT2FLSTM model in order to handles the historical time-series data. The output of this layer feeds the inputs of the deep layers.

Encoder, Hidden, and Decoder layers: The main idea is to map the entire input sequence to a vector and then use an encoder to generate the output sequence. In this layer, the encoder represents the whole input sequence in the hidden layer activities. The proposed structure can reform affords to the complexity of the training time-series.

Output Layer: The final layer is considered as the prediction layer to make the decision and perform the prediction based on the input historical time-series data were established from the previously hidden layers.

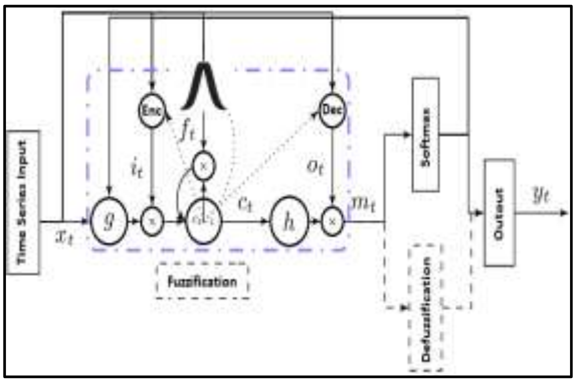

Fig.3. The architecture of the proposed model

3-2- The Mathematical Model of the HT2FLSTM

Time series prediction models specify future values of a target y_{it} for a given entity i at time t. The next step-ahead prediction is as follows:

$$\hat{y}_{i,t+1} = P_f(y_{i,t-e:t}, x_{i,t-e:t}, t) \quad (10)$$

where $\hat{y}_{i,t+1}$ is the next step predicted value, t is the time step at T, $y_{i,t-e:t}$ and $x_{i,t-e:t}$ are observations of the target and observed inputs, respectively, over a look-back window e, and P_f is the prediction function, and the final prediction is produced by Z_t:

$$P_f(y_{i,t-e:t}, x_{i,e-k:t}, t,) = g_{dec}(z_t) \quad (11)$$

$$Z_t = g_{enc}(y_{i,t-k:e}, x_{i,t-e:t}, t) \quad (12)$$

where g_{enc} and g_{dec} denote the encoder and decoder functions, respectively. The mathematical model of different parts of the LSTM network in the proposed HT2FLSTM architecture are given as follows:

Definition 1: Let N be the number of memory units of the model. In time-step t, i.e., the current time, the network keeps in memory a set of vectors by the following equations:

$$i_t = \sigma(W_{ix}x_t + W_{im}m_{t-1} + W_{ic}c_{t-1} + b_i) \quad (13)$$

$$f_t = \sigma(W_{fx}x_t + W_{fm}m_{t-1} + W_{fc}c_{t-1} + b_f) \quad (14)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g(W_{cx}x_t + W_{cm}m_{t-1}) \quad (15)$$

$$o_t = \sigma(W_{ox}x_t + W_{om}m_{t-1} + W_{oc}c_t + b_o) \quad (16)$$

$$m_t = o_t \odot h(c_t) \quad (17)$$

$$f_\varphi = (W_{hm}m_t + b_y) \quad (18)$$

where σ is the sigmoid function, W is the weight matrices, W_{ix} is a matrix of fuzzy weights from the input cell to the output gate. b is the bias vector, and i, f, o, and c are the input gates, the hidden (forget) gate, the output gate, and the cell activation function. The cell output is represented by m_t, and $\odot$ is the element-wise product in Eq. (15). f_φ is the activation function of the network. The hidden unit in this architecture is represented in memory blocks. Each block contains one or a large number of memory cells. This procedure allows these cells to preserve information for a specific time in an uncertain time-series and decide which piece of information should be stored and when to use it.

3-3- Inference Layers

The fuzzifier can be considered into two models, singleton and non-singleton, according to the number of non-zero MF values and defines the membership grade of input. In this research, the Takagi-Sugeno fuzzy rule type is considered which is more precise than the Mamdani inference engine in the literatures. The product and minimum t-norms were used in inference methods besides the singleton fuzzifier. The dropout layers' parameters and the numbers of hidden layer is investigated in the Takagi-Sugeno inference engine. The output process, contains the type-reducer and the defuzzifier, generates the crisp output, which used the well-known Karnik-Mendel (KM) algorithm [24].

3-4- Mathematical Details of Fuzzy Layers

The input variables of the HT2FLSTM model can be defined as p where *p* is the inputs of the proposed HT2FLSTM model as follows:

$$\mathrm{P}\begin{Bmatrix} \mathrm{R}_{\mathrm{p}}^{\mathrm{j}_1\mathrm{j}_2}\text{: if } \mathrm{p}_{\mathrm{t}} = \tilde{\mathrm{F}}_{\mathrm{p}}^{\mathrm{j}_1} \text{ and } \mathrm{p}_{\mathrm{t}+1} = \tilde{\mathrm{F}}_{\mathrm{p}}^{\mathrm{j}_2} \\ \text{then } \mathrm{y}_{\mathrm{p}} = \left[\underline{\mathrm{c}}_{\mathrm{p}}^{\mathrm{j}_1,\mathrm{j}_2}, \bar{\mathrm{c}}_{\mathrm{p}}^{\mathrm{j}_1,\mathrm{j}_2}\right] \end{Bmatrix}^{\mathrm{M}} \quad (19)$$

where $\tilde{F}_p^{j_1}$ and $\tilde{F}_p^{j_2}$ are the interval type-2 fuzzy sets of the inputs *xp* and *xp+1* in HT2FLSTM , respectively, M is the number of applied fuzzy rules, $\underline{c}_i^{j_1,j_2}$ and $\bar{c}_i^{j_1,j_2}$ are the consequents of a fuzzy rule $R_p^{j_1j_2}$. The product t-norm operator in Eq. (20) is applied to compute the membership interval as follows [25-27]:

$$\begin{cases} \underline{f}^{l}(p') = \underline{\mu}_{p_1^l}(p'_1) \times \cdots \times \underline{\mu}_{p'_x}(p'_x) \\ \bar{f}^{l}(p') = \bar{\mu}_{p_1^l}(p'_1) \times \cdots \times \bar{\mu}_{p'_x}(p'_x) \end{cases} \quad (20)$$

The inference part of the HT2FLSTM model can be entirely characterized by *M* fuzzy rules employed in the inference process as:

$$\tilde{R}^l\text{: if } x_1 \text{ is } \tilde{F}^l \text{ ... and } x_p \text{ is } F_p^l \text{ Then y is } \tilde{G}^l \quad (21)$$

The relation for each fuzzy rule is as:

$$\tilde{R}^l\text{: } \tilde{F}^l \times \ldots \times F_p^l \to \tilde{G}^l = \tilde{A}^l \to \tilde{G}^l \quad (22)$$

The membership function of the rule is as,

$$\mu_{\tilde{R}^l}(x,y) = \mu_{\tilde{F}_1^l}(x'_1) \cap \cdots \cap \mu_{\tilde{F}_p^l}(x'_p) \cap \mu_{\tilde{G}^l}(y) \quad (23)$$

where ∩ signifies the product t-norm operation [26].
The output of each inference procedure is $\tilde{B}^l = \tilde{A}_x \circ \tilde{R}^l$, with membership functions of $\mu_{\tilde{B}^l}(y)$ as:

$$\mu_{\tilde{B}^l}(y) = \bigcup_{x \in X} \left[\mu_{A_x}(x) \cap \mu_{\tilde{A} \to \tilde{G}}(x,y)\right] \quad (24)$$

where ∘ defines the composition operation and ∪ represents the maximum t-conorm operation [25], and $\tilde{F}^l(\acute{x})$ is the membership interval for the fuzzy rule, where $x = \acute{x}$ and F^l is as

$$F^l(x') \equiv \left[\underline{f}^l(x'), \bar{f}^l(x')\right] \quad (25)$$

The firing output set $\tilde{B}^l$ is produced through a fuzzy rule and the aggregation of the consequent of the HT2FLSTM model as:

$$\tilde{B}^l\text{: }\begin{cases} FOU(\tilde{B}^l) = \left[\underline{\mu}_{\tilde{B}'}(y \mid x'), \bar{\mu}_{\tilde{B}'}(y \mid x')\right] \\ \mu_{\tilde{B}'}(y \mid x') = \underline{f}^l(x') * \underline{\mu}_{\tilde{G}'}(y) \\ \bar{\mu}_{\tilde{B}'}(y \mid x') = \bar{f}^l(x') * \bar{\mu}_{\hat{G}}(y) \end{cases} \quad (26)$$

where ∗ represents the product t-norm operation. The final output $\tilde{B}^l$ is considered as the integration of all rule firing sets $\tilde{B}^l$ on the output:

$$\tilde{B}^l\text{: }\begin{cases} FOU(\tilde{B}) = \left[\underline{\mu}_{\tilde{B}}(y \mid x'), \bar{\mu}_{\tilde{B}}(y \mid x')\right] \\ \underline{\mu}_{\tilde{B}}(y \mid x') = \underline{\mu}_{\tilde{B}^1}(y \mid x') \vee \cdots \vee \underline{\mu}_{\tilde{B}^M}(y \mid x') \\ \bar{\mu}_{\tilde{B}}(y \mid x') = \bar{\mu}_{\tilde{B}^1}(y \mid x') \vee \cdots \vee \bar{\mu}_{\tilde{B}^M}(y \mid x') \end{cases} \quad (27)$$

where ∨ is the max operation. Then the type reduced set $Y_C(\acute{x})$ is computed using the centroid $C_{\acute{B}}$ of $\tilde{B}$:

$$Y_C(x') = C_{\tilde{B}}(x') = \frac{1}{[l_{\tilde{B}}(x'), r_{\tilde{B}}(x')]} \quad (28)$$

where the two points $l_b(\tilde{x})$ and $r_b(\tilde{x})$ are computed through the *KM* algorithm [24].

4-Experimental Results

In this section, the evaluations and experimental results of the proposed model have been presented. In Section A, datasets description, features of the study, and correlation analysis of the features have been explained, respectively. In Section B, we presented the proposed model's accuracy with different cells and depths of layers to evaluate and compare the proposed HT2FLSTM model's accuracy in the training process. In C, a ROC curve analysis and other metrics for performance evaluation have been described. Then the statistical results of the null hypothesis are described in D. The experimental results measuring

uncertainty and comparative study have been discussed in E, respectively. For air quality prediction the well-known public time-series datasets of the latest public standard pollutant data provided by the Environmental Protection Agency (EPA) were employed to evaluate the efficiency of the HT2FLSTM model in real-world scenarios. This outdoor pollutant dataset includes NO_2, SO_3, PM_{10}, $PM_{2.5}$, and CO_2 variables for one decade from June 2011 to October 2020 in Tehran and Beijing. The dataset is available on the following link:
https://data.world/datasets/air-pollution

4-1- Performance of the HT2FLSTM Model

Table 2 summarizes the information by selecting and reporting the main features, including the number of cells, depth, and accuracy during the training process. Table 2 presents the proposed model's accuracy based on different datasets with different cells and depths of layers to compare the proposed HT2FLSTM model's training process. Obtained results in Figure 4 shows different layers and cells configuration. The results reveal that the HT2FLSTM with 20 Layers and 3000 cells is the most robust configuration, which reported the best performance during the prediction process, as depicted in the next section. Also, Figure.4 proves that more training steps obtain lower uncertainty, illustrated as the lower and upper bound.

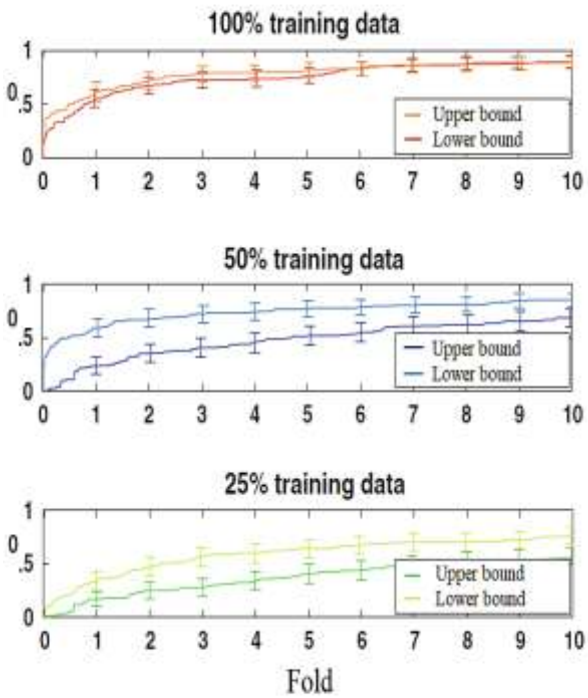

Fig.4. 10-fold cross-validation training accuracy

Table 2 :
Training Accuracy of the Ht2flstm Model

Cell	Depth	CO_2	SO_2	NO_3	$PM_{2.5}$	PM_{10}	Mean
10000	8	90	93	91	95	95	92
7000	10	91	93	92	95	95	93
5000	15	93	94	92	96	97	94
3000	20	95	96	95	97	96	96

4-2- ROC Curve Analysis and Statistical Evaluation

A ROC curve analysis is conducted to have a reliable estimate of the HT2FLSTM performance. The results were statistically verified. The following equations were used for evaluate the performance through a ROC curve analysis of the proposed model. Also, the standard metrics, such as precision and recall, were applied to evaluate the proposed model as follows:

$$\text{Precision } = \frac{\text{TP}}{(\text{TP} + \text{FP})} \times 100\% \tag{29}$$

$$\text{Recall } = \frac{\text{TP}}{(\text{TP} + \text{TN})} \times 100\% \tag{30}$$

$$\text{Accuracy } = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{FN} + \text{TN})} \tag{31}$$

$$\mu_i = \frac{1}{10} \sum_{k=1}^{10} AUCj \tag{32}$$

where μ_i is the means of the ROC curve's accuracy for the 10-fold cross-validation.

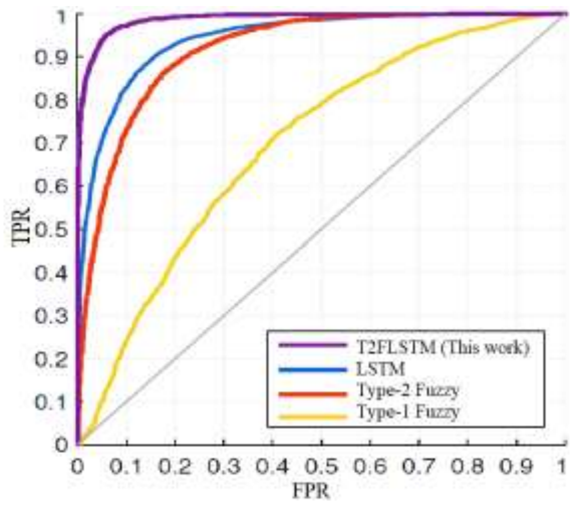

Fig.5. Comparison of ROC curve analysis

In order to show the robustness of the proposed model in comparison with counterpart models, Figure.5 shows the comparison of ROC curve analysis between the proposed

HT2FLSTM and type-1 fuzzy, type-2 fuzzy, and LSTM model.

4-3- Left Tailed T-Test and Time-Complexity Analysis

In this section, a two-sample t-test (left tailed) was applied to show the model's efficiency compared with the LSTM network. The null hypothesis is defined as $H_0 = \mu_i > \mu_j$ and $H_1: \mu_i < \mu_j$, where μ_i and μ_j are the means of the area under the ROC curve (AUC) of HT2FLSTM and LSTM for ten different runs of the cross-validation technique, respectively. The t-test results in Table 2 reveals the superiority of the proposed HT2FLSTM model for Air quality time-series prediction compared to LSTM. The t-test (according to the defined hypothesis test) failed to reject the defined null hypothesis. Also, time-series complexity is determined from the number of computation steps needed to run the algorithm as a function of input size.

Table 3 :
t-test analysis of the ht2flstm and lstm

Fold Number	HT2FLSTM	LSTM
1	0.9149	0.7724
2	0.9122	0.7111
3	0.9734	0.6351
4	0.9719	0.6221
5	0.9214	0.7128
6	0.9219	0.7815
7	0.9250	0.7641
8	0.9646	0.7401
9	0.9629	0.8001
10	0.9771	0.8111
Mean	0.9445	0.7351

Table 4:
average complexity of proposed ht2flstm model.

Pseudo Code	Average Complexity
Class Type	(8*O(1) + O(N*M) + O(N2))
Training Time	(7*O(1) + O(N2))
Calculating the Output	O(M*N)
Calculating the Error	O(1) + O(M*N)
The Output	O(1) + O(N)

4-4- Error Calculation and Technical Analysis

In this step, the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Percentage Error (MPE) have been applied to evaluate the performance of the HT2FLSTM model as follows:

$$\mathrm{MAE} = \frac{1}{2}\sum_{t=1}^{n} |(S_t - P_t)| \tag{33}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{2}\sum_{t=1}^{n} (S_t - P_t)^2} \tag{34}$$

$$\mathrm{MPE} = \frac{100\%}{n} = \sum_{t=1}^{n} \frac{(S_t - P_t)}{S_t} \tag{35}$$

The details of the obtained results in the Tehran air quality index time-series dataset are reported in Table 5. The experiments reveal that the proposed HT2FLSTM model's performance is better than its counterparts, such as the type-1 fuzzy logic system (T1FLS), type-2 fuzzy logic system (T2FLS), and the LSTM model with the same cell number and same layer. Table 5 presents a comparison of conventional methods and the proposed HT2FLSTM model. Unlike the fuzzy-inference counterparts, deep learning methods, such as the LSTM network, do not require handcrafted features. Instead, they rely on predefined structures and labeled data and can learn features on their own.

Table 5:
performance comparison of ht2flstm and counterparts

Method	AUC%	CI	Recall	Precision
T1FLS	69	[66-70] %	67%	69%
T2FLS	89	[87-89] %	89%	91%
LSTM	90	[89-91] %	91%	92%
HT2FLSTM *	97	[95-97] %	95%	98%

Figure. 6 illustrates the seasonal (long-term) comparison of the original observations and predicted values of the HT2FLSTM model for the Tehran air quality index with six different compelling features. The results show that the proposed HT2FLSTM model is 28% greater than T1FLS, 8% greater than T2FLS, 7% greater than LSTM in terms of AUC, Recall, and Precision.

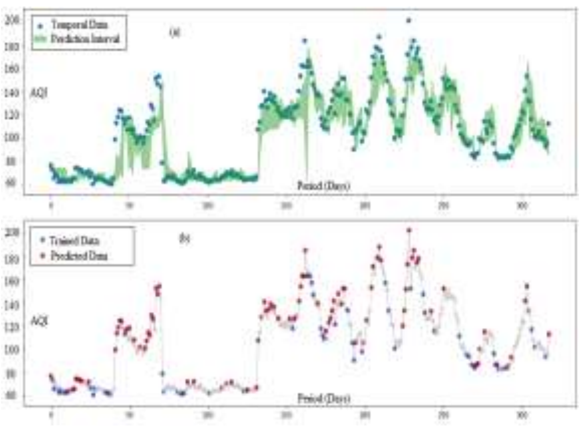

Fig.6. The HT2FLSTM applied to air quality prediction; (a) show uncertainty bounds, and (b) fit results of the time-series

According to Tables 6 and 7, the results confirm that the proposed HT2FLSTM model achieves better performance and lower error rates in terms of RMSE, MAE, and MPE, in Tehran and Beijing datasets.

Table 6:
Comparison results for HT2FLSTM in Tehran

Method	Tehran Dataset		
	MAE	*RMSE*	*MPE*
LSTM	0.068	0.041	1.81
IT1FLS	0.051	0.022	1.72
IT2FLS	0.032	0.039	1.61
HT2FLSTM (This work)	0.021	0.023	1.27

Table 7:
Comparison results for HT2FLSTM in Beijing

Method	Beijing Dataset		
	MAE	*RMSE*	*MPE*
LSTM	0.062	0.059	1.37
IT1FLS	0.041	0.51	1.26
IT2FLS	0.029	0.041	1.23
HT2FLSTM (This work)	0.021	0.025	1.14

5-Conclusion

The proposed HT2FLSTM model has more design degrees of freedom than a type-1 or other related models because of the FOU parameters in type-2 fuzzy sets and its potential to model inequality of time intervals. The proficiency of Type-2 FLSs has been proved in high-order uncertainties such as non-stationary events in time series. This study takes advantage of type-2 fuzzy logic and the LSTM network as a deep learning approach to develop the HT2FLSTM model. The experiments reveal that the performance of the proposed HT2FLSTM model is better than its counterparts. The results show that the proposed HT2FLSTM model is 28% greater than IT1FLS, 8% greater than IT2FLS, 7% greater than LSTM in terms of AUC. The proposed HT2FLSTM model has an average AUC of 97% with a 95% confidence interval [95-97] %. These results show that the proposed HT2FLSTM model can achieve reliable results in the temporal data and learning of long-term dependencies. The proposed HT2FLSTM model has additional design degrees of freedom than a type-1 or other related representations because of the FOU parameters and its budding to archetypal variation of temporal patterns on the time intervals. The experiment results confirm that the proposed model has lower error rates than other state-of-the-art algorithms. Additionally, the model can easily get updated because of its deep architecture when new cases are reported. Further studies are suggested to tune the HT2FLSTM cell structure parameters and apply them to other real-world applications associated with non-stationary uncertainties. Optimization algorithms can be used to improve the performance of the HT2FLSTM model.

References

[1] K. Greff, R.K. Srivastava, J. Koutník, B.R. Steunebrink, J. Schmidhuber, LSTM: a search space odyssey, IEEE Transactions on Neural Networks and Learning Systems 28 (10) (2017) 2222–2232.

[2] V. Georgescu. (2014). Joint propagation of ontological and epistemic uncertainty across risk assessment and fuzzy time series models, Computer Science and Information Systems. 11(2)

[3] Gaxiola F., Melin P., Valdez F., Castro J.R. (2018). Ensemble Neural Network with Type-2 Fuzzy Weights Using Response Integration for Time Series Prediction. Studies in Fuzziness and Soft Computing, vol 361.

[4] A. Safari, R. Hosseini, M. Mazinani, “A novel deep interval type-2 fuzzy LSTM (DIHT2FLSTM) model applied to COVID-19 pandemic time-series prediction”, Journal of Biomedical Informatics, Vol 123, (2021).

[5] R. Pascanu, T. Mikolov, Y. Bengio, On the difficulty of training recurrent neural networks, in: International Conference on Machine Learning (ICML), vol. 28, 2013, pp. 1310–1318.

[6] Li, R., Chen, X. and Balezentis, T. Multi-step least squares support vector machine modeling approach for forecasting short-term electricity demand with application. Neural Comput & Applic. 21(1) (2020) 1–20.

[7] Safari, A., & Hosseini, R. (2022). An Interval Type-2 Fuzzy LSTM Algorithm for Modeling Environmental Time-Series Prediction. Anthropogenic Pollution, 6(2), 62-72. doi: 10.22034/ap.2022.1963124.1133

[8] J. M. Mendel. (2013). on km algorithms for solving type-2 fuzzy set problems, IEEE Transactions on Fuzzy Systems, 21(3) 426–446.

[9] A Safari, M Mazinani, R Hosseini, A novel type-2 adaptive Neuro fuzzy inference system classifier for modelling uncertainty in prediction of air pollution disaster (research note), International Journal of Engineering, Vol. 30 (11), pp.1746-1751, 2017.

[10] A. Safari, R. Hosseini and M. Mazinani, “A Novel Type-2 Adaptive Neuro Fuzzy Inference System Classifier for Modelling Uncertainty in Prediction of Air Pollution Disaster”, IJE Transactions B: Applications, Vol. 30, No. 11, pp. 1568-1577, 2017.

[11] Safari, A., Hosseini, R., Mazinani, M. (2022). An Adaptive Intelligent Type-2 Fuzzy Logic Model to Manage Uncertainty of Short and Long Time-Series in Covid-19 Patterns Prediction: A Case Study on Iran. Computational Intelligence in Electrical Engineering.

[12] Y. Zhou, N. Zhou, L. Gong, and M. Jiang. (2020). Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine, Energy, 21(5) 117894.

[13] C. Li, C. Zhou, W. Peng, Y. Lv, and X. Luo. (2020). Accurate prediction of short-term photovoltaic power generation via a novel double input-rule-modules stacked deep fuzzy method, Energy, 11(7) 118700.

[14] N. Anh, S. Suresh, M. Pratama, and N. Srikanth. (2019). Interval prediction of wave energy characteristics using meta-cognitive interval type-2 fuzzy inference system, Knowledge-Based Systems, 169

[15] Y. Zhang, B. Yan, M. Aasma. "A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM", Expert Systems with Applications, (2020).

[16] X. Liu, Z. Zeng, D. C. Wunsch. (2020). Memristor-based LSTM network with in situ training and its applications, Neural Networks, 131, 300-311.

[17] Li, C., Wang, Z. and Rao, M Long. (2019). short-term memory networks in memristor crossbar arrays. Nat Mach Intell, 1(1) 49–57.

[18] E.Egrioglu, CH.Aladag, U.Yolcu, "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks", Expert Systems with Applications, Vol. 40, Issue 3,pp. 854-857, February 2013.

[19] SM. Chen, "Forecasting enrollments based on fuzzy time-series", Fuzzy Sets and Systems, Vol. 81, pp. 311–319, 1996.

[20] K. Ramadani and D. Devianto, The forecasting model of Bitcoin price with fuzzy time series Markov chain and chen logical method, AIP Conference Proceedings 2296, 020095 (2020).

[21] C.Cheng, G.W. Cheng, J.W. Wang, Multi-attribute fuzzy time series method based on fuzzy clustering, Expert Systems with Applications Vol.34, pp. 1235–1242, (2008)

[22] C.Cheng, T.L. Chen, H.J. Teoh, C.H. Chiang, Fuzzy time series based on adaptive expectation model for TAIEX forecasting, Expert Systems with Applications 34, pp. 1126–1132, (2008)

[23] R.C. Tsaur, J.C. Yang, H.F. Wang, Fuzzy relation analysis in fuzzy time series model, Computers and Mathematics with Application Vol.49, pp. 539–548, (2005)

[24] S.R. Singh, A simple method of forecasting based on fuzzy time series, Applied Mathematics and Computation Vol, 186, pp. 330–339, (2007)

[25] CH.Aladag, Basaran, E.Egrioglu,U.Yolcu, VR.Uslu; "Forecasting in high order fuzzy time series by using neural networks to define fuzzy relations", Expert Systems with Applications, Vol. 36, pp. 4228-4231, (2009)

[26] CH. Aladag, U. Yolcu, E.Egrioglu, Ali Z Dalar; "A new time invariant fuzzy time series forecasting method based on particle swarm optimization", Applied Soft Computing, Vol. 12, Issue 10, pp. 3291-3299, (2012)

[27] E.Egrioglu, CH.Aladag, U.Yolcu, VR.Uslu, "Fuzzy time series forecasting method based on Gustafson-Kessel fuzzy clustering" Expert Systems with Applications, Vol.38, pp. 10355-10357, (2011)

[28] E. Egrioglu, "PSO based higher order time invariant fuzzy time series method: Application to stock exchange data", Economic Modelling, Vol.38, pp.633-639, (2014)

[29] E.Egrioglu, CH.Aladag, U.Yolcu, "Fuzzy time series forecasting with a novel hybrid approach combining fuzzy c-means and neural networks", Expert Systems with Applications, Vol. 40, Issue 3,pp. 854-857, (2013)

[30] Safari, A., & Hosseini, R. (2023). Evolutionary Interval Type-2 Fuzzy Rule Learning Approaches for Uncertain Time-Series Prediction. Signal Processing and Renewable Energy, 7(1), 27-39.

[31] A. Safari, R. Hosseini and M. Mazinani, “A Type-2 Fuzzy Time Series Model for Pattern Similarity Analysis: A Case Study on Air Quality Forecasting”, in IEEE Intelligent Systems, 2021.

[32] A. Safari, R. Hosseini and M. Mazinani, " Dynamic Type-2 Fuzzy Time Warping (DT2FTW): A Hybrid Model for Uncertain Time-Series Prediction", International Journal of Fuzzy Logic and Intelligent Systems Vol. 21, pp.338-348, 2021.