Information Sciences 679 (2024) 121007

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

A new ensemble intuitionistic fuzzy-deep forecasting model: Consolidation of the IFRFs-bENR with LSTM

Ozge Cagcag Yolcu, Ufuk Yolcu *

Department of Statistics, Marmara University, Istanbul, Türkiye

ARTICLE INFO

Keywords:
Intuitionistic fuzzy regression functions
LSTM
Genetic algorithm
Elastic-net regularization
Linear and nonlinear relations
Forecasting

ABSTRACT

Among forecasting model families, the intuitionistic fuzzy-based forecasting model stands out due to its comprehensive approach to uncertainty, considering possible degrees of hesitation. This study offers a forecasting model that consolidates intuitionistic fuzzy regression functions based on elastic net regularization (IFRFs-bENR) with LSTM. The proposed consolidated model, unlike existing models, is capable of modelling both linear and nonlinear structures that coexist between inputs and outputs. Another noteworthy aspect of the consolidated forecasting model is its method of determining model hyperparameters through a systematic optimization process using GA, in contrast to the trial-and-error approach prevalent in most literature studies. The validity and consistency of the model were assessed by running the model 50 times with the optimal hyperparameter values obtained for the consolidated model. And thus, the experimental probability distributions of the forecasts were also obtained. The proposed consolidated model also outperforms its peers in this aspect. The consolidated forecasting model was applied to different sets of time series, including TAIEX, DJI, SSEC, and IstEX. The findings indicate that the proposed consolidated model produces more accurate forecasts compared to various selected benchmark models. All abbreviations used in the article are defined in Supplementary Table 1 under the List of Abbreviations.

1. Introduction — *Gaps-Deficiencies, Motivation, and contributions*

In whichever fields, medicine, economics, finance, engineering, or others, the insights derived from time series forecasting problems provide powerful and effective information for guiding economic, social, and political decision-making. In the literature, due to the importance of this subject, numerous time series forecasting models, each with its distinct properties, have been proposed in distinguished journals or various other published platforms, and the majority of them have been utilized by practitioners. These forecasting models can be broadly categorized into two groups: probabilistic and non-probabilistic models. Probabilistic models are often referred to as statistic-based and must satisfy strict assumptions to yield reliable results. However, these rigid assumptions cannot be met in many cases, and in such cases, these models may produce unsatisfactory results. This factor is considered one of the main reasons for the emergence and rapid spread of non-probabilistic models in time series forecasting literature. While most non-probabilistic models offer fuzzy-based or computational-based approaches, some hybrid approaches combine these two. Although computational-based models based on traditional and deep neural networks produce successful results, they do not offer an approach or perspective to the uncertainty/vagueness that the time series inevitably contains. At this point, hybrid models possess superior

* Corresponding author.
E-mail addresses: ozge.cagcag@marmara.edu.tr (O. Cagcag Yolcu), ufuk.yolcu@marmara.edu.tr (U. Yolcu).

https://doi.org/10.1016/j.ins.2024.121007
Received 27 November 2023; Received in revised form 7 June 2024; Accepted 8 June 2024
Available online 22 June 2024
0020-0255/

features, both in terms of their ability to closely conform to the data patterns of neural networks and thanks to the approach offered by fuzzy-based models for handling uncertainty.

Especially in recent years, fuzzy time series (FTS) models first proposed by Song and Chisom [1,2], based on fuzzy sets introduced by Zadeh [3], have dominated the literature due to their superior features. The main shortcoming these early studies in the literature, and of almost all studies so far, can be seen as their consideration of only one measure of uncertainty – *membership degree*- in the analysis process. These constraints severely limit the models' ability to comprehensively assimilate uncertainty and improve forecast accuracy. However, nearly all of the fuzzy phenomena involve a hesitation degree as well as a membership degree. In the decision-making process, hesitations are inevitable, as judgments are typically influenced by various determiners. And so, the existence of a neutrality (hesitation) degree also necessitates taking a non-membership degree into account when addressing an approach to uncertainty. As a proponent of this view, Atanassov [4] proposed the intuitionistic fuzzy sets (IFSs) to define, interpret, and model uncertainty. Especially in recent years, time series forecasting models designed based on Atanassov's intuitionistic fuzzy sets have been widely used in the literature. While some of these studies, which have been detailed in the literature review section of the study, have produced forecasts based on rule-based relations, others have generated the outputs using neural networks. In addition to this, there are also a few studies that have utilized a regression model based on the intuitionistic fuzzy regression function [5,6].

Due to their specific and distinguishing properties, various intuitionistic fuzzy set-based time series (I-FTS) forecasting models in the literature exhibit some individual superiority. However, this does not imply that each of them is devoid of shortcomings and flaws. The more these deficiencies are addressed, the higher the forecasting accuracy in time series will become. Consequently, more valid and reliable results will be achieved.

The motivation of this study is to propose a forecasting model that will fill the gaps and eliminate the deficiencies in the literature as much as possible. Thus, the aim is to produce higher accuracy and more reliable forecasts. From a detailed perspective, in contrast to the gaps in the existing literature, the features and contributions of the proposed model can be summarised as follows.

- Approach to Uncertainty

Gap(s): The existence of hesitations is inevitable because judgments in the decision-making process are influenced by various factors. Current FTS approaches overlook this phenomenon, leading to a lack of comprehensive treatment of internal uncertainty due to their failure to consider both the degree of neutrality (hesitancy) and non-membership degrees.

Contribution(s): In this respect, the proposed ensemble forecasting model offers an effective approach to uncertainty by considering the memberships and hesitation degrees together in the analysis process. In addition, it uses the information from both membership and non-membership degrees (moreover, some of their transformations) together, thereby improving forecasting performance.

- Relationship Structure

Gap(s): Almost all models introduced in the literature assume that the relevant time series has only one of the linear or nonlinear relationships. These models can only model either linear or nonlinear relationship structures. Indeed, many time series contain both linear and nonlinear relationships, subject to the forecasting problem.

Contribution(s): The proposed ensemble forecasting tool has the ability to model linear and nonlinear relationships together and simultaneously, thus it is capable of adapting the model to the solution surfaces of the data at a high level of success.

- Relationship Establishment

Gap(s): Apart from a few exceptions, whether it is fuzzy set-based or intuitionistic fuzzy set-based, current approaches determine relationships through relationship tables as based solely on cluster numbers (indices) without using membership and/or non-membership degrees. This can lead to a loss of information in the analysis process and thus negatively affect the forecasting performance.

Contribution(s): The proposed ensemble forecasting model uses memberships, non-memberships (and moreover some of their transformations) as well as lagged time series with the real values as inputs while determining relationships in both its linear and nonlinear components. In this respect, it is more competent than existing approaches in modelling and extracting the interrelationship of the time series.

- Model Assumptions – for Linear Component

Gap(s): In the limited literature on intuitionistic fuzzy regression functions-based studies that present linear structured models, an important assumption, collinearity, of the regression model has been ignored. However, this situation causes a significant issue, such as an increase in the estimation variance for the parameter estimates of the regression model. This also presents a critical problem for the consistency of the forecasts.

Contribution(s): In the proposed ensemble model, whose linear component is based on intuitionistic fuzzy regression functions (IFRFs), this vital assumption of the regression model is taken into account during the construction of the IFRFs. In this direction, with this study, intuitionistic fuzzy regression functions with elastic net regularization have been introduced for the first time in the literature. The use of IFRFs-bENR prevents the forecasts produced by the proposed forecasting model from having inconsistencies due to collinearity problems.

- Learning Strategy – for Non-Linear Component

Gap(s): As mentioned before, except for a few, existing studies use relationship tables while determining relationships. In some studies, while artificial neural network (ANN) structures are used to determine relationships, shallow learning strategies are preferred as a learning strategy. However, the relationships between model inputs and targets/outputs are generally in a complex and deep format for the time series. Shallow learning strategies may fail to model these deep and complex relationships. This situation can be considered a factor that hinders forecasting tools from achieving the desired performance.

Contribution(s): The proposed ensemble model uses LSTM, a deep neural network in its nonlinear component, to establish relationships. It is more successful in modelling and extracting relationships in complex and deep formats, owing to the used deep strategy in LSTM training. Furthermore, it is the only forecasting model in which a deep neural network – LSTM- is used to determine relationships, with its input comprising all information including memberships, non-memberships (as well as some transformations thereof), and lagged time series with real values.

- Interpretable Forecasts

Gap(s): Except for a few studies, almost all research has concentrated on enhancing forecasting performance rather than attempting to generate interpretable forecasts. However, interpretable outputs and forecasts are crucial for decision-makers to create strategies for the future. Interpretability here refers to the ability to generate a probability distribution for the forecasts. Thus, specific probabilities can be assigned regarding the range within which they can take values.

Contribution(s): In this proposed study, the consolidated model is run 50 times to obtain an experimental probability distribution for each forecast. Thus, interpretable forecasts are generated for each of which probabilities can be determined. The proposed consolidated model is also superior to its counterparts in this regard.

- Tuning of Hyperparameters and Model Selection

Gap(s): In almost all studies, the hyper-parameters and optimal models were determined by a trial-and-error approach. Two problems can be mentioned caused by this situation. First, only a predetermined subset of all possible points in the search space can be explored, which often makes reaching optimality both challenging and time-consuming. Secondly, this search is not performed through an analytical process. Moreover, in some cases, this selection is carried out based on the test set performances of the models. This means that contrary to the nature of the estimation problem, the method sees future data points.

Contribution(s): In this study, the selection of all hyperparameters and the determination of the best model are performed by an automatic mechanism during the forecasting process by a genetic algorithm (GA). This entire process is carried out on validation sets instead of test sets.

In summary, in this study, a prediction model designed to address the gaps in the literature and present the stated contributions is proposed. The proposed forecasting model consolidates IFRFs-bENR with a deep neural network, LSTM (IFRFs-bENR&LSTM). Model inputs are created using the intuitionistic fuzzy C-means (IFCM) clustering algorithm. Thanks to the use of the IFRFs-bENR and LSTM, it is capable of simultaneously modelling both linear and nonlinear structures. The forecasts produced by each component are combined with unequal weights to obtain the final forecasts for the proposed consolidated model. All hyperparameters related to the forecasting process are determined analytically by the GA through an optimization process.

The rest of the paper is organized as follows. In the second section, an in-depth literature review is presented. The third section presents information about the methods and algorithms that form the background of the proposed ensemble model, such as IFCM, IFRFs-bENR, and LSTM. Section 4 introduces the proposed ensemble forecasting model in an elaborate manner and presents its characteristics and working principles. Section 5 gives the properties of the datasets and presents scenarios and perspectives for implementation and evaluation. Section 6 presents detailed compilations and discussions related to results via some tables and figures. In Section 7, the results are evaluated from a holistic perspective. Moreover, some prospective suggestions are given in the last section.

2. Literature review

Although conventional forecasting models are widely known and used, they may not produce satisfactory output in many cases because they have to satisfy certain strict assumptions. This inadequacy and deficiency have led scientists dealing with forecasting problems to use alternative approaches based on advanced algorithms such as fuzzy-based and computational-based.

The first fuzzy-based forecasting models determined the intervals constant and arbitrarily, which they used in the fuzzification transaction [2]. While many of studies are based on fuzzy C-means (FCM) proposed by Bezdek [7], different clustering algorithms have also been used in the fuzzification stages of forecasting models [8]. In early FTS forecasting models, the fuzzy relations were determined using a fuzzy logic relation matrix [2]. For determining fuzzy relationships, fuzzy logic relationship tables have been used in the following studies [9]. In subsequent studies, for the aim of the determination of fuzzy relations, ANNs with different structures were widely used [10,11].

On the other hand, intuitionistic fuzzy set-based models have been preferred as an alternative perspective. When the fuzzy time series literature is examined, intuitionistic fuzzy methods have some advantages in obtaining satisfactory prediction results compared to other methods thanks to considering the memberships and non-memberships values together. Kumar et al. introduced the intuitionistic fuzzy time series method to figure out the nondeterminism in time series forecasting [12]. In that study, membership and non-membership were not considered in determining the relationships, and relationship and group relationship tables were used. Another deficiency of this study is that the testing set was not used in the forecasting and evaluation processes. Denghua et al. proposed an intuitionistic fuzzy time series approach particularly when there is data missing in transmission (IFTS) to predict satellite clock errors. Intuitionistic fuzzy C-means clustering algorithm was preferred to optimize universe of discourse. This study is also based on fuzzy rules and needs a defuzzification transaction, which may increase the forecasting error [13]. Bisht and Kumar showed the effect of an aggregation operator on prediction results. The proposed method produced better results when multiple valid fuzzification methods were available to fuzzify time series data, by taking advantage of considering the hesitation information based on weighted membership grades in fuzzy sets with equal and unequal interval lengths. While the degree of hesitation is used in the fuzzification of datasets, it is not used to identify relationships produced by relationship tables. In addition, the defuzzification phase is mandatory in this study, and the test set is not used in the estimation and evaluation processes [14]. Fan et al. introduced a forecasting model called a long-term intuitionistic fuzzy time series based on a curve similarity measure algorithm. On the other hand, in the analysis process, vector quantization was introduced for the pre-processing and the fuzzification stage. IFCM was used to determine optimal cluster centres [15]. Zheng et al. proposed a long-term intuitionistic fuzzy time series forecasting model. By introducing the vector forecasting technique prediction accuracy of long-term time series was improved [16]. With the idea that building a high-order model can increase prediction accuracy, Wang et al. proposed a high-order intuitionistic fuzzy time series forecasting model based on multi-dimension intuitionistic fuzzy modus ponens inference [17]. Wang et al. proposed a multi-factor high-order intuitionistic fuzzy time series forecasting model based on the usage of unequal intervals for the partition of a universe of discourse and multi-factor high-order forecast rules which make the model more sensitive to the fuzzy variation. This study did not consider membership and non-membership in determining the relationships. Also, the testing set was not used in the forecasting and evaluation processes. Another weakness of the study is the requirement for defuzzification to obtain the final forecasts [18]. Fan et al. introduced a novel IFTS forecasting model aimed at addressing both the difficulty in determining the optimal order number and the limitations of partitioning in existing models. The model can determine the partitions and intervals adaptively by forecasting the vector operator matrix. Similarly, this study uses only the cluster numbers of the observations without using the membership values while determining the relationships [19]. On the other hand, Egrioglu et al. proposed a high-order intuitionistic forecasting model by considering membership, non-membership, and hesitation margins. Principal components analysis was used to reduce the dimensions and with the usage of pi-sigma artificial neural network fuzzy relations were determined. And necessary parameters were optimised through the artificial bee colony (ABC) algorithm. Due to the nature of the neural network that determines the relationships, the estimation method proposed in this study can only model nonlinear relationships [20]. Hajek et al. introduced a new interval-valued time series (ITS) forecasting method called the Intuitionistic Fuzzy Grey Cognitive Map (IFGCM) that relies on fuzzy cognitive maps to provide accurate results with an effective representation of uncertainty in the data. In the study to be able to determine the IFGCM weights and parameters of the membership functions, the DE algorithm was preferred [21]. In addition to the usage of Long short-term memory (LSTM) in so many fields such as speech recognition, language modelling, and image processing, the effectiveness of LSTM on time series forecasting is indisputable. From this point of view, a deep high-order intuitionistic fuzzy time series forecasting method was proposed by Kocak et al. in which LSTM was preferred for determining the fuzzy relations stage. The hyperparameters were selected by trial-and-error method, and the results were given over the testing sets [22]. Fan et al. put forward a long-term intuitionistic fuzzy time series model to forecast traffic networks with multi-input multi-output structures. In the study, an improved vector-changing pattern was used to modify the IFCM clustering method [23]. A probabilistic intuitionistic fuzzy set-based model was introduced by Pattanayak et al. [24]. Differently in this study, to be able to forecast time series data set rather than original crisp observation the ratio trend variation of each crisp observation was evaluated. Additionally, a new automated trend-based discretization method was proposed to determine the length and number of intervals, making it efficiently applicable to various time series data without expert assistance. Intuitionistic fuzzy sets and self-organized direction aware (SODA) approach based computational fuzzy time series forecasting method was proposed by Pant and Kumar [25]. Unlike traditional clustering algorithm with the introduced SODA, both spatial and angular divergences for better comprehension of the data's ensemble characteristics were considered. One of the advantages of the proposed IFS and SODA based computational model is not building relational equations using sophisticated min–max operations and does not need defuzzification as well. When the introduced approaches are examined it is seen that combinations of different methods provide better forecasting results. Bas et al., put forward a new bootstrapped-based intuitionistic fuzzy time series method that uses a pi-sigma artificial neural network to model the intuitionistic fuzzy relations [26]. In the analyses process all the hyper parameters were determined by particle swarm optimization. In time-series forecasting literature, almost all used inference systems have just considered the univariate structure. Cagcag Yolcu et al. proposed a multivariate intuitionistic fuzzy inference system in that the sigma-pi neural network was used as an inference tool and not only obtained membership and non-membership values from IFCM but also lagged crisp observations of multivariable time-series were evaluated as an input of the system [27]. Wang et al. presented an interpretable intuitionistic fuzzy inference model which includes an indivisible six-layer neuro-fuzzy network model [28]. Kocak et al. proposed a method for forecasting Primary Energy Consumption for 23 countries. The method is explainable, robust, and based on high-order intuitionistic fuzzy time series [29]. Dixit and Jain aimed to present a forecasting method for non-stationary time series using the intuitionistic fuzzy time series clustering technique [30]. Egrioglu and Bas introduced a robust intuitionistic fuzzy regression function approach using the robust regression-based Welsch, Bisquare, Talwar, Huber, Logistic, and Cauchy functions instead of the ordinary least squares method [31]. Cakır recommended a hybrid methodology that combines the intuitionistic fuzzy time series method with the intuitionistic fuzzy c-means method for renewable energy generation forecasting [32]. Mohan Pattanayak et al. introduced an innovative hesitant fuzzy time series forecasting model that utilizes a support vector machine. They conducted a comparative evaluation of the results using sixteen different time series datasets [33]. Wan et al. framed the process of selecting a battery supplier as a time-series-based, multi-criteria large-scale group decision-making (LSGDM) problem, incorporating intuitionistic fuzzy information [34]. Dwivedi et al. introduced an intuitionistic fuzzified time-series prediction framework that employs logical relationships to deal with non-determinism with time-series prediction [35]. Pant et al. proposed a computational-based partitioning and Strong α, β-cut based novel intuitionistic fuzzy time series based on intuitionistic fuzzy logical relationships. Pan et al. used basic statistical parameters to determine the number of intervals and construct of intervals without specialized knowledge of the domain [36]. Moreover, Ashraf et al. introduced high-order methods based on circular intuitionistic fuzzy set theory principles, again using intuitionistic fuzzy logical relationships [37]. Despite the advancements made by these approaches in the literature, their reliance on rule-based systems and intuitive fuzzy logical relationships is considered a significant limitation. Another perspective for examining intuitionistic fuzzy models is to evaluate the relationships between inputs and outputs. Almost all studies consider only linear relationships or only nonlinear relationships. So, evaluating these two types of relations together may affect the forecasting performance. From this point of view, Cagcag Yolcu and Yolcu proposed a novel cascaded structure intuitionistic fuzzy time series prediction model which can model linear and nonlinear relationships between inputs and outputs, together [38].

3. Background

This study presents an ensemble intuitionistic fuzzy-deep time series forecasting model (E-IFD-TSFM). The proposed E-IFD-TSFM is indeed a consolidated model. It is created by consolidating intuitionistic fuzzy regression functions having elastic net regularization with LSTM, a deep neural network. A part of the input elements for the proposed forecasting model is determined via the IFCM. IFCM produces the membership and non-membership values of the respective fuzzy sets for each observation point. Membership and non-membership values, along with some transformations of both, as well as the main explanatory variables with their actual values, constitute the input matrices of the forecasting system. All hyperparameters of the proposed E-IFD-TSFM are determined analytically by the GA in an optimization process. From this point, in this section, GA, IFCM, IFRs-bENR, and LSTM, which form the background of the proposed E-IFD-TSFM, are introduced along with some of their key features.

3.1. Genetic algorithm

GA is a metaheuristic algorithm that belongs to the class of evolutionary algorithms designed for optimization. GA is inspired by the natural selection process [39]. While conventional optimization techniques, such as hill climbers, evaluate and develop a single solution, GA can develop multiple solutions called populations composed of individuals. Each chromosome /individual consists of a gene series and represents a potential solution that is an independent point of the solution space. Since GA is a widely used and almost anonymous algorithm by many researchers, no further details are given in this study about GA.

3.2. The intuitionistic fuzzy C-means clustering algorithm

The intuitionistic fuzzy sets proposed by Atanassov [4] consider both membership $\mu(x)$, $x \in X$ and non-membership $\nu(x)$ functions in handling the uncertainty and vagueness in the data sets. With this characteristic, IFSs provide a deeper and more realistic mathematical framework than Zadeh's classical fuzzy sets [3].

An IFS A in X, can be defined as:

$$A = \{x, \mu_A(x), u(x)/x \in X\} \tag{1}$$

where $\mu_A(x)\to[0,1]$ is a membership degree and $u_A(x)\to[0,1]$ of an element x in the IFS A. $\mu_A(x)$ and $u_A(x)$ are generated with the condition,

$$0 \le \mu_A(x) + u_A(x) \le 1 \tag{2}$$

When $u(x) = 1 - \mu_A(x), \forall x \in X$ in an IFS A then the set A turns into a classical fuzzy set. For each IFSs, Atanassov also introduced a hesitation degree $\pi(x)$, which represents the lack of knowledge in determining the membership degree of each element x in set A. The hesitation degree, which is defined as a different measure from membership and non-membership degrees, is given as:

$$\pi_A(x) = 1 - \mu_A(x) - u_A(x) \tag{3}$$

Here, it is clearly to be $0 \le \pi_A(x) \le 1$. If $\pi_A(x), \forall x \in X$, then the IFS A becomes a fuzzy set. From all these given, the membership values will be in the range of $[\mu_A(x), \mu_A(x) + \pi_A(x)]$.

The IFCM algorithm used in this study aims to minimize the objective function introduced by Chaira [40].

$$J_{IFCM} = \sum_{i=1}^{c}\sum_{k=1}^{N} \mu_{ik}^{*m} d_{ik}^2 + \sum_{i=1}^{c} \pi_i^* exp\left(1 - \pi_i^*\right) \tag{4}$$

The objective function J_{IFCM} has two components, the objective function of the conventional FCM algorithm and the intuitionistic fuzzy entropy (IFE).

Intuitionistic fuzzy membership, for i th IFS and k th datum point, μ_{ik}^* can be calculated as:

$$\mu_{ik}^* = \mu_{ik} + \pi_{ik} \tag{5}$$

here μ_{ik} is the conventional fuzzy membership of the k th datum point in i th fuzzy set. For the same data point and IFS, the hesitation degree π_{ik} is given as:

$$\pi_{ik} = 1 - \mu_{ik} - \left(1 - \mu_{ik}^{\alpha}\right)^{1/\alpha}, \alpha > 0 \tag{6}$$

Yager's intuitionistic fuzzy complement (IFC) also is created as:

$$N(X) = (1 - x^{\alpha})^{1/\alpha}, \alpha > 0 \tag{7}$$

So, via Yagers' IFC, the IFS becomes;

$$A_{\lambda}^{IFS} = \left\{x, \mu_A(x), (1 - \mu_A(x)^{\alpha})^{1/\alpha} \big/ x \in X\right\} \tag{8}$$

and the hesitation degree

$$\pi_A(x) = 1 - \mu_A(x) - (1 - \mu_A(x)^{\alpha})^{1/\alpha} \tag{9}$$

$$\pi_i^* = \frac{1}{N}\sum_{k=1}^{n} \pi_{ik}, k \in [1, N] \tag{10}$$

IFE, which forms the second term of the objective function, represents the amount of fuzziness or uncertainty in a set. IFE which represents the intuitionism degree is given by:

$$IFE(A) = \sum_{i=1}^{n} \pi_A(x_i) exp(1 - \pi_A(x_i)), k \in [1, N] \tag{11}$$

and

$$\pi_A(x_i) = 1 - \mu_A(x_i) - u_A(x_i) \tag{12}$$

In each iterative stage the cluster centres are modified as:

$$v_i^* = \frac{\sum_{k=1}^{n} \mu_{ik}^* x_k}{\sum_{k=1}^{n} \mu_{ik}^*} \tag{13}$$

When the maximum iteration number or $max_{ik}\left|u_{ik}^{*new} - u_{ik}^{*previous}\right| < \varepsilon$ (ε is a pre-defined value), updating of the cluster centres and memberships is stopped, and the process has been completed.

3.3. Intuitionistic fuzzy regression functions with Elastic-Net regularization

For the purpose of time series forecasting, there are a few studies using IFRFs [5,6,41]. In a model using IFRFs, the forecasting process is based on a combination of more than one multiple linear regression model generated. IFRFs, due to the nature of each function's inputs, it is inevitable to encounter a collinearity problem. Existing studies have not addressed a multicollinearity issue that may be valid for the multiple regression model, with the exception of only one study [6] in which a ridge regression-based approach. The ridge regression aims to minimize the loss function given as follows.

$$L_{Ridge}(\widehat{\beta}) = \sum_{k=1}^{N} (y_k - x'_k)^2 + \lambda \sum_{j=1}^{p} \widehat{\beta}_j^2 \tag{14}$$

Another regularization approach designed to overcome the collinearity problem is the Least Absolute Contraction and Selection Operator (Lasso). The Lasso aims to minimize the loss function given below.

$$L_{Lasso}(\widehat{\beta}) = \sum_{k=1}^{N} (y_k - x'_k)^2 + \lambda \sum_{j=1}^{p} |\widehat{\beta}_j| \tag{15}$$

Though Ridge and Lasso Regression look similar in outline, Ridge regression considers all correlated explanatory variables at a specific level, whereas Lasso regression considers only some of them. Moreover, Also, a hybrid regularization procedure called Elastic Net was proposed as a convex combination of Ridge and Lasso Regression. The Elastic Net regularization procedure is operated for the purpose of minimizing the loss function given below.

$$L_{Enet}(\widehat{\beta}) = \frac{\sum_{k=1}^{N} (y_k - x'_k\beta)^2}{2N} + \lambda \left[\frac{(1-\alpha)}{2} \sum_{j=1}^{p} \widehat{\beta}_j^2 + \alpha \sum_{j=1}^{p} |\widehat{\beta}_j|\right] \tag{16}$$

Here, N and p are # of datum points and # of explanatory variables. Also, y_k and x_{kj} represent k th response and k th datum points for j th explanatory variable, respectively. $\alpha \in [0, 1]$ and $\lambda \in [0, \infty)$ are shrinkage parameters. In terms of the features or the explanatory variables, while Lasso Regression can be seen as a selection mechanism, Ridge regression can be seen as a rating mechanism. Elastic net, by combining Ridge and Lasso Regression, has characteristics and advantages of both.

In the literature, although there is a study introducing a fuzzy regression function with elastic net regularization [42], in this study, as a linear component of the proposed ensemble forecasting model, IFRFs-bENR has been first proposed in the literature. Moreover, the shrinkage parameters α and λ for the proposed IFRFs-bENR, as another distinguishing feature, are determined through an optimization process with GA.

The inputs of the IFRFs-bENR are determined via IFCM. IFCM, for each datum point, produces the membership and the non-membership values. Some derivatives/transformations of the produced memberships and non-memberships are created. Subsequently, they are used in input matrices alongside the original explanatory variables. In the working principle of IFRFs-bENR, two different sets of multiple regression models are created. One is for memberships, while the other is for non-memberships. Also, each model set includes multiple regression models in the number of intuitionistic fuzzy sets. These models can be given with a mathematical notation as follows.

$$^{\mu}Y^{(i)} = {}^{\mu}X^{(i)}\ {}^{\mu}\boldsymbol{\beta}^{(i)} + {}^{\mu}\varepsilon^{(i)}, i = 1, 2, \cdots, c \tag{17}$$

$$^{u}Y^{(i)} = {}^{u}X^{(i)}\ {}^{u}\boldsymbol{\beta}^{(i)} + {}^{u}\varepsilon^{(i)}, i = 1, 2, \cdots, c \tag{18}$$

The input matrices and output vectors are also given below.

$$^{\mu}X^{(i)} = \begin{bmatrix} \mu_{i1} & \mu_{i1}^2 & exp(\mu_{i1}) & log\left(\frac{1-\mu_{i1}}{\mu_{i1}}\right) & x_{11} & x_{21} & \cdots & x_{p1} \\ \mu_{i2} & \mu_{i2}^2 & exp(\mu_{i2}) & log\left(\frac{1-\mu_{i2}}{\mu_{i2}}\right) & x_{12} & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \mu_{iN} & \mu_{iN}^2 & exp(\mu_{iN}) & log\left(\frac{1-\mu_{iN}}{\mu_{iN}}\right) & x_{1N} & x_{2N} & \cdots & x_{pN} \end{bmatrix}, i = 1, 2, \cdots, c \tag{19}$$

$$^{u}X^{(i)} = \begin{bmatrix} u_{i1} & u_{i1}^2 & exp(u_{i1}) & log\left(\frac{1-u_{i1}}{u_{i1}}\right) & x_{11} & x_{21} & \cdots & x_{p1} \\ u_{i2} & u_{i2}^2 & exp(u_{i2}) & log\left(\frac{1-u_{i2}}{u_{i2}}\right) & x_{12} & x_{12} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ u_{iN} & u_{iN}^2 & exp(u_{iN}) & log\left(\frac{1-u_{iN}}{u_{iN}}\right) & x_{1N} & x_{2N} & \cdots & x_{pN} \end{bmatrix}, i = 1, 2, \cdots, c \tag{20}$$

$$^{\mu}Y^{(i)} = {}^{u}Y^{(i)} = Y^{(i)} = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \end{bmatrix}^T \tag{21}$$

The Elastic-Net Regression parameter estimations are calculated as in Equations (22) for each model of membership component.

$$^{\mu}_{elastic}\widehat{\boldsymbol{\beta}}^{(i)} = \underset{\left({}^{\mu}\beta_0^{(i)}\ {}^{\mu}\beta^{(i)}\right) \in \mathbb{R}^{p+1}}{min} \sum_{k=1}^{N}\left(y_k - {}^{\mu}\beta_0^{(i)} \sum_{j=1}^{p} {}^{\mu}\beta_j^{(i)}\ {}^{\mu}x_{kj}\right)^2 + {}^{\mu}\lambda\left[(1 - {}^{\mu}\alpha)\frac{\| {}^{\mu}\beta\|_2^{\ 2}}{2} + {}^{\mu}\alpha\| {}^{\mu}\beta\|_1\right] \tag{22}$$

where;

y_k: the k th response.

$^{\mu}x_{kj}$: the k th row of j th column of the input matrix . $^{\mu}X^{(i)}$

N: the number of observations.

p: the number of columns of the input matrix . $^{\mu}X^{(i)}$

$^{\mu}\beta$: parameter vector.

Moreover, $\| {}^{\mu}\beta\|_1$ and $\| {}^{\mu}\beta\|_2$ are l_1 $-norm$ and l_2 $-norm$ ($Euclidean$ $-norm$). l_1 and l_2 norms are given in (23) and (24).

$$\| {}^{\mu}\beta\|_1 = \sum_{j=1}^{p} | {}^{\mu}\widehat{\beta}_j| \tag{23}$$

$$\| {}^{\mu}\beta\|_2 = \sqrt{\sum_{j=1}^{p} {}^{\mu}\widehat{\beta}_j} \tag{24}$$

Similarly, the Elastic-Net Regression parameter estimations are calculated as in Equations (25) for each model of non-membership component.

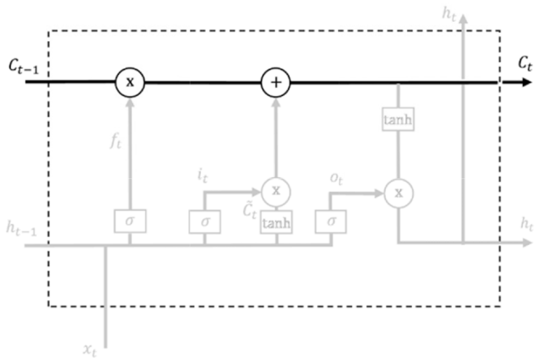


Fig. 1. LSTM cell state [44].

$$ {}^{u}_{elastic}\widehat{\beta}^{(i)} = \min_{\left({}^{u}\beta_0^{(i)}\ {}^{u}\beta^{(i)}\right)\in\mathbb{R}^{p+1}} \sum_{k=1}^{N}\left(y_k - {}^{u}\beta_0^{(i)}\sum_{j=1}^{p}{}^{u}\beta_j^{(i)}\ {}^{u}x_{kj}\right)^2 + {}^{u}\lambda\left[(1-{}^{u}\alpha)\frac{\|{}^{u}\beta\|_2^{\,2}}{2} + {}^{u}\alpha\|{}^{u}\beta\|_1\right] \tag{25} $$

Thus, the estimated models are obtained as follows.

$$ {}^{\mu}_{elastic}\widehat{Y}^{(i)} = {}^{\mu}X^{(i)}\ {}^{\mu}_{elastic}\widehat{\beta}^{(i)}, i = 1, 2, \cdots, c \tag{26} $$

$$ {}^{u}_{elastic}Y^{(i)} = {}^{u}X^{(i)}\ {}^{u}_{elastic}\widehat{\beta}^{(i)}, i = 1, 2, \cdots, c \tag{27} $$

The outputs of each component representing the forecasts are calculated, correspondence to the related membership and non-membership values, using the following formulas.

$$ {}^{\mu}_{elastic}\widehat{Y}^{*} = \frac{\sum_{i=1}^{c}{}^{\mu}_{elastic}\widehat{Y}^{(i)}\mu_{i.}}{\sum_{i=1}^{c}\mu_{i.}} \tag{28} $$

$$ {}^{u}_{elastic}\widehat{Y}^{*} = \frac{\sum_{i=1}^{c}{}^{u}_{elastic}\widehat{Y}^{(i)}u_{i.}}{\sum_{i=1}^{c}u_{i.}} \tag{29} $$

Eventually, final forecasts for IFRFs-bENR are calculated as follows.

$$ {}_{elastic}\widehat{Y} = {}^{\mu}_{elastic}\widehat{Y}^{*}(1-\pi) + {}^{u}_{elastic}\widehat{Y}^{*}\pi \tag{30} $$

3.4. Long Short-Term memory neural network

LSTM, as proposed by Hochreiter and Schmidhuber [43], is a specialized type of recurrent neural network (RNN) capable of learning long-term dependencies. Like other RNNs are in the structure of a chain composed of modules. LSTM, a type of RNN, reduces the vanishing gradient problems seen in classical RNNs, which prevent the network from learning long-term dependencies thanks to its' forget gate, input gate, and output gate. With these gates, it has cell states which represent long-term memory. The most important component of LSTM in solving the problem of long-term dependence is the "*cell state*" information given in Fig. 1.

The LSTM has structures called gates that are tasked with adding or removing information to the cell state. The working principle of LSTM can be given step by step as provided in Supplementary Table 2.

A general structure of the LSTM is given in Supplementary Fig. 1.

4. The proposed methodology

This study introduces a new consolidated time series forecasting model combining intuitionistic fuzzy regression functions having elastic net regularization with a deep neural network, LSTM. In the proposed model, both linear and non-linear relationships are simultaneously modelled. Additionally, the model hyperparameters are determined using a GA through an optimization process on the validation set. The forecasting process performed by the proposed GA-based IFRFs-bENR&LSTM is explained step by step with Algorithm 1 given below.

Algorithm 1. ***Step* 1**. *Set GA parameters.*

- Population Size: *PopS* –50
- Selection Mechanism: *SlcM* – *Roulette*
- Elite Count: *EC* –0.20 × *PopS*
- Crossover Operator: *CO* – *Single Point*
- Crossover Probability: *CP* –0.8
- Mutation Operator: *MO* – *Uniform*
- Mutation Probability: *MP* –0.01
- Natural Selection Ratio: *NSR* –0.20 × *PopS*
- Elitism Ratio: *ER* – 0.20
- Maximum Generations: *MG* – 50

Step 2. Define possible values of the parameters for the forecasting process.

- # of Lagged Variable – LV : 2 –10 / *Integer*
- # of Intuitionistic Fuzzy Cluster – *IFCl*: 3 –10 / *Integer*
- Fuzziness Index – FI : $[1.5, 3]$ / *Continuous*
- Hesitation Degree – HD : $[0.1, 0.6]$ / *Continuous*
- Shrinkage Parameters – λ_1, λ_2 : $(0, \infty]$ / *Continuous*
- Shrinkage Parameters – α_1, α_2 : $[0, 1]$ / *Continuous*
- Dropout Probability – Dp : $[0.3, 0.7]$ / *Continuous*
- # of Hidden Layer – Hl : 24 –128 / *Integer*
- Weight of Linear Component – w : $[0, 1]$ / *Continuous*

Step 3. Split the data.

The data is divided into three sub-datasets: training, validation, and testing. The possible optimal values for the parameters are determined over the training set. For the validation subset, those that minimize the prediction error are chosen as the final optimal values. The testing subset is used to measure and evaluate the out-of-sample performance.

Step 4. Randomly generate the initial population

The initial population is generated randomly from ranges containing the possible values of the parameters.

Step 5. Run the forecasting process for each individual over the training sub-dataset.

- Step 5.1. Establish a lagged variable set (*LTS*) of the real-valued time series via the corresponding genes of the individuals ($LV = p$).

$$LTS = [Y_{t-1}, Y_{t-2}, \cdots, Y_{t-p}] \tag{31}$$

- Step 5.2. Via the corresponding genes of the individuals, for the lagged variable set, apply IFCM and obtain membership and non-membership degrees.
- Step 5.3. Set the input matrices and the target vector.

Input matrices and target vectors are created for the IFRFs-bENR and LSTM components of the consolidated forecasting model. At this stage, for the intuitionistic fuzzy regression functions, $2 \times c$ ($IFCl = c$) input matrices are generated as given in Eq. (34) and Eq. (37). While c of these contain membership values and functions, the other c consists of non-membership degrees.

$${}^{\mu}X_t^{(i)} = \left[\mu_{it}\mu_{it}^2 exp(\mu_{it}) log\left(\frac{1-\mu_{it}}{\mu_{it}}\right)\right], i = 1, 2, \cdots, c \tag{32}$$

$${}^{\mu}I_t^{(i)} = \left[{}^{\mu}X_t^{(i)} y_{t-1} y_{t-2} \cdots y_{t-p}\right], i = 1, 2, \cdots, c \tag{33}$$

$${}^{\mu}I^{(i)} = \left[{}^{\mu}I_t^{(i)}\right], t = 1, 2, \cdots, N-p \tag{34}$$

In Eq. (32), ${}^{\mu}X_t^{(i)}$, for the i th intuitionistic fuzzy set and time point t, represents the membership values and some transformations thereof. ${}^{\mu}I_t^{(i)}$ given in Eq. (33), for the membership component of the IFRFs-bENR, is the input vector again for the i th intuitionistic fuzzy set and time point t.

$${}^{u}X_t^{(i)} = \left[u_{it}u_{it}^2 exp(u_{it}) log\left(\frac{1-u_{it}}{u_{it}}\right)\right], i = 1, 2, \cdots, c \tag{35}$$

$${}^{u}I_t^{(i)} = \left[{}^{u}X_t^{(i)} y_{t-1} y_{t-2} \cdots y_{t-p}\right], i = 1, 2, \cdots, c \tag{36}$$

$$^{u}I^{(i)} = \left[^{u}I_t^{(i)} \right], t = 1, 2, \cdots, N - p \tag{37}$$

In Eq. (35), $^{u}X_t^{(i)}$, for the i th intuitionistic fuzzy set and time point t, represents the non-membership values and some transformations thereof. $^{u}I_t^{(i)}$ given in Eq. (36), for the non-membership component of the IFRFs-bENR, is the input vector again for the i th intuitionistic fuzzy set and time point t.

Moreover, for the LSTM, the input matrix is created as in Eq. (39).

$$^{LSTM}X_t = \left[\, ^{\mu}X_t^{(1)} \cdots \, ^{\mu}X_t^{(c)} \, ^{u}X_t^{(1)} \cdots \, ^{u}X_t^{(c)} y_{t-1} y_{t-2} \cdots y_{t-p} \right] \tag{38}$$

$^{LSTM}X_t$ given in Eq. (38), for the LSTM, is the input vector for the time point t.

$$^{LSTM}I = \left[\, ^{LSTM}X_t \right], t = 1, 2, \cdots, N - p \tag{39}$$

Furthermore, the target vectors are set by:

$$T = {}^{\mu}T^{(i)} = {}^{u}T^{(i)} = {}^{LSTM}T = \left[\begin{matrix} y_{p+1} & y_{p+2} & \cdots & y_N \end{matrix} \right]^T \tag{40}$$

- Step 5.4. Produce the estimation/training of the parameters for each component.

Sub-Step 1. Estimate the parameters of each IFRFs-bENR for both membership and non-membership components (*use corresponding genes of the individuals;* $\lambda_1, \lambda_2, \alpha_1, \alpha_2$).

$$^{\mu}_{Enet}\widehat{\boldsymbol{\beta}}^{(i)} = min\left(^{\mu}\beta_0^{(i)}, \, ^{\mu}\beta^{(i)} \right) \in R^{p+4} \left\{ \sum_{k=1}^{N-p} \left(y_{k+p} - \, ^{\mu}\beta_0^{(i)} - \, ^{\mu}I_k^{(i)} \, ^{\mu}\beta^{(i)} \right)^2 \right\} + \lambda_1 \left[(1 - \alpha_1) \frac{\| ^{\mu}\beta^{(i)} \|_2^2}{2} + \alpha_1 \| ^{\mu}\beta^{(i)} \|_1 \right] \tag{41}$$

$$^{u}_{Enet}\widehat{\boldsymbol{\beta}}^{(i)} = min\left(^{u}\beta_0^{(i)}, \, ^{u}\beta^{(i)} \right) \in R^{p+4} \left\{ \sum_{k=1}^{N-p} \left(y_{k+p} - \, ^{u}\beta_0^{(i)} - \, ^{u}I_k^{(i)} \, ^{u}\beta^{(i)} \right)^2 \right\} + \lambda_2 \left[(1 - \alpha_2) \frac{\| ^{u}\beta^{(i)} \|_2^2}{2} + \alpha_2 \| ^{u}\beta^{(i)} \|_1 \right] \tag{42}$$

Sub-Step 2. Train the LSTM (*use corresponding genes of the individuals; Dp, Hl*).

The training of the LSTM is performed using an iterative process. The basic structure of the LSTM with the input and outputs is given in Supplementary Fig. 2.

In this sub-step, while the output of LSTM o_t is the time series data point at t time (y_t), when the number of intuitionistic fuzzy clusters is c and the number of lagged variables is p, the inputs of LSTM, $I_{t-1}, I_{t-2}, \cdots, I_{t-p}$ can be given as in Supplementary Table 3:

- Step 5.5. Over the validation subset, calculate the forecasts for each component (*use corresponding genes of the individuals*).

Sub-Step 1. Compute the forecasts – *linear component.*
The prior forecasts are obtained using Eq. (47).

$$^{\mu}_{Enet}\widehat{Y}^{(i)} = \, ^{\mu}I_k^{(i)} \, ^{\mu}_{Enet}\widehat{\boldsymbol{\beta}}^{(i)}, i = 1, 2, \cdots, c \tag{43}$$

$$^{u}_{Enet}\widehat{Y}^{(i)} = \, ^{u}I_k^{(i)} \, ^{\mu}_{Enet}\widehat{\boldsymbol{\beta}}^{(i)}, i = 1, 2, \cdots, c \tag{44}$$

In Eqs. (43) and (44), for the i th intuitionistic fuzzy set, $^{\mu}_{Enet}\widehat{Y}^{(i)}$ and $^{u}_{Enet}\widehat{Y}^{(i)}$ are the output vectors of the IFRFs-bENR for the membership and non-membership components, respectively.

$$^{\mu}_{Enet}\widehat{y}_t = \frac{\sum_{i=1}^{c} \mu_{it} \, ^{\mu}_{Enet}\widehat{y}^{(i)}}{\sum_{i=1}^{c} \mu_{it}}, t = 1, 2, \cdots, N_{val} \tag{45}$$

$$^{u}_{Enet}\widehat{y}_t = \frac{\sum_{i=1}^{c} u_{it} \, ^{u}_{Enet}\widehat{y}^{(i)}}{\sum_{i=1}^{c} u_{it}}, t = 1, 2, \cdots, N_{val} \tag{46}$$

In Eqs. (45) and (46), for the i th intuitionistic fuzzy set and time point t in the validation set, $^{\mu}_{Enet}\widehat{y}_t$ and $^{u}_{Enet}\widehat{y}_t$ are the outputs of the IFRFs-bENR for the membership and non-membership components, respectively.

$$_{Linear}\widehat{y}_t = (1 - HD) \times \, ^{\mu}_{Enet}\widehat{y}_t + HD \times \, ^{u}_{Enet}\widehat{y}_t, t = 1, 2, \cdots, N_{val} \tag{47}$$

Sub-Step 2. Compute the forecasts – *nonlinear component.*

The forecasts for the trained LSTM ${}_{Non-Linear}\hat{y}_t$ are calculated over the validation set.

- Step 5.6. Compute the final forecasts (*use corresponding genes of the individuals; w*).

$$Final\hat{y}_t = w \times {}_{Linear}\hat{y}_t + (1-w) \times {}_{Non-Linear}\hat{y}_t, t = 1,2,\cdots,N_{val} \tag{48}$$

- Step 5.7. Compute the loss function for each individual over the validation subset.

$$^{j}MSE = mean\left({}^{j}y_t - {}^{j}_{Final}\hat{y}_t\right)^2, t = 1,2,\cdots,N_{val}; j = 1,2,\cdots,PopS \tag{49}$$

Step 6. *Natural Selection* operation

The worst $NNS = NSR \times PopS$ individuals are determined. They are removed from the population and replaced with randomly generated new individuals (PoP).

Step 7. *Elitism* operation

The bets $EN = ER \times PopS$ individuals are determined. They are transferred to the next generation (save in PoP_1).

Step 8. *Crossover* operation

By using roulette wheel selection mechanism, a mating pool is formed from the $CN = (PopS - ER \times PopS)$ individuals in PoP. By mating individuals from the mating pool, $CN/2$ crossover pairs are randomly generated. A random number is generated from $Uniform(0.1)$ for each pair ($rndCrossover$). If $rndCrossover \leq CP$ then crossover operation is performed (save in PoP_2).

Step 9. *Mutation* operation

A random number is generated from $Uniform(0.1)$ for each pair ($rndMutation$). If $rndMutation \leq MP$ then mutation operation is performed (save in PoP_2).

Step 10. *Updating* operation

$$PoP = PoP_1 + PoP_2 \tag{50}$$

Step 11. Stopping criteria for optimization process is checked

If $^{j}MSE \leq \varepsilon (j = 1,2,\cdots,PopS)$ or Maximum Generations MG is reached, then stop the process. The best individual in the last iteration indicates the final parameters. Otherwise, the optimization process continues between *Step 5* and *Step 11*. The pseudo-code for Algorithm 1 is provided by Supplementary Table 4. A graphical abstract of this methodology can be given by Supplementary Fig. 3.

5. Implementation scenarios and evaluation

5.1. Data organization

The forecasting capacity and performance of the proposed GA-based IFRFs-bENR&LSTM have been examined with forty-eight financial time series applications. The datasets and their respective implementation characteristics are summarized in Supplementary Table 5. The datasets were split into three subsets: training, validation, and testing. The training sets have been used to train the forecaster model. Validation sets have been used to specify the optimal hyperparameters of the model by GA in an optimization process. The performance evaluation has been discussed on test sets. The size of the validation and test sets were chosen equally to each other in all implementations. In total, analyses were conducted for 48 different data sets.

5.2. Performance metrics

As in many studies, the comparative evaluation of performances in this study has been examined and discussed over two metrics; the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE) given in Eqs. (51) and (52) which were used as performance measures.

$$RMSE = \sqrt{mean\left((Actual_t - Predicted_t)^2\right)}; t = 1,2,\cdots,T \tag{51}$$

$$MAPE = mean\left(\left|\frac{Actual_t - Predicted_t}{Actual_t}\right| \times 100\%\right); t = 1,2,\cdots,T \tag{52}$$

As with most studies, the median relative absolute error (MdRAE) was used to compare the performance of the proposed model with a reference or benchmark model, the naïve model. MdRAE is calculated as in Eq. (54) through the relative errors given by Eq. (53).

$$r_t = \frac{Target_t - Forecasted_t}{Target_t - Forecasted_t^*} \tag{53}$$

$$MdRAE = median|r_t|; t = 1, 2, \cdots, T \tag{54}$$

Here, $Forecasted_t$ and $Forecasted_t^*$ are the forecasted values produced by the proposed and reference models, respectively. $Forecasted_t^*$ is equal to $Target_{t-1}$ for the naïve model which was chosen as reference model as in most of the studies in the literature.

Two key parameters (β and R^2) of a regression model given by Eq. (55) were used as other metrics to measure the performance of the proposed model. Here, performance success is measured by the closeness of these parameters to 1.

$$Y_t = \beta \widehat{Y}_t + \varepsilon_t \tag{55}$$

Apart from all these metrics, different graphics that present the harmony between the forecast and the target values are also used as visual metrics.

In addition to these metrics that support and prove the validity of the forecasts produced by the proposed model, the reliability/consistency of the model was also examined. For this purpose, the proposed model with the best hyperparameter values was run 50 times. The scattering of MAPE and MdRAE metrics was examined to discuss reliability/consistency.

6. Results and discussions

The outstanding forecasting performance of the GA-based IFRFs-bENR&LSTM has been proven from various perspectives, as mentioned above.

The prediction performance of the proposed ensemble model has been demonstrated from various perspectives. First, in terms of RMSE and MAPE metrics, the proposed model is compared with current methods. This comparison has been conducted for all data sets, and the superiority of the proposed model has been demonstrated.

6.1. TAIEX data sets

In general, TAIEX data sets for two different periods have been analysed in the literature so far. These are the daily observed 5-year datasets in 2000–2004 and the 11-year observed daily observed datasets in 2008–2018.

6.1.1. The implementations of the 5-years TAIEX dataset (2000–2004)

For the test sets, the RMSE values of the results obtained from the proposed method and various existing methods are given in Table 2.

Table 2 clearly presents the following results:

- The proposed IFRFs-bENR&LSTM has produced the most accurate forecasts, for the remaining 4 years, excluding 2001. In 2001, it significantly outperformed the second-best model among others and produced results that were competitive and very close to the best one.
- The best forecasts for the year 2000, among others, were produced with RMSE values of 32 ([38]), while the proposed method reduced the RMSE value to 18. This represents a remarkable 44 % enhancement compared to the best available method. This progress demonstrates an even more remarkable improvement, reaching a rate of 83 % compared to the second-best method (RMSE = 105, [45]).

Table 2
The performances and provided progress for the TAIEX 2000–2004 data sets.

Models		Time Series / TAIEX Data Sets					RMSE's	
		2000	2001	2002	2003	2004	Average	Median
[1]		293	116	76	77	82	129	82
[9]		225	116	76	77	82	115	82
[10]		131	130	80	58	67	93	80
[8]		126	113	63	51	54	81	63
[47]		128	106	65	52	54	81	65
[48]		120	113	63	49	52	79	63
[46]		209	73	22	43	54	80	54
[5]		122	110	54	51	50	77	54
[45]		105	110	60	51	50	75	60
LSTM from [45]		136	101	89	92	70	108	92
[42]		119	104	64	51	52	81	64
[38]		32	11	8	16	6	15	11
The Proposed Model		**18**	**14**	**5**	**5**	**5**	**9**	**5**
Progress (%)	**Comparison with the best of others**	**44**	**Nan**	**38**	**69**	**17**	**40**	**55**
	Comparison with the second best of others	**83**	**81**	**77**	**88**	**90**	**88**	**91**

Note: Listed and some other unlisted in the table results of the available counterparts models are available in [38].

- For 2002 again, forecast accuracy was glamorously improved by 38 % and 77 % over the best (RMSE = 8, [38]) and second-best (RMSE = 22, [46]) methods, respectively.
- In 2003, the forecasting performance of the proposed model significantly outperformed both the best (RMSE = 16, [38]) and second-best (RMSE = 43, [20]) methods, showcasing impressive improvements of 69 % and 88 %, respectively.
- In 2004, the GA-based IFRFs-bENR&LSTM model exhibited a noteworthy improvement when compared to the best method (RMSE = 6, [38]) and the second-best methods (RMSE = 50, [5,45]) achieving progress rates of 17 % and 90 %, respectively.
- Moreover, in terms of mean and median of 5-year RMSE values, the proposed model showed significant improvement over the best of others by 40 % and 55 %.

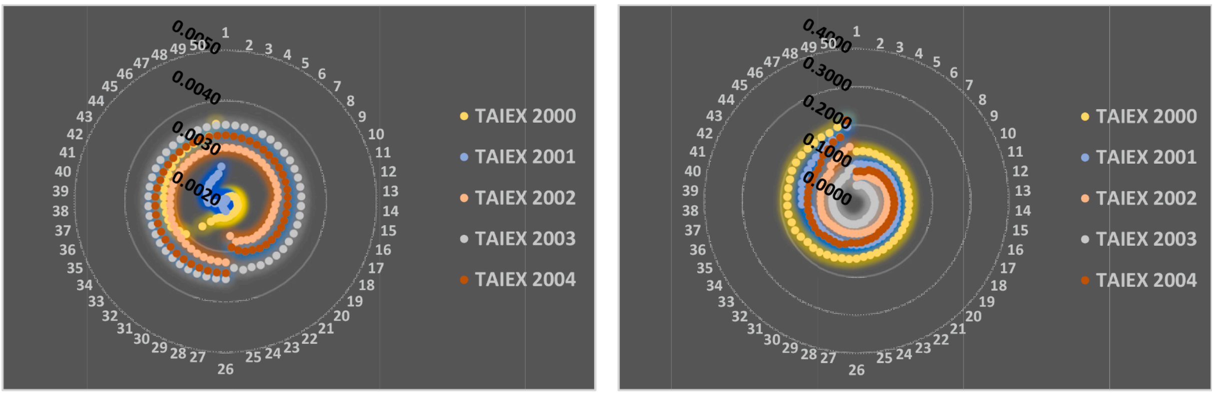
Another aspect of performance evaluation involves the analysis of two coefficients within a linear regression model, which assess the alignment between forecasts and actual observations. In this context, estimations of the established regression models for forecasts and actual observations and their two basic parameters (β and R^2) were given in Supplementary Table 6.

It can be observed that the regression coefficients, denoted as $\hat{\beta}$, for the models were estimated to be very close to 1, aligning with our expectations. Moreover, the fact that the confidence intervals for β also include 1 supports the results. Furthermore, the coefficients of determination, R^2, were also obtained very close to 1. All the findings presented in Supplementary Table 5 illustrate a strong alignment between the forecasts generated by the proposed model and the actual observations, affirming the success of the proposed method.

The demonstrated level of progress and the regression coefficients, which are obtained very close to 1, serve as statistical indicators of the superior predictive performance of the proposed model. In addition, this superior performance is also demonstrated visually by the harmony of the distributions of the forecasted and observed data points presented in Supplementary Fig. 4. Violin graphs in Supplementary Fig. 4, unlike box charts, not only present information about some statistics and distributions of data points but also about density. When these graphs are evaluated in this respect, it is seen that the median, interquartile ranges, and lower–upper adjacent statistics are almost the same for all data sets. Also, the densities for the observed and forecasted data points are almost the same. From all of these, for all-time series, as another indicator of the superior performance of the proposed model, it is clearly observed that the observed and the forecasted data points have a very high harmony with each other.

The proposed GA-based IFRs-bENT-LSTM was run 50 times with different random initial conditions and hyperparameter values determined by GA. This action was carried out for two aims. The first is the creation of an infrastructure that will allow a statistical evaluation of the forecasts for each data point and error metrics. The second objective is to assess the reliability and validity of the proposed model. Supplementary Table 7 presents descriptive statistics for MAPE and MdRAE metrics obtained from 50 runs for the TAIEX dataset. Supplementary Table 7 reveals that even at its maximum, the MAPE criterion yields relative errors of only 0.2 % to 0.3 %, highlighting the exceptional validity of the proposed GA-based IFRs-wENT-LSTM model. Furthermore, the remarkably low standard deviations of the MAPE values, ranging from 1 to 5 in ten thousand, across 50 runs, signify a high degree of reliability of the proposed forecasting tool. Beyond that, the exceptionally narrow confidence interval (CI) bounds also strongly support for the validity and reliability of the produced results. All similar interpretations can be made based on the results obtained for the distribution of the MdRAE metric. The proposed GA-based IFRs-bENT-LSTM produced approximately five times better forecasting results than the reference method for TAIEX2000, TAIEX2001, and TAIEX2004 data sets. Also, these values were roughly 6.5 times for TAIEX2002 and approximately 12 times for TAIEX2003.

Moreover, confidence intervals and desired probabilities for each data point, in addition to errors, can be produced using the experimental distribution obtained by 50 runs. Indeed, while such gains can often be achieved by many ML or AI-based forecasters, the noteworthy aspect here is the consistent and reliable performance at such a high level. Based on this, some probabilities were calculated as given below.



(a) MAPE values **(b)** MdRAE values

Fig. 4. The distribution of error criteria values from 50 runs for TAIEX 2000–2004.

Table 3
The comparative forecasting results in terms of RMSE for the TAIEX 2008–2018.

Models		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Avrg.
[2]		225.38	204.11	151.42	237.43	81.08	82.3	98.64	176.6	122.71	91.69	173.36	149.52
[9]		186.48	207.42	211.21	215.13	78.37	184.83	308.72	100.75	316.57	902.18	492.89	291.32
[10]		129.48	69.91	67.11	123.03	58.37	50.39	69.15	78.22	80.51	64.25	103.74	81.29
[49]		137.44	71.76	74.52	117.38	61.59	51.37	69.23	79.89	83.53	62.37	103.24	82.94
[50][1]		101.91	105.64	62.93	123.79	82.30	65.62	46.38	102.82	81.94	61.83	101.56	85.16
[50][2]		91.95	57.43	44.61	90.32	43.58	33.83	46.23	58.79	53.15	36.99	70.67	57.05
[38]		26.13	26.41	23.47	40.32	15.19	10.78	20.52	14.58	25.77	20.00	27.55	22.79
The Proposed Model		**12.30**	**11.42**	**7.03**	**23.08**	**13.37**	**3.9127**	**10.20**	**8.56**	**19.74**	**11.04**	**13.54**	**12.20**
Progress (%)	**with the best of others**	**53**	**57**	**70**	**43**	**12**	**64**	**50**	**41**	**23**	**45**	**51**	**46**
	with the second best of others	**87**	**80**	**84**	**74**	**69**	**88**	**78**	**85**	**63**	**70**	**81**	**79**

$$P(0.002466 \leq meanofMAPE_{TAIEX2000} \leq 0.002752) = 0.95$$

$$P(0.002297 \leq MAPE_{TAIEX2000} \leq 0.003300) = 0.80$$

$$P(5420.0298 \leq FORECAST_{TAIEX2000-DataPoint1} \leq 5428.1948) = 0.90$$

The scatter plots of error points, as depicted by the radar graphs in Fig. 4 for both criteria, visually attest to the reliability of the proposed prediction model.

From Fig. 4 (a), the MAPE values for forecasted data points of all TAIEX data sets have been observed in a narrow range. These demonstrations are indeed evidence of the reliability of the model. Even in the TAIEX2000 dataset, where the worst results were produced, MAPE values were obtained between 0.20 % and 0.35 %. In addition, MdRAE values for the same data set were obtained in the range of 0.13 and 0.21. This means that, even in the worst case, the recommended model performance is five times better than the reference model.

6.1.2. The implementations of the 11-years TAIEX dataset (2008–2018)

Secondly, the performance of the proposed GA-based IFRs-bENT-LSTM was investigated over eleven time series consisting of TAIEX datasets daily recorded between 2008 and 2018. The data sets recorded in each financial year were considered individual time series in the analysis and evaluation process. The forecasting performance of the proposed GA-based IFRs-bENT-LSTM has been presented comparatively in terms of RMSE, MAPE, and MdRAE error measures in Tables 3 and 4 and also Supplementary Table 8.

It can be reached the following results from Table 3:

- The proposed GA-based IFRs-bENT-LSTM displayed the best forecasting performance for all eleven data sets.
- Considering the best method among the others for each of the eleven datasets;
  - The rate of progress in forecast accuracy has been over 40 % for 3 data sets, TAIEX2011, TAIEX2015, and TAIEX2017.
  - The progress rates for the TAIEX 2008, TAIEX 2009, TAIEX 2014, and TAIEX 2018 data sets have exceeded 50 % and even reached 60 % and 70 % for TAIEX 2010 and TAIEX 2013, respectively.
  - For the remaining two data sets, TAEIX2012 and TAIEX216, significant progress of 12 % and 23 %, respectively, have been achieved.
  - The average improvement across the 11 datasets was a remarkable 46 %.
- Considering the second-best method among the others for each of the eleven datasets;
  - With the exception of 4 data sets, the progress rates in forecasting accuracy have reached extraordinary levels with values around or above 80 %.
  - Even for the remaining four datasets, these rates indicated significant progress, with values reaching 60 % and 70 %.
  - The average improvement over the second-best performers across the 11 datasets has also been outstanding, nearly reaching 80 %.

From Table 4, the following findings can be summarized:

- The proposed GA-based IFRFs-bENR&LSTM, in terms of MAPE, presented the best forecasting performance for all data sets.
- Considering the best method among the others for each of the eleven datasets;
  - The forecasts were obtained with average proportional errors significantly below 1 % for all datasets and even below 0.1 % for 8 out of 11 datasets. All these values affirm the exceptional forecasting capabilities of the proposed model.
  - Hence, the proposed model achieved a significant increase in forecasting accuracy, ranging from 50 % to 60 %, demonstrating superior performance across all but one of the 11 datasets.

Table 4
The comparative forecasting results in terms of MAPE for the TAIEX 2008–2018.

Models		2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Avrg.
[2]		3.87	2.04	1.53	2.65	0.89	0.78	0.86	1.78	1.10	0.71	1.29	1.59
[9]		3.49	2.41	1.90	2.61	0.86	1.77	3.31	0.97	3.20	8.45	4.73	3.06
[10]		2.34	0.66	0.64	1.35	0.59	0.50	0.61	0.77	0.63	0.47	0.81	0.85
[49]		2.43	0.68	0.72	1.23	0.59	0.52	0.61	0.80	0.63	0.46	0.81	0.86
[50][1]		1.87	1.01	0.60	1.42	0.91	0.62	0.39	0.98	0.78	0.50	0.92	0.91
[50][2]		1.57	0.55	0.43	0.94	0.42	0.33	0.44	0.59	0.43	0.27	0.53	0.59
[38]		0.37	0.24	0.24	0.45	0.16	0.10	0.18	0.14	0.22	0.15	0.22	0.22
The Proposed Model		**0.23**	**0.09**	**0.07**	**0.16**	**0.08**	**0.04**	**0.07**	**0.08**	**0.10**	**0.06**	**0.09**	**0.10**
Progress (%)	**with the best of others**	**38**	**63**	**71**	**64**	**50**	**60**	**61**	**43**	**55**	**60**	**59**	**55**
	with the second best of others	**85**	**84**	**84**	**83**	**81**	**88**	**82**	**86**	**77**	**78**	**83**	**83**

- For the only remaining time series (TAIEX2008), it exhibited a highly notable performance boost, achieving a substantial improvement of 38 %
- Once again, the average improvement in MAPE was exceptional, reaching a value of 46 %.
- Considering the second-best method among the others;
  - The progress in forecasting accuracy was again distinguished, with around 80 % and above.
  - Once more, the average improvement over the second-best performers across all datasets has been exceptional, exceeding 80 %.

Similar findings can be gleaned from Supplementary Table 9 as in Tables 3 and 4. Therefore, a detailed analysis is left to the reader.

Moreover, as another view of performance evaluation, the estimates of the regression models were created for the forecasts and actual observations. The reached findings, including R^2, were summarized in Supplementary Table 9. Regression coefficients,$\widehat{\beta}$, were estimated to be quite close to 1, as expected. Moreover, except one, the fact that the confidence intervals for all β also include 1 supports these results. It can be seen that both the lower and upper bounds of the confidence interval, which does not include the value 1, are also very close to 1. Furthermore, the coefficients of determination, R^2, were also obtained very close to 1. All findings in Supplementary Table 9 confirm a strong agreement between forecasts and actual observation.

6.2. The implementations of the 11-years DJI and SSEC dataset (2008–2018)

Third, the 11-year DOJ and SEC financial time series observed daily from 2008 to 2018 were analyzed separately for each fiscal year. The results obtained for 11 years were summarized with the mean statistics of error metrics and presented in Supplementary Table 10 for comparison. The results, as summarized in Supplementary Table 10, demonstrate that the proposed model exhibits the highest forecasting performance for both time series.

For DII and SSEC datasets, when compared to the best-performing method among others, the proposed model exhibited an average improvement in forecasting performance of 17 % and 27 %, respectively, in terms of the RMSE (34.74; 7.81) criterion. The improvement rates in forecasting accuracy, assessed using the MAPE (Mean Absolute Percentage Error) criteria, were approximately 45 % (0.22 % for DII and 0.30 % for SSEC) for both datasets. Meanwhile, for the MdRAE (Median Relative Absolute Error) criteria, the improvement rates were 73 % for DII and 58 % for SSEC (with values of 0.1134 and 0.1302, respectively). In comparison to the second-best method among the alternatives, the improvement in forecasting performance was approximately 70 % in terms of both RMSE and MAPE criteria for both datasets, while it exceeded 80 % when assessed using the MdRAE criteria.

6.3. The implementations of the 5-years IstEX dataset (2009–2013)

Finally, the proposed model has been applied to the IstEX datasets daily recorded for 5 years of 2009–2013, and the obtained findings have been discussed in detail. To offer a comparative perspective in the discussion of the findings, we have also summarized the results of several state-of-the-art models in Table 6.

ARIMA : Autoregressive integrated moving average method
ES : Exponential Smoothing.
MLP : Multilayer perceptron artificial neural network
SC : Song and Chissom's fuzzy time series prediction model [1]
ANFIS : Adaptive Neuro-Fuzzy Inference System.

Table 6
RMSE metric values for IstEX data sets.

		TIME SEIES / IstEX DATA SETS / # of Testing Set										RMSE's Average
		2009		2010		2011		2012		2013		
Prediction Model		7	15	7	15	7	15	7	15	7	15	
ARIMA		345	540	1221	1612	1058	1130	651	621	1362	1269	981
ES		345	540	1208	1612	1057	1130	651	621	1362	1269	980
MLP		325	525	1077	1603	920	1096	775	783	1315	1233	965
SC		1402	1754	1128	1742	1396	1360	1292	1047	1450	1931	1450
ANFIS		405	647	1141	2033	1007	1134	634	938	1447	1413	1080
AR-ANFIS		240	467	1136	1451	987	999	631	619	1362	1256	915
I-TSFIS		166	1046	250	251	817	384	277	228	451	1106	498
I-FRF		240	507	963	1390	658	994	296	530	690	1172	744
FF-T1-RG		319	495	1080	1575	915	1028	720	676	1251	1237	930
FF-T1-En		240	500	1045	1300	945	1000	662	762	832	1207	849
CIPM		80	309	56	89	66	89	72	128	213	397	150
The Proposed		**22**	**133**	**28**	**61**	**59**	**74**	**53**	**61**	**131**	**191**	**81**
Progress (%)	**with the best of others**	**73**	**57**	**50**	**31**	**11**	**17**	**26**	**52**	**38**	**52**	**46**
	with the second best of others	**87**	**72**	**89**	**76**	**91**	**81**	**81**	**73**	**71**	**83**	**84**

I-TSFIS : Intuitionistic time series fuzzy inference system [46]
I-FRF : Intuitionistic fuzzy regression functions approach [6]
FF-T1-RG : Type 1 fuzzy Regression function approach based on Ridge Regression [48]
FF-T1-En : Type 1 fuzzy Regression function approach based on Elastic Net [42]
CIPM : Cascaded intuitionistic prediction model [38]

From Table 6, the superior forecasting performance of the proposed model has been put forth once more. Considering the best ones among the others, the prediction performance progress ensured by the proposed model exceeded 50 % for four applications (IstEX 2009/15 = 57 %; IstEX 2010/7 = 50 %; IstEX 2012/15 = 52 %; IstEX 2013/15 = 52 %) and even 70 % for one application (IstEX 2009/7 = 73 %). In addition, while these progress rates were around 30 % or more in three applications (IstEX 2010/15 = 31 %; IstEX 2012/7 = 26 %; IstEX 2013/7 = 38 %), they were above 10 % in the other two applications (IstEX 2011/7 = 11 %, IstEX 2011/15 = 17 %). The average level of progress in 10 applications was also around 50 %. Considering the second-best ones among others, the proposed model outperformed them in seven out of ten datasets, achieving 70–80 % higher forecast accuracy based on the RMSE criterion. The progress levels demonstrated by the proposed model reached and exceeded around 90 % in three of the ten datasets (IstEX 2009/7 = 87 %; IstEX 2010/7 = 89 %; IstEX 2012/7 = 91 %). The average progress level for ten datasets was also approximately 80 %.

Fig. 5 displays a radar chart illustrating the RMSE values for the ten applications of the top 5 methods, based on their average RMSE performance. When Fig. 5 is examined, it is seen that the dark blue line representing the forecasting errors (RMSE) of the proposed model is located in the innermost part of the radar chart compared to other lines. For example, while the model proposed for IstEX 2009/7 produced the best forecasts with an RMSE value of 22, the second performance was generated by CIPM with an RMSE value of 80. For another example, the model proposed for IstEX 2013/15 produced again the best predictions, with an RMSE value of 191. These results indicate that the proposed model produces forecasts with the lowest RMSE values in all 10 applications. In conclusion, from Fig. 5, it is clearly seen that the proposed model, which produced forecasts with lower RMSE values, was distinguished as a superior model compared to other models in this respect.

7. Conclusions

This study introduces a time series forecasting model that combines (consolidates) intuitionistic fuzzy regression functions with elastic net regularization and a deep neural network, LSTM. GA allows the hyperparameters of the proposed consolidated forecasting model to be determined by an optimization process, unlike almost all studies in the literature which rely on trial and error, unlike almost all studies in the literature. Furthermore, in contrast to existing models, our approach combines IFRFs and LSTM to effectively model the coexisting linear and nonlinear structures between inputs and outputs.

Besides the exceptional qualities, the ability of the proposed model to forecast time series, which is widely used in the literature, was also noteworthy. In the 5-year TAIEX dataset (2000–2004) applications, the proposed model produced an average performance progress of 40 % compared to the best among others and nearly 90 % compared to the second best. In the other TAIEX dataset (2008–2016) implementations, the progress levels were at 46 % compared to the best of the others and almost 80 % compared to the

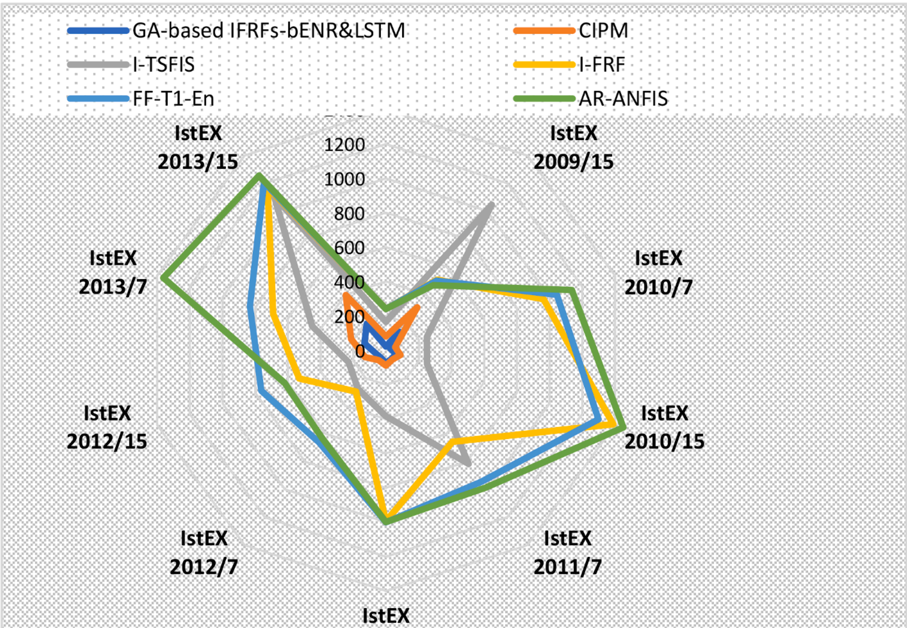


Fig. 5. Top five models on average and their RMSE values for IstEX data sets.

second best. For the 11-year DJI and SSEC time series, the progression levels produced by the proposed model were 17 % and 27 %, respectively, compared with the best of the others. These advancements reached nearly 70 % when compared to the second-best among the alternatives. Finally, the average progress of 10 applications for IstEX datasets was 46 % compared to the best and 84 % compared to the second-best of all.

8. Limitations and Discussions

The proposed GA-based IFRFs-bENR&LSTM has certain advantages and limitations. The proposed model does not consider the different uncertainty components of the data, such as indeterminacy and falsity. This may limit the determination of uncertainty in data sets and so performance. Secondly, although models formed by multiple inputs are used in modelling time series, these inputs are determined based on the historical values of the time series itself. In this context, another limitation of the proposed model is that the contributions of other time series interrelated to the focused time series are not used in the modelling process.

Apart from these few limitations, the main distinguishing and superior features of the proposed model compared to existing models can be summarized as follows:

The proposed GA-based IFRFs-bENR&LSTM

- considers memberships and degrees of hesitation together in the forecasting process and thus provides a more effective approach to handling uncertainty (*Approach to Uncertainty*).
- has the ability to model linear and non-linear relationships together and thus be adapted to the solution surfaces of the data at a high level (*Relationship Structure*).
- prevents inconsistency due to the collinearity problem in its linear component by realizing elastic net regularization first introduced in the literature in IFRFs (*Model Assumptions – for Linear Component*).
- is the first forecasting model in the literature using a deep neural network – LSTM- to determine relationships, whose input consists of all information consisting of memberships, non-memberships (as well as some transformations thereof), and real lagged time series (*Learning Strategy – for Non-Linear Component*).
- by running 50 times, generates an experimental probability distribution for each forecast and error (*Interpretable Forecasts*).
- by using an automatic mechanism, performs the selection of all hyperparameters in an optimization process over a validation set (*Tuning of Hyperparameters and Model Selection*).

In future studies, it may be worthwhile to explore the utilization of various soft sets, such as picture and neutrosophic fuzzy sets, to offer diverse perspectives on addressing uncertainty. In addition, a multivariate model can be created by using other time series interrelated to the focused time series in the modelling process. Furthermore, it's possible to explore the impact of employing different deep neural network architectures and deep learning strategies on forecasting accuracy.

CRediT authorship contribution statement

Ozge Cagcag Yolcu: Writing – review & editing, Visualization, Validation, Software, Data curation, Conceptualization. **Ufuk Yolcu:** Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ins.2024.121007.

References

[1] Q. Song, B.S. Chissom, Fuzzy time series and its models, Fuzzy Set. Syst. 54 (1993), https://doi.org/10.1016/0165-0114(93)90372-O.
[2] Q. Song, B.S. Chissom, Forecasting enrollments with fuzzy time series - part II, Fuzzy Set. Syst. 62 (1994), https://doi.org/10.1016/0165-0114(94)90067-1.
[3] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965), https://doi.org/10.1016/S0019-9958(65)90241-X.
[4] K.T. Atanassov, Intuitionistic fuzzy sets, Fuzzy Set. Syst. 20 (1986), https://doi.org/10.1016/S0165-0114(86)80034-3.

[5] O. Cagcag Yolcu, E. Bas, E. Egrioglu, U. Yolcu, A new intuitionistic fuzzy functions approach based on hesitation margin for time-series prediction, Soft. Comput. 24 (2020), https://doi.org/10.1007/s00500-019-04432-2.
[6] B. Kizilaslan, E. Egrioglu, A.A. Evren, Intuitionistic fuzzy ridge regression functions, Commun. Stat. Simul. Comput. 49 (2020), https://doi.org/10.1080/03610918.2019.1626887.
[7] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, 1981. doi: 10.1007/978-1-4757-0450-1.
[8] S.H. Cheng, S.M. Chen, W.S. Jian, Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures, Inf Sci (n Y) 327 (2016), https://doi.org/10.1016/j.ins.2015.08.024.
[9] S.M. Chen, Forecasting enrollments based on fuzzy time series, Fuzzy Set. Syst. 81 (1996), https://doi.org/10.1016/0165-0114(95)00220-0.
[10] T.H.K. Yu, K.H. Huarng, A bivariate fuzzy time series model to forecast the TAIEX, Expert Syst. Appl. 34 (2008), https://doi.org/10.1016/j.eswa.2007.05.016.
[11] C.H. Aladag, M.A. Basaran, E. Egrioglu, U. Yolcu, V.R. Uslu, Forecasting in high order fuzzy times series by using neural networks to define fuzzy relations, Expert Syst. Appl. 36 (2009), https://doi.org/10.1016/j.eswa.2008.04.001.
[12] S. Kumar, S.S. Gangwar, Intuitionistic Fuzzy Time Series: An Approach for Handling Nondeterminism in Time Series Forecasting, IEEE Trans. Fuzzy Syst. 24 (2016), https://doi.org/10.1109/TFUZZ.2015.2507582.
[13] D. Hu, L. Zan, X. Chen, W. Jie, Prediction of satellite clock errors based on deterministic intuitionistic fuzzy time series, in: International Conference on Signal Processing Proceedings, ICSP, 2016. doi: 10.1109/ICSP.2016.7877981.
[14] K. Bisht, S. Kumar, Fuzzy time series forecasting method based on hesitant fuzzy sets, Expert Syst. Appl. 64 (2016), https://doi.org/10.1016/j.eswa.2016.07.044.
[15] X. Fan, Y. Lei, Y. Wang, Y. Lu, Long-term intuitionistic fuzzy time series forecasting model based on vector quantisation and curve similarity measure, IET Signal Proc. 10 (2016), https://doi.org/10.1049/iet-spr.2015.0496.
[16] K.Q. Zheng, Y.J. Lei, R. Wang, Y.Q. Xing, Method of long-term IFTS forecasting based on parameter adaptation, Xi Tong Gong Cheng Yu Dian Zi Ji Shu/systems, Eng. Electron. 36 (2014), https://doi.org/10.3969/j.issn.1001-506X.2014.01.16.
[17] Y.N. Wang, Y.J. Lei, Y. Lei, X.S. Fan, High order intuitionistic fuzzy time series forecasting model, Tongxin Xuebao/J. Commun. 37 (2016), https://doi.org/10.11959/j.issn.1000-436x.2016099.
[18] Y. Wang, Y. Lei, Y. Lei, X. Fan, Multi-factor high-order intuitionistic fuzzy time series forecasting model, J. Syst. Eng. Electron. 27 (2016), https://doi.org/10.21629/JSEE.2016.05.13.
[19] X. Fan, Y. Lei, Y. Wang, Adaptive partition intuitionistic fuzzy time series forecasting model, J. Syst. Eng. Electron. 28 (2017), https://doi.org/10.21629/JSEE.2017.03.18.
[20] E. Egrioglu, U. Yolcu, E. Bas, Intuitionistic high-order fuzzy time series forecasting method based on pi-sigma artificial neural networks trained by artificial bee colony, Granular Computing 4 (2019), https://doi.org/10.1007/s41066-018-00143-5.
[21] P. Hajek, W. Froelich, O. Prochazka, Intuitionistic fuzzy grey cognitive maps for forecasting interval-valued time series, Neurocomputing 400 (2020), https://doi.org/10.1016/j.neucom.2020.03.013.
[22] C. Kocak, E. Egrioglu, E. Bas, A new deep intuitionistic fuzzy time series forecasting method based on long short-term memory, J. Supercomput. 77 (2021), https://doi.org/10.1007/s11227-020-03503-8.
[23] X. Fan, Y. Wang, M. Zhang, Network traffic forecasting model based on long-term intuitionistic fuzzy time series, Inf Sci (n Y) 506 (2020), https://doi.org/10.1016/j.ins.2019.08.023.
[24] R.M. Pattanayak, H.S. Behera, S. Panigrahi, A novel probabilistic intuitionistic fuzzy set based model for high order fuzzy time series forecasting, Eng. Appl. Artif. Intel. 99 (2021), https://doi.org/10.1016/j.engappai.2020.104136.
[25] S. Pant, S. Kumar, IFS and SODA based computational method for fuzzy time series forecasting, Expert Syst. Appl. 209 (2022), https://doi.org/10.1016/j.eswa.2022.118213.
[26] E. Bas, E. Egrioglu, E. Kolemen, A novel intuitionistic fuzzy time series method based on bootstrapped combined pi-sigma artificial neural network, Eng. Appl. Artif. Intel. 114 (2022), https://doi.org/10.1016/j.engappai.2022.105030.
[27] O.C. Yolcu, E. Egrioglu, E. Bas, U. Yolcu, Multivariate intuitionistic fuzzy inference system for stock market prediction: The cases of Istanbul and Taiwan, Appl. Soft Comput. 116 (2022), https://doi.org/10.1016/j.asoc.2021.108363.
[28] W. Wang, W. Lin, Y. Wen, X. Lai, P. Peng, Y. Zhang, K. Li, An interpretable intuitionistic fuzzy inference model for stock prediction, Expert Syst. Appl. 213 (2023), https://doi.org/10.1016/j.eswa.2022.118908.
[29] C. Kocak, E. Egrioglu, E. Bas, A new explainable robust high-order intuitionistic fuzzy time-series method, Soft. Comput. 27 (2023), https://doi.org/10.1007/s00500-021-06079-4.
[30] A. Dixit, S. Jain, Intuitionistic fuzzy time series forecasting method for non-stationary time series data with suitable number of clusters and different window size for fuzzy rule generation, Inf. Sci. (N Y) 623 (2023), https://doi.org/10.1016/j.ins.2022.12.015.
[31] E. Egrioglu, E. Bas, Robust intuitionistic fuzzy regression functions approaches, Inf Sci (N. Y) 638 (2023), https://doi.org/10.1016/j.ins.2023.118992.
[32] S. Çakır, Renewable energy generation forecasting in Turkey via intuitionistic fuzzy time series approach, Renew. Energy 214 (2023), https://doi.org/10.1016/j.renene.2023.05.132.
[33] R. Mohan Pattanayak, H. Sekhar Behera, S. Panigrahi, A novel high order hesitant fuzzy time series forecasting by using mean aggregated membership value with support vector machine, Inf. Sci. N Y (2023) 626, https://doi.org/10.1016/j.ins.2023.01.075.
[34] S.P. Wan, T. Rao, J.Y. Dong, Time-series based multi-criteria large-scale group decision making with intuitionistic fuzzy information and application to multi-period battery supplier selection, Expert Syst. Appl. 232 (2023), https://doi.org/10.1016/j.eswa.2023.120749.
[35] A.K. Dwivedi, U. Kaliyaperumal Subramanian, J. Kuruvilla, A. Thomas, D. Shanthi, A. Haldorai, Time-series data prediction problem analysis through multilayered intuitionistic fuzzy sets, Soft. Comput. 27 (2023), https://doi.org/10.1007/s00500-022-07053-4.
[36] M. Pant, K. Bisht, S. Negi, Computational-based partitioning and Strong α, β-cut based novel method for intuitionistic fuzzy time series forecasting, Appl. Soft Comput. 142 (2023), https://doi.org/10.1016/j.asoc.2023.110336.
[37] S. Ashraf, M. Sohail, M.S. Chohan, S. Paokanta, C. Park, Higher-order circular intuitionistic fuzzy time series forecasting methodology: Application of stock change index, Demonstratio Math. 57 (2024), https://doi.org/10.1515/dema-2023-0115.
[38] O. Cagcag Yolcu, U. Yolcu, A novel intuitionistic fuzzy time series prediction model with cascaded structure for financial time series, Expert Syst. Appl. 215 (2023), https://doi.org/10.1016/j.eswa.2022.119336.
[39] A. Yalçınkaya, U. Yolcu, B. Şenoğlu, Maximum likelihood and maximum product of spacings estimations for the parameters of skew-normal distribution under doubly type II censoring using genetic algorithm, Expert Syst. Appl. 168 (2021) 114407, https://doi.org/10.1016/J.ESWA.2020.114407.
[40] T. Chaira, A novel intuitionistic fuzzy C means clustering algorithm and its application to medical images, Appl. Soft Comput. J. (2011), https://doi.org/10.1016/j.asoc.2010.05.005.
[41] E. Bas, U. Yolcu, E. Egrioglu, Intuitionistic fuzzy time series functions approach for time series forecasting, Granul. Comput. 6 (2021), https://doi.org/10.1007/s41066-020-00220-8.
[42] N. Tak, D. İnan, Type-1 fuzzy forecasting functions with elastic net regularization, Expert Syst. Appl. 199 (2022), https://doi.org/10.1016/j.eswa.2022.116916.
[43] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, Neural Comput. 9 (1997) 1735–1780.
[44] C. Olah, Understanding LSTM Networks [Blog], (2015). http://colah.github.io/posts/2015-08-Understanding-LSTMs/ (accessed February 21, 2023) (accessed January 1, 2024).
[45] M. Kirisci, O. Cagcag Yolcu, A New CNN-Based Model for Financial Time Series: TAIEX and FTSE Stocks Forecasting, Neural Process. Lett. 54 (2022), https://doi.org/10.1007/s11063-022-10767-z.
[46] E. Egrioglu, E. Bas, O.C. Yolcu, U. Yolcu, Intuitionistic time series fuzzy inference system, Eng. Appl. Artif. Intel. 82 (2019), https://doi.org/10.1016/j.engappai.2019.03.024.

[47] N. Tak, A.A. Evren, M. Tez, E. Egrioglu, Recurrent type-1 fuzzy functions approach for time series forecasting, Appl. Intell. 48 (2018), https://doi.org/10.1007/s10489-017-0962-8.

[48] E. Bas, E. Egrioglu, U. Yolcu, C. Grosan, Type 1 fuzzy function approach based on ridge regression for forecasting, Granular Computing 4 (2019), https://doi.org/10.1007/s41066-018-0115-4.

[49] O. Cagcag Yolcu, U. Yolcu, E. Egrioglu, C.H. Aladag, High order fuzzy time series forecasting method based on an intersection operation, App. Math. Model. 40 (2016), https://doi.org/10.1016/j.apm.2016.05.012.

[50] Q. Dong, X. Ma, Enhanced fuzzy time series forecasting model based on hesitant differential fuzzy sets and error learning, Expert Syst. Appl. 166 (2021), https://doi.org/10.1016/j.eswa.2020.114056.