Deming least square regressed feature selection and Gaussian neuro-fuzzy multi-layered data classifier for early COVID prediction

Rathnamma V Mydukuri | Suresh Kallam | Rizwan Patan | Fadi Al-Turjman | Manikandan Ramachandran

1Department of Computer Science and Engineering, KSRM College Of Engineering (A), Kadapa, Andhra Pradesh, India
2Department of Computer Science & Engineering, Sree Vidyanikethan Engineering College, Tirupati, Andhra Pradesh, India
3Department of Computer Science & Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India
4Research Center for AI and IoT, Artificial Intelligence Engineering Department, Near East University, Mersin, Turkey
5School of Computing, SASTRA Deemed University, Thanjavur, Tamil Nadu, India

Correspondence
Rizwan Patan, Department of Computer Science & Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada 520007, India.
Email: prizwan5@gmail.com

Abstract
Coronavirus disease (COVID-19) is a harmful disease caused by the new SARS-CoV-2 virus. COVID-19 disease comprises symptoms such as cold, cough, fever, and difficulty in breathing. COVID-19 has affected many countries and their spread in the world has put humanity at risk. Due to the increasing number of cases and their stress on administration as well as health professionals, different prediction techniques were introduced to predict the coronavirus disease existence in patients. However, the accuracy was not improved, and time consumption was not minimized during the disease prediction. To address these problems, least square regressive Gaussian neuro-fuzzy multi-layered data classification (LSRGNF-M-LDC) technique is introduced in this article. LSRGNFM-LDC technique performs efficient COVID prediction with better accuracy and lesser time consumption through feature selection and classification. The preprocessing is used to eliminate the unwanted data in input features. Preprocessing is applied to reduce the time complexity. Next, Deming Least Square Regressive Feature Selection process is carried out for selecting the most relevant features through identifying the line of best fit. After the feature selection process, Gaussian neuro-fuzzy classifier in LSRGNFM-LDC technique performs the data classification process with help of fuzzy if-then rules for performing prediction process. Finally, the fuzzy if-then rule classifies the patient data as lower risk level, medium risk level and higher risk level with higher accuracy and lesser time consumption. Experimental evaluation is performed by Novel Corona Virus 2019 Dataset using different metrics like prediction accuracy, prediction time, and error rate. The result shows that LSRGNFM-LDC technique improves the accuracy and minimizes the time consumption as well as error rate than existing works during COVID prediction.

KEYWORDS
classification, coronavirus disease, feature selection, fuzzy technique

1 | INTRODUCTION

The occurrence of a new coronavirus called SARS-CoV-2 has infected many people worldwide. COVID-19 pandemic has resulted in more than 193,825 deaths in the last few months. Many researchers introduced different techniques to predict coronavirus before becoming worse. A long short-term memory (LSTM) method was introduced in Tomar and Gupta (2020) for COVID-19 prediction. The prediction of the number of...
positive cases and the number of recovered cases was more accurate in a certain range. Though the prediction accuracy was improved, time consumption was not minimized by the data-driven LSTM method.

A new fuzzy assisted system was designed in Adwibowo (2020) Srivastava et al. (2020) to identify the dental care safety concerning the patient and environmental conditions. The fuzzy assisted system allocated the estimation depending on patient body temperature, travel history, ventilation rate, and disinfection frequency. But the error rate was not minimized by the fuzzy assisted system. BAT model was introduced in Huang et al. (2020) for finding their order preference through similarity to ideal solution (TOPSIS) to address the best return date. However, the space complexity was not minimized by the BAT model.

A numerical model was introduced in Sujath et al. (2020) for predicting the COVID-2019 spread. Linear regression and Multilayer perceptron were introduced to anticipate the epidemiological ailment and COVID-2019 cases in India. But the prediction accuracy was not improved by the numerical model. The development of the COVID-19 pandemic was carried out in Bhardwaj (2020) Karmore et al., 2020 for different countries by the logistic model. The designed model carried out the regression analysis depending on the least-square fitting. But the computational complexity was not minimized.

A multiple ensemble neural network model was introduced in Melin et al. (2020) with fuzzy response aggregation for COVID-19 time series. Ensemble neural networks comprised of different modules to perform the prediction under diverse conditions. However, the prediction accuracy was not improved by the designed model. Non-pharmaceutical interventions (NPIs) were introduced in Lahiri et al. (2020) with prediction model for the prevention of pandemic of SARS-CoV-2 in Indian population. But the prediction time was not minimized by NPIs.

Machine learning models were introduced in Rustam et al. (2020) to forecast the upcoming patients affected by COVID-19. Regression, LASSO, SVM, and exponential smoothing were introduced to forecast the COVID-19 factor. However, the space complexity problem during prediction remained unaddressed. An early prediction of the 2019-nCoV epidemic was carried out in Zhong et al. (2020) using a mathematical model and epidemiological data in China. But the computational cost was not minimized during the COVID prediction. A hybrid artificial-intelligence (AI) model was introduced in Zheng et al. (2020). Nagasubramanian et al. (2020) for COVID-19 prediction. A susceptible model was introduced to decide the infection rate for determining the transmission law. Though the complexity was minimized, the prediction accuracy was not improved by the AI model.

The problems identified from the existing prediction methods are lesser prediction accuracy, higher prediction time, higher computational cost, higher space complexity, higher error rate, and higher computational complexity. To solve the above-said issues, least square regressive Gaussian neuro-fuzzy multi-layered data classification (LSRGNFM-LDC) technique is introduced.

1.1 | Motivation

Coronavirus disease (COVID-19) is a dangerous disease caused from SARS-CoV-2 virus. Recently, several research works are developed for COVID by using machine learning models. But the time consumption was not minimized. In conventional methods, error rate and space complexity were not minimized. In many existing methods, prediction accuracy was not improved. Motivated by this, the LSRGNFM-LDC technique uses the Deming regression and neuro-fuzzy classifier for performing prediction process. The proposed Deming regression process is applied to select the most relevant feature with time consumption. In this work, neuro-fuzzy classifier is applied to categorize the patient data depended on the risk levels through fuzzy if-then rules with higher accuracy and lesser time consumption.

1.2 | Contributions

The contributions of this work are listed below.

- To perform efficient COVID prediction with better accuracy and lesser time consumption as compared to state-of-the-art works, LSRGNFM-LDC technique is introduced. LSRGNFM-LDC technique is carried out through two processes, namely, feature selection and classification.
- To select the more relevant features with minimal of time consumption as compared to state-of-the-art works, Deming least square regressive feature selection (DLSRFS) process is introduced in LSRGNFM-LDC technique. DLSRFS process is used for identifying the line of best fit.
- To improve the prediction accuracy with minimal time consumption as compared to state-of-the-art works, the Gaussian neuro-fuzzy classifier introduced in LSRGNFM-LDC technique. Data classification process is performed to categorize the patient data with help of fuzzy if-then rules for performing the prediction process. With help of these rules, the patient data is predicted as the higher-risk patient data, medium-risk patient data, or higher-risk patient data in a more accurate manner with higher accuracy and lesser time consumption.

The paper is organized into five different sections. In Section 2, the proposed LSRGNFM-LDC technique is explained with a neat architectural diagram. Section 3 explains the related works of COVID prediction techniques. In Section 4, the experimental settings are presented with a detailed result analysis for four different parameters. The conclusion is given in Section 5.
2 | RELATED WORKS

A new approach was introduced in Marmarelis (2020). Waheed et al. (2020) depending on data-guided detection and infection waves explained by Riccati with estimated parameters. The designed approach was employed with Covid-19 daily time-series data lead to epidemic time-course decomposition. An artificial-intelligence method was introduced in Alazab et al. (2020) depending on a deep convolutional neural network (CNN) to identify COVID-19 patients with real-world datasets. But the computational cost was not minimized by the artificial-intelligence method.

A new approach was introduced in Abdel-Basset et al. (2020b). The conditioner track proceeds the couple of CT images and their tags as contribution and products a significant representation of information is transmitted is used by division path to segment the new images. Toward support actual association among both paths, in its proposed an adaptive recalibration (RR) and recombination module that permits concentrated data exchange between paths with a trivial increase in computational complexity.

A new framework was introduced in Abdel-Basset et al. (2020a) with respect to the heart rate and sleep data from wearable devices to estimate the epidemic development of COVID-19 in different cities. However, the computational complexity was not minimized. An objective approach was introduced in Petropoulos and Makridakis (2020) to forecast the COVID-19. The data was reliable and performed disease detection with an increase in confirmed COVID-19 with sizable, connected uncertainty. But the computational cost was not minimized by the objective approach.

A nonlinear machine learning method was introduced in Kavadi et al. (2020) for global pandemic prediction of COVID-19. Progressive Partial Derivative Linear Regression model was introduced to identify the best parameters efficiently. However, the prediction accuracy was not improved by the designed method. A new methodology was introduced in Fokas et al. (2020) for predicting the time evolution of individuals in each country infected with SARS-CoV-2. But the prediction time was not minimized by the designed methodology.

An online forecasting mechanism was introduced in Abdulmajeed et al. (2020) to stream the data from Nigeria Center for Disease Control. Also, the designed mechanism updated the ensemble model for providing the COVID-19 updates. However, the error rate was not minimized by an online forecasting mechanism. An innovative educational technique was introduced in DeFilippis et al. (2020) to provide experiential learning. FITs were supported in the early stage of the COVID-19 pandemic and anticipated the challenges in the United States. But the space complexity was not minimized by innovative educational techniques.

AI platform was introduced in Ke et al. (2020) to find the drugs with anti-corona virus activities. Though the error rate was minimized, the time consumption for prediction was not minimized by AI. New technology was introduced in Vaishya et al. (2020) to identify the cluster of cases and to forecast the virus through collecting and examining all previous data. But the computational cost was not minimized.

A new hybrid approach (HSMA_WOA) was introduced in Abdel-Basset et al. (2020c) depending on the SMA and WOA to resolve the image segmentation problems (ISP) for identifying the optimal threshold values. A novel meta-heuristic algorithm named slime mold algorithm (SMA) was presented for enhancing Kapur’s entropy by the whale optimization algorithm. The designed algorithm failed to solve flow shop scheduling issues, DNA fragment assembly issues, and parameter evaluation of the photovoltaic solar cell. A novel IoT system based decision-making model was developed by Abdel-Basset et al. (2020c) and Al-Turjman and Deebak (2020) to find and examine patients with type-2 diabetes. A new decision-making model was introduced on type-2 neuromorphic numbers by the VIKOR method. A decision support system was introduced for the precise prediction of type-2 diabetes risks for patients. But the accuracy was not improved.

A data analytics and visualization were introduced in Abdel-Basset et al. (2019) to examine malignant tumors and find weak spots of the tumor. However, failed to generate a greater positive impact and influence. MapReduce framework and fusion algorithm was introduced in Chang (2018a, 2018b) for medical imaging simulations. But the more varieties of gene simulations were not considered.

The disruptive technologies were introduced Abdel-Basset et al. (2020a) to analyze and restrict the spread of COVID-19. However, the space complexity was not reduced. Susceptible-Infected-Recovered (SIR) model was introduced Tutsoy et al. (2020) for the evaluation of COVID-19 casualties. But the error rate was not reduced. Mathematical and computational models were developed Torrealba-Rodriguez et al. (2020) for the estimation of the COVID-19 cases. However, it failed to reduce the prediction time. Gaussian prediction model was introduced Karaçuha et al. (2020) for forecasting the short-term future of the pandemic. However, the computational complexity was not minimized.

3 | METHODOLOGY

The coronavirus disease 19 (COVID-19) pandemic has resulted in the propagation of clinical prediction models to perform the diagnosis, disease severity estimation, and prognosis. A COVID-19 originated from β-coronavirus was reported in December 2019 at Wuhan city in China. On March 11, 2020, COVID-19 was announced as the public health crisis of international distress by World Health Organization (WHO). As of August 17, 2020, the Times of India website reported India's overall coronavirus crosses 26 lakh and 50,921 deaths. In Table 1 listed the detailed description of the variables used in this paper.
The coronavirus gets directly transmitted through cough, contact-transmission, sneeze, and respiratory globules. Patients are infectious before the beginning of clinical symptoms and contagiousness lasts up to three weeks after the recovery. Also, patients with mild symptoms were identified as infective. But the existing researchers failed to predict the COVID with higher accuracy and lesser time consumption. To address the existing problems, LSRGNFM-LDC technique is introduced. The main objective of the LSRGNFM-LDC technique is to perform an efficient COVID prediction with better accuracy and lesser time consumption. LSRGNFM-LDC technique comprises two processes, namely feature selection and classification. In this technique, first, preprocessing is carried out for removing the unwanted data from the Novel Corona Virus 2019 dataset. Preprocessing is employed for minimizing the time complexity. Next, DLSRFS process is used in the LSRGNFM-LDC technique for choosing the most relevant features. The data classification is carried out by applying neuro-fuzzy classifier to perform the prediction process, with aid of relevant features. The architecture diagram of the LSRGNFM-LDC technique is illustrated in Figure 1.

Figure 1 explains the architecture diagram of the LSRGNFM-LDC technique. Initially, the patient data is collected from the input database. The preprocessing is used for eliminating the unwanted data with lesser time complexity. After preprocessing, regression process is carried to select the most relevant feature via identifying the line of best fit. After the feature selection process, the fuzzy rules are constructed to perform the classification process. The classification process is used to classify the patient data with relevant features. In this way, the efficient COVID data prediction is carried out with higher accuracy and lesser time consumption. A brief description of the feature selection and classification process is given in the below sub-section.

### 3.1 Deming least square regressive feature selection process

DLSRFS process is introduced in LSRGNFM-LDC technique to select the more relevant features from the dataset. DR-FS algorithm is an errors-in-variables model. DLSRFS process identifies the line of best fit for performing the relevant feature selection. DLSRFS process is not the same as the linear regression where it identifies the errors on both axes. The DLSRFS process determines the maximum likelihood assessment of the error-in-variable model. The steps involved in the DLSRFS process is illustrated in Figure 2.

Table 1: Description of the variables

| Variable | Description |
|----------|-------------|
| $P_i = P_1, P_2, \ldots P_n$ | Large number of patient data |
| $F_{tj} = F_{t1}, F_{t2}, \ldots F_{tm}$ | Number of features |
| $n$ | Number of data |
| $m$ | Number of features |
| $e_j$ | Error value |
| $v_j$ | Ratio of their variance |
| $I_0$ | Intercept |
| $s_1$ | Slope |
| $F_{t1}$ and $\hat{a}_i$ | Estimate of true values of “$F_{t1}$” and “$a_i$” |
| $ln(t)$ | Input layer output |
| $P_{ai}$ | Patient data with the most relevant feature |
| weight0 | Initial weight allocated at the input layer |
| $b$ | Bias |
| Hidden(t) | Hidden layer result |
| weightih | Weight allocated between the input layer and the hidden layer |
| Out(t) | Output layer |
| weightho | Weight allocated between the hidden layer and the output layer |
| PredictionAcc | Prediction accuracy |
| PredictionTime | Prediction time |
| ErrorRate | Error rate |
| SpaceComplexity | Space complexity |

The coronavirus gets directly transmitted through cough, contact-transmission, sneeze, and respiratory globules. Patients are infectious before the beginning of clinical symptoms and contagiousness lasts up to three weeks after the recovery. Also, patients with mild symptoms were identified as infective. But the existing researchers failed to predict the COVID with higher accuracy and lesser time consumption. To address the existing problems, LSRGNFM-LDC technique is introduced. The main objective of the LSRGNFM-LDC technique is to perform an efficient COVID prediction with better accuracy and lesser time consumption. LSRGNFM-LDC technique comprises two processes, namely feature selection and classification. In this technique, first, preprocessing is carried out for removing the unwanted data from the Novel Corona Virus 2019 dataset. Preprocessing is employed for minimizing the time complexity. Next, DLSRFS process is used in the LSRGNFM-LDC technique for choosing the most relevant features. The data classification is carried out by applying neuro-fuzzy classifier to perform the prediction process, with aid of relevant features. The architecture diagram of the LSRGNFM-LDC technique is illustrated in Figure 1.

Figure 1 explains the architecture diagram of the LSRGNFM-LDC technique. Initially, the patient data is collected from the input database. The preprocessing is used for eliminating the unwanted data with lesser time complexity. After preprocessing, regression process is carried to select the most relevant feature via identifying the line of best fit. After the feature selection process, the fuzzy rules are constructed to perform the classification process. The classification process is used to classify the patient data with relevant features. In this way, the efficient COVID data prediction is carried out with higher accuracy and lesser time consumption. A brief description of the feature selection and classification process is given in the below sub-section.
features “$F_t = F_{t1}, F_{t2}, \ldots F_{tm}$.” From the expression, “$n$” denotes the number of data and “$m$” denotes the number of features in each dataset. DLSRFS process explain the available data ($a_j, Ft_j$) are calculated readings of true values ($a_j', Ft_j'$) that lie on regression line given by,

$$a_j = a_j' + e_j \quad (1)$$

$$\Delta_j = \delta_j' + v_j \quad (2)$$

From (1) and (2), “$e_j$” denotes the error value. “$v_j$” denotes the ratio of their variance. Both values are independent of each other. The intercept “$l_0$” and the slope “$s_1$” are determined as,

$$a_j = l_0 + s_1 F_t_j \quad (3)$$

From (3), “$\hat{F}_t_j$” and “$\hat{a}_j$” denotes the estimate of true values of “$F_t_j$” and “$a_j$” respectively. With help of the above equation, the DLSRFS process identifies the best fit line for every input feature. Also, it chooses the more relevant features to effectively predict the COVID disease. For achieving better outcomes, the DLSRFS process reduces the weighted sum of squared residuals. As a result, the DLSRFS process accurately selects the relevant features for disease prediction. The algorithmic processes of DLSRFS are given as,
Algorithm 1. Deming Least Square Regressive Feature Selection

Input: Number of features \( F_{t1}, F_{t2}, F_{t3}, \ldots, F_{tm} \) of patient data
Output: Select most relevant features with minimal time consumption

// Deming Least Square Regressive Feature Selection Algorithm
Step 1: Begin
Step 2: For patient data \( F_{tj} \) with number of features
Step 3: Apply Deming least square regression analysis
Step 4: Find the best fit line
Step 5: Select the most relevant features for prediction
Step 6: End For
Step 7: End

Algorithm 1 demonstrates the step-by-step process of the DLSRFS algorithm. With the above algorithmic process, the DLSRFS algorithm reduces the time consumption for relevant feature selection from the input Novel Corona Virus 2019 Dataset as compared to conventional works. Initially, the number of patient data is taken from Novel Corona Virus 2019 dataset as input. These patient data are collected from the dataset. Then, Deming least square regression analysis is used for finding the best fit line. Finally, the DLSRFS algorithm is employed to choose the most relevant features for COVID disease prediction.

3.2 Gaussian neuro-fuzzy multi-layered data classification

Gaussian neuro-fuzzy multi-layered data classification (GNFMLDC) process is carried out to classify the data points with higher accuracy and lesser time consumption. GNFMLDC process two processes, namely fuzzy logic and neural network. The fuzzy logic includes three operations, namely fuzzification operation, inference operation, and defuzzification operation. It is shown in Figure 3.

From Figure 3, the initially GNFMLDC process collects the input \( I(t) \) and determines the required performance. In the fuzzification operation, the inputs are converted to fuzzy sets. The goal of the inference operation is the efficient conversion of fuzzy input to the fuzzy output using an if-then loop. GNFMLDC process is designed with a set of rules termed fuzzy rules. The rules in the GNFMLDC process are denoted in the form of conditional statements. Finally, the defuzzification operation converts the fuzzy set to the output.

A neuro-fuzzy system is defined as a fuzzy system with a learning algorithm to determine fuzzy sets and fuzzy rules through processing the features of input data. The learning process aimed at local information and resulted in modification of the fundamental fuzzy system. The learning process of the neuro-fuzzy system includes the semantical properties of the fundamental fuzzy system. A neuro-fuzzy system is considered a multi-layer (i.e., three-layer) feedforward neural network. The first layer denotes the input layer with input variables. The hidden layer symbolizes the fuzzy rules, and the third layer denotes the output variables. Initially, the most relevant features of input data are considered as an input in the LSRGNFMLDC technique and given to the input layer. It is given by,

\[
I(n) = \sum_{i=1}^{n} P_a(t) \times \text{weight}_0 + b
\]  

From Equation (4), “In(t)” denotes the input layer output. “Pa(t)” denotes the patient data with the most relevant feature. “\text{weight}_0” represents the initial weight allocated at the input layer. “b” symbolizes the bias. Then, the input layer result is transmitted to the hidden layer. In that layer, a neuro-fuzzy system is used with the fuzzy rules and fuzzy sets. Fuzzy sets are trained as the connection weights and bias. The fuzzy set is portrayed using the membership function. It classifies the element in the fuzzy set as the continuous or the discrete form. A Gaussian membership function is the generalization of the characteristic function of a defined subset. Fuzzy if-then rules are employed for taking the correct decision based on the input data. The fuzzy rules are depending on the if-then rule condition as described in Equation (5).

\[
\text{if}(\text{age} \mu_1, \text{body temperature} \mu_2, \text{travel history} \mu_3, \text{and chronic diseases} \mu_4), \text{then } Y_i = \text{lower risk, medium risk, higher risk}
\]

In GNFMLDC, four input membership functions, and one output membership functions are employed. The four-input function are age, body temperature, travel history, and chronic disease. The three-output membership function of the GNFMLDC is lower risk, medium risk, and higher risk. After considering the membership function, the fuzzy rule is formulated as,
if $\mu_1 > 50$ $\& \mu_2 > 39^\circ C$ $\& \mu_3 < 14$ days $\& \mu_4 = 1$, then $Y_i = \text{higher risk}$  \(6\)

if $\mu_1 > 50$ $\& 37^\circ C < \mu_2 < 39^\circ C$ $\& 14 < \mu_3 < 28$ $\& \mu_4 = 1$, then $Y_i = \text{medium risk}$  \(7\)

if $\mu_1 > 50$ $\& \mu_2 = 37^\circ C$ $\& \mu_3 > 28$ days $\& \mu_4 = 0$, then $Y_i = \text{lower risk}$  \(8\)

From (6), (7), and (8), the load is determined as under load, normal load, and overload while comparing to the threshold metal weight using the if-then condition. The output triangular membership function is described in Figure 4.

The values of the output Gaussian membership function vary between 0 and 1 to categorize the fuzzy members that belong to the fuzzy set. Less than 0.25 is representing a lower risk. 0.5 is denoting a medium risk. Greater than 0.5 is representing a higher risk. The de-fuzzification process converts the fuzzy sets into output results. The hidden layer output is given by,

$$\text{Hidden}(t) = \sum_{i=1}^{n} \ln(t) * \text{weight}_{oi} + [\text{weight}_{io} * Y_i]$$ \(9\)

From Equation (9), “Hidden(t)” denotes the hidden layer result. “weight$_{oi}$” denotes the weight allocated between the input layer and the hidden layer. After that, the hidden layer results are transmitted to an output layer. An output of the GNFMLDC classifier renders the classification output for each input patient data. The result of the output layer is formulated as,

$$\text{Out}(t) = \text{weight}_{ho} * \text{Hidden}(t)$$ \(10\)

From Equation (10), “Out(t)” denotes the output layer result. “weight$_{ho}$” represents the weight allocated between the hidden layer and the output layer. In this way, the patient data is classified into three types (i.e., Lower risk, Medium risk, and Higher risk) for performing the COVID prediction. The algorithmic description of the GNFMLDC process is explained below,

Algorithm 2. Gaussian Neuro-Fuzzy Multi-Layered Data Classification

**Input:** Number of patient data with the most relevant features

**Output:** Improve prediction accuracy with minimal time consumption

//Gaussian Neuro-Fuzzy Multi-Layered Data Classification Algorithm

Step 1: Begin

Step 2: For each input patient data ‘$P_0$’

Step 3: Perform the fuzzification process

Step 4: Analyze the fuzzy interference function

Step 5: Determine the de-fuzzification process

Step 6: Obtain the classification results as higher-risk patient data, medium risk patient data, or higher risk patient data

Step 7: End for

Step 8: End
Algorithm 2 explains the step-by-step process of the Generalized Recurrent Neural Brown Boosting Classifier Algorithm. Initially, the relevant features of each patient data are collected as input. After that, a fuzzification operation is carried out to convert the input into fuzzy sets. Inference operation is the conversion of fuzzy input to the fuzzy output through the if-then loop. The fuzzy inference operation is used to minimize the prediction time. Finally, the defuzzification operation changes the fuzzy set to output. With help of these rules, the patient data is predicted as the higher risk patient data, medium risk patient data or higher risk patient data with maximum accuracy and minimum time consumption. With the above algorithmic process, LSRGNFM-LDC technique attains the better COVID prediction performance through the classification process when compared to the traditional works.

4 | EXPERIMENTAL SETTINGS

Experimental evaluation of proposed LSRGNFM-LDC technique and existing methods data-driven LSTM Tomar and Gupta (2020), fuzzy assisted system Adwibowo (2020) and disruptive technologies Abdel-Basset et al. (2020c) are implemented using Java language. The experiment of the LSRGNFM-LDC technique is conducted using Novel Corona Virus 2019 Dataset taken from the Kaggle. The URL of the mentioned dataset is given as Kaggle.com (2020). The dataset has daily level information on the number of affected cases, deaths, and recovery from the 2019 novel coronavirus. It is time-series data and the number of cases on a given day is the cumulative number. The dataset comprises eight files. The dataset totally includes 1608 columns. There are integer 1502 columns, String 63 columns, Uuid 18 columns and other 25 columns. Among the eight files, we have taken COVID19_open_line_listfile for conducting the experiments. The dataset comprises 44 features and 13,174 instances. The features are ID, age, sex, city, country, province, and so on. Among these features, relevant features are selected to perform the classification for COVID prediction. The experimental result of the proposed LSRGNFM-LDC technique is compared against three conventional methods namely data-driven LSTM, fuzzy assisted system, and disruptive technologies. For the experimental consideration, the numbers of patient data are taken from 100 to 1000. Totally ten runs are carried out for all the parameters evaluation. With the help of these numbers of patient data, the classification is done to improve the performance of COVID19 prediction. The Attributes Description in this dataset is depicted in Table 2.

Result analysis of LSRGNFM-LDC technique is compared with existing techniques with certain parameters are,

- Prediction accuracy
- Prediction time
- Error rate
- Space complexity

4.1 | Impact on prediction accuracy

Prediction accuracy is defined as the ratio of the number of patient data that are correctly predicted the risk level through the classification process to the total number of patient data taken. Consequently, the prediction accuracy is formulated as,

\[
\text{PredictionAcc} = \left( \frac{\text{Number of patient data that are correctly predicted the risk level}}{\text{Number of patient data}} \right) \times 100
\]  \hspace{1cm} (11)

From (11), the prediction accuracy is calculated. The prediction accuracy is measured in terms of percentage (%).

| Attribute       | Description                                                                 |
|-----------------|-----------------------------------------------------------------------------|
| S. no           | Serial number                                                               |
| Observation date| Date of the observation in MM/DD/YYYY                                       |
| Province/state  | Province or state of the observation                                         |
| Country/region  | Country of observation                                                      |
| Last update     | Time in UTC at which the row is updated for the given province or country.  |
| Confirmed       | Cumulative number of confirmed cases till that date                         |
| Deaths          | Cumulative number of deaths till that date                                  |
| Recovered       | Cumulative number of recovered cases till that date                         |

TABLE 2 Attributes description
Table 3 describes the prediction accuracy of a different number of patient data. For determining the prediction accuracy, the proposed and existing three methods Tomar and Gupta (2020), Adwibowo (2020), and Abdel-Basset (2020a) is implemented in the Java language. To determine the COVID prediction performance, proposed and traditional works Tomar and Gupta (2020), Adwibowo (2020), and Abdel-Basset et al. (2020a) are implemented with Novel Corona Virus 2019 Dataset. When considering the number of patient data as 300, the proposed LSRGNFM-LDC technique correctly predicts the 280-patient data whereas data-driven LSTM Tomar and Gupta (2020), Fuzzy assisted system Adwibowo (2020), and Disruptive technologies Abdel-Basset et al. (2020c) correctly predicts the 258-patient data, 269 patient data and 264 patient data respectively. Therefore, the prediction accuracy attained by the proposed LSRGNFM-LDC technique is 93% and the prediction accuracy attained by Data-driven LSTM Tomar and Gupta (2020), Fuzzy assisted system Adwibowo (2020), and Disruptive technologies Abdel-Basset et al. (2020c) is 86, 90, and 88% respectively. From the result, the prediction accuracy using the proposed LSRGNFM-LDC technique is higher when compared to other works by Tomar and Gupta (2020), Adwibowo (2020), and Abdel-Basset et al. (2020a).

The proposed LSRGNFM-LDC technique performs an accurate COVID prediction with an increasing number of patient data. LSRGNFM-LDC technique is employed to improve the prediction accuracy for selecting the most relevant features by using Deming regression. This process finds the line of best fit for performing the relevant feature selection. Then, the Gaussian neuro-fuzzy classifier is employed to classify the patient data for performing COVID prediction based on the risk levels through fuzzy if-then rules. The patient data is categorized into three types such as Lower risk, Medium risk, and Higher risk for performing the COVID prediction. This helps the LSRGNFM-LDC technique to improve prediction accuracy performance. Therefore, the proposed LSRGNFM-LDC technique illustrates that prediction accuracy is said to be minimized by 9% when compared to Tomar and Gupta (2020), 5% when compared to Adwibowo (2020) and 7% when compared to Abdel-Basset et al. (2020a) respectively.

### 4.2 Impact on prediction time

Prediction time is defined as the amount of time consumed for predicting the number of patient data as lesser risk data, medium risk data, and higher risk data. It is the product of the number of patient data and the amount of time consumed for predicting one patient data. Therefore, the prediction time is calculated as,

\[
\text{Prediction Time} = \text{Number of patient data} \times \text{time consumed for predicting one data}
\]

From (12), the prediction time is determined. The prediction time is measured in terms of milliseconds. With the help of the obtained experimental values, the graph is illustrated in Figure 5.

Figure 5 portrays the performance evaluation of the prediction time using three different methods namely data-driven LSTM Tomar and Gupta (2020) and Fuzzy assisted system Adwibowo (2020), Disruptive technologies Abdel-Basset et al. (2020c) and LSRGNFM-LDC technique. Figure 5 explains the prediction time consumed for a different number of patient data. With aim of determining the COVID prediction time, the proposed and existing two methods Tomar and Gupta (2020), Adwibowo (2020), and Abdel-Basset et al. (2020a) are implemented in Java language from a novel coronavirus dataset. When considering the number of patient data as 900, the proposed LSRGNFM-LDC technique consumes 14.85 ms prediction time whereas the Data-driven LSTM Tomar and Gupta (2020), Fuzzy assisted system Adwibowo (2020) and Disruptive technologies Abdel-Basset et al. (2020a) consume 26.1, 23.4, and 20.9 ms respectively. Therefore, it is significant that the prediction time

| Number of patient data (number) | Data-driven LSTM | Fuzzy assisted system | Disruptive technologies | LSRGNFM-LDC technique |
|--------------------------------|------------------|-----------------------|-------------------------|------------------------|
| 100                            | 80               | 85                    | 83                      | 94                     |
| 200                            | 85               | 90                    | 87                      | 95                     |
| 300                            | 86               | 90                    | 88                      | 93                     |
| 400                            | 88               | 91                    | 90                      | 95                     |
| 500                            | 89               | 92                    | 91                      | 96                     |
| 600                            | 88               | 90                    | 89                      | 95                     |
| 700                            | 87               | 89                    | 88                      | 96                     |
| 800                            | 89               | 90                    | 89                      | 96                     |
| 900                            | 89               | 91                    | 90                      | 95                     |
| 1000                           | 91               | 91                    | 92                      | 95                     |

Abbreviations: LSRGNFM-LDC, least square regressive Gaussian neuro-fuzzy multi-layered data classification; LSTM, long short-term memory.
consumed using the proposed LSRGNFM-LDC technique is lesser when compared to other works by Tomar and Gupta (2020), Adwibowo (2020), and Abdel-Basset et al. (2020a).

As described in Figure 5, the prediction time of all the methods gets increased while increasing the number of patient data. From the figure, the red color bar denotes the prediction time of the LSRGNFM-LDC technique whereas the red color bar, green color bar, and violet color bar denote the prediction time of the data-driven LSTM, Fuzzy assisted system and Disruptive technologies. The proposed LSRGNFM-LDC technique outperforms better COVID prediction with an increasing number of patient data. LSRGNFM-LDC technique significantly decreases the prediction time using applying the DLSRFS process. DLSRFS process is used to select the most relevant features through determining the line of best fit. After that, Gaussian neuro-fuzzy classifier in LSRGNFM-LDC technique performs the data classification process with help of fuzzy if-then rules for performing prediction process. Next, Gaussian neuro-fuzzy classifier is used to categorize the patient data as lower risk level, medium risk level, and higher risk level. In this way, the LSRGNFM-LDC technique minimizes the prediction time. As a result, the proposed LSRGNFM-LDC technique illustrates that prediction time is said to be minimized by 41% when compared to Tomar and Gupta (2020), 34% when compared to Adwibowo (2020) Kolhar et al. (2020) and 25% when compared to Abdel-Basset et al. (2020a) respectively.

4.3 Impact on the error rate

The error rate is defined as the ratio of the number of patient data that are incorrectly predicted through the classification process to the total number of patient data taken. As a result, the error rate is determined as,

\[ \text{Error Rate} = \left( \frac{\text{Number of patient data that are incorrectly predicted}}{\text{Number of patient data}} \right) \times 100 \]

(13)

From (13), the error rate is determined. The error rate is measured in terms of percentage (%).

The error rate comparison of three different methods, namely data-driven LSTM Tomar and Gupta (2020), fuzzy assisted system Adwibowo (2020), disruptive technologies Abdel-Basset et al. (2020c) and LSRGNFM-LDC technique is illustrated in Table 4. For a better assessment, the various number of patient data are taken as input ranging from 100, 200, 300... 1000. For the same number of input data, three various error rate results are attained as shown in Table 4. Let us consider the number of patient data as100 for conducting the experiments. Among the 100-patient data, six patient data are incorrectly predicted by the LSRGNFM-LDC technique. Also, 20, 15, and 17 patient data are incorrectly predicted by data-driven LSTM Tomar and Gupta (2020), fuzzy assisted system Adwibowo (2020), and disruptive technologies Abdel-Basset et al. (2020c) respectively. Therefore, the error rate obtained by the LSRGNFM-LDC technique, data-driven LSTM Tomar and Gupta (2020), fuzzy assisted system Adwibowo (2020), and disruptive technologies Abdel-Basset et al. (2020a) is 6, 20, 15, and 17%, respectively. The statistical discussion confirms that the proposed technique attains a lesser error rate than existing methods.

Table 4 demonstrates the error rate experimental results of three different prediction techniques with respect to the number of data collected from the COVID database. As described in Table 4, the space complexity of all the methods gets increased while increasing the number of patient data. From the obtained values, the error rate of the LSRGNFM-LDC technique is lesser than other existing works. The reason behind the performance enhancement is the application of feature selection and classification techniques for predicting the COVID patient data. The LSRGNFM-
LDC technique uses the Deming regression and neuro-fuzzy classifier for performing the prediction process. Besides, the DLSRFS process selects the most relevant feature selection for disease prediction. DLSRFS process discovers the line of best fit for performing the relevant feature selection. After that, the neuro-fuzzy classifier is applied to categorize the patient data depended on the risk levels through fuzzy if-then rules. The fuzzy if-then rule classifies the patient data into lesser risk, medium risk, and higher risk. In this manner, the proposed technique reduces the error rate during the COVID prediction. Likewise, the comparative analysis indicates that the error rate is found to be minimized by 60, 50, and 55% when compared to three existing classification algorithms Tomar and Gupta (2020), Adwibowo (2020), and Abdel-Basset et al. (2020a), respectively.

### 4.4 Impact on space complexity

Space complexity is defined as the amount of memory space required for performing COVID prediction. It is the product of the number of patient data and the space required for storing one patient data. It is given by,

\[
\text{SpaceComplexity} = \frac{\text{Number of patient data}}{\text{space consumed for storing one data}}
\]

From (14), the space complexity is computed. The space complexity is measured in terms of megabytes (MB). With attained space complexity table values, the graph is shown in Figure 6.

Figure 6 illustrates the space complexity experimental results of three prediction techniques with respect to the number of data taken from the COVID database. The various number of patient data are considered as input ranging from 100, 200, 300...1000. Let us consider the number of patient data as 500 for experiment evaluation. Among 500 patient data, the memory space consumed by one patient data is 0.0185 MB by the LSRGNFM-LDC technique. Also, 0.031, 0.028, and 0.0416 MB of memory space are consumed by data-driven LSTM Tomar and Gupta (2020), fuzzy assisted system Adwibowo (2020), and disruptive technologies Abdel-Basset et al. (2020c) respectively for one patient data prediction. Therefore, the space complexity attained by LSRGNFM-LDC technique, data-driven LSTM Tomar and Gupta (2020), fuzzy assisted system Adwibowo (2020), and disruptive technologies Abdel-Basset et al (2020c) is 10.5, 23.5, 18, and 20.8 MB, respectively.

From the figure, the red color bar denotes the space complexity of the LSRGNFM-LDC technique whereas the red color bar, green color bar, and violet color bar denotes the space complexity of data-driven LSTM, fuzzy assisted system, and disruptive technologies. With collected patient data, the space complexity of the three techniques is determined. From the attained values, the space complexity of the LSRGNFM-LDC technique is lesser than other conventional works. This is because of applying the feature selection process during the COVID prediction process. LSRGNFM-LDC technique uses the DLSRFS process for eliminating the irrelevant features. Also, the DLSRFS process chooses the most relevant feature selection through finding the line of best fit. DLSRFS process is used to effectively predict the COVID disease. This helps in minimizing the space consumption during the COVID prediction. Similarly, the result analysis denotes that space complexity is minimized by 57, 42, and 51% when compared to two existing classification algorithms Tomar and Gupta (2020), Adwibowo (2020), and Abdel-Basset et al. (2020c), respectively.

### Table 4: Tabulation for error rate

| Number of patient data (Number) | Error rate (%) |
|--------------------------------|----------------|
|                               | Data-driven LSTM | Fuzzy assisted system | Disruptive technologies | LSRGNFM-LDC technique |
| 100                            | 20              | 15              | 17              | 6               |
| 200                            | 15              | 10              | 13              | 5               |
| 300                            | 14              | 10              | 12              | 7               |
| 400                            | 13              | 9               | 10              | 5               |
| 500                            | 11              | 8               | 9               | 4               |
| 600                            | 12              | 10              | 11              | 5               |
| 700                            | 13              | 11              | 12              | 4               |
| 800                            | 11              | 10              | 11              | 4               |
| 900                            | 11              | 9               | 10              | 5               |
| 1000                           | 9               | 9               | 8               | 5               |

Abbreviations: LSRGNFM-LDC, least square regressive Gaussian neuro-fuzzy multi-layered data classification; LSTM, long short-term memory.
CONCLUSION AND FUTURE SCOPE

A new technique termed LSRGNFM-LDC technique is introduced for performing the COVID prediction with better accuracy and lesser time consumption. The preprocessing is applied on each input features to eradicate the irrelevant data. Preprocessing is used to minimize the time consumption. DLSRFS process in LSRGNFM-LDC technique selects the most relevant features through identifying the line of best fit. In LSRGNFM-LDC technique, the neuro-fuzzy classifier utilizes the fuzzy if-then rules for performing the prediction process. The fuzzy if-then rule predicts the patient data into lesser risk, medium risk, and higher risk in a more accurate manner with higher accuracy and lesser time consumption. The wide-ranging experimental evaluation is performed with the COVID database. The quantitative results are verified in terms of higher prediction accuracy and lesser time as well as space complexity when compared to other related works. In future, our proposed work is also proceed using Guided Whale Optimization Algorithm (Guided WOA) based on Stochastic Fractal Search to selects the relevant features. Weakly supervised deep learning model is applied to classify the patient data.

CONFLICT OF INTEREST

The authors declare there is no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset, below is the reference Kaggle.com. 2020. Novel Corona Virus 2019 Dataset. [online] Available at: <https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset> [Accessed 20 September 2020].

ORCID

Rizwan Patan https://orcid.org/0000-0003-4878-1988
Fadi Al-Turjman https://orcid.org/0000-0001-6375-4123

REFERENCES

Abdel-Basset, M., Chang, V., & Mohamed, R. (2020a). HSMA_WOA: A hybrid novel Slime mould algorithm with whale optimization algorithm for tackling the image segmentation problem of chest X-ray images. Applied Soft Computing (106642).
Abdel-Basset, M., Chang, V., Hawash, H., Chakrabortty, R. K., & Ryan, M. (2020b). FSS-2019-nCov: A deep learning architecture for semi-supervised few-shot segmentation of COVID-19 infection (106647). Knowledge-Based Systems.
Abdel-Basset, M., Chang, V., & Nabeeh (2020c). An Intelligent Framework using Disruptive Technologies for COVID-19 analysis. In Technological Forecasting and Social Change (120431). Elsevier.
Abdel-Basset, M., Manogaran, G., Gamal, A., & Chang, V. (2019). A novel intelligent medical decision support model based on soft computing and IoT. IEEE Internet of Things Journal, 7(5), 4160–4170.
Abdulmajeed, K., Adeleke, M., & Popoola, L. (2020). Online forecasting of COVID-19 cases in Nigeria using limited data (pp. 105683). Data in Brief.
Adiwibowo, A. (2020). Fuzzy logic assisted COVID 19 safety assessment of dental care. medRxiv.
Alazab, M., Awajan, A., Mesleh, A., Abraham, A., Jatana, V., & Alhyari, S. (2020). COVID-19 Prediction and Detection Using Deep Learning. International Journal of Computer Information Systems and Industrial Management Applications, 12, 168–181.
Al-Turjman, F., & Deebak, B. D. (2020). Privacy-aware energy-efficient framework using the internet of medical things for COVID-19. IEEE Internet of Things Magazine, 3(3), 64–68.
Bhardwaj, R. (2020). A predictive model for the evolution of COVID-19. Transactions of the Indian National Academy of Engineering, Springer, 5, 133–140.
Chang, V. (2018a). Computational intelligence for medical imaging simulations. Journal of Medical Systems, 42(1), 10.
Chang, V. (2018b). Data analytics and visualization for inspecting cancers and genes. Multimedia Tools and Applications, 77(14), 17693–17707.
DeFilippis, E. M., Schmidt, A. C. S., & Reza, N. (2020). Adapting the educational environment for cardiovascular fellows-in-training during the COVID-19 pandemic. *Journal of the American College of Cardiology*, 75(20), 2630–2634.

Fokas, A. S., Dikaix, N., & Kastis, G. A. (2020). Mathematical models and deep learning for predicting the number of individuals reported to be infected with SARS-CoV-2. *Journal of the Royal Society Interface*, 17(169), 20200494.

Huang, G., Pan, Q., Zhao, S., Gao, Y., & Gao, X. (2020). Prediction of COVID-19 outbreak in China and optimal return date for university students based on propagation dynamics. *Journal of Shanghai Jiaotong University (Science)*, 25, 140–146.

Kaggle.com. (2020). Novel Corona Virus 2019 Dataset. Retrieved from https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset

Karmore, S., Bodhe, R., Al-Turjman, F., Kumar, R. L., & Pillai, S. (2020). IoT based humanoid software for identification and diagnosis of Covid-19 suspects.

Karaçuha, E., Onal, N. O., Ergun, E., Tabatadze, V., Alkas, H., Karaçuha, K., Tontu, H. O., & Nu, N. V. N. (2020). Modeling and prediction of the Covid-19 cases with deep assessment methodology and fractional calculus. *IEEE Access*, 8, 164012–164034.

Kolhar, M., Al-Turjman, F., Alameen, A., & Abualhaj, M. M. (2020). A three layered decentralized IoT biometric architecture for city lockdown during COVID-19 outbreak. *IEEE Access*, 8, 163608–163617.

Lahiri, A., Jha, S. S., Bhattacharya, S., Ray, S., & Chakraborty, A. (2020). Effectiveness of preventive measures against COVID-19: A systematic review of In Silico modeling studies in indian context. *Indian Journal of Public Health*, 64(6), 156.

Marmarelis, V. (2020). Predictive modeling of Covid-19 data in the US: Adaptive phase-space approach. *IEEE Open Journal of Engineering in Medicine and Biology*, 1, 207–213.

Melin, P., Monica, J. C., Sanchez, D., & Castillo, O. (2020, June). Multiple ensemble neural network models with fuzzy response aggregation for predicting COVID-19 time series: The case of Mexico. In Healthcare (Vol. 8, p. 181). Multidisciplinary Digital Publishing Institute.

Nagasubramanian, G., Sankayya, M., Al-Turjman, F., & Tsaramirsis, G. (2020). Parkinson data analysis and prediction system using multi-variant stacked auto encoder. *IEEE Access*, 8, 127004–127013.

Petropoulos, F., & Makridakis, S. (2020). Forecasting the novel coronavirus COVID-19, PLoS one, 15(3), e0231236.

Rustam, F., Reshi, A. A., Ergun, E., Tabatadze, V., Alkas, H., Karaçuha, K., Tontu, H. O., & Nu, N. V. N. (2020). Modeling and prediction of the Covid-19 cases with deep assessment methodology and fractional calculus. *IEEE Access*, 8, 164012–164034.

Sujath, R., Chatterjee, J. M., & Hassanien, A. E. (2020). A machine learning forecasting model for COVID-19 pandemic in India (p. 1). Stochastic Environmental Research and Risk Assessment.

Tomar, A., & Gupta, N. (2020). Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Science of The Total Environment*, Elsevier, 728, 1–6.

Torrealba-Rodríguez, O., Conde-Gutiérrez, R. A., & Hernández-Javier, A. L. (2020). Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. *Chaos, Solitons and Fractals* (Vol. 138, 109946). Elsevier.

Tutsiy, O., Colak, S., POLAT, A., & Balikci, K. (2020). A novel parametric model for the prediction and analysis of the COVID-19 casualties. *IEEE Access*, 8, 193898–193906.

Vaishya, R., Javadi, M., Khan, I. H., & Haleem, A. (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(4), 327–329.

Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., & Pinheiro, P. R. (2020). Covidgan: Data augmentation using auxiliary classifier GAN for improved covid-19 detection. *IEEE Access*, 8, 91916–91923.

Zheng, N., Du, S., Wang, J., Zhang, H., Cui, W., Kang, Z., Yang, T., Lou, B., Chi, Y., Long, H., & Ma, M. (2020). Predicting covid-19 in china using hybrid AI model. *IEEE Transactions on Cybernetics*, 50(7), 2891–2904.

AUTHOR BIOGRAPHIES

**Dr. Rathnamma V Mydukuri** obtained his Master’s Degree and Ph.D. in Computer Science Engineering form JNTUA, Anantapur. Now she is working as an Associate Professor in Department of Computer Science Engineering, KSRM College of Engineering (A),YSR Kadapa(Dist), A.P.

- India. Her areas of interest include Computer Networks, Mobile Adhoc Networks, Trust management in MANETS and other latest trends in technology.

She has more than 11 years of experience in teaching and Research in the area of Computer Science and Engineering. She published more than 25 publications in the national and international reputed journals and conferences.

**Suresh Kallam** received the bachelor’s and master’s degrees from Jawaharlal Nehru Technological University, Hyderabad, and the Ph.D. degree from VIT University, Vellore, India. He is currently a Associate Professor at Department of Computer Science and Engineering, Sree Vidyanikethan Engineering College, Autonomous College, A. Rangampet, Tirupati, AP and previously worked as Professor, Division Head, and the Program Chair of the M.Tech. degree, School of Computer Science Engineering, Galgotias University, Greater Noida, India. He worked as a Foreign Faculty Member of the East China University of Technology, ECIT Nanchang Campus, Jiangxi, China, and a Visiting Faculty Member of Jiangxi Normal University, China. He has published over 35 national and International conference papers and 15 international journals. He
Rizwan Patan, is an Assistant Professor in the Department of Computer Science and Engineering, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, India since 2019. Former Assistant Professor at school of computing science and engineering, Galgotias University, NCR Delhi, Indian from 2017-19. Dr. Rizwan is completed his Ph.D. (Computer Science and Engineering) in 2017 at school of computer science and engineering, VIT University, Vellore, Tamil Nadu, Indian and B. Tech and M. Tech completed respectively 2012 and 2014 on Jawaharlal Nehru Technological University Anantapur, India. Dr. Rizwan has published reputed 20 SCI journals and 10 free Scopus indexed journals, and also presented paper in National/International Conferences, published book chapters in CRC Press, IGI global, Elsevier, and Edited as books. Dr. Rizwan has ten Indian patents one USA patent. Dr. Rizwan is a guest editor for International Journal of Grid and Utility Computing (Inderscience), Recent Patents on Computer Science, Information Medical Unlock (Elsevier), Neural Computing and Applications (Springer). Dr. Rizwan received award from world research council and American medical Council in the title of innovative researcher on Big Data and IoT for year 2019.

Prof. Dr. Fadi Al-Turjman received his Ph.D. in computer science from Queen's University, Kingston, Ontario, Canada, in 2011. He is a full professor and a research center director at Near East University, Nicosia, Cyprus. Prof. Al-Turjman is a leading authority in the areas of smart/intelligent, wireless, and mobile networks' architectures, protocols, deployments, and performance evaluation. His publication history spans over 250 publications in journals, conferences, patents, books, and book chapters, in addition to numerous keynotes and plenary talks at flagship venues. He has authored and edited more than 25 books about cognition, security, and wireless sensor networks' deployments in smart environments, published by Taylor and Francis, Elsevier, and Springer. He has received several recognitions and best papers' awards at top international conferences. He also received the prestigious Best Research Paper Award from Elsevier Computer Communications Journal for the period 2015-2018, in addition to the Top Researcher Award for 2018 at Antalya Bilim University, Turkey. Prof. Al-Turjman has led a number of international symposia and workshops in flagship communication society conferences. Currently, he serves as an associate editor and the lead guest/associate editor for several well reputed journals, including the IEEE Communications Surveys and Tutorials (IF 22.9) and the Elsevier Sustainable Cities and Society (IF 4.7).

Manikandan Ramachandran obtained his Ph.D. in VLSI Physical design from SASTRA Deemed University, India. He received his Bachelor of Engineering in Computer Science from Bharathidasan University and Master of Technology in VLSI Design from SASTRA Deemed University. He possesses two decades of academic and 12 years of research experience in the field of Computer Science and Engineering. He has more than 120 research contributions to his credit, which are published in referred and indexed journals, book chapters and conferences. He is presently working as Assistant Professor at SASTRA Deemed University for the last 14 years. He has delivered many lectures and has attended and presented in International Conferences in India and Abroad. He has edited more than 80 research articles to his credit, which includes his editorial experience across refereed and indexed journals, conferences and book chapters at national and international levels. His contemporary research interests include Big Data, Data Analytics, VLSI Design, IoT and Health Care Applications.