IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 2 May 2024, accepted 2 June 2024, date of publication 7 June 2024, date of current version 17 June 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3411162*

RESEARCH ARTICLE

# Predictive Analytics for Mortality: FSRNCA-FLANN Modeling Using Public Health Inventory Records

**MANISH KUMAR[1], SUSHIL KUMAR SINGH[2], (Member, IEEE), AND SUNGGON KIM[1]**

[1]Department of Computer Science and Engineering, Seoul National University of Science and Technology, Seoul 01811, South Korea
[2]Department of Computer Science and Engineering, Marwadi University, Rajkot, Gujarat 360003, India

Corresponding author: Sunggon Kim (sunggonkim@seoultech.ac.kr)

This work is financially supported by Seoul National University of Science and Technology, Seoul, South Korea.

**ABSTRACT** Predictive analytics involves the use of Artificial Intelligence (AI) and Machine Learning (ML) techniques to analyze current and historical data, identify patterns, and make predictions about future outcomes. In the context of healthcare, predictive analytics is invaluable for understanding and addressing various challenges, with a particular emphasis on predicting mortality rates in specific communities during illnesses such as diabetes, heart disease, and drug overdose. In this study, we propose a unique approach combining Feature Selection Regression based on Neighborhood Component Analysis (FSRNCA) with a less complex single-layer Functional Link Artificial Neural Network (FLANN) model to predict mortality rates using publicly available open-source inventory, i.e., Big Cities Health Inventory (BCHI). Our methodology leverages regularized FSRNCA to select relevant features from the BCHI dataset, considering factors such as demographics, socio-economic indicators, and health metrics. Subsequently, the FLANN model is trained using the selected features to predict mortality rates in diverse urban populations. The model achieves a high $R^2$ score of 91%, outperforming other competitive ML algorithms. Additionally, the proposed technique is successfully tested with datasets such as PIMA, BUPA, ECOLI, and LYMPHOGRAPHY for the classification of various other illnesses. Furthermore, FLANN exhibits minimal computational complexity and rapid training-testing times, highlighting its practicality for real-world applications. The minimal computational complexity and training-testing time between the dense multilayer neural networks reveal the excellent potential and utility of the proposed technique.

**INDEX TERMS** Functional link artificial neural network, feature selection, machine learning, healthcare, smart city, society.

## I. INTRODUCTION

This Modern engineering innovations are assisting in making all cities sustainable and intelligent. Governments are developing smart city initiatives that aim to achieve several goals, including improving performance, sustainability, public participation, and service delivery [1]. Cities with a diverse network of sensors and advanced techniques offer meaningful information. These data support the management of industries, local administrations, social services, healthcare, security, safety, and environmental sustainability benefits [2]. The Internet of Things (IoT) and Wireless Sensor Network (WSN) technologies are gaining popularity today, and devices such as smartwatches or other sensor modules can aid in the monitoring of an individual's or a community's health. Big data devices, data analytics, and AI models will be deeply integrated into smart city ecosystems. These data can be used for city planning as well as

The associate editor coordinating the review of this manuscript and approving it for publication was Mehul S. Raval.

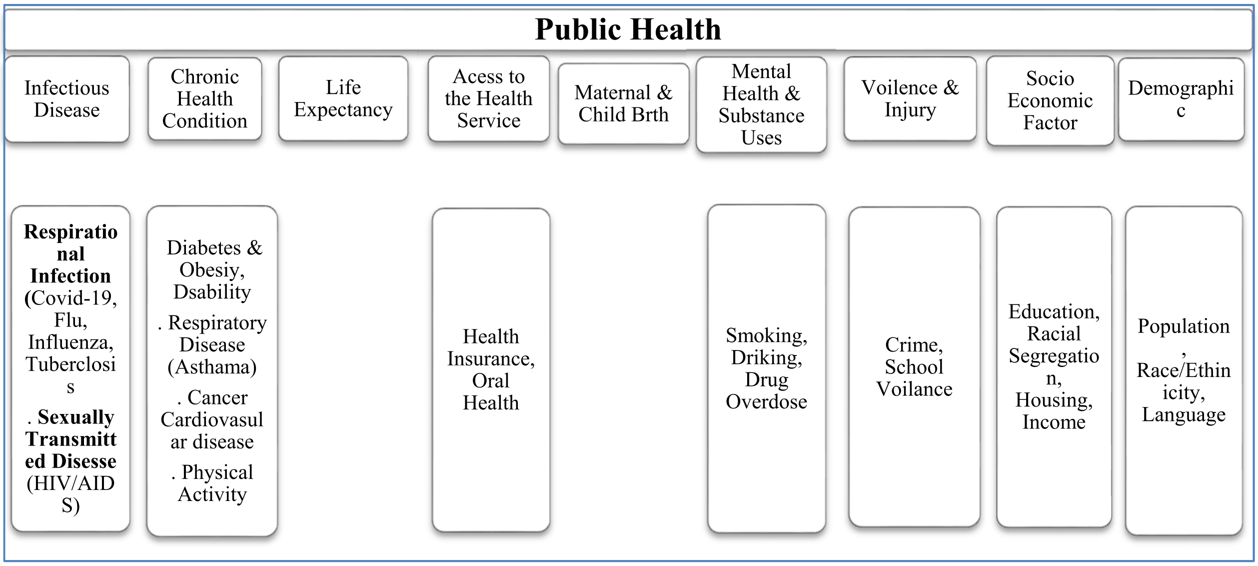



**FIGURE 1.** Parameters of the public health.

disaster preparedness. A huge influx of population is coming to metropolitan cities because of adequate job opportunities and infrastructure for communication, water, energy, education, safety, and health. Currently, healthcare services are a major concern during pandemic days, and the sudden increase in patients can stymie the entire healthcare system. Such a situation can be addressed with the help of statistical tools where early prediction is possible so that medical resources can be deployed accordingly. It is evident that the average life expectancy of a person enhances significantly with technological advancement during the last century. Providentially, Artificial Intelligence (AI) or Machine Learning (ML)-based regression model can predict such conditions, and also necessary steps can be taken before any casualties [3], [4]. It uses previous data inventory and mapped the input features and target output to build a predictive model [5]. The developed model can be deployed as a web or mobile application so that the intensity of any illness can be measured as well as managed [6].

In any megacity, people of various regions and races are living together, and their diverse population, illiteracy, poverty, and food habits are playing major roles in the public's health [7]. It may create numerous issues like specific diseases, mental illness, radicalization, narcotics, or trafficking. Figure 1 depicts the various disease conditions that influence public health. It not only ensures wellness against specific diseases, but it also reflects a community's health services, infrastructure, mental, and socioeconomic structure. These factors play a major role in the overall health of a society and need to be studied regularly so that any accident can be eradicated before any major causalities occur. Apart from that, this research article mainly focuses on illnesses like diabetes, cardiovascular disease, and drug overdose deaths. Also, what are the major factors that impact these diseases? This research also aids the local authority in allocating healthcare resources like hospitals, clinics, doctors, and nurses in the disease-prone areas.

ML-based regression techniques are very helpful in designing such an early predictive system by using available data and solving several health-related problems [8], [9], [10]. These techniques are based on the analytical thinking of humans and can solve any linear as well as non-linear problem by using inputs and target mapping. These techniques are also used in the automated diagnosis of diseases of coronary artery and heart [11], [12]. Verma and Aggarwal [13] discussed the significant role of ML and FDA in medicine and validated COVID-19 drug discovery by using ML techniques. Jangir et al. [14] have discussed the application of ML and deep learning-based techniques in the diagnosis of diabetes and discussed how this disease relies not only on blood glucose level but also on parameters like obesity, age, sex, and body mass index (BMI). Lotfi and Reazee [15] have discussed problems like the "dimensionality curse," where there are too many learning parameters in any multilayer neural network-based technique. It will not only increase the computational complexity of the regression algorithm but also make the whole network sluggish. This can be circumvented with the help of FLANN-based techniques and applied as a universal approximator. The FLANN is used in the classification of cancer gene and breast cancer [16], [17]. Apart from the disease classification using FLANN, this research focuses on the estimation of mortality in different illness conditions in various communities by incorporating parameters crucial information pertaining to ethnic demographics (e.g., Black, White, Asian, Hispanic), educational attainment (e.g., college graduation rates), and various socio-economic indicators (e.g., race-specific poverty rates, access to medical insurance, participation in public assistance programs). Additionally, we compiled city-specific data on infrastructure expenditure and demographic factors.

The proposed model helps in finding the upcoming situation in the different illness conditions, and the key contributions of this paper are listed as follows:

- We have compiled a 10-year dataset (2010–2020) from authentic sources to facilitate the analysis of early predictions and feature importance
- We propose Feature Selection Regression using Neighborhood Component Analysis (FSRNCA) to analyze the importance of input attributes on diabetes, drug overdose, and cardiovascular disease death rates.
- The proposed computationally economical model, Functional Link Artificial Neural Network (FLANN), is designed to perform parallel prediction tasks.
- We compared the proposed technique with six other state-of-the-art ML-based techniques, including Linear Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), Ensemble Learning (EL), Wide Layered Artificial Neural Network (W-ANN), and Tri-Layered Artificial Neural Network (T-ANN), based on various subjective and quantitative metrics.

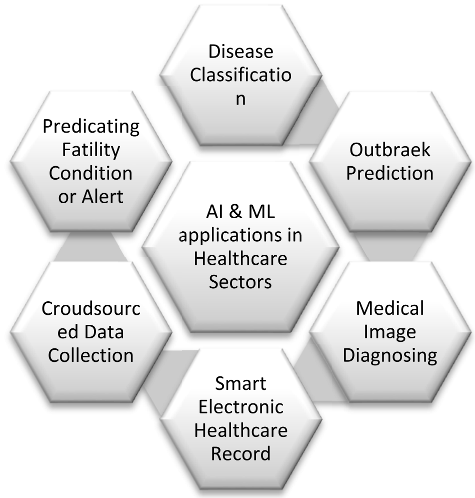
According to the author's understanding, there is currently no reported literature on the application of the FSRNCA-FLANN technique for predicting death rates related to various illnesses. The article is structured as follows: Section II provides an overview of pertinent research. Along with related work, key considerations of this research are also discussed here. Section III highlights the mathematical approach of the designed algorithm, which consists of feature selection as well as the proposed prediction model. Finally, Sections IV and V explore significant observations and their implications.

## II. RELATED WORK

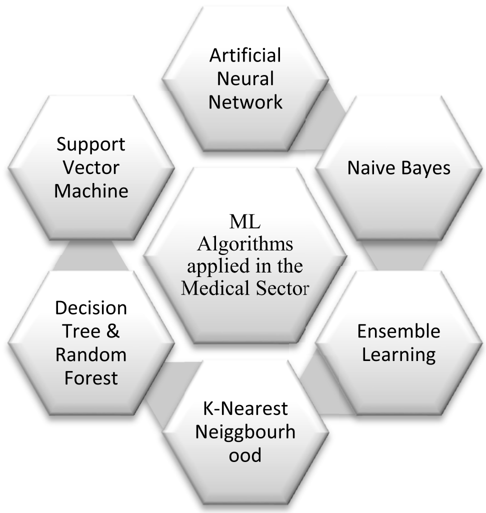
In this section, we discuss the use of ML in the various areas of the healthcare sector, including diagnosis, predicting any outbreak or alert on the basis of available data, and offering an insight and intelligent decision-making capabilities. It also highlights the requirement of current research.

### A. EXISTING RESEARCH

Javaid et al. [3] presented a comprehensive survey discussing the significance of AI and ML techniques in healthcare. They explored the potential of computer-aided diagnosis and how automation enhances treatment efficiency. Figure 2(a) illustrates the current applications of AI for medical purposes. Lyons and Lazaroiu [18] proposed the utilization of data from smart cities using IoT modules to estimate the crisis in the case of COVID-19. They included five parameters in their study: age, race or ethnicity, gender, education, and geographic region. Ghazal et al. [19] discussed available ML-based techniques for diagnosing diseases in smart cities, and Figure 2(b) showcases a few ML techniques applied in different medical applications.




(a)




(b)

**FIGURE 2. (a) AI & ML applications in healthcare sectors. (b) ML algorithms in the medical sector.**

Deep Learning (DL) networks such as Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) have been employed in disease diagnosis, and previous research reports indicate their efficiency comparable to other classification techniques [20], [21]. The convolutional filter layer is utilized for feature selection and subsequent classification purposes. Table 1 not only presents relevant research but also highlights the efficacy of various ML approaches in addressing health-related issues. ML and DL techniques are applied in both conditions, i.e., computer-aided disease classification (diagnosis) and regression problems like forecasting, where scientists estimate upcoming situations, as seen in the case of COVID-19. Since these techniques map input attributes to the target, the authenticity of the data source is crucial in such conditions. The efficiency of these techniques depends not only on the data source but also on

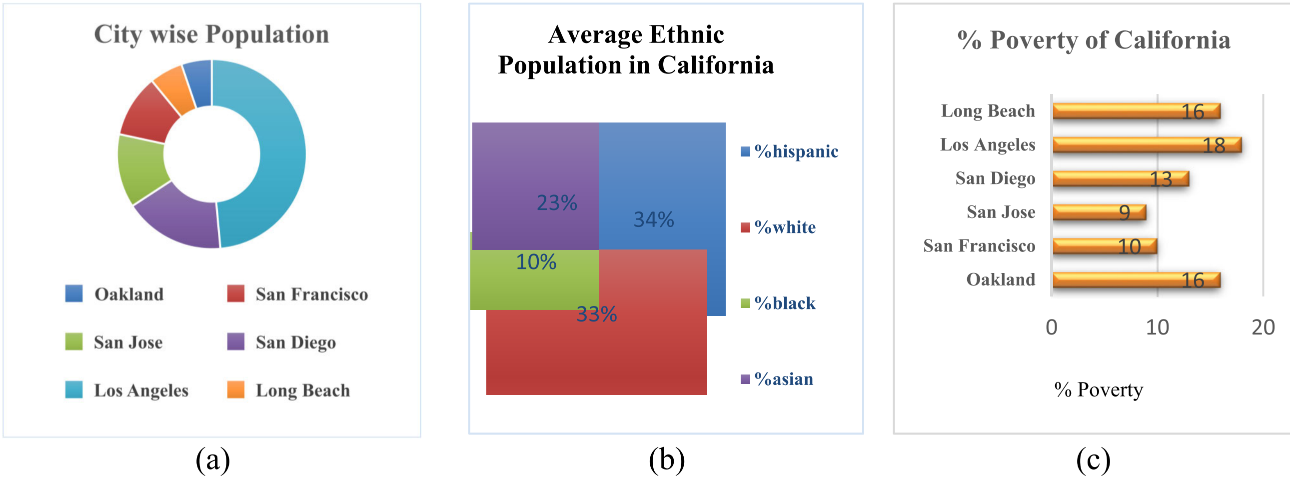



**FIGURE 3.** California graphs (a) City Population, (b) Average Ethnic Population and (C) Percentage Poverty.

parameters such as the number of features considered during training, whether those features are correlated with the target, and the selection of ML models.

AI-based tools greatly aid in disease diagnosis and can provide early alerts before any fatalities. However, they require significant computational resources during deployment, especially when computationally expensive. Data size sometimes plays a major role, and unnecessary attributes may contribute to overfitting issues. Therefore, in addition to dense layer networks like CNN and Multi-Layer Perceptron (MLP), Subrajeet et al. have implemented a simple Artificial Neural Network (ANN) variant, such as FLANN, for the segmentation of cancerous cells. This network does not utilize any hidden layers and can effectively address nonlinear problems. The researchers developed this model based on a bloodstain image dataset, achieving an accuracy of up to 98% during simulation [22]. The computationally less expensive TinyML technologies, as highlighted in [23] provide new opportunities in this field and inspire researchers to develop simpler algorithms like FLANN. This research primarily employs FSRCNA for feature selection, aiming to eliminate overfitting in the proposed prediction algorithm. Additionally, it utilizes one of the least computationally expensive algorithms, i.e., FLANN, for predicting death tolls due to various diseases in different socioeconomic conditions.

### B. KEY CONSIDERATION

#### 1) DATA COLLECTION

Data for the present work is collected from the Big Cities Health Data Inventory (BCHI) and California is considered a test case in this research [24]. This data repository is real and built on the basis of various authorized sources. These data offer a statistical view of different ethnic groups living in the United States, along with information about their health conditions across various parameters. For example, Figure 3 illustrates the population distribution and city-wise poverty level in California. The proposed predictive model was constructed using ten-year data from six major cities and four ethnic groups spanning from 2010 to 2020. The attributes mainly included for this work are city information, years, race, uninsured people, college graduates, poverty, public assistance (including food), infrastructural expenses on parks, and how these parameters impact public health.

#### 2) COMPUTATIONALLY ECONOMICAL ML MODEL

Jun et al., have discussed the problem of overfitting with ML models due to the computational complexities of ANN networks. In overfitting, any ML-based prediction model behaves efficiently; however performance will be degraded with test or verification data. In any ANN, the model complexity is determined by the hidden layer numbers. This issue can be circumvented by reducing the predictive model complexity and also by increasing the amount of data size. In this research, the size of the dataset augmented with the help of expanding input with standard mathematical functions and opting for a less complex ANN variant like FLANN. The concept of FLANN was initially presented by Pao [32], and it is a single-layer-based ANN network. FLANN's architecture makes it a less complex model than any other conventional dense feedforward neural network. FLANN design consists of mainly three processes, i.e., (mathematical functional expansion in the input layer, estimation process, and adaptive process for updating weights). The absence of hidden layers makes the whole model computationally inexpensive and also takes minimum training time than the other ANN-based predictive model.

#### 3) PARALLEL PREDICTIVE MODEL

The FLANN network is developed as a parallel approximation model to estimate death causality due to issues such as diabetes, cardiovascular disease, and drug overdose across various races. Based on the developed model, the vulnerability of another disease can also be predicted by providing relevant data to it.

**TABLE 1.** Current developments in the related work.

| Research Study | Technology Used | Dataset Information | Health Issues | | | | Feature Selection | Ethnicity/ Socio-Economical Attributes | Number of Attributes |
|---|---|---|---|---|---|---|---|---|---|
| | | | diabetes | cardiovascular | Drug-Overdose | others | | | |
| Nancy & George [18] | IoT & Machine Learning | Census Bureau's American Community Survey | × | × | × | Covid-19 | × | √ | 5 |
| Narin *et. al.*, [25] | K-Nearest Neighborhood & Genetic Algorithm | Physionet | × | √ | × | × | | × | 8 |
| Pratap & Kokil [26] | Pre-trained Convolutional Neural Network (Transfer Learning) & SVM | Structural Analysis of the Retina (STARE) & Standard Diabetic Retinopathy Database (DIARETDB0) | × | × | × | Cataract | √ | × | N/A |
| Sunil *et. al,* [14] | Functional Link Convolutional Neural Network (FL-CNN) | Localized Diabetes Dataset (LDD) | √ | × | × | × | √ | × | 12 |
| Tuli *et. al.,* [27] | Cloud Computing integrated with ensemble deep learning | Cleveland database, Gottsegen Hungarian Institute of Cardiology, Hungary | × | √ | × | × | × | × | 14 |
| Hu *et al.*, [28] | X-ray Image, CNN & Optimization Technique | COVIDetectioNet | × | × | × | Covid-19 | √ | × | N/A |
| Chen *et. al.*, [29] | CT images, SVM and ANN | UCI machine learning dataset repository | × | × | × | Brain Hemorrhage | × | × | NA |
| Wang *et. al.*; [30] | MR images and Sigle Layer Neural Network & Particle Swarm Optimization | DA-160, DA-255 | × | × | × | Pathological Brain | × | × | 1 |
| Zhang *et. al.*, [31] | Graph-CNN (GCNN) | Mini-MIAS | × | × | × | Breast Cancer | √ | × | NA |
| **Proposed Scheme** | Functional Link Artificial Neural Network | Drexel Urban Health Collaborative (https://bigcitieshealthdata.org) | √ | √ | √ | × | √ | √ | 8 |

## III. PROPOSED METHODOLOGY

This section highlights the data preprocessing and mathematical modeling of the FSRNCA and FLANN techniques. The algorithm of proposed algorithm is also discussed in this section.

### A. DATA PREPROCESSING

Mainly steps like replacing of missing data, data encoding, data normalization, and reshaping of data were performed in data preprocessing. The missing data is replaced by the median value of the each column. Again, any character like information are categorize and encoded with the help on One-Hot encoding, because during feature selection FSR-CNA accept only numerical value. After the completion of encoding, each column is normalized between [0-1]. This normalization step aims to prevent biases in prediction results arising from specific input features. The collected data was randomly shuffled and then partitioned for training, validation, and testing purposes. Additionally, data reshaping was performed to align with the requirements of the implemented ML model, as each model necessitates a specific data structure during both training and testing phases.

### B. FEATURE SELECTION REGRESSION USING NEIGHBORHOOD COMPONENT ANALYSIS (FSRNCA)

Yang et al., have discussed the importance of Neighborhood Component Analysis (NCA) in the ML and high dimensionality data. It is a non-parametric statistical technique based on nearest neighbor weighting. FSRNCA learns a feature weighting vector by maximizing the expected leave-one-out regression accuracy with a regularization term to learn feature weights for minimization of an objective function [33]. Consider a multi-class classification problem with training set $S$ and $n$ observations, aim is to predict the target i.e., $(y)$ from the $S$.

$$S = \left\{(x_i, y_i), i = 1, 2, \cdots, n\right\} \tag{1}$$

Equation (1) represents the training set. where, $y\epsilon\mathbb{R}$ are continuous and Equation (2) i.e. the probability $P(Ref\ (x) = x_j|S)$ that $x_j is\ picked\ from S$ as the reference point for $x$ is

$$P\left(Ref\ (x) = x_j \mid S\right) = \frac{k(d_w(x, x_j))}{\sum_{j=1}^{n} k(d_w(x, x_j))} \tag{2}$$

Now consider the leave-one-out application of this randomized regression model which predicts the response $x_i$ using the data $S^{-i}$, the training set S excluding the point $(x_i, y_i)$. Equation (3), the probability that point $x_j$ is picked as the reference point for $x_i$ is

$$P_{ij} = P\left(Ref\ (x_i) = x_j \mid S^{-i}\right) = \frac{k(d_w(x, x_j))}{\sum_{j=1, j\neq 1}^{n} k(d_w(x, x_j))} \tag{3}$$

where, $\hat{y}$ be the predicted response value from the regression model predicts and $y_i$ be the actual response. Let, $l : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a loss function and the average value $l(y_i, \hat{y}_i)$ is given in Equation (4).

$$l_i = E(l\left(y_i, \hat{y}_i|S^{-i}\right) = \sum_{j=1, j\neq 1}^{n} P_{ij}\left(l\left(y_i, y_j\right)\right) \tag{4}$$

After introducing the regularization parameter '$\boldsymbol{\lambda}$' in Equation(5), the objective function for minimization is:

$$\boldsymbol{f\ (w) = \frac{1}{n}\sum_{i=1}^{n} l_i + \lambda\sum_{r=1}^{p} w_r^2} \tag{5}$$

### C. FUNCTIONAL LINK ARTIFICIAL NEURAL NETWORK (FLANN) MODEL

The basic architecture of the proposed FLANN regression model is demonstrated in Figure 4. It comprises feature expansion, association of weights, network output, and an adaptive derivative algorithm which update the weights and bias respectively. This figure also demonstrates the absence of any hidden layer in the FLANN network. The features selected by the FSRNCA will serve as the input for FLANN and will be normalized using Equations (6) and (7).

$$x_n = \frac{X_k - X_{min}}{X_{max} - X_{min}} \tag{6}$$

$$t_n = \frac{X_k - t_{min}}{t_{max} - t_{min}} \tag{7}$$

Here, $x_n$, $t_n$ are normalized input and target data. These data are augmented with exponential functions. Hence, expanded data terms are shown in Equations (8) and (9),

$$X = \left\{\begin{array}{l} x_{n1}, e^{x_{n1}}, e^{2.x_{n1}}, e^{3.x_{n1}}, e^{3.x_{n1}}, e^{4.x_{n1}}, \\ x_{n2}, e^{x_{n2}}, e^{2.x_{n2}}, e^{3.x_{n3}}, e^{3.x_{n2}}, e^{4.x_{n2}}, \ldots, \\ \ldots, x_{nk}, e^{x_{nk}}, e^{2.x_{nk}}, e^{3.x_{nk}}, e^{3.x_{nk}}, e^{4.x_{nk}} \end{array}\right\} \tag{8}$$

$$T = \left\{\begin{array}{l} t_{n1}, e^{t_{n1}}, e^{2.t_{n1}}, e^{3.t_{n1}}, e^{3.t_{n1}}, e^{4.t_{n1}}, \\ t_{n2}, e^{t_{n2}}, e^{2.t_{n2}}, e^{3.t_{n3}}, e^{3.t_{n2}}, e^{4.t_{n2}}, \ldots, \\ \ldots, t_{nk}, e^{t_{nk}}, e^{2.t_{nk}}, e^{3.t_{nk}}, e^{3.t_{nk}}, e^{4.t_{nk}} \end{array}\right\} \tag{9}$$

where, $x_n$ and X are the normalized feature and feature vector respectively. Gaussian weights are set to values ranging from 0 to 1. The size of the weights will be equal to the expanded inputs, and it can be represented in Equation (10) as,

$$W = \left\{W^1, W^2, W^3, \ldots, W^N\right\} \tag{10}$$

where, $size\ (X) = size(W)$ and the network output will be represented in Equation (11).

$$\text{network output}: y = \sum_{i=1}^{i=N} (X_i W_i + b_i) \tag{11}$$

where, $i = 1, 2, \ldots, N$, $b = bias$ and $W^T$ is the initialized weights transpose. Also, Equation (11) can be simplified as,

$$y = \left[X.W^T + b_i\right] \tag{12}$$

Again, network output can be threshold by applying sigmoid(.) activation function and it is represented in Equation (13) as,

$$\hat{y}_{net} = \frac{1}{1 + e^{-y}} \tag{13}$$

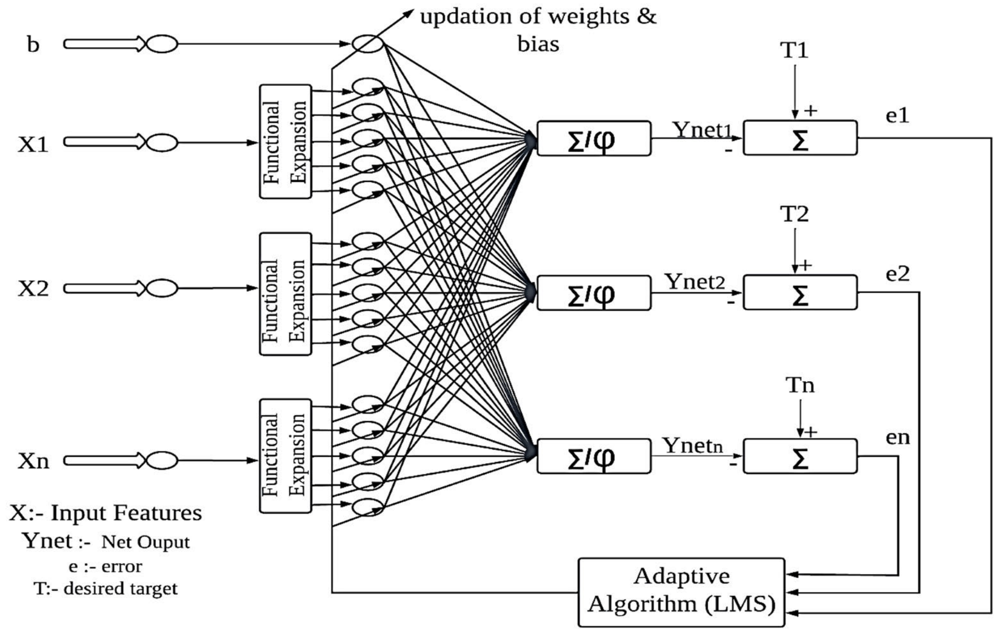



**FIGURE 4.** Architecture of the proposed FLANN network.

The proposed network error can be estimated by comparing $y_{net}$ values to the target values and it can be represented in Equation (14) as,

$$error : e = T - \hat{y}_{net} \tag{14}$$

The estimated error can be optimized by applying the Least Mean Squares (LMS) algorithm, which is an adaptive algorithm that will update the weights and biases of the FLANN until the error is at a minimum.

$$\Delta W = \frac{\sum_{k=1}^{P} 2\eta e_k X_k}{P} \tag{15}$$

$$\Delta b = \frac{\sum_{k=1}^{P} 2\eta e_k X_k}{P} \tag{16}$$

Equations (17-18) represent the updated weights and bias as,

$$W_k\left(n+1\right) = W_k\left(n\right) + \Delta W \tag{17}$$

$$b_k\left(n+1\right) = b_k\left(n\right) + \Delta b \tag{18}$$

where, $\eta$ is learning rate, $P$ total number of input pattern., $\Delta W$, $\Delta b$ are change in weight and bias and it is represented in Equations (15-16). The efficiency of the proposed models is estimated by computing statistical matrices like Root Mean Squares Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination ($R^2$). It is observed between the tested values and their respective reference or target values. Equations (19)–(22) represent the various performance metrics for evaluating the efficiency of applied regression models in this article.

$$MSE = \frac{1}{N}\sum\nolimits_{i=1}^{n}\left(O_{test} - O_{ref}\right) \tag{19}$$

$$RMSE = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{n}\left(O_{test} - O_{ref}\right)} \tag{20}$$

$$\text{MAE} = \frac{1}{N}\sum\nolimits_{i=1}^{n}\left|\left(O_{test} - O_{ref}\right)\right| \tag{21}$$

$$R^2 = \frac{\left(\sum_{i=1}^{n}\left(O_{test} - \overline{O_{test}}\right)\left(O_{ref} - \overline{O_{ref}}\right)\right)^2}{\left(\sum_{i=1}^{n}\left(O_{test} - \overline{O_{test}}\right)\right)^2\left(\sum_{i=1}^{n}\left(O_{ref} - \overline{O_{ref}}\right)\right)^2} \tag{22}$$

where, $O_{test}$ and $O_{ref}$ are the tested output from trained model and reference output respectively. The algorithm of the proposed FSRNCA-based FLANN network is summarized in Algorithm 1.

Steps 1-3 of the algorithm involve loading and preprocessing the collected dataset, saving it into the input and target arrays, and normalizing the data by scaling all array elements between 0 and 1. In Step 4, feature selection is performed by applying the FSRNCA technique, followed by data splitting in Step 5. Similarly, Steps 6-13 encompass the training phase of the FLANN network. Upon completion of training, the proposed network is tested with unknown inputs, which were not involved in the training process, as outlined in Step 14. Finally, Step 15 showcases the evaluation of the proposed model's performance in terms of various metrics to measure accuracy levels. The hyperparameters are used to design the proposed model is showcased in Table 2.

## IV. RESULTS AND DISCUSSION

All the experimentation tasks are performed in the MatLab 2022a simulation environment, and the workstation was integrated with an AMD Ryzen 5 PRO 4650U processor, 8 GB of RAM, and Radeon Graphics at a 2.10 GHz clock frequency. Figure (5) represents the FSRNCA-based feature importance in terms of weights. The regularized neighborhood component analysis can pick out important features from a larger

**Algorithm 1** FSRNCA-Based FLANN Prediction Model

**Input:** Providing year wise Big Cities Health Inventory Data as attributes as $(x_{n1}, x_{n2}, \ldots, x_{ni})$.
**Output:** Prediction of year wise casualties i.e. $(t_{n1}, t_{n2}, \ldots, t_{ni})$.

1. Start
2. Read the data and select the input as well as target data
3. Normalized the input & target data
4. Apply the regularized FSRNCA technique for the feature selection.
5. Selected features and respective target is augmented with the help of exponential function as '$X$ *and* $T$' and expanded feature is given as an input to the proposed FLANN model.

Training:

6. Initialize weights '$W$', bias '$b$', learning rate '$\eta$' and number of iterations.
7. Split augmented data as training, testing and validation set.
8. Perform steps 9-12 when stopping condition reach
9. Obtain the FLANN net input with given relation
   $y_{in} = \varphi(\sum_i^n (X_i W_i + b_i))$
   Here, activation function $\varphi$
10. Evaluate until least mean squares $(T - y_{in})$ and update the weight and bias
    $W_{new} \rightarrow W_{old} + \Delta W$
    $b_{new} \rightarrow b_{old} + \Delta b$
    Now calculate error using $=> E = (T - y_{in})^2$
11. If $y = T$ *then*
    Updating weight is not required
12. Test for stopping conditions i.e., (minimum error or maximum iteration then stop).
13. Stop Training

Testing:

14. Test the trained FLANN model with unknown test input i.e., '$X_t$' and calculate predicted casualties as,
    $y_{out} = \varphi(\sum_i^n (X_t W_{new} + b_{new}))$

Evaluate Model with Performance Measures:

15. Calculate $MSE$, $RMSE$, $MAE$, and, $R^2$.
16. Stop

pool, helping to remove irrelevant ones. Figure (5) exhibits that feature index 1, 5 and 12 have no significance in target prediction task as it attain value is 0. Hence, these features are dropped and not given to the FLANN. This step not only helps in reducing the computational time but also overcome the overfitting of FLANN by reducing feature size. This means the model will use less memory during both training and making predictions. This is especially valuable when there are limited computational resources or large datasets to work with. However, FSRNCA will take around 330 Seconds to find optimal '$\lambda$' parameter and it will minimizes the mean loss. The optimal '$\lambda$' and mean loss values are 0.0033 and 4.088. After, selecting important feature, it will be given to the proposed FLANN predictive model for training and validation. The performance of the proposed model is analyzed on the basis of subjective and quantitative measures.

Figure 6 represents the regression plots of the proposed FLANN as well as other competitive ML schemes such as SVM, LR, DT, EL, and MLP networks like Wide and Tri-layered artificial neural networks. These figures reveal that SVM shows less accuracy than the other implemented models because the predicted points are dispersed from the actual line. Similarly, the data points of LR are also far from the reference line. In fact, the plots of DT, EL, W-ANN, T-ANN, and proposed FLANN models show that the predicted data is closer to the reference line. The regression plots also demonstrate that W-ANN is predicting very well except for the lower and upper sections of the data point. This is greatly improved in T-ANN and the proposed FLANN network, except for some of the points in the middle. The key reason behind this is that the proposed FLANN and T-ANN

**TABLE 2.** Parameter of FSRNCA-based FLANN prediction model.

| Techniques | Parameters | Symbols/Details | Value/range/Type |
|---|---|---|---|
| FSRNCA | Regularization | $\lambda$ | [0 - 20] |
| | Solver | 'sgd', and 'lbfgs' | Stochastic Gradient Descent and Limited-memory BFG |
| | Iteration Limit | I1 | 30 |
| | Gradient Tolerance | GradientTolerance | 1e-4 |
| | Relative Threshold | Tol | 0.02 |
| FLANN | Functional Expansion | exponential | exp(.) |
| | Activation Function | Ø | Sigmoid(.) |
| | Learning Rate | $\eta$ | [0.001 - 0.1] |
| | Iteration Limit | I2 | 3000 |
| | Solver | LMS | Least Means Squares |

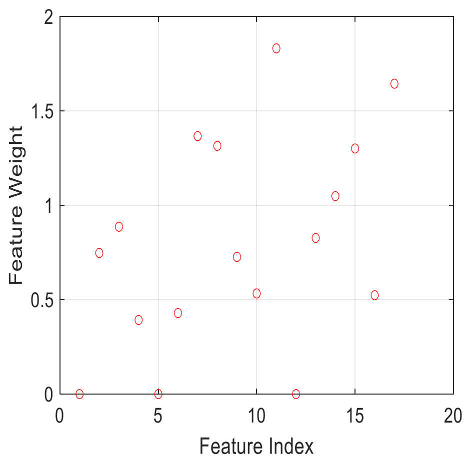
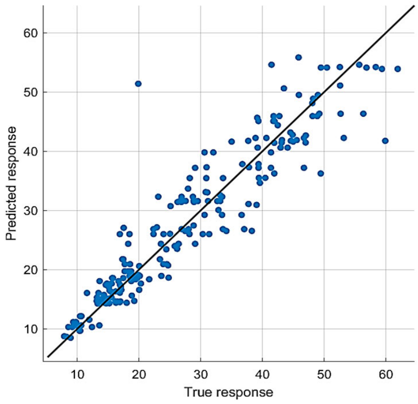
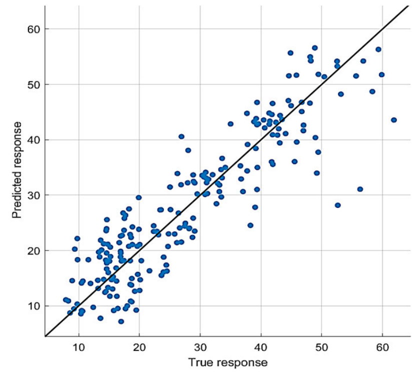


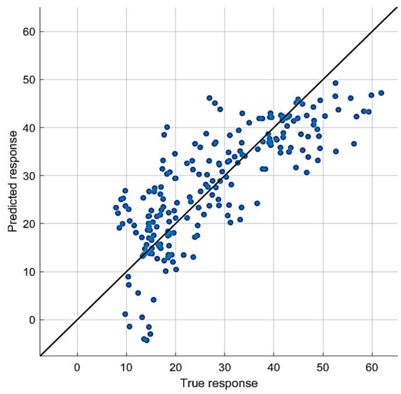

**FIGURE 5.** FSRNCA based feature importance in various illness condition.



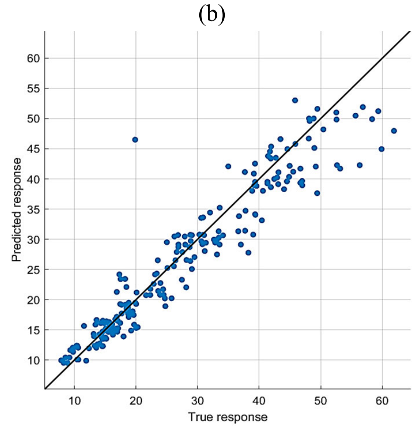

(a)



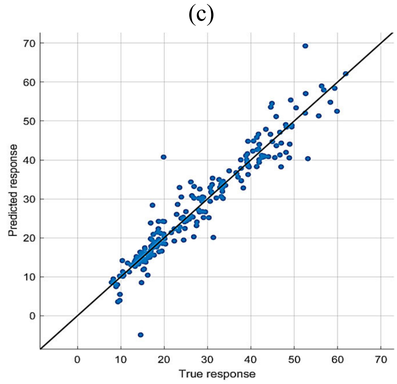

(b)



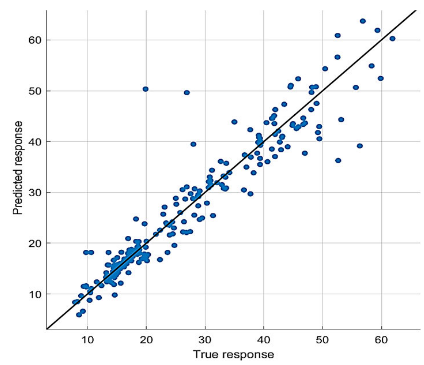

(c)



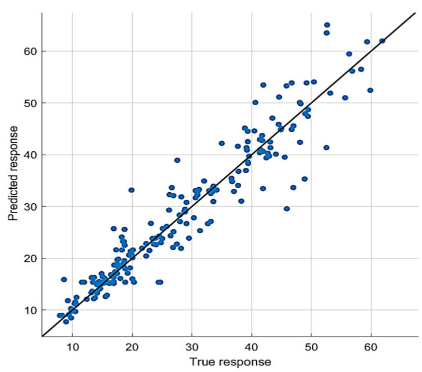

(d)




(e)




(f)




(g)

**FIGURE 6. Regression plot of various ML models: (a) Decision Tree (b) Linear Regression (c) SVM (d) Ensemble Learning (e) Wide-ANN (f) Tri-Layered ANN (g) Proposed FLANN.**

are less overfitted models than the W-ANN among the implemented neural network models. Since W-ANN consists of a fully connected layer of size 100, T-ANN uses three fully connected layers of size 10, whereas the proposed FLANN technique consists of no hidden layers, which makes the overall model less complex.

Apart from the abovementioned subjective analysis of performance, researchers suggested performance metrics like RMSE, $R^2$, MSE, and MAE in [34] to quantify the efficiency of the implemented models. The performance metrics data is provided in Table 3. Lower, RMSE, MSE, and MAE will showcase efficacy of implemented competitive algorithms. Here, $R^2$ values are considered to evaluate the performance of predictive algorithms because other parameters also change similarly to it. $R^2$ values close to 1 will help to find the efficiency of these ML techniques. The $R^2$ value of the proposed simple FLANN network is about 0.91, which demonstrates the accuracy of the proposed network, i.e., (91%). However, the accuracy of SVM was about 61%, which is lower than other ML-based schemes. Also, value of other performance metrics reveals that the proposed-FLANN model outperformed other ML techniques in terns $R^2$ value. Although, the scope of the Deep Learning (DL) network is not included in the current research because such a section is part of a dense layer scheme and can be discussed separately. The DT and EL are similar kinds of ML approaches based on the tree concept. Both techniques are recommended and perform well in classification problems compared to regression. In present research we are working on prediction of casualty which more likely a regression problem because it predicted year wise mortality in different community. Whereas, Decision Trees (DT) or DT-based ensemble learning are recommended for the classification problems. However, DT can predict a continuous outcome. In this context, the decision tree algorithm splits the data based on feature values to predict a numerical target variable. Each leaf node of the tree represents a prediction for the target variable, which is typically the average (or another summary statistic) of the target values of the training instances that reach that leaf. Only thing is we need to give proper command as well as library in Matlab or Python. The accuracy of the DT is determined by the depth of the tree, but it increases the computational complexity. Sagi and Rokach [35] have suggested the application of EL to overcome this issue. However, it may become overfitted when the feature dimensionality increases.

Similarly, in the study by Benjamins et al. [36], the researchers investigate the causes of mortality and inequities among the white and black populations using a statistical approach based on available BCHI data. Bensuic and Cheskin [37] have also examined obesity prevalence using the same data and approach, incorporating socioeconomic indicators, ethnicity, and physical activity. This research offers a platform to apply simple AI-based algorithms like FLANN to estimate mortality rates in various illnesses among different ethnicities, rather than relying solely on conventional statistical techniques. Table 4 showcase the accuracy level of proposed FLANN technique in various clinical problems where it can be utilized not only in the prediction but also provides an automation in the classification of health related issues like diabetes, liver problems, and breast cancer. The accuracy level is assessed on the basis of $R^2$ values. The proposed model demonstrates exceptional performance, achieving the highest accuracy of 95.15% when applied to the diabetes dataset and a still commendable 78.09% in the classification of breast cancer. These results underscore the significance of data clarity as a primary determinant of the efficacy of ML algorithms. Also, selecting the most appropriate model for a particular task is crucial for achieving optimal performance in ML applications. For instance, a DT-based model may perform well on a dataset with complex nonlinear relationships, while a LR model may perform better on data with a clear linear structure. Similarly, DL models image or sequential data, respectively [38]. Additionally, factors such as feature engineering, hyperparameter tuning, and the quality and quantity of data also play significant roles in determining the performance of a machine learning algorithm.

Researchers in [43] showcase the mathematical view of the computational complexity of any multilayer neural network scheme. Table 5 compares the computational complexities of the proposed FLANN and conventional multilayer neural networks such as W-ANN and T-ANN. Along with lower mathematical operations, the proposed network is computationally less complex due to limited layers 'L', nodes 'n',

**TABLE 3. Performance of implemented various ML based techniques.**

| Metrics / ML-Techniques | RMSE | $R^2$ | MSE | MAE |
|---|---|---|---|---|
| DT | 4.807 | 0.879 | 23.06 | 3.287 |
| SVM | 8.413 | 0.618 | 70.789 | 6.560 |
| LR | 5.957 | 0.808 | 35.496 | 4.574 |
| EL | 4.255 | 0.901 | 18.11 | 2.857 |
| Wide ANN | 4.045 | 0.908 | 16.367 | 2.608 |
| Tri-Layered ANN | 4.490 | 0.890 | 20.168 | 2.780 |
| **Proposed FLANN** | **4.059** | **0.915** | **16.476** | **2.693** |

**TABLE 4. Performance of Proposed-FLANN in different clinical condition.**

| Dataset | Health Issues | Features/Instance | Accuracy Level (%) |
|---|---|---|---|
| PIMA [39] | Diabetes | 9 - Attributes<br>768 - Instances | 95.155 |
| BUPA [40] | Liver | 7 - Attributes<br>345 - Instances | 87.31 |
| ECOLI [41] | Protein Localization | 7 - Attributes<br>336 - Instances | 92.5 |
| LYMPHOGRAPHY [42] | Breast-Cancer and Primary-Tumor | 19 - Attributes<br>148 - Instances | 78.09 |

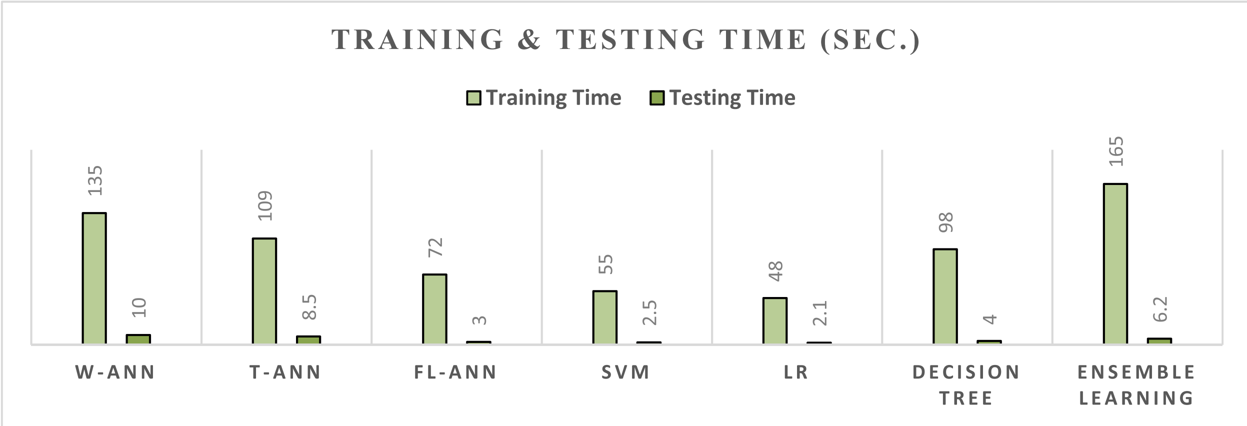



**FIGURE 7.** Training and testing time of ML based Schemes.

**TABLE 5.** Comparison of computational complexity between FLANN & Multilayer-ANN.

| Mathematical Operations | W-ANN /T-ANN | FLANN |
|---|---|---|
| Addition | $3\sum_{i=0}^{L-1} n_i n_{i+1} + 3n_L - n_0 n_1$ | $2n_1(n_0+1)+n_1$ |
| Multiplication | $\sum_{i=0}^{L-1} n_i n_{i+1} + \sum_{i=0}^{L-1} n_i - n_0 n_1 + 2n_L$ | $3n_1(n_0+1)+n_0$ |
| sigmoid(.) | $\sum_{i=1}^{L} n_i$ | $n_1$ |
| exp(.) | ... | $n_0$ |

and mathematically simpler operations utilized compared to the competitive network. The sizes of 'L' in FLANN, W-ANN, and T-ANN are 0, 100, and 3, respectively which shows the simplicity of the proposed algorithm. Figure 7 demonstrates the training and testing times of the various ML-based techniques. Here, LR takes the minimum training time of 48 seconds, whereas EL will take the maximum of 165 seconds. Similarly, FLANN will take 72 seconds to complete the training, which is the minimum for other implemented neural networks like W-ANN and T-ANN. Although, accuracy is almost as high as with the other multilayer neural network.

This research demonstrates that causality can be determined in advance using a simple FLANN-based neural network. However, the maximum level of accuracy attained by the proposed model is 91.5%, limited by our dataset spanning a period of 10 years. Furthermore, DL-based networks require further testing in future research to effectively address temporal data challenges. Additionally, the FSRNCA method requires significant time for regularization of the '$\lambda$' parameter, taking 330 seconds for feature selection, whereas the proposed FLANN model completes training in 72 seconds. This discrepancy underscores the time-intensive nature of feature selection compared to model training. Therefore, it is imperative to explore alternative feature selection techniques such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Embedded techniques in future investigations. Similarly, any issue like classification with imbalanced data will be also addressed with the help of Random Upper Sampling technique like RUSBoost technique. The proposed network is also evaluated on the basis of quantitative metrics such as training time, computational complexity and potentials to classify various other health related issues. Moreover, future research endeavors could incorporate additional health condition parameters into the current framework. Such expansions would enable a comprehensive assessment of disease impact within specific communities.

## V. CONCLUSION

This research highlights the utilization of the FLANN neural network as a predictive model for estimating mortality rates in various communities based on publicly available incorporating socioeconomic indicators. The proposed model demonstrates its capability to predict the risk of diabetes, heart disease, and drug overdose in diverse communities. Leveraging a regularized statistical method such as FSRNCA enhances the model's performance and mitigates overfitting issues. The proposed technique is also successfully tested with open source datasets for the classification of diabetes, liver disease, protein localization, and breast cancer. Future endeavors will extend the scope to include additional diseases within various ethnic groups by applying the BCHI inventory. Furthermore, the integration of other available techniques, such as wrapping or embedding, into the proposed work aims to minimize feature selection timing and enhance overall efficiency.

## REFERENCES

[1] S. Cohen, "Interconnected sensor networks and digital urban governance in data-driven smart sustainable cities," *Geopolitics, History, Int. Relations*, vol. 13, no. 1, pp. 31–41, 2021, doi: 10.22381/GHIR13120213.

[2] A. R. Javed, L. G. Fahad, A. A. Farhan, S. Abbas, G. Srivastava, R. M. Parizi, and M. S. Khan, "Automated cognitive health assessment in smart homes using machine learning," *Sustain. Cities Soc.*, vol. 65, Feb. 2021, Art. no. 102572, doi: 10.1016/j.scs.2020.102572.

[3] M. Javaid, A. Haleem, R. P. Singh, R. Suman, and S. Rab, "Significance of machine learning in healthcare: Features, pillars and applications," *Int. J. Intell. Netw.*, vol. 3, pp. 58–73, May 2022, doi: 10.1016/j.ijin.2022.05.002.

[4] P. Arora, H. Kumar, and B. K. Panigrahi, "Prediction and analysis of COVID-19 positive cases using deep learning models: A descriptive case study of India," *Chaos, Solitons Fractals*, vol. 139, Oct. 2020, Art. no. 110017, doi: 10.1016/j.chaos.2020.110017.

[5] F. Hussain, S. G. Abbas, G. A. Shah, I. M. Pires, U. U. Fayyaz, F. Shahzad, N. M. Garcia, and E. Zdravevski, "A framework for malicious traffic detection in IoT healthcare environment," *Sensors*, vol. 21, no. 9, p. 3025, Apr. 2021, doi: 10.3390/s21093025.

[6] O. P. Choudhary, P. Choudhary, and I. Singh, "India's COVID-19 vaccination drive: Key challenges and resolutions," *Lancet Infectious Diseases*, vol. 21, no. 11, pp. 1483–1484, Nov. 2021, doi: 10.1016/s1473-3099(21)00567-3.

[7] (2014). *World's Population Increasingly Urban With More Than Half Living in Urban Areas*. Accessed: Aug. 11, 2022. [Online]. Available: https://www.un.org/en/development/desa/news/population/world-urbanization-prospects-2014.html

[8] M. Jahangir, H. Afzal, M. Ahmed, K. Khurshid, and R. Nawaz, "An expert system for diabetes prediction using auto tuned multi-layer perceptron," in *Proc. Intell. Syst. Conf. (IntelliSys)*, Sep. 2017, pp. 722–728, doi: 10.1109/IntelliSys.2017.8324209.

[9] S. Ghafouri-Fard, M. Taheri, M. D. Omrani, A. Daaee, and H. Mohammad-Rahimi, "Application of artificial neural network for prediction of risk of multiple sclerosis based on single nucleotide polymorphism genotypes," *J. Mol. Neurosci.*, vol. 70, no. 7, pp. 1081–1087, Jul. 2020, doi: 10.1007/s12031-020-01514-x.

[10] D. Liu, K. Fox, G. Weber, and T. Miller, "Confederated learning in healthcare: Training machine learning models using disconnected data separated by individual, data type and identity for large-scale health system intelligence," *J. Biomed. Informat.*, vol. 134, Oct. 2022, Art. no. 104151, doi: 10.1016/j.jbi.2022.104151.

[11] L. J. Muhammad and E. A. Algehyne, "Fuzzy based expert system for diagnosis of coronary artery disease in Nigeria," *Health Technol.*, vol. 11, no. 2, pp. 319–329, Mar. 2021, doi: 10.1007/s12553-021-00531-z.

[12] S. Dutta, P. Bose, V. Goyal, and S. K. Bandyopadhyay, "Machine learning techniques and syntactic pattern recognition based heart disease prediction for smart health," *Int. J. Med. Res. Health Sci.*, vol. 10, no. 7, pp. 43–59, 2021, doi: 10.20944/preprints202106.0698.v1.

[13] A. K. Verma and R. Aggarwal, "Repurposing potential of FDA-approved and investigational drugs for COVID-19 targeting SARS-CoV-2 spike and main protease and validation by machine learning algorithm," *Chem. Biol. Drug Design*, vol. 97, no. 4, pp. 836–853, Apr. 2021, doi: 10.1111/cbdd.13812.

[14] S. K. Jangir, N. Joshi, M. Kumar, D. K. Choubey, S. Singh, and M. Verma, "Functional link convolutional neural network for the classification of diabetes mellitus," *Int. J. Numer. Methods Biomed. Eng.*, vol. 37, no. 8, p. 1, Aug. 2021, doi: 10.1002/cnm.3496.

[15] E. Lotfi and A. A. Rezaee, "A competitive functional link artificial neural network as a universal approximator," *Soft Comput.*, vol. 22, no. 14, pp. 4613–4625, Jul. 2018, doi: 10.1007/s00500-017-2644-1.

[16] S. S. Roy and S. Barman, "A non-invasive cancer gene detection technique using FLANN based adaptive filter," *Microsyst. Technol.*, vol. 27, no. 2, pp. 463–478, Feb. 2021, doi: 10.1007/s00542-018-4036-6.

[17] S. Singh, S. K. Jangir, M. Kumar, M. Verma, S. Kumar, T. S. Walia, and S. M. M. Kamal, "Feature importance score-based functional link artificial neural networks for breast cancer classification," *BioMed Res. Int.*, vol. 2022, pp. 1–8, Apr. 2022, doi: 10.1155/2022/2696916.

[18] N. Lyons and G. Lăzăroiu, "Addressing the COVID-19 crisis by harnessing Internet of Things sensors and machine learning algorithms in data-driven smart sustainable cities," *Geopolitics, Hist., Int. Relations*, vol. 12, no. 2, pp. 65–71, 2020, doi: 10.22381/GHIR12220209.

[19] T. M. Ghazal, M. K. Hasan, M. T. Alshurideh, H. M. Alzoubi, M. Ahmad, S. S. Akbar, B. Al Kurdi, and I. A. Akour, "IoT for smart cities: Machine learning approaches in smart healthcare—A review," *Future Internet*, vol. 13, no. 8, p. 218, Aug. 2021, doi: 10.3390/fi13080218.

[20] S.-H. Wang, V. V. Govindaraj, J. M. Górriz, X. Zhang, and Y.-D. Zhang, "COVID-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network," *Inf. Fusion*, vol. 67, pp. 208–229, Mar. 2021, doi: 10.1016/j.inffus.2020.10.004.

[21] Z. Rayan, M. Alfonse, and A.-B.-M. Salem, "Machine learning approaches in smart health," *Proc. Comput. Sci.*, vol. 154, pp. 361–368, Jan. 2019, doi: 10.1016/j.procs.2019.06.052.

[22] S. Mohapatra, D. Patra, S. Kumar, and S. Satpathy, "Lymphocyte image segmentation using functional link neural architecture for acute leukemia detection," *Biomed. Eng. Lett.*, vol. 2, no. 2, pp. 100–110, Jun. 2012, doi: 10.1007/s13534-012-0056-9.

[23] P. P. Ray, "A review on TinyML: State-of-the-art and prospects," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1595–1623, Apr. 2022, doi: 10.1016/j.jksuci.2021.11.019.

[24] Big Cities Health Inventory. (2022). *Health Data*. Accessed: Sep. 11, 2022. [Online]. Available: https://bigcitieshealthdata.org/

[25] A. Narin, Y. Isler, M. Ozer, and M. Perc, "Early prediction of paroxysmal atrial fibrillation based on short-term heart rate variability," *Phys. A, Stat. Mech. Appl.*, vol. 509, pp. 56–65, Nov. 2018, doi: 10.1016/j.physa.2018.06.022.

[26] T. Pratap and P. Kokil, "Computer-aided diagnosis of cataract using deep transfer learning," *Biomed. Signal Process. Control*, vol. 53, Aug. 2019, Art. no. 101533, doi: 10.1016/j.bspc.2019.04.010.

[27] S. Tuli, N. Basumatary, S. S. Gill, M. Kahani, R. C. Arya, G. S. Wander, and R. Buyya, "HealthFog: An ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IoT and fog computing environments," *Future Gener. Comput. Syst.*, vol. 104, pp. 187–200, Mar. 2020, doi: 10.1016/j.future.2019.10.043.

[28] T. Hu, M. Khishe, M. Mohammadi, G.-R. Parvizi, S. H. T. Karim, and T. A. Rashid, "Real-time COVID-19 diagnosis from X-ray images using deep CNN and extreme learning machines stabilized by chimp optimization algorithm," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102764, doi: 10.1016/j.bspc.2021.102764.

[29] H. Chen, S. Khan, B. Kou, S. Nazir, W. Liu, and A. Hussain, "A smart machine learning model for the detection of brain hemorrhage diagnosis based Internet of Things in smart cities," *Complexity*, vol. 2020, pp. 1–10, Sep. 2020, doi: 10.1155/2020/3047869.

[30] H. Wang, Y. Lv, H. Chen, Y. Li, Y. Zhang, and Z. Lu, "Smart pathological brain detection system by predator-prey particle swarm optimization and single-hidden layer neural-network," *Multimedia Tools Appl.*, vol. 77, no. 3, pp. 3871–3885, Feb. 2018, doi: 10.1007/s11042-016-4242-0.

[31] Y.-D. Zhang, S. C. Satapathy, D. S. Guttery, J. M. Górriz, and S.-H. Wang, "Improved breast cancer classification through combining graph convolutional network and convolutional neural network," *Inf. Process. Manage.*, vol. 58, no. 2, Mar. 2021, Art. no. 102439, doi: 10.1016/j.ipm.2020.102439.

[32] B. Majhi and D. Naidu, "Pan evaporation modeling in different agroclimatic zones using functional link artificial neural network," *Inf. Process. Agricult.*, vol. 8, no. 1, pp. 134–147, Mar. 2021, doi: 10.1016/j.inpa.2020.02.007.

[33] W. Yang, K. Wang, and W. Zuo, "Neighborhood component feature selection for high-dimensional data," *J. Comput.*, vol. 7, no. 1, pp. 162–168, Jan. 2012, doi: 10.4304/jcp.7.1.161-168.

[34] M. Kumar, S. K. Mishra, S. K. Choubey, S. S. Tripathy, D. K. Choubey, and D. Das, "Cat swarm optimization based functional link multilayer perceptron for suppression of Gaussian and impulse noise from computed tomography images," *Current Med. Imag. Formerly Current Med. Imag. Rev.*, vol. 16, no. 4, pp. 329–339, May 2020, doi: 10.2174/1573405614666180903115336.

[35] O. Sagi and L. Rokach, "Ensemble learning: A survey," *WIREs Data Mining Knowl. Discovery*, vol. 8, no. 4, pp. 1–18, Jul. 2018, doi: 10.1002/widm.1249.

[36] M. R. Benjamins, A. Silva, N. S. Saiyed, and F. G. De Maio, "Comparison of all-cause mortality rates and inequities between black and white populations across the 30 most populous U.S. cities," *J. Amer. Med. Assoc. Netw. Open*, vol. 4, no. 1, Jan. 2021, Art. no. e2032086, doi: 10.1001/jamanetworkopen.2020.32086.

[37] M. Benusic and L. J. Cheskin, "Obesity prevalence in large U.S. cities: Association with socioeconomic indicators, race/ethnicity and physical activity," *J. Public Health*, vol. 43, no. 1, pp. 148–154, Apr. 2021, doi: 10.1093/pubmed/fdz077.

[38] K. Thilagam, A. Beno, M. V. Lakshmi, C. B. Wilfred, S. M. George, M. Karthikeyan, V. Peroumal, C. Ramesh, and P. Karunakaran, "Secure IoT healthcare architecture with deep learning-based access control system," *J. Nanomaterials*, vol. 2022, pp. 1–8, May 2022, doi: 10.1155/2022/2638613.

[39] C. Bartley, 2016, "Replication data for: Pima Indians diabetes," *Harvard Dataverse*, doi: 10.7910/DVN/XFOZQR.

[40] Bupa Medical Research Ltd. *BUPA Liver Disorders Dataset*. UCI Mach. Learn. Repository. Accessed: Jun. 10, 2024. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Liver+Disorders

[41] K. Nakai and M. Kanehis, "ECOLI data set," UCI Mach. Learn. Repository. Accessed: Jun. 10, 2024. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Ecoli

[42] M. Zwitter and M. Soklic, "Lymphography," UCI Mach. Learn. Repository. Accessed: Jun. 10, 2024. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Lymphography

[43] J. C. Patra, R. N. Pal, R. Baliarsingh, and G. Panda, "Nonlinear channel equalization for QAM signal constellation using artificial neural networks," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 29, no. 2, pp. 262–271, Apr. 1999, doi: 10.1109/3477.752798.

**MANISH KUMAR** received the B.Tech. degree in applied electronics and instrumentation engineering from the Biju Patnaik University of Technology, Rourkela, India, in 2010, the M.Tech. degree in biomedical engineering from Manipal University, Udupi, India, in 2013, and the Ph.D. degree in electrical and electronics engineering from the Birla Institute of Technology, Ranchi, India, in 2018. He is currently a Research Professor with the Department of Computer Science and Engineering, Seoul National University of Science and Technology, Seoul, South Korea. He is working on an AI-based IoMT security solution. He is the author of more than 50 research articles, published a book on IoT, and received a patent for a virtual trial room, in May 2024. His research interests include machine and deep learning algorithm design and implementation.

**SUSHIL KUMAR SINGH** (Member, IEEE) received the M.E. degree in information technology from Karnataka State Open University, Mysore, India, the M.Tech. degree in computer science and engineering from Uttarakhand Technical University, Dehradun, India, and the Ph.D. degree from Seoul National University of Science and Technology, Seoul, South Korea. He is currently an Associate Professor with the Department of Computer Engineering, Marwadi University, Rajkot, Gujarat, India. He has published four books, such as *Computer C Programming*, *Cyber Security*, *Big Data Analytics*, and *Mobile Computing*. He has also published research papers in international journals and conferences. His current research interests include blockchain, artificial intelligence, big data, the Internet of Things, smart city security, and cyber-physical systems. He has delivered international lectures in many countries.

**SUNGGON KIM** received the B.S. degree from the Department of Computer Science, University of Wisconsin–Madison (UW-Madison), Madison, WI, USA, in 2015, and the Ph.D. degree from the Department of Computer Science and Engineering, Seoul National University (SNU), Seoul, South Korea, in 2021. He was an Intern with the Scientific Data Management Research Group (SDM Group), Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, USA, in 2018, 2019, and 2020. He was also a Postdoctoral Researcher with SNU, from 2021 to 2022. He is currently an Assistant Professor with the Department of Computer Science and Engineering, Seoul National University of Science and Technology (SeoulTech). He is also working on health and the IoT-related security. His research interests include high-performance computing (HPC), big data, file systems, virtualization, cloud computing, multi-core systems, GPU, database systems, and operating systems.