Automated outbreak prediction of epidemic diseases using Machine Learning based Global pre-emptive scheme

Ravi Kumar Suggala[1], Dr.SangramKeshari Swain[2], Dr.M. Vamsi Krishna[3]

[1]Research Scholar, Dept.of.CSE, Centurion University of Technology and Management,Odisha, India.
[2]Dept. of CSE, Centurion University of Technology and Management, Odisha, India
[3]Aditya Engineering College (A), Andhra Pradesh, India

Corresponding author: Ravi Kumar Suggala, ravi.suggala@gmail.com

Abstract. The most effective strategy to combat global disease spread is to prevent epidemics. As a result, develop a wearable safety and health monitoring device that utilizes the Internet of Things to provide authentic health testing and improve global health. However, because epidemics are nonlinear and dynamic systems, establishing preventive techniques remains a difficult task. IoT data integration is a major concern for disease prediction. To deal with this situation, a new classification and regression model (CART) has been presented to forecast data point labels. It overcomes scaling, outliers, and missing data information by traversing a binary tree according to the kernel to get the relevant labels. When the data are successfully integrated, there is a problem with noise and demographic bias in the prediction of the disease. Hence a novel Imbedded Least Square Support Regression (ILS-SVR) reduces random noise and removes drifts. In addition, to mitigate the demographic distortion of an unexpected regression problem, the theory by-law function (TBF) has been implemented. Furthermore, it helps to predict the epidemic disease better and for prevention strategy among the people, human mobility plays a significant concern. In order to tackle human mobility, a novel IoT-based global Preemptive scheme has been introduced to track the pattern and the harmful symptoms of infectious patients observed. This explores the possible role of K-NN machine learning techniques as a Hidden Markov Model (HMM). Thus automated alert systems prevent mortality, morbidity rate, and timely detection of epidemic diseases with high efficiency.

Keywords: Epidemics, disease Outbreak, Prediction, Prevention, Classification and regression model, COVID-19, Hidden Markov Model (HMM), Machine Learning,KNN.

Nomenclature list

Parameter	Explanation
CART	Classification and regression model
ILS-SVR	Imbedded Least Square Support Regression
TBF	Theory by-law function
IoT	Internet of Things
HMM	Hidden Markov Model
K-NN	K-Nearest Neighbor
COVID-19	Corona Virus Disease 19
SARS	Severe Acute Respiratory Syndrome
AIDS	Acquired Immuno-Deficiency Syndrome
WHO	World Health Organization
IoT-EHMS	IoT-Enabled Health Monitoring Systems
H-IoT	Healthcare-related Internet of Things
M2M	Machine-to-Machine

SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus2
MC	Monte Carlo
MMC	Microscopic Markov Chain
SHSM	Student Healthcare Surveillance Model
PLDB	Perceived level of noise
FBS	Fasting Blood Sugar Test
CBC	Complete Blood Count
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
ca	Axis aligned cuboid region
P	Input vector for CART
R	Output vector
p by q	Predicted output vector
1(ri=c)	Indicator function
pt	pattern
Cl(l=1,2,…,k)	classes
Xif	Inspected set
Lf	Lower fence
Uf	Upper fence
O	Outlier region
S	Higher dimensional space
Y	Input vector for ILS SVR
W	weight
B	bias
G(y)	Regression Function
Ψ(s)	Expanded Nonlinear Mapping
Γ	Regularization parameter
LG(V, e ,β)	Lagrangian Function
βj€I(j=1,2,.,l)	Lagrange
X= Ώ	Kernel matrix
K(sj, sn)	Kernel function
U	Unit matrix
M	Length of mobility Pattern
Δ	Minimum Support threshold
CMp	Candidate mobility pattern
sup(s)	Support value
L	length
D	Euclidean distance
f(n)	Estimated cost of cheapest solution
g(n)	Cost of best path
h(n)	Admissible heuristic
T	time
T	Number of prediction
Pt	Predicted value at time
Xt	Observed Value at time

1 Introduction

Health monitoring is a crucial and valuable tool for diagnosis and prevention of diseases of whatever kind [1]. Machine learning has become one of the researchers' top priorities in healthcare. In the area of applying machine learning algorithms in healthcare, there have been several experiments and these include data mining and statistics [2]. Effective decisions on diagnosis and treatment options for patients can be made by applying the appropriate machine learning methodology. In recent years, the prevention of infectious disease outbreaks has played significant role in monitoring healthcare [3]. Recall SARS outbreak in the early 2000s are waiting for COVID-19 right now, perhaps hearing the words epidemic and pandemic [4]. The epidemiology term of epidemic denotes "the occurrence in a given human population with new cases of a particular disease over a given period at a rate significantly above the estimated level based on earlier observation [6]." An epidemic's impact will spread quickly from place to place, or even from a group of nations [7]. Some of the ongoing epidemics are Acquired Immuno Deficiency Syndrome, infectious diseases, and recent epidemics are dengue fever, malaria, cholera, COVID-19, etc. The outbreak is when the epidemic becomes a pandemic, an outbreak of disease that occurs through nations or continents [8].

COVID-19 is proclaimed as a global epidemic by World Health Organization (WHO) which became evident that the epidemic was serious and was spreading rapidly through a wide scale [9]. Various countries around the world were plagued by outbreaks of epidemic disease which threatened the general public's health. In addition, disease outbreaks can cause significant socio-economic and political disruption to the point of disrupting often thriving societies. To prevent the outbreak of this epidemic disease, research studies manage machine learning algorithms for disease detection and prevention [10]. Machine learning is a data-based analytical method in which knowledge discovery is analyzed by gathering data from process of data mining and processing it via the data pipeline process. The mined data may be described, predicted, and controlled using machine learning [11]. In order to generate accurate and dependable outcomes, this learning strategy will increase pattern recognition, data classification, and feature selection. It requires the help of machine learning to become a perfectly aligned, human error-tolerant system. The system's predictive element will determine the correct input to get the predicted output.The Internet of Things, a list of recent technologies, is gaining global attention and being more commonly accessible for forecasting, avoiding, and tracking arising communicable diseases [13]. The Internet of Things is a group of networks digital sensors, detectors, and people that enables raw information to be collected and accessed through the internet for analyzing patterns or fashions. Wearable health-monitoring devices, cloud-based remote health testing, and artificial intelligence are all used to enable real-time surveillance in cutting-edge IoT-Enabled Health Monitoring Systems [14]. In real time, these monitoring systems use social networks, public information, and clinical information [14]. A more critical assessment of IoT's surveillance applications is required, as well as a more in-depth examination of the privacy problems that its usage raises [15]. In calculation to web-based surveillance parts, event-based IoT tracing gathers and distributes rare information from a variety of casual origins (newscast, social media posts, and internet queries) in an effort to discover occurrences with possible pandemic transmit sooner than more traditional methods. As a result, advancements in transferrable syndrome modelling, parasite detection, and testing have been established.

Considering the prevailing world-market, IoT-based shrewd infection monitoring schemes, rather than web-based surveillance, have the potential to be a significant advance in efforts to contain the present endemic [17]. Because most of the infrastructure (that is,smartphones, wearable devices, and internet connectivity) is already in place, the role that technology may play in restricting the spread of the pandemic is limited to data collecting and analysis [18]. Correspondingly, use a similar technique to study the influence of human mobility in forecasting a person's physical health conditions [19]. But, a person's mobility behaviorcannot convey information about her or his conditions of health. The influence of people's mobility habits on the future appearance of flu-like and cold symptoms (that is, fever, sore throat, malaise, cold, shortness of breath, headache, muscular pain andcough) was investigated in this research [20]. Healthcare-relatedInternet of Things (hereafter: H-IoT) assures to improve healthcare by providing insights of data-driven. As a consequence, the importance of H-IoT data has become clear. Specifically, the prevalent problem is “the delay in providing results and generating alerts due to data loss, buffering, network communication, monitoring, or processing”. They focused on hardware and software advancements to address the machine-to-machine (M2M) communication portion of H-IoT data challenges. As a consequence, they only shelter partiallight to identify datachallenges of H-IoT when it comes to aspects of human-to-human and human-to-machine communication [20].Thus, to initiate an effective preventive mechanism, it is must to handle the huge data integration issues experienced with the existing methodologies to guarantee précised prediction. This paper aims to propose a solution for major issues such as combining data residing at different sources and because of high sensitivity of data within these collaborative domains provokes data integration. Naive Bayes classifier outperforms in dengue leads to limitation in noise and bias. Hence to tackle the above issue, a novel technique is proposed which provides the main contribution of this paper as follows,

- ❖ Classification and regression (CART) model to reduce the data integration problem.
- ❖ Imbedded least squares support vector regression (ILS-SVR) to minimize random noise and detach data drift.
- ❖ IoT based global pre-emptive scheme to identify and detect the disease. Thus, in this strategy, using these novel techniques alert the patients whoever affected by the epidemic diseases.

The rest of the paper is organized as follows: Section 2 describes the literature survey, and Section 3 describes the proposed methodology for preventing epidemic disease outbreaks. Section 4 describes the results of the analysis and discussion. Section 5 summarizes and concludes this paper.

2 Literature Survey

Yousefpouret al. [21] discussed that recognizing the early patterns of disease spread and assessment of the feasibility of control policies play a critical role in infectious disease prevention. This paper deals with the conception of optimum management methods for the new coronavirus epidemic (COVID-19). It is considered a statistical model for coronaviruses causing severe acute respiratory syndrome2 are being transmitted in Wuhan (SARS-CoV-2).Good quality schedules for different factor rates including touch intensity and transfer rate between the symptom of infected persons in the quarantined infected classes are recommended in order to addressed the problem quickly, efficiently and effectively proposed a multi objective genetic algorithm. Two best strategies have been built effectively by modifying these conditions. This research involves groundbreaking analysis that has recommended COVID-19 policies and also covers COVID-19 and its economic effects with the help of multi-objective evolutionary algorithm. It shows that policymakers could discover useful and realistic ways to manage epidemic by optimal strategy implementation suggested. However, existing parameters, namely the number of isolated unvaccinated interactions are out into society, cannot be considered control measures for outbreak prevention.

Wang et al [22] analyzed that now the worldwide effect of COVID-19 is enormous with more than 8 million people have been contaminated in more than 100 countries. A number of countries have released protection initiatives in order to eliminate its spread. It is not known, however, when the outbreak will stop in different countries and internationally. Predicting COVID-19 patterns is a task of extreme significance. This integration into a logistic model for the epidemic trend limit and the implementation of the cover rate into the FbProphet model, an epidemic time-series machine-learning system to draw an epic curve, forecast an epidemic pattern, the most revised epidemiological data from COVID-19 before 16 June 2020. The modelling findings for Brazil, Russia, India, Peru and Indonesia are presented in three relevant fields. The worldwide epidemic is predicted to peak at the end of October, with 14.12 million cumulatively contaminated people under the statistical forecast. However, the impact of input cases and spatial control among countries have not recognized in certain countries when predicting the epidemic.

Chen et.al [23] projected a new epidemic technique focused on multi-layer multiplex grids to explore effect on infection transmission from positive and negative prevention knowledge. The one layer of this model reflects a social network with a competitive distribution of positive and negative prevention data while the other signifies the disease physical communication network. People are aware of favorable preventative measures take more successful steps to avoid being poisoned than those who are aware of negative preventative measures. The pandemic limit for the suggested pandemic model is analytically represented using the Microscopically Markov chain method. This suggested that there would be a substantial effect on the epidemic threshold on the distribution of optimistic and pessimistic preventive data and on the topology of the physical communication system. It has been shown that the performance observed with MMC and Monte Carlo (MC) simulations agree, but MMC is able to explain the dynamics of the model that has been suggested. The effect of positive and negative prevention evidence on threshold of infection and infectious diseases prevalence is illustrated by detailed simulations in the meantime. The investigation revealed that the dissemination of positive preventive information will prevent disease prevalence and epidemic outbreaks and that the spread of negative preventative information is facilitated. However, in this method, the accuracy is not evaluated whereas outbreak of epidemic cannot be stopped through promoting diffusion along with minimized the bias.

Alireza souri et al [24] implemented IoT-based student health-care monitoring model to monitor pupils'health care constantly as well as to identify changes in biological as well as behavioral through smart healthcare technology. They concentrate on vast geographical areas that enable students to monitor their health condition status. Using machine learning approaches, vital data is collected through IoT devices in this model and data analysis can be performed to predict potential dangerous of bodily and interactivevariations of a pupil. The help vector machine has obtained highest precision about 99.1 percent after testing the proposed model, which is a good outcome for the purpose.

Using machine learning and nudge theory, Daitaro Misawa et al [25] describe a new case in which the uptake rate for examinations of colorectal cancer was dramatically raised. It deals how machine learning may help patients withhealth consciousness or the general population is important for behavioral changes to occur. The article also discusses the utility of social impact bonds (SIBs) as a mechanism for implementing these applications. Using historical data from designated periodical health examinations, digitalized medical insurance receipts, and medical examination records for colorectal cancer, machine learning was used to deduce segments for whom the examination was recommended during a healthcare Social Impact Bond project conducted in the city of Hachioji, Tokyo.But, the diagnostic step, that is, discovering target populations, has received limited attention.

Wu et.al [26] presented that the overall amount of infected cases reported for China, 29 regions in China, and 33 nations around the globe that are and have been suffering large outbreaks should be calibrated according to logistical growth model, model of generalized logistical growth, generalizing Richard model and wide spreadened growth model. It discloses epidemics evolution in China and the effects both in aggregates and in each province of dramatic control steps. This paper describes four stages of the epidemic in China quantitatively, including a thorough study of the provincial hetero genes. With some instructive differences across regions, the drastic containment measures introduced by China were very successful. Based on China's past, the epidemic has grown in other countries with scenarios expected. It finds that outbreaks had stopped in 14 countries (mostly in Western Europe), with some countries experiencing a resurgence of cases. In contrast for many regions in China, in which the after-peak path is marked by a much faster collapse, the modelling effects in Western countries are clearly longer. This identified three classes of countries at various stages of breakthrough advancement and has details on the latest global pandemic. However, in this method, the lower bound for accurate prediction of disease images are not anticipated.

In [21] Other rate variables that cannot be recognized in quarantined non-infected contacts for prevention of outbreak, and in [22] the impact of input cases and geographic control between nations is not acknowledged when predict the disease similarly [23] precision is not measured and the spread of the disease can't be prevented along with reduces the bias [24] sensitivity was high for integration, [25] not mentioned the effect in the selection of subjects, likewise has limited attention to diagnosing patients and [26] the lower hurdle is not forecasted for the exact prediction of disease images. From the aforementioned issues, to reduce the danger of continuous epidemics and the chance of nativeeruptions evolving into a worldwide pandemic, fast and effective public health actions must be done in response.

3 Prevention of Epidemic Disease Outbreak

The Internet of Things (IoT) is gaining traction throughout the world, and it is becoming more accessible for predicting, preventing, and monitoring developing infectious diseases. It may help in the rapid diagnosis of infected individuals as well as the accurate prediction of disease dissemination to other sites and prevent the outbreak of diseases. However, since epidemics are non-linear and complicated processes, developing preventative techniques remains a difficult task. In IoT data integration is a major concern. Fig1 shows the architecture of the epidemic disease outbreak prevention.

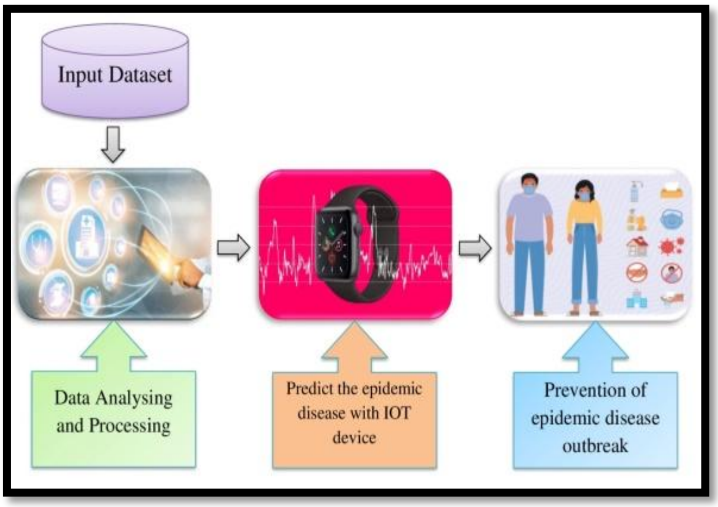

Fig. 1. Architecture of epidemic disease prevention

Data processing and analysis are complicated by the scale and variety of devices, communication networks, and protocols used in data collecting. This leads to data integration, which is defined as the challenge of merging information from several sources as well as giving a unified representation of the data to the user. Due to the high sensitivity of data in such collaborative domains, it is vital to create data integration solutions.Prevailing techniques use many machine learning algorithms to overcome data integration problem with decreased flexibility. Similarly, for Bayesian methods it is hard to assume proper class conditional distributions in each individual class; it makes the process difficult to find the required information. This would simultaneously affect the prediction accuracy due to variable scaling, missing values and outliers.

To fix these challenges, prior prediction methodologies make consume of records collected from media platforms for initial outbreaks of infectious diseases forecasting. Considering prior Naive Bayes classifier, it outperforms dengue suspected tweets however; the classifier as well as other existing methodologies had the limitations due to varied noise and demographic bias. Thus, the challenges in prediction would create a severe impact on the prevention strategies. While dealing with this preventive measure, human mobility plays a major role in instigating severe infections. Their patterns are taken into account by the mobility sensing program, which adjusts GPS sampling appropriately and avoids duplicate position fixes to conserve energy. Periodic sampling, on the other hand, ignores the diversity of actual human mobility [27]. Previous research has shown that human mobility is an effective proxy to predict mental health of people with the help of mobile phones. However, if applications have loading issues or negative effects it offers false information to the user [28]. A general-purpose open-system SEIR model simulator is designed to prevent the spread of epidemics in China. However, it only gives the preventive measures for specific city. Thus, it is crucial to tackle the data integration concern along with effective prediction to develop amechanicalalsoactiveattentive system for initial and sensible epidemic diseases findingto prevent the mortality and morbidity rate averting worldwide spread. Fig 2 shows proposed work flow of epidemic disease prevention with IoT sensor device.

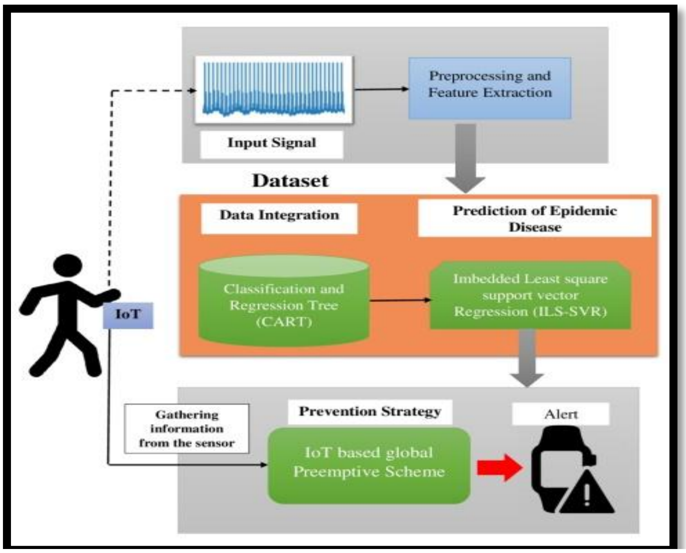

Fig. 2. Proposed diagram for epidemic disease Prevention

Thus the proposed method employed IoT based disease prediction scheme in which first the input signal is preprocessed and the extracted features is coupled with the dataset. In the dataset, to solve data integration problem classification and regression trees model is proposed which is learned from training data. To predict a label for data points, the input vector is split into cube shaped areas. The process of determining the subsequent reference identifier can be defined with the help of traversingspanning tree that corresponds to mechanism of sequential decision making with the kernel process. Besides it resolves the variable scaling, missing values and outlier’s problem. Moreover, the features are selected there is inadequacy in predicting the disease. Thus, in seismic data, to overcome the noise, and demographic bias and the timely prediction the work proposed the use of embedded least squares support vector regression (ILS-SVR) with an unconscious bias period reduces randomness and eliminates drifts. The scheme employs the weight vector and the slanted period into a set of vectors, which is then used to calculate the objective function. Also it incorporates concept regulation function to mitigate the demographic bias of unintended problem with regression mechanism. "Neighbor awareness" is the first factor people should be mindful of the neighbor that is afflicted by epidemic disease since predicting the outbreak. For healthier prevention the work implement IoT based global pre-emptive scheme in which explore the potential role of a Hidden Markov Model (HMM) with K-NN a machine learning technique. HMM detects the mobility pattern after the pattern is detected the KNN validating the effectiveness of epidemic diseases by monitoring the hazardous symptoms of infectious patients such as fever, dry cough, fatigue, trouble breathing, movement, aches and pains to stop pretentious individuals. The following section describe the overall process of the proposed method.

3.1 Classification and Regression trees (CART)

Model of Classification and Regression Trees (CART), which is learned from training results, is proposed. In order to estimate an identifier for the data sets, the feedback area is divided into cuboids. A binary tree that corresponds to a mechanism of a sequential decision-making process with the kernel process can be specified for determination of the subsequent reference label. It also solves the scaling variable, incomplete values and the dilemma of the outlier.

3.1.1 Data integration classifying Process

In CART, the information area is divided into coordinate axes cubical areas c_a, and each region is assigned its own classification or regression method to forecast an identifier for the data sets that fall into that area. A binary tree forwarding equivalent to a sequenced selection process can be used to describe the process of guessing the corresponding target label given a new, unnoticed input sequence (data point) p. A prototype that forecasts a specific class across each area is an example of classification, whereas a framework that forecasts a steady over each area is an example of regression. In the classification task, actually indicate a type parameter by a discrete probability distribution and the predicted class label for. Denote a class variable by a discrete random variable r and the predicted class label for p by q. in the classification task. The classification task is structured as follows:

$$T\left(r=c\mid k\right)=\frac{1}{\left|RE_k\right|}\sum_{i\in RE_k}1\left(r_i=c\right) \tag{1}$$

$$q=\arg\max_{c} T\left(r=c\mid p\right)=\arg\max_{c} T\left(r=c\mid k\right) \tag{2}$$

Additionally, $1\left(r_i=c\right)$ denotes indicator function. This formula denotes p that the matching area will be labelled with the most common (mode) label.

To express the regression task, denote the value of the output vector by r and the predicted output vector p by q. The regressive task can be defined as

$$q=\frac{1}{\left|RE_k\right|}\sum_{i\in RE_k} r_i \tag{3}$$

That is, the production vector for p will remains averageof the production vector of file facts in its equivalentarea.

The arrangement of the diagrammust be defined according to the preparation set while training CART. This entails defining the divided criteria and threshold parameter value for each node. Because finding the structure of optimum tree is an NP-complete problem, CART is trained using a binary decision that develops the tree top-down and picks the finestdivided node-by-node it is shown in Figure 3.

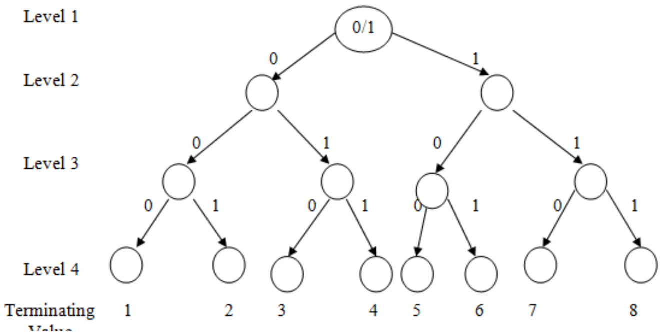

Fig. 3. Binary decision Tree

3.1.2 Decision Tree with Outliers Removal

CART entails identifying and constructing a binary classification tree from an experiment of learning algorithm that has been classified correctly. The number of enterprises in the two sub-groups formed at every binary divide, which correspond to the two branches originating from each access point, decreases over time, necessitating a large training sample to produce acceptable results.

A non-parametric strategy to object recognition can be viewed as binary classification trees for categorization. A decision tree represents the feature space in which patterns pt are allocated to classes $c_l\left(l=1,2,\ldots,k\right)$ based on the result acquired throughbelow mentioned choices taken at a series of vertices where the tree's branches diverge.

CART's decision tree considers just the bands that aid to distinguish the classes, while the rest are ignored. To minimize dimensionality, a tree is used as a feature selection approach in this research. Negligence in dealing with the outlier may have an impact on the splitting point, resulting in bias and an inaccurate tree.

The idea behind this amusing tree approach to identifying and treating outliers is to employ a strategic plan that manages and incorporates outliers while building the tree. It enables practitioners to use the approach without first performing an opening paragraph analysis to clean the records before establishing the tree.

Before performing the winsorize method, the training set can be checked for the presence of outliers using the top and bottom fence posts. From the set of variables. Where X_{if} is the inspected with the help of lower fence L_f and upper fence U_f. The value which is less than L_f or more than U_f can be defined as outlier region O.

$$O=L_f\bigcap U_f \tag{4}$$

If the entity exceeds the threshold in these outlier regions, it prevents dividing and becomes the terminal node. Otherwise, the actual document is used to re-examine the non-terminal network until the optimal cut is found (minimum impurity). The error rate is used to evaluate the finished tree.

Despite the prevalence of missing data, the system investigates simulation-based dataset to deal with missing value, which would be based on filling in (imputing) each or even more believable values for the missing value.

Thus, it reduces the data integration problem with the result of outliers and creates a talltree correctness, allowing consumers to skip the initialization step and address missing values using imputation. In particular, decision tree work will function with different features that are on entirely different scales, or a mix of binary and continuous features.Furthermore, it enables decision-makers to use proofconstructedtreatment to createscientific decisionsin difficult scenarios in finding epidemic diseases.Simultaneously in predicting the variety of epidemic disease there is shortcoming in noise and demographics bias thus which is overcome in the following section.

3.2 Imbedded least squares support vector regression (ILS-SVR)

Imbedded least squares support vector regression with an unconscious bias period to minimize random noise and eliminate seismic data deviation. The proposed solution combines the weight matrix and the slanted period into a set of vectors and incorporates the bias term into the intent function of the LS-SVR model optimal solution. In the case of the translation kernel, this allows for a smoother solution that can be calculated more easily than the Least Squares-Support Vector Regression. More relevantly, the Imbedded Least Squares-Support Vector Regression with a Ricker wavelet operating system not only suppresses random noise effectively, but also eliminates drifts in seismic surveys. The unconscious bias term and the kernel's special option, for example, the Ricker wavelet kernel, delete two primary drifts behind the suggested protocol.

3.2.1 Processing Model

ILS-SVR first expand input vector $y\in I^N$ to a higher dimensional space $s=[y,1]\in I^{M+1}$, and $\left[w,b\right]\in I^{M+1}$ in which w denotes the weight, which can controls the smoothness of the model; b is the bias. The regression function for the imbedded least squares support vector regression (ILS-SVR) can be defined by

$$G(y) = [w,b](\psi(s)) = V(\psi(s)) \tag{5}$$

Where, $\psi(s)$ is expanded nonlinear mapping. Extensive input space is mapped to another space feature. To make optimization a feasible convex optimization problem can be implemented with the slack variables. The optimization problem of ILS-SVR becomes

$$Minimum\left\{\frac{1}{2}\|V\|^2 + \frac{\gamma}{2}\sum_{j=1}^{l} e_i^2\right\} \ s.t \ z_j = V(\psi(s_j)) + e_j \tag{6}$$

Where, $j = 1, \ldots l$, γ is the regularization parameter. The corresponding Lagrangian function is

$$LG(V,e,\beta) = \frac{1}{2}\|V\|^2 + \frac{\gamma}{2}\sum_{j=1}^{l}\beta_j\left[V(\psi(s_j)) + e_j - z_j\right] \tag{7}$$

Where, $\beta_j \in I(j = 1,2,\ldots,l)$ are Lagrange multipliers. From equation (7) it declared

$$\begin{cases} \dfrac{\partial LG}{\partial V} = 0 & V = \sum_{j=1}^{l}\beta_j\psi(s_j) \\ \dfrac{\partial LG}{\partial e_j} = 0 & \beta_j = \gamma e_j, \quad j = 1,2,\ldots,l, \\ \dfrac{\partial LG}{\partial \beta_j} = 0 & V(\psi(s_i)) + e_j - z_j = 0, \quad j = 1,2,\ldots,l \end{cases} \tag{8}$$

Whichshows ILS-SVR model is used for data with noise of nonzero mean. This showspotential of ILS-SVR model to remove the drifts. After eliminating $V \ and \ e_j$ the Lagrangian multipliers is obtained by

$$\beta = \left(X + \frac{U}{\gamma}\right)^{-1} z \tag{9}$$

Where, $\beta = [\beta_1, \beta_2, \ldots \beta_l]^t$ X is a kernel matrix, $X_{j,n} = (\psi(s_j))(\psi(s_n))$, and thus the kernel function is denoted as $K(s_j, s_n)$. Here $j,n = 1,2\ldots,l$ and U is a unit matrix. Thus the regression function is followed as

$$G(y) = V.\psi(s) = \sum_{j=1}^{l}\beta_j(\psi(s_i))(\psi(s_n)) \tag{10}$$

From this regression function it is employed with $s = [y,1]$, the kernel function with RBF kernel is simulated as

$$K(y_j, y_n) = \exp\left(-\frac{\|y_j - y_n\|^2}{2\sigma^2}\right) \tag{11}$$

From the above equation (11) and the kernel matrix,it is obtained that $X_{jn} = K(s_j, s_n)$

$$\exp\left(-\frac{\|[y_j - y_n 1]\|^2}{2\sigma^2}\right) = K(y_j, y_n) = \Omega_{jn} \tag{12}$$

Then equation (9) becomes

$$\beta = \left(X + \frac{U}{\gamma} \right)^{-1} z = \left(\Omega + \frac{U}{\gamma} \right)^{-1} z \tag{13}$$

Where $\Omega = X$ is the same kernel matrix its elements are $\Omega_{jn} = K\left(y_j, y_n\right)\left(j, n = 1, 2, \ldots l\right)$. The nonlinear regression function reduces to

$$G(y) = \sum_{j=1}^{l} \beta_j \left(\psi\left(s_j\right)\right)\left(\psi(s)\right)$$

$$= \sum_{j=1}^{l} \beta_j K\left(s_{j,} s\right)$$

$$= \sum_{j=1}^{l} \beta_j K\left(y_j, y\right) \tag{14}$$

It is worth noting that any translation invariant kernel namely Ricker wavelet kernel, yields the same nonlinear regression function. In the support vector machine, the translation invariant kernel having the property $K(y,s) = K(y-s)$ is commonly employed. With varied drifts,imbedded least squares support vector regression with Ricker wavelet kernel can effectively eliminate the noise and bias term.

However, it is equally important to mitigate the unfair demographic bias. This bias is usually some sort of imbalance in the dataset. Imbalances can come from some demographic groups not having enough data, causing over fit to groups with more data also it can come from unfair representations of demographic attributes. Thus in order to mitigate the unfair bias of demographic bias the work implementsthe theory by-lawfunction process.

3.2.2 Theory By-Law Function (TBF)

Mitigating unintended demographic bias at the learning algorithm level allows one greater interpretability into machine learning itself. In this developed a mitigation technique called Concept Regularization and applied it to a realistic toxicity prediction system and overfitting. Though theory by-law caused the logistic regression algorithm to contain less unintended bias. TBF can be applied to any scenario where one would like predicted concepts from the text to be equalized between various demographic groups and controlling the overfitting issue.

In this presentation, focus on equalizing disparate toxicity predictions for different demographic groups. It formally describes logistic regression with the Perceived level of noise (PLDB) metric to create a regularization term in a classification algorithm. By minimizing the differences in toxicity predictions among identity terms within a protected group, minimize the ability for different identity terms to deferentially add toxicity to a sentence. Regularization techniques towards simpler models not only increasing test accuracy and occurred overfitting is controlled but also preventing models that work well only for the majority demographics in a dataset.

To calculate the bias in TBF is given by equation (15)

$$MAE = \frac{1}{T} \sum_{t=1}^{T} \left[P_t - X_t \right]$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} \left[P_t - X_t \right]^2} \tag{15}$$

here Pt denotes predicted value at time t, Xt denotes observed value at time t whereas T denotes number of predictions.

Thus it effectively predicts the reliable disease with certain features of epidemic diseases. However, in the prevention of this epidemic outbreak still concern a major problem. In which human mobility is a major concern thus the work implement IoT-based prevention technique it is explained in the following section.

3.3 IoT based global pre-emptive scheme

IoT - based healthcare solutions make it easier for doctors to practice, protect, and diagnose medicine. With actual information and the ability to evaluate previous therapies and diagnostic tests of a person, an IoT-enabled smart health care system can help reduce errors. Already the wearable device exists in IoT but it does not solve the human mobility of mobility pattern detection. In the benefit of IoT wearable devices, the work, therefore, implements the global pre-emptive system based on IoT to resolve the outbreak of disease due to human mobility. The IoT-based global pre-emptive system in which the potential role of a Hidden Markov Model (HMM) with K-NN a machine learning technique is explored for healthier prevention is implemented. After the pattern has been detected, HMM detects the mobility pattern, which validates the efficacy of epidemic diseases.

3.3.1 Hybrid Markov-based model

A Hybrid Markov-based model is used to forecast a user's future trends based on mobility patterns. It employs several approaches to determine the Spatio-temporal pattern of everyone’s pathway. By detecting each person's regular dynamic characteristics and taking into account the time of day when places are visited, the framework describes the set of the Markov algorithm. This method fully accounts for non-Gaussian and spatio-temporal aspects of users' pathways, which greatly improves prediction results.

Mobility Pattern detection was used to test the proposed hybrid Markov-based forecasting model for human mobility. At this stage, the mobility patterns of the discover individual, as well as the regular movement patterns, can be used to evaluate the order of the Markov determinant for each person.

3.3.1.1Mobility pattern detection

Size of the movementform $M = < i_1, i_2, \ldots i_n >$ is n when the pattern M has n hotspots. Given a minimum support threshold, δ, a candidate mobility pattern CM_p can be established as a mobility pattern if and only if CM_p contains support value satisfying: $\sup(S) \geq \delta$

This pattern alters Apriori method to mine the trajectories' frequent mobility sequences. The mobility pattern discovery method is a modified version of Apriori algorithm. Unlike the typical frequent item discovery process, the frequent mobility sequences had been found here include consecutive locations, and the maximum mobility patterns are chosen as the final result. The primary goal of this technique is to find a continuous trajectory with a support value greater than δ. Previously, each hotspot's support value was calculated first, then a set of mobility patterns with length-1 was constructed. Then, using mobility patterns with length- l, mobility patterns with length $(l-1)$ are constructed. When the set of length l is ϕ, the iteration is complete. Algorithm 1 shows the pseudo-code for the algorithm.

Algorithm 1

```
Input: Support threshold $\delta$
Mobility Trajectories $mT$
Set of hotspot $H$
Output: Mobility Patternset $S$
Formula $\left(mobilityPattern\det ection\right)\left(\delta, mT, H\right)$
$l = 1$
$CM_l = \{i \mid i \in H\}$
$S_l = \{i \mid i \in H \wedge \sup(i) > \delta\}$
$S = \{\ \}$
Repeat
    $l = l + 1$
    $for\ all\ mobility\ pattern\ M_{l-1} \in S_{l-1}\ do$
        $for\ all\ frequent\ pattern\, M_1 \in S_1\, do$
            $CM_l = \{c_l \mid c_l = M_{l-1} \bigcup S_1\}$
        $end\ for$
    $end\ for$
    $for\ all\, trajectory\ t \in T\ do$
        $CM_t = subset(CM_l, t)$
        $for\, all\, candidate\, c \in CM_t\ do$
            $count(c) = count(c) + 1$
        $end\ for$
    $end\ for$
    $S_l = \{c \mid c \in CM_l \wedge \sup(c) > \delta\}$
    $S = \bigcup S_l$
$until\ S_l = \phi$
$\mathrm{Re}turn\ S$
$end\ formula$
```

From this the detection of mobility pattern is done thus the mobility pattern is detected which help predict the person who is affected in movement of people with the neighbors. In order to predict neighbor person who is infected the system implement K-NN technique this prevent the disease outbreak which is stated in upcoming segments.

3.3.2 K-Nearest Neighbor (K-NN)

KNN is a successful machine learning data classification approach based on neighboring advanced samples of subset. K value is always positive, and an item is classed based on how many options its neighbors have. Neighbors are picked based on the symptoms that are most correctly related identifying the infected person, whereas Euclidean distance is used to compute the distance between two peoples.

Every occurrence is linked to a point in the sample space or a population. Euclidian distances formula is a common formula for calculating distances:

$$d=\sqrt{\left(D_{W1}-D_{W2}\right)^{2}+\left(D_{W21}-D_{W22}\right)^{2}+\ldots\ldots\ldots a} \quad (16)$$

Yet, it is critical to remember that all instances in the sample space must have the same scale. The distance measurement technique for qualitative data will be different which is vital to analyze if the cases are the similar or not. The qualitative items are measured at this step by assigning Boolean values to separateitem. Some strategies may be able to convert to illustrationsamong which distance may be determined. Like temperature, dry cough, fatigue, trouble breathing, movement, aches and pains to stop pretentious individuals. Closely related instances should have comparable values, whereas k-NN finds the required values from a large number of options by calculating its closest most impacted persons. Prediction of values as best heuristic distance as

$$f(n)=g(n)+h(n) \quad (17)$$

Here, $f(n)$ denotes the estimated cost of cheapest solution through n, $g(n)$ is the cost of the best path found so far to n, $h(n)$ is an admissible heuristic.

Fig. 4. Nearest neighbor Prediction

In the fig 4 depict that the circle belongsto the infected and normal person in the located area. Thus from this the K-nearest neighbor analyzes by monitoring the hazardous symptoms of infectious patients like fever, dry cough, fatigue, trouble breathing, movement, aches and pains to stop pretentious individuals from the normal person. From this method, the HMM was carried out by human mobility with individual mobility pattern detection, in addition, Aprioir algorithm can detect the individual mobility pattern and it predicts the affected person who is neighbor on the movement with the severity of disease by KNN. In addition, KNN is costly and have scaling problem but improves the accuracy of prediction model, easy to implement in IoT, good data stability for adjusting and time complexity also reduced.

From the suggested method it is highlighted that the IoT wearable device collects the data from human and the data integration has done in CART model, efficiently overcome the class distribution problem in data integration with removal of outliers, and reducing the scaling problem with missing value. Then in the prediction problem ILS-SVR technique is highly overcomes by removes the noise and demographic bias. “Prevention is better than cure” In case of predicting the disease and treat the people the outbreak of disease is not controlled. One of the main cause of disease outbreak is human mobility to prevent this disease outbreak proposed one IoT based global preemptive scheme effectively identify the affected person with the major symptoms of epidemic once the device detect the disease it alerts the person if the person with the disease affected person. In order to this the prevention of disease is highly useful for the people to better predict and prevent outbreak of the diseases.

4 Results and Discussion

The above segment explains exactly the feasibility of the proposed method by comparing and contrasting test results obtained with older techniques and performance metrics used for epidemic disease outbreak prevention. Below are specification tools and a dataset description for implementation.

4.1 Dataset description

The proposed method implemented based on benchmark data set as dengue dataset which describes the two main and secondary data were used in the dengue data collection. Primary data were gathered using the questionnaire system and secondary data were collected in the field of healthcare in medical centers. The dataset contained data observations of the attributing characteristics, such as ID, Age, Male/Female, Pulses, Acute fever, Platelet count, Rashes, Vomits, Stomach discomfort, Skin Peeling, Body ache, Cold, Constant Vomiting, Crunchiness, Gums bleeding, Headache, Nausea, Tourniquet test, Exhaustion, Fast Breathing, Antigen Dengue (NS1), IgM, Platelet deficit, IgG, NS1 Dengue.Thus, to transmit the dataset in IOT the delay will be low and the frequency of heart rate will attain better frequency

4.2 Simulation Output

Prevention of epidemic disease outbreak the system works with IoT based dataset. Initially the input signal is taken from the IoT data is gathered and preprocessing the signal with extraction of feature. Then the extracted featuresare linked with dengue dataset in which the data integration problem has overcome by CART then the integrated data send to prediction of appropriate disease where the noise and demographic bias affect the prediction process it is solved by ILS-SVR. Finally, for the prevention of diseases outbreak human mobility was monitored by IoT based global preemptive scheme which find out the neighbor person who is affected by the disease and gives the alert message.

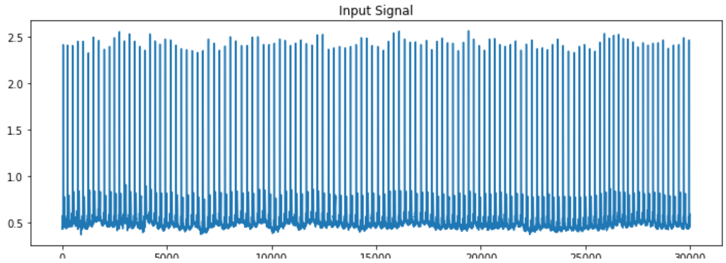

Fig. 5. Input Signal

Fig 5 shows input signal of person is taken from the IoT data which will be further go on preprocessing.

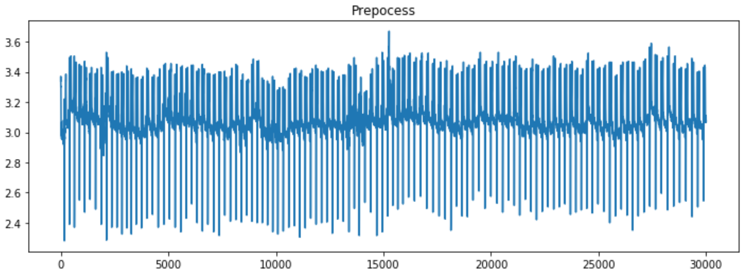

Fig. 6. Preprocessed Signal

Fig 6 shows the preprocessing of the input signal which is received from the Iot data and it is preprocessed

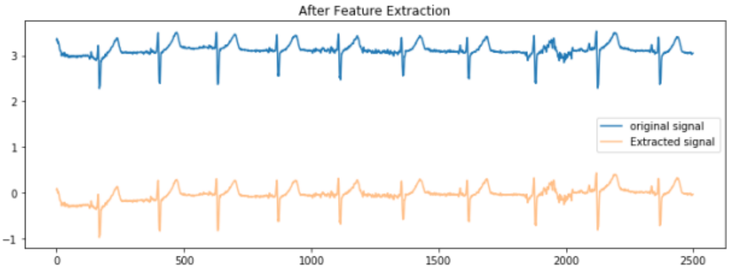

Fig. 7. Feature Extraction

Fig 7 shows the feature extracted pulse rate with the CART regression model in which highly extracts the features from the signal.

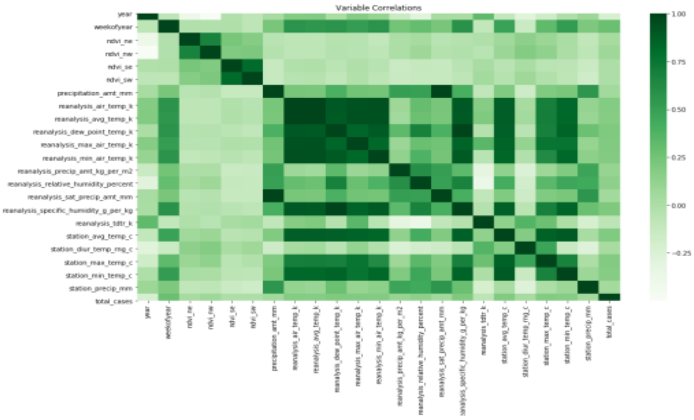

Fig. 8. Data integration for variable correlation

Fig 8shows the prediction of disease in variable correlation of data integrationin the dengue dataset which is epidemic.Thus the correlation integrates the data using CART model

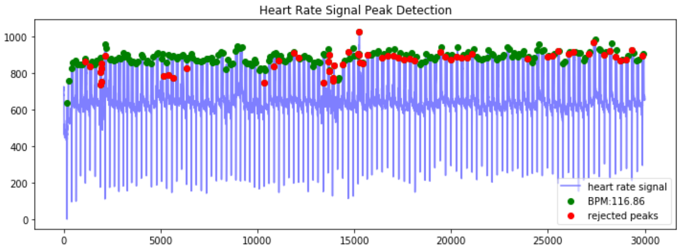

Fig. 9. Removal of noise in Heart rate signal

Fig 9 shows that the removal of noise in heart rate signal. This solves the noise from the integrated data to find the variables of correlation by rejection peaks. This rejection peaks removes the noise and attains the BPM of 116.86. Also expose the location based prediction in the removal of demographic bias

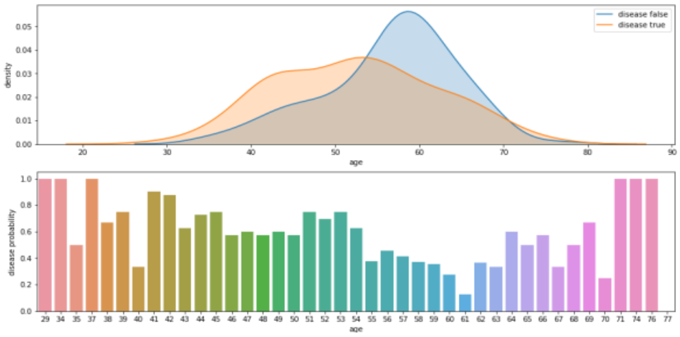

Fig.10. Disease Probability prediction using age

Fig 10 shows the disease probability prediction of the highly affected people and the density of disease varies with the public age.

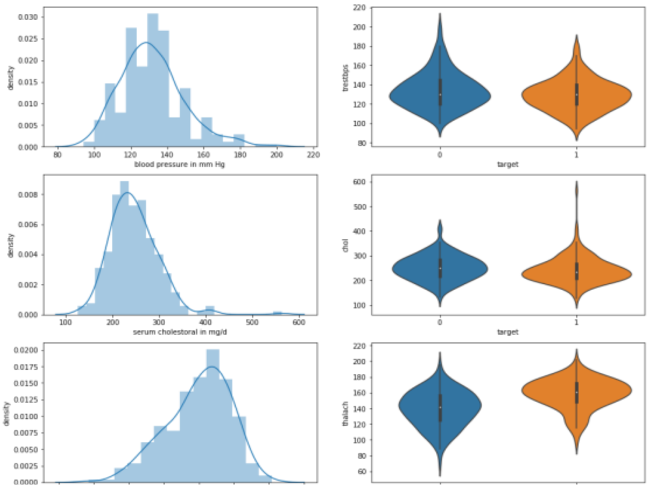

Fig.11. Mobility calculation

Fig 11shows the movement of person to analyses the dengue fever with pulse rate, blood lipids, and highest pulse rate achieved if it is exceeding the normal rate the alert message is produced to make immediate treatment to recover from the disease to prevent outbreak.

4.3Performance metrics with Comparisons

To evaluate the effectiveness of prevention of epidemic disease, some quantitative measurements have been verified the performance of such parameters. To determine superiority of proposed method a comparison is performed with existing prediction of epidemic diseases such as SHS (smart IoT based architecture of healthcare system) [24], HAFCA (Fog Cloud Assisted IoT based Hierarchical approach) [29], LDA-ANFIS (Linear Discriminant Analysis Adaptive Neuro Fuzzy Interference System) [30] and MSO-MLP (Multi Swarm Optimized Multi-Layer Perceptron) [31].

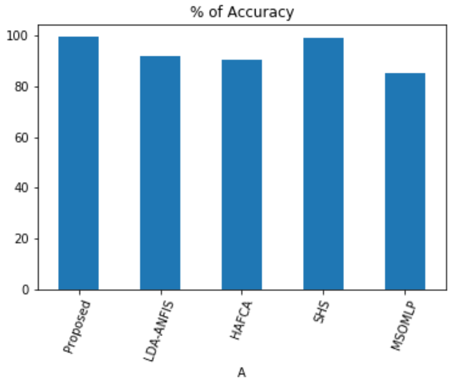

Fig. 12. Performance Metrics Comparison for Accuracy

Fig 12 shows the comparison graph for accuracy in the prediction of epidemic disease. The proposed method achieves higher accuracy of 99.5% than the state of the art methods. These better aids prevent the disease outbreak with the prediction of diseases.

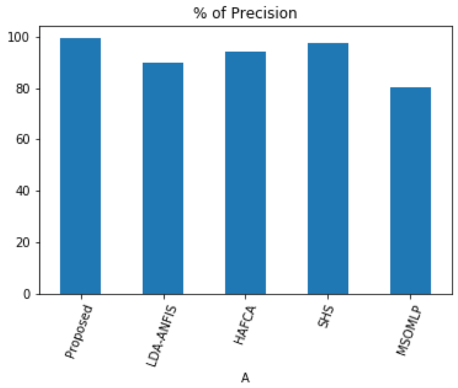

Fig. 13. Performance Metrics Comparison for Precision

Fig 13depicts that the comparison for precision with previous technique. Thus the graph reveals that the proposed method obtained 98.1% the precision will be high when compared with prior method

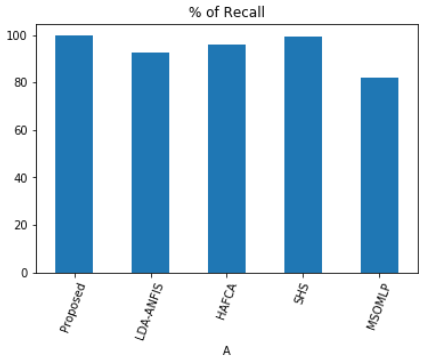

Fig.14. Performance Metrics Comparison for Recall

Fig 14 shows the comparison graph for recall, it is the fraction of relevant information which are achieved in the prediction of disease with the highest value. It is somewhat failed to predict the positive result to avoid it increase the recall, the proposed one achieve 99.7%. when compared with previous technique.

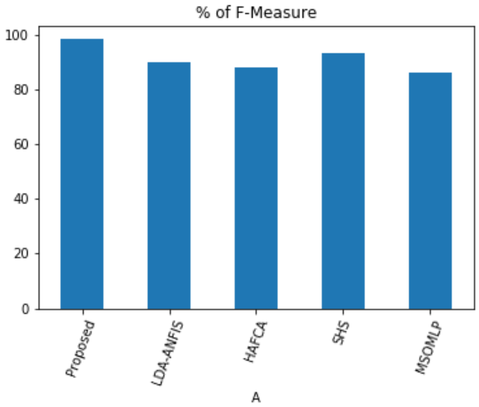

Fig.15. Performance Metrics Comparison for F-Measure

Fig 15shows the comparison graph for the F-Measure is a chromatic average of precision and recall. Thus, compared with existing technique, the proposed method achieves 97.3% in the prediction of epidemic diseases

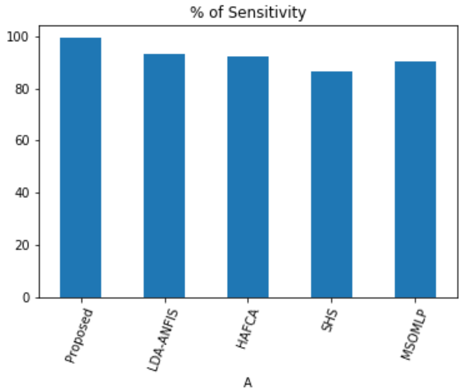

Fig.16. Performance Metrics Comparison for Sensitivity

Fig16 depicts that comparison between the prior strategy and proposed technique for sensitivity. Thus the graph reveals that the proposed method attains the sensitivity of 98% when compared with previous method.

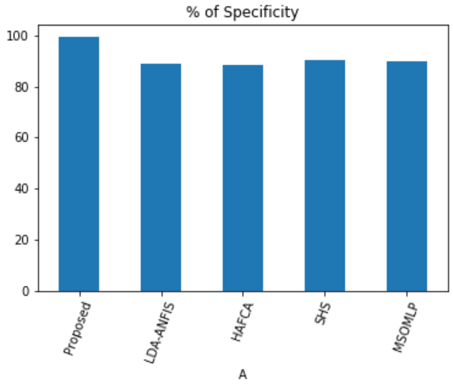

Fig.17. Performance Metrics Comparison for Specificity

Fig17 reveals the comparison graph for specificity in the epidemic disease prediction. Thus the graph represents the Specificity will be very high and attains 99% when compared with prior method and a highly test specifically if positive rules in the disease.

Table 1. shows comparison for various methodologies with various parameters for the proposed method

Methodologies	Accuracy (%)	Precision (%)	Recall (%)	F-Measure (%)	Sensitivity (%)	Specificity (%)
LDA-ANFIS	92	90	92	90	93	89
HAFCA	90.24	94.3	96.1	88.12	92.14	88.25
SHS	99	97.2	99.5	93.2	86.4	90.3
MSOMLP	85.18	80	82	86.1	90.3	89.8
Proposed	99.5	98.1	99.7	97.3	98	99

The above table 1 depicts that the comparison for various parameters with various methodologies. In Existing methodologies such as LDA-ANFIS achieves the accuracy of 92%, HAFCA attains the accuracy of 90.24%, SHS attains the accuracy of 99%, MSOMLP attain the accuracy of 85.18% and proposed attains the accuracy of 99.5%. The value of precision attains LDA-ANFIS as 90%, HAFCA as 94.3%, SHS as 97.2%, MSOMLP as 80% and proposed method has 98.1%. The value of recall achieves LDA-ANFIS as 92%, HAFCA as 96.1%, SHS as 99.5%, MSOMLP has 82% and proposed method has 99.7%. The value of F-Measure achieves LDA-ANFIS as 90%, HAFCA as 88.12%, SHS as 93.2%, MSOMLP has 86.1% and proposed method has 97.3%. The value of sensitivity achieves LDA-ANFIS as 93%, HAFCA as 92.14%, SHS as 86.4%, MSOMLP has 90.3% and proposed method has 98%. The value of specificity achieves LDA-ANFIS as 89%, HAFCA as 88.25%, SHS as 90.3%, MSOMLP has 89.8% and proposed method has 99%.

Thus from the above-obtained results,that can be revealedwhereas proposed technique attains better accuracy with the relevant parameters in the dengue dataset to analyze the disease. The proposed method achieves less time complexity and good stability to predict the epidemic disease. Comparison for proposed technology efficiently provides the better prediction of diseaseto motivate the prevention of epidemic outbreak with the quantitative assessment metrics.

5 Conclusion

Infectious diseases outbreaks have remained for long time as the challenge of public health, especially in developing countries, thus such outbreaks frequency has tremendouslyraised.Initially, the multiclass problem in the data integration, noise and demographic biasand human mobility are major issues to predict epidemic outbreak disease.Hence to vanquish the issue, Classification and regression tree is to predict the data,this technique removed the scaling and outliers of data. Furthermore, inadequacy of predictionprovokes noise and demographic bias can be overcome by Imbedded least square support vector regression (ILS-SVR) in which the accuracy of prediction conducts excellently. Finally, IoT base global preemptive scheme has prevent the neighbor people who is affected by the epidemic disease and automatically alert the public with IoT based scheme. Hence, automated alert systems prevent the mortality, morbidity rate, and detection of epidemic diseases. Thus, the proposed method not only predicts the disease but also prevent the public from epidemic outbreak achieves 99.5% accuracy with a higher prediction level. In future work includes the enhancement of proposed technique by Artificial intelligence for prediction of epidemic disease outbreak.

References

1. Gondalia, A., Dixit, D., Parashar, S., Raghava, V., Sengupta, A. and Sarobin, V.R.: IoT-based healthcare monitoring system for war soldiers using machine learning. Procedia computer science 133,1005-1013 (2018).https://doi.org/10.1016/j.procs.2018.07.075
2. Yuvaraj, N. and SriPreethaa, K.R.: Diabetes prediction in healthcare systems using machine learning algorithms on Hadoop cluster. Cluster Computing 22(1),1-9 (2019). https://doi.org/10.1007/s10586-017-1532-x
3. Kaur, P., Kumar, R. and Kumar, M.: A healthcare monitoring system using random forest and internet of things (IoT). Multimedia Tools and Applications 78(14),19905-19916 (2019). https://doi.org/10.1007/s11042-019-7327-8
4. Alhussein, M. and Muhammad, G.: Automatic voice pathology monitoring using parallel deep models for smart healthcare. IEEE Access 7,46474-46479 (2019). DOI: 10.1109/ACCESS.2019.2905597
5. Ramkumar, P.N., Haeberle, H.S., Ramanathan,D., Cantrell, W.A., Navarro, S.M., Mont, M.A., Bloomfield, M. and Patterson, B.M.: Remote patient monitoring using mobile health for total knee arthroplasty: validation of a wearable and machine learning–based surveillance platform. The Journal of arthroplasty 34(10),2253-2259 (2019).DOI: 10.1016/j.arth.2019.05.021
6. Rayan, Z., Alfonse, M. and Salem, A.B.M.: Machine learning approaches in smart health. Procedia Computer Science 154,361-368 (2019).https://doi.org/10.1016/j.procs.2019.06.052
7. Suresh, A., Udendhran, R., Balamurgan, M. and Varatharajan, R.: A novel internet of things framework integrated with real time monitoring for intelligent healthcare environment. Journal of medical systems 43(6),165 (2019). https://doi.org/10.1007/s10916-019-1302-9

8. Islam, M.M., Rahaman, A. and Islam, M.R.: Development of Smart Healthcare Monitoring System in IoT Environment. SN Computer Science 1(3) (2020). https://doi.org/10.1007/s42979-020-00195-y
9. Mani, N., Singh, A. and Nimmagadda, S.L.: An IoT Guided Healthcare Monitoring System for Managing Real-Time Notifications by Fog Computing Services. Procedia Computer Science 167,850-859 (2020).https://doi.org/10.1016/j.procs.2020.03.424
10. Rajan Jeyaraj, P. and Nadar, E.R.S.: Atrial fibrillation classification using deep learning algorithm in Internet of Things–based smart healthcare system. Health informatics journal 26(3),1827-1840 (2020).doi: 10.1177/1460458219891384
11. Mishra, A. and Mohapatro, M.: An IoT framework for bio-medical sensor data acquisition and machine learning for early detection. International Journal of Advanced Technology and Engineering Exploration 6(54),112-125 (2019). DOI:10.19101/IJATEE.2019.650027
12. Ed-daoudy, A. and Maalmi, K.: A new Internet of Things architecture for real-time prediction of various diseases using machine learning on big data environment. Journal of Big Data6(1),104 (2019). https://doi.org/10.1186/s40537-019-0271-7
13. Alfian, G., Syafrudin, M., Farooq, U., Ma'arif, M.R., Syaekhoni, M.A., Fitriyani, N.L., Lee, J. and Rhee, J.: Improving efficiency of rfid-based traceability system for perishable food by utilizing IoT sensors and machine learning model. Food Control 110,107016 (2020).https://doi.org/10.1016/j.foodcont.2019.107016
14. Gupta, A. and Katarya, R.: Social media based surveillance systems for healthcare using machine learning: A systematic review. Journal of Biomedical Informatics103500 (2020).https://doi.org/10.1016/j.jbi.2020.103500
15. Bollig, N., Clarke, L., Elsmo, E. and Craven, M.: Machine learning for syndromic surveillance using veterinary necropsy reports. PloS one 15(2), e0228105 (2020).https://doi.org/10.1371/journal.pone.0228105
16. Amini, A.,Banitsas, K. and Young, W.R.: Kinect4FOG: monitoring and improving mobility in people with Parkinson’s using a novel system incorporating the Microsoft Kinect v2. Disability and Rehabilitation: Assistive Technology 14(6),566-573 (2019).doi: 10.1080/17483107.2018.1467975
17. Kheirkhahan, M., Nair, S., Davoudi, A., Rashidi, P., Wanigatunga, A.A., Corbett, D.B., Mendoza, T., Manini, T.M. and Ranka, S.: A smartwatch-based framework for real-time and online assessment and mobility monitoring. Journal of biomedical informatics 89,29-40 (2019).https://doi.org/10.1016/j.jbi.2018.11.003
18. Jayaratne, M., Nallaperuma, D., De Silva, D., Alahakoon, D., Devitt, B., Webster, K.E. and Chilamkurti, N.: A data integration platform for patient-centered e-healthcare and clinical decision support. Future Generation Computer Systems 92,996-1008 (2019).https://doi.org/10.1016/j.future.2018.07.061
19. Suresh, A., Udendhran, R., Balamurgan, M. and Varatharajan, R.: A novel internet of things framework integrated with real time monitoring for intelligent healthcare environment. Journal of medical systems 43(6),165 (2019).
20. Ali, F., El-Sappagh, S., Islam, S.R., Kwak, D., Ali, A., Imran, M. and Kwak, K.S.: A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion 63,208-222 (2020).https://doi.org/10.1016/j.inffus.2020.06.008
21. Yousefpour, A., Jahanshahi, H. and Bekiros, S. (2020). Optimal policies for control of the novel coronavirus disease (COVID-19) outbreak. Chaos, Solitons & Fractals 136, 109883. https://doi.org/10.1016/j.chaos.2020.109883
22. Wang, P., Zheng, X., Li, J. and Zhu, B.: Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics. Chaos, Solitons & Fractals139,110058 (2020).https://doi.org/10.1016/j.chaos.2020.110058
23. Wang, Z., Xia, C., Chen, Z. and Chen, G. (2020). Epidemic propagation with positive and negative preventive information in multiplex networks. IEEE transactions on cybernetics 51(3), 1454 – 1462. DOI: 10.1109/TCYB.2019.2960605
24. Souri, A., Ghafour, M.Y., Ahmed, A.M., Safara, F., Yamini, A. and Hoseyninezhad, M.: A new machine learning-based healthcare monitoring model for student’s condition diagnosis in Internet of Things environment. Soft Computing24, 17111–17121 (2020). https://doi.org/10.1007/s00500-020-05003-6
25. Misawa, D., Fukuyoshi, J. and Sengoku, S.: Cancer Prevention Using Machine Learning, Nudge Theory and Social Impact Bond. International journal of environmental research and public health 17(3),790 (2020).doi: 10.3390/ijerph17030790
26. Wu, K., Darcet, D., Wang, Q. and Sornette, D. (2020). Generalized logistic growth modeling of the COVID-19 outbreak: comparing the dynamics in the 29 provinces in China and in the rest of the world. Nonlinear dynamics 101(3), 1561-1581. https://doi.org/10.1007/s11071-020-05862-6
27. Wang, J., Kong, X., Xia, F. and Sun, L. (2019). Urban human mobility: Data-driven modeling and prediction. ACM SIGKDD Explorations Newsletter 21(1), 1-19.https://doi.org/10.1145/3331651.3331653
28. Kiang, M.V., Santillana, M., Chen, J.T., Onnela, J.P., Krieger, N., Engø-Monsen, K.,

Ekapirat, N., Areechokchai, D., Prempree, P., Maude, R.J. and Buckee, C.O. (2021). Incorporating human mobility data improves forecasts of Dengue fever in Thailand. Scientific reports 11(1), 1-12. https://doi.org/10.1038/s41598-020-79438-0

29. Kumar Sood, S., Sood, V. and Mahajan, I. (2020). Fog–Cloud Assisted IoT-Based Hierarchical Approach for Controlling Dengue Infection. The Computer Journal 65(1), 67–79. https://doi.org/10.1093/comjnl/bxaa005
30. Sood, S.K., Kaur, S. and Chahal, K.K. (2020). An intelligent framework for monitoring dengue fever risk using LDA-ANFIS. Journal of Ambient Intelligence and Smart Environments, (Preprint)12(1), 1-16. DOI: 10.3233/AIS-200547
31. Kumar, N.K. and Sikamani, K.T. (2020). Prediction of Chronic and Infectious Diseases using Machine Learning Classifiers-A Systematic Approach 13(4), 11-20. DOI: 10.22266/ijies2020.0831.02