Open access Journal

International Journal of Emerging Trends in Science and Technology

ISSN 2348-9480

DOI: https://dx.doi.org/10.18535/ijetst/v8i9.01 || Vol:09 Issue:08

Design and Development of Artificial Intelligence Based 'Denv Shield' App for Predicting Dengue

Dr.T.Ruban Deva Prakash
Managing Director, R & D
Effyies Smart Technologies LLP
Neyyoor, India
Email:

Achshah R M
CEO, Product Development
Effyies Smart Technologies LLP
Neyyoor, India
Email: ceo@effyies.com

Abstract— Dengue fever, dengue hemorrhagic fever and dengue stroke syndrome are caused by dengue virus (DENV) belonging to the family Flaviviridae and spread through Aedes mosquitoes. Incessant mutation resulting from the proofreading inefficiency of their RNA polymerases under increasing air temperature and CO2 concentration make it thorny to diagnose or strengthening immune system against DENV. The vector competence of Aedes is influenced by diurnal temperature range (DTR). Under these circumstances predicting the genotype of the imminent DENV and vector competence in advance is momentous for vaccine development and vector surveillance. In order to make it efficient, this research aims in designing mobile app 'DENV SHIELD' which uses a novel In Vitro Replication Model (IVRM) to predict the nucleotide sequence of imminent genotype of DENV in prior. This app also makes real time prediction of the Vector Competence (VC) based on prediction of DTR corresponding to the region of the user using an Adaptive Network-Based Fuzzy Inference System (ANFIS) model. A self aware artificial intelligence model (SA-AIM) is used for forecasting air temperature and CO2 in advance. DTR is predicted using random forest algorithm which uses ARIMA time series prediction model for getting forecast of wind velocity and solar radiation. Training data are collected from National Center for Biotechnology Information, Centre for Diseases Control and Prevention and National Climate Data Center. The outcome of this research is an early warning system app which will help diagnostic centers and health workers to know genotype of forthcoming variant of virus and vector competence.

Keywords- *Early warning system, genotype, vectorial capacity, dengue fever, big data analysis, artificial intelligence*

I. INTRODUCTION

One among the dangerous arboviral diseases in India is dengue. In India first dengue outbreak is reported in 2010 followed by extensive outbreaks in subsequent years. Dengue fever (DF) or dengue hemorrhagic fever (DHF) is a mosquito-borne viral disease caused by the dengue virus of the Flaviviridae family and is responsible for colossal health and economic burden worldwide. Aedes aegypti and Aedes albopictus are the vectors responsible for the transmission of dengue viruses (DENV). The four DENV serotypes (1, 2, 3, and 4) have been co-circulating in India. Phylogenetic analyses revealed the circulation of Indian lineage of DENV-2 (genotype-IV) and DENV-3 (genotype-III) in vectors being the prevailing serotype. Dengue virus replicates itself and has viral components, such as the capsid, membrane, and envelope proteins. These viruses can evolve and gain mutations over time causing changes in viral genomes over time, thus escaping from immune system. Hence it is the need of time to predict the possible upcoming serotype and alert immune system and diagnostic centers. In addition improving diagnostics, improving dengue surveillance is an essential way to prevent and control dengue transmission. Early warnings prior to epidemics will improve the preparedness of public health officials, which can help reduce fatality rates. Vector surveillance is crucial for public health officials so that they can provide a prompt and preventative response to dengue. Vectorial capacity (VC) describes a vector's propensity to transmit dengue taking into account human, virus, and vector interactions. VC is highly temperature dependent. Recent evidence shows that diurnal temperature range (DTR) plays an important role in influencing the behavior of the primary dengue vector Aedes aegypti. A large DTR at low temperature increases the proportion of infected mosquitoes. Hence an accurate artificial intelligence (AI) model is necessary to predict VC based on DTR. Based on studies it is evident that mosquitoes lived longer and were more likely to become infected under moderate temperature fluctuations, which is typical of the high DENV transmission season than under large temperature fluctuations, which is typical of the low DENV transmission season. AI models can predict temperature and its fluctuations for a particular region based on past data and intelligence which can be used for predicting the high DENV transmission season. Climatic factors play a pivotal role in the epidemiological pattern of DF/DHF in terms of the number of cases, severity of illness, shifts in affected age groups, and the expansion of spread from urban to rural areas. Hence severity of possible illness can be predicted based on forecast of climatic factors. The recent developments in AI technology, availability of big data, extensive data analyses techniques combined with cloud storage and hinder less internet access motivates and paves way for designing and developing mobile app named DENV SHIELD to realize the above opportunities.

Basic research includes a wide range of studies focused on learning how the dengue virus is transmitted and how it infects cells and causes disease. This type of research investigates many aspects of dengue viral biology, including exploration of the interactions between the virus and humans and studies of how the dengue virus replicates itself. One important field of basic research is dengue pathogenesis, the study of the process and mechanisms of dengue in humans. Scientists want to understand how the dengue virus causes damage to the human body and how the immune system responds to a dengue infection so that they can develop new treatments for the disease. Scientists are also studying the dengue viruses to understand which factors are responsible for transmitting the virus to humans. Researchers are investigating how the dengue virus replicates itself and the structure of the viral components, such as the capsid, membrane, and envelope proteins. Other dengue research focuses on vector biology. What is vector biology? This field of dengue research studies the disease vector, Aedes mosquitoes. Vector biology studies mosquito ecology, population biology, genetics, and behaviors to understand how mosquitoes transmit the dengue viruses. Patients with severe dengue illnesses can be treated successfully if they are diagnosed as early as possible. Scientists are working on improving dengue diagnostics so that patients infected with dengue can be treated quickly. In addition to performing basic research and improving diagnostics, improving dengue surveillance is an essential way to prevent and control dengue transmission. The World Health Organization in partnership with ministries of health, research centers, and laboratories around the world has developed a dengue surveillance system called DengueNet, a database that can be continuously updated to share current and historical data on dengue cases. The goals of DengueNet are to standardize reporting of dengue cases and to improve the preparedness of public health officials by providing early warnings prior to epidemics, which can help reduce fatality rates. Monitoring mosquito populations is a first line of defense against dengue. Vector surveillance allows for a prompt response to control mosquito populations and limit dengue transmission. Vector surveillance is crucial for public health officials so that they can provide a prompt and preventative response to dengue. Hence this project aims in developing an AI based early warning system which is designed as a mobile app. Prediction of dengue parameters in advance is the key of this work.

II. LITERATURE REVIEW

Since the first half of the 20th century, all serotypes experienced an increase in genetic diversity [9], causing variations leading to new genotypes with distinct lineages [10, 11]. Some studies have revealed associations between genetic diversity and features such as clinical manifestations, virulence and epidemic potential [15, 12]. DENV-1 exists in five distinct genotypes (genotypes I–V) [13, 14]. Recently, a basal group of DENV-1 with highly divergent sequences was reported, possibly constituting a new genotype (genotype VI) [13,1]. The evolutionary capacity of RNA viruses allows them to rapidly adapt to their environment and overcome barriers to infection [7, 3]. RNA viruses, such as DENV, typically display high mutation rates as a result of the error-prone RNA polymerase activity during genome replication [16]. In addition, the genetic recombination further facilitates their opportunity for genetic novelty [17]. This ultimately results in genetically diverse virus populations consisting of multiple genotypes of monophyletic nature. Evolutionary studies of dengue virus have revealed that its genetic diversity is increasing with climate change [18, 4, 5]. This, coupled with evidence that viral strains could naturally differ in virulence, suggests that in the future we might be exposed to viruses with an expanded range of pathogenic properties due to increasing temperature and carbon content of air [8, 2]. Therefore, DENV Shield App can predict the upcoming serotype. Genotyping of all the 990 strains of DENV strains was carried out using Dengue SubTyper tool [19]. Mutation of dengue virus occurs in E-protien, 5' UTR and 3' UTR [14]. Neucleotide sequence of dengue virus can be determined by RDP4 package for varying environmental conditions [21]. For the proposed IVRM we consider only E-protien, 5' UTR and 3' UTR. Dynamics of dengue serotype 2 viruses isolated from patients with different disease severity, namely flu-like classic dengue fever (DF) and dengue shock syndrome (DSS) were studied in its mosquito vector Aedes aegypti. Comparison of isolate infectivity and vector competence (VC) among thirty-two A. aegypti-viral isolate pairs is performed. Mosquito populations from high dengue incidence area exhibited overall greater VC than those from low dengue incidence area [6]. The interaction between DENV and its principal mosquito vector, Aedes aegypti (Linnaeus), has been studied in detail in the laboratory under constant temperatures. We use this study to predict VC. Experimental transmission was observed at a range of constant temperatures from 13°C to 35°C [22]. Increasing constant temperature reduced the extrinsic incubation period (EIP), increased mosquito mortality, and resulted in a higher proportion of infected mosquitoes [23]. Thereby, proposed ANFIS model can study the mean temperature and DTR to predict the mosquitoes' vector competence. The study of Lambrechts and others also modeled the thermodynamics of virus infection and transmission by mosquitoes, under symmetric and asymmetric temperature profiles. Their model was based on empirical data from three flaviviruses (West Nile virus, St. Louis encephalitis virus, and Murray Valley encephalitis virus) and led to the following predictions. Temperature fluctuations were expected to alter the direction of the response, dependent upon the mean temperature. At a low mean temperature (< 18°C), fluctuations were predicted to increase the probabilities of infection and dissemination, and at higher mean temperatures (≥ 18°C) fluctuations reduce these probabilities. In addition, under an asymmetric temperature profile a larger DTR was predicted to shorten EIP. The expected EIP at a constant 26°C is 11–12 days, and ≤ 10 days with a DTR of 20°C [24]. In the last five decades, several studies have been conducted to estimate the surface temperature and diurnal temperature range for future years, behavior patterns of pathogen species and vectors [25]. Climatic factors play a pivotal role in the epidemiological pattern of DF/DHF in terms of the number of cases, severity of illness, shifts in affected age groups, and the expansion of spread from urban to rural areas. Hence severity of possible illness can be predicted based on forecast of climatic factors. The recent developments in AI technology, availability of big data, extensive data analyses techniques combined with cloud storage and hinder less internet access motivates and paves way for designing and developing mobile app named DENV SHIELD to realize the above opportunities.

Our proposed app can predict the DTR and surface temperature for the future years in order to study the vector competence. Hence this project aims in developing an AI based early warning system which is designed as a mobile app. Prediction of dengue parameters in advance is the key of this work.

III. STUDY OBJECTIVE

The global incidence of dengue has grown dramatically with about half of the world's population now at risk. Although an estimated 100-400 million infections occur each year, over 80% are generally mild and asymptomatic. But it is predicted that upcoming years has to face dengue epidemic throughout the world due to climate change. Researchers are expecting new variants, as DENV replicate producing varying genotype to adapt itself to changing climatic conditions. Early detection of disease progression associated with severe dengue, and access to proper medical care lowers fatality rates of severe dengueto below 1%. But new variants are expected to produce high fatality rate as they resist existing vaccines. The main objective of this study is to facilitate early diagnosis, production of vaccine and implement vector control strategies at time of need by:

- Predicting the nucleotide sequence of DENV which may evolve after a year based on forecast of environmental conditions.
- Predicting vector competence of Aedes aegypti mosquitoes for upcoming season based on forecast of DTR.
- Developing a mobile app which will provide time to time data about threat of dengue in user's location.

IV. STUDY DESIGN

This research aims at quantitative analysis of correlation between variables increase in yearly mean surface temperature, concentration of CO2 in air and change in genotype of dengue virus. Based on change in yearly mean temperature and CO2 concentration of air in past years, the mean temperature and CO2 concentration are forecasted for forthcoming years for predicting genotype of imminent variants of DENV. A novel design of intelligent model is proposed to predict genotype of imminent variants of DENV. Also correlation between distance of a region from sea, elevation, wind velocity, solar radiation and diurnal temperature range (DTR) are studied for designing intelligent model to forecast DTR for upcoming years. Using correlation design the variation in vector competence (VC) of Aedes aegypti mosquitoes with rise in DTR for past years is studied for designing intelligent model to predict VC for next season based on forecasted DTR. The outcome of research is mobile app which reveal the imminent variant of DENV and alert users about VC for next season for facilitating early diagnosis, vaccine production for predicted antigen and vector control to minimize fatality of dengue. This preliminary study is done in three phases:

Phase - I: Developing an artificial intelligence based predictive model that alerts immune system and diagnostic centers about the possibility of upcoming serotypes.

The dataset for this model is taken from Dengue Net at real time. Based on the viral components (capsid, membrane and envelope proteins) the model predicts the serotype and if any new serotype is found it will alert. Since the dataset is taken from Dengue Net data cleaning and big data analysis is vital. This model's target is to classify and predict the genotype. Ex. Genotype III, Genotype II etc. Therefore, this model comes under supervised classification problem. We can train Decision Tree, K Nearest Neighbor and XG Boost algorithms. Decision tree splits the data into smaller and smaller subset while building tree like structure. K Nearest Neighbor (KNN) assumes the similarity between the new data and already existing data. XG Boost performs sequential tree building using parallelized implementation. The performance of the algorithms is tested for the sample data set, in order to finalize and choose the best algorithm for this predictive model.

Phase - II: Artificial intelligence-basedmodel to predict vectorial capacity (VC) on a scale of one to ten.

The data set contains the fluctuation in temperature and the mean temperature based on which the VC is predicted. To compute the fluctuation in temperature the temperature of that region is collect in real time. The target of the model is to rate VC on a scale of one to ten. The model is based on supervised classification algorithms like Decision Tree, K Nearest Neighbor and XG Boost. Performance of the algorithms are analyzed using classification report and confusion metrics to choose the algorithm that is most efficient for this problem.

Phase - III: Prediction of Transmission season- High transmission season and low transmission season.

In order to predict the transmission season the same data set for VC is used. The moderate temperature fluctuation means high DENV transmission season otherwise low DENV transmission season. The seasons are binary, high transmission season and low transmission season, therefore we use supervised classification algorithms to predict. Machine learning algorithms such as Decision Tree, K Nearest Neighbor, XG Boost are used for training the data set. The performance of the algorithms are analyzed for choosing the algorithm with the best performance for developing final model.

It is found that the existing predictive algorithms analyzed during preliminary study are not efficient in serving the objective. Hence new models are proposed. The project is implemented in three phases. During phase-1 data from NASA site for SA-AIM is collected and the model is developed. Phase-2 of projects deals with collection of data from CDC and NCBI along with development of IVRM. Data collection from NCDC and development of prediction model for VC will

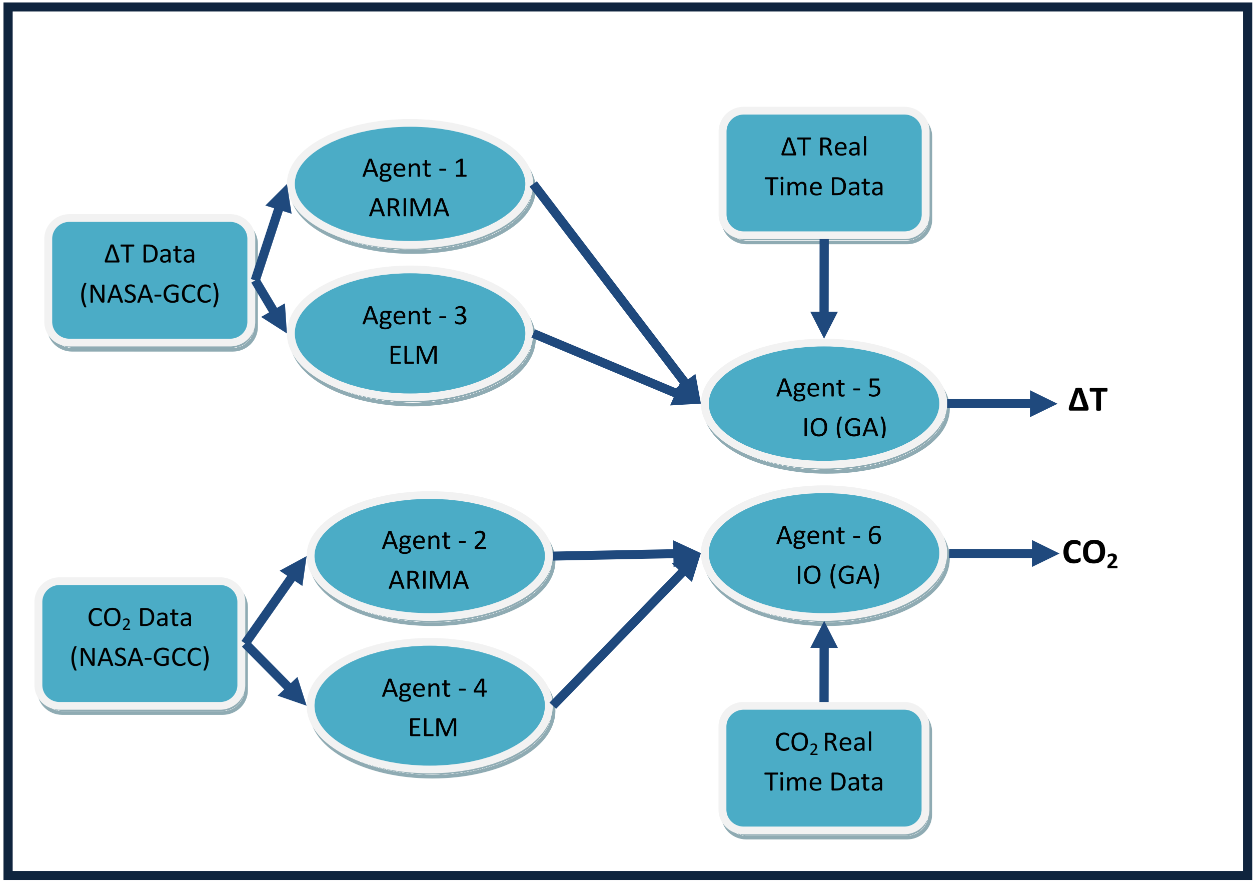

Figure 1. Architecture of SA-AIM

be completed in phase-3. Finally all the above models are integrated to develop the app 'DENVSHELD' which is an early warning system for dengue.

V. SELF AWARE ARTIFICIAL INTELLIGENCE MODEL

The first phase of project is design and development of self aware artificial intelligence model (SA-AIM) to forecast change in mean temperature and concentration of CO2. The proposed AS-AIM comprises of six intelligent agents as shown in fig. 1. The agents in first layer are pre-trained using global data called global predictors. But the user may require genotype of DENV in specific local geographical region. Hence output from layer-1 is fine tuned in layer-2 by learning with local real time data called local predictors. The agents 1 - 4 are based on ARIMA and ELM algorithms which should be pre-trained using yearly global mean temperature change and CO2 concentration data collected from NASA Global Climate Change and this forms the layer-1. Global temperature change and CO2 are predicted by both ARIMA and ELM agents of layer-1 for the demanded year by user. The predictions of both agents are inaccurate due to stohasticity and seasonality. In order to fetch region dependant predictions of variables for determining region specific nature of mutation, the global predictions are fine tuned by intelligent optimizer (IO) agents (local predictors) whom operate over weighted sum of predictions of ARIMA and ELM agents and are trained using genetic algorithm (GA) with region specific real time data. The problem statement for IO agent is:

Minimize $F = AW_1 + BW_2$

Subject to constraints:

$AW_1 + BW_2 = T$

$W_1 > 0$

$W_2 > 0$

Where A & B are outputs of ARIMA and ELM based agents respectively, T is the target and W1 and W2 are the weights to be determined corresponding to the data of local variable.

VI. IN VITRO REPLICATION MODEL

The second phase of project aims in developing In Vitro Replication Model (IVRM) which is designed specifically for outputting the nucleotide sequence of imminent genotype of DENV. The architecture of IVRM is shown in fig. 2. The 11kb RNA genome of the DENV has 10 genes which encodes three structural proteins envelope (E), membrane (M), capsid (C) and seven non-structural proteins NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 which are translated as a single polyprotein. The total length of the genomes of DENV ranges between 10,600 and 10,800 nucleotide sequences which comprises of adenine (A), cytosine (C), guanine (G), thymine (T),guanine/thymine(K)and adenine/cytosine/guanine/thymine (N). The position

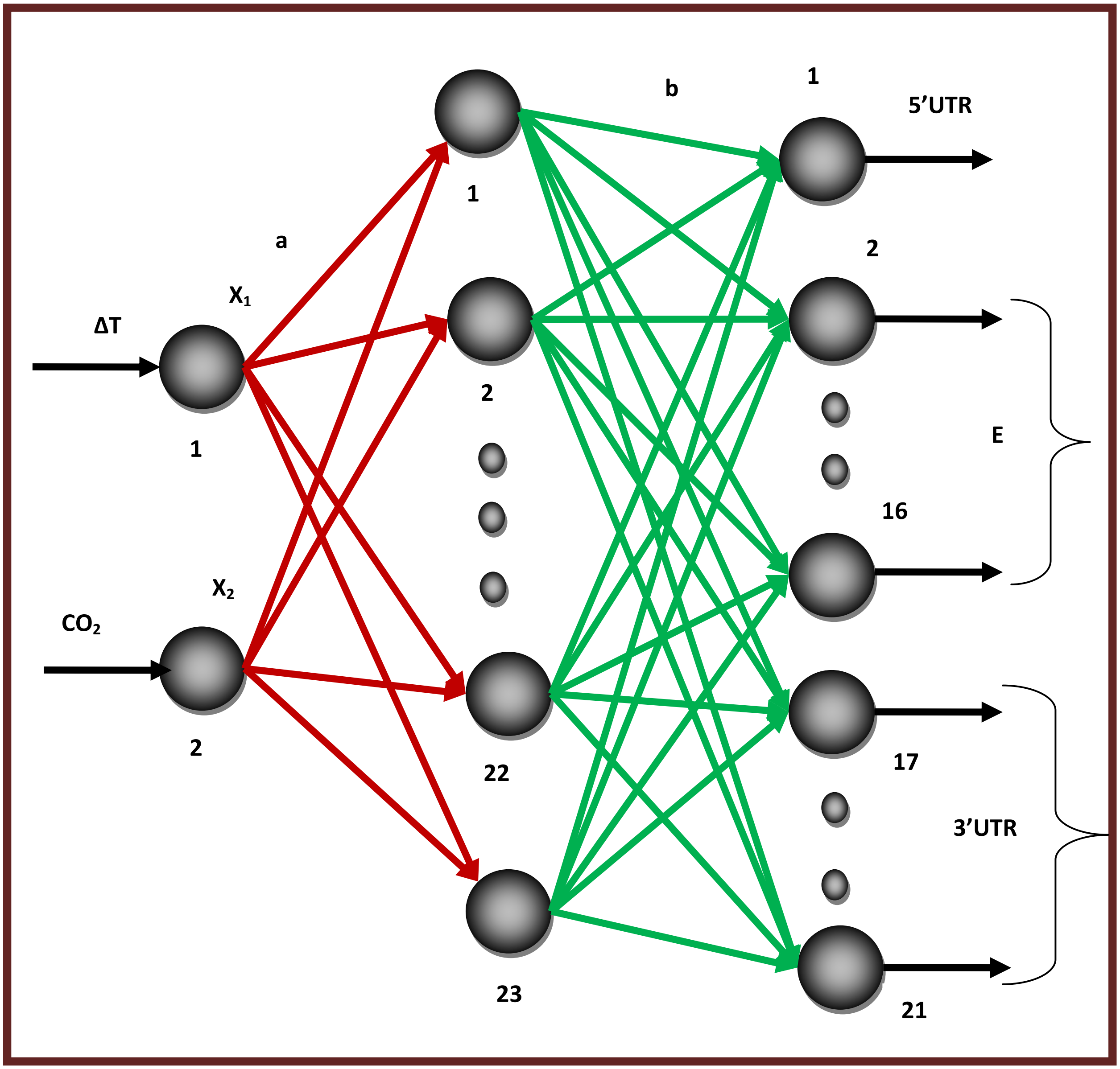

Figure 2. Architecture of IVRM

of genes in nucleotide sequence is shown in table-1. The primary determinants of virulence reside in the Envelop (E) gene with length 1485 nucleotide sequence, the 5' un-translated region (5' UTR) with length 95 nucleotide sequences and the 3' UTR with length 464 nucleotide sequences. Hence total positions which determines virulence is 2044 and these positions are critical in DENV replication cycle. Hence the IVRM is trained using coded nucleotide sequence of E gene, 5'UTR and 3'UTR which will examine mutation only in these critical positions. The codes for sequence contents are given in table-2. The output layer of IVRM has 21 neurons each neuron denote nucleotide sequence length of 100 and the input layer has 2 neurons corresponding to temperature change and CO2 concentration.

The input values in training data are standardized using (1). The hidden layer has 23 neurons with purelin activation function (y = mx).

$$\frac{(X - X_{min})}{(X_{max} - X_{min})} \qquad (1)$$

The architecture of IVRM is shown in figure-2. The target values of the training data are coded before training. During coding the contents of nucleotide sequence is replaced by corresponding numerical code and normalized by dividing with $5^{142.7}$. During decoding the output of each neuron is multiplied by $5^{142.7}$ and numbers in each digit are replaced by corresponding alphabets. The output neuron identity number is allotted as shown in table-3

TABLE -1 GENE POSITION IN NUCLEOTIDE SEQUENCE

Protein	Starting position	Ending position
5' UTR	1	95
Capsid	96	437
Membrane	438	935
Envelop	936	2420
NS1	2421	3476
NS2A	3477	4130
NS2B	4131	4520
NS3	4521	6377
NS4A	6378	6758
K	6759	6827
NS4B	6828	7574
NS5	7575	10271
3' UTR	10272	10736

TABLE -2 CODING AND DECODING

A	**C**	**G**	**K**	**N**	**T**	**NULL**
1	2	3	4	5	6	0

TABLE -3 ALLOTMENT OF OUTPUT NEURON IDENTITY

Region	Neuron Id
5' UTR	1
E	2 - 16
3' UTR	17 - 21

The output equation for IVRM is given in (2)

$$Z_i = m(W_{11}^a X_1 + W_{21}^a X_2)W_{1i}^b + m(W_{12}^a X_1 + W_{22}^a X_2)W_{2i}^b + \ldots + m(W_{1-23}^a X_1 + W_{2-23}^a X_2)W_{23-i}^b \qquad (2)$$

Where W_{ij}^a is the weight between i^{th} node of input neuron and j^{th} node of hidden layer in space-a

W_{ij}^b is the weight between i^{th} node of hidden layer and j^{th} node of output neuron in space-b

X_1 and X_2 are output of input neurons

m is the slope of purelin activation function

The error is given by (3), where T_i is the target.

$$E_i = Z_i - T_i \qquad (3)$$

The [S] matrix is

$$\begin{bmatrix} mW_{11}^b X_1 & \cdots & mW_{23-2}^b X_2 & m(W_{11}^a X_1 + W_{21}^a X_2) & \cdots & m(W_{1-23}^a X_1 + W_{2-23}^a X_2) \\ & & & \vdots & & \\ mW_{1-21}^b X_1 & \cdots & mW_{23-21}^b X_2 & m(W_{11}^a X_1 + W_{21}^a X_2) & \cdots & m(W_{1-23}^a X_1 + W_{2-23}^a X_2) \end{bmatrix}^{-1}$$

The total number of weights to be determined is 529 and the training algorithm is a batch processing algorithm. Hence 529 training set selected from entire range is fed in each batch. As we progress from batch to batch the algorithm converge in less number of epochs. The identity matrix is added to matrix

$$\left[\frac{\partial Z_i}{\partial W_{ij}^{a}}\right]$$

before taking inverse to authenticate accelerated convergence.

The training algorithm for IVRM is given below:

1. Assume all weights as [W]=0.5
2. Calculate [Z]
3. Calculate [E]
4. Print [W] if [E] < 0.1
5. Calculate [J] = [∂Z/∂W]
6. Calculate [S] = [J] + [I] and find inverse of [S].
7. Calculate [ΔW] = [S] [E]
8. Update [W] = [W] + [ΔW] and go to step-2

VII. ANFIS MODEL

The third phase of research is prediction of VC of Aedes mosquitoes in the region of user. The proposed model has two time series prediction algorithms ARIMA for forecasting the future trend of solar radiation and wind velocity. The output solar radiation and wind velocity for any future time period along with the distance of region from sea and elevation are fed to random forest algorithm. An ANFIS model is developed to output VC for the given DTR. The block diagram of model is shown in fig. 3.

VIII. SAMPLE SIZE

The term dengue fever came into general use only after 1828. Hence global climate change and weather data need to be collected for last 190 years in order to train the forecasting models. The first virologically proved epidemic of dengue fever (DF) occurred in Kolkata and Eastern Coast of India in 1963-1964. Hence it is decided to collect local weather data of seven geographical regions in India northern mountains and the northeast mountain ranges, Indo-Gangetic plains, Thar

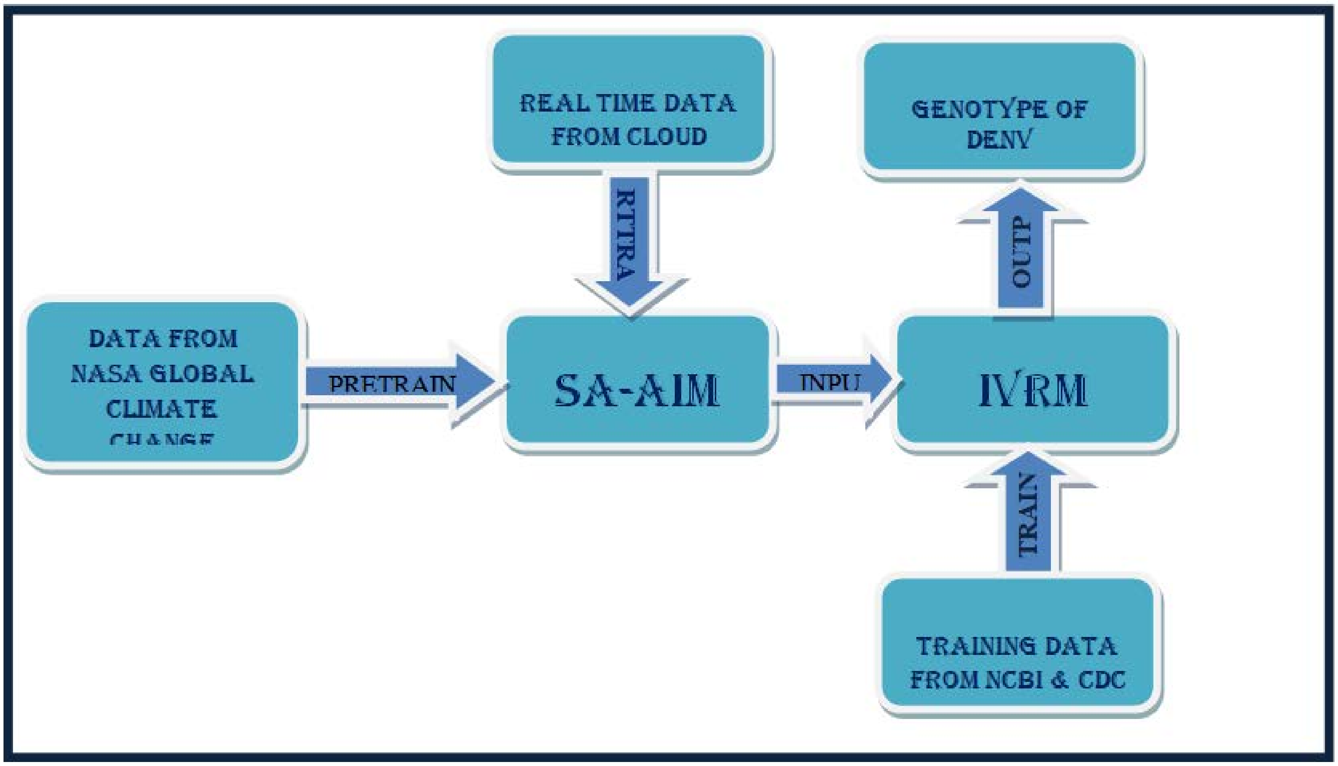

Figure 3. Architecture of IVRM

Desert, Central Highlands and Deccan Plateau, East Coast, West Coast, Bordering seas and islands for last 60 years in order to fine tune the forecasting models. Forty six variations in mutation rate corresponding to variation in temperature and concentration of CO2 is available for ΔT ranging from 0.2 to 0.7 ºC and CO2 concentration ranging from 370 to 393 ppm. Vectorial capacity values for four strains of dengue DENV 1-4 at 5, 7, and 9 days post exposure as well as the cumulative vectorial capacity (cVC) values are published by Louisiana State University School of Veterinary Medicine for different rate of change of DTR ranging from -0.036 to 0.048.

This research doesn't involve any study related to interaction between human, mosquito and virus. The weather data published in websites of NASA Global climate change and National Climate Data Center's Global Historical Climatology Network Daily is used. Mutation and genetic data is also freely available CDC website and National Center for Biotechnology Information. Experimental data required for training VC prediction model is obtained from research publication. Hence consent for using data is not required.

IX. DATA COLLECTION & STATISTICAL ANALYSIS

The data of global variables for pre-training global predictors of SA-AIM; rate of global surface temperature in ºC and variation in the concentration of carbon dioxide in parts per million can be collected from NASA Global Climate Change (https://climate.nasa.gov/vitalsigns/ global-temperature/). The real time data for fine tuning local predictors can be collected from https://openweathermap.org/api. The data obtained from research laboratories where replication of DENV is done in a laboratory setup by providing a controlled environment for experimentation is used for training IVRM. The genetic makeup data can be collected from Center for Diseases Control and Prevention (CDC) websites. Data of mutation studies are freely available from National Center for Biotechnology Information. Daily maximum and minimum air temperature for training the VC prediction model can be obtained from the National Climate Data Center's 'Global Historical Climatology Network Daily'. The Global Historical Climatology Network Daily contains records from over 75,000 stations around the world. The data of distance of the station from sea, elevation of station, wind velocity, solar radiation and DTR are tabulated and an artificial intelligence predictive model trained using above data can predict the DTR for any region in the world by fetching its distance from sea, elevation, wind velocity and solar radiation. The training data for ANFIS is collected from publication of Louisiana State University School of Veterinary Medicine (LSUSVM) whom conducted study on Aedes aegypti (Linnaeus) mosquitoes kept in KBF115 incubators maintaining experimental climatic conditions. Vector competence was calculated as the number of disseminated infections divided by the total number of mosquitoes exposed by varying temperature in incubator which is recorded by HOBO data loggers on an hourly basis.

As part of data analysis and cleaning before training the model it is required to check for null vales and duplicates as it might cause the model to be biased. In cases where the data is unbalanced we need to advanced techniques like SMOTE to make the data balanced. The correlation between the features needs to be checked and only the feature that has strong correlation with the target will be selected for training the model. It is also important to check if the features have correlation among themselves, multicollinearity. The multicollinearity should be less than 0.6. For regression model we need to check the normality of data. This is done by checking the skewness and kurtosis. The Skewness is between -0.5 and 0.5 and the kurtosis should be between -0.5 and 0.5 for the data to be normal. It is crucial to remove the outliers in the data as the accuracy of the model will be only as accurate as the data. In order find the outliers we can convert extreme data points into z scores that tells how many standard deviations away they are from the mean.

X. OUTCOME & FUTURE PLAN

Prediction about genotype of DENV helps to produce vaccines for strengthening immune system. Facilitates diagnostic center for early diagnosis and treatment thereby reducing the mortality rate.

Studying vector competence, the ability of Aedes mosquitoes to acquire and transmit the dengue virus helps health department for taking preventive measures to mitigate the spread.

IVRM model predicts only change in sequence of nucleotide in dominant proteins E, 5' UTR and 3' UTR for optimizing time and memory. The work can be extended by designing models to predict the entire nucleotide sequence. Prediction of VC is done considering extrinsic factors which mainly contribute for change in vector competence. A model can be designed to predict VC considering both extrinsic and intrinsic factor for higher accuracy. Agent based learning model can be designed to simulate the interactions between vector, virus and host. Prediction of number of cases, severity of illness and shifts in affected age groups aids in making prior lab and hospital arrangements which reduces risk of human life and discomfort.

XI. LIMITATIONS

IVRM is trained by using batch processing of training data points and maximum number of data points trained at a time is limited to number of weights in the model. The app is designed as an early warning system for benefitting health care workers, diagnostic technicians and therapist. The user should have basic knowledge about output parameters to interpret the result and take precautionary measures.

XII. CONCLUSION

The app 'DENV SHIELD' recognizes the user location and its state. It searches automatically in internet to get the past and current data of temperature and climatic factors of the region stored in cloud. The AI model is trained in real time for forecasting mean temperature, DTR and maximum temperature fluctuation. Based on this data VC, nature of transmission, number of denque cases, severity of illness and shift in affected age group are predicted for that region. The details can be obtained in prior for any period requested by user. The app get updated automatically using data from DengueNet for predicting genotype of DENV. EDA process is executed automatically by app using big data analysis techniques without human intervention. This contribution will benefit the society and health department by providing essential information to resist the disease, identify and diagnose dengue cases, and implement mosquito-control efforts in advance.

ACKNOWLEDGMENT

We thank National Center for Biotechnology Information for providing us the data of mutation studies. We also acknowledge helping hands of National Climate Data Centre and NASA Global Climate Change in getting the required training data.

REFERENCES

[1] Bellone, R., Lequime, S., Jupille, H. et al. Experimental adaptation of dengue virus 1 to Aedes albopictus mosquitoes by in vivo selection. Sci Rep 10, 18404 (2020). https://doi.org/10.1038/s41598- 020-75042-4

[2] Awang MF, Dom NC (2020) The effect of temperature on the development of immature stages of Aedes spp. against breeding containers. International Journal of Global Warming 21:215–233. https://doi.org/10.1504/IJGW.2020.108671

[3] Faruk, M.O., Jannat, S.N. & Rahman, M.S. Impact of environmental factors on the spread of dengue fever in Sri Lanka. Int. J. Environ. Sci. Technol. (2022). https://doi.org/10.1007/s13762-021-03905-y

[4] Kulkarni, M.A., Duguay, C. & Ost, K. Charting the evidence for climate change impacts on the global spread of malaria and dengue and adaptive responses: a scoping review of reviews. Global Health 18, 1 (2022). https://doi.org/10.1186/s12992-021-00793-2

[5] Data-driven methods for dengue prediction and surveillance using real-world and Big Data: A systematic review Emmanuelle Sylvestre , Clarisse Joachim, Elsa Cécilia-Joseph, Guillaume Bouzillé, Boris Campillo-Gimenez, Marc Cuggia, André Cabié, Published: January 7, 2022, https://doi.org/10.1371/journal.pntd.0010056

[6] Ronald Enrique Morales-Vargas , Dorothée Missé, Irwin F Chavez, Pattamaporn Kittayapong, Vector Competence for Dengue-2 Viruses Isolated from Patients with Different Disease Severity, Pathogens, 2020 Oct 21;9(10):859 doi: 10.3390/pathogens9100859.

[7] Dolan PT, Whitfield ZJ, Andino R. Mechanisms and concepts in RNA virus population dynamics and evolution. Annual Review of Virology. 2018;5:69–92. doi: 10.1146/annurev-virology-101416-041718.

[8] Xin-Ni Lim, Chao Shan, Jan K. Marzinek, Hongping Dong, Thiam Seng Ng, Justin S. G. Ooi, Guntur Fibriansah, Jiaqi Wang, Chandra S. Verma, Peter J. Bond, Pei-Yong Shi, Shee-mei Lok. Molecular basis of dengue virus serotype 2 morphological switch from 29°C to 37°C, PLOS Pathogens, 2019; 15 (9): e1007996 DOI: 10.1371/journal.ppat.1007996

[9] Zanotto P.M., Gould E.A., Gao G.F., Harvey P.H., Holmes E.C. Population dynamics of flaviviruses revealed by molecular phylogenies. Proc. Natl. Acad. Sci. USA. 1996;93:548–553. doi: 10.1073/pnas.93.2.548.

[10] Dos Santos F.B., De Bruycker-Nogueira F., Castro M.G., Nunes P.C.G., De Filippis A.M.B., Faria N.R.D.C., Simões J.B., Sampaio S.A., Santos C., Nogueira R.M.R. First report of multiple lineages of dengue viruses type 1 in Rio de Janeiro, Brazil. Virol. J. 2011;8:387. doi: 10.1186/1743-422X-8-387.

[11] Holmes E.C., Twiddy S.S. The origin, emergence and evolutionary genetics of dengue virus. Infect. Genet. Evol. 2003;3:19–28. doi: 10.1016/S1567-1348(03)00004-2.

[12] Guzman M.G., Gubler D.J., Izquierdo A., Martinez E., Halstead S.B. Dengue infection. Nat. Rev. Dis. Primers. 2016;2:16055. doi: 10.1038/nrdp.2016.55.

[13] Pyke A.T., Moore P.R., Taylor C., Hall-Mendelin S., Cameron J.N., Hewitson G.R., Pukallus D.S., Huang B., Warrilow D., Hurk A.F.V.D. Highly divergent dengue virus type 1 genotype sets a new distance record. Sci. Rep. 2016;6:22356. doi: 10.1038/srep22356

[14] Arenas C.J.V., Zanotto P.M.D.A. Worldwide Spread of Dengue Virus Type 1. PLoS ONE. 2013;8:e62649. doi: 10.1371/journal.pone.0062649.

[15] Cologna R., Armstrong P.M., Rico-Hesse R. Selection for Virulent Dengue Viruses Occurs in Humans and Mosquitoes. J. Virol. 2005;79:853–859. doi: 10.1128/JVI.79.2.853-859.2005.

[16] Steinhauer, D. A., Domingo, E. & Holland, J. J. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. Gene 122, 281–8 (1992).

[17] 11. Behura, S. K. & Severson, D. W. Nucleotide substitutions in dengue virus serotypes from Asian and American countries: insights into intracodon recombination and purifying selection. BMC Microbiol 13, 37 (2013).

[18] Edward C Holmes, S.Susanna Burch, The causes and consequences of genetic variation in dengue virus, March 2000, Trends in Microbiology 8(2):74-7

[19] Kolekar, Kale & Kulkarni-Kale (2012) Kolekar P, Kale M, Kulkarni-Kale U. Alignment-free distance measure based on return time distribution for sequence analysis: applications to clustering, molecular phylogeny and subtyping. Molecular Phylogenetics and Evolution. 2012;65:510–522. doi: 10.1016/j.ympev.2012.07.003.

[20] Holmes & Twiddy (2005) Holmes EC, Twiddy SS. The origin, emergence and evolutionary genetics of dengue virus. Infection, Genetics and Evolution. 2005;3:19–28.

[21] Martin et al. (2005) Martin D, Posada D, Crandall K, Williamson C. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Research and Human Retroviruses. 2005;21:98–102. doi: 10.1089/aid.2005.21.98.

[22] Watts DM, Burke DS, Harrison BA, Whitmire RE, Nisalak A. Effect of temperature on the vector efficiency of Aedes aegypti for dengue 2 virus. Am J Trop Med Hyg. 1987;36:143–152.

[23] Rohani A, Wong YC, Zamre I, Lee HL, Zurainee MN. The effect of extrinsic incubation temperature on development of dengue serotype 2 and 4 viruses in Aedes aegypti (L.) Southeast Asian J Trop Med Public Health. 2009;40:942–950.

[24] Lambrechts L, Paaijmans KP, Fansiri T, Carrington LB, Kramer LD, Thomas MB, Scott TW. Impact of daily temperature fluctuations on dengue virus transmission by Aedes aegypti. Proc Natl Acad Sci USA. 2011;108:7460–7465.

[25] Turell MJ, Lundstrom JO. Effect of environmental temperature on the vector competence of Aedes aegypti and Ae. taeniorhynchus for Ockelbo virus. Am J Trop Med Hyg. 1990;43:543–550

AUTHOR PROFILE

Dr.T.Ruban Deva Prakash has 24 years experience in teaching and research. He is the Managing Director of Effyies Smart Technologies and heading R& D wing. He also serves as Principal at Govt. College of Engineering, Munnar. He is a researcher in the power sector and has published 54 research papers in international journals, produced 7 Ph.Ds and serves as consultant to various firms.

Achshah R M is the CEO of Effyies Smart Technologies and heading Product Development wing. She is pursuing dual degree B. Tech (Artificial Intelligence) at Anna University, Chennai, and BSc (Applied AI) at IU International University, Germany. She is a certified data scientist and interested in research in the area of AI in Energy Sector and Industrial Automation. She has published four research papers in reputed international Journals in this area and successfully completed six industrial projects for her credit