TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 14856 - 14860

Spatio-Temporal Dengue Prediction System using Hybrid Algorithm

Sandra Johnson[1], A Jasmine Gilda[2], T Sethukarasi[3]
[1]Department of CSE, R.M.K. Engineering College, Kavaraipettai, India. Email: sjn.cse@rmkec.ac.in
[2]Department of CSE, R.M.K. Engineering College, Kavaraipettai, India. Email: ajg.cse@rmkec.ac.in
[3]Department of CSE, R.M.K. Engineering College, Kavaraipettai, India. Email: tsk.cse@rmkec.ac.in

Article Info
Volume 82
Page Number: 14856 - 14860
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 28 February 2020*

Abstract
A dengue outbreak is one of the major causes of the death worldwide. Many research works are being carried out to prevent and control the spread of the disease. Very few research works have been carried out to develop an efficient Spatio – Temporal prediction model for dengue outbreaks. Our work develops a hybrid prediction model using Genetic algorithm and Support Vector Machines which helps government and health workers to accurately identify regions and time of dengue outbreak. Genetic algorithm is used to identify relevant features for the dengue prediction model developed by the Support Vector Machines. This helps to mitigate the effects of dengue outbreaks and prevent them.

Keywords: *Dengue prediction, Genetic Algorithms (GA), Hybrid Algorithm, Spatio-Temporal, Support Vector Machines (SVM)*

I. INTRODUCTION

Dengue fever, a mosquito transmitted viral infection disease is an epidemic in the Southern Indian state of Tamil Nadu with 12,500 cases reported between March and August 2018. A severe form of the disease, the Dengue Haemorrhagic Fever (DHF) caused 65 deaths in the state last year, the highest in the country. The existing methods of medicine, alternate therapies and social responses to vector control are on the rise. However research-based scientific approaches to predict and mitigate dengue in the state have evolved little over the past decade. Hence, a Dengue early warning system using a hybrid algorithm is developed to predict the location and time of dengue outbreak that can help in the prevention and eradication of dengue in the state.

The major objectives of this paper are to design a predictive model using the spatial, temporal and climatic data to predict the outbreak of dengue. The work also designs a monitoring tool with help of a smart-phone based system to assist in preventive maintenance of dengue by health workers. It develops a data analytics mechanism using a hybrid algorithm, to identify the different geographical areas that are highly sensitive to dengue outbreaks and identify the factors that are favourable for the dengue outbreaks to assist health workers and government officials to take preventive measures to control the outbreak of dengue in Tamil Nadu. It also develops a surveillance system that can assist the government officials to plan the vector control strategies in advance and educate the community to mitigate dengue outbreak.

II. LITERATURE REVIEW

Atique et. al. [1] have used two space-time scan statistical analysis like geographically weighted regression models and ordinary least squares to identify the region and time of dengue outbreak in Swat, Pakistan. Yoon et. al. [2] has used Booted Regression Trees to study the relationships between land use factors and dengue cases in Malaysia. They have showed that human settlement and non-agricultural areas determine the dengue cases. M.H. Hasim et. al. [3] have used Geographical Information System and spatial statistical tools like spatial auto-correlation and average nearest neighbour analysis to identify dengue hotspots. Claude Flamand et. al. [4] have used contextual sequential pattern extraction techniques to identify the correlation of temperature, humidity and rainfall with dengue. The presence of dengue in Ecuador is

January - February 2020
ISSN: 0193 - 4120 Page No. 14856 - 14860

predicted by the authors in [5,6] by multivariate logistic regression. Relationship between weather and dengue in Thailand is developed by Wongkoon et. al. [7] using Spearman correlation analysis and Poisson regression analysis with time-series. The work by Bipin Kumar Acharya et. al. [8] have assessed and mapped spatio-temporal patterns in dengue fever in Nepal.

Chiung Ching Ho et. al. [9] have used Support Vector Machines to develop a prediction model using weather, population and land use factors to predict dengue outbreaks in Malaysia with an accuracy of 88.62%. The authors in [10] have used Susceptible-Infected-Recovered model together with Bayesian maximum entropy framework to predict the space-time spread of the dengue disease in Taiwan. Their prediction model will help the government to identify, control and prevent spread of dengue. Sumiko Anno et. al. [11] in their work have used Multi-satellite remote sensing data on rainfall, humidity and temperature to predict the dengue outbreaks in Srilanka. Srinivasa Rao Mutheneni et. al. [12] have developed a climate-based dengue prediction model for India using bivariate Pearson correlation coefficient test. Valson JS et. al. [13] has done a correlation analysis with climate variables using SatScan software. Zhu G et. al. [14] have use a machine learning algorithm Markov chain Monte Carlo method to predict the spatio-temporal patterns of disease transmission.

Dengue outbreak in Singapore is predicted and the risk is ranked in their work by Ong J et. al.[15] using Random Forest. Baquero OS et. al. [16] have used Artificial neural networks, Naïve model, Generalized Additive Models (GAM), seasonal autoregressive integrated moving average models and spatiotemporal Bayesian models for dengue prediction. They have shown that GAM has predicted large dengue epidemics efficiently for São Paulo city.Liu Keke et. al. [17] has used SatScan, ArcGis and Maxent software to perform global spatial autocorrelation. The authors [18] have done a detailed survey of the dengue epidemics prediction in their work.

Chen et. al. [19] have used least absolute shrinkage and selection operator (LASSO) based regression models for forecasting dengue. Wei Sun [20] et. al. have used SatScan software with Bernoulli method for spatio-temporal clustering analysis. Sandra et. al. [21] in their work have used a combination of genetic algorithm and SVM in their hot method prediction for compiler optimization.

III. METODOLOGY

A. Genetic Algorithms

All features are represented as either 0/1 in the feature vector which becomes the chromosomes. The initial population consists of bit vectors of all the chromosomes. The chromosomes with good fitness values are carried forward to the next generation. This evolution process is done the basic genetic operators: mutation and crossover.

B. Support Vector Machines

Support Vector Machines (SVM) is one of the supervised machine learning models. This algorithm constructs a hyper-plane separating the data into two categories. The region where the test data lies is used for the classification.

C. Proposed Hybrid System

The proposed system develops a predictive model based on the collected dengue data from geographically different regions as shown in figure 1. The collected data is pre-processed and relevant features are identified which is then used to develop a dengue predication model using a hybrid algorithm. This hybrid algorithm uses genetic algorithm and support vector machines to improve the prediction accuracy.

Initially, the feature set is given as an input to the genetic algorithm which identifies features that are relevant to the dengue prediction system. The feature set contains regional data, temporal data which includes the month and season when the dengue outbreak is found along with the other features for dengue prediction.

The feature vector constructed by the genetic algorithm uses the prediction model developed by the SVM as the fitness function. Each feature vector thus constructed is used to develop a SVM based prediction model using the ‘leave-one-out’ cross-validation technique in the dataset. The fitness value is the average of the precision value of all the prediction models constructed.

The prediction model is finally constructed using

January - February 2020
ISSN: 0193 - 4120 Page No. 14856 - 14860

the supervised machine learning algorithm SVM. The testing data set is given as an input to the prediction model to find out if the data region and month will have a dengue outbreak or not.

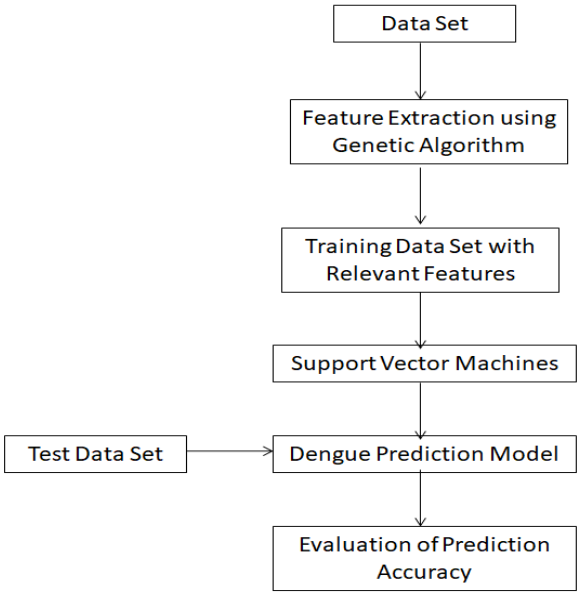

Fig. 1. A Hybrid Dengue Prediction System

IV. EXPERIMENTAL RESULTS

This work calculates the precision recall and the prediction accuracy of the dengue prediction model. Precision gives the ratio between the true positives and the actual results while recall is the ratio of the true positives with the predicted results. The prediction accuracy is calculated as the ratio between the true positives and true negatives with the total.

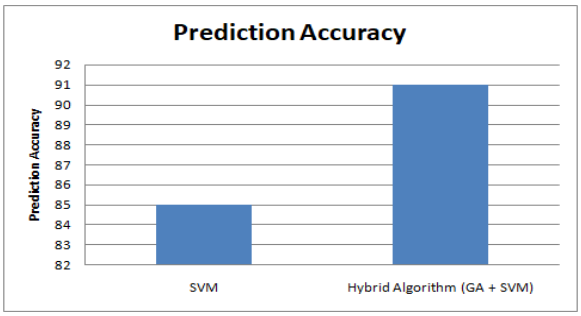

Fig. 2 : Prediction Accuracy

As shown in the Figure 2 the prediction accuracy of the SVM based classifier is 85% and that of the hybrid algorithm is 91%. The SVM based dengue prediction model constructed using the features selected by the genetic algorithm has given better prediction accuracy than the basic SVM classifier.

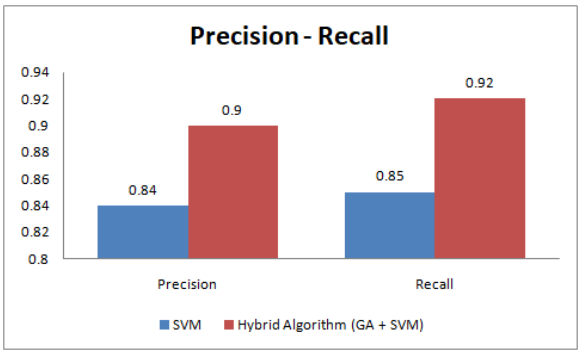

Fig. 3: Precision-Recall

The figure 3 shows precision and recall value of the SVM based prediction model and the hybrid dengue prediction model. This analysis also shows that our hybrid algorithm performs better when compared to the SVM.

V. CONCLUSION

Thus the proposed methodology predicts the outbreak of dengue with good precision and recall. The prediction system predicts the spatio-temporal data of dengue outbreaks i.e. the time and region where it is most likely to have dengue. This will help the government and health workers to prevent the spread of the disease. The work can be further studied with more machine learning algorithms to improve the precision further.

REFERENCES

[1] Atique S, Chan TC, Chen CC, Hsu CY, Iqtidar S, Louis VR, Shabbir SA, Chuang TW, "Investigating spatio-temporal distribution and diffusion patterns of the dengue outbreak in Swat, Pakistan", Journal of Infection and Public Health 11(4), pp. 550 – 557, 2018.

[2] Yoon Ling Cheong, Pedro J. Leitao, Tobia Lakes, "Assessment of land use factors associated with dengue casesin Malaysia using Boosted Regression Trees", Spatial and Spatio-temporal Epidemiology, 10, pp. 75 – 84,2014.

[3] Mohd Hazrin Hasim, Tee GuatHiong, Mohd Hatta Abdul Mutalip, Mohd Amierul Fikri Mahmud, Noor Aliza Lodz, NorzawatiYoep and Faizah Paiwai, "Spatial density of dengue incidence: A case study of a dengue outbreak in Seksyen 7, Shah Alam", International Journal of Mosquito Research, 5(2), pp. 09 - 14,2018.

[4] Claude Flamand Mickael Fabregue Sandra Bringay Vanessa Ardillon Philippe Quénel Jean-Claude Desenclos Maguelonne

January - February 2020
ISSN: 0193 - 4120 Page No. 14856 - 14860

Teisseire, "Mining local climate data to assess spatiotemporal dengue fever epidemic patterns in French Guiana", Journal of the American Medical Informatics Association, Volume 21, Issue e2, pp. e232 – e240,2014.

[5] Anna M Stewart-Ibarra, Angel G Munoz, Sadie J Ryan, Efrain Beltran Ayala, Mercy J Borbor-Cordova, Julia L Finkelstein, Raul Mejia, Tania Ordonez, G Cristina Recalde-Coronel and Keytia Rivero, "Spatiotemporal clustering, climate periodicity, and social-ecological risk factors for dengue during an outbreak in Machala, Ecuador, in 2010", BMC Infectious Diseases, 14:610, 2014.

[6] Catherine A. Lippi , Anna M. Stewart-Ibarra, Angel G. Munoz, Mercy J. Borbor-Cordova, Raul Mejia, Keytia Rivero, Katty Castillo, Washington B. Cárdenas and Sadie J. Ryan, "The Social and Spatial Ecology of Dengue Presence and Burden during an Outbreak in Guayaquil, Ecuador, 2012", International Journal Environmental Research and Public Health, 2018.

[7] Wongkoon, S, Jaroensutasinee, M and Jaroensutasinee, K., "Spatio-temporal climate-based model of dengue infection in Southern, Thailand", Tropical Biomedicine, 33(1), pp. 55 – 70, 2016.

[8] Bipin Kumar Acharya, Chun Xiang Cao, Tobia Lakes, Wei Chen and Shahid Naeem, "Spatiotemporal analysis of dengue fever in Nepal from 2010 to 2014", BMC Public Health, 16:849, 2016.

[9] Chiung Ching Ho, Choo-Yee Ting and Dhesi Baha Raja, " Using Public Open Data to Predict Dengue Epidemic: Assessment of Weather Variability, Population Density, and Land use as Predictor Variables for Dengue Outbreak Prediction using Support Vector Machine", Indian Journal of Science and Technology, Vol. 11(4), 2018.

[10] Hwa-Lung Yu , Jose M. Angulo , Ming-Hung Cheng , Jiaping Wu , George Christakos, "An online spatiotemporal prediction model for dengue fever epidemic in Kaohsiung(Taiwan)", Biomedical Journal, pp. 428 - 440, 2014.

[11] Sumiko Anno, Keiji Imaoka, Takeo Tadono, Tamotsu Igarashi, Subramaniam Sivaganesh, Selvam Kannathasan, Vaithehi Kumaran and Sinnatham, "Assessing the Temporal and Spatial Dynamics of the Dengue Epidemic in Northern Sri Lanka using Remote Sensing Data, GIS and Statistical Analysis", Geophysics and Remote Sensing, 3:4, pp. 135, 2014.

[12] Srinivasa Rao Mutheneni, Andrew P Morse, Cyril Caminade and Suryanaryana Murty Upadhyayula, "Dengue burden in India: recent trends and importance of climatic parameters", Emerging Microbes & Infections, 6, pp. e70, 2017.

[13] Valson JS, Soman B., "Spatio-temporal clustering of dengue cases in Thiruvananthapuram district, Kerala", Indian Journal of Public Health, 61, pp. 74 - 80, 2017.

[14] Zhu G, Liu J, Tan Q, Shi B, "Inferring the Spatio-temporal Patterns of Dengue Transmission from Surveillance Data in Guangzhou, China", PLoS Neglected Tropical Diseases, 10(4) , pp. e0004633, 2016.

[15] Ong J, Liu X, Rajarethinam J, Kok SY, Liang S, Tang CS, et al., "Mapping dengue risk in Singapore using Random Forest", PLoS Neglected Tropical Diseases, 12(6) , pp. e0006587, 2018.

[16] Baquero OS, Santana LMR, Chiaravalloti-Neto F, "Dengue forecasting in São Paulo city with generalized additive models, artificial neural networks and seasonal autoregressive integrated moving average models", PLoS ONE 13(4):e0195065, 2018.

[17] Liu Keke, Sun Jimin, Liu Xiaobo, Li Ruiyun, Wang Yiguan, Lu Liang, Wu Haixia, Gao Yuan, Xu Lei, Liu Qiyong, "Spatiotemporal patterns and determinants of dengue at county level in China from 2005-2017", International Journal of Infectious Diseases, 2018.

[18] P. Siriyasatien, S. Chadsuthi, K. Jampachaisri, and K. Kesorn, "Dengue Epidemics Prediction: A Survey of the State-of-the-Art Based on Data Science Processes", IEEE Access, Volume 6, pp. 53757 – 53795, 2018.

[19] Yirong Chen , Janet Hui Yi Ong, Jayanthi Rajarethinam, Grace Yap, Lee Ching Ng and Alex R. Cook, "Neighbour-hood level real-time forecasting of dengue cases in tropical urban Singapore", BMC Medicine,16:129, 2018.

[20] Wei Sun, Ling Xue and Xiaoxue Xie, "Spatial-temporal distribution of dengue and climate characteristics for two clusters in Sri Lanka from 2012 to 2016", Scientific REPORTS, 7:12884, 2017.

[21] Sandra Johnson and S. Valli, "Feature Subset Selection for Hot Method Prediction using Genetic Algorithm wrapped with Support Vector Machines", Journal of Computer Science, Vol. 7, No. 5, pp. 707-714, 2011.

AUTHORS PROFILE

Dr. Sandra Johnson is currently working as a Professor in the department of Computer Science & Engineering, R.M.K. engineering College, Chennai, Tamil Nadu, India. She received her Ph.D., from the

January - February 2020
ISSN: 0193 - 4120 Page No. 14856 - 14860

Faculty of Information & Communication Engineering, Anna University , Chennai, Tamil Nadu, India. Her area of research includes Compiler Optimization, Machine Learning, and Data Analytics.

Ms. A. Jasmine Gilda is working as Assistant Professor in the department of Computer Science and Engineering at R.M.K. Engineering College, Chennai. She received hers Masters degree in Computer Science and Engineering from Anna University, Chennai in 2012. Her research interest includes Medical Imaging, Cloud Computing and Blockchain. She is a certified as Java Programmer through the Digital Skills Readiness Program by Wipro. Acted as mentor in the Smart India Hackathon - 2018 held at Udaipur.

Dr. T.Sethukarasi is currently working as a Professor and Head in the department of Computer Science & Engineering, R.M.K. engineering College, Chennai, Tamil Nadu, India. She received her Ph.D., from the Faculty of Information & Communication Engineering , Anna University , Chennai, Tamil Nadu, India. Her area of research includes Data Mining, Soft Computing, Big data Analytics, Sensor Networks and Network Security. She has published more than 25 papers in the International Journals and Conferences. She is the recognized Research Supervisor of Anna University, Chennai and currently guiding 9 Ph.D scholars'. She produced 2 Ph.D., scholars. She is a life member in IAENG, IACSIT, IEI and ISTE.