



Machine learning for dengue outbreak prediction: An outlook

Naiyar Iqbal
Research Scholar

Dept. of Computer Science & Information Technology
Maulana Azad National Urdu University
Hyderabad, Telangana, India

Mohammad Islam
Assistant Professor

Dept. of Computer Science & Information Technology
Maulana Azad National Urdu University
Hyderabad, Telangana, India

Abstract: Dengue is a standout amongst the most well-known viral sicknesses in people. More than 33% of the total populace of world is under risk [9], including many cities of India. Timely prediction of dengue can save person's life by alerting them to take proper diagnosis and care. Prediction of infectious disease, such as Dengue, is a challenging task and most of the prediction methods are still in their infancy. Microarray and RNA-Seq data have been widely deployed for developing predictive model of various dengue. In this project, we propose to develop a machine learning model to predict Dengue. We will take various machine learning classifiers ranging from simple classifiers, like Decision Tree, Naïve Bayes, Model Tree, to complex algorithms such as Support Vector Machines, Neural Networks, Gene Expression Programming, Genetic Programming and ensemble classifiers. The algorithm giving the highest prediction accuracy will be considered for the development of Dengue Prediction Tool. We also propose to develop a novel ensemble classifier for Dengue outbreak prediction.

Keywords: Dengue fever, Machine learning algorithm, Prediction, Classification, clinical symptoms, genes

I. INTRODUCTION

Dengue (alternately known as Dengue fever or Break-bone fever) is the mostly hurriedly diffusion mosquito-borne viral disease in the world. Dengue diseases is brought about by four firmly related infections of viruses known as- DEN-1, DEN-2, DEN-3, and DEN-4. The DEN-5 has been introduced in 2013.

The initial four infections are called serotypes in light of the fact that each has distinctive associations with the antibodies in human blood serum. The four dengue infections are comparative (they share around 65% of their genomes), however even inside a solitary serotype, there is some hereditary varieties.

There are three broad classifications of dengue fever stages that is Dengue fever (DF), Dengue Hemorrhagic Fever (DHF), and Dengue Shock Syndrome (DSS). These are basically stages of dengue infection from infancy stage to critical stage shown in figure 1.

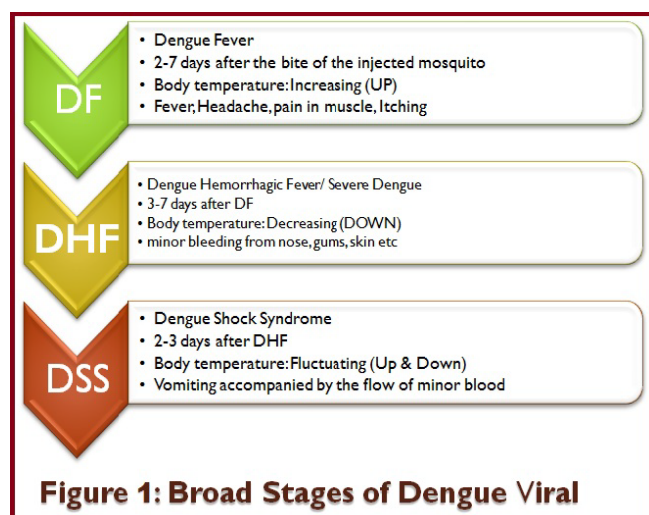


Figure 1: Broad Stages of Dengue Viral

Figure 1: Broad stages of Dengue Viral

II. WORK FLOW OF DENGUE OUTBREK PREDICTION

- Data pre-processing and normalization for minimizing noises level in the given data sets.
- Identification of differentially expressed genes (up-regulated or down-regulated) that may participate in Dengue outbreak. Looking for the possibility to incorporate other features related to Dengue disease for training the classifier.
- Doing a comprehensive evaluation of well-studies machine learning classifiers for Dengue disease prediction.
- Development of novel ensemble classifiers for Dengue outbreak predictions.
- Development of web-based tool for the classification of Dengue patient based on selected Dengue disease parameters.

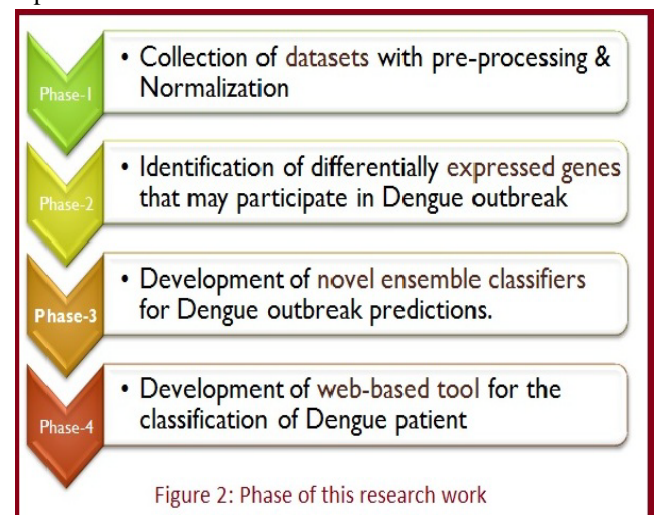


Figure 2: Phase of this research work

Figure 2: Phases of research work

III. LITERATURE SURVEY

In this review research paper, we explore different related research work of the researchers and try to show it in tabular form. Out of many research works, we conclude thirty research paper in a tabular form (Table 1) in which it shows timeline, methodologies or tool, datasets, conclusion, location and research gap or future scope of that research work.

Table 1: Related research work in Dengue viral predication

Time-line	Reference/Author(s)	Methodologies/Tools	Attribute(s) Used/Conclusion	Gap/Future Work	Location/Remarks
1995	[19] Focks, D. A., Daniels, E., Haile, D. G., & Keesling, J. E.	Stochastic Simulation Models: 1. CIMSiM (Container-Inhabiting Mosquito Simulation Model) 2. DENSiM (Dengue Simulation Model)	Climate Dataset: -Weather data Conclusion: -conclude example of simulation (entomologic, demographic, and epidemiologic)	Improvement of expectation for dengue investigation	Locations: -Caribbean, -Central America, -South America, -Southeast Asia
2001	[27] Hopp, M. J., & Foley, J. A.	CIMSiM (Container-Inhabiting Mosquito Simulation Model)	Climate Dataset: Development Rates, Survival Rates, Global-scale, Long-term Average Climate data monthly-mean temperature, Precipitation, Cloud Cover	Investigation of interannual changeability in mosquito flow	Locations: -USA -Asia -and many other locations
2002	[26] Hartley, L. M., Donnelly, C. A., & Garnett, G. P.	Mathematical Approach	Seasonal Dataset: -Recruitment rate, -Biting rate, -Vector Mortality Conclusion: 2-serotype model is considered	4-serotype model is under development	Locations: -Thailand -Bangkok -Malaysia
2004	[15] De Paula, S. O., & Fonseca, B. A. L. D.	-ELISA (Enzyme-Linked ImmunoSorbent Assay) Method -RNA (RiboNucleic Acid) Extraction	Five serological Tests Dataset: 1. HI (Hemagglutination-Inhibition) 2. CF (Complement Fixation) 3. NT (Neutralization Test) 4. MAC-ELISA (IgM Antibody Capture Enzyme-Linked Immunosorbent Assay) 5. Indirect Immunoglobulin G ELISA	molecular systems may soon accept an imperative part in dengue finding	Location: Brazil
2005	[31] Ibrahim, F., Taib, M. N., Abas, W. A. B. W., Guan, C. C., & Sulaiman, S.	-ANN -MatLab Toolbox	Neural Network Dataset: Input Neurons=9 Hidden Neurons=5 Output Neuron=1 Iteration=25 Accuracy=90% Error=10% Clinical and Epidemiological Data: Total Patients 252: (4 DF and 248 DHF Patients). 90% prediction accuracy 10% prediction error	Extended application to assist the clinician	Location: Malaysia
2007	[1] Aburas, H. M.	Statistically Developed Index: • HI (House Index) • CI (Container Index) • BI (Breteau Index) • ABURAS Index Poisson Distribution method	ABURAS Index is use as a new statistical indexing , utilizing Poisson distribution on the basis of the collection of vector populace	Easy to extend for the prediction of expected amount of confirm dengue cases	Location: Jeddah (Saudi Arabia)
2007	[18] Fink, J., Gu, F., Ling, L., Tolfvenstam, T.,	-ELISA Method	Gene expression: (IP-10 and I-TAC, A549, MG-132 and ALLN)	Suggestion for drug discovery by use of	Location: Singapore

	Olfat, F., Chin, K. C., ... & Vasudevan, S. G.		Conclusion: Unbiased gene expression investigation has distinguished new host genes connected with dengue disease, which they have approved in functional studies.	therapeutically to improves clinical result	
2008	[68] Wu, Y., Lee, G., Fu, X., & Hung, T.	Wavelet, SVM, SVR, GA	Climate/Weather Dataset: -Rainfall, Humidity, Temperature	More robust algorithm	Location: Singapore
2008	[58] Tanner, L., Schreiber, M., Low, J. G., Ong, A., Tolfvenstam, T., Lai, Y. L., ... & Simmons, C. P.	Decision Tree (C4.5)	Clinical Test Dataset: -Platelet Count -WBC/Lymphocyte -Body Temperature -Haematocrit count or -Neutrophil count No of Patient=1200 • 364-Dengue infected • 836-Non-dengue Overall Error Rate=15.7% (after k-fold validation) Conclusion: Decision algorithm could be useful	Need a huge multi-centre investigation for global utilization	Location: Singapore
2008	[10] Butt, N., Abbassi, A., Munir, S. M., Ahmad, S. M., & Sheikh, Q. H.	PCR (Polymerase Chain Reaction)	Clinical Symptoms Dataset: -Hematological -Biochemical Total Patients 104: Grade-I DHF=66 Grade-II DHF=34 Grade-III DHF=4 Grade-IV DHF=3	Recommendation of early diagnosis based on investigation	Location: Karachi
2008	[12] Choudhury, Z. M., Banu, S., & Islam, A. M.	Time series Analysis: -SARIMA (Seasonal Autoregressive Integrated Moving Average) Models	Time series Data: (January 2000 to September 2007)	More information will be added for the predication	Locations: Dhaka, Bangladesh
2008	[14] De Kruif, M. D., Setiati, T. E., Mairuhu, A. T., Koraka, P., Aberson, H. A., Spek, C. A., ... & van Gorp, E. C.	-SPSS (Statistical Package for the Social Sciences) version 13.0 -mRNA (messenger RNA) Analysis	Gene Expression Dataset: (NFKB1, NFKB2, TNFR1, IL1B, IL8, and TNFA) (TLR7, TLR4R3, TLR1, TLR2, TLR4R4, and TLR4 co-factor CD14) DHF=56 Children	Improvement of gene array profiling in future	Location: Indonesian
2009	[11] Cetiner, B. G., Sari, M., & Aburas, H. M.	ANN	Climate/Weather Dataset: (Neural Network): @Input Layers=1 (with 3 parameters): -mean temperature -mean relative humidity -total rainfall @Output Layer=1 (with 1 parameter): - number of dengue confirmed cases @ Hidden Layer=1 (with 4 parameters): - On the basis of trail & error, it chosen to be optimum with four processing components	To explore more parameters for prediction	Location: World wide Dataset Provider by: Singaporean National Environment Agency (NEA) 6 years records (January 2001 to April 2007)
2009	[13] Cordeiro, M. T., Braga-Neto, U., Nogueira, R. M. R., & Marques	ELISA Method	(IgM and IgG) 109 dengue infection patients: -59 patients: analyzed by linear discriminant analysis (LDA)	Offer the website interface with database and	Location: Brazil

	Jr, E. T.		-50 patients: Analyzed by independent test set	diagnosis tool in future	
2009	[41] Nascimento, E. J., Braga-Neto, U., Calzavara-Silva, C. E., Gomes, A. L., Abath, F. G., Brito, C. A., ... & Gil, L. H.	MAS 5.0 (MicroArray Suite) Software	Gene Expression Data: -Various genes Conclusion: Accuracy rates more than 95%	Validation and other feature will be addressed in future	Location: Brazil
2009	[52] Rissino, S., & Lambert-Torres, G.	Rough Set Theory	Symptoms Dataset: -Headache -Vomiting -Temperature Conclusion: Rough set made more familiar in recent year for dengue prediction	Only on the basis of patient attributes cannot be classify dengue or without dengue	Location: Brazil
2010	[23] Gomes, A. L. V., Wee, L. J., Khan, A. M., Gil, L. H., Marques Jr, E. T., Calzavara-Silva, C. E., & Tan, T. W.	SVM RBF	Gene Expression Data: Genes=12 (MYD88, TLR3/7/9, RIG1, IRF3/7, CLEC5A, IFN- $\alpha/\beta/\gamma$, MDA5) Most effective genes: MYD88, and TLR7 RBF kernel function (c value = 1.0 and C value = 10) Patients=28 • 15 DF • 13 DHF	Limitation: • Small size of datasets • Sampling bias	Location: Brazil
2010	[2] Aburas, H. M., Cetiner, B. G., & Sari, M.	ANN	Climate Dataset: -mean temperature -relative humidity - total rainfall Total Dataset=14,209 (Dengue confirm cases)	Explore additional parameters	Location: Malaysia, Singapore, others Dataset Provider by: Singaporean National Environment Agency (NEA)
2010	[20] Gadkari, R. A., & Srinivasan, N.	-PDB (Protein Data Bank) codes -cryoEM (Cryo-electron Microscopy)	Protein-Protein interaction: 1. E glycoprotein (Envelop Protein) 2. M protein (Membrane Protein)	More effective antiviral exploit	Location: India
2010	[25] Guzman, M. G., Halstead, S. B., Artsob, H., Buchy, P., Farrar, J., Gubler, D. J., ... & Nathan, M. B.	SVM RBF	Protein-Protein interaction: Structural Proteins=3 -Capsid (C), Membrane (M) and Envelope (E). Non- Structural Proteins=7: -(NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5)	To the development of new tool with scale up	Locations: Southeast Asia, Pacific, Americas
2010	[38] Madhu, G., Reddy, G. S., & Kiranmai, C.	Rough Set Analysis	Symptoms Attributes Dataset: -Temperature -Headache -Vomiting	Integration of rough set analysis with Fuzzy set and Neural network	Location: Hyderabad, India
2010	[46] Rachata, N., Charoenkwan, P., Yooyativong, T., Chamnongthal, K., Lursinsap, C., & Higuchi, K.	ANN	Climate/Weather Dataset: -temperature -rainfall -relative humidity (Taken Probability and maximum, minimum, average intervals by using a	Include other feature selection method like Hidden Markov	Location: Thailand

			histogram) Result: Highest accuracy 85.92%	Models.	
2010	[47] Rani, M. U., Devi, M. K., Mamatha, D. M., Seshadri, R., & Avulapti, Y. K.	ETL (Extract, Transform, Load) Tool	Clinical Dataset: -Blood sample -Number of confirm cases -Number of Death cases Conclusion: Created a data warehouse between period of 2000-2009	Extension of work	Locations: Andhra Pradesh
2011	[4] Al-Muhandis, N., & Hunter, P. R.	Comprehensive Meta-analysis Statistically Developed Index: • HI (House Index) • CI (Container Index) • BI (Breteau Index)	Conclusion: With the confirmation as of now accessible it is unrealistic to state what sorts of instructive modalities are best.	More investigation	Locations: Hundreds of countries from different continent
2011	[5] Althouse, B. M., Ng, Y. Y., & Cummings, D. A.	Three Methods used: 1. Linear regression (Step-down) 2. Generalize Boosted regression 3. negative binomial Regression Other Two method used: 1. Logistic regression 2. SVM	- Conclusion: Linear Model found superior than other model by AIC step-down At Bangkok, the model found: $r^2 \sim 0.943$, and correlation=0.869 At Singapore, the model found: $r^2 \sim 0.948$, and correlation of 0.931 - SVM Model performed better logistic regression (in both locations): Area Under Curve (AUC) for the SVM models with the use of 75th percentile cutoff =0.906 (in Singapore) and cutoff=0.960 (in Bangkok)	New disease (like chikungunya) may challenge to this model in future	Locations: -Singapore -Bangkok
2011	[24] Gomide, J., Veloso, A., Meira Jr, W., Almeida, V., Benevenuto, F., Ferraz, F., & Teixeira, M.	Spatio-temporal Analysis	Twitter Dataset: Four parameters: volume, location, time and content Conclusion: Twitter can be used to anticipate, locally and momentarily, dengue distress by cluster method	Extension of Method to other countries	Location: Brazil
2012	[8] Brasier, A. R., Ju, H., Garcia, J., Spratt, H. M., Victor, S. S., Forshey, B. M., ... & Rocha, C.	CART Method Random Forest Method	Symptoms based Dataset: - Perform 10 trials with 10-fold cross-validation Result: Average accuracy=84.0% (for DF) and 84.6% (for DHF) - AUC=0.87	More validation requires on independent population	Location: Venezuela
2012	[9] Buczak, A. L., Koshute, P. T., Babin, S. M., Feighner, B. H., & Lewis, S. H.	Fuzzy Association Rule Mining	Clinical, Meteorological, Weather, and Socio-political data: Results: -Positive Predictive value=0.686, -Negative Predictive value=0.976, -Sensitivity=0.615, -Specificity=0.982.	Enhancing the generality of the method	Location: Peru
2012	[33] Karim, M. N., Munshi, S. U., Anwar, N., & Alam, M. S.	ANOVA (Analysis of Variance)	Climate data: -Humidity, Rainfall, Minimum and Maximum temperature Result: Predicted Dengue disease (≥ 200 Cases). AUROC curve= 0.89, 95% CI=0.890.98	More improvements in model to predict for each month outbreak in future	Locations: Dhaka city, Bangladesh
2013	[30] Huy, N. T., Thao, N. T. H., Ha, T. T. N., Lan, N. T. P., Nga, P. T. T., Thuy, T. T., ... & Huong, V. T. Q.	Decision based prediction	Clinical Dataset: DSF: 444 dengue patients -318 patients (72%) with a single shock episode, -126 (28%) with recurrent shock Result: (AUC, 0.73; sensitivity and specificity, 68%)	Decision system requires to be validated using many independent studies.	Locations: Vietnam



IV. METHODOLOGIES & TOOLS

Data mining is a process of discovering various models, summaries and derived values from a given collection of data. There are the following broad phases for data mining:

1. State the problem and formulate the hypothesis
2. Collection of data
3. Pre-process the data
4. Estimate the model
5. Interpret the model and draw conclusions

There are too many prominent methodologies and tools are used to for the prediction and experiment of the diseases.

4.1) Tree-based models:

Decision Tree Learning is one of the methods of classification for approximating discrete-esteemed target operations, in which the trained functions are expressed by a Decision tree. It is one of the most widely used and practical methods for inductive inference.

Classification decision tree model is used where the objective variable can take a limited set of qualities, implies when the predicted result is to the classification to which the data associated. Regression decision tree model is applied where the target variable can be take continuous values, means at the point when the predicted consequence can be viewed as a real value.

4.2) Neural Network models:

Artificial neural network has the capability to cognize experience or observation knowledge represented through intermediate unit association qualities, and can make such learning knowledge accessible for utilization. ANNs are processing device (algorithm or actual hardware) that are loosely modelled after the neural structure of the mammalian cortex but on much smaller scales.

4.3) Evolutionary based classifiers:

Evolutionary based classification is a subset of developmental calculation a hereditary populace based

meta-heuristic streamlining algorithm. It is persuaded by biological advancement, for example reproduction, mutation, recombination and selection. There are many powerful techniques based on evolutionary such as Genetic algorithm, Neuro-evolution, Gene expression etc.

4.4) Ensemble classifier:

Ensemble model is applied to the combination of multiple models for improvement of the accuracy rate of novel predictive model. It consolidates a progression of k-learned models (M1, M2, M3...Mk) with point of making an enhanced model M* (figure 3).

Diverse and independent learners can be generated using different learning algorithm, different hyper-parameters, different input representation, different training subsets, multi-experts, and multi-stage methodologies.

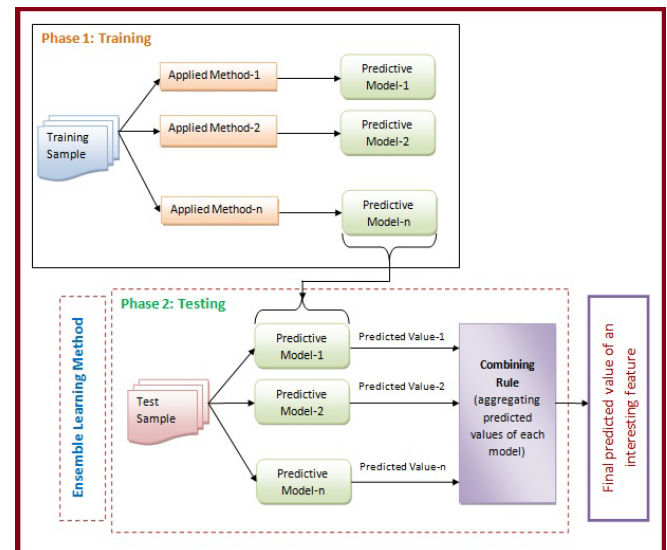


Figure 3: Ensemble Learning Model

V. DATA CLASSIFICATION

Based on the survey of this research, it is found that there are broadly four classification of dataset can be used for the dengue outbreak prediction. The classification is based on symptoms or clinical test (Table 2), climate factor or time series (Table 3), genes (Table 4) and protein (Table 5).

Table 2: List of Symptoms or Clinical Test involved in Dengue Viral

SYMPTOMS	SHORT DESCRIPTION	REFERENCE(S)
Date of fever	Month	[55]
Residence	Address	
Days	Number of days of fever	
Metallic Taste	Yes/No	[52] [16]
Joint/Muscle Pain	Yes/No	
Body Temperature/ Sudden fever	Fever	[10][16][52][55]
Severe Headache	Yes/No	
Nausea/Vomiting	Yes/No	

Diarrhea	Yes/No	[10][16][55]
Shock/ Hypotension	Fall in Blood Pressure (Yes/No)	[10] [16]
Fits/ Seizures	Yes/No	
Rash/Red Skin	Yes/No	[10][16][52]
Mouth and Nose Bleeding	Yes/No	[16]
Pleural effusion	Excess fluid	
Ascites	Abdominal Swelling	
Gastrointestinal Bleeding	Yes/No	
Altered level of Consciousness	State/level	
Itching	Yes/No	
Slow heart rate	Rate of heart beats	
Lymphocytes Count	Number of Lymphocytes	[8]
Platelet count (mm3)	Number of Platelets	[8][55]
Abdominal Pain	Yes/No	[55] [10]
Hemoglobin (gm/dl)	Range of Hemoglobin	
Hematocrit (%)	Range of Hematocrit	
WBC/ Leukocyte count 103/L	Number of White Blood Count	
Behind Eye Pain/ Retro orbital pain	Yes/No	[10]
Appetite/Anorexia	Yes/No	
Back or Body ache	Yes/No	
Depression	Low mood (Yes/No)	
Bilirubin (mg/dl)	Haematoidin test	
ALT (U/L)	Alanine Aminotransferase	
AST (U/L)	Aspartate Aminotransferase	
Gamma-GT (U/L)	Gamma-glutamyl transpeptidase	
Alkaline Phosphatase (U/L)	Alkaline Phosphatase	
PT (sec)	Prothrombin Time	
APTT (sec)	Activated Partial Thromboplastin Time	

Table 3: List of Climate factors & Time series involved in Dengue Viral

CLIMATE FACTOR	SHORT DESCRIPTION	REFERENCE(S)
Temperature	Environment temperature	[37]
Rain	Rain fall	
Wind Velocity	Speed	
Humidity	Percentage	
Period of months	Time Series	[12][65]

Table 4: List of Genes involved in Dengue Viral

GENES	SHORT DESCRIPTION	REFERENCE(S)
TNF-Gamma	Tumor Necrosis Factor	[8]
IL family (2/6/10)	Interleukin	
IP-10	Interferon gamma-induced protein 10	
MIP-1α	Macrophage Inflammatory Proteins	
IgG	Immunoglobulin G	[10]
IgM	Immunoglobulin M	
CD14	Cluster of differentiation 14	[14]
NFKB family (1/2)	Nuclear factor-kappaB	
TLR family (1/2/4/8/4R4)	Toll-like Receptor	
TNFA	Tumor Necrosis Factor α	
TNFR1	Tumor Necrosis Factor Receptor 1	
IL family (8/1B)	Interleukin	
CLEC5A	C-type lectin domain family 5 member A	[23]
IFN family (α/β)	Interferon alpha/Beta	
IRF family (3/7)	Interferon regulatory factor	
MYD88	Myeloid differentiation primary response gene 88	
IFN-Gamma	Interferon Gamma	[23] [8]
TLR family (7/9)	Toll-like Receptor	[23] [14]
RT-PCR	Reverse transcription polymerase chain reaction	[58]
MDA5	Melanoma Differentiation-Associated protein 5	[23] [42]

TLR3	Toll-like Receptor 3	[6]
RIG-1	Retinoic acid-inducible gene 1	
CYP9 P450	Cytochromes Protein	
CYP9J/ CYP9J26 family	Cytochromes genes	
ABCB4	ATP binding cassette subfamily B member 4	

Table 5: List of Protein involved in Dengue Viral

PROTEIN	SHORT DESCRIPTION	REFERENCE(S)
NS family (1/2A/2B/3/4A/4B/5)	Nonstructural protein 1	[25]
E Protein	Centromere-associated protein E	[50]

VI. CONCLUSION

Dengue is the fever which can't be identified without the indications for dengue. Dengue can begin from the fever which is mellow in circumstance yet after that it will prompt to deadly state of the casualty. Generally the symptoms effects keep going for one week however at sometimes, it begins re-emerging for 2-3 days.

As there are some prediction models are developed for dengue fever predication. The discipline is very infancy and much work has done, so more research work is needed.

So we are interested to design a new enhanced novel ensemble model for the prediction of dengue outbreak using combination of different machine learning models.

VII. FUTURE SCOPE

We would come up with three major outcomes in future:

- A list of attributes (set of genes which are differentially expressed over different samples) which are participating in Dengue disease prediction.
- A comprehensive evaluation of various machine learning classifiers will give us choice for best performing classifier for Dengue outbreak prediction.
- A novel ensemble model designed for the prediction of Dengue outbreak.

List of Abbreviations:

- DEN : Dengue
- NF : Dengue Fever
- DHF : Dengue Haemorrhage Fever
- DSS : Dengue Shock Syndrome
- RNA : Ribonucleic Acid
- MatLab : Matrix Laboratory
- ANN : Artificial Neural Network
- SVM : Support Vector Machine
- SVR : Support Vector Regression
- GA : Genetic Algorithm
- RBF : Radial Basis Function
- ELISA : Enzyme-linked immunosorbent Assay
- AUC : Area Under the Curve
- CART : Classification And Regression Tree
- ROC : Receiver Operative Characteristics
- AUROC : Area Under Receiver Operative Characteristics Curve

VIII. REFERENCES

[1] Aburas, H. M. (2007, August). ABURAS index: A statistically developed index for dengue-transmitting vector population prediction. In Proceedings of world academy of science, engineering and technology (Vol. 23, pp. 151-154).

[2] Aburas, H. M., Cetiner, B. G., & Sari, M. (2010). Dengue confirmed-cases prediction: A neural network model. *Expert Systems with Applications*, 37(6), 4256-4260.

[3] Ahmed, A. (2008). Modeling of dengue outbreak prediction in Malaysia: A comparison of Neural Network and Nonlinear Regression Model. *IEEE*,3(1036).

[4] Al-Muhandis, N., & Hunter, P. R. (2011). The value of educational messages embedded in a community-based approach to combat dengue fever: a systematic review and meta regression analysis. *PLoS Negl Trop Dis*, 5(8), e1278.

[5] Althouse, B. M., Ng, Y. Y., & Cummings, D. A. (2011). Prediction of dengue incidence using search query surveillance. *PLoS Negl Trop Dis*, 5(8), e1258.

[6] Bariami, V., Jones, C. M., Poupardin, R., Vontas, J., & Ranson, H. (2012). Gene amplification, ABC transporters and cytochrome P450s: unraveling the molecular basis of pyrethroid resistance in the dengue vector, *Aedes aegypti*. *PLoS Negl Trop Dis*, 6(6), e1692.

[7] Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., ... & Myers, M. F. (2013). The global distribution and burden of dengue. *Nature*, 496(7446), 504-507.

[8] Brasier, A. R., Ju, H., Garcia, J., Spratt, H. M., Victor, S. S., Forshey, B. M., ... & Rocha, C. (2012). A three-component biomarker panel for prediction of dengue hemorrhagic fever. *The American journal of tropical medicine and hygiene*, 86(2), 341-348.

[9] Buczak, A. L., Koshute, P. T., Babin, S. M., Feighner, B. H., & Lewis, S. H. (2012). A data-driven epidemiological prediction method for dengue outbreaks using local and remote sensing data. *BMC medical informatics and decision making*, 12(1), 1.

[10] Butt, N., Abbassi, A., Munir, S. M., Ahmad, S. M., & Sheikh, Q. H. (2008). Haematological and biochemical indicators for the early diagnosis of dengue viral infection. *J Coll Physicians Surg Pak*, 18(5), 282-5.

[11] Cetiner, B. G., Sari, M., & Aburas, H. M. (2009, May). Recognition of dengue disease patterns using artificial neural networks. In 5th International Advanced Technologies Symposium (IATS'09) (pp. 359-362).

[12] Choudhury, Z. M., Banu, S., & Islam, A. M. (2008). Forecasting dengue incidence in Dhaka, Bangladesh: A time series analysis.

[13] Cordeiro, M. T., Braga-Neto, U., Nogueira, R. M. R., & Marques Jr, E. T. (2009). Reliable classifier to differentiate primary and secondary acute dengue infection based on IgG ELISA. *PLoS one*, 4(4), e4945.

[14] De Kruif, M. D., Setiati, T. E., Mairuhu, A. T., Koraka, P., Abersson, H. A., Spek, C. A., ... & van Gorp, E. C. (2008). Differential gene expression changes in children with severe dengue virus infections. *PLoS Negl Trop Dis*, 2(4), e215.

[15] De Paula, S. O., & Fonseca, B. A. L. D. (2004). Dengue: a review of the laboratory tests a clinician must know to achieve a correct diagnosis. *Brazilian Journal of Infectious Diseases*, 8(6), 390-398.

[16] Degnue Fever [Online Available: https://en.wikipedia.org/wiki/Dengue_fever]

- [17] Fathima, A. S., Shameem, D., & Manimegalai, N. H. (2011). A review of data mining classification techniques applied for diagnosis and prognosis of the arbovirus-dengue. *IJCSI International Journal of Computer Science*, 8(6).
- [18] Fink, J., Gu, F., Ling, L., Tolfvenstam, T., Olfat, F., Chin, K. C., ... & Vasudevan, S. G. (2007). Host gene expression profiling of dengue virus infection in cell lines and patients. *PLoS Negl Trop Dis*, 1(2), e86.
- [19] Focks, D. A., Daniels, E., Haile, D. G., & Keesling, J. E. (1995). A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation, and samples of simulation results. *American Journal of Tropical Medicine and Hygiene*, 53(5), 489-506.
- [20] Gadkari, R. A., & Srinivasan, N. (2010). Prediction of protein-protein interactions in dengue virus coat proteins guided by low resolution cryoEM structures. *BMC structural biology*, 10(1), 1.
- [21] Gibbons, R. V., & Vaughn, D. W. (2002). Dengue: an escalating problem. *Bmj*, 324(7353), 1563-1566.
- [22] Githeko, A. K. (2012). Advances in developing a climate based dengue outbreak models in Dhaka, Bangladesh: challenges & opportunities. *The Indian journal of medical research*, 136(1), 7.
- [23] Gomes, A. L. V., Wee, L. J., Khan, A. M., Gil, L. H., Marques Jr, E. T., Calzavara-Silva, C. E., & Tan, T. W. (2010). Classification of dengue fever patients based on gene expression data using support vector machines. *PloS one*, 5(6), e11267.
- [24] Gomide, J., Veloso, A., Meira Jr, W., Almeida, V., Benevenuto, F., Ferraz, F., & Teixeira, M. (2011, June). Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In *Proceedings of the 3rd international web science conference* (p. 3). ACM.
- [25] Guzman, M. G., Halstead, S. B., Artsob, H., Buchy, P., Farrar, J., Gubler, D. J., ... & Nathan, M. B. (2010). Dengue: a continuing global threat. *Nature Reviews Microbiology*, 8, S7-S16.
- [26] Hartley, L. M., Donnelly, C. A., & Garnett, G. P. (2002). The seasonal pattern of dengue in endemic areas: mathematical models of mechanisms. *Transactions of the royal society of tropical medicine and hygiene*, 96(4), 387-397.
- [27] Hopp, M. J., & Foley, J. A. (2001). Global-scale relationships between climate and the dengue fever vector, *Aedes aegypti*. *Climatic change*, 48(2-3), 441-463.
- [28] Husin, N. A., & Salim, N. (2008, August). Modeling of dengue outbreak prediction in Malaysia: a comparison of neural network and nonlinear regression model. In *2008 International Symposium on Information Technology* (Vol. 3, pp. 1-4). IEEE.
- [29] Husin, N. A., Mustapha, N., Sulaiman, M. N., & Yaakob, R. (2012, September). A hybrid model using genetic algorithm and neural network for predicting dengue outbreak. In *2012 4th Conference on Data Mining and Optimization (DMO)* (pp. 23-27). IEEE.
- [30] Huy, N. T., Thao, N. T. H., Ha, T. T. N., Lan, N. T. P., Nga, P. T. T., Thuy, T. T., ... & Huong, V. T. Q. (2013). Development of clinical decision rules to predict recurrent shock in dengue. *Critical Care*, 17(6), 1.
- [31] Ibrahim, F., Taib, M. N., Abas, W. A. B. W., Guan, C. C., & Sulaiman, S. (2005). A novel dengue fever (DF) and dengue haemorrhagic fever (DHF) analysis using artificial neural network (ANN). *Computer methods and programs in biomedicine*, 79(3), 273-281.
- [32] Ibrahim, F., Taib, M. N., Sulaiman, S., & Abas, W. W. (2001). Dengue fever (DF) and dengue haemorrhagic fever (DHF) symptoms analysis from an expert system perspective. In *Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century. Proceedings. IEEE International* (pp. 212-215). IEEE.
- [33] Karim, M. N., Munshi, S. U., Anwar, N., & Alam, M. S. (2012). Climatic factors influencing dengue cases in Dhaka city: a model for dengue prediction. *The Indian journal of medical research*, 136(1), 32.
- [34] Lanciotti, R. S., Calisher, C. H., Gubler, D. J., Chang, G. J., & Vorndam, A. V. (1992). Rapid detection and typing of dengue viruses from clinical samples by using reverse transcriptase-polymerase chain reaction. *Journal of clinical microbiology*, 30(3), 545-551.
- [35] Lee, V. J., Chow, A., Zheng, X., Carrasco, L. R., Cook, A. R., Lye, D. C., ... & Leo, Y. S. (2012). Simple clinical and laboratory predictors of Chikungunya versus dengue infections in adults. *PLoS Negl Trop Dis*, 6(9), e1786.
- [36] Lima-Camara, T. N., Bruno, R. V., Luz, P. M., Castro, M. G., Lourenço-de-Oliveira, R., Sorgine, M. H., & Peixoto, A. A. (2011). Dengue infection increases the locomotor activity of *Aedes aegypti* females. *PLoS One*, 6(3), e17690.
- [37] Lu, L., Lin, H., Tian, L., Yang, W., Sun, J., & Liu, Q. (2009). Time series analysis of dengue fever and weather in Guangzhou, China. *BMC Public Health*, 9(1), 395.
- [38] Madhu, G., Reddy, G. S., & Kiranmai, C. (2010). Hypothetical description for intelligent data mining. *Int. J. Comput. Sci. Eng.* 2(7), 2349-2352.
- [39] Morens, D. M., & Fauci, A. S. (2008). Dengue and hemorrhagic fever: a potential threat to public health in the United States. *Jama*, 299(2), 214-216.
- [40] Nakhapakorn, K., & Tripathi, N. K. (2005). An information value based analysis of physical and climatic factors affecting dengue fever and dengue haemorrhagic fever incidence. *International Journal of Health Geographics*, 4(1), 1.
- [41] Nascimento, E. J., Braga-Neto, U., Calzavara-Silva, C. E., Gomes, A. L., Abath, F. G., Brito, C. A., ... & Gil, L. H. (2009). Gene expression profiling during early acute febrile stage of dengue infection can predict the disease outcome. *PloS one*, 4(11), e7892.
- [42] Nasirudeen, A. M. A., Wong, H. H., Thien, P., Xu, S., Lam, K. P., & Liu, D. X. (2011). RIG-I, MDA5 and TLR3 synergistically play an important role in restriction of dengue virus infection. *PLoS Negl Trop Dis*, 5(1), e926.
- [43] Oishi, K., Saito, M., Mapua, C. A., & Natividad, F. F. (2007). Dengue illness: clinical features and pathogenesis. *Journal of infection and chemotherapy*, 13(3), 125-133.
- [44] Phuong, C. X. T., Nhan, N. T., Kneen, R., Thuy, P. T. T., Van Thien, C., Nga, N. T. T., ... & Wills, B. (2004). Clinical diagnosis and assessment of severity of confirmed dengue infections in Vietnamese children: is the World Health Organization classification system helpful?. *The American journal of tropical medicine and hygiene*, 70(2), 172-179.
- [45] Potts, J. A., Gibbons, R. V., Rothman, A. L., Srikiatkachorn, A., Thomas, S. J., Supradish, P. O., ... & Kalayanaroop, S. (2010). Prediction of dengue disease severity among pediatric Thai patients using early clinical laboratory indicators. *PLoS Negl Trop Dis*, 4(8), e769.
- [46] Rachata, N., Charoenkwan, P., Yooyativong, T., Chamnongthai, K., Lursinsap, C., & Higuchi, K. (2008, October). Automatic prediction system of dengue haemorrhagic-fever outbreak risk by using entropy and artificial neural network. In *Communications and Information Technologies, 2008. ISCIT 2008. International Symposium on* (pp. 210-214). IEEE.
- [47] Rani, M. U., Devi, M. K., Mamatha, D. M., Seshadri, R., & Avulapati, Y. K. (2010). Clinical data warehouse on insect vector diseases to human of Andhra Pradesh. *Int. J. Comput. Sci. Inf. Secur*, 8(5), 240-244.
- [48] Ranjit, S., & Kissoon, N. (2011). Dengue hemorrhagic fever and shock syndromes. *Pediatric Critical Care Medicine*, 12(1), 90-100.
- [49] Rao, V. S. H., & Kumar, M. N. (2012). A new intelligence-based approach for computer-aided diagnosis of dengue

- fever. *IEEE transactions on information technology in biomedicine*, 16(1), 112-118.
- [50] Rey, F. A. (2003). Dengue virus envelope glycoprotein structure: new insight into its interactions during viral entry. *Proceedings of the National Academy of Sciences*, 100(12), 6899-6901.
- [51] Rigau-Pérez, J. G., Clark, G. G., Gubler, D. J., Reiter, P., Sanders, E. J., & Vorndam, A. V. (1998). Dengue and dengue haemorrhagic fever. *The Lancet*, 352(9132), 971-977.
- [52] Rissino, S., & Lambert-Torres, G. (2009). Rough set theory—fundamental concepts, principals, data extraction, and applications. *Data mining and knowledge discovery in real life applications*, 438.
- [53] Sam, S. S., Omar, S. F. S., Teoh, B. T., Abd-Jamil, J., & AbuBakar, S. (2013). Review of dengue hemorrhagic fever fatal cases seen among adults: a retrospective study. *PLoS Negl Trop Dis*, 7(5), e2194.
- [54] Schwartz, E., Weld, L. H., Wilder-Smith, A., von Sonnenburg, F., Keystone, J. S., Kain, K. C., ... & GeoSentinel Surveillance Network. (2008). Seasonality, annual trends, and characteristics of dengue among ill returned travelers, 1997–2006. *Emerg Infect Dis*, 14(7), 1081-8.
- [55] Shakil, K. A., Anis, S., & Alam, M. (2015). Dengue disease prediction using weka data mining tool. *arXiv preprint arXiv:1502.05167*.
- [56] Sylla, M., Bosio, C., Urdaneta-Marquez, L., Ndiaye, M., & Black IV, W. C. (2009). Gene flow, subspecies composition, and dengue virus-2 susceptibility among *Aedes aegypti* collections in Senegal. *PLoS Negl Trop Dis*, 3(4), e408.
- [57] Tang, E. K., Suganthan, P. N., & Yao, X. (2006). Gene selection algorithms for microarray data based on least squares support vector machine. *BMC bioinformatics*, 7(1), 1.
- [58] Tanner, L., Schreiber, M., Low, J. G., Ong, A., Tolfvenstam, T., Lai, Y. L., ... & Simmons, C. P. (2008). Decision tree algorithms predict the diagnosis and outcome of dengue fever in the early phase of illness. *PLoS Negl Trop Dis*, 2(3), e196.
- [59] Taraphdar, D., Sarkar, A., Mukhopadhyay, B. B., & Chatterjee, S. (2012). A comparative study of clinical features between monotypic and dual infection cases with Chikungunya virus and dengue virus in West Bengal, India. *The American journal of tropical medicine and hygiene*, 86(4), 720-723.
- [60] Thomas, L., Verlaeten, O., Cabié, A., Kaidomar, S., Moravie, V., Martial, J., ... & Césaire, R. (2008). Influence of the dengue serotype, previous dengue infection, and plasma viral load on clinical presentation and outcome during a dengue-2 and dengue-4 co-epidemic. *The American journal of tropical medicine and hygiene*, 78(6), 990-998.
- [61] Villegas-Rosales, P. M., Méndez-Tenorio, A., Ortega-Soto, E., & Barrón, B. L. (2012). Bioinformatics prediction of siRNAs as potential antiviral agents against dengue viruses. *Bioinformatics*, 8(11), 519-522.
- [62] Vontas, J., Kioulos, E., Pavlidi, N., Morou, E., Della Torre, A., & Ranson, H. (2012). Insecticide resistance in the major dengue vectors *Aedes albopictus* and *Aedes aegypti*. *Pesticide Biochemistry and Physiology*, 104(2), 126-131.
- [63] Wen, J. S., Jiang, L. F., Zhou, J. M., Yan, H. J., & Fang, D. Y. (2008). Computational prediction and identification of dengue virus-specific CD4+ T-cell epitopes. *Virus research*, 132(1), 42-48.
- [64] Wichmann, O., Hongsiriwon, S., Bowonwatanuwong, C., Chotivanich, K., Sukthana, Y., & Pukrittayakamee, S. (2004). Risk factors and clinical features associated with severe dengue infection in adults and children during the 2001 epidemic in Chonburi, Thailand. *Tropical Medicine & International Health*, 9(9), 1022-1029.
- [65] Wongkoon, S., Jaroensutasinee, M., & Jaroensutasinee, K. (2012). Development of temporal modeling for prediction of dengue infection in Northeastern Thailand. *Asian Pacific Journal of Tropical Medicine*, 5(3), 249-252.
- [66] Wongkoon, S., Jaroensutasinee, M., & Jaroensutasinee, K. (2013). Distribution, seasonal variation & dengue transmission prediction in Sisaket, Thailand. *The Indian journal of medical research*, 138(3), 347.
- [67] World Health Organization. (1986). *Dengue haemorrhagic fever: diagnosis, treatment and control*. World Health Organization.
- [68] Wu, Y., Lee, G., Fu, X., & Hung, T. (2008). Detect climatic factors contributing to dengue outbreak based on wavelet, support vector machines and genetic algorithm.
- [69] Wu, Y., Lee, G., Fu, X., Soh, H., & Hung, T. (2009). Mining weather information in dengue outbreak: predicting future cases based on wavelet, SVM and GA. In *Advances in Electrical Engineering and Computational Science* (pp. 483-494). Springer Netherlands.
- [70] Yusof, Y., & Mustafa, Z. (2011). Dengue outbreak prediction: A least squares support vector machines approach. *International Journal of Computer Theory and Engineering*, 3(4), 489.