



## Prediction of Dengue Fever Using Intelligent Classifier

M. Ganthimathi<sup>1</sup>, Dr. M. Thangamani<sup>2</sup>, C. Mallika<sup>3</sup>, V. Prasanna Balaji<sup>4</sup>

<sup>1</sup> Associate Professor / CSE, Muthayammal Engineering College, Namakkal, India, [gandhimanimay02@gmail.com](mailto:gandhimanimay02@gmail.com)

<sup>2</sup> Assistant Professor, Kongu Engineering College, Perundurai, India, Email: [manithangamani2@gmail.com](mailto:manithangamani2@gmail.com)

<sup>3</sup> Associate Professor/ MCA, E.G.S. Pillay Engineering College, Nagapattinam, India, [cmallikachinna@gmail.com](mailto:cmallikachinna@gmail.com)

<sup>4</sup> Student, Kongu Engineering College, Perundurai, India, [prasannabalaji014@gmail.com](mailto:prasannabalaji014@gmail.com)

### ABSTRACT

Dengue disease risk to be the leading cause of death worldwide. Early diagnosis of dengue continues to be a concern for public health in countries with a high incidence of this disease. Dengue is one of the most widespread water-borne and mosquito-borne diseases known today. Due to this Dengue disease Hepatitis creates viral infection initiation of the dengue fever. Viral infection is related to the HIV disease. However government taken many precaution to control this disease. Dengue prediction models are very important at rainfall season as the dengue cases becoming a major health issue in tropical and subtropical countries. Data Mining is a big and motivating research area in medical & healthcare department. Our system creates awareness about dengue fever and suggested treatment. Initiation of epidemic will cause dengue fever and this project aid to build a dengue incidence prediction model to avoid epidemic in real time. Experiment is conducted by using benchmark data. To process the data we are using is Data Mining techniques and Machine learning tool. It is helpful to find Profitable and successful systematic method in wellbeing data. The algorithms are used are Support Vector Machine (SVM), k-Nearest Neighbor (KNN). The dataset is downloaded from Data. World website, the algorithms are compared and SVM gives the accuracy of 100% and KNN gives the accuracy of 96%.

**Key words:** Classification, Data Mining, Dengue fever, K-Nearest Neighbor and Support Vector Machine.

### 1. INTRODUCTION

Dengue fever is the highly development virus which related issues as reported in World Health Organization (WHO). Dengue is a mosquito-borne viral infection. Dengue is found in tropical and sub-tropical climates worldwide, mostly in

urban and semi-urban areas. DHF starts abruptly with continuous high fever and headache. There are respiratory and intestinal symptoms with rashes/ red spot, pain behind eyes, low heart rate, fatigue, Nausea or vomiting, muscles and joint pain. The Dengue fever is the most common disease among a prognosis in 2019 with mainly cases from Australia and Asian Countries. The global incidence of dengue has grown dramatically in recent decades. About half of the world's population is now at risk. There are an estimated 390 million dengue virus infections per year, new cases and 120 million people deaths in 2018. Normal fever continuous on 2 to 5 days but Dengue fever attack in body 2 to 7 days incessant same symptoms. Severe Dengue is a leading cause of serious illness and death among children and adults in some Asian and Latin American countries. It demonstrated by the striking statistical numbers published every year requires management by medical professionals in hospitals. They indicate that the 5-year survival rate for patients with Dengue fever can be improved from an average of 20% up to 50% if the disease is diagnosed and treated at its early stage. Medical data as an essential part of medical diagnosis and treatment were concentrating on these data; the prognosis for dengue is usually good. These data include prosperity of unseen information that exploited by physicians in making reasoned decisions about a patient. However, extracting this relevant hidden information is a critical first step to their use. This reason motivates to use data mining techniques capabilities for efficient knowledge extraction & find hidden fever parts. The aim of this work is to create the system that will be able to identify the fever as Dengue or No Dengue at early stage of the diagnosis by applying some data technique as shown in method (techniques) proposed.

Mining Dengue data involves many processes. Medical Data Mining is a promising area of computational intelligence applied to an automatically analyze patients records aiming at the discovery of new knowledge useful for medical decision-making. Induced knowledge is anticipated

not only to increase accurate diagnosis and successful disease treatment, but also to enhance safety by reducing errors. The methods in this paper classify the digital Dengue disease in two categories: normal and abnormal. The normal ones are characterizing a No Dengue. The abnormal ones include. Type of Dengue fever; we will use a common classification method namely Support Vector Machine (SVM) and k Nearest Neighbor (KNN).

## 2. LITERATURE REVIEW

Iqra Jahangir *et. al* [1] planed a dengue prediction methodology by using the data mining technique. The Data mining technique is used Association Rule Mining (ARM) is used in this paper. The data was collected which includes patients having dengue or not. Weka tool is used for solving data mining problems in bioinformatics research. The results show 75% accuracy of this rule. The hybrid Classification model [2] is the combination of Naïve Bayes and decision tree. The Phases of Research Methodology are Pre-Processing; Apply Multiple Classifiers, Voting Method. In applying multiple classifiers, they have used Support Vector Machine (SVM), Naïve Bayes and Decision Tree. Hybrid classifier is combination of Naive Bayes and decision tree. In Voting method, the comparing the prediction of a algorithms.

The Dengue fever prediction [3] is Performed and Compared for the highest accuracy for the dataset. The techniques were used in this paper are: Naïve Bayes technique, REP tree, Random Tree, J48 and SMO. Naïve Bayes is highest accuracy among these techniques. Kashish Ara Shakil *et.al* [4] adopted the Prediction of dengue three techniques. The first technique uses explorer interface and depends on algorithms like Naïve Bayes, SMO, J48, REP Tree and RANDOM Tree. The second technique uses Experimenter interface for running algorithms Naïve Bayes, SMO, J48, REP Tree and RANDOM Tree. The third technique uses Knowledge Flow. We classified the accuracy of different algorithms Naïve Bayes, SMO, J48, REP Tree and random Tree on different dataset and compared the results for best result. Nave Bayes has 99% and J48 has 99.70 are the best performance classifier algorithm.

The Dengue predictions processed [5] in three ways are: Pre-processing the dataset, Forward selection (FS) and backward selection (BS). Applying classification techniques: Decision tree and Support Vector Machine (SVM).It shows the accuracy of classification techniques. Weka tool is used to classifying the dataset. By comparing the Decision tree and SVM algorithm results. The SVM has better accuracy. The Manivannan *et. al* [6] investigated dengue dataset in four stages namely pre-processing, attribute selection, K-Medoid Clustering and predicting the dengue fever. They have been

using R-tool. The performances of the algorithm many vary according to the method. K-Medoid clustering is increasing the efficiency of the output. The dataset consist of Fever, Bleeding, Metallic taste, Fatigue [7].

It performs various classification techniques that are REP tree, J48, SMO, ZeroR and Random tree. SMO and J48 is a highest accuracy of 84% and 76% then comparing to other algorithms. Kashish Ara Shakil *et. al* [8] used Weka tool for classification. They have taken 108 instances, in that they separated dataset like 99 rows and 18 attributes. The accuracy for Navies Bayes is 98.9% and J48 Tree is 98%. The attributes like Fever, vomiting, headache, Muscle pain and Joined pain. They are using different classifiers like Naive Bayes, J48 classifier and clustering algorithm like K-Medioids, Dbscan, k-mean accurately predicting dengue disease [9]. The author has the accuracy of 92% for Naïve Bayes and J48. The attribute are temperature, spotting, bleeding and they are using Naive Bayes and Orange tool for predicting [10]. The author has the accuracy of 77.3% for Naïve Bayes.

Naiyar Iqbal *et al* [11] have been using simple classifiers like Decision tree and Naïve Bayes and for complex algorithms they have been used SVM, Neural Networks and Ensemble classifiers. Next [12] have been using Decision tree, Naive Bayes, and Neural Network. Their Dataset consist of 1800 records. The accuracy for Decision tree is 96%, Naïve Bayes is 94% and for Neural Network is 92% . Sundari and Krishnamoorthy [13] have been using all the technique of all the paper and given the journal's list like k-means, SVM, Decision tree, etc. Their accuracy of SVM is 87% and Decision tree is 95%. Thangamani and Prasanna [14] , used Weka tool in data analytics using wine datasets . Chitrakumar, Thangamani and Premalatha [15] used Weka tool for exploring data mining application.

The attributes [16] are joint pain, muscle pain, rash, mild bleeding and applied various clustering techniques. Authors have suggested for Dengue fever. . Jorge D. Mello-Román,[17] proposed ANN and SVM, as prediction tools. Their accuracy of ANN is 96% and for SVM is 90%. The techniques like WEKA tool and their algorithms are Naive Bayes, J48 and REP tree used [16]. The accuracy for Naïve Bayes is 63.3%, J48 accuracy is 76.6% and for REP tree is 76.6%.

## 3. PROPOSED RULE

R is a language and environment for statistical computing and graphics. R provides a wide variety of statistical and graphical techniques, and is highly extensible. Classical statistical tests, time-series analysis, classification, clustering, and others. R is easily extensible through functions and

extensions, and the R community is noted for its active contributions in terms of packages.

The technique, which is used for the prediction and classification, is as follows,

1. *Support Vector Machine:*

In machine learning, support-vector machines (SVMs, also support-vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. We are collecting the bench mark dataset of dengue. We divide the dataset into two parts as training and testing data then we used the confusion matrix then have acquired the 100% accuracy.

2. *K-Nearest Neighbors:*

In pattern recognition, the value of data in k-nearest neighbors algorithm (k-NN) is a non-parametric method used for classification and regression. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression. In k-NN classification, the output is a class data of test. we used the confusion matrix then have acquired the k-value has 5 its shows 96% accuracy and when k-value as 21 its shows 90% accuracy.

4. RESULT AND ANALYSIS

The sample of 100 dengue dataset was collected for this research Shown in Table: 1. The collected data was applied in the R-Studio. Using SVM and KNN algorithm, we found an accuracy of 100% for SVM and 96% for KNN (Shown in Figure: 1)

Table 1: Attributes of Dengue dataset

	A	B	C	D	E	F	G	H	I	J	K	L
1	age	gender	marriage	period_of_rashes		pain_behi	muscles_low	heartnausea_ar	fatigue		dengue_status	
2	YA	MALE	UNMARRI	FAF	YES	YES	YES	YES	YES	YES	YES	
3	YA	FEMALE	UNMARRI	FAF	YES	YES	YES	YES	YES	YES	YES	
4	MA	MALE	UNMARRI	BF	NO	NO	YES	YES	NO	NO	NO	
5	OA	MALE	MARRIED	BF	NO	YES	YES	NO	NO	YES	NO	
6	MA	FEMALE	MARRIED	FAF	YES	YES	YES	YES	YES	YES	YES	
7	MA	FEMALE	UNMARRI	BF	YES	YES	YES	YES	YES	NO	NO	
8	OA	MALE	MARRIED	BF	YES	YES	YES	NO	NO	NO	NO	
9	OA	FEMALE	MARRIED	FAF	YES	YES	YES	YES	YES	YES	YES	
10	MA	MALE	MARRIED	BF	YES	YES	YES	NO	NO	YES	NO	
11	OA	FEMALE	MARRIED	BF	NO	YES	YES	NO	NO	NO	NO	
12	MA	FEMALE	MARRIED	FAF	YES	YES	YES	YES	YES	YES	YES	
13	OA	FEMALE	MARRIED	BF	NO	YES	NO	YES	YES	YES	NO	
14	OA	FEMALE	MARRIED	BF	NO	YES	YES	YES	NO	NO	NO	
15	OA	MALE	MARRIED	FAF	YES	YES	YES	YES	YES	YES	YES	
16	OA	MALE	MARRIED	FAF	YES	YES	YES	YES	YES	YES	YES	
17	OA	MALE	MARRIED	FAF	YES	YES	YES	YES	YES	YES	YES	
18	MA	MALE	UNMARRI	BF	YES	NO	YES	NO	YES	YES	NO	
19	MA	FEMALE	UNMARRI	BF	NO	YES	YES	NO	NO	NO	NO	
20	OA	FEMALE	MARRIED	BF	NO	YES	NO	NO	NO	NO	YES	
21	MA	MALE	UNMARRI	BF	NO	NO	NO	NO	YES	YES	NO	
22	OA	FEMALE	MARRIED	FAF	YES	YES	YES	YES	YES	YES	YES	
23	OA	MALE	MARRIED	FAF	YES	YES	YES	YES	YES	YES	YES	
24	MA	FEMALE	UNMARRI	BF	NO	NO	NO	YES	YES	YES	NO	
25	OA	FEMALE	MARRIED	BF	NO	YES	YES	YES	NO	NO	NO	

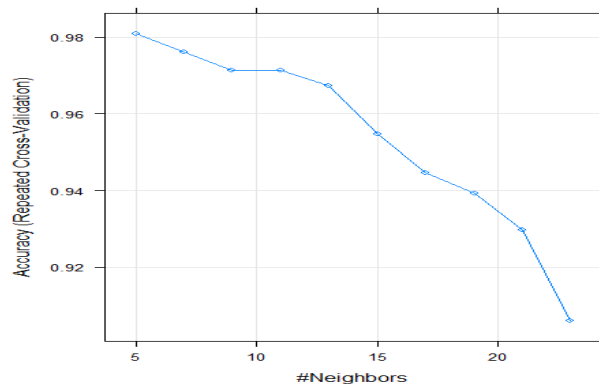
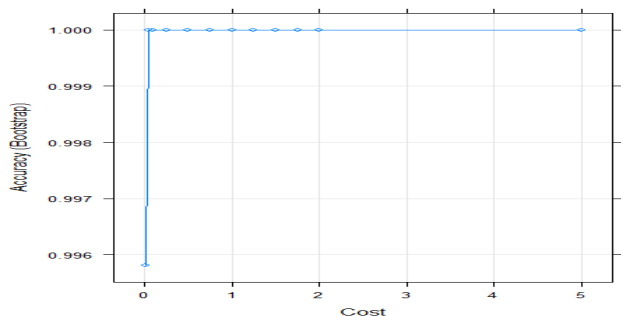


Figure 1: Above graph shows the results of SVM and K-NN classification.

5. ACKNOWLEDGEMENT

It is pleasure to acknowledge the contribution made by Kongu Engineering College Students P. Dhamotharan and K. Oviya Kumar who helped us a lot in gathering different information, collecting data time to time in making this project success.

REFERENCES

1. Iqra Jahangir and Abdul Basit. **Prediction of Dengue Disease Through Data Mining By Using Modified Apriori Algorithm**, Research Gate on Abasyn University, 2018. <https://doi.org/10.1145/3213187.3287612>
2. Prashansa Taneja and Nisha Gautam. **Hybrid Classification Method for Dengue Prediction**, International Journal of Engineering and Advanced Technology (IJEAT), 2019.
3. Kamran Shaukat, Nayyer Masood, Sundas Mehreen and Ulya Azmeen. **Dengue Fever Prediction: A Data Mining Problem**, Journal of Data Mining in Genomics & Proteomics, 2015. <https://doi.org/10.4172/2153-0602.1000181>
4. Kashish Ara Shakil, Samiya Khan, ShadmaAnis, and Mansaf Alam. **Dengue Disease Prediction Using Weka Data Mining Tool**, Proceedings of IIRAJ International Conference, 2017.
5. R. Sanjudevi, D. Savitha. **Dengue Fever Prediction Using Classification Techniques**, International Research Journal of Engineering and Technology (IRJET), 2019.
6. P. Manivannan and Dr. P. Isakki. **Dengue Fever**, Journal of Innovative Research in Computer and Communication Engineering, 2017.
7. M. Bhavani and S. Vinod Kumar. **A Data Mining Approach For Precise Diagnosis of Dengue Fever**, International Journal of Trends in Engineering and Technology, 2017.

8. Kashish Ara Shakil, Samiya Khan, ShadmaAnis and Mansaf Alam. **Dengue disease prediction using Weka data mining tool**, IIRAJ International Conference, 2017.
9. Ms. S. Freeda Jebamalar and Dr. A. Anitha. **A Survey on Prediction of Dengue Fever Using Data Mining Techniques**, International Journal of Science, Engineering and Management (IJSEM), 2017.
10. Ria Arafiyah and Fariani Hermin. **Data mining for dengue hemorrhagic fever (DHF) prediction with naive Bayes method**, 2018.  
<https://doi.org/10.1088/1742-6596/948/1/012077>
11. Naiyar Iqbal and Mohammad Islam. **Machine learning for dengue outbreak prediction**, International Journal of Advanced Research in Computer Science, 2017.
12. Dave Kaveri Atulbhai and Shilpa Serasiya. **A Survey Prediction & Detection of Dengue – Mining Methods & Techniques**, IJARIE-ISSN International Conference, 2017.
13. B. Sundari, M. Krishnamoorthy. **Factors to Predict Dengue Fever using Data Mining Techniques**, International Journal of Scientific Research and Engineering Development, 2019.
14. P. Sathya and A. Sumathi. **Predicting Dengue Fever Using Data Mining Techniques**, International Journal of Computer Science Trends and Technology (IJCST), 2018.
15. V. Prasanna and Dr. M. Thangamani. **DataAnalytics in Wine Datasets Using Weka Clustering**, International Scientific Global Journal in Engineering, Science and Applied Research (ISGJESAR), 2016.
16. T. Chithrakumar, Dr. M. Thangamani and C. Premalatha. **Exploring data Mining Classification Approach in Weka Open Source**, International Research Journal in Global Engineering and Sciences.(IRJGES), 2016.
17. D. Jorge, Mello-Román, C. Julio, Mello-Román, Santiago Gómez-Guerrero and Miguel García-Torres . **Predictive Models for the Medical Diagnosis of Dengue: A Case Study in Paraguay**, Computational and Mathematical Methods in Medicine. 2019.  
<https://doi.org/10.1155/2019/7307803>
18. N. Rajathi, S. Kanagaraj, R. Brahmanambika and K. Manjubarkavi. **Early Detection of Dengue Using Machine Learning Algorithms**, International Journal of Pure and Applied Mathematics, 2018.