

Finding the Best Feature Selection Method for Dengue Diagnosis Predictions

R Lathesparan^{1#}, RMKT Rathnayaka² and WU Wickramaarachchi¹

¹Department of Computing, Rajarata University of Sri Lanka, Mihintale, Sri Lanka

²Department of Physical Sciences and Technologies, Sabaragamuwa University of Sri Lanka Belihuloya, Sri Lanka

#lathesparan@gmail.com

Abstract— Dengue is a mosquito-borne viral disease that has dramatically increased around the world in recent years. The spread of Dengue depends on the tropics, rainfall, temperature, relative humidity and unplanned urbanization. Severe Dengue can lead to circulatory system failure, shock and even death. The development of an effective Dengue fever prediction model is therefore essential for better Dengue case management. Feature selection is the predominant phase in developing the Dengue diagnosis prediction model. It is required to identify the most crucial attributes, as not all attributes have notable effects on the results. Therefore, this study focuses on the feature selection methods such as Principal Component Analysis (PCA) and Wrapper feature selections method with Naïve Bayes, K-Nearest Neighbor (KNN), and J48 algorithms. Simple Artificial Neural Networks (ANN) were developed to validate the performance based on the accuracy of each feature selection method, since it can work well with the partial dataset. Myalgia and Retro-Ocular Pain are the most expressive features chosen by all wrapper feature selection methods. In addition, with PCA, the initial 22-dimensional system was reduced to an 8-dimensional system with a cumulative variance of 59%. ANN with PCA resulted in the higher accuracy of 72.47% and ANN with Wrapper feature selection (KNN) showed the lowest accuracy of 54.47%. In conclusion, PCA is identified as the best feature selection method for the given dataset in this study based on the accuracy of ANN. In future, multiple Dengue diagnosis prediction models can be developed with higher accuracy and efficiency using the most vital attributes.

Keywords: *feature selection, artificial neural networks, dengue diagnosis prediction models*

I. INTRODUCTION

Dengue is a mosquito-borne viral disease that has dramatically increased around the world in recent years (World Health Organization, 2016). According to the World Health Organization (WHO), the spread of Dengue is influenced by the tropics, rainfall, temperature, relative humidity, and unplanned rapid urbanization (World Health Organization, 2016). In part Dengue spread fast in urban areas when the mosquito population increase during the rainy season (Cetiner, Sari and Aburas, 2009). Dengue viruses are infected on humans through the bites of an infective mosquito, which acquires the virus while feeding on the blood of an infected person. The person infected with the Dengue virus serves as a source of the virus for uninfected mosquitoes and becomes the main carrier and multiplier of the virus, and transmits it via Aedes mosquitoes (World Health Organization, 2016). The basic symptoms of Dengue are sudden high fever, pain in the eyes, muscle & joint pain, bone pain, severe headache, and a skin rash with red spots. Severe Dengue may cause complications related to internal bleeding, organ impairment, and plasma leakage (World Health Organization, 2016). Severe Dengue situations could show abdominal pain, vomiting, diarrhea, convulsions, bruising, and uncontrolled bleeding. Dengue fever can last up to 7 days, however, this can lead to complications like circulatory system failure and shock. If the severe Dengue has not been treated appropriately, it may lead to death (Pham *et al.*, 2018).

II. LITERATURE REVIEW

Early diagnosing of Dengue is a cyclic need in the health system. It helps in improving the diagnostic time, cost, and reduces the pain (Mello-Román *et al.*, 2019). Dengue is genediagnosed checking blood pressure, heart rate, and body temperature. Comprehensive laboratory analysis must be done on the full blood count, which includes haematocrit, platelet, and white blood cell counts (Dasgupta *et al.*, 2019). Laboratory analysis depends on human work. Human involvement, however, contributes to the misidentification of Dengue and non-Dengue patients, which leads to negative medical results. Therefore, early diagnosis of an affected person without human errors is critical (World Health Organization, 2016; Mello-Román *et al.*, 2019). The development of an effective Dengue fever prediction model is therefore essential for better Dengue diagnosis (Binti Mohd Zainee and Chellappan, 2016). Feature selection is an important phase in developing the Dengue diagnosis prediction model using classification. It is a process; which selects the attributes most relevant for the prediction problem in the dataset. Clinical features of Dengue fever vary according to the age of the patient and not all attributes have significant effects on the result (Davi *et al.*, 2019). Therefore, it is required to identify the most significant attributes and avoid the least significant attributes when developing a Dengue diagnosis prediction model. The use of feature selection models has an enormous impact on developing prediction models, and it will improve the accuracy and efficiency (Dasgupta *et al.*, 2019). The purpose of this study is therefore to evaluate the performance of feature selection methods that can be used to develop a Dengue diagnosis prediction model. This study focuses on the Principal Component Analysis (PCA) and Wrapper feature selection methods with Naïve Bayes, K-Nearest Neighbour (KNN), and J48.

A. Attributes used in Dengue diagnosis prediction models

In general, Dengue-related data can be collected and categorized under four areas: environmental, epidemiological, personal information, and socio-economic data (Dharmawardana *et al.*, 2018; Pham *et al.*, 2018). Environmental features include factors such as rainfall, humidity,

temperature, and wind speed (Cetiner, Sari and Aburas, 2009; Dharmawardana *et al.*, 2018; Chovatiya *et al.*, 2019). Socio-economic features consist of aspects such as population, year, poverty, and administrative boundaries. Personal features, such as age, sex, travel, and camping near to water source also contributed to Dengue prediction. Importantly, the epidemiological features play a crucial role in predicting Dengue. Although all four categories can be used for the Dengue diagnosis prediction models, epidemiological factors are important attributes in prediction models. Epidemiological factors are the attributes that were unused-time medical situations, and it directly affects the medical diagnosis of Dengue such as Body Temperature, Retro-ocular pain, Platelet count, White blood cell count, and Hematocrit. Environmental, socio-economic, and personal information factors play a significant role in Dengue diagnosis; however, they do not directly affect it.

B. Feature selection methods used in Dengue diagnosis prediction models

Data pre-processing is a crucial step in developing prediction models as it directly affects the quality of the model and the ability of the learning process of the prediction model (Dhairya Kumar, 2018). The raw data might consist of many variations in the values of each feature, which might lead to incorrect results (Manivannan and Devi, 2018). Therefore, it is crucial to pre-process the data before it can be used (Dhairya Kumar, 2018). Researchers have been using various pre-processing steps to handle missing values, normalize the values, feature selection, and splitting data into the training set and validation set. Out of them, the feature selection is crucial to identify features that have no significant effects on the results and to remove them (Muhilthini *et al.*, 2018). Mello Roman, and others have selected the features which were only used to confirm laboratory diagnosis based on their knowledge of the domain to develop the Dengue diagnosis model (Mello-Román *et al.*, 2019). Adbiel has used the neural pathway strength feature selection method (NPSFS) to identify the most relevant inputs, by creating an ensemble of ANN's and comparing the similarities of the weighted results of the pathway. Attributes with the most similarity of pathway strengths across the whole ensemble of ANNs were selected as the most

relevant (Laureano-Rosario *et al.*, 2018). Chovatiya, has used Pearson's Correlation formula to calculate the correlation coefficient between the total numbers of cases with all the other attributes. Pearson's Correlation formula is used to calculate the correlation coefficient between attributes and target along with the ranker search method. The attributes with a maximum correlation coefficient have been selected for further processing in their study (Chovatiya *et al.*, 2019). Similarly, Nishanthini and others used a correlation analysis between each weather variable with the Dengue cases reported was carried out using the weather variables with time lags. The attributes holding the highest correlation to Dengue cases were chosen as input parameters (H.M.NishanthiHerath, Perera and Wijekoon, 2014). However, the statistical feature selection is mostly applied when the number of features is higher, and it was intended to reduce the number of attributes otherwise the domain knowledge is wildly used in feature selection. An appropriate feature selection method should be applied based on the situation uniquely for each dataset.

C. Dengue diagnosis prediction models developed
Multiple Dengue diagnosis prediction models have been developed based on selected features. Jorde and others have used ANN and Support Vector Machine (SVM) to develop a Dengue diagnosis prediction model based on epidemiological data of 4332 registered Dengue cases from the public health system of Paraguay. The missing values of data were excluded, and principal factors were replaced with the mean of adjacent data. ANN with multi-layer perceptron resulted in 96% accuracy and ANN with radial basis function resulted in 55% accuracy. SVM with

linear function given the accuracy of 64%, SVM with Gaussian function given the accuracy of 86%, and SVM with polynomial function given the accuracy of 92%. ANN with multi-layer perceptron and SVM with the polynomial function showed the highest performance in their kinds. (Mello-Román *et al.*, 2019). In another study, Norhayati and Kalaivani utilized epidemiological features collected from 60 patients and used algorithms such as Decision tree, Discriminant analysis, SVM, KNN, and Ensemble Classifiers. Out of all these methods, the Discriminant Analysis has the highest accuracy (83.3%) and SVM has the lowest accuracy (63.3%) (Binti Mohd Zainee and Chellappan, 2016). Furthermore, Anusha used 60 instances of epidemiological data that consisted of 11 attributes with algorithms, such as Rule-based classifier, KNN, ID3, Naïve Bayes, and Decision tree. Out of these algorithms, Naïve Bayes has given the highest accuracy of 72.2% and KNN showed the lowest accuracy of 50% (Anusha, 2019). In another study, conducted by Harshada and Pramod have used real-time data of Heart Rate, Body Temperature, Blood Pressure, and other epidemiological factors to diagnose Dengue. They specifically used Naïve Bayes algorithms for the probability calculations (Somwanshi and Harshada, 2018). Similarly, Manivannan and Isakki used 1910 instances that consisted of 171 attributes with K means clustering to identify the serotypes of Dengue (Manivannan and Devi, 2018). Furthermore, Sanjudevi and Savitha have used epidemiological data with 108 instances that consisted of 18 attributes to develop a Decision tree and SVM with Sequential Minimal Optimization. Here, the accuracy of the Decision tree was 87.5% and SVM with Sequential Minimal Optimization was 99% (Sanjudevi and Savitha, 2019).

III. METHODOLOGY AND EXPERIMENTAL

DESIGN

The dataset used in this study consist of epidemiological features of previous Dengue patients. Feature selection methods are applied to the given data set to determine the most relevant attributes and evaluate the impact of the feature selection process on the Dengue diagnosis prediction model. Figure 1 shows the proposed methodology of this study. First, the

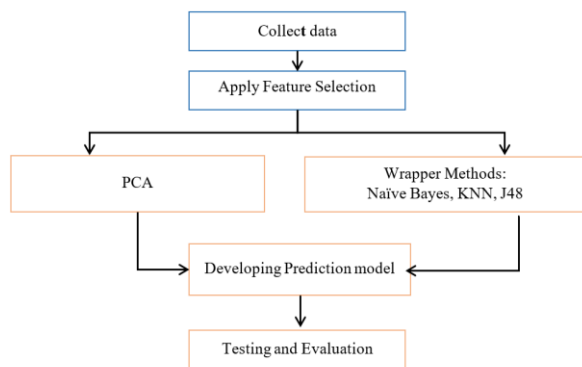


Figure 14. Proposed Methodology

dataset is collected from the scholars of a study conducted at Paraguay (Mello-Román *et al.*, 2019). Then the dataset is used for feature selection and train prediction models. PCA, Wrapper feature selection with Naïve Bayes, KNN, and J48 were applied to select the most sign noteworthy features on selected features ANN models were built for the task of Dengue diagnosis prediction problem and its accuracy was used to evaluate the performance of the feature selection method.

PCA is a statistical and unsupervised dimensionality reduction technique, which has been widely used in data mining problems. It projects the high dimensional space to a low dimensional space while improving the efficiency and accuracy of the data mining task. It is based on the eigenvectors analysis to identify critical original features for the principal component. PCA determining key variables in a high-dimensional data set and simplifies the analysis and visualization of high-dimensional data set, with less loss of information. In simple terms, PCA generates a new set of variables, called principal components. Each principal component linearly combines the original variables and works based on eigenvectors analysis to identify critical original features for the principal component. Secondly, Wrapper feature selection methods were used. As classifiers to the Wrapper feature selection method, Naïve Bayes, KNN and J48 were used along with the best-first search in the forward direction. Naïve Bayes is a probabilistic machine learning algorithm based on the Bayes theorem. It works as the probabilistic classifier of all attributes contained in the data sample individually and then classifies the data problem, KNN is an algorithm that stores all available training data and to predict, the algorithm chooses the closest data points in the training dataset, and J48 is an algorithm to generate a decision tree.

ANNs were developed to validate the performance of the feature selection methods. it can easily expand and can work with large numbers of features and datasets. More importantly, it can work well with partial data due to its structure. Therefore, in this study, ANN was chosen to evaluate the performance of feature selection algorithms. The selected features from each feature selection method were used to

develop the ANN, which evaluated the selected features based on their accuracy. ANN had input neurons and output neurons, and connections between neurons. The connections between neurons are defined by a weight associated with each connection (Khemani, 2013). This model has three layers. The first layer is the input layer and it uses the 'relu' activation function. The second layer is the dropout layer, which is used to randomly remove the interconnection between nodes with a frequency of 0.5. It helps to prevent the overfitting of the model, by not letting the model memorize the pattern, each time the modal has mustrn new patterns. The final layer is the output layer, it is used to produce the output of the model. It uses 'sigmoid' as the activation function. For every input, it keeps the output between 0 - 1. Apart from these, Adam Optimizer was used, which is computationally efficient and works well with a noisy and large amount of data. Along with that, batch sizes of 32, learning rates of 0.1, and 50 epochs were used.

D. Dataset

The medical dataset of previous Dengue patients (2012 - 2016) was obtained from a study conducted in Paraguay, Mexico titled "Predictive Models for the Medical Diagnosis of Dengue: A Case Study in Paraguay" (Mello-Román *et al.*, 2019). This dataset consists of 668 instances with 23 Dengue fever vital sign attributes. Table 1 shows the attributes and their descriptions. Only cases confirmed or discarded with the laboratory criteria were included in the dataset (Mello-Román *et al.*, 2019). Further, this data set is already preprocessed; importantly, all the missing data were imputed. Therefore, the dataset is directly applied to feature selection methods.

IV. RESULT & DISCUSSION

PCA, where it generates a new set of variables, called principal components. Each principal component linearly combines the original variables and works based on eigenvectors analysis to identify critical original features for the principal component. Table 2 shows the eigenvalues of principal components and their variance. An eigenvalue is a number that shows the spread out of the data.

Table 1. Attributes and their descriptions (Mello-Román *et al.*, 2019)

I. Attribute	II. Description
Age	Age
Sex	Sex
Headache	Symptom of headache
Myalgia	Symptom of muscle pain
Arthralgia	Symptom of joint pain
Retro-Ocular Pain	Symptom of pain around the eye
Pruritus	Symptom of unpleasant sensation in the skin
Cough	Symptom of cough
Dyspnea	Symptom of breath shortness
Epistaxis	Symptom of nose bleeding
Gingivorrhagia	Symptom of bleeding from gums
Melena	Symptom of black stool with bleeding
Black Vomiting	Symptom of black vomiting
Exantema	Symptom of rash or eruption on the skin
Conjunctive Injection	Symptom of enlargement in conjunctival vessels
Tachycardia	Symptom of high heart rate
Hepatomegaly	Symptom of enlarged liver
Splenomegaly	Symptom of enlargement of the spleen
Sensory Alteration	Symptom of changes in the sensory stimuli
Stiff Neck	Symptom of stiff neck
Petechia	Symptom of red dots in the skin
Jaundice	Symptom of Jaundice
Final Class	Laboratory confirmation of Dengue

An eigenvalue less than 1 means that the principal component explains less strength than a single original variable explained. Therefore, principal components which have an eigenvalue greater than or equal to one taken into consideration. As showed in Table 2, the first eight principal components can be taken into consideration since those principal components have eigenvalues greater than one. Thus, it can be concluded that the initial 22-dimensional system can be reduced to an 8-dimensional system with a cumulative variance of 59%. Figure 2 interprets the distribution of eigenvectors between the first and second principal components. The first principal component has a variance of 12%. Moreover, it has a large positive association with Hepatomegaly, Jaundice, Splenomegaly, and Sensory alteration. These can be categories into

severe Dengue symptoms. Large negative association with Myalgia, Arthralgia, Headache, and Retro ocular pain. Where it can be considered as the basic Dengue symptoms. However, the second principal component shows a significant difference. It has no positive associations with any features. It has a large negative association with Hepatomegaly, Splenomegaly, and Dyspnea. Similarly, Tachycardia, Dyspnea, and Cough are the large positive association of the third component, and Black vomiting, Melena, and Gingivorrhagia are the large negative association of the third component. The ANN Model developed using PCA gives an accuracy of 72.47%. Further, in the Wrapper feature selection, with the Naïve Bayes algorithm; Myalgia, Retro-Ocular Pain, Dyspnea, and Petechia were selected; ANN model developed with these features has an accuracy of 55.97%, similarly in the K nearest neighbour; Myalgia, Retro-Ocular Pain, and Exanthema were selected; ANN model developed with these features has an accuracy of 54.47%. In J48; Myalgia, Retro-Ocular Pain, Tachycardia, and Petechia were selected. ANN model developed with these features has an accuracy of 55.97%. These results can be greedy since the wrapper methods only aim to find the best possible combinations of features that result in the best performant model. Moreover, wrapper methods find the optimal feature combination by developing machine learning models. Consequently, it results in better predictive accuracy. Apart from that, Myalgia and Retro-Ocular Pain are the most significant features that have been chosen by all wrapper feature selection methods.

Table 3 shows the summary of the selected attributes and accuracy of the ANN model developed using that feature. ANN developed without feature selection has an accuracy of 54.51%. Feature selection methods other than the Wrapper feature selection with KNN have greater accuracy than the ANN without feature selection. KNN shows the lowest accuracy of 54.47% and PCA results in higher accuracy of 72.47%. Further, the Naïve Bayes and J48 have an equal accuracy of 55.97%. As the key finding of this study, for the dataset that used in this study, PCA is the most suitable feature selection method.

Table 2. Correlation of each attribute to Final Class

Component	Total	Variance	Cumulative Variance
1	2.71116	12%	12%
2	2.14864	10%	22%
3	1.76419	8%	30%
4	1.55196	7%	37%
5	1.31615	6%	43%
6	1.24483	6%	49%
7	1.07744	5%	54%
8	1.0569	5%	59%
9	0.97304	4%	63%
10	0.94682	4%	67%
11	0.87966	4%	71%
12	0.78049	4%	75%
13	0.73764	3%	78%
14	0.70948	3%	81%
15	0.66657	3%	84%
16	0.64344	3%	87%
17	0.56404	3%	90%
18	0.55597	3%	92%
19	0.50935	2%	95%
20	0.49397	2%	97%
21	0.37051	2%	99%

V. CONCLUSION

In this study two approaches were used for the feature selection, PCA and Wrapper feature selection methods with Naïve Bayes, KNN, and J48. Separate ANN models have been developed for each feature selection method. Out of the four feature selection methods PCA results in higher accuracy in developed ANN. In a conclusion, PCA performs better for the given dataset. Myalgia and Retro-Ocular Pain are the most expressive features that have been chosen by all wrapper feature selection methods. Further, with PCA the initial 22-dimensional system was reduced to the 8-dimensional system with a cumulative variance of 59%.

As the limitation of the study, the ANN model that has been developed here is not optimal, a basic model with the same hyperparameters is developed only to compare the results of each feature selection model; however, each of these ANN can be further optimized by fine-tuning the hyperparameters. Moreover, the result of feature selection methods only applicable for the dataset that is used in the study, for other datasets, these feature selection methods should be applied

again, which may output a different result than this. Furthermore, this study only used few feature selection methods, however, trying the other feature selection methods is highly welcomed and it may change the current result on the selected significant attributes that can be used to develop the Dengue diagnosis prediction model more efficiently. In the future, based on the result of the feature selection methods, Dengue diagnosis prediction models can be developed with higher accuracy and efficiency using the most significant attributes.

Table 3. Selected Features and Accuracy of ANN

Feature Selection Method	Selected features	Accuracy of ANN
No Feature Selection	All	54.51%
PCA	8 Principal Components	72.47%
Wrapper Feature Selection (Naïve Bayes)	Myalgia, Retro-Ocular Pain, Dyspnea, Petechia	55.97%
Wrapper Feature Selection (KNN)	Myalgia, Retro-Ocular Pain, Exantema	54.47%
Wrapper Feature Selection (J48)	Myalgia, Retro-Ocular Pain, Tachycardia, Petechia	55.97%

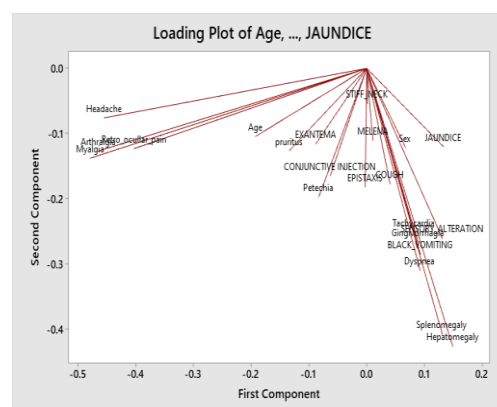


Figure 2. Loading plot of first and second principal component

REFERENCES

- Anusha, R. (2019) "Dengue Fever Prediction using Datamining Classification Technique," *International Journal of Recent Technology and Engineering*, 8(4), pp. 8685–8688. doi: 10.35940/ijrte.d8810.118419.
- Binti Mohd Zainee, N. and Chellappan, K. (2016) "A preliminary dengue fever prediction model based on vital signs and blood profile," *IECBES 2016 - IEEE-EMBS Conference on Biomedical Engineering and Sciences*, pp. 652–656. doi: 10.1109/IECBES.2016.7843530.
- Cetiner, B. G., Sari, M. and Aburas, H. M. (2009) "Recognition of Dengue Disease Patterns Using Artificial," *Neural Networks*, pp. 13–16.
- Chovatiya, M. *et al.* (2019) "Prediction of dengue using recurrent neural network," *Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019*. IEEE, 2019-April(Icoei), pp. 926–929. doi: 10.1109/icoei.2019.8862581.
- Dasgupta, S. *et al.* (2019) "Evaluating the performance of machine learning using feature selection methods on dengue dataset," *International Journal of Engineering and Advanced Technology*, 8(5), pp. 2679–2685.
- Davi, C. *et al.* (2019) "Severe Dengue Prognosis Using Human Genome Data and Machine Learning," *IEEE Transactions on Biomedical Engineering*, 66(10), pp. 2861–2868. doi: 10.1109/TBME.2019.2897285.
- Dhairya Kumar (2018) *Introduction to Data Preprocessing in Machine Learning | by Dhairya Kumar | Towards Data Science*. Available at: <https://towardsdatascience.com/introduction-to-data-preprocessing-in-machine-learning-a9fa83a5dc9d> (Accessed: October 27, 2020).
- Dharmawardana, K. G. S. *et al.* (2018) "Predictive model for the dengue incidences in Sri Lanka using mobile network big data," *2017 IEEE International Conference on Industrial and Information Systems, ICIIS 2017 - Proceedings*, 2018-Janua, pp. 1–6. doi: 10.1109/ICIINFS.2017.8300381.
- H.M.NishanthiHerath, P., Perera, A. A. I. and Wijekoon, H. P. (2014) "Prediction of Dengue Outbreaks in Sri Lanka using Artificial Neural Networks," *International Journal of Computer Applications*, 101(15), pp. 1–5. doi: 10.5120/17760-8862.
- Khemani, D. (2013) *A First Course in Artificial Intelligence*. 1st Editio. MC GRAW HILL INDIA.
- Laureano-Rosario, A. *et al.* (2018) "Application of Artificial Neural Networks for Dengue Fever Outbreak Predictions in the Northwest Coast of Yucatan, Mexico and San Juan, Puerto Rico," *Tropical Medicine and Infectious Disease*, 3(1), p. 5. doi: 10.3390/tropicalmed3010005.
- Manivannan, P. and Devi, P. I. (2018) "Dengue fever prediction using K-means clustering algorithm," *Proceedings of the 2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing, INCOS 2017*, 2018-Febru, pp. 1–5. doi: 10.1109/ITCOSP.2017.8303126.
- Mello-Román, J. D. *et al.* (2019) "Predictive Models for the Medical Diagnosis of Dengue: A Case Study in Paraguay," *Computational and Mathematical Methods in Medicine*, 2019. doi: 10.1155/2019/7307803.
- Muhilthini, P. *et al.* (2018) "Dengue Possibility Forecasting Model using Machine Learning Algorithms," *International Research Journal of Engineering and Technology*, pp. 1661–1665.
- Pham, D. N. *et al.* (2018) "How to Efficiently Predict Dengue Incidence in Kuala Lumpur," *Proceedings - 2018 4th International Conference on Advances in Computing, Communication and Automation, ICACCA 2018*. IEEE, pp. 1–6. doi: 10.1109/ICACCAF.2018.8776790.
- Sanjudevi, R. and Savitha, D. (2019) "Dengue Fever Prediction using Datamining Classification Technique," *International Journal of Recent Technology and Engineering*, 8(4), pp. 8685–8688. doi: 10.35940/ijrte.d8810.118419.
- Somwanshi, Ms Harshada, M. P. G. (2018) "Real-Time Dengue Prediction using Naive Bayes Predicator in the IoT," *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)*. IEEE, (Icirca), pp. 725–728.
- World Health Organization (2016) *Dengue and severe dengue*, World Health Organization. Available at: <http://www.who.int/mediacentre/factsheets/fs117/en/> (Accessed: October 26, 2020).

ACKNOWLEDGMENT

I wish to acknowledge the people, friends and family and others, who provided me with support in various ways throughout this study.

AUTHOR BIOGRAPHIES



R Lathesparan is an undergraduate of the Department of Computing, Rajarata University of Sri Lanka. Currently he is following an internship at

EchonLabs (Pvt) Ltd. His research interest are Artificial Intelligence and Machine Learning.



Prof. (Dr.) R.M.K.T. Rathnayake is a professor in Statistics at Sabaragamuwa university of Sri Lanka. His research interest are financial mathematics, Time Series Modelling and Data Mining and Machine Learning, Big Data Analytics, Business Models, Machine learning and Multi objective Combinatorial optimization.



Wiraj Udara Wickramaarachchi is a lecturer in department of Computing, Rajarata University of Sri Lanka. Currently reading Doctral Degree at Wuhan University of Technology, in China. His research interests are Information Security, and Privacy Protection, and Image processing.