Dengue disease prediction using machine learning algorithms: a review

Tiu Kian Siang
School of Computing
Asia Pacific University of Technology & Innovation (APU)
Kuala Lumpur, Malaysia
TP060189@mail.apu.edu.my

Chandra Reka Ramachandran
School of Computing
Asia Pacific University of Technology & Innovation (APU)
Kuala Lumpur, Malaysia
Chandra.reka@staffmail.apu.edu.my

Dr. Fatemeh Meskaran
School of Computing
Asia Pacific University of Technology & Innovation (APU)
Kuala Lumpur, Malaysia
fatemeh.meskaran@staffemail.apu.edu.my

***Abstract*— Dengue fever is a mosquito-borne viral disease spreading in tropical and subtropical regions. It affects approximately 141 countries. Currently, there is no specific medication to treat the disease and supportive care is available. Vaccine is accessible but with many limitations. Because of that, one of the best solutions is to break the dengue cycle. The study aims to design an accurate and timely prediction system that can achieve effective targeting of possible dengue outbreak areas. From the literature review, Support Vector Machine (SVM), Random Forest (RF), Bayes Network (BN) and Bayesian Ridge Regression have shown their outstanding performances in several studies for dengue outbreak forecast system. The most important and commonly used predictors in outbreak forecast are meteorological (rainfall, temperature and relative humidity), entomological (mosquito population size, dengue virus serotypes in mosquitoes) and socioeconomic (region population household income, education status, employment rate, etc.). Both regression and classification models can be employed construct dengue disease spread predictive analysis. In Malaysia, there is a lack of research to compare various machine learning algorithms on dengue outbreak prediction. The goal of this research is to build an optimized model to predict dengue outbreak, so preventive measures can be taken to break the dengue life cycle in Malaysia. The model can help government to track and forecast the outbreak of dengue fever. In future research, other machine learning algorithms and training predictors can be added or modified to obtain a model with better performance.**

Keywords*— dengue outbreak, machine learning, predictive models

I. INTRODUCTION

Dengue fever is a serious arboviral infection transmitted through female mosquito in tropical and subtropical countries with about 141 countries affected. However, 70% of the actual cases are in Asia for instance Cambodia, Laos, Philippines, and Malaysia [1]. The primary transmission vector of dengue virus is Aedes aeyti and dengue virus is a flavivirus, consisting of 4 serotypes (DENV-1, DENV-2, DENV-3, and DENV-4) with almost no cross-immunity [1]. Fever, rash, headache, and joint pain are some of the symptoms of dengue fever. A retrospective research of cases in Malaysia reveals that different dengue serotype had distinct clinical manifestations in patients [2]. The World Health Organization (WHO) received 4.2 million reported cases of dengue in 2019. A modelling estimation of total 390 million dengue infections occur annually, which 96 million develop clinical signs and 500,000 cases turn into severe dengue [3]. Severe dengue also known as Dengue Haemorrhagic Fever (DHF), is a more serious form of dengue fever which can be fatal. The number of dengue cases increased by folds since the few decades, resulting to a great impact to human healthcare management. In Malaysia, till 31 October 2020, the cumulative cases reported reaches 82,753 [3]. Most of the cases in Malaysia occur during the late monsoon season.

Coronavirus disease 2019 (COVID-19) is an infectious disease caused by a novel coronavirus. It is hard to distinguish dengue fever from COVID-19 because both share some common clinical signs and laboratory results [4]. This will increase the workload of healthcare workers to differentiate them. Besides, severe dengue can be deadly. Controlling the number of dengue infections in a country can be beneficial.

Currently, there is no specific treatment for dengue fever, only supportive treatment is available. The first approved dengue vaccine available in some countries is CYD-TVD (Dengvaxia®). However, there are many limitations to the use of the vaccine. A research shows a lower efficacy of the vaccine on DENV-2 serotype [5]. The vaccine is recommended for people ages between 9 to 45, people beyond these ages proved to have lower efficacy. People with no history of dengue infection have the risk of developing severe dengue. Until now, the efforts to develop a better vaccine are still going on. Thus, prevention is still the best approach to tackle dengue problem [6].

Dengue virus is spread through the bite of Aedes mosquito (Aedes aegypti and Aedes albopictus). These mosquitoes are found primarily in tropical and subtropical countries, including Malaysia. Few infected mosquitoes can cause major dengue outbreaks in the community and put your family at risk of becoming ill. The most effective way of dengue prevention is through vector control. Prevention to control the vector. People from dengue affected countries should protect themselves against mosquito bites through applying repellent and avoiding area vulnerable to mosquitoes. They should also remove standing water that potentially be mosquito breeding site inside and outside their house. Furthermore, aerial spraying is used as an integrated part in vector elimination program to deal with dengue epidemic in large scale.

Machine learning has been used to tackle the dengue problem in outbreak prediction, vector estimation and disease diagnosis. Timely, accurate and precise prediction and diagnosis enables people to take precautions beforehand. Various machine learning algorithms were utilized to build the dengue predictive system over the years. This research is to identify some possible machine learning algorithms used to solve dengue problem. Some research will be reviewed in the next section.

II. MACHINE LEARNING (ML) APPLICATIONS IN HEALTHCARE

As a sub discipline of Artificial Intelligence (AI), machine learning (ML) is the study of algorithms or methods to discover patterns within and between data and make prediction from the data. The use of machine learning has driven advance in many domains such as business administration, transportation (self-driving car, automatic traffic control), manufacturing (machine anomaly detection) and education (timetable scheduling). Healthcare is the maintenance of good health condition by providing medical care to people with mental or physical health impairment.

As medical records become growingly digitized, this resulting an opportunity to provide a healthcare system that is capable to deliver medical care in a better and smarter way [7].

The increasing availability and accessibility of Electronic Health Records (EHR) has promoted the applications of machine learning in healthcare [8]. In the old days, healthcare data was not collected in a systematic way or was ignored completely and there was a lack of proper data storing method, making machine learning less famous. However, healthcare datasets are typically observational and often incomplete as a result and missing values are prominent problem that can affect the effectiveness of a predictive model [9].

Data in healthcare can be categorized into clinical, omics and sensor data. Clinical data is the data collected during patient treatment process. Omics data is the molecular or genetic detail of patient. Sensor data is the data obtained from sensor device also known as biosensors and the data collected usually in time series. Different authors suggest different machine learning algorithms for different types of data in healthcare [10]. It is only the matter of time before machine learning is thoroughly used in healthcare.

The implementations of machine learning in infectious disease are varying and comprise disease diagnosis, disease risk stratification and predicting the outbreak and spread of contagious disease [8]. Davenport and Kalakota [11] stated that the most traditional application of machine learning is in precision medicine, which is applied to predict the treatment protocol that has highest predicted success rate based on patient characteristics and the treatment details. The usage of convolutional neural network in medical images for disease diagnosis also becomes increasingly popular in the past few years because of the high capability of convolutional neural network in image pattern recognition. Machine learning technologies can also be implemented in serving more patients in a shorter period and improving health quality and reducing the cost of healthcare [12].

In clinical practice, machine learning's black box nature remains a great challenge, quantitative training has not been stressed and most doctors do not have a solid understanding interpretation of machine learning [7]. According to [13], interpretability in machine learning is the ability to explain and understand the reasoning behind machine learning models. Machine learning model in healthcare has less tolerance in misclassification because it is about clinical decision making that involves life. The ability to interpret predictive models reduces the likelihood of making mistake. Hence, machine learning in healthcare has higher demands for interpretability as compared to most of the other fields.

Interpretable machine learning allows user and programmer to investigate, understand, debug, and improve the models.

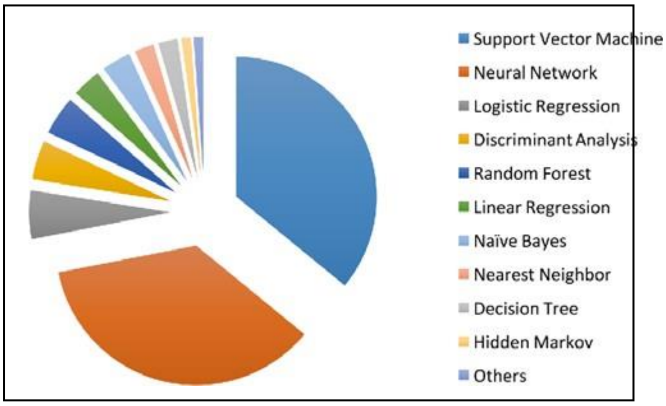

Fig. 1. The common machine learning algorithms in medical research [14]

Basically, machine learning algorithms can be categorized into supervised, unsupervised, or reinforcement learning. Most of the machine learning applications in medical field require a training dataset with outcome variable, which is known as supervised learning [11]. Fig. 1 illustrates the common machine learning algorithms found in medical literatures on PubMed. Support vector machine (SVM) appears to be the most famous algorithm followed by neural network.

III. OUTBREAK PREDICTION

Access to reliable outbreak prediction models is important to gain insight into the possible distribution and effect of infectious diseases for example in dengue outbreak. According to [15], the government and community can use the findings from predictive models to help them respond to dengue outbreak beforehand. It has been highlighted in [16] that forecasting is an important alert tool for clinical and public health providers that can assist with strategic planning and response. Ong et al. [17] further argued that outbreak or spread prediction models can help to achieve effective deployment of resources and reach maximum response impacts in mitigating the transmission of disease.

Several studies [15][18] have suggested that SVM was the optimal machine learning algorithm in dengue outbreak forecasting system. In [15], several machine learning algorithms such as generalized additive model (GAM), support vector machine (SVM) for regression and gradient boosted regression tree algorithm (GBM) were used as the candidates to develop a dengue outbreak prediction system in Guangdong, China and the SVM model has shown an outstanding performance for prediction of time series data compared to other machine learning models built with the smallest root-mean-squared error (RMSE) and higher R-squared value. In [18], the research aims to compare normalization techniques with various machine learning algorithms to predict dengue outbreak in Selangor, Malaysia and Least Squared-SVM with decimal point normalization has shown to be the best algorithm. Appropriate normalization techniques are capable to improve the performance of model.

Other recent studies [16][19] have proved that Random Forest (RF) was the ideal algorithm to build dengue outbreak prediction system. As shown in [19], random forest (RF) with delayed meteorological effects displayed the lowest RMSE (0.21) and MAE (0.15) values. [16] also demonstrated that random forest has better performance compared to artificial neural network (ANN) in predicting dengue burdens in Columbia. In reference [20], Bayes Network (BN) is the excellent machine learning algorithm for dengue outbreak prediction with a high accuracy of 92.35% and lowest RMSE (0.26). Another research [21] has shown that Bayesian Ridge Regression has the finest performance and suggested that Undecimated Fully Convolutional Neural Networks (UFCNN) may be the potential to get a better result, but further research required. Besides regression models, [22] has proposed a binary classification with only outbreak and non-outbreak to an area and has shown a good predictive performance.

The common predictors used to train the dengue outbreak models are climate factors also known as meteorological factors [15][19][20][23], search machine index [15] and entomological factors [16][17]. Moreover, [22] believed that socioeconomic factors are also significant to build the predictive system. Conversely, [24] stated that landscape factors alone produced predictability substantially better than socioeconomic factors. In addition, more diverse predictors are proposed to be added such as house index, container index and social media trending. Several studies [19][21][22] has proved the power of lagged effect in meteorological factors in improving the predictive models' performance. [22] highlighted that lagged effect 1 to 2 months proved to be effective and delayed or lagged effect can be applied separately to each climatic factor. [20] found a new risk factor, TempeRain Factor (TRF) and was used as input variable to enhance Bayesian Network (BN).

IV. DISEASE DIAGNOSIS AND PROGNOSIS

Emerging machine learning systems are starting to change medicine and healthcare and by enhancing the diagnosis of diseases. Machine learning methods can be applied for early disease diagnosis as well as severity prognosis. Dengue disease has 2 forms, a less severe dengue fever (DF) and a more severe dengue hemorrhagic fever (DHF). According to [25], in countries with infection prevalence, inexpensive, rapid, and accurate prediction of dengue severity in children can be an effective tool. It has been emphasized in [26] that these tools are very useful for disease detection and also provide an incentive for better decision-making. [27] also further expressed that machine learning is a vital part of achieving faster and more reliable results by using various algorithms.

In [25], regularized logistic regression, linear or gaussian SVM, Naive Bayes were used in early dengue severity prediction and gaussian SVM has shown to be the optimal algorithm with highest Area Under the Curve (AUC) of 0.81. In [28], ANN was applied to make prognosis using genome data and its accuracy is greater than 86%, 98% sensitivity and 51% specificity. An advantage of performing prognosis using machine learning with genome data is that it can be applied even before a person getting disease. [28] research can be extendable to other genetically related diseases.

Several studies [27][29] have built their system in dengue diagnosis instead of dengue severity prediction. In [29], the system was built using k-Nearest Neighbour (k-NN), SVM, ANN, LR, Naïve Bayes, LogitBoost; LogitBoost has shown to be the best algorithm with accuracy of 92%, sensitivity of 90%, specificity of 94% and receiver operating characteristic (ROC) area of 0.967. In [27], Classification and Regression Trees (CART) has been proved to have greatest classification efficacy with 20 out of 20 correct predictions compared to Multi-layer Perception (MLP) and C4.5. [30] has shown that Random Forest (RF) can be a potential early dengue detection algorithm with an accuracy of 83%. According to [26], the classification algorithms for dengue disease prediction are Decision Tree (DT), ANN, SVM and Rough Set Theory. Rough Set (RS) Theory appeared to be the best among these algorithms.

In disease diagnosis and prognosis, the common predictors are clinical data (symptoms such as headache, vomiting and blood test result such as white blood cells count, platelet count) [25][27][29], gene data [28] and protein data [23]. A larger volume of data [28] and appropriate feature selection can enhance the performance of a model [25].

V. VECTOR POPULATION AND BREEDING SITES ESTIMATION

Dengue fever is a vector-borne disease occurs mainly in tropical region, which transmitted by Aedes mosquitoes.

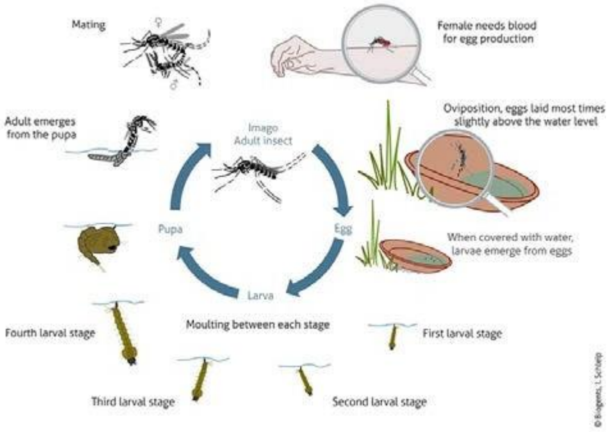

Fig. 2. Life cycle of mosquito [31]

Fig. 2 shows the life cycle of mosquito. Clear and stagnant water is the preferable breeding site for mosquito [32]. The whole life cycle from egg to adult takes about 1 week, a female mosquito manages to produce about 300 eggs in her lifespan [31]. With only a few numbers of mosquito, it can create a vast dengue outbreak. Since the transmission of dengue requires vector, one of the applications of machine learning in dengue prevention is to estimate the population of the mosquitoes or their breeding sites. According to [33], monitor vector population is important in vector-borne disease prevention by planning strategies and resources management.

[32] carried out a research on using drone images to investigate dengue vector population because of the cost effective of using drone to access dangerous places. Histogram of Oriented Gradients (HOG) algorithm was used to extract features from drone images and Support Vector Classifier (SVC) with different gamma value were built. SVC using gamma value of 1 has shown to be the fittest algorithm

with 94.33% recall and 91.14% precision. However, [34] has claimed that using drone can be relatively high cost, proper training required to appropriately control the drone and is very susceptible to bad weather. Alternatively, [34] has suggested the use of Google street view images identify potential water container through Convolutional Neural Network and found that developing container density maps from geotagged photos is a promising approach to providing accurate risk maps on a wide scale. In [33], remotely sensed data was used to estimate the dengue vector population and k-NN has shown to have the optimal performance with highest correlation (0.888), lowest MSE (0.494).

As demonstrated in [24], machine learning approach using socioeconomic and landscape pattern was applied identify mosquito abundances. A total of seven socioeconomic and seven landscape factors were used, and they found that landscape factors alone produced predictability substantially better than socioeconomic factors. Thus, they concluded that landscape factors, but not socioeconomic factors were the important variables for mosquito abundance detection.

VI. FACTORS OF DENGUE OUTBREAK

As mentioned earlier, dengue is an extensive matter in tropical and subtropical countries. Owing to ongoing global trends, including climate change and urbanization, the geographical range of dengue is expected to further increase [35]. The performance of machine learning predictive models highly depends on the selection of predictor variables. Thus, understanding of the factors affecting dengue outbreak becomes very crucial before utilizing them to construct a model.

In epidemics areas, there is a strong relation between dengue incidence and population of Aedes vector. As demonstrated in [36], climates factors such as rainfall, temperature, and wind velocity [23] play important roles in describing the temporal dynamics of dengue incidence in human populations. Similarly, [37] affirm that climate factors can influence the dynamics of the vector population and the transmission of disease, weekly temperatures above 18°C increased the mosquito abundance and humidity more than 75% reduced the abundance.

A predictive model elucidated the link between climatic factors (temperature, rainfall, and sunshine) and dengue incidence in southern Taiwan [38]. As shown by [38], these factors significantly correlated with dengue outbreak. Dengue incidence positively correlated with elevated temperatures, but inversely associated with precipitation and sunshine period.

As temperature increases, the dengue vector requires shorter time of development in their life cycle, which causes increased mosquito population. It can allow vectors to survive much longer than at lower temperatures and reach maturity faster.

Some studies suggest high amount of rainfall can lead to increase dengue incidence. The rainfall creates abundant water containers that serve as mosquito breeding sites. However, [37][38] suggested that heavy rainfall can reduce the survival rate of mosquito larvae. This claim further supported by [39], excessive precipitation will wash out these breeding sites, causing the death of larvae.

In [40], entomological (mosquito population size, dengue virus serotypes in mosquitoes), epidemiological (dengue cases, date of onset) and environmental (rainfall, temperature, humidity, and air pollution index) factors were used to build predictive system for two cities in Selangor, Malaysia and produced high accuracy by considering the lagged effect of meteorological factors. To add on, [19] has highlighted use of delayed or lagged effects for each meteorological factor in dengue outbreak prediction will produce more appropriate outcomes. This is due to the time needed to develop into adult mosquito and the fact that climate factors will not have an instant effect on dengue incidence.

Other than the factors mentioned above, there is a clear linkage between sociological factors such as population, urbanization, education, public health policy on dengue fever incidence [38]. However, further research is required to prove the connection. In machine learning as well as data mining, the classifying technique is known as one of the most crucial and critical tasks that each model should be emphasized to produce optimal results [41-44].

VII. DISCUSSION

In many fields, such as business management, transport (self-driving vehicle, automated traffic control), manufacturing (detection of computer anomaly), education (timetable scheduling) and healthcare, the use of machine learning has accelerated development. Machine learning provides alternative approaches in any of these domains. Dengue fever is a mosquito-borne disease that causing serious complication worldwide. In this paper, the applications of machine learning in tackling dengue can be divided into outbreak prediction, disease diagnosis or severity prognosis and vector population estimation.

For dengue outbreak prediction, Support Vector Machine (SVM), Random Forest (RF), Bayes Network (BN) and Bayesian Ridge Regression have shown their outstanding performances in several studies. In addition to regression models, binary classifier for outbreak and non-outbreak also exhibited excellent predictive capability in dengue disease. The predictors can be used in dengue outbreak prediction are meteorological, entomological factors and search engine or social media data. The meteorological factors such as temperature, relative humidity, and rainfall with multiple discrete lagged or delayed effects have proved to provide improved accuracy on machine learning algorithms built. Some studies also have demonstrated that appropriate normalization techniques and feature selection methods on different data and algorithms significantly improved the predictive accuracy of their models. For disease diagnosis or severity prognosis, LogitBoost, Classification and Regression Trees (CART), Random Forest (RF) and Rough Set Theory have demonstrated outstanding performance in several tests. For vector estimation, features were extracted from drone images, Google street view images and remote

sensed data. Machine learning algorithms such as Support Vector Classifier (SVC), Convolutional Neural Network and k-Nearest Neighbour (k-NN) were applied and manifested good outcomes. Different machine learning algorithms have different advantages and disadvantages. Random forest (RF) is good as it will automatically impute missing data and no data scaling required. It is very stable and insensitive to outliers and noise due to the large number of trees built. In regression, possible disadvantages are the predicted value will never exceed the highest value in training data and the computational time will be longer. Support Vector Machine (SVM) is a very flexible algorithms due to the various kernel functions available. It can be used to avoid over-fitting problem in machine learning process. However, although SVM is less sensitive to outliers but it is very sensitive to noise.

VIII. CONCLUSION

Machine learning has become increasingly popular in healthcare recently. Dengue outbreak prediction is one of the applications of machine learning in healthcare. This research is to identify the machine learning algorithms and variables used in dengue problem. Some machine learning algorithms such as Support Vector Machine (SVM), Random Forest (RF), Bayes Network (BN) and Bayesian Ridge Regression have demonstrated their power in forecast system. In the future, other relevant predictors or newly discovered risk factors can be involved in future studies. More diverse predictor data need to be added such as house index, container index and social media trending. However, difficulty in data acquisition is a limitation in machine learning. Further enhancement can be made by selecting good algorithms, predictors and design a novel ensemble classifier with the help of interpretable machine learning. We may conclude that for clinical and public health practitioners, forecasting is an alternative warning instrument that can assist with strategic planning and response.

References

[1] World Health Organization, Dengue and severe dengue (No. WHO-EM/MAC/032/E). World Health Organization. Regional Office for the Eastern Mediterranean, 2014.

[2] J. Suppiah et al., “Clinical manifestations of dengue in relation to dengue serotype and genotype in Malaysia: A retrospective observational study,” PLoS neglected tropical diseases, 12(9), e0006817, 2018.

[3] World Health Organization, Update on the dengue situation in the Western Pacific Region. Update, 2020.

[4] G. Yan et al., “Covert COVID-19 and false-positive dengue serology in Singapore,” The Lancet Infectious Diseases, 20(5), 536, 2020.

[5] E. Prompetchara, C. Ketloy, S.J. Thomas, and, K. Ruxrungtham, “Dengue vaccine: global development update,” Asian Pac J Allergy Immunol, 10, 2019.

[6] L. R. Bowman, S. Donegan and P. J. McCall, “ Is dengue vector control deficient in effectiveness or evidence?: Systematic review and meta-analysis,” PLoS neglected tropical diseases, 10(3), e0004551, 2016.

[7] M. Ghassemi, T. Naumann, P. Schulam, A. L. Beam, I. Y. Chen, and

[8] R. Ranganath, “A Review of Challenges and Opportunities in Machine Learning for Health,” AMIA Summits on Translational Science Proceedings, 2020, 191, 2020.

[9] J. Wiens, and E.S. Shenoy, “Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology,” Clinical Infectious Diseases, 66(1), 149-153, 2018.

[10] I. Y. Chen, S. Joshi, M. Ghassemi, and R. Ranganath, “Probabilistic Machine Learning for Healthcare,” arXiv preprint arXiv:2009.11087, 2020.

[11] A. Dhillon, and A. Singh, “Machine learning in healthcare data analysis: a survey,” Journal of Biology and Today's World, 8(6), 1-10, 2019.

[12] T. Davenport, and R. Kalakota, “The potential for artificial intelligence in healthcare,” Future healthcare journal, 6(2), 94, 2019.

[13] J. Akhil, S. Samreen, & R. Aluvalu, “The Future of Health care: Machine Learning,” International Journal of Engineering and Technology (UAE), 7, 23-25, 2018.

[14] M. A. Ahmad, C. Eckert, and A. Teredesai, “Interpretable machine learning in healthcare,” In Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics, pp. 559-560, 2018.

[15] F. Jiang et al., “Artificial intelligence in healthcare: past, present and future,” Stroke and vascular neurology, 2(4), 230-243, 2017.

[16] P. Guo et al., “Developing a dengue forecast model using machine learning: A case study in China,” PLoS neglected tropical diseases, 11(10), e0005973, 2017.

[17] N. Zhao et al., “ Machine learning and dengue forecasting: Comparing random forests and artificial neural networks for predicting dengue burdens at the national sub-national scale in Colombia,” bioRxiv, 2020.

[18] J. Ong et al., “ Mapping dengue risk in Singapore using Random Forest,” PLoS neglected tropical diseases, 12(6), e0006587, 2018.

[19] Z. Mustaffa, and Y. Yusof, “A comparison of normalization techniques in predicting dengue outbreak,” In International Conference on Business and Economics Research (Vol. 1, pp. 345-349), 2011.

[20] T. M. Carvajal, K. M. Viacrusis, L.F.T. Hernandez, H. T. Ho, D. M. Amalin, and K. Watanabe, “Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines,” BMC infectious diseases, 18(1), 1-15, 2018.

[21] F. Y. Nejad, and K. D. Varathan, “Identification of Significant Climatic Risk Factors and Machine Learning Models in Dengue Outbreak Prediction,” 2020.

[22] C. Sathler, and J. Luciano, “Predictive modeling of dengue fever epidemics: A neural network approach,” 2017.

[23] R. Jain, S. Sontisirikit, S. Iamsirithaworn, and H. Prendinger, “Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data,” BMC infectious diseases, 19(1), 1-16, 2019.

[24] N. Iqbal, and M. Islam, “Machine learning for Dengue outbreak prediction: An outlook,” International Journal of Advanced Research in Computer Science, 8(1), 93-102, 2017.

[25] S. Chen et al., S., “An operational machine learning approach to predict mosquito abundance based on socioeconomic and landscape patterns,” Landscape Ecology, 34(6), 1295-1311, 2019.

[26] W. Caicedo-Torres, A. Paternina., and H. Pinzón, “Machine learning models for early dengue severity prediction,” In Ibero-American Conference on Artificial Intelligence (pp. 247-258). Springer, Cham, 2016.

[27] M. Fatima, and M. Pasha, M., “Survey of machine learning algorithms for disease diagnostic,” Journal of Intelligent Learning Systems and Applications, 9(01), p.1, 2017.

[28] K. T. Swe, and P. T. Z. Tun, “Dengue Fever Classification Tool using Machine Learning,” International Journal Of All Research Writings, 2(12), 5-10, 2020.

[29] C. Davi et al., “Severe dengue prognosis using human genome data and machine learning,” IEEE Transactions on Biomedical Engineering, 66(10), 2861-2868, 2019.

[30] N. Iqbal, and M. Islam, “Machine learning for dengue outbreak prediction: A performance evaluation of different prominent classifiers,” Informatica, 43(3), 2019.

[31] N. Rajathi, S. Kanagaraj, R. Brahmanambika, and K. Manjubarkavi, “Early detection of dengue using machine learning algorithms,” International Journal of Pure and Applied Mathematics, 118(18), 3881-3887, 2018.

[32] Biogents, “Life cycle of aedes mosquitoes,” 2018. [Online]. Available: https://sea.biogents.com/life-cycle-aedes-mosquitoes/. [Accessed: Nov. 20, 2020]

[33] A. Amarasinghe et al., “A machine learning approach for identifying mosquito breeding sites via drone images,” In Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems (pp. 1-2), 2017.

[34] J. M. Scavuzzo, F. Trucco, M. Espinosa, C. B. Tauro, M. Abril, C. M. Scavuzzo, and A. C. Frery, “ Modeling Dengue vector population using remotely sensed data and machine learning,” Acta tropica, 185, 167-175, 2018.

[35] P. Haddawy et al., “Large scale detailed mapping of dengue vector breeding sites using street view images,” PLoS neglected tropical diseases, 13(7), e0007555, 2019.

[36] J. P. Messina et al., "The current and future global distribution and population at risk of dengue," Nature microbiology, 4(9), 1508-1515, 2019.

[37] L. Xu et al., "Climate variation drives dengue dynamics," Proceedings of the National Academy of Sciences, 114(1), 113-118, 2017.

[38] D. A. da Cruz Ferreira, C. M. Degener, C. de Almeida Marques-Toledo, M. M. Bendati, L. O. Fetzer, C. P. Teixeira, and Á. E. Eiras, "Meteorological variables and mosquito monitoring are good predictors for infestation trends of Aedes aegypti, the vector of dengue, chikungunya and Zika," Parasites & vectors, 10(1), 78, 2017.

[39] Y. H. Lai, "The climatic factors affecting dengue fever outbreaks in southern Taiwan: an application of symbolic data analysis," Biomedical engineering online, 17(2), 148, 2018.

[40] C. M. Benedum, O. M. Seidahmed, E. A. Eltahir, and N. Markuzon, "Statistical modeling of the effect of rainfall flushing on dengue transmission in Singapore. PLoS neglected tropical diseases, 12(12), e0006935, 2018.

[41] R. Ahmad, I. Suzilah, W. M. A. Wan Najdah, O. Topek, I. Mustafakamal, and H. L. Lee, "Factors determining dengue outbreak in Malaysia," PloS one, 13(2), e0193326, 2018.

[42] Y. C., C. Y. Chin, K. X. Chen, F. Z. K. Derick, Z. A. A. Salam., "Performance analysis of machine learning algorithms in breast cancer diagnosis" *Journal of Applied Technology and Innovation*, vol. 5, no. 2, pp. 12-22, 2021.

[43] X. Y. Chin, H. Y. Lau, Z. X. Chong, M. P. Chow, Z. A. A. Salam., "Personality prediction using machine learning classifiers" *Journal of Applied Technology and Innovation*, vol. 5, no. 1, pp. 12-22, 2021.

[44] S. W. Ling, N. K.Batcha, R. Logeswaran., "Crop Plantation Recommendation using Feature Extraction and Machine Learning Techniques" *Journal of Applied Technology and Innovation*, vol. 4, no. 4, pp. 1-4, 2020.