

LSTM-RNN Based Approach for Prediction of Dengue Cases in India

Anjelus Ronald Doni*, Thankappan Sasipraba

Faculty of Computing, Sathyabama Institute of Science and Technology, Chennai 600041, India

Corresponding Author Email: doni.cse@sathyabama.ac.in



<https://doi.org/10.18280/isi.250306>

Received: 18 March 2020

Accepted: 25 May 2020

Keywords:

deep learning, epidemic, LSTM, dengue, influenza, weather, geographical location, CNN

ABSTRACT

Data driven health care research is accentuated due to the Gargantua data available in the form of structured and unstructured. Forecasting the impact of infectious and epidemic diseases will be of major assistance to the health care industry. The objective of this work is to study the impact of dengue cases across India by applying the Deep Learning methodologies based on Long Short-Term Memory using Recurrent Neural Networks. The factors considered in the prediction method are climatic conditions, temperature, rainfall data, humidity and population considered for the period between 2014 and 2019. The activation function applied is ReLU and on training the model using LSTM the level of accuracy in forecasting the epidemic is over 89% for infection and for death it is 81%. The Root Mean Square Error values are also computed and it is observed that when the number of iterations is increased the error value decreased. The proposed methodology assists the Health care department to make safety precautions before the outbreak of the dengue fever.

1. INTRODUCTION

In the current trend of information technology, the rate of data grows exponentially. As the size of the data increases it is difficult for the human being to parse the data and build models on top of it. As on date huge volume of unstructured data is available in the health care industry. Machine learning is one of the component of data analytics which enables the automated building of models and interpret the results in an efficient manner [1]. The smart devices like watch, fit bits and other devices which collects data constantly and these data are to be processed, analyzed and interpreted for effective decision-making process [2]. As huge amount of funds is being diverted to health care industry, it is mandatory to assist the patients and physicians in their routine activities by efficiently and accurately diagnosing the symptoms with minimized cost.

Deep Learning plays a vital role in all the fields of Engineering and its applications. The deep learning has its impact in medical field, automobile industry, business forecasting/prediction, computer vision, video classification, pattern recognition, natural language processing, clustering, regression and voice recognition [3]. In machine learning it is possible to classify and predict but are not scalable. In deep learning, multiple layers are being introduced in order to arrive at a solution (in terms of classification and prediction). The concept of deep learning is derived from the structure of the brain and its neural system. The processing units are classified into input layer, hidden layer and output layer. Each layer is associated with weights and are updated in each iteration. The process of updating the weight depends on the activation functions. The activation functions can be of sigmoid, tanh or rectified linear unit functions. The results are obtained from the output layer which is considered to be the solution for the given problem statement. The steps involved in building the neural networks consists: a) training, testing and validation

data set, b) train the model, and c) predict the test data, and d) validate the model.

Neural networks are classified into feedforward neural networks, recurrent neural networks, radial basis function neural network, Kohonen self-organizing neural network and modular neural networks [4]. If the information forwarded from the input to output layer and there exist no loopbacks then the network is feedforward network. In RNN a loop back is formed i.e. the output layer is transformed as input layer for the next iteration. A network is provided with the option of memory to store the previous states and the data is trained repeatedly and the level of accuracy is improvised when compared with the feedforward model. The radial basis function neural network consists of input, hidden and output layers. The output is centered around a radial function and is implemented using Gaussian function in which each node is represented as a cluster center. Based on the weights and functions computed the approach is capable of classification or computing the inference. For organizing the input in the unsupervised learning methodology, the Kohonen algorithms is applied. The weights are computed using the Euclidean distance between the input and the output layer [5]. In modular network approach, the larger network is decomposed in to smaller and independent neural networks and each broken network processes the data and the outcome of each component is combined to provide as an output [6].

The deep learning networks are implemented using sparse autoencoders, convolutional neural networks, restricted Boltzmann machines and Long Short-Term Memory (LSTM). In the proposed model, the modified LSTM is applied in order to predict the epidemic diseases prevailing in the specific region. The neural network models can be further classified into discriminative and generative [3]. In the proposed model, the concept of discriminative approach is followed since the problem statement is prediction. In the proposed model the data is forwarded from the input layer to the output layer

through the hidden layer and is used in the classification and regression problems. The training in the deep learning networks can be broadly classified into supervised, semi-supervised and unsupervised. In the proposed model, the semi-supervised learning approach is implemented for predicting the prevalence of the epidemic disease.

In worldwide, influenza is one of the main root causes of death [7]. Soliman et al. [7] applied deep learning and machine learning algorithms for forecasting seasonal influenza in the city of Dallas, USA. Soliman et al. [7] have applied feedforward neural network approach with conventional statistical models like beta regression, autoregressive integrated moving average, least absolute shrinkage and selection operators and non-parametric multivariate regression splines. The authors have also applied a probabilistic forecasting of influenza using Bayesian model averaging method. For the feedforward network, number of hidden layers was 2 with 75 hidden nodes each and the deep learning structure is selected using cross-validation. The learning rate in terms of mean squared error is 0.005. Age, location or time are the factors that influence the impact of the non-communicable diseases such as cardio vascular disease, cancer, diabetes, hypertension, cholesterol, thyroid etc., and these factors are variable and changes continuously [8]. It has been claimed that 70% of the deaths are caused due to non-communicable diseases. The health of human and survival depends on heredity, social conditions, medical conditions and natural environment. Based on the data (can be structured or unstructured) collected, the non-communicable diseases are classified. Since unstructured data is also considered in classifying the disease, the level of accuracy has been improved. A fusion method RKRE based on ResNet and expert system has been proposed and attained an average correct proportion of 86.95% [8]. The process is divided into four phases labelled as data fusion, feature extraction, machine learning models and expert advice. In the data fusion process, the data collected from various sources (ICD-10, online, physical form filling and DADIAN 2015) are fused together into single database. The pre-processing of data takes place in this phase. In the second step, the process of feature extraction using rough sets and linear discriminant analysis (LDA) has been carried out. On selecting the feature, the data is processed using the machine learning models like Support Vector Machine (SVM), statistics based, entropy based, Artificial Neural Networks (ANN), Gradient Recurrent Unit (GRU), Dense k-Subgraph (DKS) and Residual Neural Network (ResNet) Knowledge embedding model (transE, transD, transA, transH and transR) if fed as input into the knowledge graph model (ANN, GRU, ResNet). The data is forwarded to analysis phase and are evaluated by experts.

As per the report [9], half of the population of the American citizens have one or more chronic diseases and 80% of money is spent for the treatment of the diseases. The amount spent for the chronic disease treatment is amounting to 2.7 trillion USD which is almost 18% of the entire GDP of USA. In China, the percentage of deaths due to chronic disease is 86.60. To assess the risk of chronic disease both the structured and unstructured data are merged [10] because of the following reason: incomplete information, missing data, region specific data, characteristics of the disease in different region, climate and social conditions or life style. The data which is collected from hospital is classified into S-Data, T-Data and S&T-Data for predicting the cerebral infraction disease by the applying the machine models: SVM, k-nearest neighbour and decision trees.

The data set is divided into training and test in the ratio of 6:1. The level of accuracy is approximately to 94.20% when the data set is predicted using convolutional neural network based multimodal disease risk prediction (CNN-MDRP). The authors [10] have computed defined CNN models separately for S-Data, T-Data and S&T-Data. It has been observed that the combination of structured and text data yields an increased level of accuracy, F1-measure, specificity, precision and recall. The level of convergence for the convolutional neural network algorithm exhibits an enhanced performance when the data considered is both structured and text.

The connection between the human beings and animals can spread epidemic diseases among themselves. These infectious disease data is available at large in the past decade due to the computerization of document prevailing in the hospital management systems along with timestamp [11]. As these temporal data are available with timestamp, it is possible to predict or forecast the expected epidemic diseases in the near future and thereby avoiding the calamities.

At present, combination of several models and binding them for generating the prediction by applying statistical components is referred to as ensemble learning. For the system that are noisy, complicated and dependent systems the concept of ensemble learning is applicable [12]. The prediction of infectious diseases can be broadly classified as agent-based models, compartmental models and regression-based time series models (auto regressive and seasonal terms) [13-19]. The data source includes; historical data, disease incidence time series, web crawls, Wikipedia page views, twitter, climatological tables etc. [12]. The models usually generate the following type of predictions: point predictions, associative predictive intervals or full predictive distributions [12]. Stacking is one of the ensemble models is trained using the cross-validated performance measure. The predictive density that are learnt from the component models are combined and the type of stacking is refers to as weighted density ensemble [20]. The clinical data is categorized into medical images (X-rays, computed tomography, magnetic resonance imaging, optical coherence tomography, microscopy image, positron emission tomography), clinical notes (electronic health records like discharge summaries, measurement reports, death certificates), lab results, vital signs and demographic informatics [21] and the type of diagnosis are auxiliary diagnosis, prognosis, and early warning. The authors have also discussed about the type of deep learning algorithm that can be applied based on the type of available data set. To predict the risk of disease, mortality, model patient trajectory and recommend medical dosing can be achieved by combining the data collected from the lab results and demographic information [22-25]. The monitoring devices such as electrocardiogram, multi-parameter monitors generate a large volume of data in an unstructured manner and these devices are placed at intensive care units and postoperative monitoring. Along with the above, the recent technological wearable devices, internet of things also provide live health care data on every single second of human life [26, 27]. The data collected from these devices are assisting in trajectory analyses of progress of the disease. If the prediction is carried out in advance, preventive measures can be taken in the clinical decision-making process [28].

The analysis of medical record/data set involves huge number of features. Based on the nature of the disease the impact of each feature varies accordingly. Therefore, selection of features is one of the major chores in the machine learning

frameworks. The framework in machine learning for ranking is referred to as “learning to rank” [29]. The supervised machine learning algorithms are applied for learning-to-rank approach in identifying the feature importance [30]. The learning-to-rank approach are trained using discriminative methodology and are classified into pointwise, pairwise and listwise with variant input and output space, hypothesis and loss functions [31-37]. The diagnosis for a patient is usually carried out by the empirical knowledge of the physicians which is based on the clinical reports and it sometimes might lead to misdiagnosis. Mainly in the treatment of psychiatric disorders it is a difficult task to precisely classify the type of patients and their diagnosis. A greedy deep weighted dictionary learning for classification of attention deficit hyperactivity disorder (ADHD) using the medical and big data has been carried out [38].

Precision public health is an emerging field for improving the overall health of the population by applying big data techniques [39]. Using big data methodologies, the surveillance of diseases is tracked using air pollution, antibiotic resistance, cholera, dengue, influenza and cough. Human carriers, victims or patients are tracked for the impact of epidemic diseases. Stoddard et al. [40] stated “Human movement is a critical, understudied behavioural component underlying the transmission dynamics of many vector-borne pathogens”. Detection of faults in the source code and evaluating them using random forest approach and is modelled using fuzzy entropy measure and Hurwicz criterion has been discussed. The developed model assists in identifying the behaviour of the features [41, 42]. The weights in the network are determined by Genetic Algorithm for improving the performance of the network classifier [43] and are applied using fuzzy networks. In the implementation process, in order to evaluate the correctness of the malfunction node or server the concept of aspect oriented programming can be introduced [44]. This enables to overcome some of the network anomalies in the prevailing system.

Dengue Fever is one of the most endemic disease in various cities of India, if the prediction or forecast of the number of cases in the expected regions based on the past history will assist the local government bodies to take appropriate precautionary measure and thereby the reducing the risk. In the proposed work, the forecasting of epidemic diseases specifically the cases of dengue fever are considered and are analysed by applying deep learning techniques.

2. METHODOLOGY

The principle objective of the proposed work is to restrain the spread of influenza like dengue during the peak weather conditions by predicting the estimated number of infections and location and thereby minimizing the encumber in terms of financial, health and mental suffering to both the Health officials and the individuals. The proposed model can be applied on the data collected or streamed from a specific location and able to predict the significant outbreak of the dengue fever in that region. The sequence of procedure followed in the proposed methodology is shown in Figure 1.

The data are collected from various sources are pre-processed. From the available features the attributes that are having impact are considered by applying XGBoost algorithm. The data is split into three categories: train, test and validate. The model is built and is trained by the training data set and the correctness of the proposed model is tested using the test dataset. The model is further evaluated by the validating data set. The results are then evaluated by applying the necessary measures and a recommendation engine is generated.

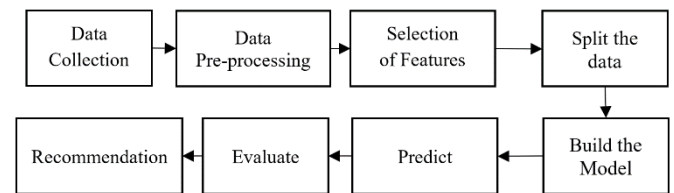


Figure 1. Flow of process for prediction of dengue fever

2.1 Data collection

The input data plays a crucial role in the recommendation system. The source of data consists of the population density in the given region, weather data of last one month, any cases that are reported with dengue positive result, any similar scenario occurred in other regions that closely match with the testing region, the species details that are responsible for causing and spreading the disease. The required data can be collected by streaming from the web scraping methodology.

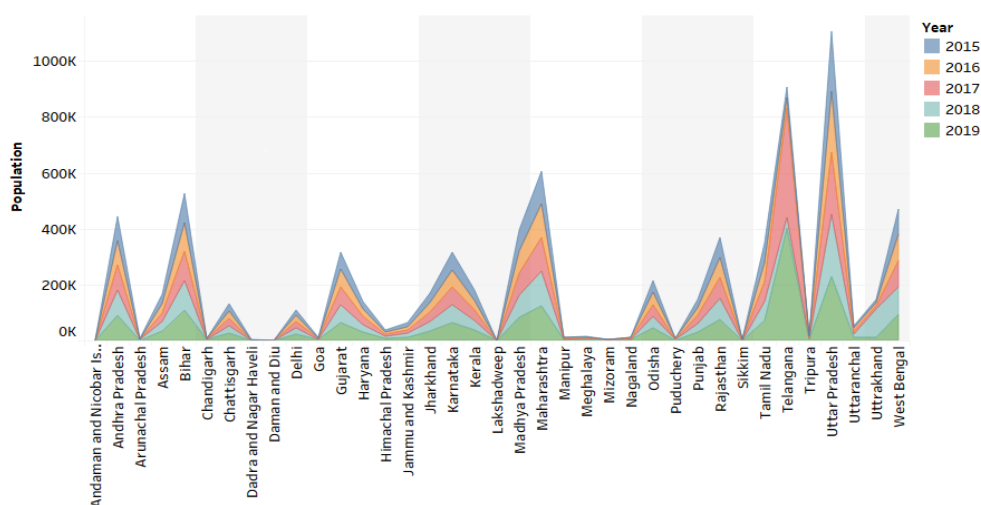


Figure 2. Projected population density in India based on 2011 census

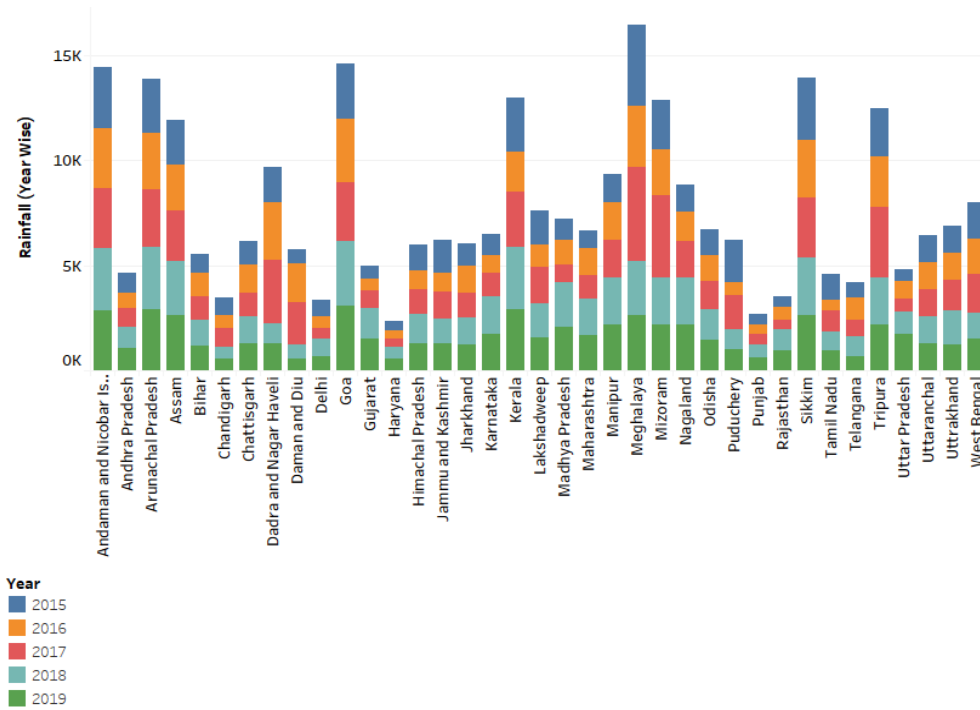
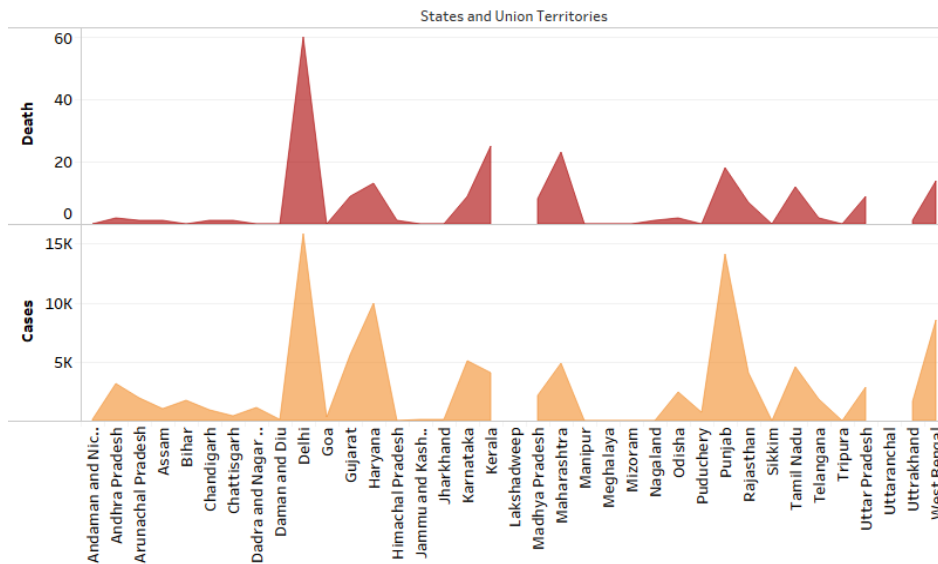
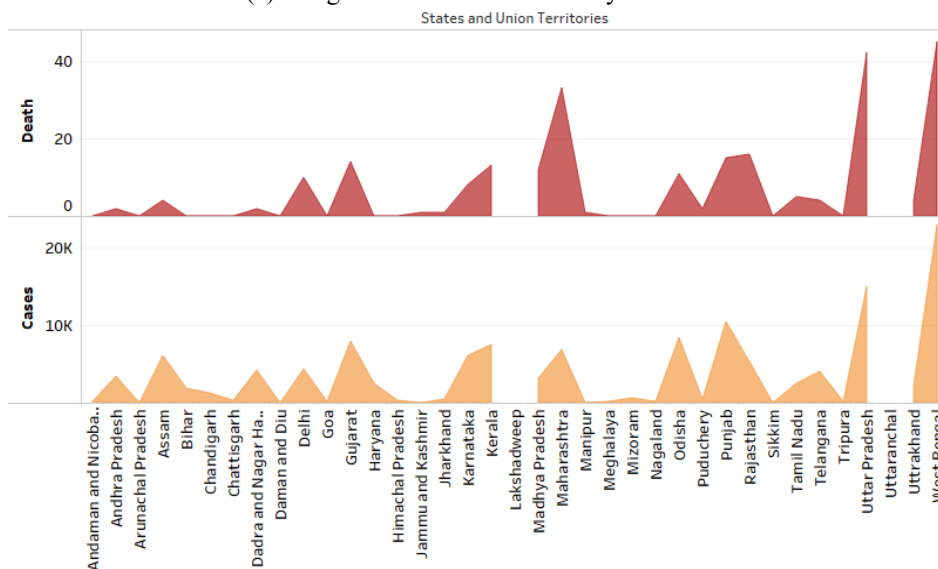


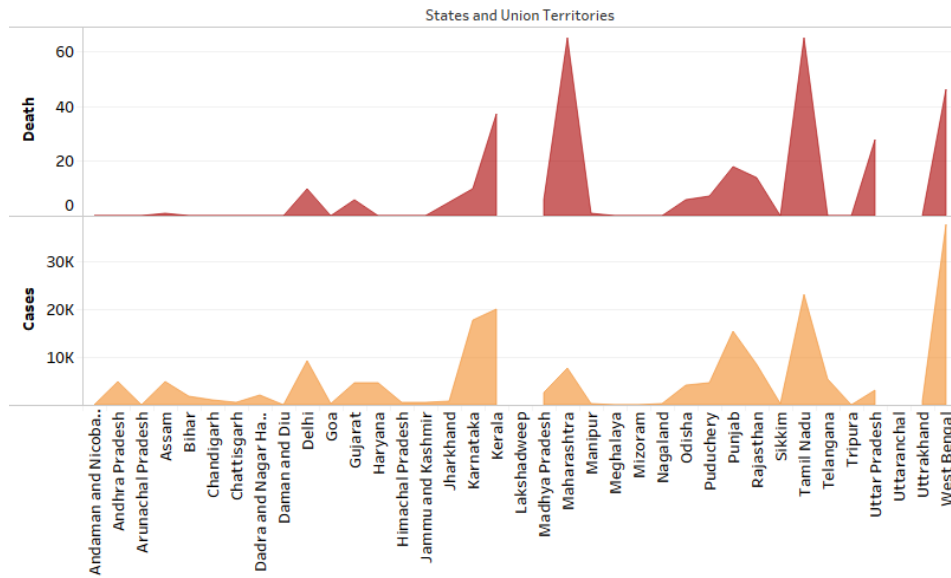
Figure 3. Average rainfall in India for the period from 2015 to 2019



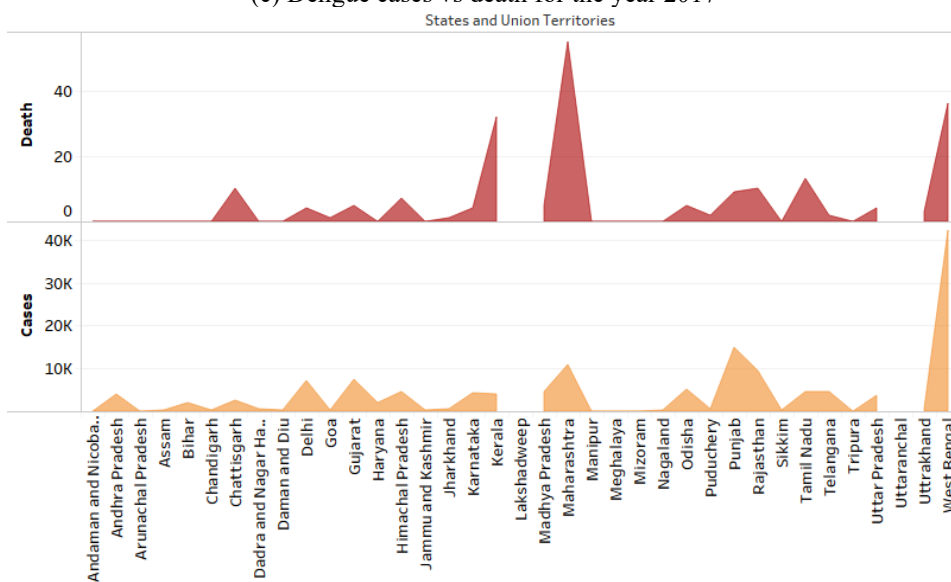
(a) Dengue cases vs death for the year 2015



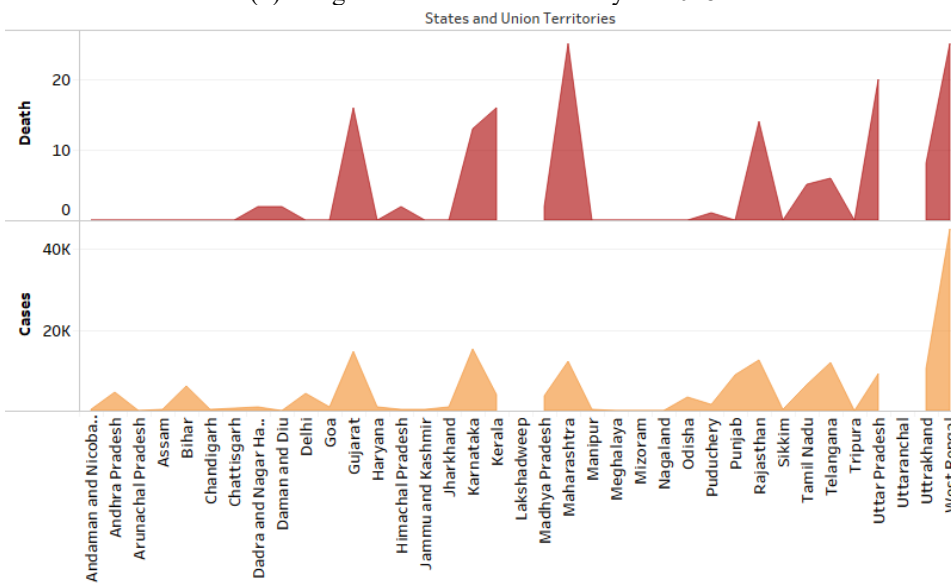
(b) Dengue cases vs death for the year 2016



(c) Dengue cases vs death for the year 2017



(d) Dengue cases vs death for the year 2018



(e) Dengue cases vs death for the year 2019

Figure 4. Dengue Cases Vs Death for the period 2015-2019

The initial source of input is the density of population in the target region and the data can be retrieved using online from the NIC portal of India. The data set provides the complete information based on the geographical location. The sample data is given in Figure 2.

The data contains the latitude and longitude of the given geographical location along with population density for the years from 2015 to 2019. The features include the name of the location, latitude, longitude, availability of water, land area measured in kilometre and the population density. Figure 3 shows the average rainfall for the period from 2015 to 2019 in India.

Figure 4a, 4b, 4c, 4d, 4e shows the dengue cases in India from the period from 2015 to 2019 across India. From the above inputs it is evident that the rate of dengue cases deviates over the years and it is observed that the variation is purely dependent on the factors of population, rainfall and humidity.

In the above charts it is evident that the rate of dengue cases (infected and death) is kept on varying consistently. It is observed that in some of the cities like Delhi, the cases have been brought down drastically whereas in case of West Bengal the cases are in rising between the period of study. In some of the places there is a constant maintenance of numbers

2.2 Data pre-processing

The transformation of data that is acceptable by the algorithm is referred to as the data pre-processing mechanism. In the input data set, some of the features are in categorical form, and few strategies are to be adopted in order to convert those features to numerical. The dengue affected regions are classified into non-occurrence and occurrence of the disease over a period of time. The features considered for the purpose of prediction is given in Table 1. The treatment of null values is carried out. For some states, the data is not available for some of the years (as those states are not formed) and those years are considered as null values and are treated accordingly.

Table 1. Features considered for the study

Feature	Notation
Population	Pi
Dengue Cases	Dc
Dengue Deaths	Dd
Temperature	Tt
Humidity	Hh
Rainfall	Rr

The features considered are the population (Pi) for the period from 2015 to 2019 (based on the 2011 census) and the data is converted to natural logarithm, number of dengue cases (Dc) that are reported across all the states and union territories of India with the data being converted to natural logarithm, number of deaths due to dengue (Dd) the data is in natural logarithm, average year rainfall (Rr) measured in centimetres, average temperature (Tt) prevailing during the entire year measured in degree centigrade and the final feature is humidity (Hh) with the percentage of RH.

2.3 ANN based model for prediction of dengue cases

The proposed model is based on the Deep learning methodology by adopting LSTM approach for the purpose of prediction and it has been compared with the existing machine learning models Back Propagation Neural Network (BPNN),

Random Forest (RF) Support Vector Machine (SVM) Classification, Gradient Boosting (XGBoost) and Stochastic Radiant Boosting and Generalized additive model (GAM). Initially the ANN architecture has been setup and it is represented in the Figure 5.

Figure 5 represents the generic structure of the proposed model for the prediction of dengue cases using artificial neural network (ANN). As per this model. The input given is population, rainfall, temperature and humidity. The history of dengue data is also being shared in line with the input and the data is processed using the hidden layers and the expected result is the prediction of number of dengue cases and the cases of death due to dengue. The above model is implemented by using the algorithms: LSTM, SVM, XGBoost, BPNN, GAM and Random Forest. The performance of the algorithms is compared.

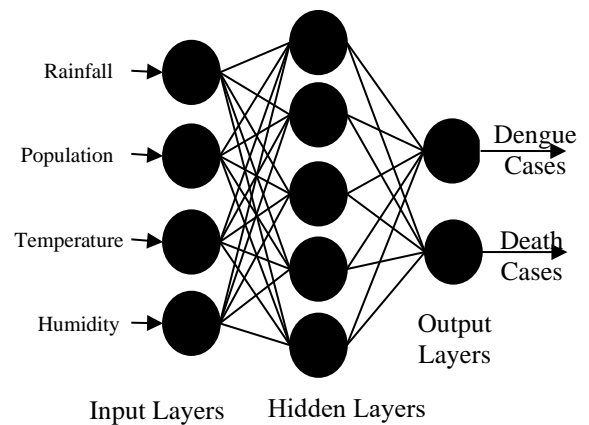


Figure 5. Structure of ANN process

In the proposed work, the data for the training is given for the period from 2015 to 2018 and the testing data considered is the year 2019. In the proposed work, the algorithm of LSTM on RNN is applied. Long Short-Term Memory (LSTM) is a part of enhanced Recurrent Neural Network (RNN) which depends on the memory cell with gating at three functions and is incorporated into the construction of the model [45]. The LSTM neural network has the capability to learn the interconnection between the attributes and is achieved by over a period of learning. From the given sample data set, the LSTM learns the network and the output is generated based on the weight adjustment (bias). In the proposed model, the features for the LSTM model are population, humidity, rainfall, temperature, dengue cases and dengue deaths for the period between 2015 and 2018. The input vector representation is given as $D_{xy} = (\ln(P_i), \ln(D_c), \ln(D_d), H_{havg}, T_{tavg}, R_{ravg})$. In a memory cell there will be 64 hidden layers. Figure 6 shows the architecture of the proposed model for the prediction of dengue using LSTM-RNN framework.

The LSTM-RNN based architecture for the prediction of dengue cases and deaths consists of a hidden layer which is the LSTM. The layer includes a set of 64 memory cells for computing the intermediate results and the resulting values are transferred to the next iteration. From the literature survey, it is provided that the initial starting value will commence from $\exp(-4)$. Since the data is available for a period of one year, the time step of LSTM is set to a value of 12. The activation function used in the proposed methodology is Rectified Linear Unit (ReLU). The performance of ReLU is comparatively good that the other activation functions like sigmoid and tanh.

To avoid overfitting of learning process, the dropout rate is set to a level of 50 percentage. The epoch was set to 100. The proposed methodology is implemented using tensorflow framework version 2.0 [45, 46] in google colab environment [47]. The Root Mean Square Error (RMSE) for the 12 iterations is shown in Figure 7.

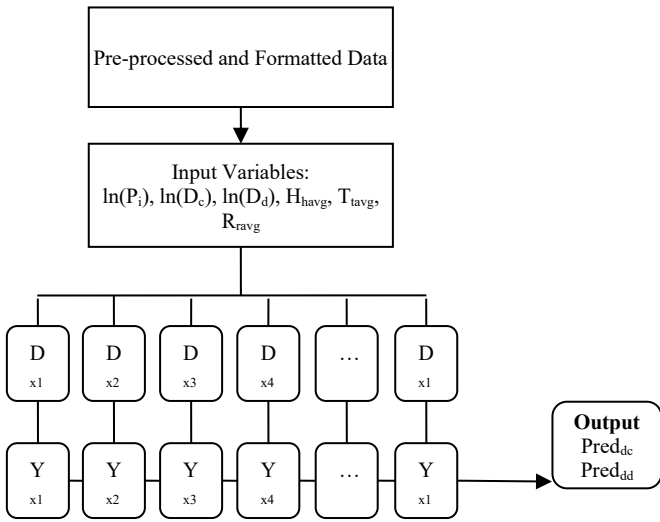


Figure 6. LSTM-RNN architecture for prediction of dengue cases and deaths

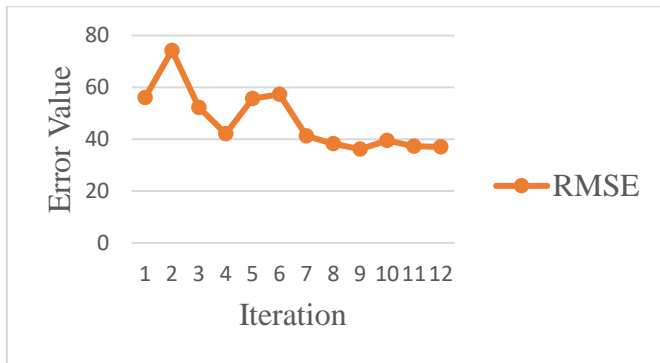


Figure 7. RMSE for hidden layer iterations

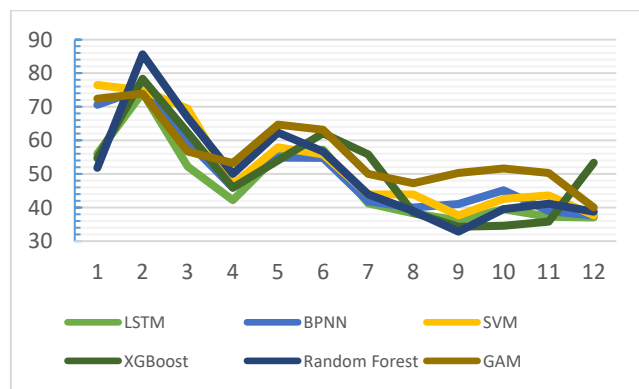


Figure 8. Comparison of RMSE of LSTM with other models

It is observed that the number of iterations is increased from the error values gets reduced. With the available data set, the prediction of dengue case is also evaluated using BPNN, SVM, Random Forest, GAM and XGBoost. Figure 8 shows the comparison of RMSE values of LSTM with the other machine learning models. It is observed that the LSTM computes with

the lower error value when compared with the other approaches.

From the trained model using LSTM, the prediction of Dengue cases for the year 2019 across India is shown in the Figure 9. Figure 9 shows the actual vs predicted cases of dengue across all the states and union territories of India. In a similar manner, Figure 10 shows the number of deaths due to dengue the actual and predicted values are shown.

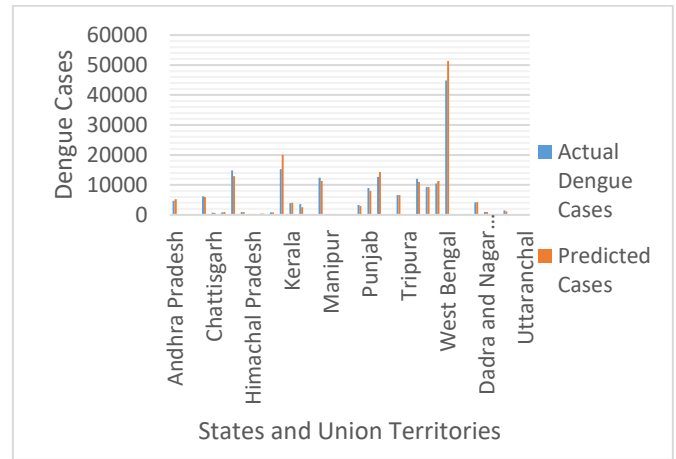


Figure 9. Actual vs predicted dengue cases for the year: 2019

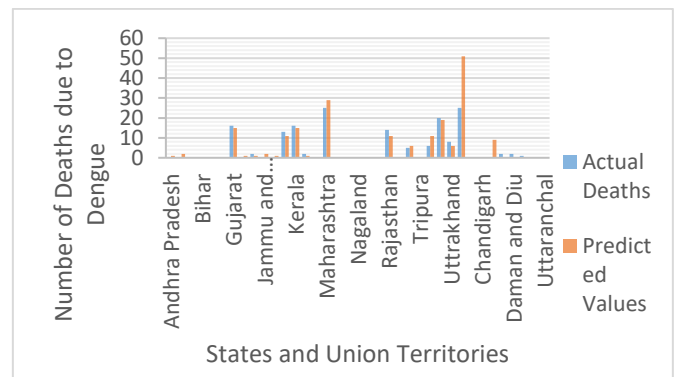


Figure 10. Actual vs predicted dengue death for the year: 2019

The analysis of the results shows that the a) in case of prediction of dengue cases the level of accuracy is close to 89% (by considering the overall prediction) b) whereas for the number of cases related to death the level of accuracy is close to 82%.

3. CONCLUSIONS

The cases of dengue in the states and union territories of India has been examined in this work using the concept of deep learning approach. The hidden layer LSTM and the activation function ReLU are applied in the developed model. The RMSE of other machine learning algorithms is also computed and it is evident that the LSTM exhibits better performance. Based on the training the model comes out with an accuracy level of 89% in case of dengue affected cases and 81% in case of deaths due to dengue. Even though the prediction level is not more than 95%, it is evident that the likeliness of the infection of dengue cases is increasing year by year and it can be used as a sign of alert to the Government to take appropriate

actions / steps towards the prevention of the disease. The trained model can be extended to transfer learning. The accuracy level of the predicted model can be improved by considering the additional datasets for over a period of 25 years. Further the model can be tested by increasing the number of epochs, other activation functions. The model is executed in a GPU environment, the performance can be tested by executing the model in TPU hardware.

REFERENCES

- [1] Bhardwaj, R., Nambiar, A.R., Dutta, D. (2017). A study of machine learning in healthcare. *Computer Software Application Conference*, 2: 236-241. <https://doi.org/10.1109/COMPSAC.2017.164>
- [2] Dolley, S. (2018). Big data's role in precision public health. *Frontiers in Public Health*, 6: 1-12. <https://doi.org/10.3389/fpubh.2018.00068>
- [3] Shrestha, A., Mahmood, A. (2019). Review of deep learning algorithms and architectures. *IEEE Access*, 7: 53040-53065. <https://doi.org/10.1109/ACCESS.2019.2912200>
- [4] Buhmann, M.D. (2003). *Radial basis functions*. Cambridge, U.K.: Cambridge University Press, p. 270.
- [5] Akinduko, A.A., Mirkes, E.M., Gorban, A.N. (2016). SOM: Stochastic initialization versus principal components. *Information Science*, 364-365: 213-221. <https://doi.org/10.1016/j.ins.2015.10.013>
- [6] Chen, K. (2015). Deep and modular neural networks. *Springer Handbook of Computational Intelligence*, Berlin, Germany, pp. 473-494. https://doi.org/10.1007/978-3-662-43505-2_28
- [7] Soliman, M., Lyubchich, V., Gel, Y.R. (2019). Complementing the power of deep learning with statistical model fusion: Probabilistic forecasting of influenza in Dallas County, Texas, USA. *Epidemics*, 28: 100345. <https://doi.org/10.1016/j.epidem.2019.05.004>
- [8] Lei, Z.F., Sun, Y., Nanekaran, Y.A., Yang, S.Y., Islam, M.S., Lei, H.Q., Zhang, D.F. (2020). A novel data-driven robust framework based on machine learning and knowledge graph for disease classification. *Future Generation Computer Systems*, 102: 534-548. <https://doi.org/10.1016/j.future.2019.08.030>
- [9] Groves, P., Kayyali, B., Knott, D., Kuiken, S. (2016). *The big data revolution in healthcare, accelerating value and innovation*. USA. Center for US Health System Reform Business Technology Office.
- [10] Chen, M., Hao, Y.X., Hwang, K., Wang, L., Wang, L. (2017) Disease prediction by machine learning over big data from healthcare communities. *IEEE Access*, 5: 8869-8879. <https://doi.org/10.1109/ACCESS.2017.2694446>
- [11] Masuda, N., Holme, P. (2013). Predicting and controlling infectious disease epidemics using temporal networks. *F1000Prime Reports*, 5: 1-12. <https://doi.org/10.12703/P5-6>
- [12] Ray, E.L., Reich, N.G. (2018). Prediction of infectious disease epidemics via weighted density ensembles. *PLoS Computational Biology*, 14(2): 1-23. <https://doi.org/10.1371/journal.pcbi.1005910>
- [13] Shaman, J., Karspeck, A. (2012). Forecasting seasonal outbreaks of influenza. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50): 20425-20430. <https://doi.org/10.1073/pnas.1208772109>
- [14] Shaman, J., Karspeck, A., Yang, W., Tamerius, J., Lipsitch, M. (2013). Real-time influenza forecasts during the 2012-2013 season. *Nature Communications*, 4: 2837. <https://doi.org/10.1038/ncomms3837>
- [15] Shaman, J., Kandula, S. (2015). Improved discrimination of influenza forecast accuracy using consecutive predictions. *PLoS Currents*, 7. <https://doi.org/10.1371/currents.outbreaks.8a6a3df285af7ca973fab4b22e10911e>
- [16] Yang, W., Karspeck, A., Shaman, J. (2014). Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. *PLoS Computational Biology*, 10(4): e1003583. <https://doi.org/10.1371/journal.pcbi.1003583>
- [17] Yang, W., Cowling, B.J., Lau, E.H., Shaman, J. (2015). Forecasting influenza epidemics in Hong Kong. *PLoS Computational Biology*, 11(7): e1004383. <https://doi.org/10.1371/journal.pcbi.1004383>
- [18] Yang, W., Olson, D.R., Shaman, J. (2016). Forecasting influenza outbreaks in boroughs and neighborhoods of New York City. *PLoS Computational Biology*, 12(11): e1005201. <https://doi.org/10.1371/journal.pcbi.1005201>
- [19] Chakraborty, P., Khadivi, P., Lewis, B., Mahendiran, A., Chen, J., Butler, P., Nsoesie, E.O., Mekaru, S.R., Brownstein, J.S., Marathe, M.V., Ramakrishnan, N. (2014). Forecasting a moving target: Ensemble models for ILI case count predictions. *The SIAM International Conference on Data Mining*, pp. 262-270. <https://doi.org/10.1137/1.9781611973440.30>
- [20] Wolpert, D.H. (1992). Stacked generalization. *Neural Networks*, 5(2): 241-259. [https://doi.org/10.1016/S0893-6080\(05\)80023-1](https://doi.org/10.1016/S0893-6080(05)80023-1)
- [21] Yu, Y., Li, M., Liu, L., Li, Y., Wang, J. (2019). Clinical big data and deep learning: Applications, challenges, and future outlooks. *Big Data Mining and Analytics*, 2(4): 288-305. <https://doi.org/10.26599/BDMA.2019.9020007>
- [22] Perotte, A., Ranganath, R., Hirsch, J.S., Blei, D., Elhadad, N. (2015). Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis. *Journal of the American Medical Informatics Association*, 22(4): 872-880. <https://dx.doi.org/10.1093%2Fjamia%2F23.4.872>
- [23] Carrara, M., Baselli, G., Ferrario, M. (2015). Mortality prediction in septic shock patients: Towards new personalized models in critical care. *37th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, pp. 2792-2795. <https://doi.org/10.1109/EMBC.2015.7318971>
- [24] Ghassemi, M.M., Richter, S.E., Eche, I.M., Chen, T.W., Danziger, J., Celi, L.A. (2014). A data-driven approach to optimized medication dosing: A focus on heparin. *Intensive Care Medicine*, 40(9): 1332-1339. <https://doi.org/10.1007/s00134-014-3406-5>
- [25] Nemati, S., Ghassemi, M.M., Clifford, G.D. (2016). Optimal medication dosing from suboptimal clinical examples: A deep reinforcement learning approach. *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, USA, pp. 2978-2981. <https://doi.org/10.1109/EMBC.2016.7591355>
- [26] Dunitz, M., Verghese, G., Heldt, T. (2015). Predicting

- hyperlactatemia in the MIMIC II database. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, pp. 985-988. <https://doi.org/10.1109/EMBC.2015.7318529>
- [27] Gunnarsdottir, K., Sadashivaiah, V., Kerr, M., Santaniello, S., Sarma, S.V. (2016). Using demographic and time series physiological features to classify sepsis in the intensive care unit. 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, Florida, USA, pp. 778-782. <https://doi.org/10.1109/EMBC.2016.7590817>
- [28] Lanata, A., Valenza, G., Nardelli, M., Gentili, C., Scilingo, E.P. (2015). Complexity index from a personalized wearable monitoring system for assessing remission in mental health. *IEEE Journal of Biomedical and Health Informatics*, 19(1): 132-139. <https://doi.org/10.1109/jbhi.2014.2360711>
- [29] Rahangdale, A., Raut, S. (2019). Deep neural network regularization for feature selection in learning-to-rank. *IEEE Access*, 7: 53988-54006. <https://doi.org/10.1109/ACCESS.2019.2902640>
- [30] Li, H. (2011). A short introduction to learning to rank. *IEICE Transaction of Information System*, 94(10): 1854-1862. <https://doi.org/10.1587/transinf.e94.d.1854>
- [31] Crammer, K., Singer, Y. (2002). Pranking with ranking. *Advances Neural Information Process System*, pp. 641-647.
- [32] Cossock, D., Zhang, T. (2006). Subset ranking using regression. *International Conference on Computational Learning Theory*, pp. 605-619. https://doi.org/10.1007/11776420_44
- [33] Li, P., Wu, Q., Burges, C.J.C. (2008). Mcrank: Learning to rank using multiple classification and gradient boosting. *Advances Neural Information Process System*, pp. 897-904.
- [34] Herbrich, R. (2000). Large margin rank boundaries for ordinal regression. *Advances in Large Margin Classifiers*. Cambridge, Massachusetts, USA: MIT Press, pp. 115-132.
- [35] Joachims, T. (2002). Optimizing search engines using click through data. 8th ACM SIGKDD International Conferences on Knowledge Discovery Data Mining. New York, New York, USA: ACM, pp. 133-142.
- [36] Burges, C.J.C., Ragno, R., Le, Q.V. (2007). Learning to rank with nonsmooth cost functions. *Advances Neural Information Process System*, pp. 193-200. <https://doi.org/10.1145/1150402.1150409>
- [37] Rigutini, L., Papini, T., Maggini, M., Scarselli, F. (2008). Sortnet: Learning to rank by a neural-based sorting algorithm. *SIGIR Workshop Learning to Rank for Information Retrieval*. (LRIR), 42(2): 76-79.
- [38] Wu, C., Luo, C., Xiong, N., Zhang, W., Kim, T.H. (2018). A greedy deep learning method for medical disease analysis. *IEEE Access*, 6: 20021-20030. <https://doi.org/10.1109/ACCESS.2018.2823979>
- [39] Dolley, S. (2018). Big data's role in precision public health. *Frontier Public Health*, 6: 1-12. <https://doi.org/10.3389/fpubh.2018.00068>
- [40] Stoddard, S.T., Morrison, A.C., Vazquez-Prokopec, G.M., Soldan, V.P., Kochel, T.J., Kitron, U., Elder, J.P., Scott, T.W. (2009). The role of human movement in the transmission of vector-borne pathogens. *PLoS Negl Trop Dis*, 3(7): e481. <https://doi.org/10.1371/journal.pntd.0000481>
- [41] Center for International Earth Science Information Network - CIESIN - Columbia University. (2018). Gridded Population of the World, Version 4 (GPWv4): Administrative Unit Center Points with Population Estimates, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). <https://doi.org/10.7927/H4BC3WMT>
- [42] Murugan, S., Ramachandran, V. (2012). Byzantine fault tolerance in SOAP communication services. *Malaysian Journal of Computer Science*, 25(2): 67-75.
- [43] Sheeba, P.T., Murugan, S. (2018). Hybrid features-enabled dragon deep belief neural network for activity recognition. *The Imaging Science Journal*, 66(6): 355-371. <https://doi.org/10.1080/13682199.2018.1483481>
- [44] Murugan, S., Ramachandran, V. (2012). Aspect oriented decision making model for byzantine agreement. *Journal of Computer Science*, 8(3): 382-388. <https://doi.org/10.3844/jcssp.2012.382.388>
- [45] Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. *Neural Computer*, 9(8): 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [46] Tensor Flow, <https://www.tensorflow.org/>.
- [47] Google Colab, <https://colab.research.google.com/>.