International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 06 | June 2021 www.irjet.net p-ISSN: 2395-0072

Classification of Leprosy using Artificial Neural Networks and Machine Learning

Professor Ms. Jalpa Mehta[1], Mr. Jaydeep Dharamsey[2], Ms. Pravalika Domal[3]

[1]Professor, Dept. Information Technology, Shah & Anchor Kutchhi Engineering College, Mumbai, India.
[2]Dept. Information Technology, Shah & Anchor Kutchhi Engineering College, Mumbai, India.
[3]Dept. Information Technology, Shah & Anchor Kutchhi Engineering College, Mumbai, India.

Abstract – *Leprosy is an irresistible sickness that causes serious, distorting skin wounds and nerve harm around the body. It is one of the significant general medical issues what's more, recorded among the ignored tropical illnesses in India. The bacteria Mycobacterium leprae causes leprosy. It's the idea that leprosy spreads through contact .This generally happens when an individual with leprosy sniffles or hacks. The various parts of the body like limbs, eyes, skin and nerves could get permanently damaged or severe if leprosy is not treated on time. Using fundamental signs of symptoms, this paper expects to portray order of leprosy cases with the help of Artificial Neural Network. Electronic Health Records (EHRs) of Leprosy Patients have been put in an Artificial Neural Network which is capable of selecting features automatically and this is the reason why it gives better results than machine learning algorithms like Support Vector Machine and Logistic Regression.*

Key Words: **Artificial Neural Network (ANN), Deep Learning (DL), Electronic Health Records (EHRs), Machine Learning (ML), Leprosy**.

1. INTRODUCTION

In medical science leprosy is one of the known diseases and can be found mostly in nations like India, Brazil, Indonesia and various parts of Africa. By the World Health Organization there were 208 617 new leprosy cases enrolled all around in 2018, as indicated by figures from 158 nations from the 6 WHO Regions. Predominance rate relates to 0.3/10 000 on the 184 213 cases toward the finish of 2018.

The WHO drafted the plan of eradication of leprosy for the upcoming 5 year from 2016 .Universally 200,000 new cases were reported by WHO in the year 2016. Close to 19,000 children were resolved to have Hansen's disease in 2016, more than 50 every day. Estimated amount of 2 to 3 million people are living with leprosy related disabilities all around the world. Many people in the world consider India to be the country for diagnosis of the leprosy cases where as many groups disproportionately get affected by the disease. 100 to 300 cases are registered of leprosy in the USA.

Worldwide G2D leprosy cases were identified and were found to have a decrease of 3.7% in the number of new cases for 2016. When compared with the previous year 2015, decrease incapacity rate was from 4.4% to 3.9% throughout the world. The Grade 2 Disability (G2D) has increased with adding new cases in India where it was 1-2% in 2005 for year then it rose to 3-4% for 2010 and in year 2014 it escalated to 4-5% according to the NLEP site.

For 2015–2016 the G2D patients were found to be 5,852 from the new leprosy cases which showed an extremely marginal decrease. With increase in percent of G2D cases it seems leprosy is lately identified within the community. Among all this G2D leprosy in kids is an add-on number to the total number of cases. Excluding India, rest of the world insights from WHO were given about G2D cases in young where, only 210 nations were reachable out of which zero kids with G2D were found in 195 nations and rest 15 nations have around 300 cases. In 2015, new 11000 leprosy kid cases, and the kids with G2D cases were 1.5% which was noticed by NLEP.

Recent findings mentioned by the National Leprosy Eradication Programme (NLEP) was that there was an imbalance found between the detecting cases and severity occurrence in the cases.Dy. General Director for leprosy (India) put forth the points which were necessary to be focused on in the month of August 2016.Many places were noticed where transmission of leprosy could take place. Indian Council for Medical Research (ICMR) put forth many undetected sections. The recognition rate for the new cases has remained nearly the same since 2005 and four, the deferral in determination made rise in the number of new handicap cases. To reduce this difficulties NLEP suggested methods which was divided into three parts. The first part was running campaign in highly prone places. Second was about spreading awareness with the help of various NGOs especially where Grade 2 Disability cases were noticed. Lastly, specific plans for difficult to arrive regions.

The constant neuritis and disability has been one of the effects that can been seen due to the postponement in finding leprosy and lepra cases. An extensive mindfulness is a requirement to know about leprosy signs and symptoms and its responses to all working medicinal staff from the early analysis till its complexities with proper administration. Considering the necessary examinations, for example, like the cutting edge PCR methods, skin smear was found to be a useful test which should be brought again into

the leprosy program of India. Recognizing smears and treating less high endemic areas and could be profoundly useful in reducing the number of people with leprosy disabilities.

The proposed system takes into consideration the key points put up by WHO and NLEP about the increase in cases of leprosy yearly is majorly due to late diagnosis. There are several ways to diagnose leprosy .Our system is designed to help doctors classify what type of leprosy that person belongs to based on the first sign and symptoms and clinical notes. Deep learning artificial neural network model is applied on the EHRs of the patients to classify leprosy using WHO and Ridley Jopling classification methodology.

2. BACKGROUND

According to WHO, in total there are two main classification of Leprosy namely PB and MB. Further leprosy is classified into 6 categories TT, BT, BB, BL, LL and PNL according to Jopling Classification. Each type of Leprosy has some unique feature which makes it different from others. Some of the signs and symptoms recommended by the doctors are the input parameters for the ANN model to classify Leprosy types.

- Swelling - Swelling is amplification of any body part, skin or organs.
- Erythematous - A greek word which represents redness of skin.
- Nerves Damage - There are a total nine nerves which can be affected by getting thickened or tendered on the left or right side. Following 9 nerves are Supra-orbital, Great Auricular, Ulnar, Median Cutaneous, Radial Cutaneous, Lateral Popliteal, Sural Nerve, Posterior Tibial and Superficial Peroneal.
- Diffuse Infiltration - Skin of the patient becomes shiny and wrinkles become invisible.
- Reddish and Punched out Lesion - Part of the skin is raised and is red in color.
- Numbness - No sensation is present in the affected region.
- Smear Value - A test taken from the blood samples collected from the affected area. This test consists of MI and BI values, any positive value can declare the patient as Leprosy positive.

(**Note**: These are very few of the many symptoms as we have used these in our system we are mentioning these. Also the parameters have different levels for e.g. patient might feel hot sensation numbness but can clearly feel cold sensation.)

Table -1: Leprosy and its classification

WHO Classification		Jopling Classification					
PB	MB	TT	BT	BB	BL	LL	PNL

3. LITERATURE REVIEW

The Rule based classification and machine learning models like SVM and Logistic Regression are applied to EHRs of leprosy data for classifying the type of leprosy and the results of all models are compared [1].Comparison between machine learning approach and artificial neural networks using various case studies and analysis [2].This paper gives an idea of how machine learning or artificial neural network algorithms and models can be designed and applied in healthcare units [3].

The severity prediction of the gallstone disease is carried out using Modcnn and ANN and comparison among both the models is done [4].This focuses on how data preprocessing can be done using Natural Language Processing techniques before feeding data to the deep learning model [5].Multiclass based classification using machine learning algorithms like Naïve Bayes, SVM, decision tree, logistic Regression and also deep learning algorithms are used to get user behavior patterns [6].

Prediction using a sequential model was developed for giving direct matrix conversion used for ac to ac conversion [7]. Early prediction of cancer using neural networks algorithms, label encoder and preprocessing techniques [8].Finding out risk prediction through deep learning architectures on highly unstructured electronic healthcare data [22].A monitoring system was developed for sickle cell disease using various machine learning algorithms like Support Vector Machine (SVM) from collected patient data and alert messages were sent to staff about patient condition [20].

The system outputs whether the person has heart disease or not by data fed through knowledge based techniques to the machine learning model like Logistic Regression [21].Information on leprosy classification according to WHO (World Health Organization) [10].A fully comprehensive information on the leprosy, its types, effects and impacts [9].Various evaluating parameters such as accuracy, f1 score, recall and precision can be calculated on artificial neural networks models or on others neural networks models [19].

Natural language processing techniques and mining options which are applied to convert unstructured data to structured form which then can be used for various algorithms [11].Significance and strategies for preprocessing data such as removal of stop words from an unstructured information for better output results [16].Creating corpus for getting clinical notes with proper clinical terms and spell check [15].Deep learning models can be trained using activation functions like softmax on pathology data generated from iot sensors [17].Diagnosis of faulty machineries using rectified linear units(ReLU) function to get advantages like better accuracy then other linear models and dissipate gradient problems [18].

Prediction on choice of travel mode using neural networks and encodings techniques like label encoder or one-hot encoder for getting better categorical features in data [14].Various trends and techniques in deep learning used for various data types whether it can be an image, text or audio and how we can incorporate it in various models or algorithms [13] .Neural networks models used for prediction of any type information from the charts or from EHRs of patients [12].

4. METHODOLOGY

4.1 Artificial Neural Network Approach

4.1.1 Labeling Data. (Preprocessing Phase)

Artificial Neural Network Model only to be fitted in a proper labeled format. Columns having not more than two classes were encoded using Label Encoder. Columns having multivariate values were labeled using OneHotEncoder.

4.1.2 Splitting Dataset.

It is always necessary to split our data into training and testing which is why we have splitted it into 80-20 ratio. Also to get a random set of pairs each time we have used a K Fold method and we have set the value of k as 3.

4.1.3 Modeling.

Sequential model was the best suitable model for this type of data. The Ann model was different for PB and MB classification (WHO) and for Jopling classification which was having 6 different classes.

- For WHO classification:

Sequential model was given an input layer with dimensions of 22 and these 22 nodes will connect to a hidden layer with an activation function relu and produce 40 nodes in the first hidden layer, second hidden layer will consist of 20 nodes which will give output of 6 nodes with relu activation function. Input layer was then connected to two hidden layers with the same relu model and finally this got compiled using binary_crossentropy and gave a single output node.

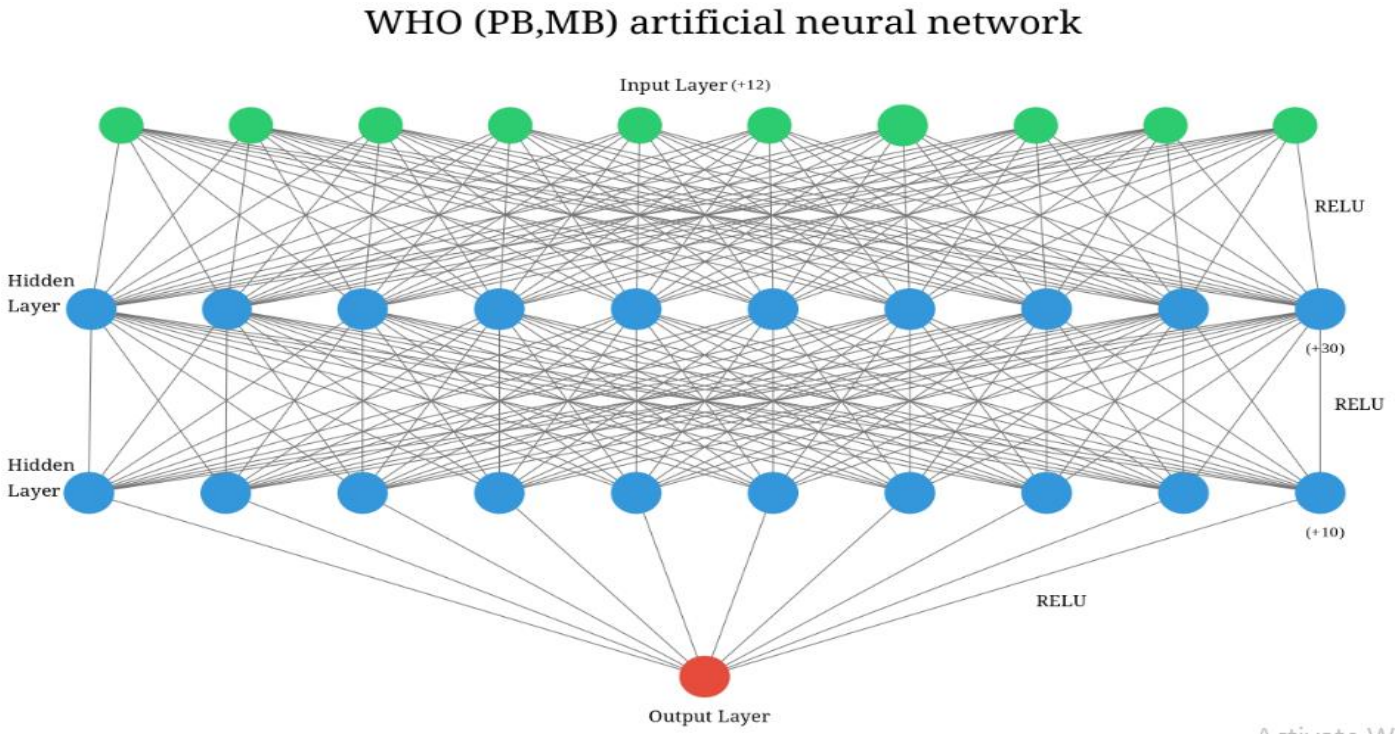

Fig -1: WHO Artificial Neural Network

- For Jopling Classification:

First input layer was added with the input of 22 different nodes and added with the activation function of relu. Hidden layer of 40 nodes is added with softmax activation function and output of 6 nodes was computed with same softmax activation and categorical loss function.

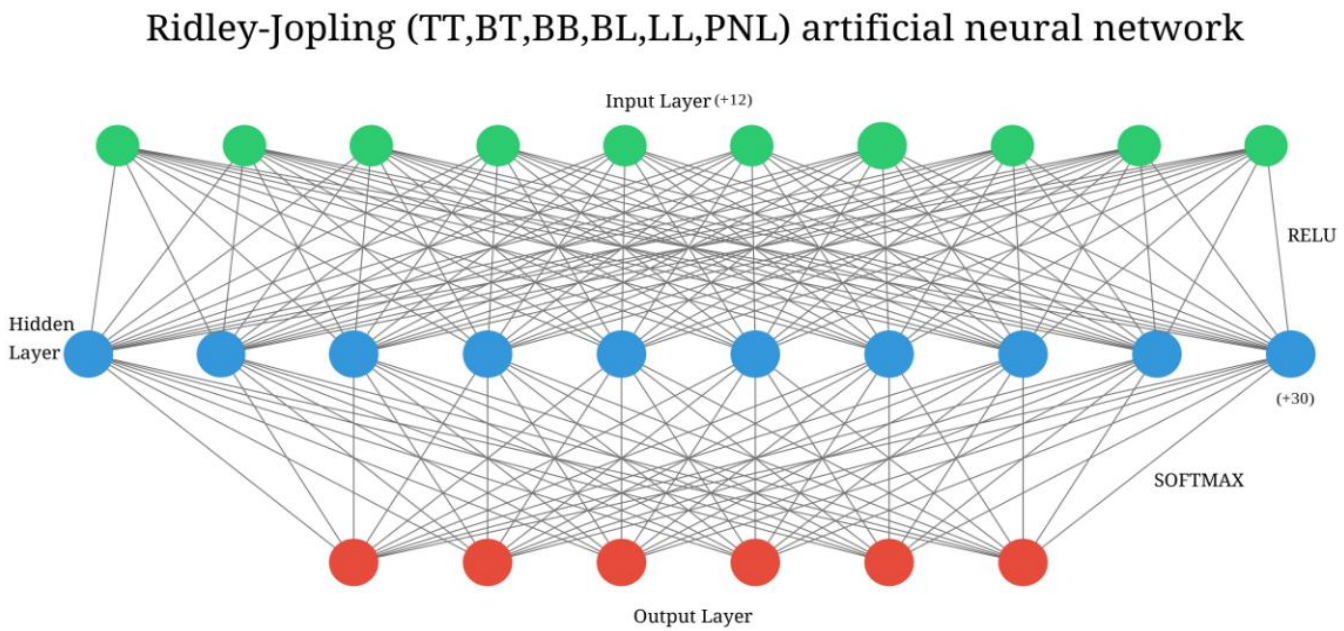

Fig -2: Ridley-Jopling Artificial Neural Network

4.2 Machine Learning Approach

4.2.1 Data Preprocessing.

Used several techniques like tokenization, lemmatization and removal of stop words and handling null values efficiently.

4.2.2 Data Modeling.

Splitting data into training and testing and applying kfold to take random samples of data. Fitting data into models of different machine learning algorithms (Logistic Regression, SVM).

4.2.3 Calculating Results

Several factors like accuracy, F1 score, precision, support and recall are computed in order to compare the final output.

3. RESULTS AND CONCLUSIONS

It can be clearly seen that the Artificial Neural Network performs better in multiclass attributes than machine learning as ANN selects parameters on its own which makes it more reliable. Although, machine learning approach gives better results in WHO classification, as the number of records will increase it is expected that ANN will give better results as compared to ML approach because of its feature of automatically selecting parameters.

Table -2: MBPB (WHO Classification)

Algorithm	Precision	F1 Score	Recall	Accuracy
Logistic Regression	95.80%	95.80%	95.80%	96.60%
SVM	97.80%	98.90%	100%	98.70%
ANN	95.56%	97.73%	100%	95.83%

Table-3: Other Types of Leprosy (Ridley-Jopling classification)

Algorithm	Precision	F1 Score	Recall	Accuracy
Logistic Regression	95.80%	95.80%	95.80%	96.60%
SVM	97.80%	98.90%	100%	98.70%
ANN	95.56%	97.73%	100%	95.83%

ACKNOWLEDGEMENT

We would like to thank Bombay Leprosy Project, Mumbai, India for providing patients' clinical record and Research cell of shah and anchor Kutchhi engineering college for providing us this opportunity to work with BLP.

REFERENCES

[1] Jalpa M., Jaydeep D., Pravalika D., Dr. Pai. Clinical Text Mining of EHRs to Classify Leprosy Cases.IJITEE. ISSN: 2278-3075, Volume-9 Issue-5, March 2020.
[2] Maity, G.N, & Das, S. Machine learning for improved diagnosis and prognosis in healthcare. IEEE Conference.2017, doi:10.1109/aero.2017.7943950
[3] Tran Le,Nguyen,Thi Thu Ha,Do.Artificial Intelligence in Healthcare:A New Technology Benefit For Both Patients and Doctors.2019 PICMET 978-1-890843-39-7.
[4] Likewin Thomas,Manoj Kumar ,Annappa. A Healthcare management using clinical decision support system 2018IEEE. 978-1-5386-6894-8/18
[5] Subiksha, K. (2018). Improvement in Analyzing Healthcare Systems using Deep learning Architecture.International 4th Conference of Computing,Communication and Automation (ICCCA). doi:10.1109/ccaa.2018.8777545
[6] Krishna Kumar Mohbey1. "Multi-class approach for user behavior prediction using a deep learning system on twitter election dataset". October 2019.Springer Nature Switzerland AG 2019, https://doi.org/10.1007/s42488-019-00013-y.
[7] Zhang, J., Li, L., Norambuena, M., Rodriguez, J., & Dorrell, D. G. Sequential Model Predictive Control of Direct Matrix Converter without Weighting Factors.Conference of the IEEE on Industrial Electronic Society.2018, doi:10.1109/iecon.2018.8592697
[8] Khuriwal, N., & Mishra, N. Breast Cancer Diagnosis Using Deep Learning Algorithms.(ICACCN) .2018,doi:10.1109/icacccn.2018.8748777
[9] Nunzi, E., & Massone, C. (Eds.). Leprosy.2012, Doi: 10.1007/978-88-470-2376-5
[10] https://www.who.int/lep/classification/en/
[11] Solangi, Y. A., Solangi, Z. A., Aarain, S., Abro, A., Mallah, G. A., & Shah, A. "Review on NLP and Sentiment Analysis and opinion Mining using toolkits". (ICETAS). 2018,doi:10.1109/icetas.2018.8629198
[12] Rajkomar, Oren, E., Chen, K., Dai, A. M., Haj, N., Hardt, M., Dean, J. Scalable, accurate deep learning using (EHRs) electronic health records. Digital Medicine, (2018)1(1). doi:10.1038/s41746-018-0029-1
[13] Hatcher, W. G., & Yu, W. A Survey under Deep Learning Applications,Platforms and Emerging Research Trend. IEEE Access, 2018, 6, 24411–24432. doi:10.1109/access.2018.2830661
[14] YIXUAN MA AND ZHENJI ZHANG. Travel Mode Choice Prediction Using Deep Neural Networks with Entity Embeddings. Received March 20, 2020, accepted March 31, 2020, date of publication April 3, 2020, date of current version April 17, 2020. Doi: 10.1109/ACCESS.2020.2985542
[15] Dalianis, H. Clinical Text Mining.2018, Doi: 10.1007/978-3-319-78503-5
[16] Ghag V., & Shah K."Comparative analysis on stopword removal for sentiment classifications". (2015).IC4.doi:10.1109/ic4.2015.7375527
[17] Amni, S. U., Hosain, M. S., Muhamad, G., Alhussein, M., & Rahman, M. A.Cognitive Healthcare Smart system for Pathology Detection and Monitoring. IEEE Access, 2019, 1–1. doi:10.1109/access.2019.2891390
[18] You, W., Shein, C., Wang, D., Jaing, X., & Zhu, Z. "An Intelligent Deep Learning model with Improved Activation Functions for Machine Fault Diagnosis".IEEE,2019 doi:10.1109/access.2019.2962734
[19] Latif, S., Zou, Z., Idrees, Z., & Ahmad, J. A Novel Attack Detection Scheme for the Industrial Internet of Things using a Lightweight Random Neural Network. IEEE Access, 2020, 1–1. doi:10.1109/access.2020.2994079.
[20] Abd, D. H., & Al-Mejibli, I. S. Monitoring System for Sickle Cell Disease Patients by Using Supervised Machine Learning. 2nd International Conference of Al-Sadiq on Communication Science and Multidisciplinary in IT and Application, (2017), doi:10.1109/aic-mitcsa.2017.8723006.
[21] Saw, M., Saxena, T., Kaithwas, S., Yadav, R., & Lal, N. Estimation of Prediction for Getting Heart Disease Using Logistic Regression Model of Machine Learning.International Conference on Computer Communication and Informatics (ICCCI), (2020). doi:10.1109/iccci48352.2020.9104210.
[22] Ayala Solres, J. R., Dileta Raimond, Rahimian, F., Canoy, D., Tran, Salimi-Khorshidi, G.Deep learning using electronic health records: A comparative review of multiple deep neural architectures. Journal of Biomedical Informatics, 101, 103337. doi:10.1016/j.jbi.2019.103337.

BIOGRAPHIES

Ms. Jalpa Mehta, Assistant Professor, Information Technology Department, Shah and Anchor Kutchhi Engineering College, Mumbai, India. She has received M.Tech. Degree in Computer Engineering from VJTI, University of Mumbai. Her research interests include Data science, Image processing, natural language processing, machine learning, recommender systems, and healthcare informatics.

Mr. Jaydeep Dharamsey, Student, Information Technology Department, Shah and Anchor Kutchhi Engineering College, Mumbai, India. He is in the final year of completing his B.E. degree. His research interest includes natural language processing, big data analytics, data science, text mining etc.

Ms. Pravalika Domal, Student, Information Technology Department, Shah and Anchor Kutchhi Engineering College, Mumbai, India. She is studying in the last year of her B.E. degree. Her research interests include machine learning, natural language processing, data mining , deep learning etc.