JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A SURVEY ON BIPOLAR OCCURRENCE PREDICTION USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES

[1]Sri Vasudev R Pyati, [2]Jayashree M Kudari, [3]Aditya Kumar, [4]AnInish S Shenai,[5]Abdullah Gafori, [6]Eze Ejiofor Isaiah
[1]UG Scholar, [2]Associate Professor, [3]UG Scholar, [4]UG Scholar, [5]UG Scholar, [6]UG Scholar
[1]School of CS and IT
[1]Jain(Deemed-To-Be) University, Bengaluru, India

Abstract—This Bipolar occurrence predictive system draws on a wide range of innovative technologies that are based on the powerful applications of Deep Learning and Machine Learning in analysing medical data and making an informative prediction to assist health professionals in treating Bipolar disease as early as possible. The types of input data available are quite helpful in making accurate forecasts. In this case, the answer to the problem statement might be implemented by the (Convoluted Neural Networks), KNN (K-nearest Neighbours), and Decision tree, which could be quite effective in identifying the disease. A wide variety of diseases have symptoms that can be found in electronic health records (EHRs), making it necessary to create models that can detect these difficulties early on and accurately diagnose the disorders. Other methods for finding hidden patterns in patient data include the use of deep neural networks and machine learning techniques including multi-layer perception, SVM, random forest, and decision trees (DT). Other methods for finding hidden patterns in patient data include the use of deep neural networks and machine learning techniques including multi-layer perception, SVM, random forest, and decision trees (DT). The study also looks at the symptoms linked with different forms of psychotic disorders and handles class imbalance from a multi-label classification standpoint. Models were assessed and compared based on their accuracy using this metric. Class imbalance accuracy was 85.17 percent better for deep neural networks compared to MLP models (58.44 percent better for MLP models). In terms of machine learning methodologies, the random forest model provided the best outcomes, with 64.1 percent and 55.87 percent, respectively, when compared to deep learning. Analysis can be improved by using a {Decision Tree Classifier–} Machine Learning algorithm

IndexTerms*: *bipolar disorder, deep learning, psychotic disorders, forecasting

I.INTRODUCTION

A lack of medical data quality reduces the accuracy of an analysis. The ability to predict disease outbreaks may be hampered by the fact that certain regional diseases have distinct characteristics that vary from region to region. Structured data, on the other hand, is what most of these existing efforts are. Semi-structured and unstructured work has no set of guidelines to follow. Structured and unstructured data will be considered by the suggested system. To improve the accuracy of the analysis, machine learning algorithms are used. CNN (Convolutional Neural Networks), KNN (K-nearest Neighbours), and Decision Tree Classifiers are built using the doctor's input to classify the ailment. To develop a web-based platform for predicting sickness.

II.OBJECTIVE

In the recent age, ML and Deep Learning (DL) have provided remedies to nearly all of these flaws. ML is a computer technique that aids in the automatic learning of a machine through the use of training datasets. DL is a subset of machine learning that uses methods based on artificial neural networks to learn features on their own based on main variation

factors in the dataset. This paper provides an overview of various bipolar classification algorithms and compares their performance to standard available datasets. In this paper, the employment of the Machine Learning (ML) classification technique, consists of several stages: (a)Acquire input data. (b)Extract features from the dataset. (c)Apply the machine learning technique used to train the model. (d)Learn by experience automatically by feeding training datasets. (e)Finally, the data is tested using a categorization based on the training. (f)Accurately predict bipolar disorder.

III.LITERATURE REVIEW

The study 'Preventive Healthcare and Corporate Female Workforce' also indicated that long hours and working under rigorous deadlines cause up to 75% of working women to suffer from depression or general anxiety disorder, compared to women with lower levels of psychological pressure at work. According to the report [1], women working in high-demand fields including media, knowledge process outsourcing, and touring are unable to take sick days because of their employment instability, which is exacerbated during times of economic crisis. For many reasons, healthcare data can't be "mined" to uncover hidden insights for better decision-making in the business. Finding underlying patterns and relationships is sometimes overlooked. This can be remedied with the use of advanced data mining tools. Data mining techniques, such as Decision Trees, Nave Bayes, and Neural Networks, were used in this study to construct a prototype HDPS. This study's findings reveal that each technique has a distinct advantage when it comes to achieving the predetermined mining objectives. [4] HDPS is able to address what-if questions that typical decision support systems cannot. Patients' risk of heart disease can be predicted based on their medical profiles, which include things like age, gender, blood pressure, and blood sugar. Research into cardiac disease can be aided by the discovery of patterns and connections between various medical variables. Web-based, user-friendly, scalable, reliable and expandable: these are the key features of the HDPS software system. [3] A classifier is a way of translating a feature space (discrete or continuous) to a set of labels (discrete or continuous) (discrete or continuous). Learning classifiers are split into supervised and unsupervised learning classifiers [2] in this article. The Support Vector Machine (SVM) was invented in the late 1980s and early 1990s as a new sort of machine learning approach (SVM) (SVM). For pattern recognition, SVM offers a variety of exciting properties, as demonstrated in the Burges' tutorial [12]. Because of increased processing power and the development of rapid learning algorithms, it is now possible to train SVM for real-world applications. SVM has garnered a lot of attention in the machine learning field, but the multi-class SVM is still a research problem.'" "Single machine" and "divide and conquer" approaches are two of the most commonly used ways for solving multi-class SVMs, but they aren't the only ones available. A commonde composition technique is the "one versus all,". One SVM is generated for each pair of classes when using the one-against-one method. More advanced decomposition algorithms based on error-correcting output codes were given by Diettrich & Bakiri [13] and Allwein et al. [14] (ECOC). Research by Hsu and Lin compared two "one machine" SVM techniques and three "divide and conquer SVM" algorithms for a multi-class classification in their study [15]. Despite the similar results, the authors conclude that "one against one" is more practical because the training approach is quicker. To the contrary, Rifkin and Klautau [16] suggest that when the SVMs are properly tuned, the so-called oneversus-all (or ECOC) approach is just as accurate as any other ECOC method, including the "one against all" strategy. Therefore, it looks impossible to decide which multi-class SVM is the best for handwriting recognition based on the literature assessment. It was for this reason this research carried out the two most common strategies, "one versus all" and "one vs one".

V. RESULTS

This paper presents a clever idea in this venture. For a powerful sickness forecast, it applies the K-Nearest Neighbour (KNN) and Decision Tree (CNN) AI calculations. An assortment of ailment side effects was expected for the sickness forecast. The living propensities for an individual and examination data are considered for dependable illness expectations in this broad infection forecast. The Decision Tree calculation has an exactness of 84.5 percent overall sickness forecast, which is higher than the KNN approach. Following a general ailment forecast, this approach can decide the gamble related to an overall sickness, which can go from low to high. Many AI applications are utilized in the proposed framework to foster a classifier that can isolate information relying upon its highlights. The informational collection is separated into at least two classes. Clinical information investigation and infection expectation are both helped by such classifiers. AI is currently omnipresent, and it is feasible to use it consistently without acknowledging it. For order, CNN utilizes both organized and unstructured information from an emergency clinic. Other AI calculations, then again, just capacity with organized information and have a long calculation time since they store every one of the information as a preparation dataset and utilize a complex working out methodology.

VI. CONCLUSION

This paper recommends using the K-Nearest Neighbour (KNN) and Decision Tree (CNN) machine learning algorithms for disease prediction. Disease prediction necessitates a list of signs and symptoms. As part of this general disease prediction, a person's lifestyle and medical history are taken into consideration. The Decision Tree method outperforms the KNN method in general disease prediction, with an accuracy rate of 84.5%. This system may compute the risk of general disease once it has predicted its occurrence and then indicate whether or not that risk is lower or higher. Many machine learning applications are used in the proposed system to build a classifier that can segregate data depending on their features. The data collection has been divided into two or more classes. Medical data analysis and disease

prediction are both aided by such classifiers. Machine learning is now so pervasive that it is possible to utilize it numerous times a day without even realizing it. For categorization, Data from a hospital, both organized and unstructured, is used by CNN. If you're looking for a machine learning method that doesn't rely on structured data and doesn't require a long computation time, then look no further than TensorFlow.

VII. ACKNOWLEDGE

I would like to acknowledge and give my warmest thanks to my guide Prof. Jayashree M kudari who made this work possible. Her guidance and advice carried me through all the stages of writing my research work. I would also like to thank my committee members for letting my defence be an enjoyable moment, and for your brilliant comments and suggestions, thanks to you.

REFERENCES

[1] Sharma, M. and Majumdar, P.K., 2009. Occupational lifestyle diseases: An emerging issue. Indian Journal of Occupational and Environmental medicine, 13(3), pp. 109– 112.

[2] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide. SPSS inc, 9, 13.

[3] Pattekari, S.A. and Parveen, A., 2012. Prediction system for heart disease using Naïve Bayes. International Journal of Advanced Computer and Mathematical Sciences, 3(3), pp. 290–294.

[4] Kanchan, B.D. and Kishor, M.M., 2016. Study of machine learning algorithms for special disease prediction using principal of component analysis. In Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), 2016 International Conference on (pp. 5–10). IEEE.

[5] Kazeminejad, A., Golbabaei, S. and Soltanian-Zadeh, H., 2017. Graph theoretical metrics and machine learning for diagnosis of Parkinson's disease using rs-fMRI. In Artificial Intelligence and Signal Processing Conference (AISP), (pp. 134–139). IEEE.

[6] Milgram, J., Cheriet, M. and Sabourin, R., 2006. "One against one" or "one against all": Which one is better for handwriting recognition with SVMs? Tenth International Workshop on Frontiers in Handwriting Recognition, La Baule (France), Suvisoft, 2006.

[7] Hossain, R., Mahmud, S.H., Hossin, M.A., Noori, S.R.H. and Jahan, H., 2018. PRMT: Predicting Risk Factor of Obesity among Middle-aged People Using Data Mining Techniques. Procedia Computer Science, 132, pp. 1068– 1076.
[8] Sayali Ambekar and Dr. Rashmi Phalnikar, 2018. Disease prediction by using machine learning, International Journal of Computer Engineering and Applications, vol. 12, pp. 1–6.

[9] Mishra, A.K., Keserwani, P.K., Samaddar, S.G., Lamichaney, H.B. and Mishra, A.K., 2018. A decision support system in healthcare prediction. In Advanced Computational and Communication Paradigms (pp. 156– 167). Springer, Singapore.

[10] V. Vasu and M. Indiramma, "A Survey on Bipolar Disorder Classification Methodologies using Machine Learning," 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 335- 340, doi: 10.1109/ICOSEC49089.2020.9215334.

[11] Mateo-Sotos, J., Torres, A. M., Santos, J. L., Quevedo, O., & Basar, C. (2022). A Machine Learning-Based Method to Identify Bipolar Disorder Patients. Circuits, Systems, and Signal Processing, 41(4), 2244-2265.

[12] Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and computing, 14(3), 199-222.

[13] Dietterich, T. G., & Bakiri, G. (1994). Solving multiclass learning problems via error-correcting output codes. Journal of artificial intelligence research, 2, 263-286.

[14] Allwein, E. L., Schapire, R. E., & Singer, Y. (2000). Reducing multiclass to binary: A unifying approach for margin classifiers. Journal of machine learning research, 1(Dec), 113-141.

[15] Lin, W. M., Wu, C. H., Lin, C. H., & Cheng, F. S. (2008). Detection and classification of multiple powerquality disturbances with wavelet multiclass SVM. IEEE transactions on power delivery, 23(4), 2575-2582.

[16] Rifkin, R., & Klautau, A. (2004). In defense of one-vsall classification. The Journal of Machine Learning Research, 5, 101-141.

VIII. REFERENCE LINKS

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC286244 1/