
Prediction of Neurological Disorders among Children Using Machine Learning Techniques

¹G. Reshma, ²Dr. P.V.S Lakshmi

¹Asst. Professor, Department of Information Technology, PVPSIT, Kanuru, Vijayawada

²Professor, Department of Information Technology, PVPSIT, Kanuru, Vijayawada

Email: reshma.gorripati@gmail.com

Received: 15th October 2018, Accepted: 29th January 2019, Published: 28th February 2019

Abstract

Diagnosis of neurological problems in children at the earliest helps the medical associates to improve the children's health conditions. Hence, there is an important need to diagnose neurological disorders that occur which may lead to critical problems if care is not taken in advance. Machine learning Techniques helps for analyzing medical data and the problem can be diagnosed in an effective way. This research has identified machine learning techniques on different measures of accuracy in diagnosing three basic neurological disorders. A Neurological data set is collected from Neuro clinic centre to evaluate the performance of Machine Learning Techniques. Among all the attributes collected some attributes were identified as vital for diagnosing the problem. It is evident from the results that the chosen ML techniques produced more accurate results and there is only a slight difference between their performances.

Keywords

Neurological Disorders, Machine Learning, Predictive Analytics, Feature Subset Selection

Introduction

Neurological disorders among children rather than adults are increasing around the world and it is estimated that one in five people in the world will be affected by neurological disorders. These are accompanied by a range of symptoms that can be structural and functional like pain, confusion, muscle weakness, unconsciousness, Depressive disorders etc. Neurological disorders diagnosis involves many steps because disorders like Epilepsy, Dementia, and Schizophrenia etc are the most compelling and unable to elucidate the mechanism of the path physiology. Actually, the diagnosis involves analyzing the symptoms, reviewing past medical history and conducting physical tests, especially in case of Neurological disorders, the test report of EEG is recorded to estimate normal disorder or abnormal disorder. Various psychological tests namely clinic neuropsychiatric observation, audio assessment, assessment of Intellectual coefficient are also conducted which are caused by the neurological disorders.

In some of the cases, the cooperation of the patient is very much essential in Diagnosing children with these types of disorders. Machine learning (ML) is the learning behaviour done by machines. Researchers have developed numerous techniques that help machines to accept the behaviour of human and tried to make inferences. Methods have also been found in machine learning techniques to deal with uncertain problems. This research has identified three machine learning techniques that are applied on neurological disorders over collected data set. A comparison is made on applied machine-learning techniques and identified the most accurate one which can be utilized in diagnosing neurological problems. Five basic neurological disorders namely, Dementia, Epilepsy, Neurological disorders associated with mal nutrition, Attention Deficit Hyperactivity Disorder (ADHD) are considered. The symptoms that are observed are given as input to the learning techniques and the disorder can be diagnosed based on the symptoms, finally the disorder based on the symptoms is retrieved as their output.

Dementia symptoms include confusion, impairments in thought, Depression, Apathy which occurs commonly in dementia. To know the exact symptoms medical professionals may conduct blood tests, memory tests, brain imaging tests and a neurological exam for this type of disorder. **Epilepsy** is a neurological chronic disorder that causes **seizures**. The seizures can be simple and complex. Simple seizures having symptoms like dizziness, tingling and twitching of limbs where as complex seizures involve unresponsiveness, performing repetitive movements. Another type of seizure namely Tonic-Clonic seizures symptoms include shaking, biting of the tongue, loss of consciousness. It is a common disorder in younger children and occurred more in males compared to females. **Neurological disorders associated with mal nutrition** symptoms include iodine deficiency disorders Delayed mental development in children, Depression, by observing Wernicke's encephalopathy coefficient test, deficiency in nutrition can be identified. Under nutrition can be assessed by the measurement of body weight and body height together with age and sex. Then it will be possible to evaluate the energy of the individual. **Attention Deficit Hyperactivity Disorder (ADHD) symptoms include Inattention and hyperactivity**. These symptoms must be present before age 12. In most of the cases. In majority of the cases, ADHD can occur based on genetic behaviour. Also certain components of food nutrients like food

additives, sugar and omega-3 fatty acids as these fats are important for brain development functioning can have effect on children behaviour. Brain injury may be one of the reasons for this type of disorder.

The aim of this research is to predict neurological disorders among children using machine learning techniques. Section 2 presents a literature survey on neurological disorders .Section 3 focus on overview of various machine learning techniques to be considered for prediction. Section 4 presents the method and the data sets collected in this research to predict the four neurological disorders. Section 5 evaluates the proposed machine learning techniques on data sets. Section 6 provides conclusion and future scope.

Literature Review

The research on applying machine learning techniques in neurological disorder diagnosis have started earlier Seixas, Flávio Luiz, et al. [1] proposed a Bayesian network decision model and combination of expert knowledge to diagnose neurological disorder Dementia. Gomuła, Jerzy, Krzysztof Pancerz, and Jarosław Szkoła.[2] proposed attribute extension approach to improve classification accuracy in the case of discretized data. Basavappa S.R. et al. [3] used depth first search method with backward search strategy to diagnose dementia. They developed a computer aided expert system using the patient's behavioural, cognitive, emotional symptoms and results of neuropsychological assessments. Chattopadhyay, S., Kaur, P., Rabhi, F., Acharya[4] proposed a hybrid system consisting a fuzzy controller on a feed forward neural network which helps in diagnosing the depression grades with an accuracy of 95.50%. Masri, R.Y.; Jani, H.M.,[5] proposed a diagnose expert system using Artificial Intelligence techniques using rule based reasoning and found out suitable treatment for each of the disorders based on various health conditions. Dabek, Filip and Jesus J. Canban[6] developed a neural network model and analyzed against dataset of 89,840 patients and have shown overall accuracy of 82.35%.In [7], Jawseef Ayoub Shaikh they have chosen parkinsons disease and tumour diseases to predict the probability among the patients. The performance of ANN, Naive Bayes, and Decision Trees Algorithms were measured. For those diseases Naive Bayes performs the best with an accuracy of 49.11%.

Overview of Machine Learning Techniques

There are many types of Machine Learning Techniques that can be applied for predicting the diseases based on symptoms. They are supervised learning, Unsupervised learning, Semi Supervised learning, Reinforcement learning and Deep Learning .These techniques are used to classify the data based on dataset chosen.

- 1) Supervised Learning Technique: learning from examples applied on a training set with given targets and on the basis of this training set, algorithms respond exactly to all inputs. Regression and classification are the types of Supervised Learning. Classification: It predicts the class of Yes or No, for example, "Is hyperactivity leads to ADHD?" Regression: It can be explained by giving the answer of test for example "How much the value is for a particular test" and "How many of them are correct?."
- 2) Unsupervised Learning Technique: Target variables or class labels are not available. Unsupervised learning technique find out the similarity among the given input data and relied upon on these data, this learning technique classify the data that it belongs to which type of class. Example for this technique is clustering.
- 3) Semi supervised learning Technique is a class of supervised learning techniques. This learning also used unlabeled class data for training purpose. In this type of learning both labelled data and unlabelled data is chosen but in variation. Minimum class label data and maximum unlabeled data.
- 4) Reinforcement Learning Technique: This learning is encouraged by behavioural properties. Algorithm will notify when the answer is not correct but it has no answer how to correct it?, It has to explore itself for finding various possibilities and test among them to find the right answer.
- 5) Deep Learning Technique: This type of machine learning is based on set of various algorithms. These learning algorithms applied on data which represents model high-level abstraction. It uses graph with more depth with various processing layers, which made up of so many linear and nonlinear transformations. The methodology of research is depicted in the figure 1.shown below.

Methodology

Initially identify the basic neurological disorders frequently in children .The procedure for diagnosing each disease should be drawn from medical associates. Then a model was built that uses machine learning techniques to diagnose four common mental health problems effectively. This model assists the medical associates to identify the problem if the known evidences or symptoms of the patient which are given as input. The data set for predicting neurological disorders are taken from a psychologist. The data set has 100 instances including the class label have been identified as important and required for analysis.

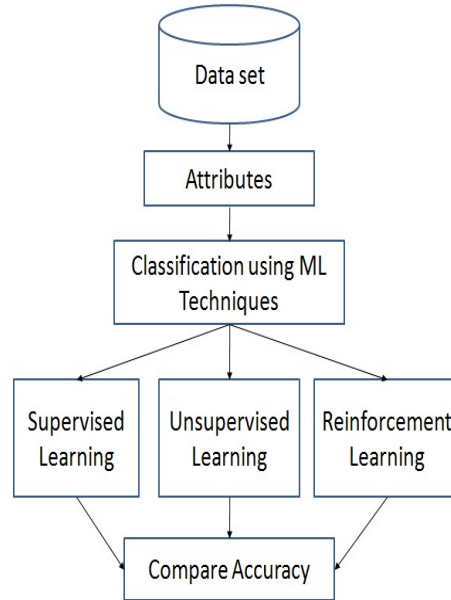


Fig. 1: Methodology

Sl.No	Attribute	Values
1	Confusion(X1)	{Y,N}
2	Impairments in thoughts(X2)	{Y,N}
3	Depression(X3)	{Y,N}
4	Apathy(X4)	{Y,N}
5	Seizures(X5)	{Y,N}
6	Tonic-Clonic (X6) Siezures	{Y,N}
7	Unresponsiveness(X7)	{A,I}
8	Repetitive movements(X8)	{A,I}
9	Shaking(X9)	{Y,N}
10	Iodine Deficiency Disorders(X10)	{Y,N}
11.	Delayed Mental development(X11)	{MEDIUM,LOW,HIGH}
12	Inattention(X12)	{Y,N}
13	Hyperactivity(X13)	{MEDIUM,LOW,HIGH}
14	Loss of Consciousness(X14)	{Y,N}
15	Dizziness(X15)	{Y,N}

Table 1: The Attributes Identified from the Data Set

The listed attributes are mentioned in the excel table from X1-X15. The dataset consisting of 100 samples and each attribute has values consisting of yes representing 1 and no represents 0. For attributes which show information like adequate(A) and inadequate(I) values are 1 for A and 0 for I. Some of the attributes are having three distinct values namely medium, low, high which can be represented as 1, 2, 3 possible values for low, medium, high respectively. The snapshot of the dataset is shown below in figure 2

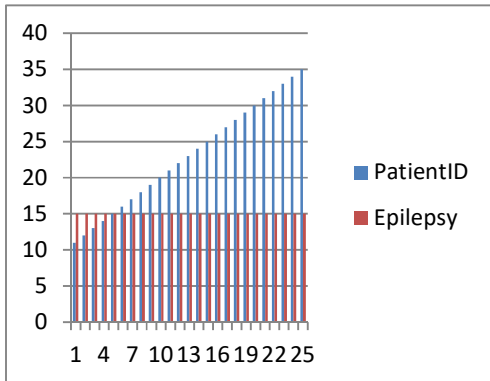
X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15	
0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	1	3	3	1	1	0	1	3	1	3	0	1	1	0	
0	1	3	3	1	1	0	1	3	1	3	0	1	1	0	
0	1	1	1	1	1	0	1	1	1	1	0	1	1	0	
1	1	2	2	1	1	1	1	2	1	2	1	1	1	1	
0	1	1	1	1	1	0	1	1	1	1	0	1	1	0	
0	1	3	3	1	1	0	1	3	1	3	0	1	1	0	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
0	2	3	3	2	2	0	2	3	2	3	0	2	2	0	
0	2	1	1	2	2	0	2	1	2	1	0	2	2	0	
1	2	1	1	2	2	1	2	1	2	1	1	2	2	1	
1	2	1	1	2	2	1	2	1	2	1	1	2	2	1	
1	2	3	3	2	2	1	2	3	2	3	1	2	2	1	
0	2	1	1	2	2	0	2	1	2	1	0	2	2	0	
1	2	1	1	2	2	1	2	1	2	1	1	2	2	1	
0	2	3	3	2	2	0	2	3	2	3	0	2	2	0	
0	2	1	1	2	2	0	2	1	2	1	0	2	2	0	
1	2	1	1	2	2	1	2	1	2	1	1	2	2	1	
1	2	1	1	2	2	1	2	1	2	1	1	2	2	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	2	2	2	2	2	1	2	2	2	2	1	2	2	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
0	2	2	2	2	2	0	2	2	2	2	0	2	2	0	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
0	2	3	3	2	2	0	2	3	2	3	0	2	2	0	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	

Fig. 2: Data Set

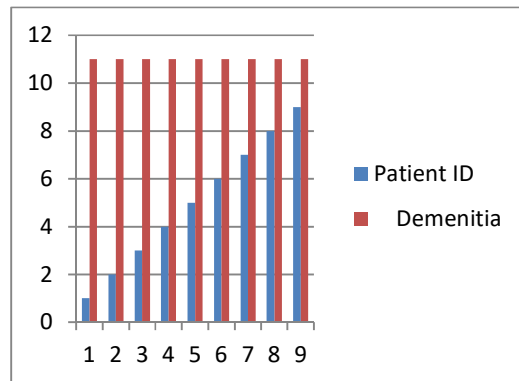
Experimental Results

The performance analysis of learning algorithms applied on dataset using R tool. Initially, the classifiers were executed on selected attributes (15). The following Figure 2 shows the distribution of patients among four neurological disorders among which patients with disease dementia occurred more frequently among all the records. The R tool provides classification measures to visualize and estimate accuracy. Accuracy level of classifiers can be measured by using Kappa Statistic, Accuracy. Kappa Statistic gave measures of prediction with the class variable. Value 1 signifies predicted value is similar to the true class and a measure of 0 signifies complete variance between prediction and the true class. Table 2 depicts that the kappa statistics of three classifiers, namely, KNN, SVM, random forest show a higher values with all attributes.

3 (a)



3 (b)



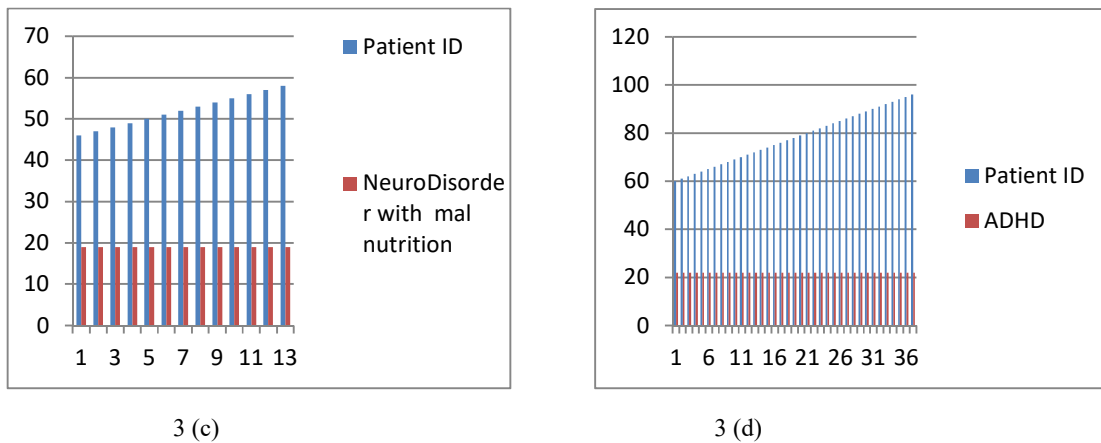


Figure 3: Distribution of Dataset with respect to Disorders

Table 2. Statistics

Accuracy				
	Min.	1st Qu.	Median	Mean
knn	0.8334	0.7167	1.0000	0.8583
svm	0.8334	0.7167	0.9168	0.8417
rf	0.8334	0.7167	0.9584	0.8500
Kappa				
	Min.	1st Qu.	Median	Mean
knn	0.752	0.8753	1.0000	0.9375
svm	0.752	0.8753	0.8750	0.9125
rf	0.753	0.8753	0.9375	0.9250

Accuracy of a classifier on a given test set is the percentage of test set instances that are correctly classified by the classifier. Obtained output which is shown in figure 3 depicts accuracy among classifiers. The accuracy of the classifier depends on how well the classifier classifies the data set being tested. So these classifiers are more accurate in predicting the neurological disorders among children.

Output:

```
100 samples
  4 predictor
  4 classes: 'Dementia', 'Epilepsy', 'ADHD', 'NDMN'
Resamplings
  Accuracy Kappa Accuracy SD Kappa SD
  0.875    0.8625 0.04025383 0.06038004
Overall Statistics
  Accuracy : 1
    90% CI : (0.8844, 1)
  No Information Rate : 0.3444
  P-Value [Acc > NIR] : 5.858e-15
```

Figure 4: Statistics Related to Accuracy

Conclusion

Prediction of diseases accurately at the beginning stage of treatment can be done by expert systems. As a part of that there are various learning techniques which helps to construct expert systems, Medical Expert systems needs to identify the best technique and to be applied for better results even in critical conditions of the patients. In this paper three machine learning techniques are compared and identified from the results that the three techniques produce more accurate results. The data set for the neurological data is not available in large and in future, the research may be applied for a large data set with more number of attributes considering values other than 0,1,2,3 because some attributes might required reports like blood tests, EEG etc. these values are of numeric type. The machine learning techniques need to be more trained for the implementation in real time prediction for all possible set of values chosen as attributes.

References

- [1] Seixas, Flávio Luiz, et al. "A Bayesian network decision model for supporting the diagnosis of dementia, Alzheimer's disease and mild cognitive impairment." *Computers in biology and medicine* 51 (2014): 140-158.
- [2] Gomuła, Jerzy, Krzysztof Pancierz, and Jarosław Szkoła. "Classification of MMPI profiles of patients with mental disorders—experiments with attribute reduction and extension." *Rough Set and Knowledge Technology*. Springer Berlin Heidelberg, 2010. 411-418.
- [3] Basavappa SR; Rao SL; Harish B; Shantala R Basavappa ; Shobini L Rao Department of Electronics & Communication, Manipal Institute of Technology, Manipal, India. Expert system for dementia / depression diagnosis. *NIMHANS Journal*. 1996 Apr; 14(2): 99-106
- [4] Chattopadhyay, S., Kaur, P., Rabhi, F., Acharya, U.R., 2011. An automated system to diagnose the severity of adult depression. In: Jana, D., Pal, P. (Ed.), *Proceedings of Second International Conference on Emerging Applications of Information Technology (CSI EAIT-2011)*, IEEE Computer Society and Conference Publishing Services, Kolkata, India, pp. 121–124.
- [5] Masri, R.Y.; Jani, H.M., "Employing artificial intelligence techniques in Mental Health Diagnostic Expert System," in *Computer & Information Science (ICCIS), 2012 International Conference on* , vol.1, no., pp.495-499,12-14 June 2012 doi:10.1109/ICCISci.2012.6297296.
- [6] Dabek, Filip, and Jesus J. Caban. "A Neural Network Based Model for Predicting Psychological Conditions." *Brain Informatics and Health*. Springer International Publishing, 2015. 252-261.
- [7] Tawseef Ayoub Shaikh., "A Prototype of Parkinson's and primary tumor disease prediction using data mining techniques", *International Journal of Engineering Science Invention*, vol 3, Issue: 4, April 2014.