Prediction of Mental Health Problems Among Children Using Machine Learning Techniques

Ms. Sumathi M.R., Research Scholar,
Dept. of Computer Science,
Bharathiar University,
Coimbatore, India

Dr. B. Poorna, Principal,
S.S.S. Jain College,
T.Nagar,
Chennai, India

Abstract—**Early diagnosis of mental health problems helps the professionals to treat it at an earlier stage and improves the patients' quality of life. So, there is an urgent need to treat basic mental health problems that prevail among children which may lead to complicated problems, if not treated at an early stage. Machine learning Techniques are currently well suited for analyzing medical data and diagnosing the problem. This research has identified eight machine learning techniques and has compared their performances on different measures of accuracy in diagnosing five basic mental health problems. A data set consisting of sixty cases is collected for training and testing the performance of the techniques. Twenty-five attributes have been identified as important for diagnosing the problem from the documents. The attributes have been reduced by applying Feature Selection algorithms over the full attribute data set. The accuracy over the full attribute set and selected attribute set on various machine learning techniques have been compared. It is evident from the results that the three classifiers viz., Multilayer Perceptron, Multiclass Classifier and LAD Tree produced more accurate results and there is only a slight difference between their performances over full attribute set and selected attribute set.**

Keywords—Mental Health Diagnosis; Machine Learning; Prediction; Feature Selection; Basic Mental Health Problems

I. INTRODUCTION

Mental illness is rising at epidemic rates around the world and WHO predicted that one in four people in the world will be affected by mental and neurological disorders at some point in their lives. Depressive disorders will become the second leading cause of the global disease burden by 2020, behind ischaemic heart disease but ahead of all other diseases [1]. The growth in number of professionals who treat mental illness is very less when compared with the growth in number of people suffering from mental illness.

Mental health diagnosis involves many steps and it is not a straightforward process. Actually, the diagnosis starts with specially designed interview with questions about the symptoms and medical history and sometimes performing a physical examination. Various psychological tests are also conducted to make sure that the symptoms are caused only by the mental health and not by any other problems. Various assessment tools are available to evaluate a person for a mental illness.

As similar factors and symptoms lead to different mental health problems, the diagnoses have become a complicated task and sometimes they may be misdiagnosed. The co-operation of the patient is very much needed in diagnosing the problem. Diagnosing children with mental health problems is more problematic than diagnosing adults. Hence, care must be taken to diagnose the mental health problem accurately.

Artificial Intelligence (AI) is the intelligence exhibited by machines. It is a field of study on creation of computer software that is capable of intelligent behavior. AI researchers have developed many machine learning techniques that help computers to imitate the human reasoning to solve puzzles or make logical deductions. Methods have also been found to deal with uncertain or incomplete information, employing concepts from probability and other fields.

This research has identified eight machine learning techniques that diagnosed mental health problems correctly over a sample data set of ten cases. A comparison is made on those eight machine-learning techniques and identified the best three which can be utilized to assist mental health professionals in diagnosing mental health problems. Five basic mental health problems of children viz., Attention problem, Academic Problem, Anxiety Problem, Attention Deficit Hyperactivity Disorder (ADHD) and Pervasive Developmental Disorder (PDD) have been considered. The factors, symptoms and various test results that are observed by the professionals are given as input to the techniques and the psychological problem diagnosed is retrieved as their output.

Attention problems in children refer to failing to pay close attention to details or making careless mistakes, not listening to when spoken directly, do not sustain attention or struggling with basic reading and math skills and so on. **Academic problems** can be very common and range from difficulties with studying and reading to anxieties about exams than practical or technical skills. **Anxiety Problems** are characterized by chronic excessive worry accompanied by restlessness, fatigue, concentration problems, irritability, muscle tension and sleep disturbance **Attention Deficit Hyperactivity Disorder (ADHD)** is one of the most common childhood disorders which may continue through adolescence and adulthood. Some of the symptoms include difficulty staying focused and paying attention, difficulty controlling behavior and hyperactivity. **Pervasive Developmental Disorder (PDD)** refers to a group of conditions that involve delays in the development of many basic skills, like, ability to socialize with others, ability to communicate and ability to use imagination.

The aim of this research is to predict basic mental health problems using machine learning techniques. Section 2 presents a literature review on mental health diagnosis using computers. Section 3 gives an overview of various machine learning techniques. Section 4 presents the methodology and the data sets used in this research to predict the five basic mental health problems. Section 5 evaluates the techniques using training and test data sets. Section 6 provides conclusion and future work.

II. Literature Review

The research on applying machine learning techniques in mental health diagnosis have started in nineteen eighties. DTREE [2] is an expert system that diagnoses DSM-IV Axis I disorders using 'Decision Tree' techniques. INTERNIST/AUTOSCID [3] is a computerized Structured Clinical interview for DSM-IV Axis II personality disorders. Yap R.H. and Clarke D.M. [4] developed an expert system called MILP (Monash Interview for Liaison Psychiatry) using constraint-based reasoning for systematic diagnoses of mental disorders based on DSM-III-R, DSM-IV and ICD-10. Constraint Logic Programming (CLP) language was used to develop the system. Kipli et al. [5] introduced new approaches for Data-mining and classification of mental disorder using Brain Imaging data.

Masri R.Y. and Jani H.M., in [7] proposed a Mental health Diagnostic Expert System to assist the psychologists in diagnosing and treating their mental patients. Three artificial intelligence techniques viz., Rule-Based Reasoning, Fuzzy Logic and Fuzzy-Genetic Algorithm were used for diagnosis and suggestion of treatment plans. Luxton, David D. [8] analyzed how artificial intelligence has been utilized in psychological practice. The current trend and future applications and their implications have been discussed.

Razzouk D., et al [9] developed a decision support system for the diagnosis of schizophrenia disorders with an accuracy of 66-82%. Chattopadhyay S., et al. [10] developed a neuro-fuzzy approach to grade adult depression. The supervised (Back Propagation Neural Network: BPNN and Adaptive Network-based Fuzzy Inference System : ANFIS) and unsupervised (Self-Organizing Map) neural net learning approaches have been utilized and then compared. It was observed that ANFIS being a hybrid system performed much better than BPNN classifier. L.C.Nunes et al. [11] introduced a hybrid method to handle diagnosis of Schizophrenia. They have integrated structured methodologies in decision support (Multi-Criteria Decision Analysis: MCDA) and structured representation of knowledge into production rules and probabilities (Artificial Intelligence: AI).

Basavappa S.R. et al. [12] used depth first search method with backward search strategy to diagnose depression or dementia. They developed an expert system using the patient's behavioral, cognitive, emotional symptoms and results of neuropsychological assessments. Rahman, Rashedur M. and Farhana Afroz [13] compared various classification techniques (Bayesian Network, Multilayer Perceptron, Decision Trees, Single Conjunctive Rule Learning and Fuzzy Inference Systems and Neuro-Fuzzy Inference System) using different data mining tools like WEKA, TANAGRA and MATLAB for diagnosing diabetes. It was observed that different techniques yielded different accuracy levels with different accuracy measures like Kappa Statistic and Error rates.

David Gil A. and Magnus jhonson B. in [14], proposed a system based on Artificial Neural Networks(ANN) and Support Vector Machines(SVM) that diagnoses Parkinson's Disease. The system has shown an increase in accuracy and a decrease in cost. In [15], Gomuła, Jerzy et al., tried to find efficient methods for classification of MMPI profiles of patients with mental disorder. They identified that Attribute Extension Approach improves classification accuracy in the case of discretized data. Flavio Luiz Seixas and Bianca Zadrozny in [16] proposed a Bayesian Network (BN) Decision Model for diagnosis of dementia, Alzheimer's disease and Mild Cognitive Impairment. The BN model was considered as it is well suited for representing uncertainty and causality. Network parameters were estimated using a supervised learning algorithm from a dataset of real clinical cases. Model was evaluated using Quantitative methods and Sensitivity Analysis and it showed better results when compared to most of the other well-known classifiers. Anchana Khemphila and Veera Boonjing in [17] used Multi-Layer Perceptron (MLP) with Back Propagation Learning to diagnose Parkinson's disease effectively with reduced number of attributes. Information Gain of all attributes is used as a measure to reduce the number of attributes. Pirooznia, Mehdi, et al., in [18] utilized data mining approaches for Genome-wide Association of Mood Disorders. Six classifiers namely Bayesian Network, Support Vector Machine, Logistic Regression, Random Forest, Radial Basis Function and Polygenic Scoring Approach were compared. It was identified that simple polygenic score classifier performed better than others and it was also found that all the classifiers performed worse with small number of Single Nucleotide Polymorphisms in the brain expressed set compared to whole genome set.

In [5], Kipli, Kuryati, Abbas Z. Kouzani, and Matthew Joordens detected depression from structural MRI scans to diagnose the mental health of patients. They investigated performances of four Feature Selection algorithms, namely, OneR, SVM, Information Gain (IG) and ReliefF. Finally, they concluded that the SVM Evaluator in combination with Expectation Maximization (EM) classifier and the IG evaluator in combination with Random Tree Classifier have achieved the highest accuracy. It had also been found that the small simple sizes limit the ability to draw firm conclusions.

Dabek, Filip, and Jesus J. Caban [19] developed a Neural Network (NN) Model with an accuracy of 82.35% for predicting the likelihood of developing psychological conditions such as anxiety, behavioural disorders, depression and post-traumatic stress disorders. In [20], Tawseef Ayoub Shaikh, compared the performance of Artificial Neural Networks, Decision Tree and Naïve Bayes in predicting Parkinson's disease and Primary Tumour Disease and found that the accuracy is high in ANN for predicting Parkinson's disease and Naïve Bayes for Primary Tumour Disease. Accuracy of Decision Tree and Naïve Bayes have further improved after reducing the size of feature set by applying Genetic Algorithm to the actual data set.

The literature review shows that, on one side, a number of research works are going on in computerizing the diagnosis of mental health disorders. On the other side, efforts are taken to diagnose the mental health problems using different machine learning techniques in an efficient way. Various combinations of machine learning techniques (Hybrid techniques) are being employed to improve the accuracy of diagnosis with reduced set of features from profiles of patients. This research is to analyse selected machine learning techniques in predicting the possibility of primary mental health problems like Attention Problems, Anxiety Problems, Academic Problems, ADHD and PDD. Early diagnosis of these types of problems among children allows early treatment and improves their quality of life. Comparison has also been made with reduced number of attributes.

III. OVERVIEW OF MACHINE LEARNING TECHNIQUES

As of now, various machine learning techniques are available and still researches are going on to techniques that produce optimal results. According to Wolpert et al. [6], there is no single learning algorithm that universally performs best across all domains. Hence, in this research, eight techniques are selected as they produced accurate results for a small dataset. The techniques selected are AODEsr, Multi Layer Perceptron (MLP), RBF Network, IB1, KStar, Multi-Class Classifier (MCC), FT, LADTree.

A. AODEsr

Averaged One-Dependence Estimator (AODE) is a probabilistic classification learning technique. It is advantageous over Naïve Bayes Classifier as it addresses the problem of attribute-independence at the cost of increased amount of computation. It supports incremental learning and handles situations where some data are missing.

B. Multilayer Perceptron

Multilayer Perceptron is a feed forward artificial neural network model that maps set of input data onto a set of appropriate outputs. It is a supervised learning technique that uses back propagation for training the network and it is used to distinguish data that are not linearly separable.

C. RBF Network

Radial Basis Function Network is a feed forward artificial neural network that uses radial basis functions as activation functions for classification. It is a supervised learning technique that trains much faster than back propagation networks.

D. IB1

IB1, an instance based classifier uses a simple distance measure to find the training instance closest to the given test instance and predicts the same class as this training instance. The training is very fast and suits to problems in which the target function is very complex, but can still be described by a collection of less complex local approximations.

E. K* or KStar

K*, an instance based classifier, is similar to K-Nearest Neighbor. It uses entropy as a distance measure for classification and handles real valued attributes and missing values.

F. Multiclass classifier

This algorithm classifies instances into one of the more than two classes i.e. each training instance belongs to one of N different classes.

G. FT

Functional Trees are classification trees that could have logistic regression functions at the inner nodes and/or leaves. It deals with binary and multi-class target variables, numeric and nominal attributes and missing values.

H. LAD Tree

Logical Analysis of Data tree is the classifier that generates multiclass alternating decision tree using logistics strategy. It combines ideas and concepts from optimization, combinatorics and Boolean functions. The patterns or rules play a vital role in classification. It has been successfully applied to data analysis problems in different domains, including biology and medicine.

IV. METHODOLOGY

As the first step for the research, the problem of diagnosing basic mental health was identified and an Interview was held with a clinical psychologist to identify the mental health problems that occur more often among children. Then, observation was made on how the diagnoses were performed by the professionals. A model was built that uses machine learning techniques to diagnose five common mental health problems effectively. This model assists the professionals to identify the problem if the known evidences of the patient are given as input.

Figure 1 shows the methodology of the research. The data set for predicting mental health problem is taken from a clinical psychologist. The data set has 60 instances in text document format. From the documents, 25 attributes including the class label have been identified manually and checked with the psychologist.
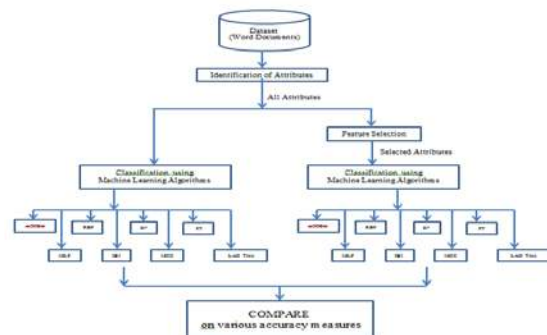

Fig. 1.   Methodology of the Research

The attributes identified from the data set are listed in table 1. All the attributes are of nominal type. As only a few attributes are relevant for classification and prediction of the problem, data set is pre-processed by eliminating irrelevant and redundant attributes using Best First Search technique.

TABLE I.   ATTRIBUTES IDENTIFIED FROM THE DATA SET. ATTRIBUTES SELECTED WITH FEATURE SELECTION ALGORITHMS ARE SHOWN IN BOLD

No.	Attribute	Meaning	Values	No.	Attribute	Meaning	Values
1	**Age**	Age group of the child, Infant / Early Childhood / Middle Childhood / Adolescent	{I,E,M,A}	14	**Attention Sustained**	Whether the child is able to direct and focus cognitive activity on specific stimuli	{Y,N}
2	**Family History**	Presence / Absence of psychological disorder in the family	{Y,N}	15	**CBCL Score**	A checklist filled by teacher/parent/self-report to identify problem behavior in children	{BC, EC, AC}
3	**Pregnancy Complication**	Presence / Absence of Complication during pregnancy	{Y,N}	16	**IQ Test Score**	A standardized test to assess the intelligence level of the child. The level may be Below Average, Average, Above Average	{BA,A,AA}
4	**Delayed Speech**	Presence / Absence of delay in development of speech	{Y,N}	17	ADHD Positive	Screening test score for Attention-Deficit/Hyperactivity Disorder	{Y,N}
5	Under Medication	Whether the child is under any type of medicinal control	{Y,N}	18	ODD Positive	Screening test score for Oppositional Defiant Disorder to measure argumentative and defiant behavior of the child	{Y,N}
6	**Academic Performance**	Whether the performance of the child is adequate/ inadequate in academics	{A,I}	19	Manic Episode Test Score	Screening test score for manic episode to check extreme elation or irritability in the mood of the child	{Y,N}
7	**Relationship Formation**	Whether the child has adequate ablity to socialize with peers, relatives and teachers	{A,I}	20	General Anxiety Disorder	Screening test score to measure the anxious level of the child	{Y,N}
8	**Behavioral Problem**	Whether the child has any problem in behavior, Ex. Cheating, telling lies	{Y,N}	21	Major Depressive Episode	Screening test score to measure the major depression level of the child	{Y,N}
9	**Concentration**	Whether the child has adequate ability to focus his attention on a particular object or activity	{A,I}	22	CDI Score	Screening test score to measure the depressive symptoms of the child	{AC,EC,BC}
10	Restless	Whether the child is able to relax or involving in constant activity	{Y,N}	23	PDD Score	Screening test score to measure the level of pervasive developmental disorder of the child	{MODERATE,NO,MILD}
11	Seizures	Whether the child has a sudden surge of electrical activity in the brain which may cause unconsciousness, rigidness, muscle stiffness, uncontrollable movements, etc.	{Y,N}	24	Autism Score	Screening test score to identify whether the child has difficulty in communicating and forming relationships with other people and in using language and abstract concepts	{MODERATE,NO,MILD}
12	Learning Difficulty	Whether the child has any difficulty in acquiring age appropriate knowledge and skills	{Y,N}	25	**Problem**	The class attribute specifying the mental health problem of the child.	{ATT_ACA, ANX_SYM, ANX_DEP_SYMP, PDD, DEV_DELAY, AUT_SYMP, ACA, ATT_BEH, ATT_EMO}
13	**Attention Aroused**	Whether the child is psychologically alert, awake and attentive	{Y,N}				

V.   EXPERIMENTAL RESULTS

In this section, the performance analysis of eight classification algorithms is carried out with a common dataset using WEKA tool. First, the classifiers were executed by including all the attributes (25) identified from the text documents and then they were executed by including only the attributes (13) selected by the feature selection algorithms. The WEKA tool provides various measures to understand the classification. Among the various measures, three measures are very important to compare the accuracy level of classifiers. They are Kappa Statistic, Accuracy and ROC Area.

Kappa Statistic measures the agreement of prediction with the true class. A measure of 1 signifies complete agreement and a measure of 0 signifies complete disagreement between prediction and the true class. Figure 2 depicts that the kappa statistics of three classifiers, namely, Multilayer Perceptron, Multilayer Classifier and LAD Tree are higher than other classifiers and selected attributes show a higher kappa value than classifier with all attributes.
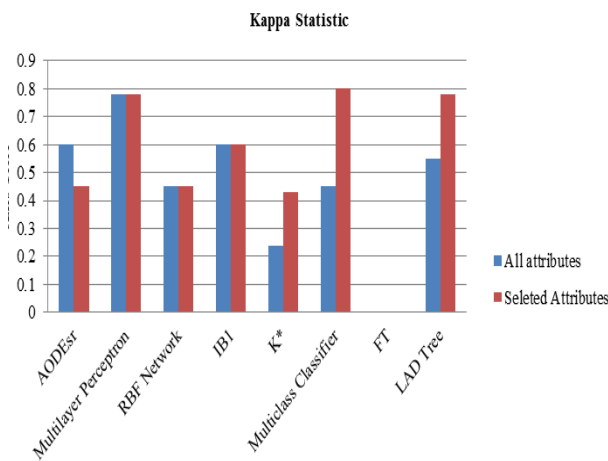

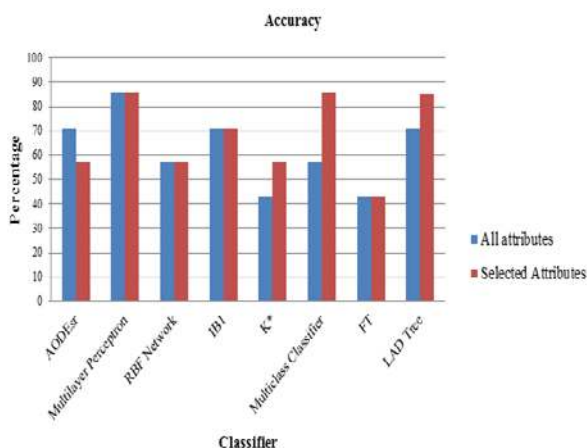Fig. 2.   Kappa Statistic for all classifiers

Fig. 3.   Accuracy measures for all classifiers

Accuracy of a classifier on a given test set is the percentage of test set instances that are correctly classified by the classifier. Fifure 3 depicts that the three classifiers, namely, Multilayer Perceptron, Multilayer Classifier and LAD Tree are more accurate than other classifiers and the selected attributes show a higher accuracy value than classifier with all attributes.

The accuracy of the classifier depends on how well the classifier classifies the data set being tested. It is measured by the area under the Receiver Operating Curve or ROC Area. An area of 1 represents a perfect test and an area of .5 represents a worthless test. The Figure 4 shows the graph on ROC Area values of the eight classifiers. The ROC Area of four classifiers, namely, Multilayer Perceptron, K*, Multiclass Classifier and LAD Tree are between 0.8 and 0.9 and so these classifiers are more accurate than other classifiers in predicting the basic mental health problems among children.
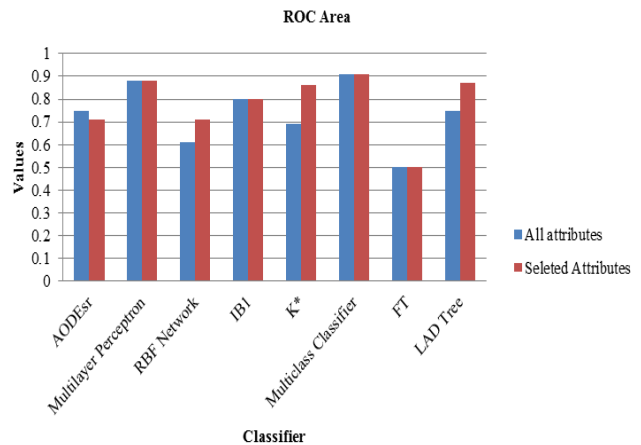

Fig. 4.   ROC Area values for all classifiers

VI.   CONCLUSION

Nowadays, a number of expert systems are utilized in medical domain to predict diseases accurately at an early stage so that treatment can be made effectively and efficiently. Also, expert systems are developed in the mental health domain to predict the mental health problem at an earlier stage. As a number of machine learning techniques are available to construct expert systems, it is necessary to compare them and identify the best that suits the domain of interest. The research has compared eight machine learning techniques (classifiers) on classifying the dataset to different mental health problems. It is evident from the results that the three classifiers viz., Multilayer Perceptron, Multiclass Classifier and LAD Tree produce more accurate results than the others. The data set is very minimal and in future, the research may be applied for a large data set to obtain more accuracy. The classifiers need to be trained prior to the implementation of any technique in real prediction.

REFERENCES

[1]   http://blogs.fortishealthcare.com/mental-health-india-wake-up-call/

[2]   Buchanan BG, Shortliffe EH. Rule-based Expert Systems: The MYCIN Experiments of the Stanford   Heuristic Programming Project. Addison-Wesley, Reading, MA, 1984.

[3]   Miller AM, Pople HE, Myers JD. "INTERNIST-I, an experimental computer-based diagnostic consultant for general internal medicine", New England Journal of Medicine, 1982:307:468-476.

[4]   Yap, R. H., & Clarke, D. M. (1996). An expert system for psychiatric diagnosis using the DSM-III-R, DSM-IV and ICD-10 classifications. Proceedings of the AMIA Annual Fall Symposium, 229–233.

[5] Kipli, Kuryati, Abbas Z. Kouzani, and Isredza Rahmi A. Hamid. "Investigating machine learning techniques for detection of depression using structural MRI volumetric features." International journal of bioscience, biochemistry and bioinformatics 3.5 (2013): 444-448.

[6] Wolpert DH, Macready WG, David H, William G. No Free Lunch Theorems for Search. Technical Report SFI-TR-95-02-010 (Santa Fe Institute)1995[Reflist]

[7] 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

[8] Luxton, David D. "Artificial intelligence in psychological practice: Current and future applications and implications." *Professional Psychology: Research and Practice* 45.5 (2014): 332.

[9] Razzouk, D., Mari, J.J., Shirakawa, I., Wainer, J., & Sigulem, D.. (2006). Decision support system for the diagnosis of schizophrenia disorders. *Brazilian Journal of Medical and Biological Research*, *39*(1), 119-128.

[10] 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.

[11] Luciano Comin Nunes, Plácido Rogério Pinheiro, Tarcísio Pequeno Cavalcante, and Mirian Calíope Dantas Pinheiro, "Handling Diagnosis of Schizophrenia by a Hybrid Method," Computational and Mathematical Methods in Medicine, vol. 2015, Article ID 987298, 13 pages, 2015. doi:10.1155/2015/987298

[12] 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

[13] Rahman, Rashedur M., and Farhana Afroz. "Comparison of various classification techniques using different data mining tools for diabetes diagnosis." *Journal of Software Engineering and Applications* 6.03 (2013): 85.

[14] Gil, David, and M. A. G. N. U. S. Johnson. "Diagnosing parkinson by using artificial neural networks and support vector machines." Global Journal of Computer Science and Technology 9.4 (2009): 63-71.

[15] Gomuła, Jerzy, Krzysztof Pancerz, 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.

[16] 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.

[17] Khemphila, Anchana, and Veera Boonjing. "Parkinsons disease classification using neural network and feature selection." *World Academy of Science, Engineering and Technology* 64 (2012): 15-18.

[18] Pirooznia, Mehdi, et al. "Data mining approaches for genome-wide association of mood disorders." *Psychiatric genetics* 22.2 (2012): 55.

[19] 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.

[20] 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