THESIS

On

An Integrated System for Diagnosis of EEG based Diseases

Submitted to

LOVELY PROFESSIONAL UNIVERSITY

In partial fulfillment of the requirements for the award of degree of

MASTER OF PHILOSOPHY (M. Phil)

IN

Computer science

Submitted by:

Deepika Kundra

Regd.no.-11312173

Supervised by:

Dr. Babita Pandey

SCHOOL OF COMPUTER APPLICATION

LOVELY PROFESSIONAL UNIVERSITY

PUNJAB

Declaration

I hereby declare that the dissertation entitled, **“An Integrated System for Diagnosis of EEG based Diseases”** submitted for the M.Phil (c.sc) Degree is entirely my original work and all ideas and references have been duly acknowledged. It does not contain any work for the award of any other degree or diploma.

Date:______________

Deepika Kundra

Regn. No: 11312173

Certificate

This is to certify that **Deepika Kundra** has completed M.Phil dissertation titled, **"An Integrated System for Diagnosis of EEG based Diseases"** under my guidance and supervision. To the best of my knowledge, the present work is the result of her original investigation and study. No part of the dissertation has ever been submitted for any other degree or diploma.

The dissertation is fit for the submission and the partial fulfilment of the conditions for the award of M.Phil Computer Science.

Date: -------

Signature:

Dr. Babita Pandey

Abstract

Medical Diagnosis System (MDS) provides facility to clinical experts for diagnosis of the different diseases. This work focus on the development of the MDS for the diagnosis of EEG based diseases, integrating Rule based reasoning (RBR) - Case based reasoning (CBR), J48 (data mining) - Case Based Reasoning (CBR) – Artificial neural network (ANN). These integrated systems reduce the error amount and degree of uncertainty. Brain is the bioelectric generator. With neurological disordering there occur some disturbances in the brain that lead some problems like muscle weakness, brain functioning. EEG (electroencephalography) is a medical imaging techniques test that measures the electric activity of the human brain. EEG signals related to the level of consciousness of the person, and contain very useful information relating to different physiological state of brain.

Acknowledgement

I would like to express my sincere gratitude to my main supervisor, Dr. Babita Pandey, for her valuable guidance and helping me with her intuitive ideas. Without her help I would have never even begun this thesis work. I am really fortunate for having opportunity to work with her. Many thanks are also forwarded to the staff at the department of computer application, Lovely Professional University for always being helpful.

Table of Content

List of Tables

List of Figures

List of Appendices

Chapter 1

INTRODUCTION

Brain is the most important organ of our body and it serve as centre of nervous system. It is a bioelectric generator. Human brain has roughly 10^{11} neurons (serendip.brynmawr.edu). It is divided into four different lobes i.e. frontal lobe, parietal lobe, occipital lobe and temporal lobe With the neurological disordering (ND) and psychological illness (PI) their occur some disruption in the working of brain (en.wikipedia.org). Neurology is related to hardware wiring of the brain and Psychology is related to programming of the brain. EEG (electroencephalography) is a medical imaging test that measures the electric activity of the human brain (Teplan, 2002). EEG signals related to the level of consciousness of the person, and contain very useful information relating to different physiological state of brain .Thus it is very effective tool to understand the complex dynamics of the brain. It is a non intensive technique that shows the electric impulses of the brain and diagnoses the different brain diseases based on different parameter of time and frequency domain. Psychological, Cognitive, Physical and EEG based parameters play vital role in the diagnosis of EEG based diseases.

Cerebral cortex, the largest part of human brain is divided into four lobes i.e. frontal lobe, parietal lobe, occipital lobe and temporal lobe and they all the lobes are associated with different functions (www .psychology.about.com) as given in the following Table 1.1 are detected by EEG signals.

Table 1.1: Brain Lobes and its functions

Brain Lobes	**Location**	**Function**
Frontal lobe (FL)	At front part of the brain	Deals with motor activity (motion and movement), reasoning, planning.
Parietal lobe (PL)	At the middle part of brain	Deals with recognition, movement, orientation, perception.

Occipital lobe (OL)	At the back portion of the brain	Deals with the visual processing.
Temporal lobe (TL)	At the bottom portion of the brain	Deals with the perception, memory and speech.

Any type of inflammation and disturbance in these brain areas can cause psychological problems such as: anxiety, abnormal behavior, stress, fear and hallucination; cognitive problem such as: unconsciousness, confusion in decision making, visual and speech disabilities and judgments and physical problems such as: problem in walking, climbing, oversleeping, speech and visualization (www.nlm.nih.gov). There are more than 600 neurological diseases and all of them occur because of infection (meningitis, encephalitis and brain abscess), stroke, tumor, trauma (head injury) seizure, injuries and damaged of nerve cell (www.nlm.nih.gov).

Human brain generates different kinds of waves such as: alpha (α), beta (β), delta (δ), and theta (Θ). These waves are measured in Hertz (Hz) (Selva Kumari and jose, 2011). Lower HZs value shows the lower brain activity. Table 1.2 describe the frequency, amplitude and different states of the all the brain waves (www.medicine.mcgill.ca).

Table 1.2: Brain Waves and associated states

Wave	**Cycle per second (Frequency)**	**Amplitude**	**Generate in state of**
Delta	0-4 Hz	High (up to 200 μV)	Dreamless sleep, unconsciousness and meditation.
Theta	4-8 Hz	Low (5 μv-20 μv)	Deep relaxation and meditation.
Alpha	8-12 Hz	High (up to 200 μV)	Physical and Mental Relaxation.
Beta	13-30 Hz	Low (less than 10μV)	Alert and normal working state. Especially when people feel agitated, tense or afraid.

Due to overlapping of the physical, cognitive, psychological and the EEG based parameter in the all brain diseases it is difficult to diagnose particular diseases (Gangwar et al., 2013). Integrated intelligent systems such as: rule based reasoning (RBR)-case based reasoning (CBR), Data Mining (DM)-CBR deal with this problem. It has been found that diagnosis problem can be considerably reduced by utilizing these systems and these systems reduce the error amount and degree of uncertainty (Cabera and edye, 2010).

1.1 Basic concepts

1.1.1 EEG signals

EEG signals generate due to electrical activity in the human brain. Four types of signal are generated named alpha, beta, theta and delta. The entire signal generate in different stages. Fig 1.1 shows the different EEG signals.

Alpha : It appears when close the eyes or in relaxation .

Beta : It is a fast activity and appeared when feel agitated, tense or afraid.

Theta : It is a slow activity and appeared in medication

Delta : It appeared in dreamless sleep (www.medicine.mcgill.ca).

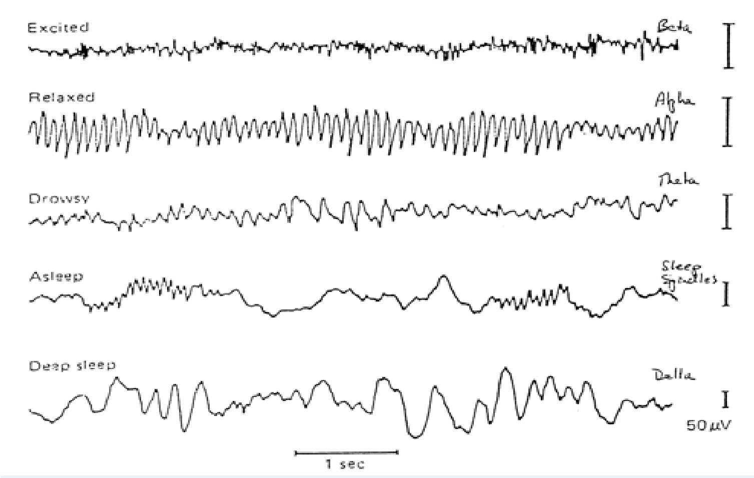


Fig 1.1: EEG signal (www.google.co.in/search?q=eeg+signal&source)

1.1.2 Psycho-physical parameters

Psychological (PSY) parameters such as: anxiety (AN), abnormal behaviour (AB), fear (FR), hyper activity (HA), stress (ST), hallucination (HL), agitation (AG) (Gangwar et al., 2012).

Cognitive (COG) parameters such as: confusion in decision making (CD),learning disability (LD), unconsciousness (UC) visual disability (VD), judgement (JD), speech disability (SD), memory loss (ML) (Gangwar et al., 2013).

Physical (PHY) parameters are problem in walking (WL), climbing (CL), hearing (HR), vision (VS), speech (SH), oversleeping (OS) (Gangwar et al., 2012).

1.1.3 EEG signal parameters

EEG signal abnormality observed in different brain lobe (BL) such as: frontal lobe (FL), parietal lobe (PL), occipital lobe (OL), temporal lobe (TL) (www .psychology.about.com).

Changing behaviour of the different brain waves (BW) such as: alpha (AL), beta (BE), theta (TH) and delta (DE) (Jaimchariyatam et al., 2011 and Ar. et al., 2011).

1.1.4 EEG based diseases

EEG based diseases are attention deficit hyperactivity disorder (ADHD), coma (C), dementia (D), epilepsy (E), mood disorder (MD), migraine (MG) and Schizophrenia (SI). All brain diseases can occur because of infection (meningitis, encephalitis and brain abscess), stroke, tumour, trauma (head injury) and seizure.

Attention deficit hyperactivity disorder (ADHD) ADHD is one of the most common mental disorder develop in children. It is a psychiatric disorder of the neurodevelopment type .The sign and symptoms of ADHD are hyperactivity, abnormal behaviour, learning and visual disabilities, difficulty staying focus (Gangwar et al., 2012). ADHD has three subtypes: predominantly hyperactivity impulsive, predominantly inattentive and combined hyperactivity impulsive and inattentive. Frontal lobe gets affected in this disease. Fig 1.2 shows the Hierarchical correlation of sign and symptoms (evidence) of Attention deficit hyper activity.

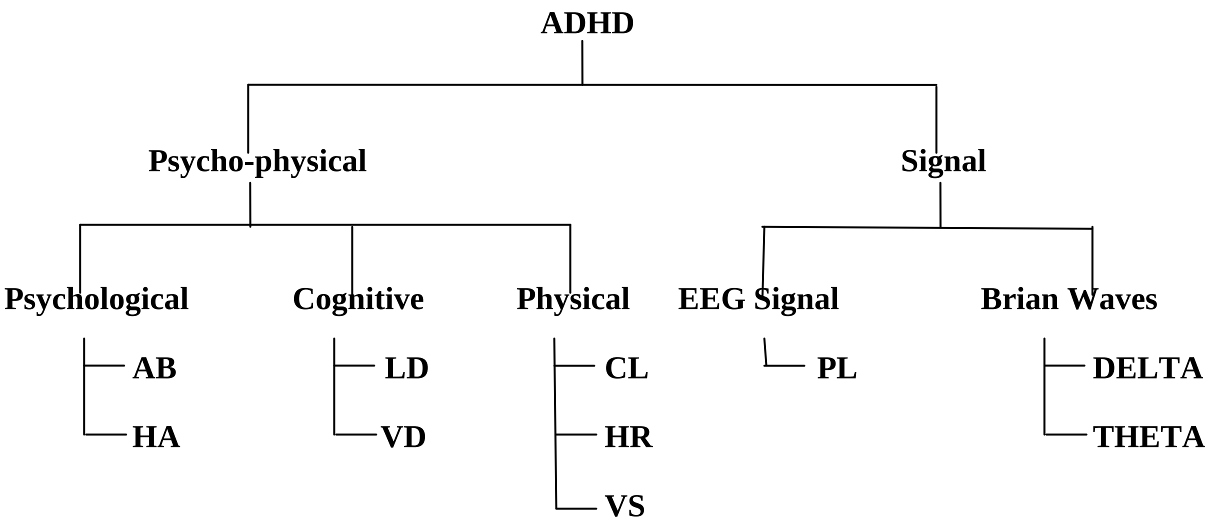


Fig 1.2: Hierarchical correlation of sign and symptoms (evidence) of the ADHD disease

Coma (C) Comma is a state of unconsciousness that can be caused by number of problems like brain injury, infection, stroke and brain tumour. It can be temporary or permanent. The sign and symptoms of C is unconsciousness, no motor activities, irregular breathing. Frontal, occipital and temporal lobes get affected in this disease (www. health communities. com). Fig 1.3 shows the Hierarchical correlation of sign and symptoms (evidence) of coma.

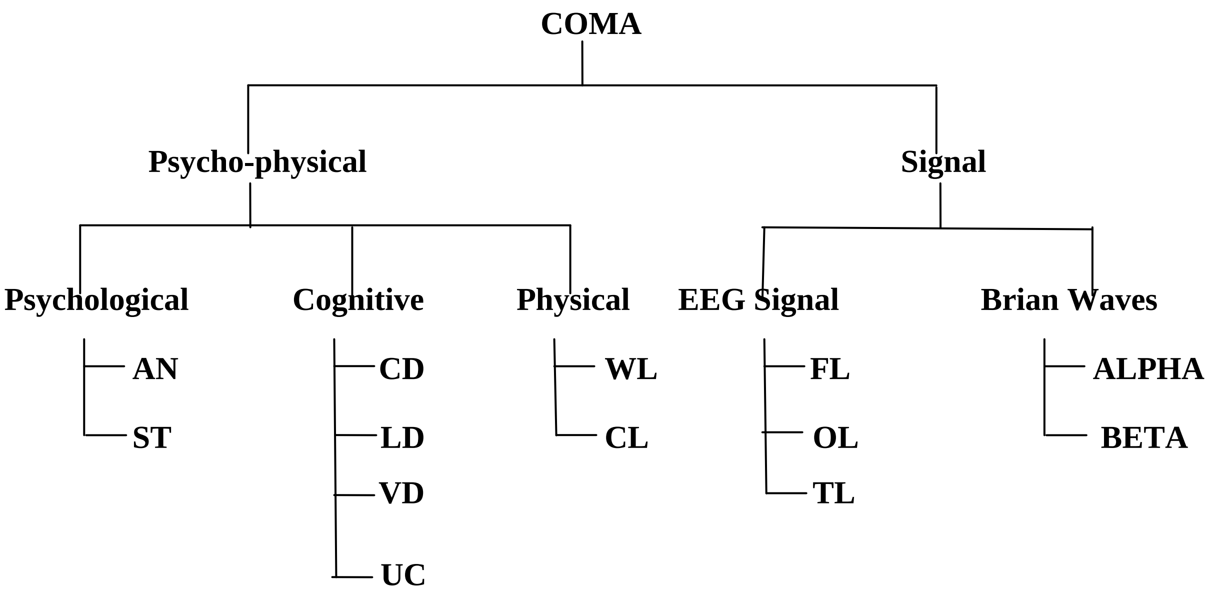


Fig 1.3: Hierarchical correlation of sign and symptoms (evidence) of the Coma disease

Dementia (D) Dementia is a problem of memory that causes the reduction in ability to learn reason or recall past experiences. The sign and symptoms of Dementia are: memory loss, poor judgment, abnormal behaviour, hallucination, confusion in decision making, problem in vision, hearing (Gangwar et al., 2013). All the brain lobes get affected in this disease. Fig 1.4 shows the Hierarchical correlation of sign and symptoms (evidence) of dementia.

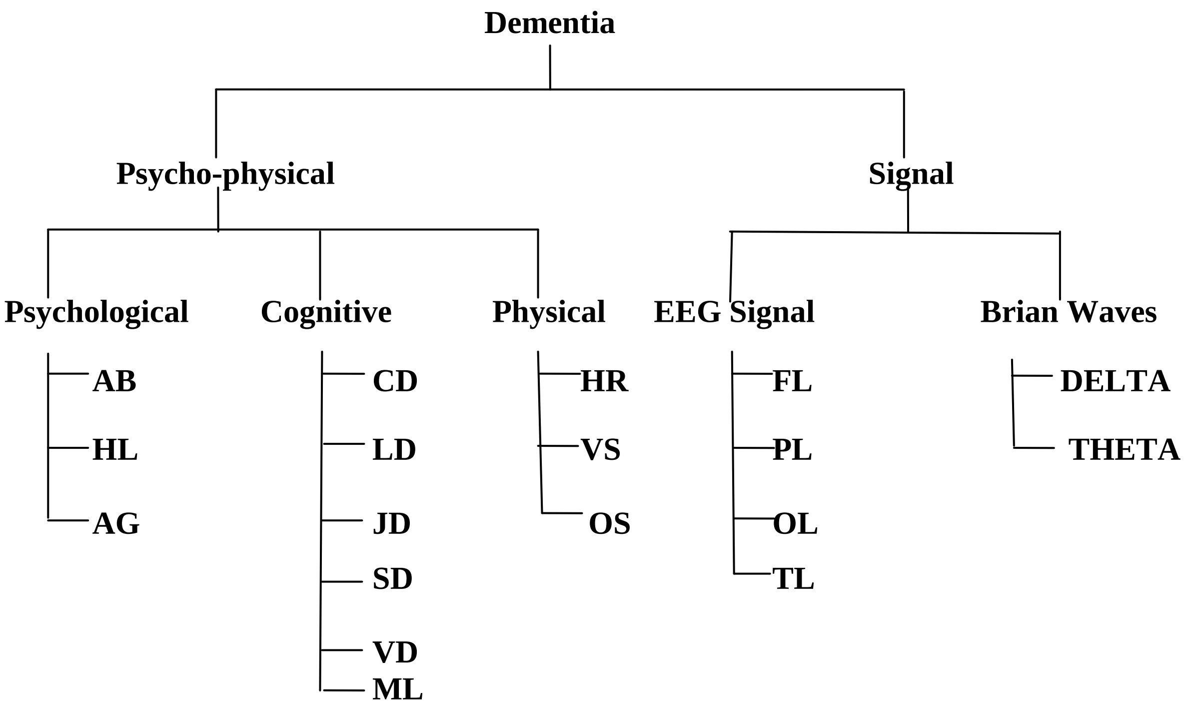


Fig 1.4: Hierarchical correlation of sign and symptoms (evidence) of the Dementia disease

Epilepsy (E) Epilepsy caused by some abnormal electrical activity in the brain. There occurs some sensory disturbance at sudden recurrent episode. And the patient has repeated seizures or convulsions over time. Temporal lobe is more affected. The sign and symptoms of epilepsy are anger, fear, anxiety, confusion in decision making, learning, vision and speech disabilities and consciousness (Gangwar et al., 2013). Fig 1.5 shows the Hierarchical correlation of sign and symptoms (evidence) of epilepsy.

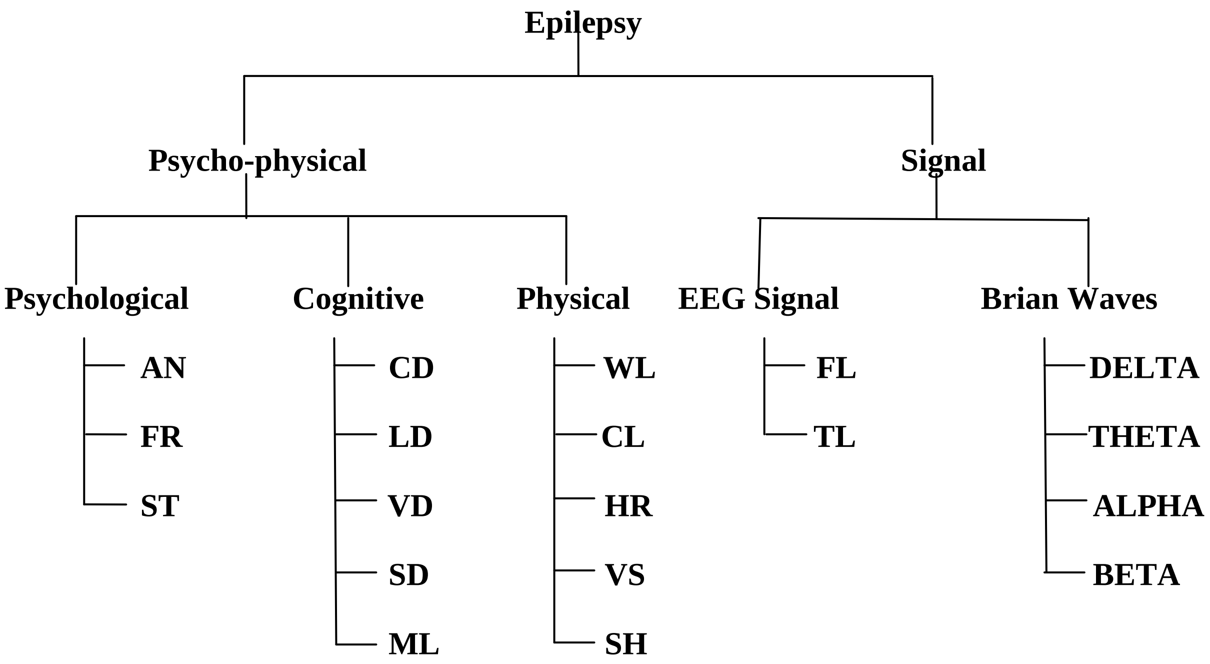


Fig 1.5: Hierarchical correlation of sign and symptoms (evidence) of the Epilepsy disease

Migraine (MG) Migraine is headache often on one side of head. The sign and symptoms of MG are depression, anger, agitation, muscular pains, oversleep. And all the brain lobes are affected in this disease (www.webmd.com). Fig 1.6 shows the Hierarchical correlation of sign and symptoms (evidence) of migraine.

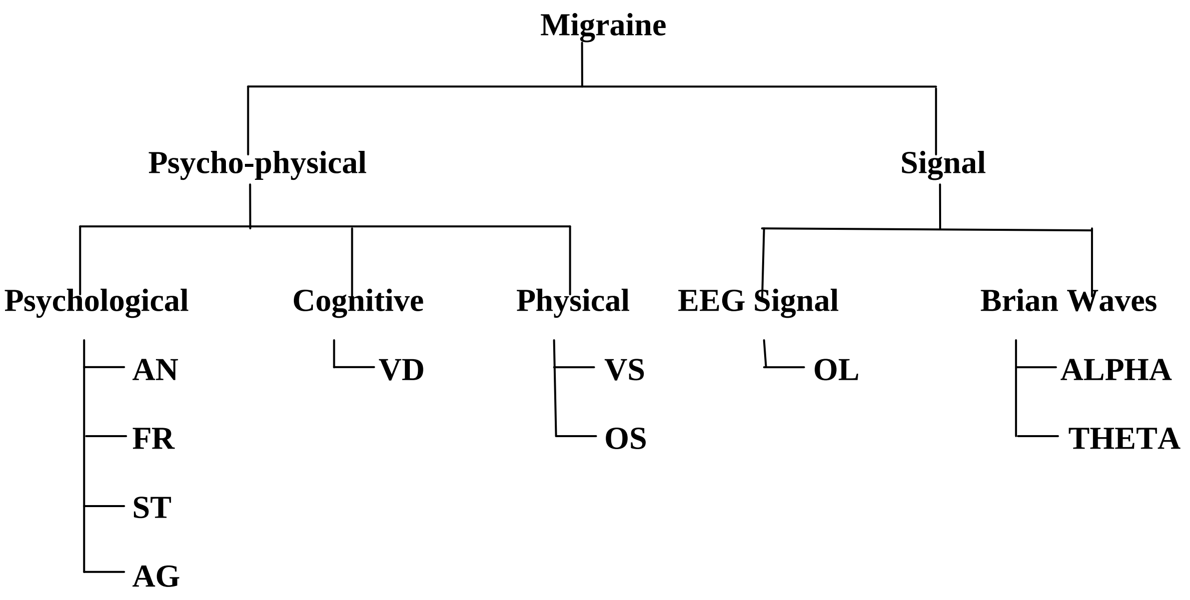


Fig 1.6: Hierarchical correlation of sign and symptoms (evidence) of the Migraine

Mood disorder (MD) Basic kinds of mood disorder are depression, cyclothymiacs, seasonal effective disorder and mania. The sign and symptoms of MD are anger, hallucination, stress, confusion in decision making, learning and vision disabilities and oversleep. Frontal and temporal lobes get affected in this disease (Gangwar et al., 2012). Fig 1.7 shows the Hierarchical correlation of sign and symptoms (evidence) of mood disorder.

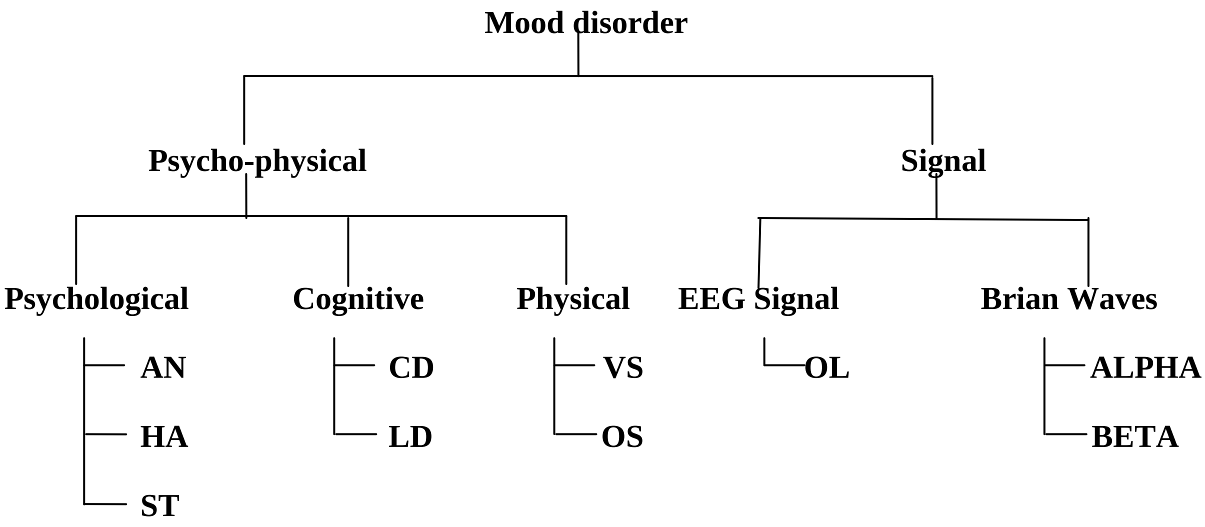


Fig 1.7: Hierarchical correlation of sign and symptoms (evidence) of the Mood disorder disease

Schizophrenia (SI) Schizophrenia is a long term mental disorder characterized by nervous breakdown it strikes in between the ages of 15 and 25. It mainly affects the thinking ability. The sign and symptoms of SI are abnormal behaviour, agitation, social withdrawal, hallucination, oversleep. Frontal lobe gets more affected in this disease (Gangwar et al., 2013). Fig 1.8 shows the Hierarchical correlation of sign and symptoms (evidence) of schizophrenia.

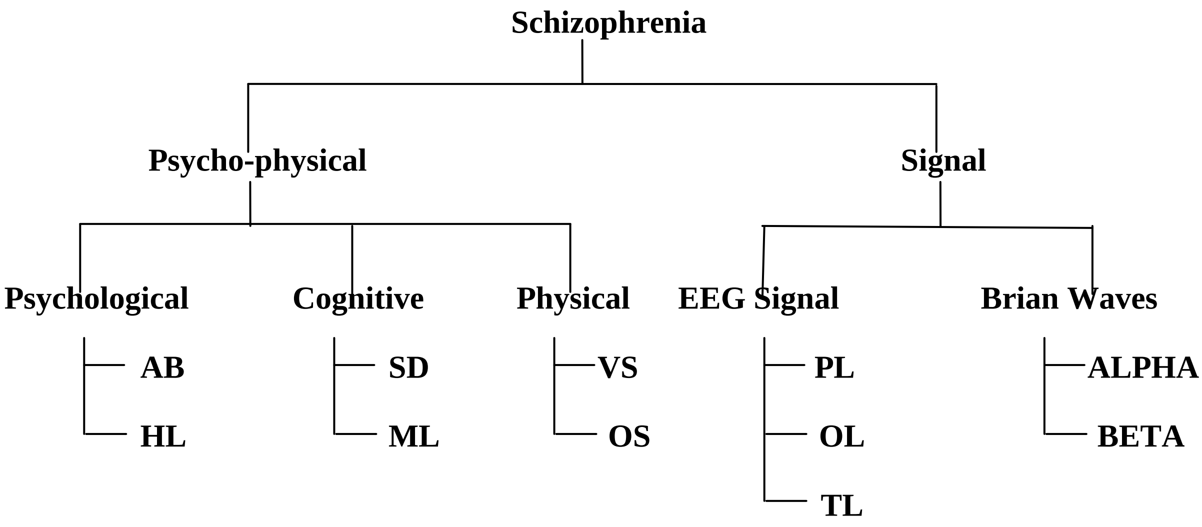


Fig 1.8: Hierarchical correlation of sign and symptoms (evidence) of the SI disease

1.2 Literature review

Abasolo et al. (2009), analyze the EEG background activity in Alzheimer's disease with deter ended moving average (DMA) method. It is a new approach to quantify correlation properties in non stationary signal. It provides an estimation of scaling information and long range correlation in time series. In this method firstly they integrate the EEG time, and then they find the moving average then subtract it from integrated profile and calculate the fluctuation for a window of size n.

Azim et al. (2010), Proposed a method for detection of sleep disorder breathing events through the analysis of electroencephalograph and Electromyography signal. The detection of these sleep disorder breathing event were based on the study of power spectral analysis of EEG and EMG signal. The variations in the various spectral bands of EEG signals are studied along with the variations in power of the corresponding EMG signals. They also discuss the probable causes of the aforementioned variations in the EEG and EMG signals.

Cohen et al. (2003), proposes a method for diagnosis of dementia using linear and non linear EEG signals. They described here, nonlinear EEG analysis based on computation of cortical potential followed by nonlinear analysis based on the computation of degree of variability is combined with imaging results and clinical parameters to form a diagnostic model for dementia diagnosis**.**

C.R. et al. (2007), propose a brain machine interface (BMI) design using five mental tasks. A Recurrent neural network is proposed for classification of EEG signals. Feature are extracted from EEG signals that are recorded during five mental tasks, namely Baseline–resting, mathematical multiplication, geometric figure rotation, letter composing and visual counting.

Exarchos et al. (2005), Presented a methodology based on the data mining algorithm which detects transient events in EEG recordings and classifies those as epileptic spikes, muscle activity, eye blinking activity and sharp alpha activity. It includes four stages 1) EEG preprocessing and transient events detection, (II) clustering of transient events and feature extraction, (III) feature discretization and (IV) association rule mining and classification.

Ganesan et al., 2007 proposes a system that integrate wavelet transform, Feature extraction and Artificial Neural Network for the detection and classification of Epileptic and non-Epileptic spike in EEG.

Glover et al. (1988), described a system for the automated interpretation of bioelectrical signals. A knowledge-based approach is taken. Providing two separate, cooperating knowledge bases, one for the signal analysis knowledge and one for the more subjective interpretation knowledge.

Gurumurthy and Tripathy (2010), propose a new Neural Network Architecture, which is a modified version of the conventional Back Propagation Network. The technique adopted by the Back Propagation Network (BPN) used in neural network. They designed software which could translate brain signal to digital text data. Two scans are used vertical and horizontal. When horizontal scan is performed signal fed in neural network which is trained to identify the bases of various brain signals. And when vertical scan is performed signal fed into other neural network which identify the brain signal on the EEG pattern.

Harischandra and perera (2012), They investigates the possibility of recognizing emotions using signal processing of Electroencephalography using discrete wavelet transform and feeding appropriate values to an adaptive neuro fuzzy inference system for classification.

Jansen et al. (1989), described knowledge based approach to automated sleep electroencephalogram (EEG) analysis. In this system, an object oriented approach is followed in which specific waveforms and sleep stages ("objects") are represented in terms of frames. The latter capture the morphological and patio-temporal information for each object. An object detection module ("frame matcher"), operating on the frames, is employed to identify what features need to be extracted from the EEG and to trigger the appropriate "specialist"-specialized signal. This leads to an opportunistic approach to EEG interpretation with quantitative information being extracted from the signal only when needed by the reasoning processes.

Li et al., 2011 developed an EEG based BCI system for classification, they apply two development Adaboost classifiers on the basis of an advanced boosting learning algorithm: AdaboostNN and Gentle Adaboost.

M. Teplan (2002) , presents an introduction into EEG measurement. The paper divides into two parts. First part is related to background of the subject, and some related search area. Second part explains EEG recording techniques such as: 1) Electrodes conductive media, 2) Amplifiers with filters, 3) A/D converter, 4) Recording device.

Pandey and Mishra (2009), Developed intelligent computing systems (ICS) for diagnosis of neuro muscular diseases. They made a study of different singular and combined methods applicable in medical domain to diagnose the disease. They had made review of the different methods of KBS, ICS and their combinations for the detection and diagnosis of different diseases. The KBS comprises RBR, CBR and MBR. The ICS consists of a ANN, GA and FL. The combined methods are RBR–CBR, CBR–MBR, RBR–CBR–MBR, ANN–CBR, ANN–GA, RBR–GA and CBR– RBR–ANN, FL–GA, FL–CBR, FL–RBR, FL–GA–ANN and FL–RBR–ANN.

Pal et al. (2010), proposes a simple method for EEG classification based on Fourier features. Parameters like energy, entropy, power, and kurtosis. After calculating the above mentioned parameters of the discussed signals, they found that without going for rigorous time-frequency domain analysis, only frequency based analysis is well suitable to classify various EEG signals.

Patil and Sherekar (2013), deployed the naive bayes and J48 algorithm on the bank data set. It based on the correct and incorrect instances of data classification. The experiments results of the work on this dataset show that the efficiency and accuracy of j48 is better than that of Naïve Bayes.

Sheniha et al. (2013), propose a method called ANFIS-DE (Adaptive Neuro Fuzzy Inference System (ANFIS) tuned by Differential Evolution (DE) algorithm) to find the artefact signal and to extract the EEG signal from stained EEG signal.DE algorithm is used to find optimum parameters of ANFIS.

Sudirman and Seow (2009), Developed an algorithm for EEG based hearing identification system using AI. Feed forward propagation neural network can identify the brain signals effectively. Back propagation algorithm is used to recognize the pattern of brain waves. And EEG signals are analyzed using Fast Fourier Transform (FFT) and filtering technique available in mat lab. System will enhance the analysis quality and accuracy of the analysis.

Sulaiman et al. (2011) designed an intelligent system to evaluate human stress level using EEG signals and Psychoanalysis tests. A graphical user interface is created. The system successfully shows the stress level, stress score and index of EEG signals.

Vijila et al. (2007), proposed a hybrid soft computing techniques called adaptive neuro –fuzzy inference system (ANFIS) to estimate interference and to separate the EEG signal from its ECG, EOG, EMG artefacts. Interference signal can be filtered out by identifying a linear model between an artefact and the immeasurable interference.

Wu et al. (2010), proposes an effective classification method on probabilistic neural network with supervised learning. Genetic algorithm is used training the neuron's smoothing parameter and hidden central vector for determining hidden neurons.

Following Table 1.5 describe the comparative study of different techniques that are used for diagnose EEG based diseases.

Table 1.3: Comparative study of different technique for diagnosis of EEG based diseases

S.No	Author's Name	Technique	Benefits
1.	Gurumurthy and Tripathy, 2010	Modified version of the conventional Back Propagation Network	This designed software could translate brain signal to digital text data.
2.	Jansen et al., 1989	Knowledge based approach	This leads to an opportunistic approach to EEG interpretation with quantitative information being extracted from the signal only when needed by the reasoning processes.
3.	C.R. et al., 2007	Recurrent neural network	Feature are extracted from EEG signals that are recorded during five mental tasks, namely Baseline–resting, mathematical multiplication, geometric figure rotation, letter composing and visual counting
4.	Sudirman and Seow, 2009	Feed forward propagation neural network, Back propagation algorithm and Fast Fourier Transform (FFT) and filtering technique available in mat lab	System will enhance the analysis quality and accuracy of the analysis.
5.	Glover et al., 1988	Knowledge-based approach, Providing two separate,	System is used for the automated interpretation of bioelectrical

		cooperating knowledge bases	signals.
6.	wu et al., 2010	Probabilistic Neural Network (PNN) with supervised learning, Genetic algorithm	System used training the neuron's smoothing parameter and hidden central vector for determining hidden neurons.
7.	Pal et al., 2010	Fourier features. Parameters like energy, entropy, power, and kurtosis are used.	System found that without going for rigorous time-frequency domain analysis, only frequency based analysis is well suitable to classify various EEG signals.
8.	Cohen et al., 2003	linear and non linear EEG signals	System will diagnose the dementia diseases.
9.	Abasolo et al., 2009	Deter ended moving average (DMA) method	System provides an estimation of scaling information and long range correlation in time series
10.	Pandey and Mishra, 2009	Different methods of KBS (knowledge base system), ICS (intelligent computing system) and their combinations	They made a study of different singular and combined methods applicable in medical domain to diagnose the disease.
11.	Vijila et al., 2007	Adaptive neuro –fuzzy inference system (ANFIS)	Estimate interference and to separate the EEG signal from its ECG, EOG, EMG artefacts.
12.	Exarchos et al., 2005	Data mining algorithm	System detects transient events in EEG recordings and classifies those as epileptic spikes, muscle activity, eye blinking activity and sharp alpha activity.
13.	Azim et al., 2010	Electroencephalograph and Electromyography signal.	System detects sleep disorder breathing events through the analysis of EEG and EMG signal. The detection of these sleep disorder breathing event were based on the study of power spectral analysis of EEG and EMG signal.
14.	Patil and Sherekar, 2013	Naive Bayes and J48	J48 gives best accuracy, efficiency and is cost efficient for classification than the Naive Bayes Algorithm.

15.	Ganesan et al., 2007	Feature extraction and ANN	System integrate wavelet transform, Feature extraction and Artificial Neural Network for the detection and classification of Epileptic and non-Epileptic spike in EEG
16.	Li et al., 2011	Boosting learning algorithm: AdaboostNN and Gentle Adaboost	System is used for classifying EEG based diseases using boosting learning algorithm.
17.	Sheniha et al,2013	ANFIS-DE (Adaptive Neuro Fuzzy Inference System (ANFIS) tuned by Differential Evolution (DE) algorithm)	System is used to estimate the artifacts and extract the EEG signals from stained EEG signals.
18.	Sulaiman et al. (2011)	EEG signals, Psychoanalysis tests and questionnaire	System successfully shows the stress level, stress score and index of EEG signals by asking questions.
19.	Harischandra and perera, (2012)	Discrete wavelet transform and adaptive neuro fuzzy	System is used to detect emotions of motor disabled people and provision a means of communication; also it is a learning tool for trainee neurologists.

1.3 Objective

As discussed above Psychological, Cognitive and Physical symptoms play important role in the diagnosis of EEG based diseases and integrated intelligent method will improve the diagnosis accuracy.

Therefore the central hypothesis: is to prove that Integrate intelligent diagnostic system based on cognitive, physical and psychological as well as EEG based systems produce more accurate result for diagnosis of EEG based disease.

Aim1: To identify physical, cognitive, psychological and EEG based parameters (symptoms) for EEG based diseases.

Aim 2: To integrate RBR-CBR for diagnosis of EEG based diseases.

Aim 3: To apply DM for feature selection and classification.

Aim 4: To integrate DM-CBR for diagnosis of EEG based diseases.

Aim5: To compare the performance of RBR-CBR and DM-CBR.

The main objective of this work is:

To develop an integrated intelligent system for diagnosis of EEG based diseases.

1.4 Methods

Artificial Intelligence techniques (AI)

Artificial techniques can be categorized in two groups (as shown in fig: 1.9): knowledge oriented such as: RBR and CBR; and data oriented such as: Artificial Neural Network (ANN) and DM. All these techniques present a complementary approach to each other. ANN complementing the knowledge base system with respect to numeric knowledge representation (Pandey and Mishra, 2009).

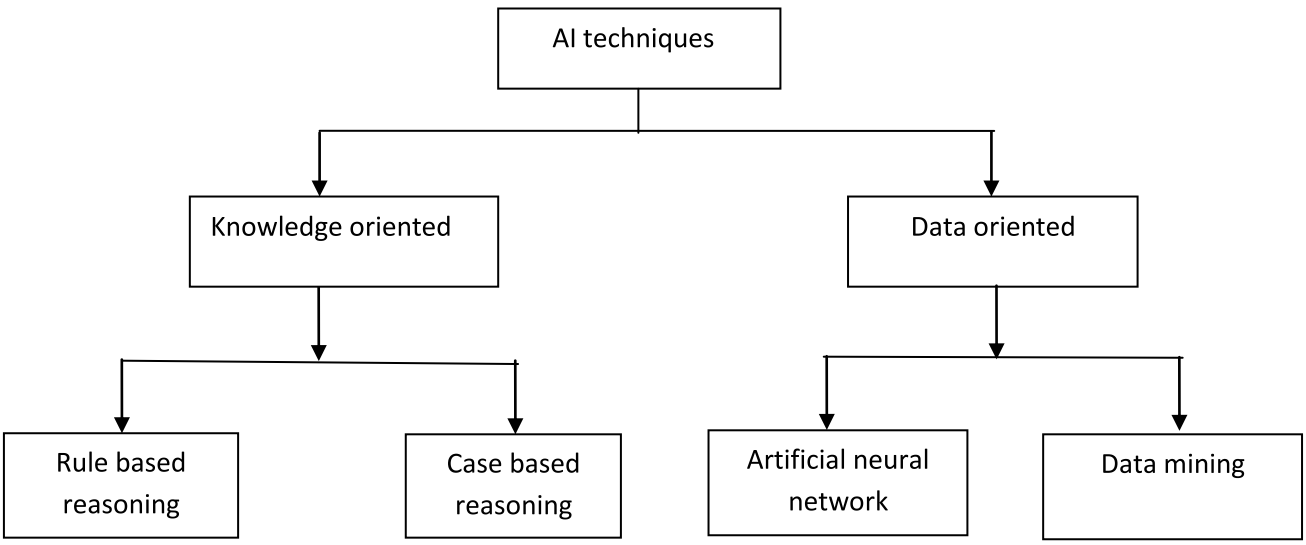


Fig 1.9: Various AI techniques

1.4.1 Rule Based Reasoning (RBR)

RBR consists of three component rule/ knowledge base, user interface, working memory and inference engine (en.wikipedia.org) as shown in fig 1.10. Knowledge base itself contains long term information elicited from the expert clinician. And short term from the data. Knowledge is gathered from many years of experience, and captured as a series of IF-THEN rules. RBS/KBS use backward chaining and forward chaining. If all the data and facts are already known then forward chaining is used ,few or no facts are known goal is to find them then prefer backward chaining. Although RBR has some advantages such as: modularity, naturalness, uniformity and compact representation (Pandey and Mishra, 2009). However it faces some problem also like difficulty in representing informal information and inference efficiency problem.

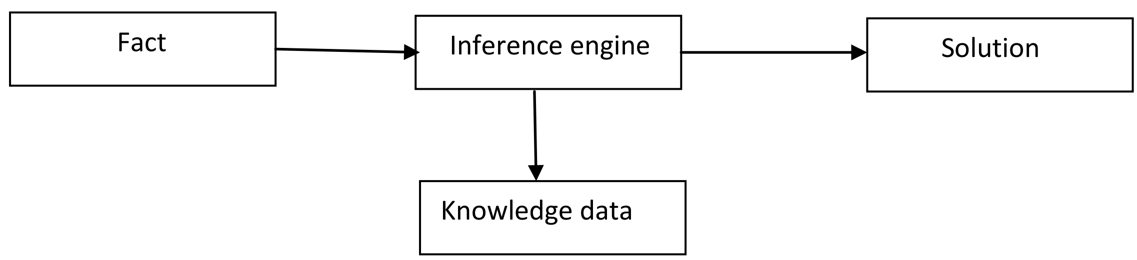


Fig 1.10: Rule Based model

1.4.2 Case Based Reasoning (CBR)

CBR is a process of solving new problems by utilizing the knowledge that is gathered from the past experience and solution. The CBR Systems consist of bunch of past cases. The cases consist of a description of problem, solution and outcome. The new case is solved by firstly matching the problem with the existing cases if matched then system retrieve them, then these cases suggest the solution, if needed system revise the solution. Finally the solution is retained as part of new case (Aamodt and Plaza, 1994). That is why these systems are self updatable. Some advantages of case based reasoning are easy knowledge acquisition, learning from experiences, inference engine (Cabrera and Edye, 2010). CBR is an approach that uses anecdotal evidence as the main operating principal. If collected data is non-statistically relevant there is no guarantee the result is correct. These knowledge based techniques are not so good in pattern recognition applications. CBR has been formalized as a four step process retrieve, reuse, revise and retain (Aamodt and Plaza, 1994) as shown in fig 1.11.

Retrieve: - For a given problem, retrieve the most similar cases from the already collected cases.

Reuse: - Use the retrieved cases to solve current problem.

Revise: -After mapping the solution now test the new solution in real world.

Retain: -When solution is successfully adapted to the current problem then store the resulting solution as new case in the memory.

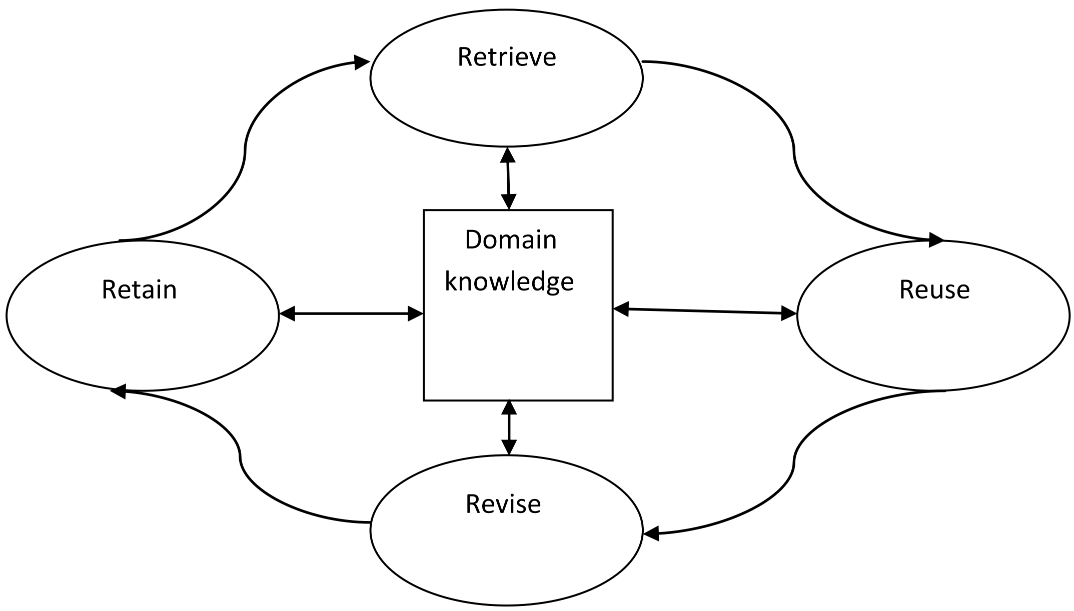


Fig 1.11: Case Based model

1.4.3 Artificial Neural Network (ANN)

ANN is inspired by the way biological nervous system. They are configured for special applications like pattern recognition, data classification, speech, vision, signal processing, and control systems (http://en.wikipedia.org/wiki/Artificial_neural_network). ANN with back propagation (BP) algorithm has been widely used in medicine because of its simplicity. ANN has some advantages over the rule based system (Pandey and Mishra, (2009). It compliment the other knowledge based systems. It has some disadvantages like structure of ANN is not transparent and neural network required training to operate, and require high processing time for large neural network.

The structure of neural network is produced with input layer, hidden layer and output layer. Input layer receive the data, then again transfer them to hidden layer. Data process and transfer result to the neuron in the next layer that is output layer as shown in fig1.12.

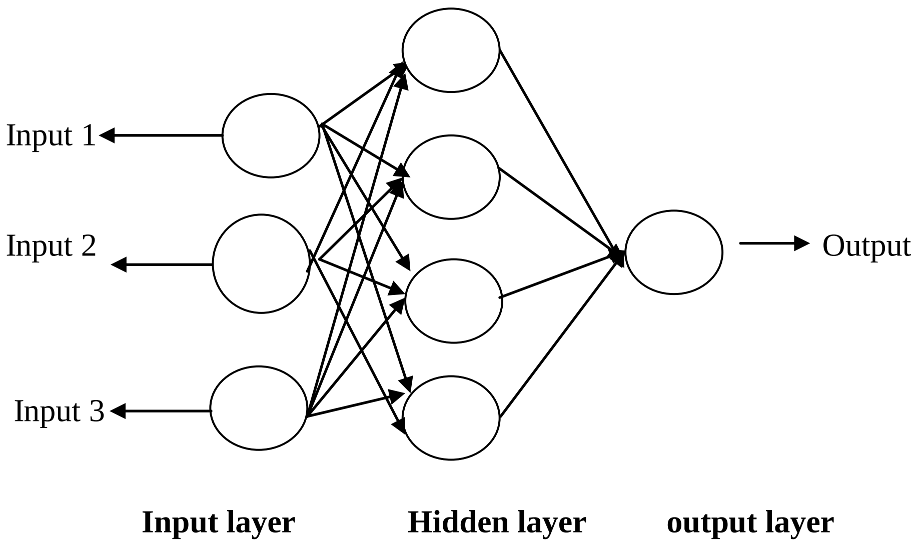


Fig 1.12: Architecture of Neural Networks

1.4.4 Data mining (J48 algorithm)

Data mining is a concept in which data is analyzed from different perspectives and summarize the information in more useful way. It is also called knowledge discovery from the database (en.wikipedia.org/wiki/Data mining).it is mostly use in business activities. For example to analyze how good your business is doing. And predict how good will it does in future. Data mining works 1) Firstly extract transform and load the data on to data ware house. 2) Store and manage the data in multidimensional DBMS.3) Business analyst analyze the data by application software .4) Present the data in graph and table forms which is useful to take (en.wikipedia.org /wiki/Extract,transform,_load).

Data mining is a non trivial extraction of implicit, previously unknown, and imaginable useful information from data. DM is used to find important information hidden in large volumes of data. Various DM techniques are Association rule, classification, regression, summarization and clustering. Classification methods such as: j48, it classify the data according to the features of the data with respect to the predefined set of classes.

J48 algorithm

J48 algorithm is the java implementation of c4.5 algorithm which generate decision tree. It is a classification technique of Data Mining implemented in WEKA (Waikato environment for knowledge analysis) tool. J48 construct univariate decision tree (Bharhava et al., 2013) with the information gain process (IGP). Entropy is used to measure the IGP with eq. 1 and eq. 2.

$$\text{Entropy(p)} = -\sum_{j=1}^{n} \frac{|pj|}{|p|} log \frac{|pj|}{|p|} \qquad \text{eq. (1)}$$

$$\text{Entropy(j/p)} = \frac{|pj|}{|p|} log \frac{|pj|}{|p|} \qquad \text{eq. (2)}$$

At each node of tree, J48 choose that attribute of data that most effectively splits its sample into subset (Patil and Shereker, 2013). The attribute with the highest information gain is selected to make decision set as node. Finally information gain is calculated by eq.3 that is subtraction of eq.1 and eq.2

$$Gain(p, j) = Entropy(p - Entropy(j\backslash p) \quad \text{eq. (3)}$$

Some advantages of RBR, CBR, ANN and DM are shown below in Table 1.4 and Table 1.5

Table 1.4: Advantages of Knowledge Oriented Techniques (RBR and CBR) (Pandey and Mishra, (2009).

Item	RBR	CBR
Advantages	1. Separation of knowledge and content 2. Modularity 3. Uniformity 4. Naturalness 5. Compact representation of knowledge	1.Rreduce knowledge acquisition 2.Learning from past experiences 3. Modularity 4. Naturalness 5. Self updatable 6.Improve as number of cases grows

Table 1.5: Advantages of Data Oriented Techniques (ANN and Data Mining)

Item	ANN	Data Mining (J48 algorithm)
Advantages	1. Implicitly detect complex non linear relationship between dependent and independent variable. 2. Run very quickly. 3. Can process huge amount of data. 4. Helpful in pattern recognition.	1. Discover the knowledge (analysis of data). 2. Helpful in prediction and Decision making. 3. Classify the data. 4. Can handle nominal and Numeric values.

1.5 Plan of thesis

Chapter 1: Deals with introduction.

Chapter 2: Describe the integration of Rule Based Reasoning (RBR) and Case Based Reasoning (CBR) for diagnosis of EEG based diseases.

Chapter 3: Describe the classification of EEG based diseases using Data Mining (J48).

Chapter 4: Describe the integration of Data Mining (J48) and Case Based Reasoning (CBR) for diagnosis of EEG based diseases.

Chapter 5: Deals with the Result and conclusion.

Chapter 2

Integration of Rule Based Reasoning and Case Based Reasoning for Diagnosis of EEG based Diseases

---

This chapter describe a method by integrating two knowledge based systems i.e. rule based system (RBR) and case based system (CBR) for classification of five type of EEG based diseases viz., attention deficit hyperactivity disorder (ADHD), dementia (D), epilepsy(E) , mood disorder (MD) and schizophrenia (SI). The analysis of diagnosis of brain activity based upon flow of electricity due to similarity and overlapping of the sign and symptoms is complex and difficult In the analysis, symptoms based upon several parameters such as physiological (muscle weakness, difficulty in climbing etc) cognitive (learning disability, speech disability etc) psychological (stress, anxiety etc) and EEG signal parameters are taken for investigation. It has been found that diagnosis problem can be considerably reduced utilizing these integrated intelligent technique.

2.1 EEG based diseases: sign and symptoms

Table 2.1 describes EEG based diseases such as: attention deficit hyper activity (ADHD), dementia, epilepsy, mood disorder and schizophrenia with their two important parameters: Psycho-physical (PP) parameters and EEG signal characteristics. Psycho-Physical parameter is further divided into three parts: Psychological, Cognitive and Physical. Psychological parameters consisting of parameter such as anxiety (AN), abnormal behaviour (AB), fear (FR), hyper activity (HA), stress (ST), hallucination (HL), agitation (AG). The cognitive parameter are confusion in decision making (CD),learning disability (LD), visual disability (VD), judgement (JD), speech disability (SD), memory loss (ML). The physical parameters are problem in walking (WL), climbing (CL), hearing (HR), vision (VS), speech (SH), oversleeping (OS) (Gangwar et al., 2012).

EEG signal characteristic is divided into two parts :1) EEG signal abnormality observed in different brain lobe such as frontal lobe (FL), parietal lobe (PL), occipital lobe (OL), temporal lobe (TL) (Gangwar et al., 2012). 2) Behaviour of the different EEG waves such as: alpha, beta, theta and delta (Selva Kumari and jose, 2011).

Here the symbols H, M, L are used as shown in Table 2.1. H=3 is high that represents these symptoms are presents at high rate. M=2 is medium that represents symptom is presents but not very much and L=1 is low that represents the absence of symptom. For example, dementia has cognitive symptom such as forgotten memory –H, problem in speech – M and hearing disability – L.

Table 2.1: Sign and Symptoms of EEG based diseases

DISEASE	**Psycho-physical Parameters**																			**Signal**							
	Psychological							**Cognitive**						**Physical**						**EEG signal**				**Brain waves**			
	AN	**AB**	**FR**	**HA**	**ST**	**HL**	**AG**	**CD**	**LD**	**VD**	**JD**	**SD**	**ML**	**WL**	**CL**	**HR**	**VS**	**SH**	**OS**	**FL**	**PL**	**OL**	**TL**	**Delta**	**Theta**	**Alpha**	**beta**
ADHD	L	H	L	H	L	L	L	L	H	M	L	L	L	L	H	H	M	L	L	L	L	L	L	M	H	L	L
Dementia	L	M	L	L	L	H	H	H	H	M	H	M	H	L	L	M	M	L	H	H	H	M	H	H	H	L	L
Epilepsy	H	L	H	L	M	L	L	H	M	M	L	M	M	H	H	M	M	H	L	H	L	L	H	M	M	M	M
Mood disorder	H	L	L	H	H	L	L	H	H	L	L	L	L	L	L	L	M	L	H	L	L	M	L	L	L	H	H
Schizophrenia	L	M	L	L	L	M	L	L	L	L	L	H	H	L	L	L	M	L	H	L	H	M	M	L	L	M	M

2.2 Rule Based Reasoning model

The rule based model is developed for all the diseases based on the all the taken parameters and their correlation. Rule Based Reasoning model generate the rules for all the diseases. It will be clear by the following Table 2.2.

Table 2.2: Rules for the Rule Based Reasoning Model

R1→ when there is same level of all the symptoms	R2→ when there is difference of one level of symptoms	R3→ when there is difference of two level of symptoms
HΦH→H	HΦM→H	HΦL→M
MΦM→H	MΦL→L	
LΦL→L		

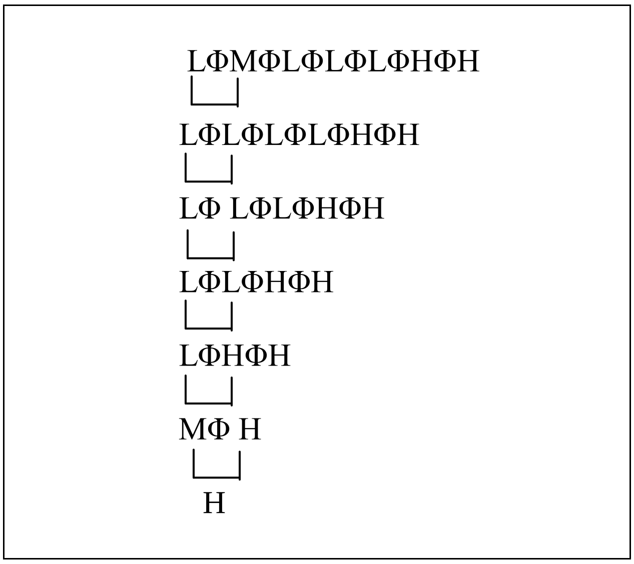
Rule model is developed for the dementia disease on the basis of Psycho-physical and EEG bases parameters and their correlation in the particular disease, using rules from Table 2.2. Similarly, rules for other diseases can be generated.

Rule model for the dementia diseases:

Psychological parameters are:	LΦMΦLΦLΦLΦHΦH→H
Cognitive parameters are:	HΦHΦMΦHΦMΦH→H
Physical parameters are:	LΦLΦMΦMΦLΦH→M
EEG signal on different brain region	HΦHΦMΦH→H
Behavior of EEG waves	HΦHΦLΦL→L

Rule Base Reasoning reduced the dimension of symptoms using (Table 2.2). Following in the fig 2.1, dimension of the psychological parameters of dementia disease is reduced**.**

Fig 2.1: Reduction of psychological parameter of dementia disease



2.3 Case Based Reasoning model

The different cases are generated to diagnose the different EEG diseases and can solve new problems by using the existing cases. Every case contains a description of different problem. The reduced form of the RBR is case based reasoning (CBR). That is why retrieval process is more efficient in CBR. Firstly retrieve the data from the case base after that adapt or use it and finally diagnose the disease as the final result.

2.3.1 Case base

The cases in a case base are generated by rules of RBR by using table 2.2. The rule reduced the dimensions of cases in a case base. Table 2.3 shows the case base used in case base reasoning.

Table 2.3: Cases of Case Based Reasoning Model

Case	Psychological	Cognitive	Physical	EEG (BR)	EEG Waves	Diagnosis
C1	M	M	H	L	L	ADHD
C2	H	H	M	H	M	Dementia
C3	L	H	M	M	M	Epilepsy
C4	H	M	M	M	M	Mood disorder
C5	M	M	M	H	M	Schizophrenia

2.3.2 Retrieval

Cases are retrieved using matching and selection. Matching is done using following similarity measure (eq.1 and eq.2)

$$S=\sum_{i=1}^{n} wi \mid U_i - C_i \mid \qquad \text{--- eq. (1)}$$

Where, Ui is ith feature of user case, Ci is ith feature of case of case base, Wi is weight assign to different parameter by domain experts.

Weight assign to psychological parameter is 0.1, cognitive parameter is 0.1, physical parameter is 0.2, EEG signal parameter is 0.3 and EEG wave's parameter is 0.3.

Selection is done using the following formula (eq.3):

$$R=\min_{i=1}^{n} (S_i) \qquad \text{--- eq. (2)}$$

The case with the smallest S_i value is the closest case. For example: If user input case entered values are M for psychological, L for cognitive, H for physical, M for EEG (BR) and M for waves. Then we subtract the all user input case values corresponding to each case (C1, C2, C3, C4 and C5) values in a case base .The minimum value among them will be our result.

2.3.3 Adaptation

Case from the case base is selected which is more similar to the user case, and the solution is used without any modification. For example suppose if the user case value is more similar to the C3. User simply concludes that symptoms are related to the ADHD disease.

2.4 Result

Table 2.5 shows the result by applying the similarity measure formulae (eq.1 and eq.2). Here weights (W) for different parameters are assigned by the experts. Ci (1<i<n) are the different previous cases stored in the case base and U is the new case entered by the user. User enters the input of new case in the range of H, M and L. Following with the help of one example it will be cleared.

Suppose user enters the user case input as shown in Table 2.4

Table 2.4: User case value

User case	Psychological	Cognitive	Physical	EEG (BR)	EEG Waves
U	M	L	H	M	M

Now, if their exist any similar case in the case base then system will simply retrieve the matched case and diagnose the diseases but If no case is matched then system will apply the similarity measure formula and subtract the value of all parameters of all cases one by one from user case. After that system diagnose the diseases by calculating the minimum different between the user case and the all cases of case based.

Table 2.5: Calculation of difference between user input and case in case base

	|Ui-Ci|					**N** **$S=\sum w_i \mid U_i-C_i \mid$** **i=1**	**Disease**
	i=1	i=2	i=3	i=4	i=5		
Weight	w1=0.1	w2=0.1	w3=0.2	w4=0.3	w5=0.3		
U	M=2	L=1	H=3	M=2	M=2	(0.1)2+(0.1)1+(0.2)3+(0.3)2+(0.3)2=2.1	
U-C1	M-M 2-2=0	L-M 1-2=1	H-H 3-3=0	M-L 2-1=1	M-L 2-1=1	(0.1)0+(0.1)1+(0.2)0+(0.3)1+(0.3)1=0.7	ADHD
U-C2	M-H 2-3=1	L-H 1-3=2	H-M 3-2=1	M-H 2-3=1	M-M 2-2=0	(0.1)1+(0.1)2+(0.2)1+(0.3)1+(0.3)0=0.8	Dementia
U-C3	M-L 2-1=1	L-H 1-3=2	H-M 3-2=1	M-M 2-2=0	M-M 2-2=0	(0.1)1+(0.1)2+(0.2)1+(0.3)0+(0.3)0=0.5	Epilepsy
U-C4	M-H 2-3=1	L-M 1-2=1	H-M 3-2=1	M-M 2-2=0	M-M 2-2=0	(0.1)1+(0.1)1+(0.2)1+(0.3)0+(0.3)0=0.4	Mood disorder
U-C5	M-M 2-2=0	L-M 1-2=1	H-M 3-2=1	M-H 2-3=1	M-M 2-2=0	(0.1)0+(0.1)1+(0.2)1+(0.3)1+(0.3)0=0.6	Schizophrenia

As clear from the table the smallest difference value is with case 4 (C4). So, diagnosis disease is mood disorder.

2.5 Conclusion

This chapter, develop integrated method based on RBR and CBR for the diagnosis of EEG based diseases. Initially, collect the different symptoms based on physical, psychological, cognitive and EEG based parameter. Then the Rule based approach used for defining rules and correlates the sign and symptoms with different disease. It helps to generate and reduce the dimension of the cases in a case base. Case base reasoning is used to diagnose the disease. The reduced dimension of cases in a case base improves the retrieval efficiency.

Chapter 3

Classification of EEG based diseases using Data Mining (J48)

In this chapter, J48 algorithm is deployed for the classification of the different EEG based diseases such as: attention deficit hyperactivity disorder (ADHD), dementia (D), epilepsy (E), mood disorder (MD) and schizophrenia (SI) using physical, psychological, cognitive and EEG based parameters brain lobes and brain waves. J48 algorithm generates decision tree, pruned tree and confusion matrix that help us to classify the different diseases. The accuracy of result is measure in term of sensitivity (SN), specificity (SF) and Receiver operating characteristics curve (ROC). SN and SF are calculated with the help of true positive rate value (TPR) and false positive rate value (FPR) calculated from confusion matrix. ROC simply compares the TPR and FPR values.

.

3.1 Implementation

The method is implemented using WEKA tool (J48). Data which is used for the experiment is collected from hospital and experts. The data set consist of 100 records and 5 attributes. In which 75 records are used for training and 25 records are used for testing. Detailed description of data set is given in Table 3.1.

Table 3.1: Detail of dataset for five diseases

Attributes	Data type
Psychological	Categorical((high ,medium ,low)
Cognitive	Categorical(high, medium, low)
Physical	Categorical(high, medium, low)
Brain lobes	Categorical(high ,medium, low)
Brain waves	Categorical(high ,medium, low)

Fig 3.1 shows the data set that is used by J48 algorithm for the classification

A	B	C	D	E	F
PSY	COG	PHY	BL	W	outcome
high	high	medium	high	medium	D
medium	medium	medium	high	medium	SI
medium	medium	high	low	low	ADHD
low	low	low	low	low	ND
low	high	medium	medium	medium	E
high	medium	medium	medium	medium	MD
medium	medium	medium	high	medium	SI
high	medium	high	low	low	ADHD
high	high	medium	high	medium	D
low	medium	low	low	low	ND
high	medium	medium	medium	medium	MD
low	medium	medium	medium	medium	E
medium	medium	low	low	medium	ND
high	medium	high	medium	medium	MD
medium	high	medium	medium	medium	E
medium	high	high	low	low	ADHD
medium	medium	medium	high	medium	SI
high	high	high	high	medium	D
high	medium	medium	high	medium	SI
medium	medium	high	low	medium	ADHD
medium	medium	medium	low	medium	E

Fig 3.1: Data set used in WEKA

3.2 Result

3.2.1 Pruned tree: J48 generate pruned tree and then rules are derived from pruned tree as shown in fig 3.2 and 3.3 respectively.

```
J48 pruned tree
------------------

brainlobe = high
|   cognitive = high: dementia (7.0/1.0)
|   cognitive = medium: SI (10.0/2.0)
|   cognitive = low: SI (1.0)
brainlobe = low
|   physical = medium: no disease (3.0/1.0)
|   physical = high: ADHD (10.0)
|   physical = low: no disease (3.0)
brainlobe = medium
|   psychology = high: mood disorder (5.0)
|   psychology = medium: epilepsy (3.0/1.0)
|   psychology = low: epilepsy (4.0)

Number of Leaves  :     9

Size of the tree :      13
```

Fig 3.2: Pruned tree generated by J48

Rule 1: If brain lobe (BL) symptoms and cognitive (COG) symptoms are high, Then disease is "dementia".

Rule 2: If brain lobe (BL) symptoms are high and cognitive (COG) symptoms are medium/low, Then disease is "SI".

Rule 3: If brain lobe (BL) symptoms are low and physical (PHY) symptoms are medium/low, Then there is no disease.

Rule 4: If brain lobe (BL) symptoms are low and physical (PHY) symptoms are high, Then disease is "ADHD".

Rule 5: If brain lobe (BL) symptoms are medium and psychological (PSY) symptoms are high, Then disease is "mood disorder".

Rule 6: If brain lobe (BL) symptoms are medium and psychological (PSY) symptoms are medium/low, Then disease is "epilepsy"

Fig 3.3: Rules generated by Pruned tree

3.2.2 Decision tree: J48 algorithm generate the decision tree as shown in fig 3.4

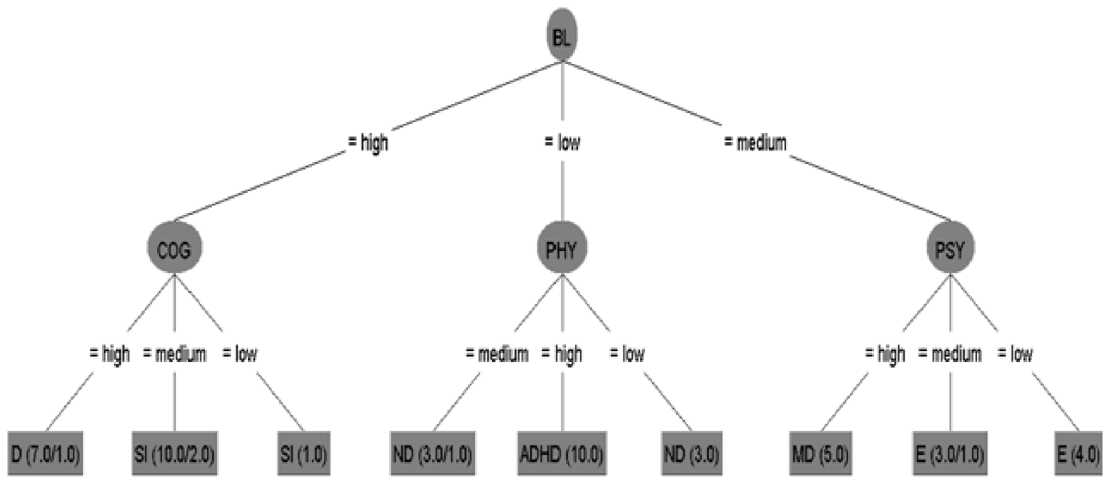


Fig 3.4: Decision tree generated by J48

3.2.3 Confusion matrix: Different attribute have been chosen randomly from the dataset. J48 is applied on the dataset and a confusion matrix is generated as shown in Fig 3.5.

```
=== Confusion Matrix ===

 a  b  c  d  e  f   <-- classified as
 6  0  0  0  0  0 |  a = dementia
 1  9  0  0  0  0 |  b = SI
 0  0 10  0  0  0 |  c = ADHD
 0  0  0  5  1  0 |  d = no disease
 0  0  0  1  6  0 |  e = epilepsy
 0  2  0  0  0  5 |  f = mood disorder
```

Fig 3.5: Confusion matrix generated by J48

From above confusion matrix, true positive for class a =”dementia” is 6 while false positive values are 0, 0, 0, 0, 0. Whereas for class b =”SI”, true positive is 9 while false positive values are 1, 0, 0, 0, 0. For class c =”ADHD”, true positive is 10 while false positive values are 0, 0,0,0,0. For class d =”no disease”, true positive is 5 while false positive values are 0, 0, 0, 1, 0. For class e =”epilepsy”, true positive is 6 while false positive values are 0, 0, 0, 1, 0. For class f =”mood disorder”, true positive is 5 while false positive values are 0, 2, 0, 0, 0. Now simply diagonal elements of the confusion matrix represents the true positive (TP) values and the rest of the elements represent the false positive (FP) values (www.opentox.org). Different operative characteristics values generated by confusion matrix are defined as follows:

True positive (TP) = when test outcome is positive and condition is positive.

False positive (FP) = when test outcome is positive and condition is negative.

True negative (TN) = when test outcome is negative and condition is negative.

False negative (FN) = when test outcome is negative and condition is positive.

Table 3.2: TPR and FPR values

For Class	Sensitivity (TPR)	FPR	Specificity (1-FPR)
A	1	0.025	0.975
B	0.9	0.056	0.944
C	1	0	1
D	0.833	0.025	0.975
E	0.857	0.026	0.974
F	0.714	0	1

The ROC curve obtained from J48 (WEKA tool) are shown in Fig 3.5, 3.6, 3.7, 3.8 and 3.9 for attention deficit hyperactivity disorder (ADHD), dementia (D), epilepsy(E) , mood disorder (MD) and schizophrenia (SI). respectively. Comparing two operating characteristic that are true positive rate and false positive rate. Receiver operating characteristics curve (ROC) and area under curve (AUC) are calculated by ThresholdCurve class in WEKA. This tool is used in many fields for diagnostic test evaluation.

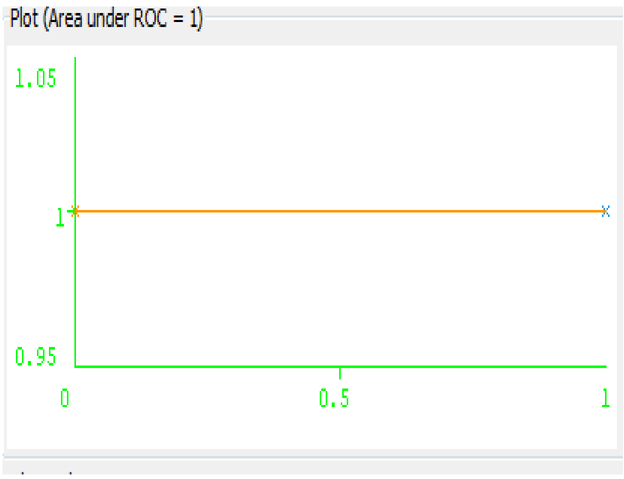


Fig 3.6: Class value ADHD

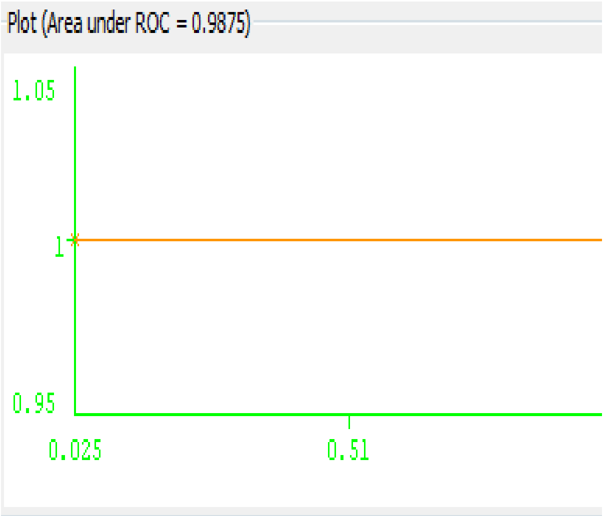


Fig 3.7: Class value D

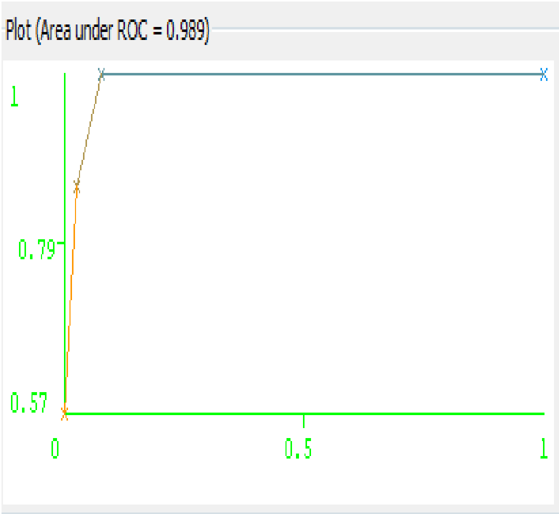


Fig 3.8: Class value E

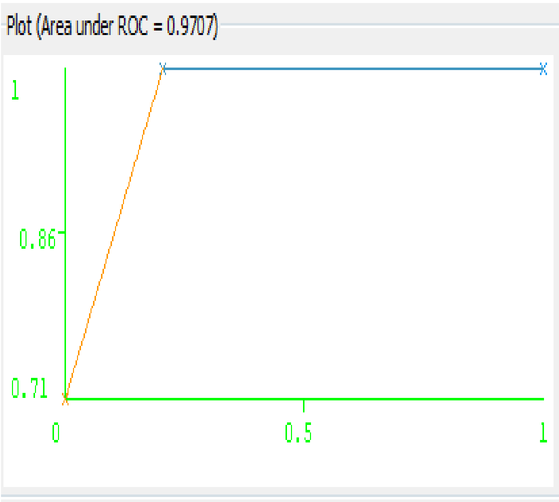


Fig 3.9: Class value MD

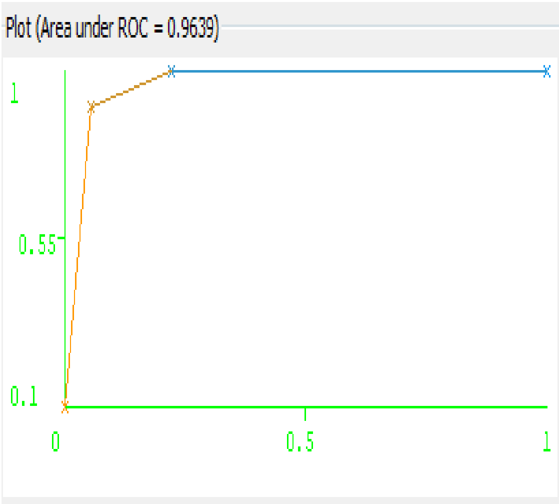


Fig 3.10: Class value SI

3.3 Conclusion

In this chapter, J48 algorithm is deployed for the classification of EEG based diseases such as: ADHD, dementia, epilepsy, mood disorder and Schizophrenia. The data is collected from hospitals and experts in the categorical form. Data Mining technique of AI is applied on the dataset using WEKA tool. J48 algorithm generates rules, confusion matrix, pruned tree and decision tree for the classification of EEG based diseases. The sensitivity for all diseases lies in the range of 70-100 % and specificity lies in the range of 94-100%. J48 algorithm in WEKA generate decision tree that help us to diagnose the disease on the basis of taken attributes/symptoms. And the ROC curve compares the two operative characteristic that are true positive rate and false positive rate.

Chapter 4

Diagnosis of EEG based Diseases using Data Mining , Case Based Reasoning and artificial neural network

This chapter focus on the development of the Medical Diagnosis System (MDS) for the diagnosis of EEG based diseases, integrating J48 (data mining), CBR and ANN. These integrated systems reduce the error amount and degree of uncertainty. Brain is the bioelectric generator. With neurological disordering there occur some disturbances in the brain that lead some problems like muscle weakness, brain functioning Firstly, J48 algorithm (WEKA) is used for reducing the dimension of parameters. After that Case Based Reasoning is implemented for diagnosis of the different EEG based diseases. ANN is used for retrieving the new cases in CBR. The integration of J48, CBR and ANN improves the accuracy of diagnosis, accuracy of retrieval and solve the problem of knowledge acquisition.

4.1 EEG based diseases

Table 4.1 describe EEG based diseases such as: attention deficit hyperactivity disorder (ADHD), dementia (DM), coma (c), epilepsy (EP), migraine (MG), mood disorder (MD) and Schizophrenia (SI). The classifications of different diseases are based on psychological (PSY), cognitive (COG), physical (PHY) and EEG signal characteristics. 1) PSY parameter: anxiety (AN), abnormal behaviour (AB), fear (FR), hyper activity (HA), stress (ST), hallucination (HL), agitation (AG). 2) COG parameter: confusion in decision making (CD), unconsciousness (UC), learning disability (LD), visual disability (VD), judgment (JD), speech disability (SD), unconsciousness (UC), memory loss (ML). 3) PHY parameters: problem in walking (WL), climbing (CL), hearing (HR), vision (VS), speech (SH), oversleeping (OS) (Gangwar et al., 2012).

EEG signal characteristic is divided into two parts :1) EEG signal abnormality observed in different brain lobe (BL) such as frontal lobe (FL), parietal lobe (PL), occipital lobe (OL), temporal lobe (TL). 2) Changing behaviour of the different EEG waves (W) that are alpha (AL), beta (BA), theta (TH) and delta (DE) (Jaimchariyatam et al., 2011 and Ar. et al., 2011). All the sign and symptoms of diseases are collected from literature review and experts suggestions.

Here the symbols H, M, L are used as shown in Table 4.1. H=3 is high that represents these symptoms are presents at high rate. M=2 is medium that represents symptom is presents but not very much and L=1 is low that represents the absence of the symptoms.

Table 4.1: Sign and Symptoms of EEG based diseases

DISEASE	**Physio-Psycho Parameters**																				**Signal**							
	Psychological							**Cognitive**							**Physical**						**EEG signal abnormality**				**EEG waves**			
	AN	**AB**	**FR**	**HA**	**ST**	**HL**	**AG**	**CD**	**LD**	**VD**	**JD**	**SD**	**UC**	**ML**	**WL**	**CL**	**HR**	**VS**	**SH**	**OS**	**FL**	**PL**	**OL**	**TL**	**D**	**T**	**A**	**B**
ADHD	L	H	L	H	L	L	L	L	H	M	L	L	L	L	L	H	H	M	L	L	H	L	L	L	M	H	L	L
Coma	L	L	L	L	L	L	L	H	H	H	H	H	H	H	H	H	M	M	L	M	H	M	H	H	L	L	H	M
Dementia	L	M	L	L	L	H	H	H	H	M	H	M	L	H	L	L	M	M	L	H	H	M	H	H	H	H	L	L
Epilepsy	H	L	H	L	M	L	L	H	M	M	L	M	M	M	H	H	M	M	H	L	H	L	L	H	M	M	M	M
Migraine	H	H	H	L	H	L	M	L	L	M	L	L	L	L	M	M	L	M	L	H	M	M	H	M	M	H	H	L
Mood disorder	H	L	L	H	H	L	L	H	H	L	L	L	L	L	L	L	L	M	L	H	H	L	M	L	L	L	H	H
Schizophrenia	L	M	L	L	L	M	L	L	L	L	L	H	L	H	L	L	L	M	L	H	H	M	M	M	L	L	M	M

4.2 Material

Data which is used for the experimental work is collected from hospitals and experts. Dataset consist of 100 records and 5 attributes. All attributes are of categorical data type.70 records are used for training and rest 30 is used for testing. Detailed description of dataset is given in following Table 4.2.

Table 4.2: Detail of dataset for seven diseases

Attributes	Data type
Psychological	Categorical((high ,medium ,low)
Cognitive	Categorical(high, medium, low)
Physical	Categorical(high, medium, low)
Brain lobes	Categorical(high ,medium, low)
Brain waves	Categorical(high ,medium, low)

4.3 Implementation

The system is implemented in two phases: firstly, J48 algorithm is used for feature selection which results into reduction of dimensions. Five parameters psychological, cognitive, physical, brain lobes and brain waves are given as the input to J48. J48 algorithms select the important parameters and skip the least important parameter by using Information Gain Process. Than the reduced set of parameters are used to classify the different diseases by applying CBR.

4.3.1 Feature selection using J48 algorithm

J48 algorithm generates decision tree, pruned tree and confusion matrix. It selects the psychological, cognitive, brain lobes and brain waves parameters and omits the physical parameters as shown fig 4.2, fig 4.3, fig 4.4 and fig 4.5.

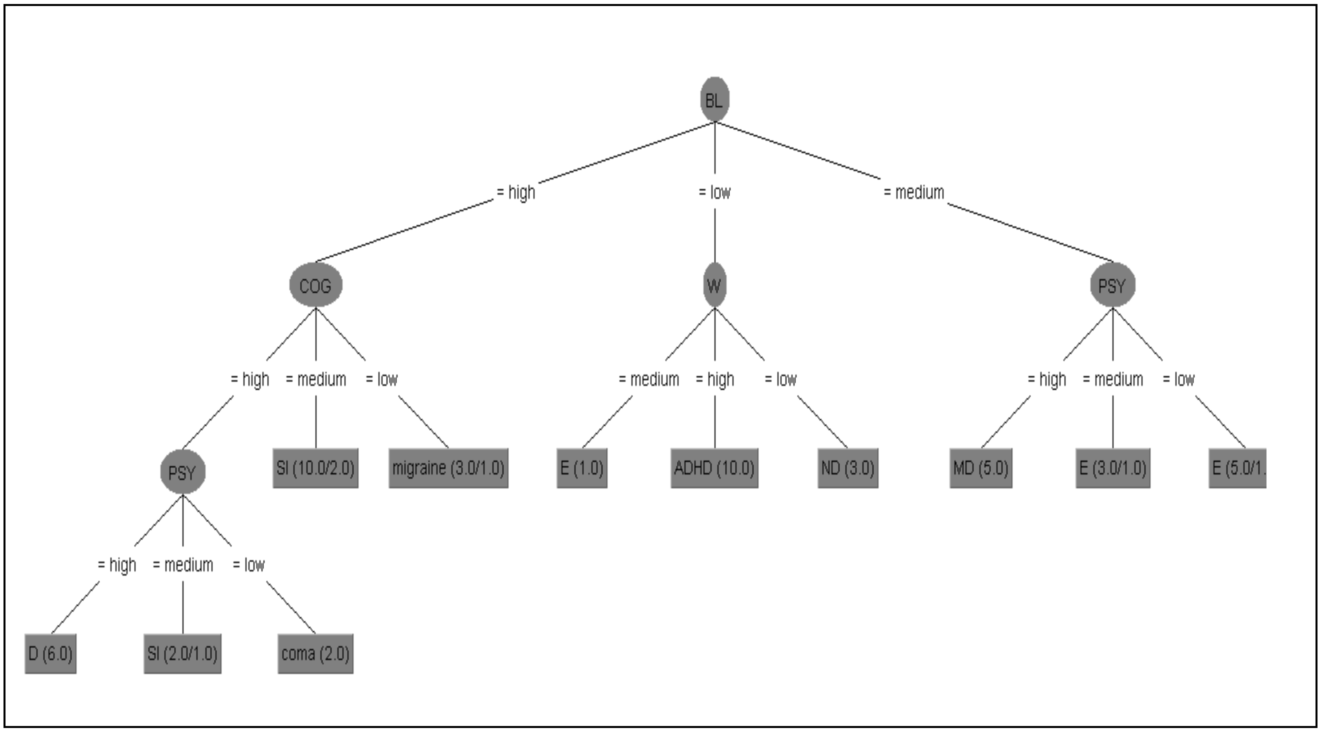


Fig 4.1: Decision tree generate by J48

```
J48 pruned tree
------------------

BL = high
|   COG = high
|   |   PSY = high: D (6.0)
|   |   PSY = medium: SI (2.0/1.0)
|   |   PSY = low: coma (2.0)
|   COG = medium: SI (10.0/2.0)
|   COG = low: migraine (3.0/1.0)
BL = low
|   W = medium: E (1.0)
|   W = high: ADHD (10.0)
|   W = low: ND (3.0)
BL = medium
|   PSY = high: MD (5.0)
|   PSY = medium: E (3.0/1.0)
|   PSY = low: E (5.0/1.0)

Number of Leaves  :     11

Size of the tree :      16
```

Fig 4.2: Pruned tree generated by J48

a	b	c	d	e	f	g	h		<-- classified as
5	0	0	0	0	0	0	1	|	a = D
2	7	0	0	0	0	0	1	|	b = SI
0	0	10	0	0	0	0	0	|	c = ADHD
0	0	0	3	0	0	0	0	|	d = ND
0	0	1	0	6	0	0	0	|	e = E
0	2	1	0	0	4	0	0	|	f = MD
0	1	1	0	0	0	2	0	|	g = coma
0	2	0	0	1	0	0	0	|	h = migraine

Fig 4.3: Confusion Matrix generated by J48

Rule 1: If brain lobe (BL) symptom , cognitive (COG) symptom is high and Psychological (PSY) symptom is high, Then disease is "dementia".

Rule 2: If brain lobe (BL) symptom, cognitive (COG) symptom is high and Psychological (PSY) symptom is medium, Then disease is "schizophrenia".

Rule 3: If brain lobe (BL) symptom, cognitive (COG) symptom is high and Psychological (PSY) symptom is low, Then disease is "coma".

Rule 4: If brain lobe (BL) symptom is high and cognitive (COG) symptom is medium, Then disease is "schizophrenia".

Rule 5: If brain lobe (BL) symptom is high and cognitive (COG) symptom is low, Then disease is "migraine".

Rule 6: If brain lobe (BL) symptom is low and wave (W) symptom is medium, Then disease is "epilepsy".

Rule 7: If brain lobe (BL) symptom is low and wave (W) symptom is high, Then disease is "ADHD".

Rule 8: If brain lobe (BL) symptom is low and wave (W) symptom is low, Then there is no disease.

Rule 9: If brain lobe (BL) symptom is medium and psychological (PSY) symptom is high, Then disease is "mood disorder".

Rule 10: If brain lobe (BL) symptom is medium and psychological (PSY) symptom is medium, Then disease is "epilepsy".

Rule 11: If brain lobe (BL) symptom is medium and psychological (PSY) symptom is low, Then disease is "epilepsy".

Fig 4.4: Rules generated by J48

4.3.2 Case Based Reasoning

In this work CBR is implemented in following phases: Knowledge acquisition, case retrieval, matching and adaption.

4.3.2.1 Knowledge Acquisition

Knowledge is acquired through dialogue session by asking question answer as shown below:

Physio-Psycho Parameters

Physio-Psycho consist psychological, cognitive and physical parameters. Physical parameters are skipped by the J48 algorithm.

Psychological parameter

Does the patient have felling of anger?	H: [AN=3], M: [AN=2], L: [AN=1]
Does the patient have abnormal behaviour	H: [AB=3], M: [AB=2], L: [AB=1]
Does the patient suffering from hallucination?	H: [HL=3], M: [HL=2], L: [HL=1]
Does the patient have felling of agitation?	H: [AG=3], M: [AG=2], L: [AG=1]
Does the patient feel fear?	H: [FR=3], M: [FR=2], L: [FR=1]
Does the patient is hyperactive?	H: [HA=3], M: [HA=2], L: [HA=1]
Does the patient feels stressed?	H: [ST=3], M: [ST=2], L: [ST=1]

Cognitive parameter

Does the patient have confusion in decision making?	H: [CD=3], M: [CD=2], L: [CD=1]
Does the patient have learning disability?	H: [LD=3], M: [LD=2], L: [LD=1]
Does the patient have visual disability?	H: [VD=3], M: [VD=2, L: [VD=1]
Does the patient have confusion in Judgment?	H: [JD=3], M: [JD=2], L: [JD=1]
Does the patient have speech disability?	H: [SD=3], M: [SD=2], L: [SD=1]
Does the patient have problem of unconsciousness?	H: [UC=3], M: [UC=2], L: [UC=1]
Does the patient have problem of memory loss?	H: [ML=3],M: [ML=2], L: [ML=1]

EEG signal abnormality

Is abnormality observed in FL?	H: [FL=3], M: [FL=2], L: [FL=1]
Is abnormality observed in PL?	H: [PL=3], M: [PL=2], L: [PL=1]
Is abnormality observed in OL?	H: [OL=3], M: [OL=2], L: [OL=1]
Is abnormality observed in TL?	H: [TL=3], M: [TL=2], L: [TL=1]

EEG Waves	
Is there change in Delta waves?	H: [Delta=3], M: [Delta=2], L: [Delta=1]
Is there change in Theta waves?	H: [Theta=3], M: [Theta=2], L: [Theta=1]
Is there change in Alpha waves?	H: [Alpha=3], M: [Alpha=2], L: [Alpha=1]
Is there change in Beta waves?	H: [Beta=3], M: [Beta=2], L: [Beta=1]

4.3.2.2 Case Retrieval

Retrieval is the most important stage in the CBR process. For solving the new problem, the problem matched against the stored cases in the case base and similar cases are retrieved.Cases are retrieved using matching and selection. Matching is done using following similarity measure:

$$S=\sum_{i=1}^{n} w_i \mid U_i-C_i \mid \qquad \text{eq.4}$$

Where U_i is ith feature of user case, C_i is ith feature of case of case base, W_i is weight assign to different parameter by domain experts.

Selection is done using the following formula:

$$R=\min_{i=1}^{n} (S_i) \qquad \text{eq.5}$$

The case with the smallest S_i value is the closest case.

4.3.2.3 Case Adaption

Case from the case base is selected which is more similar to the user case, and the solution is used without any modification. Once the problem is solved the result will be retained into case base.

4.4 Retrieval through ANN

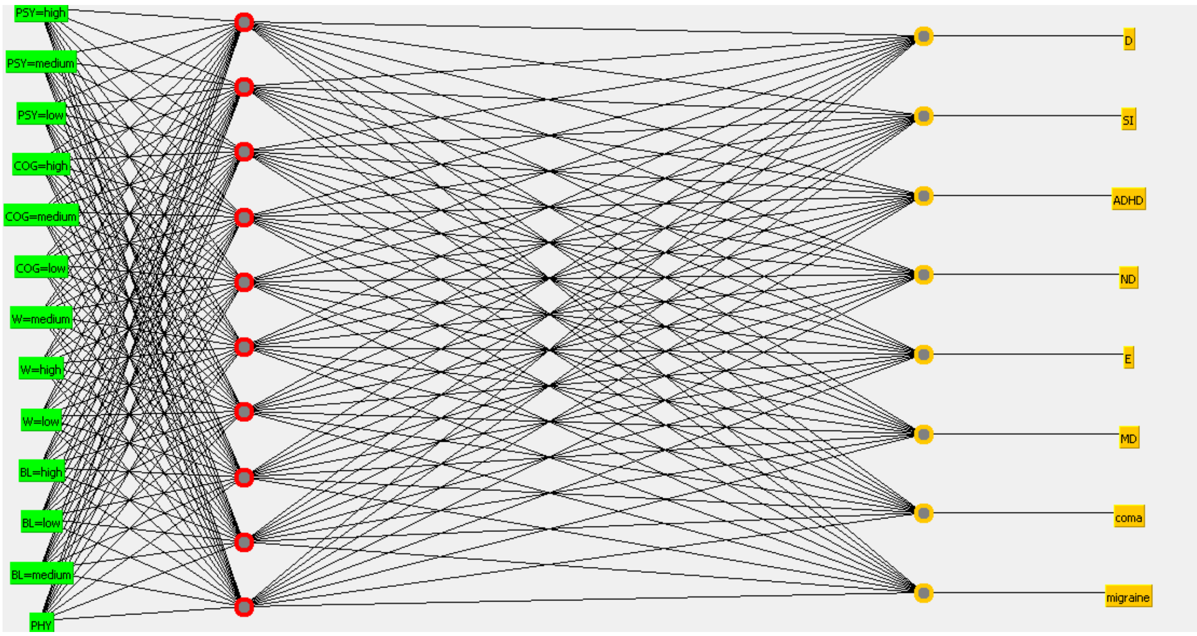


Fig 4.5: Retrieval through ANN in WEKA

Fig 4.5 shows how the ANN helps to diagnose the diseases. Four layers are shown as one input two hidden and one output layer. The input layer contain 13 neurons first hidden layer contains 10 neurons, second hidden layer contains 8 neuron and the output layer also contain 8 neurons. ANN correctly classified 94% and incorrectly classified 6%. Number of echoes generated in fig 4.5 is 500. Error per epoch is 0.0115692. Learning rate is 0.3 and momentum is 0.2

4.5 Result

EEG based diseases classified into seven types such as: ADHD, Coma, Dementia, Epilepsy, Migraine, Mood disorder and Schizophrenia. Five new user cases with their classification are shown in following table 4.3.

Performance analysis of J48-CBR-ANN integration is done by calculating the sensitivity and specificity values. Sensitivity is the true positive rate value and specificity is 1-false positive rate value. For calculating the accuracy of this integration method, 25 cases with different EEG based diseases had been tested. The sensitivity and specificity values are shown in table 6. It is observed from table 4.6 that the integration of J48-CBR-ANN provides more accurate result than J48 alone.

Table 4.3: Case-base for diagnosis of new cases

New Case	Features of case	Disease
Case1	{L,M,M,L,L,M,H,H,M,H,H,M,H,H,H,H,M,L,L,H,H}	Dementia
Case2	{L,M,M,L,L,M,L,M,L,L,M,H,M,M,H,M,M,M,H,L,L}	Schizophrenia
Case3	{L,H,L,H,M,M,H,L,H,M,L,L,L,L,M,M,L,L,L,H,M}	ADHD
Case4	{L,M,M,M,L,M,H,H,H,M,H,M,H,H,M,M,H,L,M,H,H}	Dementia
Case5	{L,M,L,M,L,M,L,M,L,L,L,H,H,M,H,M,M,M,M,M,L}	Schizophrenia

Table 4.4: Performance analysis of J48, CBR and ANN

Method	Sensitivity	Specificity	Accuracy
J48	88%	92%	92.5%
J48-CBR-ANN	95%	97%	96%

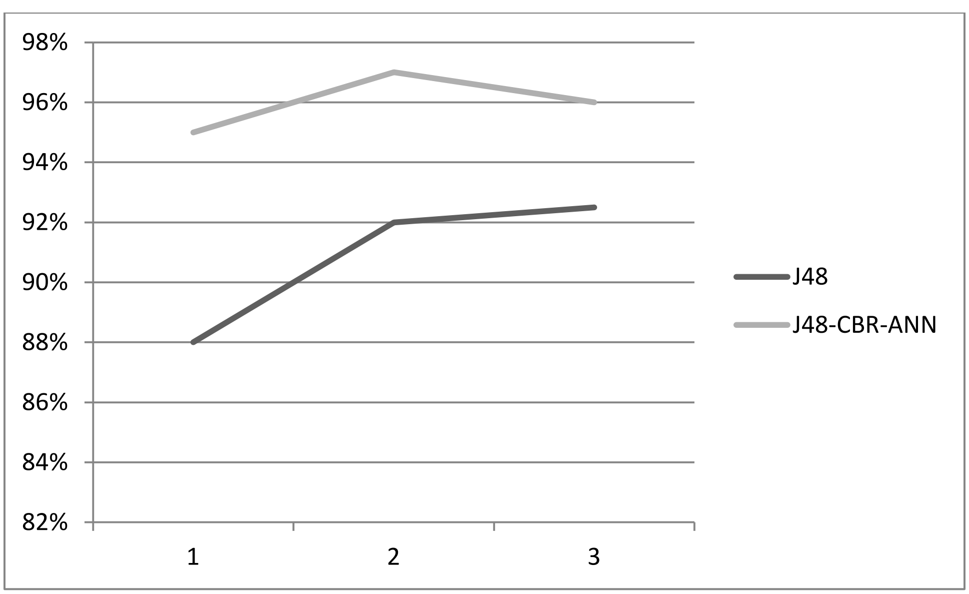


Fig 4.6: Comparative analysis of J48 and J48-CBR-ANN integrated method

As it clear from the fig 4.6, integration of J48-CBR-ANN overcome the problem that are faces by using alone J48 such as problem of interpretability and updatability. J48-CBR-ANN provides fast knowledge acquisition and system get self update.

4.6 Conclusion

In this chapter integration of J48, CBR and ANN method is deployed for the diagnosis of EEG based diseases. In the first step J48 is used for feature selection that reduces the dimension of attributes in CBR. CBR is used for diagnosis of different EEG based diseases by matching the user case data with the past cases that are stored in the knowledge base of CBR. ANN is used for retrieving the new cases in CBR. Integration of the J48, CBR and ANN method has been proven to be appropriate and convenient for diagnosis of EEG based diseases.

Chapter 5

Result and Discussion

5.1 Result and Discussion

This chapter compares the result obtained by single intelligent method such as: J48 and integrated RBR-CBR and J48-CBR-ANN method.

RBR and CBR both are kind of knowledge oriented techniques. These methods have their own advantages and disadvantages. The advantages of RBR are modularity, uniformity, separation of knowledge and content and squeezed representation of knowledge (reduce the dimension) but it lacks in fast knowledge acquisition Pandey and Mishra, (2009). On the other hand CBR is most efficient in knowledge acquisition and it is self updatable. As the new cases entered in the system it stores the new cases as the cases of case base. The integration of RBR-CBR overcomes disadvantages of each other and improves the system. The integrating of RBR-CBR improves knowledge acquisition.

J48 algorithm classifies the disease by generating the decision tree. J48 has certain advantages like it accept the nominal and continuous values and if there is missing data in the dataset it also accept that (d).

J48-CBR-ANN overcome the problem that are faces by using alone J48 such as problem of interpretability and updatability. J48-CBR-ANN provides fast knowledge acquisition and system get self update.

Table 5.1: Performance analysis of all methods

Method	Sensitivity	Specificity	Accuracy
RBR-CBR	85%	89%	87%
J48	88%	92%	92.5%
J48-CBR-ANN	95%	97%	96%

Table 5.1 shows the performance analysis of all the methods by measuring sensitivity (true positive rate), specificity (false positive rate) and accuracy. From table it is clear integration of J48-CBR-ANN provide most accurate result.

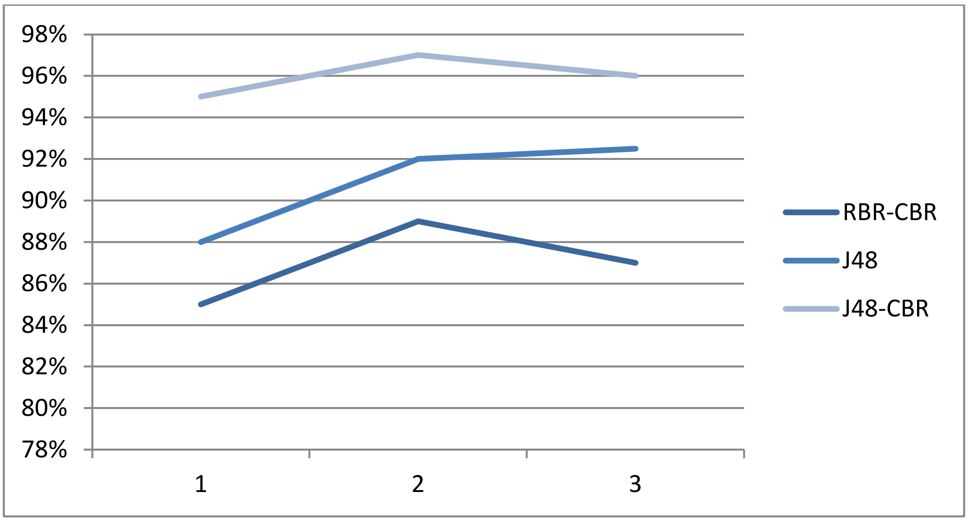


Fig 5.1: Performance analysis of all methods

Fig 5.1 shows the pictorial representation of performance analysis of all the methods.

5.2 Conclusion

This chapter concludes that integration of J48-CBR-ANN provide more accurate result. As all the AI techniques have their own advantages. But when these techniques get integrated with each other they enhance the performance because these techniques are complement to each other. If only J48 algorithm is used for diagnosis. It lacks the knowledge acquisition but if J48 integrate with CBR it overcome this problem.J48 improves the performance by feature selection, ANN improve the retrieval process and case base acquire the knowledge.

Chapter 6

Conclusion

6.1 Conclusion

Chapter 2, develop integrated method based on RBR and CBR for the diagnosis of EEG based diseases. Initially, collect the different symptoms based on physical, psychological, cognitive and EEG based parameter. Then the Rule based approach used for defining rules and correlates the sign and symptoms with different disease. It helps to generate and reduce the dimension of the cases in a case base. Case base reasoning is used to diagnose the disease. The reduced dimension of cases in a case base improves the retrieval efficiency.

Chapter 3 deployed J48 algorithm for the classification of EEG based diseases such as: ADHD, dementia, epilepsy, mood disorder and Schizophrenia. The data is collected from hospitals and experts in the categorical form. Data Mining technique of AI is applied on the dataset using WEKA tool. J48 algorithm generates rules, confusion matrix, pruned tree and decision tree for the classification of EEG based diseases. The sensitivity for all diseases lies in the range of 70-100 % and specificity lies in the range of 94-100%. J48 algorithm in WEKA generate decision tree that help us to diagnose the disease on the basis of taken attributes/symptoms. And the ROC curve compares the two operative characteristic that are true positive rate and false positive rate.

Chapter4, developed integrated method based on J48, CBR and ANN. In the first step J48 is used for feature selection that reduces the dimension of attributes in CBR. CBR is used for diagnosis of different EEG based diseases by matching the user case data with the past cases that are stored in the knowledge base of CBR. ANN is used for retrieving the new cases in CBR. Integration of the J48, CBR and ANN method has been proven to be appropriate and convenient for diagnosis of EEG based diseases.

Bibliography

Deepika Kundra is a Research Scholar in the Department of Computer Science and Application, Lovely Professional University, Punjab, India. She is conducting research in the field of Artificial Intelligent. She has published two research paper one in the peer reviewed journals of international repute and second in conference proceedings. She has been planning to conduct further research in interdisciplinary field to design new software for developing emerging technologies for the welfare of mankind.

Index

ADHD: Attention Deficit Hyperactivity Disorder

AB : Abnormal Behaviour

AG : Agitation

AI : Artificial Intelligence

AL : Alpha

AN : Anger

ANN : Artificial Intelligence

BE : Beta

BL : Brain Lobe

BW : Brain Wave

C : Coma

CBR : Case Bases Reasoning

CD : Confusion in Decision making

CL : Climbing

COG : Cognitive

D : Dementia

DM : Data Mining

E : Epilepsy

EEG : Electroencephalography

FL : Frontal Lobe

FN : False Negative

FP : False Positive

FPR : False Positive Rate

FR : Fear

H : High

HA : Hyper Activity

HL : Hallucination

HR : Hearing

JD : Judgment

L : Low

LD : Learning Disability

M : Medium

MD : Mood Disorder

MG : Migraine

ND : Neurological Disorder
OS : Oversleep
PHY : Physical
PI : Psychological Illness
PP : Psycho-Physical
PSY : Psychological
RBR : Rule Based Reasoning
ROC : Receiver Operating Characteristics
SD : Speech Disability
SF : Specificity
SH : Speech
SI : Schizophrenia
SN : Sensitivity
ST : Stress
TN : True Negative
TPR : True Positive Rate
TP : True Positive
UC : Unconsciousness
VS : Vision
VD : Visual Disability
WEKA: Waikato Environment for Knowledge Analysis
WL : Walking

Appendices

INTERFACE

The system is designed in visual studio 2008 express edition. At front end ASP.net with c# (c sharp) and at back end SQL (structured query language) server is used.

Step 1: Back End /Data Base

In 'EEG '(name of data base) 'DISEASES' table is made.100 record are entered as the training data set with 7 psychological, 6 cognitive, 4 brain lobe and 4 brain waves parameters shown in the following table.

EEG: Query(hp-hp\...\DATABASE.MDF) | Start Page | Default.aspx

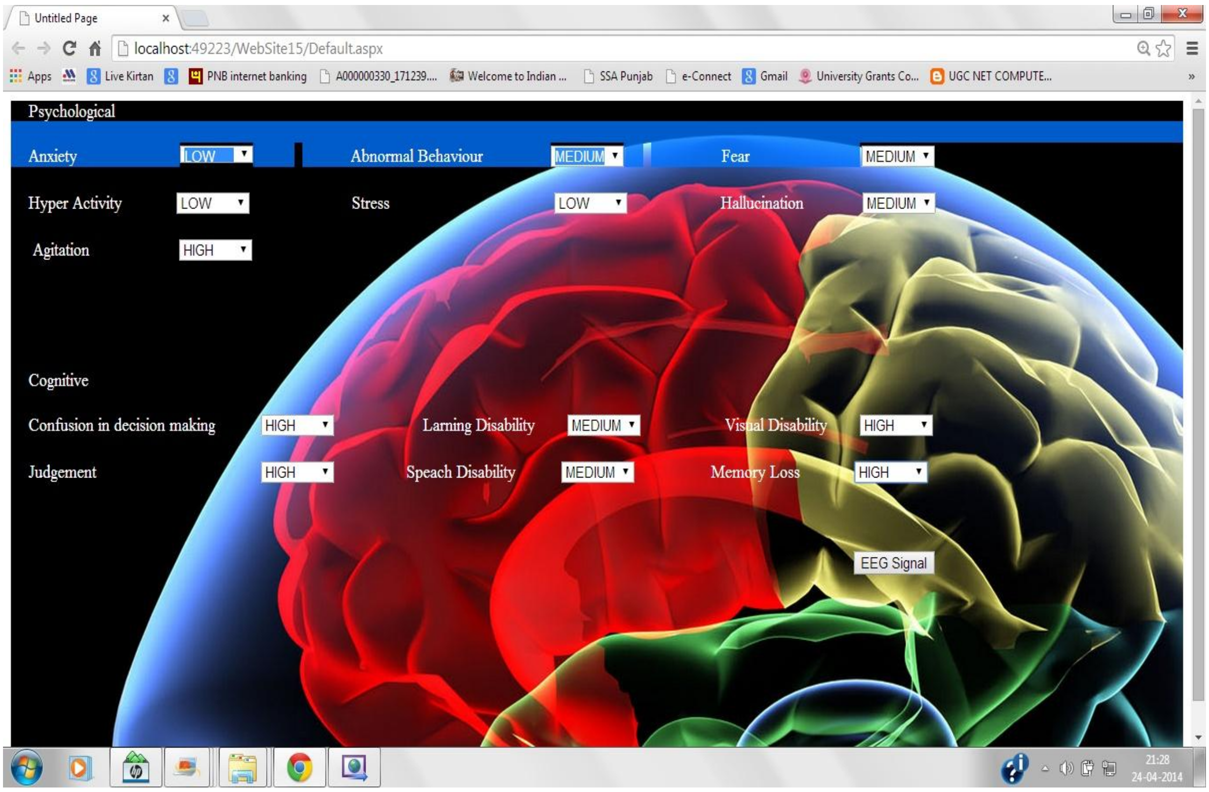
ABNORMAL_B...	FEAR	HYPER_ACTIVI...	STRESS	HALLUCINATI...	AGITATION	CONFUSION_I...	LEARNING_DIS...	VISUAL_DISABI...	JUDGEME
MEDIUM	LOW	LOW	LOW	HIGH	HIGH	HIGH	HIGH	MEDIUM	HIGH
MEDIUM	LOW	LOW	LOW	MEDIUM	LOW	LOW	LOW	LOW	LOW
HIGH	LOW	HIGH	LOW	LOW	LOW	LOW	HIGH	MEDIUM	LOW
LOW	HIGH	LOW	MEDIUM	LOW	LOW	HIGH	MEDIUM	MEDIUM	LOW
LOW	LOW	HIGH	HIGH	LOW	LOW	HIGH	HIGH	LOW	LOW
MEDIUM	MEDIUM	LOW	LOW	MEDIUM	HIGH	HIGH	MEDIUM	HIGH	HIGH
MEDIUM	LOW	MEDIUM	LOW	HIGH	HIGH	HIGH	HIGH	MEDIUM	HIGH
MEDIUM	MEDIUM	LOW	LOW	MEDIUM	LOW	LOW	MEDIUM	MEDIUM	LOW
MEDIUM	LOW	LOW	LOW	MEDIUM	LOW	MEDIUM	LOW	MEDIUM	LOW
HIGH	LOW	HIGH	MEDIUM	MEDIUM	HIGH	LOW	HIGH	MEDIUM	LOW
MEDIUM	LOW	LOW	MEDIUM	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH
MEDIUM	MEDIUM	MEDIUM	LOW	HIGH	HIGH	HIGH	HIGH	HIGH	HIGH
MEDIUM	LOW	LOW	LOW	MEDIUM	HIGH	MEDIUM	HIGH	HIGH	HIGH
MEDIUM	MEDIUM	MEDIUM	LOW	MEDIUM	HIGH	HIGH	HIGH	MEDIUM	HIGH
MEDIUM	LOW	MEDIUM	MEDIUM	HIGH	HIGH	HIGH	MEDIUM	MEDIUM	HIGH
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL

Step 2: Front End

Two pages are designed in front end.

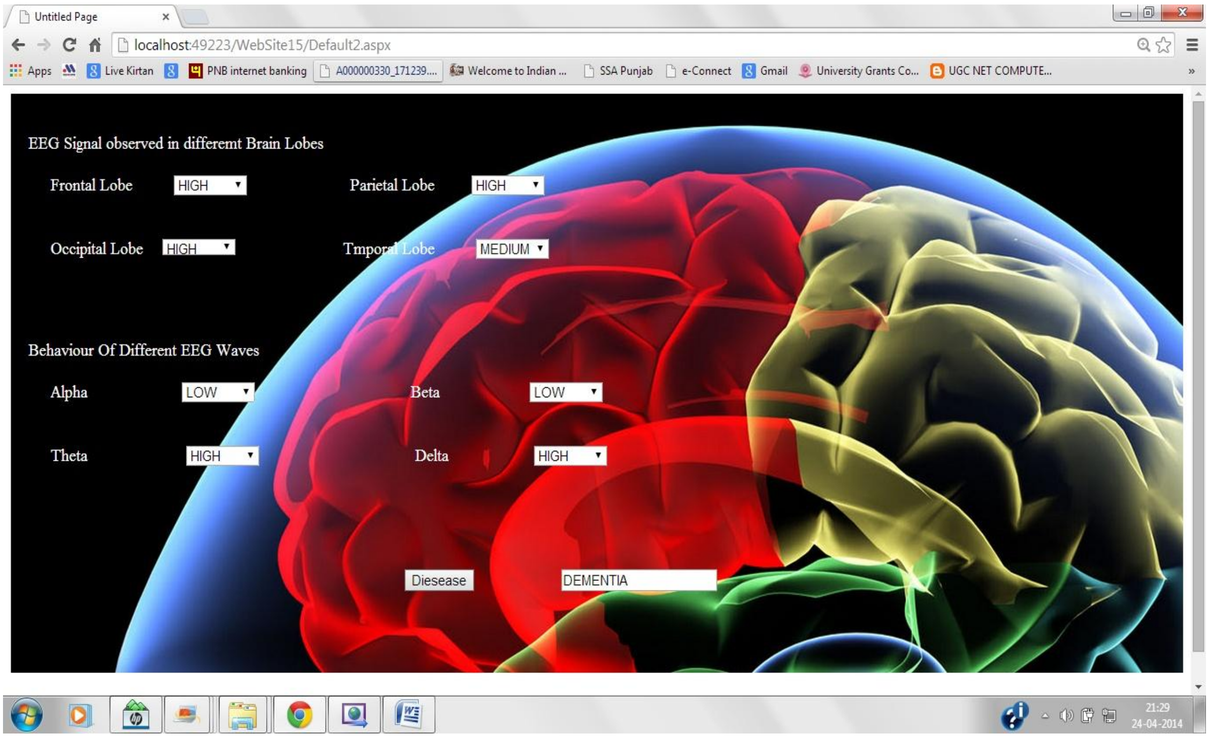
First page:

First page contains all the psychological and cognitive parameters. Users have to enter the values of all the symptoms in form of HIGH, MEDIUM and LOW by selecting from the respective dropdown list. After filling all values press the EEG PARAMETER named button that reside on the right bottom corner of the screen. By clicking on that users will redirect to next page.



Second page:

Second page contains all the EEG based parameters that are Brain lobes and Brain waves. Similar to first page Users have to enter the values of all the symptoms in form of HIGH, MEDIUM and LOW by selecting from the respective dropdown list. After filling all values press the DISEASE button. Then the diagnosed diseases will be appeared in the textbox.



Now, if the user case value is similar to the any case of case base it will simply diagnose the diseases. If not, then the system will select the more similar case stored in the data base as the resulted diseases and that new user case will be entered as a new case in the case base.

Publications

1) Kundra and Pandey ,“Diagnosis of EEG Based Disease Using Rule Based Reasoning and Case Based reasoning”, Proc of national conference of advance science and technology (NCAST) on January 25, 2014, in Gorakhpur.

2) Kundra and Pandey ,”Classification of EEG based Diseases using Data Mining “, International Journal of Computer Applications (0975 – 8887) Volume 90 – No 18, March 2014. Pp-11-15.

3) Kundra and Pandey,” Diagnosis of EEG based Diseases using Data Mining and Case Based Reasoning” is communicate in International Journal of System Design and Computing (Inder Science).

References

Aamodt,Plaza, “case based reasoning: foundational issues, methodological variation, and system approaches”, *AI communication* 7(1) (1994).pp. 39-59.

Abasolo et al.,”electroencephalogram background activity characterization with detrended moving average in Alzheimer’s disease patient”, Proc. of 6th IEEE international signal processing ,pp. 211-215.

Ar. et al., “Behavioural differences between EEG-defined subgroups of children with Attention-Deficit/Hyperactivity Disorder”, Clin Neurophysiology. 2011 Jul; 122(7): pp. 1333-41.

Azim et al., “Analysis of EEG and EMG Signals for Detection of Sleep Disordered Breathing Events”, Proc. of 2010 IEEE, pp. 646-649.

Bharhava et al., “Decision Tree Analysis on J48 Algorithm for Data Mining”, International Journal of Advanced Research in Computer Science and Software Engineering, volume 3, issue 6 ,june 2013.

Cabrera and Edye. “Integration of Rule Based Expert Systems and Case Based Reasoning in an Acute Bacterial Meningitis Clinical Decision Support System”, *(IJCSIS)* International Journal of Computer Science and Information Security, Vol. 7, No. 2, 2010, pp.112-118.

Cohen et al., “Using Clinical Information and Nonlinear EEG Analysis for Diagnosis of Dementia”, Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE Vol 3 . PP. 2299-2302.

C.R. et al., “Brain machine interface: classification of mental tasks using short-time pca and recurrent neural networks”, Proc. of International Conference on Intelligent and Advanced Systems 2007 .pp.1153-1156.

Exarchos et al., “A Data Mining based Approach for the EEG Transient Event Detection and Classification”, Proc. of 18th IEEE Symposium on Computer-Based Medical Systems ,2005.

Ganesan et al., “A novel approach for the analysis of epileptic spikes in EEG”, International Conference on Computational Intelligence and Multimedia Applications 2007, pp.297-301.

Gangwar et al., “Intelligent Computing Methods for The Interpretation of Neuropsychiatric Diseases Based on Rbr - Cbr-Ann Integration”, International journal of computers & technology 2013,Vol. 11, pp. 2490-2511.

Gangwar et al., “Intelligent Computing Method for the Interpretation of Neuropsychiatric Diseases”, International Journal of Computer Applications (0975 – 8887) Volume 55– No.17, October 2012, pp. 23-31.

Glover et al., “Knowledge based interpretation of bioelectrical signals: Examples in eeg analysis”, Proc. of the 25th Annual International Conference of the IEEE EMBS Cancun, Mexico September 17-21, 2003.

Gurumurthy and Tripathy, “Intelligent System to Recognize Human Brain Signals for finding Brain Diseases based on Neural Networks”, Proc. of the World Congress on Engineering, Vol I.WCE 2010, June 30 - July 2, 2010, London, U.K.

Harischandra and perera,” Intelligent Emotion Recognition System Using Brain Signals (EEG)”, 2012 IEEE EMBS International Conference on Biomedical Engineering and Sciences | Langkawi | 17th - 19th December 2012 .

Jaimchariyatam et al., "Prevalence and Correlates of Alpha-Delta Sleep in Major Depressive Disorders", Innovation in clinical neuroscience, 2011 july,pp.35-49.

Jansen et al., "Knowledge-Based Approach to Sleep EEG Analysis-A Feasibility Study", IEEE transactions on biomedical engineering, vol. 36, no. 5, MAY 1989**.**

Kumari and jose, "Seizure Detection In EEG Using Time Frequency Analysis and svm", proceedings of *ICETECT* 2011.Pp. 626-630.

Li et al., "An Effective Classification Approach for EEG-based BCI System", Proc. of 2011 *Sixth* International Conference on Image and Graphics.

Pal and Panda, "Classification of EEG signals for epileptic seizure evaluation", Proc. of IEEE Students' Technology Symposium 2010, pp.72-76.

Pandey and Mishra, 'Knowledge and intelligent computing system in medicine', Computers in Biology and Medicine 39 (2009), www.elsevier.com/locate/cbm. pp. 215 -230.

Pandey and Mishra, "Case-based reasoning and data mining integrated method for the diagnosis of some neuromuscular disease", Proc. of IJMEI 2011, pp.1-15.

Pandey and Mishra," An integrated intelligent computing method for the detection and interpretation of ECG based cardiac diseases", Proc. of IJKESDP 2010, pp.182-203.

Pandey and Mishra , "Data mining and CBR integrated methods in medicine: a review", Proc. of IJMEI 2010, pp.205-218.

Patil and Shereker, "Performance analysis of naïve bates and J48 classification algorithm for data classification", International Journal of Computer Science and Applications, Vol. 6, No.2, Apr 2013. Pp. 256-261.

Sheniha et al," Removal of Artifact from EEG Signal using Differential Evolution Algorithm", International conference on Communication and Signal Processing, April 3-5, 2013, India.Pp. 134-138.

Sudirman and, Seow," electroencephalographic based hearing Identification using back-propagation algorithm", Proc. Of IEEE, pp. 991-995.

Sulaiman et al. "system successfully shows the stress level, stress score and index of EEG signals". 2011 Third International Conference on Computational Intelligence, Communication Systems and Networks.Pp.362-367.

Teplan," fundamentals of EEG measurement", volume 2, section 2, 2002.

Vijila et al., 'artifacts removel in EEG signal using adaptive neuro fuzzy inference system", *Proc. of IEEE-ICSCN 2007*, MIT campus, Anna University, Chennai, india.feb.22-24, 2007, pp. 589-591.

Wu et al.," EEG classification based on artificial neural network in brain computer interface", *LSMS/ICSEE 2010*, part 1, CCIS 97, pp.154-162.

Websites

http://www.healthcommunities.com/coma/symptom.shtml

http://www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm

http://www.nlm.nih.gov/medlineplus/neurologicdiseases.html.

http://www.opentox.org/dev/documentation/components/j48.

http://psychology.about.com/od/biopsychology/ss/brainstructure_2.htm

http://serendip.brynmawr.edu/bb/kinser/Structure1.html

http://en.wikipedia.org/wiki/Rule-based_system

http://en.wikipedia.org/wiki/Neurological_disorder

http://en.wikipedia.org/wiki/Extract,_transform,_load

http://www.webmd.com/migraines-headaches/tc/migraine-headaches-symptoms