Detection of Mental Stress Using EEG Signal and Classifier Based on Reward and Punishment Processing

by

Farhana Binta Salim
16301185
Abu Md. Sadat
16301105
Maria Islam Ema
17301180
Anita Mahmud Jhara
19101670

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
January 2021

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Farhana Binta Salim

Farhana Binta Salim
16301185

Abu Md Sadat

Abu Md. Sadat
16301105

Maria Islam Ema

Maria Islam Ema
17301180

Anita Mahmud

Anita Mahmud Jhara
19101670

Approval

The thesis titled "Detection of Mental Stress Using EEG Signal and Classifier Based on Reward and Punishment Processing" submitted by

1. Farhana Binta Salim (16301185)

2. Abu Md. Sadat (16301105)

3. Maria Islam Ema (17301180)

4. Anita Mahmud Jhara (19101670)

Of Fall, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on January 11, 2021.

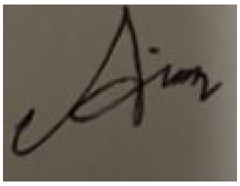
Examining Committee:

Supervisor:
(Member)

Zavid Parvez
Digitally signed by Zavid Parvez
DN: cn=Zavid Parvez, o=Brac University, ou=CSE, email=zavid.parvez@bracu.ac.bd, c=US
Date: 2021.01.15 12:52:46 +06'00'

Dr. Mohammad Zavid Parvez
Assistant Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Non-Member)

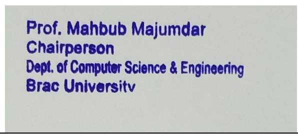


Dr Md Anisur Rahman
Academic
School of Computing and Mathematics
Charles Sturt University

Program Coordinator:
(Member)

Dr. Md. Golam Robiul Alam
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)



Mahbubul Alam Majumdar,PhD
Head of Department
Department of Computer Science and Engineering
Brac University

Abstract

Mental stress is the main well being problem worldwide today. Mental stress is by far the most common, responsible for most of all mental-brain diseases. The high incidence, high impairment and heavy risk of illness make mantel stress a major health issue facing the planet and is predicted to become the most prevalent disorder. Additionally, most of the time, suicidal people also hide true feelings and fail to communicate their psychiatric problems to physicians. Mental stress is a common as well as complex condition that has a limited explanation of its origin. Explanations regarding disease causes are centered both on their psychology and behavioral symptoms. Both factors that have led are currently being related by experiments investigating and how depressive individuals handle reward and punishment. Many studies have been proposed that people with such a disorder will be unable to utilize effective knowledge to guide actions. The specific issues that need to be addressed are, how to find an easy, reliable and realistic way to diagnose mental stress and recognize mental stress early in order to keep it from becoming a serious and irreversible condition. To avoid diseases including health issues, the primary prevention of mental stress utilizing machine learning classifiers based on reward and punishment processing is important. The nervous system of man is the primary subject of mental stress. For all of this purpose, a machine learning framework is applied to evaluate the electroencephalogram signals for fifty individuals in our proposed model. For a successful mental stress detection application, this layout has implemented a combination of features that supplies nine ML classifiers which are Support Vector Machine, Random Forests, K-Nearest Neighbors, Decision Tree classifier, AdaBoost classifier, Extra Trees classifier, Bagging classifier, Gradient Boosting classifier, Gaussian Naïve Bayes Classifier to identify Comorbidity based on reward and punishment processing. The experimental results indicate that Extra Trees classifier, Gaussian Naïve Bayes Classifier and Random Forests have higher success predicting the presence of mental stress. This implemented classifier based on reward and punishment processing systems using EEG signals has the ability to statistically detect mental stress.

Keywords: EEG, Reward and punishment process, Comorbidity, Machine learning, Support Vector Machine, Random Forests, K-Nearest Neighbors, Decision Tree classifier, AdaBoost classifier, Extra Trees classifier, Bagging classifier, Gradient Boosting classifier, Gaussian Naive Bayes Classifier

Dedication (Optional)

We would like to dedicate this thesis to our respectable parents, honorable teachers and loving siblings.

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis have been completed without any major interruption.

Secondly, to our advisor Dr. Mohammad Zavid Parvez Sir and co-advisor Dr Md Anisur Rahman for their kind support and advice in our work. We could not complete our thesis without their direction and constant support.

And finally to our parents and siblings, without their support and care it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

ANN Artificial Neural Network

BAS Behavioral Approach System

BDI Beck Depression Inventory

BIS Behavioral Inhibition System

CT Classification Tree

$dACC$ dorsal Anterior Cingulate Cortex

EEG Electroencephalogram

ICA Independent Component Analysis

KNN k-Nearest Neighbors

$LIWC$ Linguistic Inquiry and Word Count

LR Linear Regression

$MASQ$ Mood and Anxiety Symptoms Questionnaire

ML Machine Learning

OCD Obsessive Compulsive Disorder

$RLCQ$ Validation of the Recent Life Changes Questionnaire

$SPSRQ$ Sensitivity to Reward Questionnaire

SVM Support Vector Machine

TAI Test Anxiety Inventory

Chapter 1

Introduction

1.1 Introduction

Stress is our body's response which is occurred mentally, emotionally and psychologically over the changes that happens in one's external or internal environment resulting in distress, anxiety, depression etc. Mental stress occurs not only because of real situations but also for imaginary situations such as, from one's thoughts or mind-wanderings. According to medical statement, due to mental stress, "fight or flight" response is initiated which is an endocrine and neural reaction [1]. The "fight or flight" is a psychological response which generates automatically in the situation of stress. Moreover, this response prepares the body to fight or flee by triggering a response for mental stress [2]. Mental stress is a thing from which every person is affected. Almost every person has experienced mental stress but depending on their situation, it varies as not all person goes through the same situation along with different situations; and the time duration also varies as some faces mental stress for short time, however, some faces for a long time. The persons who face this stress for long time, it is very dangerous for them. Long-term mental stress harms a person's physical state, psychological stress and moreover, it leads a person to anxiety, depression, irritability, mood swing, memory and concentration problems etc. which harms a person externally and internally.

Mental stress is experienced when people face challenging situations and struggles to cope with the situation which creates a mental pressure. This stress can be positive or negative. Positive stress helps a person in a good way; such as, by motivating to achieve something good, by keeping alert, by preparing to avoid dangerous situations etc. Besides that, the stress can also be negative which hampers one's normal life by physically and psychologically; such as, by feeling threatened, frightened etc. most of the time. Human body is designed in a way that it will feel stress and respond quickly to it. In a situation of threatened, frightened or something dangerous, hypothalamus which is a tiny region at the brain's base sends an alarming signal to the body; and by combining hormonal and nerve signals, this system prompts adrenal glands to release adrenaline and cortisol hormones which are called as "stress hormones" [3]. These hormones trigger the body's "fight or flight" response which prepares the body to respond in emergency situations by reacting according to it [4]. Adrenaline hormone is responsible for increasing heart rate, blood pressure etc. and cortisol is responsible for increasing sugar in blood, enhancing the use of glucose by the brain etc. Besides, this natural alarming system of body controls mood, motivation, fear etc.; by communicating with brain [3]. This system helps in handling emergency situations which is beneficial. This system is self-limiting as once the situation has passed, the body gets back to it's normal state however, if a person experiences a mental stress excessively for a longer period of time, the "fight or flight" response will be felt permanently which is not beneficial for human survival; as rather than helping to survive, it generates the feeling that a person is not worthy to handle the situation.

The mental stress which remains for long-term is known as chronic stress. Researchers suggests that different stress impacts the brain differently. Good stress or positive stress leaves a positive impact on brain which leads to greater resilience and stronger networks. On the contrary, long-term stress or chronic stress can affect brain's structure and functionality. The brain consists of neurons and support cells, known as "grey matter", that is responsible for higher-order thinking. However, "white matter", which consists of all the axons that connect other regions of the brain to communicate, is also a part of the brain. To communicate throughout the brain, the axons are surrounded by white sheath known as myelin which speeds up the electric signals. Due to the presence of chronic stress, myelin is produced excessively which changes the brain's structure and also makes the brain vulnerable to mental disease like depression, anxiety etc [5].

It has been found that behind depression and anxiety, there is mental stress. Central nervous system is affected by mental stress which results in depression and anxiety. Besides, Researchers stated that before diagnosing with depression and anxiety, people have history of mental stress. As mental stress releases cortisol hormone, on the contrary, it reduces dopamine which has link to depression. So, in chronic stress as the "fight or flight" response does not go to the normal state, that is why dopamine also cannot go to normal state which leads to depression [6]. In addition to that, anxiety is referred as the reaction of mental stress. In chronic stress, it leads anxiety to a serious condition where a person experiences fear, tension, distress all the time; which hampers one's normal life. So, there is a direct link of depression and anxiety with mental stress.

The EEG (Electroencephalogram) is defined as an electrophysiological monitoring method. It is a medical test to evaluate the electric signals of brain and it tracks and records the wave patterns of brain. EEG consists of small metal discs with electrodes and through it, the brain signals are captured and sent to the computer to record them. Brain is made of billions of cells of which, half are neurons. To measure the electrical activity which is a result of synchronized activity of neurons, EEG is put on the scalp surface. The technique of it to record the brain activity is electrodes are used to read the signals and sends the signals to the computer to capture the result. Signals are brought from micro-volt to a range where they can be perfectly digitalized through amplifier and to convert the signals from analog to digital, an A/D converter is used [7]. EEG allows to detect the condition within cortical areas by providing better time resolution. In this paper, by analyzing EEG data, we are attempting to detect mental stress of comorbid patients (depression and anxiety) using classifiers. To develop our study we are going to use EEG classifiers like Support Vector Machine, K-Nearest Neighbors algorithm, Decision Tree classifier, AdaBoost classifier, Bagging classifier ,Extra Trees classifier, Gradient Boosting classifier, Gaussian Naïve Bayes Classifier and Random Forest classifier.

Characteristics of mental stress that effects negatively is so cruel that it tortures a person both physiologically and psychologically which leads to severe critical diseases. As different people have different situations so they handles the situation differently and faces different types of pressures which leads to mental stress. As mental stress varies from individual to individual, depends on the situations so the type and symptoms of stress will also be varied and it is a major complication.

1.2 Problem Statement

Mental stress can be blamed as a root cause of mental illness. Researchers have stated that behind every psychological problems including depression, anxiety and also insomnia, mental stress is found. It is a global challenge as according to The American Institute of Stress, about 73 percent of people due to stress have experienced mental health problems, 48 percent of people gets lack of sleep because of it. Besides America, about 450,000 Britain workers believe that they are getting ill due to it, 86 percent of Chinese workers have reported that they are suffering from mental stress [8]. In Bangladesh as well, mental stress issues have been reported. The percentage of depression, anxiety , distress that has been recorded in Bangladesh is 54.3 percent, 64.8 percent, and 59.0 percent respectively [9].

Depression, anxiety, distress, insomnia, irritability and also memory and concentration problems are the consequences of mental stress. Depression is a serious mental illness which negatively affects a person's feelings which is a common issue in today's world and it can happen to people of any age. According to the World Health Organization, more than 264 million people are affected with depression [10]. The feelings people get due to depression are negative thought patterns, changes in appetite, changes in sleep patterns, self-blame, suicidal thoughts and behaviors etc. Anxiety is a form of mental health issue that leads to excessive tension, fear and nervousness; and as a result of it, people isolates themselves from social gatherings, suffer from constant headache and chronic pain, experience difficulty in sleeping. Moreover, due to it, people can commit suicide [11]. Insomnia is a disorder that makes it difficult for a person to sleep. It is a result of most of the mental sickness including depression and anxiety and also leads to issues like heart diseases and high blood pressure [12].

Most of the mental patients final and terrible decision is committing suicide. However, it can be stopped by detecting the problem in early stage and there are treatments for them however, the major challenge which arises is detecting them. However, for the effective treatment doctors at first need to find out whether it is depression or anxiety or something else. Most of the time depression and anxiety is thought as the same and patient is prescribed with anti-depressant medicines which is very dangerous. Most of the people do not take insomnia, memory and concentration problems seriously which leads the problem in deeper state and people becomes more stressed. So, the problem needs to be detected accurately to stop the terrible consequences and suffering of the people.

1.3 Objective and Motivation

Mental stress negatively impacts the body and brain to the extent that a person takes initiative to kill himself. In 2020's March, we came to know that the German Finance Minister Thomas Schaefer has committed suicide. It is stated that he was deeply worried about the economic fallout because of Corona virus [13]. It represents that he took this decision out of depression which is a consequence of mental stress. Among people who are suffering from it are reported that 51 percent of them are depressed and 61 percent have anxiety. Moreover, among the people who have been suffering from it, 16 percent gave statement of self-harmed and 32 percent stated that they had the suicidal thought and feeling at some point of their lifetime [14].

In spite of the treatments for it, people are suffering extremely which is the most surprising fact. Most of the time, depression and anxiety is treated same and people are encouraged to take anti-depressant pills. Unfortunately, cases are found where people did suicide by taking anti-depressant medicines [15]. Most of the insomnia patients do not take it for granted and have sleeping pills which harms the body very badly. So, the major reason behind these terrible consequences can be said that people are not accurately introduced with the illness like weather it is depression, anxiety or something like this which actuates us to come up with an approach to detect illness occurred by mental stress utilizing EEG signals. Handing this innovation to the doctors and medical centers, can increase the accuracy of detecting mental health issue and spare time and people's life.

1.4 Research Methodology

Our intention is to develop a model that would detect mental stress, so that it becomes possible and easier to identify and detect anxiety and depression at an early stage for better diagnosis before it gets worse. With this target, we started to do some background study on the topic as well as relevant research articles and topics. Then we tried to collect data however, due to the Global Covid-19 pandemic, it was impossible for us to collect the data by ourselves physically. So, we started to look for publicly open data over online and we got this data set from institution that was created for Punishment and Reward learning related work. The data set was divided into two parts- training and testing. Firstly, we processed the data according to our need and extracted necessary features from the data set. After that, we have created our model and used supervised machine learning classifiers such as Support Vector Machine, K-Nearest Neighbors, Random Forest, Ada Boost classifier, Gradient Boosting Classifier, Gaussian Naïve Bayes Classifier, Extra Trees Classifier and Decision Tree classifiers were used to check the accuracy of our model in case of detecting and to calculate precision, recall, F1-score and k-fold cross validation score to validate the results.

1.5 Research Challenges

Mental stress is a huge challenge in today's world. So, for the purpose of our objective, at first, we need to collect information from individuals, hospitals however, unfortunately due to Covid-19 situation, it was impossible for us to collect data physically. That is why we had to rely on the data that was available online and it was difficult to find as well.

1.6 Thesis Overview

Comorbid Patient means that a person has more than one disease present at the same time. In this paper, we are trying to detect stress, that results depression and anxiety in a person, from their reward and punishment learning EEG signals.

Firstly, In Chapter 1 (Introduction part), we have discussed about the motivation behind this research, Recent . We also talked about our methodology and research challenges briefly in this section.

In Chapter 2 (Related Work), we have explained about the papers that have researched in detecting the diseases involving depression and anxiety due to mental stress.

In Chapter 3 (Background Study) section states the background study of comorbidity condition, reinforcement sensitivity, EEG signals and the classifiers which has been used in the detection. We mainly discussed about what we studied from online besides research papers which are necessary for our work.

In Chapter 4 (Proposed Method and Implementation), we have explained our whole workflow in this research, how we analyse data, extract features and why use many classifiers in this research. After that we have explained how we planned to extract data from EEG Signals.

In Chapter 5 (Experimental Result Analysis), we have explained our model implementation and accuracy of our data sets. Finally, we have discussed which model performs better for our data sets.

Chapter 2

Related Work

Many pieces of studies have been published on the detection of mental stress. EEG is useful for assessing brain functional behavior and a thorough evaluation of this waveform may provide critical parameters indicative of patient's mental state. In a study, researchers have projected a model in the MATLAB area unit where cross coherency, channel event based capacity, energy spectroscopy graphs were analyzed to map multi channel electroencephalogram data sets in real time, emotion-specification. The primary and most complicated problem with electroencephalogram is locating the regions of the brain and the frequencies correlated with mental stress, especially moderate depression. In one paper, [16] they had chosen a computer-aided portable set-up which consisted of a multi channel EEG signal acquisition module by which they acquired emotion-related brain signals and through the signal processing module as materials and methods. They acquired the EEG data set which was used to develop a depression management application algorithm. Similarly, in our study, we are also using EEG data set to classify the mantel stress condition using different ML classifiers.

Acharya et al [17] have written a paper on "Computer-aided diagnosis of depression using EEG signals", where FD, wavelet-based energy and entropy- these methods were applied to extract EEG signal features. It is then used with artificial neural networks to classify normal behavior and depression from EEG signals. To differentiate the MDD; from MDD, FDs are used as the input of the enhanced probabilistic neural network classifier. To learn as well as choose the right classification model with the good recognition precision, a 10-fold cross validation approach is also used. It is also found that the beta sub-band from the complexity of frontal EEG may be used for monitoring and diagnosis of MDD treatment. A survey of EEG-based determination of depression was given attention on Computer Aided Design (CAD) utilizing nonlinear techniques. The clinicians as an apparatus can make use of a CAD system for confirming the condition. It ought to be of specific incentive for the early identification of misery. The writers introduced a novel melancholy finding record utilizing the signal of EEG and the accompanying non-linear RQA and HOS strategies.

In a recent paper [18], Jennifer and Tara have worked on the topic where they reviewed 184 EEG studies that consist of improvement studies in resting-state frequency bands through numerous psychological conditions, as well as stress, depression, panic attacks, schizophrenia condition, obsessive compulsive disorder (OCD) etc. In the depression portion, they presented their discovery through research that an improvement in absolute strength in both theta and beta bands with an average magnitude of 48 % was the dominant finding for depression. However, in the above papers, they have worked on the responsible bands and classification techniques to determine the presence of depression.

The author addresses depression in the paper [19], characterizing depression by impairments in the perception of and the use of effective knowledge to direct actions. They talk through an experiment concerning several experiments on the existence and neuronal causes of these anomalies, as well as that depressive patient's response maladaptively to penalties and hypo-sensitive to incentives, and these behavior phenotypes lead to irregular performance during a circumscribed network of brain areas, substantially monoamine-innerved reinforcement networks. They concentrate on behavioral effects, activities in which compensation or punishment is provided to patients during completion of the assignment. Depending on the mission, two main observations have arisen from this literature: stressful individuals exhibit maladaptive reactions to the penalty (negative feedback) and hypo-sensible incentive responses (positive feedback). In addition, they analyze research in human brain imaging and test the neurological associations of these behavioral effects. In addition, however, they researched technical work, which has started to include a compelling account of the neuronal signals driving the behavior of suicidal individuals.

In the paper [20], the author proposed a framework to perform a quantitative survey of the connections between reinforcement sensitivity, anxiety and depression using a meta-analysis and meta-analytic Structural Equation Approach analyses on 299 samples; original copies showed the collection of information suitable for the meta-investigation, where 204 studies that were published between 1991 and 2017 are detailed either Relationship between reinforcement response and self-reported signs intensity either discrepancies for the sensitivity to reinforcement among analyzed with sound subjects (yielding 483 Effect Sizes) and extra 39 experiments documenting the relationship among reinforcement sensitivity for both anxiety and depression (yielding 156 Effect Sizes). A meta-analytical structural equation model analysis was also conducted on the subset of eligible studies that evaluated the interesting co-variance between all four components (anxiety, depression, BIS, BAS). Here reinforcement Penalty sensitivity strong detects both anxiety and depression where depression (Hedges' g =.99) and anxiety (Hedges' g =.21) were found to be elevated but compensation negatively predicting depression disorders (Hedges' g =-.21)) even when it is straightforwardly Regulation over comorbidity. Depression's effect sizes are extraordinarily delicate to clinical state and directed by strategy of clinical evaluation.

A variety of studies have identified on reinforcement sensitivity theory and other mental illnesses. The goal of the current study [21] was to see reinforcement sensitivity theory (RST) to subjects with personality disorder as opposed to subjects with severe depressive disorder and a stable individual. Here, the author used behavioral inhibition system (BIS) and behavioral approach system (BAS) and, compared with all healthy individuals, all participants completed the BAS as well as BIS. In contrast with subjects with major depression disorder and stable controls, this study aimed to test BAS and BIS sensitivities in people with borderline personality disorder. The authors found in the analysis that the borderline personality disorder group displayed hyper-activation of both BAS and BIS system relative to the borderline personality disorder group evaluated to severe depressive conditions and balanced classes. They also found that logistic regression analysis reveals that in almost 80 % of the cases, scores on sensitivity to reward and sensitivity to punishment predicted borderline personality disorder. The sample size for BDI consisted of just 222 respondents. No major variations in sex distribution were found, but differences in age were found. However, his study was based on reinforcement sensitivity theory where we did take help of different classifier in our research.

The authors [22] analyzed and researched the relationship Life tension, behavior factors of RST, including effects of anxiety and depression, where anxiety and depression are assumed to be representative of psycho-psychology. In order to assess RST, life tension, and signs of internalizing, the author used a data set of 317 samples of undergraduate participants. 285 response samples, but 32 samples were omitted on the Infrequency Scale for missed responses or high scores. For good internal consistency, full study measures by MASQ, BIS/BAS Scales, SPSRQ, RLCQ, Infrequency Scale and all scales were appropriate. In this paper, for example, it was found that the relationship between BIS, BAS, and life events were important and accounted for 1 % of the range where higher BIS sensitivity predicted higher anxiety sensitivity, lower BAS sensitivity, and stronger BIS sensitivity predicted higher hedonic depression and low BAS, high BIS reported symptoms of hedonic depression when life stress was high.

The most widely used approach is the variation of the classification algorithm for feature choice and identification. As there are not many resources for diagnosing depression, in the paper [23] they have proposed a model where in a psychophysiological database, they have recorded 213 data among which 121 were normal and 92 were depressed. In their model, they have used three EEG electrode acquisition system in the frontal part in f1, f2 and fz position. They have collected the data from electroencephalogram in resting state and by using sound track and by asking different types of questions in the resting state. They have analyzed their data by using MATLAB software (version R2014a). To remove interference signals, they have use Finite Impulse Response (FIR) using Blackman time window. Besides this, along with Discrete Wavelet Transformation Kalman derivation formula, an Adaptive Predictor Filter etc. 270 linear and non-linear features have been used in this model. They have used four classification techniques which are ANN, SVM, CT, KNN to predict which one has the maximum accuracy in predicting depression. The result of this prediction which they found out is KNN has the best accuracy level which is 76.98 percent among the four techniques. They also stated that theta wave has a potential connection with depression. Moreover, the result recommended that the absolute value of this wave may be a sound feature of discrimination against depression. However, in our paper, we also use classifier to detect mental stress via reward punishment processing.

One of the pictures developed for improving the resulting features for better results is found in Li et al paper [24], which uses the nearest k-neighbor classification method to find patients with moderate distress and normal individuals. Previously, EEG data analysis involved pre-processing, extraction of information, compilation of information and classification where the job essentially involved removal of electro-oculogram, abstraction of occurrences, removal of pseudo trace, etc. they experimented with some groups whose BDI scores were ranging from 0-13 were being stated as in normal condition and 14-28 were being stated as in mild depression state. Electroencephalogram data was collected during the 128 channel HydroCel Geodesic Detector Net (HCGSN) and data generated signals have been analyzed while using Hanning filter and extracted alpha, beta and theta frequency bands for further extraction in order to do the classification, k nearest neighbor was used, which classifies the target on the basis of its k-nearest neighbors in the data set. The outcome of their experiment was that the performance of the beta band was better than the alpha and theta bands.

In the article [25] of Shalini and Sanchit, they have studied regarding depression on different methods of EEG signals using band power and the outcome of their study is ANN, SVM, LR has the high classification accuracy. Among them, SVM for its overall margin difference between the separation of hyperplanes and the kernel technique has the best accuracy which is 98.33 %. They have also discovered that there is a conflict in the studies in presence of alpha activity during the depression.

In paper [26] authors used supervised machine learning approaches to study four types of factors such as emotional, linguistic style, temporal process and features for processing data which was received from Facebook posts to detect depression. Here, they found that classification techniques such as SVM, decision tree, KNN, and ensemble are appropriate for all the four types. They used NCapture to collected data from Facebook user's comments for exploration depressive behavioral and detection . The Facebook data was set into two parts, one was Yes and another one was No, where Yes represent depression and No represent non-depression. however, they applied total of 21 forms of features in LIWC applications used to diagnose depression and the accuracy was between 60 to 80 %.

In this research [27], the authors proposed Anxious Depression prediction model in real-time tweets. They conducted their research using online twitter posts of 100 sampled users and 5-tuple vector which were word, timing, frequency, sentiment, contrast used to define feature set based on linguistic cues and user posting patterns. The 5-tuple vector representing presence or absence of anxiety related word (word), more than 2 posts in between 12am to 6am (time), more than 3 posts in an hour (frequency), in average 25 % posts with negative polarity in a month (sentiment), the presence of 25 % more polarity contrast in posts in a day (contrast).Three classifiers Naïve Bayes, random forest, gradient boosting, are used to trained the AD prediction model and generate the final prediction using Ensemble Vote Classifier. The accuracy of AD Prediction in Real-time Social Data model is 85.09 % with an F-score of 79.68%.

The writers in the paper [28] suggested a methodology to diagnose comorbid depression and anxiety through the processing of reward and punishment. Their suggested methodology found that anxiety predicted stronger EEG features associated with punishment and depression predicted a smaller EEG feature linked to reward. Their data set included data from the train and test and there were 121 samples of individuals with anxiety, depression disorder and control. In the report, the authors conclude that a person with less than 7 BDI is a control individual, less than 13 individuals with BDI anxiety disorder and more than 13 individuals with depression disorder are BDI. Using various set of pseudo-randomly allocated across, participants conducted a forced decision learning process and testing phase. The feedback from the participants was obtained as "correct", "incorrect" or "no response detected." For the source estimation, the sLORETA toolbox was used.The authors have dimensions of mood and anxiety that are related to particular aspects of EEG responses to reward and punishment, respectively. The performance, ERP features, frequency, spatial generators, computational function for dissociated comorbid depression, and anxiety were tested in this study where we did take help of different classifier to identify the mental stress.

The authors applied a machine learning model in paper [29] to identify mental stress at different stages by investigating the variations between mental stress and management environments. In order to classify the stress levels, they performed their study using electroencephalogram signal analysis of stressed individuals. The writers used the Montreal Imaging Stress Task to trigger stress validity. The authors used forty-two samples, eleven of which were taken from women. In three classifiers that were Support Vector Machine, Logistic Regression, Naive Bayes classifiers, the structure of this ML model applied within the extracted options followed by feature choice using the receiver in service characteristic parameter, the t-test and even the Bhattacharya distance and options were added. Lastly, to prevent classifier overfitting, they use 10-fold cross-validation. Their highest accuracy was 94.0 % using t-test and Naïve Bayes for level one, 93.9 % with t-test and help vector machine for level two, 94.6 % with t-test and Naïve Bayes for level three ,91.7 % with t-test and Naïve Bayes for level four. With the t-test across each stage, it was determined that the most productivity was obtained. The proposed system, however, shows 94.6 % precision for two-level stress recognition and 83.4 % precision for multiple stage detection. Their studies recommend that electroencephalogram signals have the ability to evaluate stress level accurately.

Chapter 3

Background Study

3.1 Comorbidity Condition

The term comorbidity condition is known as a condition where more than one disease is present in a person at same time. This condition refers to a situation where for example, a person is having heart disease and psychiatric disorders in the meantime. The Corona infected patients are also observed that they have comorbidity condition . Most of the Corona infected patients are also suffering from depression or anxiety besides it. This condition occurs mostly when two disorders have overlapping risk factors. Researchers have found two types of comorbidity condition. Firstly, Homotypic comorbidity which refers to the co-occurrence of mental disease in a diagnostic group. The second one is heterotypic comorbidity which refers to the co-occurrence of two disorders from different diagnostic group [30]. Comorbidity condition occurs when a set of risk factors are same or similar for two disorders then by developing one disorder and other one may also develop as well. These factors can be biological, personality, social, environmental or a combination of them [30].This is a common phenomenon. In a study of comorbidity condition of 2009's organized in Spain, it's stated that among 7939 patients about half of them had more than one mental disorder.

People who are diagnosed with social anxiety disorder and major depressive disorder at the same time are defined as having comorbid anxiety and depressive disorders. For depression and anxiety being paired up, there are certain things behind it being said so. First of all, their biological mechanisms in the brain are same that's why the disorders are shown together most of the time. Secondly, they have symptoms that overlap with each other. For example, insomnia is a result of both depression and anxiety. In addition, in the moment of stress, these conditions are present simultaneously [31]. Furthermore, according to U.S. National Comorbidity Survey, about 51 % of major depression disorder patients were also had anxiety disorder. However, 48.6 % patients diagnosis with major depression had at least one anxiety disorder according to the Early Developmental Stages of Psychopathology Study [32].

3.1.1 Anxiety Disorder

Anxiety is a reaction to stress that alerts people in danger by reacting to it but anxiety disorder is fully different from anxiety. Anxiety disorder is a group of mental disorders that constantly makes a person frightened, overwhelmed and anxious. It is a mixed feeling of fear, worry instead of alarming one in danger, damages one's normal life. It is one of the serious psychiatric disorders and attacks people of all ages from school going children to old ages. About 8 % of children and teenagers experience this disorder before the age of 21. According to the World Health Organization, globally 1 in 13 people suffers from this disorder and reported that it is the most common mental disorders with specific and social phobias, major depressive disorder etc [33].

Anxiety is a result of interaction between different brain regions. Scientists have defined amygdala as the fear and anxiety hub which is a marble sized brain area. Amygdala is located in the emotional part of the brain and people experiences anxiety when signals from emotional part overpowers signals from cognitive part of brain. However, Dorsal Anterior Cingulate Cortex (dACC) which is a region in the frontal lobe, amplifies the signals that come from amygdala. For a research purpose when patients of anxiety disorder had shown pictures of scared faces ,the amygdala and dACC were found in producing anxiety on the contrary normal people was found to show a little of no response. Moreover, in the patients of anxiety disorder a part of frontal lobe known as ventromedial prefrontal cortex has found that it dampens the signals coming from amygdala and as a result of it, the brakes of amygdala remains lifted [34].

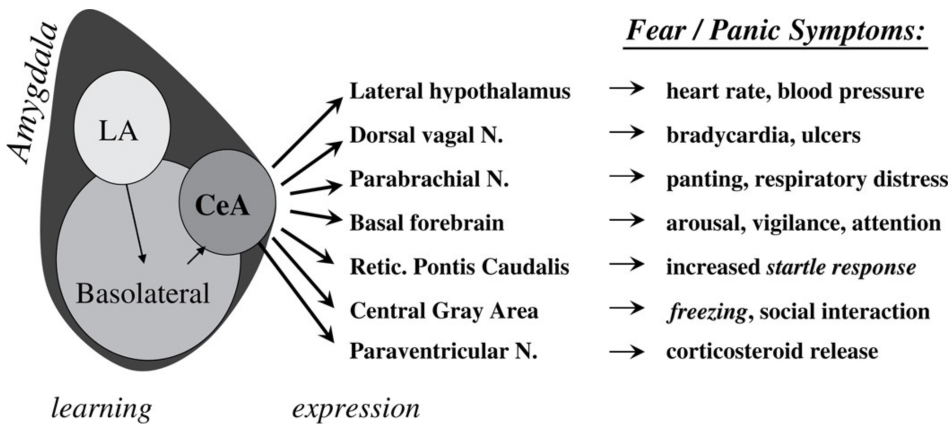


Figure 3.1: **Amygdala in anxiety and fear [35]**

Behind anxiety disorder, risk factors like genetic, psychological, environmental or combination of them works. A person can get anxiety disorder from family as well like allergies, diabetes etc. However ,the term anxiety disorder refers to psychiatric syndromes including

1. **Generalized Anxiety Disorder:** People suffering from it feels limitless anxiety for a long period of time and the least time of it is 6 months. Other symptoms of it is, having muscle tension, no control over the feelings of worry, irritated over small things etc.

2. **Panic Disorder:** In this, people have panic attacks which a intense feeling of fear which reaches at the peak very quickly within a minute.

3. **Agoraphobia:** It is a fear of being in a situation where thing might get embarrassing, no help will be available during panic attacks etc. This phobia can be so serious that a person cannot leave house because of it which affects the normal life of the person.

4. **Social Anxiety Disorder:** In this anxiety, people fears to go it the social places for the fear of being ashamed, rejected etc and for this people try to avoid social gatherings and it leads the person to lead in loneliness which increases the anxiety more and results in mental diseases also [36].

Anxiety disorder impacts on a person's personal life as well as social life. It creates full of fear, worry and makes difficulties for the person to lead a normal life. Due to the disorder, a person is scared of being a part of social events which leads to loneliness and this is a major cause of a lot of psychiatric problems. However, along with these impacts, anxiety disorder also damages to the body. Sleeping is really important for one's good health but this disorder leads a person to spent sleepless nights and days and as a result of it along with mental illness, body suffers from heart disease, high blood pressure, respiratory distress etc.

3.1.2 Major Depressive Disorder

Depression is a mood disorder which creates a constant feeling of sadness, loss of interest in things etc. for weeks, months even years. Major Depressive Disorder is a mental health condition manifesting at least five major symptoms of depression for at least two weeks. It is also known as clinical depression which affects mood and behavior as well as physical conditions. It impacts people of all age. According to the National Institute of Mental Health, 7.1 % of population of the age of 18 experiences major depressive disorder. Approximately, 1.9 million children of age of 3 to 17 have diagnosed with major depressive disorder. About 7 millions of people of the age of 65 % and more have diagnosed with this disorder [37].

Major Depressive Disorder is a result of chemical imbalance in the brain and three region of brain affects in it. They are the hippocampus, amygdala and prefrontal cortex. The hippocampus which is located near the center of the brain is in charge of storing memory and regulating cortisol hormone. Body releases cortisol in the time of stress or emotion related issues or chemical imbalance in body but excessive amount of cortisol hormone is harmful for the normal functioning of brain. In a part of the hippocampus called the dentate gyrus, brain cells are produced here. In the patients of this disorder as the cortisol releases excessively in their body ,it affects in the production of new brain cells and shrinks them in the hippocampus. Secondly, the very front of the brain is prefrontal cortex which regulates emotions, makes decisions and forms memories. Due to this disorder, body releases excessive cortisol and for this, the prefrontal cortex appears shrink. Lastly, amygdala is responsible for emotional responses. In the patients of major depressive disorder, it has been noticed that their amygdala becomes enlarged and more active which results in releasing more cortisol constantly. The abnormal amygdala can cause the body to release excessive amount of other hormones and chemical imbalance in the body [38].

In this modern era, maximum population are suffering. Researchers have stated that it is caused by the combination of genetic, environmental, psychological and biological factors. People may have depression from family as well. A person will be referred as a patient of major depressive disorder if he or she suffers from at least 5 symptoms depression including feeling worthless, unhappy, overwhelmed all the time, lack of confidence etc. There are different types of depression and some of them are,

1. **Melancholia:** It is a severe form of depression where more than one depressive symptoms are present. In it, people loses happiness completely from everything and surrounds himself by a mental environment full of depressive situations.

2. **Psychotic Depression:** This term refers to a depressive condition where people starts to have hallucinations and delusions by having thoughts that they are not good, they are responsible for anything bad etc.

3. **Antenatal and postnatal depression:** This depression is experienced by the pregnant women.10 % of women experiences depression during pregnancy and 16 % in first three months after the birth of baby [39].

Major depressive disorder impacts a person's life so badly that life becomes a burden to him or her and he or she takes the decision to finish it most of the time. According to the National Institute of Mental Health, from 2001 to 2017 the suicide rate has increased 31 % from 10.7 to 14.0 per 100,000. Among 30,000 reported suicides, depression is blamed to take lives of two-thirds [37]. Besides that it damages a person's normal life, body's normal functioning. Insomnia is a result of this disorder which harms brain cells along with body. Along with these, a person who is suffering from major depressive disorder can not eat properly, can not share his feelings of suffering with anybody, tries to do self-harm.

3.2 Reinforcement Sensitivity

A theory which proposes three brain-behavioral systems that underlie in their individual differences in reward, punishment and motivation sensitivity is known as Reinforcement Sensitivity. It is proposed by Jeffrey Gray in 1970 and then revised by him in 2000. It is one of the most influential personality theories that are biological based which describes in differences of individuals in the tendencies of approach and avoidance [40].It is a psychological process model that processes the signals that lies in the personality dimensions. The revised theory of it differentiates between fear and anxiety. This theory is built based upon psychological short-term matters which increases by time and results in long-term emotion and behavioral matters. This theory identifies the defined personality factors as variation sources that remains stable over time. This is a demand of personality facts that behaviour changes of individual with identical environments and also it's consistency over time. To identify the biological variables that are relatively static which determines the factor structures in behavior is the goal of this theory [41].

Reinforcement Sensitivity theory has a great influence in neural and psychological sector. Gray's reinforcement sensitivity theory is also known as theory of personality. From this theory, fear and anxiety can be differentiated. The study of anxiety disorder has reached in clinical settings because of this theory and along with it, this theory is also used in predicting anxiety, impulsivity, and extraversion and work performance of individuals [42]. The biological theory of personality which is renowned as reinforcement sensitivity theory is one of the major models which is based on the behavioral differences of individuals in reactions to rewarding and punishing stimuli. A group of neural structures which are responsible for associative learning, positively-valenced emotions and incentive salience are refers to as reward system . On the contrary the neural structures who are associated with unpleasant things are referred to punishment system. Reward and punishment stimuli are defined as reinforcers. Based on these two stimuli or systems, reinforcement sensitivity theory reaches to it's conclusion about psychiatric and neural matters.

3.2.1 Reward and Punishment Processing

The reward and punishment reinforcers works in shaping behaviour of individuals. With distinctive behavioral effects and neural substrates, these reinforcers affects the behaviour of individuals and that signals create a connection which helps in behavioral learning [43]. The processing which is incorporated with the response to rewarding stimuli, the ability to learn from reward the anticipation of future rewards, and engagement in goal-directed behavior towards rewards is called as reward processing. In the theories of Gray, reward processing is referred in terms of brain system which is responsible for behavioral approach of rewarding stimuli and generating incentive motivation. Reward processing has become a salient feature of reinforcement sensitivity theory. Jeffrey Gray proposed that the differences in the sensitivity involved with reward processing of neural system can form a basis partially for the validation of personality [44]. According to him, reward process is in the behavioral activation system (BAS) that approaches to motivation [42].

The processing which is incorporated with the response to punishing stimuli is referred to as punishment processing. According to Jeffrey Gray, punishment processing lies in the behavioral inhibition system (BIS) and avoid motivation [42]. This system has linked with the amygdala, septo-hippocampal system etc that are responsible for behavioral approach of punishing stimuli and in avoidance of motivation [45]. In reinforcement sensitivity theory, the other part which helps in validation of personality is punishment sensitivity.

In reinforcement sensitivity theory, reward and punishment process plays a vital role. Reinforcement sensitivity theory analyses individuals behaviour based on these two. Reward is used for indicating the increase in the probability of behavior and punishment is used in indicating the decrease in the probability of behavior [46]. By analyzing the impacts of them, mental state issues are being tried to define.

3.3 Electroencephalogram

EEG impulses provide a range of data about healthy brain activity. Application for electroencephalography that analyses brain activity can promote early detection of various mental disorders. Tiny electrodes be put at the front of the head while on an EEG and are connected to a device. Both the electrical signals that neurons trade to each other and all of the brain regions that would be at function are calculated by such a device. EEG tracks the electrodes through documenting frequency range for electrical signals that used nerves which interact via scalp-attached electrical signals. The use of massive electrode classes such 128, 256 and 512 electrode arrays connected to cranium has expanded the scope of EEG recording capabilities. However, featured, streamlined, appropriate EEG data are quite important to enhancing the accuracy of performance assessments around the same period. Through analyze a detection and recognition of EEG data, several authors have been using extraction of features as well as classification methods in different areas.

3.3.1 Measurement

The EEG headset containing 64 Ag-AgCl pin-type functional electrodes were put upon this participant's scalp with 64 channels to calculate the electroencephalogram dataset [47].

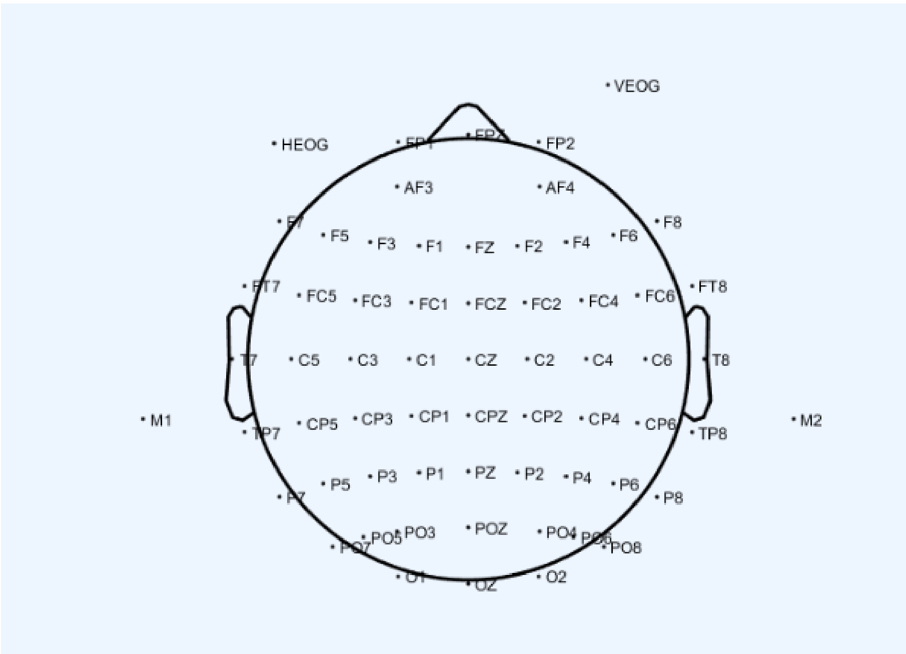


Figure 3.2: **Channel Location of 64-channel**

Position Name	Channel Name
Pre-Frontal	FP1, FPZ, FP2
Anterior-Frontal	AF3, AF4,
Frontal	F7,F5,F3,F1,FZ, F2, F4,F6, F8
Frontal-Central	FC5,FC3,FC1,FCZ,FC2,FC4,FC6
Frontal-Temporal	FT7, FT8,
Central	C5,C3,C1,CZ,C2,C4,C6
Temporal	T7,T8
Central-Parietal	CP5,CP3,CP1,CPZ,CP2,CP4,CP6
Temporal- Parietal	TP7,TP8
Parietal	P7,P5,P3,P1,PZ,P2,P4,P6,P8
Parietal-Occipital	PO7,PO5,PO3,POZ,PO4,PO6,PO8
Occipital	O1,OZ,02
Mastoid	M1, M2
Vertical and Horizontal EOG	HEOG, VEOG

Table 3.1: **List of 64-channel EEG headphone**

3.4 Classification Using Machine Learning

ML is a branch of artificial intelligence that concentrates in discovering patterns and producing predictions for the future based on training samples. It could also be as observational in real-time. In order to distinguish consumer environments, certain machine learning techniques are SVM, RF, k-NN, Decision Tree classifier, AdaBoost classifier, Extra Trees classifier, Bagging classifier, Gradient Boosting classifier, Gaussian Naïve Bayes Classifier.

3.4.1 Support Vector Machine

In 1995, on the basis of statistical learning approach, Vapnik introduced SVM, which is widely used in both detection and nonlinear regression. The basic feature of SVMs include to modify the vector to the high - dimensional feature space but also to determine the optimal identification plane such that the object can be divided linearly. The functional difficulties related by limited sizes data, nonlinear interactions, large dimensions, including local minima could be overcome by SVMs. The devices can reach predictive performance beyond 83.5% [48]. The difficulties due to the small sampling scales, nonlinear interactions, and different identifications could be overcome by helping the vector system. In order to optimize the distance among them, the concept of SVM method is to establish an ideal hyperplane as the decision surface to classify the various groups.

3.4.2 Random Forest

A random forest algorithm containing several identification trees was proposed by Breiman [49]. Each tree could have been an algorithm as well as the classification outcomes from several trees determine the general identification outcomes. Algorithms are characterized like an ensemble learning methodology that produces a few individuals. Through the toughest division among all characteristics in a regular tree where each node is separate.

3.4.3 K-Nearest Neighbors

The k-NN method is among the earliest identification algorithms in the field of machine learning. This technique requires quite basic logic to evaluate the program's performance for entry from vector space dependent on K-nearest neighbors. Since the estimation is delayed as far as necessary, each method belongs to the lazy learning applications group [49]. There are function vectors and a tag representing the type of such a function in the raw samples in the k-NN method. This same sample to the framework is number k all through identification, representing so many neighbors to also be referred to for identification as well as the unidentifiable objective function. The efficiency of such a method is based largely on the raw data and it is quite usual for input to be pre-processed before k-NN identification by any feature scaling.

3.4.4 Decision Trees Classifier

Decision tree learning has been one of the techniques to statistical analytics used during research, data processing and deep learning. The relevance of a function is evaluated by nodes in a decision tree [50]. Often, that node test compares a feature's value to a constant. Two attributes may be contrasted with one another in certain situations, or a method is added to attributes [51]. By filtering this down the tree, a new data is labeled in relation to the extracted features defined in consecutive nodes. That data is graded as per the class identified with leaf, until a leaf was already achieved [49].

3.4.5 AdaBoost Classifier

AdaBoost is an eminent community learning paradigm focused on grouping, first suggested by Freund and Schapire [52]. AdaBoost can be found with any training method, equivalent to bagging. The AdaBoost method operates by attaching the very same weight among all learning sample instances, and now by pressuring the training system to construct an algorithm for these samples and reweights each instance as per the performance of its classifier.

3.4.6 Extra-tree Classifier

Extra-Tree is an ensemble learning method. From the training dataset, it creates a large number of decision trees. It generates its output by aggregating the results of de-correlated decision trees which consists of training datasets. In the time of splitting a tree node, it is associated with randomizing strongly attribute and cut-point choice. Randomized trees who are structurally independent of the output of a sample dataset are generated by it in extreme cases . Computational efficiency and lower prediction variance are the strengths of extra-tree classifiers compared with other decision trees.

3.4.7 Bagging Classifier

Bagging Classifier is an ensemble based meta- estimator.For final prediction,it fits base classifiers on original data set's random subsets and then generates individual predictions either by voting or by averaging.It helps in reducing the overall variance and for this unstable classifiers can be made strong. The outputs of multiple classifiers of different training data set samples are combined in it.This technique is useful for both regression and statistical classification.It raises the models stability in reduction of variance and in improvement of accuracy which eliminates the problem of over fitting.

3.4.8 Gradient Boosting Classifier

Gradient boosting classifiers is a group of machine learning algorithms that constructing a strong predictive model combines several weak models. In this classifier, its predecessor's errors are corrected by each predictor. Usually decision trees are used in gradient boosting. It generates an ensemble prediction model for regression and classification problems. Loss function, weak learner and additive model are the three main components of it. It is called gradient boosting since it uses a gradient descent algorithm to reduce losses while introducing new models.

3.4.9 Gaussian Naïve Bayes Classifier

Gaussian Naive Bayes follows the normal distribution of Gaussian. Based on the Bayes theorem , a group of supervised learning classification is known as Naive Bayes. This classification technique is quite simple but the functionality of it is high. It supports features and models of continuous data. It is used in binary and multi-class classification problems. It predicts different classes' probability based on various attributes. Class and conditional probabilities are stored in this model . Mean and standard deviation are measured from training data.

Chapter 4

Proposed Method and Implementation

4.1 Workflow

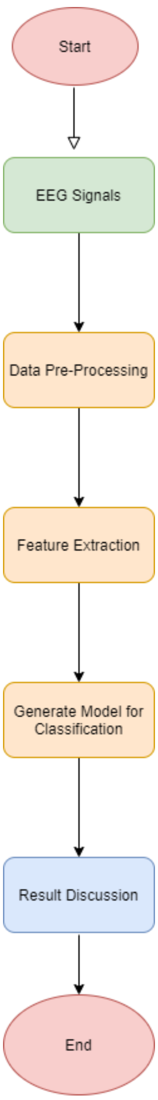


Figure 4.1: **Workflow**

4.2 Data Collection

For our paper, we use the recorded EEG data set on control and comorbid patient mentioned in [28] paper. It is publicly available for computationally informed EEG.

4.2.1 Data Description

For our venture we work on 50 participants (17 males and 33 females, ages 18 to 25) from the data set.

Gender	Count of Gender
Male	17
Female	33

Table 4.1: **Total Participants**

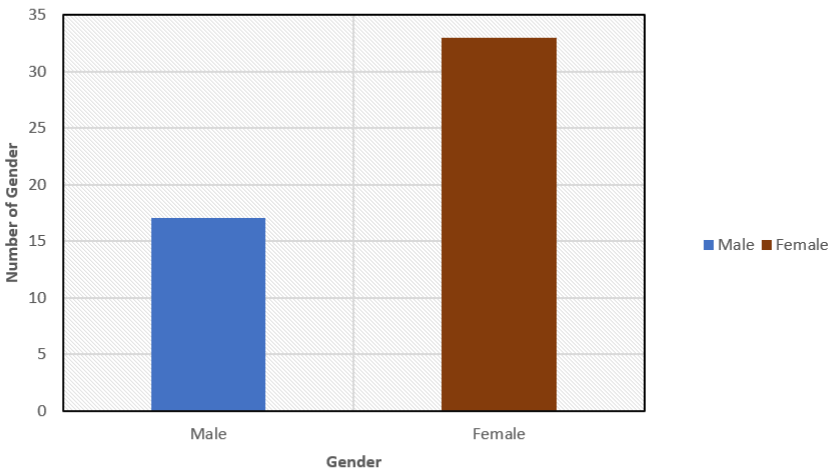


Figure 4.2: **Bar Chart of Total Participants**

A probabilistic reinforcement training assignment was carried out by the respondents. The assignment involved a process of imposed option training [53]. Three stimulus sets had been introduced to the respondents in training process. Such stimulus combinations were labelled as A or B, C or D and E or F. The parametric probability of getting "Correct" or "Incorrect" input was unique for every stimuli. The stimuli lasted for such a period of 4000 ms, but also once the decision was taken, they vanished. If within 4000-ms the individual unable to reach a decision then no response will show. The participants of Comorbid are two types in the data set. One comorbid participant had anxiety another comorbid participants had depression.

Participant	Count of Participant
Control	28
Comorbid (Anxiety and Depression)	22

Table 4.2: **Category of Participants**

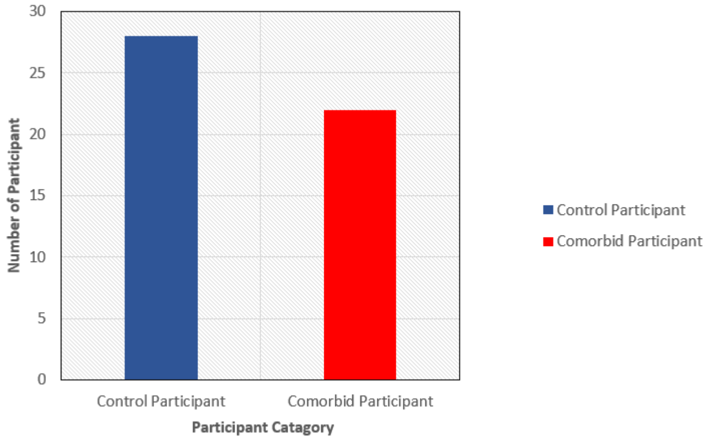


Figure 4.3: **Bar Chart of Participants**

In the data set, 22 participants were suffered from both anxiety and depression during their training phase.

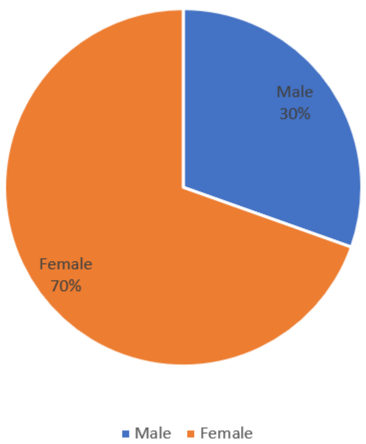


Figure 4.4: **Bar Chart (Gender) of Comorbid Patients**

The number of female comorbid participants was more compare to male comorbid participants. Which means more female participants feel anxiety or depression during training phrase than the male where the age range was 18 to 25 years.

4.3 Data Pre-processing

We have used the pipeline diagram shown in figure on our raw EEG data set.

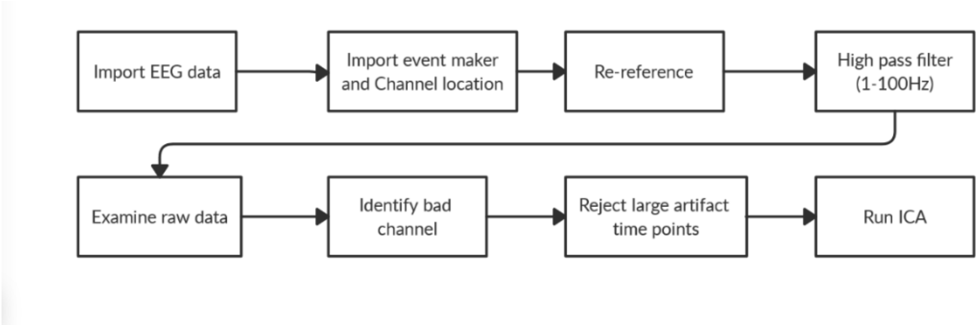


Figure 4.5: **Data Pre-processing Pipeline**

EEG recording is strongly sensitive to multiple noise sources. By processing EEG data, the noise makes it hard to detect the functions. There are several established approaches to interact successfully with noise [23]. ICA is a tool that we use to eliminate noise from a data set. The sampling rate of this data set is 500 Hz and Eye blinks were removed with Independent Components Analysis. The signals were filtered with a minimum of 1 Hz and maximum of 100 Hz using band-pass filter. In raw data, 66 channels were collected. But after pre-processing 4 channels did not improve which are T1, T2, CB1, CB2. So, we remove these 4 bad channels. In addition, M1 and M2 were used as reference.

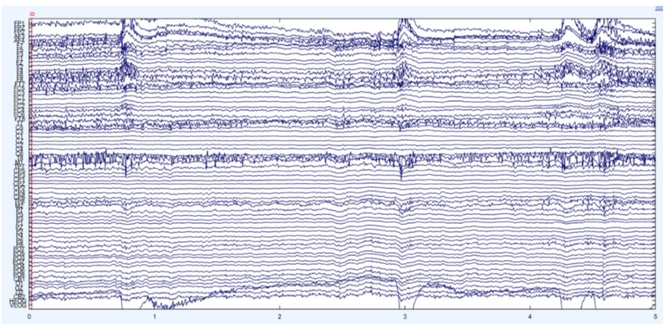


Figure 4.6: **Signal of raw data set**

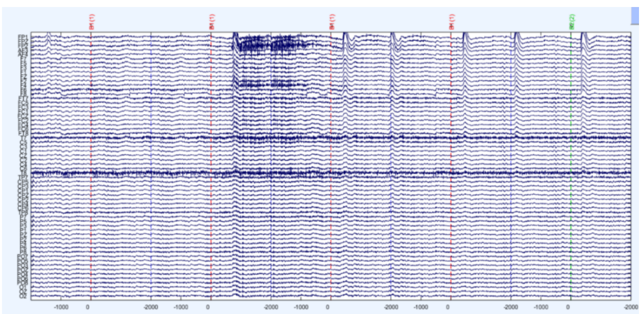
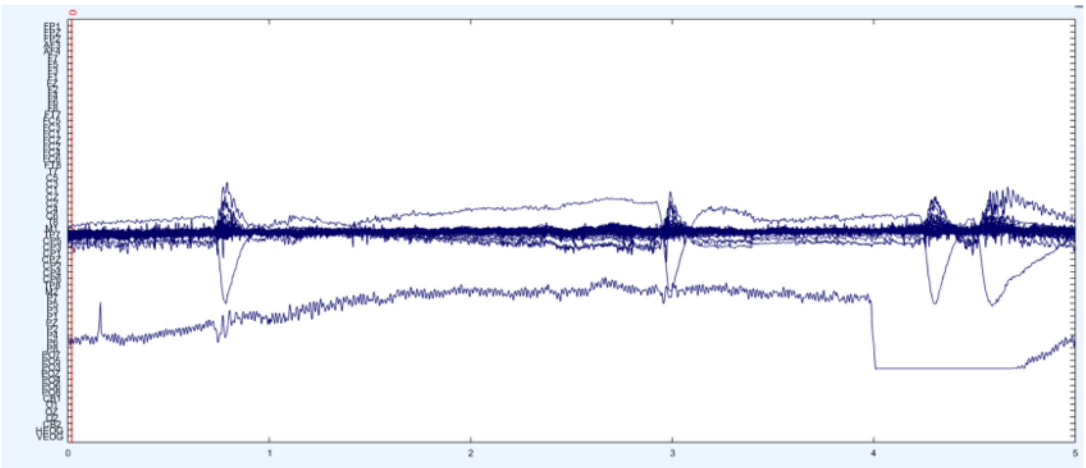


Figure 4.7: **Same signal after run ICA**

In figure 4.6 and 4.7, we did comparison of the raw channel and after pre processing channel for our data set.



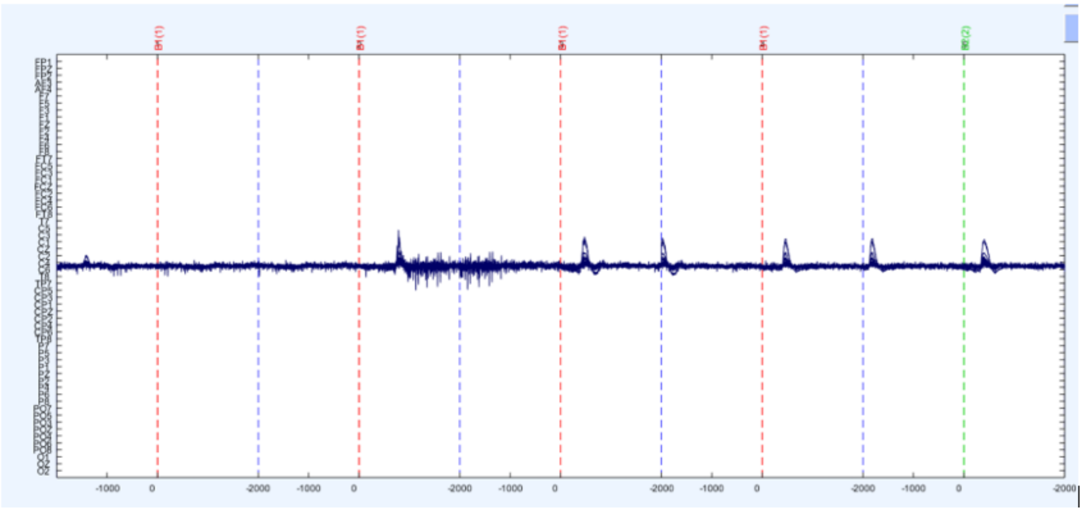


Figure 4.8: **Stack version of EEG data set**

In figure 4.8, we see the stack version of EEG signals that before pre-processing, the data set consisted of noises, eye blinks. After pre-processing, noises and eye blinks were removed.

4.4 Feature Extraction

We have split our data into two sections in this paper, correct feedback and incorrect feedback that we get in the training phrase. We label "Yes" as the correct feedback and "No" as incorrect feedback. After that, We used 7 statistical features for 60 channels which are mean, skewness, standard deviation, kurtosis, log energy entropy, shannon entropy and energy for extracting our EEG data. The description of used features and their formulas have given below:

Mean: The average of two or more is known as mean. The formula for calculating mean is, Mean = $\frac{\sum fi}{z}$

Where,
fi = i-th variable in the dataset.
z= number of variables in the dataset.

Skewness: A degree of change from the mean of a data is represented by skewness [54].It can also be referred as asymmetry in a dataset. It indicates the direction of the outliers and determines the overall shape of a distributive curve.

Standard Deviation: Standard deviation measures a dataset's dispersion from its mean. The standard deviation which indicates that the values are far from the mean, refers high standard deviation. On the contrary, low standard deviation refers to the fact that values are near to the mean.

Kurtosis: Kurtosis provides information regarding the degree of the values that are concentrated around the mean which is referred to as peakedness of a distribution of data. High kurtosis datasets have peak near mean and low kurtosis datasets have flat-top near near [55].

Log-Energy Entropy: Log energy entropy provides reliable features with a very low error [56]. Formula to calculate this,

Entropy(x) = $\sum log(x_i^2)$

where- i=number of terms

Shannon entropy: Shannon entropy is a measurement of uncertainty that is consists of random variables. Formula to calculate this [57],

Entropy = $-\sum(P \times ln(P))$

where- P = fraction of population composed of a single species

Energy: Energy is effective in detecting signal noises. In detecting noises of signal, energy is very efficient.

This analysis allows to gain a clearer and more organized view of where exactly our data is processed, and after extracting features we implemented machine learning approaches for classifying our extracted data.

4.5 Classification and Implementation

Due to pandemic, we had to rely on online available data sets. We have collected our data set from an online site (link: http://predict.cs.unm.edu/downloads.php). From the available data, we have applied cross section on training part of the data and from that have worked on 50 data sets from which 28 are of control patients and 22 are of comorbid (depression and anxiety) patients.

Patient id	Patient	Gender	BDI	TAI
507	Control	Female	0.00	23.00
509	Control	Female	7.00	44.00
510	Control	Female	1.00	27.00
511	Control	Male	1.00	23.00
512	Control	Female	1.00	26.00
514	Control	Male	5.00	37.00
515	Control	Female	5.00	36.00
516	Control	Female	0.00	28.00
517	Control	Male	0.00	23.00
518	Control	Female	1.00	36.00
519	Control	Male	6.00	40.00
520	Control	Female	3.00	28.00
521	Control	Male	2.00	24.00
522	Control	Female	0.00	30.00
523	Control	Female	1.00	30.00
524	Control	Female	3.00	38.00
525	Control	Female	2.00	30.00
526	Control	Female	0.00	30.00
527	Control	Female	1.00	31.00
528	Control	Male	1.00	33.00
529	Control	Female	0.00	27.00
530	Control	Male	1.00	34.00
531	Control	Female	1.00	33.00
532	Control	Female	1.00	42.00
533	Control	Male	2.00	24.00
534	Control	Male	5.00	37.00
535	Control	Male	2.00	36.00
536	Control	Male	2.00	37.00

Table 4.3: **List of Control Patients**

Table 4.3 has the information of Control (CTL) patients and Table 4.4 has the information of Comorbid patients. When patients have BDI score [46] is greater than equal 13 (BDI$\geq$13), then those patients have major depressive order. Also, from 45 to 80 score of patients have high anxiety (45-80=TAI). So in figure 4.3, 22 patients have both high anxiety and depression at the same time. So, these patients are known as Comorbid patients and in figure 4.2, many of the patients have no MDD and some have moderate anxiety. So, we catagorized them as Control patients.

The overall dataset is based on comorbidity condition and they have used reinforcement learning so, for the research they have gather information from individuals in a questionnaires session and stored their answer in the form of correct and incorrect. To indicate correct and incorrect they have used 94 as correct and 104 as incorrect. However, for our research we have chopped the signals based on these correct and incorrect numbers from its training part and with the chopped data, we have extracted seven features namely mean, skewness, standard deviation, kurtosis, log energy entropy, shannon entropy and energy.

After chopping and extracting the features, for the purpose of detecting depression and anxiety, we have classified the data with nine classifiers which are support vector machine (SVM), k-nearest neighbor algorithm (KNN), decision tree, adaboost, bagging, extra tree, gradient boosting, gaussian naïve bayes and also random forest. To ensure the high level of accuracy these classifiers have been used. Using these classifiers, we have calculated precision, recall, F1-score, accuracy of our correct and incorrect dataset for reward and punishment processing.

Patient id	Patient	Gender	BDI	TAI
598	Comorbid	Female	19.00	60.00
602	Comorbid	Male	27.00	65.00
603	Comorbid	Female	21.00	52.00
604	Comorbid	Female	22.00	57.00
605	Comorbid	Female	20.00	52.00
606	Comorbid	Male	20.00	47.00
607	Comorbid	Female	28.00	63.00
609	Comorbid	Female	26.00	58.00
610	Comorbid	Female	15.00	58.00
615	Comorbid	Female	30.00	68.00
616	Comorbid	Female	16.00	48.00
617	Comorbid	Female	18.00	57.00
618	Comorbid	Female	24.00	53.00
619	Comorbid	Male	23.00	50.00
620	Comorbid	Female	27.00	64.00
621	Comorbid	Female	17.00	44.00
622	Comorbid	Male	14.00	41.00
624	Comorbid	Female	23.00	60.00
625	Comorbid	Female	16.00	60.00
626	Comorbid	Female	14.00	41.00
627	Comorbid	Male	30.00	47.00
628	Comorbid	Male	19.00	56.00

Table 4.4: **List of Comorbid (both Depression and Anxiety at the same time) Patients**

Accuracy: The formula to calculate accuracy is [58],

Accuracy $= \frac{TP+TN}{TP+TN+FP+FN}$

Where,
TP=True Positive
TN= True Negative
FP= False Positive
FN= False Negative

Precision: The formula to calculate precision is [59] [60],

Precision $= \frac{TP}{TP+FP}$

Where,
P= Precision
TP=True Positive
FP= False Positive

F1-Score: The formula to calculate F1-score is [59] [60],

$F_1 = 2 \times \frac{P \times R}{P+R}$

Where,
P= Precision
R=Recall

Recall: The formula to calculate recall is[59] [60], Recall = $\frac{TP}{TP+FN}$

Where,
TP=True Positive
FN= False Negative

We have calculated precision, recall and f1-score for both correct and incorrect for control patient and comorbid patient and accuracy for overall data of the patients. Lastly, for validation of accuracy, precision and the others we have used k-fold cross validation over the data of training part.

Chapter 5

Experimental Result Analysis

5.1 Supervised Learning

Supervised learning is a sub-category of machine learning .In this learning, both labeled input data and output results are given in the program. It is called supervised learning because the algorithm learns from the training data which is associated with input data along with expected output. The objective of this learning is to predict correct label for new input data [61].As this algorithm in trained from a training data which is associated with input and output and for this in our study we have used it. Our data is consist of 64 channels which is also a reason behind using this algorithm as the training data of it, consists multiple sample data.

5.2 Support Vector Machine Algorithm Implementation

Support Vector Machine is a well known supervised learning model. Support Vector Machines are one of the most powerful prediction methods. We used Support Vector Machines to classify our data. After pre-processing our data, we split our data into a train set and a test set to a 2:1 ratio for supervised learning. The accuracy for comorbid is 59.30% and the accuracy for control is 59.51%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 59.58% and 58.16% respectively. The precision, recall and f1-score for comorbid reward was 54.86%, 44.05% and 47.81%; 59.25%, 62.06% and 61.11% for comorbid punishment; 71.58%, 15.79% and 39.37% for control reward; 58.50%, 94.49% and 57.59% for control punishment.

5.3 K-Nearest Neighbors Algorithm Implementation

K-Nearest Neighbors is another machine learning method which is non-parametric. KNN is a useful technique for classification. We took our pre-processed data and split it into a 2:1 train and test set for supervised learning. The accuracy for comorbid is 59.90% and the accuracy for control is 58.72%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 58.43% and 57.83% respectively. The precision, recall and f1-score for comorbid reward was 63.42%, 27.43% and 42.70%; 60.34%, 61.42% and 60.11% for comorbid punishment; 57.35%, 44.23% and 48.31% for control reward; 61.24%, 72.06% and 65.61% for control punishment.

5.4 Decision Tree Algorithm Implementation

Decision Tree is another predictive modeling that is used in statistical analysis or in machine learning to predict data. We took our pre-processed data and split it into a 2:1 train and test set for supervised learning. The accuracy for comorbid is 60.85% and the accuracy for control is 58.26%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 62.00% and 57.88% respectively. The precision, recall and f1-score for comorbid reward was 54.45%, 53.95% and 53.77%; 63.33%, 60.85% and 61.72% for comorbid punishment; 54.23%, 53.29% and 52.24% for control reward; 61.63%, 63.01% and 61.85% for control punishment.

5.5 AdaBoost Algorithm Implementation

Adaptive Boosting or AdaBoost is another machine learning algorithm. AdaBoost is often referred to as the best out-of-the box classifier and so we chose to use this classifier as well. We took our pre-processed data and split it into a 2:1 train and test set for supervised learning. The accuracy for comorbid is 66.47% and the accuracy for control is 62.24%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 66.34% and 63.21% respectively. The precision, recall and f1-score for comorbid reward was 61.33%, 57.85% and 59.16%; 67.47%, 70.64% and 68.72% for comorbid punishment; 59.30%, 57.58% and 57.43% for control reward; 65.03%, 65.63% and 64.73% for control punishment.

5.6 Bagging Algorithm Implementation

Bootstrap aggregating or Bagging is another machine learning algorithm that is used in statistical analysis. We used this algorithm to classify our data as well. We took our pre-processed data and split it into a 2:1 train and test set for supervised learning. The accuracy for comorbid is 63.96% and the accuracy for control is 60.95%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 64.04% and 61.47% respectively. The precision, recall and f1-score for comorbid reward was 58.47%, 59.24% and 58.04%; 68.15%, 66.87% and 67.20% for comorbid punishment; 56.73%, 61.45% and 58.19% for control reward; 66.92%, 59.87% and 62.38% for control punishment.

5.7 Extra-tree Algorithm Implementation

Extremely Randomized Trees or ExtraTrees algorithm is an extension of Random Forest algorithm, which we used on our data. We took our pre-processed data and split it into a 2:1 train and test set for supervised learning. The accuracy for comorbid is 67.98% and the accuracy for control is 63.92%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 68.52% and 64.98% respectively. The precision, recall and f1-score for comorbid reward was 67.08%, 53.73% and 57.90%; 70.72%, 71.01% and 70.03% for comorbid punishment; 61.69%, 52.04% and 55.00% for control reward; 67.17%, 74.33% and 69.98% for control punishment.

5.8 Gradient Boosting Algorithm Implementation

Gradient Boosting is a machine learning technique that is well known for its faster prediction speed. We took our pre-processed data and split it into a 2:1 train and test set for supervised learning. The accuracy for comorbid is 68.07% and the accuracy for control is 64.34%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 66.20% and 63.92% respectively. The precision, recall and f1-score for comorbid reward was 64.49%, 57.19% and 60.12%; 68.10%, 70.54% and 68.94% for comorbid punishment; 63.92%, 56.76% and 57.65% for control reward; 67.01%, 71.37% and 68.26% for control punishment.

5.9 Random Forest Algorithm Implementation

Random Decision Forest or Random Forest is another machine learning algorithm which is usually better than Decision Trees algorithm. We took our pre-processed data and split it into a 2:1 train and test set for supervised learning. The accuracy for comorbid is 67.79% and the accuracy for control is 64.86%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 68.34% and 64.78% respectively. The precision, recall and f1-score for comorbid reward was 65.88%, 53.73% and 57.39%; 69.33%, 70.04% and 69.04% for comorbid punishment; 64.61%, 52.94% and 56.04% for control reward; 66.37%, 71.66% and 68.23% for control punishment.

5.10 Gaussian Naïve Bayes Algorithm Implementation

Gaussian Naive Bayes is a simple classifier with strong independent assumptions and based on Bayes' theorem. We took our pre-processed data and split it into a 2:1 train and test set for supervised learning. The accuracy for comorbid is 68.27% and the accuracy for control is 64.13%. After that, We ran 10-fold cross validation on both comorbid and control data and the accuracy was 68.35% and 65.03% respectively.

The precision, recall and f1-score for comorbid reward was 67.10%, 53.92% and 58.12%; 69.53%, 70.85% and 69.55% for comorbid punishment; 60.38%, 52.57% and 56.90% for control reward; 66.08%, 74.52% and 69.21% for control punishment.

5.11 Result and Analysis

We evaluated the performance of our models to see how accurately each model were able to predict patient's condition. We took four parameters from confusion matrix which are TP, FP, TN and FN and observed our models precision, recall and f1-score based on these four parameters.

5.11.1 Accuracy of Comorbid and Control Patients

In figure 5.1, we see the accuracy of different classifiers on comorbid Patients data. As it is seen from the bar chart, 4 classifiers- ETC, GBC, RFC and GNBC achieved more than 67%. However, GNBC gives the highest accuracy from the list of classifiers which is 68.27%. So, GNBC is more effective for predicting comorbid patients.

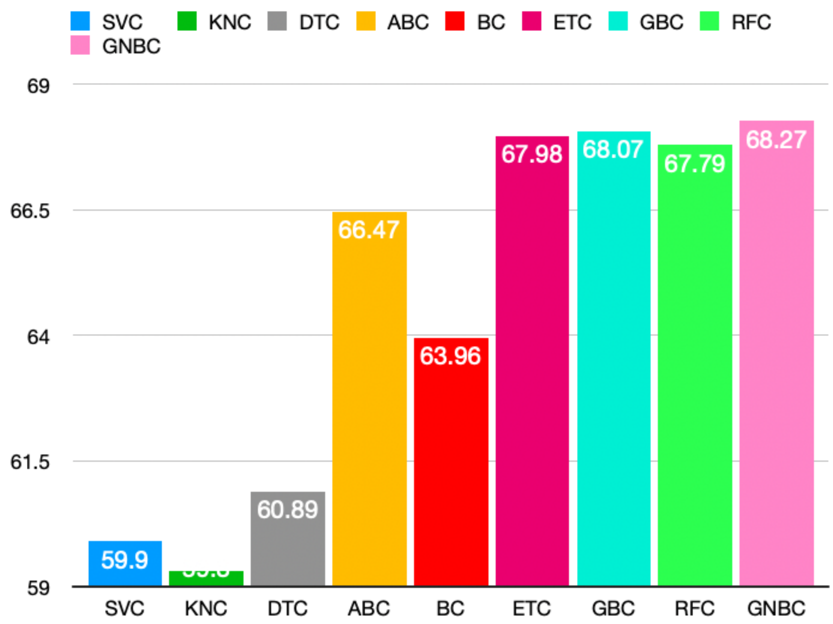


Figure 5.1: **Accuracy of Different Classifiers on Comorbid Patients**

In figure 5.2, for control patients, 4 classifiers- GNBC, ETC, GBC and RFC achieved more than 62% accuracy. However, for predicting control patients, ETC gives the most accuracy from other classifiers which is 66.5%.

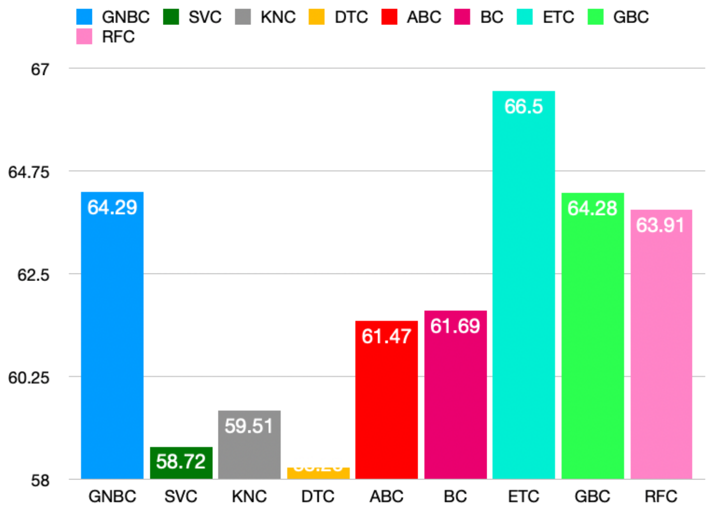


Figure 5.2: **Accuracy of Different Classifiers on Control Patients**

5.11.2 Precision (Reward and Punishment) of Comorbid and Control Patients

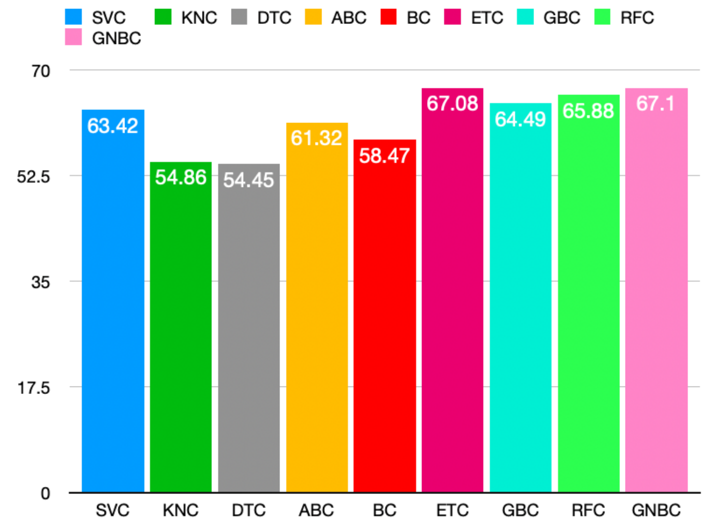


Figure 5.3: **Precision (Yes/Reward) of Different Classifiers on Comorbid Patients**

In figure 5.3, we have shown the precision for reward learning of comorbid patients which is related to positive feedback learning. Here, we see only 3 classifiers having a precision over 65% - ETC, GNBC and RFC. Among these 3, GNBC gives the highest precision for reward processing of comorbid patients which is 67.10%.

In figure 5.4, we have shown the precision for punishment learning of comorbid patients which is related to negative feedback learning. Here, we see only 6 classifiers having a precision over 65% - ABC, BC, ETC, GBC, GNBC and RFC. Among these

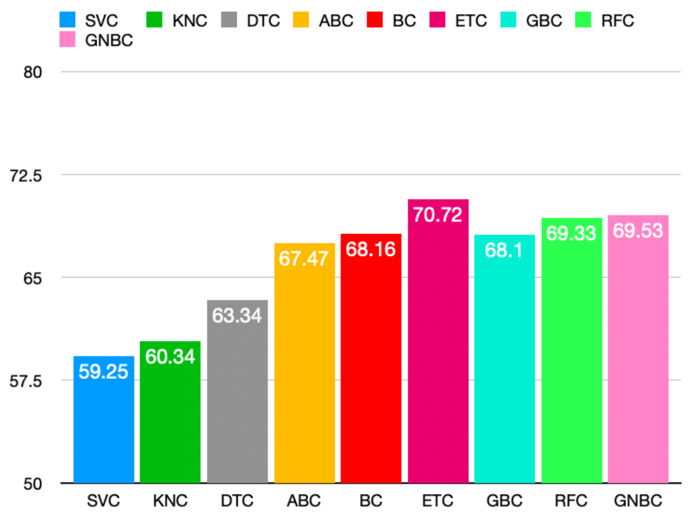


Figure 5.4: **Precision (No/Punishment) of Different Classifiers on Comorbid Patients**

6, ETC gives the highest precision for punishment processing of comorbid patients which is 70.72%.

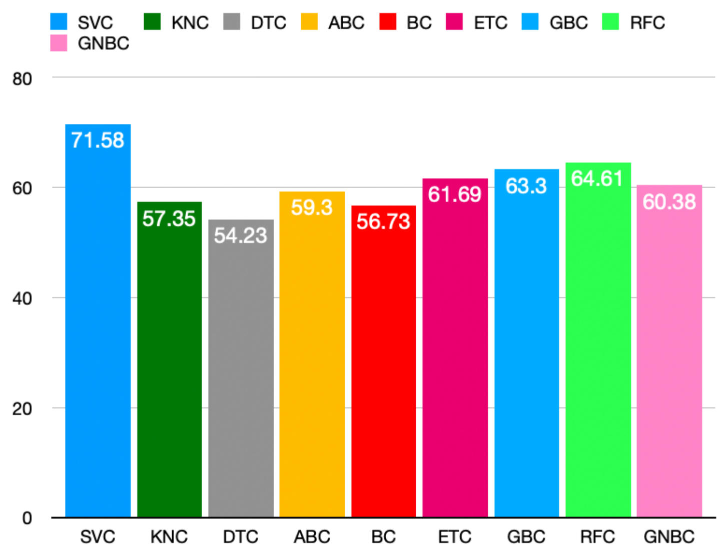


Figure 5.5: **Precision (Yes/Reward) of Different Classifiers on Control Patients**

In figure 5.5, we have shown the precision for reward learning of control patients. Here, we see 4 classifiers having a precision over 60% - SVM, ETC, GBC and RFC. Among these 4, SVM gives the highest precision for reward processing of control patients which is 71.58%.

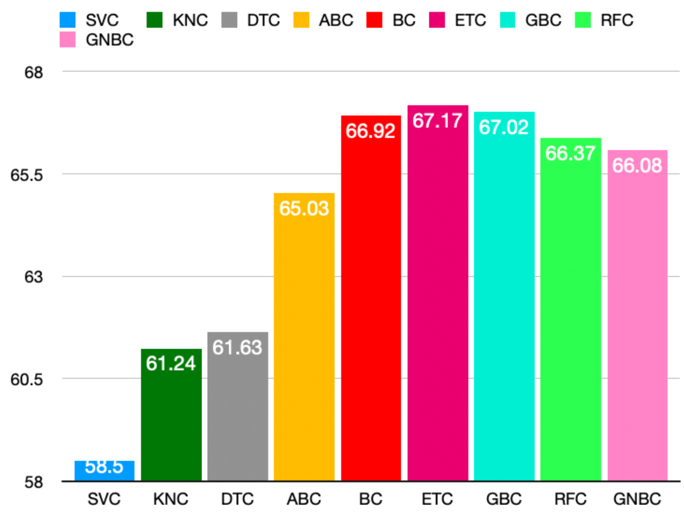


Figure 5.6: **Precision (No/Punishment) of Different Classifiers on Control Patients**

In figure 5.6, we have shown the precision for punishment learning of control patients which is related to negative feedback learning. Among these classifiers, ETC gives the highest precision for punishment processing of comorbid patients which is 67.17%.

5.11.3 Recall (Reward and Punishment) of Comorbid and Control Patients

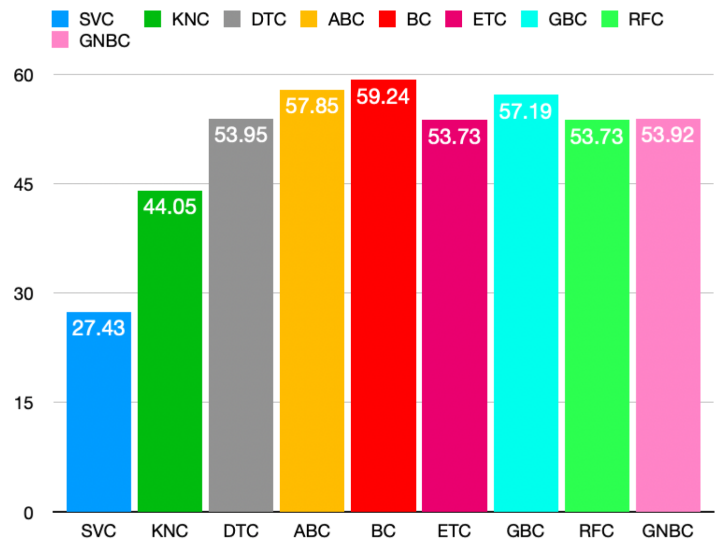


Figure 5.7: **Recall (Yes/Reward) of Different Classifiers on Comorbid Patients**

In figure 5.7, we have generated chart for reward learning of comorbid patients which is related to positive feedback learning. Among these classifiers, BC gives the highest precision for reward processing of comorbid patients which is 59.24%.

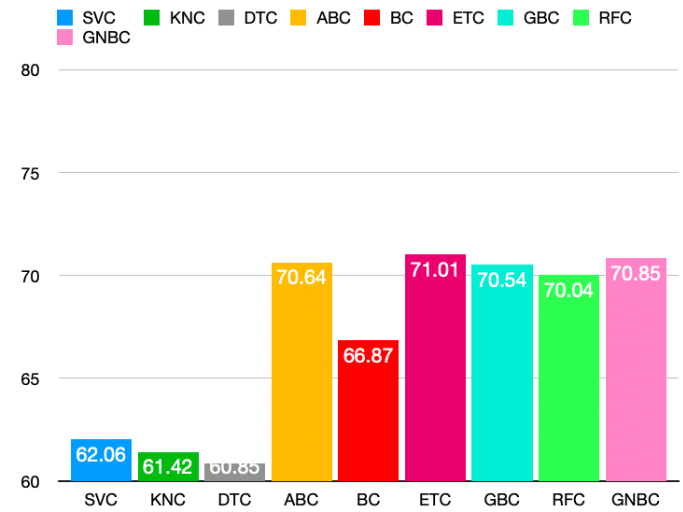


Figure 5.8: **Recall (No/Punishment) of Different Classifiers on Comorbid Patients**

In figure 5.8, we have shown the recall for punishment learning of control patients which is related to negative feedback learning. Among these classifiers, 5 classifiers gives good accuracy which is above 70%. However, ETC gives the highest precision for punishment processing of comorbid patients which is 71.01%.

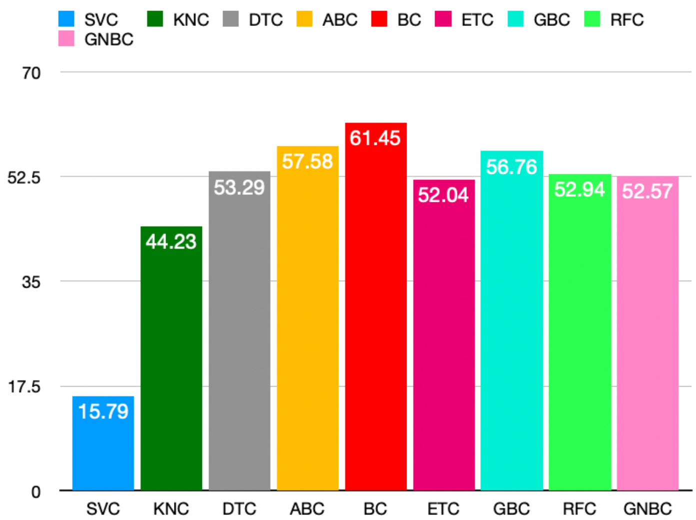


Figure 5.9: **Recall (Yes/Reward) of Different Classifiers on Control Patients**

In figure 5.9, we have shown the recall of control patients which is related to positive feedback learning. Among these classifiers, BC gives the highest precision for reward processing of comorbid patients which is 61.45%.

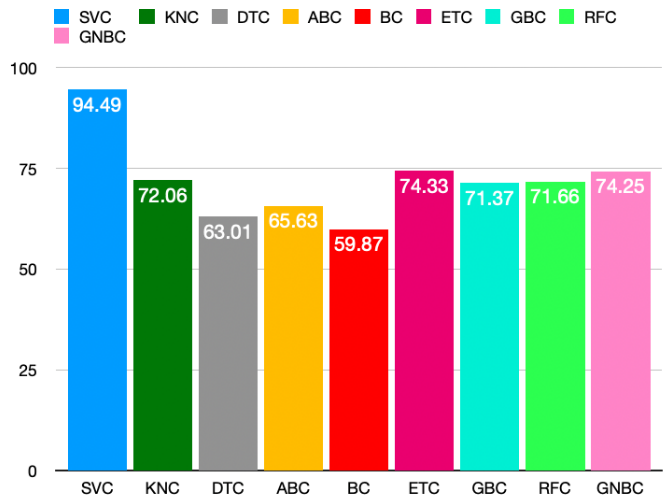


Figure 5.10: **Recall (No/Punishment) of Different Classifiers on Control Patients**

In figure 5.10, we have shown the recall for punishment learning of control patients which is related to negative feedback learning. Among the classifiers, 6 classifiers give a score above 70% - SVC, KNC, ETC, GBC, RFC, GNBC. SVC gives the highest precision for punishment processing of control patients which is 94.49%.

5.11.4 F1-Score (Reward and Punishment) of Comorbid and Control Patients

In figure 5.11, we have generated a bar chart to show F1-score of the classifiers on Comorbid Patients data on reward/ positive feedback learning. We see, only 6 classifiers have an F1-score above 55% - GNBC, ABC, BC, ETC, GBC and RFC. And among these 6 classifiers, GBC has the highest F1-score of 60.12%.

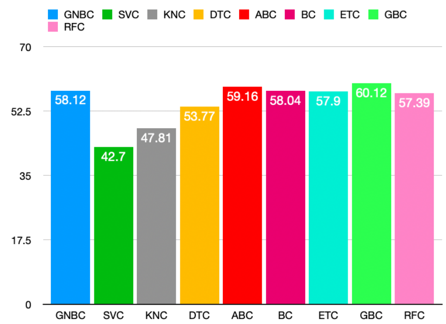


Figure 5.11: **F1-Score (Yes/Reward) of Different Classifiers on Comorbid Patients**

In figure 5.12, we have generated a bar chart to show F1-score of the classifiers on Comorbid Patients data on punishment/ negative feedback learning. We see, only 6 classifiers have an F1-score above 65% - ABC, BC, ETC, GBC, RFC and GNBC. And among these 6 classifiers, ETC has the highest F1-score of 70.03%.

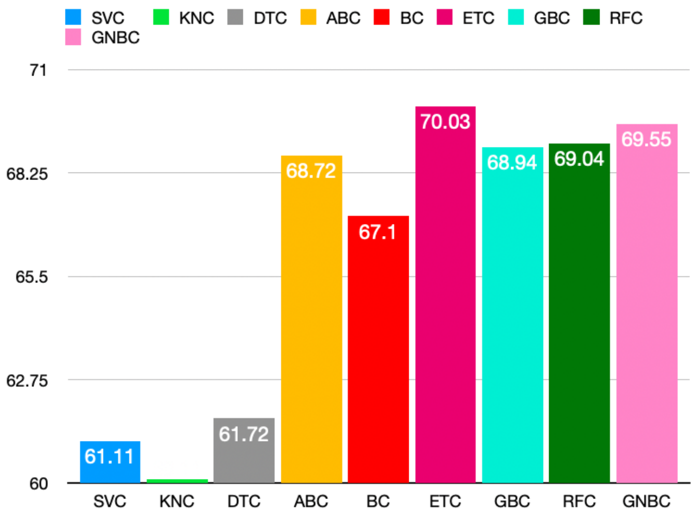


Figure 5.12: **F1-Score (No/Punishment) of Different Classifiers on Comorbid Patients**

In figure 5.13, we have created a bar chart for F1-score of the classifiers on Control Patients data on reward/ positive feedback learning. We see, only 6 classifiers have an F1-score above 55% - ABC, BC, ETC, GBC, RFC and GNBC. And among these 6 classifiers, BC has the highest F1-score of 58.19%.

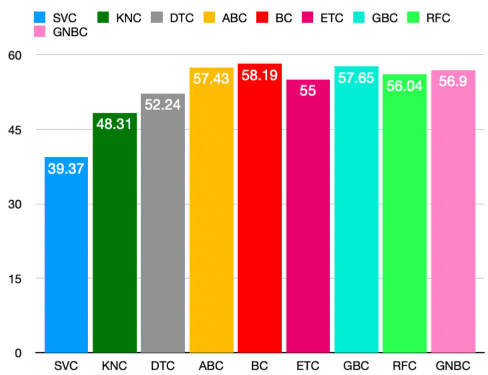


Figure 5.13: **F1-Score (Yes/Reward) of Different Classifiers on Control Patients**

In figure 5.14, we have generated a bar chart to show F1-score of the classifiers on Control Patients data on punishment/ negative feedback learning. We see, only 6 classifiers have an F1-score above 65% - SVC, KNC, ETC, GBC, RFC and GNBC. And among these 6 classifiers, SVC has the highest F1-score of 71.51%.

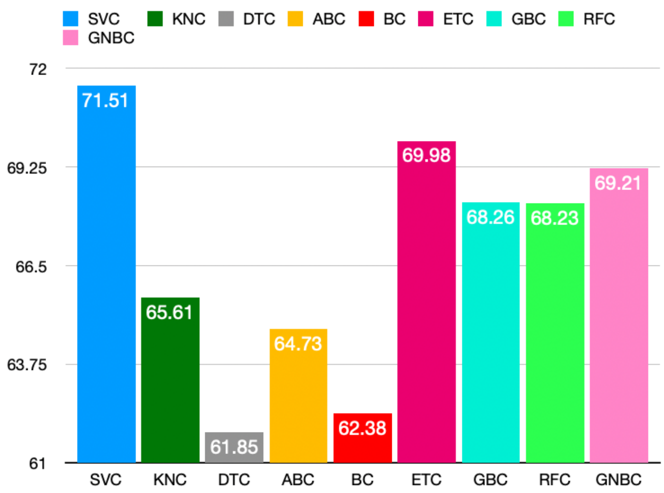


Figure 5.14: **F1-Score (No/Punishment) of Different Classifiers on Control Patients**

5.11.5 10-Fold Cross Validation for Comorbid and Control Patients

In figure 5.15, we have shown the 10-fold cross validation score of our models on the comorbid patients data where random state was 42. Here we see, 5 classifiers scored more than 65% - ABC, ETC, GBC, RFC and GNBC. Among these 5 classifiers, GNBC has the highest score 68.52%.

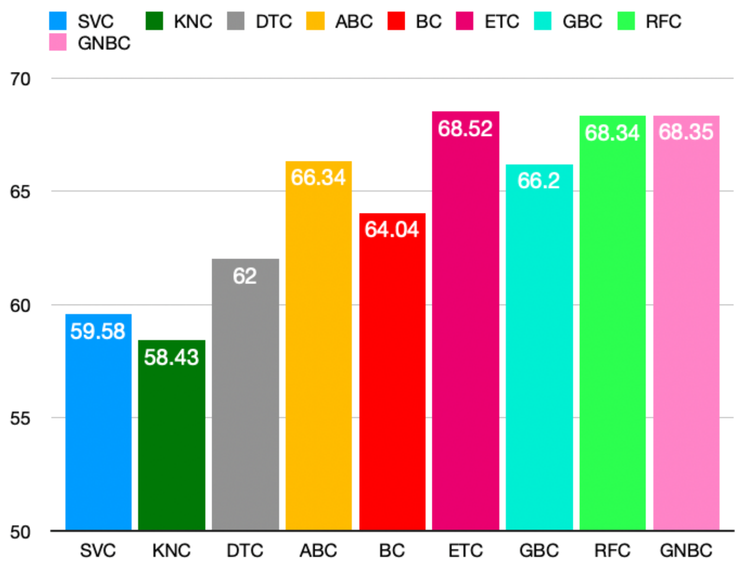


Figure 5.15: **K-10 Cross Validation score of Different Classifiers on Comorbid Patients**

In figure 5.16, we have generated bar chart of 10 fold cross validation for control patients where random state was 42. Here we see, from 9 classifiers, only 5 classifiers

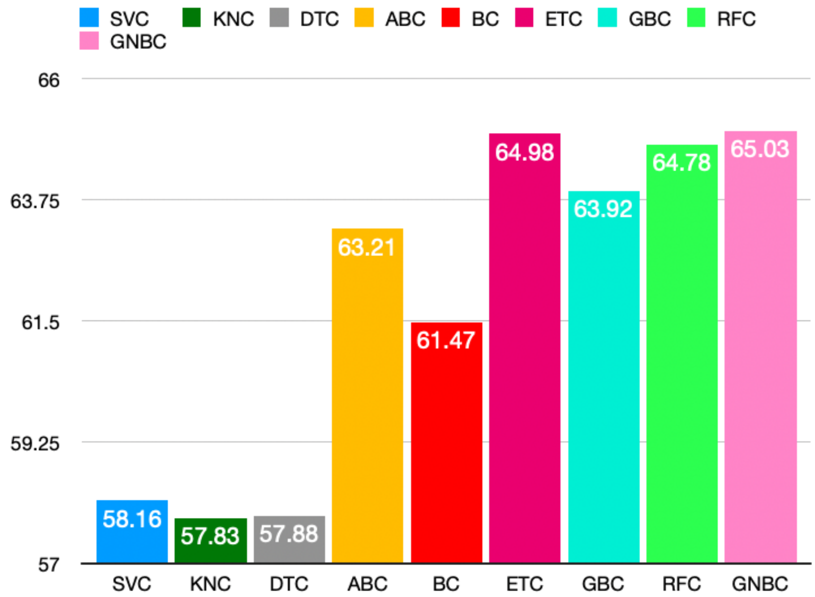


Figure 5.16: **K-10 Cross Validation score of Different Classifiers on Control Patients**

have scored more than 60% which are, ABC, BC, ETC, RFC, GBC, GNBC. howerver, among these 6 classifiers, GNBC accuracy for control patients which is 65.03%.

5.11.6 Accuracy Chart for Models

	Comorbid Patients Accuracy (%)	10-fold cross validation for Comorbid Patients(%)	Control Patients Accuracy (%)	10-fold cross validation for Control Patients (%)
SVC	59.9	59.58	58.72	58.16
KNC	59.30	58.43	59.51	57.83
DTC	60.89	62	58.26	57.88
ABC	66.47	66.34	61.47	63.21
BC	63.96	64.04	61.69	61.47
ETC	67.98	68.52	66.50	64.98
GBC	68.07	66.2	64.28	63.92
RFC	67.79	68.34	63.91	64.78
GNBC	68.27	68.35	64.29	65.03

Table 5.1: **Accuracy and Cross Validation score of all the models over Comorbid and Control patients data**

In table 5.1, We see the accuracy of all the classifiers on comorbid patients data and control patients data as well as the 10-fold cross validation score of comorbid patients data and control patients data. Here, we observe that Gaussian Naive Bayes Classifier or GNBC has the highest accuracy for comorbid patients data which is 68.27%. Though GNBC scored the highest, it is also observable that GBC, ETC and RFC was so close to GNBC with an accuracy of 68.07%, 67.98% and 67.79%

respectively. On the other hand, ExtraTrees Classifier or ETC scored the highest accuracy of 66.50% on control patients data however, GNBC, GBC and RFC were the closest with an accuracy of 64.29%, 64.28% and 63.91% respectively. It is also noticeable from the above bar charts that in case of precision, GNBC has the highest (67.10%) precision on comorbid patients on reward learning, ETC has the highest (70.72%) precision on comorbid patients on punishment learning, SVM has the highest (71.58%) precision on control patients on reward learning, ETC has highest (67.17%) precision on control patients on punishment learning. Additionally, in case of recall, BC has the highest recall of 59.24% over comorbid patients on reward learning, GNBC has the highest recall of 70.85% over comorbid patients on punishment learning, BC has the highest recall of 61.45% over control patients on reward learning, SVC has the highest recall of 94.49% over control patients on punishment learning. Moreover, in case of F1-score, GBC has the highest score of 60.12% over comorbid patients on reward learning, ETC has the highest score of 70.03% over comorbid patients on punishment learning, BC has the highest score of 58.19% over control patients on reward learning and SVC has the highest score of 71.51% over control patients on punishment learning.

We wanted to predict whether a person is experiencing mental stress from the feedback they are receiving while reward learning or punishment learning. And with our model, as we have seen above, we are able to classify whether a person is experiencing mental stress from their EEG signal at the time of reward or punishment processing.

Chapter 6

Conclusion and future Work

The main purpose of our study is to detect mental stress which is a common and serious issue in people of all age including school going children, teenagers, adults and also old people. Mental stress do not need any specific reason to be happened. It can be experienced from a very little incident to a very big issue. The consequence of it depends on how people handles it. When people loses control over his mental stress, it takes the person in another level of it. Mental stress can be declared as the root cause behind mental health issues. Depression and anxiety is one of the result of it and they are one of the major challenges in today's world. Due to the stress of family matters, personal things etc including the global challenge named Covid-19 pandemic, people are falling more into it. This issues has worked behind our study on mental stress and our objective on detecting the actual mental issue to prevent people from affected by depression and anxiety and taking major steps like suicide. However, to fulfill our target we have utilized data that have stored information on depression and anxiety from reward and punishment process and in order to detect depression and anxiety and to be ensured about the accuracy we have used nine classifiers. Due to the most dangerous risk Covid-19 pandemic, we had to collect data from online resources, so we hope that if further research is intended on this study after everything gets to normal, data will be collected physically from individuals ,schools, universities and hospitals from which information on other features of mental stress can be stored for future research purposes and we can detect mental health problems by using them. Moreover, as for data collection we had to rely on online available resources, we could collect data on depression and anxiety so if the study on this topic is intended we believe that we could detect other mental diseases occurred due to mental stress.

Bibliography

[1] W. C. S. Jr., "Medical definition of stress," DOI: https://www.medicinenet.com/stress/definition.htm.

[2] "Fight or flight response." DOI: https://www.psychologytools.com/resource/fight-or-flight-response/#:~:text=Information%20Handout,body%20to%20fight%20or%20flee..

[3] M. Clinic, "Chronic stress puts your health at risk," 2016.

[4] A. Pietrangelo, "The effects of stress on your body," Mar. 2020. DOI: https://www.healthline.com/health/stress/effects-on-body#1.

[5] K. Cherry, "5 surprising ways that stress affects your brain," Mar. 2020. DOI: https://www.verywellmind.com/surprising-ways-that-stress-affects-your-brain-2795040.

[6] K. Bruno, "Stress and depression." DOI: https://www.webmd.com/depression/features/stress-depression#1.

[7] "Eeg (electroencephalogram)." DOI: https://kidshealth.org/en/parents/eeg.html.

[8] E. Patterson, "Stress facts and statistics," Jun. 2020. DOI: https://www.therecoveryvillage.com/mental-health/stress/related/stress-statistics/#:~:text=According%20to%20The%20American%20Institute,that%20impacts%20their%20mental%20health.

[9] M. A. Mamun, M. S. Hossain, and M. D. Griffiths, "Mental health problems and associated predictors among bangladeshi students," *International Journal of Mental Health and Addiction*, pp. 1–15, 2019.

[10] H. Ritchie and M. Roser, "Mental health," *Our World in Data*, 2018.

[11] M. C. Staff, "Anxiety disorders." DOI: https://www.mayoclinic.org/diseases-conditions/anxiety/symptoms-causes/syc-20350961?p=1.

[12] ——, "Insomnia." DOI: https://www.mayoclinic.org/diseases-conditions/insomnia/symptoms-causes/syc-20355167.

[13] PTI, "German minister commits suicide after 'virus crisis worries'," Mar. 2020. DOI: https://economictimes.indiatimes.com/news/international/world-news/german-minister-commits-suicide-after-virus-crisis-worries/articleshow/74877039.cms.

[14] "Mental health statistics: Stress." DOI: https://www.mentalhealth.org.uk/statistics/mental-health-statistics-stress.

[15] M. Olfson, S. C. Marcus, and D. Shaffer, "Antidepressant drug therapy and suicide in severely depressed children and adults: A case-control study," *Archives of general psychiatry*, vol. 63, no. 8, pp. 865–872, 2006.

[16] R. Mahajan and D. Bansal, "Depression diagnosis and management using eeg-based affective brain mapping in real time," *International Journal of Biomedical Engineering and Technology*, vol. 18, no. 2, pp. 115–138, 2015.

[17] U. R. Acharya, V. K. Sudarshan, H. Adeli, J. Santhosh, J. E. Koh, and A. Adeli, "Computer-aided diagnosis of depression using eeg signals," *European neurology*, vol. 73, no. 5-6, pp. 329–336, 2015.

[18] J. J. Newson and T. C. Thiagarajan, "Eeg frequency bands in psychiatric disorders: A review of resting state studies," *Frontiers in human neuroscience*, vol. 12, p. 521, 2019.

[19] N. Eshel and J. P. Roiser, "Reward and punishment processing in depression," *Biological psychiatry*, vol. 68, no. 2, pp. 118–124, 2010.

[20] B. A. Katz, K. Matanky, G. Aviram, and I. Yovel, "Reinforcement sensitivity, depression and anxiety: A meta-analysis and meta-analytic structural equation model," *Clinical Psychology Review*, p. 101 842, 2020.

[21] J. Soler, D. Vega, M. Elices, A. Feliu-Soler, À. Soto, A. Martın-Blanco, J. Marco-Pallarés, R. Torrubia, and J. C. Pascual, "Testing the reinforcement sensitivity theory in borderline personality disorder compared with major depression and healthy controls," *Personality and Individual Differences*, vol. 61, pp. 43–46, 2014.

[22] N. E. Hundt, R. O. Nelson-Gray, N. A. Kimbrel, J. T. Mitchell, and T. R. Kwapil, "The interaction of reinforcement sensitivity and life events in the prediction of anhedonic depression and mixed anxiety-depression symptoms," *Personality and Individual Differences*, vol. 43, no. 5, pp. 1001–1012, 2007.

[23] H. Cai, J. Han, Y. Chen, X. Sha, Z. Wang, B. Hu, J. Yang, L. Feng, Z. Ding, Y. Chen, *et al.*, "A pervasive approach to eeg-based depression detection," *Complexity*, vol. 2018, 2018.

[24] Y. Li, B. Hu, X. Zheng, and X. Li, "Eeg-based mild depressive detection using differential evolution," *IEEE Access*, vol. 7, pp. 7814–7822, 2018.

[25] S. Mahato and S. Paul, "Electroencephalogram (eeg) signal analysis for diagnosis of major depressive disorder (mdd): A review," in *Nanoelectronics, Circuits and Communication Systems*, Springer, 2019, pp. 323–335.

[26] M. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, "Depression detection from social network data using machine learning techniques," *Health information science and systems*, vol. 6, no. 1, p. 8, 2018.

[27] A. Kumar, A. Sharma, and A. Arora, "Anxious depression prediction in real-time social data," in *International Conference on Advances in Engineering Science Management & Technology (ICAESMT)-2019, Uttaranchal University, Dehradun, India*, 2019.

[28] J. F. Cavanagh, A. W. Bismark, M. J. Frank, and J. J. Allen, "Multiple dissociations between comorbid depression and anxiety on reward and punishment processing: Evidence from computationally informed eeg," *Computational Psychiatry*, vol. 3, pp. 1–17, 2019.

[29] A. R. Subhani, W. Mumtaz, M. N. B. M. Saad, N. Kamel, and A. S. Malik, "Machine learning framework for the detection of mental stress at multiple levels," *IEEE Access*, vol. 5, pp. 13 545–13 556, 2017.

[30] M. Teesson *et al.*, "Comorbid mental disorders and substance use disorders: Epidemiology, prevention and treatment," 2004.

[31] B. Salcedo, "The comorbidity of anxiety and depression," *National Alliance on Mental Illness*, vol. 19, 2018.

[32] A. Cuncic, "The comorbidity of two disorders," Mar. 2020. DOI: https://www.verywellmind.com/what-is-comorbidity-3024480.

[33] "Understand the facts." DOI: https://adaa.org/understanding-anxiety.

[34] L. Gadye, "What part of the brain deals with anxiety? what can brains affected by anxiety tell us?," Jun. 2018. DOI: https://www.brainfacts.org/diseases-and-disorders/mental-health/2018/what-part-of-the-brain-deals-with-anxiety-what-can-brains-affected-by-anxiety-tell-us-062918.

[35] E. I. Martin, K. J. Ressler, E. Binder, and C. B. Nemeroff, "The neurobiology of anxiety disorders: Brain imaging, genetics, and psychoneuroendocrinology," *Psychiatric Clinics*, vol. 32, no. 3, pp. 549–575, 2009.

[36] "What are anxiety disorders?." DOI: https://www.psychiatry.org/patients-families/anxiety-disorders/what-are-anxiety-disorders.

[37] "Depression statistics." DOI: https://www.dbsalliance.org/education/depression/statistics/.

[38] E. Cirino, "The effects of depression on the brain," Feb. 2017. DOI: https://www.healthline.com/health/depression/effects-brain.

[39] "Types of depression." DOI: https://www.beyondblue.org.au/the-facts/depression/types-of-depression.

[40] M. Reuter, A. J. Cooper, L. D. Smillie, S. Markett, and C. Montag, "A new measure for the revised reinforcement sensitivity theory: Psychometric criteria and genetic validation," *Frontiers in systems neuroscience*, vol. 9, p. 38, 2015.

[41] A. D. Pickering and P. J. Corr, "Ja gray's reinforcement sensitivity theory (rst) of personality," *The SAGE handbook of personality: Theory and assessment personality measurement and testing*, vol. 2, pp. 239–255, 2008.

[42] "Reinforcement sensitivity theory," DOI: https://en.wikipedia.org/wiki/Reinforcement_sensitivity_theory.

[43] T. Wächter, O. V. Lungu, T. Liu, D. T. Willingham, and J. Ashe, "Differential effect of reward and punishment on procedural learning," *Journal of Neuroscience*, vol. 29, no. 2, pp. 436–443, 2009.

[44] L. D. Smillie, "Extraversion and reward processing," *Current Directions in Psychological Science*, vol. 22, no. 3, pp. 167–172, 2013.

[45] K. Unger, S. Heintz, and J. Kray, "Punishment sensitivity modulates the processing of negative feedback but not error-induced learning," *Frontiers in human neuroscience*, vol. 6, p. 186, 2012.

[46] J. F. Cavanagh, M. J. Frank, and J. J. Allen, “Social stress reactivity alters reward and punishment learning,” *Social cognitive and affective neuroscience*, vol. 6, no. 3, pp. 311–320, 2011.

[47] C. Mallapragada, “Classification of eeg signals of user states in gaming using machine learning,” 2018.

[48] Y. Ma, X. Ding, Q. She, Z. Luo, T. Potter, and Y. Zhang, “Classification of motor imagery eeg signals with support vector machines and particle swarm optimization,” *Computational and mathematical methods in medicine*, vol. 2016, 2016.

[49] S. Jukic, M. Saracevic, A. Subasi, and J. Kevric, “Comparison of ensemble machine learning methods for automated classification of focal and non-focal epileptic eeg signals,” *Mathematics*, vol. 8, no. 9, p. 1481, 2020.

[50] M. A. Rahman, Y. Al-Saggaf, and T. Zia, “A data mining framework to predict cyber attack for cyber security,” in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, IEEE, 2020, pp. 207–212.

[51] M. A. Rahman, B. Honan, T. Glanville, P. Hough, and K. Walker, “Using data mining to predict emergency department length of stay greater than 4 hours: Derivation and single-site validation of a decision tree algorithm,” *Emergency Medicine Australasia*, vol. 32, no. 3, pp. 416–421, 2020.

[52] J. Hu, “Automated detection of driver fatigue based on adaboost classifier with eeg signals,” *Frontiers in computational neuroscience*, vol. 11, p. 72, 2017.

[53] K. Soriano, “Depression central: Tell me all i need to know about depression.” DOI: https://www.psycom.net/depression.central.html.

[54] Z. A. Chowdhury, D. N. Alam, M. A. F. H. B. Nahid, M. A. Rahman, and M. Z. Parvez, “Detection of modulated motor cortex using anodal and cathodal tdcs based neurofeedback,” in *The 19th International Conference on Machine Learning and Cybernetics (ICMLC)*, IEEE Xplore, 2020.

[55] S. Gandhi and B. Sarkar, *Essentials of mineral exploration and evaluation.* Elsevier, 2016.

[56] S. Aydın, H. M. Saraoğlu, and S. Kara, “Log energy entropy-based eeg classification with multilayer neural networks in seizure,” *Annals of biomedical engineering*, vol. 37, no. 12, p. 2626, 2009.

[57] S. Glen, “Shannon entropy,” Aug. 2020. DOI: https://www.statisticshowto.com/shannon-entropy/.

[58] Ł. Zaborowska, “Accuracy calculator.” DOI: https://www.omnicalculator.com/statistics/accuracy.

[59] M. A. Rahman, “Automatic selection of high quality initial seeds for generating high quality clusters without requiring any user inputs,” *PhD Desertification, School of Computing and Mathematics, Charles Sturt University*, 2014.

[60] K. P. Shung, “Accuracy, precision, recall or f1?, 2018,” *URL https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9*,

[61] A. Wilson, “A brief introduction to supervised learning,” Sep. 2019. DOI: https://towardsdatascience.com/a-brief-introduction-to-supervised-learning-54a3e3932590.