Human Stress Classification Using Electroencephalogram (EEG) in Response to Standup Comedy Clips

By

Kiran Firdous

00000329163

Supervisor

Dr. Shahzad Rasool

Department of Computational Sciences and Engineering

School of Interdisciplinary Engineering and Sciences (SINES)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

March 2023

Human Stress Classification Using Electroencephalogram (EEG) in Response to Standup Comedy Clips

By

Kiran Firdous

00000329163

Supervisor

Dr. Shahzad Rasool

A thesis submitted in conformity with the requirements for the degree of *Master of Science* in Computational Science and Engineering

Department of Computational Sciences and Engineering

School of Interdisciplinary Engineering and Sciences (SINES)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

March 2023

Declaration

I, *Kiran Firdous* declare that this thesis titled "Human Stress Classification Using EEG in Response to Stand-up Comedy Clips" and the work presented in it are my own and has been generated by me as a result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a Master of Science degree at NUST
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at NUST or any other institution, this has been clearly stated
3. Where I have consulted the published work of others, this is always clearly attributed
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work
5. I have acknowledged all main sources of help
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself

Kiran Firdous,

00000329163

Copyright Notice

Dedicated to my exceptional parents and adored siblings whose tremendous support and cooperation led me to this wonderful accomplishment

Acknowledgments

- I am thankful to my Creator Allah Subhana-Watala to have guided me throughout this work at every step and for every new thought that Allah set up in my mind to improve it. Indeed I could have done nothing without Allah's priceless help and guidance. Whosoever helped me throughout the course of my thesis, whether my parents or any other individual was Allah will, so indeed none be worthy of praise but Allah.

- I am profusely thankful to my beloved parents who raised me when I was not capable of walking and continued to support me throughout every department of my life.

- I would also like to express special thanks to my supervisor Dr. Shahzad Rasool for his help throughout my thesis. I can safely say that I haven't learned any other science subject in such depth as the ones which he has taught. I appreciate his patience and guidance throughout the whole thesis.

- I would also like to thank my evaluation committee members, Dr. Muhammad Tariq Saeed and Dr. Mian Ilyas Ahmad, for being on my thesis guidance and for their support and cooperation throughout my coursework and research work.

- Finally, I am very thankful to my parents for supporting and loving me throughout my life journey. I would also like to express my gratitude to all the individuals, especially to Mr. Shah Faisal who has rendered valuable assistance to my study.

Abstract

Nearly everyone encounters stress at some point in their life. An individual's stress load can be estimated using a valid and reliable stress assessment. In this study, stand-up comedy clips in native and non-native languages are used as a stimulus to study the reduction in stress levels. The electroencephalogram (EEG), which reflects brain activity and is frequently employed in clinical diagnosis and biomedical research, serves as the main signal. An EEG dataset is generated from thirty participants using a single-channel Neurosky Mindwave 2 mobile headset. The electrical activity of the brain is captured as the participants watch various comedy clips. A state and trait anxiety questionnaire is used to obtain a subjective measure of stress level of participants.

The single-channel EEG data being an extremely noisy, non-stationary, and a non-linear signal is filtered using the Savitzky-Golay filter. Ten features from the wavelet, time-frequency, and time domains were used to classify stress using each domain. Long Short Term Memory (LSTM), Random Forest, eXtreme Gradient Boosting (XGB), and ExtraTree classifiers were used where the highest accuracy achieved was 84.29% with the ExtraTree classifier. Our findings indicate that only two classes (stressed, and Non-stressed) can be classified for a single-channel EEG device. Where non-native and native language comedy clips have obtained the maximum individual accuracy of 84.29% and 78.32%. It is evident from the results that English comedy has more influence on stress level reduction as compared to Urdu comedy.

Keywords – EEG, Machine Learning, Brain-Computer Interfaces, Comedy Clips, stress classification, personality traits

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the current era of technological advancement, "stress" is a subject of growing concern and has more recently been investigated in terms of its effects on mental health. The media has increasingly featured strategies for preventing and managing stress [1]. As stress has a significant impact on social, physical, and mental health, it continues to be a point of discussion. There are several uses for stress research, including measuring stress throughout the day, assessing stress to boost well-being and productivity at work, and delaying the development of major illnesses. This research area is equally beneficial for society and individuals. With the usage of an readily available electroencephalography (EEG) headset, this thesis encourages an objective understanding of stress by emphasizing the key features and regions of brain activity.

Stress negatively affects one's capacity to function as well as health since it makes one more susceptible to sickness and slows down the body's ability to recover [2]. A person's neurons in the hypothalamus are activated when they encounter a stressor. A corticotropin-releasing hormone (CRH) is discharged from the pituitary glands and subsequentially releases adrenocorticotropic hormone (ACTH). The blood carries ACTH, which has an impact on the adrenal glands. Cortisol, adrenaline, and norepinephrine are a few of the stress hormones that are released in response [3]. Stress causes the production of cortisol, which aids in a person's ability to deal with an immediate threat. Chronically high levels of stress may result in depressive symptoms. Human stress, which may also increase anxiety and mood problems, has been linked to depression, according to several studies [4]. Nearly 350 million individuals worldwide suffer from depression, and it appears that things are becoming worse in Pakistan [5]. In a study, 820 participants were contacted or interviewed, and out of them, about 46 percent said they had depression or symptoms associated with it [6]. There are several reasons for the rising prevalence of mental illness in Pakistan, including natural catastrophes, internal wars, economic instability, political unrest, increased rates of crimes and poverty, among others. Less than 350 psychiatrists, or 0.2 per 100,000 people, are estimated to exist in Pakistan by the World Health Organization (WHO) [7]. The condition of psychiatric wards and mental institutions is terrible, and there are only five psychiatric recognized hospitals in the nation. One of our most important goals should be to educate people about mental illness, but the harder challenge is to eliminate the stigma and guilt connected with it [3].

Psychological techniques have been used to evaluate stress, and these techniques heavily rely on human interaction. These techniques are known as observer-reported or self-reported measures in the literature [8]. When a person chooses to conceal their stress status out of shame, personal reason, or stigma, these objective measures become less successful in reflecting the amount of stress. Most people who have mental health problems go misdiagnosed, sometimes hiding their condition in shame and stigma. Technology has advanced to the point that EEG headgear is readily accessible as Commercial Off The Shelf (COTS) equipment in some form. As a result, stressors can use these EEG headsets to assess stress objectively. Such a metric might help a psychologist identify stress before it has major health consequences. It can give a direct assessment of brain activity and reduce the stigma associated with mental examinations in patients. This might also aid in studying how stress circumstances affect cognitive processes.

The study of human stress measurement is a fast-developing field in EEG data processing. There is literature on stress measurement using EEG in response to various stressors. The characterization of human stress based on baseline EEG has not received much attention. Particularly in poor nations with limited access to mental health institutions, wearable EEG stress measurement is of relevance. Thus, in these regions, the use of EEG as an objective measure for stress classification becomes more important. It is a relatively inexpensive and simple method for identifying stress. Human stress and mental states have been evaluated with success using machine learning algorithms. Generally speaking, classifiers for supervised and unsupervised learning may be trained using these techniques [2]. In classifying human stress, supervised learning algorithms have shown promising results, but they need labeling for classification, which is where stress evaluation techniques from psychology might be applicable. These evaluation techniques may include self-reporting, reporting from observers, or a mix of the two.

1.1 Motivation

Stress and anxiety are growing more and more dangerous for Pakistani youth. According to a report by Tribune Pakistan, over 34 percent of Pakistanis experience anxiety [9]. The stigma associated with common mental diseases persists in Pakistan despite the grave condition.

At various times in life, feeling stressed out for a short while is entirely acceptable. When one's productivity and health are negatively impacted by excessive stress or anxiety, problems start to occur. The way we act and think starts to change dramatically. Therefore, it is crucial to identify stress signs early on to prevent severe mental problems like depression. Therefore, to overcome this issue, we use stand-up comedy clips for the stress classification task, using EEG signals. Because EEG signals are extensively used in stress detection research and are both inexpensive and non-invasive modalities.

1.2 Problem Statement

Our research aims to profile participants by invoking stress responses through Native Language Urdu, and Non-Native language English, standup comedy clips. We use the Neurosky Mindwave Mobile 2 EEG device to capture brain signals and apply machine learning algorithms to classify stress and assess its relationship with the personality of participants.

1.3 Objectives

The primary goal of the thesis is to use a commercially accessible EEG-based brain-computer interface to evaluate and classify human stress. The objectives of this study are:

- To record a real-time dataset of EEG signals using Native Language Urdu, and Non-Native language English stand-up comedy clips.
- Utilize statistical and machine learning techniques to classify human stress using EEG-based signal recordings.
- To uncover the mapping between each participant's personality based on their level of stress.

1.4 Thesis Organization

The thesis is organized as follows: Chapter 2 details the previous research carried out for stress and personality analysis using EEG, the stimuli applied for stress classification, and machine learning algorithms and the accuracy achieved so far. Chapter 2 also identifies the research gap. Chapter 3 discusses the proposed methodology and the implementation details, including the tools used for data collection, and noise removal filters for data cleaning offered by machine learning algorithms used on the data in this research. Chapter 4 discusses the results achieved after implementing the proposed methodology, while Chapter 5 gives the conclusions drawn based on those results. It also gives future direction as to how this work can be enhanced.

Chapter 2

Literature Review

Through a discussion of various stress categories, this chapter provides the background information necessary to understand human stress. Stress is measured in a variety of ways, including psychologically, physically, and physiologically. An EEG is a physiological measurement that provides a clear indication of brain activity. EEG signal processing is a new and exciting area of brain-computer interface technology. This chapter discusses the types and paradigms of brain-computer interfaces by presenting a short survey of the BCI headsets and software tools that are now available.

2.1 What is Stress?

Unavoidable pressures enter daily life because of society's and technology's rapid advancements. Life itself presents physical and mental difficulties that are presented by life itself. Numerous pressures might cause someone to feel stressed out in their daily lives. According to Selye, stress is the non-specific physiological response to every desire for change [10]. According to the standard definition of stress, this reaction pattern is complex and frequently consists of behavioral, psychological, and cognitive components [11]. It is a description of the decline in how a body processes, evaluates, and responds to sensory stimuli [12]. An additional description is provided as Environmental pressures surpass an organism's ability for adaptation, causing biological and psychological alterations that may put people at risk for illnesses [13]. The phrase "computational stress" was very recently coined, and it implies measuring stress at a given moment in time while taking stress symptoms into account [8].

2.2 Types of Human Stress

The severity of the sickness and its symptoms are used to classify the different forms of stress. Each kind has distinct indications and effects. Chronic, Acute, episodic, and long-term stress are examples of these categories [14].

2.2.1 Chronic Stress

It can be described as an ongoing issue lasting for a minimum duration of four weeks, with at least one significant life event before commencement [15]. It can be brought on by exposure to horrific events as well as by daily tensions that aren't properly managed. Ignoring chronic stress can have serious consequences, including depression, cardiovascular disorders, and anxiety. People's health is harmed by chronic stress on a daily and even yearly basis. Chronic stress is brought on by dysfunctional families, poverty, unpleasant marriages, and undesirable jobs. When someone is confined in a bad circumstance, it is set off. It happens because of surpassing demands for what appears to be an endless amount of time.

Nine basic types of chronic stress exist i.e., demands, under-load, structural limits, unpredictability, complexity, under-reward, choice restriction, conflict, and resource deprivation [16]. The danger is the persistent possibility of damage that cannot be prevented or managed. Chronic stress develops when a person's reputation is in danger. A constant and independent aspect of the actual threat is how it is perceived. It draws attention to the stressful aspect of the issue. When a person is under the strain of several urgent, uncontrollable duties, and independence, that cannot be avoided, demands and overload occur. This kind of overload is one of several factors that contribute to chronic stress, including lack of time for oneself, a disproportionate quantity of work compared to other people, and unrealistic expectations from others [17].

Despite having distinct qualities, complexity is a type of excessive demand. Demands of this kind are interrelated, creating a level of complexity that causes chronic stress. Social duties give a complicated schedule to fulfill, which puts someone under constant stress. As an issue that is closely tied to complexity, uncertainty is said to exist. Scaling back the impacts of complexity is uncertainty. If uncertainty lasts more than expected, stress will be brought on. Uncertainty turns into chronic stress when getting the problem resolved takes priority above getting the desired outcome.

Structure-related characteristics are present in underload. When someone lacks an alternative, boredom is their only available condition. There are several types of demands, such as monotonous work that is slow-moving and frequent. Problems arise when an individual is forced to make choices from a limited range due to structural restrictions in their social context. These issues can take on a variety of shapes, from how regulations are interpreted to the reality of severe societal disadvantages like discrimination. Access restrictions are likewise seen as a type of stress. In many circumstances, discrimination is widely seen as stressful.

2.2.2 Acute Stress

The most prevalent type of stress is acute stress, which is driven by unfamiliar, frightening, unpredictable, and unmanageable conditions. Acute stress is typically brought on by the demands of the future and also the past. Examples of everyday events that can be stressful include losing a significant contract or being on the verge of a deadline. Instantaneous stress can be beneficial since it causes the body to generate stress hormones to deal with the traumatic period. However, if these stress hormones are released in abundance, the body's defense mechanisms may get worn out. Acute stress is less harmful than long-term stress since it lasts for a shorter period of time. However, excessive short-term stress can cause physical and psychological issues. such as headaches, discomfort, digestive problems, and other symptoms that indicate stress. Most acute stress patients may quickly recognize its symptoms.

Episodic Acute Stress

It is defined as acute stress that frequently occurs and has self-inflicted demands and disorganization as its main symptoms. During episodic acute stress, one might also notice short-temperedness, impatience, nervous energy, tension, and worry. Occasionally, irritation can result in aggression, which can harm social relationships. The workplace might thus become unpleasant for the sufferer. In this condition, patients forecast doom in each situation. People suffering from episodic acute stress view the world as scary and joyless, are reluctant to change, feel endless worry, over-arousal, and despair.

Under-reward is a concept that comes from Equity Theory [18]. According to this idea of social exchange, each relationship's participants should get an equivalent number of outputs relative to inputs. The phenomenon of under-reward is described as occurring when playing the victim is more stressful than playing the opponent. Every aspect of social life, business, and family is fraught with conflict. They happen because of various goals and morals. If the parties cannot agree on shared aims and goals, they tend to last for a longer amount of time. Chronic stress is caused by both the anxiety of bringing up the dispute and the problem itself.

2.3 Measures of Stress

The symptoms associated with the stress response could be used to measure stress in people in a variety of ways. Humans exhibit signs of stress reaction on the psychological, physical, and physiological levels. Measures of stress may be divided into three categories based on these symptoms. In the subsections below, the measures are explained briefly.

2.3.1 Psychological Stress

Psychological measurements are those that do not require activation and instead focus on the mental state. The measurements obtained from physiological measures or sensors are validated using these stress measures. Measurements of abstract ideas like language, intellect, and emotion are the focus of psychology. It is commonly acknowledged that psychological factors play a role in stress response. As a result, it is considered that psychological surveys can be used to measure stress [8].

Numerous surveys have been created by psychologists that can evaluate a person's degree of stress based on their reported symptoms. A relative stress scale [19], brief symptom inventory [20], positive affect and negative affect schedule [21], daily stress inventory [22], and the life stress interview, as well as the visual analog mood scales [23], the trier inventory for the evaluation of chronic stress and others may be among these questionnaires. A psychological tool called the STAI is mainly employed to measure perceived stress [24].

Although psychological questionnaires that measure perceived stress can assess stress subjectively, there are some situations when objective measurements of stress are more useful than self-reported measures since self-report may be skewed. In general, using checklists cannot approach the accuracy of a skilled interviewer in examining research questions [25]. Interviews can be used to get more accurate data than self-reports, which are produced independently. Even when responding to unfavorable questions on surveys, individuals describe pleasant occurrences [26]. Because competent interviewers may elicit descriptive information on the intensity of events across time, interview approaches offer various insights into stress assessment. Interviewers with training can make connections between similar problems and situations. Although these techniques are useful, there are cultural and language difficulties.

2.3.2 Physical Stress

Physical measurements are observable from the outside since they are related to one's behavior. Any attribute that can be seen by people with their natural eyes is referred to as a physical quality. But sophisticated tools and sensors are needed to get physical measurements. Measures of stress and anxiety associated with studying mathematics are based on gesture recognition [27]. Another research looked at pupil instances, yawning, and facial expressions during the day and night to gauge how tired the drivers were [28]. The detection of stress utilizing both internal and external stressors has been accomplished using video-recorded facial expressions [29].

The distinction between the valid pupil diameter and the blink is the most recent and may potentially be an indication of stress [30]. While a lower blink frequency is thought to be the cause of stress. Another study shows a higher blink frequency has been linked to stressful situations [31]. Recently, an objective measure of depression and stress was developed in aphasia patients by using speech features [32]. The two main speech features for detecting stress are rapid fluctuation and an expansion of the frequency range [33]. Table 2.1 presents the physical techniques used to measure stress.

S. No.	Physical Measures	Technology
1.	Eye Activity [30]	Infrared Eye Tracking
2.	Facial Expression [28]	Facial Action Coding System
3.	Body Gestures [27]	Automated Gestures Analysis
4.	Lip, Mouth, and Head tracking [29]	camera-based photoplethysmography
5.	Communication Disorder [32]	Speech Analysis

Table 2.1: Physical Techniques Used to Measure Stress

2.3.3 Physiological Stress

Physiological measurements are related to a person's typical bodily functions. Since they are automatic, it is challenging to measure them. These entail the affixing of certain instruments to a person to track alterations in the body's physiological characteristics. The autonomic nervous system (ANS) oversees controlling the body's involuntary movements. This ANS is made up of the sympathetic (SNS) and parasympathetic (PNS) neural systems. The activity rates in SNS and PNS are inversely proportional. So, an increase in SNS during stressful situations decreases the activity rate in PNS, which represents the resting state of the body. Together, SNS and PNS control blood pressure, brain waves, galvanic skin response (GSR), and heart rate variability. These are the important variables to consider while assessing stress [34].

To accurately assess stress, galvanic skin response (GSR) can be employed [35]. It is an illustration of how well electricity flows through the skin. Electrodes are positioned on the hand's first and middle fingers to measure it. The body's skin resistance diminishes when a person is under stress. This situation occurs because of the skin's increased moisture [23]. Julius et al. provide an explanation of the physiological alterations that stress causes in the skin [36]. GSR might be seen as a crucial metric for workload evaluation [37]. Mean and sum are GSR characteristics that have been used to calculate stress in a variety of scenarios [38].

It is common practice to evaluate ANS activity using heart rate variability (HRV) [39]. It is considered one of the key measures of stress [40]. An ideal way to test HRV is with an electrocardiogram (ECG), which is extremely sensitive to heartbeat detection [41]. The graph of ECG is used to depict the electrical activity of the cardiac muscles that are being recorded. It is determined by attaching electrodes to a person's body on each side of the heart. Stress is often assessed using HRV's low- and high-frequency bands. Shorter-term HRV is a good indicator of acute stress [42].

Researchers have found a link between rising stress levels and rising blood pressure. Recent research [43] found that acute stress tasks had an impact on blood pressure. Photoplethysmography (PPG), which is obtained from a finger, can be used to calculate an individual's Blood Volume Pulse (BVP). By projecting infrared light, a PPG measures the light that is reflected off the finger. Depending on how much blood is within, this light is reflected. PPG sensors are used in industrial stress monitoring systems [44]. A measurement of muscular action potentials is electromyography (EMG). To detect stress, EMG electrodes are positioned on the shoulder. Another stress indicator for people used to monitor by having a belt around their chest is respiration. However, compared to HRV and GSR, respiration monitoring is invasive and does not function well as a stress indicator [45].

Stress and brain activity are strongly correlated [46]. EEG is often employed in published studies on the human brain in the modes listed above [47]. It is the most widely utilized method for studying brain disorders and functioning economically. It is a technique that measures brain activity objectively and precisely. From electrodes positioned on the human scalp, it obtains various waves [48]. High temporal resolution and superior mobility are features of EEG sensors. Since the brain is where the stress reaction originates, EEG becomes an important signal for the analysis, recognition, and accurate measurement of human stress [49].

S. No.	Technology Name	Physiological Measures
1.	Galvanic Skin Response (GSR)[35]	Skin Response
2.	Electrocardiography (ECG) [42]	Electric Activity of Heart
3.	Photoplethysmography (PPG)[44]	Blood Volume Pressure
4.	Electrodermal Activity (EDA)[50]	Skin Response
5.	Electroencephalography (EEG) [47]	Brain Electrical Activity

Table 2.2: List of Physiological Techniques Used to Measure Stress

Numerous techniques have been created to measure and evaluate the amount of stress, whether they are based on questionnaires, surveys, or the observation of persons who measure changes in physiological signals as depicted in Table 2.2.

These two techniques are well recognized for their excellent accuracy and resolution contributions in the spatial and temporal axes. Invasive treatments have drawbacks in some cases, such as hormone analysis, which highlights the need for non-invasive, efficient, accurate, and trustworthy methodologies. Table 2.3 shows different non-invasive procedures having the ability to identify stress from human physiological signals.

S.No	Non-invasive Methods	Advantages	Limitations
1.	Functional Magnetic Resonance Imaging (fMRI)	High spatial resolution	The blood supply responds relatively slowly over time
2.	Electrocardiography (ECG)	Widely accessible equipment	sensitivity is poorer in comparison to other stress imaging techniques, poor Specificity
3.	Magnetic Encephalography (MEG)	High Resolution	Very Expensive
4.	Blood Pressure (BP)	Ease to use, reproducibility of values, the sensitivity of the measurement, and availability of normotensive data	The lack of prospective mortality data so cannot be used to decide whether treatment is indicated
5.	Electromyography (EMG)	A most suitable, reliable, viable, and accurate instrument for measuring muscle activities	Reduced clinical yield in a few cases, other technical limitations emerge with obesity and advanced age, Complexity issues in signal interpretation
6.	Galvanic Skin Response (GSR)	Immediate availability and low cost	Diurnal Fluctuations cause the time of assessment to influence results
7.	Blood Volume Pressure (BVP)	Unobtrusiveness and low costs	Need of individual calibration and drift in it over short intervals
8.	Electroencephalography (EEG)	Outstanding temporal resolution, low cost, and no real safety limitations	Low spatial resolution and elevated noise

Table 2.3: List of Non-Invasive physiological Techniques Used to identify Stress

2.4 Origin of Comedy

The production and perception of comedy clips is a biological process which is a cognitive phenotypic trait, and it is almost certainly reliant on a genetically based brain substrate Comedy has undoubtedly existed for thousands of years, if not millions of years. In their first contact with Australian aborigines, pioneering anthropologists overheard a hilarious exchange. Second, for at least 35,000 years, Australian aboriginals appear to have been genetically isolated [51]. If genetic factors determine the fundamental ability to perceive or make comedy (and convergent evolution does not apply), 35,000 years may be the minimum age for comedy.

2.4.1 Stand-up Comedians

The ability to recognize humour appears to be "instinctive", and so reliant on genetic mechanisms. Comedy is difficult to learn without the aid of a network of specialized brain pathways and a corresponding cognitive module. It is generally a combination of precise language wording and a thorough understanding of current social dynamics that determines whether something is hilarious or not. The majority of the literature focuses on internal techniques for detecting comedians' appraisals via physiological markers. Given the sophistication of comedy as a cognitive process, it is not surprising that the majority of comedy research using physiological measurements has employed brain-imaging methods. The evaluation of stand-up comedians, as well as comic production, has been widely examined utilizing fMRI tests [52]. Another approach that has been effectively employed in stand-up comic research is magnetoencephalography (MEG), It has been utilized to identify brain activations that are particular to comedy and research their temporal order [53]. One advantage of MEG over fMRI is its superior temporal resolution. While these discoveries shed light on the brain activity associated with comedian appraisal, both fMRI and MEG need room-size sensors, making them unsuitable for real-world use.

Although on Electrodermal Activity (EDA), there has been limited research on how to use one of the most frequently used signals in psychophysiology, to identify comedians' evaluations. Finkelstein et al. [54], classified emotions using EDA in conjunction with several other physiological markers, one of which was amusement. The ECG depicts the electrical activity of the heart. The heart rate and heart rate variability (HRV), which are connected to both parasympathetic and sympathetic nervous system activity and have been related to mental workload [50], are the most common parameters extracted from ECG. Fiacconi and Owen [55] looked into the temporal profile of humor evaluation, finding that it was associated with a slowing in heart rate followed by an acceleration in heart rate.

2.5 Brain-Computer Interface

An external computing device and the brain can communicate directly through a brain-computer interface. In 1924, Hans Berger made the initial finding of a scalp-based recording of the EEG signal [56]. The notion of BCI was first established in the decade of 1970, but it took nearly three decades for the BCI research area to really take off. In the literature, Brain-Computer Interface (BCI) is defined in a variety of ways. Traditionally, it is thought of as a kind of communication in which the information or orders that a person conveys to the external realm do not pass through the brain's typical output channels of muscles and peripheral nerves [57]. However, with innovation and development in this area of research, a new definition is established that considers numerous features of the diverse applications that BCI research has unlocked. Figure 2.1 represents an underestimate of the true numbers of publications on BCI technologies showing an increasing trend of BCI publications per year. BCI is a technology that tracks Central Nervous System (CNS) computing and modifies it into artificial output that replaces, improves, supplements, or explores normal CNS output, altering ongoing interactions between the CNS and its internal or external environment [58]. Different application situations can utilize BCIs applications as depicted in Figure 2.2. These can recover, exchange, improve, add to, or augment a natural CNS output that may have been diminished as a result of an illness or injury. These can be utilized in both clinical and non-clinical contexts as research tools.

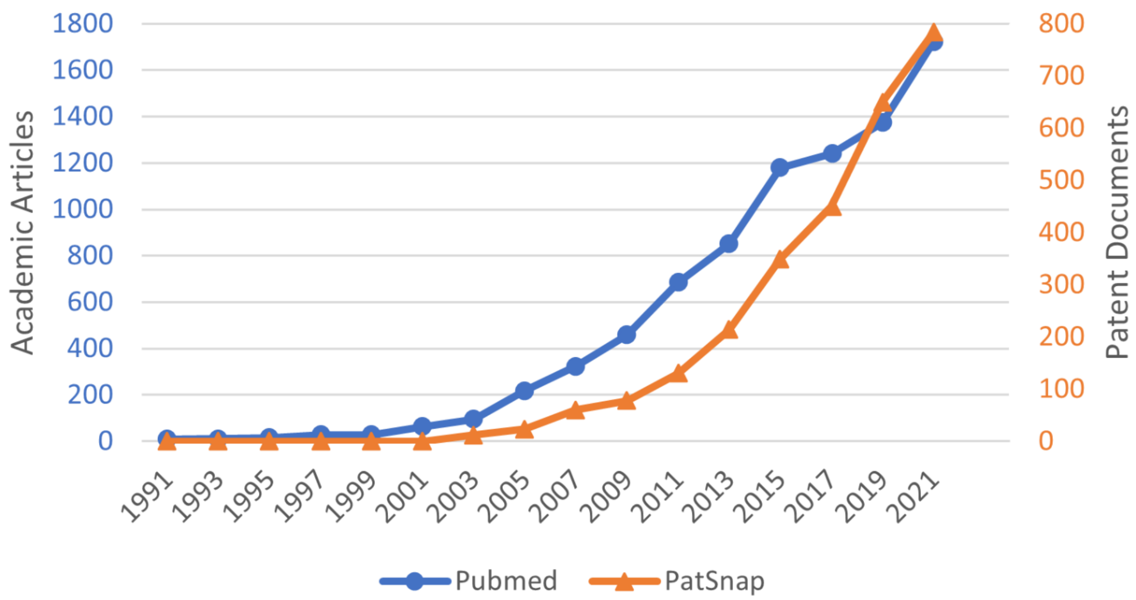

Figure 2.1: An increasing trend of BCI publications per year

Designing BCI applications for both healthy and impaired people is now viable because of the accessibility of inexpensive commercially available EEG headsets. Data collection, preprocessing, feature extraction and selection, and classification are all steps in the processing of EEG signals. The basis of a brain-computer interface system is comprised of these actions. The types of brain-computer interfaces are described in the section that follows.

2.5.1 Types of Brain-Computer Interface

BCIs use a variety of data-collecting techniques to capture a person's brain activity. These techniques can be roughly divided into invasive and non-invasive categories. Invasive techniques are operated using specialist equipment.

EEG is the modality that non-invasive BCI devices employ the most. These BCIs fall into one of three categories: spontaneous, evoked, or passive BCIs. Evoked BCIs operate on EEG signals that are produced in reaction to certain environmental stimuli.

Theoretically, electrodes placed on the scalp might simply record evoked potentials. P300 numbered-state visual evoked potential (SSVEP) has been studied the most fre-

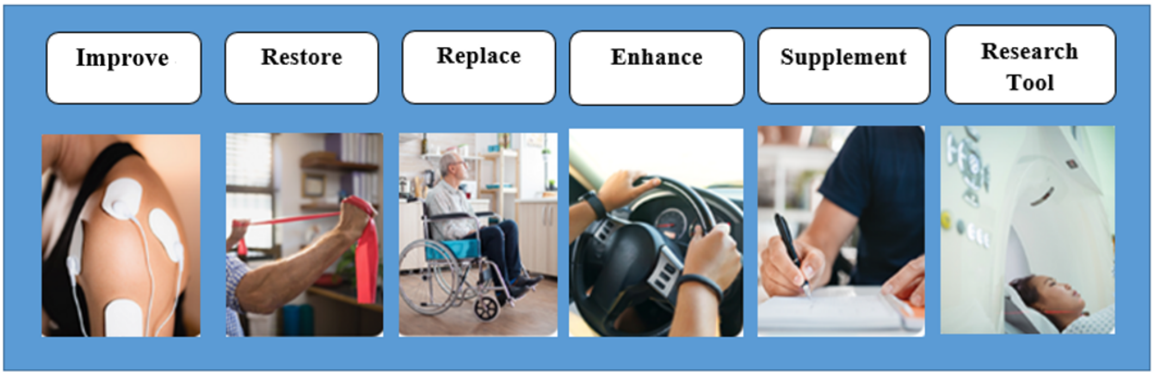

Figure 2.2: BCI Application scenarios

quently in the field of BCI. The spectrum of applications is constrained by the application of external stimuli. Brain activity may be employed for control-based applications like robotic limbs. There are two forms of brain activity: spontaneous and deliberate. These BCIs rely on naturally occurring cognitive tasks and EEG-based analysis. Both gradual potential shifts and fluctuations in rhythmic activity are kinds of brain oscillations used by these BCIs [59]. Slow potential changes can be seen in eye movement, whereas variations can be produced by neck muscle tension. Some of the basic applications of BCIs are illustrated in Figure 2.2.

- BCIs can **replace** natural Central Nervous System (CNS) output that has been lost due to illness or injury. Examples include controlling a motorized wheelchair and communicating (using a voice-synthesized spelling system).
- BCIs can **restore** lost natural CNS output. Examples include functional electrical stimulation of muscles in a paralyzed person.
- BCIs can **enhance** the CNS's natural output. Examples include keeping track of brain activity when performing prolonged activities like driving a car and spotting concentration gaps to notify and educate the person.
- BCIs can **supplement** the CNS's natural output. Examples include giving a person a third (robotic) arm and giving persons using joysticks access to a selection function.

- BCIs can **improve** the CNS's natural output. Examples include using an orthosis to increase arm movements or a BCI in stroke rehabilitation to detect and boost signals from a damaged cortical area.
- In both clinical and non-clinical research projects, BCIs can be utilized as a **research tool** to examine CNS processes.

A surgical procedure is essential for an invasive BCI system to capture a single neuron's activity. To measure a neuron's frequency, small electrodes are inserted into the brain. Such experiments are carried out on animals in a variety of ways to identify the functions the brain performs. In [55], the invasive method was used on a human subject for the first time. When one of the volunteers was able to direct a computer cursor and type messages, the experiment was deemed successful. However, the effectiveness was not significantly greater than the non-invasive method [60]. Signals from the electrodes inserted under the skull are recorded using the electrocorticogram (ECoG), a less invasive method that is less noisy than EEG signals due to the position of their electrodes, although they still need surgery [61].

2.5.2 Scalp EEG and Signal Processing

Since EEG data directly reflects the electrical activity of the brain, stress levels may be classified using this information [62]. Most often, a dense placement of EEG electrodes is necessary for an accurate assessment of stress using EEG [63]. A wearable EEG that is discrete, portable, and may be utilized for extended periods of time is needed to monitor stress in the activities of everyday life. These criteria are not met by clinically dense EEG systems with 32, 64, or 128 channels. However, it has been demonstrated that a single-channel EEG headset can also be used to detect stress satisfactorily [64].

Numerous resources are offered both for free and for a fee to aid with the advancement of BCI. Every platform offers benefits of its own based on what is needed for an experiment. For BCI development, a wide variety of tools are accessible both for free and for payment. Every platform offers benefits of its own based on what is needed for an experiment. Online processing on DataSuite, FieldTrip, BCI2000, or OpenViBE systems uses the BCILAB toolkit [65]. To do this, it might be thought of as an EEGLAB plug-in extension. A toolset called BioSig is built on MATLAB/Octave [66]. Table 2.4 gives a list of BCI headsets that are marketed and easily accessible.

S.No.	Headset Companies	Sampling Rate (samples/sec)	No. of Channel	Data Transfer Mode	Operating Time (hours)
1.	EMOTIV (Epoc+)	128 or 256 2048 (internal)	14	Wireless	12
2.	EMOTIV (Insight)	128	5	Wireless	4
3.	COGNIONICS	1000	20	Bluetooth	8
4.	Freedom 24D	300	20	Bluetooth	
5.	quasar DSI	240 or 960	12	Wireless	24

Table 2.4: Specifications of commercially available BCI headsets

Commercial BCI headset systems based on EEG are on the market. Many businesses have produced their own headset systems, including PLX Devices, Neurosky, MyndPlay, IMEC, g.tec, Emotiv, OCZ Technology, and Congnionics. These businesses provide a variety of tools and apps in addition to these headset systems.

2.6 Research Gap

The use of stand-up comedy clips to classify stressors in a diverse population is not just a novel approach to EEG research, it is an essential one. We intend to use this fact to create a sufficiently large open-source raw EEG dataset, as standup comedy clips are viewed by people of all ages and genders worldwide. The primary objective of this study is to offer a multimodal dataset that modulates both effective stress and emotions. The simultaneous regulation of effective stress and its effects on mental state and artifact generation have not yet been studied in publicly available datasets. The second objective of this research is to build a dataset that replicates real-time brain signal recording with the help of Neurosky Mindwave 2 headset and a realistic task environment (by watching stand-up comedy clips), with the help of BCIs, we can accomplish this task.

Chapter 3

Methodology and Implementation

Different classification algorithms have been used to classify EEG recordings into stress and non-stress categories based on the meaningful information learned from the extracted features. We provide a brief overview of the suggested framework for the usage of ML and signal processing in the detection and categorization of stress using standup comedy clips as depicted in Figure 3.1. The proposed methodology consists of the following main modules:

- EEG Data Acquisition
- Signal Pre-processing
- Feature Extraction
- Feature Reduction and Selection
- Data Distribution
- Classification

3.1 EEG Data Acquisition

Only a few datasets are available for stress classification because collecting EEG data takes a lot of effort. Every publically available dataset has a particular stimulus. We

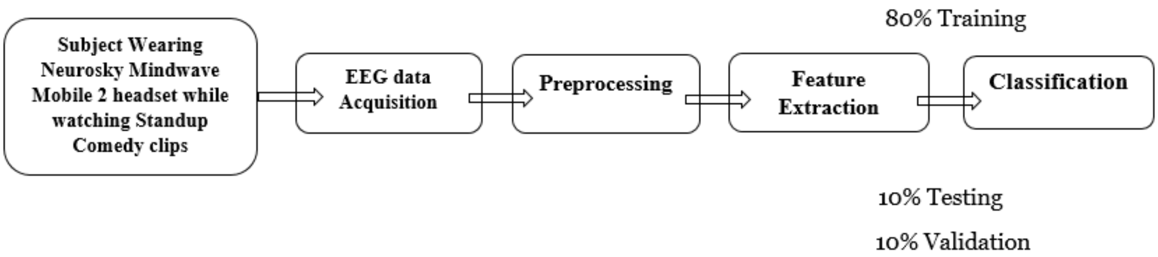

Figure 3.1: Stress Classification Network Using EEG

built a dataset from scratch for our research because there were no such datasets containing comedy clips in native and non-native language as stimuli.

3.1.1 Participants Details

In this study, data has been collected from 30 healthy volunteers with a mean age of 25.20 years and a gender split of 50% male and 50% female. All participants were students in a public sector higher education institute. They were all healthy individuals without any known mental health conditions.

3.1.2 Videos Selection

Stand-up comedy videos in two different languages, namely English and Urdu, were employed as an external stimulus for the stress classification experiment based on EEG signals. The video clip with similar content in each language was chosen depending on the YouTube rating in their countries. A clip of three minutes from each video track was utilized as a stimulant to avoid boredom during the experiment.

- For English experiment, we used a popular comedy clip of George Denis Patrick Carlin naming "Airline Announcement" [67], who was an American comedian, and social critic. He is regarded as one of the most important and influential stand-up comedians of all time in American history.
- For Urdu Experiment, We used a comedy clip of Umar Sharif regarding "Airline

Announcement" [68]. He is known professionally as one of the most acknowledged Pakistani actors and Comedians.

3.1.3 Neurosky MindWave Mobile 2

With reference to an electrode put at the earlobe and a dry electrode placed at the frontal location (FP1) of the brain, the Neurosky Mindwave device offers a single channel for EEG recording as shown in Figure 3.2. The system has made use of dry electrode technology from the Thinkgear application-specific integrated circuit module (TGAM). It has a 3-100 Hz bandwidth and operates at a minimum voltage of 2.7V. The silver TGAM electrodes are appropriate for non-hairy regions. The wearable headset is geared up to record discrete EEG data at a sampling rate of upto 512 Hz. It has a simple algorithm and the ability to capture raw EEG.

Figure 3.2: Neurosky Mindwave Mobile 2 headset

3.1.4 Experimental Procedure

Each participant was given a thorough explanation of the experiment before watching the videos. Demographic information about the participants was gathered using a survey questionnaire.

Participants first completed a consent form, Personality Test [69], and trait anxiety (Y2) form, and then baseline EEG signals were recorded for three minutes.After the baseline recording, the subject filled out the state anxiety form (Y1(I)) and then watched the stand-up comedy clip in English. The video's loudness was changed based on the participant's preferences. While watching the clips, the participant's EEG signals were captured. Following the watching of the comedy clip, the participant once again completed the state anxiety form (Y1(II)), which served as a label to differentiate the data for the stress classification task. By presenting the comedic clip in Urdu for each subject, the same technique was repeated. The experimental approach used for data collection in this investigation is depicted in Figure 3.3.

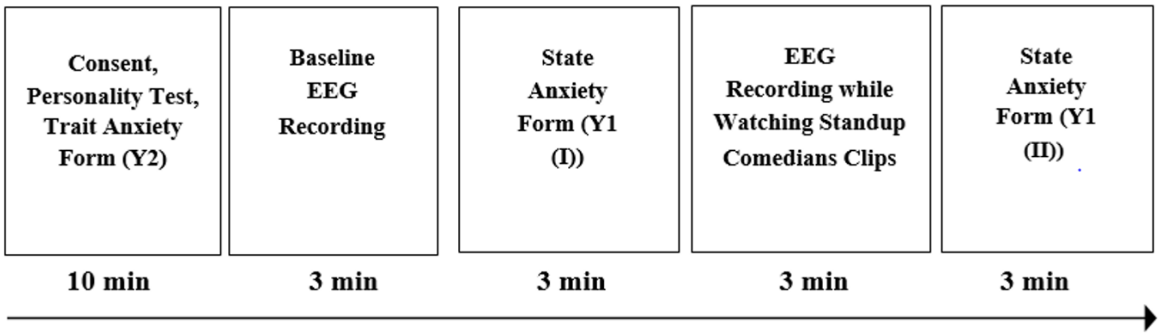

Figure 3.3: Experimental procedure

3.2 Signal Pre-Processing

The signal processing techniques used in our work are similarto the work of [68].

3.2.1 Noise Filtering

The Scipy package is used as a filter to remove muscular and ocular artifacts. These are low pass, median pass, and high pass filters as well as a notch filter with third-order filters. For smoothing and denoising, Savitzky-Golay and wavelet denoising filters are used. Firstly, to remove the background noise, third-order median filters were employed to convert the raw data into filtered data. These data were further strained using high and low pass filters. The cut-off frequencies of 0.5 and 50Hz with the fifth-order Butterworth filter were used. This is because Butterworth is used for a flat bandpass frequency response. Figure 3.4 depicts the first participant's three-second EEG signals.

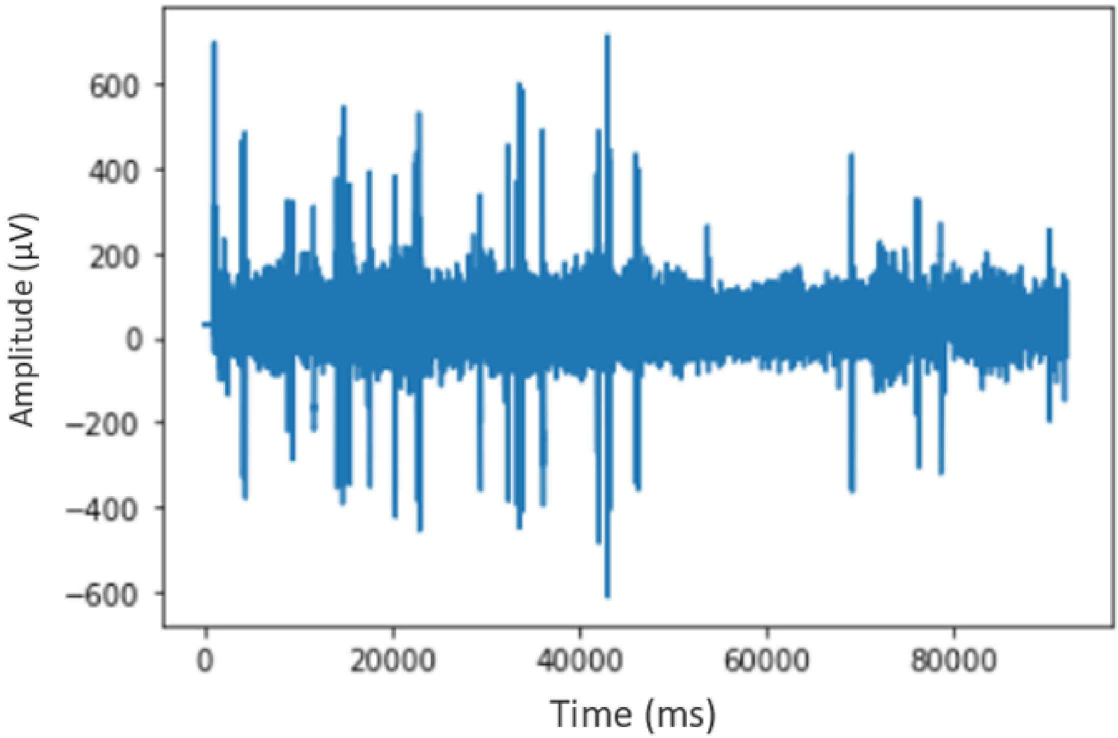

Figure 3.4: Raw EEG data of subject 01

To remove the power interference at 60Hz, a notch filter was used. Furthermore, multilevel wavelet decomposition is used to denoise these signals. In the end, the Savitzky-Golay filter was used to smooth the signals.

3.2.2 Windowing

The filtered signals are segmented using a sliding window once the data has been processed. A few things needed to be taken into account for this. For time-domain signals, a 4-sec window width was used to divide the data into small portions for the duration of 4 sec each. The width is adopted based on the literature.

Two methods were applied, to choose whether the sliding window has to be overlapping or non-overlapping, It was concluded that the overlapping technique was more effective and provided greater correctness. The maximum accuracy was attained with a 4-sliding window and 3-second overlapping after trying various values initially with 50%

overlapping.

These segments are then transferred into the frequency domain using the Fast Fourier transform. According to [70], these signals were obtained in the time-frequency domain, because the signal from that domain obtains the highest accuracy from the extracted feature. The wavelet transformation was employed for transforming the signals to the time-frequency domain.

3.3 Features Extraction

From the time, time-frequency domain and Wavelet domain, the following features are extracted:

3.3.1 Time Domain Features

1. **Mean**: The signal mean is the sum of values at each second divided by the total number of time points.

$$\mu = \frac{1}{N} \sum X_{(i)} \tag{3.1}$$

2. **Standard Deviation**: A low standard deviation (std) signifies that the amount of data is relatively near to the mean, whereas a high standard deviation suggests that the values are widely spread.

$$\sigma = \sqrt{\sum (X_{(i)} - \mu)/N} \tag{3.2}$$

3. **Median Absolute Deviation**: Median Absolute Deviation (mad) is a robust measure of the variability of a univariate sample of quantitative data.

$$MAD = |X_{(i)} - \mu| \tag{3.3}$$

3.3.2 Time-Frequency Domain Features

1. **Skewness**: A measure of an EEG signal dataset's lack of symmetry or asymmetries is called skewness.

$$Skewness = \frac{E[(X_{(n)} - \mu)^3]}{(\mu)^3} \tag{3.4}$$

2. **Kurtosis**: Kurtosis (kur) is the ratio of the fourth moment and the variance of the signal.

$$Kurtosis = \frac{m4}{m2^2} \tag{3.5}$$

3. **Interquartile Range**: The interquartile(iqr) range in descriptive statistics also called midspread, H-spread, or middle 50 percent, is a statistical dispersion measure equal to the differences between the 25th and 75th percentiles, or between lower and upper quartiles.

3.3.3 Wavelet Domain

A wavelet is an essential tool for preserving both temporal and spectral information. For the wavelet domain, entropy and energy have been estimated. Because the wavelet decomposes the signal into the required frequencies, we can also calculate energy bands from it. In this instance, the decomposition was carried out between 512 and 1 Hz. Both the detailed energy and the approximative energy were calculated, and their probabilities are expressed as follows:

1. **Entropy**:

$$H(W(a,\tau)) = -\sum_{i=1}^{N} p(W(a,\tau)) \log_2 P(W(a,\tau)) \tag{3.6}$$

where $P(W(a,\tau))$ is wavelet coefficients' probability.

2. **Energy**:

$$E_w = \sum_{i=1}^{C} W(a,\tau)^2 \tag{3.7}$$

where $(W(a,\tau))$ are the wavelet coefficients

3.4 Feature Selection

The selection of the features is an essential step once the features have been extracted since there may be instances in which certain features are linked with one another and impact how effectively the machine learning algorithm operates. High co-linearity among the features also makes ML classifiers more prone to over-fitting.

For feature selection, the variable importance function for the LGBM Classifier yielded the results given in Figure 3.5.

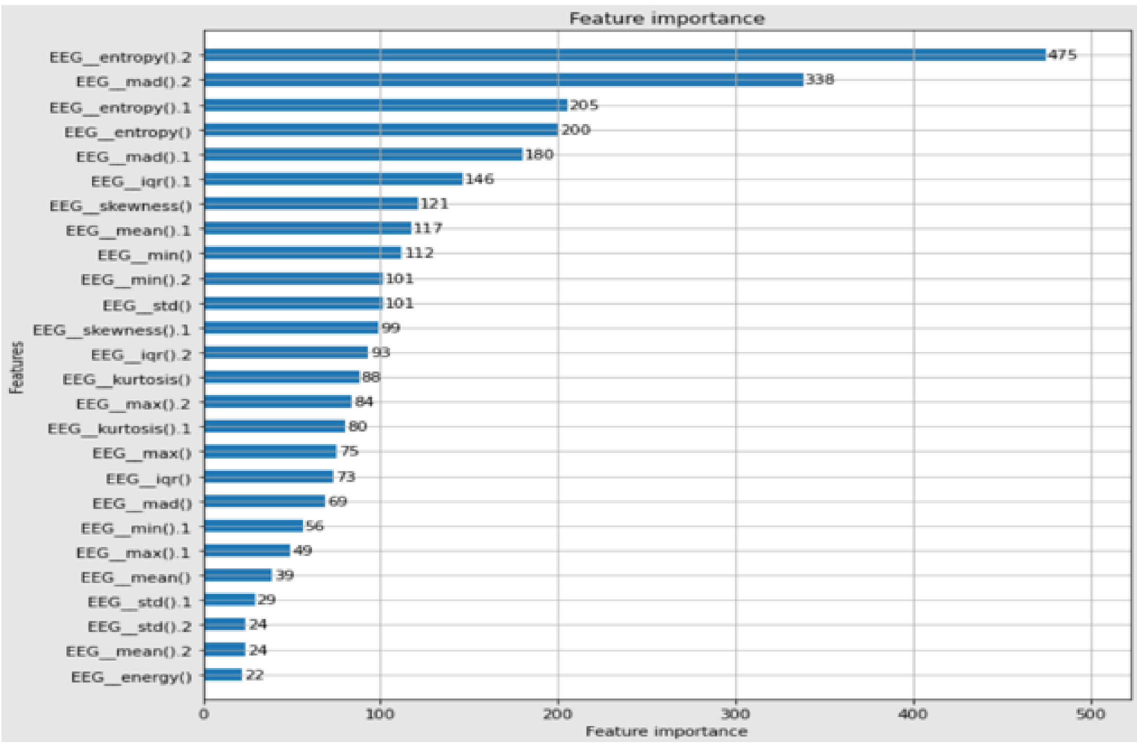

Figure 3.5: Feature Importance

It can be observed that entropy is the most noteworthy feature, followed by signal mad(), Iqr(), Skewness(), and so on. Further, we observe that the entropy feature related to the wavelet domain shows high importance, and the energy feature reduces but still maintains certain importance.

Furthermore, we observe that almost all six features mad(), iqr (), mean (), min(), max(), and std() related to the time domain show high importance. We can also observe that the features related to the time-frequency domain, such as skw() and kur(), also maintain high importance.

As shown in the figure, almost all ten features related to time, time-frequency, and wavelet domains play highly important roles in the prediction of stress classification. So we used all ten features for stress classification.

3.5 Labelling based on STAI Score

Each participant's responses are used to calculate their STAI scores. Participants' levels of mental stress are positively correlated with their STAI form score. The determined threshold T used to determine which individuals would be placed in the stress group was $T = \mu$, where μ is the mean of the STAI scores [71]. All candidates with STAI values less than or equal to T are thought to be under relatively low levels of stress, whereas those with scores equal to or higher than T are seen to be under very high levels of stress. The participants are divided into the stress and non stressed groups after the threshold. For training purposes, these labels are given to the classification methods. According to the examination of the reactions, the mean STAI of all responses for the English comedy clip was 35.97 and 30.93 for the Urdu comedy clip. For the purposes of this study, those participants with STAI values below thresholds are considered generally non-stressed, and those with scores above T are considered to be stressed.

3.6 Dataset Distribution

The dataset is split into three sections, such as Training Data, Validation Data, and Testing Data. Training, Validation, and Testing are split into (80%, 10%, 10%) percent respectively.

3.6.1 Training Data

Different machine learning models utilize 80% of the training data, although this ratio may alter according to the requirements of the project. The machine learning models, which attempt to learn from the training sample, are trained using this data.

3.6.2 Validation Data

The validation data is 10% of the original dataset and it is used to validate different machine learning algorithms' performance during training. The information obtained from this validation approach can be used to modify the model's hyperparameters and configurations. It works similarly to a critic informing us of the direction our training is heading towards. To avoid overfitting, we split the dataset into a validation dataset. Additionally, ranking the model's accuracy helps with model selection.

3.6.3 Testing Data

The ML algorithms are tested on new data using a test set that represents 10% percent of the original data. Once the model has received the necessary training, it is used for the evaluation process. It offers a concluding model performance evaluation metric in terms of accuracy precision, recall, and F1-Score. Simply, it offers a response to the question "How effectively does the model work with unseen data?".

3.7 Classification Techniques

In order to evaluate human stress, a variety of classifiers have been used. In the literature, supervised learning approaches dominate. The following subsections provide details on the five algorithms that were used to categorize stress in the present thesis.

3.7.1 Long Short Term Memory (LSTM)

Schmidhuber and Hochreiter in 1997 introduced the LSTM network [72]. The specialty of this network that differentiates it from the recurrent neural network is that they learn long-term dependencies. This makes them one of the best algorithms to work with the specialty of the memory element and enables it to remember previous sequences of steps. It is faced with Recurrent Neural Network (RNN) with a few changes in the structure, to overcome or remove the difficulty of removing gradient.

3.7.2 Light Gradient Boosting Machine (LGBM)

A gradient boosting structure that customs decision trees-based learning methods. It is effective with faster training effectiveness, distributed, and can cover a huge volume of data, but when dealing with the high dimensional features for EEG signals there exist some deficiencies the case of time consumption as well as low accuracies.

3.7.3 eXtreme Gradient Boosting (XGB)

The term "XGB" refers to a popular supervised-learning method for regression and classification on huge datasets. It uses sequentially generated shallow decision trees and a highly scalable training method to prevent overfitting in order to deliver accurate results. It functions with huge, intricate datasets.

3.7.4 Random Forest Classifier

The random forest classifier creates numerous decision trees from a randomly chosen portion of the training dataset. Then, it selects the final test class of stressed and non-stressed by voting from different decision trees. The random forest classification was generated with a reduced number of trees.

3.7.5 EXTremely RAndomized Trees classifier

Extremely Randomized (EXTRA) Trees Classifier, also called EXTRA Trees Classifier, is an ensemble learning technique that integrates the results of various de-correlated decision trees gathered in a "forest" to obtain its classification output. It differs from a Random Forest Classifier only in terms of how the decision trees in the forest are constructed theoretically.

3.8 Validation

The validation method used for all our ML classifiers is k-fold cross-validation where $k = 10$. The following steps are followed for 10-fold cross-validation.

- Data were randomized and split into ten equal parts.
- Repeat for each group:
  - Set one group as the test group.
  - Fit a model and train the rest of the groups.
  - Retain the score and overwrite the model.
- Generalize the model scores based on the above evaluation.

3.9 One-way analysis of variance (ANOVA)

To assess the variation in the amount of stress based on comedic clips, The statistical method of one-way analysis of variance (ANOVA) is employed to assess potential variations across different data sets. A bigger F-statistic in an ANOVA is correlated with a smaller p-value, which challenges the validity of the null hypothesis and recommends that there are important differences between the classes. In this study, a significance level of 0.05 is taken into account. The analysis is done using the subjects' state anxiety scores both before and after watching English and Urdu comedy clips.

H_0: The stress level is different

H_1: The stress level is the same.

With respect to Y1 (I) and Y1 (II), English comedy has a p-value of 0.013. The scores of participants who watched the comedy clip with respect to Y1 (I) and Y1 (II) in English and Urdu differed significantly, as shown by the Urdu comedy clip's p-value of 0.002.

On the other hand, the p-values for the state anxiety scores Y1 (I) of baseline English and Urdu comedy clips is 0.133, and the p-value for state anxiety score Y1 (II), after comedy clips is 0.020. It is evident from the results that no significant difference is found among the state anxiety scores of baseline English and Urdu, but a significant difference is found among the state anxiety scores of participants in response to comedy clips of English and Urdu.

3.10 Models Performance Evaluation

The performance parameters of a system must be determined in order to measure and compare performance. The accuracy and other pertinent characteristics are used to assess the classifier's performance. In the subsections that follow, we quickly cover some of the key performance metrics.

3.10.1 Confusion Matrix

It is helpful to create the $N_y \times N_y$ "confusion" matrix, whereN_y is the number of classes, to evaluate the classifier in ML. The columns of M represent the classifier's output, and the rows of M represent the actual class labels. The ratio of true positives to false positives varies as a classifier's parameter is changed. For $N_y = 2$, a confusion matrix is given in Table 3.1.

True Class	Positive	Negative
Positive Negative	True Positive False Negative	False Positives True Negative

Table 3.1: General Confusion Matrix

3.11 Precision, Recall, and Accuracy

Let's suppose that TP, TN, FP, and FN stand for true positives, true negatives, and false positives, and false negatives respectively. Eq.3.8 - Eq.3.10, may be used to determine the categorization system's accuracy, precision, and recall.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.8}$$

$$Precision = \frac{TP}{TP + FP} \tag{3.9}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.10}$$

3.11.1 F1-Score

The F1-Score is typically regarded as the harmonic mean of a system's accuracy and recall. Its value is between 0 and 1. A greater score indicates a system's accuracy and recall are higher. If the recall, and accuracy of a classification system are known, the F1-Score is determined using Eq. 3.11.

$$\text{F1-Score } = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3.11}$$

Chapter 4

Results and Discussion

A machine learning framework is proposed in this paper for stress classification using EEG signals in response to different comedy clips. The experimental results are presented using, (i) statistical analysis of human stress behavior in response to different language comedy clips and (ii) performance evaluation of two class stress classifications using EEG signals.

4.1 Statistical Inferences

For the statistical study, based on their trait anxiety score i.e., Y2, participants are divided into stressed and non-stressed groups for statistical analysis. Figure 4.1 and Figure 4.2, display the trait anxiety scores for each participant in this study. If a participant's trait anxiety score is higher than the mean trait anxiety score for all participants, they are considered to be in the stressed group. In this analysis, the average trait anxiety score for English comedy was 44 and for Urdu comedy, it was 41. For English comedy, there are 17 participants in the stressed group including 9 females and 8 males and 13 are non-stressed including 6 females and 7 males. Similarly, for Urdu comedy, the non-stressed group consists of 16 participants including 8 males and 8 females are stressed, and 14 participants are non-stressed comprising 7 males and 7 females.

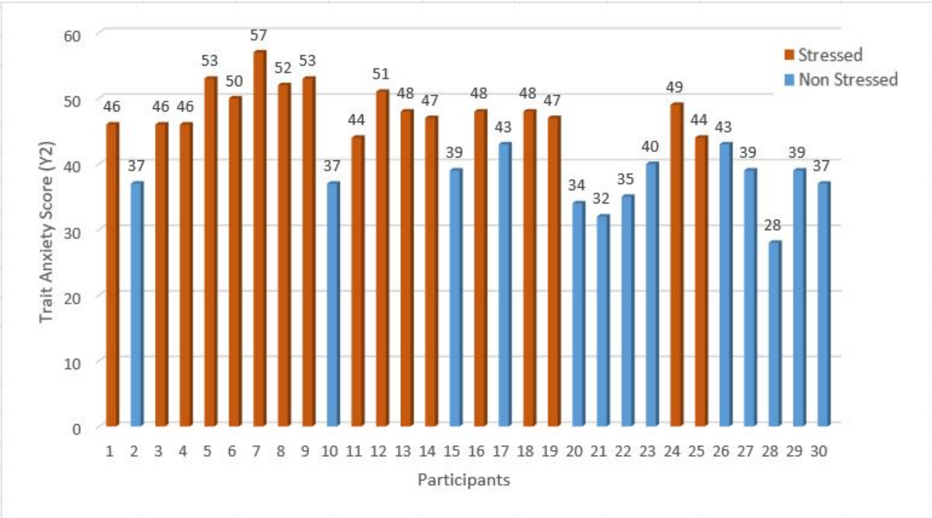

Figure 4.1: Trait Anxiety score for English Comedy

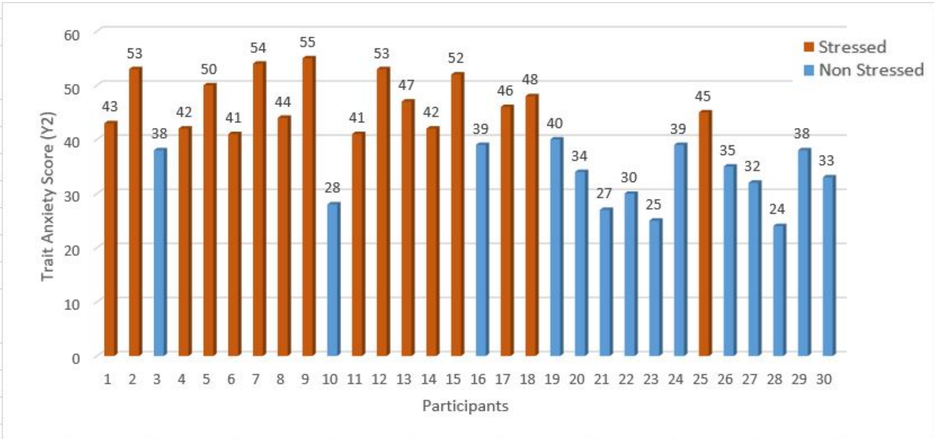

Figure 4.2: Trait Anxiety score for Urdu Comedy

Both stressed and non-stressed groups are evaluated using the ANOVA test based on state anxiety scores. i.e. Y(I). In order to better understand the language-related discrimination in stress classification in response to comedy clips we used ANOVA on Y1(I) and Y1(II) of English and Urdu Comedy Clips. The analysis is done using the subjects' state anxiety scores both before and after watching English and Urdu comedy clips.

- H_0: Stress level is different
- H_1: Stress level is the same

With respect to Y1 (I) and Y1 (II), English comedy has a p-value of 0.013. The scores of participants who watched the comedy clip in English and Urdu differed significantly, as shown by the Urdu comedy clip's p-value of 0.002 with respect to Y1 (I) and Y1 (II). The p-values for the state anxiety scores Y1 (I) of the baseline English and Urdu comedy clips, on the other hand, are 0.133, and the p-value for the state anxiety score Y1 (II), after the comedy clips, is 0.020. The results clearly show that there is no significant difference between the baseline English and Urdu state anxiety scores and that there is a significant difference between the state anxiety scores of participants.

4.2 Classification

Five different classification models were used in the experiments. The results presented herewith aim to measure the accuracy needed for each classifier throughout the data training phase.

Experiments show that EXTRA Tree outperformed all other classifiers in terms of performance with a testing accuracy of 84.29% for English and 78.32% for Urdu comedy clips. Random forest ranked second with an accuracy of 84.01% and 77.84%, followed by XGB with 83.52% and 78.19% for English and Urdu comedy clips respectively. The

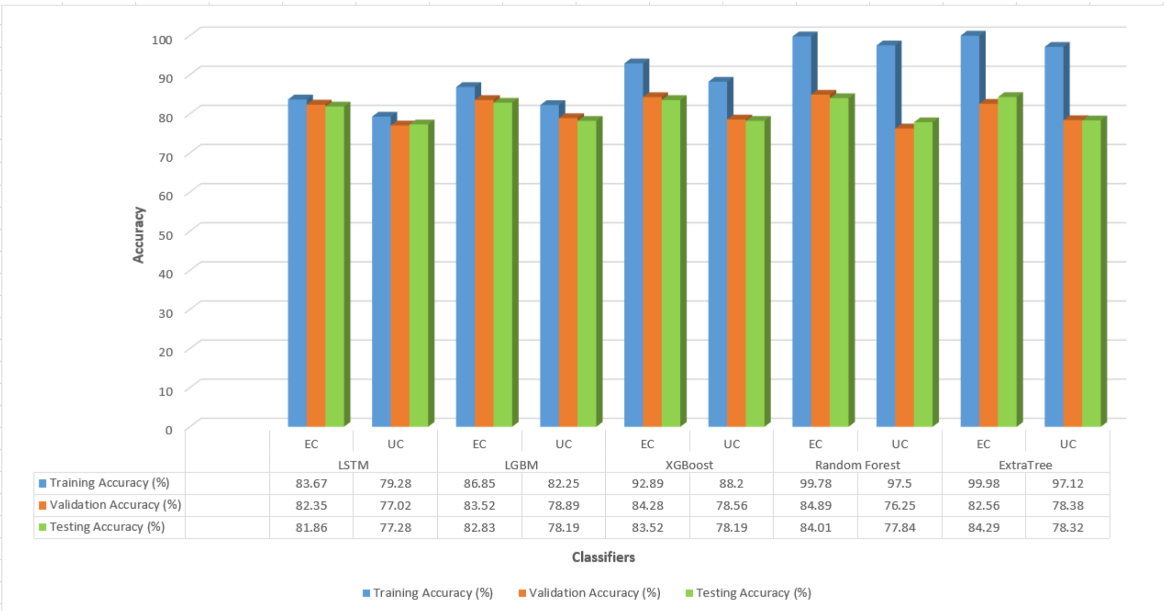

Figure 4.3: Classification Accuracy of Classifiers Used

classifier's accuracy performance is also depicted graphically in Figure 4.3.

In machine learning, the learning curve is used to predict how well models would work with different numbers of training data. A increasing number of training samples are used to monitor the training and validation accuracy. Figures 4.4 – 4.13, show a model with various training and validation accuracy score. The model's training accuracy is indicated in the figure by the red line, its validation accuracy is indicated by the green line, and the dots represent the cross-validation values of the training example, which is five. The accuracy curves for training and validation were achieved during the learning phases for English and Urdu, respectively.

Figure 4.4 demonstrates that for English Comedy data, the EXTRA Tree classifier obtained training and validation accuracy of 99.98% and 82.56%, respectively, during the training phase. The figure shows that the training and validation scores improved as the number of training examples increased. Therefore, the higher the accuracy, the better the performance of the model.

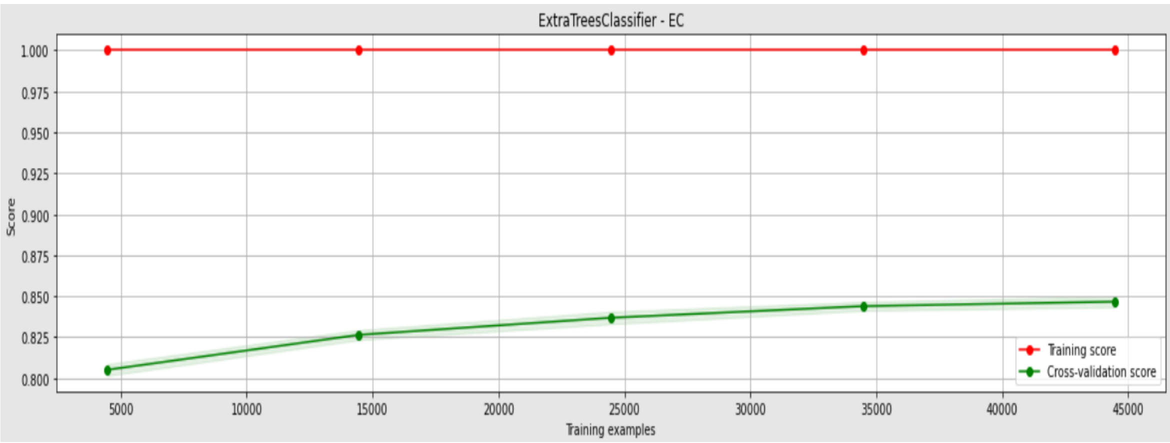

Figure 4.4: Training and Validation Accuracy of EXTRA Tree classifier for English Comedy

Figure 4.5 illustrates that for Urdu Comedy data, the EXTRA Tree classifier obtained training and validation accuracy of 97.12% and 78.38%, respectively, during the training phase.

Figure 4.6 demonstrates that for English Comedy data, the Random Forest classifier obtained training and validation accuracy of 99.78% and 84.89%, respectively, during the training phase

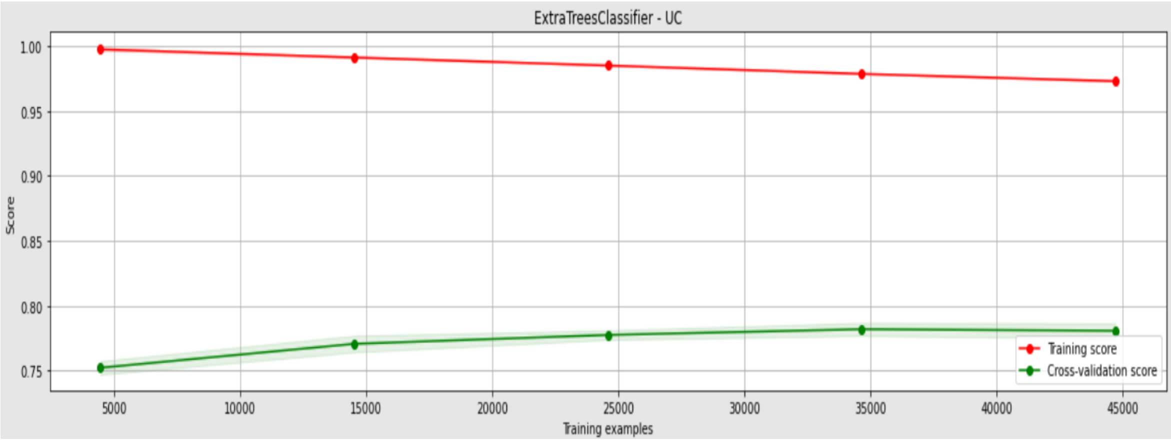

Figure 4.5: Training and Validation Accuracy of the EXTRA Tree classifier for Urdu Comedy

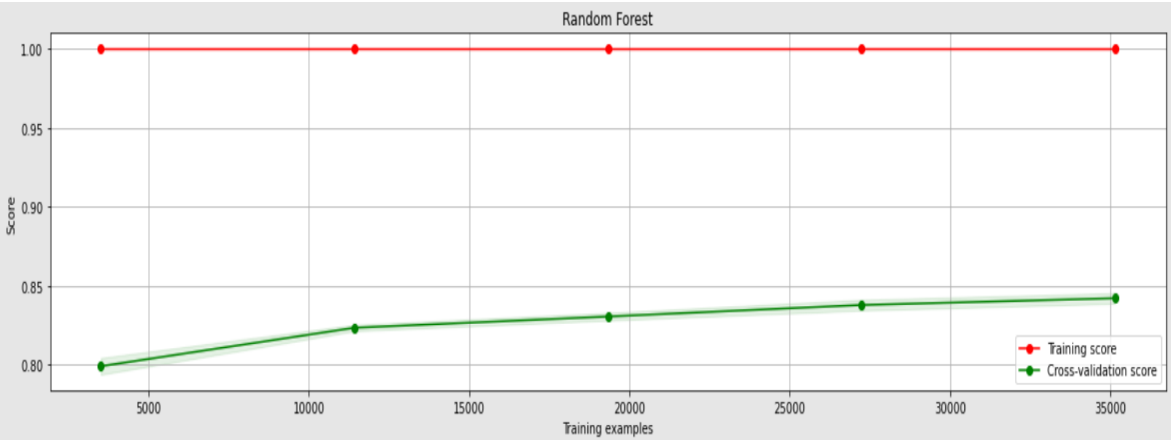

Figure 4.6: Training and Validation Accuracy of the Random forest classifier for English Comedy

Figure 4.7 demonstrates that for Urdu Comedy data, the Random Forest classifier obtained training and validation accuracy of 97.50% and 76.25%, respectively, during the training phase. The figure shows that the training and validation scores improved as the number of training examples increased.

Figure 4.8 demonstrates that for English Comedy data, the XGB classifier obtained training and validation accuracy of 92.89% and 84.28%, respectively, during the training phase.

Figure 4.9 demonstrates that for Urdu Comedy data, the XGB classifier obtained training and validation accuracy of 88.20% and 78.56%, respectively, during the training phase.

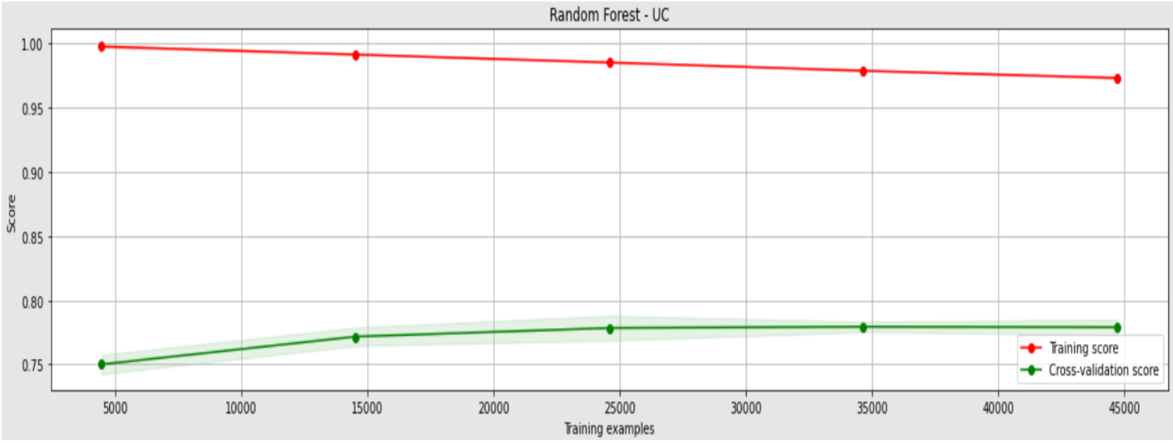

Figure 4.7: Training and Validation Accuracy of the Random forest classifier for Urdu Comedy

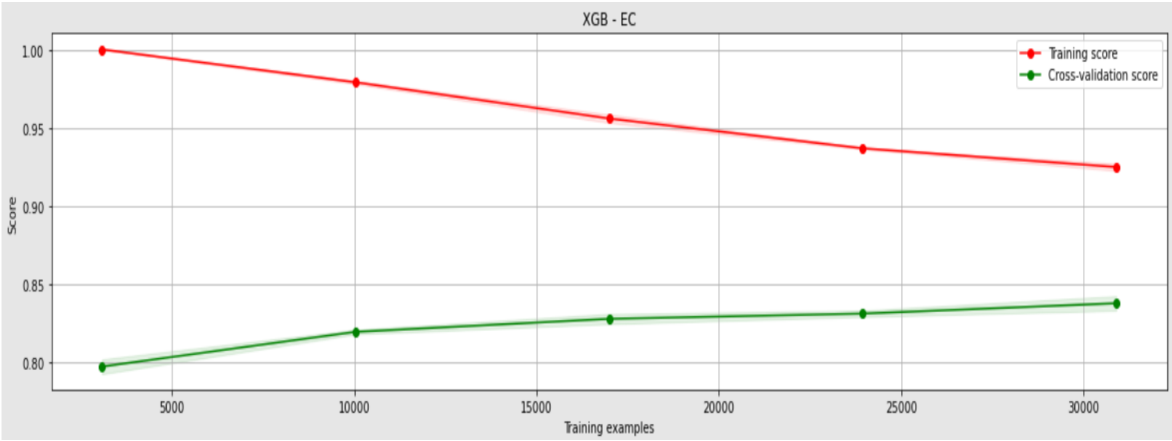

Figure 4.8: Training and Validation Accuracy of the XGB classifier for English Comedy

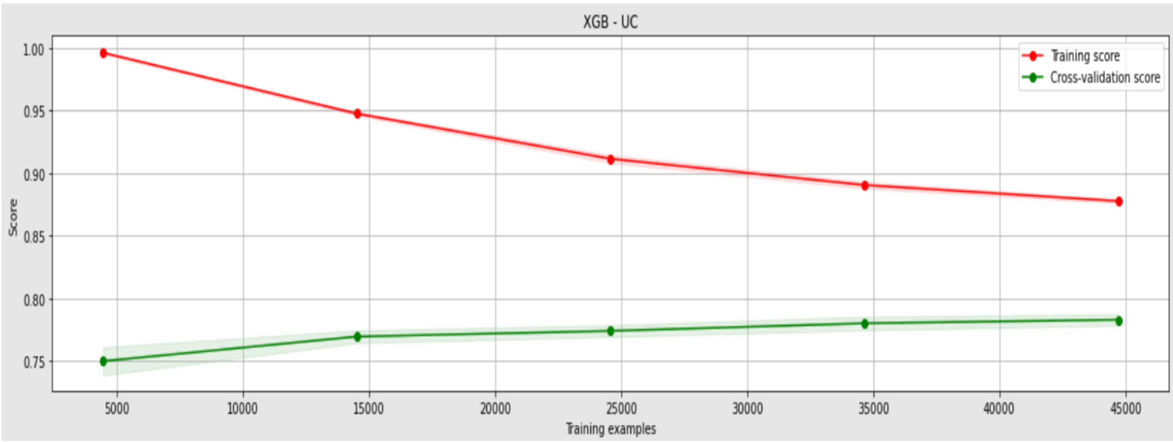

Figure 4.9: Training and Validation Accuracy of the XGB classifier for Urdu Comedy

Figure 4.10 demonstrates that for English Comedy data, the LGBM classifier obtained training and validation accuracy of 86.85% and 83.52%, respectively, during the train-

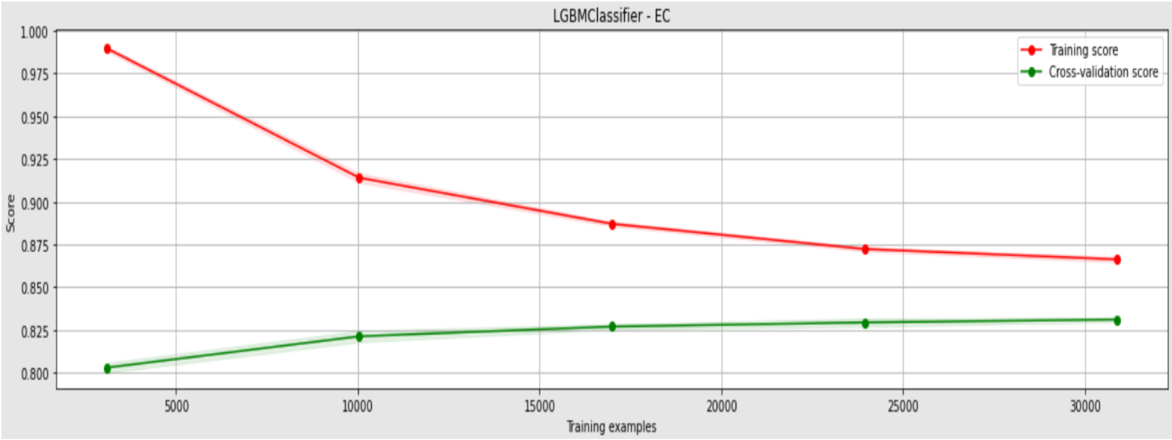

Figure 4.10: Training and Validation Accuracy of the LGBM classifier for English Comedy

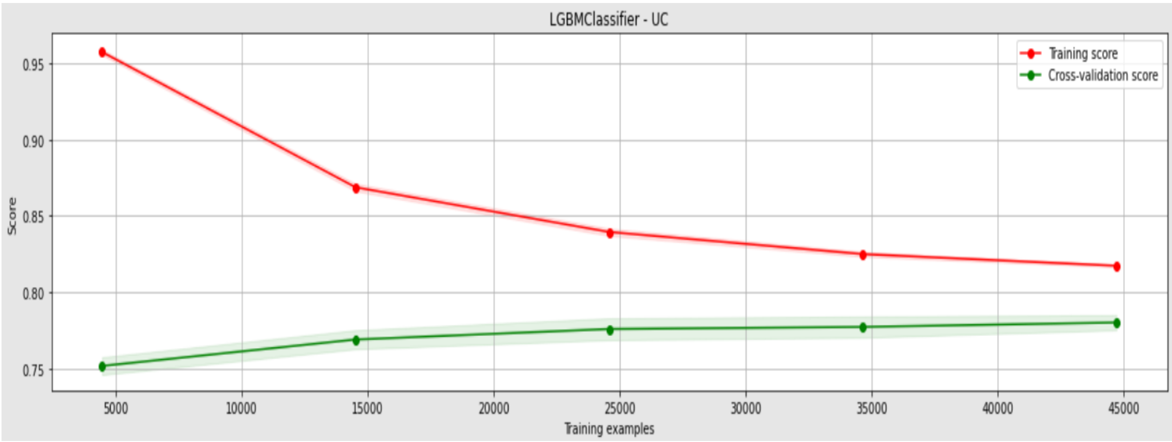

Figure 4.11: Training and Validation Accuracy of the LGBM classifier for Urdu Comedy

ing phase.

Figure 4.11 demonstrates that for Urdu Comedy data, the LGBM classifier obtained training and validation accuracy of 82.25% and 78.89%, respectively, during the training phase.

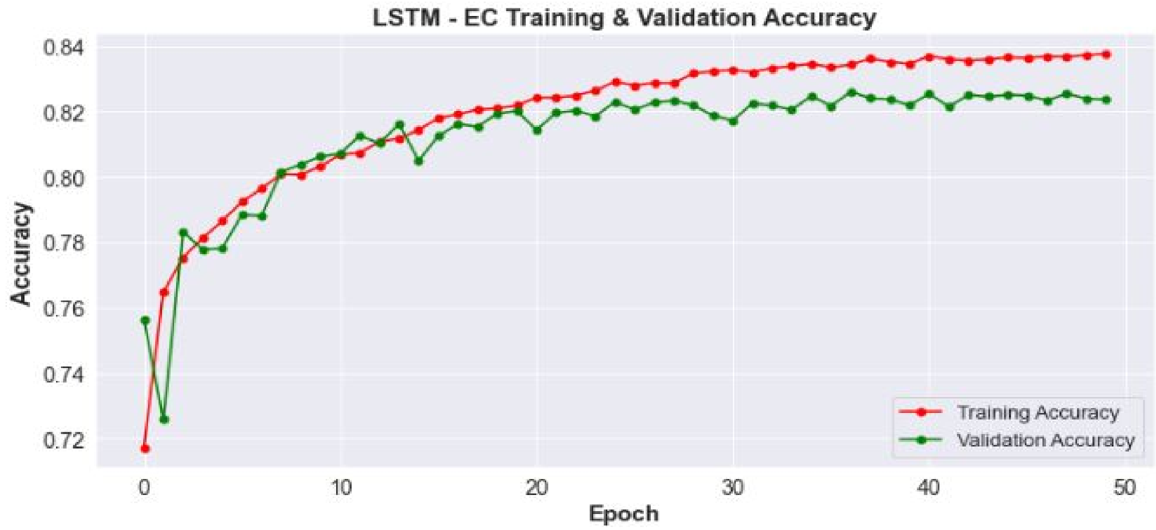

Figure 4.12: Training and Validation Accuracy of the LSTM classifier for English comedy

Figure 4.12 demonstrates that for English Comedy, the LSTM model achieved training and validation accuracy of 83.67% and 82.35%, respectively. The figure shows that the training and validation accuracy improved as the number of epochs increased. Therefore, the higher the accuracy, the better the performance of the model. Figure

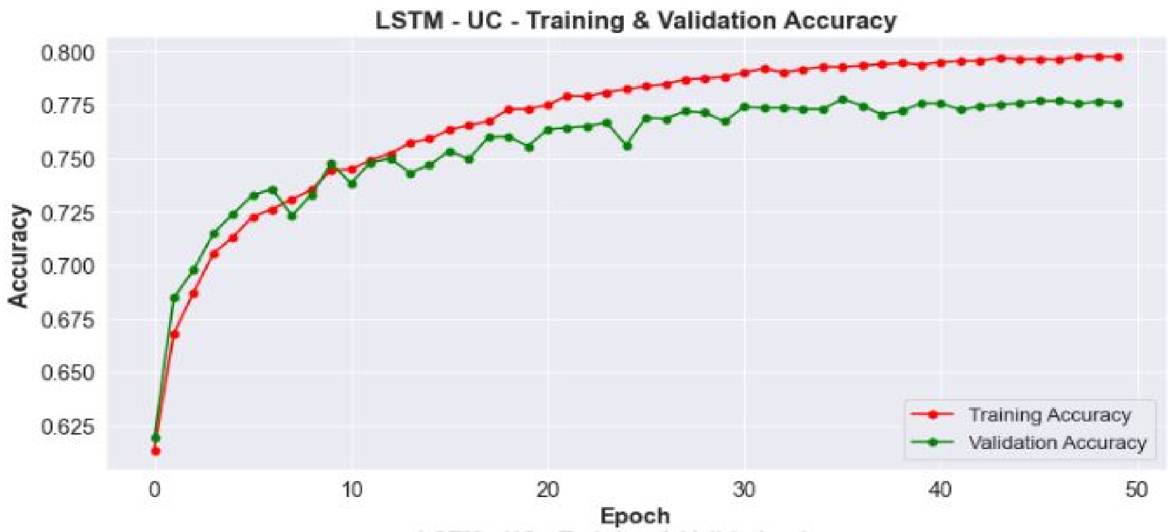

Figure 4.13: Training and Validation Accuracy of the LSTM classifier for Urdu Comedy

4.13 demonstrates that for Urdu Comedy, the LSTM model achieved training and validation accuracy of 79.28% and 77.02%, respectively. This figure also shows that the training and validation accuracy improved as the number of epochs increased.

4.3 Model Performance Evaluation

4.3.1 The LSTM Model Assessment

The performance evaluation metrics for English and Urdu Comedy clips including accuracy, precision, recall, and F1-scores for the LSTM model are given in Table 4.1 and 4.2.

LSTM: English Comedy Overall Accuracy = 81.86%			
Classes	Precision	Recall	F1-Score
Stress	0.78	0.76	0.77
Non-Stress	0.86	0.86	0.85

Table 4.1: LSTM English Comedy Accuracy

LSTM: Urdu Comedy Overall Accuracy = 77.28%			
Classes	Precision	Recall	F1-Score
Stress	0.78	0.72	0.75
Non Stress	0.77	0.81	0.79

Table 4.2: LSTM Urdu Comedy Accuracy

The results showed that the LSTM model has higher accuracy of 81.86% for the English comedy dataset as compared to Urdu comedy which is 77.28%.

The Confusion Matrix for the LSTM classifier for the English comedy clip is displayed in Figure 4.14, which shows that 76% are correctly predicted as stressed, and 24% are misclassified. 86% are correctly predicted as non-stressed while 14% are misclassified. It can be seen that the correct classification for non-stressed is higher than for the stressed one.

The Confusion Matrix for the LSTM classifier for the Urdu comedy clip is displayed in Figure 4.15, which shows that 72% are correctly predicted as stressed, and 28% are misclassified. In contrast, 81% are correctly predicted as non-stressed while 19% are misclassified. It can be seen that the correct classification for non-stressed on the LSTM classifier is again higher than for the stressed one.

It can be seen that the accuracy of non-stressed class for English comedy on the LSTM classifier is also higher than for the Urdu Comedy clip.

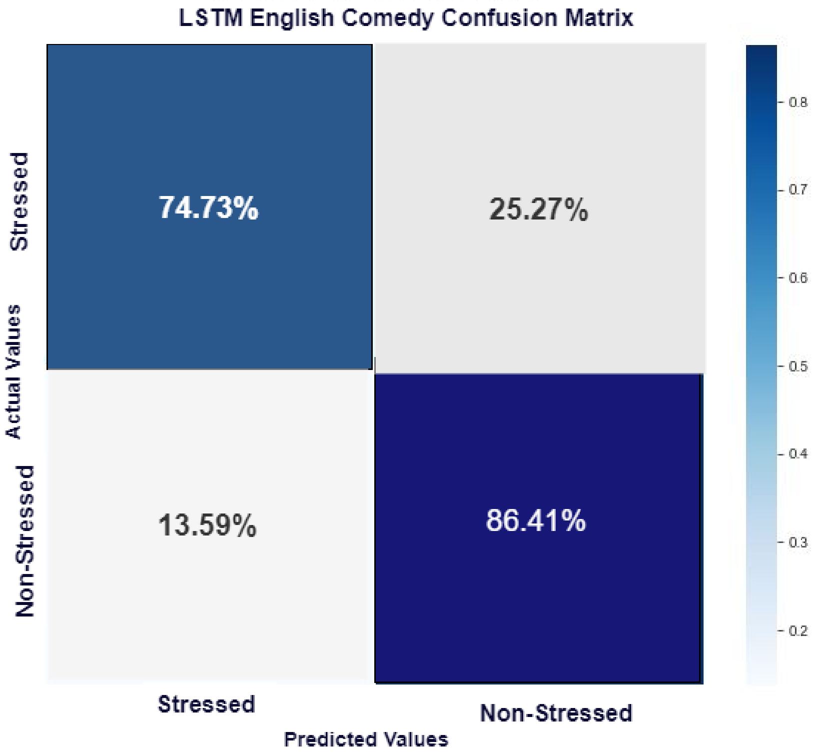

Figure 4.14: Confusion matrix of LSTM Classifier for English Comedy

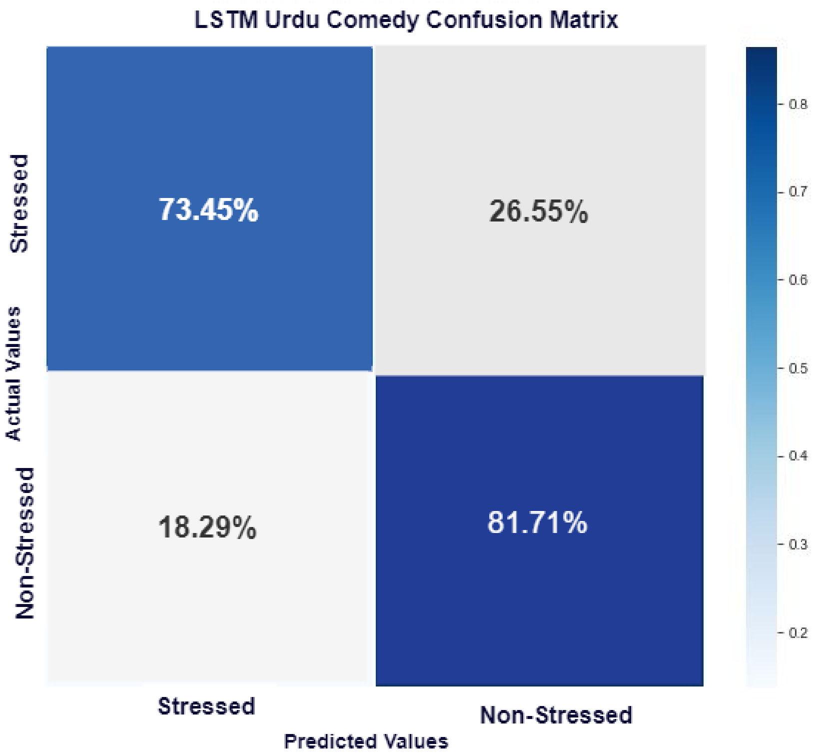

Figure 4.15: Confusion matrix of LSTM Classifier for Urdu Comedy

4.3.2 The LGBM Model Assessment

The performance evaluation metrics for English and Urdu Comedy clips including accuracy, precision, recall, and F1-scores for the LGBM model are given in Table 4.3 and Table 4.4.

LGBM: English Comedy Overall Accuracy = 82.83%			
Classes	Precision	Recall	F1 Score
Stress	0.80	0.76	0.78
Non-Stress	0.84	0.87	0.86

Table 4.3: LGBM English Comedy Accuracy

LGBM: Urdu Comedy Overall Accuracy = 78.19%			
Classes	Precision	Recall	F1-Score
Stress	0.79	0.73	0.76
Non-Stress	0.78	0.83	0.80

Table 4.4: LGBM Urdu Comedy Accuracy

The results showed that the LGBM model has higher accuracy of 82.83% for the English comedy dataset as compared to Urdu comedy which is 78.19%.

The Confusion Matrix of the LGBM classifier for the English comedy clip is displayed in Figure 4.16, which shows that 76.11% are correctly predicted as stressed, and 23.89% are misclassified. Similarly, 87.41% are correctly predicted as non-stressed while 12.59% are misclassified. It can be see that the correct classification for non-stressed on the LGBM is again higher than for the stressed one.

The Confusion Matrix of the LGBM classifier for the Urdu comedy clip is displayed in Figure 4.17, which shows that 73.18% are correctly predicted as stressed, and 26.82% are misclassified. Similarly, 82.61% are correctly predicted as non-stressed while 17.39% are misclassified.

It can be seen that the accuracy of non-stressed class for English comedy on the LGBM classifier is also higher than for the Urdu Comedy clip.

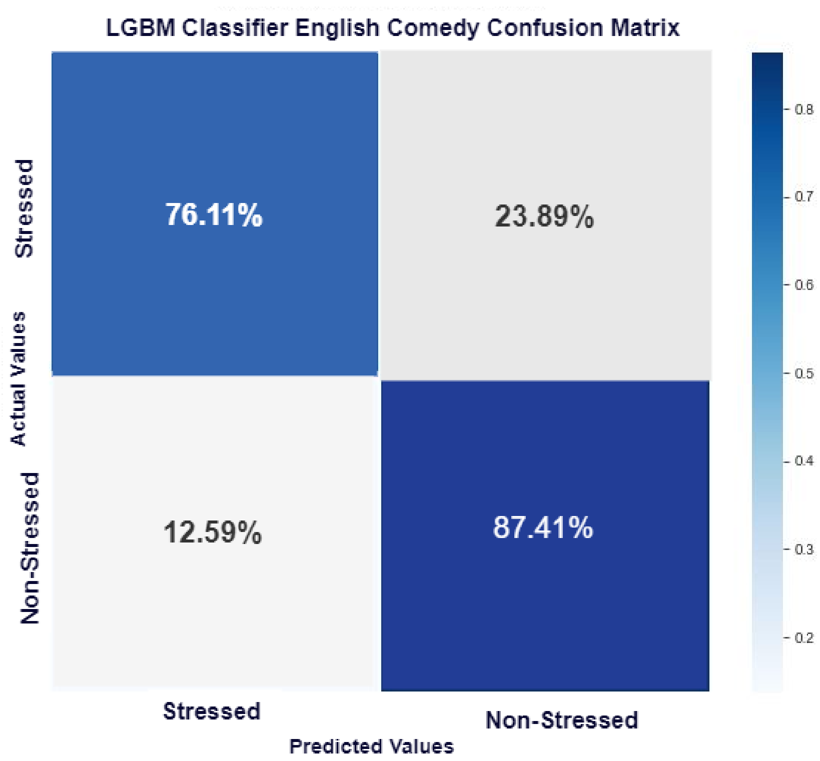

Figure 4.16: Confusion matrix of LGBM Classifier for English Comedy

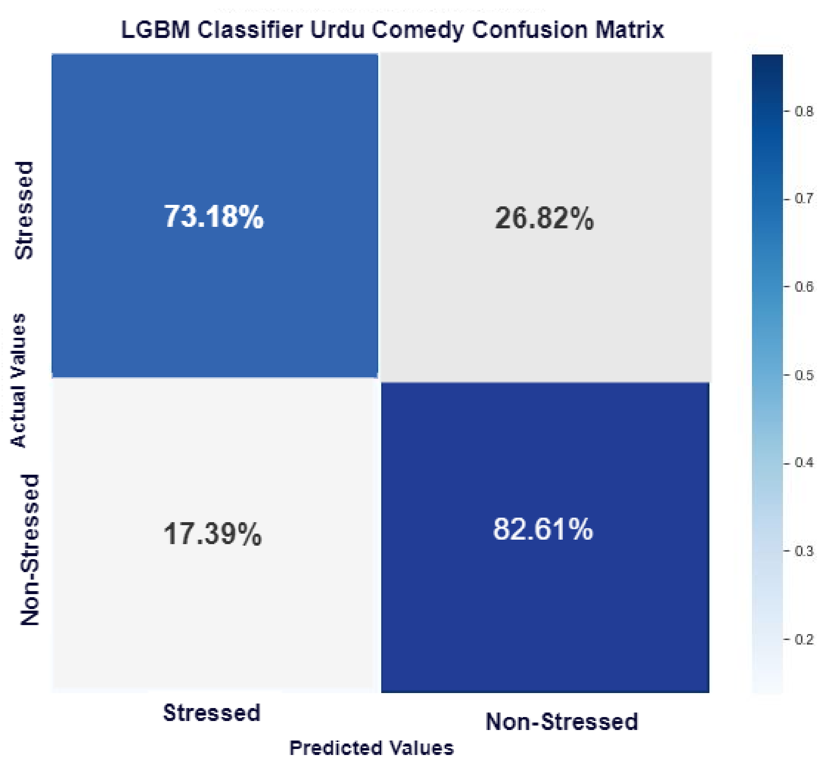

Figure 4.17: Confusion matrix of LGBM Classifier for Urdu Comedy

4.3.3 The XGB Model Assessment

The performance evaluation metrics for English and Urdu Comedy clips including accuracy, precision, recall, and F1-scores for the the XGB model are given in Table 4.5 and Table 4.6.

XGB: English Comedy Overall Accuracy = 83.52%			
Classes	Precision	Recall	F1 Score
Stress	0.81	0.78	0.80
Non-Stress	0.86	0.88	0.87

Table 4.5: XGB English Comedy Accuracy

XGB: Urdu Comedy Overall Accuracy = 78.19%			
Classes	Precision	Recall	F1-Score
Stress	0.77	0.76	0.76
Non-Stress	0.79	0.80	0.79

Table 4.6: XGB Urdu Comedy Accuracy

The results showed that the XGB model has higher accuracy of 83.52% for the English comedy dataset as compared to Urdu comedy which is 78.19%.

The Confusion Matrix of the XGB classifier for the English comedy clip is displayed in Figure 4.18, which shows that 77.68% are correctly predicted as stressed, and 22.32% are misclassified. Similarly, 87.50% are correctly predicted as non-stressed while 12.50% are misclassified. It can be seen that the correct classification for non-stressed on the XGB is again higher than for the stressed one.

The Confusion Matrix of the XGB classifier for the Urdu comedy clip is displayed in Figure 4.19, which shows that 76.47% are correctly predicted as stressed, and 23.53% are misclassified. Similarly, 79.71% are correctly predicted as non-stressed while 20.29% are misclassified. It can be seen that the correct classification for non-stressed on the XGB classifier is also higher than for the stressed one.

It can be seen that the accuracy of non-stressed class for English comedy on the XGB classifier is also higher than for the Urdu Comedy clip.

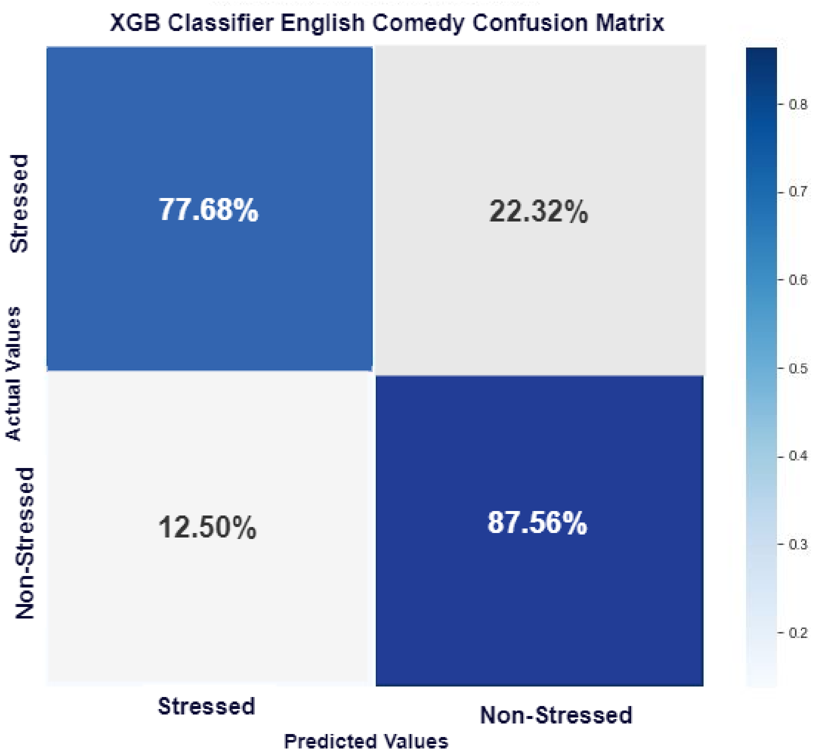

Figure 4.18: Confusion matrix of XGB Classifier for English Comedy

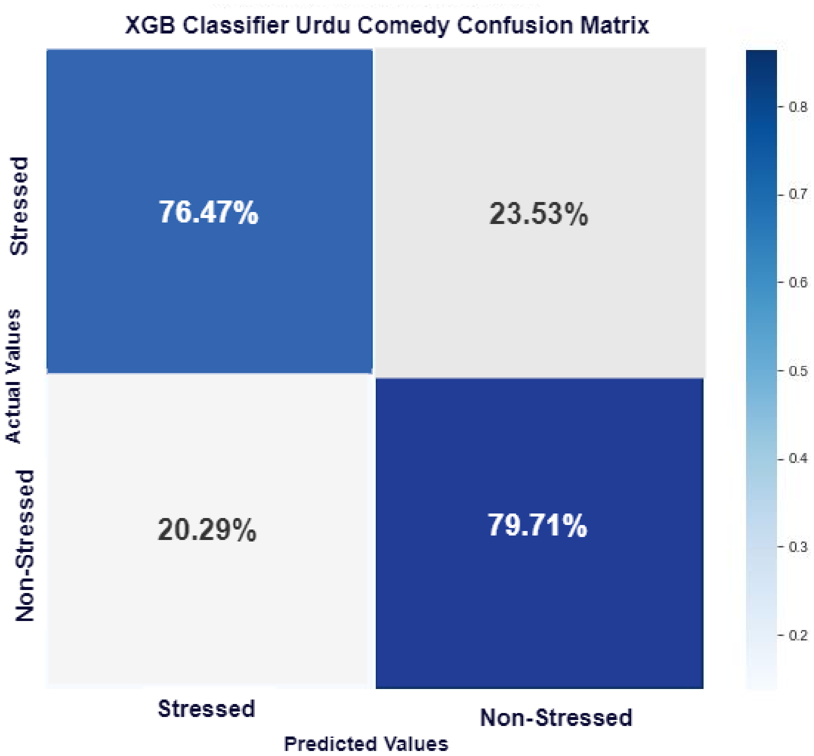

Figure 4.19: Confusion matrix of XGB Classifier for Urdu Comedy

4.3.4 The Random Forest Model Assessment

The performance evaluation metrics for English and Urdu Comedy clips including accuracy, precision, recall, and F1-scores for the random forest model are given in Table 4.7 and Table 4.8.

Random Forest: English comedy Overall Accuracy = 84.10%			
Classes	Precision	Recall	F1 Score
Stress	0.81	0.79	0.80
Non-Stress	0.86	0.88	0.87

Table 4.7: Random Forest English Comedy Accuracy

Random Forest: Urdu Comedy Overall Accuracy = 77.84%			
Classes	Precision	Recall	F1-Score
Stress	0.77	0.76	0.76
Non-Stress	0.79	0.80	0.79

Table 4.8: Random Forest Urdu Comedy Accuracy

The results showed that the Random Forest model has higher accuracy of 84.10% for the English comedy dataset as compared to Urdu comedy which is 77.84%.

The Confusion Matrix of the Random Forest classifier for the English comedy clip is displayed in Figure 4.20, which shows that 78.90% are correctly predicted as stressed, and 21.10 percent are misclassified. Similarly, 87.66% are correctly predicted as non-stressed while 12.34% are misclassified. It can be seen that the correct classification for non-stressed on the random forest is higher than for the stressed one.

The Confusion Matrix of the Random Forest classifier for the Urdu comedy clip is displayed in Figure 4.21, which shows that 75.54% are correctly predicted as stressed, and 24.46 percent are misclassified. Similarly, 79.88% are correctly predicted as non-stressed while 20.12% are misclassified. It can be seen that the correct classification for non-stressed on the random forest classifier is also higher than for the stressed one.

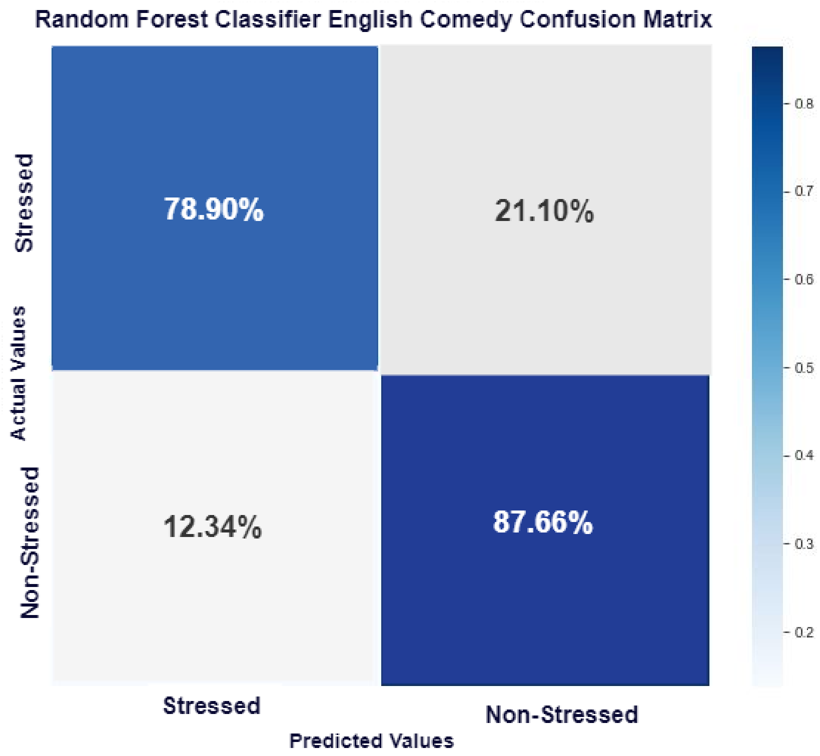

Figure 4.20: Confusion matrix of Random Forest Classifier for English Comedy

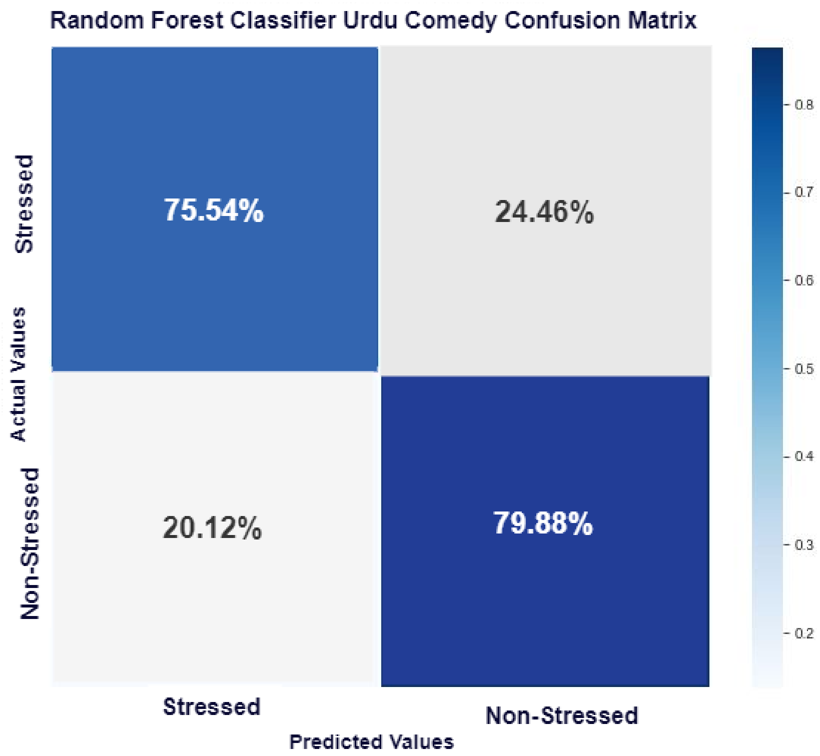

Figure 4.21: Confusion matrix of Random Forest Classifier for Urdu Comedy

4.3.5 The EXTRA Tree Model Assessment

The performance evaluation metrics for English and Urdu Comedy clips including accuracy, precision, recall, and F1-scores for the EXTRA Tree model are given in Tables 4.9 and 4.10.

ETC: EC- Overall Accuracy = 84.29%			
Classes	Precision	Recall	F1 Score
Stress	0.81	0.80	0.80
Non Stress	0.86	0.88	0.87

Table 4.9: ETC English Comedy Accuracy

ETC: UC- Overall Accuracy = 78.32%			
Classes	Precision	Recall	F1-Score
Stress	0.77	0.76	0.77
Non-Stress	0.79	0.81	0.80

Table 4.10: ETC Urdu Comedy Accuracy

The results showed that the EXTRA Tree model has higher ccuracy of 84.29% for the English comedy dataset as compared to Urdu comedy which is 78.32%

The Confusion Matrix of the EXTRA Tree classifier for the English comedy clip is displayed in Figure 4.22, which shows that 79.52% are correctly predicted as stressed, and 20.48% are misclassified. Similarly, EXTRA Tree correctly predicted 87.53% as non-stressed while 12.47% are misclassified. It can be seen that the correct classification for non-stressed on the EXTRA Tree is again higher than for the stressed one.

The Confusion Matrix of the EXTRA Tree classifier for the Urdu comedy clip is displayed in Figure 4.23, which shows that 75.77% are correctly predicted as stressed, and 24.23% are misclassified. Similarly, 80.56% are correctly predicted as non-stressed while 19.44% are misclassified. It can be seen that the correct classification for non-stressed on the EXTRA Tree classifier is also higher than for the stressed one.

It can be seen that the accuracy of non-stressed class for English comedy on the EXTRA Tree classifier is also higher than for the Urdu Comedy clip.

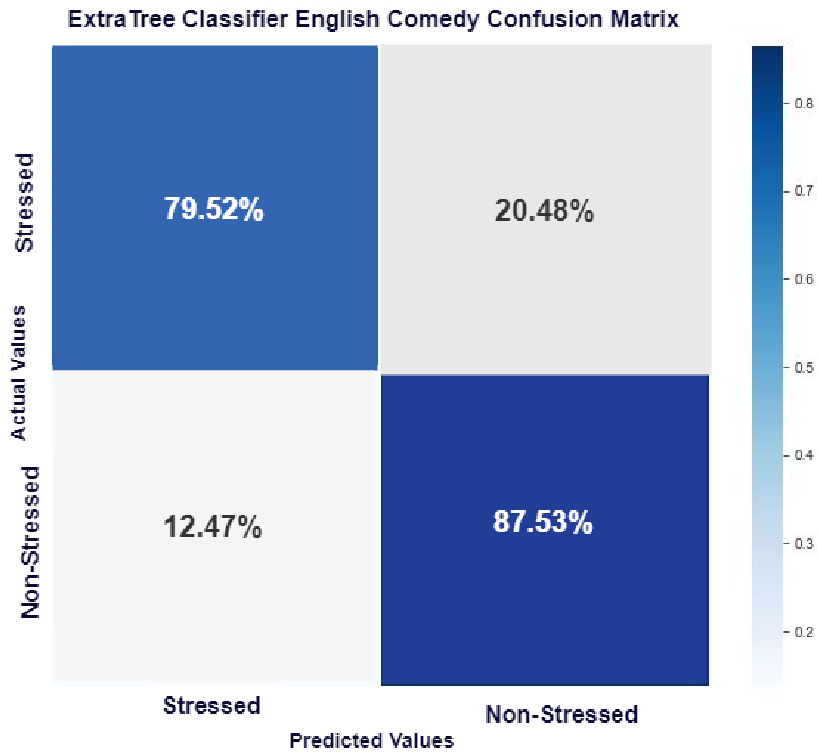

Figure 4.22: Confusion matrix of EXTRA Tree Classifier for English Comedy

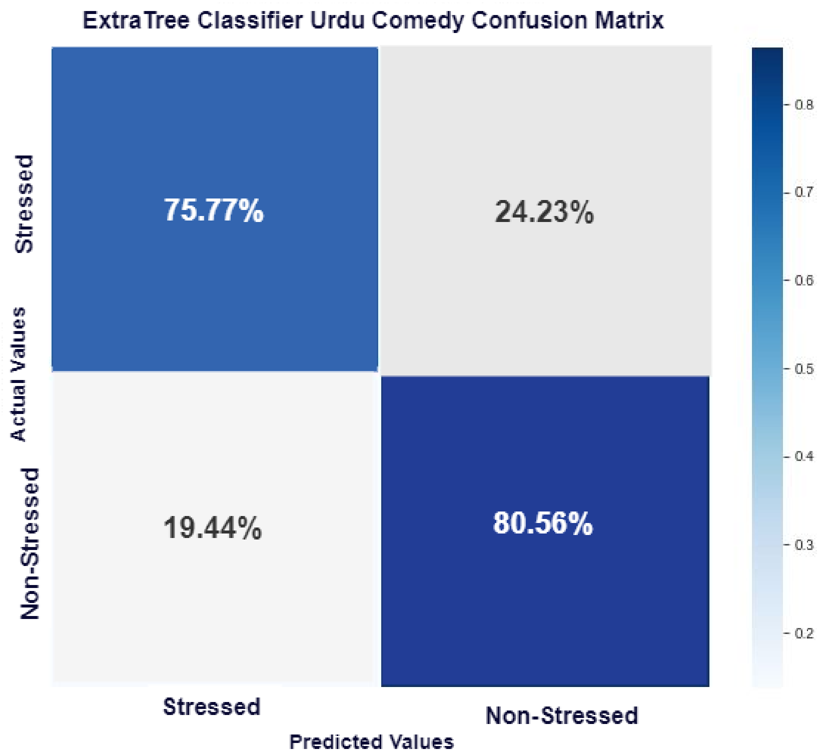

Figure 4.23: Confusion matrix of EXTRA Tree Classifier for Urdu Comedy

It is evident from the results from above five different classifiers that English comedy clip have more influence on stress reducing as compared to Urdu comedy clip, since a significant difference is reported in the results.

4.4 Comedy clips - Personality Relationship

By using ANOVA,“Agreeableness” and “Openness” provided some insights about the participants and how they respond to the comedy clips with p-values of 0.04. The rest of the traits (Extrovert, Conscientiousness, and Neuroticism) showed no significant difference between low and high levels, irrespective of the video being watched.

4.4.1 English comedy vs. Personality Trait

English vs. Agreeableness Trait			
Agreeableness	Stressed	Non-Stressed	Total
High	6	11	17
low	6	7	13
English vs. Openness Trait			
Openness	Stressed	Non-Stressed	Total
High	5	10	15
low	7	8	15

Table 4.11: Stress analysis and Personality traits assessment for English Comedy

For the English Comedy clip, it can be seen in Table 4.11, that out of 30 participants, 64% of the people with high agreeableness are non-stressed while 36% with low levels of agreeableness are stressed while watching these comedy clips. In the case of the openness to experience trait, 66.67% of people with high levels have reported being non-stressed while 33.33% of the participants with low levels are stressed while watching these videos.

It can also be concluded from Figure 4.24, and Figure 4.25 for stress vs. personality traits, that participants are generally non-stressed while watching English stand-up comedy, irrespective of the level of trait, i.e., agreeableness or openness, they possess.

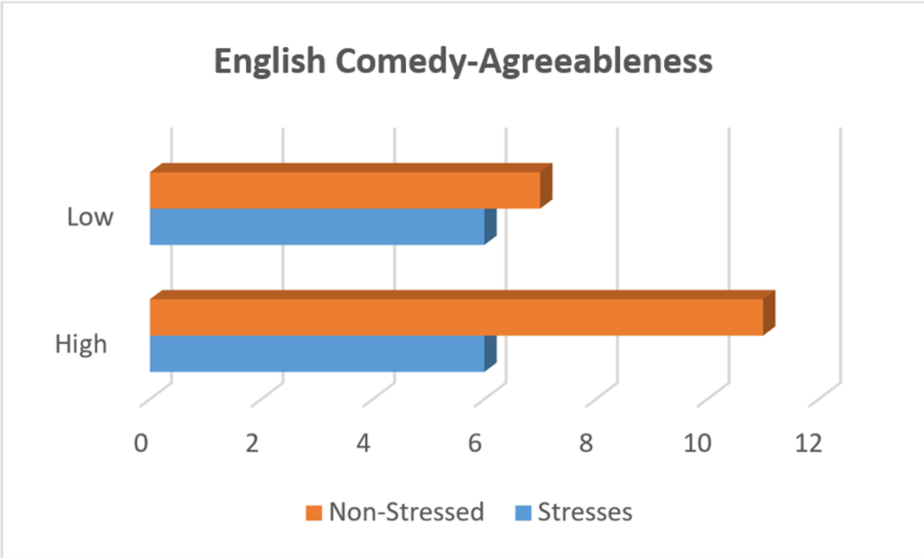

Figure 4.24: Stress analysis for English Comedy versus Low and High Scores of Agreeableness

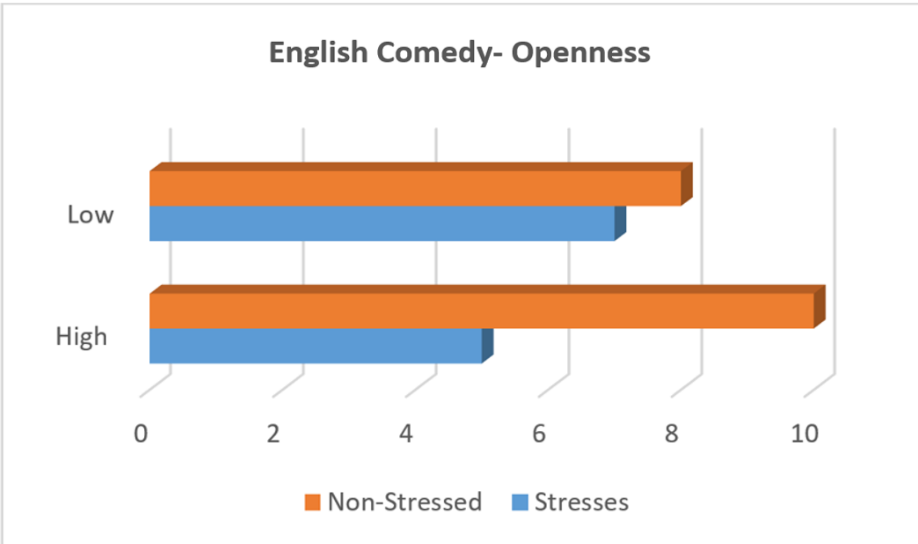

Figure 4.25: Stress analysis for English Comedy versus Low and High Scores of Openness

4.4.2 Urdu comedy vs. Personality Trait

For the Urdu comedy clip, it can be seen in Table 4.12, that out of 30 participants, people with 52% of high agreeableness scores are non-stressed while 48 percent of people with low levels of agreeableness are stressed, and in case of low level of agreeableness scores, 53% of the participants are non-stressed while 47% of participants are stressed. On the other side, Openness to Experience trait, people with a 73% of high openness trait, are non-stressed, while 27% of the participants are stressed. In the case of the low openness to experience trait, 66.67% of the participants are highly stressed while 33.33% are non-stressed.

Urdu vs. Agreeableness Trait			
Agreeableness	Stressed	Non-Stressed	Total
High	8	9	17
low	6	7	13
Urdu vs. Openness Trait			
Openness	Stressed	Non-Stressed	Total
High	4	11	15
low	10	5	15

Table 4.12: Stress analysis and Personality traits assessment for Urdu Comedy

It can also be concluded from Figures 4.26, and Figure 4.27, that for agreeableness case, participants are generally non-stressed while watching Urdu stand-up comedy clips, on the contrary, for the low level of openness trait they possess highly stress.

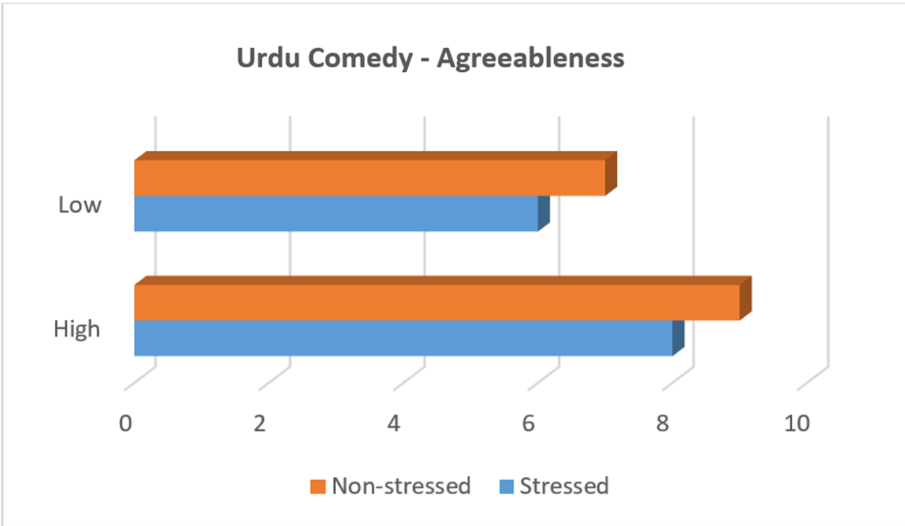

Figure 4.26: Stress analysis for Urdu Comedy versus Low and High Scores of Agreeableness

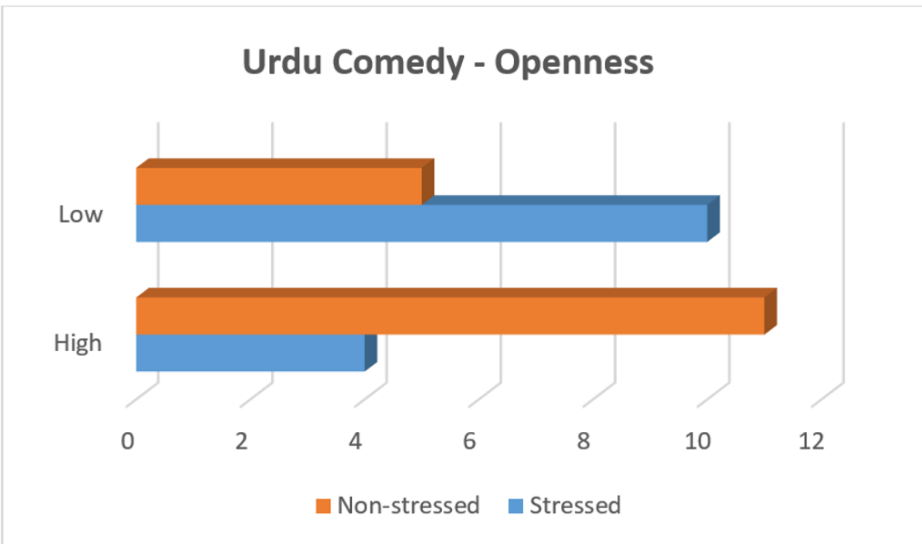

Figure 4.27: Stress analysis for Urdu Comedy versus Low and High Scores of Openness

Chapter 5

Conclusion and Recommendation

5.1 Concluding Remarks

Using standup comedy clips as stress-reducing stimuli have rarely been observed in EEG-based research. Also, diverse EEG-based brain activity datasets are not available, as taking EEG is daunting. It is now possible to handle such data and do an exploratory analysis of brain activity with cost-effective headsets available. This research aimed to build a dataset by measuring scalp EEG using a single-channel device, Neurosky Mindwave Mobile 2. non-native language English and Native language Urdu standup comedy were watched by each individual, and EEG signals were recorded. We have focused on using supervised machine learning algorithms to classify stress by using comedy clips. All Machine learning algorithms performed better for English comedy clips. EXTRA Tree Classifier outperformed with the highest accuracy of 84.29% for English and 78.32% percent for Urdu. Random forest was performed with an accuracy of 84.01% percent and 77.84%, followed by XGB with 83.52% and 78.19% for English and Urdu comedy clips respectively. It is evident from the results that English comedy has more influence on stress level reduction as compared to Urdu comedy. For personality assessment, it has been identified that from English comedy a large number of participants' stress decreased irrespective of the participants' personality traits.

5.2 Recommendations

The research can be enhanced further by taking a different population sample and testing the sample with the same conditions of the experiment as in this research to analyze how the EEG signals and resulting stress perform. Another future recommendation is to try other classifiers to test how they perform. Other than Agreeableness and Openness to Experiences traits, personality types can be tested to check how they fare for a different sample. Since we are using Time and Wavelet domains only, another area that needs to be analyzed further is the feature selection from the frequency domain also and how it can be optimized for EEG signals that help in identifying emotions induced from these comedy clips.

References

[1] Richard S. Lazarus. *Stress and emotion : a new synthesis / Richard S. Lazarus.* Springer Pub. Co., New York, N.Y, 1999 - 1999. ISBN 0826112501.

[2] Ahmad Rauf Subhani, Wajid Mumtaz, Mohamed Naufal Bin Mohamed Saad, Nidal Kamel, and Aamir Saeed Malik. Machine learning framework for the detection of mental stress at multiple levels. *IEEE access*, 5:13545–13556, 2017. ISSN 2169-3536.

[3] Eero Kajantie and David IW Phillips. The effects of sex and hormonal status on the physiological response to acute psychosocial stress. *Psychoneuroendocrinology*, 31(2):151–178, 2006.

[4] L. A. Dempsey. Stress-induced depression. *Journal of Samples*, 19(2):99, Jan 2018.

[5] Amin A. Muhammad Gadit and Gerry Mugford. Prevalence of depression among households in three capital cities of pakistan: Need to revise the mental health policy, . URL https://doi.org/10.1371/journal.pone.0000209.

[6] Amin A. Muhammad Gadit and Gerry Mugford. Prevalence of depression among households in three capital cities of pakistan: Need to revise the mental health policy. *PLOS ONE*, 2(2):1–5, 02 2007. doi: 10.1371/journal.pone.0000209. URL https://doi.org/10.1371/journal.pone.0000209.

[7] Amin A. Muhammad Gadit and Gerry Mugford. Prevalence of depression among households in three capital cities of pakistan: Need to revise the mental health policy, . URL https://doi.org/10.1371/journal.pone.0000209.

[8] Objective measures, sensors and computational techniques for stress recognition and classification: A survey. *Computer Methods and Programs in Biomedicine*, 108(3):1287–1301, 2012. ISSN 0169-2607. doi: https://doi.org/10.1016/j.cmpb.2012.07.003.

[9] 'over 34% pakistanis suffer from anxiety', Oct 2019. URL https://tribune.com.pk/story/2076914/34-pakistanis-suffer-anxiety.

[10] Anita Werner O'Toole and Sheila Rouslin Welt. *Professional Closeness*, pages 230–243. Macmillan Education UK, London, 1994. ISBN 978-1-349-13441-0. doi: 10.1007/978-1-349-13441-0_17. URL https://doi.org/10.1007/978-1-349-13441-0_17.

[11] Michael Feuerstein, Elise E. Labbé, and Andrzej R. Kuczmierczyk. *Health Behavior*, pages 237–277. Springer US, Boston, MA, 1986. ISBN 978-1-4899-0562-8. doi: 10.1007/978-1-4899-0562-8_8. URL https://doi.org/10.1007/978-1-4899-0562-8_8.

[12] N. H. Kalin T. Steckler and J. H. M. Reul. Handbook of stress and the brain part 1: The neurobiology of stress. *The Neurobiology of Stress*, 2005.

[13] T. Kamarck S. Cohen and R. Mermelstein. A global measure of perceived stress. *Journal of health and social*, pages 385–396, 1983.

[14] A. D. Smith L. H. Miller and L. Rothstein. The stress solution: An action plan to manage the stress in your life. *Pocket*, 1994.

[15] George W. Brown and Tirril Harris. Social origins of depression: a reply. *Psychological Medicine*, 8(4):577–588, 1978. doi: 10.1017/S0033291700018791.

[16] Tess Byrd O'Brien and Anita DeLongis. Coping with chronic stress. In *Coping with chronic stress*, pages 161–190. Springer, 1997.

[17] Blair Wheaton. The specification of chronic stress: Models and measurement. In *annual meeting of the Society for the Study of Social Problems, August, Cincinnati, OH*, 1991.

[18] Elaine Walster, Ellen Berscheid, and G William Walster. New directions in equity research. In *Advances in experimental social psychology*, volume 9, pages 1–42. Elsevier, 1976.

[19] Ingun Ulstein, Torgeir Bruun Wyller, and Knut Engedal. High score on the relative stress scale, a marker of possible psychiatric disorder in family carers of patients with dementia. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, 22(3):195–202, 2007.

[20] Jennifer E Wartella, Stephen M Auerbach, and Kevin R Ward. Emotional distress, coping and adjustment in family members of neuroscience intensive care unit patients. *Journal of psychosomatic research*, 66(6):503–509, 2009.

[21] Phillip J Brantley, Craig D Waggoner, Glenn N Jones, and Neil B Rappaport. A daily stress inventory: Development, reliability, and validity. *Journal of behavioral medicine*, 10(1):61–73, 1987.

[22] David Watson, Lee Anna Clark, and Auke Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063, 1988.

[23] P Schulz and W Schlotz. The trier inventory for the assessment of chronic stress (tics): scale construction, statistical testing, and validation of the scale work overload. *Diagnostica*, 45(1):8–19, 1999.

[24] Kathryn A Lee, Gregory Hicks, and German Nino-Murcia. Validity and reliability of a scale to assess fatigue. *Psychiatry research*, 36(3):291–298, 1991.

[25] CD Spielberger, F Gonzalez-Reigosa, A Martinez-Urrutia, LF Natalicio, and DS Natalicio. The state-trait anxiety inventory. revista interamericana de psicologia. *Interamerican Journal of Psychology*, 5(3 & 4), 2017.

[26] Linda C Sobell, Tony Toneatto, Mark B Sobell, Reinhard Schuller, and Maxine Maxwell. A procedure for reducing errors in reports of life events. *Journal of Psychosomatic Research*, 34(2):163–170, 1990.

[27] John R McQuaid, Scott M Monroe, John R Roberts, Sheri L Johnson, Gregory L Garamoni, David J Kupfer, and Ellen Frank. Toward the standardization of life stress assessment: Definitional discrepancies and inconsistencies in methods. *Stress Medicine*, 8(1):47–56, 1992.

[28] Surya Sari Faradiba, Cholis Sa'dijah, and Swasono Rahardjo. Identification of mathematics anxiety through gesture. In *International Conference on Education (ICE2) 2018: Education and Innovation in Science in the Digital Era*, pages 881–889, 2016.

[29] Tayyaba Azim, M Arfan Jaffar, and Anwar M Mirza. Fully automated real time fatigue detection of drivers through fuzzy expert systems. *Applied Soft Computing*, 18:25–38, 2014.

[30] Giorgos Giannakakis, Matthew Pediaditis, Dimitris Manousos, Eleni Kazantzaki, Franco Chiarugi, Panagiotis G Simos, Kostas Marias, and Manolis Tsiknakis. Stress and anxiety detection using facial cues from videos. *Biomedical Signal Processing and Control*, 31:89–101, 2017.

[31] Timo Partala and Veikko Surakka. Pupil size variation as an indication of affective processing. *International journal of human-computer studies*, 59(1-2):185–198, 2003.

[32] Martijn Haak, Steven Bos, Sacha Panic, and Léon JM Rothkrantz. Detecting stress using eye blinks and brain activity from eeg signals. *Proceeding of the 1st driver car interaction and interface (DCII 2008)*, pages 35–60, 2009.

[33] Wenhui Liao, Weihong Zhang, Zhiwei Zhu, and Qiang Ji. A real-time human stress monitoring system using dynamic bayesian network. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)-workshops*, pages 70–70. IEEE, 2005.

[34] Stephanie Gillespie, Elliot Moore, Jacqueline Laures-Gore, Matthew Farina, Scott Russell, and Yash-Yee Logan. Detecting stress and depression in adults with apha-

sia through speech analysis. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5140–5144. IEEE, 2017.

[35] Tin Lay Nwe, Say Wei Foo, and Liyanage C De Silva. Speech emotion recognition using hidden markov models. *Speech communication*, 41(4):603–623, 2003.

[36] Stevo Julius. Autonomic nervous system dysregulation in human hypertension. *The American journal of cardiology*, 67(10):B3–B7, 1991.

[37] Pedro Ferreira, Pedro Sanches, Kristina Höök, and Tove Jaensson. License to chill! how to empower users to cope with stress. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, pages 123–132, 2008.

[38] Sue C Jacobs, Richard Friedman, John D Parker, Geoffrey H Tofler, Alfredo H Jimenez, James E Muller, Herbert Benson, and Peter H Stone. Use of skin conductance changes during mental stress testing as an index of autonomic arousal in cardiovascular research. *American heart journal*, 128(6):1170–1177, 1994.

[39] Tao Lin, Masaki Omata, Wanhua Hu, and Atsumi Imamiya. Do physiological data relate to traditional usability indexes? In *Proceedings of the 17th Australia conference on computer-human interaction: Citizens online: Considerations for today and the future*, pages 1–10. Citeseer, 2005.

[40] Yu Shi, Natalie Ruiz, Ronnie Taib, Eric Choi, and Fang Chen. Galvanic skin response (gsr) as an index of cognitive load. In *CHI'07 extended abstracts on Human factors in computing systems*, pages 2651–2656, 2007.

[41] Mohit Kumar, Matthias Weippert, Reinhard Vilbrandt, Steffi Kreuzfeld, and Regina Stoll. Fuzzy evaluation of heart rate signals for mental stress assessment. *IEEE Transactions on fuzzy systems*, 15(5):791–808, 2007.

[42] Stephen BRE Brown, Jos F Brosschot, Anke Versluis, Julian F Thayer, and Bart Verkuil. Assessing new methods to optimally detect episodes of non-metabolic heart rate variability reduction as an indicator of psychological stress in everyday

life: A thorough evaluation of six methods. *Frontiers in neuroscience*, 14:564123, 2020.

[43] Gari D Clifford. *Signal processing methods for heart rate variability*. PhD thesis, Oxford University, UK, 2002.

[44] Lizawati Salahuddin and Desok Kim. Detection of acute stress by heart rate variability using a prototype mobile ecg sensor. In *2006 International Conference on Hybrid Information Technology*, volume 2, pages 453–459. IEEE Computer Society, 2006.

[45] Stanley Reisman. Measurement of physiological stress. In *Proceedings of the IEEE 23rd Northeast Bioengineering Conference*, pages 21–23. IEEE, 1997.

[46] Keri J Heilman, Mika Handelman, Gregory Lewis, and Stephen W Porges. Accuracy of the stresseraser® in the detection of cardiac rhythms. *Applied psychophysiology and biofeedback*, 33(2):83–89, 2008.

[47] Jennifer A Healey and Rosalind W Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on intelligent transportation systems*, 6(2):156–166, 2005.

[48] Z Dharmawan and LJM Rothkrantz. Analysis of computer game player stress level using eeg data. In *11th international conference on computer games: AI, animation, mobile, educational and serious games, La Rochelle, France*, pages 111–124. The University of Wolverhampton, 2007.

[49] Isabela Lobo, Liana Catarina Portugal, Ivan Figueira, Eliane Volchan, Isabel David, Mirtes Garcia Pereira, and Leticia de Oliveira. Eeg correlates of the severity of posttraumatic stress symptoms: a systematic review of the dimensional ptsd literature. *Journal of Affective Disorders*, 183:210–220, 2015.

[50] Ori Amir, Irving Biederman, Zhuangjun Wang, and Xiaokun Xu. Ha ha! versus aha! a direct comparison of humor to nonhumorous insight for determining the neural correlates of mirth. *Cerebral Cortex*, 25(5):1405–1413, 2015.

[51] James F O'Connell and Jim Allen. When did humans first arrive in greater australia and why is it important to know? *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 6(4):132–146, 1998.

[52] Saeid Sanei and Jonathon A Chambers. *EEG signal processing*. John Wiley & Sons, 2013.

[53] Pil Hyung Lee, Jae-Kwan Song, Byung Joo Sun, Hyung Oh Choi, Jeong-Sook Seo, Jin Oh Na, Dae-Hee Kim, Jong-Min Song, Duk-Hyun Kang, Jae-Joong Kim, et al. Outcomes of patients with stress-induced cardiomyopathy diagnosed by echocardiography in a tertiary referral hospital. *Journal of the American Society of Echocardiography*, 23(7):766–771, 2010.

[54] Wai King Tsang and Matilda Wong. Constructing a shared 'hong kong identity'in comic discourses. *Discourse & Society*, 15(6):767–785, 2004.

[55] Karli K Watson, Benjamin J Matthews, and John M Allman. Brain activation during sight gags and language-dependent humor. *Cerebral cortex*, 17(2):314–324, 2007.

[56] Ksenija Marinkovic, Sharelle Baldwin, Maureen G Courtney, Thomas Witzel, Anders M Dale, and Eric Halgren. Right hemisphere has the last laugh: neural dynamics of joke appreciation. *Cognitive, Affective, & Behavioral Neuroscience*, 11(1):113–130, 2011.

[57] Fatma Nasoz, Kaye Alvarez, Christine L Lisetti, and Neal Finkelstein. Emotion recognition from physiological signals using wireless sensors for presence technologies. *Cognition, Technology & Work*, 6(1):4–14, 2004.

[58] Benjamin Cowley, Marco Filetti, Kristian Lukander, Jari Torniainen, Andreas Henelius, Lauri Ahonen, Oswald Barral, Ilkka Kosunen, Teppo Valtonen, Minna Huotilainen, et al. The psychophysiology primer: a guide to methods and a broad review with a focus on human–computer interaction. *Foundations and Trends® in Human–Computer Interaction*, 9(3-4):151–308, 2016.

[59] Jonathan R Wolpaw and Elizabeth Winter Wolpaw. 25| the future of bcis: Meeting the expectations. *Brain-Computer Interfaces: Principles and Practice*, page 387, 2012.

[60] Jonathan R Wolpaw, Niels Birbaumer, William J Heetderks, Dennis J McFarland, P Hunter Peckham, Gerwin Schalk, Emanuel Donchin, Louis A Quatrano, Charles J Robinson, Theresa M Vaughan, et al. Brain-computer interface technology: a review of the first international meeting. *IEEE transactions on rehabilitation engineering*, 8(2):164–173, 2000.

[61] Clemens Brunner, Niels Birbaumer, Benjamin Blankertz, Christoph Guger, Andrea Kübler, Donatella Mattia, José del R Millán, Felip Miralles, Anton Nijholt, Eloy Opisso, et al. Bnci horizon 2020: towards a roadmap for the bci community. *Brain-computer interfaces*, 2(1):1–10, 2015.

[62] Jonathan R Wolpaw and Dennis J McFarland. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. *Proceedings of the national academy of sciences*, 101(51):17849–17854, 2004.

[63] Philip R Kennedy, Roy AE Bakay, Melody M Moore, Kim Adams, and John Goldwaithe. Direct control of a computer from the human central nervous system. *IEEE Transactions on rehabilitation engineering*, 8(2):198–202, 2000.

[64] J. d. R. Millán. Brain-computer interfaces. *The MIT Press*, 2002.

[65] Nadja Olini, Iru Rothfuchs, Damiano Azzinnari, Christopher R Pryce, Salome Kurth, and Reto Huber. Chronic social stress leads to altered sleep homeostasis in mice. *Behavioural Brain Research*, 327:167–173, 2017.

[66] Richard S Lewis, Nicole Y Weekes, and Tracy H Wang. The effect of a naturalistic stressor on frontal eeg asymmetry, stress, and health. *Biological psychology*, 75(3): 239–247, 2007.

[67] George carlin – airline announcement (part 1 of 2), Jun 2021. URL https://www.youtube.com/watch?v=nPZKz8RGeyc.

[68] Omer Shareef Emirates Airlines, Jan 2015. URL https://www.youtube.com/watch?v=GxcULJolZcA.

[69] Youngmee Kim and Larry Seidlitz. Spirituality moderates the effect of stress on emotional and physical adjustment. *Personality and Individual differences*, 32(8): 1377–1390, 2002.

[70] Omneya Attallah. An effective mental stress state detection and evaluation system using minimum number of frontal brain electrodes. *Diagnostics*, 10(5):292, 2020.

[71] Anum Asif, Muhammad Majid, and Syed Muhammad Anwar. Human stress classification using eeg signals in response to music tracks. *Computers in biology and medicine*, 107:182–196, 2019.

[72] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.