ASHESI UNIVERSITY

A PORTABLE AND LOW-COST ELECTROENCEPHALOGRAPHY DEVICE WITH AUTOMATED AUTISM DIAGNOSIS

CAPSTONE PROJECT

B.Sc. Electrical and Electronics Engineering

Hannah Boadiwaa Lormenyo

2021

ASHESI UNIVERSITY

A Portable and Low-cost Electroencephalography Device with Automated Autism Diagnosis

CAPSTONE PROJECT

Capstone Project submitted to the Department of Engineering, Ashesi University, in partial fulfillment of the requirements for the award of Bachelor of Science degree in Electrical and Electronics Engineering.

Hannah Boadiwaa Lormenyo

2021

DECLARATION

I hereby declare that this capstone is the result of my own original work and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature:

……………………………………………………………………………………………

Candidate's Name:

……………………………………………………………………………………………

Date:

……………………………………………………………………………………………

I hereby declare that preparation and presentation of this capstone were supervised in accordance with the guidelines on supervision of capstone laid down by Ashesi University.

Supervisor's Signature:

……………………………………………………………………………………………

Supervisor's Name:

……………………………………………………………………………………………

Date:

……………………………………………………………………………………………

Acknowledgements

I would like to express my gratitude to my supervisor Dr. Nathan Amanquah, whose continuous encouragement and advice have helped me throughout this journey.

Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterized by challenges in speech and communication, impairment in social skills, and repetitive behavior. Autistic children in Africa have very severe symptoms of autism due to the lateness in the diagnosis and treatment of autism on the continent. This work explores the use of a portable, low-cost electroencephalography (EEG) device with automated diagnosis as a means of expediting the process of autism diagnosis in Africa. This work compares two instrumentation amplifier designs for the EEG system. It also compares k-nearest neighbor, support vector machine, decision tree, and random forest as classifiers for providing automated diagnosis for the EEG system. The resulting design was a portable EEG system that can be interfaced with a smartphone for real-time visualization of the EEG signals and automated diagnosis with an accuracy of 85.1%.

Contents

Chapter 1: Introduction

1.1 Background

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that is very difficult to detect due to the variety of ways it presents itself. It is characterized by an impairment in social skills, challenges in speech and communication, and repetitive behavior. The various forms of autism include "classical" autism disorder, Asperger's syndrome or high-functional autism, and pervasive developmental disorder (atypical autism). Children with these different forms of autism show varying levels of intellectual disability, language delays, and difficulty in social and communication skills [17]. In Africa, specifically Ghana, there is little knowledge about the prevalence of ASD, even though research has shown that ASD is present on the continent.

It has been reported that 38.7% of children under 14 years are affected with ASD in Ghana [18]. In other African countries like Tunisia, Egypt, and Southeast Nigeria, the prevalence of ASD has been found to be 33.6%, 11.5%, and 11.4%, respectively [3][4]. Due to the little knowledge of ASD in some African countries, people still harbor the belief that it is caused by supernatural sources and seek the assistance of traditional healers before sending affected children to medical professionals for effective medical screening and testing [18]. This has led to the late detection of ASD in Africa and has increased the severity of ASD in most autistic African children. For most countries in the western world, diagnosis of ASD commonly occurs in the first three years of development. On the other hand, children in Africa are not diagnosed until they are about 8 years old.

Through research and various studies, the five-year difference has been attributed to a number of reasons. Inadequate healthcare facilities and poorly trained personnel, little

ASD knowledge or awareness, cultural beliefs and stigma, and delay in medical attention have been mentioned as the causes of the late diagnosis of ASD in Africa [5]. Due to this, various therapies such as speech and behavioral therapy for these autistic children are further delayed leading to very low IQ, severe intellectual disability, and impairments in communication and social skills amongst the children.

The conventional means of diagnosing ASD involves a series of behavioral and interview tests. During these tests, medical professionals measure the child's behavior excesses, communication, self-care, and social skills for a long time to correctly diagnose the child as autistic or not. ASD clinical and self-screening methods include Autism Diagnostic Interview-Revised (ADI-R), Autism Diagnostic Observation Schedule (ADOS), Childhood Autism Rating Scale (CARS), Joseph Picture self-concept scale, social responsiveness scale, and Diagnostic and Statistical Manual of mental disorders, 5th edition (DSM-5) [36]. DSM-5 classifies all the different forms of ASD under pervasive developmental disorders (PDD) and is the most recent and widely used technique. One of the limitations of DSM-5 is its dependability on the atypical behaviors displayed by autistic children, which eliminates the possibilities for early detection in the first three months of development. This limitation became the motivation of most research studies and led to the discovery of other biological indicators such as electroencephalography (EEG) as a means of diagnosing ASD [8].

EEG is an electrophysiological test that measures electrical activities in the human brain using metal discs attached to the scalp. Over the years, EEG has been used by doctors in diagnosing epilepsy and sleep disorders. An electroencephalogram is a special device that is usually used in recording EEG signals. This device uses 8, 16, or 32 electrodes attached to four standard positions on the scalp: the nasion, inion, and the right and left. The

electrodes are connected to the scalp using a special headgear with already installed electrodes or a special adhesive-conducting gel. An EEG signal is usually recorded as a change in voltage between two paired electrodes, and its amplitude is measured using the peak-to-peak technique [13].

1.2 Motivation

Traditional EEG devices are very bulky (weighs more than 500g), and they cost thousands of euros [21]. They generally have a short battery lifespan (2-3 days) because they consume too much power and are not portable. This does not make them ideal for deployment in rural areas where there is no or intermittent power supply and little funds to afford such equipment. Due to these limitations of traditional EEG devices, many scientists and researchers have developed numerous less cumbersome designs that consume little power. However, most of these implementations need a medical professional with domain expertise in electroencephalograms to provide a diagnosis from the recorded signals. The approach used in this work automates the process of diagnosing patients and can be used and operated by any medical personnel. This automated diagnosis feature makes this design suitable for use in almost every hospital with medical officers, even without domain expertise, across the globe.

1.3 Problem Definition

As a result of the unfortunate paucity of EEG devices in most countries in Africa, specifically Ghana, severe symptoms of autism have been quite dominant among autistic children. The disturbing rarity of EEG devices on the African continent can be attributed to the low number of highly trained neurological specialists and the high-priced traditional and commercial EEGs, which most government hospitals cannot afford. Other reasons include the high power consumption rate and the weightiness of conventional EEG devices. There

is a need for a low-cost EEG device that can automatically diagnose patients with little or no assistance from a specialist. This work aims to satisfy this need.

1.4 Objectives

The main objective of this work is to design and implement a portable and low-cost electroencephalography device that can be interfaced with a smartphone for the automated diagnosis of autism in children. To be more specific, this work is aimed at:

- designing a low power consumption unit for the device

- shortening the time used for electroencephalograms by giving an automated autism diagnosis that is at least 85 percent accurate.

- implementing a low-cost analog and digital system for the conditioning and processing of EEG signals.

- architecting a reliable and low-cost access technology network for the EEG device to ensure efficient transmission of data.

1.5 Proposed Solution

The EEG will be obtained from the scalp of the children using dry active electrodes. Dry active electrodes are easy to place on the scalp, comfortable to wear, and do not require massive hygienic procedures, unlike wet electrodes with high setup time and limited recording time due to the drying up of conductive gels [24]. Since the magnitude of the EEG is microvolt-size with a bandwidth of 0.1 Hz to 100 Hz, a high gain amplifier will be used in amplifying the signal for transmission. The EEG signal is known for a lot of artifacts, such as ocular and motor artifacts. Ocular artifacts occur in the EEG signal when a person blinks during recording, while movements during recording cause motor artifacts. There are also artifacts generated from cardiovascular activity, physical displacement of the electrodes, and power supply noise [20]. Hence, after the amplification of the signal,

4

filtering is done to remove all these artifacts. Bandpass filters will serve the purpose of eliminating very low and extremely high frequencies from the signal. The resultant signal is then converted to a digital signal using an analog-to-digital converter. The digital signal will then be transformed from the time domain to the frequency domain and then sent to the smart device using low-cost access technology. Bluetooth Low Energy (BLE) will be used for this cause because, unlike Wi-Fi, it consumes low power [21]. A wireless medium is preferred over a wired one in this case since a wired medium is more prone to errors and artifacts due to wire movements.

After the data is transferred to the smartphone, it will be viewed in real-time via a mobile application installed on the smartphone. Also, the mobile application will automatically classify the data received as autistic or neurotypical. The mobile application will use machine learning in diagnosing an individual as neurotypical or autistic using the EEG signals recorded from the electrodes on the scalp of the individual.

1.6 Scope

The scope of this work encompasses the design of a suitable signal acquisition and processing unit for the EEG device. All the design decisions for each component of the EEG system will be discussed in chapter 4. The design presented in this work uses a cross-platform mobile application that will run on both iOS and Android smartphones with a Bluetooth interface. The different features of the mobile application, including real-time visualization of signals and automated diagnosis, will also be highlighted in chapter 4. The design and implementation of the machine learning model used in the accurate diagnosis of the individual will be clearly emphasized. Furthermore, the performance of the different models used in this work will be analyzed in chapter 5. This work does not address real-life testing of the designed EEG device on neuro-typical and autistic children.

Chapter 2: Literature Review

Investigators from various research institutions have proved that EEG data can be used to effectively predict if a child has ASD as early as three months of age [30]. EEG is preferred to other emerging technologies such as eye-tracking and Magnetic Resonance Imaging (MRI) due to its relatively low cost and non-invasiveness [31]. Currently, there has been little or no work on an end-to-end system that provides fast and automated EEG interpretation. Most work that has been done in this space can be divided into three groups depending on the focus of the research. The groups include research works with a focus on:

    i.   low-cost and portability of EEG device [24][26]

    ii.   classification of recorded EEG signals [11]

    iii.   end-to-end solutions [20]

2.1 Works with focus on low-cost and portability of EEG devices

In [24], Puyol et al. presented a portable embedded system that records up to 128 EEG electrodes simultaneously using the evoked-related potentials method. The 128 Al/AlCl plated active electrodes are amplified, and every 8-set of electrodes is connected to a multiplexer amplifier system. This system channels the signal to a microcontroller with an in-built 12-bit analog-to-digital converter. The inbuilt AD converts the signal to a digital signal which is then passed to a second microcontroller to be stored on an SD flash memory or sent to a computer using a Bluetooth interface. The design presented by the authors is portable, consumes low power, and has very low noise. However, this work does not cover high-frequency sampling since Fast Fourier Transforms were not involved. Also, the design uses two Microchip PICs, 24FJ128GA306: one for recording the signals and doing an analog-to-digital conversion, and the other for saving the data onto an SD card and transmitting the data via Bluetooth. The cost could have been reduced by using one

6

microcontroller. Furthermore, all the data can be saved on the cloud after sending via Bluetooth to prevent data loss through the SD card's corruption or misplacement.

Other researchers developed a medical instrumentation device for real-time electroencephalography recording and analysis [26]. This device obtains the EEG signal using electrodes placed on the scalp. The signal is amplified and filtered using a high gain amplifier and bandpass filter, respectively. The signal is converted to a digital signal using a 12-bit analog-to-digital converter. It is then transformed into a frequency band using radix-2 fast Fourier transform on a Field Programmable Gate Array (FPGA). The resulting spectrum is then transmitted to a mobile device via Wi-Fi. This design focused on the simplicity and portability of the EEG system. However, it did not emphasize other factors such as power consumption, bandwidth efficiency, cost, and low noise. Even though the access technology they used (Wi-Fi) is faster, it consumes more power than other technologies like Bluetooth. Hence, better design decisions could have been made to cut down on cost and power consumption.

2.2 Works with focus on classification of EEG signals

Djemal et al. investigated a new autism diagnosis procedure using discrete wavelet transform (DWT) and Shannon entropy [11]. The DWT was used in decomposing the EEG data into several frequency sub-bands. Feature extraction was performed on the different sub-bands using several statistical functions (mean, variance, and standard deviation) and entropy functions (threshold entropy, Renyi entropy, log energy, and Shannon entropy). An Artificial Neural Network (ANN) was used to classify the different EEG segments as autistic or normal based on the extracted features.

Quite contrary to the design in [11], Sheikhani et al. used spectrogram and coherence values in analyzing quantitative EEG. The analysis revealed that alpha waves (8 -13 Hz)

7

could differentiate between autistic and children without autism with 96.4% accuracy in awake conditions using spectrogram criteria[32].

2.3 End-to-end solutions

While some designs have admirable features like portability, low power consumption, and low cost, other designs like the one proposed in [20] are unique. This design focuses on easing the process of EEG analysis for both specialized and non-specialized hospital personnel by converting it to music. Also, it presents an action protocol if a problem is detected. The researchers showed the feasibility of this method by using low-cost smartphone technologies with energy-efficient algorithms. The signal recorded from the electrodes with a frequency range of 0.5-13Hz is filtered and down-sampled to 32 Hz during the preprocessing stage. A phase vocoder is used to make the design suitable for real-time usage since it does not affect the time duration and preserves the spectral characteristics of the signal. The signal is transformed to the frequency domain using Discrete Fourier Transform (DFT) with a sampling rate of 32 Hz.

2.3 Relevant concepts/technologies

A typical EEG system includes electrodes, a filtering block, and an amplification unit. In [24], two amplification units were used. The first amplification was performed on the signals recorded by the electrodes. INA333 instrumentation amplifier was used because it has very low noise at low frequencies (50 nV Hz 1- 100Hz) where EEG signals can be found. Additionally, it has a very high CMRR (above 100dB), which helps eliminate common-mode voltages. OPA171 operational amplifier was used in the second amplification, which was done after filtering the EEG signals. The operational amplifier labeled as U2 in Figure 1 is connected to two low-pass filters. One of these filters acts as the anti-aliasing filter while the other removes dc components from the signal.
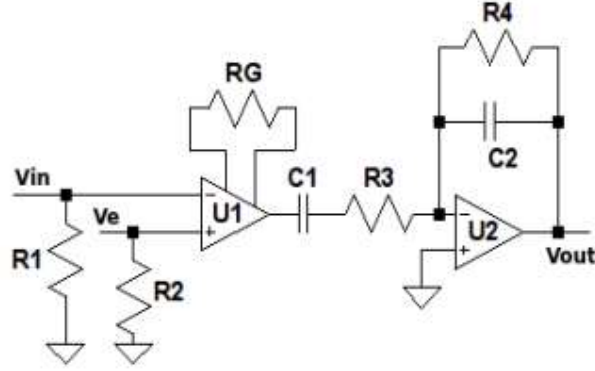
Figure 1: Amplifier unit for EEG [25]

A typical outpatient EEG takes about 20 minutes to complete [32]. This is because the recorded EEG signals are traditionally recorded and interpreted by board-certified physicians. The interpretation of the EEG signal is a time-consuming process. It can take professionals hours to weeks to provide an interpretation of the recorded EEG. Unfortunately, most of the interpretations provided by the professionals are subjective and often lead to misdiagnosis of patients [33]. Automating the interpretation of EEG data would speed up the process and eliminate cases of misdiagnosis. Since there has not been much work on an end-to-end solution that records the EEG signals and provides a diagnosis, this work exists to fill that void.

The design presented in this work will elaborate on a low-power EEG system that is very portable, low-cost, and automatically interprets the EEG data recorded. This approach will be pursued to ease the process of analysis for both specialized and non-specialized hospital personnel. Machine learning algorithms will be used in analyzing the EEG signals to provide an accurate diagnosis.

Chapter 3: Requirements

This chapter lays out the requirements for the proposed solution. The EEG system will consist of two main subsystems: the hardware and the software subsystems. The individual requirements for the hardware and software subsystems have been outlined below.

3.1 Hardware requirements

- The device must consume low power to make it suitable for use in mobile situations.

- The EEG system must have a high gain instrumentation amplifier with a large CMRR to amplify the signal and eliminate common-mode input.

- The device must be a wearable device that is comfortable to wear.

- The recording process must take less than 5 minutes.

- The device must use a low-cost access technology that supports the bitrate of the EEG system.

3.2 Software requirements

The software subsystem consists of a mobile application that will run on both Android and iOS platforms. The following are the functions of the mobile application:

- It will facilitate the real-time visualization of the EEG signals from the brain of the patient.

- It will present an accurate automated diagnosis to the user when the recording is complete.

The main objective of the application is to make it easier for professionals (with or without expertise in EEG) to interpret EEG signals. The following are the requirements of the application.

- The application should run on mobile devices with Android Jellybean v16, 4.1.x or newer, and iOS 8 or newer.

- It should generate a diagnosis from the EEG signals with an accuracy greater than or equal to 85%.

- It must have a user-friendly interface.

- The application should be able to interpret recorded signals in less than 2 minutes.

The sections below will explain the features of the application and highlight their functional requirements.

3.2.1 Real-time visualization of the EEG signals

3.2.1.1 Description and Priority

This is a high-priority feature that allows the user to view the signals in real-time as they are being recorded.

3.2.1.2 Stimulus/Response Sequences

- The user must be logged in.

- The phone must be paired with the Bluetooth of the EEG device (appears as HC-05/HC-06 with 1234 as the password).

- The user must connect to the paired HC-05/HC-06 EEG device.

3.2.1.3 Functional Requirements

- REQ-1: The software must show the data from the 4 channels in real-time.

- REQ-2: The user should be able to pause, stop and restart the recording.

- REQ-3: The data should automatically be backed up in a cloud database.

3.2.2. System Feature 2: Automated diagnosis using the recorded EEG signals.

3.2.2.1 Description and Priority

This is a high-priority feature that allows the user to quickly interpret the signals within a short time by providing an automated diagnosis that shows whether the signals are normal or abnormal.

3.2.2.2 Stimulus/Response

- The user must press the diagnose button to submit the recorded signal to the API for automated diagnosis.

3.2.2.3 Functional Requirements

- REQ-1: The API must return the diagnosis in JSON format.

- REQ-2: The software must display the diagnosis to the user in a user-friendly format.

- REQ-3: The data should automatically be backed up in a cloud database.

Chapter 4: Design and Implementation

This chapter outlines the design decisions made for each hardware and software subsystem of the EEG device.

4.1 Block Diagram

This work presents an end-to-end solution starting from the recording of the signals to its interpretation. As illustrated in the block diagram below, the EEG signals will be recorded using five electrodes placed on the scalp. One of these electrodes will be a reference for the other electrodes. The difference between the signal from each electrode and the reference electrode signal will be amplified by an instrumentation amplifier with a gain of 11 and filtered using a bandpass filter with a 5-16 Hz passband. The resulting signal from each channel is then amplified using an operational amplifier with a gain of 101. A microprocessor converts the signal into digital format and transmits it to a smartphone via Bluetooth. A mobile application that will be created in this work will be installed on the smartphone. The application makes it possible to visualize the signals in real-time. It also provides an automated diagnosis of the signals using machine learning.
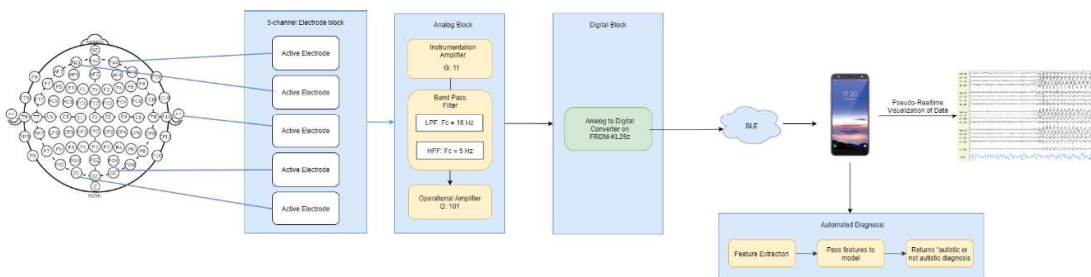

Figure 2: Block Diagram of the EEG system

4.2 Hardware Design and Implementation

The design decisions made for each component illustrated in the block diagram above will be discussed in this section.

4.2.1 Montage

A montage refers to the orderly and logical arrangement of EEG channels (electrode pairs). There are various kinds of montages, such as common electrode reference montages, bipolar montages, average reference montages, and the Laplacian montage.

i. Bipolar montage: Each channel takes the potential difference between two adjacent active electrodes. One major disadvantage of this montage is the in-phase cancellation of biological activity. This means that the differential amplifier cancels out any point in the two biological waveforms that are relatively synchronous with respect to time and amplitude.

ii. Common reference or referential montage: Each channel takes the potential difference between an active electrode and an inactive electrode (designated reference electrode). An earlobe or mastoid positions A1 and A2, scalp position Cz and the nose are the commonly used reference electrodes. The choice of a reference is critical as more electrical activity in the reference electrode might dominate the channel and not give the desired results.

iii. Average reference montage: In this montage, the outputs of all the amplifiers are averaged to produce an averaged signal. This averaged signal is then used as a common reference for each channel.

iv. Laplacian montage: Each channel in this montage is the difference between an active electrode and a weighted average of the surrounding electrodes.

14

From the various types of montages mentioned above, the significance of the reference electrode to be used is emphasized. The best reference electrode is one that cancels out most noise without canceling much of the electrocerebral activity. Due to the left-sided position of the heart, the left earlobe is often contaminated with EKG artifacts. A reference electrode on the nose may be contaminated with muscle artifact as well. Even though the midline vertex electrode Cz is free from muscle artifacts, there is a large amount of brain wave activity, especially during sleep. Also, the reference electrode must be far away from the other electrode. This is because a higher interelectrode distance results in EEG signals with higher amplitudes[35]. Considering all the stated factors, the right earlobe, A2, was chosen as the reference electrode to be used in the common-electrode reference montage.

The electrodes will be placed on FP1, FP2, O1, O2, and A2 positions on the scalp using the 10-20 international system. This means that there will be four channels with A2 as the reference electrode. The electrodes were placed based on the 10-20 international system [34]. This system is a universally accepted system used to describe the location of scalp electrodes, as shown in Figure 3.
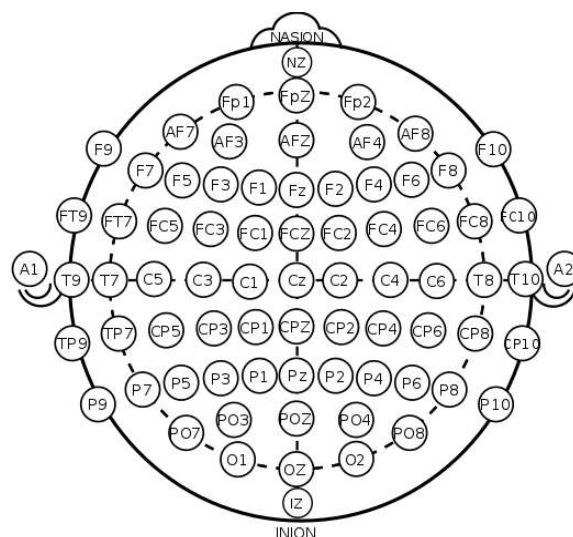

Figure 3: 10-20 EEG placement system

15

4.2.2 Electrodes

These are the metal discs placed on the scalp to measure electrical activity in the brain. There are three main types of electrodes: dry electrodes, semi-dry electrodes, and wet electrodes, that can be used for the EEG system. Semi-dry electrodes use tap water between the scalp and the electrode [26], while wet electrodes use highly conductive electrolytes like gel or saline solutions. The different types of electrodes were compared using a Pugh chart. The cost, signal-to-noise ratio, setup time, and ease of cleaning were used as the basis of comparison between the electrodes.

Table 1: Pugh chart for electrodes

Criteria	Baseline	Dry Electrodes	Semi-Dry Electrodes	Wet Electrodes
Cost	0	+2	+1	+1
Signal to Noise ratio	0	-1	+1	+2
Setup Time	0	+1	-1	-1
Easy to clean	0	+1	-1	-1
Total	0	+4	0	+1

As shown in the Pugh chart above, wet electrodes have a better signal-to-noise ratio. As such, they require simple amplifiers that are less costly and easy to design. Dry electrodes do not require any substance between the scalp and the electrode, making them very easy to install. However, dry electrodes have a small signal-to-noise ratio due to the high contact impedance between the skin and the sensor [27]. Compared to wet and semi-dry electrodes,

dry electrodes are easier to clean and set up for recording EEG signals. Dry electrodes were chosen since they satisfy the criteria highlighted in the Pugh chart.

4.2.3 Instrumentation Amplifier

This unit is essential as the EEG signal has a magnitude of 1 microvolt, with a bandwidth of 0.1 Hz to 100 Hz. The amplifier to be used must have an input range that includes the minimum and maximum values of the EEG signals. One key challenge to tackle is the offset introduced by the EOGs, EMGs, and other interferences that can cause the EEG signal to exceed the input range of the amplifier [30]. When this happens, some parts of the signal are attenuated and lost. The amplifier receives EEG signals with high impedance ranging from 1k to 1M [28]. Therefore, to prevent the loss of signal resolution, the amplifier must have high input impedance. It has been suggested that an amplifier should have a common input impedance of at least 100M to confine the signal attenuation below 1% [29]. An instrumentation amplifier would be used to amplify the EEG signals. Instrumentation amplifier ICs such as the INA333, AD8553, AD620, and AD8226 can be used. An instrumentation amplifier can be designed using resistors and operational amplifiers, as described in section 4.2.10.

4.2.4 Band Pass Filter

The filtering unit should cut out the noise and other artifacts in the signal such that only the data within the acceptable bandwidth is further processed. The acceptable bandwidth is 6-16Hz. This bandwidth captures the alpha signals well enough as they are generally within a 7.5-13 Hz band.

A bandpass filter consists of a series combination of a low-pass filter and a high-pass filter. The specifications of these filters have been outlined in the sub-sections below.

4.2.4.1 Low Pass Filter

The cut-off frequency (F_c) for the low-pass filter is 16 Hz. Consequently, frequencies above 16Hz will be filtered. The equation shows the relationship between the cut-off frequency, resistance, and capacitance of the RC filter:

$$F_c = \frac{1}{2\pi RC}$$

Assuming a capacitance $C = 0.01\mu F$, the resistance R for the low-pass filter circuit was calculated as follows:

$$R = \frac{1}{2\pi F_c C}$$

$$R = \frac{1}{2\pi * 16 * 0.01 * 10^{-6}} = 994718 \ \Omega \ \approx 1 \ M\Omega$$

4.2.4.2 High Pass filter

The cut-off frequency (F_c) for the high-pass filter was chosen to be 6 Hz. This means that frequencies below 6 Hz will be attenuated. Assuming a capacitance $C = 0.01\mu F$, the resistance R for the high-pass filter circuit was calculated as follows:

$$R = \frac{1}{2\pi F_c C}$$

$$R = \frac{1}{2\pi * 6 * 0.01 * 10^{-6}} = 2.65258 \ M\Omega \ \approx 3 \ M\Omega$$

4.2.4.3 Bandpass filter circuit

A bandpass filter circuit was constructed using the resistance and capacitance values for the low-pass and high-pass filters designed in sections 4.2.4.1 and 4.2.4.2.
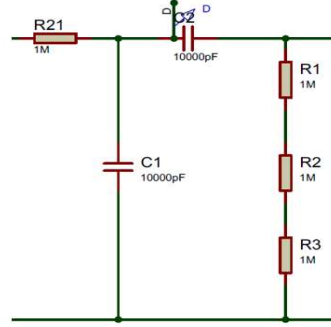
Figure 4: Bandpass Filter

Figure 4 shows the bandpass filter with a capacitance of $C = 0.01\mu F$ for both the low-pass and high-pass filters and a resistance of $1\ \text{M}\Omega$ for the low-pass filter and $3\ \text{M}\Omega$ for the high-pass filter. The $3\ \text{M}\Omega$ resistance was created using a series connection of three $1\ \text{M}\Omega$ resistors.

4.2.5 Processing unit

Since most of the computations pertain to signal processing, a Field Programmable Gate Array (FPGA), a microcontroller that supports DSP, or a DSP processor can be used. The architecture of DSP processors is optimized for signal processing. Most of these processors need a software to be loaded onto them since they lack flash program memory. Unlike microcontrollers, DSP processors are faster in handling floating-point operations and mathematical operations. However, DSP processors might not be fully optimized for Bluetooth serial communication, which is essential in this work.

The main requirements for the processing unit include cost-efficiency, processing speed, and performance. The FPGA hardware architecture is very malleable and can be constructed to meet the needs of the project. This means that the FPGA can be programmed to process more extensive data with fewer clock cycles. Unfortunately, the hardware construction might take a lot of time due to factors such as the unavailability of some

19

components. Moreover, using more hardware for the FPGA implies an undesirable increase in the form factor of the EEG device.

Microprocessors are limited by the bus, which is commonly 16-bit or 32-bit in size. They have inbuilt ADC systems for converting analog signals into digital signals. Besides signal processing, the microprocessor is also required to transfer data to a human-machine interface to visualize the EEG data. Since ARM Cortex M0+ microprocessor has an inbuilt system for ADC and can easily connect to a Bluetooth device over a serial connection, it was chosen over the FPGA and DSP processors. It was also selected due to its availability.

4.2.6 Analog-To-Digital Conversion unit

At this subsystem, the analog EEG signal is converted to a digital format. The ARM cortex M0+ has a 16-bit ADC, which does conversions using successive approximation. The ADC unit has 14 channels and can be triggered by both software and hardware.

$$minimum\ sampling\ rate = 2\ x\ maximum\ frequency$$

$$minimum\ sampling\ rate = 2\ x\ 16\ Hz = 32\ Hz$$

The sampling rate that was chosen for the ADC was 40 Hz which satisfies Nyquist conditions. The sampling frequency was determined to be slightly higher than the Nyquist frequency to prevent data loss.

4.2.7 Access Technology

It is essential to preserve the integrity of the data and consume low power in transmitting the data. Some access technologies like Wi-Fi, Bluetooth, and Zigbee have certain characteristics suitable for the EEG system. All these access technologies operate at the same frequency of about 2.4 GHz.

$$bitrate = sampling\ rate\ \times\ bit\ depth\ \times\ Number\ of\ channels$$

$$bitrate = 40\ \times 16 \times 4 = 2560\ bps$$

The access technology to be used must have a minimum transfer speed of 2.56 kbps. The transfer speeds of Wi-Fi, Zigbee, and Bluetooth are 54 Mbps, 250 kbps, and 700 kbps, respectively which are all favorable for the system. As summarised in the Pugh chart below, Wi-Fi works very well for long-range applications, while Zigbee is most appropriate for battery-powered systems. Despite this advantage that ZigBee offers, ZigBee applications cannot be controlled via a smartphone but from a unique device. Bluetooth is generally less expensive than Wi-Fi and ZigBee as it does not need any external router or network for close-range communication.

Table 2: Pugh Chart for access technology

Criteria	Baseline	Wi-Fi	ZigBee	Bluetooth
Cost	0	-2	-1	+2
Transfer speed	0	+3	+1	+2
Range	0	+3	+2	+1
Power Consumption	0	-2	+2	+1
Popularity	0	+1	-1	+1
Total	0	+3	+3	+6

From the Pugh chart analysis, Bluetooth is the most suitable access technology. It was chosen because of its popularity in mobile devices, low power consumption rate, and low cost. It also works within an appropriate range and has a suitable transfer speed.

4.2.8 Human Machine Interface

A low-cost approach for this will be a liquid crystal display (LCD). However, users will have minimal interaction with the process if an LCD is used. Also, if an LCD is to be used, all the operations would have to be completed on the microprocessor, which will demand a lot of computational power. This might be overwhelming for the microprocessor and might slow down the process. Desktops can facilitate appropriate user interaction. However, they are bulky and might eliminate the possibility of making the device portable. Also, LCD and desktops are not as popular as smartphones. Therefore, the adoption of the device may not be as widespread as it is expected to be.

Table 3: Pugh Chart for Human Machine Interface

Criteria	Baseline	Desktop	LCD	Smartphone
Cost	0	-1	+2	+1
Form Factor	0	-1	+2	+1
Ease of Analysis	0	+2	-2	+1
Power Consumption	0	-1	+2	+1
Total	0	-1	+4	+5

The Pugh chart analysis in Table 3 shows that a smartphone will be the most appropriate human-machine interface for the EEG system. It does not compromise the cost and portability of the EEG device. It also consumes a relatively low amount of power and makes it easier to view and analyze the signals.

4.2.9 Power Unit

The device must consume low power to make it suitable for use in mobile situations. Since an amplifier that operates using DC power will be used, the EEG system will be battery-powered using a 9V battery.
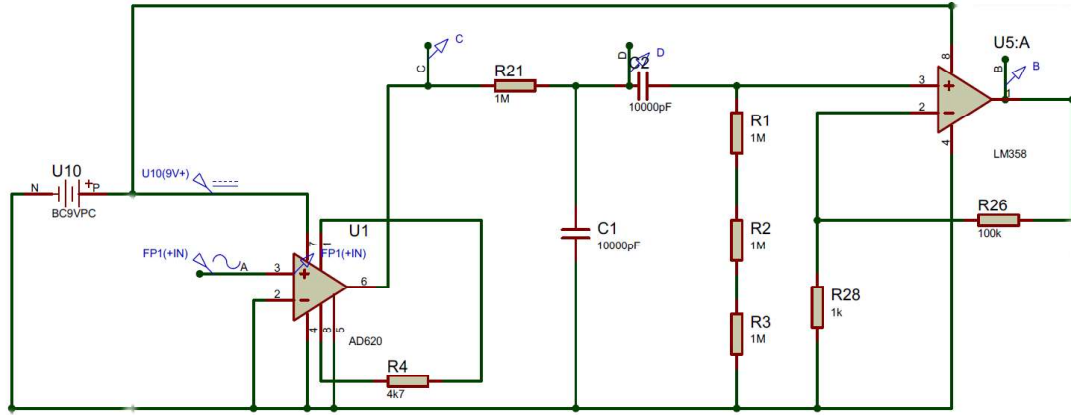
4.2.9 Full Design of Analog Block

Figure 5:1-channel EEG acquisition system with an AD620 instrumentation amplifier.

Figure 4 shows a design that uses an AD620 instrumentation amplifier. The gain of this instrumentation amplifier is controlled by the R_G resistor. The equation for the gain of the amplifier is given below:

$$Gain = \left(1 + \frac{49.4\ K}{R_G}\right)$$

A $4.7\ K\Omega$ resistor was used as R_G. The resulting gain of the amplifier was 11.51. The first amplification stage of the signal using an instrumentation amplifier is meant to boost the

signal-to-noise ratio (SNR) of the signal. Using a high gain would also amplify the noise and other artifacts in the signal, so a moderately low gain was used.

After the first stage of amplification, the signal is filtered by a first-order RC bandpass filter with a lower and an upper cut-off frequency of 6 Hz and 16 Hz. The output of the bandpass filter is a filtered and an attenuated signal. Hence, the signal is amplified using a non-inverting amplifier with a gain of 101.
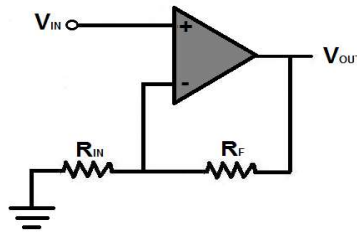

Figure 6: Non-inverting amplifier

$$Gain = 1 + \frac{R_f}{R_{IN}}$$

$$R_F = 100 \ K\Omega, \ R_{IN} = 1 \ K\Omega$$

$$Gain = 1 + \frac{100 \ K}{1 \ K} = 101$$

This gain was chosen because the resulting amplified output will have a considerable amplitude (in the kilovolts), making the signal detectable enough and preventing the amplifier from saturating (Discussed in section 4.2.11).

4.2.10 Alternative Design for Analog Subsystem

Instead of using a commercial instrumentation amplifier such as AD620, an instrumentation amplifier can be designed and created using operational amplifiers. The design consists of 2 buffer amplifiers and one difference amplifier. The gain of the instrumentation amplifier

is dictated by the values of the resistors used in conjunction with the operational amplifiers.

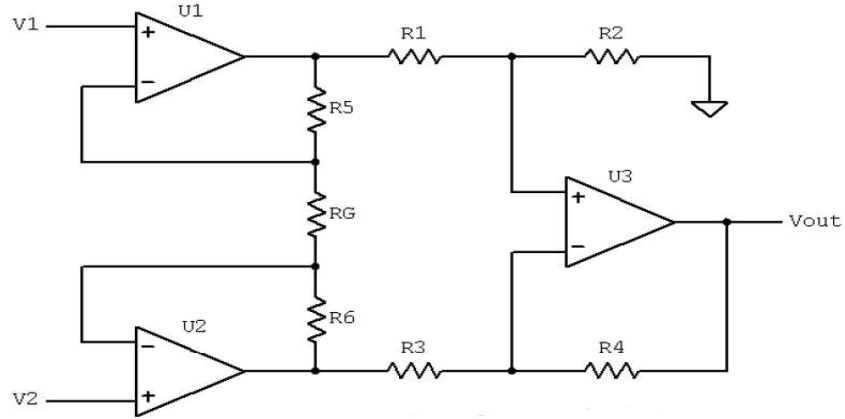A typical instrumentation amplifier consists of 7 resistors, as shown in the figure below:

Figure 7: Instrumentation Amplifier

Assuming resistors R1 to R6 are equal ($R1 = R2 = R3 = R4 = R5 = R6$), the output voltage can be calculated using the equation below:

$$V_{out} = (V_1 - V_2)\left(1 + \frac{2R}{R_G}\right)$$

In this case, the gain of the instrumentation amplifier hangs only on the value of R and R_G. The alternative design for an instrumentation amplifier used in this work had the following values for R and R_G.

$$R = 100\ K\Omega \quad R_G = 20\ K\Omega$$

$$=> Gain = \left(1 + \frac{2R}{R_G}\right) = \left(1 + \frac{2 * 100K}{20K}\right) = 11$$

The desired gain was chosen to be 11. This gain improves the signal-to-noise ratio of the EEG signal. It also prevents the noise in the signal from being over-amplified such that it offsets the EEG signal and causes it to be attenuated after the amplification.
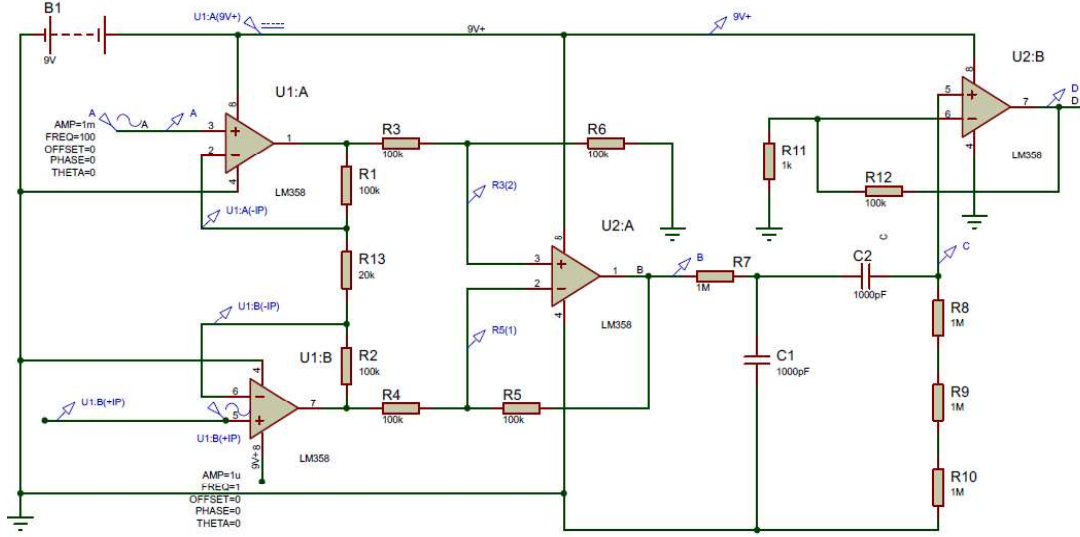
25

Figure 8: 1-channel EEG acquisition system with an instrumentation amplifier built with LM358 operational amplifiers.

Even though the gain of the amplifier in Figure 8 is very malleable, the amplifier design makes use of a lot more components. Consequently, the design has a more significant form factor compared to the AD620 design. As a result, AD620 was most preferred and was used in this work.

4.2.11 The signal amplification

The gains of the amplifiers were carefully chosen to prevent distortion of the signals. At each amplification stage, the output voltage was constrained below the saturation voltage of the amplifiers. An amplifier saturates when the input voltage exceeds the bounds of its power supply.

$$-V_s \leq G * V_{in} \leq +V_s$$

$$V_s = Supplied\ Voltage, \quad V_{in} = Input\ voltage$$

$$V_s = 9V$$

$$-9 \leq 11 * V_{in} \leq +9$$

$$-0.8181\ V \le V_{in} \le\ +0.8181\ V$$

This means that the input voltage at each amplification stage of the EEG device was constrained between -0.8181 V and +0.8181V.

4.2.12 Full Design of Circuit and 3D design of EEG headset

Figure 9 shows the circuitry for the EEG system. A PCB design was created from this schematic using Proteus 8.9. The PCB layout can be found in Appendix B.
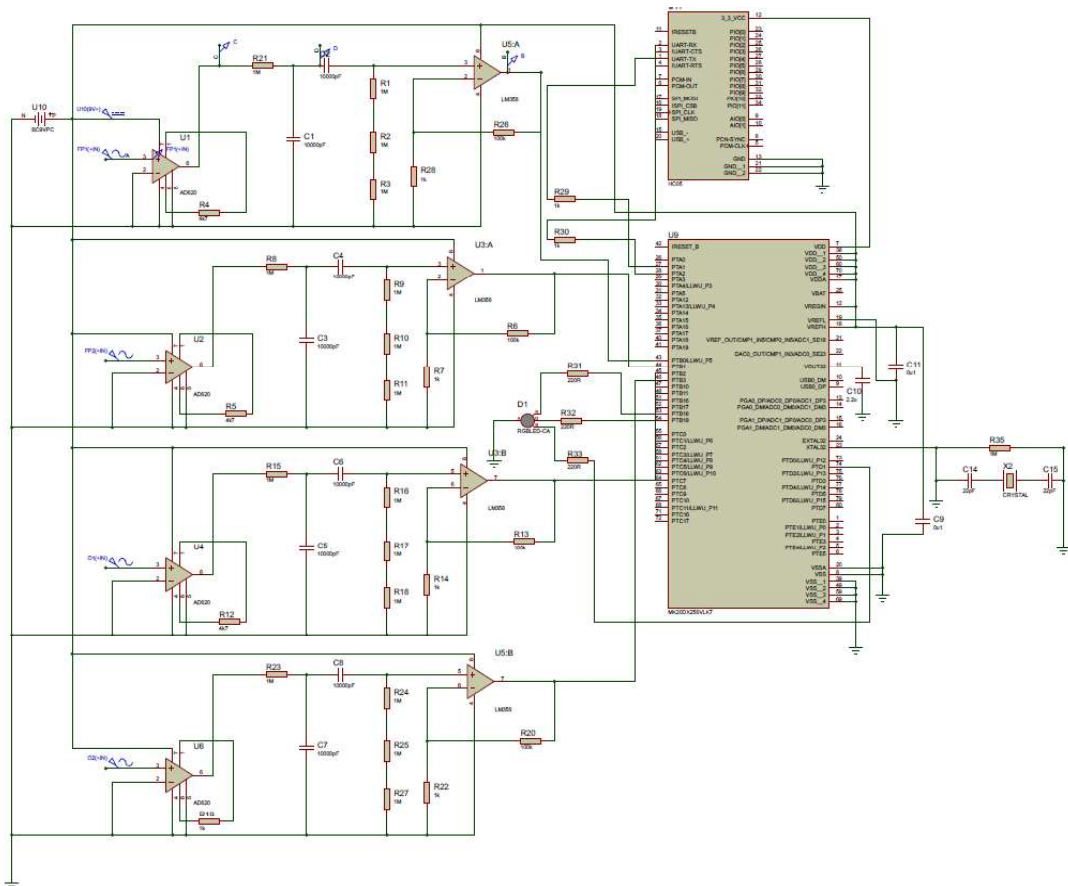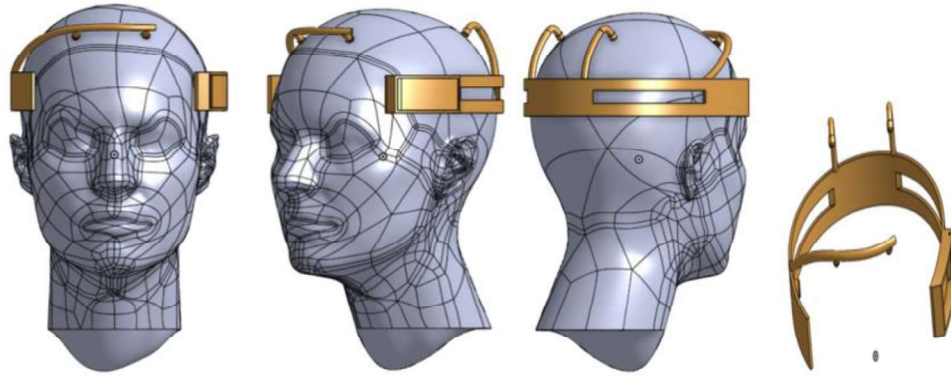

Figure 9: Full circuit for the EEG system

A 3D model was created for the entire system. Figure 10 shows the EEG headset and how it is placed on the head during recording. The parts for the assembly can be accessed here: https://bit.ly/3v0t4dE.

Figure 10: 3D model of EEG headset

4.3 Software Design and Implementation

This section contains the different features of the mobile application and how they were implemented. The pages of the mobile application can be seen in Appendix B. The software was built using flutter 1.22.6 stable version and dart version 2.10.5. Flutter is an open-source software development kit developed by Google. It can be used to create cross-platform applications from a single codebase.

4.3.1 Features of the application

The application has two main features: real-time visualization of the EEG signals and automated interpretation of the EEG signals. With these two features, a user can record EEG signals, view the signals in real-time during the EEG recording session, and get a quick diagnosis after the recording.

4.3.2 Implementation of the Automated Diagnosis

As described in section 4.3.1.2, this feature involves the use of artificial intelligence algorithms in interpreting recorded EEG signals. Also, it requires a large dataset in training a machine learning model to make predictions from the recorded data.

4.3.2.1 Data

The data used in training the model was obtained from Temple University Hospital's open-source EEG Corpus. The data repository consists of EEG signals in the European Data Format, each with about 24 to 36 channels of signal data. The signals were sampled at 250 Hz using 16 bits per sample. The dataset used contained 209 normal .edf files and 208 abnormal .edf files.

4.3.2.2 Data Pre-processing

Two data preprocessing methods were used. These methods resulted in two datasets: one was downsampled, and the other was not. These two methods have been discussed below.

4.3.2.2.1 Dataset A: No resampling

In this dataset, only the first 60 seconds of each signal was regarded in the model's training. This was done with the assumption that physicians can distinguish between normal and abnormal EEG signals using the samples recorded in the first few seconds. Since EEG signals had different recording durations, truncating the signal ensured that all the signals had uniform signal length or sample size.

15000 samples of each signal were obtained after taking the 60-second slice of each signal.

$$250\ Hz\ \times\ 60s\ =\ 15000\ samples$$

Each 15000-sample was then segmented into 15 groups using a 4-second window.

$$250\ Hz\ \times\ 4s = 1000\ samples$$

$$\frac{15000}{1000} = 15\ segments$$

29

Segmenting each signal helped in augmenting the dataset. After the data augmentation, the number of rows for normal EEG signals and abnormal EEG signals increased to 3135 and 3120, respectively.

The signals used in training the machine learning models were sampled at 250 Hz. However, the EEG device presented in this work samples signals at 40 Hz. One conflict between these two signals is the difference in the number of sample points recorded per second. One records 250 samples per second while the other records 40. Hence, segmenting each signal also helped to reduce the minimum recording time for the EEG system presented in this work.

$$40 \ Hz: 40 \ samples \ per \ second$$

$$60 \ second \ slice \ \equiv 15000 \ samples$$

$$If \ 15000 \ samples \ needed \ then:$$

$$\frac{15000}{40} = 375 \ seconds$$

Assuming the whole 60-second slice of the signals were used in training the machine learning models, the EEG system presented in this work would have to record the signals for at least 375 seconds. However, due to the 4-second window used, only 1000 samples were used in each training and testing iteration of the machine learning models. This reduced the minimum recording time from 375 seconds to 250 seconds ($\frac{1000}{40} = 250s$).

One notable merit of this method is the retention of the high-frequency components in the signals. Since the signals are not resampled, most of the features are preserved. Even though this is another conflict between the signals, the preserved features can help the

30

machine learning models make accurate predictions. However, this method also has some downsides. The cons include:

i. the preservation of noise in the signals, which can eventually lead to poor performance of the machine learning models.

ii. a mismatch in frequency components in the signals used in training the models (sampling frequency=250Hz) and the ones used in real-life testing with the EEG device (sampling frequency=40Hz).

4.3.2.2.2 Dataset B: Downsampled signals

The original sampling frequency of the signals is 250 Hz. The signals were downsampled at 40 Hz to make the frequency components uniform in both the signals used in training the model and the ones recorded with the EEG device,

After downsampling, each signal was segmented into multiple parts using a 4-minute window. This data segmentation process served the purpose of data augmentation and created a form of homogeneity (with respect to the length of the signals) in the dataset.

Downsampling the signal led to the elimination of some noise. Nevertheless, some features were lost during the preprocessing, which would have been relevant in training the machine learning models.

4.3.2.3 Feature Extraction

The same feature extraction techniques were used on both dataset A and dataset B. Fast Fourier transform (FFT) and discrete wavelet transform (DWT) were computed for each EEG signal slice in both datasets. Fast Fourier transform is an efficient algorithm used in calculating the Discrete Fourier Transform of a sequence. It is a frequency domain

transform that provides information about the frequencies present in a signal. Even though it has a high resolution in the frequency domain, it has zero resolution in the time domain. Consequently, Fast Fourier Transforms do not retain the time-related details of the signal and fail to specify the position in time where each frequency occurred in a signal. The Fast Fourier Transform $Y[k]$ of a sequence, $x[n]$ of length N is defined as:

$$Y[k] = \sum_{n=0}^{N-1} (e^{-2\pi j \frac{kn}{N}} x[n])$$

Unlike FFTs, discrete wavelet transforms store both time and frequency information contained in a signal. DWT is a time-frequency domain transform that returns approximation and detail coefficients. The approximation coefficients are outputs of a low pass filter, while a high pass filter produces the detail coefficients. An n-level DWT splits a signal into $n + 1$ frequency sub-bands. In this work, 1-level and 4-level Daubechies DWT decompositions were used.

Statistical features (mean, standard deviation, mean square, and absolute difference) were computed for coefficients obtained from the DWT and the FFT output. All complex output from the FFT were converted to real values by taking the absolute of the complex values. After the feature extraction process, 16 features were extracted from the FFT output, and 32 features were extracted for the DWT output. Figure 11 shows the feature extraction process.
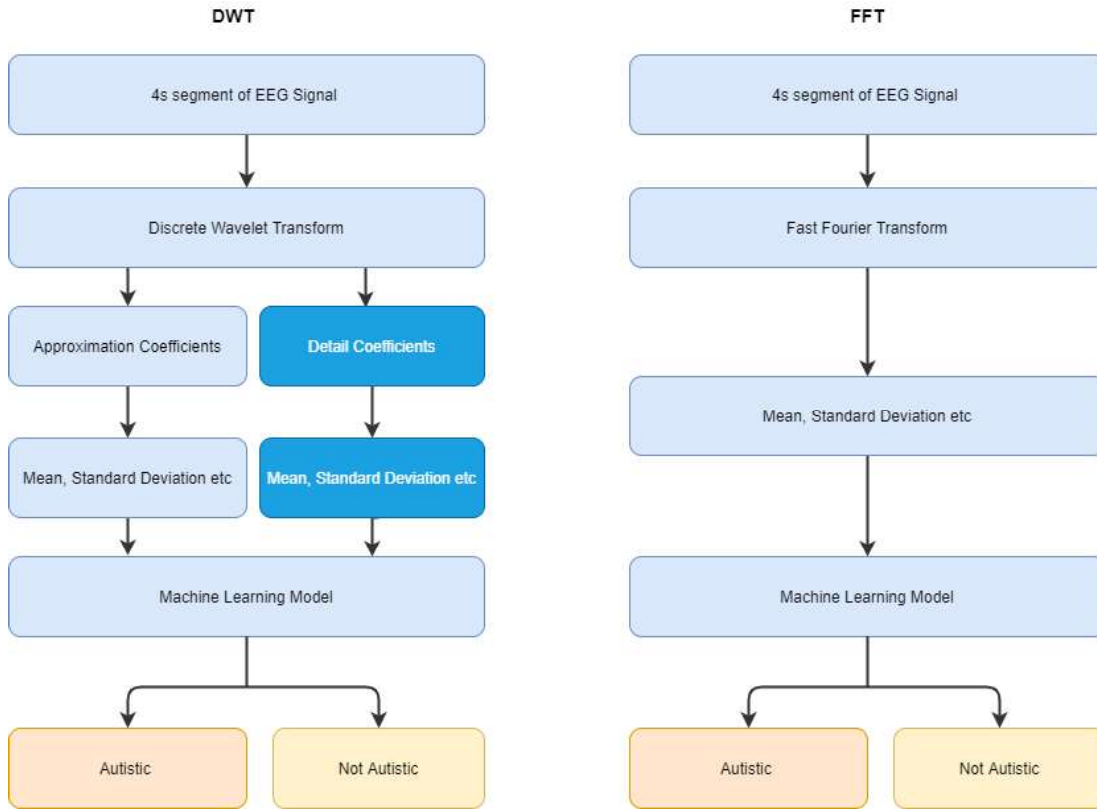
Figure 11: Feature Extraction process

4.3.2.4 Machine Learning Models

The models were trained on features from both datasets A and B. The features extracted were used in training non-parametric machine learning models such as K-nearest neighbor (KNN), decision trees, random forest (RF), and support vector machine (SVM) classifiers. These machine learning algorithms have high-performance capabilities and can fit many functional forms since they make no assumptions about the mapping function.

KNN is a distance-based supervised learning algorithm. It does not do any active learning. Instead, it memorizes the training data and classifies them using similarities between the data points. When an input vector $X = [X_0 \; X_1 \; X_2]$ is passed to a KNN

algorithm, the Euclidean distance between the input vector X and other points in the dataset is calculated.

Table 4: Sample Data

	Feature A	Feature B	Feature C	Label
EEG signal 1	xA_0	xB_0	xC_0	Normal
EEG signal 2	xA_1	xB_1	xC_1	Abnormal

Using the sample data presented in table 4, the Euclidean distance for EEG signal 1 is calculated as follows:

$$Euclidean\ distance = \sqrt{(X_0 - xA_0)^2 + (X_1 - xB_0)^2 + (X_2 - xC_0)^2}$$

After calculating the Euclidean distance for all the signals in the dataset, the distances are sorted in increasing order. The algorithm then returns the majority class in the first k items with the lowest distances to input vector X. k is the number of neighbors used in the algorithm. The KNN algorithm is straightforward and easy to implement. However, it is computationally expensive as it uses a substantial amount of memory to store training data when making predictions. Since these predictions occur on the server, it might take a few more minutes for a user to get a diagnosis after recording the signals.

SVM is also a supervised learning algorithm that classifies data by finding the best hyperplane that differentiates classes of data points plotted in n-dimensional space ($n = number\ of\ features$). The best hyperplane is the one with a maximum margin with the closest points (support vectors). A hyperplane is a straight line that can be represented by the equation below:

$$H: w^T x_n + b = 0$$

34

An SVM model with a radial basis function (RBF) kernel was used. The RBF kernel makes the data more separable by converting it from a lower dimension to a higher dimension.

Just like KNN and SVM, a decision tree classifier is also a supervised learning algorithm. It is a tree-structure classifier. It represents the features of a dataset using its internal nodes, the decision rules as its branches, and the outcome as each leaf node. A decision tree starts from the root node, which asks a simple question. Depending on a yes or no answer to the question, the tree splits into subtrees. This algorithm is easy to understand as it mimics the thinking process of a human being. However, it often has an overfitting issue.

Random forest combines several decision trees to create an ensemble that has better accuracy in classifying the data. It solves the overfitting issues in decision trees since it chooses features randomly during the training process. As a result, the random forest algorithm performs better than the decision tree classifier algorithm.

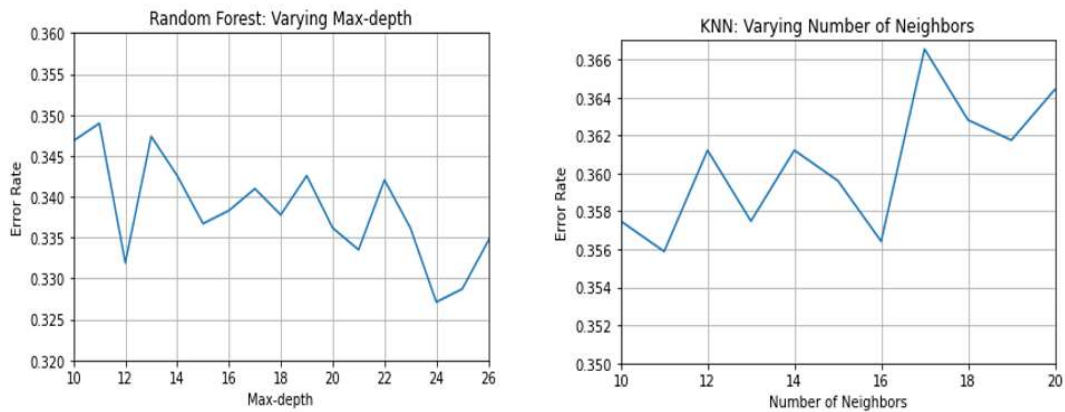## 4.3.2.4.1 Parameters used for the machine learning models trained on dataset A

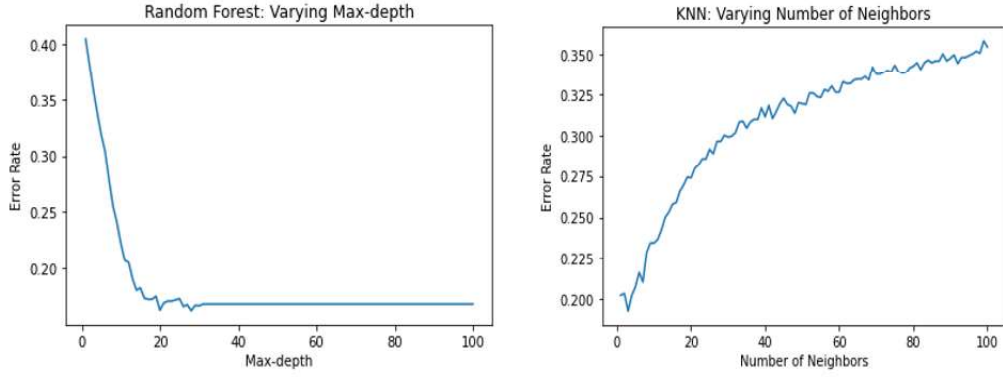*Figure 12: Parameters for Random Forest and KNN with FFT features*

Figure 13: Parameters for Random Forest and KNN with DWT features

The depth for the random forest model and the number of neighbors for the KNN model were varied to arrive at the correct parameters that give the highest accuracy. The parameters with their corresponding error rates can be seen in Figures 12 and 13. For the KNN model trained on FFT features and the KNN model trained on DWT features, the number of neighbors was determined to be 11 and 3, respectively. Also, a maximum depth of 24 resulted in the least error rate for the random forest model trained on FFT features, while that of the one with DWT features was 28.

4.3.2.4.2 Parameters used for the machine learning models trained on dataset B
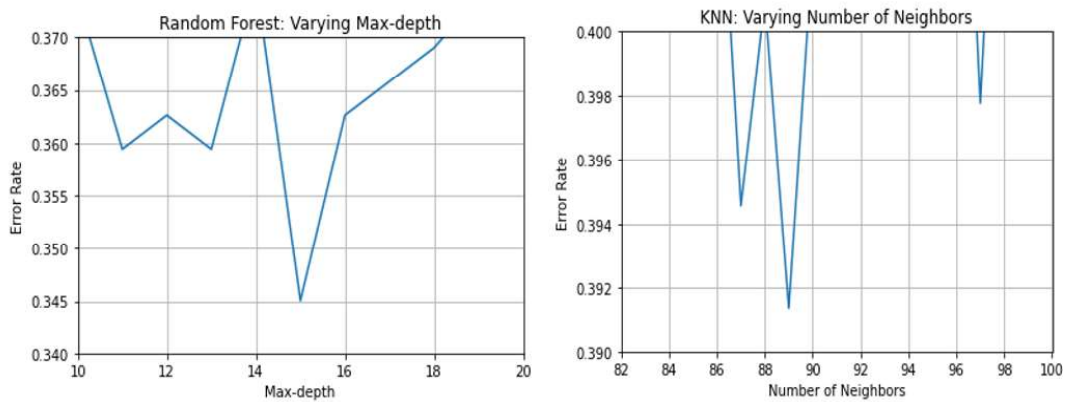

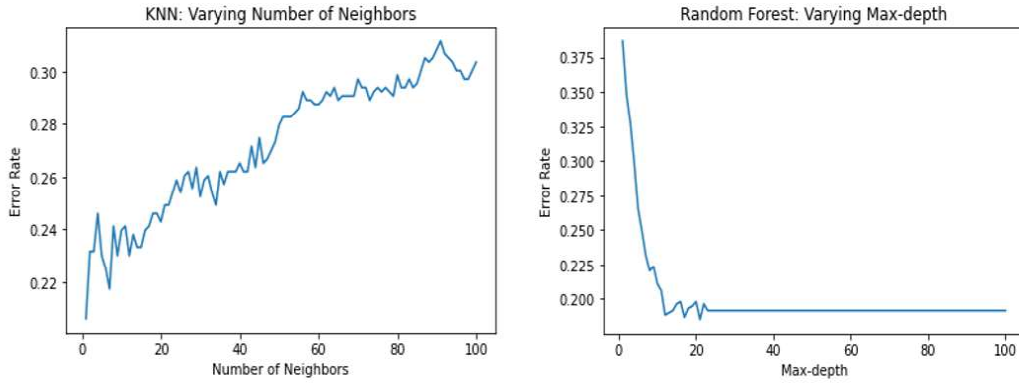*Figure 14: Parameters for Random Forest and KNN with FFT features*

Figure 15: Parameters for Random Forest and KNN with DWT features

From Figure 14, the KNN model trained on FFT features extracted from dataset B had the least error rate with 89 neighbors. The random forest model converged at a minimum error rate with a max-depth of 15. Also, from Figure 15, the KNN model and the random forest model trained on DWT features extracted from dataset B attained a minimum error rate with 1 neighbor and a max-depth of 21, respectively.

4.3.2.4.3 Deploying the machine learning model.

The machine learning model with the highest accuracy was deployed using a flask server. A post request from the mobile application after recording the EEG signals is submitted to the Flask API. The API responds with the diagnosis for the signals after passing them through the model deployed on the server.

Chapter 5: Testing and Results

This chapter delineates how the various features of the proposed solution were tested. The results of the testing are expected to meet the design requirements for the system.

5.1 Testing Real-time visualization of the EEG signals feature

FRDM-KL25z

HC-06 Bluetooth Module

Gold-plated EEG electrodes
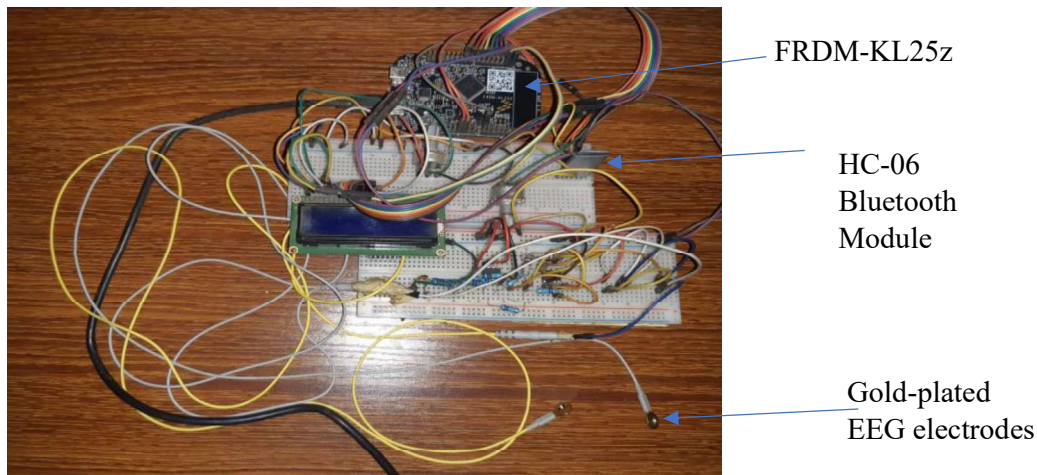
Figure 16: 1-Channel EEG circuit for testing Real-time visualization of the signals.

In testing this feature, a one-channel EEG circuit (Figure 16) was used in generating an EEG signal to be passed on to the mobile application. The one-channel circuit was connected to a FRDM-KL25z and HC-06 Bluetooth module, as shown in Figure 16. The user pairs with the HC-06 device using a smartphone. After pairing with the device, the user opens the application and connects to the HC-06 device (Figure 17: a, b). Afterward, the user clicks on the start button on the mobile application to start recording (Figure 17c). The figure below shows the user experience on the mobile application.
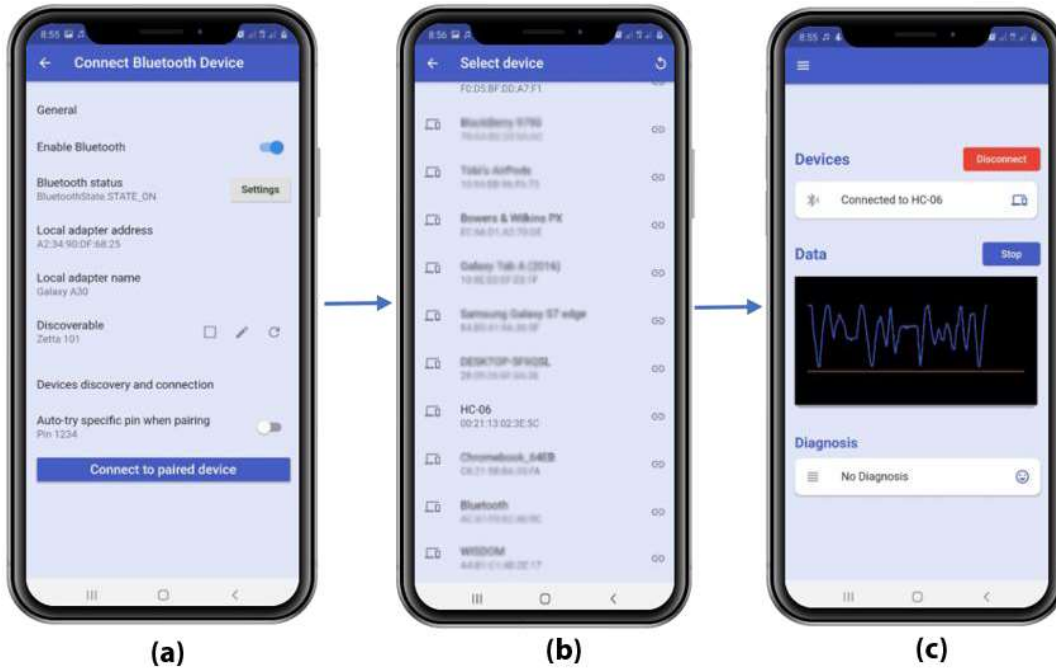
Figure 17: Real-time Visualization of the EEG signals

The primary metric used in evaluating this feature was the response time of the microprocessor in performing the ADC and the response time of the mobile application in updating the values of the signal. The mobile application generally had a fast response time. It was observed that the conversion, transmission, and representation of the signals on the mobile application were done with no latency.

5.2 Testing automated diagnosis

The traditional machine learning models were evaluated using a 10-fold cross-validation technique. The dataset was split into 10 folds, and each fold was used as testing data at some point in the cross-validation (shown in Figure 18). For each iteration of the cross-validation, the accuracy, precision, and recall scores of the models on the test data were assessed.
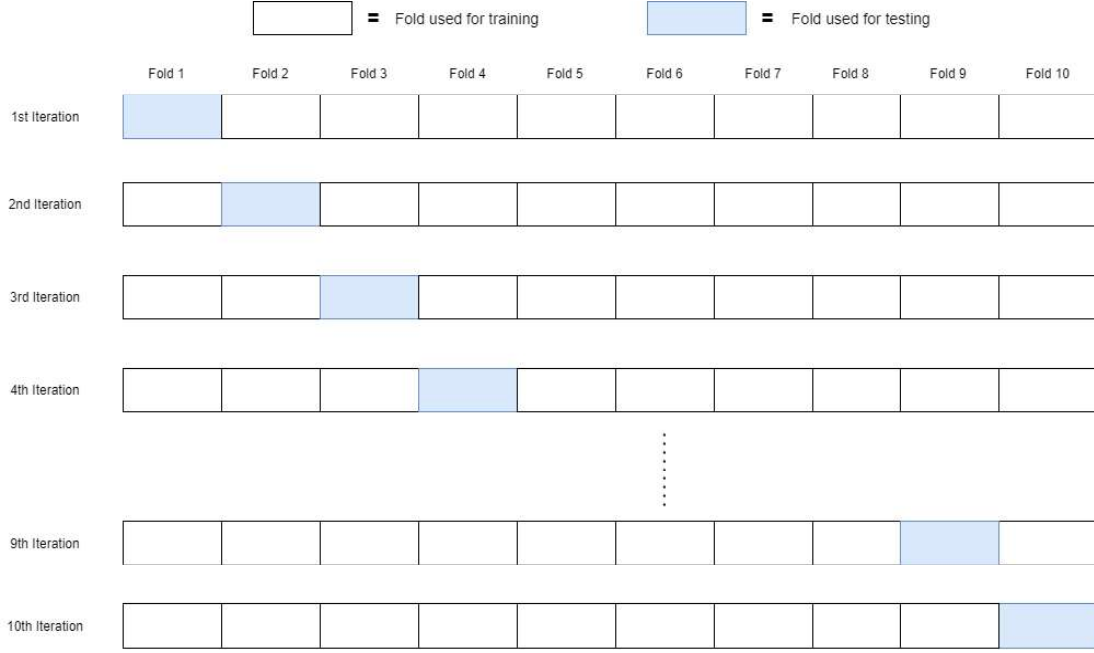
Figure 18: 10-fold Cross-Validation

The accuracy score measures the percentage of classes that were correctly predicted. Precision measures the normal cases that were correctly identified from all the predicted normal cases.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall measures the number of normal EEG signals that were correctly classified from the actual normal EEG signals.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

As shown in Figure 19 and Figure 20, the models performed better on the statistical features extracted from the discrete wavelet transforms of the signals. This can be attributed to the full resolution of discrete wavelet transforms in both time and frequency. A random forest

40

model with 28 trees trained on dataset A achieved the highest accuracy of 84 percent. Since an accuracy of at least 85 percent is required for the EEG system, more training was done to improve the accuracy score.

The models were trained on statistical features extracted from a 4-level Daubechies DWT decomposition of the EEG signal. This improved the accuracy of most of the models. The SVM model performed poorly with dataset A because of noise (overlapping classes) in the dataset. Due to reduced noise in dataset B, the SVM classifier performed better when trained on features extracted from that dataset. Unlike the other models used, the random forest model uses a group of decision tree models to make a prediction. This is why the random forest model had the highest accuracy with both FFT and DWT from the datasets. The random forest model trained on dataset A had the highest accuracy of 85.1 percent. Additionally, unlike the downsampled signals used in dataset B, dataset A has more features to differentiate normal from abnormal signals. Hence, a more accurate result from the random forest model trained on dataset A.

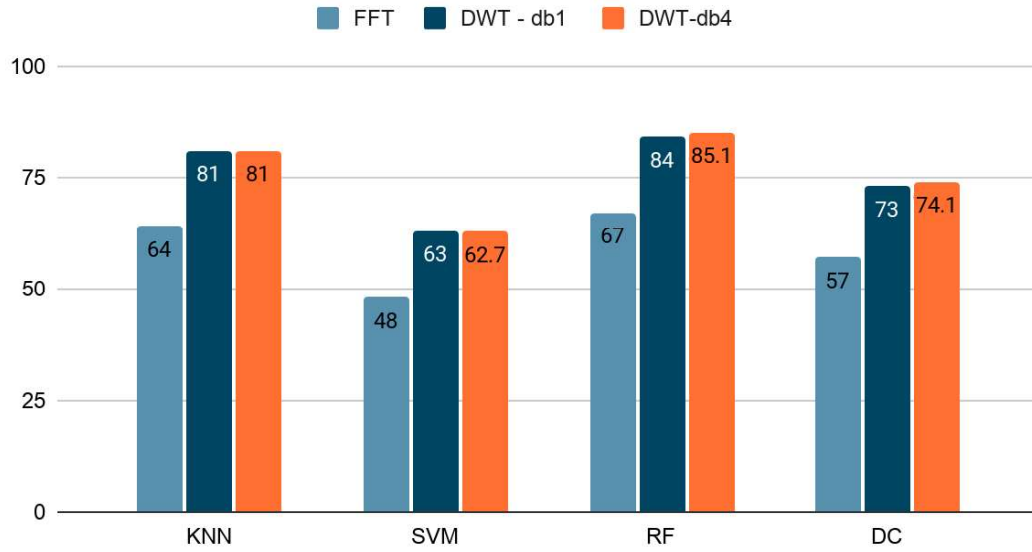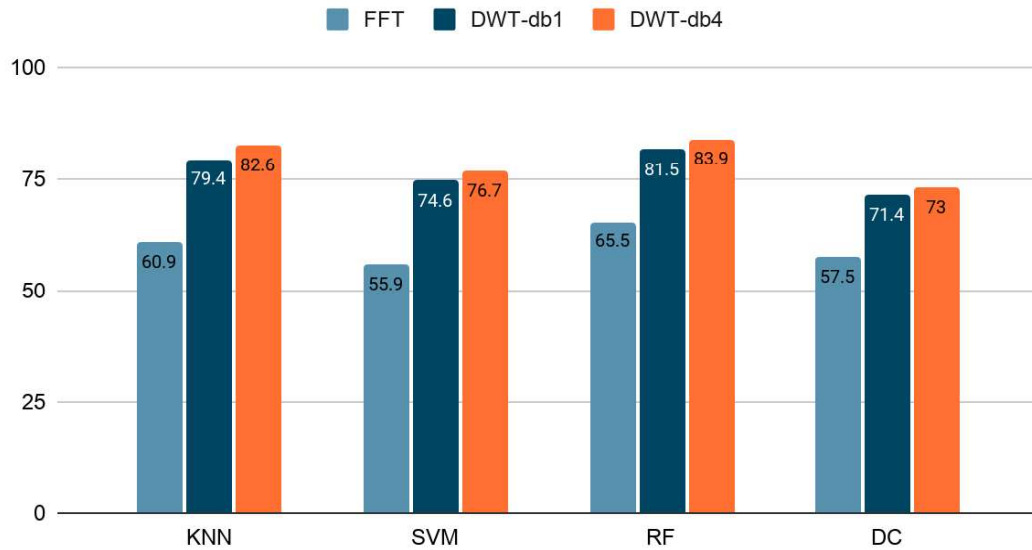*Figure 19: Accuracies of machine learning models trained using Dataset A*

Figure 20: Accuracies of machine learning models trained using Dataset B.

The other evaluation metrics for the different models have been summarised in the table below.

	KNN		Random Forest		SVM		Decision Tree	
Dataset	A	B	A	B	A	B	A	B
Accuracy	81.0	82.6	85.1	83.9	62.7	76.7	74.1	73.0
Precision	79.6	80.9	84.5	82.4	60.4	77.0	73.3	71.6
Recall	81.7	84.8	84.8	85.8	66.7	75.5	73.3	75.4

The accuracy, recall, and precision score for the random forest model trained on DWT-db4 features extracted from dataset A is approximately 85%. This means that for both normal and abnormal EEG cases, the model produces an accurate diagnosis about 85% of the time.

Figure 21: Diagnosis produced by mobile application.

Figure 21 shows a diagnosis produced by the machine learning model after recording EEG signals and querying the model.

Chapter 6: Conclusion

The problem of late detection of autism and its repercussions were laid out in the first chapter. The other chapters outlined a solution using a portable and low-cost EEG device with an automated diagnosis for early diagnosis of autism. The resulting solution developed in this work can diagnose autism using a random forest model with an accuracy of 85.1% using already existing EEG data. The device also allows users to see the signals on their smart device in real-time.

6.1 Limitations

During the design and implementation phase of the proposed EEG system, some challenges influenced some of the design choices. Some of these challenges have been listed below:

- The dataset used for training the machine learning models was one of the major limitations in this work. The dataset used was not labeled as "autistic" or "not autistic." Instead, it was labeled as "normal" and "abnormal." This implied that any classifier trained on this data would be capable of differentiating abnormal EEG signals (signals with several disorders including ictal, interictal, and autism conditions) from normal ones. As such, the output from the automated diagnosis on the mobile application only shows whether a person's EEG is abnormal or normal (not specifically autistic or not, as defined in the scope of the proposed solution).

- Due to the unavailability of all the components to construct the 4-channel EEG device proposed in this work, the automated diagnosis feature was not tested with real-life signals recorded using the EEG device.

6.2 Future Work

The design which has been proposed in this work is feasible. However, it can be enhanced in several ways. Some of the recommendations for improving the proposed solution have been enumerated below.

- A better dataset that is labeled as autistic or not autistic can be used. This will make the diagnosis given after every EEG signal recording less ambiguous.

- Time series classification using artificial neural networks can be used in the automated diagnosis. This would improve the accuracy of the model used in predicting the diagnosis.

References

[1] M. Ahmadlou, H. Adeli, and A. Adeli, "Fractality and a Wavelet-Chaos-Neural Network Methodology for EEG-Based Diagnosis of Autistic Spectrum Disorder," Journal of Clinical Neurophysiology, vol. 27, no. 5, pp. 328–333, Oct. 2010, doi: 10.1097/wnp.0b013e3181f40dc8.

[2] M. Ahmadlou, H. Adeli, and A. Adeli, "Fuzzy Synchronization Likelihood-wavelet methodology for the diagnosis of autism spectrum disorder," Journal of Neuroscience Methods, vol. 211, no. 2, pp. 203–209, Nov. 2012, doi: 10.1016/j.jneumeth.2012.08.020.

[3] M. Ahmadlou, H. Adeli, and A. Adeli, "Improved visibility graph fractality with application for the diagnosis of Autism Spectrum Disorder," Physica A: Statistical Mechanics and its Applications, vol. 391, no. 20, pp. 4720–4726, Oct. 2012, doi: 10.1016/j.physa.2012.04.025.

[4] M. Bakare and K. Munir, "Autism spectrum disorders (ASD) in Africa: a perspective," Afr. J. Psych, vol. 14, no. 3, Sep. 2011, doi: 10.4314/ajpsy.v14i3.3.

[5] M. A. Bello-Mojeed, M. O. Bakare, and K. Munir, "Identification of Autism Spectrum Disorders (ASD) in Africa: Need for Shifting Research and Public Health Focus," in Comprehensive Guide to Autism, Springer New York, 2014, pp. 2437–2453.

[6] K. Belwafi, R. Djemal, F. Ghaffari, and O. Romain, "An adaptive EEG filtering approach to maximize the classification accuracy in motor imagery," presented at the 2014 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), Dec. 2014, doi: 10.1109/ccmb.2014.7020704.

[7] W. Bosl, A. Tierney, H. Tager-Flusberg, and C. Nelson, "EEG complexity as a biomarker for autism spectrum disorder risk," BMC Med, vol. 9, no. 1, Feb. 2011, doi: 10.1186/1741-7015-9-18.

[8] W. J. Bosl, H. Tager-Flusberg, and C. A. Nelson, "EEG Analytics for Early Detection of Autism Spectrum Disorder: A data-driven approach," Sci Rep, vol. 8, no. 1, May 2018, doi: 10.1038/s41598-018-24318-x

[9] R. Coben, A. R. Clarke, W. Hudspeth, and R. J. Barry, "EEG power and coherence in autistic spectrum disorder," Clinical Neurophysiology, vol. 119, no. 5, pp. 1002–1009, May 2008, doi: 10.1016/j.clinph.2008.01.013

[10] J. A. Cruz and D. S. Wishart, "Applications of Machine Learning in Cancer Prediction and Prognosis," Cancer Inform, vol. 2, p. 117693510600200, Jan. 2006, doi: 10.1177/117693510600200030.

[11] R. Djemal, K. AlSharabi, S. Ibrahim, and A. Alsuwailem, "EEG-Based Computer-Aided Diagnosis of Autism Spectrum Disorder Using Wavelet, Entropy, and ANN," BioMed Research International, vol. 2017, pp. 1–9, 2017, doi: 10.1155/2017/9816591

[12] A. Dzedzickis, A. Kaklauskas, and V. Bucinskas, "Human Emotion Recognition: Review of Sensors and Methods," Sensors, vol. 20, no. 3, p. 592, Jan. 2020, doi: 10.3390/s20030592.

[13] N. Fauzan and N. H. Amran, "Brain Waves and Connectivity of Autism Spectrum Disorders," Procedia - Social and Behavioral Sciences, vol. 171, pp. 882–890, Jan. 2015, doi: 10.1016/j.sbspro.2015.01.204.

[14] W. O. A. S. Wan Ismail, M. Hanif, S. B. Mohamed, N. Hamzah, and Z. I. Rizman, "Human Emotion Detection via Brain Waves Study by Using Electroencephalogram (EEG)," International Journal on Advanced Science, Engineering and Information Technology, vol. 6, no. 6, p. 1005, Dec. 2016, doi: 10.18517/ijaseit.6.6.1072

[15] N. S. Khan, M. H. Muaz, A. Kabir, and M. N. Islam, "A Machine Learning-Based Intelligent System for Predicting Diabetes," International Journal of Big Data and

Analytics in Healthcare, vol. 4, no. 2, pp. 1–20, Jul. 2019, doi: 10.4018/ijbdah.2019070101.

[16] L. M. Oberman, E. M. Hubbard, J. P. McCleery, E. L. Altschuler, V. S. Ramachandran, and J. A. Pineda, "EEG evidence for mirror neuron dysfunction in autism spectrum disorders," Cognitive Brain Research, vol. 24, no. 2, pp. 190–198, Jul. 2005, doi: 10.1016/j.cogbrainres.2005.01.014.

[17] S. Phetrasuwan, M. S. Miles, G. B. Mesibov, and C. Robinson, "Defining Autism Spectrum Disorders," Journal for Specialists in Pediatric Nursing, vol. 14, no. 3, pp. 206–209, Jul. 2009, doi: 10.1111/j.1744-6155.2009.00200.x.

[18] K. Ruparelia et al., "Autism Spectrum Disorders in Africa: Current Challenges in Identification, Assessment, and Treatment," J Child Neurol, vol. 31, no. 8, pp. 1018–1026, Mar. 2016, doi: 10.1177/0883073816635748.

[19] M. G. Bleichner et al., "Exploring miniaturized EEG electrodes for brain-computer interfaces. An EEG you do not see?," Physiol Rep, vol. 3, no. 4, p. e12362, Apr. 2015, doi: 10.14814/phy2.12362.

[20] J. Poveda, M. O'Sullivan, E. Popovici, and A. Temko, "Portable neonatal EEG monitoring and sonification on an Android device," presented at the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2017, doi: 10.1109/embc.2017.8037248.

[21] T. J. Sullivan, S. R. Deiss, G. Cauwenberghs, and T.-P. Jung, "A low-noise low-power EEG acquisition node for scalable brain-machine interfaces," in Bioengineered and Bioinspired Systems III, May 2007, doi: 10.1117/12.724019.

[22] S. Iranmanesh and E. Rodriguez-Villegas, "A 950 nW Analog-Based Data Reduction Chip for Wearable EEG Systems in Epilepsy," IEEE J. Solid-State Circuits, vol. 52, no. 9, pp. 2362–2373, Sep. 2017, doi: 10.1109/jssc.2017.2720636.

[23] K. E. Mathewson, T. J. L. Harrison, and S. A. D. Kizuk, "High and dry? Comparing active dry EEG electrodes to active and passive wet electrodes," Psychophysiol, vol. 54, no. 1, pp. 74–82, Dec. 2016, doi: 10.1111/psyp.12536.

[24] R. Puyol, G. Lenzi, G. Barg and A. Arnaud, "A portable, high-density EEG acquisition system," 2013 7th Argentine School of Micro-Nanoelectronics, Technology and Applications, Buenos Aires, 2013, pp. 32-37.

[25] J. Lee, S. Kim and Y. Lee, "A potable platform to analyze EEG signal for transmitting to a mobile device via Wi-Fi," 2015 International Symposium on Consumer Electronics (ISCE), Madrid, 2015, pp. 1-2, doi: 10.1109/ISCE.2015.7177814.

[26] "N. Dey, A. Ashour, S. J. Fong, and C. Bhatt, Eds., Wearable and Implantable Medical Devices, 4th ed. Elsevier, 2020.

[27] G. Li, S. Wang, and Y. Y. Duan, "Towards conductive-gel-free electrodes: Understanding the wet electrode, semi-dry electrode and dry electrode-skin interface impedance using electrochemical impedance spectroscopy fitting," Sensors and Actuators B: Chemical, vol. 277, pp. 250–260, Dec. 2018, doi: 10.1016/j.snb.2018.08.155.

[28] J. Rosell, J. Colominas, P. Riu, R. Pallas-Areny, and J. G. Webster, "Skin impedance from 1 Hz to 1 MHz," IEEE Trans. Biomed. Eng., vol. 35, no. 8, pp. 649–651, 1988, doi: 10.1109/10.4599.

[29] A. C. Metting van Rijn, A. Peper, and C. A. Grimbergen, "High-quality recording of bioelectric events," Medical & Biological Engineering & Computing, vol. 29, no. 4, pp. 433–440, Jul. 1991, doi: 10.1007/bf02441666

[30] M. O'Neill and T. Shear, "EEG for Diagnosis of Autism Spectrum Disorder," Pediatr Neurol Briefs, vol. 32, p. 13, doi: 10.15844/pedneurbriefs-32-13.

[31] S. Smith, "EEG in the diagnosis, classification, and management of patients with epilepsy," *J Neurol Neurosurg Psychiatry*, vol. 76, no. Suppl 2, pp. ii2–ii7, Jun. 2005, doi: 10.1136/jnnp.2005.069245.

[32] A. Sheikhani, H. Behnam, M. R. Mohammadi, M. Noroozian, and M. Mohammadi, "Detection of abnormalities for diagnosing of children with autism disorders using of quantitative electroencephalography analysis," *J Med Syst*, vol. 36, no. 2, pp. 957–963, Apr. 2012, doi: 10.1007/s10916-010-9560-6.

[33] H. Azuma, S. Hori, M. Nakanishi, S. Fujimoto, N. Ichikawa, and T. A. Furukawa, "An intervention to improve the interrater reliability of clinical EEG interpretations," *Psychiatry and Clinical Neurosciences*, vol. 57, no. 5, pp. 485–489, 2003, doi: https://doi.org/10.1046/j.1440-1819.2003.01152.x.

[34] U. F. O. Themes, "Electroencephalographic Electrodes, Channels, and Montages and How They Are Chosen," *Neupsy Key*, Mar. 12, 2017. https://neupsykey.com/electroencephalographic-electrodes-channels-and-montages-and-how-they-are-chosen/ (accessed Mar. 15, 2021).

[35] Mohibullah Kakar, "10 and 20 electrode placement," 12:04:07 UTC, Accessed: Mar. 15, 2021. [Online]. Available: https://www.slideshare.net/mohibullahfazli/10-and-20-electrode-placement.

[36] R. M. Jones and C. Lord, "Diagnosing autism in neurobiological research studies," *Behavioural Brain Research*, vol. 251, pp. 113–124, Aug. 2013, doi: 10.1016/j.bbr.2012.10.037

Appendix

A. PCB Layout
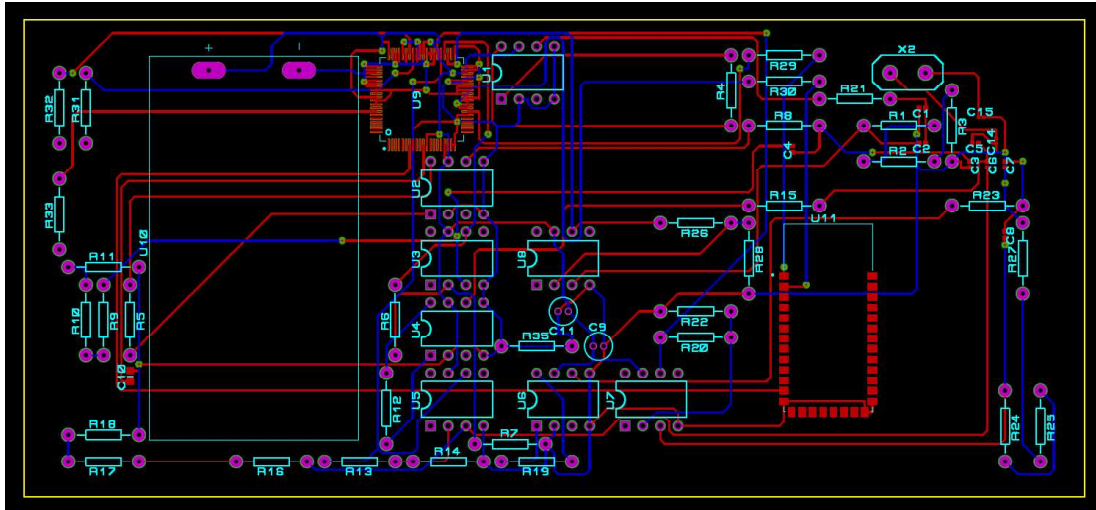
B. Mobile Application pages

Sign Up