# A Multimodel Approach for Word Familiarity Prediction

## A Dissertation

*submitted in*
*partial fulfillment of the requirements*
*for the award of degree of*

***Master of Technology***

*in*

***Computer Science & Engineering***

By:
**Vaishali Khurana**
**M.Tech CSE (16535043)**

Under the guidance of
**Dr. Partha Pratim Roy**
**Assistant Professor, CSE Department**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**
**INDIAN INSTITUTE OF TECHNOLOGY ROORKEE**
**ROORKEE,247667**

May 2018

# Candidate's Declaration

I hereby declare that the work presented in this dissertation entitled **"A Multimodel Approach for Word Familiarity Prediction"** submitted in the partial fulfillment of the requirements for the award of the Degree of **Master of Technology** in **Computer Science & Engineering** is an authentic record of my own work, carried out during the period from May 2017 to May 2018 under the guidance of **Dr. Partha Pratim Roy**, Assistant Professor, Department of Computer Science & Engineering, Indian Institute of Technology Roorkee, India. The results embodied in this report have not been submitted by me for the award of any other degree of this or any other Institute/University.

Date: .................
Place: Roorkee

**Vaishali Khurana**
M.Tech. CSE (16535043)
IIT Roorkee

---

# Certificate

This is to certify that the statement made by the candidate in the declaration is correct to the best of my knowledge and belief.

Date: .................
Place: Roorkee

**Dr. Partha Pratim Roy**
Assistant Professor
CSE Department
IIT Roorkee

# Acknowledgements

It gives me a great sense of pleasure to present this report on “A Multimodel Approach for Word Familiarity Prediction”. I have received enormous help, guidance, and advice from many people and I feel that it will not be right to mention a line about at least some of them. Firstly, I would like to express my utmost gratitude to my supervisor **Dr. Partha Pratim Roy**, Indian Institute of Technology Roorkee for his constant support and regular guidance throughout the course of my work. His veracity, thoroughness, and persistence have been a constant source of encouragement for me. I would like to express my sincere appreciation and gratitude towards all the lab colleagues, specially **Mr. Rajkumar Saini** and **Mr. Pradeep Kumar** for their consistent support and invaluable suggestions. I would also like to thank all those people who helped me record the dataset used for this research work.

I am also grateful to the Department of Computer Science and Engineering, IIT Roorkee for providing valuable resources to aid my research. A hearty thanks to my friends and family for encouraging me in good and bad times.

***Vaishali Khurana***

# Abstract

The appearance of unknown words often disturbs communication and reading. The proposed system focuses on detecting those words which are unfamiliar to the users using temporal data, Electroencephalography (EEG) and facial expressions of users. In particular, for the word where the user gazes for some time, a word-familiarity prediction approach based on time duration for which user has focused on that word, EEG signals from the user's brain waves and facial expressions of the user while reading that word, has been developed. Word-familiarity refers whether a user is familiar with the word or not while reading the text. The proposed system keeps the track of the coordinates of the gaze with the timestamp to find the duration of the fixation of the gaze at the particular word. Further, this time duration data has been fed to Stochastic Gradient Descent classifier to predict the word familiarity. Similarly, EEG signals have been processed using Wavelet decomposition technique and four features have been computed from beta and gamma frequency bands. The prediction of word-familiarity has been performed using Random Forest classifier. A decision fusion approach has also been used to boost the prediction performance. The results show that the characteristics of brain waves at the time of unknown word perception or confusion can be detected. And further facial expressions of users have been used for prediction. The video has been recorded while the user is reading the text. Image frames have been extracted from that video and from each of that frame, a total of 68 cartesian coordinate point dataset have been generated. The sequential dataset has been generated by finding the difference between the coordinate points with adjacent frame. And then word familiarity has been predicted by LSTM classifier and further results have been compared with HMM classifier. A dictionary based pop-up window has been developed to provide the meaning of the word when a user is found to be unfamiliar with the text. The dataset of 12-15 users for different models has been developed while they are reading 25 words. An accuracy of 82% has been recorded with EEG dataset using the proposed classifier combination approach, 72.9% with temporal analysis and 80.26% with facial expression dataset using LSTM classifier. Finally, a comparative study with other popular classification technique is also discussed.

# Contents

**Candidate's Declaration** **i**

**Certificate** **i**

**Acknowledgements** **ii**

**Abstract** **iii**

**Contents** **iv**

**List of Figures** **vii**

**List of Tables** **ix**

**1 Introduction** **1**
- 1.1 Problem Statement . . . . . 4
- 1.2 Research Gap and Our Contribution . . . . . 5

**2 Literature Review** **7**
- 2.1 Eye Tracking . . . . . 7
- 2.2 Electroencephalography (EEG) . . . . . 8
- 2.3 Facial Expression Recognition . . . . . 10

**3 Methodology** **12**
- 3.1 Electroencephalography (EEG) Signal Analysis for Word Familiarity Prediction 12
  - 3.1.1 System Setup and Signal Acquisition . . . . . 12
  - 3.1.2 Signal Pre-processing . . . . . 14
    - 3.1.2.1 Data Filtering & Smoothing . . . . . 14
    - 3.1.2.2 Independent Component Analysis (ICA) . . . . . 15
  - 3.1.3 Features Extraction . . . . . 16
    - 3.1.3.1 Discrete Wavelet Transform (DWT) . . . . . 16
    - 3.1.3.2 Mean . . . . . 18
    - 3.1.3.3 Standard deviation . . . . . 19
    - 3.1.3.4 Root Mean Square ( RMS ) . . . . . 20
    - 3.1.3.5 Energy and Power content . . . . . 20

3.1.4 Random Forest based Classification and Combinational Strategy . . . . 20
3.1.4.1 Random Forest (RF) Classifier . . . . . . . . . . . . . . . . . . . 21
3.1.4.2 Classifier Combination over Different Features and Frequency-bands . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
3.2 Gaze based Temporal Data Analysis for Word Familiarity Prediction . . . . . . 23
3.2.1 System Setup and Data Acquisition . . . . . . . . . . . . . . . . . . . . . 24
3.2.2 Temporal Data Pre-processing . . . . . . . . . . . . . . . . . . . . . . . . 25
3.2.3 Stochastic Gradient Descent Classifier . . . . . . . . . . . . . . . . . . . 25
3.3 Facial Expression Recognition for Word Familiarity Prediction . . . . . . . . . 27
3.3.1 System Setup and Data Acquisition . . . . . . . . . . . . . . . . . . . . . 27
3.3.2 Data Pre-processing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
3.3.2.1 Contrast Limited Adaptive Histogram Equalization (CLAHE) 28
3.3.3 Facial Coordinates Extraction . . . . . . . . . . . . . . . . . . . . . . . . 29
3.3.4 LSTM based Sequence Classification . . . . . . . . . . . . . . . . . . . . 29

**4 Experimental Results 32**
4.1 Dataset Description . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
4.1.1 EEG Dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
4.1.2 Gaze based Temporal Dataset . . . . . . . . . . . . . . . . . . . . . . . . 32
4.1.3 Facial Expression Dataset . . . . . . . . . . . . . . . . . . . . . . . . . . 34
4.2 Word Familiarity Prediction . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
4.2.1 On EEG Dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
4.2.1.1 Word Familiarity Prediction using RF . . . . . . . . . . . . . . 36
4.2.1.2 Borda Count Combination of Different Features and Frequency bands . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 36
4.2.2 On Gaze based Temporal Dataset . . . . . . . . . . . . . . . . . . . . . . 38
4.2.3 On Facial Expression Dataset . . . . . . . . . . . . . . . . . . . . . . . . 39
4.3 Word Meaning Recommendation . . . . . . . . . . . . . . . . . . . . . . . . 40
4.4 Comparative Analysis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42

**5 Conclusion and Future Scope 45**

**6 Additional Work 46**
6.1 A Survey on Neuromarketing using EEG Signals . . . . . . . . . . . . . . . . 46
6.2 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47
6.3 Computational Approaches for Neuromarketing . . . . . . . . . . . . . . . . . 49
6.3.1 Pre-processing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49
6.3.2 Feature Extraction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50
6.3.3 Classification and Results . . . . . . . . . . . . . . . . . . . . . . . . . . 51
6.3.3.1 Classification based on different brain lobes . . . . . . . . . . . 51
6.3.3.2 Classification based on User preferences . . . . . . . . . . . . . 53
6.3.3.3 Comparative Analysis . . . . . . . . . . . . . . . . . . . . . . . 54
6.4 Datasets . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55
6.4.1 Dataset Description . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56
6.4.1.1 Available Datasets . . . . . . . . . . . . . . . . . . . . . . . . . 56
6.4.1.2 Device Used . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 56
6.4.1.3 Number of Participants . . . . . . . . . . . . . . . . . . . . . . . 57
6.4.2 Dataset Type . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57

6.5 Challenges and Ethics . . . . . 58
6.5.1 Methodological Challenges . . . . . 58
6.5.2 Ethical Challenges . . . . . 59
6.6 Conclusions . . . . . 59

**Bibliography** **61**

**List of Publications** **72**

# List of Figures

1.1 User thinking about the meaning of the unfamiliar word. . . . . . . . . . . . . . . 2
1.2 Block Diagram for Multimodel Word Recognition process. . . . . . . . . . . . . 3

3.1 Block Diagram representing the steps of EEG signal processing in training & testing phase. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
3.2 EEG brain sensor: (a) Emotiv EPOC+ device and accessories (b) Placement of electrodes over skull. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 13
3.3 EEG signals (a) Raw Signal (b) Filtering and Smoothing Result (c) Two independent components after application of ICA. . . . . . . . . . . . . . . . . . . . 16
3.4 Different levels of Wavelet decomposition using DB8. . . . . . . . . . . . . . . . 18
3.5 EEG signals (a) ICA component (b) Gamma band wave (c) Beta band wave (d) Alpha band wave (e) Theta band wave (f) Delta band wave . . . . . . . . . 19
3.6 Borda Count Combination of classifiers based on their ranks. . . . . . . . . . . 22
3.7 Block Diagram showing the Borda Count combination of different classification results. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
3.8 Block Diagram representing the steps of gaze based temporal data analysis in training & testing phase. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24
3.9 Eye Tracking Device: (a) Tobii Eye Tracker 4C (b) Components of Tobii Eye Tracker 4C . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25
3.10 Block Diagram representing the steps of facial expression detection in training & testing phase. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
3.11 Kinect XBOX 360 Device . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
3.12 CLAHE Equalization: (a) Captured RGB image (b) Resultant equalized contrast image. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 29
3.13 Facial Coordinates Extraction: (a) Face detection in image (b) Coordinate marking on detected face. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 30
3.14 Long Short Term Memory Network. . . . . . . . . . . . . . . . . . . . . . . . 30

4.1 System setup where the user is reading a word during experiment. . . . . . . . 33
4.2 Words with three levels of difficulty shown to users on the computer screen. . . 33
4.3 Temporal Data Analysis: (a)Tracked Gaze on word ”worst” (b)Screen Coordinates of tracked gaze and corresponding timestamp. . . . . . . . . . . . . . . 34
4.4 Facial Expressions: (a) For Unfamiliar words (b) For Familiar words. . . . . . . 35
4.5 Extracted Facial Coordinates in various frames corresponding to unfamiliar word. 35
4.6 Confusion Matrix for the word familiarity prediction using RF classifier. . . . . 37
4.7 Accuracy prediction over different features of raw data. . . . . . . . . . . . . . 37
4.8 Accuracy prediction over different features of wavelet decomposed data. . . . . 38
4.9 Borda count combination of frequency bands and features. . . . . . . . . . . . . 38

4.10 Confusion matrix for Borda count combination of different features and frequency bands. . . . 39
4.11 Accuracy for words at different difficulty levels. . . . 39
4.12 Accuracy of classifier combination for every user. . . . 40
4.13 Confusion Matrix for the word familiarity prediction using SGD classifier. . . . 40
4.14 Confusion Matrix for the word familiarity prediction using LSTM classifier. . . 41
4.15 Meanings of unfamiliar words shown as the pop-up window on computer screen. 41
4.16 Accuracy prediction of proposed models over different datasets. . . . 42
4.17 Accuracy prediction of Power feature of Gamma frequency band over different brain portions. . . . 43
4.18 Accuracy prediction of Power feature of Gamma frequency band over different signal duration. . . . 43
4.19 Borda count combination of beta and gamma using RF and LSTM. . . . 44
4.20 Confusion Matrix for the word familiarity prediction using HMM classifier. . . 44

6.1 The process of neuromarketing where EEG response towards a product is recorded and processed to get the user's preference. . . . 47
6.2 A scenario of neuromarketing: User is watching the a product on the computer screen and EEG signals are recorded simultaneously. The BCI model predicts whether the person likes or dislike the product by analyzing brain signals. . . . 48
6.3 Another scenario of predicting video ratings using EEG signals. User is watching the video and the BCI model predicts the interest of user in terms of different ratings. . . . 49
6.4 Placement of EEG electrodes over skull as International 10-20 system: (a) 5 electrodes device [1], (b) 10 electrodes device [2], (c) 14 electrodes device [3], (d) 16 electrodes device [4], (e) 30 electrodes device [5] and (f) 61 electrodes device [6]. . . . 57
6.5 Various dataset types in Neuromarketing. . . . 58

# List of Tables


2.1 Summary of the Eye Tracking Related Work . . . . . . . . . . . . . . . . . . . . . 8
2.2 Summary of the EEG Related Work . . . . . . . . . . . . . . . . . . . . . . . . . 9
2.3 Summary of the Facial Expression Recognition Related Work . . . . . . . . . . 11

6.1 Preprocessing filters used in the field of neuromarketing . . . . . . . . . . . . . 50
6.2 Feature Extraction Techniques used in the field of neuromarketing . . . . . . . 51
6.3 Extracted Features used in the field of neuromarketing . . . . . . . . . . . . . . 51
6.4 Related work done in the field of neuromarketing considering "Product Images" as dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 52
6.5 Related work done in the field of neuromarketing considering "Advertisement Video" as dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53
6.6 Related work done in the field of neuromarketing considering "Color Visuals and 3D Virtual Products" as dataset . . . . . . . . . . . . . . . . . . . . . . . . 54

# Chapter 1

# Introduction

Word Familiarity is to detect if a person is already familiar with the given word in the text or not, as shown in Figure 1.1, where a user gets confused when he is not familiar with the word displayed on the computer screen. Unfamiliar words are referred as pseudo strings. When these words become familiar, then the reading speed increases and recognition is enhanced. For unfamiliar words, reading time increases with the word length as the user is reading it for the first time but this is not in the case with familiar words. In the process of word recognition, firstly, the letters and then their combinations to form the word are analyzed. After that, the syntactic and semantic information related to the word is retrieved. It is a very fast process which happens within the first 200 ms after looking at the word [7]. However, the effectiveness of the process of recognition depends on the properties of words and the contexts in which they appear. Semantic properties affect the speed of recognition because it is the meaning based process.

The whole process of word familiarity detection is explained in the block diagram given in Figure 1.2 where if the user is found to be unfamiliar with a word in the whole text then after searching in the dictionary, the meaning of that word will be shown on the computer screen.

For human-computer interaction systems, the computer needs to react according to the action performed by the person. And for this, the psychological state of a person needs to be determined. The psychological state includes the brain waves generated in the nervous system and corresponding facial expressions as a reaction to those brain waves. Also, the eye movement is very informative about the intentions and thoughts in a person's mind. The statement "What we are thinking is based on what we are looking" follows here. Eye gaze tracking is the measurement of eye movement with respect to the constant head position. Emile Java, a French ophthalmologist, first described the movements of the eye while reading the text, in 1879. He used the mirror to notice that the movements of the eye are not continuous but they are the combination of some rapid movements i.e., saccades and some small stops i.e., fixation


FIGURE 1.1: User thinking about the meaning of the unfamiliar word.

[8]. Eye tracker enhances the human-computer interaction. It gives information about the person's presence and focuses towards something. It can detect awareness, drowsiness and other psychological states. Eye tracking technology has various applications in the real world. These include Human-Computer interaction systems [9, 10], Market research [11] where the focus on particular brands and products shows the user's interest towards them, Diagnosis of medical issues [12] like Attention Deficit Hyperactivity Disorder, Obsessive Compulsive Disorder and Alzheimer's disease etc., drowsiness detection [13] and PC and Gaming Research where humanized user interfaces are created by combining eye tracking with other input methods such as, keyboards, mouse, touchpad etc. It makes the gaming environment more interesting and attractive.

The human brain is analogous to a network with nodes and meshes where nodes correspond to neurons and meshes correspond to pathways. Neurons in the brain communicate via electrical impulses. EEG(Electroencephalograph) signals indicate any nervous excitement by detecting brain activities. It is a graph of low voltage levels versus time. Firstly, Berger recorded the EEG signals in 1929 by applying electrodes on the human scalp [14]. EEG has been used in various fields like Neuro-marketing where the EEG signals of customers are recorded when they are exploring the stores to buy any product. These signals are further used to predict the customers' choices [15]. It can be used in prediction of advertisement preference. Gauba et al. [16] have given a fusion model where sentiment analysis of users have been combined with their EEG response to predict their advertisement preferences. Also, EEG signals can be processed to observe psychological states of human beings to predict their behavior in various situations. A major application of EEG can be seen in the field of medical treatment to monitor the improvement in brain activities with time. EEG can be used in the diagnosis of

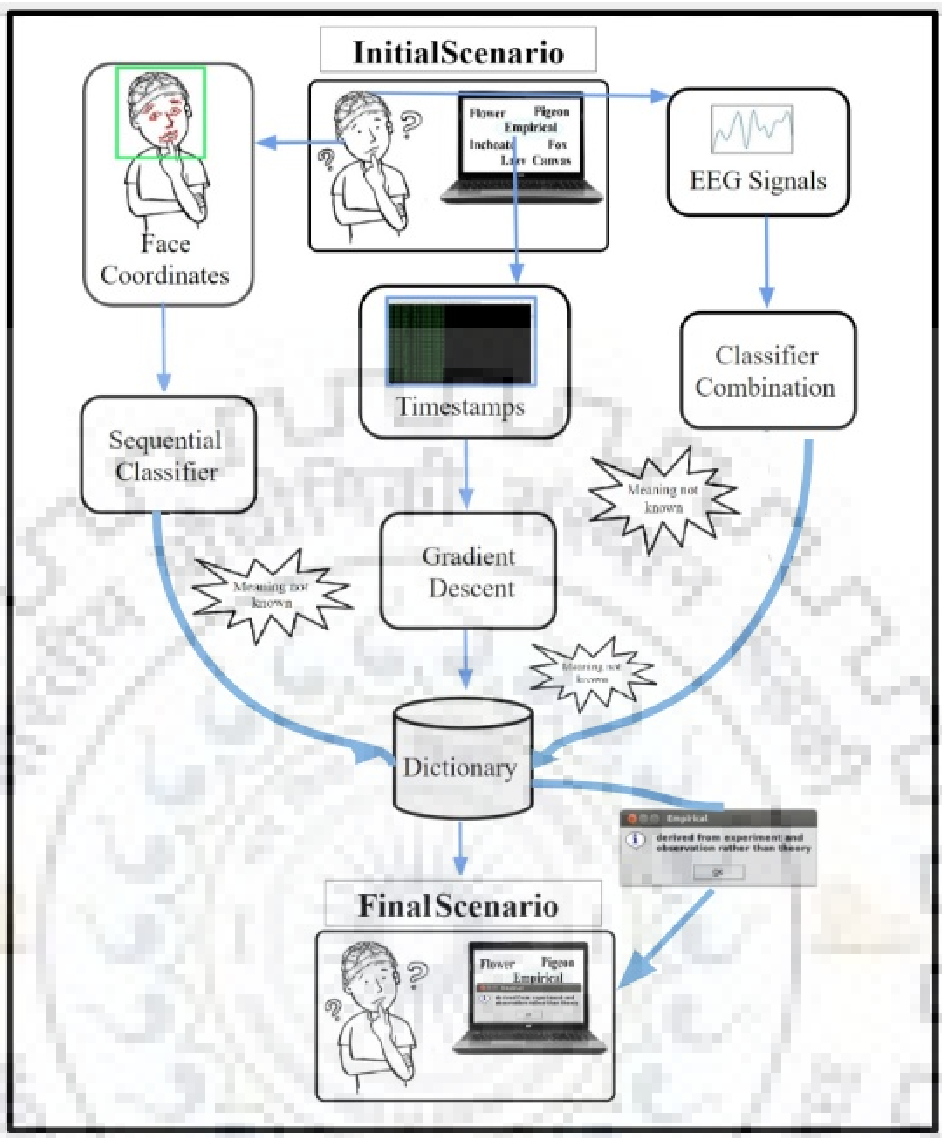


FIGURE 1.2: Block Diagram for Multimodel Word Recognition process.

epilepsy. In [17], the authors have used approximate entropy based neural network to design an automatic epileptic detection system by using EEG signals. Brain-Computer-Interface (BCI) is an emerging field for EEG where the movement of a device can be controlled using brain waves. EEG can be used for the learnability of a software. Stickel et al. [18] have proposed a learning assessment methodology with dominant alpha waves.

Facial expressions give information about a user's identity, state of mind and his intentions. These things help in creating human-computer interaction applications. The expressions can be extracted from either images or videos. Various applications of facial expression recognition include clinical purposes like understanding patient's psychological state which helps to monitor them properly [19], e-learning to detect the state of learner [20, 21], to detect the state of driver alert [22], marketing to find the taste of customers [23], security systems to uncover criminals [24].

Word familiarity using temporal data analyzes the timestamps of the change of gaze by tracking eye movements. More is the fixation of gaze at a particular word, more are the chances of the unfamiliarity with the word. In EEG based Word-familiarity, event-related potentials (ERPs) are considered. N400 component is related to Lexical activation and semantic processing. If this potential is short then it corresponds to partial activation of both dominant and subordinate targets. If the value is large then the appropriate meaning is activated where as dominant is partly activated. All the meanings in the brain lexicon are activated but the degree of activation depends on the frequency and type of context. In this study, the principle of ERPs has been followed to predict whether the user knows the meaning of the word or not based on the stimulation in the brain and EEG signals are recorded to indicate the nervous excitement. In facial expressions based word familiarity, the change of expression from a reference neutral expression is considered. The corner points of particular portions of the face, such as the eyebrows, eyes, mouth, and nose are located and extracted, and then their variations in size and orientation are calculated to predict whether the user is confused about something or not, which helps in predicting word familiarity.

## 1.1 Problem Statement

Word familiarity is the process of finding a particular word in the text and then detecting whether a user already knows the meaning of the word, he is reading. The problem can be divided into the following subproblems:

1. To propose a system that will find the desirable word from the text and then show the corresponding meaning of the word when the user is found to be unfamiliar.
2. To devise a system that will use temporal data to predict word familiarity.
3. To devise a system that will use brain signals to predict word familiarity.
4. To devise a system that will use facial expressions to predict word familiarity.
5. To perform classifiers' combinations based on the different criterion and enhance the overall performance of the system.

## 1.2 Research Gap and Our Contribution

Over the past few years, different methods for analyzing the user familiarity with the word have been developed by researchers. Researchers have used different algorithms for eye tracking. These include video or image-based processing, pupil detection, eye movement recording in terms of fixation point and duration and simultaneous quick movement of both eyes in the same direction between two or more fixation phases etc. Daugman et al. [25] and Deng et al. [26] have used iris contours, edges and outer boundaries for gaze tracking which is not that accurate as compared to combination of iris and head pose information. Morimoto et al. [27] have used pupil-corneal reflection technique. Pupil-corneal reflection is not quite accurate for general interactive applications because it needs the calibration process beforehand and does not allow free head motion. The authors have proposed systems based on the fixed head position which is actually not the case in real life applications. Ji et al. [28] have used special Infrared illumination technique where pupils appear brighter than the rest of face in the captured image. But it quite a cumbersome task to separate pupils from other bright portions in the image. Other problems with the existing models are difficulty in the system setup, controlling the illumination conditions, correct placement of the camera according to the head position, removal of noise data where the eyelashes are interfering, calibration problem, difficulty in considering users who wear glasses or contact lens, dry and watery eyes of users and head motion. Also, it is difficult to track eye gaze when there are poor light conditions. Infrared source of light can eliminate this problem which illuminates the eye steadily and enhances the pupil to make it look white.

Burgress et al. [29] have analyzed ERPs using the method of event-related desynchronization in different frequency ranges. In [30], the authors have separated familiarity-based recognition and the recollection by source unitization whereas, in [31], the familiarity of word-pairs has been considered instead of a single word. Both eye movement and EEG recordings are considered as complementary to each other in [32] to check the familiarity with a word in reading. Bentin et al. [33] has considered the negative potentials related to word-antonym, word-non-antonym, nonword-word and nonword-nonword pairs. Some of the researchers have used MEG to capture the activity of the brain that is caused by the magnetic field created by neuron activity. For example, Pylkkanen et al. [34] have proposed the use of MEG instead of EEG to track the time course of word recognition. The authors have shown that the positive or negative components of voltage potentials have their own meanings. They have conducted experiments to show that there is a correlation between the participant's brain responses and the words stored in their memory, i.e., analysis of the potential values obtained from the EEG recordings can tell about the brain activity. However, they have not trained the machines according to the response values manually given by the participant. Also, just analysis of the raw potential values without any further processing will never give correct results. However,

the cost and technical complexities are high for MEG. To overcome these limitations, an inexpensive setup has been used to examine the brain activity using EEG signals.
Zhang et al. [35] have proposed a tutorial system where the vision-based interface is recognizing facial expressions to find the degree of understanding. But, they have considered only the frontal face view. The challenge in such a tutorial system is that the facial expression recognition and then corresponding mental state inference is a very subtle task and wrong interpretations are not affordable in this case. So, it's better to fuse this model with some other model results. Loh et al. [36] have also performed facial expression recognition in the e-learning environment. Again the problem was the frontal view of images. They have directly applied the expression detection step but before expression detection, the face has to be detected first. Otherwise, the scaling problem may arise where it would be difficult to localize special facial features. Also, the coordinates may change when the face moves to different parts of the frame. So, the coordinate values need to be normalized between 0 and 1. The challenges may include non-frontal face, non-ambient light conditions, some facial feature information loss when lips are covered by mustache or forehead is covered by hair. These issues need to be considered beforehand.

Following are the contributions of the proposed framework:

1. Firstly, a framework has been proposed for predicting the word-familiarity while reading using eye gaze temporal data, brain signals and facial expressions. The eye gaze fixation, EEG signals, and video of the user have been recorded simultaneously while reading.
2. Secondly, the EEG signals have been processed in time-frequency domain for feature extraction. Image frames and then the facial coordinates have been extracted from the recorded video. Timestamp values with the screen coordinates corresponding to movement of eyes have also been extracted. Next, a classifier combination approach is used for evaluating the system performance.
3. Finally, a comparative analysis of all proposed approaches have also been presented.

Rest of the report is organized as follows. Chapter 2 provides an overview of the existing research. In Chapter 3, methodology including preprocessing, feature extraction and classification methodology have been presented. In Chapter 4, the results obtained by the combination of different classifiers have been presented. In Chapter 5, the conclusion of the report and the possible future extensions have been stated. In Chapter 6, the additional work performed with the entitled work has been presented.

# Chapter 2

# Literature Review

The recent works done in the field of word familiarity and the proposed methodology for the same have been reviewed in this chapter. Firstly, we present the survey of work done in the field of eye tracking and then we present the survey of work done in the field of EEG and facial expression detection.

## 2.1 Eye Tracking

The recent works that relate the activity of eye in accordance with the text on the screen to predict the familiarity of the user with the proposed word have been reviewed in this section. Daugman et al. [25] and Deng et al. [26] have used iris contours, edges and outer boundaries for gaze tracking with the constant head position. Morimoto et al. [27] have used pupil-corneal reflection technique. Ji et al. [28] have used special Infrared illumination technique where pupils appear brighter than the rest of face in the captured image. Hutchinson et al. [37] have used a unique user interface system, Erica where the eye tracking is performed, based on the relative position of glint and bright eye in the captured image. The glint means the fraction of reflected infrared light from the surface of the cornea. The bright eye means the image of the pupil formed when the infrared light is reflected from the retina. Firstly, they have extracted the horizontal and vertical coordinates of the center of glint and bright eye and then mapped them to the stored reference points. They have faced the problem of bright eye effect which means, not everyone has sufficient bright eyes. And also, it is difficult to operate the system without head movements which make the camera out of focus. Sibert et al. [38] have performed the experiment to measure the time duration for simple computer tasks like selecting a highlighted circle from a bunch of circles, selecting a spoken letter from a group of letters etc. using eye gaze. And the results show that the eye gaze interaction has a speed advantage over the mouse interaction. They have shown how this eye gaze interaction works and what are its advantages. Reulen et al. [39] have used infrared oculography to track the iris position. Hansen et al. [40], Orman et al. [41] and Smith et al. [42] have developed video-based eye tracker where they have used the concept that the captured image of the eye will change when there is a movement of the eye in 3D space. Some researchers have directly tracked the eye movement based on their appearance in the captured image by mapping the image data to the screen coordinates. Yang et al. [43] have mapped the coordinates by firstly processing images to grayscale unit images while Sugano et al. [44] have considered the Gaussian regression data for the same. A summary of the related work using eye tracking is presented in Table 2.1.

Table 2.1: Summary of the Eye Tracking Related Work

| **Author & Year** | **Approach** |
|---|---|
| Daugman et al. [25], 1993 | Iris contours, edges for gaze tracking with still head |
| Deng et al. [26], 1997 | Iris contours and outer boundaries for tracking with constant head position |
| Morimoto et al. [27], 2005 | Pupil Corneal Reflection Technique |
| Ji et al. [28], 2002 | Infrared Illumination Technique |
| Hutchinson et al. [37], 1989 | Relative position of glint and bright eye (UI system: Erica) |
| Reulen et al. [39], 1988 | Infrared Oculography |
| Hansen et al. [40], 2010 | Eyes position changes in video |
| Orman et al. [41], 2011 | Captured Image of eye will change with movement in 3D space |
| Yang et al.[43], 2012 | Additional pre-processing of images to grayscale in video |

## 2.2 Electroencephalography (EEG)

The recent works that relate the activity of EEG signals in predicting the familiarity of the user with the proposed word have been reviewed in this section. Diana et al. [30] have separated familiarity-based recognition and the recollection. Recollection is when the subject can recollect its memory about the story behind that particular word whereas familiarity-based recognition is just that the subject is familiar with the given word. For this, they have used the process of unitization with the source object like a color is shown in the background of the word and the subject will imagine the word object in the same color. During analysis of EEG signals, they have found that correlation of EEG signals with familiarity-based recognition is more affected by the unitization than the correlation of EEG signals with the recollection. The p-component of ERPs have been studied for this and it was found that the positivity from 1250 to 1500 ms of the parietal lobes was related to the source recollection and the positivity from 750 to 1000 ms of mid-frontal lobes was related to the familiar responses. They have recorded an accuracy of 90% for high unitization and 55% for low unitization. In [29], the authors have analyzed ERPs using desynchronization related to event potential in different frequency ranges. They have found that in the theta frequency band, there was an increase in power for a short duration, i.e., in the first 250 ms. There was also a repetition effect in theta band such that at the mid-frontal electrode, the synchronization of the new words was greater than the old words. In the alpha band, the repetition effect occurred from 750 ms at the left temporal-parietal sites. The things are interpreted using the existing models of memory recognition.

Rhodes et al. [31] have considered familiarity of word-pairs instead of a single word. Word pairs can possess relationship like association or semantic or both. Subjects were asked to give response for a word and simultaneously their ERPs are recorded and analyzed. On the basis of the values of ERPs, they found whether the given word pair is familiar or not. For the same words, the accuracy was from 75% to 83% for different types of word pairs and for the new words, the accuracy was from 85% to 87%. In [32], the authors have used both eye movement and EEG recordings as complementary to each other to check the familiarity with a word in reading. They have recorded ERPs and analyzed its N1 component for both High-Frequency and Low-Frequency words. The amplitude of the N1 component from 132–164 ms was less for High-Frequency than Low-Frequency words. The authors in [33] have conducted the experiment on word-antonym, word-nonantonym, nonword-word and nonword-nonword pairs. After analysis of ERPs, it was found that antonyms were recognized faster than words or nonwords. Negative potentials were considered. The negativity related to nonwords was larger than the negativity related to words and the negativity related to the antonym was almost non-existent. Word-antonym pair detection was more accurate than any other pair.

Table 2.2: Summary of the EEG Related Work

| **Author & Year** | **Approach** | **Word Dataset & Participants** |
|---|---|---|
| Diana et al. [30], 2011 | EEG, correlation with unitization, P component of ERP | 360 three-to-eight letter English nouns, 17 subjects |
| Burgress et al. [29], 2000 | EEG, event-related desynchronization | N.A. |
| Rhodes et al. [31], 2007 | EEG, word pairs | 408 three-to-nine letter word-pairs,25 subjects |
| Sereno et al. [32], 2003 | EEG, eye tracking, N1 component | High and low frequency words, N.A. |
| S. Bentin [33], 1987 | EEG,word antonym pairs, negative potentials | 80 pairs of antonyms and 40 additional words, 16 subjects |
| Mormann et al. [45] , 2005 | EEG, wavelet, gamma and theta bands | continuous words, 12 epilepsy patients |
| Holcomb et al. [46], 2006 | EEG, N400 | animal names |
| Hauk et al. [7], 2006 | EEG, word length, Semantic coherence | 300 nouns, 20 subjects |

Mormann et al. [45] have analyzed the power changes in various frequency bands after the wavelet transformation. And further found that there is a modulation of gamma activity by the theta cycle which is potentially related to memory encoding. Holcomb et al. [46] have shown that the N400 and other ERP components' (P150, N250, and P325) modulation gives sequential overlapping steps in words processing. The experiment was done by showing animal names followed by nonanimal items that were in complete repetitions, partial repetitions or completely unrelated to the previous word. And ERPs were recorded and analyzed for them. Similarly, the authors in [7] have conducted the experiment to find the time course to access the information related to the linguistic behavior and psychological process during word remembrance task and they have analyzed Word length, Letter n-gram frequency, Lexical frequency and Semantic coherence for the same. It was found that effect of Word length and Letter n-gram frequency can be seen around 90 ms where as the effect of Lexical frequency can be seen at 110 ms and the effect of Semantic coherence can be seen at 160 ms. A summary of the related work using EEG is presented in Table 2.2.

## 2.3 Facial Expression Recognition

Zhang et al. [35] have proposed a tutorial system where the vision-based interface is recognizing facial expressions to find the degree of understanding. Loh et al. [36] have performed facial expression recognition in e-learning environment. They have used Gabor Wavelet for the extraction of 18 facial features and then back propagation neural network for classification of those expressions. The choice of Gabor wavelet is great as it is light insensitive. They have considered 600 images(342 females and 258 males) with 4 expressions i.e., neutral, sleepy, confuse, smile. They have achieved the accuracy of 72.37% for confusing, 76.89% for smiling and 79.85% for sleepy expressions. Khan et al. [47] have used an unpopular way to recognize facial expressions. They have used an infrared camera to capture images. After normalization, the thermal intensity is calculated for facial features and then PCA is applied. They have used LDA for expression classification. Their dataset consists of 21 participants with five expressions, i.e., neutral, happiness, disgust, sadness, and fear. Smith et al. [48] have first determined the size and position of the face in the image and then the expressions are recognized. They have implemented a driver's attention recognition system where the position of the user is constant with respect to the camera. Orazio et al. [49] have devised an algorithm to detect eye in the facial image by finding the geometrical relations among various facial features. Cristinacce et al. [50] have used a multi-stage approach where the face is detected first by boosted cascaded classifier and then facial features are detected using the same classifier. They have used Pairwise Reinforcement of Feature Responses to improve localization efficiency. Ioannou et al. [51] have used SVM for face detection and then eye areas are located using a feed-forward neural network. After that eyebrows and nostrils are detected by considering eyes as reference points. And then the mouth is located. Cohn et al. [52] and Goneid et al. [53] have used image sequences or videos for facial expression detection while some researchers have used static images. Zhao et al. [54] and Terzopoulos et al. [55] have used facial features while Kotsia et al. [56] and Rosenblum et al. [57] have used image-based representation of face. But they faced the challenge of varying environmental conditions and subjects. Hammal et al. [58] and Kim et al. [59] have classified discrete facial emotions while Xiang et al. [60] have considered the facial action causing the expressions using Facial Action Coding System. A summary of the related work using Facial Expressions is presented in Table 2.3.

Table 2.3: Summary of the Facial Expression Recognition Related Work

| **Author & Year** | **Approach** |
|---|---|
| Zhang et al. [35], 2002 | Vision based interface for recognizing facial expression |
| Loh et al. [36], 2006 | Gabor Wavelet, Back propagation neural network |
| Khan et al. [47], 2006 | Infrared camera, Thermal intensity |
| Smith et al. [48], 2003 | Expressions are recognized using size and position of face with constant user position |
| Orazio et al. [49], 2004 | Detecting eye in the face for facial expression |
| Cristinacce et al. [50], 2004 | Boosted cascaded classifier |
| Ioannou et al. [51], 2007 | SVM, Feed forward neural network |
| Cohn et al. [52], 2007 | Image Sequences or videos |
| Hammal et al.[58], 2007 | Discrete facial emotions |
| Xiang et al. [60], 2008 | Facial Action Coding System |

# Chapter 3

# Methodology

The methodology proposed for word familiarity prediction has been phrased in this chapter. Firstly, we present the proposed method using EEG analysis in Section 3.1 and using gaze based temporal data analysis in Section 3.2 and facial expression recognition in Section 3.3.

## 3.1 Electroencephalography (EEG) Signal Analysis for Word Familiarity Prediction

To process the EEG signals, the steps in Figure 3.1 are followed. Firstly, the phase of the collection of EEG signals occurs. After that these signals are pre-processed using some filters and then the EEG features are extracted by various methods. At last, these features are classified based on the psychological task they depict.

### 3.1.1 System Setup and Signal Acquisition

The acquisition of EEG data is performed while users were asked to read words shown on the computer screen. A wireless device, Emotiv EPOC+ headset has been used to capture the EEG signals [61]. The device has various advantages like easy to fit over scalp due to its flexible design, wireless and rechargeable with a battery backup of 12 hours with continuous use, dense array spatial resolution which ensures the measurement of complete brain signals, salined wet sensors which ensures the avoidance of sticky gels, and compatibility with different operating systems (e.g. Windows, OSX, Linux, Android, and iOS) [62]. The device has 14 electrodes that are placed on the scalp at positions Occipital (O1, O2), Parietal (P7, P8), Temporal (T7, T8) and Frontal (AF3, AF4, F3, F4, FC5, FC6, F7, F8) as per International 10-20 system as shown in Figure 3.2.

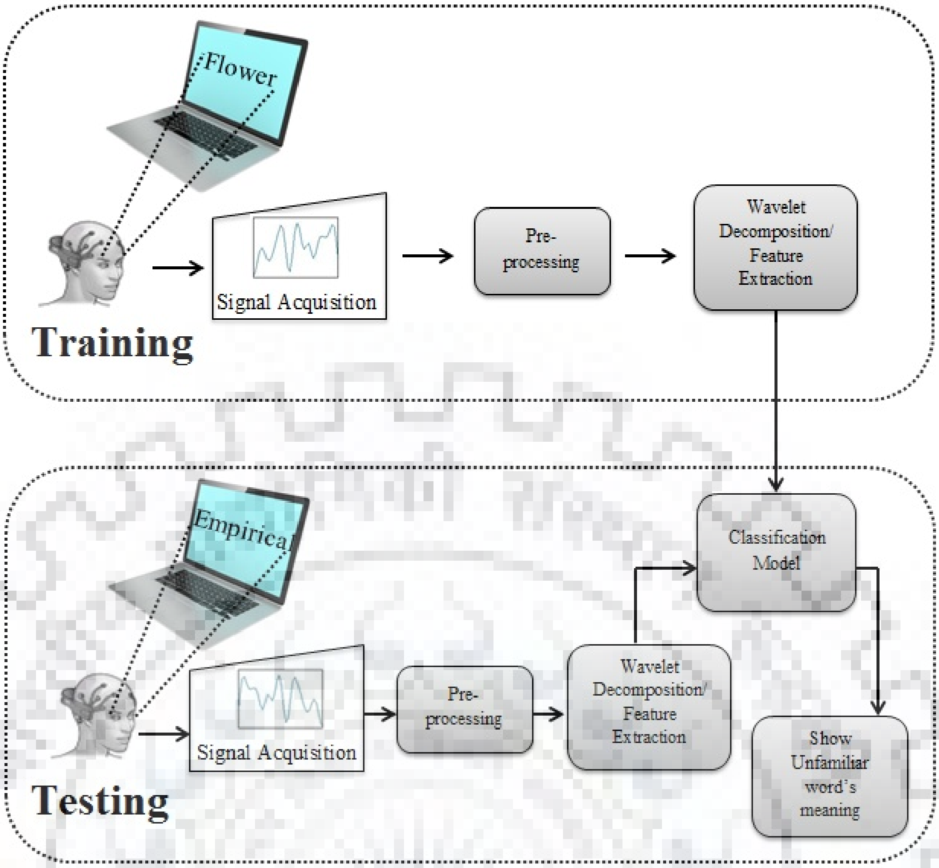


FIGURE 3.1: Block Diagram representing the steps of EEG signal processing in training & testing phase.

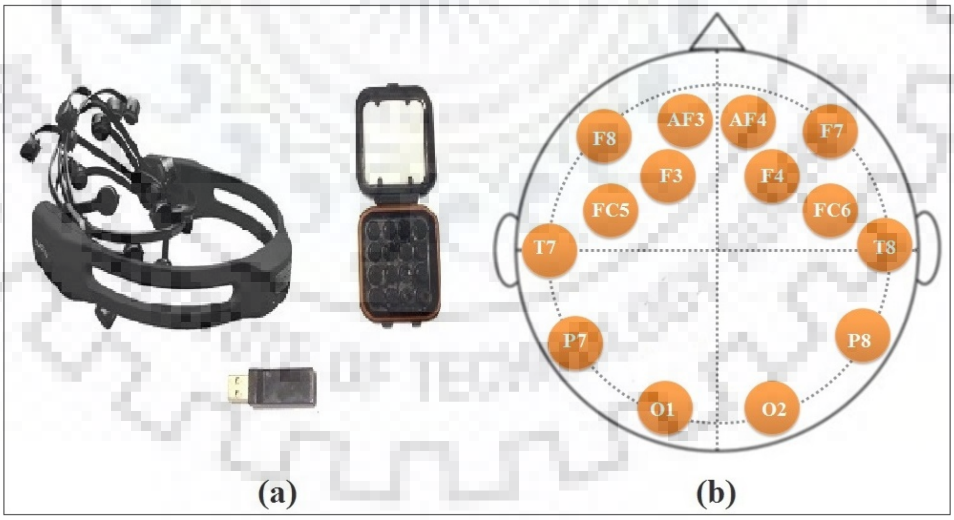


FIGURE 3.2: EEG brain sensor: (a) Emotiv EPOC+ device and accessories (b) Placement of electrodes over skull.

EEG signals were collected by placing these electrodes on the human scalp which capture the brain activity in terms of the weak (5-100$\mu$V) electrical potentials generated by the brain. The electrode contains a disk connected by wire. EEG data is captured in a laptop using Bluetooth connectivity via a USB dongle. Initially, an EPOC device is sampled at 2048 Hz frequency per channel which is then down-sampled to 128 Hz per channel. Users were asked to read words one after the other and simultaneously, EEG signals were recorded. After that, the users were asked whether they were familiar or unfamiliar with those words. The responses of the users were taken in terms of two categories, i.e., familiar or unfamiliar.

### 3.1.2 Signal Pre-processing

Various sources of artifacts like muscular activities, blinking of eyes, and electrical noise in the power line, etc. come into picture while capturing the signals[63]. These artifacts and noise very badly affect the useful features in the original signal. The magnitude of the brain-generated electrical potentials is smaller than these unwanted signals. Many solutions can be thought to remove these unwanted signals like the removal of contaminated trials and restriction of eye blinks. But that is natural and not in the hand of the user. Also, the removal of trials can cause the useful data loss [64]. So, some different methods have to be applied to eliminate these unwanted signals.

#### 3.1.2.1 Data Filtering & Smoothing

As noise can interfere with the actual signal, so it is required to separate both of them to prevent interference. Bandpass filtering is the process to limit the bandwidth of the output signal to the required band and ignore out of band signals [65]. For this, a bandpass finite impulse response digital filter has been designed. It allows frequencies within a pre-decided range and rejects other frequencies. The output of the filter is calculated one by one on each electrode signal using present inputs, past or delayed inputs, and the past or delayed outputs of that electrode and the amount of delay is given by the order of filter. A 5th-order band pass FIR filter with lower cut-off frequency 0.49 Hz and higher cut-off frequency 60 Hz with sample rate of 256 Hz has been used to achieve zero-phase digital filtering in both the forward and reverse directions.

After filtering, Smoothing of data has been performed one by one on each electrode signal. It removes the high peaks in the data by using moving average filter where every data point is replaced by the average of its neighboring data points within the defined span. The neighboring data points are decided on the basis of the data point which is at the center of the span and there are no neighboring data points for the end points, so they cannot be smoothed. For calculating the average, it uses the Eq. (3.1). A 5-point moving average filter has been used for data smoothing where the span is 5.

$$s(i) = \begin{cases} x(1), & \text{for i=1} \\ \frac{1}{3}(x(1) + x(2) + x(3)), & \text{for i=2} \\ \frac{1}{2n+1}(x(i+n) + x(i+n-1) + ... + x(i-n)), & \text{otherwise} \end{cases} \tag{3.1}$$

where '$x(i)$' is the $i^{th}$ data point to be smoothed and '$s(i)$' is its smoothed value, '$n$' is the number of neighbors on both sides of '$x(i)$' and $(2n + 1)$ is the span.

#### 3.1.2.2 Independent Component Analysis (ICA)

The generative model used to find mutually independent sources from multivariate data that is a linear mixture of various latent variables is known as ICA. It does so by finding covariance and Eigen values. The mixing system could be unknown and non-gaussian latent variables could be assumed. These independent latent variables are the end results of ICA. It can be thought similar to blind source separation which is used for parallel signals set or time series dataset [64]. The main aim of ICA is to maximize the statistical independence or non-normality of the estimated components. It is applicable on EEG data as EEG signals are additive because, at electrode level, the sources are linearly combined electrical fields. The component '$x_i$' of the dataset $x = (x_1,...,x_m)^T$ is generated as a sum of the independent components '$s_k$', where $k = 1,...,n$ by using the Eq. (3.2). Here, FastICA method has been applied [66].

$$x_i = a_{i,1}s_1 + ... + a_{i,k}s_k + ... + a_{i,n}s_n \tag{3.2}$$

where '$a_{i,k}$' represents the mixing weights. The independent components can be calculated using Eq. (3.3), where '$W$' is the transformation weight matrix.

$$s_k = (w^T * x) \tag{3.3}$$

But the mixing of the sources does not need to be linear. Using a nonlinear mixing function $f(s|\theta)$ with nonlinearity parameter '$\theta$', the nonlinear ICA model is given by Eq. (3.4), where '$\theta$' is used as $(-exp(-1/2 * y^2))$.

$$x = f(s|\theta) + n \tag{3.4}$$

Raw data, filtered data, and data after applying ICA is shown in Figure 3.3.

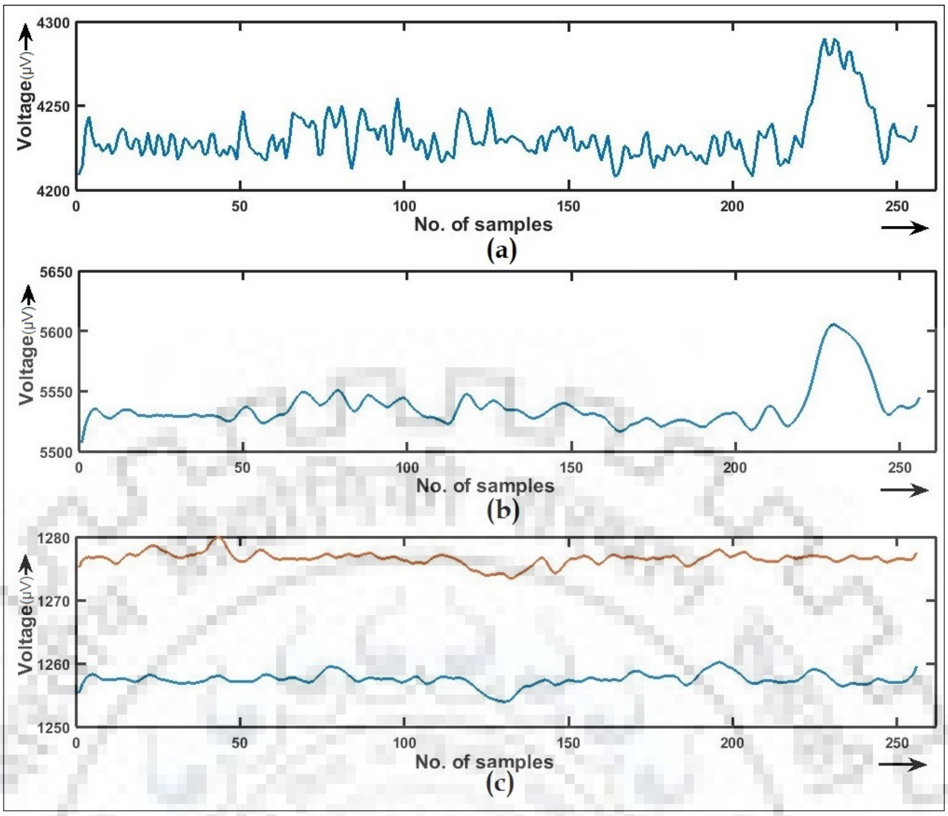


FIGURE 3.3: EEG signals (a) Raw Signal (b) Filtering and Smoothing Result (c) Two independent components after application of ICA.

### 3.1.3 Features Extraction

A feature is a characteristic or a structure extracted from a segment of the pattern. Since the EEG signal is a time domain signal, its features are buried away in the noise and to extract the useful data from them, they have to be transformed from time domain to the frequency domain. Also, highly trained professionals are required to inspect EEG signals in their unprocessed form [67]. In this work, we have applied Discrete Wavelet Transform (DWT) technique to process the signals.

#### 3.1.3.1 Discrete Wavelet Transform (DWT)

The wavelet transform provides a powerful technique for processing EEG signals before the classification process [68]. DWT applies multistage decomposition over input signal to convert it into a series of small waves. For this, the signal $S$ has to pass through digital filters, i.e., Low pass filter ($L$) and High pass filter ($H$) and then down-sampled by 2 which gives output as detail ($D$) and approximation ($A$) coefficients respectively. $A$ can be further decomposed at the next level using these filters. The wavelet function at time '$t$' is given by Eq. (3.5).

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{3.5}$$

where,

$$\psi_{m,n}(t) = {a_0^{-m/2}}\psi(a_0^{-m}t - nb_0) \tag{3.6}$$

where '$a_0$' is the scaling parameter, '$b_0$' is the translation parameter and $m = 0, 1, 2,..., M - 1$, $t = 0, 1, 2,..., T - 1$, $M = log2(T)$, $n = 0, 1, 2,..., 2^m - 1$ and '$T$' is the length of the signal. The decomposition of a signal into approximation and details coefficients can be achieved by choosing $a_0 = 2$ and $b_0 = 1$ [69]. As the characteristics of the signal depend on only half of the output of every filter, so the decomposition has halved the time resolution and hence the frequency resolution has been doubled. The down-sampling is performed using Eq. (3.7), where '$k$' is the subsampling rate and '$y[n]$' represents the signal.

$$(y \downarrow k)[n] = y[kn] \tag{3.7}$$

The approximation and detail coefficients are computed by using Eq. (3.8) and (3.9). Eq. (3.8) gives the scaling function and Eq. (3.9) gives the wavelet function. The scaling function belongs to '$L$' and wavelet function belongs to '$H$'.

$$\phi_{m,n}(t) = 2^{m/2}l(2^m t - n) \tag{3.8}$$

$$\psi_{m,n}(t) = 2^{m/2}h(2^m t - n) \tag{3.9}$$

The values of '$A_i$' and '$D_i$' at the $i^{th}$ level decomposition are computed using Eq. (3.10) and (3.11), respectively.

$$A_i = \frac{1}{\sqrt{T}} \sum_t x(t) \times \phi_{m,n}(t) \tag{3.10}$$

$$D_i = \frac{1}{\sqrt{T}} \sum_t x(t) \times \psi_{m,n}(t) \tag{3.11}$$

Eight levels of EEG signal decomposition has been performed by using the orthonormal wavelet Daubechies-8 (DB8) technique as shown in Figure 3.4. The DB8 decomposition gives first four wavelet coefficients that correspond to noise and five wavelet coefficients corresponding to frequency bands defined as:

- Gamma - ( >30) Hz : Corresponds to consciousness and memory matching of recognized objects, sounds or sensations.
- Beta - (12 - 30) Hz : Corresponds to active thinking, focus, high alert and anxiety.
- Alpha - (8 - 12) Hz : Corresponds to the resting state for the brain i.e. shows calmness and alertness.
- Theta - (4 - 8) Hz : Corresponds to dreams, fears and nightmares.
- Delta - (0.5 - 4) Hz : Corresponds to deepest meditation and dreamless sleep.

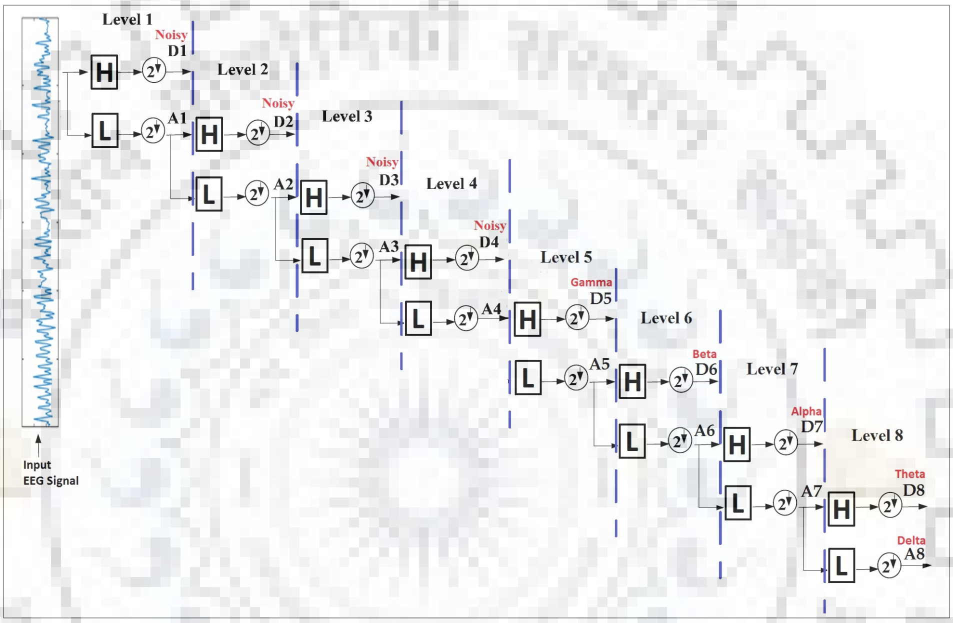


FIGURE 3.4: Different levels of Wavelet decomposition using DB8.

The wavelet decomposition is shown in Figure 3.5.

After DWT, the first order features (mean and standard deviation (SD)) have been extracted on those decomposed sub-bands or DWT coefficients one by one on every electrode signal. And the Linear features such as Root Mean Square (RMS) and Energy content of the signal uses these first order features. The four features, i.e. Mean, Standard Deviation, RMS and Power have been chosen for the experiment. These features have been used earlier by many researchers in various EEG based experiments and have given good results.[16, 70, 71].

#### 3.1.3.2 Mean

Mean is to calculate the average of all the data points and can be calculated using Eq. (3.12).

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \tag{3.12}$$

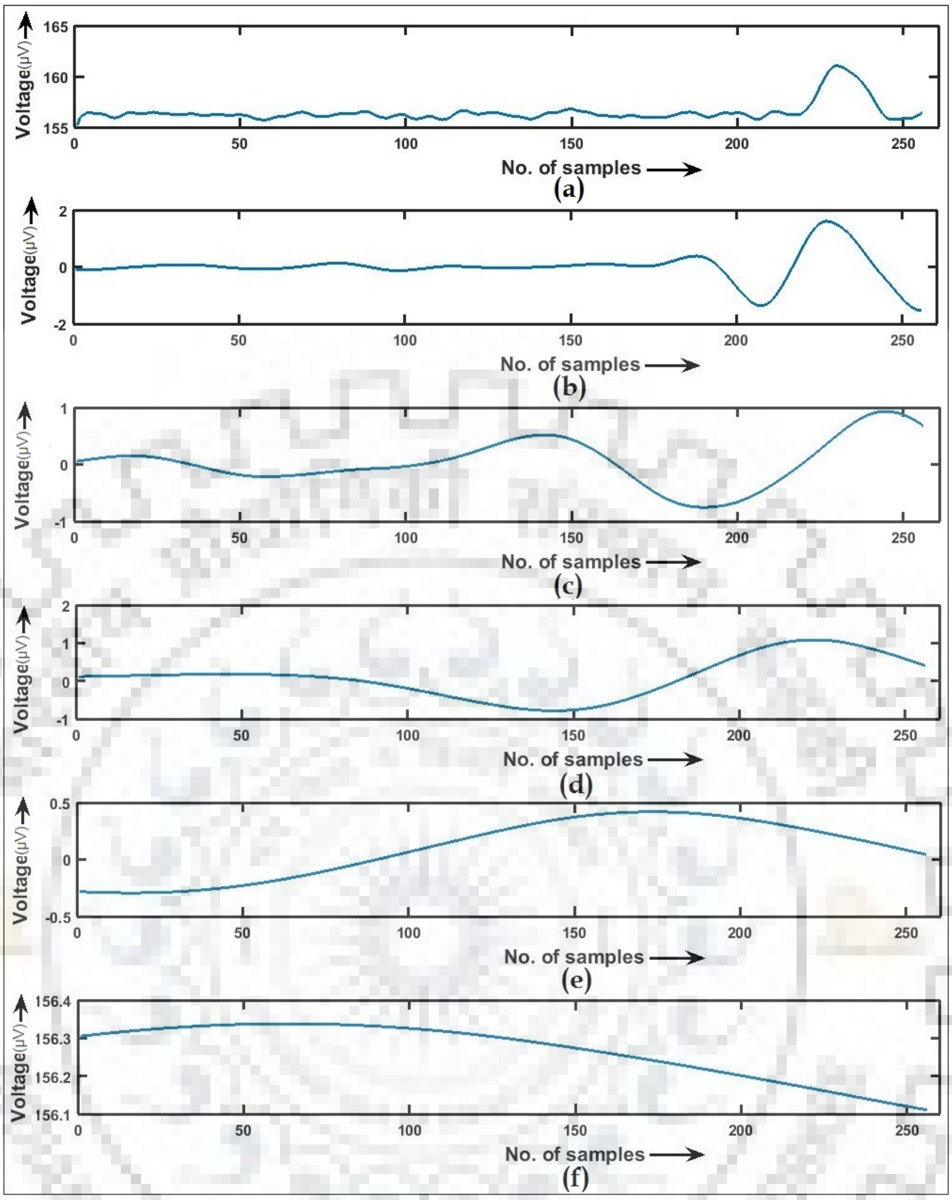


FIGURE 3.5: EEG signals (a) ICA component (b) Gamma band wave (c) Beta band wave (d) Alpha band wave (e) Theta band wave (f) Delta band wave

where, '$x_i$' are the data points and '$N$' specifies the number of samples.

#### 3.1.3.3 Standard deviation

It measures the deviation of the data from its mean. The higher deviation in the data set means there is a large gap between data points and their mean. It is calculated as the square root of variance which is the average of the sum of the squares of the difference between the data point and its mean as defined in Eq. (3.13), where, '$x_i$' is the current data point, '$\mu$' is the mean value of the signal and '$N$' is the number of samples.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}} \tag{3.13}$$

#### 3.1.3.4 Root Mean Square ( RMS )

It can be called as the quadratic mean measure. It shows the variation in the amplitude of the signal with respect to time. So, it depends on the shape of the wave and is independent of the signal frequency. That's why it is a time series statistical measure [72]. It is calculated by finding the square root of the mean of the square of the signal as defined in Eq. (3.14), where '$N$' is the number of samples and '$x$' is the amplitude value of the signal.

$$x_{RMS} = \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N}} \tag{3.14}$$

#### 3.1.3.5 Energy and Power content

The terms signal energy and signal power are used to characterize a signal and it is calculated using Eq. (3.15), where '$x_i$' is the current data point and '$N$' specifies the number of samples.

$$E = \frac{1}{2N+1}\sum_{i=1}^{N}|x_i|^2 \tag{3.15}$$

### 3.1.4 Random Forest based Classification and Combinational Strategy

In machine learning, classification means to identify to which category a new observation belongs, given a set of categories. The decision is made on the basis of a training dataset that contains those instances whose category is known beforehand. A classifier is an algorithm that implements the classification where the observational rows are termed as instances, the column variables are termed as features, to be predicted categories are termed as classes and the response value given by the user is termed as ground truth. Firstly, a classification model is built in the training phase by using the pre-processed and feature extracted recorded dataset and the response value given by the user's recognition of the word. After that, the test sample has to go through the same process of pre-processing, wavelet decomposition, feature extraction and then tested against the trained classification model in the testing phase. The results belong to two classes namely, YES (Familiar Word) and NO (Unfamiliar Word). Here, RF classifier has been used to model the features for the prediction of word-familiarity. The details are as follows.

#### 3.1.4.1 Random Forest (RF) Classifier

Random Forest has been used successfully in various classification problems based on EEG signals [73][74]. It creates a bunch of decision trees and then a vote is casted by every tree for the most suitable class. The output class of the random forest is the mode of the class's output by individual trees. As decision trees in different subspaces give their own classification results which can be the complement of each other and their combined classification can be monotonically improved, that is why RF classifier has been used [75]. A random set of features and bootstrapped samples are chosen from training data to build a classifier. It is constructed in a different way than classification trees. In decision trees, the best split among all variables is chosen to split a node where as in RF, the best subset among random predictors is chosen to split a node. The unknown samples are classified using a weighted or unweighted voting of a set of classifiers in the forest. Here, the training dataset is created by bagging technique, i.e., by randomly choosing '$N$' samples with replacement, where '$N$' is the size of the training dataset. And then to classify a test sample, the maximum voted class label from all the classifiers is assigned to it and with the predicted class label, the confidence score corresponding to that label has also been predicted. A step-by-step flow of the RF classification scheme is presented in Algorithm 1. In this work, a random forest of 100 trees for classification has been built by randomly drawing with replacement 275 examples, where 275 is the size of the original training set. Since, it is classification problem with 14 attributes, $14^{1/2}$ features (selected at random) are used in each split in the design of the decision tree.

**Algorithm 1** RandomForest

N : No. of training examples
M : No. of variables
m : No. of input variable for one node of tree, m <M
B : No. of classification Trees

**procedure** RandomForest(x[N][M],m,B)
  **for** b = 1 to B **do**
    To create tree $T_b$, draw sample of size N from N training examples with replacement.
    **while** nodesize != $n_{min}$ **do**
      Choose m variables randomly.
      Select best-split-variable among m chosen variables.
      Split node into two child nodes.
    **end while**
    Find classification result $C_b$ from $T_b$.
  **end for**
  **return** MajorityVote$[C_b]_1^B$
**end procedure**

#### 3.1.4.2 Classifier Combination over Different Features and Frequency-bands

The classifiers can be combined on several basis to produce better results. The combination can be performed over different classifiers or different frequency bands. The basis of the combination is the confidence value each classifier gives in its respective class. In this work, we have used Borda Count approach to combine the results from multiple classifiers. In Borda Count combination, for every classifier, class labels are ranked from most likely to least likely class label where the rank refers to the number of classes ranked below it. And then for a particular class label, ranks of all classifiers are combined. The output class label will be the one with maximum combined rank as shown in Figure 3.6. Eq. (3.16) is used to find the combined rank and Eq. (3.17) is used to predict output class label.

$$r_i = \sum_{j=1}^{N} {r_i}^j \tag{3.16}$$

$$c_{out} = max_{i=1}^{n}(r_i) \tag{3.17}$$

where '${r_i}^j$' is rank of classifier '$j$' for the class '$i$', '$n$' is the number of class labels and '$N$' is the number of classifiers.

Here, Borda count combination of beta and gamma bands with different features namely,

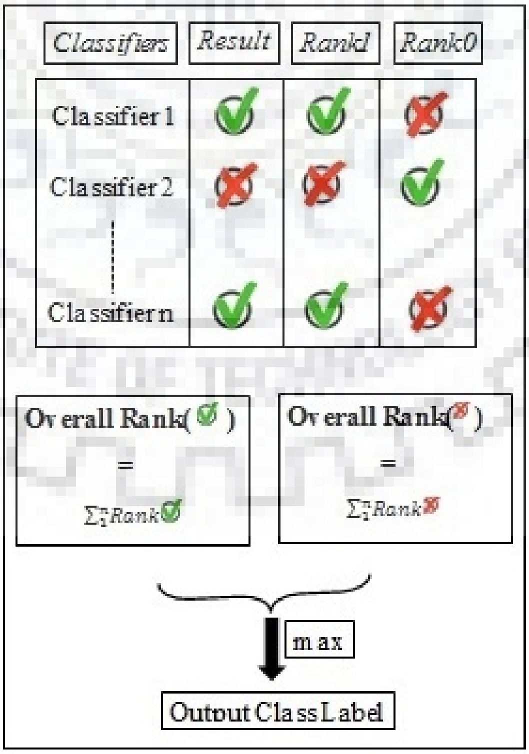


FIGURE 3.6: Borda Count Combination of classifiers based on their ranks.

mean, standard deviation, RMS and power has been performed as shown in Figure 3.7. So, a total of 8 different classification results have been combined on the basis of Borda count approach to produce better results.

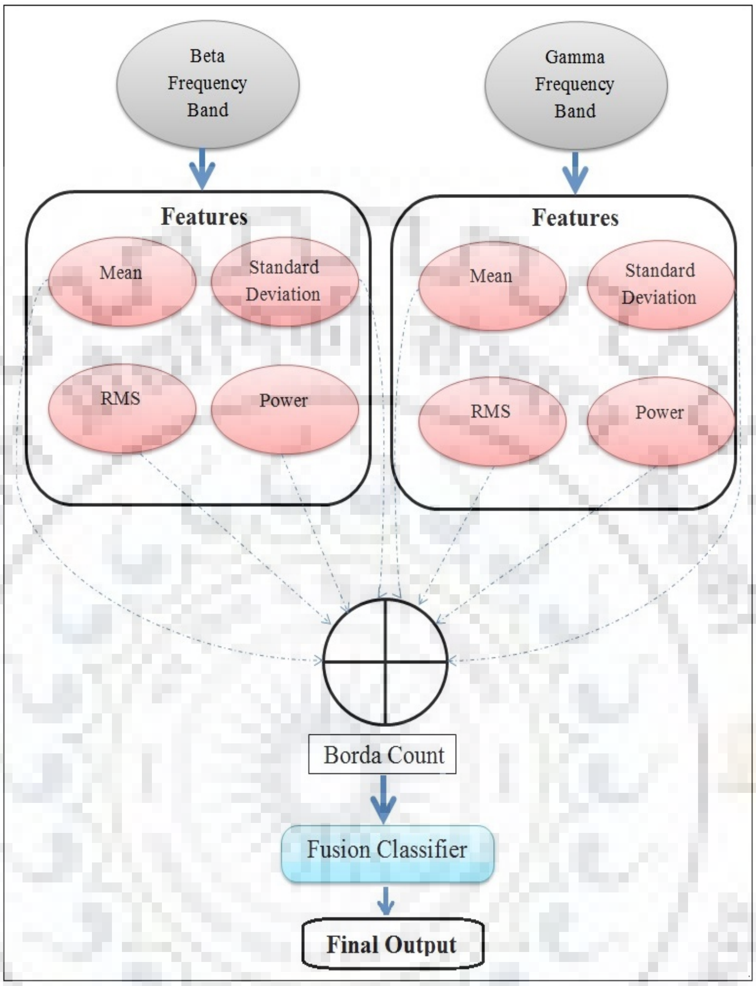


FIGURE 3.7: Block Diagram showing the Borda Count combination of different classification results.

## 3.2 Gaze based Temporal Data Analysis for Word Familiarity Prediction

To analyze the timestamp data, the steps in Figure 3.8 are followed. Firstly, the phase of recording of eye gaze occurs which generates the coordinates of gaze in reference to the screen, along with the timestamp data. After that, these coordinates and timestamp data are pre-processed to extract the information about the fixation of gaze at a word for some time. At last, this information is used to classify whether the user is familiar with the given word or not.

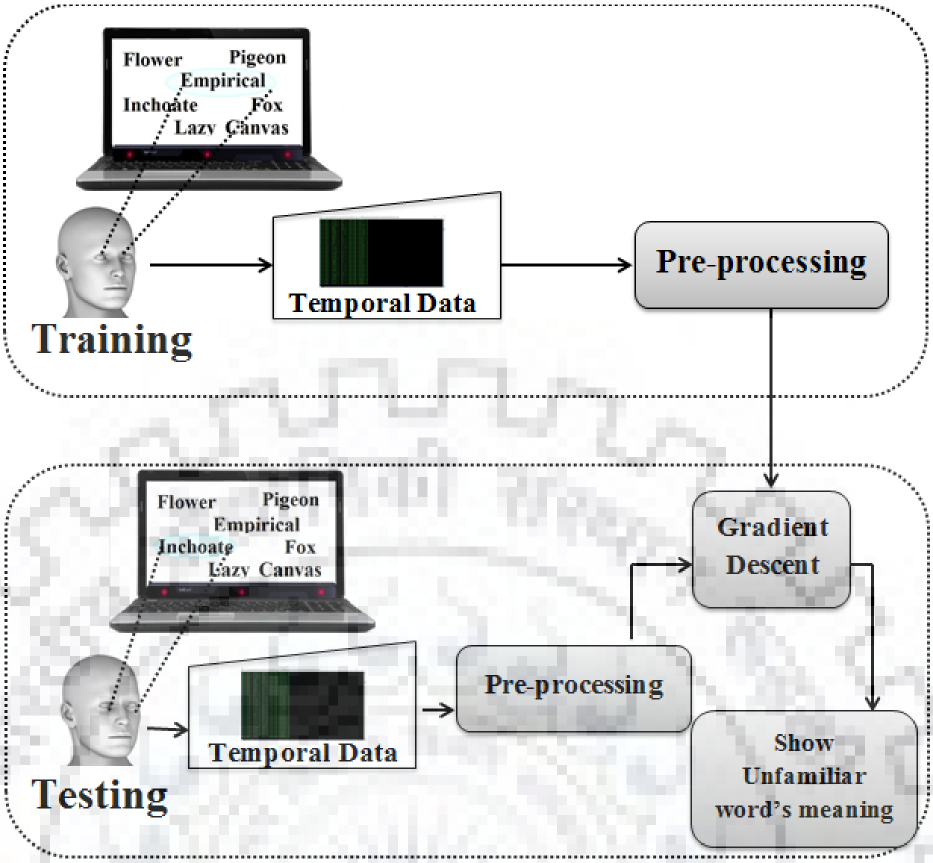


FIGURE 3.8: Block Diagram representing the steps of gaze based temporal data analysis in training & testing phase.

### 3.2.1 System Setup and Data Acquisition

The acquisition of temporal data is performed while users were asked to read words shown on the computer screen. A device, Tobii Eye Tracker 4C has been used to capture the eye gaze of the user on the computer screen. Oculography is a method to record the position and motion of eyes. The various methods to track the movements of the eye include Electro-Oculography, Sceleral Search Coils using the magnetic field, Infrared Oculography and Video Oculography [76]. The Tobii eye tracker device works on Infrafred Oculography which uses near infrared light (NIR) which is invisible and has a wavelength between 700 and 1200 nm. Tobii eye tracker gets disturbed by any other device emitting infrared light in the vicinity of it. It mainly consists of sensors with advanced optical components including both camera and projectors to capture images of eyes with high frame rate and to create reflection pattern of NIR on eyes respectively and algorithms to track the gaze by interpreting the position of eyes and size of pupils in the image stream captured by sensors. The device and its components are shown in Figure 3.9.

Temporal data was collected by placing the tracker at the bottom of the computer screen in front of the user. The tracker emits NIR light and gets the reflected image pattern at the frame

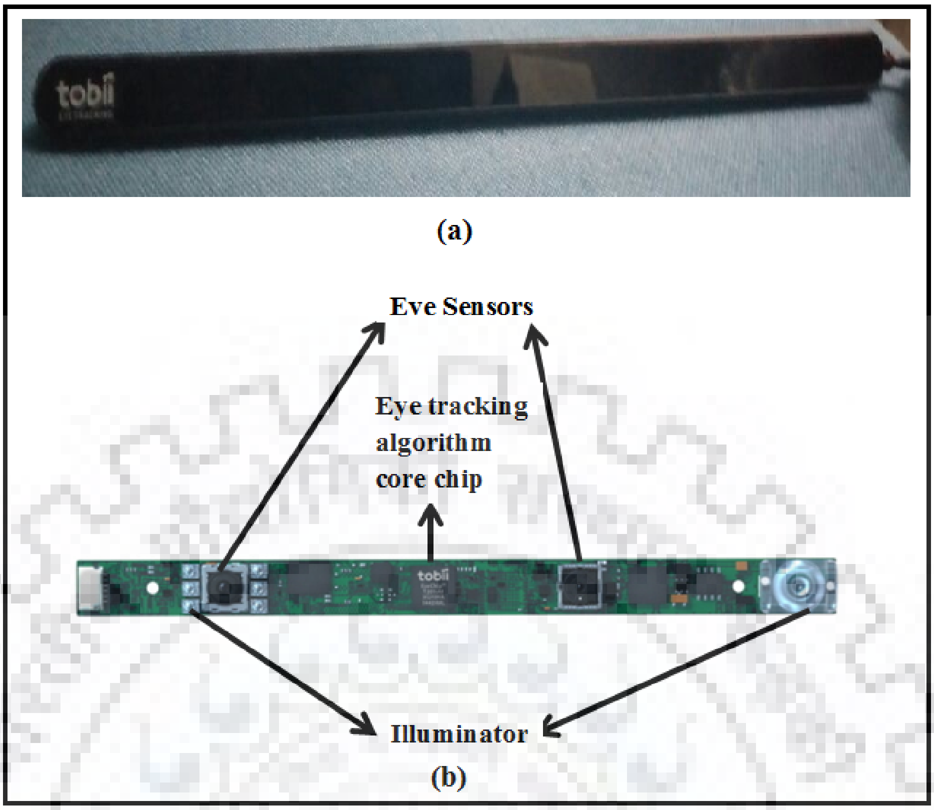


FIGURE 3.9: Eye Tracking Device: (a) Tobii Eye Tracker 4C (b) Components of Tobii Eye Tracker 4C

rate of 10 ms which generates the coordinates of gaze fixed at the screen with their timestamp. Users were asked to read words one after the other and simultaneously, eye gaze tracking was performed. After that, the users were asked whether they were familiar or unfamiliar with those words. The responses of the users were taken in terms of two categories, i.e., familiar or unfamiliar.

### 3.2.2 Temporal Data Pre-processing

The recorded coordinates and timestamps are further analyzed on the basis of coordinates. If two adjacent coordinates are found to be nearly same then their timestamps are matched to measure the time duration of fixation of gaze at that coordinate point. Finally, the dataset is reconstructed to have the coordinates and time duration for which the gaze was stuck to that coordinate.

### 3.2.3 Stochastic Gradient Descent Classifier

Stochastic Gradient descent is an iterative optimization algorithm used to find the values of coefficients(w) so as to minimize a cost function(f). It is used when it is difficult to find the values of coefficients by linear analysis. It finds the threshold value to separate any linear separable data. The equation of stochastic gradient descent is given by Eq.3.18.

$$w = w - \eta \nabla Q_i(w) \tag{3.18}$$

where $Q_i$ is the ith observation in dataset, $\eta$ is the learning rate and w is the parameter which minimizes Q(w). Here, Q(w) can be given by Eq. 3.19

$$Q(w) = \frac{1}{n} \sum_{i=1}^{N} Q_i(w) \tag{3.19}$$

where n is the size of dataset.

The algorithm performs the update specified by Eq. 3.18 for each training example in the training set. A step-by-step flow of the Stochastic Gradient Descent classification scheme is presented in Algorithm 2. In this work, a Stochastic Gradient Descent function with "hinge loss" and "$l^2$ penalty" is used for classification. Hinge loss is preferred as we need "maximum-margin" classification and $l^2$ penalty is used to minimize the sum of square errors. Maximum iterations performed are 5 with the adaptive learning rate of 0.1.

**Algorithm 2** Stochastic Gradient Descent

$\eta$ : Learning rate

w : Minimization coefficient

n : No. of samples

**procedure** SGD(w,X(n),$\eta$)
    Choose an initial vector of parameters 'w' and learning rate $\eta$ .
    **while** approximate minimum is not obtained **do**
        Randomly shuffle examples in the training set.
        **for** i=1 to n **do**
            $w = w - \eta \nabla Q_i(w)$
        **end for**
    **end while**
**end procedure**

## 3.3 Facial Expression Recognition for Word Familiarity Prediction

To perform facial expression detection, the steps in Figure 3.10 are followed. Firstly, the phase of the recording of the video occurs while the user is reading the text. Further, this video is processed to generate image frames in RGB format. After that, they are preprocessed for contrast equalization and then the coordinates of facial features are extracted which are then processed with their adjacent frames to create a feature vector. At last, this feature vector is used to classify whether the user is familiar with the given word or not.

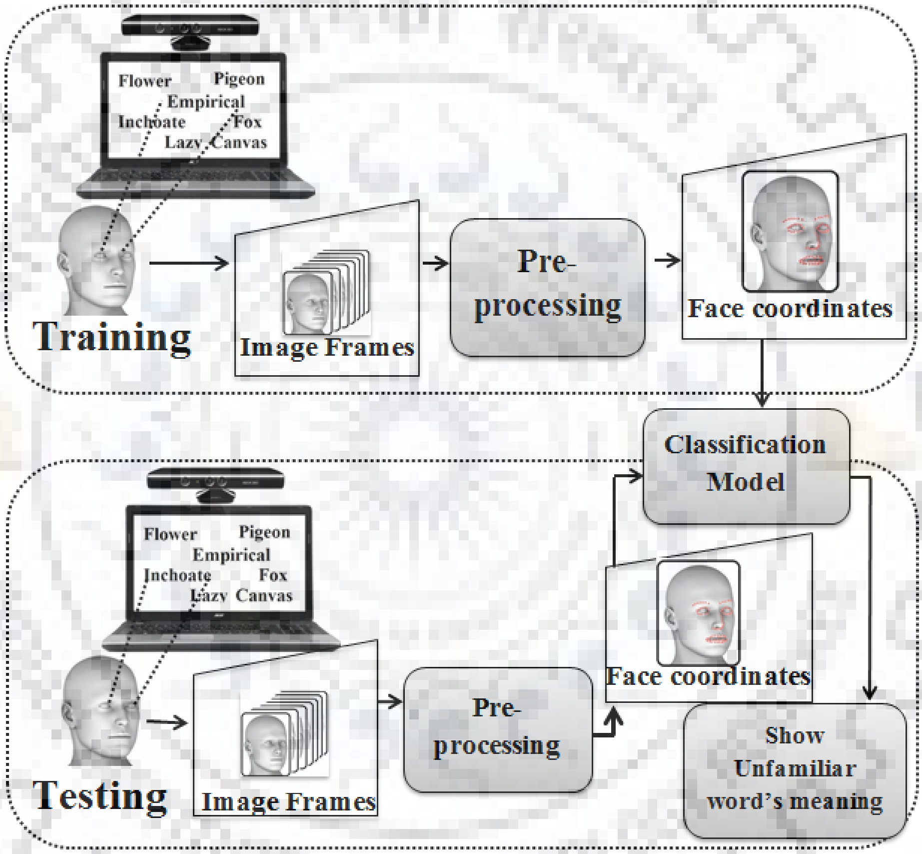


FIGURE 3.10: Block Diagram representing the steps of facial expression detection in training & testing phase.

### 3.3.1 System Setup and Data Acquisition

The detection of facial expressions is performed while users were asked to read words shown on the computer screen. A device, Kinect XBOX 360 has been used to capture the video in terms of RGB and depth images. The Kinect sensor is a horizontal bar-shaped device that can be placed in front of computer screen. It has an RGB camera and a depth sensor. Using them, the Kinect sensor can capture the motion in 3D space. It also has a multi-array microphone which can be used to voice recognition. The depth sensor works using a sensor and an infrared projector. The Kinect sensor outputs the video at a frame rate of approximately 9 Hz to 30 Hz. The RGB image capturing is performed with a resolution of 640 x 480 pixels with Bayer color filter. Similarly, the depth image is captured at the same resolution with 11-bit depth. The range of the sensor is up to 1.2 to 3.5 m. The angular view is at 43° vertically and 57° horizontally. The device for capturing RGB image sequence is shown in Figure 3.11.


FIGURE 3.11: Kinect XBOX 360 Device

RGB image frames were collected by placing the Kinect sensor in front of the computer screen which captures the image frames at a resolution of 640 x 480 pixels and framing rate of 40 frames per second. Users were asked to read words one after the other and simultaneously, their face is captured. After that, the users were asked whether they were familiar or unfamiliar with those words. The responses of the users were taken in terms of two categories, i.e., familiar or unfamiliar.

### 3.3.2 Data Pre-processing

Firstly, the images are preprocessed by converting them to grayscale and then by optimizing the contrast using Contrast Limited Adaptive Histogram Equalization (CLAHE). This step is needed to make all images with equivalent lighting conditions. Adaptive Histogram determines various histograms for different regions of the image and then use them to redistribute lightness or darkness values to the image by enhancing local contrast and edge definitions. Grayscale conversion is required because color images have various components, so histogram equalization has to be performed on each of these components. So, it is better to use uniform color intensity.

#### 3.3.2.1 Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE transforms each pixel based on the transformation function obtained from adjacent regions. The transformation function is calculated from the histograms of the squares enclosing the stated pixel. The transformation function is proportional to the slope of the neighboring Cumulative Distribution Function (CDF). The value when CDF is calculated may exceed given range, so it performs clipping before calculating CDF. So, the slope is limited during the amplification process. This clip value depends on the normalization which depends on the size of the neighborhood. The input to CLAHE is grayscale image and the output obtained is histogram equalized image. In this work, we have used clip limit of 2 and grid size of (8,8). In AHE, if the region has small intensity components then the noise or artifacts get more enhanced in that region. CLAHE limits those artifacts. CLAHE works on extremely small image blocks and then enhances the contrast of each of those blocks. The actual captured RGB image and its corresponding resultant equalized contrast image is shown in Figure 3.12.


FIGURE 3.12: CLAHE Equalization: (a) Captured RGB image (b) Resultant equalized contrast image.

### 3.3.3 Facial Coordinates Extraction

In the proposed system, face detection is performed first which determines whether there is a face in the given image or not. If there is a face then it returns the location and extent of each face in that image. After that facial feature coordinates of specific regions of the face like eyebrows, eyes, nose, and mouth are extracted. In this work, the face is detected using dlib's face detection model and then 68 landmarks points are identified on each detected face using dlib's shape prediction model which takes an object in the image as input and outputs a set of points defining the pose of that object. Figure 3.13 (a) shows the detected face in the image and Figure 3.13 (b) shows the marked coordinates on the face. Further, the difference of corresponding coordinates among the adjacent frames in a series of 120 frames per expression has been calculated.

### 3.3.4 LSTM based Sequence Classification

Lastly, the expression classification is performed. We have mainly performed the expression recognition for two basic expressions: when the user knows the meaning of the word and when the user does not know the meaning of the word. The expression of "confusion" is considered


FIGURE 3.13: Facial Coordinates Extraction: (a) Face detection in image (b) Coordinate marking on detected face.

in the not known word category. LSTM is a very special kind of Recurrent Neural Network (RNN), capable of handling long-term dependencies, i.e., remembering information for long periods of time. In a neural network, there is need to connect previous information to the present task and the gap between the relevant information and the place where it is required is very large, RNNs fails to use the past information. The problem was explored by Hochreiter and Bengio et al. [77]. Thus, when more context is required, LSTM is used. LSTM network is shown in Figure 3.14.

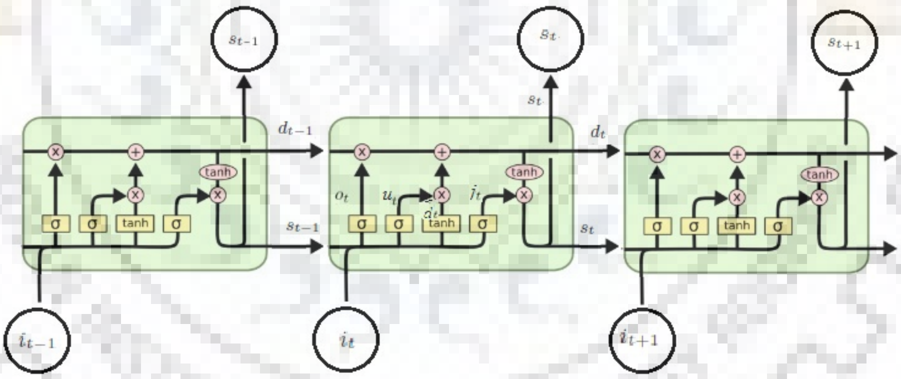

FIGURE 3.14: Long Short Term Memory Network.

LSTMs have the chain like structure like standard RNN, but the repeating module has a different structure. Instead of having a single neural network layer, there are four layers which make LSTM to add or delete information from the cell state. Gates can be a point-wise multiplication operation or a Sigmoid neural network layer. The sigmoid layer decides how much of each component should let through by outputting zero or one. The processing at each step is given by Eq. (3.20-3.25). Eq. (3.20) gives what information to throw away or omit from the cell state. It is a sigmoid function of previous state '$s_{t-1}$', current input '$i_t$' and weight '$W_o$'. Eq. (3.21) decides which values have to be updated which is again a sigmoid function of previous state '$s_{t-1}$', current input '$i_t$' and weight '$W_u$'. Eq. (3.22) creates a vector of new candidate values, '$\tilde{d}_t$', to be added to the state. Eq. (3.23) gives what new information is to be stored in the cell state which is the sum of product of forget information '$o_t$', previous candidate value '$d_{t-1}$' and product of updated value '$u_t$', new candidate value '$\tilde{d}_t$'. Eq. (3.24) decides what is to be given as output which is again a sigmoid function of previous state '$s_{t-1}$', current input '$i_t$' and weight '$W_j$'. Eq. (3.25) gives the final output.

$$o_t = \sigma(W_o.[s_{t-1}, i_t] + b_o) \tag{3.20}$$

$$u_t = \sigma(W_u.[s_{t-1}, i_t] + b_u) \tag{3.21}$$

$$\tilde{d}_t = tanh(W_d.[s_{t-1}, i_t] + b_d) \tag{3.22}$$

$$d_t = o_t * d_{t-1} + u_t * \tilde{d}_t \tag{3.23}$$

$$j_t = \sigma(W_j.[s_{t-1}, i_t] + b_j) \tag{3.24}$$

$$s_t = j_t * tanh(d_t) \tag{3.25}$$

In this work, initially, 50 neurons have been considered in the LSTM layer with the data batch size of 64. The network has been trained for 100 epochs and the loss has been calculated at each step. Adam Optimizer has been used to minimize the cross-entropy. One dense layer with sigmoid activation function has been used. Dropout of 0.8 has been used to prevent overfitting.

# Chapter 4

# Experimental Results

In this chapter, the results of the proposed framework for word familiarity prediction over collected dataset have been presented. The leave-one-out cross-validation approach has been used for the prediction. The results of the performed comparative analysis with different classifiers and their different combinations have also been presented.

## 4.1 Dataset Description

### 4.1.1 EEG Dataset

The 14 channels' EEG signals of 12 participants have been recorded while they are asked to read words on the computer screen as shown in Figure 4.1. All the participants are male and belong to the age group of 18 to 30 years and are students of Indian Institute of Technology, Roorkee, India. 25 different words have been chosen as the dataset with three levels of difficulty, i.e., easy, medium and hard as shown in Figure 4.2. The level of difficulty is based on the frequency of the words used in our everyday life. So, overall 300 (i.e. 12 x 25) EEG signals have been recorded. A feedback response in terms of familiarity/non-familiarity to a word has been collected from the subjects simultaneously with the recording of EEG signals corresponding to each word. Each word has been displayed for 2 seconds with the slide flipping time of 1 second. During the data collection, users have been instructed to provide their correct response.

### 4.1.2 Gaze based Temporal Dataset

The eye gaze screen coordinates with their respective timestamps of 15 participants have been recorded while they are asked to read text consisting of 16 words at a time, on the computer screen as shown in Figure 4.3. The participants include both males and females and belong to

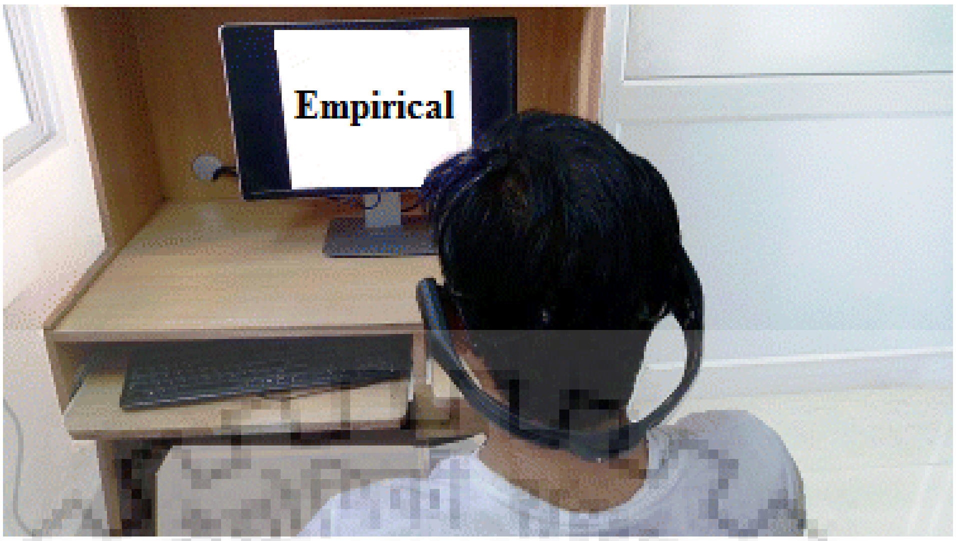


FIGURE 4.1: System setup where the user is reading a word during experiment.

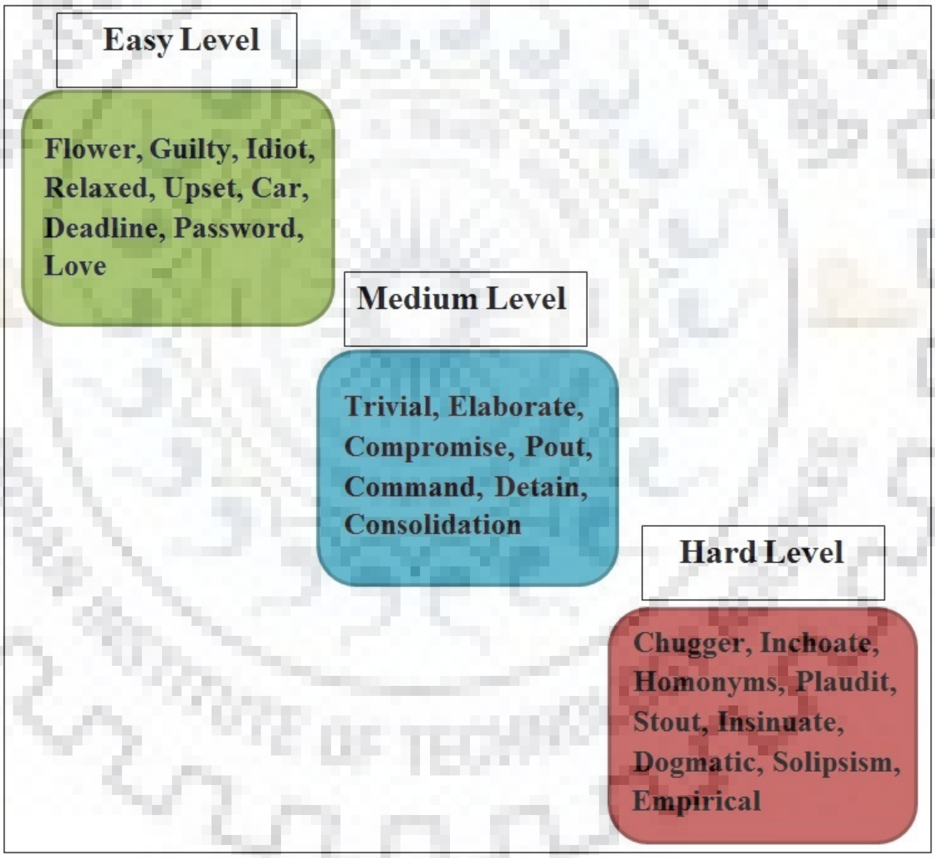


FIGURE 4.2: Words with three levels of difficulty shown to users on the computer screen.

the age group of 18 to 30 years and are students of Indian Institute of Technology, Roorkee, India. A feedback response in terms of familiarity/non-familiarity to a word in the given text has been collected from the subjects simultaneously with the tracking of eye gaze corresponding to each word. During the data collection, users have been instructed to provide their correct response.

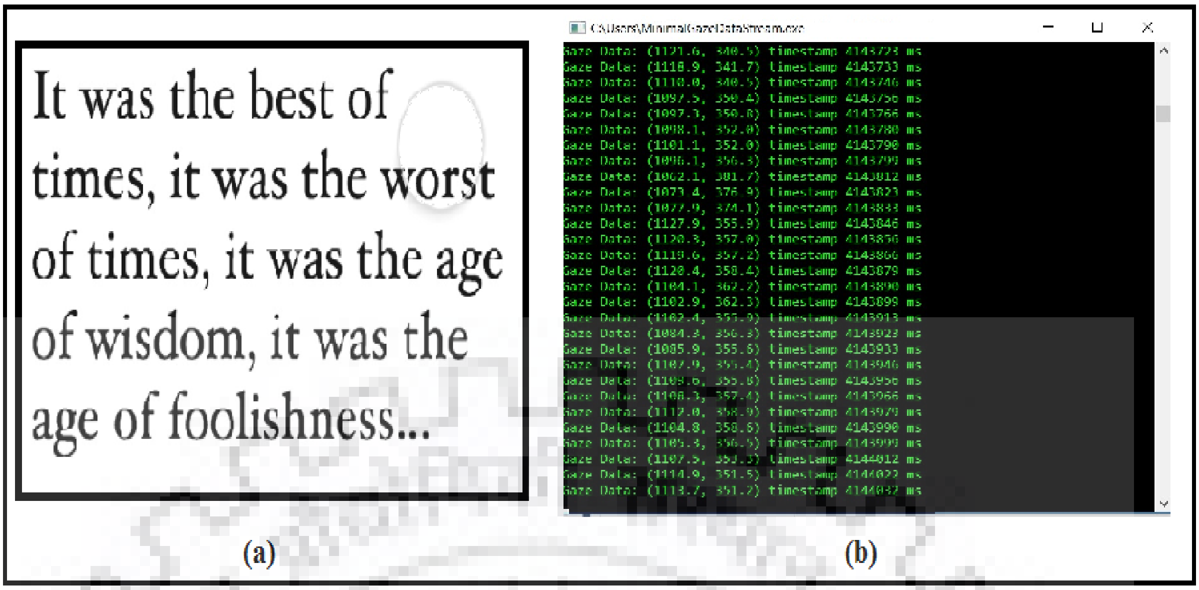


FIGURE 4.3: Temporal Data Analysis: (a)Tracked Gaze on word "worst" (b)Screen Coordinates of tracked gaze and corresponding timestamp.

### 4.1.3 Facial Expression Dataset

The facial expressions of 15 participants have been recorded while they are asked to read words on the computer screen. Figure 4.4 shows the captured RGB images for both unfamiliar and familiar words. The participants include both males and females and belong to the age group of 18 to 30 years and are students of Indian Institute of Technology, Roorkee, India. 25 different words have been chosen as the dataset with three levels of difficulty, i.e., easy, medium and hard. Each word has been displayed for 3 seconds with the slide flipping time of 1 second. The 3 seconds recording is framed at the rate of 40 frames per second. So, overall 120 frames for each word have been recorded. And then 68 (X,Y) cartesian coordinates have been extracted from each of these frames i.e., 136 sized (68 X and 68 Y) feature vector is generated. A total of (120 x 25 x 15) set of feature vectors have been created. Figure 4.5 shows extracted coordinates in various frames when the user is found to be unfamiliar with the word. A feedback response in terms of familiarity/non-familiarity to a word has been collected from the subjects simultaneously with the recording of EEG signals corresponding to each word. During the data collection, users have been instructed to provide their correct response.

## 4.2 Word Familiarity Prediction

Here, the word familiarity results using the statistical and sequential classifiers and their combinations on different datasets have been presented. The classifier model has been built from the EEG samples, sampled at 0.5 second time duration, of each of the frequency bands

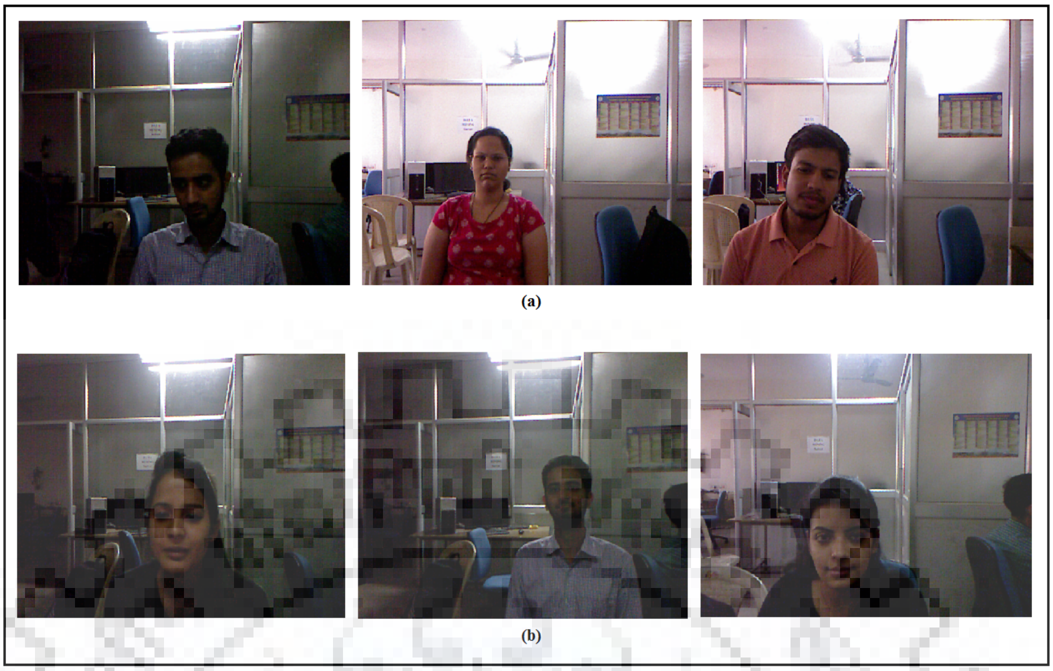


FIGURE 4.4: Facial Expressions: (a) For Unfamiliar words (b) For Familiar words.

FIGURE 4.5: Extracted Facial Coordinates in various frames corresponding to unfamiliar word.

alpha, beta, gamma, delta and theta of 11 subjects and then the trained model has been tested using the EEG signals of rest one subject. Similarly, the classifier model has been built from both temporal dataset and facial features' dataset. This strategy has been used for all the subjects one by one. Finally, the performance of proposed framework is presented in terms of accuracy using Eq. (4.1).

$$Accuracy = \frac{Count\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \tag{4.1}$$

### 4.2.1 On EEG Dataset

#### 4.2.1.1 Word Familiarity Prediction using RF

A group of 100 decision trees has been created with the split at $14^{1/2}$ attributes chosen at random and then each tree has to cast a vote for the most suitable class. The output class has been chosen by the mode of the class's output by individual trees. Each class label is predicted with some confidence value ($0 \leq confidence \leq 1$). Wherever there is confusion and the confidence of both "Familiar" and "Unfamiliar" classes is same i.e. 0.5, the priority has been given to "Unfamiliar" as per the benefit of doubt. The confusion matrix for the Random Forest classifier is shown in Figure 4.6 where 'Target Class' refers to the actual class labels and 'Output Class' refers to the class label predicted by the proposed model, Corresponding percentages of Correct and Incorrect predictions have been shown and total percentages of 'No/Yes' predicted as 'No/Yes' respectively in green color and total percentage of 'No/Yes' predicted as 'Yes/No' respectively in red color have been shown in overall columns whereas in the last blue colored block, the total percentages of correct i.e. 79.7% and incorrect results i.e. 20.3% have been shown. The accuracies over different features of raw data and of different frequency range are shown in Figure 4.7 and Figure 4.8, respectively. Here, raw data means the data acquired from the device's SDK without any preprocessing which involves different filtering and signal smoothing techniques. As evident in the bar chart shown in Figure 4.8, Random Forest over power feature of gamma frequency band is giving the maximum accuracy of 79.7%.

#### 4.2.1.2 Borda Count Combination of Different Features and Frequency bands

The Borda count combination of beta and gamma bands with different features namely, mean, standard deviation, RMS and power has been performed. So, a total of eight different classification results have been combined on the basis of Borda count to produce better results. While combining these results, wherever there is confusion and the overall rank of both "Familiar" and "Unfamiliar" classes is same i.e. 4 out of 8, the priority has been given to "Unfamiliar" as per the benefit of doubt. Otherwise, whichever class has more rank will be the output class. As evident from the plot in Figure 4.9, the accuracy of these eight results' combination is more than their individual results.

The accuracy has been improved to 82% and the confusion matrix for the same is shown in Figure 4.10 where 'Target Class' refers to the actual class labels and 'Output Class' refers to

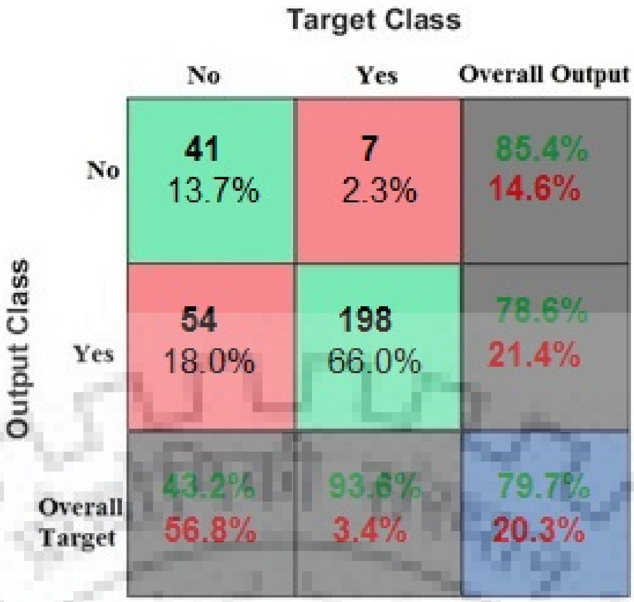


Figure 4.6: Confusion Matrix for the word familiarity prediction using RF classifier.

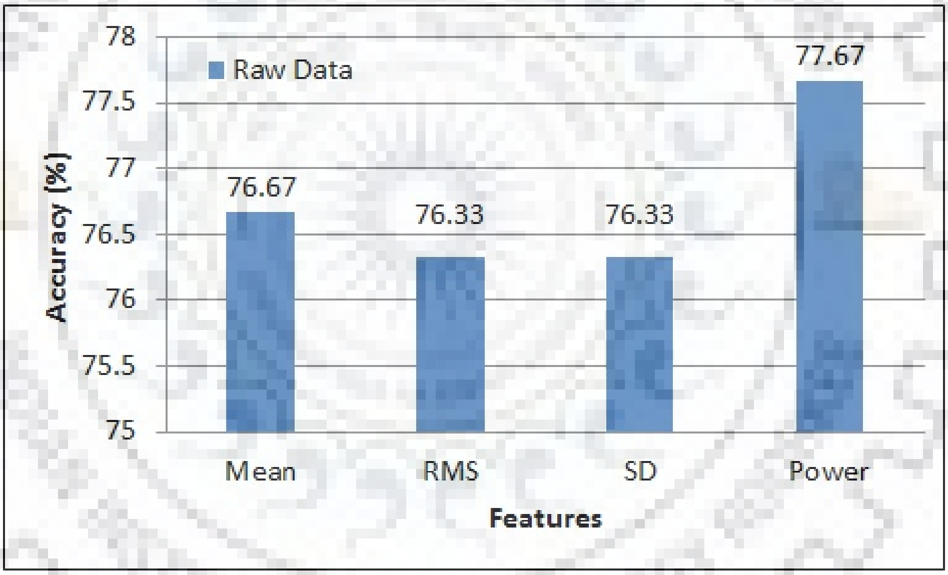


Figure 4.7: Accuracy prediction over different features of raw data.

the class label predicted by the proposed model, Corresponding percentages of Correct and Incorrect predictions have been shown and total percentages of 'No/Yes' predicted as 'No/Yes' respectively in green color and total percentage of 'No/Yes' predicted as 'Yes/No' in red color have been shown in overall columns whereas in the last blue colored block, the total percentages of correct i.e. 82% and incorrect results i.e. 18% have been shown. For different levels of word difficulty, the corresponding accuracy is shown in Figure 4.11 where Easy words are giving the maximum accuracy of 98.2%, Medium words are giving the accuracy of 91.7% and Hard level words are giving the accuracy of 50.9%. This is so because easy words are familiar to every user where as the difficult words may be familiar to some users and unfamiliar to others.

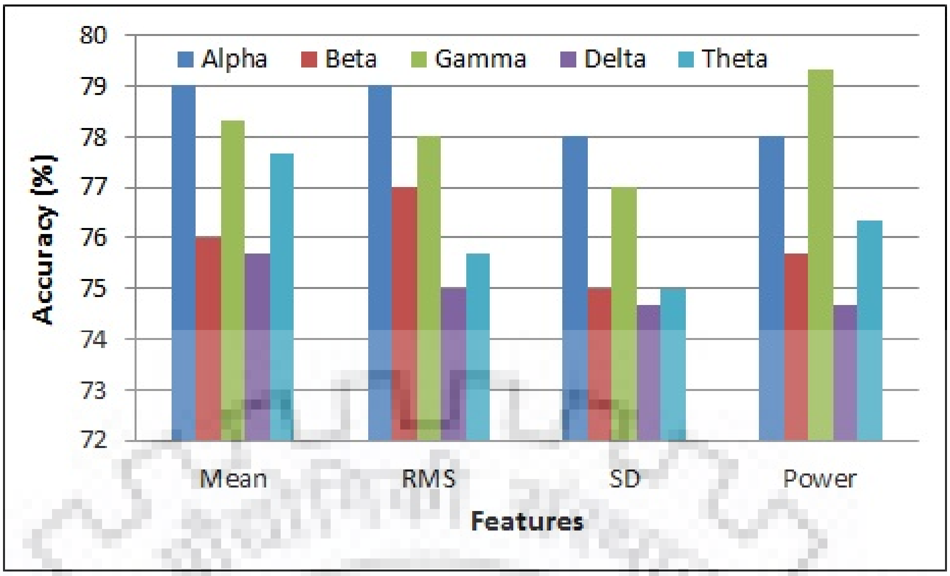


FIGURE 4.8: Accuracy prediction over different features of wavelet decomposed data.

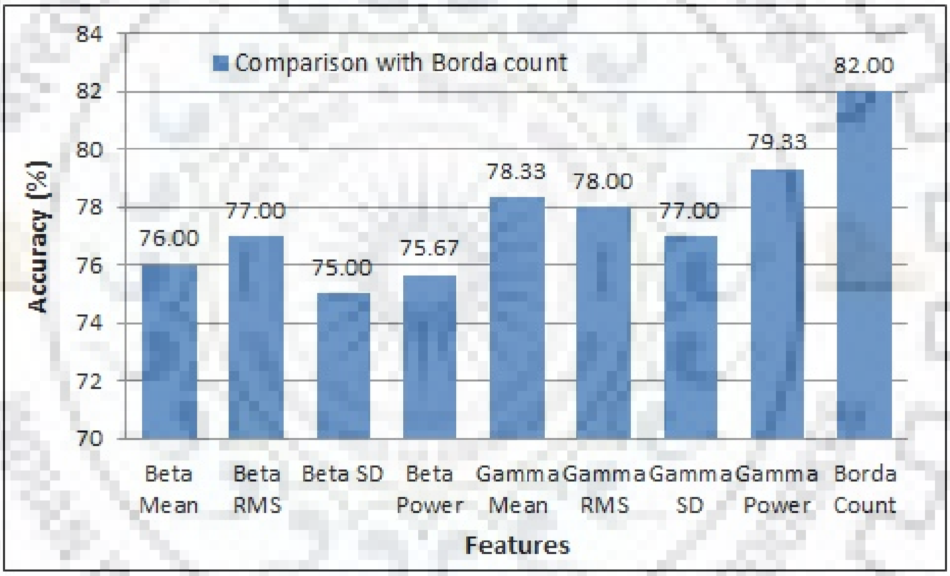


FIGURE 4.9: Borda count combination of frequency bands and features.

Accuracy for every user after applying the combination of classifiers is shown in Figure 4.12.

### 4.2.2 On Gaze based Temporal Dataset

Stochastic Gradient Descent (SGD) function with "hinge loss" and "$l^2$ penalty" with maximum iterations performed of 5 and adaptive learning rate of 0.1 have been applied on the processed temporal data. It has been found that the classifier is classifying the word to be familiar if the time duration of fixation of gaze is below 0.85 seconds and after the time duration of 1.57 seconds, it is classifying the word to be unfamiliar. But for the time gap of 0.85 seconds to


FIGURE 4.10: Confusion matrix for Borda count combination of different features and frequency bands.

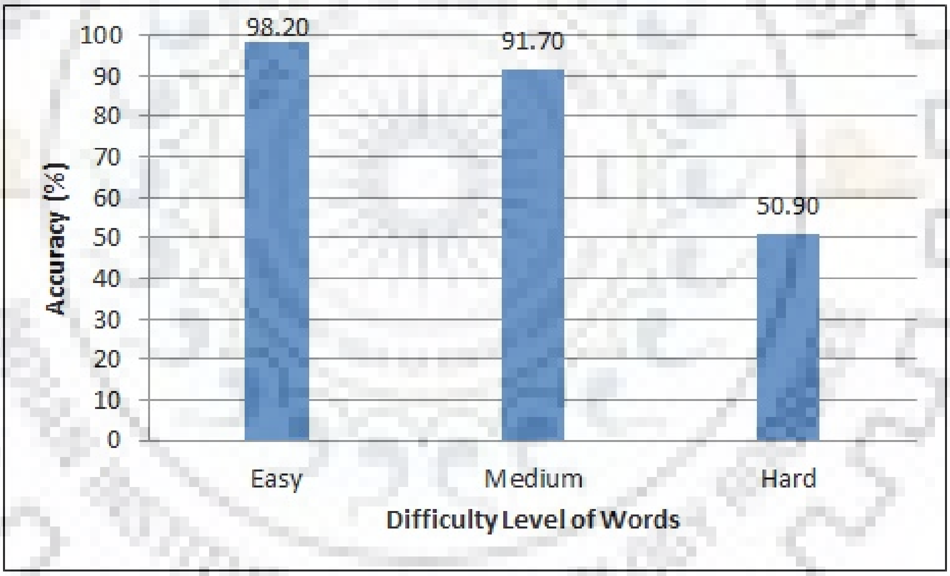


FIGURE 4.11: Accuracy for words at different difficulty levels.

1.57 seconds, there is confusion in mind. The accuracy of the proposed system is found to be 72.8%. The confusion matrix for the same has been shown in Figure 4.13

### 4.2.3 On Facial Expression Dataset

50 neurons have been considered in the LSTM layer with the data batch size of 64. The network has been trained for 100 epochs and the loss has been calculated at each step. Adam Optimizer has been used to minimize the cross-entropy. One dense layer with sigmoid activation function

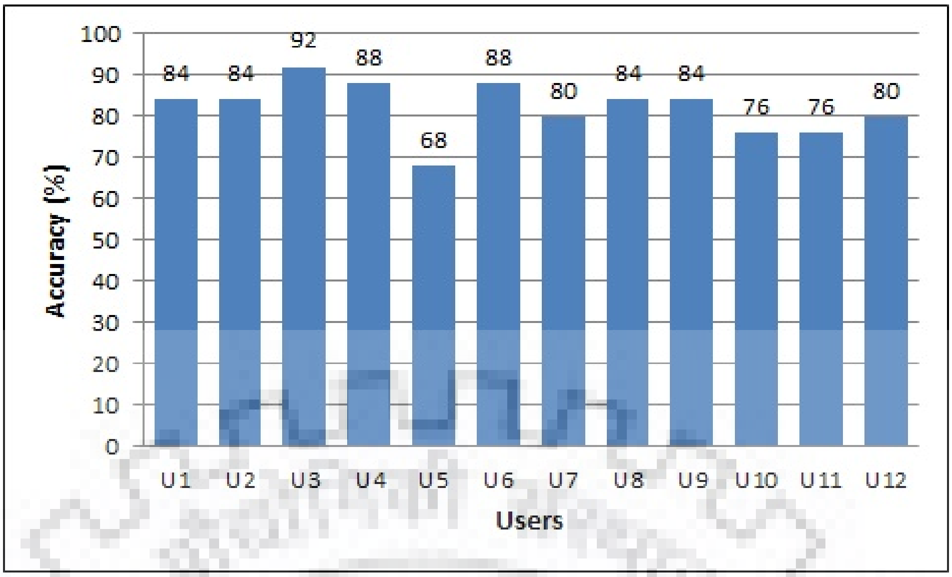


FIGURE 4.12: Accuracy of classifier combination for every user.

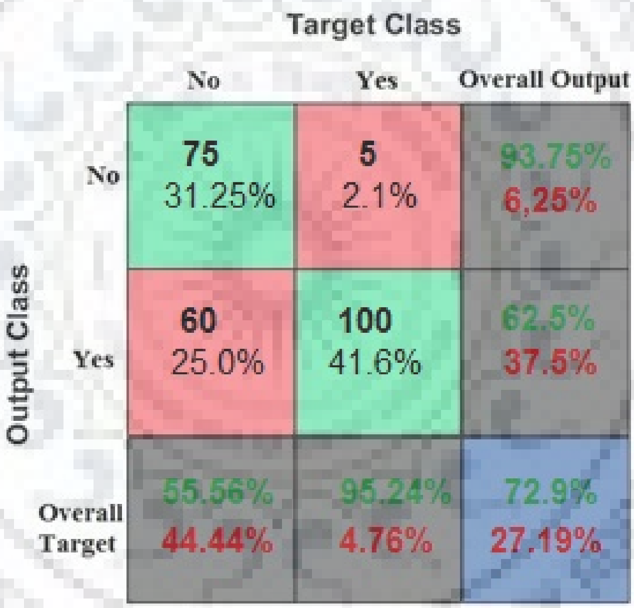


FIGURE 4.13: Confusion Matrix for the word familiarity prediction using SGD classifier.

has been used. Dropout of 0.8 has been used to prevent overfitting. The proposed system is recording an accuracy of 80.3%. The confusion matrix for the same has been shown in Figure 4.14.

## 4.3 Word Meaning Recommendation

If from the collected dataset, the user is found to be unfamiliar with a word then its meaning is shown as a pop window on the computer screen as shown in Figure 4.15 where user 4 is


FIGURE 4.14: Confusion Matrix for the word familiarity prediction using LSTM classifier.

found to be unfamiliar with a word "Empirical" and user 2 is found to be unfamiliar with a word "Inchoate". Therefore, the meanings of these words are shown to corresponding users. For the meanings of words, Wordnet[1] and its synonym sets have been referred.

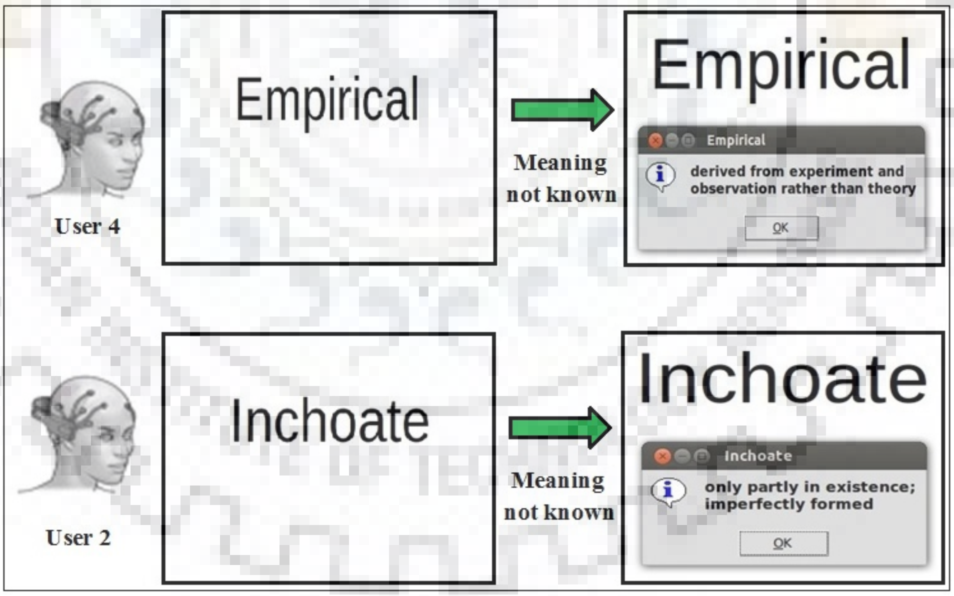


FIGURE 4.15: Meanings of unfamiliar words shown as the pop-up window on computer screen.

[1]Natural Language Toolkit(NLTK) Corpus

## 4.4 Comparative Analysis

1. **Performance Comparison of proposed models for different kinds of datasets:** The word familiarity is predicted one by one on EEG dataset, temporal gazed dataset and facial expression dataset. As evident from the bar chart shown in Figure (4.16), the Borda count combination of beta and gamma bands of EEG dataset with different features namely, mean, standard deviation, RMS, and power has achieved an accuracy of 82%, whereas the Stochastic Gradient Descent (SGD) function over temporal dataset has achieved an accuracy of 72.9% and LSTM over 68 extracted facial feature points has achieved an accuracy of 80.26%.

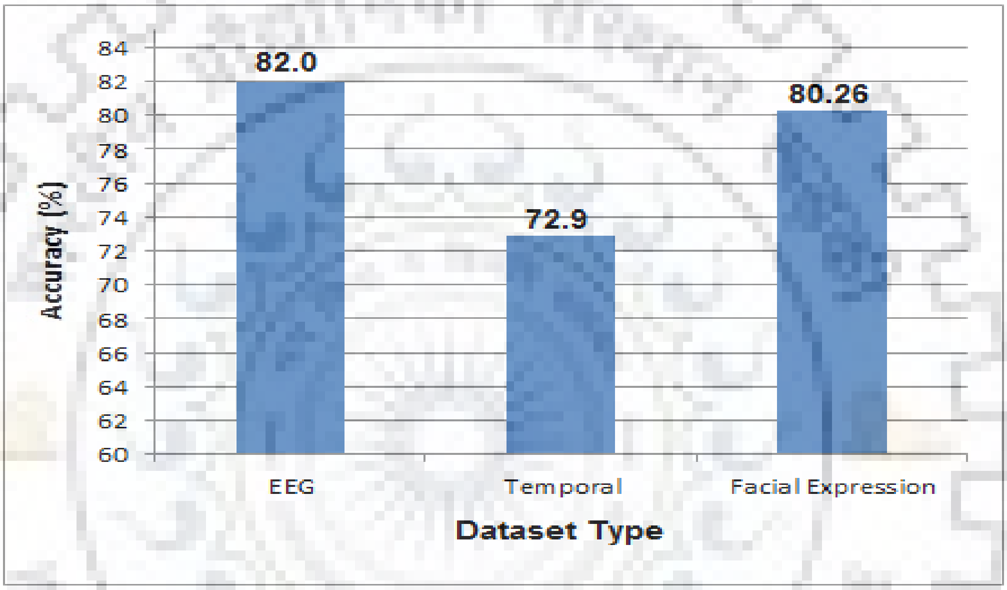


Figure 4.16: Accuracy prediction of proposed models over different datasets.

2. **Performance Comparison with Classifiers on different EEG datasets:** The classification is not only performed on the whole dataset but also to a subset of that dataset. Experiments have been performed to find the overpowering EEG channels according to different brain portions. The performance has been measured on four different brain lobes namely, Frontal lobe that corresponds to AF3, AF4, F3, F4, F7 and F8 electrodes, Parietal lobe that corresponds to P7 and P8 electrodes, Occipital lobe that corresponds to O1 and O2 electrodes and Temporal lobe that corresponds to T7 and T8 electrodes. Experiments have also been performed to find the best sampling rate. The performance has been measured on four different signal time i.e. 0.25 seconds, 0.5 seconds, 1 second and 2 seconds.

   - Classifiers over different brain lobes: The accuracies of Power feature of Gamma frequency band over different brain portions are shown in Figure 4.17. The maximum accuracy of 79.7% is achieved by all 14-channels.

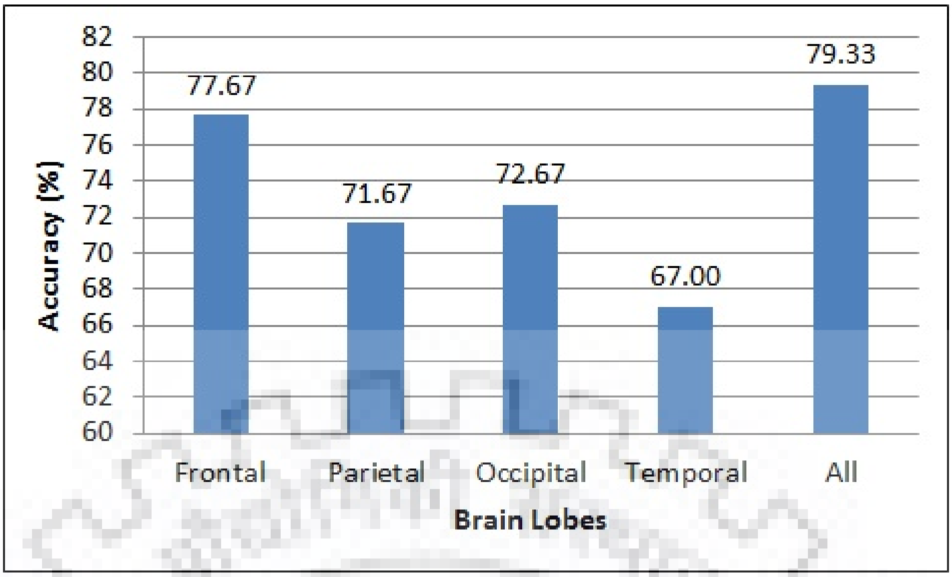


FIGURE 4.17: Accuracy prediction of Power feature of Gamma frequency band over different brain portions.

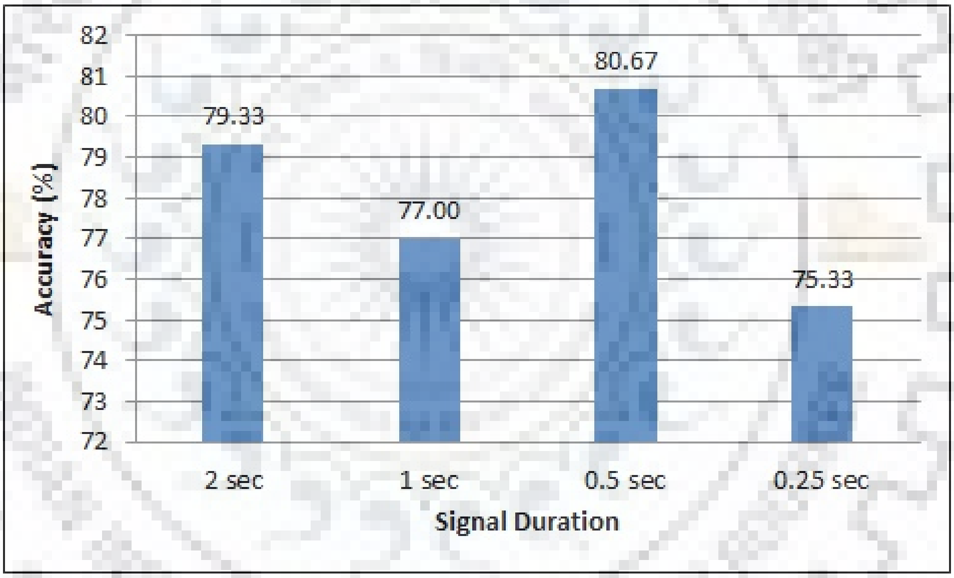


FIGURE 4.18: Accuracy prediction of Power feature of Gamma frequency band over different signal duration.

- Classifiers over different number of samples: The accuracies of Power feature of Gamma frequency band over different sampling rates are shown in Figure 4.18. The maximum accuracy of 80.667% is achieved at 0.5 seconds.

3. **Performance Comparison with Borda Count Combination of Sequential and Statistical Classifier Models over EEG dataset:**
   The Borda count combination of LSTM over beta and gamma bands with the RF over power feature of beta and gamma bands has been performed. So, a total of four different classification results have been combined on the basis of Borda count. While combining

these results, wherever there is confusion and the overall rank of both "Familiar" and "Unfamiliar" classes is same i.e. 2 out of 4, the priority has been given to "Unfamiliar" as per the benefit of doubt. Otherwise, whichever class has more rank will be the output class. As evident from the bar chart shown in Figure (4.19), the accuracy has reduced to 74% where as the accuracy of RF over power feature of gamma band has the accuracy of 79.33%.

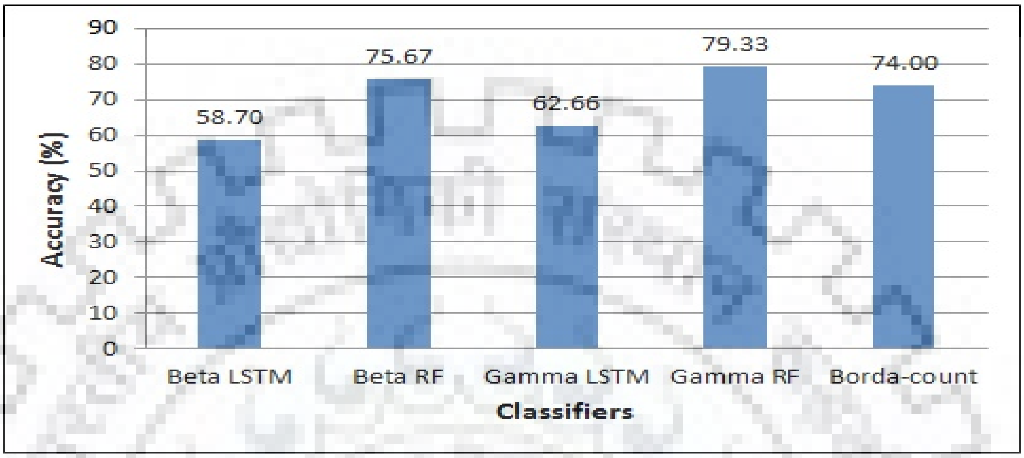


FIGURE 4.19: Borda count combination of beta and gamma using RF and LSTM.

4. **Performance Comparison with HMM Model on facial expression dataset:** The performance of proposed LSTM classifier over extracted facial coordinates is compared with the HMM classifier. The HMM Classifier has been used on the same feature vector where sequential dependencies have been modeled. The experiments have been performed by varying number of states and varying number of gaussian mixture components per state. The highest accuracy of 66.67% has been achieved at 5 states and 128 mixture components. The confusion matrix for the same has been shown in Figure 4.20.

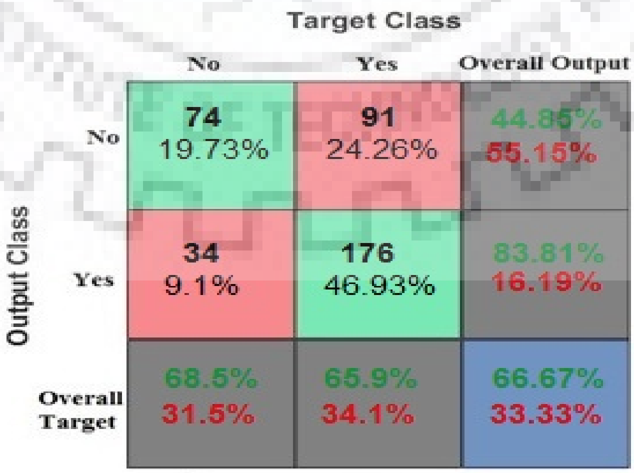


FIGURE 4.20: Confusion Matrix for the word familiarity prediction using HMM classifier.

# Chapter 5

# Conclusion and Future Scope

In this report, a word-familiarity framework to make the reading process fast and understandable using EEG signals, temporal gaze data, and facial expression recognition have been proposed. When a user gets stuck during reading, a change in neural activity and his facial expressions are also recorded along with the time he is taking to read that word. The response of all participants, including both males and females, have been recorded, while they were reading text shown on the computer screen. Next, the EEG signals have been filtered, smoothed, analyzed using Wavelet decomposition technique and modeled using RF classifier, the timestamp data is classified using Stochastic gradient descent classifier and facial expressions are recognized using LSTM classifier after extracting 68 feature points. The maximum accuracy has been recorded over EEG dataset by combining the features extracted from beta and gamma band waves using Borda count approach. If the word is found to be unfamiliar then the meaning of that unfamiliar word is shown as a pop-up window. Earlier, the features and frequency bands had not been combined. The result shows the capability of the proposed framework where Borda count combination of eight classifiers has been done. In our study, we have considered the facial expressions using only 2-dimensional coordinates. The results may improve if we also include the third depth coordinate. We have applied the proposed models separately over the dataset. If we combine them to make a hybrid system then it would be really beneficial in the field of e-learning. The approaches to tackle these problems will be considered in future.

# Chapter 6

# Additional Work

## 6.1 A Survey on Neuromarketing using EEG Signals

Neuromarketing is the application of neuroscience to understand consumer's preference towards marketing. It studies how the change in the market stimulus presentation affect the reaction of the human brain. Neuromarketing is considered as an emerging area of research. Promoters invest around 400 billion dollars every year for advertisement. Therefore, it is required to be performed efficiently by targeting focused market. Traditional approaches consider only the later feedback of the user but not at the time of purchase. The response during purchase provides the exact snapshot which cannot be altered. Various techniques can be used to learn about the decision making of the consumers. These may include brain imaging techniques (fMRI, EEG, SST, TMS) and various biometric sensors. The use of EEG in neuromarketing field is highly promising. EEG detects the sequential changes of brain activity without time delay, which is very important to know both the unconscious reaction and sensory reaction of the customer. Various types of EEG devices are available in the market. Each one has its advantages and disadvantages. Wireless EPOC+ device is highly used nowadays by various researchers. Researchers have conducted the experiment on different age group people showing different categories of products by using different EEG devices. But the neuromarketing field of research is still taken as research area with a warning from consumer protection groups that there should not be any intentional neurological effect of the advertisement on consumers. This chapter discusses about the various neuromarketing strategies, what type of information can be gathered using these strategies, how the marketing stimulus is presented to consumers, what effect it has caused to the consumer in terms of pleasantness and memorization, available machine learning techniques used in this field, various challenges faced, different ethics that must be taken care of and the applications of neuromarketing. It has been suggested that neuromarketing has the capability to improve the effectiveness of advertisements on customers.

## 6.2 Introduction

Neuropsychology studies the relationship between the brain activity of consumer and his behavior. It determines the effectiveness of a particular product to promote sale by linking the choices of consumer and his decision-making process with the marketing research. This is referred as neuromarketing. Neuromarketing is a field to understand consumers' likeliness and choices by applying the neuroscience principles to consumers' response for the marketing stimuli. It is an emerging field which relates all, neuroscience, psychology and marketing with each other [78]. Neuromarketing not only focuses on impact of small change in market stimuli on the sales but also explains how changes in the stimuli presentation affect the reaction of brain which is related to consumer's choices. Promoters invest around 400 billion dollars every year for advertisement [79]. Therefore, it is required to be done efficiently by targeting the correct market.

Traditional research methods focus only on the attitude of consumer towards products by asking the customers to fill the questionnaires which is not exactly related to the actual state of mind at the time of purchase[80]. That is an important factor for failure of various newly launched products in the market. Neuromarketing focuses on the latter by considering the brain signals at the time of purchase. Fig. 6.1 shows how the neuromarketing process workflow goes. Researchers use various techniques like functional Magnetic Resonance Imaging (fMRI), Electroencephalography (EEG), Steady State Topography (SST) and Transcranial Magnetic Stimulation (TMS) to measure changes in brain activity and biometric sensors to measure changes in physical state of customers like respiratory rate, heart rate, facial expression, skin response for emotion analysis and eye tracking for focal attention analysis, to figure out why and how customers make the decisions about the products and which brain areas are responsible for that decision [80].

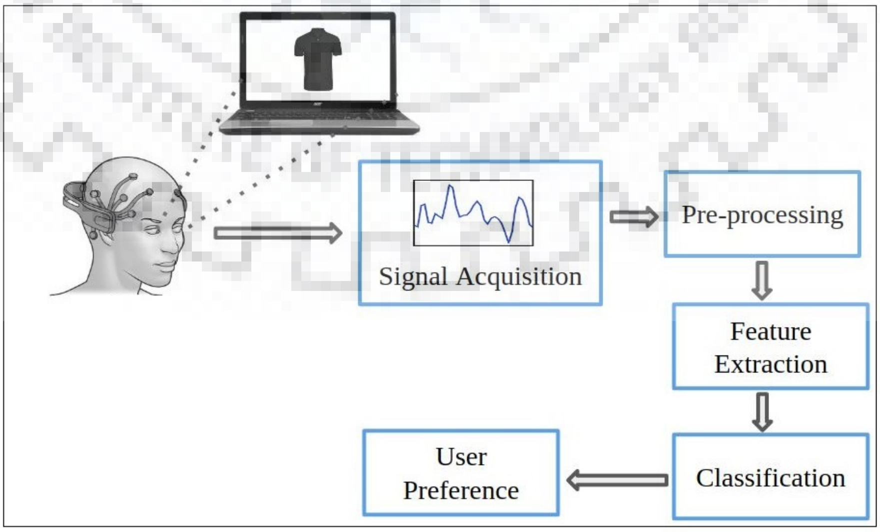


FIGURE 6.1: The process of neuromarketing where EEG response towards a product is recorded and processed to get the user's preference.

Human brain is made up of neurons and those neurons communicate with each other via electrical impulses [81]. EEG signal measurement is a practical way to detect the sequential changes of brain activity without time delay, which is very important to know both the unconscious reaction and sensory reaction of the customer. The neuromarketing field overcomes the challenge of heterogeneity within and across consumer groups which affects consumer preferences and decisions. This heterogeneity may be based on age, gender, various biological factors like hormones and genes, and various physiological factors.
Using neuromarketing, marketers can choose the best strategies like celebrity endorsement or linking with social cause for their product promotion and avoid wastage on inefficient campaign or failed celebrity endorsements. The two scenarios in Fig. 6.2 and 6.3, where EEG signals can be used to get a user's feedback for a product and video, respectively, depict the use of neuromarketing in BCI applications. In literature, researchers have focused on different marketing parameters such as brand perception [82, 83], brand evaluation decision [84–86], brand relationships [87, 88], brand preferences [89–91], pricing [92], product packaging [93, 94], brand naming [95], green consumption [96], store illumination [97], advertisement [98, 99], and new product development [100], etc.

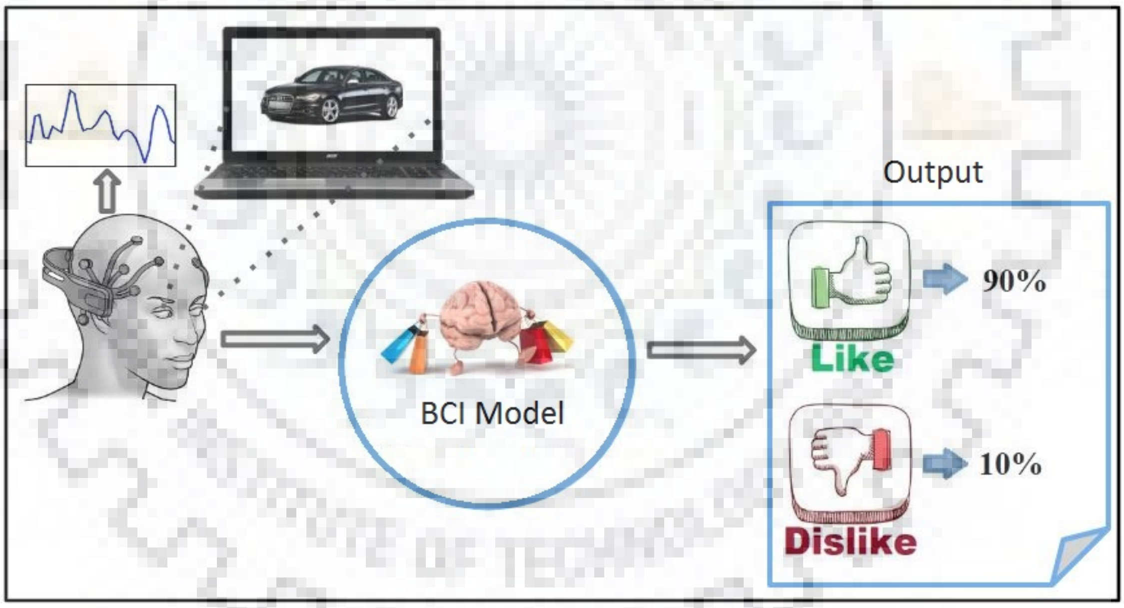


FIGURE 6.2: A scenario of neuromarketing: User is watching the a product on the computer screen and EEG signals are recorded simultaneously. The BCI model predicts whether the person likes or dislike the product by analyzing brain signals.

In this work, we have focused on neuromarketing using EEG because EEG devices are relatively inexpensive, wireless, robust, can be connected with mobile devices, can be used outside the laboratory and are comfortably wearable by users, making this EEG technology of great interest for the evaluating the marketing stimuli. Also, we discuss the effect it has caused to the consumer in terms of pleasantness and memorization, available machine learning techniques

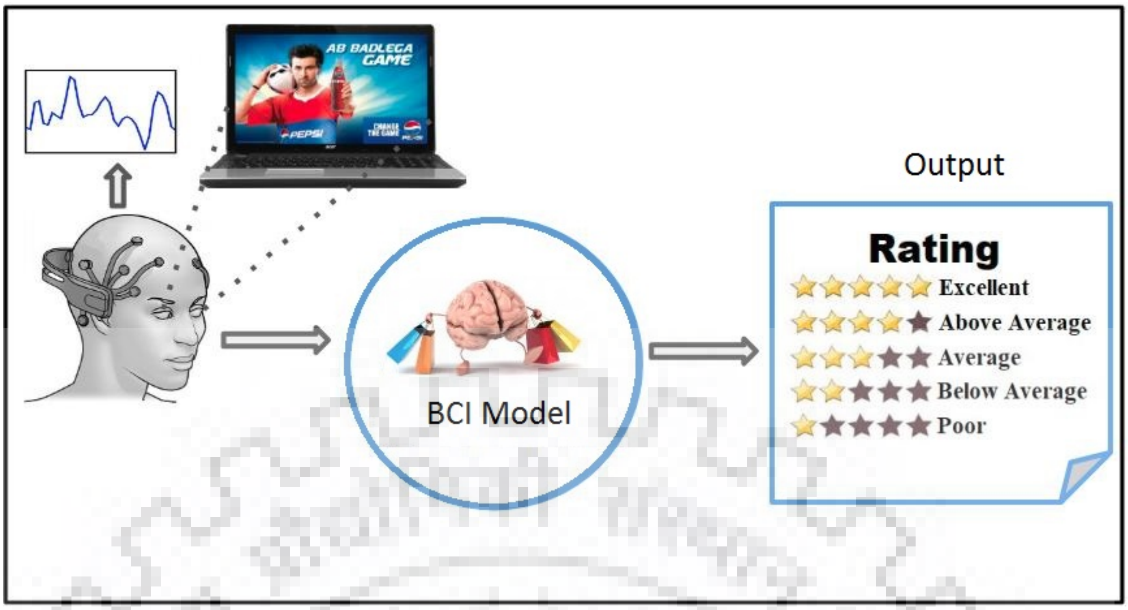


FIGURE 6.3: Another scenario of predicting video ratings using EEG signals. User is watching the video and the BCI model predicts the interest of user in terms of different ratings.

used in this field, various challenges faced, different ethics that must be taken care of and the applications of neuromarketing.

## 6.3 Computational Approaches for Neuromarketing

This section contains the detailed analysis of the techniques used for preprocessing, feature extraction and classification of EEG signals in the recent research work done in the field of neuromarketing using EEG. The Tables 6.4, 6.5 and 6.6 show the related work done in the field of neuromarketing.

### 6.3.1 Pre-processing

Various sources of artifacts like muscular activities, blinking of eyes, and electrical power line noise, etc. come into picture while capturing EEG signals [63]. These artifacts badly affect the useful features in the original signal, so they are required to be separated. Different researchers have used different pre-processing techniques as per their requirements. Table 6.1 shows different filters used in the preprocessing step by various researchers in the field of neuromarketing.

TABLE 6.1: Preprocessing filters used in the field of neuromarketing

| Filter | Details | References |
|---|---|---|
| Savitzky-Golay (S-Golay) filter | frame span = 5 with a quadratic polynomial | Yadava et al. [4] |
| Moving-Average filter | average number of points = 5 | Gauba et al. [3] |
| Notch filter | Frequency = 50 Hz in [101, 102] and 60 Hz in [103] | Teo et al. [101], Murugappan et al. [102], Lee et al. [96] |
| Surface Laplacian filter | – | Murugappan et al. [102] |
| Butterworth bandpass filter | Order = 4 with a cut off frequency between 0.5 Hz and 60 Hz | Murugappan et al. [102], Gupta et al. [104] |
| Elliptical bandpass filter | Order = 10 | Rakshit et al. [2] |
| Common average referencing spatial filter | – | Rakshit et al. [2] |
| Bandpass filter | cut-off frequency between 0.01 and 30 Hz in [6], 0.5 Hz to 40 Hz in [105], 4 to 50 Hz in [96], 0.1-45 Hz in [106] | Bastiaansen et al [6], Khushaba et al. [105], Lee et al. [96], Khushaba et al. [106], |
| FIR1 bandpass filter | 100th degree cut off frequency 1 and 45 Hz | Yilmaz et al. [91] |
| ICA(Independent Component Analysis) | – | Gauba et al. [3], Kawasaki et al. [107], Ohme et al. [108] and Khushaba et al. [106] |
| PCA(Principal Component Analysis) | – | Khushaba et al. [105] |

### 6.3.2 Feature Extraction

EEG signals are time domain signals in unprocessed form, so they are first transformed to frequency domain, otherwise they require highly trained professionals for their investigation. Different researchers have extracted different features for further evaluation as per their requirements. To convert the EEG signals from time domain to frequency domain, researchers have used various techniques and got the frequency band spectrum as Gamma (32-100 Hz), Beta (13-22 Hz), Alpha (8-13 Hz), Theta (4-8 Hz) and Delta (1-4 Hz). Table 6.2 shows various feature extraction techniques used by various researchers in neuromarketing.

Different statistical features have been extracted by researchers. Table 6.3 shows the extracted features used by various researchers in neuromarketing.

TABLE 6.2: Feature Extraction Techniques used in the field of neuromarketing

| **Feature Extraction Techniques** | **References** |
|---|---|
| DB4 (Daubechies 4) wavelet decomposition technique | Yadava et al. [4] |
| Wavelet transformation by using Morlet wavelets with a Gaussian shape | Kawasaki et al. [107] |
| FFT | Khushaba et al. [105], Djamal et al. [109] Lee et al. [96], Ohme et al. [108], Murugappan et al. [102], Khushaba et al. [106] |
| STFT | Rakshit et al. [2] |
| DFT | Teo et al. [101] |

TABLE 6.3: Extracted Features used in the field of neuromarketing

| **Extracted Features** | **References** |
|---|---|
| Statistical Mean | Yadava et al. [4], Gauba et al. [3], Bastiaansen et al. [6], Kawasaki et al. [107] |
| Standard Deviation | Yadava et al. [4] |
| Root-Mean-Square | Yadava et al. [4] |
| Relative Power | Guo et al. [110] |
| Energy | Yadava et al. [4] |
| Power Spectral Density | Rakshit et al. [2], Yilmaz et al. [91],Balconi et al. [78], Lee et al. [96], Ohme et al. [108], Vecchiato et al. [111],Khushaba et al. [105], Khushaba et al. [106] |
| Spectral Centroid and Spectral Energy | Murugappan et al. [102] |

### 6.3.3 Classification and Results

Research work in neuromarketing is focused around the study of customer's preference of images, video advertisements or color visuals and subsequent brain activations in accordance with the preference. In this section, we have listed different classification techniques being used in recent research papers and brain activations for frontal, parietal regions.

#### 6.3.3.1 Classification based on different brain lobes

Frontal region brain activations has been investigated in [108] in which the author has used mean classifier on the alpha power in ipsilateral electrodes. They have aimed to study the reaction of frontal cortex activation to different TV advertisements. The actual results of dominant reactions in only seen in one of the selected advertisements as compared to the expected left hemispheric dominance in accordance with the theoretical methods. The author has presented the detailed analysis of which advertisement's emotional content bought forth higher left frontal activation. An interesting and distinct concept is discussed in [116] where

Table 6.4: Related work done in the field of neuromarketing considering "Product Images" as dataset

| **Author, Year** | **Approach Used** | **Dataset** | **No. of subjects** | **Channels** |
|---|---|---|---|---|
| Khushaba et al. [112], 2012 | Bandpass filter, PCA, FFT, Mutual Information Classifier | Choice sets of images that vary in color and pattern | 18 participants, Aged 25 to 65 years | 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) |
| Khushaba et al. [105], 2012 | Bandpass filter, ICA and DWT for denoising, FFT with zero padding, Mutual Information Classifier | Used objects pictures to choose as screen background | 18 Participants, Aged between 25 and 65 years | 14 channels(AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, 01, 02) |
| Yilmaz et al. [113], 2013 | FIR1 and bandpass filter, Logistic regression, GLM | Powerpoint slide of images containing women's shoes in different styles and colors | 15 participants, No male and 15 females, Aged 20 to 40 | 21 channels; 19 of them used for like/dislike analysis(Fp1, Fp2, A1, A2, O1, O2, C3, Cz, C4, F3, Fz, F4, F7, F8, T3, T4, T5, T6, P3, Pz, P4) |
| Bastiaansen et al. [6], 2016 | Bandpass filter, automatic artifacts removal | photos of the tourist destination Bruges | 32 participants, 8 male and 24 females, Aged 18 to 26 | 61 electrodes |
| Yadava et al. [4], 2017 | S-Golay filter, DB4 wavelet decomposition, HMM Classifier | 14 different product images with 3 varities of each | 40 participants, 25 male and 15 females, Aged 18 to 38 | 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) |

the authors try to understand the true impact of mobile applications the brand perception and popularity. With focus on the pre-frontal cortex of brain to understand left-right alpha asymmetry with respect to the subjects' emotional response as recorded in the feedbacks that were filled before and after the experiment. As a result of the experiments the authors emphasize on the importance of clarity and simple interface of the application for better user experience. Also the results present that excessive browsing in the application leads to a negative emotional engagement possibly due to complicated interface that spoils the user experience. Another study of frontal brain waves has been done in [96] where they have introduced the concept of Green Consumers. The consumers who try to choose environment-friendly products so as to fulfill their economic responsibility are called green consumers and a study to identify green consumers is done by Lee et al. Lee et al. [96] try to identify green consumers by finding differences in the frontal theta brain waves. General Linear Model (GLM) for multivariate analysis has been used for analyzing the frontal brain waves and the results find a significant difference between the frontal theta activations of the green consumers and non-green consumers.

Apart from focus on frontal brain region, some researchers have focused on complete brain activations with respect to human preferences and the inter-dependencies between these regions. One such study is done in [106] where the authors investigate the psychological process of

Table 6.5: Related work done in the field of neuromarketing considering "Advertisement Video" as dataset

| Author, Year | Approach Used | Dataset | No. of subjects | Channels |
|---|---|---|---|---|
| Vecchiato et al. [114], 2010 | Average classifier | Different commercial video advertisements about a naturalistic documentary | A mannequin as subject | Brain Amp (61 channel system) |
| Ohme et al. [108], 2010 | ICA, FFT, Mean classifier | 3 Video advertisements from same product | 45 Participants, 21 male and 24 females, Aged 26 to 45 | 16-channel |
| Lee et al. [96], 2013 | 60 Hz Notch filter, Bandpass filter, FFT, General Linear Model (GLM) | Written description of products with their prices without visual depiction of the product | 19 university students, 12 male and 7 females, Mean age 23.4 | Niteen channel(Fp1, Fp2, A1, A2, O1, O2, C3, Cz, C4, F3, Fz, F4, F7, F8, T3, T4, T5, T6, P3, Pz, P4) |
| Murugappan et al. [102], 2014 | 50 Hz Notch filter, Butterworth 4th order bandpass Filter, Surface Laplacian filter, FFT, KNN, Probabilistic Neural Network(PNN) | Video clips of four Malaysian automotive brands | 12 Participants, 9 male and 3 females, Aged 22 to 24 | 14 channels(AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, 01, 02) |
| Gupta et al. [104], 2017 | Butterworth 4th order bandpass filter | Video clips of 4 soap brands, namely, Lux, Pears, Dove and Cinthol | 18 subjects, 9 male and 9 females, Aged 22 to 24 years | 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) |
| Esmeralda et al. [109], 2017 | FFT with windowing, non-linear SVM | TV Advertisements | 30 subjects, Aged 20 to 25 years | 4 channels (AF3, AF4, T7, and T8) |
| Gauba et al. [3], 2017 | Moving Average filter, ICA, Random Forest Regression | Video advertisements from different promotional categories(home, shopping, sports, automobiles) | 25 participants, Aged 20 to 42 years | 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) |

decision making by the participants with focus on the different brain regions' cortical activity and their inter-dependencies using mutual information analysis.

#### 6.3.3.2 Classification based on User preferences

While most of the research in neuromarketing and EEG is focused on like/dislike of the consumer, the authors here focus on the qualitative features of the product that result in the subject taking a particular decision. The results show that better cognitive processing was initiated by some particular attributes of the crackers' shape, topping or flavor. A combination of two non linear classifiers namely Probabilistic Neural Network (PNN) and k-Nearest Neighbor (KNN) is used in [102] for understanding the objective of participant's decision making

Table 6.6: Related work done in the field of neuromarketing considering "Color Visuals and 3D Virtual Products" as dataset

| Author, Year | Approach Used | Dataset | No. of subjects | Channels |
|---|---|---|---|---|
| Kawasaki et al. [107], 2012 | ICA, Wavelet Transformation, Mean classifier | Color visuals, choose color from 2 colors presented simultaneously | 19 participants, 11 male and 8 females, Aged 18 to 27 years | 60 electrodes |
| Guo et al. [115], 2013 | Adapted Collaborative Filtering for making recommendation on basis of EEG ratings | 3D virtual website where the user can easily interact with the interface | – | 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) |
| Rakshit et al. [2], 2016 | elliptical bandpass filter of order 10 and Common average referencing spatial filter, DFT, SVM, T1FS, BPTT Neural Network | visual stimuli consisting of four colors (Red, Yellow, Green, Blue) and each color appearing randomly on the screen | 7 subjects, 4 male and 3 females, Aged 22 to 30 years | 10 channels (F3; F4; Fz; P3,Pz; P4; 01; 02; T7; T8) |
| Teo et al. [101], 2017 | 50 Hz Notch filter, Automatic Artifacts removal, STFT, Deep Neural Network | 3D visual jewellery type objects stimul | 16 subjects, 8 male and 8 females, Mean age 22.44 | 9 channels (POz, Fz, Cz, C3, C4, F3, F4, P3 and P4) |

process. The authors intend to study the human behavior on basis of spectral features of alpha wave while purchasing marketing products.
Kawaski et al. [107] have studied the impact of consumer's color preference on the visual attention related section of the brain in to understand the brain activations and oscillatory activity between the left and right electrodes while the consumer focused and preferred one color over the other. Using mean classifier across single trials, nonparametric Wilcoxon signed rank test has been used for statistical analysis and the difference in alpha and theta waves as the consumer focused on the preferred color, without selecting the color has also been demonstrated by Kawaski wt al. The results show that the theta amplitude is increasing as the preferred color is being attended and selected by the consumer.

#### 6.3.3.3 Comparative Analysis

Comparison between different classifiers or combining different classifiers have been done in recent times so as to understand brain activations better or to find the most discriminative channel or feature that affects the user preference. Yilmaz et al. [113] have investigated about the better indicators of the user's preference of consumer products. The authors have used logistic regression to identify the most discriminative frequencies utilizing GLM for statistical analysis. Along with finding the most discriminative channels, the authors have also studied the timings difference between taking the like decision in female and male participants.

User preference have been studied in [112], where the authors have focused on frontal spectral activations of the brain while the subjects were recording their preferences. Khushaba et al. [112] have used mutual information measure to investigate left-to-right and front-to-back hemisphere differences. Also the authors have used eye tracker to record the eye placement on all the images presented to the subject while they clicked the most preferred image for their computers. Similar study has been done in [105], the authors have used same classifier to find out that theta bands are more relevant when extracted from symmetric occipital, frontal and parietal regions considering the information exchange between the right and left hemisphere while beta bands dominating the temporal and occipital regions and alpha band waves domineered in the parietal and frontal regions of the brain. Comparison of Interval-Type-II fuzzy classifier and other standard classifiers namely Support Vector Machine (SVM), Type 1 Fuzzy System (T1FS), Backpropogation Through time Neural Network (BPTT-NN) has been performed in [2]. The authors investigate the cognitive bias of different colors and its impact on the subject's mental arousal level which the authors have demonstrated using a brain activation map to show each colors' associated mental and emotion state.

Alpha wave and theta wave spectral feature analysis has been done in [104] for human behavior analysis on marketing stimulus. Mean value classifier have been used and then the power of mean value is determined for each of the four soap bands that the authors have used for this research.

Comparison of accuracy in frontal, parietal, occipital and temporal brain lobes in done in [4]. Yadav et al. have used multiple classifiers, namely Hidden Markov Model (HMM), SVM, Artificial Neural Network (ANN) and Random Forest (RF) to compare the performance of their proposed framework. Also, the effect of gender and age features on choice prediction have also been included. RF, Decision Trees and Linear Regression classifiers have been implemented in [3].

## 6.4 Datasets

The process of collecting EEG signals is very time consuming as each participant's recording of signals can take a significant amount of time. Further, it requires careful preparation of the environment where the signals are to be recorded as any noise or interference during the data collection may lead to erroneous results. Therefore, the lab where the EEG signals are to be collected should be insulated from outside noise and there should be very less disturbance while recording signals. Moreover, the researcher should run some initial pilot experiments so as to minimize the chances of mid failure of the experiment.

### 6.4.1 Dataset Description

#### 6.4.1.1 Available Datasets

To analyze and study EEG signals, there are multiple EEG datasets available on line which can be used directly for analysis such as LSW-neuromarketing[1] [117]. Another dataset that can be used for neuromarketing is NAS dataset[2] where the authors have shared their data of Neuro Against Smoking collected by different participants from all over the world. Another EEG dataset is published by Delorme et al. in EEG/ERP free public Dataset[3] [118] [119] that contains EEG data of 14 participants (7 females, 7 males) which is collected using the Neuroscan software. Similar dataset for EEG signals is available in EEG Database[4] that contains data measured using device with 64 electrodes in three version that can be used as per the requirements of the researcher. Multiple datasets relating to EEG signals and their usage in advertisement ratings are available online[5].

#### 6.4.1.2 Device Used

Various devices are available in the market for acquisition of EEG signals. The devices are different in terms of the way data is collected and the number of electrodes used for that. Various available devices and the placement of electrodes according to International 10-20 system have been shown in Fig. 6.4. EEG signals can be collected by placing the electrodes on human scalp which capture the brain activity in terms of the weak electrical potentials generated by the brain.

It has been observed that more is the number of electrodes collecting EEG signals, more accurate are the results. But the number should not be that large that the noisy signals are captured more than the original signals. Also the correct placement of electrodes is a major task. If the electrodes are not placed correctly, even then there can be more noisy signals. Now a days, most of the researchers are using 14 channel Emotiv EPOC+ device, as it has various advantages over the other available devices in the market like, it is easy to handle and use both by researchers and users, as it has compatibility with different operating systems (Windows, Linux, Android, iOS) and is comfortable to be worn by the users as it fits easily over scalp because of its flexible design and also it has salined wet sensors which ensures no use of sticky gels. Also, it is wireless with battery backup of 12 hours with continuous use, so can even be used outside laboratory. The dense array spatial resolution of the device makes the device to capture complete brain signals [81].

[1] `https://old.datahub.io/dataset/lsw-neuromarketing`
[2] `http://www.nmsba.com/neuro-against-smoking/data`
[3] `https://sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html`
[4] `https://archive.ics.uci.edu/ml/datasets/eeg+database`
[5] `https://sites.google.com/site/iitrcsepradeep7`

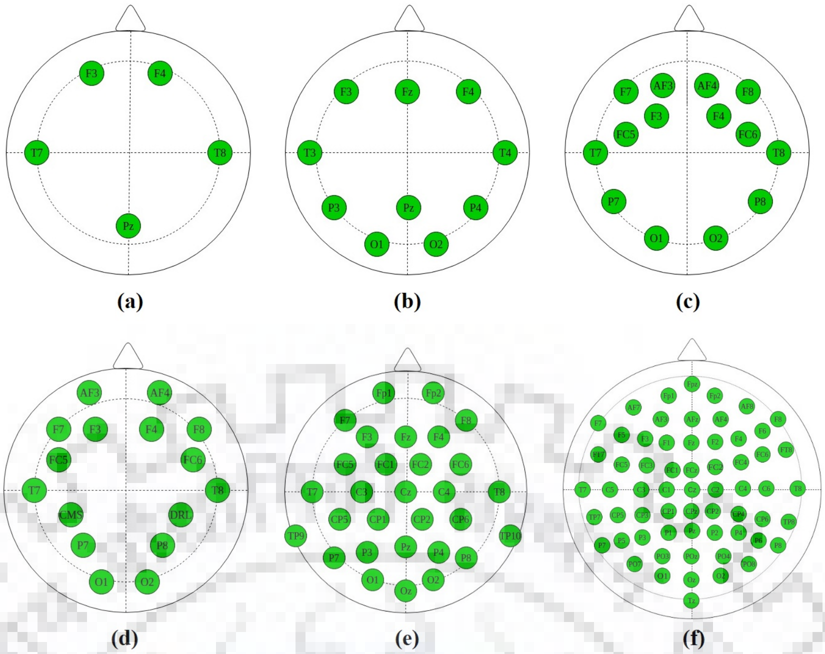


FIGURE 6.4: Placement of EEG electrodes over skull as International 10-20 system: (a) 5 electrodes device [1], (b) 10 electrodes device [2], (c) 14 electrodes device [3], (d) 16 electrodes device [4], (e) 30 electrodes device [5] and (f) 61 electrodes device [6].

#### 6.4.1.3 Number of Participants

Majority of the researchers have conducted the experiment on both males and females of different groups and concluded varying accuracies for different genders and different age groups. Researchers have targeted different age groups depending on the usability of the product type from market point of view. It is good to maintain the heterogeneity within and across consumer groups because different age groups and genders have varying preferences of product types. It has been observed that the middle age group (20-30 years) people have been targeted majorly because they are the active users of majority of the products.

### 6.4.2 Dataset Type

In this section, we discuss about the different types of data that has been used by different authors to show the participants. Most of the authors have used videos containing advertisements to record signals about how the subjects feel after watching those advertisements. Some authors have also used product images or product description for the participants to choose from and to record their subsequent feelings of like/dislike regarding the product. Another form of dataset type is using color visuals i.e. showing a screen filled with particular color to understand objective and unconscious preference of the participant along with subjective evaluation that directly affects decision making. The various kind of datasets and their subtypes have been shown in the Fig. 6.5. The dataset types are further described in detail in the next subpart.

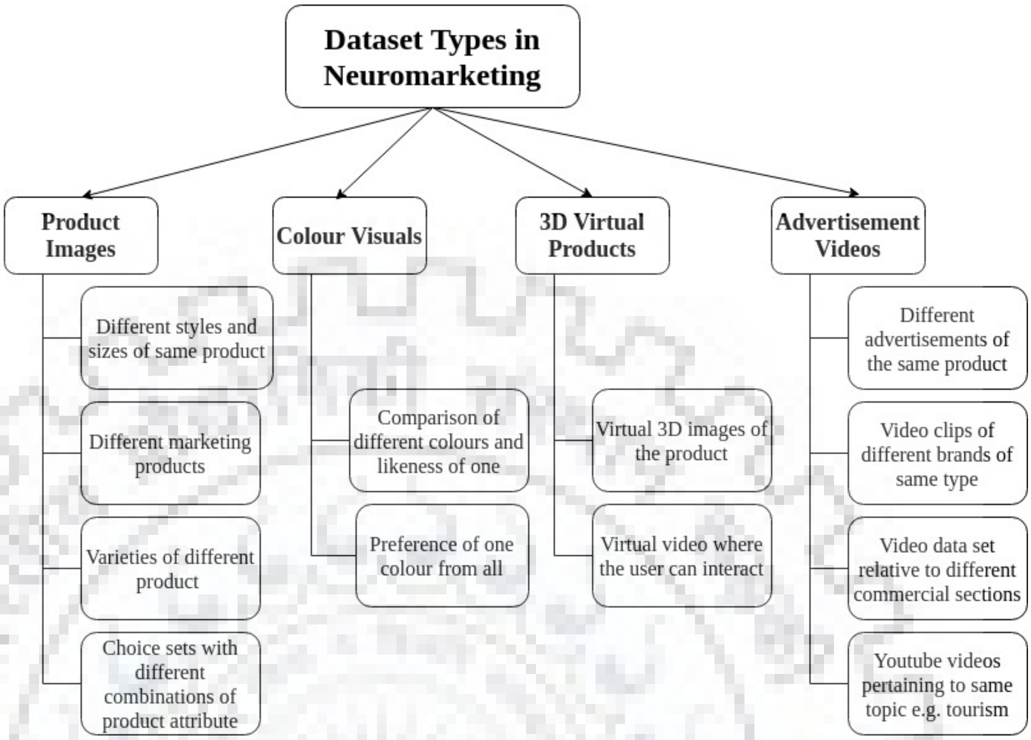


FIGURE 6.5: Various dataset types in Neuromarketing.

## 6.5 Challenges and Ethics

### 6.5.1 Methodological Challenges

In current scenario, neuromarketing lacks the deep insight of how actually the brain operates and how the operation of brain affects specialized human behavior and his decision making patterns. Currently, the techniques used for neuromarketing like EEG, MEG, fMRI, etc. are not able to detect the activity of brain at individual neuron level. Highly advanced methods are required in order to completely understand the relationship. And for this, the knowledge of human brain anatomy and high understanding of the mental abilities, and how the brain processes sensory inputs are required. Advancement in knowledge will help the neuroscientists to make such products and services that would definitely meet the demands of consumers, consciously and unconsciously.

Another big issue of the experiments presented in this chapter is that, there is no common experimental method available yet amongst the researchers who have been working in this neuromarketing field. This field is new for research and needs more time. Neuromarketing researchers are required to share their knowledge and researches. This will give the world, a

better defined experimental paradigms which will further give a shared platform to interpret outcomes and to give the direction for the future research to researchers.
The other challenge is that the current studies using fMRI or so, are mainly based on reverse inference such that the reasoning is performed in the backward direction. This means how some particular mental function is related to the activations in the specific brain areas. However, as evident from the recent studies, the broad interconnected network of the human brain is not just responsible for activating specific brain regions, rather it is responsible for the advanced mental functions.
In this chapter, mainly, EEG based neuromarketing has been targeted. One problem of EEG technique is that the collected electrical brain signals are mainly due to the activity in the cerebral portion of the brain but the electrical activity generated by the deep structures which activates emotional processing is almost impossible to get from usual EEG electrodes. However, high-resolution EEG technology has been developed which provides complete information of the brain activity with a spatial resolution of squared centimeter and time resolution of milliseconds.

### 6.5.2 Ethical Challenges

The ethical issues need to be taken care during research in the neuromarketing field in order to increase the commercial gain. The major ethical concern is that neuromarketing should not allow researchers to create such a marketing campaign that overpowers the free decision making of the user and makes him obsessive about the product. According to Consumer Alert (2003), the development in the field of neuromarketing will influence the users and will be the end for their free will. But actually it is debatable that the research in the field of neuromarketing may actually help in reducing the problems raised by the Commercial Alert (2003) [120].
The potential threats are Predicting Consumer Choice, Influencing Consumer Choice, Transparency, Quality Certification and Privacy [121]. Even after these issues, still neuromarketing field has been chosen by many researchers. This is so because, even if neuromarketing does reach critical effectiveness, the concerns of Commercial Alert (2003) would not be unfounded and may easily be targeted upon. And with these ethics, only good can come by pairing the learning principles of neuroscience and psychology with the commercial principles of economics.

## 6.6 Conclusions

Neuromarketing is an evolving field of research that helps to understand the actual logic about what goes on in the consumer's cognitive mind when they choose a particular product over

the other. This is important because of its direct implication on market as the companies can improve their marketing strategies according to what pleases users and what has adverse effect on consumers' minds. The various neuromarketing strategies, kind of information possible to be gathered with these strategies, how the marketing stimuli is presented to consumers, what effect it has caused to the consumer in terms of pleasantness and memorization, available machine learning techniques used in this field, various challenges faced, different ethics that must be taken care of and the applications of neuromarketing has all been discussed.
As the current scenario witnesses the drift from television to web that requires subsequent change in the marketing strategies which means the advertisement length needs to be shorter. This poses a challenge to show eye catching important content to users in a small time frame which further intensifies the importance of neuromarketing field. It has been observed that due to the potential ethical issues, it is not opted as a field of research by various researchers. However, in future, neuromarketing can reach critical effectiveness using modern devices and techniques.

# Bibliography

[1] Danny Oude Bos et al. Eeg-based emotion recognition. *The Influence of Visual and Auditory Stimuli*, 56(3):1–17, 2006.

[2] Amah Rakshit and Rimita Lahiri. Discriminating Different Color from EEG Signals using Interval-Type 2 Fuzzy Space Classifier (A Neuro-Marketing Study on the Effect of Color to Cognitive State).

[3] Himaanshu Gauba, Pradeep Kumar, Partha Pratim Roy, Priyanka Singh, Debi Prosad Dogra, and Balasubramanian Raman. Prediction of advertisement preference by fusing EEG response and sentiment analysis. *Neural Networks*, 2017. ISSN 18792782. doi: 10.1016/j.neunet.2017.01.013.

[4] Mahendra Yadava, Pradeep Kumar, Rajkumar Saini, Partha Pratim Roy, and Debi Prosad Dogra. Analysis of EEG signals and its application to neuromarketing. *Multimedia Tools and Applications*, 2017. ISSN 15737721. doi: 10.1007/s11042-017-4580-6.

[5] Weiping Yang, Jingjing Yang, Yulin Gao, Xiaoyu Tang, Yanna Ren, Satoshi Takahashi, and Jinglong Wu. Effects of sound frequency on audiovisual integration: An event-related potential study. *PloS one*, 10(9):e0138296, 2015.

[6] Marcel Bastiaansen, Sebastiaan Straatman, Eric Driessen, Ondrej Mitas, Jeroen Stekelenburg, and Lin Wang. My destination in your brain: A novel neuromarketing approach for evaluating the effectiveness of destination marketing. *Journal of Destination Marketing & Management*, 2016. doi: 10.1016/j.jdmm.2016.09.003.

[7] Olaf Hauk, Matthew H Davis, M Ford, Friedemann Pulvermüller, and William D Marslen-Wilson. The time course of visual word recognition as revealed by linear regression analysis of erp data. *Neuroimage*, 30(4):1383–1400, 2006.

[8] Robert Gabriel Lupu and Florina Ungureanu. A survey of eye tracking methods and applications. *Bul Inst Polit Iasi*, pages 71–86, 2013.

[9] Joseph H Goldberg, Mark J Stimson, Marion Lewenstein, Neil Scott, and Anna M Wichansky. Eye tracking in web search tasks: design implications. In *Proceedings of the 2002 symposium on Eye tracking research & applications*, pages 51–58. ACM, 2002.

[10] Robert JK Jacob. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 11–18. ACM, 1990.

[11] Harold H Greene and Keith Rayner. Eye movements and familiarity effects in visual search. *Vision research*, 41(27):3763–3773, 2001.

[12] Guangzhu Xu, Zaifeng Zhang, and Yide Ma. Improving the performance of iris recogniton system using eyelids and eyelashes detection and iris image enhancement. In *Cognitive Informatics, 2006. ICCI 2006. 5th IEEE International Conference on*, volume 2, pages 871–876. IEEE, 2006.

[13] Antoine Picot, Sylvie Charbonnier, and Alice Caplier. Drowsiness detection based on visual signs: blinking analysis based on high frame rate video. In *Instrumentation and Measurement Technology Conference (I2MTC), 2010 IEEE*, pages 801–804. IEEE, 2010.

[14] E Tamil, HM Radzi, MYI Idris, Z Razak, and AM Tamil. Electroencephalogram (eeg) brain wave feature extraction using short time fourier transform. *Faculty of Computer Science and Information Technology, University of Malaya*, 2007.

[15] Mahendra Yadava, Pradeep Kumar, Rajkumar Saini, Partha Pratim Roy, and Debi Prosad Dogra. Analysis of eeg signals and its application to neuromarketing. *Multimedia Tools and Applications*, pages 1–25, 2017.

[16] Himaanshu Gauba, Pradeep Kumar, Partha Pratim Roy, Priyanka Singh, Debi Prosad Dogra, and Balasubramanian Raman. Prediction of advertisement preference by fusing eeg response and sentiment analysis. *Neural Networks*, 2017.

[17] Vairavan Srinivasan, Chikkannan Eswaran, and Natarajan Sriraam. Approximate entropy-based epileptic eeg detection using artificial neural networks. *IEEE Transactions on information Technology in Biomedicine*, 11(3):288–295, 2007.

[18] Christian Stickel, Josef Fink, and Andreas Holzinger. Enhancing universal access–eeg based learnability assessment. *Universal Access in Human-Computer Interaction. Applications and Services*, pages 813–822, 2007.

[19] Erin B McClure, Kayla Pope, Andrea J Hoberman, Daniel S Pine, and Ellen Leibenluft. Facial expression recognition in adolescents with mood and anxiety disorders. *American Journal of Psychiatry*, 160(6):1172–1174, 2003.

[20] Eva Hudlicka. To feel or not to feel: The role of affect in human–computer interaction. *International journal of human-computer studies*, 59(1-2):1–32, 2003.

[21] Zhou Sheng, Lin Zhu-ying, and Dong Wan-xin. The model of e-learning based on affective computing. In *Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on*, volume 3, pages V3–269. IEEE, 2010.

[22] Satori Hachisuka, Kenji Ishida, Takeshi Enya, and Masayoshi Kamijo. Facial expression measurement for detecting driver drowsiness. In *International Conference on Engineering Psychology and Cognitive Ergonomics*, pages 135–144. Springer, 2011.

[23] Magali Dubosson-Torbay, Alexander Osterwalder, and Yves Pigneur. E-business model design, classification, and measurements. *Thunderbird International Business Review*, 44(1):5–23, 2002.

[24] Thierry H Pham and Pierre Philippot. Decoding of facial expression of emotion in criminal psychopaths. *Journal of personality disorders*, 24(4):445–459, 2010.

[25] John G Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE transactions on pattern analysis and machine intelligence*, 15(11): 1148–1161, 1993.

[26] Jyh-Yuan Deng and Feipei Lai. Region-based template deformation and masking for eye-feature extraction and description. *Pattern recognition*, 30(3):403–419, 1997.

[27] Carlos H Morimoto and Marcio RM Mimica. Eye gaze tracking techniques for interactive applications. *Computer vision and image understanding*, 98(1):4–24, 2005.

[28] Qiang Ji and Xiaojie Yang. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-time imaging*, 8(5):357–377, 2002.

[29] AP Burgess and JH Gruzelier. Short duration power changes in the eeg during recognition memory for words and faces. *Psychophysiology*, 37(5):596–606, 2000.

[30] Rachel A Diana, Wijnand Van den Boom, Andrew P Yonelinas, and Charan Ranganath. Erp correlates of source memory: Unitized source information increases familiarity-based retrieval. *Brain Research*, 1367:278–286, 2011.

[31] Sinéad M Rhodes and David I Donaldson. Electrophysiological evidence for the influence of unitization on the processes engaged during episodic retrieval: Enhancing familiarity based remembering. *Neuropsychologia*, 45(2):412–424, 2007.

[32] Sara C Sereno and Keith Rayner. Measuring word recognition in reading: eye movements and event-related potentials. *Trends in cognitive sciences*, 7(11):489–493, 2003.

[33] Shlomo Bentin. Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain and language*, 31(2):308–327, 1987.

[34] Liina Pylkkänen and Alec Marantz. Tracking the time course of word recognition with meg. *Trends in cognitive sciences*, 7(5):187–189, 2003.

[35] Y Zhang, G Silber, and C Kambhamettu. Facial expression driven tutorial system. In *Proceedings of the 6th World Multiconference on Systemics, Cybernetics and Informatics*, pages 287–292, 2002.

[36] May-Ping Loh, Ya-Ping Wong, and Chee-Onn Wong. Facial expression recognition for e-learning systems using gabor wavelet & neural network. In *Advanced Learning Technologies, 2006. Sixth International Conference on*, pages 523–525. IEEE, 2006.

[37] Thomas E Hutchinson, K Preston White, Worthy N Martin, Kelly C Reichert, and Lisa A Frey. Human-computer interaction using eye-gaze input. *IEEE Transactions on systems, man, and cybernetics*, 19(6):1527–1534, 1989.

[38] Linda E Sibert and Robert JK Jacob. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 281–288. ACM, 2000.

[39] JPH Reulen, JT Marcus, D Koops, FR De Vries, G Tiesinga, K Boshuizen, and JE Bos. Precise recording of eye movement: the iris technique part 1. *Medical and Biological Engineering and Computing*, 26(1):20–26, 1988.

[40] Dan Witzner Hansen and Qiang Ji. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE transactions on pattern analysis and machine intelligence*, 32(3): 478–500, 2010.

[41] Zeynep Orman, Abdulkadir Battal, and Erdem Kemer. A study on face, eye detection and gaze estimation. *IJCSES*, 2(3):29–46, 2011.

[42] Paul Smith, Mubarak Shah, and Niels da Vitoria Lobo. Monitoring head/eye motion for driver alertness with one camera. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 4, pages 636–642. IEEE, 2000.

[43] Huchuan Lu, Shipeng Lu, and Gang Yang. Robust eye tracking in video sequence. *Journal of Circuits, Systems, and Computers*, 21(01):1250012, 2012.

[44] Yusuke Sugano, Yasuyuki Matsushita, and Yoichi Sato. Appearance-based gaze estimation using visual saliency. *IEEE transactions on pattern analysis and machine intelligence*, 35(2):329–341, 2013.

[45] Florian Mormann, Juergen Fell, Nikolai Axmacher, Bernd Weber, Klaus Lehnertz, Christian E Elger, and Guillén Fernández. Phase/amplitude reset and theta–gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. *Hippocampus*, 15(7):890–900, 2005.

[46] Phillip J Holcomb and Jonathan Grainger. On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of cognitive neuroscience*, 18(10):1631–1643, 2006.

[47] Masood Mehmood Khan, Michael Ingleby, and Robert D Ward. Automated facial expression classification and affect interpretation using infrared measurement of facial skin temperature variations. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 1(1):91–113, 2006.

[48] Paul Smith, Mubarak Shah, and Niels da Vitoria Lobo. Determining driver visual attention with one camera. *IEEE transactions on intelligent transportation systems*, 4(4): 205–218, 2003.

[49] Tiziana D'Orazio, Marco Leo, Grazia Cicirelli, and Arcangelo Distante. An algorithm for real time eye detection in face images. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 3, pages 278–281. IEEE, 2004.

[50] David Cristinacce, Timothy F Cootes, and Ian M Scott. A multi-stage approach to facial feature detection. In *BMVC*, volume 1, pages 277–286, 2004.

[51] Spiros Ioannou, George Caridakis, Kostas Karpouzis, and Stefanos Kollias. Robust feature detection for facial expression recognition. *Journal on image and video processing*, 2007(2):5–5, 2007.

[52] Jeffrey F Cohn and Takeo Kanade. Use of automated facial image analysis for measurement of emotion expression. *Handbook of emotion elicitation and assessment*, pages 222–238, 2007.

[53] Amr Goneid and Rana El Kaliouby. Facial feature analysis of spontaneous facial expression. In *Proceedings of the 10th International AI Applications Conference*. Citeseer, 2002.

[54] Wenyi Zhao, Rama Chellappa, P Jonathon Phillips, and Azriel Rosenfeld. Face recognition: A literature survey. *ACM computing surveys (CSUR)*, 35(4):399–458, 2003.

[55] Demetri Terzopoulos and Keith Waters. Analysis and synthesis of facial image sequences using physical and anatomical models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6):569–579, 1993.

[56] Irene Kotsia, Ioan Buciu, and Ioannis Pitas. An analysis of facial expression recognition under partial facial image occlusion. *Image and Vision Computing*, 26(7):1052–1067, 2008.

[57] Mark Rosenblum, Yaser Yacoob, and Larry S Davis. Human expression recognition from motion using a radial basis function network architecture. *IEEE transactions on neural networks*, 7(5):1121–1138, 1996.

[58] Zakia Hammal, Laurent Couvreur, Alice Caplier, and Michele Rombaut. Facial expression classification: An approach based on the fusion of facial deformations using the transferable belief model. *International Journal of Approximate Reasoning*, 46(3): 542–567, 2007.

[59] Do Hyoung Kim, Sung Uk Jung, and Myung Jin Chung. Extension of cascaded simple feature based face detection to facial expression recognition. *Pattern Recognition Letters*, 29(11):1621–1631, 2008.

[60] Tao Xiang, Maylor KH Leung, and Siu-Yeung Cho. Expression recognition using fuzzy spatio-temporal modeling. *Pattern Recognition*, 41(1):204–216, 2008.

[61] Grant S Taylor and Christina Schmidt. Empirical evaluation of the emotiv epoc bci headset for the detection of mental actions. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 56, pages 193–197. SAGE Publications Sage CA: Los Angeles, CA, 2012.

[62] Kirill Stytsenko, Evaldas Jablonskis, and Cosima Prahm. Evaluation of consumer eeg device emotiv epoc. In *MEi: CogSci Conference 2011, Ljubljana*, 2011.

[63] Amjed S Al-Fahoum and Ausilah A Al-Fraihat. Methods of eeg signal features extraction using linear analysis in frequency and time-frequency domains. *ISRN neuroscience*, 2014, 2014.

[64] Carrie A Joyce, Irina F Gorodnitsky, and Marta Kutas. Automatic removal of eye movement and blink artifacts from eeg data using blind component separation. *Psychophysiology*, 41(2):313–325, 2004.

[65] Rahul Tandra and Anant Sahai. Snr walls for signal detection. *IEEE Journal of selected topics in Signal Processing*, 2(1):4–17, 2008.

[66] Clemens Brunner, Muhammad Naeem, Robert Leeb, Bernhard Graimann, and Gert Pfurtscheller. Spatial filtering and selection of optimized components in four class motor imagery eeg data using independent components analysis. *Pattern recognition letters*, 28 (8):957–964, 2007.

[67] Charles W Anderson and Zlatko Sijercic. Classification of eeg signals from four subjects during five mental tasks. In *Solving engineering problems with neural networks: proceedings of the conference on engineering applications in neural networks (EANN'96)*, pages 407–414. Turkey, 1996.

[68] Neep Hazarika, Jean Zhu Chen, Ah Chung Tsoi, and Alex Sergejew. Classification of eeg signals using the wavelet transform. *Signal processing*, 59(1):61–72, 1997.

[69] Hafeez Ullah Amin, Aamir Saeed Malik, Rana Fayyaz Ahmad, Nasreen Badruddin, Nidal Kamel, Muhammad Hussain, and Weng-Tink Chooi. Feature extraction and classification for eeg signals using wavelet transform and machine learning techniques. *Australasian physical & engineering sciences in medicine*, 38(1):139–149, 2015.

[70] Pradeep Kumar, Rajkumar Saini, Partha Pratim Roy, and Debi Prosad Dogra. A bio-signal based framework to secure mobile devices. *Journal of Network and Computer Applications*, 2017.

[71] Barjinder Kaur, Dinesh Singh, and Partha Pratim Roy. A novel framework of eeg-based user identification by analyzing music-listening behavior. *Multimedia Tools and Applications*, pages 1–22, 2016.

[72] P Bhuvaneswari and J Satheesh Kumar. Influence of linear features in nonlinear electroencephalography (eeg) signals. *Procedia Computer Science*, 47:229–236, 2015.

[73] Luay Fraiwan, Khaldon Lweesy, Natheer Khasawneh, Heinrich Wenz, and Hartmut Dickhaus. Automated sleep stage identification system based on time–frequency analysis of a single eeg channel and random forest classifier. *Computer methods and programs in biomedicine*, 108(1):10–19, 2012.

[74] Cristian Donos, Matthias Dümpelmann, and Andreas Schulze-Bonhage. Early seizure detection algorithm based on intracranial eeg and random forest classification. *International journal of neural systems*, 25(05):1550023, 2015.

[75] Tin Kam Ho. Random decision forests. In *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, volume 1, pages 278–282. IEEE, 1995.

[76] HR Chennamma and Xiaohui Yuan. A survey on eye-gaze tracking techniques. *arXiv preprint arXiv:1312.6410*, 2013.

[77] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2):157–166, 1994.

[78] Michela Balconi, Beniamino Stumpo, and Federica Leanza. Advertising, brand and neuromarketing or how consumer brain works. *Neuropsychological Trends*, 16(unknown): 15–21, 2014.

[79] Christophe Morin. Neuromarketing: the new science of consumer behavior. *Society*, 48 (2):131–135, 2011.

[80] Sharad Agarwal and Tanusree Dutta. Neuromarketing and consumer neuroscience: current understanding and the way forward. *Decision*, 42(4):457–462, 2015.

[81] Vaishali Khurana, Pradeep Kumar, Rajkumar Saini, and Partha Pratim Roy. Eeg based word familiarity using features and frequency bands combination. *Cognitive Systems Research*, 49:33–48, 2018.

[82] Ab Litt and Baba Shiv. Manipulating basic taste perception to explore how product information affects experience. *Journal of Consumer Psychology*, 22(1):55–66, 2012.

[83] Milica Milosavljevic, Vidhya Navalpakkam, Christof Koch, and Antonio Rangel. Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology*, 22(1):67–74, 2012.

[84] Franz-Rudolf Esch, Thorsten Möll, Bernd Schmitt, Christian E Elger, Carolin Neuhaus, and Bernd Weber. Brands on the brain: Do consumers use declarative information or experienced emotions to evaluate brands? *Journal of Consumer Psychology*, 22(1): 75–85, 2012.

[85] Zachary Estes, Michael Gibbert, Duncan Guest, and David Mazursky. A dual-process model of brand extension: Taxonomic feature-based and thematic relation-based similarity independently drive brand extension evaluation. *Journal of Consumer Psychology*, 22(1):86–101, 2012.

[86] Gad Saad and Eric Stenstrom. Calories, beauty, and ovulation: The effects of the menstrual cycle on food and appearance-related consumption. *Journal of Consumer Psychology*, 22(1):102–113, 2012.

[87] Pankaj Aggarwal and Richard P Larrick. When consumers care about being treated fairly: The interaction of relationship norms and fairness norms. *Journal of Consumer Psychology*, 22(1):114–127, 2012.

[88] Martin Reimann, Raquel Castaño, Judith Zaichkowsky, and Antoine Bechara. How we relate to brands: Psychological and neurophysiological insights into consumer–brand relationships. *Journal of Consumer Psychology*, 22(1):128–142, 2012.

[89] Vinod Venkatraman, John A Clithero, Gavan J Fitzsimons, and Scott A Huettel. New scanner data for brand marketers: how neuroscience can help better understand differences in brand preferences. *Journal of consumer psychology*, 22(1):143–153, 2012.

[90] Gregory S Berns and Sara E Moore. A neural predictor of cultural popularity. *Journal of Consumer Psychology*, 22(1):154–160, 2012.

[91] Bülent Yılmaz, Sümeyye Korkmaz, Dilek Betül Arslan, Evrim Güngör, and Musa H Asyalı. Like/dislike analysis using eeg: determination of most discriminative channels and frequencies. *Computer methods and programs in biomedicine*, 113(2):705–713, 2014.

[92] Hilke Plassmann, John O'Doherty, and Antonio Rangel. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of neuroscience*, 27(37): 9984–9988, 2007.

[93] Martin Reimann, Judith Zaichkowsky, Carolin Neuhaus, Thomas Bender, and Bernd Weber. Aesthetic package design: A behavioral, neural, and psychological investigation. *Journal of Consumer Psychology*, 20(4):431–441, 2010.

[94] Marco Stoll, Sebastian Baecke, and Peter Kenning. What they see is what they get? an fmri-study on neural correlates of attractive packaging. *Journal of Consumer Behaviour*, 7(4-5):342–359, 2008.

[95] Philipp Hillenbrand, Sarael Alcauter, Javier Cervantes, and Fernando Barrios. Better branding: brand names can influence consumer choice. *Journal of Product & Brand Management*, 22(4):300–308, 2013.

[96] Eun-Ju Lee, Gusang Kwon, Hyun Jun Shin, Seungeun Yang, Sukhan Lee, and Minah Suh. The spell of green: Can frontal eeg activations identify green consumers? *Journal of Business Ethics*, 122(3):511–521, 2014.

[97] Jakub Berčík, Elena Horská, WY Wang, Ying-Chun Chen, et al. How can food retailing benefit from neuromarketing research: a case of various parameters of store illumination and consumer response. In *143rd Joint EAAE/AAEA Seminar, March 25-27, 2015, Naples, Italy*, number 202714. European Association of Agricultural Economists, 2015.

[98] Shiree Treleaven-Hassard, Joshua Gold, Steven Bellman, Anika Schweda, Joseph Ciorciari, Christine Critchley, and Duane Varan. Using the p3a to gauge automatic attention to interactive television advertising. *Journal of Economic Psychology*, 31(5):777–784, 2010.

[99] Giovanni Vecchiato, Jlenia Toppi, Laura Astolfi, Fabrizio De Vico Fallani, Febo Cincotti, Donatella Mattia, Francesco Bez, and Fabio Babiloni. Spectral eeg frontal asymmetries correlate with the experienced pleasantness of tv commercial advertisements. *Medical & biological engineering & computing*, 49(5):579–583, 2011.

[100] Dan Ariely and Gregory S Berns. Neuromarketing: the hope and hype of neuroimaging in business. *Nature reviews neuroscience*, 11(4):284, 2010.

[101] Jason Teo, Chew Lin Hou, and James Mountstephens. Deep learning for EEG-Based preference classification. In *AIP Conference Proceedings*, 2017. ISBN 9780735415737. doi: 10.1063/1.5005474.

[102] M. Murugappan, Subbulakshmi Murugappan, Balaganapathy, and Celestin Gerard. Wireless EEG signals based Neuromarketing system using Fast Fourier Transform

(FFT). In *2014 IEEE 10th International Colloquium on Signal Processing and its Applications*, 2014. ISBN 978-1-4799-3091-3. doi: 10.1109/CSPA.2014.6805714.

[103] Eun Ju Lee, Gusang Kwon, Hyun Jun Shin, Seungeun Yang, Sukhan Lee, and Minah Suh. The Spell of Green: Can Frontal EEG Activations Identify Green Consumers? *Journal of Business Ethics*, 2014. ISSN 15730697. doi: 10.1007/s10551-013-1775-2.

[104] Ashutosh Gupta, Richa Shreyam, Ridhi Garg, and Tabassum Sayed. Correlation of Neuromarketing to Neurology. *IOP Conference Series: Materials Science and Engineering*, 2017. ISSN 1757-8981. doi: 10.1088/1757-899X/225/1/012129.

[105] Rami N. Khushaba, Luke Greenacre, Sarath Kodagoda, Jordan Louviere, Sandra Burke, and Gamini Dissanayake. Choice modeling and the brain: A study on the Electroencephalogram (EEG) of preferences. *Expert Systems with Applications*, 2012. ISSN 09574174. doi: 10.1016/j.eswa.2012.04.084.

[106] Rami N. Khushaba, Chelsea Wise, Sarath Kodagoda, Jordan Louviere, Barbara E. Kahn, and Claudia Townsend. Consumer neuroscience: Assessing the brain response to marketing stimuli using electroencephalogram (EEG) and eye tracking. *Expert Systems with Applications*, 2013. ISSN 09574174. doi: 10.1016/j.eswa.2012.12.095.

[107] Masahiro Kawasaki and Yoko Yamaguchi. Effects of subjective preference of colors on attention-related occipital theta oscillations. *NeuroImage*, 2012. ISSN 10538119. doi: 10.1016/j.neuroimage.2011.07.042.

[108] Rafal Ohme, Dorota Reykowska, Dawid Wiener, and Anna Choromanska. Application of frontal EEG asymmetry to advertising research. *Journal of Economic Psychology*, 2010. ISSN 01674870. doi: 10.1016/j.joep.2010.03.008.

[109] Esmeralda C Djamal, Reza Indrawan, Juliyanto Pratama, and Faiza Renaldi. Eeg based neuropsychology of advertising video using fast fourier transform and support vector machine. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 9(3-7):105–109, 2017.

[110] Guibing Guo and Mohamed Elgendi. A new recommender system for 3d e-commerce: An eeg based approach. *Journal of Advanced Management Science*, 1(1):61–65, 2013.

[111] Giovanni Vecchiato, F De Vico Fallani, L Astolfi, J Toppi, F Cincotti, D Mattia, S Salinari, and F Babiloni. The issue of multiple univariate comparisons in the context of neuroelectric brain mapping: an application in a neuromarketing experiment. *Journal of neuroscience methods*, 191(2):283–289, 2010.

[112] Rami N. Khushaba, Sarath Kodagoda, Gamini Dissanayake, Luke Greenacre, Sandra Burke, and Jordan Louviere. A neuroscientific approach to choice modeling: Electroencephalogram (EEG) and user preferences. In *Proceedings of the International Joint*

*Conference on Neural Networks*, 2012. ISBN 9781467314909. doi: 10.1109/IJCNN.2012.6252561.

[113] Bülent Yilmaz, Sümeyye Korkmaz, Dilek Betül Arslan, Evrim Güngör, and Musa H. Asyali. Like/dislike analysis using EEG: Determination of most discriminative channels and frequencies. *Computer Methods and Programs in Biomedicine*, 2014. ISSN 01692607. doi: 10.1016/j.cmpb.2013.11.010.

[114] G. Vecchiato, F. De Vico Fallani, L. Astolfi, J. Toppi, F. Cincotti, D. Mattia, S. Salinari, and F. Babiloni. The issue of multiple univariate comparisons in the context of neuroelectric brain mapping: An application in a neuromarketing experiment. *Journal of Neuroscience Methods*, 2010. ISSN 01650270. doi: 10.1016/j.jneumeth.2010.07.009.

[115] Guibing Guo and Mohamed Elgendi. A New Recommender System for 3D E-Commerce: An EEG Based Approach. *Journal of Advanced Management Science*, 2013. ISSN 21680787. doi: 10.12720/joams.1.1.61-65.

[116] Melody Adhami. Using neuromarketing to discover how we really feel about apps. *International Journal of Mobile Marketing*, 8(1):95–103, 2013.

[117] Clement Levallois, John A Clithero, Paul Wouters, Ale Smidts, and Scott A Huettel. Translating upwards: linking the neural and social sciences via neuroeconomics. *Nature Reviews Neuroscience*, 13(11):789, 2012.

[118] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[119] Arnaud Delorme, Guillaume A Rousselet, Marc J-M Macé, and Michele Fabre-Thorpe. Interaction of top-down and bottom-up processing in the fast visual analysis of natural scenes. *Cognitive Brain Research*, 19(2):103–113, 2004.

[120] Christopher R Madan. Neuromarketing: the next step in market research? *Eureka*, 1 (1):34–42, 2010.

[121] Steven J Stanton, Walter Sinnott-Armstrong, and Scott A Huettel. Neuromarketing: Ethical implications of its use and potential misuse. *Journal of Business Ethics*, 144(4): 799–811, 2017.

# List of Publications

1. **Vaishali Khurana**, Pradeep Kumar, Rajkumar Saini, and Partha Pratim Roy. EEG based word familiarity using features and frequency bands combination. Cognitive Systems Research, 49:33–48, 2018. *(Published)*

2. **Vaishali Khurana**, Monika Gahalawat, Pradeep Kumar and Partha Pratim Roy. A Survey on Neuromarketing using EEG Signals. IEEE Transactions on Affective Computing. *(Submitted)*