EEG-based Preference Prediction in Neuromarketing Using Objective Labeling and Investigating the Effects of Languages on the Preferences of Consumers

By

Md. Fazlul Karim Khondakar

Student ID: 19MEE001P

A thesis submitted in partial fulfilment of the requirements for the degree of MASTER of SCIENCE in Electrical and Electronic Engineering

Department of Electrical and Electronic Engineering

CHITTAGONG UNIVERSITY OF ENGINEERING AND TECHNOLOGY

February 2024

Declaration

I hereby declare that the work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is cited.

Md. Fazlul Karim Khondakar

Student ID: 19MEE001P

Department of Electrical and Electronic Engineering

Chittagong University of Engineering & Technology (CUET)

Dedication

To my parents, wife, and other family members for their continuous support.

List of Publications

Journal Article

- Publication 1: **Md. Fazlul Karim Khondakar**, Mehdi Hasan Chowdhury, Md. Hasib Sarowar, Trishita Ghosh Troyee, Mahmudul Hasan, Md. Azad Hossain, M. Ali Akber Dewan, Quazi Delwar Hossain, "EEG-Based Preference Prediction in Neuromarketing: An Objective Labeling Approach", *Nature Scientific Reports*. (Under Review)
- Publication 2: **Md. Fazlul Karim Khondakar**, Md. Hasib Sarowar, Mehdi Hasan Chowdhury, Md. Azad Hossain, M. Ali Akber Dewan, Quazi Delwar Hossain, "A Systematic Review on EEG-Based Neuromarketing: Recent Trends and Analyzing Techniques", *SN Brain Informatics*. (Under Review)

Conference

- Publication 3: **Md. Fazlul Karim Khondakar**, Trishita Ghosh Troyee, Mahmudul Hasan, Md. Hasib Sarowar, Mehdi Hasan Chowdhury, Quazi Delwar Hossain, "A Comparative Analysis of Different Pre-processing Pipelines for EEG-based Preference Prediction in Neuromarketing", *IEEE PuneCon 2023*. (Published)

Approval by the Supervisor

This is to certify that **Md. Fazlul Karim Khondakar** has carried out this work under my supervision, and that he has fulfilled relevant Academic Ordinance of the Chittagong University of Engineering and Technology, so that he is qualified to submit the following thesis in application for the degree of MASTER of SCIENCE in Electrical and Electronic Engineering.

Dr. Quazi Delwar Hossain

Professor

Department of Electrical and Electronic Engineering

Chittagong University of Engineering & Technology

Acknowledgment

I want to thank Dr. Mehdi Hasan Chowdhury, Md. Hasib Sarowar, Trishita Ghosh Troyee, Mahmudul Hasan, Dr. Md. Azad Hossain, Dr. M. Ali Akber Dewan, and Dr. Sumit Majumder for their significant support and insights that helped me to complete my research work and write my thesis. My heartfelt gratitude goes to the examination committee members for their time and patience.

Also, I am grateful to my supervisor, Dr. Quazi Delwar Hossain, for his continuous mentorship and endless support in every step of my thesis. His insights and knowledge, with a humble research approach, helped to complete my thesis.

I want to acknowledge the students of the Department of Biomedical Engineering, Chittagong University of Engineering & Technology, Bangladesh, for participating and helping me in the data collection of this study.

I am forever grateful to all the Department of Biomedical Engineering and Department of Electrical and Electronic Engineering members for their continuous support.

Also, I would like to thank my parents and wife for their patience and support.

Abstract

Neuromarketing is an emerging brain-computer interface (BCI) research field that aims to understand consumers' internal decision-making processes when choosing which products to buy. It provides valuable insights for marketers to improve their marketing strategies based on consumers' impressions. However, the current status of Electroencephalography (EEG)-based preference prediction and its classification accuracy is still below optimal. The performance of EEG-based preference detection systems depends on a suitable pre-processing pipeline selection and proper data labelling since noisy EEG data and wrong labeling are likely not to give better results. Most of the previous studies followed a traditional EEG pre-processing pipeline, and also recently, an advanced automated EEG pre-processing pipeline has been introduced in the literature. In this study, I have proposed an optimal EEG pre-processing pipeline and compared it with the other two existing methods. I collected raw EEG data and pre-processed them using the three pre-processing pipelines. I then extracted many statistical and frequency domain features and employed several machine learning (ML) classification models. After implementing my proposed pipeline to pre-process EEG data, I observed a significant improvement in classification accuracies compared to the other two existing methods. Another essential thing is to label the data properly so that ML models can be trained accurately. Previous studies have used different ML models to predict subjects' future preferences based on subjective labeling, i.e., the preferences are self-reported by the subjects. However, such labeling methods contradict the main goal of Neuromarketing research, which is to detect genuine preferences from brain data rather than rely on subjects-reported data. This leads to the necessity of an objective labeling approach where the labels can be detected directly from the brain data. In this study, I proposed an objective labeling method for EEG-based preference prediction in Neuromarketing. I compared the performance of different ML models trained on objective and subjective labeling using two datasets: a publicly available Neuromarketing dataset and a new dataset created in my experiments. My findings demonstrate that the ML models trained on objective labeling provided better classification results than the ones trained on subjective labeling. The result aligns with the goals of Neuromarketing research, as it offers an automated way of labeling data and opens up new avenues for future research. Also, among the prominent components of promotions, language has a major impact on consumers' minds. I investigated the effects of foreign languages (FL) on consumers' preferences in Neuromarketing and found that consumers show risk adoption tendencies when exposed to product ads in FL instead of their native languages (NL).

সারাংশ

নিউরোমার্কেটিং হল একটি উদীয়মান ব্রেইন-কম্পিউটার ইন্টারফেস গবেষণা ক্ষেত্র যা কোন পণ্য কিনতে হবে তা বেছে নেওয়ার সময় গ্রাহকদের অভ্যন্তরীণ সিদ্ধান্ত গ্রহণের প্রক্রিয়াগুলি বোঝার চেষ্টা করে। এটি বিপণনকারীদের ভোক্তাদের ইম্প্রেশনের উপর ভিত্তি করে তাদের বিপণন কৌশলগুলি উন্নত করার জন্য মূল্যবান অন্তর্দৃষ্টি প্রদান করে। তবে ইলেক্ট্রোএনসেফালোগ্রাফি (ইইজি)-ভিত্তিক পছন্দের পূর্বাভাসের বর্তমান অবস্থা এবং এর শ্রেণিবিন্যাস সঠিকতা এখনও আশানুরুপ পর্যায়ে পৌছাতে পারে নি। ইইজি-ভিত্তিক পছন্দ সনাক্তকরণ সিস্টেমের কর্মক্ষমতা একটি উপযুক্ত প্রাক-প্রক্রিয়াকরণ পাইপলাইন নির্বাচন এবং সঠিক ডেটা লেবেলিংয়ের উপর নির্ভর করে, কারণ বিভিন্ন সিগনাল মিশ্রিত ইইজি ডেটা এবং ভুল লেবেলিং স্বাভাবিক ভাবেই ভাল ফলাফল দেয় না। পূর্ববর্তী বেশিরভাগ গবেষণায় একটি গতানুগতিক ইইজি প্রাক-প্রক্রিয়াকরণ পাইপলাইন অনুসরণ করা হয়েছে এবং সম্প্রতি একটি উন্নত স্বয়ংক্রিয় ইইজি প্রাক-প্রক্রিয়াকরণ পাইপলাইন চালু করা হয়েছে। এই গবেষণায়, আমরা একটি ইইজি প্রাক-প্রক্রিয়াকরণ পাইপলাইন প্রস্তাব করেছি এবং এটিকে অন্য দুটি বিদ্যমান পদ্ধতির সাথে তুলনা করেছি। আমরা অশোধিত ইইজি ডেটা সংগ্রহ করেছি এবং তিনটি প্রাক-প্রক্রিয়াকরণ পাইপলাইন ব্যবহার করে সেগুলি প্রাক-প্রক্রিয়াকরণ করেছি। এরপর আমরা পরিসংখ্যানগত এবং কম্পাঙ্কগত অনেকগুলো বৈশিষ্ট্য বের করেছি এবং বেশ কয়েকটি মেশিন লার্নিং শ্রেণীবিভাগ মডেল নিযুক্ত করেছি। ইইজি ডেটা প্রাক-প্রক্রিয়া করার জন্য আমাদের প্রস্তাবিত পাইপলাইন বাস্তবায়নের পরে, আমরা অন্য দুটি বিদ্যমান পদ্ধতির তুলনায় শ্রেণীবিন্যাস নির্ভুলতার একটি উল্লেখযোগ্য উন্নতি লক্ষ্য করেছি। আরেকটি অপরিহার্য বিষয় হল ডেটা সঠিকভাবে লেবেল করা যাতে মেশিন লার্নিং মডেলগুলি সঠিকভাবে প্রশিক্ষিত হতে পারে। পূর্ববর্তী গবেষণাগুলি ব্যক্তিভিত্তিক লেবেলিংয়ের উপর ভিত্তি করে তাদের ভবিষ্যত পছন্দগুলি ভবিষ্যদ্বাণী করতে বিভিন্ন মেশিন লার্নিং মডেল ব্যবহার করেছে, অর্থাৎ, লেবেলগুলি অংশগ্রহনকারীদের দ্বারা স্ব-প্রতিবেদিত হয় । এই ধরনের লেবেলিং পদ্ধতি নিউরোমার্কেটিং গবেষণার মূল লক্ষ্যকে বিরোধিতা করে। নিউরোমার্কেটিং গবেষণার মূল লক্ষ্য হল ব্যক্তি-প্রতিবেদিত ডেটার উপর নির্ভর না করে মস্তিষ্কের ডেটা থেকে প্রকৃত পছন্দগুলি সনাক্ত করা। এটি একটি পরোক্ষ লেবেল পদ্ধতির প্রয়োজনীয়তার দিকে পরিচালিত করে যেখানে লেবেলগুলি মস্তিষ্কের ডেটা থেকে সরাসরি সনাক্ত করা যেতে পারে। এই গবেষণায়, আমরা নিউরোমার্কেটিং-এ ইইজি-ভিত্তিক পছন্দের পূর্বাভাসের জন্য একটি পরোক্ষ লেবেলিং পদ্ধতি প্রস্তাব করেছি। আমরা দুটি ডেটাসেট ব্যবহার করে পরোক্ষ এবং ব্যক্তিভিত্তিক লেবেলিংয়ের উপর প্রশিক্ষিত বিভিন্ন মেশিন লার্নিং মডেলের কর্মক্ষমতা তুলনা করেছি। ব্যবহৃত ডেটাসেট দুইটির একটি হল সর্বজনীন উন্মুক্ত বহুল ব্যবহৃত নিউরোমার্কেটিং ডেটাসেট এবং আরেকটি ডেটাসেট আমাদের নিজস্ব পরীক্ষায় তৈরি একটি নতুন ডেটাসেট। আমাদের ফলাফলগুলি দেখায় যে পরোক্ষ লেবেলিংয়ের উপর প্রশিক্ষিত মেশিন লার্নিং মডেলগুলি ব্যক্তিভিত্তিক লেবেলিংয়ের উপর প্রশিক্ষিতদের তুলনায় ভাল শ্রেণীবিভাগের ফল প্রদান করে। ফলাফলটি নিউরোমার্কেটিং গবেষণার লক্ষ্যগুলির সাথে সারিবদ্ধ, কারণ এটি ডেটা লেবেল করার একটি স্বয়ংক্রিয় উপায় সরবরাহ করে এবং ভবিষ্যতের গবেষণার জন্য নতুন পথ খুলে দেয়। এছাড়াও, বিজ্ঞাপনের প্রধান উপাদানগুলির মধ্যে, ভোক্তাদের মনে ভাষা একটি বড় প্রভাব ফেলে। আমরা নিউরোমার্কেটিং-এ ভোক্তাদের পছন্দের উপর বিদেশী ভাষার প্রভাবগুলি তদন্ত করেছি এবং দেখেছি যে ভোক্তারা তাদের স্থানীয় ভাষার পরিবর্তে বিদেশী ভাষায় পণ্যের বিজ্ঞাপনের সংস্পর্শে এলে ঝুঁকি গ্রহণের প্রবণতা দেখায়।

Table of Contents

List of Figures

List of Tables

Nomenclature

BCI: Brain-computer interface;

EEG: Electroencephalography;

fMRI: Functional magnetic resonance imaging;

MEG: Magnetoencephalography;

PET: Positron emission tomography;

fNIRS: functional near-infrared spectroscopy;

ML: Machine learning;

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses;

DNN: Deep Neural Network;

AI: Artificial intelligence;

ECG: Electrocardiogram;

EOG: Electrooculogram;

EMG: Electromyogram;

BPF: Band Pass Filter;

LPF: Low Pass Filter;

HPF: High Pass Filter;

S-Golay: Savitzky-Golay;

ICA: Independent Component Analysis;

FIR: Finite Impulse Response;

SD: Standard deviation;

ASR: Artifact Subspace Reconstruction;

FAA: Frontal Alpha Asymmetry;

DWT: Discrete Wavelet Transform;

ERP: Event Related Potential;

FFT: Fast Fourier Transform;

DFT: Discrete Fourier Transform;

PSD: Power spectral density;

GFP: Global Field Power;

LPP: Late Positive Potential;

PSW: Positive Slow Waves;

AW: Approach-Withdrawal;

DE: Differential Entropy;

HMM: Hidden Markov Model;

SVM: Support Vector Machine;

NN: Neural Network;

RF: Random Forest;

LDA: Linear Discriminant Analysis;

k-NN: k-Nearest Neighbours;

ET: Eye-tracking;

SNN: Spiking Neural Network;

ANN: Artificial Neural Network;

MLP: Multi-layer Perceptron;

NL: Native Language;

FL: Foreign Language.

1 Introduction

Neuromarketing is an emerging brain-computer interface (BCI) research field that aims to understand consumers' internal decision-making processes when choosing which products to buy. It provides valuable insights for marketers to improve their marketing strategies based on consumers' impressions. However, the current status of Electroencephalography (EEG)-based preference prediction and its classification accuracy is still below optimal. This chapter gives a brief review of Neuromarketing. Thereafter the techniques adopted in Neuromarketing are depicted. Also, the problem statement of this thesis is shared. This chapter outlines the background in Section 1.1, the present state of the problem in Section 1.2 and specific objectives in Section 1.3. Section 1.4 describes the significance and scope of this research. Finally, Section 1.5 includes an outline of the remaining chapters of the thesis.

1.1 BACKGROUND

Marketing is the process of introducing a product to the market. A good product cannot inform, engage, and sustain its target audiences without efficient marketing. Marketing serves as the conduit between a product and customers, influencing the final sale. Businesses find it challenging to expand and remain viable without receiving quantitative or qualitative feedback from their customers. In order to thrive in a highly competitive market, newly introduced products demand even more effective marketing strategies [1].

Traditional marketing largely relies on subjective focus-group assessments, interviews, and surveys. However, due to social desirability bias, consumers might not always convey their genuine thoughts [2]. They might not express their feelings in words but instead what they believe others would say in response [3]. People's emotional states or surrounding circumstances at the time of selfreporting are also considered by these post-hoc analytic techniques [4]. Marketers and academics look for supplementary or alternative methods to address the limitations of traditional marketing strategies. One such option is to use Neuromarketing which analyzes customers' spontaneous reactions to certain advertising campaigns, packaging, designs, etc. and explores how customers react to marketing stimuli.

The term "Neuromarketing" was first introduced by Ale Smidts in 2002 to refer to research on the application of neuroscience technology in the marketing sector [5]. It utilizes insights from neuroscience and cognitive science to precisely determine consumer requirements, desires, and preferences. It aids in creating marketing plans and campaigns that are appealing to the target market. By asking customers to respond to surveys after the fact, traditional research techniques concentrate mostly on the posterior attitude of consumers toward products. These replies don't accurately reflect the customer's natural state of mind at the moment of purchase since they are delayed and simplified [6]. By considering the brain signals at the moment of purchase, Neuromarketing, on the other hand, focuses on capturing the in-situ reaction.

It is well recognized that a customer's decision-making process is influenced by various sophisticated aspects, and Neuromarketing offers profound insights into customer behaviors individually. Accordingly, Neuromarketing aids marketers in formulating plans by employing in-the-moment measurements of brain activity in response to diverse marketing stimuli. It uses a direct correlational process to explain consumer responses, in contrast to traditional approaches in marketing research. This objective approach enables marketers to create more successful and efficient tactics by better comprehending consumers' complicated and constantly changing mental processes. It should be noted that Neuromarketing is not necessarily aimed at influencing the specific consumer's personal decisions; instead, it seeks to improve the understanding of prospective customers' perspectives and interests to create precise behavioral models [7].

Neuromarketing studies how people's brains react to particular ads, package designs, products, etc., using non-invasive brain-scanning techniques, such as EEG, functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), positron emission tomography (PET), functional near-infrared spectroscopy (fNIRS), etc. These neurophysiological data may now accurately represent customers' preferences, likes, and dislikes using modern feature extraction and classification algorithms [8]. Marketers utilize these results to produce advertising that customers find more attractive or encouraging. Among the neurophysiological signals, EEG has become the most popular and is widely used in the marketing sector owing to its low price and high temporal resolution. Additionally, because EEG signal changes cannot be consciously influenced, they serve as a more accurate objective measure of emotion [9]. EEG-based emotion recognition is a potential approach to understanding the mechanics behind emotional states and creating computational models for recognizing and forecasting consumer emotional responses. As a result, EEG-based techniques are frequently used in Neuromarketing to increase sales, advertising, package design, pricing, marketing campaigns, and other things.

1.2 PRESENT STATE OF THE PROBLEM

Many approaches have been done over the years to build a proper EEG-based preference prediction system for target consumers [10], [11], [12], [13], [14], [15]. However, the current status of EEG-based preference prediction and its classification accuracy is still below optimal. The performance of EEG-based preference detection systems depends on a suitable selection of pre-processing pipeline and proper labeling of the data so that machine learning algorithms can be trained accurately.

EEG data are very prone to noises and the pre-processing stage removes these noises. It is very crucial stage for preference prediction, since a noisy data is likely not to give better results. Most of the studies in the literature used a traditional pre-processing pipeline [16], [17], [18], [19], [20]. Also, an advanced automated pipeline has been introduced recently in the literature [21]. It is yet unclear, if these pre-processing pipelines are enough for making more precise predictions of customer preferences.

Another essential thing is to label the data properly so that machine learning models can be trained accurately. Neuromarketing research is motivated by the idea that people tend to conceal their true opinions in self-reports. The ultimate objective of Neuromarketing research is to identify genuine preferences from brain data. Therefore, subjective labeling methods are not appropriate as they contradict the main goal of Neuromarketing research. We require an objective labeling approach that can directly label the likes and dislikes from brain data.

The field of Neuromarketing research has evolved from predicting customer preferences to implementing neuroscience-based marketing strategies [22]. Modern studies are trying to build marketing strategies by observing how subtle changes affect the consumers' preferences in practice [23]. Among the prominent components of promotions, language has a major impact on consumers' minds [24]. In a bilingual nation like ours, which advertisement language will work better remains still unchecked.

1.3 SPECIFIC OBJECTIVES

This research work aims to develop an optimal EEG pre-processing pipeline, propose an objective labeling method for EEG-based preference prediction in Neuromarketing and compare the results with subjective labeling, and finally investigate the effects of languages on preferences of consumers in Neuromarketing. So, the specific objectives of this thesis are as follows:

1. To develop an optimal EEG pre-processing pipeline for EEG-based preference prediction in Neuromarketing.
2. To develop an objective labeling method for EEG-based preference prediction in Neuromarketing and compare the results with subjective labeling.
3. To investigate the effects of languages on preferences of consumers in Neuromarketing.

1.4 SIGNIFICANCE OF THE WORK

EEG-based preference prediction in Neuromarketing has become increasingly popular in recent years. It has the potential to help businesses refrain from marketing the wrong products or strategies and save billions of dollars annually. However, the current prediction accuracy is still below optimal levels. Adopting a proper pre-processing pipeline and labeling methods is essential to increase accuracy. Additionally, choosing the correct language is crucial for building an appropriate marketing strategy.

This work addresses these issues by building an optimal pre-processing pipeline that removes noise from the data while maintaining important information. It will also introduce an objective labeling method that avoids dependency on subjective reports and aligns with the goals of Neuromarketing research. Finally, it will examine the effects of language on consumer preferences in Neuromarketing, which can help advertisers choose the appropriate language for their brands in a bilingual nation.

1.5 THESIS OUTLINE

The overall thesis structure is presented in Fig. 1.1. The first chapter describes the background of Neuromarketing and the importance of EEG-based

Neuromarketing techniques. The problem statements, specific objectives, and the significance of the thesis outcomes are also discussed.

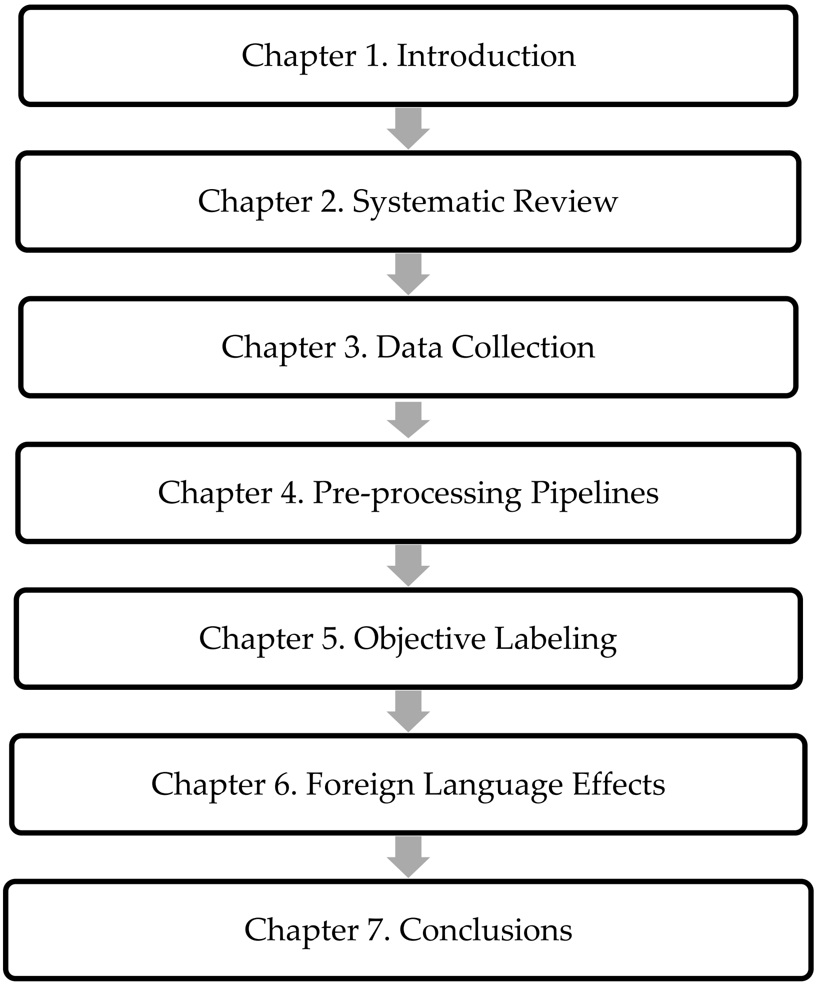


Fig. 1.1 Structure of the thesis.

Chapter 2 presents a systematic review of the previous works in Neuromarketing using cluster analysis. The current research trend in Neuromarketing has been explained. Also, special attention has been given to the active brain regions that relate to the purchase intention of the consumers, types of marketing stimuli should be used to elicit the genuine thoughts of the consumers, and pre-processing & analyzing techniques used for EEG-based Neuromarketing applications.

Chapter 3 presents the data collection scheme and discusses the data collection process of this work in detail.

Chapter 4 presents a comparative analysis of different pre-processing pipelines and proposes an optimal pre-processing pipeline for EEG-based preference prediction in Neuromarketing.

Chapter 5 introduces an objective labeling approach for EEG-based preference prediction in Neuromarketing and compares its performance with subjective labeling.

Chapter 6 uses the objective labeling method to discuss the effects of foreign languages on consumers' preferences. It suggests which language should be adopted for advertising for a bilingual nation.

Chapter 7 draws the conclusions of the study, specifies the key findings and contributions of the works, and discusses the future possibilities in this field.

2 Systematic Review

The goal of consumer neuroscience is to establish physiological neurosis for understanding consumer behavior by combining neuroscientific techniques and theories with behavioral theories, validated models and designs from consumer psychology, and allied fields like behavioral decision sciences. In this journey of consumer neuroscience, much literature has been published to contribute to understanding consumer behavioral traits. This systematic review has reviewed the works done in the last seven years. This section will explore current Neuromarketing research trends, types of marketing stimuli should be used to elicit the genuine thoughts of the consumers, and pre-processing techniques used for EEG-based Neuromarketing applications.

2.1 BACKGROUND & REVIEW QUESTIONS

Large businesses like Google, Unilever, Microsoft, and others utilize the findings from more than 150 consumer neuroscientific firms that are commercially functioning worldwide to influence their consumers in a targeted and effective manner [1]. This innovation has been made possible by academic research, particularly the tremendous analytical reliability of the engineering sector of Neuromarketing. Therefore, it is essential to look into the foundations of Neuromarketing to assess its scopes and capabilities and to offer fresh insight into this area. There have been several review articles on the theoretical aspects of consumer neuroscience, covering various disciplines such as marketing, business ethics, management, psychology, and consumer behavior [22], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34]. However, there is a lack of comprehensive engineering-focused review articles that specifically concentrate on the techniques for recording and processing brain activity and the methods used for interpreting the results in this field [1], [7], [35], [36].

Researchers will be better supported with their research work and be able to move their research careers in the right direction if they keep up to date with the literature in their field of study and are proficient at doing so. It will benefit not only their professional development but also the development of the discipline as a whole. Therefore, knowing the most recent research trends will assist researchers in constructing their research framework properly. After setting the research objectives, a typical workflow for Neuromarketing research can be as presented in Fig. 2.1.

The process of collecting brain data from subjects to analyze their responses to different types of marketing stimuli is a complex one. Researchers attempt to create a realistic buying scenario to better understand the thoughts and preferences of consumers when selecting products. The brain data should be collected from the proper brain regions which effectively possess the emotions regarding the purchase intention. Once the data is collected, it is often very noisy and requires various pre-processing techniques to clean it up. There are many pre-processing techniques available, but the key is to effectively remove the noises while maintaining all necessary information. After the data has been pre-processed, necessary features are extracted from it. There are many feature extraction techniques available for EEG data extracting different time domain, frequency domain, and time-frequency domain features.

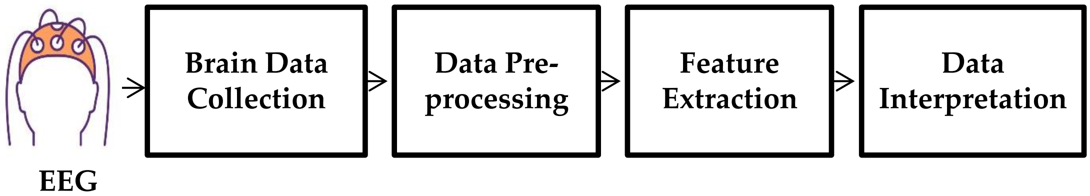


Fig. 2.1 Typical workflow for Neuromarketing research.

After extracting the features, they are utilized for classification purposes and statistical analyses. Various machine learning (ML) algorithms are used for preference classifications to predict the future preferences of consumers, while different statistical analyses are performed for various behavioral analyses.

Proper selection of the processing techniques is essential to interpret the data better. Consequently, numerous technical stages must be completed for final data interpretation, which can prove to be challenging for future researchers to comprehend. In this sense, this article is set to answer the following questions:

- What is the current research trend in the field of Neuromarketing?
- What are the active brain regions that relate to the purchase intention of the consumers?
- What types of marketing stimuli should be used to elicit the genuine thoughts of the consumers?
- Which pre-processing techniques are suitable for EEG-based Neuromarketing applications?
- How different features relate to consumers' decision-making process in EEG-based Neuromarketing applications?
- What are the techniques best suitable for interpreting the data in EEG-based Neuromarketing applications?

These questions will provide a thorough understanding of the most recent research scopes and techniques in consumer neuroscience. Following this short introduction, the methodology for this systematic review will be provided, followed by the key findings corresponding to the questions mentioned above and a discussion of the significant results.

2.2 METHODOLOGY

The systematic reviews are based on clearly stated questions, select relevant studies, evaluate their quality, and summarize the evidence using specific methods. They are distinguished from traditional reviews and commentary by their clear and systematic approach. A precise definition of the research question and a discussion of the inclusion-exclusion criteria are required for systematic reviews to determine the scope of the study. After a thorough review of the literature, papers should be chosen according to their relevance, and the findings of the selected research should be critically synthesized and evaluated to reach definite conclusions [37]. For this systematic review, I followed the instructions provided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) to select original research articles [38]. The PRISMA protocol consists of four stages, as shown in Fig. 2.2 - identification, screening, eligibility, and inclusion.

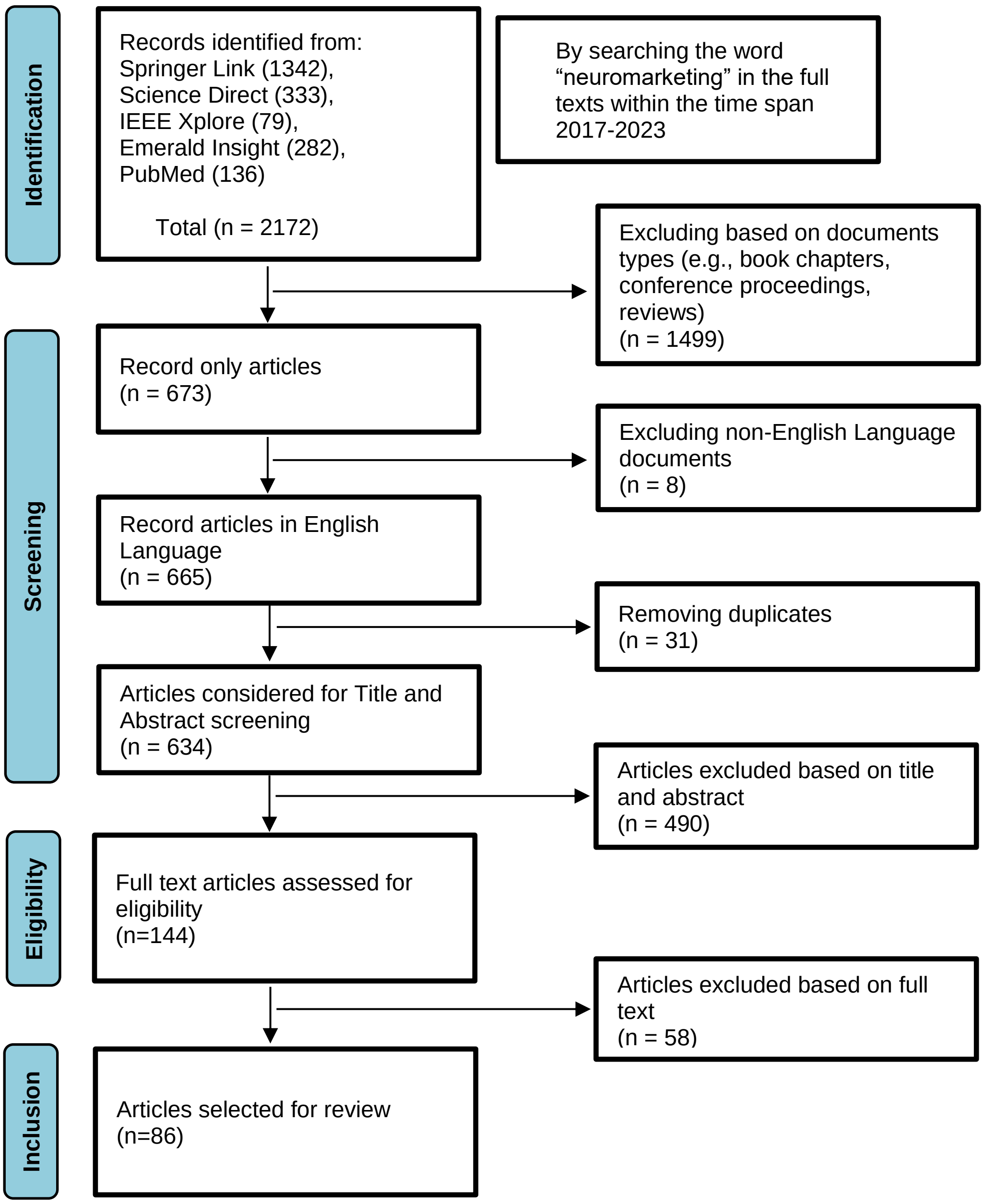


Fig. 2.2 PRISMA chart for selecting publications for the systematic review.

2.3 THE RESEARCH TREND IN NEUROMARKETING

Being a relatively new field of research, Neuromarketing has drawn the attention of researchers in recent years. Many studies have been done in this field, and the research trend is evolving. Hence upcoming researchers should pay attention to the trend.

Early studies of this field used to focus on building the theoretical base of Neuromarketing, such as using the concept of Neuroscience in marketing, economics, and customer behavior, and also to predict the preference of consumers merely. But modern studies focused on building marketing strategies utilizing Neuroscientific methods. To comprehend the context of consumer decision-making and behavior, they explored the elements influencing consumer decisions from a multidisciplinary point of view. These studies looked at how brain responses can predict consumer behavior in relation to subtle changes in marketing aspects, such as packaging design, advertising message, pricing, and brand identification other than the product.

These research aspects can be used to form different groups so that cluster analysis can be performed to track the research trend in Neuromarketing. Such cluster analysis was conducted in the study by Shahriari et al. [22]. They reviewed the research articles from valid databases between 2005 to 2017 and divided them into 6 clusters with text-mining methodology. In this study, I have considered the research articles from valid databases between 2017 to 2023 and divided them into similar 5 clusters based on the research goals of the studies. One of the clusters from Shahriari et al. [22], namely Ethical Issues, has not been considered here since this review avoided theoretical/conceptual studies and considered only the research articles that directly contribute to the Neuromarketing research's experimental findings. The names of the five clusters remain the same as in Shahriari et al. [22] to compare results between 2005-2017 and 2017-2023 and represent current research trends in Neuromarketing.

2.3.1 Cluster 1 – Ads & Video Commercials

The articles that deal with identifying emotion through brain data using different stimuli, such as e-commerce product images [10], [39], video ads [40], [41], product shapes [42] etc., have been considered under this cluster. These articles mainly focused on EEG-based emotion identification/preference prediction by adopting different ML techniques to train the models to predict the future preferences of the consumers. Yadava et al. [10] proposed a predictive model to catch consumers' intentions toward E-commerce products. Teo et al. [42] investigated several deep neural network (DNN) architecture tunings for increasing the classification rate of the preference classification task. Aldayel et al. [43] used several feature sets of EEG indices to explore preference prediction. Oikonomou et al. [44] proposed a sparse classification scheme for recognizing cognitive and affective brain mechanism in Neuromarketing. In this review, about 22.09% (19 articles) of the total research articles have been found under this cluster as compared to 25% of Shahriari et al. [22] as presented in Table 2.1.

Table 2.1 The articles under the five clusters

Cluster	Neuromarketing Studies (2017-23)	Percentage of Studies (2005-17)	Percentage of Studies (2017-23)
Ads & Video Commercials	Yadava et al. [10], Hakim et al. [39], Hakim et al. [40], Guixeres et al. [41], Teo et al. [42], Aldayel et al. [43], Oikonomou et al. [44], Aldayel et al. [13], Kumar et al. [11], Li et al. [45], Yen and Chiang [46], Zeng et al. [47], Raiesdana and Mousakhani [48], Kislov et al. [49], Al-Nafjan [50], Shah et al. [51], Hassani et al. [52], Georgiadis et al. [53], Göker [54].	25%	22.09%
Neuroscience in Marketing, Economics, and Consumer Behavior	Shen et al. [55], Uhm et al. [56], González-Morales [18], Wajid et al. [19], Gountas et al. [57], Domracheva and Kulikova [58], Wei et al. [59], Ramsøy et al. [60], Yang et al. [61], Harris et al. [62], Goto et al. [63], Eijlers et al. [64], Vozzi et al. [65], Zito et al. [66], Wang et al. [67], Ma et al. [68], Kakaria et al. [69], Lukovics et al. [70].	17%	20.93%
Marketing Strategies	Goto et al. [71], Jin et al. [72], Çakar et al. [73], Royo et al. [74], Gong et al. [12], Daugherty et al. [75], Gholami Doborjeh et al. [76], Alonso Dos Santos	32%	44.19%

	and Calabuig Moreno [77], Gordon et al. [78], García-Madariaga et al. [79], Fu et al. [80], Golnar-Nik et al. [81], Sänger [82], Alvino et al. [83], Verhulst et al. [84], Zubair et al. [85], Hsu and Chen [86], Hsu and Chen [23], Pagan et al. [87], Zhao and Wang [88], Izadi et al. [15], Robertson et al. [89], Martinez-Levy et al. [20], Wang et al. [90], Ma et al. [91], Yu et al. [92], Mengual-Recuerda et al. [93], Kim et al. [94], Wang et al. [95], Alvino et al. [96], Russo et al. [97], Wang et al. [98], Hassani et al. [99], Wei et al. [100], Damião de Paula et al. [101], Russo et al. [102], Bosshard and Walla [103], Song et al. [104].		
Advertising Message Components	Avinash et al. [105], Michael et al. [106], Hsu and Chen [107], Pennanen et al. [108], Leeuwis et al. [109], Uhm et al. [110].	6%	6.98%
Decision Making Process and Brand Selection	Ma et al. [111], Garczarek-Bąk et al. [112], Özbeyaz [113], Yang and Kim [114], Camarrone and Van Hulle [115].	10%	5.81%

2.3.2 Cluster 2 – Neuroscience in Marketing, Economics, and Consumer Behavior

The research works that used the concept of Neuroscience in marketing, economics, and customer behavior, have been incorporated into this cluster. These articles mainly focused on building the theoretical base of Neuromarketing. Such studies were motivated to find out the fundamental neural basis and psychological processing working behind the customer decision-making mechanism [55], [56]. Some studies also investigated brain activation patterns while consumers choose products to buy [18]. From such neuropsychological perspectives, Kakaria et al. [69] compared cognitive load during planned and unplanned virtual shopping and Lukovics et al. [70] investigated customer acceptance of self-driving technology. I have found about 20.93% (18 articles) of the total research articles under this cluster as compared to 17% of Shahriari et al. [22] as presented in Table 2.1.

2.3.3 Cluster 3 – Marketing Strategies

The articles under this cluster investigated the elements influencing consumer decisions in relation to subtle changes in marketing aspects, such as packaging design, advertising message, languages used for ads, and pricing other than the product. These articles mainly focused on building marketing strategies utilizing Neuroscientific methods. Royo et al. [74] investigated whether product design and follow-up advertising and marketing using a continuous narrative approach in conjunction with verbal narrative advertising could provide more positive emotional responses from future product users than using visual narrative advertising. Alonso Dos Santos and Calabuig Moreno [77] presented a pilot study that sought to examine the focus on sponsor variables by evaluating the level of consent associated with sponsors and sponsored organizations. Gordon et al. [78] studied consumer reactions to record-based videos in energy-efficient social marketing. García-Madariaga et al. [79] explored consumers' attention and preferences for three packaging attributes, image, text, and color, as separate variables. Fu et al. [80] investigated the impact of price deception on the consumer buying behavior and the neural mechanisms that underpin them. Golnar-Nik et al. [81] investigated the possibility of EEG spectral power for predicting customer preferences and interpreting changes in consumer buying behavior when the content of an advertisement, such as the backdrop color and promotions, was modified. Alvino et al. [83] assessed whether the EEG makes a significant and tangible contribution to predicting consumer behavior and preferences when using similar products of different prices. Hsu and Chen [86] looked into how hotel videos featuring a happy face emoji as a subtle subtext impact people's hotel choice. Izadi et al. [15] looked at the neuropsychological reactions of customers to promotional strategies and their choice to purchase sports goods to establish the most successful method. Hassani et al. [99] investigated the impacts of products' colors on consumers' purchase decisions.

The main objective of these studies was to investigate how subtle changes in marketing utilities affect consumers' purchase behavior. I have found out about 44.19% (38 articles) of the total research articles under this cluster as compared to 32% of Shahriari et al. [22] as presented in Table 2.1.

2.3.4 Cluster 4 – Advertising Message Components

Articles in this cluster discuss the emotional effects that various advertising message components have on consumers. The effects of music and visual stimuli were the main concerns of these articles. Avinash et al. [105] proposed a model to provide an eminent way to understand customer behavior through Neuromarketing auditory stimuli as advertisement jingles so that any company can launch the best advertisements jingle for promoting their business. Michael et al. [106] studied unconscious emotional and cognitive responses using travel images to understand the specific mental processes of travel behavior. Hsu and Chen [107] analyzed the aphrodisiac effects of musical stimuli during wine tasting. Uhm et al. [110] investigated the impact of music on viewers' responses to sports ads. In this systematic review, about 6.98% (6 articles) of the total research articles have been found under this cluster as compared to 6% of Shahriari et al. [22] as presented in Table 2.1.

2.3.5 Cluster 5 – Decision Making Process and Brand Selection

The research studies under this cluster dealt with the effects of brand familiarity on consumers' decision-making process. Ma et al. [111] examined brain characteristics impacted by brand image and product group in a brand portfolio experiment. Garczarek-Bąk et al. [112] explored the potential of forecasting brand sales based on psychophysiological responses to a retailer's television commercial. Özbeyaz [113] studied consumer judgments for branded and unbranded stimuli using machine learning algorithms. In this study, about

5.81% (5 articles) of the total research articles have been found under this cluster as compared to 10% of Shahriari et al. [22] as presented in Table 2.1.

According to the results of this study and the research of Shahriari et al. [22], the most popular topics in Neuromarketing from 2005 to 2023 are shown in Fig. 2.3. The data reveals that the number of articles on building marketing strategies has significantly increased over time, while the number of articles on other major topics has decreased. This indicates that the research trend in Neuromarketing has shifted from mere preference prediction of consumers to building marketing strategies using Neuroscience. Nevertheless, comparable research is ongoing for preference prediction in cluster 1 since understanding consumer preferences is one of the primary objectives of Neuromarketing. However, there is still no definitive method for predicting consumer preferences in Neuromarketing. With the advancements of artificial intelligence (AI) and ML techniques, researchers are increasingly interested in exploring this field of research. Another significant field is cluster 2, which aims to establish the theoretical basis of Neuromarketing. As a relatively new field of research, Neuromarketing enables researchers to develop principles for using Neuroscience concepts in marketing and economics to ensure its success.

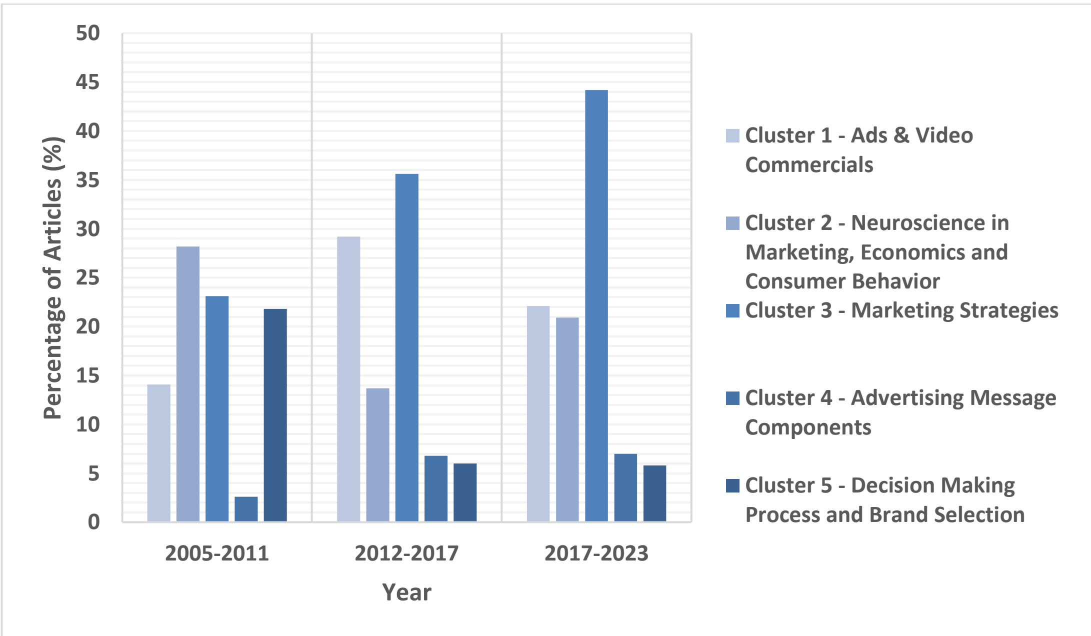


Fig. 2.3 The research trend in Neuromarketing.

This work considered the two most popular clusters – clusters 1 and 3. In cluster 1, I intend to increase the prediction accuracy of consumer preference by using an optimal pre-processing pipeline and introducing proper labeling. In cluster 3, I want to see the effects of language on consumers' preferences.

2.4 ACTIVE BRAIN REGIONS IN NEUROMARKETING APPLICATIONS

The comprehension of the human brain's structure has shown to be essential in Neuromarketing research due to the close correlation between its functionality and the interpretation of neural responses [1]. The frontal, parietal, temporal, and occipital lobes are the four lobes that make up the complex outer layer of the human brain. Each lobe has a unique function related to motor, emotional, and cognitive responses. The frontal lobe is responsible for most of our conscious reasoning and decision-making [116]. The posterior section of the frontal lobe is in charge of movement-related decisions, while the prefrontal region handles cognitive decision-making. The left frontal area plays a role in the experience of positive emotions such as joy, interest, and happiness. These emotions encourage and sustain approach motivations. On the other hand, the right frontal area is responsible for negative emotions such as fear, disgust, and sadness. These emotions support and maintain withdrawal motivations [117]. The parietal lobe processes information about taste, touch, and movement. The temporal lobe is responsible for acoustic identification, visual memory, and integrating new sensory data with old memories. Finally, the occipital brain is the primary center for visual processing [1].

The EEG is a commonly used instrument to measure brain activity. It records the electrical activity on the scalp by analyzing the voltage differences produced by firing neurons in the brain. The EEG data acquisition system utilizes the international 10-20 method to place multiple electrodes directly on the head. This method provides information about the space between electrodes, specifically within 10-20% of the scalp boundaries from front to back or left to right. The distance between adjacent electrodes is 10% or 20% of the scalp diameter. The 10-20 standard has been commonly used in various EEG systems to enhance signal reliability and minimize signal-to-noise ratio. It is necessary to position the electrodes exactly above the relevant brain region in order to record the neuronal activity associated with a particular function. The placement of electrodes and the brain regions with their primary functions are shown in Fig. 2.4.

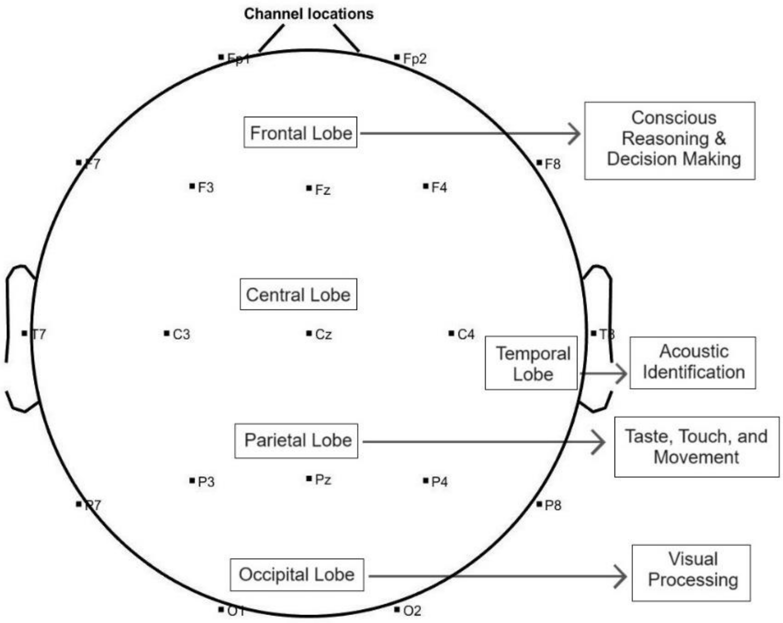


Fig. 2.4 The EEG electrodes in 10-20 system and related brain functions.

The EEG activity can be divided into different frequency bands, such as delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ). These frequency bands have a connection to emotional reactions. The theta band, located at the center of the frontal brain, shows how emotions are processed when a consumer views a product. The alpha band on the prefrontal cortex distinguishes between positive and negative emotional responses. The beta band is linked to changes in emotional arousal, while the gamma band is mainly related to the effects of arousal [118]. The brain activation is positively correlated with the beta band and negatively correlated with the alpha band. Approach motivation towards a product is indicated by higher beta and lower alpha in the left frontal than in the right frontal area. This difference in activity between the left and right frontal sides is called frontal asymmetry [66], [97], [117], [119]. Greater activation in the left prefrontal region, as indicated by higher values in the gamma band, can be correlated with consumers' willingness to pay [60]. The activity level of the frontal theta in the prefrontal cortex indicates the cognitive processing that arises from mental exhaustion. Higher levels of theta activity are associated with more challenging and complicated tasks in the frontal area [120]. The mostly utilized EEG sub-bands along with their frequency ranges, brain regions of interest, and interpretations in Neuromarketing are summarized in Table 2.2.

Table 2.2 The EEG sub-bands interpretation in Neuromarketing

EEG Band	**Frequency**	**Brain Region**	**Interpretation**	**Neuromarketing Studies**
Theta (θ)	4-8 Hz	Fronto-central	Positively correlated with more challenging and complicated tasks	Golnar-Nik et al. [81], Avinash et al. [105], Modica et al. [120].
Alpha (α)	8-13 Hz	Pre-frontal	Negatively correlated with brain activation - higher alpha in the right frontal area indicates appr oach motivation and vice versa	Zeng et al. [47], Al-Nafjan [50], Ramsøy et al. [60], Zito et al. [66], Martinez-Levy et al. [20], Russo et al. [97], Touchette and Lee [117], Al-Nafjan et al. [121].
Beta (β)	13-30 Hz	Frontal	Positively correlated with brain activation - higher alpha in the left frontal area indicates approach motivation and vice versa	Zeng et al. [47], Al-Nafjan [50], Ramsøy et al. [60], Zito et al. [66], Martinez-Levy et al. [20], Russo et al. [97], Touchette and Lee [117], Al-Nafjan et al. [121].
Gamma (γ)	30 ~ Hz	Pre-frontal	Positively correlated with brain activation - higher gamma in the left frontal area indicates more willingness-to-pay and vice versa	Zeng et al. [47], Al-Nafjan [50], Ramsøy et al. [60].

2.5 STIMULI IN NEUROMARKETING APPLICATIONS

To elucidate the decision-making process of consumers, different types of marketing stimuli have been used in the literature over the years, such as direct products, images of the products, video advertisements, normal pictures, affective images, music, movies, etc. These stimuli can be classified into three major categories: *Product, Promotion,* and *Others.*

2.5.1 Product

Actual products and images of the products are considered in this category. Most of the studies found to be using products' images in different formats rather than the actual products as presented in Table 2.3. Yadava et al. [10] used images of 14 different products having 3 different varieties of each resulting in 42 product images for stimuli. Teo et al. [42] presented 60 different bracelet-like 3D shapes to the participants to record their EEG. Kumar et al. [11] used 14 online product categories with 3 variants for each category resulting in a total of 42 different product images. Goto et al. [71] recorded participants' EEG while performing a virtual shopping task, and each participant was shown a series of 180 products, including electronics, food, beverages, and sports equipment. Gong et al. [12] used 480 product images from 4 categories: shampoo, water cup, headset, and USB flash drive. García-Madariaga et al. [79] used 63 product images from 3 categories: drinks, salty snacks, and sweet appetizers. Fu et al. [80] used colored images of 20 products with real and fake prices. Izadi et al. [15] used 40 images of sports products containing four strategies: advertising, discount, charity, and endorsement. Ma et al. [111] used different images consisting of product names and brand names. Özbeyaz [113] used 10 branded and 10 unbranded smartphone images as stimuli.

Table 2.3 Stimuli used in Neuromarketing Applications

Stimuli	Neuromarketing Studies		% of Studies
Product	Products' Images	Yadava et al. [10], Hakim et al. [39], Teo et al. [42], Aldayel et al. [43], Aldayel et al. [13], Kumar et al. [11], Yen and Chiang [46], Zeng et al. [47], Raiesdana and Mousakhani [48], Kislov et al. [49], Al-Nafjan [50], Shah et al. [51], Hassani et al. [52], Georgiadis et al. [53], Göker [54], Shen et al. [55], Domracheva and Kulikova [58], Ramsøy et al. [60], Goto et al. [63], Eijlers et al. [64], Wang et al. [67], Ma et al. [68], Kakaria et al. [69], Goto et al. [71], Jin et al. [72], Çakar et al. [73], Gong et al. [12], Gholami Doborjeh et al. [76], García-Madariaga et al. [79], Fu et al. [80], Zubair et al. [85], Zhao and Wang [88], Izadi et al. [15], Wang et al. [90], Ma et al. [91], Yu et al. [92], Kim et al. [94], Wang et al. [95], Alvino et al. [96], Wang et al. [98], Hassani et al. [99], Wei et al. [100], Damião de Paula et al. [101], Ma et al. [111], Özbeyaz [113], Yang and Kim [114], Camarrone and Van Hulle [115].	54.02%
	Actual Products	Li et al. [45], Lukovics et al. [70], Alvino et al. [83], Hsu and Chen [23], Pagan et al. [87], Robertson et al. [89], Mengual-Recuerda et al. [93], Hsu and Chen [107], Pennanen et al. [108].	10.34%
Promotion	Hakim et al. [40], Guixeres et al. [41], Oikonomou et al. [44], Wajid et al. [19], Gountas et al. [57], Wei et al. [59], Harris et al. [62], Vozzi et al. [65], Zito et al. [66], Royo et al. [74], Daugherty et al. [75], Gordon et al. [78], Golnar-Nik et al. [81], Hsu and Chen [86], Russo et al. [97], Russo et al. [102], Uhm et al. [110], Garczarek-Bąk et al. [112], Camarrone and Van Hulle [115].		21.84%
Others	Uhm et al. [56], González-Morales [18], Yang et al. [61], Alonso Dos Santos and Calabuig Moreno [77], Sänger [82], Verhulst et al. [84], Martinez-Levy et al. [20], Bosshard and Walla [103], Song et al. [104], Avinash et al. [105], Michael et al. [106], Leeuwis et al. [109].		13.79%

Some studies were found to use the actual products as stimuli as shown in Table 2.3. Li et al. [45] used different ladies' shirts. Lukovics et al. [70] used real life experience of self-driving vehicles to record EEG. Alvino et al. [83] used 2 high price wines and 2 low price wines. Some studies used different snack products [23], [108]. Pagan et al. [87] and Hsu and Chen [107] used 2 wines from two different countries of origin. Robertson et al. [89] used 20 different white wine varieties to record EEG.

2.5.2 Promotion

The studies that used only video commercials are included in this category as presented in Table 2.3. Hakim et al. [40] used 3 video commercials for 6 food products. Gountas et al. [57] used five video commercials with alcohol reduction themes. Royo et al. [74] used a video commercial of a pushchair. Gordon et al. [78] used 4 narrative videos focused on important energy use habits, including refrigerators, lighting, laundry, and appliances. Golnar-Nik et al. [81] used video ads from four popular mobile brands (Samsung, Apple, Meizu, and Nokia). Hsu and Chen [86] used hotel videos with and without subtle effects. Russo et al. [97] used two video commercials of traditional cheese as stimuli to record EEG data.

2.5.3 Others

Some studies used normal pictures, affective images, music, movies, website visiting, and texts as stimuli and related them to Neuromarketing applications. These stimuli are taken into account in this category as shown in Table 2.3. González-Morales [18] and Martinez-Levy et al. [20] used different affective images. Alonso Dos Santos and Calabuig Moreno [77] used sponsored messages and montages. Sänger [82] used "food vs. neutral" pictures for participants to perform a central oddball task. Avinash et al. [105] used different types of classic music. Michael et al. [106] used multiple photos and videos of travel destinations.

A summary of the marketing stimuli used in the literature according to the three categories is shown in Fig. 2.5. It shows that the majority of the studies in the literature used products' images as marketing stimuli to record the brain data. It is due to the popularity of online shopping where consumers typically view product images with accompanying details. Neuromarketing researchers should aim to create a realistic purchasing environment for consumers, regardless of the type of stimuli used. This will help to reveal the genuine thoughts of consumers when making purchasing decisions. However, the technology improvements have shifted the focus of marketing stimulus away from the TV commercials and toward the images of the products. While recording the data, the participants should not be given a time limit to observe the images, as this can cause pressure and does not reflect a natural purchasing environment. Studies by Yadava et al. [10] and Zeng et al. [47] allowed participants to observe product images for only 4 and 8 seconds, respectively, which is not sufficient.

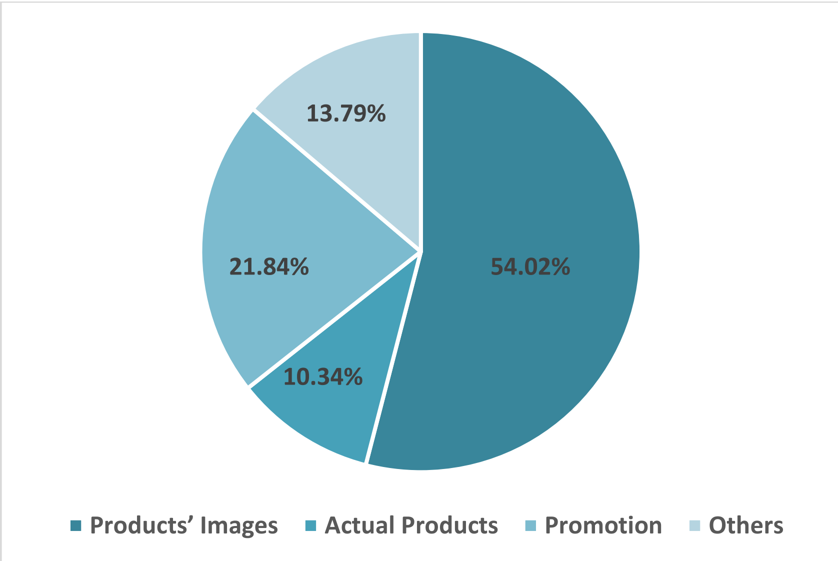


Fig. 2.5 Stimuli used in Neuromarketing applications.

2.6 PRE-PROCESSING TECHNIQUES IN NEUROMARKETING

EEG signals are very susceptible to noises, such as cardiac signals – Electrocardiogram (ECG), visual signals produced by eye movements – Electrooculogram (EOG), movement artifacts generated by muscle contractions – Electromyogram (EMG), and power line interference (50/60 Hz) caused by power line frequency. The pre-processing stage removes these noises and prepares the signal within the desired frequency band for further processing.

Properly pre-processing the raw EEG data is crucial since noisy EEG data will likely not give better results.

EEG is affected by low-frequency noises caused by various sources such as eye movement, head, electrode wires, and sweat on the scalp. The low-frequency noise is characterized by slow drifts in the EEG signal over several seconds. On the other hand, high-frequency noise is caused by eye blinks and muscle contractions, particularly in facial and neck muscles. High-frequency noise manifests itself as quick up-down changes in the EEG signal. The EEG signal has a frequency band ranging from around 0.5 to 100 Hz, leading to the common usage of Band Pass Filter (BPF) to reject low-frequency and high-frequency noises [13], [20], [57], [74], [77]. The BPF has two cutoff frequencies – lower and higher- and the signal within the two cutoff frequencies is allowed to pass through the filter. The cutoff frequencies are adjusted to fall in the desired EEG range, such as 4-45 Hz [43]. Some studies were also found to use Low Pass Filter (LPF) [18], [55], [80] and High Pass Filter (HPF) [73], [87], [88] separately, where the LPF uses only the lower cutoff frequency and the HPF uses only the higher cutoff frequency. Notch Filter, which has response characteristics quite the opposite of the BPF, is often used to eliminate power line interference (50/60 Hz) [42], [74], [112]. Another filtering technique found in the literature is the Savitzky-Golay (S-Golay) filter, which is used to smooth the signal and increase the precision of the acquired data without changing the tendency of the raw data [10], [11], [13]. This is accomplished using the convolution technique, which involves utilizing linear least squares to fit successive subsets of neighbouring data points with a low-degree polynomial [122].

Independent Component Analysis (ICA) is a blind source separation technique widely used as a pre-processing tool in Neuromarketing applications [13], [18], [19], [20], [40], [81], [87]. It is a valuable technique for analyzing multi-channel signals and separating the signal into its additive subcomponents. The noises, such as eye blink data, EMG, ECG, and EOG, are divided into different subcomponents, which can be removed to obtain the original data. Segmentation is also commonly used as a pre-processing technique to separate the signal into multiple epochs with identical statistical characteristics in terms of time and frequency [19], [55], [71], [72], [82], [83], [88], [105]. The pre-processing techniques used in Neuromarketing are summarized in Table 2.4.

Table 2.4 Pre-processing techniques used in Neuromarketing

Pre-processing Techniques	**Description**	**Neuromarketing Studies**
BPF	Allows the signal within the two cutoff frequencies to pass - to get the desired EEG range	Hakim et al. [39], Guixeres et al. [41], Aldayel et al. [43], Aldayel et al. [13], Zeng et al. [47], Raiesdana and Mousakhani [48], Al-Nafjan [50], Georgiadis et al. [53], Shen et al. [55], Wajid et al. [19], Gountas et al. [57], Domracheva and Kulikova [58], Goto et al. [63], Vozzi et al. [65], Zito et al. [66], Ma et al. [68], Lukovics et al. [70], Goto et al. [71], Royo et al. [74], Daugherty et al. [75], Alonso Dos Santos and Calabuig Moreno [77], Robertson et al. [89], Martinez-Levy et al. [20], Kim et al. [94], Russo et al. [97], Wei et al. [100], Russo et al. [102], Bosshard and Walla [103], Pennanen et al. [108], Ma et al. [111], Garczarek-Bąk et al. [112], Özbeyaz [113], Camarrone and Van Hulle [115].
LPF	Allows the signal below a cutoff frequency to pass - to avoid high frequency noises	Kislov et al. [49], Shen et al. [55], González-Morales [18], Eijlers et al. [64], Wang et al. [67], Kakaria et al. [69], Jin et al. [72], Çakar et al. [73], Fu et al. [80], Sänger [82], Zubair et al. [85], Pagan et al. [87], Zhao and Wang [88], Wang et al. [90], Ma et al. [91], Yu et al. [92], Wang et al. [98], Song et al. [104].
HPF	Allows the signal above a cutoff frequency to pass - to avoid low frequency noises	Oikonomou et al. [44], Kislov et al. [49], Hassani et al. [52], González-Morales [18], Yang et al. [61], Kakaria et al. [69], Çakar et al. [73], Pagan et al. [87], Zhao and Wang [88], Hassani et al. [99], Song et al. [104], Leeuwis et al. [109], Yang and Kim [114].
Notch Filter	Rejects the signal within the two cutoff frequencies – to	Teo et al. [42], Oikonomou et al. [44], Shah et al. [51], Hassani et al. [52], Yang et al. [61], Eijlers et al. [64], Vozzi et al. [65], Kakaria et al. [69], Royo et al. [74], Alonso Dos Santos and

	reject power line interference	Calabuig Moreno [77], Martinez-Levy et al. [20], Russo et al. [97], Hassani et al. [99], Leeuwis et al. [109], Garczarek-Bąk et al. [112], Özbeyaz [113], Yang and Kim [114].
S-Golay Filter	Uses the convolution technique utilizing linear least squares - to smooth the signal and increase the precision of the acquired data	Yadava et al. [10], Aldayel et al. [43], Aldayel et al. [13], Kumar et al. [11], Al-Nafjan [50], Shah et al. [51].
ICA	Divides the signal into its additive subcomponents, and then the noisy components can be removed	Hakim et al. [40], Guixeres et al. [41], Aldayel et al. [43], Oikonomou et al. [44], Aldayel et al. [13], Zeng et al. [47], Raiesdana and Mousakhani [48], Kislov et al. [49], Hassani et al. [52], Georgiadis et al. [53], González-Morales [18], Wajid et al. [19], Yang et al. [61], Vozzi et al. [65], Ma et al. [68], Golnar-Nik et al. [81], Pagan et al. [87], Martinez-Levy et al. [20], Kim et al. [94], Alvino et al. [96], Russo et al. [97], Hassani et al. [99], Wei et al. [100], Russo et al. [102], Yang and Kim [114].
Segmentation	Divides the signal into multiple epochs with similar statistical characteristics in terms of time and frequency	Shen et al. [55], Wajid et al. [19], Harris et al. [62], Goto et al. [63], Eijlers et al. [64], Wang et al. [67], Goto et al. [71], Jin et al. [72], Sänger [82], Alvino et al. [83], Zhao and Wang [88], Wang et al. [90], Ma et al. [91], Yu et al. [92], Song et al. [104], Avinash et al. [105].

2.7 FEATURE EXTRACTION TECHNIQUES IN NEUROMARKETING

Feature extraction is simply the process of minimizing the number of resources needed to explain a significant amount of data. It is a technique for reducing the number of features by generating new ones from the old ones such that the new set of features can summarize the majority of the information present in the original feature set. Several different feature extraction approaches and features have been utilized in the literature as listed in Table 2.5.

Table 2.5 Feature extraction techniques used in Neuromarketing

Feature Extraction Techniques	**Neuromarketing Studies**
Discrete Wavelet Transform (DWT)	Yadava et al. [10], Aldayel et al. [13], [43], Kumar et al. [11], Al-Nafjan [50], Shah et al. [51], Hassani et al. [52], [99].
Power Spectra of EEG Sub-bands	Hakim et al. [39], Hakim et al. [40], Guixeres et al. [41], Teo et al. [42], Aldayel et al. [43], Oikonomou et al. [44], Aldayel et al. [13], Li et al. [45], Zeng et al. [47], Raiesdana and Mousakhani [48], Kislov et al. [49], Al-Nafjan [50], Shah et al. [51], Hassani et al. [52], Göker [54], Uhm et al. [56], González-Morales [18], Wajid et al. [19], Gountas et al. [57], Domracheva and Kulikova [58], Wei et al. [59], Ramsøy et al. [60], Harris et al. [62], Eijlers et al. [64], Vozzi et al. [65], Zito et al. [66], Lukovics et al. [70], Çakar et al. [73], Daugherty et al. [75], Alonso Dos Santos and Calabuig Moreno [77], Gordon et al. [78], Golnar-Nik et al. [81], Hsu and Chen [23], Pagan et al. [87], Robertson et al. [89], Martinez-Levy et al. [20], Kim et al. [94], Russo et al. [97], Hassani et al. [99], Damião de Paula et al. [101], Russo et al. [102], Avinash et al. [105], Pennanen et al. [108], Leeuwis et al. [109], Uhm et al. [110].
Event Related Potential (ERP)	Shen et al. [55], Domracheva and Kulikova [58], Yang et al. [61], Goto et al. [63], Wang et al. [67], Ma et al. [68], Goto et al. [71], Jin et al. [72], Gong et al. [12], Daugherty et al. [75], Gholami Doborjeh et al. [76], Fu et al. [80], Sänger [82], Zubair et al. [85], Zhao and Wang [88], Wang et al. [90], Ma et al. [91], Yu et al. [92], Wang et al. [95], Wang et al. [98], Wei et al. [100], Bosshard and Walla [103], Song et al. [104], Ma et al. [111], Yang and Kim [114], Camarrone and Van Hulle [115].
EEG Asymmetry Indices	Hakim et al. [40], Guixeres et al. [41], Aldayel et al. [43], Oikonomou et al. [44], Zeng et al. [47], Kislov et al. [49], Vozzi et al. [65], Zito et al. [66], Kakaria et al. [69], Lukovics et al. [70], Çakar et al. [73], García-Madariaga et al. [79], Verhulst et al. [84], Martinez-Levy et al. [20], Mengual-Recuerda et al. [93], Russo et al. [97], Damião de Paula et al. [101], Russo et al. [102], Michael et al. [106], Garczarek-Bąk et al. [112].
Statistical Features	Yadava et al. [10], Aldayel et al. [43], Zeng et al. [47], Al-Nafjan [50], Hassani et al. [52], Wei et al. [59], Hassani et al. [99].

2.7.1 DWT

Many studies have been found in the literature that used DWT technique to extract wavelet coefficients as EEG features corresponding to different EEG frequency bands [10], [11], [13], [43], [50], [51], [52], [99]. The DWT is a signal analysis method that decomposes signals into different coefficients in the time-frequency domain. It can be described as a multi-resolution or multi-scale analysis where each coefficient provides a unique representation of the input signal. The DWT uses a convolution operation, which is a two-function multiplication process, to generate wavelet coefficients. The resulting inner product of each wavelet with the input signal produces a unique coefficient [43], [123], [124]. The DWT can be mathematically expressed using Equation (1).

$$D(i,j) = \sum_{n=0}^{M-1} x(n).\,\varphi_{i,j}^{*}(n) \qquad (1)$$

where $x(n)$ is a signal of length n, and $\varphi_{i,j}^{*}(n)$ is scaling wavelet function. In DWT decomposition, a filter bank consisting of a group of high- and low-pass filters is used. The low-pass filters generate approximation coefficients, while the high-pass filters produce wavelet detail coefficients. The accuracy of DWT analysis depends on the selection of wavelet technique and the number of wavelet decomposition levels [10], [43], [123], [124]. The number of wavelet decomposition levels depends on the sampling frequency of the signals. For 128 Hz sampling frequency, the four-level signal decomposition technique Daubechies 4 (DB4) was used to extract the wavelet coefficients A_4, D_4, D_3, D_2, D_1 corresponding to the EEG frequency bands δ (0.5–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), and γ (30 ~ Hz) [10], [11], [43].

2.7.2 Power Spectra of EEG Sub-bands

Power spectra is an indicator of power in a certain signal in terms of frequency [125]. It is the most commonly used feature in Neuromarketing studies as presented in Table 2.5. The asymmetry in EEG band power is an indicator of consumer preferences, which is why it is widely used in Neuromarketing [43], [47], [69], [73]. To determine the power spectra of EEG sub-bands, firstly a Fast Fourier Transform (FFT) algorithm is used to calculate the Discrete Fourier Transform (DFT) of the EEG signal sequence $x(n)$ as given by Equation (2) [45], [56], [87], [105]. Such Fourier analysis is used to convert a signal from its original time domain to frequency domain. The DFT decomposes the signal sequence $x(n)$ into its components of different frequencies.

$$X_k = \sum_{n=0}^{N-1} x_n . e^{-\frac{i2\pi kn}{N}} \qquad (2)$$

where X_k represents the frequency component at index k, x_n is the input value at index n, and N is the data size. Then different periodogram techniques such as Welch's Periodogram [126] have been used to calculate the power spectral density (PSD) of different EEG sub-bands as given by Equation (3) [13], [43], [47], [48], [81], [89].

$$P_f = \frac{1}{N} \sum_{n=0}^{N-1} |X_n(k)|^2 \qquad (3)$$

where P_f presents the PSD of $X_n(k)$ corresponding to the nth segment and the kth frequency point after windowing. The relative power of different EEG frequency bands δ (0.5–4 Hz), θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), and γ (30 ~ Hz) are calculated this way, and extensively used as features in Neuromarketing applications. Some studies used Global Field Power (GFP) to get the spatiotemporal standard deviation of activity of each EEG time point [20], [41], [65], [78].

2.7.3 ERP

ERPs are voltage variations in EEG activity that correspond to motor, cognitive, and sensory events in real-time. They use the synchronized activity of neural populations to categorize and identify linguistic, memory, and perceptual processes. ERPs are used to extract event-related activity that is hard to discern from continuous EEG activity by averaging the electrocortical responses that occur throughout each event repetition [127]. Various Neuromarketing studies have linked ERPs to consumers' purchasing behavior, as shown in Table 2.5.

ERPs can generally be expressed as small spikes in brain activity that occur in response to a specific stimulus. These spikes have extremely low amplitudes, which makes it necessary to average EEG samples over many iterations to recognize ERPs and remove noise oscillations [119]. Table 2.6 presents the most common ERPs used in Neuromarketing research.

Table 2.6 ERPs used in Neuromarketing

ERP	Brain Region	Interpretation	Neuromarketing Studies
N2/N200	Central, frontal, and fronto-central regions	Indicates consumers' conflicts during decision making process	Yang et al. [61], Goto et al. [63], Wang et al. [67], Goto et al. [71], Jin et al. [72], Gong et al. [12], Fu et al. [80], Wang et al. [90], Yu et al. [92], Yang and Kim [114].
P2/P200	Prefrontal region	Indicated positively correlation between consumers' attention and negative stimuli	Jin et al. [72], Gong et al. [12], Gholami Doborjeh et al. [76], Sänger [82], Zhao and Wang [88], Ma et al. [91], Wang et al. [95], Wei et al. [100].
P3/P300	Posterior region	Indicates both decision-making conflicts and the difficulty in reaching a decision	Gong et al. [12], Sänger [82], Zubair et al. [85].
Late Positive Potential (LPP)	Centro-parietal region	Indicates intense attention to significant stimuli	Shen et al. [55], Goto et al. [63], Goto et al. [71], Jin et al. [72], Fu et al. [80], Zhao and Wang [88], Wang et al. [90], [95], [98], Ma et al. [91], [128], Yu et al. [92], Wei et al. [100], Bosshard and Walla [103], Song et al. [104].
N400	Centro-parietal region	Indicates negative correlation between engagement and stimuli	Domracheva and Kulikova [58], Yang et al. [61], Song et al. [104], Camarrone and Van Hulle [115].
Positive Slow Waves (PSW)	Fronto-central and parietal regions	Indicates working memory functions	Goto et al. [63], Goto et al. [71].

N200

The N2/N200 is an ERP component that shows a negative voltage. It occurs around 200 ms after the stimulus onset and is mainly located in the central, frontal, and fronto-central areas of the brain [129]. N200 is an indicator of the decision maker’s conflict monitoring during the decision-making process [130]. Goto et al. [71] reported that observing consumer products that were highly preferred produced greater positive N200 amplitudes than watching products that were less preferred. Gong et al. [12] investigated the effect of different online sales promotions on perceived risk and found a positive correlation with N200 amplitude. Fu et al. [80] found that the truthful condition resulted in a weaker N2 response when compared to the deceptive condition. This indicates a lower cognitive and decisional conflict.

P200

P2/P200 is an ERP component that appears early in the decision-making process, approximately 200 ms after the stimulus onset. The prefrontal area contains the majority of this component [131]. Gong et al. [12] used the P200 component to demonstrate how consumers initially evaluate the effectiveness of the stimulus. They found that people pay more attention to negative stimuli, and there is a positive correlation between the subjects' attention to negative stimuli and the P200 amplitude. In certain situations, the components of P200 and N200 can overlap, creating a P200/N200 complex [72], [132]. To quantify the P200/N200 complex in such situations, peak-to-peak scores are utilized [63], [133]. Goto et al. [63] extracted the positive and negative peak amplitudes from two specific time windows, 130–230 ms for P200 and 200–400 ms for N200. The peak-to-peak measurements of N200 were computed for each electrode by subtracting the P200 peaks from the N200 peaks.

P300

The P3/P300 wave is a brain activity that peaks between 300 ms and 400 ms and has a generally positive voltage [134]. It can be observed in various areas of the brain, and larger amplitudes are usually found in the posterior brain [135]. P3 can indicate both decision-making conflicts and the difficulty in reaching a decision. Both N2 and P3 are related to conflict resolution and serve as indicators of the degree of conflict in decision-making processes [12], [136]. In addition, P3 also reflects the difficulty and confidence involved in making decisions [12]. As decision-making difficulty increases, the P3 amplitude decreases, while decision-making confidence correlates positively with the P3 amplitude [12], [137]. P300 also displays both attention activity in working memory and response adjustment [119]. Zubair et al. [85] observed a higher value of the P300 component in positive framing messages, indicating that non-threatening emotional information is processed with more attention.

LPP

The LPP is a positive deflection that is frequently observed between 400 and 800 ms, with its highest point generated from the centro-parietal region of the brain. It can be linked to attention toward emotional stimuli with positive or negative valence [119]. It indicates an intense attention to significant stimuli. This deflection is often observed in response to visual information that conveys emotional content [63], [138]. Ma et al. [128] reported that the larger LPP amplitudes for the prices of products were positively correlated with buying intentions. Fu et al. [80] discovered that the truthful condition elicited a greater LPP response than the deceptive condition. This suggests less cognitive and decisional conflict and a more positive evaluation of the truthful condition. The service with higher emotional arousal elicits a greater LPP amplitude compared to the service with lower emotional arousal [88].

N400

N400 is a negative deflection that reaches its maximum around 400 milliseconds after the stimulus onset and is typically observed over the centro-parietal electrode sites [58]. Researchers use this deflection to study the effects of brand familiarity and brand extension services in Neuromarketing [61], [115]. It has been found that a large N400 response indicates low association strength in the customer's mind, and vice versa [115].

PSW

The PSW is a positive deflection that occurs after 800 ms and can last up to 3 seconds after a visual stimulus is presented [71]. It is commonly observed in both fronto-central and parietal sites and indicates a relatively prolonged form of processing that may involve working memory functions [63], [139]. Goto et al. [63] observed higher prediction accuracy rates with later positivities like the LPP and PSW compared to the N200 of the frontal area.

2.7.4 EEG Asymmetry Indices

Table 2.5 shows that many studies used brain activation asymmetry indices as features to predict consumer preferences. I have found four different types of asymmetry indices in the literature that can be used as autonomic indicator of consumers' preferences – (1) Approach-Withdrawal (AW) Index, (2) Valence Index, (3) Effort Index, and (4) Choice Index.

AW Index

The AW index is a measure of frontal alpha asymmetry, which indicates the difference in activations between the left and right hemispheres. It estimates desire and motivation by assessing alpha's higher activation in the right frontal cortex [43], [50], [60], [117], [119], [140]. Many studies have exhibited the effectiveness and accuracy of FAA as a crucial factor in emotion and Neuromarketing research [20], [47], [50], [60], [66], [97], [102], [117], [121]. To calculate the AW index, we can use electrodes F4 and F3 to determine the difference between the right and left PSD using either equation (4) [50], [140] or equation (5) [50], [117].

$$AW_1 = \alpha(F4) - \alpha(F3) \tag{4}$$

$$AW_2 = \frac{\alpha(F4) - \alpha(F3)}{\alpha(F4) + \alpha(F3)} \tag{5}$$

Valence Index

Studies have found a link between frontal asymmetry and a customer's emotional state [20], [47], [50], [60], [66], [97], [117], [121]. Specifically, left frontal activation is associated with positive valence, while right frontal activation is associated with negative valence. Several studies have supported the theory that frontal EEG asymmetry is an indicator of valence [43], [47], [50], [119]. It can be calculated by either of the equations (6) [50], [121], (7) [50], [121], or (8) [47], [141].

$$Valence_1 = \frac{\beta(F3)}{\alpha(F3)} - \frac{\beta(F4)}{\alpha(F4)} \tag{6}$$

$$Valence_2 = \ln[\alpha(F3)] - \ln[\alpha(F4)] \tag{7}$$

$$Valence_3 = \frac{\alpha(F4)}{\beta(F4)} - \frac{\alpha(F3)}{\beta(F3)} \tag{8}$$

Choice Index

The choice index is a measure based on the frontal asymmetric gamma and beta oscillations, which are primarily associated with the actual decision-making process. It is also highly correlated with willingness-to-pay responses, particularly in the gamma band, which is used to assess consumer preference and choice [60]. Greater activation in the left prefrontal region is indicated by higher values in the gamma and beta bands, whereas the right prefrontal region is associated with considerably stronger activation at lower levels [47], [50], [60]. Choice index using gamma band and beta band can be expressed by equation (9) and (10) [47], [60].

$$Choice\ Index_{\gamma} = \frac{\log\left(\gamma\left(Electrode_{left}\right)\right) - \log\left(\gamma\left(Electrode_{right}\right)\right)}{\log\left(\gamma\left(Electrode_{left}\right)\right) + \log\left(\gamma\left(Electrode_{right}\right)\right)} \tag{9}$$

$$Choice\ Index_{\beta} = \frac{\log\left(\beta\left(Electrode_{left}\right)\right) - \log\left(\beta\left(Electrode_{right}\right)\right)}{\log\left(\beta\left(Electrode_{left}\right)\right) + \log\left(\beta\left(Electrode_{right}\right)\right)} \tag{10}$$

Effort Index

The activity level of the frontal theta in the prefrontal cortex is a measure that indicates the cognitive processing that arises from mental exhaustion. Higher levels of theta activity are associated with harder and more complicated tasks in the frontal area. This measure has been frequently studied in Neuromarketing research [81], [105], [120], [142], [143], [144]. We can use equation (11) to represent the effort index [105].

$$Effort\ Index = \frac{10\ \log\left(\theta\left(Electrode_{left}\right)\right) - \log\left(\theta\left(Electrode_{right}\right)\right)}{\log\left(\theta\left(Electrode_{left}\right)\right) + \log\left(\theta\left(Electrode_{right}\right)\right)} \tag{11}$$

2.7.5 Statistical Features

Different types of statistical features have been utilized in the literature as listed in Table 2.5. Mostly used statistical features are – Mean, SD, Variance, Skewness, Kurtosis, and Differential Entropy (DE) as given by the equations (12) – (17) respectively, where X_i and N represents the EEG signal and total number of samples respectively.

$$Mean\ (\mu) = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{12}$$

$$SD\ (\sigma) = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}} \tag{13}$$

$$Variance\ (var) = \sigma^2 \quad (14) \qquad Skewness = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i - \mu}{\sigma}\right)^3 \quad (15)$$

$$Kurtosis = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i - \mu}{\sigma}\right)^4 \quad (16) \qquad DE = \frac{1}{2}\log(2\pi e {\sigma^2}_i) \quad (17)$$

Apart from the above-mentioned features, some studies also used Hjorth parameters as features in Neuromarketing [47], [145]. To analyze the EEG signal in the time domain, Hjorth developed the Hjorth parameters [146]. These parameters consist of three measures: activity, mobility, and complexity as presented by equations (18) – (20) respectively. Activity measures signal amplitude deviation, mobility measures slope changes, and complexity measures amplitude standard slope count [47].

$$Activity = var(X) \quad (18) \qquad Mobility = \sqrt{\frac{var(X')}{var(X)}} \quad (19)$$

$$Complexity = \frac{Mobility(X')}{Mobility(X)} \quad (20)$$

where X represents the EEG signal, *var* presents the variance, and X' is the first derivative of the signal.

2.8 DATA INTERPRETATION TECHNIQUES IN NEUROMARKETING

The studies used either machine learning or statistical analysis to interpret data. They employed advanced machine learning algorithms to classify consumers' like/dislike decisions. On the other hand, various statistical analysis techniques were used for correlational and behavioral analyses. It is crucial to select appropriate ML or statistical analysis techniques to accurately interpret the data.

2.8.1 Machine Learning Applications in Neuromarketing

The Neuromarketing experiments utilized both supervised and unsupervised learning methods. In supervised learning, a priori ground truth, which is usually the interviewed response of the test subjects (like/dislike), is used as the labels. In the training datasets, these labels aid the classifier in recognizing the signal sequence of like and dislike. In the testing stage, a label-free dataset is used to predict the like/dislike responses. On the contrary, previous knowledge of the like/dislike labeling is not necessary for the unsupervised learning methods. The most commonly used machine learning methods are listed in Table 2.7. Supervised learning methods include the Hidden Markov Model (HMM), Support Vector Machine (SVM), Neural Network (NN), Random Forest (RF), Regression, and Linear Discriminant Analysis (LDA), while k-Nearest Neighbors (k-NN) is the only unsupervised learning method.

Table 2.7 Machine learning techniques used in Neuromarketing

Machine Learning Techniques	**Neuromarketing Studies**	**Average Classification Accuracy**
HMM	Yadava et al. [10]	70.33% (male) & 63.56% (female) [10]
SVM	Hakim et al. [40], Aldayel et al. [13], [43], Li et al. [45], Zeng et al. [47], Georgiadis et al. [53], Wei et al. [59], Gholami Doborjeh et al. [76], Golnar-Nik et al. [81], Yang and Kim [114].	80.28% [45], 75% [59], 87% [81]
k-NN	Hakim et al. [40], Aldayel et al. [13], [43], Zeng et al. [47], Raiesdana and Mousakhani [48], Georgiadis et al. [53], Avinash et al. [105].	94.22% [47], 92.4% [48], 82.75% [105]
NN	Hakim et al. [39], Guixeres et al. [41], Teo et al. [42], Aldayel et al. [43], Al-Nafjan [50], Göker [54], Gholami Doborjeh et al. [76], Özbeyaz [113].	75.09% [39], 82.9% [41], 79.76% [42], 99% [50], 96.83% [54] 90% [76], 72% [113]
RF	Aldayel et al. [13], [43], Kumar et al. [11], Al-Nafjan [50], Hassani et al. [52], [99].	100% [50], (71 .51 ± 5 .1%) [52],

		96.47% (female) & 95.32% (male) [99]
Regression	Hakim et al. [40], Kislov et al. [49], Gholami Doborjeh et al. [76], Izadi et al. [15], Garczarek-Bąk et al. [112].	61.2% [112]
LDA	Golnar-Nik et al. [81], Avinash et al. [105].	87% [81], 90% [105]

In supervised learning techniques, the HMM is a non-linear classifier that is commonly used in biomedical and spatial signals. It originated from statistical modeling and helps classify multiple classes of consecutive data in Neuromarketing research. The HMM helps researchers identify potential states for observation by using state transition probabilities when there may be a transition from one mental state to another [1]. Yadava et al. [10] proposed an HMM-based consumer preference prediction model that outperformed other common classifiers including SVM, RF, and ANN. The model achieved a classification accuracy of 70.33% for male participants and 63.56% for female participants.

SVM is another supervised machine learning algorithm that uses training data to recognize patterns and derive relationships. SVM classifies by creating a hyperplane that separates different classes. The hyperplane is based on the training data and is used to classify new data. The accuracy and simplicity of computation of SVM make it a useful tool in Neuromarketing [1]. Li et al. [45] combined EEG and eye-tracking (ET) data for product design evaluation and obtained average classification accuracy of 80.28% using SVM classifier. Numerous studies employ LDA classifiers in contrast to SVM classifiers [81], [105]. LDA organizes several data points with comparable frequencies into discrete classes, and 1D Eigen transformation establishes these classes [1]. Golnar-Nik et al. [81] used SVM along with LDA to investigate the effects of changes in promotions' contents on consumers' minds and achieved an average classification accuracy of 87%.

The unsupervised learning model k-NN can act as a regression and classification tool. Based on the K training samples that are the test sample's closest neighbors, the k-NN algorithm predicts the category of the test sample. Unlike SVM's hyperplane, k-NN establishes a decision boundary between several separate classes [1]. Using k-NN algorithm for the preference classification tasks, Raiesdana and Mousakhani [48] obtained an average classification accuracy of 92.4% for electric cars, and Zeng et al. [47] obtained 94.22% for sport shoes.

Different NN techniques have been utilized in the literature for classification purposes. The most commonly used NN techniques are Spiking Neural Network (SNN) [76], Artificial Neural Network (ANN) [113], DNN [39], [42], [50], and Multi-layer Perceptron (MLP) [76]. They use neurons or distributed nodes arranged in a stacked framework to mimic the structure of the human brain. They generate non-linear boundaries to make decisions across vast data sets. Computers can utilize this adaptive approach to learn from their errors and keep getting better. Even though they are becoming popular for data interpretation in Neuromarketing, they require a high number of features and data samples [1]. Guixeres et al. [41] explored whether NN models could be used to forecast the effectiveness of a new advertisement on digital channels, and the study's average classification accuracy was 82.9%. Teo et al. [42], Al-Nafjan [50], and Hakim et al. [39] used DNN for the preference prediction tasks and achieved average classification accuracy of 79.76%, 99%, and 75.09% respectively.

A popular ensemble machine learning approach called random forest aggregates the output of several decision trees to produce a single outcome. Its versatility and ease of use, combined with its ability to address both regression and classification problems, have driven its popularity in Neuromarketing applications. When examining how color affects consumer choices, Hassani et al. [99] employed the RF algorithm to predict preferences. For female participants, this resulted in an average classification accuracy of 96.47%, whereas for male participants, it was 95.32%. Al-Nafjan [50] used RF algorithm with optimal features and obtained an average classification accuracy of 100%.

2.8.2 Statistical Analysis in Neuromarketing

Biostatistics offers a range of statistical methods to analyze and interpret data for specific situations. It's essential to understand the assumptions and conditions of the statistical techniques to select the appropriate method for data analysis. The two main statistical methods are descriptive statistics and inferential statistics. Descriptive statistics summarize data using indexes such as mean, median, and standard deviation. Inferential statistics draw conclusions from the data using statistical tests such as t-test, Analysis of Variance (ANOVA) test, etc. Parametric and nonparametric are the two possible classifications for inferential statistical techniques. Statistical methods used for comparing means are referred to as parametric. In contrast, those used for comparing other variables like median, mean ranks, or proportions are referred to as non-parametric methods [147]. In the literature, a range of statistical analysis techniques have been employed for conducting correlational and behavioral analyses. The most commonly used statistical analysis techniques in Neuromarketing are summarized in Table 2.8.

Table 2.8 Statistical analyzing techniques used in Neuromarketing

Statistical Analyzing Techniques	**Description**	**Neuromarketing Studies**
t-test	To compare the means of two paired groups (paired samples t-test) or unpaired groups (independent samples t-test)	Zeng et al. [47], Raiesdana and Mousakhani [48], Georgiadis et al. [53], Uhm et al. [56], Wajid et al. [19], Gountas et al. [57], Ramsøy et al. [60], Yang et al. [61], Harris et al. [62], Eijlers et al. [64], Jin et al. [72], Çakar et al. [73], Gong et al. [12], Verhulst et al. [84], Hsu and Chen [86], Hsu and Chen [23], Martinez-Levy et al. [20], Ma et al. [91], Yu et al. [92], Hsu and Chen [107], Pennanen et al. [108], Yang and Kim [114].

ANOVA	To compare the means of three or more unpaired (one-way ANOVA) or paired groups (repeated-measure ANOVA)	Guixeres et al. [41], Oikonomou et al. [44], Raiesdana and Mousakhani [48], Shen et al. [55], Uhm et al. [56], Ramsøy et al. [60], Yang et al. [61], Harris et al. [62], Goto et al. [63], Vozzi et al. [65], Zito et al. [66], Wang et al. [67], Ma et al. [68], Goto et al. [71], Jin et al. [72], Royo et al. [74], Gordon et al. [78], García-Madariaga et al. [79], Fu et al. [80], Sänger [82], Alvino et al. [83], Verhulst et al. [84], Zubair et al. [85], Zhao and Wang [88], Wang et al. [90], Ma et al. [91], Yu et al. [92], Kim et al. [94], Wang et al. [95], Alvino et al. [96], Russo et al. [97], Paula et al. [101], Bosshard and Walla [103], Song et al. [104], Pennanen et al. [108], Ma et al. [111].
Chi-square test	To compare the associations between two or more independent groups	Alonso Dos Santos and Calabuig Moreno [77], Hsu and Chen [86], [107], Garczarek-Bąk et al. [112].
Friedman test	To compare the medians/ interquartile ranges of three or more paired groups	Alvino et al. [83], Izadi et al. [15].
Wilcoxon Rank test	To compare the medians/ interquartile ranges of two paired (signed-rank test) or unpaired groups (sum test)	Domracheva and Kulikova [58], Kakaria et al. [69], Alonso Dos Santos and Calabuig Moreno [77].
Mann Whitney test	To compare the medians/ interquartile ranges of two unpaired groups	Domracheva and Kulikova [58], Pagan et al. [87], Garczarek-Bąk et al. [112].
Kruskal-Wallis test	To compare the medians/ interquartile ranges of three or more unpaired groups	Alonso Dos Santos and Calabuig Moreno [77].
Levene's and Shapiro-Wilk's test	To compare the variances between samples to verify whether or not a sample fits a normal distribution	Guixeres et al. [41], Alonso Dos Santos and Calabuig Moreno [77], Damião de Paula et al. [101], Russo et al. [102], Leeuwis et al. [109].
Descriptive statistics	To interpret data using indexes such as mean, median, and SD	Uhm et al. [56], González-Morales [18], Hsu and Chen [23].

The most frequently utilized parametric inferential statistical analysis methods are the t-test and ANOVA. The t-test is used to compare the means of two paired or unpaired groups. When it is used to compare two paired groups, it is called a paired samples t-test; when it is used to compare two unpaired groups, it is called an independent samples t-test. Zeng et al. [47] used two-sample t-tests to determine the difference in the EEG's power between the consumers' like and dislike decisions. Wajid et al. [19] employed the mean t-tests to investigate the variations in the message appeals of two ads. Çakar et al. [73] performed a t-test analysis on all the FAA-based data points to mark the values at a 95% significance level. The ANOVA is used to compare the means of three or more paired or unpaired groups. When it is used to compare paired groups, it is called one-way ANOVA; when it is used to compare unpaired groups, it is called repeated-measure ANOVA. Oikonomou et al. [44] performed a one-way ANOVA test to examine the impact of the classification techniques on accuracy values. Goto et al. [71] used repeated-measure ANOVA to investigate how the ERP waveforms were affected by subsequent purchasing decisions. Zubair et al. [85] used both one-way ANOVA and repeated-measure ANOVA on the ERP components to observe the effects of three different message-framing contents on consumers' minds.

The most common nonparametric inferential statistical analysis methods are Chi-square test, Friedman test, Wilcoxon Rank test, Mann Whitney test, Kruskal-Wallis test, and Levene's and Shapiro-Wilk's test. Among them, the Chi-square test compares the proportions between different groups. In contrast, the other methods compare the medians/ interquartile ranges of different group patterns, as listed in Table 8. Kakaria et al. [69] observed the variations in cognitive stress during planned and unplanned purchases using the Wilcoxon Signed-Ranks test. Pagan et al. [87] divided study participants into two groups, recorded their EEG data, and performed the Mann-Whitney test to compare brainwave differences. Izadi et al. [15] employed the Friedman test to evaluate the difference in subjects' alpha wave amount when they were exposed to different promotions. Garczarek-Bąk et al. [112] evaluated the predictive power of their model for private label purchasing using the Chi-square test. Different studies utilized Levene’s and Shapiro-Wilk’s tests to verify whether or not their samples fit normal distributions [101], [109]. Apart from these, descriptive statistics such as mean, median, and standard deviation were used by some studies to interpret the data [18], [23], [56].

2.9 SUMMARY AND IMPLICATIONS

A synthesized summary of the reviewed research works are presented in Appendix A. The systematic review presents the current research trend, data collection, data processing, and analyzing techniques in Neuromarketing. It shows that many approaches have been made to build a proper EEG-based preference prediction system for target consumers over the years. However, the current status of EEG-based preference prediction and its classification accuracy is still below optimal. The performance of EEG-based preference detection systems depends on a suitable selection of pre-processing pipelines and proper labelling of the data so that machine learning algorithms can be trained accurately. Also, the research trend shows that Neuromarketing research has evolved from building a theoretical base to implementing neuroscience-based marketing strategies. Modern studies are trying to develop marketing strategies by observing how subtle changes affect consumers' preferences in practice. Among the prominent components of promotions, language significantly impacts consumers' minds. In a bilingual nation like ours, which advertisement language will work better remains unchecked.

Therefore, this research aims to develop an optimal EEG pre-processing pipeline, propose an objective labeling method for EEG-based preference prediction in Neuromarketing and compare the results with subjective labeling, and finally investigate the effects of languages on consumers' preferences in Neuromarketing.

3 Data Collection

This chapter discusses the datasets used in this work. I utilized two EEG databases in my research. Dataset 1 is the publicly available Neuromarketing dataset and we created the Dataset 2 during my own experiment. Section 3.1 describes the Dataset 1 and section 3.2 discusses Dataset 2.

3.1 DATASET 1

This dataset has been made publicly available since 2017 in a Neuromarketing study [10]. It one of the most commonly used public datasets for Neuromarketing research, and was used in several studies [10], [43], [51], [54], [119], [148] and can be found at https://sites.google.com/site/iitrcsepradeep7/resume. Twenty-five participants were involved in this study, where EEG data was recorded using the Emotiv EPOC+ headset. The participants viewed 14 different products on a computer screen, each of which had three variations. In total, 42 product images (14x3) were generated, and EEG data was recorded for all users, resulting in 1050 EEG data sets. The age range of the participants was between 18 and 38 years old. The EEG data were collected from 14 channels located at AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 locations on the head. The data was down-sampled to 128 Hz and then used to create 25 documents, one for each user. Participants were asked to provide feedback on each product by choosing either "like" or "dislike". EEG data was recorded while each product was displayed on the screen for four seconds, and the user's preferred choice was collected after each image was presented.

3.2 DATASET 2

I created the Dataset 2 during my own experiment. I have also made this dataset publicly available for future research works, and it can be found at https://github.com/MdFazlulKarimKhondakar. I have collected EEG data from human participants to make dataset 2. The study was approved by the proper authority of Chittagong University of Engineering & Technology, Bangladesh and conducted in accordance with the Declaration of Helsinki.

3.2.1 Participants

This study involved thirty-six undergraduate and postgraduate students, out of which eight were female. The participants were aged between 19 and 25 years old, with an average age of 22.5 years. They were all free from any neurological or psychiatric diseases and had normal or corrected-to-normal vision. Prior to the experiment, all participants provided written informed consent.

3.2.2 Stimuli

Various types of marketing stimuli have been utilized in literature to observe how consumers' brains respond when they make decisions about which products to purchase. The most frequently used stimuli are product images [10], [15], [42], actual products [89], [149], and video commercials [40], [150]. Many studies used product images as stimuli, as they have become popular in marketing due to the surge in E-commerce-based shopping. As the study involved university students who used to have a good understanding of electronic gadgets such as smartphones and laptops, I selected four smartphones and four laptops with different specifications to create marketing stimuli for this study. I created separate images for each product, displaying the specifications around the product's picture as it appears on E-commerce sites. I deliberately excluded the brand names and prices of the products as these factors have a significant impact on consumers' minds [150], [151]. For each product, I created two different images – one in the native language, Bengali and another in a foreign language – English, maintaining the same information in both languages. All images were kept at a uniform size of 1280 by 720 pixels to maintain consistency throughout the study.

3.2.3 Procedure

Prior to conducting the experiments, the participants were asked to indicate which product they were more familiar with, smartphones or laptops. Twenty-four subjects gave their consent for smartphones, while the other twelve subjects gave consent for laptops. The experiments were conducted in a dimly lit and electrically shielded room, with each subject sitting about 100 cm away from the computer screen. The participants were given a mouse and keyboard to proceed with the experiments. An experimental set-up for data collection is presented in Fig. 3.1.

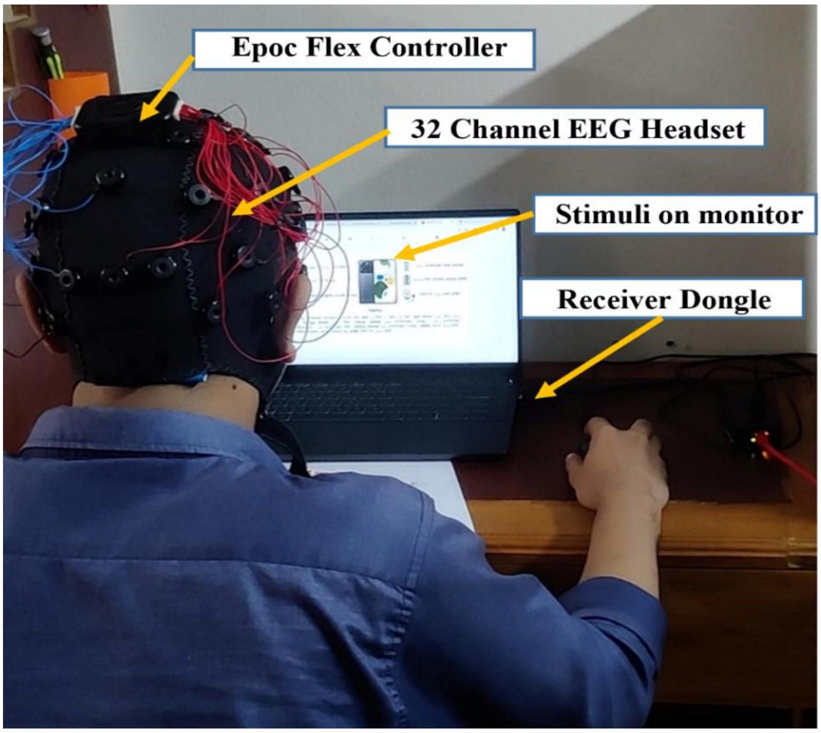


Fig. 3.1 An experimental setup for data collection.

I designed a total of sixteen experiments, eight for laptops (four in Bengali and four in English) and eight for smartphones (four in Bengali and four in English), using Emotiv Builder software (Emotiv Inc., San Francisco, California, U.S.A). The experiment's stages are shown in Fig. 3.2. During the calibration stage, I recorded the subject's brain data for 5 seconds with their eyes open looking at a fixation sign, followed by another 5 seconds with their eyes closed, and then provided general instructions. Next, the subjects viewed the stimuli image for as long as they needed to make a decision on the product. They then provided feedback on a 5-point Likert scale ranging from "strongly dislike" to "strongly like". During the study, participants were exposed to four experiments, either in their native language or a foreign language. The first experiment was considered a practice trial to familiarize them with the procedure and was excluded from the dataset. After a minimum interval of 4 weeks, the exact product stimuli were shown to the subjects in the alternate language. The order of languages among the subjects was randomized to ensure unbiased data. In total, there were 216 recordings (36*6) in the final dataset, with each subject having six recordings - three in native language stimuli and three in foreign language stimuli.

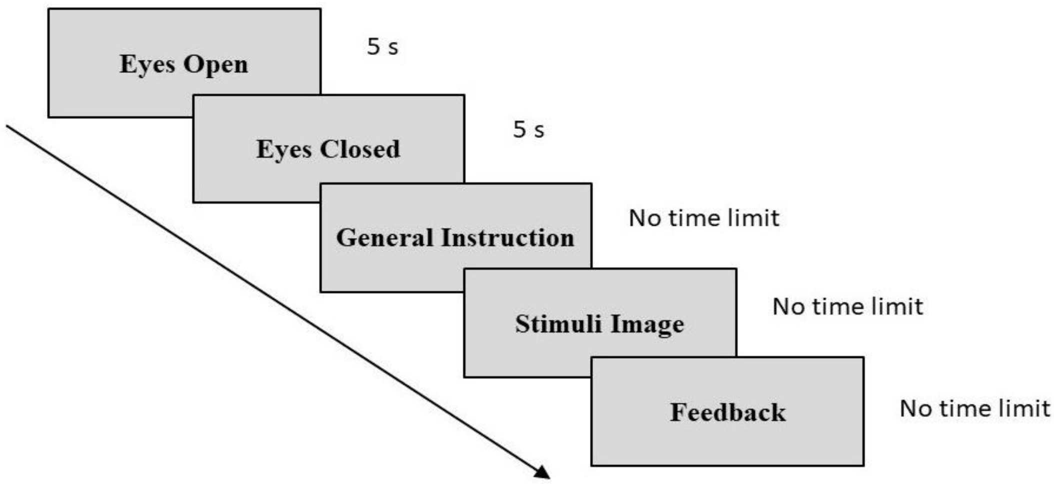


Fig. 3.2 Experiment stages for one stimulus.

3.2.4 EEG Recordings

To collect the EEG data, I used the Emotiv EPOC Flex headset (Emotiv Inc., San Francisco, California, U.S.A). It has configurable 32 Ag-AgCl electrodes according to the international 10/20 EEG system and 2 reference electrodes, CMS and DRL, located at the left and right mastoid, respectively. It has a resolution of 16 bits and a sampling frequency of 128 Hz. It provides high pass filtering of the data by 0.2 Hz, low pass filtering by 45 Hz, and high attenuation at 50 Hz and 60 Hz. I collected data from 16 channels (Cz, Fz, Fp1, F7, F3, C3, P3, O1, Pz, Oz, O2, P4, C4, F4, F8, and Fp2) which have been previously shown to be significant in preference prediction [7], [10], [47]. The channel locations are presented in Fig. 3.3.

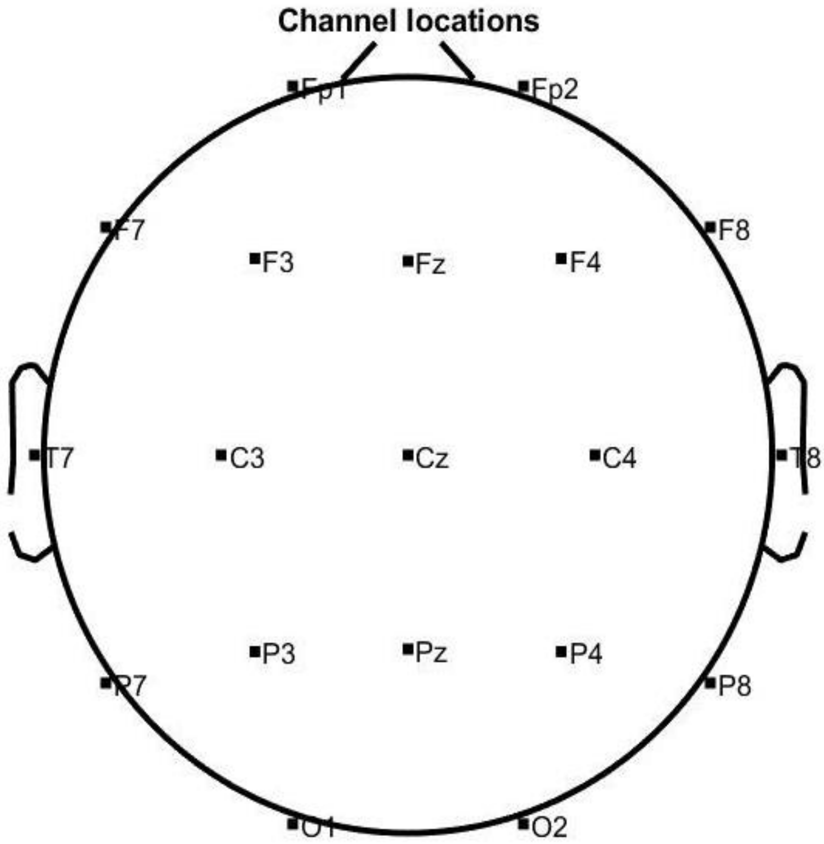


Fig. 3.3 Electrodes placement for data collection according to 10/20 system.

4 Pre-processing Pipelines

In this chapter, I draw a comparative analysis of different pre-processing pipelines and proposed an optimal pre-processing pipeline for EEG-based preference prediction in Neuromarketing. Section 4.1 describes the background, section 4.2 discusses the methodology followed, section 4.3 presents the results and discussions, and section 4.4 draws the pertinent conclusions.

4.1 BACKGROUND

EEG signals are highly sensitive to noise, such as Electrocardiogram (ECG) signals that come from the heart, Electrooculogram (EOG) signals produced by eye movements, Electromyogram (EMG) signals generated by muscle contractions, and power line interference (50/60 Hz) caused by the frequency of power lines. Pre-processing is a crucial stage that eliminates these noises and prepares the signal within the desired frequency band for further processing. In most studies found in the literature, a conventional pre-processing pipeline is utilized to eliminate artifacts from raw EEG data [16], [20], [40], [150]. An advanced automated EEG pre-processing pipeline has also been recently introduced [21]. However, it is still uncertain whether these pre-processing pipelines are sufficient in providing more accurate predictions of consumer preferences. It remains unclear which pipeline future researchers should use for EEG-based research in Neuromarketing.

In this chapter, I have developed an optimal pre-processing pipeline for EEG data and compared it with two other existing methods. After applying the three pre-processing pipelines, I extracted numerous statistical and frequency domain features, utilized various machine learning classification models, and provided recommendations for future researchers. The work has been published in [152].

4.2 METHODOLOGY

4.2.1 Data used

I used Dataset 2, described in the previous chapter, to compare the pre-processing pipelines. I used only the best and worst products' EEG data to avoid the effects of labelling and language effects; I used the data collected using English language stimuli. Therefore, the final dataset contained two recordings for each subject, yielding 72 (36*2) recordings. Fig. 4.1 presents a sample of raw data for two channels of a particular subject.

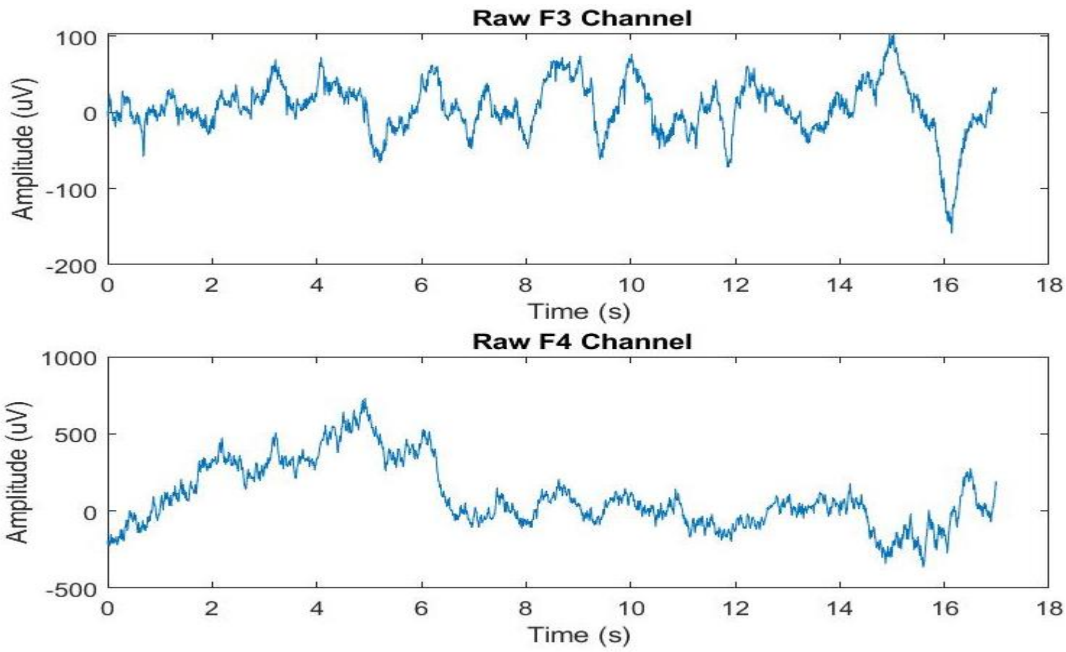


Fig. 4.1 Raw EEG data.

4.2.2 Traditional EEG pre-processing pipeline

I have used three different pre-processing pipelines to remove the noises from the data in MATLAB software. The first one is the traditional EEG pre-processing method as shown in Fig. 4.2. Most of the studies in the literature used this method to clean the raw EEG data [16], [19], [20], [43]. This pipeline starts with bandpass filtering (BPF) of the data within the desired EEG range followed by notch filtering (50 Hz/60 Hz) to remove the power line interferences, and Independent Component Analysis (ICA) to remove eye and muscle noises. The pre-processed EEG data using this method looks like the ones in Fig. 4.3.

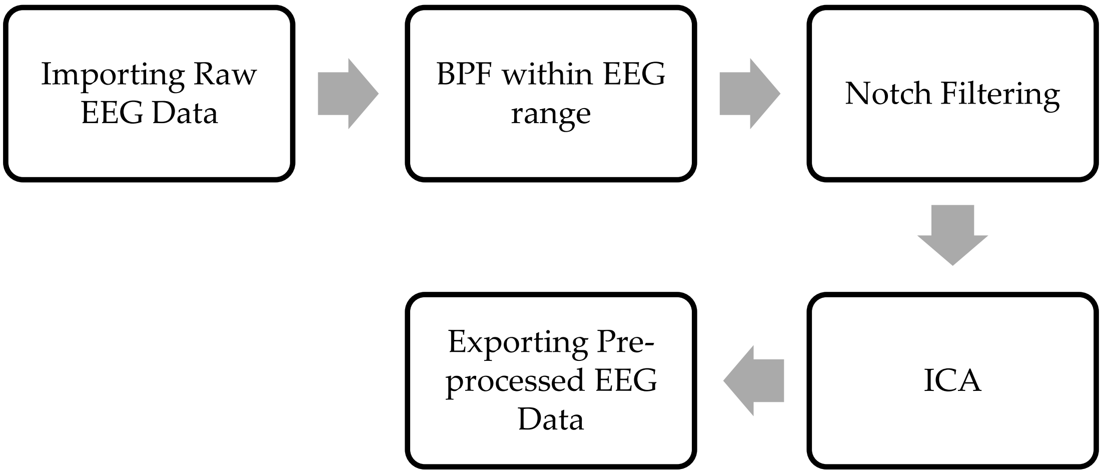


Fig. 4.2 Traditional EEG pre-processing pipeline.

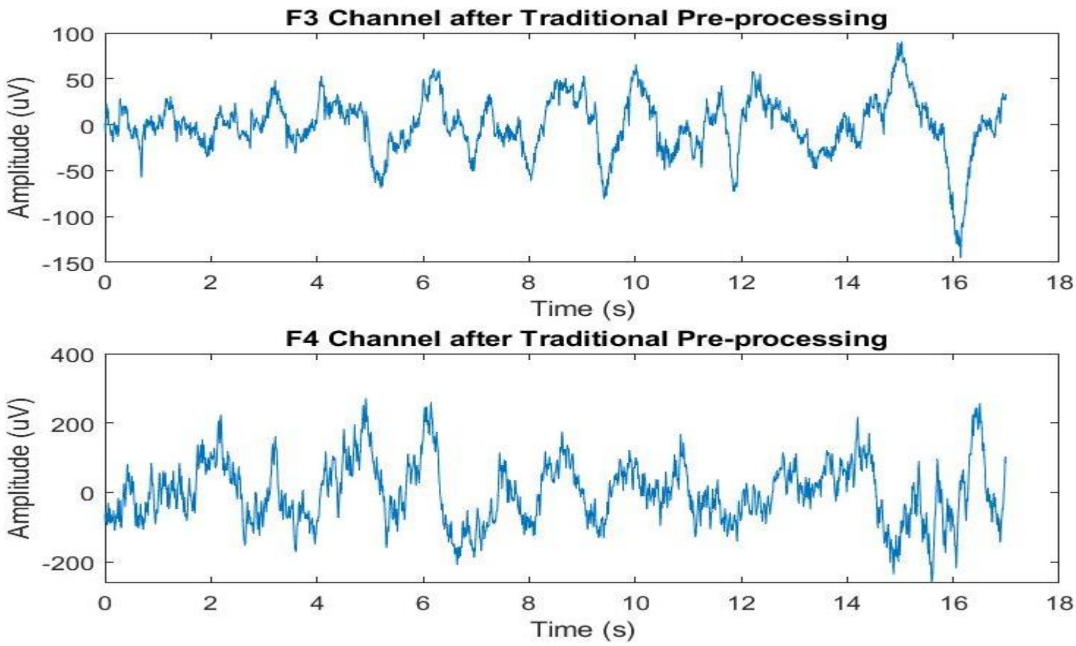


Fig. 4.3 Pre-processed data by the traditional pipeline.

The traditional method operates on a channel as a complete entity, and it is unable to filter out specific portions of the data that contain excessive noise. Since it works on the entire channel, it is unable to focus on a specific area of the channel.

4.2.3 Automated EEG pre-processing pipeline

The limitations of traditional pre-processing pipeline is addressed in the second EEG pre-processing pipeline which is an advanced automated pipeline introduced recently in the literature [21]. The pipeline uses EEGLAB [153] in MATLAB as per the workflow diagram in Fig. 4.4. I used EEGLAB 2022 in MATLAB 2021a to implement the pipeline.

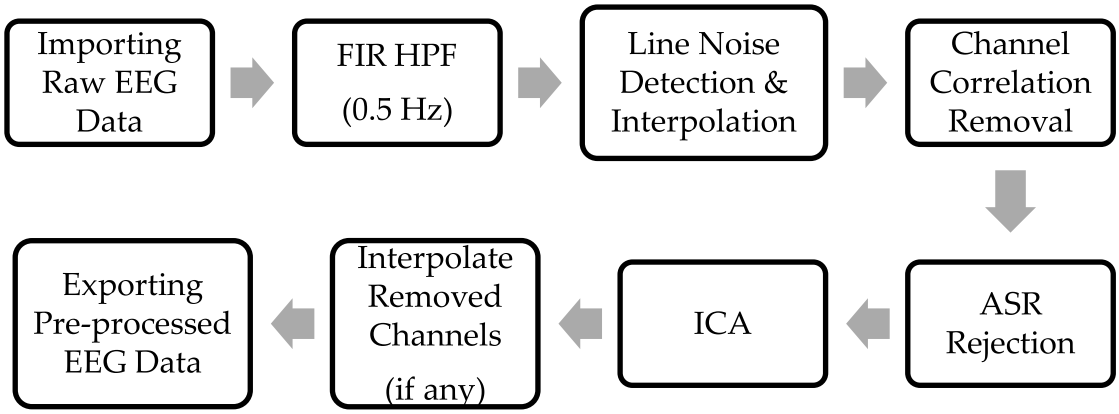


Fig. 4.4 Automated EEG pre-processing pipeline.

The pipeline starts with Finite Impulse Response (FIR) High Pass Filtering (HPF) of the data by 0.5 Hz to remove any baseline drift, followed by detecting and interpolating electrode line noise with a standard deviation (SD) threshold of 4. It removes channel correlation with a threshold of 0.9 and uses Artifact Subspace Reconstruction (ASR) rejection with a threshold of 20. It then applies ICA to remove muscle and eye artifacts. In the end, before exporting the pre-processed data, any removed channels in the data are interpolated. The pre-processed EEG data, using this method, are presented in Fig. 4.5.

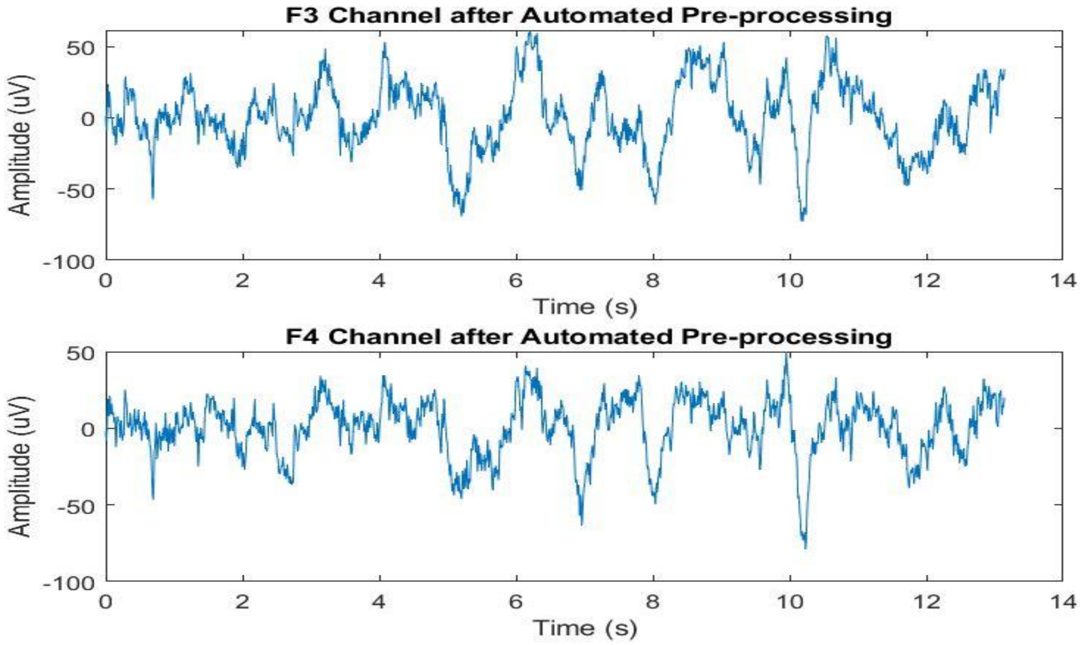


Fig. 4.5 Pre-processed data by the automated pipeline.

The automated EEG pre-processing pipeline is designed to initially remove channels with high levels of noise and interpolate them later. However, an interpolated channel doesn't contain any original information since its values are predicted from the potentials of the neighbouring electrodes. This process can result in the loss of necessary information. For instance, frontal electrodes are crucial in EEG-based preference prediction as they are used to determine essential features such as Frontal Alpha Asymmetry (FAA) indexes. If the automated pipeline removes any of the frontal electrodes and then interpolates them, the corresponding indexes will not be accurate. Moreover, the pipeline does not guarantee that the data are rank-full and ICA decomposition is free from bugs [154]. Therefore, it is important to carefully evaluate the results obtained from the automated pipeline and make necessary adjustments to ensure that the data is not compromised.

4.2.4 Optimal EEG pre-processing pipeline

In this study, I have proposed an optimal EEG pre-processing pipeline that addresses the limitations of both methods described above. The pipeline also uses EEGLAB in MATLAB, and the workflow diagram is displayed in Fig. 4.6.

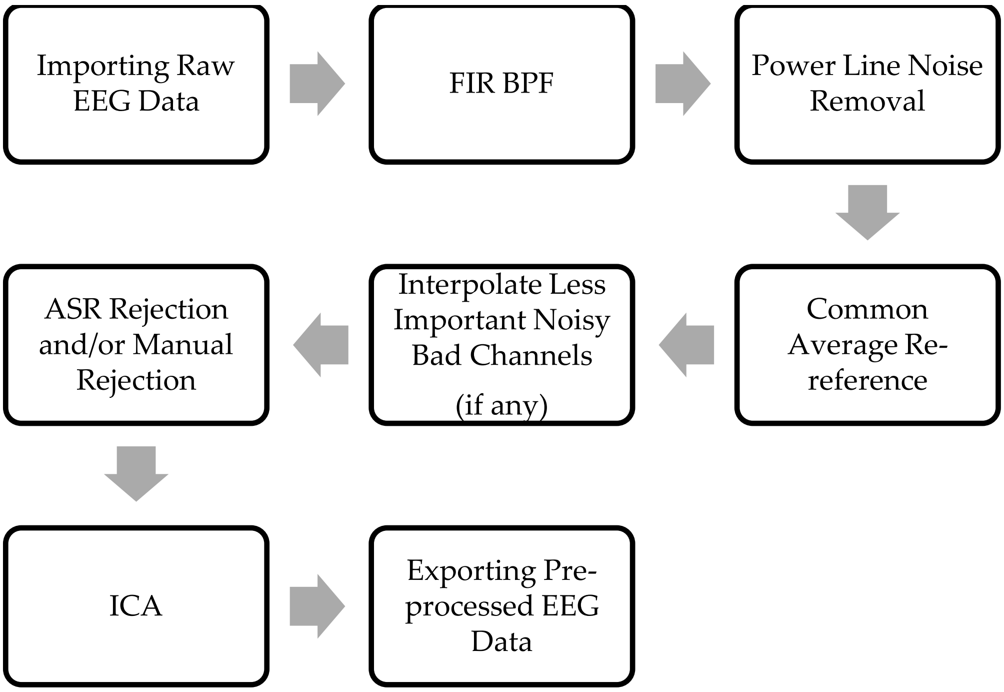


Fig. 4.6 Optimal EEG pre-processing pipeline.

I used EEGLAB 2022 in MATLAB 2021a to implement the pipeline. First, I applied the finite impulse response (FIR) bandpass filter (BPF) to the data within a range of 0.5-45 Hz and removed the power line noise (50 Hz) using the CleanLine plugin in EEGLAB[155]. This plugin uses multi-tapering and Thompson F-statistics to estimate and eliminate sinusoidal noises from the scalp channels, such as line noises[155]. Then, I re-referenced the data to the common average reference. I also checked whether any noisy bad channels are essential to the application of interest, and if not, I interpolated them subsequently. Next, I performed Artifact Subspace Reconstruction (ASR) rejection and/or manual rejection by examining the data. The ASR is a method for automatically rejecting artifacts that use principal component analysis (PCA) techniques[156]. In this method, the components with large variances are removed, and the signal is reconstructed using the remaining components. The ASR method employs a reference data set that is free of artifacts to set thresholds for determining which components to reject. This reference data set can be selected automatically. For my analysis, I used a threshold of 20 to reject components using the ASR method. If the ASR rejection marked a large portion of the data for removal, I ignored it and performed manual rejection by removing only the significantly noisy part. Finally, I applied Independent Component Analysis (ICA) to remove muscle and eye artefacts. The pre-processed EEG data obtained through this method are displayed in Fig. 4.7.

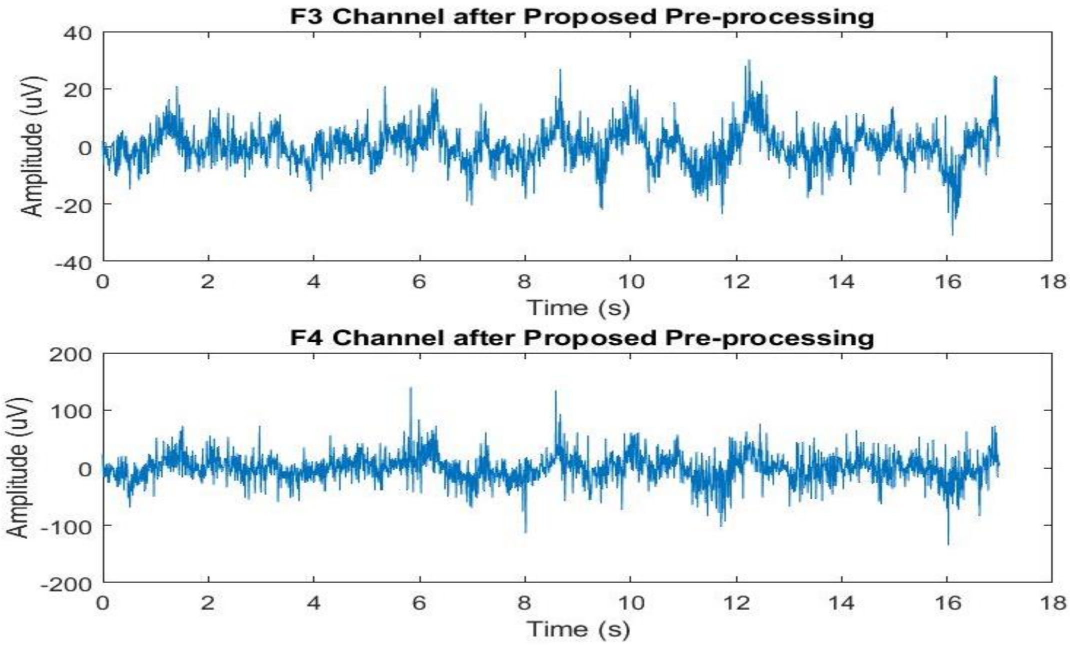


Fig. 4.7 Pre-processed data by the optimal pipeline.

4.2.5 Feature Extraction

The objective of feature extraction is to identify significant and pertinent data from EEG signals, to obtain valuable insights for further analysis. I performed segmentation of the pre-processed EEG data into 4-second epochs with 10% overlapping. I extracted 4 EEG frequency bands – theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (above 30 Hz) for each epoch from each channel of the data as shown in Fig. 4.8.

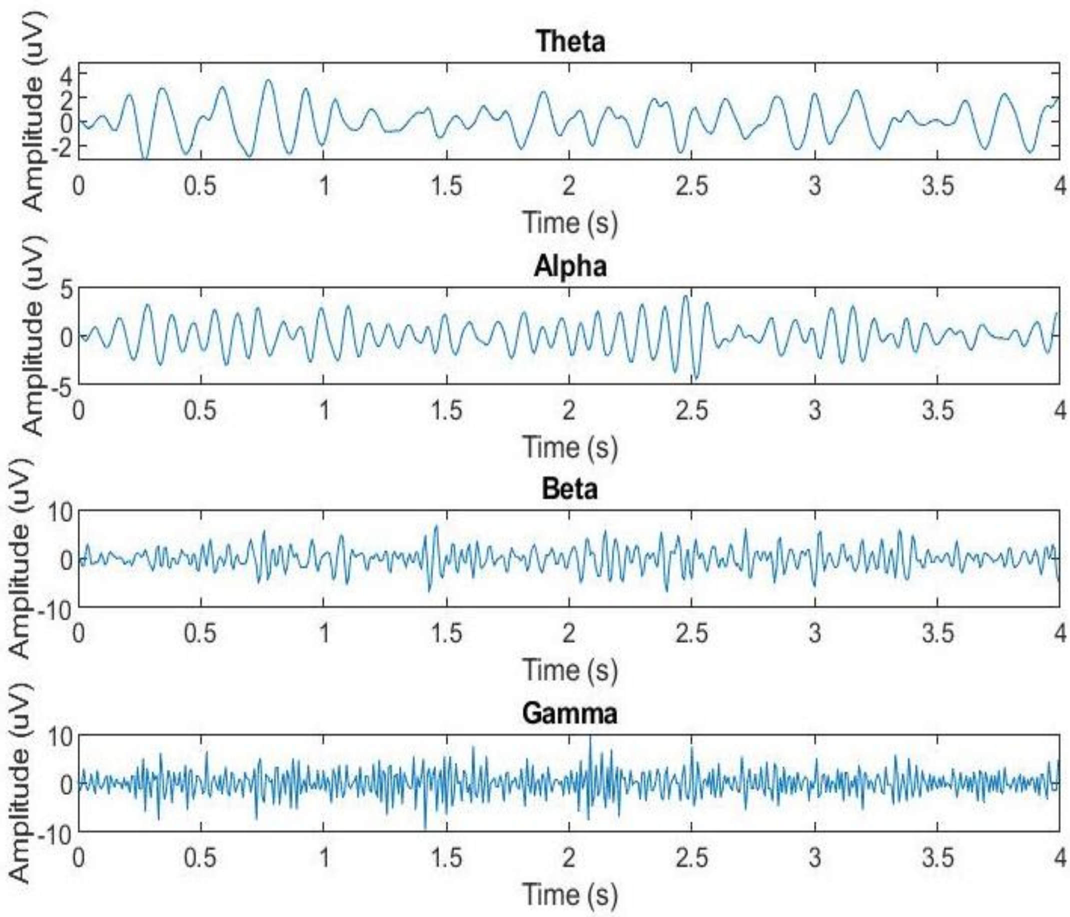


Fig. 4.8 EEG frequency bands for one epoch.

I calculated many statistical and frequency domain features commonly used for EEG-based preference prediction in Neuromarketing [43], [47]. Among the statistical features, I computed mean, SD, variance, skewness, kurtosis, Hjorth parameters- activity, complexity, mobility, spectral entropy, sample entropy, and differential entropy. I determined these 11 statistical features (F_1 – F_{11}) for each band as listed in Table 4.1.

Table 4.1 Statistical & Frequency Domain Features

Feature	Equation
Mean – F_1	$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i$
Standard Deviation – F_2	$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}}$
Variance – F_3	σ^2

Skewness – F_4	$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i - \mu}{\sigma}\right)^3$
Kurtosis – F_5	$\frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i - \mu}{\sigma}\right)^4$
Activity – F_6	$var(X)$
Mobility – F_7	$\sqrt{\frac{var(X')}{var(X)}}$
Complexity – F_8	$\frac{Mobility(X')}{Mobility(X)}$
Spectral Entropy – F_9	$\hat{p}(f) = P(f)/\sum_{f=0.5}^{45} \mathrm{P(f)}$
	$SH = -\sum_{f=0.5}^{45} \hat{p}(f)\log(\hat{p}(f))$
Differential Entropy – F_{10}	$h_i(X) = \frac{1}{2}\log(2\pi e\sigma^2{}_i)$
Sample Entropy – F_{11}	$SampEn\,(m, r) = -ln\frac{B^{m+1}(r)}{B^m(r)}$
Average Power – F_{12}	$P_f = \frac{1}{N}\sum_{n=1}^{N}\lvert X_n(k)\rvert^2$
	Avg_Pow = bandpower (P_f, f ,'psd')
AW index – F_{13}	$\frac{\alpha(F4) - \alpha(F3)}{\alpha(F4) + \alpha(F3)}$
Valence index– F_{14}	$\frac{\beta(F3)}{\alpha(F3)} - \frac{\beta(F4)}{\alpha(F4)}$

I used Welch's periodogram [126] to determine the power spectral density (PSD) of each band as presented in Fig. 4.9. I then used MATLAB's built-in *bandpower* command with *'psd'* extension to determine the average power for each band as 12th feature (F_{12}) of my study. I also calculated FAA indexes – Approach-Withdrawal (AW) index (F_{13}) [117] and Valence index (F_{14}) [121] as listed in Table 4.1. In the end, I performed feature scaling and mean normalization, using equation (21).

$$x_{normalized} = \frac{\left(x - mean(x)\right)}{SD(x)} \quad (21)$$

4.2.6 Classification

Several studies in the literature have used various machine-learning techniques to classify consumers' preferences [10], [43], [47]. In my research, I have employed Decision Tree (DT) algorithm, kNN, Support Vector Machine (SVM), Neural Network (NN), and Ensemble Learning (EL) with subspace kNN for preference prediction. I assigned label one to the best products and label zero to the worst products, and performed binary classification. I used 80% of the data for training and 20% for testing. To evaluate the performance of my models, I calculated accuracy, precision, recall, and F1 scores, which are presented in Table 4.2, Table 4.3, and Table 4.4 for traditional, automated, and proposed pipelines respectively.

Table 4.2 Performance evaluation of traditional pre-processing pipeline

Model Name	Test Accuracy	Precision	Recall	F1 Score
DT	63.9%	0.60	0.71	0.65
SVM	66.7%	0.61	0.82	0.70
kNN	83.3%	0.82	0.82	0.82
NN	69.4%	0.64	0.82	0.72
EL	75%	0.72	0.76	0.74

Table 4.3 Performance evaluation of automated pre-processing pipeline

Model Name	Test Accuracy	Precision	Recall	F1 Score
DT	65.2%	0.63	0.83	0.71
SVM	82.6%	0.83	0.83	0.83
kNN	87%	0.85	0.92	0.88
NN	73.9%	0.71	0.83	0.77
EL	78.3%	0.71	1	0.83

Table 4.4 Performance evaluation of optimal pre-processing pipeline

Model Name	Test Accuracy	Precision	Recall	F1 Score
DT	76.7%	0.77	0.71	0.74
SVM	83.3%	0.85	0.79	0.81
kNN	93.3%	1	0.86	0.92
NN	90%	0.92	0.86	0.89
EL	100%	1	1	1

4.3 RESULTS & DISCUSSION

A comparative summary of the classifiers' test accuracies for the three pipelines is presented in Table 4.5. It is evident that my proposed pre-processing pipeline outperformed both the traditional and automated pipelines for any classification models. The Ensemble learning with kNN subspace classifier obtained the best test accuracy of 100%, followed by kNN with 93.3% and NN with 90% for my proposed model.

Table 4.5 Comparison of the classifiers' test accuracies for the three pipelines

Model Name	Traditional Pipeline	Automated Pipeline	Proposed Pipeline
DT	63.9%	65.2%	**76.7%**
SVM	66.7%	82.6%	**83.3%**
kNN	83.3%	87%	**93.3%**
NN	69.4%	73.9%	**90%**
EL	75%	78.3%	**100%**

Furthermore, the Receiver Operating Characteristic (ROC) curves of the Ensemble learning classifier for the traditional, automated, and proposed pipelines are shown in Fig. 4.9, 4.10, and 4.11 respectively. The ROC analysis indicates that the classifier in my proposed pipeline produced the curve at the top-left corner of the graph and achieved the highest AUC.

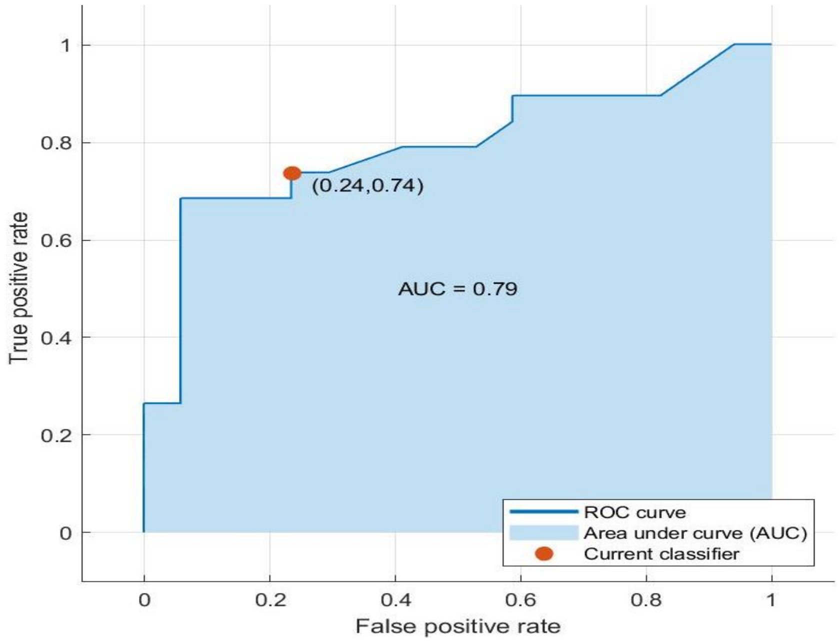


Fig. 4.9 ROC curve of the Ensemble classifier for the traditional pipeline.

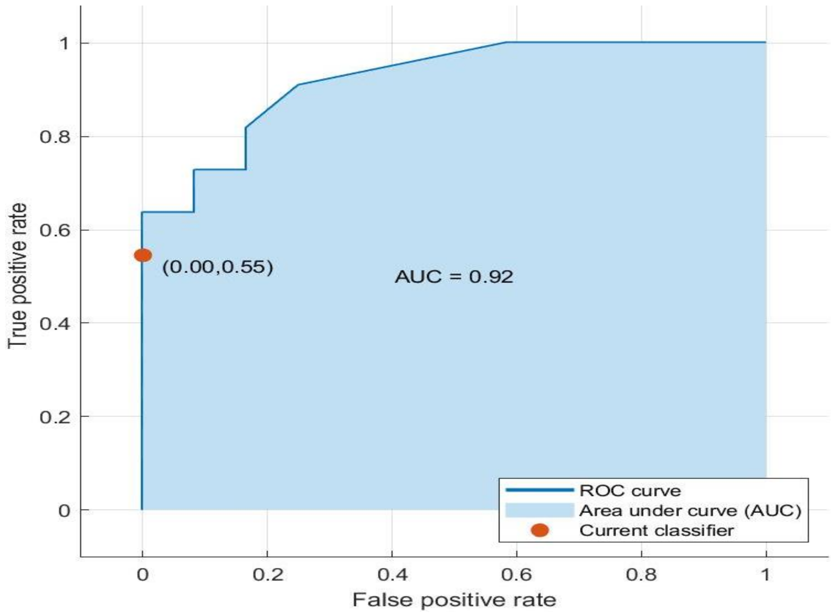


Fig. 4.10 ROC curve of the Ensemble classifier for the automated pipeline.

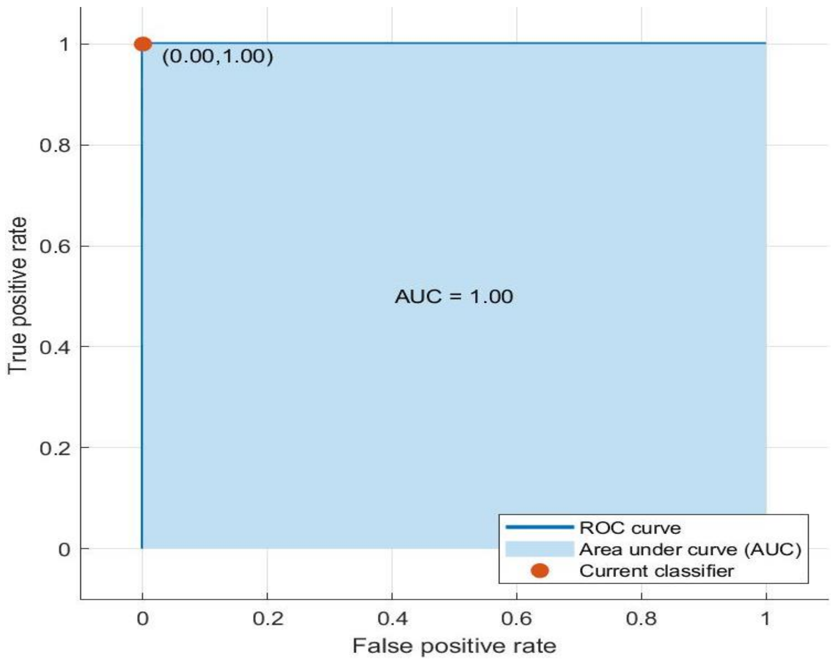


Fig. 4.11 ROC curve of the Ensemble classifier for the proposed pipeline.

Compared to the traditional method, the automated pipeline delivered better results since it filters out the specific portions of the data that contain excessive noise. However, my proposed pre-processing pipeline outperformed both the traditional and automated pipelines. My pipeline not only did a better job of cleaning the data but also retained the necessary information within the data. Therefore, I recommend future researchers pre-process the EEG data by carefully examining it. Although this may be a time-consuming process, it is worth the effort. While the automated process is fast, it may come at the cost of valuable data information. A comparison with the previous research works is presented in Table 4.6.

Table 4.6 Comparison with previous research works

Reference	Classifier	No. of Channels	No. of Different Features	Test Accuracy
Yadava et al. [10]	HMM	14	1	70.33%
Usman et al. [148]	DNN	14	16	80.89%
Alnuman et al. [16]	SVM	14	12	66.25%
Teo et al. [42]	DNN	9	1	79.76%
Aldayel et al. [43]	DNN, RF	14	16	93%
Bandara et al. [157]	RF	4	16	91.97%
Ogino & Mitsukura [158]	kNN	1	1	72.40%
Ramirez et al. [159]	CNN	12	1	91.83%
This Work	**EL**	**16**	**14**	**100%**
	kNN			**93.3%**

4.4 SUMMARY & IMPLICATIONS

In this chapter, I compared the impacts of different preprocessing pipelines on EEG-based preference prediction in Neuromarketing. I collected raw EEG data and used three pre-processing pipelines - traditional, automated, and my proposed pipeline - to clean the recorded data. I then extracted 14 different statistical and frequency domain features and used five machine learning models for preference prediction. I found that my proposed preprocessing pipeline performed better than the other two methods for every classifier. I achieved the best accuracy of 100% for Ensemble learning and 93.3% for kNN. My proposed pipeline not only cleaned the data effectively but also retained the necessary information. Therefore, I recommend that the best way to pre-process EEG data is to carefully examine the data.

5 Objective Labeling

In this chapter, I proposed an objective labeling method for EEG-based preference prediction in Neuromarketing and compared it with prospective subjective labeling. Section 5.1 describes the background, section 5.2 discusses the methodology followed, section 5.3 discusses the results, section 5.4 presents the discussions, and section 5.5 draws the pertinent conclusions.

5.1 BACKGROUND

Businesses and advertisers need to accurately predict consumer preferences to maximize profit by avoiding unpopular or unattractive products [47]. Prior studies adopted various Neuromarketing techniques to predict customers' preferences accurately [10], [11], [39], [40], [41], [42], [46], [47], [48], [49], [50], [52], [53], [54], [148]. Typically, these techniques involve collecting brain data from subjects while they are exposed to different marketing stimuli. After the experiments, the subjects give ratings based on their preferences, which researchers use to train various machine learning models to predict subjects' future preferences. Such subjective labeling methods are not appropriate as they contradict the main goal of Neuromarketing research. Neuromarketing research is motivated by the idea that people tend to conceal their true opinions in self-reports. The ultimate objective of Neuromarketing research is to identify genuine preferences from brain data. Subjective labeling differs from the genuine preferences indicated by the brain data [43], [160]. Therefore, objective labeling, where the labels should be detected directly from brain data, is essential for machine learning models to train and predict accurately.

In this chapter, I propose an objective labeling method for EEG-based preference prediction in Neuromarketing and compare the results with subjective labeling. I hypothesize that machine learning models trained by objective labeling should provide better classification results than subjective labeling because objective labels denote the genuine preferences of consumers.

5.2 METHODOLOGY

5.2.1 Data used

I used both Dataset 1 and Dataset 2, described in the "Data Collection" chapter, to compare the subjective and objective labeling. From Dataset 2, I used the data collected using English language stimuli to avoid the language effects. Therefore, the final dataset contained three recordings for each subject, yielding 108 (36*3) recordings.

5.2.2 Pre-processing

I utilized the optimal EEG pre-processing pipeline to pre-process the EEG data, as described section 4.2.4. The pipeline suggests the best way to pre-process EEG data is to look into the data. Automated pre-processing pipelines are fast, but they can cause the loss of essential information. Therefore, I took time to manually pre-process each EEG data according to the optimal pipeline. I used EEGLAB 2022 in MATLAB 2021a to implement the pipeline.

5.2.3 Subjective labeling

During the study, the participants were asked to provide their feedback on a 5-point Likert scale. The scale ranged from "strongly dislike" to "strongly like". I considered ratings 4 ("like") and 5 ("strongly like") as positive feedback, labeled with label 1 indicating that the subject liked the product. I considered ratings 1 ("strongly dislike") and 2 ("dislike") as negative feedback, labeled with label 0 indicating that the subject disliked the product. For this study, I did not take into account the data with a rating of 3 ("neutral").

5.2.4 Objective labeling

I subdivided each channel's EEG data into their constituent four frequency bands - θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), and γ (30 ~ Hz). Using Welch's periodogram [126], I determined each band's power spectral density (PSD). I then calculated the average power of each band. To label the data objectively, I used the average power of different EEG bands from various channels to determine four different asymmetry indices - Approach-Withdrawal (AW) Index, Valence Index, Choice Index (γ), and Choice Index (β). These indices can serve as autonomic indicators of consumers' preferences, as evident from previous studies [13], [20], [43], [47], [50], [60], [66], [97], [117], [121], [140].

The AW index is a metric that measures the difference in activations between the left and right hemispheres, known as frontal alpha asymmetry (FAA). It determines desire and motivation by assessing the higher activation of alpha in the right frontal cortex [43], [50], [60], [117], [119], [140]. Many studies have demonstrated the effectiveness and accuracy of the FAA as a crucial factor in emotion and Neuromarketing research [20], [47], [50], [60], [66], [97], [102], [117], [121]. To calculate the AW index, we can use electrodes F4 and F3 to determine the difference between the right and left PSD using equation (22) [50], [117].

$$AW\ Index = \frac{\alpha(F4) - \alpha(F3)}{\alpha(F4) + \alpha(F3)} \qquad (22)$$

Research has shown a correlation between a customers' emotional states and the asymmetries of their frontal lobe in terms of valence [20], [47], [50], [60], [66], [97], [117], [121]. More specifically, if the left frontal lobe is activated, it indicates a positive valence, whereas the right frontal lobe activation suggests a negative valence. Several studies have supported the theory that frontal EEG asymmetry is an indicator of valence [43], [47], [50], [119]. It can be calculated by the equation (23) [50], [121].

$$Valence\ Index = \frac{\beta(F3)}{\alpha(F3)} - \frac{\beta(F4)}{\alpha(F4)} \qquad (23)$$

The choice index is based on gamma and beta oscillations in the frontal part of the brain. These oscillations are primarily associated with the decision-making process. The gamma band is highly correlated with willingness-to-pay responses, which assess consumer preference and choice [60]. A higher value in the gamma and beta bands indicates greater activation in the left prefrontal region. On the other hand, the right prefrontal part shows considerably stronger activation at lower levels [47], [50], [60]. The choice index can be calculated using equations (24) for the gamma band and (25) for the beta band [47], [60]. To determine the choice indexes, AF3 and AF4 channels were used for dataset 1, while Fp1 and Fp2 channels were used for dataset 2.

$$Choice\ Index\ (\gamma) = \frac{\log\left(\gamma\left(Electrode_{left}\right)\right) - \log\left(\gamma\left(Electrode_{right}\right)\right)}{\log\left(\gamma\left(Electrode_{left}\right)\right) + \log\left(\gamma\left(Electrode_{right}\right)\right)} \qquad (24)$$

$$Choice\ Index\ (\beta) = \frac{\log\left(\beta\left(Electrode_{left}\right)\right) - \log\left(\beta\left(Electrode_{right}\right)\right)}{\log\left(\beta\left(Electrode_{left}\right)\right) + \log\left(\beta\left(Electrode_{right}\right)\right)} \qquad (25)$$

Previous research used Dataset 1 and utilized only the valence index to label the data objectively [13], [43]. However, relying solely on one index for labeling limits the confidence of labeling to only two specific channels and two EEG bands. Therefore, for a robust objective labeling approach using only brain data, more EEG channels and bands should be incorporated. In this study, I have used the abovementioned four indices to label the data objectively. I used a majority voting system to label the data objectively based on the four indices, thus avoiding bias toward only one index. For EEG data, I considered it a "like" and labeled it as 1 if at least three of the four indices gave a positive value. On the other hand, if at least three of the four indices gave a negative value, I marked the EEG data as "dislike" and labeled it as 0. The summary of the objective labeling for Dataset 1 and Dataset 2 can be found in Table 5.1.

Table 5.1. Objective labeling summary for Dataset 1 and Dataset 2.

Dataset	Possible objective labeling	Subjective & objective labeling match	Subjective & objective labeling mismatch
Dataset 1	711/1048 (67.84%)	360/711 (50.63%)	351/711 (49.37%)
Dataset 2	68/108 (62.96%)	33/68 (48.53%)	35/68 (51.47%)

The objective labeling percentage for Dataset 1 was 67.84%, and for Dataset 2 it was 62.96%. While labeling 100% of the data for both datasets is possible if I don't follow the majority voting system, it would reduce labeling confidence. Additionally, Dataset 1 only provides data for 4 seconds, and some subjects from Dataset 2 made decisions too quickly. These scenarios do not accurately represent the natural purchasing environment, where participants usually take more time to decide. Thus, it is not possible to achieve 100% labeling objectively.

5.2.5 Feature Extraction

I calculated various statistical and frequency domain features, as described in section 4.2.5. The statistical features I computed include mean, standard deviation (SD), variance, skewness, kurtosis, and Hjorth parameters such as activity, complexity, mobility, spectral entropy, sample entropy, and differential entropy. The frequency domain features I included consist of the average power of each EEG band and the four asymmetry indices. Dataset 2 had a comparatively higher duration than Dataset 1 and fewer EEG recordings. Therefore, I segmented the pre-processed EEG data of Dataset 2 into 4-second epochs with 10% overlapping. I extracted 4 EEG frequency bands – θ (4–8 Hz), α (8–13 Hz), β (13–30 Hz), and γ (30 ~ Hz) for each epoch from each channel of the data. Before running into the classification algorithms, I normalized the extracted features using equation (21).

5.3 RESULTS

To compare the results of subjective and objective labeling, I considered three different scenarios: case 1 - training and testing using public Dataset 1, case 2 - training and testing using Dataset 2, and case 3 - training and testing using the combination of Dataset 1 and Dataset 2. In each case, I used 80% of the data for training and 20% for testing. I used five different classification algorithms - Decision Tree (DT) algorithm, Support Vector Machine (SVM), k-Nearest Neighbor (kNN), Neural Network (NN), and Ensemble Learning (EL). I used a boosted DT algorithm for the EL classifier, and for NN, I used a narrow NN model. Since I labeled my data objectively using the four indices mentioned in equations (22), (23), (24), and (25), I ran the classification algorithms with and without these indices in the feature list to compare the results with subjective labeling. I used 5-fold cross-validation to avoid overfitting. The classification results of the three cases, including and excluding the indices in the feature list, are presented in Tables 5.2 and 5.3, respectively. It is evident that objective labeling leads to better classification results than subjective labeling for all three cases, regardless of the inclusion or exclusion of indices in the feature list.

Table 5.2 Classification results of subjective (Sbj.) and objective (Obj.) labeling excluding indices in the feature list.

Excluding indices in feature list											
Cases	**Classifier →**	**DT**		**SVM**		**kNN**		**NN**		**EL**	
	Labeling →	**Sbj.**	**Obj.**	**Sbj.**	**Obj.**	**Sbj.**	**Obj.**	**Sbj.**	**Obj.**	**Sbj.**	**Obj.**
Case - 1	Accuracy (%)	56	89.4	56.7	94.3	56	76.6	58.2	94.3	**58.9**	**95.7**
	Precision (%)	50	85.1	50.9	90	50	62.3	52.8	88.5	53.2	90.4
	Recall (%)	17.7	83.3	43.5	93.8	38.7	79.2	45.2	95.8	53.2	97.9
	F1 score (%)	26.2	84.2	47	91.8	43.6	69.7	48.7	92	53.2	94
Case - 2	Accuracy (%)	79.1	96.5	88.4	98.8	81.4	97.7	**90.7**	**100**	87.2	98.8
	Precision (%)	86.4	95.7	87	97.9	82.7	100	93.8	100	89.8	100
	Recall (%)	76.0	97.8	94	100	86	95.7	90	100	88	97.8
	F1 score (%)	80.9	96.8	90.4	98.9	84.3	97.8	91.8	100	88.9	98.9
Case - 3	Accuracy (%)	70.1	86.6	68.8	87.5	71.9	80.4	66.5	86.6	**72.8**	**91.5**
	Precision (%)	72.6	87.1	69.6	84.9	81.3	78.2	66.7	86.2	76.3	92
	Recall (%)	62.7	79.6	64.5	84.9	55.9	73.1	63.6	80.6	64.5	87.1
	F1 score (%)	67.3	83.1	67	84.9	65.9	75.6	65.1	83.3	70	89.5

Table 5.3 Classification results of subjective (Sbj.) and objective (Obj.) labeling including indices in the feature list.

Including indices in feature list											
Cases	**Classifier →**	**DT**		**SVM**		**kNN**		**NN**		**EL**	
	Labeling →	**Sbj.**	**Obj.**	**Sbj.**	**Obj.**	**Sbj.**	**Obj.**	**Sbj.**	**Obj.**	**Sbj.**	**Obj.**
Case - 1	Accuracy (%)	56.7	**99.3**	**61.7**	95	**61.7**	84.4	57.4	97.9	59.6	**99.3**
	Precision (%)	50.9	98	57.1	93.6	57.7	96.4	51.4	97.9	54.2	100
	Recall (%)	43.5	100	51.6	91.7	48.4	56.3	59.7	95.8	51.6	97.9
	F1 score (%)	47	99	54.2	92.6	52.6	71.1	55.2	96.8	52.9	98.9
Case - 2	Accuracy (%)	75.6	**100**	90.7	98.8	82.6	97.7	**91.9**	98.8	89.5	98.8
	Precision (%)	83.7	100	93.8	100	97.3	100	92.2	100	90.2	100
	Recall (%)	72	100	90	97.8	72	95.7	94	97.8	92	97.8
	F1 score (%)	77.4	100	91.8	98.9	82.8	97.8	93.1	98.9	91.1	98.9
Case - 3	Accuracy (%)	69.6	92.4	63.8	90.6	70.5	82.1	63.4	90.2	**71**	**95.5**
	Precision (%)	68.1	90.4	65.6	86.7	71.6	78.5	62.3	87.4	74.7	96.6
	Recall (%)	71.8	91.4	55.5	91.4	66.4	78.5	64.5	89.2	61.8	92.5
	F1 score (%)	69.9	90.9	60.1	89	68.9	78.5	63.4	88.3	67.7	94.5

5.3.1 Comparison of prediction accuracies for case 1

Dataset 1 has unbalanced like/dislike labels in its objectively labeled data. To balance the training data, I used the Synthetic Minority Over-sampling Technique (SMOTE) [161]. However, I separated the test data (20%) before applying SMOTE. The comparison of prediction accuracies for case 1, with both subjective and objective labeling, is presented in Fig. 5.1. The results show that each classifier had better prediction accuracy with objective labeling, regardless of the inclusion or exclusion of the indices in the feature list. When excluding indices in the feature list, the maximum accuracy obtained with the subjective labeling was 58.9%, while with the objective labeling, it was 95.7%, both using the EL classifier. When including indices in the feature list, the maximum accuracy obtained with the subjective labeling was 61.7% using the SVM classifier, while with the objective labeling, it was 99.3% using the DT classifier.

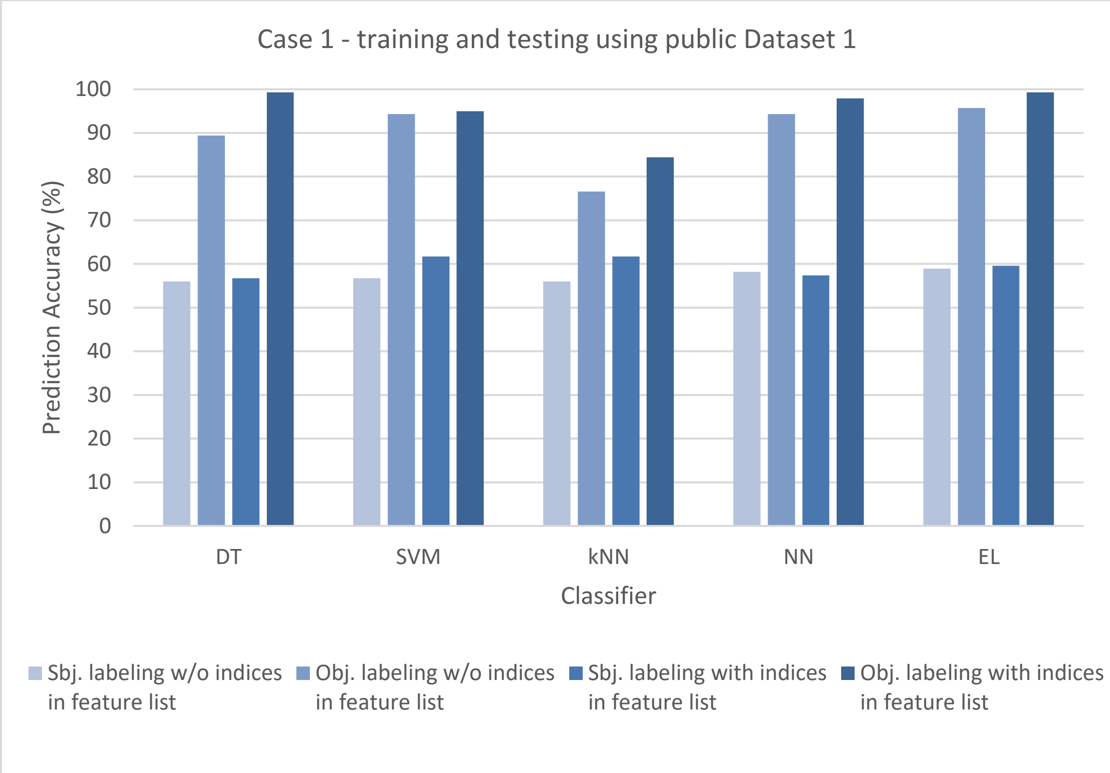


Fig. 5.1 Comparison of subjective and objective labeling prediction accuracies for case 1 - training and testing using public Dataset 1.

5.3.2 Comparison of prediction accuracies for case 2

Fig. 5.2 presents a comparison of the prediction accuracies obtained with subjective and objective labeling for case 2. Using the NN classifier and excluding indices in the feature list, the maximum accuracy achieved with subjective labeling was 90.7%, and with objective labeling, it was 100%. Including indices in the feature list, the maximum accuracy obtained with subjective labeling was 91.9% using the NN classifier, and with objective labeling, it was 100% using the DT classifier. Similar to case 1, the results of case 2 demonstrate that each classifier exhibits better prediction accuracy with objective labeling, regardless of the inclusion or exclusion of indices in the feature list. However, the difference in prediction accuracies between subjective and objective labeling is less in case 2 compared to case 1, although still significant. In the previous section, it was mentioned that the EEG data of case 1 (Dataset 1) only lasted for 4 seconds. The participants were given 4 seconds to respond to each product image. However, such a time limit does not reflect a natural buying environment where participants usually take their time to decide. It can even cause pressure on participants, leading to significant differences in prediction accuracies for subjective and objective labeling in case 1, as shown in Fig. 3. On the other hand, in case 2 (Dataset 2), the participants were given no time limit while viewing products to buy during the data recording. Despite this, the prediction accuracies with objective labeling were significantly higher than subjective labeling for both cases.

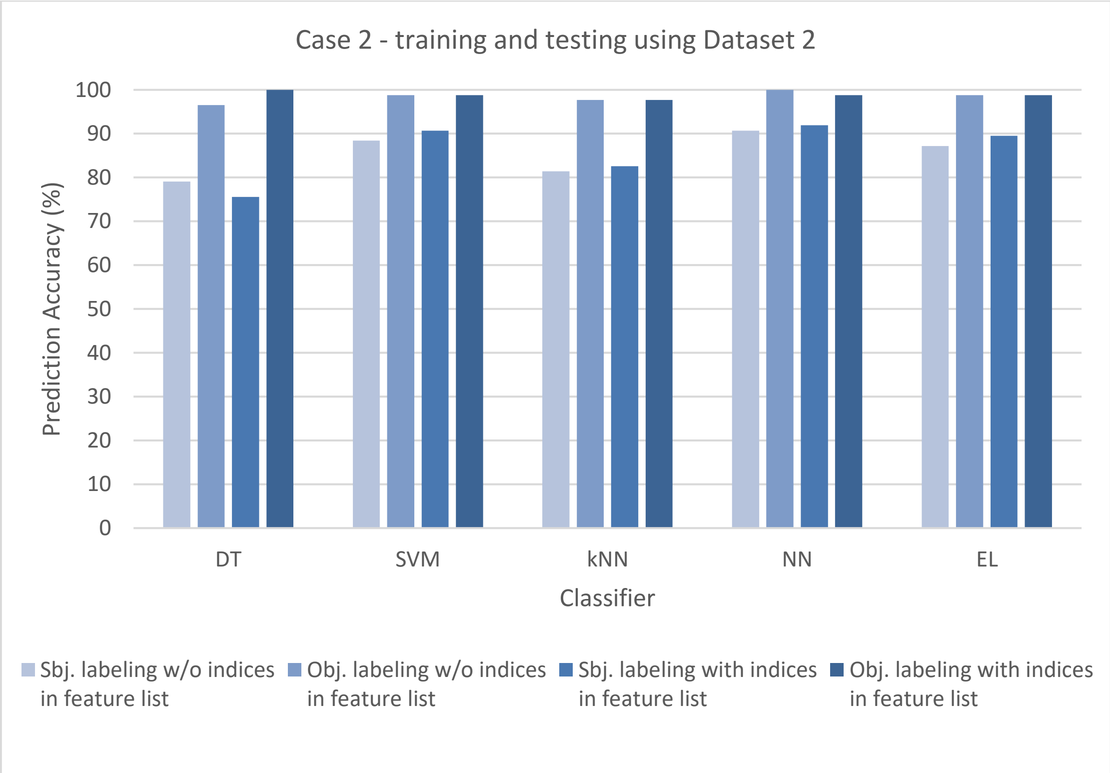


Fig. 5.2 Comparison of subjective and objective labeling prediction accuracies for case 2 - training and testing using Dataset 2.

5.3.3 Comparison of prediction accuracies for case 3

In case 3, I combined two datasets, Dataset 1 and Dataset 2, to train and test the classification models. Dataset 1 contained EEG data from 14 channels, and Dataset 2 had data from 16 channels. However, I had to remove two channels (Cz and Oz) from Dataset 2 to have the same number of features in the feature set. The comparison of prediction accuracies for subjective and objective labeling for case 3 is presented in Fig. 5.3. Excluding indices in the feature list, the maximum accuracy achieved with subjective labeling was 72.8%, while the maximum accuracy achieved with objective labeling was 91.5%, both using the EL classifier. On the other hand, including indices in the feature list, the maximum accuracy achieved with subjective labeling was 71%, while the maximum accuracy achieved with objective labeling was 95.5%, both using the EL classifier. Like in the previous cases 1 and 2, the results of case 3 indicate that each classifier showed better prediction accuracy with objective labeling, regardless of the inclusion or exclusion of the indices in the feature list.

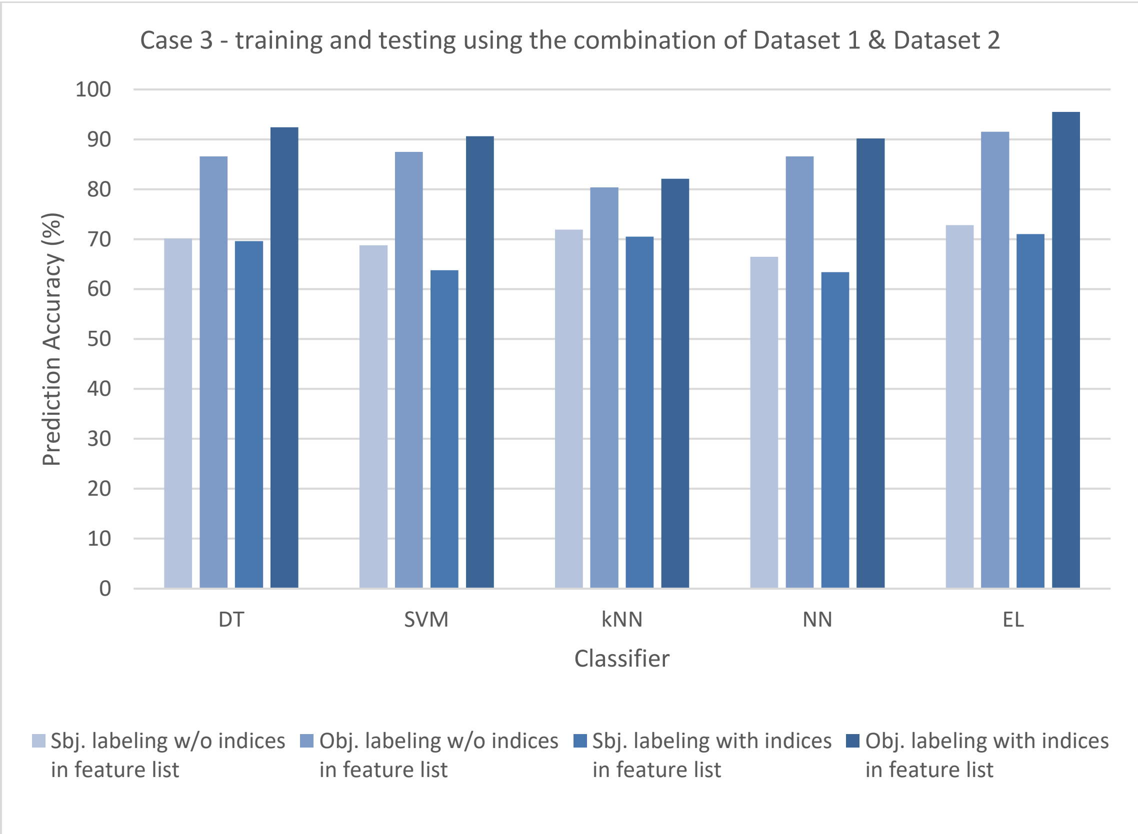


Fig. 5.3 Comparison of subjective and objective labeling prediction accuracies for case 3 - training and testing using the combination of Dataset 1 and Dataset 2.

5.4 DISCUSSION

The results show that the machine learning models trained by objective labeling provide better classification results than subjective labeling in all three cases. The previous objective labeling approach used public Dataset 1 to compare subjective and objective labeling classification results [13], [43]. In this section, I compared the classification results with the previous work using public Dataset 1 and presented the findings in Table 5.4 and Table 5.5 for subjective and objective labeling, respectively. The previous work used SVM, kNN, and EL classifiers for both subjective and objective labeling [13] and added the Deep Neural Network (DNN) classifier for objective labeling classification [43]. They used the Random Forest (RF) method as an EL classifier.

The previous study achieved a maximum accuracy of 60% through subjective labeling using the EL classifier, while the proposed study achieved 58.9%. However, when using objective labeling, the previous research achieved a maximum accuracy of 83% using the EL classifier and excluding indices in the feature list, while the proposed study achieved 95.7%. Additionally, when including indices in the feature list, the previous research achieved a maximum accuracy of 94% using the EL classifier, while the proposed study achieved 99.3%. Therefore, based on the results presented in Table 5.4 and Table 5.5, it can be concluded that although the classification results of subjective labeling in both studies are quite similar, the proposed approach with objective labeling outperforms the previous approach.

Table 5.4 Comparison of classification results of subjective labeling with the previous work.

Subjective labeling							
No. of different features	**Performance parameters**	**Classifier**					
		SVM		**kNN**		**EL**	
		Prev. work	**This work**	**Prev. work**	**This work**	**Prev. work**	**This work**
12 (Excluding indices in feature set)	Accuracy (%)	56	56.7	57	56	**60**	**58.9**
	Precision (%)	54	50.9	56	50	59	53.2
	Recall (%)	56	43.5	57	38.7	60	53.2

Table 5.5 Comparison of classification results of objective labeling with the previous work.

Objective labeling									
No. of different features	**Performance parameters**	**Classifier**							
		SVM		**kNN**		**NN**		**EL**	
		Prev. work	**This work**	**Prev. work**	**This work**	**Prev. work**	**This work**	**Prev. work**	**This work**
12 (Excluding indices in feature set)	Accuracy (%)	71	94.3	72	76.6	77	94.3	**83**	**95.7**
	Precision (%)	72	90	73	62.3	79	88.5	84	90.4
	Recall (%)	71	93.8	72	79.2	77	95.8	83	97.9
16 (Including indices in feature set)	Accuracy (%)	87	95	80	84.4	93	97.9	**94**	**99.3**
	Precision (%)	88	93.6	80	96.4	93	97.9	94	100
	Recall (%)	87	91.7	80	56.3	93	95.8	94	97.9

In Table 5.6, I have presented a comparative analysis of the classification results of my work with previous works that have used the public Dataset 1. Some studies have used automatic feature engineering techniques, such as long short-term memory (LSTM) based features, in combination with handcrafted features to achieve better classification results [51], [148]. However, my proposed objective labeling approach has outperformed the previous studies, using fewer features and without any automatic feature engineering techniques. Although the subjective labeling approach provided some good classification results, attaining proper labeling from the brain data is essential as it can lead to better classification results, as demonstrated in this research.

Table 5.6 Comparison of classification results of previous works that used the public Dataset 1.

Reference	**Labeling**	**No. of different features**	**Automated feature engineering**	**Classifier used**	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1 score (%)**
Yadava et al. [10]	Subjective	14	No	HMM	70.33	-	-	-
Shah et al. [51]	Subjective	-	Yes	EL	96.89	95.78	95.89	95.76
Göker [54]	Subjective	49	No	Bidirectional-LSTM	96.83	95	99	97
Usman et al. [148]	Subjective	-	Yes	DNN	80.89	-	80.23	-
Aldayel et al. [13], [43]	Objective	16	No	EL	94	94	94	-
This work	**Objective**	**16**	**No**	**DT/EL**	**99.3**	**98**	**100**	**99**

5.5 SUMMARY & IMPLICATIONS

In this chapter, I have proposed an objective labeling method for EEG-based preference prediction in Neuromarketing and compared the results with subjective labeling. I used two datasets and ran various classification models to compare the results in three cases. I found that the machine learning models trained by objective labeling provided better classification results than subjective labeling in all three cases. I compared my results with previous works on the same dataset using subjective and objective labeling. My proposed objective labeling approach provided better classification results than the earlier methods. These findings support the aim of Neuromarketing research, which is to accurately detect genuine preferences from brain data, rather than relying on self-reported data.

6 Foreign Language Effects

In this chapter, I *discussed the effects of foreign language on the preferences of the consumers in Neuromarketing. Section 6.1 describes the background, section 6.2 discusses the methodology followed, section 6.3 discusses the results and discussions, and section 6.4 draws the pertinent conclusions.*

6.1 BACKGROUND

The field of Neuromarketing research has evolved from predicting customer preferences to implementing neuroscience-based marketing strategies [22]. Modern studies are trying to build marketing strategies by observing how subtle changes affect the consumers' preferences in practice [23]. Among the prominent components of promotions, language has a major impact on consumers' minds [24]. In a bilingual nation like ours, which advertisement language will work better remains still unchecked.

In a recent study, researchers found that using a foreign language (FL) instead of a native language (NL) affects people's decision-making in both moral and risk-related situations [162]. The study revealed that participants were more willing to accept harm to achieve better outcomes in the moral domain when using FL. In the risk aversion domain, FL reduced people's hesitancy towards taking risks. I intended to see whether same effects can be found during purchase decision making in Neuromarketing. I intended to see whether consumers show the same risk adopting tendencies during purchase decision-making when exposed to advertisements in FL. I hypothesize that consumers will show risk adopting tendencies during purchase decision-making when exposed to ads in FL instead of NL.

6.2 METHODOLOGY

6.2.1 Data used

I used Dataset 2, described in the "Data Collection" chapter, to investigate the language effects. From Dataset 2, I used EEG data corresponding to the one product with very weak specifications and another one with comparatively very good specifications in all aspects in both NL (Bengali) and FL (English) to investigate the language effects. Therefore, the final dataset contained four recordings for each subject, yielding 144 (36*4) recordings.

6.2.2 Pre-processing

I utilized the optimal EEG pre-processing pipeline to pre-process the EEG data, as described section 4.2.4. The pipeline suggests the best way to pre-process EEG data is to look into the data. Automated pre-processing pipelines are fast, but they can cause the loss of essential information. Therefore, I took time to manually pre-process each EEG data according to the optimal pipeline. I used EEGLAB 2022 in MATLAB 2021a to implement the pipeline.

6.2.3 Objective labeling

I labeled the data objectively using the method described in chapter 5. The summary of the objective labeling for NL and FL can be found in Table 6.1.

Table 6.1 Objective labeling summary for NL and FL.

Language	Possible objective labeling
Bengali (NL)	65/72 (90.28%)
English (FL)	68/72 (94.44%)

The objective labeling percentage for NL was 90.28%, and for FL it was 94.44%. While labeling 100% of the data for both languages is possible if I don't follow the majority voting system, it would reduce labeling confidence. Additionally, some subjects from Dataset 2 made decisions too quickly. These scenarios do not accurately represent the natural purchasing environment, where participants usually take more time to decide. Thus, it is not possible to achieve 100% labeling objectively.

6.3 RESULTS & DISCUSSION

Using the objectively labeled rating, I found that 47.73% of the subjects showed purchase intentions towards the weak products when they were exposed to the stimuli in NL. On the contrary, 54.55% of the subjects showed purchase intentions towards the weak products when they were exposed to the stimuli in FL. The results are summarized in Table 6.2. It is evident from the results that the subjects were risk adopting in FL. Therefore, the hypothesis that consumers will show risk adopting tendencies during purchase decision-making when exposed to ads in FL instead of NL is true.

Table 6.2 Risk adoption tendency in NL and FL using objective labeling

Language	% of choosing the products with weak specifications
Bengali (NL)	47.73%
English (FL)	54.55%

6.4 SUMMARY & IMPLICATIONS

In this chapter, I discussed the effects of foreign languages on consumers' preferences in Neuromarketing. I found that consumers show risk adoption tendencies when exposed to product ads in FL. Therefore, companies should use NL to advertise their products in a bilingual country.

7 Conclusions

The overall methodologies and findings of this study are presented in this chapter. It summarizes the discussion of the methodology and results section. In addition, this chapter contains general summary, key findings, and limitations along with the significance of the study. The future study subsection will present the researchers interested in this topic with some suggestions to work with.

7.1 GENERAL SUMMARY

Neuromarketing is an emerging brain-computer interface research field that aims to understand consumers' internal decision-making processes when choosing which products to buy. It provides valuable insights for marketers to improve their marketing strategies based on consumers' impressions. However, the current status of EEG-based preference prediction and its classification accuracy is still below optimal.

The performance of an EEG-based preference prediction system largely depends on the suitable selection of a proper EEG preprocessing pipeline since noisy EEG data is likely not to give better results. Most of the studies followed a traditional EEG pre-processing pipeline and also recently an advanced automated EEG pre-processing pipeline has been introduced in the literature. In this study, I have proposed an optimal EEG pre-processing pipeline and compared it with the other two methods mentioned above. I collected raw EEG data and pre-processed them using the three pre-processing pipelines. I then extracted many statistical and frequency domain features and employed several machine-learning classification models. After implementing my proposed pipeline to preprocess EEG data, I observed a significant improvement in classification accuracies. Specifically, my pipeline achieved 100% accuracy for EL, and 93.3% accuracy for kNN model. In comparison, for the same classifiers the traditional and advanced automated pipelines achieved 75%, 83.3%, and 78.3%, 87% accuracies respectively.

Previous studies have used different machine learning models to predict subjects' future preferences based on subjective labeling, i.e., the preferences are self-reported by the subjects. However, such labeling methods contradict the main goal of Neuromarketing research, which is to detect genuine preferences from brain data rather than rely on subjects-reported data. This leads to the necessity of an objective labeling approach where the labels can be detected directly from the brain data. In this study, I proposed an objective labeling method for EEG-based preference prediction in Neuromarketing. I compared the performance of different machine learning models trained on objective and subjective labeling using two datasets: a publicly available Neuromarketing dataset and a new dataset created in my experiments. My findings demonstrate that the models trained on objective labeling provided better classification results than the ones trained on subjective labeling. The result aligns with the goals of Neuromarketing research, as it offers an automated way of labeling data and opens up new avenues for future research.

Among the prominent components of promotions, language has a major impact on consumers’ minds. In a bilingual nation like ours, which advertisement language will work better remains still unchecked. I investigated the effects of foreign languages on consumers' preferences in Neuromarketing. I found that consumers show risk adoption tendencies when exposed to product ads in FL. Therefore, companies should use NL to advertise their products in a bilingual country.

7.2 KEY FINDINGS

The key findings of this study are presented below:

- In this study, I presented a systematic review based on the recent EEG-based Neuromarketing studies, compared different EEG pre-processing pipelines and developed an optimal EEG-preprocessing pipeline, introduced an objective labeling method for EEG-based preference prediction in Neuromarketing, and investigated the effects of FL in purchase decision of consumers.
- The systematic review presents the current research trend, data collection, data processing, and analyzing techniques in Neuromarketing. My findings indicate that the trend in Neuromarketing research has shifted from predicting consumer preferences to developing marketing strategies. I found that the medial-frontal brain region is particularly relevant when it comes to understanding consumers' purchasing behavior. Furthermore, the focus of marketing stimuli has shifted from actual products and television ads to images of products on digital and social media.
- I have proposed an optimal EEG pre-processing pipeline and compared it with the other two existing methods. I found that my proposed preprocessing pipeline performed better than the other two existing methods for every classifier. I achieved the best accuracy of 100% for Ensemble learning and 93.3% for kNN. In comparison, for the same classifiers the traditional and advanced automated pipelines achieved 75%, 83.3%, and 78.3%, 87% accuracies respectively.
- I proposed an objective labeling method for EEG-based preference prediction in Neuromarketing. I compared the performance of

different machine learning models trained on objective and subjective labeling using two datasets: a publicly available Neuromarketing dataset and a new dataset created in my experiments. My findings demonstrate that the models trained on objective labeling provided better classification results than the ones trained on subjective labeling. The result aligns with the goals of Neuromarketing research, as it offers an automated way of labeling data and opens up new avenues for future research.

- I investigated the effects of foreign languages on consumers' preferences in Neuromarketing. I found that consumers show risk adoption tendencies when exposed to product ads in FL.

7.3 RECOMMENDATIONS FOR FURTHER STUDY

Staying current with the research trend in the field of study supports researchers' work and career advancement, benefiting both professional development and the discipline as a whole, as it aids in creating a proper research framework. It is evident from the cluster analysis that even though some studies were found to be using concepts and principles of Neuromarketing to build up a theoretical base or to predict the purchase decision of consumers in the early years of Neuromarketing research, most studies nowadays utilized Neuroscience to create marketing strategies so that businesses become successful. Before setting the research goals, future Neuromarketing researchers should follow the research trend away from simple consumer preference prediction and toward developing marketing strategies incorporating Neuroscience.

Understanding the structure of the human brain is crucial in neuromarketing research. It is closely linked to interpreting neural responses, making it necessary to position electrodes accurately over the relevant brain region to record the neuronal activity associated with a particular function.

Although many researchers have found that the medial-frontal brain region is responsible for the preference function, there is still a need for more agreement on which electrodes should be used within the same brain area. Therefore, scanning a larger brain area focusing on the frontal region would be better.

Neuromarketing researchers should strive to create an authentic purchasing setting to reveal consumers' honest thoughts when making purchasing decisions, regardless of the stimuli employed. However, as technology advances, people increasingly use digital media and social networks to compare and shop for products rather than physically visiting stores. Therefore, future marketing research should focus on using images of products as stimuli to elicit genuine emotional responses through brain data rather than relying on traditional marketing methods such as TV commercials and physical product comparisons. When collecting data, creating a realistic buying scenario for participants is essential. Allowing them ample time to view products without any time constraints will provide a more accurate representation of the purchasing environment and prevent undue pressure on their decision-making process. Using a justified sample size is also critical to obtain statistically significant results. For stimuli around 15 seconds, the sample size should be a minimum of 32, and for stimuli around 30 seconds, it should be a minimum of 24, as found in the study of [65].

EEG signals are susceptible to noises like cardiac, visual, muscle contraction, and power line interference. The preprocessing stage removes these noises and prepares the signal for further processing, ensuring better results. Studies have utilized various pre-processing techniques to create pipelines that convert raw EEG data into noise-free output, demonstrating the potential for automation to enhance efficiency. Neuromarketing researchers can use such pipelines, but they should ensure that the pipeline they choose effectively cleans the raw data without losing necessary information.

Previous studies have used different machine learning models to predict subjects' future preferences based on subjective labeling, i.e., the preferences are self-reported by the subjects. However, such labeling methods contradict the main goal of Neuromarketing research, which is to detect genuine preferences from brain data rather than rely on subjects-reported data. This leads to the necessity of an objective labeling approach where the labels can be detected directly from the brain data. Future researchers should focus on the objective labeling methods to build ML models to predict consumers' future preferences.

Language plays a crucial role in promotions and has a significant impact on consumers' minds. In countries where multiple languages are spoken, it's important to determine which language is more effective in advertising. Recent research indicates that consumers tend to take risks when they see product ads in a foreign language. Therefore, companies should use the native language to advertise their products in bilingual countries.

Bibliography

[1] F. S. Rawnaque *et al.*, “Technological advancements and opportunities in Neuromarketing: a systematic review,” *Brain Inform*, vol. 7, no. 1, p. 10, Dec. 2020, doi: 10.1186/s40708-020-00109-x.

[2] Delroy L. Paulhus, “Socially desirable responding: the evolution of a construct,” in *The Role of Constructs in Psychological and Educational Measurement*, H. I. Braun, D. N. Jackson, and D. E. Wiley, Eds., New Jersey: NJ: Lawrence Erlbaum Associates Publishers, 2002, pp. 49–69.

[3] G. A. Calvert and M. J. Brammer, “Predicting Consumer Behavior: Using Novel Mind-Reading Approaches,” *IEEE Pulse*, vol. 3, no. 3, pp. 38–41, May 2012, doi: 10.1109/MPUL.2012.2189167.

[4] M. Nilashi *et al.*, “Journal of Soft Computing and Decision Support Systems Neuromarketing: A Review of Research and Implications for Marketing,” 2020. [Online]. Available: http://www.jscdss.com

[5] S. Belden, “Science is Culture: Neuroeconomics and Neuromarketing. Practical Applications and Ethical Concerns By Sarah Rebecca Anne Belden,” *Journal of Mind Theory*. 2008. Accessed: Oct. 06, 2022. [Online]. Available: https://www.academia.edu/28024277/Science_is_Culture_Neuroeconomics_and_Neuromarketing_Practical_Applications_and_Ethical_Concerns_By_Sarah_Rebecca_Anne_Belden

[6] S. Agarwal and T. Dutta, “Neuromarketing and consumer neuroscience: current understanding and the way forward,” *DECISION*, vol. 42, no. 4, pp. 457–462, Dec. 2015, doi: 10.1007/s40622-015-0113-1.

[7] V. Khurana *et al.*, “A Survey on Neuromarketing Using EEG Signals,” *IEEE Trans Cogn Dev Syst*, vol. 13, no. 4, pp. 732–749, Dec. 2021, doi: 10.1109/TCDS.2021.3065200.

[8] E. Kroupi, P. Hanhart, J.-S. Lee, M. Rerabek, and T. Ebrahimi, “Predicting subjective sensation of reality during multimedia consumption based on EEG and peripheral physiological signals,” in *2014 IEEE International Conference on Multimedia and Expo (ICME)*, IEEE, Jul. 2014, pp. 1–6. doi: 10.1109/ICME.2014.6890239.

[9] M. Singh, M. Singh, and S. Gangwar, “Emotion Recognition Using Electroencephalography (EEG): A Review,” 2013.

[10] M. Yadava, P. Kumar, R. Saini, P. P. Roy, and D. Prosad Dogra, “Analysis of EEG signals and its application to neuromarketing,” *Multimed Tools Appl*, vol. 76, no. 18, pp. 19087–19111, Sep. 2017, doi: 10.1007/s11042-017-4580-6.

[11] S. Kumar, M. Yadava, and P. P. Roy, “Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction,”

Information Fusion, vol. 52, pp. 41–52, Dec. 2019, doi: 10.1016/j.inffus.2018.11.001.

[12] Y. Gong, W. Hou, Q. Zhang, and S. Tian, "Discounts or gifts? Not just to save money," *Journal of Contemporary Marketing Science*, vol. 1, no. 1, pp. 53–75, Sep. 2018, doi: 10.1108/JCMARS-08-2018-0009.

[13] M. Aldayel, M. Ykhlef, and A. Al-Nafjan, "Consumers' Preference Recognition Based on Brain–Computer Interfaces: Advances, Trends, and Applications," *Arab J Sci Eng*, vol. 46, no. 9, pp. 8983–8997, Sep. 2021, doi: 10.1007/s13369-021-05695-4.

[14] M. Ramirez, S. Kaheh, and K. George, "Neuromarketing Study Using Machine Learning for Predicting Purchase Decision," in *2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, IEEE, Dec. 2021, pp. 0560–0564. doi: 10.1109/UEMCON53757.2021.9666539.

[15] B. Izadi, A. Ghaedi, and M. Ghasemian, "Neuropsychological responses of consumers to promotion strategies and the decision to buy sports products," *Asia Pacific Journal of Marketing and Logistics*, vol. 34, no. 6, pp. 1203–1221, Jun. 2022, doi: 10.1108/APJML-01-2021-0026.

[16] N. Alnuman, S. Al-Nasser, and O. Yasin, "Classification of Products Preference from EEG Signals using SVM Classifier," in *2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE)*, IEEE, Oct. 2020, pp. 174–179. doi: 10.1109/ICITEE49829.2020.9271669.

[17] P. Mateusz and N. Kesra, "Cognitive neuroscience in the design process of social advertising," *Procedia Comput Sci*, vol. 176, pp. 2959–2968, 2020, doi: 10.1016/j.procs.2020.09.207.

[18] A. González-Morales, "Right evaluation of marketing stimuli with neuroscience. An electroencephalography experiment," *Computers in Human Behavior Reports*, vol. 2, p. 100030, Aug. 2020, doi: 10.1016/j.chbr.2020.100030.

[19] A. Wajid, M. M. Raziq, Q. M. Ahmed, and M. Ahmad, "Observing viewers' self-reported and neurophysiological responses to message appeal in social media advertisements," *Journal of Retailing and Consumer Services*, vol. 59, p. 102373, Mar. 2021, doi: 10.1016/j.jretconser.2020.102373.

[20] A. C. Martinez-Levy *et al.*, "Message framing, non-conscious perception and effectiveness in non-profit advertising. Contribution by neuromarketing research," *International Review on Public and Nonprofit Marketing*, vol. 19, no. 1, pp. 53–75, Mar. 2022, doi: 10.1007/s12208-021-00289-0.

[21] A. Delorme, "EEG is better left alone," *Sci Rep*, vol. 13, no. 1, p. 2372, Feb. 2023, doi: 10.1038/s41598-023-27528-0.

[22] M. Shahriari, D. Feiz, A. Zarei, and E. Kashi, “The Meta-Analysis of Neuro-Marketing Studies: Past, Present and Future,” *Neuroethics*, vol. 13, no. 3, pp. 261–273, Oct. 2020, doi: 10.1007/s12152-019-09400-z.
[23] L. Hsu and Y.-J. Chen, “Does coffee taste better with latte art? A neuroscientific perspective,” *British Food Journal*, vol. 123, no. 5, pp. 1931–1946, Apr. 2021, doi: 10.1108/BFJ-07-2020-0612.
[24] S. V. Karpova, I. V. Rozhkov, and O. E. Ustinova, “Neurolinguistic and neuromarketing effects on consumer behavior,” in *Proceedings of the International Scientific and Practical Conference on Digital Economy (ISCDE 2019)*, Paris, France: Atlantis Press, 2019. doi: 10.2991/iscde-19.2019.82.
[25] C. M. L. Cruz, J. F. De Medeiros, L. C. R. Hermes, A. Marcon, and Ã. Marcon, “Neuromarketing and the advances in the consumer behaviour studies: a systematic review of the literature,” *International Journal of Business and Globalisation*, vol. 17, no. 3, pp. 330–351, 2016, Accessed: Oct. 06, 2022. [Online]. Available: https://ideas.repec.org/a/ids/ijbglo/v17y2016i3p330-351.html
[26] M. Hsu, “Neuromarketing: Inside the Mind of the Consumer,” *http://dx.doi.org/10.1177/0008125617720208*, vol. 59, no. 4, pp. 5–22, Aug. 2017, doi: 10.1177/0008125617720208.
[27] S. D. Shaw and R. P. Bagozzi, “The neuropsychology of consumer behavior and marketing,” *Consumer Psychology Review*, vol. 1, no. 1, pp. 22–40, Jan. 2018, doi: 10.1002/arcp.1006.
[28] N. A. Vences, J. Díaz-Campo, and D. F. G. Rosales, “Neuromarketing as an Emotional Connection Tool Between Organizations and Audiences in Social Networks. A Theoretical Review.,” *Front Psychol*, vol. 11, p. 1787, 2020, doi: 10.3389/fpsyg.2020.01787.
[29] A. H. Alsharif, N. Z. M. Salleh, M. Abdullah, A. Khraiwish, and A. Ashaari, “Neuromarketing Tools Used in the Marketing Mix: A Systematic Literature and Future Research Agenda,” *Sage Open*, vol. 13, no. 1, p. 215824402311565, Jan. 2023, doi: 10.1177/21582440231156563.
[30] S. Bhardwaj, G. A. Rana, A. Behl, and S. J. Gallego de Caceres, “Exploring the boundaries of Neuromarketing through systematic investigation,” *J Bus Res*, vol. 154, p. 113371, Jan. 2023, doi: 10.1016/j.jbusres.2022.113371.
[31] A. N. McInnes, B. Sung, and R. Hooshmand, “A practical review of electroencephalography’s value to consumer research,” *International Journal of Market Research*, vol. 65, no. 1, pp. 52–82, Jan. 2023, doi: 10.1177/14707853221112622.
[32] J. Siddique, A. Shamim, M. Nawaz, and M. F. Abid, “The hope and hype of neuromarketing: a bibliometric analysis,” *Journal of Contemporary Marketing Science*, vol. 6, no. 1, pp. 1–21, May 2023, doi: 10.1108/JCMARS-07-2022-0018.

[33] A. Bazzani, S. Ravaioli, L. Trieste, U. Faraguna, and G. Turchetti, "Is EEG Suitable for Marketing Research? A Systematic Review," *Front Neurosci*, vol. 14, Dec. 2020, doi: 10.3389/fnins.2020.594566.

[34] L. Robaina-Calderín and J. D. Martín-Santana, "A review of research on neuromarketing using content analysis: key approaches and new avenues," *Cogn Neurodyn*, vol. 15, no. 6, pp. 923–938, Dec. 2021, doi: 10.1007/s11571-021-09693-y.

[35] G. Pei and T. Li, "A Literature Review of EEG-Based Affective Computing in Marketing," *Front Psychol*, vol. 12, Mar. 2021, doi: 10.3389/fpsyg.2021.602843.

[36] A. Byrne, E. Bonfiglio, C. Rigby, and N. Edelstyn, "A systematic review of the prediction of consumer preference using EEG measures and machine-learning in neuromarketing research," *Brain Inform*, vol. 9, no. 1, p. 27, Dec. 2022, doi: 10.1186/s40708-022-00175-3.

[37] K. S. Khan, R. Kunz, J. Kleijnen, and G. Antes, "Five Steps to Conducting a Systematic Review," *J R Soc Med*, vol. 96, no. 3, pp. 118–121, Mar. 2003, doi: 10.1177/014107680309600304.

[38] D. Moher *et al.*, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement," *Syst Rev*, vol. 4, no. 1, p. 1, Dec. 2015, doi: 10.1186/2046-4053-4-1.

[39] A. Hakim, I. Golan, S. Yefet, and D. J. Levy, "DeePay: deep learning decodes EEG to predict consumer's willingness to pay for neuromarketing," *Front Hum Neurosci*, vol. 17, Jun. 2023, doi: 10.3389/fnhum.2023.1153413.

[40] A. Hakim, S. Klorfeld, T. Sela, D. Friedman, M. Shabat-Simon, and D. J. Levy, "Machines learn neuromarketing: Improving preference prediction from self-reports using multiple EEG measures and machine learning," *International Journal of Research in Marketing*, vol. 38, no. 3, pp. 770–791, Sep. 2021, doi: 10.1016/j.ijresmar.2020.10.005.

[41] J. Guixeres *et al.*, "Consumer Neuroscience-Based Metrics Predict Recall, Liking and Viewing Rates in Online Advertising," *Front Psychol*, vol. 8, Oct. 2017, doi: 10.3389/fpsyg.2017.01808.

[42] J. Teo, L. H. Chew, and J. Mountstephens, "Improving Subject-Independent EEG Preference Classification Using Deep Learning Architectures with Dropouts," 2019, pp. 549–560. doi: 10.1007/978-3-030-03402-3_38.

[43] M. Aldayel, M. Ykhlef, and A. Al-Nafjan, "Recognition of Consumer Preference by Analysis and Classification EEG Signals," *Front Hum Neurosci*, vol. 14, Jan. 2021, doi: 10.3389/fnhum.2020.604639.

[44] V. P. Oikonomou, K. Georgiadis, F. Kalaganis, S. Nikolopoulos, and I. Kompatsiaris, "A Sparse Representation Classification Scheme for the Recognition of Affective and Cognitive Brain Processes in

Neuromarketing," *Sensors*, vol. 23, no. 5, p. 2480, Feb. 2023, doi: 10.3390/s23052480.

[45] B. Li, Y. Wang, K. Wang, and J. Yang, "Application of CNN Deep Learning in Product Design Evaluation," 2019, pp. 517–526. doi: 10.1007/978-981-13-2375-1_65.

[46] C. Yen and M.-C. Chiang, "Examining the effect of online advertisement cues on human responses using eye-tracking, EEG, and MRI," *Behavioural Brain Research*, vol. 402, p. 113128, Mar. 2021, doi: 10.1016/j.bbr.2021.113128.

[47] L. Zeng, M. Lin, K. Xiao, J. Wang, and H. Zhou, "Like/Dislike Prediction for Sport Shoes With Electroencephalography: An Application of Neuromarketing," *Front Hum Neurosci*, vol. 15, Jan. 2022, doi: 10.3389/fnhum.2021.793952.

[48] S. Raiesdana and M. Mousakhani, "An EEG-Based Neuromarketing Approach for Analyzing the Preference of an Electric Car," *Comput Intell Neurosci*, vol. 2022, pp. 1–19, Mar. 2022, doi: 10.1155/2022/9002101.

[49] A. Kislov, A. Gorin, N. Konstantinovsky, V. Klyuchnikov, B. Bazanov, and V. Klucharev, "Central EEG Beta/Alpha Ratio Predicts the Population-Wide Efficiency of Advertisements," *Brain Sci*, vol. 13, no. 1, p. 57, Dec. 2022, doi: 10.3390/brainsci13010057.

[50] A. Al-Nafjan, "Feature selection of EEG signals in neuromarketing," *PeerJ Comput Sci*, vol. 8, p. e944, Apr. 2022, doi: 10.7717/peerj-cs.944.

[51] S. M. A. Shah *et al.*, "An Ensemble Model for Consumer Emotion Prediction Using EEG Signals for Neuromarketing Applications," *Sensors*, vol. 22, no. 24, p. 9744, Dec. 2022, doi: 10.3390/s22249744.

[52] A. Hassani, A. Hekmatmanesh, and A. M. Nasrabadi, "Discrimination of Customers Decision-Making in a Like/Dislike Shopping Activity Based on Genders: A Neuromarketing Study," *IEEE Access*, vol. 10, pp. 92454–92466, 2022, doi: 10.1109/ACCESS.2022.3201488.

[53] K. Georgiadis, F. P. Kalaganis, V. P. Oikonomou, S. Nikolopoulos, N. A. Laskaris, and I. Kompatsiaris, "RNeuMark: A Riemannian EEG Analysis Framework for Neuromarketing," *Brain Inform*, vol. 9, no. 1, p. 22, Dec. 2022, doi: 10.1186/s40708-022-00171-7.

[54] H. Göker, "Multi-channel EEG-based classification of consumer preferences using multitaper spectral analysis and deep learning model," *Multimed Tools Appl*, Oct. 2023, doi: 10.1007/s11042-023-17114-x.

[55] Y. Shen, W. Shan, and J. Luan, "Influence of aggregated ratings on purchase decisions: an event-related potential study," *Eur J Mark*, vol. 52, no. 1–2, pp. 147–158, Feb. 2018, doi: 10.1108/EJM-12-2016-0871.

[56] J.-P. Uhm, H.-W. Lee, and J.-W. Han, "Creating sense of presence in a virtual reality experience: Impact on neurophysiological arousal and attitude towards a winter sport," *Sport Management Review*, vol. 23, no. 4, pp. 588–600, Oct. 2020, doi: 10.1016/j.smr.2019.10.003.

[57] J. Gountas, S. Gountas, J. Ciorciari, and P. Sharma, “Looking beyond traditional measures of advertising impact: Using neuroscientific methods to evaluate social marketing messages,” *J Bus Res*, vol. 105, pp. 121–135, Dec. 2019, doi: 10.1016/j.jbusres.2019.07.011.

[58] M. Domracheva and S. Kulikova, “EEG correlates of perceived food product similarity in a cross-modal taste-visual task,” *Food Qual Prefer*, vol. 85, p. 103980, Oct. 2020, doi: 10.1016/j.foodqual.2020.103980.

[59] Z. Wei, C. Wu, X. Wang, A. Supratak, P. Wang, and Y. Guo, “Using Support Vector Machine on EEG for Advertisement Impact Assessment,” *Front Neurosci*, vol. 12, Mar. 2018, doi: 10.3389/fnins.2018.00076.

[60] T. Z. Ramsøy, M. Skov, M. K. Christensen, and C. Stahlhut, “Frontal Brain Asymmetry and Willingness to Pay,” *Front Neurosci*, vol. 12, Mar. 2018, doi: 10.3389/fnins.2018.00138.

[61] T. Yang, S. Lee, E. Seomoon, and S.-P. Kim, “Characteristics of Human Brain Activity during the Evaluation of Service-to-Service Brand Extension,” *Front Hum Neurosci*, vol. 12, Feb. 2018, doi: 10.3389/fnhum.2018.00044.

[62] Harris, Ciorciari, and Gountas, “Consumer Neuroscience and Digital/Social Media Health/Social Cause Advertisement Effectiveness,” *Behavioral Sciences*, vol. 9, no. 4, p. 42, Apr. 2019, doi: 10.3390/bs9040042.

[63] N. Goto *et al.*, “Can Brain Waves Really Tell If a Product Will Be Purchased? Inferring Consumer Preferences From Single-Item Brain Potentials,” *Front Integr Neurosci*, vol. 13, Jun. 2019, doi: 10.3389/fnint.2019.00019.

[64] E. Eijlers, M. A. S. Boksem, and A. Smidts, “Measuring Neural Arousal for Advertisements and Its Relationship With Advertising Success,” *Front Neurosci*, vol. 14, Jul. 2020, doi: 10.3389/fnins.2020.00736.

[65] A. Vozzi *et al.*, “The Sample Size Matters: To What Extent the Participant Reduction Affects the Outcomes of a Neuroscientific Research. A Case-Study in Neuromarketing Field,” *Sensors*, vol. 21, no. 18, p. 6088, Sep. 2021, doi: 10.3390/s21186088.

[66] M. Zito, A. Fici, M. Bilucaglia, F. S. Ambrogetti, and V. Russo, “Assessing the Emotional Response in Social Communication: The Role of Neuromarketing,” *Front Psychol*, vol. 12, Jun. 2021, doi: 10.3389/fpsyg.2021.625570.

[67] G. Wang, J. Li, C. Zhu, S. Wang, and S. Jiang, “How Do Reference Points Influence the Representation of the N200 for Consumer Preference?,” *Front Psychol*, vol. 12, Jun. 2021, doi: 10.3389/fpsyg.2021.645775.

[68] Q. Ma, W. Mao, and L. Hu, “Electrophysiological evidence for the effects of pain on the different stages of reward evaluation under a purchasing situation,” *Front Psychol*, vol. 13, Sep. 2022, doi: 10.3389/fpsyg.2022.943699.

[69] S. Kakaria, F. Saffari, T. Z. Ramsøy, and E. Bigné, “Cognitive load during planned and unplanned virtual shopping: Evidence from a

neurophysiological perspective," *Int J Inf Manage*, vol. 72, p. 102667, Oct. 2023, doi: 10.1016/j.ijinfomgt.2023.102667.
[70] M. Lukovics *et al.*, "Combining survey-based and neuroscience measurements in customer acceptance of self-driving technology," *Transp Res Part F Traffic Psychol Behav*, vol. 95, pp. 46–58, May 2023, doi: 10.1016/j.trf.2023.03.016.
[71] N. Goto *et al.*, "Neural signals of selective attention are modulated by subjective preferences and buying decisions in a virtual shopping task.," *Biol Psychol*, vol. 128, pp. 11–20, 2017, doi: 10.1016/j.biopsycho.2017.06.004.
[72] J. Jin, W. Zhang, and M. Chen, "How consumers are affected by product descriptions in online shopping: Event-related potentials evidence of the attribute framing effect," *Neurosci Res*, vol. 125, pp. 21–28, Dec. 2017, doi: 10.1016/j.neures.2017.07.006.
[73] T. Çakar, K. Rızvanoğlu, Ö. Öztürk, D. Z. Çelik, and İ. Gürvardar, "The use of neurometric and biometric research methods in understanding the user experience during product search of first-time buyers in E-commerce," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2017, pp. 342–362. doi: 10.1007/978-3-319-58634-2_26.
[74] M. Royo, V. Chulvi, E. Mulet, and J. Galán, "Users' reactions captured by means of an EEG headset on viewing the presentation of sustainable designs using verbal narrative," *Eur J Mark*, vol. 52, no. 1–2, pp. 159–181, Feb. 2018, doi: 10.1108/EJM-12-2016-0837.
[75] T. Daugherty, E. Hoffman, K. Kennedy, and M. Nolan, "Measuring consumer neural activation to differentiate cognitive processing of advertising," *Eur J Mark*, vol. 52, no. 1/2, pp. 182–198, Feb. 2018, doi: 10.1108/EJM-10-2017-0657.
[76] Z. Gholami Doborjeh, M. G. Doborjeh, and N. Kasabov, "Attentional Bias Pattern Recognition in Spiking Neural Networks from Spatio-Temporal EEG Data," *Cognit Comput*, vol. 10, no. 1, pp. 35–48, Feb. 2018, doi: 10.1007/s12559-017-9517-x.
[77] M. Alonso Dos Santos and F. Calabuig Moreno, "Assessing the effectiveness of sponsorship messaging: Measuring the impact of congruence through electroencephalogram," *International Journal of Sports Marketing and Sponsorship*, vol. 19, no. 1, pp. 25–40, 2018, doi: 10.1108/IJSMS-09-2016-0067.
[78] R. Gordon, J. Ciorciari, and T. van Laer, "Using EEG to examine the role of attention, working memory, emotion, and imagination in narrative transportation," *Eur J Mark*, vol. 52, no. 1/2, pp. 92–117, Feb. 2018, doi: 10.1108/EJM-12-2016-0881.
[79] J. García-Madariaga, M.-F. Blasco López, I. M. Burgos, and N. R. Virto, "Do isolated packaging variables influence consumers' attention and

preferences?," *Physiol Behav*, vol. 200, pp. 96–103, Mar. 2019, doi: 10.1016/j.physbeh.2018.04.030.

[80] H. Fu, H. Ma, J. Bian, C. Wang, J. Zhou, and Q. Ma, "Don't trick me: An event-related potentials investigation of how price deception decreases consumer purchase intention," *Neurosci Lett*, vol. 713, p. 134522, Nov. 2019, doi: 10.1016/j.neulet.2019.134522.

[81] P. Golnar-Nik, S. Farashi, and M.-S. Safari, "The application of EEG power for the prediction and interpretation of consumer decision-making: A neuromarketing study," *Physiol Behav*, vol. 207, pp. 90–98, Aug. 2019, doi: 10.1016/j.physbeh.2019.04.025.

[82] J. Sänger, "Can't take my eyes off you – How task irrelevant pictures of food influence attentional selection," *Appetite*, vol. 133, pp. 313–323, Feb. 2019, doi: 10.1016/j.appet.2018.11.030.

[83] L. Alvino, R. van der Lubbe, R. A. M. Joosten, and E. Constantinides, "Which wine do you prefer? An analysis on consumer behaviour and brain activity during a wine tasting experience," *Asia Pacific Journal of Marketing and Logistics*, vol. 32, no. 5, pp. 1149–1170, Jun. 2020, doi: 10.1108/APJML-04-2019-0240.

[84] N. Verhulst, I. Vermeir, H. Slabbinck, B. Larivière, M. Mauri, and V. Russo, "A neurophysiological exploration of the dynamic nature of emotions during the customer experience," *Journal of Retailing and Consumer Services*, vol. 57, p. 102217, Nov. 2020, doi: 10.1016/j.jretconser.2020.102217.

[85] M. Zubair, X. Wang, S. Iqbal, M. Awais, and R. Wang, "Attentional and emotional brain response to message framing in context of green marketing," *Heliyon*, vol. 6, no. 9, p. e04912, Sep. 2020, doi: 10.1016/j.heliyon.2020.e04912.

[86] L. Hsu and Y.-J. Chen, "Neuromarketing, Subliminal Advertising, and Hotel Selection: An EEG Study," *Australasian Marketing Journal*, vol. 28, no. 4, pp. 200–208, Nov. 2020, doi: 10.1016/j.ausmj.2020.04.009.

[87] K. M. Pagan, J. de M. E. Giraldi, V. Maheshwari, A. L. D. de Paula, and J. H. C. de Oliveira, "Evaluating cognitive processing and preferences through brain responses towards country of origin for wines: the role of gender and involvement," *International Journal of Wine Business Research*, vol. 33, no. 4, pp. 481–501, Nov. 2021, doi: 10.1108/IJWBR-08-2020-0043.

[88] M. Zhao and X. Wang, "Perception value of product-service systems: Neural effects of service experience and customer knowledge," *Journal of Retailing and Consumer Services*, vol. 62, p. 102617, Sep. 2021, doi: 10.1016/j.jretconser.2021.102617.

[89] J. Robertson, C. Ferreira, M. Reyneke, and D. Rosenstein, "I think I like this: assessing conscious versus subconscious wine taste responses using neuroscientific techniques," *International Journal of Wine Business Research*, vol. 34, no. 1, pp. 37–51, Feb. 2022, doi: 10.1108/IJWBR-09-2020-0047.

[90] C. Wang, Y. Li, X. Luo, Q. Ma, W. Fu, and H. Fu, “The Effects of Money on Fake Rating Behavior in E-Commerce: Electrophysiological Time Course Evidence From Consumers,” *Front Neurosci*, vol. 12, Mar. 2018, doi: 10.3389/fnins.2018.00156.

[91] Q. Ma, L. Zhang, and M. Wang, “‘You Win, You Buy’—How Continuous Win Effect Influence Consumers’ Price Perception: An ERP Study,” *Front Neurosci*, vol. 12, Oct. 2018, doi: 10.3389/fnins.2018.00691.

[92] W. Yu, Z. Sun, T. Xu, and Q. Ma, “Things Become Appealing When I Win: Neural Evidence of the Influence of Competition Outcomes on Brand Preference,” *Front Neurosci*, vol. 12, Oct. 2018, doi: 10.3389/fnins.2018.00779.

[93] A. Mengual-Recuerda, V. Tur-Viñes, and D. Juárez-Varón, “Neuromarketing in Haute Cuisine Gastronomic Experiences,” *Front Psychol*, vol. 11, Aug. 2020, doi: 10.3389/fpsyg.2020.01772.

[94] Y. Kim, K. Park, Y. Kim, W. Yang, D. Han, and W.-S. Kim, “The Impact of Visual Art and High Affective Arousal on Heuristic Decision-Making in Consumers,” *Front Psychol*, vol. 11, Nov. 2020, doi: 10.3389/fpsyg.2020.565829.

[95] S. Wang, C. Xu, L. Xiao, and A. S. Ding, “The Implicit Aesthetic Preference for Mobile Marketing Interface Layout—An ERP Study,” *Front Hum Neurosci*, vol. 15, Sep. 2021, doi: 10.3389/fnhum.2021.728895.

[96] L. Alvino, E. Constantinides, and R. H. J. van der Lubbe, “Consumer Neuroscience: Attentional Preferences for Wine Labeling Reflected in the Posterior Contralateral Negativity,” *Front Psychol*, vol. 12, Oct. 2021, doi: 10.3389/fpsyg.2021.688713.

[97] V. Russo *et al.*, “The Role of the Emotional Sequence in the Communication of the Territorial Cheeses: A Neuromarketing Approach,” *Foods*, vol. 11, no. 15, p. 2349, Aug. 2022, doi: 10.3390/foods11152349.

[98] M. Wang *et al.*, “Pleasure of paying when using mobile payment: Evidence from EEG studies,” *Front Psychol*, vol. 13, Oct. 2022, doi: 10.3389/fpsyg.2022.1004068.

[99] A. Hassani, A. Hekmatmanesh, and A. M. Nasrabadi, “Gender Differences in EEG Responses to Color and Black-White Images: Implications for Neuromarketing Strategies,” *IEEE Access*, vol. 11, pp. 93739–93753, 2023, doi: 10.1109/ACCESS.2023.3308810.

[100] Q. Wei *et al.*, “The Influence of Tourist Attraction Type on Product Price Perception and Neural Mechanism in Tourism Consumption: An ERP Study,” *Psychol Res Behav Manag*, vol. Volume 16, pp. 3787–3803, Sep. 2023, doi: 10.2147/PRBM.S416821.

[101] A. L. Damião de Paula, M. Lourenção, J. de Moura Engracia Giraldi, and J. H. Caldeira de Oliveira, “Effect of emotion induction on potential consumers’ visual attention in beer advertisements: a neuroscience study,” *Eur J Mark*, vol. 57, no. 1, pp. 202–225, Jan. 2023, doi: 10.1108/EJM-06-2021-0448.

[102] V. Russo *et al.*, “Neuroselling: applying neuroscience to selling for a new business perspective. An analysis on teleshopping advertising,” *Front Psychol*, vol. 14, Oct. 2023, doi: 10.3389/fpsyg.2023.1238879.
[103] S. Bosshard and P. Walla, “Sonic Influence on Initially Neutral Brands: Using EEG to Unveil the Secrets of Audio Evaluative Conditioning,” *Brain Sci*, vol. 13, no. 10, p. 1393, Sep. 2023, doi: 10.3390/brainsci13101393.
[104] L. Song, Z. Mo, and J. Liu, “Event-related potentials evidence of how location contiguity influences consumer purchase intentions,” *Neurosci Lett*, vol. 814, p. 137472, Sep. 2023, doi: 10.1016/j.neulet.2023.137472.
[105] T. Avinash, L. Dikshant, and S. Seema, “Methods of Neuromarketing and Implication of the Frontal Theta Asymmetry induced due to musical stimulus as choice modeling,” *Procedia Comput Sci*, vol. 132, pp. 55–67, 2018, doi: 10.1016/j.procs.2018.05.059.
[106] I. Michael, T. Ramsoy, M. Stephens, and F. Kotsi, “A study of unconscious emotional and cognitive responses to tourism images using a neuroscience method,” *Journal of Islamic Marketing*, vol. 10, no. 2, pp. 543–564, Jun. 2019, doi: 10.1108/JIMA-09-2017-0098.
[107] L. Hsu and Y. jung Chen, “Music and wine tasting: an experimental neuromarketing study,” *British Food Journal*, vol. 122, no. 8, pp. 2725–2737, Jun. 2020, doi: 10.1108/BFJ-06-2019-0434.
[108] K. Pennanen, J. Närväinen, S. Vanhatalo, R. Raisamo, and N. Sozer, “Effect of virtual eating environment on consumers’ evaluations of healthy and unhealthy snacks,” *Food Qual Prefer*, vol. 82, p. 103871, Jun. 2020, doi: 10.1016/j.foodqual.2020.103871.
[109] N. Leeuwis, D. Pistone, N. Flick, and T. van Bommel, “A Sound Prediction: EEG-Based Neural Synchrony Predicts Online Music Streams,” *Front Psychol*, vol. 12, Jul. 2021, doi: 10.3389/fpsyg.2021.672980.
[110] J.-P. Uhm, H.-W. Lee, J.-W. Han, and D.-K. Kim, “Effect of background music and hierarchy-of-effects in watching women’s running shoes advertisements,” *International Journal of Sports Marketing and Sponsorship*, vol. 23, no. 1, pp. 41–58, Jan. 2022, doi: 10.1108/IJSMS-09-2020-0159.
[111] Q. Ma, M. Wang, and Q. Da, “The effects of brand familiarity and product category in brand extension: An ERP study,” *Neurosci Res*, vol. 169, pp. 48–56, Aug. 2021, doi: 10.1016/j.neures.2020.06.010.
[112] U. Garczarek-Bąk, A. Szymkowiak, P. Gaczek, and A. Disterheft, “A comparative analysis of neuromarketing methods for brand purchasing predictions among young adults,” *Journal of Brand Management*, vol. 28, no. 2, pp. 171–185, Mar. 2021, doi: 10.1057/S41262-020-00221-7/TABLES/6.
[113] A. Özbeyaz, “EEG-Based classification of branded and unbranded stimuli associating with smartphone products: comparison of several machine learning algorithms,” *Neural Comput Appl*, vol. 33, no. 9, pp. 4579–4593, May 2021, doi: 10.1007/s00521-021-05779-0.

[114] T. Yang and S.-P. Kim, "Group-Level Neural Responses to Service-to-Service Brand Extension," *Front Neurosci*, vol. 13, Jun. 2019, doi: 10.3389/fnins.2019.00676.
[115] F. Camarrone and M. M. Van Hulle, "Measuring brand association strength with EEG: A single-trial N400 ERP study," *PLoS One*, vol. 14, no. 6, p. e0217125, Jun. 2019, doi: 10.1371/journal.pone.0217125.
[116] J. D. Trobe, "The Human Brain. An Introduction to Its Functional Anatomy, 6th Edition," *Journal of Neuro-Ophthalmology*, vol. 30, no. 1, p. 107, Mar. 2010, doi: 10.1097/01.wno.0000369168.32606.54.
[117] B. Touchette and S.-E. Lee, "Measuring Neural Responses to Apparel Product Attractiveness," *Clothing and Textiles Research Journal*, vol. 35, no. 1, pp. 3–15, Jan. 2017, doi: 10.1177/0887302X16673157.
[118] S. K. Hadjidimitriou and L. J. Hadjileontiadis, "Toward an EEG-Based Recognition of Music Liking Using Time-Frequency Analysis," *IEEE Trans Biomed Eng*, vol. 59, no. 12, pp. 3498–3510, Dec. 2012, doi: 10.1109/TBME.2012.2217495.
[119] M. Aldayel, M. Ykhlef, and A. Al-Nafjan, "Deep Learning for EEG-Based Preference Classification in Neuromarketing," *Applied Sciences*, vol. 10, no. 4, p. 1525, Feb. 2020, doi: 10.3390/app10041525.
[120] E. Modica *et al.*, "Neurophysiological Responses to Different Product Experiences," *Comput Intell Neurosci*, vol. 2018, pp. 1–10, Sep. 2018, doi: 10.1155/2018/9616301.
[121] A. Al-Nafjan, M. Hosny, A. Al-Wabil, and Y. Al-Ohali, "Classification of Human Emotions from Electroencephalogram (EEG) Signal using Deep Neural Network," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 9, 2017, doi: 10.14569/IJACSA.2017.080955.
[122] Abraham. Savitzky and M. J. E. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures.," *Anal Chem*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964, doi: 10.1021/ac60214a047.
[123] L. Vega-Escobar, A. E. Castro-Ospina, and L. Duque-Munoz, "DWT-based feature extraction for motor imagery classification," in *6th Latin-American Conference on Networked and Electronic Media (LACNEM 2015)*, Institution of Engineering and Technology, 2015, pp. 3 (6 .)-3 (6 .). doi: 10.1049/ic.2015.0309.
[124] L. Chen, Y. Zhao, J. Zhang, and J. Zou, "Automatic detection of alertness/drowsiness from physiological signals using wavelet-based nonlinear features and machine learning," *Expert Syst Appl*, vol. 42, no. 21, pp. 7344–7355, Nov. 2015, doi: 10.1016/j.eswa.2015.05.028.
[125] Y. Xie and S. Oniga, "A Review of Processing Methods and Classification Algorithm for EEG Signal," *Carpathian Journal of Electronic and Computer Engineering*, vol. 13, no. 1, pp. 23–29, Sep. 2020, doi: 10.2478/cjece-2020-0004.
[126] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified

periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, Jun. 1967, doi: 10.1109/TAU.1967.1161901.

[127] V. De Pascalis, "On the Psychophysiology of Extraversion," *On the Psychobiology of Personality: Essays in Honor of Marvin Zuckerman*, pp. 295–327, Jan. 2004, doi: 10.1016/B978-008044209-9/50017-8.

[128] H. Ma, Z. Mo, H. Zhang, C. Wang, and H. Fu, "The Temptation of Zero Price: Event-Related Potentials Evidence of How Price Framing Influences the Purchase of Bundles," *Front Neurosci*, vol. 12, Apr. 2018, doi: 10.3389/fnins.2018.00251.

[129] Y.-Z. Shi, K.-M. Cheung, and G. Prendergast, "Behavioural response to sales promotion tools," *Int J Advert*, vol. 24, no. 4, pp. 469–489, Jan. 2005, doi: 10.1080/02650487.2005.11072939.

[130] M. M. Botvinick, T. S. Braver, D. M. Barch, C. S. Carter, and J. D. Cohen, "Conflict monitoring and cognitive control.," *Psychol Rev*, vol. 108, no. 3, pp. 624–652, 2001, doi: 10.1037/0033-295X.108.3.624.

[131] L. E. Martin and G. F. Potts, "Impulsivity in decision-making: An event-related potential investigation," *Pers Individ Dif*, vol. 46, no. 3, pp. 303–308, Feb. 2009, doi: 10.1016/j.paid.2008.10.019.

[132] F. Mushtaq, G. Stoet, A. R. Bland, and A. Schaefer, "Relative Changes from Prior Reward Contingencies Can Constrain Brain Correlates of Outcome Monitoring," *PLoS One*, vol. 8, no. 6, p. e66350, Jun. 2013, doi: 10.1371/journal.pone.0066350.

[133] J. Decety, K. L. Meidenbauer, and J. M. Cowell, "The development of cognitive empathy and concern in preschool children: A behavioral neuroscience investigation," *Dev Sci*, vol. 21, no. 3, May 2018, doi: 10.1111/desc.12570.

[134] J. R. Folstein and C. Van Petten, "Influence of cognitive control and mismatch on the N2 component of the ERP: A review," *Psychophysiology*, vol. 45, no. 1, pp. 152–170, 2008, doi: 10.1111/J.1469-8986.2007.00602.X.

[135] K. J. Bruin and A. A. Wijers, "Inhibition, response mode, and stimulus probability: A comparative event-related potential study," *Clinical Neurophysiology*, vol. 113, no. 7, pp. 1172–1182, 2002, doi: 10.1016/S1388-2457(02)00141-4.

[136] M. Mennes, H. Wouters, B. Van Den Bergh, L. Lagae, and P. Stiers, "ERP correlates of complex human decision making in a gambling paradigm: Detection and resolution of conflict," *Psychophysiology*, vol. 45, no. 5, pp. 714–720, Sep. 2008, doi: 10.1111/j.1469-8986.2008.00678.x.

[137] J. Qin and S. Han, "Neurocognitive mechanisms underlying identification of environmental risks," *Neuropsychologia*, vol. 47, no. 2, pp. 397–405, Jan. 2009, doi: 10.1016/j.neuropsychologia.2008.09.010.

[138] Y. Liu, H. Huang, M. McGinnis-Deweese, A. Keil, and M. Ding, "Neural Substrate of the Late Positive Potential in Emotional Processing," *The*

Journal of Neuroscience, vol. 32, no. 42, pp. 14563–14572, Oct. 2012, doi: 10.1523/JNEUROSCI.3109-12.2012.

[139] S. Watts, L. G. Buratto, E. V. Brotherhood, G. E. Barnacle, and A. Schaefer, "The neural fate of neutral information in emotion-enhanced memory," *Psychophysiology*, vol. 51, no. 7, pp. 673–684, Jul. 2014, doi: 10.1111/psyp.12211.

[140] R. J. Davidson, "What does the prefrontal cortex 'do' in affect: perspectives on frontal EEG asymmetry research," *Biol Psychol*, vol. 67, no. 1–2, pp. 219–234, Oct. 2004, doi: 10.1016/j.biopsycho.2004.03.008.

[141] H. Blaiech, M. Neji, A. Wali, and A. M. Alimi, "Emotion recognition by analysis of EEG signals," in *13th International Conference on Hybrid Intelligent Systems (HIS 2013)*, IEEE, Dec. 2013, pp. 312–318. doi: 10.1109/HIS.2013.6920451.

[142] G. Vecchiato *et al.*, "Changes in Brain Activity During the Observation of TV Commercials by Using EEG, GSR and HR Measurements," *Brain Topogr*, vol. 23, no. 2, pp. 165–179, Jun. 2010, doi: 10.1007/s10548-009-0127-0.

[143] G. Vecchiato *et al.*, "Spectral EEG frontal asymmetries correlate with the experienced pleasantness of TV commercial advertisements," *Med Biol Eng Comput*, vol. 49, no. 5, pp. 579–583, May 2011, doi: 10.1007/s11517-011-0747-x.

[144] M. G. Wisniewski, E. R. Thompson, N. Iyer, J. R. Estepp, M. N. Goder-Reiser, and S. C. Sullivan, "Frontal midline θ power as an index of listening effort," *Neuroreport*, vol. 26, no. 2, pp. 94–99, Jan. 2015, doi: 10.1097/WNR.0000000000000306.

[145] D.-W. Chen *et al.*, "A Feature Extraction Method Based on Differential Entropy and Linear Discriminant Analysis for Emotion Recognition," *Sensors*, vol. 19, no. 7, p. 1631, Apr. 2019, doi: 10.3390/s19071631.

[146] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr Clin Neurophysiol*, vol. 29, no. 3, pp. 306–310, Sep. 1970, doi: 10.1016/0013-4694(70)90143-4.

[147] P. Mishra, C. Pandey, U. Singh, A. Keshri, and M. Sabaretnam, "Selection of appropriate statistical methods for data analysis," *Ann Card Anaesth*, vol. 22, no. 3, p. 297, 2019, doi: 10.4103/aca.ACA_248_18.

[148] S. M. Usman, S. M. Ali Shah, O. C. Edo, and J. Emakhu, "A Deep Learning Model for Classification of EEG Signals for Neuromarketing," in *2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*, IEEE, Mar. 2023, pp. 1–6. doi: 10.1109/ITIKD56332.2023.10100014.

[149] W. Wichienchai and Y. Wongsawat, "EEG Feature Selection for Subjective Preference," in *2020 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, IEEE, Sep. 2020, pp. 993–997. doi: 10.23919/SICE48898.2020.9240470.

[150] M. Taghaddossi and M. H. Moradi, “Brain Effective Connectivity Comparison in Different States of Familiarity and Desiring Brands Confrontation: A Neuromarketing Study,” in *2021 29th Iranian Conference on Electrical Engineering (ICEE)*, IEEE, May 2021, pp. 907–911. doi: 10.1109/ICEE52715.2021.9544433.
[151] S. Kaheh, M. Ramirez, and K. George, “Study on the Effect of Product Brand and Pricing Using Biosignals,” in *2021 IEEE XXVIII International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, IEEE, Aug. 2021, pp. 1–4. doi: 10.1109/INTERCON52678.2021.9532800.
[152] Md. F. K. Khondakar, T. G. Troyee, M. Hasan, Md. H. Sarowar, M. H. Chowdhury, and Q. D. Hossain, “A Comparative Analysis of Different Pre-processing Pipelines for EEG-based Preference Prediction in Neuromarketing,” in *IEEE PuneCon 2023 (in press)*, Pune, India: IEEE, Dec. 2023.
[153] A. Delorme and S. Makeig, “EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” *J Neurosci Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004, doi: 10.1016/J.JNEUMETH.2003.10.009.
[154] H. Kim, J. Luo, S. Chu, C. Cannard, S. Hoffmann, and M. Miyakoshi, “ICA’s bug: How ghost ICs emerge from effective rank deficiency caused by EEG electrode interpolation and incorrect re-referencing,” *Frontiers in Signal Processing*, vol. 3, Apr. 2023, doi: 10.3389/frsip.2023.1064138.
[155] H. Bokil, P. Andrews, J. E. Kulkarni, S. Mehta, and P. P. Mitra, “Chronux: A platform for analyzing neural signals,” *J Neurosci Methods*, vol. 192, no. 1, pp. 146–151, Sep. 2010, doi: 10.1016/j.jneumeth.2010.06.020.
[156] M. Plechawska-Wójcik, P. Augustynowicz, M. Kaczorowska, E. Zabielska-Mendyk, and D. Zapała, “The Influence Assessment of Artifact Subspace Reconstruction on the EEG Signal Characteristics,” *Applied Sciences*, vol. 13, no. 3, p. 1605, Jan. 2023, doi: 10.3390/app13031605.
[157] S. K. Bandara, U. C. Wijesinghe, B. P. Jayalath, S. K. Bandara, P. S. Haddela, and L. M. Wickramasinghe, “EEG Based Neuromarketing Recommender System for Video Commercials,” in *2021 IEEE 16th International Conference on Industrial and Information Systems (ICIIS)*, IEEE, Sep. 2021, pp. 11–16. doi: 10.1109/ICIIS53135.2021.9660742.
[158] M. Ogino and Y. Mitsukura, “A Mobile Application for Estimating Emotional Valence Using a Single-Channel EEG Device,” in *2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, IEEE, Sep. 2018, pp. 1043–1048. doi: 10.23919/SICE.2018.8492583.
[159] M. Ramirez, S. Kaheh, M. A. Khalil, and K. George, “Application of Convolutional Neural Network for Classification of Consumer Preference from Hybrid EEG and FNIRS Signals,” in *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)*, IEEE, Jan. 2022, pp. 1024–1028. doi: 10.1109/CCWC54503.2022.9720831.

[160] T. Zoëga Ramsøy, N. Michael, and I. Michael, "A Consumer Neuroscience Study of Conscious and Subconscious Destination Preference," *Sci Rep*, vol. 9, no. 1, p. 15102, Oct. 2019, doi: 10.1038/s41598-019-51567-1.

[161] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.

[162] R. Circi, D. Gatti, V. Russo, and T. Vecchi, "The foreign language effect on decision-making: A meta-analysis," *Psychon Bull Rev*, vol. 28, no. 4, pp. 1131–1141, Aug. 2021, doi: 10.3758/s13423-020-01871-z.

Appendix A

Synthesized summary of the reviewed research articles chronologically from 2017 to 2023

Reference	Number of Participants	Stimuli Used	EEG Device Used	Pre-processing	Features Extraction Technique - Features	Machine Learning/ Statistical Analyzing Techniques
[10]	40	Product images in E-commerce	Emotiv EPOC+ (14 ch.)	S-Golay Filter	DWT (DB4) S-Band Oscillation (gamma, beta, alpha, theta, and delta)	HMM (70.33% for male and 63.56% for female)
[71]	38	Product images in VR shopping	ASALAB, ANT Neuro (32 ch.)	BPF (0.01-40 Hz) and Segmentation	ERP components – N200, LPP, and PSW	ANOVA
[72]	18	Product images in online shopping	Neuroscan Synamp2 (64 ch.)	LPF (30 Hz) and Segmentation	ERP components – P2, P2-N2 complex and LPP	ANOVA, t-test
[73]	19	Product images in E-commerce	g.Nautilus (32 ch.)	LPF (40 Hz), HPF (0.1 Hz)	Power of alpha band, FAA Index	t-test
[41]	35	Video Ads	REFA 32, TMSI (32 ch.)	ICA, BPF	GFP, Pleasantness Index (PI), Interest Index (II)	ANN (82.9%), Shapiro-Wilk tests, ANOVA
[74]	54	Video Ads	Emotiv EPOC+ (14 ch.)	BPF (0.2-45 Hz), Notch Filter (50 and 60 Hz)	Level of Intensity of Cognitive States	ANOVA
[105]	41	Music	Medicaid System (32 ch.)	Segmentation, Averaging	FFT – Power Spectrum, PSD (theta band)	k-NN (82.75%), LDA (90%)
[12]	20	Product images	Neuroscan (64ch.)	BPF (0.08-30 Hz)	ERP components – P2, N2, and P3	t-test

[77]	60	Images and texts	NeuroSky MindWave (1 ch.)	BPF, Notch	FFT – Power Spectrum (theta, alpha, delta, and beta)	Levene's and Shapiro-Wilk's test, Kruskal-Wallis test, Wilcoxon signed-ranks test, chi-square test
[55]	22	Product images and ratings	Net Amps 300, Geodesic Inc. (64 ch.)	BPF (0.03-100 Hz), LPF (30 Hz), Segmentation	ERP components – LPP	ANOVA
[75]	21	Video Ads	Geodesic EEG System (256 ch.)	BPF (0.3-45 Hz)	Power Spectrum (theta, alpha, beta), ERP	sLORETA
[76]	9	Product image and texts	Cognionics Instruments, Dry EEG system (32 ch.)	-	ERP components – N100, P200	MLR (15.56%), MLP (17.78%), SVM (13.33%), SNN (90%)
[78]	16	Video Ads	Neuroscan Synamp2 (64 ch.)	-	Power Spectra (theta, alpha), GFP	LORETA, ANOVA
[59]	30	Video Ads	NeuroSky MindWave (1 ch.)	-	FFT – Band Power (gamma, beta, alpha, theta, and delta), Statistical Features	SVM (75%)
[90]	21	Product images	Neuroscan Synamp2 (64 ch.)	LPF (30 Hz), Segmentation	ERP (N1, N2, LPP)	ANOVA
[91]	36	Product images	Neuroscan Synamp2 (64 ch.)	LPF (30 Hz), Segmentation	ERP (P300, LPP, P2)	t-test, ANOVA
[92]	21	Product images	Neuroscan Synamp2 (64 ch.)	LPF (30 Hz), Segmentation	ERP (N1, N2, LPP)	t-test, ANOVA

[60]	16	Product images and Price	Emotiv EPOC+ (14 ch.)	Baseline removal, Bad channel removal	Power Spectra	t-test, ANOVA
[61]	19	Images	ActiChamp, Brain Vision (9 ch.)	FIR HPF (0.5 Hz), FIR Notch (50 Hz), ICA	ERP (N2, P300, and N400)	t-test, ANOVA
[114]	19	Product images	ActiChamp, Brain Vision (32 ch.)	FIR HPF (0.5 Hz), FIR Notch (50 Hz), ICA	ERP (N2, P300)	t-test, SVM
[62]	40	Video Ads	Neuroscan SynAmpsRT (64 ch.)	Segmentation, Bad channel removal, IIR filter	FFT – Average Band Power (theta, alpha, and beta)	ANOVA, t-test, sLORETA
[115]	26	Video Ads and Product Images	BioSemi Active Two (32 ch.)	BPF (0.5-30 Hz)	ERP - N400	Statistical analysis – pairwise comparisons
[63]	36	Product images	ANT Neuro (32 ch.)	BPF (0.01–30 Hz), Segmentation	ERP (N200), LPP, PSW	ANOVA
[82]	44	Food vs. neutral pictures	QuickAmp, Brain Products (64 ch.)	LPF (200 Hz), Segmentation	ERP components – N1pc, P2, P3	ANOVA
[11]	40	Online product images	Emotiv EPOC+ (14 ch.)	S-Golay Filter	DWT – DB4 (gamma, beta, alpha, theta, and delta)	RF & ABC (RMSE= 0.29 and R^2 = 0.72)
[79]	40	Product images	EEG Bioresearch, BitBrain Tech. (16 ch.)	-	Attention Level – alpha/theta index	ANOVA
[106]	30	Images and videos	ABM X-10 EEG (10 ch.)	-	NeuroMetric scores - Emotional Arousal, Emotional Motivation, and Cognitive Load	Descriptive Statistics
[80]	17	Product images and Price	eegoAmplifier, ANT Neuro	LPF (30 Hz)	ERP components – N2 and LPP	ANOVA

			(64 ch.)			
[57]	40	Video Ads	-	BPF (0.5-40 Hz)	GFP	sLORETA Analysis, t-test
[81]	16	Video Ads	EEG8 amplifier (32 ch.)	ICA	Welch method – PSD (delta, beta, alpha, theta)	SVM and LDA (87%)
[45]	-	Product	-	Filtering, Signal Denoising	FFT – Power Spectrum, CNN – BEAM	SVM (80.28%)
[42]	16	Product Shapes	ABM X-10 EEG (10 ch.)	Notch Filter (50 Hz), Spline Interpolation	STFT – Spectral bands (Delta, Theta, Alpha, Beta, and Gamma)	DNN (79.76%)
[64]	31	Product images	BioSemi Active Two (64 ch.)	LPF (1 Hz), Notch (50 Hz), Segmentation	FFT – Band Power	Correlational analysis, t-test
[93]	60	Product	Emotiv EPOC+ (14 ch.)	-	Emotiv Pro Software indexes - Engagement, Excitement, Interest, Relaxation etc.	Behavioral Analysis
[94]	38	Product images	HydroCel GSN (128 ch.)	BPF, ICA	FFT – power spectrum (alpha, beta, theta)	ANOVA
[83]	26	Product	ActiChamp, Brain Vision (32 ch.)	Segmentation and Subtraction	Brain Wave Activity (Beta Band)	ANOVA, Correlational analysis, and Friedman tests
[56]	36	Video Clip	LXE5208, Laxtha Inc. (8 ch.)	-	FFT – Power Spectrum (gamma, beta, alpha, theta, and delta)	t-test, Descriptive Statistics, ANOVA
[107]	40	Product	NeuroSky MindWave (1 ch.)	-	Brain Wave Activity (Beta and Gamma)	Chi-square test, t-test.
[86]	20	Video Ads	NeuroSky MindWave	-	Brain Wave Activity (Theta and Beta)	Chi-square test, t-test.

			(1 ch.)			
[84]	40	-	EEG system (2 leads)	-	Frontal Brain Asymmetry (alpha-band)	t-test, ANOVA
[18]	23	Affective Images	Nation 7128C (19 ch.)	HPF (2 Hz), LPF (30 Hz), ICA	Squaring – Power values (alpha band)	Descriptive statistics
[85]	19	Product Images	NeuroScan SynAmps2 (64 ch.)	LPF (30Hz)	Mean amplitudes of ERP components – P1, N170, P3	ANOVA
[58]	18	Product Images	Neurotech, Taganrog (24 ch.)	BPF (0.1-100 Hz)	Power – gamma oscillations, ERP (N400)	Wilcoxon signed-rank test, Mann-Whitney test, Spearman correlation
[108]	18	Product	NeurOne (64 ch.)	BPF (1-40 Hz), Regression	Band Power – Alpha	t-test, ANOVA
[40]	33	Video Ads	StartStim 8 (8 ch.)	ICA	FBP (Delta, Theta, Alpha, Beta, and Gamma), Hemispheric Asymmetry (alpha-band asymmetry)	SVM, LOG, k-NN, MNR
[111]	24	Images of product name & brand name	Neuroscan Synamp2 (64 ch.)	BPF (0.05-100 Hz)	ERP (N270) and time-frequency component (theta band)	ANOVA
[112]	19	Video Ads	g.Nautilus (8 ch.)	Notch Filter (50 Hz), BPF (13-25Hz), Smoothing	Frontal Asymmetry (alpha, beta & gamma)	Mann–Whitney U test, Chi-square test, LOG (61.2%)
[19]	17	Video Ads	KT4800, Contect Medicine (21 ch.)	BPF (1-35 Hz), ICA, Smoothing, Segmentation	PSD, GFP	t-test
[13]	25	Product Images	Emotiv EPOC+ (14 ch.)	ICA, S-Golay, Averaging, BPF	Welch's method – PSD (gamma, beta, alpha, theta)	SVM, RF, k-NN
[23]	70	Product	NeuroSky MindWave	-	Power bands (alpha, beta, delta, theta, gamma)	Paired-samples t-test, Bayesian statistics

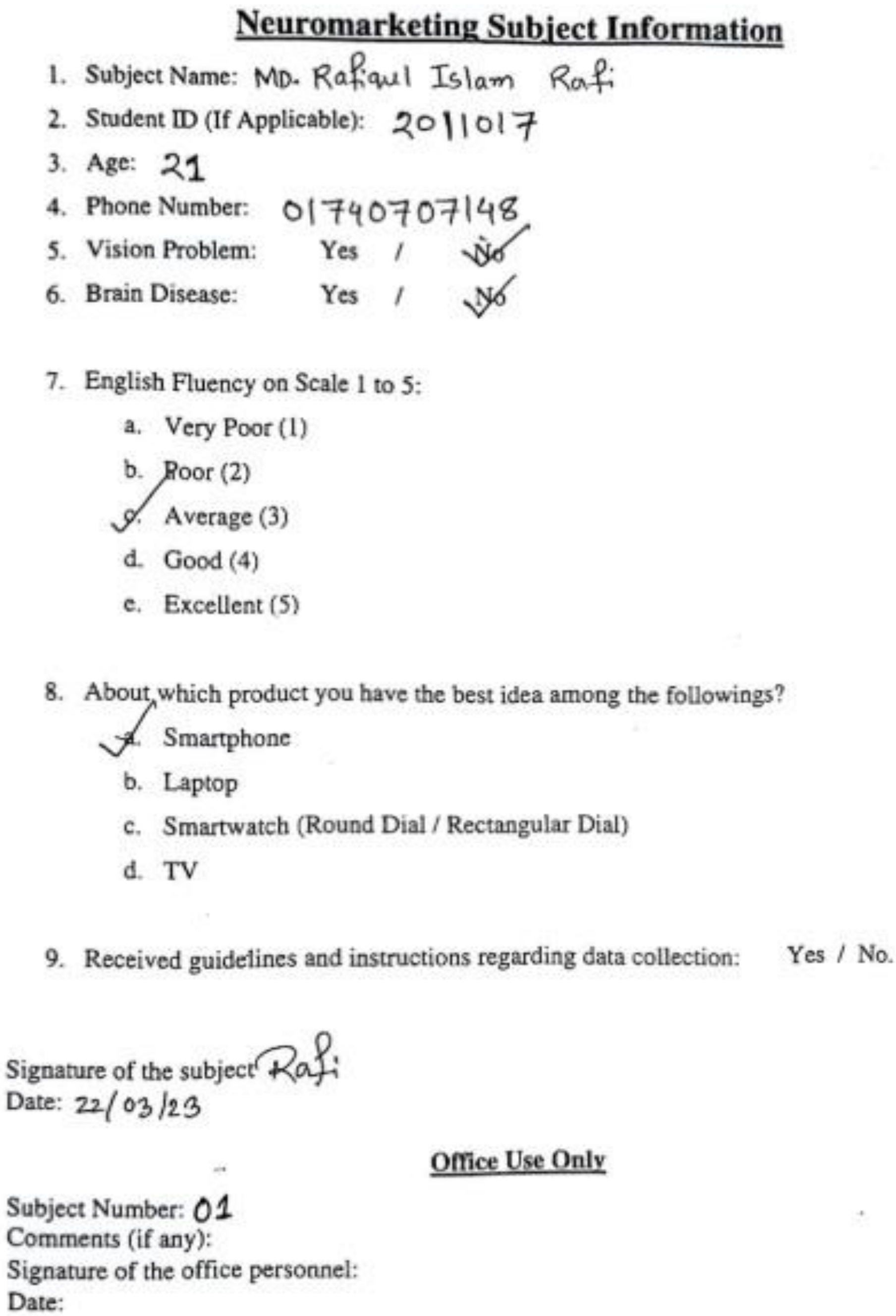
			(1 ch.)			
[87]	40	Product	Emotiv EPOC+ (14 ch.)	HPF (45 Hz), LPF (2 Hz), ICA	FFT - power spectra (beta and delta)	Mann–Whitney test
[46]	30	Product images	Neuron-Spectrum 3 (21 ch.)	-	Brain Wave Activity (alpha and beta)	sLORETA
[88]	40	Product images	Net Station EEG analyzer (9 ch.)	LPF (40 Hz), HPF, Segmentation, Averaging	ERP amplitude (P2), LPP	ANOVA
[113]	11	Product images	Emotiv EPOC+ (14 ch.)	BPF (0.16-32 Hz), Notch (50 Hz)	Autoregressive (AR) and Piecewise Constant Modeling (PCM)	ANN (72%)
[43]	25	Product images	Emotiv EPOC+ (14 ch.)	ICA, S-GOLAY, Averaging, BPF	PSD, DWT, Statistical features, Asymmetry indexes	SVM, RF, KNN, DNN
[65]	36	Video Ads	BEmicro (10 ch.)	Notch (50 Hz), BPF (2-30 Hz), ICA	GFP, FAA	ANOVA
[67]	19	Product images	Neuroscan Synamp2 (64 ch.)	Re-referencing, filtering, artifact removal	ERP – P2, LPP	ANOVA
[66]	70	Video Ads	FlexComp, T9305Z (2 ch.)	BPF (8-12 Hz)	Alpha power, AW index	ANOVA
[95]	17	Product images	Neuroscan Synamp2 (6 ch.)	Segmentation, LPF (30 Hz)	ERP – N200	ANOVA
[109]	30	Music	B-Alert X10 (9 ch.)	Notch (50 Hz), HPF (0.1 Hz)	FFT – PSD (alpha, beta, delta, theta, gamma)	Shapiro-Wilk test
[96]	35	Product images	ActiChamp, Brain Vision (32 ch.)	ICA	Brain wave activity	ANOVA

[89]	01	Product	Emotiv EPOC+ (14 ch.)	BPF (8-13 Hz)	Welch's Method – Power Spectra (alpha), Alpha Asymmetry	-
[20]	72	Affective Images	BEmicro, EBneuro (21 ch.)	Notch filter (50 Hz), BPF, ICA	GFP, AW Index	t-test
[15]	40	Product images	Enobio (20 ch.)	-	Brain wave activity – alpha wave	Friedman's statistical methods and LOG
[110]	54	Video Ads	LXE5208, Laxtha Inc. (8 ch.)	-	FFT – beta power	Multivariate analysis of covariance (MANCOVA)
[52]	20	Product images	g.USBamp (16 ch.)	HPF, Notch, Artifact subspace removal (ASR), ICA	DWT, PSD – Statistical Features	RF (71.51 ± 5.1%)
[53]	5	Product images	Enobio (8 ch.)	BPF, Wavelet ICA	-	k-NN, SVM, t-test
[48]	45	Product images	Neurowerk EEG (21 ch.)	BPF, ICA	Welch's Method – PSD	t-test, ANOVA, k-NN (92.4%)
[49]	26	Product images	NVX36 (28 ch.)	HPF (1 Hz), LPF (40Hz), ICA	FFT – band power (alpha, beta), FAA, EI	Multiple Regression Analysis (MRA)
[47]	15	Product images	BrainAmp Amplifier (32 ch.)	Re-referencing, BPF (0.5-40 Hz), ICA	PSD, AW index, EI, CI, Valence, DE, and Hjorth parameters	SVM, k-NN (94.22%), t-test
[97]	40	Video Ads	NVX-52 (38 ch.)	BPF (0.1–30 Hz), Notch (50 Hz, 100 Hz), ASR, ICA	PSD, AW index, MI	ANOVA
[68]	38	Product images	Neuroscan Synamp2 (64 ch.)	BPF (1-30 Hz), ICA	ERP – P300	ANOVA
[98]	71	Product images	Neuroscan Synamp2	Regression, LPF (30 Hz)	ERP – N300, LPP	MRA

			(64 ch.)			
[50]	25	Product images	Emotiv EPOC+ (14 ch.)	BPF(4-45 Hz), S-Golay	PSD, DWT – Statistical Features	RF (100%), DNN (99%)
[51]	25	Product images	Emotiv EPOC+ (14 ch.)	S-Golay, Notch Filter	LSTM, PSD, DWT	Ensemble learning (96.89%)
[99]	20	Product images	g.USBamp (16 ch.)	HPF, Notch, Artifact subspace removal (ASR), ICA	DWT, PSD – Statistical Features	RF (96.47% for female and 95.32% for male)
[101]	60	Product images	Emotiv EPOC+ (14 ch.)	-	Mean spectral power, FAA	Shapiro–Wilk and Kolgomorov–Smirnov tests, ANOVA
[69]	32	Product images in VR shopping	LiveAmp - Brain Products (32 ch.)	Notch (50 Hz), LPF (100 Hz), HPF (0.1 Hz)	Event-related desynchronization and synchronization (ERD/ERS index)	Wilcoxon Signed-Ranks test
[70]	25	Product	OpenBCI (4 ch.)	BPF	PSD, FAA	Behavioral Analysis
[39]	183	Product images	StartStim 8 (8 ch.)	BPF (0.5-100 Hz), Mean Subtraction	PSD	DL (75.09%)
[44]	33	Video Ads	StartStim 8 (8 ch.)	HPF (0.1 Hz), Notch (50 Hz), ICA	FBP (Delta, Theta, Alpha, Beta, and Gamma), Hemispheric Asymmetry (alpha-band asymmetry)	Sparse Representation Classification (SRC) - 64.70%, ANOVA
[100]	16	Product images and Price	Brain ActiChamp (6 ch.)	BPF, ICA	ERP (P2, LPP)	Statistical Analysis
[104]	22	Images	eegoAmplifier, ANT Neuro (15 ch.)	LPF, HPF, Segmentation	ERP (N400, LPP)	ANOVA

[103]	20	Images	BioSemi Active Two (64 ch.)	BPF (0.1-30 Hz)	ERP (LPP)	ANOVA
[102]	40	Video Ads	NVX 52 (40 ch.)	BPF (2-48Hz), ICA	PSD (Theta, Alpha, Beta, and Gamma),, AW Index (AWI), Willingness to Pay Index (WPI), Beta over Alpha plus Theta Ratio (BATR), Beta over Alpha Ratio index (BAR)	Shapiro–Wilk test, Pearson's linear correlation, Regression Models
[54]	25	Product Images	Emotiv Epoc (14 ch.)	-	Multi-taper spectral analysis – PSD	Bidirectional-LSTM deep learning (96.83%)

Appendix B



Neuromarketing Subject Information

1. Subject Name: MD. Rafiqul Islam Rafi
2. Student ID (If Applicable): 2011017
3. Age: 21
4. Phone Number: 01740707148
5. Vision Problem: Yes / No
6. Brain Disease: Yes / No
7. English Fluency on Scale 1 to 5:
   a. Very Poor (1)
   b. Poor (2)
   c. Average (3)
   d. Good (4)
   e. Excellent (5)
8. About which product you have the best idea among the followings?
   a. Smartphone
   b. Laptop
   c. Smartwatch (Round Dial / Rectangular Dial)
   d. TV
9. Received guidelines and instructions regarding data collection: Yes / No.

Signature of the subject: Rafi
Date: 22/03/23

Office Use Only

Subject Number: 01
Comments (if any):
Signature of the office personnel:
Date:

Fig. B.1 Sample subject consent form.

Fig. B.2 Sample smartphone stimuli in FL.

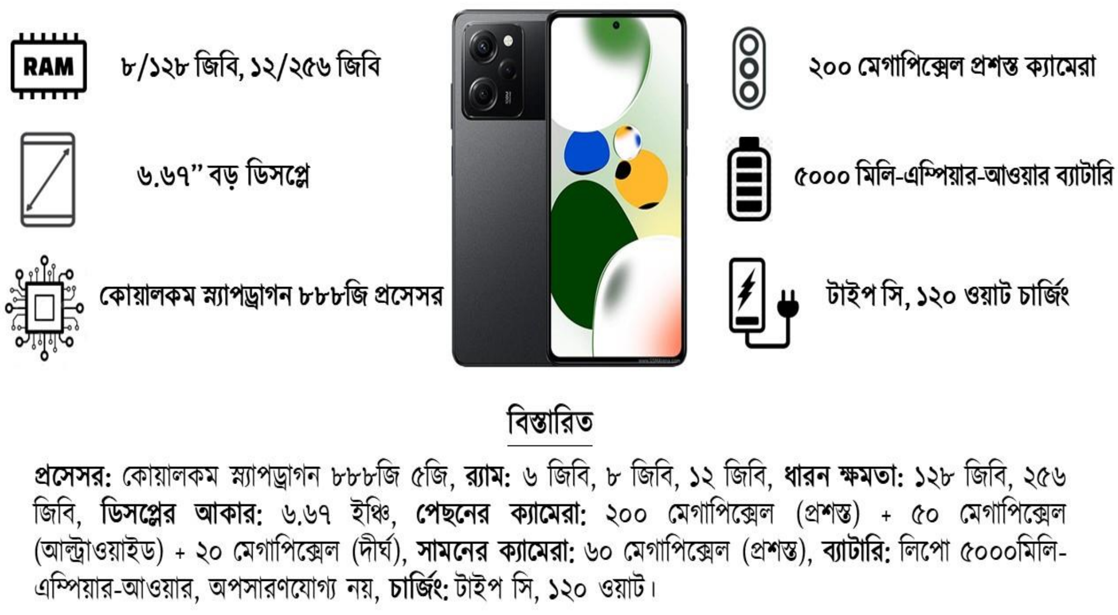


Fig. B.3 Sample smartphone stimuli in NL.

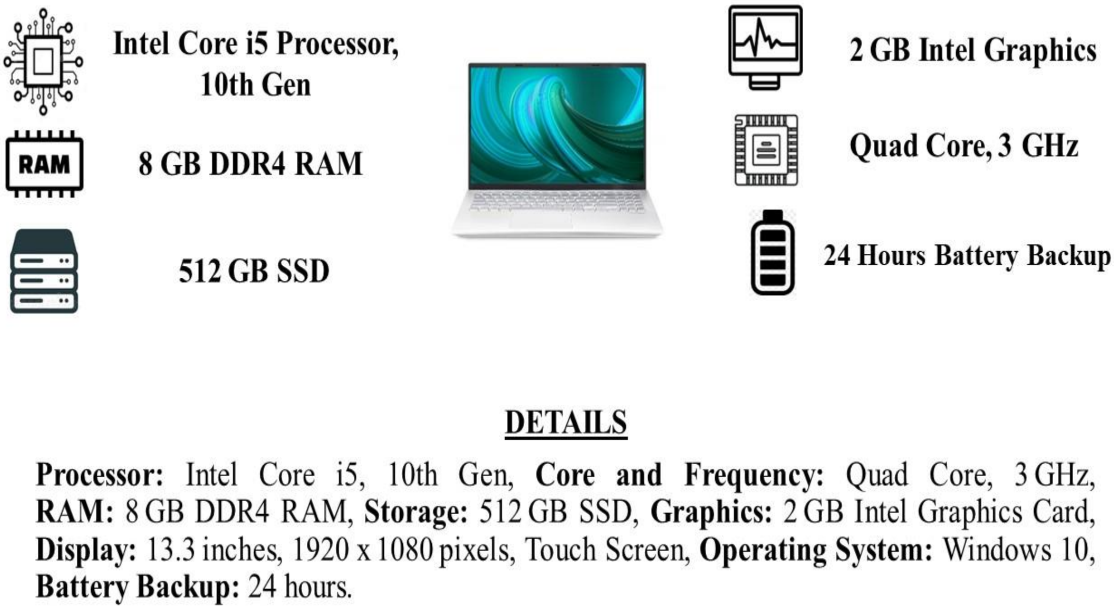


Fig. B.4 Sample laptop stimuli in FL.

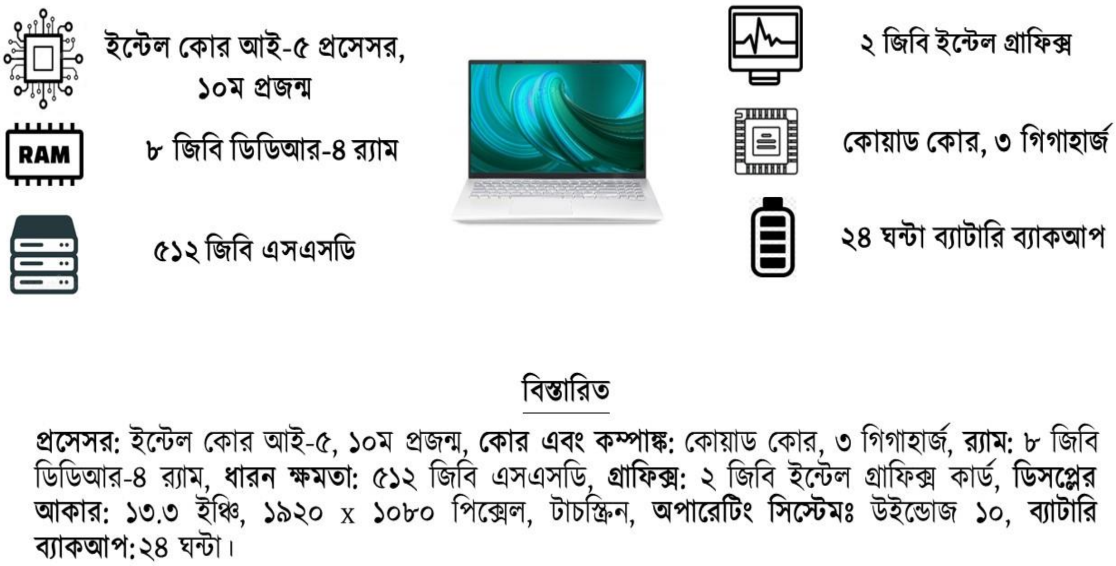


Fig. B.5 Sample laptop stimuli in NL.

Appendix C

- Similarity Report:

Msc thesis report check

ORIGINALITY REPORT

15%	10%	10%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.frontiersin.org Internet Source	1%
2	da Silva, Anderson Santos, Paul Smith, Andreas Mauthe, and Alberto Schaeffer-Filho. "Resilience support in software-defined networking: A survey", Computer Networks, 2015. Publication	1%
3	www.ncbi.nlm.nih.gov Internet Source	1%
4	Mashael ALdayel, Mourad Ykhlef, Abeer Al-Nafjan. "Deep Learning for EEG-Based Preference Classification in Neuromarketing", Applied Sciences, 2020 Publication	1%
5	Submitted to Chittagong University of Engineering and Technology Student Paper	1%
6	Ferdousi Sabera Rawnaque, Khandoker Mahmudur Rahman, Syed Ferhat Anwar, Ravi Vaidyanathan et al. "Technological	1%

Plagiarism OK