Biomedical Signal Processing and Control 87 (2024) 105469

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Neuromarketing and decision-making: Classification of consumer preferences based on changes analysis in the EEG signal of brain regions

Mounir Ouzir [a,*], Houda Chakir Lamrani [b], Rachel L. Bradley [c], Ismail El Moudden [c]

[a] High Institute of Nursing Professions and Health Techniques, (ISPITS de Beni Mellal) Beni Mellal, Morocco
[b] EMAA Business School, Av. Moulay Hassan I, Agadir 80020, Morocco
[c] Eastern Virginia Medical School, Norfolk, VA, USA

ARTICLE INFO

Keywords:
Neuromarketing
Decision-making
Consumer preference
EEG
Classification

ABSTRACT

Neuromarketing involves the study of brain responses that focuses on understanding how consumers' brains respond to products and services, and how these responses influence their choice. Evidence suggests that electroencephalography (EEG) can provide valuable insight into consumer preferences and decision-making processes. This study aims is to assess the relative importance of right/left brain regions (including hemispheres, frontal, temporal, parietal, and occipital lobes) in the consumer choice towards E-commerce products. Also, this study aims to distinguishes the EEG characteristics of consumers' preference using a classification system.

Using a publicly available EEG neuromarketing dataset, the change in EEG signals has been evaluated by a mixed model for repeated measures for all brain regions. Four classification algorithms (k-Nearest Neighbor, Random Forest, Neural Network, and Gradient Boosting) were used to distinguish like and dislike preferences.

Greater EEG activity in the right hemisphere, right parietal, right occipital, and left occipital was related to like responses. Except for both sides of the temporal lobe, all the subdivisions of the brain considered showed a significant decrease of activity at 4000 ms for like-related responses. However, no significant change in the activity was related to the dislike response. The highest AUC of the four classifiers used was as follows: 76.61% for the right parietal lobe with Neural Network, 75.33% for the left parietal lobe with Gradient Boosting, 73.55% for the right frontal lobe with k-Nearest Neighbor and 72.62% for the right frontal lobe with Random Forest. Considering the significant difference between like and dislike responses at 4000 ms, Neural Network showed the best performance followed by Gradient Boosting.

Our framework suggests that the formation of preferences (like and dislike) requires different patterns of brain activity and that Neural Network and Gradient Boosting are valuable tools for distinguishing consumer preference.

1. Introduction

Neuromarketing is a cutting-edge research field dedicated to uncovering the intricacies of decision-making, which involves the study of brain responses that explain the consumer's behavior towards products and services [1,2]. It's widely accepted that neuromarketing has the potential to create and develop more useful and pleasant products, refine pricing strategies, and implement more effective social campaigns [3]. Electroencephalography (EEG) is a widely used technique in neuromarketing, allowing the capture of electrical brain surface activity by attaching electrodes [4]. This can be used to discern the customer's choice as well as the level of attention, engagement, and memorization [5,6] Many application domains were identified including, advertisement assessment, product choice/preference, product, packaging and brand perception, and politics [7].

The use of pattern classification based on statistical techniques for predicting customer preferences has become a common standard procedure in neuromarketing. In recent years, most neuromarketing studies have been conducted based on Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Artificial Neural Network (ANN), Naïve Bayes, k-Nearest Neighbor (kNN) and Hidden Markov Model (HMM) [6]. Specifically, some algorithms have been used for the classification of EEG signals in terms of likes and dislikes with varying results. For example, KNN was employed to measure consumer preferences for

* Corresponding author.
E-mail address: m.ouzir@gmail.com (M. Ouzir).

https://doi.org/10.1016/j.bspc.2023.105469
Received 9 January 2023; Received in revised form 2 August 2023; Accepted 12 September 2023
Available online 21 September 2023

aesthetics shown as virtual three-dimensional objects [8], for sport shoes products [9], for music [10–12], for vehicles [13] and shoes and food images [14]. The accuracies varied from 74.6% to 97.99%. ANN was used to evaluate consumer preferences for shopping products [15,16], television commercials from the final of the Super Bowl 2015 viewed on an official YouTube channel [17] and TV short video advertisements [18] and was able to predict the liked and disliked products with an accuracy ranged from 60.10% to 82.90%. Only one study performed a classification by Random Forest (RF), and achieved 68.41% accuracy in predicting the liked - and disliked shopping products [16]. By contrast, no study used Gradient Boosting algorithm, which is known to make the classification more accurate and precise by reducing the overfitting problems of the traditional decision tree, for binary classifications.

Thus, the aim of this study is to build a classification system that distinguishes the EEG characteristics of consumers' preference between like and dislike based on kNN, RF, ANN, and Gradient Boosting classifiers. Before this classification step, this work will assess the relative importance of brain regions (including, the right hemisphere, left hemisphere, right frontal lobe, left frontal lobe, right temporal lobe, left temporal lobe, right parietal lobe, left parietal lobe, right occipital lobe, and left occipital lobe) in coding the consumer choice towards E-commerce products. The change in EEG signal will be evaluated using a mixed model for repeated measures between subjects (like and dislike groups) and within subjects (from baseline [second = 0] to 4000 ms) for all brain regions considered.

2. Related work

2.1. Cortical activity and decision-making process

Decision-making requires evaluating relevant sensory stimuli to make the appropriate choice. It involves several cognitive processes, such as attention, learning, working memory, and the motor system, and has been studied extensively in the context of marketing [19,20]. Neuromarketing tries to investigate and analyze certain brain processes produced by stimuli from marketing elements to explain the way of acting and the decision-making process of individuals. Prior studies in the field of neuroscience observed that the majority of our purchase decisions are made unconsciously in less than 4 s [5]. Different cortical areas have been associated with different aspects of the decision-making process in neuromarketing [6,21]. Decision-making mainly takes part in the prefrontal region, and movement-related decisions are made in the end part of the frontal lobe [22]. The parietal cortex was associated with choice information processing related to working memory, attention, and planning as well as perceptual and categorical decision-making [23,24]. It has been found that the parietal cortex encodes the probability of a possible occurrence of a reward coupled with a visual stimulus in economic decision-making [25,26].

The occipital lobe is involved in the processing of visuospatial information that allows the brain to recognize objects and patterns quickly without a significant conscious effort. Evidence suggests that the brain required more visual information processing when performing decision-making tasks [27]. Recently, the prefrontal, parietal and occipital lobes are found to play key roles in spatial decision-making [28].

According to Larsen & O'Doherty, [29] chosen value signals, which is a consequence of the decision process, appear to initially emerge predominantly in a posterior location in the brain as well as in the posterior temporal lobe cortex (early visual processing and formation of the initial decision) within 200 ms after the onset of the decision trial. The signal then shifted to a predominantly anterior locus 850 ms following the trial onset, localized to the ventromedial prefrontal cortex and lateral prefrontal cortex. Later, at approximately 1050 ms, the signal emerged from the dorsomedial prefrontal cortex, suggesting its contribution at a later stage during the process of implementation of action selection. Moreover, liking is considered one of the most important psychological components of the reward system associated with pleasure/displeasure encoding [30].

Previous studies using EEG showed that the consumer preference for one product over another was expressed by an increase in alpha, beta, and gamma activities in frontal areas [31,32]. Also, the increase in the activity of the parietal and superior prefrontal regions was related to the maintenance of highly processed representations of complex stimuli, such as TV advertisements [33].

2.2. Like/dislike product preference prediction

Table 1 shows different studies using like/dislike product preference prediction based on EEG signals. EEG was used in studies involving the evaluation of widely different products (food, clothes, vehicles, mobile, and music…) in order to associate EEG features with consumer behavioral patterns and/or predictions.

To recognize product consumer preferences for online 10 shopping products including drinks, food, cars, textiles and services, Ullah et al [15] used 5 classifiers to analyze EEG signals; SVM, Logistic Regression (LR), Decision Tree (C.5), RF, and ANN. ANN achieved a higher accuracy of 81.23% followed by SVM (80.38%). Yadava et al [16] proposed a framework for predicting consumers' choice of different shopping products based on HMM, RF, SVM, and ANN classifiers. The best accuracy was obtained by HMM classifier (70.33%). According to Zeng et al. [9] the recognition accuracy of sport shoes products reaches 94.22% based on the Differential Entropy features and kNN classifier. Also, the features of the frontal and occipital brain regions obtained a higher prediction accuracy compared to the others brain regions.

Guixeres et al [17] attempted to investigate the effectiveness of neural networks and neuroscience-based metrics to predict ad success in a digital context (YouTube). The like group showed significant differences compared to the dislike group with higher values in the delta, theta, beta, and gamma bands. The results showed that the ANN was able to precisely predict when an ad is going to be liked [17].

Yilmaz et al [34] used LR to identify the most discriminative frequencies for user preference for consumer products. The authors also explored the timings of like decisions and whether they were different between male and female subjects. in the low-frequency band (4–19 Hz), a frontal channel on the left (F7-A1) and a temporal channel on the right (T6-A2) were found to be the most discriminative channels (MDCs). In the high-frequency band (20–40 Hz), MDCs were central (Cz-A1) and occipital on the left (O1-A1) channels. The average number of like decisions for male and female subjects was the same. However, the mean time of like the decision for males was half of that for the female participants.

Using brand advertisements of four vehicles (Toyota, Audi, Proton, and Suzuki), Murugappan et al [13] showed that Toyota brand vehicles are highly preferred by the subjects compared to other brands and that the maximum mean classification rate of 96.62% is achieved using PSD feature and PNN classifier.

Chew et al [8] studied aesthetic preferences through 60 virtual three-dimensional shapes that appear like bracelets with motion and found an accuracy of 80% using alpha-theta and delta rhythms from 3 frontal channels (Fz, F3, and F4) with kNN and 79% with SVM. The left and right parietal lobes (P3 and P4) showed an accuracy of no more than 71 % for both classifiers.

To recognize subjects' preferences while watching music video clips, Koelstra et al [35] used the fast correlation-based filter as feature selection and SVM as a classifier and achieved an average accuracy of 57.9%. Koelstra et al. [35] observed a positive correlation between consumer preference (like/dislike) and the theta band and a negative correlation with the alpha band. Hadjidimitriou & Hadjileontiadis, [10,11] used kNN classifier and Hilbert–Huang spectrum-based feature vectors to predict human preference for music and obtained accuracies of up to 86.52%.

Moreover, Moon et al [12] studied the preference for video corresponding to 6 Korean pop songs using band power (BP) and asymmetry

Table 1
Summary of highlighted works on Like/Dislike products preference prediction using EEG signal.

References	Device/system	Dataset	Participants/Age	Analysis method	Accuracy (%)	Main results
[15]	4 channels Muse 2 Headband	10 different shopping product images with 3 varieties of each	15 healthy people	ANN kNN SVM LR C.5	81.23 75.72 80.38 87.21 67.86	–
[14]	14 channels EMOTIV Epoc + wireless headset	30 image advertisements (shoes and food)	49 subjects with normal or corrected vision. 24 males & 25 females. 19–28 years.	SVM kNN Neural net pattern recognition	74.2 74.6 71.0	Frontal region electrodes yielded the best selective channel performance. The performance measures for the female subjects were comparably higher than those for the males.
[9]	32-channel EEG acquisition device. BrainAmp Amplifier (Brain Product. Gilching. Germany)	25 different designs of sport shoes	15 healthy subjects. 9 men & 6 women. 22–39 years	kNN-DE kNN-PSD kNN-Hjorth kNN-CBA	94.22 88.85 86.17 82.04	The frontal and occipital brain regions obtained a higher prediction accuracy.
[41]	32-channel electrode system (Contact Precision Instruments. Uk)	4 mobile brands (Apple. Samsung. Nokia. and Meizu)	16 healthy Caucasian participants. 9 males & 7 females. 23 ± 3 years	LDA SVM	63.95 63.62	Theta band power increased in the left/right frontal regions with like/dislike preference. Difference between like/dislike decisions in F4 and Ft8 electrodes. The average brain power is higher in like condition than dislike.
[17]	32-channel system (REFA 32. TMSI hardware)	8 online ads of commercial products such as drinks. food. cars. textiles and services	36 randomly healthy volunteers. 15 females & 20 males. 25 ± 5 years.	ANN	82.9	Strong brain activity in terms of PSD in the like group than in the dislike group.
[16]	14 Channels Emotiv EPOC Inc. Headset	14 different shopping product images with 3 varieties of each	40 participants. 25 males & 15 females. 18–38 years	HMM RF SVM ANN	70.33 68.41 62.85 60.10	
[18]	14 Channels Emotiv EPOC Inc. Headset	14 TV commercials	10 subjects. 5 females &5 males. 26—60 years	ANN Ameva C4.5	80 75 69	
[38]	64-channel Bio Semi Active Two system and *ActiView* software	120 brand names in capital white letter in Tahoma font on black background and without any logo	24 subjects. 12 females & 12 males. 23.58 ± 2.39 years	–	–	Liked brands elicited significantly more positive going waveforms than disliked brands over right parietal cortical areas.
[8]	9 channels. Advanced brain monitoring (ABM). the B-alert X10 headset	60 3D bracelet-like stimuli	5 subjects. 3 females & 2 males. 22–26 years.	kNN SVM	80 79	–
[39]	64-channel EEG BioSemi system. Amsterdam	30 pairs of female shoes	40 women. 1953 years.	One-dimensionallinear classifier	80	EEG predict consumer behavior much better thanquestionnaire.
[34]	21 channels; EEG 1200. Nihon Kohden Co. Tokyo. Japan	16 images containing women's shoes in different styles and colors	15 participants. 10 females & 5 males.19–25 years.	LR	NA	4 and 5 Hz were the most discriminative frequencies. In the 4–19 Hz band. F7-A1 and T6-A2 were the most discriminative channels (MDCs). In the 20–40 Hz band. Cz-A1 and O1-A1 were the MDCs.
[13]	14 channel wireless Emotive headset	4 brand advertisements of vehicles including Toyota. Audi. Proton. and Suzuki	12 subjects. 9 males & 3 females. 22–24 years.	PNN kNN	96.62 95.19	
[12]	14 Channels Emotiv EPOC Inc. Headset	6 video segmentscorresponding to 6 Korean pop songs	15 healthy right-handed subjects. 8 males & 7 females. 23—43 years	QDA + FT kNN + FT	97.39 97.99	
[36]	2-channel frontal EEG NeuroSky-MindBand	2 sessions on separate days. For each session. participants choose 4 genres from 10 available genres	12 healthy subjects	CFP-SVM	74.77	Gamma band is essential for EEG-based music preference identification
[10,11]	14 Channels Emotiv EPOC Inc. Headset	60 musical excerpts from 4 common musical genres	9 subjects. 2 females &7 males. 23.22 ± 1.72 years	kNN	86.52	
[37]	96-channel system. BrainAmp. Brainproducts GmbH. Germany	18 commercial video clips	11 healthy undergraduate students. 8 males. 22–25 years	PSD	NA	Asymmetrical increase of frontal theta and alpha activity related to pleasant / unpleasant TV commercials in the left/right hemisphere.
[35]	32 active AgCl electrodes Biosemi ActiveTwo system	20 selected music videos clips	6 participants	FCBF-SVM	57.9	Positive correlation between consumer preference and the theta band. and a negative correlation with the alpha band.

DNN: Deep Neural Networks; kNN: k-Nearest Neighbor; PNN: Probabilistic Neural Network; FCBF: Fast Correlation Based Filter; QDA: Quadratic Discriminant Analysis; PSD: Power Spectral Density; FT: Fourier Transform; CFP: Common Frequency Pattern; CBA: choice-based asymmetry.

scores as features. the quadratic-discriminant-analysis-based model using 56 BP features achieves a classification accuracy of 97.39%. Using the Common Frequency Pattern method for feature extraction and SVM for classification, Pan et al [36] reported an accuracy of 74.77% for music preferences. The preference for music seems to be highly discriminated in the frontal cortex of the brain and dependent on gamma bands [36].

Some studies investigated the consumer's preference during the observation of TV commercials. Using Power Spectral Density maps, Vecchiato et al [37] showed that the activity in the left/right frontal hemisphere was related to the observation of ads that have been judged pleasant/unpleasant Also, an asymmetrical increase of frontal theta and alpha activities was related to observing pleasant/unpleasant TV commercials in the left/right hemisphere. Likewise, Bosshard et al [38] examined the consumer attitude toward established120 brand names. They found that liked brands are implicitly associated with increased motivational aspects compared to disliked brands, with a dominant effect over right parietal cortical areas. this result supports the fact that what generates preference for a brand is the influence on the motivational cognitive processes of a subject, which are more correlated with the right hemisphere. Soria Morillo et al [18] used the Ameva algorithm, ANN, and C4.5 to predict preference for TV short video advertisements and obtain an accuracy of approximately 80%, 75%, and 69% for ANN, Ameva, and C4.5 respectively.

It is interesting to mention that the study by Baldo et al. [39] showed that a questionnaire (self-assessment) cannot accurately predict consumer behavior, while with ECG, the prediction accuracy reached 80%. Additionally, the simulation based on the sales data showed that the ECG-based prediction would increase profits by 36.4% compared to a profit of 12.1% with the self-assessment.

Since consumers are not always aware of what motivates them to make a certain choice, as much of their mental processing is unconscious, most of them tend to overestimate their evaluation of a particular good, service, or outcome, which can lead to misleading estimates of relative value [40]. In fact, current neuromarketing research still lacks a deep and thorough understanding of how the brain operates and affects specialized human behavior and decision-making processes [4] and how physiological measures would increase the prediction power of consumers' choices.

Golnar-Nik et al [41] investigated how 4 mobile brands (Apple, Samsung, Nokia, and Meizu) affect customers' shopping preferences and showed that the difference between like and dislike decisions was observed mostly in the right frontal electrodes (F4 and Ft8) and that the average brain power when subject likes an object is higher than that of the dislike condition, with a decrease in reaction time of decision-making when a color background is added to the advertisement. Similar to other studies, the theta band activity in the right hemisphere was higher for dislike compared with the like preference, and the like preference increased the EEG power of the theta band in the left frontal region while dislike preference increased the theta band power in the right frontal region. Also, when considering the sensitivity and specificity values, the SVM classifier showed better performance for separating consumer preferences from each other [16,41].

More recently, Sourov et al [14] analyzed the EEG data of 24 males and 25 females with normal or corrected vision watching 30 image advertisements of food and shoes, using SVM, KNN, and neural network pattern recognition and obtain accuracies of 74.2%, 74.6% and 71.0% respectively. For the female subjects, the performance of the 3 classifiers was higher than for male subjects.

According to the analysis of the studies in Table 1, we observed the absence of works related to EEG signal differences within-subject (between like subject and/or between dislike subject) and between-subject (like vs dislike) without preprocessing (denoised signals), before using a classification step to predict like and dislike preference of the e-commerce products.

3. Materials & method

3.1. Dataset

The dataset used in this study was made available by Yadava et al [16], which includes EEG recordings from 25 participants, aged between 18 and 25 years. EEG signals were recorded with a 14-channel Emotiv Epoc+ (wireless device for neuro-signal data acquisition complying with the international standard 10–20 system) while participants watched different e-commerce product images for 4 s on a computer screen. The stimuli consisted of 14 different product categories (shirts, shoes, ties, school bag, muffler, belt, bracelet, gloves, sun glass, sweater, socks, wall clock, pen, and wrist watch), each containing three different images, resulting in 42 different products 42 (=14 × 3). After each image, participants had to indicate their liking or disliking of the presented product. A total of 1050 (=42 × 25) EEG data were therefore generated for all users. The EEG features were collected from 14 channels placed at AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 locations. After each image was presented, the preferred choice of the user was collected. The recording sampling rate was 128 Hz. The authors used a Savitzky-Golay (S-Golay) filter to remove unwanted noise from the acquired EEG, the Discrete Wavelet Transform to extract valuable features from the recorded signals, and the Daubechies 4 (DB4) wavelet decomposition technique to obtain the four brainwaves frequency; delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–22 Hz) and Gamma (32–100 Hz).

Table 2 displays the frequency of like and dislike-related responses in the dataset. As can be seen, more than half (55.62%) of the responses were "like" and 43.90% were "dislike". Missing responses represent 0.48%.

3.2. Brain subdivision

In the international standard 10–20 system, each letter of 14 electrodes identifies the lobe locations, while the corresponding number identifies the hemisphere location. The letters F, P, T, and O, stand for frontal, parietal, temporal, and occipital lobes, respectively. While AF3 and AF4 refer to anterior frontal locations, FC5 and FC6 refer to frontocentral locations. Odd numbers refer to the EEG channels on the right hemisphere, while even numbers refer to those located towards the left hemisphere. In this study, we considered different subdivisions of the brain as shown in Table 3.

3.3. Classification

This study aims to distinguish the EEG characteristics of consumers' preference between like and dislike based on k-Nearest Neighbor, Random Forest, Neural Network, and Gradient Boosting methods from the Yadava et al [16] dataset. The 14 EEG signals from selected segments were utilized as input features for the classifiers without any prior feature selection or extraction process. Each EEG signal captured is a time-series data, representing the electrical activity from different

Table 2
Choice preference outcome description.

Outcome	Frequency (%)
Like	584 (55.62%)
Dislike	461 (43.90%)
Missing	5 (0.48%)

Table 3
Brain regions subdivision considered in this study.

Brain regions	Electrodes
Right hemisphere	O2, P8, T8, FC6, F4, F8, AF4
Left hemisphere	AF3, F7, F3, FC5, T7, P7, O1
Right Frontal	FC6, F4, F8, AF4
Left Frontal	AF3, F7, F3, FC5
Right Temporal	T8
Left Temporal	T7
Right Parietal	P8
Left Parietal	P7
Right Occipital	O2
Left Occipital	O1

region of the brain. These signals inherently encompass the necessary temporal information which allows the classifiers to detect patterns associated to the consumer preferences.

Gradient Boosting is an ensemble model that utilizes M regression trees trained in sequence to predict the errors of previous trees. It can be thought of as a form of gradient descent in a functional space, with each iteration of Gradient Boosting involving the growth of a tree to estimate the gradient of the objective function [42].

The k-Nearest Neighbor (kNN) algorithm is a simple and effective nonparametric classification method with competitive performance, and strong statistical properties [43]. kNN has been used in pattern recognition based on their nearest neighbors. kNN aims to find the shortest distance between the data to be evaluated and the closest k-Nearest (neighbor) in the training data [44]. The 'k' refers to the count of neighbors of the new data point. For better accuracy, it is imperative to choose the accurate value of k, and this process is called parameter tuning. The standard value that K often assumes is '5′ and the 5 neighbors closest to the new data point are considered [45].

Neural networks are a preferred model for many neuromarketing tasks involving data classification, prediction, and clustering [6,46]. Neural network is a well-known tool for solving classification tasks due to its outstanding performance as a classifier [47]. Neural network is a collection of artificial neurons system composed of interconnected neurons that can classify data into multiple categories by creating non-linear decision boundaries. It is a type of supervised learning which consist of an input layer, an output layer, and one or more hidden layers. Neurons combine their weighted inputs linearly and generate an output based on their non-linear activation function.

Random Forest (RF) consists of multiple small decision trees trained on a random subset of data, each tree acts as an expert, making split decisions based on the input it has been given [48]. RF generates many decision trees independently and combines their output. Decision trees consist of internal and leaf nodes. The internal nodes use selected features to make a decision and divide the dataset into two subsets with similar responses. The features in an internal node are selected by the Gini impurity criterion. The feature that has the highest decrease in impurity is selected for the internal node.

3.4. Cross-validation and performance measures

The proposed approach was evaluated using 10-fold cross-validation, in which the subjects were randomly split into a training and a testing subset. The 10-fold cross-validation was performed strictly within the training data, ensuring that no information from the testing data was used during the training process, thus avoiding any data leakage. The test data were kept entirely separate and used only for the final evaluation of the model performance.

To assess the performance of the proposed approach, five measures are considered, including AUC, Accuracy, F1-score, Precision, and Recall or Sensitivity. Precision, recall, and F1-score measures are important criteria while evaluating the performance of classifiers. The higher the precision is, the lower the number of false positive errors committed by the classifier is. Classifiers with large recall have very few positive examples misclassified. F1-score represents a harmonic mean between recall and precision. A high value of the F1-score ensures that both precision and recall are reasonably high. The AUC value is considered a better performance measure when comparing classifiers. The flowchart of the proposed framework is shown in Fig. 1.

3.5. Sample size calculation and power analysis

The determination of sample size and conducting a power analysis is a critical step to ensure the robustness and validity of our results and to avoid a Type II error. Group sample sizes of 584 (like) and 461 (dislike) achieve 89% power to detect an EEG signal difference of 10 in a design with 8 repeated measurements having a covariance structure of autoregressive (AR) order 1 AR (1) when the standard deviation is 120, the correlation between observations on the same subject is 0.2, and the alpha level is 0.05.

3.6. Statistical analysis plan

Descriptive statistics, including mean (with 95% confidence interval), median (with interquartile range), minimum and maximum values, standard deviation (SD), and frequency were utilized to summarize the data [49]. The Shapiro-Wilk W Test was utilized to test the normality of continuous variables [50].

A mixed model for repeated measurements (MMRM) was performed to access the change in the EEG signal received from different brain regions on the whole dataset. The MMRM serves as a powerful analytical tool for the assessment of the EEG signals' changes over time and across different brain regions in our investigation. This statistical method has the advantage of accommodating both fixed and random effects, allowing us to capture the intricacies of our longitudinal data [51]. Fixed effects can be interpreted as the influences of interest that are consistent across individuals, such as experimental conditions or inherent participant characteristics. Conversely, random effects capture the individual differences in baseline levels and changes over time, which can account for the unobserved variability among subjects that might affect the EEG signals [52].

Importantly, MMRM recognizes the inherent correlation among repeated measurements from the same individual, providing a robust framework to model these within-subject dependencies. This feature is particularly crucial when dealing with longitudinal data as we have in our study. Moreover, MMRM is well-equipped to handle potential issues related to missing data and is conducive to modeling time-varying covariates - both aspects of high relevance in the context of EEG analyses [53].

Forest plots were created to illustrate the difference in EEG Least Squares Mean between the baseline (0 ms) and various time points considered in the study (250 ms, 500 ms, 750 ms, 1000 ms, 2000 s, 3000 ms, and 4000 ms). These graphical representations offer an intuitive way to visualize the differences in EEG signals over time, and enabling the identification of specific trends and patterns that might be correlated with consumer preferences. The inclusion of forest plots adds visual clarity and supplements the statistical findings of the MMRM analysis.

All statistical analyses mentioned above were conducted using SAS version 9.4 (SAS Institute, Cary, NC). The access to SAS software was provided by the EVMS Healthcare Analytics and Delivery Science Institute (HADSI) at Eastern Virginia Medical School, allowing rigorous monitoring and validation of applied statistical techniques.

4. Results

Table 4 summarizes the results of the Shapiro-Wilk normality test, as well as the median, mean, and standard deviation, and minimum and maximum values for all brain regions in the like and dislike groups. The mean values for each brain region differed between the studied groups.

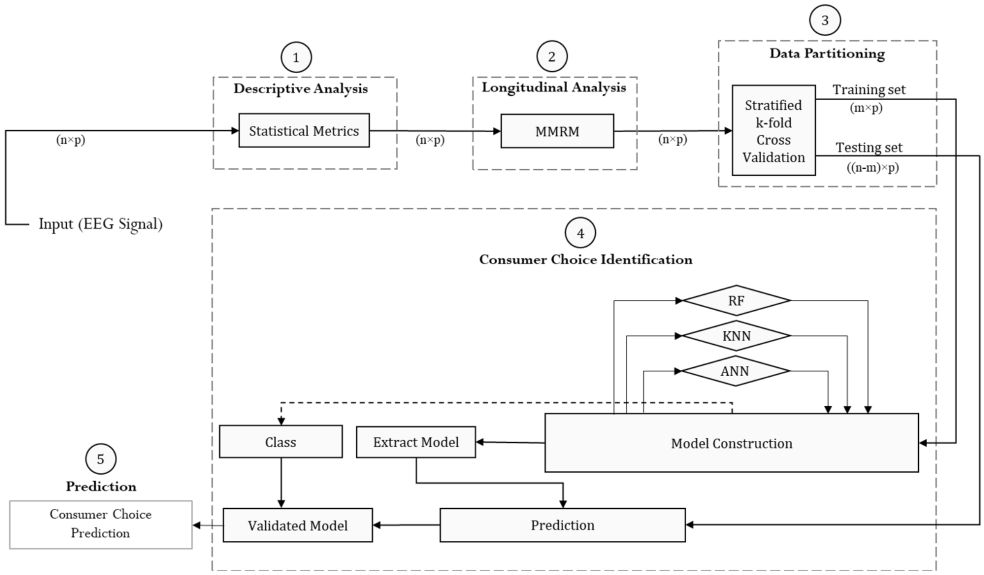

Fig. 1. Flowchart of the proposed method. n: number of images, p: number of embedded features, m: training size (m < n), n-m: testing size.

Table 4
Descriptive analysis and normality test of the EGG signal received from brain regions for Like and Dislike groups.

Brain regions	Normality*	Median	95% CI ǂ	Mean ± SD**	95% CI ǂǂ	Min - Max
Like group						
Left Hemisphere	<0.0001	4226	4225–4227	4222 ± 79.12	4219–4225	2626–4814
Right Hemisphere	<0.0001	4225	4224–4227	4221 ± 85.38	4218–4225	2310–4684
Right Frontal Lobe	<0.0001	4224	4222–4226	4220 ± 104.4	4215–4224	2204–5022
Left Frontal Lobe	<0.0001	4226	4224–4228	4222 ± 94.54	4218–4226	2701–4915
Right Temporal lobe	<0.0001	4228	4228–4228	4227 ± 4.941	4227–4228	4206–4240
Left Temporal lobe	<0.0001	4225	4225–4225	4224 ± 7.021	4224–4224	4088–4250
Right Parietal Lobe	<0.0001	4227	4225–4229	4222 ± 133.6	4217–4228	1171–4966
Left Parietal Lobe	<0.0001	4225	4224–4226	4222 ± 114.1	4217–4226	1858–5581
Right Occipital lobe	<0.0001	4227	4226–4229	4222 ± 104.7	4218–4227	1543–5187
Left Occipital lobe	<0.0001	4223	4222–4225	4218 ± 108.3	4214–4223	1492–4986
Dislike group						
Left Hemisphere	<0.0001	4226	4225–4226	4225 ± 97.11	4222–4229	1445–5580
Right Hemisphere	<0.0001	4225	4224–4227	4225 ± 95.5	4222–4229	1378–5665
Right Frontal Lobe	<0.0001	4224	4222–4225	4224 ± 110.9	4220–4228	950.5–5703
Left Frontal Lobe	<0.0001	4226	4224–4228	4226 ± 118	4222–4231	1112–5768
Right Temporal lobe	<0.0001	4228	4228–4228	4227 ± 5.052	4227–4227	4210–4247
Left Temporal lobe	<0.0001	4225	4225–4225	4224 ± 6.611	4224–4224	4058–4278
Right Parietal Lobe	<0.0001	4228	4225–4229	4228 ± 136.8	4223–4233	806.7–6551
Left Parietal Lobe	<0.0001	4226	4224–4227	4225 ± 118.8	4221–4230	803.6–5602
Right Occipital lobe	<0.0001	4227	4226–4228	4225 ± 120.5	4221–4230	804.6–6069
Left Occipital lobe	<0.0001	4224	4222–4226	4223 ± 122	4218–4227	803.6–6107

*: Shapiro-Wilk test was conducted to assess normal distribution (P-Value).
**: Standard Deviation.
ǂ: 95% Confidence Interval of the median.
ǂǂ: 95% Confidence Interval of the mean.

The Shapiro-Wilk tests showed that the EEG signals for all the regions considered in the study for both the like and dislike groups met the criteria for a normal distribution.

Fig. 2**A-J** represents the Forest plot of EEG Least Square Mean differences over time from baseline for the brain subdivisions considered in this study.

Significant differences in EEG signals were observed at 750 ms followed by 500 ms from baseline for the right and left hemispheres, the right frontal lobe, and the right and left sides of the occipital lobe **(**Fig. 2**-A, B, C, I, and J)**. The left frontal lobe showed an important difference in the EEG signals at 500 ms followed by 750 ms from baseline (Fig. 2**-D)**.

In the right parietal lobe, an important difference in EEG signals at 500 ms followed by 3000 ms from the baseline was reported (Fig. 2**-G)**. Fig. 2**-H** showed an important difference at 3000 ms followed by 500 ms in the left parietal lobe.

The forest plot in (Fig. 2**-E and F)** shows the least square mean differences over time from baseline for the right and left temporal lobe overlaps with the vertical line.

Overall, these results suggest that the most important EEG activity related to decision-making is made between 500 ms and 750 ms and even less at 3000 ms for the parietal lobe.

The results of the EGG signal variations using MMRM are presented in Figs. 3-12.

Fig. 3 illustrates the activity in the left hemisphere of the brain. The signal related to the like-response increased from the baseline to 750 ms and decreased from 1000 ms to the end of the experiment. The dislike-related response showed sustained activity from the beginning to the end of the task. Additionally, while there was no difference between the

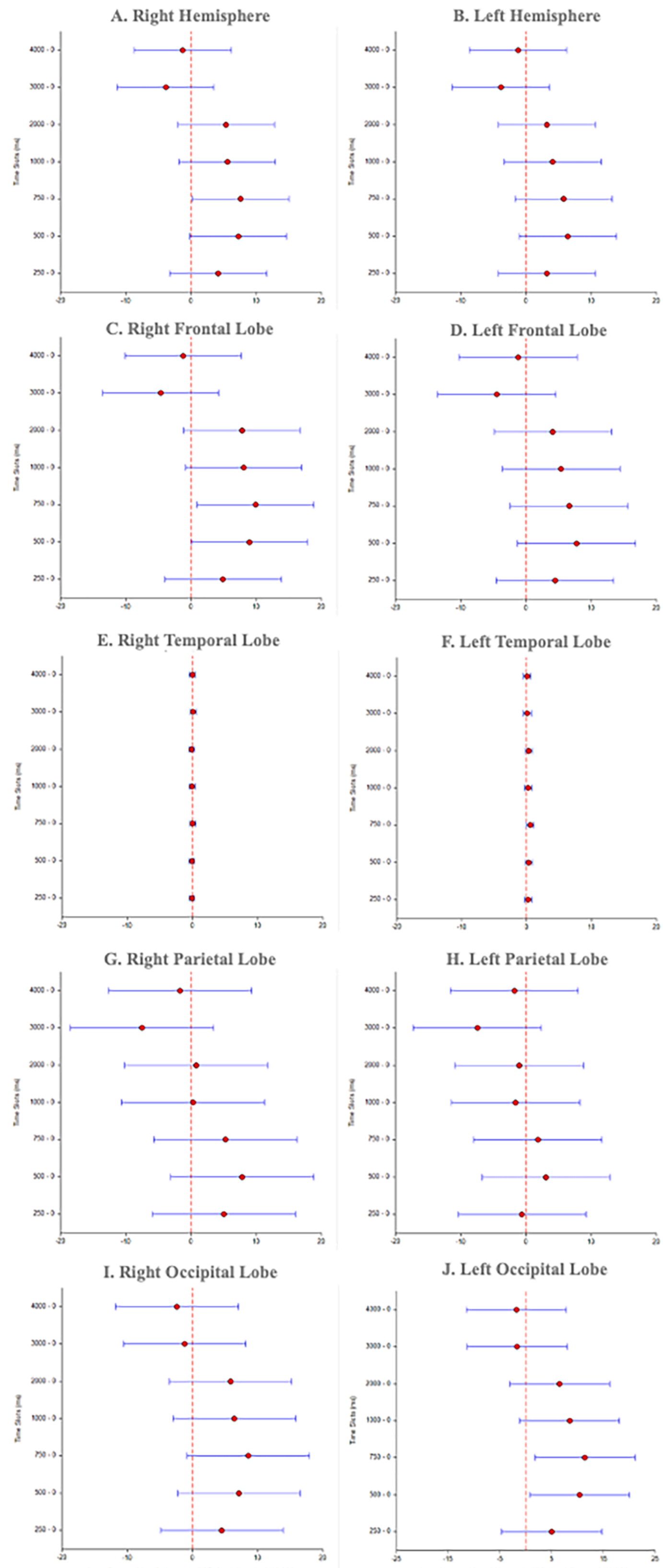

Fig. 2. Forest plot for Least Square Mean difference across time from the baseline for all the subjects included in the study (A. right hemisphere, B. Left Hemisphere, C. Right Frontal Lobe, D. Left Frontal Lobe, E. Right Temporal Lobe, F. Left Temporal Lobe, G. Right Parietal Lobe, H. Left Parietal Lobe, I. Right Occipital Lobe, and J. Left Occipital Lobe).

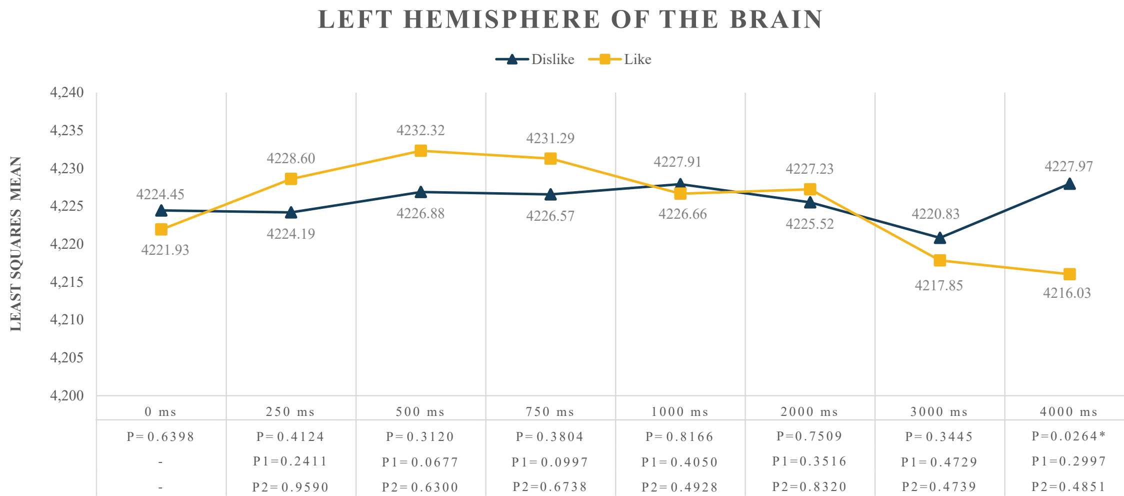

Fig. 3. EEG signal evolution across time in the left hemisphere of the brain according to like-dislike responses. The graphs show the EEG activity in the left hemisphere of the brain. The signal related to the like-response (orange line) increased from the baseline to 750 ms and decreased from 1000 ms to the end of the experiment. The dislike-related response (blue line) showed sustained activity from the beginning to the end of the task. A significant difference was observed at 4000 ms between like and dislike conditions (P = 0.0264). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

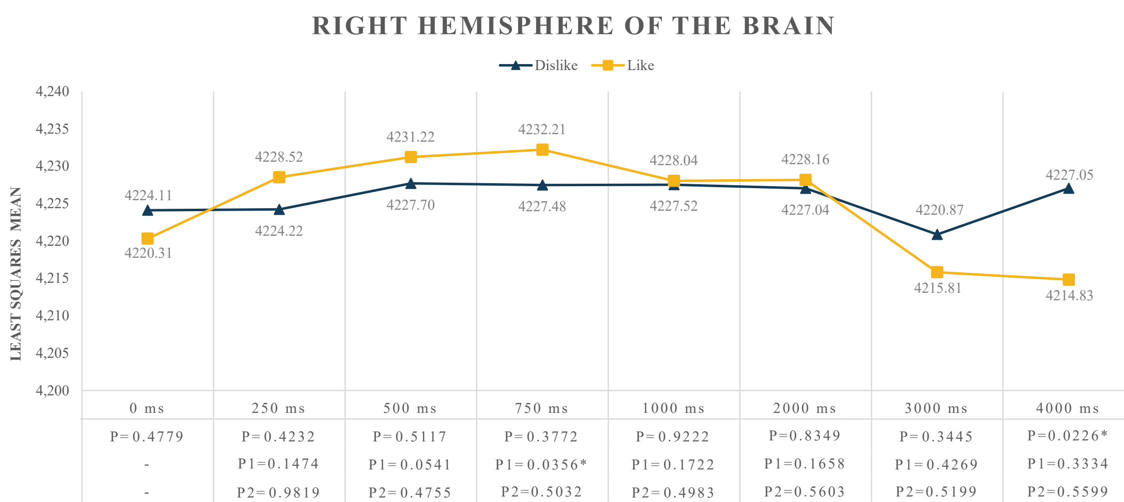

Fig. 4. EEG signal evolution across time in the right hemisphere of the brain according to like-dislike responses. The graphs show the EEG activity in the right hemisphere of the brain. For like-response (orange line), the EEG signal increased significantly from the baseline to 750 ms (P = 0.0356) and decreased from 1000 to 4000 ms. For dislike-response, (blue line), no significant changes in the signals were reported. A significant difference between like and dislike conditions was reported at 4000 ms (P = 0.0112). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

like and dislike conditions at the baseline, a statistically significant difference in the activity of the left hemisphere of the brain was observed at the end of the experiment (P = 0.0264 at 4000 ms).

In the right hemisphere of the brain (Fig. 4), the signal related to like-response increased significantly from the baseline to 750 ms (P = 0.0356) and decreased from 1000 to 4000 ms. However, no significant changes in the signals related to dislike-response from the baseline were reported. In addition, while no difference between like and dislike conditions was observed in the right hemisphere of the brain at the beginning of the test, a significant difference in the EEG signal was reported at 4000 ms (P = 0.0112).

In the left side of the frontal lobe (Fig. 5), the activity increased from the baseline to 500 ms (P = 0.060) and decreased from 750 ms to 4000 ms for the like-response condition. In contrast, the dislike-response condition did not show significant EEG changes over time. At 4000 ms, a significant difference in the activity of the left side of the frontal lobe was detected between like and dislike responses (P = 0.0457).

In the right side of the frontal lobe (Fig. 6), the activity increased from 0 ms to 750 ms and decreased from 1000 ms to 4000 ms for the like-response condition. In contrast, the activity related to the dislike-response condition remained relatively unchanged. Interestingly, when comparing EEG signal of like and dislike groups at 4000 ms, the activity was significantly lower in the right side of the frontal lobe of like-response group (P = 0.0301).

The activity in the left side of the parietal lobe (Fig. 7) showed a pattern of increasing from the baseline level until reaching a peak at 500 ms, then decreasing until 1000 ms during the like-response condition. This activity increased again at 2000 ms followed by a decrease at 3000 ms and a slight increase at 4000 ms, although these changes were not statistically significant. For the dislike-response condition, a slight

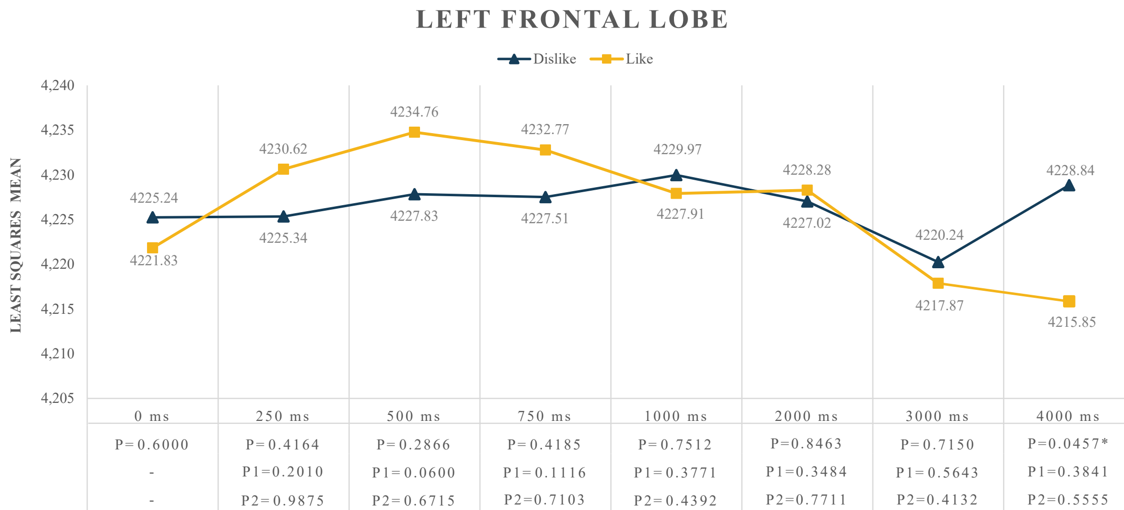

Fig. 5. EEG signal evolution across time in the left frontal lobe according to like-dislike responses. The graphs show the EEG activity in the left frontal lobe of the brain. For like-response condition (orange line), the activity on the left side increased from the baseline to 500 ms and decreased from 750 ms to 4000 ms. No significant changes over time in the dislike-response condition (blue line) was reported. At 4000 ms, a significant difference was reported between like and dislike responses (P = 0.0457). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

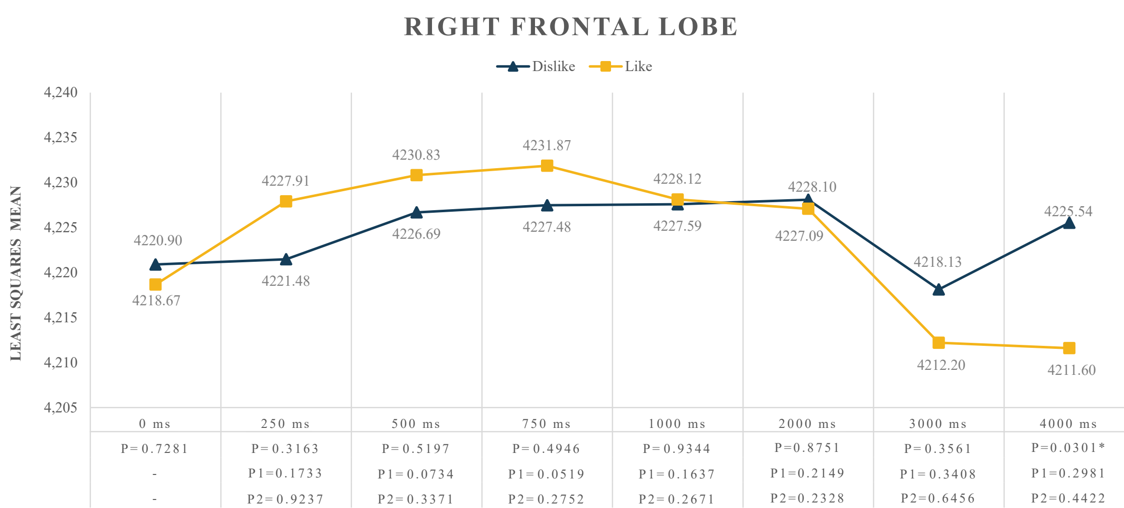

Fig. 6. EEG signal evolution across time in the right frontal lobe according to like-dislike responses. The graphs show the EEG activity in the right frontal lobe of the brain. For like-response condition (orange line), the activity increased from 0 ms to 750 ms and decreased from 1000 ms to 4000 ms for the like-response condition (statistically non-significant changes from baseline). The activity related to the dislike-response condition (blue line), remained relatively unchanged. A significant difference was observed between like and dislike responses at 4000 ms (P = 0.0301). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

oscillation in the activity was reported from 0 ms to 3000 ms followed by an increase at 4000 ms. These changes in the activity were not statistically significant. However, the comparison of the activity at 4000 ms between like and dislike responses showed a statistically significant low activity in the left side of the parietal lobe for like-response group (P = 0.0410).

In the right side of the parietal lobe (Fig. 8), the activity for the like-response condition increased from 0 ms to reach a peak at 500 ms (P = 0.0486). This was followed by a decrease until 1000 ms, a slight increase at 2000 ms, a decrease at 3000 ms, and a slight increase at 4000 ms. Nonetheless, these changes from 750 ms to 4000 ms were not statistically significant. Meanwhile, the activity related to dislike response showed a slight decrease to reach a minimum at 3000 ms followed by a slight increase at 4000 ms. These changes were not statistically significant. By contrast, the activity in the right side of the parietal lobe at 4000 ms was reported to be significantly lower for the like-response group compared to the dislike-response group (P = 0.0410).

In the left side of the occipital lobe (Fig. 9), the signal related to like-response increased significantly from 0 ms to reach a maximum at 750 ms (P = 0.00336 at 500 ms and P = 0.0148 at 750 ms), followed by a statistically non-significant tapering of the activity from 1000 ms to the end of the test. The EEG signal related to the dislike-response condition showed an increase of the activity at 500 ms that was maintained until 2000 ms, followed by a decrease at 3000 ms and a slight increase at 4000 ms. These changes were not statistically significant. The comparison of the EEG signal at 4000 ms between like and dislike responses revealed an interesting significant decrease of the activity in the left side of the occipital lobe for like-response group (P = 0.0015).

In the right side of the occipital lobe (Fig. 10), the activity related to like-response significantly increased from the baseline to 750 ms (P = 0.0379), followed by a decrease at 1000 ms and a significant increase at 2000 ms (P = 0.0485), and then a non-significant decrease from 3000

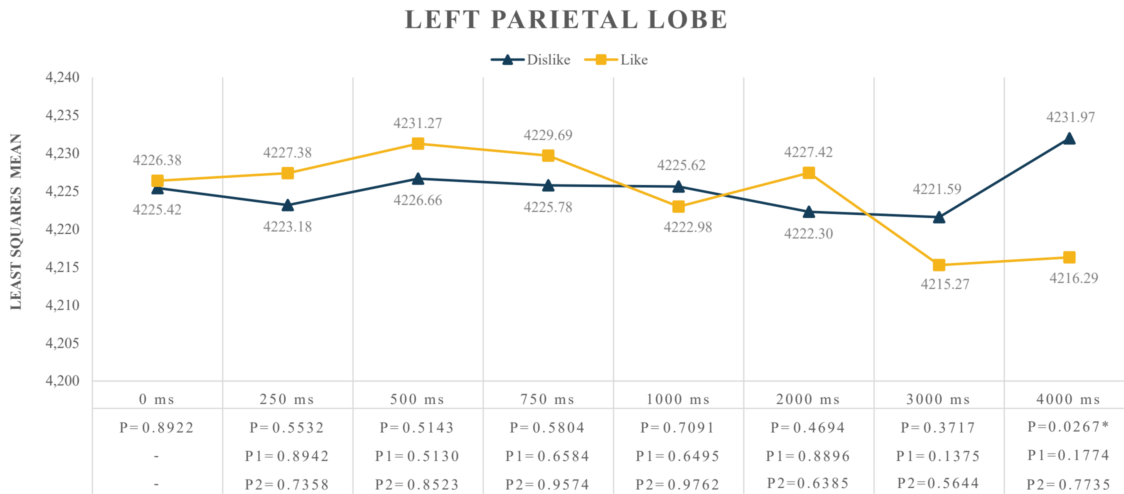

Fig. 7. EEG signal evolution across time in the left parietal lobe according to like-dislike responses. The graphs show the EEG activity in the left parietal lobe of the brain. For like-response condition (orange line), the activity increased from the baseline to 500 ms, then decreased at 1000 ms. This activity increased again at 2000 ms followed by a decrease at 3000 ms and a slight increase at 4000 ms (statistically non-significant changes). In the dislike-response condition (blue line), a slight oscillation in the activity was reported from 0 ms to 3000 ms followed by an increase at 4000 ms (statistically non-significant changes). A significant difference was found between like and dislike responses at 4000 ms (P = 0.0410). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

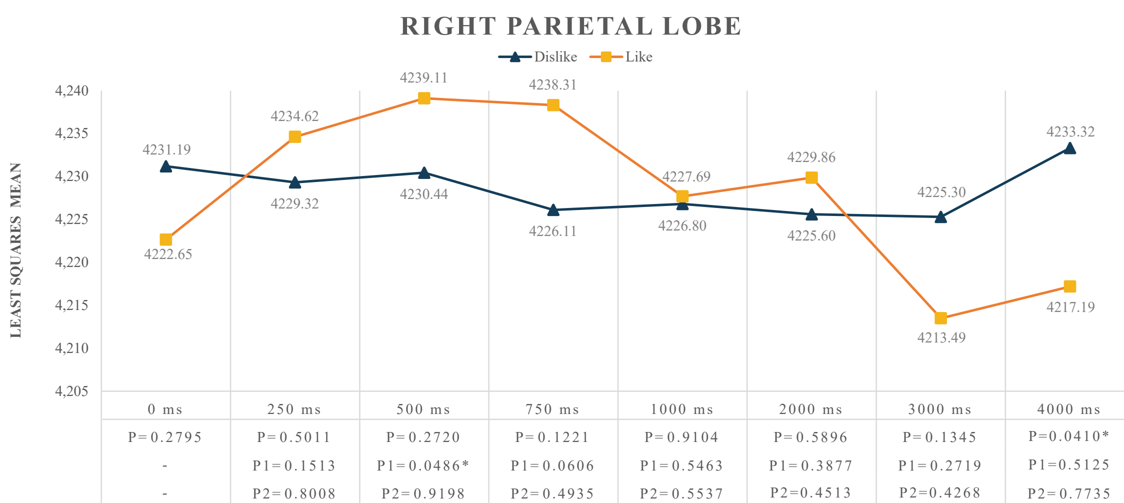

Fig. 8. EEG signal evolution across time in the right parietal lobe according to like-dislike responses. The graphs show the EEG activity in the right parietal lobe of the brain. For like-response condition (orange line), the activity increased from 0 ms to 500 ms (P = 0.0486), followed by oscillations between 750 ms and 4000 ms (changes statistically non-significant). Meanwhile, the activity related to dislike response (blue line), showed a slight decrease to reach a minimum at 3000 ms followed by a slight increase at 4000 ms (statistically non-significant changes).A significant difference was observed between like and dislike responses at 4000 ms (P = 0.0410). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ms to 4000 ms. For the dislike-response condition, a slight increase of the activity was reported reaching a maximum at 750 ms followed by a decrease that is maintained until 3000 ms and an increase at 4000 ms. This variation in the activity remains statistically insignificant.

The comparison of the activity in the right side of the occipital lobe between like and dislike responses across time revealed a significant difference in the activity at 4000 ms only (P = 0.0434).

Surprisingly, there is no significant difference in the activity between like and dislike-related responses from the beginning to the end of the test in both sides in the temporal lobe (Figs. 11 and 12). Also, no significant change in the signal was reported for like and dislike preferences across time from the baseline.

The results of the MMRM analysis were used to classify consumer preferences using the received EEG signal from the left/right hemisphere, left/right frontal lobe, left/right parietal lobe, and left/right occipital lobe at 4000 ms. The performance of four classifiers (Neural Network, Gradient Boosting, kNN, and Random Forest) was evaluated using the AUC, accuracy, F1-scores, precision, and recall, and the results are shown in Table 5. The Neural Network classifier had the highest performance for predicting consumer preference in the left and right hemispheres, with AUC scores of 76.02% and 75.95%, respectively. The Gradient Boosting classifier had the second-highest performance (73.74% and 73.59%), followed by kNN (71.81% and 71.97%) and Random Forest (69.76% and 69.69%).

The left and right sides of the frontal, parietal and occipital lobes showed similar results. For the left and right sides of the frontal lobe, the Neural Network classifier had the highest AUC at 75.94% and 75.97%, respectively, followed by the Gradient Boosting classifier, which had AUC scores of 74.57% and 74.99%, respectively. For the left and right sides of the parietal lobe, the Neural Network classifier had the highest

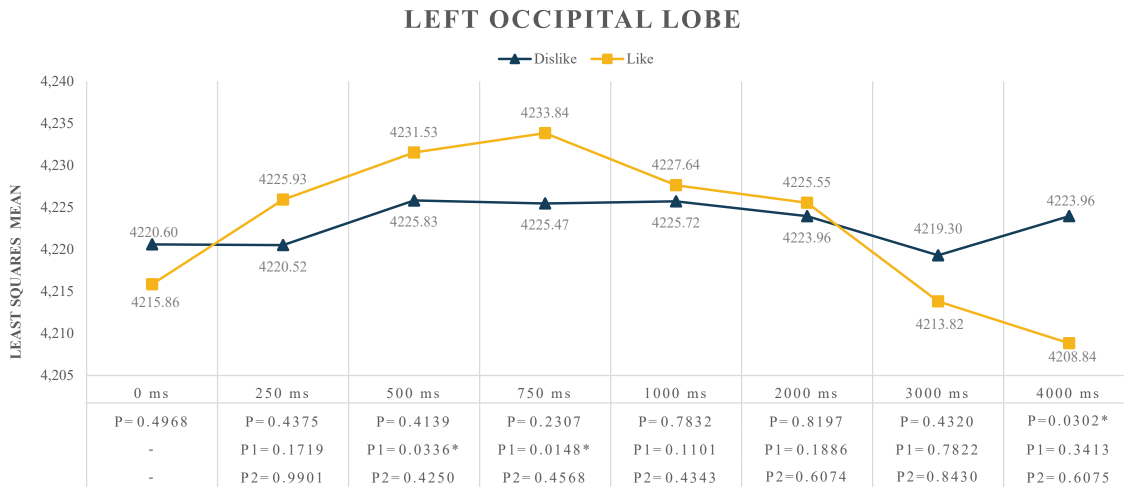

Fig. 9. EEG signal evolution across time in the left occipital lobe according to like-dislike responses. The graphs show the EEG activity in the left occipital lobe of the brain. For like-response condition (orange line), the signal increased significantly from 0 ms to reach a maximum at 750 ms (P = 0.00336 at 500 ms and P = 0.0148 at 750 ms), followed by a statistically non-significant decrease of the activity from 1000 ms to the end of the test. The signal related to the dislike-response condition (blue line) showed an increase of the activity at 500 ms that was maintained until 2000 ms, followed by a decrease at 3000 ms and a slight increase at 4000 ms (changes statistically non-significant). A significant difference was reported between like and dislike responses at 4000 ms (P = 0.0015). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

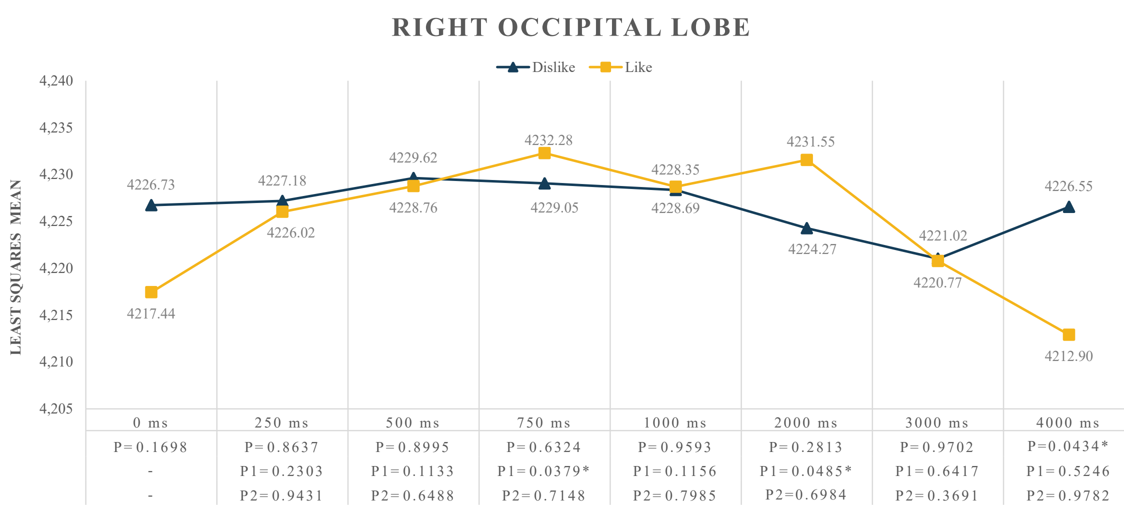

Fig. 10. EEG signal evolution across time in the right occipital lobe according to like-dislike responses. The graphs show the EEG activity in the right occipital lobe of the brain. For like-response condition (orange line), the activity significantly increased from the baseline to 750 ms (P = 0.0379), followed by a decrease at 1000 ms and a significant increase at 2000 ms (P = 0.0485), and then a non-significant decrease from 3000 ms to 4000 ms. For the dislike-response condition (blue line), a slight increase was reported reaching a maximum at 750 ms followed by a decrease till 3000 ms and an increase at 4000 ms (changes statistically non-significant). A significant difference was observed between like and dislike responses at the end of the test (P = 0.0434). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

AUC at 76.61% and 76.45%, respectively, followed by the Gradient Boosting classifier, which had AUC scores of 75.08% and 75.33%, respectively.

The Neural Network classifier had the highest AUC of 76.61% for the right parietal lobe and the lowest AUC of 75.94% for the left frontal lobe. The Gradient Boosting classifier had the highest AUC of 75.33% for the left parietal lobe and the lowest AUC of 73.59% for the right hemisphere. The kNN classifier had the highest AUC of 73.55% for the right frontal lobe and the lowest AUC of 71.05% for the left parietal lobe. The Random Forest classifier had the highest AUC of 72.62% for the right frontal lobe and the lowest AUC of 69.56% for the left frontal lobe.

5. Discussion

Previous research in the field of neuromarketing has primarily focused on implementing EEG-based machine learning systems to predict consumer preferences for products. The current study used MMRM to examine the EEG signal variations within and between subjects and then implemented a classification system to predict consumer preferences for e-commerce products. This study specifically examined differences in brain activity between subjects who liked and disliked the products, as well as differences within subjects who had either a positive (like) or negative (Dislike) preference for the products.

Our results showed a significant increase in the activity at 500 ms in the right parietal and the left occipital regions, at 750 ms in the right hemisphere, and the occipital areas, and at 2000 ms in the right occipital

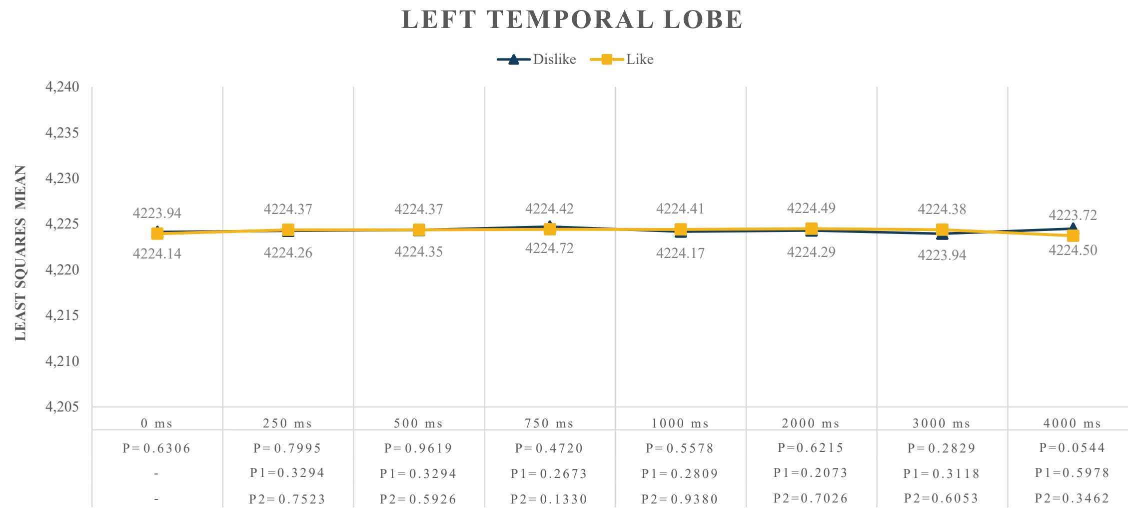

Fig. 11. EEG signal evolution across time in the left temporal lobe according to like-dislike responses. The graphs show the EEG activity in the left temporal lobe of the brain. No significant changes in the EEG signal were reported for like (orange line) and dislike (blue line) responses conditions across time from the baseline. There is no significant difference in the activity between like and dislike-related responses from the beginning to the end of the test. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

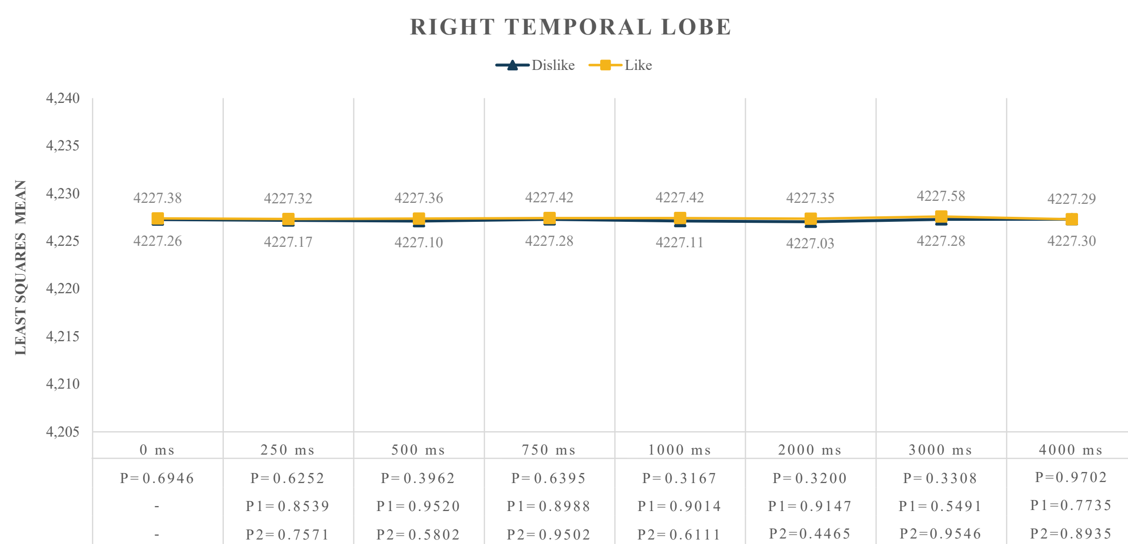

Fig. 12. EEG signal evolution across time in the right temporal lobe according to like-dislike responses. The graphs show the EEG activity in the right temporal lobe of the brain. No significant changes in the EEG signal were reported for like (orange line) and dislike (blue line) responses conditions across time from the baseline. There is no significant difference in the activity between like and dislike-related responses from the beginning to the end of the test. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

region when participants made a positive preference decision for products. However, no significant change in brain activity was observed for the negative preference condition. This may be due to differences in the brain's neural networks involved in processing positive and negative preference decisions. Previous research has shown that brain power is generally higher when a subject likes an object compared to when they dislike it [17,41]. Interestingly, the temporal lobe seems to be not involved in this process.

At 4000 ms, we observed a significant decrease in brain activity when subjects liked a product, which allowed us to identify a significant difference in brain activity between positive and negative preference responses in the right and left hemispheres, as well as the frontal, parietal, and occipital regions. This time point can potentially be used as a good predictor of consumer preference. The classification results showed that Neural Network classifier had the highest performance, with an AUC of 76.61% and 76.45% for the right and left parietal lobes, respectively, and 76.06% and 76.03% for the left and right occipital lobes, respectively. The Neural Network also had the highest AUC for the left and right hemispheres at 76.02% and 75.95%, respectively, and for the left and right frontal lobes at 75.97% and 75.94%, respectively. The Gradient Boosting also had a competitive performance. In a previous study, Yadava et al [16] reported that the maximum accuracy using an HMM classifier on EEG signals was around 70.33% for different brain regions and time durations, but did not mention the AUC. It is well known that the accuracy metric does not distinguish between false positive and false negative classification, which makes it is insufficient to assess the classification performance.

Our approach demonstrates great value for binary classification and can be improved by using other models. In fact, Sparse Representation Classification (SRC), which works by expressing a test sample as a sparse linear combination of training samples, has been reported to show superior performance in the discrimination of consumer preferences

Table 5
Performance metrics (%) of different brain regions at 4000 ms.

	Model	AUC	Accuracy	F1	Precision	Recall
Left Hemisphere	Neural Network	76.02	68.61	68.60	68.58	68.61
	Gradient Boosting	73.74	67.46	66.75	67.37	67.46
	kNN	71.81	64.40	64.48	64.63	64.40
	Random Forest	69.76	63.73	63.64	63.59	63.73
Right Hemisphere	Neural Network	75.95	68.04	67.97	67.94	68.04
	Gradient Boosting	73.59	67.27	66.67	67.11	67.27
	kNN	71.97	65.65	65.60	65.57	65.65
	Random Forest	69.69	62.68	62.66	62.65	62.68
Left Frontal	Neural Network	75.94	68.71	68.70	68.69	68.71
	Gradient Boosting	74.57	67.37	66.51	67.36	67.37
	kNN	71.81	65.55	65.59	65.63	65.55
	Random Forest	69.56	63.44	63.37	63.32	63.44
Right Frontal	Neural Network	75.97	68.13	68.07	68.04	68.13
	Gradient Boosting	74.99	66.99	66.11	66.95	66.99
	kNN	73.55	66.6	66.6	66.6	66.6
	Random Forest	72.62	64.69	64.52	64.48	64.69
Left Parietal	Neural Network	76.45	68.9	68.92	68.94	68.9
	Gradient Boosting	75.33	68.52	67.71	68.59	68.52
	kNN	71.05	65.17	65.11	65.08	65.17
	Random Forest	70.31	64.4	64.39	64.37	64.4
Right Parietal	Neural Network	76.61	69.28	69.28	69.28	69.28
	Gradient Boosting	75.08	69.00	68.21	69.10	69.00
	kNN	72.19	67.08	67.01	66.98	67.08
	Random Forest	71.92	66.51	66.41	66.37	66.51
Left Occipital	Neural Network	76.06	68.90	68.89	68.88	68.90
	Gradient Boosting	74.50	67.08	66.46	66.91	67.08
	kNN	72.16	68.42	68.46	68.52	68.42
	Random Forest	69.73	64.69	64.66	64.64	64.69
Right Occipital	Neural Network	76.03	68.33	68.30	68.28	68.33
	Gradient Boosting	73.92	67.85	67.30	67.70	67.85
	kNN	71.98	66.32	66.35	66.40	66.32
	Random Forest	70.67	65.17	65.16	65.15	65.17

Bold letters show the maximum AUC classification performance of the classifier.

between the least and most preferred products compared to SVM, kNN, typical Deep Learning Neural Network (DLNN) and Random Forest [54].

In addition, Chikara et al. [55] studied the outcomes of a hierarchical classification model with brain connectivity method to classify the neural activities of eight brain cortices related to human inhibition. In this model, phase-locking value method was utilized as the optimal features with parametric classification algorithms to classify EEG data and achieved a high accuracy of 94.44% by quadratic discriminant analysis.

In multi-class prediction, our proposed approach can be effectively adapted to handle this more complex classification task. Thus, the MMRM for longitudinal analysis of EEG signals from different brain regions will remain unchanged. Similarly, the 10-fold cross-validation procedure will continue to assures a balanced and unbiased use of all data points across the training and validation phases. The machine learning algorithms such as Random Forest, kNN, and ANN are all capable of handling multiclass classification problems by reconfiguring the output layer of each algorithm to contain neurons equivalent to the number of likability classes, each neuron corresponding to one class. Also, deep learning algorithms can be used to investigate multi-class prediction.

The current study has some limitations. First, we used one dataset with a relatively small number of subjects which may limit the performance of the machine learning methods used. Second, the dataset was used without pre-processing or feature selection or extraction steps which can lead to noisy results and other inconsistencies. Third, we did not address the difference in the four different frequency bands (delta, theta, alpha, and beta). Fourth, the current study was limited to neuromarketing binary classification (like/dislike) and should be extended to multiclass classification problems.

6. Conclusion

In this paper, we proposed a new framework for predicting consumer choice. The results of the present study clearly indicate that the brain regions considered exhibit different patterns of activity between like and dislike-related responses. While the positive preference decisions were associated with an increase in brain activity in various regions, this increase was not present in the case of negative preferences. The significant difference in the EEG signals observed between like and dislike subjects at 4000 ms is due to the significant decrease in the activity of brain regions related to the like decision at the end of the test. This highlights the importance of the longitudinal analysis using the mixed model for repeated measures which allows us to focus the classification step on the end of the test which is considered a good time for predicting subject preference. The Neural Network classifier emerged as the best classifier with improved classification performance. Our framework is expected to contribute to future research on neuromarketing. Future studies should consider using both frequency-based and time-based methods to better explore EEG data and to classify consumer preferences more effectively, as well as further explore the neuropsychological basis of consumer preferences.

Ethics Approval

Not applicable.

Funding

The authors have no relevant financial or non-financial interests to disclose.

CRediT authorship contribution statement

Mounir Ouzir: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Houda Chakir Lamrani:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing. **Rachel L. Bradley:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Ismail El Moudden:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] V. Sebastian, New Directions in Understanding the Decision-making Process: Neuroeconomics and Neuromarketing, Procedia Soc. Behav. Sci. 127 (2014) 758–762, https://doi.org/10.1016/j.sbspro.2014.03.350.

[2] G. Egidi, H.C. Nusbaum, J.T. Cacioppo, Neuroeconomics: Foundational issues and consumer relevance., in: Handbook of Consumer Psychology., Taylor & Francis Group/Lawrence Erlbaum Associates, New York, NY, 2008: pp. 1177–1214.

[3] P. Cherubino, A.C. Martinez-Levy, M. Caratù, G. Cartocci, G. Di Flumeri, E. Modica, D. Rossi, M. Mancini, A. Trettel, Consumer behaviour through the eyes of neurophysiological measures: State-of-the-art and future trends, Comput. Intell. Neurosci. 2019 (2019), https://doi.org/10.1155/2019/1976847.

[4] V. Khurana, M. Gahalawat, P. Kumar, P.P. Roy, D.P. Dogra, E. Scheme, M. Soleymani, A Survey on Neuromarketing Using EEG Signals, IEEE Trans Cogn Dev Syst. 13 (2021) 732–749, https://doi.org/10.1109/TCDS.2021.3065200.

[5] M. Royo-Vela, Á. Varga, Unveiling Neuromarketing and Its Research Methodology, Encyclopedia. 2 (2022) 729–751, https://doi.org/10.3390/encyclopedia2020051.

[6] F.S. Rawnaque, K.M. Rahman, S.F. Anwar, R. Vaidyanathan, T. Chau, F. Sarker, K. A. al Mamun, Technological advancements and opportunities in Neuromarketing: a systematic review, Brain Inform 7 (2020), https://doi.org/10.1186/s40708-020-00109-x.

[7] F.P. Kalaganis, K. Georgiadis, V.P. Oikonomou, N.A. Laskaris, S. Nikolopoulos, I. Kompatsiaris, Unlocking the Subconscious Consumer Bias: A Survey on the Past, Present, and Future of Hybrid EEG Schemes in Neuromarketing, Frontiers in Neuroergonomics. 2 (2021), https://doi.org/10.3389/fnrgo.2021.672982.

[8] L.H. Chew, J. Teo, J. Mountstephens, Aesthetic preference recognition of 3D shapes using EEG, Cogn. Neurodyn. 10 (2016) 165–173, https://doi.org/10.1007/s11571-015-9363-z.

[9] L. Zeng, M. Lin, K. Xiao, J. Wang, H. Zhou, Like/Dislike Prediction for Sport Shoes With Electroencephalography: An Application of Neuromarketing, Front. Hum. Neurosci. 15 (2022), https://doi.org/10.3389/fnhum.2021.793952.

[10] S.K. Hadjidimitriou, L.J. Hadjileontiadis, Toward an EEG-based recognition of music liking using time-frequency analysis, I.E.E.E. Trans. Biomed. Eng. 59 (2012) 3498–3510, https://doi.org/10.1109/TBME.2012.2217495.

[11] S.K. Hadjidimitriou, L.J. Hadjileontiadis, EEG-Based classification of music appraisal responses using time-frequency analysis and familiarity ratings, IEEE Trans. Affect. Comput. 4 (2013) 161–172, https://doi.org/10.1109/T-AFFC.2013.6.

[12] J. Moon, Y. Kim, H. Lee, C. Bae, W.C. Yoon, Extraction of user preference for video stimuli using eeg-based user responses, ETRI J. 35 (2013) 1105–1114, https://doi.org/10.4218/etrij.13.0113.0194.

[13] M. Murugappan, S. Murugappan, C.G. Balaganapathy, Wireless EEG signals based Neuromarketing system using Fast Fourier Transform (FFT), in: 2014 IEEE 10th International Colloquium on Signal Processing and Its Applications, 2014, pp. 25–30, https://doi.org/10.1109/CSPA.2014.6805714.

[14] I.H. Sourov, F.A. Ahmed, M.T.I. Opu, A.K. Mutasim, M.R. Bashar, R.S. Tipu, M. A. Amin, M.K. Islam, G. Ali, EEG-Based Preference Classification for Neuromarketing Application, Comput. Intell. Neurosci. 2023 (2023) 1–13.

[15] A. Ullah, G. Baloch, A. Ali, A.B. Buriro, J. Ahmed, B. Ahmed, S. Akhtar, Neuromarketing Solutions based on EEG Signal Analysis using Machine Learning, International Journal of Advanced Computer Science and Applications. 13 (2022). 10.14569/IJACSA.2022.0130137.

[16] M. Yadava, P. Kumar, R. Saini, P.P. Roy, D., Prosad Dogra, Analysis of EEG signals and its application to neuromarketing, Multimed. Tools Appl. 76 (2017) 19087–19111, https://doi.org/10.1007/s11042-017-4580-6.

[17] J. Guixeres, E. Bigné, J.M.A. Azofra, M.A. Raya, A.C. Granero, F.F. Hurtado, V. N. Ornedo, Consumer neuroscience-based metrics predict recall, liking and viewing rates in online advertising, Front. Psychol. 8 (2017), https://doi.org/10.3389/fpsyg.2017.01808.

[18] L.M. Soria Morillo, J.A. Alvarez-Garcia, L. Gonzalez-Abril, J.A. Ortega Ramírez, Discrete classification technique applied to TV advertisements liking recognition system based on low-cost EEG headsets, Biomed. Eng. Online 15 (2016), https://doi.org/10.1186/s12938-016-0181-2.

[19] J. Goulet-Kennedy, S. Labbe, S. Fecteau, The involvement of the striatum in decision making, Dialogues Clin. Neurosci. 18 (1) (2016) 55–63.

[20] M.Á. Serrano, F. Molins, A. Alacreu-Crespo, Human Decision-Making Evaluation: From Classical Methods to Neurocomputational Models, in: J.L. García Alcaraz, A. Realyvásquez Vargas (Eds.), Algorithms and Computational Techniques Applied to Industry, Springer International Publishing, Cham, 2022, pp. 163–181, https://doi.org/10.1007/978-3-031-00856-6_9.

[21] P. Renvoise, C. Morin, Neuromarketing: Understanding the Buy Buttons in Your Customer's Brain, 2007. https://www.harpercollinsleadership.com/9781418570309/neuromarketing/.

[22] P.M. Cortes, J.P. García-Hernández, F.A. Iribe-Burgos, M. Hernández-González, C. Sotelo-Tapia, M.A. Guevara, Temporal division of the decision-making process: An EEG study, Brain Res. 1769 (2021), https://doi.org/10.1016/j.brainres.2021.147592.

[23] N.Y. So, M.N. Shadlen, Decision formation in parietal cortex transcends a fixed frame of reference, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.07.019.

[24] Y. Zhou, D.J. Freedman, Posterior parietal cortex plays a causal role in perceptual and categorical decisions, Science 365 (2019) (1979) 180–185, https://doi.org/10.1126/science.aaw8347.

[25] G. Guo, M. Elgendi, A New Recommender System for 3D E-Commerce: An EEG Based Approach, J. Adv. Manage. Sci. 1 (2013) 61–65. 10.12720/joams.1.1.61-65.

[26] C. Summerfield, K. Tsetsos, Building bridges between perceptual and economic decision-making: Neural and computational mechanisms, Front. Neurosci. (2012), https://doi.org/10.3389/fnins.2012.00070.

[27] Y. Du, L. He, Y. Wang, D. Liao, The Neural Mechanism of Long-Term Motor Training Affecting Athletes' Decision-Making Function: An Activation Likelihood Estimation Meta-Analysis, Front. Hum. Neurosci. 16 (2022), https://doi.org/10.3389/fnhum.2022.854692.

[28] Y. Zhao, D. Wang, X. Wang, S.C. Chiu, Brain mechanisms underlying the influence of emotions on spatial decision-making: An EEG study, Front. Neurosci. 16 (2022), https://doi.org/10.3389/fnins.2022.989988.

[29] T. Larsen, J.P. O'Doherty, Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: A combined fmri–EEG study, Philos. Trans. R. Soc., B 369 (2014), https://doi.org/10.1098/rstb.2013.0473.

[30] H.T. Akiba, M.F. Costa, J.S. Gomes, E. Oda, P.B. Simurro, A.M. Dias, Neural correlates of preference: A transmodal validation study, Front. Hum. Neurosci. 13 (2019), https://doi.org/10.3389/fnhum.2019.00073.

[31] M.A.S. Boksem, A. Smidts, Brain responses to movie trailers predict individual preferences for movies and their population-wide commercial success, J. Mark. Res. 52 (2015) 482–492, https://doi.org/10.1509/jmr.13.0572.

[32] N. Ravaja, O. Somervuori, M. Salminen, Predicting purchase decision: The role of hemispheric asymmetry over the frontal cortex, J. Neurosci. Psychol. Econ. 6 (2013) 1–13, https://doi.org/10.1037/a0029949.

[33] G. Vecchiato, P. Cherubino, A. Trettel, F. Babiloni, The track of the electric and the magnetic brain activity, Biosystems and Biorobotics, Springer International Publishing, 2013, pp. 33–44.

[34] B. Yilmaz, S. Korkmaz, D.B. Arslan, E. Güngör, M.H. Asyali, Like/dislike analysis using EEG: Determination of most discriminative channels and frequencies, Comput. Methods Programs Biomed. 113 (2014) 705–713, https://doi.org/10.1016/j.cmpb.2013.11.010.

[35] S. Koelstra, A. Yazdani, M. Soleymani, C. Mühl, J.-S. Lee, A. Nijholt, T. Pun, T. Ebrahimi, I. Patras, Single Trial Classification of EEG and Peripheral Physiological Signals for Recognition of Emotions Induced by Music Videos, in: Y. Yao, R. Sun, T. Poggio, J. Liu, N. Zhong, J. Huang (Eds.), Brain Inform, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2010, pp. 89–100.

[36] Y. Pan, C. Guan, J. Yu, K.K. Ang, T.E. Chan, Common frequency pattern for music preference identification using frontal EEG, in: 2013 6th International IEEE/EMBS Conference on Neural Engineering (NER), 2013: pp. 505–508. 10.1109/NER.2013.6695982.

[37] G. Vecchiato, J. Toppi, L. Astolfi, F. de Vico Fallani, F. Cincotti, D. Mattia, F. Bez, F. Babiloni, Spectral EEG frontal asymmetries correlate with the experienced pleasantness of TV commercial advertisements, Med. Biol. Eng. Compu. 49 (2011) 579–583, https://doi.org/10.1007/s11517-011-0747-x.

[38] S.S. Bosshard, J.D. Bourke, S. Kunaharan, M. Koller, P. Walla, Established liked versus disliked brands: Brain activity, implicit associations and explicit responses, Cogent Psychol. 3 (2016), https://doi.org/10.1080/23311908.2016.1176691.

[39] D. Baldo, H. Parikh, Y. Piu, K.M. Müller, Brain Waves Predict Success of New Fashion Products: A Practical Application for the Footwear Retailing Industry, J. Creat. Value. 1 (2015) 61–71, https://doi.org/10.1177/2394964315569625.

[40] S. Alsmadi, K. Hailat, Neuromarketing and Improved Understanding of Consumer Behaviour through Brain-Based Neuro Activity, J. Inf. Knowl. Manag. 20 (2021), https://doi.org/10.1142/S0219649221500209.

[41] P. Golnar-Nik, S. Farashi, M.S. Safari, The application of EEG power for the prediction and interpretation of consumer decision-making: A neuromarketing study, Physiol. Behav. 207 (2019) 90–98, https://doi.org/10.1016/j.physbeh.2019.04.025.

[42] J.H. Friedman, Greedy Function Approximation: A Gradient Boosting Machine, Ann. Stat. 29 (2001) 1189–1232. http://www.jstor.org/stable/2699986.

[43] E. Fix, J.L. Hodges, Discriminatory Analysis - Nonparametric Discrimination: Consistency Properties, Int. Stat. Rev. 57 (3) (1989) 238.

[44] A.J. Gallego, J.R. Rico-Juan, J.J. Valero-Mas, Efficient k-nearest neighbor search based on clustering and adaptive k values, Pattern Recogn. 122 (2022), 108356.
[45] M. Bansal, A. Goyal, A. Choudhary, A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning, Decision Analy. J. 3 (2022), 100071, https://doi.org/10.1016/j.dajour.2022.100071.
[46] R. Gill, J. Singh, A study of neuromarketing techniques for proposing cost effective information driven framework for decision making, Mater Today Proc, Elsevier Ltd, 2020, pp. 2969–2981.
[47] M. Chakraborty, S.K. Biswas, B. Purkayastha, Rule extraction using ensemble of neural network ensembles, Cogn Syst. Res. 75 (2022) 36–52, https://doi.org/10.1016/j.cogsys.2022.07.004.
[48] L. Breiman, Random Forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[49] R.C. Littell, W.W. Stroup, R.J. Freund, SAS for Linear Models, in: Wiley, 2002. https://www.wiley.com/en-us/SAS+for+Linear+Models%2C+4th+Edition-p-9780471221746.
[50] S.S. Shapiro, M.B. Wilk, An Analysis of Variance Test for Normality (Complete Samples), Biometrika 52 (1965) 591–611, https://doi.org/10.2307/2333709.
[51] G.M. Fitzmaurice, N.M. Laird, J.H. Ware, Applied Longitudinal Analysis, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2011. 10.1002/9781119513469.
[52] A. Gałecki, T. Burzykowski, Linear Mixed-Effects Models Using R, Springer New York, New York, NY, 2013. 10.1007/978-1-4614-3900-4.
[53] G. Molenberghs, G. Verbeke, Models for Discrete Longitudinal Data, Springer-Verlag, New York, 2005. 10.1007/0-387-28980-1.
[54] V.P. Oikonomou, K. Georgiadis, F. Kalaganis, S. Nikolopoulos, I. Kompatsiaris, A Sparse Representation Classification Scheme for the Recognition of Affective and Cognitive Brain Processes in Neuromarketing, Sensors 23 (2023), https://doi.org/10.3390/s23052480.
[55] R.K. Chikara, L.-W. Ko, Neural Activities Classification of Human Inhibitory Control Using Hierarchical Model, Sensors 19 (2019) 3791, https://doi.org/10.3390/s19173791.