Independent University **Bangladesh (IUB)**

IUB Academic Repository

Electrical and Electronics Engineering Article

2023-03-01

EEG-Based Preference Classification for Neuromarketing Application

Sourov, Injamamul Haque

Hindawi

Injamamul Haque Sourov, Faiyaz Alvi Ahmed, Md. Tawhid Islam Opu, Aunnoy K. Mutasim, M. Raihanul Bashar, Rayhan Sardar Tipu, Md. Ashraful Amin, Md. Kafiul Islam, "EEG-Based Preference Classification for Neuromarketing Application", Computational Intelligence and Neuroscience, vol. 2023, Article ID 4994751, 13 pages, 2023. https://doi.org/10.1155/2023/4994751

https://ar.iub.edu.bd/handle/11348/559

Downloaded from IUB Academic Repository

Hindawi
Computational Intelligence and Neuroscience
Volume 2023, Article ID 4994751, 13 pages
https://doi.org/10.1155/2023/4994751

Research Article

EEG-Based Preference Classification for Neuromarketing Application

Injamamul Haque Sourov,[1] **Faiyaz Alvi Ahmed**,[1] **Md. Tawhid Islam Opu**,[1] **Aunnoy K. Mutasim**,[2] **M. Raihanul Bashar**,[2] **Rayhan Sardar Tipu**,[2] **Md. Ashraful Amin**,[2] **and Md. Kafiul Islam**[1]

[1]*Biomedical Instrumentation and Signal Processing Lab, Department of Electrical and Electronic Engineering, Independent University, Bangladesh, Plot 16 Block B, Bashundhara R/A, Dhaka 1229, Bangladesh*
[2]*Agency Lab, Department of Computer Science and Engineering, Independent University, Bangladesh, Plot 16 Block B, Bashundhara R/A, Dhaka 1229, Bangladesh*

Correspondence should be addressed to Md. Kafiul Islam; kafiul_islam@iub.edu.bd

Received 23 August 2022; Revised 22 January 2023; Accepted 2 February 2023; Published 1 March 2023

Academic Editor: Ghous Ali

Copyright © 2023 Injamamul Haque Sourov et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Neuromarketing is a modern marketing research technique whereby consumers' behavior is analyzed using neuroscientific approaches. In this work, an EEG database of consumers' responses to image advertisements was created, processed, and studied with the goal of building predictive models that can classify the consumers' preference based on their EEG data. Several types of analysis were performed using three classifier algorithms, namely, SVM, KNN, and NN pattern recognition. The maximum accuracy and sensitivity values are reported to be 75.7% and 95.8%, respectively, for the female subjects and the KNN classifier. In addition, the frontal region electrodes yielded the best selective channel performance. Finally, conforming to the obtained results, the KNN classifier is deemed best for preference classification problems. The newly created dataset and the results derived from it will help research communities conduct further studies in neuromarketing.

1. Introduction

Neuromarketing, a relatively new consumer behavior and marketing research technique, is used to understand the consumer mind, both conscious and subconscious, to find explicitly what triggers the "buy" button in the mind. The visual and auditory aspects of a product or marketing advertisement that influence preference making are the key focus of neuromarketing research. Researchers are interested in finding out what regions of the brain are responsible for various stimuli, how they respond, and how conscious reasoning affects the decision-making process. Neuromarketing integrates neuroscience with marketing using non-invasive modern technologies. This is done purely through the ethical use of neuroscientific approaches, where consent forms that address the materials and their usage are signed by participants to avoid any legal issues. A common tool used in neuromarketing is the electroencephalogram (EEG) for raw brain signal acquisition. Digital signal processing techniques (artifact removal, filtering, data labeling, etc.) are applied to the signal to remove irrelevant or unwanted signal components and split the data into meaningful sections or data points. Features of statistical significance are generated using the data points to train and test the classification machine learning model. The model is then used to predict consumer preferences to successfully design and produce products that sell more.

With time, neuromarketing has been adopted into several marketing campaigns of tech giants (Google, Amazon, PayPal, etc.), automobile industries (Hyundai, etc.), food commercials, etc. This technique is expected to grow more accurate in terms of understanding or classifying consumer preference (on visual-auditory aspects or merchandise pricing, packaging, and advertising) and is also

expected to lead to new heights and applications due to its enormous potential. An example of such an application would be websites where neuromarketing is used to understand and make the site more user friendly as well as visually captivating.

Figure 1 visualizes how the incorporation of research with marketing paved the way for neuromarketing. Research into pupil dilation [1, 2] while watching advertisements began in the 1970s to address consumer behavior. In the 1980s, a series of studies [3–5] were conducted at the University of Wisconsin-Madison that investigated what patterns of EEG signals emerged while watching commercials. An article in Forbes highlighted several studies [6–8] in September of 2004. This article received significant media attention, which caused the scientific community to become familiar with such concepts, inciting inspiration for future research. In 2004, a rather interesting and famous research study, the "Coke vs. Pepsi" study [9], was conducted. Here, the authors found that branding played an enormous part in determining consumer preference for soft drinks and found that the absence of branding decreased the lure of the soft drinks.

Companies evaluate pictures, words, sound effects, and music to produce branding campaigns that are more appealing to clients [10]. In [11], the authors suggested a fast Fourier transform-based model to identify the most favored automobile brand in 4 criteria. For this purpose, EEG signal of 12 participants was collected while they surveyed the advertisement videos. The preprocessing stage consisted of two filters (Butterworth and Laplacian), and a total of three statistical features from the alpha rhythm were extracted. These were the power spectral density, spectral energy, and spectral centroid, respectively. Finally, the subject intention for a brand was computed using K-nearest neighbor (KNN) and probabilistic neural network (PNN) classifiers where a maximum accuracy of 96.62% was recorded using the PSD feature with the PNN classifier.

Furthermore, in [12], the authors analyzed the EEG data of 12 males and 12 females using a classifier (2-class preference). They obtained high classifier accuracies of 84.82% and 89.36% using KNN and SVM, respectively. However, neural networks such as DNN had a maximum accuracy of 79.76%. In addition, Oon et al. [13] used an eegoSports device to collect EEG data from a group of five males and five females between ages 19 and 23. The subjects were each presented with a video comprised of four categories of merchandise, with five advertisement videos in each category, and asked for their preference. After using KNN classifier to analyze the 2-class data, the authors produced an 80% classifier accuracy using alpha band feature, 76.18% using beta band, and 72.39% using both alpha and beta as features.

The main motivation behind pursuing this research on neuromarketing is that it creates an efficient model for the advertisers and marketers, removing their dependency to rely on consumers to both accurately and willingly report their emotions on a product. In addition, the method can closely tie the physiological reactions to specific parts of the advertisement and provide insight into the automatic responses that take place at the subconscious level. These benefits shall allow the marketing industry to potentially save a significant amount of capital as well as reduce its environmental footprint that goes into making the ad.

The absence of sufficient research contents due to its age and the limited number of datasets available makes this field opportune for contributions. In addition, the results obtained from this newly created dataset can prove fruitful in contributing knowledge to this field. Furthermore, a study of the accuracies and sensitivities obtained for the several statistical features of EEG signals can help suggest the most relevant one for predicting user preference.

With the goal of building predictive models that can identify the consumers' preference only based on their EEG data, the objectives of this research were set as follows:

(i) Apply different classifier algorithms on selected dataset.

(ii) Analyze subject-wise and overall performance. Analyze selective channel performance.

(iii) Analyze predictive models for both genders.

(iv) Suggest the best classifier algorithm among the ones worked with for predicting user preference.

2. Materials and Methods

The project commenced with the selection of a database for this project. For this purpose, the dataset from a like/dislike experiment for neuromarketing was acquired, with all due permissions, from Agency Lab (currently CCDS – Center for Computational and Data Sciences) under the Department of Computer Science and Engineering, Independent University, Bangladesh. A flowchart depicting the entire process is illustrated in Figure2. The collected (raw) EEG data for each subject were then manually labeled to obtain the individual samples, to which the usual BCI steps were applied to achieve the outcomes of this project. These steps included signal preprocessing, artifact removal, feature extraction and selection, and finally classification and performance evaluation. Figure 3 summarizes the experimental procedure described below.

2.1. Dataset

2.1.1. Data Acquisition. A total of 49 subjects (24 males and 25 females) between the ages of 19 and 28 (mean age = 21.15), with normal or corrected vision, volunteered for this experiment. They were each shown a total of 30 image advertisements collected from the Internet. The glimpse of advertisement as image stimuli is illustrated in Figure 4. These advertisements were equally divided into three sets as follows:

(i) Products suited towards the female gender (such as female shoes).

(ii) Products suited towards the male gender (such as male shoes).

(iii) Products that are thought of being gender independent (such as food).

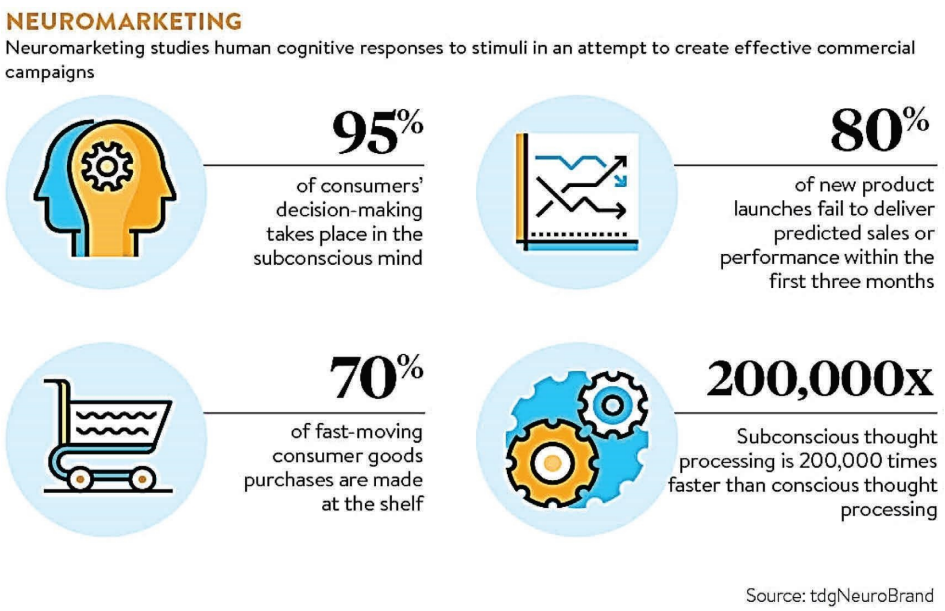


Figure 1: Neuromarketing, an attempt to generate effective campaigns (https://www.mauriziopittau.it/digital-marketing/neuromarketing-for-b2b).

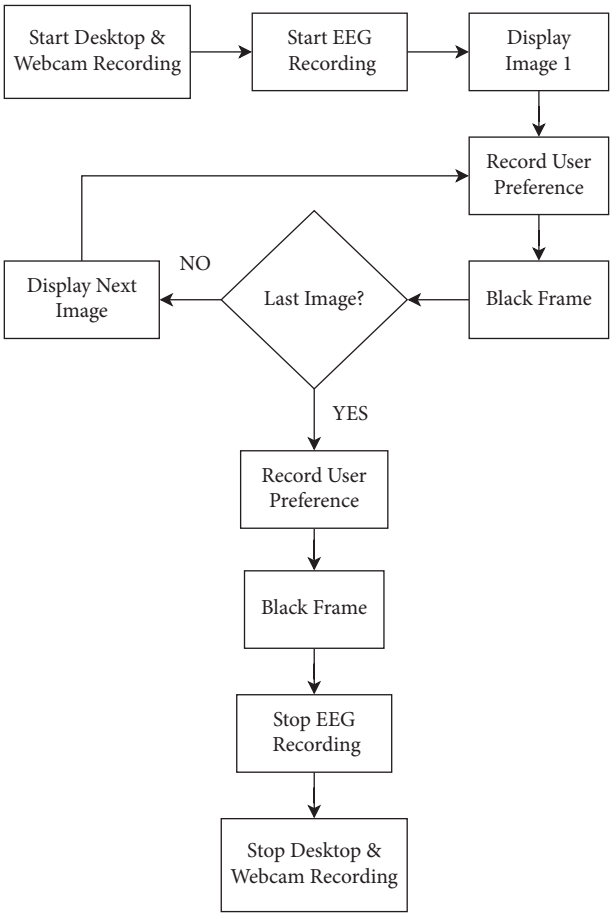


Figure 2: EEG data recording procedure.

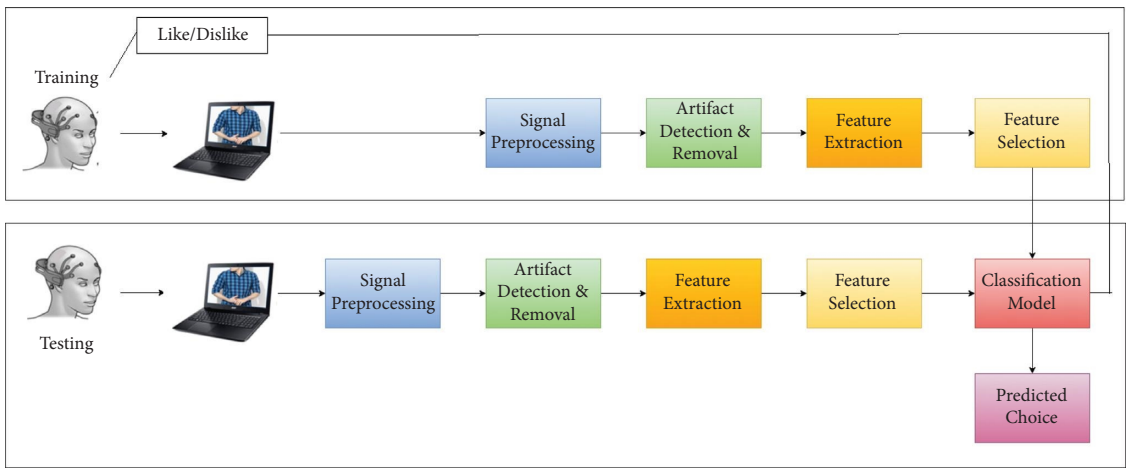


Figure 3: Experimental procedure.

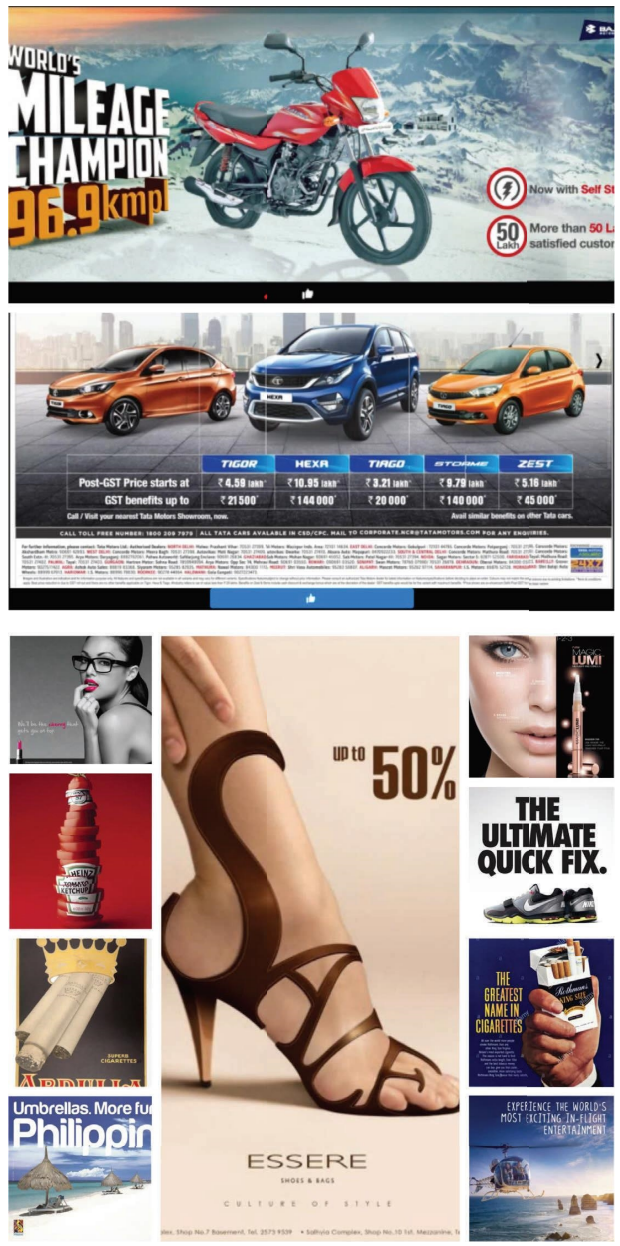


Figure 4: Examples of advert images displayed in this experiment.

The subjects were given the option to "like" the current image or skip to the next image at their convenience. Option "skip" is considered as "not liked" and thus "dislike." The experiment took on average a total of 6 minutes per subject with all the subjects being right-handed. Figure 2 illustrates how the EEG datasets were created.

A 14-channel EMOTIV EPOC+ wireless headset with sampling frequency of 128 Hz was used to record the raw EEE signals following the international standard 10–20 system of electrode montage.

The experiment data for each subject were obtained in 3 forms: (a) a CSV file that contained the raw EEG signals from all 14 channels, (b) a desktop recording that shows the time for which the images were shown and the user's preference, and (c) a webcam video of the subject interacting with the stimulus (images).

2.1.2. Dataset Labeling. Each column of the EEG file (CSV), starting in row two, represents the data collected over time. Only a few were significant to us and are discussed as follows. The first column, "COUNTER," represented the packet counter and was used as a time base. The second column stood for the "INTERPOLATED" flag, which showed if a packet was dropped (value = 1). FLAG = 0 meant the sample was good. The 14 consecutive columns labeled AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 represented the EEG channels associated with the headset, showing the voltage fluctuations over time.

The desktop recordings were used to identify the starting and ending times of a single image. These timestamps were then used to calculate the corresponding start and end row numbers in the EEG file (using the sampling frequency). Finally, the data sample was imported into the MATLAB workspace and stored alongside the user's preference for that image. A block diagram illustrating this labeling process for a particular subject is shown in Figure 5. For simplicity, the first two and the last images are shown, while the intermediate ones have been curtailed.

This process was repeated for every subject for labeling all raw EEG signals to get the total collection of approximately 1400 (neglecting 70 missing responses) feedback responses as likes/dislikes and their corresponding EEG data segments.

The uniform and non-overlapping 2 seconds of data segment for each feedback response (like/dislike) are taken into account for preference classification.

2.2. Signal Preprocessing. The acquired EEG signal is usually contaminated with noise and artifacts generated from different sources. The objective of signal preprocessing is to eliminate (or at least attenuate) these noises and artifacts from the signal. A simple and most used technique to remove both noise and artifacts is filtering. These filters are designed in such a way as to not introduce any change or distortion to the signal [14]. The Butterworth design in particular offers a maximally flat

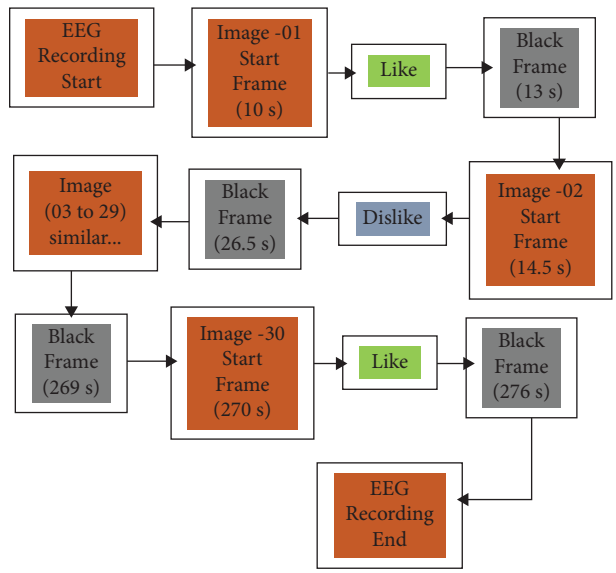


Figure 5: EEG raw data labeling procedure.

magnitude response along with a slight delayed phase response. As they have a more linear phase response for the passband when compared to Bessel and elliptic filters, they should offer better delay characteristics along with a smaller overshoot. To obtain a sharp transition in the frequency domain and avoid the generation of excess artifacts due to passband ripples, an IIR filter should be used. The authors in [15] have shown that IIR filters surpass FIR filters with a flat response characteristic as well as smaller response time.

In this project, the DC offset of the raw EEG signal was removed by subtracting the mean value of each signal from the signal itself. A high-pass 6^{th} order filter with a cutoff frequency of 0.5 Hz (produces best results after several trials) was then used to remove very low frequency components (noise or artifacts such as breathing noise). The EEG signals were downsampled to 128 Hz, and hence following the Nyquist theorem, the maximum frequency component present in the signal is 64 Hz. As a result, an additional low-pass filter was not required to remove high-frequency noises. Finally, a notch filter (very narrow band-stop filter) with a null frequency of 50 Hz was used to eliminate or at least attenuate the 50 Hz power line interference noise. The magnitude and phase responses, in MATLAB, of the high-pass filter (0.5 Hz) and the notch filter (49 Hz–51 Hz) are shown in Figures 6 and 7, respectively.

Artifacts are unwanted electrical signals with origins other than the brain itself. These unwanted signal components are captured by the EEG device along with the brain signal itself. The two chief categories of artifact are physiological/biological or non-physiological (referred to as "noise") artifacts. Physiological artifacts take various forms including cardiac pulse, respiratory, eye blink (as well as lateral eye movement), and muscle movement artifacts [16]. In this experiment, a high-pass filter was used to remove any sub-hertz frequency artifacts such as the breathing noise. In

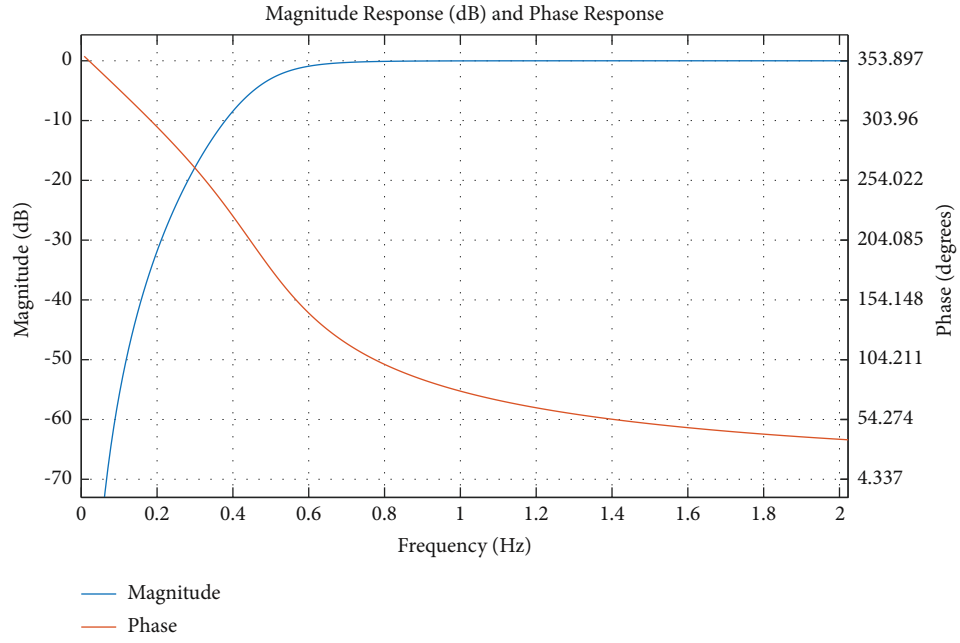


Figure 6: Magnitude and phase response of 0.5 Hz HPF.

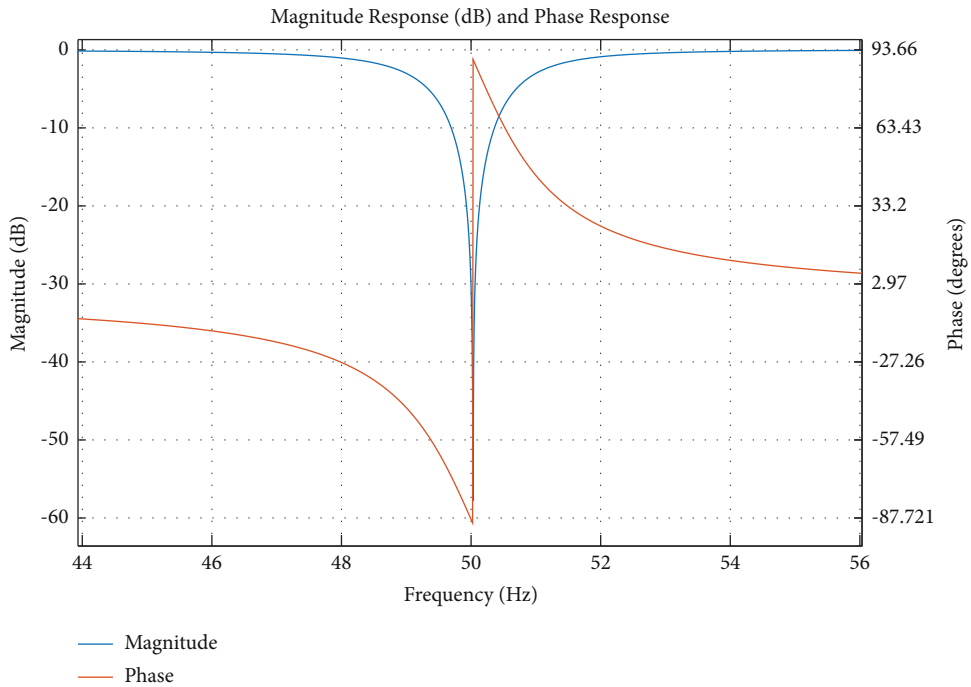


Figure 7: Magnitude and phase response of band-stop/notch filter (49 Hz–51 Hz).

addition, a visual inspection of the raw signals for each channel was performed to estimate the contamination levels and decide whether to use the channel.

2.3. Feature Extraction. A total of 27 statistical common features, based on literature reviews, were extracted. The maximum accuracy across all subjects was obtained by selecting 18 of the 27 features extracted. A selected few, that produced good results, are discussed as follows. These include a general description as well as their MATLAB implementations or interpretations.

2.3.1. Skewness. It is a measure of the deviation of a distribution, assuming a unimodal probability distribution of a random variable (real number), from its mean or symmetrical bell curve of the normal distribution. When the flag is set to 1 on MATLAB, the skewness is biased, and the following equation applies:

$$s_1 = \frac{1/n\sum_{i=1}^{n}(x_i - \overline{x})^3}{\left(\sqrt{1/n\sum_{i=1}^{n}(x_i - \overline{x})^2}\right)^3}. \quad (1)$$

2.3.2. Kurtosis. It is a measure of how susceptible a probability distribution is to its outliers; in other words, kurtosis determines the heaviness of the distribution tails. If the flag is configured to be 1 on MATLAB, then the kurtosis is biased, giving the following equation:

$$k_1 = \frac{1/n\sum_{i=1}^{n}(x_i - \overline{x})^4}{\left(1/n\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^2}. \quad (2)$$

2.3.3. Hjorth Parameters. Hjorth Activity, Hjorth Mobility, and Hjorth Complexity are all normalized slope descriptors (NSDs) used for analyzing non-stationary signals of irregular textures.

(1) Hjorth Activity. It signifies a signal's power density. In frequency domain, it represents the power spectrum surface, according to Parseval's theorem. The Hjorth Activity is usually estimated by the following expression:

$$\begin{aligned}\text{Hjorth Activity} &= \text{var}(x(t)) = \frac{1}{N-1}\\ &= \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2.\end{aligned} \quad (3)$$

(2) Hjorth Mobility. It evaluates the approximate mean frequency. The Hjorth Mobility is usually estimated by the following expression:

$$\text{Hjorth Mobility} = \sqrt{\frac{\text{var}(\mathrm{d}x(t)/\mathrm{d}t)}{\text{var}\ (x(t))}}. \quad (4)$$

(3) Hjorth Complexity. It evaluates the approximate frequency spectrum of the signal, which calculates the correlation between a sine wave and the signal. The Hjorth Complexity is usually estimated by the following expression:

$$\text{Hjorth Complexity} = \frac{\text{Mobility}(\mathrm{d}y(t)/\dot{\mathrm{d}}t)}{\text{Mobility}(y(t))}. \quad (5)$$

2.3.4. Wentropy (y, "norm," 1.1). It first converts the signal into wavelet packets indexed by position, scale (as in wavelet packet decomposition), and frequency. With power set to 1.1, the norm entropy is calculated as follows:

$$\text{norm entropy of } x(s) = \sum_i |s_i|^p = \|s\|_p^p, \quad (6)$$

where s is the signal, $(s_i)_i$ refers to the orthonormal coefficients of s, and $1 \le p$. Entropy is the average level of information or uncertainty that a random variable contains inherently in its possible outcomes.

2.4. Feature Selection. To select relevant features for classification, the obtained feature data were first sorted for preference (like/dislike) and then several statistical graphs such as box and whisker plot, scatter plot, and area plot for were visually inspected for each feature. Skewness and entropy features were the best differentiators among 27 extracted statistical features. Figures 8 and 9 demonstrate the distinguishing characteristics of the skewness and entropy features for preferences. For other features, similar comparison plots have been analyzed and the most distinguishable features have been selected such as skewness, entropy, kurtosis, Hjorth parameters, and band powers.

2.5. Classifiers and Classification. Since this research worked with a binary class problem, like and dislike, several classifier algorithms, commonly used in binary classifications, were investigated, and among them, 3 were chosen as follows: (i) support vector machine (SVM) creates a line or a hyperplane (decision boundary) that differentiates the dataset into two distinct classes; (ii) K-nearest neighbor (KNN) creates a distance function based on the K-value and separates data into two classes based on the distance; and (iii) neural network (Pattern Recognition App of MATLAB) applies random weight to nodes between each layer, and an activation function is designed to provide output that separates the data into two classes.

The feature data are split into training data, validation data, and testing data in a ratio that relies on the availability of data and the application once it has been preprocessed and is ready for training. This is done to guarantee that the model does not "overfit" or "underfit" excessively and performs equally well in the real world. This was done on classification-based MATLAB apps used, where the training-to-testing ratio was decided based on the several trials performed to determine the best ratio.

The training data are fed into the model, which are then used to train numerous classification models in the Classification Learner App in MATLAB. Their performances are compared using the validation set, which is then used to optimize the model's hyperparameters, retrain and re-evaluate the models, and consequently choose the superior model parameters. Finally, the model is employed to predict on the test dataset, yielding a final test score that more or less properly describes the model's performance on the provided dataset.

3. Results

3.1. Performance Evaluation. Ideally, a precision of 1 and a sensitivity of 1 are desired in a classifier model, implying 100% accuracy, which is not always the case for a machine learning model. So, we should aim for a high precision with a much greater sensitivity value. For this project, a greater

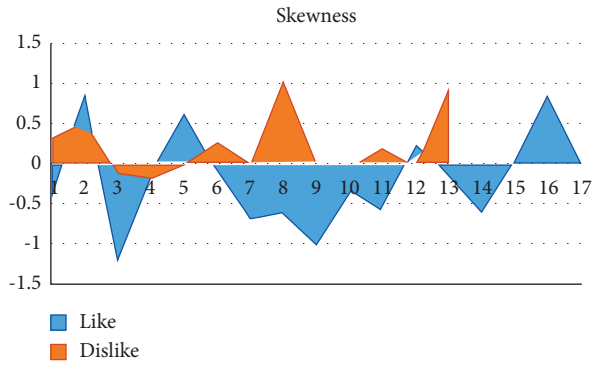


Figure 8: Comparison of skewness values for like and dislike.

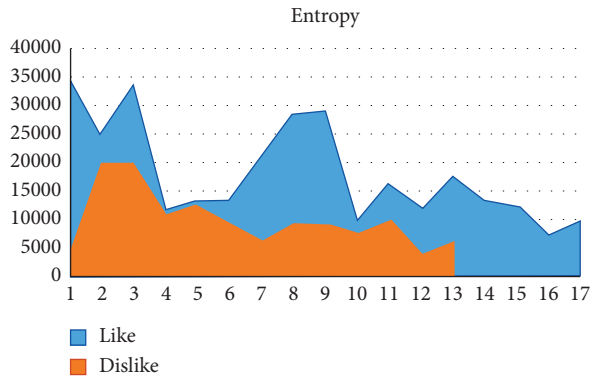


Figure 9: Comparison of entropy values for like and dislike.

significance has been attached to sensitivity before accuracy because in marketing research, the identification of a like is more important than the identification of a dislike. In addition, since the number of liked and disliked sample data points was slightly imbalanced, accuracy carries less weight when compared to sensitivity. The performance parameters are calculated using Figure 10 as follows.

The accuracy of classifier is the percentage of total samples correctly classified by the classifier, while the classification error is obtained by subtracting the accuracy from 1.

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad \text{Classification Error} = 1 - \text{ACC}. \tag{7}$$

The classifier sensitivity (true positive rate) calculates the percentage of positives that are accurately classified.

$$\text{TPR} = \frac{\text{TP}}{\text{P}} = \frac{\text{TP}}{\text{TP} + \text{FN}} = 1 - \text{FNR}. \tag{8}$$

Specificity (true negative rate) calculates the percentage of correctly identified negatives.

$$\text{TNR} = \frac{\text{TN}}{\text{N}} = \frac{\text{TN}}{\text{TN} + \text{FP}} = 1 - \text{FPR}. \tag{9}$$

Precision (positive predictive value) calculates the percentage of correctly identified positives.

		Actual Class	
		Positive	Negative
Predicted Class	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

Figure 10: Confusion matrix.

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} = 1 - \text{FDR}. \tag{10}$$

Each table that follows has one row italicized to highlight the best results.

3.2. Individual Subject Accuracies. Out of the 49 subjects in this experiment, only 35 subjects were finally selected on in this work, due to several reasons: (i) noisy or blank channels (showing no voltage fluctuations), (ii) software or hardware malfunctions, (iii) human errors as detected in the recordings (webcam and screen recording), and (iv) imbalanced dataset. The data on these subjects including their gender, age, number of likes/dislikes, and the corresponding accuracies are summarized in Table 1.

Only seven subjects had an accuracy rate that exceeds 80%. However, the high accuracy can be deceiving due to a lack of balance between "like" data points and "dislike" data points. For example, subject 25 had a classifier accuracy of 92.6%. However, the subject dataset contained 25 "like" data points, whereas, it only had 4 "dislike" data points. This indicates that the classifier lacked sufficient information due to the imbalanced dataset and most likely produced a biased classification result.

3.3. Gender-Wise Classification. Comparing the gender-wise performance as shown in Table 2, it is seen that the performance measures for the female subjects were comparably higher than those for the males. In addition, the SVM classifier had the highest accuracy and sensitivity for males, at 70.7% and 83.3%, respectively, whereas KNN had the highest accuracy and sensitivity for females at 75.7% and 95.8%, respectively.

3.4. Classification of Selective Channel Configurations. Table 3 compares the selective channel performance across all subjects. For the selected channel analysis, the highest sensitivities were seen for the frontal channels, but the highest accuracy was obtained for the frontal and occipital region for both classifiers. These results agree with the previously suggested theory that frontal regional activity is the most prominent activity for the decision-making process.

3.5. Overall Classification Performance across All Subjects. Table 4 compares the overall (male and female combined, as well as all 14 channels) performance measures obtained

Table 1: Subject-wise classification accuracy using SVM classifier.

Subject no	Gender	Age	Like data points	Dislike data points	Accuracy (%)
1	Male	23	19	9	67.9
2	Female	23	13	15	57
3	Male	21	15	13	64.3
4	Male	26	16	12	82.1
6	Female	22	25	3	89.3
7	Male	22	6	22	78.6
8	Male	20	17	11	71.4
9	Male	23	19	9	78.6
10	Male	21	20	8	75
13	Female	23	19	9	85.7
14	Male	24	16	12	67.9
15	Female	21	12	16	64.3
16	Female	22	10	18	75
17	Female	22	21	7	75
18	Male	23	19	9	67.9
19	Female	22	20	8	71.4
20	Female	19	16	12	57.1
22	Male	28	12	16	57.1
23	Female	22	17	11	60.7
24	Male	23	22	6	78.6
25	Female	27	25	2	92.6
26	Female	21	24	4	92.9
27	Male	23	13	15	71.4
28	Female	21	17	11	71.4
29	Female	21	16	12	64.3
30	Female	22	19	9	67.9
31	Male	19	20	8	71.4
37	Female	21	21	7	89.3
38	Male	23	13	15	78.6
40	Female	22	14	14	53.6
42	Male	21	16	12	67.9
43	Male	22	23	5	82.1
44	Male	21	11	17	71.4
45	Male	22	13	15	57.1
48	Male	22	13	15	64.3

Table 2: Gender-wise classification performance.

Gender	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Male	*SVM*	*70.7*	**81.33**	*56.90*	*70.93*
	KNN	69.2	76.32	59.65	71.60
	NN	68.2	76.57	57.21	70.30
Female	SVM	73.0	84.72	51.28	76.25
	KNN	*75.7*	**95.83**	*38.46*	*74.19*
	NN	73.4	91.35	40.51	73.74

Bold values represent the highest sensitivity values for both males and females.

Table 3: Selective channel classification performance.

Conf.	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Frontal (LR)	*SVM*	*67.6*	**85.81**	*39.58*	*68.65*
	KNN	*64.3*	**95.27**	*16.67*	*63.8*
Frontal and occipital	SVM	70.1	88.51	41.67	70.05
	KNN	67.6	81.08	46.88	70.18
Frontal and temporal	SVM	68.9	91.22	34.375	68.18
	KNN	70.5	82.43	52.08	72.62
Left hemisphere	SVM	67.6	87.16	37.5	68.25
	KNN	66.8	87.16	35.42	67.54

Bold values represent the two highest sensitivity values.

Table 4: Overall classification performance across all subjects.

Classifier	Configuration	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
SVM	Optimized with 150 iterations	74.2	89.19	51.04	73.74
KNN	*Optimized with 150 iterations*	*74.6*	**92.57**	*46.88*	*72.87*
NN pattern recognition	10 hidden layers and 2 output layers with retraining of 30 times	71.0	85.98	48.06	71.69

Bold values represent the highest sensitivity value.

across all subjects for all classifiers. Here, it was seen that all classifiers performed relatively well and around about the same region; however, KNN classifier had the highest accuracy and sensitivity at 74.6% and 92.57%, respectively. The optimizable SVM and KNN classifiers, on MATLAB Classification Learner App, were configured for 150 iterations, and the optimal hyperparameter was chosen to produce the best results. In contrast, the neural network pattern recognition app was configured with 10 hidden layers and 2 output layers that had been retrained a minimum of 30 times to produce the best possible results as shown below.

3.6. Confusion Matrix and Minimum Classification Error Plot for Overall Classification Performance across All Subjects. For overall performance across all subjects, the confusion matrixes are shown below for each classifier.

3.6.1. SVM Classifier. The performance matrix is shown in Figure 11 where 25% holdout validation is shown and the remaining 75% feature matrix was used to train the optimizable SVM model. Totally, 181 preferences were correctly identified out of 244 samples while validating the model, and remarkable sensitivity was found, though the classification accuracy was around 74% which is demonstrated in Table 4.

3.6.2. KNN Classifier. Similarly, the performance matrix is shown for optimizable KNN classifier in Figure 12 where similar 25% holdout method for validation was used for calculating performance parameters of the model. KNN identified one more "like" than SVM, and even the sensitivity was bit higher, but it could not identify "dislike" prominently.

3.6.3. Neural Net Pattern Recognition App. In Figure 13, the performance matrix is shown for pattern net recognition-based neural network where 15% feature matrix is used for validation, 15% is used for testing, and the remaining 70% is used for training set.

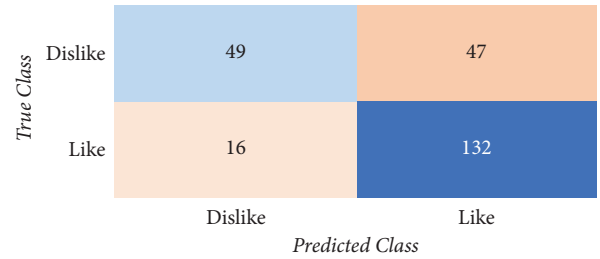


Figure 11: Confusion matrix for optimizable SVM.

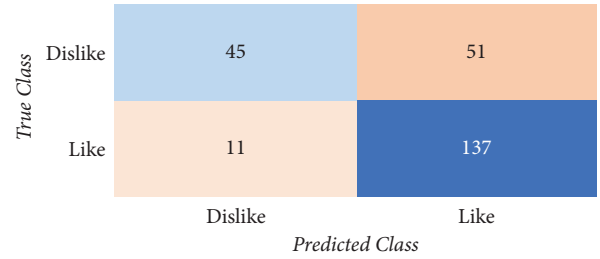


Figure 12: Confusion matrix for optimizable KNN.



Figure 13: Confusion matrix for optimizable neural net pattern recognition app.

4. Discussion

The performance measures assessed for each classifier were accuracy, sensitivity, specificity, and precision. As explained before, accuracy is the percentage of total samples that are correctly predicted. The sensitivity measure refers to the percentage of positives that are correctly predicted, whereas specificity is the percentage of negatives that are correctly predicted. For this research, a greater significance has been attached to sensitivity and then accuracy because in marketing research, the identification of a like is more important than a dislike.

The accuracies obtained with SVM, KNN, and neural network pattern recognition classifier algorithm model are around 71% to 75%. These accuracies roughly translate to seven correct predictions out of every ten predictions. When compared with the large dataset (of subjects with diverse backgrounds), the classifier performed rather well and can be considered reliable. However, sensitivity and precision are two significant parameters that should be taken into consideration for performance evaluation as explained

Table 5: Comparison of proposed work with few existing methods.

Ref	Dataset description	Classification algorithm	Channels	Rhythm	Classes	Best accuracy (%)
[20]	DEAP dataset	DTL (fine-tuning)	AF3, AF4, F3, F4, and Fz	—	Like/dislike	93
[19]	DEAP dataset	DNN	AF3, AF4, F3, F4, and Fz	—	Pleasant/ unpleasant	94
[17]	B-alert X10 headset Sampling frequency = 256 Hz (16 bits resolution) 10–20 electrode placement system 3 females and 2 males 22–26 (mean age = 22.8)	SVM	F3, F4	Alpha, theta, beta	Like/dislike	75
		KNN	Fz, F3, F4	Alpha, theta, delta		80
[18]	EMOTIV Epoc 14-channel 10–20 system Sampling rate = 128 (2048 Hz internally) (6 men, 5 women, age: 25–39) 53 image stimuli	SVM	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4	Alpha, beta, theta, gamma	Preferred/ unnoticed image	75.44
			F8, T8	Alpha		83.64
[22]	Neurosky-mindband Sampling rate = 256 Hz 12 subjects	SVM + spectral filter	FP1, FP2	Delta, theta, alpha, beta, gamma	Like/dislike	68.22
		SVM + CFP				74.77
[23]	ABM B-alert X10 Sampling rate = 256 Hz 8 female and 8 males, mean age = 22.44	DNN	POz, Fz, Cz, C3, C4, F3, F4, P3, P4	Delta, theta, alpha, beta, gamma	Like/dislike	63.99
		SVM				60.19
		KNN ($K = 5$)				56.29
[24]	EMOTIV Epoc+ 14 channels 2048 Hz downsampled to 128 Hz 25 participants Age: 18–38 42 different product images	DNN (2 layers + sigmoid)	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4	Delta, theta, alpha, beta, gamma	Like/dislike	60.10
		SVM				62.85
		RF				68.41
		HMM				70.33
This paper	*49 subjects (24 males and 25 females) 19–28 (mean age = 21.15) 30 image advertisements EMOTIV Epoc+ 14 channels 2048 Hz downsampled to 128 Hz International 10–20 system*	*SVM*	*AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4*	*Delta, theta, alpha, beta, gamma*	*Like/dislike*	*74.2*
		KNN				*74.6*
		Neural net pattern recognition				*71.0*

earlier. As a result, increased precision with a higher sensitivity value for the model is also extremely desirable. The sensitivity and precision of each classifier are as follows: 89.19% and 73.74%, 92.57% and 72.87%, and 85.98% and 71.69%. Each sensitivity parameter value is generally very high and makes up for the precision parameter that lacks behind. These performance parameters were obtained after 150 iterations. The pattern recognition neural network took a great deal of time to train and retrain several times to produce the given accuracy.

A performance comparison between few selected published works and our results is tabulated in Table 5. Though there have been better performances found in [17, 18], our model is quite convenient in performance for such own recorded big database in consideration compared to their databases. As the dataset is different, comparison between our model and deep learning based models [19, 20] is quite inconvenient. Still, few comparisons are shown in Table 5.

The innovation of this study is that the database is original, large, and self-recorded and EMOTIV 14-channel wireless device is used where the 128 Hz downsampled EEG signal is used for neuromarketing application. Our result is quite prominent for such large dataset with lowest resolution which establishes that this type of experiment can be useful for neuromarketing research and application. Even gender-based comparisons were shown that imply whether any product discrimination based on gender affects neuromarketing or not. Moreover, the database itself is a contribution of this study and getting prominent results for such large database with lowest quality of EEG signal is another major contribution of this work. Finally, another contribution of this work is the analysis of brain region based channel selection and its performance comparison for data recorded with a low number of channels (i.e., 14 here) compared to 32, 64, or even more channels .

5. Conclusion

This article presents a detailed study on the use of EEG signals for preference classification in neuromarketing. The raw EEG signal extracted from the experiment was labeled manually, where the individual images were tagged with their corresponding EEG signal durations. Several faulty datasets were identified and removed due to software glitch or hardware errors. A filter was applied to remove the DC offsets and artifacts, and from these, 18 out of 27 statistical features were used on the machine learning algorithms, and their corresponding performance parameters were analyzed.

From the three classifier algorithms, SVM had slightly better selective channel performance while KNN had a slightly better overall and gender-wise performance. Again, very high sensitivity values were obtained in nearly all classifications performed, where the minimum was around 86% and the maximum was around 95.8%. In addition, the accuracies indicate that both SVM (74.2%) and KNN (74.6%) are good classifiers for binary classifications of highly imbalanced dataset, with the latter being the superior one. Neural net pattern recognition performed roughly the same (71.0%).

The subject-wise results showed great variations where the highest individual accuracy was 92.9% and the lowest was 57%. The remaining subject varied between 64% and 82%.

The gender-wise performance showed that the performance measures for the female subjects were comparably higher than those for the males. SVM had the highest accuracy and sensitivity for males, at 70.7 and 83.3%, respectively, whereas KNN had the highest accuracy and sensitivity for females at 75.7% and 95.8%, respectively.

Finally, the results obtained from selective channel analysis agree with the previously suggested theory that frontal regional activity is the most prominent activity for the decision-making process.

The large imbalance in the number of like and dislike samples and the inconsistent length of individual samples in the dataset were the primary limitations of this work. The huge discrepancies in data length adversely affected the statistical features extracted and limited the classifier performance.

For future works, the use of dataset with consistent control parameters such as the data length and age group can be used to enhance reliability and consistency. In addition, advanced signal processing techniques such as wavelet transform and additional sets of different features can be used for classification. The models of deep learning algorithms are unexplored as this is the first attempt on database and only machine learning algorithms are intentionally explored in this research project. So, in the future work, deep learning algorithms such as convolutional neural network (CNN) and other popular models as well as further channel selection analysis can be used to investigate their effect on the classifier performance.

Data Availability

The recorded EEG data are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Injamamul Haque Sourov and Faiyaz Alvi Ahmed contributed equally to this study.

Acknowledgments

This research was sponsored by IUB Sponsored Research Grant (#2021-SETS-07) and supported by Biomedical Instrumentation and Signal Processing Lab (BISPL) of Department of EEE, IUB, and Center for Computational and Data Sciences of Department of CSE, IUB.

References

[1] B. Laeng, S. Sirois, and G. Gredebäck, "Pupillometry: a window to the preconscious?" *Perspectives on Psychological Science*, vol. 7, no. 1, pp. 18–27, 2012.

[2] D. Kahneman and J. Beatty, "Pupil diameter and load on memory," *Science*, vol. 154, no. 3756, pp. 1583–1585, 1966, http://www.jstor.org/stable/1720478.

[3] M. L. Rothschild, E. Thorson, B. Reeves, J. E. Hirsch, and R. Goldstein, "EEG activity and the processing of television commercials," *Communication Research*, vol. 13, no. 2, pp. 182–220, 1986.

[4] M. L. Rothschild, Y. J. Hyun, B. Reeves, E. Thorson, and R. Goldstein, "Hemispherically lateralized EEG as a response to television commercials," *Journal of Consumer Research*, vol. 15, no. 2, pp. 185–198, 1988.

[5] M. L. Rothschild and Y. J. Hyun, "Predicting memory for components of TV commercials from EEG," *Journal of Consumer Research*, vol. 16, no. 4, pp. 472–478, 1990.

[6] S. Erk, M. Spitzer, A. P. Wunderlich, L. Galley, and H. Walter, "Cultural objects modulate reward circuitry," *NeuroReport*, vol. 13, no. 18, pp. 2499–2503, 2002.

[7] J. R. Rossiter, R. B. Silberstein, P. G. Harris, and G. Nield, "Brain-imaging detection of visual scene encoding in long-term memory for TV commercials," *Journal of Advertising Research*, vol. 41, no. 2, pp. 13–21, 2001.

[8] T. Ambler, S. Braeutigam, J. Stins, S. Rose, and S. Swithenby, "Salience and choice: neural correlates of shopping decisions," *Psychology and Marketing*, vol. 21, no. 4, pp. 247–261, 2004.

[9] S. M. McClure, J. Li, D. Tomlin, K. S. Cypert, L. M. Montague, and P. Montague, "Neural correlates of behavioral preference for culturally familiar drinks," *Neuron*, vol. 44, no. 2, pp. 379–387, 2004.

[10] D. L. Fugate, "Marketing services more effectively with neuromarketing research: a look into the future," *Journal of Services Marketing*, vol. 22, no. 2, pp. 170–173, 2008.

[11] M. Murugappan, S. Murugappan, Balaganapathy, and C. Gerard, "Wireless EEG signals based neuromarketing system using Fast fourier transform (FFT)," in *Proceedings of the 2014 IEEE 10th International Colloquium on Signal Processing and its Applications*, pp. 25–30, Kuala Lumpur, Malaysia, March 2014.

[12] J. Teo, L. Hou, J. Tian, and J. Mountstephens, "Classification of affective states via EEG and deep learning," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 5, 2018.

[13] H. N. Oon, A. Saidatul, and Z. Ibrahim, "Analysis on nonlinear features of electroencephalogram (EEG) signal for neuromarketing application," in *Proceedings of the 2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, pp. 1–8, Kuching, Malaysia, August 2018.

[14] M. Agustina Garcés and L. L. Orosco, P. Diez, Chapter 5 - EEG signal processing in brain–computer interface," *Smart Wheelchairs and Brain-Computer Interfaces*, Academic Press, Cambridge, CA, USA, pp. 95–110, 2008.

[15] A. Milchevski and M. Gusev, "Performance evaluation of FIR and IIR filtering of ECG signals," in *ICT Innovations 2016*, G. Stojanov and A. Kulakov, Eds., vol. 665, Berlin, Germany, Springer, 2018.

[16] J. W. Britton, L. C. Frey, and J. L. Hopp, "Appendix 4. Common artifacts during EEG recording," 2016, https://www.ncbi.nlm.nih.gov/books/NBK390358/.

[17] L. H. Chew, J. Teo, and J. Mountstephens, "Aesthetic preference recognition of 3D shapes using EEG," *Cognitive neurodynamics*, vol. 10, no. 2, pp. 165–173, 2016.

[18] Y. Kim, K. Kang, H. Lee, and C. Bae, "Preference measurement using user response electroencephalogram," *Computer Science and its Applications*, vol. 330, pp. 1315–1324, 2015.

[19] 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, 2020.

[20] M. S. Aldayel, M. Ykhlef, and A. N. Al-Nafjan, "Electroencephalogram-based preference prediction using deep transfer learning," *IEEE Access*, vol. 8, pp. 176818–176829, 2020.

[21] A. Özbeyaz, "EEG-Based classification of branded and unbranded stimuli associating with smartphone products: comparison of several machine learning algorithms," *Neural Computing & Applications*, vol. 33, no. 9, pp. 4579–4593, 2021.

[22] Y. Pan, C. Guan, J. Yu, K. K. Ang, and T. E. Chan, "Common frequency pattern for music preference identification using frontal EEG," in *Proceedings of the 2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, California, CA, USA, November 2013.

[23] J. Teo, C. L. Hou, and J. Mountstephens, "Deep learning for EEG-Based preference classification," *Applied Sciences*, vol. 10, 2017.

[24] M. Yadava, P. Kumar, R. Saini, P. P. Roy, and D. Prosad Dogra, "Analysis of EEG signals and its application to neuromarketing," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 19087–19111, 2017.