Consumer Behavior Analysis using EEG Signals for Neuromarketing Application

Chowdhury Rabith Amin‡, Mirza Farhan Hasin‡, Tasin Shafi Leon‡, Abrar Bareque Aurko‡,
Tasmi Tamanna†, Md Anisur Rahman*, Mohammad Zavid Parvez‡
*School of Computing & Mathematics, Charles Sturt University, Australia
†Department of Immunology, Bangladesh University of Health Sciences, Dhaka, Bangladesh
‡Department of Computer Science and Engineering,
Brac University, Bangladesh
Email: ‡{rabith07, mfhprantik, leontasinshafi, abrar.aurko}@gmail.com,
†tasmitamanna02@gmail.com, *arahman@csu.edu.au, ‡zavid.parvez@bracu.ac.bd

***Abstract*—Neuromarketing is applying neuropsychology in marketing research studying consumer sensory-motor actions such as cognitive and affective responses to marketing stimuli with the help of modern technologies. It is one of the most recent marketing research strategies and may become the future of marketing research. Many research works have been carried out in this area to obtain better outcomes. However, literature shows that there is an opportunity for further improvement. Hence, in this study, a model is presented using data mining and machine learning algorithms for consumer behavior analysis from EEG signals. Time-frequency distribution features are extracted from EEG signals on which different classification algorithms are applied. Consumer's responses toward marketing strategies and their behavior towards purchasing or selecting goods can be studied and analyzed to understand the producer-consumer relationship. EEG signals from 25 people are collected where the participants varied in age and gender for a better understanding of consumer behavior towards a marketing policy. By analyzing the data, the reason behind how and why they like certain marketing policies was uncovered. The performance of our proposed model with an existing technique is compared. The accuracy of our model on the dataset is 95%, whereas the accuracy of the existing technique on the same dataset is 70%. We also evaluated whether neuropsychological measures can capture differences in consumer's actions according to different marketing stimuli. The experimental results on our model indicate that studies in this field can bring a change and improve marketing strategies for the betterment of both the producer and the consumer, resulting in an eventual mutual benefit.**

***Index Terms*—Neuromarketing, EEG, Neuropsychology, Marketing Strategy, Consumer, Decision Tree.**

I. Introduction

Neuromarketing is the study of the brain's response marketing strategies using medical technologies to motivate marketing policies [1]. Technologies such as eye-tracking, facial coding, functional magnetic resonance imaging (fMRI), and EEG are used by researchers to compute distinct types of brain activity in response to advertised information. By analyzing the collected information, corporations try to figure out the reasons behind consumer's decisions. They try to understand the logic behind consumer's purchasing of particular commodities as well as which part of the brain influences them to do so.

Many research works have been devoted to the neuromarketing research application. In the article of Bartels [2], it is mentioned that the modern thoughts of developing marketing strategies began in the twentieth century, which evolved with time. Producers realized that the importance of studying the effects of marketing techniques on consumers' behavior is very crucial, and they started conducting research on the market as well as on consumers.

Researches conducted on the market-consumer area helped in understanding the buying behavior of consumers as well as developing advertisement and marketing policies. Most of the time, applying traditional marketing methods to different groups of people fails to yield similar effects. Manually conducted researches on consumers can be flooded with false and made-up information. Therefore, strategies developed based on researching consumers might not always be effective as people can lie or act accordingly.

However, it is possible to determine what a person is thinking through measuring and analyzing the brain waves from EEG signals. It is where neuropsychology comes handy. With the help of neuropsychological patterns in a person, action and behavior can be understood, and thus products and services can be developed according to their choices. Also, marketing strategies can be developed, leading to mutual benefits. Patterns recognized by analyzing neuropsychology can help in building neuromarketing strategies [3].

Another study associated EEG based communication of Blankertz *et al.* [4] where results are encouraging for an EEG-based BCI system in untrained subjects that is independent of peripheral nervous system activity and does not rely on evoked potentials even when compared to results with very well-trained subjects operating other BCI systems.

In a study Heekeren *et al.* [5] examined and demonstrated the activity of different brain parts for perceptual decision making. The prefrontal area performs in basic decision making, free of external stimuli and response modalities purposes. Additionally, the authors mentioned that the Dorsolateral Prefrontal Cortex hinders the decision-making process of mon-

978-1-7281-2547-3/20/$31.00 ©2020 IEEE

orized licensed use limited to: NUST School of Electrical Engineering and Computer Science (SEECS). Downloaded on November 15,2023 at 18:24:50 UTC from IEEE Xplore. Restrictions ap

keys and people. Others have mentioned that the responsibility of the prefrontal cortex is to manage movement along with task-applicable pathways. Authors[5] also exhibited a system for how perceptual basic leadership procedures may be elicited in the human mind by utilizing a moderately straightforward subtraction instrument. Using the knowledge gathered from the studies of the brain, the authors started to understand human behavior more accurately. Thus, they proposed a marketing model combined with knowledge of neuropsychology.

In an article, Smith *et al.* [6] investigated the feasibility of inferring the choices people would make based on their neural responses to the relevant prospects when they are not engaged in actual decision making. The authors achieved an overall 68.2% success rate for prediction by using statistical methods. However, the success rate of prediction is not sufficient enough to establish the feasibility of their approach.

In another article Telpaz *et al.* [7] proposed that neuroimaging technology can predict preferences for consumer products. As these medical machinery are expensive, the authors doubted their wide use. The authors analyzed neural measurements that offered a relatively low-cost and were widely available, i.e., EEG and their research demonstrated 79% precision by using cardinal NRUM in predicting consumers' future choice.

In an article, Mahendra*et al.*,[8] proposed a predictive modeling framework to understand consumer choice towards E-commerce products in terms of "likes" and "dislikes" by analyzing EEG signals. The prediction accuracy of the model was 70% with the help of the Hidden Markov Model (HMM) classifier. However, one of the limitations in this research area is to obtain better accuracy for predicting consumer behavior.

Therefore, in this paper, we presented a prediction model based on a decision tree (DT) on EEG signals that obtains high accuracy compared to the existing methods in this area. Moreover, we extracted features based on the time-frequency distribution of the EEG signals to get better accuracy. The motivation of this paper is to reduce the gap between humans and marketers by understanding brain science. The root to leaf path (i.e., logic rule) of a decision tree on consumer data will help a producer and advertiser to make suitable marketing strategies for a particular product.

We compared the performance of our proposed method with six existing techniques in terms of area under curve, accuracy, sensitivity, and specificity. The main contributions of the paper are as follows, feature extraction using time-frequency domain distribution from EEG signals due to its superiority in fast non-redundant transforms, and we used ROC curve to determine the important part of the brain to collect EEG data.

The rest of the paper is organized as follows. In Section II, we present our proposed method for consumer behavior analysis. In Section III, the experimental results and discussion are provided. Lastly, concluding remarks are presented in Section IV.

II. Proposed Method

Our proposed method is divided into four basic steps as shown below:

Step 1: Data Collection
Step 2: Data Pre-processing
Step 3: Feature Extraction
Step 4: Classification

In our short span of research work, we found out the most precious part of the research is the proposed method. The proposed method is shown in Fig 1. The first step in the method was to collect the EEG dataset. Then pre-processing is used before extracting features. Next, we divided the features into training and testing sets for classification. We generated a model based on the training set and then predict the class label for the testing set. Then we calculated performance metrics using actual and predicted class labels.

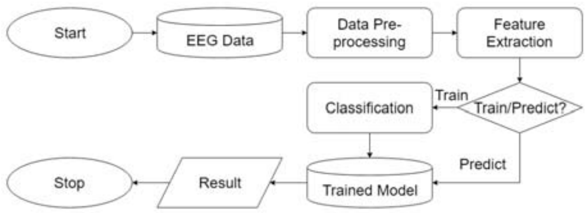

Fig. 1. A block diagram of the proposed method for consumer behaviour analysis.

A. Data Collection

The paper uses the recorded EEG dataset from 25 participants mentioned in [8]. It is publicly available for neuromarketing purposes. In this dataset, the participants were shown 42 product images of 14 categories for 4 seconds each. As 42 product images were shown to 25 participants; therefore, a total of 1050 (i.e., 42 x 25) EEG signals were recorded by them. Then participants were asked to label the products with either like or dislike, and their responses were recorded. The EEG data were recorded through 14 channels located at areas as per the International 10-20 system [9].

B. Data Pre-Processing

EEG recording is highly sensitive to different sources of noise. The noise makes it difficult to intercept the characteristics by analyzing EEG signals. There are many existing methods to deal with the noise effectively [10],[11]. Data smoothing is a method that is used to remove noise from a dataset. Using data smoothing helps pattern in signal to stand out. To remove noise from the recorded EEG data in the dataset, we used running average [12] where running average creates different subsets of the whole data set by calculating to interpret data points. A running average is commonly used with time-series data to smooth out short-term fluctuations and highlight longer-term trends or cycles. The threshold between the short-term and long-term depends on the application, and the parameters of the running average will be set accordingly.

horized licensed use limited to: NUST School of Electrical Engineering and Computer Science (SEECS). Downloaded on November 15,2023 at 18:24:50 UTC from IEEE Xplore. Restrictions ap

C. Feature Extraction

Feature extraction is a process of extracting relevant information from raw EEG signals and the reduction of dimensionality. Several feature extraction techniques based on the time domain, frequency domain, and time-frequency domain are available. For EEG signal processing, the most popular and suitable method is the Wavelet Transform (WT) method. It takes both time and frequency domains into account while extracting features. For high or low-frequency resolution, short-time and long-time windows are used respectively. In this method, EEG signals are represented using wavelets. There are two kinds of Wavelet Transform method. They are the Continuous Wavelet Transform (CWT) method and the Discrete Wavelet Transform (DWT) method. We chose to use DWT in our feature extraction due to its superiority in fast non-redundant transforms [13]. The DWT is a form of WT in which the wavelets are sampled discretely using scaling and translation parameters. The signals are decomposed into an orthogonal wavelet vector. DWT is used for multilevel feature representation of signal data. In this method, both time and frequency domains are considered.

The signal $x(t)$ is decomposed using DWT by passing it through a series of filters. At first, it's passed through a low pass filter g and then a high pass filter h as shown equation 1 and equation 2 respectively.

$$y_a = g\{x(t)\} \tag{1}$$

$$y_d = h\{x(t)\} \tag{2}$$

Then we calculated power as a feature from 5-level decomposed EEG signals, as shown in equation 3.

$$\lambda = \vartheta\{y_d\} \tag{3}$$

Where λ is a power feature, and ϑ is an equation of power.

D. Classification

For each electrode, we used 5-level decomposition of DWT and then calculated power based on decomposed signals. The total number of features per instance was the number of electrodes multiplied by 5. Therefore, the total number of instances was 1045. After feature extraction, we used a decision tree algorithm to classify the product advertising preference of a consumer. Ensemble methods like gradient boosting can improve the accuracy of decision trees but at the expense of the interpret-ability of the generated model. Additive models, such as those produced by gradient boosting, and full interaction models, such as CART, have been investigated largely in isolation.[14]

For the classification, we divided the feature set into training and testing sets and then generated the model using the training dataset. After getting the model, we passed the testing dataset into the model and calculated the predicted outcome of the classifier. Then compared the actual and predicted outcome and calculated the performance of our proposed method. In this paper, decision tree (DT) algorithm[15] is used for classification as DT is a popular data mining algorithm used for detection, prediction, and pattern analysis purposes [16].

III. Experimental Results and Discussion

In this section, we present the experimental results and discussion on our proposed method. The performance of the proposed method is compared with some existing techniques, namely KNN, DA, NB, SVM, and RF in terms of area under curve, accuracy, sensitivity, and specificity [17] The performance comparison based on the area under curve, accuracy, sensitivity, and specificity is given below.

A. Area Under Curve

Performance measurement is really important in machine learning. Through the AUC - ROC curve, it is possible to measure the performance of classifier algorithms. AUC - ROC curve is used to check or visually represent the performance of multi-class classifiers. ROC curve is best among many evaluation techniques. It is also known as AUROC (Area Under the Receiver Operating Characteristics). In different threshold settings, AUC - ROC curve can measure the performance of classification problems. AUC represents separability of measurement, and Probability is determined by the ROC curve. If the AUC is high, then it means that the model is good at predicting the right answer. When plotting the ROC curve, the x-axis is denoted as a false positive rate, and the y-axis is denoted as the true positive rate.

When the AUC of a model has a result near to 1, that means that the model will stand out, and it also means it has a good measure of separability. Whereas a poor model has AUC near 0, which means it has an unfavorable measure of separability. In fact, it means it is reciprocating the result. It is predicting 0s as 1s and 1s as 0s. And when AUC is 0.5, it means the model has no class separation capacity whatsoever [18].

The ROC curves for the decision tree algorithm and support vector machine are presented in Fig 2 and Fig 3. In Fig 2, we labeled the frontal lobe with blue color, occipital with red, parietal with yellow, temporal with violet, and cerebral cortex with color green. In Fig 3, we labeled the frontal lobe with blue color, occipital with red, parietal with yellow, temporal with violet, and cerebral cortex with color green. From Fig 2 and Fig 3, we can see that ROC for the decision tree is better than the ROC of SVM. Moreover, we also observed that the ROC curve of the decision tree is better than the other classification algorithms in every brain area as shown in Table I.

B. Accuracy

Accuracy is one of the most useful metrics for evaluating classification models. It is the fraction of correct predictions out of all predictions. In our model, higher accuracy implies the model's strength to correctly predict if the consumer will like the product or not. The accuracy can be calculated by equation 4.

horized licensed use limited to: NUST School of Electrical Engineering and Computer Science (SEECS). Downloaded on November 15,2023 at 18:24:50 UTC from IEEE Xplore. Restrictions ap

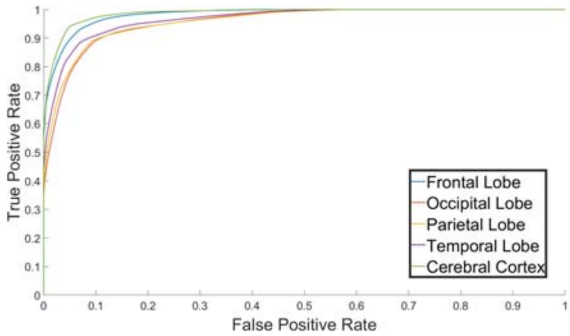

Fig. 2. Five ROC Curves of Different Brain Areas for Decision Tree Classification on Consumer Behavior Analysis Using EEG Signals.

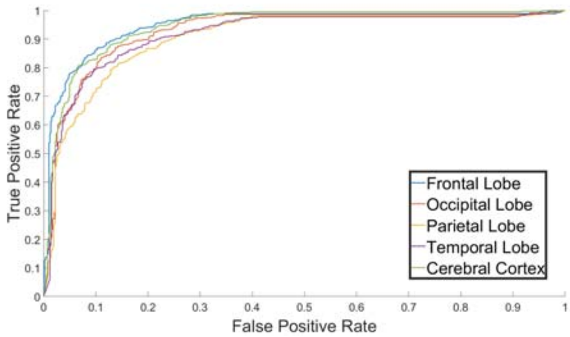

Fig. 3. Five ROC Curves of Different Brain Areas for Support Vector Machine Classification on Consumer Behavior Analysis Using EEG Signals.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \quad (4)$$

Here TN indicates true negative, TP indicates true positive, FP indicates false positive, and FN indicates false negative. II and Figure 4 show accuracy for different classification algorithms on different areas of brain. From Table II and Figure 4, we can see that the average accuracy of the decision tree on all parts of the brain is better than the average accuracy of the existing techniques. The highest accuracy of the decision tree is 95% that is obtained in the cerebral cortex. Note that the lowest accuracy of the decision tree is 90%, which is higher than the highest accuracy of any other existing classification techniques.

TABLE I
AREA UNDER CURVE FOR DIFFERENT CLASSIFICATION ALGORITHMS.

Brain Area	KNN	DA	NB	DT	SVM	RF
Frontal Lobe	83%	59%	86%	**98%**	95%	56%
Occipital Lobe	82%	56%	72%	**96%**	93%	54%
Parietal Lobe	81%	55%	75%	**96%**	91%	51%
Temporal Lobe	83%	54%	79%	**97%**	92%	55%
Cerebral Cortex	85%	66%	91%	**99%**	95%	62%

TABLE II
ACCURACY FOR DIFFERENT CLASSIFICATION ALGORITHMS.

Brain Area	KNN	DA	NB	DT	SVM	RF
Frontal Lobe	77%	60%	76%	**93%**	87%	54%
Occipital Lobe	75%	56%	63%	**90%**	85%	52%
Parietal Lobe	75%	56%	66%	**90%**	82%	52%
Temporal Lobe	76%	56%	71%	**91%**	85%	54%
Cerebral Cortex	78%	60%	81%	**95%**	87%	60%

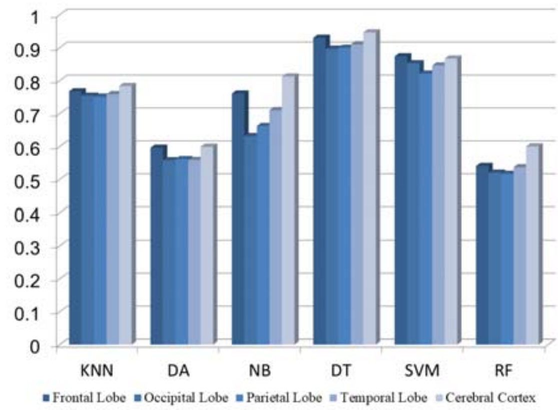

Fig. 4. Accuracy of Six Classifiers for Consumer Behaviour Analysis Using EEG Signal of Five Areas of Brain

C. Sensitivity

Sensitivity measures the proportion of correct positive predictions. In our model, higher sensitivity implies the model's strength to predict that the consumer will like the product correctly. Sensitivity can be calculated using equation 5.

$$Sensitivity = \frac{TP}{TP + FN} \quad (5)$$

Table III and Figure 5 shows sensitivity for different classification algorithms on different areas of brain. From Table III and Figure 5, we can see that the average sensitivity of the decision tree on all parts of the brain is better than the average sensitivity of the existing techniques. The highest sensitivity of the decision tree is 94% that is obtained in the cerebral cortex. Note that the lowest sensitivity of the decision tree is 89%, which is higher than the highest sensitivity of any other existing classification techniques.

TABLE III
SENSITIVITY FOR DIFFERENT CLASSIFICATION ALGORITHMS.

Brain Area	KNN	DA	NB	DT	SVM	RF
Frontal Lobe	74%	20%	67%	**93%**	79%	40%
Occipital Lobe	69%	5%	73%	**89%**	77%	38%
Parietal Lobe	68%	4%	73%	**89%**	72%	37%
Temporal Lobe	70%	3%	61%	**89%**	75%	38%
Cerebral Cortex	71%	31%	81%	**94%**	77%	46%

horized licensed use limited to: NUST School of Electrical Engineering and Computer Science (SEECS). Downloaded on November 15,2023 at 18:24:50 UTC from IEEE Xplore. Restrictions ap

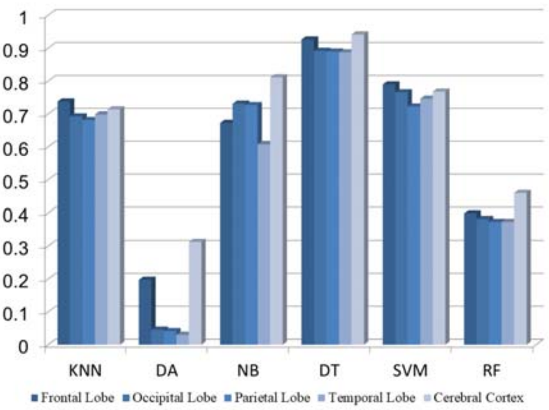

Fig. 5. Sensitivity of Six Classifiers for Consumer Behaviour Analysis Using EEG Signal of Five Areas of Brain

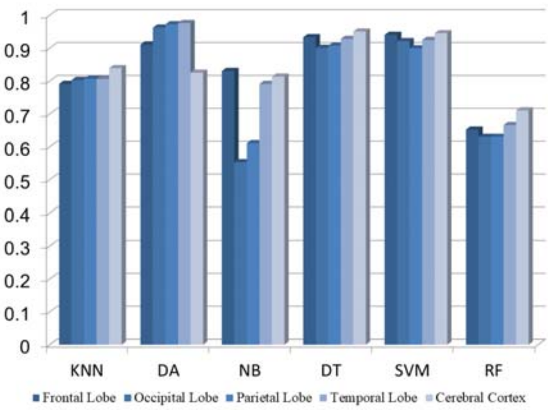

Fig. 6. Specificity of Six Classifiers for Consumer Behaviour Analysis Using EEG Signal of Five Areas of Brain

D. Specificity

Specificity measures the proportion of negative predictions that are correct. In our model, higher specificity implies model's strength to predict that the consumer will not like the product correctly. Sensitivity can be calculated using equation 6.

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

Table IV and Figure 6 shows specificity for different classification algorithms on different areas of brain. From Table IV and Figure 6, we can see that the average specificity of DA and SVM is marginally better than the average specificity of the decision tree (DT). However, the average specificity of the decision tree is better than the average specificity of the existing techniques except for DA and the support vector machine. The highest specificity of the decision tree is 95% that is obtained in the cerebral cortex. Note that the lowest specificity of the decision tree is 90%, which is higher than the highest specificity of any other existing classification techniques except DA and SVM.

TABLE IV
SPECIFICITY FOR DIFFERENT CLASSIFICATION ALGORITHMS.

Brain Area	KNN	DA	NB	DT	SVM	RF
Frontal Lobe	79%	91%	83%	93%	**94%**	65%
Occipital Lobe	80%	**96%**	55%	90%	92%	63%
Parietal Lobe	81%	**97%**	61%	91%	90%	63%
Temporal Lobe	81%	**98%**	79%	93%	92%	67%
Cerebral Cortex	84%	83%	81%	**95%**	**95%**	71%

E. Comparison with an existing technique

Table V shows the comparison of our proposed method with an existing research work of Mahendra et al.(2017)[8]. Note that we compare the performance of the techniques on the same datasets. Our proposed method achieved 95% prediction accuracy, whereas the accuracy of the existing method is 70%, demonstrating the superiority of our proposed method over the existing technique. Moreover, the root to leaf path (i.e., logic rule) of a decision tree on the data will help a producer and advertiser to understand the reasons for like and dislike of a particular product. Note that our proposed method can provide logic rules, whereas the existing work[8] is unable to provide any logic rules.

TABLE V
COMPARISON WITH PREVIOUS WORKS ON CONSUMER PREFERENCE PREDICTION

Author	Analysis Method	Prediction Rate	Imaging Tool
Mahendra et al [8] (2017)	HMM	70%	EEG
Proposed Model	Decision Tree	95%	EEG

IV. CONCLUSION

In this research work, we analyzed EEG signals to get better understandings of consumer behaviors, choices, and preferences. We have built a prediction model that can successfully predict consumer's choices (i.e., like/dislike). Our proposed model performed better than the existing techniques in terms of accuracy, sensitivity, and specificity. The highest accuracy of our proposed model is 95% for the prefrontal region of the cerebral cortex, and the lowest accuracy of the model is 90% for the occipital region. It is to be noted that the lowest accuracy of our proposed method is still above the highest accuracy of the existing techniques as our method used very rigorous processes in data pre-processing, feature extraction, classification, and analysis. The experimental results of our

Authorized licensed use limited to: NUST School of Electrical Engineering and Computer Science (SEECS). Downloaded on November 15,2023 at 18:24:50 UTC from IEEE Xplore. Restrictions apply.

proposed method represent the superiority of our proposed method over the existing techniques. The root to leaf path (i.e., logic rule) of a decision tree on product consumer data will help a producer and advertisers to understand the reasons (i.e., patterns) for like and dislike of a particular product. Our proposed method will allow advertisers to realize which of their advertisement is effective in a specific group of consumers and how they should plan their marketing strategies. Our future plan is to combine EEG signals with functional magnetic resonance images (fMRI) and facial images of consumers to produce a more accurate and effective model for predicting consumer behavior.

REFERENCES

[1] R. Dooley, *What is neuromarketing*, 2006.

[2] R. Bartels, "The history of marketing thought," 1988.

[3] Lisa, *Neuromarketing benefits "you" the consumer*, Jul. 2016. [Online]. Available: https://thebarton.org/consumer-neuromarketing-benefits/.

[4] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Muller, V. Kunzmann, F. Losch, and G. Curio, "The berlin brain-computer interface: Eeg-based communication without subject training," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, pp. 147–152, 2006.

[5] H. R. Heekeren, S. Marrett, P. A. Bandettini, and L. G. Ungerleider, "A general mechanism for perceptual decision-making in the human brain," *Nature*, vol. 431, no. 7010, p. 859, 2004.

[6] A. Smith, B. D. Bernheim, C. F. Camerer, and A. Rangel, "Neural activity reveals preferences without choices," *American Economic Journal: Microeconomics*, vol. 6, no. 2, pp. 1–36, 2014.

[7] A. Telpaz, R. Webb, and D. J. Levy, "Using eeg to predict consumers' future choices," *Journal of Marketing Research*, vol. 52, no. 4, pp. 511–529, 2015.

[8] M. Yadava, P. Kumar, R. Saini, P. P. Roy, and D. P. Dogra, "Analysis of eeg signals and its application to neuromarketing," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 19 087–19 111, 2017.

[9] R. N. Khushaba, C. Wise, S. Kodagoda, J. Louviere, B. E. Kahn, and C. Townsend, "Consumer neuroscience: Assessing the brain response to marketing stimuli using electroencephalogram (eeg) and eye tracking," *Expert Systems with Applications*, vol. 40, no. 9, pp. 3803–3812, 2013.

[10] J Mateo, A. Torres, M. A. García, and J. Santos, "Noise removal in electroencephalogram signals using an artificial neural network based on the simultaneous perturbation method," *Neural Computing and Applications*, vol. 27, no. 7, pp. 1941–1957, 2016.

[11] M. Z. Parvez and M. Paul, "Epileptic seizure detection by exploiting temporal correlation of electroencephalogram signals," *IET Signal Processing*, vol. 9, no. 6, pp. 467–475, 2015.

[12] ——, "Epileptic seizure prediction by exploiting spatiotemporal relationship of eeg signals using phase correlation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 158–168, 2015.

[13] ——, "Epileptic seizure detection by analyzing eeg signals using different transformation techniques," *Neurocomputing*, vol. 145, pp. 190–200, 2014.

[14] J. M. Luna, E. D. Gennatas, L. H. Ungar, E. Eaton, E. S. Diffenderfer, S. T. Jensen, C. B. Simone, J. H. Friedman, T. D. Solberg, and G. Valdes, "Building more accurate decision trees with the additive tree," *Proceedings of the National Academy of Sciences*, vol. 116, no. 40, pp. 19 887–19 893, 2019.

[15] J. R. Quinlan, "Improved use of continuous attributes in c4. 5," *Journal of artificial intelligence research*, vol. 4, pp. 77–90, 1996.

[16] M. A. Rahman, B. Honan, T. Glanville, P. Hough, and K. Walker, "Using data mining to predict emergency department length of stay greater than 4 hours: Derivation and single-site validation of a decision tree algorithm," *Emergency Medicine Australasia*, vol. 32, no. 3, pp. 416–421, 2020.

[17] R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," in *Proceedings of the 23rd international conference on Machine learning*, 2006, pp. 161–168.

[18] S. Narkhede, "Understanding auc-roc curve," *Towards Data Science*, vol. 26, 2018.

horized licensed use limited to: NUST School of Electrical Engineering and Computer Science (SEECS). Downloaded on November 15,2023 at 18:24:50 UTC from IEEE Xplore. Restrictions ap