NeuroRegulation

Predicting Consumer Behavior: A Critical Review of EEG-Based Neuromarketing and the Decision Tree Model

*Gayatri Kapoor Saraya**

Arizona State University, Tempe, Arizona, USA

Abstract

This critical review examines the study by Amin et al. (2020), which proposes a decision tree (DT) model for predicting consumer behavior using electroencephalogram (EEG)-based neuromarketing. The study leverages EEG signals to analyze consumer responses to marketing stimuli, employing advanced data preprocessing, feature extraction, and classification techniques. The DT model demonstrates superior performance in accuracy, sensitivity, and specificity compared to existing methods, achieving a prediction accuracy of 95%. While the study highlights the potential of EEG-based neuromarketing and the interpretability of the DT model, limitations such as sample size constraints, generalizability concerns, and trade-offs between accuracy and interpretability are noted. The review underscores the model's relevance for developing consumer-centric marketing strategies while calling for further research to address its limitations and expand its applicability across diverse populations.

Keywords: neuropsychology; neuromarketing; consumer behavior prediction; EEG signals; decision tree (DT) model

Citation: Kapoor Saraya, G. (2025). Predicting consumer behavior: A critical review of EEG-based neuromarketing and the decision tree model. *NeuroRegulation, 12*(2), 132–137. https://doi.org/10.15540/nr.12.2.132

***Address correspondence to**: Gayatri Kapoor Saraya, TG4, 1B, Orchid Gardens, Suncity, Golf Course Road, Sector-54, Gurugram, Haryana, India 122001. Email: gksaraya@gmail.com

Edited by:
Rex L. Cannon, PhD, Currents, Knoxville, Tennessee, USA

Reviewed by:
Rex L. Cannon, PhD, Currents, Knoxville, Tennessee, USA
Thomas Collura, PhD, BrainMaster Technologies, Bedford, Ohio, USA

Introduction

Neuromarketing, a groundbreaking fusion of neuroscience and marketing, leverages neuropsychological tools such as electroencephalography (EEG), functional magnetic resonance imaging (fMRI), eye tracking, or other biometrics to understand consumers' cognitive and affective responses to marketing stimuli (Sixth Factor, n.d.). The study focuses on data mining and machine-learning technologies to identify brain activities and investigate discoveries or abnormalities as consumers interact with marketing catalysts (Amin et al., 2020). EEG, used to detect electrical charges in brain waves, has been a fundamental tool in the paper by Amin et al. (2020), *Consumer Behavior Analysis using EEG Signals for Neuromarketing Application,* for capturing consumer responses and building on previous research in the field. In the study, Amin et al. (2020) propose a prediction model, the decision tree (DT) model, by harnessing EEG findings. The model is comprehensively evaluated by comparing it to existing prediction methods and pushing the boundaries to provide insights into consumer decision-making processes and thus allow advertisers to create effective marketing models. The following critical review of Amin et al.'s (2020) research work provides a thorough view of the methods and results of the paper, with a focus on the study findings, strengths, weaknesses, and relevance to further the discussion of market research methods.

Methods

Research Design

The study uses a data-driven approach, using EEG signals to explore and predict consumer conduct. Amin et al. (2020) opted for a research design involving four key steps: data collection, preprocessing, feature extraction, and classification. Time-frequency distribution features were derived from the EEG signals on which different classification algorithms were applied, ensuring all aspects of the research were thoroughly considered and measured.

Sample and Data Collection Methods

Yadava et al. (2017) previously conducted a study utilizing the hidden Markov model (HMM), building on a predictive model framework to understand consumer choices through likes and dislikes. They carried out this work by studying brain waves from EEG signals as consumers responded to marketing stimuli. The HMM achieved a 70% prediction accuracy, facilitating the DT model to build on the algorithm's precision, resulting in a high prediction accuracy of 95%.

The study uses publicly available data from Yadava et al.'s (2017) work. Data was collected from 25 subjects who were exposed to 42 images in 14 categories. Each participant labeled the image shown to them as *like* or *dislike*. Corresponding EEG signals to 1,050 responses were recorded and examined.

Data Preprocessing, Feature Extraction, and Classification

EEG signals contain noise, which can obscure the underlying patterns of the brain's activity. To address this, noise reduction methods are used in EEG analysis to smooth signal variations. In the paper, a notable noise reduction approach known as running average was employed in the preprocessing stage, which smooths data by averaging data points over a moving window to reduce noise (Amin et al., 2020). The running average method effectively deals with time-frequency datasets, ameliorating short-term fluctuations and highlighting overall trends.

For feature extraction, the wavelet transform method (WTM), a renowned signal processing technique for analyzing time-frequency representations of EEG signals, was utilized. Amin et al. (2020) specifically used the discrete wavelet transform (DWT) method, a WTM subpart known to reliably break down EEG signals into smaller parts in a fast and nonredundant manner, allowing for a detailed analysis of different frequencies. DWT is crucial for accurate classification as it uses low-pass (g) and high-pass (h) filters to analyze different frequencies (Amin et al., 2020). This offers a comprehensive view of brain activity and establishes trust in the research methods through vigorous processing methods.

Following feature extraction, the power was calculated from five-level DWT-decomposed EEG signals for each electrode. The total number of instances (i.e., single data trials) was 1,045, determined by multiplying the number of electrodes by five (since each underwent a five-level decomposition; Amin et al., 2020). To enhance model performance, the authors considered ensemble methods like gradient boosting, which improves accuracy by combining multiple weaker models into a stronger one. However, while gradient boosting improves accuracy, it is not the best choice for interpretability (Amin et al., 2020). Alternatively, additive models (predictive models that improve performance by sequentially adding corrections) and full interaction models like CART (classification and regression trees) are often studied separately (Luna et al., 2019). This is because gradient boosting prioritizes accuracy, while CART focuses on interpretability.

For classification, Amin et al. (2020) split the extracted features into a training set to build the model and a testing set to evaluate it. After creating the model, data was fed to predict consumer preferences (Amin et al., 2020). To test the proposed models' performance, the predicted outcomes were compared to the actual results. The rest of the review focuses on conveying the study findings and analyzing the advantages, drawbacks, and applicability.

Results

The proposed model's performance was evaluated using four key metrics: area under the curve (AUC), accuracy, sensitivity, and specificity (Amin et al., 2020). The authors compared their model with five existing techniques for consumer behavior prediction: k-nearest neighbors (KNN), discriminant analysis (DA), naive Bayes (NB), support vector machines (SVM), and random forests (RF). Amin et al.'s (2020) findings suggest that the DT algorithm is superior in accuracy and sensitivity testing across all brain areas, providing a more reliable method for predicting consumer decision-making. The findings of this comparison are discussed below.

Area Under Curve

An AUC-ROC (area under the receiver operating characteristics) curve, also known as ROC, evaluates classifier performance, validating the tree model's effectiveness and visually representing multiclass classifier performance (Amin et al., 2020). Among other evaluation techniques, ROC is the most noteworthy, further building on the study's validity. The proposed model's ROC curve outperformed SVM and other classification algorithms across all brain areas, with a high measure of 99% in the cerebral cortex and a low measure of 96% in the occipital and parietal lobes. The highest measure noted in other existing models was 95% (for SVM), yet lower than any other measurement for DT (Amin et al., 2020).

Accuracy

Accuracy is a key measure for assessing classification models. It represents the proportion of correct predictions among all predictions made (Amin et al., 2020). A higher accuracy indicates the model's effectiveness in determining whether a consumer will like or dislike a product. The DT model achieved a high accuracy of 95% in the prefrontal region of the cerebral cortex, with a low accuracy of 90% in the occipital region. Notably, the proposed model's lowest accuracy was still higher than the highest accuracy of any other existing technique (Amin et al., 2020).

Sensitivity

Sensitivity measures the percentage of correctly predicted positive cases. Higher sensitivity indicates the model's ability to identify accurately when a consumer will like the product. DT's sensitivity is exceptionally high, ranging from 89% (in the occipital, parietal, and temporal lobes) to 94% (in the cerebral cortex), indicating its ability to predict positive consumer preferences correctly. The proposed model's lowest sensitivity was still higher than the highest sensitivity of any other existing technique (Amin et al., 2020).

Specificity

Specificity measures the percentage of correctly predicted negative cases. Higher specificity indicates the model's ability to determine accurately when a consumer will dislike the product. While DT performs well in specificity (90%, 95%), it is marginally outperformed by DA and SVM in some brain areas. For example, DA achieves 98% specificity in the temporal lobe, compared to DT's 93% (Amin et al., 2020).

The DT model demonstrates the best overall performance, with high accuracy, AUC, sensitivity, and specificity across all brain areas. Its interpretability further enhances its practical utility for neuromarketing applications. The table below (Table 1) lists the performance of models in comparison to each other under different metrics and brain areas.

Table 1
Combined Results Table: Performance of Classification Algorithms in Predicting Consumer Preferences

Metric	Brain Area	KNN	DA	NB	DT	SVM	RF
Accuracy	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%
AUC (area under the curve)	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%
Sensitivity	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%

Table 1
Combined Results Table: Performance of Classification Algorithms in Predicting Consumer Preferences

Metric	Brain Area	KNN	DA	NB	DT	SVM	RF
Specificity	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%

Note. KNN = k-nearest neighbors; DA = discriminant analysis; NB = naive Bayes; DT = decision tree; SVM = support vector machine; RF = random forest. Data adapted from Amin et al. (2020).

Compared to the HMM proposed by Yadava et al. (2017), which had a prediction rate of 70%, the DT model significantly outperformed, achieving a 95% prediction rate using the same neurological testing tool (EEG; Table 2).

Table 2
Comparison With Previous Works

Author	Analysis Method	Prediction Rate	Imaging Tool
Yadava et al.	HMM	70%	EEG
Amin et al.	DT	**95%**	EEG

Note. HMM = hidden Markov model; DT = decision tree; EEG = electroencephalogram. Data adapted from Amin et al. (2020).

The authors conclude that their proposed method is superior to other existing techniques in terms of accuracy, sensitivity, and specificity, allowing advertisers to gain insights into consumer behavior and tailor their marketing strategies accordingly (Amin et al., 2020).

Analysis

This section highlights the study's advantages and limitations and the significance of the DT model in neuromarketing and consumer analysis.

Strengths

Needless to say, a key advantage of the study by Amin et al. (2020) is that the DT model outperforms the existing techniques by high margins, proving it to be an impactful and superior prediction model. Aside from the precise results of the study, several elements are presented that enhance the trustworthiness and robustness of its findings to improve the reliability and applicability of the results. For example, the use of technology such as data mining and machine learning helps further innovations in fields like neuropsychology. This practice involving notable data technologies assures consumers of ethical and credible research outputs.

Amin et al. (2020) often referenced well-acclaimed research to increase the potency of the study results. For example, Blankertz et al. (2006) and Heekeren et al. (2004) are cited to explain the relation between brain activities and the EEG systems, thus increasing the academic trust of the paper in review.

Using a thorough research design and reliable data processing methods (running average for noise reduction, WTM, and DWT), the study demonstrates the authors' commitment to achieving excellence in the study results. Rigorous examinations were carried out by Amin et al. (2020) by employing multiple evaluation metrics (AUC, accuracy, sensitivity, and specificity) and comparing the proposed DT model to five other existing techniques. This high standard of evaluation illustrates the robustness of the study findings.

Finally, the root-to-leaf path (logic rule) of the proposed DT algorithm makes it highly interpretable, while the study by Yadava et al. (2017) does not provide any logic rule (Amin et al., 2020). This enables a business to understand consumer cognition and its surrounding elements, making DT a more practical and desired prediction model.

Weaknesses

Despite the study's multiple advantages, a few limitations have also been observed. DT establishes superiority in the accuracy and sensitivity metrics compared to other existing techniques. However, according to the study results, DT lacks specificity (DA and SVM measure higher values), which may make it difficult to determine when a consumer will dislike a product.

A limitation of the study is the generalizability of the DT model stemming from the sample size of 25 participants. While the model proves to have high performance, it may have been overfitted to the specific dataset, meaning its high accuracy might not hold across different demographics, cultural backgrounds, or real-world consumer scenarios. Additionally, small-sample studies have reduced statistical power, thereby increasing the likelihood of spurious correlations and making it harder to detect true patterns in consumer decision-making. This also limits the ability to examine individual differences in EEG responses, which are known to vary based on age, gender, and cognitive traits (Yadava et al., 2017). To improve reliability, future studies should consider using larger and more representative samples to validate the DT model's predictive performance across different populations.

Moreover, the dataset studied by Amin et al. (2020) was derived from publicly available data from Yadava et al.'s (2017) paper. This suggests that the data was not firsthand, raising concerns about the authors' direct involvement with the subject and whether the dataset fully represents diverse consumer behaviors.

Lastly, methods such as gradient boosting and CART, as mentioned in the paper by Amin et al. (2020), are accurate in their measurement. However, a trade-off between correctness and clarity creates a potential limitation in the study due to the challenging interpretation for the reader.

Relevance of the DT Model

The DT model can potentially revolutionize how market strategies are developed and implemented today. This mutual benefit to consumers and producers can lead to a more consumer-centric and accurate approach to developing marketing strategies (Amin et al., 2020). The study's findings have significant implications in neuromarketing because the DT is a highly interpretable and applicable model. Further research in this field would help advertisers understand the reasons for consumer preferences and develop more targeted and personalized strategies by employing different neurological tools such as eye-tracking or fMRI.

On the other hand, while the use of data mining and machine learning algorithms are broadly used today in most sectors of society, ethical concerns follow regarding participant privacy and consumer manipulation. It is crucial to keep customer autonomy at the forefront when employing influential technology in machine learning.

Conclusion

The paper by Amin et al. (2020) presents a robust and interpretable model, the DT model, used to predict consumer decision-making to develop marketing strategies by harnessing EEG signals. The data presented in the paper indicates that the DT model performs better than other existing predictive models. The DT model proves superior accuracy, sensitivity, and specificity, providing valuable insights to marketing and analysis teams. The study presents several advantages and a few limitations in the overall application of the research. Amin et al. (2020) achieve high trust by using reliable techniques and robust comparisons in their work. Overall, the authors open avenues to better neuromarketing studies and offer valuable contributions to consumer behavior prediction and experience.

Author Declarations

The author declares no conflicts of interest. This research did not receive any specific grant from funding agencies. Gayatri Kapoor Saraya conceptualized the study, extracted and organized the data, and wrote the manuscript.

References

Amin, C. R., Hasin, M. F., Leon, T. S., Aurko, A. B., Tamanna, T., Rahman, M. A., & Parvez, M. Z. (2020, December). Consumer behaviour analysis using EEG signals for neuromarketing application. In *2020 IEEE symposium series on computational intelligence (SSCI)* (pp. 2061–2066). IEEE. https://doi.org/10.1109/SSCI47803.2020.9308358

Blankertz, B., Dornhege, G., Krauledat, M., Muller, K.-R., Kunzmann, V., Losch, F., & Curio, G. (2006). The Berlin brain-computer interface: EEG-based communication without subject training. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 14*(2), 147–152. https://doi.org/10.1109/TNSRE.2006.875557

Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature, 431*(7010), 859–862. https://doi.org/10.1038/nature02966

Luna, J. M., Gennatas, E. D., Ungar, L. H., Eaton, E., Diffenderfer, E. S., Jensen, S. T., Simone, C. B., Friedman, J. H., Solberg, T. D., & Valdes, G. (2019). Building more accurate decision trees with the additive tree. *Proceedings of* *the National Academy of Sciences, 116*(40), 19887–19893. https://doi.org/10.1073/pnas.1816748116

Sixth Factor. (n.d.). *Neuro marketing sensonomics.* Retrieved March 6, 2025, from https://sixthfactor.com/neuro-marketing-sensonomics/

Yadava, M., Kumar, P., Saini, R., Roy, P. P., & Dogra, D. P. (2017). Analysis of EEG signals and its application to neuromarketing. *Multimedia Tools and Applications, 76*(18), 19087–19111. https://doi.org/10.1007/s11042-017-4580-6

Received: March 10, 2025
Accepted: March 27, 2025
Published: June 27, 2025