Optimizing AI Models for Identifying Buying Intention Using Time-Frequency Domain EEG Features

Stralen Pratasik
Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
7022222003@student.its.ac.id

Adhi Dharma Wibawa
Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia | Department of Medical Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
adhiosa@te.its.ac.id (corresponding author)

Diah Puspito Wulandari
Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia | Department of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
dp.wulandari@its.ac.id

Siti Dwi Suryani
Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
suryanisitidwi@gmail.com

Yuri Pamungkas
Department of Medical Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
yuri@its.ac.id

Received: 11 June 2025 | Revised: 20 August 2025 | Accepted: 30 August 2025

 DOI: https://doi.org/10.48084/etasr.12676

ABSTRACT

This study investigates the use of Electroencephalography (EEG) to classify consumer buying intention elicited by advertising stimuli. The main objective is to identify the most relevant EEG features and evaluate the effectiveness of various machine learning algorithms in predicting consumer intent. EEG signals were collected from 28 participants using six electrodes (Fp1, Fp2, F7, F8, O1, and O2) while they viewed video advertisements. The signals underwent preprocessing involving filtering, Independent Component Analysis (ICA), and amplitude clipping. Subsequently, wavelet-based segmentation was employed to extract alpha, beta, and gamma frequency components, from which 15 statistical features per channel were computed, grouped into (1) Time-Domain Metrics, (2) Entropy and Energy Features, and (3) Fractal Measures, yielding a total of 270 features. Dimensionality reduction was performed using three feature selection techniques: Pearson Correlation (PC), Linear Discriminant Analysis (LDA), and Mutual Information (MI). Each feature selection method was evaluated in combination with five classifiers: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Random Forest (RF), Naïve Bayes (NB), and Decision Tree (DT). The RF classifier consistently achieved the highest accuracy across all selection methods, peaking at 93%. Among the feature selectors, MI proved most efficient, achieving 90% accuracy with only 20 features. The most discriminative features were primarily derived from the gamma and beta bands in frontal and occipital regions, reflecting their role in attention and decision-making processes. Overall, the findings demonstrate the potential of EEG-based approaches for compact and accurate prediction of consumer intent, offering valuable insights for applications in neuromarketing and advertising analytics.

Keywords-buying intention; neuromarketing; electroencephalography; mutual information; video advertising

I. INTRODUCTION

Understanding consumers' buying intention is essential for effective marketing, particularly in the context of digital media such as video advertising. Buying intention, defined as a consumer's willingness to purchase a product, is a strong predictor of actual purchasing behavior. Traditionally, marketing research assesses buying intentions through self-report methods like questionnaires and interviews. However, these methods are inherently subjective and prone to social desirability bias, as consumers may respond in ways they believe are expected rather than reflecting their true preferences [1]. Moreover, post-hoc surveys capture retrospective attitudes, which may not reflect the consumer's genuine real-time psychological states during stimulus exposure. In practice, this can lead to a gap between expressed intention and actual behavior.

To address these limitations, researchers have turned to neuromarketing approaches using Electroencephalography (EEG). EEG records brain electrical activity non-invasively while consumers engage with marketing stimuli, offering millisecond-level temporal resolution in a portable and cost-effective manner [2]. Most importantly, EEG signals cannot be consciously controlled, allowing for objective insight into attention, emotional engagement, and cognitive processing [1, 2] related to purchase decisions. Prior studies have associated frontal alpha asymmetry and theta power variations with consumer intent, underscoring the EEG's potential in predicting shopping behavior [3].

Recent neuromarketing research has further demonstrated the feasibility of predicting buying decisions from EEG signals. For instance, authors in [4] recorded EEG activity during an e-commerce shopping task and extracted spectral features such as Power Spectral Density (PSD) and prefrontal asymmetry indices. Using machine learning classifiers, including K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Random Forest (RF), and shallow neural networks, they achieved up to 87.1% accuracy in predicting whether a product would be purchased. Their results highlighted the importance of frontal lobe activity, related to emotion and decision-making, and occipital regions, involved in visual attention. Notably, increased frontal alpha asymmetry and elevated theta power were observed when participants viewed products they ultimately chose to buy, confirming these as neural markers of purchase intention.

Beyond decision prediction, EEG has also been used to assess preferences and emotional responses to advertisements. Authors in [5] have applied Discrete Wavelet Transform (DWT) and PSD to decompose EEG signals recorded during ad viewing into frequency bands (delta, theta, alpha, beta), extracting features linked to liking or disliking content. PSD-derived features from specific electrodes and frequency bands have been used as "preference indices" to quantify emotional engagement, while frontal asymmetry has served as an indicator of motivation and desire. Interestingly, while deep learning models yielded marginal performance gains, traditional classifiers such as SVM, KNN, and RF achieved comparable accuracy. This suggests that ensemble methods remain robust for EEG-based neuromarketing tasks, offering a balance of interpretability and efficiency.

In order to achieve EEG-based consumer state classification, a variety of machine learning methods have been applied in previous studies. For example, authors in [6] systematically compared feature selection methods such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), minimum Redundancy Maximum Relevance (mRMR), Recursive Feature Elimination (RFE), and ReliefF, combined with classifiers such as deep neural networks, SVM, KNN, and RF. The analysis revealed that feature selection consistently improved model accuracy, underscoring the importance of dimension reduction in EEG analysis.

Despite these advances, several gaps remain. Few neuromarketing studies have explicitly applied information-theoretic feature selection methods, such as Mutual Information (MI), to optimize EEG-based prediction. While authors in [6] implicitly incorporated an MI-based criterion (mRMR), most prior research has relied on unsupervised techniques (e.g., PCA) or allowed classifiers to implicitly handle redundancy. Furthermore, many state-of-the-art studies employ complex deep learning models that require extensive data and computational resources, despite evidence that simpler models like RF can achieve comparable results [5]. This reveals an opportunity to develop streamlined, interpretable, and computationally efficient frameworks for EEG-based buying intention detection.

Accordingly, this study hypothesizes that applying supervised feature selection methods, specifically Pearson Correlation (PC), LDA, and MI to EEG features in the time-frequency domain can enhance model efficiency by isolating a smaller, more relevant subset of features while maintaining or improving classification accuracy. EEG signals were decomposed using DWT to extract time-domain, entropy/energy, and fractal features from alpha, beta, and gamma frequency bands. To evaluate performance, five classification algorithms, SVM, KNN, RF, Naïve Bayes (NB), and Decision Tree (DT), were tested across varying feature subsets. The study aims to assess how different feature selection strategies and classifiers interact to optimize the detection of buying intention (high or low) from EEG data.

The key contribution of this research lies in proposing a replicable and accessible EEG-based framework that combines principled feature selection with efficient machine learning to improve the accuracy and reliability of buying intention detection. This work bridges the gap between consumer neuroscience and practical marketing analytics, offering insights that can enhance personalized advertising and predictive modeling in neuromarketing applications.

II. PROPOSED METHOD

Figure 1 outlines the proposed EEG-based framework for analyzing consumer buying intention. The process consists of five sequential stages. First, raw EEG signals are acquired while participants view marketing stimuli. Second, the signals are preprocessed to remove noise and artifacts, ensuring clean neural data. Third, the preprocessed signals are segmented into frequency bands to isolate relevant cognitive components.

Fourth, time-frequency features are extracted to capture neural characteristics associated with decision-making. Finally, feature selection and classification are applied to identify levels of buying intention using machine learning algorithms.

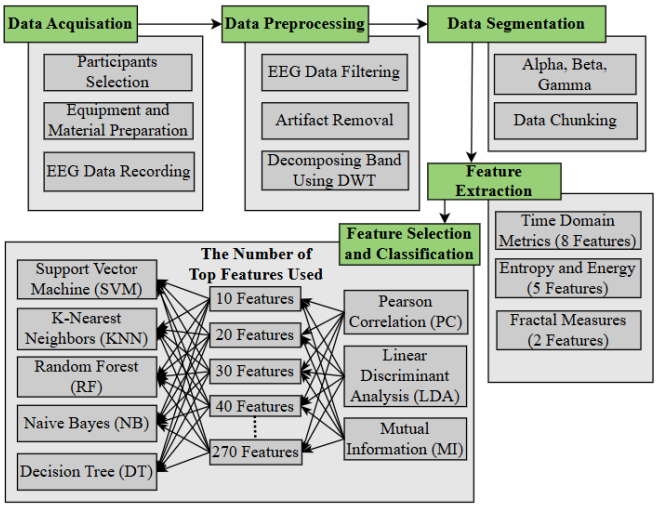

Fig. 1. The proposed EEG-based five-stage framework.

A. Data Acquisition

A total of 28 participants (18 males and 10 females; aged 18-32 years) were recruited for this study. Prior to participation, all individuals completed a self-reported eligibility questionnaire to ensure suitability for the experiment. The screening confirmed that participants were in generally good health, had no history of neurological or psychological disorders, and had obtained at least six hours of sleep the night before data collection.

EEG data were recorded using the OpenBCI system, equipped with the Ultracortex "Mark IV" headset and Cyton board, at a sampling rate of 250 Hz. The experimental stimuli consisted of two Indonesian local sports shoe advertisements retrieved from YouTube. The first video represented the brand Speed, while the second featured Pecs.

EEG acquisition involved six dry electrodes placed according to the international 10-20 system at the following scalp locations: Fp1, Fp2, F7, F8, O1, and O2. Ground electrodes were attached to both earlobes. These electrodes were selected for their sensitivity to prefrontal (cognitive and emotional) and occipital (visual attention) processes, both crucial to consumer responses during advertisement viewing [7-10]. As illustrated in Figure 2, electrodes Fp1, F7, and O1 were positioned on the left hemisphere, while Fp2, F8, and O2 were on the right hemisphere.

The experimental protocol is summarized in Figure 3. Each EEG recording session began with a baseline phase, during which participants viewed a black screen. This was followed by Stimulus 1 (the first video advertisement). After viewing, participants completed a short buying-intention questionnaire, rating their intention to purchase the advertised product. Responses were labeled as high or low intention based on the scores. Participants then rested for 30 seconds before watching Stimulus 2 (the second video), followed by the same questionnaire procedure.

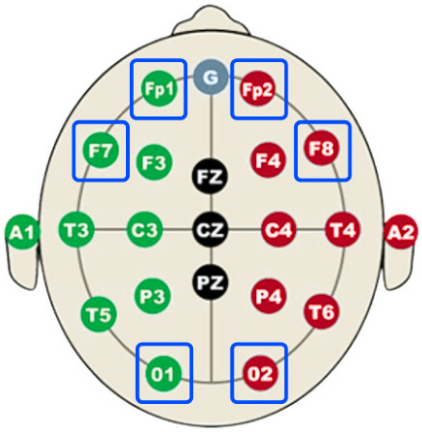

Fig. 2. Electrodes placement using the standard 10-20 placement system.

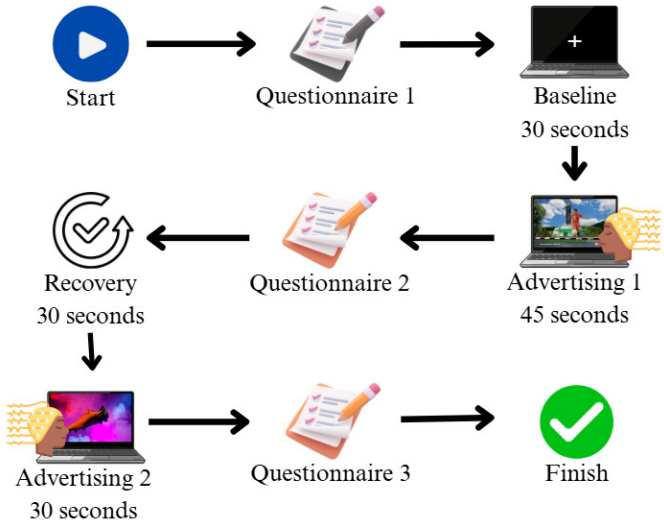

Fig. 3. Illustration of the data acquisition process.

B. EEG Data Preprocessing and Segmentation

Before analysis, the initial 5 seconds of EEG recordings were discarded, based on the prior observations indicating instability or anomalies typically present during the first 1-3 seconds of signal acquisition. Consequently, only the remaining 40 seconds from the first video stimulus and 25 seconds from the second were retained for further analysis.

Preprocessing of EEG signals began with frequency-domain filtering using a 4th-order Butterworth filter implemented via SciPy in Python. An Infinite Impulse Response (IIR) bandpass filter was applied between 1 Hz and 49 Hz to retain neural-relevant frequency components, while bandstop filters at 50 Hz and 60 Hz were used to eliminate electrical line noise. This initial step effectively suppressed low-frequency drifts and high-frequency interference, including motion-related artifacts. Next, Independent Component Analysis (ICA) was applied using the MNE-Python library to identify and remove stereotyped physiological artifacts,

particularly those generated by eye blinks and other ocular activities. Following artifact removal, amplitude clipping was performed to constrain the signal within a physiological range of ± 100 µV. This step eliminated remaining extreme values and residual spikes that might distort feature extraction. The ±100 µV threshold was selected according to established EEG amplitude standards [11]. From a neurophysiological standpoint, signals exceeding this range are typically non-neural, arising from motion, muscle activity, or electrode detachment [12]. To ensure this clipping step did not distort meaningful neural activity, visual inspections were conducted before and after clipping.

Such a multistage preprocessing pipeline is essential in EEG analysis due to the inherently noisy nature of raw signals, where artifact amplitudes can be 10-100× greater than underlying neural activity [13-16]. The combination of filtering, artifact rejection, and decomposition effectively enhances signal quality, isolating cognitively meaningful neural information [17-19].

To further refine the EEG data, the signal was decomposed into specific frequency bands using the DWT technique. Given the original sampling frequency of 250 Hz from the EEG device, the signal was downsampled to 125 Hz through this decomposition process, reducing computational complexity while maintaining relevant information. DWT is especially effective for analyzing non-stationary signals, offering superior performance compared to other spectral analysis methods [20].

The DWT operates by transforming the time-domain EEG signal $x(t)$ into its wavelet representation $\gamma(t)$, as expressed in (1) [21].

$$\gamma(t) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{2^a}} \psi\left(\frac{t-b\times 2^a}{2^2}\right) dt \quad (1)$$

This transformation involves the use of a mother wavelet function $\psi(t)$, along with scale a and shift b parameters, enabling multi-resolution analysis across both time and frequency domains. The Daubechies (db4) wavelet was chosen as the mother wavelet due to its balance between time and frequency resolution, orthogonality, and compact support, properties that make it well-suited for detecting transient, non-stationary EEG dynamics [20].

Figure 4 illustrates the band decomposition process. The resulting EEG sub-bands enabled a targeted investigation of cognitive activity. Subsequently, each signal was segmented into 0.2-second (1/5 s) chunks. This fine segmentation allowed for the detection of rapid temporal fluctuations associated with cognitive and emotional responses, enhancing the precision of feature extraction and improving the identification of EEG markers linked to buying intention.

C. Feature Extraction

After wavelet decomposition, 15 statistical and nonlinear features were extracted from each EEG band. These were grouped into three categories: (1) Time-Domain Metrics, (2) Entropy and Energy Features, and (3) Fractal Measures. Features were computed from the alpha, beta, and gamma bands, which are frequency ranges closely related to attention, emotional engagement, and decision-making processes. Table I lists each feature along with its corresponding abbreviation, grouped by each category. Features were extracted for all six electrode channels across the three frequency bands, yielding a total of 270 features (15 × 6 × 3).

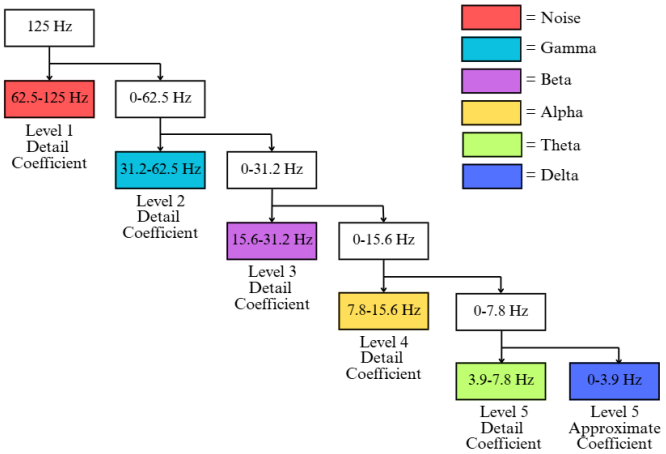

Fig. 4. Decomposition of frequency bands using 5-level DWT.

TABLE I. EXTRACTED EEG FEATURES WITH ABBREVIATIONS AND GROUP CLASSIFICATION

No	Feature Name	Abbreviation	Group
1	Mean	Mean	Time Domain Metrics
2	Mean Absolute Value	MAV	
3	Standard Deviation	Std	
4	Skewness	Skew	
5	Kurtosis	Kurt	
6	Activity	Act	
7	Mobility	Mob	
8	Complexity	Com	
9	Wavelet Entropy	WE	Entropy and Energy Features
10	Zero Crossing Rate	ZCR	
11	Shannon Entropy	SE	
12	Differential Entropy	DE	
13	Energy	En	
14	Higuchi Fractal Dimension	HFD	Fractal Measures
15	Petrosian Fractal Dimension	PFD	

D. Features Selection and Classification Stages

Three feature-ranking methods were evaluated: PC, LDA, and MI. For each, subsets of top-ranked features (10, 20, …, 270) were tested using five classifiers: SVM, KNN, RF, NB, and DT.

The PC method ranked features by the absolute correlation coefficient $|R|$ between each feature and the class label, as given in (2):

$$|R| = \left| \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}-\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \right| \quad (2)$$

where x_i and y_i represent the individual values of each variable, and $\bar{x}$ and $\bar{y}$ denote their respective mean values.

In the LDA method, features were ranked according to their discriminant coefficients $|\omega_j|$, representing each j feature's

contribution to class separation. The weight vector $\boldsymbol{w}$ is computed as the product of the inverse within-class scatter matrix Σ^{-1} and the difference between the mean vectors of the two classes $(\mu_1 - \mu_2)$.

$$|\omega_j| = |\boldsymbol{w}_j| \text{ where } \boldsymbol{w} = \Sigma^{-1}(\mu_1 - \mu_2) \quad (3)$$

The MI method quantified the dependency between each feature X and the target class Y, ranking features based on the MI score calculated as:

$$MI(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y) log\left(\frac{p(x,y)}{p(x)p(y)}\right) \quad (4)$$

where $p(x,y)$ is the joint probability distribution of X and Y, and $p(x), p(y)$ are their respective marginal probability distributions.

The five classification algorithms used in this study can be organized into three primary groups based on their methodological characteristics and strengths: RF and DT belong to tree-based classification methods. DT constructs models by recursively partitioning the dataset based on features that maximize class separation. Its intuitive structure and high interpretability make it well-suited for explaining decision logic in complex datasets [22, 23]. RF extends this principle by creating an ensemble of DTs, where the final class is determined through majority voting. This ensemble strategy improves classification accuracy and reduces overfitting, while also maintaining robustness in the presence of missing data and high-dimensional features [24-26]. KNN is an instance-based method that classifies a data point by analyzing the majority label among its k nearest neighbors. It does not require an explicit training phase, making it straightforward and adaptive to different types of data [27]. Like DT, KNN is non-parametric and handles non-linear relationships well, while its independence from assumptions about data distribution further enhances its versatility [28, 29]. NB and SVM, while fundamentally different in approach, both excel in computational efficiency and generalization ability. NB is a probabilistic classifier based on Bayes' theorem and assumes feature independence, which makes it computationally light and scalable for large datasets [30, 31]. SVM, on the other hand, identifies the optimal hyperplane that maximizes class margins, making it highly effective for high-dimensional and unstructured data, such as EEG signals, text, or images [32-36].

To evaluate and compare the performance of these classification methods, classification accuracy was used as the primary evaluation metric. Accuracy indicates how well a model correctly predicts the class labels and is defined in (5), where CP is the number of correct predictions, and DS is the total number of data samples.

$$Accuracy = \frac{CP}{DS} \cdot 100\% \quad (5)$$

To evaluate the efficiency of the feature selection method, a multicollinearity analysis using PS was conducted on the top 20 selected features. The analysis considered all possible feature pairs, resulting in 190 combinations calculated using (6), where n is the number of features. A correlation coefficient threshold of $|r| > 0.95$ was applied to identify highly correlated pairs, which indicate redundancy. In addition, a one-way Analysis of Variance (ANOVA) test was performed to examine whether the selected features exhibited statistically significant differences across classes.

$$Number\ of\ Pairs = \frac{n(n-1)}{2} \quad (6)$$

In this study, 70% of the data was allocated for training the models, while the remaining 30% was used for testing.

III. RESULTS AND DISCUSSION

The performance of the combinations of PC, LDA, and MI feature selectors and SVM, KNN, RF, NB, and DT classifiers was systematically evaluated to identify the most accurate feature-classifier pairings, assess the impact of feature relevance on performance, and determine model robustness across different EEG-derived feature sets.

Figure 5 presents the classification performance of the five classifiers when features were selected using the PC method.

The RF classifier achieved the highest and most stable accuracy, exceeding 90% with as few as 80 features, followed by KNN and DT with accuracies ranging between 80%-85%. These classifiers and SVM showed gradual improvement as the number of features increased, while NB consistently underperformed, remaining below 60% of accuracy.

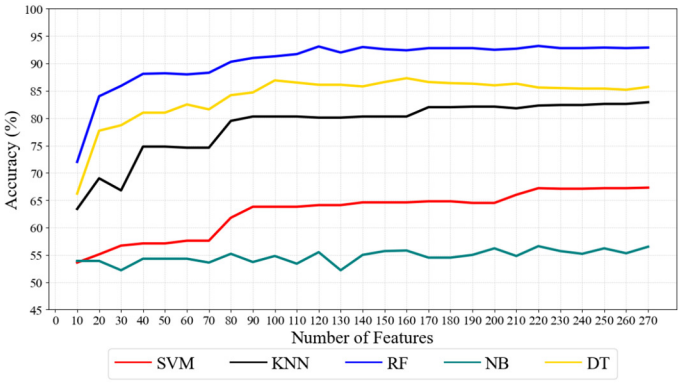

Fig. 5. Classification accuracy using PC feature selection.

Figure 6 presents the results for LDA-based feature selection. As with the PC method, RF again outperformed all other classifiers, achieving >90% accuracy with only 80 features and maintaining stability thereafter. DT performed comparably, consistently reaching accuracy levels above 80%. KNN and SVM both showed a sudden plummet in accuracy after the number of features exceeded 130, which was recovered at 190 features. Nevertheless, especially for the SVM classifier, there was an even greater plummet in accuracy at the 200 features. NB accuracy remained below 65% with very little fluctuation as the number of features increased. Compared to the PC method, for RF, KKN, and DT, the LDA-based feature selection demonstrated faster performance gains with fewer features, particularly between 30 and 50, underscoring LDA's strong discriminative power in identifying features that maximize interclass separability.

Figure 7 depicts classification results using MI-based feature selection. Once again, RF achieved the best

performance, exceeding 90% accuracy with as few as 20 features, and maintaining this high level across all feature counts. KNN and DT also showed robust performance, stabilizing above 80% beyond 20 features. SVM exhibited modest improvement with increasing feature count, plateauing near 65% after 60 features, while NB remained the lowest performer (~55-60%). Overall, the MI method effectively identified features that enhanced the performance of ensemble- and distance-based classifiers. It also proved to be the most efficient for dimensionality reduction in EEG-based consumer intent classification. Table II provides a comparative summary of the classification performance achieved using the three feature selection methods. Across all methods, RF consistently achieved the highest accuracy, confirming its robustness.

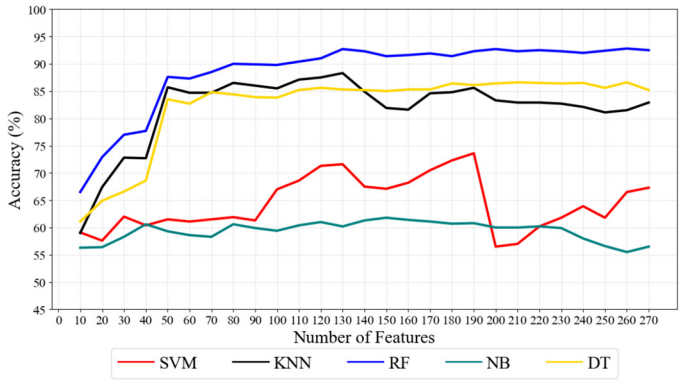

Fig. 6. Classification accuracy using LDA feature selection.

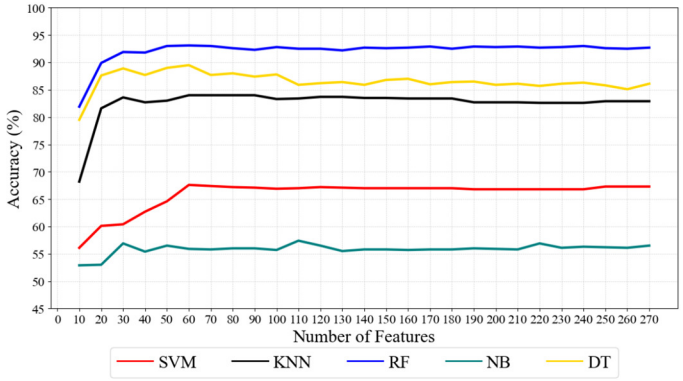

Fig. 7. Classification accuracy using MI feature selection.

TABLE II. PERFORMANCE COMPARISON OF FEATURE SELECTION METHODS

Description	PC	LDA	MI
Best classification algorithm.	RF	RF	RF
Minimum number of features required to achieve 80% accuracy.	20	50	10
Minimum number of features required to achieve 90% accuracy.	80	80	20
Number of features required to achieve the maximum accuracy.	120	130	50
Maximum accuracy.	93%	93%	93%

A one-way ANOVA performed across the five classifiers revealed a significant main effect of classifier type on accuracy ($F(4, 1345) = 4611.48$, $p < 0.001$). A Tukey Honestly Significant Difference (HSD) post hoc test showed that RF significantly outperformed all other classifiers ($p < 0.001$), with the highest mean accuracy. All pairwise comparisons were significant ($p < 0.05$), confirming the statistical superiority of RF. Among feature selection techniques, MI demonstrated the highest efficiency, reaching 80% accuracy with only 10 features and 93% with 50 features, compared to 120 and 130 features required by PC and LDA, respectively. These findings validate MI's capability to select a compact and informative feature subset for buying intention classification.

A. *Analysis of Mutual Information-Selected Features*

To evaluate the efficiency of the MI-based feature selection, a multicollinearity analysis using PC was conducted, revealing that only 27 out of 190 possible feature pairs (14.2%) exhibited high correlation ($|R| > 0.95$), indicating relatively low inter-feature redundancy. Although MI does not explicitly account for feature interdependencies, this low redundancy suggests its inherent tendency to select non-redundant features while maximizing relevance to the target class.

Furthermore, to confirm that all the features differed significantly, a one-way ANOVA analysis was performed for the top 20 selected features confirmed that most features differed significantly between classes ($p < 0.05$), validating their discriminative power. These combined findings justify that the presence of 14.2% highly correlated pairs remains acceptable, as those features retain meaningful statistical contribution to classification. Collectively, these results confirm that MI effectively produces a compact, non-redundant, and discriminative feature set.

B. *Neurophysiological Interpretation of Top Features*

Table III presents the top 20 features ranked using the MI feature selection method, with each feature labeled in the format Feature.Channel.Band. This notation reflects the type of feature, the EEG channel from which it was extracted, and the corresponding frequency band. A noticeable pattern emerges from the rankings: features derived from the gamma and beta bands overwhelmingly dominate the top positions. Notably, many of these features originate from the occipital region (O2) and the frontal region (Fp1, Fp2, F8), indicating their importance in distinguishing buying intention. The presence of O2 without O1 can be explained by hemispheric lateralization of visual processing. The right occipital region is often associated with stronger visual and attentional responses [10], which may be more informative for capturing differences in buying intention. Although O1 also reflects visual activity, O2 appears to provide richer discriminative information in this context. Similarly, the selection of F8 without F7 reflects functional asymmetry in the frontal lobes, indicating that the right frontal region is more relevant than the left in differentiating buying intention.

The dominance of gamma-band features highlights the role of high-level cognitive functions such as attention, perceptual integration, and emotional evaluation, while beta-band activity reflects active thinking and decision-making. The strong contribution of the occipital O2 channel suggests a right-hemisphere dominance in visual and attentional processing. Similarly, the prominence of right frontal sites (F8, Fp2)

indicates lateralized involvement of the prefrontal cortex in evaluative and motivational decision-making.

TABLE III. TOP 20 SELECTED FEATURES USING THE MI METHOD

No	Feature	MI Score
1	En.O2.gamma	0.1449
2	Std.O2.gamma	0.1430
3	Act.O2.gamma	0.1429
4	DE.O2.gamma	0.1427
5	Act.O2.beta	0.1324
6	DE.O2.beta	0.1323
7	Std.O2.beta	0.1314
8	MAV.O2.gamma	0.1286
9	En.O2.beta	0.1223
10	DE.Fp1.gamma	0.0993
11	Act.Fp1.gamma	0.0992
12	Std.Fp1.gamma	0.0992
13	MAV.O2.beta	0.0990
14	DE.F8.gamma	0.0965
15	SE.O2.beta	0.0964
16	En.Fp1.gamma	0.0961
17	Std.F8.gamma	0.0961
18	Act.F8.gamma	0.0958
19	En.F8.gamma	0.0904
20	DE.Fp2.gamma	0.0899

To visualize spatial patterns, Figure 8 displays scalp topographies of the two highest-ranked features, En.O2.gamma and Std.O2.gamma, averaged across participants. Clear contrasts were observed between high and low buying-intention groups, particularly at O2, where both gamma-band energy and standard deviation were higher in the low-intention group. Specifically, gamma-band energy ranged from 6.79 (min) to 20.15 (max), and standard deviation from 8.6 to 17.3, reflecting greater variability and intensity in occipital activity. Additional spatial differences at Fp2 and F8 reinforce the joint contribution of occipital and frontal regions to consumer intention classification.

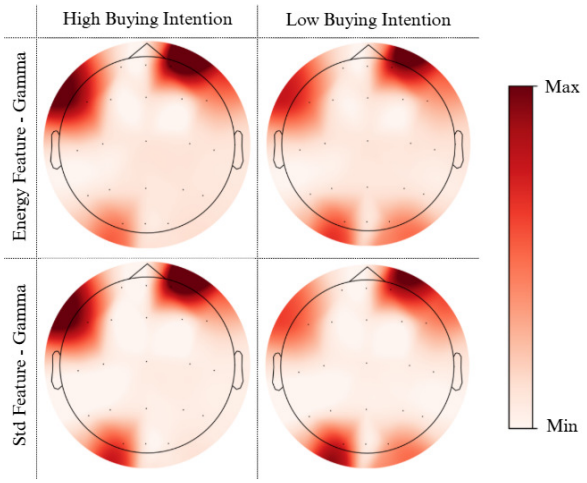

Fig. 8. Illustration of EEG patterns for buying intention.

These neurophysiological findings confirm that MI-selected features are not only statistically discriminative but also spatially and functionally meaningful, aligning with established neural correlates of visual attention and decision-related processing.

IV. CONCLUSION

This study demonstrated that Mutual Information (MI) is the most efficient feature selection method for Electroencephalography (EEG)-based classification of consumer buying intention. MI achieved over 90% accuracy with only 20 features, significantly reducing dimensionality while maintaining high performance. Among the evaluated classifiers, the Random Forest (RF) consistently outperformed others, achieving a peak accuracy of 93%, confirming its robustness and adaptability to complex EEG data structures.

The most discriminative EEG features were predominantly located in the gamma and beta frequency bands, particularly across the frontal (Fp1, Fp2, F8) and occipital (O2) regions. These findings indicate that high-frequency neural oscillations linked to attention, perceptual processing, and decision evaluation play a crucial role in differentiating levels of buying intention. These findings support the use of compact, interpretable EEG models for practical neuromarketing applications.

ACKNOWLEDGMENT

The authors would like to thank the Indonesian Education Scholarship (BPI), Center for Higher Education Funding and Assessment (PPAPT), and the Indonesian Endowment Fund for Education (LPDP) for providing scholarships and supporting this research.

REFERENCES

[1] M. F. K. Khondakar *et al.*, "A systematic review on EEG-based neuromarketing: recent trends and analyzing techniques," *Brain Informatics*, vol. 11, no. 1, Dec. 2024, Art. no. 17, https://doi.org/10.1186/s40708-024-00229-8.

[2] Z. Wei, C. Wu, X. Wang, A. Supratak, P. Wang, and Y. Guo, "Using Support Vector Machine on EEG for Advertisement Impact Assessment," *Frontiers in Neuroscience*, vol. 12, Mar. 2018, Art. no. 76, https://doi.org/10.3389/fnins.2018.00076.

[3] N. K. Horr, K. Han, B. Mousavi, and R. Tang, "Neural Signature of Buying Decisions in Real-World Online Shopping Scenarios – An Exploratory Electroencephalography Study Series," *Frontiers in Human Neuroscience*, vol. 15, Feb. 2022, Art. no. 797064, https://doi.org/10.3389/fnhum.2021.797064.

[4] Z. Xu and S. Liu, "Decoding consumer purchase decisions: exploring the predictive power of EEG features in online shopping environments using machine learning," *Humanities and Social Sciences Communications*, vol. 11, no. 1, Sep. 2024, Art. no. 1202, https://doi.org/10.1057/s41599-024-03691-1.

[5] M. Aldayel, M. Ykhlef, and A. Al-Nafjan, "Recognition of Consumer Preference by Analysis and Classification EEG Signals," *Frontiers in Human Neuroscience*, vol. 14, Jan. 2021, Art. no. 604639, https://doi.org/10.3389/fnhum.2020.604639.

[6] A. Al-Nafjan, "Feature selection of EEG signals in neuromarketing," *PeerJ Computer Science*, vol. 8, Apr. 2022, Art. no. e944, https://doi.org/10.7717/peerj-cs.944.

[7] E. M. Salamone *et al.*, "Abnormalities in Occipital Cortical Sources of Resting-State EEG Rhythms in Patients with Epileptiform Activity and Mild Cognitive Impairment due to Alzheimer's Disease," *Alzheimer's & Dementia*, vol. 19, no. S15, Dec. 2023, Art. no. e080111, https://doi.org/10.1002/alz.080111.

[8] Y. Zhang *et al.*, "Shared oscillatory mechanisms of alpha-band activity in prefrontal regions in eyes open and closed state using a portable EEG

acquisition device," *Scientific Reports*, vol. 14, no. 1, Nov. 2024, Art. no. 26719, https://doi.org/10.1038/s41598-024-78173-0.

[9] M. I. Al-Hiyali, A. J. Ishak, M. S. S. Al-Quraishi, and S. N. Mahmood, "Assessing the Relationship Between Body Mass Index and Neural Activity of Prefrontal Cortex in Overweight Adults Using EEG-Resting State Data: A Wavelet Transform Analysis," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 45, no. 1, pp. 137–153, May 2024, https://doi.org/10.37934/araset.45.1.137153.

[10] M. Iwakiri, Y. Takeo, T. Ikeda, M. Hara, and H. Sugata, "Lateralized alpha oscillatory activity in the inferior parietal lobule to the right hemisphere during left-side visual stimulation," *Neuropsychologia*, vol. 205, Dec. 2024, Art. no. 109017, https://doi.org/10.1016/j.neuropsychologia.2024.109017.

[11] A. H. Jahidin *et al.*, "Classification of intelligence quotient using EEG sub-band power ratio and ANN during mental task," in *2013 IEEE Conference on Systems, Process & Control (ICSPC)*, Kuala Lumpur, Malaysia, Dec. 2013, pp. 204–208, https://doi.org/10.1109/SPC.2013.6735132.

[12] D. L. Schomer and F. H. Lopes Da Silva, Eds., *Niedermeyer's Electroencephalography*, vol. 1. Oxford University Press, 2017.

[13] L. Hu and Z. Zhang, Eds., *EEG Signal Processing and Feature Extraction*. Singapore: Springer Singapore, 2019.

[14] Y. Li, A. Liu, J. Yin, C. Li, and X. Chen, "A Segmentation-Denoising Network for Artifact Removal From Single-Channel EEG," *IEEE Sensors Journal*, vol. 23, no. 13, pp. 15115–15127, Jul. 2023, https://doi.org/10.1109/JSEN.2023.3276481.

[15] S. Leach, G. Sousouri, and R. Huber, "'High-Density-SleepCleaner': An open-source, semi-automatic artifact removal routine tailored to high-density sleep EEG," *Journal of Neuroscience Methods*, vol. 391, May 2023, Art. no. 109849, https://doi.org/10.1016/j.jneumeth.2023.109849.

[16] M. Bullock, G. D. Jackson, and D. F. Abbott, "Artifact Reduction in Simultaneous EEG-fMRI: A Systematic Review of Methods and Contemporary Usage," *Frontiers in Neurology*, vol. 12, Mar. 2021, Art. no. 622719, https://doi.org/10.3389/fneur.2021.622719.

[17] S. Coelli *et al.*, "Selecting methods for a modular EEG pre-processing pipeline: An objective comparison," *Biomedical Signal Processing and Control*, vol. 90, Apr. 2024, Art. no. 105830, https://doi.org/10.1016/j.bspc.2023.105830.

[18] S. Abenna, M. Nahid, H. Bouyghf, and B. Ouacha, "EEG-based BCI: A novel improvement for EEG signals classification based on real-time preprocessing," *Computers in Biology and Medicine*, vol. 148, Sep. 2022, Art. no. 105931, https://doi.org/10.1016/j.compbiomed.2022.105931.

[19] N. S. Amer and S. B. Belhaouari, "EEG Signal Processing for Medical Diagnosis, Healthcare, and Monitoring: A Comprehensive Review," *IEEE Access*, vol. 11, pp. 143116–143142, 2023, https://doi.org/10.1109/ACCESS.2023.3341419.

[20] A. D. Wibawa, S. D. Suryani, and S. Pratasik, "Classifying Electroencephalogram (EEG) Signals via Brain Activity Mapping to Distinguish Identified vs Unidentified Information," *CommIT (Communication and Information Technology) Journal*, vol. 19, no. 1, pp. 101–114, Apr. 2025, https://doi.org/10.21512/commit.v19i1.12500.

[21] S. D. Suryani, A. D. Wibawa, and D. P. Wulandari, "EEG Analysis of Familiar and Unfamiliar Objects Using Wavelet Energy and Shannon Entropy," in *2024 16th International Conference on Knowledge and Smart Technology (KST)*, Krabi, Thailand, Feb. 2024, pp. 226–231, https://doi.org/10.1109/KST61284.2024.10499671.

[22] I. D. Mienye and N. Jere, "A Survey of Decision Trees: Concepts, Algorithms, and Applications," *IEEE Access*, vol. 12, pp. 86716–86727, 2024, https://doi.org/10.1109/ACCESS.2024.3416838.

[23] C. Maçãs, J. R. Campos, N. Lourenço, and P. Machado, "Visualisation of Random Forest classification," *Information Visualization*, vol. 23, no. 4, pp. 312–327, Oct. 2024, https://doi.org/10.1177/14738716241260745.

[24] M. M. Ghiasi and S. Zendehboudi, "Application of decision tree-based ensemble learning in the classification of breast cancer," *Computers in Biology and Medicine*, vol. 128, Jan. 2021, Art. no. 104089, https://doi.org/10.1016/j.compbiomed.2020.104089.

[25] S. Taylor, M. Ponzini, M. Wilson, and K. Kim, "Comparison of imputation and imputation-free methods for statistical analysis of mass spectrometry data with missing data," *Briefings in Bioinformatics*, vol. 23, no. 1, Jan. 2022, Art. no. bbab353, https://doi.org/10.1093/bib/bbab353.

[26] M. A. Alsuwaiket, "Feature Extraction of EEG Signals for Seizure Detection Using Machine Learning Algorthims," *Engineering, Technology & Applied Science Research*, vol. 12, no. 5, pp. 9247–9251, Oct. 2022, https://doi.org/10.48084/etasr.5208.

[27] S. N. Hernández Pérez, F. D. Pérez Reynoso, C. A. G. Gutiérrez, M. D. L. Á. Cosío León, and R. Ortega Palacios, "EOG Signal Classification with Wavelet and Supervised Learning Algorithms KNN, SVM and DT," *Sensors*, vol. 23, no. 9, May 2023, Art. no. 4553, https://doi.org/10.3390/s23094553.

[28] M. A. Khan, "A Comparative Study on Imputation Techniques: Introducing a Transformer Model for Robust and Efficient Handling of Missing EEG Amplitude Data," *Bioengineering*, vol. 11, no. 8, Jul. 2024, Art. no. 740, https://doi.org/10.3390/bioengineering11080740.

[29] G. Cho, J. Yim, Y. Choi, J. Ko, and S.-H. Lee, "Review of Machine Learning Algorithms for Diagnosing Mental Illness," *Psychiatry Investigation*, vol. 16, no. 4, pp. 262–269, Apr. 2019, https://doi.org/10.30773/pi.2018.12.21.2.

[30] G. Feng, H. Wang, M. Wang, X. Zheng, and R. Zhang, "A Research on Emotion Recognition of the Elderly Based on Transformer and Physiological Signals," *Electronics*, vol. 13, no. 15, Jul. 2024, Art. no. 3019, https://doi.org/10.3390/electronics13153019.

[31] G. Nguyen *et al.*, "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 77–124, Jun. 2019, https://doi.org/10.1007/s10462-018-09679-z.

[32] S. Sarkar and K. Mali, "Firefly-SVM predictive model for breast cancer subgroup classification with clinicopathological parameters," *DIGITAL HEALTH*, vol. 9, Jan. 2023, https://doi.org/10.1177/20552076231207203.

[33] M. Rippa *et al.*, "Evaluation of Machine Learning Classification Models for False-Positive Reduction in Prostate Cancer Detection Using MRI Data," *Diagnostics*, vol. 14, no. 15, Aug. 2024, Art. no. 1677, https://doi.org/10.3390/diagnostics14151677.

[34] W. Zouhri, L. Homri, and J.-Y. Dantan, "Handling the impact of feature uncertainties on SVM: A robust approach based on Sobol sensitivity analysis," *Expert Systems with Applications*, vol. 189, Mar. 2022, Art. no. 115691, https://doi.org/10.1016/j.eswa.2021.115691.

[35] D. C. Toledo-Pérez, J. Rodríguez-Reséndiz, R. A. Gómez-Loenzo, and J. C. Jauregui-Correa, "Support Vector Machine-Based EMG Signal Classification Techniques: A Review," *Applied Sciences*, vol. 9, no. 20, Oct. 2019, Art. no. 4402, https://doi.org/10.3390/app9204402.

[36] A. Miltiadous *et al.*, "An Ensemble Method for EEG-based Texture Discrimination during Open Eyes Active Touch," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12676–12687, Feb. 2024, https://doi.org/10.48084/etasr.6455.