BAREKENG: Journal of Mathematics and Its Applications
December 2026 Volume 20 Issue 4 Page 3471-3486
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss4pp3471-3486

PATTERN RECOGNITION FOR RF NEURAL SIGNAL PROCESSING USING KNN DISCRIMINANT CLASSIFICATION

Nur Irsalina Huda Nazri[1], **Muhammad Rasyid Rosli**[2], **Roshakimah Mohd Isa**[3*], **M. N. Ezzuddean Miswan Hanis**[4]

[1,2]*Bachelor of Mechanical Engineering, Faculty of Mechanical Engineering, Universiti Teknologi MARA (UiTM)*
[3]*Department of Mechatronic Engineering, Faculty of Mechanical Engineering, Universiti Teknologi MARA (UiTM)*
[4]*Bachelor of Mechatronic Engineering Technology, Faculty of Mechanical Engineering, Universiti Teknologi MARA (UiTM)*
Shah Alam, Selangor, 40450, Malaysia

*Corresponding author's e-mail: * **roshakimah@uitm.edu.my***

Article Info

Article History:
Received: 30th October 2025
Revised: 23rd April 2026
Accepted: 07th June 2026
Available online: 01st July 2026

Keywords:
Asymmetry;
Classification;
Electroencephalogram;
K-Nearest neighbors;
Radiofrequency;
Synthetic data

ABSTRACT

Radiofrequency (RF) radiation from modern wireless devices has raised concerns about its potential effects on brain activity, especially in the Alpha and Beta frequency bands. However, research in this area faces significant challenges due to small and limited electroencephalogram (EEG) datasets, which often lead to inconsistent results and hinder reliable performance evaluation. This study addresses these limitations by incorporating synthetic data generation to enhance dataset robustness while maintaining realistic signal characteristics. The main objective of this study is to optimize the Power Asymmetry Ratio (PAR) for improved brainwave classification, using a dataset of 97 participants with engineering backgrounds from Universiti Teknologi MARA (UiTM), including both Males and Females, exposed to three types of RF exposure (Left Exposure (LE), Right Exposure (RE), and Sham Exposure (SE)) across two sessions (Before and During). The analysis incorporates advanced signal processing techniques, including ANOVA based feature analysis and K-Nearest Neighbors (KNN) modeling with Mahalanobis distance metric to evaluate gender-specific neural responses. Analysis of RF exposure data reveals clear gender-based patterns in classification performance. With a 70:30 data split for training and testing, in During exposure, Female participants achieved the highest KNN accuracy with 88% for training and 66% for testing when using Mahalanobis distance at K=3, while Male participants reached 80% (training) and 58% (testing) at K=2. The combined (synthetic + actual) data consistently outperformed actual data alone in training. This highlights synthetic data's role in enhancing model robustness, especially for limited datasets. These results demonstrate that well-tuned classification algorithms can detect gender-specific brain responses to RF exposure, with Mahalanobis distance excelling in capturing subtle neural variations. The integration of synthetic data offers a scalable solution to small-data challenges, improving reliability in RF exposure studies while underscoring the need for gender- specific modeling in neural signals.

How to cite this article:

N. I. H. Nazri, M. R. Rosli, R. M. Isa and M. N. E. M. Hanis., "PATTERN RECOGNITION FOR RF NEURAL SIGNAL PROCESSING USING KNN DISCRIMINANT CLASSIFICATION", *BAREKENG: J. Math. & App.*, vol. 20, no. 4, pp. 3471-3486, Dec, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Mobile devices are important to users in today's technological world. Users of nearly every age group engage in continuous use of mobile devices. There is growing discourse on the effects of radiofrequency (RF) exposure on brain function such as memory and attention [1]. Researchers have recently begun utilizing deep neural networks to lend improved classification of Electroencephalogram (EEG) data. EEG is a non-invasive way to record neural signals data from the brain, and due to the dynamics and changes from emotions, sleep quality, and cognitive demands, it is difficult to make confident conclusions. Prior research has established EEG signals are typically described by power spectral analysis, which describes neural activity that varies in terms of frequency bands associated with various cognitive processes and capabilities. These bands are commonly presumed to be Delta waves (0.5 – 4 Hz), referred to as the waves associated with deep sleep; Theta waves (4 – 8 Hz), referred to as the waves associated with drowsiness; Alpha waves (8 – 13 Hz), and the waves primarily associated with relaxation; and Beta waves (13 – 30 Hz), refer to the waves associated with cognitive thought/attention or the general awareness of cognitive processing [2], [3], [4]. There have been many studies focused on Beta and Alpha waves, as they seem to be more related to cognitive functions and tasks in intentional surroundings (real life situations), whereas the other bands tend to be more associated with the processes behind attentiveness [5]. Therefore, EEG-based pattern recognition focusing on these frequency bands provides an approach to identify and classify brainwave changes under RF exposure conditions.

There are several concerns to be aware of when collecting EEG data. Because of variability between people, coupled with variables such as movement and eye blinks, the quality of data greatly diminishes, and the time and expenses associated with complete experiments increases [6], [7]. This high variability in EEG signals introduces non-linear and complex patterns, which require classification models capable of adapting to irregular and dynamic data distributions rather than relying on strict assumptions. The reduced quality of the data is compounded with the restricted number of large-scale EEG datasets. This issue is likely due to a lack of privacy protections, or because of limited resources, especially in studies that measure RF exposures [8], [9], [10], [11]. Small samples bias and limit the statistical analyses power and cause an even higher likelihood of machine learning (ML) models memorizing noise as opposed to patterns in training, known as overfitting. This limitation directly affects model generalization, making it necessary to adopt approaches that can perform reliably even with limited and noisy datasets. Small sample sizes also limit the ability to conclude whether the findings from one study will also be replicated in separate studies.

For this task, researchers have put forward that synthetic EEG data in the generation of EEG data can help with sample size problems. This approach is beneficial in model training and development in machine learning, particularly when learning requires large amounts of quality data, as is common for deep learning tasks. In addition, the use of synthetic data supports the proposed modelling strategy by enhancing data diversity while preserving underlying signal patterns, which improves the robustness of classification models. Research has shown that in deep neural networks for tasks of emotion recognition and Brain-Computer Interfaces (BCIs), synthetic data has resulted in improved performance by providing more training information and increasing classification accuracy and reliability [11], [12].

The usefulness of EEG data will depend primarily on the quality of the algorithms used to generate the data, if the data is generated with overly simplistic models, then it may produce unrealistic activity patterns [12]. In this study, synthetic data is specifically utilized as a data augmentation strategy rather than as a standalone modelling substitute, where the generated data serves to expand the training dataset while preserving the characteristics of the original EEG signals. Thus, feature extraction methodologies are appropriate to maintain accuracy and reliability for real and synthetic data prior to undertaking any classification. The EEG signals will go through preprocessing methodology to maximize accuracy of the signal, whilst decreasing artefacts. This will include outlier detection, identification and normalization of data. Once pre-processed, feature extraction derives meaningful metrics from the spectral data. Power Asymmetry Ratio (PAR) identifies differences in average power levels of the frequency bands. This allows assessment of differential brain activity between the hemispheres. It gauges relative strength of brain impulses in identically positioned areas of the left and right hemispheres and indicates asymmetric feature in the understanding of the functional dynamics of the brain [13], [14], [15], [16], [17]. This approach has been used by researchers in prior studies using EEG with the objective of classifying mental states, predicting motor activities, and assessing cognitive task performance.

Validation of the synthetic EEG data must be done prior to more detailed analysis to ensure proper identification and examination of data reliability, particularly to confirm that the generated data maintains physiological plausibility in terms of statistical distribution and feature behavior. As noted previously, Pearson Correlation analysis is a basic, well-recognized method. It is a measure of the strength and direction of association between two continuous variables and is appropriate for comparing the actual EEG features and synthetic data features [18], [19], [20]. A very high Pearson correlation coefficient between the actual and synthetic PAR values indicates that they bear high similarity in terms of trend or structure. In addition to correlation analysis, statistical consistency is further supported through variance-based validation methods, ensuring that the synthetic data does not introduce unrealistic deviations from expected EEG patterns.

Statistical methods such as Analysis of Variance (ANOVA), are commonly used to determine if there are statistically significant differences in PAR values across groups. In this study, ANOVA is primarily employed as a validation and analytical tool to confirm the presence of meaningful differences in neural response patterns between RF exposure conditions. One drawback of ANOVA is that it expects all groups to have equal variance, or homogeneity of variances [21]. If all groups do not have equivalent variance, then standard ANOVA tests, such as one-way ANOVA, can have misleading results through incorrect identification of differences. Therefore, based on studies, Levene's Test is a prerequisite for testing whether group variances are equal [21], [22], [23]. With unequal variances, Welch's ANOVA will yield results that are good at controlling error rates and retaining statistical power [24], [25]. This analysis ensures statistical analysis methods used are valid and accurate conclusions can be drawn concerning neural signal processing. Linear Discriminant Analysis (LDA) is used to evaluate and maximize the separability of neural response patterns. LDA has been used in previous research to improve existing BCI systems and improve emotion recognition in the health care field [26], [27]. It has also shown resilience in the face of noise and artifacts as it has been effective at allowing separation of brain signals from irrelevant information. LDA allows for classification of neural response patterns based on the PAR features.

Machine learning is important to analyze all the data effectively. Popular for simplicity and efficiency, the K-Nearest Neighbour (KNN) algorithm is a widely used classification method in machine learning. Usually employing distance measures as Euclidean, Manhattan, or Mahalanobis, KNN classification functions by comparing a new data point to previously categorized data based on similarity [28], [29]. KNN is considered a "lazy learning" method because it does not require model training, which makes it adaptable to new data without retraining [30]. Classification performance is largely influenced by the distance metric chosen as well as by the value of "K." Studies has demonstrated that KNN performs well on both noisy and clean datasets as well as consistently shows great dependability when tested using accuracy, precision, recall, and F1-score [31], [32]. Its flexibility and robustness make KNN a suitable method for classifying complex neural signals and integrating new data into dynamic systems.

2. RESEARCH METHODS

2.1 Research Design and Data Preparation

This study uses pre-processed data from previous research, which involved ninety-seven participants with engineering backgrounds from Universiti Teknologi MARA (UiTM), Shah Alam [16]. The baseline brainwave activity of engineering students is influenced by factors such as stress levels, mental workload, and their educational environment [33]. The participants included both Males and Females to represent the population. They were randomly divided into three groups: Left Exposure (LE), Right Exposure (RE), and Sham Exposure (SE). Participants in the LE and RE groups wore mobile phones attached to their left or right ears, while the SE group served as a control with no exposure [15], [16]. This controlled setup enables comparison between RF-exposed and non-exposed conditions for validating classification performance.

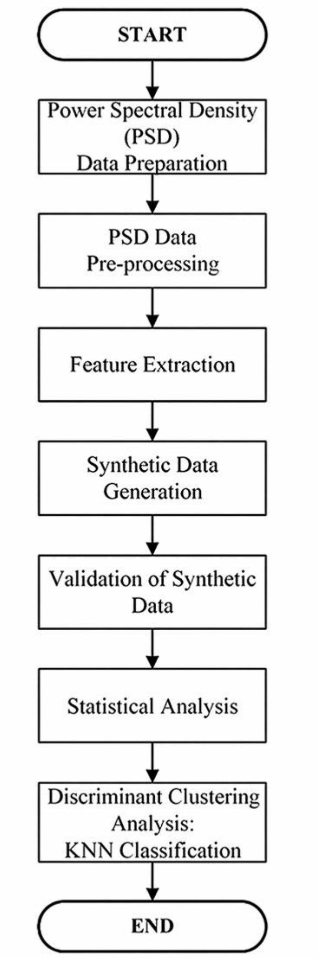


Figure 1. Flowchart of KNN Based Discriminant Classification

The experiment took place in three sessions: Before, During, and After exposure. Each session lasted 5 minutes, with 1-minute breaks in between. During the exposure session, RF exposure was operationally defined based on mobile phone usage conditions, where the researchers attached mobile phones to the participants' heads using a specialized strap. The intensity of RF exposure was controlled by maintaining consistent strap placement across all participants in the Left Exposure (LE) and Right Exposure (RE) groups. In contrast, the Sham Exposure (SE) group served as the control condition, where mobile devices were positioned without direct contact, ensuring no effective RF emission while preserving identical experimental set up conditions. All the raw EEG data was converted into Power Spectral Density (PSD) values using Fast Fourier Transform (FFT) and sorted into Beta and Alpha frequency bands. For this study, the preprocessed PSD data was organized by gender to further examine the differences in brain activity response to RF exposure. This structured data organization enhances the discriminative capability of the model by incorporating biologically relevant grouping factors. Fig. 1 shows the flowchart of KNN based discriminant classification for the EEG data analysis which outlines the important workflow: PSD data preparation, pre-processing, feature extraction, synthetic data generation and validation, statistical analysis, and KNN discriminant classification.

2.2 Data Pre-Processing

Before further analysis, extreme outliers that exceeded normal ranges were filtered out from the dataset to reduce potential distortions such as measurement artifacts and ensure analysis reliability. These outliers were identified using the Statistical Package for the Social Sciences (SPSS), which is known as IBM SPSS Statistics software, through boxplot analysis. This method is based on the Interquartile Range (IQR) technique, where the dataset distribution is divided into quartiles, and outliers are defined using a statistical threshold. Specifically, any data point that falls below $Q1 - 1.5 \times IQR$ or above $Q3 + 1.5 \times IQR$ is automatically classified as an extreme outlier and flagged for review. This criterion ensures that only statistically significant deviations from the central distribution are considered as outliers rather than normal variability in EEG signals. The cleaned PSD data were then normalized to remove scale differences across datasets. This normalization process ensured that all the data were standardized to a common scale for a consistent comparative analysis. The normalization was performed by dividing each PSD value (X) by the maximum PSD value ($Xmax$) in the dataset, as expressed by Eq. (1).

$$Normalization, X_{norm} = \frac{X}{X_{max}}. \tag{1}$$

2.3 Feature Extraction

Power Asymmetry Ratio (PAR) was computed within the processing stage to determine the activity of the right and left hemispheres of the brain. The PAR was intended to identify the shift of power between each of the hemispheres of the brain-body, inclusive of the left and right sides. Positive PAR values reflect right hemispheric dominance, whereas negative values indicate left hemispheric dominance. Eq. (2) defines an asymmetry formula to qualify the power in the left hemisphere, which is the PSD value on the left as PL, and in PSD value on the right hemisphere as PR.

$$PAR_{absolute} = \left|\frac{PR-PL}{PR+PL}\right|. \tag{2}$$

PAR ratios were converted to absolute values to stabilize variance for statistical comparisons between exposure groups (LE, RE, SE). This ensured consistent comparison across participants and reduced the variability caused by individual differences in baseline hemispheric dominance, which improved feature consistency for classification modeling. The use of PAR as a feature is particularly effective as it enhances class separability, making it suitable for comparison across different classification models such as KNN, LDA, and SVM.

2.4 Synthetic Data Generation

Synthetic data was generated by introducing ±10% random noise to the original PSD values using MATLAB R2024a to create variations that maintained the inherent patterns of the neural activity data while simulating natural biological variability. The selection of the ±10% noise range was carefully determined to balance variability and signal integrity, where lower noise levels ±5% would produce minimal variation and limit model robustness, while higher levels ±20% or more could distort the physiological characteristics of EEG signals and introduce unrealistic patterns. Therefore, the ±10% range was considered sufficient to simulate natural fluctuations without compromising the underlying signal structure. The MATLAB code implemented a reproducible randomization process to ensure consistency, where each original data point was multiplied by a random factor within the specified noise range. The synthetic dataset (containing 89 samples) and the original dataset (containing 92 samples) went through the same preprocessing steps and were compared for further validation with Pearson Correlation analysis and statistical tests, including Levene's test and Welch's ANOVA. In addition to correlation-based validation, distributional similarity between synthetic and real data was also considered to ensure that the generated data reflects realistic statistical properties, including comparable variance and spread across exposure groups. This confirmed that the patterns were preserved and that the statistical similarities were valid. The inclusion of synthetic data improves the robustness and generalization capability of the classification models, allowing a more reliable comparison of model performance.

2.5 Signal Processing Statistical Analysis

This study used ANOVA to investigate the significant differences in PAR values among the LE, RE, and SE exposure groups. Equality of variances was validated using Levene's Test. Welch's ANOVA was used if uneven variances were discovered. Additionally, LDA assessed group differences by grouping brain response patterns based on PAR behavior. In previous EEG classification studies, LDA has been widely used due to its simplicity and low computational cost. However, it assumes linear separability of data, which may limit its performance when handling complex and non-linear EEG patterns. A comparative analysis of brainwave asymmetry patterns Before and During exposure was made possible by data visualisation tools, including scatter graphs and bar charts. These approaches identified alterations in brain activity due to exposure and differentiated gender-specific responses, providing evidence of the impact of mobile phone RF radiation.

2.6 KNN Modelling Classification

Before classification stage, the synthetic and actual datasets were combined and pre-processed similarly. Research has shown that KNN can achieve high accuracy, sometimes reaching up to 98% under controlled conditions. When compared to more complex classifiers such as Support Vector Machines (SVM) and Artificial Neural Network (ANN), KNN often delivers comparable results in cases involving noisy or nonlinear signal patterns. Furthermore, its consistency across various datasets suggests strong generalization capability. Hence, these evaluation metrics not only validate the classification performance but also guide the model selection process in EEG-based systems [33], [34]. The KNN method was carried out using MATLAB R2024a and SPSS for the pattern detection of processed neural data, selected for its effectiveness in managing non-linear interactions. In this method, each training sample is represented as $x_i \in \mathrm{R}^d$ and the query sample as $x \in \mathrm{R}^d$, where d denotes the number of features. Distance metrics, including Mahalanobis, Manhattan, and Euclidean, were evaluated to optimize the model's classification performance, where the Euclidean distance was used as the baseline to measure similarity between samples as Eq. (3):

$$d(\boldsymbol{x}, \boldsymbol{x_i}) = \sqrt{\sum_{j-1}^{d} (x_j - x_{ij})^2}. \tag{3}$$

In addition, to account for correlations between features, the Mahalanobis distance was also considered denoted by Eq. (4), which incorporates the covariance structure of the dataset:

$$d(\boldsymbol{x}, \boldsymbol{x_i}) = \sqrt{(x - x_i)^T \boldsymbol{S}^{-1} (x - x_i)}\ , \tag{4}$$

where **S** is the covariance matrix of the dataset.

The dataset was split using a 70:30 training-testing split ratio to validate generalization capability while maintaining sufficient training samples. Performance was assessed through standard classification metrics, including accuracy, precision, recall, and F1-score, with the confusion matrix providing detailed error analysis. X is the input variables for the model (alpha and beta), and Y is the output of the classification (group of exposure).

Accuracy quantifies the overall proportion of correct predictions (both true positives and true negatives) among all cases. *TP* (True Positives) are correctly predicted positive cases, *TN* (True Negatives) are correctly predicted negative cases, *FP* (False Positives) are negative cases incorrectly predicted as positive, and *FN* (False Negatives) are positive cases incorrectly predicted as negative [35]. It provides a general performance overview, such as in Eq. (5).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}. \tag{5}$$

Precision evaluates the model’s ability to avoid false alarms by calculating the proportion of true positive predictions *TP* among all cases predicted as positive, while *FP* reflects negative cases wrongly classified as positive [35], such as in Eq. (6).

$$Precision = \frac{TP}{TP+FP}. \tag{6}$$

Recall quantifies the model’s capacity to detect actual positive cases by measuring the proportion of true positives *TP* among all real positive cases, while *FN* denotes positive cases the model failed to detect, incorrectly labeling them as negative [36], such as in Eq. (7).

$$Recall = \frac{TP}{TP+FN}. \tag{7}$$

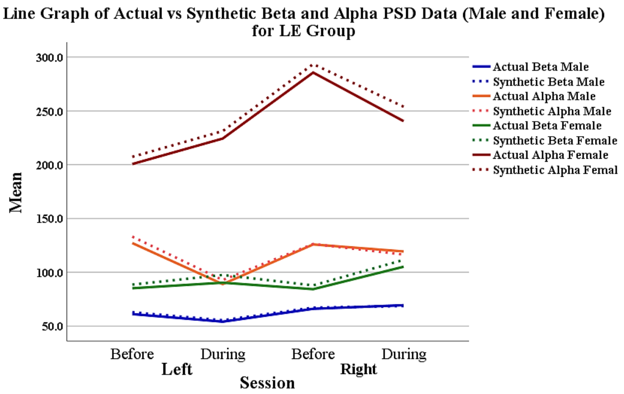
F1-score balances precision and recall through their harmonic mean value, penalizing extreme values in either metric, such as in Eq. (8).

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{8}$$

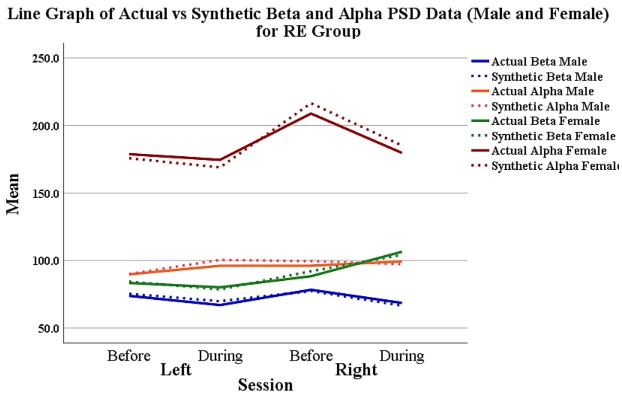
3. RESULTS AND DISCUSSION

3.1 Pearson Correlation Analysis for Data Validation

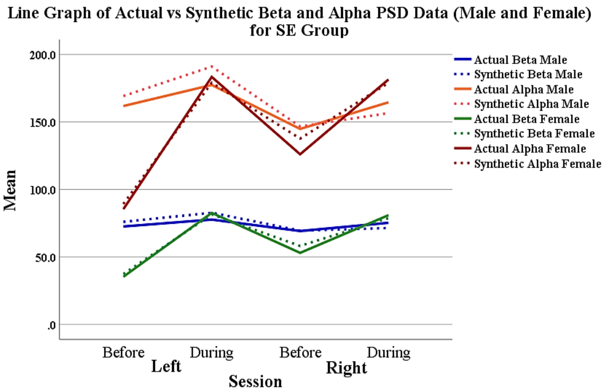
The line graphs in Fig. 2 indicate a strong correlation between the actual data (92 samples) and the synthetic data generated (89 samples) for Beta and Alpha power spectral density (PSD) across all exposure groups (LE, RE, SE) for both Male and Female participants. Although 92 samples were initially generated, three samples were removed due to outlier characteristics. The outliers were excluded from the dataset to prevent distortion of the underlying the brainwave pattern and to maintain the reliability of the PSD analysis and classification performance. A comparison assessment was then conducted between actual data with the synthetic data. It was observed that the synthetic data closely replicate the temporal and amplitude pattern of the original signals. Therefore, the synthetic signals will carry the same traits as original signals. In addition, Pearson correlation values presented in Table 1 further support this visual pattern, displaying consistently increased R-values (all > 0.93) for both frequency bands (alpha and beta) and genders (Male and Female), hence confirming a nearly perfect linear relationship. These consistently high correlations confirm an almost perfect linear relationship, showing that the synthetic data successfully capture both amplitude features and temporal dynamics of the real data across all exposure groups (LE, RE and SE) for session Before and During.



(a) Left Exposure (LE) Group



(b) Right Exposure (RE) Group



(c) Sham Exposure (SE) Group

Figure 2. Comparison of Actual and Synthetic Data for Beta and Alpha PSD (Male and Female) Across Groups (a) LE, (b) RE, and (c) SE

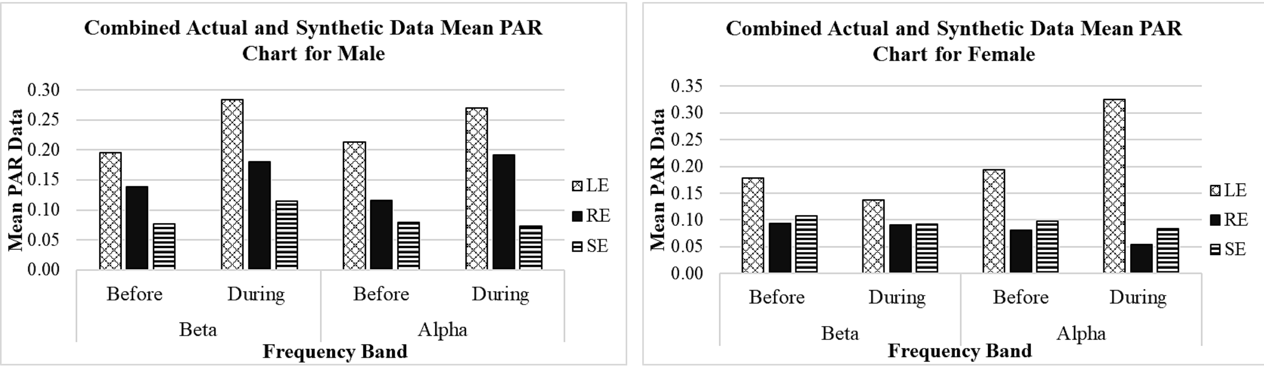
It was observed in Table 1 that the correlation R-values are all close to 1 and statistically significant as all values remain well above the threshold for strong significance ($p < 0.001$). These findings correspond with known research that quantifies the strength and direction of the linear relationship between two continuous variables, hence validating the effectiveness of this method in comparing real and synthetic EEG signal data through Pearson correlation analysis [18]. The consistent results across groups and frequency bands show the reliability of the synthetic data generation process, confirming its feasibility as an alternative for actual data sets when correlations are properly validated. Therefore, it is important to perform the data validation before proceeding with further analysis.

Table 1. Pearson Correlation Analysis Results Comparing Actual and Synthetic Data

Group	Frequency Band	Pearson R-value		Strength of Correlation
		Male	Female	
LE	Beta	0.985	0.980	Strong, Significant
	Alpha	0.992	0.984	Strong, Significant
RE	Beta	0.980	0.978	Strong, Significant
	Alpha	0.979	0.983	Strong, Significant
SE	Beta	0.939	0.980	Strong, Significant
	Alpha	0.976	0.982	Strong, Significant

3.2 Power Asymmetry Ratio (PAR) Analysis of Combined Actual and Synthetic EEG Data

The combined actual and synthetic PAR data show distinct patterns for all the three groups (LE, RE and SE), as shown in Fig. 3 (a) for Males and Fig. 3 (b) for Females across various frequency bands and exposure settings. It is observed that, for Male participants in LE, both Alpha and Beta activities show an increase from session Before to During RF exposure.



(a) Male (b) Female

Figure 3. PAR Values for Combined Actual and Synthetic Data: (a) Male, (b) Female

The increase in Alpha and Beta activity During left-side exposure supports previous studies, wherein increased Alpha activity shows increased relaxation in the left hemisphere, while increased Beta activity could be due to a delayed cognitive response or mild stress rebound [33]. On the other hand, the right hemisphere is associated with creativity, spatial skills, and emotional processing [33]. This explained its consistency in Alpha activity and an increase in Beta response. Small variations in SE results are expected and proven in studies, as this control condition contains no active RF exposure to trigger significant neural responses [37]. Besides, for Females, the pattern is a little different but shows more significant changes in both Beta and Alpha compared to Males. This difference shows a reactive adjustment to RF stimulation, where even though the absolute numerical values of brainwave measurements decreased in some conditions, the change from Before to During exposure is meaningful, as directionally opposite effects may reflect individual variability [38].

Beta activity in Females shows a moderate increase During LE but remains lower than Alpha. During RE, both Alpha and Beta display minimal changes, which reflect stability like the Male group and less sensitivity in the right hemisphere for relaxation effects. However, During SE, Females demonstrate very small differences between Before and During sessions, with the similar expectation that no actual RF exposure results in minor variations. Overall, the results confirm that LE produces more significant changes in both Alpha and Beta activity, particularly in Females, and this aligned with previous findings of hemisphere-specific effects and individual differences in RF response.

3.3 Welch's ANOVA Test

As shown in Table 2, Welch's Analysis of Variance (ANOVA) was employed to determine significant differences in Alpha and Beta brainwave activities across exposure groups (LE, RE, SE) in Before and During RF exposure sessions. This method was specifically chosen to account for potential heterogeneity of variances between groups, to ensure stronger results even when standard ANOVA assumptions were violated [23]. Results in Table 2 are from the combine brainwave dataset of both actual and synthetic signals, which shows significant differences in neural responses between Male and Female participants for both Before and During exposure conditions.

Table 2. Welch's ANOVA Results for Beta and Alpha Bands Across Sessions

Session	Frequency Band		Male	Female
			Sig.	
Before	Beta	Between Groups	0.019	0.033
	Alpha		0.052	0.008
During	Beta		0.006	0.263
	Alpha		0.003	<0.001

In Before session, significant differences between Male and Female participants are observed in both Beta and Alpha frequency bands. This validates earlier studies demonstrating that spontaneous variability in individual reactions to environmental stimuli allows brainwave variations even Before the exposure [35]. The Welch's ANOVA Before session results indicated a significant effect between groups ($p < 0.05$), confirming that these early differences are statistically meaningful among the groups. During RF exposure session, Male participants show significant alterations in both Beta and Alpha activity, while Female participants show a major significance in Alpha activity, but not in Beta activity. This pattern corresponds with a previous study, which indicated that RF exposure significantly affects Alpha activity more than Beta, especially during the exposure session [35].

The insignificance of Female Beta activity During RF exposure ($p > 0.05$) aligns with previous findings indicating that beta waves are not consistently influenced During the exposure session. However, the significant Alpha activity observed here supports the conclusion that Alpha waves are more sensitive indicators of RF-induced neural modulation [35]. Hence, this approach is very useful because this analysis helps identify important patterns early on, which is before moving to the next step, such as KNN classification. As the combined data shows strong and clear differences in neural activity, these results lead to better models, improved signal processing methods, and more reliable classification outcomes, to ensure more accurate results in engineering applications.

3.4 Linear Discriminant Analysis (LDA)

The scatter plots of Linear Discriminant Analysis (LDA) in Fig. 4 show clear patterns in group separation between the Before and During RF exposure sessions. For the Male visualization output, in the Before session, the three groups showed moderate overlap, with the left exposure (LE) group showing some spread but remaining partially mixed with right exposure (RE) and sham exposure (SE) groups, as shown in Fig. 4 (a) for Before Male. During the exposure, the LE group demonstrates a more pronounced shift, clustering further apart from the other groups, which show a stronger response to the RF signals, as shown in Fig. 4 (b) During Male. This movement aligns with the engineering background of most participants, who may be more reactive to such stimuli. It was observed that the RE group remains stable, while the SE group maintains its position, indicating minimal influence from the exposure.

From the Female visualization output, a similar pattern was observed where the LE group also shifts away from the other two groups during the exposure session. In Fig. 4 (c) Before Female, the LE group appears moderately overlapped with RE and SE, but in Fig. 4 (d) During Female, the shift becomes more pronounced. When comparing Male and Female participants, it appears that the Female group demonstrates a larger effect from the RF signals, particularly in the LE group. The LE centroid is located further along Function 1, highlighting its distinct response, while the RE and SE centroids are in close proximity showing similar average responses among the participants. This shows LE has a more significant impact on brain activity compared to RE and SE group. The shift in the Female left exposure group is more extensive, with increased dispersion and separation from the other groups. The LE group centroid also spread further away from the RE and SE group centroids. This wider spread indicates a broader range of neural response patterns among Female participants under the RF exposure thus showing the effect due to the condition.

In contrast of the of the significant pattern, the Male group shows separation in the LE group, but the clustering is tighter and less spread out. While the Male group shows consistent direction in group movement, the Female group reveals a stronger and more varied reaction. This suggests that female participants may exhibit higher sensitivity or more complex neural responses to RF signals, resulting in a greater overall effect. The LDA results visually support and reinforce the earlier findings from the PAR data and Welch's ANOVA, which both showed significant neural activity changes, especially in the Alpha and Beta bands During LE and particularly for Female participants. The clear group separations observed in LDA correspond with the statistical differences reported, confirming the neural modulation patterns identified. The consistent direction of group shifts and clustering further validates the significance of changes noted in the combined data, improving the interpretability of the results. This flow from statistical testing (Welch's ANOVA), followed by visual pattern recognition (PAR data), and finally confirmed through classification visualization (LDA), increases the robustness of the analysis by triangulating findings across multiple validated methods.

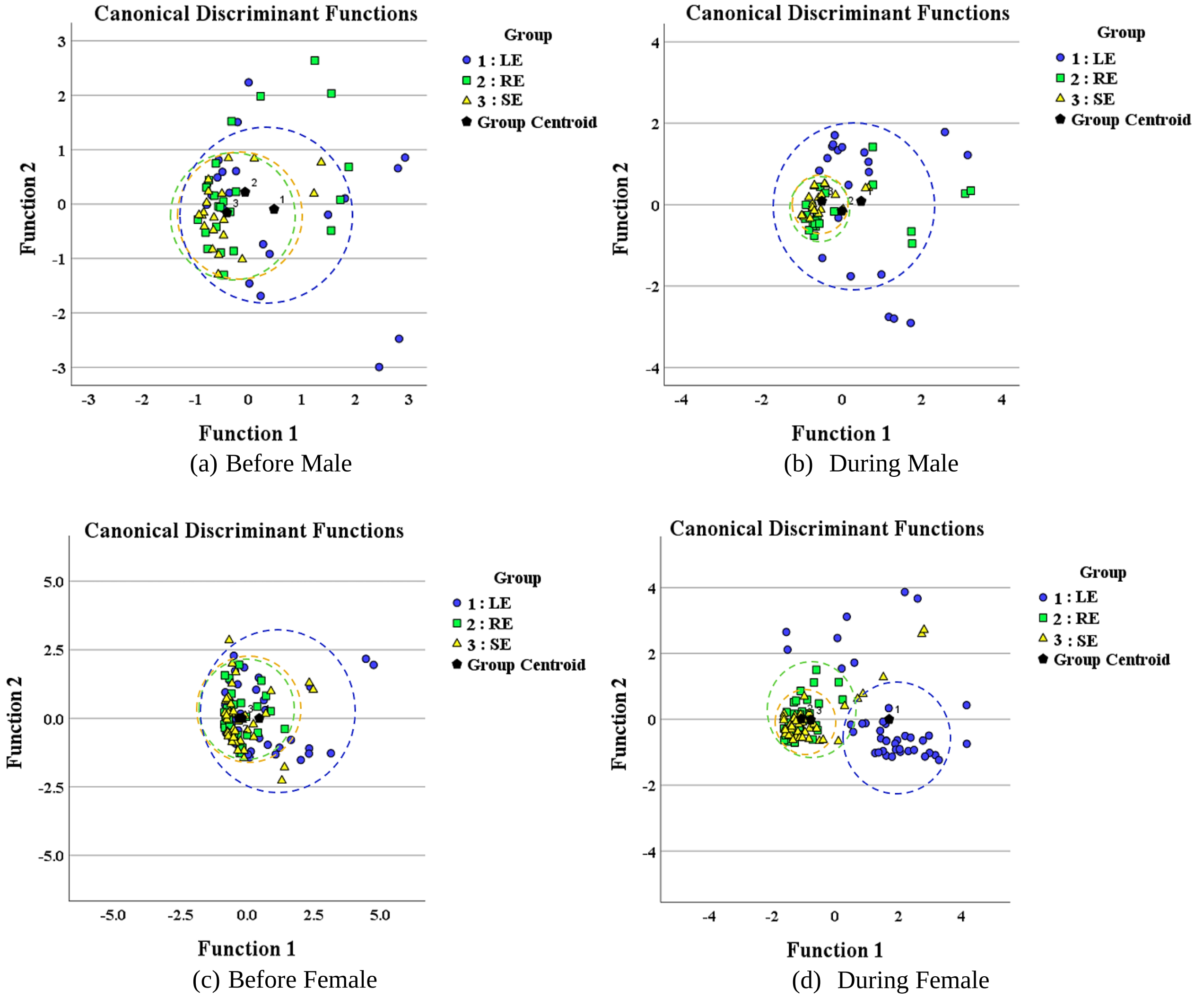


Figure 4. Canonical Discriminant Function Plots from Linear Discriminant Analysis (LDA) for (a) Before Male, (b) During Male, (c) Before Female, and (d) During Female

3.5 K-Value Optimization

The analysis in Table 3 employed a 70:30 data splitting approach for the training and testing datasets to ensure a reliable evaluation of KNN performance across different K-values. This strategy helps prevent overfitting by providing a more balanced and representative model evaluation. Identifying the optimal K-value is a crucial step before conducting classification evaluation, as it directly influences the performance and reliability of the KNN model. An appropriate K-value ensures the model does not overfit to noise (if K is too low) or underfit the data (if K is too high). The analysis of KNN classification performance across various K-values using the Mahalanobis distance metric showed distinct gender-specific patterns in accuracy. In the Before-exposure session, Male participants achieved the highest accuracy of 53% at K = 3, showing that a relatively small neighborhood size led to the best classification performance for this group. In contrast, Female participants showed their peak performance of 40% at K = 4, which suggests that a slightly higher neighborhood size is better suited to capturing their data characteristics prior to exposure. These show that optimal K-values may differ by gender, even within the same session, emphasizing the importance of tailoring the model to subgroup-specific patterns.

During the RF exposure session, both Male and Female participants showed increased classification accuracy when using Mahalanobis distance, highlighting how the RF exposure influenced brainwave patterns and model performance. Males reached their highest accuracy of 58% at K = 1, 2, 3, 5, 8, indicating consistent performance across several K-values but with a slight advantage at lower values. Females, on the other hand, peaked at 66% with K = 3 and K = 10, suggesting that a wider range of neighborhood sizes may effectively capture changes in their brain signals During exposure. These improvements compared to the baseline further support the idea that RF exposure has a stronger impact on Female brainwave patterns, as seen in the accuracy gains. Overall, the selection of the optimal K-value is determined by the point at which the model achieves the highest accuracy, as this reflects the best balance between bias and variance. This accuracy percentage acts as evidence that the selected K-value enables the model to distinguish patterns effectively without overfitting or underfitting.

Table 3. Optimal K-Value Selection in KNN Modeling: Gender Specific Accuracy Rates Using Mahalanobis Distance Metric Across Different Sessions

KNN Performance Measure (%)	Distance Metric	K	1	2	3	4	5	6	7	8	9	10
Before	Mahalanobis	Male	26	42	53	32	47	47	42	37	32	32
		Female	31	31	37	40	29	34	26	29	31	29
After		Male	58	58	58	53	58	53	53	58	53	53
		Female	60	60	66	60	63	60	63	60	63	66

3.6 KNN Classification Performance Analysis

The K-nearest neighbors (KNN) classification analysis was implemented in MATLAB with a 70:30 data splitting method, allocating 70% of the dataset for model training and 30% for performance testing. Optimal K-values were determined by evaluating classification accuracy across a range of neighbor counts (K = 1 to 10) using the Mahalanobis distance metric. Model performance was assessed through multiple evaluation metrics, including classification accuracy, precision, recall, and F1-score, providing a comprehensive assessment of predictive capability. This systematic method enabled the accurate identification of gender-specific optimum parameters while preserving model generalizability across various experimental settings. The following Tables 4 and 5 analysis provides comprehensive performance comparisons between exposure sessions, supported by confusion matrices and quantitative findings for all groups across session Before and During.

Table 4. Confusion Matrices on the Mahalanobis Distance for (a) Before Male, (b) During Male, (c) Before Female, (d) During Female

Male

True Class	LE	RE	SE
LE	10	4	2
RE	6	7	3
SE	4	0	9

Predicted class
(a) Before

True Class	LE	RE	SE
LE	16	0	0
RE	4	12	0
SE	0	5	8

Predicted class
(b) During

Female

True Class	LE	RE	SE
LE	23	5	3
RE	4	18	1
SE	6	5	17

Predicted class
(c) Before

True Class	LE	RE	SE
LE	31	0	0
RE	0	20	3
SE	0	7	21

Predicted class
(d) During

The analysis of evaluation metrics in Table 5 shows distinct performance differences between combined (synthetic and actual) data and actual data alone when using the Mahalanobis distance across different sessions and gender groups. Before exposure, Male participants reached a training accuracy of 58% at K=3 using combined data, slightly higher than 54% at K=5 achieved using actual data. However, testing accuracy remained similar between the two approaches, with 53% (combined) and 55% (actual). This shows that synthetic data helped enhance the learning process slightly during training without negatively affecting testing performance. The confusion matrix in Table 4 (a) shows that although most LE and SE samples were correctly classified, there was still noticeable misclassification between RE and SE classes, contributing to the moderate performance. Interestingly, LE consistently showed the highest correct classification rates across most sessions, suggesting it produced more distinguishable brainwave patterns. This may indicate that RF stimulation on the left side generated clearer signal responses, which the classifier could capture more effectively. For Female participants, the combined data yielded 71% training accuracy at K = 4, a significant improvement compared to 55% at K = 5 with actual data. Although the testing accuracy was lower using combined data (40%) versus actual (53%), the improved training performance suggests that synthetic data contributed positively during the learning session, especially for class pattern generalization. As seen in Table

4 (c), some overlap remained between RE and SE classifications, explaining the lower testing precision and recall.

Table 5. Classification performance metrics evaluation of the KNN model using the Mahalanobis distance

Optimum K – Value		Combined Data				Actual Data			
		Before		During		Before		During	
		Male	Female	Male	Female	Male	Female	Male	Female
		3	4	2	3	5	5	4	5
Training	Accuracy (%)	58	71	80	88	54	55	71	71
	Precision (%)	58	71	79	87	55	55	71	70
	Recall (%)	59	72	84	87	54	55	77	70
	F1-score (%)	58	71	79	87	54	55	71	70
Testing	Accuracy (%)	53	40	58	66	55	53	77	65
	Precision (%)	57	42	60	68	67	53	67	68
	Recall (%)	52	40	64	67	44	52	53	68
	F1-score (%)	53	40	69	67	53	53	58	63

Performance improved substantially During RF exposure sessions, where both genders showed higher metrics across all evaluations. Male participants achieved their highest training accuracy of 80% at K = 2 with combined data, compared to 71% at K = 4 using actual data. However, testing accuracy was lower for combined data (58%) compared to actual (77%), indicating that while synthetic data enhances training effectiveness, it may introduce slight variation affecting test generalization under certain conditions. Nonetheless, the higher recall (84% vs. 77%) and F1-score (79% vs. 58%) in training suggest that the model trained on synthetic-augmented data captured more consistent and complete patterns. Table 4(b) shows a much cleaner classification for LE and SE, indicating that the model captured more structured patterns from the synthetic-augmented dataset. For Female participants During exposure, the combined data led to the best performance overall. Training accuracy peaked at 88% at K = 3, significantly outperforming the 71% at K = 5 obtained with actual data.

More importantly, testing accuracy also improved slightly from 65% (actual) to 66% (combined), with consistent precision (68%) and recall (67%) values. This makes Female participants During exposure the group with the highest classification performance when using Mahalanobis distance and synthetic data. These results are further supported by the clearer diagonal class separations observed in Table 4 (d), which indicate fewer misclassifications and better prediction stability when synthetic data is included. Thus, testing accuracy using combined data may occasionally be lower than with actual data, especially for Male participants, synthetic data consistently enhances training performance. It helps the model better recognize patterns and increases precision and recall in several cases. The best overall result was seen in Female Participants for During RF exposure, where synthetic data improved both training and testing performance. The detailed confusion matrices in Table 4 visually support these findings by highlighting improvements in correct classifications across LE, RE, and SE classes when synthetic data is used. These results affirm the usefulness of synthetic data in EEG classification, especially for improving model robustness when real data is limited or gender-specific variability is significant.

4. CONCLUSION

In conclusion, this study demonstrates that RF exposure significantly modulates brainwave activity, with distinct gender-specific patterns in Alpha and Beta bands. The implementation of synthetic data successfully enhanced both dataset robustness and classification performance, as evidenced by strong Pearson correlations (> 0.93) between actual and synthetic signals. However, it is important to note that these findings are derived from a relatively limited participant sample and are further supported by synthetic data augmentation, which may influence the extent of generalizability to broader populations. KNN discriminant clustering with Mahalanobis distance achieved optimal accuracy (88% for Females, 80% for Males), while revealing clear separation between exposure conditions. These results indicate strong model performance within the defined experimental dataset; nevertheless, the observed classification outcomes should be interpreted within the context of the controlled dataset size and synthetic data dependency. This improvement project effectively addressed small dataset limitations through synthetic data augmentation, enabling more reliable detection of neural response patterns to RF exposure. For future work, several important improvements could advance this research. First, expanding participant diversity beyond engineering students would enhance the generalizability of findings across different populations. Second, exploring alternative machine learning approaches such as support vector machines or neural networks may better capture the complex patterns in neural responses to RF exposure. Third, developing more advanced synthetic data generation techniques could better simulate the full spectrum of individual variability in brainwave responses. Additionally, conducting longitudinal studies would provide valuable insights into the cumulative effects of prolonged RF exposure. Implementing these enhancements would significantly improve the reliability and practical relevance of the results for real-world applications.

Author Contributions

Nur Irsalina Huda Nazri: Conceptualization, Formal Analysis, and Original Draft. Muhammad Rasyid Rosli: Investigation and Methodology. Roshakimah Mohd Isa: Project Administration, Resources, and Supervised. M. N. Ezzuddean Miswan Hanis: Review, Editing and Validation. All authors reviewed and approved the final version of the manuscript prior to publication.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Acknowledgment

The authors would like to express their sincere gratitude to the staff and students of Universiti Teknologi MARA, Malaysia, for their valuable contributions and active participation in this research.

Declarations

The authors declare no conflict of interest.

Declaration of Generative AI and AI-assisted Technologies

AI used only for grammar and language polishing. Generative AI tools (ChatGPT and Gemini) were used solely for language refinement (grammar, spelling, and clarity). The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. The authors reviewed and approved all final text.

REFERENCES

[1] Y. Q. He, "IMPACTS OF RADIO FREQUENCY INTERFERENCE ON HUMAN BRAIN WAVES AND NEURO-PSYCHOLOGICAL CHANGES," *International Conference on Intelligent Informatics and Biomedical Sciences*, pp. 257–261, 2016. doi: https://doi.org/10.1109/ICIIBMS.2015.7439535.

[2] K. Singh, R. Mehra, and Suman, "MOBILE PHONE HANDSET RADIATION IMPACT STUDY ON BRAINWAVE SIGNAL USING EEG," *Int. J. Sci. Res. Dev*, vol. 3, no. 09, pp. 295–298, 2015.

[3] A. Al-Qahtani, A. Nasir, M. Z. Shakir, and K. A. Qaraqe, "COGNITIVE IMPAIRMENTS IN HUMAN BRAIN DUE TO WIRELESS SIGNALS AND SYSTEMS: AN EXPERIMENTAL STUDY USING EEG SIGNAL ANALYSIS," 2013 IEEE 15th International Conference on e-Health Networking, Applications and Services, Healthcom 2013, pp. 1–3, 2013. doi: https://doi.org/10.1109/HealthCom.2013.6720780.

[4] B. S. Zainuddin, Z. Hussain, and I. S. Isa, "ALPHA AND BETA EEG BRAINWAVE SIGNAL CLASSIFICATION TECHNIQUE: A CONCEPTUAL STUDY," Proceedings - 2014 IEEE 10th International Colloquium on Signal Processing and Its Applications, pp. 233–237, 2014. doi: https://doi.org/10.1109/CSPA.2014.6805755.

[5] A. E. Tyas, A. Dharma Wibawa, and M. H. Purnomo, "THETA, ALPHA AND BETA ACTIVITY IN THE OCCIPITAL BASED ON EEG SIGNALS FOR MENTAL FATIGUE IN HIGH SCHOOL STUDENTS," Proceeding - ICoSTA 2020: 2020 International Conference on Smart Technology and Applications: Empowering Industrial IoT by Implementing Green Technology for Sustainable Development, 2020. doi: https://doi.org/10.1109/ICoSTA48221.2020.1570614141.

[6] A. R. Nikolaev, R. N. Meghanathan, and C. van Leeuwen, "COMBINING EEG AND EYE MOVEMENT RECORDING IN FREE VIEWING: PITFALLS AND POSSIBILITIES," *Brain Cogn*, vol. 107, pp. 55–83, 2016. doi: https://doi.org/10.1016/j.bandc.2016.06.004.

[7] N. Khadijah, "SIMULATING BRAIN SIGNALS : CREATING SYNTHETIC EEG DATA VIA NEURAL-BASED GENERATIVE MODELS FOR IMPROVED SSVEP CLASSIFICATION," 2019 International Joint Conference on Neural Networks (IJCNN), no. July, pp. 1–8, 2019. doi: https://doi.org/10.1109/IJCNN.2019.8852227.

[8] S. Chaibi, C. Mahjoub, W. Ayadi, and A. Kachouri, "EPILEPTIC EEG PATTERNS RECOGNITION THROUGH MACHINE LEARNING TECHNIQUES AND RELEVANT TIME–FREQUENCY FEATURES," *Biomedizinische Technik*, vol. 69, no. 2, pp. 111–123, 2024. doi: https://doi.org/10.1515/bmt-2023-0332.

[9] M. Abufadda and K. Mansour, “A SURVEY OF SYNTHETIC DATA GENERATION FOR MACHINE LEARNING,” *22nd International Arab Conference on Information Technology, ACIT 2021*, pp. 1–7, 2021. doi: https://doi.org/10.1109/ACIT53391.2021.9677302.

[10] Md Mahbubur Rahman, “FREQUENCY-BAND BASED SYNTHETIC EEG GENERATION,” Computer, Applied Technology, Information Intelligence, Applied Artificial Faculty, Computer Science, MS Thesis, 2024.

[11] M. Goyal and Q. H. Mahmoud, “A SYSTEMATIC REVIEW OF SYNTHETIC DATA GENERATION TECHNIQUES USING GENERATIVE AI,” *Electronics (Switzerland)*, vol. 13, no. 17, 2024. doi: https://doi.org/10.3390/electronics13173509.

[12] G. Siddhad, M. Iwamura, and P. P. Roy, “ENHANCING EEG SIGNAL-BASED EMOTION RECOGNITION WITH SYNTHETIC DATA: DIFFUSION MODEL APPROACH,” *arXiv:2401.16878*, pp. 1–9, 2024. doi: https://doi.org/10.1109/TAI.2025.3641576.

[13] A. Torkan, M. Zoghi, N. Foroughimehr, A. Yavari, and S. Jaberzadeh, “EFFECTS OF MOBILE ELECTROMAGNETIC EXPOSURE ON BRAIN OSCILLATIONS AND CORTICAL EXCITABILITY: SCOPING REVIEW,” *Sensors (Basel, Switzerland)* 25.9, 2749, 2025. doi: https://doi.org/10.3390/s25092749.

[14] K. Ong and R. Paramesran, “ASYMMETRY RATIO ELECTROENCEPHALOGRAM OF SINGLE TRIAL DURING FINGER MOVEMENT,” TENCON 2005 - 2005 IEEE Region 10 Conference, pp. 0–3, 2005. doi: https://doi.org/10.1109/TENCON.2005.301169.

[15] R. M. Isa, I. Pasya, M. N. Taib, A. H. Jahidin, W. R. W. Omar, and N. Fuad, “ASSESSMENT OF BRAINWAVE ASYMMETRY AND HEMISPHERE DOMINANCE DUE TO RF RADIATION,” Proceedings - UKSim-AMSS 6th European Modelling Symposium, pp. 153–157, 2012. https://doi.org/10.1109/EMS.2012.89.

[16] R. M. Isa, “CLASSIFICATION OF BRAINWAVE ASYMMETRY INFLUENCED BY MOBILE PHONE RADIOFREQUENCY EMISSION,” Procedia Soc Behav Sci, vol. 97, pp. 538–545, 2013. doi: https://doi.org/10.1016/j.sbspro.2013.10.270.

[17] M.C. Yavas, “THE EFFECT OF SHORT-TERM ELECTROMAGNETIC FIELDS CAUSED BY MOBILE PHONES ON THE ELECTRICAL ACTIVITY OF ALPHA AND BETA BRAIN WAVES,” *Annals of Clinical and Analytical Medicine*, vol. 11, no. 05, 2020. doi: https://doi.org/10.4328/ACAM.20220.

[18] J. D. Chee and T. Queen, “PEARSON’S PRODUCT-MOMENT CORRELATION: SAMPLE ANALYSIS,” *ResearchGate*, no. May 2015, 2016.

[19] P. Schober and L. A. Schwarte, “CORRELATION COEFFICIENTS: APPROPRIATE USE AND INTERPRETATION,” *Anesth Analg*, vol. 126, no. 5, pp. 1763–1768, 2018. doi: https://doi.org/10.1213/ANE.0000000000002864.

[20] M. Shaqiri, T. Iljazi, L. Kamberi, and R. Ramani -Halili, “DIFFERENCES BETWEEN THE CORRELATION COEFFICIENTS PEARSON, KENDALL AND SPEARMAN,” *Journal of Natural Sciences and Mathematics of UT*, vol. 15–16, no. 8, pp. 392–397, 2023.

[21] F. Azizi, R. Ghasemi, and M. Ardalan, “TWO COMMON MISTAKES IN APPLYING ANOVA TEST: GUIDE FOR BIOLOGICAL RESEARCHERS,” *SSRN Electronic Journal*, vol. 800, no. July, 2022. doi: https://doi.org/10.2139/ssrn.4221619.

[22] D. Sharma and B. M. G. Kibria, “ON SOME TEST STATISTICS FOR TESTING HOMOGENEITY OF VARIANCES: A COMPARATIVE STUDY,” *J Stat Comput Simul*, vol. 83, no. 10, pp. 1944–1963, 2013. doi: https://doi.org/10.1080/00949655.2012.675336.

[23] N. F. Abdullah and N. Muda, “AN OVERVIEW OF HOMOGENEITY OF VARIANCE TESTS ON VARIOUS CONDITIONS BASED ON TYPE 1 ERROR RATE AND POWER OF A TEST,” *Journal of Quality Measurement and Analysis*, vol. 18, no. 3, pp. 111–130,qw 2022.

[24] H. Liu, “COMPARING WELCH’S ANOVA, A KRUSKAL-WALLI’S TEST AND TRADITIONAL ANOVA IN CASE OF HETEROGENEITY OF VARIANCE ANOVA,” *Theses and Dissertations*, pp. 1–46, 2015.

[25] M. Delacre, C. Leys, Y. L. Mora, and D. Lakens, “TAKING PARAMETRIC ASSUMPTIONS SERIOUSLY: ARGUMENTS FOR THE USE OF WELCH’S F-TEST INSTEAD OF THE CLASSICAL F-TEST IN ONE-WAY ANOVA,” *International Review of Social Psychology*, vol. 32, no. 1, pp. 1–12, 2020. doi: https://doi.org/10.5334/irsp.198.

[26] R. Atangana, D. Tchiotsop, G. Kenne, and L. C. DjoufackNkengfac k, “EEG SIGNAL CLASSIFICATION USING LDA AND MLP CLASSIFIER,” *Health Informatics - An International Journal*, vol. 9, no. 1, pp. 14–32, 2020. doi: https://doi.org/10.5121/hiij.2020.9102.

[27] M. Melinda, M. Oktiana, Y. Yunidar, N. H. Nabila, and I. K. A. Enriko, “CLASSIFICATION OF EEG SIGNAL USING INDEPENDENT COMPONENT ANALYSIS AND DISCRETE WAVELET TRANSFORM BASED ON LINEAR DISCRIMINANT ANALYSIS,” *International Journal on Informatics Visualization*, vol. 7, no. 3, pp. 830–838, 2023. doi: https://doi.org/10.30630/joiv.7.3.1219.

[28] H. Ghorbani, “MAHALANOBIS DISTANCE AND ITS APPLICATION FOR DETECTING MULTIVARIATE OUTLIERS,” *Facta Universitatis, Series: Mathematics and Informatics*, no. October 2019, p. 583, 2019. doi: https://doi.org/10.22190/FUMI1903583G.

[29] N. Ali, D. Neagu, and P. Trundle, “EVALUATION OF K-NEAREST NEIGHBOUR CLASSIFIER PERFORMANCE FOR HETEROGENEOUS DATA SETS,” *SN Appl Sci*, vol. 1, no. 12, pp. 1–15, 2019. doi: https://doi.org/10.1007/s42452-019-1356-9.

[30] K. Taunk, S. De, S. Verma, and A. Swetapadma, “A BRIEF REVIEW OF NEAREST NEIGHBOR ALGORITHM FOR LEARNING AND CLASSIFICATION,” *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, no. Iciccs, pp. 1255–1260, 2019. doi: https://doi.org/10.1109/ICCS45141.2019.9065747.

[31] T. K. Deo, R. K. Deshmukh, and G. Sharma, “COMPARATIVE STUDY OF PERFORMANCE OF K NEAREST NEIGHBOR AND SUPPORT VECTOR MACHINE CLASSIFIERS IN SENTIMENT ANALYSIS,” *International Research Journal of Engineering and Technology*, pp. 501–506, 2024.

[32] M. Vakili, M. Ghamsari, and M. Rezaei, “PERFORMANCE ANALYSIS AND COMPARISON OF MACHINE AND DEEP LEARNING ALGORITHMS FOR IOT DATA CLASSIFICATION,” *arXiv:2001.09636*, 2020.

[33] N. E. Ahmad, M. Rosdi, and R. M. Isa, “CLASSIFICATION OF RF-INDUCED BRAINWAVE ALTERATIONS USING SVM: AN EEG STUDY ON NEURAL ASYMMETRY IN HIGH FREQUENCY BANDS,” *Proceedings of the IEEE*

Conference on Systems, Process and Control, ICSPC, pp. 160–165, 2024. doi: https://doi.org/10.1109/ICSPC63060.2024.10862626.

[34] M. Revathi, A. B. Godbin, S. N. Bushra, and S. Anslam Sibi, "APPLICATION OF ANN, SVM AND KNN IN THE PREDICTION OF DIABETES MELLITUS," *Proceedings of the 2022 International Conference on Electronic Systems and Intelligent Computing, ICESIC 2022*, pp. 179–184, 2022. doi: https://doi.org/10.1109/ICESIC53714.2022.9783577.

[35] S. Sathyanarayanan and B. R. Tantri, "CONFUSION MATRIX-BASED PERFORMANCE EVALUATION METRICS," no. December, 2024. doi: https://doi.org/10.53555/AJBR.v27i4S.4345.

[36] P. Sedgwick, "PEARSON'S CORRELATION COEFFICIENT," Jul. 07, 2012. doi: https://doi.org/10.1136/bmj.e4483.

[37] C. Hu, H. Zuo, and Y. Li, "EFFECTS OF RADIOFREQUENCY ELECTROMAGNETIC RADIATION ON NEUROTRANSMITTERS IN THE BRAIN," *Front Public Health*, vol. 9, no. August, pp. 1–15, 2021. doi: https://doi.org/10.3389/fpubh.2021.691880.

[38] R. J. Croft., "EFFECTS OF 2G AND 3G MOBILE PHONES ON HUMAN ALPHA RHYTHMS: RESTING EEG IN ADOLESCENTS, YOUNG ADULTS, AND THE ELDERLY," *Bioelectromagnetics*, vol. 31, no. 6, pp. 434–444, 2010. doi: https://doi.org/10.1002/bem.20583.