Mobile and Forensics
Vol. 6, No. 1, March 2024, pp. 19-30
ISSN: 2714-6685 (online), 2656-6257 (print)
DOI: http://dx.doi.org/10.12928/mf.v6i1.9056

EEG Classification for Brain Response Analysis through University Website Interface in Yogyakarta Using Naive Bayes and KNN

[1]Faiq Suhail, [2,*]Ahmad Azhari
[1,2]Department of Informatics, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[1]faiq1900018305@webmail.uad.ac.id, [2*]ahmad.azhari@tif.uad.ac.id
*correspondence email

Abstract

One of the challenges high school students face is the abundant availability of information about various campuses through different media, making it difficult to accurately predict their interest in a particular campus. Electroencephalogram (EEG) technology can read human brain activity, such as when students access information on a campus website. The Naive Bayes and K-Nearest Neighbor (KNN) methods can be employed to predict student interest in a campus based on EEG signals recorded while they browse the official campus website. Naive Bayes is known for achieving high accuracy with small datasets, whereas KNN excels at classifying noisy data. These two methods offer variables that can be directly compared. Classification using Naive Bayes and KNN achieved the highest accuracy score of 92%. The most appropriate algorithm is determined by evaluating performance using a confusion matrix. In this case study, Naive Bayes slightly outperformed KNN, as evidenced by precision, recall, and f1-score matrices. The Naive Bayes method resulted in an F1-score of 94%, compared to KNN's 92%.

Keywords: EEG, KNN, Naïve Bayes, User Interface, University Website

1. INTRODUCTION

In the current educational landscape of Indonesia, the utilization of website-based applications has become a fundamental necessity for institutions. These applications serve as a centralized platform for managing various types of information and services, facilitating the connection between users and institutions for a multitude of needs, particularly in the context of new student admissions. The design of these websites is not merely a reflection of the institution's image but also plays a critical role in engaging users with the information presented. Effective website design encompasses several key elements, including visual aesthetics, User Interface (UI), and overall accessibility, which collectively enhance the user experience and maximize the utility of the provided services. [1].

However, a significant challenge faced by students in Indonesia is the limited access to high-quality information necessary for making informed decisions about their educational futures. High school students, in particular, often find it difficult to navigate through the plethora of available data to choose suitable educational institutions and majors. This difficulty is exacerbated by their susceptibility to following transient campus trends and popular majors, driven by easily accessible information across various media platforms [2].

To address these issues, this research explores the implementation of human-computer interaction tools, specifically the Electroencephalogram (EEG), to analyze user interest and engagement when accessing information on university websites. The focus is on Islamic-based private universities in Yogyakarta, where EEG technology is used to capture the brain activity of students

Article History: *Received September 18, 2023; Revised July 17, 2024; Accepted July 28, 2024*

as they interact with the websites' UI/UX designs. This study aims to assess how well these websites convey crucial information about popular majors that align with students' interests, despite the limited time available for engaging respondents.

By employing EEG analysis, this research seeks to provide insights into the effectiveness of website designs in attracting and maintaining user interest. The findings are expected to inform best practices for optimizing university websites to better serve prospective students, ensuring that essential information is readily accessible and engaging, ultimately facilitating more informed decision-making in the admission process.

An electroencephalogram (EEG) is a powerful tool that can produce data reflecting the psychological conditions of the human brain, including its responses when interacting with the user interfaces of educational institution websites. EEG technology interprets brain waves to determine activities occurring in the brain, making it possible to generate data for analyzing students' mental states while they use UI/UX features on several Islamic-based private university websites in Yogyakarta. This analysis aims to enhance user experience and assist future students in making informed decisions based on their interests in specific campuses.

Affective Computing in EEG focuses on detecting emotional signals during human-computer interactions and synthesizing emotional responses from audio-visual stimuli [3].

Prior research has classified data from UI usage derived from questionnaires or raw data from data provider services, and designed UI for websites. Classification involves finding a model or function that differentiates concepts and data classes, enabling the estimation of the class of an object with an unknown class [4]. Classification can be defined as a job that carries out training/learning on a target function that maps each set of attributes (features) to a single number of available class labels [5].

K-Nearest Neighbor (KNN) is a popular classification method for EEG data due to its resilience to noisy data and its ability to maintain accuracy in EEG signal classification [6]. Similarly, Bayesian Classification, particularly Naive Bayes, is effective in pattern recognition and can achieve high accuracy with minimal training data [7]. The Fast Fourier Transform (FFT) method is often used for extracting EEG signal features, with features determined based on the Pearson correlation coefficient [8]. Naive Bayes and KNN are two widely used classification methods for EEG signals. Naive Bayes excels with small datasets and can achieve high accuracy, while KNN is effective for noisy data [9].

Previous studies have shown varying levels of accuracy for these methods. For instance, Naive Bayes achieved a 70% accuracy rate for classifying EEG signals to detect cyber sickness, compared to 40% for KNN [4]. Another study reported that Naive Bayes had an 85.5% accuracy rate in classifying mental workload based on EEG signals, while KNN had 82.5% [10]. Additionally, in classifying sleep disorders based on EEG signals, Naive Bayes and SVM performed well, with 92.5% and 90% accuracy, respectively [11]. Deep Learning algorithms like Convolutional Neural Networks (CNN) have also been used for concentration classification based on EEG, achieving 83.3% accuracy [12].

Both Naive Bayes and KNN are supervised learning methods that use training data for prediction, and they offer comparable decision-making capabilities [7]. Despite extensive literature on UI/UX testing and user interest in websites, the implementation of EEG in this context remains underexplored. Therefore, this research provides an initial approach to understanding the classification of user interests based on EEG signals generated during website interactions. This study specifically examines several private Islamic university websites in Yogyakarta to assess students' interest using EEG data.

2. LITERATURE REVIEW

A. Data Mining

Data mining involves analyzing datasets to discover unexpected relationships and summarize data in novel ways that are both understandable and useful to the data owner. It integrates techniques from multiple disciplines, including artificial intelligence, machine learning, pattern recognition, statistics, mathematics, databases, and visualization. These techniques collectively address the challenge of extracting valuable information from large databases. Essentially, data mining uncovers hidden knowledge within a database, processing and identifying insightful information. Classification is a pivotal technique within data mining, enabling the systematic categorization of data to facilitate meaningful analysis [5].

B. Electroencephalogram

An Electroencephalogram (EEG) is a device that records and interprets electrical signals from brain cell activity. EEG data, usually digitized with a 12-bit Analog-Digital Conversion (ADC), is sampled at frequencies ranging from 100 Hz for spontaneous EEG to several kHz for short latency far-field ERP recordings [3]. The measurements are obtained by placing electrodes on the scalp, resulting in wave-like graphs that provide rich information about brain activity. EEG data must be extracted before processing, often using the Fast Fourier Transform (FFT) to convert time-domain signals into the frequency domain.

Brain waves are categorized into five types based on frequency ranges [14]:

1. **Delta (1-4 Hz):** Associated with deep, dreamless sleep, this phase facilitates physical and mental rest and self-healing.
2. **Theta (4-8 Hz):** Linked to light sleep or drowsiness, indicated by slow, deep breathing.
3. **Alpha (8-12 Hz):** Present during relaxed, conscious states with closed eyes, often during transitions to sleep.
4. **Beta (12-19 Hz):** Occurs during full consciousness and high concentration, such as during exercise or study.
5. **Gamma (19-100 Hz):** Associated with full concentration and high cognitive function, often during moments of fear or panic.

C. User Interface

A User Interface (UI) is the interactive component of a system that allows users to engage with a device or application. Effective UI design strives to achieve optimal usability, enabling users to interact with the system intuitively. UI serves as a direct communication channel between the system and the user, necessitating a design that ensures ease of use. Critical aspects of UI design include visual appearance and user accessibility, which are essential for attracting users and facilitating seamless interaction [1]. A well-designed UI not only engages users but also maximizes the functionality of the website, acting as a bridge between students and educational institutions. The UI should provide easy access to services offered by the website, ensuring that users can efficiently utilize the platform to meet their needs.

3. METHODS

A. Research Framework

The research process is organized into four key stages: Data Acquisition, Feature Extraction, Classification, and Evaluation. Each stage is meticulously designed to ensure a comprehensive analysis of EEG data to assess user interest in university websites.

1. Data Acquisition:
    a. Tool: Mindwave Mobile 2
    b. Description: The Mindwave Mobile 2 device is employed to gather EEG signal data from respondents. This tool captures the raw EEG signals necessary for subsequent analysis.
    c. Purpose: To collect foundational EEG data that reflects the brain activity of respondents as they interact with university websites.
2. Feature Extraction:
    a. Tool: MATLAB software
    b. Algorithm: Fast Fourier Transform (FFT)

   c. Description: MATLAB software is utilized to perform feature extraction on the raw EEG data using the FFT algorithm. This process converts time-domain signals into the frequency domain, highlighting significant features.
   d. Purpose: To transform and distill raw EEG signals into a format that reveals critical frequency-domain features, facilitating deeper analysis.
3. Classification:
   a. Tools: Python software
   b. Algorithms: Naive Bayes and K-Nearest Neighbors (KNN)
   c. Description: The extracted features are classified using Naive Bayes and K-Nearest Neighbors (KNN) algorithms implemented in Python. These algorithms categorize the EEG signals to predict user interest and engagement.
   d. Purpose: To classify EEG signals effectively, enabling the prediction of user interest based on their interactions with university websites.
4. Evaluation:
   a. Tool: Confusion Matrix
   b. Description: The performance of the classification algorithms is assessed using a confusion matrix. This evaluation method measures the accuracy, precision, recall, and f1-score of the classification system.
   c. Purpose: To rigorously evaluate the effectiveness of the classification algorithms, ensuring reliable prediction of user interest through detailed performance metrics.

B. Data Acquisition

The research began with an initial survey where students filled out questionnaires about their interest in a campus, which guided further research. Another survey used questionnaires to encourage EEG use on a website and included interviews about their interest levels. Respondents rated their satisfaction with "Yes" or "No", scored as 1 and 0, respectively, analyzed using the Guttman scale. EEG signal samples were then collected from final-year high school students in Yogyakarta by having them interact with the university website, producing digital graphics of brain activity. Here are the steps to produce the data:

1. The first survey involved distributing questionnaires to initial students about their interest in a campus.
2. The results of this survey will be used as preferences for continuing the research with other respondents.
3. The next survey also used a questionnaire to stimulate the use of EEG on the website and included interviews about their interest levels.
4. These interest levels (interested or not) will be used as targets in the classification process.
5. Respondents were asked to fill out a form based on User Satisfaction to implement EEG.
6. The form had choices of "Yes" and "No", with values of 1 and 0 respectively, which will be calculated using the Guttman scale.
7. The scale used is a positive scale.
8. Further data collection involved taking signal samples using an EEG tool with final year high school students in Jogjakarta.
9. EEG data was collected by directing respondents to use the university website's user interface.
10. The EEG produced electrical data on brain cell activity in the form of digital graphics during website use.

C. Fast Fourier Transform

Fast Fourier Transform (FFT) is a highly efficient algorithm used to compute the Discrete Fourier Transform (DFT) of a signal. FFT accelerates the process of transforming time-domain data into the frequency domain, making it possible to analyze the frequency components of a signal with greater efficiency compared to the traditional DFT method. This transformation is critical for extracting meaningful features from EEG signals, as it provides insight into the different frequency bands of brain activity [23]. The FFT algorithm is applied to the EEG data in this study to facilitate the analysis of the signal's frequency components, as described by Equation (1) [15].

$$X(u) = \frac{1}{N} \sum_{n=0}^{N-1} Xn = \left[\left(cos\ \frac{2\pi un}{N}\right) - j\ \left(sin\ \frac{2\pi un}{N}\right)\right] \tag{1}$$

The value from the FFT calculation is absolute to calculate the magnitude value before entering the identification stage, so that no FFT result value has a negative value using equation (2) [15].

$$|F(u)| = \sqrt{Real\ F(u)^2 + Imaginer\ F(u)^2} \tag{2}$$

D. Classification Methods

Naive Bayes

Naive Bayes is a probabilistic classification algorithm based on Bayes' Theorem, which assumes independence between features. It calculates the probability of each class given the input features and selects the class with the highest probability. This method is particularly effective with smaller datasets and can achieve high accuracy with minimal training data. In this study, Naive Bayes is used to classify the features extracted from the EEG signals, predicting user interest based on their interactions with university websites [16].

$$P(C|X) = \frac{P\ (C)}{P(X)}\ P(X|C) \tag{3}$$

The proposed method is shown in Fig. 3. for the Naive Bayes method. Read Data is a system for reading data to be processed. Partitioning is a process where data is divided into training data and test data. Predictor is predicting using naive Bayes algorithm modeling. The scorer is where the accuracy results of the predictions obtained are displayed.

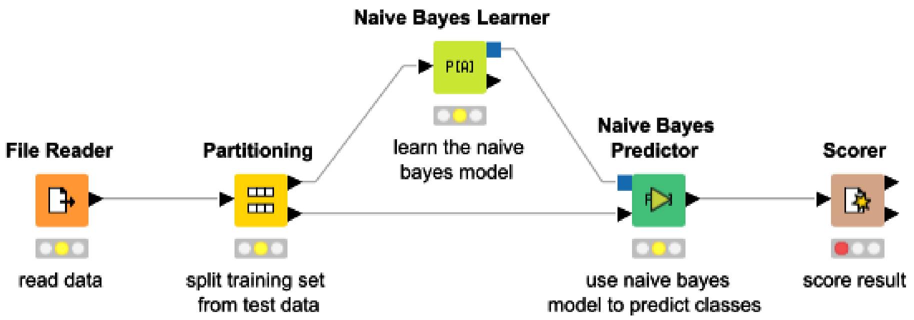

Fig. 1. The Naive Bayes Proposed Methods

K-Nearest Neighbor (KNN)

K-Nearest Neighbors is a non-parametric classification algorithm that assigns a class label to a data point based on the majority class among its 'k' nearest neighbors. The algorithm evaluates the similarity between the data point and its nearest neighbors using distance metrics, such as Euclidean distance. KNN is known for its effectiveness in handling noisy data and is used in this study to classify the features extracted from the EEG signals. By comparing the EEG signal features with those of known classes, KNN predicts user interest and engagement with the university websites.

$$d = \sqrt{(r1 - s1)^2 + (r2 + s2)^2 + \ldots + (rn - sn)^2}$$

The proposed method is shown in Fig. 2. for the K-Nearest Neighbor method. Basically, the proposed algorithm system that is built is very similar, but for the KNN modeling stage it uses the K value as the distance between grouped data. The red star is the predicted data, the small circle with k=3 is the data with the closest distance to the predicted data totaling 3. and the big circle with k=6 is the data with the closest distance to the predicted data totaling 6. if using k=3 then The star is in class B, and if k=6 then the star is in class A.

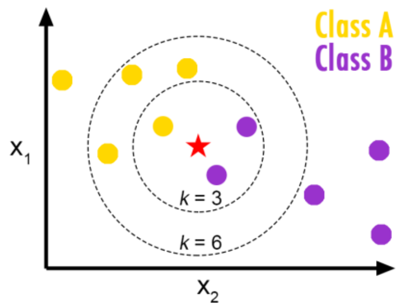

Fig. 2. The K-Nearest Neighbor Methods

4. RESULTS AND DISCUSSIONS

A. Fast Fourier Transform and Normalization

The EEG data was collected from 30 respondents from three different universities. All respondents were final-year high school students in Yogyakarta, who interacted with the university websites' user interfaces, producing digital graphics of their brain cell activity. The raw data collected was then filtered and normalized from the time domain to the frequency domain. The normalized data was extracted using Fast Fourier Transform (FFT). The final extraction results are presented in Table 1 and also illustrated in Fig.3.

Table 1. Result of Extraction

Respondents	First Campus	Second Campus	Third Campus
Respondents 1	5.246050556	10.75492833	7.577029231
Respondents 2	8.872113571	7.59096875	6.952468487
Respondents 3	4.7806625	6.400111667	5.553203824
Respondents 4	6.63555	6.33057619	6.768963889
Respondents 5	7.022932941	6.883966667	9.090531818
Respondents 6	10.75999619	5.001076667	5.09302
Respondents 7	8.571008333	7.001591912	5.986415
Respondents 8	9.856283077	10.06446136	8.475861905
Respondents 9	6.797797222	5.704198701	5.980422917
Respondents 10	5.589645789	9.4272775	5.808595395
Respondents 11	5.724164316	6.554045833	5.132457091
Respondents 12	6.499657292	8.84709	7.474335455
Respondents 13	12.254065	7.25085	6.850317857
Respondents 14	8.235916667	7.99129	5.40841413
Respondents 15	14.2624	12.95492083	13.5246875
Respondents 16	8.866665	8.6249875	6.660253676
Respondents 17	6.846838346	7.967545	13.14207
Respondents 18	9.246943333	6.013272917	5.030960256
Respondents 19	5.54114513	5.634221739	8.366204167
Respondents 20	5.623475	6.619009091	6.971983077
Respondents 21	7.297425833	5.230347619	7.584580476

Respondents 22	6.602306364	6.948446364	5.787719737
Respondents 23	6.881214286	9.189311111	12.07325357
Respondents 24	5.61538125	5.737083007	8.3459625
Respondents 25	7.726055556	5.000187308	14.778025
Respondents 26	13.78031667	6.678325	8.64772
Respondents 27	8.633685556	5.061033696	6.649188333
Respondents 28	7.9228625	4.473118182	7.903517857
Respondents 29	7.6922525	5.299368333	6.230920833
Respondents 30	7.967060769	6.418179167	4.841789211

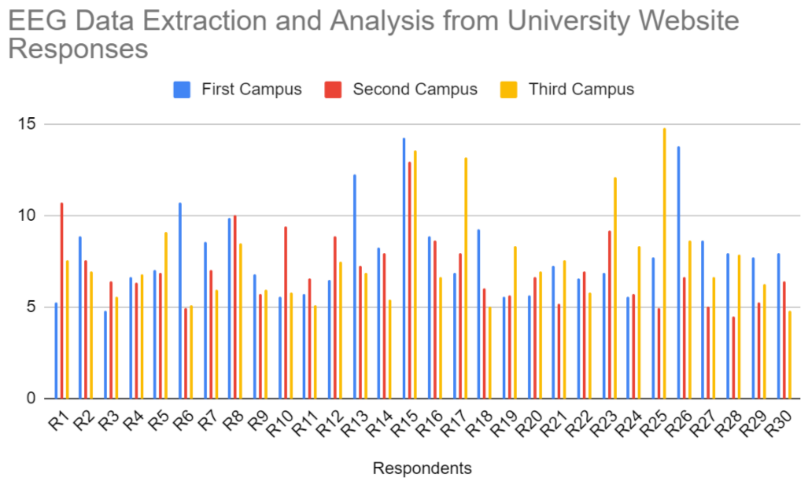

Fig.3. EEG Data Extraction from University Website Responses

B. Guttman Scale and Labeling

In this study, the Guttman scale is used for normalization and scoring of stimulus data collected via Google Forms. This process transforms raw responses into attribute data suitable for the train-test model. For binary-choice questions, responses are assigned values of 0 for "No" and 1 for "Yes." The scoring system is designed to determine the highest and lowest possible scores, where a respondent selecting "No" for all questions scores 0, and a respondent selecting "Yes" for all questions scores 100.

The assessment scale is then established based on this scoring system. A score threshold of 50% is used to categorize respondents into two groups: "Interested" and "Not Interested." Specifically, out of 15 questions, respondents must answer at least 8 questions with "Yes" to be classified as "Interested." Conversely, if fewer than 8 questions are answered with "Yes," the respondent falls into the "Not Interested" category. This 50% threshold effectively delineates the dividing line between the two categories, with scores above 50% indicating interest and scores below 50% suggesting a lack of interest.

The Guttman scale and labeling process is implemented using Python programming. This approach ensures that the scoring and categorization are both accurate and reproducible. The results of this labeling process, which are derived from respondent interviews and self-reports, are presented in Table 2 and Fig.4. This method provides a clear and systematic way to evaluate and categorize respondent interest based on their answers.

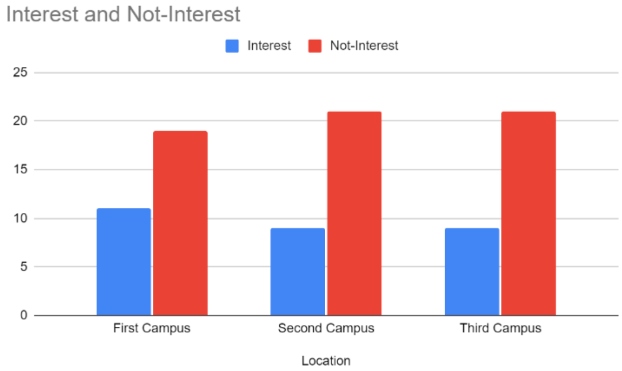

Fig.4. Labelling based on Location

Table 2. Result of Labeling

Location	Interest	Not-Interest
First Campus	11	19
Second Campus	9	21
Third Campus	9	21

C. Model Training and Testing Data

In this study, the model training and testing process involves several critical steps and variables. The variables used for model training include the extracted EEG data and the scores from questionnaires, which are computed using the Guttman scale. These scores are treated as attributes (denoted as x) and as target variables (denoted as y), where y contains labels indicating the level of interest.

To evaluate the performance of the models, different training and testing ratios are employed, specifically 0.15, 0.2, 0.25, and 0.4. These ratios determine the proportion of data used for training versus testing, allowing for a thorough assessment of model performance across various data splits. Additionally, the K-Nearest Neighbors (KNN) model is evaluated using odd values of k ranging from 1 to 10. This range of k values helps identify the optimal parameter for KNN by analyzing its effect on classification accuracy.

The accuracy of the Naive Bayes and KNN models is systematically assessed and the results are detailed in Table 3 and Fig.5 for Naïve Bayes and Table 4 and Fig.6 for KNN, respectively. These tables present the performance metrics of each model, providing a clear comparison of their effectiveness in classifying the EEG data and questionnaire scores.

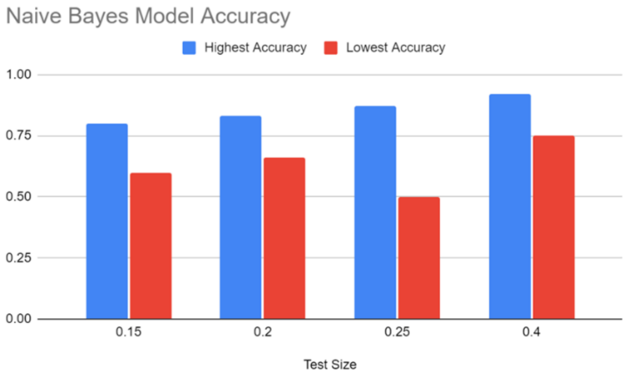

Fig.5. Accuracy based on Naïve Bayes

Table 3. Result of Naive Bayes Model Accuracy

Test Size	Highest Accuracy	Lowest Accuracy
0.15	0.8	0.6
0.2	0.83	0.66
0.25	0.87	0.5
0.4	0.92	0.75

Table 4. Result of KNN Model Accuracy

Test Size	K	Highest Accuracy	Lowest Accuracy
0.4	3	0.92	0.41
	5	0.83	0.5
	7	0.92	0.33
	9	0.83	0.41

0.25	3	0.87	0.5
	5	0.87	0.62
	7	0.87	0.25
	9	0.87	0.25
0.2	3	0.83	0.66
	5	0.83	0.5
	7	0.83	0.5
	9	0.83	0.5
0.15	3	0.8	0.6
	5	0.8	0.2
	7	0.8	0.4
	9	0.8	0.4

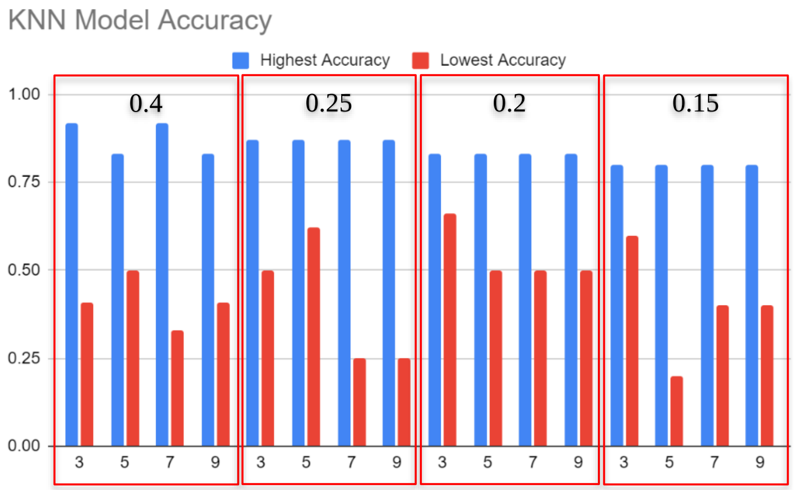

Fig.6. Accuracy based on KNN

D. Algorithm Performance Evaluation

The performance evaluation of the models in this study indicates that both models exhibit strong predictive capabilities for the case studies considered, as evidenced by their high precision and recall values. Precision measures the proportion of true positive predictions among all positive predictions, while recall measures the proportion of true positive predictions among all actual positives. For the models to be effective, it is essential that both false negative and false positive rates are low, ensuring the reliability of the predictions.

The f1-score, which is the harmonic mean of precision and recall, provides a single metric that balances these two aspects. This score is particularly useful for comparing models, as it takes both precision and recall into account, offering a comprehensive measure of a model's performance. By comparing the f1-scores, one can objectively determine the effectiveness of each model in making accurate predictions.

To facilitate the comparison and selection of the best-performing model, the precision, recall, and f1-score values are calculated and presented in Table 5 and Fig.7. These values reflect the highest accuracy achieved by each model, which is 92%. The table provides a clear and concise summary of the evaluation metrics, allowing for an informed decision on which model performs best in predicting user interest based on EEG data.

Table 5. Result of Comparison Performance Evaluation

Model	Precision	Recall	F1-Score
Naive Bayes	89%	100%	94%
K-Nearest Neighbors	100%	86%	92%

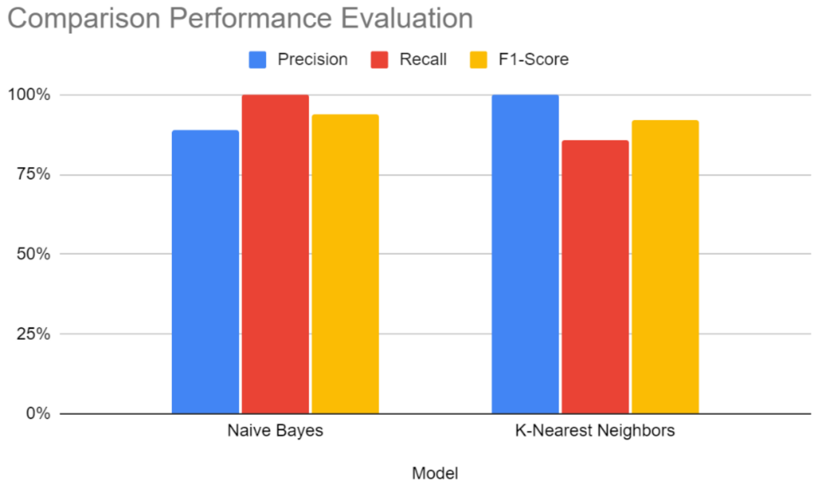

Fig.7. Comparison Results Performance Evaluation

In evaluating the performance of the models, two types of errors—False Positives (FP) and False Negatives (FN)—were considered. If the error caused by False Positives is deemed more undesirable, the model should be selected based on its precision. In this case study, high False Positive rates are undesirable because they affect the precision of the model, which assesses the accuracy of identifying respondents who are genuinely 'interested' in a campus based on their interaction with the university's profile website. Given this consideration, the K-Nearest Neighbors (KNN) model, which focuses on precision, is preferred in scenarios where minimizing False Positives is crucial.

On the other hand, if False Negatives are considered more problematic, the model should be evaluated based on its recall. False Negatives occur when respondents who are actually 'interested' in a campus are incorrectly classified as 'not interested.' Since the goal is to accurately predict interest, minimizing False Negatives is critical. In this context, the Naive Bayes model, which emphasizes recall, is more suitable as it focuses on correctly identifying all true 'interested' cases. Considering both types of errors—False Positives and False Negatives—it is important to balance precision and recall. In this study, both types of errors are deemed undesirable, and the accuracy of predictions regarding student interest in campuses is paramount. Therefore, the F1-Score, which provides a balanced measure of both precision and recall, is used to select the most appropriate model. The results indicate that the Naive Bayes model yields a better F1-Score, making it the preferred choice for this study. The F1-Score reflects a more balanced performance in predicting both 'interested' and 'not interested' categories, ensuring more accurate and reliable outcomes.

5. CONCLUSIONS

The classification of EEG signals using the Naive Bayes and K-Nearest Neighbor (KNN) methods was effective, demonstrating the capability to accurately categorize the data. The accuracy of each method varied depending on factors such as the training-to-test data ratio and the value of 'k' used in the KNN classification. The Naive Bayes model achieved a maximum accuracy of 92% and a minimum of 50% at a 0.4 training-to-test data ratio. In contrast, the KNN model also reached a maximum accuracy of 92%, but its minimum accuracy was significantly lower at 20% with a k-value of 5 and a 0.15 ratio. These results confirm that both Naive Bayes and KNN are viable methods for EEG signal classification.

When comparing the performance of the two models using the confusion matrix, the Naive Bayes model showed a slight but not significant superiority over KNN. This conclusion is based on performance metrics including precision, recall, and the F1-score. The F1-score, which balances both precision and recall, was 94% for Naive Bayes and 92% for KNN. Given the importance of balancing precision and recall in evaluating student interest through UI/UX, the Naive Bayes model, with its higher F1-score, is deemed slightly better for this case study.

The study involved 30 respondents who were tested with EEG to gauge their interest in various university websites. The results indicated that 11 respondents showed interest in the first campus, 9 in the second, and 9 in the third campus. While EEG data generally indicated high attention levels, some respondents who expressed interest in the campus websites were not identified as 'interested' by the predictive models. This discrepancy highlights that while EEG can effectively capture attention signals, it may not always align perfectly with self-reported interest. EEG data also revealed that respondents' interest was influenced by the relevance of the information provided. When respondents encountered information misaligned with their interests, EEG signals showed corresponding changes, underscoring the impact of relevant information on student engagement.

6. REFERENCES

[1] K. N. L. Mastra and R. F. Dharmawan, "Tinjauan User Interface Design Pada Website E-Commerce Laku6," *Narada*, vol. 5, no. 1, pp. 83–94, 2018.

[2] "Peran Guru BK Dalam Membantu Siswa Menentukan Pilihan Jurusan dan Kampus." https://www.quipper.com/id/blog/quipper-campus/campus-life/peran-guru-bk/ (accessed Sep. 01, 2023).

[3] H. Sanggarini, R. Purnamasari, and S. Hadiyoso, "Klasifikasi Efek Familiarity Pada Sinyal EEG Manusia Menggunakan Metode Hjorth Descriptor," vol. 6, no. 1, pp. 826–833, 2019.

[4] I. Budiman and R. Ramadina, "Penerapan Fungsi Data Mining Klasifikasi untuk Prediksi Masa Studi Mahasiswa Tepat Waktu pada Sistem Informasi Akademik Perguruan Tinggi," *Ijccs*, vol. x, No.x, no. 1, pp. 1–5, 2015.

[5] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. 2, p. 437, 2020, doi: 10.30865/mib.v4i2.2080.

[6] M. A. R. Hatmadiansyah, J. Raharjo, and ..., "Klasifikasi Sinyal Alpha Beta Terhadap Konsentrasi Diri Dalam Kondisi Mengerjakan Tes Matematikamenggunakan Metode K-nearest Neighbor (k-nn)," *eProceedings …*, vol. 8, no. 5, pp. 5090–5099, 2021, [Online]. Available: https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/15838

[7] M. A. Mawalid, "Klasifikasi Sinyal EEG untuk Mendeteksi Cybersickness melalui Time Domain Feature Extraction menggunakan Naive Bayes," 2019, [Online]. Available: https://repository.its.ac.id/60952/1/07111550052005-Master_Thesis.pdf

[8] H. J. Yoon and S. Y. Chung, "EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2230–2237, 2013, doi: 10.1016/j.compbiomed.2013.10.017.

[9] R. N. Devita, H. W. Herwanto, and A. P. Wibawa, "Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa indonesia," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 4, p. 427, 2018, doi: 10.25126/jtiik.201854773.

[10] Dessy Kusumaningrum and E. M. Imah, "Studi Komparasi Algoritma Klasifikasi Mental Workload Berdasarkan Sinyal EEG," *J. Sist. Cerdas*, vol. 3, no. 2, pp. 133–143, 2020,

doi: 10.37396/jsc.v3i2.69.
[11] S. K. Ilmiyati, "Analisis Pengolahan Data Sinyal EEG Pada Penderita Gangguan Tidur Menggunakan Metode Support Vector Machine Dan Naive Bayes," *Pros. Semin. Nas. Mhs.* ..., pp. 601–615, 2019, [Online]. Available: https://prosiding.unimus.ac.id/index.php/mahasiswa/article/view/517%0Ahttps://prosiding.unimus.ac.id/index.php/mahasiswa/article/viewFile/517/520
[12] G. A. Pamiela and A. Azhari, "Deep Learning on EEG Study Concentration in Pendemic," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 16, no. 2, p. 111, 2021, doi: 10.30872/jim.v16i2.6474.
[13] K. L. Tsui, V. Chen, W. Jiang, F. Yang, and C. Kan, "Data Mining Methods and Applications," *Springer Handbooks*, pp. 797–816, 2023, doi: 10.1007/978-1-4471-7503-2_38.
[14] I. Herdayanti, I. Wijayanto, and ..., "Klasifikasi Sinyal Eeg Saat Mendengarkan Musik Rock Dan Musik Klasik Dengan Metode Transformasi Wavelet," *eProceedings ...*, vol. 6, no. 2, pp. 4194–4201, 2019, [Online]. Available: https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/10645%0Ahttps://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/download/10645/10503
[15] K. N. Oktaviani *et al.*, "Identifikasi Neuropsikologi Emosi terhadap Video Iklan menggunakan Fast Fourier Transform dan Backpropagation Levenberg-Marquardt," *Semin. Nas. Apl. Teknol. Inf.*, pp. 11–2018, 2018.
[16] R. V. B. Vangara*, K. Thirupathur, and S. P. Vangara, "Opinion Mining Classification using Naive Bayes Algorithm," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 5, pp. 495–498, 2020, doi: 10.35940/ijitee.e2402.039520.