

Komparasi Algoritma *Machine Learning* untuk Klasifikasi Kondisi Mental Siswa Berdasarkan Sinyal EEG Wearable

Agus Waluyo^{a*}, Anton Yudhana^b, Herman Yuliansyah^c

Email: ^{a,*} 2536083022@webmail.uad.ac.id, ^b anton.yudhana@uad.ac.id,

^c herman.yuliansyah@uad.ac.id

Program Doktor Informatika, Universitas Ahmad Dahlan, Yogyakarta, Indonesia 55191^{abc}

Intisari

Kata kunci:
eeg, kondisi mental, machine learning, random forest, svm, knn, decision tree, naive bayes

Pengukuran kemampuan berpikir siswa secara objektif, real-time, dan penerapan informasi ini dalam pendidikan merupakan salah satu isu terbesar dalam pendidikan secara umum, dan khususnya dalam pembelajaran angka dan matematika. Dalam makalah penelitian ini, kami akan membandingkan kinerja lima algoritma *machine learning* seperti *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), *Decision Tree*, *Naive Bayes*, dan *random forest* dalam mengklasifikasikan tiga kondisi mental seperti *relaxed*, *neutral*, *concentrating* menggunakan sinyal EEG dari perangkat *wearable* berbiaya rendah, yaitu *headband Muse*. Data tersebut diperoleh dari dataset terbuka EEG *Brainwave Dataset: Mental State* di Kaggle, di mana data tersebut direkam menggunakan empat saluran (TP9, AF7, AF8, TP10). *Pipeline* pemrosesan sinyal melibatkan pra-pemrosesan sinyal (*filter band-pass*), ekstraksi fitur (rata-rata, *deviasi standar*, kemiringan, kurtosis) baik dalam domain waktu maupun domain frekuensi (*bandpower delta, theta, alpha, dan beta*), dan pengurangan dimensi (*Principal Component Analysis* (PCA)). Akurasi, presisi, *recall*, dan *F1-score* diuji pada 102 sampel uji. Mereka menemukan bahwa yang paling sukses adalah *Random Forest* dengan akurasi 0,84 dan *macro F1-score* 0,84, diikuti oleh SVM (0,82), KNN (0,78), *Decision Tree* (0,75), dan *Naive Bayes* (0,68). Masalah konstan dalam semua model adalah kondisi *netral* karena karakteristik sinyal kondisi netral berada di antara karakteristik sinyal dari dua kondisi lainnya. Hasil ini menunjukkan bahwa serangkaian karakteristik dasar dari perangkat EEG berbiaya rendah yang dapat dikenakan, dikombinasikan dengan algoritma *random forest* mungkin sudah cukup untuk mengklasifikasikan kondisi mental. Selain itu, pendekatan dengan model ini memiliki potensi sebagai platform untuk membangun sistem pemantauan kesiapan *kognitif* adaptif dan *real-time*.

Abstract

Tentang naskah:
-diterima: 25 Juni 2026
-diterbitkan: 1 Juli 2026

Keywords:
eeg, mental state, machine learning, random forest, svm, knn, decision tree, naive bayes

Application of this information in education is one of the biggest issues in education in general, and in learning of numbers and mathematics in particular. In this research paper, we are going to compare the performance of five machine learning algorithms such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Decision Tree, Naive Bayes and random forest in classifying three mental states such as relaxed, neutral and concentrating using EEG signals of a low cost wearable device, the Muse headband. The data had been received in the open dataset EEG Brainwave Dataset: Mental State on Kaggle where they were recorded using four channels (TP9, AF7, AF8, TP10). The signal processing pipeline involves signal pre-processing (band-pass filter), feature extraction (mean, standard deviation, skewness, kurtosis) of both time domain and frequency domain (delta, theta, alpha and beta bandpower) and dimensionality reduction (Principal Component Analysis (PCA)). Accuracy, precision, recall and F1-score were tested on 102 test samples. They found that the most successful one was Random Forest whose accuracy was 0.84 and macro F1-score was 0.84 and SVM (0.82), KNN (0.78), Decision Tree (0.75) and Naive Bayes (0.68) followed. A constant problem in all of the models was the neutral condition because the signal characteristics of the neutral condition lie between those of the other two conditions. These results imply that a set of basic characteristics of a low-cost wearable EEG device and the random forest algorithm may already be sufficient to classify mental states. Furthermore, this approach has potential as a platform on which to build adaptive, real-time cognitive preparedness monitoring systems

1. Pendahuluan

Kemampuan intelektual manusia memainkan peran penting dalam menentukan keberhasilan manusia ketika melakukan beberapa tugas intelektual seperti tugas matematika dan numerik. Meskipun demikian, penilaian kapasitas intelektual dalam lingkungan sekolah masih sangat terbatas pada cara yang lebih umum, yaitu tes tertulis, hingga saat ini, wawancara terstruktur, dan observasi. Metode-metode tersebut memiliki keterbatasan yang sulit untuk dihindari, seperti Metode-metode ini cukup lama dan terlepas dari dasar validitasnya yang cukup kuat, beberapa keterbatasan sulit diatasi. Kondisi psikologis sementara, seperti kecemasan ujian, kelelahan, kurang motivasi, dan lainnya, cenderung mengganggu skor penilaian, sehingga skor yang diperoleh belum tentu mencerminkan kemampuan kognitif siswa. Selain itu, hal ini disebabkan oleh instrumen penilaian yang masih menyetujui subjektivitas evaluator, dan kegagalan untuk mendapatkan akses ke data yang dapat terus mendokumentasikan transformasi kondisi mental siswa secara real-time, sehingga instrumen yang lebih efisien menjadi kurang efektif dalam mendokumentasikan perubahan kondisi pikiran siswa secara real-time (Al-Nafjan & Aldayel, 2022), (Wang et al., 2022).

Dekade sebelumnya menghadirkan beberapa peluang baru dengan pengukuran kondisi kognitif yang lebih objektif dan berorientasi data karena perkembangan teknologi sensor fisiologis dan komputer berkinerja tinggi yang begitu pesat. *Electroencephalography* (EEG), teknik pengukuran aktivitas listrik otak dengan menempatkan elektroda pada permukaan kulit kepala, adalah salah satu teknologi yang banyak digunakan di bidang ini. EEG dapat bermanfaat dengan resolusi temporal yang sangat tinggi yang memungkinkan untuk menangkap perubahan keadaan mental dalam hitungan milidetik (Walla, 2025a). Aspek ini berkontribusi pada fakta bahwa EEG adalah instrumen yang sensitif untuk melacak dinamika proses kognitif seseorang (Walla, 2025b). Dalam analisisnya, sinyal EEG umumnya dibedakan ke dalam beberapa pita frekuensi, yaitu delta (0,5–4 Hz), theta (4–8 Hz), alfa (8–13 Hz), beta (13–30 Hz), dan gamma (>30 Hz) (Udeberg et al., 2025). Masing-masing pita frekuensi ini dapat dikaitkan dengan proses mental yang berbeda. Gelombang theta dan alpha lebih sering dikaitkan dengan memori kerja dan keadaan rileks, sedangkan gelombang beta dan gamma lebih sering dikaitkan dengan pekerjaan mental yang serius dan pemrosesan informasi tingkat tinggi (Gkintoni & Halkiopoulou, 2025). Oleh karena itu, setiap perubahan keseimbangan pita frekuensi ini dapat digunakan sebagai

pengukuran objektif untuk menetapkan perubahan keadaan pikiran seseorang, misalnya, keadaan rileks hingga konsentrasi penuh (Do et al., 2024). Korelasi EEG dan pendidikan juga memiliki dasar yang cukup karena hasil penelitian terus menunjukkan bahwa kondisi mental siswa dalam proses pembelajaran secara langsung memengaruhi proses pembelajaran, khususnya dalam kemampuan numerik dan matematika. Berbagai pola aktivitas otak diamati dalam penelitian yang dilakukan pada EEG yang dapat dikenakan di kalangan siswa sekolah dasar saat mereka mempelajari pecahan di mana kelompok berprestasi tinggi dibandingkan dengan kelompok berprestasi rendah menunjukkan pola aktivitas otak yang berbeda (Lekati et al., 2025a). Siswa berprestasi rendah lebih cenderung mengalami dominasi bentuk gelombang delta dan theta selama beban kognitif dan siswa berprestasi tinggi memiliki gelombang beta yang lebih terkait yang terbukti memiliki kontrol kognisi dan kontrol perhatian yang lebih luas (Lekati et al., 2025b).

Secara neurofisiologis, perbedaan keadaan mental dapat diidentifikasi dalam bentuk susunan aktivitas listrik otak khususnya di daerah frontal yang memainkan peran penting dalam mempertimbangkan perhatian dan kontrol kognitif. Langkah ini akan lebih cenderung mengarah pada siswa yang berada dalam keadaan *concentrating*. Peningkatan jumlah gelombang beta dan penurunan jumlah gelombang alfa cenderung menunjukkan pertumbuhan tersebut dibandingkan dengan ketika seseorang *relaxed*, yang biasanya merupakan kebalikannya. Sementara itu, keadaan netral dapat dianggap sebagai keadaan di antara keduanya (Kotte et al., 2025). Tingkat perhatian yang diberikan kepada siswa pada kondisi ini tidak sekuat pada kondisi berkonsentrasi, dan tidak pula selonggar pada kondisi rileks. Kecenderungan perubahan rasio gelombang beta terhadap gelombang alfa ini kemudian diimplementasikan sebagai instrumen yang kurang lebih objektif untuk menilai kesiapan siswa dalam proses pembelajaran (Xie et al., 2025). Hasil penelitian menunjukkan bahwa pengembangan klasifikasi otomatis dari tiga keadaan mental, seperti: berkonsentrasi, netral, dan rileks, memiliki potensi besar untuk menjadi dasar dari alat pemantauan kesiapan kognitif yang dapat diimplementasikan untuk melacak siswa dalam lingkungan pembelajaran (Syed et al., 2025).

Masalah terbesar dalam menerapkan EEG sebagai jenis klasifikasi kondisi mental adalah kualitas sinyal. Sinyal EEG bersifat non-stasioner (bervariasi seiring waktu), non-linier (hubungan antar variabel tidak begitu jelas), dan non-dimensi sehingga sulit

diinterpretasikan hanya dengan menggunakan pengamatan manual (Khare et al., 2023). Solusi yang tepat untuk semua kekurangan ini adalah solusi *machine learning* karena pola kompleks dalam sinyal EEG dapat dideteksi dan dieksploitasi secara otomatis (Liu & Zhao, 2024). Melalui penggabungan metode ekstraksi fitur domain waktu dan domain frekuensi, *machine learning* mampu mengidentifikasi hubungan non-linier antar fitur yang seringkali sulit diidentifikasi melalui metode analisis tradisional. Algoritma klasifikasi *machine learning* lainnya telah diuji coba dalam klasifikasi kondisi mental menggunakan sinyal EEG. Diklaim bahwa *Support Vector Machine* (SVM) dapat bekerja dengan baik pada data berdimensi tinggi, dengan kecenderungan non-linier dan kemungkinan besar, dapat paling akurat dibandingkan dengan metode tradisional lainnya dalam mengklasifikasikan kondisi mental (Zuo et al., 2025). K-Nearest Neighbor (KNN) didasarkan pada gagasan pengelompokan sinyal berdasarkan kemiripannya dengan data pelatihan, dan telah terbukti mampu mendekati interaksi kompleks fitur dalam sinyal EEG [16],[17]. Model yang disediakan oleh Decision Tree mudah dipahami, untuk mengklasifikasikan kondisi mental berdasarkan fitur EEG (Ramírez-Arias et al., 2022). Naive Bayes didasarkan pada teorema Bayes, dan mengasumsikan bahwa fitur-fitur tersebut independen dan memiliki kekuatan dalam hal efisiensi komputasi jika jumlah data pelatihan terbatas [19],[20]. Sementara itu, Rand Forest yang merupakan algoritma ensemble, yang juga merupakan Decision Tree, juga telah menunjukkan dirinya sebagai salah satu model yang paling efektif untuk digunakan ketika mengklasifikasikan kondisi mental menggunakan data EEG (Jan et al., 2025). Algoritma-algoritma tersebut memiliki ciri khas masing-masing dalam hal membuat batas keputusan klasifikasi, ketahanan terhadap noise, dan efisiensi (Pramitarini et al., 2023). Studi komparasi multi-algoritma menunjukkan bahwa setiap model memiliki sensitivitas yang berbeda terhadap karakteristik dimensi data, di mana SVM sering kali memberikan batas keputusan yang lebih kokoh dibandingkan algoritma berbasis probabilitas atau kedekatan jarak (Yuliansyah et al., 2026).

Beberapa studi terbaru telah menghasilkan hasil positif dalam sinyal EEG yang dipandu untuk kondisi mental. Sinyal EEG yang diperoleh pada headband Muse dapat dianalisis dan dikategorikan menjadi tiga kondisi mental dengan fitur statistik dan fitur domain frekuensi yang memiliki tingkat akurasi lebih dari 99 persen ketika diterapkan pada algoritma Random Forest dan XGBoost (Rahman & Nasor, 2024). Namun, sebagian besar makalah ini

hanya bereksperimen dengan satu atau dua algoritma saja, dan belum melakukan perbandingan sistematis dari berbagai metode klasik, dengan data yang sama. Dengan demikian, menentukan algoritma yang paling sesuai ketika digunakan dalam sistem nyata di mana keandalan dan kinerja tinggi dibutuhkan, tetap menjadi pertanyaan yang sulit. Masalah kedua yang perlu diperbarui adalah bahwa studi sebelumnya lebih cenderung mempertimbangkan penggunaan perangkat EEG multikanal kelas klinis yang mahal yang tidak terlalu aplikatif dalam pembelajaran aktual (Orovas et al., 2025). Perangkat EEG yang lebih murah seperti headband Muse (hanya empat saluran: TP9, AF7, AF8, dan TP1) memiliki potensi yang lebih tinggi untuk digunakan di sekolah dan di rumah. Perangkat ini memiliki keterbatasan, terutama dalam aspek resolusi spasial yang berkurang dan kerentanan sinyal terhadap artefak gerakan. Laporan lain tentang penggunaan EEG yang dapat dikenakan dalam pengenalan gangguan kognitif ringan telah melaporkan bahwa tingkat akurasi tidak seragam dengan rentang akurasi 46-95 persen tergantung pada kombinasi fitur/algoritma (Lekati et al., 2025b). Dalam hal ini, diperlukan studi yang secara langsung mengukur sejauh mana kombinasi fitur sederhana dari perangkat berbiaya rendah ini bersama dengan berbagai algoritma klasifikasi klasik dapat memberikan kinerja yang wajar dalam klasifikasi kondisi mental menjadi tiga kategori.

Dataset yang dipelajari adalah dataset yang tersedia untuk umum yang disebut "EEG Brainwave Dataset: Mental State" yang ditawarkan di Kaggle. Dataset ini terdiri dari data sinyal EEG dari 4 pasien yang direkam dalam tiga kondisi mental, yaitu rileks (*relaxed*), netral (*neutral*), dan berkonsentrasi (*concentrating*) dengan bantuan headband EEG Muse. Ada dua tujuan dalam penelitian ini. Pertama, untuk membangun pipeline klasifikasi dengan beberapa langkah, termasuk langkah ekstraksi fitur di domain waktu dan frekuensi, pengurangan dimensi dengan *Principal Component Analysis* (PCA) untuk menyederhanakan data dan mencegah overfitting, dan pelatihan lima algoritma *machine learning* klasik, seperti SVM, KNN, Decision Tree, Naive Bayes, dan Random Forest. Kedua, untuk melakukan perbandingan sistematis kinerja setiap algoritma untuk melakukan tugas klasifikasi tiga kelas. Pertanyaan penelitian yang akan dijawab oleh penelitian ini dalam dua bidang minat tersebut adalah sebagai berikut: algoritma mana yang akan melakukan pekerjaan terbaik dalam mengklasifikasikan kondisi mental *relaxed*, *neutral*, *concentrating* berdasarkan sinyal EEG

yang dibuktikan oleh perangkat *wearable* berbiaya rendah?

Dua kontribusi mungkin akan dihasilkan dari studi ini. Pertama, studi ini mencakup perbandingan keseluruhan kinerja berbagai algoritma klasifikasi untuk memproses informasi EEG (sinyal aktivitas otak) yang direkam pada perangkat *wearable* yang murah. Analogi ini juga dapat bermanfaat bagi peneliti lain yang mungkin tertarik untuk membuat aplikasi serupa agar memiliki konsep yang lebih jelas tentang model mana yang paling tepat diterapkan. Kedua, studi ini akan dijadikan sebagai dasar dan sistem yang mampu memantau kesiapan kognitif dan tingkat konsentrasi siswa secara real-time dan dinamis akan dibuat. Sistem seperti itu pada akhirnya dapat mengarah pada terciptanya strategi pembelajaran yang lebih akomodatif terhadap kemampuan siswa, misalnya dalam pembelajaran operasi numerik dan matematika, pendidikan formal dan informal (Chen et al., 2024).

2. Kerangka Teori

Lima algoritma yang digunakan dalam penelitian ini dipilih karena sesuai dengan karakteristik sinyal EEG yang bersifat non-stasioner dan memiliki dimensi tinggi. Tabel dibawah ini menyajikan ringkasan kelebihan dan keterbatasan masing-masing algoritma dalam konteks klasifikasi kondisi mental berbasis EEG, yang menjadi dasar perbandingan.

Tabel 1. Perbandingan Kelebihan dan Kekurangan Algoritma Machine Learning

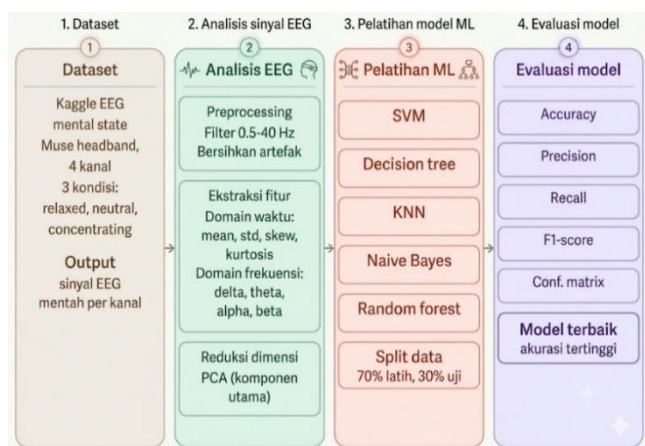
Algoritma	Kelebihan	Kekurangan
SVM	Efektif pada data berdimensi tinggi; batas keputusan yang kokoh; mendukung kernel non-linier (RBF) untuk menangani pola EEG yang kompleks	Waktu pelatihan lama pada dataset besar; pemilihan kernel dan parameter memerlukan tuning yang cermat
Random Forest	Robust terhadap noise dan overfitting, mampu menangani fitur campuran (domain waktu dan frekuensi), memberikan	Interpretabilitas rendah (black-box), memerlukan lebih banyak memori dan waktu komputasi dibanding

Algoritma	Kelebihan	Kekurangan
KNN	estimasi kepentingan fitur Sederhana dan tidak memerlukan pelatihan eksplisit; dapat mendeteksi pola lokal dalam ruang fitur EEG	model tunggal Sensitif terhadap noise dan skala fitur, performa menurun pada data berdimensi tinggi tanpa reduksi dimensi
Decision Tree	Mudah diinterpretasikan; aturan keputusan dapat divisualisasikan; tidak memerlukan normalisasi data	Rentan terhadap overfitting, tidak stabil terhadap perubahan kecil pada data pelatihan
Naive Bayes	Sangat efisien secara komputasi, bekerja baik dengan data pelatihan terbatas, mudah diimplementasikan	Asumsi independensi fitur sering tidak terpenuhi pada sinyal EEG yang saling berkorelasi antar pita frekuensi

Berdasarkan karakteristik teoritis di atas, penelitian ini secara khusus memilih kelima algoritma tersebut untuk dibandingkan kinerjanya pada dataset EEG kondisi mental.

2.1 Metode Penelitian

Desain penelitian dalam makalah ini didasarkan pada pemrosesan sinyal aktivitas otak dan klasifikasi *machine learning* dengan mempertimbangkan bahwa tujuan utamanya adalah untuk menemukan pola aktivitas otak, yang menjadi bukti keadaan mental tertentu dan mentransfernya ke nilai numerik, yang dapat dipelajari dengan bantuan algoritma klasifikasi. Data direkam di situs Kaggle dalam bentuk rekaman EEG empat saluran menggunakan ikat kepala Muse dan tiga keadaan mental diperiksa; *relaxed*, *neutral*, *concentrating*. Gambar di bawah menunjukkan tahapan penelitian ini.



Gambar 1 Tahapan Penelitian

Alur kerja penelitian akan sebagai berikut: Pertama, akuisisi data; Kedua, interpretasi sinyal yang diperoleh dengan alat ekstraksi artefak, analisis fitur domain waktu (rata-rata, deviasi standar) dan domain frekuensi (daya gelombang delta, alfa, beta dan theta) dan pengurangan dimensi, menggunakan data pelatihan dan pengujian (70: 30). Gabungan nilai fitur kerapatan spektral daya, yang digabungkan dengan prosedur transformasi Fourier dan algoritma klasik, yang dienkripsi sebagai SVM, KNN dan Random Forest, telah terbukti sangat efektif dalam pengembangan model klasifikasi perhatian berbasis EEG (Al-Nafjan & Aldayel, 2022). Relevansi headband Muse dalam penelitian ini didasarkan pada fakta bahwa perangkat ini telah terbukti mampu menghasilkan data multi-channel yang memadai dan cukup stabil untuk berguna dalam klasifikasi kondisi mental seperti yang dilaporkan oleh (Rahman & Nasor, 2024) dalam skema klasifikasi.

2.2 Pengumpulan Data

Data yang digunakan dalam makalah ini diambil dari dataset terbuka, EEG Brainwave Dataset: Mental State (Bird et al., 2018a) ang datasetnya tersedia di platform Kaggle pada tautan berikut: <https://www.kaggle.com/datasets/birdy654/eeg-brainwave-dataset-mental-state>. Dataset ini dalam format CSV dengan setiap baris berisi sebagian kecil rekaman gelombang sinyal otak seseorang dan kolom berisi angka yang diperoleh dengan memproses gelombang sinyal otak yang menunjukkan variasi gelombang otak pada waktu tertentu. Setiap baris secara longgar diakhiri dengan kolom penting tentang keadaan mental peserta ketika rekaman dibuat dalam keadaan *relaxed*, *neutral*, *concentrating*, dan kolom tersebut digunakan untuk memungkinkan algoritma *machine learning* untuk menguraikan pola otak agar sesuai dengan keadaan mental. Headband EEG Muse sebagai peralatan yang dapat dikenakan

yang relatif murah digunakan untuk mengumpulkan data ini, yang memiliki empat titik sensor yaitu, TP9, AF7, AF8 dan TP10, di dahi dan di sekitar telinga. Sampel partisipan dikenai pengumpulan data di mana sampel tersebut diinformasikan untuk menjalani tiga kondisi mental; kondisi rileks di mana tubuh dalam keadaan rileks dan di setiap kondisi tidak ada item stimulasi khusus, kondisi netral di mana partisipan diinstruksikan untuk tetap diam tanpa aktivitas berpikir khusus, dan akhirnya kondisi konsentrasi di mana partisipan diberitahu untuk memusatkan perhatian mereka pada aktivitas tertentu dan durasinya sekitar satu menit di setiap kondisi sesi (Bird et al., 2018b). Dari sinyal kasar ini, beberapa sifat statistik dan spektral diperoleh untuk menggambarkan sifat gelombang alfa, beta, theta, delta, dan gamma, yang masing-masing diketahui terkait dengan berbagai tingkat aktivitas mental. Dataset ini dipilih setelah faktor bahwa dataset ini telah banyak diterapkan dalam penelitian sebelumnya yang berkaitan dengan klasifikasi kondisi mental berbasis *machine learning* sehingga memungkinkan hasil penelitian ini untuk dibandingkan dengan hasil penelitian sebelumnya (Coman et al., 2024).

2.3 Variabel Penelitian

Dua jenis variabel dalam penelitian ini adalah variabel input seperti sinyal EEG mentah yang diambil oleh perangkat Muse dan diolah menjadi fitur statistik dan spektral, serta variabel output seperti kategori keadaan mental *relaxed*, *neutral*, *concentrating*. Idealnya, secara teoritis didasarkan pada premis bahwa aktivitas otak menghasilkan pola unik yang dapat dibedakan di antara semua keadaan mental tidak hanya pada perubahan sinyal seiring waktu tetapi juga berkaitan dengan distribusi energi gelombang pada frekuensi yang berbeda. Kombinasi sifat dari kedua domain tersebut dapat digunakan untuk menentukan keadaan relaksasi dan konsentrasi secara objektif. Hal ini sejalan dengan tinjauan oleh (Yildirim et al., 2023) yang mendukung fakta bahwa ekstraksi fitur dalam domain waktu-frekuensi, dekomposisi, waktu-frekuensi, dan spasial merupakan konstruksi inti dalam sebagian besar aplikasi *machine learning* di bidang EEG.

2.4 Instrumen dan Perangkat Penelitian

Perangkat utama yang akan digunakan dalam penelitian ini adalah headband EEG Muse, yang dapat digunakan untuk memantau aktivitas listrik otak ketika subjek dihadapkan pada situasi pengujian yang akan dikembangkan. Selain perangkat perekam, komputer atau laptop juga digunakan dalam penelitian ini sebagai perangkat penyimpanan

data, pemrosesan sinyal, ekstraksi fitur, dan pelatihan model klasifikasi. Bahasa pemrograman Python telah digunakan untuk melakukan analisis dan klasifikasi sinyal, pemrosesan data, dan evaluasi kinerja model.

2.5 Prosedur Penelitian

Proses penelitian dilakukan secara bertahap sedemikian rupa sehingga transformasi sinyal mentah menjadi keputusan klasifikasi dapat dilakukan secara sistematis dan diulang.

2.5.1 Prapemrosesan Sinyal

Sinyal EEG yang baru saja direkam tidak dapat digunakan secara langsung karena masih mengandung beberapa jenis distorsi seperti noise, pergeseran garis dasar, dan efek cahaya akibat kedipan atau gerakan otot. Oleh karena itu, praproses sinyal diperlukan dan pemrosesan lebih lanjut dapat dilakukan. Dalam penelitian ini, praproses dilakukan dengan *filter band-pass* yang memiliki rentang frekuensi antara 0,5 dan 40 Hz untuk menghilangkan frekuensi rendah dan tinggi dan diikuti dengan penghapusan artefak dan membagi sinyal menjadi segmen-segmen kecil dengan panjang konstan (Alreshidi et al., 2023). Penggunaan rentang frekuensi dan skema segmentasi seperti ini sering digunakan untuk mengambil sampel dan mempartisi rekaman EEG karena telah terbukti lebih efektif dalam mempertahankan konten fisiologis yang berpotensi bermakna, serta mengurangi gangguan yang tidak diinginkan. *Filter band-pass* 0,5-40 Hz merupakan pilihan umum dalam literatur beberapa tahun terakhir, untuk mengabaikan pergeseran garis dasar yang lambat, dan interferensi sinyal frekuensi tinggi seperti artefak otot, dan noise listrik dari lingkungan sekitar (Chaddad et al., 2023). Data kemudian dinormalisasi agar skala dapat dibandingkan antar fitur sedemikian rupa sehingga variasi dalam rentang nilai fitur statistik dan spektral tidak akan memengaruhi pelatihan model. Langkah ini akan dibenarkan oleh fakta bahwa model ini akan berada dalam posisi yang lebih baik untuk mempelajari pola sinyal dengan cara yang lebih konsisten dan objektif, setelah itu akan menciptakan klasifikasi kondisi mental yang lebih dapat diandalkan.

2.5.2 Ekstraksi Fitur

Transformasi sinyal EEG yang telah dibersihkan menjadi nilai numerik, yang dapat dimasukkan ke dalam model machine learning, disebut ekstraksi fitur. Dalam penelitian ini, domain waktu dan domain frekuensi adalah dua domain penting yang dipertimbangkan dalam ekstraksi fitur. Empat ukuran domain waktu dibuat: *mean*, *standard deviation*, *skewness*, dan *kurtosis* yang masing-masing menunjukkan

rata-rata, dispersi, asimetri, dan puncak distribusi sinyal. Penerapan transformasi sinyal berbasis domain frekuensi terbukti efektif dalam menyaring komponen gelombang otak (Yudhana et al., 2020). Karakteristik ini dipilih karena kesederhanaannya, dan cukup informatif untuk mengidentifikasi kondisi mental (Yildirim et al., 2023). *Bandpower* dalam domain frekuensi pada pita delta, theta, alfa, dan beta dicatat sebagai indikator kekuatan sinyal. Pita frekuensi juga dikaitkan dengan keadaan kognitif tertentu dan karakteristiknya membantu dalam menentukan perubahan keadaan mental (Chaddad et al., 2023). Penggabungan kedua domain didasarkan pada pengamatan bahwa keduanya saling melengkapi karena ciri-ciri domain waktu menunjukkan sifat keseluruhan sinyal dan ciri-ciri domain frekuensi menunjukkan bagaimana energi sinyal didistribusikan di sekitar durasi pita gelombang otak. Keduanya bersama-sama memberikan representasi sinyal yang lebih detail, yang biasanya meningkatkan kinerja model klasifikasi.

2.5.3 Reduksi Dimensi dengan PCA

Setelah keberhasilan ekstraksi semua fitur, dimensi diminimalkan dengan bantuan *Principal Component Analysis* (PCA). Secara sederhana, PCA berupaya meminimalkan jumlah atribut data tanpa merusak data yang bermakna dalam atribut tersebut. Hal ini dilakukan dengan mendistorsi fitur asli menjadi sejumlah komponen baru yang lebih kecil, namun sebagian besar informasi yang terkandung dalam data asli tetap ada. Studi EEG lainnya telah mengadopsi PCA untuk mengatasi masalah dimensi fitur yang tinggi, mengurangi redundansi antar fitur, dan mengurangi kemungkinan overfitting, di mana model overfitting pada data pelatihan, dan karenanya overfitting pada data baru. PCA yang diterapkan pada data EEG juga dianggap mampu mengurangi dimensi sinyal EEG multikanal, sehingga mempercepat proses ekstraksi fitur serta klasifikasi. Hal ini menunjukkan bahwa pengurangan dimensi memainkan peran utama dalam peningkatan efisiensi komputasi serta stabilitas model klasifikasi, terutama ketika ukuran fitur yang dihasilkan jauh lebih besar daripada ukuran model untuk menggeneralisasi pola terkait data (Sadegh-Zadeh, 2024).

2.5.4 Klasifikasi Data

Prosedur utama dalam makalah ini adalah bagian klasifikasi di mana fitur-fitur yang diekstrak dan disederhanakan yang dibuat dengan penerapan PCA akan dimasukkan ke dalam model untuk dilatih dan menentukan keadaan mental individu. Eksperimen ini

menerapkan lima model *machine learning* klasik yaitu *Support Vector Machine* (SVM), Random Forest, K-Nearest Neighbors (KNN), Decision Tree, dan Naive Bayes, di mana masing-masing model bekerja secara berbeda dan memiliki persamaan yang berbeda untuk sampai pada keputusan klasifikasi. Mesin vektor pendukung (*Support Vector Machine*/SVM) beroperasi dengan cara berikut: Ia mencoba menemukan garis pemisah atau bidang hiper yang membagi data berdasarkan jarak terbesar yang mungkin. Berikut ini dapat diungkapkan secara matematis sebagai model SVM:

$$f(x) = \text{sign} (\sum_i \alpha_i y_i K(x_i, x) + b)$$

dengan $K(x_i, x)$ yang merupakan fungsi kernel yang mengubah data ke ruang berdimensi lebih tinggi sehingga data yang tidak dapat dipisahkan secara linier menjadi dapat dipisahkan, - adalah koefisien pembobotan, y_i adalah label kelas dan b adalah bias. Random Forest adalah kumpulan model yang dibangun di atas Decision Tree acak, dan kemudian, output prediktif dari semua pohon dijumlahkan melalui prosedur pemungutan suara. Model matematika Random Forest terlihat seperti berikut:

$$\hat{y} = \text{mode} \{ T_1(x), T_2(x), \dots, T_n(x) \}$$

dengan $T_i(x)$ merupakan hasil prediksi dari pohon keputusan ke- i , dan $\hat{y}$ adalah mayoritas dari Decision Tree. K-Nearest Neighbors mengklasifikasikan data menggunakan data baru untuk mengklasifikasikan data berdasarkan kemiripannya dengan sekelompok tetangga terdekat. Kemiripan ini biasanya dihitung dengan bantuan jarak Euclidean. KNN dapat dimodelkan secara matematis sebagai berikut:

$$d(x, x_i) = \sqrt{(\sum_j (x_j - x_{ij})^2)}$$

Setelah jarak dihitung, label kelas diberikan berdasarkan tetangga terdekat yang paling sering muncul, yang jumlahnya ditunjukkan oleh k . Decision Tree menciptakan model Decision Tree dengan secara bertahap memberikan pembagian data berdasarkan fitur yang paling efektif. Indeks Gini adalah salah satu instrumen paling menonjol untuk membagi sesuatu ke dalam pembagian yang paling optimal. Model Decision Tree dapat direpresentasikan secara matematis sebagai berikut:

$$Gini(t) = 1 - \sum_i p_i^2$$

dengan p_i adalah proporsi data kelas ke- i pada simpul (node) t . Semakin kecil nilai Gini, semakin homogen kelas-kelas pada node tersebut. Naive Bayes membuat prediksi menggunakan teorema Bayes; sebuah persyaratan bahwa setiap fitur bersifat independen. Hal ini dapat diberikan oleh rumus berikut:

$$P(y | x_1, x_2, \dots, x_n) \propto P(y) \cdot \prod_i P(x_i | y)$$

dengan $P(y)$ adalah probabilitas awal suatu

kelas, dan $P(x_i | y)$ yaitu probabilitas fitur ke- i terjadi pada suatu kelas. Dalam penelitian ini digunakan dua skema validasi secara berurutan. Data pertama kali dibagi menggunakan train-test split dengan rasio 70:30, di mana 70% digunakan untuk data pelatihan dan 30% untuk data pengujian akhir. Pembagian data dilakukan secara random dengan stratifikasi agar distribusi setiap kelas tetap seimbang pada data latih maupun uji. Langkah selanjutnya, untuk memperoleh hasil yang lebih stabil, digunakan validasi silang 10-fold. Pada tahap ini, data pelatihan dibagi menjadi 10 bagian, kemudian model dilatih pada 9 bagian dan diuji pada 1 bagian secara bergantian hingga seluruh data mendapatkan giliran sebagai data uji. Nilai performa dari seluruh iterasi kemudian dirata-ratakan sebagai estimasi akhir. Kombinasi kedua pendekatan ini dilakukan agar evaluasi tidak hanya bergantung pada satu kali pembagian data, tetapi juga mampu memberikan gambaran yang lebih konsisten terhadap kemampuan generalisasi model serta membantu mendeteksi potensi overfitting sejak dini.

2.5.5 Evaluasi Kinerja Model

Kinerja masing-masing model klasifikasi dalam penelitian ini juga dievaluasi menggunakan berbagai ukuran, termasuk; *accuracy*, *precision*, *recall* dan *F1-score*, *Macro Average* dan *Weighted Average* serta *confusion matrix*. Semua indikator ini dijumlahkan karena merupakan serangkaian dimensi berbeda dari model untuk memberikan gambaran tentang kondisi mental seseorang. Dalam penelitian ini, kondisi mental selanjutnya dibagi menjadi tiga, yaitu *relaxed*, *neutral*, dan *concentrating*. Oleh karena itu, *accuracy*, *recall*, dan *F1-score* dari setiap kelas diestimasi menggunakan strategi one-vs-rest dengan setiap kategori diambil sebagai kategori positif terhadap dua kategori lainnya yang digunakan sebagai kategori negatif. Rasio antara prediksi yang benar dan jumlah tes positif disebut *accuracy*. Menurut matematika, *accuracy* adalah:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

Karena ini adalah klasifikasi berganda dari tiga kelas target (*relaxed*, *neutral*, dan *concentrating*, cara lain adalah satu lawan sisanya), targetnya akan berupa TP, TN, FP, dan FN dan akan ditentukan melalui metode satu lawan sisanya. Sebagai contoh, di mana kelas *relaxed* diasumsikan sebagai kelas positif, TP (True Positive) akan menunjukkan jumlah sampel berlabel *relaxed* yang diprediksi dengan benar sebagai *relaxed*, TN (True Negative) akan menjadi jumlah sampel berlabel *neutral* atau *concentrating* yang diprediksi dengan benar sebagai bukan *relaxed* (yaitu, diprediksi dengan benar sebagai *neutral* atau *concentrating*), FP

(False Positive) akan menjadi jumlah sampel yang sebenarnya *neutral* atau *concentrating* tetapi diprediksi sebagai *relaxed*, dan FN (False Negative) akan menjadi jumlah sampel *relaxed* yang sebenarnya diprediksi sebagai *neutral* atau *concentrating*. Hal ini diperoleh dengan beralih antara memperlakukan *neutral* dan *concentrating* sebagai positif dan negatif dan sebaliknya, di mana setiap kelas menawarkan nilai *precision*, *recall*, dan *F1-score* sebelum dirata-ratakan sebagai *Macro Average* atau *Weighted Average*.

Precision adalah kemampuan model untuk akurat (dalam mengklasifikasikan label). Artinya, dari semua sampel yang klasifikasinya diketahui atau diperkirakan sesuai dengan kelas tertentu, ini adalah ukuran seberapa banyak dari sampel tersebut yang sebenarnya termasuk dalam kelas tersebut. Rumusnya adalah:

$$Precision = TP / (TP + FP)$$

Kemampuan untuk mengenali semua sampel dari kelas tertentu yang seharusnya berada dalam kategori spesifik, atau berapa banyak sampel dari kategori tertentu yang telah diidentifikasi secara akurat oleh model, diukur dengan *recall*. Rumusnya adalah:

$$Recall = TP / (TP + FN)$$

F1-score adalah rata-rata harmonik dari nilai *precision* dan nilai *recall*, oleh karena itu tidak akan berada di peringkat teratas kecuali *precision* dan *recall* keduanya memiliki nilai tinggi. Ukuran ini sangat berguna ketika distribusi kelas miring karena hanya memberikan satu angka yang menyeimbangkan *precision* dan *recall* secara bersamaan. Rumusnya adalah:

$$F1 - score = 2 \times (Precision \times Recall) / (Precision + Recall)$$

Selain keempat ukuran numerik tersebut, *confusion matrix* juga dapat menganalisis pola kesalahan klasifikasi secara lebih detail dalam bentuk tabel label aktual dan label yang diprediksi model. Berdasarkan *confusion matrix*, para peneliti juga dapat mengidentifikasi kelas-kelas yang paling membingungkan dan dengan demikian dapat melihat lebih jauh untuk memahami apakah kesalahan tersebut disebabkan oleh kualitas sinyal yang buruk, kesamaan fitur kelas, atau kelemahan model dalam mengidentifikasi pola yang kompleks. Cara evaluasi yang diberikan dapat dibenarkan oleh fakta bahwa, menurut penelitian, landasan penghitungan berbagai ukuran evaluasi *accuracy*, sensitivitas, nilai prediksi, dan *F1-score* masing-masing memiliki kekhasannya dan dapat diterapkan untuk mengukur kinerja model klasifikasi.

Macro Average bekerja dengan merata-ratakan semua ukuran (*precision*, *recall*, atau *F1-score*) di semua sampel, terlepas dari ukuran sampel

di setiap kelas. Dengan cara ini, semua jenis kelas (misalnya, *relaxed*, *neutral*, *berconcentrating*) akan mendapatkan skornya masing-masing yang kemudian akan dirata-ratakan dengan bantuan rumus berikut:

$$Macro Avg = \frac{1}{N} \sum_{i=1}^N Metrik_i$$

Salah satu metode untuk menghitung metrik rata-rata (*precision*, *recall*, atau *F1-score*) adalah *Weighted Average*, di mana bobot yang digunakan proporsional dengan ukuran masing-masing kelas. Ini berarti bahwa kelas dengan lebih banyak data akan lebih memengaruhi rata-rata keseluruhan daripada *Macro Average* yang juga memperhitungkan semua kelas. Rumusnya adalah:

$$Weighted Avg = \sum_{i=1}^N \left(\frac{n_i}{n_{total}} \times Metrik_i \right)$$

III. Hasil dan Pembahasan

3.1 Hasil Klasifikasi

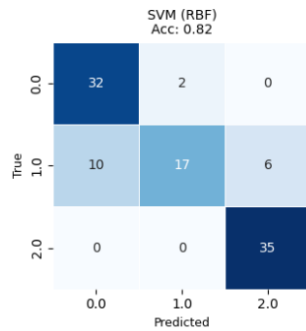
Eksperimen klasifikasi tiga kondisi mental, yaitu *relaxed*, *neutral*, dan *concentrating*, dengan bantuan sinyal EEG Muse, dilakukan pada lima algoritma *machine learning* klasik berikut: *Support Vector Machine* (SVM) dengan kernel RBF, K-Nearest Neighbor (KNN) dengan k=5, Decision Tree, Naive Bayes, dan Random Forest. 102 sampel uji memiliki ukuran kinerja sebagai berikut: *accuracy*, *precision Macro Average*, *recall Macro Average*, dan *F1-score Macro Average*.

1. Support Vector Machine (SVM)

Hasil yang diperoleh oleh *Support Vector Machine* (SVM) dengan pengujian berbasis EEG menggunakan kernel RBF menunjukkan *accuracy* tinggi dalam mengklasifikasikan tiga kondisi mental yaitu *relaxed*, *neutral*, dan *berconcentrating*. Laporan klasifikasi digunakan untuk menilai model dengan *precision*, *recall*, *F1-score*, dan skor *accuracy* dari setiap kelas kondisi mental. Keluaran kinerja model SVM dari setiap kelas diberikan pada Tabel 2 dan *confusion matrix* pada Gambar 2.

Tabel 2. Performa Model SVM

Kelas	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Support
<i>Relaxed</i>	0,76	0,94	0,84	34
<i>Neutral</i>	0,89	0,52	0,65	33
<i>Concentrating</i>	0,85	1,00	0,92	35
<i>Accuracy</i>			0,82	102
<i>Macro Avg</i>	0,84	0,82	0,81	102
<i>Weighted Avg</i>	0,84	0,82	0,81	102



Gambar 2 Confusion Matrix SVM

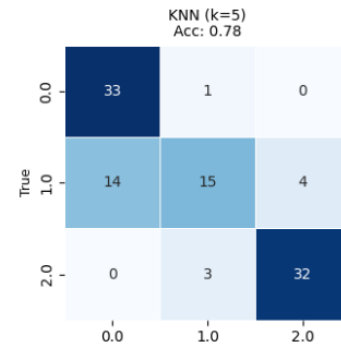
Confusion matrix SVM (RBF) menunjukkan bahwa klasifikasi sempurna dari kelas *concentrating* telah dilakukan; tidak ada satu pun kesalahan dalam mengklasifikasikan semua 35 sampel, yang berarti pola sinyal EEG selama *concentrating* penuh adalah yang paling mudah dikenali oleh model. Kinerja kelas relaksasi juga baik dengan 32/34 sampel diidentifikasi dengan benar dan 2 sampel lainnya diidentifikasi sebagai *neutral* yang salah. Sementara itu, tipe *neutral* menjadi kelemahan mendasar dari model ini: hanya 17 dari 33 sampel nyata yang dikenali dengan benar, 10 salah dikenali sebagai relaksasi dan 6 lainnya sebagai *concentrating*. Secara keseluruhan, *accuracy* model SVM (RBF) ternyata 0,82, *macro precision* 0,84, *macro recall* 0,82 dan *macro F1-score* 0,81. Hasil ini berarti bahwa SVM dapat membangun batas keputusan yang layak yang dapat mengidentifikasi dua keadaan berikut: relaksasi dan *concentrating* yang merupakan pola sinyal otak terkuat.

2. K-Nearest Neighbor (KNN)

Hasil pengujian melalui pendekatan KNN pada tiga kondisi mental yaitu *relaxed*, *neutral*, dan *berconcentrating* berdasarkan data rekaman EEG diuji dengan bantuan laporan klasifikasi yang berisi nilai *precision*, *recall*, *F1-score*, dan *accuracy* dari setiap kategori kondisi mental. Tabel 3 menunjukkan kinerja model KNN pada kinerja masing-masing kelas dan Gambar 3 menunjukkan *confusion matrix* KNN.

Tabel 3. Performa Model KNN

Kelas	Precision	Recall	F1-score	Support
Relaxed	0,70	0,97	0,81	34
Neutral	0,79	0,45	0,58	33
Concentrating	0,89	0,91	0,90	35
Accuracy			0,78	102
Macro Avg	0,79	0,78	0,76	102
Weighted avg	0,79	0,78	0,77	102



Gambar 3 Confusion Matrix KNN

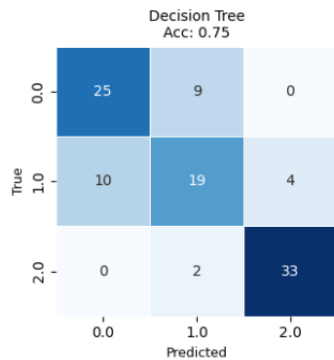
Accuracy skala besar keseluruhan dari model KNN k=5 adalah 0,78, yang berarti 80/102 sampel uji diklasifikasikan secara akurat. Angka ini tampaknya tidak buruk, tetapi hasilnya menunjukkan adanya ketidakseimbangan besar antar kelas. Kinerja kondisi *concentrating* dan relaksasi sangat baik dengan nilai *recall* masing-masing 0,91 dan 0,97. Hal ini dapat dijelaskan oleh fakta bahwa kedua kondisi tersebut memiliki karakteristik sinyal otak yang berbeda, yaitu peningkatan aktivitas gelombang beta saat dalam keadaan *concentrating* dan aktivitas gelombang alfa saat dalam keadaan relaksasi, yang dapat dideteksi oleh algoritma berbasis jarak seperti KNN yang dapat mengidentifikasinya dengan cukup mudah. Namun demikian, kebalikan dari proses tersebut menyebabkan kondisi *neutral* menjadi kelemahan utama model. Jumlah sampel yang benar-benar *neutral* adalah 33, di mana 15 diidentifikasi dengan benar dan sisanya diidentifikasi secara salah sebagai *relaxed*. Akibatnya, kategori *recall neutral* adalah yang terendah (0,45) dibandingkan dengan dua kategori lainnya, yang dapat dibandingkan dengan nilai yang jauh lebih tinggi.

3. Decision Tree

Hasil pengujian dengan algoritma Decision Tree dengan sinyal EEG ditentukan oleh *classification report* yang terdiri dari nilai *precision*, *recall*, *F1-score*, dan *accuracy*. Hasil model Decision Tree ditunjukkan pada Tabel 4 dan Gambar 4 menunjukkan *confusion matrix* Decision Tree.

Tabel 4. Performa Model Decision Tree

Kelas	Precision	Recall	F1-score	Support
Relaxed	0,71	0,74	0,72	34
Neutral	0,63	0,58	0,60	33
Concentrating	0,89	0,94	0,92	35
Accuracy			0,75	102
Macro Avg	0,75	0,75	0,75	102
Weighted Avg	0,75	0,75	0,75	102



Gambar 2 Confusion Matric Decision Tree

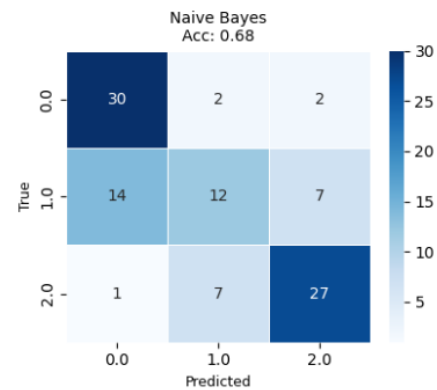
Model Decision Tree mampu mengklasifikasikan 77 dari 102 sampel dengan benar, menghasilkan *accuracy* 0,75. Namun, tidak ada pemerataan kinerja model di semua kelas di bawah angka ini. Kelas *concentrating* terbukti menjadi pengklasifikasi paling akurat dengan 33 dari 35 sampel diklasifikasikan dengan benar (*recall* 0,94). Hal ini karena tanda-tanda beta yang *terconcentrating* jauh lebih khas dibandingkan dengan tanda-tanda lainnya dan model ini dapat dengan mudah diterapkan untuk menghasilkan aturan keputusan yang mudah dibedakan. Sebaliknya, kelas *relaxed* dan *neutral* seringkali tercampur. Kelas *relaxed* memiliki tingkat ingatan terkuat sebesar 0,74 dengan 7 sampel diprediksi sebagai *neutral*, sedangkan kelas *neutral* memiliki tingkat ingatan terlemah dengan angka terendah, yaitu 0,58, karena 9 sampelnya tidak diprediksi dengan benar sebagai kelas *relaxed*.

4. Naive Bayes

Hasil pengujian dengan teknik Naive Bayes dalam pemisahan tiga kondisi mental yaitu *relaxed*, *neutral*, dan *berconcentrating* berdasarkan jejak EEG dianalisis dalam classification report yang terdiri dari nilai *precision*, *recall*, F1-score, dan *accuracy* setiap kategori kondisi mental. Tabel 5 menunjukkan output Naive Bayes ketika membuat prediksi kelas individual dan Gambar 5 menunjukkan *confusion matrix* Naive Bayes.

Tabel 5. Performa Model Naive Bayes

Kelas	Precisi	Recall	F1-score	Sup port
Relaxed	0,67	0,88	0,76	34
Neutral	0,57	0,36	0,44	33
Concentrating	0,75	0,77	0,76	35
Accuracy			0,68	102
Macro Avg	0,66	0,67	0,65	102
Weighted Avg	0,66	0,68	0,66	102



Gambar 3 Confusion Matric Naive Bayes

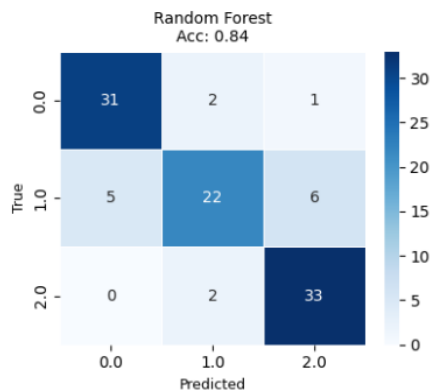
Accuracy Naive Bayes adalah yang terendah di antara semua model dan *accuracy* keseluruhan model ini adalah 0,68. Sampel uji mampu memprediksi dengan benar 69 dari 102 sampel uji dan 33 sampel sisanya salah diklasifikasikan. Kelas *Relaxed* cukup berhasil dengan modelnya dalam kemampuan untuk secara akurat mengidentifikasi sampel yang sebenarnya berada di kelas *Relaxed* dengan *recall* 0,88, atau 30 dari 34 sampel yang ditemukan. Meskipun demikian, skornya rendah yaitu 0,67 dalam hal *precision* yang menunjukkan bahwa model lebih cenderung melakukan klasifikasi *Relaxed*. Nilai terendah tercatat pada kategori *neutral*. Dari 33 sampel yang benar-benar *neutral*, hanya 12 sampel yang diidentifikasi dengan benar. Semua sampel lainnya dialokasikan ke kelas lain: 17 sampel diduga *relaxed* dan 4 sampel diduga *berconcentrating*. Nilai *recall* terendah sebesar 0,36 dan F1-score sebesar 0,44 pada model-model tersebut tampak paling lemah di antara semua model yang digunakan.

5. Random Forest

Hasil pengujian berdasarkan pendekatan Random Forest untuk mengklasifikasikan tiga kondisi mental; *relaxed*, *neutral*, dan *concentrating* menggunakan sinyal EEG diperoleh dengan bantuan classification report yang mencantumkan nilai *precision*, *recall*, F1-score, dan *accuracy* klasifikasi dari setiap kondisi mental. Tabel 6 menunjukkan hasil kinerja model Random Forest pada masing-masing kelas dan Gambar 6 menunjukkan *confusion matrix* Random Forest.

Tabel 6. Performa Model Random Forest

Kelas	Precisi	Recall	F1-score	Sup port
Relaxed	0,86	0,91	0,89	34
Neutral	0,85	0,67	0,75	33
Concentrating	0,82	0,94	0,88	35
Accuracy			0,84	102
Macro Avg	0,84	0,84	0,84	102
Weighted avg	0,84	0,84	0,84	102



Gambar 4. Confusion Matric Random Forest. Dari semua model yang diuji, yang paling efektif adalah Random Forest yang menggunakan 200 Decision Tree dan memiliki *accuracy* 0,84. Nilai macro *precision*, *recall*, dan F1-score juga selalu 0,84, yang menunjukkan bahwa model tersebut tidak menunjukkan bias dalam kinerjanya terhadap kelas tertentu. Delapan puluh enam dari 102 sampel uji berhasil diidentifikasi. Model kelas *Relaxed* telah membuktikan dirinya sangat akurat dengan 31 dari 34 sampel yang disebut benar (*recall* 0,91) dan yang terdekat dengan itu, 3 sampel disebut *neutral* dengan benar. Tidak ada pencampuran sampel *Relaxed* dengan sampel *Concentrating*, dan itu menyiratkan bahwa model tersebut mampu memperkenalkan tingkat pengambilan keputusan tertentu di antara dua kondisi yang pada kenyataannya sangat berlawanan satu sama lain. Klasifikasi yang tepat dari sampel adalah 33/35 dalam kasus *concentrating* sampel (*recall* 0,94). Terdapat 2 false negative dalam memprediksi *neutral* dan tidak ada false positive yang diabaikan. Keuntungan terbaik dicapai pada Random Forest dalam kasus kelas *neutral* yang merupakan kesalahan utama dari semua model. Kelas *neutral* adalah yang tertinggi dari model yang diuji sebelumnya dengan *recall* 0,67 (22/33) yang merupakan tertinggi di antara model-model tersebut. Semua hal ini menjadikan Random Forest sebagai yang paling menjanjikan di antara semuanya karena akan menyediakan dasar sistem pemantauan kondisi kognitif siswa berbasis EEG yang dapat dikenakan (*wearable ECG*).

Temuan yang muncul secara konsisten pada seluruh model yang telah diujikan menunjukkan bahwa performa klasifikasi kelas paling rendah terjadi pada kondisi netral. Secara neurofisiologis, kondisi tersebut bersifat ambigu, di mana sinyal EEG tidak memperlihatkan pola dominan seperti peningkatan gelombang alfa pada keadaan relaks maupun peningkatan beta saat fokus. Oleh karena itu, distribusi frekuensi pada kondisi netral cenderung saling tumpang tindih dengan dua kelas lainnya, sehingga batas

pemisah yang dibentuk oleh model menjadi kurang jelas. Untuk mengatasi keterbatasan ini, penelitian selanjutnya dapat mengarah pada penambahan fitur konektivitas fungsional antar kanal seperti coherence, atau memanfaatkan teknik augmentasi data sintetis agar variasi sampel pada kelas netral menjadi lebih representatif.

3.2 Perbandingan Kinerja Antar Algoritma

Untuk membandingkan dan menentukan algoritma mana yang paling sesuai, Tabel 7 merangkum hasil dari lima model pada subbagian di atas dan membandingkannya dalam konteks statistik evaluasi (*accuracy*, *precision*, *recall*, dan F1-score).

Tabel 7 Perbandingan Kinerja Algoritma per Metrik Pengukuran

Algoritma	Accur acy	Precis ion	Recall	F1- score
SVM (RBF)	0,82	0,84	0,82	0,81
KNN (k=5)	0,78	0,79	0,78	0,76
Decision Tree	0,75	0,75	0,75	0,75
Naive Bayes	0,68	0,66	0,67	0,65
Random Forest	0,84	0,84	0,84	0,84

Random Forest selalu lebih akurat dalam tiga dari empat ukuran: *accuracy* tertinggi (0,84), *recall* tertinggi (0,84), dan F1-score tertinggi (0,84). Dalam ukuran *precision*, random forest setara dengan SVM yang juga menerima nilai ukuran 0,84 dan keduanya merupakan yang paling akurat dalam mengklasifikasikan kelas. Meskipun demikian, SVM sedikit lebih rendah dalam *recall* (0,82) dan F1-score (0,81) daripada Random Forest. Tiga posisi berikutnya di hampir semua metrik ditempati oleh KNN, Decision Tree, dan Naive Bayes secara berturut-turut, dan nilai terendah di keempat metrik tersebut diraih oleh Naive Bayes. Kecenderungan ini membenarkan gagasan bahwa dalam eksperimen ini, Random Forest adalah algoritma yang paling seimbang dan berkinerja terbaik secara umum dalam hal tugas mengklasifikasikan tiga kondisi mental, dan SVM ternyata menjadi yang terbaik kedua dalam kasus ini.

Hasil penelitian ini menunjukkan perbedaan yang cukup mencolok dibandingkan studi Rahman dan Nasor yang melaporkan akurasi di atas 99% pada dataset dan perangkat yang serupa (headband Muse) (Rahman & Nasor, 2024). Perbedaan ini tidak dapat langsung diartikan sebagai kelemahan metode yang digunakan, melainkan lebih mungkin dipengaruhi oleh perbedaan pendekatan evaluasi, seperti skema validasi yang digunakan. Studi tersebut menerapkan metode seperti leave-one-subject-out atau cross-validation yang tentunya berbeda dengan yang

digunakan dalam penelitian ini, sehingga hasil akan berbeda. Selain itu, variasi dalam melakukan proses *ekstraksi* fitur, termasuk kemungkinan penggunaan fitur domain frekuensi yang lebih detail atau metode seperti wavelet, juga dapat memengaruhi hasil akhir. kemungkinan pula adanya potensi data *leakage* pada studi tersebut apabila proses preprocessing dilakukan sebelum pemisahan data latih dan uji. Dengan demikian, akurasi 0,84 yang diperoleh dalam penelitian ini dapat dipandang sebagai hasil yang lebih konservatif, namun lebih mencerminkan kemampuan generalisasi model dalam kondisi evaluasi yang ketat.

4. Keterbatasan dan rekomendasi

Meskipun hasil yang dicapai cukup menjanjikan, sejumlah keterbatasan perlu dipertimbangkan. Tugas-tugas untuk menghasilkan kondisi *relaxed*, *neutral*, dan *concentrating* tidak spesifik pada situasi akademis, yaitu aritmatika atau masalah matematika, melainkan bersifat umum. Ini adalah salah satu celah yang perlu diatasi dalam penelitian selanjutnya dengan mempertimbangkan bahwa tujuan utama penelitian saat ini adalah untuk mempelajari tentang keadaan mental dalam melakukan pembelajaran numerik. Oleh karena itu, penelitian generasi berikutnya akan didasarkan pada karakter tugas yang lebih matematis serta meneliti sejauh mana temuan kategorisasi tersebut terkait dengan keterampilan matematika para peserta.

5. Kesimpulan

Data EEG mentah yang direkam oleh perangkat *wearable* Muse dengan biaya murah, mampu memberikan informasi yang cukup kuat untuk mengkategorikan tiga kondisi mental yaitu : *relaxed*, *neutral*, dan *berconcentrating*. penelitian ini menggunakan lima algoritma *machine learning* klasik yaitu SVM, KNN, Decision Tree, Naive Bayes, dan Random Forest. Algoritma yang paling efektif di antara algoritma-algoritma ini adalah random forest dengan performa 0,84 berdasarkan nilai rata-rata macro *precision*, *recall*, dan F1-score masing-masing sebesar 0,84, yang berarti algoritma ini bekerja sama baiknya untuk setiap kelas. SVM dalam kernel RBF dengan *accuracy* 0,82 adalah alternatif terbaik kedua. Sementara itu, tantangan dan hambatan utama dalam penelitian ini yaitu mengklasifikasi kelas *neutral* yang paling sulit ditangani oleh semua model karena kelas ini berada antara kondisi *relaxed* dan *concentrating* yang menyebabkan batasan antar kelas menjadi kurang jelas.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa data EEG yang diperoleh

dengan perangkat *wearable* dapat menjadi dasar untuk menciptakan pemantauan real-time terhadap kondisi fungsi kognitif dalam lingkungan pembelajaran. Dengan mengintegrasikan algoritma *Random Forest* dan platform sistem pemantauan adaptif, berpotensi pengajar dapat melakukan monitoring *concentrating* siswa secara real-time yang dapat digunakan untuk mendukung proses pembelajaran adaptif di kelas.

Untuk penelitian selanjutnya, peneliti dapat meningkatkan jumlah dan keragaman subjek sehingga model yang dihasilkan menjadi lebih baik. Penggunaan stimulus yang lebih terstruktur dan kontekstual terhadap aktivitas proses pembelajaran, seperti soal pemecahan masalah matematika dapat dipertimbangkan sehingga model yang dihasilkan lebih sesuai dengan tujuan pemantauan kesiapan belajar siswa. Selain itu, eksplorasi penerapan pendekatan deep learning seperti *Long Short-Term Memory (LSTM)* maupun *Convolutional Neural Network (CNN)* juga berpotensi digunakan untuk menangkap pola gelombang sinyal EEG yang lebih kompleks, khususnya dalam mengatasi tantangan dan hambatan pada klasifikasi dengan kondisi *netral*.

Daftar Pustaka

- Al-Nafjan, A., & Aldayel, M. (2022). Predict Students' Attention in Online Learning Using EEG Data. *Sustainability (Switzerland)*, 14(11). <https://doi.org/10.3390/su14116553>
- Alreshidi, I., Moulitsas, I., & Jenkins, K. W. (2023). Multimodal Approach for Pilot Mental State Detection Based on EEG. *Sensors*, 23(17). <https://doi.org/10.3390/s23177350>
- Balconi, M., Rovelli, K., Angioletti, L., & Allegratta, R. A. (2024). Working Memory Workload When Making Complex Decisions: A Behavioral and EEG Study. *Sensors*, 24(17). <https://doi.org/10.3390/s24175754>
- Bird, J. J., Manso, L. J., Ribeiro, E. P., Ekart, A., & Faria, D. R. (2018a). A Study on Mental State Classification using EEG-based Brain-Machine Interface. *2018 International Conference on Intelligent Systems (IS)*, 795–800. <https://doi.org/10.1109/IS.2018.8710576>
- Bird, J. J., Manso, L. J., Ribeiro, E. P., Ekart, A., & Faria, D. R. (2018b). A Study on Mental State Classification using EEG-based Brain-Machine Interface. *2018 International Conference on Intelligent Systems (IS)*, 795–800. <https://doi.org/10.1109/IS.2018.8710576>
- Chaddad, A., Wu, Y., Kateb, R., & Bouridane, A. (2023). Electroencephalography Signal Processing: A Comprehensive Review and Analysis of Methods and Techniques. In *Sensors* (Vol. 23, Number 14). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/s23146434>
- Chen, Y., Wang, W., Yan, S., Wang, Y., Zheng, X., & Lv, C. (2024). Application of Electroencephalography Sensors and Artificial Intelligence in Automated Language Teaching. *Sensors*, 24(21). <https://doi.org/10.3390/s24216969>
- Coman, D. A., Ionita, S., & Lita, I. (2024). Evaluation of EEG Signals by Spectral Peak Methods and Statistical Correlation for Mental State Discrimination Induced by Arithmetic Tasks. *Sensors*, 24(11). <https://doi.org/10.3390/s24113316>
- Do, K. T., Paolizzi, S. G., & Hallquist, M. N. (2024). How

- adolescents learn to build social bonds: A developmental computational account of social explore-exploit decision-making. *Developmental Cognitive Neuroscience*, 69. <https://doi.org/10.1016/j.dcn.2024.101415>
- Gkintoni, E., & Halkiopoulos, C. (2025). Mapping EEG Metrics to Human Affective and Cognitive Models: An Interdisciplinary Scoping Review from a Cognitive Neuroscience Perspective. In *Biomimetics* (Vol. 10, Number 11). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/biomimetics10110730>
- Houssein, E. H., Hammad, A., & Ali, A. A. (2022). Human emotion recognition from EEG-based brain-computer interface using machine learning: a comprehensive review. In *Neural Computing and Applications* (Vol. 34, Number 15, pp. 12527–12557). Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1007/s00521-022-07292-4>
- Jan, D., Rico, E., Birba, A., Ravelo, Y., León-Mendez, M., Travina, K., Padrón, I., de Vega, M., & Marrero, H. (2025). Non-Invasive Classification of Mental Health Disorders Using Resting-State EEG with Dry Electrodes for Scalable Triage. *Cognitive Computation*, 17(6). <https://doi.org/10.1007/s12559-025-10529-8>
- Khare, S. K., Bajaj, V., Gaikwad, N. B., & Sinha, G. R. (2023). Ensemble Wavelet Decomposition-Based Detection of Mental States Using Electroencephalography Signals. *Sensors*, 23(18). <https://doi.org/10.3390/s23187860>
- Kotte, S., Elkhoully, A., Abd Malek, M. F., & Abohaia, Z. (2025). Deep learning-based EEG mental state classification to support mental focus in female cricketers. *Discover Artificial Intelligence*, 5(1). <https://doi.org/10.1007/s44163-025-00615-z>
- Lekati, E., Dimitrakopoulos, G. N., Lazaros, K., Giannopoulou, P., Vrahatis, A. G., Krokidis, M. G., Vlamos, P., & Doukakis, S. (2025a). Wearable EEG Sensor Analysis for Cognitive Profiling in Educational Contexts. *Sensors*, 25(20). <https://doi.org/10.3390/s25206446>
- Lekati, E., Dimitrakopoulos, G. N., Lazaros, K., Giannopoulou, P., Vrahatis, A. G., Krokidis, M. G., Vlamos, P., & Doukakis, S. (2025b). Wearable EEG Sensor Analysis for Cognitive Profiling in Educational Contexts. *Sensors*, 25(20). <https://doi.org/10.3390/s25206446>
- Liu, Z., & Zhao, J. (2024). Leveraging deep learning for robust EEG analysis in mental health monitoring. *Frontiers in Neuroinformatics*, 18. <https://doi.org/10.3389/fninf.2024.1494970>
- Orovass, C., Sapounidis, T., Volioti, C., & Keramopoulos, E. (2025). EEG in Education: A Scoping Review of Hardware, Software, and Methodological Aspects. In *Sensors* (Vol. 25, Number 1). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/s25010182>
- Pramitarini, Y., Perdana, R. H. Y., Shim, K., & An, B. (2023). DLsMR: Deep Learning-Based Secure Multicast Routing Protocol against Wormhole Attack in Flying Ad Hoc Networks with Cell-Free Massive Multiple-Input Multiple-Output †. *Sensors*, 23(18). <https://doi.org/10.3390/s23187960>
- Rahman, K. K. M., & Nasor, K. M. (2024). Classification of Mental State Using a Muse Headband and Machine Learning Algorithm. In *Advances in Signal Processing and Communication Engineering* (pp. 453–462). https://doi.org/10.1007/978-981-97-0562-7_34
- Ramirez-Arias, F. J., García-Guerrero, E. E., Tlelo-Cuautle, E., Colores-Vargas, J. M., García-Canseco, E., López-Bonilla, O. R., Galindo-Aldana, G. M., & Inzunza-González, E. (2022). Evaluation of Machine Learning Algorithms for Classification of EEG Signals. *Technologies*, 10(4). <https://doi.org/10.3390/technologies10040079>
- Sadegh-Zadeh, S.-A. (2024). Comparative analysis of dimensionality reduction techniques for EEG-based emotional state classification. *American Journal of Neurodegenerative Disease*, 13(4), 23–33. <https://doi.org/10.62347/zwry8401>
- Syed, M. K., Wang, H., Siddiqi, A. A., Qureshi, S., & Gouda, M. A. (2025). EEG-Based Attention Classification for Enhanced Learning Experience. *Applied Sciences (Switzerland)*, 15(15). <https://doi.org/10.3390/app15158668>
- Uudeberg, T., Päske, L., Hinrikus, H., Lass, J., Pöld, T., & Bachmann, M. (2025). Individual stability of single-channel EEG measures over one year in healthy adults. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-13614-y>
- Walla, P. (2025a). The Power of Time: Editorial on the Advantages of Electroencephalography (EEG) and Event-Related Potentials (ERPs) in Affective and Cognitive Neuroscience. In *Brain Sciences* (Vol. 15, Number 10). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/brainsci15101054>
- Walla, P. (2025b). The Power of Time: Editorial on the Advantages of Electroencephalography (EEG) and Event-Related Potentials (ERPs) in Affective and Cognitive Neuroscience. In *Brain Sciences* (Vol. 15, Number 10). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/brainsci15101054>
- Wang, J., Huo, S., Wu, K. C., Mo, J., Wong, W. L., & Maurer, U. (2022). Behavioral and neurophysiological aspects of working memory impairment in children with dyslexia. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-022-16729-8>
- Xie, H., Jia, C., Luo, Y., He, J., Dong, Z., Liang, D., Ren, Z., Jiang, M., Gao, X., & Chen, X. (2025). EEG analysis of brain dynamics in a simulated multi-task and multi-stage learning environment. *Npj Science of Learning*, 10(1). <https://doi.org/10.1038/s41539-025-00376-5>
- Yildirim, B., Ulkir, O., Kaya, M., Singh, A. K., & Krishnan, S. (2023). Trends in EEG signal feature extraction applications. <https://doi.org/10.3389/frai.2022.1072801>
- Yudhana, A., Muslim, A., Wati, D. E., Puspitasari, I., Azhari, A., & Mardhia, M. M. (2020). Human emotion recognition based on EEG signal using fast fourier transform and K-Nearest neighbor. *Advances in Science, Technology and Engineering Systems*, 5(6), 1082–1088. <https://doi.org/10.25046/aj0506131>
- Yudhana, A., Umar, R., & Saputra, S. (2022). Fish Freshness Identification Using Machine Learning: Performance Comparison of k-NN and Naïve Bayes Classifier. *Journal of Computing Science and Engineering*, 16(3), 153–164. <https://doi.org/10.5626/JCSE.2022.16.3.153>
- Yuliansyah, H., Ardiansyah, R., & Yudhana, A. (2026). A Machine Learning Approach for Multi-Label Classification in Candidate Election Social Media Analysis. *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, 20(1), 50–62. <https://doi.org/10.37936/ecti-cit.2026201.263596>
- Zuo, M., Chen, X., & Sui, L. (2025). Evaluation of Machine Learning Algorithms for Classification of Visual Stimulation-Induced EEG Signals in 2D and 3D VR Videos. *Brain Sciences*, 15(1). <https://doi.org/10.3390/brainsci15010075>