

## Deteksi Attention Deficit Hyperactivity Disorder (ADHD) Berbasis Electroencephalogram (EEG) : Pendekatan Klasifikasi Dengan Random Forest

Allwin M. Simarmata<sup>1</sup>, Mizanul Akbar<sup>1\*</sup>, Alexander Pakpahan<sup>1</sup>, Godlieb Sihombing<sup>1</sup>, M. Syaidin Abdiyanuarsyah Ginting<sup>1</sup>

<sup>1</sup>Program Studi Teknik Informatika, Universitas Prima Indonesia, Indonesia

---

### Artikel Info

#### Kata Kunci:

ADHD;  
Deteksi;  
Electroencephalogram;  
Klasifikasi;  
Random Forest.

#### Keywords:

ADHD;  
Detection  
Electroencephalogram;  
Classification; Random  
Forest;

#### Riwayat Artikel:

Submitted: 11 Juni 2025  
Accepted: 31 Juli 2025  
Published: 31 Juli 2025

**Abstrak:** *Attention Deficit Hyperactivity Disorder (ADHD)* merupakan gangguan perkembangan yang sering dijumpai pada anak-anak dan berdampak signifikan terhadap proses belajar serta kualitas hidup. Proses diagnosis ADHD saat ini masih didominasi oleh observasi klinis yang bersifat subjektif, sehingga diperlukan pendekatan yang lebih objektif. Penelitian ini bertujuan untuk mendeteksi ADHD menggunakan data *Electroencephalogram (EEG)* yang diproses melalui algoritma klasifikasi Random Forest. Data EEG dianalisis melalui tahapan pra-pemrosesan, ekstraksi fitur, dan klasifikasi. Hasil penelitian menunjukkan bahwa model Random Forest mampu membedakan antara data EEG penderita ADHD dan kelompok kontrol dengan akurasi sebesar 0,74. Penelitian ini berkontribusi pada upaya memperluas aplikasi machine learning di bidang kesehatan, khususnya dalam upaya diagnosis dini dan penanganan ADHD yang lebih baik, objektif, dan terintegrasi teknologi. Selain itu, aplikasi deteksi ADHD berbasis Streamlit dikembangkan untuk memfasilitasi proses deteksi secara interaktif dan visualisasi data EEG. Temuan ini menunjukkan potensi pemanfaatan algoritma Random Forest dalam deteksi ADHD berbasis EEG secara lebih akurat dan efisien, meskipun peningkatan akurasi dan eksplorasi metode klasifikasi lain masih diperlukan. Penelitian ini diharapkan dapat menjadi dasar pengembangan sistem deteksi ADHD yang lebih objektif dan konsisten.

**Abstract:** *Attention Deficit Hyperactivity Disorder (ADHD)* is a developmental disorder frequently found in children and significantly impacts learning processes and quality of life. Currently, ADHD diagnosis is still dominated by subjective clinical observation, highlighting the need for a more objective approach. This study aims to detect ADHD using *Electroencephalogram (EEG)* data processed through the Random Forest classification algorithm. EEG data were analyzed through preprocessing, feature extraction, and classification stages. The results show that the Random Forest model can distinguish between EEG data of ADHD patients and control groups with an accuracy of 0.74. This research contributes to efforts to expand the application of machine learning in the health sector, particularly in the efforts for early diagnosis and better, more objective, and integrated management of ADHD. Additionally, a Streamlit-based ADHD detection application was developed to facilitate interactive detection and EEG data visualization. These findings demonstrate the potential of utilizing the Random Forest algorithm for more accurate and efficient EEG-based ADHD detection, although further improvements in accuracy and exploration of other classification methods are still needed. This research is

*expected to serve as a foundation for developing a more objective and consistent ADHD detection system.*

---

**Corresponding Author:**

Mizanul Akbar

Email: akbarnst064@gmail.com

---

## PENDAHULUAN

Anak yang berkebutuhan khusus adalah anak yang memerlukan penanganan khusus sehubungan dengan gangguan perkembangan dan masa pertumbuhan yang dialaminya (Mulyati, 2023). Salah satu gangguan di dalam masa pertumbuhan dan perkembangan anak yaitu *Attention-deficit Hyperactivity Disorder* (ADHD) (Fitriyani et al., 2023). ADHD adalah salah satu gangguan perkembangan anak dalam peningkatan aktivitas motorik anak sehingga menyebabkan anak hiperaktif (Sadida qaulansadida et al., 2024). Hal ini ditunjukkan dengan beberapa keluhan yaitu seperti gelisah, tidak dapat diam, selalu ingin bergerak, sering meninggalkan keadaan yang tetap, tidak bisa memusatkan perhatian serta menunjukkan impulsivitas sehingga anak akan mengalami kesulitan dalam belajar. ADHD terjadi disebabkan oleh beberapa faktor, diantaranya faktor biologis dan lingkungan, paparan racun/zat kimia berbahaya, kekurangan gizi, faktor neorologis dan faktor prenatal (Maghfirah et al., 2024). Selain itu, hasil deteksi ADHD juga dapat dipengaruhi oleh variabel demografis seperti usia, jenis kelamin, dan latar belakang sosial ekonomi (Nurfadhillah et al., 2021). Anak laki-laki, misalnya, dilaporkan lebih sering terdiagnosis ADHD dibandingkan anak perempuan (Prasaja et al., 2022). Kondisi kesehatan lain seperti gangguan tidur, gangguan kecemasan, atau adanya riwayat gangguan perkembangan pada keluarga juga dapat menjadi faktor yang mempengaruhi akurasi deteksi ADHD.

Terkait dengan penentuan diagnosis dan penanganan ADHD, salah satu alat yang dapat digunakan adalah *Electroencephalogram* (EEG) (Mairita et al., 2024). EEG merupakan alat diagnostik yang penting dalam mendeteksi gangguan kejiwaan seperti ADHD karena kemampuannya untuk mendeteksi aktivitas listrik otak yang abnormal (Fasya & Sari, 2024). Sinyal EEG berupa gelombang listrik yang sangat kecil membuat pendeteksian visual secara langsung menjadi sulit. Oleh karena itu, sinyal EEG direkam untuk memberikan informasi mendalam tentang aktivitas listrik otak (Devianto et al., 2024). Dalam analisis data EEG untuk deteksi ADHD, sejumlah algoritma *machine learning* seperti *Random Forest* (RF) sebuah metode klasifikasi berbasis pohon keputusan (*decision tree*) telah digunakan untuk meningkatkan akurasi deteksi (Iskandar et al., 2024). Kelebihan utama *Random Forest* adalah kemampuannya dalam mengatasi *overfitting* dan meningkatkan akurasi dengan menggabungkan prediksi dari beberapa model (Nurohanisah et al., 2024). Hal ini sangat relevan dalam konteks analisis data EEG yang kompleks, nonlinier, dan rentan *noise*, seperti pola peningkatan gelombang theta dan penurunan aktivitas beta yang menjadi penanda ADHD (Pagiling et al., 2020).

Penelitian ini memiliki beberapa kelebihan dibandingkan penelitian-penelitian sebelumnya yang juga menggunakan model *Random Forest* untuk deteksi ADHD. Salah satu keunggulan utama adalah penggunaan dataset EEG yang jauh lebih besar dan komprehensif, yakni sebanyak 1.048.575 data dengan 19 kanal EEG pada frekuensi sampling 128 Hz, sehingga memberikan gambaran yang lebih detail dan representatif terhadap aktivitas otak penderita ADHD. Selain itu, penelitian ini tidak hanya berfokus pada deteksi keberadaan ADHD, tetapi juga mengklasifikasi tingkat keparahan gejala yang dialami oleh penderita, sementara penelitian terdahulu umumnya hanya membedakan antara ADHD dan non-ADHD. Penelitian ini juga menonjolkan aspek praktis dan *reproducibility* dengan pemanfaatan Python di Google Colab, sehingga mudah diadopsi dan dikembangkan lebih lanjut oleh peneliti lain.

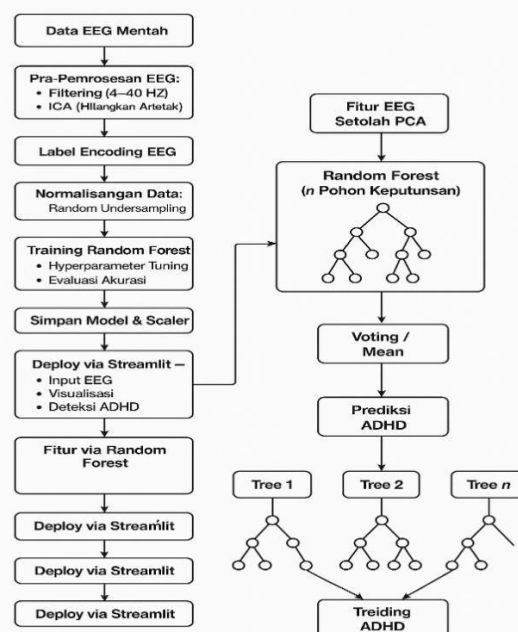
Dalam lima tahun terakhir, berbagai penelitian telah dilakukan untuk mendeteksi ADHD menggunakan data EEG dan algoritma *machine learning*. Amandeep Kaur, et al (Kaur & Kahlon, 2022)

menguji enam algoritma *machine learning* pada data aktivitas motorik *real-time* untuk identifikasi ADHD pada orang dewasa dan menemukan SVM sebagai algoritma dengan akurasi terbaik, namun penelitian ini tidak menggunakan data EEG dan tidak mengklasifikasikan tingkat keparahan ADHD. Jun Won Kim, et al (Kim et al., 2025) memanfaatkan data EEG dan XGBoost untuk diagnosis ADHD dengan hasil akurasi 90,81%, namun belum mengklasifikasikan tingkat keparahan gejala dan menggunakan dataset yang terbatas. Nishant Chauhan, et al (Chauhan & Choi, 2023) mengembangkan model *machine learning* untuk membedakan pola EEG pada ADHD, tetapi hanya membedakan ADHD dan non-ADHD tanpa klasifikasi tingkat keparahan. Hui Wen Loh, et al (Loh et al., 2024) menggunakan deep learning berbasis CNN pada data EEG dan mencapai akurasi tinggi, namun fokus pada identifikasi saluran EEG signifikan, bukan pada klasifikasi tingkat keparahan ADHD. Sementara itu, Md. Maniruzzaman, et al (Maniruzzaman et al., 2022) meneliti seleksi fitur EEG dengan LASSO dan beberapa algoritma ML, namun penelitian ini juga belum mengklasifikasikan tingkat keparahan gejala. Penelitian Hui Tian Tor, et al (Tor et al., 2021) lebih menekankan pada diferensiasi antar gangguan (ADHD, CD, dan ADHD+CD) menggunakan sinyal EEG, bukan klasifikasi tingkat keparahan ADHD. Dari berbagai penelitian tersebut, dapat disimpulkan bahwa sebagian besar penelitian masih berfokus pada deteksi keberadaan ADHD (klasifikasi ADHD vs non-ADHD) atau diferensiasi dengan gangguan lain, serta menggunakan dataset berukuran relatif kecil dan fitur terbatas. Belum banyak penelitian yang secara spesifik mengklasifikasikan tingkat keparahan ADHD menggunakan dataset EEG yang besar, multi-kanal, serta menonjolkan aspek praktis dan reproducibility pada platform terbuka seperti Google Colab.

Penelitian ini bertujuan untuk mengidentifikasi pola-pola spesifik dalam aktivitas gelombang otak yang dapat membantu dalam deteksi ADHD serta mengklasifikasi penderita ADHD yang telah terdeteksi dengan algoritma *Random Forest*. Diharapkan bahwa penelitian ini tidak hanya mampu mengidentifikasi keberadaan ADHD secara akurat, tetapi juga dapat mendeteksi tingkat keparahan gejala yang dialami oleh penderita, dengan memanfaatkan algoritma *Random Forest* yang dikenal karena kemampuannya dalam menangani data berukuran besar dan kompleks serta memberikan hasil yang *robust*, sehingga dapat meningkatkan akurasi dalam deteksi ADHD.

## METODE

Penelitian ini menggunakan pendekatan berbasis data EEG untuk mendeteksi Attention Deficit Hyperactivity Disorder (ADHD) dengan memanfaatkan algoritma Random Forest sebagai metode klasifikasi utama. Berikut adalah alur metodologi penelitian yang dilakukan ;



Gambar 1. Prosedur Kerja Penelitian

### Exploratory Data Analysis (EDA)

Sebelum membangun model, dilakukan analisis eksplorasi data (EDA) untuk memahami karakteristik dataset, meliputi penampilan beberapa baris awal data, pemeriksaan tipe data dan jumlah sampel, perhitungan statistik deskriptif seperti *mean*, median, dan standar deviasi, visualisasi distribusi kelas target untuk menilai keseimbangan data, serta pembuatan heatmap korelasi guna memahami hubungan antar fitur dalam dataset.

### Preprocessing Data

*Preprocessing* merupakan tahap penting sebelum pembangunan model, yang mencakup pemisahan fitur EEG (X) dan label status ADHD atau non-ADHD (y), pengubahan label menjadi nilai numerik menggunakan *LabelEncoder*, serta normalisasi fitur dengan *StandardScaler* agar skala data seragam. Jika terjadi ketidakseimbangan kelas, diterapkan teknik *Random Undersampling* untuk menyeimbangkan jumlah sampel antar kelas. Selain itu, reduksi dimensi dilakukan dengan *Principal Component Analysis* (PCA) guna mengurangi kompleksitas model dan mempertahankan 95% varians data, sehingga hanya fitur relevan yang digunakan dalam proses selanjutnya.

### Pemodelan

Penelitian ini menggunakan algoritma *Random Forest*, yang dikenal sebagai metode *ensemble learning* yang efektif untuk klasifikasi karena mampu menangani data berdimensi tinggi dan tahan terhadap *overfitting*. Dataset dibagi menjadi data latih (70%) dan data uji (30%) menggunakan *train\_test\_split*, di mana data latih digunakan untuk melatih model dan data uji untuk evaluasi. Model *Random Forest* dilatih dengan data latih menggunakan *hyperparameter default*, yaitu *n\_estimators=100* dan *random\_state=42* untuk memastikan hasil yang konsisten.

### Pengujian

Sebanyak 30% dari total data EEG yang telah diproses digunakan sebagai data uji untuk menguji kinerja akhir model *Random Forest* dalam mendeteksi ADHD. Data uji ini terdiri dari data EEG penderita ADHD dan non-ADHD yang dipilih secara acak dari dataset utama melalui proses *train\_test\_split*, setelah melalui tahapan *preprocessing*. Data uji ini tidak digunakan selama proses pelatihan model, sehingga hasil evaluasi pada data uji dapat mencerminkan kemampuan model dalam mengklasifikasikan data baru secara objektif dan memberikan gambaran akurasi model dalam mendeteksi ADHD pada data yang belum pernah dilihat sebelumnya.

### Evaluasi Model

Setelah model dilatih, evaluasi dilakukan menggunakan data uji sebanyak 30% dari dataset utama dengan beberapa metrik, yaitu akurasi untuk mengukur ketepatan prediksi, *classification report* yang mencakup *precision*, *recall*, dan *F1-score* untuk setiap kelas, serta *confusion matrix* untuk melihat jumlah prediksi benar dan salah pada tiap kelas. Selain itu, performa model juga dianalisis melalui ROC curve dan nilai AUC, yang menunjukkan keseimbangan antara *true positive rate* dan *false positive rate*, di mana AUC mendekati 1 menandakan performa yang baik. *Random Forest* juga memberikan informasi *feature importance*, sehingga fitur-fitur yang paling berpengaruh dalam pengambilan keputusan model dapat diidentifikasi.

### Optimasi Model

Untuk meningkatkan performa model, dilakukan optimasi *hyperparameter* menggunakan *Grid Search* atau *Manual Tuning*. Beberapa kombinasi *hyperparameter* seperti *n\_estimators*, *max\_depth*, dan *min\_samples\_split* diuji, dan kombinasi terbaik dipilih berdasarkan akurasi tertinggi. *Cross-Validation* (Validasi silang) dilakukan untuk memastikan model tidak *overfitting* dan memiliki performa yang konsisten. Data dibagi menjadi beberapa subset, dan model dilatih serta diuji pada subset yang berbeda.

### Preprocessing Khusus Data EEG

Data EEG memerlukan *preprocessing* khusus karena sifatnya yang kompleks dan rentan terhadap *noise*. Proses ini meliputi *filtering* dengan *bandpass filter* (4-40 Hz) untuk menghilangkan *noise* frekuensi rendah dan tinggi, serta penggunaan *Independent Component Analysis* (ICA) untuk memisahkan sinyal EEG menjadi komponen independen guna menghapus artefak seperti gerakan mata. Hasil *preprocessing*

kemudian divisualisasikan melalui *plot raw* EEG dan *power spectral density* (PSD) sebelum dan sesudah *filtering*, untuk memastikan *noise* berhasil dihilangkan tanpa merusak sinyal utama.

### Penyimpanan Model

Setelah model selesai dilatih dan dioptimasi, model dan objek *preprocessing* (seperti scaler, PCA, dan label encoder) disimpan menggunakan `joblib`. Ini memungkinkan model digunakan kembali tanpa perlu melatih ulang.

### Input Data EEG

Data EEG terlebih dahulu di-*scaling* untuk menormalkan rentang nilai fitur agar memiliki skala yang seragam. Proses ini penting untuk memastikan bahwa setiap fitur memberikan kontribusi yang seimbang dalam analisis dan pemodelan, serta mencegah dominasi fitur dengan nilai yang lebih besar. Setelah data dinormalisasi, dilakukan reduksi dimensi menggunakan *Principal Component Analysis* (PCA).

### Visualisasi Data EEG

Data EEG divisualisasikan dalam tiga bentuk utama untuk memberikan gambaran yang komprehensif tentang sinyal otak. Pertama, grafik sinyal EEG menampilkan nilai setiap fitur EEG dalam bentuk garis yang merepresentasikan perubahan sinyal secara temporal. Kedua, *heatmap* EEG menyajikan nilai fitur dalam bentuk peta warna yang memudahkan identifikasi pola dan intensitas aktivitas di berbagai kanal EEG. Ketiga, *spectrogram* digunakan untuk menampilkan representasi frekuensi sinyal EEG, menggambarkan bagaimana komponen frekuensi berubah seiring waktu, sehingga memberikan informasi penting terkait dinamika frekuensi gelombang otak.

### Prediksi dan Deteksi ADHD

Model *Random Forest* digunakan untuk memprediksi apakah data EEG menunjukkan kondisi ADHD atau tidak dengan cara menghasilkan probabilitas kelas (*prediction\_proba*) dan kelas prediksi (*prediction*). Kelas prediksi tersebut kemudian dikonversi kembali ke label asli, yaitu ADHD atau non-ADHD, menggunakan label encoder. Proses prediksi ini diawali dengan memasukkan fitur-fitur atau atribut data, diikuti dengan *bootstrapping* untuk mengambil sampel-sampel acak dari data latih. Dari setiap sampel *bootstrap*, dibangun sebuah pohon keputusan, dan langkah ini diulang sebanyak *n* kali untuk membuat sejumlah *n* pohon keputusan secara paralel. Setelah semua pohon selesai dibangun, dilakukan *voting* (untuk klasifikasi) atau penghitungan rata-rata (*mean*, untuk regresi) dari hasil prediksi setiap pohon, yang menghasilkan kelas prediksi atau nilai prediksi akhir.

## HASIL DAN PEMBAHASAN

### Hasil Klasifikasi

Evaluasi performa model Random Forest dalam mendeteksi ADHD berdasarkan data EEG dilakukan menggunakan beberapa metrik utama, yaitu precision, recall, f1-score, dan accuracy. Metrik-metrik ini digunakan untuk menilai kemampuan model dalam mengklasifikasikan subjek ke dalam kelas ADHD maupun kontrol secara tepat. Pada Gambar 2. merupakan hasil lengkap evaluasi model klasifikasi pada kelas ADHD dan kontrol.

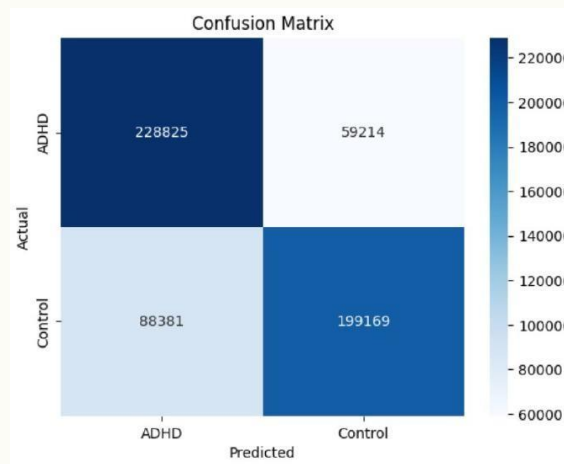
	precision	recall	f1-score	support
ADHD	0.72	0.79	0.76	288039
Control	0.77	0.69	0.73	287550
accuracy			0.74	575589
macro avg	0.75	0.74	0.74	575589
weighted avg	0.75	0.74	0.74	575589

**Gambar 2.** Hasil Evaluasi Sebuah Model Klasifikasi Pada Kelas ADHD dan Kontrol



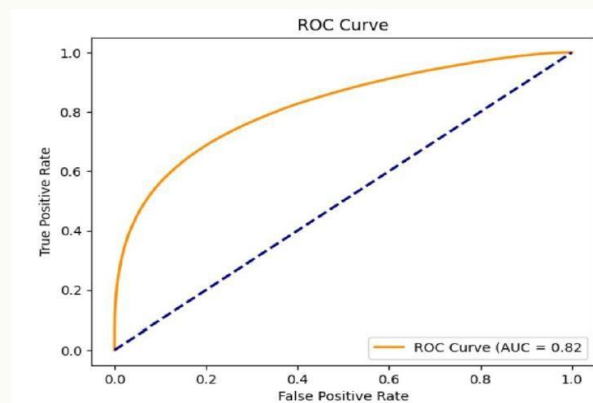
Hasil evaluasi performa model klasifikasi *Random Forest* yang dikembangkan untuk membedakan antara dua kelas, yaitu ADHD dan Control, menunjukkan akurasi sebesar 0,74. Artinya, model mampu mengklasifikasikan data dengan benar sebanyak 74% dari total data yang diuji. Evaluasi ini dilakukan menggunakan data uji yang terdiri dari 30% dari keseluruhan dataset EEG, yang diambil secara acak dari dataset utama setelah melalui tahapan preprocessing seperti pembersihan data, pemisahan fitur dan label, serta normalisasi.

Pada kelas ADHD, *precision* yang dicapai adalah 0.72, *recall* 0.79, dan *f1-score* 0.76, yang menunjukkan bahwa model cukup baik dalam mengenali kasus ADHD, dengan *recall* yang lebih tinggi daripada *precision*. Sementara itu, pada kelas Control, *precision* sebesar 0.77, *recall* 0.69, dan *f1-score* 0.73, menandakan model sedikit lebih baik dalam mengidentifikasi data Control secara tepat, namun masih kurang dalam mendeteksi seluruh kasus Control. Nilai rata-rata makro dan rata-rata berbobot untuk *precision*, *recall*, dan *f1-score* berada pada kisaran 0.74 hingga 0.75, yang menunjukkan performa model yang cukup konsisten di antara kedua kelas, dengan jumlah data yang hampir seimbang, yaitu sekitar 288 ribu untuk ADHD dan 287 ribu untuk Control dari total 575.589 data yang digunakan pada tahap evaluasi.



**Gambar 3.** Confussion Matrix Sebuah Model Klasifikasi Dalam Membedakan Kelas ADHD dan Kontrol

*Confusion matrix* yang digunakan untuk evaluasi ini memperlihatkan bahwa model berhasil mengklasifikasikan 228.825 data ADHD dengan benar (*true positive*) dan 199.169 data Control dengan benar (*true negative*), namun terdapat 59.214 data ADHD yang salah diklasifikasikan sebagai Control (*false negative*) dan 88.381 data Control yang salah diklasifikasikan sebagai ADHD (*false positive*), sehingga secara keseluruhan model menunjukkan tingkat akurasi yang cukup baik meskipun masih ada sejumlah data yang salah klasifikasi, terutama pada kelas Control yang lebih sering keliru diklasifikasikan sebagai ADHD.



**Gambar 4.** Kurva ROC Sebuah Model Klasifikasi Kelas ADHD dan Kontrol

Kurva ROC (*Receiver Operating Characteristic*) yang digunakan untuk mengukur performa model klasifikasi dalam membedakan antara dua kelas. Sumbu horizontal menunjukkan *False Positive Rate* (tingkat kesalahan positif), sedangkan sumbu vertikal menunjukkan *True Positive Rate* (tingkat keberhasilan positif). Kurva oranye pada gambar memperlihatkan performa model, sementara garis putus-putus biru merepresentasikan model acak (*baseline*) dengan AUC (*Area Under Curve*) sebesar 0,5. Pada gambar ini, nilai AUC yang diperoleh adalah 0,82, yang menunjukkan bahwa model memiliki kemampuan yang baik dalam membedakan antara kelas positif dan negatif. Semakin besar nilai AUC, semakin baik performa model; sehingga dengan AUC 0,82, model ini dapat dikatakan cukup andal dan efektif dalam melakukan klasifikasi.

### Hasil Deteksi

Aplikasi ini dirancang untuk membantu pengguna mendeteksi ADHD (*Attention Deficit Hyperactivity Disorder*) berdasarkan data EEG (*electroencephalogram*) menggunakan model *machine learning Random Forest* yang telah dilatih. Aplikasi dibangun dengan **Streamlit**, sebuah *framework* Python yang memungkinkan pembuatan aplikasi web interaktif dengan mudah. Pengguna dapat memasukkan data EEG melalui antarmuka yang disediakan, dan aplikasi akan memproses data tersebut, menampilkan visualisasi, melakukan prediksi, serta menyimpan hasilnya untuk keperluan analisis lebih lanjut. Aplikasi memuat model *Random Forest* dan objek *preprocessing* seperti *scaler*, PCA, dan *label encoder* yang telah disimpan sebelumnya. Jika proses pemuatan gagal, aplikasi akan menampilkan pesan error dan berhenti. Setelah itu, pengguna diminta memasukkan nilai untuk setiap fitur EEG (misalnya, Fp1, Fp2, F3, dll.) melalui *form input*. Nilai-nilai ini disimpan dalam sebuah list dan diupdate secara dinamis menggunakan *session\_state* untuk memastikan data tetap konsisten meskipun pengguna mengubah input. Jika pengguna ingin mengosongkan form, tombol "Reset Form" tersedia untuk mengembalikan semua input ke nilai default (0.0). Sebelum melakukan deteksi, aplikasi memvalidasi input pengguna untuk memastikan semua fitur EEG telah diisi dan tidak ada yang bernilai 0.0. Jika validasi berhasil, data EEG yang dimasukkan akan diproses lebih lanjut. Hal tersebut dapat dilihat pada **Gambar 5**.

**Deteksi ADHD menggunakan Model Hasil Klasifikasi Random Forest**

**Masukkan Data EEG**

Fp1	Fp2	F3	F4	C3
0,0 - +	0,0 - +	0,0 - +	0,0 - +	0,0 - +
C4				
0,0 - +				
F7	P3	F4	O1	O2
0,0 - +	0,0 - +	0,0 - +	0,0 - +	0,0 - +
P8	F8	T7	T8	P7
0,0 - +	0,0 - +	0,0 - +	0,0 - +	0,0 - +
Fz		Cz	Pz	
0,0 - +		0,0 - +	0,0 - +	

Reset Form

Harap isi semua 19 fitur EEG dengan nilai yang valid sebelum melakukan deteksi.

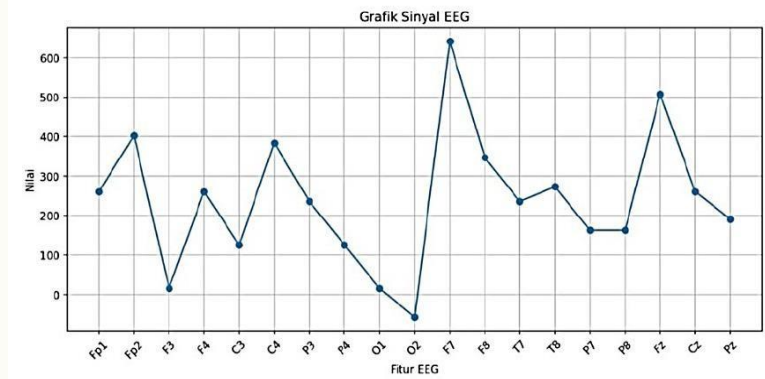
Catatan: Model ini menggunakan PCA dan Random Forest untuk deteksi ADHD.

**Gambar 5.** Aplikasi Deteksi ADHD

Data *di-scaling* menggunakan *StandardScaler* untuk menormalisasi nilai-nilai fitur, kemudian direduksi dimensinya menggunakan PCA (*Principal Component Analysis*) untuk mengurangi kompleksitas data. Setelah preprocessing, data EEG divisualisasikan dalam tiga bentuk: **grafik sinyal**

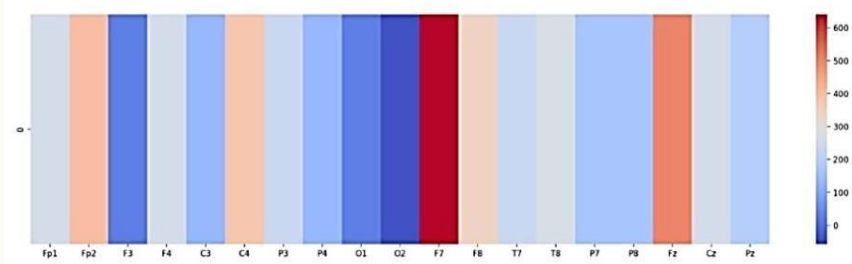
EEG (*Gambar 6*) yang menampilkan nilai setiap fitur dalam bentuk garis, **heatmap** (*Gambar 7*) yang menunjukkan distribusi nilai fitur, dan **spectrogram** (*Gambar 8*) yang merepresentasikan frekuensi sinyal EEG. Visualisasi ini membantu pengguna memahami karakteristik data yang dimasukkan.

### Visualisasi Data EEG



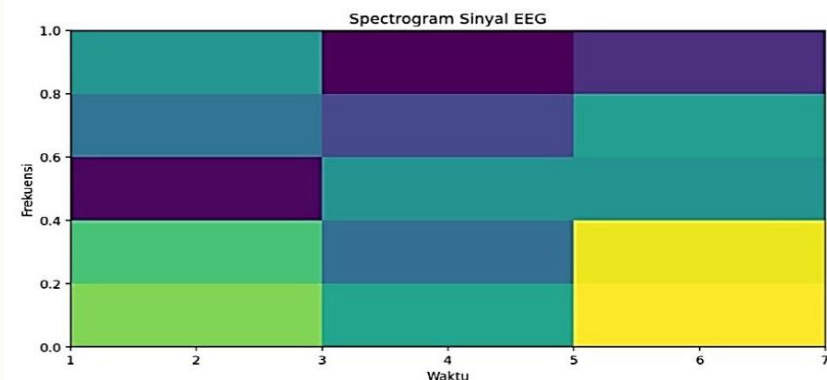
Gambar 6. Grafik Sinyal EEG

### Heatmap Data EEG

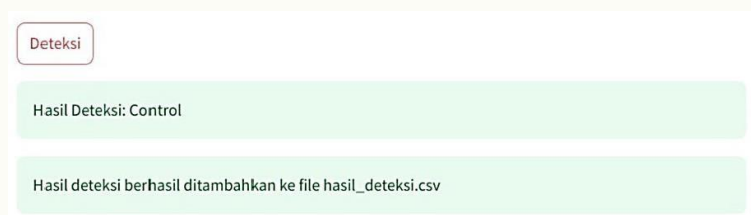


Gambar 7. Heatmap Data EEG

### Spectrogram Data EEG



Gambar 8. Spectrogram Data EEG



Gambar 9. Hasil Deteksi



Gambar 9 menjelaskan bahwa setelah data diproses dan divisualisasikan, pengguna dapat menekan tombol "Deteksi" untuk memprediksi apakah data EEG menunjukkan ADHD atau tidak. Model *Random Forest* akan memproses data yang telah di-PCA dan menghasilkan prediksi beserta *confidence* (tingkat kepercayaan). Hasil prediksi ditampilkan dalam bentuk pesan sukses, misalnya "Hasil Deteksi: Control (Confidence: 95%)". Selain itu, hasil prediksi beserta data EEG yang dimasukkan disimpan ke dalam file CSV (*hasil\_deteksi.csv*) untuk keperluan *logging* atau analisis lebih lanjut. Jika file CSV belum ada, aplikasi akan membuat file baru dan menambahkan header sebelum menyimpan data. Dengan alur ini, aplikasi ini tidak hanya memberikan prediksi ADHD yang akurat tetapi juga menyediakan visualisasi data EEG yang informatif dan kemampuan untuk menyimpan hasil deteksi. Hal ini membuat aplikasi menjadi alat yang berguna bagi profesional medis atau peneliti yang ingin menganalisis data EEG secara interaktif dan efisien.

Penelitian ini berhasil mengembangkan model *Random Forest* yang efektif untuk mendeteksi ADHD melalui analisis data EEG dengan tingkat akurasi yang tinggi. Proses pengolahan data dilakukan secara teliti melalui tahap *preprocessing*, seperti *filtering* frekuensi antara 4 hingga 40 Hz dan penerapan *Independent Component Analysis* (ICA) untuk menghilangkan *noise* dan artefak, sehingga menghasilkan data yang lebih bersih dan relevan. Selanjutnya, fitur diekstraksi berdasarkan amplitudo dari setiap kanal EEG, kemudian label kelas di-encoding agar sesuai dengan kebutuhan model. Data fitur yang diperoleh dinormalisasi menggunakan *StandardScaler* untuk memastikan distribusi yang seragam dan menghindari bias akibat perbedaan skala, serta dilakukan penyeimbangan kelas menggunakan teknik *Random UnderSampling* guna mencegah dominasi kelas mayoritas. Reduksi dimensi dengan *Principal Component Analysis* (PCA) mempertahankan 95% variansi data, sehingga kompleksitas data dapat dikurangi tanpa kehilangan informasi penting. Data kemudian dibagi menjadi subset pelatihan dan pengujian, di mana model *Random Forest* dengan 100 pohon keputusan dilatih menggunakan fitur hasil PCA dan label yang telah diproses. Model ini bekerja dengan membangun pohon keputusan secara paralel dan menentukan prediksi akhir melalui mekanisme voting mayoritas. Optimasi hyperparameter melalui validasi silang dan evaluasi ketat dilakukan untuk mencegah *overfitting* serta memastikan kemampuan generalisasi model yang baik. Performa model dievaluasi menggunakan metrik seperti akurasi, *confusion matrix*, dan kurva ROC untuk mengukur efektivitas klasifikasi ADHD. Selain itu, sistem ini dirancang dengan antarmuka interaktif yang memungkinkan pengguna memasukkan data EEG, melihat visualisasi proses analisis, dan memperoleh prediksi ADHD lengkap dengan tingkat kepercayaan. Hasil prediksi disimpan dalam database untuk mendukung evaluasi lebih lanjut atau penelitian lanjutan. Kombinasi antara model yang kuat dan antarmuka yang ramah pengguna menjadikan sistem ini alat deteksi ADHD yang andal, mudah diakses, dan informatif. Model beserta objek pendukung seperti scaler, PCA, dan label encoder disimpan dalam format *.pkl* untuk penggunaan kembali di masa depan, sehingga penelitian ini dapat menjadi fondasi pengembangan alat diagnostik berbasis AI yang lebih komprehensif dan akurat.

Visualisasi yang dihasilkan pada sistem ini memperkuat keandalan proses analisis dan deteksi. Grafik sinyal EEG pada gambar pertama menampilkan variasi nilai amplitudo di setiap kanal EEG, yang memberikan gambaran karakteristik aktivitas otak secara detail. Selanjutnya, heatmap data EEG memudahkan identifikasi pola dan distribusi nilai amplitudo antar kanal melalui perbedaan intensitas warna, sehingga potensi anomali atau pola khusus dapat dikenali dengan cepat. Spectrogram sinyal EEG pada gambar ketiga memperlihatkan dinamika perubahan frekuensi sinyal otak terhadap waktu, yang sangat penting untuk memahami aktivitas otak dalam domain waktu-frekuensi. Terakhir, hasil deteksi yang ditampilkan pada sistem menunjukkan bahwa data EEG yang dianalisis berhasil diklasifikasikan ke dalam kategori "Control", dan hasil deteksi tersebut secara otomatis disimpan ke dalam file *hasil\_deteksi.csv*. Keseluruhan visualisasi ini tidak hanya membantu dalam memahami karakteristik data EEG yang diolah, tetapi juga memperlihatkan transparansi dan akurasi proses deteksi yang dilakukan oleh sistem, sehingga mendukung keandalan model dalam proses identifikasi kondisi ADHD maupun kontrol.

## KESIMPULAN

Anak berkebutuhan khusus, seperti penderita Attention-deficit Hyperactivity Disorder (ADHD), memerlukan penanganan khusus karena gangguan perkembangan yang menyebabkan hiperaktivitas dan kesulitan belajar. Diagnosis dan deteksi ADHD dapat dibantu dengan penggunaan Electroencephalogram (EEG) yang merekam aktivitas listrik otak, meskipun sinyal EEG yang kecil dan kompleks memerlukan analisis menggunakan algoritma machine learning. Penelitian ini menonjol dengan penggunaan dataset EEG yang sangat besar dan komprehensif serta penerapan algoritma Random Forest untuk tidak hanya mendeteksi keberadaan ADHD, tetapi juga mengklasifikasikan tingkat keparahan gejala. Pendekatan ini diharapkan dapat meningkatkan akurasi dan keandalan deteksi ADHD secara praktis dan reproducible, sehingga memberikan kontribusi signifikan dalam penanganan anak berkebutuhan khusus dengan ADHD. Penelitian ini berhasil mengimplementasikan model klasifikasi menggunakan algoritma *Random Forest* untuk mendeteksi ADHD berdasarkan data EEG. Model ini mencapai akurasi sebesar 0.74 dalam membedakan antara kelas ADHD dan Control. Aplikasi deteksi ADHD yang dibangun dengan *Streamlit* memungkinkan pengguna untuk memasukkan data EEG dan mendapatkan hasil prediksi secara interaktif, dilengkapi dengan visualisasi data EEG dalam bentuk grafik sinyal, *heatmap*, dan *spectrogram*. Dibandingkan penelitian sebelumnya, penelitian ini memiliki kelebihan dalam hal ukuran dataset yang lebih besar dan komprehensif, mencapai 1.048.575 data, serta kemampuan aplikasi untuk menyajikan visualisasi data EEG yang beragam. Selain itu, penelitian ini menonjolkan aspek praktis dengan implementasi *Streamlit* untuk deteksi interaktif.

## DAFTAR PUSTAKA

- Chauhan, N., & Choi, B. J. (2023). Regional Contribution in Electrophysiological-Based Classifications. *MDPI*, 11(9), 180. <https://doi.org/10.3390/computation11090180>.
- Devianto, Y., Sedyono, E., Prasetyo, S. Y. J., & Manongga, D. (2024). Membaca Sinyal Electroencephalogram (EEG) Dalam Menangkap Tingkat Emosi (Berdasarkan Ontologi). *Faktor Exacta*, 17(2), 152-160. <https://doi.org/10.30998/faktorexacta.v17i2.20878>
- Fasya, A. N., & Sari, R. G. (2024). Peran Elektroensefalografi (EEG) Dalam Diagnostik Epilepsi : Tinjauan Efektivitas Terhadap Deteksi Awal Dan Pemantauan. *PREPOTIF*, 8(3), 7882-7891. <https://doi.org/10.31004/prepotif.v8i3.38032>
- Fitriyani, F., Oktaviani, A. M., & Supena, A. (2023). Analisis Kemampuan Kognitif dan Perilaku Sosial pada Anak ADHD (Attention-Deficit Hyperactivity Disorder). *Jurnal Basicedu*, 7(1), 250-259. <https://doi.org/10.31004/basicedu.v7i1.4331>
- Iskandar, R. F. N., Gutama, D. H., Wijaya, D. P., & Danianti, D. (2024). Klasifikasi Menggunakan Metode Random Forest untuk Awal Deteksi Diabetes Melitus Tipe 2. *Jurnal Teknik Industri Terintegrasi*, 7(3), 1620-1626. <https://doi.org/10.31004/jutin.v7i3.26916>
- Kaur, A., & Kahlon, K. S. (2022). Accurate Identification of ADHD among Adults Using Real-Time Activity Data. *Brain Sciences*, 12(7), 831. <https://doi.org/10.3390/brainsci12070831>
- Kim, J. W., Kim, B. N., Kim, J. I., Yang, C. M., & Kwon, J. (2025). Electroencephalogram (EEG) Based Prediction of Attention Deficit Hyperactivity Disorder (ADHD) Using Machine Learning. *Neuropsychiatric Disease and Treatment*, 21, 271-279. <https://doi.org/10.2147/NDT.S509094>

- Loh, H. W., Ooi, C. P., Oh, S. L., Barua, P. D., Tan, Y. R., Acharya, U. R., & Fung, D. S. S. (2024). ADHD/CD-NET: automated EEG-based characterization of ADHD and CD using explainable deep neural network technique. *Cognitive Neurodynamics*, 18(4), 1609–1625. <https://doi.org/10.1007/s11571-023-10028-2>
- Maghfirah, R., Apriliyani, Y., & khaeriyah Syukri, M. (2024). Penanganan Anak Attention Deficit Hyperactivity Disorder (Adhd) Di Kelas Therapy Slb Tncc Banda Aceh. *Anakta : Jurnal Pendidikan Islam Anak Usia Dini*, 3(1), 41-46. <https://doi.org/10.35905/anakta.v3i1.9369>
- Mairita, P., Afandi, A., & Wahyuni, E. S. (2024). Tren topik Attention Deficit Hyperactivity Disorder (ADHD): Analisis Bibliometrik menggunakan Biblioshiny. *DWIJA CENDEKIA: Jurnal Riset Pedagogik*, 8(2), 322-338. <https://doi.org/10.20961/jdc.v8i2.90153>
- Maniruzzaman, M., Shin, J., Hasan, M. A. M., & Yasumura, A. (2022). Efficient Feature Selection and Machine Learning Based ADHD Detection Using EEG Signal. *Computers, Materials and Continua*, 72(3), 5179–5195. <https://doi.org/10.32604/cmc.2022.028339>
- Mulyati, R. (2023). Upaya Guru Dalam Menangani Anak Attention Defisit Hiperactivity Disorder. *Pernik : Jurnal PAUD*, 6(1), 60-69.
- Nurfadhillah, S., Nurlaili, D. A., Syapitri, G. H., Shansabilah, L., Herni, N., & Dewi, H. (2021). Attention Deficit Hyperactive Disorder (Adhd) Pada Siswa Kelas 3 Di Sd Negeri Larangan 1. *PENSA : Jurnal Pendidikan dan Ilmu Sosial*, 3(3), 453-462.
- Nurohanisah, S., Astuti, R., & Basysyar, F. M. (2024). Deteksi Berita Palsu Menggunakan Algoritma Random Forest. *Jurnal Mahasiswa Teknik Informatika*, 8(1), 422-428. <https://doi.org/10.36040/jati.v8i1.8418>
- Pagiling, L., Dwi Nugroho H, Y., Galvanir Noor, M., & Galugu, I. (2020). Klasifikasi Sinyal EEG Menggunakan Algoritma Random Forest dan SVM pada Area Motor Cortex. *ELEKTRODA*, 5(3).
- Prasaja, P., Harumi, L., Fatmawati, R. (2022). Gambaran Demografi Anak Attention Deficit Hyperactivity Disorder (ADHD) di Yayasan Pembinaan Anak Cacat (YPAC) Surakarta. *Jurnal PROFESI (Profesional Islam) : Media Publikasi Penelitian*, 19(2), 152-157. <https://doi.org/10.26576/profesi.v19iNo.2.105>
- Sadida, Q., Tunliu, S. K., Fatimah, F., Kartikasari, N., Asmaradhani, D. T. (2024). Studi Literatur Mengenai Anak Usia Sekolah Dasar Dengan Adhd: Perspektif Neuropsikologi. *ELSE (Elementary School Education Journal)*, 8(1), 135-146. <https://doi.org/10.30651/else.v8i1.20868>
- Tor, H. T., Ooi, C. P., Lim-Ashworth, N. S., Wei, J. K. E., Jahmunah, V., Oh, S. L., Acharya, U. R., & Fung, D. S. S. (2021). Automated detection of conduct disorder and attention deficit hyperactivity disorder using decomposition and nonlinear techniques with EEG signals. *Computer Methods and Programs in Biomedicine*, 200. <https://doi.org/10.1016/j.cmpb.2021.105941>