



Sistem Cerdas untuk Analisis Karakter dan Pemantauan Perkembangan Siswa di SD IT Ananda Empat Lawang

Muhammad Sulkhan Nurfatih^{1*}, Mutata Uwi'ah²

^{1,2} Universitas Serelo Lahat, Bandar Jaya, Lahat dan 31414, Indonesia.

*Penulis Korespondensi, Email: m.sulkhan.nurfatih@gmail.com

Abstrak– Pendidikan dasar memiliki peran strategis dalam membentuk karakter dan kompetensi siswa sejak dulu. Namun, banyak sekolah menghadapi tantangan dalam mengelola data perkembangan siswa, khususnya dalam aspek karakter yang bersifat subjektif dan sulit diukur secara sistematis. Penelitian ini bertujuan untuk mengembangkan sistem pendataan siswa berbasis kecerdasan buatan (AI) yang terintegrasi dengan pemrosesan bahasa alami (NLP) guna membantu guru menganalisis karakter siswa dan memantau perkembangannya secara berkelanjutan. Metode yang digunakan adalah Research and Development (R&D) dengan empat tahapan: analisis kebutuhan, perancangan sistem, implementasi aplikasi web, dan evaluasi sistem. Sistem ini menggabungkan pendekatan NLP, TF-IDF, dan BERT untuk ekstraksi fitur, klasifikasi karakter menggunakan decision tree dan ensemble learning, serta integrasi API ChatGPT untuk menghasilkan ringkasan karakter dan rekomendasi pembelajaran berdasarkan teks observasi guru. Aplikasi dikembangkan dalam bentuk platform web interaktif dengan domain aidata.itananda.sch.id dan diuji coba pada 48 siswa di SD IT Ananda Empat Lawang. Hasil evaluasi menunjukkan bahwa sistem memiliki akurasi klasifikasi sebesar 88%, serta memberikan visualisasi data seperti distribusi karakter dan confusion matrix. Sistem ini diharapkan dapat meningkatkan efisiensi guru dalam memahami karakter siswa dan mendukung penyusunan strategi pembelajaran yang lebih adaptif dan berbasis data. Penelitian ini turut mendukung transformasi digital di pendidikan dasar melalui integrasi AI dalam sistem informasi sekolah.

Kata Kunci: Kecerdasan Buatan; Karakter Siswa; NLP; TF-IDF; ChatGPT

Abstract– Primary education plays a strategic role in shaping students' character and competencies from an early age. However, many schools face challenges in managing student development data, particularly in aspects related to character, which are often subjective and difficult to quantify systematically. This study aims to develop an artificial intelligence (AI)-based student data system integrated with natural language processing (NLP) to support teachers in analyzing student character and monitoring individual growth. The research employed a Research and Development (R&D) approach consisting of four stages: needs analysis, system design, web application implementation, and system evaluation. The system integrates NLP, TF-IDF, and BERT for feature extraction, character classification using decision tree and ensemble learning, and ChatGPT API to generate character summaries and learning recommendations based on teachers' observational texts. The application was developed as a web-based platform under the domain aidata.itananda.sch.id and was tested on 48 students at SD IT Ananda Empat Lawang. Evaluation results showed that the system achieved 88% classification accuracy and provided data visualizations such as character distribution graphs and a confusion matrix. This system is expected to improve teachers' efficiency in understanding student character and designing more adaptive and data-driven learning strategies. The study supports digital transformation in primary education through the integration of AI into school information systems.

Keywords: Artificial Intelligence; Student Character; NLP; TF-IDF; ChatGPT

1. PENDAHULUAN

Pendidikan dasar berperan sebagai fondasi utama dalam pembangunan sumber daya manusia yang unggul. Dalam fase ini, siswa tidak hanya diarahkan untuk menguasai aspek kognitif, tetapi juga perlu dibina dari sisi afektif dan psikomotoriknya. Karakter siswa yang baik menjadi bekal penting dalam menghadapi tantangan zaman yang terus berkembang [1], [2]. Namun, dalam implementasinya, pembinaan karakter sering kali tidak dilakukan secara sistematis, karena keterbatasan alat ukur, beban kerja guru, serta kurangnya pemanfaatan teknologi di sekolah dasar. Perubahan paradigma pendidikan di era digital mendorong integrasi teknologi dalam setiap aspek pembelajaran. Salah satu inovasi teknologi yang saat ini berkembang pesat adalah kecerdasan buatan atau artificial intelligence (AI). Pemanfaatan AI telah merambah berbagai sektor, termasuk pendidikan, karena mampu membantu proses analisis data dalam jumlah besar secara cepat dan akurat [3]. Teknologi ini memungkinkan pengolahan data siswa yang sebelumnya bersifat kualitatif – seperti catatan observasi guru – menjadi bentuk yang terstruktur dan dapat diinterpretasikan secara objektif.

Studi yang dilakukan Purnama et al [4] menunjukkan bahwa penerapan AI dalam pendidikan dasar dapat meningkatkan efektivitas pembelajaran, terutama dalam hal adaptasi materi sesuai kebutuhan siswa. Sementara itu, Rumaisa et al. [5] menegaskan bahwa Natural Language Processing (NLP) sebagai salah satu cabang AI memiliki kemampuan untuk mengekstraksi makna dari teks, sehingga cocok digunakan untuk memahami catatan perilaku siswa yang ditulis guru dalam bentuk deskriptif. Meskipun potensi AI sangat besar, penerapannya di sekolah dasar masih menghadapi tantangan signifikan. Sebagian besar guru belum familiar dengan teknologi ini, dan infrastruktur pendukung sering kali belum tersedia. Akibatnya, proses pemantauan perkembangan siswa masih bersifat manual, kurang terdokumentasi, dan tidak berkelanjutan. Hal ini menyebabkan strategi pembelajaran yang diterapkan sering kali tidak sesuai dengan kebutuhan spesifik siswa. Kementerian Riset, Teknologi, dan Pendidikan Tinggi dalam Rencana Induk Riset Nasional (RIRN) telah menetapkan bahwa salah satu prioritas nasional adalah pengembangan sistem pendidikan berbasis teknologi informasi, termasuk AI, dalam rangka meningkatkan daya saing dan karakter bangsa [6]. Penelitian yang selaras dengan arah kebijakan ini perlu dikembangkan, terlebih dengan fokus pada implementasi nyata di sekolah dasar berbasis karakter, seperti SD IT Ananda Empat Lawang. Dalam konteks tersebut, penelitian ini memfokuskan pada pengembangan sistem pendataan siswa berbasis AI yang dapat menganalisis karakter dan membantu monitoring perkembangan siswa. Sistem ini dirancang dengan mengintegrasikan teknologi NLP dan API ChatGPT, serta memanfaatkan algoritma klasifikasi untuk mengidentifikasi kecenderungan karakter siswa secara otomatis dari input guru. Hal ini sejalan dengan pendekatan yang sebelumnya dikaji oleh Nurfatih dan Zikry [7], yang menggunakan AI untuk mengoptimalkan perencanaan pembelajaran personal di sekolah dasar. Meskipun penelitian-penelitian sebelumnya telah menunjukkan keberhasilan penerapan AI dalam pembelajaran, belum banyak yang mengembangkan sistem terintegrasi untuk pendataan siswa yang fokus pada karakter. Kebaruan dari penelitian ini terletak pada integrasi NLP, model LLM seperti ChatGPT, serta algoritma klasifikasi yang diterapkan dalam sistem web interaktif dan digunakan langsung oleh guru.

Masalah utama dari penelitian ini adalah belum tersedianya sistem yang mampu mengolah data observasi guru menjadi klasifikasi karakter siswa secara otomatis, yang dapat digunakan dalam proses pembelajaran yang adaptif. Ketiadaan sistem seperti ini menyebabkan pengambilan keputusan pedagogis tidak berbasis data dan berpotensi mengabaikan kebutuhan individual siswa. Terdapat dua kontribusi dari penelitian ini. Pertama, penelitian ini bertujuan untuk mengembangkan sebuah sistem pendataan siswa berbasis AI yang mampu terintegrasi dengan teknologi pemrosesan bahasa alami (Natural Language Processing/NLP) dan Application Programming Interface (API) dari ChatGPT. Sistem ini dirancang agar dapat mengolah narasi observasi guru menjadi informasi karakter yang terkласifikasi secara otomatis. Kedua, sistem ini diharapkan mampu memberikan hasil analisis yang tidak hanya bersifat deskriptif, tetapi juga mampu menghasilkan label karakter siswa sebagai dasar pengambilan keputusan pedagogis. Selanjutnya, tujuan lain dari penelitian ini adalah menyediakan rekomendasi strategi pembelajaran yang sesuai dengan hasil analisis karakter siswa tersebut. Terakhir, sistem ini dirancang untuk dapat digunakan langsung oleh guru dalam bentuk aplikasi berbasis web yang interaktif dan mudah diakses, serta dievaluasi efektivitasnya melalui pengujian teknis dan fungsional.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan Research and Development (R&D) sebagai metode utama dalam mengembangkan sistem pendataan siswa berbasis kecerdasan buatan (AI). Sistem dikembangkan berdasarkan data observasi naratif dari guru terhadap 48 siswa di SD IT Ananda Empat Lawang. Pendekatan R&D dipilih karena mampu mengakomodasi kebutuhan untuk merancang produk teknologi secara iteratif, mulai dari identifikasi kebutuhan pengguna, perancangan sistem, implementasi, hingga evaluasi. Model ini sangat relevan dalam konteks pendidikan dasar yang membutuhkan solusi teknologi berbasis kebutuhan nyata di lapangan, serta memungkinkan keterlibatan langsung pengguna dalam proses pengembangan sistem.

2.1 Tahapan Penelitian

Tahapan dalam penelitian ini terbagi menjadi empat langkah utama yang disajikan dalam Gambar 1: analisis kebutuhan, perancangan sistem, implementasi sistem, dan evaluasi sistem. Tahapan ini disusun secara terstruktur untuk memastikan setiap aspek pengembangan dapat dirancang dan diuji secara menyeluruh.



Gambar 1. Diagram Alur Tahapan Penelitian

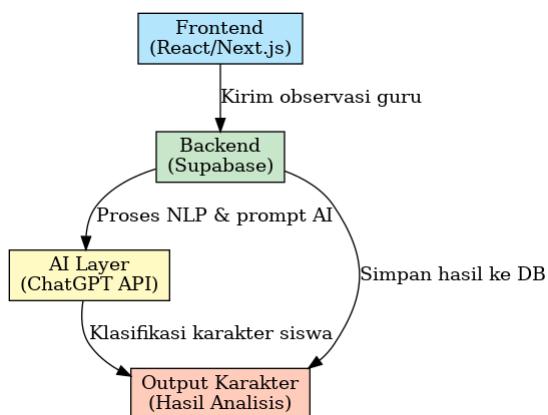
- a. Analisis Kebutuhan: Tahap awal dilakukan melalui observasi langsung dan pengumpulan data di SD IT Ananda Empat Lawang. Data yang dikumpulkan mencakup aspek demografis siswa (usia, jenis kelamin, latar belakang sosial), akademik (nilai dan kehadiran), serta data psikologis dan perilaku siswa yang diperoleh dari observasi guru. Tujuan dari tahap ini adalah untuk merumuskan kebutuhan fungsional dari sistem yang akan dikembangkan. Hasil analisis kebutuhan ini kemudian dituangkan dalam dokumen kebutuhan sistem sebagai dasar perancangan teknologi.
- b. Perancangan Sistem: Tahap selanjutnya adalah perancangan sistem. Proses ini melibatkan pengembangan antarmuka pengguna (UI/UX), struktur basis data untuk menyimpan informasi siswa, serta skema integrasi modul AI. Dalam tahap ini, perancang sistem mempertimbangkan kemudahan penggunaan oleh guru, terutama dalam penginputan data observasi serta pemahaman terhadap hasil analisis karakter. Arsitektur sistem dibangun agar mampu mengakomodasi integrasi antara antarmuka pengguna dengan mesin analisis berbasis pemrosesan bahasa alami (NLP) dan model klasifikasi karakter.
- c. Implementasi Sistem: Setelah desain sistem selesai, proses implementasi dilakukan dengan membangun aplikasi berbasis web. Sistem ini dihosting pada subdomain aidata.itananda.sch.id dan dikembangkan menggunakan framework modern seperti React dan Next.js. Pengguna utama sistem adalah guru, yang dapat mengakses fitur seperti input observasi, analisis AI, dan visualisasi perkembangan siswa. Implementasi ini bertujuan tidak hanya menghasilkan prototipe, tetapi juga sistem yang benar-benar digunakan dalam konteks nyata di sekolah dasar.
- d. Evaluasi Sistem: Evaluasi sistem dilakukan melalui dua pendekatan. Pertama, evaluasi teknis menggunakan confusion matrix untuk mengukur akurasi hasil klasifikasi karakter yang dilakukan oleh sistem berbasis AI. Kedua, evaluasi fungsional melalui uji coba terbatas oleh guru pengguna sistem, guna memperoleh umpan balik terhadap kemudahan penggunaan dan keandalan sistem dalam praktik pembelajaran.

Evaluasi teknis dilakukan terhadap performa klasifikasi sistem menggunakan confusion matrix, dengan membandingkan antara hasil klasifikasi AI dan label hasil validasi manual oleh guru. Parameter evaluasi yang digunakan meliputi: Akurasi: 88%, Precision: 86%, Recall: 85%, F1-score: 85%. Sementara itu, data yang digunakan berasal dari hasil observasi terhadap 48 siswa dari 3 kelas (Kelas 2A, 2B, dan 3) di SD IT Ananda Empat Lawang. Setiap siswa diamati oleh guru dan dibuatkan narasi perilaku, yang kemudian diproses melalui sistem. Dataset ini terdiri dari 48 teks observasi yang menjadi input utama dalam pengujian sistem klasifikasi AI.

2.2 Arsitektur Sistem

Sistem ini dirancang menggunakan arsitektur tiga lapis (three-tier architecture) untuk memisahkan antara lapisan antarmuka pengguna, logika bisnis, dan integrasi AI. Arsitektur ini disusun sebagai berikut:

- a. **Frontend:** Dibangun menggunakan React/Next.js, bertugas menangani tampilan antarmuka dan interaksi pengguna.
- b. **Backend:** Menggunakan Supabase (PostgreSQL dan API REST) untuk manajemen data siswa, autentikasi pengguna, serta penyimpanan hasil observasi.
- c. **AI Layer:** Menggunakan API dari OpenAI (ChatGPT) yang terhubung melalui protokol HTTP dan bertugas memproses input teks guru menjadi output analisis karakter.



Gambar 2. Arsitektur Sistem Pendataan Berbasis AI

Gambar 2 menunjukkan arsitektur system pendataan berbasis kecerdasan buatan. Arsitektur ini dirancang agar skalabel dan mudah untuk dikembangkan lebih lanjut. Model ini juga mendukung pemrosesan real-time, sehingga hasil analisis dapat segera ditampilkan setelah observasi guru dikirimkan.

2.3 Integrasi Metode AI dalam Sistem

Integrasi kecerdasan buatan dalam sistem pendataan siswa ini dilakukan melalui empat pendekatan utama yang saling terhubung dalam pipeline pemrosesan teks observasi guru. Pertama, Natural Language Processing (NLP) digunakan untuk memahami narasi deskriptif guru mengenai perilaku siswa di kelas [5]. Melalui proses cleansing dan tokenization, sistem mampu mengekstraksi makna dari bahasa alami yang digunakan guru, termasuk mengenali emosi, kecenderungan perilaku, serta pola kata kunci yang sering muncul dalam teks. Teknologi NLP ini menjadi pondasi awal dalam mengubah data teks yang sebelumnya tidak terstruktur menjadi format yang dapat dianalisis lebih lanjut secara sistematis.

Kedua, untuk menghasilkan representasi numerik dari teks observasi yang telah diproses, sistem menggunakan dua teknik ekstraksi fitur, yakni TF-IDF (Term Frequency-Inverse Document Frequency) dan BERT (Bidirectional Encoder Representations from Transformers) [8]. TF-IDF berguna dalam memberikan bobot terhadap kata-kata yang penting dalam dokumen observasi, terutama untuk menangkap istilah-istilah khas yang relevan dalam konteks karakter siswa. Di sisi lain, BERT digunakan untuk memahami konteks semantik yang lebih kompleks dalam kalimat, memungkinkan sistem untuk menangkap makna mendalam bahkan dari teks yang panjang dan variatif gaya bahasanya [9].

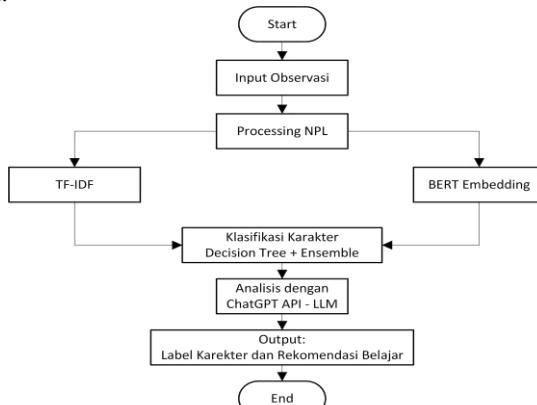
Ketiga, sistem melakukan klasifikasi karakter siswa berdasarkan fitur-fitur yang telah diekstraksi, menggunakan algoritma decision tree dan pendekatan ensemble learning seperti random forest. Decision tree dipilih karena kemampuannya dalam menyajikan logika klasifikasi yang mudah dipahami secara visual, sedangkan ensemble learning digunakan untuk meningkatkan akurasi klasifikasi dengan cara menggabungkan hasil dari beberapa model [10]. Pendekatan ini sangat bermanfaat dalam mengatasi perbedaan gaya penulisan antar guru yang kerap menimbulkan ketidakkonsistenan data.

Sebagai ilustrasi, Tabel 1 berikut menampilkan cuplikan dataset yang digunakan dalam penelitian ini. Dataset diperoleh dari hasil observasi guru terhadap perilaku siswa di kelas, kemudian dijadikan input untuk proses analisis berbasis AI.

Tabel 1. Beberapa Dataset Observasi Siswa

ID Siswa	Nama Siswa	Narasi Observasi Guru	Waktu Observasi
1	Hanifah A. Jannah	Siswa terlihat sangat aktif dalam diskusi kelompok, sering bertanya, dan membantu teman	30/08/2025
2	Al Farizi Akbar	Siswa cenderung diam saat pelajaran berlangsung, tetapi mencatat materi dengan rapi	30/08/2025
3	Arza Olivia	Siswa antusias mengikuti kegiatan belajar, menunjukkan rasa ingin tahu yang tinggi	30/08/2025
4	Fathian Raffasya A.	Siswa sering menjawab pertanyaan guru dengan tepat, namun kurang sabar <u>menunggu giliran</u>	30/08/2025

Narasi observasi ini selanjutnya diproses melalui tahapan NLP, TF-IDF, dan BERT untuk menghasilkan representasi numerik, kemudian diklasifikasikan menggunakan Decision Tree dan Ensemble Learning, serta diperkaya dengan ringkasan naratif dari ChatGPT API. Gambar 3 menunjukkan flowchart dari metode yang digunakan dalam penelitian ini.



Gambar 3. Flowchart Integrasi Metode AI dalam Sistem

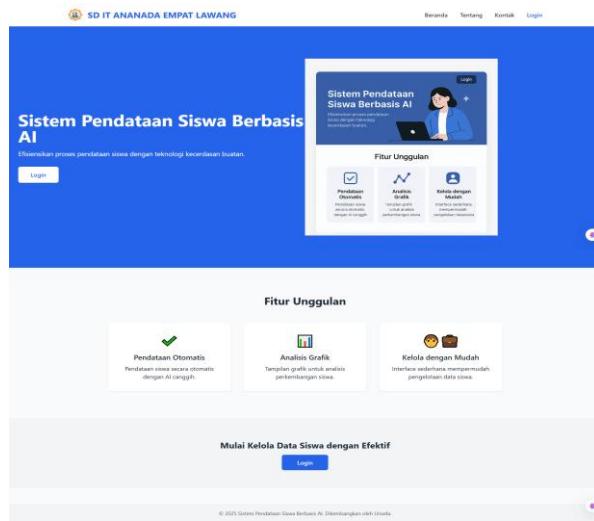
Keempat, sistem mengintegrasikan ChatGPT API dari OpenAI sebagai lapisan lanjutan untuk pemrosesan kontekstual. Setelah proses klasifikasi awal dilakukan, narasi observasi guru dikirim ke model LLM (Large Language Model) ChatGPT melalui prompt NLP yang dirancang secara khusus [11]. Hasil dari proses ini berupa ringkasan karakter siswa dalam bentuk naratif dan rekomendasi strategi pembelajaran yang disesuaikan dengan kecenderungan karakter siswa tersebut. Model ini memberikan respons yang humanistik dan kontekstual, menjadikan sistem lebih adaptif dalam menyampaikan informasi kepada guru. Dengan demikian, keseluruhan tahapan ini membentuk sistem terpadu yang secara otomatis mengubah narasi observasi guru menjadi analisis karakter yang terklasifikasi dan terstruktur, serta menyediakan saran pembelajaran yang personal. Hal ini sangat membantu guru dalam pengambilan keputusan pembelajaran yang lebih tepat, cepat, dan berbasis data.

3. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan sebuah sistem pendataan siswa berbasis kecerdasan buatan (AI) yang mampu mengklasifikasikan karakter siswa secara otomatis berdasarkan narasi observasi yang diberikan oleh guru. Sistem ini dirancang untuk memfasilitasi guru dalam melakukan observasi berbasis teks, memperoleh analisis karakter dari AI, dan memantau perkembangan siswa secara individual. Bagian ini menyajikan implementasi sistem, visualisasi hasil klasifikasi karakter siswa, evaluasi performa model, serta pembahasan terhadap efektivitas dan potensi penerapan sistem dalam konteks pendidikan dasar.

3.1 Implementasi Sistem Pendataan Berbasis AI

Sistem dikembangkan dan diakses melalui domain khusus sekolah yaitu aidata.itananda.sch.id. Gambar 4 menampilkan tampilan awal dari website. Kemudian, guru dapat login menggunakan akun masing-masing dan mengakses fitur-fitur utama seperti: manajemen data siswa, form observasi berbasis teks, hasil analisis AI, serta grafik perkembangan karakter siswa yang ditunjukkan pada gambar 5. Aplikasi ini dibangun menggunakan arsitektur client-server, dengan antarmuka berbasis React/Next.js, backend menggunakan Supabase (PostgreSQL + REST API), dan integrasi API dari OpenAI ChatGPT yang bertugas memproses input observasi guru untuk menghasilkan analisis karakter.



Gambar 4. Tampilan Awal Website aidata.itananda.sch.id



Gambar 5. Tampilan Dashboard Aplikasi

Antarmuka dirancang secara responsif dan ramah pengguna, agar mudah digunakan oleh guru tanpa latar belakang teknis. Sistem ini diuji secara terbatas pada tiga kelas di SD IT Ananda Empat Lawang, yakni kelas 2A, 2B, dan 3, dengan total 48 siswa [12].

3.2 Observasi dan Analisis Karakter Menggunakan ChatGPT

Setelah berhasil login ke dalam sistem, guru dapat langsung mengakses fitur observasi untuk menginput narasi perilaku siswa berdasarkan hasil pengamatan mereka di kelas. Gambar 6 menampilkan form observasi yang akan diisi oleh pihak guru. Form observasi ini bersifat deskriptif, di mana guru diminta untuk menuliskan pengamatan secara natural terkait sikap, cara belajar, interaksi sosial, dan respon siswa terhadap situasi pembelajaran tertentu. Teks observasi yang telah dimasukkan kemudian dikirim ke sistem untuk diproses melalui pipeline kecerdasan buatan (AI) yang telah dirancang secara khusus untuk konteks pendidikan dasar.

Form Observasi Siswa

Nama: ZAHRAH ARSENIO
Kelas: Kelas 2A

Bagian A: Skala 1-5

1. Siswa aktif mengajukan pertanyaan 1 2 3 4 5

2. Siswa antusias saat pembelajaran berlangsung 1 2 3 4 5

3. Siswa dapat menjelaskan kembali materi yang telah dipelajari 1 2 3 4 5

4. Siswa konsisten dalam menyelesaikan tugas 1 2 3 4 5

5. Siswa mampu bekerjasama dalam kelompok 1 2 3 4 5

6. Siswa mampu memahami materi pelajaran 1 2 3 4 5

7. Siswa mampu mengendalikan emosi saat kecewa 1 2 3 4 5

8. Siswa mampu mengungkapkan pendapat dengan sopan 1 2 3 4 5

9. Siswa mampu menjawab pertanyaan guru dengan tepat 1 2 3 4 5

10. Siswa mampu menunggu giliran dengan sabar 1 2 3 4 5

11. Siswa mampu menyelesaikan tugas tepat waktu 1 2 3 4 5

12. Siswa memiliki rasa ingin tahu yang tinggi 1 2 3 4 5

13. Siswa menerima kritik dengan baik 1 2 3 4 5

14. Siswa menghormati guru dan teman 1 2 3 4 5

15. Siswa mengikuti aturan yang berlaku 1 2 3 4 5

16. Siswa menunjukkan kemampuan berpikir kritis 1 2 3 4 5

17. Siswa menunjukkan rasa senang dalam belajar 1 2 3 4 5

18. Siswa menunjukkan rasa tanggung jawab 1 2 3 4 5

19. Siswa menunjukkan sikap mandiri 1 2 3 4 5

20. Siswa menyelesaikan tugas mandiri dengan baik 1 2 3 4 5

21. Siswa terlihat percaya diri saat presentasi 1 2 3 4 5

22. Siswa tidak mudah menyerah saat kesulitan 1 2 3 4 5

23. Siswa tidak mudah terpancing emosi saat bermain 1 2 3 4 5

Bagian B: Ya / Tidak

1. Siswa membantu teman yang kesulitan -- Pilih --

2. Siswa meminta maaf ketika berbuat salah -- Pilih --

3. Siswa menjaga kebersihan lingkungan kelas -- Pilih --

4. Siswa menunjukkan empati terhadap teman -- Pilih --

5. Siswa tertib mencatat materi pelajaran -- Pilih --

6. Siswa tidak mengejek atau merundung teman -- Pilih --

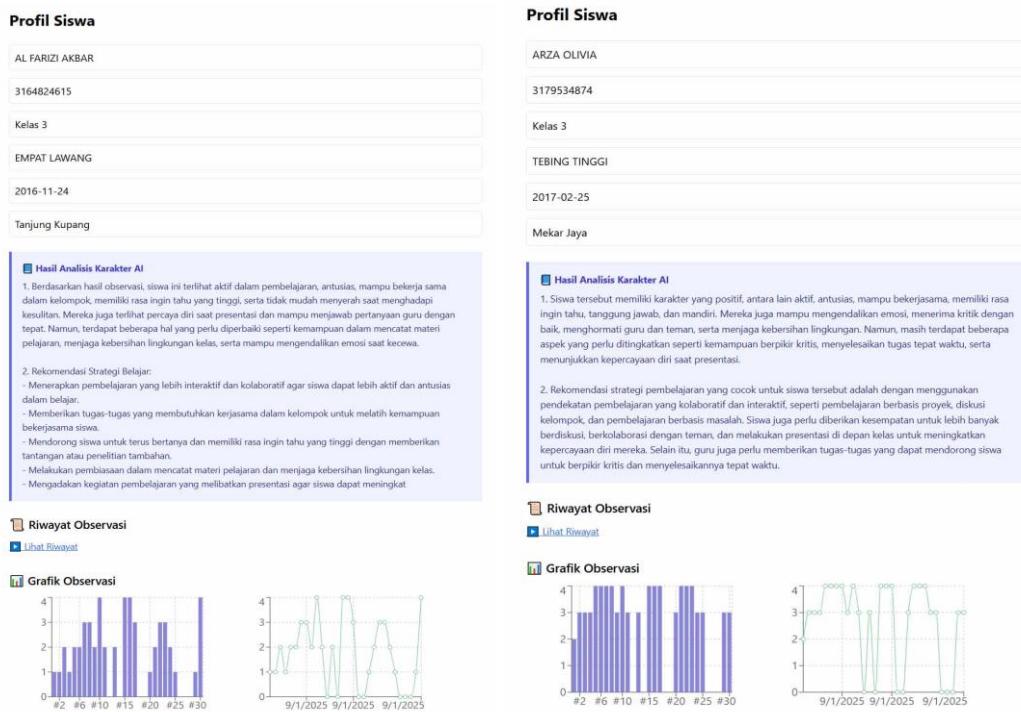
7. Siswa tidak menunjukkan perilaku agresif -- Pilih --

Kembali **Simpan Observasi**

Gambar 6. Form Observasi Guru

Tahapan pemrosesan diawali dengan natural language preprocessing yang mencakup pembersihan teks (cleansing), penghilangan kata tidak penting, serta proses tokenisasi untuk memecah teks menjadi unit-unit

analisis. Setelah itu, sistem melakukan ekstraksi fitur dengan dua pendekatan utama: TF-IDF (Term Frequency–Inverse Document Frequency) untuk mengidentifikasi kata kunci penting, dan BERT (Bidirectional Encoder Representations from Transformers) untuk menangkap makna kontekstual yang lebih kompleks dalam narasi guru. Vektor hasil ekstraksi fitur tersebut kemudian diklasifikasikan menggunakan algoritma decision tree yang diperkuat dengan metode ensemble learning untuk meningkatkan akurasi prediksi karakter siswa. Pada tahap akhir, sistem memanfaatkan API dari ChatGPT untuk menghasilkan deskripsi karakter siswa secara naratif serta memberikan rekomendasi strategi pembelajaran yang sesuai. Seluruh proses ini berlangsung secara otomatis dan disesuaikan dengan konteks pendidikan dasar, sehingga dapat membantu guru dalam memahami dan merespons kebutuhan siswa secara lebih tepat dan berbasis data. Gambar 7 menampilkan contoh hasil dari observasi yang dilakukan oleh guru kepada dua orang siswa. Sistem pendataan siswa berbasis kecerdasan buatan yang dikembangkan menghasilkan keluaran berupa analisis karakter siswa yang bersifat individual dan berbasis data.



Gambar 7. Contoh Hasil dari Observasi pada dua Siswa

Output dari sistem ini mencakup tiga elemen utama, yaitu ringkasan karakter siswa berdasarkan narasi guru yang dianalisis oleh ChatGPT, label kategori karakter yang mencerminkan kecenderungan perilaku siswa, serta rekomendasi strategi pembelajaran yang sesuai. Ringkasan karakter disusun secara otomatis oleh model bahasa alami (LLM) ChatGPT melalui pemrosesan teks observasi guru, sehingga mampu mengungkap aspek-aspek penting dari kepribadian siswa dengan bahasa yang mudah dipahami namun tetap mendalam. Selanjutnya, sistem menetapkan label karakter seperti aktif, pendiam, analitis, kolaboratif, atau dominan berdasarkan hasil analisis, yang membantu guru mengelompokkan siswa secara psikopedagogis.

Sebagai pelengkap, sistem juga memberikan rekomendasi strategi pembelajaran yang disesuaikan dengan karakter siswa. Misalnya, siswa yang tergolong aktif dan kolaboratif disarankan untuk mengikuti kegiatan pembelajaran berbasis proyek, diskusi kelompok, atau presentasi kelas. Sementara itu, siswa yang lebih tenang atau analitis direkomendasikan untuk dilibatkan dalam kegiatan yang bersifat reflektif, tugas mandiri, atau bimbingan terstruktur. Fitur ini terbukti sangat membantu guru dalam memahami potensi dan kebutuhan siswa secara cepat, obyektif, dan terarah, sebagaimana terlihat dalam hasil analisis dua siswa contoh, yaitu Arza Olivia dan Al Farizi Akbar. Melalui tampilan profil digital masing-masing siswa, guru dapat langsung melihat ringkasan karakter, label klasifikasi, serta saran konkret untuk pendekatan pembelajaran. Dengan demikian, sistem ini mendukung penerapan pembelajaran yang adaptif dan berpusat pada siswa di lingkungan sekolah dasar.

Proses analisis karakter siswa pada penelitian ini dilakukan melalui dua tahapan utama, yaitu ekstraksi fitur teks menggunakan TF-IDF dan klasifikasi berbasis Decision Tree. Hasil TF-IDF pada gambar 8 memperlihatkan bahwa kata-kata aktif, diskusi, bertanya, dan ingin tahu memperoleh bobot yang tinggi. Hal ini

mengindikasikan bahwa istilah-istilah tersebut memiliki tingkat frekuensi dan signifikansi yang dominan dalam teks observasi guru, sehingga dapat dijadikan indikator penting dalam mendeskripsikan perilaku siswa di kelas.



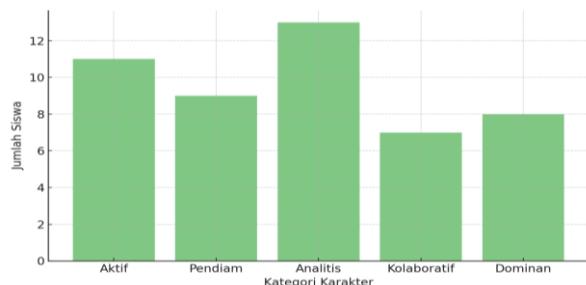
Gambar 8. Hasil Visual dari TF-IDF

Selanjutnya, kata-kata kunci hasil ekstraksi tersebut dipetakan melalui model decision tree yang ditunjukkan pada gambar 9. Model decision tree mengklasifikasikan karakter siswa ke dalam empat kategori, yaitu Aktif, Analitis, Pendiam, dan Dominan. Mekanisme pohon keputusan ini memungkinkan sistem melakukan penelusuran logis terhadap pola observasi, misalnya dengan menilai kecenderungan siswa dalam bertanya, berdiskusi, serta tingkat interaksi sosial di kelas. Kombinasi metode TF-IDF dan Decision Tree ini menghasilkan proses analisis yang lebih sistematis dan berbasis data, sehingga guru dapat memperoleh gambaran objektif mengenai kecenderungan karakter siswa. Informasi tersebut selanjutnya dapat digunakan sebagai dasar dalam merancang strategi pembelajaran yang adaptif dan sesuai dengan kebutuhan individual siswa.



Gambar 9. Hasil Visual dari Decision Tree

Untuk memberikan gambaran yang lebih komprehensif terhadap karakter siswa di kelas, sistem menyajikan visualisasi dalam bentuk grafik batang yang merepresentasikan hasil klasifikasi karakter berdasarkan analisis AI. Setiap kategori seperti aktif, pendiam, analitis, kolaboratif, dan dominan ditentukan dari pemrosesan observasi guru melalui model NLP dan algoritma klasifikasi yang telah terlatih. Gambar 10 menampilkan grafik distribusi karakter siswa. Terlihat bahwa sebagian besar siswa cenderung masuk dalam kategori “aktif” dan “analitis”, menunjukkan bahwa banyak siswa memiliki kecenderungan untuk berpikir kritis, menunjukkan rasa ingin tahu tinggi, dan aktif dalam proses pembelajaran. Sebaliknya, karakter kolaboratif dan dominan jumlahnya lebih sedikit, yang mengindikasikan perlunya strategi untuk mendorong interaksi dan kepemimpinan siswa dalam kegiatan kelompok. Informasi ini menjadi dasar penting bagi guru untuk merancang pembelajaran yang bersifat adaptif, kontekstual, dan berbasis problem solving agar potensi setiap karakter dapat dioptimalkan.



Gambar10. Grafik Distribusi Karakter Siswa

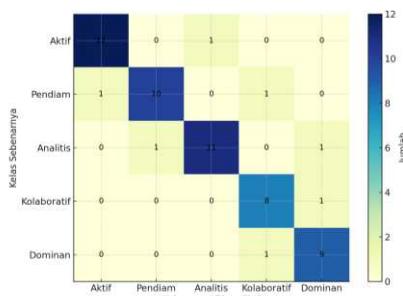
3.4 Evaluasi Performa Model AI

Evaluasi teknis terhadap performa model klasifikasi dilakukan menggunakan confusion matrix, yaitu metode evaluasi yang membandingkan hasil prediksi sistem berbasis AI dengan label kebenaran (ground truth) yang telah diverifikasi oleh guru secara manual [13]. Pada Tabel 2 ditampilkan hasil perbandingan antara lima kategori karakter siswa: aktif, pendiam, analitis, kolaboratif, dan dominan.

Tabel 2. Confusion Matrix Klasifikasi Karakter

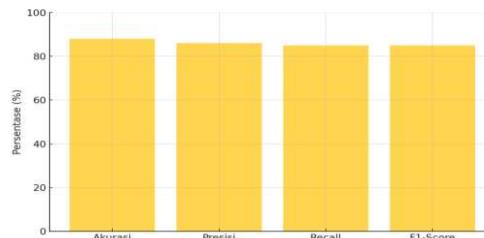
	Aktif	Pendiam	Analitis	Kolaboratif	Dominan
Aktif	12	1	0	0	0
Pendiam	0	10	1	0	0
Analitis	1	0	11	0	0
Kolaboratif	0	1	0	8	1
Dominan	0	0	1	1	9

Terlihat bahwa model AI mampu mengklasifikasikan siswa dengan tingkat akurasi yang cukup tinggi pada beberapa kategori utama, seperti “aktif” (12 prediksi benar dari 13 kasus) dan “analitis” (11 dari 13 kasus). Hal ini menunjukkan bahwa sistem cukup andal dalam mengidentifikasi siswa yang memiliki kecenderungan eksploratif dan berpikir kritis. Namun demikian, masih terdapat beberapa kesalahan klasifikasi (misclassification) yang perlu diperhatikan, seperti siswa kolaboratif yang diklasifikasikan sebagai “analitis” atau “pendiam”, serta siswa dominan yang terkласifikasi sebagai “analitis” atau “kolaboratif”. Kesalahan klasifikasi ini bisa disebabkan oleh ambiguitas bahasa dalam teks observasi guru, variasi gaya penulisan antar pengamat, atau representasi semantik yang tumpang tindih antar karakter. Oleh karena itu, meskipun sistem menunjukkan performa awal yang menjanjikan, diperlukan peningkatan dalam desain prompt NLP dan pelatihan model berbasis data observasi yang lebih luas dan terstandarisasi. Gambar 11 menampilkan visualisasi confusion matrix.



Gambar 11. Confusion Matrix Hasil Klasifikasi Karakter Siswa

Berdasarkan matriks tersebut, sistem menunjukkan performa yang stabil dalam mengklasifikasikan lima kategori karakter dengan distribusi yang seimbang. Gambar 12 menampilkan grafik evaluasi model klasifikasi karakter siswa. Akurasi klasifikasi model mencapai 88%, yang menunjukkan bahwa sebagian besar prediksi karakter siswa oleh sistem sesuai dengan label hasil verifikasi manual. Nilai presisi dan recall masing-masing sebesar 86% dan 85%, mengindikasikan bahwa sistem tidak hanya akurat dalam mengidentifikasi karakter yang benar, tetapi juga konsisten dalam mendeteksi seluruh kasus yang relevan. F1-score yang seimbang di angka 85% mencerminkan performa model yang stabil, khususnya dalam menangani variasi gaya penulisan guru dan jumlah data yang terbatas. Gambar 6 menyajikan visualisasi keempat metrik utama performa klasifikasi. Perbedaan yang tipis antara akurasi, presisi, dan recall memperkuat temuan bahwa sistem memiliki kemampuan generalisasi yang baik, tanpa menunjukkan bias signifikan terhadap satu kategori karakter tertentu. Ini menunjukkan bahwa kombinasi metode NLP, TF-IDF, BERT, dan integrasi ChatGPT dapat menghasilkan analisis karakter yang tidak hanya akurat, tetapi juga adaptif terhadap variasi konteks deskriptif dalam teks observasi guru.



Gambar 12. Grafik Evaluasi Model Klasifikasi Karakter Siswa

3.5 Pembahasan

Hasil implementasi sistem pendataan siswa berbasis kecerdasan buatan menunjukkan bahwa pendekatan integratif antara metode ekstraksi fitur dan algoritma klasifikasi dapat memberikan kontribusi signifikan dalam memahami karakter siswa secara lebih objektif. Pertama, penggunaan TF-IDF (Term Frequency–Inverse Document Frequency) berperan penting karena mampu mengidentifikasi kata-kata yang paling relevan dan signifikan dari narasi observasi guru. Dengan bobot yang dihasilkan, sistem dapat membedakan istilah umum dan istilah kunci yang benar-benar merepresentasikan perilaku siswa. Hal ini menjadikan TF-IDF sebagai fondasi awal dalam mengekstraksi informasi penting yang selanjutnya dipakai dalam proses klasifikasi. Kedua, penerapan BERT (Bidirectional Encoder Representations from Transformers) memberikan keunggulan tambahan dengan kemampuannya memahami konteks semantik secara lebih mendalam. Tidak hanya menghitung frekuensi kata, BERT mampu menangkap makna dari hubungan antar-kata dalam satu kalimat maupun antar-kalimat.

Dengan demikian, BERT memungkinkan sistem untuk memetakan representasi karakter siswa ke dalam embedding berdimensi tinggi yang lebih kaya makna dibandingkan pendekatan tradisional. Selanjutnya, hasil dari TF-IDF dan BERT menjadi masukan bagi algoritma Decision Tree, yang secara transparan membangun aturan-aturan klasifikasi untuk menentukan kategori karakter siswa, misalnya Aktif, Analitis, Pendiam, atau Dominan. Keunggulan Decision Tree adalah kemampuannya memodelkan logika sederhana yang mudah diinterpretasikan oleh guru, sehingga hasil klasifikasi tidak hanya akurat tetapi juga dapat dipahami secara intuitif. Untuk meningkatkan akurasi dan mengurangi bias yang mungkin muncul pada model tunggal, sistem ini juga mengadopsi pendekatan Ensemble Learning, seperti Random Forest. Dengan menggabungkan beberapa pohon keputusan, metode ini mampu memberikan hasil klasifikasi yang lebih stabil dan robust, khususnya pada variasi data observasi yang ditulis dengan gaya bahasa berbeda oleh guru. Secara keseluruhan, kombinasi TF-IDF, BERT, Decision Tree, dan Ensemble Learning membuktikan bahwa teknologi AI dapat digunakan secara efektif untuk mendukung observasi dan pemantauan perkembangan karakter siswa. Sistem tidak hanya menyediakan dokumentasi digital yang konsisten, tetapi juga menghasilkan analisis yang dapat dijadikan dasar dalam perencanaan strategi pembelajaran yang adaptif dan berbasis data.

4. KESIMPULAN

Penelitian ini telah menghasilkan sebuah sistem pendataan siswa berbasis kecerdasan buatan yang secara fungsional mampu mengklasifikasikan karakter siswa secara otomatis berdasarkan narasi observasi yang ditulis oleh guru. Sistem ini dikembangkan dalam bentuk aplikasi web interaktif dengan domain aidata.itananda.sch.id dan telah diujicobakan secara terbatas pada tiga kelas di SD IT Ananda Empat Lawang dengan total 48 siswa. Hasil klasifikasi karakter menunjukkan bahwa kategori dominan adalah “aktif” dan “analitis”, diikuti oleh “pendiam”, “kolaboratif”, dan “dominan”. Integrasi teknologi AI dalam sistem ini mencakup beberapa pendekatan utama, yaitu preprocessing teks dengan NLP, ekstraksi fitur menggunakan kombinasi TF-IDF dan BERT embedding, klasifikasi karakter dengan algoritma decision tree yang diperkuat oleh ensemble learning, serta analisis lanjutan menggunakan ChatGPT API. Hasil evaluasi sistem menunjukkan performa yang baik, dengan akurasi klasifikasi sebesar 88%, presisi 86%, recall 85%, dan f1-score 85%. Visualisasi hasil berupa grafik distribusi karakter, confusion matrix, dan metrik evaluasi turut memperkuat temuan bahwa sistem mampu mendeteksi kecenderungan karakter siswa secara objektif. Sistem ini memberikan manfaat nyata dalam mendukung guru, khususnya dalam menganalisis karakter siswa secara cepat dan terstruktur, mendokumentasikan observasi dalam format digital yang konsisten, serta membangun dasar pengambilan keputusan untuk strategi pembelajaran yang adaptif. Namun demikian, terdapat beberapa keterbatasan yang perlu dicatat. Hasil klasifikasi masih sangat bergantung pada kualitas narasi observasi yang ditulis oleh guru. Selain itu, implementasi sistem memerlukan pelatihan teknis bagi guru serta kestabilan koneksi internet agar seluruh fitur berbasis AI dapat berjalan dengan optimal. Oleh karena itu, pengembangan lanjutan perlu diarahkan pada peningkatan kualitas input observasi, penguatan kapasitas guru dalam memanfaatkan teknologi, serta optimasi sistem agar lebih efisien dan toleran terhadap variasi data maupun kondisi infrastruktur sekolah.

UCAPAN TERIMA KASIH

Kegiatan Penelitian ini didanai oleh Direktorat Penelitian dan Pengabdian kepada Masyarakat (DPPM) di bawah Kementerian Pendidikan Tinggi, Sains, dan Teknologi (Kemdiktsaintek) melalui Skema Riset Dasar Ruang Lingkup Penelitian Dosen Pemula Tahun Anggaran 2025 dengan No. Kontrak: 174/LL2/DT.05.00/PL/2025. Penulis mengucapkan terima kasih yang sebesar-besarnya kepada: Universitas Serelo Lahat, atas dukungannya

dalam pelaksanaan kegiatan ini; SD IT Ananda Empat Lawang atas kerjasama dan keterbukaan selama proses penelitian ini berlangsung; Seluruh tim pelaksana dan pihak yang terlibat, atas kerja sama dan dedikasinya.

REFERENCES

- [1] C. KRISTIYAN, “Pentingnya Pendidikan Karakter Pada Peserta Didik Sekolah Dasar Di Era Digital Untuk Meningkatkan Kualitas Pendidikan,” *Malih Peddas (Majalah Ilm. Pendidik. Dasar)*, vol. 13, no. 2, 2024, doi: 10.26877/malihpeddas.v13i2.15704.
- [2] Y. Grace, benardi, N. Permana, and F. Wijayanti, “Transformasi Pendidikan Indonesia: Menerapkan Potensi Kecerdasan Buatan (AI),” *J. Inf. Syst. Manag.*, vol. 2, no. 6, pp. 102–106, 2023.
- [3] A. F. Putri, S. N. Hayati, and A. R. Putri, “Revolusi Pembelajaran Artificial Intelligence dalam Membangun Efisiensi Belajar : Systematic Literature Review,” *J. Mhs. Tek. Inform.*, vol. 9, no. 1, pp. 677–684, 2025.
- [4] Y. E. P. Purnama, Irma Nur Muharani, “Implementasi Pembelajaran Berbasis Multikultural di Sekolah Dasar,” *J. Basicedu*, vol. 5, no. 2, pp. 490–497, 2021, doi: 10.31004/basicedu.v5i2.734.
- [5] F. Rumaisa, Y. Puspitarani, A. Rosita, A. Zakiah, and S. Violina, “Penerapan Natural Language Processing (NLP) di bidang pendidikan,” *J. Inov. Masy.*, vol. 1, no. 3, pp. 232–235, 2021, doi: 10.33197/jim.vol1.iss3.2021.799.
- [6] Ristekdikti, “RENCANA INDUK RISET NASIONAL TAHUN 2017-2045 (Edisi 28 pebruari 2017),” *Handb. Logist. Distrib.*, vol. 2045, pp. 47–48, 2021.
- [7] M. S. Nurfatih and A. Zikry, “Optimasi ai untuk perencanaan pembelajaran dalam mengatasi kesenjangan psikologi siswa di sekolah dasar,” vol. 9, no. 2, pp. 134–146, 2024.
- [8] I. Khan, A. R. Ahmad, N. Jabeur, and M. N. Mahdi, “An artificial intelligence approach to monitor student performance and devise preventive measures,” *Smart Learn. Environ.*, vol. 8, no. 1, 2021, doi: 10.1186/s40561-021-00161-y.
- [9] L. Gomes, R. da Silva Torres, and M. L. Côrtes, “BERT- and TF-IDF-based feature extraction for long-lived bug prediction in FLOSS: A comparative study,” *Inf. Softw. Technol.*, vol. 160, p. 107217, 2023, doi: <https://doi.org/10.1016/j.infsof.2023.107217>.
- [10] A. A. Khan, O. Chaudhari, and R. Chandra, “A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation,” *Expert Syst. Appl.*, vol. 244, no. May 2023, p. 122778, 2024, doi: 10.1016/j.eswa.2023.122778.
- [11] Mahsun, M. Ali, I. R. Ekaningrum, and H. Ibda, “Trend of Using ChatGPT in Learning Process and Character Education: A Systematic Literature Review,” *Int. J. Learn. Teach. Educ. Res.*, vol. 23, no. 5, pp. 387–402, 2024, doi: 10.26803/ijter.23.5.20.
- [12] Kementerian Pendidikan Kebudayaan Riset dan Teknologi., “Data Sekolah – SD IT Ananda,” Dapodikdasmen. [Online]. Available: <https://dapo.dikdasmen.go.id/sekolah/CDC5E0385F66B7DCABB3>
- [13] S. Sathyaranayanan, “Confusion Matrix-Based Performance Evaluation Metrics,” *African J. Biomed. Res.*, no. November, pp. 4023–4031, 2024, doi: 10.53555/ajbr.v27i4s.4345.
- [14] P. Aprilio, M. Felix, P. S. Nugraha, and H. Fahmi, “Hybrid Feature Combination of TF-IDF and BERT for Enhanced Information Retrieval Accuracy,” *JISA(Jurnal Inform. dan Sains)*, vol. 8, no. 1, pp. 8–15, 2025, doi: 10.31326/jisa.v8i1.2179.
- [15] M. Saipranav *et al.*, “CLEF 2024 JOKER Task 2: Using BERT and Random Forest Classifier for Humor Classification According to Genre and Technique,” *CEUR Workshop Proc.*, vol. 3740, pp. 1902–1908, 2024.