

ALGORITMA RANDOM FOREST UNTUK KLASIFIKASI PENENTUAN LEVEL BELAJAR SISWA BARU DI NEW CONCEPT ENGLISH EDUCATION CENTRE

Atha Zainum Muttaqin^{1*}, Nadya Safitri²

^{1,2}Program Studi Teknik Informatika, Fakultas Informatika dan Desain, Universitas Bina Insani
Jl. Siliwangi No. 6, Sepanjang Jaya, Rawalumbu, Kota Bekasi, Jawa Barat 17114

Keywords: Random Forest, Machine Learning, Classification, Learning Level, Placement Test.

Correspondent Email:
a.zainumm@gmail.com

Abstrak. Proses penentuan level belajar siswa baru di New Concept English Education Centre masih dilakukan secara manual melalui wawancara dan observasi pengajar sehingga berpotensi menimbulkan subjektivitas dalam penempatan kelas. Penelitian ini bertujuan menerapkan algoritma Random Forest untuk mengklasifikasikan level belajar siswa baru serta mengimplementasikannya ke dalam sistem berbasis website sebagai pendukung pengambilan keputusan. Penelitian menggunakan 1.522 data siswa yang terdiri atas variabel umur, tingkat kelas sekolah, nilai *placement test*, dan level belajar siswa. Tahapan penelitian meliputi preprocessing data, encoding data kategorikal, pembagian data training dan testing dengan rasio 80:20, pelatihan model Random Forest, evaluasi menggunakan confusion matrix dan classification report, serta implementasi model ke dalam aplikasi berbasis web. Hasil evaluasi menunjukkan bahwa model Random Forest memperoleh akurasi sebesar 85,57% dalam mengklasifikasikan level belajar siswa. Sistem yang dikembangkan mampu mendukung proses penentuan level belajar secara lebih konsisten, objektif, dan efisien sehingga dapat membantu pengajar dalam menentukan penempatan kelas siswa baru.



Copyright © 2026 The Author(s),
Published by Jurnal Informatika dan
Teknik Elektro Terapan. This is an
open-access article distributed under
the terms of the Creative Commons
Attribution-NonCommercial 4.0
International License (CC BY-NC
4.0).

Abstract. The student learning level placement process at New Concept English Education Centre is still conducted manually through interviews and teacher observations, which may lead to subjectivity in class placement. This study aims to implement the Random Forest algorithm to classify the learning levels of new students and integrate it into a web-based decision support system. The study employed 1,522 student records consisting of age, school grade, placement test scores, and learning level labels. The research stages included data preprocessing, categorical data encoding, an 80:20 train-test split, Random Forest model training, evaluation using a confusion matrix and classification report, and implementation of the trained model into a web-based application. The evaluation results indicate that the Random Forest model achieved an accuracy of 85.57% in classifying student learning levels. The developed system supports a more consistent, objective, and efficient learning level placement process, thereby assisting teachers in determining appropriate class placement for new students.

1. PENDAHULUAN

Perkembangan teknologi informasi dan penerapan *data mining* dalam bidang pendidikan telah membuka peluang baru dalam meningkatkan kualitas proses pembelajaran.

Pemanfaatan *machine learning* memungkinkan institusi pendidikan menganalisis data siswa secara lebih objektif dan sistematis guna mendukung pengambilan keputusan akademik. Salah satu penerapan yang banyak

dikembangkan adalah metode klasifikasi untuk memprediksi atau mengelompokkan siswa berdasarkan karakteristik dan performa akademiknya. Pendekatan ini dinilai lebih efektif dibandingkan penilaian subjektif karena mampu mengolah banyak variabel secara bersamaan dan menghasilkan keputusan berbasis data.

Lembaga kursus sebagai bagian dari pendidikan nonformal memiliki karakteristik peserta didik yang heterogen, mencakup perbedaan usia, latar belakang pendidikan, serta tingkat kemampuan awal. Kondisi tersebut menuntut adanya penentuan level belajar yang tepat guna memastikan efektivitas proses pembelajaran. Penempatan siswa pada level yang sesuai berperan penting dalam meningkatkan pemahaman materi, menjaga motivasi belajar, dan menciptakan suasana kelas yang lebih kondusif. New Concept English Education Centre merupakan salah satu lembaga kursus bahasa Inggris yang menerima siswa dengan latar belakang kemampuan yang beragam. Berdasarkan hasil observasi, proses penentuan level belajar siswa baru masih dilakukan melalui wawancara singkat mengenai usia dan tingkat kelas sekolah. Dalam beberapa kasus, siswa ditempatkan berdasarkan perkiraan pengajar atau pilihan siswa sendiri tanpa melalui *placement test* yang terstruktur. Kondisi tersebut berpotensi menimbulkan ketidaksesuaian level belajar sehingga siswa dengan kemampuan dasar dapat mengalami kesulitan ketika ditempatkan pada level yang terlalu tinggi, sedangkan siswa dengan kemampuan lebih baik dapat merasa kurang tertantang apabila berada pada level yang terlalu rendah. Selain itu, proses penentuan level yang bersifat subjektif dapat menyebabkan ketidakkonsistenan dalam penempatan kelas dan mengurangi efektivitas pembelajaran secara keseluruhan.

Berbagai penelitian sebelumnya menunjukkan bahwa algoritma Random Forest memiliki performa yang baik dalam menyelesaikan permasalahan klasifikasi di bidang pendidikan. Penelitian oleh Gunawan dan Paputungan (2025) mengembangkan model deteksi potensi kelulusan mahasiswa menggunakan Random Forest dengan tingkat akurasi sebesar 89% [1]. Putri dan Rachmatika (2025) menerapkan Random Forest dan teknik SMOTE untuk memprediksi risiko putus

sekolah pada siswa SMK [2]. Selain itu, Eirene dkk. (2025) menggunakan Random Forest dan XGBoost untuk mengklasifikasikan gaya belajar siswa [3], sedangkan Nachouki dkk. (2023) menunjukkan bahwa Random Forest efektif dalam memprediksi performa akademik mahasiswa serta mengidentifikasi variabel yang berpengaruh terhadap hasil belajar [4].

Meskipun berbagai penelitian tersebut telah membuktikan efektivitas Random Forest dalam memprediksi performa akademik, risiko putus sekolah, maupun gaya belajar siswa, penerapan metode ini dalam konteks penentuan level belajar siswa baru pada lembaga kursus bahasa Inggris masih belum banyak dilakukan. Sebagian besar penelitian berfokus pada pendidikan formal seperti sekolah dan perguruan tinggi, sedangkan implementasi pada lembaga pendidikan nonformal dengan karakteristik peserta didik yang lebih beragam masih terbatas. Oleh karena itu, penelitian ini menerapkan algoritma Random Forest untuk mengklasifikasikan level belajar siswa baru di New Concept English Education Centre serta mengimplementasikan model klasifikasi ke dalam sistem berbasis website guna mendukung proses penentuan level belajar yang lebih objektif, konsisten, dan berbasis data.

2. TINJAUAN PUSTAKA

2.1. Machine Learning

Machine Learning merupakan salah satu cabang *Artificial Intelligence (AI)* yang memungkinkan komputer mempelajari pola dari data untuk melakukan prediksi atau pengambilan keputusan tanpa diprogram secara eksplisit. Kemampuan tersebut menjadikan *machine learning* banyak diterapkan dalam berbagai bidang, termasuk pendidikan, karena mampu mengolah data dalam jumlah besar dan menemukan pola yang sulit diidentifikasi secara manual. Menurut Sutaryani dkk. (2024), *machine learning* telah banyak dimanfaatkan dalam berbagai bidang prediksi karena mampu menghasilkan model yang adaptif terhadap karakteristik data yang digunakan [5]. Dalam bidang pendidikan, penerapan *machine learning* berkembang melalui konsep *Educational Data Mining (EDM)* yang memanfaatkan teknik *data mining*, *machine learning*, dan metode statistik untuk memperoleh informasi dari data akademik.

Pratiwi dan Utami (2025) menjelaskan bahwa *Educational Data Mining* berperan dalam membantu institusi pendidikan menganalisis data siswa sehingga dapat mendukung proses pengambilan keputusan secara lebih efektif [6].

Sejalan dengan hal tersebut, Nachouki dkk. (2023) menyatakan bahwa algoritma *machine learning* mampu digunakan untuk memprediksi performa akademik mahasiswa serta mengidentifikasi faktor-faktor yang memengaruhi hasil belajar berdasarkan data historis [4]. Berdasarkan karakteristik tersebut, penelitian ini menerapkan pendekatan *supervised learning* karena data yang digunakan telah memiliki label berupa level belajar siswa. Melalui pendekatan tersebut, model dapat mempelajari hubungan antara variabel umur, tingkat kelas sekolah, nilai *placement test*, dan level belajar sehingga mampu menghasilkan klasifikasi level belajar siswa baru secara lebih objektif.

2.2. Klasifikasi

Klasifikasi merupakan salah satu teknik dalam *machine learning* yang digunakan untuk mengelompokkan data ke dalam kelas atau kategori tertentu berdasarkan atribut yang dimiliki. Dalam bidang pendidikan, teknik klasifikasi banyak dimanfaatkan untuk memprediksi performa akademik, risiko putus sekolah, maupun kategori capaian belajar siswa. Putri dan Rachmatika (2025) menjelaskan bahwa metode klasifikasi seperti *Random Forest*, *Logistic Regression*, *Support Vector Machine*, dan *XGBoost* dapat digunakan untuk mengidentifikasi siswa yang berisiko putus sekolah dengan tingkat akurasi yang tinggi [2]. Selain itu, Nachouki dkk. (2023) menyatakan bahwa pendekatan klasifikasi dalam *Educational Data Mining* dapat digunakan untuk memprediksi capaian akademik mahasiswa berdasarkan berbagai variabel prediktor, seperti nilai sebelumnya, tingkat kehadiran, dan faktor demografis [4].

Secara umum, proses klasifikasi meliputi pengumpulan dan persiapan data, pembagian data menjadi data latih dan data uji, pembangunan model, serta evaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Danil dkk. (2025) menunjukkan bahwa penggunaan metrik evaluasi tersebut mampu memberikan gambaran yang komprehensif terhadap

performa model klasifikasi pada berbagai permasalahan berbasis data [7]. Pada penelitian ini, teknik klasifikasi digunakan untuk menentukan level belajar siswa baru berdasarkan variabel umur, tingkat kelas sekolah, dan nilai *placement test* sehingga dapat mendukung proses penempatan kelas secara lebih objektif dan berbasis data.

2.3. Decision Tree

Decision Tree merupakan salah satu algoritma klasifikasi dalam *machine learning* yang membentuk struktur pohon untuk memetakan proses pengambilan keputusan berdasarkan atribut tertentu. Model ini terdiri atas *root node* sebagai titik awal, *internal node* sebagai titik percabangan berdasarkan kondisi tertentu, dan *leaf node* sebagai hasil akhir klasifikasi. Menurut Christine Eirene dkk. (2025), algoritma *Random Forest* merupakan gabungan dari sejumlah *Decision Tree*, di mana setiap pohon menghasilkan satu prediksi dan hasil akhir ditentukan berdasarkan mekanisme *majority voting* [3]. Oleh karena itu, *Decision Tree* menjadi konsep dasar dalam pembentukan algoritma *Random Forest*.

Penerapan *Decision Tree* telah banyak digunakan untuk menyelesaikan berbagai permasalahan klasifikasi, khususnya pada bidang pendidikan. Qisthiano dkk. (2023) menerapkan algoritma *Decision Tree* untuk memprediksi kelulusan mahasiswa berdasarkan atribut akademik dan karakteristik mahasiswa dengan tingkat akurasi sebesar 87,93% [8]. Meskipun memiliki keunggulan dalam kemudahan interpretasi hasil klasifikasi, *Decision Tree* memiliki kelemahan berupa kecenderungan mengalami *overfitting* apabila pohon yang terbentuk terlalu kompleks. Oleh karena itu, penelitian ini menggunakan algoritma *Random Forest* sebagai pengembangan dari *Decision Tree* untuk menghasilkan model klasifikasi yang lebih stabil dan akurat.

2.4. Algoritma Random Forest

Random Forest merupakan algoritma *machine learning* berbasis *ensemble learning* yang membangun sejumlah *Decision Tree* dan menggabungkan hasil prediksinya untuk meningkatkan akurasi serta mengurangi *overfitting*. Menurut Nachouki dkk. (2023), algoritma ini bekerja dengan melakukan

bootstrap sampling terhadap data pelatihan, kemudian setiap pohon keputusan dibangun menggunakan subset fitur yang dipilih secara acak. Hasil akhir diperoleh melalui mekanisme *majority voting* untuk proses klasifikasi sehingga mampu menangani hubungan non-linear antarvariabel dengan lebih baik dibandingkan *Decision Tree* tunggal [4]. Pratiwi dan Utami (2025) juga menjelaskan bahwa pendekatan *ensemble learning* pada *Random Forest* mampu meningkatkan stabilitas model melalui pembentukan banyak pohon keputusan yang berasal dari sampel data dan fitur yang berbeda [6].

Selain memiliki kemampuan klasifikasi yang baik, *Random Forest* juga dikenal mampu menangani data yang kompleks dan heterogen serta memiliki ketahanan terhadap *overfitting*. Cappelli dkk. (2024) menyatakan bahwa algoritma ini menyediakan mekanisme *feature importance* untuk mengidentifikasi kontribusi setiap variabel terhadap hasil prediksi [9]. Penerapan *Random Forest* dalam bidang pendidikan juga dilakukan oleh Apandi (2025) untuk mengklasifikasikan prestasi siswa berdasarkan data hasil evaluasi berkala dan menunjukkan bahwa algoritma ini mampu menangani data dengan banyak variabel secara efektif [10]. Berdasarkan karakteristik tersebut, penelitian ini menggunakan algoritma *Random Forest* untuk mengklasifikasikan level belajar siswa baru karena mampu menghasilkan model yang lebih stabil, akurat, dan mendukung proses penentuan level belajar secara objektif.

2.5. Evaluasi Model

Evaluasi model klasifikasi dilakukan untuk mengukur kemampuan model dalam menghasilkan prediksi yang sesuai dengan data aktual. Salah satu metode evaluasi yang umum digunakan adalah *Confusion Matrix*, yaitu matriks yang membandingkan hasil prediksi dengan kelas sebenarnya sehingga dapat menggambarkan performa model secara menyeluruh. Menurut Rinandi dkk. (2023), *Confusion Matrix* terdiri atas empat komponen utama, yaitu *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)*, dan *False Negative (FN)* yang menjadi dasar dalam pengukuran performa model klasifikasi [11]. Markoulidakis dan Markoulidakis (2024) juga menjelaskan bahwa penggunaan *Confusion Matrix* memungkinkan evaluasi model dilakukan

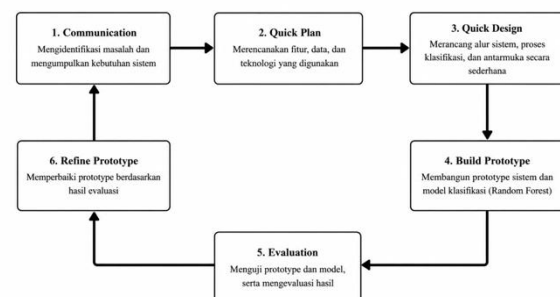
secara lebih komprehensif karena tidak hanya memperhatikan jumlah prediksi yang benar, tetapi juga jenis kesalahan klasifikasi yang dihasilkan [12].

Berdasarkan nilai pada *Confusion Matrix*, performa model dapat diukur menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Pratiwi dan Utami (2025) menjelaskan bahwa *accuracy* menunjukkan proporsi prediksi yang benar terhadap seluruh data, *precision* mengukur ketepatan model dalam memberikan prediksi pada suatu kelas, *recall* menunjukkan kemampuan model dalam mengidentifikasi seluruh data pada kelas yang benar, sedangkan *F1-score* merupakan rata-rata harmonis antara *precision* dan *recall* yang digunakan untuk memberikan gambaran keseimbangan performa model [6]. Pada penelitian ini, evaluasi model *Random Forest* dilakukan menggunakan *Confusion Matrix* dan *Classification Report* untuk mengetahui performa model dalam mengklasifikasikan level belajar siswa baru di New Concept English Education Centre.

3. METODE PENELITIAN

3.1. Alur Penelitian

Penelitian ini bertujuan membangun sistem klasifikasi untuk menentukan level belajar siswa baru menggunakan algoritma *Random Forest* yang diimplementasikan pada aplikasi berbasis web. Tahapan penelitian meliputi pengumpulan dataset, *preprocessing* data, pelatihan model *Random Forest*, evaluasi model menggunakan *confusion matrix* dan *classification report*, serta implementasi model ke dalam sistem berbasis Flask. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian
Sumber: Hasil Penelitian (2026)

3.2. Dataset Penelitian

Penelitian ini menggunakan dataset yang diperoleh dari hasil *placement test* siswa baru di New Concept English Education Centre Cabang Taman Alamanda. Dataset terdiri atas 1.522 data siswa dengan atribut umur, tingkat kelas sekolah, dan nilai *placement test* sebagai variabel input, sedangkan level belajar digunakan sebagai target klasifikasi. Ketiga variabel tersebut dipilih berdasarkan hasil observasi dan wawancara dengan pengajar sebagai faktor utama dalam proses penentuan level belajar siswa.

3.3. Pengolahan Data dan Penerapan Algoritma Random Forest

Dataset yang telah dikumpulkan selanjutnya melalui tahap *preprocessing* untuk memastikan data siap digunakan dalam proses klasifikasi. Tahap ini meliputi pembersihan data, pengecekan data yang tidak lengkap, serta transformasi data ke dalam format yang sesuai. Data kategorikal, seperti tingkat kelas sekolah dan level belajar siswa, diubah menjadi data numerik menggunakan teknik *encoding* agar dapat diproses oleh algoritma *Random Forest*. Selanjutnya, variabel umur, tingkat kelas sekolah, dan nilai *placement test* digunakan sebagai variabel input (X), sedangkan level belajar digunakan sebagai variabel output (y). Dataset kemudian dibagi menjadi data *training* sebesar 80% dan data *testing* sebesar 20% menggunakan metode *train-test split* sebagai persiapan pembangunan model klasifikasi.

3.4. Implementasi Sistem dan Teknik Pengujian

Pengembangan aplikasi dilakukan menggunakan model pengembangan Prototype karena memungkinkan proses pengembangan sistem dilakukan secara bertahap melalui tahapan komunikasi kebutuhan, perancangan prototipe, evaluasi, dan penyempurnaan sistem hingga sesuai dengan kebutuhan pengguna. Pendekatan ini dinilai efektif dalam pengembangan sistem berbasis web karena memudahkan proses validasi kebutuhan sistem sebelum implementasi akhir [13]. Selanjutnya, model Random Forest yang telah dibangun diintegrasikan ke dalam aplikasi berbasis web menggunakan bahasa pemrograman Python dan framework Flask sehingga sistem mampu

memberikan rekomendasi level belajar siswa berdasarkan data masukan pengguna.

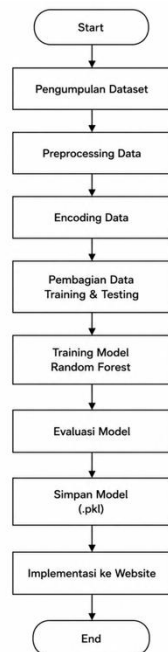
Pengujian pada penelitian ini dilakukan untuk mengevaluasi performa model klasifikasi serta memastikan sistem yang dikembangkan berfungsi sesuai dengan kebutuhan. Evaluasi model menggunakan Confusion Matrix dan Classification Report untuk memperoleh nilai *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, pengujian fungsional dilakukan menggunakan metode Black Box Testing, sedangkan pengujian pengguna dilakukan menggunakan User Acceptance Test (UAT) untuk mengetahui tingkat penerimaan pengguna terhadap sistem yang dikembangkan.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Pengembangan Model

Model klasifikasi dikembangkan menggunakan algoritma Random Forest berdasarkan data hasil *placement test* siswa baru di New Concept English Education Centre. Dataset terdiri atas 1.522 data siswa dengan atribut umur, tingkat kelas sekolah, nilai *placement test*, dan label level belajar. Variabel umur, kelas sekolah, dan nilai *placement test* digunakan sebagai fitur masukan, sedangkan level belajar digunakan sebagai target klasifikasi.

Sebelum proses pelatihan dilakukan, dataset melalui tahap *preprocessing* berupa pemeriksaan data, transformasi atribut kategorikal menggunakan teknik *encoding*, kemudian dibagi menjadi data pelatihan dan data pengujian dengan rasio 80:20. Model Random Forest selanjutnya dilatih menggunakan pustaka Scikit-learn dan disimpan dalam format *.pkl* agar dapat diintegrasikan ke dalam sistem berbasis web. Proses pembangunan model *Random Forest* ditunjukkan pada Gambar 2.



Gambar 2. Tahapan Pembangunan Model Random Forest
Sumber: Hasil Penelitian (2026)

4.2. Implementasi Sistem

Model Random Forest yang telah melalui proses pelatihan diintegrasikan ke dalam aplikasi berbasis web menggunakan framework Flask. Sistem dikembangkan sebagai pendukung proses penentuan level belajar siswa baru sehingga pengguna dapat memperoleh rekomendasi level belajar secara otomatis berdasarkan hasil *placement test*. Integrasi model dilakukan menggunakan berkas model berformat *.pkl* yang dihubungkan dengan aplikasi sehingga proses prediksi dapat dilakukan secara langsung melalui antarmuka web.

Pengguna memasukkan tiga atribut utama, yaitu umur siswa, tingkat kelas sekolah, dan nilai *placement test*. Selanjutnya sistem melakukan validasi data, proses *encoding* terhadap atribut kategorikal, serta prediksi menggunakan algoritma Random Forest. Hasil prediksi kemudian ditampilkan sebagai rekomendasi level belajar dan disimpan ke dalam basis data sebagai riwayat *placement test*. Dengan mekanisme tersebut, proses penentuan level belajar menjadi lebih cepat,

konsisten, dan berbasis data dibandingkan proses manual yang sebelumnya digunakan.

The screenshot shows a web form titled "Form Placement Test". It includes input fields for "Nama Siswa" (Student Name), "Umur" (Age), "Kelas Sekolah" (School Class), and "Nilai Placement Test" (Placement Test Score). There are also buttons for "Proses Prediksi" (Process Prediction) and "Lihat History" (View History).

Gambar 3. Halaman Input Data Siswa
Sumber: Hasil Penelitian (2026)

The screenshot shows a web page titled "Hasil Prediksi AI" (AI Prediction Result). It displays the recommended learning level as "Pre Beginner". Below this, it shows the input data: "Nama Siswa: Aif", "Umur: 12 Tahun", "Kelas Sekolah: SD 6", and "Nilai Placement Test: 35".

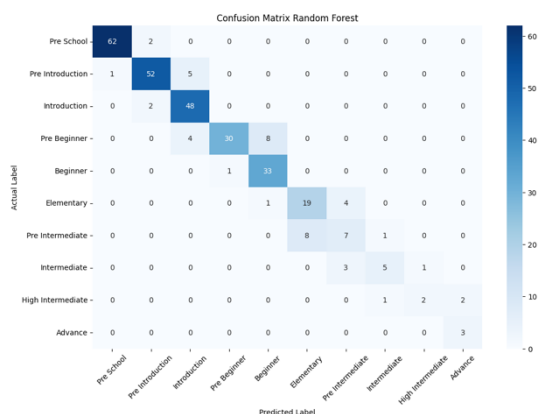
Gambar 4. Hasil Prediksi Level Belajar
Sumber: Hasil Penelitian (2026)

Implementasi model menunjukkan bahwa algoritma Random Forest dapat diintegrasikan dengan baik ke dalam aplikasi berbasis web tanpa mengubah mekanisme klasifikasi yang telah dibangun pada tahap pelatihan. Sistem mampu memberikan rekomendasi level belajar secara otomatis berdasarkan data masukan pengguna sehingga dapat membantu pengajar dalam proses penempatan siswa baru secara lebih objektif. Selain meningkatkan konsistensi penentuan level, sistem juga mempercepat proses pengambilan keputusan dibandingkan metode manual yang sebelumnya bergantung pada observasi dan wawancara pengajar.

4.3. Evaluasi Model

Evaluasi model dilakukan menggunakan *Confusion Matrix* dan *Classification Report* untuk mengukur kemampuan algoritma *Random Forest* dalam mengklasifikasikan level belajar siswa berdasarkan data *testing*. Pengujian dilakukan terhadap model yang telah

dilatih menggunakan pembagian data *training* dan *testing* sebesar 80:20 sehingga diperoleh nilai *accuracy*, *precision*, *recall*, dan *F1-score* sebagai indikator performa model klasifikasi.



Gambar 5. Confusion Matrix Model Random Forest

Sumber: Hasil Penelitian (2026)

Berdasarkan hasil pengujian, model *Random Forest* memperoleh tingkat akurasi sebesar 85,57% dalam mengklasifikasikan level belajar siswa. Nilai akurasi tersebut menunjukkan bahwa sebagian besar data uji berhasil diklasifikasikan dengan benar sehingga model layak digunakan sebagai pendukung proses penentuan level belajar siswa baru. Selain nilai *accuracy*, hasil *Classification Report* juga menunjukkan bahwa model memiliki performa yang baik berdasarkan nilai *precision*, *recall*, dan *F1-score* pada sebagian besar kategori level belajar sehingga dapat digunakan sebagai pendukung dalam proses penentuan level belajar siswa baru.

Tabel 2. Hasil Classification Report Model Random Forest

Level Belajar	Precisi on	Reca ll	F1-Score	Suppo rt
0	0.98	0.97	0.98	64
1	0.93	0.90	0.91	58
2	0.84	0.96	0.90	50
3	0.97	0.71	0.82	42
4	0.79	0.97	0.87	34
5	0.70	0.79	0.75	24
6	0.50	0.44	0.47	16
7	0.71	0.56	0.62	9
8	0.67	0.40	0.50	5
9	0.60	1.00	0.75	3

Level Belajar	Precisi on	Reca ll	F1-Score	Suppo rt
Accuracy			85,57 %	305
Macro Average	0.77	0.77	0.76	305
Weighted Average	0.86	0.86	0.85	305

Sumber: Hasil Penelitian (2026)

Hasil evaluasi menunjukkan bahwa algoritma *Random Forest* mampu menghasilkan performa klasifikasi yang baik dalam menentukan level belajar siswa baru. Meskipun demikian, masih terdapat beberapa kesalahan klasifikasi pada level belajar yang memiliki karakteristik data serupa sehingga menyebabkan sebagian data diprediksi ke level yang berdekatan. Kondisi tersebut menunjukkan bahwa tingkat kemiripan karakteristik antar siswa masih memengaruhi hasil prediksi model. Namun demikian, secara keseluruhan model telah mampu memberikan rekomendasi level belajar secara lebih konsisten dibandingkan proses penentuan level yang dilakukan secara manual.

Hasil penelitian ini sejalan dengan penelitian Amin dkk. (2025) yang menunjukkan bahwa algoritma *Random Forest* memiliki kemampuan klasifikasi yang baik pada data pendidikan dan mampu menghasilkan tingkat akurasi yang tinggi dalam proses prediksi [14]. Temuan tersebut memperkuat bahwa penerapan *Random Forest* pada penelitian ini dapat digunakan sebagai metode yang efektif untuk mendukung proses penentuan level belajar siswa baru secara objektif dan berbasis data.

4.4. Pengujian Fungsional

Pengujian fungsional dilakukan menggunakan metode Black Box Testing untuk memastikan seluruh fitur pada sistem dapat berjalan sesuai dengan kebutuhan fungsional yang telah dirancang. Pengujian meliputi proses login, pengelolaan data pengguna, input data siswa, proses klasifikasi menggunakan algoritma *Random Forest*, penyimpanan hasil prediksi, serta pengelolaan riwayat hasil *placement test*.

Berdasarkan hasil pengujian, seluruh fungsi sistem berhasil berjalan sesuai dengan hasil yang diharapkan tanpa ditemukan kesalahan fungsional. Hal ini menunjukkan bahwa sistem telah memenuhi kebutuhan fungsional dan siap digunakan pada tahap pengujian pengguna.

4.5. Pengujian Pengguna (UAT)

Pengujian pengguna dilakukan menggunakan metode User Acceptance Test (UAT) untuk mengetahui tingkat penerimaan pengguna terhadap sistem yang dikembangkan. Pengujian melibatkan 5 responden yang terdiri dari admin dan pengajar di New Concept English Education Centre Cabang Taman Alamanda. Instrumen pengujian menggunakan kuesioner dengan skala Likert 1–5 yang terdiri atas 10 pernyataan yang mencakup aspek kemudahan penggunaan, tampilan antarmuka, fungsionalitas, efektivitas, akurasi, stabilitas, dan kepuasan pengguna. Hasil rekapitulasi UAT disajikan pada Tabel 3.

Tabel 3. Rekap Hasil UAT 5 Responden

No	Pernyataan	Rata-rata	Kategori
P1	Sistem mudah digunakan saat proses penentuan level belajar siswa	4,2	Setuju
P2	Tampilan antarmuka sistem mudah dipahami	4,0	Setuju
P3	Proses login dan akses sistem berjalan dengan baik	4,4	Sangat Setuju
P4	Sistem dapat membantu menentukan level belajar siswa secara lebih cepat	4,6	Sangat Setuju
P5	Hasil prediksi level belajar yang diberikan sistem mudah dipahami	4,4	Sangat Setuju

No	Pernyataan	Rata-rata	Kategori
P6	Sistem membantu mengurangi kesalahan dalam penentuan level siswa	4,0	Setuju
P7	Proses penginputan data siswa berjalan dengan baik	3,8	Setuju
P8	Fitur pengelolaan data hasil prediksi mudah digunakan	4,4	Sangat Setuju
P9	Sistem berjalan dengan stabil tanpa kendala yang berarti	4,0	Setuju
P10	Secara keseluruhan saya puas terhadap sistem yang dikembangkan	4,2	Setuju

Sumber: Hasil Penelitian (2026)

Berdasarkan hasil rekapitulasi UAT, seluruh pernyataan memperoleh penilaian pada kategori Setuju hingga Sangat Setuju. Pernyataan dengan nilai tertinggi terdapat pada P4, yaitu sistem dapat membantu menentukan level belajar siswa secara lebih cepat dengan nilai rata-rata 4,6 (Sangat Setuju). Sementara itu, nilai terendah terdapat pada P7 mengenai proses penginputan data siswa dengan nilai rata-rata 3,8 (Setuju).

Secara keseluruhan, hasil UAT menunjukkan tingkat penerimaan pengguna sebesar 84%, yang diperoleh dari perbandingan total skor aktual terhadap skor maksimum. Hasil tersebut menunjukkan bahwa sistem mampu membantu admin dan pengajar dalam menentukan level belajar siswa secara lebih mudah, cepat, dan objektif. Selain itu, sistem dinilai memiliki tampilan yang mudah dipahami, fitur yang berjalan dengan baik, serta

mampu memberikan rekomendasi level belajar yang sesuai dengan kebutuhan pengguna. Dengan demikian, sistem yang dikembangkan memiliki tingkat penerimaan pengguna yang baik dan layak digunakan sebagai media pendukung dalam proses penentuan level belajar siswa di New Concept English Education Centre Cabang Taman Alamanda.

Pengujian pengguna dilakukan menggunakan metode User Acceptance Test (UAT) untuk mengetahui tingkat penerimaan pengguna terhadap sistem yang dikembangkan. Metode UAT digunakan untuk mengevaluasi kesesuaian sistem dengan kebutuhan pengguna berdasarkan pengalaman penggunaan secara langsung [15].

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa:

1. Algoritma *Random Forest* berhasil diterapkan untuk mengklasifikasikan level belajar siswa baru di New Concept English Education Centre berdasarkan variabel umur, tingkat kelas sekolah, dan nilai *placement test*. Model yang dibangun memperoleh tingkat akurasi sebesar 85,57%, sehingga mampu menghasilkan klasifikasi level belajar dengan performa yang baik.
2. Model *Random Forest* berhasil diimplementasikan ke dalam aplikasi berbasis web menggunakan framework Flask sehingga mampu memberikan rekomendasi level belajar siswa secara otomatis. Berdasarkan hasil pengujian fungsional, seluruh fitur sistem berjalan sesuai kebutuhan, sedangkan hasil User Acceptance Test (UAT) menunjukkan tingkat penerimaan pengguna sebesar 84%, sehingga sistem layak digunakan sebagai pendukung penentuan level belajar siswa.
3. Penelitian ini masih memiliki keterbatasan pada penggunaan atribut yang terbatas serta hanya menerapkan satu algoritma klasifikasi. Penelitian selanjutnya disarankan menggunakan dataset yang lebih beragam, menambahkan variabel lain yang relevan, serta membandingkan performa *Random Forest* dengan algoritma klasifikasi lainnya untuk memperoleh hasil yang lebih optimal.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada New Concept English Education Centre Cabang Taman Alamanda yang telah memberikan kesempatan untuk pelaksanaan observasi dan penelitian, kepada Ibu Nadya Safitri, S.T., M.Kom. selaku dosen pembimbing, serta kepada seluruh civitas akademika Universitas Bina Insani yang telah mendukung penyelesaian penelitian ini.

DAFTAR PUSTAKA

- [1] P. H. Gunawan and I. V. Paputungan, "Deteksi Tingkat Potensi Kelulusan Calon Mahasiswa menggunakan Algoritma Random Forest Detection of Graduation Potential in Prospective Students using the Random Forest Algorithm," 2025. [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [2] S. A. Putri and R. Rachmatika, "Penerapan Algoritma Random Forest dan SMOTE untuk Prediksi Risiko Putus Sekolah Siswa Sekolah Menengah Kejuruan," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 5, no. 3, pp. 903–910, Oct. 2025, doi: 10.51454/decode.v5i3.1360.
- [3] Christine Eirene, Dian Syafitri, Neny Sulistianingsih, Khasnur Hidjah, and Hairani Hairani, "Classification of Learning Styles of Junior High School Students Using Random Forest & XGBoost Algorithm," *Jurnal Bumigora Information Technology (BITE)*, vol. 7, no. 1, pp. 15–24, Jun. 2025, doi: 10.30812/bite.v7i1.4913.
- [4] M. Nachouki, E. A. Mohamed, R. Mehdi, and M. Abou Naaj, "Student course grade prediction using the random forest algorithm: Analysis of predictors' importance," *Trends Neurosci. Educ.*, vol. 33, Dec. 2023, doi: 10.1016/j.tine.2023.100214.
- [5] A. Sutaryani, S. Sunarno, and D. Djuniadi, "PERBANDINGAN PERFORMA MODEL MACHINE LEARNING DALAM PREDIKSI SUHU DI SEMARANG," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4884.
- [6] A. N. Pratiwi and E. Utami, "Sistemasi: Jurnal Sistem Informasi Prediksi Kinerja Akademik Matematika Siswa berdasarkan Kepribadian Big Five menggunakan Random Forest dengan Teknik Synthetic Minority Over-Sampling Predicting Students' Academic Performance in Mathematics based on Big Five Personality Traits using Random Forest with Synthetic Minority Over-Sampling Technique," 2025. [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>

- [7] S. Danil, N. Rahaningsih, R. D. Dana, and . M., “PENINGKATAN KLASIFIKASI KEMISKINAN INDONESIA MENGGUNAKAN METODE DECISION TREE,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 2, Apr. 2025, doi: 10.23960/jitet.v13i2.6336.
- [8] M. R. Qisthiano, P. A. Prayesy, and I. Ruswita, “Penerapan Algoritma Decision Tree dalam Klasifikasi Data Prediksi Kelulusan Mahasiswa,” *G-Tech: Jurnal Teknologi Terapan*, vol. 7, no. 1, pp. 21–28, Jan. 2023, doi: 10.33379/gtech.v7i1.1850.
- [9] F. Cappelli, G. Castronuovo, S. Grimaldi, and V. Telesca, “Random Forest and Feature Importance Measures for Discriminating the Most Influential Environmental Factors in Predicting Cardiovascular and Respiratory Diseases,” *Int. J. Environ. Res. Public Health*, vol. 21, no. 7, Jul. 2024, doi: 10.3390/ijerph21070867.
- [10] S. Apandi, “Pemodelan Klasifikasi Siswa Berprestasi dengan Random Forest: Studi Kasus pada Bimbingan Belajar,” *Faktor Exacta*, vol. 18, no. 1, p. 63, Jul. 2025, doi: 10.30998/faktorexacta.v18i1.27163.
- [11] F. R. Rinandi, H. Heryanto, and K. Wiyatmoko, “DETEKSI DINI SISWA BERMASALAH MENGGUNAKAN METODE RANDOM FOREST DI SMK MUHAMMADIYAH 3 BANJARMASIN,” *Jurnal Sistem dan Teknologi Informasi*, vol. 9, no. 1, p. 2023, 2023.
- [12] I. Markoulidakis and G. Markoulidakis, “Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis,” *Technologies (Basel)*, vol. 12, no. 7, Jul. 2024, doi: 10.3390/technologies12070113.
- [13] M. Syarif and D. Risdiansyah, “PEMANFAATAN METODE PROTOTYPE DALAM PERANCANGAN SISTEM INFORMASI PENJUALAN BERBASIS WEBSITE,” 2024.
- [14] R. Amin, A. Siwi, and F. Utami, “Prediksi Nilai Ujian Berdasarkan Kebiasaan Siswa Menggunakan Algoritma Random Forest Regressor,” *INFORMATION SYSTEM FOR EDUCATORS AND PROFESSIONALS*, vol. 10, no. 2, pp. 149–158, 2025.
- [15] V. M. Anjasmara and A. H. Sumitro, “Pengembangan Sistem Informasi Masjid Darul Arham Menggunakan Metode V-Model dan UAT (User Acceptance Testing),” *INFORMATION SYSTEM FOR EDUCATORS AND PROFESSIONALS*, vol. 8, no. 1, pp. 47–58, 2023.