**Edusaintek: Jurnal Pendidikan, Sains dan Teknologi**
Volume 12 Issue 3 2025 Pages 1448 - 1461
p-ISSN: 1858-005X e-ISSN: 2655-3392 DOI: https://doi.org/10.47668/edusaintek.v12i3.1800
website: https://journalstkippgrisitubondo.ac.id/index.php/EDUSAINTEK

# PREDIKSI PEMILIHAN PROGRAM STUDI PADA MAHASISWA KIP FAKULTAS TEKNIK MENGGUNAKAN ALGORITMA RANDOM FOREST

**Hamdani[1*], Windarsyah[2], Finki Dona Marleny[3]**
[1,2,3] Universitas Muhammadiyah Banjarmasin, Indonesia
*Corresponding author: 1955201110009@umbjm.ac.id

**Abstract:** Choosing the right study program is crucial to ensuring academic success, particularly for students receiving KIP Kuliah assistance. However, many students select programs that do not align with their interests or potential due to external influences such as parental pressure or limited access to information. This study aims to classify the study program preferences of KIP recipient students from the Faculty of Engineering at Muhammadiyah University of Banjarmasin using the Random Forest algorithm. The data were obtained from students' academic records, which include attributes such as academic grades, previous majors, and chosen study programs. The research involved several stages, including data preprocessing, model training, and performance evaluation using accuracy metrics. The evaluation results showed that the Random Forest model achieved an accuracy rate of 89%, indicating its effectiveness in recommending suitable study programs. These findings are expected to support academic guidance and assist the university in making informed student placement decisions.

**Keywords:** Classification, Kip Kuliah, Study Program Selection, Random Forest.

**Abstrak:** Pemilihan program studi yang tepat sangat penting untuk menunjang keberhasilan akademik, khususnya bagi mahasiswa penerima bantuan KIP Kuliah. Namun, banyak mahasiswa memilih program studi yang tidak sesuai dengan minat atau potensi mereka akibat pengaruh eksternal seperti dorongan orang tua atau keterbatasan informasi. Penelitian ini bertujuan untuk mengklasifikasikan pilihan program studi mahasiswa penerima KIP fakultas teknik di Universitas Muhammadiyah Banjarmasin menggunakan algoritma Random Forest. Data yang digunakan diperoleh dari informasi akademik mahasiswa, yang mencakup atribut nilai akademik, jurusan asal, dan pilihan program studi. Penelitian ini dilakukan melalui tahapan praproses data, pelatihan model, dan evaluasi performa menggunakan metrik akurasi. Hasil evaluasi menunjukkan bahwa model Random Forest mampu mencapai akurasi sebesar 89%, yang menandakan efektivitasnya dalam merekomendasikan program studi yang sesuai. Temuan ini diharapkan dapat menjadi dasar dalam proses pembinaan akademik dan pengambilan keputusan penempatan mahasiswa oleh pihak universitas.

**Kata kunci:** Klasifikasi, Kip Kuliah, Pemilihan Program Studi, Random Forest.

Copyright (c) 2025 The Authors. This is an open-access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

## PENDAHULUAN

Universitas Muhammadiyah Banjarmasin adalah sebuah lembaga pendidikan tinggi swasta di Kalimantan Selatan yang memiliki enam fakultas, termasuk Fakultas Teknik yang menaungi program studi seperti Teknik Sipil, Arsitektur, Perencanaan Wilayah dan Kota, serta Informatika. Setiap program studi diisi oleh mahasiswa dengan latar belakang yang beragam, termasuk para penerima bantuan pendidikan dari program KIP Kuliah.

KIP Kuliah di rancang untuk mendukung akses pendidikan tinggi bagi mahasiswa yang berpotensi memiliki prestasi akademik yang baik tetapi mempunyai keterbatasan ekonomi (Khairunnisa et al., 2024). Sesuai Permendikbud No.10 Tahun 2020 tentang Program Indonesia Pintar (PIP), bantuan KIP disalurkan dalam bentuk kartu elektronik yang dapat digunakan untuk pembiayaan pendidikan, seperti uang kuliah dan kebutuhan akademik lainnya(Putri et al., 2021). Berbagai ketentuan tambahan yang harus dipenuhi oleh calon penerima KIP meliputi, antara lain: (1) berasal dari keluarga yang mendapatkan bantuan sosial seperti Program Keluarga Harapan (PKH) atau yang memiliki Kartu Keluarga Sejahtera (KKS); (2) mahasiswa yang dibesarkan di panti asuhan atau berada di bawah asuhan panti sosial; dan (3) mahasiswa dengan latar belakang keluarga yang tercatat dalam Data Terpadu Kesejahteraan Sosial (DTKS) dibawah naungan Kementerian Sosial.(Kurniawijaya & Karsana, 2023). Program KIP Kuliah terbagi menjadi dua bentuk bantuan, yaitu Beasiswa KIP Kuliah dan bantuan KIP UKT. Beasiswa KIP Kuliah sendiri adalah salah satu bentuk bantuan pendidikan yang menggantikan program Bidikmisi sebelumnya(Gagan Suganda et al., 2022). Program ini sangat membantu dalam meningkatkan kualitas dan keberhasilan akademik mahasiswa penerima manfaat (Saprianto et al., 2023). KIP menjadi pendukung penting bagi mahasiswa dalam mencapai tujuan akademis dan meningkatkan kualitas pendidikan(Novrian et al., 2024). Namun, tidak sedikit mahasiswa yang merasa salah memilih program studi, baik karena tekanan eksternal, keterbatasan informasi, atau pertimbangan kuota penerimaan. Hal ini dapat mempengaruhi prestasi akademik dan motivasi belajar mahasiswa. Selain itu, kurangnya pemahaman terhadap potensi diri, minat, dan bakat juga menjadi faktor yang menyebabkan mahasiswa tidak cocok dengan program studi yang dipilih(T. Y. Pratama & Armansyah, 2024). Kondisi ini tidak hanya berdampak pada performa akademik, tetapi juga dapat menyebabkan mahasiswa merasa kurang nyaman, kehilangan semangat dalam mengikuti perkuliahan, bahkan berisiko untuk pindah jurusan atau berhenti kuliah(A. R. Pratama et al., 2023).

Untuk mengatasi masalah ini, diperlukan pendekatan berbasis teknologi yang mampu memberikan prediksi atau rekomendasi program studi sesuai dengan karakteristik mahasiswa (Aurellia & Apsiswanto, 2022). Penelitian ini memanfaatkan algoritma *Random Forest*, salah satu teknik machine learning yang efektif untuk tugas klasifikasi dan juga mampu menangani data yang tidak seimbang serta mengurangi risiko

overfitting(Hafiz, Muhamad Prayoga & Ermatita, 2024). Dalam penelitian ini, indikator yang digunakan meliputi nilai matematika, nilai bahasa Inggris, nilai seni budaya, dan nilai psikotes sebagai dasar untuk melakukan klasifikasi pemilihan program studi mahasiswa penerima KIP. Diharapkan pendekatan ini dapat mengidentifikasi pola pemilihan jurusan yang tepat berdasarkan data akademik dan latar belakang mahasiswa penerima KIP.

Penelitian sebelumnya telah banyak membahas penerapan algoritma machine learning dalam bidang pendidikan, namun belum secara spesifik mengkaji pemilihan program studi oleh mahasiswa penerima KIP. Penelitian oleh (Setyowati et al., 2025)yang berjudul "*Optimizing Random Forest Parameters with Hyperparameter Tuning for Classifying School-Age KIP Eligibility in West Java*" menggunakan algoritma *Random Forest* dengan *tuning hyperparameter* untuk mengklasifikasikan kelayakan penerima KIP, namun fokusnya terbatas pada siswa sekolah di Jawa Barat dan tidak membahas pemilihan jurusan di perguruan tinggi. Sementara itu, Mulyo & Maori (2024) dalam penelitiannya yang berjudul "Peningkatan Akurasi Prediksi Pemilihan Program Studi Calon Mahasiswa Baru melalui Optimasi Algoritma Decision Tree dan Random Forest" menerapkan algoritma klasifikasi untuk memprediksi program studi calon mahasiswa secara umum, tanpa membatasi pada kelompok penerima beasiswa dan tanpa memasukkan indikator psikotes. Penelitian lain oleh Andrianof et al., (2023) yang berjudul "Implementasi Algoritma Random Forest untuk Prediksi Kelulusan Mahasiswa" berfokus pada prediksi kelulusan mahasiswa, bukan pada proses pemilihan program studi di awal perkuliahan. Berbeda dari penelitian-penelitian tersebut, Penelitian ini secara khusus membahas pemilihan program studi oleh mahasiswa penerima KIP di Universitas Muhammadiyah Banjarmasin dengan menggunakan algoritma Random Forest dan indikator berupa nilai matematika, bahasa Inggris, seni budaya, serta nilai psikotes. Fokus pada kelompok mahasiswa KIP dan penggunaan kombinasi variabel akademik serta psikologis ini belum banyak dijumpai dalam penelitian-penelitian terdahulu, sehingga pendekatan ini diharapkan dapat memberikan kontribusi baru dalam mendukung ketepatan pemilihan program studi.

Tujuan dari penelitian ini adalah untuk menerepkan algoritma *Random Forest* dalam proses klasifikasi pemilihan program studi mahasiswa KIP serta mengevaluasi performanya berdasarkan akurasi prediksi.

## METODE

Penelitian ini memanfaatkan algoritma Random Forest sebagai pendekatan dalam metode klasifikasi. Data yang digunakan merupakan data sekunder dari mahasiswa penerima KIP Kuliah di Universitas Muhammadiyah Banjarmasin sebanyak 91 entri. Setiap entri mencakup atribut seperti jenis kelamin, niai matematika, nilai bahasa inggris, nilai seni buday, nilai pskotesrt, serta program studi yang dipilih.

**Gambar 1.** Metode Penelitian

Prosedur penelitian mencakup:

1. Pengumpulan Data

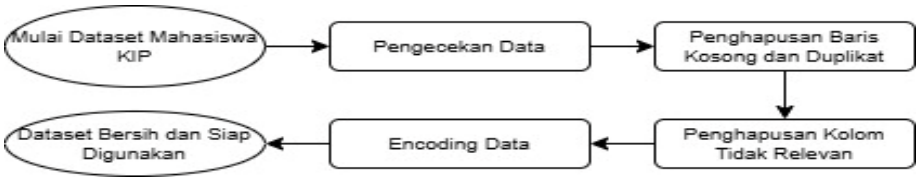
   Data diperoleh dari sistem informasi akademik, mencakup nilai akademik serta hasil tes psikologi. Data ini dimanfaatkan untuk menentukan faktor-faktor yang memengaruhi keputusan pemilihan program studi. Adapun atribut-atribut yang terkandung dalam dataset tersebut dijelaskan sebagai berikut:

**Tabel 1.** Atribut Dataset

| Nama Kolom | TipeData | Penjelasan |
|---|---|---|
| Kode Registrasi | Teks | Identifikasi unik untuk setiap mahasiswa dalam sistem akademik. |
| Nama | Teks | Nama lengkap mahasiswa yang terdaftar. |
| Jenis Kelamin | Teks | Jenis kelamin mahasiswa (Laki-laki atau Perempuan). |
| Nilai Matematika | Numerik | Nilai Akademik Matematika Mahasiswa. |
| Nilai B.Inggris | Numerik | Nilai Akademik Bahasa Inggris Mahasiswa |
| Nilai Senibudaya | Numerik | Nilai Akdemik Seni Budaya Mahasiwa |
| Nilai Psikotest | Numerik | Nilai yang diperoleh mahasiswa dari tes psikologi atau psikotest yang dilakukan. |
| Pilihan Program Studi | Teks | Program studi yang dipilih oleh mahasiswa seperti S1 Informatika, S1 Arsitektur, S1 teknik Sipil, S1 Perencanaan Wilayah dan Kota |

2. Praproses Data

Praproses Data meliputi :




**Gambar 2.** Praproses Data

Flowchart pra-premprosesan data mahasiswa KIP menggambarkan langkah-langkah sistematis yang dilakukan untuk memastikan data siap digunakan dalam proses pelatihan model. Adapun tahapan-tahapan dalam skema tersebut adalah sebagai berikut :

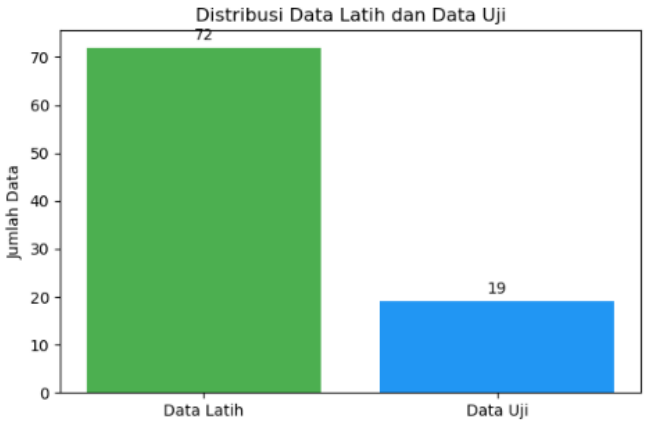
a. Proses dimulai dengan dataset mahasiswa penerima KIP yang berisi berbagai atribut seperti kode registrasi, nama, jenis kelamin, nilai matematika, nilai b.inggris, nilai senibudaya, nilai psikotest, dan pilihan program studi.
b. Pada tahap ini dilakukan pemeriksaan data untuk mengidentifikasi adanya nilai kosong (*missing values*) dan duplikasi. Tujuannya untuk memastikan data yang akan diproses lengkap dan tidak ada pengisian data ganda yang dapat mempengaruhi akurasi model.
c. Jika ditemukan nilai kosong atau data yang terduplikasi, maka baris data tersebut dihapus agar model yang dibangun akurat dan menghasilkan prediksi yang tepat.
d. Kolom identifikasi pribadi seperti 'Kode Registrasi' dan 'Nama' dihapus untuk meningkatkan efisiensi model.
e. Data non-numerik seperti 'jenis kelamin' dan 'pilihan program studi' perlu dikonversi menjadi numerik menggunakan *encoding* agar dapat diproses oleh *Random Forest*.
f. Setelah semua proses pra-pemrosesan selesai, data telah berada dalam kondisi bersih, konsisten, dan berbentuk numerik. Dataset ini kemudian siap digunakan untuk proses pembelajaran dan pengujian model klasifikasi *Random Forest.*

3. Pembagian Data

Proses pembagian data memiliki peran krusial dalam pelatihan model karena memengaruhi kemampuan model dalam mempelajari pola dan menerapkannya pada data yang belum dikenali sebelumnya. Pada penelitian ini, dataset berjumlah 91 entri mahasiswa penerima KIP dibagi ke dalam dua kategori, yaitu data latih dan data uji,

di mana 80 persen data digunakan untuk proses pelatihan, sementara sisanya 20 persen untuk pengujian. Pembagian diterapkan secara acak untuk memastikan distribusi kelas yang merata dan mencegah bias dalam proses klasifikasi.

Sebanyak 72 data atau 80% dari total dataset digunakan untuk proses pelatihan model, sementara 19 data sisanya (20%) dimanfaatkan untuk menguji kinerja model setelah dilatih. Proses pemisahan ini dilakukan untuk menilai seberapa baik model dapat memprediksi informasi yang belum pernah dikenalnya selama proses pelatihan.




**Gambar 3.** distribusi data latih dan data uji

4. Klasifikasi

Klasifikasi adalah proses pengelompokan data ke dalam kategori tertentu berdasarkan nilai-nilai variabel input, dengan tujuan memprediksi kelas atau karakteristik dari objek yang diamati (Yustanti & Rochmawati, 2022). Klasifikasi juga merupakan metode pembelajaran terawasi yang menggunakan data latih untuk menyusun aturan dalam mengelompokkan data uji ke dalam kategori tertentu, dengan pendekatan terstruktur dalam membangun model dari data masukan (Novianto et al., 2024).

5. *Decision Tree*

*Decision Tree* merupakan model berbentuk struktur pohon yang dimanfaatkan untuk klasifikasi, di mana tiap node menguji atribut, cabang menunjukkan hasilnya, dan leaf node menyatakan kelas (Doahir & Qolbi, 2022). *Root node* berada di puncak tanpa cabang masuk, sementara node lainnya memiliki satu cabang masuk dan satu atau lebih cabang keluar(Qadrini et al., 2021).

6. *Random Forest*

*Random Forest* merupakan versi lanjutan dari *Decision Tree* yang membangun pohon-pohon keputusan dengan data dan atribut yang dipilih secara acak(Marlina Haiza et al., 2023). *Random Forest* memanfaatkan sejumlah pohon keputusan sekaligus, hingga dapat mencapai akurasi yang lebih baik dibandingkan dengan penggunaan algoritma *Decision Tree* secara individu.(Saleh et al., 2022). Algoritma *Random Forest* diimplementasikan menggunakan *Python* dan pustaka *scikit-learn*. Proses pelatihan model dilakukan dengan data latih untuk mengklasifikasikan pilihan program studi berdasarkan atribut input.

7. Evaluasi Model

   Model yang sudah dilatih kemudian diuji dengan data uji dan dinilai performanya menggunakan metrik akurasi, sensivitas, dan spesifitas(Shodiqul Munir et al., 2023). Hasil akurasi yagn diperoleh sebesar 89% menunjukan bahwa model memiliki performa yang baik dalam melakukan klasifikasi.

8. Analisis Model

   Setelah model *Random Forest* selesai dilatih dan diuji, dilakukan analisis hasil untuk memahami bagaimana model mengambil keputusan klasifikasi serta seberapa besar pengaruh masing-masing atribut terhadap hasil prediksi. Analisis ini melibatkan dua aspek penting, yakni penilaian kinerja model memanfaatkan ukuran evaluasi seperti akurasi, *precision*, *recall*, dan *f1-score*, serta analisis *feature importance*. Metrik evaluasi dihitung berdasarkan hasil prediksi terhadap data uji, yang menunjukan seberapa baik model mengenali kelas yang sebenarnya. *Feature Importance* digunakan untuk mengidentifikasi atribut mana yang paling berkontribusi dalam proses klasifikasi. Analisis ini memberikan pemahaman mendalam tentang efektivitas model dan peran fitur-fitur dalam proses pengambilan keputusan, yang sangat penting untuk validasi hasil dan pengembangan model dimasa mendatang.

## HASIL DAN PEMBAHASAN

### Hasil

Model *Random Forest* diterapkan pada dataset mahasiswa penerima KIP Fakultas Teknik Universitas Muhammadiyah Banjarmasin sebanyak 91 entri. Dataset tersebut telah melalui proses praproses yang meliputi penghapusan data duplikat, penanganan *missing values*, dan encoding data kategorikal. Selanjutnya, data dibagi dengan rasio 80:20 menjadi 72 data latih dan 19 data uji, sebagaimana ditampilkan pada Gambar 3.

Setelah proses pelatihan, model diuji menggunakan data uji, dan menghasilkan akurasi sebesar 89,47%, yang menunjukkan bahwa dari 19 data uji, sebanyak 17 prediksi sesuai dengan kelas sebenarnya. Nilai akurasi ini mengindikasikan bahwa algoritma *Random Forest* bekerja cukup efektif dalam melakukan klasifikasi program studi mahasiswa berdasarkan data akademik dan psikotest.

Evaluasi dengan metrik *Precision, Recall,* dan *f1-score* menunjukkan bahwa model memiliki kinerja optimal secara keseluruhan. Berdasarkan perhitungan weighted average, diperoleh nilai *Precision* sebesar 82,00%, *Recall* sebesar 89,00%, dan *f1-score* sebesar 85,00%. Nilai-nilai ini menunjukkan keseimbangan antara kemampuan model dalam mengenali kelas (*Recall*) dan ketepatan dalam memprediksi (*Precision*), serta kemampuan umum model dalam mengklasifikasikan data (f1-score).

Sebelum pemodelan dilakukan, fitur-fitur kategorikal seperti jenis kelamin dan nama program studi diubah ke dalam bentuk numerik dengan metode label encoding. Tabel 2 dan Tabel 3 menunjukkan hasil encoding tersebut.

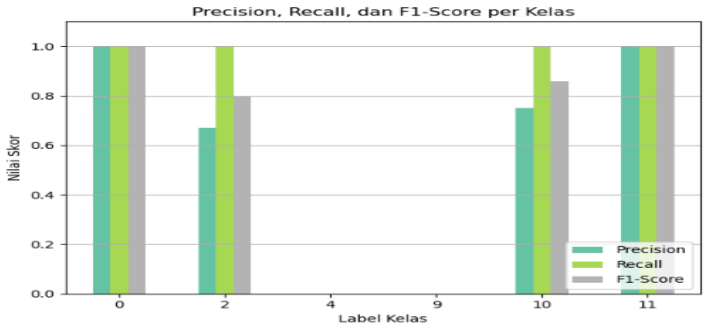
**Tabel 2.** Hasil Encoding Jenis Kelamin

| Jenis Kelamin | Label Numerik |
|---|---|
| Perempuan | 0 |
| Laki-Laki | 1 |

**Tabel 3.** Hasil Encoding Program Studi

| Nama Program Studi | Label Numerik |
|---|---|
| S1 Arsitektur | 0 |
| S1 Informatika | 1 |
| S1 Perencanaan Wilayah dan Kota | 2 |
| S1 Teknik Sipil | 3 |

Untuk memperkuat analisis performa model pada tiap kelas, dilakukan visualisasi skor evaluasi masing-masing kelas dalam bentuk grafik (Gambar 4). Grafik ini menggambarkan nilai *Precision*, *Recall*, dan *f1-score* pada setiap label program studi (kelas).

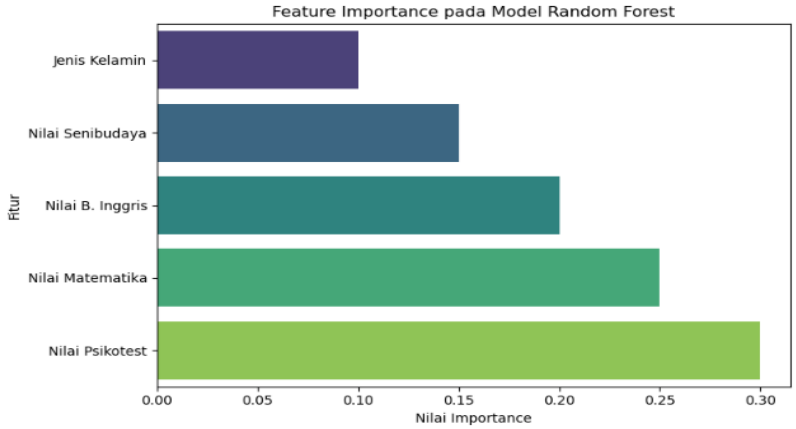



**Gambar 4.** Precision. Recall, dan F1-Score Per kelas

Visualisasi pada Gambar 4 menunjukkan performa model dalam mengenali masing-masing program studi berdasarkan tiga metrik evaluasi, yaitu *Precision*, *Recall*, dan f1-score. Model memperlihatkan hasil sempurna (skor 1.00) dalam memprediksi mahasiswa yang memilih program studi Arsitektur dan Teknik Sipil, artinya semua data uji pada dua program studi tersebut berhasil diklasifikasikan dengan benar. Sementara itu, program studi seperti Informatika dan Perencanaan Wilayah dan Kota juga menunjukkan skor yang tinggi dan stabil. Namun, terdapat pula program studi yang tidak berhasil dikenali sama sekali oleh model, ditandai dengan nilai nol pada seluruh metrik. Hal ini kemungkinan besar terjadi karena jumlah data yang sangat sedikit pada program studi tersebut atau tidak adanya representasi dalam data pelatihan.

Kegagalan model dalam memprediksi beberapa program studi kemungkinan besar disebabkan oleh jumlah data yang tidak seimbang antar program studi. Beberapa program studi hanya memiliki sedikit data dalam proses pelatihan, sehingga model tidak mendapatkan cukup informasi untuk mempelajari pola yang mewakili pilihan tersebut. Kondisi ini membuat model kesulitan dalam mengenali dan mengklasifikasikan data baru yang berasal dari program studi dengan representasi yang sangat terbatas.

Selain evaluasi perfoma model, dilakukan juga analasis terhadap kontribusi masing-masing fitur menggunakan visualisasi *feature importance*. Visualisasi ini menunjukan atribut mana yang paling berpengaruh dalam proses klasifikasi.




**Gambar 5.** Grafik *Fetaure Importance*

Gambar 5 menunjukkan grafik *feature importance* pada model *Random Forest* yang digunakan dalam penelitian ini. Grafik ini menggambarkan kontribusi relatif masing-masing fitur terhadap hasil klasifikasi pemilihan program studi mahasiswa penerima KIP. Dari grafik terlihat bahwa variabel "nilai psikotes" memiliki nilai *importance* tertinggi, yang berarti fitur ini paling berpengaruh dalam proses pengambilan keputusan model. Disusul oleh "nilai matematika" dan "nilai bahasa Inggris" yang juga memiliki kontribusi signifikan. Sementara itu, "nilai seni budaya" dan "jenis kelamin" menunjukkan pengaruh yang lebih rendah. Temuan ini menunjukkan bahwa aspek kognitif dan kemampuan akademik, khususnya dalam bidang logika dan bahasa, lebih menentukan dalam klasifikasi program studi dibandingkan aspek demografis seperti jenis kelamin. Informasi ini penting untuk pemahaman lebih lanjut terhadap faktor-faktor yang relevan dalam penentuan jurusan yang sesuai bagi mahasiswa KIP.

**Pembahasan**

Hasil klasifikasi menggunakan algoritma *Random Forest* menunjukkan performa yang baik dalam memprediksi program studi yang dipilih oleh mahasiswa penerima KIP. Model berhasil mengklasifikasikan sebagian besar data uji secara akurat dengan tingkat akurasi sebesar 89,47%. Nilai *Precision*, *Recall*, dan f1-score yang tinggi mencerminkan bahwa model mampu menjaga keseimbangan antara akurasi prediksi dan kemampuan dalam mengenali data yang sesuai.

Berdasarkan hasil visualisasi feature importance, diketahui bahwa nilai Psikotest menjadi atribut yang paling berpengaruh dalam proses klasifikasi. Artinya, hasil tes psikologi mahasiswa memiliki peran penting dalam membedakan kecenderungan pemilihan program studi. Nilai Matematika dan Bahasa Inggris juga memberikan

kontribusi besar, menunjukkan bahwa kemampuan akademik dasar turut menjadi indikator utama dalam pembentukan pola klasifikasi. Sementara itu, nilai Seni Budaya dan jenis kelamin memberikan pengaruh yang lebih rendah, namun tetap berperan sebagai pendukung dalam proses pengambilan keputusan model.

Berbeda dari beberapa penelitian sebelumnya yang berfokus pada kelayakan penerima KIP secara umum atau prediksi tanpa pendekatan psikometrik, penelitian ini mengarah pada prediksi pemilihan program studi dengan mempertimbangkan nilai psikotes sebagai variabel penting. Fokus tersebut memberikan wawasan tambahan mengenai faktor-faktor yang memengaruhi pemilihan program studi pada kelompok mahasiswa, khususnya bagi penerima KIP. Hal ini juga dapat menjadi masukan bagi penelitian serupa di masa depan untuk memperkaya atribut input dengan memasukkan variabel psikologis atau kognitif guna meningkatkan akurasi dan relevansi model.

Hasil penelitian ini sejalan dengan temuan yang dikemukakan oleh Mulyo dan Maori (2024), yang juga menerapkan algoritma Random Forest dalam memprediksi pilihan program studi mahasiswa baru. Namun demikian, pendekatan yang digunakan dalam penelitian tersebut belum mempertimbangkan indikator psikotes sebagai variabel input. Oleh karena itu, penelitian ini memberikan kontribusi yang lebih terfokus, terutama pada kelompok mahasiswa penerima KIP yang berasal dari latar belakang ekonomi terbatas, serta menambahkan dimensi psikologis yang dapat dijadikan sebagai pertimbangan penting dalam penelitian-penelitian serupa di masa mendatang.

Selain itu, jika dibandingkan dengan penelitian oleh Setyowati et al.(2025), yang memanfaatkan Random Forest untuk mengklasifikasikan kelayakan penerima KIP di tingkat siswa sekolah menengah, pendekatan dalam penelitian ini menghadirkan cakupan yang berbeda. Setyowati et al. tidak mengkaji pemilihan program studi di perguruan tinggi dan tidak menyertakan nilai psikotes sebagai faktor prediksi. Oleh karena itu, penelitian ini dapat menjadi pelengkap terhadap literatur yang ada, dengan memberikan pemahaman yang lebih mendalam mengenai faktor-faktor penentu dalam pengambilan keputusan pendidikan lanjutan bagi mahasiswa penerima bantuan.

Meskipun performa model secara keseluruhan tergolong baik, terdapat beberapa program studi yang tidak berhasil dikenali dengan baik oleh model. Hal ini terlihat dari skor evaluasi nol pada salah satu grafik, yang menunjukkan bahwa prediksi untuk program studi tersebut sepenuhnya salah. Masalah ini kemungkinan disebabkan oleh

jumlah data yang sangat terbatas, atau bahkan tidak ada representasi yang cukup dalam data latih untuk program studi tertentu. Kondisi ini menyebabkan model kekurangan informasi yang dibutuhkan untuk mempelajari karakteristik pilihan program studi tersebut, sehingga model tidak mampu memberikan prediksi yang tepat.

Temuan ini menunjukkan bahwa kualitas data dan cara data disebarkan ke dalam berbagai kategori sangat mempengaruhi keberhasilan model dalam melakukan prediksi yang tepat. Ketidakseimbangan data antar program studi dapat mempengaruhi kemampuan model dalam mengenali dan mengklasifikasikan setiap kategori dengan seimbang. Untuk penelitian selanjutnya, akan lebih baik jika dataset diperluas dan distribusi data disesuaikan, sehingga setiap program studi memperoleh representasi yang cukup. Selain itu, pendekatan seperti oversampling pada kategori dengan data terbatas dapat menjadi alternatif untuk mengurangi ketimpangan dalam data.

Secara keseluruhan, penerapan algoritma *Random Forest* dalam penelitian ini menunjukkan potensi besar untuk dijadikan alat bantu dalam memberikan rekomendasi pemilihan program studi yang sesuai dengan karakteristik dan potensi mahasiswa penerima KIP. Dengan penyempurnaan data dan pengujian lebih lanjut, model ini dapat dikembangkan menjadi sistem pendukung keputusan yang bermanfaat bagi pihak universitas.

## SIMPULAN

Penelitian ini berhasil menerapkan algoritma *Random Forest* untuk mengklasifikasikan pemilihan program studi mahasiswa penerima KIP Kuliah di Universitas Muhammadiyah Banjarmasin. Model yang dihasilkan menunjukkan akurasi 89,47% dengan metrik *Precision, Recall,* dan *f1-score* yang baik. Faktor-faktor seperti nilai Psikotest, Matematika, dan Bahasa Inggris berpengaruh besar dalam klasifikasi, sementara ketidakseimbangan data antar program studi mempengaruhi prediksi untuk beberapa kategori. Hasil ini menunjukkan bahwa *Random Forest* memiliki potensi sebagai alat bantu dalam merekomendasikan program studi yang sesuai dengan karakteristik mahasiswa, terutama jika didukung dengan dataset yang lebih luas dan berkualitas untuk meningkatkan akurasi hasil. Untuk penelitian selanjutnya, disarankan menggunakan dataset yang lebih besar dan seimbang serta menambahkan fitur-fitur tambahan seperti minat, motivasi belajar, dan data kepribadian agar hasil prediksi lebih

akurat dan bermanfaat sebagai sistem pendukung pengambilan keputusan dalam penentuan program studi yang sesuai.

## DAFTAR RUJUKAN

Andrianof, H., Gusman, A. P., & Putra, O. A. (2023). Implementasi algoritma random forest untuk prediksi kelulusan mahasiswa berdasarkan data akademik: studi kasus di perguruan tinggi indonesia. *Jurnal Sains Informatika Terapan (JSIT)*, *4*(1), 24–28. https://doi.org/https://doi.org/10.62357/jsit.v4i2.464

Aurellia, O., & Apsiswanto, U. (2022). Penerapan metode simple additive weighting (saw) untuk menyeleksi penerima bantuan langsung tunai dana desa (Blt Dd). *EDUSAINTEK: Jurnal Pendidikan, Sains Dan Teknologi*, *9*(3), 930–942. https://doi.org/10.47668/edusaintek.v9i3.624

Doahir, A., & Qolbi, A. N. (2022). Analisis potensi siswa untuk kuliah dengan classification menggunakan metode decision tree. *Jurnal POROS TEKNIK*, *14*(1), 28–32.

Gagan Suganda, Marsani Asfi, Ridho Taufiq Subagio, & Ricky Perdana Kusuma. (2022). Penentuan penerima bantuan beasiswa kartu indonesia pintar (kip) kuliah menggunakan naïve bayes classifier. *JSiI (Jurnal Sistem Informasi)*, *9*(2), 193–199. https://doi.org/10.30656/jsii.v9i2.4376

Hafiz, Muhamad Prayoga, B., & Ermatita. (2024). Analisis pemilihan jurusan pada calon siswa smk negeri 4 palembang pada faktor penentu pemilihan jurusan menggunakan association rule dan random forest analysis of major selection for prospective students of smk negeri 4 palembang on determining factors f. *Jurnal Pendidikan Dan Teknologi Indonesia (JPTI)*, *4*(12), 537–547.

Khairunnisa, Anedea, T., & Novianti, I. (2024). Evaluasi penerpan program kip-kuliah mengacu pada pasal 76 uu no.12 tahun 2012 menggunakan metode seven tools di teknik industri universitas pamulang. *AKADEMIK: Jurnal Mahasiswa Humanis*, *4*(3), 753–762.

Kurniawijaya, P. A., & Karsana, I. W. W. (2023). Implementasi metode ahp dalam sistem penunjang keputusan penerima kip kuliah. *JUKI : Jurnal Komputer Dan Informatika*, *5*(1), 22–31.

Marlina Haiza, Elmayati, Zulius Antoni, & Wijaya Harma Oktafia Lingga. (2023). Penerapan algoritma random forest dalam klasifikasi penjurusan di sma negeri tugumulyo. *penerapan kecerdasan buatan*, *4*(2), 138–143.

Mulyo, H., & Maori, N. A. (2024). Peningkatan akurasi prediksi pemilihan program studi calon mahasiswa baru melalui optimasi algoritma decision tree dengan teknik pruning dan ensemble. *Jurnal Disprotek*, *15*(1), 15–25. https://doi.org/10.34001/jdpt.v15i1.5585

Novianto, E., Suhirman, S., & Prasetyo, D. (2024). *Perbandingan metode klasifikasi random forest dan support vector machine dalam memprediksi capaian studi mahasiswa*. *9*(4), 1821–1833.

Novrian, R., Agustiani, T., Fikri, M., Hikmatulloh, M. F., Gunawan, M. E., & Firdaus, U. (2024). Penerapan algoritma random forest dalam prediksi status penerima pip pada

siswa: studi kasus pada smk amaliah 1. *Karimah Tauhid*, *3*(2), 1791–1799. https://doi.org/10.30997/karimahtauhid.v3i2.11937

Pratama, A. R., Aryanto, R. R., & Pratama, A. T. M. (2023). Sistem rekomendasi program studi sarjana berbasis machine learning untuk model klasifikasi calon mahasiswa baru. *Journal of Information Technology and Society*, *1*(1), 11–14. https://doi.org/10.35438/jits.v1i1.20

Pratama, T. Y., & Armansyah. (2024). Decision Tree C4.5 dengan teknik information gain untuk klasifikasi pemilihan program studi tingkat lanjut. *Journal of Information System Research (JOSH)*, *5*(4), 1042–1052. https://doi.org/10.47065/josh.v5i4.5643

Putri, N. A. Y., Subagio, R. T., & Asfi, M. (2021). Sistem pendukung keputusan penilaian kinerja mahasiswa kip kuliah dengan penerapan metode topsis dan promethee. *Jurnal Media Informatika Budidarma*, *5*(4), 1394. https://doi.org/10.30865/mib.v5i4.3268

Qadrini, L., Seppewali, A., & Aina, A. (2021). Decision tree dan adaboost pada klasifikasi penerima program bantuan sosial. *JIP: Jurnal Inovasi Penelitian*, *2*(7).

Saleh, J. S., M.Adrian, A., & Sanger, J. B. (2022). Sistem klasifikasi kelulusan mahasiswa dengan algoritma random forest. *Jurnal Ilmiah Realtech*, *18*(1), 10–14.

Saprianto, R., Raysharie, P. I., Hukom, A., & Takari, D. (2023). Implementasi kip kuliah pada mahasiswa/i universitas palangkaraya implementation of lecturing kip for students of palangkaraya university. *MUQADIMMAH: Jurnal Ekonomi, Manajemen, Akuntansi Dan Bisnis*, *1*(2).

Setyowati, S. L., Qalbi, A., Aristawidya, R., Sartono, B., & Firdawanti, A. R. (2025). Optimizing random forest parameters with hyperparameter tuning for classifying school-age kip eligibility in west java optimizing random forest parameters with hyperparameter tuning for classifying school-age kip eligibility in west java. *JAMBURA JOURNAL OF MATHEMATICS*, *7*(1), 40–48. https://doi.org/https://doi.org/10.37905/jjom.v7i1.28736 ©

Shodiqul Munir, A., Sunyoto, A., & Al Fatta, H. (2023). Implementasi metode seleksi fitur menggunakan artificial bee colony pada klasifikasi retinal nerve fiber layer. *EDUSAINTEK: Jurnal Pendidikan, Sains Dan Teknologi*, *10*(2), 541–553. https://doi.org/10.47668/edusaintek.v10i2.782

Yustanti, W., & Rochmawati, N. (2022). Analisis algoritma klasifikasi untuk memprediksi karakteristik mahasiswa pada pembelajaran daring. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, *8*(1), 57. https://doi.org/10.26418/jp.v8i1.50452