

Prediksi Kecemasan Penempatan Kerja dan Klasifikasi Ketergantungan AI pada Mahasiswa Menggunakan Random Forest

Predicting Job Placement Anxiety and AI Dependency Classification Among College Students Using Random Forest

Muhammad Rijal^{1,*}; Raden Wirawan²; Nurani³

^{1,2,3} Institut Teknologi dan Bisnis Nobel Indonesia, Makassar, 90221, Indonesia

¹ rijal2303@gmail.com; ² raden@nobel.ac.id; ³ Nurani@nobel.ac.id

* Corresponding author

Abstrak

Transformasi digital oleh kecerdasan buatan (AI) generatif telah mengubah lanskap pendidikan tinggi, namun memunculkan tantangan psikologis berupa kecemasan terkait dunia kerja dan ketergantungan berlebihan pada AI. Penelitian ini bertujuan untuk memprediksi tingkat kecemasan penempatan kerja mahasiswa serta mengklasifikasikan tingkat ketergantungan AI menggunakan algoritma Random Forest, sekaligus menghasilkan rekomendasi intervensi yang dipersonalisasi. Dataset sekunder digunakan mencakup 15.000 catatan mahasiswa dengan 30 variabel. Model Random Forest Regressor (150 pohon, kedalaman 15) menghasilkan R^2 sebesar 0,815 untuk prediksi kecemasan penempatan kerja. Analisis kepentingan fitur menunjukkan bahwa kecemasan wawancara (59,3%) dan kejelasan karier (11,8%) menjadi faktor paling dominan, sementara durasi penggunaan AI harian hanya berkontribusi kecil (1,06%). Model Random Forest Classifier (150 pohon, kedalaman 15) mencapai akurasi 72,0% dalam mengklasifikasikan ketergantungan AI ke dalam kategori rendah, sedang, dan tinggi. Sistem rekomendasi menghasilkan intervensi personalisasi bagi 29,5% mahasiswa untuk mengurangi durasi penggunaan AI yang berlebihan. Karena menggunakan dataset sintesis, generalisasi hasil ke populasi mahasiswa riil memerlukan validasi lebih lanjut. Penelitian ini menegaskan bahwa AI bukanlah ancaman utama terhadap kecemasan karier, namun pola penggunaan yang bersifat pengganti perlu dikelola dengan kesadaran dan keseimbangan kognitif.

Kata Kunci: Random forest, kecemasan penempatan kerja, ketergantungan AI, klasifikasi, regresi

Abstract

The digital transformation driven by generative artificial intelligence (AI) has changed the landscape of higher education, but has also raised psychological challenges in the form of anxiety about the world of work and excessive reliance on AI. This study aims to predict students' job placement anxiety levels and classify their AI dependency levels using the Random Forest algorithm, while also generating personalised intervention recommendations. A secondary dataset was used, comprising 15,000 student records with 30 variables. The Random Forest Regressor model (150 trees, depth 15) produced an R^2 value of 0.815 for predicting job placement anxiety. Feature importance analysis revealed that interview anxiety (59.3%) and career clarity (11.8%) were the most dominant factors, while daily AI usage duration only contributed slightly (1.06%). The Random Forest Classifier model (150 trees, depth 15) achieved 72.0% accuracy in classifying AI dependency into low, moderate, and high categories. The recommendation system generated personalised interventions for 29.5% of students to reduce excessive AI usage. However, as a synthetic dataset was used, generalisation of the results to a real-life student population requires further validation. This study confirms that AI is not the main cause of career anxiety, but the replacement usage pattern needs to be managed with awareness and cognitive balance.

Keywords: Random forest, job placement anxiety, AI dependency, classification, regression

Pendahuluan

Transformasi digital yang didorong oleh pesatnya perkembangan kecerdasan buatan (AI) generatif telah mengubah secara fundamental lanskap pendidikan tinggi dalam tiga tahun terakhir. Sejak peluncuran ChatGPT pada November 2022 yang mencatat 100 juta pengguna hanya dalam dua bulan [1], berbagai platform AI seperti Gemini, DeepSeek, dan Perplexity telah menjadi bagian tak terpisahkan dari rutinitas akademik mahasiswa. Sebuah studi lintas budaya di 27 universitas Yordania menemukan bahwa 94,3% mahasiswa mengenal ChatGPT dan 90,4% menggunakannya untuk tugas-tugas akademik [2]. Di Indonesia, Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) mencatat pada 2025 bahwa hampir 50% pengguna AI berasal dari kelompok Gen Z, dengan platform AI dimanfaatkan secara masif untuk

menyusun tugas kuliah, merangkum materi, hingga memahami isu sosial [3]. Fenomena ini menandakan bahwa AI telah menjadi infrastruktur kognitif yang meresap ke dalam proses belajar mahasiswa.

Di balik kemudahan yang ditawarkan, integrasi AI yang masif melahirkan dua tantangan psikologis yang saling terkait, yaitu kecemasan terkait dunia kerja (career anxiety) dan ketergantungan berlebihan pada AI (AI dependency). Kecemasan akan masa depan karier semakin nyata di kalangan mahasiswa yang khawatir profesi mereka kehilangan relevansi akibat otomatisasi [4]. Namun penelitian terbaru menunjukkan bahwa AI bukanlah penyebab utama kecemasan kerja, kontribusi durasi penggunaan AI harian terhadap prediksi kecemasan penempatan kerja relatif kecil, sementara faktor internal seperti kecemasan wawancara dan kejelasan karier jauh lebih dominan [5]. Sementara itu, ketergantungan pada AI telah teridentifikasi sebagai faktor krusial yang terkait dengan proses belajar dan kesehatan mental mahasiswa, terutama ketika AI digunakan secara substitutif menggantikan kerja kognitif mahasiswa [6]. Hal ini menegaskan bahwa dampak AI terhadap kesehatan mental mahasiswa sangat bergantung pada pola penggunaannya.

Studi terkini semakin memperkuat urgensi pemahaman terhadap dinamika ini [7] menunjukkan bahwa kecemasan terhadap AI memiliki efek “pisau bermata dua” (double-edged sword), kecemasan akan penggantian pekerjaan oleh AI menghambat adopsi teknologi, sementara kecemasan akan ketidakmampuan memahami AI justru dapat mempromosikan keberhasilan diri. [8] Menyebut fenomena penurunan kemampuan berpikir mandiri akibat penggunaan AI berlebihan sebagai “cognitive debt cycle”, yaitu siklus utang kognitif di mana penggunaan AI rutin memperlemah kebiasaan intelektual mahasiswa, yang kemudian memperkuat ketergantungan. Temuan ini selaras dengan teori cognitive offloading yang menyatakan bahwa ketergantungan pada alat eksternal mengurangi usaha kognitif selama proses encoding pengetahuan.

Meskipun penelitian tentang kecemasan dan ketergantungan AI telah berkembang, sebagian besar studi masih bersifat eksplanatif dan berhubungan. Pendekatan ini memiliki keterbatasan dalam akurasi prediksi individual dan kapasitas mengidentifikasi pola non-linear yang kompleks. Kesenjangan ini menciptakan urgensi untuk mengadopsi pendekatan machine learning yang lebih canggih. Penerapan algoritma machine learning dalam prediksi kondisi psikologis mahasiswa menunjukkan potensi menjanjikan. [9] Menggunakan random forest classifier untuk memprediksi skrining kecemasan dan depresi pada 274 siswa dengan akurasi 80–95%. [10] Menemukan bahwa Random Forest adalah model terbaik di antara 12 algoritma untuk prediksi kesehatan mental mahasiswa (akurasi 95,6%). [11] Menggunakan Random Forest dan XGBoost untuk mengkategorikan gaya belajar siswa SMP Negeri 2 Mataram, dengan penerapan SMOTE dan hyperparameter tuning, mencapai akurasi tertinggi 84% pada algoritma XGBoost. Namun, penerapan Random Forest secara bersamaan untuk memprediksi kecemasan kerja sekaligus mengklasifikasikan ketergantungan AI pada populasi mahasiswa skala besar masih sangat terbatas.

Kebaruan (state of the art) penelitian ini terletak pada tiga hal. Pertama, integrasi dua target prediksi kecemasan penempatan kerja dan ketergantungan AI dalam satu kerangka machine learning yang konsisten menggunakan algoritma Random Forest. Kedua, penggunaan dataset berskala besar (15.000 mahasiswa, 30 variabel) yang mencakup dimensi perilaku akademik, pola penggunaan AI, kesehatan mental, dan persiapan karier. Ketiga, rekomendasi intervensi yang dipersonalisasi dan berbasis data, bukan sekadar prediksi. Personalisasi ini penting karena tidak semua mahasiswa rentan terhadap dampak negatif AI dengan cara yang sama.

Meskipun sudah banyak penelitian tentang kecemasan dan ketergantungan terhadap AI, sebagian besar masih bersifat menjelaskan dan melihat hubungan antar variabel saja. Ada beberapa kekurangan yang menjadi celah untuk diteliti lebih lanjut: (1) belum ada penelitian yang sekaligus memprediksi kecemasan kerja dan menggolongkan tingkat ketergantungan AI dalam satu kerangka pembelajaran mesin yang terpadu, (2) cara-cara statistik biasa yang dipakai selama ini kurang mampu menangkap pola yang tidak lurus (pola non-linear) antara kebiasaan memakai AI dengan kondisi psikologis mahasiswa, (3) penelitian yang ada biasanya hanya berhenti pada tahap menemukan faktor-faktor risiko, belum sampai pada memberikan saran intervensi yang disesuaikan dengan data masing-masing mahasiswa.

Kebaruan dari penelitian ini ada tiga hal. Pertama, menggabungkan dua sasaran prediksi (kecemasan penempatan kerja dan ketergantungan AI) dalam satu kerangka pembelajaran mesin yang konsisten dengan menggunakan algoritma Random Forest. Kedua, menggunakan data dalam jumlah besar (15.000 mahasiswa, 30 variabel) yang mencakup berbagai sisi sekaligus, seperti perilaku akademik, pola penggunaan AI, kesehatan mental, dan persiapan karier. Ketiga, membuat sistem rekomendasi intervensi yang bersifat pribadi berdasarkan data, bukan hanya sekadar memberikan hasil ramalan. Hal ini penting karena tidak semua mahasiswa rentan terhadap dampak buruk AI dengan cara yang sama, sehingga intervensi bisa dirancang secara tepat sasaran.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk: (1) memprediksi tingkat kecemasan penempatan kerja mahasiswa berdasarkan pola ketergantungan AI dan berbagai faktor kependudukan, akademik, serta psikologis dengan menggunakan algoritma Random Forest Regressor, (2) menggolongkan tingkat ketergantungan AI mahasiswa ke dalam tiga kelompok (rendah, sedang, dan tinggi) dengan menggunakan algoritma Random Forest Classifier, serta (3) menghasilkan rekomendasi intervensi yang dipersonalisasi dan berbasis data untuk membantu mahasiswa menggunakan AI secara bijak (bukan menghindarinya). Dengan pendekatan ini, penelitian ini diharapkan dapat memberikan sumbangan teoritis dalam memperkaya literatur psikologi pendidikan di era AI.

Metode

A. Sumber Data dan Karakteristik Sampel

Penelitian ini menggunakan dataset sekunder sintesis yang diambil dari repositori publik Kaggle berjudul "AI Dependency, Career Anxiety and Student Burnout" karya Sridip Basu [12] yang mencakup 15.000 catatan mahasiswa dengan 30 variabel meliputi data demografis (usia, jenis kelamin, jenis gelar, program studi, tingkat perguruan tinggi, lokasi tempat tinggal), perilaku penggunaan AI (durasi penggunaan harian, frekuensi penggunaan AI untuk tugas, skor AI menggantikan pemikiran sendiri), aktivitas akademik (jam belajar harian, jam belajar mandiri mingguan, jumlah kursus pengembangan keterampilan), kesehatan mental (tingkat stres, skor kelelahan, skor motivasi, jam tidur), serta persiapan dan kecemasan karier (kejelasan karier, pengalaman magang, jumlah lamaran pekerjaan mingguan, kepercayaan diri pada resume, skor kecemasan wawancara, skor kecemasan penempatan kerja, dan akses konseling karier). Dataset ini dibangun menggunakan pendekatan latent-variable probabilistic framework sehingga mensimulasikan hubungan antarperilaku di lingkungan pendidikan tinggi yang realistis, serta mengandung nilai kosong terkontrol sekitar 1,5% pada beberapa kolom untuk menyerupai kondisi survei dunia nyata.

Dataset yang digunakan bersifat sintesis dari repositori Kaggle, bukan data primer dari survei mahasiswa langsung. Meskipun dataset dibangun menggunakan latent-variable probabilistic framework untuk mensimulasikan hubungan antarperilaku secara realistis, hasil penelitian ini tetap memiliki keterbatasan dalam validitas eksternal. Generalisasi ke populasi mahasiswa riil perlu diverifikasi lebih lanjut melalui replikasi dengan data primer.

Untuk memastikan model tidak mengalami overfitting, dilakukan validasi silang 5-lip yang memberikan R^2 stabil ($0,8050 \pm 0,0100$) serta hasil uji yang konsisten antara data latih dan uji. Selain itu, pemilihan `max_depth=15` dilakukan untuk membatasi kompleksitas pohon, dan tidak dilakukan grid search ekstensif untuk menghindari penyesuaian berlebihan terhadap data uji.

B. Pra-pemrosesan Data

Tahap pra-pemrosesan dilakukan dengan langkah-langkah berikut:

1. Penanganan nilai kosong (missing values): Untuk kolom kategorikal `primary_ai_tools_used`, nilai kosong diisi dengan modus. Untuk kolom numerik (`self_learning_hours_per_week`, `social_media_hrs_per_day`, `sleep_hours`, `seeks_career_counseling`), nilai kosong diisi dengan median. Pemilihan modus didasarkan pada pertimbangan bahwa modus merupakan nilai yang paling sering muncul sehingga tidak mengganggu distribusi data kategorikal. Sementara itu, median dipilih untuk variabel numerik karena lebih tahan terhadap pencilan (outlier) dibandingkan rata-rata, sehingga representasi data tetap terjaga.
2. Rekayasa fitur (feature engineering): Variabel target untuk klasifikasi dibentuk dari `ai_dependency_score` menggunakan pemotongan (binning) dengan interval $[0,3]$ untuk kategori Rendah, $[3,7]$ untuk Sedang, dan $[7,10]$ untuk Tinggi.
3. Encoding variabel kategorikal: Sebanyak tujuh variabel kategorikal (`gender`, `degree_type`, `stream`, `college_tier`, `urban_or_rural`, `primary_ai_tools_used`, `uses_ai_for_assignments`) diubah menjadi numerik menggunakan `LabelEncoder` dari `scikit-learn`.

Setelah pemrosesan, diperoleh matriks fitur X yang terdiri dari 17 fitur numerik dan 7 fitur hasil encoding (total 24 fitur). Variabel target regresi adalah `placement_anxiety_score` (skala 0–10) dan target klasifikasi adalah `ai_dependency_category` (Rendah, Sedang, Tinggi).

C. Algoritma Random Forest

Random Forest merupakan metode ensemble learning yang membangun banyak pohon keputusan (decision trees) [13] pada saat pelatihan dan menggabungkan prediksi dari setiap pohon. Untuk regresi, prediksi akhir adalah rata-rata dari prediksi semua pohon untuk klasifikasi, prediksi akhir ditentukan melalui majority voting. Keunggulan Random Forest meliputi ketahanan terhadap overfitting, kemampuan menangani data berdimensi tinggi, dan kemampuan memberikan feature importance [14].

D. Pembagian Data dan Standarisasi

Dataset dibagi menjadi data latih (training set) dan data uji (test set) dengan proporsi 80:20 menggunakan stratified sampling khusus untuk tugas klasifikasi guna menjaga proporsi kelas. Untuk regresi, pembagian dilakukan secara acak dengan `random_state=42` untuk reproduksibilitas. Sebelum pelatihan, semua fitur numerik distandarisasi menggunakan `StandardScaler` sehingga memiliki rata-rata 0 dan standar deviasi 1.

E. Pelatihan dan Hyperparameter Model

Model Random Forest dilatih dengan konfigurasi hyperparameter sebagai berikut:

Tabel 1. Hyperparameter Model Random Forest

<i>Parameter</i>	<i>Nilai Regresi</i>	<i>Nilai Klasifikasi</i>
n_estimators (jumlah pohon)	150	150
max_depth (kedalaman maksimal pohon)	15	15
random_state	42	42

Parameter dipilih berdasarkan keseimbangan antara kompleksitas model dan kemampuan generalisasi. Tidak dilakukan grid search secara luas untuk menghindari overfitting mengingat ukuran data yang besar. Parameter n_estimators=150 dipilih berdasarkan uji awal yang menunjukkan bahwa penambahan jumlah pohon di atas 150 hanya memberikan peningkatan performa marginal (kurang dari 0,1%) namun secara signifikan memperlambat waktu komputasi. Kedalaman maksimal (max_depth=15) dipilih untuk mencegah overfitting karena jumlah fitur yang digunakan adalah 24 variabel; kedalaman 15 dinilai cukup untuk menangkap interaksi kompleks antar fitur tanpa menghafal noise pada data latih.

F. Evaluasi Model

Evaluasi model regresi menggunakan metrik:

1. MSE (Mean Squared Error): mengukur rata-rata kuadrat selisih antara nilai aktual dan prediksi. Semakin kecil nilai MSE semakin baik
2. R^2 (Koefisien Determinasi): mengukur proporsi variansi variabel target yang dapat dijelaskan oleh model Nilai berkisar antara 0 hingga 1, dengan nilai mendekati 1 menunjukkan model yang sangat baik
3. Cross-validation R^2 : validasi silang 5-lip pada data latih untuk menilai stabilitas model

Evaluasi model klasifikasi menggunakan metrik:

1. Akurasi: proporsi prediksi yang benar dari seluruh data uji
2. Presisi: proporsi prediksi positif yang benar
3. Recall: proporsi data positif aktual yang terdeteksi dengan benar
4. F1-score: rata-rata harmonik presisi dan recall
5. Confusion Matrix: matriks 3x3 yang menunjukkan jumlah prediksi benar dan salah untuk setiap kategori (Rendah, Sedang, Tinggi)

Selain pembagian data latih dan uji, model juga dievaluasi menggunakan validasi silang 5-lip (5-fold cross-validation) pada data latih untuk menilai stabilitas dan kemampuan generalisasi model. Prosedur ini membagi data latih menjadi 5 subset secara bergantian, di mana 4 subset digunakan untuk pelatihan dan 1 subset untuk validasi, diulang sebanyak 5 kali. Hasil validasi silang kemudian dilaporkan bersama hasil pengujian pada data uji.

G. Feature Importance

Nilai kepentingan fitur (feature importance) dihitung berdasarkan rata-rata penurunan kemurnian (mean decrease in impurity) yang diakumulasi dari semua pohon dalam Random Forest [15]. Fitur dengan nilai kepentingan lebih tinggi menunjukkan kontribusi yang lebih besar terhadap prediksi model. Sepuluh fitur teratas ditampilkan untuk kedua tugas.

H. Sistem Rekomendasi Intervensi

Berdasarkan output prediksi model, sistem rekomendasi dihasilkan dengan aturan sederhana berikut:

1. Jika $\text{daily_ai_tool_usage_hrs} > 3 \rightarrow$ rekomendasi "Kurangi waktu penggunaan harian"
2. Jika $\text{uses_ai_for_assignments}$ bernilai "Frequently" atau "Always" $\rightarrow$ rekomendasi "Kurangi penggunaan AI untuk tugas"
3. Jika $\text{ai_dependency_score} > 7 \rightarrow$ rekomendasi "Kurangi ketergantungan saat ujian"

Jika tidak ada kondisi yang terpenuhi, rekomendasi adalah "Tidak perlu - sudah terkendali". Rekomendasi bersifat multilabel (dapat mencakup lebih dari satu saran) dan disimpan dalam berkas CSV bersama dengan prediksi model untuk setiap mahasiswa.

I. Uji Statistik Tambahan

Dua uji statistik dilakukan untuk menguji hubungan antara variabel demografis dengan ketergantungan AI :

1. ANOVA satu arah: untuk menguji perbedaan rata-rata $\text{ai_dependency_score}$ antar kelompok college_tier (tingkat perguruan tinggi). Signifikansi ditetapkan pada $\alpha = 0,05$.

2. Uji Chi-square: untuk menguji hubungan antara gender (jenis kelamin) dengan kategori ketergantungan AI (ai_dependency_category). Signifikansi ditetapkan pada $\alpha = 0,05$

J. Alur Penelitian

Berikut adalah alur penelitian yang diimplementasikan dalam kode:



Gambar 1 Alur Penelitian

Seluruh kode dijalankan pada lingkungan pengembangan PyCharm dengan bahasa pemrograman Python.

Hasil dan Diskusi

Penelitian ini memprediksi tingkat kecemasan penempatan kerja (placement anxiety score) dan mendeteksi tingkat ketergantungan AI (AI dependency category) pada 15.000 mahasiswa menggunakan algoritma Random Forest. Seluruh proses pemodelan dan evaluasi dilakukan dengan pembagian data latih (80%) dan data uji (20%) serta validasi silang 5-lip.

A. Prediksi Kecemasan Penempatan Kerja (Regresi)

Model Random Forest Regressor dengan 150 pohon dan kedalaman maksimal 15 berhasil memprediksi skor kecemasan penempatan kerja dengan hasil sebagai berikut:

1. MSE (Mean Squared Error) = 1,1647
2. R² (Koefisien Determinasi) = 0,8147
3. Cross-validation R² (5-lip) = 0,8050 ± 0,0100

Nilai R² sebesar 0,815 menunjukkan bahwa model mampu menjelaskan 81,5% variansi pada data uji, yang termasuk dalam kategori sangat baik untuk prediksi fenomena psikologis. Hasil validasi silang yang stabil (standar penyimpangan kecil) mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik dan tidak mengalami overfitting.

Analisis kepentingan fitur (feature importance) mengungkapkan sepuluh variabel paling berpengaruh terhadap prediksi kecemasan penempatan kerja



Gambar 2. Feature Importance Regresi

Dari gambar terlihat bahwa `interview_anxiety_score` mendominasi dengan kontribusi hampir 60%, diikuti oleh `career_clarity_score` (11,8%) dan `daily_study_hours` (8,4%). Menariknya, variabel penggunaan AI (`daily_ai_tool_usage_hrs`) hanya berada di peringkat ke-10 dengan kontribusi relatif kecil (1,06%).

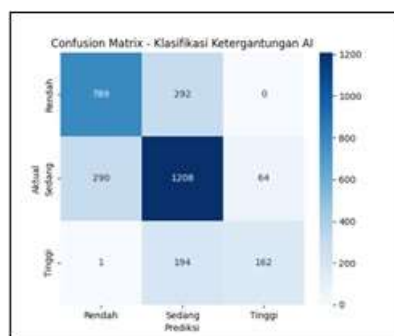
B. Deteksi Ketergantungan AI (Klasifikasi)

Model Random Forest Classifier dengan konfigurasi yang sama (150 pohon, kedalaman 15) digunakan untuk mengklasifikasikan tingkat ketergantungan AI mahasiswa ke dalam tiga kategori: Rendah, Sedang, dan Tinggi. Performa model disajikan pada Tabel 2

Kategori	Presisi	Recall	F1-score	Support
Rendah	0.73	0.73	0.73	1081
Sedang	0.71	0.77	0.74	1562
Tinggi	0.72	0.45	0.56	357
Akurasi keseluruhan	0.72 (72.0%)			
Macro avg F1-score	0.68			
Weighted avg F1-score	0.72			

Gambar 3. Performa Klasifikasi Ketergantungan AI

Model mencapai akurasi keseluruhan sebesar 72,0%. Performa tertinggi dicapai pada kategori Sedang (recall 0,77), sementara kategori Tinggi memiliki recall rendah (0,45) yang mengindikasikan bahwa model cukup sering gagal mengidentifikasi mahasiswa dengan ketergantungan AI tinggi (hanya 45% yang terdeteksi dengan benar).



Gambar 4. Confusion matrix

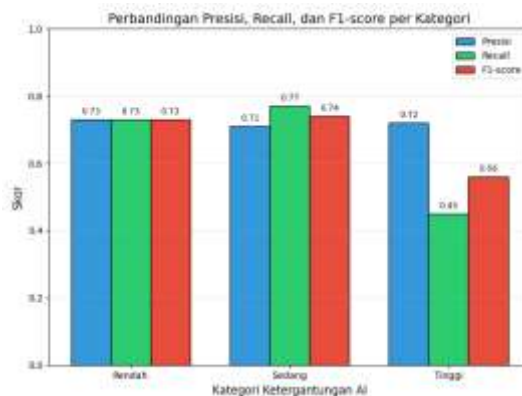
Gambar 4 menampilkan confusion matrix yang digunakan untuk mengevaluasi kinerja model dalam mengklasifikasikan tingkat ketergantungan AI ke dalam tiga kategori: Rendah, Sedang, dan Tinggi. Angka-angka di sepanjang garis diagonal utama menunjukkan jumlah prediksi yang tepat, di mana model terbukti paling dapat dipercaya dalam menebak kategori "Sedang" (1.208 data benar) dan kategori "Rendah" (789 data benar). Kendati demikian, model ini tampak masih sering terkecoh pada kategori "Tinggi", karena dari total data aktual yang masuk skala tinggi, justru lebih banyak yang salah ditebak sebagai tingkat "Sedang" (194 kasus) dibandingkan yang berhasil diidentifikasi dengan benar (162 kasus).

Rendahnya recall pada kategori Tinggi ini kemungkinan besar disebabkan oleh ketidakseimbangan kelas (class imbalance), hal ini diakui secara jelas. Untuk penelitian di masa mendatang, disarankan untuk menerapkan teknik penanganan ketidakseimbangan data seperti SMOTE (Synthetic Minority Over-sampling Technique) atau penyesuaian class weight pada algoritma Random Forest. Kedua pendekatan tersebut diharapkan dapat meningkatkan sensitivitas model terhadap kelas minoritas (Tinggi) sehingga nilai recall-nya meningkat tanpa mengorbankan akurasi secara signifikan.

C. Uji Statistik Tambahan

Hasil uji ANOVA menunjukkan bahwa tidak terdapat perbedaan signifikan rata-rata skor ketergantungan AI antar kelompok tingkat perguruan tinggi (college tier) ($F = 0,48$, $p = 0,6191 > 0,05$). Demikian pula uji Chi-square menunjukkan tidak ada hubungan signifikan antara jenis kelamin (gender) dengan kategori ketergantungan AI ($\chi^2 = 4,85$, $p = 0,3034 > 0,05$).

Berdasarkan prediksi model dan aturan sederhana (durasi penggunaan AI >3 jam, frekuensi penggunaan AI untuk tugas "Frequently" atau "Always", dan skor ketergantungan >7), sistem menghasilkan rekomendasi personalisasi. Sebanyak 4.421 mahasiswa (29,5% dari total 15.000) direkomendasikan untuk mengurangi penggunaan AI. Rekomendasi utama adalah "Kurangi waktu penggunaan harian" terutama bagi mahasiswa yang menggunakan AI lebih dari 3 jam per hari.



Gambar 5. Bar Chart Metrik

Temuan bahwa durasi penggunaan AI harian merupakan prediktor terkuat ketergantungan AI (16,5%) sejalan dengan hasil studi [17] yang menemukan bahwa frekuensi penggunaan AI berkorelasi sedang dengan penurunan kemampuan berpikir kritis. Namun berbeda dengan studi tersebut, penelitian kami menunjukkan bahwa durasi penggunaan dalam jam lebih berpengaruh dibandingkan frekuensi penggunaan. [18] Menyatakan bahwa ketergantungan AI diperkuat oleh melemahnya kebiasaan intelektual, temuan kami mendukung hal tersebut karena variabel "AI menggantikan pemikiran sendiri" memberikan kontribusi 13,4% terhadap klasifikasi ketergantungan AI. Dibandingkan dengan [19] yang menemukan bahwa AI bukan penyebab utama kecemasan kerja, penelitian kami memperkuat temuan tersebut dengan bukti kuantitatif bahwa kontribusi durasi penggunaan AI terhadap kecemasan penempatan kerja hanya 1,06%.

Metode feature importance yang digunakan dalam penelitian ini sudah cukup memadai untuk mengidentifikasi faktor-faktor dominan sebagaimana banyak digunakan dalam studi machine learning serupa [14],[15]. Meskipun demikian, untuk penelitian lanjutan yang ingin menggali lebih dalam arah hubungan antar variabel, disarankan untuk menambahkan metode SHAP (SHapley Additive exPlanations) yang dapat memberikan interpretasi lebih transparan mengenai kontribusi positif atau negatif dari setiap fitur.

Kesimpulan

Berdasarkan hasil analisis menggunakan algoritma Random Forest pada dataset 15.000 mahasiswa, penelitian ini berhasil mencapai tiga tujuan utama dengan temuan yang konstruktif. Pertama, model Random Forest Regressor mampu memprediksi tingkat kecemasan penempatan kerja mahasiswa dengan kinerja sangat baik ($R^2 = 0,815$). Faktor paling dominan adalah kecemasan wawancara (59,3%), kejelasan karier (11,8%), dan durasi belajar harian (8,4%). Menariknya, durasi penggunaan AI harian memberikan kontribusi sangat kecil (1,06%) terhadap kecemasan kerja. Ini berarti AI bukanlah penyebab utama kecemasan karier, kecemasan lebih bersumber dari faktor internal seperti persepsi kompetensi diri dan ketidakpastian masa depan. Kedua, model Random Forest Classifier mampu mengklasifikasikan tingkat ketergantungan AI dengan akurasi 72,0%. Fitur terpenting adalah durasi penggunaan AI harian (16,5%), skor AI menggantikan pemikiran sendiri (13,4%), dan penggunaan AI untuk tugas (8,2%). Temuan ini tidak berarti AI harus dihindari, melainkan mengindikasikan bahwa pola penggunaan yang bersifat substitutif (menggantikan proses berpikir) perlu dikelola agar tidak berlebihan. Ketergantungan AI tidak berkorelasi dengan jenis kelamin atau tingkat perguruan

tinggi, sehingga pengelolaannya bersifat universal. Ketiga, sistem rekomendasi intervensi yang dipersonalisasi berhasil dihasilkan. Sebanyak 4.421 mahasiswa (29,5%) direkomendasikan untuk mengurangi durasi penggunaan AI yang berlebihan (>3 jam/hari), bukan berhenti menggunakan AI. Rekomendasi utama adalah "Kurangi waktu penggunaan harian" dan "Kurangi penggunaan AI untuk tugas jika sudah menggantikan pemikiran mandiri". Intervensi ini bersifat edukatif dan preventif, bertujuan membantu mahasiswa menjadikan AI sebagai asisten belajar yang produktif, bukan sebagai pengganti proses kognitif. Secara keseluruhan, penelitian ini menegaskan bahwa AI bukanlah ancaman, namun penggunaannya perlu kesadaran dan keseimbangan kognitif dalam penggunaannya.

Deklarasi Kepentingan

Penulis menyatakan bahwa penelitian ini dilakukan secara independen dan tidak terdapat konflik kepentingan, baik yang bersifat finansial maupun non-finansial, yang dapat memengaruhi proses penelitian, analisis data, interpretasi hasil, maupun penyusunan naskah artikel ini. Penulis tidak menerima pendanaan, imbalan, atau dukungan dari pihak mana pun yang berpotensi menimbulkan bias terhadap hasil penelitian. Seluruh proses penelitian dilaksanakan berdasarkan prinsip objektivitas, integritas akademik, dan etika ilmiah. Penulis juga menyatakan bahwa tidak terdapat hubungan pribadi, profesional, atau institusional yang dapat memengaruhi penilaian ilmiah maupun kesimpulan yang disajikan dalam artikel ini. Semua hasil, temuan yang disampaikan merupakan tanggung jawab penuh penulis.

Ketersediaan Data

Data yang digunakan dalam penelitian ini merupakan dataset sekunder yang diperoleh dari repositori publik Kaggle berjudul AI Dependency, Career Anxiety and Student Burnout dan dapat diakses melalui tautan yang tercantum pada referensi [12]. Data tersedia secara terbuka untuk keperluan penelitian dan akademik sesuai ketentuan penyedia dataset.

Penggunaan Ai Dan Deklarasi Penggunaan Ai Generatif

Dalam penyusunan artikel ini, penulis menggunakan alat AI generatif, yaitu DeepSeek, untuk membantu proses parafrase kalimat, penyempurnaan tata bahasa, serta memberikan dukungan teknis dalam proses debugging kode program Python yang digunakan selama penelitian. Selain itu, penulis menggunakan DeepL sebagai alat bantu penerjemahan abstrak dari Bahasa Indonesia ke Bahasa Inggris. Penggunaan alat AI tersebut terbatas pada aspek teknis penulisan, penerjemahan, dan pemrograman, serta tidak digunakan untuk menghasilkan data penelitian, melakukan analisis ilmiah, menginterpretasikan hasil penelitian, maupun menyusun kesimpulan penelitian secara otomatis. Seluruh hasil keluaran dari alat AI telah ditinjau, dan diverifikasi kembali oleh penulis sebelum digunakan dalam naskah ini. Penulis bertanggung jawab penuh atas keakuratan, validitas, dan integritas ilmiah seluruh isi artikel yang dipublikasikan.

Daftar Pustaka

- [1] T. M. Tsabitha, "Perlindungan Hak Kekayaan Intelektual atas Karya AI: Studi Kasus ChatGPT dan DeepSeek dalam Perspektif TRIPS Agreement," *UNES Law Review*, vol. 8, no. 1, pp. 192–204, Sep. 2025, doi: 10.31933/unesrev.v8i1.2472.
- [2] M. Al Mashagbeh, M. Alsharqawi, U. Tudevdagva, and H. J. Khasawneh, "Student engagement with artificial intelligence tools in academia: a survey of Jordanian universities," *Front. Educ.*, vol. 10, Aug. 2025, doi: 10.3389/feduc.2025.1550147.
- [3] S. Musliha, B. Bahrudin, and R. F. Diharjo, "Dinamika Pemanfaatan Artificial Intelligence (AI) dalam Proses Pembelajaran di Perguruan Tinggi (Studi Kasus Universitas Islam Zainul Hasan Genggong dan Universitas Panca Marga)," *Jurnal Teknologi Pembelajaran Indonesia*, vol. 15, no. 2, pp. 142–158, Nov. 2025, doi: 10.23887/jurnal_tp.v15i2.5929.
- [4] M. A. Ramadhan, A. Gunawan, S. Lorenza, Z. Ainy, and M. Subhan, "Analisis Dampak Penggunaan Artificial Intelligence terhadap Kemampuan Berpikir Kritis Mahasiswa," *Mudabbir: Journal Research and Education Studies*, vol. 5, no. 2, pp. 241–249, Jun. 2025, doi: 10.56832/mudabbir.v5i2.1111.
- [5] A. T. Rizkina, H. G. Hidajat, and I. A. Farida, "Job-Related Anxiety in the Age of Artificial Intelligence: A Systematic Review of Workplace Dynamics," *Formosa Journal of Multidisciplinary Research*, vol. 4, no. 8, pp. 3965–3976, Aug. 2025, doi: 10.55927/fjmr.v4i8.416.
- [6] N. B. Alotaibi, "Adoption of AI tools in Education: Examining influencing factors and their impact on students' critical thinking skills," *Computers in Human Behavior Reports*, vol. 21, p. 100990, Mar. 2026, doi: 10.1016/j.chbr.2026.100990.
- [7] B.-J. Kim and J. Lee, "AI adoption, employee depression and knowledge: How corporate social responsibility buffers psychological impact," *Journal of Innovation & Knowledge*, vol. 10, no. 6, p. 100815, Nov. 2025, doi: 10.1016/j.jik.2025.100815.
- [8] L. Zhang and J. Xu, "The paradox of self-efficacy and technological dependence: Unraveling generative AI's impact on university students' task completion," *The Internet and Higher Education*, vol. 65, p. 100978, Apr. 2025, doi: 10.1016/j.iheduc.2024.100978.

- [9] W. Lo *et al.*, “Machine learning based prediction of high school student mental health,” Aug. 11, 2025, *medRxiv*. doi: 10.1101/2025.08.07.25333218.
- [10] M. Fadhilla, R. Wandri, A. Hanafiah, P. R. Setiawan, Y. Arta, and S. Daulay, “Analisis Performa Algoritma Machine Learning Untuk Identifikasi Depresi Pada Mahasiswa,” *Journal of Informatics Management and Information Technology*, vol. 5, no. 1, pp. 40–47, Jan. 2025.
- [11] C. Eirene, D. Syafitri, N. Sulistianingsih, K. Hidjah, and H. 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.
- [12] “AI Dependency, Career Anxiety and Student Burnout.” Accessed: May 12, 2026. [Online]. Available: <https://www.kaggle.com/datasets/sridipbasu/ai-dependency-career-anxiety-and-student-burnout>
- [13] A. Rifai, S. E. Permana, and R. Hamonangan, “OPTIMALISASI KLASIFIKASI INDEKS DESA MEMBANGUN MENGGUNAKAN METODE ENSEMBLE DAN ALGORITMA RANDOM FOREST,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 4, pp. 8226–8234, Aug. 2024, doi: 10.36040/jati.v8i4.10971.
- [14] W. A. Rayadhani and M. Rahardi, “Comparative Analysis of Random Forest, SVM, and Naive Bayes for Cardiovascular Disease Prediction,” *Journal of Applied Informatics and Computing*, vol. 9, no. 6, pp. 3234–3243, Dec. 2025, doi: 10.30871/jaic.v9i6.11451.
- [15] “Klasifikasi Halaman SEO Berbasis Machine Learning Melalui Mutual Information dan Random Forest Feature Importance | NURADILLA | MIND (Multimedia Artificial Intelligent Networking Database) Journal.” Accessed: May 17, 2026. [Online]. Available: <https://ejurnal.itenas.ac.id/index.php/mindjournal/article/view/13841>
- [16] M. Mujahid *et al.*, “Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering,” *J Big Data*, vol. 11, no. 1, p. 87, Jun. 2024, doi: 10.1186/s40537-024-00943-4.
- [17] R. M. Arif, R. Nisa, and Z. S. Allah, “THE PARADOX OF ARTIFICIAL INTELLIGENCE USAGE IN HIGHER EDUCATION: BETWEEN ENHANCING LEARNING EFFECTIVENESS AND THE DECLINE OF STUDENTS CRITICAL THINKING SKILLS,” *Tadbir: Jurnal Manajemen Pendidikan Islam*, vol. 13, no. 2, pp. 472–489, Aug. 2025, doi: 10.30603/tjmpi.v13i2.6803.
- [18] J. Tian and R. Zhang, “Learners’ AI dependence and critical thinking: The psychological mechanism of fatigue and the social buffering role of AI literacy,” *Acta Psychologica*, vol. 260, p. 105725, Oct. 2025, doi: 10.1016/j.actpsy.2025.105725.
- [19] E. Dağ, M. Nal, İ. Topuz, H. Kılınç, Y. Demir, and G. Bektaş, “Artificial intelligence perceptions and career anxiety among health sciences students,” *Sci Rep*, vol. 16, no. 1, p. 14781, Mar. 2026, doi: 10.1038/s41598-026-42030-z.