

Klasifikasi Tujuan Penggunaan AI Oleh Mahasiswa Dengan Algoritma *K-Nearest Neighbor*

DOI: <http://dx.doi.org/10.35889/jutisi.v14i3.3348>

Creative Commons License 4.0 (CC BY – NC)



Sabilatul Isti Karoh^{1*}, Anita Fira Waluyo²

Informatika, Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia

*e-mail Corresponding Author: Sabilatulisti23@gmail.com

Abstract

Artificial intelligence (AI) technology is developing rapidly, particularly through platforms like ChatGPT and Gemini AI, which are widely used by students for various purposes. This fact highlights the importance of a more in-depth analysis of students' AI usage patterns. This study aims to classify the purposes of AI use by students using the K-Nearest Neighbor (K-NN) algorithm. The study was conducted using two classification approaches: multi-class classification and binary classification. Data were obtained from questionnaires with 305 respondents and processed using Python on the Google Colab platform using preprocessing, normalization, and data encoding stages. Test results show that the K-NN algorithm achieved a high accuracy of 77% in the binary classification scenario (Productive/Career and Entertainment/Personal), while in the multi-class classification scenario the highest accuracy only reached 34%. This finding indicates that K-NN performance is strongly influenced by the complexity of the number of classes and is more optimally applied to classifications with a limited number of classes.

Keywords: *K-Nearest Neighbor; AI usage classification; students; Binary Classification; Multi-class Classification.*

Abstrak

Teknologi kecerdasan buatan ataupun *Artificial Intelligence* (AI) berkembang pesat, terutama melalui platform seperti ChatGPT dan Gemini AI yang banyak digunakan oleh mahasiswa untuk berbagai tujuan. Fakta ini menunjukkan pentingnya analisis yang lebih mendalam mengenai pola penggunaan AI oleh mahasiswa. Penelitian ini bertujuan untuk mengklasifikasikan tujuan penggunaan AI oleh mahasiswa memakai algoritma *K-Nearest Neighbor* (K-NN). Penelitian dilakukan melalui dua pendekatan klasifikasi, yaitu *multi-class classification* dan *binary classification*. Data didapatkan dari kuesioner terhadap 305 responden dan diolah memakai Python pada platform Google Colab menggunakan tahap pra-pemrosesan, normalisasi, dan pengkodean data. Hasil pengujian menunjukkan algoritma K-NN mencapai akurasi tinggi yaitu 77% pada skenario klasifikasi biner (Produktif/Karir dan Hiburan/Personal), sementara pada skenario klasifikasi multi-kelas akurasi tertinggi hanya mencapai 34%. Temuan ini menandakan bahwa kinerja K-NN sangat dipengaruhi oleh kompleksitas jumlah kelas dan lebih optimal diterapkan pada klasifikasi dengan jumlah kelas yang terbatas.

Kata kunci: *K-Nearest Neighbor; Klasifikasi Penggunaan AI; Mahasiswa; Klasifikasi Biner; Klasifikasi Multi-Kelas .*

1. Pendahuluan

Teknologi kecerdasan buatan berkembang sangat cepat, terutama berkat munculnya platform seperti ChatGPT dan Gemini AI. Mahasiswa kerap sekali menggunakan AI untuk berbagai tujuan, baik untuk keperluan akademik, hiburan, pekerjaan, bahkan mengembangkan diri. Menurut survei ada 2.555 mahasiswa, sebagian besar responen menyatakan menggunakan ChatGPT untuk keperluan akademik[1]. Fakta ini membuktikan bahwa dibutuhkan analisis mendalam mengenai pola penggunaan AI oleh mahasiswa. Klasifikasi tujuan penggunaan AI

mampu membantu perancang mengetahui kebutuhan mahasiswa dalam penggunaan AI dan meningkatkan kualitas platform AI.

Berbagai teknologi kecerdasan buatan AI untuk pendidikan tinggi, masih dalam proses pengembangan. AI generatif, ChatGPT, sudah tergolong banyak digunakan mahasiswa untuk keperluan perkuliahan dan pembantuan dalam penulisan ilmiah. Para peneliti menemukan bahwa pengguna ChatGPT meningkatkan keterlibatan dalam menulis sebesar 25,3% dan daya cipta sebesar 8,10%. AI menyediakan berbagai saran, ide, dan bahkan koreksi, yang mampu membantu mahasiswa untuk mengembangkan dan mempertajam tulisan yang mereka kerjakan[2][3]. Meskipun demikian, penggunaan yang berlebihan berisiko mengurangi tanggapan dari sumber-sumber terpercaya dan menurunkan daya cipta mahasiswa.

Beberapa penelitian terdahulu telah memanfaatkan algoritma klasifikasi untuk mengelompokkan perilaku pengguna berdasarkan data digital. Selain digital, algoritma K-Nearest Neighbor (K-NN) dapat mengidentifikasi kematangan buah nanas berdasarkan warna gambar[4][5]. Sistem klasifikasi citra Eurosat yang dikembangkan menggunakan metode Model mencapai akurasi 81,32%, presisi seimbang, recall, dan nilai f1-score di semua kelas[6][7][8]. Hasil penelitian tersebut membuktikan bahwa K-NN mampu mencapai akurasi tinggi dalam mengelompokkan data visual, sehingga menunjukkan fleksibilitas algoritma ini dalam berbagai bidang klasifikasi, baik berbasis teks, perilaku, maupun citra digital.

Penelitian sebelumnya lebih berfokus pada pengklasifikasian objek fisik dan sistem rekomendasi, alih-alih pada aspek perilaku penggunaan teknologi digital dalam pendidikan. Penelitian untuk mengklasifikasikan tujuan penggunaan AI oleh mahasiswa berdasarkan pola perilaku mereka memakai algoritma *K-Nearest Neighbor* (K-NN). Dengan dasar ini, Penelitian ini diharapkan dapat memberikan wawasan lebih mendalam mengenai preferensi mahasiswa dalam menggunakan AI dan membantu dalam pengembangan platform AI yang lebih relevan dengan kebutuhan pengguna.

2. Tinjauan Pustaka

Penelitian-penelitian sebelumnya yang berkaitan dengan penelitian ini dijelaskan di bawah ini. Penelitian ini mengutip makalah-makalah berikut:

Fathon et al.[9], pada penelitian “Komparasi kinerja KNN dan Naïve Bayes untuk klasifikasi sentimen pemakai aplikasi deepseek AI”. Penelitian ini menggunakan data ulasan pengguna berbahasa Inggris di Google Play Store, yang melalui tahapan *preprocessing* teks dan ekstrasi fitur menggunakan TF-IDF. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-Score* dengan pembagian data 70:30 dan 80:20 serta validasi *Stratified K-Fold Cross Validation*. Hasil membuktikan bahwa K-NN memberikan hasil yang lebih stabil dengan rata-rata akurasi mencapai 80,78%, sedikit lebih tinggi dibandingkan Naïve Bayes 80,13%.

Fitriani et al.[10], dalam penelitian “Perbandingan Kinerja metode problem transformation KNN dan Algoritma Adaptation-KNN pada klasifikasi Multi-Label Pertanyaan Kotakode”. Penelitian tersebut menggunakan variabel yang digunakan berupa penyajian teks pertanyaan yang dikelompokkan ke dalam lebih dari satu label topik. Hasil memperlihatkan metode *Label Powerset* berbasis K-NN mencapai akurasi tertinggi, yaitu sekitar 86%, yang membuktikan keberhasilan K-NN dalam mengidentifikasi masalah klasifikasi multi-label dengan kesulitan kelas yang tinggi.

Ryan Situmorang et al.[11], dengan judul “Model algoritma K-Nearest Neighbor(K-NN) dan Naïve Bayes” untuk Prediksi Kelulusan Mahasiswa berdasarkan akademik. Dalam penelitian ini menggunakan variabel yang mengkaji nilai akademik dan status kelulusan yang dikelompokkan secara biner. Hasilnya menunjukkan bahwa K-NN mampu memberikan prediksi yang lebih tepat dengan akurasi 100%, sedangkan Naïve Bayes sedikit lebih rendah pada 99,29%. Penelitian ini membuktikan keunggulan K-NN dalam menangani data akademik dan prediksi berbasis biner.

Almi Yulistia Alwanda et al.[12] dengan judul “Analisis Klasifikasi Konsentrasi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor”. Menunjukkan pemanfaatan variabel jenis kelamin, IPK, nilai mata kuliah dengan hasil penelitian yang memperlihatkan bahwa implementasi K-NN dengan nilai $K=2$ dan metode 8-fold cross-validation menghasilkan akurasi sebesar 67%.

Nurul Renanintias et al.[13], dalam studi “Studi Komparasi Algoritma Decision Tree C4.5 dan K-Nearest Neighbor pada klasifikasi Masa Studi dan Tingkat Stres Mahasiswa”. Penelitian ini membandingkan algoritma Decision Tree C4.5 dan K-Nearest Neighbor(K-NN) berdasarkan

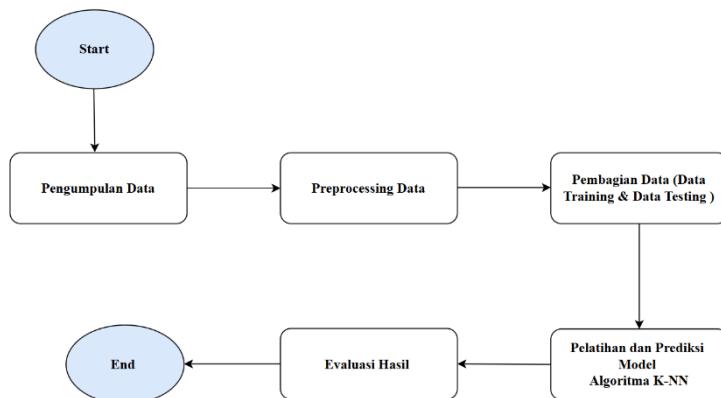
variabel akademik dan aktivitas mahasiswa. Hasil menunjukkan bahwa K-NN memiliki kinerja lebih baik dibandingkan C4.5 dengan akurasi 71,44%.

Nurul Chasanah et al.[14], pada penelitian yang berjudul “Klasifikasi Kelayakan Siswa dalam Menentukan Kelas Unggulan Menggunakan Algoritma K-Nearest Neighbor”. Penelitian tersebut menggunakan 505 data nilai siswa kelas XI dengan variabel berupa nilai Bahasa Indonesia, Matematika, Bahasa Inggris dan mata pelajaran kejuruan sebagai aktifitas prediktor, serta kelayakan siswa(Layak dan Tidak Layak) sebagai label. Hasil memperlihatkan bahwasanya akurasi tertinggi didapatkan pada nilai $K = 25$, dengan akurasi 97,00% menggunakan Google Colab, serta 95,05% menggunakan RapidMiner.

Beberapa studi sebelumnya telah menunjukkan seberapa efektif Algoritma K-Nearest Neighbor (K-NN) dalam menyelesaikan berbagai tugas klasifikasi seperti analisis sentimen pengguna AI, klasifikasi teks multi-label, prediksi akademik, dan klasifikasi siswa unggul. Kebaruan dari penelitian ini terdapat pada dua aspek utama, yaitu variabel yang dianalisis dan pendekatan klasifikasi yang digunakan. Dalam hal variabel, penelitian ini mengklasifikasi tujuan penggunaan AI oleh mahasiswa pada dimensi multi-perilaku akademik, hiburan, pekerjaan, dan pengembangan diri, yang belum banyak diteliti secara khusus menggunakan K-NN. Dalam hal pendekatan, penelitian ini menerapkan dan membandingkan kinerja K-NN pada dua skema klasifikasi, yaitu klasifikasi *multi-class classification* dan *binary classification* (produktivitas/karir dan non-produktivitas), oleh karena itu memberikan berkontribusi baru dalam menetapkan pendekatan klasifikasi optimal yang akan digunakan dalam menganalisis perilaku pengguna AI dalam lingkup pendidikan.

3. Metodologi

Penelitian memakai pendekatan kuantitatif dengan penelitian untuk melakukan klasifikasi tujuan penggunaan AI oleh mahasiswa. Proses dimulai dengan pengumpulan data, kemudian dilakukan pra-pemrosesan, pembagian data, dan pelatihan dan prediksi model menggunakan algoritma K-NN, serta evaluasi kinerja model. Berikut adalah flowchart penelitian.



Gambar 1. Metode Penelitian

3.1. Pengumpulan Data

Data dikumpulkan melalui kuesioner pada mahasiswa selama periode Mei hingga Juli 2025, dengan total 305 responden. Variabel yang dikumpulkan meliputi informasi demografi serta respons terhadap pertanyaan yang diukur menggunakan skala Likert, yang menjelaskan tujuan penggunaan AI oleh Mahasiswa. Data yang terkumpul kemudian diekspor dalam format CSV memudahkan proses lebih lanjut.

3.2. Preprocessing Data

Tahap ini untuk menjamin kualitas data sebelum pada proses pelatihan model. Tahapan pembersihan data dimana menghapus nilai yang mengandung *missing values* dan menghapus data duplikat, pengkodean data kategorikal memakai teknik one-hot encoding, serta normalisasi data numerik menggunakan metode Min–Max Scaling. Selain itu, dilakukan seleksi fitur untuk menghilangkan variabel yang tidak relevan terhadap target klasifikasi.

3.3. Pembagian Data

Data dengan tahap pra-pemrosesan selanjutnya dibagi menjadi dua bagian ialah data latih sebesar 80% dan data uji sebesar 20%. Untuk menangani ketidakseimbangan kelas pada variabel target, metode SMOTE (*Synthetic Minority Oversampling Technique*) diterapkan hanya data latih, dengan tujuan menyeimbangkan distribusi kelas tanpa memengaruhi integritas data uji.

3.4. Pelatihan dan Prediksi Model K-NN

Penelitian ini menggunakan algoritma K-Nearest Neighbor (K-NN) untuk mengklasifikasikan data berdasarkan kedekatan karena efisiensi dan efektivitasnya. K-NN mengklasifikasikan data baru menggunakan jumlah kelas tertinggi dari K tetangga terdekat dalam ruang fitur [11]. Perhitungan K-NN menggunakan Rumus Jarak Euclidean:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad \dots \dots \dots \quad (1)$$

Keterangan :

$d(x,y)$ = jarak euclidean antara dua objek x dan y

n = jumlah fitur (dimensi) yang digunakan

$(x_i - y_i)$ = komponen ke- i dari objek x dan y

Dalam penelitian ini, algoritma K-NN diterapkan pada dua skenario klasifikasi, yaitu multi-class classification untuk memprediksi empat tujuan penggunaan AI (akademik, hiburan, pekerjaan, dan pengembangan diri), serta binary classification untuk mengelompokkan tujuan penggunaan ke dalam kategori produktif/karier dan non-produktif. Kinerja model dievaluasi menggunakan metrik akurasi dan confusion matrix.

3.6. Evaluasi Hasil

Untuk menilai model klasifikasi, *confusion matrix* digunakan pada data uji. Ukuran penilaianya adalah akurasi, presisi, recall, dan skor F1. Kemampuan algoritma K-NN untuk mengidentifikasi tujuan penggunaan AI siswa dievaluasi menggunakan keempat ukuran ini.[15].

4. Hasil dan Pembahasan

Bab ini menyajikan hasil eksperimen dan pembahasan penerapan algoritma K-NN dalam mengklasifikasikan tujuan penggunaan AI oleh mahasiswa. Pembahasan difokuskan pada hasil setiap tahapan penelitian, mulai dari pengumpulan data, pra-pemrosesan, pelatihan model, hingga evaluasi hasil pengujian.

4.1 Hasil Pengumpulan Data

Data penelitian dari 305 responden mahasiswa yang mengisi kuesioner online pada periode Mei hingga Juli 2025. Dataset yang dihasilkan memuat informasi demografis serta tujuan penggunaan AI terbagi pada empat kategori, yaitu akademik, hiburan, pekerjaan, dan pengembangan diri. Contoh struktur data penelitian ini pada Gambar 2.

Gambar 2. Dataset Penelitian

4.2. Hasil Preprocessing Data

Untuk kategorisasi, 305 tanggapan kuesioner diproses terlebih dahulu. Tujuan utama penggunaan AI oleh responden dikategorikan sebagai akademis, hiburan, pekerjaan, dan pengembangan diri. Perubahan ini mengatasi variasi respons dan menetapkan label target kategorisasi yang konsisten.

```
Distribusi kelas target yang dipetakan:
Menurut Anda, apa tujuan utama Anda menggunakan AI?
Akademik      81
Pengembangan Diri  78
Pekerjaan      76
Hiburan        70
Name: count, dtype: int64
```

Gambar 3. Kelas Target

Hasil pemetaan menunjukkan pada Gambar 3. Menampilkan distribusi kelas target yang relatif seimbang. Kelas akademik memiliki jumlah data terbanyak dengan 81 responden, diikuti oleh pengembangan diri sebanyak 78 responden, pekerjaan sebanyak 76 responden, dan hiburan sebanyak 70 responden. Distribusi ini menunjukkan bahwa data tidak mengalami ketimpangan kelas yang signifikan sehingga layak digunakan dalam proses pelatihan model klasifikasi.

Selain pemetaan label, fitur numerik yang diukur menggunakan skala Likert dinormalisasi ke dalam rentang nilai 0–1 menggunakan metode *Min–Max Scaling*. Sementara itu, fitur kategorikal dikonversi ke bentuk numerik menggunakan teknik *one-hot encoding*. Hasil pra-pemrosesan menghasilkan dataset akhir yang seluruh fiturnya bersifat numerik dengan total 20 fitur dari 305 data, sehingga siap digunakan pada tahap pelatihan dan pengujian algoritma K-NN

```
Shape data setelah preprocessing: (305, 20)
```

Cek rentang hasil normalisasi (fitur numerik):

	min	max
num_Saya menggunakan AI untuk Menggerjakan tugas...	0.0	1.0
num_Saya menggunakan AI Untuk penelitian atau ...	0.0	1.0
num_Saya Menggunakan AI Untuk membantu Manajem...	0.0	1.0
num_Saya Menggunakan AI untuk Hiburan (contoh ...	0.0	1.0
num_Saya Menggunakan AI Untuk berinteraksi sec...	0.0	1.0
num_Apakah Anda setuju bahwa AI membantu efisi...	0.0	1.0
num_Saya merasa AI mempermudah tugas saya seba...	0.0	1.0
num_Saya lebih produktif saat menggunakan AI	0.0	1.0

Gambar 4. Hasil Normalisasi dan Encoding

4.3 Pembagian data

Setelah pra-pemrosesan, dataset dibagi menjadi 80% data pelatihan dan 20% data pengujian. Dari 305 dataset, dikumpulkan 244 data pelatihan dan 61 data pengujian. Data dibagi dengan mempertahankan persentase kelas yang diinginkan di setiap subset untuk menyeimbangkan distribusi kelas untuk pelatihan dan pengujian model klasifikasi.

```
X_train, X_test, y_train, y_test = train_test_split(X_processed, y_encoded, test_size=0.2, random_state=42, stratify=y_encoded)
print(f"\nData training: {X_train.shape[0]} | Data testing: {X_test.shape[0]}")
```

```
Data training: 244 | Data testing: 61
```

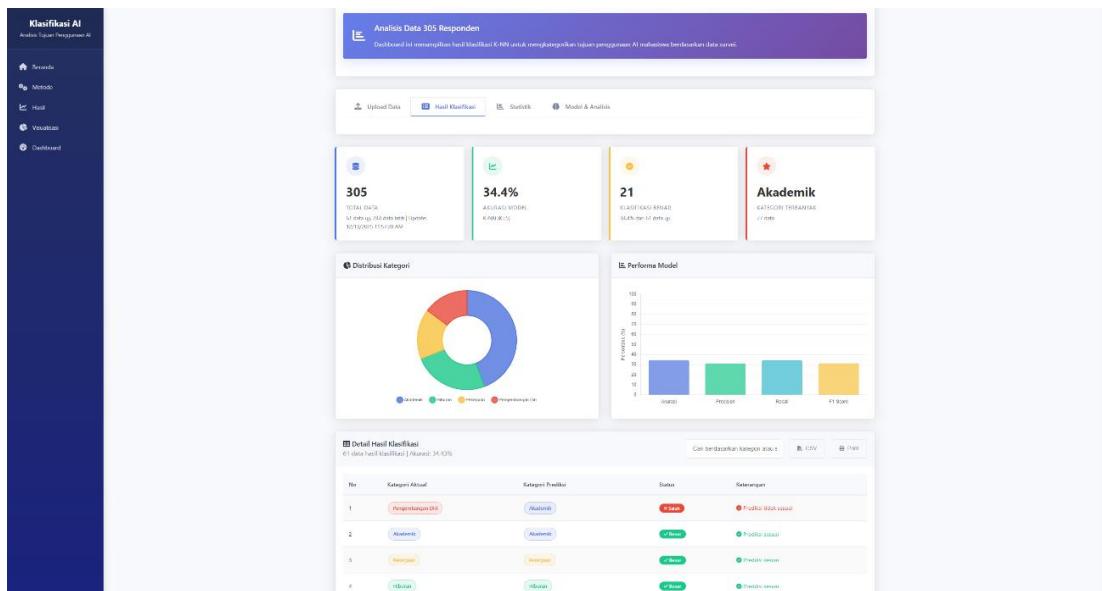
Gambar 5. Membagi dataset

4.4 Hasil Pelatihan dan Prediksi Model K-NN

Data yang telah diproses sebelumnya diolah dan dievaluasi menggunakan K-NN. Data dibagi menjadi 80% data pelatihan dan 20% data pengujian dengan nilai parameter *K* sebesar 3, 4, 5, dan 6. *Multi-Class Classification* dan *Binary Classification* digunakan untuk menguji model. Pada skenario multi-class, data diklasifikasikan ke dalam empat kategori tujuan penggunaan AI,

yaitu Akademik, Hiburan, Pekerjaan, dan Mengembangkan Diri. Sementara itu, pada skenario binary classification, kelas dikelompokkan menjadi dua kategori utama, yaitu Produktif/Karir (0) yang mencakup Akademik, Pekerjaan, dan Pengembangan Diri, serta Hiburan/Personal (1) yang mencakup penggunaan AI untuk kepentingan hiburan.

Sebelum melakukan perhitungan akurasi model secara keseluruhan, hasil Klasifikasi tujuan penggunaan AI diklasifikasikan terlebih dahulu ke dalam sebuah dasbor sistem berikut:



Gambar 6. Tampilan dasbor web

Gambar 6. menampilkan Visualisasi hasil klasifikasi tujuan penggunaan AI oleh mahasiswa. dalam sebuah dasbor web berbasis algoritma K-NN. Dasbor ini menyajikan pemetaan data responden ke dalam kategori penggunaan AI secara visual melalui grafik dan tabel, sehingga distribusi kelas dan hasil klasifikasi dapat diamati dengan jelas. Tampilan ini menekankan hasil akhir pengelompokan data, sehingga memudahkan pemahaman pola penggunaan AI oleh mahasiswa dalam format yang terstruktur dan mudah dipahami.

Sesuai dengan tujuan penelitian untuk mengklasifikasikan tujuan penggunaan AI, berikut adalah penyajian data hasil proses yang membandingkan kategori aktual dengan hasil prediksi model pada setiap skenario pengujian:

1) Multi-Class Classification

Pada skenario ini mengklasifikasikan data ke dalam empat label: Akademik, Hiburan, Pekerjaan, dan Pengembangan diri. Sebelum mengevaluasi performa secara keseluruhan, Tabel 1. Menyajikan detail hasil Prediksi per baris data:

Tabel 1. Sampel Detail Hasil Prediksi *Multi-Class Classification*

No.	Kategori Aktual	Kategori Prediksi	Status
1	Pengembangan Diri	Akademik	Salah
2	Akademik	Akademik	Benar
3	Pekerjaan	Pekerjaan	Benar
4	Hiburan	Hiburan	Benar
5	Akademik	Akademik	Benar
6	Hiburan	Pengembangan Diri	Salah
7	Akademik	Pekerjaan	Salah
8	Hiburan	Pengembangan Diri	Salah
9	Akademik	Akademik	Benar
10	Pengembangan Diri	Akademik	Salah

Berdasarkan detail hasil proses tersebut, diperoleh rangkuman akurasi untuk pengujian *Multi-Class Classification* sebagai berikut:

Tabel 2. Hasil Pengujian Akurasi *Multi-Class Classification*

Nilai K	Akurasi
3	34%
4	30%
5	33%
6	22%

Berdasarkan hasil pengujian *multi-class classification* yang disajikan pada Tabel 1. model K-NN menunjukkan tingkat akurasi yang relatif rendah, dengan akurasi tertinggi 34% pada nilai K = 3. Penurunan akurasi terlihat pada nilai K yang lebih besar, khususnya pada K = 6. Hasil ini menunjukkan bahwa klasifikasi dengan empat kelas tujuan penggunaan AI memiliki tingkat kompleksitas yang lebih tinggi. Hal tersebut disebabkan oleh adanya kemiripan pola karakteristik antar kelas, terutama antara kelas Akademik, Pekerjaan, dan Pengembangan Diri, sehingga menyulitkan model K-NN dalam membedakan masing-masing kelas secara akurat.

2) *Binary Classification*

Skenario kedua menyederhanakan target menjadikan dua label: Produktif/Karir dan Hiburan/Personal. Berikut adalah detail hasil prediksi:

Tabel 3. Sampel Detail Hasil Pprediksi *Binary Classification*

No.	Kategori Aktual	Kategori Prediksi	Status
1	Produktif/Karir(0)	Produktif/Karir(0)	Benar
2	Produktif/Karir(0)	Produktif/Karir(0)	Benar
3	Produktif/Karir(0)	Produktif/Karir(0)	Benar
4	Produktif/Karir(0)	Produktif/Karir(0)	Benar
5	Produktif/Karir(0)	Produktif/Karir(0)	Benar
6	Produktif/Karir(0)	Hiburan/Personal(1)	Salah
7	Produktif/Karir(0)	Produktif/Karir(0)	Benar
8	Produktif/Karir(0)	Produktif/Karir(0)	Benar
9	Produktif/Karir(0)	Produktif/Karir(0)	Benar
10	Hiburan/Personal(1)	Produktif/Karir(0)	Benar

Berdasarkan detail hasil proses tersebut, diperoleh rangkuman akurasi untuk pengujian *Binary Classification* sebagai berikut:

Tabel 4. Hasil Pengujian Akurasi *Binary Classification*

Nilai K	Akurasi
3	71%
4	77%
5	74%
6	77%

Berbeda dengan hasil *multi-class classification*, pengujian pada skenario *binary classification* menunjukkan peningkatan akurasi yang signifikan, sebagaimana ditampilkan pada Tabel 4. Akurasi tertinggi mencapai 77% pada nilai K = 4 dan K = 6.

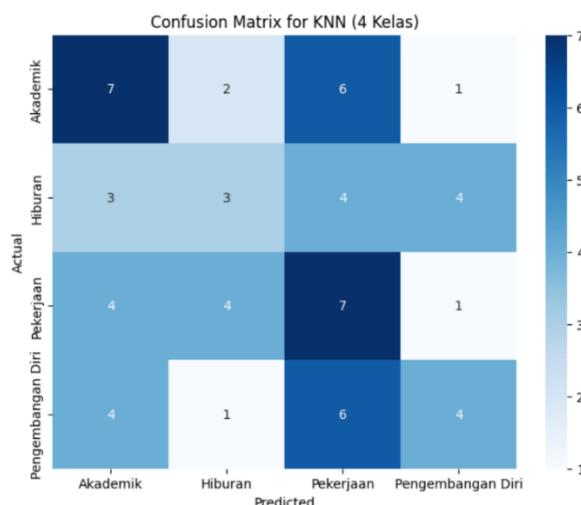
Peningkatan kinerja ini menunjukkan bahwa penyederhanaan kelas menjadi dua kategori utama mampu mengurangi kompleksitas klasifikasi dan meningkatkan kemampuan model dalam mengenali pola penggunaan AI. Dengan demikian, pendekatan *binary classification* dinilai lebih efektif dalam merepresentasikan perilaku penggunaan AI mahasiswa dibandingkan dengan pendekatan *multi-class classification*.

4.5 Hasil Evaluasi Model

Berdasarkan hasil evaluasi, terlihat bahwa model K-NN menunjukkan kinerja yang lebih unggul dalam skenario *Binary Classification* jika dibandingkan dengan *Multi-Class Classification*. Dengan tingkat akurasi yang lebih tinggi serta metrik evaluasi yang lebih memuaskan, pendekatan klasifikasi dua kelas terbukti lebih cocok untuk algoritma ini. Sedangkan, pada *Multi-Class Classification*, kinerja model cenderung menurun, yang kemungkinan besar dipengaruhi oleh adanya ketidakseimbangan distribusi data atau kemiripan fitur antar kelas yang membuat model kesulitan memisahkannya.

Tabel 5. Hasil Klasifikasi K-NN

Metrik	Binary	Multi-Class
Akurasi	77%	34%
Presisi	78%	41%
Recal	98%	69%
F1-Score	87%	51%

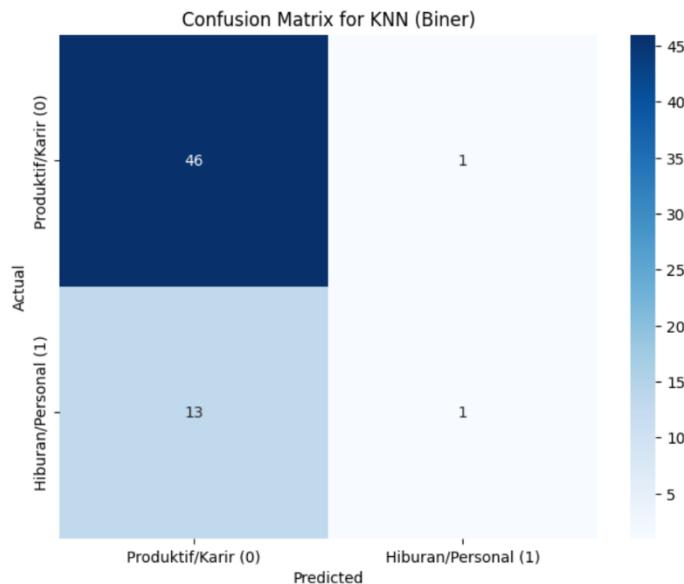


Gambar 7. Confusion Matrix K-NN Multi-Class

Pada *Binary Classification*, model K-NN memperoleh akurasi 77%, memperlihatkan keberhasilan yang lebih tinggi dalam memisahkan dua kategori utama, yaitu Produktif/Karir dan Hiburan/Personal. Dengan akurasi 78% dan recall 98%, model ini akurat dalam mendeteksi kelas Produktif/Karir (0) tetapi kesulitan dengan kelas Hiburan/Pribadi (1). Skor F1 yang tinggi (87%) menunjukkan keseimbangan presisi-recall yang kuat. Sedangkan, pada *Multi-Class Classification*, meskipun nilai recall tercatat 69%, precision hanya 41%, yang menunjukkan kesulitan model dalam membedakan kelas-kelas dengan karakteristik serupa. Akurasi yang rendah (34%) menegaskan bahwa klasifikasi dengan lebih dari dua kelas memerlukan penyesuaian lebih lanjut, seperti penanganan ketidakseimbangan data.

Gambar 7. merupakan *Confusion Matrix* untuk K-NN (4 Kelas), yang mendeskripsikan kesulitan model dalam membedakan kelas Hiburan dan Pengembangan Diri, yang memiliki

kesamaan karakteristik. Hal ini menunjukkan bahwa model kesulitan dalam memecah kelas-kelas yang memiliki ciri-ciri serupa, yang dapat hasil kinerja klasifikasi secara keseluruhan.



Gambar 8. Confusion Matrix K-NN Binary

Gambar 8. menunjukkan *Confusion Matrix* untuk K-NN (Binary), yang memperlihatkan kesuksesan model dalam mengklasifikasikan kelas Produktif/Karir (0) dengan baik, walaupun mengalami kendala dalam mengidentifikasi kelas Hiburan/Personal (1). Model ini lebih cenderung mengklasifikasikan data sebagai Produktif/Karir, yang menunjukkan bias kelas dan kesulitan dalam mendeteksi Hiburan/Personal.

4.6 Pembahasan

Hasil penelitian menunjukkan bahwa Kinerja algoritma K-Nearest Neighbor (K-NN) dalam penelitian ini terbukti sangat dipengaruhi oleh kompleksitas pengelompokan kelas. Perbedaan performa yang signifikan terlihat antara skenario *multi-class classification* dan *binary classification*, yang menegaskan bahwa jumlah kelas serta kemiripan karakteristik data berperan penting dalam menentukan efektivitas algoritma berbasis jarak seperti K-NN. Pada skenario *multi-class classification* dengan empat kelas tujuan penggunaan AI, model K-NN menghasilkan akurasi yang relatif rendah, dengan nilai tertinggi sebesar 34%. Rendahnya akurasi ini mengindikasikan adanya kesulitan model dalam membedakan kelas-kelas yang memiliki karakteristik serupa, khususnya antara Akademik, Pekerjaan, dan Pengembangan Diri. Penelitian ini sejalan dengan penelitian [10] yang menyatakan bahwa klasifikasi dengan banyak label atau kelas saling beririsan cenderung menurunkan performa algoritma K-NN karena meningkatnya kompleksitas ruang fitur dan ambiguitas antar kelas. Hal tersebut tercermin pada matriks konfusi Gambar 7. yang menunjukkan tingginya kesalahan klasifikasi antar kelas produktif. Kondisi ini menunjukkan bahwa tumpang tindih pola fitur antar kelas menjadi tantangan utama dalam *multi-class classification* menggunakan algoritma K-NN.

Sebaliknya, pada skenario *binary classification*, kinerja model meningkat secara signifikan dengan akurasi tertinggi mencapai 77%. Pengelompokan tujuan penggunaan AI ke dalam dua kategori utama, yaitu Produktif/Karir dan Hiburan/Personal, terbukti mampu mengurangi kompleksitas klasifikasi dan memperjelas batas antar kelas. Temuan ini mendukung hasil penelitian [14] dan [13] membuktikan bahwa algoritma K-NN cenderung menghasilkan performa tinggi pada klasifikasi biner, khususnya pada data numerik atau data dengan perbedaan karakteristik yang lebih jelas. Analisis matriks konfusi menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengenali kelas Produktif/Karir, ditunjukkan oleh nilai recall yang tinggi. Namun demikian, model masih mengalami kesulitan dalam mengidentifikasi kelas Hiburan/Personal, di mana sebagian data hiburan diprediksi sebagai produktif. Hal ini mengindikasikan adanya kemiripan pola perilaku penggunaan AI antara kedua kategori tersebut dalam dataset.

Penerapan metode SMOTE pada data latih berkontribusi dalam menyeimbangkan distribusi kelas sehingga model dapat dilatih secara lebih stabil. Meskipun demikian, hasil evaluasi menunjukkan bahwa penyederhanaan kelas memiliki pengaruh yang lebih besar terhadap peningkatan performa dibandingkan sekadar penyeimbangan data. Secara keseluruhan, temuan ini menegaskan bahwa algoritma K-NN lebih efektif diterapkan pada skenario klasifikasi dengan jumlah kelas yang lebih sedikit dan batas kelas yang lebih jelas.

5. Simpulan

Berdasarkan hasil pengujian dan pembahasan, penelitian ini menemukan perbedaan kinerja yang signifikan dari algoritma K-NN antara skenario multi-class classification dan *binary classification* dalam mengklasifikasikan tujuan penggunaan AI oleh mahasiswa. Pada skenario *multi-class classification*, model K-NN hanya mencapai akurasi tertinggi sebesar 34%, yang menunjukkan bahwa model mengalami kesulitan dalam membedakan empat kelas tujuan penggunaan AI akibat kemiripan karakteristik antar kelas, khususnya antara Akademik, Pekerjaan, dan Pengembangan Diri.

Meskipun, pada skenario *binary classification*, penyederhanaan kelas menjadi Produktif/Karir dan Hiburan/Personal menghasilkan kinerja yang lebih baik, dengan akurasi pengujian tertinggi sebesar 77% pada nilai $K = 4$ dan $K = 6$. Temuan ini menegaskan bahwa pola penggunaan AI mahasiswa lebih konsisten ketika diklasifikasikan berdasarkan orientasi produktif dan non-produktif.

Dengan demikian, algoritma K-NN lebih sesuai diterapkan pada skenario klasifikasi dengan jumlah kelas yang terbatas dan batas kelas yang jelas. Untuk pengembangan selanjutnya, penelitian ini dapat diperluas dengan menerapkan algoritma klasifikasi non-linear seperti *Support Vector Machine* (SVM) atau pendekatan *Deep Learning* guna meningkatkan kinerja pada skenario multi-class dan mengatasi tumpang tindih fitur antar kelas.

Daftar Referensi

- [1] H. Johnston, R. F. Wells, E. M. Shanks, T. Boey, and B. N. Parsons, "Student perspectives on the use of generative artificial intelligence technologies in higher education," *Int. J. Educ. Integr.*, vol. 20, no. 1, pp. 1–21, 2024, doi: 10.1007/s40979-024-00149-4.
- [2] Y. S. Yasmine and R. Hikmawan, "ChatGPT sebagai Alat Bantu dalam Penulisan Karya Ilmiah Mahasiswa: Analisis Keterlibatan dan Kreativitas," *Edumatic J. Pendidik. Inform.*, vol. 9, no. 1, pp. 99–108, 2025, doi: 10.29408/edumatic.v9i1.29496.
- [3] M. A. Hakim and P. Gunawan, "Implementasi Algoritma K-Nearest Neighbors untuk Menganalisis Pendapat Pakar AI tentang Kemajuan Kecerdasan Buatan," vol. 5, no. 2, pp. 1245–1258, 2024, doi: 10.37445/cisa.v6i2.456.
- [4] Y. Reswan, R. Toyib, H. Witriyono, and A. Anggraini, "Klasifikasi Tingkat Kematangan Buah Nanas Berdasarkan Fitur Warna Menggunakan Metode K-Nearest Neighbor (KNN)," *J. Media Infotama*, vol. 20, no. 1, pp. 280–287, 2024.
- [5] Ismail; Nurhikma Arifin; Prihastinur, "Klasifikasi Kematangan Buah Naga Berdasarkan Fitur Warna Menggunakan Algoritma Multi-Class Support Vector Machine," *JINTEKS (Jurnal Inform. Teknol. dan Sains)*, vol. 5, no. 1, pp. 121–126, 2023, doi: 10.51401/jinteks.v5i1.2203.
- [6] Y. I. F. Y. R. F. S. 'Uyun, "Klasifikasi Citra Eurosat Menggunakan Algoritma Kn," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 4, pp. 7754–7761, 2024, doi: 10.36040/jati.v8i4.10458.
- [7] R. F. S. Rahmadani, "Performance Comparison Analysis of Classifiers on Binary Classification Dataset," *Indones. J. Data Sci.*, vol. 4, no. 2, pp. 45–54, 2023, doi: 10.56705/ijodas.v4i2.77.
- [8] I. Setiaji and V. Lusiana, "Uji akurasi Metode KNN dan Citra HSI dalam Mengklasifikasi Batik Solo Berdasarkan Motif," *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 12, no. 3, p. 1013, 2023, doi: 10.35889/jutisi.v12i3.1377.
- [9] F. Fathoni, D. Khoiriyyah Harahap, E. Theresia Pardede, S. Nachwa, M. Ramadhani Maulizidan, and A. Ibrahim, "Komparasi Kinerja Kn Dan Naïve Bayes Untuk Klasifikasi Sentimen Pengguna Aplikasi Deepseek Ai," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 4, pp. 6021–6028, 2025, doi: 10.36040/jati.v9i4.13887.
- [10] M. P. Kotakode, E. E. Fitriani, and W. Yustanti, "Perbandingan Kinerja Metode Problem Transformation- KNN dan Algorithm Adaptation-KNN pada Klasifikasi," vol. 03, no. 03, pp. 122–130, 2022.
- [11] R. Situmorang *et al.*, "Model Algoritma K-Nearest Neighbor (K-Nn) Dan Naïve Bayes," *JATI*

- (*Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 250–254, 2023, doi: 10.36040/jati.v7i1.6342.
- [12] A. Y. A. E. U. A. Yaqin, "Klasifikasi Konsentrasi Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *Infotek J. Inform. dan Teknol.*, vol. 7, no. 2, pp. 618–682, 2024, doi: 10.29408/jit.v7i2.27084.
- [13] N. R. G. V. T. E. P. E. P. Y. S. Ritonga, "Studi Komparasi Algoritma Decision Tree C4 . 5 dan K-Nearest Neighbor pada Klasifikasi Masa Studi dan Tingkat Stres Mahasiswa," *JUTISI J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 3, pp. 1776–1785, 2024, doi: 10.35889/jutisi.v13i3.2272.
- [14] D. N. C. A. M. S. R. Rahmat, "Klasifikasi Kelayakan Siswa dalam Menentukan Kelas Unggulan Menggunakan Algoritma," *Sci. Student J. Information, Technol. Sci.*, vol. 3, no. 1, pp. 51–58, 2022.
- [15] M. H. M. P. M. B. R. W. Hadinata, "Perbandingan Algoritma K-Nearest Neighbor dan Naive Bayes untuk Klasifikasi FoMO Pengguna Media Sosial," *J. Ilm. Komput.*, vol. 21, no.2, pp.422-433, 2025, doi: <https://ojs.stmikbanjarbaru.ac.id/index.php/progresif/article/view/2784>.