



Comparison of Naive Bayes and KNN for Honey-Mumford Learning Style Classification in Interpersonal Skill

Komparasi Naive Bayes dan KNN untuk Klasifikasi Gaya Belajar Honey-Mumford pada Interpersonal Skill

Hari Moerti¹, Hamzah Setiawan²

¹) Sistem dan Teknologi Informasi, Institut Teknologi dan Bisnis Yadika Pasuruan

²) Informatika, Universitas Muhammadiyah Sidoarjo

*Email to Correspondence: hari_moerti@stmik-yadika.ac.id, hamzah@umsida.ac.id

Abstract. Developing soft skills competence, particularly interpersonal abilities, often presents a challenge for Informatics students accustomed to technical and structured thinking patterns. The mismatch between teaching methods and student learning preferences can hinder the absorption of non-technical material. This study aims to classify student learning style profiles in the Interpersonal Skill course using a Machine Learning approach based on the Honey-Mumford model (Activist, Reflector, Theorist, Pragmatist). The research methodology employs Educational Data Mining techniques by comparing the performance of Naive Bayes and K-Nearest Neighbor (KNN) algorithms in predicting learning styles based on academic history data and behavioral questionnaires. Experimental results indicate that the Naive Bayes algorithm outperforms KNN in recognizing student characteristic patterns, achieving an accuracy rate of 93.33%. These findings suggest that engineering students possess heterogeneous learning styles; therefore, adaptive and varied teaching strategies are essential to optimize the comprehension of soft skills materia.

Keywords: Educational Data Mining, Learning Style, Honey-Mumford, Interpersonal Skill, Naive Bayes

Abstrak. Pengembangan kompetensi *soft skills*, khususnya kemampuan interpersonal, seringkali menjadi tantangan tersendiri bagi mahasiswa Informatika yang terbiasa dengan pola berpikir teknis dan terstruktur. Ketidaksesuaian antara metode pengajaran dosen dengan preferensi belajar mahasiswa dapat menghambat penyerapan materi non-teknis ini. Penelitian ini bertujuan untuk mengklasifikasikan profil gaya belajar mahasiswa pada mata kuliah *Interpersonal Skill* menggunakan pendekatan *Machine Learning* berbasis model *Honey-Mumford* (*Activist, Reflector, Theorist, Pragmatist*). Metodologi penelitian menerapkan teknik *Educational Data Mining* dengan mengkomparasikan performa algoritma *Naive Bayes* dan *K-Nearest Neighbor* (KNN) dalam memprediksi gaya belajar berdasarkan data riwayat akademis dan kuesioner perilaku. Hasil eksperimen menunjukkan bahwa algoritma *Naive Bayes* terbukti lebih unggul dalam mengenali pola karakteristik mahasiswa dengan tingkat akurasi mencapai 93,33%, dibandingkan metode KNN. Temuan ini mengindikasikan bahwa mahasiswa teknik memiliki ragam gaya belajar yang heterogen, sehingga diperlukan strategi pengajaran adaptif dan variatif untuk mengoptimalkan pemahaman materi *soft skills*.

Kata kunci- Educational Data Mining, Gaya Belajar, Honey-Mumford, Interpersonal Skill, Naive Bayes

PENDAHULUAN

Di era transformasi digital dan industri 4.0, pemanfaatan teknologi informasi telah menjadi kebutuhan krusial dalam berbagai sektor. Implementasi sistem tidak hanya berfokus pada kemudahan akses informasi publik berbasis *mobile* [1], tetapi juga telah berkembang menjadi alat monitoring dan evaluasi kinerja yang efektif berbasis *web* [2]. Sejalan dengan tren tersebut, sektor pendidikan kini juga memanfaatkan teknologi komputasi untuk mengenali karakteristik

personal mahasiswa. Penguasaan *soft skills*, khususnya *Interpersonal Skill*, telah menjadi determinan utama dalam kesuksesan profesional lulusan perguruan tinggi. Kemampuan untuk berkomunikasi, berkolaborasi dalam tim, dan beradaptasi dengan lingkungan kerja yang dinamis menuntut metode pengajaran yang tidak hanya bersifat teoretis, tetapi juga personal. Namun, dalam implementasinya di kelas, pengajar seringkali menghadapi tantangan berupa heterogenitas gaya belajar mahasiswa yang berdampak pada efektivitas

penyerapan materi.

Gaya belajar merupakan preferensi individu dalam memproses dan mengasimilasi informasi baru. Salah satu model yang sangat relevan dalam pengembangan keterampilan praktis adalah model gaya belajar Honey-Mumford yang mengklasifikasikan mahasiswa ke dalam empat tipe utama, yaitu *Activist*, *Reflector*, *Theorist*, dan *Pragmatist*, yang masing-masing memiliki karakteristik belajar dan kebutuhan instruksional yang berbeda [3].

Meskipun identifikasi gaya belajar sangat penting, metode konvensional yang mengandalkan kuesioner manual seringkali memakan waktu dan sulit dilakukan secara berulang dalam skala besar. Di sinilah peran *Educational Data Mining* (EDM) menjadi relevan. EDM memungkinkan penggalian pola perilaku belajar mahasiswa secara otomatis melalui pemanfaatan data akademik dan teknik *Machine Learning* [4]. Melalui pendekatan algoritma *Machine Learning*, data perilaku dan performa mahasiswa dapat ditransformasikan menjadi model prediksi yang akurat untuk menentukan kecenderungan gaya belajar mereka tanpa intervensi manual yang subjektif.

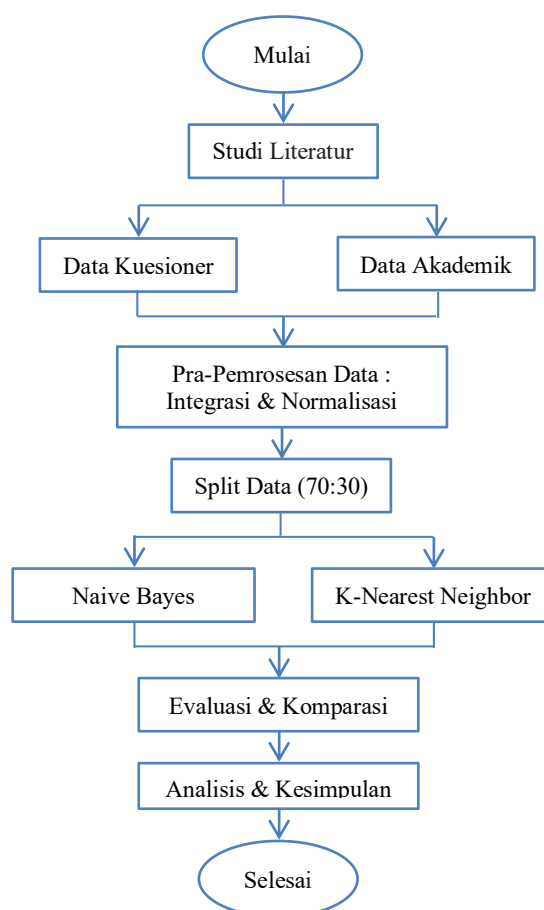
Beberapa penelitian terdahulu telah mengeksplorasi penggunaan teknik klasifikasi dalam EDM, namun mayoritas masih berfokus pada mata kuliah berbasis kognitif keras (*hard skills*) seperti matematika atau pemrograman. Masih terdapat celah penelitian (*research gap*) dalam penerapan algoritma klasifikasi pada mata kuliah pengembangan karakter dan keterampilan interpersonal. Misalnya, penelitian yang memadukan *Correlation-Based Feature Selection* dan *Naive Bayes* menunjukkan kemampuan dalam memprediksi performa akademik siswa pada mata pelajaran matematika dan bahasa Portugis [5] serta Penggunaan algoritma klasifikasi untuk memprediksi kelulusan mahasiswa pada mata kuliah *Probability & Statistics* [6]. Karakteristik data pada mata kuliah *Interpersonal Skill* memiliki keunikan tersendiri karena melibatkan aspek perilaku sosial dan performa praktis yang berbeda dari ujian tertulis pada umumnya.

Penelitian ini bertujuan untuk menerapkan dan mengevaluasi performa algoritma klasifikasi dalam mendeteksi gaya belajar Honey-Mumford pada mahasiswa mata kuliah *Interpersonal Skill*. Dengan

memanfaatkan data primer yang diintegrasikan dengan performa akademik, diharapkan penelitian ini dapat menghasilkan model yang mampu memetakan profil belajar mahasiswa secara presisi. Hasil dari klasifikasi ini diharapkan tidak hanya berkontribusi pada literatur ilmu komputer di bidang EDM, tetapi juga memberikan rekomendasi praktis bagi institusi pendidikan dalam merancang strategi pembelajaran yang lebih personal dan adaptif.

METODE PENELITIAN

Tahapan penelitian ini dirancang menggunakan pendekatan *Educational Data Mining* (EDM) untuk mengklasifikasikan gaya belajar mahasiswa [7]. Secara sistematis, alur penelitian dimulai dari studi literatur, pengumpulan data ganda, pra-pemrosesan, hingga evaluasi model seperti terlihat pada Gambar 1.xz`1



Gambar 1. Alur Penelitian Klasifikasi Gaya Belajar

A. Jenis Penelitian

Penelitian ini menggunakan pendekatan **eksperimental kuantitatif** yang berfokus pada sub-bidang *Artificial Intelligence* (AI), yaitu *Machine*

Learning. Dalam bidang pendidikan *Machine Learning* berperan penting dalam mendukung pengambilan keputusan berbasis data untuk meningkatkan kualitas pembelajaran [8]. Metode ini dipilih untuk menguji akurasi model klasifikasi dalam memprediksi gaya belajar mahasiswa dengan mengintegrasikan data psikometrik dan data akademik.

B. Lokasi Penelitian

Penelitian dilaksanakan di **Institut Teknologi dan Bisnis (ITB) Yadika Pasuruan**, Fakultas Teknologi Informasi, Program Studi Sistem dan Teknologi Informasi. Lokasi ini dipilih karena ketersediaan akses data nilai akademik serta kemudahan dalam penyebaran kuesioner kepada responden mahasiswa aktif angkatan 2023.

C. Sumber Data

Data yang digunakan dalam penelitian ini bersumber dari dua kategori utama:

1. **Data Primer:** Data gaya belajar mahasiswa yang diperoleh melalui pengisian instrumen *Learning Style Questionnaire* (LSQ) adaptasi Honey-Mumford. Setiap tipe gaya belajar Honey-Mumford menunjukkan pendekatan yang berbeda dalam menyelesaikan permasalahan, sehingga strategi pembelajaran perlu disesuaikan dengan karakteristik peserta didik [9].
2. **Data Sekunder:** Data riwayat akademik mahasiswa yang diambil dari basis data Sistem Informasi Akademik (SIKAD), mencakup atribut: Nilai Tugas, Nilai UTS, Nilai UAS, dan Nilai Kehadiran.

D. Teknik Pengumpulan Data

Teknik pengumpulan data dilakukan melalui dua tahap:

1. **Kuesioner Daring:** Menyebarkan *Google Form* berisi pertanyaan LSQ. Data mentah yang bersifat kategorikal (a, b, c, d) kemudian dikonversi menjadi label kelas target: *Activist*, *Reflector*, *Theorist*, dan *Pragmatist*.
2. **Studi Dokumentasi:** Melakukan inventarisasi data nilai dari arsip penilaian akademik pada mata kuliah yang diampu oleh peneliti. Data sekunder ini

meliputi rekapitulasi komponen nilai Tugas, UTS, UAS, dan Kehadiran mahasiswa semester 5 Tahun Akademik 2025/2026 yang tersimpan dalam basis data penilaian dosen. Data ini kemudian diintegrasikan dengan data kuesioner menggunakan atribut kunci Nomor Induk Mahasiswa (NIM).

E. Analisis Data

Tahapan analisis data mengikuti prosedur *Knowledge Discovery in Databases* (KDD) yang terdiri dari:

1. Pra-pemrosesan Data (*Preprocessing*)

Melakukan pembersihan data (*cleaning*) dan Normalisasi. Mengingat terdapat perbedaan skala antara data kuesioner (1-4) dan data nilai (0-100), dilakukan normalisasi menggunakan teknik Min-Max Scaler agar seluruh data berada dalam rentang $[0, 1]$ dengan persamaan:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Dimana:

- x' : Nilai data setelah dinormalisasi (rentang 0 sampai 1).
- x : Nilai data asli sebelum diproses.
- $\min(x)$: Nilai minimum (terendah) dari seluruh data pada fitur tersebut.
- $\max(x)$: Nilai maksimum (tertinggi) dari seluruh data pada fitur tersebut.

2. Pembagian Dataset (*Data Splitting*)

Dataset dibagi menggunakan teknik hold-out validation dengan rasio 70:30. Sebanyak 70% digunakan sebagai data latih (*training set*) dan 30% sebagai data uji (*testing set*) untuk mengukur performa model terhadap data baru.

3. Algoritma Naive Bayes

Metode ini menerapkan Teorema Bayes dengan asumsi independensi antar fitur. Naive Bayes merupakan salah satu algoritma klasifikasi yang sederhana namun memiliki tingkat akurasi yang baik dalam mengidentifikasi gaya belajar mahasiswa [10]. Probabilitas kelas gaya belajar C berdasarkan fitur X dihitung menggunakan persamaan:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)}$$

Mengingat $P(X)$ bersifat konstan untuk semua kelas, maka fungsi keputusan klasifikasi dapat disederhanakan menjadi:

$$C_{\text{map}} = \arg \max_{C \in \mathcal{C}} P(C) \prod_{i=1}^n P(x_i | C)$$

Dimana:

- $P(C|X)$ adalah *Posterior Probability* (Peluang gaya belajar C jika diketahui data X).
- $P(X|C)$ adalah *Likelihood* (Peluang data X muncul pada gaya belajar C).
- $P(C)$ adalah *Prior Probability* (Peluang awal kelas gaya belajar).
- $P(X)$ adalah *Evidence* (Peluang kemunculan data secara keseluruhan).

4. Algoritma K-Nearest Neighbor (KNN)

Algoritma KNN bekerja berdasarkan prinsip kedekatan jarak antara objek data baru dengan data latih (*training data*) pada ruang fitur. Untuk menentukan kelas gaya belajar, digunakan perhitungan jarak **Euclidean Distance** dengan formula sebagai berikut:

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

Dimana:

- $d(x, y)$ adalah jarak antara data uji x dan data latih y .
- n adalah jumlah fitur atau dimensi data.
- x_i dan y_i adalah nilai fitur ke- i pada masing-masing data.

5. Evaluasi Model

Evaluasi Model Kinerja kedua algoritma diukur menggunakan parameter Confusion Matrix untuk mengetahui seberapa baik model dalam memprediksi kelas gaya belajar. Tiga metrik utama yang digunakan adalah Akurasi (*Accuracy*), Presisi (*Precision*), dan *Recall*. Perhitungan dilakukan menggunakan persamaan berikut:

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$\text{Presisi} = \frac{TP}{TP + FP} \times 100\%$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\%$$

Dimana:

- *TP (True Positive)*: Jumlah data kelas gaya belajar yang diprediksi dengan benar (Positif) sesuai dengan kelas aslinya.
- *TN (True Negative)*: Jumlah data yang bukan kelas target dan diprediksi dengan benar sebagai bukan kelas tersebut.
- *FP (False Positive)*: Jumlah data negatif yang salah diprediksi sebagai kelas positif (Kesalahan Tipe I).
- *FN (False Negative)*: Jumlah data positif yang salah diprediksi sebagai kelas negatif (Kesalahan Tipe II/Gagal mendeteksi).

HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil eksperimen klasifikasi gaya belajar mahasiswa menggunakan algoritma Naive Bayes dan K-Nearest Neighbor (KNN). Tahapan pembahasan dimulai dari hasil pra-pemrosesan data, pengujian model, hingga analisis komparasi performa kedua algoritma.

A. Hasil Pra-pemrosesan Data

Dataset awal terdiri dari 51 data mahasiswa. Sebelum dilakukan pemodelan, tahap data cleaning dilakukan untuk menjamin validitas hasil. Sebanyak dua data dihapus karena mengandung missing values (tidak mengisi label kelas) dan outlier (anomali nilai ekstrem). Sehingga, total data bersih yang digunakan berjumlah 49 data.

Selanjutnya, dilakukan proses normalisasi menggunakan *Min-Max Scaler* untuk menyetarakan skala data akademik (rentang 0-100) dan kuesioner (rentang 1-4) menjadi rentang $[0, 1]$. Sampel hasil normalisasi dapat dilihat pada Tabel 1.

Tabel 1. Sampel Data Sebelum dan Sesudah Normalisasi

Atribut	Nilai Asli (Raw Data)	Nilai Ter-Normalisasi [0-1]
Nilai Tugas	85.0	0.615
Nilai UTS	80.0	0.600
Nilai UAS	84.0	0.733
Kehadiran	100.0	1.000
Q1 (Public Speaking)	3 (Skala 4)	0.667
Q2 (Presentasi)	4 (Skala 4)	1.000

Sumber: Data primer diolah

Seperti terlihat pada Tabel 1, transformasi ini memastikan bahwa tidak ada fitur yang mendominasi fitur lain hanya karena memiliki satuan angka yang lebih besar.

B. Hasil Implementasi dan Pengujian

Sesuai metodologi, data dibagi dengan rasio 70:30 (34 data latih dan 15 data uji). Pengujian dilakukan untuk mengukur kemampuan model dalam memprediksi kelas gaya belajar (*Activist*, *Reflector*, *Theorist*, *Pragmatist*) berdasarkan data uji yang belum pernah dilihat sebelumnya.

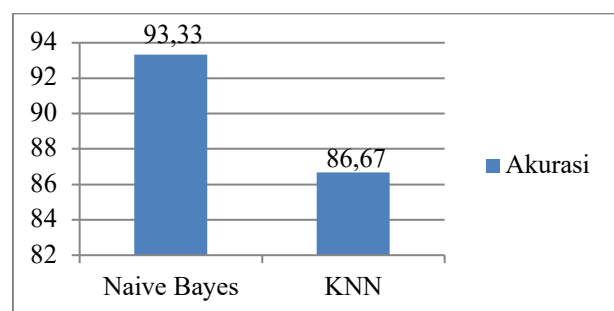
Hasil pengukuran kinerja menggunakan *Confusion Matrix* menghasilkan metrik Akurasi, Presisi, dan *Recall* sebagaimana tersaji dalam Tabel 2.

Tabel 2. Perbandingan Performa Naive Bayes dan KNN

Metrik Evaluasi	Naive Bayes	K-Nearest Neighbor (K=3)
Akurasi (Accuracy)	93.33%	86.67%
Presisi (Precision)	94.44%	88.57%
Recall	93.33%	86.67%

Sumber: Data primer diolah

Berdasarkan Tabel 2, algoritma *Naive Bayes* mencatatkan performa yang lebih unggul dibandingkan KNN pada seluruh metrik evaluasi. Algoritma *Naive Bayes* memiliki performa klasifikasi yang lebih stabil dibandingkan K-Nearest Neighbor pada dataset pendidikan berukuran kecil [11]. Visualisasi perbandingan akurasi kedua metode dapat dilihat pada Gambar 2.

**Gambar 2.** Grafik Perbandingan Akurasi

Berdasarkan hasil eksperimen, ditemukan bahwa algoritma *Naive Bayes* merupakan metode yang paling optimal untuk dataset gaya belajar mahasiswa ini dengan tingkat akurasi mencapai **93.33%**.

Keunggulan performa *Naive Bayes* dibandingkan KNN (86.67%) dapat dianalisis dari karakteristik data. Algoritma *Naive Bayes* bekerja sangat baik dan optimal pada dataset dengan jumlah sampel terbatas (49 data) karena metode ini memanfaatkan teori probabilitas yang kuat [12]. Asumsi independensi fitur pada *Naive Bayes* terbukti efektif menangani pola nilai akademik dan jawaban kuesioner, meskipun terdapat korelasi antar variabel.

Sebaliknya, sedikit penurunan performa pada KNN kemungkinan disebabkan oleh sensitivitas algoritma berbasis jarak (*distance-based*) terhadap distribusi data yang mungkin memiliki *boundary* (batas) antar kelas yang kurang tegas pada beberapa sampel mahasiswa.

Secara implikasi praktis, tingginya akurasi model ini (di atas 90%) menunjukkan bahwa pola nilai akademik (Tugas, UTS, UAS) yang dikombinasikan dengan pertanyaan psikometrik sederhana, dapat dijadikan indikator yang andal untuk mendeteksi gaya belajar mahasiswa secara otomatis tanpa perlu tes psikologi yang panjang.

C. Analisis Detail Performa per Kelas

Selain mengukur akurasi global, penelitian ini juga menganalisis kemampuan algoritma *Naive Bayes* dalam mengenali setiap kategori gaya belajar secara spesifik. Analisis ini penting untuk mengetahui kategori mana yang paling mudah atau paling sulit dideteksi oleh sistem. Pengelompokan gaya belajar mahasiswa memberikan dasar yang kuat bagi pengembangan sistem pembelajaran personal dan berbasis data [13]. Rincian performa klasifikasi untuk setiap kelas gaya belajar disajikan dalam Tabel 3.

Tabel 3. Detail Performa Naive Bayes per Kelas Gaya Belajar

Kelas Gaya Belajar	Precision	Recall	F1-Score	Keterangan
<i>Activist</i>	100%	100%	100%	Sangat Baik
<i>Pragmatist</i>	80%	100%	89%	Baik
<i>Reflector</i>	100%	75%	86%	Cukup Baik
<i>Theorist</i>	100%	100%	100%	Sangat Baik

Sumber: Data primer diolah

Berdasarkan Tabel 3, terlihat bahwa algoritma mampu mengenali gaya belajar tipe *Activist* dan *Theorist* dengan sempurna (Presisi dan Recall 100%). Namun, terdapat sedikit penurunan performa pada pengenalan tipe *Reflector*, dimana nilai *Recall* tercatat sebesar 75%. Hal ini mengindikasikan bahwa ada sebagian kecil mahasiswa tipe *Reflector* yang terdeteksi sebagai tipe lain. Deteksi gaya belajar berbasis *Machine Learning* memungkinkan sistem *e-learning* untuk memberikan pengalaman belajar yang lebih personal dan efektif [14].

D. Analisis Kesalahan (Confusion Matrix)

Untuk memahami lebih dalam distribusi kesalahan prediksi, digunakan visualisasi *Confusion Matrix* sebagaimana ditunjukkan pada Tabel 4. Matriks ini memetakan perbandingan antara Kelas Aktual (fakta) dengan Kelas Prediksi (tebakan sistem).

Tabel 4. Confusion Matrix Algoritma Naive Bayes

Aktual \ Prediksi	Prediksi <i>Activist</i>	Prediksi <i>Pragmatist</i>	Prediksi <i>Reflector</i>	Prediksi <i>Theorist</i>
Aktual <i>Activist</i>	4	0	0	0
Aktual <i>Pragmatist</i>	0	4	0	0
Aktual <i>Reflector</i>	0	1	3	0
Aktual <i>Theorist</i>	0	0	0	3

Sumber: Data primer diolah

(Cara baca Tabel 4: Diagonal angka tebal (4, 4, 3, 3) adalah tebakan benar. Angka di luar diagonal adalah tebakan salah).

Analisis terhadap *Confusion Matrix* menunjukkan bahwa:

1. **Akurasi Sempurna:** Sebanyak 4 mahasiswa *Activist*, 4 *Pragmatist*, dan 3 *Theorist* berhasil diprediksi dengan tepat tanpa kesalahan (Total 11 data benar).
2. **Misklasifikasi:** Terdapat 1 kesalahan prediksi (*Miss-classification*), dimana 1 mahasiswa yang sebenarnya memiliki gaya belajar *Reflector*, oleh sistem diprediksi sebagai *Pragmatist*.
3. **Analisis Penyebab:** Kesalahan prediksi pada kasus *Reflector* menjadi *Pragmatist* kemungkinan disebabkan oleh kemiripan pola nilai jawaban pada kuesioner butir ke-4 (Kerjasama Tim), dimana kedua

tipe ini memiliki kecenderungan preferensi kerja yang hampir beririsan pada dataset latih.

Analisis perilaku belajar mahasiswa melalui pendekatan *Machine Learning* dapat membantu dosen dalam merancang strategi pembelajaran yang lebih adaptif [15]. Analisis kluster menunjukkan bahwa mahasiswa memiliki kecenderungan gaya belajar yang heterogen dan tidak dapat diperlakukan dengan pendekatan pembelajaran yang seragam [16].

E. Implikasi Pedagogis dan Strategi Pengajaran

Berdasarkan temuan bahwa mahasiswa tersebar ke dalam empat kluster gaya belajar (*Activist*, *Reflector*, *Theorist*, *Pragmatist*), penelitian ini merekomendasikan strategi pengajaran berbasis **Siklus Belajar Kolaboratif** untuk mata kuliah *Interpersonal Skill*.

Mengingat materi UAS juga mencakup pemahaman diri (*Johari Window*), berikut adalah peta strategi yang disarankan agar materi dapat diterima oleh seluruh tipe mahasiswa:

1. **Bagi Mahasiswa *Theorist*:** Dosen perlu mengawali sesi dengan **Pemaparan Konsep Teoretis**. Contohnya, menjelaskan teori *Johari Window* secara struktural sebelum meminta mahasiswa mempraktikkannya. Kelompok ini membutuhkan kerangka logika yang jelas sebelum bertindak.
2. **Bagi Mahasiswa *Reflector*:** Diberikan porsi **Studi Kasus dan Observasi**. Sebelum praktik, mereka diminta mengamati video simulasi komunikasi. Mereka belajar paling efektif melalui pengamatan dan refleksi mendalam, bukan paksaan spontanitas.
3. **Bagi Mahasiswa *Activist*:** Metode **Roleplay (Bermain Peran)** adalah kunci. Setelah teori singkat, mahasiswa tipe ini harus segera dilibatkan dalam simulasi, misalnya mempraktikkan cara memberi umpan balik (*feedback*). Mereka belajar melalui pengalaman langsung (*learning by doing*).
4. **Bagi Mahasiswa *Pragmatist*:** Fokus pada **Implementasi Dunia Kerja**. Materi dikaitkan dengan simulasi wawancara kerja atau

penanganan keluhan pelanggan (*real-world problem solving*). Tipe ini akan termotivasi jika melihat kegunaan praktis dari materi kuliah bagi karir masa depan mereka.

Penerapan strategi campuran (*blended strategy*) ini diharapkan dapat meminimalisir kesenjangan pemahaman antar mahasiswa yang memiliki preferensi belajar berbeda. Lingkungan pembelajaran digital memperkuat kebutuhan akan strategi pembelajaran adaptif yang mampu mengakomodasi perbedaan gaya belajar mahasiswa teknik di era pembelajaran digital [17].

KESIMPULAN

Penelitian ini menyimpulkan bahwa penerapan *Educational Data Mining* menggunakan algoritma **Naive Bayes** lebih efektif dibandingkan **K-Nearest Neighbor (KNN)** dalam memetakan profil gaya belajar mahasiswa pada mata kuliah *Interpersonal Skill*, dengan capaian akurasi tertinggi sebesar **93,33%**. Hasil ini menunjukkan bahwa pemanfaatan kombinasi data riwayat akademik dan kuesioner perilaku sederhana dapat menjadi alternatif instrumen deteksi gaya belajar yang cepat dan valid tanpa memerlukan tes psikologi yang kompleks. Temuan keberagaman gaya belajar mahasiswa (*Activist, Reflector, Theorist, dan Pragmatist*) mengindikasikan perlunya strategi pembelajaran adaptif, sehingga penelitian ini merekomendasikan penerapan pendekatan pengajaran berbasis **siklus belajar kolaboratif** yang mengombinasikan ceramah konseptual, observasi studi kasus, simulasi peran (*roleplay*), dan pemecahan masalah kontekstual untuk mendukung pengembangan *soft skills* secara inklusif. Namun demikian, penelitian ini masih terbatas pada jumlah sampel yang relatif kecil, sehingga penelitian lanjutan dengan cakupan data yang lebih luas dan penambahan variabel pembelajaran disarankan guna meningkatkan generalisasi dan akurasi model.

REFERENSI

- [1] H. Moerti, K. W. Haryanto, and F. E. Andriyanto, "Aplikasi Informasi Tempat Pariwisata di Kabupaten Pasuruan Berbasis Android Menggunakan Metode RAD (Rapid Application Development)," *J. Spirit*, vol. 17, no. 1, pp. 98–106, 2025.
- [2] A. S. Illahi, M. K. Setiyansyah, W. Mualim, and H. Moerti, "Perancangan Dashboard Kinerja Berbasis Web dengan Menggunakan Metode Key Performance Indicators Sebagai Monitoring Perkembangan Peforma Karyawan," *JOSIATI (Jurnal Sist. Inf. Apl. Teknol. Informasi)*, vol. 2, no. 1, pp. 45–57, 2025, doi: 10.53567/josiati.v2i1.35.
- [3] V. Yuliza, N. Huda, and R. Junita, "Analisis kemampuan metakognitif dalam pemecahan masalah matematis ditinjau dari gaya belajar Honey-Mumford," *EMTEKA J. Pendidik. Mat.*, vol. 5, no. 2, 2024, doi: 10.24127/emteka.v5i2.5532.
- [4] S. Sumarlin and D. Angraini, "Data Mining Pendidikan: Prediksi Gaya Belajar Mahasiswa Teknik Menggunakan Machine Learning," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 3, pp. 563–572, Jun. 2025, doi: 10.25126/jtiik.2025129190.
- [5] T. Gori, A. Sunyoto, and H. Al Fatta, "Preprocessing data dan klasifikasi untuk prediksi kinerja akademik siswa," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 1, 2024.
- [6] Y. N. Rumbia and others, "Perbandingan metode KNN dan Naive Bayes untuk klasifikasi kelulusan mahasiswa pada mata kuliah Probabilitas dan Statistika," *J. PTI*, vol. 12, no. 1, 2025, doi: 10.35134/jpti.v12i1.228.
- [7] D. A. N. Wulandari, R. Annisa, L. Yusuf, and T. Prihatin, "Educational data mining for student academic prediction using K-Means and Naïve Bayes classifier," *J. Pilar Nusa Mandiri*, vol. 16, no. 2, 2020, doi: 10.33480/pilar.v16i2.1432.
- [8] M. Yağcı, "Educational data mining: prediction of students' academic performance using machine learning algorithms," *Smart Learn.*

- Environ.*, vol. 9, no. 11, 2022, doi: 10.1186/s40561-022-00192-z.
- [9] N. Citroesmi and E. Susanti, “Kemampuan pemecahan masalah matematis siswa ditinjau dari gaya belajar teori Honey-Mumford,” *JUMLAHKU J. Mat. Ilm.*, vol. 7, no. 2, 2021, doi: 10.33222/jumlahku.v7i2.1442.
- [10] S. Ramadandi and Jahring, “Klasifikasi gaya belajar mahasiswa menggunakan metode Naïve Bayes classifier,” *J. Teknol. dan Inf.*, vol. 10, no. 2, 2020, doi: 10.34010/jati.v10i2.3096.
- [11] S. T. Sianturi and U. L. Yuhana, “Student behaviour analysis to detect learning styles using decision tree, Naïve Bayes, and K-nearest neighbor method in Moodle LMS,” *IPTEK J. Teknol. dan Sains*, 2025.
- [12] H. N. Asmi, S. Risnanto, and O. B. Mohd, “Optimizing Naïve Bayes method for Felder–Silverman learning style model identification,” *J. Tek. Inform.*, vol. 18, no. 1, 2025.
- [13] L. Hanum, “Pengelompokan gaya belajar mahasiswa menggunakan metode K-Means dan validasi Davies–Bouldin Index,” *J. Manajemen, Pendidik. dan Ilmu Komput.*, vol. 2, no. 2, 2025, doi: 10.65309/1x1b7f47.
- [14] F. Rasheed and A. Wahid, “Learning style detection in E-learning systems using machine learning techniques,” *Expert Syst. Appl.*, vol. 174, p. 114774, 2021, doi: 10.1016/j.eswa.2021.114774.
- [15] Z. Meheuaoui, Y. Lafifi, and L. Zemmouri, “Learning behavior analysis to identify learner’s learning style based on machine learning techniques,” *J. Univers. Comput. Sci.*, vol. 28, no. 11, pp. 1193–1220, 2022.
- [16] I. Maulana, W. Witanti, and Melina, “Klasterisasi gaya belajar mahasiswa berbasis VARK dengan algoritma DBSCAN,” *J. Algoritm.*, vol. 22, no. 2, 2025, doi: 10.33364/algoritma/v.22-2.2980.
- [17] S. Samnur, F. R. Baharuddin, and J. R. Khairah, “Analisis gaya belajar online mahasiswa pendidikan teknik di era digital,” *J. Mediat.*, vol. 6, no. 2, 2023, doi: 10.59562/mediatik.v6i2.1412.

Conflict of Interest Statement:

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Article History:

Received: 09 August 2025 | Accepted: 30 August 2025 | Published: 30 November 2025