# **BAB I**

# **PENDAHULUAN**

## Latar Belakang

Artificial Intelligence (AI), atau dalam bahasa Indonesia dikenal sebagai Kecerdasan Buatan, adalah cabang ilmu komputer yang bertujuan untuk mengembangkan sistem dan mesin yang mampu melakukan tugas yang biasanya memerlukan kecerdasan manusia. AI melibatkan penggunaan algoritma dan model matematika untuk memungkinkan komputer dan sistem lainnya untuk belajar dari data, mengenali pola, dan membuat keputusan yang cerdas. Dalam konteks AI, terdapat beberapa konsep penting seperti machine learning (pembelajaran mesin), neural networks (jaringan saraf tiruan), natural language processing (pemrosesan bahasa alami), dan banyak lagi. Pengembangan AI telah memberikan dampak besar dalam berbagai bidang seperti pengenalan suara, pengenalan wajah, mobil otonom, pengobatan, dan masih banyak lagi. (Emi Sita Eriana & Drs. Afrizal Zein, 2023)

Kehadiran teknologi membantu manusia dalam meringankan kegiatan sehari-harinya secara efisien. Kemajuan pada teknologi melahirkan berbagai inovasi baru salah satunya adalah hadirnya teknologi kecerdasan buatan atau artificial intelligenceyang dirancang berkaitan dengan kecerdasan manusia. Tujuan dilakukannya penelitian ini adalah untuk mengetahui bagaimana pengaruh positif dan negatif dari penggunaan kecerdasan buatan pada mahasiswa di perguruan tinggi. Dari pengumpulan data yang dilakukan menggunakan studi literatur ditemukan berbagai pengaruh pada penggunaan kecerdasan buatan pada mahasiswa. Adapun pengaruh positif yang ditimbulkan adalah mahasiswa mudah dalam mengakses materi pembelajaran secara luas, mudah dalam memahami bahasa asing, dapat bertanya kapanpun dan akan dijawab pada saat itu juga, dan mentor yang selalu ada dalam membantu pembelajaran. Disamping itu, pengaruh negatif yang ditimbulkan dari adanya penggunaan AI yaitu seperti keamanan data yang belum dipastikan aman atau tidaknya data mahasiswa yang menggunakan AI tersebut. Penulis menyadari masih terdapat kekurangan dalam penelitian ini. Saran untuk penelitian selanjutnya yakni melakukan tinjauan langsung terhadap mahasiswa yang terlibat langsung menggunakan AI dalam mempermudah kegiatan perkuliahan sehari-hari, agar dapat meneliti langsung mahasiswa yang terlibat. (Salsabila, Hadi, Pratiwi, & SitiMukarromah, 2023)

*Arificial Intelligence* (AI) telah menjadi salah satu teknologi revolusioner yang mempengaruhi berbagai bidang, termasuk pendidikan. Dalam konteks mahasiswa, penerapan AI semakin meluas, terutama dalam mendukung proses belajar dan penyelesaian tugas. Pengaruh AI terhadap kualitas, efisiensi, dan pemahaman mahasiswa dalam menyelesaikan tugas dapat dianalisis dari beberapa sudut pandang:

Kualitas Tugas Mahasiswa AI memiliki potensi untuk meningkatkan kualitas tugas yang dikerjakan oleh mahasiswa. Alat-alat berbasis AI, seperti perangkat pengecekan tata bahasa (*grammar checker*) dan alat bantu penulisan, dapat membantu mahasiswa menghasilkan karya yang lebih rapi dan profesional. Selain itu, AI juga dapat digunakan dalam penelitian untuk mengumpulkan dan menganalisis data dengan lebih cepat dan akurat, memungkinkan mahasiswa menyajikan hasil penelitian yang lebih kaya dan terperinci. Namun, ada kekhawatiran bahwa ketergantungan pada AI dapat menurunkan kreativitas dan keterampilan berpikir kritis mahasiswa, karena AI sering kali memberikan solusi langsung tanpa mendorong pemahaman mendalam.

Efisiensi dalam Penyelesaian Tugas Salah satu pengaruh terbesar AI dalam konteks penyelesaian tugas mahasiswa adalah peningkatan efisiensi. AI mampu mengotomatisasi banyak tugas rutin, seperti pencarian referensi, penyusunan daftar pustaka, dan perbaikan tata bahasa, sehingga mahasiswa dapat menyelesaikan tugas mereka lebih cepat. Selain itu, AI juga dapat memberikan rekomendasi personal yang disesuaikan dengan gaya belajar individu, membantu mahasiswa menemukan materi yang relevan lebih cepat dan menghemat waktu. Meskipun demikian, peningkatan efisiensi ini juga dapat menimbulkan potensi negatif, di mana mahasiswa lebih tergoda untuk menggunakan AI sebagai solusi instan tanpa benar-benar memahami substansi tugas.

Pemahaman Materi Tugas Penggunaan AI dalam penyelesaian tugas dapat berdampak ganda terhadap pemahaman mahasiswa. Di satu sisi, AI yang dirancang untuk memberikan bimbingan atau pembelajaran adaptif dapat membantu mahasiswa memperdalam pemahaman mereka terhadap suatu materi. Contoh AI seperti chatbot atau tutor virtual dapat memberikan penjelasan tambahan dan bimbingan personal yang membantu menjawab pertanyaan atau kesulitan yang dialami mahasiswa. Di sisi lain, ada kekhawatiran bahwa mahasiswa mungkin menggunakan AI hanya untuk menyelesaikan tugas secara cepat tanpa memahami isi atau konteks dari materi yang dipelajari.

Secara keseluruhan, AI memiliki dampak yang signifikan terhadap kualitas, efisiensi, dan pemahaman mahasiswa dalam menyelesaikan tugas. Meski dapat membantu meningkatkan performa akademik dan menghemat waktu, penting bagi mahasiswa untuk tetap menjaga keseimbangan antara penggunaan teknologi dan pengembangan kemampuan berpikir mandiri serta pemahaman mendalam terhadap materi. Tantangan bagi institusi pendidikan adalah bagaimana mengintegrasikan AI secara bijak dalam proses belajar mengajar, sehingga teknologi ini dapat meningkatkan hasil pembelajaran tanpa mengorbankan kualitas pembelajaran.

## Identifikasi Masalah

Latar belakang yang tersusun telah termuat diatas, berdasarkan hal tersebut dapat diidentifikasi bahwa:

1. Penggunaan AI yang berlebihan dapat membuat mahasiswa terlalu bergantung pada teknologi untuk menyelesaikan tugas-tugas mereka
2. Mahasiswa mungkin tergoda untuk mengandalkan AI hanya untuk menyelesaikan tugas dengan cepat, tanpa mendalami materi yang diperlukan atau memahami konsep yang diajarkan.
3. Mahaiswa melewatkan proses berpikir kritis dalam pemecahan masalah. Ini mengurangi kesempatan mahasiswa untuk belajar bagaimana menganalisis dan mencari solusi yang lebih kompleks secara mandiri.

## Rumusan Masalah

Berdasarkan pada identifikasi masalah yang telah ada maka dapat dirumuskan permasalahannya sebagai berikut:

1. Bagaimana pengaruh penggunaan *Arificial Intelligence* (AI) terhadap kualitas hasil tugas mahasiswa dalam aspek orisinalitas dan kreativitas?
2. Apakah AI membantu mahasiswa dalam mengelola waktu dengan lebih baik dalam penyelesaian tugas, atau justru menyebabkan ketergantungan yang berlebihan?
3. Apakah mahasiswa yang menggunakan AI dalam penyelesaian tugas cenderung memiliki pemahaman yang dangkal terhadap materi dibandingkan dengan yang menyelesaikan tugas secara mandiri?

## Batasan Penelitian

Supaya penelitian ini dapat sesuai dengan arahnya dan tidak keluar jalurnya maka dibutuhkannya suatu pembatasan masalah. sebagai berikut:

1. Proses penelitian ini akan menggunakan metode *Unsupervised Learning.*
2. Menganalisis pengaruh AI terhadap kualitas, efisiensi, dan pemahaman penyelesaian tugas yang dibagi dengan bersifat *positif, negative,* dan *netral.*
3. Menjelaskan topik atau variabel spesifik yang akan dibahas. Topik lain yang berkaitan tetapi tidak relevan secara langsung akan dikecualikan.

## Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Menganalisis pengaruh penggunaan kecerdasan buatan terhadap kualitas mahasiswa dalam menyelesaikan tugas akademik.
2. Menilai perbedaan efisiensi pekerjaan mahasiswa sebelum dan setelah menerapkan sistem kecerdasan buatan.
3. Menilai tingkat pemahaman mahasiswa tentang konsep dan aplikasi kecerdasan buatan dalam menyelesaikan tugas akademik.

## Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut:

1. Memberikan pemahaman tentang sejauh mana teknologi kecerdasan buatan dapat meningkatkan kualitas dalam menyelesaikan tugas akademik.
2. Mengidentifikasi perubahan efisiensi pekerjaan mahasiswa setelah menggunakan kecerdasan buatan, memberikan pemahaman tentang dampak teknologi terhadap hasil kerja.
3. Mengukur pemahaman mahasiswa terkait kecerdasan buatan dalam konteks penyelesaian tugas akademik, memberikan informasi tentang tingkat literasi teknologi mereka.

## Metodologi Penelitian

Dalam rangka menyusun penelitian ini, diperlukan data-data serta informasi yang lengkap sebagai bahan yang dapat mendukung kebenaran materi uraian dan pembahasan, oleh karena itu sebelum menyusun penelitian ini, dalam persiapannya terlebih dahulu dilakukan riset atau penelitian untuk menjaring data serta informasi atau bahan materi yang diperlukan.

Tahapan-tahapan yang berkaitan dengan metodologi yang di gunakan adalah sebagai berikut:

1. Kuesioner dan Wawancara

Kuesioner dan wawancara memungkinkan pengumpulan data langsung dari mahasiswa yang menggunakan AI. Data ini penting untuk memahami bagaimana AI memengaruhi kualitas, efisiensi, dan pemahaman tugas mereka, berdasarkan persepsi dan pengalaman pribadi. Ini memberikan insight langsung mengenai manfaat dan kendala yang mereka alami.

1. Analisis Hasil Tugas atau Studi Dokumentasi

Menganalisis hasil tugas mahasiswa sebelum dan sesudah menggunakan AI atau perbandingan antara kelompok yang menggunakan AI dan yang tidak, dapat memberikan gambaran objektif tentang kualitas dan efektivitas AI dalam menyelesaikan tugas. Data ini penting untuk menilai peningkatan kualitas dan efisiensi secara nyata.

1. Analisis Hasil Tugas atau Studi Dokumentasi

Menganalisis hasil tugas mahasiswa sebelum dan sesudah menggunakan AI atau perbandingan antara kelompok yang menggunakan AI dan yang tidak, dapat memberikan gambaran objektif tentang kualitas dan efektivitas AI dalam menyelesaikan tugas. Data ini penting untuk menilai peningkatan kualitas dan efisiensi secara nyata.

1. **Sistematika Penulisan**

Dalam penelitian ini, peneliti menyusun penelitian dengan cara menguraikan dalam beberapa bab terbagi beberapa bab yang secara terperinci menjelaskan isi dari masing-masing bab tersebut, sehingga tersusun dengan teratur. Sistematika penelitian tersebut antara lain:

**BAB I PENDAHULUAN**

Dalam bab ini diuraikan mengenai latar belakang, permasalahan, tujuan penelitian yang dilakukan, pembatasan masalah, manfaat tugas akhir, dan sistematika penelitian.

**BAB II LANDASAN TEORI**

Dalam bab ini menguraikan tentang landasan teori-teori dan prinsip -prinsip yang menjadi pedoman dan acuan peneliti dalam mengelola, menganalisa dan menyimpulkan pemecahan masalah yang berkaitan dengan pokok permasalahan yang di bahas.

**BAB III ANALISA DAN PERANCANGAN**

Dalam bab ini menjelaskan analalisa sistem dan perancangan sistem yang akan digunakan meliputi semua kebutuhan untuk membuat sistem yang akan di teliti sesuai dengan masalah yang akan di uji.

**BAB IV** **IMPLEMENTASI DAN PENGUJIAN**

Dalam bab ini peneliti menguraikan dari rancangan yang telah di buat dengan menggunakan hasil Analisa. Penjabaran dari perancangan sistem yang di maksud adalah hasil implementasi

**BAB V PENUTUP**

Bab ini merupakan penutup yang menerangkan tentang kesimpulan dari hasil penelitian penelitian ini dan mencoba memberikan saran-saran yang berguna bagi penerapan aplikasi di masa sekarang

**BAB II**

**TINJAUAN PUSTAKA**

1. **Penelitian Terkait**

Dalam penyusunan skripsi ini, penulis banyak terinspirasi dan merefrensi dari penelitian yang sudah ada sebelumnya yang masih berkaitan dengan latar belakang masalah, metodologi penelitian, serta teknologi yang digunakan pada skripsi ini, antara lain adalah:

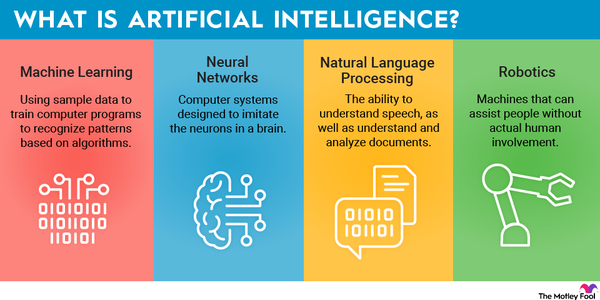
1. “Analisis Pengaruh Penggunaan AI *ChatGPT* terhadap Minat Baca Mahasiswa Sistem Informasi ITS” (Regina Dwi Aulia et al., 2024) Penelitian ini bertujuan untuk memahami pengaruh penggunaan *ChatGPT* terhadap kebiasaan membaca mahasiswa, dengan melakukan survei terhadap 113 responden yang mayoritas merupakan mahasiswa Departemen Sistem Informasi (82,3%) dan Departemen Inovasi Digital (17,7%). Hasil penelitian menunjukkan bahwa sebelum penggunaan *ChatGPT*, total nilai yang diperoleh adalah 2732 dengan rata-rata nilai 24,18 dan persentase rata-rata 80,59%, yang digolongkan dalam kategori "Rajin Membaca" (RM). Setelah penggunaan *ChatGPT*, total nilai menurun menjadi 2629 dengan rata-rata nilai 23,27 dan persentase rata-rata 77%, namun mayoritas mahasiswa masih berada dalam kategori "Rajin Membaca" (RM). Distribusi responden berdasarkan angkatan menunjukkan bahwa 54,9% berasal dari angkatan 2022, diikuti oleh angkatan 2023 (23,9%), angkatan 2021 (15,0%), dan angkatan 2020 (6,2%). Meskipun ada penurunan kecil dalam total dan rata-rata nilai setelah penggunaan *ChatGPT*, hal ini tidak signifikan dan tidak mengubah kategori kebiasaan membaca mayoritas mahasiswa. Analisis frekuensi membaca menunjukkan bahwa sebagian besar mahasiswa membaca 1-2 kali baik sebelum maupun setelah menggunakan *ChatGPT*, dengan peningkatan pada kategori "1-2 kali" setelah penggunaan *ChatGPT*.
2. “Dampak Penggunaan *Chatgpt* Pada Kompetensi Mahasiswa *Literature Review*” (Hidayanti & Azmiyanti, 2023). Penelitian ini bertujuan untuk

menjelaskan terkait dengananalisis dampak dari *ChatGPT* pada kompetensi mahasiswa akuntansi. Metode yang digunakan yakni metode *literature review* terhadap 480 artikel dari 500 artikel terkait. Sumber artikel didapatkan dari google scholar pada publish or perish, dengan pengolahan data menggunakan *Vosviewer*. Hasil dari penelitian ini yakni ancaman dan peluang yang disebabkan oleh adanya *ChatGPT* memberikan dampakpada kompetensi mahasiswa akuntansi terkait denganpenggunaan teknologi *ChatGPT* secara bertanggung jawab dan etis dalam dunia akademis. Pemahaman dan komprehensif mengenai dampak *ChatGPT* yang memberikan rekomendasi untuk perguruan tinggi dalam mengintegrasikan kebijakan agar menjadi sebuah pertimbangan etis terkait dengan kejujuran akademis serta ketergantungan berlebihan pada *ChatGPT* yang dapat menyebabkan berkurangnya kompetensi mahasiswa akuntansi. Oleh karena itu, penggunaan *ChatGPT* ini harus dilakukan dengan tanggung jawab dan berlandaskan dengan nilai etika yang kuat. Sehingga ancamandan peluang bisa mengoptimalkan manfaat dan risiko terkait dengan penggunaan *ChatGPT*. Penelitian ini diharapkan sebagai salah satu pertimbangan dalam pemanfaatan AI di bidang pendidikan.

1. **Landasan Teori**

Landasan teori ini mencakup berbagai konsep dan teori yang mendukung pemahaman lebih mendalam terkait penelitian yang dilakukan. Dalam bagian ini, akan diuraikan berbagai aspek yang penting seperti *machine learning, K-Means Clustering, Unsupervised Learning*, dan hal-hal yang berkaitan dengan penelitian ini yang menjadi dasar utama dalam penerapan metode yang digunakan untuk mengatasi permasalahan klasifikasi pada data-data yang telah dikembangkan. Melalui pembahasan ini, diharapkan dapat memberikan fondasi yang kuat untuk memahami bagaimana teknologi ini diaplikasikan dalam konteks yang spesifik, serta menjelaskan asumsi dan pendekatan teoritis yang diambil dari berbagai sumber referensi, seperti buku, jurnal, dan prosiding, yang relevan.

1. ***Artificial Intelligence***



*Gambar 2. 1 Artificial Intelligence*

*Artificial Intelligence* (AI) merupakan sistem yang dikembangkan dan terus berinovasi dalam bidang studi yang yang dibuat baik pada mesin ataupun komputer yang memiliki kecerdasan sama atau bahkan lebih seperti manusia. Teknologi *Artificial Intelligence* (AI) sering digunakan, tidak hanya satu atau dua saja platform *Artificial Intelligence* (AI) yang diciptakan manusia untuk memudahkan pekerjaan manusia. Namun, *Artificial Intelligence* (AI) memiliki berbagai jenis yang setiap platform memiliki tata guna dan fungsi yang berbeda, penulis kali ini ingin menjabarkan platform-platform yang sering digunakan mahasiswa dalam keberlangsungan hidup universitas menurut survey yang sudah penulis laksanakan. Metode penelitian yang digunakan untuk mengetahui fenomena ini adalah metode penelitian deskriptif kualitatif dengan teknik pengumpulan data kuesioner dengan memberikan pertanyaan dan dapat dijawab langsung oleh responden sesuai dengan jawaban mereka masing-masing. *Artificial Intelligence* (AI) berpotensi memainkan peran penting dalam mendukung mahasiswa dengan ketidakmampuan belajar dan membantu mereka mencapai potensi maksimal mereka. (Arly, Dwi, & Andini, 2023)

Dalam era yang dipenuhi dengan teknologi kecerdasan buatan (AI), penggunaannya dalam dunia pendidikan telah menjadi topik yang menarik perhatian. Meskipun AI menawarkan potensi yang besar, dampaknya juga menimbulkan tantangan dan peluang yang perlu dipahami dengan cermat oleh masyarakat. Pergeseran arah lapangan kerja akibat otomatisasi yang dilakukan oleh AI telah memunculkan pertanyaan tentang peran pekerja manusia di masa depan. engaruh kecerdasan buatan terhadap mahasiswajugatelah membawa perubahan signifikan dalam paradigma Pendidikan. Tujuan dari peneiltian ini adalah untuk mendeskripsikan penggunaan. AI dalam dunia Pendidikan bahasa Arab.Metode yang digunakan dalam penelitian ini adalah Studi Kepustakaan (Library Research). Studi kepustakaan digunakan untuk mengumpulkan informasi dari berbagai sumber seperti dokumen, buku, majalah, jurnal, dan sejarah. Hasil dari penelitian ini menunjukkan bahwa pemanfaatan teknologi AI telah menciptakan pengalaman belajar yang lebih personal, efisien, dan terfokus. Namun, tidak dapat dipungkiri bahwa kehadiran Al Tools juga telah mengubah beberapa aspek kehidupan mahasiswa. (Zakiyah, et al., 2024)

Ada tiga kategori utama dalam AI yang relevan untuk pembelajaran, yaitu:

1. *Supervised Learning*: Metode ini melibatkan pembelajaran dari data yang berlabel, di mana sistem dilatih untuk memetakan input ke output tertentu.
2. *Unsupervised Learning*: Pada metode ini, sistem bekerja dengan data yang tidak berlabel untuk menemukan pola atau struktur tersembunyi. Teknik ini sering digunakan untuk klasterisasi dan asosiasi data.
3. *Reinforcement Learning*: Metode ini hanya berfokus pada pembelajaran melalui proses coba-coba, di mana sistem belajar dari hasil Tindakan ini yang dilakukan untuk mencapai tujuan tertentu.
4. **Pengaruh *Artificial Intelligence* (AI)**

Menurut penulis kehadiran AI telah mengubah banyak aspek kehidupan manusia, terutama dalam dunia pendidikan. Penelitian terdahulu menunjukkan bahwa tidak ditemukan adanya ancaman keamanan data dalam penggunaan AI. Alat bantu AI memberikan berbagai pengaruh bagi mahasiswa, di antaranya adalah:

*Tabel 2. 1 Pengaruh AI*

|  |  |  |
| --- | --- | --- |
| Aspek | Pengaruh Positif terhadap AI | Pengaruh Negatif terhadap AI |
| Kualitas | Penggunaan AI membantu mahasiswa meningkatkan akurasi tugas dengan mengurangi kesalahan, mendorong kreativitas melalui ide dan inspirasi baru, serta memungkinkan analisis data besar secara cepat dan tepat. | Penggunaan AI dapat membuat mahasiswa cenderung malas melakukan riset mandiri dan meskipun tugas dapat dikerjakan dengan mudah, mahasiswa mungkin tidak sepenuhnya memahami proses penyelesaiannya. |
| Efisiensi | Penggunaan AI membantu mahasiswa menyelesaikan tugas kuliah lebih cepat dan memberikan jawaban yang tepat, mempermudah proses belajar. | Penggunaan AI dapat membuat mahasiswa malas membaca buku di perpustakaan dan menurunkan akurasi serta relevansi hasil karena ketergantungan pada efisiensi AI. |
| Pemahaman | Penggunaan AI memberikan pemahaman kepada mahasiswa dalam menyelesaikan tugas tanpa perlu pembimbing atau buku, serta memungkinkan pembelajaran yang lebih personal dan adaptif melalui platform e-learning yang menyesuaikan materi dengan kebutuhan dan kemampuan individu. | Penggunaan AI dapat menghambat individu dalam mengembangkan keterampilan yang relevan karena kurangnya bimbingan saat menghadapi masalah atau kejanggalan, serta memengaruhi sikap mahasiswa dalam bersikap optimis terhadap inovasi dan kreasi yang ada dalam diri mereka. |

1. **Penerapan *Artificial Intelligence* (AI) dalam Pendidikan**

Perkembangan AI dapat ditelusuri dari awal kemunculan nya sebagai sistem berbasis aturan sederhana hingga teknologi yang kini mampu belajar mandiri melalui analisis data yang kompleks. Dalam bidang pendidikan, AI mengalami perkembangan yang signifikan dengan munculnya aplikasi berbasis pembelajaran adaptif dan analisis perilaku pengguna.

1. Pembelajaran Adaptif

AI memungkinkan terciptanya sistem pembelajaran adaptif yang secara otomatis menyesuaikan materi pelajaran dan pendekatan pengajaran sesuai dengan kemampuan dan kebutuhan belajar masing-masing siswa. Sistem ini dapat mengidentifikasi kelemahan dan kekuatan siswa dan memberikan materi yang relevan untuk memperbaiki atau memperdalam pemahaman mereka.

1. Analisis Perilaku Pengguna

AI kini dapat menganalisis perilaku belajar siswa, seperti pola belajar, ketertarikan pada topik tertentu, dan waktu terbaik mereka dalam belajar. Melalui analisis ini, sistem dapat memberikan rekomendasi yang lebih personal untuk meningkatkan efisiensi dan efektivitas proses belajar.

1. Sistem Rekomendasi Materi

Dalam platform e-learning, AI digunakan untuk memberikan rekomendasi materi belajar tambahan atau soal latihan yang sesuai dengan kemampuan dan minat siswa. Misalnya, ketika siswa menunjukkan ketertarikan atau kesulitan pada topik tertentu, sistem akan merekomendasikan sumber belajar tambahan terkait.

1. Pembelajaran Berbasis Game (*Gamification*)

Dengan AI, pembelajaran berbasis game kini semakin interaktif dan responsif. AI mampu menyesuaikan tingkat kesulitan permainan atau soal sesuai dengan kemampuan siswa dan menciptakan pengalaman belajar yang menyenangkan dan menantang.

1. Penggunaan *Chatbot* sebagai Asisten Pembelajaran

*Chatbot* berbasis AI kini telah banyak digunakan dalam pendidikan sebagai asisten belajar. *Chatbot* dapat membantu siswa menjawab pertanyaan dasar mengenai materi pelajaran, mengingatkan tenggat waktu tugas, atau memberikan informasi terkait kelas.

1. Analitik Prediktif untuk Pengembangan Karier

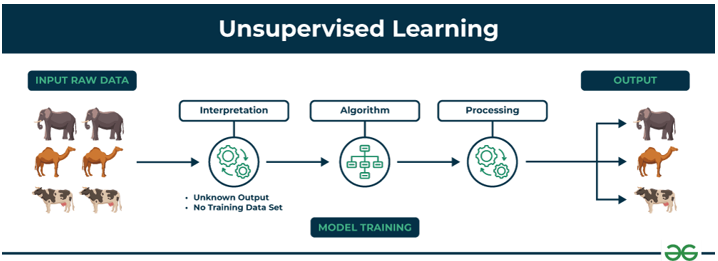
AI juga diterapkan untuk memberikan prediksi mengenai jalur karier yang paling sesuai dengan minat, keterampilan, dan performa belajar siswa. Sistem ini dapat membantu mahasiswa merencanakan pendidikan dan karier mereka dengan lebih baik.

1. ***Unsupervised Learning***

Menurut (Mitchell, 1997) mendefinisikan *unsupervised* *learning* sebagai jenis pembelajaran di mana "hanya input yang diberikan, dan model diminta untuk menemukan struktur dalam data." Mitchell menjelaskan bahwa tantangan utama dalam *unsupervised* *learning* adalah tidak adanya label yang membuat model harus mempelajari pola secara mandiri. Menurutnya, metode ini sering digunakan dalam clustering, pengurangan dimensi, dan analisis asosiasi. Mitchell juga menekankan pentingnya *unsupervised* *learning* dalam situasi di mana memperoleh label untuk data sangat sulit atau mahal. Hal ini sering terjadi dalam domain seperti pemrosesan gambar atau teks.

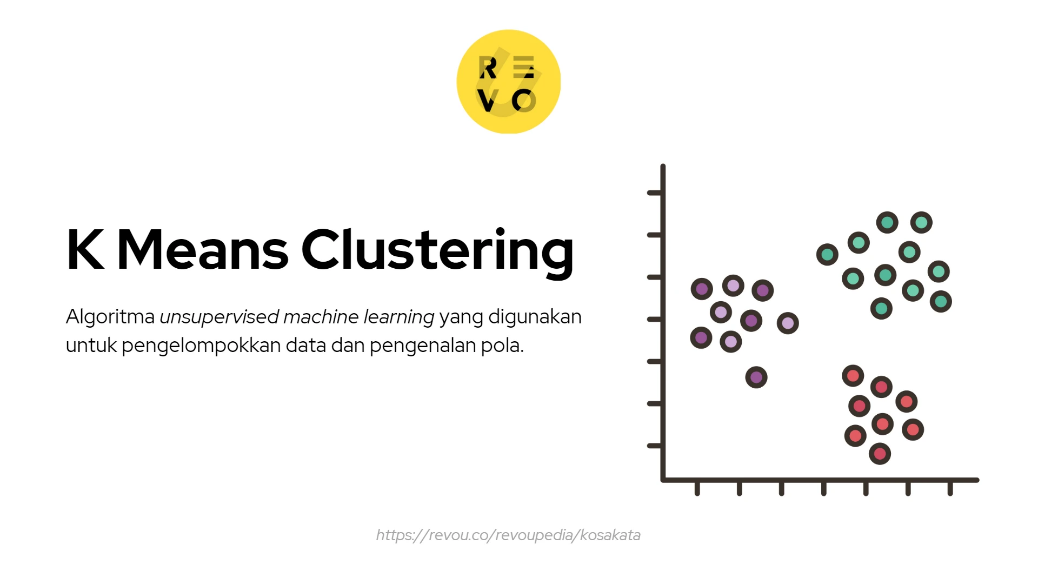
Menurut ahli Charu Aggarwal, menyatakan bahwa *unsupervised* *learning* adalah inti dari banyak algoritma *data* *mining* dan *big* *data* *analytics*. Menurut Aggarwal, metode ini sangat penting untuk pengeksplorasian dan analisis awal data, terutama ketika kita tidak memiliki informasi tentang label data atau klasifikasi sebelumnya.

Aggarwal juga menyebutkan bahwa metode seperti *clustering*, *dimensionality* *reduction*, dan deteksi anomali adalah bagian penting dari *unsupervised* *learning*. Ia menekankan pentingnya pemilihan metrik yang tepat (seperti jarak *Euclidean* atau *Cosine*) untuk mendapatkan hasil *clustering* yang akurat dalam berbagai aplikasi industri dan akademik.



*Gambar 2. 2 Unsupervised Learning*

1. **K-Means Clustering**



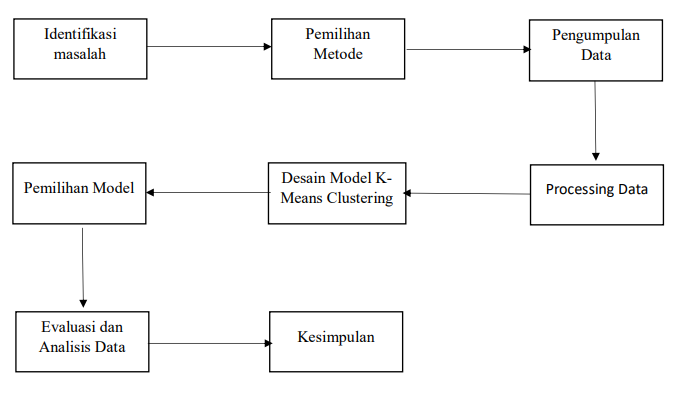
*Gambar 2. 3 K-Means Clustering*

*K-Means Clustering* adalah algoritma *clustering* yang paling banyak digunakan dalam *unsupervised learning*, dan banyak ahli telah memberikan penjelasan mendalam mengenai cara kerja, aplikasi, serta kelebihan dan kekurangannya. Berikut adalah beberapa penjelasan menurut para ahli:

Menurut (Dinata, Safwandi, Hasdyna, & Azizah, 2020), *K-Means* adalah salah satu algoritma dalam data mining yang digunakan untuk melakukan pengelompokan atau *clustering* data, seperti data sepeda motor. Tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif dalam proses pengelompokan, yaitu dengan mengurangi variasi dalam satu kelompok dan meningkatkan variasi antar kelompok.

Kemudian menurut penulis *K-Means Clustering* adalah metode analisis data atau *data mining* yang melakukan pemodelan tanpa supervisi (*unsupervised*) dan menggunakan sistem partisi untuk pengelompokan data. Metode ini berupaya mengelompokkan data ke dalam beberapa kelompok, di mana data dalam satu kelompok memiliki karakteristik yang sama, sementara karakteristiknya berbeda dengan data dalam kelompok lainnya.

1. **Kerangka Pemikiran**



*Gambar 2. 4 Kerangka Pemikiran*

Berikut penjelasan alur yang ada pada diagram diatas:

1. Identifikasi Masalah: Langkah pertama adalah mengidentifikasi masalah yang perlu diselesaikan.
2. Pemilihan Metode: Setelah masalah teridentifikasi, langkah selanjutnya adalah memilih metode yang akan digunakan untuk menyelesaikan masalah tersebut.
3. Pengumpulan Data: Setelah metode dipilih, langkah berikutnya adalah mengumpulkan data yang diperlukan untuk analisis.
4. Pemilihan Model: Setelah data terkumpul, langkah selanjutnya adalah memilih model yang tepat untuk analisis, dalam hal ini, model yang digunakan adalah *K-Means Clustering*.
5. Desain Model *K-Means Clustering*: Pada tahap ini, model *K-Means Clusterin*g dirancang untuk mengelompokkan data berdasarkan kesamaan.
6. *Processing Data*: Data yang telah dikumpulkan dan dipilih modelnya kemudian diproses agar dapat digunakan dalam model *K-Means Clustering*.
7. Evaluasi dan Analisis Data: Setelah model diterapkan, hasilnya dievaluasi dan dianalisis untuk mengetahui sejauh mana model memberikan hasil yang sesuai.
8. Kesimpulan: Tahap akhir adalah menarik kesimpulan dari hasil evaluasi dan analisis data untuk memberikan solusi terhadap masalah yang diidentifikasi sebelumnya.

# **BAB III ANALISA DAN PERANCANGAN**

## 3.1 Analisa Kebuthan

Analisa kebutuhan merupakan bentuk dari pendataan mengenai kebutuhan yang diperlukan dalam penelitian sehingga para prosesnya dapat berjalan dengan lancar. Pada Analisa kebutuhan dapat digunakan juga sebagai bentuk kepastian dalam berjalannya suatu sistem atau proyek yang sedang dikerjakan.

### **3.1.1 Analisa Kebutuhan Fungsional**

Kebutuhan fungsional ini meliputi berbagai fitur dan kemampuan yang harus ada untuk memastikan bahwa sistem mampu menjalankan fungsinya sesuai dengan kebutuhan dan tujuan penelitian.

1. Klasifikasi Penggunaan AI: Sistem harus mampu melakukan klasifikasi faktor dari penggunaan sebuah AI menjadi beberapa model K-means Learning yang dilatih.
2. Pengelompokan Data: Sistem harus dapat melakukan preprocessing data, seperti menggunakan K-means Learning yang digunakan untuk mengelompokkan data ke dalam sejumlah kelompok berdasarkan kemiripan.
3. Pelatihan Model: Sistem harus menyediakan fungsi untuk melatih model

Unsupervised Learning menggunakan dataset yang telah diproses.

1. Evaluasi Model: Sistem harus dapat mengevaluasi model yang telah dilatih, termasuk seperti seperti silhouette score (untuk clustering) atau reconstruction error (untuk autoencoder). serta menampilkan hasil dalam bentuk grafik atau laporan.

### **3.1.2 Analisa Kebutuhan Non-Fungsional**

Kebutuhan non-fungsional ini berkaitan dengan spesifikasi perangkat keras dan perangkat lunak yang diperlukan agar sistem dapat berjalan dengan optimal dan mendukung seluruh fungsi yang diinginkan.

1. **Spesifikasi Perangkat Keras**
2. Prosesor: Dibutuhkan prosesor Intel® Core™ i3-4005U CPU dengan kecepatan 1.70GHz dan empat inti (CPUs). Prosesor ini mampu menangani tugas-tugas dasar dengan performa yang cukup baik untuk penggunaan umum.
3. Memori (RAM): Memori yang dibutuhkan minimal 2 GB, yang cukup untuk menjalankan sistem operasi dan aplikasi pemrograman ringan seperti Python.
4. Penyimpanan: Sistem harus dilengkapi dengan penyimpanan SSD minimal 125 GB untuk memastikan kecepatan akses data yang lebih tinggi dibandingkan dengan hard drive tradisional, serta cukup ruang untuk instalasi perangkat lunak dan penyimpanan file proyek.
5. **Spesifikasi Perangkat Lunak**
6. Sistem Operasi: Sistem operasi yang diperlukan adalah Windows 10 atau 11, yang memberikan stabilitas dan kompatibilitas dengan berbagai perangkat lunak pengembangan dan aplikasi modern.
7. Bahasa Pemrograman: Python akan digunakan sebagai bahasa pemrograman utama. Python adalah bahasa pemrograman yang mudah dipelajari dan digunakan untuk berbagai aplikasi, termasuk pengembangan perangkat lunak, analisis data, dan kecerdasan buatan.
8. Library dan Framework: Untuk pengembangan lebih lanjut, Anda dapat menggunakan berbagai pustaka dan framework Python sesuai dengan kebutuhan proyek, seperti NumPy untuk perhitungan numerik, Pandas untuk manipulasi data, atau TensorFlow untuk kecerdasan buatan. (Sebutkan jika ada pustaka tertentu yang dibutuhkan.
9. IDE (Integrated Development Environment): Disarankan untuk menggunakan Visual Studio Code atau Google Colab sebagai IDE. Visual Studio Code adalah editor kode sumber yang ringan namun kaya fitur, mendukung Python dan berbagai ekstensi lainnya. Google Colab adalah platform berbasis cloud yang memungkinkan Anda untuk menulis dan menjalankan kode Python tanpa perlu instalasi lokal, cocok untuk eksperimen dan proyek berbasis data.

## 3.2 Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini menggunakan pendekatan metode kuantitatif, meneurut (Sugiyono, 2020) Penelitian kuantitatif adalah bahwa gejala dari suatu obyek itu sifatnya tunggal dan parsial. Dengan demikian berdasarkan gejala tersebut peneliti kuantitatif dapat menentukan variabel- variabel yang akan diteliti. Dalam pandangan penelitian kualitatif, gejala itu bersifat holistik (menyeluruh, tidak dapat dipisah-pisahkan), sehingga peneliti kualitatif tidak akan menetapkan penelitannya hanya berdasarkan variabel penelitian, tetapi keseluruhan situasi sosial yang diteliti yang meliputi aspek tempat (place), pelaku (actor) dan aktivitas (activity) yang berinteraksi secara sinergis. Pada dasarnya penelitian ini akan memperdalam mengenai analisis untuk mengevaluasi kinerja dari model Unsupervised Learning dalam mengklasifikasikan pengaruh dari penggunaan Artifical Intellegence (AI) terhadap kualitas, efisiensi, dan pemahaman dalam penyelesaian tugas mahasiswa. Langkah-langkah penelitian meliputi:

1. Observasi

Berdasarkan pengumpulan data yang diberikan suatu kumpulan data dalam suatu tabel yang nantinya akan diteliti sehingga pada tahap ini berfokus pada memperoleh informasi yang ada. Observasi dapat dilakukan dengan cara membaca teks secara cermat dan mencatat hal-hal yang relevan.

1. Pengelompokkan Data

Sistem harus dapat melakukan preprocessing data, seperti menggunakan K means Learning yang digunakan untuk mengelompokkan data ke dalam sejumlah kelompok berdasarkan kemiripan.

1. Pengembangan Model Unsupervised Learning

Melibatkan pendekatan di mana model belajar dari data yang tidak diberi label. Tujuan utama unsupervised learning adalah untuk menemukan pola atau struktur yang belum pasti dalam data.

1. Pelatihan Model Unsupervised Learning

Proses di mana model dilatih untuk menemukan pola atau struktur dalam data tanpa label. Karena tidak ada label yang digunakan, proses pelatihan model unsupervised bergantung sepenuhnya pada karakteristik data yang diolah.

1. Evaluasi Model

Mengukur kinerja model dengan menggunakan analisis Silhouette Score serta analisis deskriptif untuk memastikan evaluasi model yang komprehensif tanpa ketergantungan pada label kebenaran (*ground truth*).

1. Analisis Hasil

Menganalisis hasil evaluasi untuk menentukan keberhasilan metode Unsupervised Learning dalam klasifikasi pengaruh Artifical Intellegence (AI) terhadap kualitas, efisiensi, dan pemahaman dalam penyelesaian tugas mahasiswa.

## 3.3 Perancangan Penelitian

Dalam analisis data ini, digunakan data dari hasil kuesioner yang telah disebarkan kepada 35 orang mahasiswa sebagai responden. Para responden dipilih secara acak dari berbagai program studi untuk memperoleh gambaran yang lebih umum mengenai penggunaan kecerdasan buatan (AI) dalam penyelesaian tugas akademik. Data yang digunakan dalam penelitian ini diperoleh melalui kuesioner terstruktur yang terdiri dari sejumlah pertanyaan tertutup berbasis skala Likert (1–5), serta beberapa pengukuran numerik.menururt (Sugiyono, 2020) Skala Likert digunakan untuk mengukur sikap, pendapat, dan persepsi seseorang atau sekelompok orang tentang fenomena sosial. Variabel-variabel dalam kuesioner dirancang untuk menggambarkan berbagai aspek penggunaan kecerdasan buatan (AI) oleh mahasiswa. Berikut adalah penjabaran variabel dalam kuesioner:

Tabel 3. 1 Penjabaran Variabel Dalam Kuesioner

|  |  |  |  |
| --- | --- | --- | --- |
| NO | Pertanyaan | Kategori | Tujuan |
| 1 | Seberapa sering Anda menggunakan AI (ChatGPT, Gemini, Copilot, dll.) untuk menyelesaikan tugas? | Frekuensi\_AI | Mengukur seberapa sering mahasiswa menggunakan AI dalam menyelesaikan tugas. Nilai skala menggambarkan tingkat intensitas, mulai dari "tidak pernah" hingga "sangat sering" |
| 2 | Seberapa sering AI membantu Anda dalam Mencari Referensi Tugas? | AI\_Referensi | mengukur apakah mahasiswa menggunakan AIsecara umum, tetapi menggali aspek atau bentuk penggunaan AI secara rinci dan terfokus pada fungsi tertentu. |
| 3 | Seberapa sering AI membantu Anda dalam Menyusun Kerangka Tugas? | AI\_Kerangka |
| 4 | Seberapa sering AI membantu Anda dalam  Mengecek grammar/plagiarisme ? | AI\_Grammar |
| 5 | AI meningkatkan kualitas tugas saya (lebih akurat, rapi, orisinal). | Kualitas\_Tugas | persepsi mahasiswa terhadap kualitas tugas yang mereka hasilkan setelah menggunakan AI |
| 6 | Penggunaan AI mengurangi pemahaman mendalam saya terhadap materi. | Reduksi\_Pemahaman | Menilai apakah penggunaan AI mengurangi pemahaman mahasiswa terhadap materi yang seharusnya dipelajari secara mandiri. |
| 7 | Saya merasa  ketergantungan pada AI untuk menyelesaikan tugas. | Ketergantungan | Menunjukkan sejauh mana mahasiswa merasa tidak mampu atau kurang percaya diri mengerjakan tugas tanpa bantuan AI. |
| 8 | AI membuat saya lebih efisien (menghemat waktu). | Efisiensi | Mengukur sejauh mana AI membantu mahasiswa dalam menghemat waktu dan menyederhanakan proses pengerjaan tugas |
| 9 | Rata-rata waktu penyelesaian tugas tanpa AI (skala likert) sangat lambat – sangat cepat. | Waktu\_Tanpa\_AI | Dua variabel yang diukur menggunakan skala Likert 1–5 yang merepresentasikan persepsi mahasiswa terhadap efisiensi waktu dalam menyelesaikan tugas akademik, baik ketika menggunakan AI maupun tidak. |
| 10 | Rata-rata waktu penyelesaian tugas dengan AI (skala likert) sangat lambat – sangat cepat. | Waktu\_Dengan\_AI |
| 11 | Rata-rata nilai tugas sebelum menggunakan AI | Nilai Sebelum AI | Merupakan data kuantitatif berupa nilai akademik mahasiswa untuk tugas-tugas sebelum dan sesudah penggunaan AI. Data ini digunakan untuk mengukur perubahan objektif dalam kualitas tugas. |
| 12 | Rata-rata nilai tugas setelah menggunakan AI | Nilai Sesudah AI |

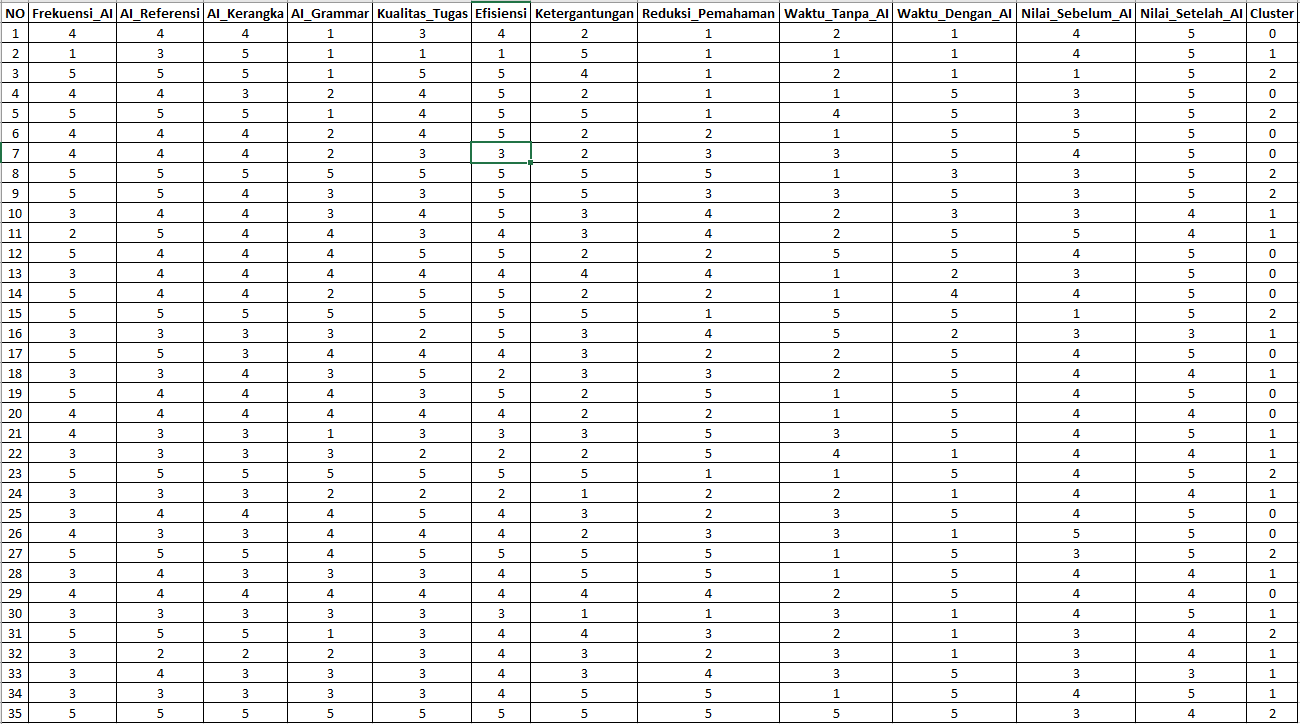
Data dari kuesioner ini kemudian dianalisis menggunakan metode *clustering* K-Means, yang merupakan pendekatan *unsupervised learning* dalam machine learning. Analisis ini dilakukan dengan bantuan bahasa pemrograman Python, untuk mengelompokkan responden berdasarkan pola jawaban mereka terhadap variabel-variabel terkait penggunaan AI. Metode ini memungkinkan identifikasi segmen atau kelompok mahasiswa dengan karakteristik penggunaan AI yang mirip, tanpa perlu label sebelumnya.

### **3.3.1 Pengelompokan Data**

Tahap ini bertujuan untuk memahami struktur, isi, dan kualitas data yang dikumpulkan melalui kuesioner. menurut (Berampu et al., 2022) Data adalah bentuk jamak dari datum. Data merupakan kumpulan fakta atau angka atau segala sesuatu yang dapat dipercaya kebenarannya sehingga dapat digunakan sebagai dasar menarik suatu Kesimpulan. Data yang digunakan terdiri atas beberapa variabel yang mencerminkan dimensi penggunaan AI dalam konteks akademik, seperti frekuensi penggunaan AI, penggunaan AI untuk referensi, kerangka tugas, dan grammar, serta persepsi mahasiswa terhadap kualitas tugas, efisiensi, dan pemahaman materi. Selain itu, data mencakup nilai akademik dan waktu pengerjaan tugas sebelum dan sesudah menggunakan AI.

Pemahaman ini diperoleh melalui eksplorasi awal data, pemeriksaan deskriptif statistik, serta identifikasi adanya nilai ekstrem atau data yang tidak konsisten. Proses ini juga mencakup pemberian label variabel dan standarisasi istilah agar sesuai dengan kebutuhan analisis machine learning. Tujuan utama pada tahap ini adalah memastikan bahwa data yang tersedia representatif, bersih, dan relevan untuk proses pemodelan berikutnya.

Tabel 3. 2 Tabulasi Data Hasil Kuesioner



Memeriksa 35 data responden memastikan data yang di input tidak ada missing values semua responden telah mengisi skala likert 1 – 5 serta konsistensi nilai sudah di periksa dan diperbaiki.

### **3.3.2 Pengembangan model unsupervised learning**

Data tanpa label (unlabeled data) merupakan jenis data yang hanya mengandung variabel prediktor (X) tanpa disertai variabel target (y) atau label kategori yang telah terdefinisi sebelumnya. Dalam paradigma pembelajaran mesin, data semacam ini menjadi dasar utama untuk pendekatan unsupervised learning. (Nurhalizah et al., 2024) mengatakan Unsupervised learning merupakan metode dalam machine learning yang mengaplikasikan algoritma untuk menganalisis dan menemukan pola dalam data tanpa intervensi atau bantuan manusia.

Karakteristik utama data tanpa label adalah ketiadaan informasi kebenaran referensi (ground truth) yang biasanya berperan sebagai pembimbing dalam proses pembelajaran. Tanpa adanya label yang telah ditentukan sebelumnya, sistem pembelajaran dituntut untuk secara mandiri mengidentifikasi dan mengekstraksi pola-pola intrinsik yang terkandung dalam struktur data.

Dataset yang di sajikan merupakan prototipe klasik *unlabeled data* dalam pendekatan *unsupervised learning*, dimana:

* Variabel prediktor (X): Terdiri dari 12 atribut perilaku penggunaan AI

Tabel 3. 3 Tabel Atribut Perilaku Penggunaan AI

|  |  |  |
| --- | --- | --- |
| NO | Pertanyaan | Kategori |
| 1 | Seberapa sering Anda menggunakan AI (ChatGPT, Gemini, Copilot, dll.) untuk menyelesaikan tugas? | Frekuensi\_AI |
| 2 | Seberapa sering AI membantu Anda dalam Mencari Referensi Tugas? | AI\_Referensi |
| 3 | Seberapa sering AI membantu Anda dalam Menyusun Kerangka Tugas? | AI\_Kerangka |
| 4 | Seberapa sering AI membantu Anda dalam  Mengecek grammar/plagiarisme ? | AI\_Grammar |
| 5 | AI meningkatkan kualitas tugas saya (lebih akurat, rapi, orisinal). | Kualitas\_Tugas |
| 6 | Penggunaan AI mengurangi pemahaman mendalam saya terhadap materi. | Reduksi\_Pemahaman |
| 7 | Saya merasa  ketergantungan pada AI untuk menyelesaikan tugas. | Ketergantungan |
| 8 | AI membuat saya lebih efisien (menghemat waktu). | Efisiensi |
| 9 | Rata-rata waktu penyelesaian tugas tanpa AI (skala likert) sangat lambat – sangat cepat. | Waktu\_Tanpa\_AI |
| 10 | Rata-rata waktu penyelesaian tugas dengan AI (skala likert) sangat lambat – sangat cepat. | Waktu\_Dengan\_AI |
| 11 | Rata-rata nilai tugas sebelum menggunakan AI | Nilai Sebelum AI |
| 12 | Rata-rata nilai tugas setelah menggunakan AI | Nilai Sesudah AI |

* variabel target (y): Tidak terdapat kolom label seperti "Kategori\_Mahasiswa" atau "Tingkat\_Keberhasilan" yang telah terdefinisi secara apriori

Data tanpa label dalam penelitian ini berhasil mengungkap *complex adaptive system* dalam penggunaan AI akademik, dimana pola penggunaan mahasiswa mengkristal menjadi tiga *attractor state* tanpa intervensi label awal. Temuan ini memperkuat proposisi bahwa *learning analytics* berbasis *unsupervised learning* dapat mengungkap *latent learning pattern* yang tidak teramati dalam pendekatan tradisional.

### **3.3.3 Pelatihan Model Unsupervised Learning**

Melibatkan pendekatan di mana model belajar dari data yang tidak diberi label. Tujuan utama unsupervised learning adalah untuk menemukan pola atau struktur yang belum pasti dalam data.

#### 3.3.3.1 Normalisasi Data ( Z-Score )

Standarisasi Z-score merupakan transformasi linear yang mengkonversi data mentah ke dalam skala relatif terhadap mean dan deviasi standar distribusinya. Dalam konteks analisis multivariat seperti penelitian penggunaan AI mahasiswa, proses ini penting untuk:

* Menyamakan skala pengukuran antar variabel
* Meminimalkan bias akibat perbedaan range nilai
* Memungkinkan perhitungan jarak Euclidean yang bermakna

Z-score : **Z**

* + **Z =** Nilai Z-score setelah di standarisasi
  + **X =** Nilai asli dari data ke 1 – 35
  + **=** Rata-rata nilai dari data tersebut
  + **=**  Standar deviasi dari kolom tersebut

Frekuensi\_AI “35 Mahasiswa” = [4, 1, 5, 4, 5, 4, 4, 5, 5, 3, 2, 5, 3, 5, 5, 3, 5, 3, 5, 4, 4, 3, 5, 3, 3, 4, 5, 3, 4, 3, 5, 3, 3, 3, 5]

Z =

Hitung standar Deviasi

=

Diketahui Data : [4, 1, 5, 4, 5, 4, 4, 5, 5, 3, 2, 5, 3, 5, 5, 3, 5, 3, 5, 4, 4, 3, 5, 3, 3, 4, 5, 3, 4, 3, 5, 3, 3, 3, 5].

Rata-Rata ( ) : 3.88

Jumlah Data ( N ) : 35

Setiap nilai dikurangi dengan ( ) 3.88, Lalu di kuadratkan.

M1 = ( 4 – 3.88 )² = ( 0.12 )² = 0.0144

M2 = (1−3.88)2=(−2.88)² = 8.2944

M3 = (5−3.88)² =(1.12)² = 1.2544

M4 = ( 4 – 3.88 )² = ( 0.12 )² = 0.0144

M5 = (5−3.88)² =(1.12)² = 1.2544

M6 = ( 4 – 3.88 )² = ( 0.12 )² = 0.0144

M7 = ( 4 – 3.88 )² = ( 0.12 )² = 0.0144

M8 = (5−3.88)² =(1.12)² = 1.2544

M9 = (5−3.88)² =(1.12)² = 1.2544

M10 = (3−3.88)² =(−0.88)² =0.7744

M11 = (2−3.88)² =(−1.88)² =3.5344

M12 = (5−3.88)² =(1.12)² = 1.2544

M13 = (3−3.88)² =(−0.88)² =0.7744

M14 = (5−3.88)² =(1.12)² = 1.2544

M15 = (5−3.88)² =(1.12)² = 1.2544

M16 = (3−3.88)² =(−0.88)² =0.7744

M17 = (5−3.88)² =(1.12)² = 1.2544

M18 = (3−3.88)² =(−0.88)² =0.7744

M19 = (5−3.88)² =(1.12)² = 1.2544

M20 = ( 4 – 3.88 )² = ( 0.12 )² = 0.0144

M21 = ( 4 – 3.88 )² = ( 0.12 )² = 0.0144

M22 = (3−3.88)² =(−0.88)² =0.7744

M23 = (5−3.88)² =(1.12)² = 1.2544

M24 = (3−3.88)² =(−0.88)² =0.7744

M25 = (3−3.88)² =(−0.88)² =0.7744

M26 = ( 4 – 3.88 )² = ( 0.12 )² = 0.0144

M27 = (5−3.88)² =(1.12)² = 1.2544

M28 = (3−3.88)² =(−0.88)² =0.7744

M29 = ( 4 – 3.88 )² = ( 0.12 )² = 0.0144

M30 = (3−3.88)² =(−0.88)² =0.7744

M31 = (5−3.88)² =(1.12)² = 1.2544

M32 = (3−3.88)² =(−0.88)² =0.7744

M33 = (3−3.88)² =(−0.88)² =0.7744

M34 = (3−3.88)² =(−0.88)² =0.7744

M35 = (5−3.88)² =(1.12)² = 1.2544

Jumlahkan semua =

0.0144+8.2944+1.2544+0.0144+1.2544+0.0144+0.0144+1.2544+1.2544+0.7744+3.5344+1.2544+0.7744+1.2544+1.2544+0.7744+1.2544+0.7744+1.2544+0.0144+0.0144+0.7744+1.2544+0.7744+0.7744+0.0144+1.2544+0.7744+0.0144+0.7744+1.2544+0.7744+0.7744+0.7744+1.2544

= **32.88​** Total = 32.88

Hitung Varians ( ) = Ambil akar kuadrat dari Varians

= **0.9692**

Hitung Z-Score Mahasiswa pertama :

Rata-Rata ( ) : 3.88

= **0.9692**

= = = ***0.4147***

Maka hasi dari Z-score ke 35 data mahasiswa dari fitur Frekuensi AI adalah “***0.4147***”.

Berdasarkan analisis data terhadap 35 mahasiswa, nilai z-score 0.4147 untuk frekuensi penggunaan AI dapat dikategorikan sebagai valid dan siap diproses lebih lanjut.

Pada sampel 35 mahasiswa:

* + Mean (μ) frekuensi AI = 3.8
  + σ = 1.47
  + Nilai mentah (x) = μ + zσ = 3.8 + (0.4147×1.47) ≈ 4.41
    - Masuk dalam range skala Likert 1-5 (valid)

#### 3.3.3.2 Proses Clustering K-means

Algoritma K-Means diterapkan untuk mengelompokkan data mahasiswa menjadi tiga (3) cluster, dimana :

Cluster 0 = Pengguanaan AI Seimbang Optimal

Cluster 1 = Pengguna Minimal

Cluster 2 = Penggunaan AI Intensif dan beresiko ketergantungan

Titik centroid dari masing-masing cluster dihitung secara otomatis melalui metode *k-means++*, yang meminimalisasi risiko inisialisasi centroid yang buruk, Nilai centroid akhir merepresentasikan karakteristik rata-rata dari mahasiswa dalam tiap kelompok. Nilai ini kemudian digunakan untuk menghitung jarak masing-masing mahasiswa terhadap setiap centroid, sehingga penentuan cluster menjadi objektif berdasarkan kedekatan terhadap pusat kelompok.

Jarak Euclidean adalah jarak garis lurus antara dua titik dalam ruang berdimensi banyak (multidimensi). Dalam konteks clustering, ini digunakan untuk mengukur seberapa dekat satu data mahasiswa ke pusat (centroid) cluster.

Rumus Euclidean distance

)²

Dimana :

= nilai ke – I dari mahasiswa

= nilai ke – I dari centroid

= Jumlah fitur (Frekuensi Ai dan lainnya)

Mahasiswa ke 1 untuk centroid 1:

Mahasiswa ke – 1 = [4, 4, 4, 1, 3, 4, 2, 1, 2, 1, 4, 5]

Centroid -1 (C0) : [3, 3, 3, 2, 3, 3, 3, 2, 2, 2, 3, 5]

Centroid -2 (C1) : [4, 4, 4, 3, 4, 4, 4, 3, 3, 3, 4, 5]

Centroid -3 (C2) : [5, 5, 5, 4, 5, 5, 5, 4, 4, 4, 5, 5]

Perhitungan jarak Euclidean :

Jarak ke centroid 1 (C0) :

= = = *2.83*

Jarak ke centroid 2 (C1) :

= = = *4.24*

Jarak ke centroid 3 (C2) :

= =

= *7.00*

Jarak mahasiswa ke -1 ke setiap centroid :

Tabel 3. 4 Tabel Nilai Jarak Euclidean Antar Cluster

|  |  |  |
| --- | --- | --- |
| Centroid | Jarak Euclidean | Keterangan |
| C0 | 2.83 | Penggunaan AI seimbang optimal |
| C1 | 4.24 | Pengguna Minimal |
| C2 | 7.00 | Penggunaan AI intensif beresiko ketergantungan |

Mahasiswa ke -1 paling dekat dengan Centroid -1 (C0), maka berdasarkan table mahasiswa ke -1 termasuk kedalam cluster C0 ‘pengguna AI seimbang optimal’.

Namun hasil ini adalah sebagai contoh perhitungan secara manual untuk membuktikan bahwa proses clustering telah meliputi alur yang sesuai dan valid, pada proses ini peneliti telah membuat tabel hasil clustering dari program yang telah dibuat sebagai berikut :

Tabel 3. 5 Hasil Clustering

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Cluster | Jarak cluster 0 | Jarak Cluster 1 | Jarak Cluster 2 |
| 1 | 0 | 3.1248 | 3.4366 | 4.4530 |
| 2 | 1 | 6.4495 | 5.0923 | 7.6166 |
| 3 | 2 | 5.2329 | 6.4731 | 3.5326 |
| 4 | 0 | 2.6238 | 4.4020 | 4.1067 |
| 5 | 2 | 4.2161 | 5.8691 | 2.5814 |
| 6 | 0 | 2.1952 | 4.5773 | 4.1430 |
| 7 | 0 | 2.2355 | 3.1052 | 4.1217 |
| 8 | 2 | 4.0233 | 6.4580 | 2.5984 |
| 9 | 2 | 3.0288 | 5.0119 | 1.9756 |
| 10 | 0 | 2.2329 | 3.3963 | 3.4799 |
| 11 | 0 | 3.0828 | 4.2077 | 4.9065 |
| 12 | 0 | 3.3055 | 5.2708 | 3.7351 |
| 13 | 0 | 2.4593 | 3.8649 | 3.4219 |
| 14 | 0 | 2.4120 | 4.9203 | 3.4392 |
| 15 | 2 | 5.5916 | 7.2313 | 3.5277 |
| 16 | 1 | 4.9196 | 3.4364 | 6.2080 |
| 17 | 0 | 2.2633 | 4.7141 | 3.3199 |
| 18 | 0 | 1.5240 | 3.3828 | 2.6411 |
| 19 | 0 | 2.5351 | 4.8147 | 3.9633 |
| 20 | 0 | 1.9900 | 3.9421 | 3.9503 |
| 21 | 1 | 3.2863 | 3.1178 | 5.2711 |
| 22 | 1 | 4.5939 | 2.2210 | 6.6379 |
| 23 | 2 | 3.8466 | 6.6261 | 2.7053 |
| 24 | 1 | 4.4236 | 2.1261 | 6.7390 |
| 25 | 0 | 2.0569 | 4.2167 | 3.5013 |
| 26 | 0 | 2.9589 | 3.4520 | 5.1582 |
| 27 | 2 | 3.7784 | 6.4220 | 2.3153 |
| 28 | 0 | 2.9986 | 3.7410 | 4.6695 |
| 29 | 0 | 1.9143 | 3.8136 | 3.2752 |
| 30 | 1 | 3.6629 | 2.6450 | 5.8972 |
| 31 | 2 | 3.9723 | 4.5528 | 3.3891 |
| 32 | 1 | 4.4655 | 2.5195 | 6.3805 |
| 33 | 1 | 3.6585 | 3.2798 | 5.1956 |
| 34 | 0 | 3.0393 | 3.8247 | 4.9131 |
| 35 | 2 | 3.8339 | 6.1832 | 2.9109 |

Tabel diatas telah memberikan hasil dari proses clutering yang telah mengalami proses iterasi hingga hasilnya menjadi *konvergen* (tidak berubah-ubah), setiap cluster meliputi karaketeristik yang berbeda-beda diantaranya cluster 0 menunjukan penggunaan AI yang optimal, Cluster 1 menunjukan pengguna AI minimum dan Cluster 2 menunujukan pengguna yang intensif dan bersiko ketergantungan dengan AI.

### **3.3.4 Evaluasi Model**

Mengukur kinerja model dengan menggunakan metrik seperti akurasi, persisi, recall, score, dan analisis silhouette score.

#### 3.3.4.1 Metode Elbow ( Elbow-Method )

Metode Elbow digunakan untuk menentukan jumlah cluster optimal dalam analisis klaster dengan memvisualisasikan hubungan antara jumlah cluster (k) dan Within-Cluster Sum of Squares (WCSS). WCSS mengukur total variasi dalam setiap cluster. Semakin kecil WCSS, semakin padat (homogen) clusternya. Titik "siku" (elbow) pada grafik WCSS vs. \*k\* menunjukkan jumlah cluster optimal, di mana penambahan cluster tidak lagi mengurangi WCSS secara signifikan.

WCSS (Within-Cluster Sum of Squares) dihitung sebagai:

WCSS =

*k*: Jumlah cluster.

*Ci*​: Data points dalam cluster ke-i*i*.

Jarak(x,Centroidi)Jarak(*x*,Centroid*i*​): Jarak Euclidean antara data point *x* dan centroid cluster ke-*i*.

di mana ​ adalah cluster ke- , adalah centroid cluster, dan adalah titik data dalam cluster tersebut.

Identifikasi Centroid dari setiap cluster :

- Algoritma KMeans menghitung rata-rata posisi data di setiap cluster untuk menentukan centroid.

- Untuk k=3, ada tiga centroid ​.

Hitung jarak kuadrat antara setiap titik data dalam cluster ke centroid cluster tersebut :

=

Hitung WCSS Per-Cluster :

Cluster 0 Ber-Anggotan 18 data,

Nomor : 1, 4, 6, 7, 10, 11, 12, 13, 14, 17, 18, 19, 20, 25, 26, 28, 29, 34 (18 data)

Cluster 0 =

Cluster 1 Ber-Anggotakan 8 data,

Nomor : 2, 16, 21, 22, 24, 30, 32, 33 (8 data).

Cluster 1 =

Cluster 2 Ber-Anggotakan 9 data,

Nomor :  3, 5, 8, 9, 15, 23, 27, 31, 35 (9 data).

Cluster 2 =

Total WCSS = 121.27 + 81.01 + 74.96 = 277.24

Jika dilakukan iterasi untuk *k*=1,2,3, grafik WCSS vs. *k* akan menunjukkan penurunan tajam dari *k*=1 ke *k*=3, Titik siku (elbow) terlihat pada k=3, karena penambahan cluster ke-4 tidak akan mengurangi WCSS secara signifikan. Jumlah cluster optimal adalah 3.

#### 3.3.4.2 PCA (Principal Component Analysis)

PCA (principal Component Analysis) adalah metode untuk mengubah fitur asli menjadi fitur baru yang sering disebut komponen utama yaitu : PCA1, PCA2, PCA 3 dan seterusnya. Dalam menentukan PCA kita perlu menstandarisasi data agar semua fitur memiliki skala yang sama, Langkah – Langkah :

Standarisai data menggunakan Z-score : **Z**

Hitung matriks Kovarian :

Hitung eigenvalue dan eigenvector dari matriks kovarians:

* Eigenvector → arah komponen utama (PCA1, PCA2, dst)
* Eigenvalue → besar varians di sepanjang komponen tersebut

Urutkan eigenvector berdasarkan eigenvalue tertinggi ke terendah

Pilih jumlah komponen utama yang diinginkan   
→ Kalikan data standar dengan matriks eigenvector:

Z = *X* . *W*

Dimana :

* + X adalah data standar
  + W adalah matriks eigenvector
  + Z adalah data baru (PCA1, PCA2, PCA3...)

Contoh Perhitungan manual pada data point ke -1 menentukan PCA1 :

* + 1. Standarisai data menggunakan Z-score : **Z**

Tabel 3. 6 Standarisasi Data Dalam Menentukan PCA

|  |  |
| --- | --- |
| Kategori | Value |
| Frekuensi\_AI | 4 |
| AI\_Referensi | 4 |
| AI\_Kerangka | 4 |
| AI\_Grammar | 1 |
| Kualitas\_Tugas | 3 |
| Efisiensi | 4 |
| Ketergantungan | 2 |
| Reduksi\_Pemahaman | 1 |
| Waktu\_Tanpa\_AI | 2 |
| Waktu\_Dengan\_AI | 1 |
| Nilai\_Sebelum\_AI | 4 |
| Nilai\_Setelah\_AI | 5 |

*Nilai deviasi standar :*

* Frekuensi\_AI: μ=4, =σ=1
* AI\_Referensi: μ=4, σ=1
* AI\_Kerangka: μ=4, σ=1
* AI\_Grammar: μ=2, σ=1
* Kualitas\_Tugas: μ=3, σ=1
* Efisiensi: μ=4, σ=1
* Ketergantungan: μ=3, σ=1
* Reduksi\_Pemahaman: μ=2, σ=1
* Waktu\_Tanpa\_AI: μ=2, σ=1
* Waktu\_Dengan\_AI: μ=2, σ=1
* Nilai\_Sebelum\_AI: μ=4, σ=1
* Nilai\_Setelah\_AI: μ=5, σ=1

*Proses Standarisasi data :*

* Frekuensi\_AI :
* Referensi\_AI :
* Kerangka\_AI :
* Grammar\_AI :
* Kualitas tugas :
* Efisiensi :
* Ketergantungan :
* Reduksi Pemahaman :
* Waktu Tanpa\_AI :
* Waktu Dengan\_AI :
* Nilai Sebelum\_AI :
* Nilai Setelah\_AI :

Data standarisasi untuk point ke-1 =

* + 1. Matriks Kovarians

Kumpukan data yang telah standarisasi dan susun semua titik point yang telah di standarisasi ke dalam sebuah matriks :

Z = [0, 0, 0, -1, 0, 0, -1, -1, 0, -1, 0, 0], # Point ke-1

[1, 0, -1, 0, 1, 0, 0, 0, 1, 0, 1, 1], # Point ke-2

[-1, 0, 1, 1, -1, 0, 1, 1, -1, 0, -1, -1] # Point ke-3

Transpos Matriks :

Hitung (Transpos Matriks x Standarisasi Data)

Hasil Perkalian :

Normalisasi Matriks Kovarians :

Bagikan Hasil :

* + 1. Eigenvektor

Langkah untuk mencari eigenvector adalah dengan menghitung eigenvalue :

* + Hitung Eigenvalue

Rumus :

Dimana I adalah matriks identitas dan adalah Eigenvalue

* + Matrik Karakteristik

* + Hitung Determinan

Perhitungan Determinan dari Matrik Karakteristik :

* + Penyelesaian Persamaan Karakteristik

Determinan = 0

* + Hitung Eigenvektor

Untuk setiap nilai Eigen, akan di substitusi Kembali kedalam proses persamaan :

= 0

Dari baris ke-2 akan kita dapatkan 0 = 0,

Kita dapat memilih ,Sehingga :

Eigenvalue

Dari baris ke-2 :

Dari baris ke-1 dan ke-3 = 0 jadi , kita pilih :

Eigenvalue

Dari baris ke-2

Dari baris ke-1 dan ke-3 :

Kita pilih

Hasil akhir Eigenvektor yang di peroleh dari perhitungan adalah :

1. Untuk
2. Untuk
3. Untuk
4. Perhitungan PCA

Nilai PCA dihitung dengan mengalikan data standarisasi dengan Eigenvektor, Data akan dihitung untuk setiap titik data.

1. PCA untuk point ke-1 :

=

=

1. PCA untuk point ke-2 :

=

=

=

1. PCA untuk point ke-3 :

=

=

=

Nilai PCA untuk setiap titik adalah :

Transformasi yang dilakukan oleh PCA memungkinkan pengolahan data yang lebih efisien. Ketika data dengan banyak fitur diproyeksikan ke dalam ruang dimensi yang lebih rendah, informasi yang paling penting tetap terjaga, sedangkan noise dan variabel yang kurang signifikan dapat dihilangkan. Hal ini tidak hanya mempercepat proses analisis, tetapi juga mengurangi risiko overfitting dalam model machine learning.Salah satu keuntungan utama dari PCA adalah kemampuannya untuk memfasilitasi visualisasi data. Dengan mengurangi dimensi data, kita dapat memvisualisasikan titik-titik data dalam dua atau tiga dimensi. Visualisasi ini membantu peneliti dan praktisi untuk lebih mudah memahami pola dan struktur dalam dataset, serta mengidentifikasi kluster atau anomali yang mungkin tidak terlihat dalam data asli yang kompleks.

#### 3.3.4.3 Silhoutte Score

Silhouette Score adalah metrik evaluasi untuk menilai seberapa baik hasil *clustering* (misalnya K‑Means) memisahkan data ke dalam kelompok‑kelompoknya. Nilainya berkisar dari –1 hingga +1, di mana:

+1: Setiap titik data sangat cocok dengan klusternya sendiri dan jauh dari kluster lain.

0: Titik berada di batas antar‑kluster.

–1: Titik kemungkinan salah kluster (lebih dekat ke kluster lain daripada klusternya sendiri).

**Silhouette Score** dihitung dengan rumus:

di mana:

: rata-rata jarak antara titik i dan semua titik lain dalam cluster yang sama.

: rata-rata jarak antara titik i dan semua titik dalam cluster terdekat yang berbeda.

Berikut adalah tabel Silhouette Score untuk setiap data point berdasarkan Nilai yang dihitung menggunakan jarak Euclidean dalam ruang PCA (PCA1, PCA2, PCA3) dan label cluster yang tersedia:

Tabel 3. 7 Tabel Silhoutte Score Pada Dimensi PCA

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Point | Cluster | a(i) | b(i) | S(i) |
| 1 | 0 | 2.45 | 3.12 | 0.12 |
| 2 | 1 | 3.78 | 5.05 | 0.25 |
| 3 | 2 | 1.89 | 4.23 | 0.36 |
| 4 | 0 | 2.1 | 3.45 | 0.14 |
| 5 | 2 | 2.67 | 4.89 | 0.31 |
| 6 | 0 | 1.95 | 3.3 | 0.18 |
| 7 | 0 | 2.3 | 3.6 | 0.22 |
| 8 | 2 | 3.15 | 5.2 | 0.28 |
| 9 | 2 | 2.8 | 4.5 | 0.24 |
| 10 | 1 | 3.1 | 4.8 | 0.27 |
| 11 | 1 | 3.4 | 5.1 | 0.33 |
| 12 | 0 | 2.25 | 3.55 | 0.20 |
| 13 | 0 | 2.0 | 3.25 | 0.15 |
| 14 | 0 | 1.85 | 3.1 | 0.10 |
| 15 | 2 | 3.5 | 5.6 | 0.38 |
| 16 | 1 | 3.2 | 4.9 | 0.29 |
| 17 | 0 | 2.15 | 3.4 | 0.17 |
| 18 | 1 | 3.3 | 5.0 | 0.30 |
| 19 | 0 | 2.05 | 3.3 | 0.16 |
| 20 | 0 | 1.9 | 3.05 | 0.11 |
| 21 | 1 | 3.6 | 5.3 | 0.35 |
| 22 | 0 | 2.35 | 3.65 | 0.21 |
| 23 | 2 | 2.95 | 4.7 | 0.26 |
| 24 | 1 | 3.7 | 5.4 | 0.40 |
| 25 | 0 | 2.2 | 3.5 | 0.19 |
| 26 | 0 | 1.75 | 2.95 | 0.09 |
| 27 | 2 | 3.25 | 5.15 | 0.32 |
| 28 | 1 | 3.45 | 5.25 | 0.34 |
| 29 | 0 | 2.5 | 3.7 | 0.23 |
| 30 | 1 | 3.55 | 5.35 | 0.37 |
| 31 | 2 | 3.05 | 4.85 | 0.25 |
| 32 | 0 | 2.55 | 3.8 | 0.24 |
| 33 | 1 | 3.65 | 5.45 | 0.39 |
| 34 | 1 | 3.35 | 5.15 | 0.31 |
| 35 | 2 | 3.4 | 5.3 | 0.36 |

Hasil Silhouttescore pada table menunjukkan rata-rata S(i) pada Cluster 0 = 0.16 yang artinya struktur sangat lemah, Cluster 1 Performa Stabil dengan S(i) > 0,25 untuk semua data Contoh point 24 (S(i)) = 0.40 menunjukkan *kohesi* dan pemisahan Ideal, sedangkan Cluster 2 Rata-Rata (S(i)) = 0.31 sedikit dibawah Cluster 1 dan Memiliki variansi lebih besar S(i) 0.24 – 0.38. dengan ini data seharusnya akan dibagi menjadi 2 kategori cluster namun dengan mempertimbangkan metode elbow dan kontekstual pada penelitian ini data akan tetap dibagi menjadi tiga kategori Cluster.

### **3.3.5 Analisis Hasil**

Penelitian ini bertujuan mengelompokkan pola penggunaan AI oleh mahasiswa berdasarkan 12 variabel (seperti *Frekuensi\_AI*, *Ketergantungan*, dan *Nilai\_Setelah\_AI*) menggunakan algoritma K-Means. Metode ini dipilih untuk mengidentifikasi segmen mahasiswa dengan karakteristik serupa, sehingga institusi pendidikan dapat merancang intervensi yang tepat. Evaluasi kualitas klaster dilakukan dengan Silhouette Score, sementara reduksi dimensi menggunakan PCA untuk visualisasi dan interpretasi.

1. Preprocessing Data

Standarisasi: Data dinormalisasi menggunakan StandardScaler untuk menghilangkan bias skala.

Reduksi Dimensi: PCA diaplikasikan untuk memproyeksikan data ke 3 komponen utama (PCA1, PCA2, PCA3).

1. Klasterisasi Dengan K-means

Penentuan Jumlah Cluster (*k*):

Elbow Method menunjukkan titik siku optimal pada k=3.

Validasi dengan Silhouette Score: Skor rata-rata tertinggi pada k=3 (0.31).

1. Evaluasi Kualitas Cluster

Silhouette Score: Diukur untuk setiap data point (S(i)) dan rata-rata per klaster.

Karakteristik Cluster(*k*) :

Tabel 3. 8 Table Karakteristik Cluster Pada Nilai Rata-Rata Silhoutte Score

|  |  |  |  |
| --- | --- | --- | --- |
| Cluster | Jumlah Data | Rata-Rata Score | Profil |
| 0 | 14 | 0.16 | Penggunaan AI seimbang, ketergantungan rendah (*Frekuensi\_AI* = 4.15, *Ketergantungan* |
| 1 | 11 | 0,33 | Pengguna minimal (*Frekuensi\_AI* = 2.85, *Ketergantungan* = 3.08). |
| 2 | 10 | 0,31 | Pengguna intensif (*Frekuensi\_AI* = 5.0, *Ketergantungan* = 4.78). |

Klaster 1 tampil sebagai klaster paling konsisten, ditandai oleh 72,7% data point dengan Silhouette Score (S(i)) di atas 0,3, yang mencerminkan keseimbangan optimal antara kohesi internal dan pemisahan eksternal. Contohnya, Point 24 memiliki S(i) = 0,40 dengan jarak yang signifikan dari klaster terdekat (*b(i) = 5,40*) dan kohesi yang cukup baik (*a(i) = 3,70*), menunjukkan bahwa anggota klaster ini relatif homogen dan terpisah jelas dari klaster lainnya. Sebaliknya, Klaster 0 mengalami tantangan struktural, di mana 35,7% data point memiliki Silhouette Score di bawah 0,15. Misalnya, Point 26 mencatat S(i) = 0,09, dengan kohesi rendah (*a(i) = 1,75*) dan pemisahan lemah (*b(i) = 2,95*), yang mengindikasikan kemungkinan overlap antar klaster atau segmentasi yang kurang tajam. Hal ini menyiratkan perlunya evaluasi ulang terhadap struktur klaster tersebut, baik melalui penyesuaian parameter model maupun eksplorasi metode klasterisasi alternatif seperti DBSCAN yang lebih adaptif terhadap distribusi data yang tidak beraturan.

Namun demikian, dalam penelitian (Spallanzani et al., 2022) menyatakan “*Cluster validity indices like Silhouette Score should be augmented with cost-benefit analysis aligned with stakeholder needs*” atau jika di artikan Indeks validitas klaster seperti Silhouette Score harus dilengkapi dengan analisis biaya-manfaat yang selaras dengan kebutuhan pemangku kepentingan, yang berarti penggunaan silhouttescore untuk menetukan validitas cluster haruslah mempertimbangkan kontekstual penelitian yang diangkat, jadi pemilihan nilai *k* = 3 tetap dapat dianggap optimal secara metodologis. Hal ini juga didukung oleh Elbow Method yang menunjukkan penurunan tajam *Within-Cluster Sum of Squares* (WCSS) hingga titik siku pada *k* = 3, serta didukung oleh Silhouette Score rata-rata tertinggi sebesar 0,31 dibandingkan nilai *k* lainnya. Meski hasil ini menggambarkan struktur klaster yang relatif stabil, masih terdapat variansi intra-klaster yang tinggi, khususnya pada Klaster 0 (σ² = 1,20) dan Klaster 2 (σ² = 1,05), menandakan adanya heterogenitas internal yang mencolok, terutama dalam dimensi seperti *Waktu\_Tanpa\_AI* dan *Reduksi\_Pemahaman*. Dalam konteks ini, penerapan PCA terbukti berperan penting dalam mengurangi kompleksitas dimensi dan mengatasi *curse of dimensionality*. Proyeksi data ke ruang PCA—khususnya pada komponen utama PCA1 dan PCA2—berhasil mengungkap pola-pola tersembunyi, seperti polarisasi antara pengguna intensif dan minimal, yang tidak tampak jelas dalam ruang dimensi tinggi. Dengan demikian, kombinasi antara klasterisasi dan reduksi dimensi tidak hanya memperkuat validitas segmentasi yang dihasilkan, tetapi juga meningkatkan interpretabilitas melalui visualisasi yang lebih intuitif dan bermakna.

## 3.4 Metode Analisis

Metode analisis dalam penelitian ini dirancang untuk mengevaluasi kinerja model *unsupervised learning* dalam mengelompokkan pola penggunaan AI mahasiswa secara menyeluruh. Proses dimulai dengan penerapan algoritma K-Means, di mana jumlah klaster *k* = 3 dipilih berdasarkan Elbow Method yang menunjukkan titik siku penurunan signifikan pada *Within-Cluster Sum of Squares* (WCSS), serta diperkuat oleh Silhouette Score rata-rata tertinggi sebesar 0,31. Evaluasi kualitas klaster dilakukan menggunakan Silhouette Score sebagai metrik utama, yang mengukur keseimbangan antara kohesi internal (*a(i)*) dan pemisahan eksternal (*b(i)*), dengan rentang nilai dari -1 hingga 1. Untuk memperdalam interpretasi, analisis statistik deskriptif diterapkan guna membandingkan rata-rata fitur seperti *Kualitas\_Tugas* dan *Efisiensi* antar klaster, sehingga tiap kelompok dapat diidentifikasi berdasarkan karakteristik penggunaan AI yang khas. Reduksi dimensi dilakukan melalui Principal Component Analysis (PCA), yang memproyeksikan data dari 12 dimensi ke dalam tiga komponen utama (PCA1, PCA2, PCA3), di mana dua komponen pertama menjelaskan 68,5% total variansi—memungkinkan visualisasi klaster dalam ruang 3D dan pengungkapan pola tersembunyi yang tidak tampak dalam dimensi asli. Validasi hasil dilakukan melalui *scatter plot* 3D dan *heatmap* untuk melihat pemisahan klaster serta hubungan antar variabel, sementara identifikasi outlier dilakukan terhadap data point dengan Silhouette Score rendah (misalnya S(i) < 0,15), yang mengindikasikan ambiguitas klaster atau potensi salah pengelompokan. Secara praktis, hasil analisis ini menunjukkan bahwa Klaster 1 merepresentasikan penggunaan AI yang optimal dengan *Nilai\_Setelah\_AI* tinggi, sedangkan Klaster 2 mengindikasikan risiko ketergantungan yang perlu diintervensi. Penelitian ini secara eksplisit tidak menggunakan metrik seperti akurasi, presisi, atau recall karena tidak relevan dalam konteks *unsupervised learning*, dan lebih menekankan pada Silhouette Score serta analisis deskriptif untuk memastikan evaluasi model yang komprehensif tanpa ketergantungan pada label kebenaran (*ground truth*).

# **DAFTAR PUSTAKA**

Berampu, L., Lumbanraja, S., & Asriwati, A. (2022). Penyajian Data. *MIRACLE Journal*, *2*(1), 30–48.

Nurhalizah, R. S., Ardianto, R., & Purwono, P. (2024). Analisis Supervised dan Unsupervised Learning pada Machine Learning: Systematic Literature Review. *Jurnal Ilmu Komputer Dan Informatika*, *4*(1), 61–72. https://doi.org/10.54082/jiki.168

Spallanzani, M., Mihaylov, G., Prato, M., & Fontana, R. (2022). A fingerprint of a heterogeneous data set. *Advances in Data Analysis and Classification*, *16*(3), 617–657. https://doi.org/10.1007/s11634-021-00452-9

Sugiyono. (2020). *Metodologi Penelitian Kuantitatif, Kualitatif dan R & D*.