

ANALISIS POLA KEGIATAN BELAJAR MAHASISWA TERHADAP KEBERHASILAN AKADEMIK MENGGUNAKAN ALGORITMA APRIORI

Helmi Saputra¹⁾, Rivaldo Aldino Uneputty²⁾, Leonardo Agustinus Simyapen³⁾, Mohammad Farhan Binalawan Mustamir⁴⁾, Fellisia Ayu Ikawanti⁵⁾, Surya Putri Kusumawati⁶⁾

Program Studi Teknik Informatika, Universitas Papua
Email Korespondensi: helmi@polbanganmanokwari.ac.id

Article Info

Article history:

Received: May 2, 2025
Revised: May 20, 2025
Accepted: May 30, 2025
Published: June 30, 2025

Keywords:

Data mining;
Algoritma
Apriori;
Keberhasilan Akademik;
Pendidikan Tinggi;
Mahasiswa

ABSTRACT

Permasalahan dalam penelitian ini adalah belum optimalnya pemanfaatan data perilaku belajar mahasiswa untuk memahami faktor-faktor yang memengaruhi keberhasilan akademik. Penelitian ini menawarkan pendekatan baru dengan menerapkan algoritma Apriori untuk mengidentifikasi pola belajar yang berasosiasi dengan IPK tinggi. Metode yang digunakan mengikuti tahapan Knowledge Discovery in Databases (KDD) dan melibatkan 408 mahasiswa dari berbagai perguruan tinggi di Kabupaten Manokwari. Temuan utama menunjukkan bahwa kombinasi motivasi akademik yang tinggi, kehadiran $\geq 90\%$, manajemen waktu yang baik, dan pembelajaran daring secara konsisten terkait dengan keberhasilan akademik. Sebanyak delapan aturan asosiasi paling menonjol berhasil diidentifikasi berdasarkan evaluasi metrik *support*, *confidence*, *lift*, dan *Zhang's metric*. Kebaruan dari penelitian ini terletak pada tiga aspek utama: (1) penerapan algoritma Apriori dalam konteks pendidikan tinggi di wilayah Indonesia Timur yang belum banyak dieksplorasi dalam studi sebelumnya; (2) integrasi metrik evaluasi ganda untuk menyaring aturan asosiasi yang paling bermakna; dan (3) visualisasi jaringan asosiasi sebagai pendekatan interpretatif dalam mengungkap keterkaitan perilaku belajar mahasiswa. Dengan demikian, penelitian ini tidak hanya memberikan pemahaman baru terhadap pola keberhasilan akademik, tetapi juga merekomendasikan penerapan strategi pembelajaran berbasis data sebagai upaya peningkatan kualitas pendidikan tinggi di masa depan.



This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY SA 4.0)

1. PENDAHULUAN

Keberhasilan akademik merupakan indikator utama dalam menilai kualitas pendidikan tinggi, yang umumnya diukur melalui Indeks Prestasi Kumulatif (IPK). IPK tidak hanya mencerminkan pencapaian akademis, tetapi juga menjadi syarat penting untuk kelulusan, penerimaan beasiswa, dan peluang karier. Oleh karena itu, memahami faktor-faktor yang memengaruhi IPK sangatlah penting dalam upaya meningkatkan kualitas pendidikan [1].

Berbagai penelitian telah menunjukkan bahwa keberhasilan akademik tidak semata-mata ditentukan oleh kemampuan intelektual, tetapi juga dipengaruhi oleh faktor-faktor lain seperti motivasi, disiplin, manajemen waktu, dan strategi belajar. Sebagai contoh, [2] menunjukkan bahwa motivasi intrinsik dan kegigihan (*grit*) memiliki korelasi positif yang signifikan dengan pencapaian akademik siswa sekolah menengah atas.

Demikian pula, [3] menemukan bahwa manajemen diri dan efikasi diri berperan penting

dalam menentukan keberhasilan akademik mahasiswa. Studi ini menunjukkan bahwa kemampuan mahasiswa dalam mengelola diri dan percaya pada kemampuan sendiri dapat meningkatkan pencapaian akademik mereka. Selain faktor-faktor individu, dukungan dari lingkungan belajar juga memainkan peran penting. [4] menyoroti bahwa resiliensi akademik dan dukungan dari dosen memiliki hubungan positif dengan kinerja akademik mahasiswa. Penelitian ini menekankan pentingnya peran dosen dalam mendukung mahasiswa untuk mencapai keberhasilan akademik.

Perkembangan teknologi informasi telah memungkinkan pengumpulan dan analisis data dalam skala besar, termasuk data terkait perilaku belajar mahasiswa. Namun, data tersebut seringkali belum dimanfaatkan secara optimal untuk mengidentifikasi pola-pola yang dapat mendukung peningkatan kinerja akademik. Dalam konteks ini, teknik penambangan data (*data mining*) [5], khususnya Algoritma Apriori, dapat digunakan untuk menemukan aturan-aturan asosiasi antara berbagai variabel yang memengaruhi keberhasilan akademik.

Algoritma Apriori merupakan salah satu metode dalam *data mining* yang digunakan untuk menemukan pola-pola asosiatif dalam dataset yang besar. Dengan menerapkan algoritma ini pada data perilaku belajar mahasiswa, dapat diidentifikasi kombinasi faktor-faktor yang sering muncul bersamaan dengan pencapaian IPK tinggi. Sebagai contoh, [6] menggunakan algoritma Apriori untuk menganalisis hubungan antara keterlibatan siswa dalam pembelajaran daring dan kinerja akademik mereka, dan menemukan bahwa tingkat keterlibatan yang tinggi berasosiasi dengan pencapaian akademik yang lebih baik.

Penerapan algoritma Apriori dalam konteks pendidikan tinggi di Provinsi Papua Barat khususnya Kabupaten Manokwari masih relatif terbatas. Padahal, pendekatan ini dapat memberikan wawasan baru

dalam memahami pola-pola perilaku belajar yang efektif dan mendukung pengambilan keputusan dalam perencanaan pembelajaran. Dengan demikian, penelitian ini bertujuan untuk menganalisis pola kegiatan belajar mahasiswa yang berasosiasi dengan keberhasilan akademik menggunakan algoritma Apriori, dengan harapan dapat memberikan kontribusi dalam upaya peningkatan kualitas pendidikan tinggi di Kabupaten Manokwari.

2. TINJAUAN PUSTAKA

2.1 *Data mining* dalam Pendidikan

Data mining dalam pendidikan, atau *Educational Data Mining* (EDM), merupakan pendekatan analitis yang digunakan untuk mengeksplorasi dan menganalisis data pendidikan guna meningkatkan proses pembelajaran dan pengambilan keputusan di institusi pendidikan. EDM bertujuan untuk memahami cara mahasiswa belajar, mendeteksi risiko akademik, serta mendukung keputusan manajemen kampus [7]. Proses data mining mencakup tahap-tahap dalam *Knowledge Discovery in Databases* (KDD), seperti pemahaman domain, seleksi dan pembersihan data, transformasi data, pemilihan algoritma, penambangan data, hingga interpretasi dan pemanfaatan hasil [8].

Dalam konteks global, [9] menunjukkan bahwa EDM telah berkembang pesat dalam dekade terakhir, dengan fokus utama pada analisis perilaku belajar, prediksi hasil akademik, dan personalisasi pembelajaran. Penelitian terbaru menekankan bahwa pemanfaatan EDM secara strategis dapat mendorong pembelajaran berbasis data untuk meningkatkan mutu pendidikan, terutama dalam ekosistem digital.

2.2 Algoritma Apriori

Algoritma Apriori adalah salah satu metode *association rule mining* dalam data mining yang digunakan untuk menemukan pola keterkaitan atau aturan asosiasi antara item dalam dataset besar. Algoritma ini bekerja berdasarkan prinsip bahwa jika suatu itemset sering muncul bersama dalam sebuah dataset, maka subset dari itemset tersebut juga sering

muncul. Dalam konteks pendidikan, algoritma ini digunakan untuk menganalisis keterkaitan antara perilaku mahasiswa (seperti aktivitas e-learning, kehadiran, atau pengumpulan tugas) dengan keberhasilan akademik [10].

Dalam penelitian [11] mengungkapkan bahwa peningkatan nilai minimum support dan *confidence* dalam algoritma Apriori berpengaruh signifikan terhadap kekuatan aturan asosiasi yang dihasilkan. Meskipun penelitian mereka berfokus pada data transaksi ritel, pendekatan ini dapat diadaptasi untuk pendidikan dalam menganalisis hubungan antar aktivitas belajar dan nilai akademik. Sementara itu, [12] menunjukkan bahwa algoritma Apriori dapat digunakan untuk menilai keterlibatan siswa dalam pembelajaran daring dan menghubungkannya dengan pencapaian akademik.

2.3 Kegiatan Belajar Mahasiswa

Kegiatan belajar mahasiswa meliputi aktivitas seperti mengikuti perkuliahan, membaca bahan ajar, berdiskusi dengan teman sejawat, mengerjakan tugas, dan memanfaatkan teknologi pembelajaran [13]. Menurut [14], keterlibatan aktif mahasiswa dalam berbagai kegiatan belajar berdampak positif terhadap Indeks Prestasi Kumulatif (IPK).

Perkembangan teknologi memungkinkan analisis lebih dalam terhadap aktivitas belajar, seperti penggunaan Learning Management System (LMS). [15] membuktikan bahwa partisipasi aktif mahasiswa dalam LMS seperti Moodle atau Google Classroom berhubungan erat dengan keberhasilan akademik. Hal ini membuka peluang pemanfaatan data aktivitas digital mahasiswa untuk menganalisis pola belajar yang efektif.

2.4 Keberhasilan Akademik

Keberhasilan akademik mahasiswa sering kali diukur melalui IPK, namun pencapaian ini dipengaruhi oleh banyak faktor, baik internal maupun eksternal. Faktor internal meliputi motivasi, strategi belajar, manajemen waktu, dan efikasi diri. [16]

menyatakan bahwa mahasiswa dengan strategi belajar terstruktur dan kemampuan manajemen waktu yang baik cenderung memiliki IPK lebih tinggi.

Di sisi lain, faktor eksternal seperti dukungan dosen, suasana kelas, dan ketersediaan fasilitas juga berperan penting. [17] menunjukkan bahwa kombinasi kedua jenis faktor tersebut mampu menjelaskan variasi dalam performa akademik mahasiswa. Oleh karena itu, analisis pola kegiatan belajar yang komprehensif menjadi penting untuk mendukung pencapaian akademik yang optimal.

3. METODOLOGI

3.1 Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan fokus pada pengolahan data numerik melalui algoritma Apriori.

3.2 Metode Penelitian

Metodologi yang digunakan dalam penelitian ini mengikuti kerangka Knowledge Discovery in Databases (KDD), sebuah proses iteratif dan sistematis untuk mengekstraksi pengetahuan yang bermakna dari data [18]. Tahapan-tahapan yang akan dilalui meliputi:

1. Seleksi Data (Data Selection):

Tahap awal ini berfokus pada identifikasi dan pemilihan data yang relevan dari database mahasiswa yang ada. Data yang akan dipilih secara spesifik meliputi:

- a) Indeks Prestasi Kumulatif (IPK) mahasiswa: Sebagai indikator utama keberhasilan akademik.
- b) Catatan kehadiran: Menunjukkan tingkat partisipasi dan disiplin dalam proses pembelajaran.
- c) Kebiasaan belajar: Meliputi frekuensi belajar, metode belajar yang diterapkan (misalnya, belajar mandiri, belajar kelompok), dan durasi belajar.
- d) Preferensi penggunaan media belajar: Contohnya, penggunaan *e-learning*, buku teks, jurnal, atau platform online.

- e) Tingkat motivasi untuk meraih IPK tinggi: Data ini dapat diperoleh melalui kuesioner yang mengukur aspek psikologis mahasiswa.
 - f) Aspek manajemen waktu: Meliputi kemampuan mahasiswa dalam mengatur jadwal belajar, mengerjakan tugas, dan menyeimbangkan kegiatan akademik dan non-akademik. Pemilihan variabel-variabel ini didasarkan pada studi literatur yang menunjukkan potensi korelasinya dengan keberhasilan akademik.
2. Pra-pemrosesan (Data Preprocessing)
- Setelah data terseleksi, tahap ini krusial untuk memastikan kualitas dan kebersihan data sebelum analisis. Aktivitas utama yang dilakukan meliputi:
- a) Normalisasi data: Mengubah skala data numerik (misalnya, IPK, durasi belajar) ke dalam rentang standar tertentu (misalnya, 0-1) untuk mencegah dominasi atribut dengan nilai numerik yang lebih besar dalam proses analisis. Hal ini penting agar semua atribut memiliki bobot yang setara.
 - b) Pengisian data hilang (*missing values*): Mengidentifikasi dan menangani entri data yang kosong. Metode pengisian dapat bervariasi, seperti imputasi dengan nilai rata-rata/median, modus, atau menggunakan algoritma tertentu untuk memprediksi nilai yang hilang. Tujuannya adalah untuk menjaga kelengkapan dataset.
 - c) Penanganan *outlier* (jika diperlukan): Mengidentifikasi dan memutuskan strategi untuk mengelola data yang sangat berbeda dari sebagian besar data lainnya, karena *outlier* dapat memengaruhi hasil analisis.
3. Transformasi (Data Transformation):
- Pada tahap ini, data yang telah bersih dan normalisasi akan diubah ke dalam format yang sesuai untuk algoritma association rule mining. Secara khusus, data akan dikonversi menjadi representasi transaksi (itemset). Setiap baris data mahasiswa akan dianggap sebagai satu "transaksi," dan setiap karakteristik (misalnya, "IPK Tinggi", "Rajin Belajar", "Sering Menggunakan E-learning") akan

direpresentasikan sebagai "item". Konversi ini seringkali melibatkan diskretisasi variabel numerik (misalnya, IPK menjadi kategori "Tinggi", "Sedang", "Rendah") dan binerisasi variabel kategorikal.

4. Penambangan Pola (Pattern Mining):

Inti dari proses KDD, tahap ini melibatkan aplikasi algoritma data mining untuk menemukan pola atau hubungan tersembunyi dalam data. Dalam penelitian ini, algoritma Apriori akan digunakan.

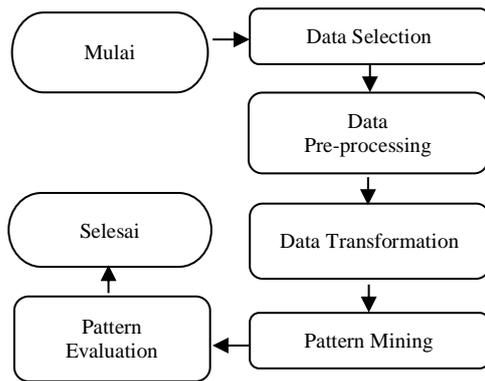
- a) Algoritma Apriori akan mengidentifikasi itemset frekuensi (kombinasi item yang sering muncul bersama) dan kemudian menghasilkan aturan asosiasi (association rules) dalam bentuk "IF A THEN B" [19].
- b) Penentuan parameter minimum *support* dan minimum *confidence* sangat penting. *Support* mengukur seberapa sering sebuah itemset muncul dalam dataset, sedangkan *confidence* mengukur probabilitas itemset konsekuen (B) muncul ketika itemset anteseden (A) sudah ada. Pengaturan nilai-nilai ini akan menentukan jumlah dan relevansi aturan asosiasi yang ditemukan [20].

5. Evaluasi (Pattern Evaluation):

Tahap akhir ini berfokus pada penilaian dan interpretasi pola asosiasi yang telah ditemukan oleh algoritma Apriori.

- a) Aturan-aturan yang dihasilkan akan disaring dan dievaluasi berdasarkan metrik tambahan (misalnya, *lift*, *conviction*) serta relevansi domain.
- b) Interpretasi akan dilakukan untuk menjelaskan makna dari setiap pola asosiasi yang signifikan, terutama yang menunjukkan hubungan antara pola kegiatan belajar mahasiswa dengan tingkat keberhasilan akademik mereka.
- c) Pola-pola ini kemudian akan dianalisis untuk menarik kesimpulan yang bermakna dan memberikan wawasan yang dapat ditindaklanjuti, seperti rekomendasi bagi

mahasiswa atau institusi pendidikan untuk meningkatkan keberhasilan akademik.



Gambar 1. Diagram Alur Penelitian

3.3 Sumber dan Teknik Pengumpulan Data

Target pengumpulan data responden pada penelitian ini adalah mahasiswa aktif dari berbagai perguruan tinggi di wilayah Kabupaten Manokwari, Kabupaten Manokwari menjadi target pengumpulan data responden pada penelitian ini karena merupakan salah satu wilayah di Indonesia bagian timur yang memiliki karakteristik sosial, budaya, dan pendidikan yang khas. Wilayah ini dipilih karena adanya kemudahan akses untuk menjangkau responden mahasiswa melalui jaringan akademik yang telah terbentuk sebelumnya. Hal ini mendukung kelancaran distribusi kuesioner dan pelaksanaan dokumentasi akademik. Data dikumpulkan menggunakan kuesioner online. Untuk memastikan validitas kuesioner, proses validasi isi dilakukan dengan melibatkan dosen di bidang pendidikan. Selain itu, reliabilitas data diuji melalui uji coba lapangan yang melibatkan responden pada kelas uji yang lebih kecil. Dokumentasi akademik (IPK) juga disertakan dalam kuesioner untuk menambah kedalaman dan keakuratan informasi yang diperoleh, jumlah total responden yang terkumpul adalah 408 responden. Rincian demografinya bisa dilihat pada table 1.

Variabel	Kategori	Frekuensi	Presentase
	Institut agama kristen ambon	1	0,2%

	Pendidikan Guru Sekolah Dasar Polbangtan Manokwari	1	0,2%
	Poltekkes kemenkes sorong prodi manokwari	89	21,8%
	Sekolah Tinggi Ilmu Hukum Manokwari Sekolah Tinggi Sains Dan Teknologi Manokwari	1	0,2%
	Sekolah Tinggi Teologi Erikson Tritt	11	2,7%
	STKIP Muhammadiyah Manokwari	1	0,2%
	STMIK Kreatindo Manokwari	4	1,0%
	Universitas Caritas Indonesia	35	8,6%
	Universitas Kristen Indonesia Maluku	16	3,9%
	Universitas Papua	39	9,6%
	UNIVERSITAS PATTIMURA	2	0,5%
	Universitas Pattimura Ambon	1	0,2%
	Laki-Laki	2	0,5%
	Perempuan	158	38,7%
	Semester 2	250	61,3%
	Semester 4	102	25,0%
	Semester 6	99	24,3%
	Semester 8 atau lebih	131	32,1%
		76	18,6%

Tabel 1. Demografi Responden

3.4 Teknik Analisis Data

Algoritma Apriori dijalankan dengan minimum support sebesar $\geq 30\%$ untuk memastikan bahwa hanya pola-pola yang cukup sering muncul dalam data (setidaknya pada 30% dari seluruh transaksi) yang dianggap signifikan. Selain itu, nilai *confidence* sebesar $\geq 70\%$ digunakan untuk menjamin bahwa aturan asosiasi yang dihasilkan memiliki tingkat keandalan yang tinggi, sehingga pola kegiatan belajar yang teridentifikasi cenderung diikuti oleh indikator keberhasilan akademik tertentu [21].

Algoritma Apriori bekerja berdasarkan prinsip bahwa jika sebuah kombinasi item (itemset) sering muncul dalam dataset, maka semua subset-nya juga cenderung sering muncul. Proses pencarian dilakukan secara iteratif dengan membentuk itemset dari tingkat

1 (1-itemset), kemudian meningkatkan levelnya menjadi 2-itemset, 3-itemset, dan seterusnya, hingga tidak ditemukan lagi kombinasi yang memenuhi nilai ambang batas (threshold) minimum.

Dua metrik utama yang digunakan dalam algoritma ini adalah support dan confidence, support mengukur proporsi transaksi dalam dataset yang mengandung kombinasi item tertentu. Rumus support didefinisikan sebagai.

$$\text{Support}(A \rightarrow B) = \frac{\text{Jumlah transaksi yang mengandung } A \cup B}{\text{Jumlah total transaksi}}$$

Dalam penelitian ini, batas minimum support ditetapkan sebesar 30% ($\geq 0,3$), artinya hanya kombinasi item yang muncul pada setidaknya 30% dari keseluruhan data yang akan dipertimbangkan.

Confidence mengukur kekuatan implikasi suatu aturan, yaitu seberapa besar kemungkinan item B muncul dalam suatu transaksi jika item A sudah ada. Rumus confidence ditulis

$$\text{Confidence}(A \rightarrow B) = \frac{\text{Support}(A \cup B)}{\text{Support}(A)}$$

Nilai confidence minimum yang digunakan dalam penelitian ini adalah 70% ($\geq 0,7$), yang berarti bahwa aturan hanya dianggap valid jika item B muncul pada setidaknya 70% dari transaksi yang mengandung item A.

Pemilihan nilai minimum *support* sebesar **30%** dan *confidence* sebesar **70%** dalam penelitian ini dilakukan dengan mempertimbangkan dua pendekatan utama yaitu studi literatur dan uji coba awal terhadap data yang tersedia. Secara literatur, batas *support* minimal 30% dan *confidence* minimal 70% telah digunakan secara luas dalam penelitian association rule mining pada konteks pendidikan dan non-pendidikan. Penelitian oleh [11] dan [21] menunjukkan bahwa nilai-nilai ini cenderung menghasilkan aturan asosiasi yang cukup representatif namun tidak terlalu banyak, sehingga tetap relevan secara interpretatif dan aplikatif dalam konteks pengambilan Keputusan, pada pengujian awal,

dilakukan uji coba dengan beberapa skenario threshold berbeda, yakni Support 20%, 25%, dan 30% Confidence 60%, 65%, dan 70%, hasil uji coba menunjukkan bahwa nilai *support* di bawah 30% menghasilkan terlalu banyak aturan yang kurang relevan secara praktis (noise), sementara nilai *confidence* di bawah 70% menghasilkan aturan dengan kekuatan prediktif rendah. Nilai 30% dan 70% memberikan kompromi terbaik antara jumlah aturan, kejelasan pola, dan kemudahan interpretasi.

Dengan demikian, pemilihan threshold ini bersifat terukur dan berbasis data, hal ini penting untuk memastikan bahwa aturan-aturan yang dihasilkan tidak hanya valid secara statistik tetapi juga memiliki daya guna dalam praktik pendidikan tinggi, khususnya dalam konteks wilayah Indonesia Timur.

Dalam penelitian ini, proses analisis data dilakukan dengan menggunakan bahasa pemrograman Python melalui lingkungan interaktif Jupyter Notebook, yang memungkinkan eksplorasi data secara fleksibel dan sistematis. Untuk implementasi algoritma Apriori, digunakan pustaka (library) dalam data mining, yaitu mlxtend (machine learning extensions), yang menyediakan fungsi siap pakai untuk penambangan aturan asosiasi.

4. HASIL DAN PEMBAHASAN

4.1 Seleksi Data

Acad	IPK_Terakhir	Kehadiran	kebiasaan_belajar	kepan_belajar	media_belajar	penyebutan_online	motivasi_ikg	manajemen_waktu	organisasi	
0	Perguruan Tinggi	≥ 3,50	≥ 90%	1-2 Kali	Malam	E-book atau jurnal online	3	4	4	Aktif Organisasi
1	Universitas Papua	≥ 3,50	≥ 90%	1-2 Kali	Tidak tentu	video pembelajaran (Zoom, lab, Ruangrupa, dll)	4	3	3	Tidak Aktif
2	Universitas Papua	3,00 - 3,49	75-89%	1-2 Kali	Malam	video pembelajaran (Zoom, lab, Ruangrupa, dll)	4	4	4	Kerja Sambilan
3	Universitas Papua	≥ 3,50	≥ 90%	1-2 Kali	Malam	Grup diskusi online (WhatsApp, Telegram, dll)	3	4	4	Tidak Aktif
4	Universitas Papua	3,00 - 3,49	≥ 90%	3-4 kali	Tidak tentu	LMS kampus (Google Classroom, Moodle, dll) E-	4	4	4	Tidak Aktif

Gambar 2. Hasil Seleksi Data

Pada tahap seleksi data, memilih atribut atau variabel yang relevan dari *database* mentah yang dimiliki. Gambar 4.1 memperlihatkan cuplikan data mahasiswa yang telah diseleksi dan akan menjadi

input untuk proses selanjutnya. Kolom-kolom (atribut) dalam tabel tersebut adalah data yang spesifik dipilih, antara lain:

- Asal Perguruan Tinggi:** Informasi demografi yang mungkin memengaruhi pola belajar.
- IPK Terakhir:** Indikator utama keberhasilan akademik, yang akan menjadi variabel target atau variabel yang akan dianalisis korelasinya.
- Kehadiran:** Menunjukkan tingkat partisipasi.
- Kebiasaan Belajar:** Menggambarkan frekuensi atau pola belajar mahasiswa.
- Kapan Belajar:** Aspek waktu belajar (misalnya, malam, tidak tentu).
- Media Belajar:** Jenis media yang digunakan mahasiswa dalam belajar.
- Pembelajaran Online:** Indikator keterlibatan dalam metode pembelajaran daring.
- Motivasi IPK:** Tingkat motivasi untuk mencapai IPK tertentu.
- Manajemen Waktu:** Kemampuan mengatur waktu.
- Organisasi:** Keterlibatan mahasiswa dalam kegiatan organisasi.

Data yang akan dipakai untuk tahap selanjutnya adalah data mahasiswa yang memiliki nilai IPK tinggi

4.2 Pra-pemrosesan Data

Data yang telah di peroleh melalui kuesioner mahasiswa dipersiapkan agar sesuai dengan format Algoritma Apriori. Proses ini dilakukan untuk membersihkan dari data yang duplikasi atau data yang tidak sesuai, mengubah nilai numerik menjadi kategori.

	IPK_Terakhir	IPK_Tinggi
0	≥ 3.50	1
1	≥ 3.50	1
2	3.00 - 3.49	0
3	≥ 3.50	1
4	3.00 - 3.49	0
...
397	3.00 - 3.49	0
398	3.00 - 3.49	0
399	≥ 3.50	1
400	2.50 - 2.99	0
401	3.00 - 3.49	0
402 rows x 2 columns		

Gambar 3. Pra-pemrosesan Data

Tujuan pra-pemrosesan ini untuk memastikan data yang akan digunakan sudah konsisten, bersih, dan terstruktur dengan baik sehingga hasil aturan asosiasi yang dihasilkan menjadi lebih akurat dan bermakna.

4.3 Transformasi Data

Transformasi data dilakukan dengan mengubah setiap. baris data menjadi format transaksi.

Gambar 4. Transformasi Data

Dari setiap respon atau atribut mahasiswa akan diubah menjadi item yang membentuk satu set transaksi. Kemudian Data direpresentasikan dalam format biner (0 dan 1) menggunakan teknik one-hot encoding.

4.4 Pattern Mining (Pencarian Pola Asosiasi)

	support	itemsets
0	0.850394	(pembelajaran_online)
1	0.976378	(motivasi_ipk)
2	0.944882	(manajemen_waktu)
3	0.937008	(Kehadiran_≥ 90%)
4	0.464567	(kebiasaan_belajar_1-2 Kali)
...
58	0.370079	(motivasi_ipk, Kehadiran_≥ 90%, manajemen_wakt...
59	0.314961	(organisasi_Aktif Organisasi, motivasi_ipk, Ke...
60	0.330709	(motivasi_ipk, organisasi_Tidak Aktif, manajem...
61	0.362205	(Kehadiran_≥ 90%, manajemen_waktu, motivasi_ip...
62	0.314961	(Kehadiran_≥ 90%, manajemen_waktu, motivasi_ip...

Gambar 5. Item Set (pola yang terbentuk)

Algoritma **Apriori** digunakan untuk menemukan kombinasi jawaban yang sering muncul dari data kuesioner mahasiswa, seperti tentang kebiasaan belajar, tingkat kehadiran, dan motivasi.

Setelah ditemukan kombinasi data yang sering muncul dan memenuhi batas minimum kemunculan, kemudian membuat aturan hubungan antara data (**asosiasi**) berdasarkan seberapa besar kemungkinan suatu hal terjadi jika hal lainnya terjadi (**confidence**) [22].

Algoritma Apriori diterapkan untuk menemukan pola itemset yang sering muncul dengan nilai minimum support sebesar 0.3. Aturan asosiasi kemudian dibentuk dari itemset tersebut berdasarkan nilai *confidence* minimum 0.7, untuk memastikan aturan yang diperoleh memiliki kekuatan prediktif yang cukup tinggi

4.5 Pattern Evaluation (Evaluasi Pola)

	antecedents	consequents	support	confidence	li
0	(pembelajaran_online)	(motivasi_ipk)	0.834646	0.981481	1.00522
1	(motivasi_ipk)	(pembelajaran_online)	0.834646	0.854839	1.00522
2	(pembelajaran_online)	(manajemen_waktu)	0.818898	0.962963	1.01913
3	(manajemen_waktu)	(pembelajaran_online)	0.818898	0.866667	1.01913
4	(pembelajaran_online)	(Kehadiran_≥ 90%)	0.803150	0.944444	1.00793
...
213	(Kehadiran_≥ 90%, kebiasaan_belajar_3-4 kali)	(pembelajaran_online, motivasi_ipk, manajemen_waktu)	0.314961	0.833333	1.03758
214	(motivasi_ipk, kebiasaan_belajar_3-4 kali)	(pembelajaran_online, Kehadiran_≥ 90%, manajemen_waktu)	0.314961	0.816327	1.05786
215	(pembelajaran_online, kebiasaan_belajar_3-4 kali)	(manajemen_waktu, motivasi_ipk, Kehadiran_≥ 90%)	0.314961	0.930233	1.05481
216	(manajemen_waktu, kebiasaan_belajar_3-4 kali)	(pembelajaran_online, motivasi_ipk, Kehadiran_≥ 90%)	0.314961	0.800000	1.00594
217	(kebiasaan_belajar_3-4 kali)	(pembelajaran_online, motivasi_ipk, Kehadiran_≥ 90%)	0.314961	0.800000	1.04742

Gambar 6. Pola Asosiasi

Gambar 6 merupakan hasil proses dari evaluasi pola yang dijalankan menggunakan algoritma apriori, evaluasi pola dilakukan untuk menilai sejauh mana aturan asosiasi yang dihasilkan benar-benar relevan, kuat, dan bermakna dalam konteks keberhasilan akademik mahasiswa. Dengan menggunakan berbagai metrik evaluasi seperti support, confidence, lift, dan Zhang's metric. Dari seluruh aturan yang terbentuk, dipilih delapan aturan asosiasi terbaik berdasarkan kombinasi tertinggi dari keempat metrik evaluasi tersebut. Aturan-aturan ini dianggap paling signifikan dalam menjelaskan hubungan antar variabel kegiatan belajar mahasiswa yang mendukung keberhasilan akademik (IPK tinggi).

No	Antecedents	Consequents	Evaluasi (Support, Confidence, Lift, Zhang's Metric)
1	motivasi_ipk	Kehadiran_≥ 90%	Support: 0.929, Confidence: 0.952, Lift: 1.016, Zhang: 0.650
2	motivasi_ipk	Kehadiran_≥ 90%, manajemen_waktu	Support: 0.882, Confidence: 0.903, Lift: 1.015, Zhang: 0.631
3	motivasi_ipk	pembelajaran_online, Kehadiran_≥ 90%	Support: 0.795, Confidence: 0.815, Lift: 1.014, Zhang: 0.591

4	motivasi_ipk	manajemen_waktu	Support: 0.929, Confidence: 0.952, Lift: 1.007, Zhang: 0.299
5	Kehadiran_≥ 90%	motivasi_ipk	Support: 0.929, Confidence: 0.992, Lift: 1.016, Zhang: 0.244
6	motivasi_ipk	pembelajaran_online, Kehadiran_≥ 90%, manajemen_waktu	Support: 0.764, Confidence: 0.782, Lift: 1.014, Zhang: 0.574
7	manajemen_waktu	motivasi_ipk	Support: 0.929, Confidence: 0.983, Lift: 1.007, Zhang: 0.128
8	manajemen_waktu	pembelajaran_online	Support: 0.819, Confidence: 0.867, Lift: 1.019, Zhang: 0.341

Tabel 2. Aturan Asosiasi

Setelah proses pencarian, pola yang terbentuk dapat diuraikan sebagai berikut :

- Motivasi_ipk tinggi → kehadiran $\geq 90\%$ menunjukkan, Confidence: (95,2%) Lift (1,016), Mahasiswa yang memiliki motivasi akademik tinggi hampir selalu hadir di kelas, sehingga confidence sangat tinggi. Namun, karena kehadiran $\geq 90\%$ adalah perilaku umum dalam dataset (hampir semua mahasiswa rajin hadir), maka lift menjadi rendah. Ini menunjukkan bahwa meskipun hubungan itu konsisten, tidak berarti motivasi menyebabkan kehadiran—karena keduanya sama-sama tinggi dalam populasi.
- Motivasi_ipk → Manajemen_waktu menunjukkan Confidence tinggi (95,2%), Lift sangat rendah (1,007), mahasiswa yang termotivasi cenderung memiliki kedisiplinan (kehadiran tinggi) dan keterampilan organisasi diri (manajemen waktu). Namun karena kedua atribut tersebut juga sering muncul secara umum, lift-nya tetap rendah. Artinya, meskipun kombinasi ini sering terjadi, pola ini masih belum menunjukkan peningkatan kekuatan prediktif yang tinggi secara statistik.

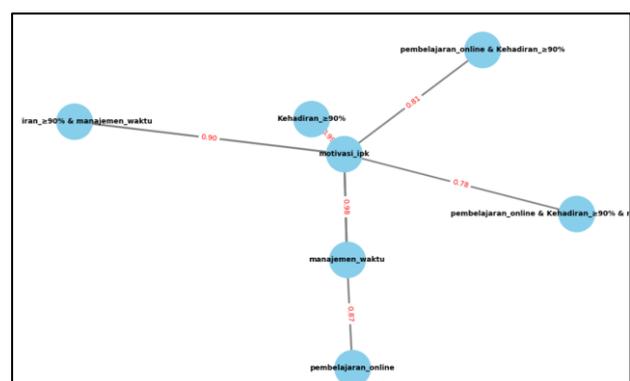
- c) Motivasi_ipk → pembelajaran_online dan Kehadiran $\geq 90\%$ menunjukkan Support (0,795), Confidence (81,5%) Lift (1,014), mahasiswa dengan motivasi tinggi cenderung belajar secara daring dan tetap hadir secara aktif. Pola ini relevan dalam konteks pembelajaran modern. Meskipun demikian, karena perilaku daring dan kehadiran tinggi sudah umum, lift tetap rendah. Ini menyiratkan bahwa kombinasi ini menggambarkan tren umum, bukan hubungan yang memperkuat secara signifikan.
- d) Motivasi_ipk → manajemen_waktu menunjukkan Support (0,929), Confidence (95,2%), Lift (1,007), sangat banyak mahasiswa yang termotivasi juga memiliki kemampuan manajemen waktu. Namun nilai lift yang mendekati 1 menunjukkan bahwa manajemen waktu baik juga terjadi pada mahasiswa lain yang mungkin kurang termotivasi. Dengan demikian, pola ini sangat umum tetapi kurang eksplosif sebagai penentu tunggal kesuksesan.
- e) Kehadiran $\geq 90\%$ → motivasi_ipk menunjukkan Support (0,929), Confidence (99,2%) dan Lift (1,016), Confidence tertinggi dari seluruh aturan. Ini berarti hampir semua mahasiswa yang rajin hadir memiliki motivasi akademik tinggi. Tapi sekali lagi, lift yang rendah menunjukkan bahwa motivasi tinggi tidak eksklusif bagi kelompok ini, sehingga hubungan ini kuat secara frekuensi, tetapi tidak secara prediktif murni.
- f) Motivasi_ipk → pembelajaran_online, Kehadiran $\geq 90\%$, manajemen_waktu menunjukkan Support (0,764), Confidence (78,2%) dan Lift (1,014), ini adalah pola kompleks dengan tiga item dalam consequents. Confidence cukup tinggi, menunjukkan banyak mahasiswa termotivasi memiliki ketiga perilaku tersebut. Namun karena masing-masing atributnya umum, nilai lift tetap tidak tinggi. Ini memberi sinyal bahwa kombinasi ini

menggambarkan “profil ideal” mahasiswa berprestasi, tetapi bukan pola yang langka.

- g) Manajemen_waktu → motivasi_ipk menunjukkan Support (0,929), Confidence (98,3%) dan Lift (1,007), mahasiswa yang bisa mengatur waktu hampir selalu memiliki motivasi IPK tinggi. Namun hubungan ini tidak terlalu memperkuat probabilitas terjadinya motivasi IPK jika dibandingkan dengan kejadian acak, karena lift mendekati 1. Ini menegaskan bahwa manajemen waktu bukan satu-satunya pendorong motivasi tinggi.
- h) manajemen_waktu → pembelajaran_online menunjukkan Support (0,819), Confidence (86,7%) dan Lift: 1,019, Ini adalah aturan dengan lift tertinggi dalam tabel. Artinya, mahasiswa yang pandai mengatur waktu lebih cenderung menggunakan pembelajaran daring dibandingkan kelompok lainnya. Confidence cukup tinggi, dan lift > 1 menunjukkan bahwa hubungan ini tidak hanya umum, tapi juga memperkuat kemungkinan keterkaitan nyata antar atribut. Ini bisa dimanfaatkan sebagai indikator awal keterlibatan digital mahasiswa.

4.6 Visualisasi Pola Asosiasi

Untuk meningkatkan pemahaman terhadap hubungan antar atribut kegiatan belajar mahasiswa, visualisasi dalam bentuk diagram jaringan asosiasi disusun berdasarkan delapan aturan asosiasi terkuat yang telah diidentifikasi sebelumnya.



Gambar 7. Visualisasi jaringan asosiasi

Gambar diatas memperlihatkan jaringan antar atribut, di mana setiap simpul (node) merepresentasikan variabel perilaku belajar mahasiswa seperti motivasi akademik, kehadiran, manajemen waktu, dan pembelajaran online. Arah panah menunjukkan aturan asosiasi dari antecedent → consequent, sedangkan ketebalan garis merepresentasikan kekuatan asosiasi berdasarkan nilai confidence. Ukuran node menunjukkan frekuensi kemunculan variabel dalam dataset (proxy dari support).

5. KESIMPULAN

Mahasiswa dengan IPK tinggi dalam penelitian ini menunjukkan pola perilaku yang konsisten, terutama pada aspek motivasi akademik yang tinggi, manajemen waktu yang baik, kehadiran $\geq 90\%$, dan penggunaan pembelajaran daring. Aturan-aturan asosiasi yang dihasilkan memiliki nilai confidence tinggi (sebagian besar di atas 90%), yang mengindikasikan bahwa perilaku-perilaku ini sering muncul bersamaan. Namun, sebagian besar nilai lift mendekati 1, yang menunjukkan bahwa meskipun hubungan antar variabel tersebut kuat secara frekuensi, kekuatannya tidak selalu signifikan dalam meningkatkan probabilitas terjadinya variabel konsekuen secara kausal. Hal ini penting untuk diperhatikan agar interpretasi hasil tetap proporsional dan tidak berlebihan.

Temuan ini menunjukkan bahwa keberhasilan akademik mahasiswa tidak hanya dipengaruhi oleh satu faktor tunggal, tetapi merupakan hasil dari sinergi beberapa perilaku positif yang saling mendukung. Kombinasi seperti motivasi tinggi, kehadiran yang konsisten, manajemen waktu yang terstruktur, dan pemanfaatan pembelajaran daring membentuk profil

mahasiswa berprestasi yang mampu menavigasi tantangan akademik dengan efektif.

Dalam konteks geografis dan sosial Papua Barat, pola-pola ini mencerminkan strategi adaptif mahasiswa dalam menghadapi hambatan akses, keterbatasan fasilitas, serta kondisi infrastruktur pendidikan yang belum merata. Sebagai contoh, hubungan yang kuat antara motivasi dan kehadiran tinggi dapat mencerminkan usaha mahasiswa untuk tetap hadir secara fisik di tengah keterbatasan transportasi atau jarak tempuh yang jauh. Sementara itu, penggunaan pembelajaran daring berasosiasi dengan manajemen waktu menunjukkan bahwa teknologi dimanfaatkan sebagai solusi atas keterbatasan fisik ruang dan waktu belajar. Jika studi serupa dilakukan di wilayah perkotaan atau dengan kondisi infrastruktur yang berbeda, kemungkinan besar pola-pola ini akan bergeser mengikuti karakteristik lokal masing-masing.

Temuan ini memiliki implikasi kebijakan yang penting bagi institusi pendidikan tinggi, khususnya dalam pengembangan layanan dan program pendampingan mahasiswa berbasis data. Institusi dapat memanfaatkan pola-pola perilaku ini untuk menyusun sistem peringatan dini (early warning system) atau platform bimbingan belajar adaptif, yang secara otomatis mengidentifikasi mahasiswa dengan kombinasi perilaku yang mendukung atau berisiko terhadap keberhasilan akademik. Selain itu, program pelatihan manajemen waktu, peningkatan literasi pembelajaran daring, serta monitoring motivasi belajar dapat dirancang secara lebih terarah, berbasis pada data konkret perilaku mahasiswa di lapangan. Pendekatan ini mendukung transformasi institusi pendidikan menuju ekosistem akademik yang berbasis data (data-driven decision making), inklusif, dan responsif terhadap konteks lokal.

Referensi:

- [1] Ira Veronika Girsang *et al.*, “Analisis Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa/i Jurusan Ekonomi Pembangunan FEB Universitas Palangka Raya,” *Nian Tana Sikka: Jurnal ilmiah Mahasiswa*, vol. 2, no. 1, pp. 145–156, Jan. 2024, doi: 10.59603/niantanasikka.v1i1.298.
- [2] A. Alhadabi and A. C. Karpinski, “Grit, self-efficacy, achievement orientation goals, and academic performance in University students,” *Int J Adolesc Youth*, vol. 25, no. 1, pp. 519–535, Dec. 2020, doi: 10.1080/02673843.2019.1679202.
- [3] A. Guerrero-Lucendo, F. García-Orenes, J. Navarro-Pedreño, and D. Alba-Hidalgo, “General Mapping of the Environmental Performance in Climate Change Mitigation of Spanish Universities through a Standardized Carbon Footprint Calculation Tool,” *Int J Environ Res Public Health*, vol. 19, no. 17, p. 10964, Sep. 2022, doi: 10.3390/ijerph191710964.
- [4] Z. Cai and Q. Meng, “Academic resilience and academic performance of university students: the mediating role of teacher support,” *Front Psychol*, vol. 16, Apr. 2025, doi: 10.3389/fpsyg.2025.1463643.
- [5] S. A. A. Kharis and A. H. A. Zili, “Learning Analytics dan Educational Data Mining pada Data Pendidikan,” *JURNAL Riset PEMBELAJARAN MATEMATIKA SEKOLAH*, vol. 6, no. 1, pp. 12–20, Mar. 2022, doi: 10.21009/jrpms.061.02.
- [6] Z. Tang, Z. Jiang, Y. Li, H. Yuan, J. Han, and C. Chen, “Research on the Association Analysis of Online Learning Behaviors Based on the Apriori Algorithm,” *Frontiers in Computing and Intelligent Systems*, vol. 9, no. 2, pp. 18–22, Aug. 2024, doi: 10.54097/4a3h3p03.
- [7] M. Yağcı, “Educational data mining: prediction of students’ academic performance using machine learning algorithms,” *Smart Learning Environments*, vol. 9, no. 1, p. 11, Dec. 2022, doi: 10.1186/s40561-022-00192-z.
- [8] P. Ihantola *et al.*, “Educational Data Mining and Learning Analytics in Programming,” in *Proceedings of the 2015 ITiCSE on Working Group Reports*, New York, NY, USA: ACM, Jul. 2015, pp. 41–63. doi: 10.1145/2858796.2858798.
- [9] F. Kurniawati, L. Ambarwati, and Lukman El Hakim, “Meningkatkan Kemampuan Pemecahan Masalah dan Disposisi Matematis Siswa dengan Model Pembelajaran Problem Based Learning Berbasis E-Learning,” *JURNAL Riset PEMBELAJARAN MATEMATIKA SEKOLAH*, vol. 6, no. 1, pp. 1–11, Mar. 2022, doi: 10.21009/jrpms.061.01.
- [10] R. Agrawal and R. Srikant, *Fast Algorithms for Mining Association Rules in Large Databases*. In *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB)*. Santiago, 1994.
- [11] N. A. Hibnastiar, A. F. Setiawan, and E. H. Susanto, “Penerapan Algoritma Apriori dalam Menentukan Rekomendasi Paket Produk,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 1, pp. 321–331, Jan. 2025, doi: 10.57152/malcom.v5i1.1782.
- [12] M. Ibrahim, R. Riana, and S. Soraya, “Evaluasi Keterlibatan Siswa dalam Lingkungan Pembelajaran Daring: Tinjauan Sistematis Literatur,” *Jurnal Ulul Albab*, vol. 28, no. 2, p. 112, Aug. 2024, doi: 10.31764/jua.v28i2.26058.
- [13] F. Aldresti, E. Erviyenni, and S. Haryati, “Pengembangan Lembar Kegiatan Mahasiswa Elektronik (e-LKM) berbasis Collaborative Learning Untuk Mata Kuliah Dasar-Dasar Pendidikan MIPA,” *PENDIPA Journal of Science Education*, vol. 5, no. 3, pp. 292–299, Mar. 2021, doi: 10.33369/pendipa.5.3.292-299.
- [14] Ira Veronika Girsang *et al.*, “Analisis Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa/i Jurusan Ekonomi Pembangunan FEB Universitas Palangka Raya,” *Nian Tana Sikka: Jurnal ilmiah Mahasiswa*, vol. 2, no. 1, pp. 145–156, Jan. 2024, doi: 10.59603/niantanasikka.v1i1.298.
- [15] I. S. MAHMUDI, M. MAS’ULA, and P. PURNAMAWATI, “EFEKTIVITAS MANAJEMEN PEMBELAJARAN DENGAN METODE BLENDED LEARNING MELALUI JEJARING MOODLE DAN GOOGLE CLASS ROOM PASCA COVID 19,” *EDUTECH: Jurnal Inovasi Pendidikan Berbantuan Teknologi*, vol. 2, no. 2, pp.

- 165–174, Jun. 2022, doi: 10.51878/edutech.v2i2.1208.
- [16] Y. Fu *et al.*, “Unlocking academic success: the impact of time management on college students’ study engagement,” *BMC Psychol.*, vol. 13, no. 1, p. 323, Apr. 2025, doi: 10.1186/s40359-025-02619-x.
- [17] Y. Yulyanah and H. Khotimah, “ANALYSIS OF INTERNAL AND EXTERNAL FACTORS AFFECTING STUDENTS LEARNING OUTPUT CASE STUDY D3 ACCOUNTING, PAMULANG UNIVERSITY, 2019/2020 ACADEMIC YEAR,” *JIM UPB (Jurnal Ilmiah Manajemen Universitas Putera Batam)*, vol. 9, no. 2, pp. 158–165, Jul. 2021, doi: 10.33884/jimupb.v9i2.3654.
- [18] N. Ameliana, N. Suarna, and W. Prihartono, “ANALISIS DATA MINING PENGELOMPOKKAN UMKM MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING DI PROVINSI JAWA BARAT,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 3261–3268, May 2024, doi: 10.36040/jati.v8i3.9655.
- [19] H. A. Mahdiraji, E. Razghandi, and A. Hatami-Marbini, “Overlapping coalition formation in game theory: A state-of-the-art review,” *Expert Syst Appl*, vol. 174, p. 114752, Jul. 2021, doi: 10.1016/j.eswa.2021.114752.
- [20] M. Das, “Analyzing impact of parental occupation on child’s learning performance: a semantics-driven probabilistic approach,” *Int J Data Sci Anal*, vol. 12, no. 1, pp. 31–44, Jun. 2021, doi: 10.1007/s41060-020-00220-6.
- [21] M. I. P. Wijaya, N. L. Aqromi, and S. N. Afyah, “Implementasi Data Mining Pola Pembelian Pada Toko Santoso Tiga Sumenep Dengan Menerapkan Algoritma Apriori,” *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 17, no. 2, pp. 97–108, Feb. 2023, doi: 10.32815/jitika.v17i2.909.
- [22] I. Irawan, Sunardi, and S. Harlina, “Implementasi Algoritma Apriori Pada Aplikasi Penjualan Buah Berbasis Web,” *Jurnal JTik (Jurnal Teknologi Informasi dan Komunikasi)*, vol. 9, no. 1, pp. 234–243, Nov. 2024, doi: 10.35870/jtik.v9i1.2971.