

Sistem Informasi Identifikasi Faktor yang Mempengaruhi Prestasi Siswa MAN Karo Menggunakan Pendekatan Data Mining

Ai Siti Mariam¹, Suendri²

^{1,2}Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara, Indonesia

¹stmrym122@gmail.com, ²suendri@uinsu.ac.id

Info Artikel

Riwayat Artikel:

Received 2025-12-27

Revised 2026-02-13

Accepted 2026-02-20

Corresponding Author:

Ai Siti Mariam

Email:

stmrym122@gmail.com



This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.

Abstract – Student academic achievement is a key indicator of learning success and is influenced by various internal factors (such as motivation, learning style, and attendance) as well as external factors (such as parental support and learning facilities). This study aims to identify the main factors that determine students' academic achievement by implementing a web-based information system using the Apriori algorithm. The study applied a quantitative descriptive approach using data from 300 students of MAN Karo, collected through questionnaires and academic records. The Apriori algorithm was implemented with a minimum support threshold of 0.05 and a minimum confidence threshold of 0.6, producing 1,621 association rules. The results indicate that a combination of internal and external factors is strongly associated with high academic achievement. The strongest association rule shows that students who attend tutoring classes tend to achieve high report card scores (≥ 85) and have complete learning facilities, with a confidence value of 67.6% and a lift ratio of 3.12. Overall, internal factors such as motivation (77.33%) and high attendance (63.67%), along with external factors including parental support, learning facilities, and tutoring participation, significantly contribute to improving students' academic performance. These findings demonstrate that the Apriori algorithm can effectively support schools in identifying dominant academic factors and can be used as a basis for data-driven decision making.

Keywords: Academic Achievement; Apriori Algorithm; Association Rules; Data Mining; Information Systems.

Abstrak – Prestasi akademik siswa merupakan indikator penting keberhasilan proses pembelajaran yang dipengaruhi oleh berbagai faktor internal, seperti motivasi, gaya belajar, dan kehadiran, serta faktor eksternal, seperti dukungan orang tua dan fasilitas belajar. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor dominan yang memengaruhi prestasi akademik siswa melalui penerapan sistem informasi berbasis web menggunakan algoritma Apriori. Penelitian ini menggunakan pendekatan deskriptif kuantitatif dengan data dari 300 siswa MAN Karo yang dikumpulkan melalui kuesioner dan catatan akademik. Algoritma Apriori diterapkan dengan nilai minimum support sebesar 0,05 dan minimum confidence sebesar 0,6 sehingga menghasilkan 1.621 aturan asosiasi. Hasil penelitian menunjukkan bahwa kombinasi faktor internal dan eksternal memiliki hubungan yang kuat terhadap prestasi akademik siswa. Aturan asosiasi terkuat menunjukkan bahwa siswa yang mengikuti bimbingan belajar cenderung memperoleh nilai rapor tinggi (≥ 85) dan memiliki fasilitas belajar lengkap, dengan nilai confidence sebesar 67,6% dan lift ratio sebesar 3,12. Secara umum, motivasi belajar (77,33%), tingkat kehadiran tinggi (63,67%), serta dukungan orang tua dan fasilitas belajar berkontribusi signifikan terhadap peningkatan prestasi siswa. Hasil ini menunjukkan bahwa algoritma Apriori dapat digunakan untuk membantu sekolah dalam menentukan kebijakan peningkatan prestasi siswa secara lebih tepat berbasis data.

Kata Kunci: Algoritma Apriori; Aturan Asosiasi; Data Mining; Prestasi Akademik; Sistem Informasi.

I. PENDAHULUAN

Perkembangan teknologi informasi (TI) saat ini telah memberikan dampak yang luas dalam berbagai bidang, termasuk pendidikan. Transformasi digital dalam pendidikan memungkinkan institusi untuk mengoptimalkan proses pembelajaran melalui analisis data yang lebih mendalam [1]. Dalam dunia pendidikan, teknologi informasi memungkinkan pengelolaan data akademik yang lebih efektif dan efisien, serta memberikan wawasan baru dalam pengambilan keputusan berbasis data. Salah satu pendekatan yang banyak digunakan dalam menganalisis data akademik adalah *data mining*, yang dapat mengidentifikasi pola-pola tersembunyi dalam data besar guna membantu pengambilan keputusan strategis [2]. Algoritma Apriori merupakan salah satu teknik *data mining* yang telah banyak diterapkan dalam berbagai domain, termasuk analisis pola pembelian [3], penentuan persediaan barang [4], dan korelasi perilaku siswa dengan tingkat prestasi [5]. *Data mining* dapat diterapkan untuk mengidentifikasi faktor-faktor yang mempengaruhi prestasi siswa, sehingga pihak sekolah dapat merancang kebijakan yang lebih tepat guna meningkatkan kualitas pembelajaran. Permasalahan ini erat kaitannya dengan bidang analisis data pendidikan, khususnya dalam mengidentifikasi faktor-faktor yang berkontribusi terhadap prestasi akademik siswa.

Madrasah Aliyah Negeri (MAN) Karo merupakan salah satu lembaga pendidikan agama yang berlokasi di Kelurahan Gung Negeri, Kabanjahe, Kabupaten Karo, dan berkomitmen untuk meningkatkan kualitas akademik siswanya. Sekolah ini menghadapi kesulitan dalam mengidentifikasi faktor-faktor yang berpengaruh terhadap prestasi akademik siswa, baik dari aspek internal seperti metode belajar dan kehadiran, maupun eksternal seperti lingkungan keluarga dan fasilitas sekolah [6]. Berbagai penelitian telah mengidentifikasi bahwa faktor internal seperti kesiapan belajar dan dukungan sosial terbukti berpengaruh signifikan terhadap prestasi mahasiswa [7]. Sementara itu, faktor eksternal seperti kondisi sosial ekonomi keluarga dan kegiatan di luar sekolah juga memiliki korelasi dengan prestasi siswa [8]. Hal ini disebabkan oleh keterbatasan sekolah dalam memanfaatkan data akademik yang telah tersedia secara sistematis dan analitis. Selama ini, data siswa hanya digunakan untuk keperluan administratif, bukan sebagai dasar dalam pengambilan keputusan strategis. Meskipun data akademik siswa telah tersedia, data tersebut belum dimanfaatkan secara optimal untuk analisis mendalam, sehingga pihak sekolah belum bisa menyusun strategi yang berbasis data untuk meningkatkan prestasi siswa. Akibatnya, pihak sekolah belum memiliki gambaran menyeluruh mengenai faktor-faktor yang paling signifikan terhadap pencapaian akademik siswa, sehingga intervensi yang dilakukan bersifat umum dan tidak didasarkan pada pola data yang objektif dan terukur [9]. Oleh karena itu, diperlukan pendekatan berbasis *data mining* untuk mengidentifikasi hubungan antara berbagai faktor dan prestasi akademik siswa, sehingga sekolah dapat merancang strategi peningkatan kualitas pendidikan yang lebih efektif dan berbasis data.

Berdasarkan permasalahan tersebut, solusi yang dapat diusulkan adalah penerapan metode *data mining* dengan algoritma Apriori untuk menganalisis pola hubungan antara faktor-faktor yang mempengaruhi prestasi akademik siswa di MAN Karo. Algoritma Apriori dikenal sebagai salah satu metode dalam *data mining* yang dapat mengidentifikasi hubungan asosiasi antar variabel dalam dataset yang besar [10]. *Association rule mining* dengan algoritma Apriori telah terbukti efektif dalam menemukan pola hubungan antar variabel dengan tingkat *support* dan *confidence* yang tinggi [11]. Teknik ini juga telah digunakan dalam seleksi penjurusan siswa dengan hasil yang memuaskan [12]. Dengan menerapkan algoritma ini, sekolah dapat menemukan pola faktor yang sering muncul bersama dan memiliki pengaruh signifikan terhadap prestasi akademik siswa. Sistem informasi berbasis algoritma Apriori diharapkan mampu membantu sekolah dalam menyusun kebijakan pendidikan yang lebih akurat, relevan, dan berbasis data nyata dari siswa.

Penelitian terdahulu telah banyak membahas berbagai metode analisis untuk mengidentifikasi faktor-faktor yang memengaruhi prestasi akademik siswa maupun mahasiswa. Salah satu metode yang digunakan adalah regresi logistik biner yang diterapkan untuk mengidentifikasi faktor yang memengaruhi Indeks Prestasi Kumulatif (IPK) mahasiswa, dengan hasil bahwa variabel seperti motivasi, kedisiplinan, dan kehadiran memiliki pengaruh signifikan terhadap pencapaian akademik [13]. Selain itu, pendekatan *Structural Equation Modeling* (SEM) juga digunakan untuk menganalisis hubungan kompleks antar variabel akademik, seperti motivasi, fasilitas belajar, dan dukungan sosial, yang menunjukkan bahwa faktor internal dan eksternal saling berhubungan dalam menentukan prestasi siswa [14]. Namun, metode SEM membutuhkan asumsi statistik yang cukup ketat dan analisis yang lebih kompleks sehingga tidak selalu mudah diterapkan pada data sekolah dengan kategori sederhana.

Metode lain yang cukup populer adalah *Decision Tree* C4.5 yang digunakan dalam prediksi prestasi belajar siswa dengan menghasilkan klasifikasi berdasarkan atribut dominan seperti nilai tugas, absensi, dan sikap belajar [15]. Kelebihan metode ini adalah mampu menghasilkan aturan keputusan yang mudah dipahami, tetapi kelemahannya adalah hasil model sering bergantung pada atribut dominan tertentu dan kurang efektif dalam menggambarkan hubungan antar variabel yang muncul bersamaan. Selain itu, metode CHAID juga diterapkan untuk mengklasifikasikan faktor prestasi akademik mahasiswa, dengan hasil bahwa variabel sosial ekonomi, kehadiran, dan aktivitas akademik menjadi faktor utama dalam pengelompokan prestasi [16]. Meskipun efektif dalam klasifikasi, CHAID lebih berfokus pada segmentasi data daripada menggali pola asosiasi antar faktor.

Pendekatan lain seperti algoritma *K-Nearest Neighbor* (KNN) telah digunakan untuk klasifikasi data pendidikan dengan menghasilkan akurasi yang cukup baik dalam memprediksi kategori prestasi siswa [17]. Namun, KNN memiliki keterbatasan dalam interpretasi pola karena hanya menghasilkan hasil klasifikasi tanpa menjelaskan keterkaitan antar faktor secara eksplisit. Selain itu, metode clustering seperti K-Means juga digunakan untuk mengelompokkan siswa berdasarkan karakteristik prestasi akademik, sehingga sekolah dapat mengenali kelompok siswa berprestasi tinggi dan rendah [18]. Penelitian lain bahkan mengkombinasikan clustering dengan analisis korelasi untuk menghasilkan intervensi pembelajaran yang lebih tepat sasaran [19]. Akan tetapi, pendekatan clustering hanya membentuk kelompok tanpa memberikan aturan hubungan sebab-akibat atau asosiasi antar variabel secara detail.

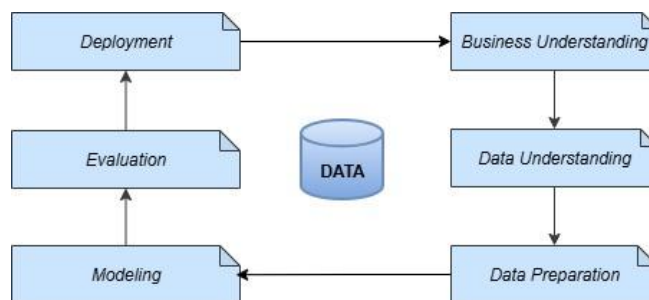
Selain metode klasifikasi dan clustering, penelitian lain juga menggunakan analisis faktor untuk mengetahui penyebab rendahnya prestasi siswa, khususnya pada mata pelajaran matematika. Hasil penelitian tersebut menunjukkan bahwa faktor kemampuan dasar, metode belajar, serta dukungan lingkungan belajar berperan besar dalam memengaruhi hasil ujian [20]. Walaupun mampu mereduksi variabel menjadi beberapa faktor utama, analisis faktor tidak secara langsung memberikan aturan hubungan antar variabel yang dapat digunakan sebagai dasar rekomendasi kebijakan sekolah.

Berdasarkan berbagai penelitian tersebut, dapat disimpulkan bahwa sebagian besar metode yang digunakan lebih berfokus pada prediksi dan klasifikasi prestasi akademik. Namun, penelitian yang secara spesifik menggali hubungan kombinasi faktor internal dan eksternal dalam bentuk aturan asosiasi (*association rules*), terutama pada

konteks madrasah, masih relatif terbatas. Oleh karena itu, penelitian ini menggunakan algoritma Apriori karena mampu menghasilkan pola keterkaitan antar faktor secara lebih jelas melalui nilai *support*, *confidence*, dan *lift*, sehingga dapat membantu pihak sekolah dalam memahami kombinasi faktor dominan yang memengaruhi prestasi akademik siswa.

II. METODE

Penelitian ini menggunakan metode *data mining* dengan algoritma Apriori untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap prestasi akademik siswa Madrasah Aliyah Negeri (MAN) Karo. Data yang digunakan terdiri dari data primer berupa hasil kuesioner yang diberikan kepada 300 siswa dan data sekunder berupa nilai akademik, data absensi, serta catatan kegiatan belajar siswa. Variabel penelitian mencakup faktor internal dan faktor eksternal serta indikator prestasi akademik. Model proses penelitian yang digunakan mengacu pada CRISP-DM (*Cross Industry Standard Process for Data Mining*), yang terdiri dari enam tahapan utama: *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment* [21].



Gambar 1. Tahapan Penelitian Berdasarkan Model CRISP-DM

Pada tahap *business understanding*, peneliti melakukan observasi dan mengidentifikasi masalah utama di MAN Karo, yaitu data akademik siswa belum dimanfaatkan untuk menemukan pola faktor yang berpengaruh terhadap prestasi. Setelah itu ditentukan tujuan penelitian, yaitu menghasilkan aturan asosiasi yang menunjukkan hubungan antara faktor internal dan eksternal terhadap prestasi siswa. Tahap ini juga menetapkan output penelitian berupa nilai *support*, *confidence*, dan *lift* sebagai dasar pengambilan keputusan.

Tahap *data understanding* dilakukan dengan mengumpulkan dan memeriksa data siswa yang diperoleh dari kuesioner, rapor, serta absensi. Pada tahap ini dilakukan eksplorasi data untuk mengetahui karakteristik dan distribusi setiap variabel, serta mendeteksi adanya data kosong atau tidak konsisten. Tahap ini penting agar data yang digunakan benar-benar sesuai dengan kebutuhan analisis Apriori.

Tahap *data preparation* dilakukan dengan membersihkan data dari duplikasi dan nilai kosong, kemudian mengubah data menjadi bentuk kategorikal agar dapat diproses dengan Apriori. Kategori nilai rapor ditetapkan menjadi <75, 75–84, dan ≥85. Kehadiran dikategorikan menjadi tinggi (90–100%), sedang (75–89%), dan rendah (<75%). Variabel lain seperti motivasi, fasilitas belajar, dukungan orang tua, dan les dikategorikan menjadi Ya dan Tidak. Selanjutnya data ditransformasikan ke dalam bentuk transaksi biner (*True/False*) menggunakan *Transaction Encoder* sehingga siap dianalisis.

Tahap modeling dilakukan dengan menerapkan algoritma Apriori menggunakan minimum *support* sebesar 0,05 dan minimum *confidence* sebesar 0,6. Nilai *support* digunakan untuk mengetahui seberapa sering suatu itemset X muncul dalam seluruh transaksi sesuai Rumus (1). Nilai *confidence* digunakan untuk mengukur kekuatan hubungan antara itemset X (*antecedent*) terhadap itemset Y (*consequent*) sesuai Rumus (2). Sementara itu, nilai *lift* digunakan untuk mengetahui tingkat signifikansi hubungan X dan Y sesuai Rumus (3). Dalam penelitian ini, X merupakan kombinasi faktor internal dan eksternal siswa, sedangkan Y merupakan indikator prestasi akademik seperti nilai rapor tinggi. Tiga parameter utama yang digunakan adalah:

1. *Support*

Mengukur probabilitas kemunculan itemset X:

$$Support(X \rightarrow Y) = \frac{\text{Jumlah transaksi yang mengandung } (X \cup Y)}{\text{Total Transaksi}} \quad (1)$$

2. Confidence

Mengukur kekuatan hubungan antara *antecedent* X dan *consequent* Y.

$$Confidence(X \rightarrow Y) = \frac{Support(X \cup Y)}{Support(X)} \quad (2)$$

3. Lift

Menunjukkan seberapa besar pengaruh X terhadap Y.

$$Lift(X \rightarrow Y) = \frac{Confidence(X \rightarrow Y)}{Support(Y)} \quad (3)$$

Tahap *evaluation* dilakukan dengan menilai aturan asosiasi yang dihasilkan berdasarkan nilai *support*, *confidence*, dan *lift*. Aturan yang memenuhi batas minimum *support* dan *confidence* kemudian dipilih, lalu aturan dengan nilai *lift* tertinggi dianggap sebagai aturan paling kuat. Evaluasi juga dilakukan dengan membandingkan hasil sistem dengan perhitungan manual menggunakan Rumus (1), (2), dan (3) untuk memastikan hasil yang diperoleh valid.

Tahap *deployment* dilakukan dengan mengimplementasikan hasil analisis ke dalam sistem informasi berbasis web menggunakan *framework* Laravel dan *database* MySQL [22]. *Output* hasil analisis Apriori dari Python diekspor ke format CSV dan ditampilkan dalam sistem agar pihak sekolah dapat melihat pola hubungan faktor prestasi siswa secara otomatis. Perancangan sistem menggunakan UML seperti *use case diagram*, *activity diagram*, *sequence diagram*, dan *class diagram* [23].

III. HASIL DAN PEMBAHASAN

A. Data Observasi

Penelitian ini melibatkan 300 siswa MAN Karo sebagai responden. Data diperoleh melalui kuesioner yang memuat 16 atribut dan dikelompokkan menjadi tiga kategori utama, yaitu faktor internal, faktor eksternal, serta indikator prestasi akademik. Faktor internal mencakup waktu belajar, motivasi, gaya belajar, dan tingkat kehadiran siswa. Faktor eksternal meliputi dukungan orang tua, pendidikan orang tua, fasilitas belajar, serta keikutsertaan les atau bimbingan belajar. Sementara itu, indikator prestasi akademik mencakup nilai rapor, partisipasi lomba, dan penerimaan beasiswa. Distribusi karakteristik responden berdasarkan variabel-variabel tersebut tersaji pada Tabel 1.

TABEL 1
DISTRIBUSI KARAKTERISTIK RESPONDEN

Variabel	Kategori	Jumlah	Persentase
Motivasi	Ya	232	77,33%
	Tidak	68	22,67%
Kehadiran	Tinggi (90-100%)	191	63,67%
	Sedang (75-89%)	82	27,33%
	Rendah (<75%)	27	9,00%
Les/Bimbingan	Ya	34	11,33%
	Tidak	266	88,67%
Fasilitas	Lengkap	184	61,33%
	Cukup	116	38,67%
Dukungan Ortu	Tinggi	87	29,00%
	Sedang	156	52,00%
	Rendah	57	19,00%
Nilai Rapor	≥85	84	28,00%
	75-84	174	58,00%
	<75	42	14,00%
Beasiswa	Ya	55	18,33%
	Tidak	245	81,67%

Berdasarkan Tabel 1, terlihat bahwa sebagian besar siswa memiliki motivasi belajar tinggi dengan persentase 77,33%, sedangkan siswa yang tidak termotivasi hanya sebesar 22,67%. Dari aspek kehadiran, mayoritas siswa berada pada kategori kehadiran tinggi (63,67%), yang menunjukkan bahwa sebagian besar siswa memiliki tingkat kedisiplinan yang baik dalam mengikuti pembelajaran. Namun, hanya sebagian kecil siswa yang mengikuti les atau bimbingan belajar, yaitu sebesar 11,33%, sedangkan 88,67% lainnya tidak mengikuti les. Selain itu, fasilitas belajar siswa sebagian besar berada pada kategori lengkap (61,33%), dan prestasi rapor siswa mayoritas berada pada kategori 75–84 sebesar 58%. Hal ini menunjukkan bahwa masih terdapat peluang peningkatan prestasi siswa agar lebih banyak mencapai nilai rapor ≥85.

B. Transformasi Data

Tahapan transformasi ini penting karena algoritma Apriori hanya dapat bekerja pada data yang berbentuk itemset atau transaksi yang bernilai benar/salah (True/False). Proses transformasi dilakukan dengan cara mengubah setiap atribut kategorikal menjadi item-item yang berdiri sendiri. Sebagai contoh, atribut motivasi diubah menjadi item “Motivasi=Ya” atau “Motivasi=Tidak”, atribut les menjadi “Les=Ya” atau “Les=Tidak”, dan atribut nilai rapor menjadi “Rapor $\geq$ 85”, “Rapor=75–84”, atau “Rapor<75”. Setelah data dikonversi menjadi bentuk item, setiap responden direpresentasikan sebagai satu transaksi yang berisi kumpulan item sesuai kondisi siswa tersebut. Selanjutnya, data transaksi tersebut diubah menjadi matriks biner menggunakan TransactionEncoder dari pustaka mlxtend pada Python. Pada matriks ini, nilai True menunjukkan bahwa item muncul dalam transaksi siswa, sedangkan nilai False menunjukkan bahwa item tidak muncul. Hasil transformasi data dalam bentuk transaksi biner dapat dilihat pada Tabel 2, yang menunjukkan contoh beberapa transaksi siswa setelah proses encoding dilakukan.

TABEL 2
DATA TRANSAKSI SETELAH DITRANSFORMASIKAN

ID	Les=Tidak	Les=Ya	Motivasi=Ya	Kehadiran=Tinggi	Fasilitas=Lengkap	Rapor $\geq$ 85
0	TRUE	FALSE	TRUE	TRUE	TRUE	FALSE
1	TRUE	FALSE	TRUE	FALSE	FALSE	FALSE
2	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
3	TRUE	FALSE	FALSE	TRUE	TRUE	FALSE
4	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE

C. Hasil Analisis Apriori

Analisis data dilakukan menggunakan algoritma Apriori dengan parameter minimum support sebesar 0,05 dan minimum confidence sebesar 0,6. Hasil analisis menghasilkan frequent itemsets dan aturan asosiasi yang menggambarkan hubungan antar faktor internal maupun eksternal terhadap prestasi akademik siswa. Aturan asosiasi yang diperoleh kemudian dihitung nilai support, confidence, dan lift untuk mengetahui tingkat kekuatan hubungan antar variabel.

Salah satu aturan asosiasi terkuat yang diperoleh adalah aturan {Les=Ya} $\rightarrow$ {Rapor $\geq$ 85, Fasilitas=Lengkap}. Berdasarkan hasil analisis, aturan tersebut memiliki nilai support sebesar 7,73%. Nilai ini menunjukkan bahwa kombinasi siswa yang mengikuti les, memiliki fasilitas lengkap, dan memperoleh nilai rapor $\geq$ 85 muncul pada 7,73% dari total transaksi siswa. Perhitungan confidence berdasarkan Rumus (2) menghasilkan nilai sebesar 67,65%, yang berarti bahwa siswa yang mengikuti les memiliki peluang 67,65% untuk memperoleh nilai rapor $\geq$ 85 dan memiliki fasilitas belajar lengkap. Selanjutnya, nilai lift yang diperoleh sebesar 3,12 yang menunjukkan hubungan positif yang sangat kuat, karena peluang siswa mencapai rapor tinggi meningkat lebih dari tiga kali lipat dibandingkan kondisi acak.

Aturan asosiasi lainnya yang juga signifikan adalah {Les=Ya, DukunganOrtu=Tinggi} $\rightarrow$ {Rapor $\geq$ 85}. Aturan ini memiliki confidence sebesar 85%, yang menunjukkan bahwa siswa yang mengikuti les dan mendapatkan dukungan orang tua tinggi sangat berpeluang mencapai nilai rapor $\geq$ 85. Nilai lift pada aturan ini sebesar 3,04, yang menunjukkan bahwa kombinasi dukungan orang tua dan les merupakan faktor eksternal yang sangat kuat dalam mendukung prestasi akademik siswa.

Selain itu, aturan {Kehadiran=Tinggi, Les=Ya} $\rightarrow$ {Rapor $\geq$ 85} juga menunjukkan hasil yang kuat. Confidence yang diperoleh sebesar 83,33% menunjukkan bahwa siswa dengan kehadiran tinggi yang juga mengikuti les memiliki peluang besar untuk mencapai rapor $\geq$ 85. Nilai lift sebesar 2,98 memperkuat bahwa hubungan kedua faktor tersebut terhadap prestasi siswa bersifat signifikan dan tidak terjadi secara kebetulan.

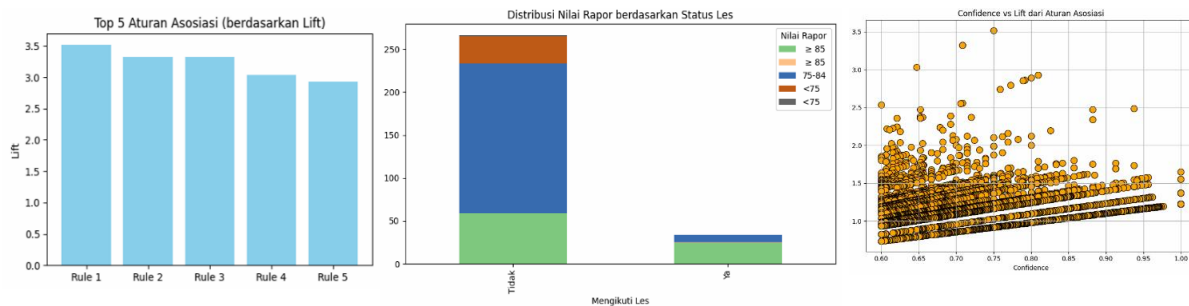
Hasil aturan asosiasi dengan nilai lift tertinggi secara keseluruhan disajikan pada Tabel 4. Berdasarkan Tabel 4, aturan dengan lift tertinggi adalah {Les=Ya} $\rightarrow$ {Rapor $\geq$ 85, Fasilitas=Lengkap} dengan lift 3,1206. Hal ini menunjukkan bahwa keikutsertaan les dan dukungan fasilitas belajar merupakan kombinasi faktor dominan yang paling berpengaruh terhadap prestasi akademik siswa. Dengan demikian, hasil analisis membuktikan bahwa faktor eksternal seperti les dan fasilitas belajar dapat meningkatkan peluang siswa mencapai prestasi tinggi, terutama jika didukung faktor internal seperti kehadiran yang tinggi.

TABEL 4
TOP 5 ATURAN ASOSIASI BERDASARKAN LIFT

No	Antecedent (IF)	Consequent (THEN)	Support	Confidence	Lift
1	Les=Ya	Rapor $\geq$ 85, Fasilitas=Lengkap	0,0773	0,6765	3,1206
2	Les=Ya, DukunganOrtu=Tinggi	Rapor $\geq$ 85	0,0567	0,85	3,0422
3	Kehadiran=Tinggi, Les=Ya	Rapor $\geq$ 85	0,0667	0,8333	2,9819
4	Lomba=1-2 kali/tahun, Les=Ya	Rapor $\geq$ 85	0,0533	0,8	2,8614
5	Les=Ya, Fasilitas=Lengkap	Rapor $\geq$ 85	0,0773	0,7931	2,8349

Hasil perhitungan manual di atas konsisten dengan hasil yang dihasilkan oleh sistem berbasis Python, membuktikan akurasi implementasi algoritma Apriori dalam penelitian ini.

D. Visualisasi Hasil Analisis



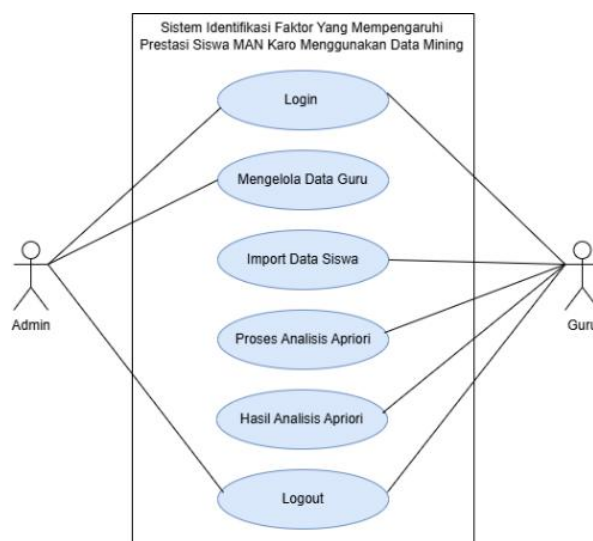
Gambar 2. Visualisasi Nilai (a) Support, (b) Confidence, dan (c) Lift

Hasil aturan asosiasi yang diperoleh kemudian divisualisasikan untuk mempermudah interpretasi nilai *support*, *confidence*, dan *lift* dari aturan-aturan yang terbentuk. Visualisasi pertama ditampilkan pada Gambar 2, yang memperlihatkan perbandingan nilai support dari aturan asosiasi utama. Berdasarkan Gambar 2 (a), aturan yang melibatkan variabel “Les=Ya” cenderung memiliki support lebih tinggi dibandingkan aturan lain, yang menunjukkan bahwa les merupakan faktor yang cukup sering muncul pada siswa dengan prestasi tinggi.

Visualisasi kedua ditampilkan pada Gambar 2 (b), yang menunjukkan perbandingan nilai *confidence* dari aturan asosiasi. Berdasarkan Gambar 2 (b), aturan {Les=Ya, DukunganOrtu=Tinggi} → {Rapor≥85} memiliki *confidence* tertinggi, yang menunjukkan bahwa dukungan orang tua yang tinggi memperkuat pengaruh les terhadap pencapaian prestasi akademik. Selanjutnya, visualisasi ketiga ditampilkan pada Gambar 2 (c), yang menunjukkan nilai *lift* dari aturan asosiasi. Berdasarkan Gambar 2 (c), aturan dengan *lift* tertinggi adalah aturan yang melibatkan les dan fasilitas belajar lengkap. Hal ini menegaskan bahwa kombinasi faktor eksternal tersebut memberikan peningkatan peluang yang signifikan bagi siswa untuk mencapai nilai rapor tinggi. Visualisasi ini membantu menunjukkan bahwa aturan dengan *confidence* tinggi belum tentu memiliki *lift* tertinggi, sehingga *lift* menjadi indikator penting untuk menentukan signifikansi hubungan antar variabel.

E. Diagram UML

Perancangan sistem informasi pada penelitian ini dimodelkan menggunakan *Unified Modeling Language* (UML) untuk menggambarkan struktur dan interaksi pengguna terhadap sistem. Use case diagram digunakan untuk menjelaskan aktor yang terlibat dan fungsi utama sistem, sebagaimana ditunjukkan pada Gambar 3.



Gambar 3. Use Case Diagram Sistem Informasi Identifikasi Faktor Prestasi Siswa

Berdasarkan Gambar 3, terdapat dua aktor utama yaitu admin dan guru. Admin memiliki hak akses untuk melakukan login, mengelola data siswa, dan mengelola data pendukung lainnya. Sementara itu, guru berperan dalam menjalankan proses analisis Apriori serta melihat hasil aturan asosiasi yang terbentuk. Diagram UML ini membantu menggambarkan alur kerja sistem secara jelas dan memudahkan proses implementasi sistem.

F. Implementasi Sistem Informasi

Implementasi sistem dilakukan untuk memudahkan pihak sekolah dalam memanfaatkan hasil analisis data mining secara otomatis melalui sistem berbasis web. Sistem dibangun menggunakan framework Laravel dan database MySQL. Proses utama dalam sistem adalah modul analisis Apriori, yaitu modul yang menjalankan proses pengolahan data dan menampilkan hasil aturan asosiasi.

No	Nama	Kelas	Jurusan	JK	Waktu Belajar	Motivasi	Gaya Belajar	Kehadiran	Dukungan Ortu	Pendidikan Ortu	Fasilitas	Les	Raport	Lomba	Beasiswa	Aksi
1	Arbaim Zahrir Ramadhan	XI	IPS	L	1-2 jam	Ya	Kinestetik/prak...	Tinggi (100%)	Rendah	SMA	Lengkap/luar b...	Tidak	75-84	1-2 kali/tahun	Tidak	[Edit] [Hapus]
2	Ramliansyah	XI	IPS	P	1-2 jam	Ya	Kinestetik/prak...	Rendah (<60%)	Sedang	Perguruan Tinggi	Lengkap/luar b...	Tidak	75-84	Tidak pernah	Tidak	[Edit] [Hapus]
3	Inayah Tusaqilah	XI	IPS	P	<1 jam	Tidak	Visual/metah...	Sedang (60-80%)	Sedang	SMA	Tidak ada	Tidak	<75	Tidak pernah	Tidak	[Edit] [Hapus]
4	Rea Ramadani	XI	IPS	P	1-2 jam	Ya	Visual/metah...	Tinggi (100%)	Sedang	SMP	Lengkap/luar b...	Tidak	≥ 85	Tidak pernah	Ya	[Edit] [Hapus]

Gambar 4. Form Input Data Siswa

Form *input* data siswa digunakan untuk memasukkan atribut penelitian seperti kehadiran, motivasi belajar, dukungan orang tua, fasilitas, keikutsertaan les, serta nilai rapor. Data ini tersimpan di database dan akan diproses dalam analisis *data mining*.

Parameter Apriori

Minimum Support (0.01 - 1.0)
0.05

Threshold untuk itemset yang sering muncul (default: 0.05 = 5%)
0.05

Minimum Confidence (0.1 - 1.0)
0.6

Threshold kepercayaan aturan asosiasi (default: 0.6 = 60%)

Info Data:
Total siswa yang akan diproses: 300 siswa
Jumlah atribut: 14 variabel

Mulai Proses Apriori

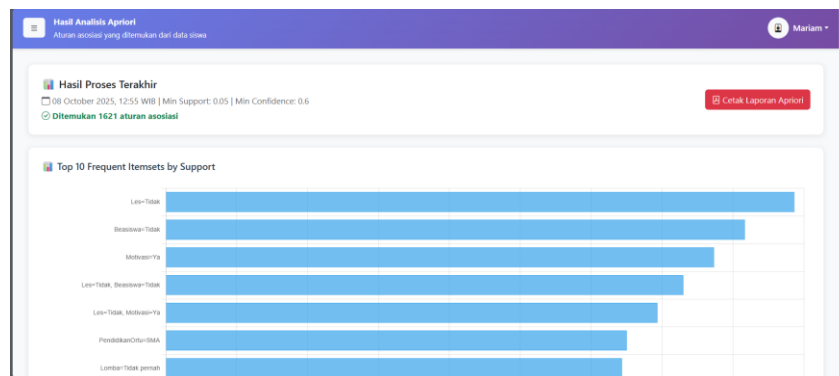
Reset Hasil Sebelumnya

Progress
Siap untuk memproses...

Log Proses
[2025-10-08 12:55:06] Sistem siap...

Gambar 5. Proses Penerapan Algoritma Apriori

Gambar diatas menunjukkan proses penerapan algoritma Apriori pada dataset siswa MAN Karo. Sistem mengeksekusi perhitungan *support*, *confidence*, dan *lift* menggunakan data yang telah ditransformasi ke bentuk boolean. Proses ini berjalan otomatis setelah pengguna menekan tombol “Proses Analisis”, dan hasilnya akan ditampilkan dalam tabel aturan asosiasi.



Gambar 6. Hasil Analisis Aturan Asosiasi Menggunakan Algoritma Apriori

Pada Gambar 9 terlihat tampilan hasil analisis berupa tabel aturan asosiasi dengan nilai *support*, *confidence*, dan *lift*. Nilai *lift* tertinggi menunjukkan kombinasi faktor yang paling berpengaruh terhadap prestasi siswa, seperti kombinasi *Les=Ya* dan *Dukungan Orang Tua=Tinggi*.

Penelitian ini memperkuat temuan studi sebelumnya bahwa faktor internal seperti motivasi dan kehadiran memiliki peran besar dalam membentuk prestasi, sementara faktor eksternal seperti dukungan keluarga, fasilitas belajar, dan bimbingan tambahan memperkuat efek positif dari faktor internal. Kombinasi *Les=Ya* dengan *Dukungan Orang Tua Tinggi* menghasilkan Rapor ≥ 85 dengan *confidence* 85% dan *lift* 3,04, menandakan dampak yang sangat signifikan. Pendekatan *data mining* terbukti berguna untuk merumuskan kebijakan berbasis fakta empiris di lingkungan pendidikan.

IV. SIMPULAN

Berdasarkan hasil penelitian dan implementasi sistem informasi identifikasi faktor yang memengaruhi prestasi siswa MAN Karo menggunakan pendekatan data mining dengan algoritma Apriori, dapat disimpulkan bahwa algoritma Apriori berhasil menemukan hubungan signifikan antara faktor internal dan eksternal siswa terhadap prestasi akademik. Kombinasi faktor *Les = Ya* dan *Dukungan Orang Tua = Tinggi* menunjukkan hubungan paling kuat terhadap pencapaian nilai rapor ≥ 85 , dengan nilai *confidence* sebesar 0,85 dan *lift* sebesar 3,04. Selain itu, faktor kehadiran tinggi, motivasi belajar, serta fasilitas belajar yang lengkap juga berkontribusi dalam meningkatkan peluang siswa memperoleh prestasi akademik yang lebih baik. Sistem informasi berbasis web yang dikembangkan menggunakan framework Laravel dan basis data MySQL mampu mengintegrasikan hasil analisis Apriori dari Python sehingga proses identifikasi faktor prestasi dapat dilakukan secara otomatis, cepat, dan efisien. Dengan demikian, sistem ini dapat mendukung pihak sekolah dalam pengambilan keputusan berbasis data untuk meningkatkan mutu pembelajaran dan strategi peningkatan prestasi siswa. Sebagai pengembangan penelitian selanjutnya, penelitian ini dapat diperluas dengan menambahkan jumlah atribut dan dataset yang lebih besar agar aturan asosiasi yang dihasilkan semakin akurat dan representatif. Selain itu, penelitian lanjutan dapat mengkombinasikan algoritma Apriori dengan metode klasifikasi seperti Decision Tree atau *Random Forest* untuk membandingkan hasil prediksi prestasi siswa. Pengembangan lainnya dapat dilakukan dengan menerapkan algoritma association rule mining lain seperti FP-Growth untuk meningkatkan efisiensi proses pencarian frequent itemsets. Sistem juga dapat dikembangkan dengan menambahkan fitur visualisasi interaktif dan rekomendasi otomatis berbasis aturan asosiasi sehingga pihak sekolah dapat memperoleh saran strategi peningkatan prestasi siswa secara lebih spesifik.

DAFTAR PUSTAKA

- [1] A. Dima, M. A. Kleden, and A. Atti, "Analisis Faktor-Faktor yang Mempengaruhi Prestasi Siswa Menggunakan Metode Structural Equation Modeling (SEM)," *Statistika*, vol. 23, no. 2, pp. 132–146, 2023, doi: 10.29313/statistika.v23i2.2642.
- [2] A. Anas and A. J. Zebua, "Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Mendapatkan Pola Kelulusan Mahasiswa," *Jurnal Ilmiah MEDIA SISFO*, vol. 16, no. 1, pp. 54–61, 2022.
- [3] Amrin Amrin, "Data Mining Dengan Algoritma Apriori untuk Penentuan Aturan Asosiasi Pola Pembelian Pupuk," *Paradigma*, vol. XIX, no. 1, pp. 74–79, 2017, doi: <https://doi.org/10.31294/p.v19i1.1836>.
- [4] F. Zahra, M. A. Ridla, and N. Azise, "Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus: Toko Sinar Harahap)," *JUSTIFY: Jurnal Sistem Informasi Ibrahimi*, vol. 3, no. 1, pp. 55–65, 2024, doi: 10.35316/justify.v3i1.5335.
- [5] P. Studi, S. Informasi, and S. K. Binjai, "Korelasi Perilaku Siswa dengan Tingkat Prestasi Siswa Menggunakan Metode Apriori," no. 4, 2024.
- [6] M. Saleh, T. Chamidy, and Suhartono, "Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Kepuasan Siswa Mts Surya Buana Menggunakan Metode Regresi Logistik," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 3463–3470, 2024, doi: 10.36040/jati.v8i3.9696.
- [7] U. Wahidah, Y. Hermawan, and G. Gumilar, "Pengaruh Kesiapan Belajar Dan Dukungan Sosial Terhadap Prestasi Belajar Mahasiswa Pasca Pandemi," *Jurnal Sains Student Research*, vol. 1, no. 2, pp. 926–932, 2023.
- [8] 2020 Purwodaminta Hendro, "Korelasi Untuk Mengetahui Prestasi Siswa Terhadap Sosial Ekonomi Keluarga, Kegiatan Siswa Diluar Lingkungan Sekolah dan Tingkat Motivasi Belajar Siswa Menggunakan Metode Apriori," vol. 2, no. 4, 2024.
- [9] K. Hulu *et al.*, "IDENTIFIKASI FAKTOR-FAKTOR YANG MEMPENGARUHI MENGGUNAKAN METODE CHAID," vol. 5, no. 3, pp. 1718–1727, 2024.
- [10] A. H. As, "Penerapan Data Mining untuk Menemukan Pola Asosiasi Aktivitas Belajar dan Prestasi Santri Menggunakan Algoritma Apriori," vol. 5, no. 4, pp. 653–662, 2024, doi: 10.33650/trilogi.v5i4.9919.
- [11] Y. A. Ünvan, "Market basket analysis with association rules," *Commun. Stat. Theory Methods*, vol. 50, no. 7, pp. 1615–1628, 2021, doi: 10.1080/03610926.2020.1716255.
- [12] R. P. Karina, S. Lestanti, and F. Febrinita, "Penerapan Algoritma Apriori Dalam Seleksi Penjurusan Calon Peserta Didik Baru Di Smak Diponegoro Blitar," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 2, pp. 716–724, 2022, doi: 10.36040/jati.v6i2.5630.
- [13] Andi Mariani, A. Sauddin, Adiatma, and H. Nufus, "Analisis Regresi Logistik Biner Untuk Mengidentifikasi Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif Mahasiswa Matematika Universitas Islam Negeri Alauddin Makassar," *Jurnal MSA (Matematika dan Statistika serta Aplikasinya)*, vol. 11, no. 1, pp. 6–11, 2023, doi: 10.24252/msa.v11i1.34860.
- [14] P. A. Setyawan, S. H. Wijoyo, and F. Amalia, "Analisis Faktor-Faktor yang Mempengaruhi Prestasi Belajar Siswa Kelas 11 Rekeyasa Perangkat Lunak pada Mata Pelajaran Basis Data di SMK National Media Center Malang," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 2, pp. 863–871, 2023.

- [15] H. Lestari, A. Irma Purnamasari, and T. Suprpti, "Penerapan Data Mining Menggunakan Algoritma C4.5 Untuk Prediksi Prestasi Belajar Siswa Di Mts Yamuallim Panongan," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 2, pp. 1992–1999, 2024, doi: 10.36040/jati.v8i2.8312.
- [16] J. J. R. Faggidae, "Klasifikasi Faktor-faktor yang Mempengaruhi Prestasi Akademik Mahasiswa Pendidikan Matematika FKIP Undana dengan Metode CHAID," *Fraktal: Jurnal Matematika Dan Pendidikan Matematika*, vol. 2, no. 1, pp. 23–33, 2021, doi: 10.35508/fractal.v2i1.4018.
- [17] R. K. Dinata, H. Akbar, and N. Hasdyna, "Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, pp. 104–111, 2020, doi: 10.33096/ilkom.v12i2.539.104-111.
- [18] S. N. Br Sembiring, H. Winata, and S. Kusnasari, "Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means," *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, vol. 1, no. 1, p. 31, 2022, doi: 10.53513/jursi.v1i1.4784.
- [19] C. Renita *et al.*, "DENGAN TEKNIK CLUSTERING DAN ANALISIS KORELASI," vol. 8, no. 5, pp. 10286–10292, 2024.
- [20] E. Nasri, T. H. Setiawan, H. Warianto, A. Aden, and I. Ilmadi, "Faktor-Faktor Yang Mempengaruhi Rendahnya Prestasi Belajar Ujian Matematika Siswa Dengan Menggunakan Metode Analisis Faktor," *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, vol. 3, no. 1, pp. 12–28, 2022, doi: 10.46306/lb.v3i1.64.
- [21] Y. S. Siregar, D. Handoko, M. Khairani, N. I. Syahputri, H. Harahap, and H. Artikel, "Implementasi Data Mining Klasifikasi Algoritma Chaid Dalam Menentukan Pola Penerima Mahasiswa Baru," *Digital Transformation Technology (Digitech) | e*, vol. 3, no. 2, pp. 978–989, 2023.
- [22] F. Rinaldy, "Sistem Informasi Monitoring Siswa pada Madrasah Aliyah Negeri (MAN) 1 Kabanjahe Berbasis Web," Universitas Islam Negeri Sumatera Utara, 2024.
- [23] Suendri, "Implementasi Diagram UML (Unified Modelling Language) Pada Perancangan Sistem Informasi Remunerasi Dosen Dengan Database Oracle (Studi Kasus: UIN Sumatera Utara Medan)," *Jurnal Ilmu Komputer dan Informatika*, vol. 3, no. 1, pp. 1–9, 2018.