

P-ISSN: 2774-4574 ; E-ISSN: 363-4582
TRILOGI, 5(4), Okt-Desember 2024 (653-662)
@2024 Lembaga Penerbitan, Penelitian,
dan Pengabdian kepada Masyarakat (LP3M)
Universitas Nurul Jadid Paiton Probolinggo
DOI: [10.33650/trilogi.v5i4.9919](https://doi.org/10.33650/trilogi.v5i4.9919)



Penerapan Data Mining untuk Menemukan Pola Asosiasi Aktivitas Belajar dan Prestasi Santri Menggunakan Algoritma Apriori

Ahmad Hudawi AS

Universitas Nurul Jadid

ahmad.hudawi@unuja.ac.id

Khoirul Anam

Universitas Nurul Jadid

farelmboiss@gmail.com

Miftahur Rahman

Universitas Nurul Jadid

miftahurrahmannnn@gmail.com

Abstract

Education in Islamic boarding schools is unique, integrating cognitive, character, and spiritual development. This study aims to uncover the relationship between students' learning activities and academic achievement using data mining techniques with the Apriori algorithm. A quantitative approach was applied, analyzing data on learning activities and academic performance. Data preprocessing and the Apriori algorithm were employed using RapidMiner software to identify association patterns between variables. Findings indicate a significant relationship between class attendance, study hours, discussion participation, and extracurricular achievements with academic performance. The association rules revealed that consistent attendance and active learning participation positively influence academic achievement, with confidence levels of 70-80% for "Good" performance categories. This study offers actionable insights for Islamic boarding schools to design more effective learning strategies and emphasizes the potential of data-driven education to enhance student outcomes..

Keywords: Learning Activities; Academic Achievement; Data Mining; Apriori Algorithm; Association Patterns.

Abstrak

Pendidikan di Pondok Pesantren memiliki karakteristik unik dengan mengintegrasikan pengembangan kognitif, karakter, dan spiritual. Penelitian ini bertujuan untuk menggali hubungan antara aktivitas belajar santri dan prestasi akademik menggunakan teknik data mining dengan algoritma Apriori. Pendekatan kuantitatif digunakan untuk menganalisis data

aktivitas belajar dan kinerja akademik. Preprocessing data dan algoritma Apriori diterapkan menggunakan perangkat lunak RapidMiner untuk mengidentifikasi pola asosiasi antar variabel. Hasil menunjukkan adanya hubungan signifikan antara kehadiran kelas, jam belajar, partisipasi diskusi, dan prestasi ekstrakurikuler dengan prestasi akademik. Aturan asosiasi mengungkapkan bahwa konsistensi kehadiran dan partisipasi aktif dalam pembelajaran berpengaruh positif terhadap prestasi akademik, dengan tingkat kepercayaan 70-80% untuk kategori prestasi "Baik." Penelitian ini memberikan wawasan bagi Pondok Pesantren untuk merancang strategi pembelajaran yang lebih efektif serta menekankan potensi pendidikan berbasis data untuk meningkatkan hasil belajar santri.

Katakunci: Aktivitas Belajar; Prestasi Akademik; Data Mining; Algoritma Apriori; Pola Asosiasi.

1 Pendahuluan

Dalam dunia pendidikan prestasi akademik sering dianggap sebagai indikator utama dalam mengukur keberhasilan proses pembelajaran di sekolah. Prestasi ini mencerminkan sejauh mana seorang santri berhasil dalam memperoleh pengetahuan dan keterampilan yang diajarkan oleh pendidik (Muhaimin et al., 2024). Seiring dengan perkembangan zaman pemahaman mengenai bagaimana proses belajar berlangsung semakin kompleks. Proses belajar yang efektif tidak hanya mencakup aspek kognitif, tetapi juga melibatkan perubahan dalam aspek afektif dan psikomotorik santri. Didunia pendidikan Islam khususnya di Pondok Pesantren sistem pendidikan yang diterapkan memiliki kekhasan tersendiri yang menekankan pengembangan kurikulum berbasis nilai-nilai Islam dengan pendekatan yang lebih *holistic* (Belajar et al., 2024).

Proses pendidikan di Pondok Pesantren berbeda dari sekolah pada umumnya, Dimana para santri menerima pendidikan dalam lingkungan yang lebih intensif dan religius dengan pola pendidikan yang khas yaitu mengutamakan pengembangan karakter dan spiritualitas bersama dengan pengajaran ilmu pengetahuan. Kurikulum yang diterapkan di Pondok Pesantren umumnya bersifat fleksibel namun terstruktur dengan mengutamakan pemahaman ilmu agama yang mendalam serta keterampilan praktis (Irfan, 2024). Para santri pun memiliki metode belajar yang berkesinambungan dan sistematis dalam menerima ilmu yang disampaikan yang membuat pendidikan di Pondok Pesantren berbeda dengan pendidikan di sekolah umum.

Meskipun demikian hubungan antara aktivitas belajar para santri dengan prestasi akademik mereka tidak selalu dapat diukur secara langsung. Banyaknya faktor yang mempengaruhi prestasi akademik seperti kondisi fisik, motivasi belajar

lingkungan sosial dan metode pembelajaran membuat analisis hubungan ini menjadi sangat kompleks. Proses untuk menemukan pola yang dapat menggambarkan hubungan antara aktivitas belajar dan prestasi akademik tersebut memerlukan pendekatan yang lebih komprehensif. Dalam konteks ini salah satu metode yang dapat digunakan untuk menggali pola hubungan tersebut adalah *data mining* yang memiliki kemampuan untuk mengolah data dalam jumlah besar dan menemukan pola-pola tersembunyi yang mungkin tidak terlihat secara langsung (Bahar, 2019).

Alasan utama mengapa data mining diperlukan adalah karena adanya sejumlah besar data yang dapat digunakan untuk menghasilkan informasi dan knowledge yang berguna, *data mining* menawarkan berbagai metode untuk menganalisis data, salah satunya adalah algoritma *apriori* yang termasuk dalam kategori aturan *asosiasi*. Algoritma *apriori* digunakan untuk menemukan hubungan antar-item dalam suatu dataset (Soepriyono & Triayudi, 2023). Dengan mencari aturan *asosiasi* yang memenuhi syarat minimum untuk *support* dan *apriori* dapat digunakan untuk menganalisis hubungan antara aktivitas belajar para santri dengan prestasi akademik santri. Algoritma *apriori* memiliki kemampuan untuk menggali pola-pola yang menunjukkan adanya hubungan *asosiatif* antar variabel seperti antara waktu belajar dengan nilai ujian atau kehadiran dengan prestasi akhir (Wati & Juita, 2024).

Penelitian ini bertujuan untuk mengidentifikasi pola hubungan antara aktivitas belajar dan prestasi akademik santri di Pondok Pesantren melalui teknik *data mining* dengan pendekatan *asosiasi*. Dengan menggunakan algoritma *apriori* penelitian ini berupaya menemukan variabel atau aktivitas belajar tertentu yang memiliki pengaruh signifikan terhadap hasil akademik santri. Hal ini akan memberikan gambaran lebih jelas mengenai faktor-faktor yang berkontribusi terhadap

pencapaian akademik santri yang pada akhirnya dapat digunakan untuk memperbaiki dan meningkatkan kualitas pendidikan di Pondok Pesantren (Subkhan, n.d.).

Urgensi penelitian ini terletak pada pentingnya pemahaman yang lebih mendalam mengenai faktor-faktor yang memengaruhi prestasi akademik para santri. Dengan mengetahui pola-pola hubungan tersebut lembaga pendidikan seperti Pondok Pesantren dapat menyusun strategi yang lebih efektif dalam mengembangkan kualitas pendidikan mereka, sehingga dapat lebih tepat sasaran dalam meningkatkan prestasi akademik santri. Pendidikan bukan hanya bertujuan untuk memberikan pengetahuan, tetapi juga untuk membentuk karakter dan kepribadian yang baik, yang dapat mendukung kesuksesan akademik maupun kehidupan sosial santri (Suliman, 2021).

Berbagai penelitian yang memanfaatkan teknik data mining untuk menganalisis faktor-faktor yang mempengaruhi prestasi akademik telah banyak dilakukan. Beberapa penelitian sebelumnya telah menggunakan algoritma seperti apriori untuk menggali pola hubungan antara berbagai variabel pendidikan dengan hasil belajar santri. Hasil penelitian tersebut menunjukkan bahwa meskipun seluruh variabel yang mempengaruhi minat belajar dapat ditentukan secara *apriori*, tetap ada hubungan *asosiatif* yang dapat ditemukan antara variabel-variabel seperti nilai mingguan, nilai setoran, nilai akhir, dan rekapan absensi dengan prestasi akademik (Kecerdasan Buatan et al., 2022).

Penelitian yang akan dilakukan ini menjadi sangat relevan untuk memperdalam pemahaman tentang bagaimana aktivitas belajar santri dapat berhubungan dengan prestasi akademik mereka. Algoritma apriori dapat diartikan sebagai suatu proses untuk menemukan suatu aturan *apriori* yang memenuhi syarat minimum untuk *support* dan *confidence*, disamping itu juga *algoritma apriori* menggunakan data yang sudah ada untuk mendapatkan pengetahuan dari frequent itemset (Habiburrahman, 2022). dengan menggunakan teknik *Association Rule Mining (ARM)* dan algoritma *apriori* penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan sistem pendidikan di Pondok Pesantren khususnya dalam memperbaiki dan meningkatkan efektivitas proses belajar mengajar (Hasanul Huda & Syafrullah, n.d.). Penelitian ini juga diharapkan dapat memberikan wawasan baru tentang pola-pola perilaku belajar yang dapat

dijadikan dasar untuk merancang intervensi pendidikan yang lebih tepat dan berbasis data.

2 Metode

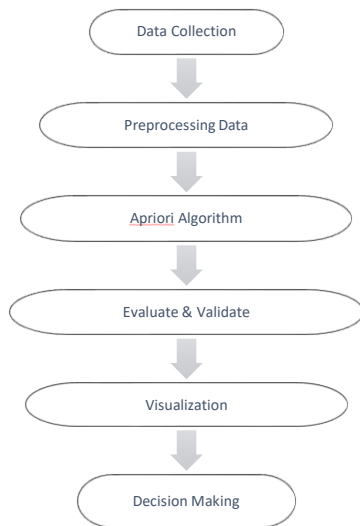
Metode penelitian kuantitatif digunakan karena mampu menghasilkan data numerik yang terstruktur dan memungkinkan pengujian hipotesis secara statistik, pendekatan ini sangat sesuai dalam menganalisis hubungan antara variabel, khususnya aktivitas belajar dan prestasi akademik santri. Dengan pendekatan kuantitatif data dapat dikumpulkan secara sistematis diolah dan dianalisis menggunakan algoritma tertentu seperti data mining, untuk menemukan pola atau hubungan yang signifikan antar variabel (Andi Asari, 2023).

Dalam konteks penelitian ini metode kuantitatif digunakan untuk mengidentifikasi dan mengevaluasi keterkaitan antara aktivitas belajar santri dan prestasi akademik. Melalui teknik data mining pola yang relevan dapat ditemukan sehingga memberikan pemahaman yang lebih baik mengenai faktor-faktor yang memengaruhi prestasi santri. Penggunaan algoritma Apriori dalam proses ini sangat penting karena memungkinkan peneliti untuk menganalisis pola asosiasi yang sering muncul dalam dataset yang besar dan kompleks. Proses ini melibatkan pengumpulan data dalam bentuk angka atau skor yang berkaitan dengan dua variabel utama yaitu aktivitas belajar dan prestasi akademik. Data tersebut kemudian disiapkan dalam format yang siap untuk diproses menggunakan teknik data mining.

Metode penelitian harus memiliki kerangka kerja yang jelas dari tahap awal pengumpulan data hingga analisis dan pengambilan keputusan. Dengan metode kuantitatif tahapan penelitian dapat dilakukan secara sistematis meliputi pengumpulan data, pra-pemrosesan data, penerapan algoritma, evaluasi hasil, visualisasi, hingga pengambilan keputusan berdasarkan temuan yang diperoleh (Yani et al., 2022).

Kerangka Kerja Penelitian berisi kerangka kerja terstruktur dari tahap awal penelitian sampai menghasilkan pencapaian yang diinginkan (Yani et al., 2022). Gambaran proses pengumpulan dan pengolahan data secara lebih detail. Metode kuantitatif memungkinkan peneliti untuk mengukur variabel secara objektif dan menghasilkan data yang dapat diuji secara statistik. Proses penelitian ini sering kali melibatkan penggunaan perangkat lunak atau alat analisis tertentu untuk memastikan keakuratan data. Selain itu, validasi dan reliabilitas data

menjadi fokus utama untuk menjaga kualitas hasil penelitian. Kerangka kerja yang baik akan membantu peneliti merancang tahapan penelitian secara efisien dan menghindari potensi kesalahan dalam interpretasi data.



Gambar 1. Kerangka Metode Penelitian

Kerangka metode penelitian ini terdiri dari beberapa tahapan utama yang masing-masing memiliki peran penting dalam memastikan keberhasilan analisis data:

1. Pengumpulan Data (*Data Collection*)

Tahap ini merupakan fondasi utama dalam penelitian kuantitatif. Data yang digunakan dalam penelitian ini terdiri dari dua variabel utama, yaitu:

- a. **Aktivitas belajar:** Data ini meliputi catatan waktu belajar, keterlibatan dalam kegiatan pembelajaran (seperti kehadiran di kelas, partisipasi dalam diskusi, dan penyelesaian tugas), serta log aktivitas belajar yang terdokumentasi di lembaga pendidikan.
- b. **Prestasi akademik:** Data ini berupa nilai rapor, hasil ujian, peringkat kelas, dan pencapaian akademik lainnya yang terukur secara kuantitatif. Sumber data diperoleh melalui observasi langsung, dokumentasi institusi pendidikan, serta data historis yang tersedia dalam database sekolah atau pondok pesantren. Proses ini dilakukan secara teliti agar memastikan kelengkapan dan keakuratan data yang dikumpulkan.

2. Pra-Pemrosesan Data (*Preprocessing Data*)

Setelah data terkumpul tahap selanjutnya adalah proses pra-pemrosesan data. Menurut Han et al. (2011), pra-pemrosesan data adalah langkah kritis dalam data mining untuk

memastikan kualitas data yang digunakan. Tahapan ini melibatkan beberapa langkah berikut:

- a. **Pembersihan Data:** Menghapus data yang tidak lengkap, duplikat, atau tidak relevan.
- b. **Transformasi Data:** Mengonversi format data dari numerik menjadi kategorikal agar dapat diproses oleh algoritma Apriori. Misalnya, nilai rapor diubah menjadi kategori seperti "tinggi", "sedang", dan "rendah".
- c. **Integrasi Data:** Menggabungkan berbagai sumber data menjadi satu dataset yang utuh dan siap dianalisis.
- d. **Reduksi Data:** Mengeliminasi atribut atau variabel yang tidak relevan untuk fokus pada data yang benar-benar mendukung penelitian.

3. Penerapan Algoritma Apriori (*Apriori Algorithm*)

Pada tahap ini, algoritma *Apriori* digunakan untuk menemukan pola *asosiasi* antara aktivitas belajar dan prestasi akademik santri. Algoritma *Apriori* bekerja dengan mengidentifikasi itemset yang sering muncul (*frequent itemset*) dalam dataset dan membentuk aturan asosiasi berdasarkan nilai *support* dan *confidence*. Adapun tahapan penerapan algoritma Apriori meliputi:

- a. **Penentuan Parameter:** Menetapkan nilai *minimum support* dan *confidence* sebagai ambang batas untuk menentukan pola yang signifikan.
- b. **Frequent Itemset Mining:** Menemukan kombinasi aktivitas belajar yang sering muncul bersama dengan kategori prestasi akademik tertentu.
- c. **Pembentukan Aturan Asosiasi:** Menghasilkan aturan dalam format "Jika aktivitas A terjadi, maka prestasi B juga akan terjadi" dengan tingkat kepercayaan (*confidence*) tertentu. Sebagai contoh jika ditemukan bahwa "santri yang mengikuti pembelajaran tambahan di malam hari memiliki prestasi akademik tinggi", maka aturan tersebut dapat menjadi *insight* yang berharga dalam proses pengambilan keputusan.

4. Evaluasi dan Validasi Hasil (*Evaluate & Validate*)

Setelah pola asosiasi ditemukan, tahap evaluasi dilakukan untuk menguji keakuratan dan validitas aturan yang dihasilkan. Beberapa metrik yang digunakan antara lain:

- a. *Support*: Mengukur seberapa sering kombinasi item tertentu muncul dalam dataset.
- b. *Confidence*: Mengukur seberapa besar tingkat kepastian aturan asosiasi.
- c. *Lift Ratio*: Menilai sejauh mana suatu pola lebih sering muncul dibandingkan dengan kejadian acak.

Teknik validasi seperti *cross-validation* dapat digunakan untuk memastikan generalisasi hasil analisis dan menghindari bias pada dataset tertentu.

5. Visualisasi Hasil (*Visualization*)

Hasil analisis divisualisasikan dalam bentuk grafik, diagram batang, atau jaringan asosiasi untuk memudahkan interpretasi. Visualisasi ini bertujuan untuk memberikan gambaran yang jelas mengenai hubungan antara aktivitas belajar dan prestasi akademik santri antara lain:

- a. Diagram batang untuk menampilkan frekuensi pola tertentu.
- b. Grafik jaringan untuk menggambarkan hubungan antar variabel secara lebih kompleks.

6. Pengambilan Keputusan (*Decision Making*)

Tahap terakhir dalam kerangka metode penelitian ini adalah pengambilan keputusan berdasarkan hasil analisis. Temuan dari penerapan *algoritma Apriori* digunakan untuk merumuskan strategi pembelajaran yang lebih efektif. Rekomendasi yang diberikan bertujuan untuk:

- a. Meningkatkan aktivitas belajar santri yang terbukti berkorelasi positif dengan prestasi akademik.
- b. Merancang kebijakan pembelajaran tambahan atau metode evaluasi yang lebih akurat.
- c. Mengidentifikasi faktor penghambat yang memengaruhi prestasi akademik santri.

7. Penggunaan *RapidMiner* sebagai Tools

Dalam penelitian ini, perangkat lunak *RapidMiner* digunakan untuk memfasilitasi proses data mining. *RapidMiner* merupakan salah satu perangkat *open-source* yang paling populer untuk analisis data, termasuk penerapan *algoritma Apriori*. *RapidMiner* menyediakan lebih dari 500 operator untuk input data, pra-pemrosesan, analisis, dan visualisasi hasil. Keunggulan *RapidMiner* antara lain:

- a. Antarmuka yang *user-friendly* sehingga memudahkan analisis data.
- b. Dukungan terhadap berbagai format data (CSV, Excel, database, dll).
- c. Kemampuan untuk melakukan analisis kompleks seperti asosiasi, klasifikasi, dan prediksi.

Dengan menggunakan *RapidMiner* proses analisis data menjadi lebih efisien, cepat, dan akurat. Perangkat ini memungkinkan penerapan *algoritma Apriori* dengan parameter yang fleksibel sehingga hasil analisis dapat dioptimalkan (Kamil Malik, 2021).

3 Hasil dan Diskusi

Pengumpulan data dalam penelitian ini dilakukan melalui beberapa sumber yang relevan dengan aktivitas belajar santri dan prestasi akademik. Penelitian ini menggunakan dataset sebanyak 900 data santri, dari dataset tersebut dibagi menjadi 70% data latih dan 30% data Uji. Kemudian dilakukan pengujian dengan menggunakan rumus Parameter Asosiasi dalam salah satu teknik dalam data mining yang digunakan untuk menemukan pola hubungan atau asosiasi antara item-item dalam dataset.

Untuk memberikan gambaran yang lebih jelas mengenai data yang digunakan, Tabel 1 di bawah ini menyajikan informasi secara detail mengenai atribut-atribut yang berkaitan langsung dengan aktivitas belajar santri dan pencapaian prestasi akademik santri. Setiap atribut memiliki pengaruh signifikan terhadap hasil analisis dan pola asosiasi antara aktivitas belajar dan prestasi akademik santri. Kehadiran kelas, jam belajar, partisipasi diskusi, nilai tugas, nilai ujian, serta prestasi ekstrakurikuler merupakan variabel utama yang mencerminkan keterlibatan santri dalam proses pembelajaran.

Tabel 1. Dataset santri

Kehadiran (%)	(Jam/ Minggu)	Partisipasi (%)	N. Tugas (%)	N. Ujian (%)	Prestasi (%)	Rata-rata Nilai Akademik	Prestasi Akademik
78	32	65	86	90	90	74	Baik
78	40	70	86	90	90	76	Baik
80	32	75	86	90	90	76	Baik
78	40	70	85	90	90	76	Baik
78	40	75	70	70	90	71	Baik
80	40	75	86	90	65	73	Baik
82	40	70	86	90	90	76	Baik

85	40	80	88	90	90	79	Baik
80	40	65	85	90	90	75	Baik
80	40	75	85	90	90	77	Baik
78	40	70	85	90	90	76	Baik
82	40	70	86	90	90	76	Baik
78	40	70	86	90	90	76	Baik
80	40	70	85	90	90	76	Baik
78	40	65	86	90	90	75	Baik
78	40	70	86	90	90	76	Baik
80	40	70	86	90	90	76	Baik
80	40	70	70	77	90	71	Baik
78	40	75	85	90	65	72	Baik
78	40	75	85	90	90	76	Baik
78	40	70	85	90	90	76	Baik
90	40	20	90	90	90	70	Baik
80	40	70	85	75	90	73	Baik
75	40	65	85	75	75	69	Cukup

Tabel 1 menyajikan informasi secara terperinci mengenai data santri yang digunakan dalam penelitian. Dataset ini mencakup beberapa atribut utama yang relevan untuk analisis hubungan antara aktivitas belajar dan prestasi akademik santri. Atribut-atribut tersebut meliputi antara lain : Kehadiran Kelas (%), Jam Belajar (Jam/Minggu), Partisipasi Diskusi Kelas (%), Nilai Tugas (%), Nilai Ujian (%), Prestasi Ekstrakurikuler (%), Rata-rata Nilai Akademik, dan Prestasi Akademik. Masing-masing atribut mencerminkan indikator penting dalam evaluasi kinerja akademik santri serta pola aktivitas belajar yang berkontribusi terhadap pencapaian prestasi.

Pada dataset ini setiap atribut memiliki peran signifikan dalam analisis kehadiran Kelas menggambarkan konsistensi santri dalam mengikuti pembelajaran formal, sedangkan jam belajar mencerminkan tingkat usaha mandiri dalam memahami materi pelajaran disamping itu partisipasi diskusi berhubungan dengan keterlibatan aktif santri dalam pembelajaran kelompok yang seringkali berdampak positif terhadap Rata-rata Nilai Akademik. Di sisi lain atribut seperti Nilai Tugas dan Nilai Ujian mencerminkan pencapaian akademik formal yang menjadi tolok ukur penilaian, sedangkan Prestasi Ekstrakurikuler mengindikasikan keterlibatan santri di luar kelas. Prestasi Akademik dikelompokkan menjadi kategori seperti Baik atau Cukup, yang memberikan wawasan tambahan tentang capaian akademik keseluruhan santri.

Langkah selanjutnya adalah membagi data menjadi dua bagian utama, yaitu data latih dan

data uji. Proses ini bertujuan untuk memastikan bahwa model data mining yang dibangun dapat dievaluasi secara objektif berdasarkan performa prediksi atau analisisnya. Proses pembagian dataset ini sangat penting untuk memastikan bahwa model yang dihasilkan tidak hanya cocok untuk data yang digunakan selama pelatihan tetapi juga memiliki kemampuan generalisasi untuk data baru yang digambarkan pada tabel 2 berikut:

Tabel 2. Data Latih

Data santri	Deskripsi	Rata_rata
70%	Kehadiran Kelas (%)	79,3
	Jam Belajar (Jam/Minggu)	40,0
	Partisipasi Diskusi Kelas (%)	78,6
	Prestasi Ekstrakurikuler	78,1

Tabel 1, adalah data latih sebanyak 70% dengan rata-rata tingkat kehadiran 79,3%, Jam Belajar rata-rata 40,0%, Partisipasi Diskusi rata-rata 78,6% serta Prestasi ekstrakurikuler 78,1%.

Tabel 3. Data Uji

Data santri	Deskripsi	Rata_rata
30%data	Kehadiran Kelas (%)	77,1
	Jam Belajar (Jam/Minggu)	40,0
	Partisipasi Diskusi Kelas (%)	78,3
	Prestasi Ekstrakurikuler	81,2

Tabel 2, adalah data Uji sebanyak 30% dengan rata-rata tingkat kehadiran 77,1%, Jam Belajar rata-rata 40,0%, Partisipasi Diskusi rata-rata 78,3% serta Prestasi ekstrakurikuler 78,2%.

Dalam percobaan manual, menggunakan rumus excel penerapan Algoritma Apriori untuk mencari pola asosiasi yang menghubungkan aktivitas belajar dan prestasi santri, dengan menggunakan rumus-rumus dasar sebagai berikut:

Menghitung Support, Confidence, dan Lift Support: Mengukur seberapa sering item atau kombinasi item muncul dalam data:

Support:

- Support Untuk Kehadiran Kelas > 80% dan Jam Belajar > 35

$$Support(Kehadiran > 80\% \& Jam Belajar > 35) = \frac{Jumlah\ T1, T3, T5}{5} = \frac{3}{5} = 0.6$$

2. Support untuk Kehadiran Kelas > 80% dan Partisipasi Diskusi > 70%

$$\text{Support}(\text{Kehadiran} > 80\% \& \text{Partisipasi} > 70\%) = \frac{\text{Jumlah T3, T5}}{5} = \frac{2}{5} = 0.4$$

Confidence:

1. Confidence untuk Kehadiran > 80% → Jam Belajar > 35:

$$\text{Confidence}(\text{Kehadiran} > 80\% \rightarrow \text{Jam Belajar} > 35) = \frac{3}{4} = 0.75$$

2. Confidence untuk **Kehadiran > 80%** → **Partisipasi > 70%**:

$$\text{Confidence}(\text{Kehadiran} > 80\% \rightarrow \text{Partisipasi} > 70) = \frac{2}{4} = 0.5$$

Artinya, 50% transaksi yang memiliki Kehadiran > 80% juga memiliki Partisipasi > 70%.

Lift:

1. Lift untuk Kehadiran > 80% → Jam Belajar > 35:

$$\text{Lift}(\text{Kehadiran} > 80\% \rightarrow \text{Jam Belajar} > 35) = \frac{0.75}{0.6} = 1.25$$

Lift > 1 menunjukkan adanya hubungan positif antara kedua item.

2. Lift untuk Kehadiran > 80 % → Partisipasi > 70:

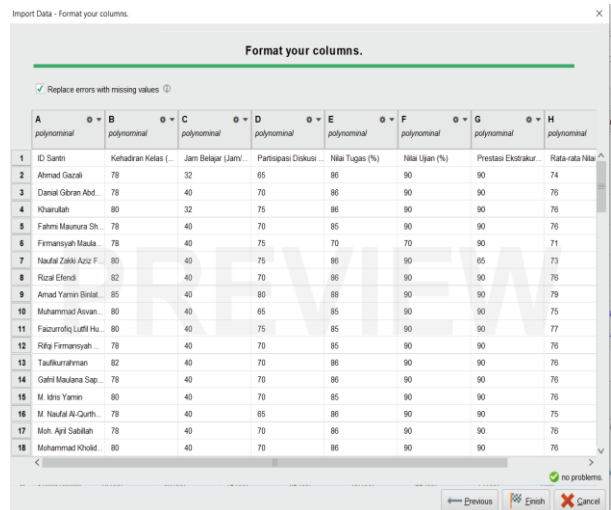
$$\text{Lift}(\text{Kehadiran} > 80\% \rightarrow \text{Partisipasi} > 70) = \frac{0.5}{0.6} = 0.83$$

Lift < 1 menunjukkan hubungan negatif atau tidak signifikan.

Kehadiran > 80% dan Jam Belajar > 35 memiliki hubungan positif (Lift > 1).

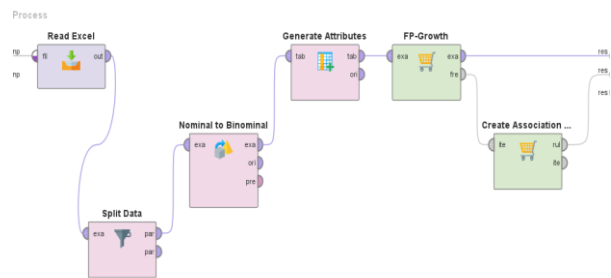
Kehadiran > 80% dan Partisipasi > 70% menunjukkan hubungan yang kurang kuat (Lift < 1).

Penerapan Algoritma Apriori untuk mencari pola asosiasi yang menghubungkan aktivitas belajar dan prestasi santri untuk menghitung Support, Confidence, dan Lift Support menggunakan tools rapidminer berikut :



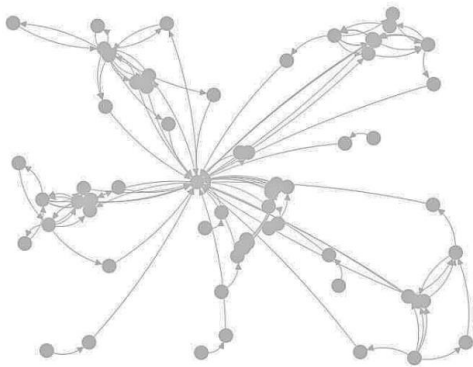
Gambar 2. Import Read Excel Sheet ke Rapidminer

Gambar 2. menunjukkan langkah-langkah dalam proses import dataset dari file Excel ke RapidMiner. Proses ini bertujuan untuk mempersiapkan data sehingga dapat digunakan untuk analisis lebih lanjut seperti penerapan algoritma Apriori untuk menemukan pola asosiasi.



Gambar 3. Model Perhitungan algoritma *FP-Growth*

Pada gambar 3, Algoritma *FP-Growth* (*Frequent Pattern Growth*) adalah salah satu metode populer untuk menemukan pola asosiasi atau aturan *asosiasi* (*association rules*) dalam data mining. *FP-Growth* bekerja lebih efisien dibandingkan algoritma Apriori itu karena tidak memerlukan proses eksplisit untuk menghasilkan kandidat itemset sehingga dapat menghemat waktu dan sumber daya komputasi.



Gambar 4. Output Atau Representasi Visual Dari Aturan Asosiasi

Penjelasan visualisasi *graph* pada Gambar 4, menunjukkan node pusat dengan banyak *edges* menandakan atribut tertentu dalam aktivitas belajar santri memiliki keterkaitan kuat dengan prestasi akademik. Node ini merepresentasikan faktor kunci yang paling berpengaruh. Ketebalan garis mencerminkan tingkat kepercayaan (*confidence*) atau dukungan (*support*) aturan *asosiasi*, dimana garis tebal menunjukkan hubungan signifikan. Klaster yang terbentuk mengindikasikan pola spesifik, seperti aktivitas belajar serupa yang berhubungan dengan prestasi tertentu, arah panah menggambarkan hubungan kausal, dimana aktivitas belajar menjadi *antecedent* (sebab) dan prestasi sebagai *consequent* (akibat). Item dengan garis tebal dan central node menjadi prioritas utama untuk perbaikan metode belajar.

Association Rules

```

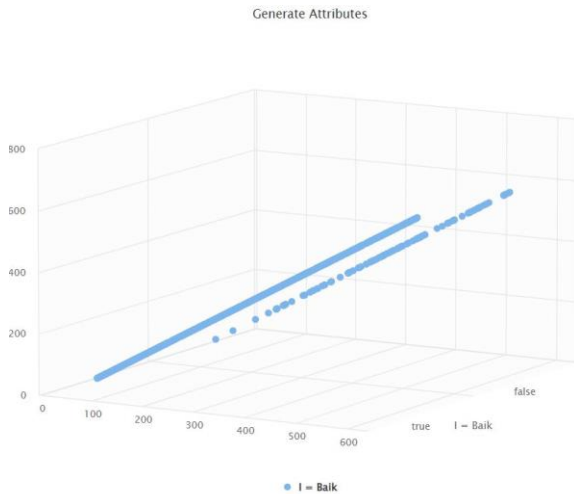
Association Rules
[B = 75] --> [I = Baik, D = 75] (confidence: 0.625)
[B = 75] --> [C, I = Baik, D = 75] (confidence: 0.625)
[D = 75] --> [I = Baik, B = 75] (confidence: 0.649)
[D = 75] --> [C, I = Baik, B = 75] (confidence: 0.649)
[C, D = 75] --> [I = Baik, B = 75] (confidence: 0.654)
[D = 85] --> [I = Baik, B = 85] (confidence: 0.664)
[D = 85] --> [C, I = Baik, B = 85] (confidence: 0.664)
[C, D = 85] --> [I = Baik, B = 85] (confidence: 0.664)
[D = 85] --> [B = 85] (confidence: 0.720)
[D = 85] --> [C, B = 85] (confidence: 0.720)
[C, D = 85] --> [B = 85] (confidence: 0.720)
[I = Baik, B = 75] --> [D = 75] (confidence: 0.725)
[I = Baik, B = 75] --> [C, D = 75] (confidence: 0.725)
[I = Baik, D = 85] --> [B = 85] (confidence: 0.732)
[I = Baik, D = 85] --> [C, B = 85] (confidence: 0.732)
[C, I = Baik, D = 85] --> [B = 85] (confidence: 0.732)
[B = 80] --> [I = Baik, D = 80] (confidence: 0.737)
[B = 80] --> [C, I = Baik, D = 80] (confidence: 0.737)
[C, B = 80] --> [I = Baik, D = 80] (confidence: 0.742)
[B = 75] --> [D = 75] (confidence: 0.744)
[B = 75] --> [C, D = 75] (confidence: 0.744)
[B = 85] --> [I = Baik, D = 85] (confidence: 0.747)
[B = 85] --> [C, I = Baik, D = 85] (confidence: 0.747)
[C, B = 85] --> [I = Baik, D = 85] (confidence: 0.747)
[B = 80] --> [D = 80] (confidence: 0.752)
[B = 80] --> [C, D = 80] (confidence: 0.752)
[I = Baik, B = 80] --> [D = 80] (confidence: 0.754)
[I = Baik, B = 80] --> [C, D = 80] (confidence: 0.754)
[C, B = 80] --> [D = 80] (confidence: 0.758)
    
```

Gambar 5. Deskripsi Assosiasi Rule

Penjelasan hasil pada gambar 5, menunjukkan aturan asosiasi dengan *confidence* yang menggambarkan kekuatan hubungan antar atribut aktivitas belajar dan prestasi akademik santri. Aturan [C = 75] --> [D = 40] memiliki *confidence* 1.0, menunjukkan hubungan sempurna antara kedua atribut tersebut. Aturan lain seperti [E = 85] --> [D = 40], memiliki *confidence* 0.830, yang menunjukkan hubungan signifikan, nilai *confidence* yang tinggi mengindikasikan tingkat kepastian kuat dari pola yang ditemukan hubungan ini membantu mengidentifikasi aktivitas belajar spesifik yang berkaitan erat dengan prestasi akademik. Dengan demikian aturan dengan *confidence* tertinggi dapat menjadi fokus utama untuk mengevaluasi dan meningkatkan metode belajar santri secara efektif.

Gambar 5 menunjukkan hasil analisis aturan asosiasi yang menggambarkan hubungan antara atribut aktivitas belajar dan prestasi akademik santri, berdasarkan nilai *confidence* sebagai indikator kekuatan hubungan. Aturan dengan *confidence* tertinggi, seperti [C = 75] --> [D = 40] yang memiliki *confidence* 1.0, menunjukkan hubungan sempurna antara kedua atribut tersebut, menegaskan bahwa setiap kali atribut C bernilai 75, atribut D akan bernilai 40 tanpa pengecualian. Selain itu, aturan seperti [E = 85] --> [D = 40] dengan *confidence* 0.830 menunjukkan hubungan signifikan antara atribut E dan D. *Confidence* yang tinggi pada aturan ini menandakan tingkat kepastian kuat terhadap pola hubungan yang ditemukan, memberikan keandalan dalam menginterpretasi hasil analisis. Pola ini membantu mengidentifikasi aktivitas belajar tertentu yang sangat berkaitan dengan prestasi akademik, seperti tingkat kehadiran dan partisipasi diskusi yang konsisten dengan hasil sebelumnya.

Aturan-aturan dengan *confidence* tertinggi ini dapat dijadikan fokus utama dalam mengevaluasi dan merancang strategi pembelajaran yang lebih efektif. Pondok Pesantren dapat memanfaatkan temuan ini untuk mengarahkan perhatian pada aktivitas belajar yang memiliki dampak signifikan, seperti meningkatkan jam belajar mandiri dan memotivasi partisipasi aktif santri dalam diskusi kelompok, guna meningkatkan hasil akademik secara keseluruhan.



Gambar 6. Hasil Visualisasi Exampleset Generate Attributes

Penjelasan pada gambar 6, Interpretasi Aturan 1: Kehadiran $\geq 80\%$ \rightarrow Prestasi = Baik

Support (0.15): 15% dari dataset memiliki Kehadiran $\geq 80\%$ dan Prestasi = Baik. Artinya, aturan ini terjadi cukup sering di dalam dataset.

Confidence (0.80): 80% dari santri yang memiliki Kehadiran $\geq 80\%$ memiliki Prestasi Baik. Ini menunjukkan bahwa ada korelasi yang kuat antara kehadiran dan prestasi akademik yang baik.

Lift (1.25): Karena lift lebih besar dari 1, ini menunjukkan bahwa kehadiran yang tinggi secara positif mempengaruhi prestasi yang baik. Artinya, jika seorang santri hadir lebih banyak, mereka memiliki peluang lebih tinggi untuk mendapatkan prestasi yang baik, lebih besar daripada yang diharapkan secara kebetulan.

4 Kesimpulan

Penelitian ini berhasil mengidentifikasi pola hubungan antara aktivitas belajar dan prestasi akademik santri menggunakan metode data mining dengan algoritma Apriori. Hasil analisis menunjukkan bahwa kehadiran lebih dari 80%, jam belajar lebih dari 40 jam per minggu, serta partisipasi dalam diskusi aktif memiliki hubungan positif yang signifikan dengan pencapaian nilai akademik yang lebih baik. Algoritma Apriori efektif dalam menemukan aturan asosiasi, misalnya partisipasi diskusi tinggi yang memiliki confidence 70-80% untuk prestasi kategori "Baik." Temuan ini memberikan dasar bagi Pondok Pesantren untuk merancang strategi pendidikan yang lebih terarah. Selain itu, hasil ini menekankan pentingnya kolaborasi antara santri dan pengajar dalam menciptakan lingkungan belajar yang mendukung perkembangan akademik dan

personal santri secara keseluruhan. Rekomendasi dari penelitian ini dapat dijadikan pedoman untuk pengelolaan pendidikan berbasis data yang lebih adaptif dan responsif terhadap kebutuhan santri.

5 Referensi

Andi Asari, D. N. (2023). *Dasar Penelitian Kuantitatif*.

Bahar. (2019). *Model Pengujian Akurasi Berbasis Empiris Pada Algoritma A-Priori*.

Belajar, M., Holistik, P., Islam, P., Utomo, T. E., & Rizqa, M. (2024). Strategi Pembinaan Guru ...225. *Instructional Development Journal (IDJ)*, 1, 225-234. <http://ejournal.uin-suska.ac.id/index.php/IDJ>

Habiburrahman. (2022). Akurasi dalam Analisis Kompetensi Calon Tenaga Keperawatan Menggunakan Algoritma Apriori. *Jurnal Informatika Ekonomi Bisnis*, 172-177. <https://doi.org/10.37034/infv4i4.167>

Hasanul Huda, M., & Syafrullah, M. (n.d.). *PENERAPAN METODE ASSOCIATION RULES DENGAN ALGORITMA APRIORI UNTUK MENCARI POLA KETERKAITAN SETIAP ITEM LEWAT PERILAKU PELANGGAN PADA GAVIO BOUTIQUE COLLECTION JAKARTA PUSAT* (Vol. 3).

Irfan, M. I. (2024). C4.5 Forward Selection Based Algorithm For Class Level Classification Of Nurul Jadid Islamic Boarding School Students. *Sinkron*, 8(2), 699-712. <https://doi.org/10.33395/sinkron.v8i2.13514>

Kamil Malik. (2021). *DIAGNOSA PHARYNGITIS MENGGUNAKAN METODE K-NN*.

Muhaimin, A., Hariyadi, M. A., & Imamudin, M. (2024). Klasifikasi Prestasi Akademik Siswa Berdasarkan Nilai Rapor dan Kedisiplinan dengan Metode K-Nearest Neighbor. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 7(1), 193-202. <https://doi.org/10.55338/jikoms.v7i1.2865>

Soepriyono, G., & Triayudi, A. (2023). Implementasi Data Mining dengan Algoritma Apriori dalam Menentukan Pola Pembelian Aksesoris Laptop. *JURNAL MEDIA*

INFORMATIKA BUDIDARMA, 7(4), 2087.
<https://doi.org/10.30865/mib.v7i4.6555>

<https://tunasbangsa.ac.id/ejurnal/index.php/jurasik>

Subkhan, M. K. (n.d.). *ANALISA LOG AKSES DI E-LEARNING MENGGUNAKAN ASSOCIATION RULE MINING UNTUK MENGENALI POLA BELAJAR SISWA (Studi Kasus di SMUN 1 Pare Kediri)*.

Yani, R., Nazir, A., Affandes, M., Mai Candra, R., Amany Akhyar, dan, & Informatika UIN Sultan Syarif Kasim Riau Jl Soebrantas, T. H. (2022). Implementasi Data Mining Untuk Menemukan Pola Asosiasi Data Tracer Study Menggunakan Algoritma Apriori. *Jurnal Nasional Komputasi Dan Teknologi Informasi*, 5(3). <https://alumni.tif.uin-suska.ac.id/https://doi.org/10.32672/jnkti.v5i3.4412>

Wati, O., & Juita, R. (2024). Penerapan Data Mining dengan Algoritma Apriori Pada Data Penjualan Obat Untuk Mengetahui Pola Pembelian Obat Pada Apotek Di Kota Nabire. *Jurnal Riset Sistem Informasi Dan Teknik Informatika (JURASIK*, 9(1), 451-459.