

P-ISSN: 2774-4574 ; E-ISSN: 2774-4582  
TRILOGI, 5(4), September-Desember 2024 (5-24)  
©2020 Lembaga Penerbitan, Penelitian,  
dan Pengabdian kepada Masyarakat (LP3M)  
Universitas Nurul Jadid Paiton Probolinggo  
DOI: <https://doi.org/10.33650/trilogi.v5i4.9919>

**TRILOGI**  
JURNAL  
Ilmu Teknologi, Kesehatan, dan Humaniora

## **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](mailto:ahmad.hudawi@unuja.ac.id)

**Khoirul Anam**

Universitas Nurul Jadid  
[farelmoiss@gmail.com](mailto:farelmoiss@gmail.com)

**Miftahur Rahman**

Universitas Nurul Jadid  
[miftahurrahmannnn@gmail.com](mailto:miftahurrahmannnn@gmail.com)

### **Abstract**

Education in Islamic boarding schools is characterised by a holistic approach that not only prioritises cognitive aspects, but also character development and spirituality, where the relationship between students' learning activities and their academic achievement often cannot be measured directly. Many factors affect academic achievement such as motivation, social environment, and learning methods. Therefore, this research uses data mining techniques with the Apriori algorithm to explore the pattern of the relationship between learning activities and students' academic achievement. The main objective of this research is to identify the relationship patterns of learning activities that are significant in influencing students' academic achievement. This research uses a quantitative approach by collecting data in the form of records of learning activities and students' academic performance. The data is then analysed using the Apriori algorithm with the help of rapidminer tools to find association patterns between variables related to learning activities and academic results with research stages including data collection, pre-processing, application of the Apriori algorithm, evaluation, and visualisation of analysis results. The results show that there is a significant relationship between variables such as class attendance, study hours, discussion participation, and extracurricular achievements with the academic achievement of students, this is evidenced by the support and confidence values in finding an association rule of one quarter of the student population with had a confidence of 70-80% for achievement in the

'Good' category. The findings show that factors such as consistency of attendance and study intensity have a positive impact on academic achievement. This research provides new insights for Pondok Pesantren to develop more effective learning strategies. Further research can be conducted to examine the influence of other external factors, such as socio-economic conditions, on the academic achievement of santri.

**Keywords:** Learning Activity; Academic Achievement; Data Mining; Apriori Algorithm; Association Pattern..

### Abstrak

Pendidikan di Pondok Pesantren memiliki ciri khas dengan pendekatan holistik yang tidak hanya mengutamakan aspek kognitif, tetapi juga pengembangan karakter dan spiritualitas, dimana hubungan antara aktivitas belajar santri dan prestasi akademik mereka sering kali tidak dapat diukur secara langsung. Banyak faktor yang mempengaruhi prestasi akademik seperti motivasi, lingkungan sosial, dan metode pembelajaran. Oleh karena itu penelitian ini menggunakan teknik data mining dengan algoritma Apriori untuk menggali pola hubungan antara aktivitas belajar dan prestasi akademik santri. Tujuan utama penelitian ini adalah untuk mengidentifikasi pola hubungan aktivitas belajar yang signifikan dalam mempengaruhi prestasi akademik santri. Penelitian ini menggunakan pendekatan kuantitatif dengan mengumpulkan data berupa rekaman aktivitas belajar dan prestasi akademik santri. Data tersebut kemudian dianalisis menggunakan algoritma Apriori dengan bantuan tools rapidminer untuk menemukan pola asosiasi antar variabel yang terkait dengan aktivitas belajar dan hasil akademik dengan tahapan-tahapan penelitian meliputi pengumpulan data, pre-processing, penerapan algoritma Apriori, evaluasi, dan visualisasi hasil analisis. Hasil penelitian menunjukkan adanya hubungan signifikan antara variabel seperti kehadiran kelas, jam belajar, partisipasi diskusi, dan prestasi ekstrakurikuler dengan prestasi akademik santri, ini dibuktikan dengan nilai support dan confidence dalam menemukan aturan asosiasi seperti partisipasi diskusi tinggi yang memiliki confidence 70-80% untuk prestasi kategori "Baik." Temuan ini memperlihatkan bahwa faktor-faktor seperti konsistensi kehadiran dan intensitas belajar memiliki dampak positif terhadap pencapaian akademik. Penelitian ini memberikan wawasan baru bagi Pondok Pesantren untuk mengembangkan strategi pembelajaran yang lebih efektif. Penelitian lebih lanjut dapat dilakukan untuk menguji pengaruh faktor eksternal lainnya, seperti kondisi sosial-ekonomi, terhadap prestasi akademik 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

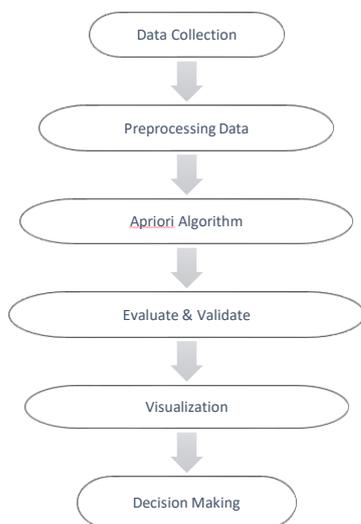
Metode penelitian kuantitatif digunakan karena mampu menghasilkan data numerik yang terstruktur dan memungkinkan pengujian hipotesis secara statistik, pendekatan ini sangat sesuai dalam menganalisis hubungan antara variabel, khususnya aktivitas belajar dan prestasi akademik santri. Dengan pendekatan 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.



**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 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 3 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:

1. Support Untuk Kehadiran Kelas > 80% dan Jam Belajar > 35

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

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:

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

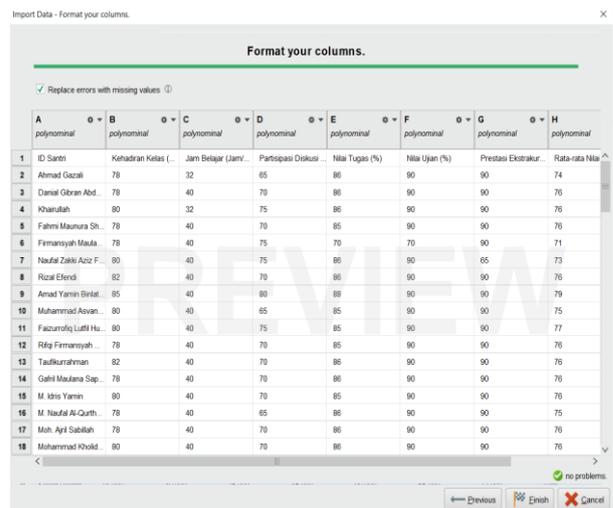
Lift < 1 menunjukkan hubungan negatif atau tidak signifikan.

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

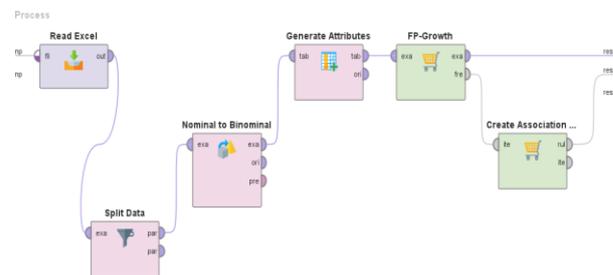
Kehadiran > 80% dan Partisipasi > 70% menunjukkan hubungan yang kurang kuat (Lift < 1).

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

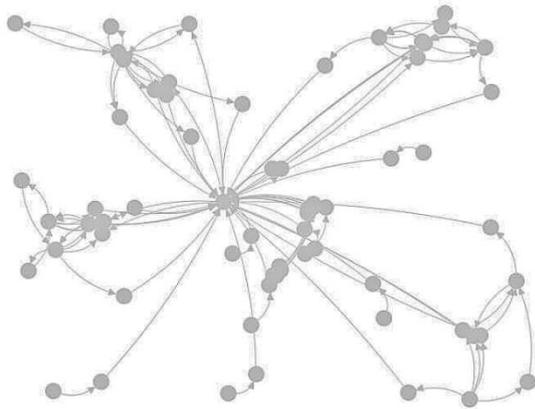


Gambar 2. menunjukkan langkah-langkah dalam proses import dataset dari file Excel ke RapidMiner. Proses ini bertujuan untuk mempersiapkan data sehingga dapat digunakan untuk analisis lebih lanjut seperti penerapan algoritma Apriori untuk menemukan pola asosiasi.



**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.

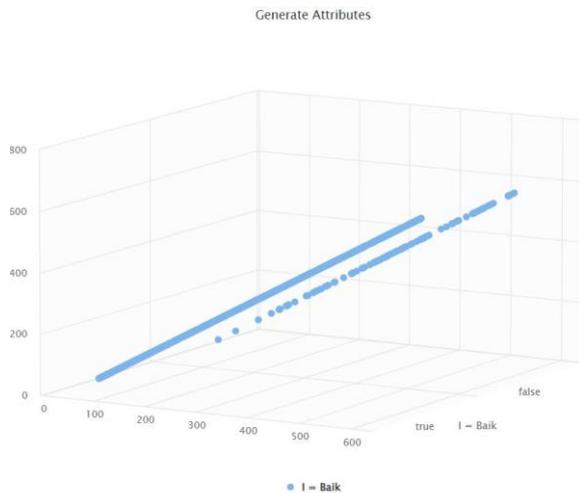
**Association Rules**

```

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 6.** Hasil Visualisasi Exampleset Generate Attributes

Penjelasan pada gambar 6, Interpretasi Aturan 1: Kehadiran ≥ 80% → Prestasi = Baik

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

Confidence (0.80): 80% dari santri yang memiliki Kehadiran ≥ 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* dengan hasil analisis menunjukkan bahwa kehadiran lebih dari 80% jam belajar lebih, dari 40 jam per minggu serta partisipasi diskusi aktif sangat berhubungan positif dengan pencapaian nilai akademik yang lebih baik, hasil ini dibuktikan dengan nilai *support* dan *confidence* signifikan. *Algoritma Apriori* efektif dalam menemukan aturan *asosiasi* seperti 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 fokus pada kehadiran, diskusi aktif, dan jam belajar mandiri guna meningkatkan prestasi 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/infeb.v4i4.167>
- Hasanul Huda, M., & Syafrullah, M. (n.d.). *PENERAPAN METODE ASSOCIATION RULES DENGAN ALGORITME 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.
- 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>
- 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)*.
- 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. <https://tunasbangsa.ac.id/ejurnal/index.php/jurasik>
- 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/>