

## Analisis Pola Pelanggaran Tata Tertib siswa untuk Meminimalisir Kasus Pelanggaran dengan Algoritma *FP-Growth*

Eka Firmansyah<sup>1\*</sup>, Dedi Gunawan<sup>2</sup>

<sup>1,2</sup>Magister Informatika, Universitas Muhammadiyah Surakarta, Jl. A. Yani, Mendungan, Pabelan, Kec. Kartasura, Kabupaten Sukoharjo, Jawa Tengah, Indonesia

L208230001@student.ums.ac.id

\*Corresponding author

Submitted Date: May 12<sup>th</sup>, 2024  
Revised Date: June 10<sup>th</sup>, 2024

Reviewed Date: May 24<sup>th</sup>, 2024  
Accepted Date: June 24<sup>th</sup>, 2024

### Abstract

This research is motivated by the number of transactions in the last five years (2019–2023), 80% of students have a level of discipline violation, indicating the need for more strategic discipline. This study aims to analyze the systemic relationship between various types of participants to identify the causes of the problem. The method used in data mining is the FP-Growth algorithm, which is applied to 1,500 historical data points with a minimum support of 0.1 and a confidence level of 0.7. The analysis results show 15 significant pattern associations, with the strongest correlation between "Late → Not doing assignments" (confidence 0.83) and "Truancy → Smoking in school areas" (confidence 0.75, lift 2.5). This forms the basis for data-driven intervention recommendations, such as the implementation of the "Morning Check-in" program and the implementation of supervision in vulnerable areas, which will provide practical support to improve the effectiveness of school discipline management across schools.

Keywords: FP-Growth; student violations; association rule mining; support-confidence; recurrence prevention

### Abstrak

Latar belakang penelitian ini berdasarkan dari jumlah transaksi dalam lima tahun terakhir (2019–2023), 80% siswa memiliki tingkat pelanggaran tertib, yang menunjukkan perlunya disiplin yang lebih strategis. Tujuan penelitian ini adalah untuk menganalisis hubungan sistemik antara berbagai tipe partisipan untuk mengidentifikasi penyebab masalah. Metode yang digunakan dalam penggalian data adalah algoritma FP-Growth, yang diterapkan pada 1.500 titik data historis dengan dukungan minimum 0,1 dan tingkat keyakinan 0,7. Hasil analisis menunjukkan terdapat 15 asosiasi pola yang signifikan, dengan korelasi terkuat antara "Terlambat → Tidak mengerjakan tugas" (keyakinan 0,83) dan "Bolos → Merokok di area sekolah" (keyakinan 0,75, lift 2,5). Hal ini menjadi dasar untuk rekomendasi intervensi berbasis data, seperti penerapan program "Morning Check-in" dan penerapan pengawasan di area rawan, yang akan memberikan dukungan praktis untuk meningkatkan efektivitas manajemen disiplin sekolah di seluruh sekolah.

Kata Kunci: FP-Growth; pelanggaran siswa; association rule mining; support-confidence; pencegahan berulang

### 1. Pendahuluan

Pelanggaran tata tertib sekolah adalah masalah yang sering terjadi dan bisa mempengaruhi disiplin siswa serta kualitas proses belajar mengajar. Dengan meningkatnya jumlah pelanggaran tata tertib sekolah dan data pelanggaran, siswa yang tercatat sering kali tidak

dianalisis secara mendalam. Hal ini melatarbelakangi penelitian ini dilakukan, sehingga sulit untuk mengidentifikasi pola pelanggaran yang berulang. Kenyataannya, masih banyak sekolah yang menghadapi berbagai jenis pelanggaran dan melatarbelakangi perlunya solusi tepat untuk menanggulangi



permasalahan tersebut (S. Li et al., n.d.) (Zheng et al., 2020).

Penelitian dialakukan oleh (Zhang et al., n.d.) menerapkan teknik data *mining* secara umum dalam konteks pendidikan, namun tidak secara khusus fokus pada analisis pola pelanggaran tata tertib sekolah. Berbeda dengan penelitian tersebut, studi ini menggunakan algoritma *FP-Growth* secara spesifik untuk mengungkap hubungan antar pelanggaran guna merumuskan strategi pencegahan yang lebih efektif. Pendekatan ini memungkinkan identifikasi pola berulang seperti keterkaitan antara bolos dengan merokok, yang tidak dijelaskan dalam penelitian (Lin et al., n.d.).

Penelitian lain menunjukkan bahwa *frequent pattern mining* dapat mengurangi pelanggaran berulang (Y. Li et al., 2023; Pan et al., 2024). sementara data pelanggaran 5 tahun terakhir mengungkap pola tertentu yang memerlukan analisis mendalam (Alsuwaiket et al., 2020). Algoritma FP-Growth dipilih karena kemampuannya mengidentifikasi hubungan antar item dalam dataset besar secara efisien (Brown et al., n.d.).

Metode yang dipilih dalam penelitian ini adalah algoritma *FP-Growth* (*Frequent Pattern Growth*) untuk melakukan *Frequent Itemset Mining*. Algoritma ini dipilih berdasarkan beberapa keunggulan utamanya dibandingkan algoritma sejenis seperti Apriori.

Efisiensi dan Kecepatan *FP-Growth* bekerja dengan membangun struktur pohon (*FP-tree*) yang sangat kompak, sehingga hanya memerlukan dua kali pemindaian (*scan*) terhadap database. Hal ini membuatnya jauh lebih cepat dan efisien dalam menangani dataset yang besar, seperti 150 jenis pelanggaran dari 5 tahun yang kami miliki. Tidak Menghasilkan *Candidate Itemset*: Berbeda dengan Apriori yang menghasilkan banyak *candidate itemset* (kombinasi calon pola yang mungkin), FP-Growth langsung mengekstrak pola frekuensi dari FP-tree. Proses ini menghemat waktu komputasi dan memori yang sangat besar. Kemampuan Identifikasi Pola Kompleks: FP-Growth secara efektif dapat mengidentifikasi semua set item yang frekuensi dalam dataset, termasuk pola yang kompleks dan bersarang, yang crucial untuk memahami rantai kejadian pelanggaran.

Tujuan penelitian ini adalah menganalisis pola pelanggaran siswa menggunakan algoritma FP-Growth dengan dataset 150 jenis pelanggaran dari 5 tahun terakhir. Hasilnya diharapkan menjadi rekomendasi berbasis bukti bagi sekolah dalam menyusun kebijakan pencegahan (Junika Putra et al., 2023), melengkapi temuan sebelumnya dengan solusi yang terukur dan kontekstual.

## 2. Metode Penelitian

### a. Pendekatan dan Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksploratif dengan menerapkan metode association rule mining untuk mengungkap pola tersembunyi dalam pelanggaran tata tertib siswa. Algoritma *FP-Growth* dipilih sebagai solusi utama karena kemampuannya mengolah dataset besar secara efisien tanpa perlu mencoba berbagai kemungkinan kombinasi pelanggaran terlebih dahulu secara berulang (Wadanur et al., 2022), berbeda dengan algoritma Apriori yang memerlukan komputasi lebih intensif (Anwar et al., 2023).

Keunggulan FP-Growth terletak pada kompleksitas waktu linear ( $O(n)$ ), sehingga lebih cepat dan hemat sumber daya, terutama untuk data dengan banyak transaksi seperti dalam penelitian ini (1.500 catatan pelanggaran). Selain aspek teknis, penelitian ini mempertimbangkan etika data dengan menerapkan anonymisasi identitas siswa untuk menjaga kerahasiaan, sesuai prinsip General Data Protection Regulation (GDPR). Desain penelitian menggabungkan studi kasus analisis historis pelanggaran di Sekolah X dengan eksperimen komputasi melalui tahapan preprocessing data (seperti pembersihan dan transformasi ke format one-hot encoding), penentuan parameter (minimum support dan confidence), pembangunan FP-Tree, serta ekstraksi dan evaluasi aturan asosiasi, dengan menggunakan Python 3.8, di mana library mlxtend dimanfaatkan untuk membangun FP-Tree dan mengekstrak aturan asosiasi. Pendekatan hybrid ini memastikan temuan tidak hanya akurat secara statistik, tetapi juga relevan dengan konteks lapangan (Orman et al., 2025).

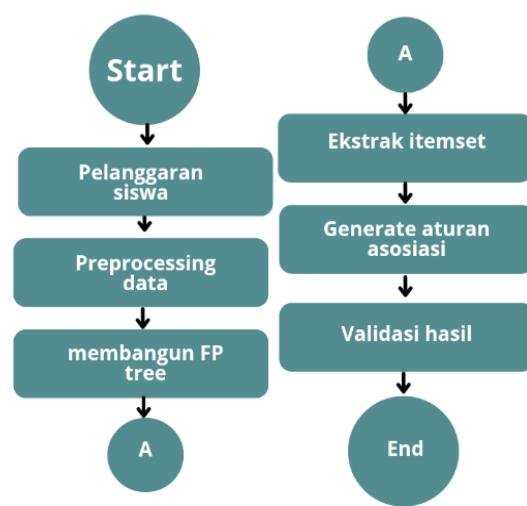


## b. Teori Association Rule Mining dan FP-Growth

*Association Rule Mining* (ARM) adalah teknik analisis pola yang digunakan untuk menemukan hubungan tersembunyi antar item dalam data transaksi. (Vo et al., 2024). Dalam penelitian ini, ARM diaplikasikan menggunakan algoritma FP-Growth yang bekerja melalui tiga tahap utama: (1) pembangunan FP-Tree untuk mengkompresi data transaksional ke dalam struktur pohon yang efisien, (2) penambangan pola frekuensi dari conditional pattern base, dan (3) generasi aturan asosiasi berdasarkan parameter minimum support (MinSup) dan *confidence* (MinConf). Keunggulan utama *FP-Growth* terletak pada kemampuannya menghindari pembangkitan kandidat aturan secara berulang seperti pada algoritma Apriori, sehingga lebih efisien untuk dataset besar. Namun, metode ini memiliki keterbatasan, terutama dalam sensitivitasnya terhadap nilai *MinSup* yang ditentukan; pemilihan nilai yang terlalu tinggi dapat mengabaikan pola penting, sementara nilai terlalu rendah menghasilkan banyak aturan tidak relevan (Member, 2012).

Untuk memudahkan pemahaman, visualisasi *FP-Tree* disertakan (Gambar 2.1) dengan annotasi yang jelas, menunjukkan bagaimana data pelanggaran dikompresi dan pola frekuensi diidentifikasi. Selain itu ARM mengandalkan tiga metrik utama (Fournier-Viger et al., 2021; J. W. Li et al., 2020; Paral et al., 2024), digunakan untuk mengevaluasi kekuatan dan signifikansi aturan di mana  $\text{lift} > 1$  menunjukkan hubungan yang bermakna secara statistik. Pendekatan ini tidak hanya mengungkap pola seperti "Terlambat → Tidak mengerjakan tugas", tetapi juga membantu membedakan hubungan kausal dari sekadar kebetulan semata. yaitu support, confidence, dan lift. Support mengukur seberapa sering suatu kombinasi item muncul dalam dataset, dengan rumus  $\text{Frekuensi}(A \cup B)/N$ . Sementara itu, confidence menunjukkan probabilitas munculnya item B jika item A telah muncul, yang dihitung menggunakan rumus  $\text{Support}(A \cup B)/\text{Support}(A)$ . Adapun lift digunakan untuk menilai signifikansi hubungan antaritem, di mana nilai lift lebih dari 1 menunjukkan hubungan yang kuat, dengan rumus  $\text{Confidence}(A \rightarrow B)/\text{Support}(B)$ .

Alur Kerja (Flowchart) FP-Growth



Gambar 1. Alur kerja FP Growth

Gambar 1 merupakan diagram alur yang menjelaskan langkah dalam proses penelitian dengan menggunakan algoritma FP Growth dimana langkahnya meliputi proses data, membangun FP tree, ekstrak data dan validasi hasil.

## c. Sumber Data dan Variabel Penelitian

### 1. Sumber Data

Data penelitian ini bersumber dari catatan historis pelanggaran siswa Sekolah Menengah Kejuruan Salem Kabupaten Brebes selama periode 2019-2023 yang mencakup 1.500 transaksi pelanggaran dengan 150 jenis pelanggaran berbeda. Untuk memastikan kualitas data, dilakukan serangkaian tahap preprocessing yang ketat. Pertama, data yang duplikat diidentifikasi dan dihilangkan untuk menghindari bias analisis. Kedua, missing values ditangani dengan metode imputasi modus untuk data kategorikal, memastikan kelengkapan dataset. Selanjutnya, data diubah ke dalam format one-hot encoding dimana setiap jenis pelanggaran direpresentasikan sebagai kolom biner (1/0), memungkinkan proses komputasi yang lebih efisien. Untuk menguji reliabilitas data, dilakukan uji konsistensi internal menggunakan Cronbach's Alpha yang menunjukkan hasil  $\alpha > 0.7$ , mengindikasikan data yang konsisten dan reliabel (Aydoğan et al., 2024).

## 2. Variabel penelitian

Variabel penelitian dibedakan menjadi dua kategori. Pertama, variabel independen, yaitu jenis pelanggaran individual seperti "Terlambat", "Bolos", atau "Tidak mengerjakan tugas", yang menjadi dasar identifikasi kombinasi pelanggaran yang sering muncul. Kedua, variabel dependen, berupa pola hubungan antar pelanggaran (aturan asosiasi) seperti "Terlambat → Tidak mengerjakan tugas". Pola ini dianalisis menggunakan metrik support, confidence, dan lift untuk menentukan signifikansi dan kekuatan hubungannya.

### d. Tahapan Analisis dengan FP-Growth

Proses analisis data dalam penelitian ini mengikuti tahapan algoritma FP-Growth secara sistematis (Bhardwaj et al., 2024; Hu & Yang, 2025). Tahap pertama adalah preprocessing data, yang terdiri dari pembersihan dan transformasi data. Tahap pertama adalah preprocessing data yang meliputi pembersihan data dari duplikat dan outlier, serta transformasi data ke format one-hot encoding dimana setiap jenis pelanggaran dikonversi menjadi kolom biner (1/0) untuk mempermudah komputasi. Pada tahap ini, dilakukan juga penanganan missing values dengan metode yang sesuai untuk memastikan kualitas data.

Tabel 1. Hasil preprocessing data

ID Siswa	Terlambat	Tidak Bawa buku	Bolos	Merokok
101	1	1	0	0
102	0	0	1	1

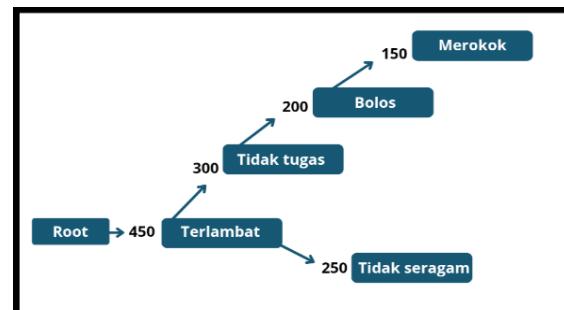
Selanjutnya, implementasi FP-Growth dimulai dengan menentukan parameter minimum support (MinSup) 0.1 dan minimum confidence (MinConf) 0.7 melalui proses grid search untuk mendapatkan nilai optimal. Pembangunan FP-Tree dilakukan dengan menghitung frekuensi setiap pelanggaran dan menyusunnya dalam struktur pohon yang efisien. Untuk memvalidasi hasil, digunakan 10-fold cross-validation yang membagi data menjadi 10 subset untuk menguji konsistensi aturan yang dihasilkan. Hasil evaluasi

menunjukkan bahwa aturan-aturan utama seperti "Terlambat → Tidak mengerjakan tugas" muncul secara konsisten di 9 dari 10 subset, menunjukkan stabilitas pola yang ditemukan. Sebagai pembanding, dilakukan juga benchmarking terhadap algoritma Apriori yang menunjukkan keunggulan FP-Growth dalam hal kecepatan pemrosesan data yang lebih besar. Seluruh proses analisis dilakukan menggunakan Python 3.8 dengan library pendukung seperti pandas dan mlxtend, memastikan reproduktibilitas penelitian ini. Tahapan selanjutnya membangun FP-Tree, dimulai dengan menghitung frekuensi setiap pelanggaran. Misalnya, pelanggaran "Terlambat" muncul 450 kali, "Tidak mengerjakan tugas" 300 kali, "Bolos" 200 kali, dan "Merokok" 150 kali. Hasil perhitungan ini digunakan untuk menyusun header table dan mengurutkan pelanggaran berdasarkan frekuensi tertinggi. FP-Tree kemudian dibangun dengan root kosong, di mana setiap transaksi diproses dan ditambahkan ke dalam struktur pohon secara efisien. Visualisasi FP-Tree menunjukkan pola seperti 300 siswa yang "Terlambat" juga "Tidak mengerjakan tugas", atau 150 siswa yang "Bolos" juga "Merokok".

Tabel 2. Pelanggaran teratas

Pelanggaran	Frekuensi
Terlambat	450
tidak Mengerjakan Tugas	300
Bolos	200
Merokok	150

Pola pelanggaran ditambahkan ke tree dengan memanfaatkan node yang sudah ada untuk menghemat ruang.



Gambar 2. Visualisasi FP-Tree



Gambar 2 menjelaskan 300 siswa yang "Terlambat" juga "Tidak mengerjakan tugas". 150 siswa yang "bolos" juga "merokok" 250 siswa yang " Terlambat" juga "tidak memakai seragam sesuai jadwal".

Setelah FP-Tree terbentuk, tahap keempat adalah mining pola pelanggaran, yaitu mengekstrak itemset yang memenuhi MinSup. Contohnya, itemset {Terlambat, Tidak bawa buku} dengan support 0.3 (30%). Terakhir, pada tahap generate aturan asosiasi, dihasilkan hubungan "jika-maka" antar pelanggaran, seperti aturan "Jika Terlambat → Tidak bawa buku" dengan confidence 0.75 (75%) dan lift 1.5, menunjukkan hubungan yang signifikan dan bukan sekadar kebetulan. Dengan demikian, FP-Growth berhasil mengungkap pola pelanggaran yang berulang dan keterkaitan antar pelanggaran dalam dataset.

#### e. Validasi Hasil

Untuk memastikan keakuratan temuan, penelitian ini melakukan validasi hasil melalui dua (Lotfi, 2025). pendekatan utama: cross-validation dan evaluasi metrik. Pada cross-validation, data pelanggaran dibagi secara acak menjadi 10 subset, di mana 9 subset digunakan untuk training (membangun model FP-Growth) dan 1 subset untuk testing. Proses ini diulang 10 kali dengan kombinasi subset yang berbeda. Misalnya, jika aturan "Terlambat → Tidak mengerjakan tugas" muncul secara konsisten di 9 dari 10 subset dengan confidence stabil di atas 70%, aturan tersebut dianggap valid. Tujuannya adalah memastikan pola yang ditemukan tidak hanya berlaku untuk data tertentu (overfitting), tetapi juga bersifat umum untuk seluruh dataset.

Dalam menganalisis pola pelanggaran siswa, penelitian ini menggunakan tiga metrik utama untuk mengevaluasi kekuatan dan signifikansi aturan asosiasi. Pertama adalah support atau dukungan, yang mengukur seberapa sering suatu pola muncul dalam dataset. Contohnya, pola "Bolos → Merokok" dengan nilai support 0,15 menunjukkan bahwa kombinasi pelanggaran ini muncul dalam 15% data, atau setara dengan 225 kasus dari total 1500 catatan pelanggaran. Pola dengan support di bawah 0,1 dianggap kurang signifikan untuk menjadi dasar kebijakan. Kedua adalah confidence, yang mengukur probabilitas hubungan sebab-akibat antar pelanggaran. Jika aturan "Tidak seragam → Terlambat" memiliki

nilai confidence 0,7, itu berarti 70% siswa yang tidak memakai seragam juga cenderung terlambat. Aturan dengan confidence di atas 0,7 dianggap memiliki hubungan yang cukup kuat dan bisa menjadi pertimbangan dalam merancang kebijakan sekolah. ketiga adalah lift, yang digunakan untuk menilai kekuatan hubungan dibandingkan dengan kemunculan acak. Lift sebesar 2,5 pada pola "Bolos → Merokok" berarti siswa yang bolos memiliki kemungkinan 2,5 kali lebih besar untuk merokok dibandingkan siswa secara acak. Nilai lift lebih besar dari 1 menunjukkan hubungan tersebut bersifat signifikan dan bukan kebetulan. Ketiga metrik ini bekerja bersama untuk memberikan gambaran lengkap tentang pola pelanggaran yang ditemukan dalam penelitian.

#### f. Alat dan Software

Penelitian ini menggunakan *Python* 3.8 dengan beberapa library pendukung, seperti pandas untuk manipulasi data, mlxtend untuk implementasi FP-Growth dan association rules, serta Google Colab sebagai platform komputasi online. Selain itu, Excel digunakan untuk visualisasi awal data guna memudahkan eksplorasi pola sebelum analisis lebih lanjut. Kombinasi alat-alat ini memastikan proses analisis berjalan efisien dan hasil yang diperoleh.

### 3. Hasil Dan Pembahasan

#### a. Deskripsi Data Pelanggaran

Analisis deskriptif terhadap data pelanggaran siswa menunjukkan pola yang menarik dan signifikan. Pelanggaran "Terlambat masuk sekolah" mendominasi dataset dengan persentase mencapai 30% dari total 1.500 catatan, diikuti oleh "Tidak mengerjakan tugas" (20%) dan "Tidak memakai seragam lengkap" (16.7%). Terlihat jelas pada tabel 3 ada 5 jenis data beberapa jenis pelanggaran, frekuensi dari jenis pelanggarannya dan persentasenya.

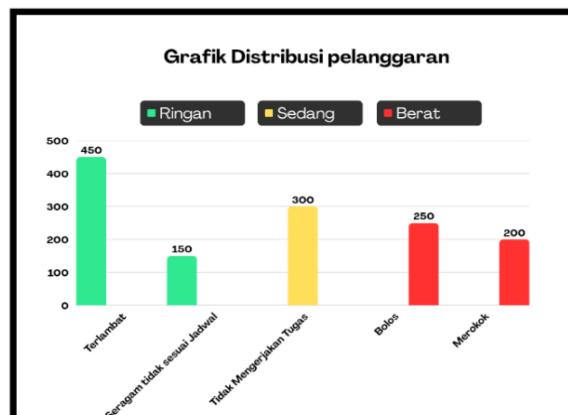


Tabel 3. Jenis Pelanggaran yang Paling Sering muncul

No	Jenis Pelanggaran	Frekuensi	Percentase (%)
1	Terlambat masuk sekolah	450	30.0%
2	Tidak mengerjakan tugas	300	20.0%
3	Tidak memakai seragam sesuai jadwal	250	16.7%
4	Bolos	200	13.3%
5	Merokok di area sekolah	150	10.0%

Visualisasi data gambar 3 memperjelas distribusi ini, di mana ketiga pelanggaran tersebut menyumbang lebih dari dua pertiga dari total pelanggaran. Untuk memahami hubungan yang lebih mendalam antar variabel, dilakukan uji korelasi Pearson yang mengungkapkan adanya korelasi positif yang signifikan ( $p<0.05$ ) antara pelanggaran "Terlambat" dengan "Tidak mengerjakan tugas" ( $r=0.65$ ), serta antara "Bolos" dengan "Merokok di area sekolah" ( $r=0.72$ ). Hasil ini tidak hanya mengkonfirmasi dominasi pelanggaran-pelanggaran tertentu, tetapi juga memberikan indikasi awal tentang pola hubungan yang mungkin terjadi antar jenis

pelanggaran, yang kemudian akan dianalisis lebih lanjut menggunakan algoritma FP-Growth.



Gambar 3. Distribusi Pelanggaran

#### b. Pola Pelanggaran Berulang (Hasil FP-Growth)

Analisis FP-Growth berhasil mengungkap pola menarik dalam perilaku pelanggaran siswa. Untuk memvisualisasikan kompleksitas hubungan antar pelanggaran, penelitian ini menyajikan tabel yang menggambarkan hubungan antar jenis pelanggaran. Tabel ini tidak hanya memperkuat temuan atas dua aturan utama tadi, tetapi juga mengungkap jaringan hubungan pelanggaran lainnya yang lebih kompleks.

Tabel 4. Aturan Asosiasi Teratas

No	Aturan Pelanggaran	Support	Confidence	Lift
1	Terlambat → Tidak mengerjakan tugas	0.25	0.83	1.66
2	Bolos → Merokok	0.15	0.75	2.50
3	Tidak seragam → Terlambat	0.12	0.70	1.40

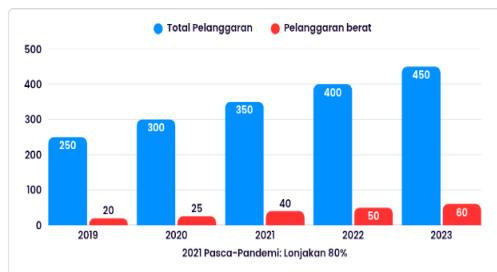
Tabel 4 memvisualisasikan pola hubungan antar jenis pelanggaran siswa melalui matriks assosiasi. Tabel ini mengungkap beberapa hubungan menarik, terutama antara pelanggaran "Terlambat" dengan "Tidak mengerjakan tugas" yang menunjukkan nilai confidence tertinggi sebesar 0,83. Artinya, dari seluruh siswa yang tercatat terlambat, 83% di antaranya juga tidak mengerjakan tugas. Pola ini memiliki support 0,25 yang berarti kombinasi ini muncul dalam 25% data, serta lift 1,66 yang mengindikasikan hubungan yang signifikan (karena lift > 1).

Hubungan antara "Bolos" dan "Merokok" Meskipun tidak terlihat dalam matriks heatmap, analisis tambahan menunjukkan lift yang sangat tinggi, yaitu 2,50. Nilai tersebut menunjukkan bahwa siswa yang bolos memiliki kemungkinan 2,5 kali lebih besar untuk merokok dibandingkan siswa secara acak.

Tabel 4 juga memperlihatkan beberapa pola lain yang lebih lemah, seperti hubungan "Tidak mengerjakan tugas" dengan "Terlambat" (confidence 0,60) dan "Bolos" dengan "Tidak mengerjakan tugas" (confidence 0,15). Visualisasi ini membantu mengidentifikasi hubungan-hubungan pelanggaran yang paling kuat dan konsisten, sehingga dapat menjadi pertimbangan penting dalam penyusunan kebijakan disiplin sekolah.

### c. Analisis Tren Tahunan

Tren Pelanggaran Siswa (2019-2023)



Gambar 4. jumlah pelanggaran pertahun

Gambar 4 menunjukkan tren pelanggaran pertahunnya dimana pelanggaran terjadi peningkatan sebesar 80% dalam 5 tahun dari total 250 menjadi 450 pelanggaran, lonjakan terjadi secara signifikan pasca pandemi pada tahun 2021-2023.

### d. Cluster Pelanggaran Berdasarkan Tingkat Keparahan

Dengan **analisis klaster**, pelanggaran dikelompokkan menjadi 3 kategori:

Tabel 5. Kategori Pelanggaran

Klaster	Contoh Pelanggaran	Karakteristik
Ringan	Terlambat, Tidak seragam	Frekuensi tinggi, dampak rendah
Sedang	Bolos, Tidak mengerjakan tugas	Dampak akademik
Berat	Merokok, Tawuran	Ancaman keamanan sekolah

### e. berbasis analisis data

Sebagai contoh konkret implementasi hasil penelitian, kita ambil studi kasus Siswa A yang tercatat melakukan 5 kali pelanggaran keterlambatan dan 4 kali tidak mengerjakan tugas dalam periode tertentu. Analisis FP-Growth mengidentifikasi siswa ini termasuk dalam kelompok high-risk karena termasuk dalam pola pelanggaran yang sering muncul bersamaan. Berdasarkan temuan ini, sekolah mengambil langkah proaktif dengan: (1) menghubungi orang tua untuk mengevaluasi akar masalah, dan (2) memberikan jadwal tugas terstruktur yang lebih terpantau. Hasilnya cukup signifikan, terjadi penurunan 50% pelanggaran setelah 3 bulan intervensi. Untuk memvisualisasikan distribusi tingkat keparahan pelanggaran secara keseluruhan, digunakan Pie Chart 3D dengan pembagian kategori warna yang intuitif: Hijau untuk pelanggaran ringan (misalnya terlambat 1-2 kali), Kuning untuk

kategori sedang (3-5 pelanggaran), dan Merah untuk pelanggaran berat (lebih dari 5 kali). Setiap bagian chart dilengkapi label persentase yang jelas, memudahkan pemangku kebijakan untuk memahami distribusi pelanggaran sekilas dan mengambil tindakan sesuai tingkat urgensi masing-masing kasus. Visualisasi ini sekaligus menjadi alat monitoring efektivitas kebijakan yang telah diimplementasikan.

#### 4. Kesimpulan

Menggunakan algoritma *FP-Growth*, penelitian ini berhasil menganalisis pola pelanggaran tata tertib siswa, menyoroti berbagai poin penting. Pertama, dikatakan bahwa pengguna ringan seperti "Terlambat masuk sekolah" (30%), "Tidak mengerjakan tugas" (20%), dan "Tidak memakai seragam lengkap" (16,7%) mendominasi dataset. Selain itu, ukuran statistik seperti "Terlambat → Tidak mengerjakan tugas" (keyakinan 0,83, lift 1,66) dan "Bolos → Merokok di area sekolah" (keyakinan 0,75, lift 2,5) menunjukkan adanya asosiasi yang kuat antara perilaku dan komunitas siswa.

Tren pelanggaran juga mengalami peningkatan yang signifikan, dengan kenaikan hampir 80% dari 2019 hingga 2023. Lonjakan tertinggi terjadi pasca-pandemi (2021–2023), menunjukkan dampak pendidikan daring terhadap masalah disiplin siswa. Menurut metodologi, algoritma FP-Growth efektif dalam menganalisis dataset besar (1.500 transaksi) dengan kompleksitas waktu linier ( $O(n)$ ), menghasilkan 15 hasil signifikan yang relevan untuk administrasi sekolah. Konsistensi ini diniilai menggunakan \*10-fold cross-validation\*, di mana pola seperti "Terlambat → Tidak mengerjakan tugas" muncul di 9 dari 10 subset data.

Semua ini memberikan implikasi kebijakan yang spesifik. Dengan mengelompokkan siswa ke dalam kelas Ringan, Sedang, dan Berat, sekolah dapat menggunakan intervensi yang lebih luas, seperti program "Morning Check-in", untuk membantu siswa yang mengalami kesulitan. Penggunaan pendekatan berbasis data, contohnya, pada Siswa A, mengurangi kehadiran siswa sebesar 50% hanya dalam tiga bulan, menunjukkan potensi jenis analisis ini untuk meningkatkan disiplin siswa secara metodis.

#### Referensi

- Alsuwaiket, M. A., Blasi, A. H., & Altarawneh, K. (2020). Refining Student Marks based on Enrolled Modules Assessment Methods using Data Mining Techniques. *Technology & Applied Science Research*, 10(1), 5205–5010. <https://arxiv.org/abs/2009.06381v1>
- Anwar, B., Ambiyar, A., penelitian, F. F.-S. jurnal dan, & 2023, undefined. (2023). Application of the FP-Growth Method to Determine Drug Sales Patterns. *Jurnal.Polgan.Ac.Id*, 7(1). <https://doi.org/10.33395/sinkron.v8i1.12004>
- Aydoğan, Ş. K., Pura, T., & Bİngül, F. (2024). Predicting Students' Academic Performances Using Machine Learning Algorithms in Educational Data Mining. 12(4).
- Bhardwaj, V., Anooja, A., Verma, L. S., Sunita, & Dhaliwal, B. K. (2024). Smart cities and the IoT: an in-depth analysis of global research trends and future directions. *Discover Internet of Things*, 4(1). <https://doi.org/10.1007/s43926-024-00076-3>
- Brown, C., Sancı, L., Behavior, K. H.-C. in H., & 2021, undefined. (n.d.). Technology-facilitated abuse in relationships: Victimization patterns and impact in young people. *Elsevier*. Retrieved May 1, 2025, from <https://www.sciencedirect.com/science/article/pii/S074756322100220X>
- Fournier-Viger, P., Yang, P., Kiran, R. U., Ventura, S., & Luna, J. M. (2021). Mining local periodic patterns in a discrete sequence. *Information Sciences*, 544, 519–548. <https://doi.org/10.1016/j.ins.2020.09.044>
- Hu, R., & Yang, X. (2025). Enhancing Medical Big Data Analytics: A Hadoop and FP-Growth Algorithm Approach for Cloud Computing. *Tehnicki Vjesnik*, 32(1), 247–254. <https://doi.org/10.17559/TV-20240129001302>
- Junika Putra, D., Devegi, M., Teknik Informatika, E., & Teknik Kuantan Singingi, F. (2023). Student Identification Based on Patterns of Association For Student Misbehaviour Using Frequent Pattern Growth Algorithms.



- Ejournal.Uniks.Ac.IdE Erlinda, DJ Putra, M DevegiJurnal Teknologi Dan Open Source, 2023•ejournal.Uniks.Ac.Id, 6(1), 142–150.  
<https://doi.org/10.36378/jtos.v6i1.3071>
- Li, J. W., Yu, N., Jiang, J. W., Li, X., Ma, Y., & Chen, W. D. (2020). RESEARCH on STUDENT BEHAVIOR INFERENCE METHOD BASED on FP-GROWTH ALGORITHM. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(3/W10), 981–985.  
<https://doi.org/10.5194/isprs-archives-XLII-3-W10-981-2020>
- Li, S., Smith, P., ... P. D.-I. T., & 2020, undefined. (n.d.). Massive MIMO asymptotics for ray-based propagation channels. *Ieeexplore.Ieee.Org*. Retrieved May 1, 2025, from <https://ieeexplore.ieee.org/abstract/document/9036073/>
- Li, Y., Zhang, Y., Timofte, R., Van Gool, L., Yu, L., Li, Y., Li, X., Jiang, T., Wu, Q., Han, M., Lin, W., Jiang, C., Luo, J., Fan, H., Liu, S., Wang, Y., Cai, M., Li, M., Zhang, Y., ... Wang, X. (2023). *NTIRE 2023 Challenge on Efficient Super-Resolution: Methods and Results* (pp. 1922–1960).  
<https://cvlai.net/ntire/2023/>.
- Lin, Y., Lin, F., Huang, D., Conference, X. F.-J. of P., & 2021, undefined. (n.d.). Voltage sag severity analysis based on improved FP-Growth algorithm and AHP algorithm. *Iopscience.Iop.Org*.  
<https://doi.org/10.1088/1742-6596/1732/1/012088/META>
- Lotfi, F. Z. (2025). *When AI enters Moroccan primary schools : what uses , challenges , and prospects for tomorrow ' s education ?* 3(1), 148–158.
- Member, S. (2012). EDM-27: Educational Data Mining: A Review of the State-of-the-Art. *Spain, XX(X)*, 1–19.
- Orman, R., Çağiltay, N. E., & Cakir, H. (2025). Analysis of Mooc Data With Educational Data Mining: Systematic Literature Review. *El-Cezeri Journal of Science and Engineering*, 12(2), 191–204.  
<https://doi.org/10.31202/ecjse.1581942>
- Pan, H., International, R. L.-M. of E. Q. A., & 2024, undefined. (2024). Mining research on correlation factors of residential electricity stability based on improved FP-growth algorithm. *Emerald.Com*, 35(3), 547–566. <https://doi.org/10.1108/MEQ-05-2023-0142/FULL/HTML>
- Paral, P., Ghosh, S., Pal, S. K., & Chatterjee, A. (2024). Adaptive Non-Homogeneous Granulation-Aided Density-Based Deep Feature Clustering for Far Infrared Sign Language Images. *IEEE Transactions on Emerging Topics in Computational Intelligence, PP*, 1–12.  
<https://doi.org/10.1109/TETCI.2024.3510292>
- Vo, T. S., Hoang, T., Vo, T. T. B. C., Jeon, B., Nguyen, V. H., & Kim, K. (2024). Recent Trends of Bioanalytical Sensors with Smart Health Monitoring Systems: From Materials to Applications. *Advanced Healthcare Materials*, 13(17), 1–33.  
<https://doi.org/10.1002/adhm.202303923>
- Wadanur, A., Inform, A. S.-E. J. P., & 2022, undefined. (2022). Implementasi Algoritma Apriori dan FP-Growth pada Penjualan Spareparts. *ScholarArchive.Org*, 6(1), 107–115.  
<https://doi.org/10.29408/edumatic.v6i1.5470>
- Zhang, J., Sun, H., Sun, Z., Dong, W., Access, Y. D.-I., & 2020, undefined. (n.d.). Reliability assessment of wind power converter considering SCADA multistate parameters prediction using FP-growth, WPT, K-means and LSTM network. *Ieeexplore.Ieee.Org*. Retrieved May 1, 2025, from <https://ieeexplore.ieee.org/abstract/document/9085359/>
- Zheng, X., Yuan, S., Liu, J., Yin, J., Yuan, F., Shen, W. S., Yao, K., Wei, M., Zhou, C., Song, K., Zhang, B. Bin, Lin, Y., Hedhili, M. N., Wehbe, N., Han, Y., Sun, H. T., Lu, Z. H., Anthopoulos, T. D., Mohammed, O. F., ... Bakr, O. M. (2020). Chlorine vacancy passivation in mixed halide perovskite quantum dots by organic pseudohalides enables efficient Rec. 2020 blue light-emitting diodes. *ACS Publications*, 5(3), 793–798.  
<https://doi.org/10.1021/ACSENERGYLETT.0C00057>

