

Implementasi *Data Mining* Dalam Menentukan Korelasi Antara Pelanggaran Sikap Siswa Di Lingkungan Sekolah Lanjutan Advent Purwodadi Menggunakan Algoritma *Association Rule*

Gomgom Yosua Balutaro Sihotang
PT. Kosada Group Indonesia
e-mail: gomgombalutaro@gmail.com

Abstrak

Saat ini banyak sekali kenakalan yang menjadi pelanggaran di lingkungan sekolah, dan yang diketahui oleh anak itu sendiri bahwa jika perbuatannya itu sempat diketahui oleh petugas hukum ia bisa dikenai hukuman. Kenakalan yang mengakibatkan sebuah pelanggaran biasanya dimulai dengan sebuah perilaku yang menyimpang di antara siswa. Pada penelitian ini memungkinkan peneliti untuk mencari nilai korelasi antara data pelanggaran di antara siswa Sekolah Lanjutan Advent Purwodadi, dengan cara menggunakan metode *Association Rule* dan juga menggunakan *FP-Growth* sebagai algoritmanya. Adapun dengan menggabungkan data pelanggaran berdasarkan atribut yakni jenis-jenis pelanggaran yang dilakukan. Data yang digunakan adalah data pelanggaran siswa dari tahun 2016-2020 di Sekolah Lanjutan Advent Purwodadi. Hasil pengujian menunjukkan, bahwa implementasi korelasi pelanggaran dengan menggunakan metode *Association Rule* mendapatkan 27 aturan asosiasi dengan nilai *minimum support* adalah 0.08 atau 8% dan nilai *minimum confidence* adalah 0.7 atau 70%.

Kata Kunci: *Data mining, Association Rule, Pelanggaran, FP-Growth*

Data Mining Implementation in Determining the Correlation Between Student Attitude Violations in the Purwodadi Adventist High School Environment Using Algorithm Association Rules

Abstract

Currently, there are a lot of delinquency that has become a violation in the school environment, and it is known by the child himself that if his actions were discovered by the legal officer he could be subject to punishment. Delinquency that results in an offense usually begins with a deviant behavior among students. This study allows researchers to find the correlation value between data violations among students of the Purwodadi Adventist High School, by using the Association Rule method and also using FP-Growth as the algorithm. As for combining violation data based on attributes, namely the types of violations committed. The data used is data on student violations from 2016-2020 at the Purwodadi Adventist High School. The results show that the implementation of the correlation violation using the Association Rule method gets 27 association rules with the minimum value of support being 0.08 or 8% and the minimum confidence value is 0.7 or 70%.

Keywords: *Data mining, Association Rule, violations committed, FP-Growth*

1. Pendahuluan

“Saat ini banyak sekali kenakalan yang menjadi pelanggaran di lingkungan sekolah, dan yang diketahui oleh anak itu sendiri bahwa jika perbuatannya itu sempat diketahui oleh petugas hukum ia bisa dikenai hukuman” [1]. Kenakalan yang mengakibatkan sebuah pelanggaran biasanya dimulai dengan sebuah perilaku yang menyimpang di antara siswa. Siswa dapat dianggap sebagai seseorang “subjek didik” di mana nilai kemanusiaan sebagai individu, yakni makhluk sosial yang memiliki identitas moral, perlu dikembangkan untuk mencapai tingkatan suatu proses untuk mencapai hasil yang ideal dan kriteria kehidupan sebagai manusia diharapkan oleh bangsa dan negara [2].

Menurut Willis (2010) Pada masa remaja ini tidak jarang remaja melakukan tingkah laku yang dianggap melanggar aturan yang ada. Tindak perbuatan para remaja yang juga bertentangan dengan hukum, agama dan norma-norma masyarakat sering sekali merugikan orang lain, mengganggu ketenteraman umum dan juga merusak dirinya sendiri. Adapun remaja yang melakukan hal-hal tersebut dapat dikatakan dengan kenakalan remaja [3].

Kun dan juju menyampaikan pengertian Perilaku menyimpang (*deviance*) secara umum adalah apabila Tindakan itu tidak sesuai nilai-nilai norma sosial yang berlaku dalam masyarakat, contohnya berkelahi, mencuri, menodong dan lain-lain dan dalam skala yang lebih kecil, perilaku menyimpang juga termasuk pelanggaran terhadap kebiasaan atau kepantasan, seperti siswa yang tidak berada di sekolah pada jam-jam sekolah atau jam pelajaran [4].

Visi Sekolah Lanjutan Advent Purwodadi adalah berkarakter seperti Yesus Kristus. berprestasi dalam Ilmu pengetahuan dan teknologi serta mandiri dengan mendidik siswa-siswa yang bertanggung jawab terhadap lingkungan. Melalui penelitian ini diharapkan mampu membantu pihak sekolah mengambil tindakan dalam menyikapi pelanggaran sikap yang sering terjadi, agar sesuai dengan visi sekolah dengan cara menganalisis pelanggaran siswa/i-nya. Pendeteksian korelasi antar jenis pelanggaran sikap di sekolah disebut dengan metode *Association Rule*. Penelitian ini akan mencari hubungan atau korelasi, antara jenis pelanggaran sikap yang ada di lingkungan Sekolah Lanjutan Advent Purwodadi dengan menggunakan metode data mining selama mereka bersekolah di tempat tersebut.

Menurut Han, Jiawei data mining adalah proses mengekstraksi pola-pola yang menarik, dan pengetahuan dari data yang berjumlah besar [5]. Penerapan data mining dengan menggunakan teknik aturan asosiasi sudah beberapa kali dilakukan. Penelitian terdahulu dilakukan oleh Alfannisa, dkk mengenai “Penerapan Data Mining Untuk Analisa Pola Pembelian Konsumen Dengan Algoritma *FP-Growth* pada Data Transaksi Penjualan *Spare Part Motor*” penelitian ini menemukan hubungan beberapa itemset hingga 100% di antara pola pembelian pelanggan. Diantaranya pelanggan yang membeli *Screw Valve adjusting* dipastikan akan membeli *oil seal*, dan pembelian *oil seal*, dan juga *graphic 1* dipastikan membeli *axle* karena memiliki nilai *confidence*: 1.000 [6].

Penelitian yang dilakukan oleh Yera yang berjudul “Implementasi Data Mining Untuk Analisa Tingkat Pelanggaran Lalu Lintas Dengan Algoritma *Association Rule*”. Di mana ditemukan hubungan asosiasi tertinggi adalah pelanggaran motor di mana tidak memiliki SIM yang memiliki *confidence* hingga 100%, dan diikuti oleh pelanggaran remaja yang kurangnya memenuhi perlengkapan, dengan memiliki *confidence* 83%. Data ini diharapkan dapat membantu polisi dalam menganalisis pelanggaran yang sering melanggar dan menentukan kebijakan ke depannya [7].

Berdasarkan beberapa penelitian yang telah dilakukan, penggunaan data mining dapat dikatakan membantu perusahaan atau suatu lembaga, yang dapat digunakan dalam pengambilan sebuah keputusan [8]. Selanjutnya penulis akan mengimplementasikan penggunaan data mining dengan tujuan yang sama tapi dalam kasus yang berbeda.

Dalam penelitian ini, algoritma yang digunakan adalah Frequent *Pattern-Growth (FP-Growth)* yaitu pengembangan dari metode Apriori yang merupakan salah satu alternatif untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data dengan membangkitkan struktur data Tree atau disebut dengan Frequent Pattern Tree (*Fp-Tree*) [9].

Diharapkan dengan adanya penelitian ini Pihak sekolah mampu memprediksi korelasi antar jenis pelanggaran sikap siswa, sehingga dapat mencegah siswa untuk melakukan pelanggaran lainnya yang bisa jauh lebih buruk baik pembangunan karakter siswa tersebut.

2. Landasan Teori

Untuk menambah pemahaman dan memperkuat landasan berfikir dalam penelitian ini, berikut berbagai konsep teoritis yang penulis gunakan adalah sebagai berikut

Siswa

Siswa menurut ketentuan umum undang-undang RI tentang sistem pendidikan nasional adalah anggota masyarakat yang berusaha mengembangkan potensi diri melalui proses pembelajaran yang tersedia pada jalur, jenjang, dan jenis pendidikan tertentu [10]. Sebagaimana yang diamanatkan oleh UU No. 20 tahun 2003 tentang sistem pendidikan nasional bab II pasal 3 dijelaskan mengenai fungsi dan tujuan Pendidikan Indonesia; Fungsi pendidikan yaitu mengembangkan kemampuan dan membentuk watak serta peradaban bangsa. Hal ini bertujuan untuk berkembangnya potensi peserta didik agar menjadi manusia seutuhnya yang beriman dan bertakwa kepada Tuhan YME, berakhlak mulia, sehat berilmu, cakap, kreatif, mandiri, dan menjadi warga negara yang demokratis serta bertanggungjawab

Pelanggaran Sikap Siswa

Sebuah Pelanggaran sikap siswa di sekolah dapat dikategorikan sebuah kenakalan remaja. Pada masa remaja ini tidak jarang remaja melakukan tingkah laku yang dianggap melanggar aturan yang ada. Dalam hal ini biasa disebut dengan kenakalan remaja. Dikatakan sebagai kenakalan remaja, apabila tindak perbuatan sebagian para remaja yang bertentangan dengan hukum, agama dan norma-norma masyarakat sehingga akibatnya dapat merugikan orang lain, mengganggu ketenteraman umum dan juga merusak dirinya sendiri [2].

Data Mining

Menurut Turban et al dalam buku Kusri, data mining adalah : "Data mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam database. Data mining merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dan berguna yang bermanfaat yang tersimpan di dalam database besar" [11].

Tahap-tahap Data mining

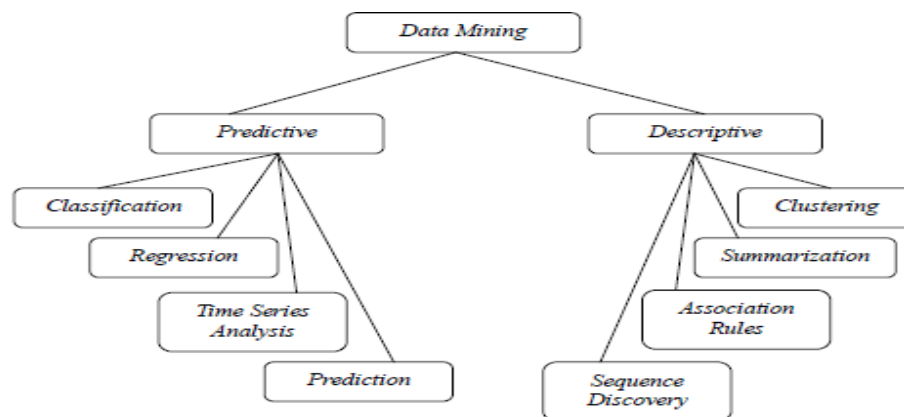
1. Pembersihan data (*data cleaning*)
Pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan.
2. Integrasi data (*data integration*)
Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru.
3. Seleksi data (*data selection*)
Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database.
4. Transformasi data (*data transformation*)
Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining.
5. Proses mining
Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.
6. Evaluasi pola (*pattern evaluation*)
Untuk mengidentifikasi pola-pola menarik ke dalam knowledge based yang ditemukan.

7. Presentasi pengetahuan (*knowledge presentation*)
Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna

Teknik-teknik Data mining

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Perlu diingat bahwa kata mining sendiri berarti usaha untuk mendapatkan sedikit data berharga dari sejumlah besar data dasar. Karena itu data mining sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), machine learning, statistik dan basis data.

Teknik data mining biasanya terbagi dalam dua kategori, prediksi dan deskripsi. Teknik prediksi menggunakan data historis untuk menyimpulkan sesuatu tentang kejadian pada masa depan. Sedangkan teknik deskripsi bertujuan untuk menemukan pola dalam data yang menyediakan beberapa informasi tentang hubungan interval yang tersembunyi.



Gambar 1 Pembagian teknik data mining menurut kategorinya.

Association Rule

Analisis asosiasi atau *Association Rule* mining adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Aturan asosiasi ingin memberikan informasi tersebut dalam bentuk hubungan if then atau jika maka. Aturan ini dihitung dari data yang sifatnya probabilistik [12]. Analisis asosiasi juga dapat didefinisikan suatu proses untuk menemukan semua aturan asosiatif yang memenuhi syarat *minimum* untuk support (*minimum support*) dan syarat *minimum* untuk confidence (*minimum confidence*) [13]. *Association Rule* mining adalah suatu prosedur untuk mencari hubungan antar item dalam suatu data set yang ditentukan [5].

Association rule meliputi dua tahap [14]:

1. Mencari kombinasi yang paling sering terjadi dari suatu item set.
2. Mendefinisikan condition dan result (untuk conditional *association Rule*)

FP-Growth

FP-Growth adalah algoritma dalam teknik asosiasi pada data mining. *Frequent Pattern Growth (FP Growth)* merupakan salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data [9]. Algoritma *FP-Growth* adalah pengembangan dari algoritma Apriori. Sehingga kekurangan dari algoritma Apriori diperbaiki oleh algoritma *FP-Growth*.

Algoritma apriori dalam penggunaannya memerlukan *generate candidate* untuk mendapatkan *frequent itemsets*. Akan tetapi, bila di algoritma *FP-Growth generate candidate* tidak dilakukan karena *FP-Growth* menggunakan konsep pembangunan tree dalam pencarian *frequent itemsets*.

Hal tersebutlah yang menyebabkan algoritma *FP-Growth* lebih cepat dari algoritma Apriori. Metode *FP-Growth* dapat dibagi menjadi 3 tahapan utama yaitu [5]:

1. Tahap pembangkitan *conditional pattern base*.
Conditional Pattern Base merupakan subdatabase yang berisi *prefix path* (lintasan prefix) dan *suffix pattern* (pola akhiran). Pembangkitan *conditional pattern base* didapatkan melalui *Fp-Tree* yang telah dibangun sebelumnya.
2. Tahap pembangkitan *conditional Fp-Tree*.
Pada tahap ini, *support count* dari setiap item pada setiap *conditional pattern base* dijumlahkan, lalu setiap item yang memiliki jumlah *support count* lebih besar sama dengan *minimum support count* ξ akan dibangkitkan dengan *conditional Fp-Tree*.
3. Tahap pencarian *frequent itemset*.
Apabila *Conditional Fp-Tree* merupakan lintasan tunggal (*single path*), maka didapatkan *frequent itemset* dengan melakukan kombinasi item untuk setiap *conditional Fp-Tree*. Jika bukan lintasan tunggal, maka dilakukan pembangkitan *FP-Growth* secara rekursif.

3. Metode Penelitian

Tahapan Penelitian

- a. Perumusan Masalah
Pada penelitian ini masalah yang dirumuskan adalah bagaimana menganalisis data pelanggaran di antara siswa di sekolah lanjutan Advent Purwodadi mahasiswa dengan menggunakan metode *Association Rule* untuk menentukan aturan dari korelasi antara pelanggaran di Sekolah Lanjutan Advent Purwodadi.
- b. Penentuan Tujuan
Tujuan yang akan dicapai dalam penelitian ini adalah menemukan aturan pelanggaran di Sekolah Lanjutan Advent Purwodadi dan memanfaatkan informasi yang didapatkan untuk membantu pihak sekolah mengatur siswa/i nya
- c. Mempelajari Literatur
Mempelajari literatur-literatur yang dapat mencapai tujuan penelitian, literatur-literatur bersumber dari jaringan internet. Literatur-literatur yang digunakan nanti dilampirkan dalam daftar pustaka.
- d. Pengumpulan Data dan Informasi
Dalam pengumpulan data, penulis mendapatkan data melalui bagian Kesiswaan Sekolah Lanjutan Advent Purwodadi
- e. Proses Data Mining
Data Mining adalah proses pengekstrasian *knowledge* yang tersimpan dalam database bervolume besar. Untuk mendapatkan *knowledge* dalam database digunakanlah teknik *Clustering K-Mean*.

Data

Dokumen yang diperoleh adalah data pelanggaran - pelanggaran sikap siswa yang dimiliki Sekolah Lanjutan Advent Purwodadi dari tahun 2015-2019. Data berikut yang akan dianalisis korelasinya. *Dataset* yang akan digunakan memiliki atribut berjumlah 30 atribut. Atribut yang dipilih berdasarkan buku Pedoman Peraturan Umum Kehidupan Siswa/I SLA Purwodadi.

Preprocessing Data

Data yang telah berhasil dikumpulkan selanjutnya akan dilakukan tahapan *preprocessing*. Pada tahapan ini dilakukan *cleaning* dan *transformation* data agar dapat digunakan pada aturan asosiasi. Berikut adalah proses yang dilakukan pada tahapan *cleaning* dan *transformation*.

1. *Cleaning*

Data siswa yang memiliki pelanggaran yang sama lebih dari 1 kali, akan dilakukan *cleaning data*, ini untuk menghindari adanya redudansi data. contoh: seorang anak melakukan pelanggaran rokok setelah didata, seminggu lagi dia melakukan pelanggaran yang sama merokok. Pada penelitian ini pelanggaran yang akan didata hanya rokok satu kali saja. *Cleaning* dilakukan juga pada atribut yang tidak digunakan, sehingga atribut itu akan dihapuskan. Contoh: dari tahun 2016 hingga tahun 2020 tidak ditemukan data pelanggaran pelanggaran berzina, sehingga pada penelitian ini atribut berzina tidak di ikutsertakan.

2. *Transformation*

Data pelanggaran yang telah didapatkan selanjutnya akan di transformasi menjadi format tabular, yakni format yang dapat diolah oleh tools Rapidminer. Data pelanggaran siswa dalam bentuk tabular dapat dilihat pada tabel 1.

Tabel 1 Data Pelanggaran Siswa Tahun 2016 - 2020

Perkelahiaan(p1)	Pemukulan(p2)	Pencurian(p3)	Rokok(p4)	Minuman keras(p5)
0	0	0	1	0
1	0	0	1	0
0	0	0	0	0
0	0	0	1	0
0	0	0	1	0
0	0	0	0	0
0	0	0	1	0

FP-Growth

Tahapan selanjutnya pada pada program *Rapidminer* adalah memilih algoritma yang akan digunakan. Algoritma pada penelitian ini menggunakan algoritma *FP-Growth*. Pada tahapan ini dihasilkan sebuah *frequent itemset* dari pelanggaran siswa/i di Sekolah Lanjutan Advent Purwodadi. Pada tabel 2 adalah *frequent Item Set* yang didapatkan dengan menggunakan algoritma *FP-Growth*.

Tabel 2 *Frequent Item Sets* - FP-Growth menggunakan Rapid Miner

Size	Support	Item1	Item2	Item3	Item4
1	0.545	Rokok			
2	0.291	Rokok	Keluar tanpa izin		
3	0.091	Rokok	Pemukulan	Pengancaman	
4	0.082	Rokok	Pemukulan	Pengancaman	Pembangngangan

Penentuan *Minimum support*

Penentuan *minimum support* dengan menggunakan *FP-Growth* terjadi sebelum perhitungan algoritma *FP-Growth* dimulai. Pada program Rapidminer terdapat parameter *min support* yang terlebih dahulu harus diisi. Untuk default yang diberikan dari program adalah 0.95. pada penelitian ini penulis menentukan *minimum support* yang akan digunakan adalah 0.08 atau 8%.

Berdasarkan wawancara dengan pihak terkait, ditemukan bahwa pihak sekolah ingin lebih mengetahui pelanggaran, yang terdapat pada golongan A, terutama dari golongan A adalah pelanggaran merokok dari pada jenis pelanggaran lainnya. Pihak terkait juga menyatakan bahwa pelanggaran yang sering dilakukan berulang-ulang oleh siswa yang sama adalah pelanggaran akan penggunaan rokok. Oleh sebab itu rokok akan menjadi acuan dalam penentuan *minimum support* ini.

Nilai support tertinggi pada pelanggaran akan rokok, dengan nilai confidence diatas nilai *minimum confidence* yang telah ditentukan adalah 12.7%. lalu dari 12.7% penulis menganalisis bahwa 8% menjadi landasan terbaik dalam menentukan *minimum support*. Dikarenakan jika dilihat secara berurutan pada tabel 4.5, terdapat nilai support pada pelanggaran penggunaan rokok dengan nilai *support* 12.7%, 10.9 %, 10%, 9,1%, 8,2 % yang nilai *confidence*-nya diatas nilai *minimum confidence* yang telah ditentukan. Sedangkan pada kisaran 5% - 7% tidak ditemukan lagi adanya nilai support yang memenuhi nilai *minimum confidence* yang telah ditentukan, Sehingga 8% ditentukan sebagai *minimum support* untuk menentukan aturan pada pelanggaran siswa. Berikut pada tabel 3 penjelasan nilai *support* pada beberapa data pelanggaran akan penggunaan rokok dengan *confidence* melebihi *minimum confidence* yang telah ditentukan.

Adapun penentuan nilai *support* menggunakan persamaan berikut:

$$Support (A, B) = P(A \cap B) = \frac{Jumlah\ transaksi\ yang\ mengandung\ A\ dan\ B}{Jumlah\ total\ transaksi}$$

Tabel 3 Nilai *support data* pelanggaran penggunaan rokok

premises	conclusion	support	confidence
Keluar tanpa izin, pemukulan	Rokok	1.9%	100%
Gaya rambut, Sepatu tidak sesuai standar	Rokok	2.7%	75%
Keluar tanpa izin, absen melebihi maximum	Rokok	4.5%	71%
Pengancaman, Pemukulan,	Rokok, Pembangkangan	8.2%	81%
Pemukulan, pengancaman	Rokok,	9.1%	90.9%
Pembangkangan	Rokok	10%	78%
Pengancaman	Rokok	10.9%	80%
Minuman Keras	Rokok	12.7%	93%

Association Rule

Selanjutnya adalah menentukan *Association Rules* dari *Frequent Item Sets* yang sudah didapatkan melalui proses *FP-Growth* dengan menggunakan program *Rapid Miner*. pada tahapan tersebut dihasilkan aturan (*rule*) sesuai dengan *minimum confidence* yang sudah ditentukan yakni 0.7 atau 70%. Pada tabel 4 akan ditampilkan hasil dari perhitungan *Association Rule* oleh *Rapid Miner*.

Tabel 4 *Association Rule*

No	Premises	Conclusion	Support	Confidence	Lift
1	Pembangkangan	Pengancaman	0.091	0.714	5.238
2	Pembangkangan	Pemukulan, pengancaman	0.091	0.714	7.143
3	Pengancaman	Pemukulan	0.100	0.733	3.507
4	Rokok, Pengancaman	pembangkangan	0.082	0.750	5.893
5	Rokok, Pengancaman	Pemukulan, pembangkangan	0.082	0.750	6.875
6	Pemukulan Pembangkangan	Rokok, pengancaman	0.082	0.750	6.875
7	Pembangkangan	Rokok	0.100	0.786	1.440
8	Pembangkangan	Rokok, pemukulan	0.100	0.786	5.402
9	Pengancaman	Rokok	0.109	0.800	1.467
10	Rokok, Pembangkangan	Pengancaman	0.082	0.818	6
11	Pemukulan, Pengancaman	Rokok, pembangkangan	0.082	0.818	8.182
12	Rokok, pembangkangan	Pemukulan, pengancaman	0.082	0.818	8.182
13	Rokok, Pemukulan, Pembangkangan	Pengancaman	0.082	0.818	6
14	Rokok, Pengancaman	Pemukulan	0.082	0.833	3.986
15	Pemukulan, pembangkangan	Pengancaman	0.091	0.833	6.111
16	Pembangkangan	Pemukulan	0.091	0.857	4.099
17	Pengancaman, pembangkangan	rokok	0.109	0.900	1.650
18	Rokok, pemukulan, pengancaman	Pembangkangan	0.082	0.900	7.071
19	Pengancaman, pembangkangan	Rokok, pemukulan	0.082	0.900	6.188
20	Pemukulan, pengancaman, pembangkangan	Rokok	0.082	0.900	1.650

21	Pemukulan, pengancaman	Rokok	0.091	0.909	1.667
22	Pemukulan, pengancaman	Pembangangan	0.091	0.909	7.143
23	Pemukulan, pembangkangan	Rokok	0.100	0.917	1.681
24	Minuman keras	Rokok	0.127	0.933	1.711
25	Rokok pembangkangan	Pemukulan	0.100	1	4.783
26	Pengancaman, pembangkangan	Pemukulan	0.091	1	4.783
27	Rokok, pengancaman, pembangkangan	Pemukulan	0.082	1	4.783

Penentuan *Minimum Confidence*

Penulis menetapkan 0.7 atau 70% sebagai *minimum confidence*. 70 % juga sering menjadi angka yang penting dalam dunia pendidikan dalam menentukan standar kelulusan. Seperti yang dikutip dari website resmi kemdikbud menyatakan berdasarkan rekap KKM mata pelajaran dari setiap kelas diatas maka KKM satuan pendidikan adalah 70. 70 merupakan nilai terendah dari semua KKM mata pelajaran yang ada [15]. Pada penelitian ini penulis juga mengambil 70% sebagai *minimum confidence* dikarenakan penelitian ini juga dilakukan pada institusi pendidikan sehingga, aturan yang didapatkan diharapkan dapat diterima oleh pihak sekolah selaku lembaga pendidikan, yang nantinya akan mengimplementasikan aturan tersebut untuk mengatur siswanya.

Penulis juga menilai bahwa penelitian yang menyangkut psikologi anak didik memerlukan ketepatan tinggi, disebabkan ini akan menyangkut masa depan anak didik. Oleh sebab itu penelitian ini nantinya akan membantu pihak sekolah dalam menyelidiki sebuah kasus permasalahan siswa/i-nya, ataupun penelitian ini dapat sebagai bahan pertimbangan dalam pengambilan keputusan anak bermasalah.

Adapun untuk pencarian nilai *confidence* menggunakan persamaan berikut:

$$Confidence (A \rightarrow B) = \frac{Jumlah\ transaksi\ yang\ mengandung\ A\ dan\ B}{Jumlah\ total\ transaksi\ yang\ mengandung\ A}$$

Lift Ratio

Dalam menentukan *lift ratio* diperlukan nilai *confidence* dari sebuah korelasi terlebih dahulu. Rumus mencari *lift ratio* menggunakan persamaan berikut:

$$Lift\ ratio = \frac{Confidence(A,C)}{Benchmark\ confidence(A,c)}$$

Benchmark confidence adalah jumlah transaksi dengan item yang menjadi *consequent (Nc)* dibagi oleh jumlah transaksi basis data (*N*), yang dalam hal ini adalah total data siswa/i pelanggaran di sekolah Lanjutan Advent Purwodadi. Rumus mencari *Benchmark confidence* terdapat pada persamaan berikut:

$$Benchmark\ confidence = \frac{Nc}{N}$$

Lift ratio digunakan untuk mengetahui valid atau tidaknya aturan asosiasi yang telah dibentuk. Adapun 3 kemungkinan yang akan dihasilkan saat menghitung *lift ratio*, yaitu:

- a. Jika *lift* < 1, maka kemunculan item A berkorelasi negatif dengan berkorelasi negatif dengan kemunculan item B. hal ini berarti bahwa kemunculan satu kemungkinan mengarah pada ketiadaan yang lain
- b. Jika *lift* > 1, maka item A dan B adalah berkorelasi positif. Hal ini berarti bahwa satu menyiratkan terjadinya yang lain.
- c. Jika *lift* = 1, maka item A dan B adalah independen dan tidak ada korelasi di antara kedua item. Nilai lift yang dibutuhkan untuk mengetahui korelasi antara kombinasi item valid yaitu perhitungan *lift* > 1 [16].

Proses Koneksi Data Mining

Berikut pada gambar 2 adalah proses koneksi data menggunakan algoritma *FP-Growth* pada tools Rapidminer.



Gambar 2 Proses Koneksi Data Mining

Proses ini menghasilkan 27 aturan asosiasi. Aturan ini telah disesuaikan dengan *min support* dan *min confidence* yang telah ditetapkan. Pada Tabel 23 ditampilkan aturan yang telah dihasilkan.

4. Hasil

Hasil *Association Rule*

Berdasarkan penentuan aturan asosiasi dengan menggunakan algoritma *FP-Growth* dalam menganalisis data pelanggaran siswa/i, ditemukan 27 aturan dari 110 data pelanggaran siswa yang didapatkan dari tahun 2016 hingga 2020.

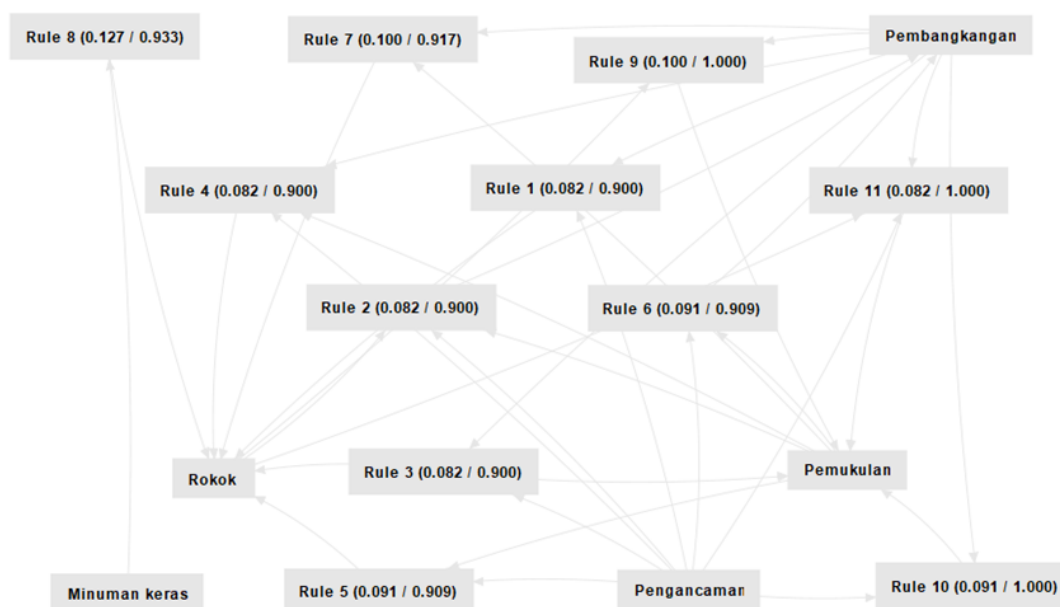
Pembahasan

Penjelasan mengenai hasil dan aturan final berdasarkan tabel 4 pada penelitian ini adalah.

- 1) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pembangkangan kemudian dia akan melakukan pelanggaran pengancaman, menghasilkan nilai confidence sebesar 71%.
- 2) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pembangkangan, kemudian dia akan melakukan pelanggaran pemukulan dan pengancaman, menghasilkan nilai confidence sebesar 71%.
- 3) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pengancaman kemudian dia akan melakukan pelanggaran pemukulan, menghasilkan nilai confidence sebesar 73%.
- 4) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran akan penggunaan rokok dan pengancaman kemudian dia akan melakukan pelanggaran pembangkangan, menghasilkan confidence sebesar 75%.
- 5) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran akan penggunaan rokok dan pengancaman kemudian dia akan melakukan pelanggaran pemukulan dan pembangkangan menghasilkan nilai confidence sebesar 75%.
- 6) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran Pemukulan dan pembangkangan kemudian dia akan melakukan pelanggaran akan penggunaan rokok dan pengancaman menghasilkan confidence sebesar 75%.
- 7) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pembangkangan kemudian dia akan melakukan pelanggaran penggunaan akan rokok menghasilkan nilai confidence sebesar 78%.
- 8) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pembangkangan kemudian dia akan melakukan pelanggaran dan pemukulan menghasilkan nilai confidence sebesar 78%.

- 9) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pengancaman kemungkinan dia akan melakukan pelanggaran rokok menghasilkan nilai confidence sebesar 80%.
- 10) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran rokok dan pembangkangan kemudian dia akan melakukan pelanggaran pengancaman menghasilkan nilai confidence sebesar 81%.
- 11) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pemukulan dan pengancaman kemudian dia akan melakukan pelanggaran rokok dan pembangkangan menghasilkan nilai confidence sebesar 81%.
- 12) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran rokok dan pembangkangan kemudian dia akan melakukan pelanggaran pemukulan dan pengancaman menghasilkan nilai confidence sebesar 81%.
- 13) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran rokok, pemukulan, dan pengancaman kemudian dia akan melakukan pelanggaran pengancaman menghasilkan nilai confidence sebesar 75%.
- 14) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran Rokok dan pengancaman kemudian dia akan melakukan pelanggaran pemukulan menghasilkan nilai confidence sebesar 83%.
- 15) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pemukulan dan pembangkangan kemudian dia akan melakukan pelanggaran pengancaman menghasilkan nilai confidence sebesar 83%.
- 16) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pembangkangan kemudian dia akan melakukan pelanggaran pemukulan menghasilkan nilai confidence sebesar 85%.
- 17) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pengancaman dan pembangkangan kemungkinan dia akan melakukan pelanggaran rokok nilai confidence sebesar 81%.
- 18) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran Rokok ,pemukulan, dan pengancama, kemungkinan dia akan melakukan pelanggaran pembangkangan menghasilkan nilai confidence sebesar 90%.
- 19) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pengancaman dan pembangkangan kemudian dia akan melakukan pelanggaran rokok dan pemukulan menghasilkan nilai confidence sebesar 90%.
- 20) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pemukulan, pengancaman, dan pembangkangan kemudian dia akan melakukan pelanggaran rokok menghasilkan nilai confidence sebesar 90%.
- 21) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pemukulan dan pengancaman kemudian dia akan melakukan pelanggaran rokok menghasilkan nilai confidence sebesar 90,9%.
- 22) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pemukulan dan pengancaman kemudian dia akan melakukan pelanggaran pembangkangan menghasilkan confidence sebesar 90,9%.
- 23) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pemukulan dan pembangkangan kemudian dia akan melakukan pelanggaran rokok menghasilkan nilai confidence sebesar 91,7%.
- 24) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran minuman keras kemudian akan melakukan menghasilkan nilai confidence sebesar 93%.

- 25) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran rokok dan pembangkangan kemudian dia akan melakukan pelanggaran pemukulan menghasilkan nilai confidence sebesar 100%.
- 26) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran pengancaman dan pembangkangan kemudian dia akan melakukan pelanggaran pemukulan menghasilkan nilai confidence sebesar 100%.
- 27) Ketika seorang anak selama dia bersekolah pernah melakukan pelanggaran rokok, pengancaman, dan pembangkangan kemungkinan dia akan melakukan pelanggaran pemukulan kemudian nilai confidence sebesar 100%.



Gambar 3 Graph rules dengan nilai confidence 90% ke atas.

Dari penelitian ini dapat dilihat semakin tinggi dalam penentuan nilai confidence maka aturan (rule) yang akan terbentuk semakin sedikit. Hasil pengelolaan data juga mendapatkan bahwa nilai confidence tertinggi yang dapat di temukan melalui penelitian ini adalah 100%. Terdapat pada gambar 3 untuk aturan yang memiliki nilai confidence 90% hingga 100%.

Implementasi *Association Rule*

Aturan yang didapatkan dari penelitian yang dilakukan. Dapat digunakan pihak sekolah untuk mengatur siswannya. Aturan ini dapat digunakan di antaranya dengan cara:

1. Aturan ini dapat digunakan dalam menyelidiki sebuah kasus pelanggaran di antara siswa/i yang bersekolah di sekolah tempat penelitian
2. Aturan dapat digunakan menentukan tindakan terbaik dalam pengambilan keputusan dari pelanggaran yang terjadi.
3. Aturan (*rules*) yang didapatkan juga dapat digunakan sebagai landasan dalam mengembangkan peraturan sekolah, yang terdapat pada bab 4 buku peraturan sekolah yang berjudul "Pelanggaran dan Sanksinya".

5. Kesimpulan

Dari hasil penelitian dan pengujian yang dilakukan penulis, maka metode Association Rule dapat digunakan menganalisa data pelanggaran-pelanggran sikap siswa disekolah. Algoritma *FP-Growth* dalam menganalisis data pelanggaran siswa/i, ditemukan 27 aturan dari 110 data pelanggaran siswa yang didapatkan dari tahun 2016 hingga 2020. Dengan nilai *minimum support* yang ditentukan adalah 0.08 atau 8% dan nilai *minimum confidence* yang ditentukan adalah 0.7 atau 70%.

6. Referensi

- [1] W. S. Sarlito, Psikologi Remaja, Jakarta: Raja Grafindo Pustaka, 2012.
- [2] M. Strategi Belajar Mengajar, Surabaya: Citra Media, 2005.
- [3] S. S. Willis, Remaja dan Masalahnya, Bandung: Alfabeta, 2010.
- [4] M. Kun and S. Juju, Sosiologi untuk SMA dan MA Kelas X, Jakarta: Esis, 2007.
- [5] H. Jiawei and M. Kamber, Data Mining: Concepts and Techniques, 2nd Edition, San Fransisco: Morgan Kaufmann Publishers, 2006.
- [6] A. A. Fajrin and A. Maulana, "Penerapan Data Mining Untuk Analisa Pola Pembelian Konsumen Dengan Algoritma *FP-Growth* Pada Data transaksi Penjualan Spare Part Motor," *Kumpulan Jurnal Ilmu Komputer (KLIK)*, pp. 27-36, 2018.
- [7] Y. W. Wahdi, "Implementasi Data Mining Untuk Analisa Tingkat Pelanggaran Lalu Lintas Dengan Algoritma *Association Rule*," *Ruang Teknik Journal*, pp. 105-112, 2018.
- [8] Y. S. Nugroho, "Data Mining Menggunakan Algoritma Naive Bayes Untuk Klasifikasi Kelulusan Universitas Dian Nuswantoro," pp. 1-10, 2014.
- [9] D. Samuel, "Penerapan Struktur *Fp-Tree* dan Algoritma *FP-Growth* dalam Optimasi Penentuan *Frequent itemset*," *Institut Teknologi Bandung*, 2008.
- [10] I. Ali, Manajemen Peserta Didik Berbasis Sekolah, Jakarta: Bumi Aksara, 2011.
- [11] E. T. Luthfi and Kusriani, Algoritma Data Mining, Yogyakarta: Andi Offset, 2009.
- [12] S. Budi, Data Mining Teknik Pemanfaatan Data Untuk Keperluan Bsinis, Yogyakarta: Graha Ilmu, 2007.
- [13] I. Pramudiono, "http://datamining.japati.net," 16 Januari 2007. [Online]. Available: <http://datamining.japati.net/cgi-bin/indodm.cgi?bacaarsip&1155527614&artikel>.
- [14] U. David, "Mining an Online Auctions Data Warehouse," in *Proceedings of MASPLAS'02 the Mid-Atlantic Student Workshop on Programming Languages and System Pace University*, 2002.
- [15] kemdikbud, "PENETAPAN KRITERIA KETUNTASAN MINIMAL," [Online]. Available: https://pklk.gtk.kemdikbud.go.id/webpage/show_pdf_article/artikel/2e01e17467891f7c933dbaa00e1459d23db3fe4f.
- [16] J. Han, J. Pei and M. Kamber, Data mining, concepts and techniques., Elsevier, 2011.