



Analisis Tingkat Kejahatan pada Anak Dibawah Umur Menggunakan Metode FP-Growth

Angga Putra Juledi^{1✉}, Sarjon Defit², Yuhandri Yunus³
^{1,2,3}Universitas Putra Indonesia YPTK Padang

anggaj19@gmail.com

Abstract

Crimes in minors are a series of negligence by parents who endanger or pose a dangerous threat to the child. The purpose of this study is to implement Data Mining, Association rule, and the FP-Growth Algorithm in cases of juvenile crime so that it can extract knowledge and important and interesting information from the database. The data source used is raw data that has not been processed and is a crime data on minors which are summarized in the form of reports from the West Sumatra Regional Police. The results of this study are in the form of software by analyzing data collected using the FP-Growth Algorithm and using the concept of FP-Tree development in searching for Frequent Itemset, for testing the results carried out with applications that have been designed namely the Php programming language. The results of testing are obtained from associations of crime cases that often occur in minors. So it can be seen that data mining using the FP-Growth Algorithm can be used to analyze cases of crime in minors as a material consideration for the police in order to know the ins and outs of crime in children so that it can assist the investigation process.

Keywords: Data Mining, Association Rule, Algoritma FP-Growth, Frequent Itemset, FP-Tree.

Abstrak

Kejahatan pada anak dibawah umur merupakan serangkaian kelalaian oleh orang tua yang membahayakan atau memberikan ancaman yang berbahaya kepada anak. Tujuan penelitian ini untuk mengimplementasikan *Data Mining*, Association rule, dan Algoritma *FP-Growth* pada kasus kejahatan anak dibawah umur sehingga bisa mengekstrak ilmu pengetahuan serta informasi penting dan menarik dari database. Sumber data yang digunakan merupakan data mentah yang belum diolah dan merupakan data kejahatan pada anak di bawah umur yang dirangkup berupa laporan dari Polda Sumbar. Hasil penelitian ini adalah berupa suatu perangkat lunak dengan menganalisa data yang dikumpulkan menggunakan Algoritma *FP-Growth* dan menggunakan konsep pembangunan *FP-Tree* dalam mencari *Frequent Itemset*, untuk pengujian hasil dilakukan dengan aplikasi yang telah dirancang yaitu bahasa pemrograman Php. Hasil dari pengujian didapatkan dari asosiasi kasus kejahatan yang sering terjadi pada anak dibawah umur. Sehingga dapat diketahui bahwa data mining menggunakan Algoritma *FP-Growth* bisa digunakan untuk menganalisa kasus kejahatan pada anak dibawah umur sebagai bahan pertimbangan untuk pihak kepolisian agar dapat mengetahui seluk beluk kejahatan pada anak sehingga dapat membantu proses penyidikan.

Kata kunci: *Data Mining*, *Association Rule*, *Algoritma FP-Growth*, *Frequent Itemset*, *FP-Tree*.

© 2020 JSisfotek

1. Pendahuluan

Knowledge Discovery in Database (KDD) adalah proses menentukan informasi yang berguna terhadap pola-pola yang ada dalam data. Informasi yang dihasilkan bersumber dari basis data yang berukuran besar, dimana informasi ini tidak diketahui dan bermanfaat sebelumnya [1]. Untuk mengetahui informasi yang terkandung tersebut, maka diperlukan metode data mining. Proses data mining ini dapat menemukan informasi baru dengan mencari pola atau aturan tertentu dari data yang sangat besar tersebut [2]. KDD dan data mining sering digunakan secara bersamaan untuk menjelaskan proses pencarian informasi yang tersembunyi di dalam database yang besar. Keduanya ini memiliki konsep yang berbeda tetapi saling berkaitan satu sama lainnya.

Salah satu proses dalam data mining adalah *Association Rules*. Proses ini merupakan suatu proses untuk menentukan semua aturan asosiatif yang memenuhi persyaratan minimum dalam support (minsup) dan

confidence (minconf) pada sebuah database [3]. Salah satu algoritma yang digunakan dalam data mining adalah *FP-Growth*.

Algoritma *FP-Growth* merupakan algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sekumpulan data [4]. Algoritma *FP-Growth* berfungsi untuk melakukan identifikasi dan mengekstraksi informasi yang bermanfaat dan pengetahuan yang terkait dalam database yang besar. Algoritma *FP-Growth* melakukan pemanggilan database yang sama sebanyak dua kali [5].

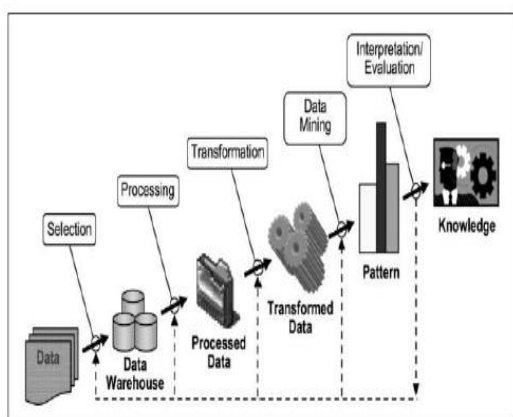
FP-Growth menerapkan konsep pembangunan akar (*tree*) dalam pencarian itemset. Hal tersebutlah yang menyebabkan algoritma *FP-Growth* lebih cepat dari algoritma lainnya. Sehingga algoritma *FP-Growth* disebut juga dengan *FP-Tree*. Dengan menggunakan *FP-Tree* terhadap algoritma *FP-Growth* dapat secara langsung mengekstrak *frequent itemset* data yang besar [6].

Salah satu data yang besar adalah data yang berhubungan dengan masyarakat. Item data yang berhubungan dengan masyarakat ini sangat terpengaruh oleh perkembangan zaman dan peningkatan jumlah penduduk. *Item* data ini dapat berupa mempengaruhi perilaku masyarakat, seperti pengangguran dan kemiskinan. Dampak dari perilaku ini mengakibatkan tingginya angka kriminalitas terutama kejahatan terhadap anak di bawah umur.

Kejahatan terhadap anak dari tahun ketahun semakin meningkat dan menjadi perhatian di kalangan masyarakat. Menurut hasil penelitian pada Badan Keluarga Berencana dan Pemberdayaan Perempuan Kota Padang pada tahun 2018, terdapat banyaknya kasus kejahatan yang dialami oleh lingkup rumah tangga terutama pada anak dibawah umur. Untuk itu, dilakukan penelitian ini dalam menggali informasi tentang kasus kejahatan pada anak. Penelitian ini mencari faktor-faktor yang paling sering menyebabkan tindak kekerasan pada anak dibawah umur. Hasil dari informasi ini dapat menentukan keputusan yang tepat dalam menanggulangi kejahatan pada anak. Dalam penelitian ini dilakukan pengujian menggunakan metode algoritma *FP-Growth* [9]. Hasil dari penelitian ini dapat memberikan informasi dalam membantu pihak kepolisian dan masyarakat untuk mengantisipasi kejahatan kekerasan pada anak dibawah umur yang sering muncul.

2. Metodologi Penelitian

Dalam penelitian ini menggunakan KDD untuk mengekstraksi data yang potensial, implicit dalam mengetahui informasi terjadinya kejahatan pada anak. Proses penggalian informasi ini melibatkan proses data mining dalam mendapatkan akurasi yang tinggi dan mudah dipelajari [6]. Tahapan-tahapan proses dalam KDD disajikan pada Gambar 1.



Gambar 1. Tahap-tahap KDD

Salah satu tahapan dalam gambar 1 diatas adalah *data mining*. Tahapan ini merupakan sebuah pengembangan atau menghasilkan informasi yang baru dengan cara menentukan pola atau aturan tertentu dari kumpulan data yang sangat banyak. *Data mining* disebut juga

sebagai rangkaian proses untuk mencari nilai tambah dari suatu data, sehingga menemukan pengetahuan yang selama ini tidak diketahui [10].

Dalam data mining terdapat proses *association rule* (aturan asosiatif). Proses ini berfungsi untuk menemukan *pattern* atau pola-pola dari sebuah kumpulan data [7]. *Association rule* bekerja untuk menemukan aturan-aturan tertentu yang mengasosiasikan sebuah data dengan data yang lainnya. Proses dalam menemukan *association rule* dari suatu kumpulan data harus diawali dengan pencarian kemunculan bersamaan (*frequent itemset*). Nilai yang dihasilkan berkisar antara 0% sampai dengan 100% yang menjadi pendukung (*support*) dari *association rule*. Rumus dalam proporsi dari transaksi dalam database adalah yang mengandung A dan B, dengan persamaan:

$$Support = P(A \cap B) \quad (1)$$

Confidence dari *association rule* adalah ukuran ketepatan suatu *rule*, yaitu persentase transaksi dalam database yang mengandung A dan mengandung B.

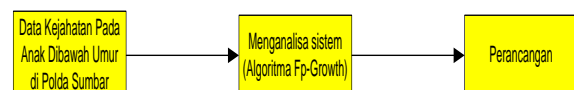
$$Confidence = \frac{A \cap B}{T} \quad (2)$$

Algoritma *FP-Growth* yang digunakan dalam penelitian ini merupakan suatu teknik pada data mining untuk menemukan *frequent itemset* tanpa menggunakan *candidate generation*. *FP-Growth* menggunakan pendekatan yang berbeda dari paradigma [8].

3. Hasil dan Pembahasan

Pada tahap ini data yang telah dikumpulkan dianalisis. Analisis menggunakan *association rule* yang dilakukan melalui mekanisme penghitungan *support* dan *confidence* dari suatu hubungan item. Sebuah *rule* asosiasi dikatakan *interesting* jika nilai *support* adalah lebih besar dari *minimum support* dan juga nilai *confidence* adalah lebih besar dari *minimum confidence*. Dengan menggunakan teknik *FP-Growth* yang menghasilkan *frequent itemset* tanpa melakukan *candidates generation*. Tujuannya adalah untuk mendapatkan pengetahuan yang baru (*knowledge*) yang berupa informasi tentang tindakan kejahatan pada anak di bawah umur yang sering terjadi.

Selanjutnya dilakukan beberapa tahapan algoritma *FP-Growth* untuk mengolah data. Tahapan proses disajikan pada Gambar 2. Tujuannya adalah untuk memudahkan dalam menganalisa dan merancang sistem.

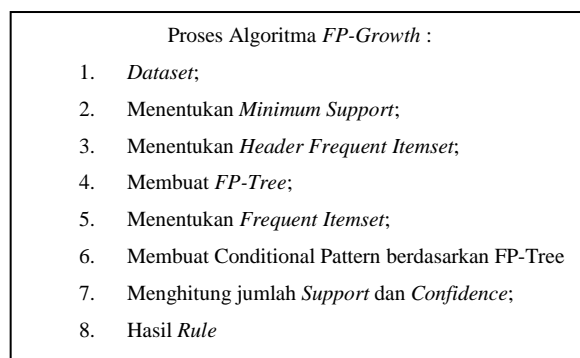


Gambar 2. Bagan Alir Analisa dan Perancangan

Data yang diolah dalam penelitian ini adalah data Kejahatan Anak Dibawah Umur pada Polisi Daerah Sumatera Barat. Data sumber masih berbentuk mentah dan belum diolah. Data kejahatan terdiri atas data kekerasan pada anak dibawah umur. Data ini mencakup laporan kriminalitas tentang kasus kejahatan pada anak dibawah umur di Polda Provinsi Sumatera Barat. Data yang diambil memiliki atribut yaitu pelaku, korban, kasus dan tempat kejadian perkara (TKP). Waktu kejahatan yang terjadi pada bulan Januari sampai dengan Desember 2018. Jumlah data yang diolah sebanyak 56 kasus.

Keterangan data yang telah diperoleh, maka dilanjutkan dengan proses penggolong berdasarkan kategori. Keterangan kategori yang pertama yaitu Tersangka dengan nilai data berisi Pelajar, Swasta, Pegawai (Gurun dan Guru Honor), Pengangguran, IRT (Ibu Rumah Tangga), Sopir, Pedagang, Petani, Buruh, dan Nelayan. Kategori yang kedua adalah Korban dengan nilai data berisi Pelajar, Balita, Pengangguran. Kategori yang ketiga adalah Kasus dengan nilai data berisi Pencurian (Jambret, Copet, Cubis, Curat, Curas, Curanmor, dan Pencurian), Penipuan (Penggelapan dan Penipuan), Penganiayaan (Kekerasan, Penganiayaan), Penculikan (Melarikan anak dibawah umur), Cabul (Pelecehan Sexual, Pemerkosaan, dan Cabul), dan Pengeroyokan. Berikutnya keterangan kategori yang keempat yaitu TKP (Tempat Kejadian Perkara) sesuai kejadian di Provinsi Sumatera Barat dengan nilai data berisi Padang, Kuranji, Lubeg, Pauh, Koto Tengah, Bungus, Lubuk Kilangan, Nanggalo, Pasaman, Pariaman, Solok, Payakumbuh, dan Padang Panjang.

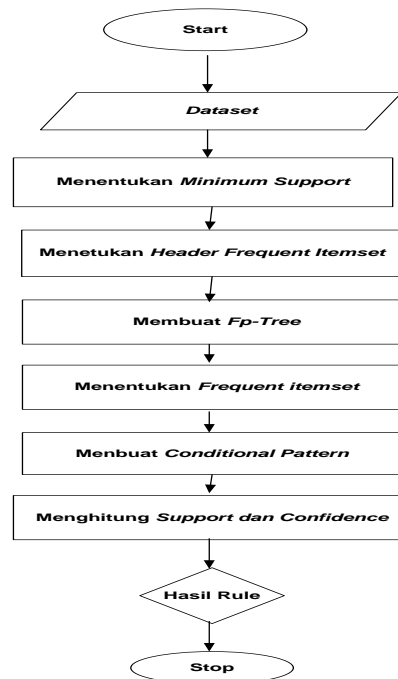
Proses selanjutnya adalah menganalisa sistem. Proses ini menggunakan algoritma *FP-Growth* dalam menentukan *frequent itemset*. Untuk menemukan sebuah pola transaksi yang sering muncul adalah dengan perluasan menggunakan pohon prefix atau *FP-Tree* sehingga memudahkan dalam pencarian *frequent itemset* dengan menggunakan struktur data yang ada.



Gambar 3. Algoritma *FP-Growth*

Gambar 3 merupakan algoritma *FP-Growth* dalam penelitian ini yang diterjemahkan kedalam bentuk *flowchart* atau diagram alir. *Flowchart* ini menggambarkan tahapan atau langkah-langkah dalam

klasifikasi menggunakan algoritma *FP-Growth* yang disajikan pada Gambar 4.



Gambar 4. Proses *Flowchart* pada Algoritma *FP-Growth*

3.1. Dataset

Data yang penelitian tidak dapat langsung diolah. Data ini terlebih dahulu ditransformasikan dan diberikan kode agar dapat memudahkan dalam melakukan proses pengolahan *data mining*. Atribut yang digunakan adalah tersangka, korban, kasus, dan TKP (tempat kejadian perkara).

3.2. Menentukan *Minimum Support*

Tabel 1 menyajikan hasil pemindaian yang memiliki frekuensi diatas *support count* $\xi=12$, yaitu A1, K5, T4, T1, dan K3.

Tabel 1. Data Kasus Hasil Pemindaian

Item	Frequent
A1	49
K5	23
T4	17
T1	16
K3	15

Setelah prose pemindaian, maka didapatkan item yang memiliki frekuensi diatas *support count* $\xi=12$, yaitu Korban Pelajar (A1), Kasus Cabul (K5), Tersangka Pengangguran (T4), Tersangka Pelajar (T1), dan Kasus Penganiayaan (K3). Ke 5 item inilah yang mempengaruhi dan juga digunakan untuk melakukan pencarian *FP-tree*. Item lainnya diabaikan karena tidak berpengaruh.

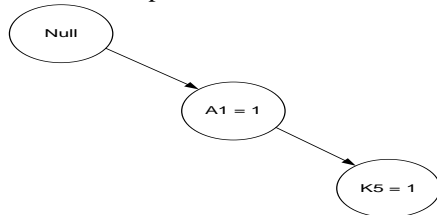
3.3. Menentukan Header Frequent Itemset

Dalam menentukan *Header Frequent Itemset* dalam data kasus kejahatan anak dibawah umur ditentukan atas nilai minimum *support* yang dibandingkan dengan dataset.

3.4. Menentukan FP-Tree

Langkah menentukan *FP-Tree* merupakan pembentukan *FP-Tree* berdasarkan *headen frequent itemset*. *Headen frequent* sebagai berikut:

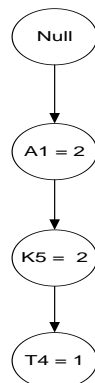
1. Gambar yang memberikan ilustrasi pembentukan *FP-Tree* setelah pembacaan tindakan 1.



Gambar 5 . Hasil Pembentukan *FP-Tree* Setelah Pembacaan Tindakan 1

Gambar 5 merupakan pembentukan *FP-Tree* setelah dilakukan pembacaan pada tindakan 1 yang berisi Null-A1=1 (Korban Pelajar), K5=1 (Kasus Cabul).

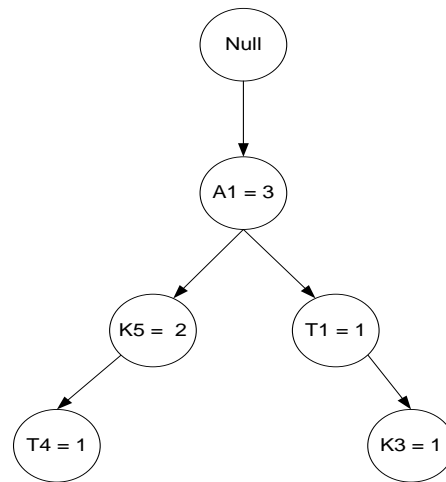
2. Gambar yang memberikan ilustrasi pembentukan *FP-Tree* setelah pembacaan tindakan 2.



Gambar 6. Hasil Pembentukan *FP-Tree* Setelah Pembacaan Tindakan 2

Gambar 6 merupakan pembentukan *FP-Tree* setelah melakukan pembacaan pada tindakan 2 yang berisi Null-A1=2 (Korban Pelajar), K5=2 (Kasus Cabul), T4=1 (Tersangka Pengangguran).

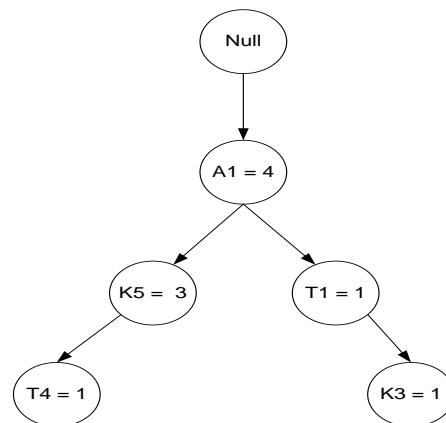
3. Gambar yang memberikan ilustrasi pembentukan *FP-Tree* setelah pembacaan tindakan 3.



Gambar 7. Hasil Pembentukan *FP-Tree* Setelah Pembacaan Tindakan 3

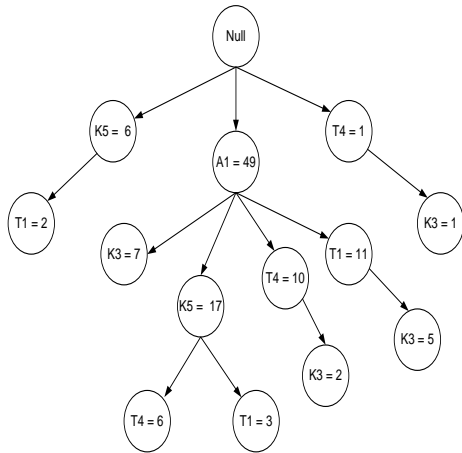
Gambar 7. merupakan pembentukan *FP-Tree* setelah melakukan pembacaan pada tindakan 3 yang berisi Null-A1=3 (Korban Pelajar), K5=2 (Kasus Cabul), T4=1 (Tersangka Pengangguran), T1=1 (Tersangka Pengangguran), K3=1 (Kasus Penganiayaan).

4. Berikut gambar yang memberikan ilustrasi pembentukan *FP-Tree* setelah pembacaan tindakan 4.



Gambar 8. Hasil Pembentukan *FP-Tree* Setelah Pembacaan Tindakan 4

Gambar 8 merupakan pembentukan *FP-Tree* setelah melakukan pembacaan pada tindakan 4 yang berisi Null-A1=4 (Korban Pelajar), K5=3 (Kasus Cabul), T4=1 (Tersangka Pengangguran), T1=1 (Tersangka Pengangguran), K3=1 (Kasus Penganiayaan). Gambar 8 merupakan ilustrasi pembentukan *FP-Tree* setelah pembacaan tindakan 5.

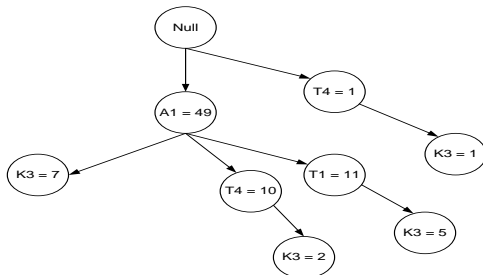


Gambar 9. Hasil Pembentukan *FP-Tree* Setelah Pembacaan Tindakan 5

Gambar 9 merupakan pembentukan *FP-Tree* setelah melakukan pembacaan pada tindakan 56 yang berisi Null- A1=49 (Korban Pelajar), K5=23 (Kasus Cabul), T4=17 (Tersangka Pengangguran), T1=16 (Tersangka Pelajar), dan K3=15 (Kasus Penganiayaan).

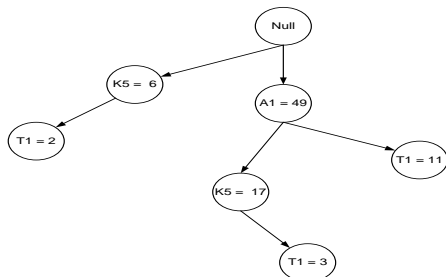
5. Menentukan *Frequent Itemset*.

Untuk menentukan *frequent itemset* dari headen *frequent itemset*, terlebih dahulu ditentukan lintasan yang berakhir dengan *support count* yang paling kecil. Nilai yang paling kecil adalah K3 serta diikuti oleh T1, T4, K5, dan A1. Tahap pembuatan masing-masing node disajikan pada Gambar 9.



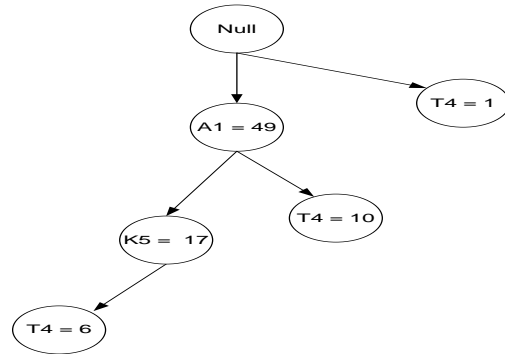
Gambar 10. Lintasan yang Mengandung Simpul K3

Gambar 10 merupakan penjelasan lintasan yang memiliki akhiran K3, dimana K3 tersebut memiliki jumlah terkecil dengan *frequent kemunculan* 15 kali.



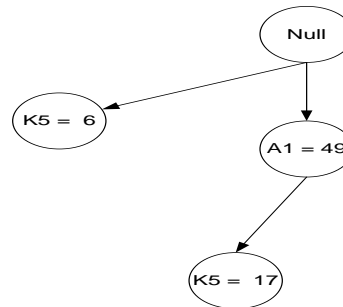
Gambar 11. Lintasan Yang Mengandung Simpul T1

Gambar 11 merupakan penjelasan dari lintasan yang berakhir T1 yang mempunyai *frequent kemunculan* sebanyak 16 kali.



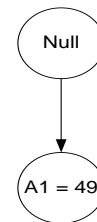
Gambar 12. Lintasan yang Mengandung Simpul T4

Gambar 12 merupakan penjelasan dari lintasan yang berakhir T4 yang mempunyai *frequent kemunculan* sebanyak 17 kali.



Gambar 13. Lintasan yang Mengandung Simpul K5

Gambar 13 merupakan penjelasan dari lintasan yang berakhir K5 yang mempunyai *frequent kemunculan* sebanyak 23 kali.



Gambar 14. Lintasan yang Mengandung Simpul A1

Gambar 14 merupakan penjelasan dari lintasan berakhir A1 yang mempunyai *frequent kemunculan* yang paling banyak dari seluruh lintasan yaitu sebanyak 49 kali.

6. Membuat *Conditional Pattern* dari Bentuk *FP-Tree*

Setelah mendapatkan *frequent itemset*, maka diakhiran *suffix* dirangkum pada tabel 2.

Tabel 2. Tabel *Suffix*

<i>Suffix</i>	<i>Frequent itemset</i>
K3	{A1,K3},{A1,T4,K3},{A1,T1,K3},{T4,K3}
T1	{A1,T1},{A1,K5,T1},{K5,T1}
T4	{T4},{A1,T4},{A1,K5,T4}
K5	{K5},{A1,K5}
A1	{A1}

7. Menghitung *Support* dan *Confidence*

Setelah membuat *Conditional Pattern* dari bentuk *FP-Tree* selanjutnya menghitung *support* dan *confidence* dari hasil pembentukan *FP-Tree*.

$$\text{support} = \frac{\sum \text{jumlah Kasus}}{\sum \text{jumlah seluruh kasus}} \times 100\%$$

Item pada kasus Korban Pelajar (A1), Kasus Cabul (K5), Tersangka Pengangguran (T4), Tersangka Pelajar (T1), dan Kasus Penganiayaan (K3) ada 5 dari 56 kasus sehingga nilai *Support* adalah $5/56 * 100\% = 9\%$.

$$\text{Confidence} = \frac{\sum \text{jumlah Kasus}}{\sum \text{jumlah Kasus Pada Antecedent}} \times 100\%$$

Item pada kasus korban pelajar (A1), kasus cabul (K5), tersangka pengangguran (T4), Tersangka Pelajar (T1), dan kasus penganiayaan (K3) ada 5 dan jumlah kasus pada *Antecedent* yang berarti nilai *support* adalah 9 sehingga nilai *confidence* adalah $5/9 * 100\% = 56\%$.

8. Hasil *Rule*

Setelah menghitung nilai *support* dan *confidence* maka didapatkan hasil *rule* yang disajikan pada tabel 3.

Tabel 3. Hasil *Rule*

<i>If antecedent then consequent</i>	<i>Support</i>	<i>Confidence</i>
If Penganiayaan (K3) Then Korban Pelajar (A1)	14/56=25%	14/15=93%
If Cabul (K5) Then Korban Pelajar (A1)	17/56=30%	17/23=74%
If Tersangka Pelajar (T1) Then Korban Pelajar (A1)	14/56=25%	14/16=88%
If Tersangka Pengangguran (T4) Then Korban Pelajar (A1)	16/56=29%	16/17=94%
If Tersangka Pelajar (T1) Penganiayaan (K3) Then Korban Pelajar (A1)	5/56=9%	5/5=100%
If Cabul (K5) Tersangka Pengangguran (T4) Then Korban Pelajar (A1)	6/56=11%	6/6=100%

4. Kesimpulan

Hubungan antara jenis tersangka, kasus, dan korban dengan menggunakan *Association Rule* dengan *Metode FP-Growth* menghasilkan *rule* dan pengetahuan baru

(*knowledge*) pada kasus kejahatan anak dibawah umur. Hasil pengujian penghitungan nilai *minimum support* $\geq 9\%$ dan *confidence* $\geq 56\%$ maka *rule* yang terpenuhi adalah jika kasus cabul dan tersangka pengangguran maka korbanya adalah pelajar dengan nilai *support* 11% dan *confidence* 100% . Sehingga penelitian ini dapat menentukan faktor-faktor kejahatan pada anak dengan tepat.

Daftar Rujukan

[1] Ikhwan, A., Nofriansyah, D., & Sriani. (2015). Penerapan Data Mining dengan Algoritma FP-Growth untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus Kampus STMIK Triguna Dharma). *Jurnal SAINTIKOM*, 15(3), 211-226.

[2] Meilani, B. D., & Azinar, A. W. (2015). Penentuan Pola Yang Sering Muncul Untuk Penerima Kartu Jaminan Kesehatan Masyarakat (JAMKESMAS) Menggunakan Metode Fp-Growth. *Seminar Nasional "Inovasi dalam Desain dan Teknologi"-IdeaTech 2015*.

[3] Rusdianan, D., & Setiyono, A. (2018). Algoritma FP-Growth Dalam Penempatan Lokasi Barang Di Gudang PT.XYZ. *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, 4(1), 63-70.

[4] Herasmus, H. (2017). Analisa Costumer Service System Menggunakan Metode Data Mining Dengan Algoritma FP-Growth (Studi Kasus Di PT Batamindo Investment Cakrawala). *Jurnal Teknik Ibnu Sina (JT-IBSI)*.

[5] Afdal, M. (2018). Implementasi Data Mining dalam Pencarian Daerah Strategis untuk Pengenalan Sekolah Swasta dengan Metode FP-Growth. *JURNAL INSTEK (Informatika Sains Dan Teknologi)*, 3(2).

[6] Maulana, A., & Fajrin, A. A. (2018). Penerapan Data Mining untuk Analisis Pola Pembelian Konsumen dengan Algoritma FP-Growth pada Data Transaksi Penjualan Spare Part Motor. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 5(1).

[7] Elwani, E. (2017). Penentuan Aturan Asosiasi pada Transaksi Peminjaman Buku Menggunakan Algoritma FP-Growth. *Jurnal Informatika, Manajemen Dan Komputer*, 9(1). DOI: <http://dx.doi.org/10.36723/juri.v9i1.59>.

[8] Fitriyani, F. (2015). Implementasi Algoritma FP-Growth Menggunakan Association Rule pada Market Basket Analysis. *Jurnal Informatika*, 2(1). DOI: <https://doi.org/10.31311/ji.v2i1.85>.

[9] Hilmy, N. F., & Andoko, B. S. (2016). Rancang Bangun Aplikasi Data Mining Analisis Tingkat Kelulusan Menggunakan Algoritma FP-Growth (Studi Kasus Di Politeknik Negeri Malang). *Jurnal Informatika Polinema*, 2(4). DOI: <https://doi.org/10.33795/jip.v2i4.76>.

[10] Fauzy, M., W, K. R. S., & Asror, I. (2016). Penerapan Metode Association Rule Menggunakan Algoritma Apriori pada Simulasi Prediksi Hujan Wilayah Kota Bandung. *Jurnal Ilmiah Teknologi Informasi Terapan (JITTER)*, 2(3).