

# Penerapan Data Mining Dalam Penentuan Pengambilan Semester Pendek Menggunakan Algoritma Fp-Growth (Studi Kasus Di Stmik Triguna Dharma Medan)

Milfa Yetri<sup>#1</sup>, Sarjon Devit<sup>#2</sup>, Gunadi Widi Nurcahyo<sup>#3</sup>

<sup>#1</sup> Program Studi Sistem Informasi, STMIK Triguna Dharma

<sup>#2,3</sup> Program Pascasarjan Magister Komputer, Universitas Putra Indonesia-YPTK Padang

---

## Article Info

### Article history:

Received Jan 13<sup>th</sup>, 2018

Revised Jan 22<sup>th</sup>, 2018

Accepted Jan 29<sup>th</sup>, 2018

---

### Keyword:

Biro Administrasi Akademik

Semester Pendek

FP-Growth Algorithm

---

## ABSTRACT

Semester pendek bertujuan untuk meningkatkan nilai siswa yang gagal di setiap semester. Sejauh ini dalam pelaksanaan semester pendek, otoritas akademik terkadang mengalami kesulitan dalam menentukan kelayakan siswa untuk mengikuti. Ini karena ada beberapa persyaratan akademis. Jadi khususnya di Biro Administrasi Akademik (BAAK) sempat meragukan proses penerimaan dan persetujuan aplikasi untuk semester pendek. Salah satu cara yang bisa dilakukan untuk mengetahui mana siswa yang bisa mengikuti semester pendek adalah dengan menggunakan teknik data mining. Teknik yang digunakan dalam kasus ini adalah FP-Growth Algorithm. FP-Growth adalah salah satu alternatif algoritma yang bisa digunakan untuk menentukan kumpulan data yang paling sering muncul (frequent item set) dalam satu set data. Algoritma FP-Growth merupakan perpanjangan dari algoritma Apriori. Penelitian ini dilakukan dengan mengamati beberapa variabel penelitian, yaitu tingkat kehadiran, nilai, administrasi, dan jumlah SP yang dipertimbangkan oleh perguruan tinggi khususnya Biro Administrasi Akademik (BAAK) dalam menentukan siswa mana yang dapat mengikuti Semester Pendek.

*Copyright © 2018 STMIK Triguna Dharma.  
All rights reserved.*

---

### First Author

Nama : Milfa Yetri, S.Kom, M.Kom

Kantor : STMIK Triguna Dharma

Program Studi : Sistem Informasi

E-Mail : airputih.girls@mail.com

---

## 1. PENDAHULUAN

STMIK Triguna Dharma merupakan suatu instansi pendidikan komputer di Sumatera Utara. Dalam pelaksanaan kegiatan operasional tidak lepas dengan hal-hal akademik. Di dalam kegiatan akademik terdapat elemen-elemen penilaian diantaranya adalah semester pendek. Semester Pendek merupakan kegiatan akademik yang ditujukan kepada mahasiswa STMIK Triguna Dharma yang memperoleh nilai gagal pada suatu mata kuliah di setiap semester.

Selama ini dalam pelaksanaan semester pendek, pihak akademik terkadang mengalami kesulitan dalam hal penentuan kelayakan mahasiswa yang dapat mengikutinya. Hal ini disebabkan karena ada beberapa persyaratan akademik yang banyak. Sehingga khususnya di Biro Administrasi Akademik Kemahasiswaan (BAAK) mengalami kebingungan di dalam proses penerimaan dan persetujuan terhadap permohonan semester pendek.

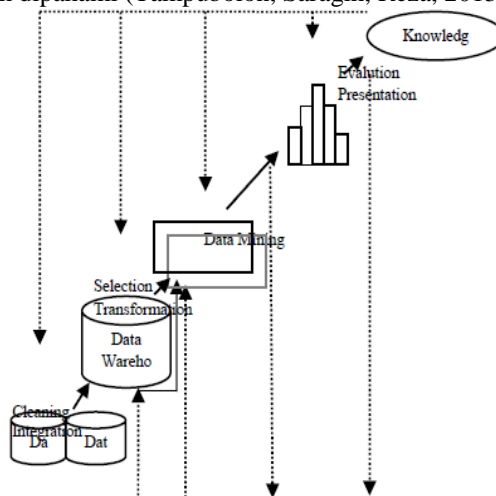
Analisa yang dilakukan nantinya akan mengadopsi Algoritma Fp-Growth yang terdiri atas beberapa elemen di antaranya adalah fitur untuk memproses kelayakan mahasiswa mengikuti semester pendek dengan Algoritma Fp-Growth beserta dengan laporan keputusannya.

## 2. METODE PENELITIAN

### 2.1 Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses Knowledge Discovery in Database melibatkan hasil proses

Data Mining (proses pengekstrak kecenderungan suatu pola data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami (Tampubolon, Saragih, Reza, 2013).



Gambar 1. Konsep Dasar Data Mining Dalam Proses KDD

## 2.2 Data Mining

Data Mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis. Definisi lain diantaranya adalah pembelajaran berbasis induksi (induction-based learning) adalah proses pembentukan definisi-definisi konsep umum yang dilakukan dengan cara mengobservasi contoh-contoh spesifik dari konsep-konsep yang akan dipelajari (Fajar Astuti Hermawati, 2013).

Istilah Data Mining lainnya memiliki beberapa padanan, seperti knowledge discovery ataupun pattern recognition. Kedua istilah tersebut sebenarnya memiliki ketepatannya masing-masing. Istilah knowledge discovery atau penemuan pengetahuan tepat digunakan karena tujuan utama dari Data Mining memang untuk mendapatkan pengetahuan yang masih tersembunyi di dalam bongkahan data. Istilah pattern recognition atau pengenalan pola pun tepat digunakan karena pengetahuan yang hendak digali memang berbentuk pola-pola yang mungkin juga masih perlu digali dari dalam bongkahan data yang tengah dihadapi (Sani Susanto dan Dedy Suryadi, 2010).

## 2.3 Algoritma FP-Growth

FP-growth adalah metode yang sering itemset pertambahan tanpa Generasi calon. Ini membangun sebuah struktur data yang sangat padat (FP-tree) untuk kompres database transaksi asli (Jiawei and Kamber Data Mining and Concept, 2006). Algoritma FP-Growth generate candidate tidak dilakukan karena FP-Growth menggunakan konsep pembangunan tree dalam pencarian frequent itemset. Berdasarkan analisis algoritma association rule di atas maka dalam penelitian ini algoritma FP-Growth akan digunakan untuk mempercepat proses penentuan frequent itemset sebelum men-generate rule sebagai rekomendasi keputusan.

Struktur data yang digunakan untuk mencari frequent itemset dengan algoritma FP-Growth adalah perluasan dari sebuah pohon prefix, yang biasa disebut FP-Tree. Dengan menggunakan FP-Tree, algoritma FP-Growth dapat langsung mengekstrak frequent itemset dari FP-Tree yang telah dibentuk dengan menggunakan prinsip divided and conquer.

### 2.3.1 Analisis Pola Frekuen Tinggi Dengan FP-Growth

Pada penentuan frequent itemset terdapat 2 tahap proses yang dilakukan yaitu: pembuatan FP-Tree dan penerapan algoritma FP-Growth untuk menemukan frequent itemset. Struktur data yang digunakan untuk mencari frequent itemset dengan algoritma FP-growth adalah perluasan dari penggunaan sebuah pohon prefix, yang biasa disebut adalah FP-tree. Dengan menggunakan FP-tree, algoritma FP-Growth dapat langsung mengekstrak frequent itemset dari FP-tree yang telah terbentuk. (Goldie, 2012).

### 2.3.2 Pembuatan FP-Tree

FP-tree merupakan struktur penyimpanan data yang dimanfaatkan. FP-tree dibangun dengan memetakan setiap data transaksi ke dalam setiap lintasan tertentu dalam FP-tree. Karena dalam setiap transaksi yang dipetakan, mungkin ada transaksi yang memiliki item yang sama, maka lintasannya memungkinkan untuk saling menimpa. Semakin banyak data transaksi yang memiliki item yang sama, maka proses pemampatan dengan struktur data FP-tree semakin efektif. (Goldie, 2012).

Setelah tahap pembangunan FP-tree dari sekumpulan data transaksi, akan diterapkan algoritma FP-growth untuk mencari frequent itemset yang signifikan. Rama et all (2014). Algoritma FP-growth dibagi menjadi tiga langkah utama, yaitu:

1. Tahap Pembangkitan Conditional Pattern Base
2. Tahap Pembangkitan Conditional FP-tree
3. Tahap Pencarian Frequent Itemset

### 2.3.3 Bentuk Algoritma FP-Growth

Bentuk algoritma FP-Growth sebagai berikut:

Input : FP-Tree Tree

Output : Rt Sekumpulan lengkap pola frequent

Proses : FP-Growth (Tree,  $\alpha$ )

```

If tree mengandung single path p Then untuk tiap kombinasi
(dinotasikan  $\beta$ ) dari node-node dalam  $\beta$  Else untuk tiap a1 dalam header dari tree
Do
  {
    Bangkitkan pola Bangun  $\beta = a1 \alpha$  dengan support = a1
    support
  }
If tree  $\beta = \theta$ 
  Then panggil FP-Growth (Tree,  $\beta$ )
}}
```

## 3. ANALISIS DAN HASIL

### 3.1. Analisis

Data yang akan digunakan dalam membentuk pohon keputusan untuk menganalisis kelayakan mahasiswa mengikuti semester pendek. Data semester pendek tersebut selanjutnya akan dilakukan pra-proses untuk menghasilkan data kasus yang siap untuk dibentuk menjadi sebuah pohon keputusan. Karakteristik algoritma FP-Growth adalah struktur data yang digunakan adalah tree yang disebut dengan FP-Tree. Dengan menggunakan FP-Tree, algoritma FP-growth dapat langsung mengekstrak frequent Itemset dari FP-Tree. Penggalan itemset yang frequent dengan menggunakan algoritma FP-Growth akan dilakukan dengan cara membangkitkan struktur data tree.

Metode FP-Growth dapat dibagi menjadi 3 tahapan utama yaitu sebagai :

1. Tahap pembangkitan conditional pattern
2. Tahap pembangkitan conditional FP-Tree, dan
3. Tahap pencarian frequent itemset.

Ketiga tahap tersebut merupakan langkah yang akan dilakukan untuk mendapat frequent itemset.

Input : FP-Tree Tree

Output : Rt Sekumpulan lengkap pola frequent

Methode: FP-Growth (Tree, null)

Procedure : FP-Growth (Tree,  $\alpha$ )

```

{
  01: if Tree mengandung single path P;
  02: then untuk tiap kombinasi (dinotasikan  $\beta$ ) dari node-node dalam path do
  03: bangkitkan pola  $\beta \alpha$  dengan support dari node-node dalam path do  $\beta$ ;
  04: else untuk tiap a1 dalam header dari tree do
  }
  05: bangkitkan pola
  06: bangun  $\beta = a1 \alpha$  dengan support = a1 support
  07: if Tree  $\beta =$ 
```

Pada penulisan ini, menggunakan proses asosiasi dalam menganalisa mahasiswa yang hendak mengikuti semester pendek yang akan diproses dengan pembangunan FP-Tree dalam pencarian frequent itemsets untuk menghasilkan suatu output. Data data tersebut mempunyai beberapa kriteria yang merupakan syarat dalam pembangunan FP-Tree.

Kumpulan frequent itemsets dicari dengan melakukan scanning database kemudian hasilnya diakumulasikan dengan tiap items lainnya dan dikelompokkan untuk mendapatkan frequent itemsets yang memenuhi nilai minimum support. Setelah didapatkan frequent itemsets, lalu frequent itemsets yang memiliki nilai kurang dari minimum support akan dipangkas atau dibuang. Sehingga akan didapatkan frequent itemsets yang memenuhi minimum support dan disusun dengan urutan menurun.

Untuk mencari pola hubungan daerah yang strategis dengan beberapa atribut pada data siswa dengan menggunakan algoritma frequent pattern growth (FP-Growth) melalui empat tahapan data mining yaitu selection, preprocessing, transformation dan proses mining.

#### 3.1.1 Selction

Menfilter data Absensi, kemudian didapat hasil filter data absen maka dilakukan filter data nilai. Kemudian melakukan pemilihan atribut yang akan digunakan pada tahapan selanjutnya yaitu administrasi dan jumlah SP dari 8 atribut yang tersedia pada data mahasiswa STMIK TRIGUNA DHARMA MEDAN

### 3.1.2 Preprocessing

Pada tahapan ini dilakukan pembersihan data terhadap data yang mengandung nilai yang kosong (missing value) dan nilai yang tidak konsisten (inconsistent), jika jumlah data yang digunakan besar maka proses ini bisa diabaikan (dihapus) dan jika jumlah data yang digunakan kecil atau terbatas, maka data yang hilang bisa digantikan dengan nilai rata-rata dari atribut yang berkaitan.

### 3.1.3 Selection (Iterasi)

Setelah didapatkan hasil dari tahapan preprocessing maka dilakukan tahapan selection (iterasi) untuk memilih atribut yang hanya digunakan pada tahapan selanjutnya.

### 3.1.4 Transformation

Transformasi data dilakukan sesuai dengan kebutuhan dan penyesuaian dengan aplikasi (perangkat lunak) yang digunakan dalam implementasi proses mining, langkah pertama yang dilakukan yaitu melakukan inisialisasi atribut dengan tujuan agar mudah dipahami dan mempermudah dalam implementasi pada aplikasi yang digunakan. Selanjutnya melakukan konversi data menjadi tabular data dengan tujuan agar bisa diimplementasi pada aplikasi yang digunakan yaitu Rapidminer 5.2.

### 3.1.5 Proses Mining Dengan Algoritma FP-Growth

Berdasarkan data atribut yang telah didapatkan pada proses sebelumnya, selanjutnya bagaimana penggunaan Algoritma FP-Growth untuk mendapatkan frequent itemset yang memenuhi minimum support dan minimum confidence dari data hasil transformation. Nilai minimum support dan minimum confidence diperoleh dari the best rule karena proses penentuan minimum support dan minimum confidence tidak bisa terlalu rendah dan tinggi. Jika menentukan nilai yang terlalu rendah maka nilai the best rule akan banyak muncul. Begitu juga sebaliknya, jika nilai support dan confidence terlalu tinggi maka nilai the best rule tidak akan muncul.

Tabel 1. Dataset Mahasiswa STMIK TRIGUNA DHARMA MEDAN

No	Absensi	Nilai	Administrasi	Jumlah SP
1	CK	E	SL	A5
2	CK	D	SL	A5
3	CK	D	SL	A1
4	CK	C	SL	A5
5	TC	C	SL	A5
6	CK	C	SL	A5
7	CK	E	SL	A4
8	CK	C	SL	A1
9	CK	E	SL	A4
10	CK	C	SL	A1
11	CK	C	SL	A5
12	CK	C	SL	A5
13	CK	C	SL	A1
14	CK	E	SL	A4
15	CK	C	SL	A1
16	CK	C	SL	A3
17	CK	D	SL	A1
18	CK	C	SL	A2
19	TC	D	SL	A2
20	TC	D	SL	A3

1. Dilakukan perhitungan terhadap frekuensi kemunculan tiap item dan nilai support menggunakan rumus

Tabel 2. Frekuensi dan support tiap item

No	Item	Frekuensi	Support	Support (%)
1	SL	20	$20/20 = 1$	100%
2	CK	17	$17/20 = 0.85$	85%
3	C	11	$11/20 = 0.55$	55%
4	A5	7	$7/20 = 0.35$	35%
5	A1	6	$6/20 = 0.30$	30%
6	D	5	$5/20 = 0.25$	25%
7	E	4	$4/20 = 0.20$	20%
8	TC	3	$3/20 = 0.15$	15%
9	A4	3	$3/20 = 0.15$	15%
10	A2	2	$2/20 = 0.10$	10%
11	A3	2	$2/20 = 0.10$	10%
12	TL	0	$0/20 = 0$	0%

2. Menentukan nilai support count. Pada penelitian ini diambil nilai support count = 20%. Nilai support count akan mempengaruhi item yang akan dianalisa ke tahap pembuatan FP-Tree. Berdasarkan nilai support count 20%, maka item yang akan digunakan adalah item yang memiliki frekuensi di  $\geq 20\%$  yaitu pada table 4.6 sedangkan item TC, A4, A2, A3, dan TL tidak memenuhi minimum support dengan nilai support 15%, 15%, 10%, 10%, dan 0% sehingga bisa dibuang.

Tabel 3. Item Memenuhi Minimum Support

Item	Frekuensi	Support	Support (%)
SL	20	$20/20 = 1$	100%
CK	17	$17/20 = 0.85$	85%
C	11	$11/20 = 0.55$	55%
A5	7	$7/20 = 0.35$	35%
A1	6	$6/20 = 0.30$	30%
D	5	$5/20 = 0.25$	25%
E	4	$4/20 = 0.20$	20%

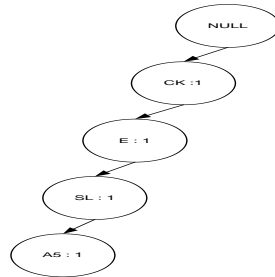
3. Melakukan pemindaian data atau menyusun kembali data yang memenuhi minimum support. Data diurutkan berdasarkan nilai frekuensi yang paling tinggi.

Tabel 4. Data Transaksi

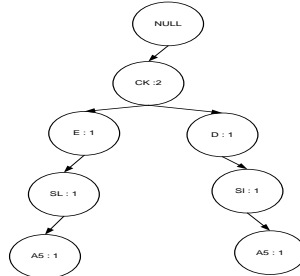
No	SL, CK, C, A5, A1, D, E
1	CK, E, SL, A5
2	CK, D, SL, A5
3	CK, D, SL, A1
4	CK, C, SL, A5
5	TC, C, SL, A5
6	CK, C, SL, A5
7	CK, E, SL
8	CK, C, SL, A1
9	CK, E, SL
10	CK, C, SL, A1

11	CK, C, SL, A5
12	CK, C, SL, A5
13	CK, C, SL, A1
14	CK, E, SL
15	CK, C, SL, A1
16	CK, C, SL
17	CK, D, SL, A1
18	CK, C, SL
19	TC, D, SL
20	TC, D, SL

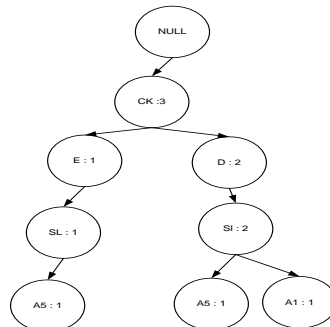
4. Pembangunan *FP-tree* dari data transaksi yang telah ada.



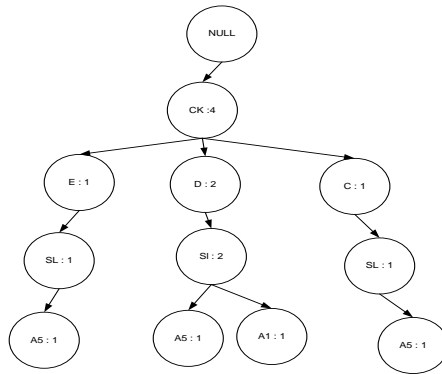
Gambar 2. Hasil Pembentukan *FP-Tree* TID 1 Setelah Pembacaan CK,E,SL,A5



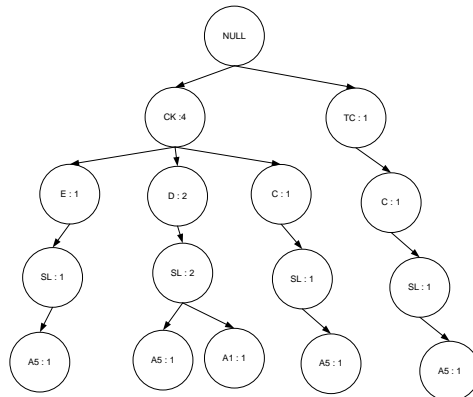
Gambar 3. Hasil Pembentukan *FP-Tree* TID 2 Setelah Pembacaan CK,D,SL,A5



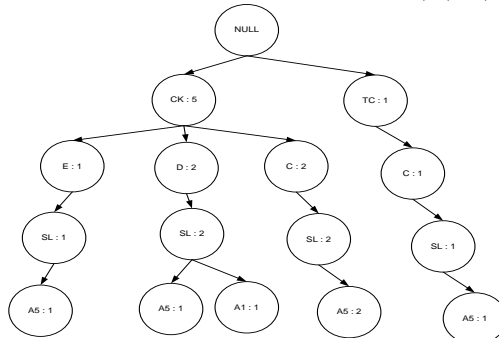
Gambar 4. Hasil Pembentukan *FP-Tree* TID 3 Setelah Pembacaan CK,D,SL,A1



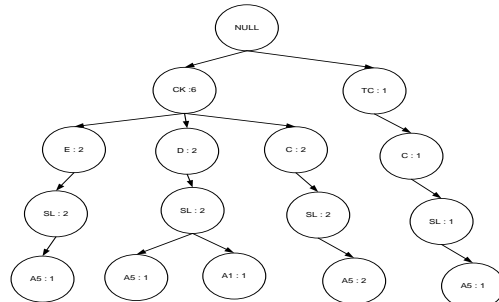
Gambar 5. Hasil Pembentukan *FP-Tree* TID 4 Setelah Pembacaan CK,C,SL,A5



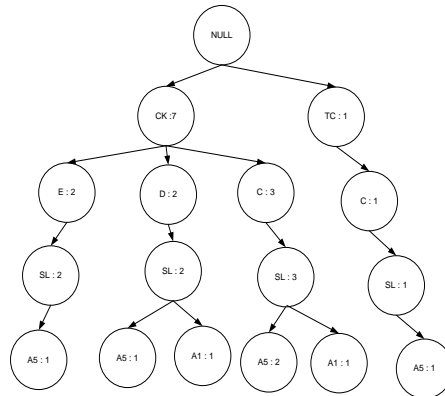
Gambar 6. Hasil Pembentukan *FP-Tree* TID 5 Setelah Pembacaan TC,C,SL,A5



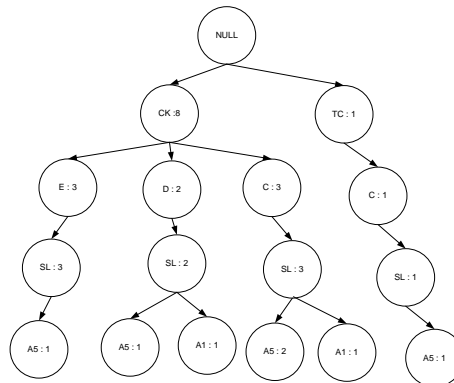
Gambar 7. Hasil Pembentukan *FP-Tree* TID 6 Setelah Pembacaan CK,C,SL,A5



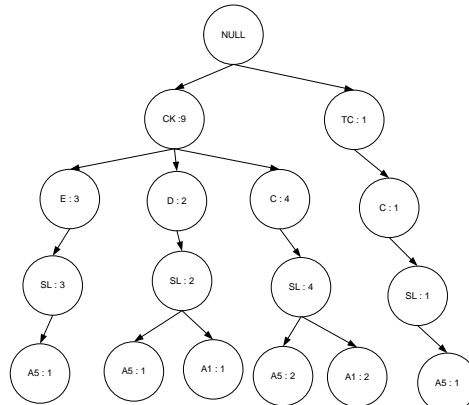
Gambar 8. Hasil Pembentukan *FP-Tree* TID 7 Setelah Pembacaan CK,E, SL



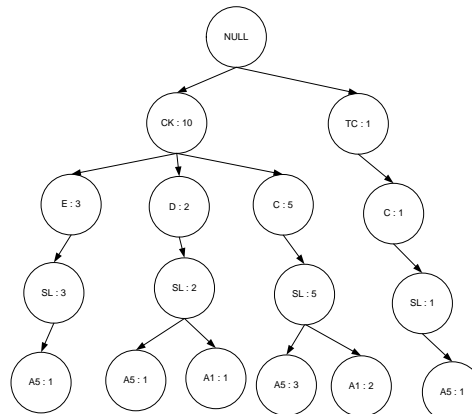
Gambar 9. Hasil Pembentukan *FP-Tree* TID 8 Setelah Pembacaan CK,C,SL, A1



Gambar 10. Hasil Pembentukan *FP-Tree* TID 9 Setelah Pembacaan CK, E, SL

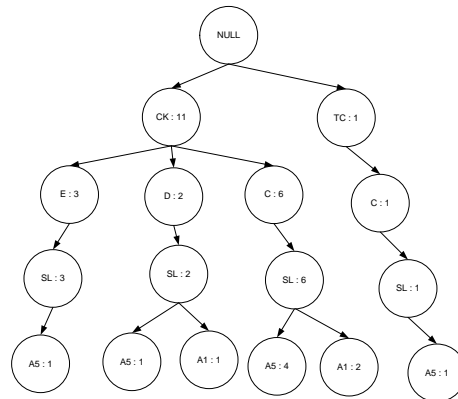


Gambar 11. Hasil Pembentukan *FP-Tree* TID 10 Setelah Pembacaan CK, C, SL, A1

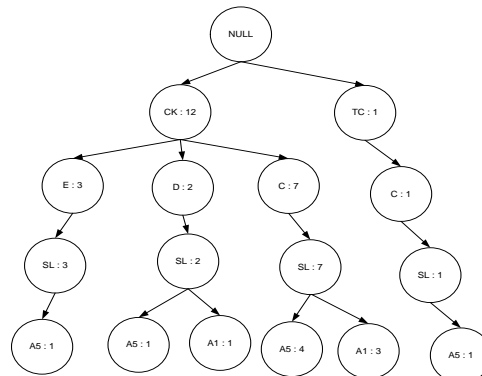


Gambar 12. Hasil Pembentukan *FP-Tree* TID 11 Setelah Pembacaan CK, C, SL, A5

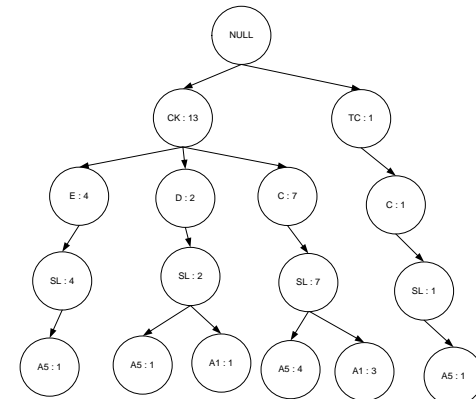




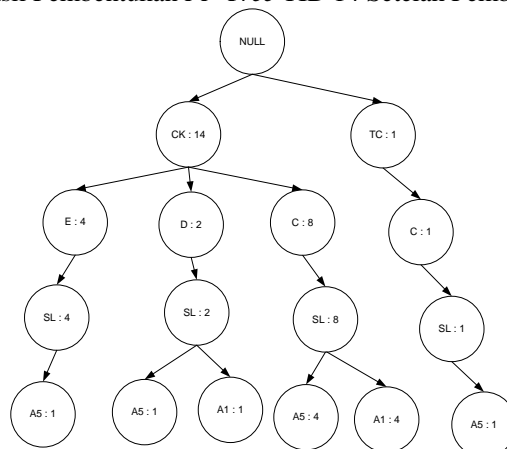
Gambar 13. Hasil Pembentukan *FP-Tree* TID 12 Setelah Pembacaan CK, C, SL, A5



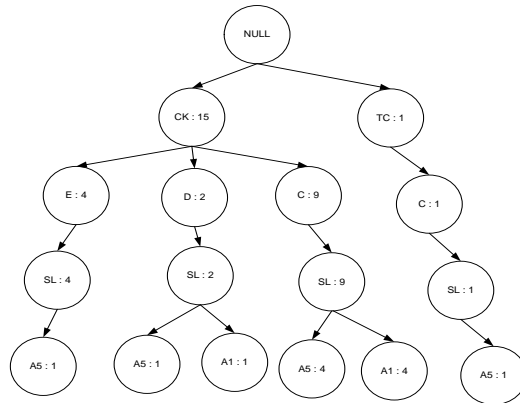
Gambar 14. Hasil Pembentukan *FP-Tree* TID 13 Setelah Pembacaan CK, C, SL, A1



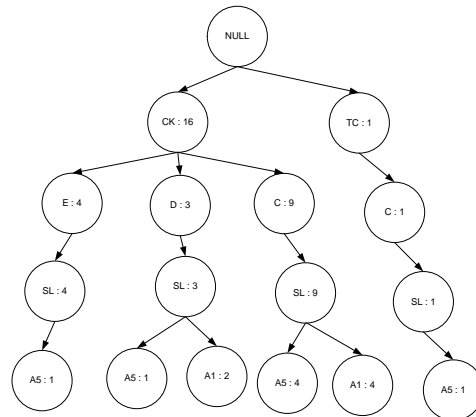
Gambar 15. Hasil Pembentukan *FP-Tree* TID 14 Setelah Pembacaan CK, E, SL



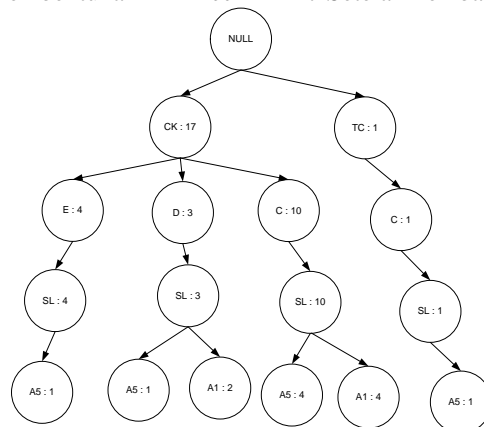
Gambar 16. Hasil Pembentukan *FP-Tree* TID 15 Setelah Pembacaan CK, C, SL, A1



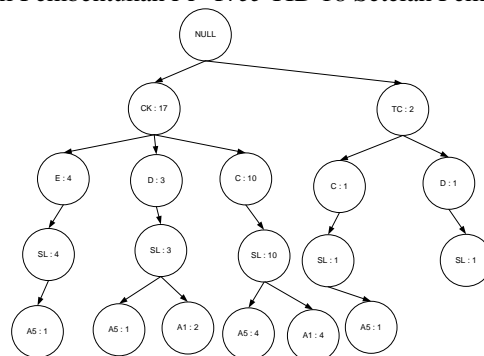
Gambar 17. Hasil Pembentukan *FP-Tree* TID 16 Setelah Pembacaan CK, C, SL



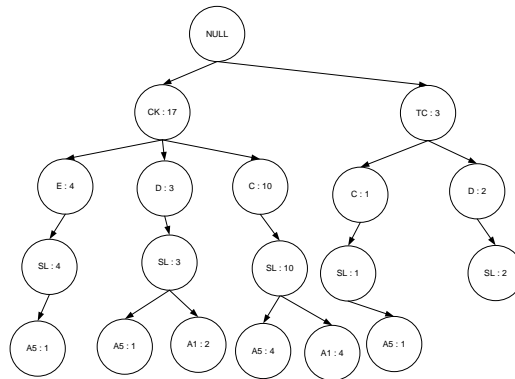
Gambar 18. Hasil Pembentukan *FP-Tree* TID 17 Setelah Pembacaan CK, D, SL, A1



Gambar 19. Hasil Pembentukan *FP-Tree* TID 18 Setelah Pembacaan CK, C, SL



Gambar 20. Hasil Pembentukan *FP-Tree* TID 19 Setelah Pembacaan TC, D, SL



Gambar 21. Hasil Pembentukan *FP-Tree* TID 20 Setelah Pembacaan TC, D, SL

5. Penerapan *FP-Growth* untuk mencari *frequent itemset*. Pada penelitian ini mencari hubungan antar atribut dalam menentukan daerah strategis.
  - a. Pembangkitan *conditional pattern base*  
 Pembangkitan *conditional pattern base* dilakukan dengan cara menentukan *upapohon* dengan lintasan yang berakhiran dengan nilai *support* terkecil yaitu E, D, A1, A5, C, CK, SL.
  - b. Pembangkitan *Condition FP-Tree*  
*Conditional FP-Tree* dapat dibentuk dengan cara:
    1. *Support count* pada semua item pada setiap *conditional pattern base* dibandingkan dengan nilai *conditional* berdasarkan nilai *minimum support* yang ditentukan yaitu 20%. Item A1 memiliki nilai *support* 30% sehingga item A1 besar dari nilai *minimum support* Sehingga subproblem selanjutnya adalah menemukan *frequent itemset* pada item yang berakhir dengan (A1,SL), (A1,C) dan (A1,CK) dengan membentuk *conditional FP-tree* dengan cara:
      - a. Membuang setiap lintasan yang tidak mengandung A1. Pada lintasan yang berakhir di A1 terdapat lintasan yang tidak mengandung simpul A1 yaitu null→CK→D→C→SL.
      - b. Semua lintasan berakhir di A1, maka simpul A1 dapat dibuang. Seperti pada gambar 4.30. Permasalahan selanjutnya menemukan lintasan *frequent itemset* yang berakhir di (A1,SL), (A1,C) dan (A1,CK).
      - c. Berdasarkan *conditional FP-tree* dari A1 didapat bahwa CK,D,C dan SL memnuhi *minimum support* sehingga (A1, (A1,C), (A1,CK), (A1,SL,CK), (A1,C,CK), dan (A1,SL,C,CK) adalah *frequent itemset*.

Tabel 5. Hasil *Frequent Itemset*

Suffix	Frequent Itemset
E	(E), (E, CK)
A1	(A1), (A1, SL), (A1, CK), (A1, SL, CK), (A1, C, CK), (A1, SL, C, CK)
A5	(A5), (A5, CK), (A5, E), (A5, SL), (A5, C), (A5, SL, E), (A5, SL, CK), (SL, E, CK), (A5, SL, C), (SL, C, CK), (A5, SL, E, CK), (A5, SL, C, CK)
C	(C), (C, CK)
CK	(CK)
SL	(SL), (SL, CK), (SL, E), (SL, C), (SL, D), (SL, E, CK), (SL, C, CK), (SL, D,CK)

Setelah didapat *frequent itemset*, selanjutnya membuat *rules* dengan cara menghitung *confidence* dari tiap kombinasi *rule*. Dari 31 *itemset* yang dihasilkan pada table 3.5 tidak semua dihitung karena *rule* yang dihasilkan adalah jika A maka B, sehingga *itemset* yang dihitung minimal berisi dua *item*. Maka yang dihitung *confidence*-nya adalah 25 *subsets*, yaitu : (E, CK), (A1, SL), (A1, CK), (A1, SL, CK), (A1, C, CK), (A1, SL, C, CK), (A5, CK), (A5, E), (A5, SL), (A5, C), (A5, SL, E), (A5, SL, CK), (SL, E, CK), (A5, SL, C), (SL, C, CK), (A5, SL, E, CK), (A5, SL, C, CK), (C, CK), (SL, CK), (SL, E), (SL, C), (SL, D), (SL, E, CK), (SL, C, CK), (SL, D,CK)

Hanya kombinasi yang lebih besar atau sama dengan *minimum confidence* yang akan diambil atau *strong association rule* saja. Rumus untuk menghitung *confidence* menggunakan rumus 2.3. Berikut perhitungan *confidence* dari 25 *subset* yaitu (A1, CK, SL) :

$$A1 \wedge SL \rightarrow CK = \frac{6}{6} = 100\% \quad Support\ \% \frac{6}{20} = 30\%$$

$$\begin{aligned}
 CK \rightarrow A1 \wedge SL &= \frac{6}{17} = 35\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 A1 \wedge CK \rightarrow SL &= \frac{6}{6} = 100\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 A1 \rightarrow SL \wedge CK &= \frac{6}{6} = 100\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 SL \wedge CK \rightarrow A1 &= \frac{6}{17} = 0.35\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 SL \rightarrow A1 \wedge CK &= \frac{6}{20} = 30\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 A1 \rightarrow SL &= \frac{6}{20} = 30\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 SL \rightarrow A1 &= \frac{6}{20} = 30\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 A1 \rightarrow CK &= \frac{6}{6} = 100\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 CK \rightarrow A1 &= \frac{6}{17} = 35\% \quad \text{Support \% } \frac{6}{20} = 30\% \\
 SL \rightarrow CK &= \frac{17}{20} = 85\% \quad \text{Support \% } \frac{17}{20} = 85\% \\
 CK \rightarrow SL &= \frac{17}{20} = 85\% \quad \text{Support \% } \frac{17}{20} = 85\%
 \end{aligned}$$

Jika dimisalkan *minimum confidence* adalah 50% maka dari 20 data penelitian ini yang termasuk *strong association rule* adalah  $SL \rightarrow CK$  dan  $CK \rightarrow SL$ . Berikut tabel yang berisi hasil pola-pola yang dihasilkan atau *rules* yang dihasilkan :

Tabel 6. *Strong Association Rule*

NO	JIKA	MAKA	SUPP ORT %	CONFID ENCE %
1	SL	CK^C	50 %	50 %
2	SL	C	55 %	55 %
3	A5	CK^C	20 %	57 %
4	A5	SL^CK^C	20 %	57 %
5	SL^A5	CK^C	20 %	57 %
6	CK	C	50 %	59 %
7	CK	SL^C	50 %	59 %
8	SL^CK	C	50 %	59 %
9	D	CK	15 %	60 %
10	D	SL^CK	15 %	60 %
11	SL^D	CK	15 %	60 %
12	A1	C	20 %	67 %
13	A1	SL^C	20 %	67 %
14	SL^A1	C	20 %	67 %
15	CK^A5	C	20 %	67 %
16	A1	CK^C	20 %	67 %
17	CK^A1	C	20 %	67 %
18	CK^D	A1	10 %	67 %
19	CK^A5	SL^C	20 %	67 %
20	SL^CK ^A5	C	20 %	67 %
21	A1	SL^CK^C	20 %	67 %
22	SL^A1	CK^C	20 %	67 %
23	CK^A1	SL^C	20 %	67 %
24	SL^CK ^A1	C	20 %	67 %
25	CK^D	SL^A1	10 %	67 %

26	SL <sup>CK</sup> <sup>D</sup>	A1	10 %	67 %
27	A5	C	25 %	71 %
28	A5	SL <sup>C</sup>	25 %	71 %
29	SL <sup>A5</sup>	C	25 %	71 %
30	C <sup>A5</sup>	CK	20 %	80 %
31	C <sup>A5</sup>	SL <sup>CK</sup>	20 %	80 %
32	SL <sup>C</sup> <sup>A5</sup>	CK	20 %	80 %
33	SL	CK	85 %	85 %
34	A5	CK	30 %	86 %
35	A5	SL <sup>CK</sup>	30 %	86 %
36	SL <sup>A5</sup>	CK	30 %	86 %
37	C	CK	50 %	91 %
38	C	SL <sup>CK</sup>	50 %	91 %
39	SL <sup>C</sup>	CK	50 %	91 %
40	CK	SL	85 %	100 %
<b>NO</b>	<b>JIKA</b>	<b>MAKA</b>	<b>SUPP ORT %</b>	<b>CONFID ENCE %</b>
41	C	SL	55 %	100 %
42	A5	SL	35 %	100 %
43	A1	SL	30 %	100 %
44	D	SL	25 %	100 %
45	E	SL	20 %	100 %
46	A1	CK	30 %	100 %
47	E	CK	20 %	100 %
48	CK <sup>C</sup>	SL	50 %	100 %
49	CK <sup>A5</sup>	SL	30 %	100 %
50	A1	SL <sup>CK</sup>	30 %	100 %
51	SL <sup>A1</sup>	CK	30 %	100 %
52	CK <sup>A1</sup>	SL	30 %	100 %
53	CK <sup>D</sup>	SL	15 %	100 %
54	E	SL <sup>CK</sup>	20 %	100 %
55	SL <sup>E</sup>	CK	20 %	100 %
56	CK <sup>E</sup>	SL	20 %	100 %
57	C <sup>A5</sup>	SL	25 %	100 %
58	C <sup>A1</sup>	SL	20 %	100 %
59	A5 <sup>D</sup>	SL	5 %	100 %
60	A5 <sup>E</sup>	SL	5 %	100 %
61	A1 <sup>D</sup>	SL	10 %	100 %
62	C <sup>A1</sup>	CK	20 %	100 %
63	A5 <sup>D</sup>	CK	5 %	100 %
64	A5 <sup>E</sup>	CK	5 %	100 %
65	A1 <sup>D</sup>	CK	10 %	100 %

66	CK <sup>^</sup> C <sup>^</sup> A5	SL	20 %	100 %
67	C <sup>^</sup> A1	SL <sup>^</sup> CK	20 %	100 %
68	SL <sup>^</sup> C <sup>^</sup> A1	CK	20 %	100 %
69	CK <sup>^</sup> C <sup>^</sup> A1	SL	20 %	100 %
70	A5 <sup>^</sup> D	SL <sup>^</sup> CK	5 %	100 %
71	SL <sup>^</sup> A5 <sup>^</sup> D	CK	5 %	100 %
72	CK <sup>^</sup> A5 <sup>^</sup> D	SL	5 %	100 %
73	A5 <sup>^</sup> E	SL <sup>^</sup> CK	5 %	100 %
74	SL <sup>^</sup> A5 <sup>^</sup> E	CK	5 %	100 %
75	CK <sup>^</sup> A5 <sup>^</sup> E	SL	5 %	100 %
76	A1 <sup>^</sup> D	SL, CK	10 %	100 %
77	SL <sup>^</sup> A1 <sup>^</sup> D	CK	10 %	100 %
78	CK <sup>^</sup> A1 <sup>^</sup> D	SL	10 %	100 %

Setelah didapat *rules* dari *frequent itemset*, maka langkah selanjutnya adalah mengevaluasi kuat tidaknya sebuah aturan asosiasi yang didapatkan dengan menggunakan *lift ratio*. Jika nilai *lift ratio* lebih besar atau sama dengan 1 maka aturan tersebut bisa digunakan. Semakin tinggi nilai *lift ratio* dari sebuah *rule* maka semakin besar kekuatan asosiasinya.

Berdasarkan perhitungan *lift ratio* dari tiap *rule* yang didapatkan maka didapatkan 6 *rules* yang nilai *lift ratio* di bawah 1 yang artinya keenam *rules* tersebut tidak bisa digunakan. Setelah didapatkan *strong association rules* yang memenuhi nilai *lift ratio* maka dilakukan pemilihan *rule* yang memenuhi ketentuan pada penelitian ini yaitu hubungan antara absensi, nilai, administrasi, dan jumlah SP. Sehingga terdapat 72 *rules* yang memenuhi yang bisa mengambil semester pendek (SP) yaitu yang nilai *lift ratio* 1 keatas. Sebagai contoh, untuk no. 1 pada tabel 3.6, terdapat SL → CK<sup>^</sup>C yang dapat diartikan jika sudah membayar administrasi (SL), absensi cukup (CK) kemudian nilai yang mau di semester pendek adalah C maka mahasiswa berhak mengikuti semester pendek.

#### 4. KESIMPULAN

Dari yang sudah dibahas sebelumnya dapat ditarik kesimpulan :

1. Metode dalam pencarian *Frequent Itemset* pohon keputusan menggunakan algoritma *FP-Growth* bekerja sangat baik dalam melakukan *Frequent Itemset* dengan proses pembentukan *FP-Tree* dengan menghasilkan *rule* dari data sampel mahasiswa baru.
2. Penentuan data variabel sangat menentukan tingkat akurasi *FP-Growth* yang dibuat dan besarnya presentase dalam menentukan *minimum support* dan *minimum confidence* dipengaruhi oleh data variabel yang digunakan untuk mencari *frequent itemset* yang saling berhubungan untuk menemukan data variabel yang akan menentukan mahasiswa/i yang bisa mengikuti Semester Pendek.
3. Dari penelitian yang dilakukan ada beberapa atribut yang tidak digunakan dalam *rule* yang dihasilkan, sehingga pemilihan atribut di dalam *dataset* sangat penting.
4. Algoritma *FP-Growth* dapat diterapkan untuk mendukung menentukan mahasiswa yang bisa mengikuti Semester Pendek pada Perguruan Tinggi. Adapun Informasi yang berkaitan dengan pelaksanaan SP dapat tersedia dengan cepat.
5. Pelaksanaan Semester Pendek pada perguruan tinggi sangat dibantu dengan adanya penerapan algoritma *FP-Growth* ini sehingga diharapkan efektifitas pelaksanaan SP akan dapat semakin ditingkatkan.
6. Untuk menemukan *rule* atau *knowledge* yang sangat berarti dalam menganalisa strategi dalam menentukan Semester Pendek, dibutuhkan banyak data sampel yang digunakan untuk proses ekstraksi *rule* atau *knowledge* tersebut.

#### DAFTAR PUSTAKA

- [1] Alghamdi, A. S. A. (2011). "Efficient Implementation of FP Growth Algorithm-Data Mining on Medical Data." *International Journal of Computer Science and Network Security* 11(12): 7-16.

- [2] Buulolo, E. (2013). "Implementasi Algoritma Apriori pada Sistem Persediaan Obat (Studi Kasus: Apotik Rumah Sakit Estomih Medan)." *Pelita Informatika Budi Darma*: 1-13.
- [3] D. Magdalene Delight Angeline.,(2013). "Association Rule Generation For Student Performance Analysis Using Apriori Algorithm". *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA)*. 1 (1). 12-16.
- [4] Efori Buulolo.,(2013). "Implementasi Algoritma Apriori Pada Sistem Persediaan Obat (Studi Kasus : Apotik Rumah Sakit Estomih Medan)". *Pelita Informatika Budi Darma*. IV (1). 71-83.
- [5] Eko Wahyu Tyas D.,(2008). "Penerapan Metode Association Rule Menggunakan Algoritma Apriori Untuk Analisa Pola Data Hasil Tangkapan Ikan." *E-Indonesia Initiative* 2008.
- [6] Fu, D., S. Zhou, et al. (2009). *The Design and Implementation of a Distributed Network Intrusion Detection System Based on Data Mining*. *Software Engineering, 2009. WCSE'09. WRI World Congress on, IEEE*.
- [7] Goldie Gunadi., dan Dana Indra Sensuse.,(2012). "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : Studi Kasus Percetakan PT. Gramedia." *Jurnal TELEMATIKA MKOM*. 1 (1). 118-132
- [8] Hermawati, F. A. F., Immatul (2014). *Association-Rule Analysis Application For Course Scheduling*. *Seminar on Intelligent Technology and Its Applications 2014*.
- [9] Kennedy, T., Hoga, S. dan Bobby, R.,(2013). "Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan." *Informasi dan Teknologi Ilmiah (INTI)*. 1 (1). 93-106.
- [10] Kumar, B. S. and K. Rukmani (2010). "Implementation of web usage mining using APRIORI and FP growth algorithms." *Int. J. of Advanced Networking and Applications* 1(06): 400-404.
- [11] Miraldi, R. N., A. R. Christanto, et al. (2015). "Implementasi Algoritma FP-Growth untuk Sistem Rekomendasi Buku di Perpustakaan UKDW." *Jurnal Informatika* 10(1).
- [12] Padmavathi, K. and R. A. Kirithika "Performance Based Study of Association Rule Algorithms On Voter DB."
- [13] Rama Novta Rinaldi., Antonius Rachmat., dan Budi Santoso.,(2014). "Implementasi Algoritma FP-GROWTH untuk Sistem Rekomendasi Buku di Perpustakaan UKDW." *Jurnal INFORMATIKA* . 10 (1). 29-39.
- [14] Ririanti.,(2014). "Implementasi Algoritma Fp-Growth Pada Aplikasi Prediksi Persediaan Sepeda Motor (Studi Kasus Pt. Pilar Deli Labumas)." *Pelita Informatika Budi Darma*. VI (1). 139-144.
- [15] Sensuse, G. G. D. I. (2012). "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : Studi Kasus Percetakan Pt. Gramedia." *Jurnal TELEMATIKA MKOM* 4.
- [16] Tampubolon, K., H. Saragih, et al. (2013). "Implementasi Data Mining Algoritma Apriori pada Sistem Persediaan Alat-alat Kesehatan." *Majalah Ilmiah*: 1-14.
- [17] Yaohui Zeng., Zijun Zhang., and Andrew Kusiak.,(2015). "Predictive modeling and optimization of a multi-zone HVAC system with data mining and firefly algorithms." *Journal Elsevier Energy* (86). 393-402.
- [18] Yuli Asriningtias dan Rodhyah Mardhiyah.,(2014). "Aplikasi Data Mining Untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa." *Jurnal Informatika* (8).

## BIOGRAFI PENULIS



**Milfa Yetri, S.Kom., M.Kom.**, Perempuan kelahiran Medan 09 Maret 1988 Saat ini menduduki Jabatan Sekretaris Program Studi Manajemen Informatika, beberapa mata kuliah diampu diantaranya : Teknik Presentasi, Arsitektur Komputer, Animasi, Pengantar Teknologi Informasi, Perakitan dan Perawatan Komputer, tamat Strata 1 Program Studi Sistem Informasi STMIK Triguna Dharma, tamat Strata 2 Magister Komputer Universitas Putra Indonesia (UPI-YPTK) Padang