

**PENERAPAN ALGORITMA APRIORI DAN ALGORITMA
FP-GROWTH DALAM MENEMUKAN HUBUNGAN DATA
NILAI IJAZAH MATEMATIKA DAN BAHASA INGGRIS
DENGAN NILAI MATA PELAJARAN PEMROGRAMAN
DAN WEB PROGRAMMING
(STUDI KASUS SMK SANDIKTA KELAS X BEKASI)**

MAULANA

mul_spider@yahoo.com

Program Studi Teknik Informatika Universitas Bina Insani

ABSTRACT

In comparing objects to analyze the relationship between objects, in computer technique, we can use the methods called data mining. This system is very familiar created with many steps. In this research, algorithm ways used is algoritmaprioryandalgorithm FP-Growth. To decide the relationship between scores in Math and English and the final result of programming and web programming study, we use two methods above. So, we can see the relationship between three indicators mentioned. The connection of math and programming, the highest confidence score was 54.5%. The connection between English and programming was 66.7% and the connection between English and web programming was 60%. For the result of FP-Growth Algorithm process, the high difference was seen in simpler and faster way, compared if using apriory. This is caused by the reason Algorithm FP-Growth, not all data scanned to get the score support based on frequencies from each data given.

Key Words: Data Mining, AprioryAlgorithm, Algorithm FP-Growth, Curriculum, Programming, Math

1. PENDAHULUAN

Penelitian ini bertujuan untuk membandingkan proses dari dua *Algoritma* yang sangat umum dalam Data Mining guna mendapatkan informasi yang dibutuhkan. Informasi tersebut berasal dari data-data yang dikumpulkan. Data dapat berasal dari beberapa sumber, seperti: hasil wawancara, observasi dan juga *database*. *Database* merupakan media penyimpanan dari sebuah program, yang berisi data-data mengenai institusi/organisasi yang menggunakan program tersebut. Data-data tersebut tidak akan bermanfaat hingga data tersebut

dapat diubah menjadi informasi yang sangat berguna.

Dalam melakukan atau mengubah data-data tersebut menjadi sebuah informasi yang bermanfaat, dibutuhkan suatu konsep untuk mencari atau membantu menggali informasi tersebut. Salah satu konsep yang berhubungan dengan hal ini adalah *Data Mining*. *Data Mining* adalah suatu konsep yang tujuannya menemukan data yang tersembunyi pada *database* dan merupakan data yang berjumlah besar. Konsep inilah yang juga akan digunakan pada SMK Sandikta Bekasi, untuk

menemukan data-data yang terkait dengan data nilai ijazah Matematika dan Bahasa Inggris dengan data mata pelajaran Bahasa Pemrograman dan *Web Programming*.

Adapun data yang diambil merupakan data dari murid kelas X-A SMK Sandikta, tahun ajaran 2013-2014. Dalam hal ini kelas X tahun ajaran 2013-2014 SMK Sandikta menggunakan kurikulum 2013 yang memuat dalam kurikulum tersebut mata pelajaran Bahasa Pemrograman dan *Web Programming* yang diajarkan pada kelas X.

Penggunaan Data Mining ini diharapkan dapat membantu keterhubungan antara nilai ijazah Matematika dan Bahasa Inggris dengan mata pelajaran Bahasa Pemrograman dan *Web Programming*, sehingga hasilnya nanti dapat menjadi referensi yang berguna dan bermanfaat bagi pihak SMK Sandikta Bekasi. Selain itu dengan membandingkan dua proses *Data Mining* dengan menggunakan *Algoritma Apriori* dan *Algoritma FP-Growth*, akan didapat kekurangan dan kelebihan masing-masing proses *Algoritma* tersebut.

2. LANDASAN TEORI

A. Data Mining.

Data mining adalah suatu konsep yang digunakan untuk menemukan pengetahuan atau informasi 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 tersimpan didalam *database* besar (Turban, 2005).

Menurut Gartner Group, *Data Mining* adalah suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyampaian dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Larose, 2005)

B. Algoritma Apriori

Algoritma Apriori adalah salah satu algoritma pertama yang terbaik untuk melakukan pencarian semua *frequent itemset* dengan menggunakan teknik *Association Rule* (Erwin, 2009).

Algoritma Apriori juga bisa berfungsi sebagai algoritma untuk membentuk dan menghasilkan kandidat secara lebih baik, karena algoritma ini menggunakan cara pencarian dari data pertama hingga kandidat terhitung.

Keuntungan dari *Algoritma Apriori* adalah:

1. Dapat mengakomodasi data yang sangat besar
2. Sangat mudah dan sederhana dalam mengaplikasikannya

C. Algoritma FP-Growth

Algoritma FP-Growth merupakan pengembangan dari *Algoritma Apriori*. *Frequent Pattern Growth (FP-Growth)* adalah salah satu alternatif algoritma yang digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemsets*) dalam sebuah kumpulan data (David, 2008).

Pada *Algoritma FP-Growth* menggunakan konsep pembangunan *tree*, yang disebut *FP-Tree*, dalam pencarian *frequent itemset* bukan menggunakan *generate candidate* seperti yang dilakukan pada *Algoritma Apriori*. Dengan menggunakan konsep tersebut, *Algoritma FP-Growth* menjadi lebih cepat dibandingkan dengan *Algoritma Apriori* (Erwin, 2009).

Metode *FP-Growth* dibagi menjadi tiga tahapan utama, yaitu (Micheline, 2006):

- Tahap pembangkitan *conditional pattern base*.
- Tahap pembangkitan *conditional pattern Tree*.
- Tahap pencarian *frequent itemset*.

D. Association Rule

Association Rule merupakan suatu proses pada data mining untuk menentukan semua aturan asosiatif yang memenuhi syarat minimum untuk *support* (minsup) dan *confidence* (minconf) pada sebuah database. Kedua syarat tersebut akan digunakan untuk *interesting association rule* dengan dibandingkan dengan batasan yang telah ditentukan, yaitu minsup dan minconf (Rosadi, 2013).

Association Rule Mining adalah suatu prosedur untuk mencari hubungan item dalam dataset. Dimulai dengan mencari *frequent itemset*, yaitu kombinasi yang paling sering terjadi dalam suatu *itemset* dan harus memenuhi minsup (Rosadi, 2013).

E. Support and Confidence

Di dalam *Association Rule Mining* ada dua parameter penting yang berfungsi untuk pembentukan *rules* yaitu *support* dan *confidence*, sehingga dihasilkan aturan asosiasi yang kuat (*strong rules*).

Support adalah persentase kombinasi suatu *item* dalam *database*, sedangkan *confidence* adalah kuatnya hubungan antar *item* dalam aturan asosiasi. Dalam analisa asosiasi mempunyai metodologi dasar yang terbagi menjadi dua (kusrini, 2009):

a. Analisa Pola Frekuensi Tinggi

Pada tahap ini untuk mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam database. Nilai *support* sebuah *item* diperoleh dengan rumus berikut:

$$\text{Support (A, B)} = P(A \cap B)$$
$$\text{Support (A, B)} = \frac{\sum \text{Transaksi } A \cap B}{\sum \text{Total Transaksi}} \times 100\%$$

b. Pembentukan Aturan Asosiasi

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk

confidence, dengan menghitung aturan asosiatif $A \rightarrow B$. Nilai *confidence* dari aturan diperoleh dari rumus berikut:

$$\text{Confidence} = \frac{\text{Confidence} = P(B|A)}{\sum \text{Transaksi NENGANDung A dan B}}$$
$$\text{Confidence} = \frac{\sum \text{Transaksi NENGANDung A dan B}}{\sum \text{Transaksi NENGANDung A}} \times 100\%$$

3. METODE PENELITIAN

a. Lokasi dan Objek Penelitian

Lokasi penelitian dilakukan di SMK Sandikta Bekasi dengan objek penelitian yaitu siswa-siswi kelas X-A. Dari objek tersebut dibutuhkan data awal dari nilai ijazah SMP mata pelajaran Matematika dan Bahasa Inggris serta data nilai akhir mata pelajaran Bahasa Pemrograman dan *Web Programming* untuk mencari keterhubungan kedua data tersebut.

b. Metode Pengumpulan Data.

1) Data Primer

Data Primer adalah data yang dikumpulkan langsung dari responden. Teknik pengumpulan data primer yang dilakukan menggunakan satu cara, yaitu wawancara. Wawancara dilakukan kepada kepala sekolah SMK Sandikta Bekasi. Tujuannya adalah meminta izin melakukan riset dan meminta data murid yang dibutuhkan.

2) Data Sekunder

Data Sekunder adalah data yang sudah tersedia sehingga data bisa langsung diberikan oleh kepala sekolah. Data berbentuk *softcopy* yang sudah diolah dengan *microsoft excel*, yang sudah diolah oleh Tata Usaha SMK Sandikta Bekasi.

Apriori dan *Algoritma FP-Growth* pada mata pelajaran Matematika dan Bahasa Inggris dengan data nilai Mata Pelajaran Pemrograman dan *Web Programming*.

4. HASIL PENELITIAN

Setelah mendapatkan data, maka dilakukan proses penerapan *Algoritma*

a. Algoritma Apriori

Dari data-data tersebut telah ditentukan *itemset-itemset* yang akan digunakan.

<i>Itemset Data</i> yang digunakan	Inisialisasi
Matematika	M
Bahasa Inggris	I
Bahasa Pemrograman	P
Web Programming	W

Tabel 1

Berikut data siswa berdasarkan *itemset* di atas:

No	M	I	P	W		No	M	I	P	W
1	5	7	8	8		11	6	7	7	6
2	4	7	7	6		12	6	8	7	7
3	6	6	8	6		13	5	7	6	8
4	6	6	8	8		14	5	6	8	8
5	5	7	6	6		15	5	6	8	8
6	5	7	5	7		16	5	6	8	8
7	5	7	6	7		17	5	6	6	7
8	4	6	6	7		18	6	8	7	6
9	5	6	8	7		19	4	7	5	6
10	5	6	7	6		20	4	7	5	6
						21	4	6	6	7

Tabel 2

1) Nilai Matematika dan Bahasa Pemrograman

Langkah pertama yang dilakukan adalah penginisialisasian. Tujuannya adalah memudahkan saat melakukan proses penghitungan. Nilai Matematika dibagi menjadi 3: Nilai 4 berinisial M4,

Nilai 5 berinisial M5 dan Nilai 6 berinisial M6.

Nilai Bahasa Pemrograman dibagi menjadi 4: Nilai 5 berinisial P5, Nilai 6 berinisial P6, Nilai 7 berinisial P7 dan Nilai 8 berinisial P8.

Iterasi 1.1.A.

1. Tabulasi Data

<i>Itemset</i>	Jumlah
M4	5
M5	11
M6	5
P5	3
P6	6
P7	5
P8	7

Tabel 3

2. *Frequent Itemset*, dengan mengoleksi data berdasarkan *minimum support* 25%.

Itemset	Jumlah	Support
M4	5	$\frac{5}{21} \times 100\% = 23.8\%$
M5	11	$\frac{11}{21} \times 100\% = 52.4\%$
M6	5	$\frac{5}{21} \times 100\% = 23.8\%$
P5	3	$\frac{3}{21} \times 100\% = 14.3\%$
P6	6	$\frac{6}{21} \times 100\% = 28.6\%$
P7	5	$\frac{5}{21} \times 100\% = 23.8\%$
P8	7	$\frac{7}{21} \times 100\% = 33.3\%$

Tabel 4

Data yang tidak memenuhi minimum *support* akan dieliminasi.

Pola Kombinasi 2 *itemset* yang memenuhi minimum *support* di Iterasi 1.1.A.

Iterasi 2.1.A.

Pola 2 itemset	Jumlah (K)	Support	K+1	Support
M5,P6	4	$\frac{4}{21} \times 100\% = 19.8\%$	5	$\frac{5}{21} \times 100\% = 23.8\%$
M5,P8	5	$\frac{5}{21} \times 100\% = 23.8\%$	6	$\frac{6}{21} \times 100\% = 28.6\%$
P6,P8	0	0	0	0

Tabel 5

Proses pengombinasian hanya sampai pola 2 *itemset*.

Tabel Final Association Rule

Pola 2 <i>itemset</i>	Support	Confidence
M5,P8	$\frac{5}{21} \times 100 = 23.8\%$	$\frac{6}{11} \times 100\% = 54.5\%$

Tabel 6

1) Nilai Bahasa Inggris dan Bahasa Pemrograman

Langkah pertama yang dilakukan adalah penginisialisasian. Tujuannya adalah memudahkan saat melakukan proses penghitungan. Nilai Bahasa Inggris

dibagi menjadi 3: Nilai 5 berinisial I6, Nilai 7 berinisial I7 dan Nilai 8 berinisial I8. Nilai Bahasa Pemrograman dibagi menjadi 4: Nilai 5 berinisial P5, Nilai 6 berinisial P6, Nilai 7 berinisial P7 dan Nilai 8 berinisial P8.

Iterasi 1.2.A.

1. Tabulasi Data

<i>Itemset</i>	Jumlah
I6	9
I7	10
I8	2
P5	3
P6	6
P7	5
P8	7

Tabel 7

2. *Frequent Itemset*, dengan mengoleksi data berdasarkan minimum *support* 25%

<i>Itemset</i>	Jumlah	<i>Support</i>
I6	9	$\frac{9}{21} \times 100\% = 42.8\%$
I7	10	$\frac{10}{21} \times 100\% = 47.6\%$
I8	2	$\frac{2}{21} \times 100\% = 9.5\%$
P5	3	$\frac{3}{21} \times 100\% = 19.3\%$
P6	6	$\frac{6}{21} \times 100\% = 28.6\%$
P7	5	$\frac{5}{21} \times 100\% = 23.8\%$
P8	7	$\frac{7}{21} \times 100\% = 33.3\%$

Tabel 8

Data yang tidak memenuhi minimum *support* akan dieliminasi.

Iterasi 2.2.A.

Pola Kombinasi 2 itemset yang memenuhi minimum *support* di Iterasi 1.2.A

Pola 2 <i>itemset</i>	Jumlah	<i>Support</i>
I6,P6	2	$\frac{2}{21} \times 100\% = 9.5\%$
I6,P8	6	$\frac{6}{21} \times 100\% = 28.6\%$
I7,P6	4	$\frac{4}{21} \times 100\% = 19.1\%$
I7,P8	1	$\frac{1}{21} \times 100\% = 4.8\%$
I6,I7	0	0
P6,P8	0	0

Tabel 9

Proses pengombinasian hanya sampai pola 2 *itemset*

Tabel Final Association Rule

Pola 2 <i>itemset</i>	<i>Support</i>	<i>Confidence</i>
I6,P8	$\frac{6}{21} \times 100 = 28.6\%$	$\frac{6}{9} \times 100\% = 66.7\%$

Tabel 10

1) Nilai Matematika dan Web Programming

Langkah pertama yang dilakukan adalah penginisialisasian. Tujuannya adalah memudahkan saat melakukan proses penghitungan. Nilai Matematika dibagi menjadi 3: Nilai 4 berinisial M4, Nilai 5 berinisial M5 dan Nilai 6

berinisial M6. Nilai *Web Programming* dibagi menjadi 3: Nilai 6 berinisial W6, Nilai 7 berinisial W7 dan Nilai 8 berinisial W8.

Iterasi 1.3.A.

1. Tabulasi Data

<i>Itemset</i>	Jumlah
M4	5
M5	11
M6	5
W6	8
W7	7
W8	6

Tabel 11

2. *Frequent Itemset*, dengan mengoleksi data berdasarkan *minimum support* 25%

<i>Itemset</i>	Jumlah	<i>Support</i>
M4	5	$\frac{5}{21} \times 100\% = 23.8\%$
M5	11	$\frac{11}{21} \times 100\% = 52.4\%$
M6	5	$\frac{5}{21} \times 100\% = 23.8\%$
W6	8	$\frac{8}{21} \times 100\% = 38.1\%$
W7	7	$\frac{7}{21} \times 100\% = 33.3\%$
W8	6	$\frac{6}{21} \times 100\% = 28.6\%$

Tabel 12

Data yang tidak memenuhi *minimum support* akan dieliminasi.

1. Pola Kombinasi 2 *itemset* yang memenuhi *minimum support* di Iterasi 1.2.A.

Iterasi 2.3.4.

Pola 2 <i>itemset</i>	Jumlah (K)	<i>Support</i>	K+1	<i>Support</i>
M5,W6	2	$\frac{2}{21} \times 100\% = 9.5\%$	3	$\frac{3}{21} \times 100\% = 14.7\%$
M5,W7	4	$\frac{4}{21} \times 100\% = 19\%$	5	$\frac{5}{21} \times 100\% = 23.9\%$
M5,W8	5	$\frac{5}{21} \times 100\% = 23.9\%$	6	$\frac{6}{21} \times 100\% = 28.7\%$
W6,W7	0	0	0	0
W6,W8	0	0	0	0
W7,W8	0	0	0	0

Tabel 13

Proses pengombinasian hanya sampai pola 2 *itemset*.

Tabel Final Association Rule

Pola 2 itemset	Support	Confidence
M5,W8	$\frac{6}{21} \times 100 = 28.7\%$	$\frac{6}{11} \times 100\% = 54.5\%$

Tabel 14

2) **Nilai Bahasa Inggris dan Web Programming.**

Langkah pertama yang dilakukan adalah penginisialisasian. Tujuannya adalah memudahkan saat melakukan proses penghitungan. Nilai Bahasa Inggris dibagi menjadi 3: Nilai 5 berinisial I6,

Nilai 7 berinisial I7 dan Nilai 8 berinisial I8. Nilai Web Programming dibagi menjadi 3: Nilai 6 berinisial W6, Nilai 7 berinisial W7 dan Nilai 8 berinisial W8.

Iterasi 1.4.A

1. Tabulasi Data

Itemset	Jumlah
I6	9
I7	10
I8	2
W6	8
W7	7
W8	6

Tabel 15

2 *Frequent Itemset*, dengan mengoleksi data berdasarkan minimum support 25%.

Itemset	Jumlah	Support
I6	9	$\frac{9}{21} \times 100\% = 42.8 \%$
I7	10	$\frac{10}{21} \times 100\% = 47.6 \%$
I8	2	$\frac{2}{21} \times 100\% = 9.5 \%$
W6	8	$\frac{8}{21} \times 100\% = 38.1 \%$
W7	7	$\frac{7}{21} \times 100\% = 33.3 \%$
W8	6	$\frac{6}{21} \times 100\% = 28.6 \%$

Tabel 16

Data yang tidak memenuhi minimum support akan dieliminasi.

Iterasi 2.4.A

1. Pola kombinasi 2 itemset yang memenuhi minimum support di Iterasi 1.4.A

Pola 2 itemset	Jumlah (K)	Support	K+1	Support
I6,W6	2	$\frac{2}{21} \times 100\% = 9.5 \%$	3	$\frac{3}{21} \times 100\% = 14.7\%$
I6,W7	3	$\frac{3}{21} \times 100\% = 14.3 \%$	4	$\frac{4}{21} \times 100\% = 19\%$
I6,W8	4	$\frac{4}{21} \times 100\% = 19 \%$	5	$\frac{5}{21} \times 100\% = 23.8\%$
I7,W6	5	$\frac{5}{21} \times 100\% = 23.8 \%$	6	$\frac{6}{21} \times 100\% = 28.6\%$
I7,W7	3	$\frac{3}{21} \times 100\% = 14.3 \%$	4	$\frac{4}{21} \times 100\% = 19\%$
I8,W8	2	$\frac{2}{21} \times 100\% = 9.5\%$	3	$\frac{3}{21} \times 100\% = 14.3\%$

Tabel 17

Proses pengombinasian hanya sampai 2 pola itemsets

Tabel Final Association Rule

Pola 2 itemset	Support	Confidence
I7,W6	$\frac{6}{21} \times 100 = 28.7\%$	$\frac{6}{10} \times 100\% = 60\%$

Tabel 18

b. Algoritma FP-Growth

Berikut *dataset* yang akan digunakan dalam pengolahan data menggunakan *Algoritma FP-Growth*.

Itemset Data yang digunakan	Inisialisasi
Matematika	M
Bahasa Inggris	I
Bahasa Pemrograman	P
Web Programming	W

Tabel 19

Berikut data siswa berdasarkan item set di atas:

No	M	I	P	W	No	M	I	P	W
1	5	7	8	8	11	6	7	7	6
2	4	7	7	6	12	6	8	7	7
3	6	6	8	6	13	5	7	6	8
4	6	6	8	8	14	5	6	8	8
5	5	7	6	6	15	5	6	8	8
6	5	7	5	7	16	5	6	8	8
7	5	7	6	7	17	5	6	6	7
8	4	6	6	7	18	6	8	7	6
9	5	6	8	7	19	4	7	5	6
10	5	6	7	6	20	4	7	5	6
					21	4	6	6	7

Tabel 20

Untuk selanjutnya *dataset* tersebut akan *diinisialisasi* untuk memudahkan dalam membuat data frekuensi kemunculan itemset tersebut. Berikut inisialisasi dari *itemset-itemset* tersebut:

Nilai Matematika dibagi menjadi 3: Nilai 4 berinisial M4, Nilai 5 berinisial M5 dan Nilai 6 berinisial M6. Nilai Bahasa Inggris dibagi menjadi 3: Nilai 5 berinisial I6, Nilai 7 berinisial I7 dan Nilai 8 berinisial I8.

Nilai Bahasa Pemrograman dibagi menjadi 4: Nilai 5 berinisial P5, Nilai 6 berinisial P6, Nilai 7 berinisial P7 dan Nilai 8 berinisial P8. Nilai Web Programming dibagi menjadi 3: Nilai 6 berinisial W6, Nilai 7 berinisial W7 dan Nilai 8 berinisial W8.

Berikut adalah tabel inisial dan frekuensi dari itemset-itemset pada dataset awal:

Itemset	Jumlah
M4	5
M5	11
M6	5
I6	9
I7	10
I8	2
W6	8
W7	7
W8	6
P5	3
P6	6
P7	5
P8	7

Tabel 21

1. Hubungan Mata Pelajaran Matematika dengan Bahasa Pemrograman

mata pelajaran Pemrograman, berdasarkan pada *dataset* awal.

a. Dataset

Tabel TID merupakan betuk dari *dataset* kombinasi 2 *itemset* mata pelajaran Matematika dan Bahasa Pemrograman.

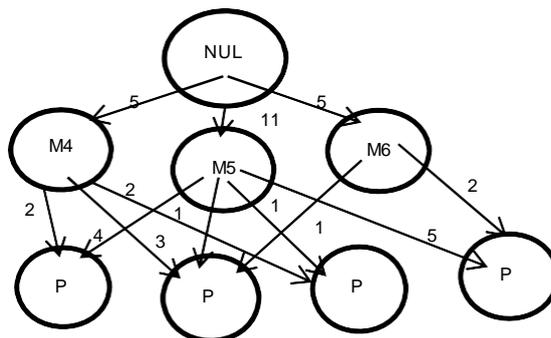
Berikut adalah tabel keterhubungan antara mata pelajaran Matematika dengan

No	M	P	Itemset	No	M	P	Itemset
1	5	8	M5,P8	11	6	7	M6,P7
2	4	7	M4,P7	12	6	7	M6,P7
3	6	8	M6,P8	13	5	6	M5,P6
4	6	8	M6,P8	14	5	8	M5,P8
5	5	6	M5,P6	15	5	8	M5,P8
6	5	5	M5,P5	16	5	8	M5,P8
7	5	6	M5,P6	17	5	6	M5,P6
8	4	6	M4,P6	18	6	7	M6,P7
9	5	8	M5,P8	19	4	5	M4,P5
10	5	7	M5,P7	20	4	5	M4,P5
				21	4	6	M4,P6

b. Pembentukan FP-Tree

Berikut bentukan *FP-Tree* dari kombinasi 2 *itemset* bredasarkan tabel TID di atas.

Dengan hasil *dataset* yang sudah diolah tersebut, selanjutn akan dibentuk *FP-Tree* dengan kombinasi 2 *itemset*.



c. FP-Growth

Pada tahapan ini dapat dilakukan dengan melihat kembali *FP-Tree* yang pernah dibuat sebelumnya.

Untuk langkah selanjutnya adalah tahap pembangkitan *Conditional Patern Base*, pembangkitan *Conditional FP-Tree* dan tahap pencarian *frequent itemset*.

Item	Conditional Patern Base
P5	{{M4:2}, {M5:1}}
P6	{{M4:2}, {M5:4}}
P7	{{M4:1}, {M5:1}, {M6:3}}
P8	{{M5:5}, {M6:2}}

Item	Conditional FP-Tree
P5	{{M4:2}, {M5:1}}
P6	{{M4:2}, {M5:4}}
P7	{{M4:1}, {M5:1}, {M6:3}}
P8	{{M5:5}, {M6:2}}

Item	Frequent Itemset
P5	{{M4,P4:2}, {M5,P5:1}, {P5:3}}
P6	{{M4,P6:2}, {M5,P6:4}, {P6:6}}
P7	{{M4,P7:1}, {M5,P7:1}, {M6,P7:3}, {P7:5}}
P8	{{M5,P8:5}, {M6,P8:2}, {P8:7}}

a. Association Rule

Setelah mendapat pola selanjutnya kita akan menghitung *confidence* berdasarkan dari tabel-tabel dan *FP-Tree* di atas.

$$\text{Sueeort} = \frac{5}{21} \times 100\%$$

$$\text{Sueeort} = 23.8 \%$$

Dalam hal ini syarat minimum yang dibutuhkan adalah 25%, sehingga hasil yang didapat dari penghitungan berdasarkan rumus *support* adalah:

Untuk nilai Matematika:

$$\text{Sueeort} = \frac{11}{21} \times 100\%$$

$$\text{Sueeort} = 52.4 \%$$

Untuk nilai Bahasa Pemrograman:

$$\text{Sueeort} = \frac{3}{21} \times 100\% = 19.3 \%$$

$$\text{Sueeort} = \frac{6}{21} \times 10\% = 28.6 \%$$

$$\text{Sueeort} = \frac{5}{21} \times 100\% = 23.8 \%$$

$$\text{Sueeort} = \frac{7}{21} \times 100\% = 33.3 \%$$

Dengan didapatkannya seluruh nilai *support*, maka langkah selanjutnya yaitu dengan mencari nilai *confidence* dari

seluruh itemset yang memenuhi kriteria minimum *support*:

Berikut hasil perhitungan *confidence*:

Pola 2 itemset	Confidence
M5,W8	$\frac{5}{11} \times 100\% = 45.5\%$

2. Hubungan Mata Pelajaran Matematika dengan Web Programming

dengan mata pelajaran *Web Programming*, berdasarkan pada *dataset* awal. Tabel

a. Dataset

Berikut adalah tabel keterhubungan antara mata pelajaran Matematika

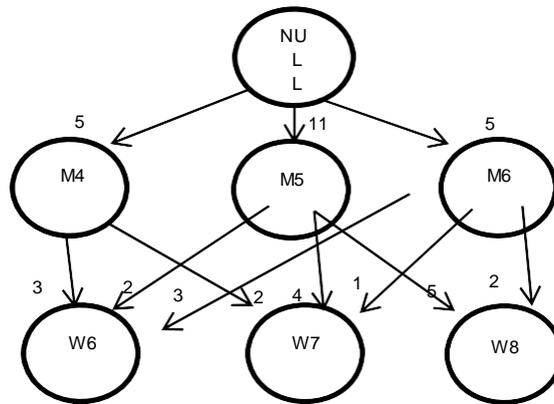
TID merupakan bentuk dari *dataset* kombinasi 2 *itemset* mata pelajaran Matematika dan *Web Programming*.

No	M	W	Itemset	No	M	W	Itemset
1	5	8	M5,W8	11	6	6	M6,W6
2	4	6	M4,W6	12	6	7	M6,W7
3	6	6	M6,W6	13	5	8	M5,W8
4	6	8	M6,W8	14	5	8	M5,W8
5	5	6	M5,W6	15	5	8	M5,W8
6	5	7	M5,W7	16	5	8	M5,W8
7	5	7	M5,W7	17	5	7	M5,W7
8	4	7	M4,W7	18	6	6	M6,W6
9	5	7	M5,W7	19	4	6	M4,W6
10	5	6	M5,W6	20	4	6	M4,W6
				21	4	7	M4,W7

b. Pembentukan FP-Tree

Dengan hasil *dataset* yang sudah diolah tersebut, selanjutnya akan dibentuk *FP-Tree* dengan kombinasi 2 itemset. Berikut bentuk

FP-Tree dari kombinasi 2 *itemset* berdasarkan tabel TID di atas.



c. FP-Growth

Untuk langkah selanjutnya adalah tahap pembangkitan *Conditional Patern Base*, pembangkitan *Conditional FP-Tree* dan tahap pencarian *frequent itemset*.

Pada tahapan ini dapat dilakukan dengan melihat kembali *FP-Tree* yang pernah dibuat sebelumnya

Item	Conditional Patern Base
W6	{{M4:3}, {M5:2}, {M6:3}}
W7	{{M4:2}, {M5:4}, {M6:1}}
W8	{{M5:5}, {M6:2}}

Item	Conditional FP-Tree
W6	{{M4:3}, {M5:2}, {M6:3}}
W7	{{M4:2}, {M5:4}, {M6:1}}
W8	{{M5:5}, {M6:2}}

Item	Frequent Itemset
W6	{{M4,W6:3}, {M5,W6:2}, {M6,W6:3}}
W7	{{M4,W7:2}, {M5,W7:4}, {M6,W7:1}}
W8	{{M5,W8:5}, {M6,W8:2}}

d. Association Rule

Setelah mendapat pola selanjutnya kita akan menghitung *confidence* berdasarkan dari tabel-tabel dan *FP-Tree* di atas. Dalam hal ini syarat minimum

yang dibutuhkan adalah 25%. Sehingga hasil yang didapat dari penghitungan berdasarkan rumus *support* adalah:

Untuk nilai matematika:

$$\text{Support} = \frac{5}{21} \times 100\%$$

$$\text{Support} = 23.8 \%$$

$$\text{Support} = \frac{11}{21} \times 100\%$$

$$\text{Support} = 52.4 \%$$

Untuk nilai bahasa pemrograman:

$$\text{Support} = \frac{8}{21} \times 100\% = 38.1\%$$

$$\text{Support} = \frac{7}{21} \times 100\% = 33.3\%$$

$$\text{Support} = \frac{6}{21} \times 100\% = 28.6\%$$

Dengan didaptkannya seluruh nilai *support*, maka langkah selanjutnya yaitu dengan mencari nilai *confidence* dari

seluruh *itemset* yang memenuhi kriteria minimum *support*. Berikut hasil perhitungan *confidence*:

Pola 2 <i>itemset</i>	<i>Confidence</i>
M5,W8	$\frac{5}{11} \times 100\% = 45.5\%$

3. Hubungan Mata Pelajaran Bahasa Inggris dengan Bahasa Pemrograman

dengan mata pelajaran Pemrograman, berdasarkan pada *dataset* awal:

Tabel TID berikut merupakan bentuk dari *dataset* kombinasi 2 *itemset* mata pelajaran Bahasa Inggris dan Bahasa Pemrograman

a. Dataset

Berikut adalah tabel keterhubungan antara mata pelajaran Bahasa Inggris

No	I	P	<i>Itemset</i>	No	I	P	<i>Itemset</i>
1	7	8	I7,P8	11	7	7	I7,P7
2	7	7	I7,P7	12	8	7	I8,P7
3	6	8	I6,P8	13	7	6	I7,P6
4	6	8	I6,P8	14	6	8	I6,P8
5	7	6	I7,P6	15	6	8	I6,P8
6	7	5	I7,P5	16	6	8	I6,P8
7	7	6	I7,P6	17	6	6	I6,P6
8	6	6	I6,P6	18	8	7	I8,P7
9	6	8	I6,P8	19	7	5	I7,P5
10	6	7	I6,P7	20	7	5	I7,P5
				21	6	6	I6,P6

b. Pembentukan FP-Tree

Dengan hasil *dataset* yang sudah diolah tersebut, selanjutnya akan dibentuk *FP-Tree* dengan kombinasi 2 *itemset*.

Berikut bentuk *FP-Tree* dari kombinasi 2 *itemset* berdasarkan tabel TID di atas.

c. FP-Growth

Untuk langkah selanjutnya adalah tahap pembangkitan *Conditional Patern Base*, pembangkitan *Conditional FP-Tree*

dan tahap pencarian frequent itemset. Pada tahapan ini dapat dilakukan dengan melihat kembali *FP-Tree* yang pernah dibuat sebelumnya

Item	Conditional Patern Base
P5	{{I7:3}}
P6	{{I6:3}, {I7:3}}
P7	{{I6:1}, {I7:2}, {I8:2}}
P8	{{I6:6}, {I7:1}}

Item	Conditional FP-Tree
P5	{{I7:3}}
P6	{{I6:3}, {I7:3}}
P7	{{I6:1}, {I7:2}, {I8:2}}
P8	{{I6:6}, {I7:1}}

Item	Frequent Itemset
P5	{{I7,P5:3}}
P6	{{I6,P6:3}, {I7,P6:3}}
P7	{{I6,P7:1}, {I7,P7:2}, {I8,P7:2}}
P8	{{I6,P8:6}, {I7,P8:1}}

d. Association Rule

Setelah mendapat pola selanjutnya kita akan menghitung *confidence* berdasarkan dari tabel-tabel dan *FP-Tree* di atas. Dalam hal ini syarat minimum

$$S_{\text{ueeort}} = \frac{10}{21} \times 100\%$$

$$S_{\text{ueeort}} = 47.8\%$$

$$S_{\text{ueeort}} = \frac{2}{21} \times 100\%$$

yang dibutuhkan adalah 25%. Sehingga hasil yang didapat dari perhitungan berdasarkan rumus *support* adalah:

Untuk nilai Bahasa Inggris:

$$S_{\text{ueeort}} = \frac{9}{21} \times 100\%$$

$$S_{\text{ueeort}} = 42.8\%$$

$$S_{\text{ueeort}} = 9.5\%$$

Untuk nilai Bahasa Pemrograman:

$$S_{\text{ueeort}} = \frac{3}{21} \times 100\% = 19.3\%$$

$$S_{\text{ueeort}} = \frac{6}{21} \times 10\% = 28.6\%$$

$$S_{\text{ueeort}} = \frac{5}{21} \times 100\% = 23.8\%$$

$$S_{\text{ueeort}} = \frac{7}{21} \times 100\% = 33.3\%$$

Dengan didapatkannya seluruh nilai *support*, maka langkah selanjutnya yaitu dengan mencari nilai *confidence* dari

seluruh *itemset* yang memenuhi kriteria minimum *support*. Berikut hasil perhitungan *confidence*:

Pola 2 itemset	Confidence
I6,P8	$\frac{6}{9} \times 100\% = 66.7\%$

4. Hubungan Mata Bahasa Inggris dengan Web Programming

a. Dataset

Berikut adalah tabel keterhubungan antara mata pelajaran Bahasa Inggris

dengan mata pelajaran *Web Programming*, berdasarkan pada *dataset* awal.

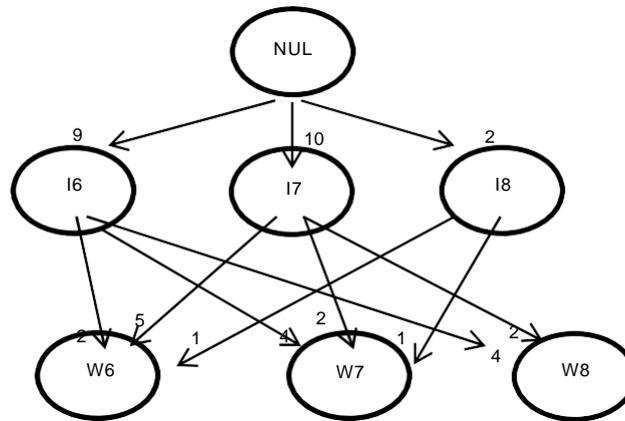
Tabel TID merupakan bentuk dari dataset kombinasi 2 *itemset* mata pelajaran Bahasa Inggris dan *Web Programming*

No	I	W	Itemset	No	I	W	Itemset
1	7	8	I7,W8	11	7	6	I7,W6
2	7	6	I7,W6	12	8	7	I8,W7
3	6	6	I6,W6	13	7	8	I7,W8
4	6	8	I6,W8	14	6	8	I6,W8
5	7	6	I7,W6	15	6	8	I6,W8
6	7	7	I7,W7	16	6	8	I6,W8
7	7	7	I7,W7	17	6	7	I6,W7
8	6	7	I6,W7	18	8	6	I8,W6
9	6	7	I6,W7	19	7	6	I7,W6
10	6	6	I6,W6	20	7	6	I7,W6
				21	6	7	I6,W7

b. Pembentukan FP-Tree

Dengan hasil dataset yang sudah diolah tersebut, selanjutnya akan dibentuk *FP-Tree* dengan kombinasi 2 *itemset*.

Berikut bentuk *FP-Tree* dari kombinasi 2 *itemset* berdasarkan tabel TID di atas.



c. FP-Growth

Untuk langkah selanjutnya adalah tahap pembangkitan *Conditional Patern Base*, pembangkitan *Conditional FP-Tree*

dan tahap pencarian *frequent itemset*. Pada tahapan ini dapat dilakukan dengan melihat kembali *FP-Tree* yang pernah dibuat sebelumnya

Item	Conditional Patern Base
W6	{{M4:3}, {M5:2}, {M6:3}}
W7	{{M4:2}, {M5:4}, {M6:1}}
W8	{{M5:5}, {M6:2}}

Item	Conditional FP-Tree
W6	{{M4:3}, {M5:2}, {M6:3}}
W7	{{M4:2}, {M5:4}, {M6:1}}
W8	{{M5:5}, {M6:2}}

Item	Frequent Itemset
W6	{{M4,W6:3}, {M5,W6:2}, {M6,W6:3}}
W7	{{M4,W7:2}, {M5,W7:4}, {M6,W7:1}}
W8	{{M5,W8:5}, {M6,W8:2}}

d. Association Rule

Setelah mendapat pola selanjutnya kita akan menghitung *confidence* berdasarkan dari tabel-tabel dan *FP-Tree* di atas. Dalam hal ini syarat minimum

yang dibutuhkan adalah 25%. Sehingga hasil yang didapat dari penghitungan berdasarkan rumus *support* adalah:

Untuk nilai Bahasa Inggris:

$$\text{Sueeort} = \frac{10}{21} \times 100\%$$

$$\text{Sueeort} = 47.8\%$$

$$\text{Sueeort} = \frac{9}{21} \times 100\% \quad \text{Sueeort} = 42.8\%$$

$$\text{Sueeort} = \frac{2}{21} \times 100\% \quad \text{Sueeort} = 9.5\%$$

Untuk nilai Bahasa Pemrograman:

$$\text{Sueeort} = \frac{8}{21} \times 100\% = 38.1$$

$$\text{Sueeort} = \frac{7}{21} \times 100\% = 33.3\%$$

$$\text{Sueeort} = \frac{6}{21} \times 100\% = 28.6\%$$

Dengan didapkannya seluruh nilai *support*, maka langkah selanjutnya yaitu dengan mencari nilai *confidence* dari

seluruh *itemset* yang memenuhi kriteria minimum *support*. Berikut hasil perhitungan *confidence*:

Pola 2 itemset	Confidence
I6,P8	$\frac{6}{9} \times 100\% = 66.7\%$

5. KESIMPULAN DAN SARAN

A. Kesimpulan

Kesimpulan dari hasil penelitian tersebut adalah:

- 1) Berdasarkan hasil keterhubungan nilai Matematika dengan nilai Bahasa Pemrograman yang memiliki nilai *confidence* tertinggi adalah 54.5%, dengan nilai matematika 5 dan Bahasa Pemrograman 8
- 2) Berdasarkan hasil keterhubungan nilai Bahasa Inggris dengan nilai Bahasa Pemrograman yang memiliki nilai *confidence* tertinggi adalah 66.7%, dengan nilai Bahasa Inggris 6 dan Bahasa Pemrograman 8
- 3) Berdasarkan hasil keterhubungan nilai Matematika dengan nilai *Web Programming* yang memiliki nilai *confidence* tertinggi adalah 54.5%, dengan nilai Matematika 5 dan Bahasa Web Programming 8.
- 4) Berdasarkan hasil keterhubungan nilai Bahasa Inggris dengan nilai *Web Programming* yang memiliki nilai *confidence* tertinggi adalah 60%, dengan nilai Bahasa Inggris 7 dan Bahasa Pemrograman 6

- 5) Untuk hasil dari proses *Algoritma FP-Growth*, didapatkan nilai *confidence* yang bervariasi, dan terlihat perbedaan proses yang lebih sederhana dan cepat dibandingkan dengan menggunakan Apriori. Alasan ini dikarenakan untuk *Algoritma FP-Growth* tidak semua data di-*scan* untuk menentukan nilai *support* berdasarkan frekuensi dari masing-masing data tersebut.

B. Saran

Adapun saran dari penulis adalah sebagai berikut:

- 1) Object dari penelitian bisa diperbanyak dengan menambah kelas lain di kelas X. Agar mendapatkan variasi data sebagai bahan acuan memberikan bahan ajaran kepada murid siswa kelas X pada mata pelajaran Pemrograman dan Bahasa Pemrograman
- 2) Membandingkan kedua proses dalam bentuk aplikasi, agar dapat terlihat perbedaan dari sisi kecepatan dan keakuratan hasil dari kedua *algoritma* tersebut, sebagai aplikasi penerapan dari kedua *algoritma* tersebut.

DAFTAR PUSTAKA

- Arifin, Nurul Riza. (2009). *Implementasi Algoritma FP-Growth Menentukan Asosiasi antar Produk (Studi Kasus Nadia Mart)*. Semarang: Universitas Dian Nuswantoro
- Hernawati, Astuti. (2013). *Pengenalan Data Mining*. Jakarta: Media Kita
- Kadir, Abdul. (2005). *Konsep dan Tuntutan Praktis Basis Data*. Yogyakarta: Andi
- Kusrini dan Lutfi, Taufik. (2009). *Algoritma Data Mining*. Jakarta: Penerbit Andi
- Larasati, Dyah Prameti. (2015). *Analisis dan Implementasi Algoritma FP-Growth Pada Aplikasi Smart Untuk Menentukan Market Basket Analisis Pada Usaha Retai*. Bandung: Universitas Telkom Bandung.
- Mythili, M.S. (2013). *Performance Evaluation of Apriori and FP-Growth Algorithms*. Tiruchirappalli Bishop Herber College.