

APLIKASI *HINTS* SOAL UJIAN MATEMATIKA DENGAN MENGGUNAKAN ALGORITMA *FREQUENT PATTERN GROWTH*

Siti Lia Sisimawati, Nurul Hasanah, Dewi Istiqomah

Program Manajemen Informatika, Universitas Trunojoyo Madura

Jl. Raya Telang, PO BOX 2, Kamal, Bangkalan - 69162

E-mail: Lea_ledys@yahoo.com¹, aLeafriili@gmail.com², bainkk@gmail.com³

ABSTRAK

Perkembangan *e-learning* pada saat ini sudah terbilang pesat, hampir semua pihak telah mengerti *e-learning*. MTS Al-FALAH adalah contoh sekolah yang sempat menggunakan *elearning* dalam suatu mata pelajaran. Namun, pengajar hanya memanfaatkan untuk mengambil nilai dari ujian yang diberikan. Oleh karenanya, dalam penelitian ini akan menggunakan data-data yang dianggap tidak penting untuk membantu para pengajar dalam mengukur perkembangan dan membantu siswa dalam mengerjakan soal yang dianggap sulit. Mata pelajaran yang digunakan dalam *e-learning* ini adalah matematika. Matematika merupakan salah satu mata pelajaran yang dianggap sulit oleh sebagian siswa, karenanya dengan aplikasi ini akan sedikit membantu siswa dalam mengerjakan soal. Algoritma yang digunakan dalam aplikasi ini adalah *frequent pattern growth*, yaitu algoritma dalam *association rules* untuk mencari hubungan kombinasi dari setiap data. Data yang digunakan adalah jawaban salah dari 120 siswa. Soal terdiri dari 200 nomor dari 20 materi yang ada. Dalam penelitian ini, data dipecah menjadi 5, yaitu data dari 5 murid, 25, 50, 75, 100 dan 120 murid. Hasil penelitian dari pembagian data tersebut, diketahui bahwa semakin banyak data, maka kombinasi itemset yang terbentuk juga akan semakin banyak, dan waktu proses juga akan semakin lama. Namun, semakin besar *minimum support* yang diinputkan, maka kombinasi yang terbentuk akan semakin sedikit.

Kata Kunci: *Association Rules, FP-Growth, minimum support*

ABSTRACT

The progress of e-learning was fairly fast, almost everyone understood about e-learning. MTS Al-FALAH was an example of schools that could use elearning in one of the subjects. However, the teachers only used to take the value of the test. Therefore, in this research would use data that are considered essential to assist teachers measured students's progress and help students answer the difficult questions. Subjects that used in this e-learning was math. Math was one subject that was considered difficult by most students. So, this application would help students solve a problems. The algorithm that used in this application was a frequent pattern growth, it was algorithm of association rules to find combinations of data's relations. The data were the wrong answers of 120 students. Questions consist of 200 numbers and 20 learning materials. In this research, the data was divided into 5, the data from 5 students, 25 students, 50 student, 75 students, 100 students and 120 students. The results of this reseacrh was if there were more data, the combination formed itemset also be more and more, and the processing time will also be longer. However, if the greater the minimum support was entered, then the combination was formed will be less.

Keywords: *Association Rules, FP-Growth, minimum support.*

PENDAHULUAN

Teknologi Informasi telah mengalami perkembangan pesat seiring dengan berjalannya waktu. Kebutuhan akan suatu mekanisme dan konsep dalam pendidikan juga mendapatkan dampak yang sangat baik. Proses belajar mengajar dengan berbasis TI menjadi suatu hal yang memberikan banyak kemudahan dan manfaat bagi tiap-tiap individu. Konsep yang kemudian terkenal dengan sebutan *e-learning* ini membawa pengaruh terjadinya proses transformasi pendidikan konvensional ke dalam bentuk digital, baik secara isi dan sistemnya. Kini, proses pembelajaran dapat dilakukan walau tanpa adanya tatap muka antara pengajar dan peserta didik. Dengan proses pembelajaran seperti itu, memungkinkan pelajar untuk belajar melalui komputer di tempat mereka masing-masing tanpa harus secara fisik pergi mengikuti pelajaran di kelas. Pengajar juga dapat memberikan tugas dan ujian secara *online*.

MTS AL-FALAH merupakan salah satu lembaga pendidikan yang ada di Kabupaten Bangkalan MTS Negeri ini sudah mulai mencicipi penggunaan *e-learning* sejak beberapa tahun lalu. Namun, sampai saat ini penggunaan *e-learning* hanya digunakan saat siswa melakukan ujian, dan hanya sebatas memperoleh nilai. Padahal, seiring penggunaan *e-learning* tersebut akan mendapatkan informasi lebih tentang perkembangan para siswa apabila memanfaatkan data-data yang telah ada dan memberikan sedikit bantuan bagi para siswa dalam pengerjaan soal ujian.

Pada penelitian ini dibangun aplikasi untuk mencari informasi tentang perkembangan siswa dengan melihat jawaban-jawaban siswa yang salah dalam suatu ujian. Dengan data tersebut akan diperoleh suatu kombinasi keterkaitan materi yang dirasa menyulitkan para siswa, dan akan dijadikan sarana untuk membantu para siswa dalam mengerjakan soal yang akan berupa *Hints*. Aplikasi ini menggunakan algoritma *FP-Growth*, yakni sebuah algoritma dalam *Association Rule* untuk menemukan relasi antar data.

Dalam penelitian yang lain, Anna Katrina Dominguez bersama dengan Kalina dan James C. Curran telah mencoba membangun *hints* serupa dalam sebuah ujian. Proses *Association Rules* yang dipakai menggunakan Algoritma *FP-Growth*[1]. Penelitian selanjutnya dilakukan oleh Erwin, penelitiannya tentang perbandingan kinerja dua algoritma dalam *Association Rules*, yaitu Apriori dan *FP-Growth*. Dalam penelitian ini memberikan informasi bahwa Algoritma *FP-Growth* lebih baik dari Algoritma Apriori[2]. Penggunaan algoritma *FP-Growth* pernah dilakukan oleh Andreas Handojo dalam penelitiannya untuk menentukan kombinasi dalam berdasarkan pola pemesanan menu oleh para pelanggan dalam sebuah restoran[3].

Berdasarkan ketiga penelitian tersebut, dapat dilihat bahwa penelitian serupa telah pernah dilakukan sebelumnya. Namun, juga telah diketahui telah ada algoritma yang lebih baik lagi. Oleh karena itu dalam penelitian ini akan menggunakan algoritma *FP-Growth* yang lebih baik dari algoritma Apriori untuk mencari kombinasi berdasarkan pola jawaban yang salah. Hasil akhir dari penelitian ini diharapkan dapat memperoleh informasi berdasarkan kombinasi terbentuk. *Rule* yang dihasilkan juga akan digunakan sebagai fitur *hints* untuk membantu siswa yang kesulitan dalam

memahami soal, sekaligus sebagai informasi bagi pengajar tentang materi-materi yang sulit dikuasai oleh para peserta didik.

TINJAUAN PUSTAKA

Profil MTS

Sekolah Madrasah Stanawiyah Al-Falah merupakan lembaga pendidikan negeri yang berada di Jl. Raya Buluh Dakiring Kec.Socah Kab.Bangkalan. MTS ini menjadi Sekolah Bertandar Internasional pada tahun 2008 dan mempunyai nilai akreditasi A.

Association Rules

Association rules adalah salah satu teknik yang terkenal dalam data mining. *Association rules* merupakan teknik data mining untuk mencari hubungan antar item dalam suatu *data set* [4]. Konsep ini sendiri diturunkan dari terminologi *market basket analysis*, yaitu pencarian hubungan dari beberapa produk di dalam transaksi pembelian [5]. *Association rule* meliputi dua tahap, yaitu mencari kombinasi yang paling sering terjadi dari suatu itemset. Dan mendefinisikan *Condition* dan *Result* (untuk *conditional association rule*) [6].

Association Rules juga bermanfaat untuk pemakaian data web yang berdasarkan personalitas. Pendekatan ini diadopsi dari hubungan dengan *collaborative filtering* [7]. Dalam menentukan suatu *association rules*, terdapat suatu *interestingness measure* (ukuran kepercayaan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu [8]. Umumnya ada dua ukuran, yaitu:

- *Support* : suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item/itemset dari keseluruhan transaksi. Ukuran ini menentukan apakah suatu item/itemset layak untuk dicari *confidence*-nya (misal, dari keseluruhan transaksi yang ada, seberapa besar tingkat dominasi yang menunjukkan bahwa item A dan B dibeli bersamaan).

$$\text{Support} = P(A \cup B)$$

$$= \frac{\text{jumlah transaksi yang mengandung A dan B}}{\text{total transaksi}} \dots\dots\dots (i)$$

- *Confidence*: suatu ukuran yang menunjukkan hubungan antar 2 item secara conditional (misal, seberapa sering item B dibeli jika orang membeli item A)

$$\text{Confidence} = P(A \cap B)$$

$$= \frac{\text{jumlah transaksi yang mengandung A dan B}}{\text{total transaksi yang mengandung A}} \dots\dots\dots (ii)$$

METODE

Frequent Pattern Growth

Algoritma *FP-Growth* merepresentasikan transaksi dengan menggunakan struktur data *FP-Tree* [9]. Cara kerja dari algoritma ini adalah melalui *scan database* yang dilakukan hanya dua kali saja. Data kemudian ditampilkan dalam bentuk *FP-tree*. Ketika selesai membuat *FP-Tree*, kita

tidak begitu saja bisa mendapatkan *frequent itemset* yang terdapat dalam *dataset*. Suatu kombinasi *itemset* bisa saja berada di beberapa *path* yang berbeda.

Untuk mendapatkan suatu *pattern* dalam *FP-Growth* langkah yang lebih mudah adalah mencari arah dari ujung suatu *path*, kemudian mencari mulai dari *header* untuk item di ujung tersebut, barulah kemudian dibuat berdasarkan tiap *node* berisi *item* tersebut dicari arah *path node* ke atas. Hal ini tentu lebih cepat dari pada *up-down* karena pointer langsung yang dimiliki tiap *node* adalah *pointer* ke *parent*. *Path-path* yang dieksplorasi hanyalah *path-path* yang memiliki *node* yang sedang dicari.

Berikut adalah tahapan yang lebih jelas dari *algoritma FP-Growth* [10]:

a. Membuat *Header Item*

Header dalam hal ini selain sebagai *header* suatu item ke *FP-Tree* juga sebagai jenis item dasar yang memenuhi *minimum support*. Setelah mendapatkan item dan nilai *support*-nya, maka item yang tidak *frequent* dibuang dan item diurutkan berdasarkan nilai *support*-nya.

b. Membuat *FP-Tree*

FP-Tree dibangun dengan mencari item sesuai urutan pada item yang *frequent*. Sesudah membuat *root*, tiap item yang ditemukan dimasukkan berdasarkan *path* pada *FP-Tree*. Jika item yang ditemukan sudah ada, maka nilai *support item* tersebut yang ditambahkan. Namun jika *path* belum ada, maka dibuat *node* baru untuk melengkapi *path* baru pada *FP-Tree* tersebut. Item-item yang ditemukan dalam transaksi akan berurutan memanjang ke bawah. Dalam struktur *FP-Tree*, diterapkan alur *path* dari *child* hingga ke *root*. Jadi, suatu *path* utuh dalam *FP-Tree* adalah dari *child* terbawah hingga ke *root*. Tiap *node* pada *FP-Tree* memiliki *pointer* ke *parent*, sehingga pencarian harus dimulai dari bawah.

c. *Pattern Extraction*

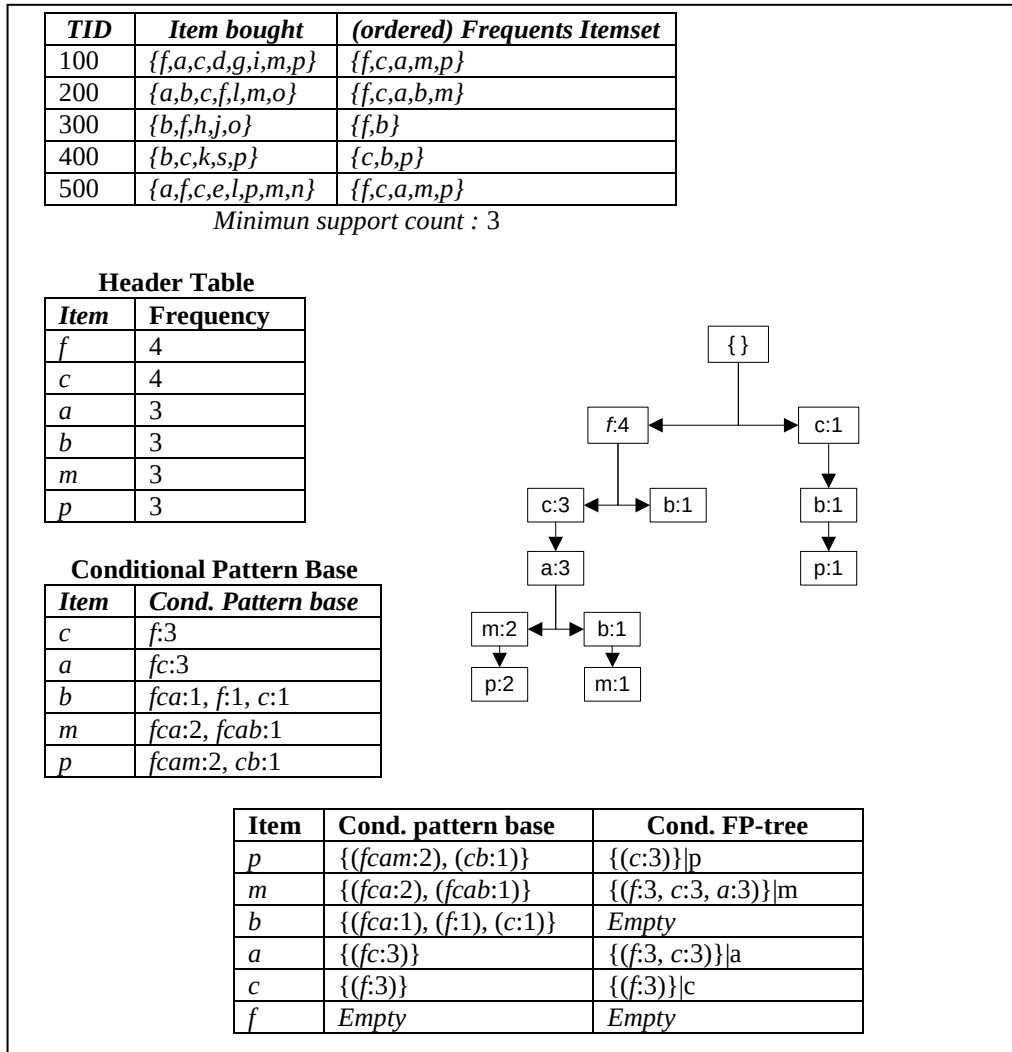
Pattern extraction dilakukan berdasarkan keterlibatan item pada suatu *path*. Di setiap *path*, diperiksa semua kombinasi yang mungkin dimana item tersebut terlibat. Di iterasi berikutnya dilakukan dengan melibatkan item berikutnya, tanpa melibatkan item sebelumnya, sehingga *pattern* yang sama tidak akan ditemukan dua kali pada *path* yang sama.

d. Memasukkan setiap *pattern* yang ditemukan ke dalam *Pattern tree*

Setelah mengolah *FP-Tree* menjadi *pattern-pattern*, diperlukan proses akumulasi *pattern-pattern* yang ditemukan mengingat *pattern* yang sama dapat ditemukan pada *path* yang berbeda. Untuk itu digunakan struktur data *Pattern tree*. Setiap *node* di *Pattern tree* merepresentasikan dan menyimpan frekuensi suatu *pattern*. *Pattern tree* terdiri atas *Pattern tree node* yang menyimpan nilai item dan nilai *support*.

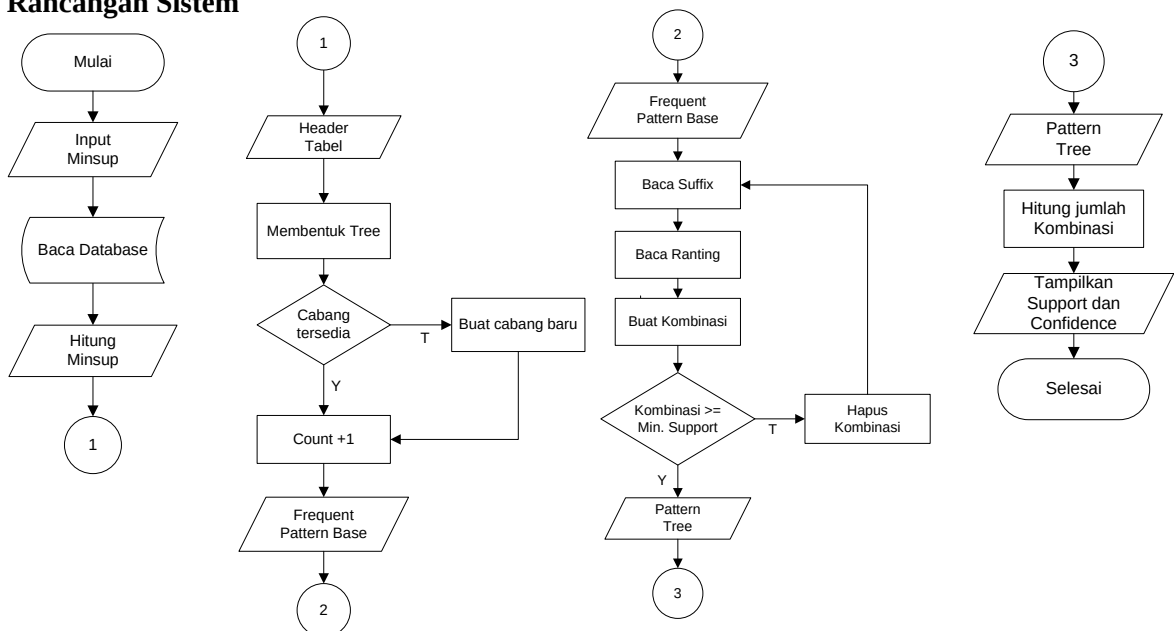
e. Mengurutkan dan Menyeleksi *Pattern*

Pattern yang tidak memenuhi *minimum support*, dihapus dari daftar *pattern*. *Pattern-pattern* yang tersisa kemudian diurutkan untuk memudahkan pembuatan *rules*



Gambar 1. Contoh Perhitungan *FP-Growth*

Rancangan Sistem



Gambar 2. Flowchart sistem dan *Fp-Growth* .

Perhitungan dalam aplikasi ini dimulai dengan memasukkan *minimum support* yang selanjutnya dilakukan proses oleh sistem berupa pemangkasan data yang tidak memenuhi *minimum support*, sehingga diperoleh *header table*. Dari *header tabel* inilah yang kemudian dilakukan pembentukan *tree*, pencarian kombinasi dan kemudian didapatkan nilai *support* dan *confidence* berdasarkan jumlah kombinasi yang terbentuk.

HASIL DAN PEMBAHASAN

Data yang digunakan dalam tugas akhir ini merupakan data dari jawaban salah siswa kelas IX yang diperoleh dari MTS Al-falah . Data didapatkan dari siswa sebanyak 120 orang yang telah diberikan soal sebelumnya. Dalam implementasinya, jawaban-jawaban salah tersebut yang akan dijadikan sebagai data mentah/transaksi untuk mencari kombinasi keterkaitan dari seluruhnya. Selanjutnya, data akan dipecah menjadi 5 dengan jumlah murid yang berbeda. Hal ini untuk melihat perbedaan kinerja algoritma FP-Growth dan perbedaan kombinasi yang muncul untuk data yang jumlahnya berbeda dan *minimum support* yang berbeda pula.

Dalam hal ini, akan digunakan 5 data awal sebagai data ujicoba untuk mencoba sistem, seperti pada tabel dibawah. *Minimum support* yang diinputkan adalah 50% dari seluruh transaksi data.

Tabel 1. Data Percobaan.

Murid	Materi yang tidak dikuasai
1	101,116,113,109,102,106,104,103
2	119,115,113,106,107,103,102,101
3	120,116,110,108,106,102
4	118,117,116,111,103,102,109
5	102,113,112,106,105,103,101,116

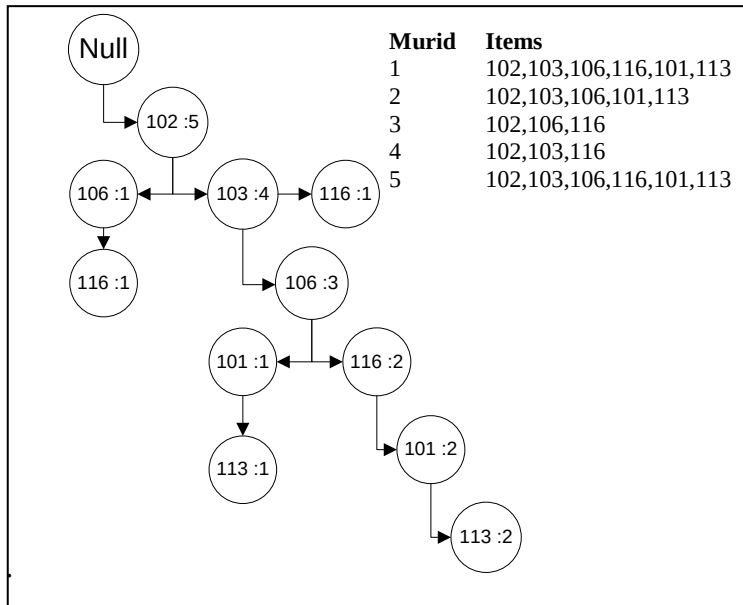
FP-Tree

Selanjutnya adalah mencari frekuensi tiap materi berdasarkan *minimum support* yang diminta, materi yang tidak memenuhi akan dihapus dari data karena tidak berpengaruh pada perhitungan selanjutnya. Dan materi yang tersisa, dalam hal ini adalah 102, 103, 106, 116, 101, 113 akan diurutkan berdasarkan frekuensi kemunculan.

Tabel 2. Data Terurut.

Murid	Materi yang sulit
1	102,103,106,116,101,113
2	102,103,106,101,113
3	102,106,116
4	102,103,116
5	102,103,106,116,101,113

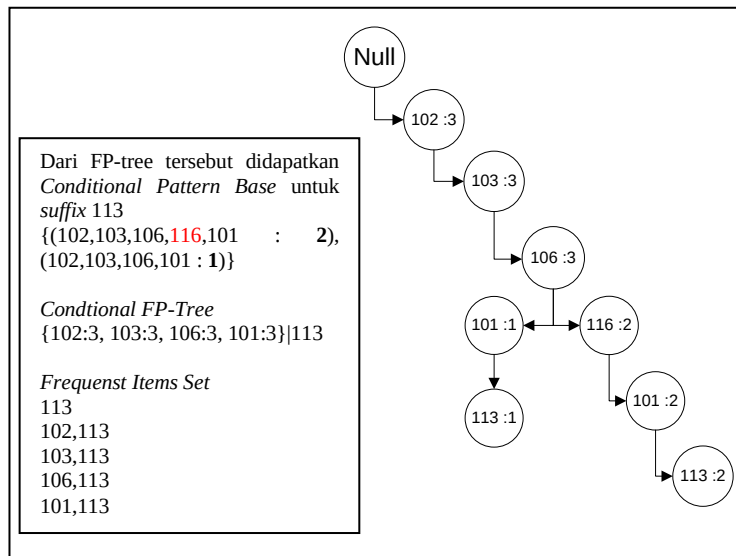
Berdasarkan data yang telah terurut diatas, langkah selanjutnya adalah membentuk *tree*. Pembuatan *tree* ini dimulai dari membaca murid pertama hingga terakhir. *Tree* yang berhasil dibentuk berdasarkan table di atas adalah seperti pada gambar berikut.



Gambar 3. Tree Sampai Murid 5

Conditional Pattern Tree

Proses dari Algoritma *FT-Growth* selanjutnya adalah *Conditional Pattern Tree* dari *Conditional Pattern Growth* yang di dapat untuk setiap *path*. Membangkitkan *conditional FP-Tree* dilakukan dari bawah ke atas dari item yang jumlah frekuensi kemunculannya terkecil. Berikut adalah *conditional pattern tree* dengan *suffix* 113. Proses pencarian *conditional pattern tree* untuk *suffix* yang lain tidak ada perbedaan, yaitu hanya menghapus kombinasi yang tidak memenuhi *minimum support count*.



Gambar 4. Conditional Pattern Tree untuk suffix 113

Support dan Confidence

Proses terakhir pada aplikasi ini adalah perhitungan *support* dan *confidence*. Setelah sebelumnya ditemukan kombinasi-kombinasi *itemset* yang memenuhi *minimum support*. Sekarang hanya perlu menggunakan rumus *support* dan *confidence*.

Tabel 3. Tabel *support* dan *confidence*

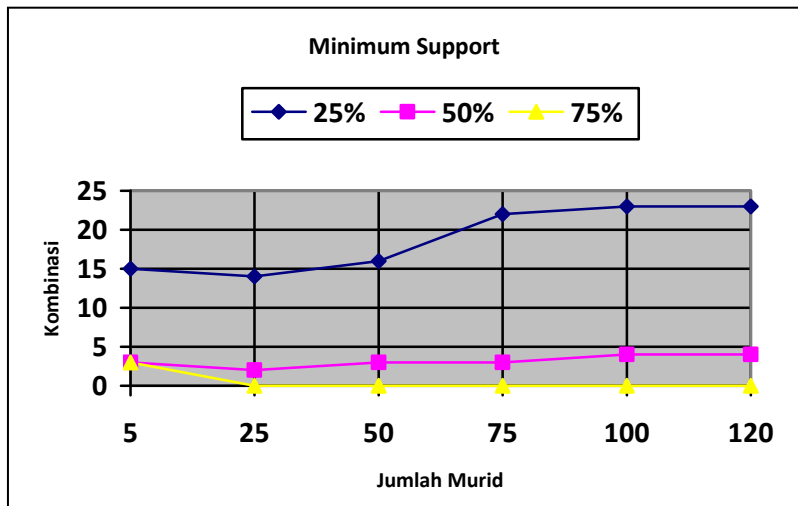
Kombinasi Itemset	Support	Confidence
{102,113}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{5} \times 100\% = 60\%$
{103,113}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{4} \times 100\% = 75\%$
{106,113}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{4} \times 100\% = 75\%$
{101,113}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{3} \times 100\% = 100\%$
{102,101}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{5} \times 100\% = 60\%$
{103,101}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{4} \times 100\% = 75\%$
{106,101}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{4} \times 100\% = 75\%$
{102,116}	$\frac{4}{5} \times 100\% = 80\%$	$\frac{4}{5} \times 100\% = 80\%$
{103,116}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{4} \times 100\% = 75\%$
{106,116}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{4} \times 100\% = 75\%$
{102,106}	$\frac{4}{5} \times 100\% = 80\%$	$\frac{4}{5} \times 100\% = 80\%$
{103,106}	$\frac{3}{5} \times 100\% = 60\%$	$\frac{3}{4} \times 100\% = 75\%$
{102,103}	$\frac{4}{5} \times 100\% = 80\%$	$\frac{4}{5} \times 100\% = 80\%$
13	Kombinasi	

Analisa

Setelah proses perhitungan *support* dan *confidence* dari semua kombinasi berhasil didapatkan. Selanjutnya adalah proses perbandingan untuk jumlah data yang berbeda dan *minimum support* yang berbeda pula.

Tabel 4. Tabel Perbandingan

No	Jumlah Murid	Minimum Support	Jumlah Kombinasi
1	5	25%	15
		50%	3
		75%	3
2	25	25%	14
		50%	2
		75%	0
3	50	25%	16
		50%	3
		75%	0
4	75	25%	22
		50%	3
		75%	0
5	100	25%	23
		50%	4
		75%	0
6	120	25%	23
		50%	4
		75%	0



Gambar 5. Grafik hasil analisa perbandingan jumlah kombinasi

SIMPULAN DAN SARAN

Simpulan

Beberapa kesimpulan yang dapat diambil dari hasil penelitian ini adalah :

1. Jumlah kombinasi yang terbentuk berbanding lurus dengan jumlah data, untuk nilai *minimum support* yang sama.
2. Semakin besar *minimum support*, semakin sedikit kombinasi yang terbentuk, dan semakin cepat waktu proses yang dibutuhkan.

Saran

Saran-saran yang bisa disampaikan adalah sebagai berikut:

1. Aplikasi ini masih bisa dikembangkan untuk algoritma *association rules* lainnya.
2. Dalam aplikasi ini, kombinasi yang terbentuk hanyalah terdiri dari dua item. Hal ini bisa dikembangkan lagi untuk kombinasi yang lebih.
3. Tampilan aplikasi ini masih bisa dibuat lebih menarik lagi.

DAFTAR PUSTAKA

- [1] Dominguez, A.K., Yacef, K., Curran, J.R. Data Mining for *Individualised Hints in eLearning*. *Educational Data Mining Conference Proceedings*. 1: 91-100. 2010.
- [2] Handojo, A., Budhi, G.S., Dwiyo, N.A. *A Decision Support System for "De Joglo" Restaurant Using Frequent Pattern Tree Data Mining*. *Proceedings Makasar International Conference on Electrical Engineering and Informatics*. 1: 118-124. 2010.
- [3] Erwin. *Analisis Market Basket Dengan Algoritma Apriori dan FP-Growth*. *Jurnal Generic* Vol 4: 26-30. 2019.
- [4] Hand, D., Mannila, H., Smyth, P. *Principles of Data Mining*. Cambridge: A Bradford Book. The MIT Press. 2001
- [5] Witten, I. H. and Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques 2nd Edition*. Morgan Kaufmann Publisher. 2005

- [6] Ulmer, D. 2002. *Mining an Online Auctions Data Warehouse*. The Mid-Atlantic Student Workshop on Programming Languages and Systems. New York: Pace University.
- [7] Samuel, D. *Penerapan Struktur FP-Tree dan Algoritma FP-Growth dalam Optimasi Penentuan Frequent Itemset*. Institut Teknologi Bandung. 2008.
- [8] Wahyuni, F. S., Siahaan, D. O., Fatichah. C. *Penggunaan Cluster-Based Sampling Untuk Penggalan Kaidah Asosiasi Multi Obyektif*. KURSOR. 5: 59-67. 2009.
- [9] Han, J., Pei, J. 2004. *Mining Frequent Patterns without Candidate Generation: A Frequent-Itemset Tree Approach*. Proceedings of MASPLAS. 8: 1-10. 2002
- [10] Ruldeviyani, Y., Fahrian, M. *Implementasi Algoritma-Asosiasi, Sebagai Bagian Dari Pengembangan Data Mining Algorithms Collection*. Konferensi Nasional Sistem dan Informatika. 43: 244-248. 2008.