

PENGGALIAN POLA JAWABAN SOAL UJIAN DENGAN MENGGUNAKAN ALGORITMA *FREQUENT PATTERN GROWTH*

Budanis Dwi Meilani, Abdus Tomi

Jurusan Teknik Informatika, Fakultas Teknologi Informasi

Institut Teknologi Adhi Tama Surabaya

Jl. Arief Rahman Hakim 100 Surabaya

Email : dwimeilanibudanis@yahoo.com , tomymachoo@yahoo.co.id

ABSTRAK

E-learning pada saat ini sudah terbilang pesat, hampir semua pihak telah mengerti tentang *e-learning*. Dan juga telah banyak sekolah yang menerapkannya. Namun, pengajar memanfaatkan hanya untuk menampilkan materi dan mengambil nilai dari ujian yang diberikan. Pengajar belum bisa mengetahui materi yang sebenarnya sulit untuk diterima dan dipahami semua siswa. Oleh karenanya, dalam penelitian ini akan menentukan pola-pola materi yang kurang dipahami siswa sehingga pihak pengajar akan menambah materi dan latihan soal pada materi tersebut. Data-data yang digunakan adalah data-data jawaban yang salah dari siswa yang melakukan ujian. Mata pelajaran yang digunakan dalam penelitian ini adalah matematika. Matematika merupakan salah satu mata pelajaran yang dianggap sulit oleh sebagian siswa. Algoritma yang digunakan dalam aplikasi ini adalah *Frequent Pattern Growth*, yaitu algoritma dalam *association rules* untuk mencari hubungan kombinasi dari setiap data. Hasil yang dicapai adalah pola materi yang kurang dipahami siswa. Dari pola itu akan membantu para pengajar dalam mengukur perkembangan siswa dan membantu siswa sendiri dalam mengerjakan soal yang dianggap sulit. Data yang diujicobakan adalah jawaban salah dari 122 siswa. Soal yang diujicobakan 200 nomor dari 10 materi yang ada. Hasil penelitian menghasilkan pola yang sering muncul adalah materi persamaan kuadrat, persamaan linier dua perubahan dan Aritmetika Sosial. Selain itu menghasilkan kesimpulan 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 : *Data mining, Association Rules, FP-Growth, Tree*

ABSTRACT

E-learning at the moment is fairly rapid, almost all parties have to understand about e-learning. And also has a lot of schools that implement it. However, teachers utilize only to showcase the material and take the value of the test is given. Teachers can not know the material sebenarnya difficult to be accepted and understood by all students. Therefore, in this study will determine the patterns are poorly understood materials so that the student teacher will add material and exercises on such material. The data used is the data wrong answers from students who do the exams. The subjects used in this study is mathematics. Mathematics is one of the subjects that are considered difficult by most students. The algorithm used in this application is the Frequent Pattern Growth, namely the association rules algorithm to search for combinations of any data connection. The results achieved are poorly understood patterns of student material. From the pattern that will assist teachers in measuring student progress and assist students in working on their own are considered difficult. Data were tested was one of 122 students answer. Tested about 200 numbers of 10 existing material. The results of the study resulted in a pattern that often arises is the material quadratic equations, linear equations and two changes of Social Arithmetic. In addition, lead to the conclusion that the more data, the combination formed itemset also be more and more, and the processing time will be longer. However, the greater the minimum support is entered, then the combination is formed will be less. Keyword : Data Mining, Association Rules, FP-Growth, Tree

Keyword : *Data mining, Association Rules, FP-Growth, Tree*

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 teknologi informasi 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 (*contents*) dan sistemnya. Hal tersebutlah yang mendorong penulis untuk meneliti dan membuat aplikasi untuk memanfaatkan *e-learning* lebih baik lagi. Pada penelitian ini akan 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 dan pengajar untuk lebih mempelajari materi yang sulit. Aplikasi ini menggunakan algoritma FP-Growth, yakni sebuah algoritma dalam Association Rule untuk menemukan relasi antar data. Beberapa hal yang menjadi permasalahan dalam penelitian ini dapat diuraikan sebagai berikut:

1. Bagaimana merancang sistem untuk menghasilkan pola jawaban siswa yang terbentuk dengan algoritma FP-Grwoth?
2. Bagaimana membuat aplikasi dengan menggunakan *Frequent Pattern-Growth (FP-Growth)*?

TINJAUAN PUSTAKA

Data Mining

Data Mining (DM) adalah salah satu bidang yang berkembang pesat karena besarnya kebutuhan akan nilai tambah dari database skala besar yang makin banyak terakumulasi sejalan dengan pertumbuhan teknologi informasi. Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam database. Dilihat dari sudut pandang bahasa, data minning diartikan sebagai penambangan data. Definisi umum dari DM itu sendiri adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data.

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. Konsep ini sendiri diturunkan dari terminologi market basket analysis, yaitu pencarian hubungan dari beberapa produk di dalam transaksi pembelian .

Association rule meliputi dua tahap, yaitu mencari kombinasi yang paling sering terjadi dari suatu itemset. Dan mendefinsikan *Condition* dan *Result* (untuk *conditional association rule*). *Association Rules* juga bermanfaat untuk pemakaian data web yang berdasarkan personalitas. Pendekatan ini diadopsi dari hubungan dengan *collaborative filtering*.

Dalam menentukan suatu *association rule*, terdapat suatu *interestingness measure* (ukuran kepercayaan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu . 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).

$$\begin{aligned} \text{Support} &= P(A \cup B) \\ &= \frac{\text{jumlah transaksi yang mengandung } A \text{ dan } B}{\text{total transaksi}} \end{aligned}$$

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

$$\begin{aligned} \text{Confidence} &: \\ &= \frac{\text{jumlah transaksi yang mengandung } A \text{ dan } B}{\text{jumlah transaksi yang mengandung } A} \end{aligned}$$

FP-Growth

Algoritma FP-Growth merepresentasikan transaksi dengan menggunakan struktur data *FP-Tree*. Cara kerja dari algoritma ini adalah melalui *scan database* yang dilakukan hanya dua kali saja. Data kemudian ditampilkan dalam bentuk *FP-tree*. Algoritma *FP-Growth* menggunakan struktur data *FP-Tree*. Ketika selesai membuat *FP-Tree*, tidak begitu saja bisa mendapatkan *frequent itemset* yang terdapat dalam *dataset*. Suatu kombinasi *itemset* bisa saja berada di beberapa *path* yang berbeda.

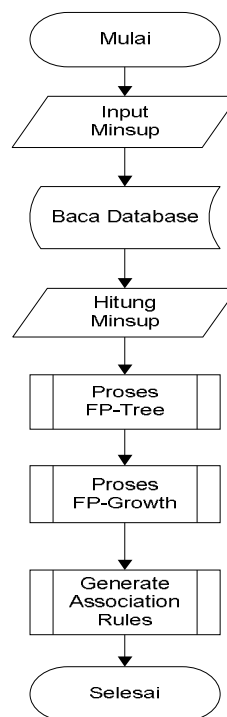
Untuk mendapatkan suatu *pattern* dalam *FP-Growth* langkah yang lebih mudah adalah mencari arah dari ujung suatu *path*, kemudian kita mencari mulai dari *header* untuk item di ujung tersebut, barulah kemudian dibuat berdasarkan tiap node berisi item tersebut dicari arah *path* node ke atas.

ANALISA DAN PERANCANGAN SISTEM

Deskripsi Sistem

Sistem yang dirancang pada penelitian ini adalah sistem untuk mencari kombinasi dari record yang ada dalam tabel database. Sistem yang dibangun menggunakan metode *Association Rule* dengan menggunakan algoritma FP-Growth dalam penentuan pola jawaban siswa.

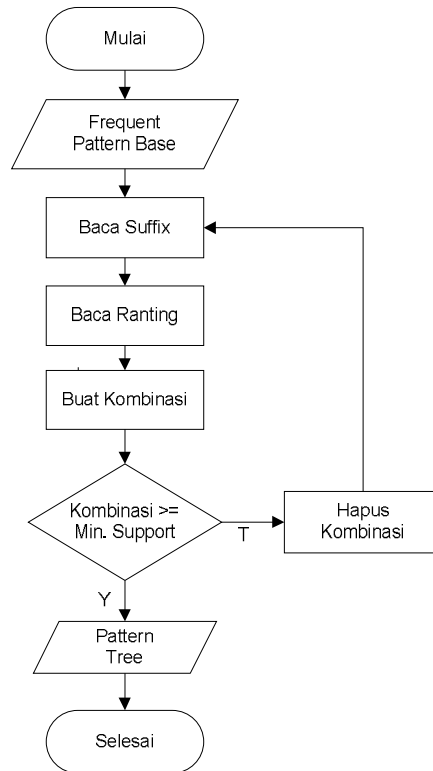
Flowchart Algoritma :



Gambar 1. *Flowchart* sistem utama

Gambar 1 adalah *flowchart* sistem utama. *Flowchart* sistem utama menjelaskan bagaimana proses *association rules* dari awal, mulai dari menginputkan *minimum support* yang kemudian akan dilanjutkan dengan membangkitkan *FP-Tree*, *FP-Growth*, dan terakhir dengan penghitungan *support* dan *confidence*. *flowchart* algoritma *frequent pattern growth* dari aplikasi yang akan dibangun.

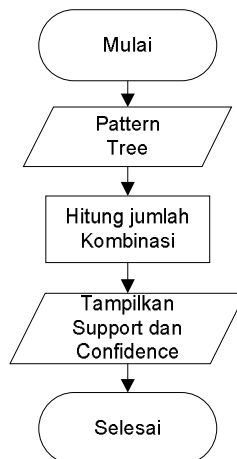
Hasil pembentukan *FP-Tree* diproses dengan dibangkitkan kembali menggunakan *FP-Growth* untuk kemudian dikombinasikan menurut *suffix* (akhiran) dari *FP-Tree* yang dihasilkan dalam setiap transaksi sehingga *output* berupa *subsets* (kombinasi) dapat digunakan dan dihitung *support* dan *confidence*-nya. Proses pembangkitan *FP-Growth* dapat dilihat pada Gambar 2.



Gambar 2. Flowchart FP-Growth

Proses Association Rules

Setelah didapat *subsets* (kombinasi) melalui pembangkitan *FP-Growth*, kemudian *subsets* (kombinasi) tersebut dihitung *support* atau nilai penunjang, frekuensi kemunculan item dalam suatu transaksi, dan *confidence* atau nilai penunjang kemunculan item secara bersamaan dalam suatu transaksi. Proses perhitungan *support* dan *confidence* dapat dilihat pada Gambar 3.



Gambar 3. Flowchart Association Rules

HASIL DAN PENGUJIAN SISTEM

Uji Coba Program

Perhitungan *asociatiom rules* yang ada dalam e-leraning ini hanya bisa digunakan oleh guru. Sedangkan siswa hanya bisa mendapatkan *feedback* berupa *hints* yang nanti akan muncul di setiap pertanyaan yang salah

dijawab. *Hints* yang muncul berupa link materi yang mempunyai kedekatan pola yang telah dihitung dengan *association rules* sebelumnya. Berikut adalah skenario analisa perhitungan dari 122 orang siswa.



Gambar 1. Tampilan Utama Halaman Admin

Hasil Uji Coba

Pola	Support	Confidence
Persamaan Kuadrat-Persamaan Linier Dua Peubah	63.636363636364	79.381443298969
Persamaan Kuadrat-Aritmetika Sosial	63.636363636364	79.381443298969
Persamaan Linier Dua Peubah-Relasi dan Fungsi	62.809917355372	80
Persamaan Linier Dua Peubah-Aritmetika Sosial	61.98347107438	78.947368421053
Persamaan Kuadrat-Relasi dan Fungsi	61.98347107438	77.319587628866
Aritmetika Sosial-Persamaan Linier Dua Peubah	61.98347107438	78.947368421053
Aritmetika Sosial-Relasi dan Fungsi	60.330578512397	76.842105263158
Persamaan Kuadrat-Himpunan	60.330578512397	75.257731958763
Relasi dan Fungsi-Himpunan	57.851239669421	75.268817204301
Garis-garis Sejajar-Operasi Bentuk Aljabar	57.02479338843	77.52808988764
Persamaan Kuadrat-Persamaan Linier Dua Peubah-Relasi dan Fungsi	51.239669421488	80.519480519481

Gambar 4. Tampilan semua nilai *support* dan *confidence*

Setelah dilakukan perhitungan terhadap beberapa skenario ujicoba berdasarkan jumlah data yang berbeda, yaitu 5, 25, 50, 75, 100 dan 120 siswa, maka hasil analisa terlihat pada tabel 4.1. Partisi data dilakukan secara urut sesuai dengan id siswa, dimulai dari id siswa yang terkecil hingga yang terbesar. Sehingga dari partisi data yang dilakukan, akan diperoleh suatu perbandingan seperti yang terlihat pada table 4.1.

Table 2. Tabel Perbandingan

No	Jumlah Murid	Minimum Support	Jumlah Kombinasi
1	5	25%	4
		50%	3
		75%	1
2	25	25%	24
		50%	24
		75%	12
4	50	25%	49
		50%	49
		75%	22
	75	25%	74
		50%	74
		75%	8

5	100	25%	99
		50%	99
		75%	34
6	120	25%	120
		50%	120
		75%	17

Apabila jumlah murid yang melakukan kesalahan semakin banyak, maka kombinasi *itemset* yang terbentuk juga akan semakin banyak. Namun, apabila *minimum support* yang diinputkan semakin besar, maka kombinasi yang terjadi akan berbanding terbalik, yakni akan semakin sedikit.

KESIMPULAN

Setelah menyelesaikan perancangan dan pembuatan *elearning* serta implementasi sistem pada penelitian ini, maka dapat ditarik kesimpulan sebagai berikut:

1. Algoritma fp-tree dapat digunakan untuk menarik kesimpulan atau pola jenis materi yang tidak dikuasai siswa.
2. Dengan memanfaatkan *tree* penyimpanan dan pembentukan kombinasi lebih efisien.
3. FP-Tree akan bekerja dengan lebih cepat jika support dari kasus yang diminta bernilai tinggi, sehingga kasus yang diproses juga makin sedikit.
4. Dari hasil uji coba 122 siswa. Soal yang diujicobakan 200 nomor dari 10 materi yang ada. Hasil penelitian menghasilkan pola yang sering muncul adalah materi persamaan kuadrat, persamaan linier dua perubahan dan Aritmetika Sosial.

DAFTAR PUSTAKA

- Dominguez, A.K., Yacef, K., Curran, J.R. Data Mining for *Individualised Hints in eLearning*. *Educational Data Mining Conference Proceedings*. 1: 91-100. 2010.
- Erwin. *Analisis Market Basket Dengan Algoritma Apriori dan FP-Growth*. *Jurnal Generic Vol 4*: 26-30. 2019.
- Gunadi, G., Sensune, D. I. *Penerapan Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori dan Frequent Pattern Growth*. *Jurnal Telematika Mkom Vol. 4*: 1228-132. 2012.
- Handojo, A., Budhi, G.S., Dwiyono, 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.
- Larose, D. T. *Discovering Knowledge In Data: An Introduction to data Mining*. Hoboken: John Wiley & Sons. 2005
- Samuel, D. *Penerapan Struktur FP-Tree dan Algoritma FP-Growth dalam Optimasi Penentuan Frequent Itemset*. Institut Teknologi Bandung. 2008.
- Witten, I. H. and Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques 2nd Edition*. Morgan Kaufmann Publisher. 2005.