

PENERAPAN ASSOCIATION RULE MINING DENGAN ALGORITMA A-PRIORI UNTUK REKOMENDASI PEMINJAMAN BUKU PADA PERPUSTAKAAN DAERAH PROVINSI JAWA TENGAH

Sodikun

Abstract— Perpustakaan daerah Provinsi Jawa Tengah memiliki sistem informasi perpustakaan yang dapat memudahkan pengguna fasilitas untuk meminjam buku. Namun data peminjaman buku yang tercatat pada database sistem masih belum mampu diolah kembali secara maksimal. Hal ini dikarenakan belum adanya modul pengolahan kembali data peminjaman sehingga hanya akan terakumulasi dan tertimbun didalam database sistem dan tidak menghasilkan kembali informasi dan pengetahuan yang bermanfaat untuk mendukung kemajuan manajemen perpustakaan. Teknik Data mining telah banyak digunakan untuk mengatasi permasalahan yang ada salah satunya dengan penerapan algoritma A-Priori untuk menemukan aturan asosiasi yang terbentuk dari dataset peminjaman buku. Sehingga diketahui kaitan asosiasi antar judul buku yang dipinjam. Sehingga mempermudah peminjam untuk memilih buku yang akan dipinjam.

Kata kunci: Perpustakaan Jateng, peminjaman buku, data mining, Algoritma A-Priori, aturan asosiasi.

INTRODUCTION

anyaknya jumlah mahasiswa yang masuk tentu menjadi kebanggaan dari suatu Universitas, namun dengan jumlah mahasiswa yang banyak dimana jumlah mahasiswa masuk tidak sama dengan mahasiswa yang lulus. Akan mengakibatkan banyak masalah antara lain turunnya akreditasi pada perguruan tinggi tersebut, kualitas dari lulusan mahasiswa, serta menurunnya kinerja program studi. Dengan banyaknya mahasiswa yang lulus tidak tepat waktu akan menghambat akreditasi program. Oleh karena itu, pada penelitian ini penulis akan menerapkan Algoritma Naive Bayes Classifier untuk memprediksi masa studi mahasiswa Fakultas Ilmu Komputer Program Studi Teknik Informatika S1 dengan fitur seleksi menggunakan Algoritma Genetika untuk menentukan variabel yang akan digunakan.

I. TINJAUAN PUSTAKA

Penelitian tentang perekomendasi buku pada transaksi peminjaman di perpustakaan bukanlah pertama kali dilakukan. Sebelumnya sudah pernah dilakukan penelitian mengenai penerapan metode data mining dalam penelitian tersebut.

Berikut adalah penelitian-penelitian yang relevan dengan masalah tersebut adalah penelitian yang dilakukan oleh Thomas Richard Stif Wijaya [1], dengan judul “Aplikasi Data Mining Menggunakan Algoritma A-Priori Untuk Menentukan

Aturan Asosiasi Pada Transaksi Peminjaman Buku Perpustakaan Universitas Dian Nuswantoro Semarang”. Dalam penelitiannya membahas mengenai penggunaan algoritma A-Priori untuk menghasilkan aturan asosiasi yang terbentuk dari transaksi peminjaman buku yang ada di perpustakaan UDINUS Semarang yang telah diolah menjadi dataset pada periode 3-September-2013 sampai 22-Oktober-2013 (50 Hari) mempunyai jumlah transaksi sebanyak 827 transaksi peminjaman.

Nugroho Wandi, Rully A. Hendrawan, dan Ahmad Mukhlason [2], dengan judul “Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalan Association Rules Menggunakan Algoritma A-Priori (Studi Kasus Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur)”. Dalam penelitiannya membahas mengenai penggunaan algoritma A-Priori untuk menghasilkan rekomendasi buku dari transaksi peminjaman buku yang ada. Pada penelitian yang dilakukan, digunakan 45.108 data transaksi tahun 2009 sampai 2012. Batas minimum support yang diberikan sebesar 1 hingga 4, dan panjang itemset berbanding terbalik dengan jumlah minimum support yang diberikan, semakin besar nilai minimum support yang diberikan semakin kecil panjang jumlah itemset yang terbentuk.

Kennedi Tampubolon, Hoga Saragih dan Bobby Reza [3], dengan judul “Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan”, dalam penelitian membahas mengenai penyediaan alat-alat kesehatan pada apotek kelambir 2 Medan yang menghasilkan pola kombinasi itemsets yang bisa dijadikan sebagai informasi yang sangat berharga dalam pengambilan keputusan untuk mempersiapkan stok barang apa yg di perlukan kemudian. Penerapan Algoritma Apriori pada data teknik data mining sangat efisien dan dapat mempercepat proses pembentukan pola kombinasi itemset hasil penjualan alat-alat kesehatan di apotek kelambir 2 Medan yaitu dengan support dan confidence tertinggi adalah Stick Asam Urat – Stick Gula dan Stick Colestrol-Stick Gula.

II. DATA MINING

Menurut Tan (2006) mendefinisikan *data mining* sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah *data mining* kadang disebut juga *knowledge discovery* [4].

Data Mining adalah proses analisis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari database besar [5].

“Data mining merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database, dan visualisasi untuk pengenalan 56 permasalahan pengambilan informasi dari database yang besar.” Kemajuan luar biasa yg terus berlanjut dalam bidang data mining didorong oleh beberapa faktor, antara lain [6]:

1. Pertumbuhan data yang sangat cepat dalam pengumpulan data.
2. Penyimpangan data dalam data warehouse, sehingga seluruh perusahaan memiliki akses kedalam database yang handal.
3. Adanya peningkatan akses data melalui navigasi web dan intranet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk data mining (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

Dari definisi-definisi yang telah disampaikan di atas, hal penting yang berkaitan dengan *data mining* adalah :

1. *Data mining* merupakan suatu proses mencari atau mengali sebuah informasi yang bermanfaat terhadap data yang sudah ada.
2. Data yang digunakan untuk proses *mining* merupakan sebuah data yang sangat besar.
3. Tujuan *data mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat bagi pemilik data.

III. ALGORITMA A-PRIORI

Algoritma a-priori adalah suatu algoritma dasar yang diusulkan oleh Agrawal dan Srikant pada tahun 1994 untuk menentukan *Frequent itemsets* untuk aturan asosiasi Boolean. *Algoritma A-priori* termasuk jenis Aturan Asosiasi pada data mining. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut *affinity analysis* atau *market basket analysis*. Analisis asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan suatu kombinasi item. Salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi tinggi (*frequent pattern mining*). Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolok ukur, yaitu : *support* dan *confidence*. *Support* (nilai penunjang) adalah persentase kombinasi item tersebut dalam database, sedangkan *confidence* (nilai kepastian) adalah kuatnya hubungan antara-item dalam aturan asosiasi [8].

1. Pola frekuensi tinggi

Adalah tahap dimana pembentukan kombinasi antar item yang telah memenuhi syarat *minimum support* yang telah

ditentukan sebelumnya. Nilai *support* pada satu itemset dapat diperoleh dengan menggunakan sebuah rumus di bawah ini:

$$\text{Support}(W) = \frac{\sum \text{Transaksi yang mengandung } W}{\sum \text{transaksi yang telah terjadi}} \quad (1)$$

Rumus diatas memiliki arti untuk menentukan nilai *support* pada satu itemset jumlah transaksi yang mengandung item *W* dibagi dengan jumlah transaksi yang ada pada database.

Sedangkan pada dua itemset atau lebih diperoleh dengan rumus:

$$\text{Support}(W, X) = \frac{\sum \text{Transaksi mengandung } W \text{ dan } X}{\sum \text{Total transaksi yang terjadi}} \quad (2)$$

Pada rumus diatas untuk menentukan nilai *support* pada dua itemset atau lebih, jumlah transaksi yang mengandung item *W* dan *X* dibagi dengan total transaksi yang terjadi pada dataset.

2. Pembentukan Aturan Asosiasi

Langkah yang dilakukan setelah menentukan nilai *support* pada itemset berfrekuensi tinggi lalu dibentuk aturan asosiasi yang menyatakan kuatnya hubungan kombinasi itemset pada transaksi. Untuk menentukan aturan asosiasi yang terbentuk minimal itemset harus memiliki dua kandidat *W* dan *X*. Pada rules yang terbentuk berlaku hukum asosiatif $W \rightarrow X$ tidak berlaku $X \rightarrow W$. Untuk menentukan aturan $W \rightarrow X$ digunakan rumus:

$$\text{Confidence}(W \rightarrow X) = \frac{\sum \text{Transaksi mengandung } W \text{ dan } X}{\sum \text{Transaksi mengandung } W}$$

Pada rumus diatas untuk menentukan nilai *confidence* pada aturan asosiasi, jumlah transaksi yang mengandung item *W* dan *X* dibagi dengan jumlah transaksi yang mengandung item *W* [7].

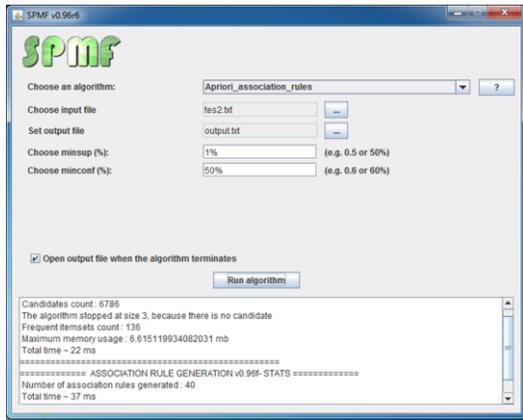
IV. SEQUENTIAL PATTERN MINING FRAMEWORK (SPMF)

Sequential Pattern Mining Framework (SPMF) adalah fungsi pustaka (library) yang ditulis pada bahasa pemrograman Java untuk menangani tugas data mining dengan lisensi open-source GPL v3. SPMF memiliki 52 koleksi algoritma yang dapat digunakan untuk:

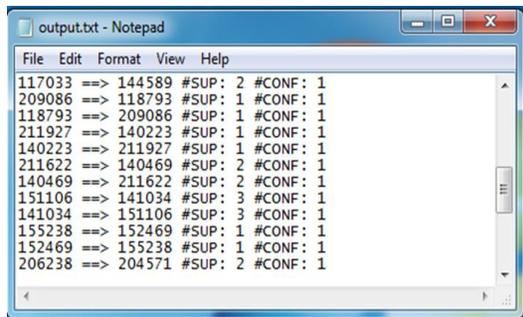
- Sequential Pattern Mining.
- Association Rule Mining.
- Frequent Itemset Mining.
- Sequential Rule Mining
- Clustering

Source code untuk tiap algoritma yang tersedia pada fungsi pustaka dapat diintegrasikan pada program Java yang dibuat oleh pengembang software.

V. PENGUJIAN SPMF



Gambar 1 Pengujian SPMF



Gambar 2 Hasil Training

Tabel 1 hasil Pengujian

Kode Judul Buku Aturan Asosiasi	Nilai Support	Nilai Confidence
63515 → 61788 Jika meminjam buku Jalan ketiga pembaharuan demokrasi sosial maka juga meminjam buku Protokol Kyoto : Implikasinya bagi negara berkembang	2/90 2,22%	0.5 (50%)
148284 → 61788 Jika meminjam buku Manajemen Publik maka juga meminjam buku Protokol Kyoto : Implikasinya bagi negara berkembang	2/90 2,22%	0.5 (50%)
65849 → 63810 Jika meminjam buku Ada rindu di mata peri maka juga meminjam buku Konspirasi hari kiamat = the doomsday conspiracy	2/90 2,22%	1 (100%)
211622 → 140469 Jika meminjam buku Gokil abiiz ! maka juga meminjam buku Seni Membaca Wajah : Membaca Karakter Orang	2/90 2,22%	0,66667 66,667%

dengan Wajah / Adhe Firmansyah		
210308 → 148244 Jika meminjam buku Totto- Chan maka juga meminjam buku Pencari harta karun	2/90 2,22%	1 (100%)

VI. KESIMPULAN

Berdasarkan Dari permasalahan di atas tentang data peminjaman buku perpustakaan daerah provinsi Jawa Tengah dapat diselesaikan dengan teknik data mining menggunakan aturan asosiasi untuk menentukan buku-buku yang sering dipinjam bersamaan. Software data mining yang penulis gunakan dapat memudahkan penyelesaian tugas data mining dengan menggunakan algoritma A-Priori Association Rules. Sehingga software data mining tersebut bisa di gunakan di perpustakaan daerah provinsi Jawa Tengah.

VII. SARAN

Jumlah buku yang dapat ditraining ditingkatkan agar pola mining yang terbentuk semakin bervariasi dan banyak. Sehingga prose mining lebih akurat dalam menentukan keterkaitan antara buku – buku yang dipinjam.

Beberapa ide gagasan yang dapat digunakan untuk penelitian selanjutnya adalah sebagai berikut:

1. Penelitian selanjutnya dapat dikembangkan pada jenis data yang sama dengan menggunakan metode yang lebih baik seperti algoritma FP-GROWTH.
2. Hasil aturan asosiasi yang terbentuk dapat dikembangkan menjadi basis pengetahuan untuk membuat aplikasi yang lebih bagus mengenai hubungan antara buku – buku yang ada dalam perpustakaan.

REFERENCES

- [1] T. R. S. Wijaya, "Aplikasi Data Mining Menggunakan Algoritma A-Priori Untuk Menentukan Aturan Asosiasi Pada Transaksi Peminjaman Buku Perpustakaan Universitas Dian Nuswantoro Semarang," Skripsi Teknik Informatika Universitas Dian Nuswantoro, Semarang, 2014.
- [2] W. e. Nugroho, "Pengembangan Sistem Rekomendasi Penelusuran Buku dengan Penggalan Association Rule Menggunakan Algoritma Apriori (Studi Kasus Badan Perpustakaan dan Kearsipan Provinsi Jawa Timur)," *JURNAL TEKNIK ITS Vol. 1, (Sept, 2012) ISSN: 2301-9271*, pp. 1-5, 2012.
- [3] K. e. Tampubolon, "Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan," *Implementasi Data Mining Algoritma Apriori Pada Sistem Persediaan Alat-Alat Kesehatan*, pp. 1-14, 2013.
- [4] E. Prasetyo, *Data Mining Konsep dan Aplikasi Menggunakan MATLAB*, Yogyakarta: C.V ANDI OFFSET, 2012.
- [5] E. e. Turban, *Decision Support Systems and Intelligent Systems*, Yogyakarta: Andi Offset, 2005.

- [6] D. T. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining*, John Willey & Sons, Inc, 2005.
- [7] Kusrini and E. T. Luthfi, *Algoritma Data Mining*, Yogyakarta: Andi Offset, 2009.
- [8] B. Gregorius S., "Algoritma Porter Stemmer For Bahasa Indonesia Untuk Pre-Processing Text Mining Berbasis Metode Market Basket Analisis," *Algoritma Porter Stemmer For Bahasa Indonesia Untuk Pre-Processing Text Mining Berbasis Metode Market Basket Analisis*, vol. 63, pp. 1-10, 2006.