

BAB II

LANDASAN TEORI

2.1 Perpustakaan

Menurut UU Perpustakaan pada Bab I pasal 1 menyatakan Perpustakaan adalah institusi yang mengumpulkan pengetahuan tercetak dan terekam, mengelolanya dengan cara khusus guna memenuhi kebutuhan intelektualitas para penggunanya melalui beragam cara interaksi pengetahuan.

Dalam arti tradisional, perpustakaan adalah sebuah koleksi buku dan majalah. Walaupun dapat diartikan sebagai koleksi pribadi perseorangan, namun perpustakaan lebih umum dikenal sebagai sebuah koleksi besar yang dibiayai dan dioperasikan oleh sebuah kota atau institusi, dan dimanfaatkan oleh masyarakat yang rata-rata tidak mampu membeli sekian banyak buku atas biaya sendiri. Tetapi, dengan koleksi dan penemuan media baru selain buku untuk menyimpan informasi, banyak perpustakaan kini juga merupakan tempat penyimpanan dan/atau akses ke map, cetak atau hasil seni lainnya, mikrofilm, mikrofiche, tape audio, CD, LP, tape video dan DVD, dan menyediakan fasilitas umum untuk mengakses gudang data CD-ROM dan internet.

Perpustakaan dapat juga diartikan sebagai kumpulan informasi yang bersifat ilmu pengetahuan, hiburan, rekreasi, dan ibadah yang merupakan kebutuhan hakiki manusia. Oleh karena itu perpustakaan modern telah didefinisikan kembali sebagai tempat untuk mengakses informasi dalam format apa pun, apakah informasi itu disimpan dalam gedung perpustakaan tersebut atau tidak. Dalam perpustakaan modern ini selain kumpulan buku tercetak, sebagian buku dan koleksinya ada dalam perpustakaan digital (dalam bentuk data yang bisa diakses lewat jaringan komputer).

2.2 Peran Perpustakaan

Perpustakaan merupakan upaya untuk memelihara dan meningkatkan efisiensi dan efektifitas proses belajar-mengajar. Perpustakaan yang

terorganisir secara baik dan sistematis, secara langsung atau pun tidak langsung dapat memberikan kemudahan bagi proses belajar mengajar di sekolah tempat perpustakaan tersebut berada. Hal ini, terkait dengan kemajuan bidang pendidikan dan dengan adanya perbaikan metode belajar-mengajar yang dirasakan tidak bisa dipisahkan dari masalah penyediaan fasilitas dan sarana pendidikan.

2.3 Tujuan Perpustakaan

Tujuan perpustakaan adalah untuk membantu masyarakat dalam segala umur dengan memberikan kesempatan dengan dorongan melalui jasa pelayanan perpustakaan agar mereka :

- a. Dapat mendidik dirinya sendiri secara berkesimbangan.
- b. Dapat tanggap dalam kemajuan pada berbagai lapangan ilmu pengetahuan, kehidupan sosial dan politik.
- c. Dapat memelihara kemerdekaan berfikir yang konstruktif untuk menjadi anggota keluarga dan masyarakat yang lebih baik.
- d. Dapat mengembangkan kemampuan berfikir kreatif, membina rohani dan dapat menggunakan kemampuannya untuk dapat menghargai hasil seni dan budaya manusia.
- e. Dapat meningkatkan tarap kehidupan sehari-hari dan lapangan pekerjaannya.
- f. Dapat menjadi warga negara yang baik dan dapat berpartisipasi secara aktif dalam pembangunan nasional dan dalam membina saling pengertian antar bangsa.
- g. Dapat menggunakan waktu senggang dengan baik yang bermanfaat bagi kehidupan pribadi dan social.

2.4 Jenis- Jenis Perpustakaan

Menurut Sutarno NS (2006) jenis-jenis perpustakaan di Indonesia adalah sebagai berikut:

1. Perpustakaan Nasional RI

Merupakan Perpustakaan Nasional yang berkedudukan di Ibu Kota Negara Indonesia yang mempunyai jangkauan dan ruang lingkup secara Nasional dan merupakan salah satu Lembaga Pemerintah Non Departemen (LPND) yang bertanggung jawab kepada Presiden.

2. Badan Perpustakaan Daerah

Badan perpustakaan daerah atau lembaga lain yang sejenis adalah yang berkedudukan di tiap provinsi di Indonesia yang mengelola perpustakaan.

3. Perpustakaan Umum

Perpustakaan umum diibaratkan sebagai Universitas Rakyat atau Universitas Masyarakat, maksudnya adalah bahwa perpustakaan umum merupakan lembaga pendidikan bagi masyarakat umum.

4. Perpustakaan Perguruan Tinggi

Perpustakaan yang berada di Perguruan Tinggi, baik berbentuk Universitas, Akademi, Sekolah Tinggi, ataupun Institut. Keberadaan, tugas dan fungsi perpustakaan tersebut adalah dalam rangka melaksanakan Tri Dharma Perguruan Tinggi, meliputi pendidikan, penelitian / riset dan pengabdian kepada masyarakat.

5. Perpustakaan Sekolah

Perpustakaan sekolah berada di sekolah, dikelola sekolah, dan berfungsi untuk sarana kegiatan belajar mengajar, penelitian sederhana, menyediakan bahan bacaan, dan tempat rekreasi.

6. Perpustakaan Khusus

Perpustakaan khusus berada pada lembaga-lembaga pemerintahan dan swasta. Perpustakaan tersebut diadakan sebagai sumber informasi dan ilmu pengetahuan yang berkaitan baik langsung maupun tidak langsung dengan instansi induknya.

7. Perpustakaan Lembaga Keagamaan

Merupakan perpustakaan yang dimiliki dan dikelola oleh lembaga-lembaga keagamaan, misalnya perpustakaan, masjid, gereja.

8. Perpustakaan Internasional

Perpustakaan Internasional Merupakan perpustakaan internasional yang memiliki koleksi yang menyangkut negara-negara anggota atau negara-negara yang berafiliasi kepada lembaga dunia tersebut. Perpustakaan ini dikelola dan diselenggarakan lembaga internasional.

9. Perpustakaan Kantor Perwakilan Negara-negara Asing

Merupakan perpustakaan yang dimiliki dan diselenggarakan oleh lembaga / kantor perwakilan Negara masing-masing. Contohnya perpustakaan lembaga kebudayaan amerika dan pusat kebudayaan jepang

10. Perpustakaan Pribadi / Keluarga

Merupakan perpustakaan yang dimiliki dan dikelola oleh perorangan atau orang-orang tertentu bersama anggota keluarganya.

11. Perpustakaan Digital

Perpustakaan digital bukan merupakan salah satu jenis perpustakaan yang berdiri sendiri, tetapi merupakan pengembangan dalam system pengelolaan dan layanan perpustakaan.

2.5 Data Mining

Secara sederhana, *data mining* merupakan ekstraksi informasi yang tersirat dalam sekumpulan data. *Data mining* merupakan sebuah proses untuk menggali kumpulan data dan menemukan informasi di dalamnya. (Turban, E., dkk. 2005). *Data mining* merupakan proses pengekstrakan informasi dari jumlah kumpulan data yang besar dengan menggunakan algoritma dan teknik gambar dari statistik, mesin pembelajaran dan sistem manajemen *database*. Penggalian data ini dilakukan pada sekumpulan data yang besar untuk menemukan pola atau hubungan yang ada dalam kumpulan data tersebut (Kusrini dan E.T. Lutfi. 2009). Hasil penemuan yang diperoleh setelah proses penggalian data ini, kemudian dapat digunakan untuk analisis yang lebih lanjut.

Data mining yang disebut juga dengan *Knowledge-Discovery in Database* (KDD) adalah sebuah proses secara otomatis atas pencarian data di dalam sebuah memori yang amat besar dari data untuk mengetahui pola

dengan menggunakan alat seperti klasifikasi, hubungan (*association*) atau pengelompokan (*clustering*). Proses KDD ini terdiri dari langkah-langkah sebagai berikut :

1. *Data Cleaning*, proses menghapus data yang tidak konsisten dan kotor.
2. *Data Integration*, penggabungan beberapa sumber data.
3. *Data Selection*, pengambilan data yang akan dipakai dari sumber data.
4. *Data Transformation*, proses dimana data ditransformasikan menjadi bentuk yang sesuai untuk diproses dalam data mining.
5. *Data Mining*, suatu proses yang penting dengan melibatkan metode untuk menghasilkan suatu pola data.
6. *Pattern Evaluation*, proses untuk menguji kebenaran dari pola data yang mewakili *knowledge* yang ada didalam data itu sendiri.
7. *Knowledge Presentation*, proses visualisasi dan teknik menyajikan *knowledge* digunakan untuk menampilkan *knowledge* hasil *mining* kepada *user*.

2.6 Metode *Data Mining*

Pada umumnya metode *data mining* dapat dikelompokkan kedalam dua kategori yaitu *deskriptif* dan *prediktif*. Metode *deskriptif* bertujuan untuk mencari pola yang dapat dimengeti oleh manusia yang menjelaskan karakteristik dari data. Metode *prediktif* menggunakan ciri-ciri tertentu dari data untuk melakukan prediksi. Metode-metode yang ada dalam *data mining* adalah sebagai berikut:

1. *Classification*

Klasifikasi (*Classification*) merupakan proses untuk menemukan sekumpulan model yang menjelaskan dan membedakan kelas-kelas data, sehingga model tersebut dapat digunakan untuk memprediksi nilai suatu kelas yang belum diketahui pada sebuah objek. Untuk mendapatkan model, kita harus melakukan analisis terhadap data latih (*training set*). Sedangkan data uji (*test set*) digunakan untuk mengetahui tingkat akurasi dari model yang telah dihasilkan. Dalam proses klasifikasi pohon

keputusan tradisional, fitur (atribut) dari tupel adalah kategorikal atau numerikal. Biasanya definisi ketepatan nilai (*point value*) sudah didefinisikan di awal. Pada banyak aplikasi nyata, terkadang muncul suatu nilai yang tidak pasti. Klasifikasi dapat digunakan untuk memprediksi nama atau nilai kelas dari suatu objek data. Metode inilah yang digunakan dalam tugas akhir ini.

2. *Clustering*

Pengelompokkan (*Clustering*) merupakan proses untuk melakukan segmentasi. Digunakan untuk melakukan pengelompokkan secara alami terhadap atribut suatu set data. Termasuk kedalam *unsupervised task*. Contoh *clustering* seperti mengelompokkan dokumen berdasarkan topiknya.

3. *Association*

Tujuan dari metode ini yaitu untuk menghasilkan sejumlah *rule* yang menjelaskan sejumlah data yang terhubung kuat satu dengan yang lainnya. Sebagai contoh *association analysis* dapat digunakan untuk menentukan produk yang datang dibeli secara bersamaan oleh banyak pelanggan atau bisa juga disebut dengan *market basket analysis*.

4. *Regression*

Regression mirip dengan klasifikasi. Perbedaan utamanya adalah terletak pada atribut yang diprediksi berupa nilai yang kontinyu.

5. *Forecasting*

Prediksi (*Forecasting*) berfungsi untuk melakukan prediksi kejadian yang akan datang berdasarkan data sejarah yang ada.

6. *Sequence Analysis*

Tujuan dari metode ini adalah untuk mengenali pola dari data diskrit. Sebagai contoh adalah menemukan kelompok *gen* dengan tingkat ekspresi yang mirip.

7. *Deviation Analysis*

Tujuan dari metode ini adalah untuk menemukan penyebab perbedaan antara data yang satu dengan data yang lain dan biasa disebut

sebagai *oulier detection*. Sebagai contoh adalah apakah sudah terjadi penipuan terhadap pengguna kredit dengan melihat catatan transaksi yang tersimpan dalam basis data perusahaan kartu kredit. (Santosa, Budi. 2007).

2.7 Assosiation Rules

Analisis asosiasi berguna untuk menemukan hubungan penting yang tersembunyi di antara set data yang sangat besar. Hubungan yang sudah terbuka dipresentasikan dalam bentuk aturan asosiasi atau set aturan *item* yang sering muncul.

Isu penting dalam analisis asosiasi adalah bagaimana cara menemukan pola tertentu dari data yang berjumlah sangat besar, yang membuat biaya komputasi menjadi sangat mahal. Isu penting lainnya adalah bahwa diantara pola-pola yang harus dipelajari dalam teknik analisis asosiasi, ada banyak pola yang akhirnya dibuang karena tidak termasuk dalam kriteria pola yang dicari.

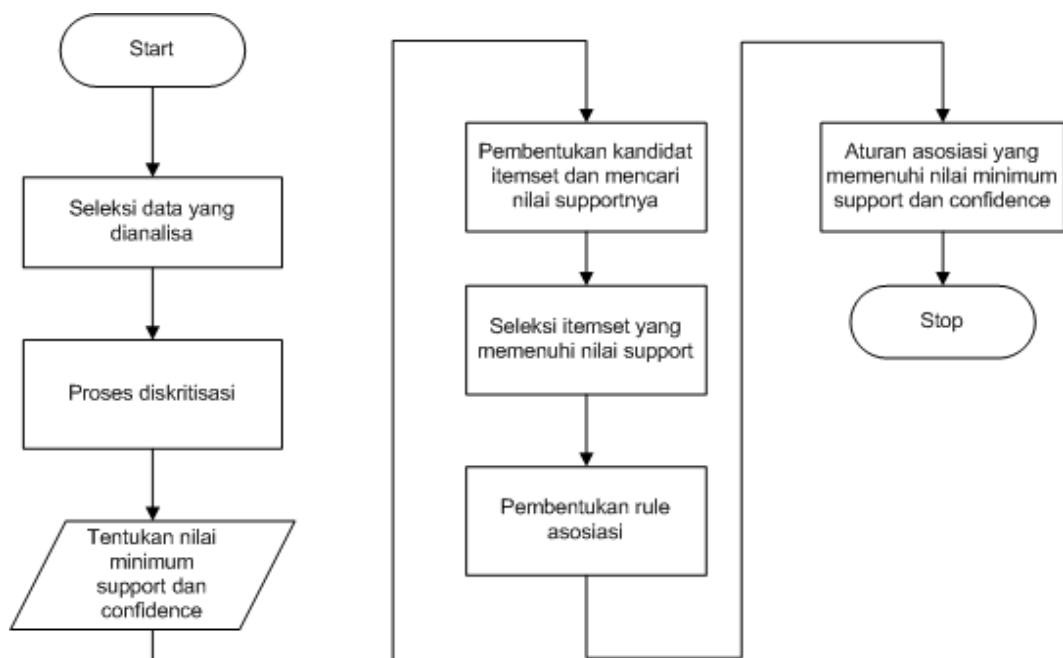
Algoritma Apriori menggunakan pendekatan level yang baik dalam pembangkitan asosiasi, dimana setiap level berhubungan dengan sejumlah *item* yang dimiliki konsekuena aturan. Awalnya, aturan dengan *confidence* yang tinggilah yang mempunyai satu *item* diekstrak. Aturan ini kemudian digunakan untuk membangkitkan kandidat aturan yang baru (Prasetyo, 2012).

Dalam algoritma apriori semakin kecil nilai minimum *support*, semakin banyak *rule* yang dihasilkan tetapi tidak semua *rule* yang dihasilkan valid, yang menunjukkan bahwa nilai minimum *support* berpengaruh pada pembentukan *rule* tetapi tidak banyak berpengaruh terhadap kevalidan *rule*. Semakin kecil nilai minimum *confidence*, semakin banyak *rule* yang dihasilkan dan besar kemungkinan *rule* tersebut valid, yang menunjukkan bahwa nilai minimum *confidence* berpengaruh pada kevalidan *rule*. Dari *rule-rule* yang memenuhi nilai *confidence* dihitung nilai lift atau interest factor rasionya, *rule* yang memiliki nilai lift rasio atau interest factor lebih dari 1 maka *rule* tersebut valid dan bisa dijadikan rekomendasi.

Confidence = jumlah transaksi dengan *item* dalam *antedent* (A) dan *concequent* (B) dibagi jumlah transaksi dengan *item* dalam *antedent* (A) dan biasanya pada istilah *antedecedent* untuk mewakili bagian “jika” dan *concequent* untuk mewakili bagian “maka”.

2.8 Tahapan *Assosiation Rule*

Analisis asosiasi dikenal juga sebagai salah satu teknik data mining yang menjadi dasar dari berbagai teknik data mining lainnya. Khususnya salah satu tahap dari analisis asosiasi yang disebut analisis pola frekuensi tinggi (*frequent pattern mining*) menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien. Berikut *flowchart* Apriori pada gambar 2.1.



Gambar 2.1 *Flowchart* Apriori

Dari gambar 2.1 dapat dijelaskan bahwa proses pertama adalah menyeleksi data sebagai data uji kemudian melakukan proses diskritisasi dimana merubah representasi atribut kategorikal menjadi biner pada setiap

interval. Selanjutnya metodologi dasar analisis asosiasi terbagi menjadi dua tahap :

1. Analisa pola frekuensi tinggi

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

- *Support* (dukungan)

Support merupakan suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu *item* atau *itemset* dari keseluruhan transaksi.

Nilai *support* 1 *item* diperoleh dengan rumus 1 berikut :

$$Support(A) = \frac{Jumlah\ Transaksi\ mengandung\ A}{Jumlah\ transaksi} \dots\dots\dots(2.1)$$

Nilai *support* 2 *item* diperoleh dengan rumus 2 berikut :

Sebagai contoh, *support* dari aturan "A=>B" adalah probabilitas kemunculan barang A dan B secara bersamaan.

$$Support(A,B) = \frac{Jumlah\ Transaksi\ mengandung\ A\ dan\ B}{Jumlah\ transaksi} \dots\dots\dots(2.2)$$

2. Pembentukan Aturan Asosiasi

- *Confidence* (tingkat kepercayaan)

Probabilitas kejadian beberapa barang dibeli bersamaan dimana salah satu produk sudah pasti dibeli. Sebagai contohnya, jika terdapat *n* transaksi dimana barang A dibeli dan ada *m* transaksi dimana barang A dan B dibeli secara bersamaan, maka *confidence* dari aturan *if A then B* adalah *m/n*.

Nilai *confidence* 2 *item* diperoleh dengan rumus 1 berikut :

$$Confident(A \rightarrow B) = \frac{Support\ A \cup B}{Support\ A} \dots\dots\dots(2.3)$$

3. Pada proses evaluasi sistem ini digunakan metode korelasi *Lift*. Dimana metode ini menguji korelasi dari *rule* yang telah dibentuk. Adapun korelasi *Lift* dapat dihitung dengan rumus seperti dibawah ini:

$$Lift(X, Y) = \frac{P(X \cup Y)}{P(X)P(Y)} \dots\dots\dots(2.4)$$

Dimana :

$Lift(X,Y) = korelasi\ Lift\ X,Y$

$P(XUY) =$ Jumlah kemunculan X dan Y dibagi dengan total transaksi

$P(X)P(Y) =$ Jumlah kemunculan X dikali dengan jumlah kemunculan Y pada total transaksi

Lift ratio digunakan untuk mengevaluasi kuat tidaknya sebuah aturan asosiasi (Santoso, 2007). *Lift ratio* adalah parameter penting selain *support* dan *confidence* dalam asosiasi *rule*. *Lift ratio* merupakan kevalidan proses bahwa dalam transaksi tersebut, set X dan set Y benar-benar muncul secara bersamaan. Apabila dari perhitungan tersebut menghasilkan nilai dibawah 1 maka terdapat korelasi *negatif*, untuk perhitungan yang menghasilkan nilai diatas 1 maka terdapat korelasi *positif*. Namun apabila menghasilkan nilai sama dengan 1 maka tidak ada korelasi antara X dan Y.

Berikut ini adalah langkah-langkah yang digunakan dalam algoritma apriori:

1. Pembentukan kandidat *itemset*.

Kandidat *k-itemset* dibentuk dari kombinasi $(k-1)$ -*itemset* yang didapat dari iterasi sebelumnya. Satu cara dari algoritma apriori adalah pemangkasan kandidat *k-itemset* yang subsetnya berisi $k-1$ *item* tidak termasuk dalam pola frekuensi tinggi dengan panjang $k-1$.

2. Penghitungan *support* dari tiap kandidat *k-itemset*.

Support dari tiap kandidat *k-itemset* didapat dengan menscan database untuk menghitung jumlah transaksi yang memuat semua *item* didalam kandidat *k-itemset* tersebut. Ini adalah juga ciri dari algoritma apriori dimana diperlukan penghitungan dengan cara seluruh database sebanyak *k-itemset* terpanjang.

3. Tetapkan pola frekuensi tinggi.

Pola frekuensi tinggi yang memuat k *item* atau *kitemset* ditetapkan dari kandidat *k-itemset* yang *supportnya* lebih besar dari minimum *support*.

2.9 Penelitian sebelumnya

Pada penelitian sebelumnya yang menggunakan metode Apriori adalah penelitian yang berjudul “Penerapan Metode Asosiasi Menggunakan Algoritma Apriori Pada Aplikasi Analisa Pola Belanja Konsumen (Studi Kasus Toko Buku Gramedia Bintaro)” oleh Imroatul Ulum. Adapun data yang diambil dalam penelitian ini adalah data penjualan Toko Buku Gramedia tahun 2016. Atribut yang digunakan adalah nomor transaksi, id *item*, jumlah, nama kategori. Pada penelitian tersebut menggunakan nilai *min support* = 5% dan *min conf* = 15% menghasilkan 7 pola dengan *strong rule* (pola yang memiliki nilai *support* dan nilai *confidence* relative tinggi) yang didapatkan adalah *Schoolbooks* Indonesia *Curriculum* → *Childrens Books* atau bisa dibaca jika konsumen membeli buku *Schoolbooks* Indonesia *Curriculum* maka konsumen juga membeli *Childrens Books* dengan nilai *support* 11,22% dan nilai *confidence* 30,66%. Setelah menemukan kombinasi pembelian, dalam penelitian tersebut juga dilakukan pengujian waktu. Pengujian waktu yang dilakukan menghasilkan kesimpulan bawa semakin banyak data yang diproses maka akan membutuhkan waktu eksekusi yang lebih lama.

Penelitian serupa dilakukan oleh Azwar Anas, mahasiswa Pendidikan Informatika STKIP PGRI Sumbar. Penelitian yang dilakukan berjudul “Analisa Pola Peminjaman Buku Perpustakaan Menggunakan Algoritma Apriori.” Data yang digunakan adalah data peminjaman Perpustakaan SMPN 24 Padang. Pada penelitian ini *itemset* yang dicari adalah 2 *itemset* dan 3 *itemset* dengan nilai *min support* sebesar 5% dan nilai *min confidence* sebesar 10%. Pada aturan *asosiasi* 2 *itemset* menghasilkan 9 *rule* yaitu {(MTK, IPA), (Agama, Penjaskes), (TIK, Agama), (TIK, IPA), (TIK, Penjaskes), (Penjaskes, IPS), (IPA, Agama), (IPA, TIK), (B. Inggris, TIK)}. Sedangkan untuk aturan 3 *itemset* menghasilkan {(Agama, TIK, Penjaskes), (TIK, IPA, Agama), (TIK, IPA, Penjaskes), (IPA, TIK, Agama)}.