

BAB II

TINJAUAN PUSTAKA

2.1. Kajian Pustaka

Kajian pustaka dilakukan untuk menyempurnakan konsep dan kerangka pemikiran, penelitian sebelumnya mendeskripsikan perkembangan hasil penelitian yang mempunyai kaitan dengan penelitian yang akan dilakukan. Dari pengkajian pustaka dapat diperoleh masalah-masalah yang belum terpecahkan atau belum terjawab dari penelitian sebelumnya.

Penelitian yang dilakukan oleh Moh. Sholik dan Abu Salam dalam jurnal yang berjudul “Implementasi Algoritma Apriori untuk Mencari Asosiasi Barang yang Dijual di E-commerce OrderMas“, Mengangkat masalah pengadaan stok barang oleh supplier OrderMas yang tidak terprediksi. Penelitian ini menggunakan metode Algoritma Apriori untuk membantu menyelesaikan masalah tersebut dengan cara mencari keterkaitan pembelian antar *item*. Data yang digunakan diperoleh dari set data trans nasional yang berisi transaksi toko ritel online di Inggris, setelah diproses menggunakan aplikasi RapidMiner dengan menentukan minimal support 50%, minimal confidence 50%, dan menentukan jumlah rule yang ditampilkan yaitu 10 rules, diperoleh 4 barang paling sering dibeli[9].

Penelitian yang dilakukan oleh Fitri Rahmawati dan Nita Merlina dalam jurnal yang berjudul “Metode *Data mining* Terhadap Data Penjualan Sparepart Mesin Fotocopy Menggunakan Algoritma Apriori”, mengangkat masalah penumpukan informasi rekap penjualan dan arsip. Penelitian ini menggunakan metode Algoritma Apriori untuk memanfaatkan informasi rekap penjualan dan arsip sehingga diperoleh informasi baru yang menjadi acuan atau rekomendasi penyusunan strategi pemasaran. Data diperoleh dari penjualan yang dilakukan oleh CV. Nusantara Copy. Setelah perhitungan menggunakan Algoritma Apriori selesai dilakukan, peneliti menguji perhitungan manual nya menggunakan aplikasi Tanagra. Kedua metode

yaitu manual dan digital menunjukkan nilai yang hampir sama, yaitu didapatkan 2 rule terkuat sesuai dengan minimal confidence dan minimal support yang ditentukan oleh peneliti[10].

Penelitian yang dilakukan oleh Nur Fitrianti Fahrudin dalam jurnal yang berjudul “Penerapan Algoritma Apriori untuk Market Basket Analysis”, mengangkat masalah penggunaan algoritma apriori terhadap Market Basket Analysis. Data yang digunakan adalah sampel barang yang biasa dibeli di supermarket lokal. Setelah dilakukan perhitungan manual menggunakan Algoritma Apriori dan Pengujian dengan perhitungan otomatis pada aplikasi WEKA, diperoleh 4 aturan yang dapat digunakan sebagai rekomendasi kepada manajer toko dalam menyusun strategi pemasarn yang lebih baik[11].

Penelitian yang dilakukan oleh Agus Nursikuwagus dan Tono Hartono dalam jurnal yang berjudul “Implementasi Algoritma Apriori Untuk Analisis Penjualan Dengan Berbasis Web”, mengangkat masalah penumpukan data transaksi penjualan yang tidak dimanfaatkan dengan baik. Data yang digunakan diperoleh dari Microsoft Adventure Work yang berupa transaksi penjualan perlengkapan hobi sepeda gunung. Dari data tersebut dihitung confidence menggunakan Algoritma Apriori, menghasilkan 2 aturan kuat yang mewakili 3 barang populer dibeli oleh pelanggan[12].

Penelitian yang dilakukan oleh Fricles Ariwisanto Sianturi dalam jurnal yang berjudul “Penerapan Algoritma Apriori untuk Penentuan Tingkat Pesanan”, mengangkat masalah pemanfaatan data pesanan untuk mengembangkan strategi pemasaran. Data diperoleh dari observasi pada perusahaan CV. Mentari Persada, serta Wawancara terhadap staff administrasi dan staff persediaan barang. Peneliti melakukan pemrosesan data dengan merubah data pesanan dari bentuk data transaksi menjadi data tabular, lalu dengan menggunakan aplikasi Tanagra dicari aturan terkuat, yang menghasilkan 1 aturan terkuat terdiri dari 2 barang[13].

Tabel 1.1. Kajian Pustaka

No.	Judul	Comparing	Contrasting	Criticize	Synthesize	Summarize
1	Implementasi Algoritma Apriori untuk Mencari Asosiasi Barang yang Dijual di E-Commerce OrderMas[9]	Penelitian ini dilakukan terhadap data transaksi dari E-Commerce OrderMas	Penelitian ini dilakukan untuk mengetahui barang yang akan di suplai pada periode selanjutnya	Penambahan nilai minimal support dan nilai minimal confidence dapat memisahkan lebih detail barang apa saja yang paling sering dibeli.	Penelitian dimulai dari mengumpulkan dan merubah format data transaksi, lalu melakukan penghitungan secara manual, lalu diuji menggunakan rapid miner dan w-apriori	Penelitian menghasilkan 10 rule yang berguna bagi perusahaan untuk merancang strategi suplay barang periode selanjutnya
2	Metode <i>Data mining</i> Terhadap Data Penjualan	Penelitian ini dilakukan	Penelitian ini dilakukan untuk	Perubahan data transaksi dari	Penelitian ini dimulai dari	Penelitian ini menghasilkan 2

No.	Judul	Comparing	Contrasting	Criticize	Synthesize	Summarize
	Sparepart Mesin Fotocopy Menggunakan Algoritma Apriori[10]	terhadap data transaksi dari CV. Nusantara Copy	memanfaatkan data transaksi untuk mendapatkan informasi baru.	perbulan menjadi perhari, dan penambahan dataset	mengumpulkan data, lalu mencari frekuensi 1, 2, dan 3 <i>itemset</i> , lalu dicari final asosiasi, kemudian diujikan dengan aplikasi Tanagra	buah rule yang saling berhubungan.
3	Penerapan Algoritma Apriori untuk Market Basket Analysis[11]	Penelitian ini dilakukan terhadap data transaksi sampel	Penelitian ini dilakukan untuk mengeksplorasi analisis asosiasi dengan menggunakan	Perlu adanya implementasi ke studi kasus sehingga mendapatkan	Penelitian ini dimulai dari pengambilan data transaksi yang dibuat oleh peneliti lalu	Penelitian ini menghasilkan 5 aturan yang dapat memberikan rekomendasi

No.	Judul	Comparing	Contrasting	Criticize	Synthesize	Summarize
		yang ditentukan oleh peneliti	Algoritma Apriori dan FP-Growth	hasil yang lebih realistis	menghitung support untuk <i>n-itemset</i> , setelah itu dihitung nilai confidence dan lift ratio untuk menilai seberapa kuat aturan yang ada.	terhadap manajer untuk mengambil strategi pemasaran antara lain adalah menyimpan barang yang saling berhubungan secara berdekatan.
4	Implementasi Algoritma Apriori untuk Analisis	Penelitian ini dilakukan terhadap data	Penelitian ini dilakukan untuk menerapkan	Perlu adanya fitur untuk mengupload	Penelitian ini dimulai dari melakukan	Penelitian ini menghasilkan 2 aturan yang kuat

No.	Judul	Comparing	Contrasting	Criticize	Synthesize	Summarize
	Penjualan dengan Berbasis Web[12]	transaksi toko hobi sepeda yang diperoleh dari internet	analisis keranjang pasar dalam bentuk website	data sehingga dapat dilakukan analisis berulang kali.	pencarian <i>itemset</i> berfrekuensi tinggi dan menghitung confidence. Dilanjutkan dengan membuat website.	lalu dilanjutkan dengan visualisasi kedalam website sehingga hasil dari analisis keranjang belanja dapat dipahami oleh orang awam.
5	Penerapan Algoritma Apriori untuk Penentuan Tingkat Pesanan[13]	Penelitian ini dilakukan terhadap data transaksi CV. Mentari Persada	Penelitian ini dilakukan untuk mendapatkan tingkat pesanan	Perlu adanya optimalisasi metode mengingat semakin besar data transaksi	Penelitian ini dimulai dengan menghitung secara manual support dan confidence	Penelitian ini menghasilkan 1 aturan terkuat yang telah

No.	Judul	Comparing	Contrasting	Criticize	Synthesize	Summarize
			studi kasus terkait	nilai minimal support dan nilai minimal confidence cenderung susah untuk di tentukan	sehingga ditemukan prediksi aturan terkuat, lalu diujikan dengan aplikasi Tanagra	diujikan di aplikasi Tanagra

Berdasarkan penelitian rujukan tersebut, dapat disimpulkan bahwa faktor – faktor yang memberikan keharusan dalam melakukan market basket analysis atau analisis penjualan antara lain, kekosongan stok barang yang tidak diprediksi terhadap tren pembelian pelanggan, strategi pemasaran yang belum memiliki sistem rekomendasi, hingga kurangnya pemanfaatan data transaksi toko swalayan. Dengan adanya penelitian tentang market basket analysis atau analisis penjualan dapat memberikan informasi yang bernilai sebagai rekomendasi strategi pemasaran hingga rekomendasi pengadaan stok sesuai dengan tren pembelian pelanggan. Informasi yang diberikan berupa barang – barang yang saling terhubung satu sama lainnya dalam bentuk aturan – aturan asosiatif yang valid diukur dari nilai Lift Ratio.

2.2. Dasar Teori

2.2.1. Data mining

Data mining merupakan proses eksekusi satu atau lebih *machine learning* atau pembelajaran mesin secara otomatis, dengan tujuan melakukan analisis dan melakukan ekstraksi pengetahuan yang tersirat. *Data mining* juga memiliki sebutan lain yaitu pembelajaran dengan basis induksi adalah pembelajaran dengan cara melakukan observasi terhadap contoh spesifik dari pengetahuan yang akan dipelajari, untuk membentuk definisi baru. *Data mining* juga merupakan salah satu langkan dalam proses pencarian pengetahuan dalam database atau *Knowledge Discovery in Databases*. *Data mining* terbagi dalam beberapa tahap yang bersifat interaktif atau pengguna secara langsung berinteraksi. Terdapat 6 tahap dalam melakukan *data mining*[14], yaitu:

- a. *Data Cleaning* adalah proses pembersihan data merupakan proses pembersihan data dari data – data yang tidak diperlukan, atau data – data yang tidak relevan terhadap *data mining*.
- b. *Data Integration* adalah proses penggabungan data dari lebih dari 1 database, dalam *data mining* terkadang memerlukan lebih dari 1

database agar mendapatkan 1 database utama yang mencakup kebutuhan *data mining*.

- c. *Data Selection* atau pemilihan atribut dari suatu database dan menghapus atribut – atribut data yang tidak terpakai.
- d. *Data Transformation* merupakan perubahan data menjadi format yang sesuai
- e. *Process Mining* merupakan proses utama dalam menerapkan metode sehingga menghasilkan pengetahuan baru yang tersembunyi dan berharga
- f. *Pattern Evaluation* merupakan proses evaluasi terhadap proses *mining* yang telah dilakukan, yaitu dengan mengidentifikasi pola bermunculan
- g. *Knowledge Presentation* merupakan proses menyajikan informasi tentang metode yang telah digunakan dan mengemukakan kesimpulan atas proses *mining* yang terjadi

2.2.2. *Association Rules*

Association rule merupakan salah satu teknik yang dapat digunakan dalam mengolah data dengan cara melakukan asosiasi terhadap kandidat data yang akan diproses. Teknik yang dilakukan pada proses association rule juga memakai konsep relasi antara satu *item* atau kandidat dengan *item* lainnya. Dalam menentukan association rule, terdapat *interestingness* measure (ukuran ketertarikan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu. Perhitungan yang dimaksud adalah perhitungan nilai *support* dan nilai *confidence*. Nilai *support* adalah nilai penunjang antara *item* sedangkan nilai *confidence* adalah kepastian terhadap *item* yang terbentuk. Nilai *support* dan nilai *confidence* merupakan nilai yang akan dijadikan acuan atau pedoman dalam pengolahan terhadap data yang akan diproses. Berikut ini adalah rumus yang digunakan dalam perhitungan untuk mendapatkan nilai *support*, A adalah barang yang terlebih dahulu dibeli dan B adalah barang yang dibeli setelah A adalah sebagai berikut:

$$Support(A, B) = \frac{Jumlah\ item\ yang\ mengandung\ A\ dan\ B}{Jumlah\ keseluruhan\ item} \times 100\% \quad (2.1)$$

Sedangkan rumus yang digunakan dalam melakukan perhitungan untuk mendapatkan nilai *confidence* adalah sebagai berikut:

$$Confidence(A, B) = \frac{Jumlah\ item\ yang\ mengandung\ A\ dan\ B}{Jumlah\ keseluruhan\ item\ yang\ mengandung\ A} \times 100\% \quad (2.2)$$

Baik nilai yang terdapat pada nilai *support* ataupun nilai *confidence* memiliki keterkaitan antara satu dengan yang lainnya. Dalam melakukan perhitungan untuk mendapatkan nilai *confidence* haruslah melakukan perhitungan terlebih dahulu untuk nilai *support*, begitu juga sebaliknya. Nilai *support* yang didapatkan dapat dikatakan sesuai apabila didukung dengan adanya perhitungan dari nilai *confidence*[15].

Jumlah aturan asosiasi yang hanya menggunakan nilai *support* dan *confidence* dapat menjadi sangat besar dalam menentukan *rule*. *Lift ratio* adalah salah satu langkah tambahan dalam mengatasi masalah tersebut. *Lift ratio* dalam aturan asosiasi menunjukkan tingkat kebergantungan antara *antecedent* dan *consequent*. Apabila *lift ratio* bernilai > 1 , artinya tingkat ketergantungan antara *antecedent* dan *consequent* tinggi. Cara menghitung *lift ratio* adalah dengan membagi nilai *confidence* dengan *Expected Confidence* yang dinyatakan dengan rumus[16]:

$$Expected\ Confidence = \frac{Jumlah\ item\ yang\ mengandung\ B}{Jumlah\ keseluruhan\ item} \times 100\% \quad (2.3)$$

$$Lift\ Ratio = \frac{Confidence}{Expected\ Confidence} \quad (2.4)$$

2.2.3. Market Basket Analysis

Market basket analysis atau analisis keranjang belanja merupakan sebuah analisis terhadap perilaku pelanggan dalam melakukan kegiatan transaksi jual beli. Analisis keranjang belanja menghasilkan temuan asosiasi atau keterhubungan satu barang dengan barang lainnya, dimana barang - barang tersebut berada dalam satu keranjang pelanggan. Hasil dari analisis keranjang belanja dapat menjadi suatu informasi yang dapat digunakan untuk membuat rencana susunan barang dalam sebuah lemari pajang toko swalayan. Selain menjadi pertimbangan rencana susunan barang, analisis keranjang belanja juga dapat dimanfaatkan untuk membuat strategi baru

dalam menarik perhatian pelanggan untuk berbelanja di toko swalayan. Contoh pemanfaatan lainnya dalam analisis keranjang belanja adalah rencana pemberian potongan harga terhadap barang - barang yang dinilai populer dibeli oleh pelanggan toko swalayan. Analisis *market basket* dapat dikombinasikan dengan metode - metode *data mining* lainnya seperti prediksi atau klasifikasi, sehingga menghasilkan sebuah produk pembelajaran mesin yang relevan terhadap kebutuhan toko swalayan. Contoh dari analisis keranjang belanja adalah mayoritas produk Roti dibeli bersamaan dengan susu kental manis, Produk roti memiliki frekuensi pembelian paling tinggi, dan contoh lainnya adalah mencari tingkat frekuensi pembelian atau pembayaran jasa tertentu[17][16][17].

2.2.4. Algoritma Apriori

Algoritma apriori dikenal sebagai algoritma yang memiliki popularitas tinggi dalam bidang menemukan pola dengan tingkat frekuensi yang tinggi. Pola frekuensi tinggi adalah pola-pola *item* di dalam suatu database yang memiliki frekuensi atau *support* di atas ambang batas tertentu yang disebut dengan istilah minimum *support*. Algoritma apriori dibagi menjadi beberapa tahap yang disebut iterasi atau pass yaitu:

Pertama, Penyusunan kandidat *itemset* dengan cara melakukan kombinasi k-1 dari hasil yang telah dilakukan pada iterasi sebelumnya. Algoritma apriori melakukan ini dengan teknik *pruning* atau pemangkasan terhadap kandidat *itemset* yang memiliki subset k-1

Kedua, Penghitungan nilai *support* untuk tiap - tiap *item* dari kandidat *itemset*. Support dihitung dengan cara melakukan scan berulang terhadap database yang memiliki semua *item* dalam kandidat *itemset* tersebut.

Ketiga, Penetapan pola frekuensi tinggi yang memiliki kandidat *item* atau *itemset*, dimana nilai *support* dari kandidat tersebut lebih besar dari nilai minimal *support* yang telah ditetapkan.

Keempat, apabila ditemukan pola frekuensi tinggi yang baru maka ulangi langkah pertama dengan kombinasi k-2 dan seterusnya. Apabila tidak

ditemukan pola frekuensi tinggi yang baru, maka hentikan semua iterasi yang dilakukan oleh algoritma apriori[20]

2.2.5. *Flask*

Flask merupakan sebuah kerangka kerja mikro atau *library* python yang digunakan untuk mengembangkan sebuah aplikasi web[21]. Merujuk pada arti kerangka kerja mikro, *flask* memiliki struktur awal yang ringkas dan kecepatan eksekusi yang tinggi[22]. Meskipun diberikan struktur yang ringkas seperti tidak adanya fungsi validasi *form*, database dan fungsi lainnya, *flask* memiliki sifat fleksibel dan skalabel yang artinya dengan tambahan pihak ketiga dan librari python *flask* dapat mengimbangi kerangka kerja web lainnya[23].