

BAB 2

TINJAUAN PUSTAKA DAN DASAR TEORI

Pada bab ini akan dijelaskan mengenai tinjauan terhadap penelitian terdahulu dan penelitian sekarang, serta dasar-dasar teori yang diperlukan.

2.1. Tinjauan Pustaka

2.1.1. Penelitian Terdahulu

Untuk menghindari terjadinya kesamaan pada penelitian sebelumnya, perlu dilakukan peninjauan terhadap penelitian-penelitian yang sudah ada. Selain untuk menghindari terjadinya kesamaan, dengan meninjau penelitian-penelitian sebelumnya dapat dilakukan pengembangan terhadap penelitian-penelitian yang telah ada tersebut.

Tinjauan yang pertama adalah tentang *data mining*. Vazan dkk (2017) melakukan prediksi terhadap perilaku proses manufaktur dengan menggunakan data produksi, dimana informasi dan pengetahuan tentang perilaku masa depan sistem manufaktur didapatkan dari tugas prediktif data mining. Perilaku proses yang diprediksi diverifikasi menggunakan model simulasi. Para penulis menganalisa metode yang berbeda. Metode jaringan saraf dipilih untuk menyebarkan data baru oleh file PMML di fase akhir. Tujuan dari penelitian ini adalah untuk merancang dan memverifikasi alat penambangan data untuk mendukung kontrol sistem manufaktur dengan bertujuan meningkatkan proses pengambilan keputusan. Berdasarkan prediksi hasil produksi tujuan, strategi pengendalian yang sebenarnya dapat dimodifikasi secara tepat. Kemudian mereka dapat digunakan dalam sistem manufaktur nyata tanpa risiko.

Sedangkan Pitchayadejanant dan Nakpathom (2017) menggunakan *data mining* untuk mengelompokkan beberapa aktivitas turis di Orchard untuk merangsang lebih banyak lagi turis yang datang berkunjung. Metode yang digunakan adalah *Association Rule* dengan bantuan *tool* Rapid Miner. Hasil yang didapat adalah menuai dan mencicipi buah merupakan aktivitas yang paling direkomendasikan.

Tinjauan yang kedua adalah tentang *Market Basket Analysis*. Heydari dan Yousefli (2017) telah membuat model optimisasi untuk *Market Basket Analysis* dengan pendekatan solusi algoritma genetik. Pendekatan algoritma genetik diterapkan untuk memecahkan masalah pemrograman biner non-linear yang

diformulasikan dan contoh numerik yang digunakan untuk menggambarkan model yang disajikan.

Davies dan Worrall (1998) juga melakukan *Market Basket Analysis* dengan melihat perilaku konsumen Inggris. Venkatachari (2016) juga melakukan *Market Basket Analysis* hanya saja obyek yang dipelajari berbeda, yaitu kebiasaan membeli konsumen India di pasar Spanyol.

Selanjutnya ada Kamakura (2012) yang telah melakukan penelitian tentang metode *Sequential Market Basket Analysis*. *Sequential Market Basket Analysis* merupakan pengembangan dari metode *Market Basket Analysis* tradisional. Pada penelitian tersebut, dilakukan perbandingan antara *Market Basket Analysis* tradisional dengan *Sequential Market Basket Analysis* di sebuah swalayan untuk menentukan relasi antara 2 *item*.

Kemudian ada Loraine dkk (2012) yang telah sukses melakukan penelitian, yaitu *Market Basket Analysis for a Supermarket Based on Frequent Itemset Mining*. Pada penelitian ini dilakukan analisis tentang frekuensi *itemset* untuk mengatasi permasalahan pada supermarket. Salah satu metode yang digunakan untuk menyelesaikan permasalahan ini adalah *Market Basket Analysis*.

Sedangkan Samuel dkk (2014) melakukan penelitian untuk mengetahui hubungan *Product, Place, Promotion and Price* dalam *Market Basket Analysis*.

Tinjauan ketiga adalah tentang *Association Rule*. Bathla dan Kathuria (2015) membandingkan dua algoritma aturan asosiasi, yaitu: Algoritma Apriori dan Filter Associator. Mereka menganalisis pembuatan *frequent itemset* dan jumlah siklus yang dilakukan menggunakan algoritma Apriori dan Filter Associator dalam cakupan analisis asosiasi. Berdasarkan perbandingan di atas, mereka menyimpulkan bahwa Filter Associator adalah algoritma yang efisien daripada algoritma Apriori berdasarkan dua faktor di atas (besar *itemset* dan jumlah siklus) karena algoritma Apriori menghasilkan jumlah siklus yang lebih banyak dan menghasilkan *itemsets* yang sangat besar sehingga menurunkan kinerja algoritma.

Sedangkan Al-maolegi dan Arkok (2014) berhasil mengembangkan algoritma apriori yang dinilai membuang-buang waktu untuk pemindaian seluruh pencarian database pada *frequent itemsets*. Hasil eksperimen menunjukkan bahwa Apriori yang mereka buat dapat mengurangi waktu yang dikonsumsi sebesar 67,38%

dibandingkan dengan Apriori asli, sehingga membuat Algoritma Apriori lebih efisien dan menghabiskan waktu lebih sedikit.

2.1.2. Penelitian Sekarang

Penelitian yang kini dilakukan adalah penelitian *Market Basket Analysis* pada toko X yang terletak di kota Kudus, Jawa Tengah. Penelitian ini menggunakan metode *Association Rules* dengan algoritma Apriori. Data yang digunakan adalah data per-transaksi 3 bulan terakhir tahun 2016. Pengolahan data dilakukan dengan *software* Excel dan Tanagra untuk menganalisis keterkaitan produk sehingga dapat diketahui produk mana yang sering dibeli secara bersamaan oleh konsumen. Hasil dari penelitian ini berguna untuk memberikan usulan perbaikan dalam penataan *display* produk yang saling terkait sehingga memudahkan konsumen dalam berbelanja dan meningkatkan omzet penjualan toko X.

2.2. Dasar Teori

2.2.1. Display

Sopiah dan Syhabuddin (2008) menjelaskan bahwa *display* adalah usaha untuk menata barang yang berfokus agar pembeli tertarik untuk melihat dan memutuskan pembelian.

Sedangkan Alma (2011) menjelaskan bahwa *display* adalah upaya meningkatkan perhatian dan mendorong minat beli konsumen melalui daya tarik penglihatan langsung.

Jadi dapat disimpulkan bahwa *display* merupakan salah satu aspek penting pada ritel, dimana *display* atau penataan barang dagangan ini dapat mendorong keinginan beli konsumen pada saat datang ke ritel jika ditata dengan tepat. Penerapan *display* yang efektif dapat meningkatkan omzet penjualan dengan merangsang keputusan beli konsumen secara seketika.

2.2.2. Data Mining

Menurut Santoso (2007) *data mining* sering juga disebut sebagai *knowledge discovery in database* (KDD). *Knowledge discovery in database* adalah kegiatan mengumpulkan, memakai data, historis untuk menemukan keteraturan, hubungan atau pola dalam set data berukuran besar. *Knowledge discovery in database* terdiri dari tiga proses, yaitu:

a. *Preprocessing*

Ini merupakan tahap awal dari KDD yang mencakup proses pembersihan data, seleksi, penggabungan maupun transformasi.

b. *Data mining*

Data mining merupakan proses utama dari KDD, pada tahap ini algoritma diterapkan dengan tujuan utama mendapatkan informasi atau pengetahuan dari sumber data.

c. *Post Processing*

Tahap ini merupakan tahap evaluasi dari hasil *data mining* yang telah didapat.

Larose (2005) menyebutkan bahwa berdasarkan tugasnya *data mining* dapat dikelompokkan menjadi:

a. Deskripsi

Berkaitan dengan menggambarkan kecenderungan suatu data ke dalam sebuah pola.

b. Klasifikasi

Berkaitan dengan penggolongan data dengan variabel target kategori.

c. Prediksi

Hampir sama dengan klasifikasi namun untuk menghasilkan sebuah nilai pada masa yang akan datang.

d. Estimasi

Hampir sama dengan klasifikasi namun estimasi lebih mengacu ke arah numerik daripada kategori.

e. Pengklasteran

Pengelompokan baris data (*record*) kedalam beberapa kelas yang memiliki kemiripan / ciri khasnya sendiri-sendiri.

f. Asosiasi

Berkaitan dengan mencari atribut-atribut yang muncul secara bersamaan. Salah satu contoh penerapan asosiasi adalah *market basket analysis*.

Jadi dapat disimpulkan bahwa secara umum data mining adalah mencari dan menggali sekumpulan data dalam jumlah besar yang telah ada untuk mendapatkan informasi dan pengetahuan yang bermanfaat.

2.2.3. Market Basket Analysis

Market Basket Analysis adalah sebuah asosiasi dalam *data mining* untuk menemukan atribut-atribut yang muncul secara bersamaan/ dalam satu waktu. Proses ini menganalisis pola pembelian konsumen untuk mengetahui hubungan atau relasi antara satu *item* dengan *item* lainnya yang diletakkan dalam satu keranjang belanja. Biasanya hasil dari *Market Basket Analysis* akan digunakan untuk membuat sebuah kebijakan pada ritel atau toko (Nataliani dkk, 2012).

2.2.4. Association Rules

Dion (2007) menjelaskan bahwa analisis asosiasi atau *association rules* adalah suatu cara untuk menentukan hubungan antar-*item* dari bayaknya *itemset* (*large set of variables*) dalam data transaksi (*data set*). Sedangkan Fauzy dkk (2016) menjelaskan bahwa *association rule mining* atau analisis asosiasi adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi *item*. Tiga pengukuran atau parameter yang biasa digunakan dari sebuah analisis asosiasi adalah *support*, *confidence*, dan *lift ratio*.

Support dalam analisis asosiasi adalah suatu nilai yang menjelaskan banyaknya kemunculan atau dominasi suatu *item* ataupun kombinasi *item* dalam keseluruhan transaksi, contohnya, dari keseluruhan transaksi yang ada, seberapa tingkat dominasi yang menunjukkan bahwa *item* A dan *item* B dibeli dalam waktu bersamaan. Nilai *support* sebuah *item* dalam suatu transaksi dapat diperoleh dengan menggunakan rumus berikut (Deepak Y, 2018):

$$\text{Support (A)} = P(A) = \frac{\Sigma \text{Transaksi mengandung A}}{\Sigma \text{Transaksi}} \quad (2.1)$$

Apabila transaksi terdiri dari kombinasi 2 *item* pengukuran nilai *support* dapat menggunakan rumus:

$$\text{Support (A, B)} = P(A \cap B) = \frac{\Sigma \text{Transaksi mengandung A dan B}}{\Sigma \text{Transaksi}} \quad (2.2)$$

Confidence dalam analisis asosiasi merupakan nilai untuk mengukur kuatnya hubungan antar-*item* dalam *association rule*. Contohnya, seberapa sering *item* B dibeli jika orang membeli *item* A atau sebaliknya. Nilai *confidence* dari suatu *item* bisa diperoleh dengan rumus berikut (Deepak Yadav, 2018):

$$\text{Confidence (A, B)} = P(A | B) = \frac{\Sigma \text{Transaksi mengandung A dan B}}{\Sigma \text{Transaksi mengandung B}} \quad (2.3)$$

Lift Ratio adalah suatu parameter untuk mengukur kekuatan dari aturan asosiasi yang terbentuk. Biasanya *lift ratio* digunakan untuk menilai *valid* atau tidaknya suatu aturan asosiasi. Aturan asosiasi bersifat kuat dan *valid* jika memiliki nilai *lift ratio* lebih besar dari 1 (*lift ratio* > 1). *Lift ratio* diperoleh dengan rumus berikut:

$$Lift\ Ratio = \frac{Confidence(A, B)}{(\sum Transaksi\ B / \sum Transaksi)} \quad (2.4)$$

2.2.5. Klasifikasi ABC

Park dkk (2014) menjelaskan bahwa klasifikasi ABC adalah teknik yang digunakan untuk mengklasifikasikan/ mengkategorikan *item* persediaan ke dalam tiga kelompok sesuai dengan prioritas dan kepentingan, berdasarkan prinsip Pareto: kelas A berarti sangat penting, kelas B berarti cukup penting, dan kelas C berarti kurang penting. Kelas A biasanya mencakup barang-barang yang mencapai 80% dari total nilai namun hanya 20% dari total *item*. Kelas B mewakili kira-kira 15% dari total nilai dan mewakili sekitar 30% dari total *item*. Sedangkan kelas C mewakili hanya 5% dari total nilai namun merupakan 50% dari total *item*. Guvenir dan Erel (1998) menyarankan bahwa kontrol manajemen harus diperketat dan prakiraan permintaan individu harus dibuat untuk *item* kelas A, tingkat moderat untuk *item* kelas B, paling sedikit untuk kelas C.

2.2.6. Program Tanagra 1.4.50

Tanagra merupakan *software* yang dapat digunakan untuk membantu pengolahan data statistik dan beberapa metode *data mining*, seperti: Regression, Clustering, Scoring, Association, dan lain sebagainya. Tanagra yang digunakan oleh peneliti adalah Tanagra versi 1.4.50. Program Tanagra 1.4.50 adalah program *data mining* yang dikembangkan di negara Perancis dan dirilis pada tahun 2004. Program ini bisa diunduh dari situs aslinya dengan bebas biaya.



Gambar 2.1. Logo Program Tanagra