

BAB II
LANDASAN TEORI

2.1 Tinjauan Pustaka

Berikut ini adalah beberapa tinjauan pustaka yang dilakukan oleh peneliti pada penelitian untuk menjadi pendukung penelitian yang sedang di lakukan :

Tabel 2.1 Daftar Literatur

No	Nama Peneliti	Tahun	Judul
1	Fajar Adhinda Kusuma Wardani, Titin Kristiana	2020	Implementasi Data Mining Penjualan Produk Kosmetik Pada PT. Naturan Nusantara Menggunakan Algoritma Apriori
2	Arhfan Prasetyo, Numan Musyaffa, Ricki Sastra	2020	Implementasi Data Mining Untuk Analisis Data Penjualan Dengan Menggunakan Algoritma Apriori.
3	Surti Kanti & Richardus Eko Indrajit	2017	Implementasi Data Mining Penjualan Handphone Oppo Store SDC Tangerang dengan Algoritma Apriori
4	Dewi Anggraeni, Sukmawati Anggraeni Putri, Lilyani Asri Utami	2020	Implementasi Algoritma Apriori Dalam Menentukan Penjualan Mobil Yang Paling Banyak Diminati Pada Honda Permata Serpong
5	Nurul Adha, Lince Tomorola & Edward Robinson Slaglan	2017	Implementasi Data Mining Penjualan Sabun Dengan Menggunakan Metode Apriori

2.1.1 Literatur 1

Algoritma apriori dalam penelitian ini digunakan untuk mengetahui pola kombinasi penjualan produk kosmetik sehingga dapat membantu dalam strategi pemasaran pada PT. Natural Nusantara. Data transaksi penjualan yang terus bertambah dan menyebabkan penyimpanan data yang sangat besar. Kebanyakan data transaksi penjualan hanya dijadikan arsip tanpa dimanfaatkan dengan baik. Permasalahan tersebut bisa diselesaikan dengan data mining menggunakan algoritma apriori yang bertujuan untuk mengetahui bagaimana pola kombinasi item dan itemset frekuensi tinggi dan menggunakan *software RapidMiner 5.3*.

Setelah dilakukan dengan cara manual dan perhitungan menggunakan *software RapidMiner 5.3* dapat diketahui bahwa hasil yang diperoleh dari perhitungan manual sama dengan perhitungan menggunakan *software RapidMiner 5.3*. Pola asosiasi yang terbentuk dengan nilai *minimum support* 10% dan nilai *minimum confidence* 50% menghasilkan 7 aturan asosiasi. Dan *strong rules* yang didapatkan adalah jika membeli CHARCO maka akan membeli EBS dengan nilai *support* 15.5% dan nilai *minimum confidence* 61.4%. Dengan diketahui produk yang sering dibeli maka perusahaan dapat mengembangkan strategi pemasaran dan mengatur stok barang agar tidak terjadi penumpukan barang.

2.1.2 Literatur 2

Usaha rumahan Dapoerin's merupakan usaha pada bidang kuliner yang menjual makanan kudapan yang dibuat sendiri oleh pemiliknya. Algoritma apriori dalam penelitian ini digunakan untuk membantu menganalisis data transaksi untuk mengoptimalkan penjualan dan juga agar dapat mengurangi banyaknya sisa

kudapan yang tidak terbeli oleh pembeli. Berdasarkan permasalahan tersebut algoritma apriori dapat membantu mengembangkan strategi dalam memasarkan makanan kudapan, hal ini dapat dilihat pada hasil kajian yang menunjukkan bahwa penjualan makanan kudapan yang paling laris dapat diketahui dengan melihat makanan kudapan yang memenuhi *minimal support* dan *minimal confidence*, makanan kudapan yang laris terjual adalah Lontong, Putu Ayu, Risoles dan Piscok. Jika membeli kudapan putu ayu maka akan membeli lontong dengan nilai support 50% dan *confidence* 88%. Jika membeli kudapan risoles maka akan membeli lontong dengan nilai *support* 35,71 dan *confidence* 100. Jika membeli kudapan piscok maka akan membeli lontong dengan nilai *support* 35,71% dan *confidence* 83%.

2.1.3 Literatur 3

Masalah yang dihadapi pada Store SDC Tangerang adalah bagaimana menerapkan algoritma apriori untuk mengetahui penjualan tipe oppo yang paling banyak terjual. Karena barang yang sudah menjadi stock di *showroom* harus terjual semua tidak dapat dikembalikan ke pusat. Maka dari itu diperlukan algoritma apriori untuk pemanfaatan data penjualan dan tidak hanya digunakan sebagai arsip sehingga oppo store SDC dapat menyusun strategi pemasaran untuk memasarkan produk dengan tipe lain dengan meneliti apa kelebihan produk yang paling banyak terjual tersebut dengan produk lainnya dan dapat menambah persediaan stok di *store*.

Dapat di ketahui produk item yang paling banyak terjual, jika membeli NEO 5 maka akan membeli F1 Plus dengan jumlah *confidence* 75% dan jika membeli

R7S maka akan membeli F1F dengan jumlah *confidence* 86% sehingga *showroom* oppo *store* semarang dapat menyusun strategi pemasaran untuk memasarkan produk dengan tipe lain.

2.1.4 Literatur 4

Honda Permata Serpong merupakan salah satu dealer yang bergerak dalam bidang penjualan suku cadang dan bengkel resmi. Penumpukan barang pada Honda Permata Serpong yang tidak laku terjual dan sulit menentukan produk mana yang paling diminati konsumen. Untuk mengantisipasi hal ini, Honda Permata Serpong perlu mengolah data transaksi menjadi sebuah informasi yang berguna dan dapat dimanfaatkan sebagai strategi bisnis. Dengan data mining di harapkan dapat membantu para pebisnis dalam pengambilan keputusan terhadap apa yang berhubungan dengan persediaan barang.

Hasil data penjualan Mobil Honda Permata Serpong 2018, yaitu dengan nilai *support* dan nilai *confidence* tertinggi adalah Mobil Mobilio-Mobil Brio Satya (50%-100%). Mobil Jazz-Mobil HRV (33,33%-100%), dan Mobil HRV-Mobil Mobilio-Mobil Brio Satya (33,33%-100%). Hasil perhitungan manual dengan Algoritma Apriori Hasil yang didapatkan 100% sama, sehingga menghasilkan Hipotesa H0 yang berarti tidak terdapat perbedaan antara perhitungan manual dengan perhitungan algoritma apriori.

2.1.5 Literatur 5

Banyaknya persaingan dalam dunia bisnis khususnya dalam industri penjualan, menuntut para pengembang untuk menemukan suatu strategi yang dapat meningkatkan penjualan dan pemasaran produk yang dijual, salah satunya adalah

dengan pemanfaatan data penjualan produk sabun. Perlu sebuah kreativitas dan inovasi dari produsen agar penjualan produknya bisa mengalami peningkatan setiap bulannya, produk yang ditawarkan berbagai macam merek. Untuk mengetahui merek dengan penjualan terbanyak salah satu caranya adalah dengan memanfaatkan teknik data mining dalam hal ini menggunakan algoritma apriori (asosiasi data mining).

2.2 Pengertian Parfum

Parfum atau minyak wangi yaitu sebuah campuran minyak esensial dan aroma senyawa, fiksatif, dan pelarut yang dipakai dalam memberikan rasa wangi untuk tubuh manusia, objek, atau sebuah ruangan. Jumlah dan juga tipe pelarut yang di campur dengan minyak wangi menjelaskan apakah suatu parfum dianggap sebagai ekstrak parfum, *Eau de parfum*, *Eau de toilette*, atau *Eau de Cologne*. (Wikipedia)

Perbedaan antara *Eau de parfum*, *Eau de toilette* dan *Eau de Cologne* adalah sebagai berikut :

2.2.1 Eau de Perfume (EDP)

Jenis wewangian ini memiliki sedikit kandungan alkohol dan mengkonsentrasikan 816% dari bahan wewangian. Oleh karena itu, wanginya akan bertahan cukup lama selama 4 - 6 jam. *Genesis Eau de Perfume* lebih cocok untuk orang-orang yang berkegiatan desetiap harinya. Aromanya tak kalah menarik dengan parfumnya, dan harganya pun lebih terjangkau.

2.2.2 Eau De Toilette (EDT)

Eau de Toilette adalah wewangian tubuh yang lebih ringan dan menyegarkan dibandingkan dengan parfum yang terkesan memiliki wangi yang kuat. Karena kandungan alkohol dari parfum ini lebih tinggi dari pada wewangian. Konsentrasi *essence* dalam EDT hanya 5% sampai 15%. Inilah mengapa harga *eau de toilette* lebih murah dari parfum biasa. Karena kandungan alkohol yang tinggi dan konsentrasi biji parfum yang rendah, *eau de toilette* hanya dapat bertahan sekitar 3 hingga 5 jam.

Meskipun daya tahan wangi *eau de toilette* tidak bertahan lama namun jenis wewangian tubuh inilah yang paling populer karena harganya yang cenderung murah. *Eau de toilette* sendiri berasal dari bahasa Prancis *faire sa toilette* yang artinya bersiap-siap.

2.2.3 Cologne (EDC)

Ea De Cologne atau biasa dikenal dengan *cologne* adalah *deodorant* tubuh yang hanya mengandung 5% konsentrat minyak atsiri. Dibandingkan dengan dua parfum di atas, EDC memiliki kadar air dan alkohol yang lebih tinggi. Aromanya tentu lebih cepat menguap dan hanya bertahan sekitar 2-3 jam saja. *Cologne* berasal dari kota Koln, Jerman, tempat parfum ini pertama kali dibuat. Di tempat asalnya, ada beberapa campuran parfum yang termasuk dalam kategori khusus *cologne*, dengan aroma buah yang ringan dan segar seperti lemon, bergamot, dan jeruk. Namun belakangan ini, *cologne* atau *cologne* diklasifikasikan sebagai parfum encer. Cologne memang jarang yang mahal, namun tetap menjadi pilihan

masyarakat karena identik dengan jenis parfum untuk anak muda. EDC umumnya dikenal sebagai *Body Mist*, *Body Spray* atau *Body Splash* di pasaran.

2.3 Sejarah Parfum

Parfum telah dikenal selama ribuan tahun lalu, kata "parfum" berasal dari bahasa Latin *per fume*, yang berarti "melalui asap". Salah satu penggunaan parfum yang tertua adalah dupa dan aroma yang digunakan oleh para leluhur dalam upacara keagamaan. Orang Mesir adalah orang pertama yang memasukkan parfum ke dalam budaya mereka, diikuti oleh orang Cina kuno, India, Israel, Kartago, Arab, Yunani, dan Romawi. Penggunaan botol parfum paling awal adalah di daerah Mesir kuno sekitar 1000 SM. Orang Mesir menemukan bahwa kaca dan botol parfum adalah salah satu penggunaan umum paling awal dari sebuah kaca..

2.4 Data Mining

Data mining adalah sebuah cara yang menggabungkan teknik analisis data dan menemukan pola-pola penting dalam data. Secara sederhana, data mining atau pengembangan data dapat didefinisikan sebagai proses memilih, mengeksplorasi, dan memodelkan sejumlah besar data untuk menemukan sebuah tren yang cukup tidak dikenali. Dalam data mining, data disimpan dengan cara elektronik dan diproses secara otomatis oleh komputer dengan memakai teknik dan perhitungan tertentu. Data mining akan menjadi sebuah acuan untuk mengambil keputusan. Pengolahan data bisa dengan banyak cara di data mining, diantaranya adalah *rough set*, *fuzzy*, *association rule*, algoritma C4.5, dan lain-lain (Maharani et al., 2017: 6).

2.4.1 Tugas Data Mining

Data mining dibagi berdasarkan beberapa kelompok dan tugas yang dapat dilakukan, yaitu (Larose, Daniel T. 2005).

1. Deskripsi

Deskripsi yaitu menerangkan sebuah pola hubungan dan juga kecenderungan yang terdapat dalam data secara sederhana. Deskripsi ini juga merupakan pola dan kecenderungan yang sering memberikan kemungkinan penerangan untuk kecenderungan sebuah pola.

2. Klasifikasi

Merupakan cara dengan melihat pada kelakuan dan atribut dari kelompok yang sudah diartikan. Cara ini juga dapat memberikan klasifikasi pada sebuah data baru dengan memanipulasi data yang telah diklasifikasi dan dengan menggunakan hasilnya untuk memberikan sejumlah aturan. Klasifikasi menggunakan *supervised*.

3. Estimasi

Estimasi mirip dengan klasifikasi, yang berbeda adalah variabelnya target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun dengan menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi.

4. Prediksi

Prediksi ini memiliki kemiripan dengan klasifikasi dan estimasi, bedanya yaitu hasil dari prediksi akan ada dimasa yang akan datang. Ada beberapa teknik yang digunakan dalam klasifikasi dan estimasi yaitu dapat juga dimanfaatkan untuk keadaan yang tepat untuk diprediksi.

5. Klustering

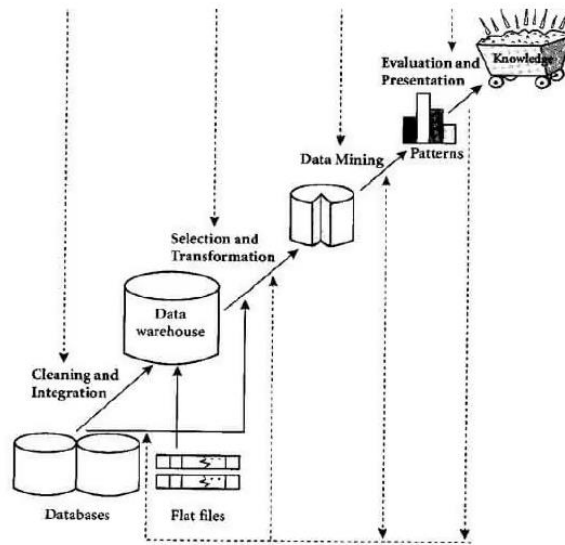
Klustering adalah pengelompokan data sebelumnya, pengamatan, atau memperhatikan dan membentuk kelas sebuah objek yang memiliki kesamaan satu dengan yang lainnya dan memiliki ketidak miripan dengan *record* dalam suatu kluster. Klustering menggunakan *unsupervised learning*.

6. Asosiasi

Tugas asosiasi atau sering disebut juga sebagai *market basket analysis* dalam data mining adalah menemukan hubungan atau sebuah korelasi diantara himpunan item dan mendapatkan atribut yang muncul dalam satu waktu. Asosiasi menggunakan *unsupervised learning*. Penting tidaknya suatu aturan assosiatif dapat diketahui dengan dua parameter, *support* dan *confidence*.

2.4.2 Tahapan Data Mining

Berikut beberapa tahapan dalam data mining menurut (Mabrur & Lubis 2012) diantaranya adalah :



Gambar 2.2 Proses Data Mining

1. Pembersihan Data (*Data Cleaning*)

Proses cleaning dengan cara membuang data duplikat, dan memeriksa data yang berbeda, dan menangani kesalahan pada sebuah data seperti tipografi atau salah cetak.

2. Integrasi Data (*Data Integration*)

Pada tahap data integration dilakukan pengecekan kombinasi data terhadap data yang berasal dari banyak sumber

3. Seleksi Data (*Data Selection*)

Pemilihan (seleksi) data dari sebuah kelompok data operasional perlu dilakukan sebelum pada tahap penggalian informasi. Data hasil penyeleksian yang akan dimanfaatkan dalam proses data mining lalu disimpan ke dalam suatu penyimpanan terpisah dari basis data operasional.

4. Transformasi Data (*Data Transformation*)

Adalah sebuah tahapan transformasi bagi data yang terpilih, sehingga didapatkan data yang sesuai untuk proses data mining. Proses ini adalah

sebuah tahapan yang sangat bergantung pada jenis atau pola informasi yang akan dicari pada basis data. Terkadang dalam data *warehousing* transformasi dan juga konsolidasi dilakukan sebelum proses seleksi data. Reduksi data yang dilakukan untuk memperoleh representasi yang lebih kecil. Dari data asli tanpa mengorbankan integritasnya.

5. Data Mining (Data Mining)

Proses data mining merupakan tindakan mencari sebuah pola hubungan atau informasi yang menarik dalam data yang sudah terpilih menggunakan Teknik dan tahapan tertentu. Pemilihan tahapan tersebut dan algoritma yang tepat akan sangat bergantung dari tujuan dan proses KDD tersebut secara data keseluruhan.

6. Evaluasi Pola (*Pattern Evaluation*)

Proses yang mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan berdasarkan beberapa ukuran tindakan meliputi hipotesa sebelumnya

7. Evaluasi Pola (*Knowledge Presentation*)

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah di mengerti oleh pihak yang berkepentingan.

2.5 Association Rules

Association Rule atau aturan Asosiasi adalah Teknik mining untuk menemukan aturan asosiatif atau pola kombinasi dari suatu item. Bila kita mengambil contoh aturan. Dalam menentukan *Association Rule*, ada dua ukuran yang digunakan yaitu:

1. *Minimal Support*

Minimal Support adalah ukuran yang harus dipenuhi sebagai batas frekuensi kejadian (*support count*) dari seluruh itemset (*support*) dalam keseluruhan transaksi.

Untuk mencari nilai *support* sebuah item (X) diperoleh dengan rumus :

$$\text{Support}(x) = \frac{\text{Jumlah Transaksi X}}{\text{Total Transaksi}}$$

Sedangkan untuk nilai *support* untuk itemset (X,Y) diperoleh dari rumus :

$$\text{Support}(x) = \frac{\text{Jumlah Transaksi X dan Y}}{\text{Total Transaksi}}$$

2. Minimal Confidence

Minimal Confidence adalah parameter seberapa sering item dalam Y muncul di transaksi yang mengandung X.

$$\text{Confidence}(X|Y) = \frac{\text{Jumlah Transaksi X dan Y}}{\text{Total Transaksi X}}$$

2.6 Algoritma Apriori

R. Agrawal dan R. Srikant adalah orang pertama yang mengenalkan algoritma apriori pada tahun 1994. Algoritma apriori yaitu salah satu algoritma yang dapat melakukan proses pencarian *frequent itemset* dengan menggunakan teknik *association rule*. Untuk mencari *association rule* dari suatu kumpulan data, tahap pertama yang harus dilakukan adalah mencari *frequent itemset* terlebih dahulu. *Frequent itemset* adalah sekumpulan item yang sering muncul secara bersamaan.

Penting tidaknya suatu asosiasi dapat diketahui dengan dua tolak ukur, yaitu *support* dan *confidence*. *Support* adalah nilai penunjang atau persentase kombinasi sebuah item dalam database, sedangkan *confidence* adalah nilai kepastian yaitu kuatnya hubungan antar item dalam aturan asosiasi. Proses utama yang dilakukan dalam algoritma apriori untuk mendapat *frequent itemset*.

2.7 Market Basket Analyst

Market basket analysis yaitu suatu metodologi untuk melakukan analisis *buying habit* konsumen dalam upaya menemukan asosiasi di beberapa barang yang berbeda, yang diletakan konsumen dalam *shopping basket* (keranjang belanja) yang dibeli pada suatu transaksi tertentu. Han, Cheng dan Xin (2007) mengatakan bahwa *Market Basket Analysis* dapat meningkatkan penjualan dengan melakukan pemasaran selektif serta mengatur ruang rak. Tujuan dari *market basket analysis* adalah untuk mengetahui produk - produk mana yang mungkin akan dibeli secara bersamaan.

2.8 Google Colab

Bahasa pemrograman *python* adalah bahasa pemrograman Dibuat oleh Guido van Rossum dan memiliki Kemampuan untuk menjalankan instruksi secara bersamaan Berbagai fungsi langsung menggunakan metode Object berorientasi pada pemrograman. Bahasa pemrograman ini juga menggunakan semantik dinamis untuk memberikan level Sintaksnya jelas dan mudah dibaca. Alat untuk melakukan perhitungan Menggunakan bahasa pemrograman Python adalah *Tools* Google Colab. Alat Google Colab adalah produk Google *Research* Akses mudah melalui

browser. Google Colab juga adalah file yang dapat dieksekusi untuk Tulis, simpan, dan bagikan program tertulis melalui Google Drive.