

BAB II

TINJAUAN PUSTAKA

2.1 Teori

Pada sub bab ini membahas teori, teori merupakan perkembangan dari bagian atau faktor yang saling terkait, definisi dan rekomendasi yang menyajikan perspektif yang efisien tentang kekhasan dengan memutuskan hubungan antara faktor-faktor, bertekad untuk mengklarifikasi kekhasan normal.

2.1.1 Teori Umum

Teori umum definisi dari **teori** pokok yang dibutuhkan dalam pembuatan penelitian ini dan sebagai landasan untuk pembuatan penelitian.

2.1.1.1 Data

Data adalah kenyataan pahit tentang individu, kesempatan dan sesuatu yang penting untuk dikoordinasikan. Informasi terdiri dari titik data statistik yang ditangani ke dalam data. Dapat dianggap bahwa informasi adalah sekelompok realitas atau angka dan dapat ditangani menjadi data yang bermanfaat. Dapat disimpulkan bahwa data adalah sekelompok realitas atau angka dan dapat ditangani menjadi data berharga.

2.1.1.2 Data Mining

Data Mining kini telah menjadi salah satu fokus pertimbangan skolastik seperti halnya para ahli. *Data mining* adalah metode yang terlibat dengan mendapatkan data berharga dari kumpulan data besar yang harus dihapus untuk menjadi data baru dan dapat membantu dalam navigasi. *Data mining* adalah cara paling umum untuk menyelidiki informasi dari berbagai dan menduganya menjadi data atau informasi atau contoh yang sangat penting untuk memperluas manfaat, membatasi biaya, atau bahkan keduanya.

Data Mining is one of the fastest growing fields due to the huge demands for added value of large-scale databases that are in line with the growth of information technology (Wahyuni, 2018). Koleksi data mining tidak hanya mengumpulkan informasi tetapi mencakup penelitian dan perkiraan data yang ingin Anda tunjukkan. Informasi yang dikumpulkan disimpan dalam kumpulan data dan kemudian ditangani sehingga dapat digunakan untuk pengambilan keputusan dengan mempertimbangkan informasi yang akan digunakan. (Sunge, 2018).

Data Mining adalah strategi atau teknik tertentu yang melacak bentuk atau data menarik berkaitan dengan informasi yang dipilih (Harman, 2020). Data mining sendiri memiliki beberapa metode, salah satunya adalah pengelompokan. Pengaturan adalah cara paling umum untuk melacak properti serupa dalam banyak artikel dalam kumpulan data dan mencirikannya ke dalam berbagai kelas sesuai model pesanan yang dicirikan (Harapan & Rismayanti, 2018).

2.1.1.3 Tipe Data Mining

1. Tipe *Data* Numerik adalah dat yang mewakili *data* atau nilai berjenis bilangan, makanya juga dapat disebut tipe data bilangan.
2. Tipe *Data* Kategorial adalah *data* dimana variabel – variabelnya dapat dikelompokkan menjadi beberapa kelompok atau kategori.
3. Tipe *Data* Rentang Waktu adalah tipe yang diperoleh dengan cara menunjukkan beberapa objek yang berbeda.

2.1.1.4 Algoritma C4.5

Algoritma merupakan sistem kerja komputer yang digunakan untuk menyusun langkah yang tepat untuk digunakan untuk mengatasi masalah (Cholil et al., 2020). Algoritma C4.5 untuk melakukan proses klasifikasi data dengan teknik pohon keputusan. Algoritma C4.5 merupakan ekstensi dari algoritma ID3 dan menggunakan prinsip decision tree yang mirip (Iswati, 2022). *Algoritma* C4.5 adalah kemajuan perhitungan ID3 yang dibuat oleh J. Rose Quinlan.

Sebagai aturan, *Algoritma* C4.5 untuk membangun pohon pilihan adalah sebagai berikut:

- a. pilih karakteristik sebagai akar
- b. membuat cabang untuk setiap nilai
- c. bagi kasus dalam cabang
- d. ulangi siklus untuk setiap cabang sampai semua kasus di cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, itu tergantung pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan persamaan

$$1. \text{ Gain}(S, A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy}(S_i) \quad \text{Rumus 2.1} \quad \text{Perhitungan}$$

gain

Keterangan :

S = himpunan kasus

A = atribut

n = jumlah partisi atribut A

|S_i| = jumlah kasus pada partisi ke-i

|S| = jumlah kasus dalam S

Sementara itu, perhitungan nilai *entropy* dapat dilihat pada persamaan. Entropy adalah ukuran dari teori informasi yang dapat mengetahui karakteristik (Arif & Risal, 2016).

$$2. \text{ Entropy}(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad \text{Rumus 2.2} \quad \text{Perhitungan } \textit{entropy}$$

2.1.1.5 *Decision Tree*

Metode *decision tree* adalah Struktur *flowchart* seperti struktur pohon, setiap titik pohon adalah karakteristik yang telah dicoba, setiap cabang adalah konsekuensi dari tes, dan titik akhir adalah pembagian kelas berikutnya. Pohon keputusan adalah penggambaran dasar dari prosedur pengaturan untuk berbagai kelas tak terbatas, di mana hub ke dalam dan hub akar dipisahkan oleh nama-nama sifat, tulang rusuk diberi nama kualitas yang mungkin dan ikatan daun dipisahkan dengan berbagai kelas (Eska, 2016).

Aturan dapat secara efektif dirasakan dengan bahasa biasa dan mereka juga dapat dikomunikasikan dalam struktur bahasa set data seperti SQL (Structured Query Language) untuk mencari catatan dalam klasifikasi tertentu Hal ini dapat dengan mudah dirasakan dengan bahasa normal dan mereka juga dapat dikomunikasikan dalam struktur bahasa set data seperti SQL (Structured Query Language) untuk mencari catatan di kelas tertentu. Pohon keputusan atau decision tree adalah pohon yang digunakan sebagai metodologi berpikir untuk menemukan solusi untuk masalah yang disertakan (Febriani et al., 2018).

Di *decision tree* ada 3 jenis *node*, yaitu :

1. *Root node*, merupakan *node* adalah *node* teratas, di *node* ini tidak ada *input* dan tidak dapat memiliki *output* atau memiliki lebih dari satu *output*.
2. *Internal node*, merupakan *node* percabangan, hanya ada satu *input* dan memiliki *output* dari sesuatu seperti dua.
3. *Leaf node* atau *terminal node*, merupakan *node* akhir, di hub ini hanya ada satu *input* dan tidak memiliki *output*.

2.1.2 Teori Khusus

Teori khusus adalah teori yang berkaitan dengan sejumlah fakta - fakta partikular tertentu. Teori yang menjelaskan fakta – fakta itu dalam hubungannya yang satu dengan yang lainnya.

2.1.2.1 Prediksi

Prediksi adalah proses menilai secara metodis apa yang mungkin akan terjadi di kemudian hari mengingat lebih dari data rentang waktu yang signifikan,

sehingga kesalahan (perbedaan antara sesuatu yang terjadi dan hasil perkiraan) dapat dikurangi. Ramalan penjualan metode yang terlibat dengan menilai beberapa persyaratan masa depan yang mengingat kebutuhan untuk proporsi jumlah, kualitas, pengaturan umum yang diperlukan untuk memenuhi kebutuhan akan tenaga kerja dan produk. (Rosita Dewi & Farouq Mauladi, 2020). Pengertian Prediksi setara dengan angka atau dugaan.

Sesuai dengan referensi kata besar bahasa Indonesia, prediksi adalah konsekuensi dari tindakan meramalkan atau mengantisipasi atau menilai nilai-nilai di kemudian hari dengan memanfaatkan informasi masa lalu. Perkiraan menunjukkan apa yang akan menimpa situasi tertentu dan berkontribusi pada persiapan dan siklus dinamis.

Peramalan (*forecasting*) adalah Teknik untuk membuat data yang dapat diverifikasi tentang keadaan sosial di masa depan berdasarkan data yang ada tentang masalah pengaturan. Dugaan memiliki tiga struktur prinsip: proyeksi, prediksi, dan perkiraan.

1. Suatu proyeksi adalah ukuran dalam terang ekstrapolasi pada berbagai pola waktu untuk apa yang akan datang. Proyeksi membuat pertanyaan tegas mengingat argument yang didapat dari metode eksplisit dan kasus yang sama.
2. Sebuah prediksi adalah ramalan mengingat anggapan hipotetis yang kuat. Anggapan ini dapat muncul sebagai hukum hipotetis (misalnya hukum mengurangi nilai uang tunai), saran hipotetis (misalnya saran bahwa pemisahan masyarakat umum hasil dari lubang antara asumsi dan

kapasitas), atau kesamaan (misalnya kesamaan antara pengembangan asosiasi legislatif dan pengembangan bentuk kehidupan alami).

3. Suatu perkiraan (*conjecture*) adalah dugaan mengingat evaluasi instruktif atau penilaian utama keadaan masa depan masyarakat.

Peramalan kebijakan juga penting untuk dikendalikan, dapat dikatakan, upaya untuk merencanakan dan menetapkan cara untuk memberikan pilihan aktivitas terbaik untuk melihat di antara hasil yang mungkin disajikan oleh apa yang akan datang. Apa yang akan datang juga di sana-sini dipengaruhi oleh masa lalu. Dengan menyinggung penyelidikan pendekatan masa depan harus memiliki pilihan untuk menilai nilai dari apa yang dapat atau harus mengarahkan kegiatan di masa depan.

2.1.2.2 Penjualan

Penjualan adalah pekerjaan terkoordinasi untuk mendorong rencana penting yang dipandu di tempat kerja untuk memenuhi kebutuhan dan kebutuhan pembeli, untuk mendapatkan penawaran yang menciptakan manfaat. Penawaran juga dapat diuraikan sebagai upaya yang dilakukan oleh orang-orang untuk meneruskan barang dagangan yang dibutuhkan oleh pelanggan. Dengan tawaran orang dapat bertemu dari dekat dan pribadi untuk membuat, meningkatkan, atau mengikuti koneksi perdagangan sehingga keuntungan umum.

Pengertian penjualan lainnya, Penjualan adalah tindakan yang menunjuk ke arah melacak pembeli, mempengaruhi dan mendorong pembeli untuk mengubah kebutuhan mereka untuk penciptaan yang ditawarkan dan masuk ke penyelesaian pada biaya yang bagus.

Penjualan dipahami sebagai suatu kegiatan yang secara sistematis merupakan adalah suatu rangkaian usaha yang secara terpadu dilakukan dalam menjalankan rencana organisasi untuk melakukan pemenuhan kebutuhan dan keinginan konsumen dalam rangka mendapatkan selisih antara harga jual dan harga beli yang disebut sebagai laba atau keuntungan (Izyuddin & Wibisono, 2020).

2.1.2.3 Produk

Produk adalah semua yang dapat diusulkan ke pasar untuk memenuhi keinginan atau kebutuhan, termasuk produk aktual, administrasi, pertemuan, acara, orang, tempat, properti, asosiasi, data dan pikiran. Produk merupakan titik konvergensi menampilkan latihan mengingat fakta bahwa produk adalah efek samping dari sebuah organisasi yang dapat diusulkan ke pasar untuk pemanfaatan dan merupakan perangkat dari sebuah organisasi untuk mencapai tujuan organisasinya. (Tusarwenda, 2018).

2.1.2.4 Keputusan Pembelian

Keputusan pembelian adalah Siklus di mana pembeli merasakan masalah, mencari data tentang item atau tanda tertentu dan menilai dengan tepat semua opsi ini dapat mengatasi masalah ini, mendorong opsi pembelian.. Dan ada juga menurut sumber lain yaitu menurut Kotler & Amstrong (2014), keputusan pembelian adalah tahap dalam proses pengambilan keputusan pembeli di mana konsumen benar-benar membeli.

2.1.2.5 Sistem Informasi

Menurut Husein dan Wibowo (2006), Sistem Informasi adalah sekelompok bagian yang saling terkait yang mampu mengumpulkan, berinteraksi, menyimpan, dan mengedarkan data untuk membantu *navigasi* dan pengawasan diasosiasi. Sistem informasi terdiri dari data tentang individu, tempat, kemudian sesuatu dalam asosiasi atau iklim yang menyiratkannya. Seperti yang ditunjukkan oleh Laudon dan Laudon (2017), sistem informasi sebenarnya adalah sekelompok bagian yang saling terkait yang mengumpulkan, menyimpan, memproses, dan menyebarkan data untuk membantu pengarahan dan pengawasan independen dalam asosiasi.

Sistem Informasi data juga membantu administrator dan perwakilan dalam memecah masalah, menggambarkan hal-hal yang kompleks, juga membuat item atau pengembangan baru. Kerangka data berisi data penting sebagai individu, tempat/wilayah, dan hal-hal penting lain yang berhubungan dengan pergaulan dan iklim di luar pergaulan itu. Dari pengertian tersebut, dalam keseharian sistem informasi adalah perkembangan dari komponen atau bagian saling berhubungan dan mempunyai tugas untuk mengumpulkan, menyimpan, memproses, dan data yang sesuai yang nantinya bisa dijadikan alasan untuk membuat pilihan.

2.1.2.6 RapidMiner

RapidMiner adalah pemrograman terbuka (open document). *RapidMiner* adalah jawaban untuk membedah informasi penambangan, penambangan teks, dan pemeriksaan *prescient*. *RapidMiner* menggunakan yang berbeda menjelaskan dan mengetahui metode dalam memberikan pengetahuan kepada klien sehingga Anda dapat menentukan pilihan terbaik. *RapidMiner* memiliki sekitar 500 administrator

penambahan informasi, termasuk *administrator* untuk input, hasil, pemrosesan dan *persepsi* informasi.

RapidMiner adalah program *independent* sendiri untuk penyelidikan informasi dan sebagai motor penambahan informasi yang dapat dikoordinasikan ke dalam item yang sebenarnya. *RapidMiner* dibuat dengan memanfaatkan Bahasa *Java* sehingga dapat menangani setiap kerangka kerja. *RapidMiner* memberikan *GUI (UI Realistic)* untuk perencanaan pemeriksaan pipa. *GUI* ini akan membuat *file XML (Extensible Markup Bahasa)* yang mencirikan arah keinginan *klien* untuk diterapkan pada informasi. Dokumen ini kemudian digunakan oleh *RapidMiner* untuk dijalankan ujian terprogram.

2.1.2.7 Hari Special Shopee

Sebagai salah satu pusat *komersial* yang memiliki banyak *klien* di Indonesia, Indonesia, Shopee serius menyelesaikan berbagai *promosi* untuk menarik tambahan lagi-lagi minat *klien* untuk melakukan belanja internet. *Promosi* Shopee adalah salah satu komponennya, kemajuan paling *persuasif* di *Merchant Place* dibuat khusus untuk membantu *klien* mendapatkan barang dagangan dengan biaya rendah. Kemajuan ini biasanya letakkan *banner* yang terlihat dan tersedia secara *efektif* di halaman utama atau halaman kelas Shopee.

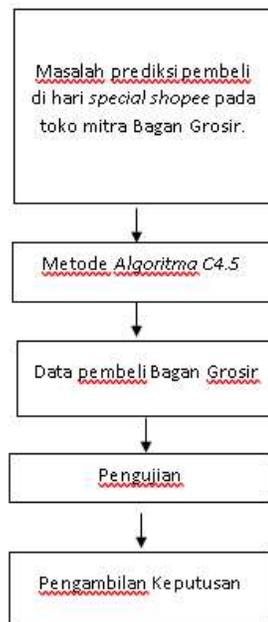
Dibawah ini merupakan salah satu contoh hari special shopee :

- 3.15 Hari belanja konsumen.
- 11.11 Big Sale.
- 25 – 27 Shopee gajian sale setiap bulan.

- 12.12 Birthday sale.
- 8.8 mid year sale.
- 2.2 Cod sale.
- 10.10 Brand festival.
- Dll.

2.2 Kerangka Pemikiran

Pada sub bab ini menjelaskan tahapan yang dilakukan untuk menyelesaikan masalah berdasarkan penelitian yang di buat, yaitu sebagai berikut tahapannya :



Gambar 2.1 Kerangka Pemikiran

Sumber: (Penulis,2021)

2.3 Hipotesis Penelitian

Hasil hipotesis perkiraan sementara dari penulisan ini adalah sebagai berikut :

1. *Metode data mining* dengan perhitungan algoritma C4.5 mampu menganalisa prediksi pembeli di hari *special shopee* pada toko Bagan Grosir.
2. Hasil dari penelitian ini dapat mengetahui prediksi minat pembeli pada hari *special promo shopee* .

2.4 Penelitian Terdahulu

Berikut informasi dan refrensi dari penelitian terdahulu yang memiliki keterkaitan tentang penelitian ini :

1. Berjudul “Penerapan Data mining untuk prediksi penjualan readymix menggunakan metode *algoritma C4.5* pada PT Remicon Widyaprima”. Yang dilakukan oleh (Hendra,2020). Penelitian ini berisikan tentang pengujian klasifikasi dan *algoritma C4.5* dalam wujud penerapan dari *data mining* yang di implementasikan dalam menentukan penjualan Readymix.
2. Berjudul “Penerapan *Data Mining* Dengan *Algoritma C4.5* Dalam Prediksi Penjualan Botol Pada CV.Seribukilo”. Yang dilakukan oleh (Tri Bagus tusarwenda,2018). Penelitian ini terkait dengan tingkat presisi dalam meramalkan penawaran botol pada CV.Seribukilo menggunakan *algoritma C4.5*.
3. Berjudul “ Prediksi Penjualan Produk Roti Menggunakan *Algoritma C4.5* pada PT. Prima Top Boga”. Yang dilakukan oleh (Diah

4. Rahmawati,2018). Penelitian ini berisikan tentang tujuan penulis Supaya perencanaan produksi diperusahaan sesuai permintaan konsumen dan pengirimannya tepat waktu.
5. Berjudul “Prediksi Kasus Positif covid-19 di Indonesia dengan menggunakan Metode Backpropagation”. Yang dilakukan oleh (Henzo Sabiq,2020). Penelitian ini memiliki tujuan untuk mengetahui apakah metode *Backpropagation* dapat digunakan untuk memprediksi kasus positif *COVID-19*.
6. Berjudul Pengaruh *Promo Shopee* “ Tanggal-Tanggal Istimewa”. Yang dilakukan oleh (Namira Shafinazh,2020). Penelitian ini memiliki tujuan untuk mengetahui pengaruh *promo Shopee* “tanggal-tanggal istimewa” terhadap minat beli mahasiswa Fakultas Ilmu Sosial dan Ilmu Politik Universitas Sumatera Utara.
7. Berjudul “Penerapan *Data Mining* Untuk Prediksi Penjualan *Wallpaper* Menggunakan *Algoritma C4.5*”. Yang dilakukan oleh (Juna Eska,2016). Penelitian ini penulis mengambil kesimpulan bahwa Membeli *wallpaper* bmenggunakan strategi mining informasi, terutama *Algoritma C4.5* akan sangat berharga sekali dalam interaksi dinamis dalam membeli *Wallpaper*.
8. Berjudul “Analisa *Algoritma C4.5* Untuk Memprediksi Penjualan *Motor* Pada PT. CAPELLA DINAMIK NUSANTARA Cabang Muka Kuning”. Yang dilakukan oleh (Nurul Azwanti,2018). Penulis mengambil kesimpulan bahwa *algoritma C4.5* dengan metode *decission*

tree bisa menghasilkan informasi *rule* prediksi untuk menggambarkan proses yang terkait dengan prediksi penjualan *motor*.