

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

2.1.1 Analisis Penjualan

Analisis penjualan merupakan rencana komprehensif, terintegrasi, dan konsisten di bidang pemasaran, yang berfungsi sebagai panduan untuk menjalankan berbagai aktivitas guna mencapai tujuan pemasaran perusahaan (Sidiq and Rohman, 2024). Dalam industri rokok, termasuk produk Clasmild, penjualan dipengaruhi oleh berbagai faktor seperti preferensi konsumen, strategi promosi, lokasi distribusi, serta event-event tertentu seperti *Musicverse*. Memahami perilaku konsumen secara mendalam dapat membantu perusahaan merancang strategi pemasaran yang lebih efektif untuk meningkatkan penjualan.

2.1.2 Event Marketing

Event marketing adalah bentuk promosi di mana perusahaan atau merek terhubung dengan sebuah acara atau kegiatan bertema, dengan tujuan menciptakan pengalaman bagi konsumen sekaligus mempromosikan produk atau layanan. Event *Musicverse*, sebagai studi kasus dalam penelitian ini, berperan penting dalam menarik perhatian konsumen, khususnya di kalangan anak muda. Dengan menganalisis data penjualan dan preferensi konsumen selama event ini, perusahaan dapat mengidentifikasi pola pembelian dan mengoptimalkan strategi promosi produk.

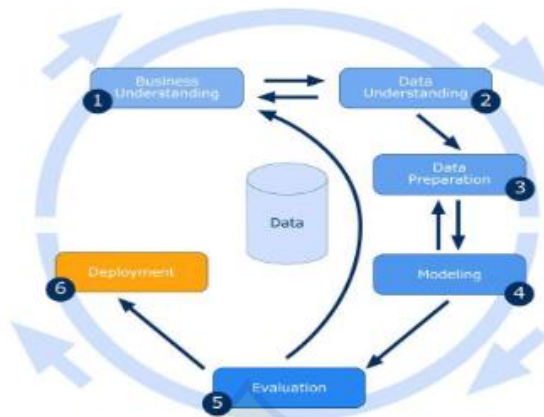
2.1.3 Data Mining

Data Mining adalah istilah yang merujuk pada proses penemuan pengetahuan dalam sebuah database. Proses ini melibatkan penggunaan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstrak serta mengidentifikasi informasi dan pengetahuan yang berharga dari kumpulan data yang besar. Berdasarkan definisi tersebut, data mining dapat diartikan sebagai upaya untuk menemukan pengetahuan tersembunyi dalam database.

Proses ini bertujuan untuk menemukan pola-pola tertentu dengan memanfaatkan metode statistik, matematika, kecerdasan buatan, serta machine learning guna mengekstrak dan mengenali informasi penting dari database yang dianalisis (Utomo and Mesran, 2020).

2.1.3.1 Metode Pengolahan Data (CRISP-DM)

CRISP-DM (Cross Industry Standard Process for Data Mining) adalah standar yang dirancang untuk proses data mining, di mana data yang tersedia akan melalui setiap tahap yang terstruktur, terdefinisi dengan jelas, dan efisien. Metode ini mengarahkan alur kerja dalam proyek data mining dari awal hingga akhir dengan pendekatan yang sistematis (Hasanah et al., 2021). Metodologi ini terdiri dari enam tahapan yaitu *Business Understanding*, *Data Understanding*, *Data Preparation*, *Modelling*, *Evaluation*, dan *Deployment*. Penjelasan lebih lanjut akan dijelaskan sebagai berikut.



Gambar 2. 1 Crisp-DM

Pada gambar 2.1 merupakan visualisasi dari proses crisp-dm yang meliputi *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation* dan *deployment*.

1. *Business Understanding* (Pemahaman Bisnis)

Yang dilakukan pada tahap ini yaitu memahami kebutuhan dan tujuan dari sudut pandang bisnis selanjutnya mengartikan pengetahuan kedalam sebuah pendefinisian masalah kedalam data mining selanjutnya menentukan rencana serta strategi untuk mencapai tujuan dari data mining itu sendiri.

2. *Data Understanding* (Pemahaman Data)

Pada tahap ini diawali dengan pengumpulan data, pendeskripsian data dan mengevaluasi kualitas data.

3. *Data Preparation* (Persiapan Data)

Dalam tahap ini membangun dataset akhir dari dataset mentah. Ada beberapa hal yang perlu dilakukan untuk pembersihan data (*Data Cleaning*), pemilihan data

(*Data Selection*), record, atribut- atribut, dan melakukan transformasi terhadap data (*Data Transformation*) untuk dijadikan masukan dalam tahap pemodelan.

4. *Modelling* (Pemodelan)

Di dalam tahap pemodelan ini secara langsung melibatkan *Machine Learning* untuk penentuan Teknik *data mining*, alat bantu data mining serta algoritma data mining.

5. *Evaluation* (Pengujian)

Tahapan ini melakukan pemahaman terhadap hasil dari data mining yang dihasilkan pada proses pemodelan dari tahap sebelumnya. Tujuan evaluasi ini untuk model yang ditentukan dapat sesuai dengan tujuan yang ingin di capai pada tahapan awal.

6. *Deployment* (Penyebaran)

Tahap ini melibatkan implementasi solusi data mining yang sudah dikembangkan kedalam lingkungan operasional atau produksi. Pada tahap ini hasil dari data mining yang dikembangkan pada tahap sebelumnya akan di integrasikan ke dalam sistem. Tujuannya adalah agar hasil dapat digunakan secara efisien untuk mendukung pengambilan keputusan.

2.1.3.2 Time Series

Time series adalah sekumpulan observasi yang dilakukan pada suatu variabel yang diukur secara berurutan dalam interval waktu tertentu, seperti tahunan, bulanan, triwulanan, atau periode lainnya. Metode *time series* bertujuan untuk mengidentifikasi pola dalam data historis dan menggunakan pola tersebut untuk memproyeksikan atau memprediksi kondisi di masa mendatang (Lubis, 2019)

2.1.3.2.1 Komponen Time Series

Data time series terdiri dari beberapa komponen utama, yaitu tren (trend), musiman (seasonality), siklus (cyclical), dan fluktuasi acak (irregularities). Tren menunjukkan arah umum perubahan data dalam jangka panjang, sedangkan musiman menggambarkan pola berulang dalam interval tertentu, seperti lonjakan penjualan pada akhir pekan atau selama event tertentu

2.1.3.2.2 Metode Time Series yang digunakan

Regresi linear merupakan teknik statistik yang digunakan untuk memahami hubungan antara satu atau lebih variabel independen (prediktor) dengan variabel dependen (respon). Persamaan regresi linear berfungsi untuk memperkirakan nilai variabel dependen berdasarkan perubahan yang terjadi pada variabel independen. Bentuk umum dari model ARIMA adalah sebagai berikut:

Bentuk umum dari rumus regresi linear sebagai berikut:

$$Y=a+bX$$

Keterangan:

Y = variabel dependen (hasil yang diprediksi)

X = variabel independen (faktor yang memengaruhi Y)

a = **intercept** (konstanta), yaitu nilai Y saat X=0

b = **koefisien regresi**, yang menunjukkan seberapa besar perubahan Y untuk setiap kenaikan satu unit pada X

2.1.3.3 Decision Tree

Decision Tree adalah metode klasifikasi yang menggunakan struktur pohon untuk pengambilan keputusan. Setiap node dalam pohon merepresentasikan atribut tertentu, cabang mewakili nilai dari atribut tersebut, dan daun (leaf node) menunjukkan kelas atau hasil keputusan. (Partogi and Pasaribu, 2022)

2.1.3.3.1 Struktur dan Prinsip Kerja Decision Tree

Struktur decision tree terdiri dari simpul (node) yang merepresentasikan atribut, cabang (branch) yang menunjukkan pilihan keputusan berdasarkan atribut tersebut, serta daun (leaf) yang menampilkan hasil klasifikasi atau keputusan akhir.

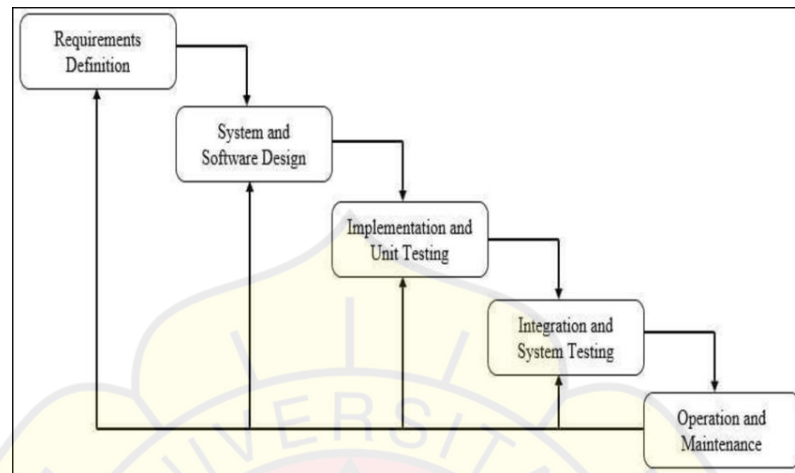
2.1.3.3.2 Penerapan Decision Tree pada Penelitian

Dalam penelitian ini, Decision Tree digunakan untuk mengklasifikasikan pengunjung event Musicverse berdasarkan preferensi mereka terhadap produk rokok, sehingga dapat memberikan insight bagi strategi pemasaran yang lebih tepat. Pembentukan pohon keputusan dilakukan dengan memilih atribut terbaik sebagai pemisah, menggunakan metrik seperti Information Gain atau Gini Index. Target sasaran.

2.1.4 Metode Pengembangan Sistem (Waterfall)

Metode yang digunakan untuk merancang dan membuat aplikasi sistem analisis dengan menggunakan metode *waterfall*. Metode waterfall adalah suatu metode yang melakukan pendekatan secara sistematis dan terurut pada pengembangan perangkat lunak. (Ula, 2021) Metode ini terdapat 5 (lima) tahap untuk mengembangkan aplikasi sistem disposisi yaitu *requirement definition*,

system and software design, implementation and unit testing, integration and system testing, dan operation and maintenance.



Gambar 2. 2 Model Perancangan Sistem (Waterfall)

Pada gambar 2.2 merupakan visualisasi dari metode waterfall yang tahapannya meliputi *requirement analyst, system design, implementation, integration and unit testing, dan operation and maintenance.*

Metode *Waterfall* memiliki tahapan-tahapan sebagai berikut :

1. *Requirements Definition*

Proses pengumpulan kebutuhan dilakukan secara matang untuk menyesuaikan kebutuhan perangkat lunak agar dapat dipahami perangkat lunak seperti apa yang dibutuhkan oleh pengguna.

2. *System and Software Designs*

Dalam tahap ini, digunakan untuk merancang fitur dari sistem baru yang dikembangkan.

3. *Implementation and Unit Testing*

Pembuatan kode program merupakan tahap terjemah *program code* untuk desain sistem yang sudah dibuat kedalam perintah perintah yang dimengerti komputer

4. *Integration and system Testing*

Pada pengujian program, seluruh unit yang dikembangkan dalam tahap implementasi digabungkan atau di integrasikan guna diuji kesesuaian program dengan kebutuhan dan tidak ada kesalahan.

5. *Operation and maintenance*

Tidak menutup kemungkinan sebuah program yang sudah selesai dibuat akan timbul kesalahan – kesalahan yang tidak ditemukan pada langkah sebelumnya. Oleh karena itu, setelah program berhasil dibuat perlu dilakukan pemeliharaan dengan baik untuk menanggulangi kesalahan – kesalahan yang tidak ditemukan tersebut.

2.1.5 Bahasa Pemrograman

Berikut beberapa bahasa pemrograman yang penulis pakai:

2.1.5.1 PHP (Perl Hypertext Preprocessor)

PHP merupakan bahasa pemrograman web berbasis server (server-side) yang dapat memproses kode PHP dari file web dengan ekstensi .php, sehingga menghasilkan tampilan website yang dinamis di sisi klien (browser). PHP adalah bahasa skrip yang sangat sesuai untuk pengembangan web dan dapat disisipkan ke dalam HTML.(Agustini and Kurniawan, 2019).

2.1.5.2 Codeigniter

Framework codeigniter adalah sebuah framework PHP yang dapat membantu mempercepat developer dalam pengembangan aplikasi web berbasis PHP(Anggraini et al., 2020)

2.1.6 Peralatan Pendukung

2.1.6.1 UML





Unified Modeling Language (UML) adalah bahasa pemodelan grafis yang digunakan untuk memvisualisasikan dan mendeskripsikan sistem perangkat lunak. UML merupakan salah satu standar bahasa yang umum digunakan di industri untuk mendefinisikan kebutuhan (requirement), melakukan analisis dan desain, serta menggambarkan arsitektur dalam pemrograman berorientasi objek (Julian Gerung, 2022).

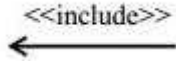
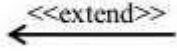
Berikut adalah beberapa jenis diagram UML beserta penjelasannya

1. UseCase Diagram

Use case diagram digunakan untuk memodelkan perilaku sistem informasi yang akan dibangun. Diagram ini menggambarkan interaksi tipikal antara pengguna (user) dan sistem melalui skenario yang mendeskripsikan bagaimana sistem digunakan. Komponen-komponen pembentuk *usecase* dijelaskan pada tabel 2.1 dibawah ini.

Tabel 2. 1 Komponen *Use Case Diagram*

Komponen UseCase	Keterangan
 Aktor	Merupakan sebuah komponen yang menggambarkan seseorang atau sesuatu (seperti perangkat atau sistem lainnya) yang berinteraksi dengan sistem.
 <i>Use Case</i>	<i>Use case</i> adalah gambaran fungsional dari suatu sistem, sehingga pengguna sistem paham dan mengerti mengenai kegunaan sistem yang akan dibangun.
 <i>Association</i>	Abstraksi hubungan antara aktor dan <i>Use Case</i> .
 Generalisasi	Menunjukkan spesialisasi aktor untuk berpartisipasi dengan <i>Use Case</i> .

	<p>Menunjukkan bahwa suatu <i>Use Case</i> seluruhnya merupakan fungsionalitas dari usecase lainnya.</p>
	<p>Menunjukkan bahwa suatu usecase merupakan tambahan fungsional dari <i>Use Case</i> lainnya jika suatu kondisi sudah terpenuhi.</p>

2. Class Diagram

Class diagram menggambarkan struktur sistem dari segi pendefinisian kelas-kelas yang akan dibuat untuk membangun sistem. Diagram ini berisi atribut dan operasi dalam kelas, yang bertujuan membantu pengembang untuk menghubungkan dokumentasi perancangan dengan perangkat lunak yang dikembangkan.



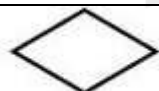
3. Sequence Diagram




Sequence diagram menggambarkan interaksi antar objek dalam use case, dengan mendeskripsikan urutan waktu hidup objek dan pesan yang dikirimkan serta diterima antar objek. Jumlah sequence diagram yang dibutuhkan bertambah seiring dengan banyaknya use case yang didefinisikan, dimana setiap use case dengan proses tersendiri memerlukan diagram ini.

4. Activity Diagram

Activity diagram menggambarkan alur kerja (workflow) atau aktivitas dari sistem dalam perangkat lunak. Diagram ini digunakan untuk menunjukkan bagaimana aktivitas-aktivitas dalam sistem berjalan secara berurutan dan saling terkait. Komponen-komponen pembentuk diagram *activity* dijelaskan pada table 2.2 dibawah ini.

Tabel 2. 2 Komponen Activity Diagram

Activity Diagram	Keterangan
 Status Awal	Start State, sebagai tanda awal proses dari <i>activity</i> diagram
 <i>Activity</i>	Menampung event atau aktivitas pada proses sistem.
 <i>Decision</i>	<i>Decision</i> , digunakan ketika terjadi pemilihan 2 kondisi event pada diagram

 <i>Join</i>	Penggabungan dimana lebih dari satu aktivitas lalu digabungkan menjadi satu
<i>Swimlane</i> 	Memisahkan organisasi bisnis yang bertanggung jawab terhadap aktivitas yang terjadi
 Status Akhir	<i>End State</i> , sebagai tanda akhir dari <i>activity diagram</i> .

2.2 Tinjauan Literatur/Kajian Literatur

2.2.1 Paper 1

Peramalan Penjualan Obat Generik Melalui *Time Series* Forecasting Model Pada Perusahaan Farmasi di Tangerang: Studi Kasus

Penulis : Hernadewita , Yan Kurnia Hadi, Muhammad Julian Syaputra , Donny Setiawan

Publikasi: Vol. 1 No. 2 : JULI 2020

Klasifikasi Jurnal : *Google Scholar*

Tujuan Penelitian: Tujuan dari penelitian ini adalah mengetahui metode peramalan *time series* terbaik untuk meramalkan penjualan obat generik, dan juga

memperoleh hasil peramalan penjualan obat generik untuk periode selanjutnya, yaitu terhitung bulan April 2020 sampai dengan Maret 2021

Metodologi : Penelitian ini merupakan penelitian deskriptif yang menggunakan metode peramalan *time series*. Teknik analisis data dilakukan dengan bantuan Microsoft Excel dan Minitab 17. Untuk mengetahui tingkat kesalahan pada metode peramalan, dihitung beberapa parameter, yaitu mean absolute deviation (MAD), mean squared error (MSE), dan mean absolute percentage error (MAPE). Berdasarkan hasil analisis, metode peramalan *time series* yang paling efektif untuk meramalkan penjualan obat generik ini adalah metode tren musiman.

Hasil Penelitian : Hasil peramalan penjualan obat generik dapat dijadikan acuan untuk memperkirakan kemungkinan yang akan terjadi di masa depan. Berdasarkan hasil prediksi, penjualan tertinggi tercatat pada bulan Mei 2020. Namun, perlu diperhatikan bahwa permintaan obat generik dipengaruhi oleh faktor musiman serta hari libur besar, seperti Lebaran yang jatuh pada awal Mei 2020. Oleh karena itu, peningkatan penjualan obat generik diperkirakan akan terjadi pada bulan April hingga Mei 2020. Selain itu, hari libur besar lainnya seperti Natal dan Tahun Baru juga perlu diperhitungkan karena berpotensi menyebabkan kenaikan permintaan. Dengan adanya prediksi ini, perusahaan dapat menggunakan informasi tersebut untuk merencanakan jumlah produksi yang sesuai dan memastikan ketersediaan stok yang mencukupi guna mengantisipasi lonjakan permintaan obat generik pada bulan-bulan tertentu.

2.2.2 Paper 2

Klasifikasi Data Penjualan Untuk Memprediksi Tingkat Penjualan Produk Menggunakan Metode Decision Tree

Penulis : Demira Intan Suranda, Adi Nugroho

Publikasi: Volume 7, Nomor 1, Juni 2024

Klasifikasi Jurnal : *Google Scholar*

Tujuan Penelitian: Penelitian ini bertujuan untuk menerapkan metode decision tree ID3 guna memperoleh informasi terkait klasifikasi data penjualan produk gamis dan kerudung, baik yang memiliki tingkat penjualan tinggi maupun rendah di Butik Aruna. Dengan penerapan metode ini, diharapkan Butik Aruna dapat mengelola persediaan produk gamis dan kerudung secara lebih optimal, mengurangi risiko terjadinya kelebihan atau kekurangan stok, serta meningkatkan pelayanan kepada pelanggan dengan menyediakan produk yang sesuai dengan kebutuhan mereka.

Metodologi : Penelitian dilakukan di sebuah butik yang berada di Ambarawa yaitu Butik Aruna. Pada penelitian ini menggunakan data penjualan produk gamis dan kerudung sebanyak 257 data. Metode yang digunakan yaitu metode decision tree algoritma ID3 dalam mengolah data.

Hasil Penelitian: Proses klasifikasi menggunakan metode decision tree ID3 pada data penjualan Butik Aruna dengan 257 data menghasilkan aturan berdasarkan jenis produk. Produk yang laris meliputi Gamis 2 dan Kerudung 2, sedangkan produk yang kurang laris adalah Gamis 1 dan Kerudung 3. Penelitian ini mencatat akurasi

sebesar 88,24%, menunjukkan bahwa metode ID3 dengan tools RapidMiner dapat digunakan untuk mendukung pengambilan keputusan strategis bisnis.

Butik Aruna dapat memanfaatkan hasil klasifikasi ini untuk memastikan ketersediaan stok produk laris, menghindari kekurangan, serta mengelola stok produk yang kurang laris agar tidak berlebihan. Strategi lain seperti menempatkan produk laris di posisi yang mudah diakses oleh konsumen juga dapat meningkatkan penjualan. Dengan memanfaatkan informasi ini, efektivitas dan tingkat penjualan perusahaan diharapkan meningkat.

2.2.3 Paper 3

ANALISIS PENGGUNAAN ALGORITMA KLASIFIKASI DALAM PREDIKSI KELULUSAN MENGGUNAKAN ORANGE DATA MINING

Penulis : Dinda Safitri, Shofa Shofiah Hilabi, Fitria Nurapriani

Publikasi: Volume 8 No. 1, Januari 2023

Klasifikasi Jurnal : *Google Scholar*

Tujuan Penelitian: Penelitian ini menganalisis tiga metode, yaitu Naive Bayes, K-NN, dan Neural Network, dengan tujuan untuk menentukan metode yang paling efektif dalam memprediksi kelulusan. Selain itu, penelitian ini membandingkan ketiga metode tersebut dan menghasilkan temuan bahwa metode K-NN adalah yang terbaik, dengan tingkat akurasi sebesar 89%.

Metodologi : Penelitian ini menggunakan beberapa algoritma dalam metode klasifikasi, yaitu Naive Bayes, K-NN, dan Neural Network, dengan memanfaatkan

Orange Data Mining sebagai alat bantu analisis. Data yang digunakan adalah tingkat stres mahasiswa dalam menghadapi skripsi, yang dianggap sebagai salah satu faktor yang memengaruhi kelulusan. Tujuan penelitian ini adalah untuk mengevaluasi tingkat akurasi masing-masing algoritma guna menentukan metode yang paling akurat atau tepat, sekaligus mengidentifikasi pola kelulusan mahasiswa.

Hasil Penelitian: Penelitian ini menggunakan tiga metode, yaitu Naive Bayes, K-NN, dan Neural Network, untuk memprediksi kelulusan mahasiswa. Hasil analisis menunjukkan bahwa algoritma K-NN memiliki performa terbaik dengan akurasi 89%, presisi 89%, dan recall 100%. Algoritma ini juga berhasil memprediksi 68 mahasiswa lulus tepat waktu pada kategori stres level 2.

Sebagai perbandingan, metode Naive Bayes mencapai akurasi 84% (presisi 90%, recall 99%), sementara Neural Network memiliki akurasi 88% (presisi 89%, recall 98%). Data menunjukkan bahwa mayoritas mahasiswa yang lulus tepat waktu berada pada level stres 2, mengindikasikan bahwa tingkat stres sedang tidak menjadi penghambat utama untuk kelulusan.

Penelitian ini merekomendasikan kajian lanjutan dengan membandingkan algoritma clustering dan klasifikasi untuk analisis yang lebih mendalam terkait akurasi kedua pendekatan tersebut.