

**SISTEM *POINT OF SALE* DAN PREDIKSI PENJUALAN
BRAND SMARTPHONE TERLARIS MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR***

PROPOSAL TUGAS AKHIR



Oleh

RAFAEL DWIKY NOVIAN HERNANDO

NIM E31222467

**PROGRAM STUDI MANAJEMEN INFORMATIKA
JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI JEMBER
2024**

**SISTEM *POINT OF SALE* DAN PREDIKSI PENJUALAN
BRAND SMARTPHONE TERLARIS MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR***

PROPOSAL TUGAS AKHIR



sebagai salah satu syarat untuk memperoleh gelar Ahli Madya (A.Md.)
di Program Studi Manajemen Informatika
Jurusan Teknologi Informasi

Oleh

RAFAEL DWIKY NOVIAN HERNANDO

NIM E31222467

**PROGRAM STUDI MANAJEMEN INFORMATIKA
JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI JEMBER
2024**

KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI
POLITEKNIK NEGERI JEMBER
JURUSAN TEKNOLOGI INFORMASI

HALAMAN PENGESAHAN PROPOSAL

- | | | |
|-------------------------------|---|--|
| 1. Judul | : | SISTEM <i>POINT OF SALE</i> DAN PREDIKSI
PENJUALAN BRAND SMARTPHONE
TERLARIS MENGGUNAKAN METODE
<i>K-NEAREST NEIGHBOR</i> |
| 2. Identitas Pelaksana | | |
| a. Nama Lengkap | : | Rafael Dwiky Novian Hernando |
| b. NIM | : | E31222467 |
| c. Jurusan/Program Studi | : | Teknologi Informasi/Manajemen Informatika |
| 3. Lokasi | : | Politeknik Negeri Jember |
| 4. Identitas Dosen Pembimbing | | |
| Dosen Pembimbing | | |
| a. Nama Lengkap | : | Pramuditha Shinta Dewi Puspitasari,
S.Kom., M.Kom. |
| NIP | : | NIP 19880404 202012 2 013 |
| Jurusan/Program Studi | : | Teknologi Informasi/Manajemen Informatika |
| 5. Lama Kegiatan | : | Enam (6) Bulan |
-

Jember, 19 Juli 2024

Menyetujui,
Dosen Pembimbing



Pramuditha Shinta Dewi P., S.Kom., M.Kom.
NIP. 198804042020122013

Pelaksana,



Rafael Dwiky N.H
E31222467

Mengetahui,
Koordinator Program Studi Manajemen Informatika



Taufiq Rizaldi, S.ST., MT
NIP. 198903292019031007

DAFTAR ISI

	Halaman
PROPOSAL TUGAS AKHIR.....	i
HALAMAN PENGESAHAN PROPOSAL	ii
DAFTAR ISI	iii
DAFTAR TABEL.....	v
DAFTAR GAMBAR	vi
BAB I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	2
1.5 Batasan Masalah	3
BAB II. TINJAUAN PUSTAKA	4
2.1 Tinjauan Pustaka	4
2.1.1 Penelitian Terdahulu	4
2.1.2 State of The Art.....	7
2.2 Landasan Teori.....	10
2.2.1 Smartphone	10
2.2.2 Penjualan.....	10
2.2.3 Data	10
2.2.4 Data Mining	11
2.2.5 KDD.....	12
2.2.6 Metode K-Nearest Neighbor.....	14
2.2.7 Rapid Miner	15
2.2.8 PHP	16
2.2.9 MySQL	16
2.2.10 Xampp.....	16
2.2.11 CRISP-DM.....	16
2.2.12 User Acceptance Testing (UAT)	19

2.2.13	Black Box Testing.....	20
BAB III. METODE PENELITIAN.....		22
3.1	Tempat dan Waktu Pelaksanaan.....	22
3.1.1	Tempat Pelaksanaan.....	22
3.1.2	Waktu Pelaksanaan	22
3.2	Alat dan Bahan.....	22
3.2.1	Alat yang digunakan	22
3.2.2	Bahan yang digunakan.....	23
3.3	Metode Pengumpulan Data.....	23
3.3.1	Studi Literatur	23
3.3.2	Observasi.....	24
3.3.3	Wawancara	24
3.3.4	Dokumentasi	24
3.4	Metode Kegiatan.....	24
3.4.1	Business Understanding.....	25
3.4.2	Data Understanding	26
3.4.3	Data Preparation.....	26
3.4.4	Modeling.....	26
3.4.5	Evaluation	35
3.4.6	Deployment.....	35
3.5	Jadwal Kegiatan.....	36
DAFTAR PUSTAKA		37

DAFTAR TABEL

Tabel 2.1 State of The Art	7
Tabel 3.1 Jadwal Kegiatan	36

DAFTAR GAMBAR

Gambar 2.1 Langkah-langkah KDD	13
Gambar 2.2 Eclidean	15
Gambar 2.3 Metode CRISP-DM	17
Gambar 3.1 Metode CRISP-DM	25
Gambar 3.2 Alur Penjualan Smartphone	25
Gambar 3.3 Metode Enclidean	27
Gambar 3.4 Flowchart KNN	28
Gambar 3.5 Flowchart Sistem	29
Gambar 3.6 Halaman Login	30
Gambar 3.7 Halaman Dashboard	30
Gambar 3.8 Halaman Data Barang	31
Gambar 3.9 Halaman Supplier	31
Gambar 3.10 Halaman Transaksi Pembelian	32
Gambar 3.11 Halaman Transaksi Penjualan	32
Gambar 3.12 Halaman Laporan	33
Gambar 3.13 Halaman User	33
Gambar 3.14 Halaman Data Training	34
Gambar 3.15 Halaman Data Testing	34
Gambar 3.16 Halaman Hasil Uji	34
Gambar 3.17 Usecase Diagram	35

BAB I. PENDAHULUAN

1.1 Latar Belakang

Industri *smartphone* mengalami perkembangan pesat dalam beberapa tahun terakhir. Dengan banyaknya *brand* dan model yang beredar di pasaran, persaingan untuk menjadi *brand smartphone* terlaris pun semakin ketat. Perkembangan teknologi yang pesat membawa perubahan dan kemudahan bagi khalayak dalam menjalankan aktivitas. Beragam kecanggihan teknologi dihadirkan, salah satunya *smartphone* yang dilengkapi berbagai fitur dengan sifat yang lebih fleksibel (Adisty, 2022). Dalam keadaan seperti ini, kemampuan memprediksi penjualan sangat penting bagi produsen untuk mengoptimalkan strategi pemasaran dan produksi mereka. Salah satu strategi yang dapat digunakan adalah dengan memprediksi penjualan *brand smartphone* paling laris. Prediksi ini dapat membantu produsen dalam mengambil keputusan strategis, seperti menentukan jenis *smartphone* yang akan diproduksi, target pasar, dan strategi marketing yang tepat sehingga diperlukan data *mining*.

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam *database*. *Data mining* merupakan proses yang menggunakan teknik statistik, Matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakit dari berbagai *database* besar. Dengan menganalisis data penjualan, karakteristik produk, dan preferensi konsumen, *data mining* dapat memberikan wawasan yang berguna untuk memprediksi penjualan *brand smartphone* di masa depan. Salah satu metode *data mining* yang populer adalah *K-Nearest Neighbor* (KNN). Disini penulis memilih metode KNN dibanding metode yang lain dikarenakan metode KNN memiliki keakuratan yang baik dibandingkan dengan metode *Naive Bayes*, *K-Nearest Neighbor* mencapai akurasi sebesar 88,46% dan *AUC* sebesar 99,4% termasuk dalam kategori *Excellent*, namun metode klasifikasi *Naive Bayes* memiliki akurasi sebesar 73,08% dan *AUC* sebesar 79,4%, yang masuk dalam kategori fair menurut kriteria *AUC* (Priatmojo, Saputra, Prasetyo, Puspitasari, & Nurlaela, 2023). Kemudian dari data tersebut dapat

diperkuat dengan kelebihan dari metode KNN yaitu tahan terhadap noise data pelatihan, tidak memerlukan pra-pemrosesan data yang rumit, mudah diimplementasikan dan dipahami, serta efisien jika data pelatihannya rumit sehingga dengan adanya metode ini dapat membantu mengetahui keakuratan pada label laris dan tidak laris (Rismala., 2023). KNN adalah algoritma klasifikasi yang bekerja dengan membandingkan data baru dengan data lama yang tersimpan dalam basis data. Algoritma ini mencari k tetangga terdekat dari data baru berdasarkan kemiripan fitur, dan kemudian mengklasifikasikan data baru berdasarkan kelas mayoritas dari k tetangga terdekat tersebut.

Penjualan merupakan suatu hal yang sangat penting bagi suatu perusahaan, yaitu menghasilkan pendapatan dan keuntungan sehingga perusahaan dapat terus berjalan. Galaxy Phone Jember merupakan toko *smartphone* di Jember yang menjual *smartphone* baru dan bekas. Masalah yang umum sering terjadi adalah bagaimana cara menjual barang di masa depan. Untuk memudahkan toko dalam melakukan stok barang, diperlukan prediksi berdasarkan data historis sebelumnya dengan menggunakan *data mining*. Berdasarkan uraian latar belakang diatas, penulis mengusulkan sebuah judul “Sistem *Point Of Sale* dan Prediksi Penjualan *Brand Smartphone* Terlaris Menggunakan Metode *K-Nearest Neighbor*”.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, rumusan masalahnya adalah bagaimana membangun sebuah sistem *point of sale* yang mampu membantu dalam pengambilan keputusan dengan prediksi *brand smartphone* terlaris di toko Galaxy Phone Jember.

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah untuk mengetahui penjualan produk *brand smartphone* terlaris dengan studi kasus Galaxy Phone Jember.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini yaitu:

1. Membantu pemilik Galaxy Phone Jember dalam mengambil keputusan dengan mengetahui produk *brand smartphone* yang paling banyak dibeli.

2. Membantu dan mempermudah pemilik toko dalam perencanaan penyediaan stok.

1.5 Batasan Masalah

Batasan masalah penelitian ini yaitu:

1. Penelitian ini hanya berfokus pada penerapan metode *K-Nearest Neighbor* (KNN) untuk memprediksi penjualan brand *smartphone* paling laris, tidak mencakup metode *data mining* lainnya.
2. Data yang digunakan dalam penelitian ini terbatas pada data penjualan *smartphone* yang dimiliki oleh Galaxy Phone Jember dari Januari 2024 hingga Desember 2024 (1 tahun).
3. *Tools* yang dipakai di dalam penerapan *data mining* ini adalah *Rapidminer*.
4. Sistem *point of sale* cerdas toko *smartphone* berbasis web.

BAB II. TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

2.1.1 Penelitian Terdahulu

Penelitian yang dilakukan orang lain merupakan penelitian terdahulu pada program maupun pengaplikasian sebelum di buat oleh peneliti. Sebagai pemerolehan wawasan yang berharga untuk membangun penelian agar penyempurnaan analisis yang dihasilkan. Berikut beberapa contoh karya tulisan sebelumnya.

2.1.1.1 Penerapan *Data Mining* Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT.Nippon Indosari Corpindo Tbk Menggunakan Metode *K-Nearest Neighbor* (Yolanda, 2021).

Jurnal ini membahas penerapan *data mining* dengan metode *K-Nearest Neighbor* (KNN) untuk memprediksi penjualan produk roti terlaris di PT. Nippon Indosari Corpindo Tbk, yang menghadapi kesulitan dalam menentukan produk terlaris dari berbagai jenis roti yang diproduksi. KNN adalah metode klasifikasi yang mengelompokkan data baru berdasarkan kemiripan dengan data lama melalui perhitungan jarak terdekat. Penelitian menggunakan data penjualan roti selama 3 bulan terakhir, dengan atribut kuantitas produk dan kuantitas terjual, yang ditransformasi ke dalam bentuk bobot dan dihitung jarak *euclidean*-nya untuk menemukan tetangga terdekat berdasarkan nilai K yang ditentukan. Hasil penelitian menunjukkan bahwa metode KNN efektif untuk memprediksi penjualan produk roti terlaris, membantu perusahaan menentukan jumlah produksi sesuai permintaan pasar.

2.1.1.2 Penerapan *Data Mining* Untuk Prediksi Penjualan Smartphone Paling Laris Menggunakan Metode *K-Nearest Neighbor* (Studi Kasus: Pusat Ponsel dan Laptop) (Garang, 2022).

Penelitian ini dilakukan di Pusat Ponsel & Laptop di Palangka Raya untuk memprediksi penjualan *smartphone* terlaris menggunakan metode *Data Mining* dan algoritma *K-Nearest Neighbor* (KNN). Data

penjualan smartphone selama enam bulan, dari Oktober 2021 hingga Maret 2022, dengan total 1519 unit, dianalisis menggunakan teknik observasi, dokumentasi, dan studi kepustakaan. Proses analisis mengikuti tahapan *Knowledge Discovery in Database* (KDD) dengan memanfaatkan aplikasi *RapidMiner*. Hasil prediksi menunjukkan bahwa penjualan smartphone Vivo paling tinggi dengan 210 unit, diikuti oleh Oppo (130 unit), Samsung (53 unit), Xiaomi (23 unit), dan Realme (44 unit). Evaluasi algoritma dengan Root Mean Square Error (RMSE) menunjukkan nilai terkecil pada Realme (3.325), menunjukkan hasil prediksi yang mendekati nilai sebenarnya, sementara Vivo menjadi smartphone terlaris dengan nilai *RMSE* 11.468. Penelitian ini membantu Pusat Ponsel & Laptop dalam pengelolaan stok smartphone berdasarkan prediksi penjualan.

2.1.1.3 Implementasi *Data Mining* Untuk Prediksi Penjualan Kusen Terlaris Menggunakan Metode *K-Nearest Neighbor* (Bahtiar, 2022).

Penelitian ini dilakukan di Toko Kusen Kembar Djaya, yang telah beroperasi sejak 2003, untuk memprediksi penjualan kusen terlaris menggunakan metode *Data Mining* dan algoritma *K-Nearest Neighbor* (KNN). Dengan meningkatnya pembangunan rumah, kebutuhan akan kusen juga meningkat, sehingga prediksi penjualan diperlukan untuk mengelola stok bahan baku secara efektif. Data penjualan dari Juni 2019 hingga Mei 2022 dianalisis menggunakan langkah-langkah dalam *Knowledge Discovery in Database* (KDD), termasuk pemilihan, pembersihan, transformasi data, dan interpretasi hasil. Metode KNN menghasilkan akurasi prediksi sebesar 88.89% untuk data penjualan dan 80.00% untuk data bahan terpakai, membuktikan efektivitasnya dalam membantu toko mengantisipasi permintaan dan mengatur stok bahan baku agar tidak kekurangan.

2.1.1.4 Penerapan *Data Mining* Untuk Prediksi Penjualan Produk Pangan Hewan Menggunakan Metode *K-Nearest Neighbor* (Iriane, 2023).

Penelitian ini menggunakan metode *K-Nearest Neighbor* (KNN) untuk memprediksi penjualan produk pakan hewan, bertujuan membantu pemilik usaha dalam pengambilan keputusan stok produk. Dengan menggunakan 208 dataset penjualan yang dibagi menjadi 80% data training dan 20% data *testing*, penelitian ini melakukan *preprocessing* data, menentukan nilai K, dan menghitung *euclidean distance* untuk klasifikasi produk. Hasilnya menunjukkan akurasi 80,4%, dengan produk terlaris "Whiskas 80gr Junior Tuna" (prediksi 6963 terjual) dan yang paling sedikit terjual "Whiskas Adult 1.2kg" (prediksi 8 terjual). Ini menunjukkan bahwa metode KNN efektif dalam memprediksi penjualan dan dapat digunakan untuk mengoptimalkan stok produk pakan hewan.

2.1.1.5 Penerapan *Data Mining* Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode *K-Nearest Neighbor* (Dewi, 2022).

Penelitian oleh Sri Puspita Dewi, Nurwati, dan Elly Rahayu membahas penerapan data mining menggunakan metode *K-Nearest Neighbor* (KNN) untuk memprediksi penjualan produk terlaris di UD Andar, sebuah usaha dagang yang menjual berbagai produk seperti jamu, kantong plastik, bahan makanan dan minuman, serta makanan beku. Sistem prediksi ini dibutuhkan untuk menggantikan metode manual yang tidak efisien dan kurang akurat, sehingga dapat membantu perencanaan penyediaan stok. Penelitian ini melibatkan pengumpulan data penjualan, analisis masalah sistem manual, penerapan algoritma KNN, perancangan dan pembangunan sistem informasi, uji coba, serta implementasi sistem baru. Hasilnya, sistem prediksi penjualan yang dihasilkan mampu meningkatkan akurasi dan efisiensi dalam menentukan produk terlaris, membantu optimalisasi keuntungan dan perencanaan stok di UD Andar.

2.1.2 State of The Art

Tabel 2.1 State of The Art

No	Penulis	Judul	Metode Penelitian	Hasil
1.	Ike Yolanda & Hasanul Fahmi (2021)	Penerapan <i>Data Mining</i> Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT.Nippon Indosari Corpindo Tbk Menggunakan Metode <i>K-Nearest Neighbor</i>	<i>K-Nearest Neighbor</i>	Hasil perhitungan jarak, maka akan didapatkan suatu hasil keputusan yaitu Naik = 3 dan Menurun = 2. Dapat dilihat pada Tabel 4.11 mayoritas klasifikasi yang memiliki jumlah paling banyak adalah “>1,5” dimana variable “>1,5” merupakan kategori “Naik”, dan variable “
2.	Braen Dwiatmajaya Garang (2022)	Penerapan <i>Data Mining</i> Untuk Prediksi Penjualan <i>Smartphone</i> Paling Laris Menggunakan Metode <i>K-Nearest Neighbor</i> (Studi Kasus: Pusat Ponsel dan Laptop).	<i>K-Nearest Neighbor</i>	Hasil prediksi untuk setiap <i>brand smartphone</i> yang paling laris adalah Vivo sebanyak 210 unit, Oppo sebanyak 130 unit, Samsung sebanyak 53 unit, Realme sebanyak 44 unit, dan Xiaomi sebanyak 23 unit <i>smartphone</i> .

3.	Rino Bahtiar (2022)	Implementasi <i>Data Mining</i> Untuk Prediksi Penjualan Kusen Terlaris Menggunakan Metode <i>K-Nearest Neighbor</i>	<i>K-Nearest Neighbor</i>	Dengan menerapkan data mining menggunakan metode <i>K-Nearest Neighbor</i> pada prediksi didapatkan hasil 88.89% menggunakan data penjualan dan 80.00% menggunakan data bahan terpakai. Dengan hasil akurasi yang cukup besar artinya metode <i>K-Nearest Neighbor</i> dapat digunakan untuk prediksi pada Toko Kusen Kembar Djaya.
4.	Rara Iriane dan Nurfaizah (2023)	Penerapan <i>Data Mining</i> Untuk Prediksi Penjualan Produk Pangan Hewan Menggunakan Metode <i>K-Nearest Neighbor</i>	<i>K-Nearest Neighbor</i>	Didapatkan prediksi produk pada bulan Oktober 2022 yaitu Whiskas 80gr junior tuna sebagai produk yang diprediksi paling banyak terjual sebanyak 6963 yang berkategori laris. Dan prediksi produk paling sedikit terjual yaitu Whiskas Adult Mackarel 1,2kg diprediksi terjual sebanyak 4 yang berkategori tidak laris.

5.	Sri Puspita Dewi, Nurwati, dan Elly Rahayu (2022)	Penerapan <i>Data Mining</i> Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode <i>K-Nearest Neighbor</i>	Dengan menerapkan metode <i>K-Nearest Neighbor</i> kedalam sebuah sistem aplikasi maka dapat membantu UD tersebut dalam menyelesaikan masalah yang dihadapi, sehingga teknik <i>data mining</i> dan metode algoritma <i>K-Nearest Neighbor</i> ini dapat diimplementasikan untuk memprediksi penjualan produk terlaris pada UD Andar.
6.	Rafael Dwiky Novian Hernando (2024)	Sistem <i>Point Of Sale</i> dan Prediksi Penjualan <i>Brand Smartphone</i> Terlaris Menggunakan Metode <i>K-Nearest Neighbor</i>	Pada penelitian ini akan dihasilkan sebuah sistem <i>point of sale</i> dan sistem penghitungan prediksi penjualan <i>brand smartphone</i> terlaris pada toko Galaxy Phone Jember. Kemudian kriteria yang digunakan meliputi data nama <i>smartphone</i> , <i>brand smartphone</i> , dan harga <i>smartphone</i> .

2.2 Landasan Teori

2.2.1 *Smartphone*

Ponsel pintar (*smartphone*) adalah telepon genggam yang mempunyai fitur dan fungsi tambahan yang mirip dengan komputer. Bagi sebagian orang, ponsel cerdas adalah ponsel yang menjalankan perangkat lunak sistem operasi apa pun yang menyediakan konektivitas standar dan dasar bagi pengembang aplikasi. Bagi yang lain, ponsel pintar hanyalah sebuah telepon yang menawarkan fitur-fitur canggih seperti email (*email*), Internet dan kemampuan membaca buku elektronik (*e-book*), atau memiliki keyboard (baik yang terpasang atau terhubung) dan konektor *VGA* (Kaswandari, 2017). Dengan kata lain, *smartphone* adalah sebuah komputer kecil dengan fungsi telepon yang berguna bagi manusia.

2.2.2 Penjualan

Penjualan adalah suatu kegiatan tambahan atau kegiatan pelengkap pembelian, sehingga memungkinkan terjadinya transaksi (Abdullah, 2017). Oleh karena itu, transaksi jual beli merupakan bagian penting dalam peralihan hak dan transaksi. Sedangkan Assrianti (2016) menyatakan bahwa penjualan adalah produk yang diproduksi dan dijual oleh suatu perusahaan. Penjualan adalah kegiatan yang dilakukan penjual untuk menjual barang atau jasa dengan harapan memperoleh keuntungan dari transaksi tersebut. Musa (2016) dalam jurnal Bauran Promosi, Sari Kuliah, Satu Nusa menyatakan bahwa penjualan adalah kegiatan menjual barang atau jasa secara kredit atau tunai.

2.2.3 Data

Data adalah suatu istilah majemuk yang berarti fakta atau bagian dari suatu fakta yang mengandung arti, lambang, gambar, angka, huruf, atau lambang nyata yang berhubungan dengan suatu gagasan, benda, kondisi atau situasi (Garang, 2022).

Sebaran data dibedakan menjadi dua jenis menurut sumber pengumpulannya, yaitu:

- 1) Data primer: yaitu data yang diperoleh atau dikumpulkan oleh penulis penelitian atau pihak berkepentingan yang memerlukannya. Data primer disebut juga informasi awal atau informasi baru.

- 2) Data sekunder: yaitu data yang diperoleh atau dikumpulkan dari sumber yang ada. Informasi tersebut biasanya diperoleh dari perpustakaan atau laporan / dokumen peneliti sebelumnya. Data sekunder disebut juga data tersedia.

Sebaran data menurut sifatnya terbagi menjadi dua bagian yaitu:

- 1) Data kualitatif: yaitu data yang tidak berbentuk angka. Data kualitatif berupa pernyataan verbal, simbol atau gambar. Contoh: warna kulit, jenis kelamin, status perkawinan, dll.
- 2) Data kuantitatif: yaitu data yang berbentuk angka atau data kualitatif yang berbentuk angka. Contoh: tinggi badan, umur, jumlah, titik belajar, suhu, dll.

2.2.4 Data Mining

Data mining merupakan serangkaian proses untuk mengekstrak nilai tambah dari suatu *database* berupa informasi yang tidak diketahui secara manual (Tarigan, Hardinata, Qurniawan, Safii, & Winanjaya, 2022). Penambangan data muncul dari tahun 1990-an sebagai cara yang valid dan mudah untuk mengekstrak pola dan data digunakan untuk menemukan hubungan antara data untuk mengelompokkan menjadi satu atau lebih cluster sehingga objek objek berada dalam satu dalam sebuah *cluster* dapat kesamaan yang tinggi antara satu sama lain. *Knowledge mining* merupakan bagian dari proses pencarian informasi penemuan pengetahuan dalam *database*. *Data mining* dibagi menjadi beberapa kelompok sesuai dengan tugas yang dilakukan, yaitu:

- 1) Deskripsi
Peneliti dan analis biasanya mencoba mencari cara untuk mendeskripsikan pola dan tren tersembunyi dalam data.
- 2) Evaluasi
Evaluasi mirip dengan klasifikasi, hanya saja variabel target evaluasi bersifat numerik, bukan kategoris. Model dibangun menggunakan catatan lengkap yang memberikan nilai variabel target sebagai prediktor. Kemudian pada revisi berikutnya, nilai variabel target dievaluasi berdasarkan nilai variabel prediksi.
- 3) Prediksi
Prediksi mirip dengan klasifikasi dan estimasi, hanya saja pada prediksi, nilai

hasilnya adalah di masa mendatang. Beberapa algoritma dan teknik yang digunakan dalam klasifikasi dan estimasi juga dapat digunakan (dalam kondisi yang sesuai) untuk prediksi.

4) **Klasifikasi**

Dalam klasifikasi variabel, tujuannya bersifat kategorikal. Misalnya mengklasifikasi kan gudang menjadi tiga kategori: gudang besar, gudang sedang, dan gudang kecil.

5) **Clustering**

Clustering adalah teknik untuk mengelompokkan catatan, observasi, atau kasus ke dalam kategori serupa. *Cluster* adalah sekumpulan *record* yang mirip satu sama lain dan berbeda dari record lain dalam *cluster*.

6) **Hubungan**

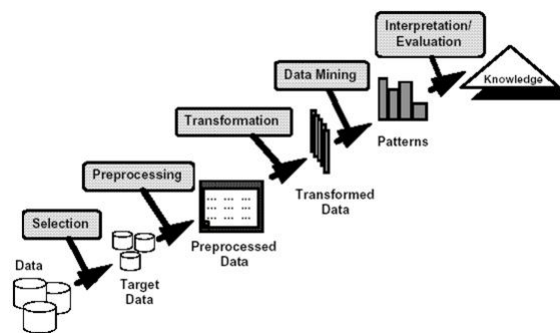
Identifikasi hubungan antara berbagai peristiwa yang terjadi secara bersamaan. Dalam dunia bisnis, hal ini lebih sering disebut dengan analisis keranjang belanja. *Data mining* memiliki kelebihan seperti:

- a) *Data mining* ini dapat menangani volume data yang besar dan kompleks.
- b) *Data mining* juga dapat memproses data dengan atribut yang berbeda-beda.
- c) *Data mining* dapat mencari dan mengolah data secara otomatis. Disebut semi-otomatis karena beberapa teknik data mining memerlukan parameter yang harus dimasukkan pengguna secara manual.
- d) *Data mining* dapat menggunakan pengalaman atau kesalahan masa lalu untuk meningkatkan kualitas dan hasil analisis untuk mendapatkan hasil terbaik. Ketika mencari kelemahan dalam mining, data dicari tidak secara individual, melainkan dari himpunan yang bersifat individualistis, dengan kata lain dikelompokkan berdasarkan kriteria tertentu.

2.2.5 KDD

Knowledge Discovery in Database adalah tentang mencari informasi tentang database. Ini juga didefinisikan sebagai proses non-sepele untuk mengidentifikasi data yang valid, baru, berpotensi berguna, dan pada akhirnya memiliki pola yang dapat dimengerti. Pencarian informasi sering kali terhambat oleh tantangan dalam mengintegrasikan dan menavigasi berbagai jenis informasi. Selain itu, seiring

bertambahnya jumlah dimensi data, pendekatan baru diperlukan untuk menemukan pola. Berdasarkan definisi tersebut, dapat disimpulkan bahwa Penemuan *Knowledge Discovery in Database* (KDD) adalah proses mengeksplorasi dan menganalisis kumpulan data yang besar serta mengekstraksi pengetahuan dan informasi yang berguna (Arta, 2019).



Gambar 2.1 Langkah-langkah KDD

Langkah-langkah penting dari proses KDD ditunjukkan pada gambar 2.1, yang terdiri dari langkah-langkah berikut:

- 1) *Data cleaning* adalah proses menghapus data duplikat, memeriksa data yang bertentangan, dan kesalahan yang ada pada data, seperti kesalahan ketik. Biasanya informasi yang diperoleh dari *database* perusahaan atau hasil pengujian memiliki isi yang tidak lengkap, misalnya informasi yang hilang, informasi yang salah, atau sekadar kesalahan ketik. Selain itu, terdapat juga atribut data yang tidak berhubungan dengan hipotesis *data mining*. *data cleaning* juga mempengaruhi kinerja data teknik *data mining* karena jumlah dan kompleksitas data yang diproses berkurang.
- 2) *Data Integration* adalah menambah data yang ada atau informasi lain yang relevan atau bisa disebut data adalah menggabungkan data dari *database* yang berbeda ke dalam satu *database* baru yang dibutuhkan KDD. Pada fase *cleaning* dan integrasi KDD, integrator data diharapkan dapat menghilangkan data asli dengan mengintegrasikan beberapa kumpulan data secara paralel. Penambahan data sebagai langkah dalam proses KDD.

- 3) *Data selection* adalah pemilihan data relevan yang dapat dianalisis dari data penggunaan. Informasi hasil pemilu disimpan dalam *database* terpisah.
- 4) *Data transformation* adalah proses mengubah data ke format tertentu sehingga data tersebut sesuai untuk proses *data mining*. Misalnya, beberapa metode standar seperti analisis asosiasi dan pengelompokan hanya dapat menerima data kategorikal sebagai masukan.
- 5) *Data Mining* adalah proses menemukan pola atau informasi menarik menggunakan teknik, metode, atau algoritma tertentu.
- 6) *Pattern evaluation* adalah proses untuk mengidentifikasi pola yang sangat menarik berdasarkan hasil *data mining*. Pada langkah ini, hasil teknik *data mining* dievaluasi dalam bentuk pola representatif dan model prediktif untuk mengevaluasi apakah hipotesis yang ada telah tercapai atau belum.
- 7) *Knowledge presentation* adalah proses menampilkan model informasi yang merupakan hasil proses *data mining*, visualisasi ini membantu hasil *data mining* dalam bentuk yang mudah dipahami.

2.2.6 Metode K-Nearest Neighbor

K-Nearest Neighbor (KNN) merupakan metode yang menggunakan algoritma terawasi dimana hasil *queryinstance* baru diklasifikasikan berdasarkan pengidentifikasi kelas mayoritas K-NN. Algoritma *K-Nearest Neighbor* (KNN) merupakan metode untuk mengklasifikasikan objek baru berdasarkan (K) tetangga terdekatnya. KNN adalah algoritma pembelajaran terawasi dimana hasil *query* baru diklasifikasikan berdasarkan mayoritas kelas KNN. Kelas yang paling umum adalah kelas hasil dari klasifikasi (Amalia, 2018).

Pada algoritma KNN terdapat 5 (lima) cara untuk mencari tetangga terdekat yaitu:

- 1) Jarak *Euclidean*
- 2) Jarak *Manhattan*
- 3) Jarak *Cosine*
- 4) Jarak *Correlation*
- 5) Jarak *Hamming*

Pada penelitian ini akan menggunakan jarak *Euclidean*, maka rumus perhitungan jarak dengan *Euclidean* seperti persamaan 1:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Gambar 2.2 Eclidean

x_i dan y_i merupakan subjek yang akan dibandingkan sebanyak n . Dimana nilai x_i merupakan nilai yang ada pada data *training*, sedangkan nilai y_i merupakan nilai yang ada pada data *testing*, nilai n merupakan banyaknya data *testing*.

Parameter k merupakan nilai yang menentukan berapa banyak tetangga yang akan dipilih untuk algoritma KNN. Pilihan k yang tepat berdampak signifikan pada kinerja diagnostik algoritma K-NN. K besar mengurangi dampak varians yang disebabkan oleh kesalahan acak, tetapi berisiko mengabaikan pola kecil tapi penting. Kunci untuk memilih nilai k yang sesuai dengan mencapai keseimbangan antara *overfitting* dan *underfitting* (Yanosma, Johar, & Anggriani, 2016).

2.2.7 *Rapid Miner*

Rapid Miner adalah platform perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Ini menyediakan berbagai alat untuk persiapan data, pemodelan, evaluasi dan implementasi. *Rapid Miner* dirancang agar mudah digunakan dan memungkinkan pengguna dengan mudah membangun dan menguji model yang berbeda, bahkan tanpa pengalaman pemrograman. *Rapid Miner* menyediakan antarmuka *drag-and-drop* yang memungkinkan pengguna membuat alur kerja untuk pemrosesan dan analisis data. Ini mendukung berbagai sumber data, termasuk file datar, *database*, dan platform data besar seperti *Hadoop* dan *Spark*. Perangkat lunak juga mencakup beberapa operator siap pakai yang merupakan blok penyusun alur kerja yang mencakup semua langkah proses penambangan data, seperti pembersihan data, pemilihan fitur, dan pemodelan (Nahjan, Heryana, & Voutama, 2023).

2.2.8 PHP

PHP Menurut Supono dan Putratama (Supono & Putratama, 2018) menyatakan bahwa “PHP (PHP: *Hypertext Preprocessor*) adalah bahasa pemrograman yang digunakan untuk menerjemahkan basis kode suatu program menjadi kode mesin yang dapat dipahami oleh komputer yang bersifat *server-side* yang ditambahkan ke HTML”. *Hypertext Preprocessor* (PHP) adalah bahasa pemrograman untuk membuat *website* dinamis yang dapat berinteraksi dengan pengunjung atau pengguna (PHP) adalah bahasa pemrograman yang memanipulasi *database*, konten *website* sehingga *website* yang dibuat adalah *web* dinamis dan PHP merupakan bahasa pemrograman yang digabungkan dengan HTML.

2.2.9 MySQL

MySQL adalah jenis server database yang terkenal. Popularitasnya karena MySQL menggunakan *SQL* sebagai bahasa dasar untuk mengakses *database*. *Mysql* adalah jenis RDBMS (*Relational Database Management System*). Di MySQL, database berisi satu atau beberapa tabel. Sebuah tabel terdiri dari beberapa baris, dan setiap baris berisi satu atau lebih kolom. Database MySQL dapat dikelola dengan beberapa cara yaitu dengan DOS (*tool command line*) (Maulana, 2016).

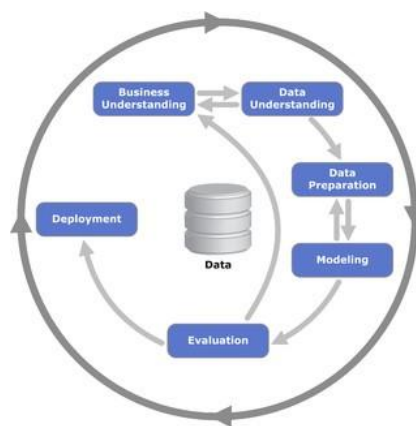
2.2.10 Xampp

Menurut Adi Baskoro, Xampp merupakan software server yang dapat berjalan di sistem operasi seperti Windows, Apple, dan Linux. XAMPP dapat menjalankan website atau aplikasi CMS, termasuk Joomla, Drupal, WordPress dan lain-lain. XAMPP sendiri mendukung dua sistem operasi yaitu Windows dan Linux. Di Linux proses instalasinya menggunakan command line sedangkan di Windows instalasi menggunakan GUI, jadi XAMPP mempunyai 3 komponen inti yang tertanam yaitu webserver Apache, PHP dan MySQL (Andani, 2021).

2.2.11 CRISP-DM

Metode *CRISP-DM* merupakan salah satu *framework* yang banyak digunakan untuk menganalisis data pada proses *data mining*. Kerangka kerja ini terdiri dari enam langkah, yaitu pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi dan implementasi. Metode *CRISP-DM* adalah pendekatan sistematis dan komprehensif untuk penambangan data. Metode ini dapat membantu

usaha mikro, kecil, dan menengah untuk mengatur dan menganalisis data secara efektif, sekaligus memberikan informasi berharga untuk mengambil keputusan bisnis yang akurat dan strategis, karena merupakan metode yang dapat diterapkan di banyak industri. Metode penelitian dengan menggunakan metode *CRISP-DM* merupakan gabungan antara metode kualitatif dan kuantitatif yang digunakan untuk mendeskripsikan data yang diteliti kemudian memberikan rekomendasi (Asyraf, 2024). Sebagai metode penelitian, siklus proses *CRISP-DM* dapat dijelaskan pada setiap tahapannya sebagai berikut:



Gambar 2.3 Metode CRISP-DM

1) Pemahaman bisnis

Tahap ini berfokus pada pemahaman menyeluruh terhadap tujuan bisnis yang ingin dicapai melalui data mining. proses. Pemahaman bisnis yang mendalam diperlukan untuk mengelola seluruh proses data mining agar hasilnya bermakna dan berguna bagi organisasi. Salah satu cara untuk memperoleh pemahaman kewirausahaan yang memadai sesuai tujuan penelitian adalah dengan membuat peta konsep. Pemetaan konsep dibuat berdasarkan informasi yang diperoleh melalui wawancara dan proses bisnis, serta konfirmasi yang diterima dari entitas yang menjadi subjek penelitian.

2) Pemahaman Data

Kemudian pada tahap pemahaman data, analis melakukan studi pendahuluan terhadap proses data mining. Dengan memiliki pemahaman mendalam terhadap data sejak dini, analis dapat memastikan bahwa data yang digunakan dalam proses hilir terkendali dan konsisten dengan tujuan bisnis,

bahwa hasil *data mining* akan menjadi lebih akurat dan dapat diandalkan untuk dipahami. data dan melakukan penelitian pendahuluan, analisis, terutama analisis TI latar belakang, biasanya membuat struktur *database* yang memfasilitasi proses di sini. Struktur database dapat digambarkan dengan diagram hubungan entitas. Diagram koneksi secara umum didasarkan pada data peta konsep dan wawancara terhadap unit-unit yang menjadi objek penelitian untuk memperjelas keakuratan informasi dan data.

3) Persiapan Data

Selanjutnya pada langkah persiapan data, data yang dipahami pada langkah sebelumnya sudah siap untuk digunakan dalam proses data mining. Data tersebut dimasukkan ke dalam format digital, setelah itu data yang terkumpul dibersihkan dan diolah sehingga siap digunakan dalam proses pemodelan menggunakan algoritma pohon keputusan. Dengan langkah persiapan data yang menyeluruh dan komprehensif, analisis dapat memastikan bahwa data yang digunakan dalam proses *data mining* sudah siap digunakan dan dapat menghasilkan hasil analisis bisnis yang akurat. Hasil dari langkah persiapan data ini dapat berupa laporan yang disajikan dalam tabel digital (*spreadsheet*). Dalam penelitian ini analisis menggunakan Microsoft Excel.

4) Pemodelan

Setelah tahap pemodelan, analisis menerapkan teknik penambangan data yang sesuai untuk mengembangkan model deskriptif dan prediktif berdasarkan data yang telah disiapkan sebelumnya. Fase pemodelan ini adalah inti dari proses penambangan data di mana informasi dan wawasan berharga ditemukan dalam data, yang kemudian membantu membuat keputusan yang lebih cerdas dan strategis untuk bisnis. Ada beberapa perangkat lunak untuk penambangan data, analisis dan visualisasi, seperti SPSS dan Dataminer, serta yang terbaru seperti Tableau, Power Query, Power BI dan KNIME. Penggunaan perangkat lunak ini dapat dikombinasikan sesuai dengan kebutuhan penelitian dan kemudahan penggunaan. Dalam penelitian ini, alat data mining dan visualisasi adalah KNIME. Perangkat lunak yang digunakan memiliki alat yang

membantu pembelajar algoritma yang dipilih, yaitu pohon keputusan dari node tersebut.

5) Evaluasi

Tahap evaluasi mengevaluasi secara menyeluruh hasil model dan *data mining* untuk mengukur kualitas dan relevansinya dengan tujuan bisnis yang telah ditetapkan sebelumnya. Fase evaluasi memastikan bahwa hasil *data mining* berguna, dapat diandalkan dan relevan dengan kebutuhan bisnis, yang menjamin keberhasilan implementasi dan pengambilan keputusan berbasis data yang lebih baik. Hasil pemodelan dan data mining serta analisis data bisnis dapat disajikan dalam bentuk laporan beserta wawasan yang diperoleh. Pada tahap evaluasi, kami melihat lebih luas model mana yang paling sesuai dengan perusahaan dan apa yang harus dilakukan selanjutnya baik untuk perusahaan maupun proyek penelitiannya.

6) Penyebaran

Pada fase ini, hasil model dan analisis *data mining* diterapkan dalam lingkungan bisnis nyata. Analisis berkolaborasi dengan tim bisnis untuk memastikan implementasi yang efektif dan integrasi model atau visi ke dalam proses operasi perusahaan, dimana pemilik perusahaan dapat mengambil keputusan berdasarkan hasil analisis jenis produk yang dijual. barang-barang. dapat melanjutkan dengan teknik *data mining* yang telah dievaluasi pada tahap sebelumnya. Namun, sebagai sebuah siklus hidup, hasil penerapan model ini atau pemahamannya dapat digunakan untuk perencanaan ulang pengembangan bisnis jangka panjang dan tujuan penambahan data di masa depan.

2.2.12 User Acceptance Testing (UAT)

User Acceptance Testing (UAT) adalah pengujian langsung interaksi antara *end-user* dan sistem, yang tujuannya adalah untuk memastikan bahwa fungsi berfungsi sesuai dengan kebutuhan pengguna (Chamida, 2021). Pengujian UAT merupakan tahap akhir dari proses pengujian sistem dimana sistem telah menyelesaikan tahap pengembangan. UAT adalah salah satu rangkaian pengujian perangkat lunak terakhir dan dilakukan sebelum dikembangkan dan dirilis.

Menurut Black (2002), *acceptance testing* biasanya mencoba untuk menunjukkan bahwa suatu sistem memenuhi persyaratan tertentu. Dalam perangkat lunak dan perangkat keras komersial, pengujian penerimaan juga biasanya disebut "*alpha tests*" (dilakukan oleh pengguna internal perusahaan) dan "*beta tests*" (dilakukan oleh pengguna yang menggunakan atau akan menggunakan sistem). Tes *alpha* dan *beta* juga biasanya menunjukkan bahwa produk tersebut siap untuk dijual atau dipasarkan. Pengujian penerimaan melibatkan data, lingkungan, dan skenario yang sama atau hampir sama dengan waktu nyata, yang biasanya berfokus pada skenario penggunaan produk tertentu.

Berikut adalah rumus menghitung skor pengujian *User Acceptance Testing* (UAT), yaitu:

$$\text{Skor} = \frac{\text{Jumlah jawaban}}{(\text{Jumlah Pertanyaan} \times \text{Jumlah Responden})} \times 100\%$$

Keterangan:

Jumlah jawaban: Jumlah jawaban responden ya / tidak

Jumlah pertanyaan: Jumlah pertanyaan yang diajukan kepada responden

Jumlah responden: Jumlah responden yang mengisi kuisioner

2.2.13 Black Box Testing

Black Box Testing menguji perangkat lunak sesuai spesifikasi secara fungsional tanpa menguji desain dan kode program. Tujuan pengujian adalah untuk mengetahui apakah fungsi perangkat lunak, input dan output memenuhi spesifikasi yang dibutuhkan. Pengujian *black box* dilakukan dengan membuat kasus pengujian yang menguji seluruh fungsi perangkat lunak untuk melihat apakah memenuhi spesifikasi yang diperlukan.

Menurut Perry (2006), pengujian fungsional bisa juga disebut pengujian *black box*, karena tidak ada informasi tentang logika internal sistem, yang digunakan dalam penyusunan kasus uji. Umumnya pengujian fungsional menggunakan teknik validasi untuk melakukan pengujian. Tim pengujian memvalidasi tombol fungsi yang ada dan memverifikasi hasilnya.

Keuntungan pengujian fungsional:

- a. Simulasi penggunaan sistem yang sebenarnya.
- b. Jangan membuat asumsi tentang struktur sistem.

Kekurangan Pengujian Fungsional:

- a. Berpotensi menghapus logika perangkat lunak yang salah.
- b. Mengaktifkan pengujian redundan.

Menurut Lewis (2009), dalam pengujian *black-box* atau pengujian fungsional, kondisi pengujian dikembangkan sesuai dengan fungsionalitas program atau sistem yang diuji, sehingga penguji memerlukan informasi tentang input dan output data diamati tetapi tidak mengetahui cara kerja program atau sistem atau tidak perlu mengetahui cara menjalankan struktur internal program. Penguji fokus pada pengujian fungsionalitas program terhadap spesifikasi. Menurut definisi di atas, pengujian *black-box* menguji fungsionalitas suatu sistem, terutama masalah perilaku dan bisnis, dimana penguji memerlukan informasi tentang data input dan output yang diamati, tetapi tidak perlu mengetahui struktur internal sistem.

BAB III. METODE PENELITIAN

3.1 Tempat dan Waktu Pelaksanaan

3.1.1 Tempat Pelaksanaan

Lokasi dari penelitian adalah Toko Handphone Galaxy Phone Jember Jl. Kalimantan No.77, Tegal Boto Lor, Sumbersari, Kec. Sumbersari, Kabupaten Jember, Jawa Timur 68121.

3.1.2 Waktu Pelaksanaan

Waktu pelaksanaan penelitian akan dilakukan pada Juli 2024 hingga Desember 2024.

3.2 Alat dan Bahan

3.2.1 Alat yang digunakan

Berikut adalah alat yang digunakan untuk melakukan prediksi penjualan *brand smartphone* paling laris menggunakan *data mining*. Alat yang digunakan pada penelitian ini terdiri dari perangkat keras (*hardware*) dan perangkat lunak (*software*) sebagai berikut:

1. Perangkat keras (*hardware*) yang digunakan:
 - a. Laptop *Lenovo*
 - b. Smartphone *Huawei*

Spesifikasi laptop yang digunakan untuk melakukan analisis sebagai berikut:

- 1) *Processor AMD Ryzen 5 5600H*
- 2) *RAM 16 GB dual channel*
- 3) *DDR 4 memory*
- 4) *Ideapad Gaming 3*
- 5) *2 TB SSD Nvme Gen 4*

Spesifikasi *smartphone* yang digunakan untuk melakukan analisis sebagai berikut:

- 1) *Huawei Nova 5T*
- 2) *RAM 4 GB*
- 3) *Baterai 3650 mAh*
2. Perangkat lunak (*software*) yang digunakan:

- a. Sistem operasi *Windows 11 Home*
- b. *Orange / Rapid Miner* sebagai tools
- c. *MySQL* sebagai database
- d. Bahasa pemrograman *PHP*
- e. *Xampp* sebagai server localhost
- f. *Visual Studio Code* sebagai *IDE* aplikasi
- g. *Microsoft Word 2021*
- h. *Microsoft Excel 2021*

3.2.2 Bahan yang digunakan

Adapun bahan yang dibutuhkan dalam pembuatan Tugas Akhir “Sistem *Point Of Sale* dan Prediksi Penjualan *Brand Smartphone* Terlaris Menggunakan Metode *K-Nearest Neighbor*”, yaitu berupa:

1. Data penjualan *smartphone* pada toko Galaxy Phone Jember dari bulan Juli 2024 hingga bulan Desember 2024.
2. Materi dan referensi berbentuk jurnal maupun naskah publikasi yang terkait dengan pembuatan sistem.

3.3 Metode Pengumpulan Data

Teknik pengumpulan data digunakan untuk mengumpulkan data sesuai tata cara penelitian sehingga diperoleh data yang dibutuhkan. Mustahil peneliti dapat menghasilkan temuan, jika tidak memperoleh data. Selanjutnya metode pengumpulan data yang dapat digunakan sesuai dengan judul tugas akhir penulis “Sistem *Point Of Sale* dan Prediksi Penjualan *Brand Smartphone* Terlaris Menggunakan Metode *K-Nearest Neighbor*” adalah sebagai berikut:

3.3.1 Studi Literatur

Studi literatur merupakan metode pengumpulan data dengan cara mengkaji sumber-sumber tertulis seperti buku, jurnal ilmiah, artikel, laporan penelitian terdahulu, dan dokumen lainnya yang relevan dengan topik penelitian. Penelitian ini menggunakan beberapa jurnal dan skripsi dari beberapa penulis yang berkaitan dengan penggunaan KNN pada penjualan terlaris sebagai referensi penulis untuk menentukan metode yang tepat pada data penjualan terlaris.

3.3.2 Observasi

Observasi adalah pengumpulan data secara langsung di objek yang diteliti untuk mengambil data yang dibutuhkan oleh peneliti. Observasi ini dilakukan secara langsung mengetahui proses penjualan *smartphone* pada toko Galaxy Phone Jember.

3.3.3 Wawancara

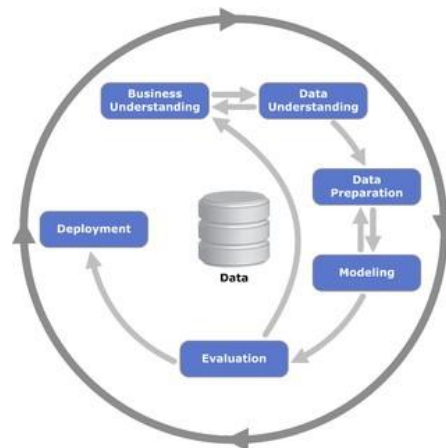
Wawancara digunakan sebagai teknik untuk memperoleh data langsung dari sumbernya baik mengenai pandangan atau pendapat maupun fenomena yang dilihat. Wawancara ini dilakukan dengan cara tanya jawab pada pemilik toko Galaxy Phone Jember. Hasil yang didapat pada wawancara tersebut menemukan sebuah permasalahan yaitu, belumnya penggunaan sistem *point of sale* pada penjualan barang sehingga penjualan masih menggunakan sistem manual, dan pemilik toko kesulitan untuk mengetahui *brand smartphone* paling laris dan tidak laku.

3.3.4 Dokumentasi

Dokumentasi ditujukan untuk memperoleh data langsung dari tempat penelitian, meliputi proses jual penjualan *smartphone* pada toko Galaxy Phone Jember, foto produk yang dijual, harga produk yang dijual, data yang relevan dengan penelitian.

3.4 Metode Kegiatan

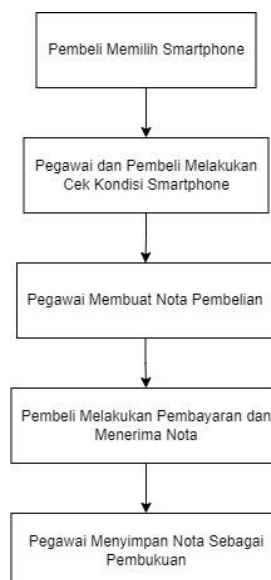
Agar tahapan analisis bisa lebih terstruktur peneliti menerapkan *Cross-Industry Standard Process for Data Mining (CRIDP-DM)* yang merupakan salah satu framework yang banyak digunakan untuk menganalisis data pada proses data mining. Metode penelitian dengan menggunakan metode CRISP Data Mining merupakan gabungan antara metode kualitatif dan kuantitatif yang digunakan untuk mendeskripsikan data yang diteliti kemudian memberikan rekomendasi (Asyraf, 2024). Kerangka kerja ini terdiri dari enam langkah yang akan diimplementasikan seperti pada gambar 3.1.



Gambar 3.1 Metode CRISP-DM

3.4.1 *Business Understanding*

Tahapan awal ini adalah tahapan dimana peneliti melakukan pemahaman terhadap objek masalah pada penelitian pada toko Galaxy Phone Jember. Tahapan tersebut dilakukan untuk mengetahui permasalahan apa yang dibutuhkan pada objek tersebut. Hal tersebut dapat dilakukan dengan cara wawancara untuk mendapatkan informasi tentang proses penjualan *smartphone* pada toko Galaxy Phone Jember. Dibawah ini gambar 3.2 merupakan alur penjualan *smartphone*.



Gambar 3.2 Alur Penjualan Smartphone

Dapat disimpulkan alur penjualan *smartphone* yang dilakukan pada toko Galaxy Phone Jember masih menggunakan sistem manual. Dari hasil penelitian

berikut nantinya dapat membantu pemilik toko Galaxy Phone Jember menjual produknya menggunakan sistem otomatis (*point of sale*).

3.4.2 Data Understanding

Tahap ini dilakukan pengumpulan data, penggambaran data, atau deskripsi data, selanjutnya mengidentifikasi masalah yang berkaitan dengan data yang tersedia. Untuk data yang tersedia berupa *brand*, tipe, dan harga. Pada penelitian ini data didapatkan dari toko Galaxy Phone Jember berdasarkan laporan penjualan dari bulan Juli 2024 hingga Desember 2024.

3.4.3 Data Preparation

Tahap ini merupakan pemilihan data yang sesuai dengan kebutuhan dari toko Galaxy Phone Jember. Salah satu metodenya yaitu *cleaning data* yang merupakan proses memperbaiki atau menghapus kesalahan, ketidakkonsistenan, dan ketidakakuratan dalam kumpulan data. Selanjutnya mengubah sebuah data manual menjadi data digital. Setelah melakukan pemilihan data yang sesuai, data akan diproses dan dianalisis yang dilakukan pada tahap *modelling*.

3.4.4 Modeling

Modeling merupakan tahap menentukan pemodelan apa yang akan digunakan kemudian menerapkannya. Biasanya berupa metode dan algoritma, kemudian diimplementasikan. Algoritma yang akan digunakan yaitu KNN (*K-Nearest Neighbor*). Setelah menentukan pemodelan, penulis melakukan perancangan terhadap sistem yang akan dibuat dengan menjabarkannya kedalam flowchart untuk mengetahui alur dari sistem. Selain itu, dilakukan pula perancangan database dengan menggunakan database MySQL. Selanjutnya dilakukan implementasi metode serta algoritma ke dalam program.

3.4.4.1 Algoritma K-Nearest Neighbor

a. Input data

Pada tahap pertama melakukan pengimputan data yang berisikan data yang telah disiapkan dan dipilih sebelumnya pada proses *Data Preperation*.

b. Menentukan Nilai K.

Nilai K ini menentukan jumlah tetangga yang akan diambil sebagai pertimbangan dalam menentukan klasifikasi titik kueri tertentu. Pemilihan

nilai K merupakan langkah penting karena dapat memengaruhi kinerja algoritma. Pilihan K harus disesuaikan dengan data yang digunakan, dan seringkali disarankan untuk memilih nilai K yang ganjil untuk menghindari kebingungan dalam klasifikasi. Strategi *cross-validation* juga dapat digunakan untuk membantu menentukan nilai K yang optimal untuk dataset tertentu.

- c. Menghitung jarak Euclidean antara data uji dengan data latih.

Menghitung jarak menggunakan rumus jarak *euclidean distance* pada persamaan 1. Berikut pada gambar 3.3 merupakan penghitungan *Euclidean distance*:

$$dis(x_1, x_2) = \sqrt{\sum_{i=0}^n (x_{1i} - x_{2i})^2}$$

Gambar 3.3 Metode Enclidean

Yang dimana:

D : jarak kedekatan

x : data training

y : data testing

n : jumlah atribut individu antara 1 s.d. n

f : fungsi similitary atribut n antara kasus x dan kasus y

i : Atribut individu antara 1 sampai dengan n

- d. Mengurutkan hasil perhitungan jarak *Euclidean* secara *ascending*.

Dari hasil perhitungan jarak, hasil tersebut diurutkan secara *ascending* atau dari terkecil ke terbesar untuk mengetahui kedekatan jarak dengan data uji.

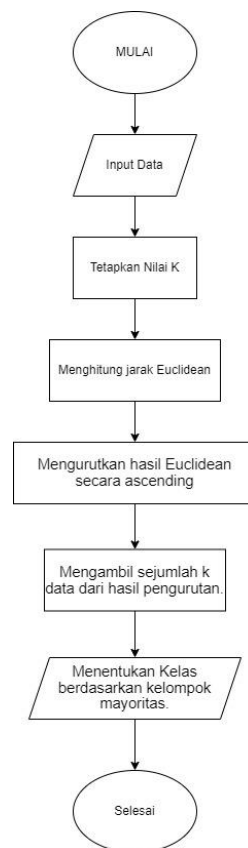
- e. Mengambil sejumlah k data dari hasil pengurutan.

Berdasarkan hasil jarak yang telah diurutkan secara *ascending* diambil sebanyak K tetangga terdekat.

- f. Menentukan Kelas berdasarkan kelompok mayoritas.

Hasil dari pemilihan tetangga terdekat sebanyak k tetangga menjadi penentu kelas dari data uji.

Berikut ini gambar 3.4 yang merupakan proses perhitungan dengan menggunakan algoritma *K-Nearest Neighbor* dalam menentukan prediksi penjualan *brand smartphone* terlaris pada toko Galaxy Phone Jember.

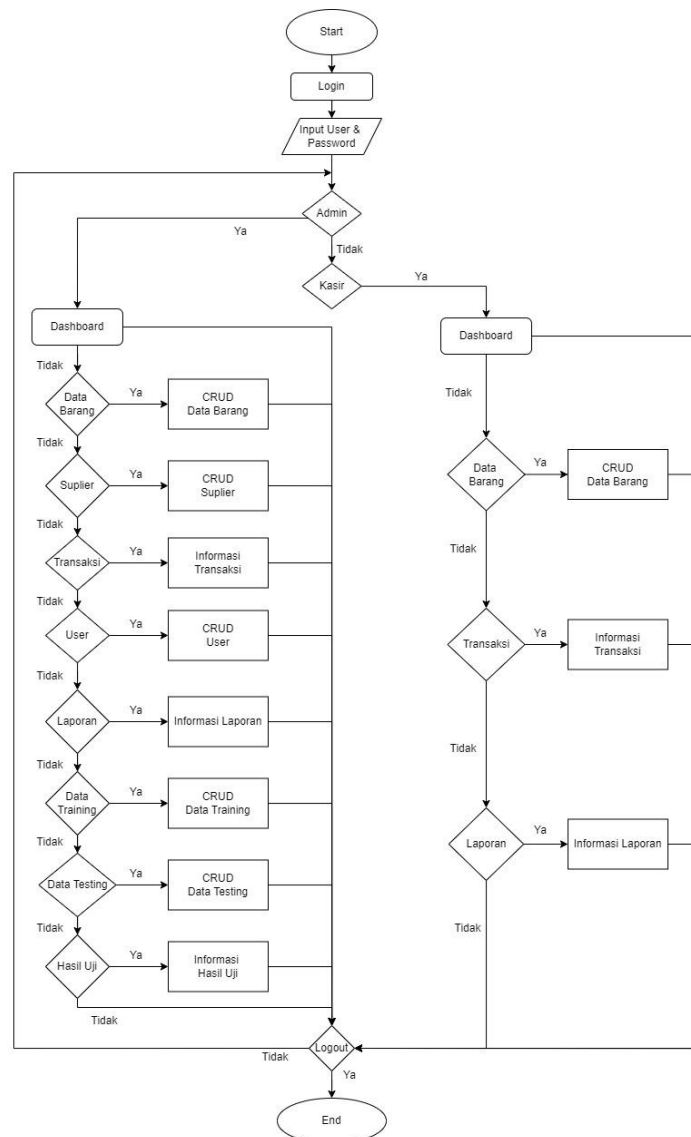


Gambar 3.4 Flowchart KNN

3.4.4.2 Model Sistem

a. *Flowchart* Sistem

Pada flowchart sistem bertujuan untuk memberikan gambaran mengenai alur sistem yang akan di lakukan. *Flowchart* berisi gambaran singkat mengenai langkah-langkah yang dilakukan sistem saat menjalankan sistem *point of sale* serta mengolah data guna mendapatkan hasil prediksi menggunakan algoritma *K-Nearest Neighbor*. Berikut ini gambar 3.5 yang merupakan flowchart sistem:



Gambar 3.5 Flowchart Sistem

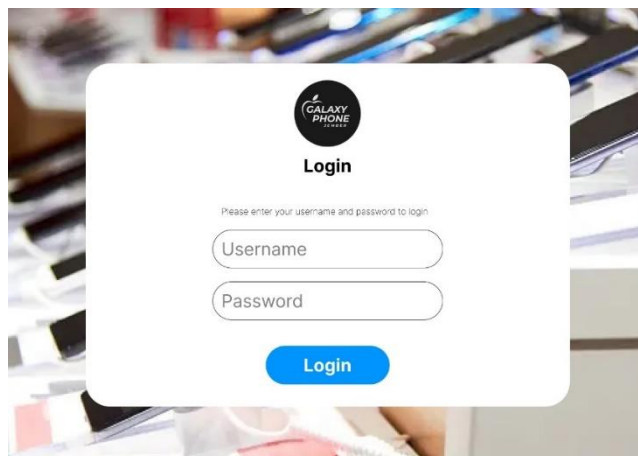
b. *Mockup* Sistem

Pada mockup sistem bertujuan untuk memberikan perencanaan desain aplikasi yang akan dibuat sesuai dengan alur sistem pada *flowchart* sistem. Berikut gambar beserta penjelasan dari *mockup* sistem:

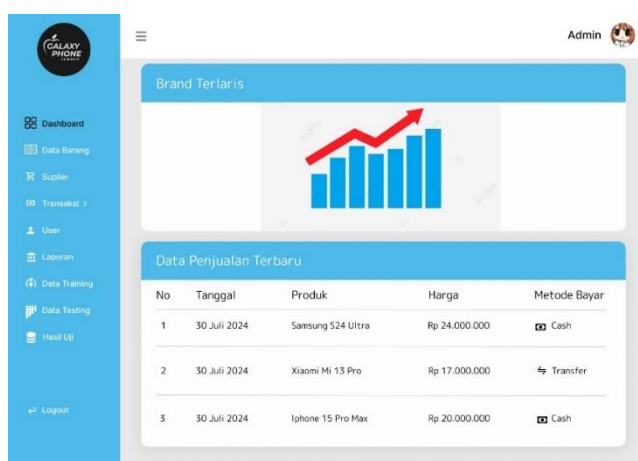
1) Halaman *Login* dan *Dashboard*

Pada gambar 3.6 merupakan *mockup* dari tampilan halaman *login*. Halaman *login* memiliki untuk menginputkan data *username* dan *password* untuk masuk ke bagian halaman *dashboard*. Nantinya pada bagian *login* terdapat 2 *role* yaitu Admin dan Kasir.

Kemudian pada gambar 3.7 merupakan *mockup* dari tampilan halaman *dashboard*. Halaman *dashboard* memiliki fungsi untuk menampilkan data *brand* terlaris serta data penjualan terbaru.



Gambar 3.6 Halaman Login

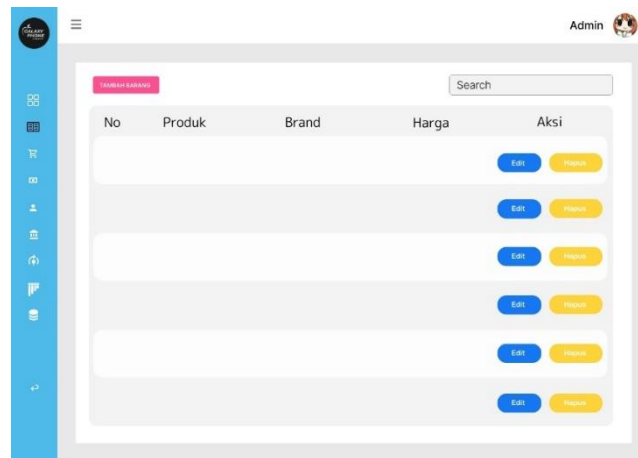


Gambar 3.7 Halaman Dashboard

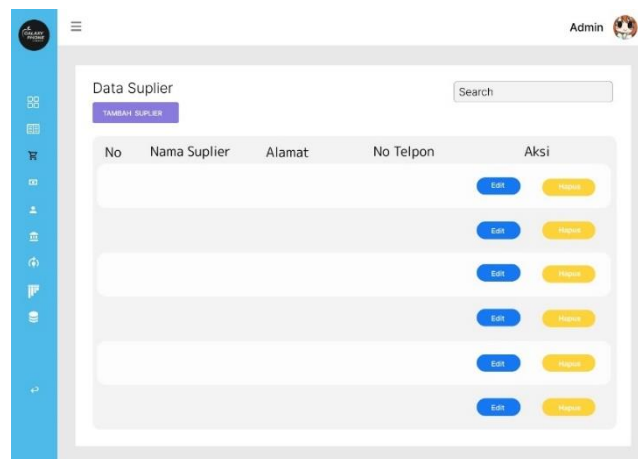
2) Halaman Data Barang dan Supplier

Pada gambar 3.8 merupakan *mockup* dari tampilan halaman data barang. Halaman data barang memiliki fungsi untuk CRUD data barang yang berisikan produk, *brand*, dan harga smartphone. Pada halaman data barang juga memiliki fungsi untuk menampilkan keseluruhan data barang yang ada.

Kemudian pada gambar 3.9 merupakan *mockup* dari tampilan halaman supplier. Halaman supplier memiliki fungsi untuk CRUD data supplier yang berisikan nama supplier, alamat, dan nomor telepon. Pada halaman supplier juga memiliki fungsi untuk menampilkan keseluruhan data supplier yang ada.



Gambar 3.8 Halaman Data Barang



Gambar 3.9 Halaman Supplier

3) Halaman Transaksi Pembelian dan Transaksi Penjualan

Pada gambar 3.10 merupakan mockup dari tampilan halaman transaksi pembelian. Halaman transaksi pembelian memiliki fungsi untuk proses pembelian *smartphone* pada supplier yang telah ditentukan. Pada halaman transaksi pembelian terdapat input nama supplier, tanggal, kode barang, nama barang, jumlah, dan harga.

Kemudian pada gambar 3.11 merupakan mockup dari tampilan halaman transaksi penjualan. Halaman transaksi penjualan memiliki fungsi untuk proses penjualan *smartphone* kepada pembeli. Pada halaman transaksi penjualan terdapat input kode barang, input nama barang, input stock, input harga, input jumlah, dan hasil transaksi penjualan.

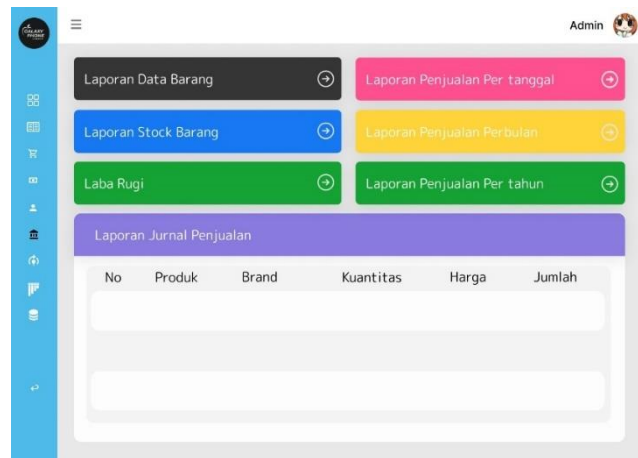
Gambar 3.10 Halaman Transaksi Pembelian

Gambar 3.11 Halaman Transaksi Penjualan

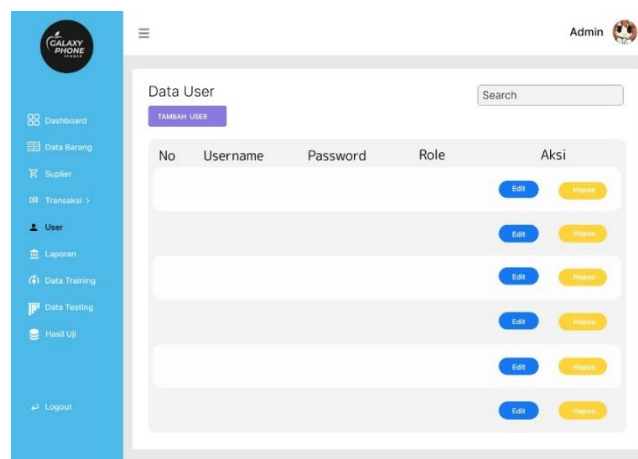
4) Halaman Laporan dan User

Pada gambar 3.12 merupakan mockup dari tampilan halaman laporan. Halaman laporan memiliki fungsi untuk menampilkan laporan yang ada pada toko Galaxy Phone Jember. Pada halaman laporan terdapat laporan data barang, laporan penjualan per tanggal, laporan stok barang, laporan penjualan perbulan, laporan penjualan pertahun, laporan laba dan rugi, dan laporan jurnal penjualan.

Kemudian pada gambar 3.13 merupakan mockup dari tampilan halaman data user. Halaman data *user* memiliki fungsi untuk CRUD data user yang berisikan *username*, *password*, *role*, dan edit *user*.



Gambar 3.12 Halaman Laporan



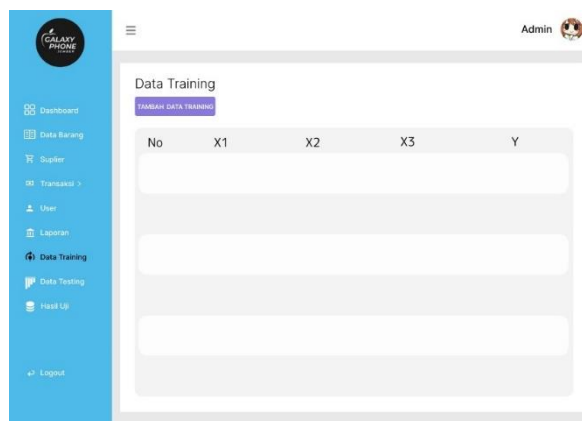
Gambar 3.13 Halaman User

5) Halaman Data *Training*, Data *Testing* dan Hasil Uji

Pada gambar 3.14 merupakan mockup dari tampilan halaman data training. Halaman data training memiliki fungsi untuk CRUD data training yang berisikan nilai x_1 , x_2 , x_3 , dan y yang merupakan data dari penjualan *brand smartphone*. Pada halaman data training juga bisa menggunakan data penjualan pertahun, perbulan, maupun pertanggal.

Kemudian pada gambar 3.15 merupakan *mockup* dari tampilan halaman data *testing*. Halaman data testing memiliki fungsi untuk CRUD data *testing* yang berisikan nilai x_1 , x_2 , x_3 , dan y yang merupakan data dari penjualan *brand smartphone*. Pada halaman data testing juga bisa menggunakan data penjualan pertahun, perbulan, maupun pertanggal.

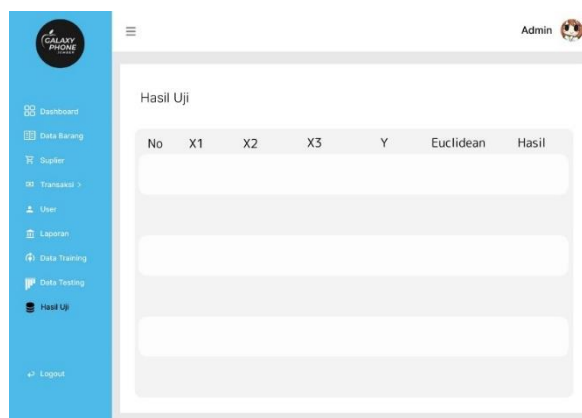
Selanjutnya pada gambar 3.16 merupakan mockup dari tampilan halaman hasil uji. Halaman hasil uji memiliki fungsi untuk menampilkan hasil uji dari data testing yang telah ditentukan. Pada halaman hasil uji berisikan nilai x_1 , x_2 , x_3 , y , Euclidean, dan hasil berupa data hasil penghitungan metode *K-Nearest Neighbor*.



Gambar 3.14 Halaman Data Training



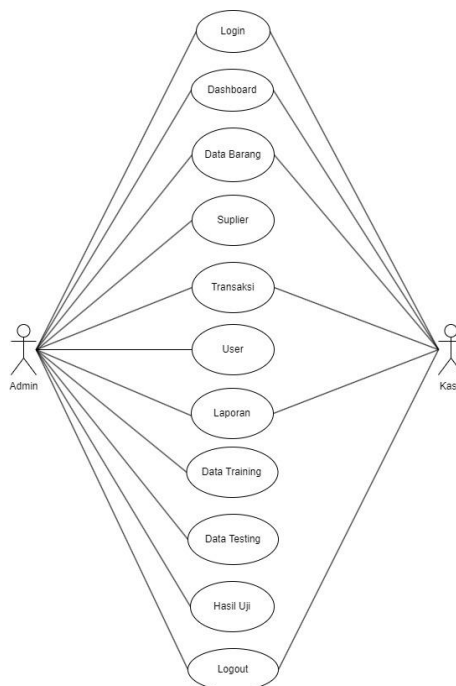
Gambar 3.15 Halaman Data Testing



Gambar 3.16 Halaman Hasil Uji

c. *Usecase Diagram*

Pada *usecase* diagram bertujuan untuk menggambarkan interaksi antara pengguna dan sistem, mendefinisikan kebutuhan fungsional, memberikan gambaran umum fitur utama, memfasilitasi komunikasi antar pemangku kepentingan, dan menjadi dasar perencanaan proyek pengembangan perangkat lunak. Berikut ini gambar 3.17 yang merupakan *usecase* diagram yang sesuai.



Gambar 3.17 Usecase Diagram

3.4.5 *Evaluation*

Tahap *evaluation* dilakukan pengujian terhadap model apakah sudah memenuhi tujuan. Kemudian hasil yang telah dihasilkan melalui model yang digunakan akan dievaluasi keakuratannya. Apakah sudah sesuai dengan tujuan penelitian. Melakukan evaluasi terhadap algoritma dan sistem yang telah dibuat sebelumnya. Tahap ini dinyatakan berhasil jika hasil telah memenuhi kebutuhan awal.

3.4.6 *Deployment*

Tahap *deployment* merupakan result dari data yang telah diolah. Biasanya berupa *final report* dari hasil evaluasi. Terdapat juga penghitungan data *mining* menggunakan tools *Rapidminer*. Serta pada tahap ini dimulainya proses

pengembangan aplikasi menggunakan yang akan dibuat menggunakan bahasa pemrograman PHP.

Kemudian dilakukannya pengujian sistem atau testing yang terbagi menjadi 2 sistem yaitu *Black Box* dan *UAT (User Acceptance Testing)*. *Black Box* testing berfungsi untuk mengevaluasi fungsionalitas eksternal sistem perangkat lunak tanpa memeriksa kode internal, dengan fokus pada validasi *input-output*, deteksi kesalahan, dan kesesuaian dengan spesifikasi yang ditentukan. Sedangkan *UAT (User Acceptance Testing)* berfungsi untuk memverifikasi bahwa sistem atau perangkat lunak memenuhi kebutuhan dan harapan pengguna akhir dalam lingkungan operasional yang sebenarnya sebelum implementasi penuh. Pada tahap ini juga nantinya dihasilkan pembuatan sistem *point of sale* beserta penghitungan data mining menggunakan metode *K-Nearest Neighbor*.

3.5 Jadwal Kegiatan

Berikut merupakan jadwal kegiatan peneliti yang direncanakan selama penelitian berlangsung:

Tabel 3.1 Jadwal Kegiatan

NO	KEGIATAN	Februari	Maret	April	Mei	Juni
1	Studi Literatur					
2	Pengumpulan Data					
3	Pengolahan Data					
4	Perancangan dan Pembuatan Sistem					
5	Pengujian					

DAFTAR PUSTAKA

- Abdullah. (2017). Bauran Promosi, Sari Kuliah, Satu Nusa.
- Adisty, N. (5. 11 2022). Mengulik Perkembangan Penggunaan Smartphone di Indonesia.
- Amalia, Y. R. (2018). Penerapan Data Mining Untuk Prediksi Penjualan Produk Elektronik Terlaris Menggunakan Metode K-Nearest Neighbor (Studi Kasus: Pt. Bintang Multi Sarana Palembang).
- Andani, M. S. (2021). Sistem Informasi Pelayanan Kependudukan Desa Lecah Berbasis Web Menggunakan PHP Dan MySQL. *Jurnal Sistem Informasi Mahakarya (JSIM) Vol. 4(1)*, 15-17.
- Arta, I. K. (2019). Data Mining Rekomendasi Calon Mahasiswa Berprestasi Di STMIK Denpasar Menggunakan Metode Technique For Others Reference By Similarity To Ideal Solution.
- Assrianti, B. (2016). Bauran Promosi, Sari Kuliah, Satu Nusa. Bandung.
- Asyraf, H. P. (2024). Implementasi Metode CRISP DM dan Algoritma Decision Tree Untuk Strategi Produksi Kerajinan Tangan pada UMKM A. *Jurnal Media Informatika Budarma Vol. 8(1)*, 94-105.
- Bahtiar, R. (2022). Implementasi Data Mining Untuk Prediksi Penjualan Kusen Terlaris Menggunakan Metode K-Nearest Neighbor.
- Black, R. (2002). Mengelola Proses Pengujian: Langkah-langkah Praktis dan Teknik Mengelola Pengujian Perangkat Keras dan Perangkat Lunak.
- Chamida, M. A. (2021). Analisa User Acceptance Testing Terhadap Sistem Informasi Pengelolaan Bedah Rumah di Dinas Perumahan Rakyat dan Kawasan Permukiman Kabupaten Jepara. *Indonesian Journal of Technology, Informatics and Science (IJTIS)*.
- Dewi, S. P. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor.
- Garang, B. D. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Smartphone Paling Laris Menggunakan Metode K-Nearest Neighbor (Studi Kasus: Pusat Ponsel & Laptop).
- Iriane, R. N. (2023). Penerapan Data Mining Untuk Prediksi Penjualan Produk Pangan Hewan Menggunakan Metode K-Nearest Neighbor.
- Kaswandari, Y. (2017). Rancang Bangun Sistem Kendali Tirai Vertical Blind dan Lampu Ruangan Melalui Smartphone Android Berbasis Arduino.

- Lewis, W. (2009). Pengujian Perangkat Lunak dan Peningkatan Kualitas Berkelanjutan.
- Maulana, H. (2016). Analisis dan Perancangan Sistem Replikasi Database MySQL dengan Menggunakan Vmware Pada Sistem Operasi Open Source. *Jurnal Nasional Informatika dan Teknologi Jaringan Vol 1(1)*.
- Musa. (2016). Bauran Promosi, Sari Kuliah, Satu Nusa. Bandung.
- Nahjan, M. R.;Heryana, N.;& Voutama, A. (2023). Implementasi Rapidminer dengan Metode Clustering K-Means untuk Analisa Penjualan Pada Toko OJ Cell. *JATI (Jurnal Mahasiswa Teknik Informatika)*.
- Perry, W. E. (2006). Effective Methods for Software Testing.
- Priatmojo, H.;Saputra, F.;Prasetyo, M. H.;Puspitasari, D.;& Nurlaela, D. (2023). Perbandingan Klasifikasi Tingkat Penjualan Buah di Supermarket dengan Pendekatan Algoritma Decision Tree, Naive Bayes dan K-Nearest Neighbor. *Jurnal INSAN (Journal of Information Systems Management Innovation), 3(1)*, 8.
- Rismala., A. I. (2023). PENERAPAN METODE K-NEAREST NEIGHBOR UNTUK PREDIKSI. *JATI (Jurnal Mahasiswa Teknik Informatika)*.
- Supono;& Putratama, V. (2018). *Pemrograman Web dengan Menggunakan PHP dan Framework Codeigniter*. Yogyakarta: Deepublish.
- Tarigan, P. M.;Hardinata, J. T.;Qurniawan, H.;Safii, M.;& Winanjaya, R. (2022). Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus: Toko Sinar Harahap).
- Wahid, A. A. (2020). Analisis Metode Waterfall Untuk Pengembangan Sistem. *Jurnal Ilmu-ilmu Informatika dan Manajemen STMIK*.
- Yanosma, D.;Johar, A.;& Anggriani, K. (2016). Implementasi Metode K-Nearest Neighbor (KNN) dan Simple Additive Weighting (SAW) dalam Pengambilan Keputusan Seleksi Penerimaan Anggota Paskibraka (Studi Kasus: Dinas Pemuda dan Olahraga Provinsi Bengkulu).
- Yolanda, I. F. (2021). Penerapan Data Mining Untuk Prediksi Penjualan Produk Roti Terlaris Pada PT.Nippon Indosari Corpindo Tbk Menggunakan Metode K-Nearest Neighbor.