

BAB III

METODE PENELITIAN

A. Jenis Penelitian

Penelitian ini menggunakan pendekatan Research and Development (R&D) sebagai metode utama karena penelitian tidak hanya berfokus pada analisis permasalahan, tetapi juga menghasilkan produk berupa sistem informasi pengelompokan pelanggan berbasis web yang dapat diterapkan pada Earnest Apparel. Metode R&D terdiri atas dua tahapan utama, yaitu tahap *research* (penelitian) dan *development* (pengembangan). Tahap *research* difokuskan pada analisis data transaksi pelanggan dengan menggunakan pendekatan *Knowledge Discovery in Databases* (KDD), yang meliputi proses seleksi data, *pre-processing*, transformasi data, serta penerapan algoritma *K-Means Clustering* untuk mengelompokkan pelanggan berdasarkan pola pembelian.

Hasil dari proses klusterisasi tersebut kemudian dianalisis dan dievaluasi untuk menghasilkan segmentasi pelanggan yang dapat dijadikan sebagai dasar informasi dalam pengambilan keputusan. Selanjutnya, tahap *development* diarahkan pada proses perancangan dan pembangunan sistem informasi pengelompokan pelanggan berbasis *web* yang mengintegrasikan hasil analisis algoritma *K-Means*. Dalam pengembangannya, penelitian ini menerapkan model *Waterfall* karena memiliki tahapan yang terstruktur dan sistematis, meliputi analisis kebutuhan, perancangan sistem, implementasi, pengujian, serta pemeliharaan. Sistem yang dihasilkan dirancang untuk mengelola data transaksi pelanggan sekaligus menampilkan visualisasi hasil segmentasi pelanggan sebagai pendukung pengambilan keputusan bisnis pada Earnest Apparel.

1. Tahapan Knowledge Discovery in Databases (KDD)

Pada penelitian ini, metode Knowledge Discovery in Database (KDD) digunakan sebagai tahapan dalam mengolah data mentah transaksi pelanggan yang diperoleh dari Earnest Apparel. Tahapan KDD dilakukan untuk mempersiapkan data agar siap dianalisis menggunakan metode *K-Means Clustering*. Melalui tahapan KDD, data mentah yang awalnya belum terstruktur dapat diolah menjadi data yang

berkualitas sehingga menghasilkan pengelompokan pelanggan yang lebih akurat dan mudah diinterpretasikan.

Tahapan KDD yang diterapkan dalam penelitian ini terdiri dari *selection*, *preprocessing*, *transformation*, *data mining*, *evaluation*, dan *knowledge presentation*, yang dijelaskan sebagai berikut.

a. Seleksi data

Tahap seleksi data dilakukan untuk memilih data yang relevan dengan tujuan penelitian. Data yang digunakan berasal dari basis data transaksi Earnest Apparel, yang meliputi data pelanggan, data transaksi penjualan, dan data produk. Seleksi data bertujuan untuk memastikan bahwa hanya data yang berkaitan langsung dengan aktivitas pembelian pelanggan yang digunakan dalam proses analisis, sehingga pola yang dihasilkan lebih akurat dan relevan.

b. *Preprocessing data* (pembersihan data)

Tahap *preprocessing* merupakan tahap penting dalam KDD yang bertujuan untuk meningkatkan kualitas data sebelum dilakukan proses analisis. Pada penelitian ini, *preprocessing* data dilakukan melalui beberapa tahapan sebagai berikut.

1) *Data cleaning*

Data cleaning dilakukan untuk menghilangkan atau memperbaiki data yang tidak *valid*, seperti data transaksi ganda (duplikasi data), data kosong atau tidak lengkap (*missing value*), serta kesalahan pencatatan data transaksi atau pelanggan. Proses ini bertujuan untuk menghindari kesalahan perhitungan dan bias hasil *clustering*.

2) *Data integration*

Apabila data transaksi, data pelanggan, dan data produk tersimpan pada tabel yang terpisah, maka dilakukan proses penggabungan data (integrasi) agar seluruh informasi yang dibutuhkan dapat dianalisis secara terpadu.

3) *Data reduction*

Tahap data *reduction* dilakukan untuk menyederhanakan data tanpa menghilangkan informasi penting yang dibutuhkan dalam proses *clustering*. Atribut yang dipertahankan hanyalah atribut yang berpengaruh

terhadap pola pembelian pelanggan, sehingga proses *clustering* menjadi lebih efektif dan efisien.

c. *Transformation* data (transformasi data)

Tahap transformasi data dilakukan untuk mengubah data mentah menjadi bentuk numerik yang dapat diproses oleh algoritma *K-Means*. Pada tahap ini, data transaksi diolah untuk menghasilkan variabel analisis sebagai berikut: frekuensi pembelian pelanggan, total jumlah produk yang dibeli, dan total nilai transaksi pelanggan. Selanjutnya, dilakukan normalisasi data menggunakan metode *Min-Max Normalization* untuk menyamakan skala nilai antar variabel, sehingga tidak ada satu variabel yang mendominasi perhitungan jarak pada proses *clustering*.

Metode normalisasi yang digunakan dalam penelitian ini adalah metode *Min-Max Normalization*, dengan rumus sebagai berikut:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3.1)$$

Keterangan:

- X' : nilai hasil normalisasi
- X : nilai data asli
- X_{min} : nilai minimum
- X_{max} : nilai maksimum

Hasil normalisasi berada pada rentang nilai 0 hingga 1.

d. *Data mining (clustering)* menggunakan algoritma *K-Means*

Tahap data mining merupakan tahap inti dalam proses KDD. Pada tahap ini diterapkan algoritma *K-Means Clustering* untuk mengelompokkan pelanggan berdasarkan kesamaan pola pembelian. Tahapan algoritma *K-Means* yang digunakan dalam penelitian ini adalah sebagai berikut:

1) Menentukan jumlah *cluster*

Penentuan jumlah *cluster* (k) dalam penelitian ini ditetapkan sebanyak tiga *cluster*. Penetapan nilai k tersebut didasarkan pada kebutuhan bisnis

Earnest Apparel dalam mengelompokkan pelanggan ke dalam tiga kategori utama, yaitu pelanggan baru, pelanggan langganan, dan pelanggan loyal. Pembagian kategori ini telah digunakan oleh pihak Earnest Apparel sebagai dasar dalam memahami karakteristik pelanggan berdasarkan intensitas dan nilai pembelian. Oleh karena itu, penggunaan tiga *cluster* dianggap paling relevan untuk merepresentasikan segmentasi pelanggan yang sesuai dengan kondisi dan tujuan bisnis perusahaan.

2) Menentukan *centroid* awal

Setelah nilai k ditentukan, langkah selanjutnya adalah menentukan *centroid* awal. Pada penelitian ini, *centroid* awal dipilih secara acak dari data yang ada sebagai titik pusat awal masing-masing *cluster*.

3) Menghitung jarak data ke *centroid*

Setiap data dihitung jaraknya terhadap seluruh *centroid* menggunakan rumus *Euclidean Distance* menggunakan persamaan 2.1. Hasil perhitungan jarak digunakan sebagai dasar untuk menentukan kedekatan data terhadap masing-masing *cluster*.

4) Mengelompokkan data ke *cluster* terdekat

Berdasarkan hasil perhitungan jarak, setiap data ditempatkan ke dalam *cluster* yang memiliki nilai jarak terdekat terhadap *centroid*. Proses ini menghasilkan pembagian data ke dalam beberapa *cluster* awal.

5) Menghitung *centroid* baru

Setelah seluruh data terkelompok, *centroid* baru dihitung berdasarkan rata-rata data pada masing-masing *cluster* dengan menggunakan rumus pada persamaan 2.2. *Centroid* baru digunakan sebagai acuan pada iterasi selanjutnya.

6) Iterasi hingga konvergen

Langkah perhitungan jarak, pengelompokan data, dan pembaruan *centroid* diulangi secara iteratif hingga memenuhi kriteria berhenti, yaitu:

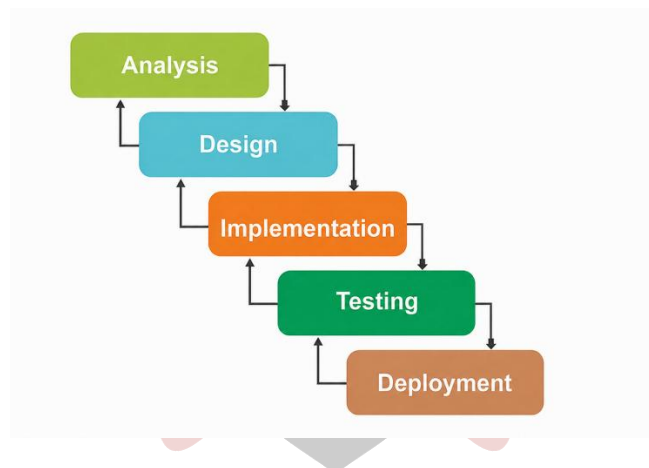
- 1) Tidak terjadi perubahan nilai *centroid*, atau
- 2) Tidak terdapat perpindahan data antar *cluster*.

e. *Evaluation*

Tahap evaluasi dilakukan untuk menilai kualitas hasil *clustering* yang dihasilkan oleh algoritma *K-Means*. Pada penelitian ini, evaluasi dilakukan menggunakan *Davies–Bouldin Index* (DBI). DBI digunakan untuk mengukur tingkat kohesi dalam *cluster* dan separasi antar *cluster*. Nilai DBI yang lebih kecil menunjukkan bahwa hasil pengelompokan pelanggan semakin optimal.

2. Model Pengembangan Sistem

Model pengembangan sistem dalam penelitian ini dapat dilihat pada Gambar 3.1 sebagai berikut.



Gambar 3. 1 Model pengembangan sistem

a. *Analysis*

Tahap analisis kebutuhan dilakukan untuk mengidentifikasi kebutuhan fungsional sistem dalam mendukung proses pengelompokan pelanggan pada Earnest Apparel. Sistem ini berbasis *web* dan melibatkan dua aktor utama, yaitu admin dan pemilik, yang memiliki peran serta hak akses berbeda. Admin berperan sebagai pengelola sistem dengan hak akses penuh, meliputi proses *login*, pengelolaan data produk, data pelanggan, dan data transaksi, serta menjalankan proses analisis menggunakan algoritma *K-Means*. Admin juga bertanggung jawab dalam melihat dan mengevaluasi hasil pengelompokan pelanggan ke dalam beberapa klaster, seperti pelanggan loyal, pelanggan langganan, dan pelanggan baru, sebagai dasar pengambilan keputusan bisnis. Sementara itu,

pemilik memiliki hak akses terbatas, yaitu hanya untuk melihat hasil segmentasi pelanggan yang ditampilkan dalam bentuk laporan dan visualisasi. Pemilik tidak memiliki kewenangan untuk mengubah data maupun menjalankan proses analisis. Dengan pembagian peran tersebut, sistem dapat berjalan secara terstruktur, aman, dan efektif dalam mendukung proses pengambilan keputusan berdasarkan pola pembelian pelanggan.

b. *Design*

Tahap ini berfokus pada pembuatan rancangan sistem yang akan menjadi acuan dalam proses pengkodean. Desain dilakukan menggunakan beberapa diagram pendukung, yaitu *Use Case Diagram*, *Activity Diagram*, dan *Entity Relationship Diagram*. *Use Case Diagram* digunakan untuk menggambarkan interaksi admin dengan fitur-fitur sistem, seperti *input* data transaksi, menjalankan proses *K-Means*, dan melihat hasil analisis. *Activity Diagram* menggambarkan alur proses sistem, mulai dari admin memasukkan data penjualan hingga sistem menghasilkan kluster pelanggan. ERD digunakan untuk merancang struktur *database* yang menyimpan data pelanggan, produk, transaksi, dan hasil klusterisasi. Hasil desain ini menjadi pedoman dalam proses implementasi sistem agar terstruktur dan sesuai kebutuhan.

c. *Implementation*

Tahap ini merupakan proses implementasi desain sistem ke dalam bentuk aplikasi web. Sistem pengelompokan pelanggan dikembangkan menggunakan bahasa pemrograman PHP dengan dukungan HTML, CSS, dan *JavaScript* sebagai antarmuka, serta *MySQL* sebagai basis data. Pada tahap ini, seluruh proses perhitungan algoritma *K-Means* diimplementasikan langsung menggunakan PHP. Data transaksi pelanggan diambil dari *database*, kemudian diproses di dalam sistem menggunakan logika perhitungan *K-Means* yang telah diintegrasikan ke dalam kode program PHP. Proses perhitungan meliputi penentuan *centroid* awal, perhitungan jarak setiap data terhadap *centroid*,

pengelompokan data ke dalam *cluster* terdekat, serta pembaruan *centroid* hingga mencapai kondisi konvergen. Hasil dari proses klasterisasi kemudian disimpan kembali ke dalam *database* dan ditampilkan pada *dashboard* sistem sebagai informasi segmentasi pelanggan.

d. *Testing*

Pengujian dilakukan untuk memastikan bahwa seluruh fitur sistem informasi pengelompokan pelanggan berbasis *web* berjalan sesuai dengan kebutuhan Earnest Apparel. Pengujian menggunakan dua metode:

- 1) *Black Box Testing*, yang digunakan untuk memeriksa apakah setiap fitur sistem seperti *input* transaksi, proses klasterisasi, dan tampilan hasil berfungsi dengan benar tanpa memeriksa struktur kode internal.
- 2) *User Acceptance Testing (UAT)*, yaitu pengujian oleh admin untuk menilai kemudahan penggunaan, keakuratan hasil klasterisasi, serta kesesuaian sistem dalam membantu proses pengelolaan data pelanggan dan pemahaman perilaku pelanggan.

Pengujian ini memastikan bahwa sistem siap digunakan dan memberikan manfaat nyata bagi Earnest Apparel.

Berdasarkan tahapan perancangan dan implementasi sistem yang telah dijelaskan, sistem yang dikembangkan menghasilkan *output* berupa informasi hasil pengelompokan pelanggan berdasarkan pola pembelian menggunakan algoritma *K-Means*. *Output* sistem menampilkan pembagian pelanggan ke dalam tiga kelompok, yaitu pelanggan loyal, pelanggan langganan, dan pelanggan baru, yang disajikan dalam bentuk tabel dan visualisasi grafik pada *dashboard* sistem. Informasi ini digunakan sebagai dasar pendukung pengambilan keputusan dalam penyusunan strategi penjualan yang lebih tepat sasaran.

e. *Deployment*

Pada tahap *deployment*, sistem tidak dilakukan proses *hosting* ke *server* karena aplikasi ini hanya digunakan oleh dua pengguna utama di lingkungan toko. Oleh karena itu, implementasi sistem cukup dilakukan

dengan cara menginstall aplikasi langsung pada komputer yang tersedia di toko. Sistem kemudian digunakan secara lokal untuk melakukan proses klasterisasi data pelanggan. Pengoperasian sistem dilakukan secara berkala, yaitu setiap dua minggu atau satu bulan sekali, sesuai kebutuhan analisis data pelanggan. Dengan cara ini, sistem tetap dapat berjalan efektif tanpa memerlukan infrastruktur server yang kompleks, namun tetap mampu menghasilkan informasi pengelompokan pelanggan secara rutin dan akurat.

B. Setting Penelitian

1. Lokasi Penelitian

Penelitian ini dilakukan di Percetakan Earnest Apparel yang berlokasi di Kota Padang. Pemilihan lokasi ini didasarkan pada kebutuhan untuk memperoleh informasi yang mendalam terkait proses penjualan produk yang sedang berjalan, alur pengolahan data pelanggan, serta permasalahan yang muncul dalam pengelolaan transaksi dan identifikasi pola pembelian pelanggan.

Melalui penelitian ini, penulis berupaya mengembangkan sistem informasi pengelompokan pelanggan berbasis web yang memanfaatkan algoritma *K-Means* untuk melakukan segmentasi pelanggan berdasarkan pola pembelian. Diharapkan, hasil implementasi sistem ini dapat meningkatkan efisiensi operasional Earnest Apparel, khususnya dalam pengelolaan data penjualan, rekomendasi produk, serta pengambilan keputusan terkait strategi pemasaran yang lebih tepat sasaran.

2. Jadwal Penelitian

Rangkaian perencanaan waktu penelitian dapat dilihat pada gambar 3.1

Tabel 3. 1 Jadwal penelitian

Kegiatan/Waktu	Bulan 1				Bulan 2				Bulan 3				Bulan 4			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Studi literatur																
Pengumpulan Data																

Kegiatan/Waktu	Bulan 1				Bulan 2				Bulan 3				Bulan 4			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Data <i>preprocessing</i> (<i>Cleaning, integration, reduction</i>)																
Penerapan <i>K-Means Clustering</i> (<i>perhitungan manual</i>)																
Analisis hasil <i>clustering</i>																
Analisis kebutuhan sistem																
Perancangan sistem																
Implementasi sistem																
Pengujian sistem																
Deployment sistem																
Penyusunan laporan skripsi																

C. Alat/Instrumen Penelitian

1. Perangkat Keras

Perangkat keras (*hardware*) merupakan komponen fisik komputer yang digunakan untuk mendukung proses pengembangan, pemrosesan data, dan pengujian aplikasi. Dalam penelitian “Implementasi Sistem Informasi Pengelompokan Pelanggan Berbasis *Website* Menggunakan Algoritma *K-Means* Berdasarkan Identifikasi Pola Pembelian Pelanggan Di Earnest Apparel”, perangkat keras berperan penting untuk memastikan proses komputasi, pemodelan algoritma, serta pengujian sistem berjalan dengan baik.

Pada penelitian ini, penulis menggunakan laptop pribadi sebagai perangkat utama dalam merancang dan membangun sistem informasi pengelompokan

pelanggan berbasis web. Adapun spesifikasi minimum perangkat keras yang digunakan sebagai berikut:

Tabel 3. 2 Perangkat keras

Perangkat	Deskripsi
<i>CPU: AMD® Ryzen™ 5 7520U 2.8GHz</i>	Prosesor ini memiliki kinerja yang cukup baik untuk mendukung pengembangan dan pengujian sistem. Kecepatan prosesor nya mendukung aplikasi berbasis <i>web</i> .
<i>RAM: 8.00 GB</i>	RAM yang digunakan cukup besar untuk mendukung <i>multitasking</i> dan menjalankan berbagai aplikasi bersamaan.
<i>Storage: 477GB SSD</i>	Penyimpanan SSD akan membantu dalam mempercepat proses <i>rendering</i> aplikasi serta pengelolaan data dalam sistem yang sedang dikembangkan.

2. Perangkat Lunak

Perangkat lunak yang digunakan bertujuan untuk memfasilitasi pengolahan data dan pengembangan sistem informasi.

Tabel 3. 3 Perangkat lunak

Perangkat	Deskripsi
Sistem Operasi: <i>Windows 11 Home Single Language</i>	<i>Operating System</i> yang digunakan adalah versi terbaru dari <i>windows</i> yang memiliki banyak fitur serta mendukung dalam pengembangan sistem <i>website</i> .

Perangkat	Deskripsi
<i>Web Browser: Google Chrome, Microsoft Edge</i>	Browser yang berguna untuk menampilkan hasil dari visual yang dibuat dalam pengkodean sebelumnya.
Aplikasi Desain: <i>Balsamiq, Figma</i>	Aplikasi yang digunakan berguna untuk mendesain atau merancang sistem yang akan dibuat (UI/UX).
<i>Visual Studio Code</i>	Aplikasi yang digunakan berguna untuk penulisan pengkodean dalam pengembangan <i>website</i> .
<i>MySQL</i>	<i>Software</i> yang berguna untuk penyimpanan <i>database</i> sistem.
XAMPP	Aplikasi yang berguna untuk menjalankan hasil maupun testing sistem yang diperlukan dengan menggunakan <i>localhost</i> sebagai <i>server</i> yang digunakan.
Microsoft Excel	perangkat lunak pengolah data berbasis <i>spreadsheet</i> yang digunakan untuk memasukkan, mengelola, dan mengolah data dalam bentuk tabel.

D. Teknik Pengumpulan Data

Teknik pengumpulan data digunakan untuk memperoleh informasi yang relevan dalam proses pengembangan sistem informasi pengelompokan pelanggan berbasis *web* pada Earnest Apparel. Data yang dikumpulkan digunakan sebagai dasar dalam memahami kebutuhan pengguna, menggambarkan alur bisnis yang berjalan, serta menyediakan data set untuk proses analisis pola pembelian

menggunakan algoritma *K-Means*. Dalam penelitian ini, penulis menggunakan empat teknik pengumpulan data, yaitu wawancara, observasi, dokumentasi, dan studi pustaka.

1. Wawancara

Wawancara dilakukan secara langsung dengan pemilik dan staf administrasi Earnest Apparel yang terlibat dalam aktivitas penjualan maupun pengelolaan data transaksi. Melalui wawancara ini, penulis menggali informasi mengenai kendala yang dialami dalam pengelolaan data pelanggan, serta kebutuhan fitur yang diharapkan pada sistem informasi pengelompokan pelanggan berbasis web. Wawancara ini juga memberikan gambaran mengenai bagaimana sistem baru dapat membantu meningkatkan efektivitas penjualan dan mendukung pengambilan keputusan berdasarkan perilaku pelanggan.

2. Observasi

Observasi dilakukan dengan mengamati langsung aktivitas operasional Earnest Apparel, seperti pencatatan transaksi, pengaturan data pelanggan, dan alur proses penjualan. Observasi ini bertujuan untuk memahami sistem manual yang berjalan saat ini, mengidentifikasi kekurangan yang perlu diperbaiki, dan memastikan bahwa fitur-fitur pada sistem yang dikembangkan benar-benar sesuai dengan kebutuhan pengguna. Hasil observasi memberikan gambaran nyata mengenai bagaimana data transaksi dikumpulkan dan digunakan dalam kegiatan bisnis sehari-hari.

3. Dokumentasi

Dokumentasi dilakukan dengan mengumpulkan data historis penjualan yang dimiliki Earnest Apparel, termasuk data total transaksi, frekuensi pembelian, jumlah pembelian, dan data pelanggan. Data tersebut menjadi bahan utama untuk dianalisis menggunakan algoritma *K-Means* sehingga dapat menghasilkan segmentasi pelanggan berdasarkan pola pembelian mereka. Teknik dokumentasi ini sangat penting karena kualitas data yang dikumpulkan akan memengaruhi hasil klusterisasi dan efektivitas sistem yang dikembangkan.

4. Studi Pustaka

Studi pustaka dilakukan dengan menelaah berbagai sumber seperti jurnal ilmiah, buku, artikel penelitian, dan referensi terkait lainnya yang membahas sistem informasi, *data mining*, algoritma *K-Means*, dan sistem informasi pengelompokan pelanggan berbasis *web*. Studi pustaka bertujuan untuk memperkuat landasan teori yang digunakan dalam penelitian, memahami konsep dan teknologi yang relevan, serta memperoleh informasi mengenai penelitian terdahulu yang dapat menjadi acuan dalam merancang dan membangun sistem yang lebih optimal.

