

PROPOSAL PENELITIAN

PENERAPAN ALGORITMA K-MEANS PADA KLASTERISASI PENGELUARAN KEUANGAN PT. TASPEN (PERSERO) KANTOR CABANG BOGOR

Oleh:
Putri Rezekika Br Ginting
065121117



PROGRAM STUDI ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PAKUAN
BOGOR
2025

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, karena rahmat dan hidayah- Nya penulis dapat menyelesaikan skripsi ini yang berjudul **“Penerapan Algoritma K-Means Pada Klasterisasi Pengeluaran Keuangan PT. Taspen (Persero) Kantor Cabang Bogor ”**. Penulisan tugas akhir ini merupakan salah satu syarat memperoleh gelar Sarjana Komputer di Program Studi Ilmu Komputer FMIPA UNPAK Bogor.

Dalam penulisan tugas akhir ini, penulis dengan senang hati ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Dr. Hermawan Taher, selaku Pembimbing Utama yang telah memberikan dorongan moril dan motivasi kepada penulis.
2. Boldson H.S., S.Kom., MMSI., selaku pembimbing Pendamping yang telah memberikan bimbingan, semangat dan motivasi.
3. Arie Qur'ania, M.Kom., Selaku Ketua Program Studi Ilmu Komputer FMIPA Universitas Pakuan Bogor.
4. Keluarga saya yang tanpa henti memberikan perhatian, motivasi, dan dukungan dalam setiap langkah yang saya tempuh, sehingga saya dapat menyelesaikan Proposal ini dengan sebaik-baiknya.
5. Rekan-rekan serta sahabat yang telah memberikan dukungan, semangat, dan kebersamaan dalam menghadapi setiap tantangan akademik.

Penulis menyadari bahwa laporan ini masih jauh dari kesempurnaan. Oleh karena itu, penulis dengan senang hati menerima segala saran dan kritik yang membangun demi penyempurnaan tugas akhir ini. Semoga Allah SWT membalas segala kebaikan semua pihak yang telah membantu dalam penyusunan laporan ini. Akhir kata, semoga laporan ini dapat memberikan manfaat bagi kita semua. Aamiin.

Bogor, 24 Maret 2025

Putri Rezekika Br Ginting
065121117

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pada era digital saat ini, sistem informasi berbasis web semakin banyak diadopsi oleh berbagai instansi, termasuk lembaga pemerintahan dan perusahaan, untuk menggantikan serta mengotomatisasi proses kerja yang sebelumnya dilakukan secara manual. PT. TASPEN (Persero) merupakan Badan Usaha Milik Negara (BUMN) yang bergerak di bidang asuransi tabungan hari tua dan dana pensiun bagi Aparatur Sipil Negara (ASN) serta Pejabat Negara. Sebagai entitas yang berfokus pada layanan keuangan, PT. TASPEN memiliki tanggung jawab besar dalam memastikan pengelolaan transaksi keuangan berjalan secara akurat, efisien, dan dapat dipertanggungjawabkan.

Salah satu bagian yang berperan penting dalam mendukung akuntabilitas tersebut adalah divisi Rumah Tangga, yang bertugas menyusun Lembar Pertanggungjawaban Transaksi (LPT) sebagai dokumen resmi atas pengeluaran kantor. Namun, proses penyusunan Lembar Pertanggungjawaban Transaksi (LPT) yang masih dilakukan secara manual melalui Microsoft Excel memiliki berbagai keterbatasan, terutama dalam menangani data berjumlah besar dan kompleks. Prosedur ini tidak hanya rentan terhadap kesalahan input, tetapi juga membutuhkan waktu lebih lama karena tidak terintegrasi secara otomatis, sehingga mengurangi efisiensi kerja dan meningkatkan beban administratif.

Untuk mengatasi permasalahan tersebut, dikembangkan sebuah sistem informasi berbasis web bernama Laporanku. Sistem ini dirancang untuk mempermudah proses pencatatan dan pengelolaan Lembar Pertanggungjawaban Transaksi (LPT) secara digital. Melalui formulir yang terstruktur, pengguna dapat menginput data secara cepat dan akurat, serta mengunduh hasil laporan dalam berbagai format seperti PDF dan Excel. Selain itu, sistem ini juga menyediakan fitur pengiriman laporan yang terintegrasi dengan sistem tanda tangan digital, sehingga mendukung proses verifikasi dan otorisasi yang lebih praktis.

Lebih dari sekadar otomasi administratif, sistem Laporanku juga dilengkapi dengan algoritma K-Means Clustering untuk mengelompokkan data pengeluaran ke dalam beberapa klaster berdasarkan kemiripan nilai, seperti kategori pengeluaran kecil, sedang, atau besar. K-Means merupakan salah satu algoritma unsupervised learning dalam data mining yang terbukti efektif dalam mengidentifikasi pola tersembunyi pada data numerik. Dengan adanya fitur klasterisasi ini, pengguna dapat memperoleh wawasan tambahan mengenai pola dan karakteristik pengeluaran, yang dapat dimanfaatkan sebagai dasar dalam pengambilan keputusan keuangan yang lebih tepat sasaran.

Efektivitas metode K-Means dalam pengelompokan data keuangan telah dibuktikan oleh berbagai penelitian sebelumnya. Penelitian oleh Setianingsih dan Ali (2023) yang berjudul "Klasterisasi Pengeluaran Kas di Desa Pamengkang Menggunakan Metode K-Means" berhasil membentuk klaster pengeluaran yang valid dengan pendekatan KDD dan evaluasi menggunakan Davies-Bouldin Index sebesar 0,012. Sementara itu, Arofah dan Hartati (2020) dalam penelitiannya "Penerapan Data Mining untuk Klasterisasi Data Anggaran Pendapatan dan Belanja

Daerah Menggunakan Algoritma K-Means" menunjukkan bahwa metode ini mampu memetakan distribusi anggaran pendidikan ke dalam beberapa klaster berdasarkan nominal. Penelitian lain oleh Anggraeni dan Kuswanto (2024) berjudul "Implementasi Algoritma K-Means Data Mining untuk Clustering Data Transaksi Pengeluaran Toko Sepatu" juga menunjukkan bahwa metode ini efektif dalam menyusun data pengeluaran menjadi kelompok yang lebih terstruktur dan informatif.

Berdasarkan hasil-hasil tersebut, dapat disimpulkan bahwa algoritma K-Means memiliki potensi yang besar untuk diterapkan dalam berbagai bidang, termasuk keuangan dan manajemen operasional. Oleh karena itu, penelitian ini difokuskan pada pengembangan sistem informasi berbasis web yang mendukung digitalisasi penyusunan Lembar Pertanggungjawaban Transaksi (LPT) sekaligus menerapkan algoritma K-Means untuk melakukan klasterisasi data pengeluaran pada PT. TASPEN (Persero) Kantor Cabang Bogor. Diharapkan sistem ini tidak hanya mampu meningkatkan efisiensi dalam dokumentasi transaksi, tetapi juga memberikan nilai tambah melalui analisis data keuangan yang lebih mendalam dan informatif.

1.2. Tujuan

Penelitian ini bertujuan untuk memberikan solusi dalam pengelompokan data pengeluaran secara efektif melalui pengembangan sistem informasi. Secara khusus, tujuan dari penelitian ini adalah sebagai berikut:

1. Membangun sistem informasi berbasis web untuk mendukung digitalisasi proses penyusunan Lembar Pertanggungjawaban Transaksi (LPT) di PT. TASPEN (Persero) Kantor Cabang Bogor.
2. Menerapkan algoritma K-Means Clustering pada data pengeluaran untuk mengelompokkan transaksi.
3. Menghasilkan informasi klaster pengeluaran yang dapat dijadikan dasar dalam evaluasi dan perencanaan anggaran secara lebih terstruktur.

1.3. Ruang Lingkup

Untuk memperjelas batasan dan fokus dari penelitian ini, ruang lingkup yang dikaji meliputi beberapa aspek berikut:

1. Data yang digunakan dalam penelitian ini berasal dari PT. TASPEN (Persero) Kantor Cabang Bogor, khususnya bagian Rumah Tangga yang bertugas menyusun Lembar Pertanggungjawaban Transaksi (LPT). Jumlah data yang dianalisis sebanyak 1.234 transaksi, yang dikumpulkan dalam rentang waktu Mei 2020 hingga Oktober 2023.
2. Penelitian ini berfokus pada klasterisasi data Lembar Pertanggungjawaban Transaksi (LPT) berdasarkan nominal transaksi, dengan cakupan wilayah terbatas pada transaksi di PT. TASPEN (Persero) Kantor Cabang Bogor. Sistem yang dikembangkan hanya mencakup digitalisasi proses Lembar Pertanggungjawaban Transaksi (LPT) , tanpa melibatkan keseluruhan sistem keuangan perusahaan.
3. Pengembangan sistem dilakukan menggunakan Python dengan framework Flask atau FastAPI untuk backend, serta React.js dan CSS untuk frontend. Penyimpanan data dilakukan menggunakan MySQL sebagai basis data. Untuk proses pengolahan dan klasterisasi data, digunakan Google Collab sebagai lingkungan komputasi yang mendukung penerapan algoritma K-Means.

4. Sistem ini dirancang untuk digunakan oleh staf bagian Rumah Tangga PT. TASPEN (Persero) Kantor Cabang Bogor dalam menyusun, mengelola, dan mencetak Lembar Pertanggungjawaban Transaksi (LPT) secara digital.

1.4. Manfaat Penelitian

Penelitian ini dilaksanakan dengan harapan dapat memberikan kontribusi positif, baik secara teoritis maupun praktis, melalui manfaat sebagai berikut:

1. Memberikan solusi teknologi informasi berupa sistem berbasis web yang mendukung digitalisasi proses penyusunan Lembar Pertanggungjawaban Transaksi (LPT), sehingga dapat meningkatkan efisiensi, akurasi, dan akuntabilitas pencatatan LPT di PT. TASPEN (Persero) Kantor Cabang Bogor.
2. Menyediakan kemudahan penggunaan sistem bagi pengguna, khususnya staf Rumah Tangga, melalui antarmuka yang user-friendly serta fitur-fitur yang mendukung otomatisasi dalam proses penginputan dan pelaporan data transaksi.
3. Memberikan kontribusi akademik dan praktis berupa penerapan algoritma K-Means Clustering yang dapat dijadikan referensi dalam pengembangan sistem analisis data Lembar Pertanggungjawaban Transaksi (LPT) serta menjadi pertimbangan dalam pengambilan keputusan manajerial berbasis data.

BAB II

TINJAUAN PUSTAKA

2.1. Landasan Teori

2.1.1. Struktur Biaya Pengeluaran Perusahaan

Struktur biaya (*cost structure*) dalam konteks perusahaan merujuk pada komposisi dan proporsi berbagai jenis biaya yang dikeluarkan dalam operasional bisnis. Memahami struktur biaya sangat penting bagi manajemen untuk mengendalikan pengeluaran, meningkatkan efisiensi, dan memaksimalkan profitabilitas.

Menurut Tati Sulastri dan Purwanti (2024), struktur biaya terdiri dari biaya tetap dan biaya variabel, yang keduanya memiliki dampak signifikan terhadap kemampuan perusahaan dalam mencapai laba. Biaya tetap tidak berubah dengan fluktuasi volume penjualan, sementara biaya variabel berhubungan langsung dengan tingkat produksi atau penjualan. Pengelolaan yang efisien terhadap kedua jenis biaya ini, bersama dengan peningkatan volume penjualan, dapat meningkatkan profitabilitas perusahaan.

Selain itu, penelitian oleh Ahmad Nur Fuad Chalimi dan Tegar Syaifuloh (2024) menekankan pentingnya struktur organisasi, perencanaan, pelaksanaan, dan sistem pelaporan biaya dalam pengendalian biaya standar. Mereka menemukan bahwa perencanaan memiliki pengaruh dominan terhadap pengendalian biaya, menunjukkan bahwa perencanaan yang efektif dapat membantu perusahaan dalam mengendalikan biaya secara lebih efisien.

Dari pendapat tersebut, dapat disimpulkan bahwa struktur biaya pengeluaran perusahaan adalah komposisi biaya tetap dan variabel yang dikeluarkan selama operasional. Struktur ini penting dalam pengambilan keputusan manajerial karena mempengaruhi efisiensi dan profitabilitas. Biaya tetap tidak berubah meski aktivitas berfluktuasi, sedangkan biaya variabel mengikuti tingkat produksi atau penjualan. Pengelolaan yang efektif bergantung pada perencanaan, struktur organisasi, dan pelaporan biaya yang baik. Pemahaman yang tepat atas struktur biaya membantu perusahaan mengendalikan pengeluaran dan mencapai tujuan finansial secara optimal.

2.1.2. Algoritma K-Means

Menurut Maulana et al. (2024), Algoritma K-Means merupakan salah satu metode analisis data yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kesamaan atribut. Metode ini bekerja dengan cara mendefinisikan pusat kluster (centroid) secara acak, kemudian menghitung jarak antara setiap data dengan centroid tersebut. Selanjutnya, setiap data akan dikelompokkan ke dalam kluster yang memiliki centroid terdekat. Setelah proses pengelompokan selesai, centroid baru akan dihitung berdasarkan rata-rata dari seluruh data dalam masing-masing kluster.

2.1.3. Klasterisasi

Menurut Sugiarto dan Wibowo (2020), klasterisasi adalah metode pengelompokan data ke dalam beberapa kluster atau kelompok, di mana data dalam satu kluster memiliki tingkat kemiripan maksimum, sedangkan data antar kluster

memiliki kemiripan minimum. Klasterisasi merupakan proses partisi dari suatu himpunan objek data ke dalam kelompok-kelompok yang disebut klaster. Objek-objek dalam satu klaster memiliki karakteristik yang mirip satu sama lain dan berbeda dengan objek di klaster lainnya. Proses partisi ini tidak dilakukan secara manual, melainkan menggunakan algoritma klasterisasi tertentu.

2.1.4. Lembar Pertanggung Jawaban Transaksi (LPT)

Menurut Malihah, L. (2023), lembar pertanggungjawaban bertujuan untuk menyajikan informasi yang akurat dan jelas mengenai pencapaian tujuan kegiatan serta evaluasi terhadap proses yang telah dilaksanakan. Sementara itu, menurut Muthaqqin, R. (2022) Transaksi adalah sebuah kesepakatan antara penjual dan pembeli dalam menukar barang atau jasanya

Dari kedua pendapat tersebut, dapat disimpulkan bahwa Lembar Pertanggungjawaban Transaksi (LPT) merupakan laporan yang memuat rincian mengenai kegiatan, penggunaan dana, serta transaksi yang dilakukan dalam suatu proyek atau aktivitas tertentu. Laporan ini bertujuan untuk memastikan bahwa seluruh pengeluaran dan penerimaan dana tercatat secara jelas dan dapat dipertanggungjawabkan.

2.1.5. Website

Menurut Yuhefizar (2021), Website adalah metode untuk menampilkan informasi di internet dalam bentuk gambar, video, teks, suara, atau interaktif, yang menghubungkan dokumen satu dengan dokumen lainnya melalui hyperlink (hypertext) yang dapat diakses melalui browser.

2.1.6. Google Collab

Menurut Oliver (2022), Google Collab merupakan sebuah executable document yang memungkinkan pengguna untuk menyimpan, menulis, dan membagikan program secara langsung melalui Google Drive. Platform ini juga digunakan sebagai lingkungan eksekusi untuk machine learning dan pemrograman Python.

2.1.7. Python

Python merupakan bahasa pemrograman yang dikembangkan oleh Guido van Rossum dan pertama kali dirilis pada tahun 1991 di Belanda. Dalam beberapa tahun terakhir, Python telah menjadi salah satu bahasa pemrograman paling populer karena sintaksisnya yang sederhana dan mudah dipahami. Menurut Suharto (2023), salah satu keunggulan Python adalah kemudahan dalam mempelajarinya dibandingkan dengan bahasa pemrograman lainnya, sehingga sangat diminati oleh berbagai kalangan, baik pemula maupun profesional.

2.1.8. React JavaScript Library

Menurut Iswari (2021), React merupakan library JavaScript bersifat open source yang bersifat deklaratif, efisien, dan fleksibel, serta dirancang khusus untuk pengembangan antarmuka pengguna (user interface). Dengan React, pengembang dapat membangun antarmuka pengguna yang kompleks menggunakan kumpulan kode yang ringkas dan terstruktur, yang dikenal sebagai komponen. React JS banyak digunakan dalam pengembangan situs web maupun aplikasi berbasis web karena kemampuannya dalam membangun user interface yang interaktif dan responsif.

2.1.9. Mysql

Menurut Nu'man et al. (2020), MySQL sangat populer dalam pengembangan web karena kecepatan dan ukuran yang kecil, menjadikannya lebih ideal untuk website. MySQL merupakan layanan database management system yang bersifat open source, multi-user, dan multi-threaded, serta mampu menyimpan data relasional dengan menggunakan bahasa SQL. Selain itu, menurut Arief (2022), MySQL adalah jenis database server yang sangat terkenal dan banyak digunakan untuk membangun aplikasi web yang mengandalkan database sebagai sumber dan pengolahan datanya.

2.1.10. XAMPP

Menurut Aprilian (2020), XAMPP merupakan paket server web yang populer di kalangan pengembang karena menggunakan PHP dan MySQL sebagai database utama.

2.2. Penelitian Terdahulu

Berikut ini disajikan beberapa penelitian terdahulu yang relevan dengan topik yang diangkat dalam penelitian ini yaitu

1. Nama : Nur Adzilla Aulia dan Muhammad Yasin Simargolang
Judul : Implementasi Algoritma K-Means Data Mining untuk Clustering Data Transaksi Pengeluaran Toko Sepatu
Tahun : 2024
Isi : Penelitian ini bertujuan untuk mengembangkan sistem berbasis algoritma K-Means yang digunakan untuk melakukan clustering pada data transaksi pengeluaran di Toko Sepatu Ldr. Secondbrand. Sistem ini dirancang untuk mengolah data pengeluaran seperti biaya sewa gedung, gaji karyawan, air, listrik, dan pengeluaran lainnya. Data tersebut kemudian divisualisasikan dalam bentuk web agar dapat memberikan informasi yang lebih terstruktur dan mudah dipahami oleh pengguna. Tujuan utama dari penelitian ini adalah mengimplementasikan teknik data mining pada transaksi pengeluaran toko sepatu guna menganalisis pola pembelian pelanggan dan memberikan informasi yang bermanfaat bagi toko dalam pengambilan keputusan bisnis. Dengan menggunakan algoritma K-Means, penelitian ini berupaya untuk mengelompokkan data transaksi ke dalam beberapa cluster berdasarkan karakteristik tertentu. Hasil dari penelitian menunjukkan bahwa sistem yang dirancang berhasil mengelompokkan data pengeluaran ke dalam empat cluster. Cluster pertama memiliki pengeluaran terbesar sebesar Rp72.000.000 per tahun, cluster kedua sebesar Rp6.600.000, cluster ketiga sebesar Rp2.940.000, dan cluster keempat sebesar Rp105.000.000. Sistem ini terbukti efektif dalam menyelesaikan permasalahan clustering data transaksi pengeluaran toko sepatu serta mampu memberikan informasi klasifikasi.

2. Nama : Widiyanto, M. Rifqy Zakaria, Irvan
- Judul : Penerapan Data Mining untuk Klasterisasi Data Anggaran Pendapatan dan Belanja Daerah Menggunakan Algoritma K-Means
- Tahun : 2025
- Isi : Penelitian ini membahas penerapan data mining dengan menggunakan Algoritma K-Means untuk meng klasterisasi data anggaran pendapatan dan belanja daerah (APBD). Latar belakangnya adalah kebutuhan menyusun Rencana Pembangunan Daerah (RPD) selama transisi kepemimpinan daerah. Algoritma K-Means digunakan untuk menganalisis potensi pendapatan dan belanja daerah berdasarkan kluster yang terbentuk. Tujuan utama penelitian ini adalah mengelompokkan data anggaran pendapatan dan belanja daerah di Kabupaten Tolikara. Hasil pengelompokan diharapkan memberikan wawasan baru dalam analisis potensi pendapatan dan belanja daerah. Penelitian menunjukkan bahwa data dapat dikelompokkan menjadi dua kluster, dengan kluster 1 memiliki centroid 192973008 dan 167000000, serta kluster 2 dengan centroid 7000000 dan 225000000. Kluster 1 diidentifikasi sebagai potensi belanja tinggi, sedangkan kluster 2 sebagai potensi belanja rendah. Model K-Means yang diterapkan berhasil memberikan pengelompokan data yang informatif bagi Badan Pengelolaan Keuangan dan Aset Daerah (BPKAD) Kabupaten Tolikara.
3. Nama : Muhamad Rizki Nugroho, Iwansyah Edo Hendrawan, Puwantoro
- Judul : Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI
- Tahun : 2022
- Isi : Penelitian ini membahas penerapan algoritma K-Means untuk klasterisasi data obat di Rumah Sakit ASRI Purwakarta. Latar belakangnya adalah masalah pengelolaan stok obat yang sering mengalami kekurangan atau kelebihan meskipun tidak signifikan. Teknik clustering diharapkan dapat membantu pengelompokan obat berdasarkan tingkat pemakaian. Tujuan utama penelitian ini adalah menerapkan K-Means untuk pengelompokan obat berdasarkan pemakaian. Hasil penelitian menunjukkan bahwa data obat dibagi menjadi dua cluster: cluster pertama dengan pemakaian tinggi (6 obat) dan cluster kedua dengan pemakaian rendah (933 obat). Dengan demikian, algoritma K-Means dapat membantu rumah sakit dalam mengidentifikasi obat-obatan yang sering digunakan dan yang kurang diminati, sehingga pengelolaan stok obat menjadi lebih efisien.

4. Nama : Indri Setianingsih dan Irfan Ali
- Judul : Klasterisasi Pengeluaran Kas di Desa Pamengkang Menggunakan Metode K-Means
- Tahun : 2023
- Isi : Penelitian ini membahas klasterisasi pengeluaran kas di Desa Pamengkang menggunakan metode K-Means. Latar belakangnya adalah masalah pengelolaan pengeluaran kas yang sering mengalami kesalahan perhitungan dan ketidaksesuaian data, yang mengakibatkan keterlambatan laporan. Tujuan penelitian ini adalah untuk mengelompokkan data pengeluaran desa dengan metode K-Means. Data yang digunakan adalah 142 data pengeluaran kas tahun 2021, diolah menggunakan aplikasi RapidMiner. Hasil penelitian menunjukkan bahwa klasterisasi menggunakan K-Means menghasilkan nilai Davies Bouldin Index sebesar 0,012, dengan cluster 1 memiliki nilai 0,014, cluster 2 sebesar 0,017, dan cluster 3 sebesar 0,027. Hasil ini menunjukkan bahwa cluster 1 adalah kelompok pengeluaran paling signifikan. Penelitian ini menyimpulkan bahwa metode K-Means dapat memberikan wawasan dalam pengelolaan pengeluaran kas desa.
5. Nama : Agung Riyanto dan Mutaqin Akbar
- Judul : Pengelompokan Data Pengeluaran Bulanan pada Aplikasi Pencatatan Keuangan Pribadi menggunakan Metode K-Means Clustering
- Tahun : 2024
- Isi : Penelitian ini membahas pengelompokan data pengeluaran bulanan pada aplikasi pencatatan keuangan pribadi menggunakan metode K-Means Clustering. Latar belakangnya adalah pentingnya pengelolaan keuangan pribadi di era modern, di mana gaya hidup berlebihan dan kurangnya pemahaman finansial membuat aplikasi pencatatan keuangan semakin relevan. Penelitian ini bertujuan untuk mengimplementasikan pengelompokan data pengeluaran bulanan menjadi kelompok pengeluaran rendah, sedang, dan tinggi. Data yang digunakan terdiri dari 30 data training dan 25 data uji dari aplikasi Pencatatan Keuangan Pribadi (CaKeP). Proses clustering menghasilkan tiga kelompok dengan skor Silhouette Score sebesar 72%. Penentuan jumlah cluster menggunakan Elbow Method menunjukkan 3 sebagai jumlah cluster optimal. Hasil penelitian menunjukkan bahwa pengelompokan data pengeluaran dengan K-Means berhasil dilakukan dengan baik, di mana setiap sampel terklasifikasi sesuai dengan jarak yang baik terhadap kluster lainnya.

Tabel berikut menyajikan perbandingan beberapa penelitian terdahulu yang relevan dan menjadi acuan dalam penelitian ini :

Tabel 1. Tabel Perbandingan Penelitian

| No | Nama Penelitian, Tahun | Judul | Jenis Perbandingan | | | |
|----|--|--|--------------------|---------------|-----|---------|
| | | | Platform | | | Metode |
| | | | Rapid Miner | Google Collab | Web | K-Means |
| 1. | Nur Adzilla Aulia dan Muhammad Yasin Simargolang, (2024) | Implementasi Algoritma K-Means Data Mining untuk Clustering Data Transaksi Pengeluaran Toko Sepatu | | | ✓ | ✓ |
| 2. | Widianto, M. Rifqy Zakaria, Irvan, (2025) | Penerapan Data Mining untuk Klasterisasi Data Anggaran Pendapatan dan Belanja Daerah Menggunakan Algoritma K-Means | | | ✓ | ✓ |
| 3. | Muhamad Rizki Nugroho, Iwansyah Edo Hendrawan, Puwantoro, (2022) | Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI | ✓ | | | ✓ |
| 4. | Indri Setianingsih dan Irfan Ali, (2023) | Klasterisasi Pengeluaran Kas di Desa Pamengkang Menggunakan Metode K-Means | ✓ | | | ✓ |
| 5. | Agung Riyanto dan Mutaqin | Pengelompokan Data Pengeluaran | | ✓ | | ✓ |

| | | | | | | |
|----|-----------------------------------|---|--|---|--|---|
| | Akbar, (2024) | Bulanan pada Aplikasi Pencatatan Keuangan Pribadi menggunakan Metode K-Means Clustering | | | | |
| 6. | Putri Rezekika Br Ginting, (2025) | Penerapan Algoritma K-Means Pada Klasterisasi Pengeluaran Keuangan PT. Taspen (Persero) Kantor Cabang Bogor | | ✓ | | ✓ |

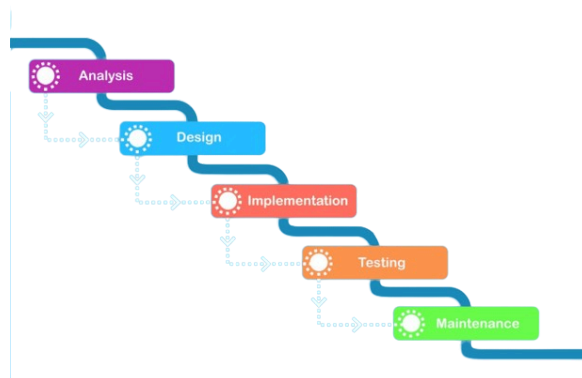
BAB III

METODE PENELITIAN

3.1. Metodologi Penelitian

Dalam penelitian ini, metode pengembangan yang digunakan adalah metode Waterfall, yang dikenal sebagai pendekatan sistem informasi yang sistematis dan berjalan secara bertahap. Menurut Wahid (2020), metode Waterfall merupakan salah satu metode yang paling umum digunakan dalam proses pengembangan perangkat lunak. Pendekatan ini merupakan bentuk awal dari Software Development Life Cycle (SDLC) dan sering disebut sebagai model tradisional atau klasik. Waterfall dikenal pula dengan istilah model linier sekuensial, karena prosesnya berlangsung secara berurutan mulai dari tahap analisis, perancangan, pengkodean, pengujian, hingga pemeliharaan..

Model ini cocok digunakan dalam penelitian ini karena proses pengembangannya membutuhkan alur kerja yang jelas dan terdokumentasi dengan baik, terlebih dengan adanya integrasi antara algoritma K-Means dan sistem informasi berbasis web bernama Laporanku. Adapun tahapan-tahapan dalam model Waterfall dapat dilihat pada Gambar 1 berikut.



Gambar 1 Metode Waterfall

3.1.1 Tahap Analisis (*Analysis*)

Tahap Analisis bertujuan untuk memahami kebutuhan pengguna dan batasan sistem melalui wawancara, observasi, dan diskusi langsung. Dari hasil analisis, ditemukan bahwa proses pembuatan Lembar Pertanggung Jawaban Transaksi (LPT) di PT. Taspen (Persero) Kantor Cabang Bogor masih dilakukan secara manual menggunakan Excel. Proses tersebut rentan terhadap kesalahan, tidak efisien, dan memerlukan waktu yang lama. Oleh karena itu, dibutuhkan sistem berbasis web yang terintegrasi dan mampu mengelola serta mengelompokkan data pengeluaran menggunakan algoritma K-Means.

3.1.2 Tahap Desain (*Design*)

Tahap desain merupakan proses perancangan sistem secara menyeluruh sebelum proses pembangunan dilakukan. Desain sistem ini mencakup Diagram Konteks, Data Flow Diagram (DFD), Entity Relationship Diagram (ERD), dan desain antarmuka pengguna (UI). Tujuannya adalah memastikan sistem memiliki alur yang jelas, mudah digunakan, serta mampu mendukung proses klasterisasi pengeluaran menggunakan algoritma K-Means secara efektif.

3.1.3 Tahap Implementasi (*Implementation*)

Tahap implementasi merupakan proses di mana sistem mulai direalisasikan berdasarkan rancangan yang telah disusun sebelumnya. Pada tahap ini, pengembangan backend dilakukan menggunakan bahasa pemrograman Python dengan framework Flask atau FastAPI. Untuk bagian frontend, digunakan React.js dan CSS guna menciptakan tampilan yang responsif dan user-friendly. Sistem penyimpanan data dibangun menggunakan MySQL sebagai basis data utama. Sementara itu, untuk proses klasterisasi data pengeluaran, digunakan platform Google Colab dengan penerapan algoritma K-Means guna memperoleh hasil klaster yang optimal.

3.1.4 Tahap Pengujian (*Testing*)

Tahapan ini dilakukan untuk memastikan bahwa sistem berfungsi dengan baik dan sesuai dengan kebutuhan pengguna. Pengujian dilaksanakan dalam beberapa bagian berikut:

1. Pengujian Struktural

Pengujian ini berfokus pada arsitektur sistem secara internal, mencakup struktur kode program dan interaksi antar komponen sistem, untuk memastikan bahwa sistem dibangun dengan struktur yang baik dan efisien.

2. Pengujian Fungsional

Pada tahap ini dilakukan pengujian terhadap seluruh fitur dan fungsi sistem untuk memastikan bahwa setiap bagian berjalan sesuai dengan kebutuhan dan spesifikasi yang telah ditetapkan, termasuk proses input, pemrosesan data klasterisasi, hingga output laporan.

3. Pengujian Validasi

Pengujian dilakukan dengan melibatkan pengguna secara langsung untuk menilai sejauh mana sistem sesuai dengan kebutuhan pengguna, mudah digunakan, serta mampu berfungsi secara optimal.

3.1.5 Tahap Pemeliharaan (*Maintenance*)

Tahap pemeliharaan dilakukan setelah sistem digunakan secara resmi oleh pengguna. Pada tahap ini, sistem dipantau dan dievaluasi secara berkala untuk memastikan kinerjanya tetap optimal serta sesuai dengan kebutuhan yang terus berkembang. Oleh karena itu, kegiatan pemeliharaan mencakup perbaikan bug serta penyempurnaan fitur berdasarkan masukan dari pengguna.

3.2. Alur Pengolahan Data

Pengolahan data merupakan tahapan penting dalam penelitian ini, yang bertujuan untuk menyiapkan data agar dapat digunakan dalam proses klasterisasi menggunakan algoritma K-Means, serta diintegrasikan ke dalam sistem informasi yang dikembangkan, yaitu Laporanku. Alur pengolahan dilakukan melalui beberapa tahapan utama di platform Google Collab.

Proses dimulai dengan mengimpor dataset menggunakan pustaka pandas. Setelah data berhasil dimuat, dilakukan pemeriksaan struktur dataset untuk memahami kolom-kolom yang tersedia, jenis data, serta jumlah entri yang ada seperti tanggal, akun/deskripsi, deskripsi kegiatan, dan realisasi (rp).

Langkah selanjutnya adalah pembersihan data (data cleaning) yang bertujuan untuk menghilangkan duplikasi penuh dan nilai-nilai yang tidak relevan. Baris dengan nilai realisasi (rp) sebesar nol dihapus karena tidak berkontribusi terhadap

analisis pengeluaran. Setelah itu, dilakukan transformasi format tanggal, di mana data pada kolom tanggal diubah ke format datetime untuk mempermudah analisis berdasarkan periode waktu tertentu.

Kemudian dilakukan normalisasi data, yaitu proses penyetaraan skala nilai pada kolom realisasi (rp) agar fitur ini tidak mendominasi saat dianalisis menggunakan algoritma k-means. Setelah data dinormalisasi, dilanjutkan ke tahap klusterisasi menggunakan algoritma k-means. Penentuan jumlah cluster dilakukan dengan metode elbow untuk memperoleh hasil yang optimal. Selanjutnya, data dikelompokkan ke dalam beberapa kategori pengeluaran seperti biasa, sedang, dan boros berdasarkan kemiripan nilai realisasi.

Pada tahap akhir, hasil klusterisasi disimpan kembali dalam format .csv atau dimasukkan langsung ke dalam basis data MySQL. Data ini kemudian diintegrasikan ke dalam sistem Laporanku untuk ditampilkan dalam laporan bulanan, serta menjadi dasar bagi fitur notifikasi kategori pengeluaran. Dengan mengikuti alur pengolahan data ini, sistem mampu menyajikan informasi pengeluaran yang lebih terstruktur, efisien, dan bermakna bagi pengguna.

3.3. Sumber Data

Data yang digunakan dalam penelitian ini berasal dari laporan pengeluaran bulanan PT. Taspen (Persero) Kantor Cabang Bogor, yang mencakup periode waktu dari bulan Mei 2020 hingga Oktober 2023. Data tersebut diperoleh dalam format .csv dan berisi informasi terkait aktivitas pengeluaran Lembar Pertanggungjawaban Transaksi (LPT) , yang terdiri dari beberapa kolom utama seperti Tanggal, Akun/Deskripsi, Deskripsi Kegiatan, Dan Realisasi (Rp).

Data ini dipilih karena mencerminkan kondisi aktual pengeluaran perusahaan dalam periode yang cukup panjang, sehingga dianggap representatif untuk dianalisis menggunakan algoritma K-Means. Dengan karakteristik tersebut, data ini sangat relevan dalam proses klusterisasi pengeluaran serta pengembangan sistem informasi Laporanku yang bertujuan menyajikan Lembar Pertanggungjawaban Transaksi (LPT) berbasis kategori pengeluaran.