

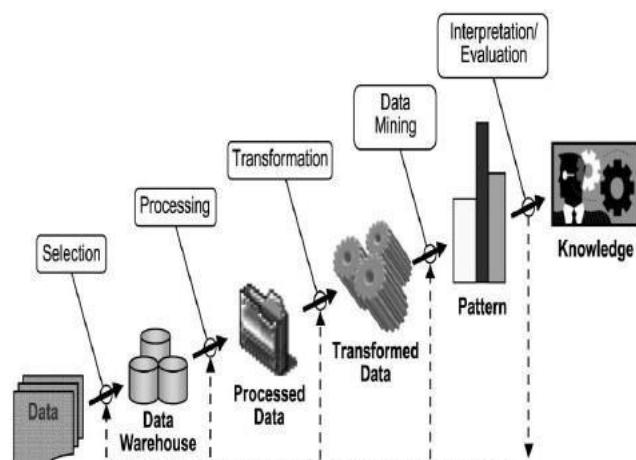
BAB II KERANGKA TEORITIS

A. Landasan Teori

1. Data mining

Dalam proses menemukan informasi yang menarik terdapat sebuah data yang akan diolah dengan menggunakan metode tertentu. Menurut (Muhammad Arhami, 2020, pp. 2-5) *data mining* merupakan proses untuk menganalisa sebuah pola pada data menjadi sebuah informasi yang berguna sehingga dapat membantu dalam membuat keputusan yang lebih baik. *Data mining* memiliki tujuan untuk menemukan sebuah pola yang memiliki arti dan maksud tertentu yang dapat digunakan untuk menggali dan menganalisis dari sebagian besar data yang ada.

Istilah *data mining* digunakan sebagai proses penguraian penemuan informasi dalam sebuah data. (Putra, et al, 2023, p. 30) mengutarakan bahwa *data mining* memiliki hubungan di berbagai bidang keilmuan yang dapat membantu untuk mengidentifikasi pola ataupun hubungan tersembunyi dari sebuah data yang dapat mengungkapkan sebuah informasi berharga dari data. Proses yang terjadi seperti pengambilan informasi, analisa pola dan ringkasan pengetahuan yang memiliki pemahaman mengenai data serta menggiring pengukuran secara konstruktif dari area yang terlibat.



Gambar 2. 1 Tahapan Proses KDD

Data mining termasuk kedalam salah satu rangkaian *Knowledge Discovery in Database* (KDD). Menurut (Amna et al, 2023, p. 10) *Knowledge Discovery in Database* (KDD) merupakan tahap penggalian dan menganalisa guna mendapatkan suatu informasi yang berguna dari sejumlah data. *Data mining* dapat digunakan untuk mendapatkan informasi pada bisnis dan analisis data. Gambar 2. 1 menunjukkan proses tahapan *data mining* menurut

Fayyad, Shapiro, dan Smith pada tahun 1996 dalam (Yessi Jollyta, Dedy Hartama Ramdhan, 2020, pp. 48-49) sebagai berikut:

- (1) *selection*, tahapan awal dengan memilih atau menyeleksi beberapa data yang dibutuhkan dari sekumpulan data untuk diseleksi kembali pada tahap selanjutnya;
- (2) *preprocessing*, tahap ini dilakukan pembersihan data pada field yang dianggap tidak memadai untuk menjamin format data konsisten;
- (3) *transformation*, pada tahap ini data akan disesuaikan dengan teknik *data mining* yang digunakan agar kualitas data tepat dan sesuai untuk digunakan pada tahap selanjutnya, pada penelitian ini normalisasi data yang digunakan dalam penelitian ini. Tujuan normalisasi adalah untuk mengubah nilai suatu variabel untuk memiliki wilayah yang seragam atau skala seragam. Ini mencegah variabel dalam kisaran nilai yang menyesuaikan perhitungan atau analisis. Persamaan untuk normalisasi data Min-Max dapat ditemukan pada Gambar 2.2 di bawah ini: (Rahayu et al, 2018, p. 52).

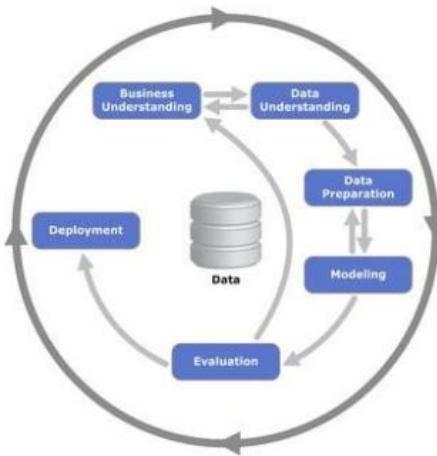
Rumus Normalisasi Min-Max:

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

- (4) *data mining*, tahap ini menggunakan sebuah teknik atau metode khusus untuk menganalisis data guna mendapatkan hasil yang berisi informasi penting yang tersembunyi dalam data;
- (5) *interpretation & evaluation*, tahap dimana menginterpretasikan pola menjadi suatu pengetahuan yang dapat digunakan untuk membantu mendukung pengambilan keputusan.

2. Cross Industry Standard Process for Data Mining (CRISP-DM)

Dalam upaya memahami dan memecahkan permasalahan yang kompleks dalam bisnis maupun penelitian, Cross Industry Standard Process for Data Mining hadir sebagai kerangka kerja yang terstruktur. Sebagaimana dijelaskan oleh (Kusrini, 2009, p. 9), CRISP-DM menyediakan sebuah proses standar yang sistematis dalam melakukan penambangan data.



Gambar 2. 2 Alur CRISP-DM

Sumber : (Kusrini, 2009)

Gambar 2. 2 menunjukkan proses data *mining* pada CRISP-DM atau Cross-Industry Standard Process for Data Mining yang digunakan sebagai strategi dalam pemecahan masalah pada bisnis maupun penelitian dengan menyediakan sebuah proses standar dari data *mining*. Menurut (Kusrini et al, 2009, p. 8) data *mining* memiliki enam fase siklus hidup dalam CRISP-DM, dimana pada fase siklus hidup tersebut memiliki sifat adaptif secara berurutan yang digambarkan dengan panah untuk menunjukkan hubungan penting pada fase siklus hidup. Adapun fase CRISP-DM menurut Laroze dalam (Kusrini et al, 2009, p. 9) sebagai berikut:

- (1) fase pemahaman bisnis (*Business Understanding Phase*); fase ini terdiri dari psenentuan tujuan proyek dan kebutuhan secara spesifik secara keseluruhan, lalu menerjemahkan tujuan dan batasan menjadi sebuah acuan dari permasalahan dan pada akhirnya menyiapkan strategi awal untuk mencapai awal tujuan;
- (2) fase pemahaman data (*Data Understanding Phase*); fase ini dimulai dari mengumpulkan data, lalu melakukan analisis pada data dan mengevaluasi kualitas pada data;
- (3) fase persiapan data (*Data Preparation Phase*); fase ini dimulai dari menyiapkan data, memilih variabel yang akan digunakan dan mengubahnya jika ada beberapa variabel yang diperlukan;
- (4) fase pemodelan (*Modeling Phase*); fase ini menerapkan teknik data *mining* yang sesuai dengan permasalahan yang diambil;
- (5) fase evaluasi (*Evaluation Phase*); fase ini dilakukan evaluasi kembali model yang dipilih agar permasalahan yang terdapat terpecahkan dengan baik;

(6) fase penyebaran (*Deployment Phase*); fase ini menyebarkan model yang telah dihasilkan dari teknik data *mining* yang berupa laporan.

Data *mining* memiliki beberapa metode diantaranya klasifikasi, *clustering*, asosiasi, estimasi, prediksi dan regresi. Pada penelitian ini menggunakan metode *clustering*. (Arhami and Nasir, 2020, p. 149) mengutarakan bahwa *clustering* merupakan data atau nilai yang belum memiliki label pada kelasnya perlu diprediksi ke dalam sebuah kelas untuk mengetahui objek tersebut akan dimasukan pada kelas yang sesuai berdasarkan kesamaan pola atau karakteristik pada kelompoknya. Salah satu metode pada data mining yaitu pengelompokan data ke dalam kelompok berdasarkan dengan karakteristik tertentu.

3. *K-Means Clustering*

1) *Clustering*

Clustering yaitu pengelompokan suatu set objek yang mirip ke dalam kelas atau kelompok. Data dikelompokkan ke dalam kelas-kelas yang memiliki kemiripan tinggi satu sama lain tetapi memiliki perbedaan dengan cluster lainnya (Bhatia, 2019, p. 206). Objek-objek dikelompokkan berdasarkan prinsip memaksimalkan kesamaan dalam satu kelas (*intra-cluster similarity*) dan meminimalkan kesamaan antar kelas (*inter-cluster similarity*) (Reddy et al, 2014, p. 4). Analisis cluster dapat digunakan sebagai alat eksplorasi mandiri untuk memahami distribusi data, mengamati karakteristik masing-masing cluster, serta menjadi dasar untuk analisis lebih lanjut terhadap data (Han et al, 2012, p. 443).

Menurut (Putra et al, 2023, pp. 86-87), clustering termasuk dalam unsupervised learning, yaitu metode pembelajaran tanpa pengawasan di mana data tidak bergantung pada label tertentu untuk bisa diproses. Clustering memungkinkan ditemukannya pola atau struktur alami dalam data, dengan cara mengelompokkan data berdasarkan karakteristik tertentu yang serupa. Pada penelitian ini, untuk memetakan aset perangkat perusahaan, digunakan metode clustering dengan pendekatan partisi clustering, yakni data dibagi secara langsung menjadi beberapa kelompok berdasarkan kedekatan nilai-nilai atributnya.

2) *Algoritma K-Means*

Dalam penelitian ini, pemetaan aset perangkat perusahaan diperlukannya sebuah algoritma. Menurut (Santosa, 2007, p. 42) K-Means merupakan sebuah teknik klastering yang umum dikenal dan paling sederhana, dengan K-Means dapat mengelompokkan suatu objek kedalam

sebuah klaster. (Eko Prasetyo, 2014, p. 189) mengutarakan bahwa algoritma K-Means sederhana untuk diimplementasikan, mudah dipahami, umum dalam penggunaannya dan relatif cepat. Dengan algoritma K-Means dapat dilakukan pembagian data ke dalam sebuah kelompok yang memiliki karakteristik yang serupa dikelompokkan ke dalam satu set yang serupa, sebaliknya jika memiliki karakteristik berbeda, maka akan dikelompokkan ke dalam set yang berbeda pula, hal ini dilakukan untuk meminimalkan nilai variatif pada suatu kelompok (Putra et al, 2023, pp. 87-88).

Menurut (Arhami and Nasir, 2020, pp. 148-149) berikut ini langkah-langkah dasar pada algoritma K-means yaitu:

- (1) Menentukan nilai K klaster sesuai dengan yang diinginkan;
- (2) pilih titik-titik atau sampel yang menjadi anggota klaster secara acak;
- (3) tentukan nilai centroid atau titik tengah dari klaster tersebut dengan rumus;

$$M_k = \left(\frac{1}{n_k} \right) \sum_{i=1}^{n_k} X_{ik}$$

- (4) hitung *square error* untuk setiap klaster C_k yang merupakan jumlah kuadrat dari jarak Euclidean antara tiap sampel dalam C_k dan titik tengahnya (*centroid*), error dikenal juga dengan nama *within cluster variation (WCV)*, yaitu;

$$e_k^2 = \sum_{i=1}^{n_k} (X_{ik} - M_k)^2$$

- (5) selanjutnya jumlah dari keseluruhan error dari k -cluster juga dihitung dengan rumus;

$$E_k^2 = \sum_{k=1}^K e_k^2$$

- (6) kelompokkan kembali semua sampel berdasarkan jarak minimum dari masing-masing pusat M_1, M_2, \dots, M_k sehingga diperoleh distribusi baru dari sampel sesuai kelasternya, untuk memperoleh distribusi baru dari sampel baru tersebut dapat dilakukan dengan menghitung jarak masing-masing titik pusat dengan keseluruhan sampel $d(M_1, x_1) \dots d(M_k, x_k)$; perhitungan jarak dari masing-masing titik tersebut dapat menggunakan beberapa metode, contohnya *Euclidean Distance*, dimana jarak antara dua titik dalam satu, dua, tiga ataupun dimensi n dapat dihitung sebagai berikut;

$$d(p, q) = \sqrt{((p_1 - q_1)^2 + ((p_2 - q_2)^2 + \dots + ((p_n - q_n)^2))}$$

- (7) tuliskan hasil anggota klaster baru sesuai dengan hasil yang diperoleh pada langkah ke-5 dan ulangi langkah ke-3 sampai beberapa iterasi yang nantinya akan ditemukan nilai total *square error* turun secara signifikan.

Menurut (Arhami and Nasir, 2020, pp. 149-152), berikut ini perhitungan K-means yang dapat dipaparkan dengan data contoh sebagai berikut;

x_k	p	q
x ₁	0	2
x ₂	0	0
x ₃	1.5	0
x ₄	5	0
x ₅	5	2

(a) misal pada kasus seperti ini, nilai *K* yang diambil yaitu *K* = 2;

(b) klaster diambil secara acak yaitu sebagai berikut;

$$C_1 = \{x_1, x_2, x_4\}$$

$$C_2 = \{x_3, x_5\}$$

(c) pada langkah ke-2 dapat ditentukan titik tengah dari kedua klaster tersebut, yaitu;

$$M_1 = \left\{ \left(\frac{0+0+5}{3} \right), \left(\frac{2+0+0}{3} \right) \right\} = \{1.66, 0.66\}$$

$$M_2 = \left\{ \left(\frac{1.5+5}{2} \right), \left(\frac{0+2}{2} \right) \right\} = \{2.25, 1.00\}$$

(d) *square error* diperoleh;

$$e_1^2 = [(0 - 1.66)^2 + (2 - 1.66)^2] + [(0 - 1.66)^2 + (0 - 1.66)^2] \\ + [(5 - 1.66)^2 + (0 - 1.66)^2] = 19.36$$

$$e_2^2 = [(1.5 - 3.25)^2 + (0 - 1)^2] + [(5 - 3.25)^2 + (2 - 1)^2] = 8.12$$

Sehingga diperoleh *square error* yaitu;

$$E^2 = e_1^2 + e_2^2 = 19.36 + 8.12 = 27.48$$

(e) restribusi sampel ke dalam kelas yang sesuai dengan hasil perhitungan jarak yang dimiliki oleh tiap sampel, dimana jarak minimum dari titik pusat klaster yang ada akan menjadi klaster baru bagi sampel tersebut; berikut hasil yang diperoleh;

$$d(M_1, x_1) = \sqrt{(1.66 - 0)^2 + (0.66 - 2)^2} = 2.14$$

$$d(M_2, x_1) = \sqrt{(3.25 - 0)^2 + (1 - 2)^2} = 3.40$$

$$d(M_1, x_2) = \sqrt{((1.66 - 0)^2 + (0.66 - 0)^2)} = 1.79$$

$$d(M_2, x_2) = \sqrt{((3.25 - 0)^2 + (1 - 2)^2)} = 3.40$$

$$d(M_1, x_3) = \sqrt{((1.66 - 1.5)^2 + (0.66 - 0)^2)} = 0.83$$

$$d(M_2, x_3) = \sqrt{((3.25 - 1.5)^2 + (1 - 0)^2)} = 2.01$$

$$d(M_1, x_4) = \sqrt{((1.66 - 5)^2 + (0.66 - 0)^2)} = 3.41$$

$$d(M_2, x_4) = \sqrt{((3.25 - 5)^2 + (1 - 0)^2)} = 2.01$$

$$d(M_1, x_5) = \sqrt{((1.66 - 5)^2 + (0.66 - 2)^2)} = 3.60$$

$$d(M_2, x_5) = \sqrt{((3.25 - 5)^2 + (1 - 2)^2)} = 2.01$$

berikut hasil perhitungan yang telah dilakukan dan jarak minimum dari titik tengah klaster ke sampel menjadi kelas sampel;

<i>X_k</i>	<i>d(M₂, x_{k2})</i>	<i>d(M₂, x_{k2})</i>	<i>Klaster</i>
<i>X₁</i>	2.14	3.40	<i>C₁</i>
<i>X₂</i>	1.79	3.40	<i>C₁</i>
<i>X₃</i>	0.83	2.01	<i>C₁</i>
<i>X₄</i>	3.41	2.01	<i>C₂</i>
<i>X₅</i>	3.60	2.01	<i>C₂</i>

sehingga klaster baru dapat dituliskan $C_1 = \{x_1, x_2, x_3\}$, $C_2 = \{x_4, x_5\}$, pada kedua klaster dihitung kembali titik tengah dan total *square error*-nya untuk memastikan perlu dilanjutkan atau tidak ke iterasi berikutnya; titik tengah dari kedua klaster dapat dihitung sebagai berikut:

$$M_1 = \left\{ \left(\frac{0+0+1.5}{3} \right), \left(\frac{2+0+0}{3} \right) \right\} = \{0.5, 0.66\}$$

$$M_2 = \left\{ \left(\frac{5+5}{2} \right), \left(\frac{0+2}{2} \right) \right\} = \{5, 1.00\}$$

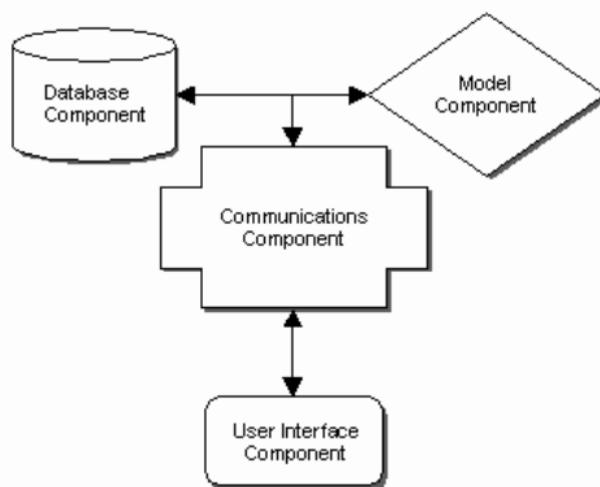
dan *square error*-nya yaitu; $e_1^2 = 4.17$ dan $e_2^2 = 2$ sehingga total yang diperoleh adalah;

$$E^2 = e_1^2 + e_2^2 = 4.17 + 2 = 6.17$$

jika dilihat dari total *square error*-nya yang turun begitu signifikan dari 27.48 ke 6.17, maka untuk kasus ini pada iterasi pertama sudah cukup untuk mengambil kesimpulan bahwa klaster yang dihasilkan adalah $C_1 = \{x_1, x_2, x_3\}$, $C_2 = \{x_4, x_5\}$ karena jika analisis dilanjutkan berdasarkan dengan titik tengah klaster baru maka klaster yang akan diperoleh juga akan sama seperti klaster yang dihasilkan pada iterasi pertama.

4. Sistem Pendukung Keputusan

Menurut (Power, 2002, p.12), Sistem Pengambilan Keputusan (*Decision Support Systems/DSS*) adalah kombinasi dari teknologi, model analisis, dan basis data yang bertujuan untuk membantu pengambilan keputusan dalam situasi yang kompleks. Menurut (Kusrini, 2007, p. 7) Sistem Pendukung Keputusan (SPK) atau *Decision Support System* (DSS) merupakan sebuah sistem yang mampu memberikan kemampuan pemecahan masalah maupun kemampuan pengkomunikasian untuk masalah dengan kondisi terstruktur maupun tidak terstruktur, sistem merupakan kumpulan elemen yang saling berkaitan yang bertanggung jawab memproses masukan (input) sehingga menghasilkan keluaran (output). Sistem pendukung keputusan juga adalah sistem yang digunakan dalam pengambilan keputusan didefinisikan suatu sistem informasi yang membantu mengidentifikasi kesempatan, membuat keputusan atau menyediakan informasi untuk membantu membuat keputusan (Whitten, Jeffery L, 2004, p.38).



Gambar 2. 3 Komponen DSS

Sumber : (Power, 2002, p. 88)

Menurut (Power, 2002, pp. 12-15), komponen-komponen utama dari sistem pendukung keputusan (SPK) meliputi:

- (1) Database: Database merupakan gudang data yang menyimpan informasi yang diperlukan dalam pengambilan keputusan. Database harus dirancang dengan baik dan terstruktur agar dapat digunakan secara efektif dalam SPK.
- (2) Model: Model dalam SPK digunakan untuk menganalisis data dan membantu dalam pengambilan keputusan. Model dapat berupa model

matematis, model statistik, atau model simulasi, tergantung pada jenis keputusan yang dihadapi.

- (3) Antarmuka Pengguna: Antarmuka pengguna dalam SPK berfungsi untuk menghubungkan pengguna dengan sistem. Antarmuka yang mudah digunakan, intuitif, dan menyediakan informasi yang relevan sangat penting untuk memastikan penggunaan yang sukses dari SPK.
- (4) Indikator Performa: Indikator performa digunakan untuk mengukur dan mengevaluasi hasil dari keputusan yang dibuat. Indikator performa ini membantu pengguna melihat efektivitas keputusan dan memperbaikinya jika perlu.
- (5) Pengolah Keputusan: Pengolah keputusan dalam SPK bertanggung jawab untuk menjalankan prosedur yang diperlukan untuk menghasilkan rekomendasi keputusan. Pengolah keputusan ini dapat menggunakan logika, algoritma optimasi, atau teknik pengambilan keputusan lainnya.

5. Aset Perangkat diperusahaan

Aset perangkat di perusahaan merupakan komponen penting dalam operasional bisnis, mencakup berbagai jenis perangkat keras dan lunak yang digunakan untuk mendukung aktivitas perusahaan. Menurut (Andry Alamsyah, 2015, p. 45), aset teknologi informasi dalam suatu perusahaan dapat dikategorikan sebagai sumber daya strategis yang berkontribusi terhadap efisiensi, efektivitas, dan keunggulan kompetitif.

Perusahaan mengelola aset perangkat dengan tujuan meningkatkan produktivitas dan mengoptimalkan biaya pemeliharaan. (Kadir dan Triwahyuni, 2013, p. 78) menjelaskan bahwa aset teknologi meliputi perangkat komputer, server, jaringan, perangkat lunak, serta sistem keamanan yang harus dipelihara agar tetap berfungsi dengan baik. Pengelolaan aset yang buruk dapat menyebabkan peningkatan biaya operasional, penurunan kinerja, serta risiko keamanan data.

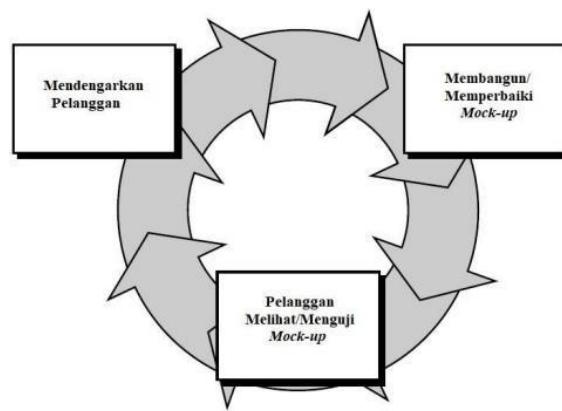
Selain itu, keberlanjutan penggunaan aset juga dipengaruhi oleh faktor usia perangkat, biaya perawatan, serta kebijakan perusahaan terkait siklus hidup perangkat. Oleh karena itu, perusahaan perlu menerapkan strategi manajemen aset yang efektif untuk memastikan bahwa aset yang digunakan tetap efisien dan sesuai dengan kebutuhan operasional.

6. Pengembangan *System Development Life Cycle*

Dalam hal ini penelitian yang dilakukan menggunakan sebuah model pengembangan agar proses yang terjadi dalam penelitian ini lebih terstruktur. Menurut (A.S. and Shalahuddin, 2013, p. 26) *System Development Life Cycle*

(SDLC) merupakan suatu pengembangan pada sistem yang sudah ada sebelumnya dengan memiliki tahapan-tahapan yang harus dilalui agar terciptanya sebuah sistem yang berkualitas. Pada *System Development Life Cycle* (SDLC) terdapat beberapa metode, salah satu metode yang digunakan pada penelitian ini yaitu metode *prototyping*. Menurut (Khadir, 2014, p. 29) *prototyping* merupakan metode dalam pengembangan sebuah sistem yang dilakukannya pendekatan untuk menghasilkan spesifikasi yang terperinci sebelum pengguna melakukan evaluasi pada sistem tersebut.

Pada penelitian ini, *prototyping* digunakan untuk menjabarkan kebutuhan pengguna secara rinci. Menurut (A.S. and Shalahuddin, 2013, pp. 30-31) model *prototyping* dibutuhkan untuk memberi kejelasan mengenai spesifikasi kebutuhan yang diinginkan pada sebuah sistem dengan awal mula apa saja yang dibutuhkan oleh user mengenai sistem yang dibutuhkan, *prototyping* menyediakan sebuah gambaran tampilan dari alur kerja sebuah sistem yang pada nantinya akan di evaluasi oleh pengguna hingga ditemukan spesifikasi yang sesuai dengan keinginan pengguna. Berikut ini model *prototyping* menurut (A.S. and Shalahuddin, 2013, p. 32) yang terdiri dari 3 tahapan ditunjukkan pada gambar 2.4, yaitu;



Gambar 2. 4 Pola Melingkar dari Siklus Hidup Sistem

Sumber: (A.S. and Shalahuddin, 2013, p. 32)

- (1) mendengarkan pelanggan untuk melakukan pengumpulan kebutuhan dari sebuah sistem bagaimana sistem yang sedang berjalan dengan melakukan analisis kebutuhan dari pengguna;
- (2) dilakukannya perancangan dan membangun *prototyping* sesuai dengan kebutuhan pengguna yang telah dijabarkan sebelumnya;
- (3) dilakukannya pengujian *prototyping* oleh pengguna kemudian pengguna memberikan evaluasi kekurangan dan kebutuhan apa saja yang dibutuhkan oleh pengguna sehingga pada saat pengujian masih terdapat

kekurangan dan kebutuhan maka kembali kepada tahap mendengarkan pelanggan untuk memperbaiki *prototyping* yang sudah tersedia; *prototyping* sebuah sistem yang dapat menyediakan dan mendemonstrasikan fungsi-fungsi yang terdapat serta mungkin dapat dilakukan pengujian desain sistem tersebut dapat dikatakan sebagai *mock-up*.

7. Metode *Prototype*

Metode pengembangan perangkat lunak yang dikenal sebagai *prototyping* melibatkan pembuatan model fisik kerja sistem yang berperan sebagai versi awal dari sistem; prototype akan diubah atau diperluas pada bagian-bagian agar sejalan dengan rencana dan analisis yang dikerjakan oleh pengembang, sampai pengujian dilakukan secara bersama dengan pengembang (Ogedebe and Jacob, 2012, p. 124). Sedangkan menurut menurut (Pressman, 2012, p. 50) suatu metode dalam pengembangan perangkat lunak yang memungkinkan pengguna dan pengembang untuk mengevaluasi kemungkinan desain serta merespons secara cepat terhadap umpan balik.

Pengembangan *prototype* mempunyai beberapa tahapan menurut (Pressman, 2012, pp. 77-78) yaitu;

- (a). Identifikasi Kebutuhan Dasar: Tahap ini melibatkan pengumpulan kebutuhan umum dan fungsi dasar yang perlu diimplementasikan dalam sistem. Kebutuhan yang rumit dan tidak dipahami dengan baik menjadi fokus utama.
- (b). Pembuatan Prototipe Awal: Berdasarkan kebutuhan yang telah diidentifikasi, sebuah prototipe awal dibuat. Prototipe ini biasanya berfokus pada aspek-aspek sistem yang paling sulit dimengerti atau paling sulit diprediksi hasilnya.
- (c). Evaluasi Pengguna: Prototipe yang dibuat kemudian dievaluasi oleh pengguna. Pengguna berinteraksi dengan prototipe dan memberikan umpan balik mengenai fungsionalitas dan kemungkinan peningkatan.
- (d). Perbaikan Prototipe: Berdasarkan umpan balik dari pengguna, perubahan dan peningkatan dilakukan pada prototipe. Hal ini dapat melibatkan penambahan fitur baru, modifikasi fitur yang ada, atau bahkan penghapusan fitur yang tidak perlu.
- (e). Iterasi: Tahap evaluasi dan perbaikan ini diulangi berkali-kali sampai semua kebutuhan pengguna terpenuhi dan semua keraguan tentang fungsi dan kinerja sistem telah dihilangkan.

(f). Implementasi Sistem: Setelah prototipe mencapai tingkat kepuasan yang memadai, perangkat lunak final dikembangkan dengan menggunakan pengetahuan dan pemahaman yang diperoleh dari proses *prototyping*.

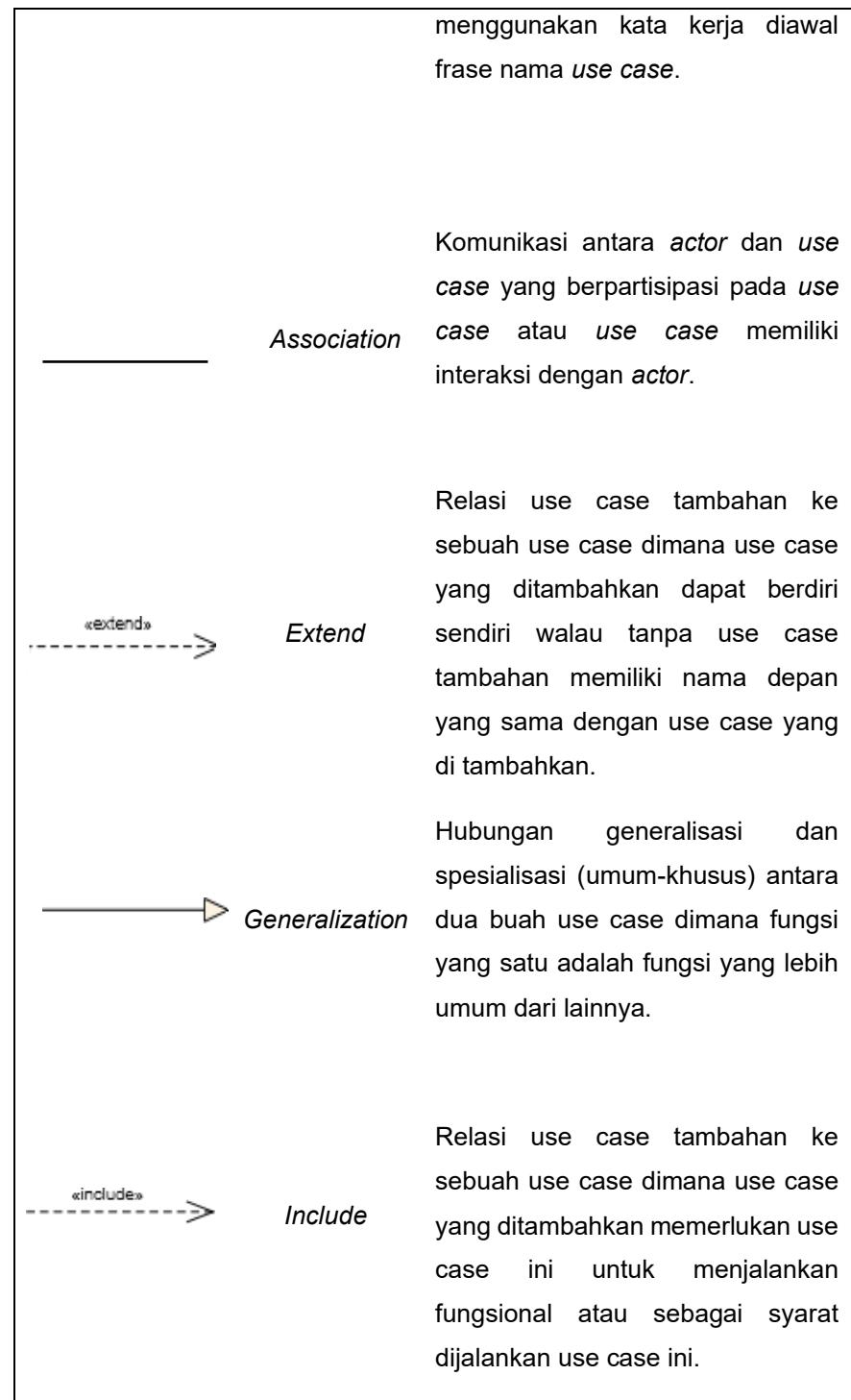
7. UML (*Unified Modeling Language*)

(Nugroho, 2010, p. 6) UML atau *Unified Modeling Language* merujuk pada suatu bahasa pemodelan yang dirancang khusus untuk sistem atau perangkat lunak yang mengadopsi paradigma “berorientasi objek”; pemodelan atau proses pemodelan sebanarnya bertujuan untuk menyederhanakan berbagai tantangan yang kompleks, sehingga pada akhirnya dapat lebih mudah diajarkan dan dipahami. Dari penelitian UML berarti sebuah symbol yang memiliki peran dalam perancangan sistem untuk menggambarkan hubungan antar aktivitas atau kegiatan yang akan terjadi pada sistem. UML memiliki beberapa jenis, setiap jenis tentunya memiliki tujuan yang berbeda tergantung dari apakah jenis diagram tersebut dirancang sebelum implementasi atau bahkan setelahnya. Berikut jenis-jenis UML:

(a) *Usecase Diagram*

Dikemukakan menurut penelitian dari (Murad dkk, 2013, pp. 44-58) diagram *usecase* adalah representasi visual yang berupa gambar kondisi yang menunjukkan kelompok *usecase* dan *actor* (sejelas kelas khusus). Berikut ini merupakan notasi dan symbol dari *usecase* diagram :

Simbol	Nama	Keterangan
	Actor	Orang proses, atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat di luar sistem itu sendiri, jadi walaupun simbol dari <i>actor</i> adalah gambar orang, biasanya dinyatakan menggunakan kata benda di awal frase nama <i>actor</i>
	Use Case	Fungsionalitas yang disediakan system sebagai unit-unit yang saling bertukar pesan antara unit aktif, biasanya dinyatakan dengan



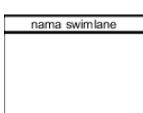
Gambar 2. 5 Simbol Usecase Diagram

Sumber : (Murad dkk, 2013)

(b) Activity Diagram

Menurut (Hasanah, 2020, p. 80) Activity Diagram digunakan untuk memvisualisasikan aliran aktivitas dalam suatu sistem, mulai dari titik awal, keputusan, hingga titik akhir, termasuk proses yang dapat berlangsung paralel. Diagram ini menyerupai flowchart dan biasanya dibuat berdasarkan use case untuk menggambarkan proses bisnis serta urutan aktivitas secara jelas. Beberapa simbol yang digunakan dalam Activity Diagram dijelaskan pada Gambar 2.6.

Simbol	Keterangan	Deskripsi
	Status awal	Initial State pada Activity Diagram merupakan titik awal atau kondisi permulaan dari suatu aktivitas dalam sistem.
	Aktivitas	Aktivitas merupakan suatu tindakan atau kegiatan yang dijalankan oleh sistem, yang umumnya dinyatakan menggunakan kata kerja.
	Percabangan / <i>decision</i>	Decision menggambarkan percabangan dalam aliran aktivitas, yaitu saat terdapat lebih dari satu jalur aktivitas yang dapat dipilih.
	Penggabungan / <i>join</i>	Penggabungan menunjukkan Penggabungan dua atau lebih aliran aktivitas menjadi satu jalur tunggal.
	Status akhir	Status akhir pada Activity Diagram menunjukkan titik akhir atau kondisi akhir dari suatu aktivitas yang dijalankan oleh sistem.

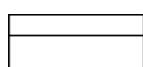
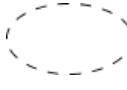
Simbol	Keterangan	Deskripsi
 Atau 	Swimlane	Swimlane digunakan untuk memisahkan entitas bisnis atau aktor yang memiliki tanggung jawab atas pelaksanaan aktivitas tertentu dalam sistem.

Gambar 2. 6 Activity Diagram

Sumber : (Hasanah, 2020, p. 82)

(c) Class Diagram

Berdasarkan sumber dari (Munawar, 2018, pp. 50-51) *Class Diagram* merupakan kumpulan entitas yang memiliki karakteristik serupa, setiap entitas memiliki kondisi saat ini (state) dan tingkah laku (behavior), kondisi saat ini entitas tercermin dalam atribut atau sifat yang dimilikinya. Berikut ini merupakan notasi dalam symbol class diagram:

Simbol	Nama	Keterangan
<hr/>	<i>Generalization</i>	Hubungan dimana objek anak (<i>descendent</i>) berbagi perilaku dan struktur data dari objek yang ada di atasnya objek induk (<i>ancestor</i>).
	<i>Nary Association</i>	Upaya untuk menghindari asosiasi dengan lebih dari 2 objek.
	<i>Class</i>	Himpunan dari objek-objek yang berbagi atribut serta operasi yang sama.
	<i>Collaboration</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem yang menghasilkan

		suatu hasil yang terukur bagi suatu actor
<-----	<i>Realization</i>	Operasi yang benar-benar dilakukan oleh suatu objek.
----->	<i>Dependency</i>	Hubungan dimana perubahan yang terjadi pada suatu elemen mandiri (<i>independent</i>) akan mempengaruhi elemen yang bergantung padanya elemen yang tidak mandiri.
_____	<i>Association</i>	Apa yang menghubungkan antara objek satu dengan objek lainnya

Gambar 2. 7 Simbol Class Diagram

Sumber : (Munawar, 2018, pp. 50-51)

(d) Sequence Diagram

Diagram Sequence mengilustrasikan cara sistem menanggapi tindakan pengguna, *diagram sequence* yang dibuat adalah yang terkait secara langsung dengan aktivitas pokok dari sistem informasi yang berkaitan dengan anggaran pendapatan dan belanja desa berbasis objek (Irmayani & Susyatih, 2017, pp. 10-17). Berikut ini notasi dan symbol dari sequence diagram :

Simbol	Nama	Keterangan
	<i>Actor</i>	Menggambarkan orang yang sedang berinteraksi dengan sistem

	<i>Entity Class</i>	Menggambarkan hubungan yang akan dilakukan
	<i>Boundary Class</i>	Menggambarkan sebuah gambaran dari form
	<i>Control Class</i>	Menggambarkan penghubung antara boundary dengan tabel
	<i>A focus of Control & A Life Line</i>	Menggambarkan tempat mulai dan berakhinya message.
	<i>A message</i>	Menggambarkan pengiriman pesan.

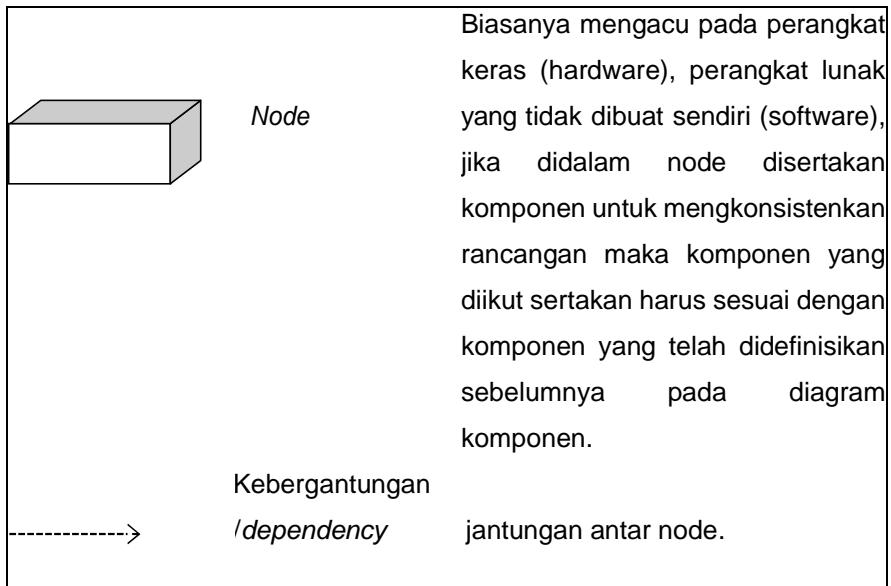
Gambar 2. 8 Simbol Sequence Diagram

Sumber : (Irmayani & Susyatih, 2017, pp. 10-17)

(e) Development Diagram

Dalam diagram penyebaran atau *development diagram*, disajikan pengaturan komponen saat aplikasi dijalankan, *sequence diagram* juga mampu mewakili aspek-aspek seperti sistem tambahan serta hubungan *client/server* dalam sebuah sistem (Shalahuddin, 2013, p. 154). Berikut ini notasi dan symbol dari development diagram :

Simbol	Nama	Keterangan
	<i>Package</i>	Merupakan sebuah bungkusan dari satu atau lebih mode



Gambar 2. 9 Simbol Development Diagram

Sumber : (Shalahuddin, 2013, p. 154)

8. Database

Dalam sebuah sistem tentunya akan ada proses pengolahan database, database akan diolah menggunakan MySQL, menurut (Nita Syafarina, 2016, p. 5) database atau basis data merupakan pengelompokan data yang terhubung untuk memfasilitasi kegiatan dalam mendapatkan informasi, tujuan dari basis data adalah mengatasi masalah pada sistem yang mengandalkan pendekatan berkas.

9. Web Server

Dalam membangun sebuah sistem berbasis website, perlu menggunakan web server sebagai salah satu tools untuk menjalankannya, Web server adalah software yang menjadi tulang punggung dalam arsitektur web. Fungsinya adalah menerima permintaan layanan dari browser web dan membendasnya dengan mengirimkan konten yang diminta. Proses komunikasi antara web browser dan web server ini berlangsung menggunakan protokol HTTP (Hypertext Transfer Protocol) (Shalahuddin, 2013, p. 250).

10. Bahasa Pemrograman

(a) Pengertian Python

Dalam membangun sebuah sistem tentunya memerlukan bahasa pemrograman. Untuk kebutuhan penelitian ini, di mana fokus utama adalah pemrosesan data dan implementasi algoritma K-Means, bahasa pemrograman Python dipilih sebagai alat utama. Python merupakan bahasa pemrograman interpretatif tingkat tinggi yang serbaguna. Dibuat

oleh Guido van Rossum dan pertama kali dirilis pada tahun 1991, Python dirancang dengan filosofi yang menekankan pada keterbacaan kode (Mark Lutz, 2013, p. 3)

Bahasa ini dipilih karena kemampuannya yang sangat fleksibel dan dikenal mendukung pengolahan data secara efektif. Fleksibilitas ini menjadikannya sangat sesuai untuk berbagai aplikasi seperti pengembangan web, analisis data, dan kecerdasan buatan, yang relevan dengan tujuan penelitian untuk membangun prototipe sistem pemetaan aset.

(b) Pengertian Hypertext Markup Language (HTML)

Dalam proses pembuatan sistem berbasis website sering kali HTML menjadi salah satu bahasa yang sering dipake oleh programmer, berdasarkan penjelasan dari (Achmad, 2016, p. 10) menerangkan bahwa “*Hypertext Markup Language (HTML)* merupakan bahasa pemrograman web yang memberikan intruksi kepada web browser tentang cara mengatur dan menampilkan konten pada halaman web. Secara esensial, HTML bertindak sebagai dasar bagi situs web. HTML mampu merender elemen-elemen seperti teks, tabel, tautan, gambar, dan video.

(c) Intranet

Dalam dunia digital saat ini lalu lintas data untuk mendapatkan informasi memerlukan internet sebagai penghubung untuk berkomunikasi. *intranet* merupakan jaringan komputer di dalam suatu perusahaan yang mengadopsi protocol komunikasi data standar yang digunakan dalam internet. Dengan kata lain, *intranet* dapat dianggap memiliki koneksi internet di dalam lingkungan internal perusahaan(Khoe Yao Tung, 2001, p. 5).

B. Tinjauan Studi

Dalam pelaksanaan penelitian ini, dilakukan proses pengumpulan informasi berdasarkan penelitian-penelitian terdahulu yang relevan. Kajian mengenai klasifikasi daerah rawan tindak kekerasan terhadap perempuan dan anak telah banyak dilakukan sebelumnya, begitu pula penelitian yang menerapkan teori Algoritma *K-Means*. Penelitian-penelitian terdahulu ini berkontribusi sebagai bahan perbandingan serta sumber referensi dalam pelaksanaan penelitian ini. Berikut merupakan beberapa penelitian yang menjadi rujukan dalam penelitian ini:

(1) Algoritma *K-Means* Untuk Peningkatan Model Segmentasi Data Aset Tetap

Pada PT XYZ (Dwi Prayogo, Rudi Kurniawan, and Yudhistira Arie Wijaya, 2025). Pengelolaan aset tetap merupakan salah satu elemen penting dalam operasional

suatu Organisasi ataupun perusahaan, termasuk di PT. XYZ. Namun demikian, metodologi manual sering menimbulkan tantangan seperti inefisiensi, kesalahan klasifikasi, dan kekurangan analitis yang menghambat dalam pengambilan keputusan yang sangat strategis. Penelitian ini berusaha untuk menerapkan algoritma K-Means dalam peningkatan model segmentasi data aset tetap, dengan penekanan terkonsentrasi pada pengoptimalan parameter nilai K, yang bertujuan meningkatkan ketepatan hasil pengelompokan. Berbagai studi sebelumnya, termasuk yang dilakukan oleh Yuan & Yang (2019), Ahmed et al. (2020), dan Ikotun et al. (2023), menunjukkan bahwa algoritma K-Means telah digunakan secara luas, namun, masih ada kekurangan dalam pengoptimalan parameter dan penerapannya pada data aset tetap. Penelitian ini mengadopsi metodologi eksperimental yang memanfaatkan data aset tetap dari PT. XYZ sebagai objek penelitian. Algoritma K-Means diterapkan untuk mengkategorikan data berdasarkan karakteristik tertentu, seperti kategori aset, umur ekonomis, kondisi dan tingkat pemanfaatan. Hasil penelitian menunjukkan bahwa pendekatan ini mampu meningkatkan efisiensi proses pengelolaan aset tetap yang diperoleh nilai Davies Bouldin Index (DBI) = 0,395 melalui pengujian parameter dengan nilai K Optimal = 4. Dengan demikian, penelitian ini memberikan kontribusi teoretis dalam pengembangan algoritma K-Means dan manfaat praktis bagi perusahaan dalam pengelolaan aset tetap mereka.

- (2) **Penerapan Data Mining untuk Pemetaan Siswa Berprestasi menggunakan Metode Clustering K-Means** (Rahayu, Anshor and Afriantoro, 2024). Pendidikan memiliki peranan penting sebagai modal utama membangun karakter bangsa. Keberhasilan murid dievaluasi berdasarkan pelajaran teoritis dan praktis, serta kehadiran murid selama di dalam kelas. Jumlah pemrosesan data yang terus meningkat mengharuskan penggunaan strategi dan metode sehingga dapat ditransformasikan menjadi informasi dan pengetahuan yang dapat dimanfaatkan oleh pendidik dalam proses pembuatan kebijakan. Tujuan penelitian ini yaitu mengelompokan prestasi siswa berdasarkan nilai siswa menggunakan data mining dengan menerapkan algoritma K-Means. Dalam penelitian ini data yang digunakan adalah nilai raport PAS Ganjil SMK TON'S periode tahun 2023/2024. Perhitungan klasterisasi dengan algoritma K-Means berhasil mengelompokkan 24 siswa menjadi 3 cluster. Dimana siswa dengan kategori tinggi sebanyak 6 siswa, siswa dengan kategori cukup sebanyak 8 siswa, dan siswa dengan kategori rendah sebanyak 10 siswa. Berdasarkan hasil evaluasi, didapat nilai Davies Bouldin Index (DBI) 0.144 atau mendekati dengan 0 sehingga dapat disimpulkan bahwa hasil cluster yang dihasilkan cukup baik.

Dengan hasil clustering yang telah diterapkan diharapkan bisa membantu dalam mengambil keputusan dengan tepat untuk menentukan siswa berprestasi.

- (3) **Penerapan K-Means Clustering Untuk Pemetaan Wilayah Rawan Bencana Alam Kota Malang** (I Kadek Riski Dwi Putra, and Ahmad Faisol, 2024). Bencana alam merupakan salah satu fenomena alam yang membahayakan dan dapat mengancam kelangsungan hidup manusia. Pemetaan wilayah rawan bencana yang dilakukan oleh BPBD Kota Malang dapat menjadi jawaban dari permasalahan tersebut. Namun pada halaman resmi website BPBD Kota Malang belum memadainya pengelompokan data bencana. Dalam ilmu sciene dan teknologi telah banyak permasalahan yang dapat diselesaikan oleh disiplin ilmu data mining untuk mengklasterkan data tertentu. Penggunaan Algoritma K-Means Clustering salah satu metode analisis data yang bisa digunakan dalam pengelompokan wilayah berdasarkan karakteristik yang mirip. Tujuan pengelompokan tersebut diharapkan dapat membantu tugas BPBD Kota Malang untuk mengidentifikasi wilayah dengan tingkat resiko yang sama. Menurut hasil pengujian menggunakan metode K-Means Clustering dalam pemetaan wilayah bencana di Kota Malang pada sistem dan dibandingkan dengan perhitungan manual memiliki tingkat akurasi sebesar 77,19%. Pengembangan pada aplikasi untuk perhitungan K-Means Clustering data ditambahkan untuk perhitungan untuk pemilihan optimasi pada pusat centroid dengan tujuan untuk menetapkan titik pusat agar jarak yang dihitung tidak berubah - ubah.
- (4) **Implementasi Algoritma K-Means Pada Pemetaan Daerah Terdampak Tanah Longsor Di Jawa Tengah** (Rahmawati, Nurina Sari and Jajuli, 2024). Karakteristik wilayah dan letak geografis Indonesia menjadikan wilayah Indonesia menjadi daerah rawan bencana tanah longsor. Salah satu wilayah yang berpotensi tinggi mengalami bencana tanah longsor adalah Jawa Tengah. Dampak dari bencana tanah longsor di Jawa Tengah mengancam kehidupan masyarakat karena mengakibatkan rusaknya infrastruktur, merugikan harta benda dan menghilangkan nyawa manusia. Untuk meminimalisir dampak yang terjadi, diperlukan mitigasi bencana tanah longsor. Salah satu langkah awal dalam mitigasi bencana tanah longsor adalah pemetaan daerah. Pada penelitian ini dilakukan pengklasteran daerah berdasarkan dampak bencana tanah longsor di Jawa Tengah menggunakan algoritma K-Means dan pemetaan menggunakan QGIS. Dari metode elbow diperoleh nilai k yang paling optimal adalah K=2. Dari 2 Cluster yang terbentuk, sebanyak 24 daerah dengan kategori rendah dan sebanyak 7 daerah dengan kategori tinggi. Hasil pemodelan tersebut dievaluasi menggunakan Silhouette Index (SI) untuk mengetahui kualitas dari cluster. Dari

hasil pengujian didapatkan nilai SI 0,48 dengan struktur lemah. Penelitian ini diharapkan dapat memberikan penyajian data pemetaan daerah terdampak tanah longsor di Jawa Tengah agar dapat digunakan sebagai informasi untuk meningkatkan kewaspadaan dan kesiapsiagaan akan terjadinya bencana tanah longsor.

- (5) **Sistem Informasi Geografis Pemetaan Wilayah Kelayakan Tanaman Cabai Menggunakan Metode K-Means** (Rahmawati, Nurina Sari and Jajuli, 2024). Penelitian ini bertujuan untuk menerapkan Sistem Informasi Geografis (SIG) dalam pemetaan wilayah kelayakan tanaman cabai di Kabupaten Batu Bara menggunakan metode K-means. Metode ini digunakan untuk mengidentifikasi dan mengelompokkan area-area yang paling layak untuk budidaya tanaman cabai berdasarkan sejumlah faktor yang relevan. Data spasial dan atribut wilayah, termasuk kondisi tanah, suhu tanah, pH tanah, kelembaban, curah hujan, dan faktor-faktor lain yang memengaruhi pertumbuhan tanaman cabai, dikumpulkan dan dianalisis. Hasil pemetaan ini diintegrasikan ke dalam sistem yang akan menampilkan informasi mengenai wilayah kelayakan tanaman cabai yang ada di Kabupaten Batu Bara, dan juga penggunaan openstreetmap yang diimplementasikan pada aplikasi agar dapat memberikan informasi visual berupa titik lokasi lahan yang sesuai. Berdasarkan pengujian yang telah dilakukan dengan menggunakan 3 cluster C1 = Sesuai, C2 = Cukup Sesuai, C3= Kurang Sesuai, dan 12 data. Didapatkan hasil clustering dengan jumlah C1 (Sesuai) = 5 data, C2(Cukup Sesuai) = 4 data, C3 (Kurang Sesuai) = 3 data, dan hasil pemetaan telah ditandai berdasarkan masing-masing warna, hasil “sesuai” ditandai dengan warna hijau, hasil “cukup sesuai” warna kuning, dan ‘kurang sesuai’ warna merah.
- (6) **Pemetaan Penyebaran Covid-19 Pada Tingkat Kabupaten/Kota Di Pulau Jawa Menggunakan Algoritma K-Means Clustering** (Lidia Gayatri, 2021). Coronavirus atau Covid-19 adalah virus yang ditemukan di Wuhan, China pada Desember 2019, memiliki kemampuan penyebaran yang cukup cepat diseluruh dunia termasuk di Indonesia. Kasus Covid-19 mulai masuk di Indonesia pada 2 Maret 2020 dan kasus Covid-19 terus meningkat setiap hari, tercatat 1.492.002 kasus Covid-19 pada 27 Maret 2021 di Indonesia. Kasus Covid-19 yang terus meningkat menyebabkan perlunya pemetaan tingkat kerawanan penyebaran Covid- 19 khususnya di Pulau Jawa menggunakan data dari website resmi pemerintah pada tingkat provinsi dengan menggunakan 3 parameter, yaitu jumlah kasus dirawat, sembuh, dan meninggal. Untuk menentukan banyaknya *cluster* digunakan algoritma *K-Means* dan metode *Davies Bouldin Index* (DBI).

Algoritma K-Means adalah salah satu metode *clustering* yang dapat membagi data ke dalam beberapa kelompok atau *cluster*, sedangkan DBI digunakan untuk menghitung kemiripan setiap *cluster*. Hasil pengujian menunjukkan *cluster* terbaik pada ukuran *cluster* 3 dengan nilai DBI 0.609. Terdapat 3 tingkat kerawanan, yaitu kerawanan rendah terdapat pada *cluster* 0 yang memiliki 105 kabupaten/kota, kerawanan sedang terdapat pada *cluster* 2 yang memiliki 7 kabupaten/kota, dan kerawanan tinggi terdapat pada *cluster* 1 yang memiliki 7 kabupaten/kota. Hasil pemetaan menunjukkan kabupaten/kota yang berada pada tingkat kerawanan tinggi berada di Kota Jakarta Utara, Kota Jakarta Barat, Kota Jakarta Pusat, Kota Jakarta Selatan, Kota Jakarta Timur, Kota Surabaya, dan Kota Semarang. Hasil dari penelitian diharapkan dapat digunakan sebagai acuan oleh masyarakat maupun pemerintah dalam melihat setiap daerah yang memiliki tingkat kerawanan rendah, sedang, dan tinggi dalam penyebaran Covid-19 di Pulau Jawa, agar pemerintah memfokuskan perhatian lebih kepada kabupaten/kota yang memiliki tingkat kerawanan tinggi.

- (7) **Metode K-Means Untuk Pemetaan Persebaran Usaha Mikro Kecil Dan Menengah** (Ramdhhan Pebriyanto, 2024). Perkembangan di era globalisasi saat ini sangat bergantung pada sektor ekonomi yang menjadi tolok ukur keberhasilan yang dilakukan oleh pemerintah. Peran masyarakat dalam pembangunan nasional di bidang ekonomi adalah adanya Usaha Mikro, Kecil, dan Menengah (UMKM). Untuk meningkatkan peran UMKM sebagai tolok ukur keberhasilan sektor ekonomi tersebut, maka harus ada dukungan dari pemerintah seperti pemberian bantuan bagi para pelaku usaha yang memiliki keterbatasan biaya. Tujuan dari penelitian ini adalah untuk mengetahui kelompok usaha masyarakat sebagai tolok ukur tingkat usaha sehingga memudahkan pemerintah dalam memberikan bantuan. Metode K-Means Clustering merupakan metode yang digunakan untuk pengelompokan tingkat usaha berdasarkan pendapatan yang ada di masyarakat saat ini. Hasil dari penelitian ini adalah sistem pengelompokan tingkat usaha berbasis website yang digunakan oleh Dinas Koperasi dan UKM dengan mengelompokkannya menjadi usaha mikro, kecil, dan menengah berdasarkan pendapatan/asset.
- (8) **Penerapan Metode K-Means Clustering untuk Pemetaan Pengelompokan Lahan Produksi Tandan Buah Segar** (Hidayati et al., 2021). Di Perkebunan Sei Lukut, Desa Maredan Barat, Kecamatan Tualang, Kabupaten Siak, Provinsi Riau, PT. Surya Intisari Raya, sebuah perusahaan swasta, mengelola perkebunan kelapa sawit. Memiliki 4 bagian lahan kelapa sawit yang terdiri dari 216 blok dengan total sekitar 4.000 Ha. Blok kelapa sawit biasanya mencakup 20 hektar

dan berisi 28.000 pohon kelapa sawit, dengan kapasitas produksi bulanan sebesar 57 ton. Pemetaan klaster produksi tandan buah segar berupaya membantu pelaku usaha memutuskan kebijakan apa yang akan diterapkan untuk meningkatkan akurasi dan produktivitas produksi minyak sawit. Metode K-Means merupakan komponen dari metode clustering, yang merupakan subset dari kelompok Unsupervised Learning dan digunakan untuk mempartisi data ke dalam berbagai kategori. Untuk mengelompokkan blok lahan berdasarkan delapan data variabel luas pokok, panjang panen, daun lepas, curah hujan, pupuk, tujuan, dan persentase keberhasilan, penelitian ini akan menerapkan Indeks Davies Bouldin dengan alat RapidMiner. Kesimpulan akhir dari penelitian ini adalah sebuah aplikasi yang dapat memetakan pengelompokan areal produksi tandan buah segar dengan menerapkan metode K-Means Clustering, dengan nilai Davies Bouldin Index terkecil sebesar 0,921 pada jumlah cluster 3 yang termasuk Cluster C1 (Produktivitas Sedang). Terdiri dari 96 blok tanah, Cluster C2 (Produktivitas Rendah) terdiri dari 41 blok tanah, dan Cluster C3 (Produktivitas Tinggi) terdiri dari 79 blok tanah.

- (9) **Sistem Informasi Geografis Pemetaan Warga Kurang Mampu Di Kelurahan Karang Besuki Menggunakan Metode K-Means Clustering** (Ali Hasymi, Faisol and Ariwibisono, 2021). Kemiskinan menjadi salah satu permasalahan di negara berkembang, bantuan sosial merupakan salah satu cara untuk mengurangi tingkat kemiskinan. Dalam penerapannya masih ada penerima bantuan sosial yang kurang tepat sasaran dikarenakan penggunaan data yang kurang maksimal serta minimnya informasi dan letak geografis, termasuk di Kelurahan Karangbesuki. Sebagai solusi masalah ini, diperlukan sistem informasi geografis yang menampilkan informasi mengenai pemetaan warga kurang mampu, diharapkan dengan adanya sistem informasi geografis tersebut dapat memberikan informasi dan mengelompokkan warga kurang mampu, serta mengetahui lokasi warga kurang mampu. Sistem yang akan dibangun yaitu sistem informasi geografis pemetaan warga kurang mampu meng-gunakan metode clustering serta dikombinasikan dengan QGIS 2.18. Metode clustering yang digunakan dalam sistem ini adalah metode K-Means clustering, kriteria yang digunakan pada pengelompokan data berupa umur, pekerjaan, pendidikan dan tanggungan keluarga. Dari hasil pengujian sistem informasi geografis menggunakan K-Means Clustering dengan 3 cluster untuk pengelompokan warga di kelurahan Karangbesuki, dengan menggunakan 325 sampel data, terdapat hasil pengelompokan sebanyak 178 (55%) warga yang tidak layak menerima bantuan, 99 (30%) warga yang kurang layak menerima bantuan dan

48 (15%) warga yang layak menerima bantuan. Berdasarkan hasil pengujian kinerja sistem dengan membandingkan pengelompokan secara manual atau data lama dengan pengelompokan pada sistem, sistem ini cukup baik dalam mengklasifikasi warga kurang mampu dengan tingkat akurasi 91%..

(10) Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means

Clustering (Hidayati *et al.*, 2021). *Clustering* merupakan proses pengelompokan sekumpulan data ke dalam klaster yang memiliki kemiripan. Kemiripan dalam satu klaster ditentukan dengan perhitungan jarak. Untuk melihat perfoma beberapa perhitungan jarak, dalam penelitian ini penulis menguji pada 6 data yang memiliki atribut berbeda, yakni 2, 3, 4, dan 6 atribut. Dari hasil uji perbandingan rumus jarak pada *K-Means clustering* menggunakan *Silhouette coefficient* dapat disimpulkan bahwa: 1) *Chebyshev distance* memiliki perfoma yang stabil baik untuk data dengan sedikit atribut maupun banyak. 2) *Average distance* memiliki hasil *Silhouette coefficient* paling tinggi dibandingkan dengan pengukuran jarak lain untuk data yang memiliki outliers seperti data 3. 3) *Mean Character Difference* mendapatkan hasil yang baik hanya untuk data dengan sedikit atribut. 4) *Euclidean distance*, *Manhattan distance*, dan *Minkowski distance* menghasilkan nilai baik untuk data yang memiliki sedikit atribut, sedangkan untuk data yang banyak atribut mendapatkan nilai cukup yang mendekati 0,5.

Tabel 2. 1 Tinjauan Studi Penelitian

No	Nama Peneliti	Judul	Sumber	Kontribusi
1.	Dwi Prayogo, Rudi Kurniawan, Yudhistira Arie Wijaya, Umi Hayati (2025)	Algoritma K- Means Untuk Peningkatan Model Segmentasi Data Aset Tetap Pada PT XYZ	Universitas Lampung (http://dx.doi.org/10.23960/jite.v13i1.5936)	Berkontribusi dalam optimalisasi parameter nilai K untuk meningkatkan ketepatan hasil pengelompokan aset.

No	Nama Peneliti	Judul	Sumber	Kontribusi
2	Nia Dwi Rahayu, Abdul Halim Anshor, Irfan Afriantoro (2024)	Penerapan Data Mining untuk Pemetaan Siswa Berprestasi menggunakan Metode Clustering K-Means	Universitas Pelita Bangsa, Vol. 6 No. 1 (2024): JUKI : Jurnal Komputer dan Informatika, Edisi Mei 2024 https://ioinformatic.org/index.php/JUKI/article/view/474	Berkontribusi dalam segmentasi data dan evaluasi hasil clustering
3	I Kadek Riski Dwi Putra, Ahmad Faisol, Mira Orisa (2024)	Penerapan K-Means Clustering Untuk Pemetaan Wilayah Rawan Bencana Alam Kota Malang	Institut Teknologi Nasional Malang, (Jurnal Mahasiswa Teknik Informatika) Vol. 8 No. 5, Oktober 2024 https://ejournal.itn.ac.id/index.php/jati/article/view/10586	Berkontribusi dalam pemetaan data berbasis clustering dengan algoritma K-Means untuk analisis wilayah.
4	Erika Rahmawati, Betha Nurina Sari, Mohamad Jajuli (2024)	Implementasi Algoritma K-Means Pada Pemetaan Daerah Terdampak Tanah Longsor Di Jawa Tengah	Universitas Singaperbangsa Karawang, Vol. 8 No. 3 (2024): JATI Vol. 8 No. 3, https://ejournal.itn.ac.id/index.php/jati/article/view/10586	Berkontribusi dalam metode evaluasi clustering dengan Silhouette sebagai alat

No	Nama Peneliti	Judul	Sumber	Kontribusi
			php/jati/article/view/9487	ukur efektivitas klasterisasi.
5	Angga Pratama, Cinta Amalia, Rizki Putra Phonna (2024)	Sistem Informasi Geografis Pemetaan Wilayah Kelayakan Tanaman Cabai Menggunakan Metode K-Means	Sistem Informasi, Universitas Malikussaleh, Vol. 8 No. 3 (2024): JATI Vol. 8 No. 3 https://ejournal.itn.ac.id/index.php/jati/article/view/9155	Berkontribusi dalam integrasi hasil clustering untuk visualisasi data.
6	Lidia Gayatri, Hendry (2021)	Pemetaan Penyebaran Covid-19 Pada Tingkat Kabupaten/Kota Di Pulau Jawa Menggunakan Algoritma K-Means Clustering Lidia	Sebatik Vol. 25 No.2 (2021) https://jurnal.wicida.ac.id/index.php/sebatik/article/view/1307	Kontribusinya yaitu untuk meghitung Min-Max Scalling untuk Normalisasi Data
7	Dicky Jordan Aji Putra, Dwi Remawati, Tri Irawati (2021)	Metode K-Means Untuk Pemetaan Persebaran Usaha Mikro Kecil Dan Menengah Metode K-	Jurnal TIKomSiN, Vol. 9, No. 2, (2021) https://p3m.sinus.ac.id/jurnal/index.php/TIKO	Kontribusinya yaitu untuk meghitung Min-Max Scalling untuk Normalisasi Data

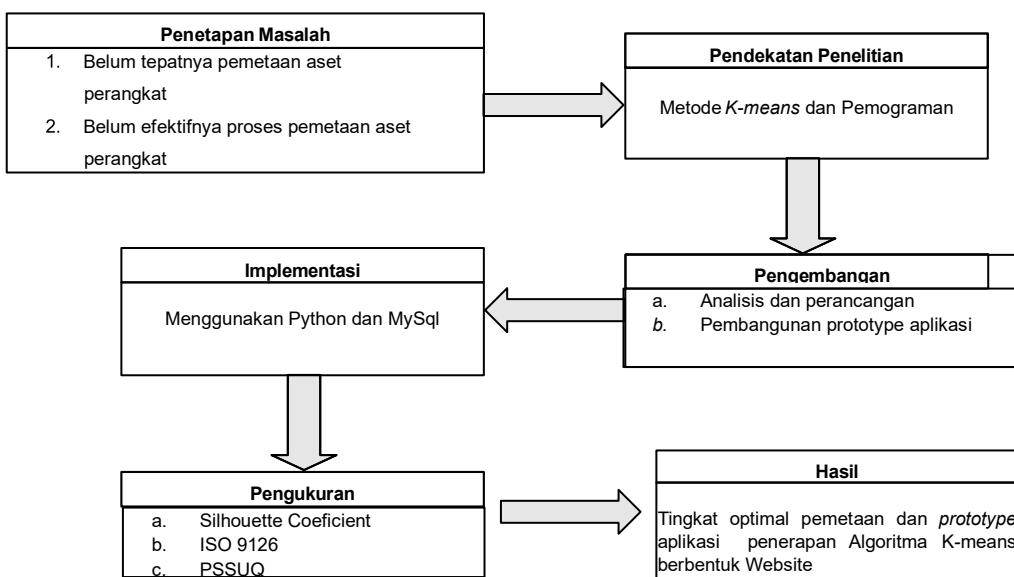
No	Nama Peneliti	Judul	Sumber	Kontribusi
		Means Untuk Pemetaan Persebaran Usaha Mikro Kecil Dan Menengah	mSiN/article/vi ew/574	
8	Dina Nuriyah Ramadhani, Aris Trijaka Harjanta, Aptanang Tyogi, Serli Agnes Ajhara	Penerapan Metode K-Means Clustering untuk Pemetaan Pengelompokan Lahan Produksi Tandan Buah Segar	JURNAL INFORMATIK A, Vol.10 No.1. (2023) https://ejournal.bsi.ac.id/ejournal .bsi.ac.id/ejournal al/index.php/ji/ article/view/15 621	Berkontribusi Dalam Pembuatan Prototyping Produk
9	Muhammad Ali Hasymi, Ahmad Faisol, FX. Ariwibisono	Sistem Informasi Geografis Pemetaan Warga Kurang Mampu Di Kelurahan Karang Besuki Menggunakan Metode K-Means Clustering	Jurnal Mahasiswa Teknik Informatika (JATI), Vol. 5 No. 1, 2021 https://ejournal .itn.ac.id/index. php/jati/article/ view/3269	Berkontribusi Dalam Pembuatan Prototyping Produk

No	Nama Peneliti	Judul	Sumber	Kontribusi
10	Rahmatina Hidayati, Anis Zubair, Aditya Hidayat Pratama, Luthfi Indana (2021)	Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering	Techno.COM, Vol. 20, No. 2, (2021) https://publikasi.i.dinus.ac.id/index.php/technoc/article/view/4556	Berkontribusi Sebagai Acuan Dalam Proses Perhitungan Silhouette Coeficient

Berdasarkan penelitian sebelumnya yang dilakukan oleh (Dwi Prayogo dkk, 2025) mengenai optimalisasi parameter nilai K dalam algoritma K-Means untuk segmentasi aset tetap, dapat dilihat bahwa penerapan metode K-Means telah terbukti efektif dalam membantu proses pengelompokan aset. Namun, penelitian ini memiliki perbedaan orisinalitas, yaitu fokus pada aset perangkat perusahaan pengembang aplikasi dengan variabel khusus seperti Status, Usia, *Checkout*, *Warranty Expired*, dan *Purchase Cost*, serta menghasilkan kategori keputusan berupa diperbaiki, ditingkatkan, atau diganti. Dengan demikian, penelitian ini memberikan kontribusi baru dalam konteks pengelolaan aset perangkat IT yang belum banyak diteliti sebelumnya.

C. Kerangka Pemikiran

Kerangka pemikiran pemecahan masalah penelitian ini digambarkan pada gambar 2.9 kerangka pemikiran untuk memecahkan masalah dalam penelitian ini data digambarkan pada gambar dibawah ini:



Gambar 2. 10 Kerangka Pemikiran

Kerangka pemikiran pada Gambar 2.9 dapat dijelaskan sebagai berikut:

- (1) Penelitian ini diawali dengan melakukan penetapan masalah belum tepat dan efektif yang terjadi saat proses pemetaan aset perangkat perusahaan.
- (2) Metode merupakan algoritma yang digunakan dalam proses penelitian, pada permasalahan ini pendekatan yang dilakukan dengan menggunakan metode *K-Means* dan pemograman.
- (3) Pengembangan dilakukan dengan menganalisis data desain, dan pengembangan *prototype* aplikasi untuk pemetaan dengan variabel yang ditentukan.
- (4) Implementasi dilakukan dengan melakukan menggunakan Python dan MySql
- (5) Pengukuran kesesuaian *prototype* dengan masalah dan kebutuhan yang sudah ditentukan dilakukan evaluasi dengan menerapkan algoritma *Silhouette Coefficient* untuk mengukur tingkat optimal hasil clustering dari algoritma *K-Means*. Selain itu, pengujian ISO 9126 sebagai uji ahli digunakan untuk menilai fungsionalitas sistem, sementara uji *PSSUQ* diterapkan guna mengevaluasi aspek kebergunaan dan kepuasan pengguna terhadap aplikasi yang dikembangkan.
- (6) Hasil yang dikeluarkan yaitu sebuah sistem informasi rekomendasi asset untuk investasi sudah tepat dan efektif.

D. Hipotesis Penelitian

Algoritma *K-Means* merupakan salah satu metode dalam *clustering* yang digunakan untuk mengelompokkan data ke dalam beberapa klaster berdasarkan karakteristik tertentu. Pada penelitian sebelumnya yang berjudul “Algoritma *K-means* Untuk Peningkatan Model Segmentasi Data Aset Tetap Pada PT XYZ (Dwi Prayogo, dkk 2025), penggunaan algoritma *K-Means* menunjukkan tingkat keberhasilan penggunaan algoritma K-Means dalam memetakan data aset tetap dari PT. XYZ sebagai objek penelitian. Basis kerja pada algoritma K-Means yaitu melakukan suatu pengelompokan, Bawa pada penelitian ini juga berkaitan dengan pengelompokan data aset perangkat perusahaan. maka dapat ditetapkan hipotesis penelitian ini adalah bahwa penerapan algoritma K-Means di duga dapat melakukan pemetaan data aset perangkat perusahaan pengambang aplikasi