

BAB III METODOLOGI PENELITIAN

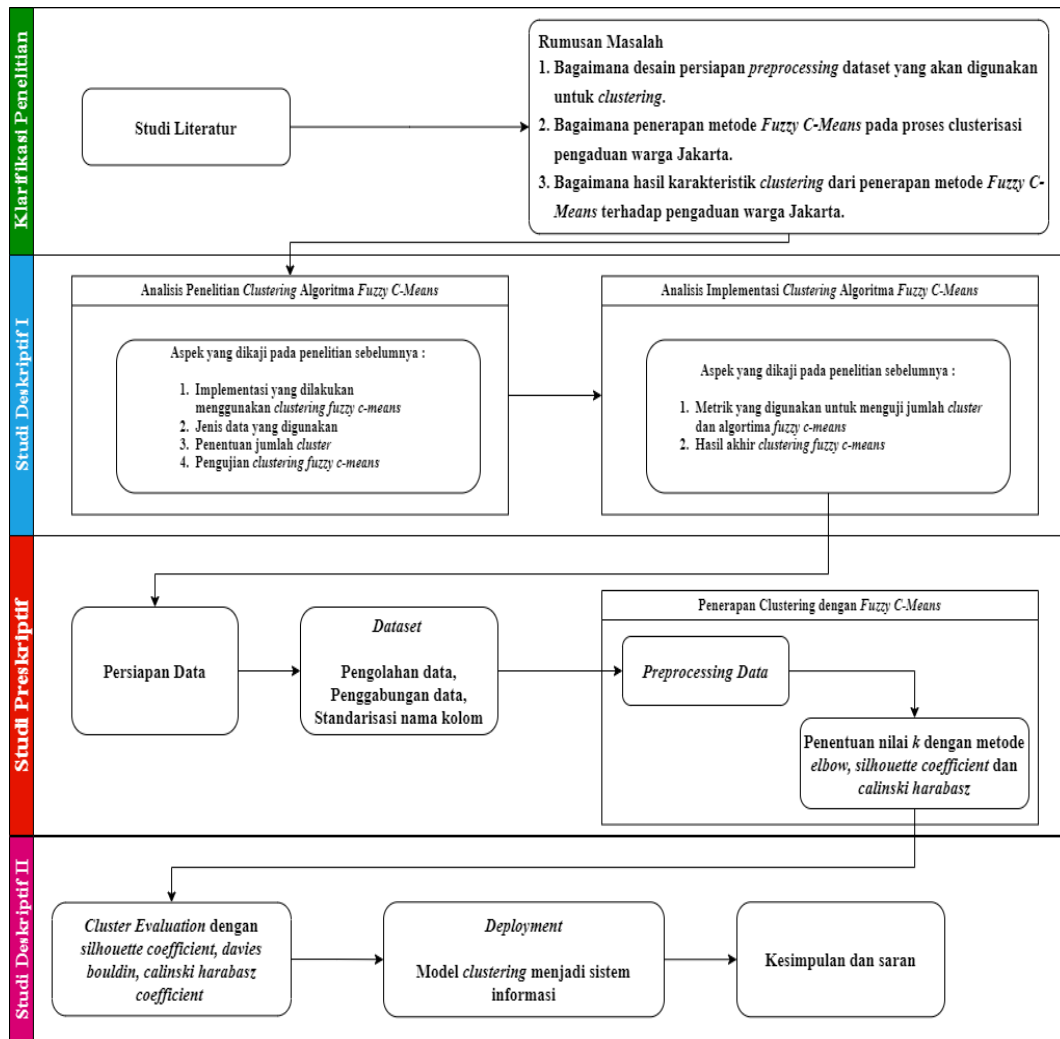
3.1 Desain Penelitian

Desain penelitian disusun secara runtut untuk memudahkan peneliti dalam melaksanakan penelitian serta memberikan gambaran dalam memudahkan proses penelitian. Metode penelitian yang diterapkan ialah *Design Research Methodology* (DRM). DRM salah satu metode penelitian yang efektif dan efisien karena mendukung pendekatan secara praktis, bermanfaat dan realistis dengan mengidentifikasi area penelitian, proyek dan program secara konsisten terdapat 4 tahapan yakni klarifikasi penelitian, studi deskriptif I, studi preskriptif dan studi deskriptif II dengan masing-masing *output* pada tiap tahapannya (Blessing dan Chakrabarti, 2009). Penulis memilih desain penelitian ini karena relevan dan cocok digunakan untuk penelitian yang berkaitan area penelitian, proyek dan program. Adapun Gambar 3.1 Desain Penelitian yang dibuat menggunakan metode *Design Research Methodology* (DRM) berikut tahapan yang dilaksanakan selama penelitian secara berurutan (Blessing dan Chakrabarti, 2009).

3.1.1 Klarifikasi Penelitian

Pada tahap awal penelitian dilakukan Klarifikasi Penelitian. Dalam tahap ini langkah untuk mengidentifikasi inti masalah. Dibantu dengan studi literatur bersumber dari berbagai jurnal, buku, skripsi dan bacaan dari internet. Dengan maksud untuk mempelajari dan memperkuat ketepatan solusi yang ditentukan. Selama proses ini menjadi acuan untuk menyelesaikan dan menemukan perumusan masalah. Beberapa literatur yang dilakukan mulai dari mencari penelitian terkait, teoritis tentang pengaduan warga Jakarta, Open data Jakarta, data mining, *machine learning*, *clustering*, *preprocessing data*, *fuzzy c-means* dan *cluster evaluation*.

Tahapan Klarifikasi Penelitian ini memiliki *output* awal penemuan topik permasalahan mengenai pengaduan warga Jakarta, selanjutnya dilakukan kajian literatur terhadap penelitian yang relevan dan perumusan atas identifikasi masalah. Sehingga tahap ini mengambil fokus pemecahan permasalahan dengan penerapan *clustering* menggunakan algoritma *Fuzzy C-Means*.



Gambar 3.1 Desain Penelitian

3.1.2 Studi Deskriptif I

Selanjutnya tahap studi deskriptif I dilakukan untuk meningkatkan pemahaman mengenai permasalahan dan menginvestigasi fenomena yang memfokuskan awal pendukung solusi pengembangan berdasarkan hasil studi literatur dan perumusan masalah tahap sebelumnya. capaian yang dituju pada tahap ini ialah proses awal pengukuran dan penentuan pembaharuan dari solusi penelitian. Penulis ditahap ini melakukan analisis seputar penelitian terdahulu yang berkaitan dengan implementasi *clustering* algoritma *fuzzy c-means* dengan tujuan memperdalam pemahaman tentang masalah yang diteliti terkait implementasi *clustering* algoritma *fuzzy c-means* serta analisisnya. Mulai dari mengkaji jenis data yang digunakan, penentuan jenis *cluster* dan pengujian *clustering fuzzy c-means*. Selanjutnya

melakukan menganalisis beberapa kajian mengenai metrik untuk menguji jumlah *cluster* dan algoritma *fuzzy c-means* serta melihat bagaimana hasil dari *clustering fuzzy c-means*.

3.1.3 Studi Preskriptif

Pada tahapan studi preskriptif ini merupakan penerapan *clustering* menggunakan algoritma *fuzzy c-means* dalam penelitian. Pada tahap ini bertujuan untuk pengembangan secara sistematis atas *output* pada tahap studi deskriptif hingga proses evaluasi atas pengukuran yang ditemukan. Dalam tahapan ini penulis melakukan persiapan data yaitu mengambil kumpulan data-data pengaduan warga Jakarta melalui kanal informasi resmi milik pemerintah Provinsi Jakarta di data.jakarta.go.id. Data tersebut, dapat diperoleh dengan judul pengaduan masyarakat yang disampaikan melalui aplikasi CRM (Cepat Respon Masyarakat). Dari sumber tersebut didapatkan data di beberapa wilayah Jakarta dimulai dari tahun 2019-2021. Tahap ini akan berdampak kepada cara pengolahan data yang tersedia. Sehingga setelah itu dataset diolah menjadi satu kesatuan, karena dataset warga Jakarta terbagi menjadi beberapa wilayah dan tahun pengaduan. Maka dilakukan standarisasi data agar sesuai antar nama-nama kolom dan juga penambahan jenis kolom yang bertujuan memudahkan proses *preprocessing data*. Selanjutnya melakukan penerapan *clustering* dengan *fuzzy c-means* yang diuraikan dalam tahapan berikut:

A. *Preprocessing Data*

Pada tahap ini akan dilakukan proses *cleaning data* supaya data tepat, seralasi, lengkap, konsisten dan berkualitas, dengan tujuan untuk mendeteksi, memperbaiki dan menghapus data yang rusak, ganda serta tidak akurat yang terdapat pada dataset. *Data cleaning* diidentifikasi apakah data tiap kolom pada tabel lengkap, isi kolom yang benar, bila tidak sesuai maka data dapat diganti, dihapus atau dimodifikasi tanpa menghilangkan esensial dari dataset tersebut. selanjutnya redundansi data dilakukan untuk menyesuaikan isi data yang bernilai sama atau ganda menjadi satu kesatuan agar data berkualitas.

Serta *exploratory data* sebagai pemahaman mendalam terhadap data baik berupa gambaran visual, untuk memberikan pemahaman baru bagi peneliti. Setelah

Athoillah Sholahuddin, 2023

PENERAPAN CLUSTERING PENGADUAN WARGA JAKARTA MENGGUNAKAN ALGORITMA FUZZY C-MEANS

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

itu dilakukan *feature engineering* untuk proses membentuk fitur-fitur, parameter dan atribut pada dataset untuk dalam pemudahan penerapan algoritma. Untuk teknik yang digunakan dalam *feature engineering* ini ialah *feature construction* dengan tujuan melakukan *label encoding* pada jenis data kategorikal yang menjadi *raw data* baru untuk mengoptimalkan proses *learning* pada algoritma.

B. Penentuan nilai *cluster* (*k*)

Penentuan nilai *cluster* (*k*) dilakukan guna melihat dan mencari jumlah *cluster* yang ditetapkan untuk membentuk pola data dengan menggunakan *fuzzy c-means* berdasarkan analisis metode *elbow*, *silhouette coefficient* dan *calinski harabasz* sehingga didapatkan jumlah *cluster* yang terbaik dan memudahkan proses *clustering fuzzy c-means*. Setelah mendapatkan nilai *k* maka selanjutnya menerapkan *fuzzy c-means* dengan mencari jarak antar *cluster*, menentukan *fuzzier* atau derajat keanggotaan, melakukan iterasi serta terakhir menemukan nilai *centroids* pada masing-masing *cluster* guna melihat hasil label data berdasarkan titik *centroids* yang menjadi hasil data *clusterisasi*.

3.1.4 Studi Deskriptif II

Pada tahapan studi Deskriptif II ini merupakan evaluasi dari pengembangan *clustering* dan mendukung pembuatan sistem yang didokumentasikan. Untuk menguji algoritma *Fuzzy C-Means* pada data yang di pelajari mesin memiliki akursi optimal dari penerapan *cluster* pengaduan warga Jakarta. Dengan menggunakan pengujian *cluster evaluation*, nilai *silhouette coefficient*, *David Bouldin*, dan *calinski harabasz coefficient*. Bertujuan melihat hasil pembelajaran algoritma yang di terapkan baik, cukup atau kurang mampu dalam proses penerapan *learning clustering*. Setelah melakukan *cluster evaluation* algoritma, tahap berikutnya ialah mengkaji dan mendokumentasikan dari keberhasilan maupun ketidakberhasilan selama proses tindakan atau tahapan sebelumnya. Dalam hal ini peneliti melakukan *output* hasil *learning* kedalam bentuk sistem informasi berupa *website*. Sistem informasi berisi visualisasi dashboard yang tujuannya untuk memberikan informasi kepada warga ataupun pemerintah dalam pengaduan warga. setelah itu peneliti akan menganalisis kesimpulan yang didapat selama proses penelitian serta membuat laporan skripsi.

3.2 Alat dan Bahan Penelitian

Pada pelaksanaan di dalam penelitian yang akan berlangsung, peneliti memerlukan beberapa perangkat untuk pendukung proses tercapainya penelitian. Perangkat tersebut terdiri atas perangkat keras dan perangkat lunak dengan spesifikasi sebagai berikut.

A. Perangkat Keras

1. 11th Gen Intel® Core™ I5-11357 @ 2.42 GHz.
2. RAM 16 GB, 64-bit.
3. SSD 512 GB.
4. Resolusi 1920x 1080, 14 inch.

B. Perangkat Lunak

1. Jupyter notebook.
2. Google *Colab*.
3. Portal Open data Jakarta.
4. Microsoft Windows 11 Home Single Language.
5. Microsoft Excel.
6. *Looker Studio*.
7. *HTML, CSS dan Javascript*.
8. Python language programming.

3.3 Instrumen Penelitian

Instrumen penelitian yang dilakukan ialah proses pengumpulan data untuk dapat melanjutkan ke tahap berikutnya. Data diambil dari Portal Open Data Jakarta, Pemerintah Provinsi Jakarta dengan jenis dataset data pengaduan masyarakat yang disampaikan melalui aplikasi CRM yang mulai dari tahun 2019-2021. Dataset yang terkumpul dan tersedia berjumlah 42.925 pengaduan yang terdiri dari wilayah Jakarta Pusat, Jakarta Barat, Jakarta Selatan dan Kepulauan Seribu. Dari wilayah tersebut menjadi sample dari penelitian sebab dari keempat wilayah tersebut memiliki dominasi kepadatan penduduk terpadat di Jakarta mulai selama kurun 2019-2021. Selanjutnya Dengan metode secara kuantitatif data akan dilakukan proses dan pemecahan masalah dengan *clustering fuzzy c-means*. Kemudian

digunakan proses validitas hasil *cluster* dengan *cluster evaluation silhouette coefficient*, *davies bouldin* dan *calinski-harabasz coefficient*.