

BAB I PENDAHULUAN

1.1. Latar Belakang

Revolusi industri 4.0 memberikan dampak yang signifikan terhadap kehidupan masyarakat sehari-hari, dengan digitalisasi sebagai penggerak utamanya. Era ini ditandai dengan pesatnya perkembangan teknologi dan informasi, menjadikan konektivitas perangkat serta digitalisasi layanan menjadi bagian tak terpisahkan dalam kehidupan. Hasil survei yang dilakukan Asosiasi Pengguna Jasa Internet Indonesia (APJII) pada tahun 2022 menunjukkan bahwa dari 262 juta masyarakat Indonesia, sekitar 210 juta aktif menggunakan Internet. Selain itu, penggunaan jejaring sosial juga semakin meningkat, menjadi wadah bagi masyarakat untuk menyampaikan pendapat, keluhan, dan komentar terkait dengan kegiatan dan pelayanan publik.. Laporan masyarakat ini dikelola oleh Pusat Pengelolaan Aduan Masyarakat (P3M) yang menerima, mengelola, mengklarifikasi, memantau, mengevaluasi dan mengkonfirmasi aduan dari masyarakat (Tarigan & Yetri, 2022) . Dari berbagai macam laporan yang masuk pemerintah perlu lebih teliti dalam mengklasterisasi/ mengkategorikan jenis laporan masyarakat. Tujuannya adalah agar laporan dapat dengan cepat direspon dan mudah dengan sesuai dengan kategorinya yaitu, pengduan tentang jalan rusak, aduan tentang lampu apil, aduan tentang sampah.

Penambangan teks adalah metode yang umum digunakan untuk pengelompokan teks. *Text mining* adalah metode mengubah teks tidak terstruktur menjadi data teks semi terstruktur, bertujuan untuk menemukan pola antar teks dan memecahkan masalah. *Text mining* merupakan metode yang cocok diterapkan dalam proses pencarian informasi di jejaring sosial. Ada beberapa

algoritma yang umum diterapkan pada *Text Mining*, salah satu yang memiliki tingkat akurasi yang baik adalah algoritma K-Means (Dwi Darmawan S, 2022). Algoritma K-Means merupakan algoritma yang paling umum digunakan dalam proses *clustering*. K-Means juga merupakan metode *clustering* yang cukup sederhana, dimana hasil clustering dihitung kelas dengan frekuensi tertinggi. Dengan menggunakan metode Text Mining dengan algoritma K-Means (Vafky, 2022), diharapkan Pemerintah Kota Yogyakarta dapat mengelola pengaduan dengan lebih baik dengan mengelompokkan berbagai jenis laporan yang disampaikan masyarakat (Smith, 2020). K-Means dapat berguna sebagai algoritma pencarian pengelompokan teks, mudah digunakan, dan cocok untuk lingkungan yang cukup kompleks (Johnson & Williams, 2019).

Dalam penelitian ini, metode *clustering*, khususnya menggunakan algoritma K-Means, diaplikasikan untuk mengelompokkan keluhan masyarakat (Rizki M, 2021) menjadi berbagai kategori yang telah ditentukan sebelumnya. Pendekatan ini dipilih karena cocok untuk data keluhan yang bersifat konkret dan memungkinkan fleksibilitas dalam menentukan jumlah kelompok yang ingin dibuat. Algoritma K-Means adalah algoritma non-hirarki yang memulai dengan pembentukan partisi kluster awal, dan kemudian secara iteratif mengoptimalkan partisi tersebut hingga mencapai kesamaan yang signifikan di antara keluhan-keluhan dalam kluster yang sama. Metode K-Means, menjadi alat yang sangat berguna dalam berbagai bidang analisis data dan pengelompokan informasi (Falentino Sembiring, 2020).

Beberapa penelitian terdahulu menerapkan metode algoritma *K-Means Clustering* pengelompokan data di antaranya, (Rofiqo & Windarto, 2018)

meneliti tentang penduduk yang mempunyai keluhan kesehatan dengan menggunakan algoritma K-Means yang bertujuan untuk melakukan pengelompokan jumlah penduduk yang memiliki keluhan kesehatan dengan memanfaatkan dokumen-dokumen keterangan jumlah penduduk dari Badan Pusat Statistik Nasional. Hasil yang diperoleh dari penelitian tersebut yang di olah menggunakan *Ms. Excel* untuk ditentukan nilai centroid dalam 3 cluster yaitu cluster tingkat keluhan kesehatan tinggi, cluster tingkat keluhan kesehatan sedang, cluster tingkat keluhan kesehatan rendah. Data centroid untuk cluster 37,48 dengan tingkat keluhan kesehatan tinggi, data centroid untuk cluster 37,48 Tingkat klaster sakit adalah 27,08, dan data centroid untuk klaster sakit tingkat rendah 14.89. Ini akan menjadi peringkat berbasis indeks penduduk dengan masalah kesehatan di 7. Provinsi dengan keluhan kesehatan tinggi antara lain Jawa Tengah, Yogyakarta, Bali, Nusa Tenggara Barat dan Nusa Tenggara Timur-Tenggara, Kalimantan Selatan, Gorontalo, 18 dan 9 provinsi dengan kesehatan sedang Lainnya termasuk gangguan kesehatan tingkat rendah. Hasil yang diperoleh dari penelitian dapat dijadikan sebagai masukan Pemerintah, negara bagian harus lebih memperhatikan penduduk dengan kondisi kesehatan yang serius.

Algoritma K-Means, yang telah terbukti efektif dalam klasterisasi data, menjadi pilihan yang potensial. Dalam konteks ini, penelitian ini merujuk pada aplikasi sebelumnya dari algoritma K-Means dalam identifikasi penyakit (Sembiring, 2021) dan klasterisasi tingkat kehadiran dosen (Virgo et al., 2020). Dengan demikian, penelitian ini bertujuan untuk mengadopsi algoritma K-Means sebagai alat yang efektif untuk mengklasterisasi dan menganalisis keluhan masyarakat dengan lebih baik.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diraikan, maka permasalahan yang diangkat bagaimana membangun suatu model klasterisasi menggunakan metode *K-Means* untuk melakukan klasterisasi teks aduan.

1.3. Batasan Masalah

Agar mendapatkan hasil penelitian yang sesuai dengan yang di harapkan, maka dibutuhkan pembatasan dalam penggunaan metode dalam penelitian in, antara lain:

1. Penulis hanya melakukan klasterisasi teks aduan dalam 3 kategori yaitu, aduan tentang jalan rusak, aduan tentang sampah, aduan tentang lampu apil.
2. Proses klasterisasi dilakukan hanya pada teks aduan.
3. Bahasa pemrograman yang digunakan bahasa *Python*.
4. Data yang digunakan adalah data *Scraping*.

1.4. Tujuan Penelitian

Tujuan yang hendak di capai dalam penelitian ini adalah membangun suatu model klasterisasi menggunakan metode *K-Means* untuk melakukan klasterisasi teks aduan berdasarkan jenisnya.

1.5. Manfaat Penelitian

Manfaat yang didapat dari penulisan penelitian ini adalah diharapkan mampu memudahkan pihak terkait guna mengklasterisasi aduan masyarakat.