

BAB I

PENDAHULUAN

1.1 Latar Belakang

Akhir tahun 2019 dunia dikejutkan dengan kemunculan virus corona yang mematikan dan menyebar dengan sangat cepat bahkan ke berbagai negara. *Coronavirus disease 2019* (Covid-19) merupakan virus yang menyerang sistem pernafasan pada manusia dengan gejala batuk, demam, pilek, sesak nafas, sakit tenggorokan, dan lemas yang menyebar dengan cepat. Virus ini dapat Virus ini ditularkan dari orang ke orang melalui droplet (tetesan kecil/percikan) yang dihasilkan penderita saat batuk, bersin, atau berbicara. Itu juga dapat menyebar melalui udara atau melalui benda-benda yang terkontaminasi tetesan kemudian disentuh dengan tangan dan tangan tersebut menyentuh bagian mata, hidung, ataupun mulut maka orang tersebut dapat tertular Covid-19. Virus ini pertama kali muncul di kota Wuhan, China pada awal Desember 2019, hingga pada tanggal 20 Januari 2020 WHO sebagai Badan Kesehatan Dunia menetapkan status *Public Health Emergency of International Concern* (PHEIC) sebagai bentuk peringatan bahwa virus ini sangat berisiko (Suni, 2020).

Kasus pertama Covid-19 di Indonesia terjadi pada awal Maret 2020. Sejak kasus pertama sampai Januari 2021, Indonesia terus mengalami peningkatan kasus Covid-19. Berdasarkan pantauan laman *worldometers.info* pada 22 Januari 2021 Indonesia menempati urutan ke 19 kasus terbanyak dengan total kasus 951,651 dan total kematian 27,203. Berbagai cara telah dilakukan untuk mengurangi kasus Covid-19 di Indonesia dengan beberapa kebijakan seperti memberlakukan *social distancing* atau Pembatasan Sosial Berskala Besar (PSBB) di seluruh daerah di Indonesia dengan kebijakan masing-masing pemerintahan daerah. Pemerintah pusat juga telah menyediakan *website* yang memberikan informasi lengkap untuk jumlah kasus terkonfirmasi positif, sembuh, meninggal, dll secara *realtime*, bahkan telah tersedia informasi pengelompokan daerah dengan risiko tinggi,

sedang, rendah, dan daerah tidak ada kasus Covid-19 yang dapat diakses secara bebas oleh

siapa pun melalui alamat : <http://covid19.go.id>. Pengelompokan zonasi ditentukan menggunakan skor dan pembobotan diperbaharui mingguan berdasarkan faktor kesehatan masyarakat. Namun, peng-*cluster* an ini dianggap masih kurang akurat karena tidak diketahui secara jelas proses perhitungan untuk mendapatkan *cluster-cluster* tersebut. Bahkan menurut (Farizi & Harmawan, 2020), dalam penelitiannya menyebutkan bahwa melalui *website* resmi Satgas Percepatan Penanganan Covid-19, pemerintah pusat terlihat tidak transparan dalam menampilkan data kepada publik termasuk proses untuk melakukan klasterisasi terhadap daerah-daerah yang ada di Indonesia.

Masing-masing pemerintahan daerah, terkhusus tingkat provinsi juga telah memiliki *website* sebagai pusat informasi Covid-19. Namun, pada laman *website* provinsi Aceh ini belum ada pengelompokan zona-zona dengan kasus Covid-19 tertinggi sampai terendah, padahal menurut data yang ada pada laman tersebut seluruh kabupaten/kota yang ada di Aceh telah terdampak Covid-19. Pengelompokan zona ini sangat dibutuhkan untuk masyarakat agar tetap waspada, melakukan *sosial distancing* dan mematuhi protokol dalam menjalankan aktivitas ataupun saat akan bepergian ke daerah lain yang masuk dalam pengelompokan tinggi kasus Covid-19. Sehingga diperlukan *website* yang mampu melakukan pengelompokan terhadap daerah-daerah dengan kasus Covid-19 tinggi, sedang, dan rendah. Dengan banyak nya data yang dibutuhkan untuk melakukan proses pengelompokan ini, dibutuhkan suatu teknik *clustering* dalam data mining.

Data mining sering dianggap sebagai bagian dari *knowledge discovery in database* (KDD), yaitu sebuah proses mencari pengetahuan yang bermanfaat dari data atau sekumpulan prosedur untuk menemukan nilai tambah berupa data yang sebelumnya tidak diketahui dari database. Pengetahuan yang diperoleh berasal dari penggalian dan pengenalan pola yang relevan atau menarik dalam data database. Klasterisasi adalah sebuah proses untuk pengelompokan data kedalam beberapa *cluster*. Sehingga dengan adanya klasterisasi ini di harapkan dapat dapat memberikan informasi yang bermanfaat untuk menjadi landasan-landasan dalam proses pengambilan keputusan sebagai upaya dari penekanan kasus Covid-19.

Beberapa peneliti terdahulu juga telah melakukan pengelompokan kasus Covid-19. Seperti yang dilakukan (Noviyanto, 2020) dengan judul “Penerapan Data Mining Dalam Mengelompokkan Jumlah Kematian Penderita Covid-19 Berdasarkan Negara di Benua Asia”. Pada penelitian ini hanya menggunakan parameter jumlah kematian akibat Covid-19 dari negara-negara di benua asia menggunakan algoritma K-Means dan software RapidMiner sehingga didapatkan hasil *clustering* jumlah kematian akibat Covid-19 adalah, 4 negara dengan *cluster* tinggi yaitu Turki, Iran, India, dan China. *Cluster* tingkat sedang sebanyak 4 negara, yaitu Pakistan, Jepang, Indonesia, dan Philipina. Kemudian *cluster* tingkat rendah adalah 41 negara lainnya. Adapun peneliti lainnya, (Solichin & Khairunnisa, 2020) dengan judul “Klasterisasi Persebaran Virus Corona (Covid-19) di DKI Jakarta Menggunakan Metode K-Means”. Dalam penelitian ini klasterisasi dilakukan berdasarkan parameter jumlah ODP, PDP, kasus positif, pasien sembuh, dan pasien meninggal dengan menggunakan metode K-Means dan metode pengukuran jarak *Euclidean*. Hasil dari penelitian ini adalah metode K-Means dapat digunakan dengan baik untuk klasterisasi kasus Covid-19 di DKI Jakarta dengan jumlah klaster yang paling baik adalah 9, namun untuk mengurangi risiko kekeliruan digunakan 3 klaster karena klaster 3 juga memiliki kualitas yang baik menurut perhitungan metode Elbow. Hasil klasterisasi Covid-19 per kabupaten di DKI Jakarta, terdapat 19 kecamatan pada klaster C1, 23 kecamatan pada klaster C2, dan 2 kecamatan pada klaster C3. Dalam penelitian ini juga menyarankan penelitian lanjutan menggunakan metode klasterisasi lain. Berdasarkan hasil penelitian (Muslimatin, 2011) dengan judul “Perbandingan Metode K-Means dan Metode Fuzzy C-Means (FCM) untuk *Clustering* Data (Studi Kasus pada Data Saham Harian PT. Astra, Tbk.)”. Penelitian ini mencoba untuk membandingkan hasil *clustering* metode K-Means dan Fuzzy C-Means berdasarkan parameter harga saham sebelum (*open*) dan harga saham tertinggi (*high*). Hasil dari penelitian ini didapatkan 5 kelompok harga saham, yaitu pada metode K-Means: (31947 32434), (24502 24999), (19266 20032), (14370 14776), (11547 11818) dan pada metode Fuzzy C-Means: (33133 33596), (29454 29954), (23037 23546), (15995 16466), (11958 12261) dan menyimpulkan bahwa metode

Fuzzy C-Means kemungkinan kegagalan konvergen lebih kecil dibandingkan metode K-Means. Hal ini dikarenakan pada metode K-Means setiap data dialokasikan secara tegas (*hard*).

Berdasarkan saran peneliti sebelumnya untuk mengembangkan penelitian menggunakan metode klasterisasi lainnya dan belum ada penelitian tentang persebaran Covid-19 di Aceh menggunakan metode Fuzzy C-Means. Maka dari itu, penulis membuat skripsi ini dengan judul **“Implementasi Fuzzy C-Means Untuk Klasterisasi Persebaran Covid-19 di Aceh Berbasis Web.”**

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka dapat dirumuskan permasalahan sebagai berikut :

1. Bagaimana merancang suatu sistem untuk klasterisasi persebaran Covid-19 di Provinsi Aceh?
2. Bagaimana menerapkan metode Fuzzy C-Means untuk klasterisasi persebaran Covid-19 di Provinsi Aceh?

1.3 Batasan Masalah

Agar pembahasan dan analisa masalah dalam penelitian ini jelas dan terarah, maka masalah dibatasi pada hal-hal berikut :

1. Sistem ini hanya digunakan untuk pengklasterisasian kasus Covid-19 di Provinsi Aceh yang terdiri dari 23 kabupaten/kota.
2. Sistem ini dibangun menggunakan bahasa pemrograman PHP dengan *database MySQL*.
3. Sistem ini menggunakan algoritma Fuzzy C-Means untuk proses klasterisasi.
4. Data diambil dari *website* Covid-19 Provinsi Aceh :
<https://covid19.acehprov.go.id/>
 Data yang diambil dari 23 kabupaten/kota yang ada di Aceh.
5. Parameter yang digunakan untuk melakukan klasterisasi persebaran Covid-19 di Aceh adalah:

- a. Jumlah terkonfirmasi
 - b. Jumlah pasien dalam perawatan
 - c. Jumlah pasien sembuh
 - d. Jumlah meninggal
 - e. Jumlah suspek, dan
 - f. Jumlah *probable*
6. Data yang telah diperoleh akan dikelompokkan kedalam 3 *cluster*, yaitu:
- a. *Cluster* 1 (zona merah dengan kasus Covid-19 yang tinggi)
 - b. *Cluster* 2 (zona kuning dengan kasus Covid-19 yang sedang), dan
 - c. *Cluster* 3 (zona hijau dengan kasus Covid-19 yang rendah).

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk :

- 1. Membangun sistem klasterisasi persebaran kasus Covid-19 di Provinsi Aceh menggunakan algoritma Fuzzy C-Means.
- 2. Mengetahui bagaimana penerapan algoritma Fuzzy C-Means dalam pengklasteran kasus Covid-19 di Provinsi Aceh.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah :

- 1. Untuk membangun sistem klasterisasi menggunakan metode Fuzzy C-Means, sehingga dapat menjadi landasan dalam pengambilan keputusan demi mengurangi kasus Covid-19 di Provinsi Aceh.
- 2. Untuk menentukan daerah mana yang masuk *cluster* 1 (zona merah dengan kasus Covid-19 yang tinggi), *cluster* 2 (zona kuning dengan kasus Covid-19 yang sedang), dan *cluster* 3 (zona hijau dengan kasus Covid-19 yang rendah).