

BAB I

PENDAHULUAN

1.1 Latar Belakang

Lingkungan adalah segala sesuatu yang ada di sekitar makhluk hidup memiliki hubungan timbal balik dan berpengaruh terhadap kelestarian lingkungannya, terdapat berbagai komponen lingkungan hidup antara lain manusia, hewan, tumbuhan, air, udara dan lain sebagainya. Makhluk hidup akan terus berinteraksi dengan lingkungan sekitar tempat tinggalnya (Welianto, 2020). Pengertian lingkungan kehidupan juga ditegaskan berdasarkan UU No 32 Tahun 2009, lingkungan adalah suatu kesatuan ruang dari semua benda, daya, kondisi, makhluk hidup, termasuk manusia dan makhluk hidup lainnya. Berdasarkan pengertian di atas, dapat disimpulkan secara sederhana bahwa lingkungan hidup merupakan suatu kesatuan yang mencakup segala aspek kehidupan. Kondisi lingkungan memiliki peranan yang sangat penting dan saling berhubungan satu sama lain dalam kelestarian ekosistem ini. Terciptanya lingkungan hidup yang berkualitas akan menjadikan manusia sehat. Ada banyak manfaat bagi lingkungan yang baik, ada beberapa contoh manfaat lingkungan hidup bagi manusia seperti udara untuk keperluan bernafas karena tidak ada manusia yang dapat bertahan hidup tanpa adanya bantuan udara, air untuk minum dan pembangkit tenaga listrik (Ulfiyah, 2013).

Menurut Anjelita M, dkk (2019) tingginya jumlah penduduk menyebabkan terjadinya peningkatan akan berbagai kebutuhan. Kebutuhan - kebutuhan ini mengakibatkan kerusakan – kerusakan lingkungan yang terjadi diberbagai belahan dunia termasuk indonesia. Dari data PBB pada hari kependudukan dunia yang jatuh pada tanggal 11 juli dan dengan populasi penduduk yang terus menerus meningkat setiap tahunnya. Data Perserikatan Bangsa – Bangsa (PBB) bahwa jumlah penduduk dunia ditahun 2017 tercatat 7,6 miliar dan akan meningkat menjadi 8,6 miliar pada tahun 2030, 9,8 miliar pada tahun 2050. Data The Spector Index ditahun 2018 dari 20 negara dengan penduduk terbanyak di urutan nomer empat di dunia berdasarkan data Worldometers,

Indonesia di tahun 2019 jumlah penduduk mencapai 269 juta jiwa atau 3,49% dari total populasi dunia hal tersebut menyebabkan kerusakan lingkungan yang terjadi di Indonesia. Salah satu kerusakan yang ditimbulkan adalah pencemaran air di lingkungan. Air merupakan bagian penting dalam kehidupan sehari-hari, tidak hanya bagi manusia tetapi juga bagi hewan dan tumbuhan, Berdasarkan data Badan Pusat Statistik (BPS) jumlah pencemaran lingkungan dari (BPS) menunjukkan masih banyak terjadi pencemaran lingkungan hidup pada tahun 2018 yang mencapai nilai tertinggi ada pada Provinsi Jawa Tengah dengan tingkat pencemaran air sebanyak 1900 pencemaran. Pencemaran tanah sebanyak 380 pencemaran dan pencemaran udara sebanyak 1336 pencemaran.

Pengelompokan wilayah perlu dilakukan untuk memberi masukan pemerintah untuk segera dapat menanggulangi pencemaran lingkungan, mengingat banyaknya dampak buruk yang dapat terjadi akibat pencemaran lingkungan. Ada beberapa algoritma pengelompokan yang dapat digunakan, salah satunya adalah *Fuzzy C-Means* (FCM). Analisa yang dilakukan dengan mengelompokan Provinsi di Indonesia berdasarkan Pencemaran Lingkungan Hidup. Pencemaran Lingkungan Hidup yang dinyatakan sebagai variabel dan Provinsi di Indonesia dinyatakan sebagai objek yang dikelompokan berdasarkan kemiripan karakteristiknya.

Pada penelitian sebelumnya yang dilakukan oleh Anjelita (2019) dengan studi kasus “Pengembangan Data Mining Klastering Pada Kasus Pencemaran Lingkungan Hidup” Data yang digunakan adalah data pencemaran lingkungan hidup pada tahun 2018, Peneliti menggunakan dua metode data mining clustering yaitu metode pengelompokan *k-means* dengan pengelompokan *medoid k-means* menggunakan *software rapidminer* versi 5.3. Pada penelitian tersebut peneliti tidak menggunakan optimasi *cluster* dalam melakukan *clustering* dan telah berhasil mengumpulkan data pencemaran lingkungan hidup pada tahun 2018 menggunakan algoritma *K-means clustering* menjadi 2 cluster, yaitu cluster 1 berjumlah 4 provinsi, dan 30 lainnya berada di cluster 2. Penelitian lainnya dilakukan oleh Saputra dan

Rikosakomara (2018) dengan studi kasus “Implementasi *Fuzzy C-Means* dan Model RFM untuk segmentasi Pelanggan (Studi Kasus: PT. XYZ)” menggunakan algoritma *Fuzzy C-Means* dan metode *Elbow* untuk optimasi *cluster*. Analisis *cluster* adalah teknik yang bertujuan untuk mengklasifikasikan objek berdasarkan karakteristiknya. Saat ini analisis *cluster* telah banyak diaplikasikan dalam berbagai bidang penelitian dan jurnal. Pada metode *cluster* konsep yang digunakan adalah pencarian titik pusat (*centroid*) *cluster* secara iteratif, titik pusat pada masing-masing *cluster* ditentukan berdasarkan jarak minimum setiap data pada *cluster*.

Pada analisis *Cluster*, metode yang sering digunakan adalah metode *K-Means*, algoritma ini memiliki kelebihan yaitu mudah untuk diimplementasikan dan mudah untuk diadaptasi, akan tetapi pada beberapa studi penelitian metode ini memiliki kelemahan yaitu, jika nilai random untuk inisialisasi kurang baik, maka pengelompokan yang dihasilkan kurang optimal. Bila titik data sampel hanya sedikit, maka cukup mudah untuk menghitung dan menemukan titik terdekat yang diinisialisasi secara acak ke titik tersebut. Namun, bila ada banyak titik data (misalnya, satu miliar keping data). Maka penghitungan dan pencarian titik terdekat akan memakan waktu lama. Proses ini dapat dipercepat, tetapi struktur data yang lebih kompleks seperti *kD-tree* atau *hashing* diperlukan.

Logika *fuzzy* adalah salah satu teori pengelompokkan data yang selisih data tidak terlalu jauh. Teori himpunan *fuzzy* lebih luas dari logika *fuzzy* dalam arti sempit dan memiliki lebih dari satu cabang. Di antara cabang-cabang tersebut adalah aritmatika *fuzzy*, topologi *fuzzy*, teori grafik *fuzzy*, dan analisis data *fuzzy*. *Fuzzy clustering* merupakan salah satu teknik yang dapat digunakan untuk menentukan *cluster* optimal dalam satu ruang vektor yang didasarkan pada bentuk normal Euclidian untuk jarak antara vector. Ada beberapa algoritma pengelompokan data, salah satu adalah *Fuzzy C-Means* (FCM). *Fuzzy C-Means* (FCM) adalah salah satu diantaranya mengclusterkan data dimana keberadaan setiap titik dalam suatu *cluster* ditentukan oleh derajat keanggotaan. (Kusuma Dewi, 2002:159).

Kelebihan *Fuzzy C-Means* adalah dapat mengelompokkan lebih dari satu variabel dalam satu waktu, berdasarkan keunggulan ini maka melakukan penelitian terhadap pencemaran lingkungan hidup di Indonesia ditahun 2018 dari data BPS karena jumlah pencemaran lingkungan hidup masih menunjukkan banyak terjadinya pencemaran lingkungan oleh karna itu penulis melakukan penelitian data pencemaran lingkungan dengan judul “ Clustering Kasus Pencemaran Lingkungan Hidup Di Indonesia Menggunakan Lagoritma *Fuzzy C-Means*”. Pada penelitian ini terdapat 4 atribut yang di gunakan yaitu, pencemaran air, pencemaran tanah, pencemaran udara, tidak ada pencemaran. Selain menggunakan *Fuzzy C-Means* penulis juga melakukan kombinasi menggunakan dengan metode *Elbow*. Metode ini adalah metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah *cluster* terbaik dengan melihat presentase hasil perbandingan antara jumlah *cluster* yang akan membentuk siku pada suatu titik.

1.2 Rumusan Masalah

Rumusan masalah yang dapat di simpulkan dari latar belakang diatas, yaitu:

1. Berapa jumlah *cluster* optimum yang didapatkan untuk pengelompokan wilayah provinsi di Indonesia berdasarkan data Pencemaran Lingkungan Hidup menggunakan algoritma *Fuzzy C-Means* ?
2. Berapa provinsi yang terdapat dalam masing - masing *cluster* optimum?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Data yang digunakan merupakan data dari Badan Pusat Statistik yaitu Pencemaran Lingkungan Hidup.
2. Data terdiri dari 34 provinsi berdasarkan tahun 2018.
3. Penentuan *cluster* optimum berdasarkan 2 *cluster* sampai 10 *cluster*.
4. Pengukuran *cluster* optimum menggunakan *Elbow*.
5. *Tools clustering* menggunakan *software* RStudio.

1.4 Tujuan Penelitian

Tujuan Penelitian yang ingin dicapai adalah:

1. Untuk menentukan jumlah *cluster* terbaik pada kasus Pencemaran Lingkungan Hidup setiap provinsi di Indonesia.
2. Untuk mengetahui kelompok provinsi di Indonesia yang terdapat dalam masing-masing *cluster* yang terbaik.

1.5 Manfaat Penelitian

Adapun manfaat penelitian sebagai berikut:

1. Menjadi alternatif metode pengelompokan pada data berdasarkan Pencemaran Lingkungan Hidup di Indonesia.
2. Menjadi referensi mengelompokan data menggunakan algoritma *Fuzzy C-Means* dengan metode *Elbow*.

