

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada 30 tahun lalu, ilmuwan sosial serta ahli biologi sudah mulai mencari bagaimana cara yang sistematis (algoritma) dalam menemukan kelompok yang ada dalam data. Pada tahun 1990, Kaufman dan Rousseeuw membuat analisis klaster (*cluster analysis*) yang praktis, sehingga mampu mengelompokkan data dengan bantuan teknologi komputer (Kaufman dan Rousseeuw, 2005).

Analisis klaster merupakan seni dalam menemukan suatu kelompok dalam data. Analisis klaster berbeda dengan analisis diskriminan dalam hal menetapkan kelompok. Analisis diskriminan meletakkan objek dalam kelompok yang telah didefinisikan sebelumnya, sedangkan analisis klaster meletakkan objek dalam kelompok yang belum didefinisikan sebelumnya (Kaufman dan Rousseeuw, 2005).

Analisis klaster mengelompokkan data dengan mempertimbangkan sebuah pendekatan yang berarti untuk mencari kesamaan dalam data dan membuat data yang sama menjadi suatu kelompok. Klasterisasi (*clustering*) bertujuan untuk menemukan cara yang paling alami untuk mengelompokkan sebuah dataset. Klasterisasi merupakan proses pengelompokan data yang memiliki kemiripan ke dalam suatu klaster dan menyajikan kepada pengguna dalam bentuk yang lebih ringkas (Garcia dkk., 2019). Secara umum, analisis klaster merupakan metode pengelompokan objek-objek sedemikian rupa sehingga objek-objek dalam suatu kelompok (klaster) serupa satu sama lain dan berbeda dengan objek-objek dalam kelompok lain. Sehingga, data dalam suatu klaster memiliki tingkat kemiripan yang besar sedangkan data antar-klaster memiliki tingkat kemiripan yang kecil.

Secara garis besar analisis klaster terdiri atas analisis hierarki (*hierarchial clustering*) dan analisis partisi (*partional clustering*). Analisis hirarki merupakan suatu cara pengelompokan data menggunakan suatu bagan hierarki yang disebut dendogram. Dendogram secara iteratif membagi data menjadi himpunan yang lebih kecil sampai

setiap himpunan bagian hanya terdiri satu objek, proses ini dilakukan secara bertahap (Kaufman dan Rousseuw, 2005).

Partional clustering mengelompokkan data dengan membuat partisi di mana setiap kluster memiliki pusat gravitasi (*centroid*). Algoritma partisi ini dimulai dengan partisi awal dan dilakukan iterasi agar mendapatkan kluster yang optimal. Input parameter dalam algoritma ini adalah k yang menyatakan jumlah kluster yang dibentuk. Perbedaan antara kedua metode tersebut sangat nampak pada metode partisi banyaknya kluster ditentukan terlebih dahulu sedangkan pada hirarki jumlah kluster tidak ditentukan terlebih dahulu (Kaufman dan Rousseuw, 2005).

Jain (1988) telah mengembangkan algoritma *clustering* dengan pendekatan berdasarkan kepekatan (*density-based approach*). Algoritma *density based-approach* mengelompokkan data ke dalam beberapa sel. Selain itu, histogram juga dibangun untuk menggambarkan jumlah anggota pada setiap sel. Sel dengan jumlah anggota yang relatif tinggi merupakan pusat kluster. Algoritma *density based-approach* mampu mengidentifikasi kluster dalam bentuk apapun. Namun, algoritma tersebut memerlukan ruang penyimpanan yang besar dan membutuhkan waktu pemrosesan yang lama, sehingga Ester dkk. (1996) mengusulkan algoritma DBSCAN untuk mengatasi kelemahan tersebut.

DBSCAN merupakan suatu metode *clustering* yang menentukan kluster berdasarkan kepadatan di mana kepadatan di dalam kluster akan lebih besar dibandingkan kepadatan di luar kluster. DBSCAN menggunakan dua parameter sebagai *input* yaitu “epsilon” (jarak maksimum dalam suatu kluster dan “minPts” (jumlah titik minimum dalam suatu kluster). Sehingga, suatu kluster setidaknya memiliki anggota sebanyak “minPts” dalam radius “epsilon” (Schubert dkk., 2017).

Konsep kepadatan DBSCAN menghasilkan tiga macam jenis dari setiap datanya, yaitu inti (*core*), batas (*border*) dan *noise*. Sebuah data akan dikatakan inti jika jumlah data tetangga dan dirinya sendiri berada pada radius “epsilon” berjumlah lebih besar sama dengan minPts. Nilai radius “epsilon” dan “minPts” ini ditetapkan secara sendiri. Jika datayang jumlah tetangga dan dirinya sendiri dalam radius “epsilon” kurang dari “minPts”, tetapi tetangganya menjadi inti karena kehadirannya maka data tersebut

Pitria Sari, 2022

PENGELOMPOKAN PROVINSI DI INDONESIA PERIODE 2020/2021 DENGAN METODE DENSITY BASED SPATIAL CLUSTERING APPLICATION WITH NOISE (DBSCAN)

(Studi Kasus: Data Perekonomian Provinsi di Indonesia Periode 2020/2021)

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

dikatakan sebagai batas. Apabila jumlah tetangga dan dirinya dalam radius “epsilon” kurang dari minPts tetapi tidak ada tetangga yang menjadi inti karena ada dirinya, maka data tersebut dikatakan sebagai *noise* (Schubert dkk., 2017).

Ada beberapa keuntungan menggunakan DBSCAN, yaitu tidak memerlukan informasi tentang banyaknya kelompok yang akan dibentuk, dapat memperoleh kelompok dalam bentuk sembarang, serta dapat menemukan kelompok yang benar-benar dikelilingi oleh (tetapi tidak terhubung) pada klaster yang berbeda. Kemudian DBSCAN juga mempunyai informasi tentang data *outlier* (*noise*), dan juga hanya membutuhkan dua parameter yang sebagian besar tidak sensitif terhadap urutan titik dalam basis data (Mumtaz dkk., 2010).

White (2002) menjelaskan bahwa kondisi setiap kelompok maupun individu paska krisis moneter dapat berbeda satu sama lainnya. Hal ini didorong oleh perbedaan lokasi, kelompok penduduk berdasarkan umur dan jenis kelamin, kelompok sosial, sektor ekonomi, dan etnis. Tambunan (2005) mendukung bahwa kebijakan ekonomi khususnya terkait Usaha Kecil Menengah (UKM) dalam penanggulangan paska krisis moneter dapat diterapkan dengan bantuan analisis klaster. Hal ini dikarenakan analisis klaster dapat menentukan kelompok data dengan karakteristik tertentu sehingga dapat digunakan sebagai dasar pengambilan keputusan atau pemecahan masalah.

Salah satu permasalahan yang melanda Indonesia saat ini ialah pandemi covid-19. Pandemi covid-19 yang melanda Indonesia sejak awal Maret 2020 menyebabkan kasus pasien positif semakin meningkat pesat. Banyak usaha yang telah ditempuh pemerintah untuk mengurangi tingkat penyebaran virus corona atau disebut dengan istilah “*flatten the curve*” (Steven, 2020). Salah satu kebijakan yang dilakukan pemerintah ialah menerapkan Pembatasan Sosial Berskala Besar (PSBB) yang dikeluarkan pemerintah sejak tanggal 31 Maret 2020. Melalui Peraturan Menteri Kesehatan No. 9 Tahun 2020 tentang Pedoman Percepatan Penangan Covid-19, ditetapkan beberapa kebijakan PSBB di antaranya: 1) Peliburan sekolah dan tempat kerja; 2) Pembatasan kegiatan keagamaan; 3) Pembatasan kegiatan di tempat umum; 4) Pembatasan kegiatan sosial budaya; 5) Pembatasan moda transportasi; 6) Pembatasan kegiatan lainnya terkait pertahanan dan keamanan (Kementrian Luar Negeri, 2020).

Pitria Sari, 2022

PENGELOMPOKAN PROVINSI DI INDONESIA PERIODE 2020/2021 DENGAN METODE DENSITY BASED SPATIAL CLUSTERING APPLICATION WITH NOISE (DBSCAN)

(Studi Kasus: Data Perekonomian Provinsi di Indonesia Periode 2020/2021)

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

Kebijakan yang diterapkan membuat perekonomian di Indonesia terdampak di berbagai bidang, terutama bidang perekonomian. Penerapan kebijakan PSBB membuat perekonomian di Indonesia semakin menurun, Deputy Bidang Ekonomi Kementerian PPN/Bappenas memperkirakan Indonesia akan sulit untuk kembali ke lintasan ekonomi sebelum pandemi covid-19 jika tidak melakukan perubahan yang mendasar. Perekonomian negara tidak dapat terus menerus berada di posisi ini, sehingga perlu adanya tindakan pemulihan agar roda perekonomian di Indonesia berputar sebagaimana mestinya (BNPB, 2020). Tindakan pemulihan seperti pemberian bantuan sosial harus dilakukan dengan perencanaan yang tepat. Sehingga diperlukan pengelompokan wilayah agar diberikan tindakan pemulihan yang sesuai untuk setiap kelompok.

Pengelompokan wilayah dapat dilakukan menggunakan analisis kluster. Penelitian yang pernah dilakukan oleh Pusdiktasari, dkk., (2021) tentang Pengelompokan Provinsi di Indonesia dengan Ekonomi Terdampak Covid-19 Menggunakan analisis kluster dan telah menggunakan beberapa metode *linkage*, dan mendapatkan hasil bahwa metode *Average Linkage* memiliki pengelompokan yang teliti dan memiliki nilai koefisien korelasi *cophenetic* yang lebih tinggi dibandingkan metode lainnya.

Berdasarkan latar belakang yang telah diuraikan tersebut, peneliti tertarik untuk mengelompokkan provinsi di Indonesia yang terdampak pandemi dengan menggunakan metode DBSCAN berdasarkan aspek ekonomi dan aspek sosial ekonomi setiap provinsinya. Aspek ekonomi diukur dengan indikator Produksi Pertanian Padi dan Produk Domestik Regional Bruto (PDRB) perkapita. Aspek sosial ekonomi diukur dengan indikator rata-rata upah perbulan, persentase penduduk yang bekerja di sektor formal (Babenco, dkk., 2021). Metode DBSCAN dipilih karena algoritma DBSCAN bisa mengatasi data yang *noise*, random dan tidak linear, sesuai dengan karakteristik data perekonomian yang datanya cenderung random, tidak mempunyai pola tertentu dan ada data *outlier(noise)*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dipaparkan, maka rumusan masalah dalam penelitian ini adalah bagaimana hasil pengelompokan provinsi di Indonesia dengan menggunakan metode DBSCAN pada kasus perekonomian akibat pandemi Covid-19?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian ini adalah memperoleh hasil pengelompokan provinsi di Indonesia dengan menggunakan metode DBSCAN pada kasus perekonomian akibat pandemi Covid-19.

1.4 Manfaat Penelitian

Hasil penelitian ini bermanfaat secara teoritis dan praktis, sebagai berikut:

1.4.1 Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan pengetahuan dan referensi baru mengenai implementasi algoritma DBSCAN untuk mengetahui pengelompokan provinsi di Indonesia dengan menggunakan algoritma DBSCAN periode 2020/2021. Selain itu, dengan adanya penelitian ini diharapkan dapat menjadi masukan untuk pengembangan keilmuan bahan kajian lainnya.

1.4.2 Manfaat Praktis

- Bagi peneliti, penelitian ini dibuat supaya dapat menyelesaikan skripsi ini sebagai syarat memperoleh gelar sarjana matematika dan menerapkan materi yang sudah dipelajari selama perkuliahan.
- Bagi Pemerintah, pemerintah dapat menentukan prioritas wilayah yang memerlukan bantuan agar lebih cepat memulihkan perekonomian akibat pandemi Covid-19.
- Bagi pembaca, pembaca dapat menambah wawasan tentang metode DBSCAN dan perekonomian setiap provinsi di Indonesia periode 2020/2021.

1.5 Sistematika Penulisan

Penulisan skripsi ini terdiri atas lima bab yaitu:

1. BAB I: Pendahuluan

Pada bab ini peneliti menguraikan latar belakang yang mendasari perlunya penelitian ini, rumusan masalah, tujuan penelitian, manfaat penelitian secara teoritis dan praktis, serta sistematika penulisan dalam penelitian ini.

2. BAB II: Kajian Pustaka

Pada bab ini peneliti menguraikan kajian pustaka yang mendasari penelitian terkait dengan pengelompokan provinsi di Indonesia periode 2020/2021 dengan metode *density based clustering application with noise* (DBSCAN), seperti *clustering*, DBSCAN, algoritma DBSCAN, ukuran jarak, indeks validitas dan kerangka berpikir.

3. BAB III: Metodologi Penelitian

3.1 Jenis dan Sumber Data

Pada subbab ini dijelaskan mengenai jenis data dan teknik pengumpulan data yang akan digunakan. Data yang digunakan dalam penelitian ini adalah data yang diperoleh dari Badan Pusat Statistik (BPS) yang berjudul “Perkembangan Beberapa Indikator Utama Sosial-Ekonomi Indonesia” pada periode 2020/2021.

3.2 Pendekatan Penelitian dan Variabel Penelitian

Pada subbab ini dijelaskan pendekatan penelitian yang digunakan pada penelitian ini, serta mengetahui variabel bebas dan variabel terikat yang digunakan dan mengenai variabel apa saja yang digunakan dalam penelitian ini.

3.3 Metode Analisis

Pada subbab ini dijelaskan metode analisis data dari mulai pengambilan data hingga visualisasi hasil pengklasteran.

3.4 Tahapan Penelitian

Pada subbab ini dijelaskan mengenai semua tahapan yang digunakan untuk mengelompokkan provinsi menggunakan DBSCAN.

4. BAB IV: Studi Kasus

Pada bab ini peneliti menguraikan tentang hasil dan pembahasan dari pengelompokan provinsi di Indonesia periode 2020/2021 yang didapatkan dengan menggunakan metode DBSCAN.

5. BAB V: Penutup

Pada bab ini berisikan mengenai kesimpulan dari penelitian yang telah didapatkan dari penerapan metode DBSCAN pada kasus pengelompokan provinsi di Indonesia periode 2020/2021 dan saran untuk pengembangan penelitian selanjutnya.