

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Cluster

Analisis *cluster* merupakan salah satu teknik data mining yang bertujuan untuk mengidentifikasi sekelompok obyek yang mempunyai kemiripan karakteristik tertentu yang dapat dipisahkan dengan kelompok obyek lainnya. Obyek yang berada dalam kelompok yang sama relatif lebih homogen daripada obyek yang berada pada kelompok yang berbeda. Tujuan utama dari proses ini adalah untuk pengelompokan sejumlah data/obyek ke dalam sebuah *cluster* sehingga dalam setiap *cluster* akan berisi data yang semirip mungkin (Nurjanah, dkk, 2014).

Pada analisis *cluster* memungkinkan untuk melakukan pengelompokan pada data yang besar sehingga penggunaannya menjadi lebih praktis (Kucherov & Kurenkov, 2017). Baskoro (2010) menyatakan bahwa *Clustering* adalah salah satu alat bantu pada data mining yang bertujuan mengelompokkan objek-objek ke dalam *cluster-cluster*. Objek akan dikelompokkan ke dalam satu atau lebih *cluster* sehingga objek-objek yang berada dalam satu *cluster* akan mempunyai kesamaan yang tinggi antara satu dengan lainnya. Objek-objek dikelompokkan berdasarkan prinsip memaksimalkan kesamaan objek pada *cluster* yang sama dan memaksimalkan ketidaksamaan pada *cluster* yang berbeda.

2.2 Fuzzy Geographically Weighted Clustering (FGWC)

Analisis *Cluster Fuzzy Geographically Weighted Clustering* (FGWC) merupakan salah satu analisis yang dapat mengatasi kelemahan metode *clustering* yang lainnya, yaitu metode *Fuzzy C-Means*. Analisis *Fuzzy Geographically Weighted Clustering* pertama kali diperkenalkan oleh (Mason & Jacobson, 2006). FGWC merupakan perbaikan dari algoritma *Fuzzy Geodemographics* yang diusulkan oleh (Feng dan Flowerdew, 1998). Pengaruh satu area terhadap area yang lainnya dianggap oleh FGWC sebagai produk dari populasi pada daerah tersebut (Wijayanto & Purwarianti, 2014).

Pada setiap iterasi di pengelompokkan *fuzzy clustering*, penentuan keanggotaan kelompok menggunakan rumus :

$$\mu'_i = \alpha\mu_i + \beta \frac{1}{A} \sum_j^n w_{ij} \mu_j \quad (2.1)$$

dengan :

μ'_i : Nilai keanggotaan baru dari objek i

μ_i : Nilai keanggotaan lama dari objek i

w_{ij} : Ukuran penimbang sejumlah interaksi antar wilayah

A : Nilai untuk memastikan nilai penimbang tidak lebih dari 1

α dan β merupakan faktor pengali untuk nilai keanggotaan yang lama dan nilai penimbang dari rerata keanggotaan unit observasi lain. Nilai α dan β didefinisikan sebagai berikut.

$$\alpha + \beta = 1$$

Penimbang keanggotaan (w_{ij}) didefinisikan sebagai berikut.

$$w_{ij} = \frac{(m_i m_j)^2}{d_{ij}^\alpha} \quad (2.2)$$

dengan:

m_i : Jumlah populasi dari wilayah i

m_j : Jumlah populasi dari wilayah j

d_{ij} : Jarak antara wilayah i dan wilayah j

α dan β merupakan sebuah parameter yang ditentukan oleh pengguna (*user definable parameter*). Jika diasumsikan bahwa interaksi spasial memiliki dampak yang sama seperti fitur demografi perilaku masyarakat, maka :

$$\alpha = \beta = \frac{1}{2}$$

FGWC merupakan integrasi dari metode *fuzzy clustering* klasik dan unsur geo-demografi. FGWC memasukkan unsur geografis dalam analisisnya sehingga cluster yang terbentuk akan sensitif terhadap efek lingkungan dan berpengaruh pada nilai-nilai pusat *cluster* untuk menciptakan *cluster* yang “*geographically aware*”. FGWC merupakan perbaikan dari algoritma *Fuzzy Geodemographics*. Metode *Fuzzy Geodemographics* masih memiliki beberapa keterbatasan, antara lain : *Neighborhood Effect (NE)* mengabaikan efek daerah yang tidak memiliki batas umum dan efek populasi yang merupakan sebuah pertimbangan kunci dari analisis geodemografis (Mason & Jacobson, 2006).

2.3 Fungsi Objektif FGWC

Metode FGWC dapat mengatasi kelemahan dari salah satu metode *clustering* yaitu metode *Fuzzy C-Means*. Algoritma *Fuzzy C-Means* masih memiliki kelemahan dalam tahap inialisasinya. Untuk mengatasi kelemahan dan keterbatasan dalam algoritma FCM, algoritma FGWC digunakan untuk menentukan *cluster* yang memiliki efek geografis didalamnya.

Fungsi objektif dari FGWC adalah :

$$J_{FGWC}(U, V; X) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|v_i - x_k\|^2 \quad (2.3)$$

dengan :

U : Matriks keanggotaan

V : Matriks untuk pusat cluster

X : Matriks untuk data

v_i : Pusat cluster untuk objek i

u_{ik} : Elemen dari matriks keanggotaan

x_k : Titik data

m : *Fuzziness*, parameter yang digunakan untuk mengukur tingkat kesamaran dari hasil pengelompokkan, $m > 1$

Adapun pusat *cluster* dapat didefinisikan sebagai berikut :

$$V_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m} \quad (2.4)$$

Mendefinisikan matriks keanggotaan juga diperlukan sebelum melakukan algoritma FGWC. Matriks keanggotaan dapat dihitung menggunakan rumus :

$$U_{ik} = \frac{1}{\sum_{j=1}^c \left(\frac{\|v_i - x_k\|}{\|v_j - x_k\|} \right)^{\frac{2}{m-1}}} \quad (2.5)$$

2.4 Indeks Validitas

Pada konsep *fuzzy clustering*, suatu anggota dapat menjadi anggota beberapa *cluster* sekaligus menurut derajat keanggotaannya. Dalam proses *clustering* selalu mencari solusi terbaik untuk parameter yang didefinisikan. Akan tetapi, dalam beberapa hal terdapat *cluster* yang tidak sesuai dengan data. Untuk menentukan jumlah *cluster* yang optimal maka perlu adanya pengukuran indeks validitas. Adapun indeks validitas yang sering digunakan untuk *fuzzy clustering* yaitu *IFV Index*.

Penggunaan Indeks IFV telah dilakukan seperti pada penelitian (Hadi, 2017) dan (Putra, 2016). Indeks IFV sering digunakan untuk memvalidasi pengelompokan *fuzzy* dengan data spasial, karena sifatnya yang kuat dan stabil (Chunchun dkk., 2008). Ketika nilai IFV maksimum maka kualitas *cluster* semakin baik. Ukuran persamaannya diuraikan sebagai berikut :

$$IFV = \frac{1}{c} \sum_j \left\{ \frac{1}{N} \sum_{k=1}^N \mu_{kj}^2 \left[\log_2 c - \frac{1}{N} \sum_{k=1}^N \log_2 \mu_{kj} \right]^2 \right\} \frac{SD_{max}}{\sigma_D} \quad (2.9)$$

Jarak maksimum antara pusat *cluster* diuraikan sebagai berikut :

$$SD_{max} = \max_{k \neq j} \|V_k - V_j\|^2$$

dimana :

μ_{ij} : keanggotaan titik data j di dalam *cluster* i

N : Jumlah titik data (*records*)

c : Jumlah *cluster*

V_k : pusat *cluster* ke- k

2.5 Indikator Indeks Pembangunan Manusia

Konsep pembangunan manusia seutuhnya merupakan konsep yang menghendaki peningkatan kualitas hidup penduduk baik secara fisik, mental maupun secara spiritual. Bahkan secara eksplisit disebutkan bahwa pembangunan yang dilakukan menitikberatkan pada pembangunan sumber daya manusia yang seiring dengan pertumbuhan ekonomi. Pembangunan sumber daya manusia secara fisik dan mental mengandung makna peningkatan kapasitas dasar penduduk yang kemudian akan memperbesar kesempatan untuk dapat berpartisipasi dalam proses pembangunan yang berkelanjutan (BPS, 2014).

Lembaga UNDP telah mempublikasikan laporan pembangunan sumber daya manusia dalam ukuran kuantitatif yang disebut *Human Development Index* (HDI) atau dalam Bahasa Indonesia adalah Indeks Pembangunan Manusia (IPM) yang diperkenalkan pertama sekali pada tahun 1990. Meskipun IPM merupakan alat ukur pembangunan Sumber Daya Manusia (SDM) yang dirumuskan secara konstan, diakui tidak akan pernah menangkap gambaran pembangunan SDM secara sempurna (BPS, 2014). Adapun indikator yang dipilih untuk mengukur dimensi

Indeks Pembangunan Manusia adalah sebagai berikut (UNDP, *Human Development Report* 1993):

- a) *Longevity*, diukur dengan variabel harapan hidup saat lahir atau *life expectancy of birth* dan angka kematian bayi per seribu penduduk atau *infant mortality rate*.
- b) *Educational Achievement*, diukur dengan dua indikator, yakni harapan lama sekolah bagi penduduk usia 7 tahun ke atas (*expected years of schooling*) dan tahun rata-rata bersekolah bagi penduduk usia 25 tahun ke atas (*the mean years of schooling*).
- c) *Access to resource*, dapat diukur secara makro melalui PDB riil perkapita dengan terminologi *purchasing power parity* dan dapat dilengkapi dengan tingkatan penduduk usia kerja.

2.5.1 Angka Harapan Hidup

Angka Harapan Hidup saat Lahir (AHH) didefinisikan sebagai rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang sejak lahir. AHH mencerminkan derajat kesehatan suatu masyarakat. Angka Harapan Hidup merupakan alat untuk mengevaluasi kinerja pemerintah dalam meningkatkan kesejahteraan penduduk pada umumnya, dan meningkatkan derajat kesehatan pada khususnya. Angka Harapan Hidup yang rendah di suatu daerah harus diikuti dengan program pembangunan kesehatan, dan program sosial lainnya termasuk kesehatan lingkungan, kecukupan gizi dan kalori termasuk program pemberantasan kemiskinan.

2.5.2 Angka Kematian Bayi

Angka Kematian Bayi (AKB) adalah banyaknya kematian bayi yang belum berusia tepat satu tahun per 1000 kelahiran hidup dalam satu tahun. Angka kematian bayi merupakan indikator yang penting untuk mencerminkan keadaan derajat kesehatan di suatu masyarakat. Karena bayi yang baru lahir sangat sensitif terhadap keadaan lingkungan tempat tinggal bayi tersebut. Untuk dapat menurunkan angka kematian bayi akan lebih mudah tercapai apabila ditunjang dengan sumber daya manusia yang berkualitas dan ketersediaan standar, pedoman, sistem pencatatan pelaporan serta logistik yang memadai dan bermutu.

2.5.3 Harapan Lama Sekolah

Harapan lama sekolah didefinisikan sebagai lamanya sekolah (dalam tahun) yang diharapkan akan dirasakan oleh anak pada umur tertentu di masa mendatang. Harapan lama sekolah dihitung untuk penduduk berusia 7 tahun ke atas. Indikator ini dapat digunakan untuk mengetahui kondisi pembangunan sistem pendidikan di berbagai jenjang yang ditunjukkan dalam bentuk lamanya pendidikan (dalam tahun) yang diharapkan dapat dicapai oleh setiap anak. Harapan lama sekolah menggunakan batasan yang dipakai sesuai kesepakatan UNDP. Batas maksimum untuk harapan lama sekolah adalah 18 tahun, sedangkan batas minimumnya 0 (nol).

2.5.4 Rata-Rata Lama Sekolah

Rata-rata lama sekolah menggambarkan jumlah tahun yang digunakan oleh penduduk usia 25 tahun ke atas dalam menjalani pendidikan formal. Penghitungan rata-rata lama sekolah menggunakan dua batasan yang dipakai sesuai kesepakatan UNDP. Rata-rata lama sekolah memiliki batas maksimumnya 15 tahun dan batas minimum sebesar 0 tahun. Rata-rata lama sekolah mengindikasikan makin tingginya pendidikan yang dicapai oleh masyarakat di suatu daerah. Semakin tinggi rata-rata lama sekolah berarti semakin tinggi jenjang pendidikan yang dijalani.

Asumsi yang berlaku secara umum bahwa semakin tinggi tingkat pendidikan seseorang maka semakin tinggi pula kualitas seseorang, baik pola pikir maupun pola tindaknya. Atmanti (2005), mengemukakan bahwa orang yang memiliki tingkat pendidikan yang lebih tinggi, diukur dengan lamanya waktu untuk sekolah akan memiliki pekerjaan dan upah yang lebih baik dibandingkan dengan orang yang pendidikannya lebih rendah.

2.5.5 Paritas Daya Beli

Pengukuran daya beli masyarakat kabupaten/kota ataupun provinsi, menggunakan rata-rata konsumsi yang dianggap paling dominan dari hasil Survei Sosial Ekonomi Nasional (SUSENAS) yang telah di standarkan agar dapat digunakan sebagai perbandingan antar daerah dan waktu sesuai indeks daya beli (*Purchasing Power Parity – PPP*). Terdapat 96 komoditi yang dipilih, terdiri dari 66 komoditi adalah jenis makanan sedangkan 30

komoditi lainya adalah jenis non makanan. Rata-rata pengeluaran per kapita dibuat konstan/riil dengan tahun dasar 2012=100. Perhitungan paritas daya beli (PPP) menggunakan metode Rao.

Teori Paritas Daya Beli diperkenalkan oleh Gustav Cassel pada tahun 1918 yang menjelaskan hubungan antara harga komoditi dalam mata uang domestik (lokal) dengan dengan nilai tukar.. Teori ini menyatakan bahwa nilai tukar akan menyesuaikan diri dari waktu ke waktu untuk mencerminkan selisih inflasi antara dua negara, akibat adanya daya beli konsumen untuk membeli produk domestik akan sama dengan daya beli untuk membeli produk luar negeri (Madura, 2000).

2.5.6 Tingkat Pengangguran Terbuka

Menurut BPS Indonesia tingkat pengangguran terbuka (TPT) adalah persentase jumlah pengangguran terhadap jumlah angkatan kerja. Tingkat pengangguran dapat di ukur dari banyaknya jumlah orang yang menganggur. Semakin banyak jumlah pengangguran maka akan mengakibatkan permasalahan sosial, diantaranya semakin banyak pengangguran maka semakin minim pendapatan masyarakat tersebut.

Menurut Agus (2010) pengangguran adalah golongan angkatan kerja yang secara aktif mencari pekerjaan. Pada suatu pekerjaan tertentu sedangkan masih belum mendapatkan pekerjaan. Ataupun sedang dalam kondisi dalam mencari pekerjaan yang terus berusaha tetapi tidak mendapatkan pekerjaan.