

BAB III

LANDASAN TEORI

Pada tahap ini, penulis akan mendeskripsikan tentang teori apa saja yang akan digunakan sebagai acuan untuk menganalisis data rumah tangga miskin menggunakan metode *clustering* dengan algoritma *Fuzzy C-Means*. Adapun pembahasannya meliputi komponen-komponen yang nantinya disertakan pada penelitian ini antara lain kemiskinan, rumah tangga, metode *Fuzzy C-Means*, *clustering*, data mining dan bahasa pemrograman R.

3.1. Kemiskinan

3.1.1. Pengertian Kemiskinan

Umumnya pengertian kata “miskin” adalah “serba ketidakcukupan”, sedangkan pengertian “kemiskinan” adalah sebuah kondisi di mana selalu berkekurangan baik itu dalam bentuk fisik maupun non-fisik. Kemiskinan dalam arti yang sempit adalah sebuah kondisi serba kekurangan untuk menjamin kelangsungan hidup. Kemiskinan dalam arti yang luas adalah suatu rancangan yang terdiri dari lima komponen yaitu kemiskinan, tidak berdaya, kerentanan mengalami kondisi darurat, saling ketergantungan, dan keterasingan baik secara geografis maupun sosiologis [14].

Di dalam persoalan kemiskinan rendahnya penghasilan sering terjadi sehingga jika dipandang lebih detail lagi kemiskinan merupakan suatu keadaan di

mana tidak dipenuhinya kebutuhan pokok masyarakat, yaitu pangan, papan, dan sandang. Salah satunya yaitu pengertian kemiskinan menurut Badan Pusat Statistik, di mana dijelaskan bahwa kemiskinan ialah ketidakmampuan seseorang untuk memenuhi kebutuhan pokok masyarakat untuk kehidupan yang cukup [15].

3.1.2. Penyebab Kemiskinan

Penyebab kemiskinan terbagi dalam beberapa pengertian, yaitu [15]:

1. Kemiskinan yang diakibatkan oleh globalisasi di mana globalisasi itu dapat memberikan dua pengertian negara yaitu negara pemenang dan negara kalah. Yang disebut negara pemenang adalah negara-negara maju, sedangkan yang disebut negara-negara berkembang adalah negara yang sering kali disudutkan oleh persaingan dan pasar bebas yang merupakan salah satu syarat dalam kemiskinan yang diakibatkan oleh globalisasi. Karena pada umumnya negara-negara berkembang itu tersudutkan dan terpinggirkan, maka jumlah kemiskinan di negara-negara berkembang sangat besar dibandingkan negara-negara maju contohnya adalah Indonesia.
2. Kemiskinan yang diakibatkan oleh pembangunan. Pada model pembangunan ini ada beberapa bentuk kemiskinan yang terbangun, seperti **kemiskinan pedesaan** di mana keadaan penduduk desa yang menghadapi permasalahan kemiskinan yang diakibatkan oleh sistem pengembangan yang membuat masyarakat pedesaan merasa terpinggirkan; **kemiskinan perkotaan** di mana permasalahan kemiskinan yang diakibatkan oleh hakikat serta kelajuan pertumbuhan ekonomi sehingga beberapa kelompok tidak mendapatkan laba.

3. Kemiskinan yang berkaitan dengan sosial di mana kondisi ini dilihat dari aktivitas sosial masyarakat yang tidak menguntungkan dalam kehidupan bermasyarakat. Contoh kemiskinan yang diakibatkan oleh kondisi sosial yang tidak menguntungkan adalah kemiskinan yang dialami oleh perempuan, anak-anak dan kelompok minoritas. Keadaan sosial yang dimaksud seperti bias gender, diskriminasi, atau eksploitasi ekonomi.
4. Kemiskinan yang berkaitan dengan konsekuensial. Di mana pada bagian ini biasanya yang menjadi penyebab terjadinya kemiskinan adalah faktor-faktor eksternal atau faktor dari luar. Yang dimaksud dari faktor eksternal yang menimbulkan kemiskinan di masyarakat yaitu konflik, bencana alam, kerusakan lingkungan, dan tingginya jumlah penduduk.

3.1.3. Garis Kemiskinan

Dalam pengukuran sebuah kemiskinan, Badan Pusat Statistik (BPS) menggunakan konsep kemampuan memenuhi kebutuhan dasar (basic needs approach). Berdasarkan pendekatan yang dipakai ini, kemiskinan dilihat sebagai ketidakmampuan dalam bidang ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran. Metode yang digunakan adalah dengan menghitung Garis Kemiskinan (GK) [6]. Menurut Badan Pusat Statistik, Garis Kemiskinan (GK) mencerminkan nilai rupiah pengeluaran minimum yang diperlukan seseorang untuk memenuhi kebutuhan pokok hidupnya selama sebulan, baik kebutuhan makanan maupun non-makanan [1].

Garis Kemiskinan terdiri dari Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non-Makanan (GKNM). Garis Kemiskinan Makanan (GKM) merupakan nilai pengeluaran minimum untuk kebutuhan makanan yang disetarakan dengan 2100 kilokalori per kapita per hari. Paket komoditi kebutuhan dasar makanan diwakili oleh 52 jenis komoditi (padi-padian, umbi-umbian, ikan, daging, telur dan susu, sayuran, kacang-kacangan, buah-buahan, minyak dan lemak, dll). Garis Kemiskinan Non-Makanan (GKNM) merupakan nilai pengeluaran minimum untuk kebutuhan non-makanan berupa perumahan, sandang, pendidikan dan kesehatan. Paket komoditi kebutuhan dasar non-makanan diwakili oleh 51 jenis komoditi di perkotaan dan 47 jenis komoditi di perdesaan [1].

Rumus Penghitungan :

$$GK = GKM + GKNM \quad (2.1)$$

Keterangan :

GK = Garis Kemiskinan

GKM = Garis Kemiskinan Makanan

GKNM = Garis Kemiskinan Non Makan

3.1.4. Kriteria Pengukuran Kemiskinan

Kriteria kemiskinan berdasarkan Badan Pusat Statistik yaitu [16] :

Tabel 3.1. Tabel Variabel dan Kriteria Kemiskinan

No.	Variabel Kemiskinan	Kriteria Kemiskinan
-----	---------------------	---------------------

1	Luas lantai bangunan tempat tinggal	< 8m ² per kapita
2	Jenis lantai yang digunakan dalam tempat tinggal	Jenisnya : tanah, bambu, atau kayu murahan
3	Jenis bangunan yaitu dinding pada tempat tinggal	Jenisnya : bambu, rumbia, kayu berkualitas rendah, atau tembok tanpa plester
4	Fasilitas yang digunakan untuk tempat buang air besar	Jenisnya : tidak memiliki / dipakai bersama rumah tangga lain
5	Sumber penerangan yang digunakan dalam rumah tangga	Jenisnya : bukan listrik
6	Sumber air minum yang digunakan dalam rumah tangga	Jenisnya : mata air tidak terlindungi, sungai, atau air hujan
7	Bahan bakar yang digunakan untuk memasak sehari-hari	Jenisnya : kayu bakar, arang, atau minyak tanah
8	Mengonsumsi daging/ayam dalam seminggu	Jenisnya : tidak pernah, atau satu kali dalam seminggu
9	Pembelian pakaian baru setiap anggota rumah tangga dalam setahun	Jenisnya : tidak pernah membeli, atau satu/dua kali setahun
10	Ukuran makan dalam sehari untuk setiap anggota rumah tangga	Jenisnya : satu kali, atau dua kali makan dalam sehari
11	Kesanggupan dalam membayar untuk berobat ke Puskesmas/Poliklinik	Jenisnya : tidak mampu berobat
12	Lapangan pekerjaan utama seorang kepala rumah tangga	Jenisnya : Petani dengan luas lahan <0,5 Ha, buruh tani, nelayan, buruh bangunan, buruh perkebunan, atau pekerjaan lain dengan pendapatan rumah tangga < Rp 600.000,-
13	Pendidikan terakhir seorang kepala rumah tangga	Jenisnya : tidak sekolah, tidak tamat SD, atau tamat SD
14	Kepemilikan aset, harta bergerak, atau harta tidak bergerak	Jenisnya : tidak mempunyai tabungan, atau barang yang mudah dijual dengan nilai > Rp 500.000,-, seperti sepeda motor, emas, dll

3.2. Rumah Tangga

Secara umum dapat diketahui bahwa rumah tangga merupakan organisasi terkecil dalam masyarakat yang berbentuk karena adanya ikatan perkawinan. Biasanya rumah tangga terdiri atas ayah, ibu dan anak-anak. Namun di Indonesia seringkali dalam rumah tangga juga ada sanak-saudara yang ikut bertempat tinggal, misalnya orang tua, baik dari suami atau istri, saudara kandung/tiri dari kedua belah pihak, kemenakan dan keluarga yang lain, yang mempunyai hubungan darah [17].

Istilah rumah tangga dan keluarga sering dicampur adukkan dalam kehidupan sehari-hari. Pengertian rumah tangga lebih mengacu pada sisi ekonomi, sedangkan keluarga lebih mengacu pada hubungan kekerabatan, fungsi sosial dan lain sebagainya. Keluarga didefinisikan sebagai sekumpulan orang yang tinggal dalam satu rumah yang masih mempunyai hubungan kekerabatan/hubungan darah karena perkawinan, kelahiran, adopsi dan lain sebagainya [18].

Badan Pusat Statistik membagi rumah tangga menjadi dua yaitu rumah tangga biasa dan rumah tangga khusus. Rumah tangga biasa adalah seseorang atau sekelompok orang yang mendiami sebagian atau seluruh bangunan fisik atau sensus dan umumnya tinggal bersama serta makan dari satu dapur. Yang dimaksud dengan satu dapur adalah bahwa pembiayaan keperluan jika pengurusan kebutuhan sehari-hari dikelola bersama-sama. Rumah tangga khusus adalah sekelompok orang yang tinggal di asrama atau tempat tinggal yang pengurusan sehari-harinya diatur oleh yayasan atau badan, misalnya asrama mahasiswa, lembaga pemasyarakatan, orang-orang yang berjumlah lebih dari 10 orang yang kos dengan makan, asrama ABRI dan lain sebagainya [18].

3.3. *Fuzzy C-Means*

Algoritma Fuzzy C-Means diusulkan pertama kali oleh Dunn pada tahun 1973 dan kemudian diperbaharui oleh Bezdek pada tahun 1981. Algoritma ini merupakan salah satu teknik *soft clustering* yang paling populer dengan menggunakan pendekatan data *point* dimana titik pusat *cluster* akan selalu diperbaharui sesuai dengan nilai keanggotaan dari data yang ada dan selain itu algoritma *fuzzy c-means* juga merupakan algoritma yang bekerja dengan menggunakan model *fuzzy* sehingga memungkinkan semua data dari semua anggota kelompok terbentuk dengan derajat keanggotaan yang berbeda antara 0 dan 1. Metode *Fuzzy C-Means* pada dasarnya memiliki tujuan meminimalisasikan fungsi serta mendapatkan pusat *cluster* yang nantinya akan digunakan untuk mengetahui data yang masuk ke dalam sebuah *cluster* [19].

Algoritma ini dimulai dengan menentukan jumlah *cluster* yang diinginkan serta menginisialisasikan nilai keanggotaan yang berisikan semua data kemudian akan dikelompokkan berdasarkan *clusternya*. Pusat-pusat *cluster* dihitung dari jarak terdekat ke titik-titik yang memiliki nilai keanggotaan lebih besar. Dengan kata lain, nilai-nilai keanggotaan tersebut akan bertindak sebagai nilai bobot sementara pada suatu *cluster* [19].

Tingkatan keberadaan informasi dalam sesuatu kelas ataupun *cluster* ditetapkan dari derajat keanggotaannya. Langkah-langkah algoritma *Fuzzy C-Means* yaitu [11][19]:

1. Memasukkan data yang akan di*cluster* berupa matriks X, yaitu matriks

berukuran $n \times m$ (n = banyaknya sampel data dan m = banyaknya variabel setiap data).

X_{ij} = data sampel ke- i ($i = 1, 2, \dots, n$), variabel ke- j ($j = 1, 2, \dots, m$).

2. Menentukan :

- Banyaknya *cluster* yang akan dibentuk (c)
- Pangkat pembobot (w)
- Maksimum iterasi ($MaxItr$)
- Error terkecil (ξ)
- Fungsi objektif awal ($P_0 = 0$)
- Iterasi awal ($t = 1$)

3. Bangkitkan matriks partisi awal $U_{n \times c} = [\mu_{ik}]$, μ_{ik} yaitu bilangan random yang menyatakan suatu derajat keanggotaan.

4. Menghitung pusat *cluster* ke- k (V_{kj}) dengan $k = 1, 2, \dots, c$; dan $j = 1, 2, \dots, m$ sebagai berikut :

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (2.2)$$

5. Menghitung fungsi objektif pada iterasi ke- t , P_t , yang menggambarkan jumlah jarak data ke pusat *cluster*.

$$P_t = \sum_{i=1}^n \sum_{k=1}^c \left(\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right] (\mu_{ik})^w \right) \quad (2.3)$$

Dengan :

P_t = fungsi obyektif;

X_{ij} = elemen X baris i, kolom j,

V_{kj} = pusat *cluster*

6. Menghitung perubahan matriks partisi

$$\mu_{ik} = \frac{\left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left[\sum_{j=1}^m (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{w-1}}}, \quad (2.4)$$

Dengan :

$i = 1, 2, \dots, n$

$k = 1, 2, \dots, c$

X_{ij} = sampel data ke-i, variabel ke-j

V_{kj} = pusat cluster ke-k untuk variabel ke-j

w = pangkat pembobotan

7. Mengecek keadaan berhentinya di mana

- Jika : $(|P_t - P_{t-1}| < \xi)$ atau $(t > \text{MaxIter})$ maka berhenti;
- Jika tidak : $t = t + 1$, maka ulangi langkah ke-4.

Algoritma *Fuzzy C-Means* memiliki keuntungan, yaitu [19] :

1. Dalam implementasi menyelesaikan masalah algoritma *Fuzzy C-Means* dapat memahami karakteristik data yang kabur atau data yang tidak terdefinisikan.
2. Memiliki kemampuan dalam mengelompokkan data yang besar.
3. Lebih kokoh terhadap data *outlier*/data dengan karakter yang berbeda atau *value* yang berbeda dalam satu atau beberapa variabel.

4. Penentuan titik *cluster* yang optimal.
5. Dapat melakukan *clustering* lebih dari satu variabel secara sekaligus.

Adapun kelemahan dari algoritma *Fuzzy C-Means* ini, yaitu [19]:

1. Pada algoritma *Fuzzy C-Means* user memerlukan lebih banyak waktu untuk proses perhitungan komputasinya dalam menentukan *cluster* pada setiap anggota di suatu dataset.
2. Masih terpengaruh terhadap cara pembagian data yang sering dipergunakan pada data yang sama dan sangat sensitif terhadap kondisi awal seperti jumlah *cluster* dan titik pusat *cluster* pada pengelompokan data.

3.4. *Clustering*

Teknik yang paling dikenali dalam *Data Mining* adalah *clustering*. Pengertian *clustering* dalam *Data Mining* yaitu bagaimana cara mengelompokkan sejumlah data atau objek ke dalam *cluster* (grup) sehingga di dalam setiap *cluster* tersebut akan berisi data yang serupa dan juga berbeda dengan objek dalam sebuah *cluster* lainnya [20].

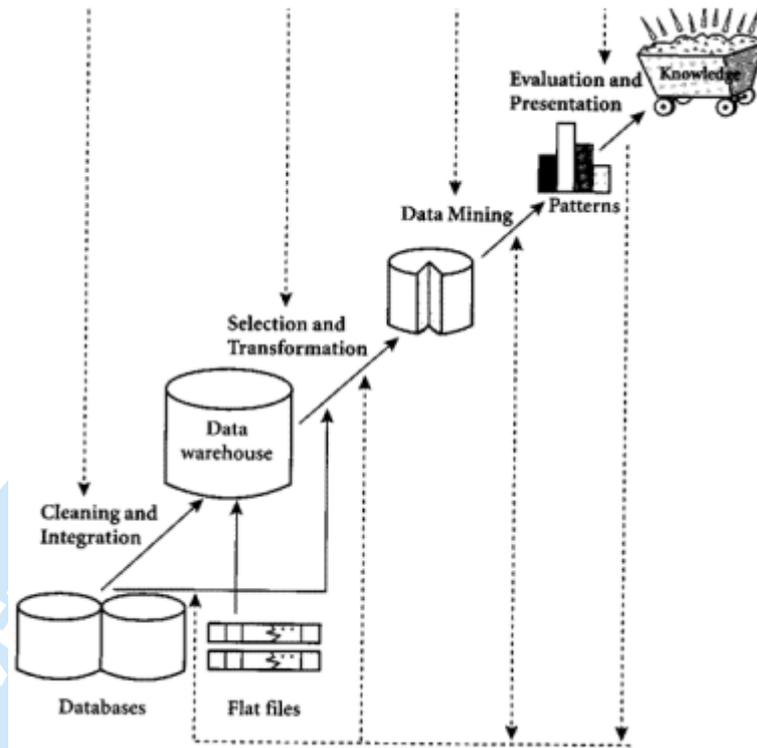
Tata cara yang dipakai agar dapat memecah rangkaian sebuah data sehingga membentuk sebagian grup sesuai dengan keselarasan yang sudah ditunjukkan terdahulu disebut *Clustering*. *Cluster* merupakan kelompok atau kumpulan bahan-bahan di mana data yang sejenis antara satu dengan yang lainnya dalam sebuah *cluster* yang serupa dan diselaraskan mengikuti bahan-bahan yang berbeda *cluster* / kelompok. Bahan-bahan tersebut selanjutnya digabungkan menjadi satu atau lebih sebuah *cluster* kemudian bahan-bahan yang ada pada suatu *cluster* dapat memiliki

persamaan yang tinggi diantara satu dan yang lain [21].

3.5. *Data Mining*

Data mining adalah suatu bagian dan merupakan sebuah langkah yang sangat diperlukan didalam suatu prosedur KDD (*Knowledge Discovery in Database*) yang berhubungan dengan pemisahan dan perhitungan bentuk-bentuk yang telah diperiksa [22]. KDD merupakan proses pengolahan data yang memiliki beberapa tahapan. Tahapan inti pada proses KDD yaitu penambangan data. Fungsi utama dari penambangan data ialah mengekstrak pola data yang tersimpan dengan menerapkan berbagai metode dan algoritma [23].

Data mining adalah proses menganalisis data dengan berbagai perspektif dan meringkasnya ke dalam informasi yang berguna, dimana informasi tersebut bisa digunakan untuk meningkatkan pendapatan, memotong biaya atau keduanya. Tugas dari *data mining* adalah yang pertama **deskriptif**, yaitu menemukan gambaran pola yang menarik dari data, dan yang kedua **prediktif**, yaitu memprediksi perilaku dari model berdasarkan data yang ada. Karena *data mining* adalah suatu rangkaian proses, *data mining* dapat dibagi menjadi beberapa tahap. Tahapan-tahapan tersebut bersifat interaktif dimana pemakai terlibat langsung atau dengan perantara basis pengetahuan [23].



Gambar 3.1. Tahap-Tahap Data Mining

Dari **Gambar 3.1** di atas dapat dijelaskan tahapan-tahapan *data mining* tersebut sebagai berikut: (1) Pembersihan data (untuk membuang data yang tidak konsisten dan *noise*), (2) Integrasi data (penggabungan data dari beberapa sumber), (3) Transformasi data (data diubah menjadi bentuk yang sesuai untuk *data mining*), (4) Aplikasi teknik *data mining*, (5) Evaluasi pola yang ditemukan (untuk menemukan yang menarik/bernilai), dan (6) Presentasi pengetahuan [23].

3.6. Bahasa Pemrograman R

R adalah salah satu perangkat lunak yang digunakan untuk memanipulasi data, perhitungan, simulasi, penampilan grafik, dan juga sebagai salah satu bahasa pemrograman yang bersifat interpreter. Sebelum menjadi bahasa pemrograman R,

bahasa ini dikenal dengan sebutan bahasa S, yaitu bahasa pemrograman yang telah disempurnakan pada Laboratorium Bell. Dikarenakan R berlisensi *open source* oleh karena itu bahasa R mampu didapat dan disebarakan dengan percuma pada lisensi umum GNU. Bahasa R mampu dipakai dalam sistem operasi Windows, Mac OS X, Unix, ataupun Linux. Bahasa R ialah bahasa pemrograman tingkat tinggi dan juga merupakan bahasa yang dapat digunakan untuk analisis data dan grafik [24].

