

BAB II

LANDASAN TEORI

2.1 *Data Mining*

Data mining awalnya didefinisikan sebagai proses menemukan informasi secara otomatis yang berguna dalam *repository* data yang besar (Christian, 2016). Secara tradisional, *data mining* hanya langkah dari penemuan pengetahuan dalam *database*, yang disebut proses *Knowledge Discovery in Database* (KDD) untuk mengkonversi data mentah menjadi ilmu yang bermanfaat. Proses KDD terdiri dari rangkaian langkah transformasi, *data preprocessing*, yang mengubah sumber data yang baku ke dalam bentuk yang tepat untuk analisis selanjutnya (Christian, 2016).

Data mining adalah suatu istilah yang digunakan untuk menemukan pengetahuan yang tersembunyi di dalam kumpulan data. *Data mining* merupakan proses semi otomatis yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial yang berguna dan bermanfaat yang tersimpan di dalam *big database* (Bhuvanewari, 2012). *Data mining* merupakan proses pemilihan, mengeksplorasi dan pemodelan data dalam jumlah besar untuk menemukan pola atau hubungan yang jelas dan berguna dalam proses analisis data (Bellazzi dan Zupan, 2008).

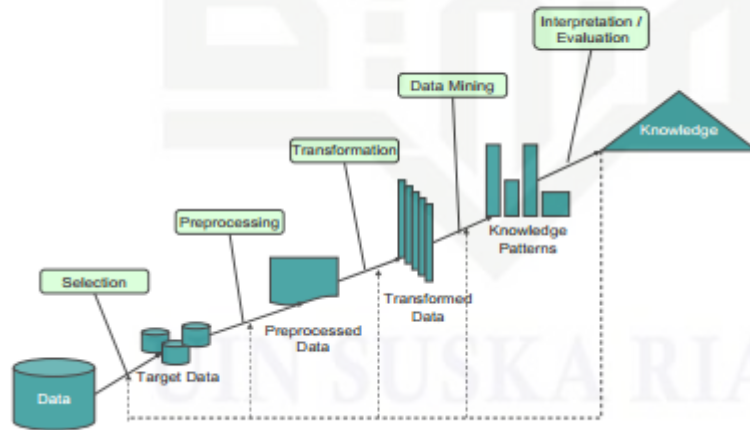
Kemajuan luar biasa yang terus berlanjut dalam bidang *data mining* didorong oleh beberapa faktor, antara lain (Leidiyana, 2006):

1. Pertumbuhan yang cepat dalam kumpulan data.
2. Penyimpanan data dalam *data warehouse*, sehingga seluruh perusahaan memiliki akses ke dalam *database* yang baik.
3. Adanya peningkatan akses data melalui navigasi *web* dan *intranet*.

4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk *data mining* (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

2.1.1 Proses *Data Mining*

Istilah *data mining* dan KDD sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining* (Gilbert dkk, 2008). Gambar 2.1 adalah proses KDD yang paling umum disajikan dalam peta konseptual yang membuat lebih mudah proses seleksi.



Gambar 2.1 Proses Knowledge
 (Sumber: Gilbert dkk, 2008)

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. *Data Selection*

Dalam tahap ini membuat himpunan data target, pemilihan himpunan data, atau memfokuskan pada subset variabel atau sampel data, dimana penemuan (*discovery*) akan dilakukan. Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Data Preprocessing*

Proses ini bertujuan untuk mentransformasikan data input ke dalam format yang sesuai untuk kemudian dianalisa. Dalam tahap ini dilakukan proses penggabungan data dari berbagai sumber, pembersihan data untuk menghilangkan *noise data* dan *duplicate data*, serta memilih atribut data yang diperlukan bagi proses *data mining*.

3. *Data Transformation*

Data transformation adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *mining* dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data Mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Pattern Evolution*

Dalam tahap ini hasil dari teknik *data mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa, ada beberapa *alternative* yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses *data mining*, mencoba teknik *data mining* lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat. Ada beberapa teknik *data mining* yang menghasilkan hasil analisa berjumlah besar seperti analisa prediksi. Visualisasi hasil analisis akan sangat membantu untuk memudahkan pemahaman dari hasil *data mining*.

2.1.2 Pengelompokan *Data Mining*

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu (Kusrini dkk, 2009):

1. Deskripsi

Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh: petugas pengumpulan suara mungkin tidak dapat menemukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden. Deskripsi dari pola dan kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

2. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan *record* lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya pada peninjauan berikutnya, estimasi nilai dari variabel target dibuat berdasar nilai variabel prediksi. Sebagai contoh: akan dilakukan estimasi tekanan darah *systolic* dari pasien rumah sakit berdasarkan umur pasien, jenis kelamin, *index* berat badan dan level

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

sodium darah. Hubungan antara tekanan darah *systolic* dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya. Contoh lain yaitu estimasi nilai *index* prestasi kumulatif mahasiswa program pascasarjana dengan melihat nilai *index* prestasi mahasiswa tersebut pada saat mengikuti program sarjana.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa mendatang. Contoh dari prediksi dalam bisnis dan penelitian yaitu sebagai berikut:

- a. Prediksi harga beras dalam tiga bulan yang akan datang.
- b. Prediksi persentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikan.

Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang tepat) untuk prediksi.

4. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Contoh penggolongan pendapatan, dapat dipisahkan dalam tiga kategori yaitu pendapatan tinggi, pendapatan sedang dan pendapatan rendah. Contoh lain klasifikasi dalam bisnis dan penelitian yaitu:

- a. Menentukan apakah suatu transaksi kartu kredit merupakan transaksi yang curang atau bukan.
- b. Memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit yang baik atau buruk.
- c. Mendiagnosa penyakit seorang pasien untuk mendapatkan termasuk kategori penyakit apa.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

5. Clustering

Clustering merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. *Cluster* adalah kumpulan dari *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam *cluster* lain. Peng-*cluster*-an berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam peng-*cluster*-an. Peng-*cluster*-an tidak mencoba untuk melakukan klasifikasi, mengestimasi atau memprediksi nilai dari variabel target. Akan tetapi algoritma peng-*cluster*-an mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (*homogen*), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal. Contoh peng-*cluster*-an dalam bisnis dan penelitian adalah:

- a. Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
- b. Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku *financial* dalam baik dan mencurigakan.
- c. Melakukan peng-*cluster*-an terhadap ekspresi dari *gen*, untuk mendapatkan kemiripan perilaku dari *gen* dalam jumlah besar.

6. Asosiasi

Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut sebagai analisis keranjang belanja. Contoh asosiasi dalam bisnis dan penelitian adalah:

- a. Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respon positif terhadap penawaran

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

upgrade layanan yang diberikan.

- b. Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli secara bersamaan.

2.2 Clustering

Analisis kelompok adalah pekerjaan mengelompokkan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan di antaranya (Prasetyo, 2012).

Clustering adalah proses membuat pengelompokan sehingga semua anggota dari setiap partisi mempunyai persamaan berdasarkan matriks tertentu. Objek data yang terletak didalam *cluster* harus memiliki kemiripan sedangkan yang tidak berada dalam satu *cluster*, tidak mempunyai kemiripan. Pengelompokan dapat dibedakan menurut struktur kelompoknya pengelompokan dibagi menjadi dua, yaitu pengelompokan hirarki dan sekatan (*partitioning*). Dalam pengelompokan hirarki, satu data tunggal bisa dianggap sebuah kelompok, dua atau lebih kelompok kecil dapat bergabung menjadi sebuah kelompok besar dan begitu seterusnya. Pada sisi lain, metode pengelompokan sekatan membagi set data ke dalam sejumlah kelompok yang tidak tumpang tindih antara satu kelompok dengan kelompok lain (Praseyto, 2012).

Algoritma *clustering* dapat dikategorikan ke dalam kategori hirarkis, partisi berbasis, berbasis *density*, *grid-based* dan model berbasis algoritma pengelompokan (Lu, 2013).

2.3 K- Means Clustering

Clustering merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*). Ada dua jenis data *clustering* yang sering dipergunakan dalam proses pengelompokan data yaitu *hierarchical* (hirarki) data *clustering* dan *non-hierarchical* (nonhirarki) data *clustering*. *K-means* merupakan salah satu

metode data *clustering* non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster* kelompok (Mustakim, 2012). Metode ini dikembangkan oleh Mac Queen pada tahun 1967, dasar algoritma *K-Means* adalah sebagai berikut:

1. Menentukan nilai k sebagai jumlah *cluster* yang ingin di bentuk.
2. Bangkitkan k *centroid* (titik pusat *cluster*) awal secara random.
3. Hitung jarak setiap data ke masing-masing pusat *cluster* menggunakan *Euclidean Distance*.
4. Kelompokan setiap data berdasarkan jarak terdekat antara data dengan pusatnya.
5. Tentukan posisi pusat *cluster* baru (C_{kj}) dengan cara menghitung nilai rata-rata dari data-data yang ada pada pusat *cluster* yang sama.

2.4 *Fuzzy C-Means* (FCM)

Metode FCM pertama kali dikenalkan oleh Jim Bezdek pada tahun 1981 (Ramadhana, 2013). Metode FCM memungkinkan satu data milik dua atau lebih kelompok. Dengan kemampuan FCM untuk memisahkan data terhadap kelompok yang berbeda dengan nilai tertentu, sehingga nilai yang hilang dapat diperoleh dengan menghitung jarak dari *dataset* yang lengkap dan menggunakannya (Samat, 2016). Metode FCM merupakan pengembangan dari metode nonhirarki *K-Means cluster*, karena pada awalnya ditentukan dahulu jumlah kelompok atau *cluster* yang akan dibentuk. Kemudian dilakukan iterasi sampai mendapatkan keanggotaan kelompok (Purnamasari dkk, 2014).

FCM adalah salah satu teknik pengelompokan data yang mana keberadaan tiap titik data dalam suatu kelompok (*cluster*) ditentukan oleh derajat keanggotaan. Metode FCM termasuk metode *supervised clustering* dimana jumlah pusat *cluster* ditentukan di dalam proses *clustering* (Ramadhan, 2013). Algoritma FCM adalah sebagai berikut:

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak milik UIN Suska Riau
Site: <http://www.uin-suska-riau.ac.id>
Library of Sigitas Syarifuddin Riasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Input data yang akan di-*cluster* X , berupa matriks berukuran $n \times m$. n adalah jumlah sampel data, m adalah atribut setiap data. X_{ij} adalah data sampel ke- i dan atribut ke- j .

2. Tentukan:

- a. Jumlah *cluster* = c .
- b. Bobot = w .
- c. Maksimum iterasi = MaxIter .
- d. Error terkecil = ϵ .
- e. Fungsi objektif awal = $P_0 = 0$.
- f. Iterasi awal = $t = 1$.

3. Bangkitkan bilangan random μ_{ik} , $i=1,2,\dots,n$; $k=1,2,\dots,c$; sebagai elemen-elemen matriks partisi awal U . Hitung jumlah setiap kolom dengan Persamaan 2.1.

$$Q_i = \sum_{k=1}^c \mu_{ik} \quad (2.1)$$

Dengan $j=1,2,\dots,m$,

μ_{ik} adalah derajat keanggotaan yang merujuk pada seberapa besar kemungkinan suatu data menjadi anggota ke dalam suatu *cluster*, lihat pada Persamaan 2.2.

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i} \quad (2.2)$$

4. Hitung pusat *cluster* ke- k : V_{kj} , dengan $k=1,2,\dots,c$; dan $j=1,2,\dots,m$ dengan Persamaan 2.3.

$$V_{kj} = \frac{\sum_{i=1}^n (\mu_{ik})^w \times X_{ij}}{\sum_{i=1}^n (\mu_{ik})^w} \quad (2.3)$$

5. Menghitung fungsi objektif pada iterasi ke- t dengan Persamaan 2.4.

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

6. Hitung perubahan matriks partisi menggunakan Persamaan 2.5.

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ij} - V_{kj})^2]^{-\frac{1}{w-1}}} \quad (2.5)$$

7. Cek kondisi berhenti

- Jika $|P_t - P_{t-1}| < \epsilon$ atau $t > \text{MaxIter}$ maka berhenti.
- Jika tidak $t = t+1$, ulangi langkah k-4.

2.5 Validitas Algoritma

Bezdek mengusulkan validasi *fuzzy clustering* dengan menghitung koefisien partisi atau *Partition Coefficient Index* (PCI) sebagai evaluasi nilai keanggotaan data pada setiap *cluster*. Nilai PCI hanya mengavaluasi nilai derajat keanggotaan, tanpa memandang vektor (data) yang biasanya mengandung informasi geometrik. Nilai dalam rentang $[0,1]$, nilai yang semakin besar mendekati 1 mempunyai arti bahwa kualitas *cluster* yang didapat semakin baik (Prasetyo, 2014). Berikut ini adalah formula untuk menghitung PCI pada Persamaan 2.6.

$$PCI = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k u_{ij}^2 \quad (2.6)$$

2.6 Kemiskinan

Definisi yang umum digunakan adalah dari World Bank, dalam perhitungan dan kajian-kajian akademik pengertian kemiskinan adalah sebagai ketidakmampuan mencapai standar hidup minimum. Sedangkan *United Nations Development Program* (UNDP) mendefinisikan kemiskinan sebagai ketidakmampuan untuk memperluas pilihan-pilihan dalam hidup (Listiyono, 2012). Kemudian BPS mendefinisikan kemiskinan sebagai ketidakmampuan untuk memenuhi standar minimum kebutuhan dasar yang meliputi kebutuhan makanan maupun non-makanan. Inti dari model ini adalah membandingkan tingkat konsumsi

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

(SUSENAS). Garis kemiskinan menurut BPS ditetapkan berdasarkan tingkat pendapatan yang diperlukan untuk memenuhi kebutuhan konsumsi 2100 kalori perorang perhari ditambah dengan beberapa kebutuhan non pangan makanan lain seperti sandang, pangan, dan jasa.

- c. Kriteria kesejahteraan yang disebut *index* Kebutuhan Fisik Minimum (KFM). KFM adalah nilai barang dan jasa minimum yang diperlukan oleh suatu keluarga perbulan. Ukuran garis kemiskinan didasarkan pada nilai pengeluaran konsumsi perkapita.

2. Kemiskinan dipahami dalam pengertian relatif (*relativer deprivation*). Indikator kemiskinan ini yaitu:

- a. Deprivasi material yang diukur dari kurangnya pemenuhan kebutuhan pangan, sandang, kesehatan, papan, dan kebutuhan konsumsi dasar lainnya.
- b. Isolasi yang dicerminkan oleh lokasi geografisnya maupun oleh marjinalisasi rumah tangga miskin secara sosial dan politik. Mereka sering tinggal di daerah terpencil, hampir tanpa sarana transportasi dan komunikasi.
- c. Alineasi, yaitu perasaan tidak punya identitas dan kontrol atas diri sendiri. Hal ini diakibatkan isolasi dan hubungan sosial yang eksploitatif.
- d. Ketergantungan yang selama ini menyebabkan lemahnya posisi tawar orang miskin dalam hubungan sosial yang sudah timpang antara pemilik dan penggarap, antara majikan dan buruh atau antara pandega dan ponggawa. Buruh tidak punya kemampuan untuk menetapkan upah, petani dan buruh nelayan tidak bisa menetapkan harga hasil produksi yang dihasilkannya.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- e. Ketidakmampuan membuat keputusan sendiri dan tidak adanya kebebasan memilih dalam produksi, konsumsi dan kesempatan kerja, serta kurangnya perwakilan sosial politik mereka yang tercermin dari tidak adanya fleksibilitas dan berkurangnya kesempatan bagi kaum miskin di desa.
- f. Kelangkaan aset yang membuat penduduk miskin desa bekerja dengan tingkat produktivitas yang sangat rendah.
- g. Kerentanan terhadap guncangan eksternal dan konflik-konflik sosial internal. Hal ini terjadi karena faktor alamiah, perubahan pasar, maupun kondisi kesehatan.
- h. Tidak adanya jaminan keamanan akibat status sosial rendah yang disebabkan oleh posisi yang lemah, faktor agama, ras, dan etnik.

2.7 Kemiskinan di Provinsi Riau

Kemiskinan adalah ketidaksamaan kesempatan untuk mengakumulasi basis kekuasaan sosial yang meliputi aset (tanah, perumahan, peralatan, kesehatan), sumber keuangan (pendapatan dan kredit yang memadai), organisasi sosial politik yang dapat dimanfaatkan untuk mencapai kepentingan bersama, jaringan sosial untuk memperoleh pekerjaan, barang atau jasa, pengetahuan dan keterampilan yang memadai, dan informasi yang berguna (Friedman, 2002).

Jumlah dan persentase penduduk miskin di Riau pada periode 2011-2016 menunjukkan kecenderungan meningkat, dimana jumlah penduduk miskin mengalami peningkatan sebanyak 33,35 ribu jiwa yaitu dari 482,05 ribu jiwa pada tahun 2011 menjadi 515,40 ribu jiwa pada tahun 2016. Jumlah dan persentase penduduk miskin Riau pada Maret 2016 adalah 515,40 ribu atau 7,98 persen dari total penduduk. Peningkatan jumlah penduduk miskin tersebut tidak sepenuhnya terjadi akibat krisis ekonomi, tetapi juga dikarenakan perubahan standar yang digunakan (BPS, 2016).

2.8 Dana Bantuan Pemerintah

Pendapatan Daerah dan Pembiayaan (PAD) merupakan sumber-sumber dana belanja daerah. Apabila PAD terbatas untuk membiayai belanja daerah maka diperlukan adanya bantuan dana transfer dari pemerintah pusat untuk membantu pendanaan belanja daerah dan menggunakan dana Pembiayaan (SiLPA, Pinjaman) bila terjadi defisit anggaran (Zega, 2014).

2.9 WebGIS

Web mapping system adalah sebuah sistem yang digunakan untuk menampilkan peta secara digital. Sistem ini dapat memadukan kekuatan GIS sebagai sebuah alat bantu yang canggih, terutama dalam menangani analisis secara keruangan dengan kekuatan internet sebagai media penyampaian informasi.

WebGIS merupakan sistem informasi geografis yang berbasis *web*. WebGIS standar dibangun menggunakan *Web Map Services* (WMS). Ciri khas WebGIS dalam menampilkan data spasial atau *map* adalah menarik langsung dari *server* dengan tidak memperhitungkan banyaknya lapisan yang diminta. Padahal proses menampilkan data spasial atau *map* ini tidak bisa dipaksakan menarik dalam jumlah besar karena akan memperlambat waktu respon *web*. Hal ini berbanding terbalik dengan harapan pengguna akan WebGIS yang mudah dioperasikan, memiliki tampilan yang ramah, pengaksesan yang halus dan data cukup *update* atau mendekati *near real time* (Dewi, 2015).

2.10 Quantum GIS

Quantum Geographical Information System (QGIS) merupakan *software* yang berbasis *open source* (tidak memerlukan lisensi). Pada QGIS dapat dilakukan proses pengolahan data baik itu spasial maupun nonspasial. Selain itu di dalam QGIS juga dapat dilakukan suatu penambahan fungsi, yang tidak dapat dilakukan pada *software* pemetaan lain seperti Arc GIS. QGIS memiliki fitur-fitur yang pada

umumnya terdapat di dalam ArcGIS, sehingga pada QGIS juga dapat dilakukan proses *georeferencing*, proses pembuatan peta tematik, menghitung luasan dari suatu daerah/ wilayah, dan proses pengolahan pemetaan lainnya yang berhubungan dengan data spasial maupun nonspasial. QGIS sendiri dapat dijalankan pada banyak *Operating System*, seperti *Windows*, *Linux*, maupun *MAC*.

QGIS menyediakan semua fungsionalitas dan fitur-fitur yang dibutuhkan oleh pengguna GIS pada umumnya. Menggunakan *plugins* dan fitur inti (*core features*) dimungkinkan untuk memvisualisasi (meragakan) pemetaan (*maps*) untuk kemudian diedit dan dicetak sebagai sebuah peta yang lengkap.

2.11 PHP

PHP merupakan bahasa *server-side* yang cukup handal, yang akan disatukan dengan HTML dan berada di *server*. Artinya, sintaks dan perintah yang diberikan akan sepenuhnya dijalankan di *server* sebelum dikirim ke komputer *client*. Seluruh aplikasi yang berbasis *web* dapat dibuat menggunakan PHP. Salah satu kelebihan PHP adalah kemampuan untuk dapat melakukan koneksi dengan berbagai *database*, seperti MySQL, PostgreSQL, dan Access. Selain itu PHP juga bersifat *open source*.

Pada awalnya PHP merupakan kependekan dari *Personal Home Page* (Situs personal). PHP pertama kali dibuat oleh Rasmus Lerdorf pada tahun 1995. Pada waktu itu PHP masih bernama *Form Interpreted* (FI), yang wujudnya berupa sekumpulan skrip yang digunakan untuk mengolah data formulir dari *web*. Pada tahun 1997, sebuah perusahaan bernama Zend menulis ulang interpreter PHP menjadi lebih bersih, lebih baik, dan lebih cepat. Kemudian pada Juni 1998, perusahaan tersebut merilis interpreter baru untuk PHP dan meresmikan rilis tersebut sebagai PHP 3.0 dan singkatan PHP diubah menjadi akronim berulang PHP: *Hypertext Preprocessing* (www.php.net, 2018).

Pada Juni 2004, Zend merilis PHP 5.0. Dalam versi ini, inti dari interpreter PHP mengalami perubahan besar. Versi ini juga memasukkan model pemrograman berorientasi objek ke dalam PHP untuk menjawab perkembangan bahasa pemrograman ke arah paradigma berorientasi objek. Pada 17 Desember 2014 Zend kembali merilis PHP terbaru dengan versi 5.6.4 dimana menambahkan beberapa fitur dari versi PHP 6. Tidak membutuhkan waktu yang lama Zend kembali merilis versi terbaru dari PHP yaitu PHP 7. PHP 7 berkembang dari sebuah proyek dengan nama PHPNG (*PHP Next Generation*), PHPNG bertujuan untuk menyusun ulang kode PHP dan untuk meningkatkan performa dari PHP itu sendiri (Sidik, 2016).

2.12 MySQL

MySQL adalah sebuah aplikasi *Relational Database Management Server* (RDBMS) yang sangat cepat dan kokoh serta bersifat *open source*. *MySQL* merupakan salah satu jenis *database server* yang banyak digunakan di dunia maya, yang menggunakan *SQL* sebagai bahasa dasar untuk mengakses *database*. *MySQL* dapat digunakan pada berbagai *platform* sistem operasi.

2.13 Riset Terdahulu

Pada penelitian sebelumnya tentang Implementasi FCM dalam Menganalisa Kemiskinan Desa, Herliana menyebutkan bahwa untuk mempermudah dalam proses evaluasi penyebab kemiskinan suatu desa, maka penentuan kemiskinan suatu desa dapat dilakukan secara berkelompok (*cluster*) dengan mencari tingkat kemiripan dan ketidakmiripan antar desa pada suatu *cluster*. Metode FCM dipilih karena setiap desa dimungkinkan untuk menjadi anggota dari masing-masing *cluster* dengan derajat keanggotaan yang berbeda antara 0 dan 1 (Harliana, 2014). Adapun hasil dari penelitian tersebut adalah menghasilkan empat *cluster* dengan perbandingan pengeluaran konsumsi dengan tingkat pendapatan bersih dari pekerjaan masyarakat. Pada *cluster* pertama dikategorikan desa miskin dilihat dari

pendapatan mencapai Rp. 66.214,- dengan 91,22% pengeluaran digunakan untuk konsumsi. Pada *cluster* ketiga dikategorikan desa cenderung miskin dilihat dari tingkat pendapatan bersih Rp. 58.796,- dengan 89,17% pengeluaran untuk konsumsi. Pada *cluster* keempat dikategorikan desa sangat miskin dilihat dari pendapatan bersih Rp. 52.678,- dengan 91,49% pengeluaran digunakan untuk konsumsi.

Berikut adalah daftar riset lain yang digunakan sebagai acuan atau referensi dalam penelitian ini, lihat pada Tabel 2.1.

Tabel 2.1 Penelitian Terdahulu

Penulis	Judul	Hasil Penelitian
Herliana (2014)	Implementasi <i>Fuzzy C-Means</i> Dalam Menganalisa Kemiskinan Desa	Pengelompokan desa-desa yang berada pada 3 kecamatan, yaitu kecamatan gebang, kapetakan dan weru. Hasil menggunakan metode FCM menjadi 4 kelompok. Secara umum <i>cluster</i> 1 dikategorikan menjadi desa miskin. <i>Cluster</i> 3 dikategorikan menjadi desa cenderung miskin dan <i>cluster</i> 4 dikategorikan menjadi desa sangat miskin. Hasil <i>cluster</i> di lihat dari persentasi pendapatan, pengeluaran dan hasil matapencaharian dari masyarakat yang sudah di proses menggunakan metode FCM.
Redjeki, dkk. (2014)	Perancangan Sistem Indikasi dan Pemetaan Potensi Kemiskinan untuk Optimalisasi Program Kemsकिनan	Metode yang digunakan adalah AHP dalam sistem pendukung keputusan kemiskinan yang dapat membantu memberi pertimbangan bagi pengambilan keputusan dalam penentuan status warga miskin. Sistem yang dihasilkan pada penelitian di implementasikan pada sistem geografis.

- Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 2.1 Penelitian Terdahulu (Lanjutan)

Penulis	Judul	Hasil Penelitian
Ahmadi, dkk. (2013)	Penerapan <i>Fuzzy C-Means</i> dalam Sistem Pendukung Keputusan Untuk Penentuan Penerima Bantuan Langsung Masyarakat (BLM) PNPM-Mpd	Dari hasil pengujian algoritma FCM dalam penelitian Ahmadi didapat bahwa input di beberapa masukan pada poses FCM berpengaruh pada hasil <i>cluster</i> . Input yang dimasukan adalah Maxiter, <i>fuzzier</i> , nilai error terkecil, fungsi objektif awal dan iterasi awal. Hasil perangkangan dalam usulan kegiatan PNPM Mandiri Pedesaan di Kecamatan Ngadirojo berjalan dengan baik, namun nilai input awal sangat berpengaruh pada hasil <i>cluster</i> FCM. Selain itu nilai <i>random</i> yang terbentuk juga sangat mempengaruhi proses perhitungan dalam menentukan kedekatan usulan terhadap pusat <i>cluster</i> .
Wulandari, dkk. (2017)	Model Penanggulangan Kemiskinan Berbasis Wilayah Menggunakan <i>Fuzzy C-Means</i> dan Simple Matching Coefficient	Pengelompokan kemiskinan dibedakan berdasarkan beberapa faktor kemiskinan yaitu: (1) kesejahteraan, (2) pendidikan, (3) lingkungan ekonomi, (4) infrastruktur dan pelayanan dan (5) lingkungan pendukung. Bobot kemiskinan yang ditentukan adalah (1) Miskin = 0-38.54, (2) Sedang = 38.55-61.45 dan (3) Sejahtera = 61.46-100. Dari proses FCM didapat jumlah kluster adalah 4 (empat) kuluster.
Selviana (2016)	Analisis Penerapan <i>K-Means</i> dan <i>Fuzzy C-Means</i> untuk <i>Clustering</i> Pemetaan Motivasi Belajar Mahasiswa Terhadap Strategi Pembelajaran	Hasil validasi <i>cluster</i> dari algoritma <i>k-means</i> dan FCM, menunjukkan bahwa algoritma FCM lebih baik dibandingkan dengan algoritma <i>k-means</i> , hal ini dibuktikan hasil validasi <i>k-means</i> 0,2896 dan FCM 0,5098. Berdasarkan ketentuan dari validasi <i>cluster</i> bahwa nilai yang lebih mendekati satu mempunyai kualitas <i>cluster</i> yang semakin baik.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 2.1 Penelitian Terdahulu (Lanjutan)

Penulis	Judul	Hasil Penelitian
Simbolon, dkk. (2013).	Lulusan Mahasiswa Matematika FMIPA Untan Pontianak Menggunakan Algoritma <i>Fuzzy C-Means</i>	FCM digunakan untuk mengelompokkan lulusan jurusan Matematika FMIPA Universitas Tangjungpura (UNTAN). Lulusan dibagi kedalam empat <i>cluster</i> berdasarkan IPK dan lama studi. Hasil dari penelitian ini adalah diperoleh empat pusat <i>cluster</i> atau <i>center</i> . Hasil yang diperoleh menunjukkan bahwa masih banyak lulusan jurusan Matematika di Fakultas MIPA UNTAN yang menempuh lama studi lebih dari 5 tahun
Reza, dkk. (2017)	Penentuan Jurusan Siswa Sekolah Menengah Atas Disesuaikan Dengan Minat Siswa Menggunakan Algoritma <i>Fuzzy C-Means</i>	Dari Hasil pengujian algoritma <i>Fuzzy C-Means</i> (FCM) dalam penentuan jurusan di Sekolah Menengah Atas pada 278 data siswa yang diuji dalam penelitian ini, menunjukkan bahwa Algoritma FCM memiliki tingkat akurasi yang baik (yaitu rata-rata 82,01%) dengan memasukkan nilai tes minat dibandingkan dengan cara manual berdasarkan pilihan individu siswa hanya 63,67%, sehingga penjurusan dengan FCM lebih tinggi 18,34%. Proses klastering pada penelitian ini hanya dibagi ke dalam 2 (dua) klaster, karena mengikuti jumlah penjurusan yang di sekolah tempat pengambilan data (jurusan IPA dan IPS).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.