

BAB II

LANDASAN TEORI

2.1. Data Mining

Data Mining merupakan salah satu ilmu yang paling penting dan mendukung penelitian dengan tujuan untuk menemukan informasi yang berarti dari data yang sangat besar. Di era sekarang, *Data Mining* menjadi populer di bidang kesehatan karena ada kebutuhan metodologi analitis efisien untuk mendeteksi informasi yang tidak diketahui dan berharga dalam data kesehatan. Dalam industri kesehatan, *Data Mining* memberikan beberapa manfaat seperti deteksi penipuan di asuransi kesehatan, ketersediaan solusi medis untuk pasien dengan biaya yang lebih rendah, deteksi penyebab penyakit dan identifikasi metode pengobatan medis (Tomar, 2013). Menurut Gartner Group, *Data Mining* adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika (Larose, 2005).

Kemajuan luar biasa yang terus berlanjut dalam bidang *Data Mining* didorong oleh beberapa faktor, antara lain: (Larose, 2005 dalam Nofriansyah 2014)

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.

Berdasarkan definisi-definisi yang telah disampaikan, hal penting yang terkait dengan *Data Mining* adalah:

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 Mining* merupakan suatu proses otomatis terhadap data yang sudah ada.

2. Data yang akan diproses berupa data yang sangat besar.

Tujuan *Data Mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

2.1.1 Jenis Permasalahan Data Mining

Aplikasi yang menggunakan *Data Mining* dapat digunakan untuk menyelesaikan berbagai permasalahan dengan membangun model berdasarkan data yang sudah digali untuk diterapkan pada data yang lain. Secara umum ada dua jenis tipologi aplikasi dalam *Data Mining* yaitu (Widodo, 2013):

1. Metode prediksi, yang bermaksud memprediksi nilai yang akan datang berdasarkan data-data yang telah ada variabelnya seperti klasifikasi, regresi, deteksi anomaly dan lain-lain.
2. Metode deskriptif, yang bermaksud membantu *user* agar dapat dengan mudah melihat pola-pola yang berasal dari data yang ada.

2.1.2 Pengelompokan Data Mining

Data Mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat di lakukan, yaitu (Kusrini, 2009):

1. Deskripsi

Terkadang peneliti dan analisis 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

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.

berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya.

3. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Contoh prediksi dalam bisnis dan penelitian adalah:

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

4. Klasifikasi

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

- 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 apa.

5. Pengklasteran

Pengklasteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Klaster adalah kumpulan *record* yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan *record*

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.

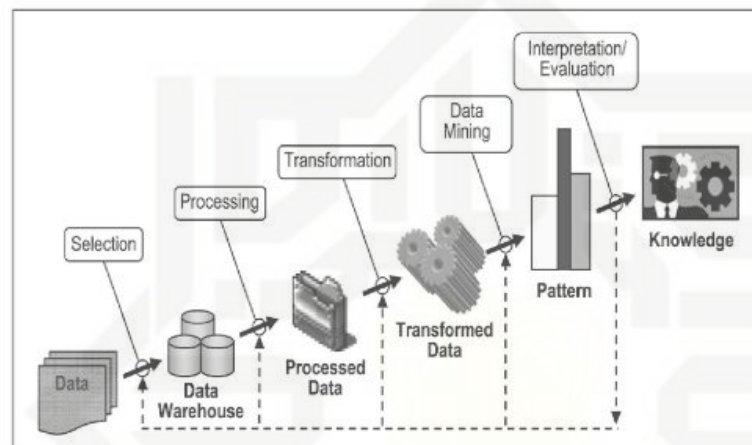
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dalam klaster lain. Pengklasteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklasteran. Pengklasteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan dengan *record* dalam kelompok lain akan bernilai minimal. Contoh pengklasteran 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 akutansi, yaitu melakukan pemisahan terhadap perilaku finansial dalam baik dan mencurigakan.
 - c. Melakukan pengklasteran terhadap ekspresi dari gen, dalam jumlah besar.
6. Asosiasi
- Tugas asosiasi dalam *Data mining* adalah menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut 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 *upgrade* layanan yang diberikan.
 - b. Menemukan barang dalam supermarket yang dibeli secara bersamaan dan barang yang tidak pernah dibeli bersamaan.

2.1.3 Proses Data Mining

Istilah *data mining* dan *Knowledge Discovery in Database* (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*. Proses KDD secara garis besar dapat dijelaskan sebagai berikut (Nofriansyah, 2014) :



Gambar 2.1 Proses *Knowledge Discovery in Database* (Nofriansyah, 2014)

a. Data Cleaning

Data-data yang tidak relevan itu juga lebih baik dibersihkan karena keberadaannya bisa mengurangi mutu atau akurasi dari hasil *data mining* merupakan istilah yang sering dipakai untuk menggambarkan tahapan ini. Pembersihan data juga akan mempengaruhi performasi dari sistem *data mining* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

b. Data Integration

Data mining tidak hanya berasal dari suatu *database* tetapi juga berasal dari beberapa *database* atau file teks, integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan sebagainya.

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.

c. *Data Transformation*

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

d. *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.

e. *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 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.

f. *Knowledge*

Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisa yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining.

2.2. Pengelompokan (*Clustering*)

Clustering atau klusterisasi adalah salah satu alat bantu pada data mining yang bertujuan mengelompokkan objek-objek ke dalam *cluster* - *cluster*. *Cluster* adalah sekelompok atau sekumpulan objek - objek data yang similar satu sama lain dalam cluster yang sama dan dissimilar terhadap objek-objek yang berbeda 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 (Defiyanti, 2017).

Analisis kelompok (*cluster analisis*) adalah pekerjaan mengelompokkan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan diantaranya (Tan, 2006 dalam buku Prasetyo, 2012). Dengan menggunakan klusterisasi, dapat mengidentifikasi daerah yang padat, menemukan pola-pola distribusi secara keseluruhan, dan menemukan keterkaitan yang menarik antara atribut - atribut data.

2.3. Algoritma K-Means

Dari beberapa teknik klastering yang paling sederhana dan umum dikenal adalah klastering *K-Means*. Untuk melakukan klastering ini, nilai *k* harus ditentukan terlebih dahulu. Biasanya *user* atau pemakai sudah memiliki informasi awal tentang objek yang sedang dipelajari, termasuk berapa jumlah klaster (*cluster*) yang paling tepat. Secara detail bisa menggunakan ukuran ketidakmiripan untuk mengelompokkan objek. Ketidakmiripan bisa diterjemahkan kedalam konsep jarak. Jika jarak dua objek atau data titik cukup dekat, maka dua objek itu mirip. Semakin dekat berarti semakin tinggi kemiripannya. Semakin tinggi nilai jarak, semakin tinggi ketidakmiripannya (Selviana dan Mustakim, 2016).

Proses *clustering* dimulai dengan mengidentifikasi data yang akan di *cluster*, X_{ij} ($i=1,...,n$; $j=1,...,m$) dengan n adalah jumlah data yang akan di *cluster* dan m adalah jumlah variabel. Pada awal iterasi, pusat setiap *cluster* ditetapkan secara bebas (sembarang), C_{kj} ($k=1,...,k$; $j=1,...,m$). Kemudian dihitung jarak antara setiap data dengan setiap pusat *cluster*. Untuk melakukan penghitungan jarak ke- i (x_i) pada pusat *cluster* ke- k (c_k), diberi nama (d_{ik}), dapat digunakan formula *Euclidean* seperti Rumus 2.1.

$$d_{ik} \sqrt{\sum_{j=1}^m (C_{ij} - C_{kj})^2} \quad (2.1)$$

Suatu data akan menjadi anggota dari *cluster* ke-*k* apabila jarak data tersebut ke pusat *cluster* ke-*k* bernilai paling kecil jika dibandingkan dengan jarak ke pusat *cluster* lainnya. Hal ini dapat dihitung dengan menggunakan Rumus 2.2 kelompokkan data-data yang menjadi anggota pada setiap *cluster*.

$$m \sum_{k=1}^k d_{ik} \sqrt{\sum_{j=1}^m (C_{ij} - C_{kj})^2} \quad (2.2)$$

Nilai pusat *cluster* yang baru dapat dihitung dengan cara mencari nilai rata-rata dari data yang menjadi anggota pada *cluster* tersebut, dengan menggunakan Rumus 2.3.

$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p} \quad (2.3)$$

Dimana $x_{ij} \in \text{Cluster ke-}k$, P adalah banyaknya anggota *cluster* ke-*k*.

Algoritma *K-means clustering* adalah sebagai berikut:

1. Pilih jumlah *cluster* *k*.
2. Inisialisasi *k* pusat *cluster* ini bisa dilakukan dengan berbagai cara. Yang paling sering dilakukan adalah dengan cara random. Pusat-pusat *cluster* diberi nilai awal dengan angka-angka *random*.
3. Tempatkan setiap data atau objek ke *cluster* terdekat kedekatan dua objek ditentukan berdasar jarak kedua objek tersebut. Demikian juga kedekatan suatu data *cluster* tertentu ditentukankan jarak antara data dengan pusat *cluster*. Ini menggunakan Rumus 2.1.
4. Hitung kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang pusat *cluster* adalah rata-rata dari semua data atau objek dalam *cluster* tertentu. Ini menggunakan Rumus 2.2.
5. Tugaskan lagi setiap objek dengan memakai pusat *cluster* yang baru. Jika pusat *cluster* sudah tidak berubah lagi, maka proses peng-*cluster*-an selesai. Ini menggunakan Rumus 2.3.
6. Kembali lagi ke langkah tiga sampai pusat *cluster* tidak mengalami perubahan signifikan dari nilai SSE (*Sum of Squared Errors*).

2.4. Metode Asosiasi Rule

Analisis asosiasi atau *association rule mining* adalah teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Contoh aturan asosiatif dari analisa pembelian di suatu pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu. Analisis asosiasi menjadi terkenal karena aplikasinya untuk menganalisa isi keranjang belanja di pasar swalayan. Analisis asosiasi juga sering disebut dengan istilah *market basket analysis*.

Salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien adalah analisis pola frekuensi tinggi (*frequent pattern mining*) atau biasa dikenal dengan *support*. Kombinasi item yang digunakan adalah yang memenuhi syarat minimum dari nilai *support*. Nilai *support* sebuah item diperoleh dengan menggunakan persamaan 2.4 berikut (Srikanti dkk, 2018):

$$Support(A) = \frac{Jumlah\ responden\ mengandung\ A}{Jumlah\ total\ responden} \quad (2.4)$$

Setelah semua pola frekuensi tinggi atau *support* ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan asosiatif A U B. Nilai *confidence* dari aturan A U B diperoleh dengan persamaan 2.5 berikut (Srikanti dkk, 2018).

$$Confidence = \frac{Jumlah\ responden\ mengandung\ A\ dan\ B}{Jumlah\ responden\ mengandung\ A} \quad (2.5)$$

2.5. Kota Pekanbaru

Kota Pekanbaru mempunyai luas wilayah 632,26 Km² terdiri dari 12 (dua belas) Kecamatan dan 58 Kelurahan. Penduduk Kota Pekanbaru tahun 2015 berdasarkan data dari Kementerian Kesehatan dan BPS adalah berjumlah ±1.038.118 jiwa, dengan kepadatan penduduk 1.642 Km². Kota Pekanbaru terletak antara garis 101,14' – 101,34' Bujur Timur dan 0,25' – 0,45' Lintang Utara dengan batas wilayah sebagai berikut :

1. Sebelah Utara berbatasan dengan Kabupaten Siak.
2. Sebelah Selatan berbatasan dengan Kabupaten Kampar.
3. Sebelah Barat berbatasan dengan Kabupaten Kampar.
4. Sebelah Timur berbatasan dengan Kabupaten Kampar dan Siak.

2.5.1. Kependudukan Kota Pekanbaru

Berdasarkan data dari Keputusan menteri kesehatan Republik Indonesia Nomor HK.02.02/MENKES/117/2015 tentang Data penduduk sasaran program pembangunan kesehatan tahun 2015-2019 dan data dari Badan Pusat Statistik (BPS) tentang Proyeksi Penduduk Kabupaten/Kota Provinsi Riau, jumlah penduduk kota Pekanbaru $\pm 1.038.118$ Jiwa. Kelompok umur yang memiliki proporsi terbesar yaitu jumlah penduduk berusia antara 20 sampai 24 tahun yaitu 123.302 orang (12%). Dari struktur penduduk terlihat bahwa kelompok umur balita yaitu 107.889 orang (10%), sedangkan kelompok umur >65 tahun terdapat 26.470 orang (2,5%). *Sex ratio* antara laki-laki dan perempuan, ditemukan laki-laki lebih besar dari pada perempuan yaitu rasio 105,61.

Kepadatan penduduk Kota Pekanbaru 1.642 Km^2 . Adapun beban tanggungan yaitu beban yang ditanggung oleh penghasilan golongan produktif (usia 15 – 64 tahun) untuk dikeluarkan bagi memenuhi kebutuhan mereka yang tidak produktif (usia 0 – 14 tahun dan usia > 65 tahun) .

2.5.2. Sarana Kesehatan

Pada tahun 2015 di kota Pekanbaru terdapat 28 Rumah Sakit. Rumah Sakit tersebut terdiri atas:

Tabel 2.1 Rumah Sakit di Kota Pekanbaru

No	Nama Rumah Sakit	Kecamatan	Alamat
1	RS. Arifin Ahmad	Pekanbaru Kota	Jl.Diponegoro No.2
2	RS. Polda Riau	Pekanbaru Kota	Jl.Kartini No.14
3	RS. TNI AD	Senapelan	Jl.Kesehatan No.2
4	RS. TNI AU	Marpojan Damai	Jl.Adi Sucipto
5	RS. Islam Ibnu Sina	Sukajadi	Jl.Melati No.60

- 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 Rumah Sakit di Kota Pekanbaru (Lanjutan)

No	Nama Rumah Sakit	Kecamatan	Alamat
6	RS. Islam Yayasan Abdurrah	Marpoayan Damai	Jl.Sudirman No.410
7	RS. Santa Maria	Sukajadi	Jl.A.Yani No.68
8	RS. Bina Kasih	Senapelan	Jl.Samanhudi No 3-5
9	RS. Pekanbaru Medical Center	Sail	Jl. Lembaga Pemasyarakatan
10	RS. Lancang Kuning	Sail	JL.Ronggowarsito No.25
11	RS. Eka Hospital	Marpoayan Damai	Jl. Soekarno Hatta Km 6,5
12	RS. A.Yani	Pekanbaru Kota	Jl. A.Yani
13	RS. Awal Bros	Bukit Raya	Jl. Jend.Sudirman No117
14	RS. Petala Bumi	Lima Puluh	Jl. S.Syarif Kasim
15	RS. Andini	Rumbai Pesisir	Jl. Sekolah
16	RS. Budi Mulia	Marpoayan Damai	Jl. Soekarno Hatta
17	RS. Awal Bros Panam	Tampan	Jl. Subrantas
18	RS. Sansani	Marpoayan Damai	Jl. Soekarno Hatta
19	RS. Universitas Riau	Tampan	Komplek UNRI
20	RSIA Eria Bunda	Sukajadi	Jl.KH.Ahmad Dahlan
21	RSIA Zainab	Sail	Jl. Ronggowarsito I
22	RSIA Syafira	Marpoayan Damai	Jl. Jend.Sudirman No 134
23	RSIA Andini	Marpoayan Damai	Jl. Tuanku Tambusai No 55
24	RSIA Labuh Baru	Payung Sekaki	Jl. Durian
25	RSIA Annisa	Marpoayan Damai	Jl. Garuda
26	RSIA Budhi Mulya	Marpoayan Damai	Jl. Soekarno Hatta
27	RS Mata Eye Center	Marpoayan Damai	Jl. Soekarno Hatta
28	Rumah Sakit Jiwa Tampan	Tampan	Jl. Subrantas

(Sumber: Dinas Kesehatan Kota Pekanbaru, 2015)

Dinas Kesehatan Kota Pekanbaru mempunyai 20 Puskesmas yang tersebar di setiap kecamatan terdiri dari 15 Puskesmas Rawat Jalan dan 5 Puskesmas Rawat Inap. Daftar puskesmas yang ada di Kota Pekanbaru dapat dilihat pada Tabel 2.2.

- 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 2.2 Puskesmas di Kota Pekanbaru

No	Nama Puskesmas	Kecamatan	Alamat
1	Pusk. Langsung	Sukajadi	Jl. Langsung
2	Pusk. Melur	Sukajadi	Jl. Melur
3	Pusk. Payung Sekaki	Payung Sekaki	Jl. Fajar No.21
4	Pusk. Lima Puluh	Lima Puluh	Jl. Sumber Sari No.118
5	Pusk. Pekanbaru Kota	Pekanbaru Kota	Jl. Teuku Umar
6	Pusk. Senapelan	Senapelan	Jl. Jati No.4
7	Pusk. Rumbai	Rumbai Pesisir	Jl. Sekolah
8	Pusk. Karya Wanita (RI)	Rumbai Pesisir	Jl. Gabus Raya
9	Pusk. Umban Sari	Rumbai	Jl. Purnama Sari
10	Pusk. Muara Fajar	Rumbai	Jl. Lintas Duri-Pekanbaru
11	Pusk. Simpang Tiga (RI)	Marpoyan Damai	Jl. Kaharudin Nasution
12	Pusk. Harapan Raya (RI)	Bukit Raya	Jl. Harapan Raya No.40
13	Pusk. Rejosari	Tenayan Raya	Jl. Hangtuah Komp. Indah Sari
14	Pusk. Sail	Sail	Jl. Hang Jebat
15	Pusk. Tenayan Raya (RI)	Tenayan Raya	Jl. Budi Luhur Pekanbaru
16	Pusk. Garuda	Marpoyan Damai	Jl. Garuda No.12
17	Pusk. Sidomulyo (RI)	Tampan	Jl. Delima Panam
18	Pusk. Sidomulyo	Tampan	Jl. HR. Subrantas
19	Pusk. Simpang Baru	Tampan	Jl. Kamboja – Panam
20	Pusk. Rumbai Bukit	Rumbai	Jl. Sri Palas

(Sumber: Dinas Kesehatan Kota Pekanbaru, 2015)

2.6. Dinas Kesehatan Kota Pekanbaru

Merupakan sebuah instansi/badan yang dibentuk pemerintah untuk membantu Kepala Daerah dibidang Kesehatan sesuai dengan Peraturan Menteri Kesehatan RI Tahun 2016 tentang pedoman teknis pengorganisasian Dinas Kesehatan Provinsi dan Kabupaten/Kota. Kantor Dinas Kesehatan Kota Pekanbaru beralamat di Jalan Melur Nomor 103 Kelurahan Padang Bulan Kecamatan Senapelan.

2.6.1. Tugas Dinas Kesehatan

Dinas mempunyai tugas pokok menyelenggarakan perumusan kebijakan, koordinasi, fasilitasi, pemantauan, evaluasi dan pelaporan pada Sekretariat, Bidang Sumber Daya Kesehatan dan Kefarmasian, Bidang Pelayanan Kesehatan, Bidang Pencegahan, Pengendalian Penyakit dan Penyehatan Lingkungan dan Bidang Promosi Kesehatan dan Kesehatan Keluarga serta menyelenggarakan kewenangan yang dilimpahkan Pemerintah kepada Gubernur sesuai dengan peraturan perundang-undangan.

2.6.2. Fungsi Dinas Kesehatan

Fungsi Dinas Kesehatan adalah sebagai berikut:

1. Penyelenggaraan perumusan kebijakan pada Sekretariat, Bidang Sumber Daya Kesehatan dan Kefarmasian, Bidang Pelayanan Kesehatan, Bidang Pencegahan, Pengendalian Penyakit dan Penyehatan Lingkungan dan Bidang Promosi Kesehatan dan Kesehatan Keluarga;
2. Penyelenggaraan koordinasi dan fasilitasi pada Sekretariat, Bidang Sumber Daya Kesehatan dan Kefarmasian, Bidang Pelayanan Kesehatan, Bidang Pencegahan, Pengendalian Penyakit dan Penyehatan Lingkungan dan Bidang Promosi Kesehatan dan Kesehatan Keluarga;
3. Penyelenggaraan pemantauan, evaluasi dan pelaporan pada Sekretariat, Bidang Sumber Daya Kesehatan dan Kefarmasian, Bidang Pelayanan Kesehatan, Bidang Pencegahan, Pengendalian Penyakit dan Penyehatan Lingkungan dan Bidang Promosi Kesehatan dan Kesehatan Keluarga;
4. Penyelenggaraan tugas dan fungsi lain sesuai dengan tugas dan fungsinya berdasarkan peraturan perundang-undangan.

2.6.3. Struktur Organisasi Dinas Kesehatan

Struktur Organisasi Dinas Kesehatan yaitu sebagai berikut:

1. Kepala Dinas
2. Sekretariat

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.

- a. Sub Bagian Perencanaan Program
- b. Sub Bagian Keuangan dan Perlengkapan
- c. Sub Bagian Umum
3. Bidang Sumber Daya Kesehatan dan Kefarmasian
 - a. Seksi Farmasi, Makanan dan Minuman
 - b. Seksi Pengembangan dan Pendayagunaan Tenaga Kesehatan
 - c. Seksi Pengembangan Sarana Kesehatan
4. Bidang Pelayanan Kesehatan
 - a. Seksi Kesehatan Dasar
 - b. Seksi Rujukan dan Kesehatan Khusus
 - c. Seksi Pembiayaan Kesehatan
5. Bidang Pencegahan, Pengendalian Penyakit dan Penyehatan Lingkungan
 - a. Seksi Pengendalian dan Pemberantasan Penyakit
 - b. Seksi Penyehatan Lingkungan
 - c. Seksi Surveilans dan Kesehatan Matra
6. Bidang Promosi Kesehatan dan Kesehatan Keluarga
 - a. Seksi Promosi Kesehatan
 - b. Seksi Gizi Masyarakat
 - c. Seksi Kesehatan Keluarga

2.6. Penyakit

Penyakit pada manusia merupakan gangguan kesehatan yang dapat mempengaruhi fungsi yang ada pada tubuh menjadi tidak normal sehingga dapat mengganggu manusia dalam beraktifitas (Hardiyanto dkk, 2016). Berbagai penyakit yang ada di kota Pekanbaru berdasarkan data Dinas Kesehatan Kota Pekanbaru tahun 2014 menyebutkan bahwa terdapat 10 penyakit tidak menular terbesar diantaranya *hipertensi*, *diabetes*, asma, stroke dan lainnya. Sedangkan berdasarkan hasil survey penelitian menunjukkan ada beberapa penyakit yang umum terjadi dikalangan masyarakat kota Pekanbaru diantaranya alergi, batuk, flu, demam, diare, magh, sakit kepala, sakit perut, *hipertensi* dan *hipotensi*.

Berdasarkan penelitian ada beberapa penyakit yang saling berhubungan, diantaranya penyakit *hipertensi* dan *stroke*. Responden dengan riwayat *hipertensi* memiliki resiko 2 kali lebih besar terkena *stroke* (Hafid, 2014). Selanjutnya penyakit demam dan *influenzha*, semakin tinggi demam yang diderita maka semakin besar kemungkinan gejala terserang penyakit *influenzha* (Komnas FBPI, 2009).

2.7.1. Alergi

Definisi Reaksi Alergi (Reaksi *Hipersensitivitas*) adalah reaksi-reaksi dari sistem kekebalan yang terjadi ketika jaringan tubuh yang normal mengalami cedera/terluka. Mekanisme dimana sistem kekebalan melindungi tubuh dan mekanisme dimana reaksi *hipersensitivitas* bisa melukai tubuh adalah sama. Karena itu reaksi alergi juga melibatkan antibodi, limfosit dan sel-sel lainnya yang merupakan komponen dalam system imun yang berfungsi sebagai pelindung yang normal pada sistem kekebalan. Alergi bisa berupa partikel debu, serbuk tanaman, obat atau makanan, yang bertindak sebagai antigen yang merangsang terajdinya respon kekebalan (Hikmah, 2010).

2.7.2. Batuk

Batuk adalah pengeluaran sejumlah volume udara secara mendadak dari rongga toraks melalui epiglottis dan mulut. Melalui mekanisme tersebut dihasilkan aliran udara yang sangat cepat yang dapat melontarkan keluar material yang ada di sepanjang saluran respiratorik, terutama saluran yang besar. Dengan demikian batuk mempunyai fungsi penting sebagai salah satu mekanisme utama pertahanan respiratorik (Setyanto, 2004). Mekanisme lain yang bekerja sama dengan batuk adalah bersihan *mukosilier* (*mucociliary clearance*). Batuk akan mencegah aspirasi makanan padat atau cair dan berbagai benda asing lain dari luar. Batuk juga akan membawa keluar sekresi berlebihan yang diproduksi di dalam saluran respiratorik, terutama pada saat terjadi radang oleh berbagai penyebabnya..

2.7.3. Influenzha

Influenzha atau “flu” disebabkan oleh infeksi virus *influenzha* A, B, dan lebih jarang, C. Penyakit ini terutama berdampak terhadap tenggorok dan paruparu, tetapi juga dapat mengakibatkan masalah jantung dan bagian lain tubuh, terutama

di kalangan penderita masalah kesehatan lain. Kebanyakan penderita sembuh dalam waktu seminggu. Dibandingkan dengan banyak infeksi lain (misalnya pilek), *influenza* cenderung mengakibatkan gejala dan komplikasi yang lebih parah. Komplikasi dapat termasuk pneumonia, kegagalan jantung atau semakin parahnya penyakit lain (Sumber: www.health.nsw.gov.au).

2.7.4. Demam

Demam adalah keadaan dimana temperatur rektal $>38^{\circ}\text{C}$. Demam mengacu pada peningkatan suhu tubuh yang berhubungan langsung dengan tingkat *sitokin pirogen* yang diproduksi untuk mengatasi berbagai rangsang, misalnya terhadap *toksin* bakteri, peradangan, dan rangsang pirogenik lain (Atiq, 2009).

2.7.5. Diare

Penyakit diare atau *Gastroenteritis* adalah sebuah penyakit disaat tinja atau feses berubah jadi lembek yang biasanya terjadi paling sedikit tiga kali dalam 24 jam. Pada tahun 2015 jumlah penderita diare sebanyak 7.051 orang dan seluruhnya (100%) ditolong oleh tenaga kesehatan maupun kader kesehatan di posyandu. (Dinas Kesehatan Kota Pekanbaru, 2015)

2.7.6. Maag (Gastritis)

Gastritis atau istilah yang sering dikenal masyarakat sebagai maag atau penyakit lambung adalah kumpulan gejala yang dirasakan sebagai nyeri terutama diulu hati, orang yang terserang penyakit ini biasanya sering mual, muntah, rasa penuh, dan merasa tidak nyaman (Misnadiarly, 2009). Beberapa hal yang mempengaruhi timbulnya Maag antara lain:

- Pengeluaran asam lambung yang berlebihan,
- Pertahanan dinding lambung yang lemah,
- Infeksi *helicobakter pylori* (sejenis bakteri yang hidup didalam lambung, dalam jumlah kecil) ketika asam lambung yang dihasilkan lebih banyak sehingga pertahanan lambung menjadi lemah, bakteri ini bisa bertambah banyak jumlahnya, apalagi kebersihan makanan yang kurang.
- Gangguan gerakan saluran pencernaan,
- Stress psikologi.

2.7.7. Sakit Kepala

Sakit kepala atau nyeri kepala adalah rasa nyeri atau rasa tidak mengenakan diseluruh daerah kepala dengan batas bawah dari dagu sampai ke belakang kepala. Berdasarkan penyebabnya sakit kepala digolongkan sakit kepala primer dan sakit kepala sekunder. Sakit kepala primer adalah nyeri kepala yang tidak jelas kelainan anatomi atau kelainan struktur, yaitu migrain, nyeri kepala tipe tegang, nyeri kepala klaster dan nyeri kepala primer lainnya. Sakit kepala sekunder adalah nyeri kepala yang jelas terdapat kelainan anatomi maupun kelainan struktur dan bersifat kronis progresif, antara lain meliputi kelainan non vaskuler (Hidayati, 2016).

2.7.8. Tekanan Darah Tinggi (*Hipertensi*)

Menurut definisi Organisasi Kesehatan Dunia, *hipertensi* ditandai dengan pembacaan tekanan darah yang melebihi nilai lebih dari 140 (tekanan darah tinggi) dan 90 (tekanan darah rendah) mmHg dengan pengukuran berulang saat orang tersebut sedang dalam kondisi istirahat. *hipertensi* bisa dikategorikan menjadi dua jenis, primer dan sekunder, dengan penyebab yang berbeda yaitu sekitar 90% dari kasus *hipertensi* adalah *hipertensi* primer, tidak ada penyebab yang bisa diidentifikasi dan sebagian besar terkait dengan faktor keturunan, kepribadian, stres mental, dan obesitas. Sedangkan *hipertensi* sekunder yaitu *hipertensi* yang umumnya disebabkan oleh penyakit lain atau perubahan pada kondisi kesehatan, misalnya penyakit ginjal dan gangguan endokrin (gangguan kelenjar endokrin yang mengeluarkan hormon). Tekanan darah pasien akan kembali normal jika penyakit tersebut sudah disembuhkan (Pusdatin Kementrian Kesehatan RI, 2014).

2.7.9. Tekanan Darah Rendah (*Hipotensi*)

Menurut Balai Informasi Teknologi LIPI (2009) *Hipotensi* atau tekanan darah rendah adalah suatu keadaan dimana tekanan darah lebih rendah dari 90/60 mmhg atau tekanan darah cukup rendah sehingga menyebabkan gejala-gejala seperti pusing dan pingsan. Penyebab terjadinya *Hipotensi* dikelompokkan menjadi 3 yaitu:

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.

- a. Curah jantung berkurang, penyebabnya irama jantung abnormal, kerusakan atau kelainan fungsi otot jantung, penyakit katup jantung, emboli pulmoner.
- b. Volume darah berkurang, penyebabnya perdarahan hebat, diare, keringat berlebihan, berkemih berlebihan.
- c. Meningkatnya kapasitas pembuluh darah, penyebabnya syok septik, pemaparan oleh panas, diare, obat-obat vasodilator (nitrat, penghambat kalsium, penghambat ACE).

2.7. Rapid Miner

Rapid miner merupakan sebuah software data mining yang digunakan untuk menganalisis atau mengintegrasikan data sebagai mesin data mining kedalam berbagai bentuk (Zainal dkk, 2015). Tools Rapid miner mempunyai beberapa fitur unik seperti:

1. Data integrasi, *Extract Transform Load* (ETL) analitis, menganalisis dan melaporkan data satu persatu.
2. Menggunakan intuisi *Graphical User Interface* (GUI) yang maksimal untuk proses analisis desainnya
3. Adanya *Repository* untuk proses, data dan manajemen metadata.
4. Mendeteksi eror dan penyelesaian yang cepat
5. Fleksibel dan lengkap dengan ratusan metode tersedia untuk mengintegrasikan data, transformasi data, permodelan dan visualisasi.



Gambar 2.2 Tampilan Tools Rapid Miner 8.1

2.8. Davies Bouldien Index (DBI)

Konsep dari DBI ini adalah meminimalkan jarak rata-rata antar cluster yang terdekat (Ansari dkk, 2011). Tujuan dari DBI ini adalah untuk mengidentifikasi kumpulan kelompok yang mirip tetapi terpisah jauh, adapun formula DBI dapat dilihat pada Persamaan 2.6 berikut (Bhatia dan Dixit, 2012):

$$DB = 1/K \sum_{i=1}^K \max \left[\frac{\text{diam}(c_i) + \text{diam}(c_j)}{d(c_i, c_j)} \right] \quad (2.6)$$

Adapun K merupakan jumlah cluster, d_i dan d_j adalah diameter dari cluster C_i dan C_j , sedangkan d adalah jarak rata-rata antar cluster. Nilai rata-rata yang lebih kecil menunjukkan kualitas cluster yang baik (Bhatia dan Dixit, 2012).

2.9. Penelitian Terdahulu

Melihat pentingnya kesehatan masyarakat, di Indonesia sudah banyak melakukan penelitian, adapun teknik yang diambil dalam penyelesaian masalah adalah teknik data mining, selain berfokus pada topik permasalahan penelitian terdahulu juga mengambil penerapan algoritma *K-Means* dan *Association Rule* pada pemecahan kasus yang lain untuk melihat bagaimana performanya, sehingga algoritma ini layak diterapkan ke dalam studi kasus, adapun penelitian terdahulu yang menjadi dasar tugas akhir ini adalah:

1. Penelitian Sulastris dan Gufroni (2017) dengan judul penerapan data mining dalam pengelompokan penderita penyakit *Thalassaemia* berhasil mengelompokan Penderita *Thalassaemia* menggunakan Algoritma *K-Means* menjadi 3 kelompok.
2. Penelitian Wardhani (2016) berjudul implementasi algoritma *k-means* untuk pengelompokan penyakit pasien pada puskesmas kaje pekalongan. *K-Means* digunakan untuk menentukan konsistensi data kesehatan serta mampu menggali informasi tersembunyi dari kumpulan data multidimensi yang telah diperoleh, selain itu pengestrakan data yang terhubung dengan data lain juga dapat dilakukan menggunakan algoritma ini.

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.

3. Penelitian Meisida (2014) dengan judul *K-means* untuk klasifikasi penyakit karies gigi. Pada penelitiannya *K-means* digunakan sebagai metode pengklasifikasian untuk *cluster* kelas karies gigi anatomi Black karena memiliki nilai eror yang kecil, tetapi data harus dinormalisasikan terlebih dahulu.
4. Budiman dkk (2017) yang menerapkan metode klasifikasi dan *Association rule* untuk klasifikasi kecenderungan penyakit mata.
5. Tampubolon (2013) mengimplementasikan algoritma apriori pada sistem persediaan alat-alat kesehatan untuk menemukan kecenderungan pola kombinasi *itemsets* sehingga dapat dijadikan sebagai informasi dalam pengambilan keputusan untuk mempersiapkan stok jenis barang apa yang diperlukan kemudian.
6. Mustajib dan Mustakim (2016) yang menggunakan *association rule mining* sebagai pendukung keputusan tata letak dan perencanaan ketersediaan obat di apotek rumah sakit umum daerah tembilahan

Dari beberapa penelitian yang telah dijabarkan diatas maka dapat diambil kesimpulan bahwasannya metode data mining menggunakan teknik *clustering analysis* menggunakan algoritma *K-Means* dan analisis *Association Rule* dapat digunakan untuk membantu penyelesaian ataupun sebagai analisis pola pengelompokan data kesehatan masyarakat.