

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

BAB II LANDASAN TEORI

2.1. Diabetes

Diabetes adalah suatu penyakit dimana tubuh tidak dapat menghasilkan insulin (hormon pengatur gula darah) atau insulin yang dihasilkan tidak mencukupi atau insulin tidak bekerja dengan baik. Oleh karena itu akan menyebabkan gula darah meningkat saat diperiksa. Pada penderita DM, terjadi gangguan keseimbangan transportasi glukosa ke dalam sel glukosa yang disimpan di hati. Akibatnya kadar glukosa darah meningkat, dan keluar melalui urin, sehingga urin mengandung gula. Hiperglikemia terjadi karena dua penyebab, pertama pankreas tidak mampu lagi memproduksi insulin. Kedua, sel tidak memberi respon pada kerja insulin sebagai kunci untuk membuka pintu sel sehingga glukosa tidak dapat masuk ke dalam sel. Diagnosis DM diperkirakan dengan adanya gejala khas berupa polifagia, poliuria, polidipsia, lemas, dan berat badan turun. Gejala yang mungkin diketahui oleh pasien adalah kesemutan, gatal, mata kabur, dan gatal pada daerah kemaluan (WHO, 2016).

Bila penderita DM tidak patuh terhadap apa yang menjadi pantangan dan jika tidak rajin memeriksakan kadar gula darah maka ada kemungkinan akan menjadi dasar dari komplikasi penyakit. Komplikasi akut antara lain hipoglikemia, ketoasidosis, koma hiperosmolar non ketotik (Anna Maria Sirait, 2015). Di otak, misalnya terjadi stroke dan di mata menimbulkan katarak. Komplikasi kronik DM antara lain adalah kerusakan saraf, penyakit jantung, penyakit hati, penyakit ginjal, gangguan saluran pencernaan, penyakit kulit dan rentan terhadap bakteri dan virus. Kaki penderita akan mudah kesemutan, bila ada luka proses sembuhnya akan lama dan bisa menimbulkan infeksi, bahkan bisa berdampak amputasi. Seorang penderita DM sangat dianjurkan untuk memeriksakan diri dan berobat ke dokter secara teratur untuk membantu dalam pengobatan dan pencegahan komplikasi.

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.

Kontrol yang rutin dan sesuai jadwal akan membuat kadar glukosa darah selalu berada dalam batas yang normal. Ada 2 jenis tipe utama dalam Diabetes yaitu:

1. Diabetes tipe 1

Diabetes tipe 1 disebabkan oleh kerusakan sel β pankreas sehingga produksi insulin tidak memadai. Gambaran klinis dan laboratorium telah banyak dipelajari. Sehingga penderita harus menggunakan suntikan insulin dalam mengatur gula darahnya. Sebagian besar penderitanya adalah anak-anak & remaja. (Haryudi Aji, 2009)

2. Diabetes tipe 2

Diabetes tipe 2 ini terjadi ketika pankreas tidak cukup membuat insulin untuk menjaga level gula darah tetap normal, seringkali disebabkan tubuh tidak merespon dengan baik terhadap insulin tersebut. Kebanyakan orang tidak menyadari telah menderita diabetes tipe 2, walaupun keadaannya sudah menjadi sangat serius. Diabetes tipe 2 sudah menjadi umum dialami didunia maupun di Indonesia, dan angkanya terus bertambah akibat gaya hidup yang tidak sehat, kegemukan dan malas berolahraga. (Awad, 2011)

Kontrol kadar gula darah merupakan suatu pengendalian glukosa pasien Diabetes Melitus. Kontrol ini minimal dilakukan setiap 3 bulan sekali yang meliputi pemeriksaan kontrol kadar gula darah puasa, kadar gula acak, trigliserida, HDL dan LDL. Sehingga tidak bisa dikontrol kadar kandungannya dalam darah (Rachmawati, 2015). Sehingga tidak melewati ambang batas normal sebagai berikut :

1. GDA kurang dari 140 mg/dL
2. GDP kurang dari 100 mg/dL
3. Trigliserida kurang dari 150 mg/dL
4. HDL lebih dari 60 mg/dL
5. LDL kurang dari 100 mg /dL

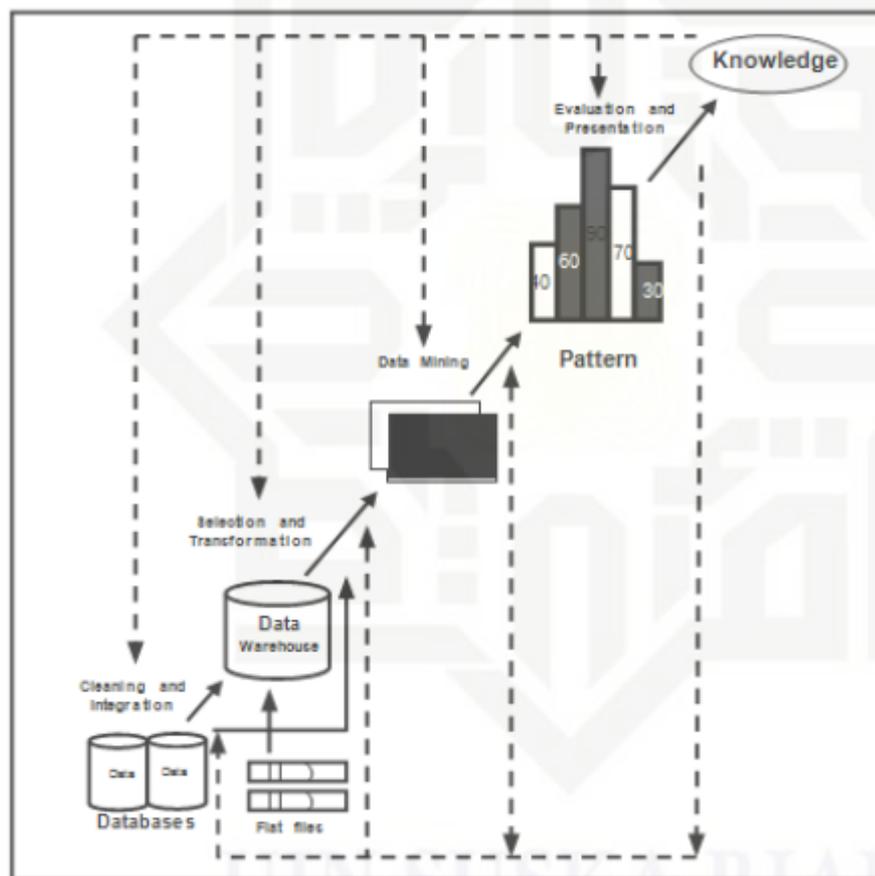
Hak Cipta Dilindungi Undang-Undang

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.

2.2 KDD (*Knowledge Discovery In Database*)

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 (Marisa, 2013). Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah data mining.

Tahapan-tahapan KDD



Gambar 2.1 Tahapan Tahapan dari KDD (Marissa, 2013)

Proses KDD secara garis besar dapat dijelaskan sebagai berikut:

1. Data Selection

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

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.

dari basis data. Dimana data yang relevan dengan tugas analisis dikembalikan ke dalam database.

2. Pre- processing

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses pembersihan pada data yang menjadi focus KDD. Proses pembersihan mencakup antara lain membuang duplikasi data, memeriksa data yang inkosisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi).

3. Transformation

Transformasi dilakukan pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses Transformasi 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. Interpretation / Evaluation

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

2.3 Data Mining

Data Mining merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut. Data Mining dapat diterapkan pada berbagai bidang yang mempunyai sejumlah data. Menurut Daryl Pregibon bahwa “Data mining

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.

adalah campuran dari statistic, kecerdasan buatan, dan riset basis data” yang masih berkembang (Sulastri, 2017)

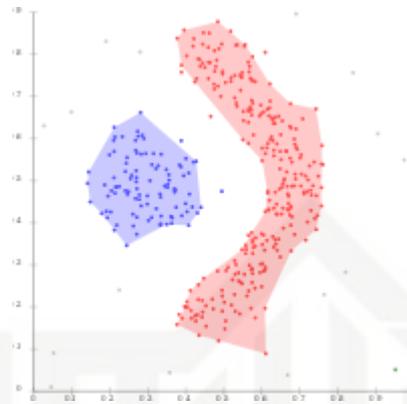
Secara sederhana, data mining atau penambangan data dapat didefinisikan sebagai proses seleksi, eksplorasi, dan pemodelan dari sejumlah besar data untuk menemukan pola atau kecenderungan yang biasanya tidak disadari keberadaannya. Data mining dapat dikatakan sebagai proses mengekstrak pengetahuan dari sejumlah besar data yang tersedia. Pengetahuan yang dihasilkan dari proses data mining harus baru, mudah dimengerti, dan bermanfaat. Dalam data mining, data disimpan secara elektronik dan diproses secara otomatis oleh komputer menggunakan teknik dan perhitungan tertentu (Pramadhani, 2014)

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai basis data besar. Data mining, sering juga disebut Knowledge Discovery in Database atau disingkat menjadi KDD, adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar (Anggraeni, 2013).

Berdasarkan maksud yang didapat dari beberapa pengertian diatas tujuan akhir dari *Data Mining* adalah sebuah informasi, yang mana informasi tersebut bisa dalam bentuk pola atau aturan tertentu. Agar mendapatkan pola atau aturan tersebut maka diterapkan sebuah metode pada prosesnya. Sebuah proses data mining dapat dibagi menjadi dua kategori utama, tergantung pada target hasil dan metode belajar, kategori tersebut adalah proses yang diawasi (*Supervised*) dan tanpa pengawasan (*Unsupervised*).

Unsupervised atau tanpa pengawasan merupakan proses analisis data dimana data yang dianalisis tidak mempunyai atribut label. Oleh karena itu, data mining *unsupervised* bertujuan untuk menemukan pola berulang dan kedekatan dan kemiripan data. Dengan kata lain metode belajar tanpa adanya latihan dan label.

2.4 Clustering



Gambar 2.2 Clustering

Pada dasarnya clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output. Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering* (Ong, 2013).

Clustering dengan pendekatan partisi atau sering disebut dengan *non-hierarchical clustering* mengelompokkan data dengan memilah-milah data yang dianalisa ke dalam cluster-cluster yang ada. Clustering dengan pendekatan hirarki atau sering disebut dengan *hierarchical clustering* mengelompokkan data dengan membuat suatu hirarki berupa dendogram dimana data yang mirip akan ditempatkan pada hirarki yang berdekatan dan yang tidak pada hirarki yang berjauhan.

2.5. K-Means

K-Means adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan

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.

sistem partisi. Metode k-means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Dengan kata lain, metode ini berusaha untuk meminimalkan variasi antar data yang ada di dalam suatu cluster dan memaksimalkan variasi dengan data yang ada di cluster lainnya.

Langkah – langkah K-Means sebagai berikut :

1. Menentukan banyaknya *cluster*

Untuk melakukan *Clustering* dengan algoritma *K-Means* langkah yang pertama kali yaitu menentukan banyak *cluster* yang akan dibentuk. Pada penelitian ini, terdapat dua *cluster*.

2. Menentukan titik pusat (*centroid*)

Setelah menentukan banyak *cluster* yang akan dibentuk, langkah selanjutnya yaitu dengan menentukan titik pusat (*centroid*) dari tiap-tiap *cluster* dan pengambilan titik pusat dilakukan secara acak.

3. Menghitung jarak setiap objek ke titik pusat (*centroid*)

Selanjutnya setelah menentukan titik pusat adalah dengan menghitung jarak setiap data dengan titik pusat yang sudah ditentukan sebelumnya. Rumus untuk menghitung jarak setiap objek ke titik pusat adalah :

$$D(X_j - C_j) = \sqrt{\sum_{j=0}^n (X_j - C_j)^2}$$

Keterangan

X = adalah dataset Pasien.

C = titik pusat (*centroid*) yang sudah ditentukan secara acak.

4. Kelompokkan data berdasarkan jarak terpendeknya antara data dengan *centroid* menjadi sebuah kelompok *cluster* dalam tabel baru.

5. Hitung rata-rata tiap kelompok *cluster* yang terbentuk untuk dijadikan sebagai *centroid* yang baru dan diulangi mencari jarak terpendek antara data dan *centroid* apabila *centroid* berubah dan perhitungan akan berhenti apabila *centroid* tidak mengalami perubahan.

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.

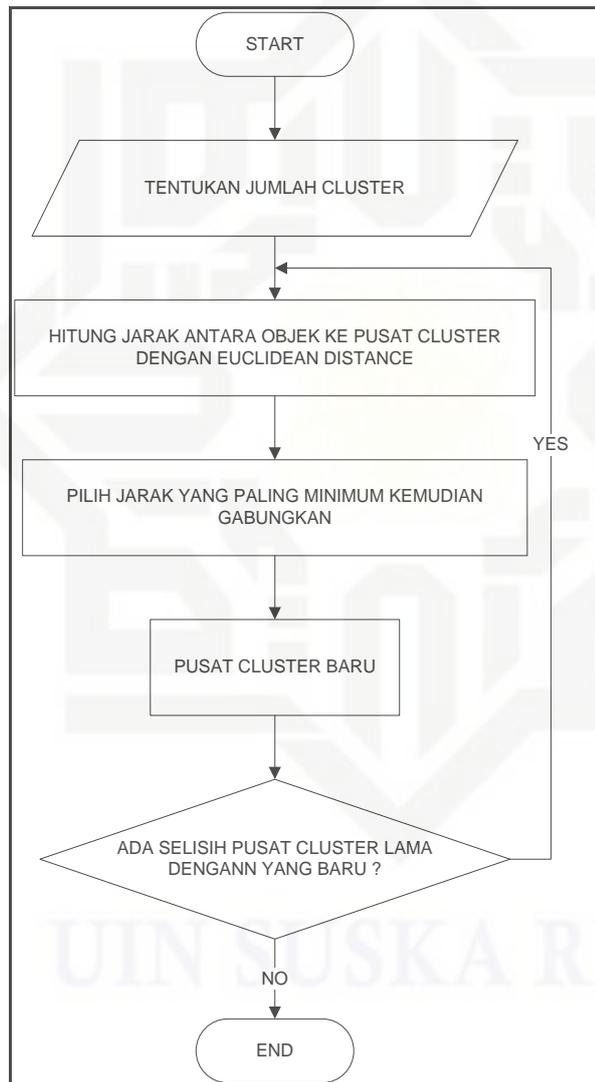
$$C_{kj} = \frac{\sum_{i=1}^p x_{ij}}{p}$$

Keterangan

x_{ij} = \in kluster ke - k

p = banyaknya anggota kluster ke k

Flowchart K-Means



Gambar 2.3 Flowchart K-Means

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.

2.6 Penelitian Terkait

Sebelum melanjutkan penelitian ini, penulis melihat beberapa referensi untuk dijadikan acuan. Adapun beberapa penelitian terdahulu yang menjadi bahan dalam penelitian ini adalah:

Tabel 2.1 Penelitian terkait

Penulis	Judul	Kesimpulan	Tahun
Debora C Sihombing	Matematika untuk Mendeteksi Diabetes Mellitus Tipe 1	Konsentrasi glukosa darah dipengaruhi dan dikendalikan oleh berbagai jenis hormon, diantaranya faktor dominan adalah hormon insulin. Hormon insulin disekresi oleh pankreas.	2007
Fuji Rahmawati	Skrining Diabetes Mellitus Gestasionan dan Fator yang Mempengaruhinya	Prevalensi angka kejadian DM Gestasional di wilayah kerja Puskesmas Simpang Timbangan Kabupaten Ogan Ilir Tahun 2016 sebanyak 5,6%.	2016
Silvi Agustina	Clustering kualitas beras secara berdasarkan ciri fisik menggunakan Metode K-Means	Didapatan 3 cluster yang menunjukkan kualitas beras dengan akurasi hingga 92%	2013

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.

Fina Nasari	Penerapan Algoritma K-Means Clustering Untuk Pengelompokkan Penyebaran Diare Di Kabupaten Langkat	Didapatkan penyebaran penyakit diare di kabupaten langkat sehingga dapat di tangani penyebarannya.	2016
-------------	---	--	------

