

BAB II

LANDASAN TEORI

1.1 Tinjauan Studi

Beberapa penelitian tentang deteksi penyakit Jantung sebagai berikut :

- 1) Penelitian yang dilakukan oleh Subbalakashmi (2011), rumah sakit sudah mulai menggunakan sistem informasi pendukung keputusan, tetapi sebagian besar terbatas. Sistem informasi yang dibuat digunakan untuk penagihan, mengingat catatan pasien, memprediksi kemungkinan pasien yang mendapatkan penyakit jantung. Keputusan klinis sering dibuat berdasarkan pada intuisi dokter dan pengalaman daripada pengetahuan yang kaya data tersembunyi dalam database. Praktek ini menyebabkan bias yang tidak diinginkan, kesalahan dan biaya medis yang berlebihan yang mempengaruhi kualitas pelayanan yang diberikan kepada pasien. Penelitian ini telah mengembangkan sistem pendukung keputusan dalam prediksi penyakit jantung dengan menggunakan teknik pemodelan data mining, yaitu *naive bayes*. Dengan menggunakan profil medis seperti *age, sex, chest pain type, fasting blood sugar, rest ecg, exercise, slope, CA, Threst blood pressure, Thal, Chol, Thalach, old peak, smoking* dapat memprediksi probabilitas pasien mendapatkan penyakit jantung. Hal ini diimplementasikan sebagai aplikasi web berbasis kuesioner dan dapat digunakan sebagai media pelatihan untuk melatih perawat dan mahasiswa kedokteran untuk mendiagnosa pasien dengan penyakit jantung.

- 2) Rajkumar dan Reena (2010) dalam penelitiannya mengklasifikasi data didasarkan pada algoritma *supervised machine learning* yang menghasilkan akurasi, dan waktu yang dibutuhkan untuk membangun algoritma. *Tanagra tool* digunakan untuk mengklasifikasikan data dan data yang dievaluasi dengan menggunakan 10 kali lipat validasi silang dan hasilnya dibandingkan. Algoritma yang dibandingkan adalah *naive bayes, decision list, k-nn* dan

parameter yang digunakan meliputi: *id, age, sex, chest pain type, painexer, relrest, chest pain location, resting blood pressure, chol, famhist, restecg, ekgmo, thaldur, thalach, thalrest, num, vessel*. Algoritma *naive bayes* adalah waktu terbaik untuk data set pengolahan dan menunjukkan kinerja yang lebih baik dalam prediksi akurasi. Waktu yang dibutuhkan untuk menjalankan data untuk hasil yang cepat bila dibandingkan dengan algoritma lainnya. Atribut ini sepenuhnya diklasifikasikan oleh algoritma *naive bayes* dan memberikan 52,33% dari hasil yang akurat, decision list (52%), k-nn (45,67%). Dari hasil di atas algoritma *naive bayes* memainkan peran penting dalam membentuk peningkatan akurasi klasifikasi dataset.

- 3) Setelah Rajkumar dan Reena, penelitian yang dilakukan oleh Salem (2010), deteksi penyakit jantung dengan algoritma *K-Nearest Neighbour* dengan metode *Case Based Reasoning* (CBR) yang mengembangkan sistem pakar berbasis kasus prototipe untuk mendukung diagnosis penyakit jantung. Sebanyak 110 kasus telah dikumpulkan untuk dikelompokkan ke dalam 4 kategori penyakit jantung yaitu; mitral stenosis, sisi kiri gagal jantung, angina pectoris stabil dan hipertensi. Setiap kasus mengandung 207 atribut tentang kedua data demografi dan klinis. Setelah menghapus kasus diduplikasi, sistem telah diuji coba dari 42 kasus untuk pasien jantung di Mesir. Analisis statistik telah dilakukan untuk menentukan pentingnya nilai-nilai dari fitur kasus. dua pengambilan strategi yang diselidiki yaitu; induksi dan K-nn. Hasil menunjukkan bahwa K-nn adalah lebih baik daripada strategi induksi. Ahli jantung mengevaluasi kinerja sistem dengan menguji praktis untuk 13 kasus baru dan sistem berhasil dalam memperkirakan diagnosis yang benar.
- 4) Pada tahun 2011 penelitian yang dilakukan oleh Srinivas (2011) deteksi penyakit jantung dengan menggunakan algoritma *decision tree, multi layer perceptron (MLP), bayesian model*, dan *support vektor machine* (SVM) diteliti dari keakuratan dan sensitivitasnya dengan parameter : *id, age, sex, chest pain type, fasting blood sugar, restecg, exercise, slope, CA, trest blood pressure, cholestrol, thalach, old peak*. Hasil penelitian menunjukkan

decision tree (akurasinya 82%, sensitivitasnya 88%), *multi layer perceptron* (MLP) (akurasinya 89,2%, sensitivitasnya 92%), *baysian model* (akurasinya 82%, sensitivitasnya 87%), SVM (akurasinya 83,5 %, sensitivitasnya 88%).

Dari penelitian yang pernah dilakukan dengan menggunakan parameter yang ada, algoritma yang digunakan dan hasil akurasi yang didapat. Maka, pada penelitian ini akan mendeteksi penyakit jantung menggunakan algoritma data mining *naive bayes* untuk meningkatkan akurasi dalam deteksi penyakit jantung. Data set yang akan digunakan dari *hasil Medical Check Up* di laboratorium klinik dengan parameter yang digunakan, meliputi usia, jenis kelamin, kolesterol total, HDL, LDL, trigliserid, rasio kolestrol, irama, frekuensi/menit, zone transisi, axis (posisi listrik), gelombang P, gelombang QRS, interval P-R, interval QRS, voltage , segmen ST, gelombang T, gelombang U, cepat capek disertai nyeri dada dan sesak, merokok, olahraga, hipertensi, tekanan darah.

1.2 Tinjauan Pustaka

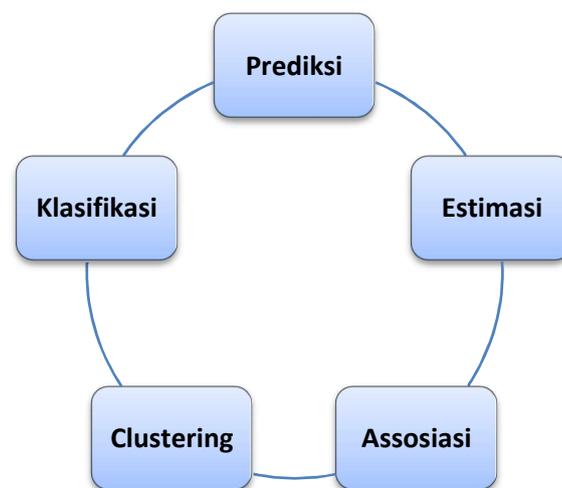
1.2.1 Data Mining

Data mining menurut Tan dalam Prasetyo (2012:2) adalah proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Sedangkan menurut Budi Santoso (2012) Data mining sering disebut *knowledge discovery in database* (KDD), yaitu kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data set berukuran besar . Keluaran dari data mining bisa dipakai untuk memperbaiki pengambilan keputusan di masa yang akan datang berdasarkan informasi yang diperoleh dari masa lalu. Data yang ada bisa berupa data pasien, mahasiswa, pemohon kredit, gambar atau pengukuran lain yang perlu diklasifikasikan atau relasi atributnya.



Bagan 2.1 Proses Data Mining

Data mining dapat dikelompokkan menjadi: prediksi, estimasi, asosiasi, clustering, dan klasifikasi. Data mining erat kaitannya dengan *machine learning*. *Machine learning* adalah suatu area dalam *artificial intelligence* atau kecerdasan buatan yang berhubungan dengan pengembangan teknik-teknik yang bisa diprogramkan dan belajar dari data masa lalu. *Machine learning* menjadi alat analisis dalam data mining. Gambar 1 merupakan pengelompokan data mining. Kasus klasifikasi yang bersifat linear maupun non linear menjadi kajian yang menarik. Banyak metode yang dikembangkan untuk kelinearan data. Metode yang demikian tentu akan mempunyai performansi yang sesuai dengan kasusnya linear atau non linear.



Gambar 2. 1 Pengelompokan *Data Mining*

Menurut Kusri (2009:11) data mining dibagi beberapa kelompok sebagai berikut :

1) Estimasi

Estimasi hampir sama dengan klasifikasi kecuali variabel target estimasi lebih numerik daripada kategori. Model dibangun *records* lengkap menyediakan nilai dari variabel target sebagai nilai prediksi. Contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks massa tubuh, 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.

Contoh lain yaitu estimasi nilai indeks prestasi kumulatif mahasiswa dengan melihat indeks prestasi mahasiswa tersebut saat mengikuti program sarjana.

2) Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa yang akan datang. Contoh prediksi : Prediksi harga beras dalam tiga bulan yang akan datang, prediksi persentase kenaikan kecelakaan lalu lintas tahun depan jika batas bawah kecepatan dinaikkan.

3) 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 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 penyakit apa.

4) Clustering

Klustering merupakan pengelompokkan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek – objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan dengan *record-record* dalam kluster lain.

Pengklusteran berbeda dengan klasifikasi yaitu tidak adanya variabel target dalam pengklusteran. Pengklusteran tidak mencoba untuk melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target. Akan tetapi, algoritma pengklusteran mencoba untuk melakukan pembagian terhadap keseluruhan data

menjadi kelompok – kelompok yang memiliki kesamanaan (homogen), yang mana kemiripan *records* dalam stu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *records* dalam kelompok lain akan bernilai minimal. Contoh pengklusteran :

- 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 finansial dalam baik dan mencurigakan.

5) Assosiasi

Tugas assosiasi dalam data mining adalah menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Contoh assosiasi:

- 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 secara bersamaan.

1.2.2 Naive Bayes

Bayes menurut Kusri,dkk (189:2009) merupakan pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. *Bayes* memiliki akurasi dan kecepatan yang sangat tinggi saat diaplikasi ke dalam database dengan data yang besar. *Naive bayes* merupakan algoritma yang dapat meminimalkan tingkat kesalahan dibandingkan dengan semua pengklasifikasi lainnya. Namun, dalam praktek ini tidak selalu terjadi, karena untuk ketidak akuratan dalam asumsi yang dibuat untuk penggunaannya class yang tidak utuh dan kurangnya data probabilitas yang tersedia.

Pengklasifikasi Bayesian digunakan untuk pembenaran teoritis untuk pengklasifikasi lain yang tidak secara eksplisit menggunakan teorema Bayes. Berikut *teorema bayes* :

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

Naive bayes adalah penyederhanaan dari *teorema bayes*. Berikut rumus *naive bayes* :

$$P(H|X) = P(H) \prod P(X_i|H) \quad (2)$$

Keterangan :

- X : data dengan class yang belum diketahui
- H : hipotesis data x merupakan suatu class spesifik
- P(H|X) : probabilitas hipotesis H berdasarkan kondisi X (posteriori probability)
- P(H) : probabilitas hipotesis H (prior probability)
- P(X|H) : probabilitas X berdasar kondisi pada hipotesis H
- P(X) : probabilitas dari X

Klasifikasi dengan *naive bayes* bekerja berdasarkan teori probabilitas yang memandang semua fitur dari data sebagai bukti dalam probabilitas. Menurut Prasetyo (2012:72) Berikut beberapa keuntungan dari algoritma klasifikasi *naive bayes* :

- 1) Kuat terhadap pengisolasi gangguan pada data
- 2) Jika terjadi kasus missing value ketika proses komputasi sedang berlangsung, maka objek tersebut akan diabaikan
- 3) Tangguh terhadap data tidak relevan

Sebagai contoh, data hasil rekap medical check up tahun 2010 sejumlah 15 pasien (terlampir).

Terdapat dua class dari klasifikasi yang dibentuk, yaitu:

C1 = kesimpulan = normal

C2 = kesimpulan = kelainan

Misal terdapat data X (belum diketahui *class*-nya)

X = (usia = "52", JK = "1", kolesterol total = "199", HDL = "38", LDL = "140", trigliserid = "86", rasio kolesterol = "5.1", irama = "sinus", frekwensi = "55", zona transisi = "114", Axis = "LAD", Gel. P = "normal", Gel. QRS = "normal", segmen ST = "ISOELECTRIC", Gel. T = "UPWARD", cepat capek disertai nyeri dada dan sesak = "tidak", merokok = "ya", olahraga = "3 kali", hipertensi = "tidak", tekanan darah = "140/80")

Penyelesaian:

Dibutuhkan untuk memaksimalkan $P(X|C_i)P(C_i)$ untuk $i = 1,2$

$P(C_i)$ merupakan *prior probability* untuk setiap class berdasar data contoh.

$$P(\text{kesimpulan} = \text{"normal"}) = 10/15 = 0.67$$

$$P(\text{kesimpulan} = \text{"kelainan"}) = 5/15 = 0.33$$

Hitung $P(X|C_i)$, untuk $i=1,2$

$$P(\text{usia} = \text{"52"} \mid \text{kesimpulan} = \text{"normal"}) = 1/10 = 0.1$$

$$P(\text{usia} = \text{"52"} \mid \text{kesimpulan} = \text{"kelainan"}) = 2/5 = 0.4$$

$$P(\text{jenis kelamin} = \text{"1"} \mid \text{kesimpulan} = \text{"normal"}) = 10/10 = 1$$

$$P(\text{jenis kelamin} = \text{"1"} \mid \text{kesimpulan} = \text{"kelainan"}) = 5/5 = 1$$

$$P(\text{kolesterol total} = \text{"199"} \mid \text{kesimpulan} = \text{"normal"}) = 1/10 = 0.1$$

$$P(\text{kolesterol total} = \text{"199"} \mid \text{kesimpulan} = \text{"kelainan"}) = 0/5 = 0$$

$$P(\text{HDL} = \text{"38"} \mid \text{kesimpulan} = \text{"normal"}) = 1/10 = 0.1$$

$$P(\text{HDL} = \text{"38"} \mid \text{kesimpulan} = \text{"kelainan"}) = 1/5 = 0.2$$

$$P(\text{LDL} = \text{"140"} \mid \text{kesimpulan} = \text{"normal"}) = 1/10 = 0.1$$

$$P(\text{LDL} = \text{"140"} \mid \text{kesimpulan} = \text{"kelainan"}) = 0/5 = 0$$

$$P(\text{trigliserid} = \text{"86"} \mid \text{kesimpulan} = \text{"normal"}) = 1/10 = 0.1$$

$$P(\text{trigliserid} = \text{"86"} \mid \text{kesimpulan} = \text{"kelainan"}) = 0/5 = 0$$

$$P(\text{rasio kolesterol} = "5.1" \mid \text{kesimpulan} = "normal") = 2/10 = 0.2$$

$$P(\text{rasio kolesterol} = "5.1" \mid \text{kesimpulan} = "kelainan") = 0/5 = 0$$

$$P(\text{irama} = "normal" \mid \text{kesimpulan} = "normal") = 10/10 = 1$$

$$P(\text{irama} = "normal" \mid \text{kesimpulan} = "kelainan") = 5/5 = 1$$

$$P(\text{frekwensi/menit} = "55" \mid \text{kesimpulan} = "normal") = 2/10 = 0.2$$

$$P(\text{frekwensi/menit} = "55" \mid \text{kesimpulan} = "kelainan") = 0/5 = 0$$

$$P(\text{zona transisi} = "V4" \mid \text{kesimpulan} = "normal") = 1/10 = 0.1$$

$$P(\text{zona transisi} = "V4" \mid \text{kesimpulan} = "kelainan") = 1/5 = 0.2$$

$$P(\text{Axis} = "LAD" \mid \text{kesimpulan} = "normal") = 1/10 = 0.1$$

$$P(\text{Axis} = "LAD" \mid \text{kesimpulan} = "kelainan") = 3/5 = 0.6$$

$$P(\text{Gelombang P} = "normal" \mid \text{kesimpulan} = "normal") = 10/10 = 1$$

$$P(\text{Gelombang P} = "normal" \mid \text{kesimpulan} = "kelainan") = 5/5 = 1$$

$$P(\text{Gelombang QRS} = "normal" \mid \text{kesimpulan} = "normal") = 8/10 = 0.8$$

$$P(\text{Gelombang QRS} = "normal" \mid \text{kesimpulan} = "kelainan") = 5/5 = 1$$

$$P(\text{interval P-R} = "0.14" \mid \text{kesimpulan} = "normal") = 2/10 = 0.2$$

$$P(\text{interval P-R} = "0.14" \mid \text{kesimpulan} = "kelainan") = 0/5 = 0$$

$$P(\text{voltage} = "1" \mid \text{kesimpulan} = "normal") = 10/10 = 1$$

$$P(\text{voltage} = "1" \mid \text{kesimpulan} = "kelainan") = 5/5 = 1$$

$$P(\text{segmen ST} = "ISOELECTRIC" \mid \text{kesimpulan} = "normal") = 10/10 = 1$$

$$P(\text{segmen ST} = "ISOELECTRIC" \mid \text{kesimpulan} = "kelainan") = 5/5 = 1$$

$$P(\text{Gelombang T} = "UPWARD" \mid \text{kesimpulan} = "normal") = 9/10 = 0.9$$

$$P(\text{Gelombang T} = "UPWARD" \mid \text{kesimpulan} = "kelainan") = 5/5 = 1$$

$$P(\text{cepat capai disertai nyeri dada dan sesak} = \text{"tidak"} \mid \text{kesimpulan} = \text{"normal"}) = 9/10 = 0.9$$

$$P(\text{cepat capai disertai nyeri dada dan sesak} = \text{"tidak"} \mid \text{kesimpulan} = \text{"kelainan"}) = 3/5 = 0.6$$

$$P(\text{merokok} = \text{"ya"} \mid \text{kesimpulan} = \text{"normal"}) = 2/10 = 0.2$$

$$P(\text{merokok} = \text{"ya"} \mid \text{kesimpulan} = \text{"kelainan"}) = 1/5 = 0.2$$

$$P(\text{olahraga} = \text{"3 kali"} \mid \text{kesimpulan} = \text{"normal"}) = 3/10 = 0.3$$

$$P(\text{olahraga} = \text{"3 kali"} \mid \text{kesimpulan} = \text{"kelainan"}) = 1/5 = 0.2$$

$$P(\text{hipertensi} = \text{"tidak"} \mid \text{kesimpulan} = \text{"normal"}) = 9/10 = 0.9$$

$$P(\text{hipertensi} = \text{"tidak"} \mid \text{kesimpulan} = \text{"kelainan"}) = 3/5 = 0.6$$

$$P(\text{tekanan darah} = \text{"140/80"} \mid \text{kesimpulan} = \text{"normal"}) = 1/10 = 0.1$$

$$P(\text{tekanan darah} = \text{"140/80"} \mid \text{kesimpulan} = \text{"kelainan"}) = 0/5 = 0$$

$$P(X \mid \text{kesimpulan} = \text{"normal"})$$

$$= 0.1 \times 1 \times 0.1 \times 0.1 \times 0.1 \times 0.1 \times 0.1 \times 0.2 \times 1 \times 0.2 \times 0.1 \times 0.1 \times 1 \times 0.8 \times 0.2 \times 0.4 \times 1 \times 1 \times 0.9 \times 0.9 \times 0.2 \times 0.3 \times 0.2 = 0.12$$

$$P(X \mid \text{kesimpulan} = \text{"kelainan"})$$

$$= 0.4 \times 1 \times 0 \times 0.2 \times 0 \times 0 \times 0.2 \times 0 \times 1 \times 0 \times 0.2 \times 0.6 \times 1 \times 0 \times 0.2 \times 1 \times 1 \times 1 \times 0.6 \times 0.2 \times 0.2 \times 0.6 \times 0 = 0$$

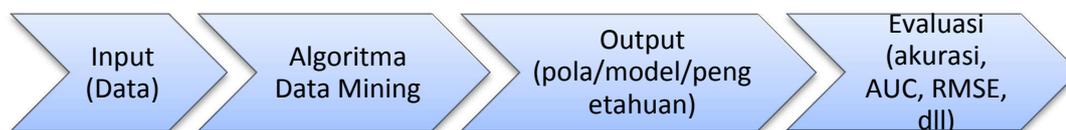
$$P(X \mid \text{kesimpulan} = \text{"normal"}) P(\text{kesimpulan} = \text{"normal"})$$

$$= 0.12 \times 0.67 = 0.08$$

Kesimpulan = "normal"

1.2.3 Evaluasi Kinerja Klasifikasi

Dalam permasalahan klasifikasi, sangat penting untuk menentukan nilai klasifikasi yang benar dan tidak benar. Klasifikasi akurasi mengukur cara kerja klasifikasi dalam mengidentifikasi objek. Dengan menggunakan tabel 2.1 yang merupakan hasil dari proses klasifikasi. Oleh sebab itu, evaluasi kinerja model klasifikasi berdasarkan objek yang benar dan prediksi yang tidak benar dengan menggunakan *confusion matrix* yang menjelaskan detail break down dari klasifikasi yang terlewat (*misclassifications*) (Gronescu dalam Widiastuti (2011:22) . Class prediksi ditampilkan pada matriks yang paling atas dan class observasi ditampilkan paling bawah sebelah kiri. Masing-masing angka menunjukkan kasus yang diberikan di class observasi ditandai dengan model yang diberikan di class prediksi.



Bagan 2.2 Tahapan Proses Data mining

Confusion matrix pertama kali dikenalkan oleh kohavi dan Provost pada tahun 1998 dengan membagi sistem klasifikasi ke dalam dua kelas yaitu class observed dan class prediksi. Berikut tabel 2.1 tabel *confusion matrix* untuk klasifikasi menurut Gronescu.

Tabel 2. 1 *Confusion Matrix* untuk 2 class

CLASSIFICATION		Predicted Class	
		Yes	No
Observed Class	Yes	A <i>(true positive-TP)</i>	B <i>(false negative-FN)</i>
	No	C <i>(false positive-FP)</i>	D <i>(true negative-TN)</i>

Berikut rumus akurasi :

$$= \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

Keterangan :

- ✓ TP(*true positive*)= jumlah parameter yang benar yang termasuk ke dalam penyakit jantung (terdeteksi).
- ✓ FN(*false negative*) = jumlah parameter yang tidak terdeteksi
- ✓ FP(*false positive*) = Jumlah parameter yang salah tetapi terdeteksi
- ✓ TN (*true negative*) = jumlah parameter yang salah deteksi tetapi salah

Precision adalah rasio jumlah parameter yang benar yang termasuk ke dalam penyakit jantung (*true positive*) dengan jumlah parameter yang benar yang termasuk ke dalam penyakit jantung (*true positive*) dan jumlah parameter yang salah tetapi benar (*false positive*). *Recall rate* adalah rasio jumlah parameter yang benar terdeteksi sebagai penyakit jantung (*true positive*) dengan jumlah parameter yang benar terdeteksi sebagai penyakit jantung (*true positive*) dan jumlah parameter yang tidak terdeteksi sebagai penyakit jantung (*false negative*).

Adapun rumus untuk menghitung precision dan recall, sebagai berikut:

$$= \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

$$= \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

1.2.4 Penyakit Jantung

Jantung merupakan suatu organ otot berongga yang terletak di pusat dada. Pada bagian kanan dan kiri jantung memiliki ruang sebelah atas (*atrium*) yang berfungsi untuk mengumpulkan darah dan ruang sebelah bawah (*ventrikel*) yang berfungsi mengalirkan darah keluar. Fungsi utama jantung adalah menyediakan oksigen ke seluruh tubuh dan membersihkan tubuh dari hasil metabolisme berupa karbondioksida.

Penyakit jantung merupakan gangguan yang terjadi pada sistem pembuluh darah besar sehingga menyebabkan jantung dan peredaran darah tidak berfungsi sebagaimana mestinya. Penyakit-penyakit yang berhubungan dengan organ jantung dan pembuluh darah antara lain: gagal jantung, jantung koroner, dan jantung rematik (Sutanto : 2010). Penyakit jantung koroner adalah penyakit jantung dan pembuluh darah yang disebabkan karena penyempitan arteri koroner. Penyempitan pembuluh darah terjadi karena proses *aterosklerosis* atau *spasme* atau kombinasi keduanya. *Aterosklerosis* yang terjadi karena timbunan kolesterol dan jaringan ikat pada dinding pembuluh darah secara perlahan-lahan, hal ini sering ditandai dengan keluhan nyeri pada dada. Pada waktu jantung harus bekerja lebih keras terjadi ketidakseimbangan antara kebutuhan dan asupan oksigen, hal inilah yang menyebabkan nyeri dada. Kalau pembuluh darah tersumbat sama sekali, pemasokan darah ke jantung akan terhenti dan kejadian inilah yang disebut dengan serangan jantung (Supriyono:2011). Menurut Djohan (2004), penyakit jantung merupakan pengerasan pembuluh darah nadi menyempit karena terjadi endapan-endapan lemak (*atheroma* dan *plaques*) pada dindingnya.

Seseorang kemungkinan mengalami serangan jantung, karena terjadi iskemia miokard atau kekurangan oksigen pada otot jantung, yaitu jika mengeluhkan adanya nyeri dada atau nyeri hebat di ulu hati (*epigastrium*) yang bukan disebabkan oleh trauma, terjadi pada laki-laki berusia 35 tahun atau perempuan berusia di atas 40 tahun. Sindrom koroner akut ini biasanya berupa nyeri seperti tertekan benda berat, rasa tercekik, ditinju, ditikam, diremas, atau rasa seperti terbakar pada dada. Umumnya rasa nyeri dirasakan dibelakang tulang dada (sternum) disebelah kiri yang menyebar ke seluruh dada. Rasa nyeri dapat menjalar ke tengkuk, rahang, bahu, punggung dan lengan kiri. Keluhan lain dapat berupa rasa nyeri atau tidak nyaman di ulu hati yang penyebabnya tidak dapat dijelaskan. Sebagian kasus disertai mual dan muntah, disertai sesak nafas, banyak berkeringat, bahkan kesadaran menurun. Tiga bentuk penyakit jantung ini adalah serangan jantung, angina pectoris, serta gangguan irama jantung.

Secara klinis, menurut Supriyono untuk mendiagnosis *infark miokard* diperlukan 3 (tiga) kriteria, sebagai berikut :

- 1) Terdapat riwayat klinis : perasaan tertekan dan nyeri pada dada (ulu hati), selama 30 menit atau lebih.
- 2) Perubahan gambaran EKG : segmen ST elevasi lebih dari 0,2 mV paling sedikit 2 (dua) *precordial leads*, depresi segmen ST lebih besar dari 0,1 mV paling sedikit 2 (dua) *leads*, ketidak normalan gelombang Q atau inversi gelombang T paling sedikit 2 (dua) *leads*.
- 3) Peningkatan konsentrasi serum kreatinin kinase 2 (dua) kali lebih besar dari nilai normal pada pemeriksaan laboratorium.

Penyumbatan pembuluh darah koroner terjadi akibat adanya proses aterosklerosis, yang diawali dengan penimbunan lemak pada lapisan-lapisan pembuluh darah tersebut. Proses aterosklerosis sebenarnya sudah dimulai sejak masa anak-anak, tetapi baru manifes pada usia dewasa, usia pertengahan atau usia lanjut. Selain proses aterosklerosis ada juga proses lain, yakni spasme (penyempitan) pembuluh darah koroner tanpa adanya kelainan anatomis, yang secara tersendiri atau bersama-sama memberikan gejala iskemia.

Beberapa faktor yang bisa menimbulkan penyakit jantung antara lain:

- a) bertambahnya usia,
- b) gaya hidup
- c) stress
- d) kurangnya waktu istirahat
- e) kurangnya berolah raga
- f) merokok
- g) obesitas
- h) dislipidemia
- i) permasalahan dalam diagnosa klinis penyakit jantung

Berikut parameter yang akan digunakan dalam penelitian:

- 1) usia

usia mempengaruhi seseorang untuk terkena penyakit jantung. Usia yang semakin tua akan meningkatkan resiko penyakit jantung, hal tersebut telah diketahui pada pria atau wanita. Hal ini mungkin merupakan pencerminan faktor resiko digabung dengan kecenderungan bertambah beratnya derajat tiap-tiap faktor resiko dengan pertambahan usia. Ketika usia 45 tahun, sangat penting bagi kaum pria untuk menyadari kerentanan mereka dan mengambil tindakan positif untuk mencegah penyakit serangan jantung. Ketika memasuki usia 55 tahun atau mulai mengalami *menopause* dini, baik alami maupun sebagai akibat operasi, kaum wanita mulai menyusul kaum pria dalam hal resiko penyakit jantung. Meskipun demikian bukan berarti lansia pasti akan mengalami masalah jantung.

2) jenis kelamin

Laki-laki memiliki risiko lebih besar terkena serangan jantung dan kejadiannya lebih awal dari pada wanita. Morbiditas penyakit jantung adalah laki-laki dua kali lebih besar dibandingkan dengan wanita dan kondisi ini terjadi hampir 10 tahun lebih dini pada laki-laki daripada perempuan. Estrogen endogen bersifat protektif pada perempuan, namun setelah menopause insiden penyakit jantung meningkat dengan pesat, tetapi tidak sebesar insiden penyakit jantung pada laki-laki. Perokok pada wanita mengalami menopause lebih dini daripada bukan perokok.

Faktor resiko penyakit jantung untuk jenis kelamin laki-laki dan perempuan sama, tetapi pria biasanya menderita penyakit jantung 10 sampai 15 tahun lebih awal daripada wanita. Hingga berusia 60 tahun, di Amerika Serikat, hanya 1 dari 17 wanita yang sudah mengalami kelainan koroner, sedangkan pria 1 dari 5. Sesudah usia 60 tahun, penyakit jantung menjadi penyebab utama kematian wanita, sama dengan pria [16].

3) kolesterol total

Kolesterol ialah molekul yang ditemukan dalam sel. Merupakan sejenis lipid yang merupakan molekul lemak atau yang menyerupainya. Kolesterol ialah

jenis khusus lipid yang disebut steroid. Steroids ialah lipid yang memiliki struktur kimia khusus. Struktur ini terdiri atas 4 cincin atom karbon.

Steroid lain termasuk steroid hormon seperti kortisol, estrogen, dan testosteron. Nyatanya, semua hormon steroid terbuat dari perubahan struktur dasar kimia kolesterol. Saat tentang membuat sebuah molekul dari perubahan molekul yang lebih mudah, para ilmuwan menyebutnya *sintesis*. Hiperkolesterolemia berarti bahwa kadar kolesterol terlalu tinggi dalam darah.

4) High Density Lipid (HDL)

High Density Lipid (HDL) disebut juga dengan kolesterol baik karena mempunyai efek anti aterogenik yaitu mengangkut kolesterol bebas dari pembuluh darah dan jaringan lain menuju hati selanjutnya mengeluarkannya lewat empedu. Kadar HDL yang tinggi cenderung disertai kadar trigliserida yang rendah. Sebaliknya, penurunan kadar HDL diikuti peningkatan kadar trigliserida.

5) Low Density Lipid (LDL)

Low Density Lipid (LDL) merupakan lipoprotein yang terutama terdiri atas kolesterol bersirkulasi dalam tubuh dan dibawa ke sel-sel otot, lemak, dan sel-sel lain. Reseptor LDL yang ada dalam hati akan mengeluarkan LDL dari sirkulasi. Jadi, hati berperan dalam pengaturan kadar kolesterol darah, sehingga jika hati mengalami gangguan, kadar kolesterol darah dapat meningkat. Kolesterol LDL berpotensi menyebabkan terjadinya penyumbatan dan pengendapan di arteri (aterosklerosis) yang berujung penyakit jantung koroner.

6) trigliserid

Contoh trigliserida lemak tak jenuh. Terdiri dari gliserol, asam palmitat, asam oleat, asam alfa-linolenat. Trigliserida atau lebih tepatnya triasilgliserol atau triasilgliserida adalah sebuah gliserida, yaitu ester dari gliserol dan tiga asam

lemak. Trigliserida merupakan penyusun utama minyak nabati dan lemak hewani.

Sebagian besar atom karbon yang berasal dari glukosa dan asam amino yang berlebihan akan disintesis menjadi trigliserida. Trigliserida ini akan dikeluarkan ke dalam aliran darah sebagai very low density lipoprotein (VLDL) yang digunakan untuk menghasilkan energi atau disimpan dalam sel-sel lemak. Apabila energi tidak segera digunakan tubuh, dapat mengakibatkan kegemukan dan peningkatan kadar trigliserida dalam tubuh. Trigliserida dan lipida besar lainnya (kolesterol dan fosfolipida) yang terbentuk di dalam usus halus dikemas untuk diabsorpsi secara aktif dan ditransportasi oleh darah. Bahan-bahan ini bergabung dengan protein-protein khusus dan membentuk alat angkut lipida yang dinamakan lipoprotein.

7) rasio kolestrol

rasio kolesterol adalah perbandingan antara trigliserid, HDL, dan LDL dalam darah.

8) irama

Gangguan irama jantung dapat menimbulkan kematian secara mendadak. Gejalanya berupa hilangnya kesadaran dengan cepat, yang sering kali didahului nyeri dada. Disamping itu, secara umum penderita biasanya tampak cemas, gelisah, pucat dan berkeringat dingin. Denyut nadi umumnya cepat (takhikardi), irama tidak teratur, tetapi dapat pula denyut nadi lambat (bradikardia). Hipertensi maupun hipotensi dapat terjadi pada penderita ini. Meskipun kadang-kadang kurang jelas pada pemeriksaan fisik pada jantung, tetapi pemeriksaan menggunakan EKG (*elektrokardiografi*) akan sangat membantu memberi informasi.

9) frekuensi/menit

frekuensi denyut jantung ketika di EKG. Saat istirahat biasanya sekitar 70 kali per menit. Frekuensi denyut jantung dikontrol terutama oleh reduksi

dalam stimulasi melalui serat nervus parasimpatis (vagus), pengaruh yang lebih kecil oleh stimulasi melalui serat nervus simpatis.

10) zone transisi, axis (posisi listrik)

Arus listrik yang dihasilkan dalam jantung selama dipolarisasi dan repolarisasi akan dihantarkan ke seluruh permukaan tubuh. Muatan listrik tersebut dapat dicatat menggunakan elektroda yang ditempelkan pada kulit. Berbagai kombinasi – kombinasi elektroda ini akan menghasilkan 12 sadapan standar. Masing – masing sadapan mencatat peristiwa listrik dari seluruh siklus jantung, namun masing – masing hantaran meninjau jantung dari sudut pandang yang agak berbeda. Oleh sebab itu, bentuk gelombang yang dibentuk oleh sadapan sedikit berbeda.

11) gelombang P, gelombang QRS, gelombang S, gelombang T

Gelombang P menggambarkan depolarisasi otot atrium, normalnya setinggi 2,5 atau kurang dan durasinya 0,11 detik atau kurang. Defleksi negatif pertama setelah gelombang P adalah gelombang Q, yang normalnya berdurasi kurang dari 0,03 detik dan amplitudonya kurang dari 25 % gelombang R, sedangkan gelombang S adalah defleksi negatif pertama setelah gelombang R. Kompleks QRS (dimulai oleh gelombang Q, atau gelombang R bila tidak ada gelombang Q, diakhiri oleh gelombang S) menggambarkan depolarisasi otot ventrikel. Kompleks QRS normalnya berdurasi 0,04 sampai 0,10 detik. Gelombang T menggambarkan repolarisasi otot ventrikel. Gelombang ini mengikuti kompleks QRS dan biasanya mempunyai defleksi yang sama dengan kompleks QRS. Gelombang U diperkirakan menggambarkan repolarisasi serat purkinje tetapi kadang-kadang ditemukan pada pasien dengan hipokalemia (kadar kalium rendah). Gelombang U terjadi setelah gelombang T dan kurang lebih ukurannya sama dengan gelombang P. Gelombang ini sering disalah tafsirkan sebagai gelombang P ekstra.

12) interval P-R

Interval PR diukur mulai dari permulaan gelombang P sampai permulaan gelombang Q atau R dan menggambarkan waktu yang diperlukan untuk depolarisasi atrium dan perlambatan impuls di nodulus AV sebelum depolarisasi ventrikel.

13) interval QRS

Interval QT yang menggambarkan waktu total repolarisasi dan depolarisasi ventrikel, diukur dari awal gelombang Q atau R jika tidak ada gelombang Q diakhiri dengan gelombang T. Interval QT bervariasi sesuai dengan frekuensi jantung, biasanya kurang dari setengah interval RR (diukur dari permulaan satu gelombang R sampai awal gelombang R berikutnya), dan biasanya durasinya 0,32 sampai 0,40 detik apabila frekuensi jantungnya 65 sampai 95 denyut per menit.

14) voltage

Tegangan yang digunakan alat EKG terhadap kondisi tubuh pasien.

15) segmen ST

Segmen ST yang menggambarkan repolarisasi ventrikel awal, berlangsung dari akhir gelombang S sampai permulaan gelombang T

16) cepat capek disertai nyeri dada dan sesak

Gejala nyeri biasanya timbul ketika penderita melakukan aktivitas dan akan mereda setelah beristirahat. Pemicu timbulnya nyeri ini antara lain udara dingin dan stres psikologik. Penyebab sakit dada berhubungan dengan pengisian arteri koronaria sewaktu diastole. Setiap keadaan yang akan meningkatkan denyut jantung akan meningkatkan juga kebutuhan jantung yang tidak bisa dipenuhi oleh pasok aliran darah koroner dan akan mengakibatkan sakit. Sakit sering terjadi sesudah suatu keadaan emosi, latihan fisik, makan banyak, perubahan suhu, bersenggama, dan lain-lain. Sakit akan menghilang bila kecepatan denyut jantung diperlambat, relaksasi,

istirahat, atau makan obat *glyceril trinitrat*. Sakit biasanya menghilang dalam waktu 5 menit.

17) merokok

Merokok merupakan faktor risiko mayor untuk terjadinya penyakit jantung. Merokok sigaret menaikkan risiko serangan jantung sebanyak 2 sampai 3 kali. Sekitar 24 % kematian akibat penyakit jantung pada laki-laki dan 11 % pada perempuan disebabkan kebiasaan merokok. Meskipun terdapat penurunan yang progresif proporsi pada populasi yang merokok. Salah satu hal yang menjadi perhatian adalah prevalensi kebiasaan merokok yang meningkat pada remaja, terutama pada remaja perempuan. Orang yang tidak merokok dan tinggal bersama perokok (perokok pasif) memiliki peningkatan risiko sebesar 20 – 30 % dibandingkan dengan orang yang tinggal dengan bukan perokok. Risiko terjadinya penyakit jantung akibat merokok berkaitan dengan dosis dimana orang yang merokok 20 batang rokok atau lebih dalam sehari memiliki risiko sebesar dua hingga tiga kali lebih tinggi daripada populasi umum untuk mengalami kejadian penyakit jantung. Peran rokok dalam patogenesis penyakit jantung merupakan hal yang kompleks, diantaranya :

- a. Timbulnya aterosklerosis.
- b. Peningkatan trombogenesis
- c. Peningkatan tekanan darah dan denyut jantung.
- d. Provokasi aritmia jantung.
- e. Peningkatan kebutuhan oksigen miokard.
- f. Penurunan kapasitas pengangkutan oksigen.

Risiko terjadinya penyakit jantung akibat merokok turun menjadi 50 % setelah satu tahun berhenti merokok dan menjadi normal setelah 4 tahun berhenti. Rokok juga merupakan faktor risiko utama dalam terjadinya : penyakit saluran nafas, saluran pencernaan, *cirrhosis hepatis*, kanker kandung kencing dan penurunan kesegaran jasmani. Manfaat penghentian kebiasaan merokok lebih sedikit kontroversinya dibandingkan dengan diet dan olah

raga. Oleh karena itu saran penghentian kebiasaan merokok merupakan komponen utama pada program rehabilitasi jantung koroner

18) olahraga

Pada latihan fisik akan terjadi dua perubahan pada sistem kardiovaskuler, yaitu peningkatan curah jantung dan redistribusi aliran darah dari organ yang kurang aktif ke organ yang aktif. Aktivitas aerobik secara teratur menurunkan risiko penyakit jantung, meskipun hanya 11 % laki-laki dan 4 % perempuan memenuhi target pemerintah untuk berolah raga. Disimpulkan juga bahwa olah raga secara teratur akan menurunkan tekanan darah sistolik, menurunkan kadar katekolamin di sirkulasi, menurunkan kadar kolesterol dan lemak darah, meningkatkan kadar HDL lipoprotein, memperbaiki sirkulasi koroner dan meningkatkan percaya diri.

Diperkirakan sepertiga laki-laki dan dua per tiga perempuan tidak dapat mempertahankan irama langkah yang normal pada kemiringan gradual (3 mph pada gradient 5 %). Olah raga yang teratur berkaitan dengan penurunan insiden penyakit jantung sebesar 20 – 40 %. Dengan berolah raga secara teratur sangat bermanfaat untuk menurunkan faktor risiko seperti kenaikan HDL-kolesterol dan sensitivitas insulin serta menurunkan berat badan dan kadar LDL-kolesterol.

19) hipertensi

Riwayat hipertensi adalah keadaan dimana pasien dinyatakan pernah menderita hipertensi (*sistole* > 140 mmHg dan *diastole* > 90 mmHg) oleh dokter, pernah atau masih mendapatkan pengobatan anti hipertensif sebelum terdiagnosis penyakit jantung. Data riwayat hipertensi diperoleh dari rekam medis rumah sakit dan hasil wawancara dengan pasien atau keluarga.

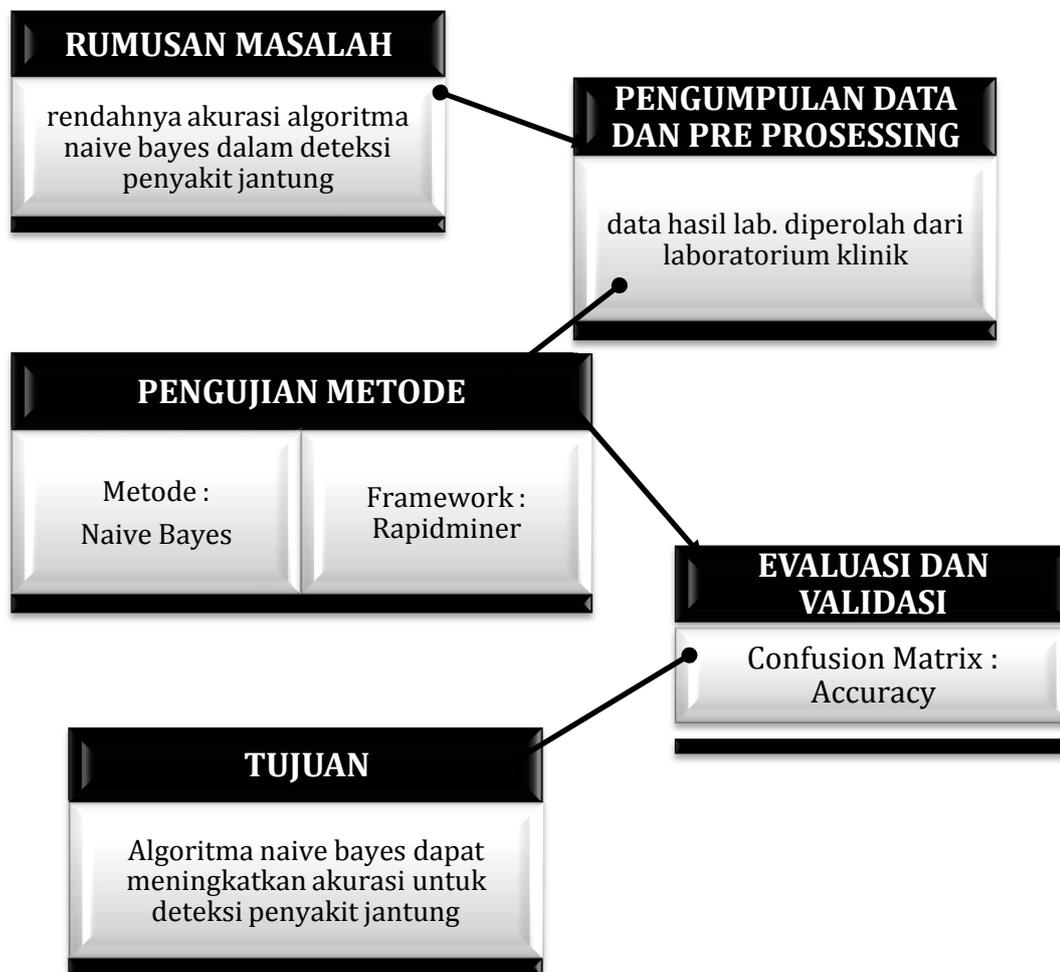
20) tekanan darah

Jika Anda memeriksakan tekanan darah maka akan didapatkan dua angka yaitu tekanan sistolik/tekanan diastol. Angka lebih tinggi yang diperoleh ketika jantung berkontraksi disebut tekanan sistol, sedangkan angka yang

lebih rendah disebut tekanan diastolik diperoleh ketika jantung berileksasi. Seseorang dikatakan sebagai penderita darah tinggi ketika terjadi kenaikan tekanan sistolik dan tekanan diastolik di atas normal.

1.3 Kerangka Pemikiran

Kerangka pemikiran dalam penelitian ini adalah rendahnya akurasi algoritma naive bayes dalam deteksi penyakit jantung. Oleh sebab itu, penelitian ini menggunakan algoritma klasifikasi data mining *naive bayes* dengan menggunakan data set hasil laboratorium yang akan dimasukkan ke dalam rapidminer dan akan dihitung akurasi. Dari hasil akurasi akan didapat algoritma yang akurat. Adapun kerangka pemikiran sebagai berikut:



Gambar 2. 2 kerangka pemikiran