BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Angka Kematian Ibu (AKI)

Kematian ibu adalah jumlah kematian ibu selama periode waktu tertentu per 100.000 kelahiran hidup. Kematian ibu adalah kematian seorang wanita saat hamil atau dalam 42 hari pengakhiran kehamilan, terlepas dari durasi dan tempat kehamilan, dari setiap penyebab yang berhubungan dengan atau diperburuk oleh kehamilan atau penanganannya tetapi bukan dari penyebab kecelakaan atau *incidental*[10].

Kematian ibu menunjukkan lingkup yang luas, tidak hanya terkait dengan kematian yang terjadi saat proses persalinan, tetapi mencakup kematian ibu yang sedang dalam masa hamil dan nifas[11].

2.1.1.1 Penyebab Kematian Maternal

Kematian ibu dibagi menjadi kematian langsung dan tidak langsung[12].

a. Kematian ibu langsung adalah sebagai akibat komplikasi kehamilan, persalinan, atau masa nifas dan segala intervensi atau penanganan tidak tepat dari komplikasi tersebut[12].
b. Kematian ibu tidak langsung merupakan akibat dari penyakit yang sudah ada atau penyakit yang timbul sewaktu kehamilan yang berpengaruh terhadap kehamilan, misalnya malaria, anemia, HIV/AIDS, dan penyakit kardiovaskuler[13].

Secara global, 80% kematian ibu tergolong pada kematian ibu langsung, pola penyebab langsung dimana-mana, yaitu perdarahan (25%, biasanya perdarahan pasc persalinan), sepsis (15%), hipertensi dalam kehamilan (12%), partus macet (8%), komplikasi aborsi tidak aman (13%) dan sebab-sebab lain (8%)[13].

2.1.1.2 Faktor Risiko yang Mempengaruhi Kematian Maternal

Faktor-faktor yang mempengaruhi kematian ibu yaitu[14]:

a. Determinan dekat yang berhubungan langsung dengan kematian ibu merupakan gangguan obstetrik seperti perdarahan, preeklamsi/eklamsi, dan infeks atau penyakit yang diderita ibu sebelum atau selama kehamilan yang dapat memperburuk kondisi kehamilan seperti jantung, malaria, tuberkulosis, ginjal, dan *acquired immunodeficiency syndrome*.
b. Determinan dekat secara langsung dipengaruhi oleh determinan antara yang berhubungan dengan faktor kesehatan, seperti status kesehatan ibu, status reproduksi, akses terhadap pelayanan kesehatan, dan perilaku penggunaan fasilitas kesehatan.
c. Determinan jauh berhubungan dengan faktor demografi dan sosiokultural. Kesadaran masyarakat yang rendah tentang kesehatan ibu hamil, pemberdayaan perempuan yang tidak baik, latar belakang pendidikan, sosial ekonomi keluarga, lingkungan masyarakat dan politik, serta kebijakan secara tidak langsung diduga ikut berperan dalam meningkatkan kematian ibu.

2.1.2 Faktor Risiko Kehamilan

Terdapat 20 faktor risiko yang dibagi menjadi 3 kelompok faktor risiko pada penilaian KSPR[15].

1. Kelompok Faktor Risiko I (Ada Potensi Gawat Obstetrik)
   a. Primi muda

      Ibu hamil pertama pada umur ≤ 16 tahun, rahim dan panggul belum tumbuh mencapai ukuran dewasa. Akibatnya diragukan keselamatan dan kesehatan janin dalam kandungan. Selain itu mental ibu belum cukup dewasa.

      Bahaya yang mungkin terjadi antara lain:

      1) Bayi lahir belum cukup umur
      2) Perdarahan bisa terjadi sebelum bayi lahir

3) Perdarahan dapat terjadi sesudah bayi lahir.

b. Primi Tua

1) Pada umur ibu ≥ 35 tahun

Ibu yang hamil pertama pada umur ≥ 35 tahun. Pada usia tersebut mudah terjadi penyakit pada ibu dan organ kandungan yang menua. Jalan lahir juga tambah kaku. Ada kemungkinan lebih besar ibu hamil mendapatkan anak cacat, terjadi persalinan macet dan perdarahan. Bahaya yang terjadi antara lain:

a) Hipertensi / tekanan darah tinggi
b) Pre-eklamsia
c) Ketuban pecah dini: yaitu ketuban pecah sebelum persalinan
d) Persalinan tidak lancar atau macet: ibu mengejan lebih dari satu jam, bayi tidak dapat lahir dengan tenaga ibu sendiri melalui jalan lahir biasa.
e) Perdarahan setelah bayi lahir

Bayi lahir dengan berat badan lahir rendah (BBLR) < 2500 gr

2) Lama perkawinan ≥ 4 tahun, Ibu hamil pertama setelah kawin 4 tahun atau lebih dengan kehidupan perkawinan biasa.

Bahaya yang terjadi pada primi tua :

a) Selama hamil dapat timbul masalah, faktor risiko lain oleh karena kehamilannya, misalnya pre-eklamsia.
b) Persalinan tidak lancar.

c. Primi Tua Sekunder

Ibu hamil dengan persalinan terakhir ≥ 10 tahun yang lalu. Ibu dalam kehamilan dan persalinan ini seolah-olah menghadapi persalinan yang pertama lagi. Kehamilan ini bisa terjadi pada:

1) Anak pertama mati, janin didambakan dengan nilai sosial tinggi.

2) Anak terkecil hidup umur 10 tahun lebih, ibu tidak ber-KB.

Bahaya yang dapat terjadi:

1) Persalinan dapat berjalan tidak lancar
2) Perdarahan pasca persalinan
3) Penyakit ibu: Hipertensi (tekanan darah tinggi), diabetes, dan lain-lain.

d. Anak terkecil < 2 tahun

Ibu hamil yang jarak kelahiran dengan anak terkecil kurang dari 2 tahun. Kesehatan fisik dan rahim ibu masih butuh cukup istirahat. Ada kemungkinan ibu masih menyusui. Selain itu anak tersebut masih butuh asuhan dan perhatian orang tuanya.

Bahaya yang dapat terjadi pada ibu hamil antara lain :

1) Perdarahan setelah bayi lahir karena kondisi ibu masih lemah.
2) Bayi prematur/lahir belum cukup bulan, sebelum 37 minggu.
3) Bayi dengan berat badan lahir rendah BBLR <2500 gram

e. Grande multi

Ibu pernah hamil / melahirkan 4 kali atau lebih. Karena ibu sering melahirkan maka kemungkinan akan banyak ditemui keadaan:

1) Kesehatan terganggu: anemia, kurang gizi
2) Kekendoran pada dinding perut
3) Tampak ibu dengan perut menggantung
4) Kekendoran dinding Rahim

Bahaya yang dapat terjadi:

1) Kelainan letak, persalinan letak lintang
2) Robekan rahim pada kelainan letak lintang
3) Persalinan lama
4) Perdarahan pasca persalinan.

f. Umur ibu ≥ 35 tahun

Ibu hamil berumur 35 tahun atau lebih, dimana pada usia tersebut terjadi perubahan pada jaringan alat-alat kandungan dan jalan lahir tidak lentur lagi. Selain itu ada kecenderungan didapatkan

penyakit lain dalam tubuh ibu.

Bahaya yang dapat terjadi:

1) Tekanan darah tinggi dan pre-eklamsia
2) Ketuban pecah dini
3) Persalinan tidak lancar / macet
4) Perdarahan setelah bayi lahir.

g. Tinggi badan ≤ 145 cm

Terdapat tiga batasan pada kelompok risiko ini:

1) Ibu hamil pertama sangat membutuhkan perhatian khusus. Luas panggul ibu dan besar kepala janin mungkin tidak proporsional, dalam hal ini ada dua kemungkinan yang terjadi:
   a) Panggul ibu sebagai jalan lahir ternyata sempit dengan janin / kepala tidak besar.
   b) Panggul ukuran normal tetapi anaknya besar / kepala besar
2) Ibu hamil kedua, dengan kehamilan lalu bayi lahir cukup bulan tetapi mati dalam waktu (umur bayi) 7 hari atau kurang.
3) Ibu hamil kehamilan sebelumnya belum penah melahirkan cukup bulan, dan berat badan lahir rendah < 2500 gram.

Bahaya yang dapat terjadi: persalinan berjalan tidak lancar, bayi sukar lahir, dalam bahaya.

h. Pernah gagal kehamilan

Dapat terjadi pada ibu hamil dengan:

1) Kehamilan kedua, dimana kehamilan yang pertama mengalami:
   a) Keguguran
   b) Lahir belum cukup bulan
   c) Lahir mati
   d) Lahir hidup lalu mati umur ≤ 7 hari
2) Kehamilan ketiga atau lebih, kehamilan yang lalu pernah mengalami keguguran ≥ 2 kali.

3) Kehamilan kedua atau lebih, kehamilan terakhir janin mati dalam kandungan

Bahaya yang dapat terjadi:

1) Kegagalan kehamilan dapat berulang dan terjadi lagi, dengan tanda-tanda pengeluaran buah kehamilan sebelum waktunya keluar darah, perut kencang.
2) Penyakit dari ibu yang menyebabkan kegagalan kehamilan, misalnya: Diabetes mellitus, radang saluran kencing, dll

i. Pernah melahirkan dengan tindakan

Persalinan yang ditolong dengan alat melalui jalan lahir biasa atau per-vaginam:

1) Tindakan dengan cunam / forcep / vakum. Bahaya yang dapat terjadi:
   a) Robekan / perlukaan jalan lahir
   b) Perdarahan pasca persalinan
2) Uri manual, yaitu: tindakan pengeluaran plasenta dari rongga rahim dengan menggunakan tangan. Tindakan ini dilakukan pada keadaan bila:
   a) Ditunggu setengah jam uri tidak dapat lahir sendiri
   b) Setelah bayi lahir serta uri belum lahir terjadi perdarahan banyak > 500 cc

Bahaya yang dapat terjadi:

a) Radang, bila tangan penolong tidak steril
b) Perforasi, bila jari si penolong menembus rahim
c) Perdarahan

3) Ibu diberi infus / tranfusi pada persalinan lalu. Persalinan yang lalu mengalami perdarahan pasca persalinan yang banyak lebih dari 500 cc, sehingga ibu menjadi syok dan membutuhkan infus, serta transfusi darah

j. Pernah operasi sesar

Ibu hamil, pada persalinan yang lalu dilakukan operasi sesar. Oleh karena itu pada dinding rahim ibu terdapat cacat bekas luka operasi.

Bahaya pada robekan rahim : kematian janin dan kematian ibu, perdarahan dan infeksi.

2. Kelompok Faktor Risiko II (Ada Gawat Obstetrik)
   a. Penyakit ibu : anemia, malaria, TBC paru, payah jantung, dan penyakit lain
   b. Preeklampsia ringan
   c. Hamil kembar/gemeli : Perut ibu sangat besar, gerak anak terasa sangat dibanyak tempat
   d. Hamil kembar air/Hidramnion : air ketuban terlalu banyak
   e. Hamil serotinus : hamil lebih bulan (≥ 9 bulan belum melahirkan)
   f. Janin mati di dalam rahim ibu
   g. Letak sungsang
   h. Letak Lintang

3. Kelompok Faktor Risiko III (Ada Gawat Darurat Obstetrik)
   a. Perdarahan Antepartum : dapat berupa solusio plasenta, plasenta previa, atau vasa previa
   b. Preeklampsia berat/eklampsia

2.1.3 Umur

Usia seorang wanita pada saat hamil sebaiknya tidak terlalu muda dan tidak terlalu tua. Umur yang kurang dari 20 tahun atau lebih 35 tahun, berisiko tinggi untuk melahirkan. Kesiapan seorang perempuan untuk hamil harus siap fisik, emosi, psikologi, social dan ekonomi[16].

Berikut kategori umur menurut Depkes RI (2009) :

1. Masa balita : 0-5 tahun
2. Masa kanak- kanak : 5-11 tahun

3. Masa remaja awal : 12-16 tahun
4. Masa remaja akhir : 17-25 tahun
5. Masa dewasa awal : 26-35 tahun
6. Masa dewasa akhur : 36-45 tahun
7. Masa Lansia Awal : 46-55 tahun
8. Masa lansia akhir : 56-65 tahun
9. Masa manula : > 65 tahun

2.1.4 Pemanfaatan Data MIK

2.1.4.1 Informasi

Informasi adalah data yang sudah diolah, dibentuk atau dimanipulasi sesuai dengan kebutuhan tertentu. Pengolahan data menjadi informasi yang antara lain berbentuk laporan, model, deskriptif dan bentuk statistik[17].

Transformasi informasi adalah komponen proses dalam pengelolaan sistem informasi yang berfungsi memproses data menjadi informasi sehingga dapat dihasilkan produk informasi yang diperlukan bagi para pemakai informasi. Terdiri dari[18] :

a. Pengumpulan data

   Pengumpulan data dilaksanakan sesuai dengan jenis data objek dan sumber data serta persiapan pengumpulan data cara memperoleh data ialah dia secara langsung ataupun tidak langsung.

b. Pengolahan data

   Pengolahan data dapat dilakukan secara manual ataupun dengan bantuan komputer. Hasil pengolahan data berupa keterangan-keterangan

c. Penyajian dan penyebarluasan data informasi

   Penyajian data dan informasi dilakukan baik secara visual maupun dalam bentuk publikasi dengan metode komunikasi langsung atau tak langsung.

d. Penataan dokumentasi

Pendokumentasian dapat dilakukan dengan cara yang lama (file) dan cara baru (komputerisasi) contohnya perpustakaan bertalian dengan upaya pengumpulan pemeliharaan penyimpanan pengaturan dan pendayagunaan informasi.

2.1.5 *Data mining*

2.1.5.1 Pengertian *Data mining*

Data mining sering juga disebut *knowledge discovery in database* (KDD), adalah kegiatan yang meliputi pengumpulan, pemakaina data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari *data mining* ini bisa di pakai untuk memperbaiki pengambilan keputusan di masa depan [19]. *Data mining* adalah proses informasi yang berguna secara otomatis dalam repositori data besar. Teknik *data mining* dikerahkan untuk mencari database besar untuk menemukan pola baru dan berguna yang mungkin masih belum diketahui [20].

Data mining adalah langkah analisa terhadap proses penemuan pengetahuan di dalam basisdata atua *knowledge discovery in databases* yang disingkat KDD. Pengetahuan bisa berupa pola data atau relasi antar data yang *valid. Data mining* merupakan gabungan sejumlah disiplin ilmu komputer, yang didefenisikan sebagai proses penemuan pola-pola baru dari kumpulan-kumpulan data sangat besar, meliputi metode-metode yang merupakan irisan dari *artificial intelligence, machine learning, statistics, dan database systems. Data mining* diajukan untuk mengeksrak (mengambil intisari) pengetahuan dari sekumpulan data sehingga didaptkan struktur yang dapat dimengerti manusia serta meliputi basis data dan managemen data, prapemrosesan data, pertimbangan model dan inferensi, ukuran ketertarikan, pertimbangan kompleksitas, pascapemrosesan terhadap struktur yang ditemukan, visualisasi, dan *online updating* [21].

2.1.5.2 Kegunaan *Data Mining*

Secara umum, kegunaan *data mining* dapat di bagi menjadi dua: deskriptif dan prediktif. Deskriptif berarti *data mining* digunakan untuk mencari pola-pola yang dapat dipahami manusia yang menjelaskan karakteristik data. Sedangkan prediktif berarti *data mining* digunakan untuk membentuk sebuah model pengetahuan yang akan digunakan untuk melakukan prediksi. Berdasarkan fungsionalitasnya, tugas-tugas *data mining* bisa dikelompokkan kedalam enam kelompok berikut ini[21] :

a. Klasifikasi (*classificationi*) : men-generalisasi struktur yang diketahui untuk duplikasikan pada data-data baru. Misalkan klasifikasi penyakit ke dalam sejumlah jenis, klasifikasi email ke dalam spam atau bukan.
b. Klasterisasi (*clustering*) : mengelompokkan data, yang tidak diketahui label kelasnya, kedalam sejumlah kelompok tertentun sesuai dengan ukuran kemiripannya.
c. Regresi (*regression*) : menemukan suatu fungsi memodelkan data dengan galat (kesalahan prediksi) seminimal mungkin.
d. Deteksi anomaly (*anomaly detection*) : mengidentifikasi data yang tidak umum, bisa berupa *outlier* (pencilan), perubahan atau deviasi yang mungkin sangat penting dan perlu investigasi lebih lanjut.
e. Pembelajaran aturan asosiasi (*association rule learning*): atau pemodelan kebergantungan (*dependency modelling*): mencari relasi atar variabel.
f. Perangkuman (*summarization*) : menyediakan representasi data yang lebih sederhana, meliputi visualisasi dan pembuatan laporan.

2.1.5.3 *Prediction / Forecasting*

Prediction pada umumnya di anggap sebagai tindakan yang menjelaskan mengenai masa mendatang. Hal ini berbeda dengan menebak secara sederhana dengan mempertimbangakn pengalaman, opini dann informasi lainnya dalam melakukan peramalan. Istilah yang

umumnya dikaitkan dengan *prediction* adalah *forecasting*[22].

Ada perbedaan tipis namun sangat penting di antara keduanya. *Prediction* pada umumnya berbasis opini dan pengalaman, *forecasting* berbasis data dan model. Dalam terminologi *data mining prediction / forecasting* digunakan secara sinonim, dan istilah prediksi digunakan sebagai penyajian yang umum. Bergantung pada sifat alami yang akan di prediksikan, "*prediction*" bisa disebut secara spesifik sebagai *classification* (dimana hal yang diprediksi, seperti ramalan esok, diberi label *class* misalnya '*rainy*' atau '*sunny*') atau regresi (dimana hal yang diprediksi, misalnya suhu esok, adalah angka riil mislanya '65˚C')[22].

2.1.6 *Decision Tree*

2.1.6.1 Pengertian *Decision Tree*

Decision tree adalah salah satu metode klasifikasi yang pouler dan banyak digunakan secara praktis. Metode ini berusaha menemukan model klasifikasi yang tahan terhadap derau. Salah satu metode *decision tree* yang sangat populer *iterative dichotomizer version* 3 (ID3). Dua varian lain yang popular adalah C4.5 dan ASSISTANT[21].

Metode ID3 berusaha membangun model kalsifikasi yang berupa *decision tree* (pohon keputusan) secara *top-down* (dari atas ke bawah). Caranya adalah dengan mengevaluasi semua atribut menggunakan suatu ukuran statistik, biasanya berupa *information gain,* untuk mengukur efektifitas suatu atribut dalam mengklasifikasikan himpunan sampel data.

2.1.6.2 Kekuatan dan Kelemahan *Decision Tree*

Menurut (Suyatno, 2017), *decision tree* dapat dikatakan memiliki tiga kekuatan, yaitu:

a. Mempunyai konsep yang jelas dan mudah dipahami. Pohon keputusan yang dihasilkan dari proses pelatihan dapat menjelaskan bagaimana model klasifikasi data bekerja.
b. Mudah di implementasikan dengan menggunakan algoritma rekursif.

Sementara itu, kelemahan *decision tree* yang sangat terlihat adalah:

a. Sulit diaplikasikan untuk himpunan data sangat besar dengan ribuan atribut dan milyaran objek data.
b. Mudah mengalami *overfit* karena proses pelatihan *greedy* tidak menjamin dihasilkannya pohon keputusan yang optimum.

2.1.6.3 Algoritma C4.5

C4.5 merupakan pengembangan dari ID3, dimana ID3 adalah model *decision tree* yang lain. Dalam ID3 kita gunakan kriteria *information gain* untuk memilih atribut yang akan digunakan untuk pemisahan obyek. Atribut yang mempunyai *information gain* paling tinggi di banding atribut yang lain relatif terhadap set dalam suatu data, dipilih untuk melakukan pemecahan. Perbaikan dilakukan dalam hal[19] :

a. Bisa mengatasi missing data
b. Bisa mengatasi data kontinyu
c. Pruning
d. Aturan
e. Menggunakan

Algoritma C4.5. *Decision Tree* menyerupai struktur *flowchart*, yang masing-masng internal *node*-nya dinyatakan sebagai atribut pengujian, setiap cabang mewakili *output* dari pengujian, dan setiap *node* daun (*terminal node*) menentukan label *class*. *Node* paling atas dari sebuah pohon adalah *node* akar [23].

Algoritma C4.5 menggunakan konsep *information gain* atau *entropy reduction* untuk memilih pembagian yang optimal (Larose, 2005). Tahapan dalam membuat pohon keputusan dengan algoritma C4.5 yaitu[24] :

1. Mempersiapkan data *training*, dapat diambil dari data histori yang pernah terjadi sebelumnya dan sudah dikelompokkan dalam kelas-kelas tertentu.

2. Menentukan akar dari pohon dengan menghitung nilai *gain* yang tertinggi dari masing-masng atribut atau berdasarkan nilai *index entropy* terendah.

   Secara istilah, *entropy* adalah keberbedaan atau keberagaman. Dalam *data mining, entropy* didefinisikan sebagai suatu parameter untuk mengukur heterogenitas (keberagaman) dalam suatu himpunan data. Semakin heterogen suatu himpunan data, semakin besar pula nilai *entropy*-nya. Secara matematis, *entropy* dirumuskan sebagai :

   Entropy : $\sum_i^c -Pi\ Log2\ Pi$

   Dimana c adalah jumlah nilai yang terdapat pada atribut target (jumlah kelas). Sedangkan P*i* menyatakan porsi atau rasio antara jumlah sampel di kelas *i* dengan jumlah semua sampel pada himpunan data.

3. Hitung nilai *gain* dengan rumus:

   Secara istilah, *information gain* adalah perolehan informasi. Dalam *data mining, information gain* didefinisikan sebagai ukuran efektivitas suatu atribut dalam mengklasifikasikan data. Secara sistematis, *information gain* dari suatu atribut A, dituliskan sebagai

   Gain(S,A) =Entropy(S)-$\sum_{Values\ (A)} \frac{|s1|}{|s|}$ * Entropy(Sv)

   Dimana :

   A : atribut
   V : menyatakan suatu nilai yang mungkin untuk atribut A
   Values(A) : himpunan nilai-nilai yang mungkin untuk atribut A
   |Sv| : jumlah sampel untuk nilai v
   |S| : jumlah seluruh sampe data
   Entropy(Sv) : entropy untuk sampel-sampel yang memiliki nilai

4. Ulangi langkah ke-2 hingga semua *record* terpartisi. Proses partisi pohon keputusan akan berhenti disaat:

a. Semua tupel dalam *record* dalam simpul N mendapat kelas yang sama.
b. Tidak ada atribut dalam *record* yang dipartisi lagi.
c. Tidak ada *record* di dalam cabang yang kosong.

2.1.6.4 Fungsi C4.5 dalam Bidang Kesehatan

Untuk mendukung aturan-aturan yang terbentuk dari data pasien yang diperoleh kedalam sistem maka dibentuklah pohon keputusan, selain berfungsi sebagai penentuan *rules* atau aturan klasifikasi penyakit, sistem pohon keputusan ini juga mempresentasikan bagaimana seorang pasien bisa terserang penyakit tersebut dari beberapa variabel yang tersedia dari data pasien[25]. Selain itu, algoritma C4.5 juga dapat digunakan sebagai GUI (*graphical user interface*) atau aplikasi yang digunakan untuk membantu dokter dan pasien dalam mendiagnosa penyakit[26].

2.1.7 Model *Evaluation*

2.1.7.1 *Confusion Matrix* (Tabel Kebenaran)

Kinerja klasifikasi paling baik dideskripsikan dengan tepat oleh alat yang disebut matriks kebingungan. Memahami tabel kebenaranmembutuhkan pengenalan terhadap beberapa definisi. Tetapi sebelum memperkenalkan definisi, kita harus melihat tabel kebenarandasar untuk klasifikasi biner atau binomial di mana ada dua kelas (katakanlah, Y atau N). Keakuratan klasifikasi contoh spesifik dapat dilihat salah satu dari empat cara yang mungkin :

1. Kelas yang diprediksi adalah Y, dan kelas yang sebenarnya juga Y -> ini adalah "Benar Positif" atau TP.
2. Kelas yang diprediksi adalah Y, dan kelas yang sebenarnya adalah N -> ini adalah "Salah Positive" atau FP.
3. Kelas yang diprediksi adalah N, dan kelas yang sebenarnya adalah Y -> ini adalah "False Negative" atau FN.
4. Kelas yang diprediksi adalah N, dan kelas yang sebenarnya juga N -> ini adalah "Benar Negatif" atau TN.

Tabel 2.1
Confusion Matrix

		Actual class (observation)	
		Y	N
Prediksi Kelas	Y	TP (True Positive) Hasil yang benar	FP (False Positive) Tidak terduga
	N	FN (False Negative) Hasil yang hilang	TN (True Negative) Tidak ada nilai yang benar

Tabel kebenaran disusun sebagai matrix 2 x 2 seperti yang ditunjukkan pada tabel diatas. Kelas-kelas yang diprediksi disusun secara vertikal dalam kolom, meskipun kadang-kadang urutan ini dibalik.

Sensitivity adalah kemampuan alat penggolong untuk memilih semua kasus yang perlu dipilih. Penggolong sempurna akan memilih semua Y yang sebenarnya dan tidak akan melewatkan Y yang sebenarnya. Dengan kata lain itu tidak akan memiliki negatif palsu. Pada kenyataannya, setiap pengklasifikasi akan kehilangan beberapa Y yang benar dan karenanya memiliki beberapa negative yang palsu. *Sensitivity*

dinyatakan sebagai rasio (atau persentase) dihitung sebagai berikut: TP / (TP + FN). Namun *Sensitivity* saja tidak cukup untuk mengevaluasi penggolong.

Specificity adalah kemampuan pengklasifikasi untuk menolak semua kasus yang perlu ditolak. Pengklasifikasi yang sempurna akan menolak semua N yang sebenarnya dan tidak akan memberikan hasil yang terpegaruh. Dengan kata lain itu tidak akan memiliki positif palsu. Pada kenyataannnya, setiap penggolong akan memilih beberapa kasus yang perlu ditolak dan dengan demikian memiliki beberapa kesalahan positif. *Specificity* dinyatakan sebagai rasio (atau persentase) yang dihitung sebagai berikut: TN / (TN + FP).

Pecision di definisikan sebagai tingkat ketepatan antara inforasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh system. Oleh karena itu nilai ketepatannya adalah TP / (TP + FP).

Recall yaitu tingkat keberhasilan system dalam menemukan kembali sebuah informasi. *Recall* sama dengan *sensitivity* karena *recall* juga diberikan oleh TP / (TP + FN).

Accuracy yaitu tingkat kedekatan antara nilai prediksi dengan nilai akurasi. *Accuracy* dihitung dengan rumus (TP + TN) / (TP + FP + TN + FN).

2.1.7.2 Receiver Operator Characteristic (ROC) Curves And Area Under The Curve (AUC)

Kurva karakteristik operator penerima (ROC) memenuhi kebutuhan ini dan pada awalnya dikembangkan di bidang deteksi sinyal. Kurva ROC dibuat dengan memplot fraksi true positive (tingkat TP) versus fraksi false positive (tingkat FP). Ketika kita menghasilkan tab dengan nilai seperti itu, kita dapat memplot laju FP pada sumbu horizontal dan laju TP (sama seperti sensitivitas atau penarikan kembali) pada sumbu vertikal.

Performance keakurasian AUC dapat diklasifikasikan menjadi lima kelompok yaitu[24] :

0,90 – 1,00 = *Excellent Classification*

0,80 – 0,90 = *Good Classification*

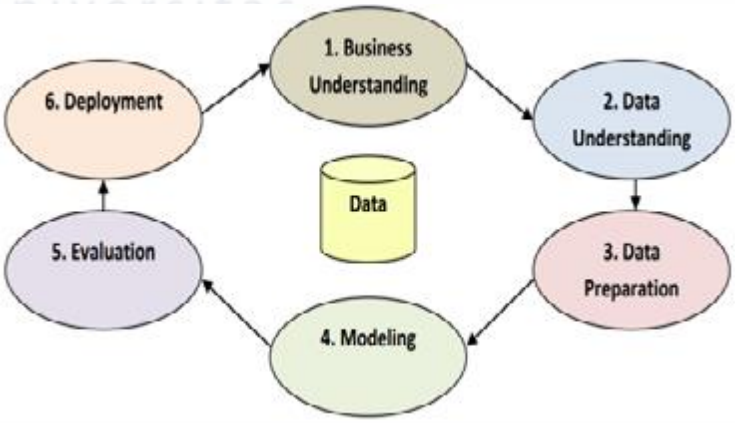
0,70 – 0,80 = *Fair Classification*

0,60 – 0,70 = *Poor Classification*

0,50 – 0,60 = *Failure Classification*

2.1.8 Metode *Data Mining Standar Process* (CRISP-DM)

Cross-Industry Standard Process for Data Mining (CRISP-DM) yang dikembangkan tahun 1996 oleh analis dari beberapa industri menyediakan standar proses data mining sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian :



Gambar 2.1
Langkah Tahapan pada CRISP-DM

Dalam CRISP-DM, siklus hidup data mining yang terbagi dalam 6 (enam) fase yaitu:

1. Fase pemahaman bisnis (*Business Understanding Phase*)
   a. Penentuan tujuan proyek dan kebutuhan secara detail dalam hidup bisnis atau unit penelitian.
   b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan data mining.
   c. Menyiapkan strategi awal untuk mencapai tujuan
2. Fase pemahaman data (*Data Understanding Phase*)
   a. Mengumpulkan data.
   b. Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
   c. Mengevaluasi kualitas data.
   d. Jika diinginkan, pilih sebagian kecil grup data yang mungkin mengandung pola dari permasalahan
3. Fase pengolahan data (*Data Preparation Phase*)
   a. Siapkan data awal, kumpulkan data yang akan digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.
   b. Pilih kasus dan variabel yang ingin dianalisis dan yang sesuai analisis yang akan dilakukan.

c. Lakukan perubahan pada beberapa variabel jika dibutuhkan.
d. Siapkan data awal sehingga siap untuk perangkat pemodelan.

4. Fase pemodelan (*Modeling Phase*)
   a. Pilh dan aplikasikan teknik pemodelan yang sesuai.
   b. Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan data mining yang sama.
   c. Jika diperlukan, proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik data mining tertentu.
5. Fase evaluasi (Evaluation Phase)
   a. Pengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan untuk mendapatkan kualitas dan efektivitas sebelum disebarkan untuk digunakan.
   b. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.
   c. Menentukan apakah terdapat permasalah penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
   d. Mengambil keputusan yang berkaitan dengan penggunaan hasil dari data mining.
6. Fase penyebaran (*Deployment Phase*)
   a. Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesaikannya proyek.
   b. Contoh sederhana penyebaran: pembuatan laporan.
   c. Contoh kompleks penyebaran: penerapan proses data mining secara paralel pada departemen lain.

2.2 Penelian Terdahulu

Tabel 2.2 Penelitian Terdahulu

No	Peneliti, Tahun Peneliti	Judul Penelitian	Variabel Penelitian	Desain Penelitian	Kesimpulan Penelitian
1	L Thurn, PG Lindqvist, dkk. 2015	Plasenta invasif abnormal — prevalensi, faktor risiko, dan kecurigaan antenatal: hasil dari kohort kehamilan berbasis populasi yang banyak diteliti di negara-negara Nordik	Umur ibu hamil, perokok, berat badan, kondisi sebelumnya, Komplikasi pada kehamilan ini	Dalam studi NOSS, dokter melaporkan kasus AIP dari bangsal ibu dan data divalidasi terhadap pendaftar kesehatan nasional.	Temuan kami menunjukkan bahwa tingkat CS yang lebih rendah dalam populasi mungkin merupakan cara paling efektif untuk menurunkan kejadian AIP. Penilaian ultrasonografi terfokus pada wanita berisiko tinggi kemungkinan akan meningkatkan kecurigaan antenatal. PPH sebelumnya adalah faktor risiko baru yang terkait dengan peningkatan prevalensi AIP.
2	Dian Gusta Anggraini, Rise Marisa Satri, 2015	Kehamilan risiko tinggi di Puskesmas lubuk gadang Kabupaten solok selatan	Umur Paritas Kehamilan sebelumnya Persalinan sebelumnya	Penelitian ini menggunakan desain *cross* *Sectional study*.	Dapat disimpulkan Bahwa lebih dari separuh responden dengan kehamilan risiko tinggi. Lebih dari separuh responden berada pada kategori umur tidak beresiko,

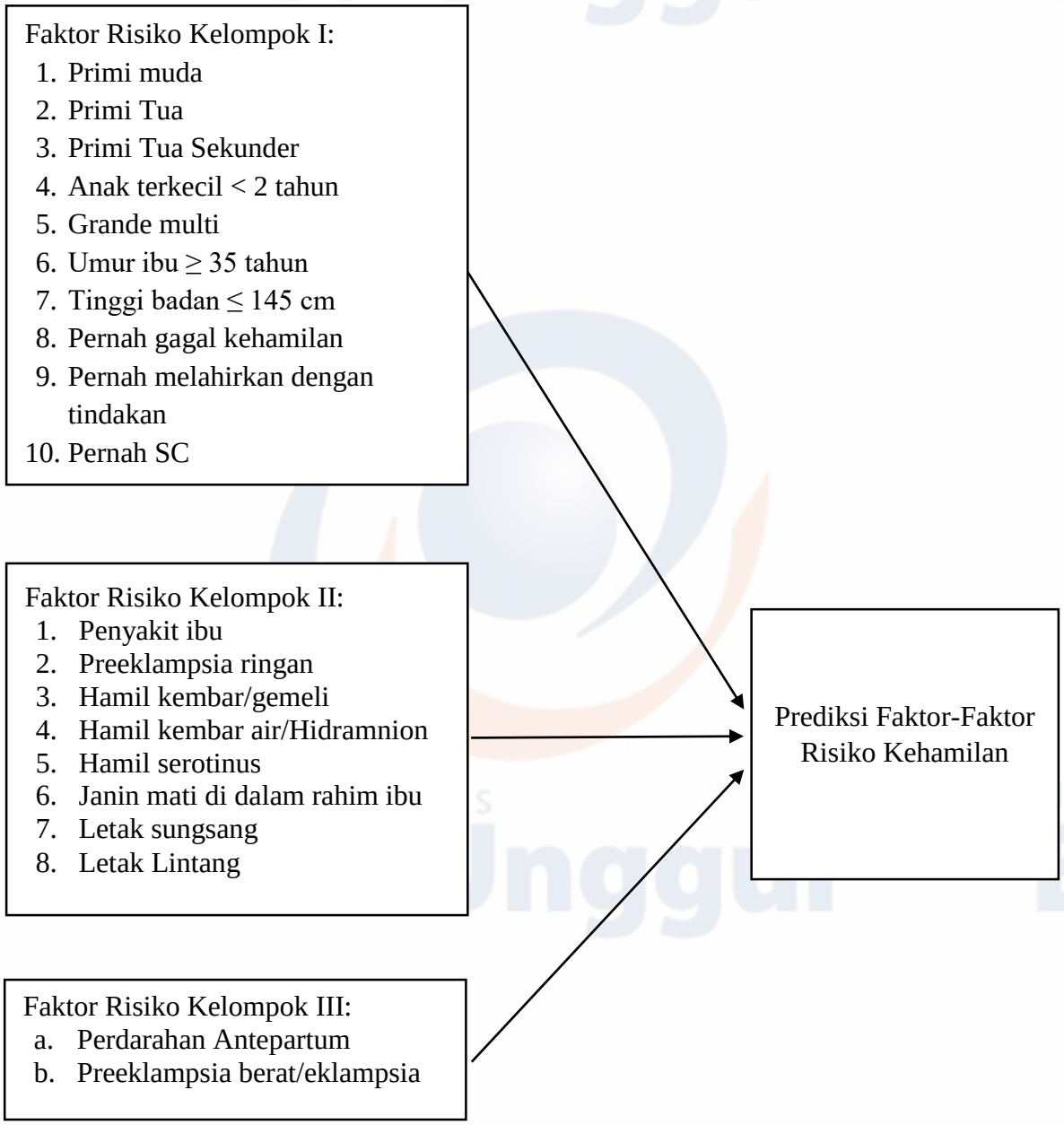
No	Peneliti, Tahun Peneliti	Judul Penelitian	Variabel Penelitian	Desain Penelitian	Kesimpulan Penelitian
					paritas berisiko, riwayat kehamilan sebelumnya berisiko, dan sebagian besar responden dengan riwayat persalinan sebelumnya berisiko.
3	Prima Hidayah, Heni Puji Wahyuningsih, dan Kusminatun 2018	Hubungan Tingkat Risiko Kehamilan dengan Kejadian Komplikasi Persalinan di RSUD Panembahan Senopati Bantul	Variabel bebas penelitian ini adalah tingkat risiko kehamilan dan variabel terikatnya adalah kejadian komplikasi persalinan.	Metode penelitian menggunakan survei analitik dengan rancangan cross sectional corelasional . Penelitian ini dianalisis secara univariabel, bivariabel dengan uji Chi Square dan Ratio Prevalence (RP) dan multivariabel dengan uji regresi logistik.	Kesimpulan bahwa ada hubungan antara tingkat risiko kehamilan dengan kejadian komplikasi persalinan.
4	Noviandi, 2018	Implementasi Algoritma Decision Tree C4.5 untuk Prediksi Penyakit Diabetes	Variabel dalam penelitian ini meliputi : Kadar gula darah, tekanan darah, insulin *body mas index* dan usia.	Metode yang digunakan dalam penelitian ini adalah metode klasifikasi *data mining* dengan	Dalam penelitian yang dilakukan adalah melakukan model prediksi penyakit diabetes. Prediksi penyakit diabetes dilakukan dengan

No	Peneliti, Tahun Peneliti	Judul Penelitian	Variabel Penelitian	Desain Penelitian	Kesimpulan Penelitian
				algoritma C4.5	menggunakan algoritma decision tree C4.5. Akurasi model prediksi adalah 70,32% dengan 9 rule dengan 4 rule dengan class tidak dan 5 rule dengan class tidak.
5	Dwi Meylitasari Br. Tarigan, Dr. Dian Palupi Rini, M.T, Verlly Puspita, 2017	Perancangan *Data mining* untuk Klasifikasi Prediksi Penyakit ISPA dengan Algoritma C4.5	Data yang diambil meliputi : Nama, usia, jenis kelamin dan penderita hipertensi	Algoritma C4.5 menggunakan gain ratio sebagai penentu simpul akar, internal dan daun. Algoritma C4.5 dapat menangani data numenrk (kontinyu) dan diskret.	Pada penelitian ini mengusulkan perancangan *data mining* untuk memprediksi penderita penyakit ISPA untuk kategori Akut dan Tidak Akut menggunakan Klasifikasi Algoritma C4.5, karena rules yang terbentuk sederhana. Untuk pengukuran atau memilih atribut sebagai akar berdasarkan nilai gain tertinggi dari atribut-atribut yang ada.
6	Moghimipour, Ida dan Ebrahimpou,	Membandingkan Metode Pohon Keputusan Dari Tiga Perangkat Lunak	Umur ibu, jenis kelamin bayi, tahun lahir, metode persalinan, berat badan,	Selain itu, deskripsi singkat dari tiga perangkat lunak	Akurasi tertinggi decision tree

No	Peneliti, Tahun Peneliti	Judul Penelitian	Variabel Penelitian	Desain Penelitian	Kesimpulan Penelitian
	2014	*Data mining*	tinggi badan.	penambangan data, yaitu SPSS-Clementine, RapidMiner dan Weka juga disediakan.	

2.3 Kerangka Berpikir

Berikut merupakan kerangka berfikir yang digunakan dalam penelitian ini:



Gambar 2.2 Kerangka Berfikir

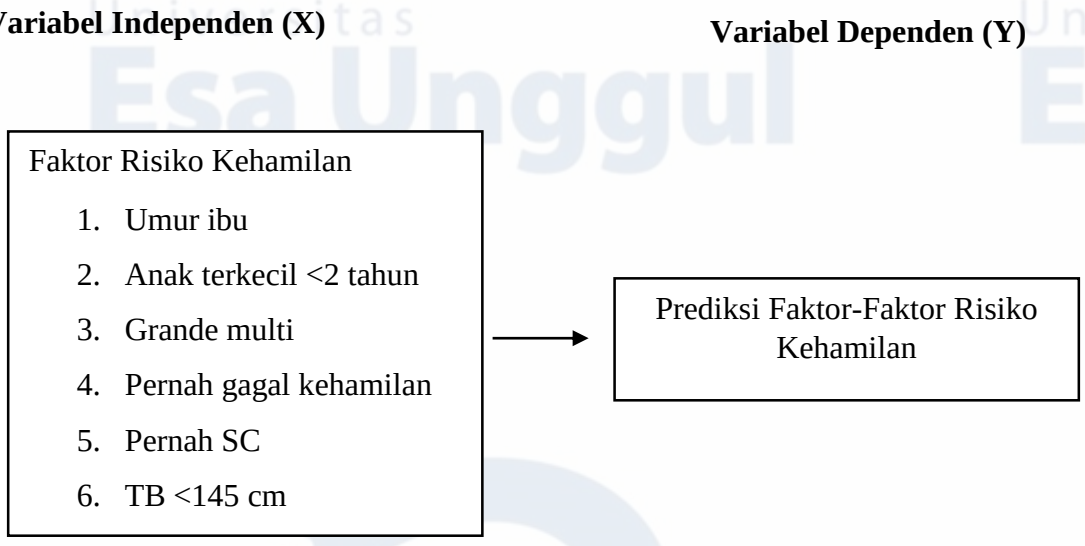
(Sumber: Rochyati, 2011)

Berdasarkan bagan kerangka berpikir diatas, untuk mengetahui prediksi faktor-faktor risiko kehamilan, maka kita perlu untuk mengetahui apa saja yang menjadi faktor-faktor dari risiko kehamilan tersebut. Faktor risiko kehamilan terdiri dari 3 kelompok, Faktor Risiko kelompok I adalah, primi muda, primi tua, primi tua sekunder, anak terkecil < 2 tahun, grande multi, umur ibu ≥ 35 tahun, tinggi badan ≤ 145 cm, pernah gagal kehamilan, pernah melahirkan dengan tindakan dan pernah sc. Faktor risiko kelompok II yaitu: penyakit ibu, preeklampsia ringan, hamil kembar/gemeli, hamil kembar air/hidramnion, hamil serotinus, janin mati di dalam rahim ibu, letak sungsang dan letak lintang. Sedangkan faktor risiko kelompok III yaitu perdarahan antepartum dan preeklampsia berat/eklampsia. Dari faktor-faktor risiko tersebut, akan berdampak pada tingginya angka kematian ibu.

Diharapkan dengan adanya prediksi faktor-faktor risiko kehamilan, maka akan menjadi dasar pengambilan keputusan tata laksana penanganan pasien antenatal di rumah sakit yang dapat membantu meningkatkan *outcome* pasien.

2.4 Kerangka Konsep

Berikut kerangka konsep dari penelitian ini:

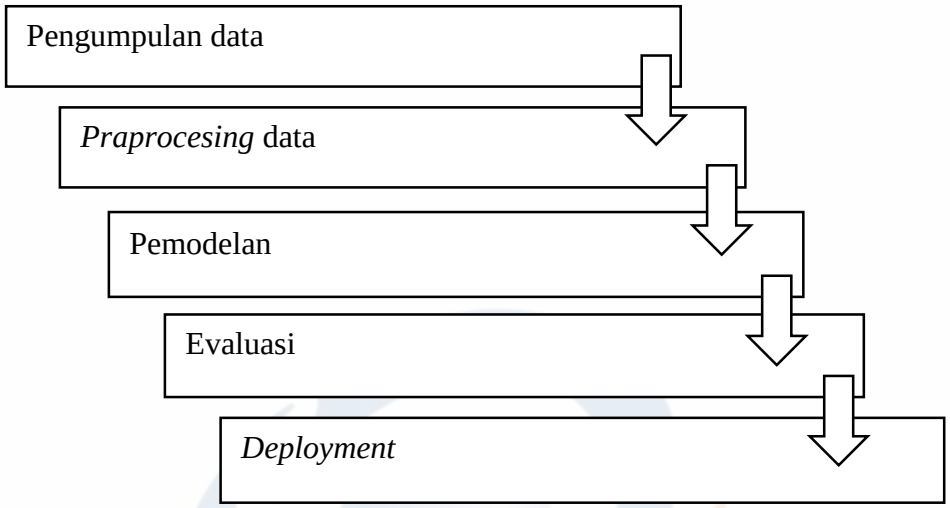


Gambar 2.3
Kerangka Konsep

Dari bagan kerangka konsep di atas, dapat dilihat bahwa dalam penelitian ini hanya fokus meneliti faktor-faktor risiko kehamilan pada kelompok 1 yang terdiri dari Umur Ibu, Anak terkecil <2 tahun, Grande multi, Pernah gagal kehamilan, Pernah SC dan TB <145 cm. Hal itu karena berdasarkan teori Pudji Rochyati, dari ke-3 kelompok faktor risiko yang ada, pada kelompok I baru ada potensi gawat obstetrik atau merupakan kehamilan dengan risiko rendah. Dibandingkan dengan faktor risiko yang ada di kelompok II sudah ada gawat obstetrik atau kehamilan risiko tinggi, dan pada kelompok III ada gawat darurat obstetrik atau merupakan kehamilan dengan risiko sangat tinggi. Oleh karena itu peneliti tidak melakukan penelitian pada ke-3 kelompok faktor tersebut, penelitian hanya dilakukan pada kelompok I, karena apabila peneliti melakukan penelitian faktor-faktor risiko pada ke-3 kelompok tersebut, maka sudah jelas terlihat bahwa yang paling beresiko pada kehamilan adalah pada kelompok III. Sehingga peneliti memutuskan untuk melakukan penelitian hanya pada 1 kelompok, yaitu kelompok I.

2.5 Kerangka Operasional

Dalam penelitian ini dilakukan beberapa langkah-langkah atau tahapan penelitian seperti pada gambar dibawah ini :



Gambar 2.4
Kerangka Operasional

Berikut adalah penjelasan mengenai kerangka operasional diatas :

1. Pengumpulan data

   Data pada penelitian ini adalah data sekunder yang diperoleh dari data rekam medis di Rumah Sakit Patria IKKT. Data yang dikumpulkan adalah data kunjungan antenatal pada tahun 2017-2018 dengan jumlah 3.295.

2. *Praprocesing* data

   Untuk mendapatkan data berkualitas, beberapa teknik yang perlu dilakukan adalah sebagai berikut :

   a. Pembersihan Data (*Data Cleaning*)

      Pembersihan data dilakukan dengan memilih data pasien antenatal yang relevan sesuai dengan krtiteria faktor-faktor risiko kehamilan yaitu umur ibu, anak terkecil <2 tahun, grande multi, pernah gagal kehamilan, pernah SC dan TB <145 cm.

b. Transformasi Data (*Data Transformation*)

Untuk meningkatkan akurasi dan efisiensi algritma dalam *data mining decision tree* C4.5. Data yang digunakan dalam penelitian ini bernilai kategorikal.

3. Pemodelan

Menggunakan Algoritma yang paling mudah dipahami pada Proses *Data mining* yaitu Algoritma *Decision Tree* C4.5 yang berfungsi untuk mendapatkan pohon keputusan kelompok faktor-faktor risiko kehamilan. Pada tahap ini dilakukan eksperimen dan pengujian metode yang digunakan yaitu menghitung dan mendapatkan *rule-rule* yang ada pada algoritma yang diusulkan.

4. Evaluasi

Setelah data diolah maka dapat diuji tingkat akurasinya untuk melihat kinerja dari metode algoritma C4.5. Pengujian tingkat akurasi dilakukan dengan menggunakan *confussion matrix* dan kurva ROC/AUC (*Area Under Cover*).

5. *Deployment*

Merupakan tahapan membuat laporan hasil kegiatan *data mining*. Laporan akhir mengenai pengatahuan yang didapat atau pengenalan pola pada data dalam proses *data mining* dan dipresentasikan dalam bentuk grafik atau deskripsi yang mudah dipahami.

2.6 Definisi Operasional

Tabel 2.1
Definisi Operasional

No	Variabel	Definisi Variabel	Indikator	Alat Ukur	Hasil Ukur	Skala Ukur
1.	Faktor Risiko Kehamilan	Prediksi faktor-faktor/kondisi pada ibu hamil yang dapat menyebabkan kemungkinan risiko/bahaya pada persalinan yang dapat menyebabkan kematian atau kesakitan pada ibu dan bayinya.	Faktor-faktor/kondisi pada ibu hamil yang dapat menyebabkan kemungkinan risiko/bahaya pada persalinan yang dapat menyebabkan kematian atau kesakitan pada ibu dan bayinya.	Observasi	1. Risiko 2. Tidak Berisiko	*Data mining* dengan klasifikasi decision tree C4.5
2.	Umur Ibu	Umur adalah lama hidup ibu hamil dalam tahun sejak lahir sampai dengan ulang tahun terakhir.	Umur ibu yang dikelompokkan menjadi : 1. Remaja akhir 2. Dewasa awal 3. Dewasa akhir	Data sekunder	1. Remaja akhir 17-25 Tahun 2. Dewasa awal 26-35 Tahun 3. Dewasa akhir 36-45 Tahun	Ordinal

No	Variabel	Definisi Variabel	Indikator	Alat Ukur	Hasil Ukur	Skala Ukur
			4. Lansia awal		4. Lansia awal 46-55 Tahun	
3.	Anak terkecil umur < 2 tahun	Anak terkecil umur < 2 tahun yaitu jarak kelahiran dengan anak terkecil kurang dari 2 tahun.	Selisih umur bayi yang dilahirkan.	Data sekunder	1. Anak terkecil umur < 2 tahun 2. Anak terkecil umur ≥ 2 Tahun	Nominal
4.	Grande multi	Grande multi adalah wanita yang telah melahirkan anak lebih dari 4 kali.	Jumlah anak yang dimiliki ibu.	Data sekunder	1. Melahirkan ≥ 4 kali 2. Melahirkan < 4 kali	Nominal
5.	Pernah gagal kehamilan	Pernah gagal kehamilan adalah kehamilan kedua/ketiga, dimana kehamilan yang sebelumnya mengalami keguguran, lahir belum	Kondisi ibu pada kehamilan sebelumnya yang pernah mengalami kegagalan.	Data sekunder	1. Ya 2. Tidak	Nominal

No	Variabel	Definisi Variabel	Indikator	Alat Ukur	Hasil Ukur	Skala Ukur
		cukup bulan, lahir mati, lahir hidup lalu mati umur ≤ 7 hari.				
6.	Pernah SC	Pernah SC adalah suatu tindakan insisi perut dan uterus dengan tujuan untuk mengeluarkan janin.	Tindakan operasi sesar yang pernah dilakukan pada kehamilan sebelumnya.	Data sekunder	1. Ya 2. Tidak	Nominal
7.	TB <145 CM	Ukuran tinggi badan ibu yang diukur dengan menggunakan miknotase dalam satuan cm dengan batas minimal tinggi badan 145 cm.	Ukuran tinggi badan pada ibu hamil <145 cm atau >145 cm.	Data sekunder	1. TB <145 CM 2. TB ≥ 145 CM	Nominal