

**IMPLEMENTASI DATA MINING UNTUK DIAGNOSIS PENYAKIT HIPERTENSI
PADA KEHAMILAN DENGAN MENGGUNAKAN DECISION TREE
(STUDY KASUS: RSIA YK MADIRA PALEMBANG)**

RIKA ANISA WULANDARI NIM : 10142152

PEMBIMBING AKADEMIK: P.H SAKSONO, S.T., M.Sc., Ph.D

1. Masalah & Penyebab

Hipertensi atau peningkatan tekanan darah adalah keadaan dimana supply oksigen dan nutrisi dibawa oleh darah terhambat dalam proses pengirimannya ke jaringan tubuh yang membutuhkan. Penyakit ini dapat menimbulkan gangguan pada organ tubuh yang lain karena menyebabkan organ-organ tersebut harus bekerja lebih keras. Pada awal menderita penyakit ini tidak ada tanda-tanda yang muncul sehingga tidak dapat dideteksi. Hipertensi baru akan terdeteksi ketika dilakukan pemeriksaan untuk penyakit yang berkaitan dengan hipertensi seperti pemeriksaan diabetes, stroke, dan sebagainya. Oleh sebab itu, penyakit ini dikenal dengan nama *silent killer*.

Penyakit hipertensi juga tidak bisa terlepas dari kehamilan yang tengah dialami oleh para ibu. Kurangnya kontrol dan pengawasan terhadap tekanan darah pada ibu hamil dapat menyebabkan hipertensi yang tidak diketahui. Hipertensi dapat berpengaruh perkembangan janin pada kehamilan, bahkan tekanan darah yang berlebih juga dapat mengakibatkan keguguran pada janin. Sehingga penyakit hipertensi sangat berbahaya jika menyerang ibu hamil, terutama bagi janin itu sendiri.

2. Alternatif Solusi

Dari masalah yang ditimbulkan oleh penyakit hipertensi pada kehamilan, diperlukan kontrol dan pengawasan sedini mungkin terhadap ibu hamil yang terdeteksi penyakit hipertensi agar tidak terjadi komplikasi terhadap kehamilan yang dapat membahayakan janin dan kesehatan ibu hamil. Diagnosis dilakukan dengan menggunakan data mining dalam menambang data-data hasil pemeriksaan ibu hamil. Teknik yang digunakan adalah metode klasifikasi Decision Tree (Pohon Keputusan), karena metode tersebut memiliki kemudahan dalam interpretasi hasil yang diharapkan.

3. Hasil yang Diharapkan

Hasil yang diharapkan dari penerapan data mining dalam diagnosis penyakit hipertensi pada kehamilan adalah mendapatkan hasil diagnosis yang akurat pada data rekam medis (kehamilan) terhadap penyakit hipertensi. Sehingga dapat dilakukan tindakan medis lebih

lanjut terhadap ibu hamil yang terdeteksi penyakit hipertensi, agar tidak terjadi komplikasi pada saat kehamilan yang dapat membahayakan kesehatan janin dan dirinya.

3. Kebutuhan Data dan Cara Mendapatkan

Data yang dibutuhkan dalam diagnosis penyakit hipertensi pada kehamilan yaitu data source. Data source adalah sumber data yang bersifat real yang diambil dari suatu transaksi. Dalam hal ini, data source yang digunakan adalah data rekam medis dan data hasil pemeriksaan kehamilan yang juga bertindak sebagai variabel input.

Data source tersebut akan dikonversikan pada sebuah database atau dibangun data warehouse sebagai gudang penampung data. Dari database atau data warehouse itulah akan dilakukan ekstrak data yang akan membentuk suatu pola dan memberikan nilai atau informasi yang menarik dalam data terpilih, proses tersebut disebut juga data mining.

Hasil dari informasi yang menarik tersebut adalah data diagnosis penyakit hipertensi pada kehamilan, yang disebut juga dengan variabel output.

Cara mendapatkan data yang dibutuhkan untuk diagnosis penyakit hipertensi pada kehamilan yaitu dengan melakukan:

1. Pengamatan (Observasi)

Mengumpulkan data dengan cara mengamati dan mencatat secara langsung.

2. Wawancara (Interview)

Mengadakan tanya jawab secara langsung dengan pihak terkait, seperti dokter, petugas administrasi, dan staf lainnya.

3. Dokumentasi

Menggunakan data-data dari rumah sakit dan data-data yang berhubungan dengan data rekam medis dari pasien yang berobat dirumah sakit.

4. Metode Pelaksanaan

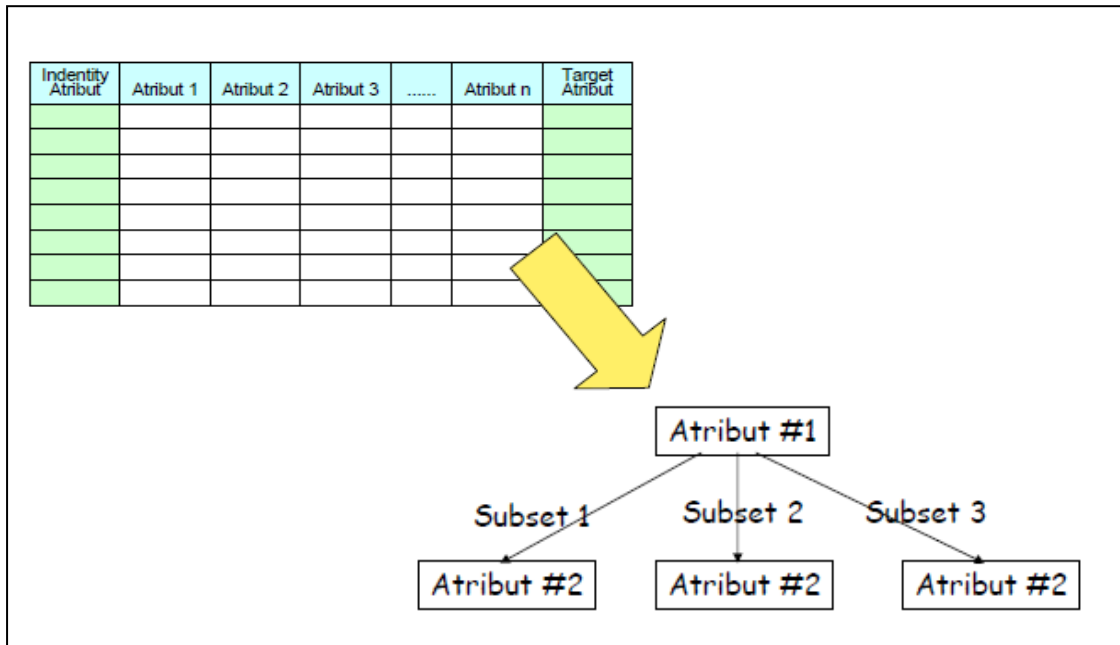
Metode yang digunakan dalam penelitian ini adalah metode klasifikasi dengan menggunakan Decision Tree. *Decision tree* memiliki beberapa keunggulan yaitu mudah dalam pengembangan sebuah model, mudah dipahami oleh pengguna, dan mampu menangani *noisy data* dan *unknown data*.

Metode Decision Tree mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami (dapat diekspresikan dalam *Structured Query Language*). Decision tree adalah struktur yang dapat digunakan membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian aturan keputusan.

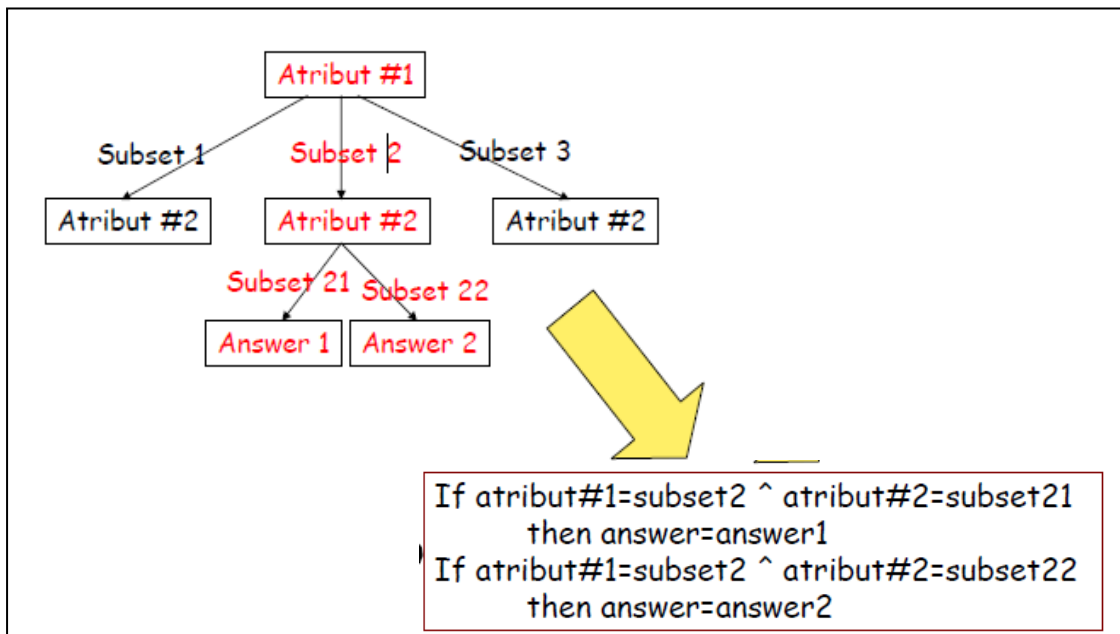
Adapun proses dalam decision tree adalah sebagai berikut:

- Mengubah bentuk data (tabel) menjadi model tree.
- Mengubah model tree menjadi rule
- Menyederhanakan Rule (Pruning)

Proses data menjadi tree :



Proses mengubah tree menjadi rules :



Algoritma yang digunakan pada metode decision tree adalah C4.5. Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- Pilih atribut sebagai akar
- Buat cabang untuk tiap-tiap nilai
- Bagi kasus dalam cabang
- Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama

Untuk memilih atribut sebagai akar, didasarkan pada *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain* digunakan rumus:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Dengan:

S : Himpunan kasus
A : Atribut
n : Jumlah partisi atribut A
|S_i| : Jumlah kasus pada partisi ke i
|S| : Jumlah kasus dalam S

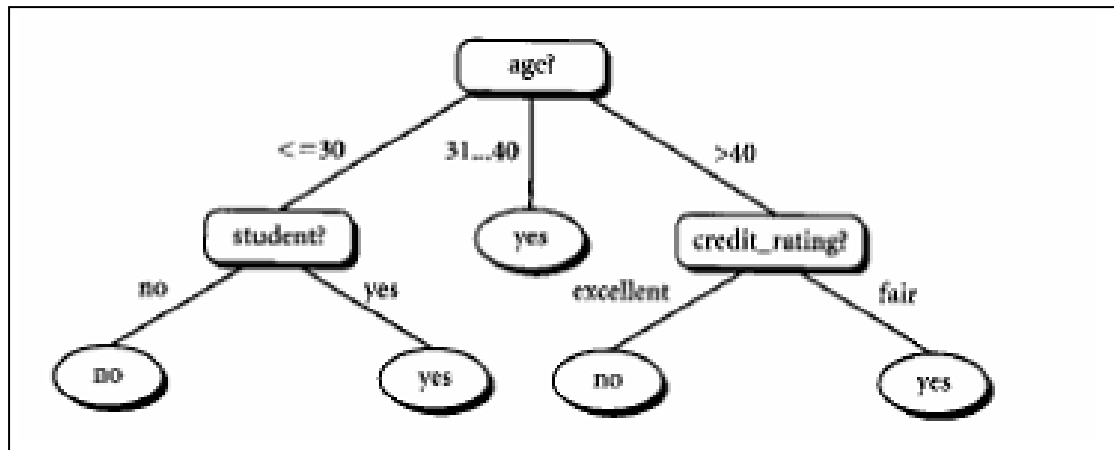
Sedangkan untuk menghitung nilai entropy adalah sebagai berikut:

$$Entropy(S) = \sum_{i=1}^n - p_i * \log_2 p_i$$

dengan:

S : Himpunan Kasus
A : Fitur
n : Jumlah partisi S
p_i : Proporsi dari S_i terhadap S

Sehingga terbentuklah suatu model decision tree (Pramudiono, 2008) :



Aplikasi yang digunakan pada proses pembentukan pola decision tree adalah Weka. Weka merupakan aplikasi data mining yang berbasis opens source (GPL) dan berengine java. Aplikasi ini dikembangkan pertama kali oleh sebuah universitas selandia baru yang bernama Universitas Waikato sebelum menjadi bagian di pentaho. Weka terdiri dari koleksi algoritma machine learning yang dapat digunakan untuk melakukan generalisasi satu formulasi dari sekumpulan data sampling. Kekuatan Weka sebenarnya terletak pada algoritma yang makin lengkap dan canggih, kesuksesan data mining tetap terletak pada factor pengetahuan manusianya sendiri. Hal tersebut bisa dilakukan jika adanya pengumpulan data yang berkualitas sehingga akan menjamin keakuratan formulasi yang diharapkan.

Mengetahui,
Pembimbing Akademik

Palembang, 21 Oktober 2013
Penulis

P.H. Saksono, S.T., M.Sc., Ph.D

Rika Anisa Wulandari