



KILAT

JURNAL KAJIAN ILMU DAN TEKNOLOGI

Rakhmadi Irfansyah Putra

PERANCANGAN APLIKASI PENGEMBALIAN BERKAS TERHAPUS PADA NTFS

Rizqia Cahyaningtyas
Awit Lela Sigi

PERANCANGAN APLIKASI PENGELOLAAN RUMAH TANGGA
LABORATORIUM KOMPUTER STT-PLN

Dian Hartanti

MODEL CLUSTERING MENGGUNAKAN ALGORITMA K-MEANS PADA DATA
KELUHAN PELANGGAN PT PLN (PERSERO)
(STUDI KASUS : PT PLN (PERSERO) DISTRIBUSI JAKARTA DAN TANGERANG)

Abdul Haris; Nina Nirmaya

APLIKASI KONVERSI AKSARA SUNDA KE BAHASA INDONESIA BERBASIS WEB
MENGGUNAKAN PHP MYSQL

Riki Ruli A. Siregar
Alwi Baraqbah

SISTEM KONTROL PADA PENGGUNAAN ENERGI LISTRIK BERBASIS ANDROID
STUDI KASUS : SEKOLAH MENENGAH PERTAMA NEGERI (SMPN) 7 PONTIANAK

Yessy Asri

ANALISA PERBANDINGAN KEPUTUSAN METODE KLASIFIKASI DECISION TREE
DAN NAÏVE BAYES DALAM PENENTUAN DIAGNOSA HIPERTENSI

Marliana Sari

PUSAT INFORMASI KEMAHASISWAAN DENGAN MENGGUNAKAN PHP, MYSQL
DAN METODE MVC

Inge Handriani

FLOWCHART SISTEM PENAGIHAN PADA PERUSAHAAN JASA KONSULTAN

Herman Bedi Agtriadi

RANCANG BANGUN SISTEM MONITORING PADA BANGUNAN PENGENDALI
SEDIMEN BERBASIS ANDROID DI PLTA

Indah Handayasari
Aziz Maulana

DESAIN ALTERNATIF JEMBATAN MENGGUNAKAN PLAT GIRDER
(STUDI KASUS JEMBATAN RSUD KOTA TANGERANG)

Irma Wirantina Kustanrika

PERHITUNGAN SINYAL PADA SIMPANG DENGAN METODE WEBSTER

Mukhlis Akhadi

MEMPRODUKSI BAHAN SEMIKONDUKTOR DI DALAM TERAS REAKTOR NUKLIR

ISSN 2089-1245



9 772089 124519

SEKOLAH TINGGI TEKNIK - PLN (STT-PLN)

KILAT

VOL.4

NO.1

HAL. 1 - 119

APRIL 2015

ISSN 2089 - 1245

ANALISA PERBANDINGAN KEPUTUSAN METODE KLASIFIKASI DECISION TREE DAN NAÏVE BAYES DALAM PENENTUAN DIAGNOSA HIPERTENSI

Yessy Asri ST., MMSI
Jurusan Teknik Informatika
Sekolah Tinggi Teknik PLN Jakarta
e-mail : yesfar2@gmail.com

Abstrak

Dunia bisnis yang penuh persaingan membuat para pelakunya harus selalu memikirkan strategi-strategi terobosan yang dapat menjamin kelangsungan bisnis mereka. Salah satu aset utama yang dimiliki oleh perusahaan masa kini adalah data bisnis dalam jumlah yang luar biasa banyak. Ini melahirkan kebutuhan akan adanya teknologi yang dapat memanfaatkannya untuk membangkitkan "pengetahuan-pengetahuan" baru, yang dapat membantu dalam pengaturan strategi bisnis. Teknologi data mining hadir sebagai solusi.

Makalah ini akan mengulas permasalahan bisnis yang ada dalam mendeteksi potensi hipertensi dari data-data historis atau data set yang ada selama ini. Metode yang digunakan dalam penulisan ini adalah Decision Tree dan Naive-Bayes yang sama-sama merupakan bagian dari metode Data Mining Klasifikasi (Supervised Learning). Metode Decision Tree dan Metode Naive mengubah data training yang dinyatakan dalam bentuk tabel dengan atribut dan record menjadi menjadi pohon keputusan (decision tree) dan aturan-aturan keputusan (rule). Metode Naive Bayes data dinyatakan dalam tabel dengan atribut dan record terutama untuk data-data-data yang tidak konsisten maupun bias untuk mendapatkan hipotesis dari suatu keputusan.

Dari data training yang diolah terdapat perbedaan keputusan yang dihasilkan antar kedua metode tersebut, hal ini terlihat pada proses pembentukan pohon / tree pada tree yang dikembangkan untuk atribut usia dengan instance tua terdapat 1 data training yang menyatakan ya dan 1 data menyatakan tidak, maka untuk menentukan pilihan memerlukan campur tangan seorang pakar, sehingga dapat disimpulkan bahwa pada proses pembentukan tree jika terdapat kondisi leaf node memiliki nilai entropy yang sama, maka membutuhkan inputan pengetahuan dari sang pakar. Hasil keputusan dari metode Naive Bayes lebih akurat dari pada hasil keputusan metode Decision Tree.

Kata Kunci : Keputusan, Klasifikasi, Decision Tree, Naive Bayes, Diagnosa Hipertensi.

Abstract

The business world is full of competition makes the perpetrators must always think of breakthrough strategies that can ensure the continuity of their business. One of the main assets owned by the company today is business data in an incredible amount of many. This gave rise to the need for technology that can use it to generate "knowledge-knowledge", which can help in setting business strategy. Data mining technology is present as a solution.

This paper will review the existing business problems in detecting potential hypertension historis of data or data sets that exist for this. The method used in this paper is the Decision Tree and Naive-Bayes equally a part of Data Mining method Classification (Supervised Learning). Method of Decision Tree and Naive method to change the training data set forth in tabular form with the attributes and records become into a decision tree (decision tree) and decision rules (rule). Naive Bayes methods of data expressed in tables with attributes and records primarily for data-data that is inconsistent and biased to get the hypothesis of a decision.

From the training data are processed there are differences between the two decisions that produced these methods, it is seen in the process of forming the tree / tree on a tree that was developed for the attribute of old age with the instance there is one training data stating the data stated yes and 1 no, then to determine option requires the intervention of an expert, so it can be concluded that the process of establishing if there is a condition tree leaf node has the same entropy value, then it requires the input of knowledge from the experts. The decision of the Naive Bayes method is more accurate than the results of decisions Decision Tree method.

Keywords: Decision, Classification, Decision Tree, Naive Bayes, Diagnosis Hypertension.

1. Pendahuluan

1.1 Latar Belakang Masalah

Pada era informasi digital saat ini, dalam berbagai bidang kehidupan telah terjadi pengumpulan dan penumpukan data dalam skala besar, misalnya: data transaksi penjualan pada sebuah swalayan, data

pasien pada rumah sakit, data rekening pada bank, dan sebagainya. Ukuran basis data meningkat baik dalam jumlah record (baris data) maupun jumlah atribut pada record [FAY, 1996]. Hal ini didukung oleh perkembangan perangkat keras dan teknologi basis data yang memungkinkan penyimpanan dan

pengaksesan data secara efisien dan murah [FAY,1996].

Kumpulan data jika dibiarkan begitu saja tidak dapat memberikan nilai tambah berupa pengetahuan yang bermanfaat. Pengetahuan yang bermanfaat ini misalnya, dari basis data Penyakit Hipertensi pada pasien, dapat diperoleh pengetahuan tentang hubungan antara usia pasien, berat badan pasien serta jenis kelamin pasien dengan kemungkinan seseorang terkena penyakit hipertensi. Pengetahuan ini dapat digunakan untuk mengetahui bagaimana pengaruh usia, berat badan serta jenis kelamin seseorang terhadap kemungkinan hipertensi atau tidak.

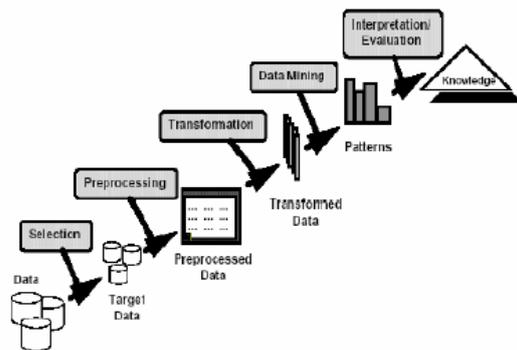
1.2 Tujuan Penelitian

Tujuan utama dari penulisan ini adalah melakukan studi perbandingan hasil keputusan yang didapatkan dari data training yang bias terhadap mini data mining dalam penentuan kemungkinan seseorang menderita hipertensi atau tidak dengan metode klasifikasi decision tree dan naïve bayes.

2. Landasan Teori

2.1. Pengertian Data Mining

Data Mining adalah keseluruhan proses untuk mencari dan mengidentifikasi pola (pattern) dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti. Data Mining memiliki arsitektur sebagai berikut :



Gambar 1. Arsitektur Data Mining

2.2. Konsep Data Mining

Dengan kecanggihan teknologi yang semakin meningkat, kini database mampu untuk menyimpan data berkapasitas terabytes atau samadengan 1,000,000,000,000 bytes data. Dalam kumpulan data yang sangat banyak ini, tersimpan informasi tersembunyi yang merupakan strategi penting.



Gambar 2. Ilustrasi Kaya Data Tetapi Miskin Informasi

Konsekuensinya, keputusan-keputusan yang penting sering dibuat tidak berdasarkan data yang kaya informasi yang disimpan di database tapi cukup pada intuisi pembuat keputusan. Pelebaran jarak antara data dan informasi untuk sebuah pengembangan sistematis dari tools data mining yang akan merubah kuburan data menjadi "tambang emas" dari pengetahuan



Gambar 3. Data mining - searching for knowledge in your data.

Data Mining Skala Kecil

Pada sisi lain, ada suatu konsep data mining dalam skala yang lebih kecil dimana konsep ini mempunyai perbedaan dalam input data. Dalam data mining, data input didapat dari file penyimpanan data. Tetapi pada konsep data mining dalam skala yang lebih kecil, data input hanya bisa didapatkan dari analisis bisnis.

2.3. Teknik-Teknik Data Mining

Beberapa teknik Data Mining yang sering digunakan dalam dunia industri antara lain:

2.3.1. Clustering

Analisis cluster merupakan salah satu teknik data mining yang bertujuan untuk mengidentifikasi sekelompok obyek yang mempunyai kemiripan karakteristik tertentu yang dapat dipisahkan dengan kelompok obyek lainnya, sehingga obyek yang berada dalam kelompok yang sama relatif lebih homogen daripada obyek yang berada pada kelompok yang berbeda.

2.3.2. Association Rule (Market Basket Analysis)

Association Rule adalah teknik data mining yang digunakan untuk menemukan aturan assosiatif antara suatu kombinasi item. Salah satu contoh penerapan Association Rule adalah Market Basket Analysis.

2.3.3. Classification

Klasifikasi merupakan proses untuk menemukan sekumpulan model yang menjelaskan dan membedakan kelas-kelas data, sehingga model tersebut dapat digunakan untuk memprediksi nilai suatu kelas yang belum diketahui pada sebuah objek.

2.3.4. Regression (Predictive)

Analisis regresi merupakan teknik untuk membantu menganalisis hubungan antara suatu peristiwa atau keadaan yang terjadi akibat peristiwa yang lain.

2.3.5. Classification

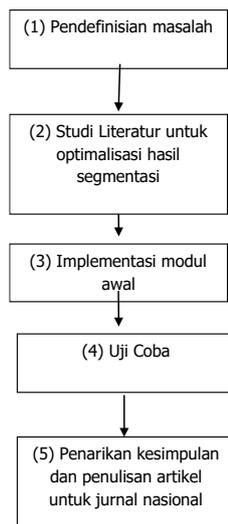
Klasifikasi merupakan proses untuk menemukan sekumpulan model yang menjelaskan dan membedakan kelas-kelas data, sehingga model tersebut dapat digunakan untuk memprediksi nilai suatu kelas yang belum diketahui pada sebuah objek.

2.4.6 Regression (Predictive)

Analisis regresi merupakan teknik untuk membantu menganalisis hubungan antara suatu peristiwa atau keadaan yang terjadi akibat peristiwa yang lain.

3. Metodologi Penelitian

Metodologi dalam penelitian akan menghasilkan perbandingan jumlah keputusan yang akan dihasilkan dari kedua metode klasifikasi, yaitu decision tree dan naive bayes. Berikut adalah metodologi penelitian yang akan dilakukan :



Gambar 4. Diagram metodologi penelitian

- Mendefinisikan masalah, yaitu melakukan identifikasi masalah, tujuan dan fungsi dalam analisa perbandingan dua metode klasifikasi.
- Studi literatur untuk optimalisasi hasil. Dalam tahap ini, dipelajari tahapan-tahapan algoritma dari setiap metode menggunakan data training dalam skala kecil / mini data mining.
- Melakukan implementasi algoritma dari kedua metode klasifikasi data mining dengan data training "Hipertensi" dengan atribut usia, berat badan dan jenis kelamin.

- Melakukan serangkaian percobaan, pembelajaran, dan pengujian.
- Melaporkan hasil penelitian, dan mempublikasikan hasilnya pada jurnal nasional.

4. Analisa Kebutuhan Sistem

Analisa kebutuhan data merupakan tahapan mengidentifikasi tentang data-data yang diperlukan dalam menganalisa perbandingan keputusan metode klasifikasi Decision Tree dengan Naïve Bayes dalam penentuan diagnosa Hipertensi. Tujuannya adalah untuk melakukan studi perbandingan hasil keputusan/rule yang didapatkan dari data training yang bias terhadap mini data mining dalam penentuan kemungkinan seseorang menderita hipertensi atau tidak dengan metode klasifikasi decision tree dan naïve.

5. Hasil dan analisis

5.1. Pohon Keputusan (Decision Tree) – ID3

Metode decision tree memiliki tahapan dalam membentuk keputusan-keputusan, tahapan tersebut ditunjukkan dalam gambar berikut :



Gambar 5. Tahapan Decision Tree

Langkah yang digunakan untuk mengubah data menjadi tree :

- Menentukan node yang terpilih
Menggunakan nilai Entropy dari setiap kriteria dengan sampel data yang ditentukan. Node yang terpilih adalah kriteria dengan nilai entropy yang paling kecil.
 $Entropy(S) = -p+ \log_2 p+ -p- \log_2 p-$
- Menyusun tree
Sesuai dengan aturan-aturan yang ada pada konsep Tree (Pohon). Data diambil dengan 8 sampel dengan pemikiran bahwa yang mempengaruhi seseorang menderita hipertensi atau tidak Usia, Berat badan dan jenis kelamin.

Tabel 1. Data Training

Nama	Usia	Berat	Kelamin	Hipertensi
Ali	muda	overweight	pria	ya
Edi	muda	underweight	pria	tidak
Annie	muda	average	wanita	tidak
Budi	tua	overweight	pria	tidak
Herman	tua	overweight	pria	ya
Didi	muda	underweight	pria	tidak
Rina	tua	overweight	wanita	ya
Gatot	tua	average	pria	tidak

Usia	Hipertensi	Jumlah
muda	Ya (+)	1
muda	Tidak (-)	3
tua	ya	2
tua	tidak	2

Usia = Muda

$$q_1 = -1/4 \log_2 1/4 - 3/4 \log_2 3/4 = 0,81$$

Usia = Tua

$$q_2 = -2/4 \log_2 2/4 - \log_2 2/4 = 1$$

Nilai Entropy untuk Usia :

$$E = \frac{4}{8} \log_2 \frac{4}{8} + \frac{4}{8} \log_2 \frac{4}{8} = \frac{4}{8} (0,81) + \frac{4}{8} (1) = 0,91$$

Dengan cara yang sama, entropy untuk berat dan jenis kelamin didapatkan hasil sebagai berikut :

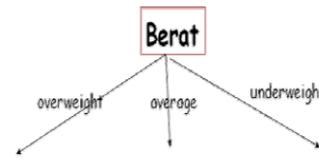
Berat	Hipertensi	Jumlah
overweight	ya	3
overweight	tidak	1
average	ya	0
average	tidak	2
underweight	tidak	2

Entropy = 0,41

Kelamin	Hipertensi	Jumlah
pria	ya	2
pria	tidak	4
wanita	ya	1
wanita	tidak	1

Entropy = 0,94

Kemudian disusun tree sebagai berikut :



Data training untuk berat-overweight, dihitung nilai entropy sebagai berikut :

Nama	Usia	Kelamin	Hipertensi
Ali	muda	pria	ya
Budiman	tua	pria	tidak
Herman	tua	pria	ya
Rina	tua	wanita	ya

Usia	Hipertensi	Jumlah
muda	ya	1
muda	tidak	0
tua	ya	2
tua	tidak	1
Entropy =		0,69

Kelamin	Hipertensi	Jumlah
pria	ya	2
pria	tidak	1
wanita	ya	1
wanita	tidak	0
Entropy =		0,69

Dari Tree yang dihasilkan maka dapat diubah ke dalam rule :

R1 : IF berat = average v berat = underweight THEN hipertensi = tidak

R2 : IF berat = overweight ^ kelamin = wanita THEN hipertensi = ya

R3 : IF berat= overweight ^ kelamin= pria ^ usia= muda THEN hipertensi = ya

R4 : IF berat= overweight ^ kelamin= pria ^ usia= tua THEN hipertensi = tidak

Hasil prediksi pada data training dievaluasi dengan mencocokkan antara tree yang didapatkan sebagai model terhadap hasil klasifikasi yang terdapat dalam data set. Dari dataset berikut, didapatkan 1 kesalahan dari 8 sampel yang ada. Sehingga nilai akurasi dapat dihitung dari 1 - prosentase kesalahan = 1 - 12.5% = 87.5%

Nama	Usia	Berat	Kelamin	Hipertensi	Prediksi
Ali	muda	overweight	pria	ya	ya
Edi	muda	underweight	pria	tidak	tidak
Annie	muda	average	wanita	tidak	tidak
Budiman	tua	overweight	pria	tidak	tidak
Herman	tua	overweight	pria	ya	tidak
Didi	muda	underweight	pria	tidak	tidak
Rina	tua	overweight	wanita	ya	ya
Gatot	tua	average	pria	tidak	tidak

5.2 Naïve Bayes

Tahapan dalam metode Naïve Bayes :

1. Pengecekan data training apakah data set adalah data yang konsisten, bias, atau tidak konsisten

Usia	Berat	Kelamin	Hipertensi
------	-------	---------	------------

muda	overweight	pria	ya
tua	overweight	pria	ya
tua	overweight	wanita	ya

Usia	Berat	Kelamin	Hipertensi
muda	underweight	pria	tidak
muda	average	wanita	tidak
tua	overweight	Pria	tidak
muda	underweight	pria	tidak
tua	average	pria	tidak

- Setelah dilakukan analisa terhadap data training, maka dapat disimpulkan bahwa data set adalah data yang tidak konsisten maupun bias sehingga kita tidak dapat membuat hipotesa langsung dengan menggunakan algoritma FIND-S.
- Metode Bayes menggunakan probabilitas bersyarat sebagai dasarnya. Probabilitas bersyarat dinyatakan sebagai :

$$P(X|Y) = \frac{P(X \cap Y)}{P(Y)}$$

$$P(\{X\} | Y) = \frac{P(\{X_1\}Y * \{X_2\}Y * \dots * \{X_n\}Y)}{P(Y)}$$

$$P(\{X\} | \sim Y) = \frac{P(\{X_1\} \sim Y * \{X_2\} \sim Y * \dots * \{X_n\} \sim Y)}{P(\sim Y)}$$

Bila nilai $P(\{X\}|Y) > P(\{X\}|\sim Y)$, maka keputusannya adalah nilai target Y

Asumsi :

Y = Hipertensi
 X1 = usia
 X2 = berat
 X3 = Kelamin

Fakta menunjukkan :

$$P(Y=ya) = 3/8 = 0.4$$

$$P(Y=tdk) = 5/8 = 0.6$$

Fakta :

$$P(X1=muda|Y=ya) = 1/3 = 0.3$$

$$P(X1=muda|Y=tdk) = 3/5 = 0.6$$

$$P(X1=tua|Y=ya) = 2/3 = 0.7$$

$$P(X1=tua|Y=tdk) = 2/5 = 0.4$$

$$P(X2=underw|Y=ya) = 0$$

$$P(X2=underw|Y=tdk) = 2/5 = 0.4$$

$$P(X2=average|Y=ya) = 0$$

$$P(X2=average|Y=tdk) = 2/5 = 0.4$$

$$P(X2=overw|Y=ya) = 3/3 = 1$$

$$P(X2=overw|Y=tdk) = 1/5 = 0.2$$

$$P(X3=pria|Y=ya) = 2/3 = 0.7$$

$$P(X3=pria|Y=tdk) = 4/5 = 0.8$$

$$P(X3=wanita|Y=ya) = 1/3 = 0.3$$

$$P(X3=wanita|Y=tdk) = 1/5 = 0.2$$

HMAP dari keadaan ini : (jika dilihat dari atribut dengan instance dari rule-uleang dihasilkan dengan decision tree sebelumnya) :

- $P(X2=average|Y=ya) = 0$
 $P(X2=average|Y=tdk) = 0.2$
 Karena $P(X2=average|Y=tdk) = 0.2 > P(X2=average|Y=ya) = 0$, maka keputusannya adalah jika berat badan average makaseseorang tidak berpotensi hipertensi
- $P(X2=underweight|Y=tdk) = 0.4$
 Karena $P(X2=underweight|Y=tdk) = 0.4 > P(X2=underweight|Y=ya) = 0$, maka keputusannya adalah jika berat badan seseorang underweight maka seseorang tidak berpotensi hipertensi
- $P(X2=overweight, X3=wanita|Y=ya)$
 $= \{P(X2=overweight|Y=ya) * P(X3=wanita|Y=ya)\} * P(Y=ya)$
 $= \{ (1) * (0.3) \} * (0.4) = 0.12$
 $P(X2=overweight, X3=wanita|Y=tdk)$
 $= \{P(X2=overweight|Y=tdk) * P(X3=wanita|Y=tdk)\} * P(Y=tdk)$
 $= \{ (0.2) * (0.2) \} * (0.6) = 0.024$
 Karena $P(X2=overweight, X3=wanita|Y=ya) > P(X2=overweight, X3=wanita|Y=tdk)$ maka keputusannya adalah jika berat badan overweight dan jenis kelamin wanita maka seseorang akan berpotensi hipertensi
- $P(X2=overweight, X3=pria, X1=muda|Y=ya)$
 $= \{ P(X2=overweight|Y=ya) * P(X3=pria|Y=ya) * P(X1=muda|Y=ya) \} * P(Y=ya)$
 $= \{ (1) * (0.7) * (0.3) \} * (0.4) = 0.084$
 $P(X2=overweight, X3=pria, X1=muda|Y=tdk)$
 $= \{ P(X2=overweight|Y=tdk) * P(X3=pria|Y=tdk) * P(X1=muda|Y=tdk) \} * P(Y=tdk)$
 $= \{ (0.2) * (0.8) * (0.6) \} * (0.6) = 0.0576$
 Karena $P(X2=overweight, X3=pria, X1=muda|Y=ya) > P(X2=overweight, X3=pria, X1=muda|Y=tdk)$ maka keputusannya adalah jika berat badan overweight dan jenis kelamin pria dan usia muda maka seseorang akan berpotensi hipertensi
- $P(X2=overweight, X3=pria, X1=tua|Y=ya)$
 $= \{ P(X2=overweight|Y=ya) * P(X3=pria|Y=ya) * P(X1=tua|Y=ya) \} * P(Y=ya)$
 $= \{ (1) * (0.7) * (0.7) \} * (0.4) = 0.196$
 $P(X2=overweight, X3=pria, X1=tua|Y=tdk)$
 $= \{ P(X2=overweight|Y=tdk) * P(X3=pria|Y=tdk) * P(X1=tua|Y=tdk) \} * P(Y=tdk)$
 $= \{ (0.2) * (0.8) * (0.4) \} * (0.6) = 0.0384$
 Karena $P(X2=overweight, X3=pria, X1=tua|Y=ya) > P(X2=overweight, X3=pria, X1=tua|Y=tdk)$ maka keputusannya adalah jika berat badan

overweight dan jenis kelamin pria dan usia tua maka seseorang akan berpotensi hipertensi

Tabel 2. Hasil Perbandingan Metode Decision Tree-ID3 dengan Naive Bayes

Rule	Keputusan (Hipertensi)	
	Decision Tree – ID3	Naïve Bayes
R1	Tidak	Tidak
R2	Tidak	Tidak
R3	ya	ya
R4	ya	Ya
R5	tidak	ya

6. Kesimpulan

1. Metode Decision Tree dan metode Bayes digunakan untuk persoalan klasifikasi dengan supervised learning dan data-data kategorikal.
2. Metode Decision Tree dan metode Bayes memerlukan pengetahuan awal untuk dapat mengambil suatu keputusan. Tingkat keberhasilan metode ini sangat tergantung pada pengetahuan awal yang diberikan.
3. Pada metode Decision Tree jumlah rule / keputusan yang dihasilkan lebih terbatas jika dibandingkan dengan jumlah rule/keputusan yang dihasilkan oleh metode Naïve Bayes.
4. Dari data training yang diolah terdapat perbedaan keputusan yang dihasilkan antar kedua metode tersebut, hal ini terlihat pada

proses pembentukan pohon / tree pada tree yang dikembangkan untuk atribut usia dengan instance tua terdapat 1 data training yang menyatakan ya dan 1 data menyatakan tidak, maka untuk menentukan pilihan memerlukan campur tangan seorang pakar, sehingga dapat disimpulkan bahwa pada proses pembentukan tree jika terdapat kondisi leaf node memiliki nilai entropy yang sama, maka membutuhkan inputan pengetahuan dari sang pakar.

5. Hasil keputusan dari metode Naïve Bayes lebih akurat dari pada hasil keputusan metode Decision Tree.

Daftar Pustaka

1. Sundaram, Aurobindo. "An Introduction to Intrusion Detection". whitepaper. 1996
2. FAY96, Fayyad, Usama. "Advances in Knowledge Discovery and Data Mining". MIT Press. 1996
3. Thearling, Kurt. "An Introduction To Data Mining". Whitepaper. <http://www3.shore.net/~kht/dmwhite/dmwhite.htm>
4. Han, Jiawei. "Data Mining Concept and Techniques". Presentation. <http://www.cs.sfu.ca/~han/dmbook>
5. Lee, Wenke. Salvatore, J. Stolvo. "A Framework for Constructing Features and Models for Intrusion Detection System". ACM Transaction Information and System Security Vol.3, No.4, November 2000.
6. Lee, Wenke. Salvatore, J. Stolvo. Mok, Kui W. "Mining Audit Data to Build Intrusion Detection Models". <http://www.aaai.org>. 1998