

# PERBANDINGAN PERFORMANSI ANTARA METODE *NAÏVE BAYES* DAN *ITERATIVE DICHOTOMIZER 3 (ID3)* DALAM KLASIFIKASI DIAGNOSIS PENYAKIT DEMAM BERDARAH DENGUE (DBD)

<sup>1</sup>Fitri Andiani <sup>2</sup>Yesy Diah Rosita <sup>3</sup>Yunita Prastyaningsih  
Program Studi Teknik Informatika Fakultas Teknik Universitas Islam Majapahit  
E-Mail : [fitriandiani1202@gmail.com](mailto:fitriandiani1202@gmail.com)

## ABSTRAK

Pernah terjadi kasus penderita DBD di desa Dadapan, karena minimnya pemahaman tentang gejala awal penyakit DBD, keluarga penderita tersebut menganggap bahwa demam yang dialami hanyalah demam biasa dan bisa disembuhkan dengan obat demam biasa. Karena kondisi semakin lemah dan demam tidak turun, akhirnya penderita di bawah ke Puskesmas Jetis. Disana penderita langsung dalam penanganan dokter dan ternyata setelah dilakukan beberapa poses tes lab penderita tersebut positif terkena DBD. Karena kondisi penderita yang sangat lemah akhirnya penderita tidak dapat menahan sakit dan akhirnya meninggal. Berdasarkan permasalahan diperlukan sebuah aplikasi pendukung yang dapat membantu dokter dalam mendiagnosis dini penyakit DBD, sehingga dapat mengurangi tingkat kematian. Dalam pembuatan aplikasi peneliti menerapkan metode yang berfungsi sebagai pengklasifikasian yaitu metode *Naïve Bayes* dan metode *Iterative Dichotomizer 3 (ID3)*. Jumlah data yang digunakan adalah sebanyak 50 data sampel, 35 data sebagai data *training* dan 15 data sebagai data *testing*, dari data *testing* ke lima belas dengan metode *Naïve Bayes* dihasilkan hasil diagnosis positif penyakit DBD dengan nilai akurasi sebesar 67%, sedangkan dengan metode *Iterative Dichotomizer 3 (ID3)* dihasilkan hasil diagnosis positif penyakit DBD dengan nilai akurasi sebesar 73%. Maka dapat disimpulkan perbandingan tingkat akurasi performansi antara metode *Naïve Bayes* dan *Iterative Dichotomizer 3 (ID3)* adalah 1,0896%.

**Kata kunci :** *Data Mining, Klasifikasi, Naïve Bayes, Iterative Dichotomizer 3 (ID3), Demam Berdarah Dengue*

## PENDAHULUAN

Penyakit Demam Berdarah Dengue (DBD) merupakan salah satu masalah kesehatan masyarakat di Indonesia yang jumlah penderitanya cenderung meningkat dan penyebarannya semakin luas (Mamik Sudarmiati, 2014). Gejala penyakit DBD hingga sekarang memang tidak terduga,

namun secara umum penyakit ini memiliki ciri-ciri seperti panas tinggi, pusing dan bahkan muntah darah. Sedangkan, gejala yang sama sering ditemukan pada penyakit lain, akibatnya sampai sekarang sering terjadi salah diagnosis. Oleh sebab itu, kita harus lebih waspada dan mengenali gejala lainya pada penyakit DBD seperti tekanan dari trombosit dan leukosit (Susanto, 2010).

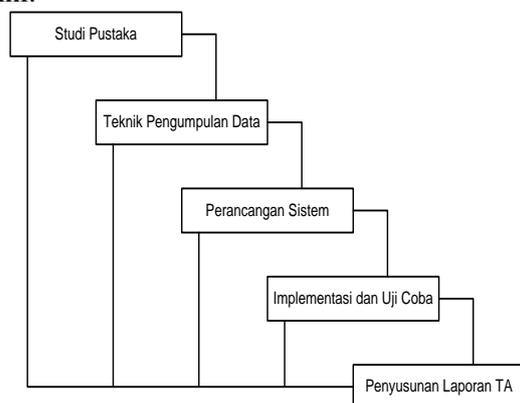
Disampaikan oleh bapak Sugeng Gianto warga Dusun Dadapan Desa Mojolebak Kecamatan Jetis lebih tepatnya RT 03 RW 01 ada salah satu keluarga yaitu putranya sendiri yang menderita penyakit DBD, karena minimnya pemahaman tentang gejala DBD, keluarga penderita tersebut menganggap bahwa demam yang diderita hanyalah demam biasa dan bisa disembuhkan dengan obat demam biasa. Karena kondisi badan semakin lemah dan demam tidak lekas turun, akhirnya penderita di bawah ke Puskesmas Jetis. Disana penderita langsung dalam penanganan dokter dan ternyata setelah dilakukan beberapa poses tes lab penderita tersebut positif terkena DBD. Karena kondisi penderita yang sangat lemah akhirnya penderita tidak dapat menahan sakit dan akhirnya meninggal.

Berdasarkan permasalahan di atas diperlukan sebuah aplikasi pendukung yang dapat membantu atau mempercepat dokter dalam mendiagnosis dini penyakit DBD, sehingga dapat mempercepat penanganan dan mengurangi tingkat kematian akibat DBD. Dalam pembuatan aplikasi tersebut peneliti menerapkan sebuah metode yang berfungsi sebagai pengklasifikasian yaitu algoritma *Naïve Bayes*, merupakan metode yang memanfaatkan nilai probabilitas. Penggunaan metode *Naïve Bayes* dalam pembuatan aplikasi ini dikarenakan hasil probabilitas nilai akurasi mendekati nilai keakuratan (Mamik Sudarmity, 2014).

Selain menggunakan metode *Naive Bayes* dalam penelitian pembuatan aplikasi juga menggunakan Metode *Iterative Dichotomizer 3* (ID3), alasan mengapa dipilih metode *Iterative Dichotomizer 3* (ID3) adalah karena memiliki kinerja yang unggul, telah mampu hampir kurang lebih 91,46% dalam mengklasifikasikan data berdasarkan kelasnya (Dr. Ahmed Tariq Sadiq, 2012). Dengan itu peneliti mengangkat judul **Perbandingan Performansi Antara Metode *Naive Bayes* Dan *Iterative Dichotomizer 3* (Id3) Dalam Klasifikasi Diagnosis Penyakit Demam Berdarah Dengue (Dbd).**

## METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah menggunakan model *waterfall* seperti pada gambar di bawah ini:



Gambar 1 Model *Waterfall*

### 1. Studi Pustaka

Penelitian yang dilakukan berdasarkan pada metode klasifikasi *Naive Bayes* dan *Iterative Dichotomizer 3* (ID3) dan dengan materi-materi tambahan sebagai bahan rujukan, diantaranya adalah beberapa jurnal yang berkaitan dengan kasus maupun metode yang peneliti lakukan dan juga buku yang menunjang penelitian. Jurnal yang digunakan merupakan jurnal yang dipublikasi dalam kurun waktu lima tahun terakhir dan terindeks secara nasional maupun internasional buku yang digunakan merupakan buku dengan terbitan terbaru.

### 2. Teknik Pengumpulan Data

Pada tahap ini peneliti membutuhkan data riil dari Puskesmas Jetis dimana data tersebut adalah berupa data pasien yang telah didiagnosis oleh dokter, yang nantinya akan dijadikan data *training* oleh peneliti dalam

pembuatan aplikasi ini, Teknik pengumpulan data yang digunakan dalam penelitian ini adalah sebagai berikut:

#### a. Observasi

Observasi atau pengamatan yang dilakukan kurang lebih dua bulan (April-Mei 2018) di Puskesmas Jetis. Informasi yang dibutuhkan dalam penelitian ini adalah bagaimana sistem dalam mendiagnosis penyakit, khususnya penyakit DBD dan juga contoh data pasien sebelumnya yang pernah didiagnosis antara negatif maupun positif DBD untuk uji coba diagnosis menggunakan sebuah aplikasi.

#### b. Wawancara

Wawancara dilakukan pada hari Selasa dan Rabu tanggal 15 -16 Mei 2018 di Puskesmas Jetis yang dilakukan dengan kepala program DBD yaitu Eka Yuli Setyawan, S.KM, wawancara berupa tanya jawab. Hasil dari wawancara terdapat pada lampiran dalam penelitian ini.

### 3. Perancangan Sistem

Penelitian yang akan dilakukan peneliti membutuhkan beberapa data dan beberapa kriteria yang digunakan dengan inputan data *testing* dari semua kriteria yang telah ditentukan peneliti. Banyak data *training* yang akan digunakan adalah 50 data dengan jumlah kriteria tujuh, dengan demikian data *testing* sesuai dengan ketujuh kriteria tersebut. Kemudian uji coba akan dilakukan dengan menggunakan metode *Naive Bayes* dan *Iterative Dichotomizer 3* (ID3).

### 4. Implementasi dan Uji Coba

Pada tahap ini meliputi implementasi program yang terdiri dari beberapa *button* dan *form* yang telah dirancang. Uji validitas dilakukan untuk menguji tingkat kehandalan *software* yang dirancang. Hal ini dilakukan untuk memastikan kehandalan dan keberhasilan sistem.

### 5. Penyusunan Laporan TA

Penyusunan laporan ini terdiri dari beberapa bab diantaranya yaitu bab 1 pendahuluan, bab 2 landasan teori, bab 3 analisa dan perancangan sistem, bab 4 pembahasan dan implementasi terakhir yaitu bab 5 penutup.

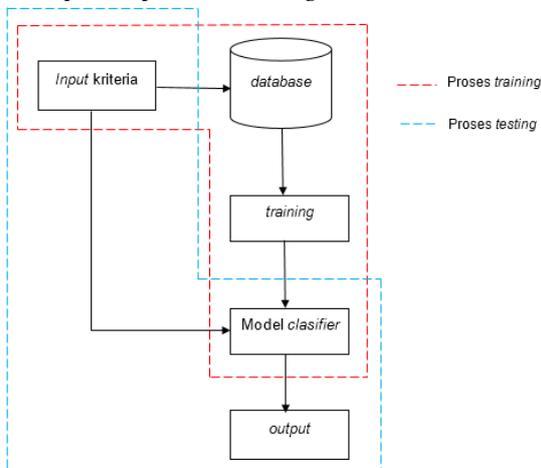
## Proses Klasifikasi

Klasifikasi merupakan proses pembelajaran suatu fungsi tujuan (target) yang memetakan tiap hubungan atribut dari label kelas yang didefinisikan sebelumnya. Fungsi target disebut juga model klasifikasi.

Ada dua jenis pemodelan klasifikasi, yaitu:

1. Pemodelan deskriptif model klasifikasi dapat bertindak sebagai alat yang bersifat menjelaskan untuk membedakan objek dari kelas yang berbeda.
2. Pemodelan prediktif model klasifikasi juga dapat digunakan untuk memprediksi label kelas dari *record* yang telah diketahui.

Pada gambar 2 di bawah dapat dijelaskan kriteria yang diinputkan akan disimpan dalam *database*, dari database akan terdapat proses *training* data untuk melihat tingkat probabilitas atau kemunculan data pada setiap kriteria terhadap atribut kelas, dari proses *training* tersebut akan menghasilkan suatu model *classifier*. Proses-proses tersebut merupakan proses *training*.



Gambar 2 Proses Klasifikasi

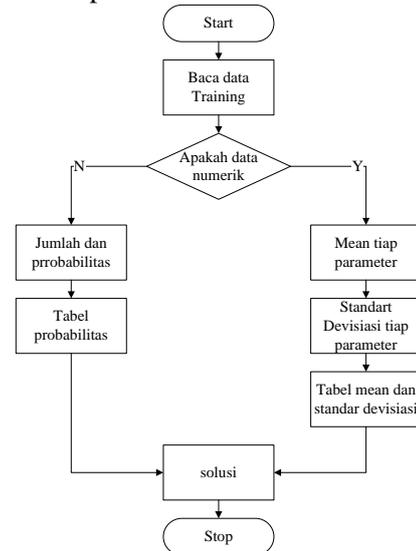
Proses *testing* dimulai ketika menginput data kriteria, dari data kriteria yang diinputkan akan disesuaikan dengan model *classifier* yang tepat, dari model *classifier* tersebut akan diperoleh *output* berupa hasil penentuan *class*. Proses *training* dan *testing* dapat dilihat pada gambar 2 di atas.

Pada gambar 3 di bawah akan dijelaskan alur *flowchart* dari metode *Naïve Bayes*, dan adalah penjelasannya,

1. Baca data *training*
2. Hitung Jumlah dan probabilitas, namun apabila data numerik maka:

- a. Cari nilai mean dan standar deviasi dari masing masing parameter yang merupakan data numerik.
- b. Cari nilai probabilitistik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.

Mendapatkan nilai dalam tabel mean, standart deviasi dan probabilitas.



Gambar 3 Flowchart Algoritma *Naïve Bayes* (Bustami, 2014)

## Iterative Dichotomizer 3 (ID3)

*Iterative Dichotomizer 3* (ID3) adalah algoritma *decision tree learning* (algoritma pembelajaran pohon keputusan) yang paling dasar. Algoritma ini melakukan pencarian secara rakus /menyeluruh (*greedy*) pada semua kemungkinan pohon keputusan. Salah satu algoritma induksi pohon keputusan yaitu *Iterative Dichotomizer 3* yang dikembangkan oleh *J. Ross Quinlan* (Wahyudin, 2013).

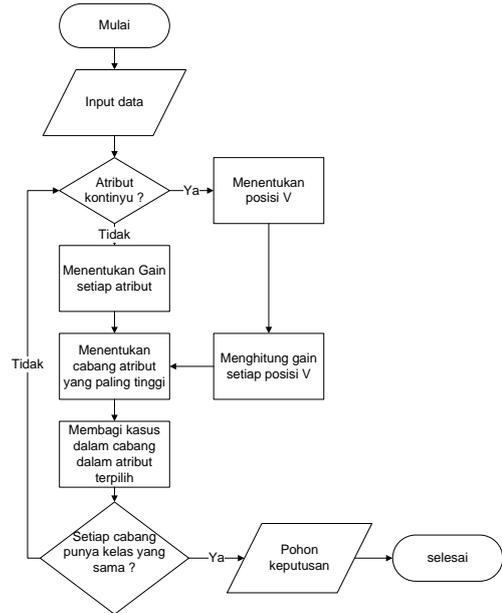
Algoritma ID3 dapat diimplementasikan menggunakan fungsi *rekursif* (fungsi yang memanggil dirinya sendiri). Algoritma ID3 berusaha membangun *decision tree* (pohon keputusan) secara *top-down* (dari atas ke bawah), mulai dengan pertanyaan : “atribut mana yang pertama kali harus dicek dan diletakkan pada *root*?” pertanyaan ini dijawab dengan mengevaluasi semua atribut yang ada dengan menggunakan suatu ukuran statistik (yang banyak digunakan adalah *information gain*) untuk mengukur efektivitas suatu atribut dalam mengklasifikasikan kumpulan sampel data (Efendi Subhan, 2018).

*Decision Tree* adalah sebuah struktur pohon, dimana setiap node pohon

merepresentasikan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan node daun (*leaf*) merepresentasikan kelompok kelas tertentu. Level node teratas dari sebuah *decision tree* adalah node akar (*root*) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu (Wahyudin, 2013). Kemampuannya untuk mem-*break down* proses pengambilan keputusan yang kompleks menjadi lebih simpel sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan adalah manfaat utama dari penggunaan pohon keputusan.

Dari *flowchart* di bawah ini dapat diketahui proses alur dari sistem metode ID3,

1. Input data
2. Terdapat kondisi yang menyatakan inputan tersebut termasuk atribut kontinyu atau tidak.
  - a. Jika iya maka proses menentukan posisi V.
  - b. Setelah itu mengitung nilai *gain* dari setiap posisi V.
  - c. Menentukan cabang atribut yang paling tinggi.
  - d. Membagi kasus dalam cabang dalam atribut terpilih.
  - e. Kondisi setiap cabang mempunyai yang sama atau tidak, jika iya pohon keputusan dan selesai, sedangkan jika maka kembali lagi pada kondisi pada proses kedua.
3. Jika data tidak termasuk atribut kontinyu maka
  - a. Menentukan nilai *gain* setiap atribut.
  - b. Menentukan cabang atribut yang paling tinggi.
  - c. Membagi kasus dalam cabang dalam atribut terpilih.
  - d. Kondisi setiap cabang mempunyai yang sama atau tidak, jika iya pohon keputusan dan selesai, sedangkan jika maka kembali lagi pada kondisi pada proses kedua.



Gambar 4 Flowchart Algoritma ID3

## HASIL DAN PEMBAHASAN

Dalam sub bab ini akan menjelaskan langkah-langkah penggunaan aplikasi diagnosis penyakit DBD Puskesmas jetis ini. Aplikasi diagnosis ini hanya untuk petugas atau dokter sebagai pengoperasi aplikasi ini. Pada aplikasi ini dapat menghasilkan atau *output* yaitu hasil diagnosis pasien dan rekam pasien.

### a. Form Login

Gambar 5 Form Login

Langkah awal untuk memasuki aplikasi ini, setiap dokter memiliki username dan password masing-masing, username adalah nama dokter dan password adalah NID (Nomor Induk Dokter), jika username dan *password* yang dimasukkan benar.

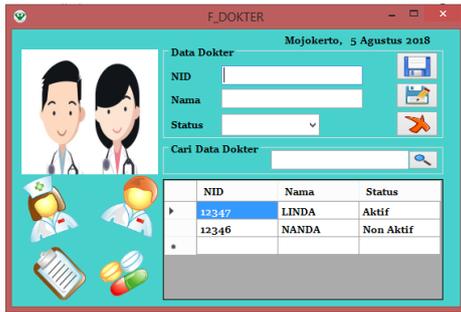
### b. Form Home

Merupakan tampilan awal (*home*) dari sistem. Tampilan ini akan muncul apabila pengguna berhasil masuk ke dalam sistem dengan data login yang benar. Tampilan form login dapat dilihat pada gambar 6.



Gambar 6 Form Home

c. Form Dokter



Gambar 7 Form Dokter

Pada form ini admin atau dokter dapat melakukan *input* data dokter baru, hapus data dokter yang sudah tidak aktif, mencari data dokter dan melakukan edit data.

d. Form Pasien

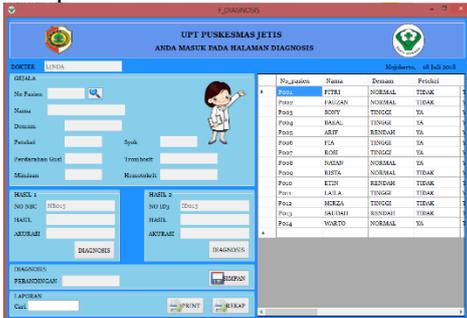
Adalah halaman yang digunakan untuk *input* data pasien dan gejala yang dirasakan oleh pasien, data-data ini nantinya akan sebagai data uji atau data *testing* yang akan didiagnosis oleh dokter menggunakan sistem aplikasi ini.



Gambar 8 Form Pasien

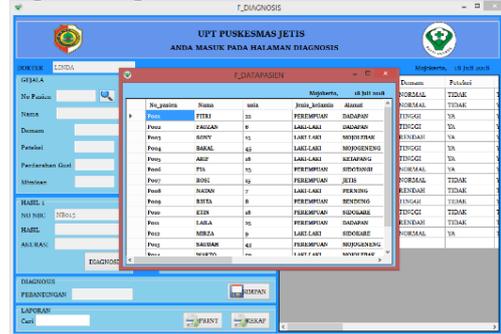
e. Form Diagnosis

berikut adalah proses diagnosis. Masuk pada halaman awal.



Gambar 9 Halaman Awal Form Diagnosis

Selanjutnya yaitu pilih data pasien dengan mengambil data melalui *button* cari



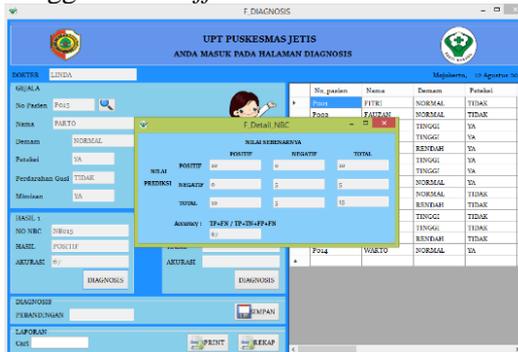
Gambar 10 Halaman Data Gejala Pasien

Selanjutnya proses diagnosis kedua metode dengan menekan *button* diagnosis.



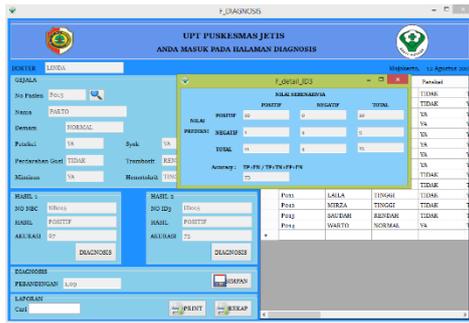
Gambar 11 Halaman Diagnosis

Otomatis nilai perbandingan dari kedua metode terhitung, pada saat proses diagnosis pada gambar 12 adalah form untuk menampilkan detail perhitungan nilai keakuratan dari metode NBC, dimana menggunakan *confusion matrix*.



Gambar 12 Form Detail Perhitungan NBC

Pada gambar 13 adalah form detail perhitungan keakuratan dari metode *Iterative Dichotomizer 3 (ID3)*, sama dengan metode *Naive Bayes* perhitungan ini juga menggunakan *confusion matrix*.



Gambar 13 Form Detail Perhitungan NBC

### Pembahasan Metode Naive Bayes

Perhitungan manual metode *Naive Bayes* akan dicontohkan untuk satu macam data testing. Data testing yang akan digunakan adalah demam  $\leq 37,7^{\circ}\text{C}$ , petepei = ya, perdarahan gusi = tidak, mimisan = ya, syok = ya, trombosit  $< 100.000/\mu\text{l}$  dan hemotokrit  $> 48\%$ . Langkah-langkah untuk menghitung klasifikasi menggunakan metode *Naive Bayes* adalah sebagai berikut:

- Menghitung probabilitas hipotesis (prior) atau jumlah kelas  
Kelas diagnosis adalah positif dan negatif

$$P(\text{Status} = \text{"positif"}) = \frac{28}{50} = 0,56$$

$$P(\text{Status} = \text{"negatif"}) = \frac{22}{50} = 0,44$$

- Menghitung probabilitas berdasarkan kondisi pada hipotesis (probabilitas kelas per kasus)

- Kriteria demam variabel  $\leq 37,7^{\circ}\text{C}$  dengan diagnosis positif

$$P(\text{demam} = \leq 37,7^{\circ}\text{C} | \text{status} = \text{positif}) = \frac{7}{28} = 0,25$$

- Kriteria demam variabel  $\leq 37,7^{\circ}\text{C}$  dengan diagnosis negatif

$$P(\text{demam} = \leq 37,7^{\circ}\text{C} | \text{status} = \text{negatif}) = \frac{8}{22} = 0,36$$

- Kriteria trombosit variabel  $< 100.000/\mu\text{l}$  dengan diagnosis positif

$$P(\text{demam} = < 100.000/\mu\text{l} | \text{status} = \text{"positif"}) = \frac{2}{28} = 0,07$$

- Kriteria trombosit variabel  $< 100.000/\mu\text{l}$  dengan diagnosis negatif

$$P(\text{demam} = < 100.000/\mu\text{l} | \text{status} = \text{"positif"}) = \frac{9}{22} = 0,41$$

Perhitungan untuk probabilitas berdasarkan kondisi untuk kriteria yang lain sama seperti pada poin 2.a dan 2.b.

- Mengalikan semua variabel kelas  
Untuk kelas diagnosis adalah positif

$$P(x | \text{diagnosis} = \text{"positif"}) = P(\text{demam} = \text{"normal"}) \times P(\text{trombosit} = \text{"rendah"})$$

$$= 0,25 \times 0,07$$

$$= P(\text{petepei} = \text{"ya"}) \times P(\text{perdarahan gusi} = \text{"tidak"})$$

$$= 0,79 \times 0,43$$

$$= P(\text{mimisan} = \text{"ya"}) \times P(\text{syok} = \text{"ya"})$$

$$= 0,68 \times 0,50$$

$$= P(\text{hemotokrit} = \text{"tinggi"}) = 0,61$$

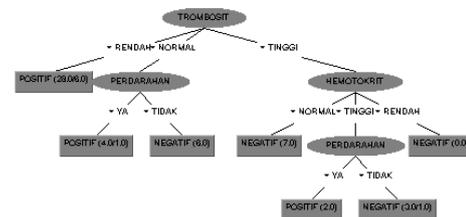
$$= 0,001239$$

Membandingkan hasil per kelas

Dari perhitungan di atas diperoleh hasil bahwa kelas positif DBD memiliki nilai 0,001239 dan kelas negatif DBD memiliki nilai 0,001114. Diketahui bahwa nilai dari kelas positif lebih besar dari pada kelas negatif. Sehingga untuk data testing yang ke lima belas untuk kriteria yang telah disebutkan diklasifikasikan ke dalam diagnosis positif DBD.

### Pembahasan Metode Iterative Dichotomizer 3 (ID3)

Di bawah ini akan di jelaskan hasil dari pohon keputusan yang dihasilkan dengan bantuan menggunakan *tools* Weka versi 3.8, dengan terdapat empat kategori tersusun.



Gambar 14 Pohon Keputusan

Dari gambar 14 peneliti mendapatkan model klasifikasi yang akan diterapkan pada sistem aplikasi diagnosis ini,

- If ((trombosit = 'rendah') then diagnosis = 'positif')
- If ((trombosit = 'normal' and (perdarahan gusi = 'ya')) then diagnosis = 'positif')
- If ((trombosit = 'normal' and (perdarahan gusi = 'tidak')) then diagnosis = 'negatif')
- If ((trombosit = 'tinggi' and (hemotokrit = 'normal')) then diagnosis = 'negatif')
- If ((trombosit = 'tinggi' and (hemotokrit = 'tinggi' and (perdarahan gusi = 'ya')) then diagnosis = 'positif')
- If ((trombosit = 'tinggi' and (hemotokrit = 'tinggi' and (perdarahan gusi = 'tidak')) then diagnosis = 'negatif')

7. If ((trombosit='tinggi' and (hemotokrit = 'rendah')) then diagnosis='negatif')

Dari hasil pohon keputusan yang terbuat untuk data testing yang akan digunakan adalah demam  $\leq 37,7^{\circ}\text{C}$ , peteeki = ya, perdarahan gusi = tidak, mimisan= ya, syok = ya, trombosit  $< 100.000/\mu\text{l}$  dan hemotokrit  $> 48\%$  adalah mendapatkan hasil klass positif.

#### Perhitungan Nilai Akurasi ke dua Metode

##### 1. Metode Naive Bayes

$$accuracy = \frac{TP+FN}{TP+TN+FP+FN} \quad (1)$$

Keterangan:

TP: True Positive

FP: False Positive

FN: False Negative

TN: True Negative

$$accuracy = \frac{TP + FN}{TP + TN + FP + FN} = \frac{10 + 0}{10 + 5 + 0 + 0} = 0,67\%$$

Sedangkan nilai akurasi metode *Iterative Dichotomizer 3 (ID3)* adalah

$$accuracy = \frac{TP+FN}{TP+TN+FP+FN} = \frac{10+1}{10+4+0+1} = 0,73 \%$$

#### KESIMPULAN

1. Terdapat tujuh parameter *input* yang sebagai data *testing* atau data uji, diantaranya yaitu demam, peteeki (bintik merah), perdarahan gusi, mimisan, syok, trombosit dan hemotokrit.
2. Klass *output* yang dihasilkan adalah dua tipe klass, yaitu klass positif dan negatif.
3. Dari data *testing* ke lima belas dengan metode *Naive Bayes* dihasilkan hasil diagnosis positif penyakit DBD dengan nilai akurasi sebesar 67%, sedangkan dengan metode *Iterative Dichotomizer 3 (ID3)* dihasilkan hasil diagnosis positif penyakit DBD dengan nilai akurasi sebesar 73%. Maka dapat disimpulkan perbandingan tingkat akurasi performansi antara metode *Naive Bayes* dan *Iterative Dichotomizer 3 (ID3)* adalah 1,0896%.

#### DAFTAR PUSTAKA

- Basuki. (2013). *Metode Bayes*. PENS-ITS
- Bustami. (2015). *Penerapan Algoritma Naive Bayes Untuk Mengklasifikasi Data Nasabah Asuransi*.

- Universitas Malikussaleh
- Efend. S. (dkk). (2012). *Prediksi Penyakit Diabetes Menggunakan Algoritma ID3 dengan Pemilihan Atribut Terbaik*. Universitas Diponegoro
- El-Bialy, Randa. (dkk). (2015). *Feature Analysis of Coronary Artery Heart Disease Data Sets*. ScienceDirect
- Fajar Astuti. (2013). *Data Mining*. Yogyakarta:ANDI
- Mamik Sudarmiati. (dkk). (2014). *Diagnosa Demam Berdarah Menggunakan Metode Naive Bayes*. Universitas Muhammadiyah Jember
- Susanto, Agus. (2010). *Waspada! Gigitan Nyamuk*. Jakarta:Sunda Kepala Pustaka
- Wahyudin. (2013). *Metode Iterative Dichotomizer 3 (ID3) Untuk Penerimaan Mahasiswa Baru*. Universitas Pendidikan Indonesia
- WHO. (2011). *Haemoglobin Concentration for the Diagnosis of Anaemia and Assessment of Severity*. World Health Organizat

