

# Penerapan Data Mining Menggunakan Metode Naïve Bayes Untuk Klasifikasi Tindakan Jenis Abortus Di Rsud Duta Mulya

Indri Dwi Lestari<sup>a,1</sup>, Tedy Setiadi<sup>a,2</sup>, Lisna Zahrotun<sup>a,3</sup>

<sup>a</sup>Program Studi Teknik Informatika, Universitas Ahmad Dahlan,  
Prof. Dr. Soepomo, S.H., Janturan, Umbulharjo, Yogyakarta 55164

<sup>1</sup>Indridwilestari777@gmail.com; <sup>2</sup>tedy.setiadi@tif.uad.ac.id; <sup>3</sup>lisna.zahrotun@tif.uad.ac.id

## Abstrak

Masalah kesehatan merupakan masalah penting yang tengah dihadapi masyarakat saat ini, apalagi yang tengah menimpa kaum wanita. Mortalitas dan morbiditas pada wanita merupakan masalah yang perlu mendapatkan perhatian yang serius. Di Negara berkembang seperti Indonesia, angka kematian ibu (AKI) adalah jumlah wanita yang meninggal dari suatu penyebab kematian terkait dengan gangguan kehamilan atau penanganannya. Abortus adalah ancaman atau pengeluaran hasil konsepsi sebelum janin sebelum janin dapat hidup diluar kandungan. Banyak faktor resiko yang berkontribusi terhadap kejadian abortus, yaitu usia ibu, usia kehamilan, HB/anemia, jarak kehamilan, riwayat penyakit, riwayat abortus dan gejala. Salah satu rumah sakit yang memberikan perhatian khusus pada abortus adalah RSUD Duta Mulya, kejadian abortus tertinggi terjadi pada tahun 2014. Oleh karena itu, penelitian ini berusaha membantu sebuah rumah sakit untuk merancang sebuah aplikasi data mining yang berfungsi untuk mengklasifikasikan tindakan dari jenis abortus.

Dalam penelitian ini metode yang digunakan dalam melakukan proses klasifikasi adalah algoritma naïve bayes. Sedangkan untuk pengujian akurasi hasil klasifikasi digunakan metode confusion matrix.

Hasil penelitian ini adalah sebuah aplikasi yang dapat mengklasifikasikan tindakan jenis abortus dengan menggunakan metode naïve bayes. Sedangkan berdasarkan pengujian dari confusion matrix dihasilkan akurasi sebesar 99,26%. Hasil tersebut didapatkan dari 14 data training dan 121 data testing. Dan akurasi sebesar 97,78% dengan 108 data training dan 28 data testing. Dari kedua percobaan yang dilakukan dalam penelitian ini, naïve bayes lebih bagus digunakan dengan menggunakan data training dengan jumlah sedikit.

Kata Kunci: Klasifikasi, Algoritma *Naïve Bayes*, *Confusion Matrix*, *Abortus*.

## 1. Pendahuluan

Masalah kesehatan merupakan masalah penting yang tengah dihadapi oleh masyarakat saat ini, apalagi yang tengah menimpa kaum wanita. Mortalitas dan morbiditas pada wanita merupakan masalah yang perlu mendapatkan perhatian yang serius. Di Negara berkembang seperti Indonesia, angka kematian ibu (AKI) adalah jumlah wanita yang meninggal dari suatu penyebab kematian terkait dengan gangguan kehamilan atau penanganannya. Berjuta-juta wanita setiap tahunnya mengalami kehamilan yang tidak diinginkan. Beberapa kehamilan berakhir dengan kelahiran tetapi beberapa diantaranya diakhiri dengan abortus [1]. Faktor resiko yang memengaruhi terjadinya abortus adalah usia ibu, usia kehamilan, HB/anemia, jarak kehamilan, riwayat penyakit, riwayat abortus dan gejala (Sastrawinata, dkk, 2005). Data kejadian abortus di RSUD Duta Mulya tertinggi terjadi pada tahun 2014. Prosentase tertinggi mencapai 3,9% yaitu sebanyak 347 kasus dibanding tahun-tahun sebelumnya.

Klasifikasi merupakan peran utama dalam teknik dalam data mining. Klasifikasi termasuk kedalam supervised learning karena dalam proses klasifikasi terdapat proses pembelajaran lampau. Proses ini digunakan algoritma untuk menganali pola dari data yang nantinya akan diterapkan kepada baru yang belum diketahui kelompoknya (Bala, 2012). Salah satu algoritma yang umum dipakai

dalam klasifikasi adalah algoritma naïve bayes. Algoritma naïve bayes dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Hal tersebut dapat diterapkan pada masalah klasifikasi seperti peramalan cuaca, deteksi gangguan, diagnosis penyakit dan lain sebagainya. Naïve bayes dipilih karena memiliki tingkat ketelitian dan kecepatan yang tinggi saat diaplikasikan untuk jumlah data yang besar [3] (Kabir, et , all, 2011).

Dengan melihat permasalahan dari data pasien mengalami abortus dan kelebihan dari metode naïve bayes, maka akan dilakukan pengklasifikasian tindakan jenis abortus dengan metode naïve bayes.

## 2. Kajian Pustaka

### 2.1. Kajian Terdahulu

Kajian terdahulu yang digunakan sebagai refrensi penelitian mengacu pada penelitian yang sudah cukup banyak membahas data mining dengan metode klasifikasi. Penelitian yang dilakukan oleh Almira Syawli, dkk yang berjudul “ Diagnosa Penyakit Diabetes Mellitus Dengan Metode Naïve Bayes Berbasis Desktop Application” penelitian ini menghasilkan pengujian kinerja system penentuan terkena diabetes mellitus dilihat dari probabilitas kemunculan dari setiap gejala indikator sebagai kunci dari diabetes mellitus [2].

Penelitian lainnya yang dilakukan oleh Royani Chairiyah dengan judul “ Faktor –Faktor Yang Mempengaruhi Ibu Hamil Dengan Abortus Incomplit Di RSUD Kota Bekasi ” penelitian ini bertujuan untuk distribusi frekuensi faktor- faktor yang mempengaruhi ibu hamil dengan abortus inkomplit berdasarkan blighted ovum, umur, riwayat abortus dan riwayat penyakit [4].

### 2.2. Landasan Teori

#### a. Data Mining

Data Mining sering disebut Knowledge Discovery in Database (KDD), adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keterkaitan, pola atau hubungan dalam set data berukuran besar [5]. (Budi Santosa, 2007).

#### b. Klasifikasi (classification)

Klasifikasi termasuk ke dalam supervised learning karena dalam proses klasifikasi terdapat proses pembelajaran lampau. Proses ini digunakan algoritma untuk mengenali pola dari data yang nantinya akan diterapkan kepada data baru yang belum diketahui kelompoknya (Bala,2012).

#### c. Algoritma Naïve bayes

Algoritma naïve bayes adalah metode yang paling populer yang digunakan untuk mengklasifikasikan data dalam jumlah yang besar dan dapat digunakan untuk memprediksi keanggotaan suatu kelas. Naïve bayes dipilih karena memiliki tingkat ketelitian dan kecepatan yang tinggi saat diaplikasikan untuk jumlah data yang besar [3] (Kabir, et, all, 2011). Klasifikasi naive bayes mempunyai persamaan yang ditunjukkan dalam persamaan 1 (Han, et, al, 2012)

$$P(C_i | X) = \frac{P(X|C_i) \cdot P(C_i)}{P(X)} \dots\dots\dots(1)$$

#### d. Confusion Matrix

Confusion Matrix adalah matrik yang berbentuk 2x2 untuk mendapatkan hasil klasifikasi dengan tepat terdapat pada weka classifier. Untuk mendapatkan hasil klasifikasi dengan tepat tiap kelas yang diprediksi mempunyai kemungkinan keluaran yang berbeda-beda seperti diantaranya true positive yaitu menunjukkan ketepatan klasifikasi dan false positive adalah jika nilai aslinya dan nilai yang diprediksi berbeda [6]. Tabel Confusion Matrix ditunjukkan dalam Tabel 3.

Tabel 3 Confusion matrix

		Predicted Class	
		C1	C2
Actual Class	C1	True Positive	False negative
	C2	False positive	True negative

- e. Faktor-faktor yang mempengaruhi kejadian abortus menurut beberapa ahli antara lain :

1) Umur

Umur adalah lama waktu hidup (DepDikNas, 2001). Umur ibu adalah waktu hidup ibu bersalin sejak lahir sampai melahirkan. Saat terbaik untuk seorang wanita hamil adalah 20-35 tahun, karena pada usia tersebut seorang wanita sudah mengalami kematangan fungsi organ-organ reproduksi dan secara psikologis sudah dewasa [7]. Ibu hamil dengan usia kurang dari 20 tahun organ reproduksinya belum matang dan belum berfungsi secara optimal untuk hamil sehingga dapat merugikan kesehatan ibu maupun pertumbuhan janin. Sedangkan pada usia lebih dari 35 tahun dimana organ-organ tubuh sudah mengalami penurunan fungsi.

2) Usia kehamilan

Kehamilan dengan usia 0-12 minggu masih sangat rentan mengalami abortus, karena pada awal kehamilan sebelum 3 bulan, Keadaan ini disebabkan karena pada masa tersebut rentan terjadi kelainan pertumbuhan janin atau malformasi (prawirohardjo, 2008).

3) HB/anemia

Wanita hamil atau dalam nifas dinyatakan menderita anemia apabila kadar hemoglobin < 11gr%. Penurunan kadar hemoglobin pada wanita sehat yang hamil disebabkan ekspansi volume plasma yang lebih besar daripada volume sel darah merah dan hemoglobin terutama terjadi pada trimester II [8].

4) Jarak dua kehamilan yang dekat

Bila jarak dengan anak sebelumnya  $\leq 2$  tahun, rahim dan kesehatan ibu belum pulih dengan baik. Kehamilan dalam keadaan ini perlu diwaspadai karena ada kemungkinan pertumbuhan janin kurang baik, pendarahan kemudian abortus. Indikasi abortus meningkat pada wanita yang hamil dalam 3 bulan setelah melahirkan aterm (Kenneth J. Leveno et al, 2009).

5) Riwayat Penyakit

Riwayat penyakit yang diderita ibu seperti hipertensi, diabetes, tbc serta pneumonia, tifus abdominalis, malaria, dan lainnya dapat menyebabkan abortus. Toksin, bakteri, virus atau plasmodium dapat melalui plasenta masuk ke janin, sehingga menyebabkan kematian janin kemudian terjadi abortus [9].

6) Riwayat Abortus

Data dari beberapa studi menunjukkan bahwa setelah 1 kali abortus pasangan mempunyai resiko 15% untuk mengalami keguguran lagi, sedangkan bila pernah 2 kali, risikonya akan meningkat 25%. Setelah 3 kali abortus meningkat 30-45%. Penderita dengan riwayat abortus akan menunjukkan adanya pertumbuhan janin yang terhambat pada kehamilan berikutnya melahirkan bayi premature (Suryadi, 2009).

7) Gejala

Komplikasi yang berbahaya pada abortus ialah pendarahan. Kejadian abortus yang terjadi dapat menimbulkan komplikasi dan dapat menyebabkan kematian. Komplikasi abortus yang dapat menyebabkan kematian ibu antara lain karena pendarahan dan infeksi. Pendarahan yang terjadi selama abortus dapat mengakibatkan pasien menderita anemia, sehingga dapat meningkatkan resiko kematian ibu (Donny, 2015).

### 3. Metodologi

Metodologi dalam penelitian ini terdiri dari:

Dalam penelitian ini yang menjadi subyek penelitian adalah meramalkan tindakan yang diambil pada pasien yang mengalami keguguran atau abortus menggunakan metode naive bayes. Sumber data penelitian diambil dari bagian administrasi pada poli kandungan RSUD DUTA MULYA, yaitu sebanyak 135 data rekam medis berupa Nomor registrasi, nama pasien, alamat, usia Ibu bersalin, usia kelahiran, HB, jarak kehamilan, riwayat penyakit, riwayat abortus, gejala, jenis abortus dan tindakan.

### 3.1. Metode Pengumpulan Data

Metode pengumpulan data yang dilakukan pada penelitian ini adalah sebagai berikut:

#### A. Studi Pustaka

Studi pustaka ini adalah mempelajari teori pendukung data mining algoritma naïve bayes untuk penelitian seperti buku, literatur, e-book, dan penelitian lain yang berkaitan untuk membandingkan hasil penelitian.

#### B. Wawancara

Wawancara adalah menyusun daftar pertanyaan yang diajukan oleh dokter Tatang sebagai dokter kandungan di RSUD Duta Mulya untuk penelitian tentang pasien mengalami keguguran (abortus). Hasil dari wawancara tersebut dijadikan sebagai salah satu metode pengumpulan yang digunakan untuk melakukan penelitian.

### 3.2. Analisis Kebutuhan

#### A. Perangkat Keras (Hardware)

Perangkat keras yang digunakan dalam pengembangan klasifikasi menggunakan algoritma naïve bayes ini adalah: Prosesor Intel(R) Core(TM) i3-2330M @2.2 Ghz, Memori 2,00 GB, Mouse.

#### B. Perangkat Lunak (Software)

Perangkat lunak yang digunakan dalam klasifikasi pasien mengalami abortus dengan naïve bayes adalah: Sistem operasi Windows 7, VB 06, Database SQL, Excel.

#### C. Kebutuhan Input

Analisis kebutuhan Input aplikasi mampu melakukan pengambilan data rekamedis yang merupakan file data excel.

#### D. Kebutuhan Proses

Dari analisis kebutuhan input maka ditentukan algoritma yang akan digunakan dalam proses klasifikasi pasien mengalami abortus dengan menggunakan algoritma naïve bayes yang dimulai dari proses mengklasifikasikan masing-masing data atribut yang bernilai diskrit

#### E. Kebutuhan Output

Output dari penerapan data mining ini adalah hasil prediksi dengan data atribut yang paling berpengaruh pada faktor yang mempengaruhi tindakan dari jenis abortus kedalam klasifikasi.

### 3.3. Tahap-tahap Data Mining

Tahap-tahap data mining [10].

#### 1. Pembersihan Data (cleaning)

Pada tahap ini dilakukan pembersihan data dengan menghapus data yang tidak lengkap dari data rekamedis, namun pada penelitian ini tidak ada data yang dibersihkan karena data sudah lengkap.

#### 2. Integrasi Data (integration)

Pada penelitian ini, data murni dari excel, sehingga tidak perlu dilakukan proses integrasi data.

#### 3. Seleksi Data (selection)

Pada tahap ini dilakukan pemilihan atribut yang akan digunakan pada proses mining. Dari 12 atribut yang ada, akan dipilih 9 atribut yang digunakan yaitu usia ibu, usia kehamilan, hb/anemia, jarak kehamilan, riwayat penyakit, riwayat abortus, dan gejala.

#### 4. Transformasi Data (transformation)

Pada tahap ini data diubah dari numerik akan dikategorikan sesuai dengan format yang sesuai.

#### 5. Proses mining

Pada proses mining dilakukan pengklasifikasian dengan menggunakan algoritma naïve bayes untuk mengklasifikasikan faktor yang mempengaruhi terjadinya abortus.

#### 6. Evaluasi Pola

Pola dapat diketahui setelah tahapan-tahapan data mining. Hal ini tentunya dilakukan dilakukan setelah proses mining diterapkan pada aplikasi data mining yang telah dibangun guna menemukan pola menarik untuk direpresentasikan ke dalam knowledge based.

### 3.4. Pengujian Validasi

#### Pembagian Data Training dan Testing

Dari 135 data rekamedis. Jumlah data untuk training dan testing masing-masing adalah untuk mengetahui keberan dari data, digunakan confusion matrix yang dibagi dalam data training dan testing.

Percobaan 1:

$$N_{\text{training}} = 1/5 \times 135 = 27 \text{ data}$$

$$N_{\text{testing}} = 4/5 \times 135 = 108 \text{ data}$$

Percobaan 2

$$N_{\text{training}} = 1/10 \times 135 = 14 \text{ data}$$

$$N_{\text{testing}} = 9/10 \times 135 = 121 \text{ data}$$

## 4. Hasil Dan Pembahasan

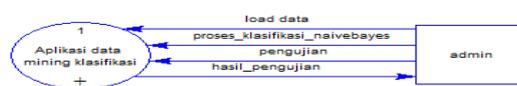
### 4.1. Analisis Kebutuhan Aplikasi

Analisis kebutuhan aplikasi bertujuan untuk mengetahui kebutuhan pengguna aplikasi yang akan dibangun. Di dalam analisis kebutuhan, aplikasi yang diharapkan dapat memecahkan masalah yang memiliki kemampuan sebagai berikut:

- Aplikasi mampu melakukan pengambilan data rekamedis yang merupakan file data excel.
- Aplikasi mampu melakukan proses pembersihan data, dimana data yang tidak lengkap akan dibersihkan.
- Aplikasi mampu melakukan proses seleksi data, dimana data yang diseleksi adalah data yang digunakan untuk proses mining seperti usia ibu, usia kehamilan, hb, jarak kehamilan, riwayat penyakit, riwayat abortus, gejala, jenis abortus dan tindakan.
- Aplikasi mampu melakukan proses transformasi data, data diubah kedalam format yang sesuai atau diubah menjadi kategori tertentu untuk mempermudah dalam proses mining.

### 4.2. Perancangan Sistem

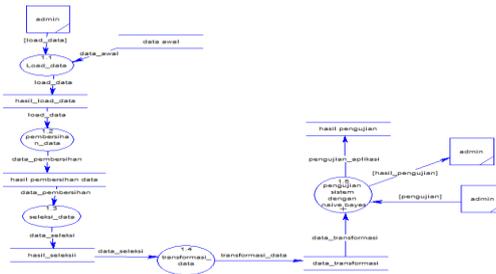
#### a) DFD Level 0



Gambar 1. DFD Level-0 Aplikasi Data Mining Klasifikasi

Pada gambar 1 merupakan DFD level-0 yang menjelaskan proses data mining klasifikasi dengan input load data berupa excel, proses klasifikasi naïve bayes dan pengujian aplikasi.

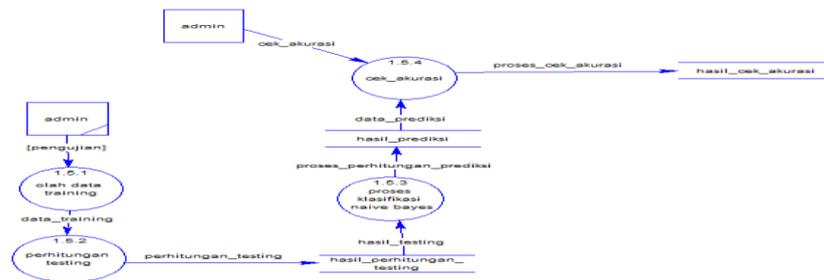
#### b) DFD Level 1



Gambar 2. DFD Level-0 Aplikasi Data Mining Klasifikasi

DFD level-1 melakukan proses pembersihan data, admin mengambil data rekamedis lalu dilakukan proses pembersihan data. Data dari proses pembersihan data tersebut kemudian dilakukan proses seleksi, yaitu memilih atribut mana saja yang akan digunakan dalam proses mining klasifikasi. Dari hasil seleksi lalu dilakukan proses tranformasi data, lalu dilakukan pengujian sistem. kemudian admin akan memperoleh hasil dari pengujian tersebut.

c) DFD Level 2



Gambar 3. DFD Level-2 Proses Pengujian Aplikasi

DFD level-2 admin melakukan proses pengujian data training dan testing dari data untuk mendapatkan sebuah tingkat akurasi dalam klasifikasi.

1. Implementasi

Implementasi program adalah proses dari perancangan antar muka atau user interface kemudian dilanjutkan pada sebuah aplikasi program, bahasa pemrograman yang digunakan adalah Visual Basic. Implementasi program pada perancangan yang pertaman adalah sebagai berikut:

a. Form Menu Utama

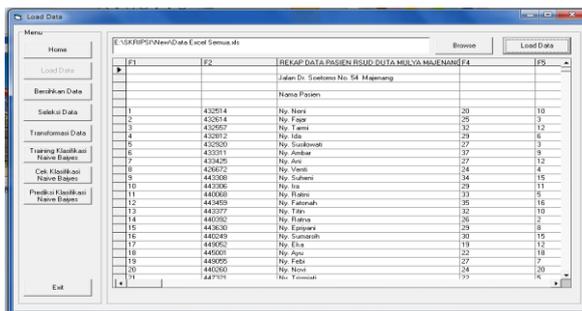
Form menu utama terdapat beberapa menu yaitu load data,bersihkan data, seleksi data, transformasi data, training klasifikasi, cek klasifikasi, dan prediksi. Ditunjukkan pada gambar 4.



Gambar 4 Form Menu Utama

b. Form load data

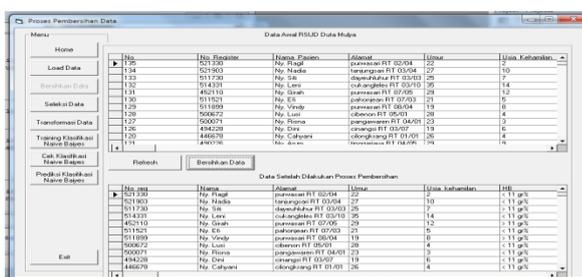
Load data akan mengambil data dalam bentuk file excel sesuai dengan kebutuhan aplikasi.



Gambar 5 Form Load Data

c. Form Bersihkan Data

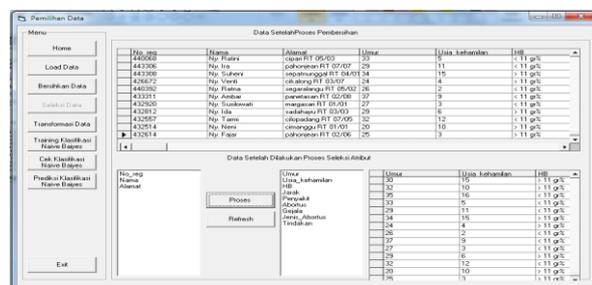
Bersihkan data digunakan untuk menghilangkan data yang tidak relevan atau tidak lengkap untuk proses mining. Ditunjukkan pada gambar 6.



Gambar 6 Form Bersihkan Data

d. Seleksi Data

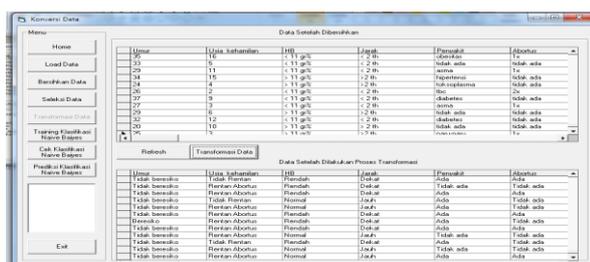
Form seleksi digunakan untuk menyeleksi data atribut yang akan digunakan dalam proses mining. Ditunjukkan pada gambar 7



Gambar 7 Form Seleksi Data

e. Transformasi Data

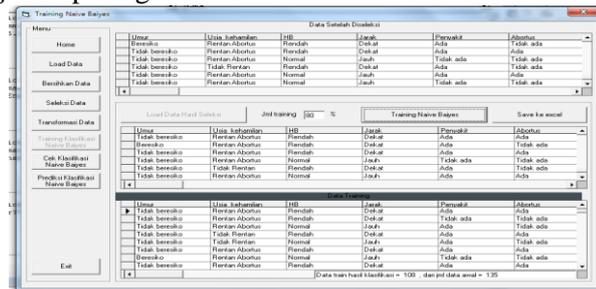
Form transformasi data digunakan untuk mengkategorikan data kedalam format yang sesuai. Ditunjukkan pada gambar 8.



Gambar 8. Form Transformasi Data

f. Training Klasifikasi

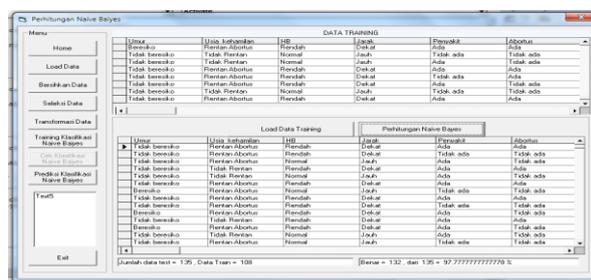
Proses training klasifikasi merupakan proses pembagian data training dan testing dari data untuk menghitung akurasi. Ditunjukkan pada gambar 9



Gambar 9. Form Training Klasifikasi

**g. Cek klasifikasi naïve bayes**

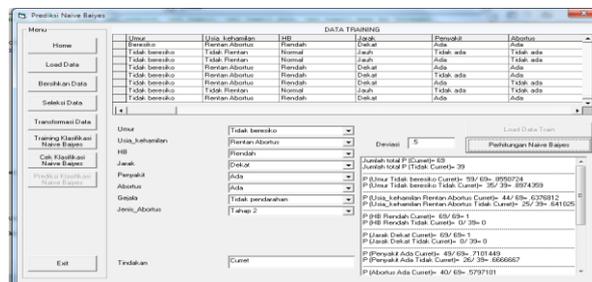
Proses cek klasifikasi merupakan hasil akurasi dari pembagian training dan testing dalam data. Ditunjukkan pada gambar 10



Gambar 10 Form Cek Klasifikasi

**h. Prediksi Klasifikasi Naïve Bayes**

Proses ini dilakukan untuk melakukan prediksi pada data pasien baru sebagai gambaran sebuah tindakan yang akan didapatkan pasien. Ditunjukkan pada gambar 11.



Gambar 11 Form Prediksi Klasifikasi

**5. Kesimpulan**

Berdasarkan hasil pengembangan dan pembahasan maka diperoleh kesimpulan sebagai berikut:

1. Dari hasil penelitian yang telah dilakukan, maka dihasilkan perangkat lunak (software) baru tentang penerapan data mining menggunakan metode naïve bayes untuk klasifikasi tindakan jenis abortus di RSUD Duta Mulya.
2. Perangkat lunak yang dihasilkan mampu menerapkan proses data mining untuk mengklasifikasikan tindakan jenis abortus, dengan menghasilkan rule atau kamus data. Kebenaran yang terbaik dari 2 percobaan adalah dengan menggunakan data training sebanyak 10% yaitu sebanyak 14 data dan testing 90% sebanyak 121 data, nilai kebenarannya adalah 99,26%.

## 6. Saran

Saran yang berkaitan dengan aplikasi untuk menerapkan data mining dalam mengklasifikasikan tindakan jenis abortus adalah :

1. Aplikasi yang sudah dibuat masih bisa dikembangkan lagi, sehingga perlu penelitian lebih lanjut untuk mengoptimalkan metode naïve bayes classifier.

### Daftar Pustaka

- [1] Saifuddin, Abdul Bari. (2008). Acuan Nasional Pelayanan Kesehatan Maternal dan Neonatal. Jakarta : Yayasan Bina Pustaka Sarwono Prawirohardjo.
- [2] Almira Syawli, Gopi Yudoyo, M. Ali Fahmi, Silvia AS, dan Zulkarnaein. 2015. “Diagnosa Penyakit Diabetes Mellitus Dengan Metode Naïve Bayes Berbasis Desktop Application”. Teknik Informatika, Program Teknologi Informasi dan Ilmu Komputer, Universitas Brawijaya, Malang, 65145, Indonesia .
- [3] Kabir, M, F, Rachman, C.M.,Hossain,A.,& Dahal, K(2011) Enhanced Classification Accuracy on Naïve Bayes Data Mining Models. International Journal of Computer Application, Volume 28. No.3
- [4] Royani Chairiyah. (2010) “Faktor-Faktor Yang Mempengaruhi Ibu Hamil Dengan Abortus Inkomplit Di Rsud Kota Bekasi”.
- [5] Budi santosa, “Data Mining : Teknik Pemanfaatan Data Untuk Keperluan Bisnis”. Graha Ilmu, Yogyakarta, 2007.
- [6] Neila, R., Djamaludin, A., Yuliardi, S., Peno, S., 2012. Teacher Quality Improvement Program: Empowering teachers to increasing a quality of Indonesian’s education, Indonesia. Journal of Procedia-Social and Behavioral Sciences 69 (2012 ) 1836 – 1841
- [7] Winkjosastro. 2008. Ilmu Kebidanan, edisi 3, cetakan 5, Jakarta. Yayasan Bina Pustaka Sarwono Prawirohardjo.
- [8] Bobak, 2010. “Buku Ajar Keperawatan Maternitas “ Edisi 4. Penerbit Buku Kedokteran EGC, Jakarta
- [9] Mochtar R. Sinopsis Obstetri Jilid II. Edisi II. Jakarta: 1998.
- [10] Han, J. And Kamber, M, 2006. “Data Mining Concept and Techniques Second Edition”. Morgan Kauffman, San Fransisco.