

PENERAPAN ALGORITMA C4.5 PADA KLASIFIKASI KELAHIRAN BAYI PREMATUR DI DESA SETIA MEKAR

Aswan Supriyadi Sunge, Ana Angelia Aditasari

Program Studi Teknik Informatika Sekolah Tinggi Teknologi Pelita Bangsa

aswan.s@pelitabangsa.ac.id

Disetujui, 05 Maret 2018

Abstrak

Setiap ibu pasti ingin melahirkan anak dengan sempurna dan sehat. banyak hal yang menyebabkan bayi baru lahir meninggal, beberapa diantaranya adalah kekurangan gizi pada saat dalam kandungan, janin yang memiliki kelainan pada tubuh, faktor kelahiran prematur. Kematian akibat kehamilan prematur tentunya menjadi masalah yang serius. Ada beberapa faktor yang mempengaruhi persalinan prematur, seperti faktor plasenta dan janin, faktor ibu atau ibu, dan faktor gaya hidup. Permasalahan yang terjadi di kawasan Tambun Selatan khususnya di Desa Setia Mekar pada tahun 2017 dimana terdapat bayi yang lahir secara spontan akibat mengancam persalinan prematur sehingga kehamilan tidak dapat dipertahankan sehingga menambah jumlah kelahiran prematur. Oleh karena itu penulis tertarik dan ingin mengulas untuk mengetahui ciri-ciri ibu bersalin prematur dan normal. Dari hasil pengujian yang dilakukan terhadap 130 data dengan menggunakan metode Rapidminer dan Algoritma C4.5 menghasilkan matriks kounfusi dan kurva ROC dengan nilai akurasi 80.77%. Nilai akurasi tersebut dapat dikatakan cukup baik untuk mengetahui karakteristik ibu melahirkan prematur dan normal.

Kata kunci: Data Mining, algoritma C4.5, kelahiran prematur

Abstract

Every mother would want to give birth to children perfectly and healthy. many things that cause a newborn baby to die, some of which are lack of nutrition at the time in the womb, a fetus that has a defect in the body, premature birth factor. Death due to premature pregnancy certainly becomes a serious problem. There are several factors that influence preterm labor, such as placental and fetal factors, maternal or maternal factors, and lifestyle factors. Problems that occur in the South Tambun area, especially in the village of Setia Mekar in 2017 where there are babies born spontaneously due to threatening premature labor that can not be maintained pregnancy so that adds to the number of premature births. Therefore, the authors are interested and want to review to find out the characteristics of premature and normal maternity mothers. From the results of tests conducted on 130 data using Rapidminer and C4.5 Algorithm methods that produce a counfusion matrix and ROC curve with an accuracy value of 80.77%. The accuracy value can be said to be good enough to know the characteristics of premature and normal delivery mothers.

Keyworad: Data Mining, C4.5 algorithm, premature birth

1. Pendahuluan

Setiap ibu pastinya ingin melahirkan anak dengan sempurna dan sehat. banyak hal yang menyebabkan bayi yang baru lahir meninggal, beberapa diantaranya adalah kekurangan nutrisi pada saat dalam kandungan, janin yang memiliki cacat dalam tubuhnya, faktor kelahiran usia prematur. Kematian akibat kelahira prematur tentunya menjadi masalah yang cukup serius. Persalinan prematur adalah persalinan dengan usia kehamilan antara 20 minggu sampai 37 minggu. Hampir setiap tahun, tercatat sekitar lima belas juta bayi lahir dalam keadaan prematur di seluruh dunia, sekitar satu juta di antaranya meninggal akibat komplikasinya, dan sebagian lainnya mengalami kecacatan seumur hidup, termasuk ketidakmampuan belajar dan gangguan penglihatan dan pendengaran. Lebih dari 60% kelahiran prematur terjadi di negara-negara Afrika dan Asia bagian selatan (WHO. 2013).

Terdapat beberapa faktor yang berpengaruh terhadap persalinan prematur, seperti faktor plasenta dan janin, faktor ibu atau maternal, serta faktor gaya hidup. Faktor plasenta dan janin yang dapat menyebabkan kelahiran prematur adalah pertumbuhan janin terlambat dan dapat menimbulkan kecil untuk masa kehamilan, faktor inkompatibilitas darah, serta terdapat pemicu persalinan prematur seperti solusio plasenta, plasenta previa, dan infeksi yang dapat menyebabkan karioamnionitis. (Manuaba dkk. 2007).

Permasalahan yang terjadi di daerah Tambun Selatan khususnya di daerah desa Setia Mekar pada tahun 2017 dimana terdapat bayi lahir dengan spontan akibat persalinan prematur mengancam yang tidak dapat dipertahankan kehamilannya sehingga menambah angka kelahiran prematur. Pada kasus persalinan prematur mengancam dapat diupayakan untuk mengurangi angka kematian ibu dan bayi dengan menentukan faktor-faktor penyebabnya. Kemajuan teknologi informasi telah menyebabkan banyak orang memperoleh data dengan mudah. Banyaknya data yang dimiliki instansi kesehatan yang bisa menyebabkan kesulitan dalam mengklasifikasi data. Kegiatan pengklasifikasian oleh manusia masih memiliki keterbatasan, salah satunya cara mengatasi masalah pengklasifikasian ini dengan menggunakan Data Mining. Selama beberapa dekade terakhir, cukup banyak algoritma Data Mining yang telah diusulkan oleh peneliti kecerdasan komputasi untuk memecahkan masalah klasifikasi di dunia nyata.

Berdasarkan permasalahan yang telah disebutkan diatas, perlu untuk kita mengetahui faktor-faktor apa saja yang mempengaruhi usia kelahiran pada bayi agar tidak terjadi hal yang tidak diinginkan kedepannya. Karena belum adanya suatu prediksi dalam klasifikasi kelahiran bayi prematur. Penelitian ini bertujuan untuk memprediksi usia kelahiran menjadi 2 kategori berdasarkan faktor-faktor yang dialami oleh ibu hamil dengan menggunakan prediksi. Kategori usia kelahiran yang ada di penelitian ini diantaranya adalah Prematur dan Normal atau cukup bulan. Berdasarkan uraian diatas maka penulis tertarik dan ingin mengkaji ulang untuk mengetahui gambaran karakteristik ibu bersalin Prematur dan Normal,

2. Tinjauan Studi

teori yang digunakan untuk mendukung penelitian tugas akhir dalam mengklasifikasi persalinan prematur atau Normal dengan menggunakan algoritma C4.5.

2.1 Pengertian Data Mining

Data Mining adalah suatu proses logis yang digunakan untuk mencari dan menemukan pola melalui sejumlah besar data tujuan dari teknik ini adalah untuk menemukan pola yang sebelumnya tidak diketahui. Dalam Data Mining tumpukan data masa lalu ini dianggap sebagai tambang yang dapat diolah untuk menghasilkan sebuah pengetahuan yang berharga (Bharati, 2011). Data Mining didefinisikan sebagai sebuah proses untuk menemukan hubungan, pola dan tren baru yang bermakna dengan menyaring data yang sangat besar, yang tersimpan dalam penyimpanan, menggunakan teknik pengenalan pola seperti teknik Statistik dan Matematika (Larose, 2005).

Process Data Mining bertujuan untuk menghasilkan patterns yang berguna dari koleksi text. Aktivitas text mining untuk step data mining terdiri dari pemilihan mining teknik yang benar, penentuan mining model dan parameters. Term Clustering dan Parsial Parsing bisa dimasukkan dalam step ini. Parsial parsing atau robust parsing bertujuan untuk mengidentifikasi relationship yang lebih dalam antar kata-kata dalam kalimat. Parsial parsing memerlukan hasil dari POS Tagging dan biasanya di gunakan secara bersamaan. Teknik penggunaan rule based system, memory based system, statistical method, atau kombinasi antar teknik banyak di gunakan untuk parsial parsing. (adiwijaya, 2006).

Berdasarkan definisi-definisi yang telah disampaikan, hal penting yang terkait dengan Data Mining adalah (Larose, 2005) :

1. Data Mining merupakan suatu proses untuk menemukan hubungan.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan Data Mining adalah mendapatkan hubungan atau pola statistik dan Matematika yang mungkin memberikan indikasi yang bermanfaat.

2.2 Pengertian Klasifikasi

Klasifikasi merupakan bagian dari prediksi, dimana nilai yang diprediksi berupa label. Klasifikasi menentukan class atau grup untuk tiap contoh data, input dari model klasifikasi adalah atribut dari contoh data (Data samples) dan outputnya adalah class dari Data samples itu sendiri, dalam machine learning untuk membangun model klasifikasi digunakan metode supervised learning. (Moghimpour, dkk, 2014) Klasifikasi merupakan pengelompokan sampel berdasarkan ciri-ciri persamaan dan perbedaan dengan menggunakan variabel target sebagai kategori. Ada beberapa macam pengklasifikasian dalam Data Mining yaitu decision tree, naive Bayes, svm, dan lain-lain (Larose, 2005). Klasifikasi merupakan suatu pekerjaan menilai objek daya untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia (Prasetyo, 2012).

Klasifikasi merupakan pengelompokkan Data *sample* dengan ciri-ciri persamaan dan perbedaan dengan memasukan kedalam kelas tertentu.

2.3 Pengertian Pohon Keputusan (*Decision Tree*)

Pohon keputusan merupakan metode klasifikasi dari prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah *fakta yang sangat besar menjadi pohon keputusan yang mempresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti Structured Query Language untuk mencari record pada kategori tertentu. Pohon keputusan juga berguna untuk mengeksplorasi data menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target (Kusrini, 2009).*

Decision Tree (Pohon Keputusan) adalah pohon klasifikasi yang digunakan sebagai prosedur penalaran untuk mendapatkan jawaban dari masalah yang dimasukkan. Pohon yang dibentuk tidak selalu berupa pohon biner. Jika semua fitur dalam data menggunakan dua macam nilai kategorik maka bentuk pohon yang didapatkan berupa pohon biner namun jika dalam fitur berisi lebih dari dua macam nilai kategorikal atau menggunakan tipe numerik maka bentuk pohon yang didapatkan biasanya tidak berupa pohon biner (Prasetyo, 2012).

Pohon keputusan (decision tree) merupakan salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasikan. Pohon keputusan adalah model prediksi menggunakan struktur pohon atau struktur berhirarki. Pohon keputusan adalah struktur flowchart yang menyerupai pohon (tree), dimana setiap percabangan menyatakan kondisi yang harus dipenuhi dan tiap ujung pohon menyatakan kelas data. Pada pohon keputusan setiap simpul daun menandai label kelas. Proses dalam pohon keputusan yaitu mengubah bentuk data (tabel) menjadi model pohon (tree) kemudian mengubah model pohon tersebut menjadi aturan (rule) (Pramudiono, 2007).

Pohon keputusan atau Decision Tree merupakan proses pelatihan data set yang memiliki atribut dengan besaran nominal, yaitu bersifat kategoris dan setiap nilai tidak bisa dijumlahkan atau dikurangkan. Pada umumnya, ciri kasus berikut cocok untuk diterapkan pada decision tree:

- a. Data/example dinyatakan dengan pasangan atribut dan nilainya.
- b. Label/output data biasanya berniali diskrit.
- c. Data mempunyai missing value.

2.4 Pengertian Algoritma C4.5

Algoritma C4.5 adalah pengembangan dari ID3 (*Iterative Dichotomiser 3*) yang merupakan salah satu algoritma pohon keputusan (Fairuz El Said, 2009). Algoritma C4.5 merupakan salah satu solusi pemecahan kasus yang sering digunakan dalam pemecahan masalah pada teknik klasifikasi. Keluaran dari algoritma C4.5 itu berupa sebuah decision tree layaknya teknik klasifikasi lain. Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan *record* yang lebih kecil dengan menerapkan serangkaian aturan keputusan. Dengan masing-masing rangkaian pembagian, anggota himpunan hasil menjadi mirip satu dengan yang lain (Nofriansyah, 2017). Secara umum Algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut (Jefri, 2013):

- a. Pilih atribut sebagai akar
- b. Buat cabang untuk masing-masing nilai
- c. Bagi kasus dalam cabang
- d. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama. Algoritma C4.5 merupakan perbaikan dari ID3 menggunakan *Gain Ratio* untuk diperbaharui *information gain* maka dengan rumus :

$$GainRatio(S, A) = \frac{Gain(S, A)}{SplitInfo(S, A)}$$

Dimana;

S = Ruang/Data Sample yang dipergunakan untuk Data *Training*

A = Atribut

Gain(S,A) = information gain pada atribut A SplitInfo(S,A) = split information pada atribut A
 Dengan atribut yang nilai *Gain Ration* yang paling tertinggi dipilih sebagai atribut test untuk simpul.

Pendekatan ini menerapkan Normalisasi pada information gain dengan menggunakan apa yang disebut dengan *split information*, dengan rumus :

$$\dots\dots\dots(2.2) \quad SplitInfo(S,A) = - \sum_{i=1}^i \frac{S_i}{S} \log_2 \frac{S_i}{S}$$

Dimana:

S = Ruang (Data) sample yang digunakan untuk Training. A = Atribut.
 Si = Jumlah sample untuk atribut i

Pada saat membangun pohon keputusan, mungkin adanya data yang noise maupun kosong pada *Training Data*. Pemangkasan pohon dapat dilakukan untuk mengenali dan menghilangkan cabang tersebut agar pohon lebih kecil dan lebih mudah dipahami agar lebih baik dalam melakukan klasifikasi.

Ada dua metode dalam melakukan pemangkasan pohon keputusan yaitu:

- a. Dengan Prepruning yaitu menghentikan pembangunan lebih awal pada *subtree* agar tidak sampai lebih jauh dalam pemangkasan Data Training. Rumus Prepruning :

$$c = \frac{r + \frac{Z^2}{2n} + Z \sqrt{\frac{r}{n} - \frac{r^2}{n} + \frac{Z^2}{4n^2}}}{1 + \frac{Z^2}{n}} \dots\dots\dots(2.3)$$

Dimana :

r = nilai perbandingan *error rate*

n = total samplez = $\Phi^{-1}(c)$

c= *confidence level*

- b. Dengan Postpruning yaitu menyederhanakan pohon dengan membuang beberapa cabang *subtree* setelah selesai dibangun. Metode ini sebagai bagian standard untuk Algoritma C4.5.

2.5 Pengertian Persalinan

Menurut Manuaba (2010), persalinan adalah proses pengeluaran hasil konsepsi (janin dan urin) yang telah cukup bulan atau dapat hidup diluar kandungan atau melalui jalan lain dengan bantuan atau tanpa bantuan. Menurut Sarwono (2011), persalinan adalah proses pengeluaran janin yang terjadi pada kehamilan cukup bulan (37-42 minggu), lahir spontan dengan presentasi belakang kepala yang berlangsung dalam 18 jam tanpa komplikasi baik pada ibu maupun pada janin. Menurut Yanti (2010), persalinan adalah serangkaian kejadian yang berakhir dengan pengeluaran bayi cukup bulan atau hampir cukup bulan, disusul dengan pengeluaran plasenta dan selaput janin dari tubuh ibu. Dari definisi diatas dapat disimpulkan bahwa persalinan merupakan suatu kejadian yang dialami oleh wanita hamil dengan tujuan mengeluarkan bayi yang ada dalam kandungan dengan usia cukup bulan, baik dengan bantuan atau tanpa bantuan.

2.6 Definisi bayi prematur

Menurut WHO, bayi prematur adalah bayi lahir hidup sebelum usia kehamilan minggu ke-37 (dihitung dari hari pertama haid terakhir). *The American Academy of Pediatric*, mengambil batasan 38 minggu untuk menyebut prematur. Bayi prematur adalah bayi yang lahir di bawah dari 37 minggu atau berat bayi kurang dari 2.500 gram (Manuaba, 2009). Bayi prematur merupakan bayi yang lahir pada usia kehamilan kurang atau sama dengan 37 minggu, tanpa memperhatikan berat badan lahir (Manuaba,

2009). Bayi prematur adalah bayi yang lahir setelah 24 minggu dan sebelum 37 minggu kehamilan, dengan berat badan 2500 gram atau kurang saat lahir, terlepas dari usia kehamilan tepat atau dibawah 37 minggu (Brooker, 2008). Secara patofisiologis menurut Nelson (2010), bayi BBLR ini berhubungan dengan usia kehamilan yang belum cukup bulan (prematur) disamping itu juga disebabkan dismaturitas. Bayi lahir cukup bulan (usia kehamilan 38 minggu), tapi berat badan (BB) lahirnya lebih kecil dari masa kehamilannya, yaitu tidak mencapai 2.500 gram.

2.7 Penyebab Kelahiran Bayi Prematur

Faktor *predisposisi* terjadinya kelahiran prematur diantaranya:

- 1) Faktor ibu yaitu riwayat kelahiran prematur sebelumnya, perdarahan antepartum, *malnutrisi*, kelainan uterus, hidromion, penyakit jantung /penyakit kronik lainnya, hipertensi, umur ibu kurang dari 20 tahun atau lebih dari 35 tahun, jarak dua kehamilan yang terlalu dekat, infeksi, trauma, kebiasaan, yaitu pekerjaan yang melelahkan, merokok (Prawirohardjo, 2006).
- 2) Faktor janin yaitu cacat bawaan, kehamilan ganda, hidramion, ketuban pecah dini cacat bawaan dan infeksi (Prawirohardjo, 2006).
- 3) Faktor Plasenta Kelahiran prematur yang disebabkan oleh faktorplasenta meliputi: plasenta previa, dan solutio plasenta (Surasmi, Handayani, dan Kusuma, 2003)
- 4) Faktor Khusus : serviks inkompeten Persalinan prematur berulang, overistensi uterus, kehamilan ganda, kehamilan dengan hidramnion (Manuaba dkk, 2007)
- 5) Terjadi produksi prostaglandin. Secara anatomis kutub bawah persambungan selaput janin dengan desidua yang menutupi koralis servikalis tersambung dengan vagina. Meskipun demikian susunan anatomis ini menyediakan jalan masuk bagi penyebaran mikroorganisme ke dalam jaringan intrauteri dan kemudian menginvasi kantong amnion. Mikroorganisme ini menginduksi pembentukan sitokin yang memicu produksi prostaglandin dan mendorong terminasi kehamilan lebih dini (Cunningham, 2006).
- 6) Terjadi pada wanita multipara, karena adanya jaringan parut uterus akibat kehamilan dan persalinan sebelumnya (berulang). Yang menyebabkan tidak adekuatnya persediaan darah ke plasenta sehingga plasenta menjadi lebih tipis dan mencakup uterus lebih luas. Plasenta yang melekat tidak adekuat ini mengakibatkan isoforitin yang merupakan protein hasil produksi sel limfosit T untuk menghambat reaktivitas uterus dan melindungi buah kehamilan diproduksi sedikit. Sehingga dengan keadaan demikian risiko untuk mengalami persalinan prematur menjadi lebih besar (Raymond, 2006).
- 7) Wanita yang pernah melahirkan lebih dari 1 kali atau yang termasuk paritas tinggi mempunyai risiko lebih tinggi mengalami partus prematur karena menurunnya fungsi alat reproduksi dan meningkatkan pula risiko terjadinya perdarahan antepartum yang dapat menyebabkan terminasi kehamilan lebih awal (Saifudin, 2008).

Prawirohardjo (2011) menyatakan bahwa kondisi yang terjadi selama kehamilan dapat berisiko terhadap kejadian persalinan prematur yang dibagi dalam dua faktor, yaitu:

1. Janin dan plasenta
 - a. perdarahan trimester awal
 - b. perdarahan antepartum (plasenta previa, solution plasenta, vasa previa)
 - c. ketuban pecah dini (KPD)
 - d. pertumbuhan janin terhambat
 - e. cacat bawaan janin
 - f. kehamilan ganda/gemeli
 - g. polihidramnion
2. Ibu
 - a. penyakit berat pada ibu
 - b. diabetes mellitus
 - c. preeklamsia/hipertensi
 - d. infeksi saluran kemih/genital/intrauterin
 - e. penyakit infeksi dengan masalah kehamilan

- f. stress psikologik
- g. kelainan bentuk uterus/serviks
- h. riwayat persalinan prematur/abortus berulang
- j. pemakaian obat narkotik
- k. trauma perokok berat
- l. kelainan imunologik/kelainan resus

3. Hasil Penelitian Dan Pengujian

Hasil penelitian dilakukan dengan menggunakan aplikasi *Rapidminer* 8.2.1 dengan jumlah data sebanyak 130 data ibu hamil. Berikut ini adalah hasil dari pengujian yang dilakukan.

4.1.1 Evaluasi dan Validasi hasil Algoritma C4.5

Data yang digunakan pada penelitian ini adalah sebanyak 130 data yang diambil dari data pasien. Langkah pembuatan pohon keputusan (*Desicion Tree*), pada langkah awal dari proses evaluasi algoritma C4.5. Data Ibu hamil yang akan diproses adalah *dataset* keseluruhan, untuk melakukan pengelolaan data di *Rapidminer*.

4.2 Perhitungan Algoritma C4.5

Menghitung nood (akar) terlebih dahulu pada *dataset* untuk mendapatkan hasil *entropy* total pada pohon keputusan, dapat dilihat pada tabel berikut ini:

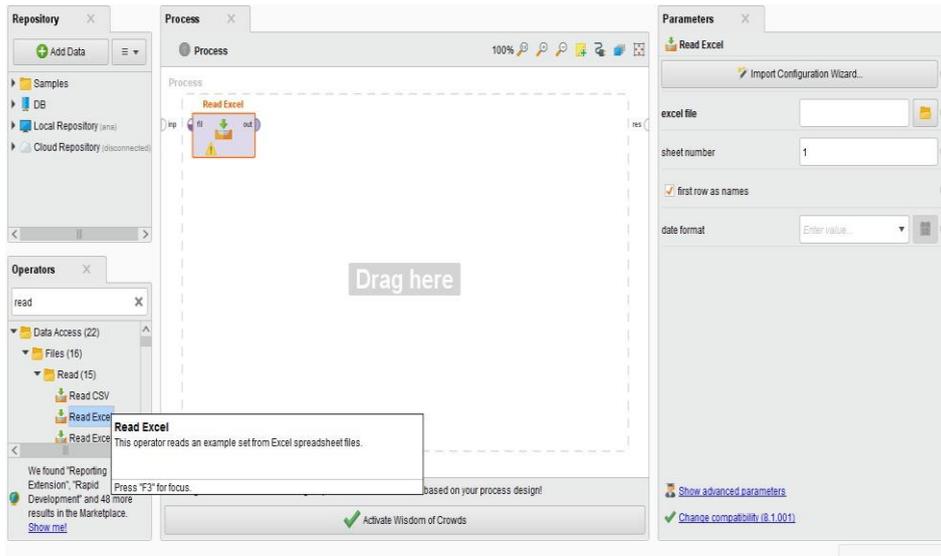
Tabel 4. 1 *Entropy Total*

Hasil	Frekuensi	PJ	LOG 2*PJ	-Pj*LOG2*Pj
<i>Premature</i>	71	0.546154	-0.8726215	0.47659
Normal	59	0.45384	-1.13972	0.51726
Total	130			0.99384

Setelah hasil *entropy* total diketahui maka selanjutnya mencari hasil *entropy* dan *gain* pada masing-masing atribut untuk menentukan akar pohon pada *dataset* Ibu Hamil. Berikut ini tabel hasil perhitungan *entropy* atribut untuk menentukan akar pohor pada *Decision Tree C4.5*:

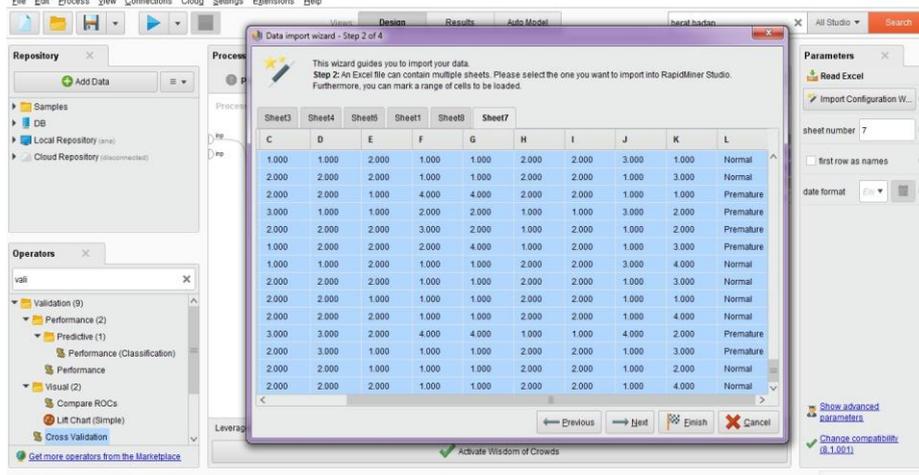
4.3 Pengelolaan Dataset

Langkah awal pengelolaan *dataset* adalah *read excel* cari pada kolom *operators*. Lalu drag ke kolom *process* gambar *read excel dataset* yang disajikan pada gambar 4.1 berikut ini:



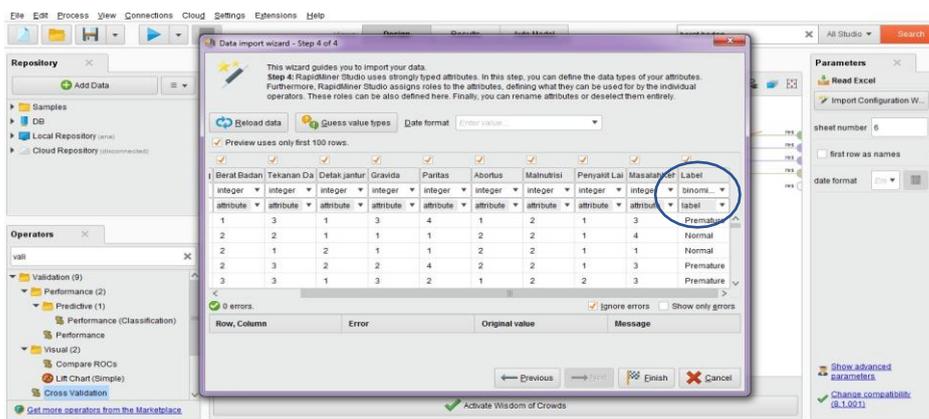
Gambar 4. 1 Read Excel

Klik *read excel* lalu pada pojok kanan atas klik *import configuration wizard*. Setelah itu pilih *dataset*, pemilihan *dataset* disajikan pada gambar 4.2 sebagai berikut ini:



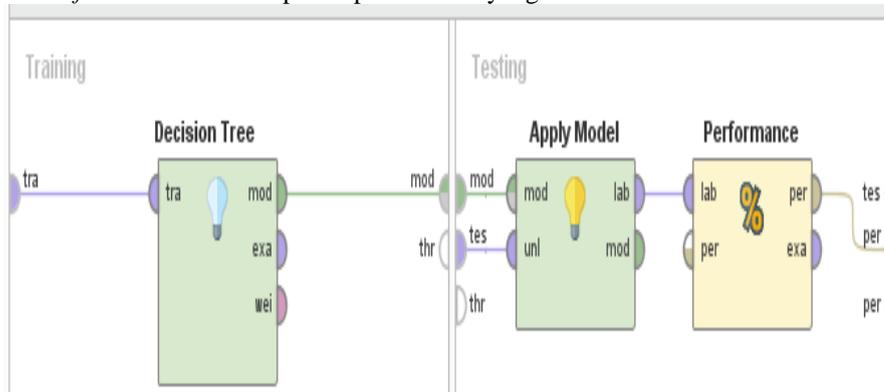
Gambar 4. 2 Pemilihan Dataset

Klik tombol *next*, lalu *next* lagi, setelah itu pilih hasil keputusan *binominal* dan *label*, dari 12 atribut yang digunakan 11 atribut menjadi atribut *class* dan 1 atribut menjadi atribut. Sedangkan yang lainnya sebagai *integer* dan *atribut*. Setelah itu klik *finish*. Lalu pilih *cross validation* pada kolom operators, dan pada kolom *cross validation* klik dua kali untuk proses pada *cross validation*. Gambar pemilihan hasil keputusan sebagai tabel (*dataset* Ibu Hamil) disajikan pada gambar 4.3 sebagai berikut ini:



Gambar 4. 3 Pemilihan Hasil Keputusan Sebagai Tabel (Dataset)

Pada penelitian ini dilakukan proses validasi menggunakan fitur *cross validation* dengan nilai $k = 10 \text{ folds}$. Berikut merupakan permodelan yang ada didalam *cross validation*.



Gambar 4. 4 Model Cross Validation dengan Decision Tree C4.5

Pada permodelan *cross validation* didalamnya terdapat dua bagian, yaitu bagian *training* dan *testing*. Bagian *training* digunakan untuk algoritma klasifikasi *Decision Tree* dan bagian *testing* menggunakan fitur *Apply Model* untuk mengaplikasikan model pada data *testing* dan fitur *Performance* untuk menampilkan *confusion table*, yang digunakan untuk menampilkan hasil dari *accuracy*, *recall*, *precision*, dan nilai AUC.

Setelah itu klik proses untuk masuk ke gambar read *excel* dan *validation*. Tarik garis dan sambung pada setiap kotaknya, pada kotak *validation* sambung empat garis. Setelah proses dilakukan klik *start the execution* (segitiga biru diatas) yang ada pada *border* untuk proses *run*.
Pembahasan

Penelitian dilakukan dengan menggunakan algoritma klasifikasi *Decision Tree C4.5* dengan menghasilkan nilai *accuracy*, *recall*, *precision*, dan AUC. Berikut tabel hasil dari penelitian yang telah dilakukan:

Tabel 4.1 Hasil Penelitian

Criterion	Pengujian
	Decision Tree
Accuracy	80,77%
Recall	77,67%
Precision	81,45%
AUC	0,805

Berdasarkan hasil penelitian diperoleh nilai *accuracy*, *recall*, *precision*, dan nilai AUC. Dari hasil yang telah didapat penelitian mengenai klasifikasi dalam kelahiran bayi *premature* menggunakan algoritma C4.5 dengan kurva *ROC* mendapatkan nilai akurasi *good classification* sebesar 0,805.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan menggunakan penambahan Data algoritma C4.5 untuk mengklasifikasi persalinan Normal atau Prematur, dapat diambil kesimpulan sebagai berikut :

1. Penambahan Data untuk mengklasifikasi persalinan Normal atau Prematur dapat

diimplementasikan dengan baik menggunakan algoritma C4.5 pada *rapidminer*.

2. Pengujian yang dilakukan terhadap 130 Data menggunakan *counfusion matrix* maupun *ROC curve* menghasilkan tingkat keakuratan sebesar 80.77%. Nilai keakuratan tersebut dapat dikatakan cukup baik untuk mengetahui gambaran karakteristik ibu bersalin Prematur dan Normal.

Saran

Saran yang dapat diberikan untuk mengembangkan penelitian yang akan Datang adalah :

3. Pendekatan klasifikasi yang mampu menerima masukan dengan jumlah atribut yang bervariasi di penelitian mendatang.
4. Penelitian selanjutnya diharapkan dapat mengembangkannya dalam suatu aplikasi.

Daftar Pustaka

- Adiwijaya, igg. 2006. *“Text Mining dan Knowledge Discovery”*. Jakarta. Data mining indonesia
- Brooker, Cris. 2008. *“Ensiklopedia Keperawatan”*. Jakarta : EGC
- Carolina, Ramanda. 2018. *“Klasifikasi Kelahiran Prematur Menggunakan Algoritma C.4”* STMIK Nusa Mandiri, Jakarta Cunningham, G. 2006. *“Obstetri William vol.1”*. Jakarta: EGC
- Fairuz,E. S. (2009). *“data-mining-konsep-pohon-keputusan”*. Jakarta: fairuzelsaid.wordpress.com/2009/11/24/data-mining-konsep-pohon- keputusan/ diakses 11 juli 2018
- Jefri. 2013. *“Implementasi Algoritma C4.5 Dalam Aplikasi Untuk Memprediksi Jumlah Mahasiswa Yang Mengulang Mata Kuliah Di STMIK AMIKOM Yogyakarta”*, Yogyakarta Kusri, Luthfy, E.T. 2009. *“Algoritma Data Mining.”* Yogyakarta : Andi Larose, 2005. *“Discovering Knowledge in Data: An Introduction to Data Mining”*, John Willey & Sons, Inc.
- Manuaba, I. B. G. 2001. *“Ilmu Kebidanan Penyakit Kandungan dan Keluarga Berencana”*. Jakarta. EGC
- Manuaba, I. B. G., I.A Chandranita Manuaba. 2007. *Pengantar Kuliah Obstetri*. Jakarta : EGC.
- Manuaba, AC., 2009. *“Buku Ajar Patologi Obstetri Untuk Mahasiswa Kebidanan.”* Penerbit Buku Kedokteran. Jakarta. EGC
- Manuaba, Ida Ayu Chandranita. 2010. *“Gadar Obstetri & Ginekologi &Obstetric Ginekologi Sosial Untuk Profesi Bidan.”* Jakarta. EGC
- Manuaba, IAC., I Bagus, dan IB Gde. 2013. *“Ilmu Kebidanan, Penyakit Kandungan dan KB untuk Pendidikan Bidan”*. Edisi kedua. Jakarta: EGC.
- Maria, Laurensia. 2017. *“Klasifikasi persalinan Normal atau caesar menggunakan algoritma C4.5”*, Yogyakarta
- Moghimpour, Ida dan Ebrahimou.2014. *“Comparing Decision Tree Method Over Three Data Mining Software.”* International Journal of Statistics and Probability;
- Nelson. 2010. *“Ilmu Kesehatan Anak”*. Volume 2. Jakarta: EGC
- Nofriansyah, Dicky. 2017 *“Modul Data Mining”*. Medan. STMIK Triguna dharma.
- Prawirohardjo, S. 2006. *“Buku Acuan Nasional Pelayanan Kesehatan Maternal dan Neonatal”*. Jakarta : Yayasan Bina Pustaka.
- Prawirohardjo, Sarwono. 2008. *“Ilmu Kebidanan”*. Jakarta: Yayasan Bina pusaka. Prawirohardjo, Sarwono. 2009. *“Ilmu Kebidanan”*. Jakarta:Yayasan Bina Pustaka. Prawirohardjo, Sarwono. 2011. *“Ilmu Kebidanan”*. Jakarta:BP-SP
- Puspita, Ari. 2016. *“Prediksi Kelahiran Bayi Secara Prematur Dengan Menggunakan Algoritma C4.5 Berbasis Particle Swarm Optimazion “*
- Pramudiono, Iko. *“Pengantar Data Mining: Menambang Permata Pengetahuan di Gunung Data.”* <http://www.ilmukomputer.com>. Diakses tanggal 18 juli 2018
- Prasetyo, E., 2012. *“Klasifikasi: Metode-Metode Pilihan.”* Yogyakarta: Penerbit ANDI
- R. M. Bharati, *“Data Mining Technique and Applications,”* Indian Journal of Computer Science and Engineering,
- Raymond, 2006. *“Obstetric and Goinekology”*. Jakarta: Hipokrates.
- Sunge, Aswan. 2018. *“Prediksi Kompetensi Karyawan Menggunakan Algoritma C4.5(Studi Kasus :PT Hankook Tire Indonesia)”*
- Surasmi A., Handayani S., Kusuma H. (2003). *“Perawatan Bayi Resiko Tinggi”*. Jakarta: EGC.
- Yanti. 2010. *“Buku Ajar Asuhan Kebidanan Persalinan.”* Yogyakarta: Pustaka Rihama
- Word Health Organization (WHO), 2013. *“Born too soon preterm birth in europe trends, causes and prevention”* www.searo.who.int. Diakses 02 juli 2018

World Health Organization (WHO), 2016. “*Over half a million premature deaths annually in the European Region attributable to household and ambient air pollution*”. www.searo.who.int.
Diakses 02 juli 2018
<https://afidburhanuddin.wordpress.com/2013/09/24/tahapan-penelitian/> (18 juli 2018).