

Penerapan Klasifikasi dengan Algoritma C4.5 pada Antenatal Care Terhadap Status Risiko Preeklamsia Ibu Hamil di Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh Jakarta Timur

Sugiyono¹, Suhartini², Rasiban³, Untung Surapati⁴

^{1, 3, 4}Informatika, Fakultas Ilmu Komputer, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika

²Sistem Informasi, Fakultas Ilmu Komputer, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika

*Email: inosoguy007@gmail.com¹, suhartinish410@gmail.com², rasiban@gmail.com³, kisuro2003@gmail.com⁴

Abstrak

Perawatan kehamilan atau antenatal care adalah serangkaian layanan medis penting yang dirancang untuk menjamin kesehatan ibu dan janin selama masa kehamilan. Preeklamsia merupakan komplikasi yang menyerang ibu dan janin selama kehamilan. Preeklamsia dibagi menjadi dua kategori, preeklamsia berat dan preeklamsia ringan. Data pemeriksaan kehamilan atau kunjungan antenatal ibu hamil yang tercatat di Tempat Praktik Bidan Mandiri (TPMB) Lilik Faiqoh digunakan untuk mengklasifikasikan faktor risiko terjadinya preeklamsia. Gejala preeklamsia pada ibu hamil ditandai dengan tekanan darah tinggi, kelebihan protein dalam urin, dan ada tidaknya edema. Penyebab terjadinya preeklamsia adalah kurangnya kesadaran terhadap status kesehatan ibu hamil, tidak mengetahui faktor penyebab terjadinya preeklamsia, dan tidak menyadari gejala awal risiko kehamilan. Untuk mengatasi masalah ini, kita perlu meningkatkan kualitas pelayanan terhadap ibu hamil. Teknologi data mining dengan metode algoritma C4.5 digunakan untuk pengolahan data pada penelitian ini. Tujuan penelitian ini adalah mengembangkan model klasifikasi algoritma C4.5 untuk deteksi dini status preeklamsia pada ibu hamil. Penerapan penelitian ini menggunakan metode CRISP-DM yang melalui tahapan pemahaman bisnis, pemahaman data, penyiapan data, pemodelan, evaluasi, serta pengembangan. Pengolahan dataset sebanyak 300 data menggunakan *Rapidminer* untuk mengklasifikasikan status risiko preeklamsia ibu hamil yang menghasilkan akurasi sebesar 95,00%.

Kata kunci: Algoritma C4.5, CRISP-DM, Data Mining, Klasifikasi, Risiko Preeklamsia

Abstract

Pregnancy care or antenatal care is a set of essential medical services designed to guarantee the health of the mother and fetus during pregnancy. Preeclampsia is a complication that affects mothers and fetuses during pregnancy. Preeclampsia is divided into two categories, severe and mild. Data from pregnancy examinations or prenatal visits of pregnant mothers recorded at the Lilik Faiqoh Independent Childhood Practice Centre (TPMB) are used to classify risk factors for preeclampsia. Preeclampsia symptoms in pregnant women are characterized by high blood pressure, excess protein in the urine, and lack of edema. Preeclampsia is caused by a lack of awareness of the health status of the pregnant woman, the lack of knowledge of the causes of preeclampsia, and the failure to recognize the early symptoms of the risk of pregnancy. To overcome this problem, we need to improve the quality of service to the pregnant mother. Data mining technology with the C4.5 algorithm method is used for data processing in this study. The aim of this study is to develop a C4.5-algorithmic classification model for early detection of preeclampsia status in pregnant mothers. The application of this research uses CRISP-DM methods that go through the stages of business understanding, data comprehension, data preparation, modeling, evaluation, and development. Processing 300 data sets using Rapidminer to classify pregnant women's preeclampsia risk status yielded an accuracy of 95.00%.

Keywords: *Algorithm, CRISP-DM, Data Mining, Classification, Preeclampsia Risk*

PENDAHULUAN

Kehamilan adalah sebuah keadaan yang membahagiakan untuk kehidupan dalam rumah tangga setiap manusia. Kehamilan tentu memiliki risiko baik untuk ibu maupun janin, maka dari itu untuk menghindarinya ibu hamil harus di perhatikan dengan baik selama kehamilan (Hamzah *et al.*, 2021). Dengan mengidentifikasi faktor risiko kehamilan sejak dini membantu tenaga kesehatan dalam mengetahui pengobatan yang tepat dan akurat. Pemeriksaan rutin oleh ibu hamil ke bidan atau dokter dapat menghindari dan mencegah terjadinya risiko pada kehamilan. Pemeriksaan rutin ini dapat dilakukan di rumah sakit, puskesmas, klinik atau tempat praktik bidan daerah sekitar. Penanganan risiko kehamilan adalah kunci keberhasilan dalam menurunkan angka kematian ibu dan bayi.

Deteksi dini penanganan risiko kehamilan dengan rutin menjalani pemeriksaan selama kehamilan. Antenatal Care (ANC) atau disebut perawatan antenatal merupakan sebuah layanan kesehatan yang diberikan untuk ibu hamil. Pemeriksaan antenatal bertujuan untuk memantau kondisi kesehatan fisik dan mental ibu hamil, sehingga siap ketika menghadapi persalinan, masa nifas, masa pemberian ASI, serta kembalinya alat reproduksi secara alami dan bertahap. Pemeriksaan kehamilan yang sesuai prosedur diharapkan dapat mendeteksi lebih dini ibu hamil apabila ada kelainan dan risiko komplikasi kehamilan seperti preeklamsia, selain itu dengan menjaga pola hidup sehat, mengkonsumsi makanan yang bergizi lengkap dan seimbang, berolahraga, membatasi makanan tinggi garam, mengonsumsi suplemen, vitamin dan mineral untuk ibu hamil juga sangat penting untuk mencegah preeklamsia (Kementrian Kesehatan, 2023).

Preeklamsia dipengaruhi oleh beberapa faktor. Umur, pekerjaan, dan kunjungan antenatal care (ANC) adalah faktor risiko yang memengaruhi kejadian preeklamsia (Nurul Selviana Dewi *et al.*, 2024). Hipertensi atau tekanan darah tinggi menjadi penyebab urutan pertama angka kematian mencapai 33% di Indonesia (SRS Litbangkes, 2016). Nilai kejadian preeklamsia di Indonesia cenderung

besar yakni 3,4% - 8,5% (Kementrian Kesehatan Republik Indonesia, 2021). Preeklamsia merupakan suatu kondisi kelainan pada ibu hamil, dimana terjadinya komplikasi. Komplikasi pada ibu hamil ditandai dengan tekanan darah tinggi, kelebihan protein dalam urin serta dengan atau tanpa edema (Haslan and Trisutrisno, 2022). Preeklamsia termasuk komplikasi dalam kehamilan yang berkelanjutan dengan penyebab yang sama. Preeklamsia umumnya terjadi pada ibu hamil dengan usia rentan 40 tahun keatas. Usia kandungan ibu hamil yang mengalami terjadinya preeklamsia dengan kandungan lebih dari 20 minggu, dan dapat terjadinya persalinan prematur (Retnaningtyas *et al.*, 2021). Faktor yang dapat mempengaruhi terjadinya preeklamsia yaitu usia ibu, paritas, usia kehamilan, status ekonomi dan hiperplasia plasenta.

Hasil dari data pemeriksaan antenatal ibu hamil di Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh nantinya akan di olah dalam penelitian ini. Pengolahan data dilakukan dengan teknik data mining dengan memperkirakan adanya kemungkinan yang terjadi berdasarkan kejadian sebelumnya. Data mining mengelola data yang sudah ada sehingga menghasilkan informasi baru yang diharapkan membantu dalam data baru (Ain Banyal, Doktor Husni Ingratubun Papua and Raya Abepura Kotaraja Jayapura Papua, 2023). Salah satu metode data mining yang cocok untuk penelitian ini yaitu klasifikasi. Klasifikasi merupakan proses data mining yang bertujuan untuk menemukan pola yang berharga dari data yang berukuran relatif besar sampai sangat besar (Mir'aatunnas Pratiwi *et al.*, 2022). Klasifikasi merupakan sebuah proses dalam menempatkan suatu objek ke dalam suatu kategori yang telah ditentukan (Badriah *et al.*, 2021). Metode ini memiliki beberapa algoritma, salah satunya algoritma C4.5 dengan teknik pohon keputusan (*Decision Tree*). Algoritma data mining C4.5 adalah salah satu algoritma yang berguna untuk melakukan pengklasifikasian atau segmentasi atau bisa disebut mengelompokkan yang sifatnya prediktif (Islam *et al.*, 2022). Algoritma data mining C4.5 sendiri merupakan model pengembangan dari algoritma ID3, dimana dapat menangani atribut *kontinu* dan hilang (Satria Tambunan *et al.*, 2022). Metode pohon keputusan adalah metode yang pengklasifikasian dengan mengubah fakta yang besar menjadi pohon keputusan yang merepresentasikan

aturan. Dengan metode ini dirasa dapat digunakan untuk memecahkan masalah pada klasifikasi status risiko preeklamsia pada ibu hamil dengan algoritma C4.5.

Penelitian terkait pada data mining dalam prediksi resiko kesehatan ibu hamil dengan menggunakan metode *decision tree* pernah dilakukan oleh (Amalia et al., 2023), menghasilkan bahwa dengan memanfaatkan metode data mining menggunakan algoritma C4.5 dapat digunakan sebagai penunjang keputusan dapat melakukan tindakan pencegahan kondisi kesehatan yang tidak diinginkan yakni risiko tinggi dalam kehamilan yaitu kematian ibu dan anak.

Berdasarkan penelitian pada pemodelan klasifikasi untuk menentukan penyakit diabetes dengan faktor penyebab menggunakan *decision tree* c4.5 pada wanita oleh (Habibah et al., 2023). Hasil klasifikasi penderita diabetes dengan ciri-ciri yang terdapat pada database diabetes yaitu. kehamilan, glukosa, tekanan darah, ketebalan kulit, BMI insulin, riwayat keluarga diabetes, usia, dapat digunakan sebagai informasi dalam klasifikasi diabetes yang mendapatkan akurasi sebesar 76,67%.

METODOLOGI PENELITIAN

Data Penelitian

Data yang digunakan merupakan data Antenatal Care (ANC) semasa kehamilan ibu hamil di Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh Buaran Jakarta Timur. Data yang didapatkan berjumlah 300 *record*.

1. Data Primer

Data yang didapatkan secara langsung dari sumber yang bersangkutan, yaitu Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh, dengan 2 teknik yaitu :

a). Observasi

Teknik ini dilakukannya observasi langsung ke Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh, dengan mengamati kegiatan pemeriksaan pada ibu hamil (diluar pemeriksaan *vital*).

b). Wawancara

Pengumpulan data dengan teknik wawancara yang dilakukan dengan bu bidan Lilik Faiqoh. Data yang didapat yaitu informasi untuk penelitian dan data

antenatal care yang digunakan sebagai atribut.

2. Data Sekunder

Data sekunder didapatkan dari mengumpulkan informasi yang bersumber dari penelitian terdahulu, untuk mendapatkan data sekunder yaitu :

a). Studi Pustaka

Studi Pustaka merupakan suatu kegiatan, dimana kegiatan ini mempelajari dari berbagai sumber referensi seperti jurnal yang berkaitan dengan penelitian.

b). *Textbook*

Textbook adalah teknik pengumpulan informasi yang dilakukan dengan membaca buku, buku yang dibaca berkaitan dengan penelitian.

Atribut Data Awal

Pada tabel 1 merupakan data awal antenatal care dengan 17 atribut beserta keterangannya.

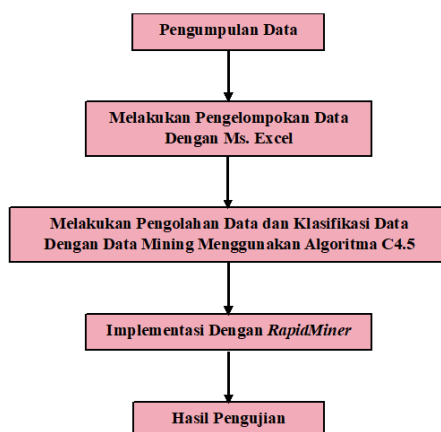
Tabel 1. Atribut Awal

Atribut	Keterangan
No.Reg	Nomor registrasi kunjungan pasien
Nama Pasien	Nama pasien istri dan suami
Usia Ibu	Usia ibu hamil
Alamat / No. Tlp	Alamat pasien dan nomor handphone pasien
Usia Kehamilan	Usia kehamilan ibu yang dihitung dengan minggu
Gpa	Gravida (jumlah kehamilan) Paritas (jumlah kelahiran) Abortus (jumlah keguguran)
Bb/Tb	Berat badan dan Tinggi badan ibu hamil
Td	Tekanan darah pada ibu hamil
Hb	Hemoglobin ibu hamil
Lila	Lingkar atas lengan

HIV	Reaktif atau non reaktif (NR) HIV
Hbsag	Reaktif atau non reaktif (NR) hepatitis B pada darah
Gol. Darah	Golongan darah ibu hamil
Sifilis	Reaktif atau non reaktif (NR) sifilis
Resti	Resiko tinggi pada ibu hamil
Pasien	Pasien baru atau lama
Terapi	Obat yang diberikan untuk ibu hamil

Penerapan Metodologi

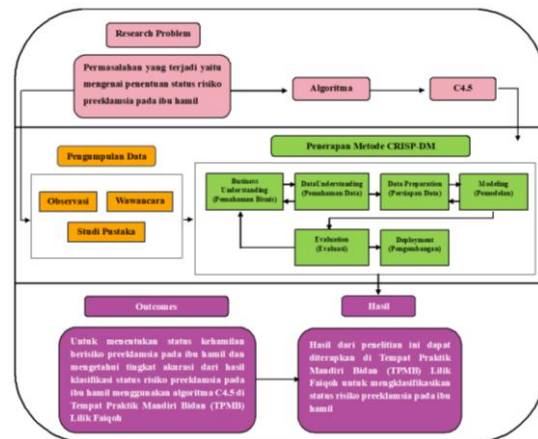
Penelitian ini menggunakan sebuah metode klasifikasi dengan algoritma C4.5. Pada gambar 1 diagram penerapan algoritma C4.5 berikut :



Gambar 1. Diagram Algoritma C4.5

Selanjutnya dilakukan penerapan metode *Cross Standar Process Industrial Data Mining* (CRISP-DM). CRISP-DM merupakan suatu standarisasi pemrosesan pada data mining yang telah dikembangkan dengan data yang ada akan melewati setiap fase terstruktur serta terdefinisi secara jelas dan efisien. Dimana banyak model prosedural yang telah melakukan standarisasi proses penambangan data dengan pendekatan CRISP-DM (Hasanah, Soim and Handayani, 2021). Dimana pada proses penelitian mengacu dengan tahapan yang ada pada CRISP-DM yaitu pemahaman bisnis (*business understanding*),

pemahaman data (*data understanding*), persiapan data (*data preparation*), pemodelan (*modelling*), evaluasi (*evaluation*) sampai pengembangan (*deployment*). Pada gambar 2 tahapan metodologi penelitian yang dapat dilihat sebagai berikut :



Gambar 2. Tahapan Metodologi

Rancangan Pengujian

Proses dalam penelitian ini terdapat 6 (enam) tahapan atau fase yang dapat dijelaskan sebagai berikut :

1. *Business Understanding* (Pemahaman Bisnis)

Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh merupakan pelayanan bidan kepada masyarakat di daerah Buaran Jakarta Timur. Tujuan dalam penelitian di Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh yaitu untuk mengetahui status risiko preeklamsia pada ibu hamil dan dapat melihat kategori status preeklamsia dengan pengklasifikasian algoritma C4.5.

a). Menentukan Tujuan Bisnis

Tujuan bisnis dilakukannya penelitian yaitu untuk menentukan status risiko preeklamsia ibu hamil berdasarkan kategori berisiko dan tidak berisiko. Penanganan preeklamsia membutuhkan perhatian lebih dan perlu adanya tindakan yang efektif apabila ibu berisiko preeklamsia.

b). Menilai Situasi

Pemeriksaan kehamilan di Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh salah satunya yaitu kunjungan antenatal dengan minimal 6 kunjungan selama kehamilan ibu. Dalam penentuan status preeklamsia ibu hamil sulit ditentukan dengan efektif,

dimana dilihat dari penyebab preeklamsia belum dipastikan serta faktor yang mempengaruhi preeklamsia.

2. Data Understanding (Pemahaman Data)

Berdasarkan data antenatal care (ANC) yang didapatkan untuk penelitian ini, tahap selanjutnya yaitu pemahaman data, dimana data dipahami terkait kebutuhan data dengan pencapaian tujuan untuk lebih efektif serta efisien. Data antenatal care (ANC) yang didapatkan merupakan dataset yang berjumlah 300 *record*.

a). Pengumpulan Data

Data awal dikumpulkan dengan melakukan pengumpulan data yang dibutuhkan untuk mendukung dalam melakukan pemahaman data. Data yang digunakan pada penelitian ini yaitu dataset antenatal care Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh pada Januari 2023 sampai April 2024 sebanyak 300 *record*.

b). Pendeskripsian Data

Pada tahap pendeskripsian dilakukan dengan pemahaman mengenai data antenatal care yang mendeskripsikan data dari Januari 2023 sampai April 2024 yang diperoleh dari total 300 *record* dan memiliki 17 atribut. Atribut dataset terdiri dari atribut nomor registrasi, nama pasien (istri dan suami), usia ibu, alamat pasien, usia kehamilan, gravida paritas abortus (Gpa), berat badan/tinggi badan, tekanan darah, lingkaran lengan atas, hemoglobin, HIV, Hbsag, golongan darah, sifilis, resiko tinggi, pasien, dan terapi. Dari semua keseluruhan atribut tidak semua digunakan tetapi hanya 6 atribut yang dapat digunakan.

c). Evaluasi Pemilihan Data

Sebelum proses pengolahan data, tahap evaluasi dalam pemilihan data dilakukan dengan tujuan agar dalam proses pengolahan data lebih mudah pelaksanaannya.

d). Pemilihan Atribut

Atribut pada dataset antenatal care tidak seluruhnya digunakan dalam pengolahan data. Penyesuaian atribut dilakukan untuk fokus terhadap penelitian. Atribut yang digunakan untuk proses data *mining* antara lain usia ibu, usia kehamilan,

paritas, berat badan/tinggi badan, tekanan darah, dan resiko tinggi.

3. Data Preparation (Persiapan Data)

Tahap ini mempersiapkan data dengan meliputi dataset antenatal care, dimana data tersebut akan diolah dengan alat pemodelan. Dataset antenatal care sebelumnya melalui proses data mentah awal sampai nantinya dilakukan proses data *mining*.

a). Seleksi Data

Dataset mentah pada penelitian ini menggunakan atribut usia ibu, usia kehamilan, paritas, berat badan/tinggi badan, tekanan darah, dan resiko tinggi di Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh. Berikut atribut yang digunakan pada penelitian ini beserta keterangannya, pada tabel 2 yaitu :

Tabel 2. Atribut yang digunakan

Atribut	Keterangan
Usia Ibu	Usia ibu hamil dalam tahun
Usia Kehamilan	Usia kehamilan ibu dalam minggu
Paritas	Paritas / Kelahiran bayi hidup
BB/TB	Berat badan / tinggi badan ibu
TD	Tekanan darah ibu hamil
Resti	Resiko tinggi pada ibu hamil

b). Pemrosesan Data Mentah

Tahap selanjutnya ini, merupakan tahapan untuk memastikan data antenatal care tersebut efektif untuk diolah. Setelah melakukan pengecekan data satu per satu serta memastikan tidak ada data yang bermasalah sesudah pembersihan data. Kemudian dilakukan pembersihan atribut dengan menghapus atribut yang tidak diperlukan lagi pada tahap pembuatan model klasifikasi. Dan setelahnya dilakukannya data *transformation* yaitu dengan mengubah data nominal menjadi sebuah data numerik serta menormalisasikan atribut tertentu.

4. *Modelling* (Pemodelan)

Dalam tahap pemodelan pada metode CRISP-DM di penelitian ini dilakukannya pemodelan pada data yang ada. Tahap pemodelan merupakan tahapan yang secara langsung menggunakan teknik data *mining*, dimana menentukan algoritma C4.5 yang digunakan. Pada tahap ini juga perlu adanya pembentukan simpul akar berdasarkan nilai atribut dengan nilai *gain* tertinggi dan pembentukan cabang tiap nilai berdasarkan *entropy*.

5. *Evaluation* (Evaluasi)

Tahapan evaluasi pada metode CRISP-DM ini dilakukan analisa atau pengukuran ketepatan terhadap pemodelan yang telah dilakukan. Tahap evaluasi dilakukan dengan adanya validasi, dimana pada pengukuran akurasi dari pemodelan algoritma yang digunakan dengan memanfaatkan *tools rapidminer*. Evaluasi dilakukan bertujuan untuk mengetahui pemodelan yang dilakukan pada *tools* apakah sudah sesuai dan tepat pada penerapan kasus penelitian ini. Hasil akhir pada evaluasi untuk menentukan langkah selanjutnya dalam kelanjutan atau pengulangan pada kesesuaian penelitian ini.

6. *Deployment* (Pengembangan)

Tahap terakhir pada metode CRISP-DM yaitu pengembangan, dimana tahap ini digunakan untuk pengembangan atau penyebaran hasil penelitian yang sudah dilakukan. Hasil penelitian berupa laporan ataupun presentasi dari pengetahuan yang sudah didapatkan dari pemodelan dan evaluasi data *mining*.

HASIL DAN PEMBAHASAN

Implementasi Dan Pengujian

Adapun langkah – langkah dalam pengerjaan algoritma data *mining* C4.5 membentuk pohon keputusan, antara lain :

1. Memilih atribut sebagai akar atau menentukan *root node*-nya
2. Membuat cabang untuk masing – masing nilai atribut
3. Membagi kasus dalam cabang untuk mencari *node* berikutnya
4. Ulangi proses untuk masing – masing cabang sampai kasus pada cabang memiliki kelas yang sama atau dengan kata lain setelah simpul menghasilkan solusi

Untuk penyelesaian kasus menggunakan algoritma data *mining* C4.5 perlu menghitung

elemen yang diketahui yaitu (Algoritma *et al.*, 2023) :

1. Gain

Memilih atribut sebagai akar harus menggunakan Gain yang didasarkan pada nilai Gain tertinggi dari setiap perhitungan atribut yang ada. Gain (S, A) merupakan hasil dari atribut A *relative* terhadap output S. Hasil yang diterima dari informasi output data atau *variable dependent* S yang telah dijadikan satu oleh atribut A dinotasikan dengan Gain (S, A).

Rumus :

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

Keterangan :

S : Himpunan Kasus

A : Atribut Kasus

n : Jumlah partisi atribut A

|S_i|: Jumlah Kasus pada partisi ke – i

|S| : Jumlah kasus dalam S

2. Entropy

Entropy berguna sebagai alat mengukur ketidakakasian S.

Rumus :

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

Keterangan :

S : Himpunan Kasus

n : Jumlah partisi kasus S

P_i : Proporsi dari S_i terhadap S

Hasil yang didapat pada perhitungan mencari nilai gain tertinggi dan entropy, dapat dilihat pada tabel 3 :

Tabel 3. Hitungan Entropy dan Gain

	KATEGORI	JUMLAH KASUS (S)	JUMLAH BERISIKO (S1)	JUMLAH TIDAK BERISIKO (S2)	ENTROPHY	GAIN
TOTAL		300	147	153	0,9997114	
USIA IBU					0,8994837	
	< 20 TAHUN	6	2	4	0,9182958	0,100227693
	20 - 35 TAHUN	183	64	119	0,9338238	
	> 35 TAHUN	111	81	30	0,8418522	
USIA KEHAMILAN					0,7694106	
	< 20 MINGGU	101	13	88	0,5538976	0,230300865
	20 - 38 MINGGU	188	123	65	0,9302102	
	> 38 MINGGU	11	11	0	0	
PARITAS					0,9614645	
	NOL	109	42	67	0,9617136	0,038246942
	SATU KALI	100	46	54	0,9953784	
	DUA KALI	66	41	25	0,9571814	
	TIGA KALI	24	17	7	0,8708645	
	EMPAT KALI	1	1	0	0	
BB/TB					0,9980677	
	NORMAL	262	126	136	0,9989489	0,001643705
	TIDAK IDEAL	38	21	17	0,9919924	
					2,989009	
TD					0,2880387	
	TINGGI	155	143	12	0,3930202	0,711672766
	NORMAL	124	4	120	0,2055925	
	RENDAH	21	0	21	0	
RESTI					0,9680738	
	TIDAK ADA RIWAYAT	281	133	148	0,9979435	0,031637662
	CAESAR	10	5	5	1	
	OBESITAS	8	8	0	0	
	GULA	1	1	0	0	

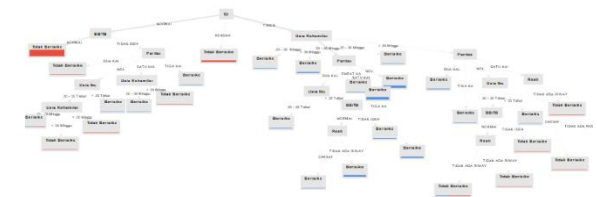
$$Entropy (S2) = \left(-\frac{4}{124} \times \log_2 \left(\frac{4}{124} \right) \right) + \left(-\frac{120}{124} \times \log_2 \left(\frac{120}{124} \right) \right) = 0,2055925$$

$$Entropy (S3) = \left(-\frac{0}{21} \times \log_2 \left(\frac{0}{21} \right) \right) + \left(-\frac{21}{21} \times \log_2 \left(\frac{21}{21} \right) \right) = 0$$

$$Gain (S, A) = 0,9997114 \times \left(\frac{155}{300} \times 0,3930202 \right) - \left(\frac{124}{300} \times 0,2055925 \right) - \left(\frac{21}{300} \times 0 \right) = 0,711672766$$

Hasil Pengujian

Pada gambar 4 ditampilkan hasil pohon keputusan yang telah terbentuk dari Rapidminer.



Gambar 4. Pohon Keputusan

Hasil perhitungan gain tertinggi didapatkan pada atribut tekanan darah (TD) dengan perhitungan sebagai berikut :

Entropy Total

$$= \left(-\frac{147}{300} \times \log_2 \left(\frac{147}{300} \right) \right) + \left(-\frac{153}{300} \times \log_2 \left(\frac{153}{300} \right) \right) = 0,9997114$$

Entropy (S1)

$$= \left(-\frac{143}{155} \times \log_2 \left(\frac{143}{155} \right) \right) + \left(-\frac{12}{155} \times \log_2 \left(\frac{12}{155} \right) \right) = 0,3930202$$

Untuk hasil dari *performance* berupa akurasi, kappa dan AUC dapat dilihat pada gambar berikut, yaitu didapatkan hasil :

1. Nilai *Accuracy* sebesar 95.00% +/- 5.03% (*micro average* : 95.00%)

accuracy: 95.00% +/- 5.03% (micro average: 95.00%)

	true Berisiko	true Tidak Berisiko	class precision
pred. Berisiko	138	6	95.83%
pred. Tidak Berisiko	9	147	94.23%
class recall	93.88%	96.08%	

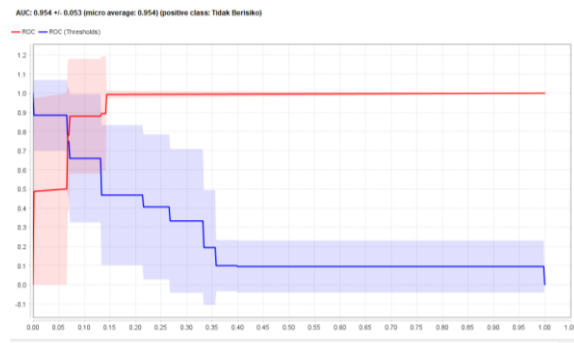
Gambar 5. Nilai Akurasi

2. Nilai *Kappa* sebesar 0.900 +/- 0.101 (*micro average* : 0.900)

kappa: 0.900 +/- 0.101 (micro average: 0.900)			
	True Berisiko	True Tidak Berisiko	class precision
pred Berisiko	138	6	95.63%
pred Tidak Berisiko	9	147	94.23%
class recall	93.88%	96.08%	

Gambar 6. Nilai Kappa

3. Nilai AUC sebesar 0.954 +/- 0.053 (micro average : 0.954) (positive class: Tidak Berisiko)



Gambar 7. Nilai AUC

Hasil dari *PerformanceVector* yang di deskripsikan pada gambar 8, yaitu :

PerformanceVector

```

PerformanceVector:
accuracy: 95.00% +/- 5.03% (micro average: 95.00%)
ConfusionMatrix:
True: Berisiko    Tidak Berisiko
Berisiko: 138    6
Tidak Berisiko: 9    147
kappa: 0.900 +/- 0.101 (micro average: 0.900)
ConfusionMatrix:
True: Berisiko    Tidak Berisiko
Berisiko: 138    6
Tidak Berisiko: 9    147
AUC: 0.954 +/- 0.053 (micro average: 0.954) (positive class: Tidak Berisiko)
  
```

Gambar 8. Deskripsi *PerformanceVector*

KESIMPULAN

Berdasarkan hasil penelitian dalam penerapan data mining dengan algoritma C4.5 yang dilakukan di Tempat Praktik Mandiri Bidan (TPMB) Lilik Faiqoh, maka dapat ditarik kesimpulan yaitu, analisa data menggunakan algoritma C4.5 dalam menentukan faktor – faktor apa saja yang berpengaruh dalam status risiko preeklamsia ibu hamil dan dapat melihat kategori status risiko preeklamsia. Didapatkan bahwa gain tertinggi ada pada perhitungan tekanan darah (TD).

Data yang digunakan sebanyak 300 record, yang kemudian dipilih dengan 6 yaitu usia ibu, usia kehamilan, paritas, BB/TB, tekanan darah dan resiko tinggi. Selanjutnya dataset yang ada pada *Microsoft Excel*

dikoneksikan ke aplikasi *rapidminer* untuk diolah dengan algoritma C4.5 yang mendapatkan hasil akurasi sebesar 95.00%.

DAFTAR PUSTAKA

Ain Banyal, N., Doktor Husni Ingratubun Papua, U. and Raya Abepura Kotaraja Jayapura Papua, J. (2023) ‘IMPLEMENTASI DATA MINING UNTUK PREDIKSI PENYAKIT MALARIA TROPIKA DENGAN ALGORITMA C4.5’, *Jurnal Ilmiah MATRIK*, 25(3), p. 2023.

Algoritma, P.C. *et al.* (2023) ‘Penerapan Algoritma C.45 Pada Klasifikasi Status Gizi Balita di Posyandu Desa Sukalilah Cibatu Kabupaten Garut Jawa Barat’, *Jurnal Sains dan Teknologi*, 5(1), pp. 177–182. Available at: <https://doi.org/10.55338/saintek.v5i1.1375>.

Amalia, H. *et al.* (2022) ‘Prediksi Resiko Kesehatan Ibu Hamil Dengan Menggunakan Metode Decision Tree’, *JURNAL SWABUMI*, 11(1), p. 2023.

Badriah, S. *et al.* (2021) ‘JIP (Jurnal Informatika Polinema) Klasifikasi Algoritma C4.5 dalam Menentukan Penerima Bantuan Covid-19 (Studi Kasus: Desa di Karawang)’.

Habibah, N.N. *et al.* (2023) ‘Pemodelan Klasifikasi Untuk Menentukan Penyakit Diabetes dengan Faktor Penyebab Menggunakan Decision Tree C4.5 Pada Wanita’, *Jurnal Sistem Komputer dan Informatika (JSON)*, 4(4), p. 654. Available at: <https://doi.org/10.30865/json.v4i4.6202>.

Hamzah, S.T.R. *et al.* (2021) ‘Antenatal care parameters that are the risk factors in the event of preeclampsia in primigravida’, *Gaceta Sanitaria*, 35, pp. S263–S267. Available at: <https://doi.org/10.1016/j.gaceta.2021.10.073>.

- Hasanah, M.A., Soim, S. and Handayani, A.S. (2021) *Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir, Journal of Applied Informatics and Computing (JAIC)*. Available at: <http://jurnal.polibatam.ac.id/index.php/JAIC>.
- Haslan, H. and Trisutrisno, I. (2022) 'Dampak Kejadian Preeklamsia dalam Kehamilan Terhadap Pertumbuhan Janin Intrauterine', *Jurnal Ilmiah Kesehatan Sandi Husada*, pp. 445–454. Available at: <https://doi.org/10.35816/jiskh.v11i2.810>.
- Islam, H.I. et al. (2022) 'Penerapan Algoritma C4.5 dalam Klasifikasi Status Gizi Balita', *Jurnal Ilmiah Wahana Pendidikan*, 8(10), pp. 116–125. Available at: <https://doi.org/10.5281/zenodo.6791722>.
- Kemkes. *Pentingnya Pemeriksaan Kehamilan di Fasilitas Kesehatan atau Puskesmas*. ayosehat.kemkes.go.id. Diakses : 7 Juli 2024, jam 14.20.
- Kemkes. *Peringatan Hari Preeklamsia Sedunia 2021*. ayosehat.kemkes.go.id. Diakses : 7 Juli 2024, jam 14.27.
- Kemkes. *Mengenal Preeklamsia*. yankes.kemkes.go.id. Diakses : 7 Juli 2024, jam 14.35.
- Mir'aatunnas Pratiwi, U. et al. (2022) 'KLASIFIKASI FAKTOR YANG BERPENGARUH DALAM KEHAMILAN TIDAK DIINGINKAN MENGGUNAKAN METODE ALGORITMA DECISION TREE', 3(2). Available at: <https://doi.org/10.46306/lb.v3i2>.
- Nurul Selviana Dewi, I. et al. (2024) *FAKTOR-FAKTOR YANG BERHUBUNGAN DENGAN KEJADIAN REEKLAMPSIA*. Available at: <http://journal.stikeskendal.ac.id/index.php/PSKM>.
- Retnaningtyas, E., Manila No, J. and Kota Kediri Jawa Timur, S. (2021) *PREEKLAMPSI & ASUHAN KEBIDANAN PADA PREEKLAMPSI STRADA PRESS*.
- Satria Tambunan, H. et al. (2022) *Penerapan Data Mining Klasifikasi Gizi Bayi Dengan Algoritma Decision Tree C4.5*. Available at: <https://ejurnal.pdsi.or.id/index.php/zahra/index>.

SENGAJA DIKOSONGKAN