

Klasifikasi Hasil Cardiotocography (CTG) Ibu Hamil untuk Memprediksi Kesehatan Janin

Aprindita Dwi Monica¹, Sulastri²

¹Teknik Informatika, Universitas Stikubank, Jl. Tri Lomba Juang, Mugassari, Kec. Semarang Selatan, Kota Semarang, Jawa Tengah, Indonesia, 50241

e-mail: ¹aprinditadwimonica@mhs.unisbank.ac.id, ²sulastri@edu.unisbank.ac.id

Submitted Date: July 03rd, 2023

Reviewed Date: July 12th, 2023

Revised Date: July 28th, 2023

Accepted Date: July 29th, 2023

Abstract

A future mother must want the fetus in her womb to be in a healthy condition. A healthy fetus is one that has optimal growth and has sufficient nutrients at the time it is in the womb. The risk of miscarriage and maternal and fetal deaths can be reduced by monitoring fetal health and staying alert when necessary. CTG is carried out in cases where there is a risk of pregnancy and a relatively worrying birth. It will be easier to make further decisions to reduce health risks at birth if the CTG examination shows a good fetal condition. The study aims to study the comparison of different classification algorithms used in determining the exact values of the fetal health datasets used through Knowledge Discovery in Databases (KDD). The researchers used a data collection of 2126 data points from Kaggle's website that had 22 variables divided into three categories Normal, Suspect, and Pathological. *Prolonged decelerations*, *abnormal short-term variability*, and the *percentage of time with abnormally long-term variability* are the most significant variables. The results of the study were obtained by dividing the data set into training and trial data, which were then divided into three trials. The results showed that the KNN algorithm had the best accuracy value of 91% in the second trial, the SVM algorithm had the most accurate value of 87% in the first and second trials, the Logistic Regression algorithm had the Best Accuracy Value of 84% in the Second Trial, the Naive Bayes algorithm had the Best Accuracy value of 84%, and the Decision Tree Algorithm had 89% in the First and Second Trials.

Keywords: Classification; Cardiotocography; Data Mining

Abstrak

Seorang calon ibu pasti menginginkan janin dalam kandungannya dalam kondisi sehat. Janin yang sehat adalah janin yang memiliki pertumbuhan yang optimal dan memiliki nutrisi yang cukup pada saat di dalam kandungan. Resiko keguguran dan kematian ibu dan janin dapat dikurangi dengan memantau kesehatan janin dan tetap waspada saat diperlukan. CTG dilakukan dalam kasus di mana ada risiko kehamilan dan kelahiran bayi yang cukup mengkhawatirkan. Akan lebih mudah untuk membuat keputusan lanjut untuk mengurangi risiko kesehatan pada saat melahirkan jika pemeriksaan CTG menunjukkan kondisi janin baik. Penelitian ini bertujuan untuk mempelajari perbandingan algoritma klasifikasi yang berbeda yang digunakan dalam menentukan nilai yang tepat dari dataset kesehatan janin yang digunakan melalui pendekatan Knowledge Discovery in Database (KDD). Peneliti menggunakan kumpulan data 2126 data dari website Kaggle yang memiliki 22 variabel yang dibagi menjadi tiga kategori Normal, Suspect, dan Pathological. *Prolonged decelerations*, *abnormal short term variability*, dan *percentage of time with abnormal long term variability* adalah variabel yang paling signifikan. Hasil penelitian ini didapatkan dengan membagi dataset menjadi data pelatihan dan data uji coba, kemudian dibagi menjadi tiga kali uji coba. Hasil menunjukkan bahwa algoritma KNN memiliki nilai akurasi terbaik sebesar 91% pada uji coba kedua, Algoritma SVM memiliki nilai akurasi terbaik sebesar 87% pada uji coba pertama dan kedua, Algoritma Logistic Regression memiliki nilai akurasi terbaik sebesar 84% pada uji coba kedua, Algoritma

Naive Bayes memiliki nilai akurasi terbaik sebesar 84% pada uji coba pertama, Algoritma Decision Tree memiliki nilai akurasi terbaik sebesar 89% pada uji coba pertama dan kedua.

Keywords: Klasifikasi; Cardiotocography; Data Mining

1 Pendahuluan

Janin merupakan bakal manusia yang mengalami fase pertumbuhan di dalam rahim ibunya. Janin yang sehat diperoleh dari ibu yang memenuhi cukup nutrisi di dalam rahim. Setiap ibu hamil pasti ingin memiliki anak yang terlahir dengan kondisi sehat. Kondisi janin tidak sehat atau memiliki kelainan, dapat mengakibatkan ibu mengalami keguguran, koma, bahkan kematian pasca melahirkan (Sulihati et al., 2022).

Pada tahun 2015, World Health Organization (WHO) memperkirakan bahwa sekitar 3,82-22,14 juta orang di seluruh dunia mengalami keguguran. Menurut POGI (Persatuan Obstetri Ginekologi Indonesia), ketidaktahuan sang ibu tentang kondisi dan kesehatan janinnya menyebabkan 15% hingga 20% dari semua ibu hamil mengalami keguguran. Ketidaktahuan ibu dapat menyebabkan komplikasi selama kehamilan, persalinan, dan kelahiran bayi. Setiap ibu hamil harus memperhatikan pentingnya memantau kesehatan janinnya setiap saat. Akibat semakin banyaknya ibu hamil yang mengalami keguguran, dilakukan pencegahan keguguran ibu hamil dengan melakukan pengumpulan data kesehatan janin menggunakan alat Cardiotocography (CTG).

Cardiotocography (CTG) adalah alat untuk memantau denyut jantung dan kontraksi rahim pada bayi saat di dalam rahim. Hasil dari CTG digunakan untuk melihat bagaimana denyut jantung bayi dipengaruhi oleh kontraksi. CTG dilakukan apabila ada resiko bahaya pada kehamilan dan melahirkan pada bayi dan ibu yang cukup mengkhawatirkan (Santoso & Musa, 2021). Jika pemeriksaan CTG menunjukkan kondisi janin baik, maka akan lebih mudah untuk mengambil keputusan lebih lanjut untuk meminimalkan risiko kesehatan pada saat melahirkan. Pemantauan dengan CTG juga memiliki beberapa hasil tidak konsisten, terutama pada kehamilan dengan resiko rendah. Hasil dari CTG ini bergantung pada hasil denyut jantung janin apakah janin dalam kondisi baik atau tidak (Arif et al., 2020).

Namun, permasalahan hasil dari CTG perlu pemeriksaan kembali oleh dokter atau tenaga kesehatan untuk meminimalkan kesalahan pemeriksaan. Dengan memanfaatkan software, hasil pemeriksaan CTG akan dievaluasi ulang untuk merekomendasikan pemeriksaan lebih lanjut terkait kondisi janin. Berdasarkan permasalahan tersebut, penelitian ini menggunakan teknik data mining untuk memeriksa dan mengolah dataset CTG.

Penelitian terkait kesehatan janin oleh (Rahmayanti et al., 2021) membandingkan 7 algoritma yang terdiri dari Artificial Neural Network (ANN), Long Short Term Memory (LSTM), XG Boost (XGB), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Light GBM (LGBM), dan Random Forest menggunakan 3 skenario menunjukkan algoritma bekerja sangat akurat sebesar 89-99% yaitu algoritma XGB, SVM, KNN, LGBM, dan Random Forest. LGBM menjadi paling baik dari kelima algoritma dengan 3 skenario.

Menurut (Arif et al., 2020) pohon keputusan melakukan penyeleksian dari dataset cardiotocography dengan fitur seleksi dan mengklasifikasikan risiko kehamilan dengan tingkat akurasi yang tinggi. Fitur seleksi terbukti cukup efektif dan dapat meningkatkan akurasi hasil hingga 98,7%. Tahapan penelitian tersebut terdiri dari menghitung angka pendapatan informasi tertinggi untuk menentukan akar dan kemudian menghitung angka pendapatan informasi tertinggi untuk menentukan cabang. Simpul akan dihasilkan oleh node cabang yang diklasifikasikan pada kelas tertentu, dan klasifikasi benar atau salah akan dilakukan menggunakan Confusion Matrix.

Penelitian ini bertujuan untuk mengetahui prediksi kesehatan janin dengan metode klasifikasi menggunakan algoritma machine learning seperti K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Logistic Regression, Naive Bayes, dan Decision Tree. Penggunaan algoritma machine learning akan dilakukan untuk perbandingan antar algoritma

untuk mengetahui kinerja hasil algoritma dengan dataset yang diambil dari website Kaggle.

2 Tinjauan Literatur

Penelitian ini membutuhkan hasil dari penelitian sebelumnya yang relevan dengan penelitian yang dilakukan saat ini. Peneliti pertama menggunakan Algoritma C5.0 untuk menganalisis kesehatan janin setiap ibu hamil. Penelitian ini menggunakan dataset cardiotography yang berkaitan dengan kondisi janin yang terdiri dari 2.126 daftar, yang masing-masing memiliki 22 kolom atribut. Klasifikasi terdiri dari normal, suspect, dan pathological. Confusion matrix juga digunakan sebagai pembanding akurasi untuk perhitungan dengan entropy, gain ratio, split information, dan information gain. Data dibagi menjadi data pelatihan, data uji, dan data pengujian dengan perbandingan 60:40, 70:30, 80:20, dan 90:10. Hasil penelitian menunjukkan bahwa perbandingan 90%:10% menghasilkan akurasi sebesar 93,40% dengan jumlah aturan sebanyak 257, perbandingan 80%:20% menghasilkan akurasi sebesar 91,29% dengan jumlah aturan sebanyak 239, dan perbandingan 60%:40 menghasilkan akurasi sebesar 88,12% dengan jumlah aturan sebanyak 204. Hasil menunjukkan bahwa data pelatihan yang lebih akurat sebanding dengan lebih banyak aturan yang membantu dalam pengambilan keputusan (Santoso & Musa, 2021).

Peneliti kedua melakukan penelitian bahwa stroke adalah penyakit ketiga yang paling umum di Indonesia, setelah kanker dan penyakit jantung. Karena kurangnya tenaga medis di Indonesia, banyak orang tidak menyadari penyakit stroke dan sulit bagi masyarakat untuk mendeteksi secara dini. Penelitian ini bertujuan untuk menggunakan pengujian Algoritma Support Vector Machine (SVM) dan pengklasifikasian menggunakan Matriks Konflik. Hasilnya menunjukkan bahwa Algoritma SVM dengan *Relief-f* menggunakan Kernel Linear yang dibagi menjadi 2398 data pelatihan dan 1028 data pengujian mendapatkan akurasi 100% (Amelia et al., 2022).

Peneliti ketiga melakukan penelitian uji performa Algoritma Naive Bayes untuk memprediksi masa studi ribuan siswa informatika di Universitas AMIKOM Yogyakarta. Karena

hanya 50% dari ribuan siswa lulus tepat waktu. Studi ini menggunakan data alumni tahun 2012 dan 2013, yang memiliki 150 data yang menunjukkan kelulusan tepat waktu dan tidak tepat waktu. Dalam penelitian ini, Algoritma Naive Bayes digunakan untuk membuat sistem klasifikasi dengan menggunakan metode validasi cross-fold dan matriks konfusi. Hasil penelitian menunjukkan bahwa Algoritma Naive Bayes memiliki nilai kaurasi sebesar 68%, nilai ketepatan sebesar 61,3%, nilai recall sebesar 65,3%, dan nilai f1-score sebesar 61%. (Saputro & Sari, 2020).

Peneliti keempat melakukan penelitian bahwa diabetes yang tidak dikontrol dengan baik dicirikan oleh kadar gula dalam darah yang tinggi dan dapat menjadi kronis. Tujuan dari penelitian ini adalah untuk menguji kinerja metode klasifikasi yang menggunakan Algoritma K-Nearest Neighbor (KNN). Metode ini mengklasifikasikan objek yang paling dekat dengannya. Hasil penelitian disesuaikan. Nilai tertinggi untuk akurasi sebesar 39% pada K=3, nilai tertinggi untuk presisi sebesar 65% pada K=3 dan K=5, nilai tertinggi untuk recall sebesar 36% pada K=3, dan nilai F-Measure tertinggi sebesar 46% pada K=3 (Argina, 2020).

Peneliti kelima menemukan bahwa penyakit jantung adalah penyakit tidak menular (PTM) yang paling umum terjadi pada pria dewasa di bawah usia 60 tahun. Studi ini bertujuan untuk membandingkan algoritma klasifikasi seperti Naive Bayes, Support Vector Machine (SVM), C.45, Logistic Regression, dan Back Propagation ketika menggunakan dataset data statlog penyakit jantung yang terdiri dari 270 rekaman data. Selanjutnya, algoritma Support Vector Machine (SVM) diuji melalui validasi berulang untuk mengevaluasi kinerja mereka dalam hal recall, akurasi, dan presisi. Hasil penelitian menunjukkan bahwa algoritma SVM memiliki nilai recall tertinggi sebesar 94,67%, nilai akurasi tertinggi sebesar 80,087 persen, dan nilai presisi tertinggi sebesar 86,16%. (Putra & Rini, 2019).

Peneliti keenam melakukan penelitian tentang penyakit katarak pada mata. Mata adalah panca indera utama manusia yang memungkinkan mereka melihat keindahan dunia, melakukan tugas, dan berinteraksi dengan orang-orang di sekitar mereka. Salah satu kondisi mata

yang disebut penyakit katarak menyebabkan penderitanya mengalami penglihatan seperti tertutup air karena lensa yang keruh. Penyakit katarak dapat menyebabkan penyakit kronis jika terus dibiarkan. Dengan menggunakan Algoritma K-Nearest Neighbor (KNN), yang mencari tetangga terdekat, penelitian ini bertujuan untuk menentukan jenis penyakit katarak yang berbeda. Studi ini mengklasifikasikan berbagai jenis katarak berdasarkan gejala pasien yang dirawat di Rumah Sakit Anutapura Palu dari Januari hingga Maret 2018. Hasil penelitian ini menunjukkan bahwa Algoritma KNN memiliki akurasi 91,76% dari 170 data (Safaat et al., 2020).

Peneliti ketujuh menyelidiki diabetes, penyakit yang tidak menular dan berlangsung lama yang berdampak pada kesehatan seseorang secara keseluruhan. Pada tahun 2021, Indonesia memiliki 19,5 juta penderita diabetes, menempati urutan kelima terbanyak di seluruh dunia dan keenam tertinggi dalam hal kematian. Dataset PIMA Diabetes India (PIDD) adalah dataset yang dapat digunakan untuk pembelajaran mesin. Penelitian ini bertujuan untuk mengevaluasi fitur-fitur penting dalam database PIMA India dengan menggunakan Algoritma K-Nearest Neighbor untuk mengklasifikasi. Penelitian ini menemukan bahwa algoritma KNN dengan nilai $K=22$ menghasilkan nilai akurasi tertinggi sebesar 83,12% ketika menemukan glucose, usia, insulin, tekanan darah, indeks massa tubuh, kehamilan, dan ketebalan kulit (Perdana et al., 2023).

Menurut hasil penelitian sebelumnya, dapat disimpulkan bahwa ada persamaan dan perbedaan antara penelitian yang sedang dilakukan dan penelitian sebelumnya. Penelitian terdahulu sama-sama membahas prediksi penyakit. Salah satu hal yang membedakan penelitian ini adalah metode dan algoritma yang akan digunakan dan salah satu penelitian memprediksi masa studi.

3 Metodologi

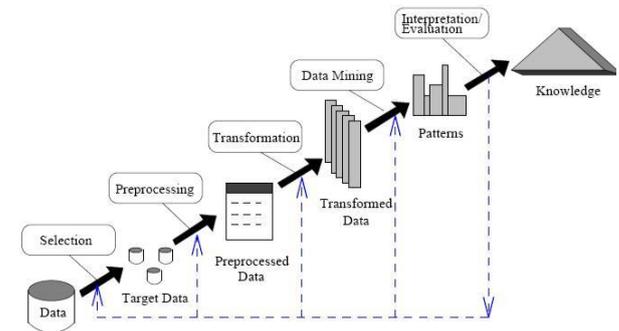
Penelitian ini menggunakan dataset terbuka yang diperoleh dari website <https://www.kaggle.com/> yang memberikan informasi tentang kesehatan janin untuk mencegah kematian ibu dan anak. Berdasarkan dari data yang diperoleh terdapat 3 kelas yang terdapat pada kolom *fetal_health* yaitu:

Table 1. Kelas Kesehatan Janin

No	Kelas
1	Normal
2	Suspect
3	Pathological

3.1 Tahap Penelitian

Penelitian ini memanfaatkan metode Knowledge Discovery in Database (KDD), yang terdiri dari lima tahapan, yaitu :



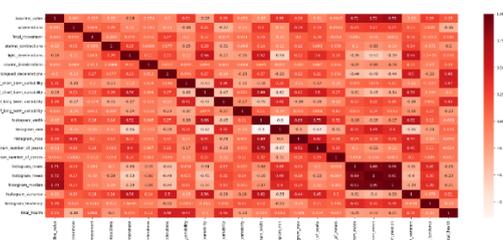
Gambar 1. Tahap Penelitian

3.1.1 Selection

Selection atau seleksi data adalah tahap sebelum penggalian informasi, seleksi data perlu dilakukan dari sekumpulan dataset. Hasil dari seleksi ini digunakan dalam pemrosesan data mining. Berikut ini adalah potongan dari dataset kesehatan janin :

13070	0.0	0.0	0.0	0.0	0.0	71.0	9.5	43.0	1.4	84.0	62.0	130.0	2.0	0.0	130.0	137.0	121.0	79.0	1.0	2.0	
13120	0.000	0.0	-0.000	0.000	0.0	0.0	27.0	2.1	0.0	10.0	100.0	66.0	100.0	6.0	1.0	141.0	136.0	140.0	12.0	0.0	1.0
13130	0.000	0.0	-0.000	0.000	0.0	0.0	16.0	2.1	0.0	13.0	100.0	66.0	100.0	5.0	1.0	141.0	135.0	139.0	13.0	0.0	1.0
13140	0.000	0.0	-0.000	0.000	0.0	0.0	18.0	2.4	0.0	10.0	107.0	71.0	170.0	13.0	0.0	137.0	139.0	137.0	13.0	1.0	1.0
13150	0.007	0.0	-0.008	0.0	0.0	0.0	18.0	2.4	0.0	10.0	107.0	71.0	170.0	9.0	0.0	137.0	139.0	139.0	13.0	1.0	1.0
13160	0.002	0.0	-0.011	0.000	0.0	-0.002	19.0	2.0	0.0	10.0	100.0	70.0	200.0	5.0	1.0	140.0	137.0	137.0	13.0	1.0	1.0
13170	0.002	0.0	-0.013	0.008	0.0	-0.003	20.0	4.3	0.0	0.0	100.0	50.0	200.0	6.0	1.0	137.0	139.0	139.0	13.0	0.0	0.0
13220	0.0	0.0	0.0	0.0	0.0	60.0	0.5	0.0	13.0	60.0	62.0	100.0	0.0	0.0	122.0	122.0	122.0	1.0	1.0	1.0	
13230	0.0	0.0	-0.002	0.0	0.0	64.0	0.5	0.0	13.0	60.0	62.0	100.0	0.0	0.0	122.0	122.0	122.0	1.0	1.0	1.0	
13240	0.0	0.0	-0.003	0.0	0.0	60.0	0.5	0.0	13.0	60.0	62.0	100.0	0.0	0.0	122.0	122.0	122.0	1.0	1.0	1.0	
13110	0.0	0.0	-0.001	0.001	0.0	0.0	64.0	1.9	0.0	27.0	100.0	56.0	100.0	2.0	0.0	100.0	140.0	151.0	9.0	1.0	2.0
10010	0.0	0.0	-0.001	0.001	0.0	0.0	64.0	1.9	0.0	27.0	100.0	56.0	100.0	2.0	0.0	100.0	140.0	151.0	9.0	1.0	2.0
1110	0.005	0.072	-0.008	0.003	0.0	0.0	38.0	1.4	0.0	12.0	60.0	86.0	150.0	5.0	0.0	105.0	134.0	137.0	7.0	1.0	1.0
11110	0.000	0.122	-0.006	0.001	0.0	0.0	38.0	1.4	0.0	11.0	67.0	77.0	180.0	2.0	0.0	100.0	132.0	141.0	10.0	1.0	1.0
10010	0.006	0.408	-0.004	0.005	0.0	0.001	21.0	2.3	0.0	7.0	107.0	47.0	174.0	7.0	0.0	145.0	135.0	133.0	7.0	0.0	1.0
10010	0.006	0.398	-0.004	0.005	0.0	0.001	21.0	2.3	0.0	8.0	107.0	47.0	174.0	7.0	0.0	145.0	135.0	133.0	7.0	0.0	1.0
10010	0.006	0.441	-0.005	0.005	0.0	0.0	24.0	2.1	0.0	0.0	115.0	53.0	178.0	5.0	0.0	143.0	138.0	139.0	7.0	1.0	1.0
11110	0.002	0.080	-0.001	0.000	0.0	0.002	18.0	2.4	0.0	13.0	107.0	67.0	174.0	5.0	0.0	104.0	120.0	120.0	4.0	0.0	2.0
10010	0.003	0.451	-0.006	0.006	0.0	0.001	21.0	1.9	0.0	8.0	90.0	60.0	160.0	6.0	0.0	110.0	124.0	129.0	16.0	1.0	1.0
10010	0.005	0.469	-0.005	0.004	0.0	0.001	21.0	1.7	0.0	7.0	112.0	65.0	177.0	6.0	1.0	110.0	120.0	120.0	17.0	0.0	2.0
12010	0.0	0.34	-0.006	0.002	0.0	-0.003	19.0	2.1	0.0	8.0	120.0	54.0	182.0	2.0	0.0	120.0	126.0	129.0	10.0	0.0	1.0
12010	0.005	0.425	-0.003	0.003	0.0	0.002	26.0	1.7	0.0	6.7	141.0	57.0	180.0	9.0	0.0	119.0	125.0	131.0	14.0	0.0	1.0
12010	0.0	0.336	-0.003	0.003	0.0	-0.003	16.0	2.3	0.0	6.0	140.0	54.0	180.0	11.0	1.0	119.0	125.0	131.0	14.0	0.0	1.0
12010	0.0	0.0	0.0	0.0	0.0	80.0	0.5	0.0	6.0	10.0	114.0	101.0	0.0	0.0	126.0	124.0	125.0	1.0	1.0	1.0	

Gambar 2. Dataset Cardioctophraphy



Gambar 3. Corelation Matrix

Pada penelitian ini variabel-variabel diatas tidak semua digunakan dalam proses data mining. Variabel yang akan digunakan yaitu :

Tabel 2. Variabel Kesehatan Janin

Variabel	Keterangan	Penjelasan
accelerations	Prediktor	Jumlah akselerasi per detik
prolongued_decelerations	Prediktor	Jumlah PD per detik
abnormal_short_term_variability	Prediktor	Persentase waktu dengan variabilitas jangka pendek abnormal
percentage_of_time_with_abnormal_long_term_variability	Prediktor	Persentase waktu dengan variabilitas jangka panjang abnormal
fetal_health	Target	Kesehatan Janin : 1 - Normal, 2 - Suspect, 3 - Pathologica l

3.1.2 Preprocessing

Preprocessing atau pemrosesan data adalah komponen penting dari metode Knowledge Discovery in Database (KDD) sebelum data mining dilakukan. Prapemrosesan data mencakup menghilangkan duplikat data, memeriksa data yang masih berubah atau tidak konsisten, dan memperbaiki kesalahan cetak. Dataset ini siap untuk digunakan dalam proses data mining karena memiliki 2126 daftar data dan 22 variabel dan tidak ditemukan kesalahan dalamnya.

3.1.3 Transformation

Untuk mendapatkan fitur yang berguna untuk proses data mining, transformasi atau transformasi data dilakukan. Proses ini sangat bergantung pada dataset yang akan digunakan untuk proses dan pola yang dihasilkan dari informasi data tersebut (Yuli Mardi, 2019). Pada penelitian ini, dataset yang digunakan tidak perlu diubah karena datasetnya sudah sesuai.

3.1.4 Data Mining

Data mining adalah proses pengumpulan data untuk menemukan pola tertentu (Antoni et al., 2023). Metode klasifikasi digunakan dalam metode Knowledge Discovery in Database (KDD), yang menggunakan algoritma K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Logistic Regression, Naive Bayes, dan Decision Tree. Python adalah bahasa pemrograman. Setelah perbandingan data pelatihan dan pengujian selama tiga percobaan, lima algoritma akan digunakan untuk menganalisis data yang akan diolah (Sulastri et al., 2020).

3.1.5 Interpretation/Evaluation

Interpretation/Evaluation adalah metode yang digunakan untuk mengidentifikasi pola-pola yang muncul selama proses data mining. Untuk membuat dataset yang telah diolah lebih mudah dipahami, diagram yang dibuat akan digunakan untuk membandingkan nilai akurasi, precision, recall, dan f1-score. Selanjutnya, evaluasi model klasifikasi menggunakan matriks kecacauan untuk menemukan algoritma terbaik.

3.2 Confusion Matrix

Pada penelitian ini hasil evaluasi data mining menggunakan *Confusion Matrix* untuk model klasifikasi. *Confusion Matrix* digunakan untuk mengevaluasi model klasifikasi menggunakan tabel matriks (Athalla et al., 2018). Juga digunakan untuk menguji coba dan mengetahui kinerja algoritma yang digunakan berdasarkan kumpulan data dengan membaginya ke dalam data pelatihan dan data ujicoba (Handayani et al., 2021). False positive, true negative, dan false negative termasuk dalam kelas positif dan negative Confusion Matrix.

Tabel 3. Confusion Matrix

Confusion Matrix		Prediksi	
		Positive	Negative
Aktual	Positive	TP (True Positive)	FN (False Negative)
	Negative	FP (False Positive)	TN (True Negative)

Berdasarkan tabel diatas TP (True Positive) merupakan data positive dan diprediksi positive, FP (False Positive) merupakan data negative dan diprediksi positive, FN (False Negative) merupakan data positive dan diprediksi negative, TN (True Negative) merupakan data negative dan diprediksi positive (Sulihati et al., 2022). Tingkat keberhasilan algoritma pada model klasifikasi dilihat dari perhitungan nilai akurasi, precision, recall, dan f1-score yaitu :

1. Akurasi
 Perhitungan akurasi digunakan untuk melihat hasil kerja algoritma pada model klasifikasi (Handayani et al., 2021). Persamaan akurasi :

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

2. Precision
 Perhitungan precision digunakan untuk melihat hasil kerja data positive dari hasil data yang diprediksi positive (Handayani et al., 2021). Persamaan precision :

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

3. Recall
 Perhitungan recall digunakan untuk melihat hasil kerja dari data yang diprediksi positive dari hasil data positive (Handayani et al., 2021). Persamaan recall :

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

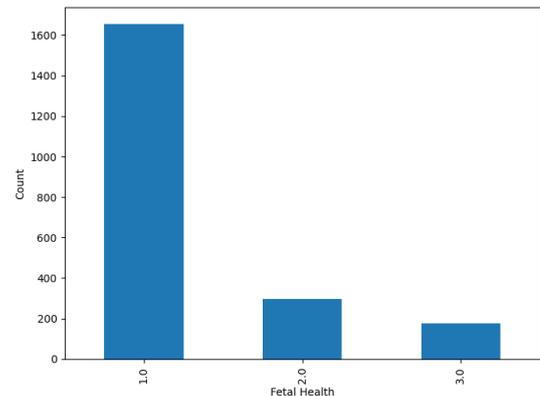
4. F1-Score
 Perhitungan f1-score digunakan untuk melihat hasil kerja rata-rata dari hasil perbandingan precision serta recall (Handayani et al., 2021). Persamaan f1-score :

$$F1 - Score = 2 * \frac{Recall * Precision}{Recall + Precision} \quad (4)$$

4 Hasil dan Pembahasan

Proses menganalisis dataset kesehatan janin yang terdiri dari tiga kelas: normal, suspect,

dan pathological. Kemudian dataset dianalisis menggunakan algoritma K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Logistic Regression, Naive Bayes, dan Decision Tree. Dataset yang dihasilkan terdiri dari 2126 record menggunakan metode Knowledge Discovery in Database (KDD).



Gambar 4. Jumlah Kelas Normal, Suspect, dan Pathological

Berdasarkan Gambar 4 jumlah kelas 1 (Normal) sebanyak 1655 record, kelas 2 (Suspect) sebanyak 295 record, dan kelas 3 (Pathological) sebanyak 176 record. Penelitian kali ini menggunakan 3 percobaan dengan pembagian data pelatihan dan data ujicoba sebesar 70:30, 80:20, dan 90:10.

4.1 K-Nearest Neighbors (KNN)

Pengujian dilakukan dengan algoritma K-Nearest Neighbor (KNN) pada dataset kesehatan janin yang berisi 2126 record. Data ini dibagi menjadi data pelatihan dan pengujian sebagai berikut :

Tabel 4. Pembagian Data Pelatihan dan Data Pengujian

Uji Coba	Jumlah Record		
	Probabilitas	Pelatihan	Pengujian
1	70:30	1488	628
2	80:20	1700	426
3	90:20	1913	213

4.1.1 Uji Coba 1 (Probabilitas 70:30)

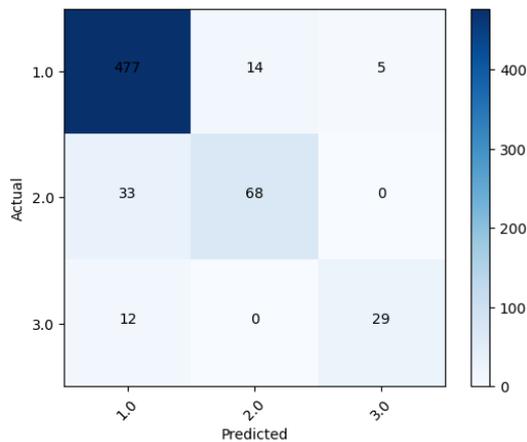
Dengan percobaan 1 menggunakan algoritma K-Nearest Neighbor (KNN), hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.91	0.96	0.94	496
2.0	0.83	0.67	0.74	101
3.0	0.85	0.71	0.77	41
accuracy			0.90	638
macro avg	0.87	0.78	0.82	638
weighted avg	0.90	0.90	0.90	638

Accuracy : 0.8996865203761756

Gambar 5. Klasifikasi KNN (70:30)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 70:30 menunjukkan hasil sebesar 90%, hasil precision sebesar 90%, hasil recall sebesar 90%, dan hasil f1-score sebesar 90%.



Gambar 6. Confusion Matrix KNN (70:30)

4.1.2 Uji Coba 2 (Probabilitas 80:20)

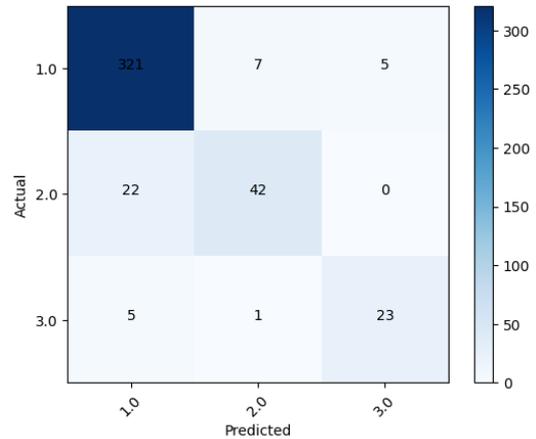
Dengan percobaan 2 menggunakan algoritma K-Nearest Neighbor (KNN), hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.92	0.96	0.94	333
2.0	0.84	0.66	0.74	64
3.0	0.82	0.79	0.81	29
accuracy			0.91	426
macro avg	0.86	0.80	0.83	426
weighted avg	0.90	0.91	0.90	426

Accuracy : 0.9061032863849765

Gambar 7. Klasifikasi KNN (80:20)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 80:20 menunjukkan hasil sebesar 91%, hasil precision sebesar 90%, hasil recall sebesar 91%, dan hasil f1-score sebesar 90%.



Gambar 8. Confusion Matrix KNN (80:20)

4.1.3 Uji Coba 3 (Probabilitas 90:10)

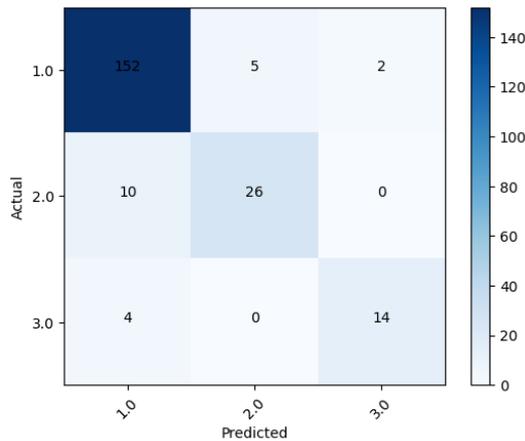
Dengan percobaan 3 menggunakan algoritma K-Nearest Neighbor (KNN), hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.92	0.96	0.94	159
2.0	0.84	0.72	0.78	36
3.0	0.88	0.78	0.82	18
accuracy			0.90	213
macro avg	0.88	0.82	0.85	213
weighted avg	0.90	0.90	0.90	213

Accuracy : 0.9014084507042254

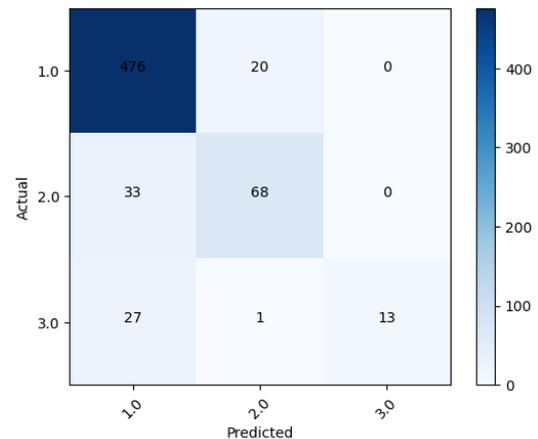
Gambar 9. Klasifikasi KNN (90:10)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 90:10 menunjukkan hasil sebesar 90%, hasil precision sebesar 90%, hasil recall sebesar 90%, dan hasil f1-score sebesar 90%.



Gambar 10. Confusion Matrix KNN (80:20)

sebesar 88%, hasil recall sebesar 87%, dan hasil f1-score sebesar 86%.



Gambar 12. Confusion Matrix SVM (70:30)

4.2 Support Vector Machine (SVM)

Pengujian dilakukan dengan algoritma Support Vector Machine (SVM) pada dataset kesehatan janin yang berisi 2126 record. Data ini dibagi menjadi data pelatihan dan pengujian sebagai berikut :

Tabel 5. Pembagian Data Pelatihan dan Data Pengujian

Uji Coba	Jumlah Record		
	Probabilitas	Pelatihan	Pengujian
1	70:30	1488	628
2	80:20	1700	426
3	90:20	1913	213

4.2.1 Uji Coba 1 (Probabilitas 70:30)

Dengan percobaan 1 menggunakan algoritma Support Vector Machine (SVM), hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.89	0.96	0.92	496
2.0	0.76	0.67	0.72	101
3.0	1.00	0.32	0.48	41
accuracy			0.87	638
macro avg	0.88	0.65	0.71	638
weighted avg	0.88	0.87	0.86	638

Accuracy : 0.8730407523510971

Gambar 11. Klasifikasi SVM (70:30)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 70:30 menunjukkan hasil sebesar 87%, hasil precision

4.2.2 Uji Coba 2 (Probabilitas 80:20)

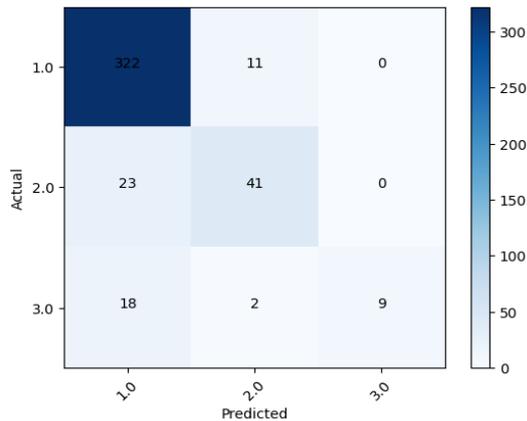
Dengan percobaan 2 menggunakan algoritma Support Vector Machine (SVM), hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.89	0.97	0.93	333
2.0	0.76	0.64	0.69	64
3.0	1.00	0.31	0.47	29
accuracy			0.87	426
macro avg	0.88	0.64	0.70	426
weighted avg	0.88	0.87	0.86	426

Accuracy : 0.8732394366197183

Gambar 13. Klasifikasi SVM (80:20)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 80:20 menunjukkan hasil sebesar 87%, hasil precision sebesar 88%, hasil recall sebesar 87%, dan hasil f1-score sebesar 86%.



Gambar 14. Confusion Matrix SVM (80:20)

4.2.3 Uji Coba 3 (Probabilitas 90:10)

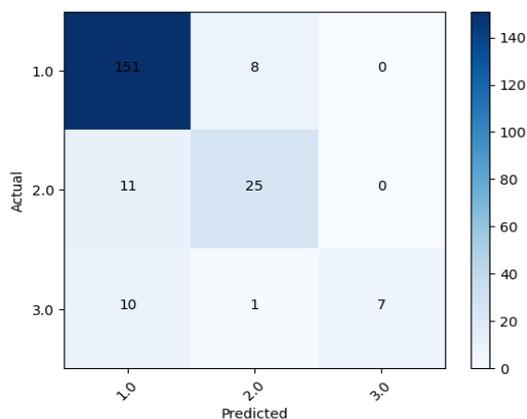
Dengan percobaan 3 menggunakan algoritma Support Vector Machine (SVM), hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.88	0.95	0.91	159
2.0	0.74	0.69	0.71	36
3.0	1.00	0.39	0.56	18
accuracy			0.86	213
macro avg	0.87	0.68	0.73	213
weighted avg	0.86	0.86	0.85	213

Accuracy : 0.8591549295774648

Gambar 15. Klasifikasi SVM (90:10)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 90:10 menunjukkan hasil sebesar 86%, hasil precision sebesar 86%, hasil recall sebesar 86%, dan hasil f1-score sebesar 85%.



Gambar 16. Confusion Matrix SVM (90:10)

4.3 Logistic Regression

Pengujian dilakukan dengan algoritma Logistic Regression pada dataset kesehatan janin yang berisi 2126 record. Data ini dibagi menjadi data pelatihan dan pengujian sebagai berikut :

Tabel 6. Pembagian Data Pelatihan dan Data Pengujian

Uji Coba	Jumlah Record		
	Probabilitas	Pelatihan	Pengujian
1	70:30	1488	628
2	80:20	1700	426
3	90:20	1913	213

4.3.1 Uji Coba 1 (Probabilitas 70:30)

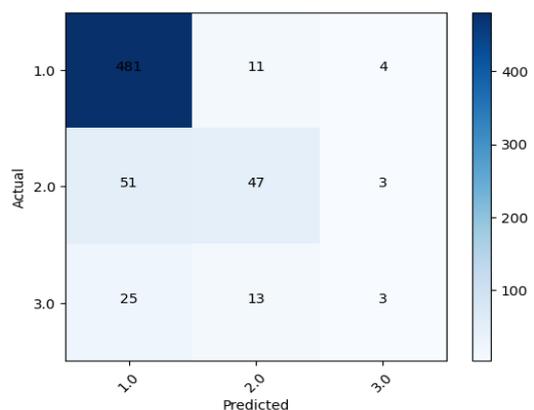
Dengan percobaan 1 menggunakan algoritma Logistic Regression, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.86	0.97	0.91	496
2.0	0.66	0.47	0.55	101
3.0	0.30	0.07	0.12	41
accuracy			0.83	638
macro avg	0.61	0.50	0.53	638
weighted avg	0.80	0.83	0.80	638

Accuracy : 0.8322884012539185

Gambar 17. Klasifikasi Logistic Regression (70:30)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 70:30 menunjukkan hasil sebesar 83%, hasil precision sebesar 80%, hasil recall sebesar 83%, dan hasil f1-score sebesar 80%.



Gambar 18. Confusion Matrix Logistic Regression (70:30)

4.3.2 Uji Coba 2 (Probabilitas 80:20)

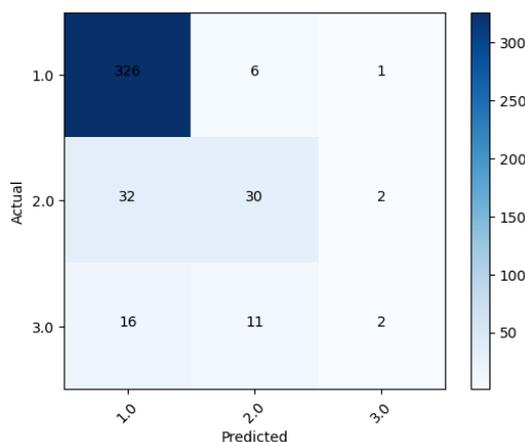
Dengan percobaan 2 menggunakan algoritma Logistic Regression, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.87	0.98	0.92	333
2.0	0.64	0.47	0.54	64
3.0	0.40	0.07	0.12	29
accuracy			0.84	426
macro avg	0.64	0.51	0.53	426
weighted avg	0.80	0.84	0.81	426

Accuracy : 0.8403755868544601

Gambar 19. Klasifikasi Logistic Regression (80:20)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 80:20 menunjukkan hasil sebesar 84%, hasil precision sebesar 80%, hasil recall sebesar 84%, dan hasil f1-score sebesar 81%.



Gambar 20. Confusion Matrix Logistic Regression (80:20)

4.3.3 Uji Coba 3 (Probabilitas 90:10)

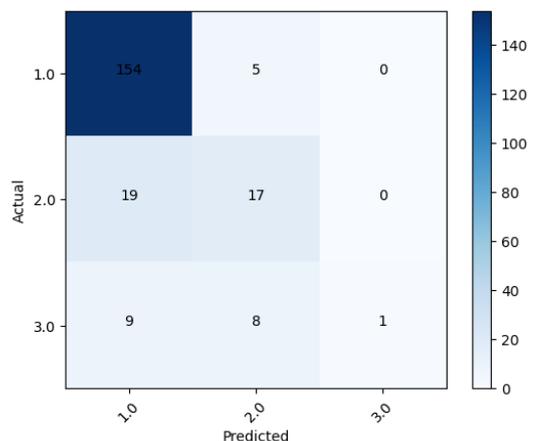
Dengan percobaan 3 menggunakan algoritma Logistic Regression, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.85	0.97	0.90	159
2.0	0.57	0.47	0.52	36
3.0	1.00	0.06	0.11	18
accuracy			0.81	213
macro avg	0.80	0.50	0.51	213
weighted avg	0.81	0.81	0.77	213

Accuracy : 0.8075117370892019

Gambar 21. Klasifikasi Logistic Regression (90:10)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 90:10 menunjukkan hasil sebesar 81%, hasil precision sebesar 81%, hasil recall sebesar 81%, dan hasil f1-score sebesar 77%.



Gambar 22. Confusion Matrix Logistic Regression (90:10)

4.4 Naive Bayes

Pengujian dilakukan dengan algoritma Naive Bayes pada dataset kesehatan janin yang berisi 2126 record. Data ini dibagi menjadi data pelatihan dan pengujian sebagai berikut :

Tabel 7. Pembagian Data Pelatihan dan Data Pengujian

Uji Coba	Jumlah Record		
	Probabilitas	Pelatihan	Pengujian
1	70:30	1488	628
2	80:20	1700	426
3	90:20	1913	213

4.4.1 Uji Coba 1 (Probabilitas 70:30)

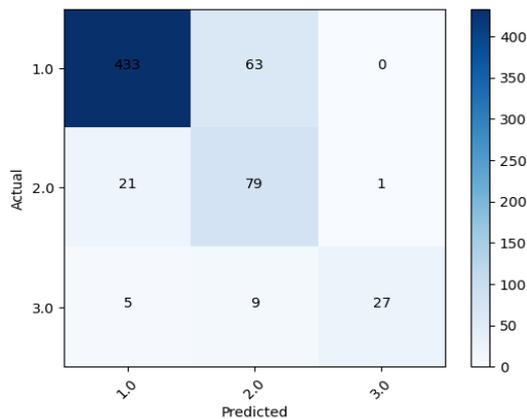
Dengan percobaan 1 menggunakan algoritma Naive Bayes, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.94	0.87	0.91	496
2.0	0.52	0.78	0.63	101
3.0	0.96	0.66	0.78	41
accuracy			0.84	638
macro avg	0.81	0.77	0.77	638
weighted avg	0.88	0.84	0.85	638

Accuracy : 0.8448275862068966

Gambar 23. Klasifikasi Naive Bayes (70:30)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 70:30 menunjukkan hasil sebesar 84%, hasil precision sebesar 88%, hasil recall sebesar 84%, dan hasil f1-score sebesar 85%.



Gambar 24. Confusion Matrix Naive Bayes (70:30)

4.4.2 Uji Coba 2 (Probabilitas 80:20)

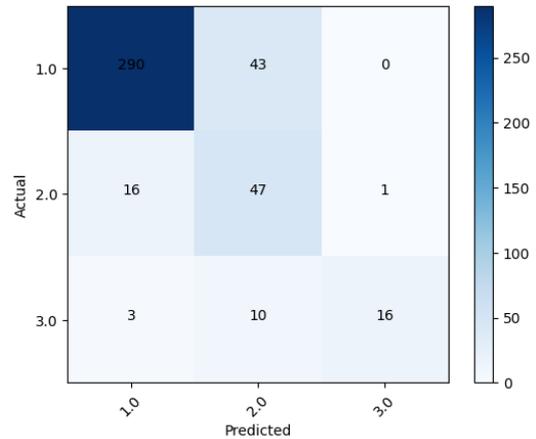
Dengan percobaan 2 menggunakan algoritma Naive Bayes, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.94	0.87	0.90	333
2.0	0.47	0.73	0.57	64
3.0	0.94	0.55	0.70	29
accuracy			0.83	426
macro avg	0.78	0.72	0.72	426
weighted avg	0.87	0.83	0.84	426

Accuracy : 0.8286384976525821

Gambar 25. Klasifikasi Naive Bayes (80:20)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 80:20 menunjukkan hasil sebesar 83%, hasil precision sebesar 87%, hasil recall sebesar 83%, dan hasil f1-score sebesar 84%.



Gambar 26. Confusion Matrix Naive Bayes (80:20)

4.4.3 Uji Coba 3 (Probabilitas 90:10)

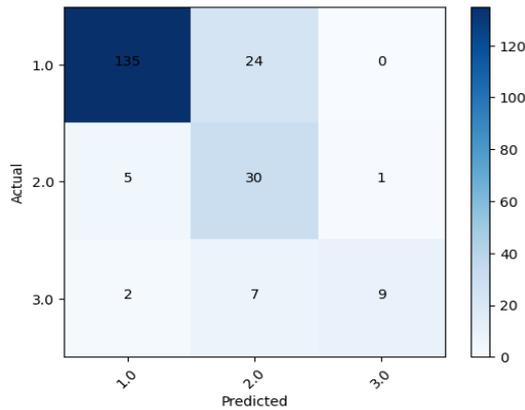
Dengan percobaan 3 menggunakan algoritma Naive Bayes, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.95	0.85	0.90	159
2.0	0.49	0.83	0.62	36
3.0	0.90	0.50	0.64	18
accuracy			0.82	213
macro avg	0.78	0.73	0.72	213
weighted avg	0.87	0.82	0.83	213

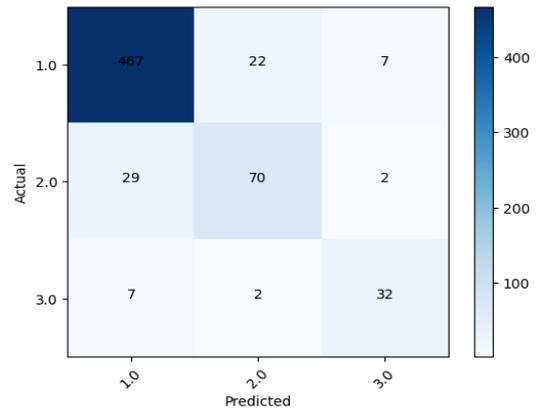
Accuracy : 0.8169014084507042

Gambar 27. Klasifikasi Naive Bayes (90:10)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 90:10 menunjukkan hasil sebesar 82%, hasil precision sebesar 87%, hasil recall sebesar 82%, dan hasil f1-score sebesar 83%.



Gambar 28. Confusion Matrix Naive Bayes (90:10)



Gambar 30. Confusion Matrix Decision Tree (70:30)

4.5 Decision Tree

Pengujian dilakukan dengan algoritma Decision Tree pada dataset kesehatan janin yang berisi 2126 record. Data ini dibagi menjadi data pelatihan dan pengujian sebagai berikut :

Tabel 8. Pembagian Data Pelatihan dan Data Pengujian

Uji Coba	Jumlah Record		
	Probabilitas	Pelatihan	Pengujian
1	70:30	1488	628
2	80:20	1700	426
3	90:20	1913	213

4.5.1 Uji Coba 1 (Probabilitas 70:30)

Dengan percobaan 1 menggunakan algoritma Decision Tree, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.93	0.94	0.93	496
2.0	0.74	0.69	0.72	181
3.0	0.78	0.78	0.78	41
accuracy			0.89	638
macro avg	0.82	0.81	0.81	638
weighted avg	0.89	0.89	0.89	638

Accuracy : 0.8918495297805643

Gambar 29. Klasifikasi Decision Tree (70:30)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 70:30 menunjukkan hasil sebesar 89%, hasil precision sebesar 89%, hasil recall sebesar 89%, dan hasil f1-score sebesar 89%.

4.5.2 Uji Coba 2 (Probabilitas 80:20)

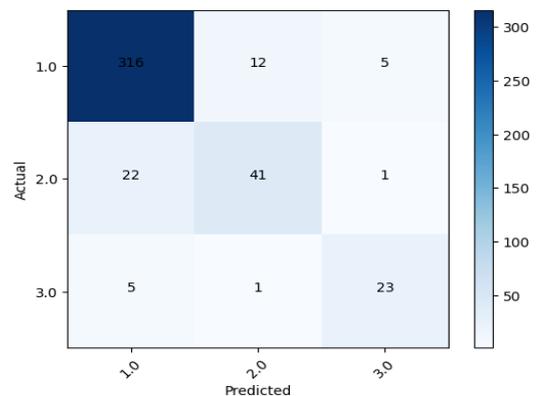
Dengan percobaan 2 menggunakan algoritma Decision Tree, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.92	0.95	0.93	333
2.0	0.76	0.64	0.69	64
3.0	0.79	0.79	0.79	29
accuracy			0.89	426
macro avg	0.82	0.79	0.81	426
weighted avg	0.89	0.89	0.89	426

Accuracy : 0.892018779342723

Gambar 31. Klasifikasi Decision Tree (80:20)

Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 80:20 menunjukkan hasil sebesar 89%, hasil precision sebesar 89%, hasil recall sebesar 89%, dan hasil f1-score sebesar 89%.



Gambar 32. Confusion Matrix Decision Tree (80:20)

4.5.3 Uji Coba 3 (Probabilitas 90:10)

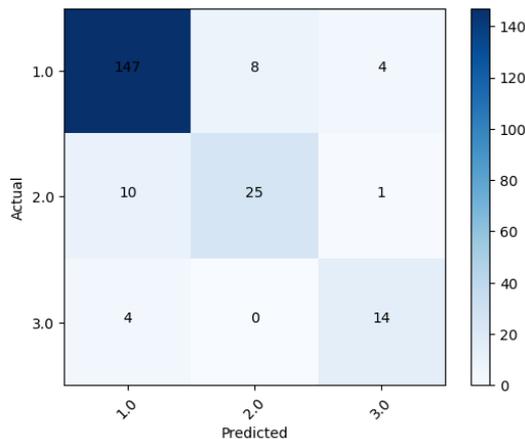
Dengan percobaan 3 menggunakan algoritma Decision Tree, hasil yang didapatkan yaitu :

	precision	recall	f1-score	support
1.0	0.91	0.92	0.92	159
2.0	0.76	0.69	0.72	36
3.0	0.74	0.78	0.76	18
accuracy			0.87	213
macro avg	0.80	0.80	0.80	213
weighted avg	0.87	0.87	0.87	213

Accuracy : 0.8732394366197183

Gambar 33. Klasifikasi Decision Tree (90:10)

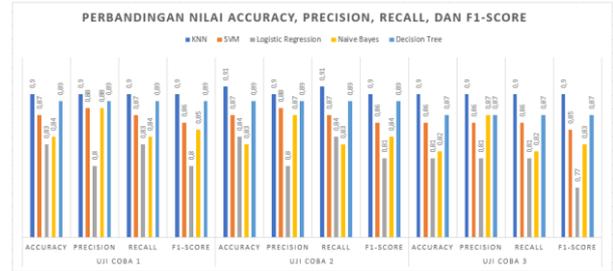
Hasil akurasi dengan membagi data pelatihan dan pengujian menjadi 90:10 menunjukkan hasil sebesar 87%, hasil precision sebesar 87%, hasil recall sebesar 87%, dan hasil f1-score sebesar 87%.



Gambar 34. Confusion Matrix Decision Tree (90:10)

4.6 Analisis Perbandingan antar Algoritma

Setelah melakukan 15 uji coba pada masing-masing algoritma terhadap dataset kesehatan janin maka dapat dilihat untuk hasil akurasi, precision, recall, dan f1-score dari Gambar 35.



Gambar 35. Diagram Perbandingan Antar Algoritma

Tabel 9. Tabel Perbandingan Antar Algoritma

Confusion Matrix	Algoritma					
	KNN	SVM	Logistic Regression	Naive Bayes	Decision Tree	
Uji Coba 1	Accuracy	0,9	0,87	0,83	0,84	0,89
	Precision	0,9	0,88	0,8	0,88	0,89
	Recall	0,9	0,87	0,83	0,84	0,89
	F1-Score	0,9	0,86	0,8	0,85	0,89
Uji Coba 2	Accuracy	0,91	0,87	0,84	0,83	0,89
	Precision	0,9	0,88	0,8	0,87	0,89
	Recall	0,91	0,87	0,84	0,83	0,89
	F1-Score	0,9	0,86	0,81	0,84	0,89
Uji Coba 3	Accuracy	0,9	0,86	0,81	0,82	0,87
	Precision	0,9	0,86	0,81	0,87	0,87
	Recall	0,9	0,86	0,81	0,82	0,87
	F1-Score	0,9	0,85	0,77	0,83	0,87

Menurut hasil dari tiga ujicoba yang dilakukan pada masing-masing algoritma dengan data pelatihan dan pengujian yang berbeda dan dengan berbagai algoritma, lima algoritma yang digunakan untuk membandingkan pembagian dataset pada uji coba pertama, yaitu 70:30. Algoritma KNN memiliki nilai akurasi terbaik sebesar 90%, precision terbaik sebesar 90%, recall terbaik sebesar 90%, dan skor f1 terbaik sebesar 90%. Dalam uji coba kedua, algoritma KNN memiliki nilai akurasi terbaik sebesar 91%, nilai ketepatan terbaik sebesar 90%, nilai recall terbaik sebesar 90%, dan nilai f1-score terbaik sebesar 90%. Dalam uji coba ketiga, algoritma KNN memiliki nilai akurasi terbaik sebesar 90%, nilai ketepatan terbaik sebesar 90%, nilai recall terbaik sebesar 90%, dan nilai f1-score terbaik sebesar 90%.

5 Kesimpulan

Prolongued decelerations, abnormal short-term variability, dan percentage of time with abnormal long-term variability adalah variabel yang paling signifikan dalam penelitian ini. Percobaan dilakukan tiga kali untuk setiap algoritma dan dengan berbagai hasil. Hasil menunjukkan bahwa algoritma KNN memiliki nilai akurasi tertinggi sebesar 91% pada uji coba

ke 2, nilai ketepatan tertinggi sebesar 90% pada uji coba ke 1,2, dan 3, nilai recall tertinggi sebesar 91% pada uji coba ke 2, dan nilai f1-score tertinggi sebesar 90% pada uji coba ke 1,2, dan 3. Perbandingan tingkat performa dalam melakukan prediksi kesehatan janin, algoritma KNN menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sebesar 91%, 90%, 91%, 90% dengan rata-rata performa algoritma sebesar 90,5%, algoritma SVM menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sebesar 87%, 88%, 87%, 86% dengan rata-rata performa algoritma sebesar 87%, algoritma Logistic Regression menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sebesar 84%, 80%, 84%, 81% dengan rata-rata performa algoritma sebesar 82,25%, algoritma Naive Bayes menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sebesar 84%, 88%, 84%, 85% dengan rata-rata performa algoritma sebesar 85,25%, dan algoritma Decision Tree menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sebesar 89%, 89%, 89%, 89% dengan rata-rata performa algoritma sebesar 89%. Keseluruhan, hasil penelitian ini menunjukkan bahwa, dibandingkan dengan algoritma lainnya, algoritma KNN cenderung stabil untuk semua nilai. Sebaliknya, algoritma regresi logistik dan pohon keputusan tidak stabil untuk semua nilai. (Gambar 35).

References

- Amelia, U., Indra, J., & Masruriyah, A. F. N. (2022). Implementasi Algoritma Support Vector Machine (SVM) Untuk Prediksi Penyakit Stroke Dengan Atribut Berpengaruh. *Scientific Student Journal for Information, Tecnology and Sciene*, III(2), 254–259. <http://journal.ubpkarawang.ac.id/mahasiswa/index.php/ssj/article/view/447>
- Rahmi, Antoni, D., Syaputra, H., Fatoni & Kurniawan, T. B. (2023). Metode Klasifikasi Gejala Penyakit Coronavirus Disease 19 (COVID-19) Menggunakan Algoritma Neural Network. *Sistem Informasi Dan Komputer*, 12(1), 16–23. <https://doi.org/10.32736/sisfokom.v12i1.1406>
- Argina, A. M. (2020). Penerapan Metode Klasifikasi K-Nearest Neighbor pada Dataset Penderita Penyakit Diabetes. *Indonesian Journal of Data and Science*, 1(2), 29–33. <https://doi.org/10.33096/ijodas.v1i2.11>
- Arif, M. Z., Ahmed, R., Sadia, U. H., Tultul, M. S. I., & Chakma, R. (2020). Decision Tree Method Using for Fetal State Classification from Cardiotography Data. *Journal of Advanced Engineering and Computation*, 4(1), 64-73. <https://doi.org/10.25073/jaec.202041.273>
- Atthalla, I. N., Jovandy, A., & Habibie, H. (2018). Klasifikasi Penyakit Kanker Payudara Menggunakan Metode K Nearest Neighbor. *Prosiding Annual Research Seminar*, 4(1), 148-151.
- Handayani, F., Sari Kusuma, K., Leoni Asbudi, H., Guines Purnasiwi, R., Kusuma, R., Sunyoto, A., & Mega Pradnya, W. (2021). *JEPIN (Jurnal Edukasi dan Penelitian Informatika) Komparasi Support Vector Machine, Logistic Regression Dan Artificial Neural Network dalam Prediksi Penyakit Jantung*. 7(3), 329–334.
- Perdana, A., Hermawan, A., & Avianto, D. (2023). Analyze Important Features of PIMA Indian Database For Diabetes Prediction Using KNN. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 12(1), 70–75. <https://doi.org/10.32736/sisfokom.v12i1.1598>
- Putra, P. D., & Rini, D. P. (2019). Prediksi Penyakit Jantung dengan Algoritma Klasifikasi. *Prosiding Annual Research Seminar 2019*, 5(1), 978–979.
- Rahmayanti, N., Pradani, H., Pahlawan, M., & Vinarti, R. (2021). Comparison of machine learning algorithms to classify fetal health using cardiotocogram data. *Procedia Computer Science*, 197(2021), 162–171. <https://doi.org/10.1016/j.procs.2021.12.130>
- Rian Santoso, M., & Musa, P. (2021). *Rekomendasi Kesehatan Janin Dengan Penerapan Algoritma C5.0 Menggunakan Classifying Cardiotocography Dataset*. 9(2), 65-76.
- Safaat, M., Sahari, A., & Lusiyanti, D. (2020). Implementasi Metode K-Nearest Neighbor Untuk Mengklasifikasi Jenis Penyakit Katarak. *Jurnal Ilmiah Matematika Dan Terapan*, 17(1), 92–99. <https://doi.org/10.22487/2540766x.2020.v17.i1.15184>
- Saputro, I. W., & Sari, B. W. (2020). Uji Performa Algoritma Naive Bayes untuk Prediksi Masa Studi Mahasiswa. *Creative Information Technology Journal*, 6(1), 1–11. <https://doi.org/10.24076/citec.2019v6i1.178>
- Sulastri, S., Hadiono, K., & Anwar, M. T. (2020). Analisis Perbandingan Klasifikasi Prediksi Penyakit Hepatitis Dengan Menggunakan Algoritma K-Nearest Neighbor, Naive Bayes



- Dan Neural Network. *Dinamik*, 24(2), 82–91.
<https://doi.org/10.35315/dinamik.v24i2.7867>
- Sulihati, I., Syukur, A., & Marjuni, A. (2022). Deteksi Kesehatan Janin Menggunakan Decision Tree dan Feature Forward Selection. *Building of Informatics, Technology and Science (BITS)*, 4(3), 1658-1664.
<https://doi.org/10.47065/bits.v4i3.2672>
- Yuli Mardi. (2019). Data Mining: Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika*. *Jurnal Edik Informatika*, 2(2), 213–219.

