

# Prediksi Tingkat Resiko Kesehatan Ibu Saat Hamil Menggunakan Algoritma C4.5

Muhammad Zidan  
Universitas Trilogi  
JL.TMP Kalibata no 1, Kampus Trilogi  
[muzidan69@gmail.com](mailto:muzidan69@gmail.com)

Penulis Korespondensi : Muhammad Zidan

**Abstrak**— Sejak Organisasi Kesehatan Dunia membentuk komisinya pada determinan sosial kesehatan (SDOH) lebih dari satu dekade yang lalu, sejumlah besar penelitian telah menunjukkan bahwa determinan sosial didefinisikan sebagai kondisi di mana orang dilahirkan, tumbuh, hidup, bekerja, dan usia merupakan pendorong yang signifikan terhadap risiko dan kerentanan penyakit dalam perawatan klinis dan sistem kesehatan masyarakat. Kehamilan adalah waktu khusus ketika perubahan biologis mungkin membuat Anda lebih sensitif terhadap paparan bahan kimia. Wanita hamil terpapar berbagai racun lingkungan, termasuk penghambat api, plasticizer, dan pestisida, melalui udara, makanan, air, dan barang-barang konsumen. Meskipun sebagian besar bahaya kesehatan bahan kimia bagi wanita kurang dipahami, paparan timbal meningkatkan kemungkinan masalah hipertensi akibat kehamilan. Hasil Penelitian terdapat tingkat akurasi 67.36%. dengan nilai class precision high risk sebesar 86.61% dan class precision low risk sebesar 58,35%. Class recall untuk high risk sebesar 80.88% dan class recall untuk low risk sebesar 99.01%. Dari hasil penelitian yang telah dilakukan oleh peneliti, maka dapat disimpulkan. Tingkat resiko kesehatan ibu saat hamil dapat diprediksi dengan memanfaatkan teknik data mining menggunakan algoritma C4.5.

**Kata Kunci**—Data Mining, Ibu Hamil, Prediksi, Algoritma C4.5

**Abstract**— Since the World Health Organization established its commission on the social determinants of health (SDOH) more than a decade ago, a large number of studies have shown that social determinants defined as the conditions in which people are born, grow up, live, work, and age are significant drivers of risk. and disease susceptibility in clinical care and public health systems. Pregnancy is a special time when biological changes may make you more sensitive to chemical exposure. Pregnant women are exposed to a variety of environmental toxins, including flame retardants, plasticizers, and poisons, through air, food, air and consumer goods. Although most of the health hazards of chemicals for women are poorly understood, exposure to lead increases the possibility of hypertension problems due to pregnancy. The results of the study showed an accuracy rate of 67.36%. With a class precision high risk of 86.61% and a class precision low risk of 58.35%. Class recall for high risk is 80.88% and class recall for low risk is 99.01%. From the results of research that has been conducted by researchers, it can be concluded. The level of maternal health risk during pregnancy can be predicted by utilizing data mining techniques using the C4.5 algorithm.

**Keywords**— Data Mining, Pregnant Women, Prediction, C4.5



## I. PENDAHULUAN

Teknologi Data Mining dapat mencari informasi yang berpotensi penting dalam jumlah data yang sangat besar. Ini dipisahkan menjadi persiapan data dan penambahan data, serta ekspresi dan analisis hasil.

Ini adalah teknik pemrosesan informasi canggih yang memanfaatkan teknologi basis data. Teknologi basis data adalah cabang ilmu komputer yang mempelajari, mengelola, dan mengimplementasikan basis data. Data database diproses dan dianalisis dengan meneliti teori yang mendasari dan teknik implementasi struktur database, penyimpanan, desain, administrasi, dan aplikasi.

Pohon keputusan adalah alat klasifikasi dan prediksi yang kuat dan terkenal. Pendekatan pohon keputusan mengubah sejumlah besar informasi menjadi pohon keputusan yang mewujudkan aturan. Bahasa alami membuatnya mudah untuk memahami aturan. Keuntungan utama dari menggunakan pohon keputusan adalah kapasitasnya untuk menyederhanakan proses pengambilan keputusan yang rumit sehingga pembuat keputusan dapat lebih memahami solusi masalah dengan mengubah data (tabel) menjadi model pohon keputusan dan kemudian menerjemahkan model pohon keputusan menjadi aturan. Pohon keputusan juga dapat digunakan untuk mengeksplorasi data, khususnya untuk menemukan korelasi tersembunyi antara sejumlah faktor input potensial dan variabel target.

Kehamilan adalah satu-satunya waktu dari proses yang dimediasi hormon yang terkoordinasi dengan erat yang sangat mengubah fisiologi ibu untuk mengakomodasi bayi yang sedang tumbuh dan mempersiapkan persalinan dan menyusui. Perubahan fisiologi vaskular, metabolisme, organ reproduksi, aktivitas endokrin, dan sistem kekebalan mungkin membuat ibu lebih rentan terhadap paparan bahan kimia dan masalah kesehatan yang menyertainya. Timbal, misalnya, menumpuk di tulang dari waktu ke waktu dan dimobilisasi selama kehamilan dengan pelepasan kalsium.

Untuk menjawab permasalahan diatas peneliti menggunakan teknik data mining dengan metode algoritma C4.5 untuk memprediksi keselamatan pasien gagal jantung berdasarkan umur, penurunan sel darah merah atau hemoglobin, jika pasien menderita hipertensi, tingkat enzim CPK

dalam darah, jika pasien menderita diabetes, persentase darah yang keluar jantung pada setiap kontraksi, jenis kelamin, trombosit dalam darah, tingkat kreatinin dalam darah, Tingkat natrium dalam darah, jika pasien merokok, periode tindak lanjut

Penelitian ini dilandasi oleh peneliti terdahulu oleh menyampayakan hasil penelitiannya mengenai Penerapan Algoritma C4.5 Untuk Klasifikasi Pola Kepuasan Pelayanan E-Ktp Di Kantor Camat Pematang Bandar. Penelitian ini menerapkan algoritma C4.5 dalam memutuskan kepuasan pelayanan. Hasil penelitian terdapat tingkat akurasi 95,24%

## II. METODE PENELITIAN

Ada 2 tahap metode penelitian yaitu :

### 1. Pengumpulan Data

Data yang digunakan di penelitian ini diambil dari UCI Repisitory. Ada 1014 sampel dari data yang diambil, atribut-atribut yang ada di data adalah umur, nilai atas tekanan darah (SystolicBP), nilai bawah tekanan darah (DiastolicBP), kadar glucose darah (BS), detak jantung (Heartrate).

### 2. Rumus Algoritma C4.5

Rumus yang digunakan dalam penelitian ini adalah

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

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

### 3. Pemodelan Metode

Algoritma C4.5 digunakan dalam penelitian ini. Teknik C4.5 memodelkan kumpulan data sampel, lalu menghitung entropi (S) dari semua karakteristik. Setelah entropi (S) ditentukan, Gain maksimum semua atribut dihitung, dan atribut dengan Gain terbesar digunakan sebagai root/node. Kemudian, untuk setiap nilai, bangun sebuah cabang, dan untuk setiap kasus di cabang, ulangi perhitungan Gain sampai semua data termasuk dalam kelas yang sama. Ketika tidak ada lagi atribut yang dipartisi dan semua tupel di node N memiliki kelas yang sama, proses pembuatan pohon keputusan

berakhir, dan atribut yang dipilih tidak lagi diperlukan dalam perhitungan.

### III. HASIL DAN PEMBAHASAN

Peneliti menghitung Algoritma C4.5 dan menghasilkan 5 model rule atau aturan untuk tingkat resiko ibu hamil. Grafik berikut menggambarkan model aturan yang dihasilkan oleh peneliti dalam bentuk pohon keputusan:

#### Tree

```

SystolicBP > 132.500
| BS > 6.750: high risk (high risk=128, low risk=0, mid risk=0)
| BS ≤ 6.750: mid risk (high risk=0, low risk=0, mid risk=5)
SystolicBP ≤ 132.500
| BS > 9.500
| | Age > 38.500
| | | BodyTemp > 101.500: mid risk (high risk=0, low risk=0, mid risk=2)
| | | BodyTemp ≤ 101.500: high risk (high risk=52, low risk=4, mid risk=)
| | | Age ≤ 38.500: high risk (high risk=40, low risk=0, mid risk=0)
| | BS ≤ 9.500
| | | SystolicBP > 129.500: mid risk (high risk=0, low risk=0, mid risk=35)
| | | SystolicBP ≤ 129.500
| | | | BS > 8.500: mid risk (high risk=10, low risk=0, mid risk=19)
| | | | BS ≤ 8.500: low risk (high risk=42, low risk=402, mid risk=245)
    
```

Gambar 3. Model Aturan

accuracy: 67.36%

	true high risk	true low risk	true mid risk	class precision
pred. high risk	220	4	30	86.61%
pred. low risk	42	402	245	58.35%
pred. mid risk	10	0	61	85.92%
class recall	80.88%	99.01%	18.15%	

Gambar 4. Hasil Akurasi

Berdasarkan Pengolahan data diatas didapat nilai akurasi sebesar 67,36%. Dengan nilai class precision class recal untuk high risk sebesar 86,61% dan 80,88%, untuk low risk sebesar 58,35% dan 99,01% dan untuk mid risk sebesar 85,92% dan 18.15%

### IV. KESIMPULAN

Berdasarkan pembahasan sejauh ini, dapat disimpulkan bahwa data mining dengan algoritma C4.5 dapat digunakan untuk meramalkan besarnya risiko kesehatan pada ibu hamil. Hasil pengujian pendekatan Algoritma C4.5 ke dalam RapidMiner menghasilkan 5 (sepuluh) rule dengan tingkat akurasi sebesar 67,36%.

### V. DAFTAR PUSTAKA

J. Yang et al., “Brief introduction of medical database and data mining technology in big data era,” *Journal of Evidence-Based*

*Medicine*, vol. 13, no. 1. Blackwell Publishing, pp. 57–69, Feb. 01, 2020. doi: 10.1111/jebm.12373.

I. S. Damanik, A. P. Windarto, A. Wanto, Poningsih, S. R. Andani, and W. Saputra, “Decision Tree Optimization in C4.5 Algorithm Using Genetic Algorithm,” in *Journal of Physics: Conference Series*, Sep. 2019, vol. 1255, no. 1. doi: 10.1088/1742-6596/1255/1/012012.

J. Varshavsky et al., “Heightened susceptibility: A review of how pregnancy and chemical exposures influence maternal health,” *Reproductive Toxicology*, vol. 92. Elsevier Inc., pp. 14–56, Mar. 01, 2020. doi: 10.1016/j.reprotox.2019.04.004.

P. Nuraini, J. Tata Hardinata, Y. Pranayama Purba Program Studi Sistem Informasi, S. A. Tunas Bangsa Jalan Jendral Sudirman Blok, and S. Utara, “RESOLUSI: Rekayasa Teknik Informatika dan Informasi Penerapan Algoritma C4.5 Untuk Klasifikasi Pola Kepuasan Pelayanan E-Ktp Di Kantor Camat Pematang Bandar,” *Media Online*, vol. 3, no. 2, pp. 138–144, 2022, [Online]. Available: <https://djournals.com/resolusi>

J. Crear-Perry, R. Correa-De-Araujo, T. Lewis Johnson, M. R. Mclemore, E. Neilson, and M. Wallace, “Social and Structural Determinants of Health Inequities in Maternal Health,” *J Womens Health*, vol. 30, no. 2, pp. 230–235, Feb. 2021, doi: 10.1089/jwh.2020.8882.

U. M. Kudus, J. Ganesha, and P. Kudus, “Fida Maisa Hana.”

N. Sunanto and G. Falah, “PENERAPAN ALGORITMA C4.5 UNTUK MEMBUAT MODEL PREDIKSI PASIEN YANG MENGIDAP PENYAKIT DIABETES,” *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 7, no. 2, pp. 208–216, Jul. 2022, doi: 10.36341/rabit.v7i2.2435.

J. S. Lee, “AUC4.5: AUC-Based C4.5 Decision Tree Algorithm for Imbalanced Data Classification,” *IEEE Access*, vol. 7, pp. 106034–106042, 2019, doi: 10.1109/ACCESS.2019.2931865.

T. H. B. de Lima, M. M. Amorim, S. Buainain Kassar, and L. Katz, “Maternal near miss determinants at a maternity hospital for high-risk pregnancy in northeastern Brazil: A prospective study,” *BMC Pregnancy*

---

Childbirth, vol. 19, no. 1, Aug. 2019, doi:  
10.1186/s12884-019-2381-9.  
S. Pant, S. Koirala, and M. Subedi, "Access to  
Maternal Health Services during COVID-

19," *Europasian Journal of Medical  
Sciences*, vol. 2, pp. 46–50, Jul. 2020, doi:  
10.46405/ejms.v2i2.110

