

# Penilaian Kinerja Berbasis Pohon Keputusan

## Pengklasifikasi Prediktif untuk Perawatan Kehamilan Risiko

Mário WL Moreira<sup>1,2</sup>, Joel JPC Rodrigues<sup>1,3,4</sup>, Neeraj Kumar<sup>5</sup>, Jianwei Niu<sup>6</sup>, Ishak Woungang<sup>7</sup>

<sup>1</sup> Instituto de Telecomunicações, Universidade da Beira Interior, Covilhã,

<sup>2</sup> Portugal Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), Aracati, Brasil

<sup>3</sup> Institut Telekomunikasi Nasional (Inatel), Santa Rita do Sapucaí, MG, Brasil

<sup>4</sup> Universitas Fortaleza (UNIFOR), Fortaleza, CE, Brasil

<sup>5</sup> Departemen Ilmu dan Teknik Komputer, Universitas Thapar, Patiala (Punjab), India

<sup>6</sup> Sekolah Ilmu dan Teknik Komputer, Universitas Beihang, Beijing, Cina

<sup>7</sup> Departemen Ilmu Komputer, Universitas Ryerson, Ontario, Toronto, Kanada

mario.moreira@ieee.org; joelj@ieee.org; neeraj.kumar@thapar.edu; niujianwei@buaa.edu.cn; iwoungan@scs.ryerson.ca

**Abstrak**— Konsep inti e-Health mencakup penggunaan Web secara terintegrasi dengan alat dan layanan untuk kesehatan. Definisi ini meningkatkan akses, efisiensi, dan proses kualitas perawatan klinis yang diperlukan untuk peningkatan pemberian layanan. Sistem pendukung keputusan (DSS) milik sejumlah besar dimensi konsep e-Health. Untuk konstruksi sistem ini, penting untuk menemukan mekanisme cerdas yang andal yang mampu mengidentifikasi penyakit yang dapat memperburuk kondisi klinis pasien. Dengan demikian, makalah ini mengusulkan penggunaan teknik data mining (DM) berbasis pohon untuk prediksi gangguan hipertensi pada kehamilan berisiko. Ini menyajikan pemodelan, evaluasi kinerja, dan perbandingan antara pengklasifikasi berbasis pohon ID3 dan NBTtree. Metode validasi silang 5 kali lipat mewujudkan perbandingan kinerja. Hasil menunjukkan bahwa pengklasifikasi NBTtree memperoleh kinerja yang lebih baik, menyajikan ukuran-F 0,609, area ROC 0,753, dan statistik Kappa 0,4658. Pengklasifikasi ini dapat menjadi kunci pengembangan sistem cerdas yang mampu memprediksi kejadian risiko pada kehamilan. Oleh karena itu, DSS adalah solusi utama untuk mengurangi kematian ibu dan janin.

**Kata Kunci**— Sistem Pendukung Keputusan; Pohon keputusan; Algoritma klasifikasi; Kehamilan; Hipertensi; Pengukuran; Standar

### I. PENDAHULUAN

Rekomendasi kelima yang disampaikan oleh United Nations Millennium Development Goals (MDGs) berkaitan dengan peningkatan kesehatan ibu. Untuk mencapai tujuan ini, diusulkan pengurangan angka kematian ibu sekitar tiga perempat, dalam 25 tahun [1]. Menurut laporan PBB yang diterbitkan pada tahun 2014, angka kematian ibu sekitar 45% antara tahun 1990 dan 2015, menunjukkan tanda-tanda kemajuan terkait peningkatan kesehatan ibu selama kehamilan dan persalinan (di beberapa negara). Namun, rasio kematian ibu tetap tidak dapat diterima tinggi. Uang muka jauh dari penurunan tahunan 5,5% yang dibutuhkan untuk mencapai tujuan. Perdarahan, sepsis, abortus dalam kondisi berisiko, obstruksi persalinan, dan gangguan hipertensi dalam kehamilan merupakan penyebab lebih dari 80% kematian ibu. Sayangnya, sebagian besar dari semua kerugian dapat dicegah. Akses terhadap pelayanan kesehatan yang memadai,

peralatan, dan perbekalan serta staf kesehatan yang berkualitas dapat mencegah kematian ini [2]. Tinjauan komprehensif terhadap tujuan milenium yang diwujudkan oleh Organisasi Kesehatan Dunia (WHO) telah menunjukkan perlunya cakupan kesehatan universal.

Cakupan ini menyiratkan bahwa semua orang memiliki akses ke layanan kesehatan berkualitas tanpa mempertaruhkan situasi keuangan mereka. Proposal tersebut mencakup akses ke pencegahan, perlindungan, pengobatan, dan perawatan paliatif. Untuk itu diperlukan pengembangan smart health system [3]. Dari sudut pandang e-Health, satu register harus menggabungkan semua orang dari negara tertentu dan memberikan akses informasi yang relevan kepada mereka. Tujuan ini menentukan kebutuhan untuk mengembangkan strategi dan aplikasi e-Health yang mendukung pencegahan dan promosi kesehatan, seperti platform dan layanan yang memungkinkan pasien menangani kesehatan pribadinya. Dalam konteks ini, ternyata kebijakan kesehatan layak untuk daerah terpencil dan penggunaan solusi seluler, telemedis, teknologi yang dapat dikenakan, sensor, catatan kesehatan elektronik (EHR), dan diagnosis adalah hal mendasar. Penggunaan teknologi informasi dan komunikasi (TIK) yang difokuskan pada e-Health melibatkan integrasi data dan transformasinya untuk aplikasi dalam pendukung keputusan [4]. Data mining (DM) adalah solusi yang baik untuk memproses volume data yang besar dari sistem informasi kesehatan [5]. Metode ini dapat mengenali pola yang berguna, memungkinkan dukungan dalam analisis kompleks data klinis.

Untuk mengidentifikasi kemungkinan kasus berisiko tinggi yang dapat menyebabkan kematian ibu hamil dan/atau janin, mekanisme inferensi yang andal ini merupakan alat yang berharga. Kemudian, menggunakan pendekatan DSS ini dapat mendukung penurunan angka kematian ibu dengan cara yang andal, mudah diakses, dan berbiaya rendah. Kemudian, makalah ini mengusulkan penggunaan teknik DM berbasis pohon untuk prediksi gangguan hipertensi pada kehamilan berisiko. Ini menyajikan pemodelan, evaluasi kinerja, dan perbandingan antara pengklasifikasi berbasis pohon ID3 dan NBTtree.

Sisa kertas ini disusun sebagai berikut. Bagian II menyajikan karya-karya terkait membahas metode yang digunakan untuk mengidentifikasi gangguan hipertensi pada kehamilan berisiko tinggi. Bagian III menunjukkan proposal pemodelan yang menggunakan pengklasifikasi berbasis pohon yang mampu mengidentifikasi gangguan hipertensi berdasarkan gejala yang disajikan oleh pasien. Evaluasi kinerja, metode

perbandingan, dan analisis hasil dengan mempertimbangkan rencana yang diusulkan ditunjukkan pada Bagian IV. Akhirnya, Bagian V memberikan kesimpulan dan saran untuk pekerjaan selanjutnya.

## II. IDENTIFIKASI GANGGUAN HIPERTENSI PADA PT KEHAMILAN RISIKO TINGGI

Penelitian ini menggunakan data dari 25 ibu hamil hipertensi yang diperoleh dari dokter spesialis kandungan/kebidanan. Tabel I menunjukkan atribut yang digunakan dan menyertakan penjelasan singkat terkait masing-masing atribut. Variabel input adalah gejala dan faktor risiko yang ditimbulkan oleh ibu hamil sedangkan variabel output (kelas) adalah penyakit hipertensi primer dalam kehamilan. Tabel II memberikan kelas yang mewakili gangguan hipertensi utama dalam kehamilan berdasarkan studi yang disajikan dalam [6].

Tabel I. GEJALA UTAMA DAN FAKTOR RISIKO : PENJELASAN SINGKAT .

<b>Gejala utama</b>	<b>Penjelasan singkat</b>
Sakit kepala	Migrain dapat dikaitkan dengan peningkatan risiko preeklampsia.
Nyeri epigastrium	Nyeri epigastrium adalah gejala lanjut, yang dikaitkan dengan peregangan kapsul hepatic karena edema dan/atau perdarahan. Biasanya dimulai pada bulan pertama kehamilan dan melewati minggu ke-14 atau ke-16 dengan peningkatan produksi hormon.
Mual atau muntah	Beberapa ibu hamil mengalami krisis muntah parah, yang dikenal sebagai hiperemesis, yang menyebabkan tubuh kehilangan cairan dan mineral.
Penglihatan kabur	Dimungkinkan untuk mengidentifikasi beberapa tanda preeklampsia dengan mata. Pada 8% wanita hamil dapat terjadi kehilangan penglihatan sementara, peningkatan kepekaan terhadap cahaya, penglihatan kabur, dan lingkaran cahaya atau kilatan.
Pusing	Pusing terus-menerus adalah tanda anemia, migrain, tekanan darah tinggi atau rendah, dan penyakit tertentu lainnya. Ini terjadi bersamaan dengan penglihatan kabur dan sakit kepala.
Hiperrefleksia	Kehadiran sakit kepala, gejala visual dan neurologis (hyperreflexia) memiliki korelasi tinggi dengan peningkatan tekanan perfusi serebral, yang akan memprediksi peningkatan risiko perubahan serebral (eklampsia).
Busung	Preeklampsia biasanya disertai dengan edema tangan dan wajah. Kehadiran edema tidak dianggap sebagai sinyal yang dapat diandalkan untuk diagnosis preeklampsia karena dapat terjadi pada wanita hamil tanpa perubahan tingkat tekanan.
Oliguria	Penurunan produksi urin (volume urin <500ml/24 jam) merupakan kriteria untuk identifikasi preeklampsia berat.
Hipertensi	Tingkat ketegangan mengklasifikasikan hipertensi dalam ringan (TD sekitar 140/90mmHg) dan berat (TD lebih besar dari 160/110mmHg). Klasifikasi tersebut mengidentifikasi pasien dengan risiko lebih besar atau lebih kecil selama kehamilan, selain memandu pelaksanaan pengobatan.
Proteinuria	Proteinuria terjadi dalam tiga hingga empat minggu sebelum perubahan perkembangan janin dan/atau memburuknya kondisi klinis ibu. Batas proteinuria 2g/24 jam adalah kriteria keparahan. Tingkat intensitas dapat berhubungan dengan prognosis maternal dan perinatal yang lebih buruk.
<b>Faktor risiko</b>	<b>Penjelasan Singkat</b>
Gestasional usia	Usia kehamilan dapat memprediksi risiko berkembangnya bentuk preeklampsia berat, seperti eklampsia dan Sindrom HELLP.

Dimungkinkan untuk membuat pohon keputusan untuk model yang diusulkan berdasarkan informasi ini. Dari gejala yang ditunjukkan oleh seorang wanita hamil dan usia kehamilannya, model yang disarankan dapat menyimpulkan untuk mengidentifikasi gangguan hipertensi dalam kehamilan dan tingkat keparahannya.

Tabel II. GANGGUAN HIPERTENSI KEHAMILAN UTAMA .

Kehamilan Gangguan Hipertensi	Penjelasan singkat
Hipertensi yang sudah ada sebelumnya mempersulit kehamilan, persalinan, dan masa nifas (CH)	Ini adalah hipertensi yang muncul sebelum kehamilan atau didiagnosis sebelum usia kehamilan 20 minggu. Tekanan darah sistolik (TD) > 140mmHg dan/atau TD diastolik > 90mmHg, diukur dalam dua kali pemeriksaan dengan interval 4 jam mendefinisikan hipertensi. Diagnosis seperti itu lebih sulit dilakukan pada wanita hipertensi tanpa diagnosis sebelumnya, karena adanya penurunan tekanan darah secara fisiologis, yang terjadi pada paruh pertama kehamilan. Itu juga dianggap CH yang didiagnosis pertama kali selama kehamilan dan tidak menjadi normal pada periode postpartum.
Gangguan hipertensi yang sudah ada sebelumnya dengan proteinuria yang dilapiskan (PS)	Munculnya proteinuria (>0,3g/24jam) setelah usia kehamilan 20 minggu pada pasien CH. Indikator lain didasarkan pada peningkatan proteinuria pada mereka yang sudah mengalami peningkatan sebelumnya, peningkatan TD mendadak pada mereka dengan kadar terkontrol, karakteristik kelainan klinis atau laboratorium dari
Hipertensi gestasional (diinduksi kehamilan) tanpa proteinuria (GH) yang signifikan	PE. Ini adalah peningkatan tekanan darah yang terjadi setelah usia kehamilan 20 minggu tanpa adanya proteinuria. Ini bisa mewakili PE yang belum sempat mengembangkan proteinuria. Definisi lain adalah hipertensi transien, yaitu jika TD kembali ke nilai normal setelah 12 minggu persalinan, atau CH jika TD tetap ada.
Preeklampsia atau eklampsia (PE)	Preeklampsia adalah suatu sindrom yang ditandai dengan kelainan klinis <del>komponis</del> <sup>komponen</sup> berturut-turut. Sindrom ini dapat dinyatakan sebagai sindrom hipertensi, proteinuria, dan/atau berbagai gejala) atau sebagai sindrom janin, atau keduanya.  Ini terjadi pada 5% hingga 8% kehamilan dan merupakan penyebab utama kematian ibu dan perinatal di negara berkembang. Eklampsia adalah timbulnya kejang pada pasien dengan PE atau GH. Hal ini terkait dengan peningkatan kematian ibu dan sering diikuti oleh pendarahan otak. Meskipun penelitian besar di bidang ini, penyebabnya masih belum diketahui.

## AKU AKU AKU. KLASIFIKASI ID3 DAN NBTREE DALAM KESEHATAN

Pohon keputusan adalah model yang menggunakan pelatihan terawasi untuk klasifikasi dan prediksi data. Algoritma berbasis pohon keputusan adalah salah satu pendekatan inferensi yang paling populer. Beberapa bidang pengetahuan telah menggunakan teknik ini, seperti di bidang kesehatan, untuk diagnosis medis. Kishore *dkk.* mengusulkan varian algoritma ID3 dalam set data pasien anemia untuk mengurangi kompleksitas komputasi dan waktu yang dibutuhkan untuk membangun pohon keputusan menggunakan dataset besar [7]. Bashir *et al.* [8] membandingkan beberapa teknik klasifikasi menggunakan dua repositori data

benchmark untuk meningkatkan akurasi diagnosis diabetes menggunakan pohon keputusan. Hasil penelitian menunjukkan bahwa penggunaan algoritma ID3 bersama dengan teknik ansambel dapat meningkatkan kinerja pohon keputusan dalam hal akurasi. Gomathi dan Narayani [9] menggunakan teknik klasifikasi ID3 untuk memprediksi lupus eritematosus sistemik. Hasil menunjukkan bahwa algoritma mampu mengurangi kompleksitas dan meningkatkan kinerja komputasi.

Elaborasi algoritma ID3 didasarkan pada sistem inferensi dan konsep sistem pembelajaran. Untuk membangun pohon keputusan, pengklasifikasi ID3 memisahkan set pelatihan menjadi subset yang berisi contoh kelas tunggal. Algoritme membagi set ini melalui satu atribut yang dipilih dari properti statistik, yang disebut perolehan informasi, yang mengukur seberapa informatif atribut itu. Algoritma I menunjukkan langkah-langkah ini.

ALGORITMA I. PSEUDO-CODE UNTUK ALGORITMA ID3 .

ID3 Algoritma 1:
node LearnTree(contoh, targetAttribute, atribut) 2: mulai 3: jika semua contoh memiliki nilai targetAttribute yang sama,
kembalikan daun dengan nilai lain
4: jika set atribut kosong kembalikan daun
dengan nilai targetAttribute yang paling umum di antara contoh lain mulai
5:
6: A = atribut terbaik diantara atribut yang memiliki range nilai
v1, v2, ..., vk
7: Contoh partisi menurut nilainya untuk A menjadi himpunan S1, S2, ..., Sk
Buat node keputusan B dengan nilai A dan himpunan S sebagai anak
8:
9:
10:
11:
12:
13: Lampirkan B ke LearnTree(Si, targetAttribute, atribut –
(SEBUAH))
14: kembalikan B
15: Lampirkan B ke simpul daun dengan targetAttribute 16
paling umum: akhir 17: kembalikan simpul keputusan 18: akhir 19: akhir

Tingkat kemurnian suatu himpunan menentukan entropinya. Konsep ini mewakili ukuran kekurangan informasi. Persamaan (1) menunjukkan rumus perhitungan entropi yang diberikan himpunan S, dengan instance milik kelas i, dengan probabilitas pi.

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

Keuntungan informasi (IG) mendefinisikan pengurangan entropi.  $H(S) - H(S_A)$  berarti penurunan entropi yang diharapkan dalam S, diurutkan oleh atribut A. Berikut ini, persamaan (2) menghitung IG.

$$IG(A) = H(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} H(S_i)$$
 (2)

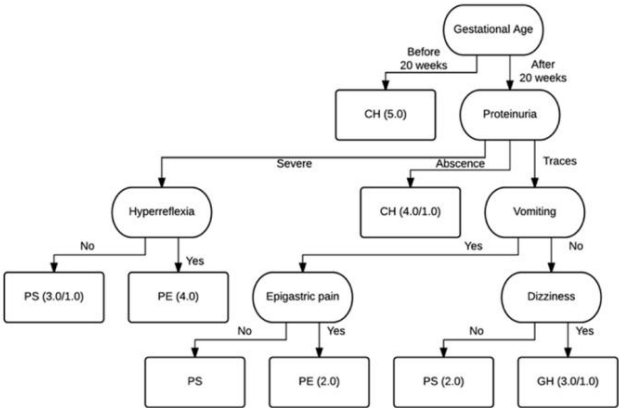
Pengklasifikasi NBTtree adalah model klasifikasi hibrida yang menggabungkan pohon keputusan dengan pengklasifikasi Naïve-Bayes. Node berisi divisi yang mempertimbangkan atribut tunggal, seperti pada

pohon keputusan biasa, tetapi simpul daun berisi pengklasifikasi berdasarkan jaringan Bayesian. Dalam pemilihan atribut kategorikal sebagai simpul pohon, setiap nilai atribut yang mungkin menerima sebuah keunggulan. Dalam hal atributnya numerik, menurut ambang (nilai yang membagi contoh), perlu untuk mewujudkan pembagian biner menggunakan teknik standar untuk meminimalkan entropi, ditunjukkan pada (2). Di bidang kesehatan, pendekatan ini dapat ditemukan dalam penelitian yang melibatkan pengoptimalan klasifikasi penyakit hati [10], kesejahteraan lansia [11], dan pada metode pemilihan protein es. Metode ini hadir dalam Algoritma II.

ALGORITMA II. PSEUDO-CODE UNTUK NBTREE CLASSIFIER .

NBTree Classifier [13] 1: untuk
setiap atribut Xi, evaluasi perolehan informasi IG(Xi), dari pemisahan pada atribut Xi
2: biarkan J = AttMax(Ui). Atribut dengan perolehan informasi tertinggi 3: jika Uj tidak lebih baik dari perolehan informasi dari simpul saat ini, buat pengklasifikasi Naïve Bayes untuk simpul saat ini dan kembalikan
4: Partisi T menurut analisis pada Xj. untuk semua nilai yang mungkin, lakukan pemisahan multi-arah 5: Untuk setiap node anak, panggil algoritme secara rekursif pada porsi T yang cocok dengan tes yang mengarah ke anak.

Gambar 1 dan 2 menunjukkan pohon keputusan yang dibuat untuk kedua metode menggunakan algoritma yang disajikan.

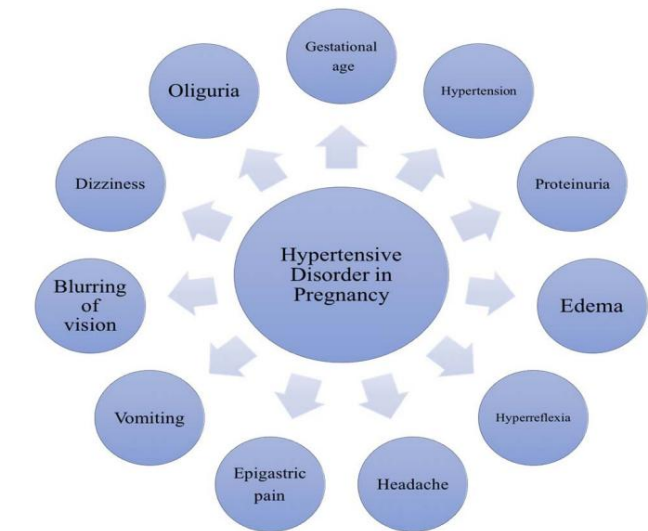


Gambar 1. Pohon keputusan yang dibuat berdasarkan algoritma ID3.

Karena ada sejumlah besar kemungkinan hipotesis untuk pembuatan pohon keputusan (Gbr. 1), penelitian ini memangkas pohon tersebut, menghilangkan atribut yang memiliki perolehan informasi mendekati nol, yaitu atribut yang tidak relevan secara jelas. Pertama, seluruh pohon dibuat dan, kemudian, dipangkas.

IV. EVALUASI KINERJA DAN ANALISIS HASIL

Studi evaluasi kinerja menggunakan database yang sama yang dipertimbangkan dalam Moreira *et al.* [14] menambahkan atribut "usia kehamilan". Studi ini juga membandingkan classifier pohon keputusan J48. Database berisi informasi dari 25 kasus simulasi ibu hamil yang menderita gangguan hipertensi dalam kehamilan. Dokter kebidanan/ginekolog berkolaborasi dengan pekerjaan ini membangun kumpulan data ini melalui pengetahuan dan pengalaman mereka untuk memvalidasi model yang diusulkan.



Gambar 2. Grafik yang merepresentasikan hubungan antar classifier NBTtree atribut.

A. Pengukuran metrik standar

Evaluasi kinerja penelitian ini menggunakan metode validasi silang 5 kali lipat [15]. Teknik ini membagi kumpulan data menjadi 5 subset (fold). Model belajar, mengecualikan data dari setiap subset, satu kali hingga menguji semua lipatan. Lipatan terpisah, ketika diterapkan pada model untuk menguji/mengevaluasi, memperkirakan kesalahan. Metrik untuk evaluasi kinerja juga mempertimbangkan matriks konfusi. Mengklasifikasikan seluruh kasus model ke dalam kategori memungkinkan konstruksinya, menentukan apakah nilai prediksi sesuai dengan nilai sebenarnya. Matriks konfusi adalah alat konvensional untuk evaluasi model statistik.

Jenis kesalahan klasifikasi yang paling signifikan adalah klasifikasi contoh negatif sebagai positif (false positive (FP)), juga dikenal sebagai alarm palsu (*misalnya*, mengidentifikasi kehamilan berisiko tinggi ketika wanita hamil itu sehat). Tipe kedua mengklasifikasikan kasus positif sebagai negatif (false negative (FN)), *misalnya*, mendiagnosis kehamilan berisiko tinggi sebagai orang yang sehat. Berdasarkan informasi ini, dimungkinkan untuk menghitung indikator kinerja, seperti presisi, daya ingat, ukuran-F, dan kurva ROC. Presisi menunjukkan persentase kasus positif yang diprediksi model sebagai negatif. Sementara penarikan mewakili tingkat yang mengukur jumlah kasus negatif, model menganggap positif. F-measure menghitung rata-rata harmonik antara daya ingat dan presisi. Ini mengukur efisiensi sistem mengingat kesalahan di kedua kelas. Persamaan (3) menghitung indikator ini.

$$\frac{2 \times \text{Presisi} \times \text{Daya Ingat}}{\text{Presisi} + \text{Daya Ingat}} \quad (3)$$

B. Hasil Eksperimen Tabel III

menunjukkan evaluasi kinerja metode yang diusulkan dibandingkan dengan algoritma berbasis pohon keputusan J48 yang terkenal.

Tabel III. NILAI PRESISI, RECALL, F-MEASURE, DAN ROC AREA DARI ALGORITMA ID3 DAN NBTREE UNTUK SETIAP KELAS .

Kelas Tingkat TP Tingkat FP Presisi Ingat ukuran-F						ROC Daerah
Algoritma ID3						
CH	0,6	0,105	0,6	0,6	0,6	0,697
PS	<b>0,333</b>	0,278	0,286	<b>0,333</b>	0,308	0,535
GH	0,333	<b>0,167</b>	0,4	0,333	0,364	0,588
pe	0,571	0,185	0,571	0,571	0,571	0,702
pengklasifikasi NBTtree						
CH	<b>0,667</b>	<b>0</b>	<b>1</b>	<b>0,667</b>	<b>0,8</b>	<b>0,904</b>
PS	<b>0,333</b>	<b>0,211</b>	<b>0,333</b>	<b>0,333</b>	<b>0,333</b>	<b>0,553</b>
GH	<b>0,667</b>	0,211	<b>0,5</b>	<b>0,667</b>	<b>0,571</b>	<b>0,719</b>
pe	<b>0,714</b>	<b>0,111</b>	<b>0,714</b>	<b>0,714</b>	<b>0,714</b>	<b>0,825</b>

Pengklasifikasi NBTtree menyajikan kinerja terbaik untuk semua metrik yang dipertimbangkan dan di semua kelas daripada algoritma ID3. Hanya indikator FP Rate (di mana nilai terbaiknya sekitar nol) untuk kelas PS, algoritma ID3 menunjukkan kinerja yang lebih baik. Tabel IV menunjukkan metrik standar rata-rata untuk semua metode berbasis pohon keputusan.

Tabel IV. NILAI METRIK STANDAR RATA-RATA METODE ID3, NBTREE , DAN J48 .

Metrik Standar ID3 NBTtree J48 Prec. 0,463 0,64 0,449			
Rek.	0,458	<b>0,6</b>	0,44
F-Ukur	0,459	<b>0,609</b>	0,441

Rata-rata, pengklasifikasi NBTtree juga berkinerja lebih baik. Hasil ini menunjukkan bahwa metode ini merupakan prediktor yang sangat baik untuk semua jenis gangguan hipertensi selama kehamilan. Algoritma ID3 menyajikan kinerja yang sedikit lebih baik daripada algoritma J48.

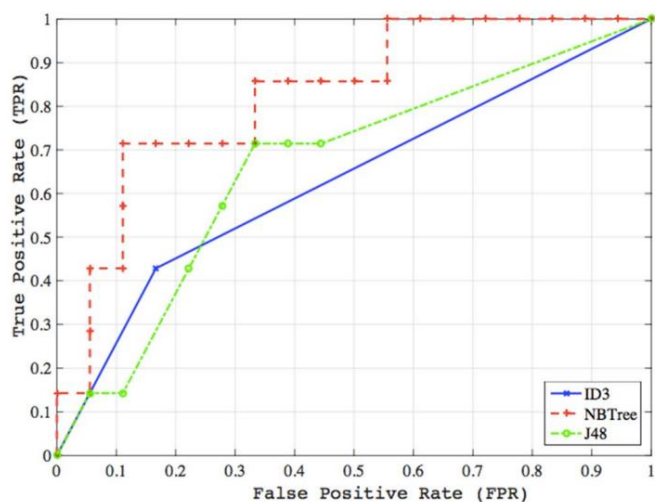
Kurva Receiver Operating Characteristic (ROC) mewakili metode grafis yang kuat, mewakili variasi spesifisitas (FP Rate) dan sensitibilitas (TP Rate).

Tabel V menunjukkan nilai area di bawah kurva ROC.

Tabel V. AREA ROC YANG MENCIRIKAN KINERJA MODEL KLASIFIKASI .

Metrik Standar ID3 NBTtree J48			
Tingkat TP	0,417	<b>0,56</b>	0,52
Tingkat FP	0,198	<b>0,145</b>	0,169
Daerah ROC	0,603	<b>0,758</b>	0,738

Gambar 3 menunjukkan hubungan antara tingkat FP dan tingkat TP dengan kurva ROC. Pengklasifikasi NBTtree menyajikan kurva ROC yang paling dekat dengan titik (0, 1) dibandingkan dengan metode lainnya. Kurva yang mendekati titik ini menunjukkan kinerja prediktif yang sangat baik.



Gambar 3. Kurva ROC untuk pendekatan berbasis pohon keputusan.

Statistik Kappa mengungkapkan ukuran kesepakatan yang digunakan dalam skala nominal yang memberikan gambaran tentang seberapa jauh perbedaan pengamatan dari yang diharapkan. Tabel IV menunjukkan statistik k untuk metode yang diusulkan.

Tabel VI. PERHITUNGAN STATISTIK KAPPA UNTUK KEANDALAN INTER-RATER .

Statistik Kappa	ID3	NBTee	J48
k	0,2186	0,4124	0,3548

Hasil menunjukkan bahwa pengklasifikasi NBTee sedang dalam hal keandalan dan pengklasifikasi ID3 dan J48 memberikan hasil yang dapat diterima.

## V. KESIMPULAN

Makalah ini menyajikan pengklasifikasi berbasis pohon yang dikenal sebagai ID3 dan NBTee, membandingkan perilakunya dengan pengklasifikasi pohon keputusan J48. Pengklasifikasi NBTee menunjukkan kinerja yang luar biasa untuk indikator ukuran-F (0,609), kinerja yang sangat baik di Area ROC (0,758), dan memberikan kinerja rata-rata dalam statistik Kappa (0,4124) dibandingkan dengan pendekatan lain yang disajikan.

Perspektif kerja lebih lanjut menunjukkan studi tentang metode klasifikasi baru untuk meningkatkan keandalan sistem, seperti jaringan syaraf tiruan. Algoritme berbasis pencarian lainnya juga perlu dievaluasi, seperti pohon keputusan dan regresi, dan aturan keputusan (antara lain algoritma OneR, Coverage, Top-Down, dan Bottom up). Pengklasifikasi Bayesian mewakili topik penting lainnya dalam masalah klasifikasi. Pendekatan ini telah menghadirkan banyak aplikasi dalam penelitian akademik dan industri.

Konsep e-Health adalah istilah yang relatif baru mengenai praktik perawatan kesehatan, tetapi solusinya dapat menjangkau wanita yang tidak dibantu, kebanyakan di negara berkembang di mana sebagian besar kematian ibu terjadi. Kontribusi yang disajikan pada makalah ini menegaskan pentingnya dan kegunaannya.

## UCAPAN TERIMA KASIH

Pekerjaan ini sebagian didukung oleh *Instituto de Telecomunicações*, Next Generation Networks and Applications Group (NetGNA), Portugal, oleh Pendanaan Nasional dari FCT - *Fundação para a Ciência e Tecnologia* melalui Proyek UID/EEA/50008/2013, oleh Finep, dengan sumber dari Funttel, Hibah No. 01.14.0231.00, di bawah proyek *Centro de Referência em Radiocomunicações - CRR* dari *Instituto Nacional de Telecomunicações* (Inatel), Brasil, dan oleh *Ciência sem Fronteiras* dari CNPq, Brasil, melalui nomor proses 207706 /2014 - 0.

## REFERENSI

- [1] DA Campbell, "Pembaruan tentang Tujuan Pembangunan Milenium Perserikatan Bangsa-Bangsa." *Jurnal Keperawatan Obstetri, Ginekologi & Neonatal*, vol. 46, tidak. 43, hlm. e48-e55, Mei-Jun. 2017.
- [2] L. Say, D. Chou, A. Gemmill, Ö. Tunçalp, AB Moller, J. Daniels, A. M. Gülmezoglu, M. Temmerman, dan L. Alkema, "Penyebab global kematian ibu: Analisis sistematis WHO," *Lancet Global Health*, vol. 2, tidak. 6, hlm. 323–333, Juni 2014.
- [3] C. Dye, JC Reeder, dan RF Terry, "Penelitian untuk Cakupan Kesehatan Universal," *WHO*, vol. 5, tidak. 199, hlm. 1–3, Agustus 2013.
- [4] J. Andreu-Perez, CC Poon, RD Merrifield, ST Wong, dan GZ, "Big data for health," *IEEE journal of biomedis and health informatics*, v. 19, no. 4, hal. 1193–1208, Juli 2015.
- [5] W. Fan dan A. Bifet, "Mining Big Data: Current Status, and Forecast to the Future," *ACM SIGKDD Exploration Newsletter*, vol. 14, tidak. 2, Desember 2012, hlm. 1–5.
- [6] LA Magee, A. Pels, M. Helewa, E. Rey, P. von Dadelszen, dan SOGC, "Diagnosis, Evaluasi, dan Pengelolaan Gangguan Hipertensi Kehamilan: Ringkasan Eksekutif," *Jurnal Obstetri dan Ginekologi Kanada*, vol. 36, tidak. 7, hlm. 575–576, Mei 2014.
- [7] CR Kishore, KP Rao, dan GRS Murthy, "Evaluasi Kinerja Entropi dan Gini menggunakan Threaded dan Non Threaded ID3 pada Dataset Anemia," *Konferensi Internasional Kelima tentang Sistem Komunikasi dan Teknologi Jaringan (CSNT 2015)*, Gwalior, MP, India, 4-6 April 2015, hlm. 1080–1084.
- [8] S. Bashir, U. Qamar, FH Khan, dan MY Javed, "Klasifikasi Diabetes Berbasis Aturan yang Efisien Menggunakan ID3, C4.5 & CART Ensembles," *Konferensi Internasional ke-12 tentang Perbatasan Teknologi Informasi (FIT 2014)*, Islamabad, Pakistan, 17-19 Desember 2014, hlm. 226–231.
- [9] S. Gomathi dan V. Narayani, "Systemic Lupus Erythematosus Manifestation using ID3 Algorithm – A Clinical Analysis," *2014 International Conference on Data Mining and Intelligent Computing (ICDMIC 2014)*, Delhi, India, Sep. 5-6, 2014, hlm. 1–5.
- [10] N. Novita dan T. Mantoro, "Data Mining Techniques For Optimatization of Liver Disease Clasification," *International Conference on Advanced Computer Science Applications and Technologies (ACSAT 2013)*, Sarawak, Malaysia, 22-24 Desember 2013, hlm. 379–384.
- [11] A. Cufoglu dan J. Chin, "Towards an understanding classification of well-being for care of elder people," *IEEE 13th International Conference on Industrial Informatics (INDIN 2015)*, Cambridge, United Kingdom, 22-24 Juli, 2013, hlm. 1494–1499.
- [12] J. Zhong, J. Wang, W. Peng, Z. Zhang, dan M. Li, "Metode Seleksi Fitur untuk Prediksi Protein Esensial," *Sains dan Teknologi Tsinghua*, vol. 20, tidak. 5, hlm. 491–499, Oktober 2015.
- [13] M. Hussein, "Menganalisis Algoritma Deteksi Intrusi NB, DT dan NBTee," *Jurnal Zankoy Sulaimani-Bagian A*, vol. 16, tidak. 1, hal. 69–76, Februari 2014.
- [14] MWL Moreira, JJPC Rodrigues, AMB Oliveira, K. Saleem, dan A. Neto, "Evaluasi Kinerja Pengklasifikasi Prediktif untuk Perawatan Kehamilan," *IEEE Global Communications Conference (GLOBECOM 2016)*, 4-8 Desember 2016, Washington, DC, AS, hlm. 1–5.
- [15] C. Bernau, M. Riester, A. Boulesteix, G. Parmigiani, C. Huttenhower, L. Waldron, dan L. Trippa, "validasi studi silang untuk penilaian algoritma prediksi," *Bioinformatika*, vol. 30, tidak. 12, hlm. 105–112, Juni 2014.