

DIAGNOSIS DINI PENYAKIT PREEKLAMPSIA PADA IBU HAMIL DENGAN METODE *K-NEAREST NEIGHBOR* (KNN)

Dwi Normawati^{1*}, *Rizki Akbari*², *Arfiani Nurhusna*³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknologi Industri, Universitas Ahmad Dahlan

e-mail: ¹dwi.normawati@tif.uad.ac.id, ²rizki1700018075@webmail.uad.ac.id, ³arfiani.khusna@tif.uad.ac.id

ABSTRAK

Menurut data Kementerian Kesehatan, di Indonesia Angka Kematian Ibu (AKI) masih cukup tinggi. Terdapat 4.999 kasus di tahun 2015, yang kemudian mengalami penurunan menjadi 4.912 kasus di tahun 2016, kemudian terjadi penurunan drastis menjadi 1.712 kasus di tahun 2017. Namun jumlah kita masih sangat tinggi dibandingkan negara-negara di kawasan ASEAN. Pre-eklampsia salah satu permasalahan kesehatan di Indonesia yang sering terjadi pada ibu hamil yang disebabkan oleh peningkatan tekanan darah yang diikuti dengan tanda gejala hipertensi dan peningkatan protein urin. Gejala Pre-eklampsia biasanya muncul pada usia kehamilan 20 minggu atau paling umum di usia 24-26 minggu. Tingginya angka kematian ibu serta banyaknya dampak terhadap kesehatan ibu memerlukan upaya pencegahan dan penanganan kasus-kasus preeklampsia serta melakukan analisis penyebab peningkatan hipertensi terhadap kematian ibu hamil disetiap tahun. Penelitian menggunakan metode *K-Nearest Neighbor* (KNN) yang memiliki kelebihan mudah diterapkan dan efisien dengan data yang banyak. Objek penelitian merupakan data rekam medis ibu hamil di RSUD Ratu Zalecha berjumlah 100 data dengan atribut pekerjaan, pendidikan, umur, usia kehamilan, tekanandarah, riwayat kehamilan hingga sekarang (G), paritas, riwayat abortus, riwayat persalinan, riwayat penyakit, proteinuria dan kelas yang berisi Pre-eklampsia atau tidak Pre-eklampsia. Penelitian ini menghasilkan sistem yang dapat mendiagnosis dini penyakit Pre-eklampsia pada ibu hamil. Pengujian pada metode *K-Nearest Neighbor* (KNN) menggunakan *confusion matrix* mendapatkan hasil akurasi 88%, *precision* 85,7%, *recall* 89% dan *f1-score* 85,4% ini menunjukkan bahwa performa algoritma KNN sangat baik dan layak digunakan untuk prediksi dini penyakit Pre Eklampsia pada Ibu Hamil. Sedangkan hasil pengujian sistem menggunakan *blackbox* menunjukkan bahwa sistem baik secara fungsionalitas sedangkan dengan metode *System Usability Scale* (SUS) nilai yang diperoleh sebesar 81,64 mendapat Grade B berarti sistem yang dibuat sudah cukup bagus dan layak untuk digunakan.

Kata Kunci: *Data Mining*, *K-Nearest Neighbor* (KNN), *Pre-eklampsia*, *Confusion Matrix*, *System Usability Scale* (SUS).

1. PENDAHULUAN

Bagi ibu hamil dan janinnya kesehatan menjadi faktor penting untuk menjamin keselamatan proses persalinan. Perlu pemantauan dan perhatian khusus terhadap upaya peningkatan kesehatan ibu hamil. Menurut data Kementerian Kesehatan, di Indonesia Angka Kematian Ibu (AKI) masih cukup tinggi. Terdapat 4.999 kasus di tahun 2015, yang kemudian mengalami penurunan menjadi 4.912 kasus di tahun 2016, kemudian terjadi penurunan drastis menjadi 1.712 kasus di tahun 2017. Namun jumlah kita masih sangat tinggi dibandingkan negara-negara di kawasan ASEAN [1]. Penyakit pendarahan, hipertensi dalam kehamilan, infeksi merupakan faktor utama penyebab kematian ibu terbanyak pada tahun 2019 [2]. Hipertensi dalam kehamilan merupakan gejala utama dari penyakit Pre-eklampsia. Pre-eklampsia merupakan gangguan dimasa kehamilan dengan gejala hipertensi dan tanda-tanda kerusakan organ, seperti kandungan protein tinggi dalam urin (*proteinuria*) yang mengindikasikan kerusakan ginjal. Gejala biasanya muncul saat usia kehamilan mencapai 20 minggu atau usia kehamilan yang paling umum adalah 24-26 minggu [3]. Preeklampsia adalah faktor utama penyebab terjadinya eklampsia yang mengancam kehidupan ibu bersalin [4].

Faktor penyebab Preeklampsia pada ibu hamil dilihat dari status kesehatan ibu, status reproduksi, status gizi, akses ke pelayanan kesehatan dan perilaku kesehatan pada ibu hamil. Perilaku kesehatan pada ibu hamil merupakan suatu respon terhadap terjadinya rangsangan. Perilaku kesehatan terdiri dari 3 aspek salah satunya adalah perilaku gizi atau perilaku makan yang dilakukan ibu hamil selama masa kehamilan. Ada 2 dampak yang dapat dialami ibu hamil yang menderita Penyakit Preeklampsia yaitu dampak bagi janin dan bagi ibu sendiri. Dampak bagi janin menyebabkan gangguan pertumbuhan janin dan kematian janin dalam kandungan, sedangkan dampak bagi ibu hamil dapat menyebabkan kondisi ibu menjadi tidak sadar (koma) sampai meninggal dunia. Pengaruh preeklampsia terhadap tingginya AKI dan banyaknya dampak terhadap kesehatan ibu, maka diperlu dilakukan upaya pencegahan dan penanganan kasus-kasus Preeklampsia dan melakukan analisis penyebab penyebab peningkatan hipertensi terhadap kematian ibu hamil di Indonesia setiap tahunnya [5].

Metode diagnosis *computer aided* banyak memanfaatkan metode *data mining* dengan memanfaatkan database riwayat rekam medis pasien. *Data mining* dengan algoritma cerdas dapat digunakan untuk mengatasi masalah diagnosis dengan *dataset* medis yang melibatkan beberapa inputan data [6]. *Data mining* dalam dunia medis memiliki potensi besar untuk mengetahui pola tersembunyi pada *dataset* medis. Pola tersebut dapat dimanfaatkan untuk membantu dokter secara signifikan meningkatkan kualitas keputusan medis untuk mengungkapkan ada atau tidaknya suatu penyakit [7].

Pada penelitian ini mencoba mengimplementasikan metode datamining KNN untuk membangun suatu sistem diagnosis dini Penyakit Pre-eklamsia pada ibu hamil dengan menggunakan data pasien ibu hamil sebanyak 100 record dimana data didapat dari RSUD Ratu Zalecha, Kec. Martapura, Banjar, Kalimantan Selatan, dengan harapan sistem dapat membantu para tenaga medis mengambil keputusan dini yang tepat untuk tindakan yang tepat dan lebih lanjut pada para pasien ibu hamil penderita Pre-eklamsia, dengan menggunakan bahasa pemrograman *Python*.

2. METODE PENELITIAN

Pada penelitian ini menggunakan metode klasifikasi datamining yaitu KNN. Penelitian ini dilakukan dengan beberapa tahapan yaitu Studi Literatur, Pengumpulan data, Preprocessing data, Split Data, Implementasi metode KNN dan Pengujian Sistem.

2.1 Studi Literatur

Pada tahapan ini adalah tahap mengumpulkan, mencari dan mempelajari literature berupa jurnal-jurnal, e-book dan buku-buku kesehatan tentang penyakit Pre-eklamsia untuk mendukung penelitian ini.

Pada penelitian pertama melakukan klasifikasi terhadap kehamilan resiko tinggi pada ibu hamil menggunakan metode *naive bayes*. Pada penelitian ini dipilih beberapa *class* antara lain umur ibu hamil, umur kehamilan, tekanan darah *sistole* dan *diastole*, serta *proteinuria*. Sistem yang dibuat mampu menunjukkan hasil prediksi dalam risiko kehamilan ibu hamil tersebut mengalami preeklamsia ringan atau preeklamsia berat. Hasil klasifikasi didapatkan akurasi 95% dari 100 *data training* yang dimasukkan ke dalam sistem [8].

Penelitian kedua mengklasifikasi kehamilan beresiko menggunakan metode *K-Nearest Neighbor* (KNN). Sistem dari penelitian ini menghasilkan akurasi 93% dengan nilai $k=5$. Kelemahan metode ini terdapat pada penentuan nilai K . Jika K berupa angka genap maka menghasilkan nilai ganda sehingga tidak bisa membandingkan hasil prediksi terbaik berdasarkan pengurutan hasil terkecil jarak *euclidean* [9].

Penelitian lain memprediksi penyakit hipertensi pada kehamilan menggunakan metode *decision tree* dengan algoritma *c4.5*. Penelitian ini menggunakan *data training* berjumlah 286 *instances* menghasilkan kesalahan 7,3427% dan akurasi 92,6573%. Kelebihan metode *decision tree* adalah kesederhanaan dalam algoritmanya namun menghasilkan klasifikasi dan prediksi yang sangat kuat. Kelemahan penelitian ini yaitu belum dibuat ke dalam sistem atau *prototype* namun masih berupa model *data mining* [10].

Penelitian yang ditulis oleh [11] dengan judul “Analisis Pola Asosiasi Data Penyakit Hipertensi pada Ibu Hamil Menggunakan Algoritma *Apriori* (Studi Kasus: RSIA Melinda Kediri)”, menggunakan 10 data kriteria yaitu usia produktif, usia lanjut, riwayat persalinan normal, riwayat persalinan cesar, riwayat kehamilan sehat, riwayat kehamilan keguguran, riwayat diabetes, riwayat diabetes, trombosit normal, dan trombosit kurang dengan 372 data diagnosis pasien membutuhkan waktu 15 menit dalam proses analisis. Hasil penelitian menunjukkan tiga kombinasi dengan kombinasi 1 menghasilkan nilai 79,62% dengan kriteria usia produksi, kombinasi 2 sebesar 61,13% dengan kriteria kehamilan sehat dan trombosit normal, dan kombinasi 3 sebesar 47,45% dengan kriteria kehamilan sehat, trombosit normal, serta usia produksi.

Penelitian lainnya yang ditulis oleh [12] dengan judul “Metode *Fuzzy ID3* untuk Klasifikasi Status Preeklamsi Ibu Hamil”. Penelitian ini menggunakan algoritma *Fuzzy ID3* yang diterapkan dalam proses klasifikasi status preeklamsi pada ibu hamil berdasar atribut tekanan darah sistolik, tekanan darah diastolik, usia ibu, usia kehamilan, protein urine, dan odema. Atribut yang menjadi *root* atau akar dari pohon keputusan dalam penelitian ini adalah protein urine, sebagai parameter penentu utama yang ditentukan untuk klasifikasi status preeklamsi ibu hamil. Berdasarkan hasil uji coba, *k-fold 5* mempunyai hasil akurasi yang tertinggi yaitu 98,44%, presisi terbesar 96,66%, dan *recall* terbesar 97,61% dengan 17 aturan, sehingga menjadi aturan terbaik yang dapat digunakan untuk klasifikasi status preeklamsi ibu hamil. *K-fold 5* dengan aturan 4 *fold* sebagai *data training* dan 1 *fold* sebagai *data testing*, mempunyai *data training* yang paling banyak dibanding *k-fold 1* sampai *k-fold 4*. Hal ini menunjukkan bahwa semakin besar *data training* maka semakin tinggi akurasi klasifikasi. Penelitian ini dapat dikembangkan dengan melakukan proses pemangkasan pohon (*pruning*) pada *tree* yang terbentuk. *Pruning* dapat dilakukan juga menggunakan *pre pruning* dan *post pruning* untuk menentukan akurasi terbaik

2.2 Pengumpulan Data

Pada tahapan ini adalah proses mengumpulkan data rekam medik pasien ibu hamil. Data yang didapat sebanyak 100 *record* data. Data diambil dari data rekam medik pasien ibu hamil tahun 2019 di RSUD Ratu Zalecha, Kec. Martapura, Banjar, Kalimantan Selatan. Atribut yang digunakan sebanyak sembilan atribut, yaitu data usia

ibu, usia kehamilan, tekanan darah, riwayat persalinan, jenis persalinan, paritas, riwayat abortus, riwayat penyakit, proteinuria, serta 2 class, yaitu Pre-eklamsia dan Tidak Pre-eklamsia.

2.3 Preprocessing Data

Tahap ini adalah tahap untuk mengolah data yang sudah didapatkan untuk membuat sistem. Datayang akan diolah merupakan data rekam medik ibu hamil yaitu pekerjaan, pendidikan, umur, usiakehamilan, tekanan darah, riwayat kehamilan hingga sekarang (G), paritas, riwayat abortus, riwayat persalinan, riwayat penyakit, proteinuria dan kelas yang terbagi menjadi kelas positif Pre-eklamsia dan kelas normal.

Proses Preprocessing Data yang dilakukan pada penelitian ini yaitu pertama melakukan Seleksi Data dimana dilakukan penghapusan atribut yang dianggap tidak penting untuk proses selanjutnya, yang kedua melakukan Transformasi data yaitu merubah data kategori menjadi data neumerik berdasarkan interval tertentu dengan ketentuan berdasarkan hasil wawancara dengan pakar yaitu salah satu bidan di Puskesmas Rembang. Terdapat 10 atribut yang akan di transformasi yaitu pendidikan, pekerjaan, usia ibu, usia kandungan, paritas, riwayat abortus, riwayat persalinan, riwayat penyakit, pretoinuria dan kelas. Contoh pengolahan data adalah mengubah data kategorikal menjadidata numerik agar bisa diproses oleh metode KNN.

2.4 Split Data

Setelah preprocessing data yang dihasilkan kemudian dilakukan split data untuk proses selanjutnya. Penelitian ini menggunakan metode *K-fold cross validation*. *K-fold cross validation* merupakan metode untuk membagi data sekaligus untuk validasi. Validasi dengan *cross validation* adalah tahap uji preforma dari suatu algoritma dengan cara mengulang kembali proses perhitungan berdasarkan banyaknya *k-fold*. *Cross validation* adalah proses membagi *dataset* menjadi data pelatihan dan data pengujian. Banyak peneliti menggunakan metode ini dengan memecah data menjadi 10 bagian, 90% digunakan untuk data latih dan sisanya digunakan sebagai data uji. Proses ini diulang hingga 10 kali sampai semua *record* data mendapatkan bagian menjadi data uji. Banyak peneliti yang menggunakan proses ini karena performa dari algoritma yang digunakan terbukti menjadi lebih stabil [13].

2.5 Implementasi Metode KNN

Tahapan ini merupakan tahapan dimana metode *K-nearest Neighbor* (KNN) diimplementasikan pada sistem dengan menggunakan bahasa pemrograman *Python*. KNN merupakan teknik supervised learning yang memanfaatkan teknik pembelajaran berbasis kesamaan. Algoritma non-parametrik ini adalah algoritma pembelajaran berbasis instance yang banyak digunakan yang secara akurat memprediksi kasus uji. Algoritma KNN menetapkan kelas sampel baru dengan menghitung jarakantara sampel dalam sampel saat ini (*k* - tetangga terdekat) dan nilai *k* sampel yang diberikan. Iamenemukan *k* sampel "dekat" dengan memanfaatkan ukuran jarak seperti Euclidean, Manhattan, Chebyshev, dan Levenshtein [14]. Metode KNN dipilih karena mudah untuk diimplementasikan, selain itu metode KNN memiliki sifat yang sangat non-linier sehingga efektif untuk kasus data dengan banyak atribut.

Tahapan klasifikasi pada KNN [15], yaitu:

1. Menentukan parameter *k*.
2. Menghitung jarak antara data uji dan data latih.
3. Mengurutkan jarak yang terbentuk (diurutkan naik).
4. Menentukan jarak terdekat sampai urutan *k*.
5. Pasangkan kelas yang bersesuaian.
6. Tentukan jumlah kelas dari tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelasdata yang diuji.

Rumus untuk algoritma KNN adalah sebagai berikut:

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (1)$$

Penjelasan:

x = contoh data

y = data pengujian

d = jarak terdekat

Kasus dalam dunia nyata menggunakan lebih dari 2 independent variable sehingga perlu menggunakan rumus lain seperti euclidean distance ,rumus ini mirip seperti rumus pada 2.1 hanya saja memiliki dimensi lebih dari 2:

$$d_i = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Penjelasan:

x = contoh data

y = data pengujian

i = indeks data ke-i

n = jumlah data latih

di = jarak terdekat berdasarkan indeks ke-i

2.6 Pengujian Sistem

Tahap pengujian sistem dilakukan untuk mengevaluasi sistem yang akan dibuat baik dari sisi algoritma dan dari sisi sistem itu sendiri. Metode pengujian yang digunakan dalam penelitian ini ada 3 yaitu *confusion matrix*, *blackbox testing* dan *system usability scale* (SUS).

2.6.1 Confusion Matrix

Confusion matrix dilakukan untuk mengukur performa dari algoritma KNN yang diimplementasikan pada sistem. *Confusion matrix* merupakan suatu cara untuk mengetahui seberapa baik hasil klasifikasi, kelas yang diprediksi berada dibagian atas sedangkan kelas yang diamati berada dibagian kiri. Setiap isi dari kolom menunjukkan jumlah kasus dari kelas yang diprediksi seperti terlihat pada Tabel 1 [16].

Tabel 1 *Confusion Matrix*

Hasil Prediksi	Hasil Sebenarnya	
	TP	FN
FP	TN	

Tabel 1 merupakan tabel *Confusion Matrix*. Komponen tabel *Confusion Matrix* berisi nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP) dan *False Negative* (FN). *True Positive* (TP), yaitu jumlah dokumen dari kelas 1 yang benar dan diklasifikasikan sebagai kelas 1. *True Negative* (TN), yaitu jumlah dokumen dari kelas 0 yang benar diklasifikasikan sebagai kelas 0. *False Positive* (FP), yaitu jumlah dokumen dari kelas 0 yang salah diklasifikasikan sebagai kelas 1. Dan *False Negative* (FN) yaitu jumlah dokumen dari kelas 1 yang salah diklasifikasikan sebagai kelas 0 [16].

Dari tabel *Confusion Matrix* dilakukan perhitungan evaluasi performa algoritma KNN dengan menghitung nilai Akurasi, *Precision*, *Recall*, dan *F1-Score*. Akurasi adalah proporsi total dari prediksi semua data yang benar. *Precision* adalah ukuran keakuratan hasil model. Persamaan *precision* adalah perbandingan antara *true positive* dengan total data dengan label *positive*. *Recall* adalah ukuran integritas model. Persamaan *recall* adalah perbandingan antara *true positive* terhadap total contoh yang benar-benar *positive* [17].

2.6.2 Blackbox

Black-box Testing merupakan pengujian sistem yang fokus pada fungsional sistem tanpa memperhatikan kode program yang digunakan. Terdapat beberapa teknik untuk *black-box testing* dan teknik yang akan digunakan dalam penelitian ini adalah *Equivalent Partitions*. Teknik *Equivalent Partitions* merupakan pengujian berdasarkan fungsi input. Setiap fungsi inputan pada sistem akan dikelompokkan berdasarkan nilainya yaitu valid atau tidak valid. Hasil pengujian yang tidak sesuai akan diperbaiki sehingga sistem terbebas dari kesalahan dan memenuhi persyaratan yang ditetapkan [18].

2.6.3 System Usability Scale (SUS)

System Usability Scale (SUS) merupakan salah satu metode untuk pengujian kelayakan *software*. Caranya dengan cara membiarkan pengguna menguji coba sistem kemudian diberikan kuisioner yang terdiri dari 10 pernyataan dan 5 jawaban dengan skala 1-5 atau sangat tidak setuju sampai sangat setuju [19].

3. HASIL DAN PEMBAHASAN

3.1 Preprocessing Data

Tahap preprocessing data merupakan tahap awal untuk mempersiapkan data sebelum melakukan perhitungan dengan algoritma KNN. Proses Preprocessing Data yang dilakukan yaitu Seleksi Data dan Transformasi Data.

Yang pertama dilakukan adalah Seleksi Data yaitu menghapus atribut jenis kehamilan karena memiliki nilai yang sama. Kemudian merubah atribut kelas yang awalnya terdapat tiga kelas yaitu normal (n), Pre-eklampsia (pe) dan Pre-eklampsia berat (peb) menjadi dua kelas sehingga hasilnya menjadi normal (n) dan Pre-eklampsia (pe). Hasil Seleksi Data terlihat pada Tabel 1.

Tabel 2 Hasil Seleksi Data

Pendidikan	Pekerjaan	Usia	Usia Kandungan	Tekanan Darah		G	Paritas	Riwayat Abortus	Riwayat Persalinan	Riwayat Penyakit	Proteinuria	kelas
SLTA	SWASTA	31	37	90	60	1	0	0	belum	tidak ada	POSITIF 1	pe
SD	IRT	22	40	143	86	1	0	0	belum	tidak ada	NEGATIF	pe
SMA	PELAJAR	18	38	137	99	1	0	0	spt	tidak ada	POSITIF 1	pe
UNIVERSITAS	IRT	35	37	190	130	3	2	0	sc	tidak ada	POSITIF 1	pe
SD	IRT	19	35	142	79	2	1	0	sc	tidak ada	POSITIF 1	pe
SD	SWASTA	16	39	162	99	1	0	0	belum	tidak ada	NEGATIF	pe
SMP	IRT	24	36	165	95	1	0	0	sc	tidak ada	NEGATIF	pe
LAIN2	IRT	19	42	110	60	1	0	0	belum	HT	POSITIF 1	n
SMA	IRT	19	41	118	79	1	0	0	belum	ASMA	NEGATIF	n
UNIVERSITAS	GURU	27	42	163	108	1	0	0	belum	tidak ada	NEGATIF	n

Proses transformasi merubah data kategori menjadi data numerik berdasarkan interval tertentu dengan ketentuan berdasarkan hasil wawancara dengan pakar yaitu salah satu bidan di Puskesmas Rembang. Terdapat 10 atribut yang akan di transformasi yaitu pendidikan, pekerjaan, usia ibu, usia kandungan, paritas, riwayat abortus, riwayat persalinan, riwayat penyakit, proteinuria dan kelas. Aturan Transformasi data penelitian ini terdapat pada Tabel 2. Sedangkan Tabel 3 adalah Hasil Transformasi Data berdasarkan Aturan Tabel 2.

Tabel 3 Aturan Transformasi Data

Atribut	Transformasi	Keterangan
Pendidikan	Pendidikan ditransformasi menjadi 0 jika pendidikan rendah dan 1 jika pendidikan tinggi	0 : Pendidikan Rendah \leq SLTA/SLTP/SMA 1 : Pendidikan Tinggi $>$ SLTA/SLTP/SMA
Pekerjaan	Pekerjaan ditransformasi menjadi 0 jika tidak bekerja dan 1 jika bekerja	0 : IRT, Mahasiswa, Pelajar 1 = Guru, Apoteker, PNS
Usia Ibu	Usia ibu ditransformasi menjadi 0 jika berusia antara 20 sampai 35 tahun dan 1 jika kurang atau lebih dari itu	0 : 20-35 1 : $<$ 20 atau $>$ 35
Usia Kandungan	Usia kandungan ditransformasi menjadi 0 jika usia kandungan 27 minggu atau kurang dari itu dan 1 jika lebih dari 27 minggu	0 : \leq 27 minggu 1 : $>$ 27 minggu
Paritas	Paritas ditransformasi menjadi 0 jika jumlah paritas 2 atau 3 selain itu akan diberi nilai 1	0 : 2 atau 3 1 : $<$ 2 atau $>$ 3

Riwayat Abortus	Riwayat abortus ditransformasi menjadi 0 jika belum pernah dan 1 jika pernah melakukan aborsi, berapapun jumlah aborsinya akan ditransformasi menjadi 1	0 : belum pernah 1 : jika pernah
Riwayat Persalinan	Riwayat persalinan ditransformasi menjadi 0 jika belum pernah, 1 jika persalinan terakhir adalah SC, 2 jika persalinan adalah SPT dan 3 jika persalinan adalah VE	0 : belum 1 : SC (Sectio Caesarea) 2 : SPT (Spontan) 3 : VE (Ekstraksi Vacum)

Tabel 4 Hasil Transformasi Data

Pendidikan	Pekerjaan	Usia	Usia Kandungan	Tekanan Darah Sistolik	Tekanan Darah Diastolik	G	Paritas	Riwayat Abortus	Riwayat Persalinan	Riwayat Penyakit	Proteinuria	kelas
0	1	0	1	90	60	1	1	0	0	0	2	1
0	0	0	1	143	86	1	1	0	0	0	0	1
0	0	1	1	137	99	1	1	0	2	0	2	1
1	0	0	1	190	130	3	0	0	1	0	2	1
0	0	1	1	142	79	2	1	0	1	0	2	1
0	1	1	1	162	99	1	1	0	0	0	0	1
0	0	0	1	165	95	1	1	0	1	0	0	1
0	0	1	1	110	60	1	1	0	0	1	2	0
0	0	1	1	118	79	1	1	0	0	1	0	0
1	1	0	1	163	108	1	1	0	0	0	0	0

3.2 Split Data

Tahap *split data* atau membagi data menggunakan *k-fold cross validation* untuk membagi data dan juga mengevaluasi hasil dari algoritma yang digunakan. Data akan dibagi sebanyak jumlah k 'subset' dengan jumlah data yang sama, pada penelitian ini akan menggunakan nilai 10 sebagai nilai *kfold*, sehingga data yang berjumlah 100 akan dibagi 10 sehingga menjadi 10 *fold* dan dari 1-*fold* sampai 10-*fold* bagian tadi akan bergantian menjadi data latih dan data uji agar mendapatkan hasil validasi terhadap algoritma KNN seperti pada Gambar 3 Hasil *K-Fold Cross Validation*.

	Iterasi 1	Iterasi 2	Iterasi 3	Iterasi 4	Iterasi 5	Iterasi 6	Iterasi 7	Iterasi 8	Iterasi 9	Iterasi 10		
Subset 1	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		Data Uji
Subset 2	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		Data Latih
Subset 3	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		
Subset 4	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		
Subset 5	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		
Subset 6	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		
Subset 7	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		
Subset 8	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		
Subset 9	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		
Subset 10	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100		

Gambar 1 Hasil *K-Fold Cross Validation*

3.3 Pengujian

Penelitian ini melakukan 3 (tiga) pengujian yaitu pengujian *Confusion Matriks*, *Blackbox Testing* dan *System Usability Scale (SUS)*.

3.3.1 Confusion Matriks

Pengujian *confusion matrix* bertujuan untuk mengevaluasi metode yang digunakan terhadap dataset. Berdasarkan hasil pengujian dari sistem didapatkan hasil bahwa nilai $K = 5$ memiliki tingkat akurasi paling yang besar. Hasil pengujian confusion matrix dapat dilihat pada Tabel 2 Hasil Pengujian Confusion Matrix.

Tabel 5 Hasil *Confusion Matrix*

Iterasi Ke	Akurasi	Presisi	Recall	F1-Score
1	70,0	71,0	83,0	77,0
2	90,0	50,0	100,0	67,0
3	90,0	100,0	67,0	80,0
4	80,0	67,0	100,0	80,0
5	100,0	100,0	100,0	100,0
6	90,0	86,0	100,0	92,0
7	90,0	100,0	67,0	80,0
8	100,0	100,0	100,0	100,0
9	80,0	83,0	83,0	83,0
10	90,0	100,0	90,0	95,0
Rata-rata	88,0	85,7	89,0	85,4

3.3.2 Blackbox Testing

Pengujian dilakukan secara langsung mencoba aplikasi yang telah dibuat apakah hasil sesuai dengan yang diinginkan. Hasil pengujian dapat dilihat pada Tabel 3 Hasil *Blackbox Testing*.

Tabel 6 Hasil *Blackbox Testing*

No	Rincian Pengujian	Hasil yang Diinginkan	Hasil yang Didapatkan	Keterangan
1	Membuka aplikasi	Menampilkan tabel data rekam medis	Menampilkan tabel data rekam medis	Sesuai
2	Mengosongkan file kemudian klik <i>Submit</i>	Menampilkan pesan "Tidak ada file yang dipilih"	Menampilkan pesan "Tidak ada file yang dipilih"	Sesuai
3	Mengisi file dengan ekstensi selain file csv kemudian klik <i>Submit</i>	Menampilkan pesan "File harus berekstensi csv!"	Menampilkan pesan "File harus berekstensi csv!"	Sesuai
4	Mengisi file dengan ekstensi csv kemudian klik <i>Submit</i>	Menampilkan pesan "File berhasil ditambahkan ke database"	Menampilkan pesan "File berhasil ditambahkan ke database"	Sesuai
5	Mengosongkan nilai K kemudian klik Latih	Menampilkan pesan "Nilai K tidak boleh kosong!"	Menampilkan pesan "Nilai K tidak boleh kosong!"	Sesuai
6	Mengisi nilai K dengan angka "5" kemudian klik Latih	Menampilkan hasil akurasi dan <i>confusion matrix</i> dari perhitungan KNN berdasarkan nilai K yang diisi sebelumnya	Menampilkan hasil akurasi dan <i>confusion matrix</i> dari perhitungan KNN berdasarkan nilai K yang diisi sebelumnya	Sesuai
7	Mengosongkan bagian pendidikan, pekerjaan, riwayat persalinan, riwayat penyakit dan proteinuria kemudian klik <i>submit</i>	Menampilkan pesan "Please select an item in the list"	Menampilkan pesan "Please select an item in the list"	Sesuai
8	Mengosongkan bagian usia, usia kandungan, tekanan darah sistolik, tekanan darah diastolik, jumlah kehamilan hingga sekarang, paritas, riwayat abortus kemudian klik <i>submit</i>	Menampilkan pesan "Please fill out this field"	Menampilkan pesan "Please fill out this field"	Sesuai
9	Mengisi form data rekam medis dengan isi pendidikan "SLTA", pekerjaan "IRT", usia "27", usia kandungan "37", tekanan darah sistolik "135", tekanan	Menampilkan pesan untuk hasil prediksi	Menampilkan pesan "Data ditambahkan ke database", pesan	Sesuai

darah diastolik “88”, jumlah kehamilan hingga sekarang “1”, paritas “0”, riwayat abortus “0”, riwayat persalinan “belum”, riwayat penyakit “tidak ada”, proteinuria “TRACE” kemudian klik submit		untuk prediksi, dan hasil prediksinya	
--	--	---------------------------------------	--

Setelah dilakukan pengujian black-box seperti pada Tabel 3, selanjutnya dapat dilakukan perhitungan persentase kelayakan aplikasi dengan menggunakan rumus berikut:

$$\begin{aligned} \text{Presentase Kelayakan} &= \frac{\text{Nilai Hasil Pengujian}}{\text{Nilai Maksimal}} \times 100\% \\ &= \frac{9}{9} \times 100\% \\ &= 100\% \end{aligned}$$

3.3.3 System Usability Scale (SUS)

Responden akan menjalankan sistem dan kemudian diberi 10 pertanyaan dengan 5 pilihan jawaban. Pertanyaan diserahkan melalui *google form*, hasil jawaban seluruh responden dapat dilihat pada Tabel 4 Hasil Pengujian SUS.

Tabel 7 Hasil Pengujian SUS

NO	1	2	3	4	5	6	7	8	9	10	Jumlah	(x 2,5)
1	4	2	4	2	4	2	4	1	5	5	33	82.5
2	4	2	3	2	3	2	3	2	3	3	27	67.5
3	3	2	4	1	5	2	5	1	4	2	29	72.5
4	3	2	3	2	3	3	3	2	3	4	28	70
5	5	1	5	2	5	1	5	1	5	4	34	85
6	3	2	4	3	3	2	3	2	3	3	38	70
7	5	3	5	4	5	2	4	2	5	4	39	97.5
8	4	2	4	3	3	2	5	2	3	4	32	80
9	5	2	5	2	5	2	5	1	5	4	36	90
10	3	3	5	5	2	4	1	4	2	4	33	82.5
11	4	4	4	4	4	4	4	4	4	4	40	100
12	5	3	5	2	5	3	4	2	4	3	36	90
13	4	2	4	3	4	2	3	3	4	2	31	77.5
14	4	2	5	3	5	1	4	2	4	3	33	82.5
15	5	2	5	3	5	2	5	2	5	4	38	95
16	4	2	4	2	4	1	4	3	4	2	30	75
17	4	1	5	1	4	1	4	1	4	2	27	67.5
18	4	2	4	2	5	3	4	2	4	2	32	80
19	4	2	5	2	4	1	5	1	4	4	32	80
20	5	5	5	4	4	2	4	4	5	5	43	107.5
21	5	2	5	1	5	2	4	1	5	2	32	80
22	3	2	4	2	4	2	4	2	4	2	29	72.5
23	5	1	4	2	5	1	4	1	5	1	29	72.5
24	5	1	4	2	5	2	3	1	4	1	28	70
25	5	2	5	3	5	2	5	2	5	2	36	90
26	3	2	3	3	3	2	3	3	3	3	28	70
27	5	4	4	4	4	3	3	3	4	4	38	95
28	4	1	5	1	4	2	4	2	4	2	29	72.5
29	4	3	3	3	3	3	3	2	4	5	33	82.5
30	5	1	5	1	5	3	1	1	5	1	28	70
31	5	4	4	4	4	4	4	4	5	4	42	105

32	3	2	4	2	4	2	4	2	4	3	30	75
33	3	3	3	4	4	2	2	2	1	4	28	70
34	4	2	5	2	5	2	4	2	4	3	33	82.5
35	5	1	5	5	5	2	3	4	5	4	39	97.5
Jumlah											2857.5	
Rata-rata											81.64	

Berasarkan Tabel 4 Hasil Perhitungan SUS, nilai yang didapatkan dengan 35 responden adalah 81,64, artinya sistem yang telah dibuat sudah cukup bagus dan diinterpretasikan kedalam *Grade B* yang berarti termasuk kedalam kategori *Good-Exelent* secara *adjective* dan *acceptable* secara *acceptability ranges*.

4. KESIMPULAN

Metode KNN yang diimplementasikan kedalam sistem untuk mendiagnosis secara dini penyakit pre-eklampsia pada ibu hamil dengan jumlah data 100 dan 11 atribut terbukti dapat melakukan klasifikasi dengan baik dengan nilai $k = 5$. Berdasarkan pengujian evaluasi dari metode *confusion matrix* dan validasi *cross validation* dengan $k\text{-fold} = 10$ dengan nilai rata-rata akurasi 88%, rata-rata *precision* 85.7%, rata-rata *recall* 89% dan rata-rata *f1-score* 85.4% ini menunjukkan bahwa performa algoritma KNN sangat baik dan layak untuk memprediksi diagnosis Penyakit Pre-eklamsia pada Ibu Hamil.

Pengujian sistem dengan *blackbox testing* mendapatkan hasil 100% sedangkan dengan metode system usability scale terhadap 35 responden mendapatkan hasil rata-rata 81.64 atau *Grade B*. Artinya sistem yang telah dibuat dapat diterima dari segi fungsional dan *usabilisy* serta layak untuk digunakan masyarakat.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada **LPPM UAD** yang telah memberi “dukungan finansial” terhadap penelitian ini.

DAFTAR PUSTAKA

- [1] A. Utantoro, “Angka Kematian Ibu di Indonesia Masih Tinggi,” *Media Indonesia*, 2019. .
- [2] G. Prabhakara, *Health Statistics (Health Information System)*. 2010.
- [3] Dr. Marianti, “Preeklamsia,” *ALODOKTER*, 2017. .
- [4] D. Magetan, *Rentra Dinas Kesehatan Kabupaten Magetan 2018-2023*. 2018.
- [5] C. H. Transyah, “Hubungan Umur Dan Paritas Ibu Bersalin Dengan Kejadian Pre-Eklampsia,” *Hum. Care J.*, vol. 3, no. 1, 2018.
- [6] R. A. Pramunendar, I. N. Dewi, and H. Asari, “Penentuan Prediksi Awal Penyakit Jantung Menggunakan Algoritma Back Propagation Neural Network dengan Metode Adaboost,” vol. 2013, no. November, pp. 298–304, 2013.
- [7] D. Normawati and S. Winarti, “Feature selection with combination classifier use rules-based data mining for diagnosis of coronary heart disease,” *Proceeding 2018 12th Int. Conf. Telecommun. Syst. Serv. Appl. TSSA 2018*, pp. 2–7, 2019.
- [8] I. Septiani and N. Rijiati, “Abstrak,” 2016.
- [9] D. I. Annisa, R. Ariyanto, and A. T. R. H. Ririd, “KLASIFIKASI KEHAMILAN BERESIKO DENGAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR,” vol. 3, no. November, pp. 34–39, 2016.
- [10] A. Muzakir and R. Anisa Wulandari, “Model Data Mining sebagai Prediksi Penyakit Hipertensi Kehamilan dengan Teknik Decision Tree Model Data Mining sebagai Prediksi Penyakit Hipertensi Kehamilan dengan Teknik Decision Tree,” no. May, 2016.
- [11] D. T. Purwanti, “ANALISIS POLA ASOSIASI DATA PENYAKIT HIPERTENSI PADA IBU HAMIL MENGGUNAKAN ALGORITMA APRIORI,” vol. 01, no. 04, 2017.
- [12] Y. Kustiyahningsih, Mula’ab, and N. Hasanah, “Metode Fuzzy ID3 Untuk Klasifikasi Status Preeklamsi Ibu Hamil,” *Teknika*, vol. 9, no. 1, pp. 74–80, 2020.
- [13] Indriyanti, D. Sugianti, and M. A. Al Karomi, “Peningkatan Akurasi Algoritma KNN dengan Seleksi Fitur Gain Ratio untuk Klasifikasi Penyakit Diabetes Mellitus,” *IC-Tech*, vol. 7, no. 2, pp. 1–6, 2017.
- [14] “Tampilan Komparasi Performansi Algoritma Pengklasifikasi KNN, Bagging Dan Random Forest Untuk Prediksi Kanker Payudara.pdf.” .
- [15] R. Hidayat and T. Astuti, “Diagnosis Preeklamsia pada Ibu Hamil Berdasarkan Algoritme K- Nearest Neighbour,” vol. 14, no. 2, 2020.
- [16] R. K. Dinata, H. Akbar, and N. Hasdyna, “Algoritma K-Nearest Neighbor dengan Euclidean Distance dan

- Manhattan Distance untuk Klasifikasi Transportasi Bus,” *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 104–111, 2020.
- [17] M. R. A. Nasution and M. Hayaty, “Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter,” *J. Inform.*, vol. 6, no. 2, pp. 226–235, 2019.
- [18] R. B. Trengginaz, A. Yusup, D. S. Sunyoto, M. R. Jihad, and Y. Yulianti, “Pengujian Aplikasi Pemesanan Tiket Kereta berbasis Website Menggunakan Metode Black Box dengan Teknik Equivalence Partitioning,” *J. Teknol. Sist. Inf. dan Apl.*, vol. 3, no. 3, p. 144, 2020.
- [19] A. W. Soejono, A. Setyanto, and A. F. Sofyan, “Evaluasi Usability Website UNRIYO Menggunakan System Usability Scale (Studi Kasus: Website UNRIYO),” *J. Teknol. Inf.*, vol. XIII, no. 1, pp. 29–37, 2018.