LINIER
LITERATUR INFORMATIKA & KOMPUTER
e-ISSN: 3063-2218
DOI : 10.33096/linier.v1i4.2532
Literatur Informatika & Komputer
Vol 1, No 4 (2024), pp 328-335

Risiko Kesehatan Bagi Pasien Hamil Menggunakan Metode KNN

Miftahul Haq[a], Harlinda [b], A.Ulfah Tenripada Sahar[c]
Universitas Muslim Indonesia, Makassar, Indonesia
[a]13020200288@umi.ac.id ; [b]harlinda@umi.ac.id ; [c]a.ulfah@umi.ac.id

Received: 21-07-2024 | Revised: 20-08-2024 | Accepted: 10-11-2024 | Published: 29-12-2024

Abstrak

Kematian ibu hamil dalam melahirkan merupakan permasalahan kesehatan yang sudah lama terjadi. Ibu hamil memerlukan pemantauan terus menerus selama kehamilan hingga proses melahirkan agar dapat diketahui faktor risiko yang dialami. Ibu hamil perlu mengetahui penyakit yang mungkin bisa diderita, dan untuk mendapatkan pengetahuan serta melakukan diagnosa dini penyakit kehamilan, diperlukan sistem yang dapat membantu. Untuk mengatasi hal ini, pemantauan selama kehamilan penting untuk perkembangan janin yang sehat dan memastikan persalinan yang aman. Oleh karena itu, klasifikasi risiko kesehatan ibu hamil menjadi sangat penting untuk mencegah dan mengurangi risiko kesehatan yang serius bagi ibu dan bayi yang dikandungnya. Klasifikasi merupakan proses menemukan model untuk menggambarkan data dengan tujuan model tersebut dapat digunakan prediksi yang belum diketahui dari suatu objek. Salah satu metode dalam klasifikasi data adalah metode *K-Nearest Neighbor* (*KNN*), teknik ini menggunakan algoritma terawasi untuk mengklasifikasikan hasil *query instance* yang berdasar pada mayoritas dalam kategori KNN. Tujuan dari karya ilmiah ini yaitu mengklasifikasikan resiko kesehatan ibu hamil ke dalam kategori resiko tinggi seperti resiko menengah dan resiko rendah seperti resiko kesehatan yang dapat menganggu kesehatan ibu hamil yang dapat mengakibatkan adanya keguguran. Pada karya ilmiah ini mengklasikasikan data berdasarkan dari usia, tekanan darah sistolik, tekanan darah diastolik, gula darah, dan denyut jantung berdasarkan nilai K-Nearest yang di inputkan. Berdasarkan hasil klasifikasi KNN dengan menggunakan nilai K yang berbeda yaitu K=3, K=5, K=9 dan K=11 didapatkan hasil dengan nilai akurasi, presisi, recall, f1-score paling tinggi yaitu nilai K=11 dengan nilai 80%.

Kata kunci: Hamil, Klasifikasi, Kesehatan, K-Nearest Neighbor, Resiko

Pendahuluan

Kesehatan ibu hamil merupakan aspek penting yang perlu diperhatikan. Menurut *World Health Organization* (WHO), setiap tahunnya terdapat sekitar 800 wanita meninggal dunia akibat komplikasi kehamilan dan persalinan [1]. Masih tingginya angka kematian ibu disebabkan oleh banyak faktor. Salah satunya adalah faktor keterlambatan penanganan kasus [2]. Kematian ibu hamil dan melahirkan merupakan permasalahan di bidang kesehatan yang sudah lama terjadi, khususnya di negara-negara berkembang [3]. Ibu hamil memerlukan pemantauan terus menerus selama periode kehamilan sampai proses melahirkan supaya dapat diketahui faktor risiko yang dialami selama kehamilan [4]. Kurangnya pengetahuan dan informasi mengenai penyakit dengan berbagai gejala dapat menimbulkan keterlambatan ibu hamil untuk mengetahui terjadinya penyakit pada kehamilannya yang dapat mengakibatkan keguguran. [5].

Ibu hamil perlu menambah pengetahuan untuk mengetahui penyakit apa saja yang mungkin bisa diderita olehnya untuk mendapatkan pengetahuan tersebut dan agar bisa melakukan diagnosa dini penyakit kehamilan, diperlukan suatu sistem yang dapat membantu ibu hamil [6]. Untuk mengatasi hal tersebut harus dipantau selama kehamilan untuk perkembangan janin yang sehat dan untuk memastikan persalinan yang tidak berbahaya. Oleh karena itu, klasifikasi risiko kesehatan ibu hamil menjadi sangat penting untuk mencegah dan mengurangi risiko kesehatan yang serius bagi ibu dan bayi yang dikandungnya. Klasifikasi merupakan proses menemukan model untuk menggambarkan data dengan tujuan model tersebut dapat digunakan prediksi yang belum diketahui dari suatu objek [7] [8]. Salah satu metode dalam klasifikasi data adalah metode *K-Nearest Neighbor* (KNN). Teknik *K-Nearest Neighbor* (KNN) menggunakan algoritma terawasi untuk mengklasifikasikan hasil query instance yang berdasar pada mayoritas dalam kategori KNN [9]. Metode ini merupakan contoh pembelajaran berbasis instance dimana data latih disimpan, sehingga proses pengklasifikasian data yang tidak terklasifikasi dapat dengan mudah ditemukan dengan cara membandingkan data tersebut dengan data yang paling mirip pada data uji yang ada [10]. Pada tahap ini mengklasifikasikan resiko kesehatan ibu hamil ke dalam kategori resiko tinggi, menengah dan resiko rendah yang dilihat dari usia, tekanan darah sistolik, tekanan darah diastolik, gula darah, dan denyut jantung berdasarkan nilai K-Nearest yang di inputkan.

Penelitian sebelumnya telah melakukan penelitian mengenai klasifikasi kesehatan ibu hamil dengan judul Sistem Prototype Klasifikasi Risiko Kehamilan Dengan Algoritma *K-Nearest Neighbor* (K-NN). Pada penelitian tersebut menunjukkan bahwa sistem K-NN yang dikembangkan mampu membantu proses komputasi lebih cepat dengan dilengkapi hasil klasifikasi dengan tingkat akurasi mencapai 80% [11]. Penelitian selanjutnya yaitu Klasifikasi Kehamilan Beresiko Dengan Menggunakan Metode *K-Nearest Neighbor* (Studi Kasus Dinas Kesehatan Kabupaten Malang). Hasil uji dalam mengukur akurasi metode ini menggunakan metode validasi dengan membandingkan data yang diperoleh dari dinas kesehatan (puskesmas) dengan sistem dan menghasilkan tingkat akurasi nilai sebesar 93 % dengan menggunakan nilai K = 5, maka metode ini dapat dikategorikan baik dalam mengklasifikasi kehamilan beresiko ini [12]. Berdasarkan permasalahan diatas maka penulis membuat karya ilmiah dengan judul Klasifikasi Risiko Kesehatan Bagi Pasien Hamil Menggunakan Metode KNN, sehingga setelah proses klasifikasi tersebut diharapkan kehamilan tersebut dapat di awasi, jika di butuhkan akan diberikan rujukan untuk perawatan lebih lanjut sehingga dapat meminimalisir kematian ibu dan anak.

Metode

A. *K-Nearest Neighbors* (KNN)

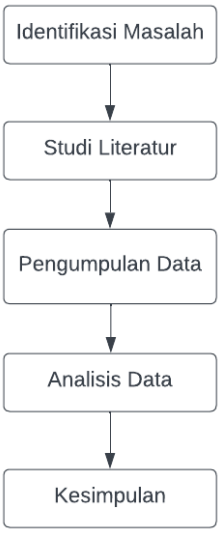
Algoritma KNN merupakan metode yang digunakan untuk melakukan klasifikasi data berdasarkan jarak terpendek terhadap objek data [13]. Penentuan nilai K yang terbaik untuk algoritma ini berdasarkan pada data yang ada. Nilai K yang tinggi dapat mengurangi efek noise pada klasifikasi, bisa juga membuat batasan antara setiap klasifikasi [14]. Tujuan dari algoritma KNN adalah mengklasifikasikan objek baru berdasarkan atribut dan training data. Algoritma KNN berkerja berdasarkan jarak terpendek dari *query instance* ke training data untuk menentukan k-NN-nya [15]

$$D_{xy} = \sqrt{\sum_i^n = 1 \, (x_i - \, y_i \,)^2} \quad (1)$$

Keterangan :

D : jarak kedekatan
x : data training
y : data testing
n : jumlah atribut individu antara 1 s.d. n
f : fungsi similitary atribut i antara kasus X dan kasus Y
i : Atribut individu antara 1 sampai dengan n

Pada tahap ini penulis memberikan gambaran terhadap alur penelitian yang dilakukan. Alur penelitian secara menyeluruh dapat dilihat pada visualisasi yang disajikan pada gambar 1 berikut ini:

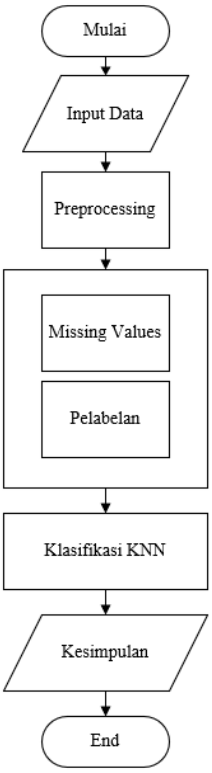


Gambar 1. Tahapan Penelitian

1. Identifikasi Masalah
   Tahap pertama dalam penelitian ini adalah identifikasi masalah. Penulis mengidentifikasi dalam mengklasifikasi Risiko Kesehatan Bagi Pasien Hamil.
2. Studi Literatur
   Tahap selanjutnya adalah studi literatur. Tahap ini dilaksanakan dengan memahami penelitian serupa yang pernah dilakukan sebelumnya untuk mendapatkan teori dan informasi yang mendukung penelitian literatur yang dipelajari yaitu mengenai klasifikasi resiko rendah, resiko sedang dan resiko tinggi.
3. Pengumpulan Data
   Pengumpulan data dilakukan menggunakan *kaagle* untuk memperoleh data yang diperlukan seperti data kesehatan ibu hamil dari usia, tekanan darah sistolik, tekanan darah diastolik, gula darah, dan denyut jantung.
4. Analisis Data
   Selanjutnya dilakukan analisis data yang terdiri dari klasifikasi KNN untuk mengklasifikasikan data. Library yang dimasukkan pada *google colab* yaitu stopwords bahasa Indonesia dan *stemming list* bahasa Indonesia, dan nilai akurasi dari metode tersebut.

Perancangan

A. Desain Perancangan
   Desain Perancangan ini berisikan tahapan-tahapan yang akan dilakukan. Gambar 2 berikut ini adalah desain perancangan yang akan dilakukan**.**



Gambar 2. Desain Perancangan

1. Input Data
   Pada penelitian ini mengambil data di *kaagle* mengenai data resiko kesehatan ibu hamil yang kemudian di klasifikasikan ke dalam resiko tinggi, resiko sedang dan resiko rendah
2. Preprocessing
   Pada tahap preprocessing ini, data yang telah diambil di *kaagle* akan diolah menjadi data yang dapat digunakan pada tahapan selanjutnya. Dengan melakukan pendeskripsian data, dapat diketahui dataset ini terdapat *Missing Value* pada variabel *RiskLevel*. Sehingga dilakukan pengisian data pada nilai yang

kosong dengan menggunakan operator *replace missing values* untuk mengisi data yang kosong pada RiskLavel dengan nilai *Average* atau rata-rata pada setiap variabel yang ada pada dataset.

3. Klasifikasi,
   Dataset " *Maternal Health Risk Data* " terdiri dari atribut kategorikal yaitu *RiskLavel* dan juga numerikal seperti usia, Tekanan darah Sistolik, tekanan darah Diastolik, gula darah, denyut jantung. KNN cenderung lebih tahan terhadap data noise karena keputusan klasifikasi didasarkan pada mayoritas kelas dari K tetangga terdekat. Titik-titik data yang berada jauh dari pusat klaster tidak akan secara signifikan mempengaruhi keputusan klasifikasi. Pada penelitian ini jumlah splitdata yang optimalkan algoritma k-NN yaitu (80:20) dimana data training 80% untuk pengaplikasian pada algoritma k-NN dan data testing 20% untuk mengetes model algoritma k-NN. Pada penelitian ini membandingkan nilai k terbaik antara nilai k=3, k=5, k=9 dan k=11
4. Evaluasi
   Setelah melakukan klasifikasi tahapan selanjutnya yaitu evaluasi, Ada beberapa hal yang bisa digunakan untuk mengevaluasi kinerja klasifikasi. Penelitian ini menggunakan *Accuracy* sebagai parameter evaluasi untuk mengetahui seberapa baik hasil yang akan diberikan.

Tabel 1. Definisi Operasional Variabel

Variabel	Definisi	Ket.
Umur, Sistol, Diastolic,BodyTemp,HeartRate	Parameter dalam klasifikasi resiko kesehatan pasien hamil	*Input*
Klasifikasi	Mengklasifikasikan resiko kesehatan pasien hamil	Proses
RiskLevel	Penentuan klasifikasi resiko kesehatan pasien hamil	*Output*

Pada tabel 1 ini menunjukkan bahwa penelitian atau sistem ini bertujuan untuk mengklasifikasikan risiko kesehatan pada ibu hamil. Data-data yang ada yaitu usia, tekanan darah, suhu tubuh, dan detak jantung digunakan sebagai input untuk model klasifikasi. Model ini kemudian akan menghasilkan output berupa tingkat risiko kesehatan ibu hamil.

Pemodelan

A. Input Data
   Memasukkan data yang diambil di *kaagle* mengenai data **resiko kesehatan pasien hamil** ke dalam *google colab*

	Age	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate	RiskLevel
0	25	130	80	15.0	98.0	86	high risk
1	35	140	90	13.0	98.0	70	high risk
2	29	90	70	8.0	100.0	80	high risk
3	30	140	85	7.0	98.0	70	high risk
4	35	120	60	6.1	98.0	76	low risk
...	...	...	...	...	...	...	...
1009	22	120	60	15.0	98.0	80	high risk
1010	55	120	90	18.0	98.0	60	high risk
1011	35	85	60	19.0	98.0	86	high risk
1012	43	120	90	18.0	98.0	70	high risk
1013	32	120	65	6.0	101.0	76	mid risk

1014 rows × 7 columns

Gambar 3. Data Pasien

Pada gambar 3 merupakan data pasien yang telah berhasil dimasukkan yang diambil dari *kaagle* yang terdiri dari parameter usia, tekanan darah sistolik, tekanan darah diastolik, gula darah, dan denyut jantung

B. *Pre-processing*
   1. *Missing Values*

Melakukan *missing values* dengan menghilangkan data yang kosong atau duplikat dalam memudahkan dalam mengolah data

	Age	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate	RiskLevel
0	25	130	80	15.0	98.0	86	high risk
1	35	140	90	13.0	98.0	70	high risk
2	29	90	70	8.0	100.0	80	high risk
3	30	140	85	7.0	98.0	70	high risk
4	35	120	60	6.1	98.0	76	low risk
...	...	...	...	...	...	...	...
1009	22	120	60	15.0	98.0	80	high risk
1010	55	120	90	18.0	98.0	60	high risk
1011	35	85	60	19.0	98.0	86	high risk
1012	43	120	90	18.0	98.0	70	high risk
1013	32	120	65	6.0	101.0	76	mid risk

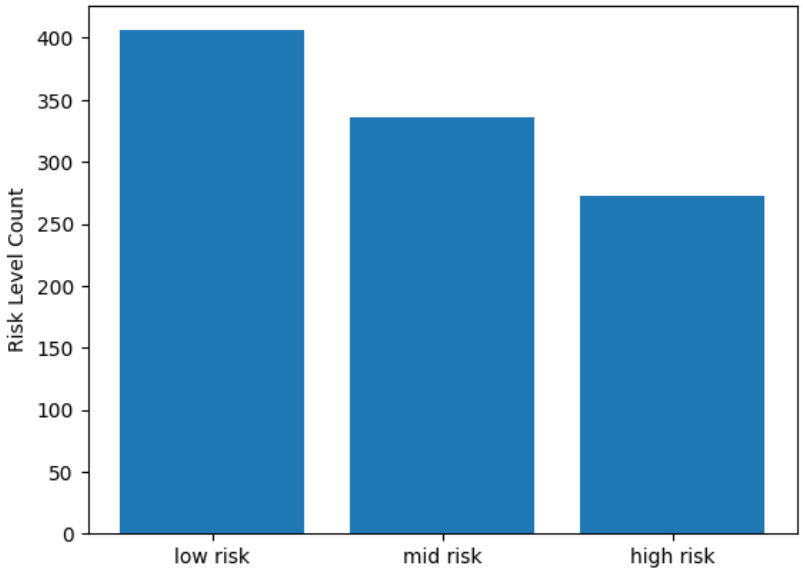
1014 rows × 7 columns

Gambar 4. Hasil *Missing Values*

Pada gambar 4 merupakan proses *missing values* yaitu menghilangkan data yang duplikat atau kosong yang ada pada data yang telah dimasukkan

2. Pelabelan Data

Proses pelabelan data sentimen analisis hasil klasifikasi risiko kesehatan bagi pasien hamil diberikan label berdasarkan data hasil klasifikasi pada ibu hamil



Gambar 5. Visualisasi Jumlah Data Klasifikasi

Pada gambar 5 merupakan hasil jumlah data klasifikasi yang digunakan yang terdiri dari kategori risiko rendah dengan 406 data, risiko menengah dengan 336 data dan risiko tinggi 272 data.

C. Klasifikasi KNN

Berikut adalah hasil dari klasifikasi dari metode KNN dengan menggunakan data testing sebanyak 20 data nilai value K=3

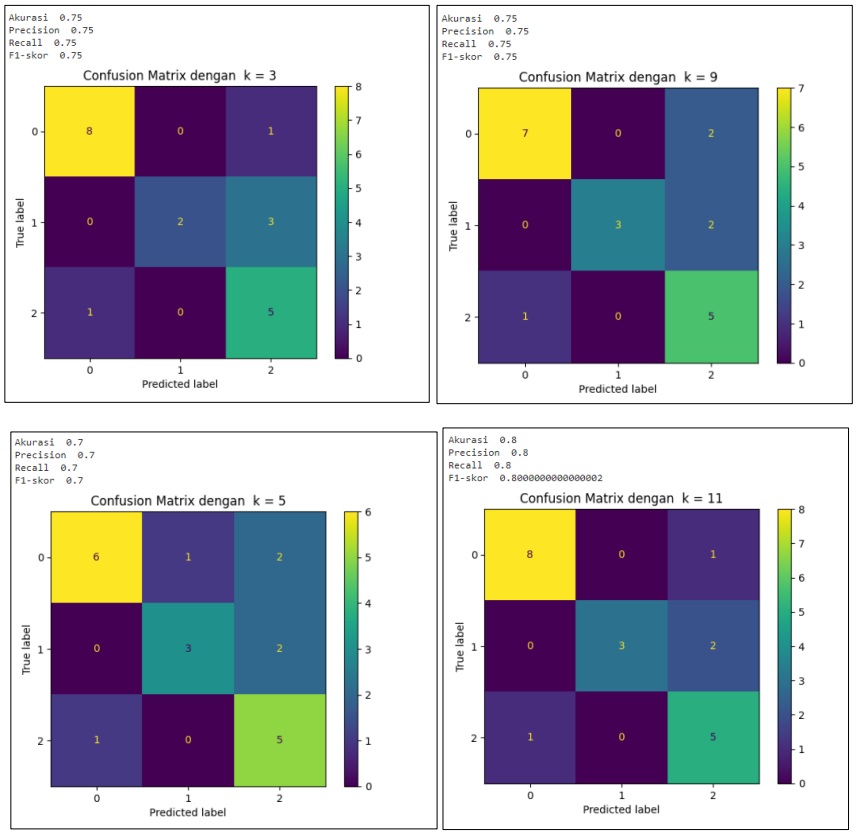
Enter Value for K:3

	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate	RiskLevel	Predicted
0	130	80	15.00	98	86	high risk	high risk
1	140	90	13.00	98	70	high risk	high risk
2	90	70	8.00	100	80	high risk	high risk
3	140	85	7.00	98	70	high risk	high risk
4	120	60	6.10	98	76	low risk	mid risk
5	140	80	7.01	98	70	high risk	high risk
6	130	70	7.01	98	78	mid risk	mid risk
7	85	60	11.00	102	86	high risk	high risk
8	120	90	6.90	98	70	mid risk	mid risk
9	130	80	18.00	98	70	high risk	mid risk
10	90	60	7.01	98	76	low risk	low risk
11	120	80	7.00	98	70	mid risk	mid risk
12	110	89	7.01	98	77	low risk	mid risk
13	120	75	7.01	100	70	mid risk	mid risk
14	120	80	11.00	98	88	mid risk	high risk
15	120	80	7.01	98	70	low risk	low risk
16	140	90	15.00	98	90	high risk	high risk
17	140	100	7.01	98	80	high risk	high risk
18	120	80	6.90	101	76	mid risk	mid risk
19	70	50	6.90	98	70	low risk	mid risk

Gambar 6. Hasil Klasifikasi KNN

Pada gambar 6 merupakan hasil klasifikasi KNN dengan nilai K adalah 3 yang memprediksikan data keseluruhan pasien dengan data hasil prediksi

D. Evaluasi



Gambar 7. Hasil Evaluasi Model KNN

Pada gambar 7 merupakan hasil klasifikasi KNN dengan menggunakan nilai K yang berbeda didapatkan hasil dengan nilai akurasi, presisi, recall, f1-score paling tinggi yaitu nilai K=11 dengan nilai 80%. Pada nilai K=3 didapatkan akurasi, presisi, recall, f1-score 75%, sedangkan pada nilai K=5 didapatkan didapatkan akurasi, presisi, recall, f1-score 70%, serta nilai K=9 didapatkan akurasi, presisi, recall, f1-score 75%.

Kesimpulan

Dengan ini kita dapat mengklasifikasikan prediksi risiko kesehatan bagi pasien hamil ke dalam kelompok beresiko tinggi, beresiko menengah, dan beresiko rendah berdasarkan parameter usia, tekanan darah, kadar glukosa, subu tubuh dan denyut jantung sehingga diberikan rujukan untuk perawatan lebih lanjut dan dapat meminimalisir kematian ibu dan anak. Berdasarkan hasil klasifikasi metode KNN didapatkan hasil dengan nilai akurasi, presisi, recall, f1-score paling tinggi yaitu nilai K=11 dengan nilai 80%.

Daftar Pustaka

[1] R. Aryanti, T. Misriati, and R. Hidayat, "Klasifikasi Risiko Kesehatan Ibu Hamil Menggunakan Random Oversampling Untuk Mengatasi Ketidakseimbangan Data," *KLIK: Kajian Ilmiah Informatika dan Komputer* , vol. 3, no. 5, pp. 409–416, 2023, [Online]. Available: https://djournals.com/klik

[2] Yunita Permatasari, Umi Salamah, and Ristu Saptono, "Klasifikasi Risiko Bahaya Kehamilan dengan Metode Fuzzy C-Means," *JURNAL ITSMART* , vol. 2, no. 1, pp. 8–15, 2023.

[3] R. Aryanti, T. Misriati, and R. Hidayat, "Klasifikasi Risiko Kesehatan Ibu Hamil Menggunakan Random Oversampling Untuk Mengatasi Ketidakseimbangan Data," *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 3, no. 5, pp. 409–416, Apr. 2023, [Online]. Available: https://djournals.com/klik

[4] M. Ahmed, M. A. Kashem, M. Rahman, and S. Khatun, "Review and Analysis of Risk Factor of Maternal Health in Remote Area Using the Internet of Things (IoT)," in *Lecture Notes in Electrical Engineering*, Springer, 2020, pp. 357–365. doi: 10.1007/978-981-15-2317-5_30.

[5] Aprilia Manganti, Saifulloh, and Andria, "SISTEM PAKAR DIAGNOSA PENYEBAB KEGUGURAN PADA IBU HAMIL MENGGUNAKAN METODE FORWARD CHAINING," *Jurnal Sistem Informasi dan Sains Teknologi*, vol. 3, no. 2, pp. 1–13, 2021.

[6] Y. Pinanda, W. Firdaus Mahmudy, and E. Santoso, "Klasifikasi Risiko Penyakit pada Ibu Hamil menggunakan Metode Modified K-Nearest Neighbor (MKNN)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 5, pp. 2116–2121, 2022, [Online]. Available: http://j-ptiik.ub.ac.id

[7] P. R. Sihombing and I. F. Yuliati, "Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 2, pp. 417–426, May 2021, doi: 10.30812/matrik.v20i2.1174.

[8] Amaliah Faradibah, Dewi Widyawati, A Ulfah Tenripada Syahar, and Sitti Rahmah Jabir, "Comparison Analysis of Random Forest Classifier, Support Vector Machine, and Artificial Neural Network Performance in Multiclass Brain Tumor Classification," *Indonesian Journal of Data and Science*, vol. 4, no. 2, pp. 54–63, Jul. 2023, doi: 10.56705/ijodas.v4i2.73.

[9] A. W. Widodo, D. Hernando, and W. F. Mahmudy, "Mangrove Forest Classification in Drone Images Using HSV Color Moment and Haralick Features Extraction with K-Nearest Neighbor," *Signal and Image Processing Letters*, vol. 1, no. 3, pp. 1–12, Nov. 2019, doi: 10.31763/simple.v1i3.6.

[10] Y. Pinanda, W. Firdaus Mahmudy, and E. Santoso, "Klasifikasi Risiko Penyakit pada Ibu Hamil menggunakan Metode Modified K-Nearest Neighbor (MKNN)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 6, no. 5, pp. 2116–2121, May 2022, [Online]. Available: http://j-ptiik.ub.ac.id

[11] A. Deharja, M. W. Santi, M. Yunus, and E. Rachmawati, "Sistem Prototype Klasifikasi Risiko Kehamilan Dengan Algoritma k-Nearest Neighbor (k-NN)," *JTIM : Jurnal Teknologi Informasi dan Multimedia*, vol. 4, no. 1, pp. 66–72, May 2022, doi: 10.35746/jtim.v4i1.229.

[12] D. Itsna Annisa, R. Ariyanto, and A. Tri Retno Hayati Ririd, "Klasifikasi Kehamilan Beresiko Dengan Menggunakan Metode K-Nearest Neighbor (STUDI KASUS DINAS Kesehatan Kabupaten Malang)," *Jurnal Informatika Polinema* , vol. 3, no. 1, pp. 34–39, 2019.

[13] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, "Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa," *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 6, no. 2, pp. 118–127, 2021.

[14] Nefa Mutiara Shandhini Maylita, Hani Zulfia Zahro, and Nurlaily Vendyansyah, "Penerapan Metode K-Nearest Neighbor (KNN) Umtuk Menentukan Status Gizi Balita," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 6, no. 2, pp. 953–956, 2022.
[15] U. Muldayawati Asmudi, Harlinda, and N. Kurniati, "Penentuan Tingkat Kelulusan Mahasiswa di Fakultas Ekonomi dan Bisnis Universitas Muslim Indonesia Berbasis Web," *Buletin Sistem Informasi dan Teknologi Islam*, vol. 2, no. 3, pp. 208–212, 2021.