



OPTIMASI METODE RANDOM FOREST UNTUK KLASIFIKASI RISIKO OBESITAS BERDASARKAN POLA MAKAN

Nafiati Rosidah¹, Rastri Prathivi^{2*}, Susanto³

^{1,2,3}Program Studi Teknik Informatika, Universitas Semarang

email: vivi@usm.ac.id^{2*}

Abstrak: Pola makan tidak sehat, seperti mengonsumsi makanan cepat saji dan mengurangi asupan sayur serta buah, merupakan salah satu penyebab utama obesitas, yang menjadi masalah kesehatan global. Obesitas dapat meningkatkan risiko penyakit tidak menular seperti diabetes, hipertensi, dan penyakit jantung. Dengan menggunakan algoritma Random Forest, penelitian ini bertujuan untuk mengembangkan model prediksi risiko obesitas berbasis pola makan. Data yang digunakan terdiri dari 1.610 record dengan 15 atribut, yang diambil dari dataset publik. Tahapan penelitian meliputi pengumpulan data, preprocessing, pembagian data menjadi data latih dan uji, implementasi model, optimasi hyperparameter dengan Grid Search, dan evaluasi model. Hasil penelitian menunjukkan bahwa model Random Forest yang dioptimasi mampu mencapai akurasi sebesar 85,4%. Frekuensi mengonsumsi makanan cepat saji, jumlah makanan utama setiap hari, kebiasaan ngemil, dan konsumsi sayur adalah beberapa variabel penting yang memengaruhi prediksi. Model ini diharapkan dapat membantu pencegahan dan penanganan obesitas secara lebih efektif sekaligus memberikan wawasan tambahan untuk pengembangan sistem prediksi kesehatan berbasis data.

Kata Kunci : Obesitas, Pola Makan, *Random Forest*, *Grid Search*, Klasifikasi

PENDAHULUAN

Perubahan gaya hidup yang terjadi di negara maju dan berkembang telah meningkatkan prevalensi obesitas, sebuah kondisi kelebihan berat badan yang berdampak negatif pada kesehatan, hal ini terjadi karena adanya ketidakseimbangan antara masukan energi dan keluaran energi. [1]. Kondisi terjadinya obesitas ditentukan berdasarkan parameter Indeks Massa Tubuh (IMT) [2]. Salah satu jenisnya, yaitu obesitas berat normal (*Normal Weight Obesity/NWO*) memiliki definisi yang berbeda berdasarkan studi, populasi, dan jenis kelamin; namun, biasanya didefinisikan sebagai persentase lemak tubuh di atas 30% [3]. Asupan kalori yang berlebihan dibandingkan dengan aktivitas pembakaran kalori akan mengakibatkan kelebihan itu menumpuk dalam bentuk lemak [4]. *World Health Organization* (WHO), menyatakan bahwa lebih dari 1,9 miliar orang dewasa berusia 18 tahun ke atas mengalami kelebihan berat badan, 600 juta di antaranya obesitas. [5].

Obesitas dapat dipengaruhi oleh berbagai faktor, salah satunya adalah pola makan yang tidak sehat. Perubahan pola makan ini sering kali mengarah pada kebiasaan makan berlebihan, terutama dengan konsumsi makanan berkalori tinggi [6]. Misalnya, konsumsi makanan cepat saji yang berlebihan menyebabkan peningkatan asupan energi, lemak, dan gula dalam tubuh, di mana makanan cepat saji biasanya mengandung banyak sodium tetapi sedikit serat. Semakin sering mengonsumsinya, maka semakin tinggi risiko mengalami obesitas [7]. Selain itu, konsumsi rendah sayur dan buah yang merupakan sumber serat, vitamin, dan mineral, dapat menyebabkan obesitas dan penyakit tidak menular seperti diabetes, kanker kolon, hipertensi, stroke, dan jantung [8]. Tidak hanya itu, camilan yang biasanya dimakan di luar waktu makan utama sering kali dikonsumsi beberapa kali sehari, seperti di pagi, siang, sore, atau malam hari [9]. Jika konsumsi camilan ini tidak terkendali, kalori harian bisa meningkat dan berpotensi menyebabkan kenaikan berat badan.

Perkembangan ilmu pengetahuan dan teknologi ada peluang besar untuk membangun model data mining yang dapat membantu mengidentifikasi tingkat obesitas dengan lebih efisien. Data mining merupakan sebuah teknik yang dapat dipergunakan untuk membantu penyelesaian permasalahan tersebut. Pada data mining banyak digunakan untuk melakukan proses pengolahan data [10]. Klasifikasi adalah proses pembelajaran fungsi atau model pada sekumpulan data latih sehingga model tersebut dapat digunakan untuk memprediksi klasifikasi data uji [11]. Penelitian ini menggunakan algoritma Random memiliki proses pemilihan kumpulan data, penelitian ini menggunakannya untuk meningkatkan kinerja model klasifikasi. Dengan fitur ini, Random Forest pasti dapat bekerja secara efisien pada kumpulan data yang besar dengan parameter yang sangat kompleks, yang akan sangat memudahkan pengklasifikasiannya.

Penelitian ini bertujuan untuk mengidentifikasi faktor pola makan yang berkontribusi pada obesitas, seperti konsumsi makanan cepat saji, frekuensi konsumsi sayuran, jumlah makanan utama per hari, serta asupan makanan di antara waktu makan (camilan) dan juga mengembangkan model prediksi obesitas berbasis algoritma Random Forest. Hasil penelitian diharapkan dapat mendukung upaya pencegahan dan penanganan obesitas secara efektif.

TINJAUAN PUSTAKA

Penelitian yang dilakukan oleh Apriliah dkk., 2021, berusaha untuk mengidentifikasi risiko diabetes pada tahap awal dengan menggunakan algoritma klasifikasi Random Forest. Data ini dikumpulkan dari dataset 520 kasus dari rumah sakit di Sylhet, Bangladesh. Hasil evaluasi menunjukkan bahwa algoritma Random Forest memiliki akurasi tertinggi sebesar 97,88% dibandingkan dengan algoritma Support Vector Machine (SVM) dan algoritma Naive Bayes yang juga diuji dalam penelitian ini [12]. Demikian dengan Mualfah dkk.,2022, pada penelitiannya yang membahas penggunaan



SMOTE untuk mengatasi ketidakseimbangan data dalam deteksi stroke dengan Random Forest. Data yang digunakan dari 43.400 pasien, dengan 783 kasus stroke. Random Forest tanpa SMOTE menghasilkan akurasi 0.98, presisi 0.69, recall 0.51, dan f1-score 0.51. Dengan SMOTE, akurasi menjadi 0.91, presisi 0.92, recall 0.91, dan f1-score 0.91. Dengan demikian Random Forest efektif untuk deteksi stroke pada data besar, dan SMOTE efektif mengatasi ketidakseimbangan data [13]. Penelitian Arifin Yusuf Permana dkk.,2023, menyatakan bahwa pada penelitian tentang penggunaan algoritma Random Forest untuk prediksi kanker paru. Random Forest menghasilkan akurasi 90.61% dan AUC 0.941. Setelah optimasi fitur evolusioner, akurasi meningkat menjadi 94.17% dengan AUC 0.942. Dengan optimasi bobot evolusioner, akurasi mencapai 93.20% dengan AUC 0.946 [14]. Selanjutnya, seperti yang dinyatakan oleh Dwi dkk.,2024, dalam penelitian mereka yang membahas bagaimana algoritma Random Forest dapat digunakan untuk memprediksi tingkat obesitas di Meksiko. Dengan akurasi 93.38%, Model Hutan Random memiliki hasil terbaik pada kelas Obesity_Type_I dan Obesity_Type_III. Namun, memiliki sedikit penurunan pada kelas Overweight_Level_I, Model Random Forest masih dapat memprediksi tingkat obesitas dan menunjukkan kemampuan untuk mencegah dan menangani obesitas di Meksiko [15].

METODE

Tahapan Penelitian

Tahapan penelitian adalah kumpulan langkah-langkah yang harus dilakukan untuk mencapai tujuan penelitian. Diagram 1 menunjukkan langkah-langkah yang diambil dalam penelitian ini.

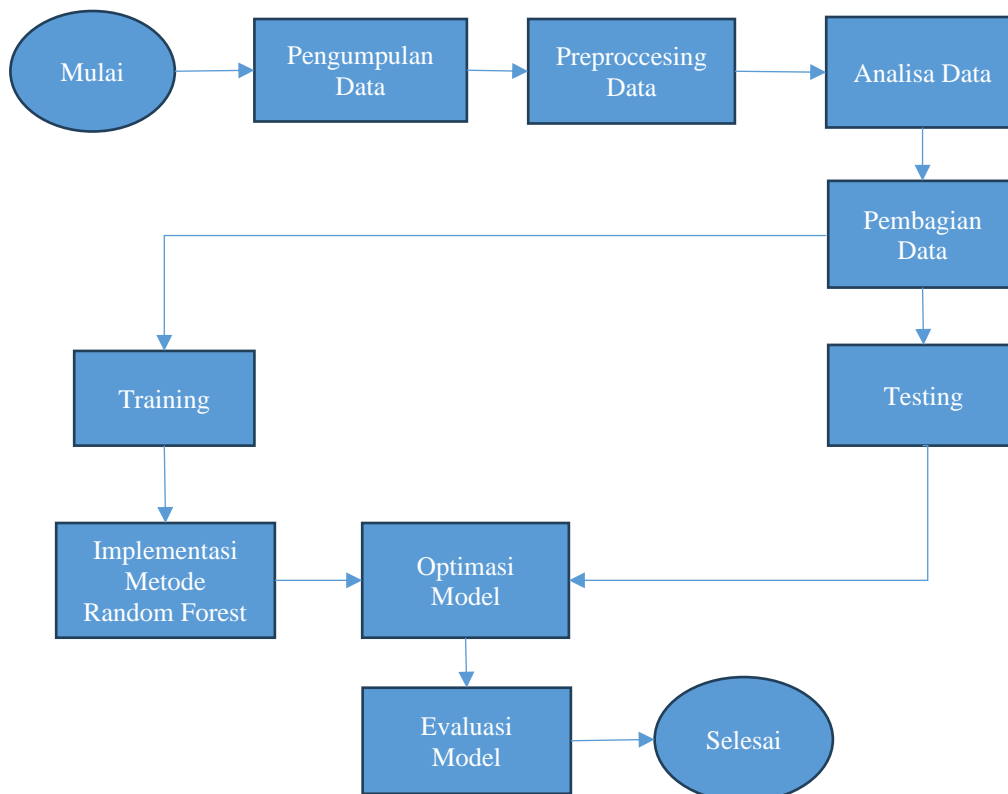


Diagram 1. Alur Penelitian

Pengumpulan Data

Metode pengumpulan data sekunder digunakan dalam penelitian ini untuk mengumpulkan sumber data. Data yang digunakan untuk penelitian ini berasal dari dataset yang diakses dari repositori data Kaggle.

Preprocessing Data

Preprocessing mempersiapkan data yang dikumpulkan dari tahap sebelumnya sebelum digunakan pada tahap berikutnya. Dalam tahap ini, nilai yang hilang diganti dengan nilai null atau kosong.

Analisis Data

Metode data mining digunakan untuk menemukan pola dan wawasan dalam data penelitian ini. Random Forest digunakan untuk menentukan tingkat obesitas. Untuk menilai, akurasi, ketepatan, recall, dan skor F1 digunakan.



Pembagian Data (Split Data)

Dataset dibagi menjadi dua bagian: data pelatihan dan data pengujian, dan rasio pemisahannya kira-kira 80% untuk data pelatihan sebanyak 1288 data dan 20% untuk data pengujian sebanyak 332 data.

Implementasi Model

Dalam penelitian ini, model Random Forest digunakan. Random Forest adalah salah satu algoritma pembelajaran mesin yang paling efektif, yang menggabungkan beberapa pohon keputusan untuk meningkatkan akurasi prediksi. Algoritma ini melatih setiap pohon dengan memilih sebagian atribut dan subset data secara acak. Untuk menghasilkan klasifikasi akhir, hasil prediksi dari setiap pohon digabungkan [16].

Optimasi Model

Salah satu cara untuk meningkatkan akurasi model klasifikasi adalah dengan mengatur hyperparameter [17]. Nilai parameter yang ideal untuk subproses operator dapat ditemukan dengan menggunakan Optimize Parameter Grid, yang melibatkan penilaian kinerja untuk setiap kombinasi nilai parameter yang dipilih [18].

Evaluasi Model

Evaluasi model yang akan dievaluasi menggunakan indikator evaluasi yaitu akurasi, presisi, recall, dan f1-score. Selain itu, confusion matrix digunakan untuk melihat kesalahan prediksi, dan cross-validation untuk menguji model pada data yang berbeda.

HASIL DAN PEMBAHASAN

Peneliti dalam penelitian ini memeriksa klasifikasi risiko obesitas berdasarkan pola makan. Dataset publik yang tersedia di Kaggle digunakan untuk mengumpulkan dataset pertama. Dataset ini terdiri dari 1610 record dan 15 atribut, yang dijelaskan dalam tabel 1.

Tabel 1. Atribut Kumpulan Data

Atribut	Fitur	Value
Jenis Kelamin	1. Laki-laki 2. Perempuan	712 898
Usia	Usia dalam tahun	
Tinggi badan	Tinggi badan dalam cm	
Riwayat Keluarga Kelebihan Berat Badan/Obesitas	1. Ya 2. Tidak	266 1344
Konsumsi Makanan Cepat Saji	1. Ya 2. Tidak	436 1174
Frekuensi Konsumsi Sayuran	1. Jarang 2. Terkadang 3. Selalu	400 708 502
Jumlah Makan Utama per Hari	1. 1-2 2. 3 3. 3+	444 928 238
Asupan Makanan di Antara Waktu Makan	1. Jarang 2. Terkadang 3. Biasanya 4. Selalu	346 564 417 283
Merokok	1. Ya 2. Tidak	492 118
Asupan Cairan Harian	1. Kurang dari 1 liter 2. 1 hingga 2 liter 3. Lebih dari 2 liter	456 523 631
Perhitungan Asupan Kalori	1. Ya 2. Tidak	286 1324

Aktivitas Fisik/Olahraga	1. Tidak ada	206
	2. 1-2 hari	290
	3. 3-4 hari	370
	4. 5-6 hari	358
	5. 5. 6+ hari	386
Waktu yang Didedikasikan untuk Teknologi	1. 0 dan 2 jam	382
	2. 3 dan 5 jam	826
	3. 3. 5+ jam	402
Jenis Transportasi yang Digunakan	1. Mobil	660
	2. Sepeda Motor	94
	3. Sepeda	116
	4. Angkutan umum	602
	5. Berjalan	138
Kategori	1. Kurang	73
	2. Biasa	658
	3. Kegemukan	592
	4. Obesitas	287

Data yang telah didapat terlebih dahulu akan langsung dilakukan preprocessing, dimana pada tahap ini data akan disiapkan sebelum masuk dalam tahapan selanjutnya.

Preprocessing Data

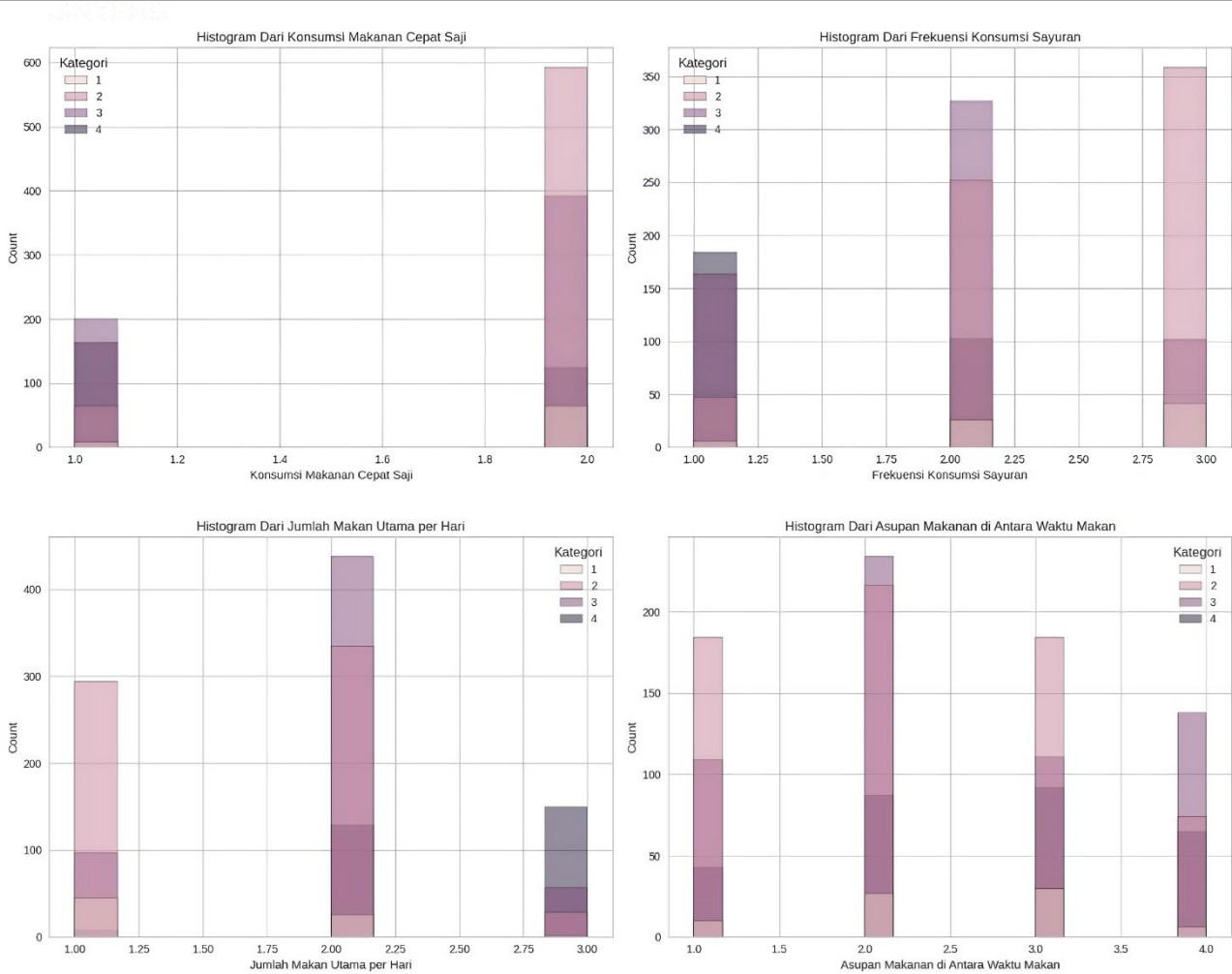
Data preprocessing dilakukan untuk memastikan tidak ada nilai kosong, seperti yang terlihat pada Gambar 2. Kemudian, dataset dibagi menjadi dua bagian yaitu data training dan data testing. Hal ini penting untuk menguji kemampuan model untuk menangani data baru dengan baik dan menghasilkan hasil yang dapat dipercaya [19]. Selanjutnya pada tahapan ini hanya menggunakan beberapa atribut untuk menentukan tingkat obesitas berdasarkan pola makan saja yang mencakup konsumsi makanan cepat saji, frekuensi konsumsi sayuran, jumlah makanan utama per hari, serta asupan makanan di antara waktu makan (camilan). Seperti yang terlihat pada Tabel 2 berikut:

Tabel 2. Atribut Yang digunakan

Atribut	Fitur	Value
Konsumsi Makanan Cepat Saji	1. Ya	436
	2. Tidak	1174
Frekuensi Konsumsi Sayuran	1. Jarang	400
	2. Terkadang	708
	3. Selalu	502
Jumlah Makan Utama per Hari	1. 1-2	444
	2. 3	928
	3. 3+	238
Asupan Makanan di Antara Waktu Makan (Ngemil)	1. Jarang	346
	2. Terkadang	564
	3. Biasanya	417
	4. Selalu	283

Analisis Data

Untuk mengetahui gambaran data pada analisis data dari setiap atribut di visualisasikan ke dalam bentuk diagram yang terlihat pada Gambar 2 berikut.



Gambar 2. Visualisasi Diagram

Pada Gambar visualisasi diatas menunjukkan pada diagram Konsumsi Makanan Cepat Saji sebagian besar responden (1.174 orang) mengonsumsi makanan cepat saji dengan frekuensi tinggi (Ya, nilai 2.0 pada sumbu X), Sebaliknya, hanya 436 responden yang tidak memiliki kebiasaan mengonsumsi makanan cepat saji (Tidak, nilai 1.0). Sementara itu pada diagram Frekuensi Konsumsi Sayuran, responden dengan konsumsi sayuran rendah (Jarang) berjumlah 400 orang. Sebagian besar responden memiliki konsumsi sayuran yang lebih baik, dengan 708 orang terkadang mengonsumsi sayuran dan 502 orang selalu mengonsumsinya. Untuk diagram jumlah makan utama per hari, sebagian besar responden (928 orang) makan utama sebanyak 2 kali sehari (puncak pada nilai 2.0 di sumbu X). Sebanyak 444 orang makan 1-2 kali sehari, sedangkan 238 orang memiliki pola makan utama lebih dari 3 kali sehari. Pada diagram asupan makanan di antara waktu makan, 346 responden jarang ngemil, sedangkan 564 responden terkadang ngemil, 417 responden biasanya ngemil, dan 283 responden selalu ngemil. Frekuensi ngemil tertinggi berada di nilai 2.0, tetapi ada juga responden dengan kebiasaan ngemil lebih tinggi, mencapai nilai 4.0.

Implementasi Model Random Forest

Hasil pengujian menunjukkan Random Forest merupakan model yang paling efektif diantara model yang diuji, dengan akurasi tertinggi sebesar 41,3%. Meskipun akurasi ini mungkin tidak cukup tinggi dalam beberapa konteks, hal ini menunjukkan bahwa Random Forests lebih mampu menangkap pola dalam data dibandingkan model lain yang diuji. Hasil akurasi dari Random Forest bisa dilihat pada Tabel 3.

Tabel 3. Hasil Akurasi Random Forest Berdasarkan Data Uji

No	Metric	Nilai (%)
1	Accuracy	0.41
2	Precision	0.19



3	Recall	0.41
4	F1-Score	0.25

Optimasi Model

Untuk membantu meningkatkan kinerja model Random Forest, tuning hyperparameter digunakan dengan GridSearchCV. Selama proses tuning hyperparameter pada tahap pengembangan model klasifikasi, kombinasi parameter yang diprediksi diperoleh, yang menghasilkan kinerja terbaik [20]. Hasil dari Grid Search menunjukkan bahwa kombinasi hyperparameter dengan max_depth 20 dan n_estimators 200 memberikan kinerja terbaik untuk model Random Forest dengan akurasi mencapai sekitar 84.24%.

Evaluasi Model

Setelah dilakukan optimasi dengan Grid Search, model Random Forest menunjukkan akurasi 85.40%, artinya prediksi benar. Precision untuk kelas 0 mencapai 1.00, yang berarti model sangat akurat dalam memprediksi kelas ini, sementara untuk kelas lainnya, precision masing-masing mencapai 88% (kelas 1), 85% (kelas 2), dan 77% (kelas 3). Recall menunjukkan bahwa model mendeteksi 92% dari data kelas 0, 90% untuk kelas 1, 80% untuk kelas 2, dan 85% untuk kelas 3. F1-score tertinggi pada kelas 0 (96%) dan terendah pada kelas 3 (80%). Total support adalah 322, yang merupakan jumlah seluruh data uji yang digunakan untuk mengevaluasi model. Rata-rata makro dan weighted menunjukkan kinerja yang seimbang dan dapat diandalkan untuk semua kelas, model ini memiliki kinerja yang baik dalam klasifikasi data.

Tabel 4. Hasil Optimasi Berdasarkan Data Uji

Class	Precision	Recall	F1-Score	Support
0	1.00	0.85	0.92	13
1	0.89	0.90	0.90	125
2	0.83	0.80	0.82	131
3	0.78	0.85	0.81	53
accuracy			0.85	322
macro avg	0.87	0.85	0.86	322
Weighted avg	0.85	0.85	0.85	322

KESIMPULAN DAN SARAN

Dengan menggunakan algoritma Random Forest, penelitian ini menghasilkan model prediksi risiko obesitas berdasarkan pola makan yang berhasil. Model ini, yang dioptimasi melalui pencarian grid, mencapai akurasi sebesar 85,4%, menunjukkan bahwa algoritma ini sangat efektif dalam menemukan elemen yang berkontribusi pada obesitas. Frekuensi konsumsi makanan cepat saji, jumlah makanan utama yang dikonsumsi setiap hari, kebiasaan ngemil, dan konsumsi sayur adalah faktor penting yang memengaruhi prediksi risiko obesitas. Diharapkan bahwa hasil penelitian ini akan memberikan informasi tambahan tentang bagaimana membuat sistem prediksi kesehatan berbasis data dan mendukung upaya pencegahan dan penanganan obesitas yang lebih efektif. Disarankan agar penelitian selanjutnya menggunakan dataset yang lebih besar dan lebih beragam untuk meningkatkan kemampuan generalisasi model. Selain itu, penelitian lebih lanjut dapat dilakukan dengan menguji algoritma tambahan untuk membandingkan kinerja dan menemukan metode terbaik untuk klasifikasi risiko obesitas. Selain itu, pengembangan aplikasi berbasis teknologi informasi (TI) juga dapat dilakukan untuk mengajarkan orang tentang pola makan sehat dan cara mencegah obesitas.

DAFTAR PUSTAKA

- [1] S. K. Saraswati *et al.*, "Literature Review : Faktor Risiko Penyebab Obesitas," *Media Kesehat. Masy. Indones.*, vol. 20, no. 1, pp. 70–74, 2021, doi: 10.14710/mkmi.20.1.70-74.
- [2] L. Setiyani, A. N. Indahsari, and R. Roestam, "Analisis Prediksi Level Obesitas Menggunakan Perbandingan Algoritma Machine Learning dan Deep Learning," *JTERA (Jurnal Teknol. Rekayasa)*, vol. 8, no. 1, p. 139, 2023, doi: 10.31544/jtera.v8.i1.2022.139-146.
- [3] N. Mohammadian Khonsari *et al.*, "Normal Weight Obesity and Cardiometabolic Risk Factors: A Systematic Review and Meta-Analysis," *Front. Endocrinol. (Lausanne)*, vol. 13, no. March, 2022, doi: 10.3389/fendo.2022.857930.
- [4] S. Y. Sibi and A. R. Widiarti, "Klasifikasi Tingkat Obesitas Mempergunakan Algoritma KNN," *Semin. Nas. Corisindo*, vol. 7, no. 2, pp. 370–375, 2022.
- [5] H. T. Santoso, F. A. Felmidi, A. Nur, A. Ristyawan, and E. Daniati, "Analisis Kinerja Algoritma Data Mining pada Klasifikasi Tingkat Obesitas dengan K-Fold Cross Validation dan AUC," vol. 8, pp. 113–122, 2024.



- [6] A. B. Putri and A. Makmun, "Pola Makan terhadap Obesitas," *Indones. J. Heal.*, vol. xx, no. xx, pp. 68–76, 2021, doi: 10.33368/inajoh.v2i1.39.
- [7] D. Alfora, E. Saori, and L. N. Fajriah, "Pengaruh konsumsi makanan cepat saji terhadap gizi remaja," *FLORONA J. Ilm. Kesehat.*, vol. 2, no. 1, pp. 43–49, 2023, doi: 10.55904/florona.v2i1.688.
- [8] E. Yuniarti, "Hubungan Konsumsi Sayur dan Buah dengan Kegemukan Remaja di Kota Padang," *J. Sehat Mandiri*, vol. 18, no. 1, pp. 137–145, 2023, doi: 10.33761/jsm.v18i1.974.
- [9] D. I. Gizi, F. Kedokteran, and U. Diponegoro, "ASUPAN ENERGI CEMILAN, DURASI DAN KUALITAS TIDUR PADA REMAJA OBESITAS DAN NON OBESITAS," vol. 7, no. 2010, pp. 147–154, 2018.
- [10] D. Librado and A. H. Nasyuha, "Penerapan Data Mining Untuk Klasifikasi Penerima Kredit Dengan Perbandingan Algoritma Naive Bayes dan Algoritma C4.5," *J. Media Inform. Budidarma*, vol. 7, no. 4, p. 1952, 2023, doi: 10.30865/mib.v7i4.6907.
- [11] M. Azhari, Z. Situmorang, and R. Rosnelly, "Perbandingan Akurasi, Recall, dan Presisi Klasifikasi pada Algoritma C4.5, Random Forest, SVM dan Naive Bayes," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 640, 2021, doi: 10.30865/mib.v5i2.2937.
- [12] W. Apriliah, I. Kurniawan, M. Baydhowi, and T. Haryati, "Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest," *Sistemasi*, vol. 10, no. 1, p. 163, 2021, doi: 10.32520/stmsi.v10i1.1129.
- [13] D. Mualfah, W. Fadila, and R. Firdaus, "Teknik SMOTE untuk Mengatasi Imbalance Data pada Deteksi Penyakit Stroke Menggunakan Algoritma Random Forest," *J. CoSciTech (Computer Sci. Inf. Technol.)*, vol. 3, no. 2, pp. 107–113, 2022, doi: 10.37859/coscitech.v3i2.3912.
- [14] Arifin Yusuf Permana, Hari Noer Fazri, M. Fakhri Nur Athoilah, Mohammad Robi, and Ricky Firmansyah, "Penerapan Data Mining Dalam Analisis Prediksi Kanker Paru Menggunakan Algoritma Random Forest," *J. Ilm. Tek. Inform. dan Komun.*, vol. 3, no. 2, pp. 27–41, 2023, doi: 10.55606/juitik.v3i2.472.
- [15] E. Dwi *et al.*, "Penggunaan Data Mining untuk Prediksi tingkat Obesitas di Meksiko Menggunakan Metode Random Forest," *Agustus*, vol. 8, pp. 2549–7952, 2024.
- [16] N. H. Alfajr and S. Defiyanti, "METODE RANDOM FOREST DAN PENERAPAN PRINCIPAL COMPONENT ANALYSIS (PCA)," vol. 12, no. 3, 2024.
- [17] A. Toha, P. Purwono, and W. Gata, "Model Prediksi Kualitas Udara dengan Support Vector Machines dengan Optimasi Hyperparameter GridSearch CV," *Bul. Ilm. Sarj. Tek. Elektro*, vol. 4, no. 1, pp. 12–21, 2022, doi: 10.12928/biste.v4i1.6079.
- [18] D. Iskandar, "Optimasi Parameter Random Forest menggunakan Grid Search Untuk Analisis Time series," *Petir*, vol. 16, no. 2, pp. 267–277, 2023, doi: 10.33322/petir.v16i2.2084.
- [19] D. Benaya, "Implementasi Random Forest dalam Klasifikasi Kanker Paru-Paru," *JOINTER J. Informatics Eng.*, vol. 5, no. 01, pp. 27–31, 2024, doi: 10.53682/jointer.v5i01.331.
- [20] Z. Maisat, E. Darmawan, and A. Fauzan, "Implementasi Optimasi Hyperparameter GridSearchCV Pada Sistem Prediksi Serangan Jantung Menggunakan SVM Implementation of GridSearchCV Hyperparameter Optimization in Heart Attack Prediction System Using SVM," *Unipdu*, vol. 13, no. 1, pp. 8–15, 2023.