

Penerapan SVM dalam Prediksi Resiko Obesitas berdasarkan Gaya Hidup dan Pola Makan

Implementation of SVM in Predicting Obesity Risk Based on Lifestyle and Dietary Patterns

Adinda Febiola^{1*}, Fahriya Arдинингрум², Michael Orlando A. Purba³, Fernando Siahaan⁴, Victor Asido Elyakim P⁵
^{1,2,3,4,5}STIKOM Tunas Bangsa, Indonesia

Article Info

Genesis Artikel:

Diterima, 18 Februari 2025
 Direvisi, 06 Maret 2025
 Disetujui, 20 Maret 2025

Kata Kunci:

Obesitas
 Gaya Hidup
 Support Vector Machine
 Klasifikasi
 Kecerdasan Buatan

Keywords:

Obesity
 Lifestyle
 Support Vector Machine
 Classification
 Artificial Intelligence

ABSTRAK

Obesitas merupakan salah satu isu kesehatan global yang mengalami peningkatan signifikan dalam beberapa dekade terakhir. Kondisi ini berkaitan erat dengan pola hidup modern yang tidak seimbang, seperti kurangnya aktivitas fisik, pola makan tidak sehat, serta kebiasaan merokok dan konsumsi alkohol. Penelitian ini bertujuan untuk menganalisis hubungan antara gaya hidup dan risiko obesitas, serta mengevaluasi efektivitas metode Support Vector Machine (SVM) dalam memprediksi tingkat risiko obesitas. Dataset yang digunakan diperoleh dari platform Kaggle, mencakup berbagai variabel seperti usia, jenis kelamin, indeks massa tubuh (BMI), kebiasaan makan, pola tidur, dan aktivitas fisik. Proses preprocessing dilakukan melalui normalisasi data dan pengkodean variabel kategorikal untuk memastikan kesiapan data sebelum dimasukkan ke dalam model. Model SVM dilatih menggunakan berbagai rasio pembagian data pelatihan dan pengujian, dan menunjukkan tingkat akurasi yang sangat tinggi, bahkan mencapai 100% pada beberapa skenario. Hasil ini menunjukkan bahwa SVM mampu mengidentifikasi pola dalam data gaya hidup yang berkontribusi terhadap obesitas secara efektif. Dengan demikian, penerapan SVM dapat menjadi alat prediktif yang bermanfaat bagi tenaga kesehatan dalam merancang strategi pencegahan obesitas berbasis data yang lebih akurat dan efisien.

ABSTRACT

Obesity is one of the global health issues that has seen a significant increase in recent decades. This condition is closely related to an unbalanced modern lifestyle, such as lack of physical activity, unhealthy eating patterns, and habits of smoking and alcohol consumption. This study aims to analyze the relationship between lifestyle and obesity risk, as well as to evaluate the effectiveness of the Support Vector Machine (SVM) method in predicting the level of obesity risk. The dataset used was obtained from the Kaggle platform, covering various variables such as age, gender, body mass index (BMI), eating habits, sleep patterns, and physical activity. Preprocessing was carried out through data normalization and encoding of categorical variables to ensure data readiness before being input into the model. The SVM model was trained using various training and testing data split ratios and showed a very high accuracy rate, even reaching 100% in some scenarios. These results demonstrate that SVM can effectively identify patterns in lifestyle data that contribute to obesity. Thus, the application of SVM can be a useful predictive tool for healthcare professionals in designing more accurate and efficient data-driven obesity prevention strategies.

This is an open access article under the [CC BY-SA](#) license.



Penulis Korespondensi:

Adinda Febiola,
 Program Studi Teknik Informatika,
 STIKOM Tunas Bangsa, Indonesia
 Email: adindafebiola2005@gmail.com

1. PENDAHULUAN

Obesitas merupakan masalah kesehatan global yang terus meningkat dalam beberapa dekade terakhir. WHO mendefinisikannya sebagai akumulasi lemak tubuh berlebihan yang dapat meningkatkan risiko penyakit kronis seperti diabetes tipe 2, hipertensi, dan gangguan metabolismik lainnya[1]. Data WHO menunjukkan bahwa pada tahun 2022, sekitar 2,5 miliar orang dewasa mengalami kelebihan berat badan, dengan 890 juta di antaranya mengalami obesitas[2]. Angka ini meningkat drastis dibandingkan tahun 1990, ketika hanya 25% orang dewasa mengalami kelebihan berat badan. Prevalensi obesitas bervariasi di berbagai wilayah, dengan angka terendah di Asia Tenggara dan Afrika (31%) serta tertinggi di Amerika (67%). Kasus obesitas

pada anak-anak juga meningkat, dengan 37 juta anak di bawah 5 tahun mengalami kelebihan berat badan pada 2022. Di Afrika, jumlah anak-anak yang kelebihan berat badan naik hampir 23% sejak 2000, sementara di Asia, hampir separuh kasus obesitas anak dunia terjadi[3]. Pada kelompok usia 5–19 tahun, jumlah anak dan remaja dengan kelebihan berat badan meningkat dari 8% pada 1990 menjadi 20% pada 2022, dengan kasus obesitas melonjak dari 2% menjadi 8% dalam periode yang sama. Fenomena ini juga terjadi di Indonesia, di mana prevalensi obesitas terus meningkat setiap tahunnya. Penyebab obesitas bersifat kompleks, dipengaruhi oleh faktor genetik, lingkungan, dan perilaku individu. Perubahan pola makan dan gaya hidup modern yang kurang aktif berkontribusi besar terhadap peningkatan obesitas[4]. Oleh karena itu, memahami faktor risiko obesitas menjadi kunci dalam merancang strategi pencegahan dan intervensi yang lebih efektif.

Pola makan dan gaya hidup memiliki peran krusial dalam menentukan kondisi kesehatan seseorang, termasuk risiko mengalami obesitas[5]. Konsumsi makanan yang tinggi kalori, rendah serat, dan kaya akan lemak jenuh menjadi salah satu faktor utama yang mendorong peningkatan berat badan[6]. Selain itu, kebiasaan makan yang tidak teratur, sering mengonsumsi makanan cepat saji, serta asupan gula yang berlebihan turut berkontribusi terhadap terjadinya obesitas. Pola makan yang tidak seimbang dapat menyebabkan ketidakseimbangan energi dalam tubuh, di mana jumlah kalori yang dikonsumsi lebih besar dibandingkan dengan yang digunakan, sehingga kelebihan energi tersebut disimpan sebagai lemak tubuh yang dapat memicu obesitas dalam jangka panjang. Selain pola makan, gaya hidup juga memiliki pengaruh besar terhadap perkembangan obesitas[7]. Minimnya aktivitas fisik akibat gaya hidup yang cenderung pasif, seperti bekerja dalam posisi duduk dalam waktu yang lama, kurangnya olahraga, serta kebiasaan tidur yang tidak teratur, telah terbukti berdampak pada peningkatan berat badan. Selain itu, kemajuan teknologi yang mendorong penggunaan perangkat elektronik seperti televisi, komputer, dan ponsel dalam kehidupan sehari-hari telah mengurangi aktivitas fisik, sehingga semakin meningkatkan risiko obesitas[8]. Kurangnya aktivitas fisik membuat tubuh tidak mampu membakar kalori secara optimal, yang pada akhirnya menyebabkan penumpukan lemak. Selain faktor pola makan dan gaya hidup, aspek psikologis serta lingkungan sosial juga turut memengaruhi kebiasaan seseorang[9]. Stres yang berkepanjangan dapat mengubah pola makan menjadi tidak sehat, misalnya dengan kecenderungan mengonsumsi makanan berkalori tinggi sebagai bentuk pelampiasan atau emotional eating. Faktor sosial, seperti lingkungan keluarga dan pergaulan, juga dapat membentuk pola makan dan tingkat aktivitas fisik seseorang. Individu yang tumbuh dalam lingkungan yang memiliki kebiasaan makan kurang sehat serta kurangnya aktivitas fisik cenderung memiliki kemungkinan lebih besar untuk mengalami obesitas.

Dalam beberapa tahun terakhir, banyak penelitian telah dilakukan untuk mengembangkan metode yang lebih akurat dalam memprediksi risiko obesitas. Salah satu pendekatan yang banyak digunakan adalah pemanfaatan kecerdasan buatan dan pembelajaran mesin, khususnya Support Vector Machine (SVM). Beberapa penelitian terdahulu yang telah menerapkan SVM dalam analisis obesitas yaitu Penelitian oleh Rahmawati et al. (2021), studi ini menunjukkan bahwa SVM memiliki akurasi sekitar 82% dalam memprediksi risiko obesitas di kalangan mahasiswa berdasarkan pola makan dan aktivitas fisik[10]. Namun, penelitian ini memiliki keterbatasan dalam jumlah sampel yang relatif kecil serta kurangnya variasi dalam data yang digunakan, yang dapat mempengaruhi generalisasi model. Selanjutnya penelitian oleh Sari & Nugroho (2020), penelitian ini menemukan bahwa faktor konsumsi makanan cepat saji dan kebiasaan kurang tidur memiliki pengaruh signifikan terhadap obesitas, dengan SVM menunjukkan performa yang lebih baik dibandingkan metode regresi logistik. Namun, penelitian ini belum mempertimbangkan faktor psikologis dan genetik yang juga berperan dalam obesitas[10]. Dan Terakhir penelitian oleh Yulianto et al. (2019): Studi ini mengevaluasi efektivitas SVM dalam mendeteksi obesitas pada anak-anak di Indonesia berdasarkan pola makan dan aktivitas fisik mereka. Model yang dikembangkan menunjukkan akurasi lebih dari 80%, tetapi penelitian ini masih terbatas pada dataset yang bersumber dari satu wilayah tertentu, sehingga hasilnya mungkin tidak dapat digeneralisasi ke populasi yang lebih luas[11].

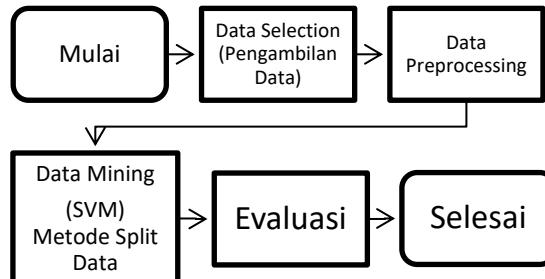
Salah satu solusi untuk mengatasi tantangan dalam memprediksi risiko obesitas adalah penerapan kecerdasan buatan, khususnya algoritma Support Vector Machine (SVM). SVM merupakan teknik machine learning yang banyak diterapkan dalam tugas klasifikasi dan regresi, termasuk di bidang kesehatan[12]. Dengan menganalisis data mengenai kebiasaan hidup dan pola makan individu, SVM dapat mengidentifikasi pola tertentu dari data historis untuk menentukan tingkat risiko obesitas seseorang[13]. Dalam analisis prediksi obesitas, SVM dapat mengklasifikasikan individu berdasarkan berbagai faktor seperti indeks massa tubuh (BMI), pola konsumsi makanan, tingkat aktivitas fisik, asupan gula dan lemak, serta kualitas tidur[14]. Dengan mengolah informasi ini, SVM mampu menghasilkan prediksi yang lebih akurat, sehingga dapat menjadi alat bantu bagi tenaga medis atau ahli gizi dalam merancang intervensi yang tepat bagi individu yang memiliki risiko tinggi mengalami obesitas. Keunggulan utama SVM dibandingkan metode lain terletak pada kemampuannya dalam menangani data berdimensi tinggi dan melakukan generalisasi yang lebih baik terhadap data baru[15]. Selain itu, SVM lebih tahan terhadap overfitting, terutama pada dataset dengan jumlah sampel yang terbatas, dibandingkan algoritma seperti decision tree atau artificial neural network[16]. SVM juga dapat bekerja dengan baik meskipun terdapat data yang tidak terdistribusi secara linier, berkat penggunaan fungsi kernel yang memungkinkan pemetaan ke ruang fitur berdimensi lebih tinggi[17].

Penggunaan Support Vector Machine (SVM) dalam memprediksi risiko obesitas berdasarkan gaya hidup dan pola makan memiliki potensi besar dalam mendukung upaya pencegahan sejak dulu. Dengan kemampuannya dalam menganalisis data kompleks dan mengidentifikasi pola tersembunyi, SVM dapat membantu tenaga kesehatan dalam mengambil keputusan yang lebih akurat[18]. Diharapkan, hasil penelitian ini dapat menjadi acuan bagi tenaga medis dan pembuat kebijakan dalam merancang strategi pencegahan yang lebih efektif, sehingga dapat menekan angka obesitas di masyarakat.

2. METODE PENELITIAN

2.1 Tahapan Penelitian

Pendekatan penelitian ini menggunakan metode eksperimen berbasis Data Mining, dengan tahapan sistematis mulai dari pengumpulan data, pra-pemrosesan, hingga pengembangan dan evaluasi model, yang divisualisasikan pada gambar 1.



Gambar 1. Kerangka Penelitian

2.2 Dataset Penelitian

Dataset yang digunakan dalam penelitian ini adalah dataset publik yang diambil dari laman Kaggle (<https://www.kaggle.com/code/pranita1202/obesityanalysis/notebook>). Dataset ini berisi record dengan 15 variabel yang mencakup karakteristik demografis (jenis kelamin, usia, tinggi badan), faktor keluarga (riwayat obesitas), pola makan (konsumsi makanan cepat saji, frekuensi makan sayur, jumlah makanan utama per hari, konsumsi makanan di luar waktu makan utama, dan asupan cairan harian), serta gaya hidup (status merokok, frekuensi aktivitas fisik, waktu yang dihabiskan untuk aktivitas berbasis teknologi, dan jenis transportasi yang digunakan). Variabel target adalah obesitas, yang diklasifikasikan ke dalam empat kategori: underweight, normal, overweight, dan obese.

2.3 Pembagian Dataset

Setelah data dibersihkan, langkah berikutnya adalah membagi dataset menjadi fitur X dan target y. Fitur X mencakup semua variabel yang digunakan untuk memprediksi target, sementara target Y adalah variabel yang akan diprediksi. Pembagian dataset dilakukan dengan membagi data menjadi dua bagian, di mana data pelatihan dan data pengujian. Tujuannya adalah untuk melatih model menggunakan sebagian besar data dan menguji kinerjanya pada data yang belum pernah dilihat sebelumnya.

Penelitian ini juga memastikan keseimbangan antara data pelatihan dan pengujian, mengingat dataset terdiri dari 100 record. Pembagian data pelatihan untuk memberikan cukup data untuk model mempelajari pola dengan baik, sementara data pengujian untuk menyediakan jumlah data yang cukup untuk mengevaluasi kemampuan generalisasi model di semua kelas. Pendekatan ini didukung oleh praktik umum dalam penelitian machine learning, di mana lebih banyak data pengujian membantu menghasilkan metrik evaluasi yang lebih stabil dan andal. Selain itu, validasi performa model dilakukan menggunakan cross-validation pada data pelatihan untuk mengurangi risiko bias akibat pembagian data, sehingga rasio ini memberikan keseimbangan yang optimal antara akurasi pelatihan dan kemampuan generalisasi model.

2.4 Algoritma SVM

Support Vector Machine (SVM) adalah algoritma yang banyak digunakan dalam data mining untuk tugas klasifikasi dan regresi, meskipun lebih sering diterapkan untuk klasifikasi[19]. SVM bekerja dengan mencari hiperplan terbaik yang dapat memisahkan dua kelas data pada ruang fitur. Tujuan utamanya adalah menemukan batas pemisah (decision boundary) yang memaksimalkan margin, yaitu jarak terdekat antara titik-titik data dari kedua kelas yang berbeda ke hiperplan pemisah[20]. Titik-titik data yang berada di dekat atau pada margin disebut support vectors, dan titik-titik inilah yang menentukan posisi dari hiperplan tersebut[13].

Jika data tidak bisa dipisahkan secara linear, SVM menggunakan fungsi kernel untuk memetakan data ke ruang dimensi yang lebih tinggi, di mana data dapat dipisahkan secara linear. Kernel yang umum digunakan termasuk kernel linear, polynomial, dan radial basis function (RBF) [18]. SVM sangat efektif untuk menangani data berdimensi tinggi dan tetap bekerja dengan baik meski ukuran dataset kecil[19]. SVM digunakan dalam berbagai aplikasi seperti pengenalan pola, teks mining, dan bioinformatika.

Rumus dasar SVM untuk menemukan hiperplan optimal adalah:

$$w \cdot x - b =$$

Di mana:

- *w* adalah vektor bobot,
- *x* adalah vektor fitur,
- *b* adalah bias atau intercept.

SVM memaksimalkan margin, dengan syarat untuk setiap data:

$$y_i(w \cdot x_i - b) \geq 1$$

Dimana y_i adalah label kelas (1 atau -1) x_i adalah data, dan w dan b adalah parameter model yang harus dioptimalkan

2.5 Matriks Evaluasi

Matriks evaluasi kinerja model dapat diperoleh dari confusion matrix. Dalam confusion matrix, baris menunjukkan jumlah sampel aktual pada set pengujian, sedangkan kolom mewakili hasil prediksi model. Confusion matrix memiliki empat parameter utama: true positive (TP), true negative (TN), false positive (FP), dan false negative (FN). Kinerja model dihitung berdasarkan parameter-parameter ini, seperti yang ditunjukkan pada persamaan 1, 2, 3, dan 4.

Rumus:

$$Recall = \frac{TP}{TP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$F1 - Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (3)$$

$$Akurasi = \frac{\text{Correctly classified Values}}{\text{total number of values}}$$

2.6 Dataset

Dataset yang digunakan berisi berbagai variabel yang berkaitan dengan obesitas dan gaya hidup individu. Dalam konteks penelitian, dataset ini memiliki potensi besar untuk digunakan dalam analisis faktor risiko obesitas. Berikut datasetnya:

Tabel 1. Dataset Obesitas

Age	Gender	ObesityCategory
46	Female	Obese
32	Female	Normal
25	Female	Normal
38	Male	Normal
56	Female	Normal
36	Male	Obese
28	Female	Normal
28	Female	Normal
41	Male	Normal
53	Male	Obese
57	Female	Obese
41	Female	Normal
....
26	Male	Normal
26	Male	Normal

3. HASIL DAN PEMBAHASAN

3.1 Pengolahan Data (Preprocessing)

Setelah pengumpulan dan kategorisasi data selesai, langkah berikutnya adalah pengolahan data atau preprocessing. Preprocessing dilakukan untuk membersihkan dan mempersiapkan data agar siap digunakan dalam model klasifikasi Naïve Bayes. Berikut adalah langkah-langkah preprocessing yang diterapkan pada dataset ini.

1. Pembersihan Data (Pembersihan Data)

Dataset yang digunakan terdiri dari 100 entri dengan 11 kolom. Hasil pemeriksaan awal menunjukkan bahwa tidak ada nilai yang hilang dalam dataset, sehingga tidak diperlukan proses imputasi data.

2. Pengkodean Variabel Kategorikal

Beberapa kolom dalam dataset berisi kategori data, yaitu: Jenis Kelamin (Jenis kelamin), Smoking (Merokok), Konsumsi Alkohol (Konsumsi alkohol), Kategori Obesitas (Kategori obesitas, yang berfungsi sebagai label untuk klasifikasi).

Agar dapat digunakan dalam algoritma Support Vector Machine (SVM), variabel-variabel ini dikonversi ke dalam bentuk numerik menggunakan metode Label Encoding. Proses ini mengubah setiap kategori menjadi angka sehingga dapat dipahami oleh model pembelajaran mesin.

3. Normalisasi Data

Algoritma SVM sangat dipengaruhi oleh skala data, sehingga fitur numerik perlu dinormalisasi agar memiliki distribusi yang lebih seragam. Kolom yang dinormalisasi meliputi: Age (Usia), BMI (Indeks Massa Tubuh), PhysicalActivity (Frekuensi aktivitas fisik), SleepHours (Jumlah jam tidur), FastFoodIntake (Frekuensi konsumsi makanan cepat saji), VegetableIntake (Jumlah konsumsi sayuran), DailyCalorieIntake (Asupan kalori harian)

Normalisasi dilakukan menggunakan metode StandardScaler, yang mentransformasikan setiap fitur agar memiliki distribusi dengan mean = 0 dan standar deviasi = 1. Tahapan normalisasi meliputi:

- Menghitung rata-rata (mean) dan standar deviasi dari setiap fitur numerik pada data latih.
- Mengurangkan nilai setiap fitur dengan rata-rata (mean centering) agar memiliki pusat distribusi di nol.
- Membagi hasil pengurangan dengan standar deviasi sehingga nilai berada dalam skala yang seragam.
- Menerapkan transformasi yang sama pada uji data, menggunakan mean dan standar deviasi dari data latih agar skala tetap konsisten.

Dengan normalisasi ini, fitur dengan skala yang berbeda dapat dibandingkan secara lebih adil dalam model SVM, sehingga meningkatkan efektivitas proses klasifikasi.

3.2 Pembagian Data Training dan Data Testing

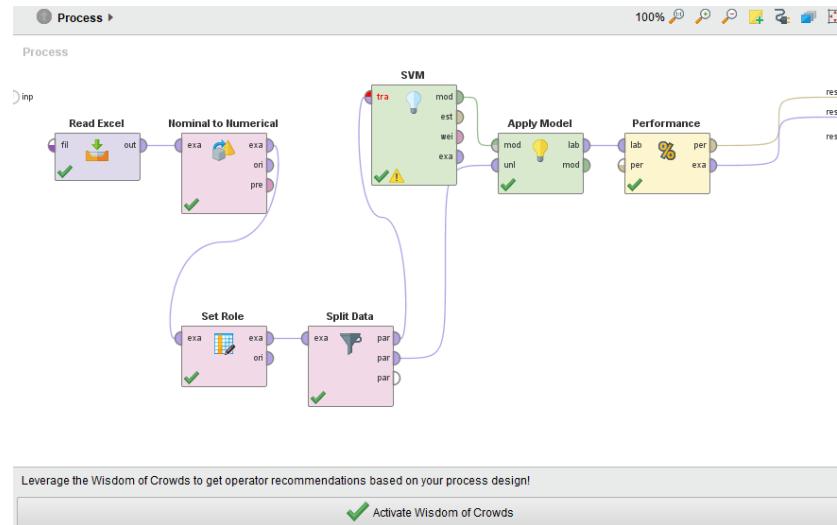
Data training merupakan data yang digunakan dalam proses pelatihan model klasifikasi. Sementara itu, data testing adalah data yang digunakan untuk menguji aturan klasifikasi yang telah dibangun. Dalam penelitian ini, pembagian antara data training dan data testing dilakukan dengan rasio 60:40, 70:30, dan 80:20, sebagaimana terlihat pada Tabel 2. Metode yang digunakan untuk pembagian ini adalah stratified sampling.

Tabel 2. Sebaran Data Training dan Data Testing pada Dataset

Perbandingan 50:50			
Kategori	Training	Testing	Total
Normal	20	20	40
Obeses	30	30	60
Total	50	50	100
Perbandingan 60:40			
Kategori	Training	Testing	Total
Normal	24	16	40
Obeses	36	24	60
Total	60	40	100
Perbandingan 70:30			
Kategori	Training	Testing	Total
Normal	28	12	40
Obeses	42	18	60
Total	70	30	100
Perbandingan 80:20			
Kategori	Training	Testing	Total
Normal	32	8	40
Obeses	48	12	60
Total	80	20	100

Pada Tabel 2, ditampilkan sebaran antara data pelatihan dan data pengujian dalam dataset. Data yang digunakan untuk pelatihan (training) memiliki proporsi yang seimbang antara data Baik dan data Bermasalah untuk setiap model yang dibandingkan. Pengambilan sampel dilakukan secara proporsional (stratified sampling) agar distribusi kelas tetap terjaga keseimbangannya.

3.3 Modelling Menggunakan Split Data



Gambar 2. Metode SVM Menggunakan Split Data

- Read Excel, membaca dataset dari file Excel.
- Nominal to Numerical, mengonversi data nominal (kategorikal) menjadi angka.
- Set Role, menentukan atribut target sebagai label untuk klasifikasi.
- Split Data, membagi dataset menjadi training dan testing sesuai rasio tertentu.
- Naïve Bayes, melatih model menggunakan data training.
- Apply Model, menerapkan model yang telah dilatih pada data testing.
- Performance, mengevaluasi hasil prediksi model dengan metrik seperti accuracy, precision, dan recall.

3.4 Hasil Pemodelan Menggunakan Split Data

Setelah menyelesaikan tahap pemodelan pada data pelatihan dan data pengujian yang telah melalui proses preprocessing, nilai akurasi yang dihasilkan dapat dilihat pada Tabel 3. yang menunjukkan akurasi berdasarkan pemisahan data.

Tabel 3. Akurasi Menggunakan Split Data

Model	Akurasi	Precision	Precision	Recall	Recall
		(Normal)	(Obeses)	(Normal)	(Obeses)
50:50:00	91.89%	100%	88.89%	76.92%	100%
60:40:00	100%	100%	100%	100%	100%
70:30:00	100%	100%	100%	100%	100%
80:20:00	100%	100%	100%	100%	100%

3.5 Pengujian dan Evaluasi

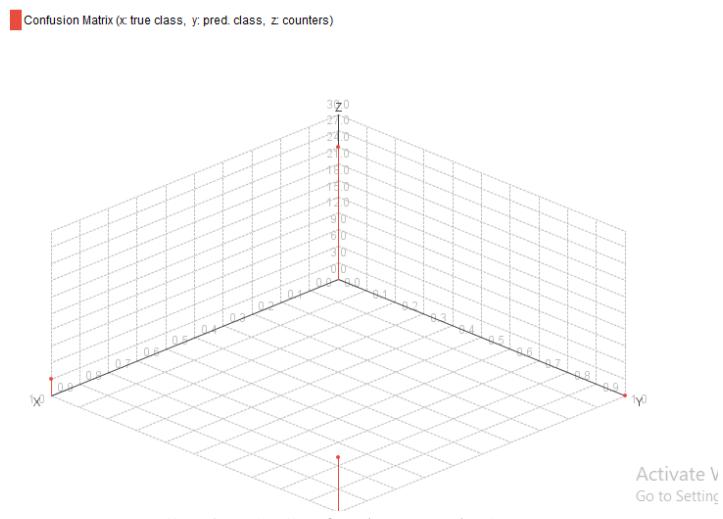
Dalam penelitian ini, pengujian akan dilakukan dengan memanfaatkan confusion matrix yang dihasilkan dari tahapan pemodelan menggunakan metode Naïve Bayes. Tabel 4 akan menyajikan akurasi berdasarkan pembagian data yang telah dilakukan. Matrix yang disajikan pada Tabel 5. adalah hasil dari evaluasi dan pengukuran klasifikasi Naïve Bayes dengan dataset yang berjumlah 364, terdapat 218 data training dan 146 data testing.

Tabel 4. Confusion Matrix

	TRUE	TRUE	Class
	Obeses	Normal	precision
<i>Pred.Normal</i>	0	26	100%
<i>Pred. Obeses</i>	48	0	100%
<i>Class recall</i>	100%	100%	

Tabel 4 menyajikan confusion matrix yang digunakan untuk mengevaluasi kinerja model. Model ini menunjukkan kemampuan yang baik dalam mengklasifikasikan sebagian besar sampel, dengan tingkat precision sebesar 100% untuk kelas "Normal" dan 100% untuk kelas "Obeses". Selain itu, nilai recall untuk kelas "Normal" tercatat sebesar 100%, sementara untuk kelas "Obeses" angkanya lebih tinggi, yaitu 100%. Meskipun demikian, jumlah prediksi yang salah masih cukup signifikan, terutama pada kelas "Baik", dan hal ini dapat memengaruhi akurasi keseluruhan model.

- a) Precision kelas Baik = $\frac{TP}{TP+FP} = \frac{38}{38+17} = 69.09\%$
- b) Precision kelas Bermasalah = $\frac{TP}{TP+FP} = \frac{71}{71+19} = 78.89\%$
- c) Recall kelas Baik = $\frac{TP}{TP+FP} = \frac{38}{38+19} = 66.67\%$
- d) Recall kelas Bermasalah = $\frac{TP}{TP+FP} = \frac{71}{71+17} = 80.68\%$
- e) Accuracy = $\frac{TP+TN}{TP+TN+FP+FN} = \frac{71+38}{71+38+19+17} = 75.17\%$



Gambar 3. Confussion Matrix 3D

Gambar 3. menampilkan confusion matrix dalam bentuk visualisasi 3D, dengan sumbu X sebagai kelas asli, sumbu Y sebagai kelas prediksi, dan sumbu Z menunjukkan jumlah sampel. Dari tampilan grafik, terlihat bahwa sebagian besar data berada di titik tertentu, mengindikasikan ketidakseimbangan prediksi model. Hal ini menunjukkan bahwa model cenderung lebih sering memprediksi satu kelas dibandingkan kelas lainnya, yang dapat memengaruhi akurasi dalam mengenali kedua kelas secara seimbang.

4. KESIMPULAN

Penelitian ini menyoroti pentingnya identifikasi dini terhadap risiko obesitas yang dipengaruhi oleh faktor gaya hidup dan pola makan. Dengan menggunakan pendekatan kecerdasan buatan, khususnya algoritma Support Vector Machine (SVM), penelitian ini berhasil memprediksi tingkat risiko obesitas berdasarkan variabel-variabel seperti usia, jenis kelamin, indeks massa tubuh (BMI), aktivitas fisik, pola tidur, kebiasaan merokok, konsumsi alkohol, serta asupan makanan seperti fast food dan sayuran. Hasil pemodelan menunjukkan bahwa SVM mampu mengklasifikasikan risiko obesitas dengan tingkat akurasi yang tinggi, bahkan mencapai 100% pada pembagian data tertentu. Ini menunjukkan bahwa SVM sangat efektif dalam mendeteksi pola kompleks dalam data kesehatan dan gaya hidup, serta memiliki ketahanan terhadap data berdimensi tinggi dan ukuran sampel yang relatif kecil. Faktor-faktor gaya hidup terbukti memiliki pengaruh signifikan terhadap peningkatan risiko obesitas, terutama aktivitas fisik yang rendah dan pola makan tidak sehat. Oleh karena itu, implementasi metode SVM dapat menjadi alat bantu yang potensial bagi tenaga kesehatan dan pengambil kebijakan untuk menyusun program pencegahan obesitas yang lebih akurat, efisien, dan berbasis data. Ke depan, penelitian ini dapat dikembangkan dengan melibatkan lebih banyak data dari berbagai wilayah serta mempertimbangkan faktor tambahan seperti kondisi psikologis dan genetik, guna menghasilkan model prediksi yang lebih komprehensif dan aplikatif dalam skala nasional.

REFERENSI

- [1] 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.
- [2] A. I. Putri, Y. Syarif, P. Jayadi, F. Arrazak, and F. N. Salisah, "Implementasi Algoritma Decision Tree dan Support Vector Machine (SVM) untuk Prediksi Risiko Stunting pada Keluarga," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 3, no. 2, pp. 349–357, 2024, doi: 10.57152/malcom.v3i2.1228.
- [3] N. Istanti, Y. Ernawati, and A. N. Antara, "Faktor-faktor yang mempengaruhi risiko obesitas pada remaja di Panti Asuhan Darun Najah Sleman Yogyakarta," vol. 12, no. 2, pp. 206–217, 2024, [Online]. Available: <https://ejournal.unsrat.ac.id/v3/index.php/jkp/article/view/55870/47465>
- [4] L. Fitria, Y. Yarmaliza, and Z. Zalmaliza, "Evaluasi Perilaku Masyarakat Terhadap Faktor Resiko Kejadian Hipertensi Desa Purwodadi Tahun 2022," *J-KESMAS J. Kesehat. Masy.*, vol. 8, no. 1, p. 73, 2022, doi: 10.35329/jkesmas.v8i1.2858.
- [5] Azmy Mu'thia Hanum, "FAKTOR-FAKTOR PENYEBAB TERjadinya OBESITAS PADA REMAJA Azmy Mu' thia Hanum Program Studi Gizi, Fakultas Kesehatan Masyarakat, Universitas Airlangga," vol. 9, no. 2, pp. 137–147, 2023.
- [6] R. Rosmiati, N. R. Haryana, H. Firmansyah, and R. Purba, "Pola Makan, Aktivitas Fisik dan Obesitas pada Pekerja Urban di Indonesia," *Amerta Nutr.*, vol. 7, no. 2SP, pp. 164–170, 2023, doi: 10.20473/amnt.v7i2sp.2023.164-170.
- [7] A. Khikam, N. M. Anggadimas, M. Udin, T. Informatika, U. M. Pasuruan, and K. Pasuruan, "IMPLEMENTASI DECISION TREE UNTUK KLASIKASI OBESITAS," vol. 9, no. 3, pp. 3946–3952, 2025.
- [8] A. S. . Meiliana, F. Bachtiar, Condrowati, and F. Nazhira, "Hubungan antara Aktivitas Fisik dengan Indeks Massa Tubuh pada Situasi Pandemi Covid-19," *Sains Olahraga dan Pendidik. Jasmani*, vol. 4, no. 2, pp. 1–14, 2022.
- [9] 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.
- [10] S. A. Utiahman, A. Mulawati, and M. Pratama, "Penerapan Support Vector Machine dan Random Forest Classifier Untuk Klasifikasi Tingkat Obesitas," vol. 14, no. 3, pp. 754–760, 2024.
- [11] U. Muhammadiyah, M. Aceh, and U. B. Nusantara, "Penggunaan Algoritma Support Vector Machine (SVM) Untuk Deteksi Penipuan pada Transaksi Online," vol. 13, pp. 1627–1632, 2024.
- [12] W. Athira Luqyana, I. Cholissodin, and R. S. Perdana, "Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 11, pp. 4704–4713, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [13] N. G. Ramadhan and A. Khoirunnisa, "Klasifikasi Data Malaria Menggunakan Metode Support Vector Machine," *J. Media Inform. Budidarma*, vol. 5, no. 4, p. 1580, 2021, doi: 10.30865/mib.v5i4.3347.
- [14] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021.
- [15] M Riski Qisthiano, "Klasifikasi Terhadap Prediksi Kelulusan Mahasiswa Dengan Menggunakan Metode Support Vector Machine (Svm)," *Semin. Nas. Teknol. dan Multidisiplin Ilmu*, vol. 2, no. 2, pp. 203–207, 2022, doi: 10.51903/semnastekmu.v2i1.170.
- [16] R. Nanda, E. Haerani, S. K. Gusti, and S. Ramadhani, "Klasifikasi Berita Menggunakan Metode Support Vector Machine," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 2, pp. 269–278, 2022, doi: 10.32672/jnkti.v5i2.4193.
- [17] Septi Putri, Yohanes Agung Apriyanto, and Andri Wijaya, "Analisis Sentimen Ulasan Aplikasi DeepL Pada Google Play Dengan Metode Support Vector Machine (Svm)," *J. Sist. Inf.*, vol. 4, no. 2, pp. 59–66, 2023, doi: 10.32546/jusin.v4i2.2368.
- [18] M. . Imelda A.Muis & Muhammad Affandes, "Penerapan Metode Support Vector Machine (SVM) Menggunakan Kernel Radial Basis Function (RBF) Pada Klasifikasi Tweet," *Sains, Teknol. dan Ind. Sultan Syarif Kasim Riau*, vol. 12, no. 2, pp. 189–197, 2015.
- [19] M. G. Pradana, P. H. Saputro, and D. P. Wijaya, "Komparasi Metode Support Vector Machine Dan Naïve Bayes Dalam Klasifikasi Peluang Penyakit Serangan Jantung," *Indones. J. Bus. Intell.*, vol. 5, no. 2, p. 87, 2022, doi: 10.21927/ijubi.v5i2.2659.
- [20] N. M. Farhan and B. Setiaji, "Indonesian Journal of Computer Science," *Indones. J. Comput. Sci.*, vol. 12, no. 2, pp. 284–301, 2023, [Online]. Available: <http://ijcs.stmikindonesia.ac.id/ijcs/index.php/ijcs/article/view/3135>