

Pengembangan Sistem Klasifikasi Makanan Sehat bagi Penderita Obesitas Menggunakan Machine Learning Support Vector Machine

Wildan Hayatul Fikri, Ratih Kumalasari, Danar Putra Pamungkas¹

E-mail: ¹kuliahwildan8@gmail.com, ²ratih.workmail@gmail.com, ³danar@unpkediri.ac.id

Penulis Korespondens : Wildan Hayatul Fikri

Abstrak

Obesitas merupakan salah satu masalah kesehatan yang terus meningkat dan dapat menyebabkan berbagai penyakit kronis seperti diabetes, hipertensi, dan penyakit jantung. Pemilihan makanan sehat menjadi langkah penting dalam menjaga pola hidup penderita obesitas. Penelitian ini bertujuan mengembangkan sistem klasifikasi makanan sehat menggunakan metode *Machine Learning Support Vector Machine* (SVM). Dataset yang digunakan berisi informasi kandungan gizi makanan seperti kalori, lemak, protein, karbohidrat, dan gula. Tahapan penelitian meliputi pengumpulan data, preprocessing, pembagian data training dan testing, pelatihan model SVM, serta evaluasi performa model. Hasil penelitian menunjukkan bahwa metode SVM mampu melakukan klasifikasi makanan sehat dengan tingkat akurasi yang baik dalam membedakan kategori makanan sehat dan tidak sehat bagi penderita obesitas. Sistem yang dikembangkan dapat membantu masyarakat dalam memilih makanan yang lebih sehat secara cepat dan efektif. **Kata Kunci:** klasifikasi, machine learning, makanan sehat, obesitas, *support vector machine*

Abstract

Obesity is one of the increasing health problems that can cause various chronic diseases such as diabetes, hypertension, and heart disease. Choosing healthy food is an important step in maintaining a healthy lifestyle for obese patients. This study aims to develop a healthy food classification system using the Support Vector Machine (SVM) machine learning method. The dataset used contains nutritional information such as calories, fat, protein, carbohydrates, and sugar. The research stages include data collection, preprocessing, training and testing data division, SVM model training, and model performance evaluation. The results show that the SVM method is able to classify healthy foods with good accuracy in distinguishing healthy and unhealthy food categories for obese patients. The developed system can assist people in selecting healthier foods quickly and effectively.

Keywords: *classification, healthy food, machine learning, obesity, support vector machine*

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

Obesitas merupakan kondisi kelebihan berat badan akibat penumpukan lemak tubuh yang berlebihan dan menjadi salah satu masalah kesehatan utama di dunia [1]. Pola makan yang tidak sehat, konsumsi makanan cepat saji, serta kurangnya aktivitas fisik menjadi faktor utama meningkatnya angka obesitas [2]. Kondisi ini dapat meningkatkan risiko berbagai penyakit seperti diabetes melitus, hipertensi, penyakit jantung, dan gangguan metabolisme lainnya [3].

Pemilihan makanan sehat menjadi salah satu langkah penting dalam mengontrol berat badan penderita obesitas [4]. Namun, masih banyak masyarakat yang kesulitan menentukan jenis makanan yang sesuai berdasarkan kandungan gizinya. Oleh karena itu, diperlukan suatu sistem yang mampu membantu pengguna dalam mengidentifikasi makanan sehat secara otomatis dan cepat.

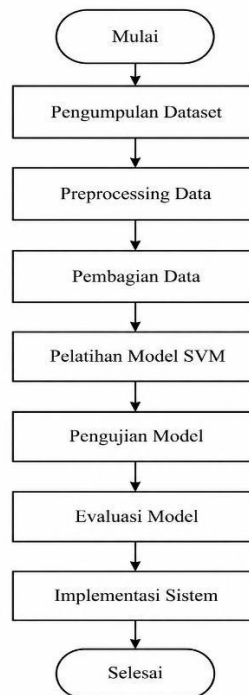
Perkembangan teknologi *Machine Learning* memungkinkan komputer melakukan klasifikasi data secara otomatis berdasarkan pola tertentu [5]. Salah satu algoritma yang sering digunakan dalam klasifikasi adalah *Support Vector Machine* (SVM) [6]. Algoritma SVM memiliki kemampuan yang baik dalam menangani data klasifikasi dengan tingkat akurasi tinggi serta mampu bekerja pada data multidimensi.

Penelitian ini bertujuan untuk mengembangkan sistem *klasifikasi* makanan sehat bagi penderita obesitas menggunakan metode *Support Vector Machine*. Sistem ini diharapkan dapat membantu pengguna dalam memilih makanan berdasarkan kandungan gizi sehingga dapat mendukung pola hidup sehat bagi penderita obesitas.

II. METODE

Penelitian ini menggunakan metode Machine Learning dengan algoritma *Support Vector Machine* (SVM) untuk mengklasifikasikan makanan sehat bagi penderita obesitas. Dataset yang digunakan terdiri dari 1.346 data makanan yang memiliki atribut nama makanan (*name*), kalori (*calories*), protein (*proteins*), lemak (*fat*), dan karbohidrat (*carbohydrate*). Dataset diperoleh dari Indonesian *Food and Drink Nutrition Dataset* yang tersedia pada platform *Kaggle* dan selanjutnya dilakukan proses pelabelan ke dalam tiga kategori kelas, yaitu Sangat Sehat, Sehat, dan Hindari sesuai kriteria penelitian. Dataset tersebut dapat diakses melalui platform *Kaggle* yang menyediakan informasi kandungan nutrisi makanan dan minuman Indonesia [7].

Tahapan penelitian meliputi pengumpulan data, *preprocessing* data, pembagian data training dan testing, pelatihan model SVM, serta *evaluasi* model. Alur penelitian ditunjukkan pada Gambar 1.

Gambar 1. *Flowchart* alur penelitian

2.1 Pengumpulan Data

Tahap pengumpulan data dilakukan dengan menggunakan dataset makanan yang berisi informasi kandungan gizi. Dataset terdiri dari 1.346 data makanan dengan atribut *calories*, *proteins*, *fat* dan *carbohydrate*. Data tersebut digunakan sebagai fitur utama dalam proses klasifikasi makanan sehat bagi penderita obesitas.

2.2 *Preprocessing* Data

Preprocessing dilakukan untuk meningkatkan kualitas data sebelum digunakan dalam proses pelatihan model. Tahapan *preprocessing* yang dilakukan meliputi:

a) Seleksi Fitur

Fitur yang digunakan dalam penelitian ini adalah *calories*, *proteins*, *fat* dan *carbohydrate*. *Variabel* name digunakan sebagai identitas makanan dan tidak digunakan dalam proses pelatihan model.

b) Pemeriksaan *Missing Valu*

Dataset diperiksa untuk memastikan tidak terdapat data kosong yang dapat memengaruhi proses *klasifikasi*.

c) Pemeriksaan Data Duplikat

Data duplikat dihapus agar tidak menyebabkan bias pada proses pembelajaran model.

d) Normalisasi Data

Seluruh fitur numerik dinormalisasi menggunakan metode *StandardScaler* agar memiliki skala yang seragam dan meningkatkan performa algoritma SVM.

e) Pelabelan Data

Data makanan dikelompokkan ke dalam tiga kategori kelas yaitu Sangat Sehat, Sehat, dan Hindari berdasarkan kandungan gizinya.

2.3 Pelatihan Model SVM

Setelah proses *preprocessing* selesai, dataset dibagi menjadi data training sebesar 80% dan data testing sebesar 20% menggunakan metode *train-test split*. Data *training* digunakan untuk membangun model klasifikasi menggunakan algoritma *Support Vector Machine* (SVM).

SVM bekerja dengan mencari *hyperplane* optimal yang mampu memisahkan data dari setiap kelas dengan margin terbesar [8]. Pada penelitian ini digunakan kernel yang memberikan performa terbaik berdasarkan hasil pengujian model.

Fungsi keputusan pada SVM dituliskan sebagai berikut:

$$f(x) = w \cdot x + b \dots\dots\dots(1)$$

dengan:

$f(x)$ = fungsi keputusan

w = bobot (*weight*) x =

vektor fitur b = bias

Model yang telah dilatih kemudian digunakan untuk melakukan prediksi terhadap data testing.

2.4 Evaluasi Model

Evaluasi model dilakukan menggunakan *confusion matrix* untuk mengetahui kemampuan model dalam melakukan klasifikasi [9]. Kinerja model diukur menggunakan beberapa metrik evaluasi yaitu *accuracy*, *precision*, *recall*, dan *F1-score*.

Nilai *accuracy* dihitung menggunakan persamaan berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \dots\dots\dots(2)$$

dengan:

TP (*True Positive*) : jumlah prediksi positif yang benar

TN (*True Negative*) : jumlah prediksi negatif yang benar

FP (*False Positive*) : jumlah prediksi positif yang salah

FN (*False Negative*) : jumlah prediksi negatif yang salah

Nilai *precision* dihitung menggunakan persamaan:

$$Precision = \frac{TP}{TP+FN} \dots\dots\dots(3)$$

Nilai *recall* dihitung menggunakan persamaan:

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots(4)$$

Nilai *F1-Score* dihitung menggunakan persamaan:

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \dots\dots\dots(5)$$

Precision digunakan untuk mengukur ketepatan prediksi model, *recall* digunakan untuk mengukur kemampuan model menemukan seluruh data pada kelas tertentu, sedangkan *F1score* digunakan untuk mengukur keseimbangan antara *precision* dan *recall* [10].

2.5 Variabel Penelitian

Tabel 1 menunjukkan *variabel* yang digunakan dalam proses *klasifikasi* makanan sehat.

Tabel 1. *Variabel* Dataset Penelitian

<i>Variabel</i>	<i>Deskripsi</i>	<i>Tipe Data</i>
<i>name</i>	Nama makanan	<i>String</i>
<i>calories</i>	Kandungan kalori makanan	<i>Numerik</i>
<i>proteins</i>	Kandungan protein makanan	<i>Numerik</i>
<i>fat</i>	Kandungan lemak makanan	<i>Numerik</i>
<i>carbohydrate</i>	Kandungan <i>karbohidrat</i> makanan	<i>Numerik</i>
<i>label</i>	Kategori makanan (Sangat Sehat, Sehat, Hindari)	Kategorikal

Tabel 2. Tahapan Penelitian

Tahap	<i>Aktivitas</i>	<i>Output</i>
1	Pengumpulan data	<i>Dataset</i> makanan
2	<i>Preprocessing</i> data	Data siap digunakan
3	Pembagian data <i>training</i> dan <i>testing</i>	<i>Dataset training</i> dan <i>testing</i>
4	Pelatihan model SVM	<i>Model klasifikasi</i>
5	Pengujian model	Hasil <i>prediksi</i>
6	<i>Evaluasi</i> model	<i>Accuracy, Precision, Recall, F1-Score</i>
7	<i>Implementasi</i> sistem	Sistem <i>klasifikasi</i> makanan sehat

III. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan model *klasifikasi* makanan sehat bagi penderita obesitas menggunakan algoritma *Support Vector Machine* (SVM). Model dibangun menggunakan dataset makanan yang telah melalui tahap *preprocessing* dan pembagian data

menjadi data training dan data testing. Pengujian dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan makanan ke dalam kategori Sangat Sehat, Sehat, dan Hindari.

3.1 Hasil Pelatihan Model

Setelah proses pelatihan menggunakan data *training*, model SVM berhasil membentuk *hyperplane* yang digunakan untuk memisahkan setiap kelas berdasarkan karakteristik kandungan gizi makanan. Fitur yang digunakan dalam proses klasifikasi meliputi *calories*, *proteins*, *fat* dan *carbohydrate*.

Model yang telah dilatih kemudian digunakan untuk melakukan prediksi terhadap data testing. Hasil *prediksi* menunjukkan bahwa model mampu mengenali pola data makanan berdasarkan kandungan gizinya.

3.2 Hasil Evaluasi Model

Kinerja model SVM dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* sesuai persamaan (2), (3), (4) dan (5). Hasil evaluasi model ditunjukkan pada Tabel 3.

Tabel 3. Hasil *Evaluasi Model SVM*

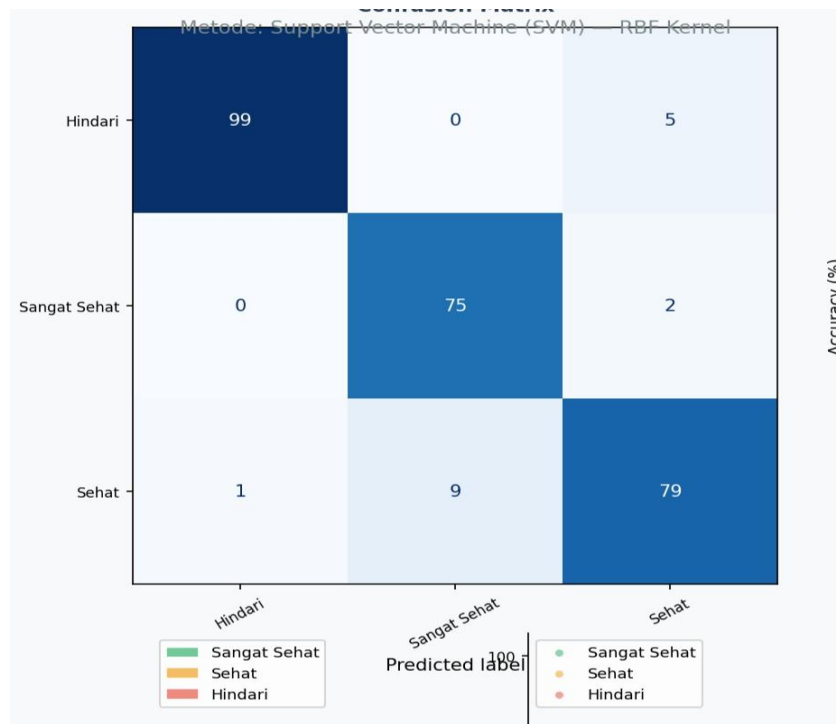
<i>Metrik</i>	<i>Nilai</i>
<i>Accuracy</i>	93.70%
<i>Precision</i>	93.88 %
<i>Recall</i>	93.70 %
<i>F1-Score</i>	93.72%

Berdasarkan Tabel 3, model SVM menunjukkan performa yang baik dalam melakukan klasifikasi makanan sehat. Nilai *accuracy* yang tinggi menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar oleh model. Selain itu, nilai *precision*, *recall*, dan *F1-score* yang baik menunjukkan bahwa model memiliki kemampuan yang seimbang dalam mengenali setiap kategori makanan.

Hasil tersebut menunjukkan bahwa model tidak hanya mampu menghasilkan prediksi yang akurat, tetapi juga memiliki kemampuan yang baik dalam mengidentifikasi setiap kategori makanan berdasarkan kandungan gizinya. Dengan demikian, algoritma SVM dapat digunakan sebagai metode klasifikasi yang efektif untuk mendukung pemilihan makanan sehat bagi penderita obesitas.

3.3 Confusion Matrix

Untuk mengetahui distribusi hasil klasifikasi secara lebih rinci digunakan *confusion matrix* seperti ditunjukkan pada Gambar 2.

Gambar 2. *Confusion Matrix* Model SVM

Berdasarkan *confusion matrix*, sebagian besar data berhasil *diprediksi* sesuai dengan kelas sebenarnya. Namun masih terdapat beberapa kesalahan klasifikasi yang terjadi pada data dengan karakteristik gizi yang mirip antar kelas. Hal ini menunjukkan bahwa kandungan gizi tertentu memiliki pola yang saling berdekatan sehingga memengaruhi kemampuan model dalam membedakan kategori makanan.

Meskipun masih terdapat beberapa kesalahan klasifikasi, jumlah prediksi yang benar jauh lebih banyak dibandingkan prediksi yang salah. Hal ini menunjukkan bahwa *hyperplane* yang dibentuk oleh SVM mampu memisahkan sebagian besar data makanan sesuai dengan kelas sebenarnya sehingga model memiliki kemampuan generalisasi yang baik terhadap data uji.

3.4 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma *Support Vector Machine* (SVM) mampu digunakan untuk mengklasifikasikan makanan sehat bagi penderita obesitas berdasarkan kandungan kalori, protein, lemak, dan karbohidrat. Model yang dibangun berhasil mengenali pola pada data nutrisi makanan sehingga dapat mengelompokkan makanan ke dalam kategori Sangat Sehat, Sehat, dan Hindari dengan tingkat *akurasi* sebesar 93,7%.

Kemampuan SVM dalam membentuk *hyperplane* optimal memungkinkan model memisahkan data dari setiap kelas secara efektif. Proses normalisasi data yang dilakukan pada tahap *preprocessing* juga berkontribusi dalam meningkatkan kinerja model karena seluruh fitur memiliki skala yang seragam sebelum dilakukan pelatihan.

Pada penelitian ini, setiap data makanan direpresentasikan sebagai vektor fitur yang terdiri dari nilai *calories*, *proteins*, *fat* dan *carbohydrate*. Algoritma SVM kemudian

membentuk *hyperplane* optimal untuk memisahkan data ke dalam kategori Sangat Sehat, Sehat, dan Hindari berdasarkan karakteristik kandungan gizinya. Proses normalisasi menggunakan *StandardScaler* membantu menyamakan skala antar fitur sehingga tidak terjadi dominasi atribut tertentu dalam proses pembentukan *hyperplane*. Dengan demikian, model mampu mempelajari pola data secara lebih optimal dan menghasilkan tingkat klasifikasi yang tinggi.

Keunggulan metode SVM pada penelitian ini terlihat dari kemampuannya menangani data multidimensi serta menghasilkan performa klasifikasi yang baik. Hasil penelitian ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa SVM merupakan salah satu algoritma klasifikasi yang efektif untuk data numerik karena mampu menemukan batas pemisah (*decision boundary*) yang optimal antar kelas data [11].

Karakteristik data nutrisi makanan yang terdiri dari beberapa atribut numerik menjadikan SVM cocok digunakan dalam penelitian ini. Setiap atribut memberikan kontribusi terhadap proses pemisahan kelas sehingga model dapat mengenali pola hubungan antar kandungan gizi makanan dengan kategori kesehatan yang telah ditentukan. Kemampuan tersebut menjadi salah satu faktor yang mendukung tercapainya nilai akurasi sebesar 93,7%.

Berdasarkan hasil *evaluasi*, model memperoleh nilai *accuracy* sebesar 93,7%, *precision* sebesar 93,88%, *recall* sebesar 93,70%, dan *F1-score* sebesar 93,72%. Nilai tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam melakukan klasifikasi serta mampu menjaga keseimbangan antara ketepatan prediksi dan kemampuan mengenali setiap kelas makanan.

Meskipun demikian, penelitian ini masih memiliki beberapa keterbatasan. Dataset yang digunakan hanya memanfaatkan empat fitur utama, yaitu kalori, protein, lemak, dan karbohidrat. Faktor nutrisi lainnya seperti kandungan gula, serat, vitamin, dan mineral belum digunakan dalam proses klasifikasi. Selain itu, penelitian ini hanya menggunakan algoritma SVM sehingga belum dapat dibandingkan secara langsung dengan algoritma klasifikasi lainnya seperti *Decision Tree*, *Random Forest*, *K-Nearest Neighbor* (KNN), maupun *Naïve Bayes*.

Hasil penelitian ini menunjukkan bahwa penerapan *Machine Learning*, khususnya algoritma *Support Vector Machine*, dapat dimanfaatkan sebagai sistem pendukung keputusan dalam pemilihan makanan sehat bagi penderita obesitas. Sistem yang dikembangkan mampu membantu pengguna memperoleh informasi mengenai kategori kesehatan makanan secara cepat dan objektif berdasarkan kandungan gizi yang dimiliki.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem klasifikasi makanan sehat bagi penderita obesitas menggunakan algoritma *Support Vector Machine* (SVM). Sistem dibangun dengan memanfaatkan data kandungan gizi makanan yang terdiri dari kalori, protein, lemak, dan karbohidrat sebagai fitur utama dalam proses *klasifikasi*. Melalui tahapan *preprocessing*, pembagian data, pelatihan model, dan *evaluasi*, model SVM mampu mengklasifikasikan makanan ke dalam kategori Sangat Sehat, Sehat, dan Hindari dengan performa yang baik.

Hasil pengujian menunjukkan bahwa algoritma SVM efektif dalam mengenali pola kandungan gizi makanan dan dapat digunakan sebagai metode klasifikasi untuk mendukung pemilihan makanan yang sesuai bagi penderita obesitas. Sistem yang dikembangkan dapat membantu pengguna dalam memperoleh informasi mengenai kategori kesehatan suatu makanan secara lebih cepat dan objektif berdasarkan nilai gizi yang dimiliki.

Penelitian selanjutnya dapat dilakukan dengan menambahkan jumlah dataset serta fitur gizi lainnya seperti kandungan gula, serat, vitamin, dan mineral untuk meningkatkan akurasi model. Selain itu, pengembangan sistem berbasis web atau mobile dapat dilakukan agar sistem lebih mudah diakses dan dimanfaatkan oleh masyarakat luas.

DAFTAR PUSTAKA

- [1] N. Andriani, A. Nurdin, U. Fitria, and K. A. Dinen, "Perilaku konsumsi makanan cepat saji pada remaja dan dampaknya bagi kesehatan," *Public Heal. J.*, vol. 1, no. 2, 2024.
- [2] S. R. N. Syarifah, A. Nurdin, U. Fitria, and K. A. Dinen, "RELEVANSI GIZI DAN KESEHATAN," *Public Heal. J.*, vol. 1, no. 2, 2024.
- [3] K. Fahmi *et al.*, "HUBUNGAN POLA MAKAN DAN AKTIFITAS FISIK TERHADAP KEJADIAN OBESITAS PADA REMAJA DI SMK NEGERI 4 MEULABOH," *J. Healthc. Technol. Med.*, vol. 11, no. 1, pp. 614–621, 2025.
- [4] A. S. Pratama, S. M. Sari, M. F. Hj, M. Badwi, and M. I. Anshori, "Pengaruh Artificial Intelligence, Big data dan otomatisasi terhadap kinerja SDM di Era digital," *J. Publ. Ilmu Manaj.*, vol. 2, no. 4, pp. 108–123, 2023.
- [5] R. O. Banjarnahor, F. F. Banurea, J. O. Panjaitan, R. S. P. Pasaribu, and I. Hafni, "Faktor-faktor risiko penyebab kelebihan berat badan dan obesitas pada anak dan remaja: Studi literatur," *TROPHICO*, vol. 2, no. 1, pp. 35–45, 2022.
- [6] U. Arifin, and N. Sulistiyowati, "Penerapan algoritma support vector machine (SVM) dengan TF-IDF n-gram untuk text classification," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 6, no. 2, pp. 129–136, 2021.
- [7] A. F. Hanif, "Indonesian Food and Drink Nutrition Dataset," Kaggle. Accessed: Jun. 04, 2026. [Online]. Available: <https://www.kaggle.com/datasets/anasfikrihanif/indonesian-food-and-drinknutrition-dataset>
- [8] S. S. T. Anis Ervina *et al.*, *TRANSFORMASI GIZI DAN KESEHATAN: SOLUSI BERBASIS DATA UNTUK GENERASI EMAS*. Nuansa Fajar Cemerlang, 2024.

- [9] A. Ridwan, “Penerapan algoritma Naïve Bayes untuk klasifikasi penyakit diabetes mellitus,” *J. SISKOM-KB (Sistem Komput. dan Kecerdasan Buatan)*, vol. 4, no. 1, pp. 15–21, 2020.
- [10] B. D. S. Septiani *et al.*, *BUKU AJAR DIETETIK PENYAKIT TIDAK MENULAR*. Nuansa Fajar Cemerlang, 2025.
- [11] K. Arumsari *et al.*, *GIZI DALAM KESEHATAN REPRODUKSI*. PT Penerbit Qriset Indonesia, 2025.