

ObeCheck Sebagai Platform Penerapan Metode *K-Nearest Neighbors* untuk Klasifikasi Obesitas Berbasis Website

Lailatul Fariah*, Puji Subekti

Fakultas Teknologi dan Desain, Teknik Informatika, Institut Teknologi dan Bisnis Asia, Malang, Indonesia

Email: ^{1,*}lailatulfaridah278@gmail.com, ²puji.subekti@asia.ac.id

Email Penulis Korespondensi: lailatulfaridah278@gmail.com

Submitted: 14/01/2025; Accepted: 26/01/2025; Published: 01/03/2025

Abstrak—Peningkatan obesitas menjadi salah satu tantangan utama dalam bidang kesehatan, sehingga memerlukan solusi yang cepat dan efektif untuk klasifikasi dan diagnosis sejak dini. Penelitian ini bertujuan untuk mengembangkan sistem berbasis website dengan menggunakan metode *K-Nearest Neighbors* (KNN) untuk mengklasifikasikan obesitas berdasarkan data pengguna, sehingga dapat membantu masyarakat dalam mendeteksi obesitas sejak dini. Dataset yang digunakan dalam penelitian ini terdiri dari 2111 data dan 17 atribut yang mencakup berbagai faktor terkait obesitas, seperti berat badan, tinggi badan, usia, jenis kelamin, faktor genetik, serta gaya hidup, termasuk pola makan dan aktivitas fisik. Dataset ini diperoleh dari website *UCI Repository*. Data ini akan diolah menggunakan metode *K-Nearest Neighbors* (KNN) untuk menghasilkan model klasifikasi obesitas yang akurat dan relevan. Untuk mengevaluasi kinerja model *K-Nearest Neighbors* (KNN), dataset dibagi menjadi data pelatihan dan data pengujian dengan rasio 80:20, kemudian dievaluasi menggunakan *Confusion Matrix*, menghasilkan akurasi sebesar 89%. Karena model menunjukkan performa yang baik dalam mengklasifikasikan data uji, model ini selanjutnya dapat diimplementasikan sebagai sistem berbasis website untuk menguji data baru. Sistem ini akan menghasilkan hasil klasifikasi kategori berat badan, yaitu “Berat Badan Kurang”, “Berat Badan Normal”, “Berat Badan Berlebih”, dan “Obesitas”. Dengan demikian, masyarakat dapat dengan mudah dan akurat mengklasifikasikan obesitas menggunakan sistem ini.

Kata Kunci: Data Mining; Klasifikasi; K-Nearest Neighbors (KNN); Obesitas; Website

Abstract—The increase in obesity has become one of the major challenges in the healthcare sector, requiring quick and effective solutions for early classification and diagnosis. This study aims to develop a web-based system using the K-Nearest Neighbors (KNN) method to classify obesity based on user data, thereby assisting the public in early detection of obesity. The dataset used in this research comprises 2,111 records and 17 attributes, covering various factors related to obesity, such as weight, height, age, gender, genetic factors, and lifestyle, including dietary habits and physical activity. This dataset was obtained from the UCI Repository website. The data is processed using the K-Nearest Neighbors (KNN) method to generate an accurate and relevant obesity classification model. To evaluate the performance of the K-Nearest Neighbors (KNN) model, the dataset was split into training and testing data with a ratio of 80:20 and evaluated using a Confusion Matrix, resulting in an accuracy of 89%. Since the model demonstrates good performance in classifying test data, it can be implemented as a web-based system to test new data. This system will produce weight classification results, including categories such as "Underweight," "Normal Weight," "Overweight," and "Obesity." Thus, the public can easily and accurately classify obesity using this system.

Keywords: Data Mining; Classification; K-Nearest Neighbors (KNN); Obesity; Website

1. PENDAHULUAN

Kesehatan tubuh seseorang dapat dinilai baik apabila terdapat keseimbangan antara berat badan dan tinggi badan. Ketidakseimbangan ini dapat menyebabkan kondisi kelebihan berat badan atau obesitas [1]. Obesitas dapat diartikan sebagai suatu kelainan atau penyakit yang ditandai dengan penimbunan lemak tubuh secara berlebihan, akibat dari ketidakseimbangan antara asupan energi dengan energi yang digunakan [2]. Kondisi ini telah menjadi masalah kesehatan global yang semakin mengkhawatirkan, termasuk di Indonesia. Berdasarkan data Riset Kesehatan Dasar (Riskesdas) tahun 2018, prevalensi obesitas pada orang dewasa di Indonesia mencapai 21,8%. Angka tersebut terus meningkat setiap tahunnya, menggarisbawahi pentingnya upaya penanganan yang lebih efektif dalam mengatasi masalah obesitas.

Obesitas dapat memicu berbagai masalah kesehatan serius, seperti penyakit jantung, diabetes, hipertensi, dan bahkan kanker [3]. Penyakit-penyakit ini tidak hanya berdampak pada kesehatan individu, tetapi juga menimbulkan beban ekonomi yang signifikan karena membutuhkan perawatan medis yang intensif dan berkelanjutan [4]. Dari perspektif ekonomi, pencegahan obesitas jauh lebih efisien dibandingkan dengan biaya pengobatan jangka panjang. Salah satu langkah kunci dalam pencegahan adalah mengidentifikasi individu yang berisiko mengalami obesitas, sehingga tindakan pencegahan dapat diterapkan sedini mungkin untuk mengurangi risiko berkembangnya komplikasi kesehatan lebih lanjut.

Identifikasi obesitas umumnya dilakukan menggunakan Indeks Massa Tubuh (IMT), yang mengukur berat badan relatif terhadap tinggi badan [5]. Kelebihan IMT adalah kemudahan penggunaannya. Namun, metode konvensional ini memiliki keterbatasan karena tidak mempertimbangkan faktor-faktor lain yang turut mempengaruhi obesitas. Obesitas dipengaruhi oleh berbagai faktor, termasuk genetik, lingkungan, dan gaya hidup. Mengingat kompleksitas penyebab obesitas, masih banyak masyarakat yang belum sepenuhnya memahami risiko dan cara mengenali obesitas secara akurat, sehingga pencegahan dini menjadi sulit dilakukan. Oleh karena itu, diperlukan teknologi yang mampu mengidentifikasi obesitas dengan lebih komprehensif, untuk membantu masyarakat dalam upaya pencegahan sejak dini. Salah satu pendekatan yang dapat memfasilitasi identifikasi obesitas secara lebih efektif adalah melalui klasifikasi obesitas berbasis teknologi.

Salah satu teknologi yang dapat diterapkan untuk mengidentifikasi obesitas secara lebih efektif adalah melalui penerapan data mining. Data mining merupakan proses menemukan pola atau hubungan tersembunyi antar data yang sebelumnya tidak diketahui, dan menyajikannya dalam bentuk yang mudah dipahami untuk mendukung pengambilan keputusan [6]. Data mining mencakup beberapa metode, seperti estimasi, prediksi, klasifikasi, clustering, dan asosiasi [7]. Dalam penelitian ini, digunakan metode klasifikasi dengan metode *K-Nearest Neighbors* (KNN) yang merupakan sebuah metode untuk melakukan klasifikasi terhadap obyek baru berdasarkan (K) tetangga terdekatnya [8]. *K-Nearest Neighbors* (KNN) dikenal sebagai salah satu algoritma pembelajaran mesin yang efektif untuk tugas klasifikasi karena simple dan mudah interpretasinya serta tangguh terhadap data training yang memiliki banyak noise dan data dalam jumlah yang besar.

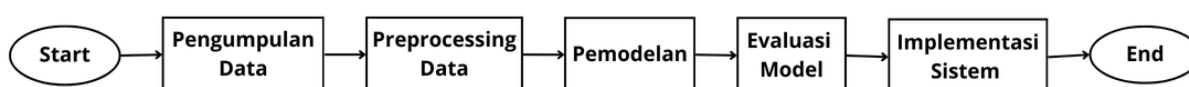
Metode KNN sudah digunakan untuk proses klasifikasi diberbagai obyek penelitian. Penelitian untuk klasifikasi materi pembelajaran menggunakan metode *K-Nearest Neighbors* (K-NN), yang menggunakan 60 materi pembelajaran sebagai dataset dengan label visual, aural, read/write & kinesthetic, dengan pembagian 48 data training dan 12 data testing mendapatkan akurasi sebesar 91% [9]. Penelitian tentang klasifikasi penerima bantuan program rehabilitasi rumah tidak layak huni dengan menggunakan metode *K-Nearest Neighbors* (KNN) menghasilkan akurasi sebesar 97,93% [10]. Klasifikasi dengan metode *K-Nearest Neighbors* (KNN) juga diterapkan untuk menentukan kepribadian siswa berdasarkan nilai raport, kuesioner *The Big Five Personality*, nilai pengetahuan, dan nilai keterampilan sebagai parameter penelitian. Pada penelitian tersebut, menghasilkan akurasi sebesar 95% [11]. Penelitian dalam komparasi metode *K-Nearest Neighbor* (KNN) dan *Naive Bayes* pada klasifikasi tingkat kualitas udara kota Tangerang Selatan, dapat disimpulkan bahwa metode *K-Nearest Neighbor* (KNN) memiliki akurasi lebih tinggi dibandingkan Naive Bayes dengan nilai akurasi 94.44% [12]. Penelitian yang menerapkan metode *K-Nearest Neighbor* (KNN) dalam klasifikasi penyakit jantung, hasil dari pengujian mendapatkan nilai akurasi sebesar 92% [13]. Dari penelitian-penelitian sebelumnya menunjukkan hasil yang baik dengan menerapkan metode *K-Nearest Neighbor* (KNN).

Meskipun sudah banyak penelitian dengan menggunakan metode *K-Nearest Neighbor* (KNN), masih terdapat keterbatasan pada penelitian sebelumnya, seperti penggunaan dataset yang relatif kecil, sehingga menghasilkan model yang kurang akurat. Sebagian besar penelitian sebelumnya cenderung berfokus pada hasil evaluasi model, tanpa mengimplementasikan model tersebut kedalam sebuah sistem. Sehingga model yang telah dikembangkan tidak dapat langsung dimanfaatkan dalam kehidupan sehari-hari. Oleh karena itu, penelitian ini bertujuan mengembangkan sistem berbasis website dengan menggunakan metode *K-Nearest Neighbors* (KNN) untuk mengklasifikasikan obesitas berdasarkan data pengguna. Sistem ini diharapkan dapat membantu masyarakat mendeteksi obesitas sejak dini. Dengan memanfaatkan dataset yang mencakup berbagai faktor terkait obesitas seperti jenis kelamin, usia, berat badan, tinggi badan, faktor genetik, serta gaya hidup seperti pola makan dan aktivitas fisik diharapkan klasifikasi obesitas dapat dilakukan dengan lebih akurat dan relevan. Dengan demikian, masyarakat dapat dengan mudah dan akurat mengklasifikasikan obesitas menggunakan sistem ini, sehingga dapat segera melakukan upaya pencegahan dan penanganan yang tepat. Pencegahan dan penanganan obesitas sejak dini dapat dilakukan melalui perbaikan gaya hidup, seperti mengadopsi pola makan yang sehat dan meningkatkan aktivitas fisik, sehingga tercipta keseimbangan antara berat badan dan tinggi badan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini mengadopsi pendekatan kuantitatif dengan metode eksperimen untuk membangun dan menguji model klasifikasi obesitas dengan metode *K-Nearest Neighbors* (KNN). Pendekatan kuantitatif dipilih karena penelitian ini berorientasi pada analisis data numerik guna menghasilkan model yang dapat diukur dan dipercaya. Metode eksperimen ini dimanfaatkan untuk mendesain, melatih, serta mengevaluasi kinerja model KNN dalam melakukan klasifikasi data obesitas berdasarkan atribut-atribut yang relevan. Selain itu, penelitian ini juga mencakup integrasi model KNN ke dalam sebuah sistem berbasis website. Sistem ini dirancang untuk mempermudah akses dan penerapan, sehingga masyarakat dapat menggunakan teknologi ini secara mandiri untuk mendeteksi obesitas secara dini dengan akurat dan efisien. Hal ini diharapkan dapat meningkatkan kesadaran serta membantu upaya pencegahan terhadap masalah kesehatan yang berkaitan dengan obesitas. Penelitian ini dilaksanakan melalui serangkaian tahapan tertata dan terstruktur dan sistematis untuk memastikan bahwa setiap langkah yang diambil menghasilkan hasil yang akurat serta dapat diandalkan. Tahapan-tahapan ini dirancang dengan cermat untuk menjawab tujuan penelitian secara efektif, dimulai dari pengumpulan data, *preprocessing data*, pemodelan, evaluasi model, hingga implementasi sistem. Tahapan penelitian tersebut ditunjukkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Gambar 1 menampilkan diagram alur penelitian yang terdiri dari lima langkah utama untuk mengembangkan sistem klasifikasi obesitas berbasis website. Tahap awal adalah pengumpulan data, di mana dataset obesitas dikumpulkan dari sumber terpercaya untuk menjamin relevansi dan kualitas. Selanjutnya, data diproses melalui tahap pra-pemrosesan, yang mencakup pembersihan, konversi data numerik, pembagian data menjadi pelatihan dan pengujian serta seleksi data input dan output. Setelah itu, dilakukan pemodelan menggunakan metode KNN untuk melatih model klasifikasi. Pada tahap berikutnya, evaluasi model dilakukan dengan menggunakan metrik seperti akurasi untuk mengukur kinerja model pada data pengujian. Langkah terakhir adalah implementasi sistem, yaitu integrasi model ke dalam platform berbasis website yang dirancang untuk memberikan aksesibilitas dan kemudahan kepada masyarakat dalam mendeteksi obesitas secara mandiri dan akurat.

2.2 Pengumpulan Data

Dataset yang digunakan pada penelitian ini diambil dari website *UCI Repository*. Dataset ini berisi tingkat obesitas pada masyarakat dari negara Meksiko, Peru dan Kolombia, dengan usia antara 14 dan 61 tahun serta kebiasaan makan dan kondisi fisik yang beragam. Dataset ini mencakup berbagai faktor terkait obesitas, seperti berat badan, tinggi badan, usia, jenis kelamin, faktor genetik, serta gaya hidup, termasuk pola makan dan aktivitas fisik. Dataset ini berisi 17 atribut dan 2111 data. Berikut atribut, indikator data serta tipe data dari masing masing atribut pada dataset yang disajikan dalam Tabel 1.

Tabel 1. Kriteria Atribut

No.	Atribut	Indikator Data	Tipe Data
1.	Jenis Kelamin	a. Wanita b. Pria	Nominal
2.	Usia	Nilai numerik (14 – 61)	Numerik
3.	Tinggi Badan	Nilai numerik dalam meter	Numerik
4.	Berat Badan	Nilai numerik dalam kilogram	Numerik
5.	Genetik Obesitas	a. Ya b. Tidak	Nominal
6.	Kalori Tinggi	a. Ya c. Tidak	Nominal
7.	Sayuran	a. Tidak Pernah = 1 b. Kadang – Kadang = 2 a. Selalu = 3	Numerik
8.	Makanan Utama	a. Antara 1/2 = 1 b. Tiga = 2 c. Lebih dari 3 = 3	Numerik
9.	Camilan	a. Tidak b. Kadang - Kadang c. Sering d. Selalu	Nominal
10.	Merokok	a. Ya b. Tidak	Nominal
11.	Air Putih	a. Kurang dari 1 liter = 1 b. Antara 1 dan 2 liter = 2 c. Lebih dari 2 liter = 3	Numerik
12.	Pemantauan Kalori	a. Ya = 1 b. Tidak = 2	Numerik
13.	Aktivitas Fisik	a. Tidak Pernah = 1 b. 1 atau 2 hari = 2 c. 2 atau 4 hari = 3 d. 4 atau 5 hari = 4	Numerik
14.	Perangkat Teknologi	a. 0-2 jam = 1 b. 3-5 jam = 2 c. Lebih dari 5 jam = 3	Numerik
15.	Alkohol	a. Tidak minum b. Kadang – kadang c. Sering d. Selalu	Nominal
16.	Transportasi	a. Mobil b. Sepeda Motor c. Sepeda d. Transportasi Umum e. Jalan Kaki	Nominal

17. Kategori	a. Berat Badan Kurang	Nominal
	b. Berat Badan Normal	
	c. Berat Badan Berlebih	
	d. Obesitas	

2.3 Preprocessing Data

Preprocessing data dilakukan untuk memastikan kualitas data baik sebelum digunakan waktu menganalisis data [14]. Pada tahap ini akan dilakukan sejumlah proses untuk persiapan data agar data siap diolah sesuai kebutuhan. Berikut dijelaskan mengenai proses-proses yang akan dilakukan pada tahap *preprocessing data*:

a. *Data Selection* (Seleksi Data)

Seleksi data (*data selection*) merupakan data yang tidak sesuai tidak diambil sedangkan data yang sesuai untuk dianalisis yang akan diambil [15]. Pada penelitian ini semua atribut akan digunakan, yang akan dibagi menjadi dua jenis atribut yaitu atribut untuk masukan (*input*) dan atribut untuk keluaran (*output*) atau label. Pada dataset ini atribut *input* yaitu Jenis Kelamin, Usia, Tinggi Badan, Berat Badan, Genetik Obesitas, Kalori Tinggi, Sayuran, Makanan Utama, Camilan, Merokok, Air Putih, Pemantauan Kalori, Aktivitas Fisik, Perangkat Teknologi, Alkohol dan Transportasi. Sedangkan atribut *output* atau label yaitu Kategori.

b. *Data Cleaning* (Pembersihan Data)

Pembersihan data merupakan proses menghilangkan noise dan data yang tidak relevan [16]. Pada tahap *data cleaning* melibatkan beberapa langkah yaitu penanganan nilai yang hilang (*missing values*) yaitu dengan menghapus atau mengisi *missing values* serta penanganan duplikasi yaitu dengan menghapus baris yang memiliki duplikasi. Pada dataset yang digunakan tidak ada nilai yang hilang (*missing values*) sehingga tidak perlu dilakukan penanganan lebih lanjut. Namun, terdapat baris duplikat sebanyak 27 data sehingga perlu dilakukan penghapusan baris yang duplikat, jadi dataset yang awalnya berisi 2111 data menjadi 2084 data.

c. *Encoding*

Encoding yaitu mengubah data non-numerik menjadi data numerik dengan tujuan supaya dapat diolah untuk kebutuhan penelitian ini dan dapat dipahami oleh algoritma pemodelan [17]. Ada 7 atribut yang perlu dilakukan proses *encoding* yaitu Jenis Kelamin, Genetik Obesitas, Kalori Tinggi, Camilan, Merokok, Alkohol, Transportasi karena masih berupa data non-numerik sehingga perlu diubah menjadi data numerik, dengan proses konversi data numerik seperti pada Tabel 2 berikut.

Tabel 2. Konversi Data Numerik

Nama Atribut	Konversi Data Numerik				
	1	2	3	4	5
Jenis Kelamin	Wanita	Pria			
Genetik Obesitas	Ya	Tidak			
Kalori Tinggi	Ya	Tidak			
Camilan	Ya	Tidak			
Merokok	Tidak Pernah	1 atau 2 hari	2 atau 4 hari	4 atau 5 hari	
Alkohol	Tidak Minum	Kadang - Kadang	Sering	Selalu	
Trasnportasi	Mobil	Sepeda Motor	Sepeda	Transportasi Umum	Jalan Kaki

d. *Split Data*

Split Data yaitu proses membagi dataset menjadi *data training* (data pelatihan) dan *data testing* (data pengujian), untuk melatih dan menguji model [18]. Pembagian data pelatihan dan data pengujian pada penelitian ini akan dibagi menjadi 80% untuk data pelatihan dan 20% untuk data pengujian.

2.4 Pemodelan

Pada tahap ini dilakukan proses pemodelan menggunakan metode *K-Nearest Neighbors* (KNN). *K-Nearest Neighbors* (KNN) adalah sebuah metode untuk melakukan klasifikasi terhadap objek baru berdasarkan (K) tetangga terdekatnya. Kelas yang paling banyak muncul, yang akan menjadi kelas hasil klasifikasi [19]. Berikut untuk alur perhitungan metode *K-Nearest Neighbors* (KNN) [20]:

a. Menentukan Nilai K

Pada penelitian ini akan menggunakan nilai K = 5.

b. Menghitung Jarak dengan Rumus *Euclidean Distance*

Salah satu cara untuk menghitung jarak dekat atau jauhnya tetangga menggunakan metode *Euclidean Distance* dengan rumus yang ditunjukkan pada persamaan 1 [19]:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{1}$$

Keterangan: $d(x, y)$ adalah jarak antara data x ke data y , x_i sebagai nilai x pada data training, y_i sebagai nilai y pada data testing, dan n sebagai batas jumlah banyaknya data.

c. Mengurutkan nilai data yang terkecil ke terbesar

d. Pemberian label berdasarkan data mayoritas dalam rentang nilai K

2.5 Evaluasi Model

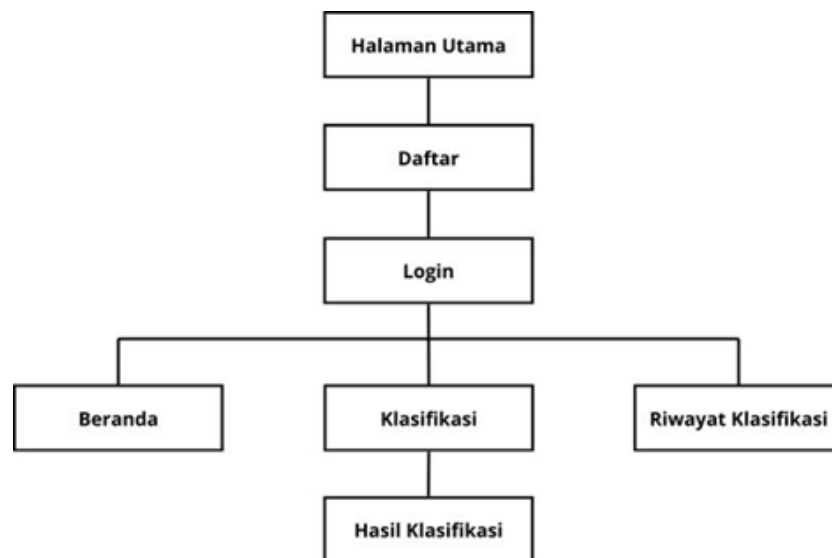
Metode pengujian pada penelitian ini menggunakan akurasi dari evaluasi *confusion matrix*. *Confusion matrix* merupakan pegujian yang dapat digunakan untuk menghitung kinerja atau tingkat kebenaran proses klasifikasi [21]. Untuk mengevaluasi suatu model klasifikasi dapat menggunakan akurasi. Akurasi adalah perbandingan jumlah prediksi yang benar dengan total prediksi [22]. Rumus akurasi ditunjukkan pada persamaan 2 berikut [22].

$$Akurasi = \frac{TP+TN}{TP+FP+TN+FN} \times 100\% \tag{2}$$

Keterangan: *True Positive (TP)* mengacu pada total data positif yang berhasil diklasifikasikan dengan benar oleh sistem, sementara *True Negative (TN)* adalah total data negatif yang juga terklasifikasi dengan benar oleh sistem. Sebaliknya, *False Negative (FN)* mencakup total data positif yang salah diklasifikasikan sebagai data negatif oleh sistem, sedangkan *False Positive (FP)* adalah total data negatif yang justru diklasifikasikan sebagai data positif.

2.6 Implementasi Sistem

Implementasi sistem merupakan tahap penerapan sekaligus pengujian sistem berdasarkan hasil analisis dan perancangan yang telah dilakukan [9]. Sistem ini diimplementasikan menggunakan bahasa pemrograman PHP dan database MySQL. Pada tahap ini, model yang telah dirancang akan diintegrasikan ke dalam sistem berbasis website, dengan struktur halaman yang ditampilkan pada Gambar 2.



Gambar 2. Site Map Website

3. HASIL DAN PEMBAHASAN

3.1 Hasil Evaluasi Model *K-Nearest Neighbors (KNN)*

Pada penelitian ini, telah dilakukan serangkaian eksperimen untuk mengevaluasi kinerja algoritma pembelajaran mesin dalam mengklasifikasikan obesitas. Dataset yang digunakan untuk membuat model *K-Nearest Neighbors (KNN)* merupakan data yang sudah melalui tahap *preprocessing data* yang meliputi pembersihan data serta *encoding* sehingga data siap diolah sesuai kebutuhan. Dataset yang sudah melalui tahap *preprocessing data* berisi 17 atribut dan 2084 data. Dataset mencakup informasi mengenai berbagai faktor terkait obesitas, seperti berat badan, tinggi badan, usia, jenis kelamin, faktor genetik, serta gaya hidup, termasuk pola makan dan aktivitas fisik, serta label yang menunjukkan kategori berat badan yang terdiri dari “Berat Badan Kurang”, “Berat Badan Normal”, “Berat Badan Berlebih”, dan “Obesitas”. Sampel dataset yang digunakan pada penelitian ini ditampilkan pada Tabel 3.

Tabel 3. Sampel Dataset

No.	Jenis Kelamin	Usia	Tinggi Badan	Berat Badan	Genetik Obesitas	Kalori Tinggi	Sayuran	Makanan Utama
1	1	21	1.520000	56.000000	1	2	3	3
2	2	27	1.800000	87.000000	2	2	3	3
...
2083	1	24	1.738836	133.472641	1	1	3	3
2084	2	19	1.800000	35.000000	2	2	2	1

Tabel Lanjutan 4. Sampel Dataset

Camilan	Merokok	Air Putih	Pemantauan Kalori	Aktivitas Fisik	Perangkat Teknologi	Alkohol	Transportasi	Kategori
2	1	3	1	4	1	2	4	berat badan normal
2	2	2	2	3	1	3	5	berat badan berlebih
...
2	2	3	2	2	2	2	4	obesitas
2	1	1	2	1	3	4	5	berat badan kurang

Model yang diuji yaitu *K-Nearest Neighbors* (KNN), pengujian dilakukan dengan rasio pembagian *data training* dan *data testing* sebesar 80% : 20% dengan nilai K = 5. Nilai K yang dipilih adalah ganjil karena jumlah kelas atau label dalam klasifikasinya adalah genap. Pemodelan serta evaluasi model menggunakan Bahasa pemrograman python di platform Google Colab. Hasil evaluasi kinerja model ditunjukkan pada Gambar 3.

	precision	recall	f1-score	support
berat badan berlebih	0.82	0.92	0.87	109
berat badan kurang	0.81	0.98	0.89	56
berat badan normal	0.90	0.46	0.61	61
obesitas	0.96	0.99	0.98	191
accuracy			0.89	417
macro avg	0.87	0.84	0.83	417
weighted avg	0.90	0.89	0.88	417

Akurasi KNN : 0.8920863309352518

Gambar 3. Hasil Evaluasi Model KNN

Gambar 3 menunjukkan hasil evaluasi kinerja model *K-Nearest Neighbors* (KNN) menunjukkan akurasi keseluruhan yang baik sebesar 89%. Nilai akurasi model KNN ditampilkan untuk memberikan informasi seberapa baik model tersebut dalam memprediksi data uji. Performa sangat baik pada kelas obesitas, dengan *precision*, *recall*, dan *f1-score* yang sangat tinggi. Namun, performa model kurang optimal pada kelas berat badan normal, yang memiliki *recall* rendah (0.46), menandakan bahwa model kesulitan dalam mengidentifikasi data dengan berat badan normal. Meskipun demikian, model ini memiliki keseimbangan yang baik pada sebagian besar kelas, dengan *weighted average precision*, *recall*, dan *f1-score* yang cukup tinggi, namun masih ada ruang untuk perbaikan dalam mengatasi kelas tertentu. Karena model memiliki performa yang cukup baik dalam mengklasifikasikan data uji, selanjutnya model dapat disimpan dan digunakan untuk menguji data baru tanpa perlu melatih ulang model tersebut.

3.2 Implementasi Model K-Nearest Neighbors (KNN) Dalam Website

Model *K-Nearest Neighbors* (KNN) yang telah dirancang dan dievaluasi, diintegrasikan ke dalam sistem website yang dikembangkan menggunakan bahasa pemrograman PHP dan database MySQL. Model KNN ini diterapkan pada website dengan tampilan antarmuka yang dirancang khusus untuk memberikan kemudahan bagi pengguna. Tampilan antarmuka mencakup beberapa menu utama, yaitu Beranda, Klasifikasi, dan Riwayat Klasifikasi. Saat pertama kali mengakses website, pengguna akan diarahkan ke Halaman Utama. Di sana, mereka diminta untuk membuat akun melalui Halaman Daftar dan login melalui Halaman Login sebelum dapat mengakses menu Beranda, Klasifikasi, dan Riwayat Klasifikasi. Berikut penjelasan dari masing-masing halaman:

a. Halaman Utama

Ketika pertama kali mengakses website, pengguna akan diarahkan ke Halaman Utama yang berisi judul, deskripsi singkat, dan dua tombol yaitu “Daftar” dan “Login”. Pengguna yang belum memiliki akun dapat mendaftar melalui tombol “Daftar”, sedangkan pengguna yang sudah memiliki akun dapat login melalui tombol “Login”. Tampilan Halaman Utama ditunjukkan pada Gambar 4 berikut.

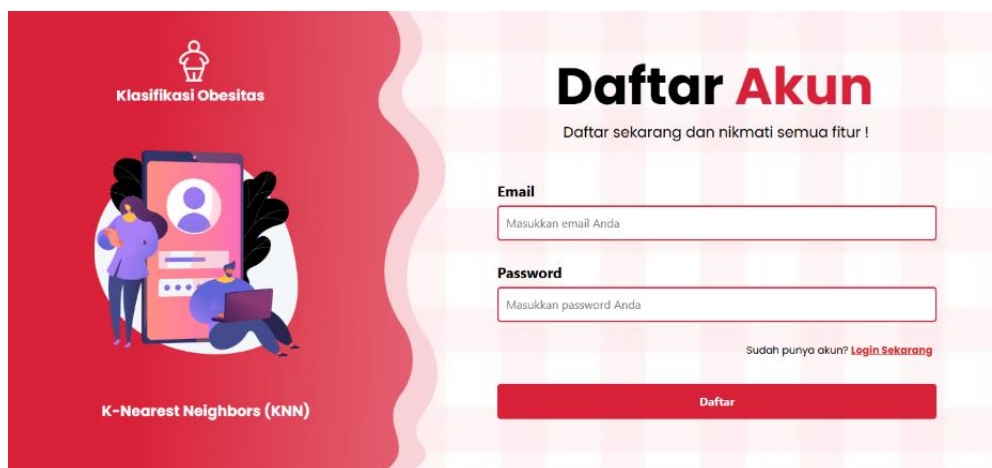


Gambar 4. Halaman Utama

Tampilan awal pada website ini berfungsi untuk memperkenalkan fitur utama, yaitu klasifikasi obesitas menggunakan metode *K-Nearest Neighbors* (KNN). Pada bagian tengah terdapat judul besar "Klasifikasi Obesitas" yang disertai deskripsi singkat tentang penggunaan metode KNN untuk menganalisis data obesitas. Sebagai pelengkap, ilustrasi di sebelah kanan menunjukkan seseorang yang berdiri di atas timbangan dengan pita pengukur, melambangkan proses pengukuran berat badan. Di sekitarnya, terdapat label seperti "BB Kurang", "BB Normal", "BB Lebih" dan "Obesitas" yang merepresentasikan kategori hasil klasifikasi. Untuk navigasi, disediakan dua tombol utama, yaitu "Daftar" untuk pengguna baru dan "Login" bagi pengguna yang telah memiliki akun. Dengan desain latar belakang berpola kotak-kotak berwarna pink dan elemen tombol berwarna merah, tampilan ini memberikan kesan ramah, sederhana, dan menarik perhatian, sesuai dengan tujuan aplikasinya yang berfokus pada kesehatan.

b. Halaman Daftar

Pada halaman Daftar, pengguna memasukkan email dan password, lalu klik tombol "Daftar". Jika email sudah terdaftar, akan muncul peringatan untuk mengganti email. Setelah data valid, pengguna akan diarahkan ke halaman Login. Tampilan halaman Daftar ditunjukkan pada Gambar 5 berikut.

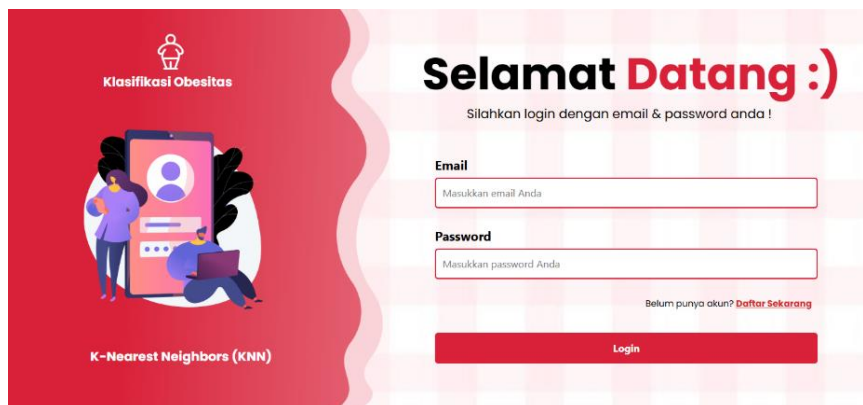


Gambar 5. Halaman Daftar

Tampilan ini merupakan halaman pendaftaran akun untuk website "Klasifikasi Obesitas". Pada bagian kiri, terdapat logo aplikasi, ilustrasi dua orang dengan perangkat ponsel, dan teks "*K-Nearest Neighbors* (KNN)" yang menggambarkan teknologi yang digunakan. Di bagian kanan, terdapat formulir pendaftaran dengan judul "Daftar Akun" yang diikuti subjudul ajakan, "Daftar sekarang dan nikmati semua fitur!". Formulir ini terdiri dari dua kolom input untuk memasukkan email dan password pengguna. Di bawahnya, terdapat tautan "Login sekarang" untuk pengguna yang sudah memiliki akun, serta tombol besar bertuliskan "Daftar" untuk melanjutkan proses pendaftaran. Dengan desain yang sederhana dan terorganisir, halaman ini mempermudah pengguna untuk memulai akses ke fitur-fitur website.

c. Halaman Login

Pada halaman Login, pengguna memasukkan email dan password yang terdaftar, lalu klik tombol "Login". Jika data tidak sesuai, akan muncul peringatan untuk mengulanginya. Setelah data valid, pengguna diarahkan ke halaman Beranda. Tampilan halaman Login ditunjukkan pada Gambar 6 berikut.



Gambar 6. Halaman Login

Tampilan ini merupakan halaman login untuk website "Klasifikasi Obesitas". Pada bagian kiri, terdapat logo aplikasi, ilustrasi dua orang dengan perangkat ponsel, dan teks "K-Nearest Neighbors (KNN)" yang menggambarkan teknologi yang digunakan. Di bagian kanan, terdapat formulir login dengan judul "Selamat Datang :)" yang diikuti subjudul ajakan, "Silahkan login dengan email & password Anda!". Formulir ini terdiri dari dua kolom input untuk memasukkan email dan password pengguna. Di bawahnya, terdapat tautan "Daftar Sekarang" untuk pengguna yang belum memiliki akun, serta tombol besar bertuliskan "Login" untuk melanjutkan proses masuk. Dengan desain yang sederhana dan modern, halaman ini mempermudah pengguna untuk mengakses website.

d. Halaman Beranda

Halaman Beranda adalah halaman utama setelah login, dengan menu navigasi ke "Beranda", "Klasifikasi", "Riwayat Klasifikasi", dan tombol "Log Out". Di tengah halaman terdapat judul, deskripsi singkat, dan tombol "Mulai Klasifikasi" untuk memulai proses klasifikasi. Tampilan halaman Beranda ditunjukkan pada Gambar 7 berikut.



Gambar 7. Halaman Beranda

Halaman utama aplikasi "Klasifikasi Obesitas" ini dirancang untuk mengelompokkan kategori obesitas menggunakan metode *K-Nearest Neighbors* (KNN). Bagian atas layar menampilkan menu navigasi seperti "Beranda," "Klasifikasi," dan "Riwayat Klasifikasi", dilengkapi dengan tombol "Log Out" di sudut kanan atas untuk keluar dari aplikasi. Judul halaman "Klasifikasi Obesitas" terlihat menonjol, disertai uraian singkat mengenai tujuan utama aplikasi. Pada area tengah terdapat ilustrasi seseorang yang menggunakan alat pengukur berat badan, menggambarkan tema obesitas secara visual. Selain itu, tersedia indikator kategori seperti "BB Kurang", "BB Normal", "BB Lebih" dan "Obesitas". Di bawah ilustrasi, terdapat tombol aksi berwarna merah bertuliskan "Mulai Klasifikasi" yang memudahkan pengguna untuk memulai proses klasifikasi. Latar belakang berpola kotak-kotak dengan dominasi warna lembut memberikan tampilan yang ramah dan menarik bagi pengguna.

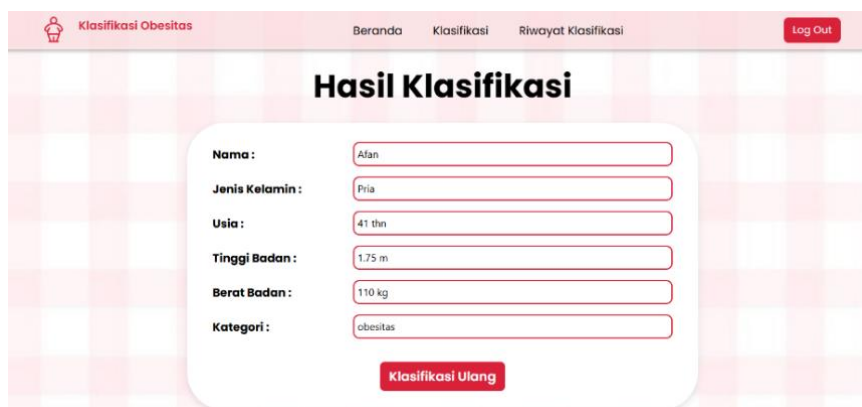
e. Halaman Klasifikasi dan Hasil Klasifikasi

Ketika pengguna mengklik tombol "Mulai Klasifikasi" atau menu "Klasifikasi", mereka akan diarahkan ke halaman Klasifikasi. Halaman ini berisi formulir untuk menentukan kategori berat badan pengguna. Tampilan halaman Klasifikasi terdapat pada Gambar 8, dan halaman Hasil Klasifikasi pada gambar 9. Pada halaman Hasil Klasifikasi, pengguna dapat melakukan klasifikasi ulang dengan tombol Klasifikasi Ulang, yang mengarahkan ke halaman Klasifikasi. Satu akun dapat melakukan klasifikasi berkali-kali.



Gambar 8. Halaman Klasifikasi

Halaman Klasifikasi menampilkan formulir input yang dirancang untuk mengumpulkan data pribadi pengguna, seperti nama, jenis kelamin, usia, tinggi badan, berat badan, faktor genetik, serta berbagai faktor terkait gaya hidup, termasuk pola konsumsi makanan dan aktivitas fisik. Semua informasi pada formulir ini wajib diisi, dan input usia, tinggi badan, serta berat badan harus berupa angka. Jika terdapat data yang tidak lengkap atau salah, sistem akan menampilkan peringatan dan proses klasifikasi tidak dapat dilanjutkan. Sebaliknya, jika data telah diisi dengan benar, tombol "Proses Klasifikasi" akan memproses data menggunakan metode K-Nearest Neighbors (KNN) untuk mengelompokkan pengguna ke dalam kategori berat badan, yaitu "Berat Badan Kurang," "Berat Badan Normal," "Berat Badan Berlebih," atau "Obesitas". Selain itu, halaman ini menyediakan navigasi menuju "Beranda," "Klasifikasi," dan "Riwayat Klasifikasi" dan tombol "Log Out", memberikan kemudahan bagi pengguna dalam mengakses fitur-fitur yang tersedia.



Gambar 9. Halaman Hasil Klasifikasi

Halaman ini menampilkan "Hasil Klasifikasi" dari website "Klasifikasi Obesitas". Pada bagian atas, tersedia menu navigasi seperti "Beranda," "Klasifikasi," dan "Riwayat Klasifikasi," serta tombol "Log Out" di sudut kanan atas untuk keluar dari aplikasi. Bagian tengah layar menampilkan formulir hasil klasifikasi yang memuat data pengguna yang telah diinputkan pada halaman Klasifikasi, termasuk Nama, Jenis Kelamin, Usia, Tinggi Badan, Berat Badan, dan Kategori. Setiap informasi ditampilkan di kolom yang tersedia, dengan hasil analisis berupa kategori berat badan, yaitu "Berat Badan Kurang," "Berat Badan Normal," "Berat Badan Berlebih," atau "Obesitas". Di bagian bawah formulir, terdapat tombol aksi berwarna merah bertuliskan "Klasifikasi Ulang" yang memungkinkan pengguna untuk memulai proses klasifikasi kembali dan akan diarahkan ke halaman Klasifikasi. Tampilan ini dirancang dengan latar belakang berpola kotak-kotak berwarna lembut, menciptakan suasana yang sederhana dan ramah bagi pengguna.

f. Halaman Riwayat Klasifikasi

Riwayat klasifikasi dapat dilihat di Halaman Riwayat Klasifikasi melalui menu navigasi. Pengguna dapat mengedit atau menghapus data riwayat. Tombol "Edit" mengarahkan ke halaman Klasifikasi untuk mengubah formulir, dan data akan otomatis diperbarui. Tombol "Hapus" akan menghapus data. Tampilan halaman Riwayat Klasifikasi ditunjukkan pada gambar 20.



Nama	Jenis Kelamin	Usia	Tinggi Badan	Berat Badan	Kategori	Aksi
Afan	Pria	41 thn	1.75 m	110 kg	obesitas	Edit Hapus
Zein	Pria	16 thn	1.84 m	45 kg	kekurangan berat badan	Edit Hapus

Gambar 10. Halaman Riwayat Klasifikasi

Halaman ini, menampilkan tabel yang berisi data riwayat hasil klasifikasi yang telah dilakukan oleh pengguna. Setiap baris tabel menampilkan informasi penting seperti nama pengguna, jenis kelamin, usia, tinggi badan, berat badan, dan hasil kategori klasifikasi (misalnya, "Berat Badan Kurang," "Berat Badan Normal," "Berat Badan Berlebih," atau "Obesitas"). Tabel ini juga dilengkapi dengan kolom Aksi, yang menyediakan tombol "Edit" untuk memperbarui data dan "Hapus" untuk menghapus entri riwayat tertentu. Navigasi di bagian atas memungkinkan pengguna berpindah ke halaman lainnya, seperti "Beranda" dan "Klasifikasi", serta tombol "Log Out" untuk keluar dari sistem. Desain tabel ini sederhana dan terorganisir untuk memudahkan pengguna dalam melihat atau mengelola riwayat klasifikasi mereka.

4. KESIMPULAN

Pada penelitian ini dilakukan pemodelan menggunakan metode *K-Nearest Neighbors* (KNN) untuk klasifikasi obesitas menggunakan 2111 data dan 17 atribut yang mencakup berbagai faktor terkait obesitas, seperti berat badan, tinggi badan, usia, jenis kelamin, faktor genetik, serta gaya hidup, termasuk pola makan dan aktivitas fisik. Tahapan yang dilakukan dalam penelitian ini, meliputi pengumpulan data, *preprocessing data*, pemodelan, evaluasi model dan implementasi sistem. Penerapan metode *K-Nearest Neighbors* (KNN) dalam klasifikasi obesitas bekerja dengan sangat baik ini dibuktikan dengan menggunakan nilai $K = 5$ dengan bobot data training sebanyak 80% dan 20% data testing mendapatkan *score accuracy* sebesar 89%. Performa sangat baik pada kategori obesitas, dengan *precision*, *recall*, dan *f1-score* yang sangat tinggi. Namun, performa model kurang optimal pada kategori berat badan normal, yang memiliki recall rendah (0.46), menandakan bahwa model kesulitan dalam mengidentifikasi data dengan kategori berat badan normal. Sistem yang dikembangkan dalam bentuk website dapat berfungsi dengan baik dan menghasilkan hasil klasifikasi kategori berat badan, yaitu "Berat Badan Kurang", "Berat Badan Normal", "Berat Badan Berlebih", dan "Obesitas". Dengan demikian, masyarakat dapat dengan mudah dan akurat mengklasifikasikan obesitas menggunakan sistem ini. Namun, pada penelitian ini sistem yang dikembangkan hanya berfokus untuk klasifikasi obesitas, tanpa memberikan rekomendasi penanganan atau pencegahan yang sesuai berdasarkan hasil klasifikasi pengguna.

REFERENCES

- [1] E. T. Lestari, Mustakim, dan J. Adhiva, "Implementasi Algoritma Naive Bayes Classifier dan K-Nearest Neighbor untuk Klasifikasi Status Gizi Obesitas Anak Disabilitas," *SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat*, vol. 1, no. 1, hlm. 1–11, 2022.
- [2] M. Heri, P. I. S. Dewi, N. M. D. Y. Astriani, P. A. Ariana, I. M. Sudayana, dan K. Y. Aryawan, *Obesitas Pada Anak Dalam Perspektif Orang Tua*. Pasuruan: Penerbit Qiara Media, 2022.
- [3] A. Ariyanto, T. Y. Fatmawati, F. Chandra, dan N. Efni, "Perilaku Mahasiswa dalam Pencegahan Obesitas," *Jurnal Akademika Baiturrahim Jambi*, vol. 12, no. 1, hlm. 201–206, Mar 2023, doi: 10.36565/jab.v12i1.696.
- [4] D. Hermawan dkk., *Mengenal Obesitas*. Yogyakarta: Penerbit Andi, 2020.
- [5] A. H. Santoso dkk., "Pengabdian Masyarakat - Pengukuran Indeks Massa Tubuh dan Lingkar Perut dalam Upaya Pemetaan Obesitas Sentral pada Warga Masyarakat di Desa Dalung, Serang, Banten," *SEWAGATI: Jurnal Pengabdian Masyarakat Indonesia*, vol. 2, no. 2, hlm. 01–08, Jun 2023, doi: 10.56910/sewagati.v2i2.596.
- [6] H. F. Putro, R. T. Vulandari, dan W. L. Y. Saptomo, "Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan," *Jurnal Teknologi Informasi dan Komunikasi (TIKomsin)*, vol. 8, no. 2, hlm. 19–24, Okt 2020, doi: 10.30646/tikomsin.v8i2.500.
- [7] R. F. Putra dkk., *DATA MINING : Algoritma dan Penerapannya*. Jambi: PT. Sonpedia Publishing Indonesia, 2023.
- [8] I. L. Putra, "Implementasi Algoritma Particle Swarm Optimization(Pso) Dan K-Nearest Neighbor(K-Nn) Dalam Memprediksi Keberhasilan Anak Smk Mendapatkan Kerja," *Technologia : Jurnal Ilmiah*, vol. 13, no. 4, hlm. 339–350, Okt 2022, doi: 10.31602/tji.v13i4.8167.
- [9] A. Agung Nugraha dan U. Budiyanto, "Adaptive E-Learning System Berbasis Vark Learning Style dengan Klasifikasi Materi Pembelajaran Menggunakan K-NN (K-Nearest Neighbor)," *Technomedia Journal*, vol. 7, no. 2, hlm. 248–261, Sep 2022, doi: 10.33050/tmj.v7i2.1900.



- [10] A.-N. S. Na'iem, H. Mulyo, dan N. A. Widiastuti, "Klasifikasi penerima bantuan program rehabilitasi rumah tidak layak huni menggunakan algoritme K-Nearest Neighbor," *Jurnal Teknologi dan Sistem Komputer*, vol. 10, no. 1, hlm. 32–37, Jan 2022, doi: 10.14710/jtsiskom.2021.14110.
- [11] M. T. Silaen, "Klasifikasi Karakteristik Kepribadian Siswa Berdasarkan The Big Five Personality Dengan Menggunakan Metode K- Nearest Neighbor (Knn)," *Jurnal Informatika dan Rekayasa Elektronik*, vol. 6, no. 1, hlm. 121–129, Apr 2023, doi: 10.36595/jire.v6i1.860.
- [12] A. Budianita, N. Iman, F. M. Hana, dan C. B. Hakim, "Komparasi Algoritma K-Nearest Neighbor dan Naive Bayes pada Klasifikasi Tingkat Kualitas Udara Kota Tangerang Selatan," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, hlm. 320–327, 2024, doi: <http://dx.doi.org/10.36499/jinrpl.v6i1.10956>.
- [13] M. F. A. Sularno, W. Wiyanto, D. Ardiatma, dan A. T. Zy, "Penerapan Algoritma K-Nearest Neighbor Dalam Klasifikasi Penyakit Jantung," *Journal of Computer System and Informatics (JoSYC)*, vol. 4, no. 4, hlm. 850–860, Agu 2023, doi: 10.47065/josyc.v4i4.4071.
- [14] D. Utami dan P. A. R. Devi, "Klasifikasi Kelayakan Penerima Bantuan Program Keluarga Harapan (Pkh) Menggunakan Metode Weighted Naive Bayes Dengan Laplace Smoothing," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 7, no. 4, hlm. 1373–1384, Des 2022, doi: 10.29100/jupi.v7i4.3592.
- [15] L. Iryani, "Penerapan Datamining Menentukan Minat Baca Mahasiswa Di Perpustakaan Universitas Bina Darma Palembang Menggunakan Metode Clustering," *INTECOMS: Journal of Information Technology and Computer Science*, vol. 3, no. 1, hlm. 82–89, Jun 2020, doi: 10.31539/intecom.v3i1.1251.
- [16] N. Khasanah, A. Salim, N. Afni, R. Komarudin, dan Y. I. Maulana, "PREDIKSI KELULUSAN MAHASISWA DENGAN METODE NAIVE BAYES," *Technologia: Jurnal Ilmiah*, vol. 13, no. 3, hlm. 207–214, Agu 2022, doi: 10.31602/tji.v13i3.7312.
- [17] C. Herdian, A. Kamila, dan I. G. Agung Musa Budidarma, "Studi Kasus Feature Engineering Untuk Data Teks: Perbandingan Label Encoding dan One-Hot Encoding Pada Metode Linear Regresi," *Technologia: Jurnal Ilmiah*, vol. 15, no. 1, hlm. 93–108, Jan 2024, doi: 10.31602/tji.v15i1.13457.
- [18] W. Astuti, R. Kurniawan, dan Y. A. Wijaya, "Analisis Data Sentimen Ulasan Aplikasi Dana di Google Play Store Menggunakan Algoritma Naive Bayes," *Jurnal Informatika dan Rekayasa Perangkat Lunak*, vol. 6, no. 1, hlm. 158–163, Jan 2024, doi: <http://dx.doi.org/10.36499/jinrpl.v6i1.10272>.
- [19] A. A. Putri, "RESOLUSI: Rekayasa Teknik Informatika dan Informasi Penerapan Data Mining Untuk Memprediksi Penjualan Buah Dan Sayur Menggunakan Metode K-Nearest Neighbor (Studi Kasus: PT. Central Brastagi Utama)," *RESOLUSI: Rekayasa Teknik Informatika dan Informasi*, vol. 1, no. 6, hlm. 354–361, 2021, doi: <https://doi.org/10.30865/resolusi.v1i6.187>.
- [20] S. Y. Riska dan L. Farokhah, "Klasifikasi Bumbu Dapur Indonesia Menggunakan Metode K-Nearest Neighbors (K-NN)," *SMATIKA JURNAL*, vol. 11, no. 01, hlm. 37–42, Jun 2021, doi: 10.32664/smatika.v11i01.568.
- [21] S. Proboningrum dan Acihmah Sidauruk, "SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN SUPPLIER KAIN DENGAN METODE MOORA," *JSiI (Jurnal Sistem Informasi)*, vol. 8, no. 1, hlm. 43–48, Mar 2021, doi: 10.30656/jsii.v8i1.3073.
- [22] V. A. D. Hidayatullah, A. Nilogiri, dan H. A. Al Faruq, "Klasifikasi Siswa Berprestasi Menggunakan Metode K-Nearest Neighbor (KNN) Pada SMA Negeri 2 Situbondo Classification Of Achieving Students Using K-Nearest Neighbor (KNN) Method At SMA Negeri 2 Situbondo," *Jurnal Smart Teknologi*, vol. 1, no. 1, hlm. 100–102, 2021.