IMPLEMENTASI ALGORITMA XGBOOST dan SVM UNTUK MEMPREDIKSI TINGKAT OBESITAS PADA KALANGAN ORANG DEWASA

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika

disusun oleh

Syandi Aulia Aqila

20.11.3545

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2025

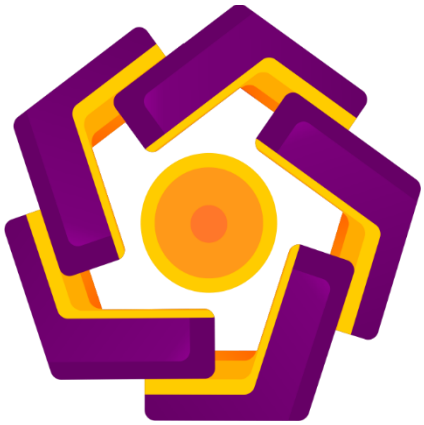
IMPLEMENTASI ALGORITMA XGBOOST dan SVM UNTUK MEMPREDIKSI TINGKAT OBESITAS PADA KALANGAN ORANG DEWASA

HALAMAN JUDUL

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana

Program Studi Informatika



disusun oleh

Syandi Aulia Aqila

20.11.3545

Kepada

FAKULTAS ILMU KOMPUTER

UNIVERSITAS AMIKOM YOGYAKARTA

YOGYAKARTA

2025

HALAMAN PERSETUJUAN

SKRIPSI

IMPLEMENTASI ALGORITMA XGBOOST dan SVM UNTUK MEMPREDIKSI TINGKAT OBESITAS PADA KALANGAN ORANG DEWASA

yang disusun dan diajukan oleh

Syandi Aulia Aqila

20.11.3545

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 23 Agustus 2024

Dosen Pembimbing,

Kusnawi, S.Kom., M.Eng

NIK: 19030202112

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

SKRIPSI

IMPLEMENTASI ALGORITMA XGBOOST dan SVM UNTUK MEMPREDIKSI TINGKAT OBESITAS PADA KALANGAN ORANG DEWASA

yang disusun dan diajukan oleh

Syandi Aulia Aqila

20.11.3545

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Agustus 2024

Susunan Dewan Penguji

Nama Penguji	Tanda Tangan
Rizqi Sukma Kharisma, M.Kom **NIK. 190302215**	
Ria Andriani, M.Kom **NIK. 190302458**	
Kusnawi, S.Kom, M. Eng. **NIK. 190302112**	

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Agustus 2024

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, M.Kom., Ph.D.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,
Nama mahasiswa : Syandi Aulia Aqila
NIM : 20.11.3545

Menyatakan bahwa Skripsi dengan judul berikut:

IMPLEMENTASI ALGORITMA XGBOOST dan SVM UNTUK MEMPREDIKSI TINGKAT OBESITAS PADA KALANGAN ORANG DEWASA

Dosen Pembimbing: Kusnawi, S.Kom., M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 23 Agustus 2024

Yang Menyatakan,

Syandi Aulia Aqila

HALAMAN PERSEMBAHAN

Alhamdulilah segala puji Syukur atas rahmat dan karunia Allah SWT yang telah memberikan petunjuk dan kemudahan bagi penulis dalam menyelesaikan skripsi ini dengan sebaik-baiknya. Dengan ini penulis mempersembahkan skrispsi ini kepada semua pihak yang telah membantu dengan memberikan motivasi serta dukungan yang besar sehingga skripsi ini terselesaikan dengan sebaik-baiknya.

Melalui skripsi ini penulis mempersembahkan terima kasih kepada :

Dengan penuh rasa Syukur kepada Allah SWT yang telah memberikan rahmat dan karuniaNya sehingga penulis memiliki kemampuan untuk menyelesaikan skripsi ini.

1. Keluarga penulis Bapak Suriadi dan Ibu Syafaruni yang tak hentinya memberikan doa, cinta kasih dan dukungannya untuk penulis.
2. Teman-teman saya yang telah menemani selama masa perkuliahan yaitu: Ririn, Alan, Zalza, Rafli, Faris, Mardiah, Mugen, Ageng, Ulin, Tobit, Teddy, Adi, Natan, Tyas, Joko, Faiq, Ryan, Sindi, Topik, Dias, Ryo. Serta seluruh teman teman saya selama di universitas Amikom yang tidak dapat saya sebutkan satu persatu.
3. Sahabat saya di grup “Warsagafi” yaitu Rehan, Rafi, Gelegar, Wawan, yang memberi dukungan dan menjadi tempat berkeluh kesah penulis.
4. Terimakasih untuk anak anak”lapas rozez” karena telah mau menampung dan bermain bersama saya ketika saya tidak memiliki teman kelas.
5. Terakhir, untuk diri sendiri terima kasih telah berjuang dalam menyelesaikan skripsi.

KATA PENGANTAR

Alhamdulillah rabbil ’alamin, segala puji dan Syukur penulis panjatkan atas kehadirat Allah subhanahu wa ta ‘ala segala rahmat dan karunia-Nya sehingga penulis mampu menyelesaikan skripsi dengan judul “IMPLEMENTASI ALGORITMA XGBOOST dan SVM UNTUK MEMPREDIKSI TINGKAT OBESITAS PADA KALANGAN ORANG DEWASA” dengan tepat waktu. Skripsi ini ditulis untuk memenuhi syarat kelulusan untuk mencapai Gelar Sarjana pada Fakukltas Ilmu Komputer Program Studi Informatika Universitas Amikom Yogyakarta.

Dalam menyelesaikan penulisan skripsi ini, penulis banyak memperoleh bantuan baik, pengajaran dan arahan dari berbagai pihak secara langsung maupun tidak langsung. Dengan ini penulis menyampaikan terimakasih kepada:

1. Bapak prof. Dr. M. Suyanto, M.M. selaku rector Universitas Amikom Yogyakarta.
2. Bapak Hanif Al fatta, S.Kom.,M.Kom.,Ph.D. selaku Dekan fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Windha Mega Pradnya D, M. Kom.selaku ketua program Studi S1 Informatika Universitas Amikom Yogyakarta.
4. Bapak Kusnawi, S. Kom., M. Eng selakun dosen pembimbing saya yang telah memberikan waktu, saran, kritik,motivasi dan bimbingan dalam skripsi ini sehingga dapat menyelesaikan skripsi selesai dengan baik.
5. Bapak Suriadi dan Ibu Syafaruni selaku orang tua penulis yang memberikan dukungan, doa dan motivasi kepada penulis.

Penulis menyadari bahwa skripsi ini masih banyak kekurangan karena keterbatasan pengetahuan yang dimiliki oleh penulis. Oleh karena, itu kritik dan saran sangat diharapkan bagi perbaikan penelitian selanjutnya.

Yogyakarta, 23 Agustus 2024

Syandi Aulia Aqila

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PERSETUJUAN ii
HALAMAN PENGESAHAN iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI iv
HALAMAN PERSEMBAHAN v
KATA PENGANTAR vi
DAFTAR ISI vii
DAFTAR TABEL ix
DAFTAR GAMBAR x
DAFTAR ISTILAH xi
INTISARI xii
ABSTRACT xiii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 2
1.3 Batasan Masalah 3
1.4 Tujuan Penelitian 3
1.5 Manfaat Penelitian 3
1.6 Sistematika Penulisan 4
BAB II TINJAUAN PUSTAKA 7
2.1 Studi Literatur 7
2.2 Dasar Teori 14
BAB III METODE PENELITIAN 24

3.1 Alur Penelitian 24
3.2 Pengumpulan data 25
3.3 Alat dan Bahan 27
BAB IV HASIL DAN PEMBAHASAN 28
4.1 Load Data 28
4.2 Exploratory Data Analayst 29
4.2.1 Pengubahan nama variabel 29
4.2.2 Transformasi Data 29
4.2.3 Pembulatan nilai kolom 30
4.2.4 Pemetaan nilai variabel 31
4.3 Preprocessing 36
4.3.1 Pembagian kategori variabel 36
4.3.2 Proses One Hot Encoder 36
4.4 Pembagian Data Latih dan Data Uji 39
4.5 Evaluasi 40
4.5.1 XGBoost 40
4.5.2 SVM 42
4.5.1 SVM rbf 42
4.5.2 SVM Linear 44
4.6 Perbandingan Model XGBoost dan SVM 45
BAB V PENUTUP 47
5.1 Kesimpulan 47
5.2 Saran 47
REFERENSI 48

DAFTAR TABEL

Tabel 2. 1 Tabel Keaslian.....10
Tabel 2. 2 Parameter Algoritma XGBoost.....19
Tabel 2. 3 Ilustrasi Confusion Matrix 22
Tabel 3. 1 Deskripsi Variabel pada Dataset 26
Tabel 4. 1 Dataset Penyakit Obesitas.....28
Tabel 4. 2 Contoh Perubahan Nama pada Variabel 29
Tabel 4. 3 Korelasi antar Variabel 33
Tabel 4. 4 Splitting Data 40

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....24
Gambar 4. 1 5 Data Teratas.....28
Gambar 4. 2 Transformasi Data.....30
Gambar 4. 3 Hasil Transformasi Data.....30
Gambar 4. 4 Kode Pembulatan Nilai.....31
Gambar 4. 5 Hasil Kode Pembulatan Nilai.....31
Gambar 4. 6 Mapping Nilai.....31
Gambar 4. 7 Penginputan Mapping Nilai.....32
Gambar 4. 8 Dataframe yang telah di Mapping Nilai.....32
Gambar 4. 9 Kategori Berat.....33
Gambar 4. 10 Perbandingan Kategori Berat Pria dan Wanita.....34
Gambar 4. 11 Histogram berat.....35
Gambar 4. 12 Kode Label Encoder.....36
Gambar 4. 13 Hasil Kode Label Encoder.....36
Gambar 4. 14 Kode Pembulatan nilai Unik.....37
Gambar 4. 15 Pembuatan Pipeline Onehot Encoder.....38
Gambar 4. 16 Hasil kode pembuatan Pipeline.....38
Gambar 4. 17 Pembuatan Coloumn Onehot Encoder.....39
Gambar 4. 18 Hasil Onehot Encoder.....39
Gambar 4. 19 Confusion Matrix XGBoost.....41
Gambar 4. 20 Model Score Algoritma XGBoost.....42
Gambar 4. 21 Confusion Matrix SVM rbf.....43
Gambar 4. 22 Model Score SVM rbf.....43
Gambar 4. 23 Confusion Matrix SVM Linear.....44
Gambar 4. 24 Model Score SVM Linear.....45

DAFTAR ISTILAH

Akurasi	tingkat ketepatan prediksi model dengan nilai sebenarnya
Data Training	sampel yang digunakan untuk melatih model
Data Testing	sampel yang digunakan untuk menguji model
Target	variabel yang ingin diprediksi oleh model
Algoritma	prosedur yang diterapkan oleh model untuk mempelajari pola dari data
Korelasi	ukuran hubungan statistic antar variabel
Dataset	kumpulan data yang digunakan untuk analisis
Kardiovaskular	penyakit yang disebabkan adanya gangguan pada jantung dan pembuluh darah
pravalensi	proporsi populasi yang memiliki karakteristik tertentu dalam periode waktu tertentu
osteoarthritis	peradangan kronis di sendi akibat kerusakan pada tulang rawan
obstructive sleep apnea	gangguan pernapasan yang terjadi saat tidur
BMI	cara menghitung berat badan ideal berdasarkan tinggi dan berat badan

INTISARI

Obesitas, yang ditentukan oleh Indeks Massa Tubuh (IMT), berdampak negatif terhadap kesehatan dan dipengaruhi oleh faktor-faktor seperti jenis kelamin, pendidikan, pekerjaan, status pernikahan, keturunan, aktivitas fisik, merokok, dan asupan kalori. Obesitas mempengaruhi semua kelompok usia, termasuk anak-anak, dan dapat menyebabkan masalah kesehatan serius seperti Sindrom Metabolik, penyakit kardiovaskular, dan diabetes. Penelitian ini menggunakan algoritma XGBoost dan dua SVM (RBF dan Linear) untuk memprediksi tingkat obesitas, yang dikategorikan menjadi tujuh tahap dari Berat Badan Kurang hingga Obesitas Tipe III. XGBoost, yang dikenal dengan akurasinya dan kemampuannya menangani dataset besar, dan SVM, yang diakui efektif dalam ruang berdimensi tinggi, dipilih untuk tugas ini.

Penelitian ini bertujuan untuk mengevaluasi kinerja algoritma tersebut dalam hal akurasi, sensitivitas, dan spesifisitas. Dataset yang mencakup berbagai faktor kesehatan dan gaya hidup diolah dan dibagi untuk pelatihan dan pengujian model. XGBoost dan SVM (RBF dan Linear) kemudian diterapkan untuk memprediksi tingkat obesitas. Akurasi yang dicapai oleh XGBoost adalah 98%, sedangkan SVM RBF dan SVM Linear masing-masing mencapai 87% dan 95% akurasi. Metrik kinerja dibandingkan untuk menentukan model yang lebih efektif. Temuan penelitian ini diharapkan dapat meningkatkan strategi kesehatan masyarakat dengan menyediakan alat yang lebih akurat untuk prediksi dan pencegahan obesitas.

Kata kunci: Prediksi Obesitas, XGBoost, Support Vector Machine (SVM), Pembelajaran Mesin, Kesehatan Masyarakat

ABSTRACT

Obesity, determined by Body Mass Index (BMI), negatively affects health and is influenced by factors like gender, education, employment, marital status, heredity, physical activity, smoking, and calorie intake. It impacts all age groups, including children, and can lead to severe health issues such as Metabolic Syndrome, cardiovascular disease, and diabetes. This study employs XGBoost and two SVM algorithms (RBF and Linear) to predict obesity levels, categorizing them into seven stages from Insufficient Weight to Obesity Type III. XGBoost, known for its accuracy and handling of large datasets, and SVM, recognized for its effectiveness in high-dimensional spaces, were chosen for this task.

The research aims to evaluate the performance of these algorithms in terms of accuracy, sensitivity, and specificity. The dataset, comprising various health and lifestyle factors, was preprocessed and split for training and testing the models. XGBoost and SVM (RBF and Linear) were then applied to predict obesity levels. The accuracy achieved by XGBoost was 98%, while SVM RBF and SVM Linear achieved 87% and 95% accuracy, respectively. The performance metrics were compared to determine the more effective model. This study's findings are expected to enhance public health strategies by providing more accurate tools for obesity prediction and prevention.

***Keyword*:** Obesity Prediction, XGBoost, Support Vector Machine (SVM), Machine Learning, Public Health