

**IMPLEMENTASI ALGORITMA NAÏVE BAYES
UNTUK MENGLASIFIKASI TINGKAT OBESITAS**



SKRIPSI

**diajukan untuk memenuhi salah satu syarat
memperoleh gelar sarjana komputer**

Oleh :

**IKA NURUL LAILIAH
2017.69.04.0073**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS YUDHARTA PASURUAN
2021**

**IMPLEMENTASI ALGORITMA NAÏVE BAYES
UNTUK MENGLASIFIKASI TINGKAT OBESITAS**



SKRIPSI

**diajukan untuk memenuhi salah satu syarat
memperoleh gelar sarjana komputer**

Oleh :

**IKA NURUL LAILIAH
2017.69.04.0073**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS YUDHARTA PASURUAN
2021**

PERNYATAAN PENULIS

JUDUL : PENERAPAN ALGORITMA NAÏVE BAYES
UNTUK MENGLASIFIKASI TINGKAT
OBESITAS
NAMA : IKA NURUL LAILIAH
NIM : 2017.69.04.0073

“Saya menyatakan dan bertanggungjawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”.

Pasuruan, 04 Agustus 2021



Ika Nurul Lailiah
Penulis

PERSETUJUAN SKRIPSI

JUDUL : IMPLEMENTASI ALGORITMA NAÏVE
BAYES UNTUK MENGLASIFIKASI
TINGKAT OBESITAS
NAMA : IKA NURUL LAILIAH
NIM : 2017.69.04.0073

Skripsi ini telah diperiksa dan disetujui
Pasuruan, 04 Agustus 2021



Pembimbing,

Arif Fauzin M.Kom
NIK. 0690707002

PENGESAHAN SKRIPSI

JUDUL : IMPLEMENTASI ALGORITMA NAÏVE BAYES
UNTUK MENGLASIFIKASI TINGKAT OBESITAS
NAMA : IKA NURUL LAILIAH
NIM : 2017.69.04.0073

Skripsi ini diujikan dan dipertahankan di depan Dewan Penguji
pada Sidang Skripsi 04 Agustus 2021. Menurut pandangan kami,

Skripsi ini memadai dari segi kualitas untuk tujuan
penganugerahan gelar Sarjana Komputer (S.Kom)
Pasuruan, 04 Agustus 2021

Pembimbing,

Arif Fakin M.Kom
NIP. Y 069107002

Penguji Utama,

Moch. Lutfi M.Kom
NIK.Y 0710048900

Kaprosdi,

M. Imron Roshadi M.Kom
NIK.Y 06907213121

Penguji Anggota,

M. Faishol Amrulloh M.Kom
NIK.Y 0691709007

Dekan Fakultas Teknik,

Misbach Munir ST.,MT
NIK.Y 0690201015

MOTTO

فَاسْتَبِقُوا الْخَيْرَاتِ

Berlomba-lomba dalam kebaikan

*Skripsi ini saya tunjukkan kepada
Orang Tua Tercinta,
Pamanku M. Imam Zubairi S.Pd
Adik, Keluarga & Saudara*

IMPLEMENTATION OF NAVE BAYES ALGORITHM TO CLASSIFY OBESITY LEVEL

Ika Nurul Lailiah

Informatics Engineering Study Program, Yudharta University Pasuruan

ABSTRACT

Patients with obesity in Indonesian adolescents continue to increase from year to year, due to improper diet and lifestyle. In obese people this must be considered. People with obesity can also be at higher risk for health problems, such as heart disease, stroke, diabetes, and several other diseases. Therefore, this study was conducted with the aim of analyzing the type of obesity disease using a classification algorithm in data mining techniques. The data studied were based on the results of public datasets obtained from UCI Machine Learning, namely Obesity data based on eating habits and physical conditions, with a total of 200 data with a ratio of 70% training data and 30% testing data. The criteria or attributes used to classify in determining the type of obesity include Gender, Age, Height, Weight, Family_history_With_Overweight, and NObeyesdad. Therefore, a data mining analysis system is needed using the Naïve Bayes Classifier (NBC) Algorithm method. The NBC algorithm is a method of probabilistic reasoning. The NBC algorithm aims to perform a data classification in certain classes. The Naïve Bayes method is one of the methods that can be applied in the classification system, where the system is very good at classifying based on previous events. The results of the Confusion Matrix test obtained an accuracy of 70% on the amount of data as much as 200.

Keywords: *Data Mining, Classification, Naïve Bayes Classifier (NBC), Obesity*

IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK MENGLASIFIKASI TINGKAN OBESITAS

Ika Nurul Lailiah

Program Studi Teknik Informatika, Universitas Yudharta Pasuruan

ABSTRAK

Penderita obesitas pada remaja Indonesia terus meningkat dari tahun ke tahun, dikarenakan pola makan dan gaya hidup yang tidak benar. Pada para penderita obesitas ini harus diperhatikan. Penderita obesitas juga bisa beresiko lebih tinggi untuk masalah kesehatan, seperti penyakit jantung, stroke, diabetes, dan beberapa jenis penyakit lainnya. Oleh karena itu, penelitian ini dilakukan dengan tujuan untuk menganalisa jenis penyakit obesitas menggunakan Algoritma klasifikasi pada teknik data mining. Data yang diteliti berdasarkan hasil dari dataset public yang diperoleh dari *UCI Machine Learning* yaitu data Obesitas Berdasarkan Kebiasaan Makan dan Kondisi Fisik, dengan jumlah data sebanyak 200 dengan perbandingan 70% *data training* dan 30% *data testing*. Kriteria atau atribut yang digunakan untuk mengklasifikasi dalam menentukan jenis penyakit obesitas diantaranya yaitu *Gender, Age, Height, Weight, Family_history_With_Overweight*, dan *NObesidad*. Oleh karena itu, diperlukan sistem analisa data mining dengan metode *Algoritma Naïve Bayes Classifier (NBC)*. *Algoritma NBC* merupakan salah satu metode pada *probabilistic reasoning*. *Algoritma NBC* bertujuan untuk melakukan suatu klasifikasi data pada kelas-kelas tertentu. Metode *Naïve bayes* adalah salah satu metode yang dapat diterapkan dalam sistem klasifikasi, dimana sistem tersebut sangat baik dalam melakukan klasifikasi berdasarkan kejadian sebelumnya. Hasil pengujian *Confusion Matrix* diperoleh hasil akurasi sebesar 70% pada jumlah data sebanyak 200.

Kata Kunci: *Data Mining, Klasifikasi, Naïve Bayes Classifier (NBC), Obesitas*

KATA PENGANTAR

Puji syukur saya panjatkan kepada Allah SWT, yang telah memberikan kesehatan, rahmat, hidayah serta inayahnya. Sehingga saya masih diberikan kesempatan untuk menyelesaikan skripsi ini, dengan judul **“IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK MENGLASIFIKASI TINGKAT OBESITAS”** sebagai salah satu syarat untuk mencapai gelar Sarjana Komputer Starata Satu pada Program Studi Teknik Informatika, Universitas Yudharta Pasuruan. Walaupun jauh dari kata sempurna, namun saya bangga telah mencapai pada titik ini, yang akhirnya skripsi ini bisa selesai diwaktu yang tepat. Dengan penuh rasa ikhlas dan rasa syukur kepada Allah SWT yang Maha Esa.

Penulis menyadari bahwa dalam penyusunan Skripsi ini bukanlah dari jerih payah sendiri, melainkan daribimbingan berbagai pihak. Oleh sebab itu penulis mengucapkan banyak terimakasih kepada semua pihak yang turut membantu dalam proses penulisan Skripsi ini, yaitu kepada:

1. Puji syukur kehadiran Allah SWT. Yang telah memberikan rahmat dan hidayahnya sehingga saya dapat menyelesaikan skripsi ini.
2. Nabi Muhammad SAW. Yang telah menjadi sumber inspirasi dalam berbagai macam hal dan tindakan dalam

hidup saya.

3. KH.Sholeh Bahrudin, selaku Pembina Yayasan Darut Taqwa yang selalu memberikan doa restunya.
4. Gus Muhammada dan Ning Siti Khurrotin, selaku pengasuh Asrama I yang selalu memberikan motivasi, dukungan dan do'a tanpa henti.
5. Kedua orangtua saya yang senantiasa mendoakan dan selalu tak henti-hentinya memberi dorongan semangat dan moral atau material demi tercapinya cita-cita saya.
6. Untuk Paman saya Muhammad Imam Zubairi yang selalu memberi kasih sayang, support, motivasi, dukungan, do'a dan cinta kasihnya.
7. Bapak Dr. H. Kholid Murtadlo, SE., M.E., selaku Rektor Universitas Yudharta Pasuruan.
8. Bapak Misbach Munir, ST., MT., selaku Dekan Fakultas Teknik Universitas Yudharta Pasuruan.
9. Muhammad Imron Rosyadi, S.Kom., M.Kom., selaku Ketua Prodi Teknik Informatika.
10. Arif Faizin S.Kom., M.Kom., selaku Dosen Pembimbing yang telah membimbing saya dalam menyelesaikan laporan skripsi ini.
11. Seluruh Dosen Program Studi Teknik Informatika Universitas Yudharta Pasuruan yang telah mengajarkan

dan memberikan ilmu yang bermanfaat dan berguna selama 8 semester ini untuk bekal di masa depan.

12. Teman-teman seperjuangan Teknik Informatika angkatan 2017, dan semua pihak yang telah membantu saya yang tidak dapat disebutkan satu persatu atas bantuan dan dorongan semangat dalam menyelesaikan skripsi ini.
13. Teman-temanku kamar i.7 yang selalu memberikan semangat, dorongan dan do'a atas terselesainya skripsi ini.
14. Almamaterku.

Pasuruan, 04 Agustus 2021

Ika Nurul Lailiah

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN KEASLIAN PENULIS.....	ii
PERSETUJUAN SKRIPSI.....	iv
PENGESAHAN SKRIPSI	v
ABSTRACT	viii
ABSTRAK.....	ix
KATA PENGANTAR.....	x
DAFTAR ISI.....	xiii
DAFTAR TABEL.....	xvi
DAFTAR GAMBAR.....	xvii
BAB 1 PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Batasan Penelitian	7
1.3 Rumusan Masalah	7
1.4 Tujuan Penelitian.....	7
1.5 Manfaat Penelitian.....	8
1.5.1 Manfaat Teoritis	8
1.5.2 Manfaat Praktisi.....	8

1.6 Sistematika Penulisan.....	9
BAB II TINJAUAN PUSTAKA	
2.1 Penelitian Terkait	12
2.2 Data Mining.....	30
2.3 Peran Utama Data Mining	31
2.4 <i>Knowledge Discovery in Database (KDD)</i>	32
2.5 Klasifikasi.....	34
2.6 Naïve Bayes Classifier (NBC).....	35
2.7 Obesitas	36
2.8 Sistem Informasi.....	38
2.9 Analisis Sistem	39
2.10 Bahasa Pemrograman	40
2.10.1 <i>Python</i>	40
2.11 <i>Flowchart</i>	41
BAB III METODE PENELITIAN	
3.1 Kerangka Pemikiran	43
3.2 Tahapan Perancangan.....	44
3.3 Metode Penelitian.....	45
3.4 Tahap Pengumpulan Data.....	47
3.5 Jenis Data	48
3.6 Teknik Analisa Data.....	49
3.7 Alat Pendukung Yang Dibutuhkan.....	50
3.7.1 Perangkat Lunak (<i>Software</i>).....	50

3.7.2 Perangkat Keras (<i>Hardware</i>).....	50
3.8 Tahap Klasifikasi.....	51
3.8.1 <i>Flowchart</i> Algoritma NBC.....	51
3.9 Tahapan Penelitian	53
BAB IV HASIL DAN PEMBAHASAN	
4.1 Analisis Sistem.....	54
4.2 Analisis Studi Kasus.....	54
4.3 Analisis Metode.....	55
4.4 Pengumpulan Data	56
4.5 Penentuan Atribut dan Kelas data	57
4.6 <i>Preprocessing</i> Data.....	57
4.7 <i>Cleaning</i> Data.....	57
4.8 <i>Transformasi</i> Data.....	59
4.9 Klasifikasi dengan Algoritma Naïve Bayes.....	60
4.10 Klasifikasi Algoritma <i>Naïve Bayes Classifier</i> (NBC)	61
4.10.1 Perhitungan Manual Algoritma NBC	61
4.10.2 Implementasi Algoritma <i>Naïve Bayes Classifier</i> (NBC) menggunakan <i>python</i>	67
BAB V PENUTUP	
5.1 Kesimpulan.....	79
5.2 Saran.....	79
DAFTAR PUSTAKA	
LAMPIRAN	

DAFTAR TABEL

Tabel 2.1 Tinjauan Pustaka Terkait Masalah Penelitian	23
Tabel 2.2 Simbol <i>Flowchart</i>	42
Tabel 4.1 Data Obesitas.....	56
Tabel 4.2 Hasil <i>Cleaning Data</i>	58
Tabel 4.3 <i>Transformasi Data</i>	59
Tabel 4.4 <i>Data Training</i>	60
Tabel 4.5 <i>Data Testing</i>	61
Tabel 4.6 Perhitungan <i>data testing</i>	62
Tabel 4.7 Nilai Probabilitas Tiap Kelas	63
Tabel 4.8 Nilai Mean.....	64
Tabel 4.9 Nilai Standar Deviasi	64
Tabel 4.10 Nilai Distribusi Normal	65
Tabel 4.11 Probabilitas dari Tiap-tiap Kelas.....	66
Tabel 4.12 Hasil Akhir Klasifikasi.....	66

DAFTAR GAMBAR

Gambar 2.1 Flowchart Penelitian Yang Diusulkan Wiwik Muslehatin. 2017	13
Gambar 2.2 Flowchart Penelitian Yang Diusulkan Hartatik. 2017	14
Gambar 2.3 Flowchart Penelitian Yang Diusulkan Wahyu Hadi Kristianto. 2019	15
Gambar 2.4 Flowchart Penelitian Yang Diusulkan Dewi Simanjuntak. 2019	16
Gambar 2.5 Flowchart Penelitian Yang Diusulkan Rhyzoma Grannata Rafsanjaji. 2018	17
Gambar 2.6 Flowchart Penelitian Yang Diusulkan Nadia Tiara Rahman. 2020	18
Gambar 2.7 Flowchart Penelitian Yang Diusulkan Muhammad Subaeri	19
Gambar 2.8 Flowchart Penelitian Yang Diusulkan Ayu Made Surya Dewi. 2020	20
Gambar 2.9 Flowchart Penelitian Yang Diusulkan Riffat Hossain. 2018	21
Gambar 2.10 Flowchart Penelitian Yang Diusulkan Sri Astuti Thamrin. 2018	23
Gambar 2.11 Data Mining dalam Proses <i>Knowlegde Discovery</i> .	34
Gambar 2.12 Python	40

Gambar 3.1 Kerangka Konsep Pemikiran	43
Gambar 3.2 Tahapan Data Mining	45
Gambar 3.3 <i>Flowchart NBC</i>	51
Gambar 3.5 Tahapan Penelitian	53
Gambar 4.1 Hasil <i>Precision, Recall, F1-Score</i> dan <i>Accuracy</i> dari Klasifikasi Naive Bayes Pada Pengujian-1	69
Gambar 4.2 <i>Output</i> Nilai Real dan Nilai Prediksi	70
Gambar 4.3 Hasil <i>Precision, Recall, F1-Score</i> dan <i>Accuracy</i> dari Klasifikasi Naive Bayes Pada Pengujian-1	71
Gambar 4.4 <i>Output</i> Nilai Real dan Nilai Prediksi	72
Gambar 4.5 Hasil <i>Precision, Recall, F1-Score</i> dan <i>Accuracy</i> dari Klasifikasi Naive Bayes Pada Pengujian-1	73
Gambar 4.6 <i>Output</i> Nilai Real dan Nilai Prediksi	74
Gambar 4.7 Hasil <i>Precision, Recall, F1-Score</i> dan <i>Accuracy</i> dari Klasifikasi Naive Bayes Pada Pengujian-1	75
Gambar 4.8 <i>Output</i> Nilai Real dan Nilai Prediksi	76
Gambar 4.9 Hasil <i>Precision, Recall, F1-Score</i> dan <i>Accuracy</i> dari Klasifikasi Naive Bayes Pada Pengujian-1	77
Gambar 4.10 <i>Output</i> Nilai Real dan Nilai Prediksi	78

