



Article DOI: 10.36350/jbs.v13i2.200
Received: May; Accepted: June; Published: July

Penerapan Algoritma *Naïve Bayes* Untuk Penentuan Diagnosa Obesitas Pada Peserta Sosialisasi Deteksi Dini Penyakit Tidak Menular (PTM)

Ida Maryani¹, Irmayansyah^{2*}

¹ Universitas Binaniaga Indonesia/Sistem Informasi
Email: ida.maryani85@gmail.com

² Universitas Binaniaga Indonesia/Sistem Informasi
Email: irma@unbin.ac.id

*) *Corresponding Author*

ABSTRACT

Obesity is excess fat accumulation due to an imbalance between energy intake and energy expenditure for a long time. Obesity can cause various non-communicable diseases, including heart disease, stroke, diabetes, high blood pressure, gallstones, respiratory problems, cancer, osteoarthritis, including infertility. The Regional Government through Posbindu Non-Communicable Diseases (PTM) in every UPTD Puskesmas carries out socialization as an early detection of PTM risk. This activity is carried out every two or three months, by conducting routine health checks for employees. Measurement of obesity diagnosis is usually by using the Body Mass Index (BMI). BMI is a measurement method by calculating the height in meters and weight in kilograms. However, BMI also has drawbacks, namely it cannot distinguish between muscle mass and fat mass. Apart from BMI, other factors also affect the diagnosis of obesity, including gender, age, abdominal circumference, risky behaviors such as lack of physical activity, eating patterns of excess sugar, excess salt, excess fat, not eating enough fruits and vegetables. With so many determinants of the diagnosis of obesity, the Naïve Bayes Algorithm is used to predict the diagnosis of obesity more effectively and accurately. The application built is in the form of a prototype that utilizes the PHP programming language. This study used data from the health examination of participants in the socialization of PTM risk early detection with a total of 60 participants. There are 10 variables used, namely Gender, Age, Height, Weight, Abdominal Circumference, Lack of Physical Activity, Eating Patterns of Excess Sugar, Excess Salt, Excess Fat and Undereating and Fruit while there are 2 classes, namely obesity and normal diagnoses. Based on the results of calculating accuracy with the Confusion Matrix, an accuracy value of 86.6% is obtained.

Keywords: *Obesity, Body Mass Index (BMI), Naïve Bayes Algorithm, Prediction*

ABSTRAK

Obesitas merupakan penumpukan lemak yang berlebih akibat ketidakseimbangan asupan energi (*energy intake*) dengan energi yang digunakan (*energy expenditure*) dalam waktu lama. Obesitas dapat menyebabkan berbagai penyakit tidak menular, diantaranya penyakit jantung, stroke, diabetes, tekanan darah tinggi, batu empedu, masalah pernapasan, kanker, osteoarthritis, termasuk infertilitas. Pemerintah Daerah melalui Posbindu Penyakit Tidak Menular (PTM) di setiap UPTD Puskesmas melaksanakan sosialisasi sebagai deteksi dini resiko PTM. Kegiatan ini dilakukan setiap dua atau tiga bulan sekali, dengan melakukan pemeriksaan kesehatan rutin bagi pegawai. Pengukuran diagnosa obesitas biasanya dengan menggunakan Indeks Massa Tubuh (IMT). IMT

adalah metode pengukuran melalui penghitungan antara tinggi badan dalam satuan meter dan berat badan dalam satuan kilogram. Akan tetapi, IMT juga memiliki kekurangan yaitu tidak dapat membedakan antara massa otot dan massa lemak. Selain IMT, faktor lain juga mempengaruhi diagnosa obesitas diantaranya Jenis Kelamin, Usia, Lingkar Perut, Perilaku Beresiko seperti Kurangnya Aktivitas Fisik, Pola Makan Gula Berlebih, Garam Berlebih, Lemak Berlebih, Kurang Makan Buah dan Sayur. Dengan banyaknya faktor penentu diagnosa obesitas maka digunakan Algoritma Naïve Bayes untuk prediksi penentuan diagnosa obesitas lebih efektif dan akurat. Aplikasi yang dibangun berupa prototype yang memanfaatkan Bahasa pemrograman PHP. Penelitian ini menggunakan data pemeriksaan kesehatan peserta sosialisasi deteksi dini resiko PTM dengan jumlah peserta 60 orang. Variabel yang digunakan ada 10 yaitu Jenis Kelamin, Usia, Tinggi Badan, Berat Badan, Lingkar Perut, Kurang Aktifitas Fisik, Pola Makan Gula Berlebih, Garam Berlebih, Lemak Berlebih dan Kurang Makan dan Buah sedangkan kelasnya ada 2 yaitu diagnosa obesitas dan normal. Berdasarkan hasil perhitungan akurasi dengan Confusion Matrix didapatkan nilai akurasi sebesar 86,6%.

Kata Kunci: Obesitas, Indeks Massa Tubuh (IMT), Algoritma *Naïve Bayes*, Prediksi.

A. PENDAHULUAN

1. Latar Belakang

Menurut Riset Kesehatan Dasar (Riskesdas) 2018 memperkirakan satu dari lima orang dewasa, satu dari lima anak berusia 5-12 tahun, dan satu dari tujuh remaja berusia 13-18 tahun di Indonesia mengalami obesitas. Angka prevalensi obesitas pada balita sebanyak 3,8 persen dan usia 18 tahun keatas sebanyak 21,8 persen. Jika kondisi ini terus dibiarkan, penelitian memprediksi angka obesitas dapat mencapai 40 persen pada tahun 2030. Kebijakan yang dirumuskan pemerintah pusat melalui Kementerian Kesehatan Republik Indonesia untuk mengatasi obesitas diantaranya dimulai pada tahun 2017 melaksanakan Gerakan Nusantara Tekan Angka Obesitas (GENTAS). Dimana bentuk kegiatan diturunkan kepada pemerintah daerah dengan melibatkan berbagai lintas program dan lintas sektor yaitu SKPD, Institusi Pendidikan, Lembaga Swadaya Masyarakat, organisasi profesi organisasi masyarakat dan lainnya. Bentuk kegiatannya antara lain: Sosialisasi deteksi dini penyakit tidak menular, Seminar dampak bahaya obesitas, penyelenggaraan Pos Pembinaan Terpadu (Posbindu) Penyakit Tidak Menular (PTM), penyelenggaraan Cek Kesehatan Secara Berkala (CERDIK) di sekolah, kantin sehat di institusi Pendidikan dan perkantoran.

Era teknologi informasi saat ini telah menyentuh berbagai bidang dan aspek kehidupan, termasuk bidang kesehatan. Teknologi informasi hari ini, sebagai pendukung untuk pengambilan keputusan medis, sangat banyak difokuskan pada pengembangan alat bantu yang secara otomatis melakukan diagnosis penyakit. Kelebihan sistem teknologi informasi kesehatan dibanding system konvensional terletak pada potensi untuk mengurangi diagnosis yang terlambat (delay), terlupakan (neglected) atau bahkan salah diagnosis (misdiagnosed). Sebagai alat bantu pengumpul informasi dasar dari penyakit, teknologi dapat dengan tepat dan cepat mengumpulkan data-data dasar yang dibutuhkan untuk menegakkan diagnosis. Terlebih lagi dengan penggunaan artificial intelligence (AI), sehingga data yang diperoleh dari pasien dapat diklasifikasi dan dipelajari oleh mesin belajar (machine learning) untuk dapat menentukan proses diagnostic selanjutnya. (Oski Illiandri et al, 2021)

2. Permasalahan

Selama ini mengukur obesitas dengan menggunakan Indeks Massa Tubuh (IMT). IMT adalah metode pengukuran melalui penghitungan antara tinggi badan dalam satuan meter dan berat badan dalam satuan kilogram. IMT merupakan metode yang cepat dan sangat mudah digunakan. Hasil penghitungan tersebut akan dikategorikan obesitas jika nilai IMT lebih dari 27 kg/m².

Menurut dr. Sienny Agustin dikutip dari laman alodokter ada beberapa alasan mengapa tidak dianjurkan untuk bergantung pada IMT diantaranya IMT tidak mempertimbangkan faktor lain seperti usia, jenis kelamin, genetik, gaya hidup, dan riwayat kesehatan. IMT juga tidak

mempertimbangkan distribusi lemak pada tubuh manusia, padahal lokasi lemak di tubuh dapat mengidentifikasi sehat atau tidaknya seseorang. Pengukuran dengan IMT menganggap semua bobot itu sama, sehingga mengesampingkan kemungkinan seseorang memiliki massa otot yang lebih tinggi dibandingkan orang lain, IMT juga menyebabkan bias berat badan dan sering kali malah mengabaikan masalah medis yang lebih serius. Alasan yang terakhir adalah bahwa nilai IMT tidak sama untuk seluruh populasi manusia di dunia. Pada sebuah penelitian di Amerika Serikat, kelompok olahragawan American Football didapatkan memiliki rata-rata IMT lebih dari 31,35 kg/m², menurut IMT mereka termasuk obesitas padahal disebabkan oleh massa otot yang besar. Maka dari itu, selain pengukuran IMT diperlukan juga pemeriksaan lain seperti pengukuran lingkar perut untuk mengetahui kadar lemak sentral untuk membedakan massa otot dan massa lemak. Dimana batas aman lingkar perut untuk pria adalah 90 cm dan wanita adalah 80 cm. Berdasarkan uraian permasalahan yang telah dijelaskan, maka identifikasi masalah antara lain:

- a. Belum akuratnya dalam menentukan diagnosa obesitas dengan IMT;
- b. Belum efektifnya proses penentuan diagnosa obesitas karena masih menggunakan hitungan manual.

3. Tujuan dari penelitian ini adalah:

- a. Dapat menemukan proses yang akurat dalam menentukan diagnosa obesitas yang tepat bersumber dari hasil pemeriksaan kesehatan yang terdiri dari Jenis Kelamin, Usia, Tinggi Badan, Berat Badan, Lingkar Perut, Perilaku Beresiko seperti Kurangnya Aktivitas Fisik, Pola Makan Gula Berlebih, Garam Berlebih, Lemak Berlebih, Kurang Makan Buah dan Sayur;
- b. Mendapatkan proses yang lebih efektif dalam menentukan diagnosa obesitas dengan algoritma Naïve Bayes;
- c. Mengembangkan *prototype* aplikasi penentuan diagnosa obesitas pada peserta sosialisasi deteksi dini penyakit tidak menular di Lingkungan Pemerintah Daerah Kota Bogor;

4. Tinjauan Pustaka

a. Klasifikasi

Menurut Prasetyo (2013, p.45) klasifikasi adalah penilaian suatu objek data yang akan dimasukkan dalam kelas kategoris dari jumlah-jumlah kelas yang sudah ada. Klasifikasi terdapat dua pekerjaan utama yang dilaksanakan, yang pertama adalah membangun suatu model menjadi *prototype* yang akan disimpan sebagai *memory* dan yang kedua mempergunakan model itu sebagai identifikasi, klasifikasi atau prediksi suatu objek data lain agar terlihat dalam kelas mana objek data tersebut berada di model yang sudah tersimpan. Klasifikasi diketahui dengan cara yang rinci untuk pekerjaan yang melaksanakan *training* atau pengkajian terhadap tugas target yang memvisualkan tiap-tiap set atribut (fitur) x ke satu dari sejumlah label kelas y yang telah ada. Pekerjaan *training* itu melahirkan model tertentu yang selanjutnya disimpan sebagai *memory*.

Dalam klasifikasi, model mempunyai arti yang sama dengan kotak hitam, dimana ada suatu model yang menerima masukan, kemudian mampu melakukan pemikiran terhadap masukan tersebut dan memberikan jawaban sebagai keluaran dari hasil pemikirannya. Model yang telah dirancang ketika *training* selanjutnya dipergunakan memprediksi label kelas data yang baru dan belum teridentifikasi sebelumnya. Pada konteks perancangan suatu model disaat prosedur pelatihan dibutuhkan algoritma guna membangun rancangan yang disebut algoritma pelatihan. Terdapat algoritma pelatihan yang sudah dikembangkan oleh peneliti terdahulu seperti *C4.5*, *Nearest Neighbor*, *Bayesian Classification*, *Neural Network* dan sebagainya.

b. Algoritma Naïve Bayes

Menurut Muhammad Arhami dan Muhammad Nasir (2020, p.104) Algoritma Naïve Bayes (NB) dapat digunakan untuk masalah klasifikasi biner dan multikelas. NB membuat dan menilai model dengan sangat cepat dan skala secara linear dalam jumlah prediksi dan baris.

NB merupakan klasifikasi yang mempresentasikan setiap kelas objek berdasarkan kesimpulan atau rekapitulasi probabilistik dan menemukan kemungkinan besar kelas yang sesuai untuk tiap objek-objek yang akan ditentukan kelasnya dari objek-objek uji yang ada berdasarkan atribut-atribut atau variabel yang telah diketahui nilainya. Klasifikasi Naïve Bayes juga digolongkan ke dalam pembelajaran yang terawasi (*supervised learning*) dimana untuk penentuan kelas objek uji didasarkan dari data latihnya dan setiap objek didata latih telah diketahui kelasnya masing-masing. Pada teorema Bayes, *probabilitas* dijelaskan sebagai berikut:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Dengan :

H = hipotesisnya

X = kejadian yang merelasikan ke hipotesis

P(H) = peluang hipotesis

P(X) = peluang kejadian

P(H|X) = jumlah probabilitas hipotesis H berdasarkan kondisi X

P(X|H) = jumlah probabilitas posterior X berdasarkan kondisi H

c. Metode Prototype

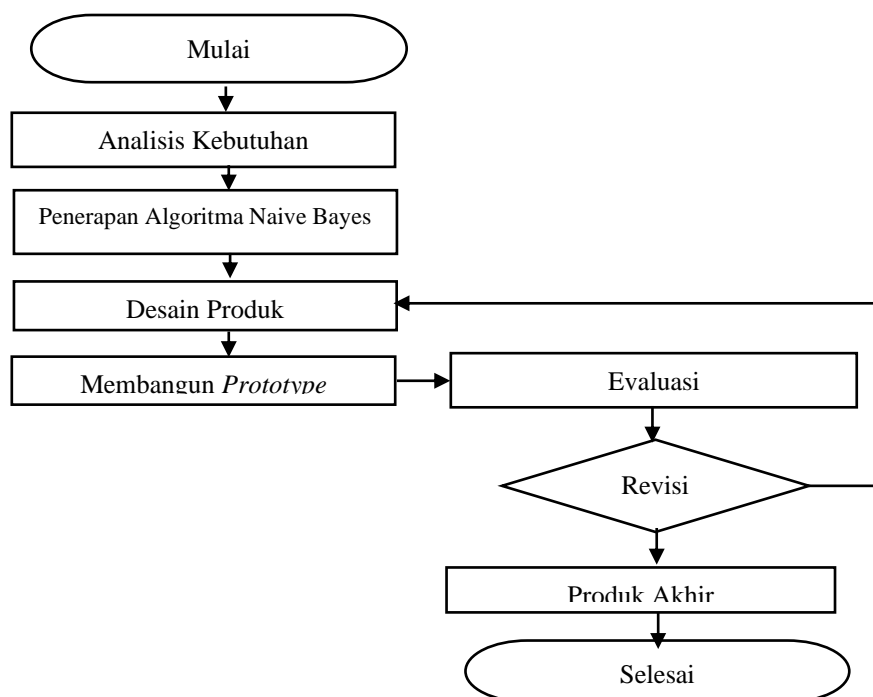
Menurut (Ogedebe, 2012), mengungkapkan bahwa prototyping adalah suatu metode pengembangan perangkat lunak yang membentuk suatu model berupa bentuk pertama dari sistem. Dengan metode prototype ini akan menghasilkan suatu prototipe sistem sebagai penghubung antara pengembang dan pengguna sehingga dapat berinteraksi dalam proses pengembangan sistem informasi. *Prototype* dimulai dengan dilakukannya komunikasi antar tim pengembang dengan pelanggan atau *user*". Pertemuan dengan *user* akan menghasilkan identifikasi spesifikasi kebutuhan, iterasi pembuatan *prototype* direncanakan dengan cepat dan pemodelan dalam bentuk rancangan cepat dilakukan. Rancangan tersebut berupa *prototype* rancangan antar muka pengguna (*user interface*) atau format tampilan. *Prototype* kemudian diserahkan kepada pengguna untuk dievaluasi yang kemudian akhirnya akan memberikan umpan balik yang akan digunakan untuk memperhalus spesifikasi kebutuhan. Iterasi akan terjadi saat *prototype* diperbaiki untuk memenuhi kebutuhan pengguna, sementara pada saat yang sama

B. METODE

1. Model Yang Diusulkan

- Dalam penelitian ini terdapat dua tujuan yang hendak dicapai, tujuan yang pertama adalah menerapkan algoritma Naïve Bayes untuk menentukan diagnosa obesitas, tujuan kedua adalah mengembangkan *prototype* aplikasi penentuan diagnosa obesitas. Untuk dapat mewujudkan tujuan tersebut digunakan dua buah metode atau model, yaitu metode atau model konseptual dan procedural. Metode konseptual merupakan metode pemecahan masalah secara konsep atau teori. Metode konseptual yang digunakan dalam penelitian ini yaitu metode Naïve bayes. Metode prosedural merupakan metode yang digunakan dalam rangka mengembangkan sistem informasi, Adapun metode procedural yang digunakan adalah metode prototype. Kedua metode tersebut di terapkan kedalam prosedur pengembangan penelitian yang akan dilakukan sebagaimana dapat dilihat pada gambar 1. Adapun penjelasan langkah-langkah prosedur pengembangan penelitian tersebut adalah :
- Analisis kebutuhan, yang mengidentifikasi semua kebutuhan, gambaran umum dari aplikasi yang akan dikembangkan dan tujuan membangun aplikasi. Analisis dilakukan dengan cara Pengumpulan dokumen, data yang dikumpulkan adalah data Hasil Pemeriksaan Kesehatan bulan Juli 2022. Dari data tersebut dapat diperoleh diagnosa obesitas atau normal dari pegawai di Lingkungan Sekretariat Daerah Kota Bogor. Selain

- pengumpulan dokumen juga dilakukan wawancara dengan narasumber petugas puskesmas dari UPTD Bogor Tengah dan Bagian Kepegawaian Sekretariat Daerah Kota Bogor
- Penerapan algoritma Naive Bayes terdiri dari penerapan metode dari algoritma Naive Bayes yang digunakan dengan memasukkan data sampel dan secara bertahap menghitung per kelas;
 - Product Design*, yang melakukan perancangan pada fase-fase aplikasi yang akan dilakukan, guna mencapai tujuan dari aplikasi yang sesuai dengan kebutuhan pengguna atau *user*;
 - Membangun model, yaitu membuat rancangan *prototype* sesuai dengan aplikasi *rule* akan dikembangkan;
 - Evaluasi*, yaitu menguji produk jadi yang dilakukan oleh ahli sistem informasi dan pengguna untuk melihat keberhasilan aplikasi berdasarkan kebutuhan dan kesalahan aplikasi;
 - Revisi* yaitu melakukan perbaikan dan pengecekan apakah aplikasi sudah sesuai kebutuhan atau tidak, jika sudah sesuai maka akan menjadi produk akhir, namun jika terjadi kesalahan saat pengujian maka akan kembali ke tahap desain produk;
 - Produk jadi, yaitu produk yang telah melewati tahap evaluasi oleh pakar sistem, pakar materi, dan pengguna yang sesuai.
 - Evaluasi*, yaitu menguji produk jadi dengan ahli sistem, ahli materi, dan pengguna
- Langkah-langkah prosedur pengembangan penelitian ditunjukkan pada gambar 1 berikut



Gambar 1. Prosedur Pengembangan

2. Instrumen Penelitian

Instrumen pengumpulan data yang digunakan adalah kuesioner yang diisi oleh responden. Berdasarkan peran dan status responden dalam penelitian, instrument yang dihimpun terdiri atas:

a. Instrumen untuk Ahli Sistem Informasi

Instrumen untuk ahli sistem informasi menggunakan kuisisioner yang berisi pertanyaan terkait pengujian blackbox, Pengujian black box yaitu menguji perangkat lunak dari segi spesifikasi fungsional tanpa menguji desain dan kode program (Rosa A.S dan M. Shalahudin, 2011). Pengujian Blackbox dilakukan di oleh 2 orang ahli sistem informasi yang hasilnya akan di analisa untuk mengetahui ketepatan fungsionalitas pada aplikasi

b. Instrumen Untuk Pengguna

Instrumen untuk pengguna menggunakan kuesioner yang berisi pertanyaan menggunakan paket kuesioner PSSUQ (Post-Study Sistem Usability Questionnaire) yang diolah dengan

menilai rata-rata dan melakukan uji signifikansi penilaian. Pengujian dilakukan oleh 4 orang pengguna yang hasilnya akan dianalisis untuk mengetahui kebergunaan aplikasi yang telah dikembangkan

3. Teknik Analisis Data

a. Uji Produk

Dalam penelitian ini, menggunakan presentase kelayakan untuk analisis kelayakan. Adapun rumus yang digunakan adalah sebagai berikut:

$$\text{persentase kelayakan (\%)} = \frac{\text{skor yang diobservasi}}{\text{skor yang diharapkan}} \times 100\%$$

Hasil presentase digunakan untuk memberikan jawaban atas kelayakan dari aspek-aspek yang diteliti. Terdapat 5 kategori kelayakan yang menunjukkan rentang dari bilangan presentase. Nilai minimum 0% dan maksimal yang diharapkan adalah 100% kategori kelayakan menurut Arikunto (2009:44) dapat dilihat pada tabel berikut:

Tabel 1. Kategori Kelayakan

Presentase Pencapaian	Interpetasi
<21%	Sangat Tidak Layak
21% - 40%	Tidak Layak
41% - 60%	Cukup Layak
61% - 80%	Layak
81% - 100%	Sangat Layak

Untuk mengetahui kelayakan digunakan tabel di atas sebagai acuan penilaian data yang dihasilkan dari validasi pengguna.

b. Uji Hasil

Cara yang berguna untuk menganalisis seberapa baik *classifier* mengenali *tuple* dari kelas yang berbeda (Han et al., 2012, p.365). Sebuah matrik dari prediksi yang akan dibandingkan dengan kelas yang asli dari inputan atau dengan kata lain berisi informasi nilai aktual dan prediksi pada klasifikasi.

Tabel 2. Confusion Matrix

<i>Classification</i> <i>Actual Class</i>	<i>Predicted class</i>	
	<i>Yes</i>	<i>No</i>
<i>yes</i>	<i>a (true positive-TP)</i>	<i>b (false negative-FN)</i>
<i>no</i>	<i>c (false positive-FP)</i>	<i>d (true negative-TN)</i>

Akurasi adalah perbandingan kasus yang diidentifikasi benar dengan jumlah semua kasus. Rumus untuk menghitung tingkat akurasi pada matrik adalah:

$$\text{accuracy} = \frac{TP+TN}{TP+FP+TN+FN} = \frac{A+D}{A+B+C+D}$$

Keterangan:

A = jika hasil prediksi positif dan data sebenarnya positif

B = jika hasil prediksi negatif dan data sebenarnya positif

C = jika hasil prediksi positif dan data sebenarnya negatif

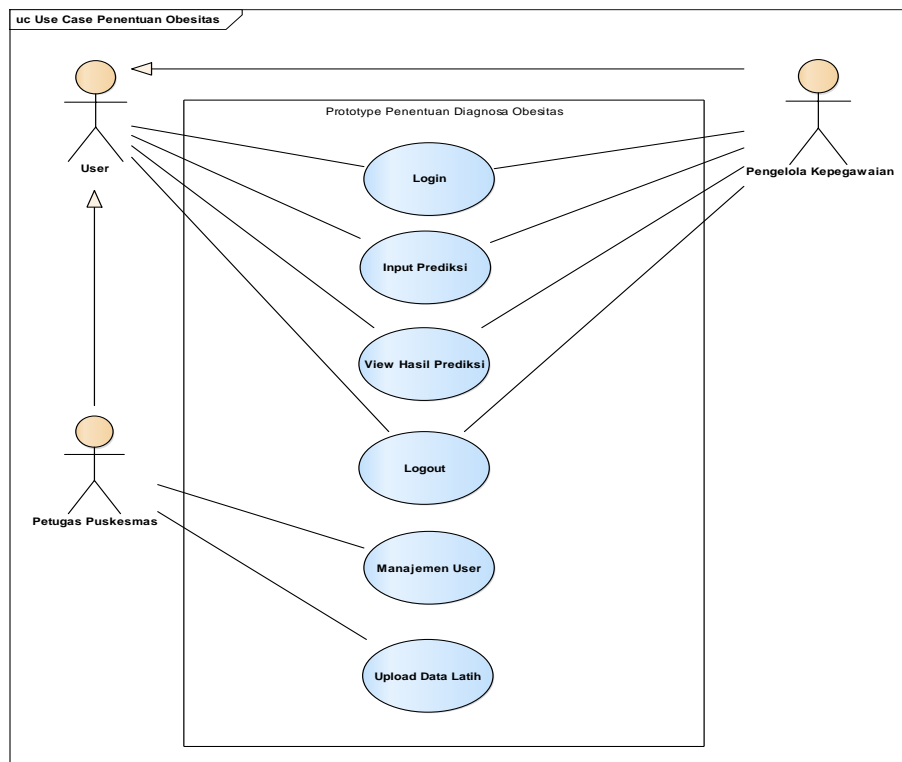
D = jika hasil prediksi negatif dan data sebenarnya negatif

C. HASIL DAN PEMBAHASAN

1. HASIL

a. Analisa Produk

untuk mengetahui fungsi apa saja yang ada di dalam sebuah sistem informasi *prototype* penentuan diagnosa obesitas dan siapa saja yang berhak menggunakan fungsi-fungsi tersebut dapat dilihat pada diagram usecase yang disajikan pada gambar 2.



Gambar 2 Use case Diagram

Pada gambar 2 dijelaskan bahwa terdapat 2 aktor dalam sistem penentuan diagnosa obesitas pada peserta deteksi dini penyakit tidak menular yaitu petugas puskesmas dan pengelola kepegawaian setda yang diharuskan login terlebih dahulu agar dapat mengakses sistem tersebut. Untuk petugas puskesmas dapat mengakses manajemen user, upload dataset atau data latih, kemudian input prediksi, dan melihat hasil prediksi. Sedangkan pengelola kepegawaian setda dapat mengakses input prediksi dan melihat hasil prediksi.

b. Analisa Metode

Metode yang digunakan adalah metode Naïve Bayes yang akan menghasilkan prediksi diagnosa obesitas atau normal pada peserta sosialisasi deteksi dini penyakit tidak menular. Dalam data mining ada beberapa tahapan sebelum penerapan Algoritma Naïve bayes yaitu :

a) Data Selection;

Data yang akan digunakan adalah data pemeriksaan kesehatan pada bulan Juli tahun 2022. Data pemakaian obat tersebut sebanyak 600 record dengan 60 jumlah pegawai. Data tersebut dalam bentuk excel. Dalam tahap data *selection* ini data yang dikumpulkan diseleksi atau dipilih data yang siap digunakan untuk prediksi diagnosa obesitas. Jumlah atribut yang digunakan sebanyak 10 atribut yaitu atribut nama, usia, jenis kelamin, tinggi badan, berat badan, lingkar perut, kurangnya aktivitas fisik, pola makan gula berlebih, garam berlebih, lemak berlebih, kurang makan buah dan sayur serta diagnosa sebagai atribut label. Dilakukan pengecekan data apakah ada duplikasi data, data yang kosong atau tidak konsisten. Data pemeriksaan kesehatan yang akan dijadikan sebagai data training atau data latih dapat dilihat pada tabel 3 berikut ini:

Tabel 3. Tabel Data Training

Pegawai	JK	Usia	TB	BB	LP	Kurang Aktifitas Fisik	Gula Berlebih	Garam Berlebih	Lemak Berlebih	Kurang Makan Buah dan Sayur	Diagnosa
1	P	41	156	93.6	112	YA	YA	TIDAK	TIDAK	TIDAK	OBESITAS
2	P	40	157	52.2	79	YA	TIDAK	TIDAK	TIDAK	TIDAK	NORMAL
3	L	54	168.5	68.8	85	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	NORMAL
4	P	50	153	60.5	87	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	OBESITAS
5	P	50	162	61	87	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	OBESITAS
6	L	49	178	86.4	102	TIDAK	TIDAK	TIDAK	YA	YA	OBESITAS
7	P	43	149	67	85	TIDAK	YA	YA	YA	TIDAK	OBESITAS
8	P	36	165	55	70	YA	YA	TIDAK	TIDAK	YA	NORMAL
9	P	41	159	63.8	80	YA	YA	TIDAK	YA	TIDAK	OBESITAS
10	L	43	165	65	88	YA	TIDAK	YA	YA	TIDAK	NORMAL
...
60	P	30	165	65.6	78	YA	TIDAK	TIDAK	TIDAK	TIDAK	NORMAL

b) Transformasi Data

Tahap selanjutnya setelah seleksi data adalah transformasi data. proses *transformation* dengan cara mengubah pengetahuan dari data asli (*real*) ke dalam bentuk yang sesuai untuk penambangan. Proses *transformation* adalah dapat dilihat pada tabel 4 berikut:

Tabel 4. Transformasi Data

No	Atribut	Nilai
1	Jenis Kelamin	P = Perempuan; L= Laki-Laki
2	Usia	Dewasa Awal = 25-35 Tahun ; Dewasa Akhir = 36-46 Tahun ; Lansia Awal = 47-57 Tahun
3	Tinggi Badan	Pendek = 140 - \leq 152 cm ; Sedang = $>$ 152 - \leq 165 cm, Tinggi = $>$ 165 - \leq 178 cm
4	Berat Badan	Kecil = 40 - \leq 58 kg ; Sedang = $>$ 58 - \leq 77 kg Besar = $>$ 77 - \leq 96 kg
5	Lingkar Perut	Kecil = 62 - \leq 79 cm ; Sedang = $>$ 79 - \leq 97 cm Besar = $>$ 97 - \leq 115 cm
6	Kurang Aktifitas Fisik	Ya ; Tidak
7	Pola Makan Gula Berlebih	Ya ; Tidak
8	Pola Makan Garam Berlebih	Ya ; Tidak
9	Pola Makan Lemak Berlebih	Ya; Tidak
10	Pola Makan Kurang Makan Buah dan Sayur	Ya ; Tidak
11	Diagnosa	Obesitas ; Normal

c) Tahap Pre-Processing

Setelah transformasi data dari numerik ke kategorik, maka hasil data transformasinya adalah sebagai berikut:

Tabel 5 Data Training Setelah Ditransformasi

Pegawai	JK	Usia	TB	BB	LP	Kurang Aktifitas Fisik	Gula Berlebih	Garam Berlebih	Lemak Berlebih	Kurang Makan Buah dan Sayur	Diagnosa
1	P	Dewasa Akhir	Sedang	Besar	Besar	YA	YA	TIDAK	TIDAK	TIDAK	OBESITAS
2	P	Dewasa Akhir	Sedang	Kecil	Kecil	YA	TIDAK	TIDAK	TIDAK	TIDAK	NORMAL
3	L	Lansia Awal	Tinggi	Sedang	Sedang	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	NORMAL
4	P	Lansia Awal	Sedang	Sedang	Sedang	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	OBESITAS
5	P	Lansia Awal	Sedang	Sedang	Sedang	TIDAK	TIDAK	TIDAK	TIDAK	TIDAK	OBESITAS
10	L	Dewasa Akhir	Sedang	Sedang	Sedang	YA	TIDAK	YA	YA	TIDAK	NORMAL
...
60	P	Dewasa Awal	Sedang	Sedang	Sedang	YA	TIDAK	TIDAK	TIDAK	TIDAK	NORMAL

d) Penerapan Algoritma Naïve Bayes

Dari Data *Training* pada Tabel 5 diatas diterapkan Algoritma Naïve Bayes untuk memprediksi data *testing* yaitu:

Tabel 6 Data Testing/Data Uji

Pegawai	JK	Usia	TB	BB	LP	Kurang Aktifitas Fisik	Gula Berlebih	Garam Berlebih	Lemak Berlebih	Kurang Makan Buah dan Sayur	Diagnosa
11	L	DEWASA AKHIR	TINGGI	BESAR	SEDANG	TIDAK	TIDAK	YA	TIDAK	TIDAK	?
12	P	DEWASA AWAL	PENDEK	KECIL	KECIL	YA	YA	TIDAK	TIDAK	TIDAK	?

Maka dengan rumus (1) perhitungan Naïve Bayes sebagai berikut :

- a. Menghitung nilai probabilitas (peluang) dari kasus baru setiap kelas (label) yang ada "P(XK|C_i) Tahap pertama diawali dengan menghitung nilai probabilitas data kasus baru setiap kelas dengan menghitung keseluruhan kelas diagnosa "Obesitas" dengan diagnosa "Normal". Perhitungannya antara lain:
 1. $P(Y=Obesitas) = 37/60 = 0,62$ atau jumlah data diagnosa obesitas dibagi seluruh data
 2. $P(Y=Normal) = 23/60 = 0,38$ atau jumlah data diagnosa normal dibagi seluruh data
 3. $P(\text{Jenis Kelamin} = "P" | \text{Diagnosa} = "Obesitas") = 27/37 = 0,73$
 4. $P(\text{Jenis Kelamin} = "P" | \text{Diagnosa} = "Normal") = 16/23 = 0,7$
 5. $P(\text{Jenis Kelamin} = "L" | \text{Diagnosa} = "Obesitas") = 10/37 = 0,27$
 6. $P(\text{Jenis Kelamin} = "L" | \text{Diagnosa} = "Normal") = 7/23 = 0,3$
 7. $P(\text{Usia} = "Dewasa Awal" | \text{Diagnosa} = "Obesitas") = 5/37 = 0,13$
 8. $P(\text{Usia} = "Dewasa Awal" | \text{Diagnosa} = "Normal") = 8/23 = 0,35$
 9. $P(\text{Usia} = "Dewasa Akhir" | \text{Diagnosa} = "Obesitas") = 22/37 = 0,60$
 10. $P(\text{Usia} = "Dewasa Akhir" | \text{Diagnosa} = "Normal") = 11/23 = 0,48$
 11. $P(\text{Usia} = "Lansia Awal" | \text{Diagnosa} = "Obesitas") = 10/37 = 0,27$
 12. $P(\text{Usia} = "Lansia Awal" | \text{Diagnosa} = "Normal") = 4/23 = 0,17$

13. $P(\text{Tinggi Badan} = \text{"Pendek"} \mid \text{Diagnosa} = \text{"Obesitas"}) = 6/37 = 0,16$
14. $P(\text{Tinggi Badan} = \text{"Pendek"} \mid \text{Diagnosa} = \text{"Normal"}) = 3/23 = 0,13$
15. $P(\text{Tinggi Badan} = \text{"Sedang"} \mid \text{Diagnosa} = \text{"Obesitas"}) = 25/37 = 0,68$
16. $P(\text{Tinggi Badan} = \text{"Sedang"} \mid \text{Diagnosa} = \text{"Normal"}) = 16/23 = 0,70$
17. $P(\text{Tinggi Badan} = \text{"Tinggi"} \mid \text{Diagnosa} = \text{"Obesitas"}) = 6/37 = 0,16$
18. $P(\text{Tinggi Badan} = \text{"Tinggi"} \mid \text{Diagnosa} = \text{"Normal"}) = 4/23 = 0,17$
19. $P(\text{Berat Badan} = \text{"Kecil"} \mid \text{Diagnosa} = \text{"Obesitas"}) = 5/37 = 0,13$
20. $P(\text{Berat Badan} = \text{"Kecil"} \mid \text{Diagnosa} = \text{"Normal"}) = 10/23 = 0,43$
21. $P(\text{Berat Badan} = \text{"Sedang"} \mid \text{Diagnosa} = \text{"Obesitas"}) = 23/37 = 0,62$
22. $P(\text{Berat Badan} = \text{"Sedang"} \mid \text{Diagnosa} = \text{"Normal"}) = 13/23 = 0,57$
23. $P(\text{Berat Badan} = \text{"Besar"} \mid \text{Diagnosa} = \text{"Obesitas"}) = 9/37 = 0,24$
24. $P(\text{Berat Badan} = \text{"Besar"} \mid \text{Diagnosa} = \text{"Normal"}) = 0/23 = 0$
25. $P(\text{Lingkar Perut} = \text{"Kecil"} \mid \text{Diagnosa} = \text{"Obesitas"}) = 2/37 = 0,05$
26. $P(\text{Lingkar Perut} = \text{"Kecil"} \mid \text{Diagnosa} = \text{"Normal"}) = 16/23 = 0,7$

Menghitung nilai akumulasi probabilitas dari setiap kelas pada setiap kejadian pada data uji. Maka perhitungannya sebagai berikut:

Pegawai 11

1. $P(\text{Diagnosa} = \text{"Obesitas"}) \times P(Y = \text{"Obesitas"}) = 0,27 \times 0,60 \times 0,16 \times 0,24 \times 0,7 \times 0,30 \times 0,78 \times 0,35 \times 0,6 \times 0,54 = 0,00011$
2. $P(\text{Diagnosa} = \text{"Normal"}) \times P(Y = \text{"Normal"}) = 0,3 \times 0,48 \times 0,17 \times 0 \times 0,30 \times 0,35 \times 0,74 \times 0,04 \times 0,83 \times 0,78 = 0$

Pegawai 12

1. $P(\text{Diagnosa} = \text{"Obesitas"}) \times P(Y = \text{"Obesitas"}) = 0,73 \times 0,13 \times 0,16 \times 0,13 \times 0,05 \times 0,7 \times 0,21 \times 0,65 \times 0,6 \times 0,54 = 0,000003$
2. $P(\text{Diagnosa} = \text{"Normal"}) \times P(Y = \text{"Normal"}) = 0,7 \times 0,35 \times 0,13 \times 0,43 \times 0,7 \times 0,65 \times 0,26 \times 0,95 \times 0,82 \times 0,78 = 0,00098$

- b. Menghitung nilai probabilitas akhir setiap kelas (label) " $P(X|C_i) \times P(C_i)$ ", yaitu menghitung nilai probabilitas dari seluruh nilai variabel pada setiap kelas sebagai berikut:

Pegawai 11

1. $P(\text{Diagnosa} = \text{"Obesitas"}) \times P(Y = \text{"Obesitas"}) = 0,00011 \times 0,62 = \mathbf{0,000072}$
2. $P(\text{Diagnosa} = \text{"Normal"}) \times P(Y = \text{"Normal"}) = 0 \times 0,38 = 0$

Pegawai 12

1. $P(\text{Diagnosa} = \text{"Obesitas"}) \times P(Y = \text{"Obesitas"}) = 0,000003 \times 0,62 = 0,000002$
2. $P(\text{Diagnosa} = \text{"Normal"}) \times P(Y = \text{"Normal"}) = 0,00098 \times 0,38 = \mathbf{0,00037}$

- c. Tahap terakhir adalah menentukan probabilitas akhir terbesar dari tiap-tiap kelas yang berarti membandingkan nilai terbesar dalam nilai probabilitas akhir kelas diagnosa, hasilnya adalah pada tabel dibawah ini:

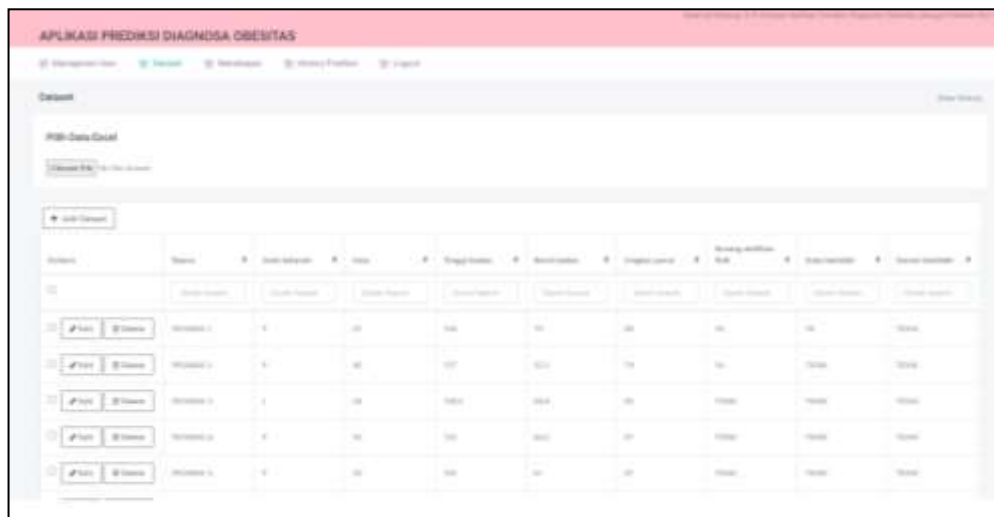
Tabel 7. Hasil Prediksi

Pegawai	JK	Usia	TB	BB	LP	Kurang Aktifitas Fisik	Gula Berlebih	Garam Berlebih	Lemak Berlebih	Saxur dan Buah	Kurang Makan	Diagnos ^a
11	L	DEWASA AKHIR	TINGGI	BESAR	SEDANG	TIDAK	TIDAK	YA	TIDAK	TIDAK	TIDAK	OBESITAS
12	P	DEWASA AWAL	PENDEK	KECIL	KECIL	YA	YA	TIDAK	TIDAK	TIDAK	TIDAK	NORMAL

c. Produk

1) Halaman Dataset

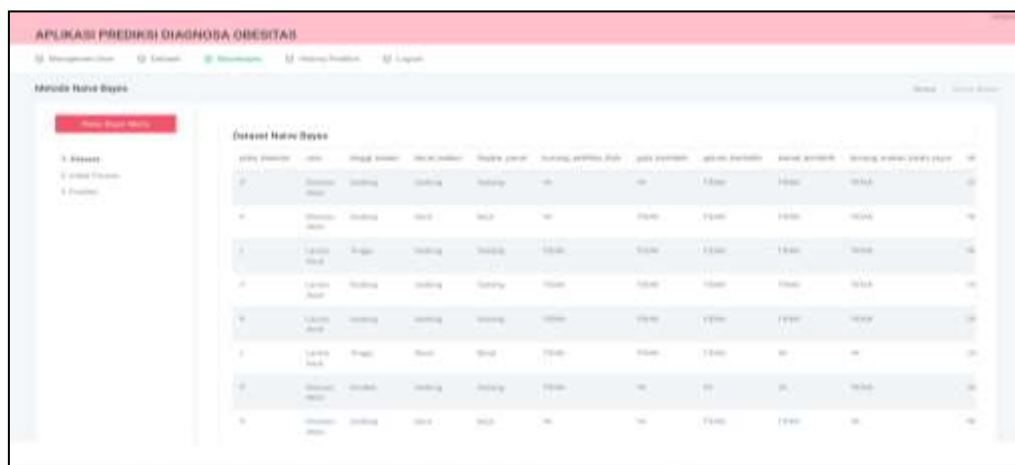
Pada laman Dataset, admin dapat melakukan *upload* data dalam format excel, *edit*, *delete* dan *add* data pada laman ini.



Gambar 3. Halaman Data Set

2) Halaman proses inisialisasi

Halaman Proses inisialisasi berisi dataset yang sudah ditransformasi



Gambar 4. Halaman Proses Inisialisasi

3) Halaman Prediksi

Halaman ini menampilkan hasil prediksi Algoritma Naïve bayes.



Gambar 5. Halaman Prediksi

2. PEMBAHASAN

a. Hasil Kuisisioner Ahli Sistem Informasi

Prototype system di uji oleh 2 ahli sistem informasi, informasi yang dihasilkan dari pengolahan kuisisioner tersebut merupakan gambaran pendapat ahli terhadap kelayakan fungsionalitas prototype sistem prediksi diagnosa obesitas, skala yang digunakan untuk pengujian tersebut digunakan skala guttman. Adapun hasil pengolahan data kuisisioner yang disebut seperti yang ditunjukkan pada tabel 8:

Tabel 8. Hasil Kuisisioner Ahli Sistem Informasi

No	Skenario Pengujian	Proses yang diuji / Test Case	Hasil Yang diharapkan	Hasil	
				Ahli 1	Ahli 2
1	User name benar sedangkan password salah	Login	Akan menampilkan tampilan log in kembali	1	1
2	User name dan Password Benar	Log in	Akan masuk ke dalam system dan tampil menu utama	1	1
3	Memiliki empat menu utama yaitu, manajemen user, dataset, naïve bayes dengan sub menu data set, initial process, prediksi dan history	Hak akses	Setiap menu akan tampil sesuai dengan apa yang klik	1	1
4	Memilih Menu manajemen user	Data User	Hanya menampilkan data user	1	1
5	Memilih Menu data set	Data set	Menampilkan data set, upload dataset dan edit, delete dataset	1	1
6	Memilih menu naïve bayes dengan sub menu dataset, initial process dan prediksi	Data latih, Data uji, Prediksi	Menampilkan data set, initial process dan prediksi	1	1
7	Memilih menu history	Data Hasil prediksi	Akan menampilkan data hasil prediksi	1	1
Total Skor				7	7
Total				14	

$$\text{Persentase kelayakan} = \frac{\text{skor yang diobservasi}}{\text{skor yang diharapkan}} \times 100\%$$

$$\text{Persentase kelayakan} = \frac{14}{14} \times 100\%$$

$$\text{Persentase kelayakan} = 100\%$$

Dari perhitungan diatas maka berdasarkan skala kelayakan Arikunto, masuk ke dalam katagor "Sangat Layak".

b. Hasil Kuisisioner Pengguna

uji coba pengguna bertujuan untuk mengetahui pendapat pengguna tentang kebergunaan *prototype* aplikasi yang telah dibangun. Kuesioner yang digunakan yaitu *Post-Study Sistem Usability Questionnaire* (PSSUQ). Responden pengguna adalah 4 orang yaitu 2 orang sebagai admin dari petugas puskesmas dan 2 orang sebagai user dari pengelola kepegawaian Sekretariat Daerah Kota Bogor. Tabel hasil uji pengguna adalah sebagai berikut :

Tabel 9. Hasil Kuisisioner Pengguna

No.	Pertanyaan	R1	R2	R3	R4
1	Secara Keseluruhan, saya puas dengan kemudahan penggunaan aplikasi ini	6	6	7	6
2	Aplikasi ini mudah digunakan	6	6	6	6
3	Saya secara efektif dapat menyelesaikan tugas -tugas dan scenario menggunakan aplikasi ini	6	6	7	7
4	Saya bisa menyelesaikan tugas tugas dan scenario menggunakan aplikasi ini	7	6	6	7
5	Saya dengan efisien dapat menyelesaikan tugas-tugas dan scenario menggunakan aplikasi ini	6	6	7	7
6	Saya merasa nyaman menggunakan aplikasi ini	7	6	7	6
7	Mudah untuk belajar menggunakan aplikasi ini	6	6	7	7
8	Saya percaya saya bisa menjadi produktif dengan cepat menggunakan aplikasi ini	6	6	7	7
9	Aplikasi ini memberikan pesan kesalahan yang jelas memberitahu saya bagaimana untuk memperbaiki masalah	6	5	7	7
10	Setiap kali saya melakukan kesalahan , dengan aplikasi ini bisa pulihdengan mudah dan cepat	7	5	7	7
11	Informasi (seperti online pesan bantuan pada layer, dan dokumentasi lainnya) disediakan dengan jelas oleh aplikasi ini	7	6	7	7
12	Mudah untuk menemukan informasi yang saya butuhkan	6	7	6	6
13	Informasi efektif dalam membantu menyelesaikan tugas-tugas dan skenario	7	7	6	6
14	Informasi efektif dalam membantu menyelesaikan tugas-tugas dan skenario	6	6	7	6
15	Informasi efektif dalam membantu menyelesaikan tugas-tugas dan scenario	6	6	7	7
16	Antarmuka aplikasi ini menyenangkan	6	6	7	6
17	Saya suka menggunakan antarmuka aplikasi ini	7	6	7	6
18	Aplikasi ini memiliki semua fungsi dan kemampuan yang saya harapkan	6	7	7	7
19	Secara keseluruhan, saya puas dengan aplikasi ini	6	6	7	7
	Total Skor	120	115	129	125
	Total keseluruhan	489			

$$\text{Persentase kelayakan} = \frac{\text{skor yang diobservasi}}{\text{skor yang diharapkan}} \times 100\%$$

$$\text{Persentase kelayakan} = \frac{(120+115+129+125)}{532} \times 100\%$$

$$\text{Persentase kelayakan} = \frac{489}{532} \times 100\%$$

$$\text{Persentase kelayakan} = 91\%$$

Dari data hasil pengujian diatas diperoleh persentase kelayakan sebesar 91 %, berdasarkan skala kelayakan arikunto masuk ke dalam katagori” Sangat layak”,

c. Uji Hasil

Uji hasil dilakukan dengan pengukuran akurasi menggunakan Confusion Matrix. Pengukuran akurasi Pengukuran akurasi dilakukan dengan membandingkan hasil

rekomendasi data nyata sebanyak 60 data berdasarkan variabel yang telah ditetapkan dengan data yang seharusnya atau data nyata. Hasil perbandingan dapat dilihat pada tabel 10

Tabel 10 Perhitungan Confusion Matrix

Aktual	Prediksi	
	Obesitas	Normal
Obesitas	TP (35)	FP (2)
Normal	FN (6)	TN (17)

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} =$$

$$Akurasi = \frac{35 + 17}{35 + 2 + 17 + 6} \times 100\% = 86,6\%$$

Berdasarkan perhitungan akurasi, maka didapatkan hasil akurasi sebesar 86,6%.

D. KESIMPULAN

1. Kesimpulan

Dari penelitian dan pembahasan yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut :

- Bahwa Algoritma Naïve Bayes dapat digunakan secara akurat untuk penentuan diagnosa obesitas pada peserta sosialisasi deteksi dini penyakit tidak menular di Lingkungan Pemerintah dengan menggunakan variabel sebanyak 10 yaitu Jenis Kelamin, Usia, Tinggi Badan, Berat Badan, Lingkar Perut, Kurang Aktifitas Fisik, Pola Makan Gula Berlebih, Garam Berlebih, Lemak Berlebih dan Kurang Makan dan Buah;
- Berdasarkan hasil perhitungan akurasi dengan *Confusion Matrix* didapatkan nilai akurasi sebesar 86,6%.

2. Saran

Saran untuk pengembangan lebih lanjut kedepan agar lebih baik, diantaranya:

- Pada penelitian ini data yang digunakan sebanyak 60 pegawai, dengan jumlah variabel 10 yaitu Jenis Kelamin, Usia, Tinggi Badan, Berat Badan, Lingkar Perut, Kurang Aktifitas Fisik, Pola Makan Gula Berlebih, Garam Berlebih, Lemak Berlebih dan Kurang Makan dan Buah. Agar penelitian menjadi lebih baik, maka dapat digunakan lebih banyak data dan menambahkan variabel yang lain misalkan kebiasaan merokok, kebiasaan minum alkohol.
- Penelitian selanjutnya dapat menggunakan algoritma prediksi yang lain seperti Algoritma C.45 yang dapat menangani atribut kontinyu dan diskrit.

E. DAFTAR PUSTAKA

- [1] Arikunto, Suharsimi. 2006. *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta
- [2] Ghaniy, R., & Sihotang, K. (2019). Penerapan Metode Naïve Bayes Classifier Untuk Penentuan Topik Tugas Akhir. *TeknoIS : Jurnal Ilmiah Teknologi Informasi dan Sains*, 9(1), 63-72. doi:<https://doi.org/10.36350/jbs.v9i1.7>
- [3] Han, J. 2012. *Data Mining Concepts and Techniques*. In Morgan Kaufmann is an imprint of Elsevier
- [4] Lubis, D., & Gusti, G. (2023). Penerapan Algoritma Naïve Bayes Untuk Penentuan Balita Penerima Makanan Tambahan (PMT) Berdasarkan Status Gizi Di Pos Pelayanan Terpadu (POSYANDU). *TeknoIS : Jurnal Ilmiah Teknologi Informasi dan Sains*, 13(1), 58-66.

doi:<https://doi.org/10.36350/jbs.v13i1.177>

- [5] Muhammad Arhami dan Muhammad Nasir. 2020. Data Mining Algoritma dan Implementasi. Penerbit Andi
- [6] Ogedebe, P.M., & Jacob, B.P. (2012). Software Prototyping: A Strategy to Use when User Lacks Data Processing Experience. ARPN Journal of Systems and Software. VOL.2, No.6, 2012, http://scientificjournals.org/journalofsystemsandsoftware/archive/vol2no6/vol2no6_4.pdf
- [7] Oski Illiandri, 2021. Informasi teknologi di dunia ilmu kesehatan, Penerbit Media Sains Indonesia
- [8] Prasetyo, Eko. 2013. Data Mining Konsep Dan Aplikasi Menggunakan Matlab. Yogyakarta: Penerbit Andi
- [9] Triawan, A., & Lintang Melinda, D. (2020). Penerapan Metode Naïve Bayes Untuk Rekomendasi Topik Tugas Akhir Berdasarkan Daftar Hasil Studi Mahasiswa di Perguruan Tinggi. *TeknoIS : Jurnal Ilmiah Teknologi Informasi dan Sains*, 10(2), 58-70. doi:<https://doi.org/10.36350/jbs.v10i2.91>
- [10] Utari, L., & Ulfah, Y. (2021). Penerapan Metode Naïve Bayes untuk Prediksi Minat Baca Berdasarkan Usia. *TeknoIS : Jurnal Ilmiah Teknologi Informasi dan Sains*, 11(1), 67-74. doi:<https://doi.org/10.36350/jbs.v11i1.104>