

**ANALISIS PREDIKSI AWAL PENYAKIT DIABETES MELLITUS
MENGUNAKAN ALGORITMA NAÏVE BAYES**

SKRIPSI



disusun oleh

Ahmad Nashruddin

17.11.1240

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**ANALISIS PREDIKSI AWAL PENYAKIT DIABETES MELLITUS
MENGUNAKAN ALGORITMA NAÏVE BAYES**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Ahmad Nashruddin

17.11.1240

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

**ANALISIS PREDIKSI AWAL PENYAKIT DIABETES MELLITUS
MENGUNAKAN ALGORITMA NAÏVE BAYES**

yang dipersiapkan dan disusun oleh

Ahmad Nashruddin

17.11.1240

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 3 Mei 2021

Dosen Pembimbing,

Rumini, M.Kom.

NIK. 190302246

PENGESAHAN
SKRIPSI
ANALISIS PREDIKSI AWAL PENYAKIT DIABETES MELLITUS
MENGGUNAKAN ALGORITMA NAÏVE BAYES

yang dipersiapkan dan disusun oleh

Ahmad Nashruddin

17.11.1240

telah dipertahankan di depan Dewan Penguji
pada tanggal 17 Juni 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Sri Ngudi Wahyuni, S.T., M.Kom.
NIK. 190302060

Ikamah, M.Kom
NIK. 190302282

Rumini, M.Kom
NIK. 190302246

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 17 Juni 2021

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, M.Kom
NIK. 190302096

PERNYATAAN

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 22 Juni 2021



Ahmad Nashruddin

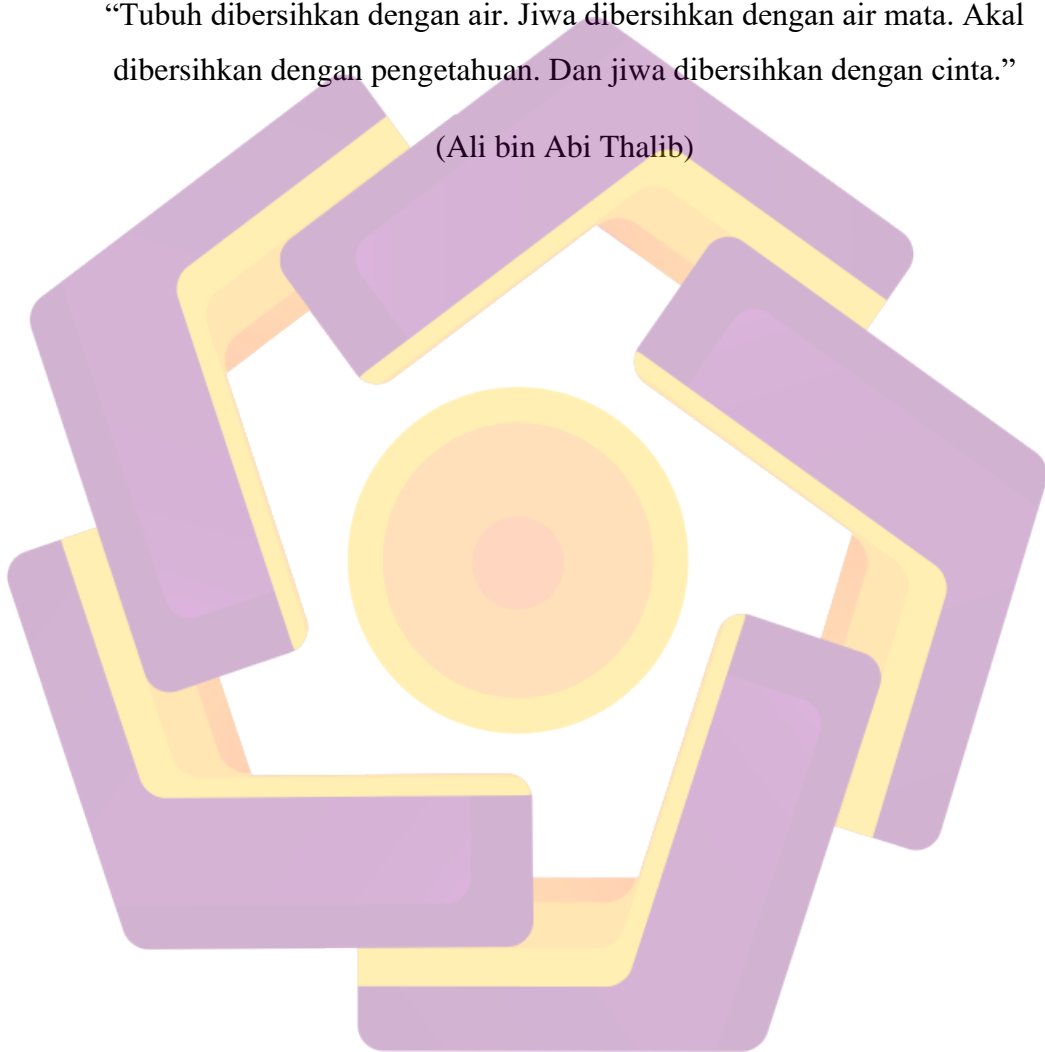
NIM. 17.11.1240

MOTTO

“Amalan yang lebih dicintai Allah adalah amalan yang terus menerus dilakukan walaupun sedikit.” (HR. Bukhari dan Muslim)

“Tubuh dibersihkan dengan air. Jiwa dibersihkan dengan air mata. Akal dibersihkan dengan pengetahuan. Dan jiwa dibersihkan dengan cinta.”

(Ali bin Abi Thalib)



PERSEMBAHAN

Dengan penuh rasa syukur saya persembahkan skripsi ini kepada semua pihak yang terlibat secara langsung maupun tidak langsung dalam proses pembuatan skripsi.

1. Allah SWT yang memberikan segala nikmat dan kasih sayangnnya sampai sejauh ini.
2. Kedua orang tua saya dan keluarga, yang selalu memberikan do'a, semangat, dan dukungan tak henti – hentinya kepada saya.
3. Ibu Rumini, M.Kom yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
4. Dosen – dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu selama perkuliahan.
5. Teman – teman kelas 17 IF 05 yang selalu menemani perkuliahan, mendukung dan memberikan semangat sampai saat ini.
6. Nurul Latifah yang selalu menemani, memberikan semangat, dan berbagi keluh kesah kepada saya.
7. Orang – orang baik yang senantiasa mendoakan, memberi semangat dan bantuan kepada saya.

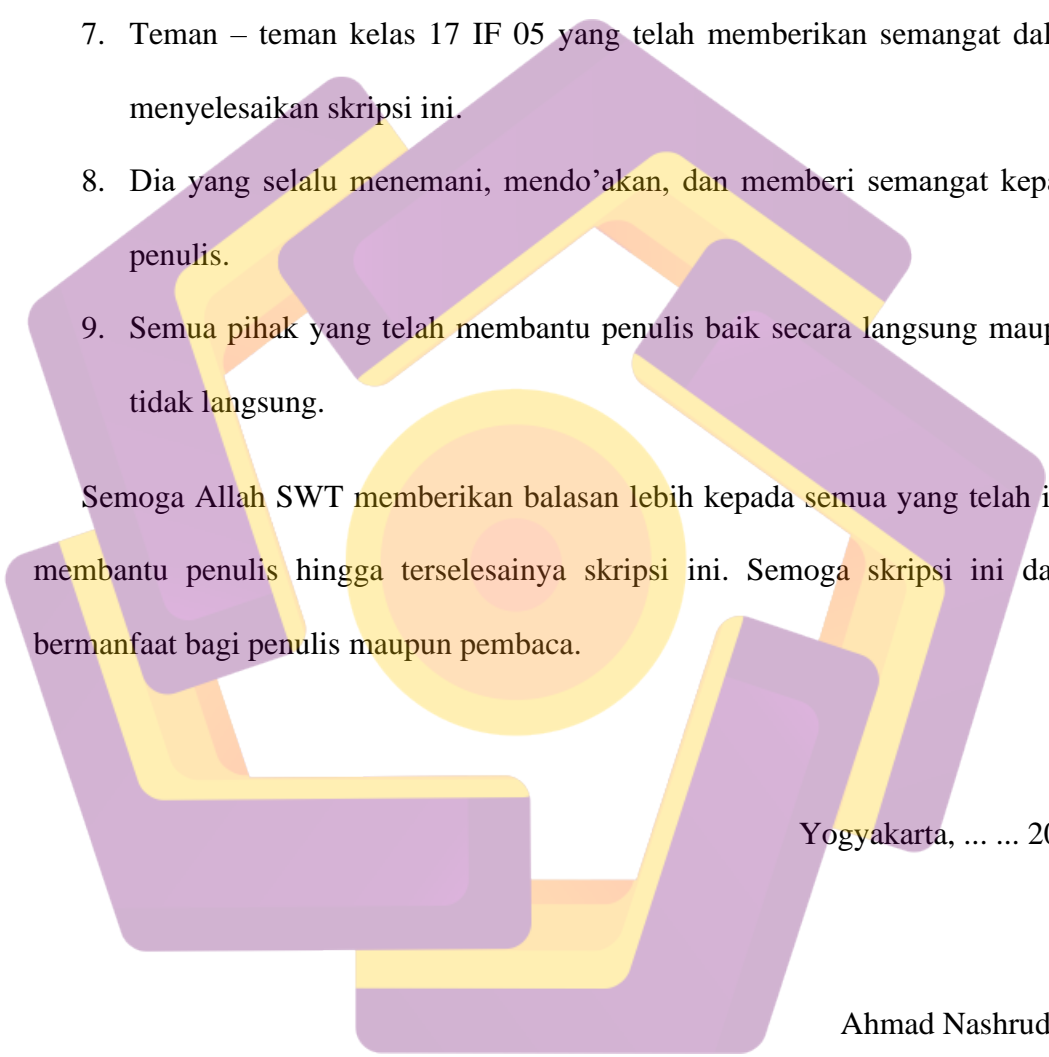
KATA PENGANTAR

Puji dan syukur kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, dan hidayah-Nya kepada kita semua serta shalawat dan salam penulis curahkan kepada junjungan kita Nabi Muhammad SAW sehingga penulis dapat menyelesaikan skripsi yang berjudul “Analisis Prediksi Awal Penyakit Diabetes Mellitus menggunakan Algoritma Naive Bayes”.

Skripsi ini disusun guna menyelesaikan studi jenjang Strata Satu (S1) pada program studi Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga sebagai bukti bahwa mahasiswa telah memenuhi salah satu syarat untuk memperoleh gelar Sarjana Komputer.

Dalam penyusunan skripsi ini, penulis menyadari masih banyak kekurangan dan jauh dari kata sempurna. Selain itu dengan terselesaikannya skripsi ini, penulis ingin menyampaikan rasa terima kasih kepada:

1. Allah SWT atas segala berkah dan ridho-Nya penulis dapat menyelesaikan skripsi ini dengan baik.
2. Nabi Muhammad SAW yang selalu menjadi panutan dan suri tauladan.
3. Kedua orang tua dan keluarga penulis sebagai sosok luar biasa yang selalu memberikan dukungannya dalam bentuk apapun.
4. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.

- 
5. Ibu Rumini, M.Kom, selaku dosen pembimbing yang telah memberikan bimbingan, arahan, dan masukan kepada penulis.
 6. Seluruh staff pengajar dan karyawan Universitas AMIKOM Yogyakarta yang telah memberikan ilmu yang bermanfaat.
 7. Teman – teman kelas 17 IF 05 yang telah memberikan semangat dalam menyelesaikan skripsi ini.
 8. Dia yang selalu menemani, mendo'akan, dan memberi semangat kepada penulis.
 9. Semua pihak yang telah membantu penulis baik secara langsung maupun tidak langsung.

Semoga Allah SWT memberikan balasan lebih kepada semua yang telah ikut membantu penulis hingga terselesainya skripsi ini. Semoga skripsi ini dapat bermanfaat bagi penulis maupun pembaca.

Yogyakarta, 2021

Ahmad Nashruddin

DAFTAR ISI

JUDUL	i
PERSETUJUAN	ii
PENGESAHAN	iii
PERNYATAAN.....	iii
MOTTO	v
PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xiv
INTISARI.....	xvii
<i>ABSTRACT</i>	xviii
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	3
1.4. Maksud dan Tujuan Penelitian	3
1.5. Manfaat Penulisan	4

1.6.	Metode Penelitian.....	4
1.6.1.	Metode Pengumpulan Data.....	4
1.6.2.	Metode Analisis	5
1.6.3.	Metode Perancangan	5
1.6.4.	Metode Implementasi.....	5
1.6.5.	Metode Pengujian.....	5
1.7.	Sistematika Penulisan.....	5
BAB II LANDASAN TEORI.....		7
2.1.	Kajian Pustaka.....	7
2.2.	Dasar Teori.....	9
2.2.1.	<i>Data Mining</i>	9
2.2.2.	<i>Diabetes Mellitus</i>	9
2.2.3.	<i>Naïve Bayes</i>	13
2.2.4.	<i>Laplace Correction</i>	15
2.2.5.	Transformasi Data.....	15
2.2.6.	<i>Confusion Matrix</i>	16
2.2.7.	<i>Flowchart</i>	17
2.2.8.	<i>Hypertext Preprocessor (PHP)</i>	18
2.2.9.	MySQL.....	18
2.2.10.	XAMPP.....	19

2.2.11.	Visual Studio Code	20
BAB III ANALISIS DAN PERANCANGAN		21
3.1.	Analisis Kebutuhan	21
3.1.1.	Analisis Kebutuhan Fungsional	21
3.1.2.	Analisis Kebutuhan Non Fungsional	21
3.2.	<i>Flowchart</i> Proses	22
3.3.	Analisis Prediksi dengan Algoritma Naïve Bayes.....	23
3.3.1.	Transformasi Data.....	25
3.3.2.	Naïve Bayes	27
3.3.3.	<i>Laplace Correction</i>	30
3.4.	Evaluasi	31
3.5.	Analisis Kebutuhan Data.....	34
3.5.1.	Perancangan Basis Data	34
3.5.2.	Perancangan <i>Dataset</i>	37
3.6.	Perancangan Antarmuka Pengguna.....	38
3.6.1.	Perancangan Antarmuka Halaman Beranda.....	38
3.6.2.	Perancangan Antarmuka Halaman Informasi	39
3.6.3.	Perancangan Antarmuka Halaman Tentang.....	40
3.6.4.	Perancangan Antarmuka Halaman Diagnosis.....	40
3.6.5.	Perancangan Antarmuka Halaman <i>Login Admin</i>	42

3.6.6.	Perancangan Antarmuka Halaman <i>Dashboard Admin</i>	43
3.6.7.	Perancangan Antarmuka Halaman <i>Data Admin</i>	44
3.6.8.	Perancangan Antarmuka Halaman <i>Data Training</i>	45
3.6.9.	Perancangan Antarmuka Halaman <i>Data Testing</i>	46
3.6.10.	Perancangan Antarmuka Halaman Hitung Naïve Bayes	47
3.6.11.	Perancangan Antarmuka Halaman Hasil Prediksi	48
3.6.12.	Perancangan Antarmuka Halaman Performa.....	48
BAB IV IMPLEMENTASI DAN PEMBAHASAN		50
4.1.	Deskripsi Implementasi	50
4.2.	Implementasi Pembuatan Basis Data	50
4.3.	Implementasi <i>Dataset</i>	53
4.4.	Pengujian Data	54
4.4.1.	Naïve Bayes	55
4.4.2.	Evaluasi	61
4.5.	Implementasi Antarmuka	65
BAB V PENUTUP.....		75
5.1.	Kesimpulan.....	75
5.2.	Saran.....	75
DAFTAR PUSTAKA		77

DAFTAR TABEL

Tabel 2. 1 Perbandingan penelitian	8
Tabel 2. 2 <i>Confusion matrix</i>	16
Tabel 3. 1 Model data <i>training</i>	24
Tabel 3. 2 Model data <i>training</i> setelah transformasi data	26
Tabel 3. 3 Model data <i>testing</i>	27
Tabel 3. 4 Perhitungan probabilitas	27
Tabel 3. 5 Perhitungan probabilitas setelah <i>laplace correction</i>	30
Tabel 3. 6 Contoh hasil perbandingan prediksi dengan aktual	31
Tabel 3. 7 Model <i>confusion matrix</i>	32
Tabel 3. 8 Struktur tabel <i>admin</i>	34
Tabel 3. 9 Struktur tabel data <i>training</i>	34
Tabel 3. 10 Struktur tabel data <i>testing</i>	35
Tabel 3. 11 Struktur tabel hasil prediksi	36
Tabel 4. 1 <i>Confusion matrix</i> pengujian pertama	62
Tabel 4. 2 <i>Confusion matrix</i> pengujian kedua	63
Tabel 4. 3 <i>Confusion matrix</i> pengujian ketiga	64

DAFTAR GAMBAR

Gambar 3. 1 <i>Flowchart</i> proses	22
Gambar 3. 3 Dataset untuk <i>training</i> dan <i>testing</i>	38
Gambar 3. 4 Rancangan antarmuka halaman beranda	39
Gambar 3. 5 Rancangan antarmuka halaman informasi	39
Gambar 3. 6 Rancangan antarmuka halaman tentang	40
Gambar 3. 7 Rancangan antarmuka halaman diagnosis	41
Gambar 3. 8 Rancangan antarmuka halaman proses diagnosis	41
Gambar 3. 9 Rancangan antarmuka halaman hasil diagnosis	42
Gambar 3. 10 Rancangan antarmuka halaman <i>login admin</i>	43
Gambar 3. 11 Rancangan antarmuka halaman <i>dashboard admin</i>	44
Gambar 3. 12 Rancangan antarmuka halaman data <i>admin</i>	45
Gambar 3. 13 Rancangan antarmuka halaman data <i>training</i>	46
Gambar 3. 14 Rancangan antarmuka halaman data <i>testing</i>	47
Gambar 3. 15 Rancangan antarmuka halaman hitung naïve	47
Gambar 3. 16 Rancangan antarmuka halaman hasil prediksi	48
Gambar 3. 17 Rancangan antarmuka halaman performa	49
Gambar 4. 1 Pembuatan basis data	51
Gambar 4. 2 Data <i>training</i>	54
Gambar 4. 3 Data <i>testing</i>	54
Gambar 4. 4 <i>Script import data testing</i>	55
Gambar 4. 5 <i>Script perhitungan prior probability</i>	56

Gambar 4. 6 <i>Script</i> perhitungan <i>likelihood</i> umur	56
Gambar 4. 7 <i>Script</i> perhitungan <i>likelihood</i> jenis kelamin	56
Gambar 4. 8 <i>Script</i> perhitungan <i>likelihood polyuria</i>	57
Gambar 4. 9 <i>Script</i> perhitungan <i>likelihood polydipsia</i>	57
Gambar 4. 10 <i>Script</i> perhitungan <i>likelihood</i> bb turun	57
Gambar 4. 11 <i>Script</i> perhitungan <i>likelihood</i> keletihan	57
Gambar 4. 12 <i>Script</i> perhitungan <i>likelihood polyphagia</i>	57
Gambar 4. 13 <i>Script</i> perhitungan <i>likelihood</i> iritasi genital	58
Gambar 4. 14 <i>Script</i> perhitungan <i>likelihood</i> pandangan kabur	58
Gambar 4. 15 <i>Script</i> perhitungan <i>likelihood</i> gatal	58
Gambar 4. 16 <i>Script</i> perhitungan <i>likelihood</i> mudah marah	58
Gambar 4. 17 <i>Script</i> perhitungan <i>likelihood</i> sembuh lambat	58
Gambar 4. 18 <i>Script</i> perhitungan <i>likelihood partial paresis</i>	59
Gambar 4. 19 <i>Script</i> perhitungan <i>likelihood</i> otot kaku	59
Gambar 4. 20 <i>Script</i> perhitungan <i>likelihood</i> rambut rontok	59
Gambar 4. 21 <i>Script</i> perhitungan <i>likelihood</i> obesitas	59
Gambar 4. 22 <i>Script</i> perhitungan <i>likelihood probability</i>	60
Gambar 4. 23 <i>Script</i> perhitungan <i>posterior probability</i>	60
Gambar 4. 24 <i>Script</i> perbandingan nilai <i>prior probability</i>	60
Gambar 4. 25 <i>Script</i> pelabelan <i>confusion matrix</i>	61
Gambar 4. 26 <i>Script</i> perhitungan total label	61
Gambar 4. 27 <i>Script</i> perhitungan <i>confusion matrix</i>	61
Gambar 4. 28 Halaman beranda	66

Gambar 4. 29 Halaman informasi	67
Gambar 4. 30 Halaman tentang.....	68
Gambar 4. 31 Halaman diagnosis	69
Gambar 4. 32 Halaman proses diagnosis	69
Gambar 4. 33 Halaman hasil diagnosis.....	70
Gambar 4. 34 Halaman <i>login admin</i>	70
Gambar 4. 35 Halaman <i>dashboard admin</i>	71
Gambar 4. 36 Halaman data <i>admin</i>	71
Gambar 4. 37 Halaman data <i>training</i>	72
Gambar 4. 38 Halaman data <i>testing</i>	72
Gambar 4. 39 Halaman hitung naive.....	73
Gambar 4. 40 Halaman hasil prediksi.....	74
Gambar 4. 41 Halaman performa.....	74

INTISARI

Kesehatan merupakan hal terpenting dalam kehidupan ini. Jika kesehatan manusia sedang terganggu, maka aktivitas sehari - harinya juga akan terganggu. Salah satu penyakit yang mengganggu aktivitas manusia adalah *Diabetes Mellitus*. Dari data *International Diabetes Federation*, Indonesia menduduki peringkat ke - 6 dunia dengan jumlah penderita *Diabetes Mellitus* lebih dari 10 juta jiwa. *World Health Organization* memprediksi jumlah penderita *Diabetes Mellitus* di Indonesia akan mencapai 21 juta jiwa di tahun 2030. Angka penderita *Diabetes Mellitus* yang semakin meningkat menandakan tingkat kesadaran penduduk Indonesia yang masih rendah. Untuk itulah, diperlukan pemahaman dan pengetahuan tentang gejala-gejala awal *Diabetes Mellitus* guna mencegah secara dini penyakit tersebut. Karena semakin dini penyakit terdeteksi akan semakin besar juga kesempatan untuk sembuh. Untuk melakukan itu, peneliti akan memanfaatkan informasi dari situs *UCI Machine Learning* untuk mendapatkan dataset yang nantinya dapat diproses dengan algoritma Naïve Bayes dalam memprediksi penyakit *Diabetes Mellitus*.

Algoritma Naïve Bayes merupakan salah satu algoritma yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu *class*. Algoritma ini memiliki tingkat akurasi dan kecepatan yang tinggi saat diaplikasikan pada data yang besar. Dengan algoritma Naïve Bayes, akan digunakan untuk memprediksi penyakit *Diabetes Mellitus* berdasarkan gejala - gejala yang dialami. Hasil prediksi akan berupa kelas positif dan negatif.

Pengujian sistem ini dilakukan sebanyak 3 kali. Total jumlah data yang digunakan yaitu sebanyak 500 data. Dari hasil pengujian didapatkan akurasi tertinggi pada pengujian ketiga yaitu dengan akurasi 89 %. Pada pengujian ketiga menggunakan perbandingan data training dan data testing sebesar 80% dan 20% (400 data *training* dan 100 data *testing*).

Kata Kunci: Prediksi, Algoritma Naïve Bayes, Diabetes Mellitus

ABSTRACT

Health is the most important thing in this life. If human health is being disturbed, then their daily activities will also be disrupted. One of the diseases that interfere with human activities is Diabetes Mellitus. Based on data from the International Diabetes Federation, Indonesia is ranked 6th in the world with more than 10 million people with Diabetes Mellitus. The World Health Organization predicts the number of people with Diabetes Mellitus in Indonesia will reach 21 million in 2030. The increasing number of people with Diabetes Mellitus indicates the level of awareness of the Indonesian population is still low. For this reason, it is necessary to understand and know about the early symptoms of Diabetes Mellitus in order to prevent the disease early. Because the earlier the disease is detected, the greater the chance for recovery. To do that, researchers will use information from the UCI Machine Learning website to obtain a dataset that can later be processed with the Naïve Bayes algorithm in predicting Diabetes Mellitus.

Naïve Bayes algorithm is one of the algorithms that can be used to predict the probability of membership of a class. This algorithm has a high level of accuracy and speed when applied to large data. With the Naïve Bayes algorithm, it will be used to predict Diabetes Mellitus based on the symptoms experienced. Prediction results will be positive and negative classes.

Testing of this system was carried out 3 times. The total amount of data used is 500 data. From the test results obtained the highest accuracy in the third test with an accuracy of 89%. In the third test using a comparison of training data and testing data of 80% and 20% (400 training data and 100 testing data).

Keywords: *Prediction, Naïve Bayes Algorithm, Diabetes Mellitus*