

TUGAS AKHIR
SKEMA SKRIPSI
KLASIFIKASI PENYAKIT HEPATITIS MENGGUNAKAN
METODE NAÏVE BAYES



Disusun Oleh :

ATIKAH DYAH CAHYANI

205410085

PROGRAM STUDI INFORMATIKA

PROGRAM SARJANA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA

YOGYAKARTA

2025

TUGAS AKHIR
SKEMA SKRIPSI
KLASIFIKASI PENYAKIT HEPATITIS MENGGUNAKAN METODE
NAÏVE BAYES

Diajukan sebagai salah satu syarat untuk menyelesaikan studi pada



Disusun Oleh

ATIKAH DYAH CAHYANI

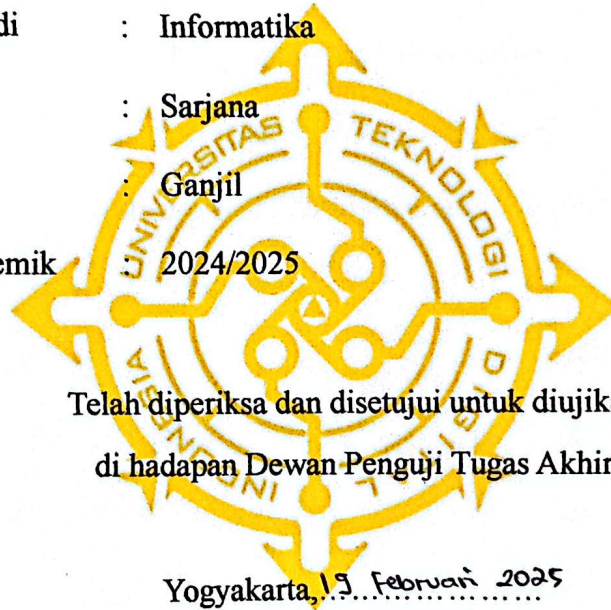
NIM : 205410085

PROGRAM STUDI INFORMATIKA
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA

2025

HALAMAN PERSETUJUAN UJIAN TUGAS AKHIR

Judul : Klasifikasi Penyakit Hepatitis Dengan Menggunakan Metode Naïve Bayes
Nama : Atikah Dyah Cahyani
NIM : 205410085
Program Studi : Informatika
Program : Sarjana
Semester : Ganjil
Tahun Akademik : 2024/2025



Telah diperiksa dan disetujui untuk diujikan
di hadapan Dewan Penguji Tugas Akhir

Yogyakarta, 15 Februari 2025

Dosen Pembimbing,

Ariesta Damayanti, S.Kom.,M.Cs.

NIDN : 0020047801

HALAMAN PENGESAHAN

**KLASIFIKASI PENYAKIT HEPATITIS MENGGUNAKAN METODE
NAÏVE BAYES**

**Telah dipertahankan di depan Dewan Penguji dan dinyatakan diterima
untuk memenuhi sebagian persyaratan guna memperoleh**



Dewan Penguji

1. Dini Fakta Sari, S.T., M.T

2. Ariesta Damayanti, S.Kom., M.Cs.

NIDN

0507108401

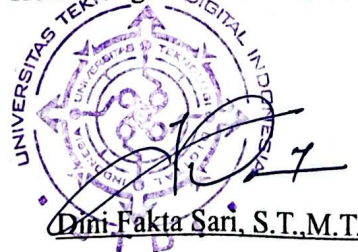
0020047801

Tandatangan



Mengetahui

Ketua Program Studi Informatika



Dini Fakta Sari, S.T., M.T.

NIDN : 0507108401

PERNYATAAN KEASLIAN TUGAS AKHIR

Dengan ini saya menyatakan bahwa naskah Tugas Akhir ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 19 Februari 2025



Atikah Dyah Cahyani

NIM: 205410085

HALAMAN PERSEMBAHAN

Dengan menyebut nama Allah yang maha pengasih lagi maha penyayang. Saya mempersembahkan segala sesuatu yang tertuang pada karya ilmiah ini untuk :

1. Kedua orang tua tercinta, Ayahanda Suyadi dan Ibunda Prihatin, dua orang paling berjasa dalam hidup saya, yang tidak ada henti – hentinya memberikan kasih sayang dengan penuh cinta dan selalu memberikan motivasi. Terimakasih selalu berjuang untuk kehidupan saya. Terimakasih untuk semuanya berkat do'a dan dukungan dari ayah dan mamah saya bisa berada dititik ini. Terimakasih ayah, mamah putri kecilmu ini sudah tumbuh besar dan siap melanjutkan mimpi yang lebih tinggi lagi. Sehat selalu dan hiduplah lebih lama lagi, ayah dan mamah harus selalu ada disetiap perjalanan dan pencapaian hidup saya.
2. Seluruh keluarga besar terimakasih atas segala doa yang kalian panjatkan sehingga saya bisa sampai berada dititik ini.
3. Sahabat penulis, Ainun Annisa, Debora Holi Laiku, Sulastri terimakasih selalu kebersamaan penulis dari awal perkuliahan sampai tugas akhir. Tanpa bantuan, dukungan, dan motivasi yang kalian berikan perjalanan ini terasa berat untuk dilalui. *Thank you for being a best friend and see you on top, guys!*
4. *The author Favorite Person*, terimakasih telah menjadi support sytem dalam segala hal, yang menemani penulis dari awal perkuliahan sampai saat ini,

tempat berbagi suka cita, memberi dukungan, kasih sayang dan semangat, serta perhatian kepada penulis dalam penyusunan skripsi dari awal hingga selesainya skripsi ini.

5. Diri saya sendiri, Atikah Dyah Cahyani. Apresiasi sebesar – besarnya karena telah bertanggung jawab atas apa yang telah dimulai. Terimakasih karena terus selalu berusaha dan tidak menyerah,serta senantiasa menikmati setiap prosesnya yang bisa dibbilang tidak mudah. Terimakasih sudah bertahan.
6. Kepada seluruh rekan - rekan prodi informatika angkatan 2020, terimakasih atas suka dan duka yang telah kita lalui, semoga kelak kita semua menjadi orang yang sukses.

PRAKATA

Alhamdulillah, dengan mengucapkan puji syukur saya panjatkan kehadiran Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Penyakit Hepatitis dengan Menggunakan Metode Naïve Bayes”. Penulisan skripsi ini dibuat sebagai salah satu syarat kelulusan program Sarjana Komputer dari Program Studi Informatika Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia. Penulis menyadari bahwa tanpa bimbingan dan dorongan dari semua pihak, maka penulisan ini tidak akan berjalan lancar. Oleh karena itu pada kesempatan ini, izinkanlah penulis menyampaikan ucapan terimakasih kepada :

1. Ibu Sri Redjeki, S.Si., M.Kom, Ph.D selaku Rektor Universitas Teknologi Digital Indonesia.
2. Dr. Bambang Purnomosidi DP, S.E. Akt., S.Kom., MMSI. selaku Dekan Fakultas Teknologi Informasi Universitas Teknologi Digital Indonesia.
3. Ibu Dini Fakta Sari, S.T.,M.T dan Ibu Femi Dwi Astuti,S.Kom., M.Cs. selaku Ketua dan Sekretaris Program Studi Informatika Universitas Teknologi Digital Indonesia.
4. Ibu Ariesta Damayanti, S.Kom., M.Cs. selaku dosen pembimbing yang telah memberikan arahan dan bimbingan selama pengerjaan skripsi.
5. Seluruh dosen Fakultas Teknologi Informasi Program Studi Informatika Universitas Teknologi Digital Indonesia yang telah memberikan ilmu dan pengalamannya serta mendidik penulis selama masa kuliah.

6. Seluruh Rekan – Rekan Mahasiswa/I Informatika angkatan 2020.

Terimakasih atas kenangan dan pengalamannya. Semoga kelak kita akan menjadi orang yang sukses.

Akhir kata, penulis dapat menyadari tanpa Ridho dan pertolongan dari Allah SWT, serta bantuan,dukungan, motivasi dari segala pihak skripsi ini tidak dapat diselesaikan. Kepada semua pihak yang telah memberikan bantuan dalam penulisan ini, penulisa ucapkan banyak terimakasih dan semoga Allah SWT membalas segera kebaikan kalian. *Aamiin Yarabbal'alam.*

Yogyakarta,.....

Atikah Dyah Cahyani

NIM : 205410085

DAFTAR ISI

TUGAS AKHIR.....	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN.....	iv
HALAMAN PERSEMBAHAN	v
PRAKATA.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xiii
INTISARI.....	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Ruang Lingkup.....	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	6
2.1 Tinjauan Pustaka	6
2.2 Dasar Teori.....	11
2.2.1 Hepatitis	11
2.2.2 Data Mining	12
2.2.3 Algoritma Naïve Bayes	13
2.2.4 Confusion Matrix	14
2.2.4 Flask.....	15
BAB III METODE PENELITIAN.....	17
3.1 Data	17
3.2 Peralatan.....	18
3.2.1 Kebutuhan Perangkat Keras	18

3.2.2	Kebutuhan Perangkat Lunak	18
3.2.3	Kebutuhan Input	19
3.2.4	Kebutuhan Proses	19
3.2.5	Kebutuhan Output	23
3.3	Analisis dan Rancangan Sistem	23
3.3.1	Use Case Diagram	23
3.3.2	Perancangan Antarmuka	23
BAB IV IMPLEMENTASI DAN PEMBAHASAN		27
4.1	Hasil	27
4.1.1	Pengumpulan Data	27
4.1.2	Preprocessing Data	28
4.1.3	Membagi Data Training dan Data Testing	39
4.1.4	Algoritma Naïve Bayes	42
4.1.5	Confusioun Matrix	46
4.2	Impelementasi Aplikasi	50
4.2.1	Antarmuka Aplikasi Web	51
BAB V PENUTUP		55
5.1	Kesmipulan	55
5.2	Saran	55
DAFTAR PUSTAKA		57
Lampiran		58

DAFTAR GAMBAR

Gambar 3. 1 Flowchart Data Training	20
Gambar 3. 2 Flowchart Data Testing	22
Gambar 3. 3 Use Case Diagram	23
Gambar 3. 4 Halaman Home.....	24
Gambar 3. 5 Halaman Prediksi	25
Gambar 3. 6 Hasil Prediksi	26
Gambar 3. 7 Halaman About.....	26
Gambar 4. 1 Kode Menampilkan Data	27
Gambar 4. 2 Hasil Menampilkan Data.....	28
Gambar 4. 3 Kode Mengecek Missing Value.....	29
Gambar 4. 4 Hasil Mengecek Missing Value	29
Gambar 4. 5 Kode Mencari Nilai Modus	30
Gambar 4. 6 Nilai Modus Kolom Steroid	30
Gambar 4. 7 Nilai Modus Kolom Fatigue.....	30
Gambar 4. 8 Nilai Modus Kolom Malaise	31
Gambar 4. 9 Nilai Modus Kolom Anorexia	31
Gambar 4. 10 Nilai Modus Kolom Liver Big	31
Gambar 4. 11 Nilai Modus Kolom Liver Firm	31
Gambar 4. 12 Nilai Modus Kolom Spleen Palpable	32
Gambar 4. 13 Nilai Modus Kolom Spiders.....	32
Gambar 4. 14 Nilai Modus Kolom Ascites	32
Gambar 4. 15 Nilai Modus Kolom Varices	32
Gambar 4. 16 Gambar Mencari Nilai Median	33
Gambar 4. 17 Hasil Mencari Nilai Median	34
Gambar 4. 18 Kode Menginputkan Nilai Modus.....	34
Gambar 4. 19 Kode Menginputkan Nilai Median.....	35
Gambar 4. 20 Kode Memanggil Data Frame	36
Gambar 4. 21 Missing Value Sebelum diisi Nilai Modus dan Nilai Median	36
Gambar 4. 22 Missing Value Setelah diisi Nilai Modus Dan Nilai Median	37
Gambar 4. 23 Kode Memeriksa Data Duplikat.....	37
Gambar 4. 24 Hasil Memeriksa Data Duplikat.....	37
Gambar 4. 25 Kode Memeriksa Data.....	37
Gambar 4. 26 Hasil Memeriksa Data	38
Gambar 4. 27 Kode Label Encoder.....	38
Gambar 4. 28 Hasil Label Encoding	39
Gambar 4. 29 Kode Mendefinisikan Atribut dan Target	40
Gambar 4. 30 Kode Membagi Data Training dan Data Testing.....	40
Gambar 4. 31 Kode Menampilkan Data Training dan Data Testing	41
Gambar 4. 32 Hasil Data Traing 75%	41
Gambar 4. 33 Hasil Data Testing 25%	42
Gambar 4. 34 Kode Melatih Model	42

Gambar 4. 35 Kode Menentukan Hasil Prediksi.....	43
Gambar 4. 36 Hasil Prediksi	43
Gambar 4. 37 Kode Menghitung Probabilitas	43
Gambar 4. 38 Hasil Menghitung Probabilitas.....	44
Gambar 4. 39 Kode Mengambil Data Secara Random.....	44
Gambar 4. 40 Kode Mencetak Perbandingan Antara Label Asli dan Prediksi	45
Gambar 4. 41 Hasil Perbandingan Label	45
Gambar 4. 42 Kode Confusion Matrix.....	47
Gambar 4. 43 Hasil Confusion Matrix	47
Gambar 4. 44 Kode Classification Report	48
Gambar 4. 45 Hasil Classification Report	49
Gambar 4. 46 Tampilan Terminal.....	51
Gambar 4. 47 Halaman Home.....	51
Gambar 4. 48 Halaman Prediksi	52
Gambar 4. 49 Halaman About.....	53
Gambar 4. 50 Memasukkan Nilai Input Prediksi.....	53
Gambar 4. 51 Halaman Result Prediksi	54

DAFTAR TABEL

Tabel 2. 1 Perbandingan dengan Penelitian Sebelumnya	6
Tabel 2. 2 Confusion Matrix	14
Tabel 3. 1 Deskripsi Variable Dataset	17

INTISARI

Salah satu penyakit yang berkembang pada saat ini adalah penyakit hepatitis. Penyakit hepatitis merupakan penyakit peradangan hati karena infeksi virus yang menyerang dan menyebabkan kerusakan pada sel – sel dan fungsi organ hati. Dengan memanfaatkan teknologi yang semakin berkembang pesat ini, para peneliti mencoba untuk memprediksi atau mengklasifikasi apakah pasien tersebut menderita penyakit hepatitis berdasarkan dengan hasil tes yang sudah dilakukan sebelumnya.

Dengan menggunakan metode Naive Bayes untuk melakukan klasifikasi penyakit hepatitis dengan menggunakan bahasa pemrograman Python. Jumlah data yang digunakan sebanyak 155 dataset yang diambil dari website Kaggle, dengan menggunakan pembagian data 75% untuk data latih dan 25% untuk data uji. Hasil *accuracy* yang didapat menggunakan evaluasi model sebesar 84.61%, *precision* sebesar 96%, *recall* sebesar 83% dan *f1-score* sebesar 89%. Dalam hal ini menunjukkan bahwa model memiliki performa yang cukup baik dalam mengklasifikasi suatu data.

Kata Kunci : *Data Mining, Klasifikasi, Hepatitis, Python*

ABSTRACT

One of the diseases that is currently developing is hepatitis. Hepatitis is an inflammatory disease of the liver caused by a viral infection that attacks and causes damage to the cells and function of the liver. By utilizing technology that is increasingly developing rapidly, researchers are trying to predict or classify whether the patient is suffering from hepatitis based on the results of tests that have been carried out previously.

Using the Naive Bayes method to classify hepatitis diseases using the Python programming language. The amount of data used was 155 datasets taken from the Kaggle website, using a data division of 75% for training data and 25% for test data. The accuracy results obtained using model evaluation were 84.61%, precision 96%, recall 83% and f1-score 89%. In this case, it shows that the model has quite good performance in classifying data.

Keywords: *Data Mining, Classification, Hepatitis, Python*