

TUGAS AKHIR

**PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK
KLASIFIKASI PENYAKIT LIVER**



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2021**

HALAMAN JUDUL

PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENYAKIT LIVER

Disusun untuk Melegkapi dan Memenuhi Syarat Kelulusan Program Sastra 1
Program Studi Teknik Informatika Fakultas Teknik
Universitas Muhammadiyah Jember



OLEH:

EGA YUSNI HABIBIE

1610651037

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH JEMBER
2021**

HALAMAN PERSETUJUAN

PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENYAKIT LIVER

Oleh:

Ega Yusni Habibie

1610651037

Telah disetujui bahwa Laporan Tugas Akhir ini untuk diajukan pada Sidang
Tugas Akhir sebagai salah satu syarat kelulusan dan mendapatkan gelar
Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

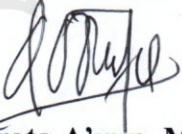
Disetujui oleh:

Pembimbing I


Hardian Oktavianto, M.Kom.

NIDN. 0722108105

Pembimbing II


Qurrota A'yun, M.Pd.

NIDN. 0703069002

HALAMAN PENGESAHAN

PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENYAKIT LIVER

Oleh:

Ega Yusni Habibie

1610651037

Telah mempertanggung jawabkan Laporan Tugas Akhirnya pada Sidang Tugas Akhir tanggal ... sebagai salah satu syarat kelulusan dan mendapatkan gelar Sarjana Komputer (S.Kom)

di

Universitas Muhammadiyah Jember

Disetujui Oleh,

Penguji I

Ir. Dewi Lusiana Pater, M.T.

NIDN. 0712086702

Penguji II

Wiwik Suharso, S.Kom., M.Kom.

NIDN. 0006097601

Pembimbing I

Hardian Oktavianto, M.Kom.

NIDN. 0722108105

Pembimbing II

Qurrota A'yun, M.Pd.

NIDN. 0703069002

Mengesahkan,

Dekan Fakultas Teknik



Dr. Nanang Saiful Rizal, S.T., M.T.

NIDN. 0705047806

Mengetahui,

Ketua Program Studi Teknik
Informatika

Triawan Adi Cahyanto, M.Kom.

NIDN. 0702098804

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

NAMA : EGA YUSNI HABIBIE

NIM : 1610651037

INSTITUSI : Strata-1 Program Studi Teknik Informatika Fakultas Teknik,
Universitas Muhammadiyah Jember.

Menyatakan bahwa Tugas Akhir yang berjudul "**Penerapan Algoritma K-Nearest Neighbor untuk Klasifikasi Penyakit Liver**" bukan merupakan Tugas Akhir orang lain baik sebagian maupun keseluruhan kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar penulis bersedia mendapatkan sanksi akademik.

Jember,...



Menyatakan,
Ega Yusni Habibie

NIM. 1610651037

UNGKAPAN TERIMA KASIH

Puji syukur kepada Allah S.W.T atas rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan laporan tugas akhir ini. Atas segala upaya, bimbingan dan arahan dari semua pihak, penulis mengucapkan terimakasih sebesar-besarnya kepada :

1. Allah SWT berkat segala ridho, rahmat dan hidayah-Nya penulis dapat menyelesaikan segala urusan dalam menyusun laporan Tugas Akhir dan diberikan kesempatan mendapatkan gelar Sarjana Komputer.
2. Bapak Dr. Nanang Saiful Rizal, S.T., M.T selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember.
3. Bapak Triawan Adi Cahyanto, M.Kom selaku Kepala Program Studi Teknik Informatika di Universitas Muhammadiyah Jember.
4. Bapak Hardian Oktavianto, S.Kom selaku dosen pembimbing 1 tugas akhir yang telah memberikan bimbingan dari awal sampai akhir dengan penuh kesabaran sehingga tugasakhir ini dapat terselesaikan.
5. Ibu Qurrota A'yun, M.Pd selaku dosen pembimbing 2 tugas akhir yang telah memberikan bimbingan dari awal sampai akhir dengan penuh kesabaran sehingga tugas akhir ini dapat terselesaikan.
6. Ibu Ir. Dewi LusianaPater, M.T selaku dosen penguji 1 yang telah memberikan kritik, saran dan masukkan yang membangun dalam penelitian ini.
7. Bapak Wiwik Suharso, S.Kom., M.Kom selaku dosen penguji 2 yang telah memberikan kritik, saran dan masukkan yang membangun dalam penelitian ini.
8. Kedua orang tua saya Bapak Jayusman .S.P. dan Ibu Nur Aini, terimakasih atas *support* dan doa nya dalam penyusunan laporan tugas akhir ini.
9. Teman-teman seperjuangan yang senasib dan setanah air, baik “GRUP TIMBANGANE MENENG, GRUP PKL, KONTRAKAN B26 , KOS KARIMATA, KONCO KENTEL” terimakasih atas doa dan dukungannya selama ini, dan terima kasih atas kebersamaannya.

MOTTO

“Kau tak dapat meraih sesuatu dalam hidup tanpa pengorbanan sekecil apapun”

Shakira

“Kesuksesan tidak terwujud karena tidak pernah membuat kesalahan. Tapi, tidak pernah membuat kesalahan yang sama untuk kedua kalinya”

George Bernanard Shaw

“Lakukan yang terbaik hingga aku tidak bisa menyalahkan diri sendiri atas semua yang terjadi”

Magdalena Neurer

KATA PENGANTAR

Alhamdulillah segala puji bagi Allah yang senantiasa memberikan nikmat yang sangat luar biasa dalam hidup ini kepada penukisan berupa nikmat untuk selalu bersujud dan meminta ampunan kepada-Nya. Serta berkat rahmat-Nya penulis diberikan kemudahan dalam menyelesaikan studi di kampus Universitas Muhammadiyah Jember.

Atas segala upaya, bimbingan dan arahan dari semua pihak, penulisan mengucapkan banyak terimakasih. Dalam kesempatan ini penulis mengucapkan terima kasih sebanyak-banyak kepada :

1. Allah SWT atas segala rahmat, nikmat dan hidayah-Nya. Terimakasih atas kehendak-Mu telah memberikan kesempatan ini untuk mendapatkan gelar Sarjana Komputer.
2. Bapak Nanang Saiful Rizal, S.T., M.T., selaku Dekan Fakultas Teknik Universitas Muhammadiyah Jember
3. Bapak Triawan Adi Cahyanto, M.Kom, selaku Ketua Program Studi Teknik Informatika Universitas Muhammadiyah Jember.
4. Bapak Hardian Oktavianto, M.Kom., selaku dosen pembimbing I dan Ibu Qurrota A'yun, M.Pd, selaku dosen pembimbing II yang telah meluangkan waktu untuk memberi arahan dan membimbing penulis dalam menyelesaikan Tugas Akhir ini.
5. Ibu Dewi Lusiana,Ir.MT, selaku dosen penguji 1 dan Bapak Wiwik Suharso S.Kom.,M.Kom selaku dosen penguji II yang telah memberikan saran dan masukan dalam penyelesaian Tugas Akhir ini.
6. Kepada Bapak dan Ibu Dosen Program Studi Teknik Informatika Universitas Muhammadiyah Jember yang telah memberikan ilmunya kepada penulis.

7. Kepada Bapak dan Ibu Staf dan Laborab Program Studi Teknik Informatika Universitas Muhammadiyah Jember yang telah membantu penulis dalam melaksanakan tugas akhir.
8. Kepada orang tua penulis Bapak Jayusman S.P. dan Ibu Nur Aini yang selalu mendoakan, memberikan motivasi, dukungan dan memberikan kasih sayang sehingga penulis dapat menyelesaikan tugas akhir ini.
9. Kepada teman-teman yang telah membantu untuk memberikan semangat, dukungan, bantuan, nasihat, saran, dan doa ke penulis untuk menyelesaikan tugas akhir ini.
10. Serta kepada teman-teman Program Studi Teknik Informatika Angkatan 2016 yang telah memberikan doa, dukungan, dan bantuannya.

Dan untuk semuanya, semoga Allah menunjukkan kepada kita pada jalan yang telah diridhoi oleh-Nya, diberikan ilmu yang barokah, serta diampuni dosa dosa kita, dan senantiasa tercurah limpahkan ramat, taufiq dan hodayah-Nya kepada kita semua.

Aamin... Yaa Robbal Alamiin...

PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENYAKIT LIVER

Ega Yusni Habibie¹, Hardian Oktavianto²,Qurrota A'yun³

Program Studi Teknik Informatika Informatika, Fakultas Teknik, Universitas
Muhammadiyah Jember

egayusni001@gmail.com, hardian@unmuhjember.ac.id,

qurotta.ayun@unmuhjember.ac.id

ABSTRAK

Liver merupakan organ yang sangat penting dalam tubuh manusia. Penyakit liver disebab oleh berbagai faktor yang merusak hati. Permasalahan yang terjadi yaitu sulitnya mengenali penyakit liver sejak dini, bahkan saat penyakit liver ini sudah menyebar masih sulit untuk diteksi. tujuan dari penelitian tugas akhir untuk mengetahui berapa tingkat akurasi, presisi dan recall. pada diagnosa penyakit liver menggunakan algoritma *K-Nearst Neighbor*. Algoritma *K-Nearst Neighbor* merupakan sebuah metode untuk melakukan klasifikasi terhadap objek yang berdasarkan dari data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Pada penelitian ini diterapkan metode untuk klasifikasi penyakit liver menggunakan algoritma K-Nears Neighbor karena dianggap cukup fleksibel. Data yang digunakan adalah data Indian Liver Patients Dataset (ILDP) yang diambil dari website UCI Machine Learning. Tujuan dari penelitian ini untuk mengetahui berapa tingkat akurasi, presisi dan recall pada diagnosa penyakit liver menggunakan algoritma K-Nearst Neighbor. Hasil yang didapatkan dari pengujian algoritma *K-Nearst Neighbor* dengan nilai akurasi tertinggi sebesar 67,20% , presisi tertinggi sebesar 59,36%, dan recall tertinggi sebesar 58,57% . Jadi, Alogaritma *K-Nearest Neighbor*, cukup akurat di atas 50% untuk mengklasifikasi data pasien penyakit liver.

Kata Kunci : *Klasifikasi, Penyakit Liver, K-Nearest Neighbor*

APPLICATION OF K-NEAREST NEIGHBOR ALGORITHM FOR LIVER DISEASE CLASSIFICATION

Ega Yusni Habibie¹, Hardian Oktavianto², Qurotta A'yun³

Informatics Engineering Study Program, Faculty of Engineering,

University of Muhammadiyah Jember

egayusni001@gmail.com, hardian@unmuahjember.ac.id,

qurotta.ayun@unmuahjember.ac.id

ABSTRACT

The liver is a very important organ in the human body. Liver disease is caused by various factors that damage the liver. The problem that occurs is the difficulty in recognizing liver disease from an early age, even when the liver disease has spread, it is still difficult to detect. The purpose of this final project research is to determine the level of accuracy, precision and recall. in diagnosing liver disease using the K-Nearst Neighbor algorithm. The K-Nearst Neighbor algorithm is a method for classifying objects based on learning data that is closest to the object. In this study, a method for classification of liver disease was applied using the K-Nearst Neighbor algorithm because it was considered quite flexible. The data used is the Indian Liver Patients Dataset (ILDP) which is taken from the UCI Machine Learning website. The purpose of this study was to determine the level of accuracy, precision and recall in diagnosing liver disease using the K-Nearst Neighbor algorithm. The results obtained from testing the K-Nearst Neighbor algorithm with the highest accuracy value of 67.20%, the highest precision of 59.36%, and the highest recall of 58.57%. So, the K-Nearest Neighbor Algorithm, is quite accurate above 50% for classifying liver disease patient data.

Keywords: Classification, Liver Disease, K-Nearest Neighbor

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN.....	iv
UNGKAPAN TERIMA KASIH	v
MOTTO	vi
KATA PENGANTAR.....	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL	xv
BAB I PENDAHULUAN 1	
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA 4	
2.1. Penyakit Liver 4	
2.1.1. Definisi Penyakit Liver	4
2.2. <i>Data Mining</i>	7
2.2.1. Pengertian <i>Data Mining</i>	7

2.2.2.	Tahap-Tahap <i>Data Mining</i>	8
2.3.	Klasifikasi	10
2.3.1.	Definisi Klasifikasi	10
2.4.	K-Nearest Neighbor (KNN).....	11
2.5.	<i>K-Fold Cross Validation</i>	12
2.6.	Normalisasi Data.....	12
2.7.	Confusion Matrix	13
2.8.	RapidMiner Studio.....	14

BAB III METODOLOGI PENELITIAN 16

3.1.	Tahap Penelitian.....	16
3.2.	Studi Literatur	16
3.3.	Pengumpulan Data	17
3.4.	Praproses Data	17
3.5.	Perancangan	19
3.5.1.	Flowchart Algoritma K-Nearest Neighbor.....	19
3.6.	Skenario Uji	20
3.7.	Implementasi Algoritma <i>K-nearest Neighbor</i>	21
3.7.1.	Menentukan Data Latih dan Data Uji	21
3.7.2.	Menghitung Jarak Euclidean antar latih dan data uji	23
3.7.3.	Menentukan Nilai K dan Menentukan Mayoritas Data	25

BAB IV HASIL DAN PEMBAHASAN 27

4.1	PengumpulanData.....	27
4.2	Pre-processing Data	27
4.3	Hasil Klasifikasi.....	28
4.3.1	Pengujian menggunakan <i>Kfold 2</i>	28
4.3.2	Pengujian menggunakan <i>Kfold 4</i>	32
4.3.3	Pengujian menggunakan <i>Kfold 5</i>	35

4.3.4	Pengujian menggunakan <i>Kfold 10</i>	38
-------	---	----

BAB V KESIMPULAN DAN SARAN 42

5.1.	Kesimpulan	42
------	------------------	----

5.2.	Saran	42
------	-------------	----

DAFTAR PUSTAKA 43



DAFTAR GAMBAR

Gambar 3. 1 Tahapan Metodologi Penelitian	18
Gambar 3. 2 <i>Flowchart Algoritma K-Nearest Neighbor</i>	19
Gambar 4. 1 Hasil Pengujian <i>Kfold 2 K3</i>	28
Gambar 4. 2 Hasil Pengujian <i>Kfold 2 K5</i>	28
Gambar 4. 3 Hasil Pengujian <i>Kfold 2 K7</i>	29
Gambar 4. 4 Hasil Pengujian <i>Kfold 2 K9</i>	29
Gambar 4. 5 Hasil Pengujian pada <i>Kfold 2</i>	29
Gambar 4. 6 Hasil Pengujian <i>Kfold 4 K3</i>	31
Gambar 4. 7 Hasil Pengujian <i>Kfold 4 K5</i>	31
Gambar 4. 8 Hasil Pengujian <i>Kfold 4 K7</i>	31
Gambar 4. 9 Hasil Pengujian <i>Kfold 4 K9</i>	32
Gambar 4. 10 Hasil Pengujian pada <i>Kfold 4</i>	32
Gambar 4. 11 Hasil Pengujian <i>Kfold 5 K3</i>	34
Gambar 4. 12 Hasil Pengujian <i>Kfold 5 K5</i>	34
Gambar 4. 13 Hasil Pengujian <i>Kfold 5 K7</i>	34
Gambar 4. 14 Hasil Pengujian <i>Kfold 5 K9</i>	35

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	13
Tabel 3. 1 Praproses Data.....	17
Tabel 3. 2 Normalisasi Data.....	18
Tabel 3. 3 Data Latih.....	21
Tabel 3. 4 Data Uji.....	22
Tabel 3. 5 Hasil jarak Euclidean data latih dan data uji.....	24
Tabel 3. 6 Mayoritas Data.....	25
Tabel 3. 7 Hasil Klasifikasi.....	26
Tabel 4. 1 Data Pasien Liver	27
Tabel 4. 2 Normalisasi Data.....	28
Tabel 4. 3 Confuxion Matrix pada Algoritma KNN menggunakan <i>Kfold</i> 2.....	29
Tabel 4. 4 Nilai Akurasi, Presisi dan Recall menggunakan <i>Kfold</i> 2	30
Tabel 4. 5 <i>Confuxion Matrix</i> pada Algoritma KNN menggunakan <i>Kfold</i> 4	32
Tabel 4. 6 Nilai Akurasi, Presisi dan Recall pada <i>Kfold</i> 4	33
Tabel 4. 7 Confuxion Matrix pada Algoritma KNN menggunakan <i>Kfold</i> 5.....	35
Tabel 4. 8 Nilai Akurasi, Presisi dan Recall pada <i>Kfold</i> 5	36
Tabel 4. 9 Confuxion Matrix pada Algoritma KNN menggunakan <i>K-fold</i> 10	39
Tabel 4. 10 Nilai Akurasi, Presisi dan Recall pada <i>Kfold</i> 10.....	40