

SKRIPSI
PENERAPAN K-NEAREST NEIGHBOR UNTUK
KLASIFIKASI PENYAKIT TBC PARU
(Studi Kasus: Puskesmas Bumiayu)



**Diajukan Untuk Memenuhi Satu Syarat Guna Memperoleh
Gelara Sarjana Komputer**

Oleh:
Amira Nabila Zakiyyah
42417029

PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PERADABAN
2023

PERNYATAAN PENULIS

JUDUL : PENERAPAN K-NEAREST NEIGHBOR UNTUK
KLASIFIKASI PENYAKIT TBC PARU
NAMA : AMIRA NABILA ZAKIYYAH
NIM : 42417029

Saya yang bertanda tangan dibawah ini:

Menyatakan dengan sebenar-benarnya bahwa skripsi ini merupakan hasil karya saya sendiri dan bukan merupakan pengambilan tulisan atau pikiran orang lain yang saya akui sebagai tulisan dan pikiran saya, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka, apabila dikemudian hari terbukti atau dapat dibuktikan bahwa skripsi ini hasil plagiat maka saya bersedia menerima sanksi berupa pencabutan gelar akademik yang saya peroleh terkait dengan skripsi ini.

Bumiayu, 4 September 2023

Yang membuat pernyataan


Amira Nabila Zakiyyah

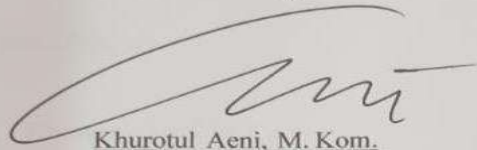
42417029

PERSETUJUAN SKRIPSI

JUDUL : PENERAPAN K-NEAREST NEIGHBOR UNTUK
KLASIFIKASI PENYAKIT TBC PARU (STUDI KASUS
PUSKESMAS BUMIAYU)
NAMA : AMIRA NABILA ZAKIYYAH
NIM : 42417029

Skripsi ini telah diperiksa dan disetujui
Bumiayu, 4 September 2023

Pembimbing I



Khurotul Aeni, M. Kom.
NIDN.0618098802

Pembimbing II



Sorikhi, M. Kom.
NIDN. 0608087902

Ketua Jurusan



Khurotul Aeni, M. Kom.
NIDN.0618098802

PENGESAHAN SKRIPSI

JUDUL : PENERAPAN K-NEAREST NEIGHBOR UNTUK
KLASIFIKASI PENYAKIT TBC PARU
NAMA : AMIRA NABILA ZAKIYYAH
NIM : 42417029

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 20 September 2023. Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom)

Bumiayu, 24 September 2023


Nama Penguji

Tanda Tangan

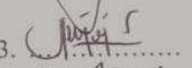
1. Khurotul Aeni, M.Kom

1. 

2. Sorikhi, M.Kom

2. 

3. Fathulloh, S.T., M.Kom

3. 

4. Nurul Mega Saraswati, M.Kom

4. 

Dekan Fakultas Sains dan Teknologi
Universitas Peradaban



DR. Pudjono, SU., Apt
NIDN. 9990000424

Ketua Program Studi



Khurotul Aeni, M.Kom
NIDN. 0618098802

ABSTRACT

Pulmonary Tuberculosis (PTB) is a disease caused by an infection of the bacterium Mycobacterium tuberculosis. PTB can spread through the air and infect the human respiratory system. Therefore, early detection and accurate diagnosis are crucial for controlling the spread of this disease and providing proper treatment to patients. PTB remains a global health issue, with Indonesia ranking third in the number of PTB cases after India and China. PTB cases are more prevalent in developing countries, such as Indonesia. PTB cases in Bumiayu are relatively high, and the transmission of PTB is quite easy, making it quickly transmissible to healthy individuals. Easy and rapid transmission can lead to an increase in PTB patients and a higher risk of mortality, underscoring the importance of early diagnosis to reduce transmission and mortality risks. It can be concluded that solutions to this problem have been actively sought, not only in the field of healthcare but also in technology, with the role of information technology and data mining, particularly the application of K-Nearest Neighbor algorithm for PTB disease classification. K-Nearest Neighbor algorithm is highly suitable for PTB disease classification, as it is utilized to classify and improve early diagnosis of PTB based on rapidly emerging and accurate symptoms. The accuracy rate obtained from the implementation of K-Nearest Neighbor algorithm for PTB disease classification at the Bumiayu Health Center, using a confusion matrix, resulted in an accuracy of 89%.

Keywords: Pulmonary Tuberculosis, Data Mining, K-Nearest Neighbor Algorithm, Confusion Matrix

ABSTRAK

Penyakit Tuberkulosis Paru (TBC Paru) adalah penyakit disebabkan oleh infeksi bakteri *Mycobacterium tuberculosis*. TBC Paru dapat menyebar melalui udara dan dapat menginfeksi sistem pernapasan manusia. Oleh karena itu, deteksi dini dan diagnosis yang akurat sangat penting untuk mengendalikan penyebaran penyakit ini dan memberikan perawatan yang tepat kepada pasien. TBC paru masih menjadi masalah kesehatan di dunia, Indonesia menduduki posisi ketiga jumlah penderita TBC paru setelah India dan Cina. Kasus TBC paru lebih banyak di negara berkembang contohnya Indonesia. Kasus TBC paru di Bumiayu terbilang tinggi serta penularan penyakit TBC paru sangat mudah sehingga dapat dengan cepat menular kepada orang yang sehat. Penularan yang mudah dan cepat dapat menyebabkan bertambahnya penderita TBC paru lebih cepat dan memiliki resiko kematian yang tinggi sehingga perlunya diagnosis dini untuk mengurangi penularan dan mengurangi resiko kematian. Masalah fenomena tersebut dapat ditarik kesimpulan bahwa solusi untuk masalah tersebut telah banyak dilakukan pencarian solusinya tidak hanya dalam bidang kesehatan saja akan tetapi dalam bidang teknologi pun turut ikut andil, yaitu peran teknologi informasi *data mining* penerapan algoritma *K-Nearest Neighbor* untuk klasifikasi penyakit TBC paru. Algoritma *K-Nearest Neighbor* sangat baik digunakan untuk klasifikasi penyakit TBC paru, dimana algoritma *K-Nearest Neighbor* dimanfaatkan untuk mengklasifikasi dan meningkatkan diagnosis dini penyakit TBC paru berdasarkan gejala-gejala yang timbul secara cepat dan akurat. Tingkat akurasi yang didapatkan dari implementasi algoritma *K-Nearest Neighbor* untuk klasifikasi penyakit TBC paru di Puskesmas Bumiayu menggunakan *confusion matrix* menghasilkan akurasi sebesar 89%.

Kata kunci: Penyakit TBC Paru, *Data Mining*, Algoritma *K-Nearest Neighbor*, *Confusion Matrix*

KATA PENGANTAR

Segala puji bagi Allah SWT yang telah melimpahkan kesehatan keselamatan dan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Penerapan K-Nearest Neighbor untuk klasifikasi penyakit TBC paru”.

Penulisan skripsi ini merupakan syarat wajib yang harus dilakukan setiap mahasiswa dalam rangka untuk meraih gelar Sarjana Komputer Fakultas Sains dan Teknologi Universitas Peradaban Bumiayu, Penulis menyadari bahwa dalam penulisan skripsi ini, penulis sangat membutuhkan motivasi, masukan dan saran dari berbagai pihak sehingga skripsi ini bisa diselesaikan, oleh karena itu, penulis mengucapkan terima kasih yang sangat dalam kepada:

1. (Alm) Bapak Muslih dan (Almh) Ibu Sa’atun serta keluarga dan saudara penulis yang telah memberikan doa, dukungan baik moril maupun material sehingga penulis dapat menyelesaikan penulisan skripsi ini.
2. Prof Dr. Muh. Kadarisman. SH. M. Si, selaku rektor Universitas Peradaban Bumiayu.
3. Dr, Pudjono, S.U., Apt, selaku dekan Fakultas Sains dan Teknologi Universitas Peradaban Bumiayu.
4. Ibu Khurotul Aeni, M. Kom, selaku Kepala Program Studi Informatika Universitas Peradaban Bumiayu, sekaligus dosen pembimbing pertama yang telah memberikan saran, nasehat, dukungan dan arahan yang sangat dibutuhkan oleh penulis dalam menyelesaikan skripsi ini.
5. Bapak Sorikhi, M. Kom, selaku pembimbing kedua yang senantiasa memberikan, saran, dukungan dan bimbingan dengan setulus hati yang sangat dibutuhkan oleh penulis dalam penyusunan skripsi ini.
6. Teman-teman saya Febrilania Nur Islami, Khaerunisa, Silvani Margita saputri atas dukungan, bantuan dalam menyelesaikan skripsi ini.

Bumiayu, 24 September 2023

Amira Nabila Zakiyyah

DAFTAR ISI

PERNYATAAN PENULIS	ii
PERSETUJUAN SKRIPSI.....	iii
PENGESAHAN SKRIPSI.....	iv
<i>ABSTRACT</i>	v
ABSTRAK.....	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
1.6 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1. Penelitian Terkait	7
2.2. Landasan teori	13
2.2.1. Data Mining	13
2.2.2. Klasifikasi	14

2.2.3.	Algoritma <i>K-Nearest Neighbor (KNN)</i>	15
2.2.4.	<i>Tuberculosis</i>	16
2.2.5.	<i>Confusion Matrix</i>	18
2.2.6.	Tkinter	19
2.3.	Kerangka Pemikiran	20
BAB III METODE PENELITIAN		24
3.1	Tahapan Penelitian	24
3.1.1	Identifikasi Masalah	25
3.1.2	Penentuan Tujuan.....	25
3.1.3	Pengambilan Data	26
3.1.4	Analisis Kebutuhan	27
3.1.5	Pengolahan Data.....	28
3.1.6	<i>Data Mining</i>	29
3.1.7	Kesimpulan Dan Saran.....	29
3.2	Proses Data Mining	30
3.2.1	Pembagian <i>Data Training</i> dan <i>Data Testing</i>	31
3.2.2	Proses <i>K-Nearest Neighbor</i>	32
3.2.3	Proses <i>Confusion Matrix</i>	33
BAB IV PEMBAHASAN		35
4.1	Persiapan Data Pasien TBC Paru	35
4.2	Praproses Data Pasien TBC Paru (Transformasi data).....	38
4.3	Proses Data Mining	40
4.3.1	Pembagian <i>Data Training</i> dan <i>Data Testing</i>	41
4.3.2	Penerapan algoritma <i>K-Nearest Neighbor</i>	42

4.3.3	Pengujian Data	43
4.4	Penerapan GUI <i>Tkinter</i>	43
BAB V KESIMPULAN DAN SARAN		47
5.1.	Kesimpulan.....	47
5.2.	Saran.....	47
DAFTAR PUSTAKA		48

DAFTAR GAMBAR

Gambar 2. 1 Alur Data Mining	13
Gambar 2. 2 Kerangka Pemikiran	20
Gambar 3. 1 Kerangka Pemikiran	24
Gambar 3. 2 Proses Data Mining	31
Gambar 3. 3 Proses KNN.....	33
Gambar 4. 1 Pair Plot Hasil Klasifikasi	42
Gambar 4. 2 Tampilan Awal.....	44
Gambar 4. 3 Tampilan Menu	44
Gambar 4. 4 Tampilkan Data Pasien TBC.....	45
Gambar 4. 5 Cek Missing Value.....	45

DAFTAR TABEL

Tabel 3. 1 Confussion Matrics	34
Tabel 4. 1 Data Asli Pasien TBC	35
Tabel 4. 2 Indikator Penyakit TBC	38s
Tabel 4. 3 Data Sebelum Transformasi.....	39
Tabel 4. 4 Data Setelah Transformasi	40
Tabel 4. 5 Pembagian Data Training dan Data Testing	41
Tabel 4. 6 Tabel Confusion Matrix	43
Tabel 4. 7 Hasil Akurasi KNN.....	46