



**KLASIFIKASI PENYAKIT TUBERKULOSIS (TBC)
MENGUNAKAN METODE *UNIFORM MANIFOLD
APPROXIMATION AND PROJECTION* (UMAP) DAN *K-
NEAREST NEIGHBORS* (K-NN) DI PUSKESMAS
PRABUMULIH TIMUR**

SKRIPSI

Abel Destria

2021.11.0084

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN SAINS
UNIVERSITAS INDO GLOBAL MANDIRI**

2025

**KLASIFIKASI PENYAKIT TUBERKULOSIS (TBC)
MENGUNAKAN METODE *UNIFORM MANIFOLD
APPROXIMATION AND PROJECTION* (UMAP) DAN *K-
NEAREST NEIGHBORS* (K-NN) DI PUSKESMAS
PRABUMULIH TIMUR**



SKRIPSI

**Diajukan Sebagai Syarat untuk Menyelesaikan
Pada Pendidikan Program Strata-1
Program Studi Teknik Informatika**

Oleh:

**Abel Destria
2021.11.0084**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER DAN SAINS
UNIVERSITAS INDO GLOBAL MANDIRI
2025**

LEMBAR PENGESAHAN SKRIPSI

Klasifikasi Penyakit Tuberkulosis (TBC) Menggunakan Metode *Uniform Manifold Approximation And Projection (UMAP)* Dan *K-Nearest Neighbors (K-NN)* Di Puskesmas Prabumulih Timur

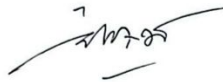
Oleh

Abel Destria

NPM : 2021.11.0084

Palembang , 20 Februari 2025

Pembimbing I



Dr. Rendra Gustriansyah, S.T., M.Kom
NIK : 1999.01.0006

Pembimbing II



Ir Nazori Suhandi, M.M
NIK:1999.01.0008

Mengetahui,

Dekan Fakultas Ilmu Komputer dan Sains
FAKULTAS ILMU KOMPUTER DAN SAINS




Rudi Heriansyah, S.T., M.Eng. Ph.D
NIK:2022.01.0315

LEMBAR PERSETUJUAN DEWAN PENGUJI

Pada hari Senin tanggal 10 Febuari 2025 telah dilaksanakan ujian sidang skripsi :

Nama : Abel Destria

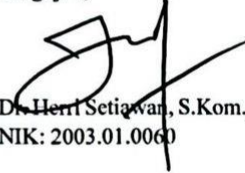
NPM : 2021.11.0084

Judul : Klasifikasi Penyakit Tuberkulosis (TBC) Menggunakan Metode *Uniform Manifold Approximation And Projection (UMAP)* dan *K-Nearest Neighbors (K-NN)* di Puskesmas Prabumulih Timur

Oleh Prodi Teknik Informatika Fakultas Ilmu Komputer dan Sains Universitas Indo Global Mandiri Palembang

Palembang, 20 Febuari 2025

Penguji 1,



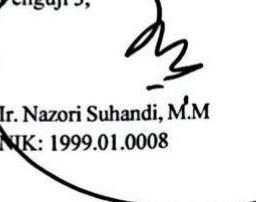
Dr. Henri Setiawan, S.Kom., M.Kom
NIK: 2003.01.0060

Penguji 2,



Lastris Widya Astuti, S.Kom., M.Kom
NIK: 2003.01.0063

Penguji 3,



Ir. Nazori Suhandi, M.M
NIK: 1999.01.0008

Menyetujui,
Ka. Prodi Teknik Informatika



Zaid Romegar Mair, S.T., M.Cs
NIK : 2021.01.0307



SURAT KETERANGAN REVISI SKRIPSI
PROGRAM STUDI TEKNIK INFORMATIKA (SI)
FASILKOM DAN SAINS UNIVERSITAS INDO GLOBAL MANDIRI

Kami yang bertanda tangan dibawah ini, menerangkan bahwa :

Nama : Abel Destria
NPM : 2021.11.0084
Judul : Klasifikasi Penyakit Tuberkulosis (TBC) Menggunakan Metode
Uniform Manifold Approximation And Projection (UMAP) dan *K-Nearest Neighbors (K-NN)* di Puskesmas Prabumulih Timur

Mahasiswa yang namanya tercantum diatas, telah selesai merevisi penulisan SKRIPSI

Palembang, 20 Februari 2025

Penguji 1,

Dr. Herri Setiawan, S.Kom., M.Kom
NIK: 2003.01.0060

Penguji 2,

Lastri Widya Astuti, S.Kom., M.Kom
NIK: 2003.01.0063

Penguji 3,

Ir. Nazori Suhandi, M.M
NIK: 1999.01.0008

Menyetujui,
Ka. Prodi Teknik Informatika

Zaid Romegar Mair, S.T., M.Cs
NIK: 2021.01.0307

**KLASIFIKASI PENYAKIT TUBERKULOSIS (TBC) MENGGUNAKAN
METODE *UNIFORM MANIFOLD APPROXIMATION AND PROJECTION*
(UMAP) DAN *K-NEAREST NEIGHBORS* (K-NN) DI PUSKESMAS
PRABUMULIH TIMUR
ABSTRAK**

Tuberkulosis (TBC) adalah penyakit infeksi yang membutuhkan diagnosis cepat dan akurat. Penelitian ini mengembangkan sistem klasifikasi TBC berbasis UMAP untuk reduksi dimensi dan K-NN untuk klasifikasi guna meningkatkan akurasi dan efisiensi diagnosis di Puskesmas Prabumulih Timur. Dataset yang digunakan terdiri dari 278 rekam medis pasien dengan atribut gejala klinis seperti batuk berdahak, suhu tubuh, dan penurunan berat badan. Model diuji dengan membandingkan K-NN tanpa UMAP dan K-NN dengan UMAP untuk mengukur efektivitas reduksi dimensi.

Hasil penelitian menunjukkan bahwa *accuracy* K-NN tanpa UMAP adalah 94,57%, sedangkan setelah menggunakan UMAP meningkat menjadi 100%. Nilai *precision* tetap 100%, nilai *recall* meningkat dari 90,20% menjadi 100% dan *F1-score* juga meningkat dari 94,85% menjadi 100%. Selain meningkatkan performa, UMAP juga mempercepat komputasi, membuat sistem lebih efisien untuk digunakan dalam layanan kesehatan. Dengan hasil ini, model klasifikasi dari metode UMAP dan K-NN dapat menjadi alat bantu yang lebih cepat dan akurat dalam mendukung tenaga medis untuk diagnosis dini TBC, serta berpotensi diterapkan dalam sistem kesehatan yang lebih luas.

Kata kunci: Tuberkulosis, UMAP, K-NN, klasifikasi, Reduksi Dimensi, Puskesmas Prabumulih Timur

**CLASSIFICATION OF TUBERCULOSIS (TBC) DISEASE USING
UNIFORM MANIFOLD APPROXIMATION AND PROJECTION (UMAP)
AND K-NEAREST NEIGHBORS (K-NN) METHODS AT PRABUMULIH
TIMUR COMMUNITY HEALTH CENTER**

ABSTRACT

Tuberculosis (TBC) is an infectious disease that requires fast and accurate diagnosis. This study develops a TB classification system based on UMAP for dimension reduction and K-NN for classification to improve the accuracy and efficiency of diagnosis at the Prabumulih Timur Health Center. The dataset used consists of 278 patient medical records with clinical symptom attributes such as cough with phlegm, body temperature, and weight loss. The model was tested by comparing K-NN without UMAP and K-NN with UMAP to measure the effectiveness of dimension reduction.

The results showed that the accuracy of K-NN without UMAP was 94.57%, while after using UMAP it increased to 100%. The precision value remained at 100%, the recall value increased from 90.20% to 100% and the F1-score also increased from 94.85% to 100%. In addition to improving performance, UMAP also speeds up computation, making the system more efficient for use in health services. With these results, the classification model of the UMAP and K-NN methods can be a faster and more accurate tool in supporting medical personnel for early diagnosis of TB, and has the potential to be applied in a wider health system.

Keywords: *Tuberculosis, UMAP, K-NN, classification, Dimension Reduction, Prabumulih Timur Health Center*

KATA PENGANTAR

Puji dan syukur saya ucapkan atas kehadiran Allah Subhanahu Wata'ala berkat rahmat dan hidayahnya akhirnya penulis dapat menyelesaikan penelitian ini dengan baik tepat pada waktunya, tidak lupa shalawat serta salam selalu dilimpahkan kepada junjungan kita Nabi besar Nabi Muhammad SAW beserta keluarga sahabat para pengikut dan insyaallah kita semua hingga akhir zaman.

Skripsi yang penulis buat dengan judul “Klasifikasi Penyakit Tuberkulosis (TBC) Menggunakan Metode *Uniform Manifold Approximation And Projection* (UMAP) Dan *K-Nearest Neighbors* (K-NN) Di Puskesmas Prabumulih Timur” disusun guna memenuhi syarat kelulusan dalam memperoleh gelar Sarjana (S1) pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer dan Sains, Universitas Indo Global Mandiri (UIGM) Palembang.

Tidak lupa penulis mengucapkan terimakasih atas bantuan yang diberikan selama penyusunan skripsi ini kepada

1. Bapak Dr. Marzuki Alie, SE., MM, selaku Rektor Universitas Indo Global Mandiri Palembang.
2. Bapak Rudi Heriansyah, ST., M.Eng. Ph.D selaku Dekan Fakultas Ilmu Komputer dan Sains Universitas Indo Global Mandiri.
3. Bapak Zaid Romegar Mair, S.T., M.Cs sebagai Ketua Program Studi Teknik Informatika Universitas Indo Global Mandiri.
4. Bapak Dr. Rendra Gustriansyah S.T., M.Kom sebagai Dosen Pembimbing I.
5. Bapak Ir. Nazori Suhandi, M.M, sebagai Dosen Pembimbing II.
6. Bapak/Ibu Dosen Fakultas Ilmu Komputer dan Sains serta Karyawan/Karyawati Universitas Indo Global Mandiri.
7. Kedua Orang Tua saya dan Keluarga yang telah memberikan dukungan, serta Do'a sehingga penulis bisa menyelesaikan skripsi ini.
8. Semua teman-teman seperjuangan Teknik Informatika Angkatan 2021.

Dengan segala kerendahan hati penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna, oleh karena itu dibutuhkan kritik dan saran untuk perbaikan dan pengembangan skripsi ini sangat diharapkan. Akhir kata, semoga skripsi ini bermanfaat bagi semua pihak terima kasih.

Palembang, 2025

Penulis

Abel Destria

2021.11.0084

DAFTAR ISI

HALAMAN JUDUL LUAR	i
HALAMAN JUDUL DALAM	ii
LEMBAR PENGESAHAN SKRIPSI	iii
LEMBAR PERSETUJUAN DEWAN PENGUJI.....	iv
SURAT KETERANGAN REVISI SKRIPSI	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvi
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat	4
1.4.1 Tujuan Penelitian.....	4
1.4.2 Manfaat Penelitian.....	5
1.5 Sistematika Penulisan	5
BAB II	7
LANDASAN TEORI.....	7
2.1 Puskesmas Prabumulih Timur.....	7
2.2 Tuberkulosis.....	8
2.3 <i>Machine Learning</i>	10

2.4	Reduksi Dimensi.....	12
2.5	Klasifikasi.....	12
2.6	Metode <i>Elbow</i>	14
2.7	Metode UMAP.....	16
2.8	Metode K-NN.....	18
2.9	<i>Confusion Matrix</i>	19
2.10	Penelitian Terdahulu.....	22
BAB III.....		27
METODOLOGI PENELITIAN.....		27
3.1	Tahapan Penelitian.....	27
3.2	Pengumpulan Data.....	28
3.3	Pra-Pemrosesan Data.....	30
3.3.1	Penghapusan Data Duplikat.....	30
3.3.2	Penghapusan Kolom Yang Tidak Digunakan.....	31
3.3.3	<i>Encoding Data</i>	32
3.3.4	Penanganan Nilai Hilang.....	33
3.3.5	Pembagian Dataset.....	34
3.4	Penerapan Metode <i>Elbow</i>	35
3.5	Penerapan Metode UMAP.....	36
3.6	Penerapan Metode K-NN.....	39
3.6.1	K-NN tanpa UMAP.....	39
3.6.2	K-NN dengan UMAP.....	41
3.7	Pengujian Data Dummy.....	44
3.7.1	Pengujian Menggunakan R-studio.....	44
3.7.2	Pengujian Menggunakan Excel.....	48

3.8 Hasil	50
BAB IV	52
HASIL DAN PEMBAHASAN	52
4.1 Pengumpulan Data	52
4.2 Pra – Pemrosesan Data.....	52
4.2.1 Penghapusan Data Duplikat	54
4.2.2 Penghapusan Kolom Yang Tidak Digunakan.....	54
4.2.3 <i>Encoding Data</i>	55
4.2.4 Penanganan Nilai Hilang.....	56
4.2.5 Pembagian Dataset	57
4.3 Penerapan Metode <i>Elbow</i>	57
4.4 Penerapan Metode UMAP	59
4.5 Penerapan Metode KNN	62
4.6 Pengujian.....	63
4.7 Hasil	65
BAB V.....	67
KESIMPULAN DAN SARAN	67
5.1 Kesimpulan	67
5.2 Saran	67
DAFTAR PUSTAKA.....	69
LAMPIRAN.....	77

DAFTAR GAMBAR

Gambar 2.1 Puskesmas Prabumulih Timur	8
Gambar 3.1 Tahapan Penelitian	27
Gambar 3.2 Tahapan Pra-Pemrosesan Data.....	30
Gambar 3.3 Proses Penghapusan Data Duplikat	30
Gambar 3.4 Proses Penghapusan Kolom Yang Tidak Digunakan	31
Gambar 3.5 Hasil Penghapusan Kolom Atribut	32
Gambar 3.6 Proses <i>Encoding Data</i>	32
Gambar 3.7 Hasil <i>Encoding Data Pasien</i> Penyakit TBC	33
Gambar 3.8 Proses Penanganan Nilai Hilang.....	33
Gambar 3.9 Pembagian Dataset.....	34
Gambar 3.10 Penerapan Metode <i>Elbow</i>	35
Gambar 3.11 Grafik Hasil Dari Metode <i>Elbow</i>	35
Gambar 3.12 Proses Menentukan Jumlah Dimensi.....	36
Gambar 3.13 Hasil Kumulatif Variansi Dimensi	37
Gambar 3.14 Grafik Ambang Hasil Kumulatif Variansi Dimensi.....	37
Gambar 3.15 Penerapan Metode UMAP	38
Gambar 3.16 Hasil Penerapan Metode UMAP	38
Gambar 3.17 Penerapan Metode K-NN Tanpa UMAP	39
Gambar 3.18 Penerapan Metode K-NN Tanpa UMAP	42
Gambar 3.19 Proses Pengujian Kinerja Model dengan <i>Confusion Matrix</i>	45
Gambar 3.20 Hasil <i>Confusion Matrix</i> K-NN Tanpa UMAP	46
Gambar 3.21 Hasil <i>Confusion Matrix</i> K-NN Dengan UMAP.....	46
Gambar 3.22 Proses Perhitungan Metrik Evaluasi	47
Gambar 3.23 Hasil Pengujian <i>Confusion Matrix</i>	48
Gambar 3.24 <i>Confusion Matrix</i> K-NN Tanpa UMAP Excel.....	48
Gambar 3.25 <i>Confusion Matrix</i> K-NN UMAP Excel	49
Gambar 3.26 Hasil <i>Confusion Matrix</i> Excel	50
Gambar 4.1 Hasil Penghapusan Data Duplikat	54
Gambar 4.2 Hasil Penghapusan Kolom Yang Tidak Digunakan	55

Gambar 4.3 Hasil <i>Encoding Data</i>	56
Gambar 4.4 Hasil Penanganan Nilai Hilang.....	57
Gambar 4.5 Hasil Metode <i>Elbow</i>	58
Gambar 4.6 Hasil Kumulatif Variansi Dimensi.....	60
Gambar 4.7 Hasil Diagram Ambang Batas Kumulatif Variansi Dimensi	60
Gambar 4.8 Hasil Reduksi UMAP	61
Gambar 4.9 <i>Confusion Matrix</i> K-NN Tanpa UMAP	63
Gambar 4.10 <i>Confusion Matrix</i> K-NN Dengan UMAP	64
Gambar 4.11 Hasil Perhitungan Metrik Evaluasi	65
Gambar 4.12 Grafik Perhitungan Metrik Evaluasi.....	665

DAFTAR TABEL

Tabel 2.1 Contoh Model Confusion Matrix.....	20
Tabel 2.2 Penelitian Terdahulu	22
Table 3.1 Data Dummy Pasien TBC Puskesmas Prabumulih Timur	29
Table 3.2 Data Latih	39
Table 3.3 Data Uji.....	40
Table 3. 4 Hasil Perhitungan K-NN Excel	41
Table 3.5 Hasil Target K-NN Dengan Excel	41
Table 3.6 Data Latih	42
Table 3.7 Data Uji.....	42
Table 3.8 Hasil Perhitungan K-NN Excel	44
Table 3.9 Hasil Target K-NN Dengan Excel	44
Table 4.1 Data Pasien TBC Puskesmas Prabumulih Timur.....	53

DAFTAR LAMPIRAN

Lampiran 1. Biografi Penulis.....	77
Lampiran 2. Pernyataan Tidak Plagiat.....	78
Lampiran 3. Kartu Bimbingan Skripsi.....	79
Lampiran 4. Surat Izin Pengambilan Data Dinkes.....	80
Lampiran 5. Surat Izin Pengambilan Data Puskesmas Prabumulih Timur.....	81
Lampiran 6. Surat Balasan Pengambilan Data Dinkes.....	82
Lampiran 7. Surat Balasan Pengambilan Data Puskesmas Prabumulih Timur.....	83