

**DETEKSI PENYAKIT ANEMIA MENGGUNAKAN METODE
MODIFIED K-NEAREST NEIGHBOR (MKNN)**

SKRIPSI



Oleh :

PUTRA DWI WIRA GARDHA YUNIAHANS

20081010103

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
2024**

LEMBAR PENGESAHAN SKRIPSI

**Judul : DETEKSI PENYAKIT ANEMIA MENGGUNAKAN METODE
MODIFIED K-NEAREST NEIGHBOR (MKNN)**

Oleh : Putra Dwi Wira Gardha Yuniahans

NPM : 20081010103

**Telah Diseminarkan Dalam Ujian Skripsi Pada :
Hari Kamis, Tanggal 4 Juli 2024**

Mengetahui

1. **Dosen Pembimbing**

Dr. Eng. Ir. Anggraini P S, ST., MT.
NIP. 222198 60 816400

1. **Dosen Penguji**

Budi Nugroho, S.Kom., M.Kom.
NIP. 19800907 2021211 005

2.

Yisti Vita Via, S.ST., M.Kom.
NIP. 19860425 2021212 001

2.

Andreas Nugroho S, S.Kom., M.Kom.
NPT. 211199 00 412271

Menyetujui

**Dekan
Fakultas Ilmu Komputer**

Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

**Koordinator Program Studi
Informatika**

Fetty Tri Anggraeny, S.Kom., M.Kom
NIP. 19820211 2021212 005

SURAT PERNYATAAN BEBAS DARI PLAGIASI

Saya, mahasiswa Program Studi Sarjana Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur, yang bertanda tangan dibawah ini:

Nama : Putra Dwi Wira Gardha Yuniahans

NPM : 20081010103

Menyatakan dengan sesungguhnya bahwa Skripsi/Tugas Akhir yang saya kerjakan berjudul:

“DETEKSI PENYAKIT ANEMIA MENGGUNAKAN METODE MODIFIED
K-NEAREST NEIGHBOR (MKNN)”

bukan merupakan plagiasi sebagian atau keseluruhan dari Skripsi/Tugas Akhir/Penelitian orang lain dari juga bukan merupakan produk dan software yang saya beli dari pihak lain. Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini secara keseluruhan adalah pekerjaan Saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional “Veteran” Jawa Timur maupun di Institut Pendidikan lain. Bukti hasil pengecekan plagiasi dokumen ini dapat ditelusuri melalui QR Code di bawah.

Apabila di kemudian hari terbukti bahwa dokumen ini merupakan plagiasi karya orang lain, saya sanggup menerima sanksi sesuai aturan yang berlaku.

Demikian atas perhatiannya disampaikan terima kasih.

Surabaya, 19 Juli 2024

Hormat saya,



Putra Dwi Wira Gardha Yuniahans

NPM. 20081010103

DETEKSI PENYAKIT ANEMIA MENGGUNAKAN METODE MODIFIED K-NEAREST NEIGHBOR (MKNN)

Nama Mahasiswa : Putra Dwi Wira Gardha Yuniahans
NPM : 20081010103
Program Studi : Teknik Informatika
Dosen Pembimbing : Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT.
Yisti Vita Via, S.ST., M.Kom.

Abstrak

Anemia merupakan sebuah kondisi dimana sel darah merah atau konsentrasi hemoglobin lebih rendah dari kadar normalnya. Berdasarkan hasil Riset Kesehatan Dasar (Riskesdas) Kementerian Kesehatan Indonesia (KEMENKES) dari tahun 2017, 2013, hingga 2018 angka prevalensi penyakit anemia terhadap usia 0 hingga 24 tahun selalu mengalami tren kenaikan. Dampak buruk juga dirasakan oleh kalangan-kalangan tersebut dari keterlambatan perkembangan psikomotor, prestasi di sekolah yang semakin menurun hingga pada perempuan juga bisa mengakibatkan kelahiran yang prematur. Maka dari itu bantuan teknologi seperti *machine learning* untuk mendeteksi dini penyakit anemia tersebut sangat penting. Metode yang digunakan dalam mendeteksi penyakit anemia tersebut adalah metode Modified K-Nearest Neighbor (MKNN). MKNN sendiri merupakan sebuah metode modifikasi dari metode K-Nearest Neighbor (KNN). Pada metode MKNN sendiri terdapat tambahan proses yang tidak dimiliki oleh metode KNN yakni pada proses perhitungan validasi pada data latih dan perhitungan *weight voting*. Sehingga dalam penelitian ini metode MKNN menghasilkan nilai akurasi 90% dengan parameter split data dengan rasio 60:40 serta nilai $k=3$. Hasil tersebut menyatakan bahwa metode MKNN dapat meningkatkan nilai akurasi dari metode KNN sendiri.

Kata kunci : MKNN, Penyakit Anemia, Data Mining, Machine Learning

KATA PENGANTAR

Dengan rasa syukur, penulis menghaturkan puji dan syukur kepada Allah Subhanahu wa Ta'ala atas segala berkah dan rahmat-Nya, sehingga penulis dapat menyelesaikan skripsi dengan judul "DETEKSI PENYAKIT ANEMIA MENGGUNAKAN METODE MODIFIED K-NEAREST NEIGHBOR (MKNN)" tepat waktu. Penulis menyadari bahwa penyusunan skripsi ini memerlukan banyak usaha, pemikiran, dan waktu.

Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang konstruktif untuk penyempurnaan di masa mendatang. Akhir kata, penulis berharap skripsi ini dapat bermanfaat bagi para pembaca serta menjadi referensi yang berguna untuk penelitian selanjutnya.

Surabaya, 19 Juli 2024

Penulis

UCAPAN TERIMA KASIH

Penelitian ini dapat terselesaikan berkat berbagai faktor, termasuk dukungan dari beberapa individu yang selalu memberikan dampak positif kepada penulis sepanjang proses penelitian. Bantuan dan dorongan mereka sangat berarti, memberikan motivasi tambahan saat menghadapi tantangan yang muncul. Penulis ingin mengucapkan terima kasih yang tulus kepada mereka yang telah memberikan dukungan moral maupun material, sehingga penelitian ini dapat berjalan dengan lancar dan mencapai hasil yang diharapkan.

Maka dengan ini, penulis ingin mengucapkan terima kasih kepada:

1. Ibu Suwanah, Bapak Hadi Sulistiono selaku orang tua dan Putri Fardha Asa selaku kakak saya yang selalu memberikan dukungan penuh kepada penulis dalam bentuk material maupun non-material.
2. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Dr. Novirina Hendrasarie, S.T, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Ibu Fetty Tri Anggraeny, S.Kom, M.Kom. selaku Koordinator Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
5. Ibu Dr. Eng. Ir. Anggraini Puspita Sari, ST., MT. selaku dosen pembimbing pertama yang telah dengan sepenuh hati memberikan arahan selama penulis melakukan penelitian.
6. Ibu Yisti Vita Via, S.ST., M.Kom. selaku dosen pembimbing kedua yang telah mengarahkan penulis dalam melakukan penelitian hingga selesai.
7. Seluruh Bapak dan Ibu Dosen Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur yang telah memberikan ilmu selama kegiatan belajar mengajar berlangsung.
8. Teman-teman penulis yang telah memberikan motivasi kepada penulis selama penelitian hingga penelitian mencapai hasil.

DAFTAR ISI

LEMBAR PENGESAHAN.....	ii
SURAT PERNYATAAN BEBAS DARI PLAGIASI.....	iii
KATA PENGANTAR.....	v
UCAPAN TERIMAKASIH.....	vi
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	viii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Tujuan.....	3
1.4. Manfaat.....	3
1.5. Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Penelitian Terdahulu.....	5
2.2. Penyakit Anemia.....	7
2.3. Data Mining.....	8
2.4. KNN.....	9
2.5. Modified KNN.....	12
2.6. Confusion Matrix.....	13
2.7. Python.....	16
2.8. Pandas.....	16
2.9. Matplotlib.....	17
2.10. Jupyter Notebook.....	18
BAB III METODOLOGI.....	19

3.1. Gambaran Penelitian.....	19
3.2. Tahapan Penelitian.....	19
3.3. Studi Literatur.....	20
3.4. Pengumpulan Data.....	20
3.4.1. Sumber Data.....	21
3.4.2. Rancangan Struktur Data.....	21
3.5. Pembuatan Model.....	23
3.5.1. Pengolahan Data.....	24
3.5.2. Implementasi Metode MKNN.....	25
3.6. Skenario Pengujian.....	35
BAB IV HASIL DAN PEMBAHASAN.....	37
4.1. Pengolahan Data.....	37
4.1.1. Pengecekan Data Duplikasi.....	37
4.1.2. Pengecekan Data Missing Value.....	39
4.1.3. Perubahan Tipe Data.....	40
4.1.4. Split Data dan Visualisasi Data.....	41
4.2. Implementasi Modified KNN (MKNN).....	43
4.2.1. Perhitungan Validasi.....	43
4.2.2. Perhitungan Weight Voting.....	47
4.2.3. Perhitungan Confusion Matrix.....	50
4.3. Skenario Pengujian.....	51
4.3.1. Pengujian Pada Parameter Split Data.....	51
4.3.2. Skenario Pengujian Kedua.....	62
4.3.3. Perbandingan Metode MKNN dan KNN.....	74
4.3.4. Analisa Prediksi Metode MKNN.....	78
BAB V KESIMPULAN DAN SARAN.....	80

5.1. Kesimpulan.....	80
5.2. Saran.....	80
DAFTAR PUSTAKA.....	82

DAFTAR TABEL

Tabel 3.1 Struktur Dataset.....	21
Tabel 3.2 Contoh Dataset.....	22
Tabel 3.3 Dataset Perhitungan Manual.....	27
Tabel 3.4 Data Latih Perhitungan Manual.....	27
Tabel 3.5 Data Uji Perhitungan Manual.....	28
Tabel 3.6 Hasil Perhitungan Jarak Data Latih Baris Pertama.....	29
Tabel 3.7 Hasil Pengurutan Nilai Jarak.....	29
Tabel 3.8 Hasil Pencocokan Kelas.....	30
Tabel 3.9 Hasil Keseluruhan Perhitungan Validitas Data Latih.....	31
Tabel 3.10 Hasil Perhitungan Jarak Data Uji Baris Pertama dan Data Latih.....	32
Tabel 3.11 Hasil Perhitungan Weight Voting Data Uji Baris Pertama.....	33
Tabel 3.12 Hasil Pengurutan Nilai Weight Voting Data Uji Baris Pertama.....	33
Tabel 3.13 Hasil Prediksi Data Uji.....	35
Tabel 4.1 Hasil Pengujian Split Data 60:40.....	52
Tabel 4.2 Hasil Pengujian Split Data 70:30.....	54
Tabel 4.3 Hasil Pengujian Split Data 80:20.....	57
Tabel 4.4 Hasil Pengujian Split Data 90:10.....	59
Tabel 4.5 Hasil Pengujian Split Data.....	61
Tabel 4.6 Hasil Pengujian k=5.....	63
Tabel 4.7 Hasil Pengujian k=7.....	65
Tabel 4.8 Hasil Pengujian k=9.....	68
Tabel 4.9 Hasil Pengujian k=11.....	70
Tabel 4.10 Hasil Keseluruhan Pengujian Nilai k.....	73
Tabel 4.11 Hasil Perbandingan MKNN dan KNN.....	75
Tabel 4.12 Data Kesalahan KNN k=3.....	76

Tabel 4.13 Analisa Data Hasil Prediksi MKNN dan KNN k=3.....	77
Tabel 4.14 Nilai Weight Voting MKNN.....	78
Tabel 4.15 Analisa MKNN 60:40 k=3.....	78

DAFTAR GAMBAR

Gambar 2.1 Cara Kerja Metode KNN.....	10
Gambar 2.2 Penjabaran Tabel Confusion Matrix.....	14
Gambar 3.1 Langkah Langkah Penelitian.....	20
Gambar 3.2 Diagram Alir Pembuatan Model.....	24
Gambar 3.3 Diagram Alir Data Preprocessing.....	24
Gambar 3.4 Diagram Alir Implementasi Metode MKNN.....	26
Gambar 4.1 Hasil Pengecekan Data Duplikasi.....	38
Gambar 4.2 Hasil Pengecekan Data Missing Value.....	39
Gambar 4.3 Visualisasi Data Penelitian.....	43
Gambar 4.4 Hasil Perhitungan Validasi.....	47
Gambar 4.5 Contoh Klasifikasi Data.....	50
Gambar 4.6 Hasil Prediksi.....	50
Gambar 4.7 Hasil Akurasi Pengujian MKNN 60:40.....	52
Gambar 4.8 Confusion Matriks Pengujian Split Data 60:40.....	53
Gambar 4.9 Data Prediksi Salah Pengujian Split Data 60:40.....	53
Gambar 4.10 Data Prediksi Benar Pengujian Split Data 60:40.....	54
Gambar 4.11 Hasil Pengujian Split Data 70:30.....	55
Gambar 4.12 Confusion Matriks Pengujian Split Data 70:30.....	55
Gambar 4.13 Data Prediksi Salah Pengujian Split Data 70:30.....	56
Gambar 4.14 Data Prediksi Benar Pengujian Split Data 70:30.....	56
Gambar 4.15 Hasil Pengujian Split Data 80:20.....	57
Gambar 4.16 Confusion Matriks Pengujian Split Data 80:20.....	58
Gambar 4.17 Data Prediksi Benar Pengujian Split Data 80:20.....	58
Gambar 4.18 Data Prediksi Salah Pengujian Split Data 80:20.....	59
Gambar 4.19 Hasil Pengujian Split Data 90:10.....	59

Gambar 4.20 Confusion Matriks Pengujian Split Data 90:10.....	60
Gambar 4.21 Data Prediksi Salah Pengujian Split Data 90:10.....	60
Gambar 4.22 Data Prediksi Benar Pengujian Split Data 90:10.....	60
Gambar 4.23 Hasil Pengujian k=5.....	63
Gambar 4.24 Confusion Matriks Pengujian Split Data 60:40 dan k=5.....	64
Gambar 4.25 Data Prediksi Salah Pengujian k=5.....	64
Gambar 4.26 Data Prediksi Benar Pengujian k=5.....	65
Gambar 4.27 Hasil Pengujian k=7.....	66
Gambar 4.28 Confusion Matriks Pengujian Split Data 60:40 dan k=7.....	66
Gambar 4.29 Data Prediksi Salah Pengujian k=7.....	67
Gambar 4.30 Data Prediksi Benar Pengujian k=7.....	67
Gambar 4.31 Hasil Pengujian k=9.....	68
Gambar 4.32 Confusion Matriks Pengujian Split Data 60:40 dan k=9.....	69
Gambar 4.33 Data Prediksi Salah Pengujian k=9.....	69
Gambar 4.34 Data Prediksi Benar Pengujian k=9.....	70
Gambar 4.35 Hasil Pengujian k=11.....	71
Gambar 4.36 Confusion Matriks Pengujian Split Data 60:40 dan k=11.....	71
Gambar 4.37 Data Prediksi Salah Pengujian k=11.....	72
Gambar 4.38 Data Prediksi Benar Pengujian k=11.....	72
Gambar 4.39 Visualisasi Hasil Pengujian Nilai k.....	73
Gambar 4.40 Hasil Pengujian KNN 60:40 k=3.....	75
Gambar 4.41 Hasil Perbandingan MKNN dan KNN 60:40 k=3.....	75
Gambar 4.42 Data Kesalahan KNN k=3.....	76