

**PERBANDINGAN KINERJA ALGORITMA DECISION TREE
DAN LOGISTIC REGRESSION DALAM MEMPREDIKSI
RISIKO INSOMNIA BERDASARKAN KEBIASAAN TIDUR**



SKRIPSI

**WILCOUSTINE QHRISTMAS PNIEL WIJAYA
00000056960**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**PERBANDINGAN KINERJA ALGORITMA DECISION TREE
DAN LOGISTIC REGRESSION DALAM MEMPREDIKSI
RISIKO INSOMNIA BERDASARKAN KEBIASAAN TIDUR**



SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)

WILCOUSTINE QHRISTMAS PNIEL WIJAYA
00000056960

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Wilcoustine Qhristmas Pniel Wijaya

Nomor Induk Mahasiswa : 00000056960

Program Studi : Informatika

Skripsi dengan judul:

Analisis Kebiasaan Tidur dan Prediksi Risiko Insomnia Menggunakan Machine Learning dengan Metode Decision Tree dan Logistic Regression

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan **TIDAK LULUS** untuk mata kuliah yang telah saya tempuh.

Tangerang, 30 Juni 2025



(Wilcoustine Qhristmas Pniel Wijaya)

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

HALAMAN PENGESAHAN

Skripsi dengan judul

PERBANDINGAN KINERJA ALGORITMA DECISION TREE DAN LOGISTIC REGRESSION DALAM MEMPREDIKSI RISIKO INSOMNIA BERDASARKAN KEBIASAAN TIDUR

oleh

Nama : Wilcoustine Qhristmas Pniel
Wijaya
NIM : 00000056960
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Senin, 14 Juli 2025

Pukul 13.00 s/s 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang

Penguji

(David Agustriawan, S.Kom., M.Sc.,
Ph.D.)
NIDN: 0525088601

(Dr. Maria Irmina Prasetyowati, S.Kom.,
M.T.)
NIDN: 0725057201

Pembimbing

(Dr. Ivransa Zuhdi Pane, B.Eng., M.Eng.)

NIDN: 8812520016

Ketua Program Studi Informatika,

(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Wilcoustine Qhristmas Pniel Wijaya
NIM : 00000056960
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Perbandingan Kinerja Algoritma
Decision Tree dan Logistic Regression
dalam Memprediksi Risiko Insomnia
Berdasarkan Kebiasaan Tidur

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
- Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 30 Juni 2025

Yang menyatakan



Wilcoustine Qhristmas Pniel Wijaya

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PERSEMBAHAN / MOTTO

"Commit your way to the Lord; trust in him, and he will act"

Psalm 22:1 (NIV)



KATA PENGANTAR

Dengan memanjatkan puji syukur kehadiran Tuhan Yang Maha Esa, skripsi saya yang berjudul "Analisis Kebiasaan Tidur dan Prediksi Risiko Insomnia Menggunakan Machine Learning dengan Metode Decision Tree dan Logistic Regression" telah selesai. Karya ini menjadi salah satu syarat untuk memperoleh gelar Sarjana Komputer Jurusan Informatika dari Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya sangat terbantu oleh bimbingan dan dukungan berbagai pihak sejak masa kuliah hingga rampungnya skripsi ini. Oleh karena itu, saya ingin menyampaikan terima kasih kepada:

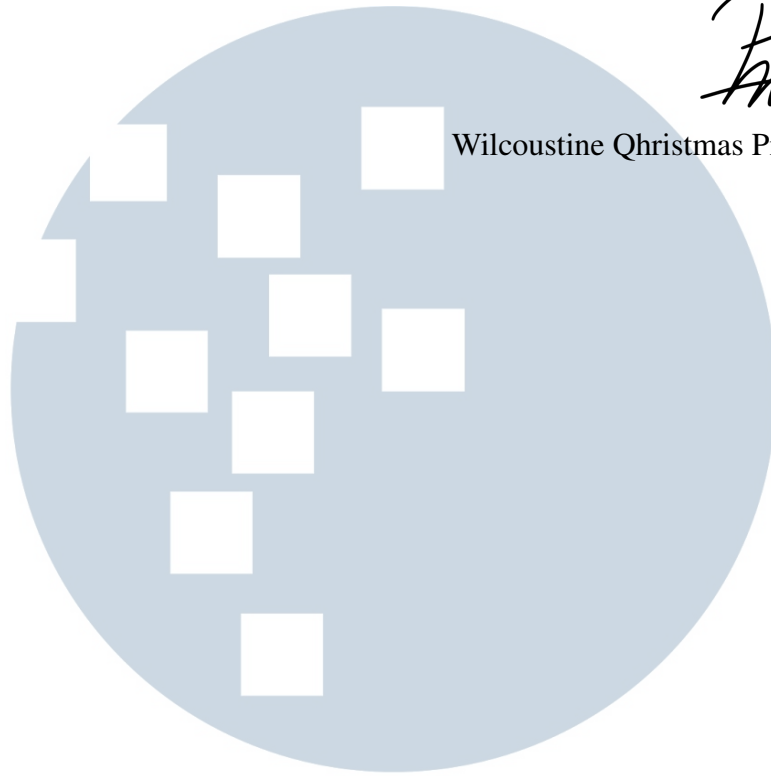
1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Dr. Ivrensa Zuhdi Pane, B.Eng., M.Eng., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Teman saya Anak *Kost*, Motivasi May, Abel, Nisa, Mishel, Keren, dan Jessica yang telah menjadi teman bercerita, berbagi ide, membantu baik dalam bentuk diskusi ilmiah maupun dukungan emosional serta dorongan yang memotivasi selama saya menyelesaikan tesis ini.
7. Semua pihak yang telah membantu demi lancarnya penulisan laporan skripsi yang tidak bisa disebutkan satu persatu.

Meskipun tesis ini belum sepenuhnya sempurna, saya sangat berharap hasil penelitiannya dapat memberikan informasi dan inspirasi yang berguna bagi para pembaca.

Tangerang, 30 Juni 2025



Wilcoustine Qhristmas Pniel Wijaya



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

PERBANDINGAN KINERJA ALGORITMA DECISION TREE DAN LOGISTIC REGRESSION DALAM MEMPREDIKSI RISIKO INSOMNIA BERDASARKAN KEBIASAAN TIDUR

Wilcoustine Qhrstmas Pniel Wijaya

ABSTRAK

Insomnia merupakan gangguan tidur yang memiliki dampak signifikan terhadap kesehatan fisik dan mental. Penelitian ini bertujuan mengembangkan model prediksi risiko insomnia menggunakan algoritma *machine learning* berdasarkan data kebiasaan tidur dan aktivitas harian. Dataset numerik dari Kaggle diproses menggunakan *StandardScaler* dan diperluas dengan *data augmentation*. Dua algoritma klasifikasi, yaitu *Decision Tree (DT)* dan *Logistic Regression (LR)*, diterapkan dan dievaluasi menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil evaluasi menunjukkan bahwa *Logistic Regression* memberikan performa terbaik pada data uji, dengan akurasi sebesar 0,91, presisi 0,910112, *recall* 0,987805, dan *F1-score* 0,947368. Sebaliknya, *Decision Tree* mencatat akurasi 0,89, presisi 0,9863, *recall* 0,878, dan *F1-score* 0,929. Analisis *Permutation Importance* mengidentifikasi *Caffeine Intake* sebagai fitur paling berpengaruh terhadap risiko insomnia pada kedua model. Penelitian ini menyimpulkan bahwa *Logistic Regression* merupakan model yang andal untuk prediksi risiko insomnia, serta menegaskan pentingnya pengendalian konsumsi kafein sebagai langkah preventif.

Kata kunci: *Decision Tree*, Insomnia, *Logistic Regression*, *Machine Learning*, Prediksi Risiko

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

**PERFORMANCE COMPARISON OF DECISION TREE AND LOGISTIC
REGRESSION ALGORITHMS IN PREDICTING INSOMNIA RISK BASED
ON SLEEP HABITS**

Wilcoustine Qhristmas Pniel Wijaya

ABSTRACT

Insomnia is a sleep disorder with significant implications for both physical and mental health. This study aims to develop a machine learning-based prediction model for insomnia risk using data on sleep habits and daily activities. A numerical dataset from Kaggle was preprocessed using StandardScaler and expanded through data augmentation. Two classification algorithms, Decision Tree (DT) and Logistic Regression (LR), were implemented and evaluated using accuracy, precision, recall, and F1-score metrics. The results indicate that Logistic Regression outperformed Decision Tree on the test set, achieving an accuracy of 0.91, precision of 0.910112, recall of 0.987805, and F1-score of 0.947368. In contrast, Decision Tree achieved an accuracy of 0.89, precision of 0.9863, recall of 0.878, and F1-score of 0.929. Permutation Importance analysis revealed that Caffeine Intake was the most influential feature in predicting insomnia risk for both models. The study concludes that Logistic Regression is a robust and reliable model for predicting insomnia risk, with caffeine intake identified as a key risk factor to consider in early detection and intervention strategies.

Keywords: *Decision Tree, Insomnia, Logistic Regression, Machine Learning, Risk Prediction*

U M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	vi
KATA PENGANTAR	vii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR KODE	xv
DAFTAR RUMUS	xvi
DAFTAR LAMPIRAN	xvii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB 2 LANDASAN TEORI	6
2.1 Penelitian Pendahulu	6
2.2 Tidur	9
2.3 Kebiasaan Tidur	11
2.4 Insomnia	14
2.4.1 Prevalensi Insomnia	14
2.4.2 Gejala dan Tanda-Tanda Klinis	15
2.4.3 Dampak Insomnia	15
2.5 Machine Learning	16
2.6 Decision Tree	18
2.7 Vector Autoregression (VAR)	19
2.8 Standard Scaler	20
2.9 Logistic Regression	21
2.10 Evaluasi: Confusion Matrix	22
BAB 3 METODOLOGI PENELITIAN	25
3.1 Metodologi Penelitian	25
3.1.1 Telaah Literatur	26
3.2 Analisis Kebutuhan	26
3.2.1 Pengumpulan Data (<i>Data Gathering</i>)	26
3.2.2 Pemahaman Data (<i>Data Understanding</i>)	27
3.3 Praproses Data (<i>Data Preprocessing</i>)	28
3.3.1 Eksplorasi Data (<i>Data Exploration</i>)	28
3.4 Perancangan Kode dan Model	30
3.5 Evaluasi Model	31
3.6 <i>Hyperparameter Tuning</i> dengan <i>RandomizedSearchCV</i>	31
3.6.1 Logistic Regression	31
3.6.2 Decision Tree	33

3.6.3	Hasil Tuning	35
3.7	Penentuan Fitur Paling Berpengaruh	35
BAB 4	HASIL DAN DISKUSI	37
4.1	<i>Testing Environment</i>	37
4.2	Perancangan Model	38
4.2.1	Import Library	38
4.2.2	Praproses Data (<i>Data Preprocessing</i>)	40
4.2.3	Perancangan Model	50
4.2.4	Hyperparameter Tuning	54
4.2.5	Evaluasi Model	55
4.3	Penentuan Fitur Paling Berpengaruh	56
BAB 5	SIMPULAN DAN SARAN	60
5.1	Simpulan	60
5.2	Saran	60
DAFTAR PUSTAKA		62



DAFTAR TABEL

Tabel 2.1	Tabel Penelitian Pendahuluan	6
Tabel 2.2	Siklus Tidur dalam Bentuk Tabel	11
Tabel 2.3	Dampak Insomnia Berdasarkan Aspeknya	15
Tabel 2.4	Tabel <i>Confusion Matrix</i>	22
Tabel 3.1	Data Kebiasaan Tidur serta Kualitas Tidur	27
Tabel 4.1	Spesifikasi Perangkat	37
Tabel 4.2	Distribusi Kelas Variabel Target	44
Tabel 4.3	Perbandingan Hasil Evaluasi Model	55



DAFTAR GAMBAR

Gambar 2.1	Siklus Tidur Manusia	10
Gambar 2.2	<i>Decission Tree Classifier</i>	18
Gambar 3.1	<i>Flowchart</i> Metodologi Penelitian	25
Gambar 3.2	Flowchart Proses Logistic Regression	33
Gambar 3.3	<i>Decission Tree Classifier</i>	34
Gambar 4.1	Tampilan Isi <i>Dataset</i>	41
Gambar 4.2	Distribusi Kelas Target <i>Insomnia</i>	43
Gambar 4.3	Barplot Skala Fitur Sebelum dan Sesudah Standardisasi	45
Gambar 4.4	Histogram Distribusi Fitur Sebelum dan Sesudah Standardisasi	46
Gambar 4.5	Histogram Distribusi Fitur Sebelum dan Sesudah Standardisasi	46
Gambar 4.6	Histogram Distribusi Fitur Sebelum dan Sesudah Standardisasi	46
Gambar 4.7	Pembagian Data <i>Train</i> , <i>Val</i> , dan <i>Test Split</i>	47
Gambar 4.8	Visualisasi Data <i>Training</i>	47
Gambar 4.9	Visualisasi Data <i>Validation</i>	48
Gambar 4.10	Visualisasi Data <i>Testing</i>	49
Gambar 4.11	<i>Classification Report</i> pada Data Validasi	52
Gambar 4.12	<i>Classification Report</i> pada Data Uji	53
Gambar 4.13	<i>Confusion Matrix</i> pada Data Validasi	56
Gambar 4.14	Visualisasi Struktur Pohon <i>Decision Tree</i>	56
Gambar 4.15	Perbandingan <i>Feature Importance</i>	58
Gambar 4.16	Grafik Perbandingan Kepentingan Fitur	58
Gambar 4.17	Plot Perbandingan Kepentingan Fitur	59



DAFTAR KODE

Kode 4.1	Proses <i>Import Library</i>	38
Kode 4.2	Mengunduh <i>Dataset</i> dari <i>Kaggle</i>	40
Kode 4.3	Menunjukkan Folder <i>Dataset</i> Berada	40
Kode 4.4	Membaca <i>Dataset</i>	40
Kode 4.5	Pembagian Data <i>Train, Validation, dan Test Split</i>	41
Kode 4.6	Augmentasi Load original data df	42
Kode 4.7	Membuat <i>Dataframe</i> <i>Insomnia</i>	44
Kode 4.8	Memproses <i>Dataframe</i> <i>Insomnia</i>	44
Kode 4.9	Proses Normalisasi Data	45
Kode 4.10	Proses <i>Training dan Validation</i> Model <i>Decision Tree</i>	50
Kode 4.11	Proses <i>Training dan Validation</i> Model <i>Logistic Regression</i>	50
Kode 4.12	Evaluasi Model Menggunakan Data <i>Validation</i>	50
Kode 4.13	<i>Feature Importance</i>	57
Kode 4.14	<i>Print Perbandingan Feature Importance</i>	57



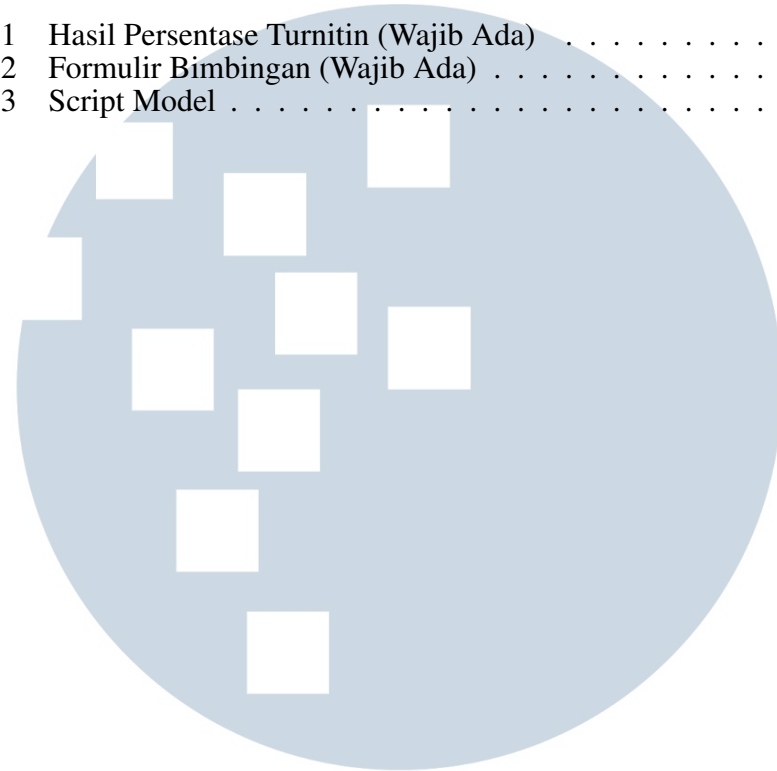
DAFTAR RUMUS

Rumus 2.1	<i>Gini Index</i>	19
Rumus 2.3	<i>Standard Scaler</i>	20
Rumus 2.4	<i>Mean</i>	21
Rumus 2.5	<i>Standar Deviasi</i>	21
Rumus 2.6	<i>Fungsi Sigmoid Logistic Regression</i>	22
Rumus 2.7	<i>Accuray</i>	23
Rumus 2.8	<i>Precision</i>	23
Rumus 2.9	<i>Recall</i>	23
Rumus 2.10	<i>F1-Score</i>	24



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin (Wajib Ada)	70
Lampiran 2	Formulir Bimbingan (Wajib Ada)	80
Lampiran 3	Script Model	83



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA