

**PERBANDINGAN ALGORITMA *RANDOM FOREST*,
GRADIENT BOOSTING, DAN *SUPPORT VECTOR
MACHINE* DALAM KLASIFIKASI PENYAKIT DEMAM
BERDARAH BERBASIS DATA HEMATOLOGI**

Tugas Akhir

**Diajukan Untuk Memenuhi Salah Satu Persyaratan Mencapai Derajat Sarjana
Komputer**



Dibuat oleh:

BAGUS PRAMONO AJI

210711394

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA YOGYAKARTA
2025**

HALAMAN PENGESAHAN

Tugas Akhir Berjudul
**PREDIKSI PENYAKIT DEMAM BERDARAH MENGGUNAKAN
RANDOM FOREST, GRADIENT BOOSTING, DAN SUPPORT
VECTOR MACHINE BERBASIS DATA HEMATOLOGI**

yang disusun oleh
BAGUS PRAMONO AJI
210711394

dinyatakan telah memenuhi syarat pada tanggal 28 Juli 2025

Keterangan

Dosen Pembimbing 1	: Herlina, S.Kom., M.Eng	Telah Menyetujui
Dosen Pembimbing 2	: Aloysius Gonzaga Pradnya Sidhawara, S.T., M.Eng.	Telah Menyetujui
Tim Penguji		
Penguji 1	: Herlina, S.Kom., M.Eng	Telah Menyetujui
Penguji 2	: Prof. Dr. Ir. Alb. Joko Santoso, M.T.	Telah Menyetujui
Penguji 3	: Joanna Ardhyanti Mita N, S.Kom., M.Kom	Telah Menyetujui

Yogyakarta, 28 Juli 2025

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. Ir. Parama Kartika Dewa SP., S.T., M.T.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH

Saya yang bertanda tangan di bawah ini :

Nama Lengkap : BAGUS PRAMONO AJI

NPM : 210711394

Program Studi : Informatika

Fakultas : Teknologi Industri

Judul Penelitian : PERBANDINGAN ALGORITMA *RANDOM FOREST*,
GRADIENT BOOSTING, DAN *SUPPORT VECTOR*
MACHINE DALAM KLASIFIKASI PENYAKIT DEMAM
BERDARAH BERBASIS DATA HEMATOLOGI

Menyatakan dengan ini:

1. Tugas Akhir ini adalah benar tidak merupakan Salinan sebagian atau keseluruhan dari karya penelitian lain.
2. Memberikan kepada Universitas Atma Jaya Yogyakarta atas penelitian ini, berupa Hak untuk menyimpan, mengelola, mendistribusikan, dan menampilkan hasil penelitian selama tetap mencantumkan nama penulis.
3. Bersedia menanggung secara pribadi segala bentuk tuntutan hukum atas pelanggaran Hak Cipta dalam pembuatan Tugas Akhir ini.

Demikian pernyataan ini dibuat dan dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 2025

Yang Menyatakan,



BAGUS PRAMONO AJI

210711394

HALAMAN PERSEMBAHAN

Dengan segala rasa syukur dan kerendahan hati, karya tugas akhir ini penulis persembahkan kepada pihak-pihak yang telah memberikan dukungan, doa, bimbingan, dan kasih sayang dalam setiap langkah perjalanan pendidikan penulis hingga akhirnya dapat menyelesaikan tugas akhir ini. Ucapan terima kasih yang tulus penulis haturkan kepada:

1. Ibu Nurwidyaningtyas, S.H., sosok ibu sekaligus ayah bagi penulis, yang dengan penuh perjuangan dan kasih sayang telah membimbing serta mendukung penulis dalam menuntut ilmu setinggi-tingginya.
2. Ibu Nanu Nuryanti, nenek tercinta penulis, yang dengan penuh kasih sayang telah merawat dan mendidik penulis sejak kecil.
3. Kepada Reka Natalia, selaku sahabat dalam berbagi pikiran, sekaligus penyemangat yang selalu setia mendampingi dan membimbing penulis dalam proses penyusunan tugas akhir ini.
4. Seluruh teman – teman penulis yang selalu membantu dan memberikan semangat selama penulis menjalani kegiatan perkuliahan.
5. Seluruh pihak – pihak lainnya yang telah memberikan dukungan dan doa sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Terakhir, untuk diri penulis sendiri yang telah berjuang tanpa lelah, bersabar dalam setiap ujian, dan terus melangkah demi mengukir kebanggaan serta mengangkat derajat keluarga tercinta. Semoga setiap tetes peluh dan langkah kecil ini menjadi pijakan menuju masa depan yang lebih baik.

” Kami pernah disitu, diposisimu, helakkan kesahmu ”

33x - Perunggu

KATA PENGANTAR

Puji dan syukur penulis haturkan kepada Tuhan Yang Maha Esa karena berkat rahmat dan karunia-Nya penulis dapat menyelesaikan pembuatan tugas akhir “*PERBANDINGAN ALGORITMA RANDOM FOREST, GRADIENT BOOSTING, DAN SUPPORT VECTOR MACHINE DALAM KLASIFIKASI PENYAKIT DEMAM BERDARAH BERBASIS DATA HEMATOLOGI*” ini dengan baik. Penulisan tugas akhir ini bertujuan memenuhi salah satu syarat untuk mencapai derajat sarjana computer dari Program Studi Informatika, Fakultas Teknologi Industri di Universitas Atma Jaya Yogyakarta.

Penulis menyadari bahwa dalam pembuatan tugas akhir ini penulis telah mendapatkan bantuan, bimbingan, dan dorongan dari banyak pihak. Untuk itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada:

7. Bapak Dr. Ir. Parama Kartika Dewa SP., S.T., M.T. IPU, selaku Dekan Fakultas Teknologi Industri, Universitas Atma Jaya Yogyakarta.
8. Bapak Thomas Adi Purnomo Sidhi, S.T., M.T., selaku Kepala Program Studi Informatika, Universitas Atma Jaya Yogyakarta.
9. Ibu Herlina S.Kom., M.Eng., selaku dosen pembimbing I yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.
10. Bapak Aloysius Gonzaga Pradnya S., S.T., M.Eng., selaku dosen pembimbing II yang telah membimbing dan memberikan masukan serta motivasi kepada penulis untuk menyelesaikan tugas akhir ini.

Demikian laporan tugas akhir ini dibuat, dan penulis ucapkan terima kasih kepada semua pihak. Semoga laporan ini dapat bermanfaat bagi pembaca.

Yogyakarta, 2025



BAGUS PRAMONO AJI
210711394

DAFTAR ISI

LEMBAR PENGESAHAN	ii
PERNYATAAN ORISINALITAS & PUBLIKASI ILMIAH.....	ii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR	ix
DAFTAR KODE.....	x
DAFTAR TABEL.....	xi
INTISARI.....	xii
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Rumusan Masalah	2
C. Batasan Masalah	2
D. Tujuan Penelitian	4
E. Metode Penelitian	4
F. Sistematika Penulisan	6
BAB II TINJAUAN PUSTAKA.....	9
BAB III LANDASAN TEORI.....	18
A. Demam Berdarah <i>Dengue</i> (DBD).....	18
B. Hematologi.....	18
C. <i>Machine learning</i>	19
D. <i>Classification</i>	22
E. <i>Preprocessing</i>	22
1. <i>Data Cleaning</i>	23
2. <i>Data Scaling</i>	23
3. <i>Data Balancing</i>	24

F. <i>Feature Selection</i>	26
G. Pipeline.....	28
H. <i>Random Forest</i>	28
I. <i>Gradient Boosting</i>	30
J. Support Vector Machine (SVM).....	32
K. <i>Grid Search Cross Validation</i>	36
L. Metrik Evaluasi	38
BAB IV DATASET DAN PENGEMBANGAN MODEL.....	41
A. Deskripsi Problem.....	41
B. Dataset.....	41
1. Penyesuaian Dataset	43
2. Preprocessing.....	47
C. Pengembangan Model.....	51
1. Pelatihan menggunakan <i>Random Forest</i>	51
2. Pelatihan menggunakan <i>Gradient Boosting</i>	56
3. Pelatihan menggunakan <i>Support Vector Machine</i>	62
D. Hasil Pelatihan Model	66
1. Hasil Pelatihan Model <i>Random Forest</i>	66
2. Hasil Pelatihan Model <i>Gradient Boosting</i>	70
3. Hasil Pelatihan Model SVM.....	74
BAB V IMPLEMENTASI MODEL DAN HASIL	83
A. Deskripsi Penelitian	83
B. Implementasi Model kedalam Aplikasi Website	83
1. Pemuatan Model.....	84
2. Tampilan Antarmuka.....	84
3. Validasi Input	88
4. Proses Prediksi dan Hasil Prediksi	89
5. Uji coba menggunakan data baru	90
BAB VI PENUTUP	97
A. Kesimpulan	97
B. Saran.....	98

DAFTAR PUSTAKA	99
----------------------	----

DAFTAR GAMBAR

Gambar 1.1 Alur Penelitian	5
Gambar 3.1 Ilustrasi <i>Supervised Learning</i> (A) dan <i>Unsupervised Learning</i> (B)	20
Gambar 3.2 Ilustrasi <i>Reinforcement Learning</i>	21
Gambar 3.3 Ilustrasi metode <i>filter</i> , <i>wrapper</i> , dan <i>embedded</i>	27
Gambar 3.4 Ilustrasi <i>K-Fold Cross Validation</i>	36
Gambar 3.5 Ilustrasi menggunakan <i>GridSearchCV</i>	37
Gambar 3.6 <i>Confusion Matrix</i> pada <i>Multiclass</i> (tiga Kelas)	38
Gambar 4.1 Alur penyesuaian dataset	43
Gambar 4.2 Alur Preprocessing dataset	47
Gambar 5.1 Antarmuka inputan dalam <i>website</i>	87
Gambar 5.2 Tampilan Keseluruhan Antar Muka	88
Gambar 5.3 Validasi Inputan masih Kosong	89
Gambar 5.4 Pengujian pada data non-dbd 1	92
Gambar 5.5 Pengujian pada data non-dbd 2	93
Gambar 5.6 Pengujian pada data DF-1	94
Gambar 5.7 Pengujian pada data DF-2	94
Gambar 5.8 Pengujian terhadap data DHF-1	95
Gambar 5.9 Pengujian terhadap data DHF-2	96

DAFTAR KODE

Kode 4.1 Import library, memuat dataset, dan pemeriksaan dataset	44
Kode 4.2 Kode pengecekan Nan dan Duplikat lalu di drop	44
Kode 4.3 Penyesuaian Dataset menggunakan parameter WHO.....	46
Kode 4.4 Potongan kode pengecekan terhadap data duplikat dan Nan.....	48
Kode 4.5 Encode pada fitur Gender	48
Kode 4.6 Potongan kode pemisahan fitur dan target.....	49
Kode 4.7 Potongan kode split data training dan testing	49
Kode 4.8 Kode untuk visualisasi pengecekan keseimbangan data.....	49
Kode 4.9 Kode resampling menggunakan ADASYN	50
Kode 4.10 kode Pipeline dengan GridSearchCV untuk <i>Random Forest</i> Skenario 1	54
Kode 4.11 kode Pipeline dengan GridSearchCV untuk <i>Random Forest</i> Skenario 2	55
Kode 4.12 Kode Pipeline dengan GridSearchCV untuk <i>GradientBoosting</i> Skenario 1.....	59
Kode 4.13 Kode Pipeline dengan GridSearchCV untuk <i>GradientBoosting</i> Skenario 2.....	60
Kode 4.14 Kode Pipeline dengan GridSearchCV untuk SVC Skenario 1	64
Kode 4.15 Kode Pipeline dengan GridSearchCV untuk SVC Skenario 2	65
Kode 5.1 Kode untuk mengunduh model.....	84
Kode 5.2 Kode Mengimport ke <i>Streamlit</i>	84
Kode 5.3 Kode Pembuatan Inputan pada <i>Streamlit</i>	86
Kode 5.4 Kode validasi setiap inputan	89
Kode 5.5 <i>Encode Gender</i> menjadi Numerik.....	89
Kode 5.6 Model memprediksi sesuai Inputan User.....	90

DAFTAR TABEL

Tabel 2.1 Tabel Perbandingan Tinjauan Pustaka	16
Tabel 3.1 Tabel Perbandingan DF & DHF berdasarkan WHO (2011)	19
Tabel 4.1 Keterangan Dataset Hematologi	42
Tabel 4.2 Kolom <i>Result</i> sebelum disesuaikan	46
Tabel 4.3 Kolom <i>Result</i> setelah disesuaikan	46
Tabel 4.4 Distribusi kelas sebelum & sesudah resampling	50
Tabel 4.5 Perbandingan Skenario 1 dan Skenario 2 untuk <i>Random Forest</i>	52
Tabel 4.6 Perbandingan Skenario 1 dan Skenario 2 untuk <i>Gradient Boosting</i>	57
Tabel 4.7 Tabel Perbandingan Skenario 1 dan Skenario 2 untuk SVM	62
Tabel 4.8 Tabel Model Evaluasi Skenario 1 & 2 untuk <i>Random Forest</i>	68
Tabel 4.9 Tabel <i>Confusion Matrix Random Forest</i> Skenario 1	69
Tabel 4.10 Tabel <i>Confusion Matrix Random Forest</i> Skenario 2	69
Tabel 4.11 Tabel Model Evaluasi Skenario 1 & 2 untuk <i>Gradient Boosting</i>	71
Tabel 4.12 Tabel <i>Confusion Matrix Gradient Boosting</i> Skenario 1	72
Tabel 4.13 Tabel <i>Confusion Matrix Gradient Boosting</i> Skenario 2	73
Tabel 4.14 Tabel Model Evaluasi Skenario 1 & 2 untuk SVM	75
Tabel 4.15 Tabel <i>Confusion Matrix SVM</i> Skenario 1	76
Tabel 4.16 Tabel <i>Confusion Matrik SVM</i> Skenario 2	77
Tabel 4.17 Hasil Keseluruhan Pengujian Skenario Model	79
Tabel 5.1 Tabel Pengujian Data Acak	90

INTISARI

PERBANDINGAN ALGORITMA *RANDOM FOREST*, *GRADIENT BOOSTING*, DAN *SUPPORT VECTOR MACHINE* DALAM KLASIFIKASI PENYAKIT DEMAM BERDARAH BERBASIS DATA HEMATOLOGI

Intisari

BAGUS PRAMONO AJI

210711394

Demam Berdarah *Dengue* (DBD) merupakan penyakit menular yang masih menjadi tantangan kesehatan di Indonesia. Proses diagnosis DBD umumnya masih dilakukan secara manual oleh tenaga medis berdasarkan hasil pemeriksaan hematologi, yang rentan terhadap kesalahan interpretasi dan membutuhkan waktu lama. Deteksi dini sangat penting untuk mencegah komplikasi serius seperti perdarahan internal dan kebocoran plasma.

Penelitian dilakukan dengan memodifikasi dataset hematologi menjadi tiga kelas dan menerapkan teknik *oversampling ADASYN*(*Adaptive Synthetic Sampling Approach*) untuk mengatasi ketidakseimbangan data. Penelitian ini bertujuan mengembangkan model klasifikasi DBD berbasis machine learning untuk membedakan Non-DBD, *Dengue Fever* (DF), dan *Dengue Hemorrhagic Fever* (DHF) menggunakan data hematologi (trombosit, leukosit, hemoglobin, hematokrit). Tiga algoritma dibandingkan, yaitu *Random Forest*, *Gradient Boosting*, dan *Support Vector Machine* (SVM).

Hasil penelitian menunjukkan *Random Forest* pada skenario 1 memiliki kinerja terbaik dengan akurasi 0,73 dan F1-score 0,82 untuk DF, 0,77 untuk DHF, serta 0,46 untuk Non-DBD. Algoritma ini direkomendasikan untuk diagnosis dini DBD berbasis data hematologi.

Kata kunci: Demam Berdarah *Dengue*, Hematologi, Klasifikasi DBD, *Machine Learning*, *Random Forest*.

Dosen Pembimbing I : Ibu Herlina S.Kom., M.Eng.
Dosen Pembimbing II : Bapak Aloysius Gonzaga Pradnya S., S.T., M.Eng.
Jadwal Sidang Tugas Akhir : Jumat, 25 Juli 2025