

**KLASIFIKASI PENYAKIT DEMAM BERDARAH
MENGUNAKAN ALGORITMA STACKING
ENSEMBLE LEARNING**

TESIS

Oleh

I KADEK YOGI PRAYOGA

NIM 2129101052



**PROGRAM STUDI ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS PENDIDIKAN GANESHA
2025**

KLASIFIKASI PENYAKIT DEMAM BERDARAH
MENGUNAKAN ALGORITMA STACKING
ENSEMBLE LEARNING



PROGRAM STUDI ILMU KOMPUTER
PROGRAM PASCASARJANA
UNIVERSITAS PENDIDIKAN GANESHA
2025

LEMBAR PERSETUJUAN PEMBIMBING

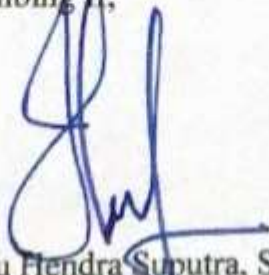
Tesis oleh I Kadek Yogi Prayoga telah diperiksa dan disetujui untuk mengikuti Ujian Tesis.

Singaraja, Juni 2025
Pembimbing I,



Dr. I Made Gede Sunarya, S.Kom., M.Cs.
NIP. 196710131994031001

Pembimbing II,



Dr. Putu Hendra Suputra, S.Kom., M.Cs.
NIP. 198212222006041001

LEMBAR PERSETUJUAN TIM PENGUJI

Tesis oleh I Kadek Yogi Prayoga ini telah dipertahankan di depan tim penguji dan dinyatakan diterima sebagai salah satu persyaratan untuk memperoleh gelar Magister Pendidikan di Program Studi Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha.

Disetujui pada tanggal: Juni 2025

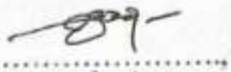
oleh
Tim Penguji



Ketua (Dr. I Made Gede Sunarya, S.Kom., M.Cs.)
NIP. 196710131994031001



Anggota (Dr. I Nyoman Sukajaya, M.T.)
NIP. 196711151993031001



Anggota (Dr. Gede Suweken, M.Sc.)
NIP. 196111111987021001



Anggota (Dr. Putu Hendra Suputra, S.Kom., M.Cs.)
NIP. 198212222006041001

Mengetahui Direktur
Program Pascasarjana Undiksha



Prof. Dr. I Nyoman Jampel, M.Pd.
NIP. 195910101986031003

LEMBAR PERNYATAAN

Saya menyatakan dengan sesungguhnya bahwa tesis yang saya susun sebagai syarat untuk memperoleh gelar Magister Pendidikan dari Pascasarjana Universitas Pendidikan Ganesha seluruhnya merupakan hasil karya saya sendiri. Bagian-bagian tertentu dalam penulisan tesis yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas dan sesuai dengan norma, kaidah, serta etika akademis.

Apabila di kemudian hari ditemukan seluruh atau sebagian tesis ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku di wilayah Negara Kesatuan Republik Indonesia.

Singaraja, Juni 2025
Penulis



(I Kadek Yogi Prayoga)

PRAKATA

Puji dan syukur penulis panjatkan kehadapan Tuhan Yang Maha Esa atas anugerah-Nya, sehingga tesis yang berjudul “Klasifikasi Penyakit Demam Berdarah Menggunakan Algoritma Stacking Ensemble Learning”, dapat diselesaikan sesuai dengan yang direncanakan.

Tesis ini disusun untuk memenuhi salah satu persyaratan untuk memperoleh gelar Magister Pendidikan Pascasarjana Universitas Pendidikan Ganesha pada program studi Ilmu Komputer. terselesaikannya tesis ini telah banyak memperoleh uluran tangan dari berbagai pihak. Untuk itu, ijin penulis menyampaikan terima kasih dan penghargaan kepada pihak-pihak berikut.

1. Dr. I Made Gede Sunarya, S.Kom., M.Cs. selaku Pembimbing I yang telah memberikan bimbingan dan arahan sehingga penulis tesisi ini dapat diselesaikan sesuai dengan harapan.
2. Dr. Putu Hendra Suputra, S.Kom., M.Cs. selaku Pembimbing II yang telah memberikan motivasi, semangat, bimbingan, dan saran kepada penulis sehingga mampu melewati berbagai hambatan dalam penyelesaian tesis ini.
3. dr. I Made Ariyoga Budiana selaku mitra penelitian di tempat melakukan penelitian yang telah banyak membantu penulis dalam melakukan penelitian.
4. Kepala Puskesmas I Kecamatan Mengwi yang telah memberikan izin kepada penulis untuk melakukan penelitian di Puskesmas.
5. Koordinator program studi Ilmu Komputer dan staf dosen pengajar yang telah banyak membantu dan memotivasi penulis dalam penyusunan ini.
6. Direktur Pascasarjana Undiksha dan staf, yang telah membantu penulis dalam menyelesaikan tesis ini.
7. Rektor Universitas Pendidikan Ganesha, yang telah memberikan bantuan secara moral dan memfasilitasi berbagai kepentingan penulis dalam menyelesaikan tesis ini.

8. Rekan-rekan seangkatan di Program Studi Ilmu Komputer yang telah banyak berkontribusi membentuk kemandirian penulis selama menjalani studi dan menyelesaikan tesis ini.
9. Bapak I Nyoman Rastika dan Ibu Ni Ketut Murni selaku orang tua penulis, serta saudara-saudara dan keluarga yang telah banyak membantu secara material dan moral.

Semoga semua bantuan yang telah mereka berikan dalam menyelesaikan studi ini, mereka diberkati imbalan yang sepadan oleh Tuhan Yang Maha Esa, kesehatan, dan keharmonian dalam menjalani kehidupan.

Penulis menyadari bahwa tesis ini belum sempurna. Namun, kehadirannya dalam konstelasi masyarakat akademis akan menambah perbendaharaan ilmu dalam perkembangan ilmu pengetahuan. Semoga tesis ini bermanfaat bagi masyarakat akademis, terutama mereka yang menyatakan diri bernaung di bawah kebesaran panji-panji pendidikan.

Singaraja, Juni 2025

Penulis

ABSTRAK

Prayoga, I Kadek Yogi (2025), Klasifikasi Penyakit Demam Berdarah Menggunakan Algoritma Stacking Ensemble Learning. Tesis, Ilmu Komputer, Program Pascasarjana, Universitas Pendidikan Ganesha

Tesis ini telah disetujui dan diperiksa oleh Pembimbing I: Dr. I Made Gede Sunarya, S.Kom., M.Cs. dan Pembimbing II: Dr. Putu Hendra Suputra, S.Kom., M.Cs.

Kata-kata kunci: demam berdarah, *decision tree*, *naïve naves*, *KNN*, *stacking ensemble learning*

Ensemble Learning digunakan untuk meningkatkan kinerja klasifikasi dengan mengintegrasikan beberapa metode klasifikasi. Ada banyak teknik yang terkait dengan metode ini, seperti *Stacking*, *Boosting* dan *Bagging*. Dalam *Stacking Ensemble Learning*, *stacking* adalah teknik yang menggabungkan beberapa metode dasar (*base-learner*) yang dilatih untuk menghasilkan hasil akhir menjadi metode akhir (*meta-learner*). Tujuan dari penelitian ini adalah untuk mengevaluasi efektivitas klasifikasi menggunakan pengklasifikasi singular dibandingkan dengan *Stacking Ensemble Learning*. Penelitian ini menerapkan banyak prosedur, termasuk pengumpulan data, prapemrosesan data, pemodelan, dan evaluasi. Data yang akan digunakan diperoleh dari rekam medis pasien di Puskesmas I, Kecamatan Mengwi. Selanjutnya, data tersebut menjalani pemrosesan tambahan (*preprocessing*), membuang data yang memiliki nilai kosong. Data tersebut kemudian diperiksa untuk mengidentifikasi ketidakkonsistenan dan memperbaiki kesalahan, termasuk kesalahan ketik dan ketidakkonsistenan format. Teknik yang telah ditentukan sebelumnya digunakan untuk melakukan tahap pemodelan, dan hasilnya kemudian dinilai menggunakan *confusion-matrix* sebagai alat ukur. Metode *Stacking Ensemble Learning* digunakan dalam proses klasifikasi, dan selanjutnya dinilai menggunakan *confusion-matrix*. Pengujian ini menggunakan dua skenario: Klasifikasi menggunakan *single-classifier* dan *Stacking Ensemble Learning* dengan Logistic Regression. *Cross Validation* digunakan untuk langsung menerapkan algoritma ini pada dataset, dengan nilai k-fold 5 dan 10. Hasil evaluasi menunjukkan bahwa metode *Stacking Ensemble Learning* telah meningkat sebesar 8,89% dari metode *single-classifier* terendah, *Decision tree* dengan k-fold 5, dan sebesar 11,08% dari metode *single-classifier* terendah, *Decision tree* dengan k-fold 10, pada dataset 650 data.

ABSTRACT

Prayoga, I Kadek Yogi (2025), *Classification of Dengue Fever Disease Using Stacking Ensemble Learning Algorithm*. Thesis, Computer Science, Master Program, Ganesha University of Education.

This thesis has been approved and examined by Supervisor I: Dr. I Made Gede Sunarya, S.Kom., M.Cs., and Supervisor II: Dr. Putu Hendra Suputra, S.Kom., M.Cs..

Keywords: dengue fever, decision tree, naïve naves, KNN, stacking ensemble learning

Ensemble Learning is used to improve classification performance by integrating multiple classification methods. There are many techniques associated with this method, such as Stacking, Boosting and Bagging. In Stacking Ensemble Learning, stacking is a technique that combines several base methods (base-learners) that are trained to produce the final result into the final method (meta-learner). The purpose of this study is to evaluate the effectiveness of classification using a singular classifier compared to Stacking Ensemble Learning. This study applies many procedures, including data collection, data preprocessing, modeling, and evaluation. The data to be used were obtained from patient medical records at Puskesmas I, Mengwi District. Furthermore, the data undergoes additional processing (preprocessing), discarding data that has empty values. The data is then checked to identify inconsistencies and correct errors, including typos and format inconsistencies. A predetermined technique is used to perform the modeling stage, and the results are then assessed using a confusion matrix as a measuring tool. The Stacking Ensemble Learning method is used in the classification process and is then assessed using a confusion matrix. This test uses two scenarios: Classification using a single-classifier and Stacking Ensemble Learning with Logistic Regression. Cross Validation is used to directly apply this algorithm to the dataset, with k-fold values of 5 and 10. The evaluation results show that the Stacking Ensemble Learning method has increased by 8.89% from the lowest single-classifier method, Decision tree with k-fold 5, and by 11.08% from the lowest single-classifier method, Decision tree with k-fold 10, on a dataset of 650 data.

DAFTAR ISI

	<i>(halaman)</i>
HALAMAN SAMPUL	i
LEMBAR PERSETUJUAN PEMBIMBING	iii
LEMBAR PERSETUJUAN TIM PENGUJI	iv
LEMBAR PERNYATAAN	v
PRAKATA	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABLE	xiii
DAFTAR GAMBAR	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Batasan Masalah	4
1.4 Rumusan Masalah	4
1.5 Tujuan Penelitian	5
1.6 Manfaat Penelitian	5
BAB II KAJIAN PUSTAKA	6
2.1 Kajian Teori	6
2.1.1 Demam Berdarah	6
2.1.2 Metode Naïve Bayes	10
2.1.3 Metode K-Nearest Neighbor (KNN)	12
2.1.4 Metode Decision Tree	13
2.1.5 Metode Logistic Regression	15
2.1.6 Ensemble Learning	16
2.1.7 Stacking	16
2.1.8 Confusion Matrix	18

2.2	Tinjauan Hasil Penelitian yang Berkaitan	21
2.3	Dasar Pemikiran	27
BAB III METODE PENELITIAN.....		29
3.1	Pengumpulan Data.....	29
3.2	Data Preprocessing	29
3.3	Modeling	33
3.1.1	Konversi Atribut Data Training dan Testing	35
3.1.2	Perhitungan Algoritma Naïve Bayes.....	36
3.1.3	Perhitungan Algoritma KNN (K-Nearest Neighbor)	43
3.1.4	Perhitungan Algoritma Decision Tree	48
3.1.5	Perhitungan Algoritma Stacking Ensemble Learning.....	56
3.4	Evaluasi	66
3.5	Jadwal Penelitian.....	68
BAB IV HASIL DAN PEMBAHASAN		69
4.1	Data Penelitian.....	69
4.2	Data Preprocessing	70
4.3	Justifikasi atau Landasan Pemilihan dari Masing – Masing Base Learner 73	
4.3.1	Kelebihan dan Kekurangan dari Base Learner yang Digunakan	74
4.4	Hasil Uji	83
4.5	Detail Eksperimen dengan Metode Single Classifier dan Stacking Ensemble Learning dengan Cross Validation	87
4.4.1	Pengujian Menggunakan K-Fold 5	87
4.4.2	Pengujian Menggunakan K-Fold 10	91
4.5	Hasil Eksperimen Menggunakan Kombinasi antara Base Learner	95
4.5.1	Pengujian Menggunakan 450 Data	96
4.5.2	Pengujian Menggunakan 650 Data	101
4.6	Pengujian Korelasi antar Fitur.....	107
BAB V PENUTUP.....		113
5.1	Kesimpulan.....	113
5.2	Saran	114

DAFTAR PUSTAKA	115
LAMPIRAN.....	123



DAFTAR TABLE

Tabel 2. 1 Confusion matrix.....	18
Tabel 2. 2 Penelitian Terkait	21
Tabel 3. 1 Segmentasi Atribut.....	30
Tabel 3. 2 Detail Kriteria	35
Tabel 3. 3 Perhitungan Jumlah Data pada Setiap Kriteria	36
Tabel 3. 4 Data Testing	39
Tabel 3. 5 Data Nilai Kriteria.....	39
Tabel 3. 6 Hasil Perhitungan Probabilitas Parsial Negative	41
Tabel 3. 7 Hasil Perhitungan Logaritma Natural untuk Probabilitas Parsial Negatif	41
Tabel 3. 8 Hasil Perhitungan Logaritma Natural dari Probabilitas Parsial Positive	42
Tabel 3. 9 Hasil Perhitungan Normalisasi Probabilitas	43
Tabel 3. 10 Konversi Nilai Fitur	44
Tabel 3. 11 Konversi Data Testing	45
Tabel 3. 12 Data Uji KNN	46
Tabel 3. 13 Jarak Data Uji dengan Data Latih	47
Tabel 3. 14 Jarak Tetangga Terdekat	48
Tabel 3. 15 Data Latih Decision Tree	48
Tabel 3. 16 Decision Tree Node Pertama	51
Tabel 3. 17 Dataset Setelah Perhitungan Node Pertama.....	53
Tabel 3. 18 Decision Tree Perhitungan Node Selanjutnya	54
Tabel 3. 19 Decision Tree Node Terakhir.....	55
Tabel 3. 20 Data Latih Stacking Ensemble Learning	56
Tabel 3. 21 Pembobotan awal	57
Tabel 3. 22 Hasil iterasi pertama.....	57
Tabel 3. 23 Hasil iterasi kedua.....	60
Tabel 3. 24 Pembobotan iterasi 2	62
Tabel 3. 25 Hasil iterasi ketiga.....	63
Tabel 3. 26 Data Uji	65
Tabel 3. 27 Pembobotan iterasi 3.....	66
Tabel 3. 28 Confusion Matrix	66
Tabel 3. 29 Jadwal Penelitian.....	68
Tabel 4. 1 Informasi mengenai penyakit DBD dari Puskesmas Mengwi	69
Tabel 4. 2 Konversi Atribut umur ke Skala Interval	71
Tabel 4. 3 Konversi Atribut sistole ke Skala Interval	71
Tabel 4. 4 Konversi Atribut diastole ke Skala Interval.....	72
Tabel 4. 5 Konversi Atribut trombosit ke Skala Interval	72

Tabel 4. 6 Konversi Atribut leukosit ke Skala Interval.....	72
Tabel 4. 7 Konversi Atribut hematokrit ke Skala Interval	73
Tabel 4. 8 Alasan Pemilihan Base Learner	77
Tabel 4. 9 Pengujian 450 Data dengan Cross Validation.....	83
Tabel 4. 10 Pengujian 650 Data dengan Cross Validation.....	85
Tabel 4. 11 Hasil Cross Validation dengan K-Fold 5 Metode Single Classifier KNN	87
Tabel 4. 12 Hasil Cross Validation dengan K-Fold 5 Metode Single Classifier Naive Bayes	88
Tabel 4. 13 Hasil Cross Validation dengan K-Fold 5 Metode Single Classifier Decision Tree	89
Tabel 4. 14 Hasil Cross Validation dengan K-Fold 5 Metode Single Classifier Logistic Regression.....	89
Tabel 4. 15 Hasil Cross Validation dengan K-Fold 5 Metode Stacking Ensemble Learning	90
Tabel 4. 16 Hasil Cross Validation dengan K-Fold 10 Metode Single Classifier KNN.....	91
Tabel 4. 17 Hasil Cross Validation dengan K-Fold 10 Metode Single Classifier Naive Bayes	92
Tabel 4. 18 Hasil Cross Validation dengan K-Fold 10 Metode Single Classifier Decision Tree	93
Tabel 4. 19 Hasil Cross Validation dengan K-Fold 10 Metode Single Classifier Logistic Regression.....	93
Tabel 4. 20 Hasil Cross Validation dengan K-Fold 10 Metode Stacking Ensemble Learning	94
Tabel 4. 21 Pengujian Kombinasi KNN dengan Naive Bayes Menggunakan 450 Data	96
Tabel 4. 22 Pengujian Kombinasi KNN dengan Decision Tree Menggunakan 450 Data	97
Tabel 4. 23 Pengujian Kombinasi Naive Bayes dengan Decision Tree Menggunakan 450 Data	99
Tabel 4. 24 Pengujian Kombinasi KNN dengan Naive Bayes Menggunakan 650 Data	101
Tabel 4. 25 Pengujian Kombinasi KNN dengan Decision Tree Menggunakan 650 Data	103
Tabel 4. 26 Pengujian Kombinasi Naive Bayes dengan Decision Tree Menggunakan 650 Data	105
Tabel 4. 27 Pasangan Fitur Perhitungan Korelasi.....	108
Tabel 4. 28 Sampel Data Untuk Perhitungan Nilai Koefisien Korelasi.....	109
Tabel 4. 29 Nilai Selisih dan Perkalian antar Fitur	110

Tabel 4. 30 Fitur dengan Korelasi Tertinggi Terhadap Class	111
--	-----



DAFTAR GAMBAR

Gambar 2. 1 Proses Stacking Ensemble Learning	18
Gambar 2. 2 Kerangka Berpikir	27
Gambar 3. 1 Data Sebelum Pemilahan	32
Gambar 3. 2 Contoh dataset demam berdarah	33
Gambar 3. 3 Modeling Stacking Ensemble Learning	34
Gambar 4. 1 Heatmap Korelasi antar Fitur	111



DAFTAR LAMPIRAN

Lampiran 1 Grafik Pengujian Fold 5	124
Lampiran 2 Grafik Pengujian Fold 10	124
Lampiran 3 Hasil Heatmap Korelasi Antar Fitur	125

