

**COMPARISON OF ENSEMBLE METHODS FOR DECISION
TREE MODELS IN CLASSIFYING E. COLI BACTERIA**

JALUR SCIENTIST

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Sistem Informasi



disusun oleh

ALVIN RAHMAN AL MUSYAFFA

21.12.1844

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**COMPARISON OF ENSEMBLE METHODS FOR DECISION
TREE MODELS IN CLASSIFYING E. COLI BACTERIA**

JALUR NON REGULER - SCIENTIST

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1 Sistem Informasi



disusun oleh

ALVIN RAHMAN AL MUSYAFFA

21.12.1844

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

JALUR NON REGULER - SCIENTIST

COMPARISON OF ENSEMBLE METHODS FOR DECISION TREE
MODELS IN CLASSIFYING E. COLI BACTERIA

yang disusun dan diajukan oleh

Alvin Rahman Al Musyaffa

21.12.1844

telah disetujui oleh Dosen Pembimbing
pada tanggal 22 Januari 2025

Dosen Pembimbing,



Yoga Pristyanto, S.Kom, M.Eng

NIK. 190302412

HALAMAN PENGESAHAN

JALUR SCIENTIST

**COMPARISON OF ENSEMBLE METHODS FOR DECISION TREE
MODELS IN CLASSIFYING E. COLI BACTERIA**

yang disusun dan diajukan oleh

Alvin Rahman Al Musyaffa

21.12.1844

Telah dipertahankan di depan Dewan Penguji
pada tanggal 12 Februari 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Irma Rofni Wulandari, S.Pd., M.Eng.
NIK. 190302329

Ika Nur Fajri, S.Kom., M.Kom.
NIK. 190302268

Yoga Pristyanto, S.Kom., M.Eng.
NIK. 190302412



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 12 Februari 2025

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom. Ph.D
NIK. 190302096

HALAMAN PERNYATAAN KONTRIBUSI KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Alvin Rahman Al Musyaffa
NIM : 21.12.1844

Menyatakan bahwa karya dengan judul berikut:

COMPARISON OF ENSEMBLE METHODS FOR DECISION TREE MODELS IN CLASSIFYING E. COLI BACTERIA

Dosen Pembimbing : Yoga Pristyanto, S.Kom, M.Eng

1. Karya adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya merupakan gagasan, rumusan maupun penelitian yang orisinal dan SAYA memiliki KONTRIBUSI terhadap karya tersebut.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka atau Referensi pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 22 Januari 2025

Yang Menyatakan,



Alvin Rahman Al Musyaffa

HALAMAN PERSEMBAHAN

Puji Syukur kehadiran Allah SWT atas Rahmat dan hidayah-Nya sehingga penulis diberi kesempatan untuk dapat menyelesaikan skripsi ini. Dengan penuh rasa Syukur, penulis persembahkan skripsi ini kepada orang-orang yang selalu memberikan dukungan penuh kepada penulis.

Skripsi ini penulis persembahkan kepada orang tua penulis, Ibu Walyati dan Bapak Sri Agung atas segala lantunan doa, dukungan moral, dan materiil yang tiada henti. Terimakasih atas segala bentuk perjuangan, pengorbanan, serta kasih sayang yang selalu diberikan kepada penulis hingga saat ini.

Selain itu skripsi ini penulis persembahkan kepada kakak tercinta, Nurul Eka Hidayah, dan seluruh keluarga besar yang selalu memberikan semangat, kebahagiaan, dan dukungan tanpa batas dalam setiap perjalanan hidup penulis.

Skripsi ini penulis persembahkan kepada diri penulis sendiri yang telah berusaha tidak menyerah dan senantiasa menikmati proses perkuliahan dari awal hingga akhir. *For every late night, every moment of doubt, and every small victory along the way I celebrate the strength, perseverance, and passion that have brought me to this milestone.*

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir ini dengan lancar. Tugas akhir ini disusun untuk memenuhi salah satu syarat kelulusan Program Sarjana pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Selain itu penulis dengan segala kerendahan hati ingin menyampaikan rasa terima kasih kepada semua pihak yang telah berjasa memberikan dukungan dan bantuan untuk menyelesaikan laporan tugas akhir ini. Untuk itu penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M. selaku Rektor Universitas Amikom Yogyakarta
2. Bapak Hanif Al Fatta, S.Kom., M.Kom. selaku Dekan Program Fakultas Ilmu Komputer
3. Bapak Anggit Dwi Hartanto., M.Kom. selaku Ketua Program Studi Sistem Informasi
4. Bapak Yoga Pristyanto, S.Kom, M.Eng. selaku dosen pembimbing yang memberikan arahan, saran, dan motivasi terhadap penulis
5. Ibu Acihmah Sidauruk, M.Kom selaku dosen wali yang turut membantu dalam proses perkuliahan penulis.
6. Seluruh dosen Universitas Amikom Yogyakarta, yang telah memberikan ilmu, pengetahuan, dan dukungan yang sangat berarti selama penulis menempuh pendidikan di Universitas Amikom Yogyakarta.
7. Orang tua, kakak, dan keluarga besar, yang selalu memberikan dukungan moril, materiil, serta doa yang tiada henti, sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Terima kasih atas segala cinta, perhatian, dan semangat yang telah diberikan selama ini.
8. Nia Mauliza, Aisha Shakila, Faza Dhika, Bagas Ari, David Nur, Jibrán Tauf, Andi Salam, Nabrasca, Ferry Agung, Anggito, I Wayan Aditya,

Isma Magfirotul, Ahda Zulfaa, Yuditya Astri yang telah memberikan dukungan, semangat, dan kebersamaan. Terima kasih atas persahabatan dan kerja sama yang luar biasa.

9. Kepada kontrakan syahdu Faza dan Bagas yang sangat membantu proses penulis dalam membuat tugas akhir ini, karena kontrakan ini tempat berkumpul penulis untuk menemukan ide atau sekedar meringankan pikiran dengan bermain bersama.
10. Seluruh teman kelas 21SI01 yang tidak bisa disebutkan satu persatu yang telah memberikan banyak dukungan dan bantuan kepada penulis selama mengerjakan tugas akhir. Terimakasih juga telah membersamai perjalanan akademik penulis.
11. Semua pihak yang telah memberikan dukungan, bantuan, dan semangat sepanjang proses penyusunan skripsi ini.
12. Untuk diri sendiri, yang sudah berjuang dan tidak menyerah meskipun banyak tantangan. Terima kasih sudah tetap semangat dan terus maju sampai akhirnya tugas akhir ini bisa selesai.

Yogyakarta, 22 Januari 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KONTRIBUSI KARYA	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
INTISARI	ix
ABSTRACT	x
BAB I IDENTITAS PUBLIKASI	1
BAB II PROSES SUBMIT	2
2.1 Lembar Review	2
2.2 Review Round 1	3
2.3 Review Round 2	6
2.4 Review Round 3	8
2.5 Review Round 4	10
2.6 Review Round 5	13
2.7 Lembar Persetujuan (LoA)	14
BAB III ISI KARYA ILMIAH	15
3.1 Intisari	15
3.2 Pendahuluan	16
3.3 Metode	18
3.4 Classification Models	19
3.5 Model Evaluation	23
3.6 Hasil dan Pembahasan	24
BAB IV	31
KESIMPULAN	31
Referensi	32
Lampiran	35

INTISARI

Strain *Escherichia coli* (*E. coli*) tertentu dapat menyebabkan penyakit serius, sehingga mengidentifikasi strain berbahaya dengan akurasi tinggi merupakan prioritas dalam mendukung kesehatan masyarakat dan keamanan pangan. Namun, metode pembelajaran mesin tradisional, seperti Decision Trees, seringkali tidak cukup tangguh untuk menangani kompleksitas data biologis. Penelitian ini menyajikan solusi dengan mengevaluasi secara sistematis tujuh metode ensemble, yaitu Adaboost, Gradient Boosting, XGBoost, LightGBM, Random Forest, Bagging, dan Stacking, menggunakan dataset yang mencakup 336 sampel *E. coli* dengan delapan fitur biologis. Model-model ini dievaluasi berdasarkan akurasi, presisi, recall, dan skor F1, dengan optimasi parameter untuk mendapatkan hasil terbaik. Hasil penelitian menunjukkan bahwa XGBoost lebih unggul dengan akurasi, recall, dan skor F1 sebesar 88% dan presisi sebesar 87%, mengungguli metode lain. Penelitian ini memiliki keunggulan berupa pendekatan komprehensif dalam membandingkan berbagai metode ensemble secara bersamaan, disertai dengan penerapan evaluasi berbasis matriks kebingungan untuk memastikan keakuratan hasil. Selain itu, pendekatan ensemble terbukti lebih efektif dalam menangani pola data yang kompleks dan mengurangi bias dalam klasifikasi strain bakteri. Temuan ini memberikan kontribusi yang signifikan, yaitu kerangka kerja praktis untuk meningkatkan diagnostik laboratorium dan pengawasan kesehatan masyarakat, dengan solusi berbasis pembelajaran mesin yang lebih cepat, lebih andal, dan dapat diterapkan untuk lingkungan industri dan klinis. Penelitian ini memperluas pemahaman tentang potensi metode ensemble dalam klasifikasi data mikrobiologi dan memberikan arah baru untuk teknologi diagnostik modern.

Kata Kunci: kinerja klasifikasi, decision tree, metode ensemble, klasifikasi *escherichia coli*, machine learning.

ABSTRACT

Certain strains of Escherichia coli (E. coli) can cause serious illness, so identifying dangerous strains with high accuracy is a priority in supporting public health and food safety. However, traditional machine learning methods, such as Decision Trees, are often not robust enough to handle the complexity of biological data. This research presents a solution by systematically evaluating seven ensemble methods, namely Adaboost, Gradient Boosting, XGBoost, LightGBM, Random Forest, Bagging, and Stacking, using a dataset that includes 336 E. coli samples with eight biological features. These models are evaluated based on accuracy, precision, recall, and F1 score, with parameter optimization to obtain the best results. The results show that XGBoost is superior with accuracy, recall, and F1 score of 88% and precision of 87%, outperforming other methods. This research has the advantage of a comprehensive approach in comparing various ensemble methods simultaneously, accompanied by the application of confusion matrix-based evaluation to ensure the accuracy of the results. Additionally, the ensemble approach proved to be more effective in handling complex data patterns and reducing bias in bacterial strain classification. These findings provide a significant contribution, namely a practical framework for improving laboratory diagnostics and public health surveillance, with machine learning-based solutions that are faster, more reliable, and applicable for both industrial and clinical environments. This research expands understanding of the potential of ensemble methods in microbiological data classification and provides new directions for modern diagnostic technology.

Keywords: *classification performance, decision tree, ensemble methods, escherichia coli classification, machine learning*