

**MENGINTEGRASIKAN EXPLAINABLE AI DALAM
DETEKSI BERITA HOAKS MELALUI METODE ENSEMBLE
LEARNING: RANDOM FOREST DAN ADABOOST**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi SI-Informatika



disusun oleh

WAYAN SURYA ADNYANA

21.11.4160

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

**MENGINTEGRASIKAN EXPLAINABLE AI DALAM
DETEKSI BERITA HOAKS MELALUI METODE ENSEMBLE
LEARNING: RANDOM FOREST DAN ADABOOST**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1-Informatika



disusun oleh

WAYAN SURYA ADNYANA

21.11.4160

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**MENGINTEGRASIKAN EXPLAINABLE AI DALAM DETEKSI BERITA
HOAKS MELALUI METODE ENSEMBLE LEARNING: RANDOM
FOREST DAN ADABOOST**

yang disusun dan diajukan oleh

WAYAN SURYA ADNYANA

21.11.4160

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 26 Mei 2025

Dosen Pembimbing,



Arifivanto Hadinegoro, S.Kom, M.T

NIK. 190302289

HALAMAN PENGESAHAN
SKRIPSI
MENGINTEGRASIKAN EXPLAINABLE AI DALAM DETEKSI BERITA
HOAKS MELALUI METODE ENSEMBLE LEARNING: RANDOM
FOREST DAN ADABOOST

yang disusun dan diajukan oleh

WAYAN SURYA ADNYANA

21.11.4160

Telah dipertahankan di depan Dewan Penguji
pada tanggal 26 Mei 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Ainul Yaqin, S.Kom., M.Kom.
NIK. 190302255

Uyock Anggoro Saputro, S.Kom., M.Kom.
NIK. 190302419

Arifiyanto Hadinegoro, S.Kom., M.T.
NIK. 190302289

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 26 Mei 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriati, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Wayan Surya Adnyana
NIM : 21.11.4160

Menyatakan bahwa Skripsi dengan judul berikut:

MENGINTEGRASIKAN EXPLAINABLE AI DALAM DETEKSI BERITA HOAKS MELALUI METODE ENSEMBLE LEARNING: RANDOM FOREST DAN ADABOOST

Dosen Pembimbing : Arifiyanto Hadinegoro, S.Kom, M.T

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
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 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, 26 Mei 2025

Yang Menyatakan,



Wayan Surya Adnyana

HALAMAN PERSEMBAHAN

Terselesaikannya skripsi ini merupakan hasil dari dukungan dan bantuan berbagai pihak yang terus memberikan semangat dan motivasi kepada penulis selama proses penyusunannya. Untuk itu, dengan rasa hormat penulis menyampaikan terima kasih kepada :

1. Orang tua, yang selalu memberikan dukungan moral dan materi yang tak pernah putus, menjadi sumber kekuatan utama bagi penulis dalam menyelesaikan skripsi ini.
2. Bapak Arifiyanto Hadinegoro, S.Kom, M.T, selaku dosen pembimbing yang telah memberikan bimbingan, masukan, dan motivasi dengan penuh kesabaran serta dedikasi hingga skripsi ini dapat diselesaikan dengan baik.
3. Teman-teman seperjuangan, yang senantiasa memberikan semangat, kebersamaan, dan dorongan selama masa studi hingga penyelesaian karya ini, khususnya kepada: Mahendra Bayu Prayoga, Bagas Restya Ernawan, Aldino Marsel Pratama, Reihansyah Maulana, dan Hafid Afnan.

KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Skripsi ini diajukan sebagai salah satu syarat untuk menyelesaikan program studi pada tingkat pendidikan tinggi.

Dalam proses penyusunan skripsi ini, penulis mendapatkan banyak dukungan dan bimbingan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis mengucapkan terima kasih yang sebesar-besarnya kepada Bapak Arifiyanto Hadinegoro, S.Kom, M.T, selaku Dosen Pembimbing yang telah memberikan arahan, masukan, dan motivasi selama penyusunan skripsi ini. Ucapan terima kasih juga penulis sampaikan kepada Tim Dosen Penguji atas saran dan kritik yang membangun untuk memperbaiki karya ini.

Penulis juga mengucapkan terima kasih kepada orang tua yang senantiasa memberikan doa, dukungan moral, dan materi selama proses studi hingga penyelesaian skripsi ini. Tidak lupa, ucapan terima kasih disampaikan kepada seluruh teman-teman dan orang-orang terdekat yang telah membantu, baik secara langsung maupun tidak langsung, dalam penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, saran dan kritik yang membangun sangat diharapkan untuk perbaikan di masa yang akan datang. Semoga skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan pembaca sekalian.

Yogyakarta, 23 Maret 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	I
HALAMAN PERSETUJUAN	II
HALAMAN PENGESAHAN	III
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	IV
HALAMAN PERSEMBAHAN	V
KATA PENGANTAR	VI
DAFTAR ISI	VII
DAFTAR TABEL	X
DAFTAR GAMBAR	XI
DAFTAR LAMPIRAN	XIII
DAFTAR LAMBANG DAN SINGKATAN	XIV
DAFTAR ISTILAH	XV
INTISARI	XVI
ABSTRACT	XVII
BAB I PENDAHULUAN	1
1.1 LATAR BELAKANG MASALAH	1
1.2 RUMUSAN MASALAH	2
1.3 BATASAN MASALAH	2
1.4 TUJUAN PENELITIAN	3
1.5 MANFAAT PENELITIAN	3
1.6 SISTEMATIKA PENULISAN	4
BAB II TINJAUAN PUSTAKA	5
2.1 STUDI LITERATUR	5
2.2 DASAR TEORI	13

2.2.1 Berita Hoaks	13
2.2.2 Term Frequency-Inverse Document Frequency (TF-IDF)	13
2.2.3 Machine Learning	14
2.2.4 Random Forest	15
2.2.5 AdaBoost	16
2.2.6 Ensemble Voting	18
2.2.7 Explainable AI (XAI)	20
2.2.8 Confusion Matrix	21
BAB III METODE PENELITIAN	23
3.1 ALUR PENELITIAN	23
3.1.1 Pengumpulan Data	23
3.1.2 Pengolahan Data	24
3.1.3 Pengembangan & Implementasi Metodologi	24
3.1.3.1 Data Gathering	25
3.1.3.2 Preprocessing	26
3.1.3.3 Modeling	31
3.1.4 Analisis Hasil	34
3.1.5 Dokumentasi	34
3.2 ALAT DAN BAHAN	34
BAB IV HASIL DAN PEMBAHASAN	36
4.1 PENGUMPULAN DATA	36
4.2 PENGOLAHAN DATA	37
4.3 PENGEMBANGAN & IMPLEMENTASI METODOLOGI	37
4.3.1 Data Gathering	37
4.3.2 Preprocessing	40
4.3.2.1 Case-Folding	41
4.3.2.2 Text Cleaning	42
4.3.2.3 Stopwords Removal	43
4.3.2.4 Tokenizing	45
4.3.2.5 Stemming	46

4.3.3	<i>Modeling</i>	49
4.3.3.1	<i>Feature Extraction With TF-IDF</i>	50
4.3.3.2	<i>Data Splitting</i>	52
4.3.3.3	Hasil Pemodelan <i>Random Forest</i>	52
4.3.3.4	Hasil Pemodelan <i>AdaBoost</i>	55
4.3.3.5	Hasil Pemodelan <i>Ensemble Voting</i>	57
4.3.3.6	Hasil <i>Explainable AI</i> dengan <i>LIME</i>	59
4.4	ANALISIS HASIL	62
4.5	DOKUMENTASI	66
BAB V PENUTUP		67
5.1	KESIMPULAN	67
5.2	SARAN	67
REFERENSI		69
LAMPIRAN		72

DAFTAR TABEL

TABEL 2.1 KEASLIAN PENELITIAN	8
TABEL 3.1 CONTOH PROSES <i>CASE-FOLDING</i>	27
TABEL 3.2 CONTOH PROSES <i>TEXT CLEANING</i>	28
TABEL 3.3 CONTOH PROSES <i>STOPWORD REMOVAL</i>	29
TABEL 3.4 CONTOH PROSES <i>TOKENIZING</i>	30
TABEL 3.5 CONTOH PROSES <i>STEMMING</i>	31
TABEL 4.1 HASIL EKSPLORASI AKURASI PADA STUDI LITERATUR	37
TABEL 4.2 PERBANDINGAN AKURASI MODEL KLASIFIKASI	64
TABEL 4.3 PERBANDINGAN AKURASI HASIL PENELITIAN SAAT INI DENGAN STUDI SEBELUMNYA	65

DAFTAR GAMBAR

GAMBAR 2.1 TINGKAT ERROR RANDOM FOREST BERDASARKAN JUMLAH POHON.....	16
GAMBAR 2.2 PERSENTASE TINGKAT <i>ERROR ADABOOST</i> BERDASARKAN ITERASI.....	18
GAMBAR 2.3 ILUSTRASI <i>ENSEMBLE VOTING</i>	19
GAMBAR 2.4 CONFUSION MATRIX 2X2.....	21
GAMBAR 3.1 ALUR PENELITIAN.....	23
GAMBAR 3.2 ALUR PERANCANGAN MODEL <i>MACHINE LEARNING</i>	25
GAMBAR 4.1 DATASET BERITA HOAX DAN VALID.....	38
GAMBAR 4.2 HASIL SCRAPPING DARI WEBSITE TURNBACKHOAX.ID.....	38
GAMBAR 4.3 DATA DARI KAGGLE.....	39
GAMBAR 4.4 SEBELUM PEMBERSIHAN DATA DUPLIKAT.....	40
GAMBAR 4.5 SETELAH PEMBERSIHAN DATA DUPLIKAT.....	40
GAMBAR 4.6 HASIL <i>CASE-FOLDING</i>	41
GAMBAR 4.7 HASIL <i>TEXT-CLEANING</i>	43
GAMBAR 4.8 HASIL STOPWORDS <i>REMOVAL</i>	45
GAMBAR 4.9 HASIL PROSES <i>TOKENIZING</i>	46
GAMBAR 4.10 HASIL <i>STEMMING</i>	48
GAMBAR 4.11 DISTRIBUSI LABEL SETELAH PREPROCESSING.....	49
GAMBAR 4.12 DATASET BERITA HOAKS.....	51
GAMBAR 4.13 PERBANDINGAN NILAI <i>TF-IDF</i>	51
GAMBAR 4.14 HASIL EVALUASI MODEL <i>RANDOM FOREST</i>	53
GAMBAR 4.15 CONFUSION MATRIX MODEL <i>RANDOM FOREST</i>	54

GAMBAR 4.16 HASIL CROSS-VALIDATION MODEL RANDOM FOREST	54
GAMBAR 4.17 HASIL EVALUASI MODEL <i>ADABOOST</i>	55
GAMBAR 4.18 CONFUSION MATRIX MODEL <i>ADABOOST</i>	56
GAMBAR 4.19 HASIL <i>CROSS-VALIDATION</i> MODEL <i>ADABOOST</i>	57
GAMBAR 4.20 HASIL EVALUASI MODEL <i>ENSEMBLE VOTING</i>	58
GAMBAR 4.21 CONFUSION MATRIX MODEL <i>ENSEMBLE VOTING</i>	59
GAMBAR 4.22 HASIL CROSS-VALIDATION MODEL <i>ENSEMBLE VOTING</i>	59
GAMBAR 4.22 VISUALISASI <i>XAI</i> PADA DATA UJI MODEL <i>ENSEMBLE VOTING</i>	60
GAMBAR 4.23 HASIL PENGUJIAN MODEL <i>ADABOOST</i> MENGGUNAKAN <i>XAI</i>	61
GAMBAR 4.24 HASIL PENGUJIAN MODEL <i>ENSEMBLE VOTING</i> MENGGUNAKAN <i>XAI</i>	62

DAFTAR LAMPIRAN

LAMPIRAN 1 KODE <i>SCRAPPING DATA</i>	72
LAMPIRAN 2 KODE PENGGABUNGAN DATASET	75
LAMPIRAN 3 KODE DETEKSI BERITA HOAKS	77
LAMPIRAN 4 PENGGUNAAN MODEL DETEKSI HOAX	88



DAFTAR LAMBANG DAN SINGKATAN

ξ	Slack Variable
Ω	konstanta penalti
LIME	Local Interpretable Model-agnostic Explanations
SHAP	SHapley Additive exPlanations
XAI	Explainable Artificial Intelligence
RF	Random Forest
AdaBoost	Adaptive Boosting
TF-IDF	Term Frequency-Inverse Document Frequency
CSV	Comma-Separated Values
CV	Cross Validation

DAFTAR ISTILAH



Soft Voting	Metode ensemble yang memilih kelas berdasarkan rata-rata probabilitas klasifikasi dari semua model
Hard Voting	Metode ensemble yang memilih kelas berdasarkan suara terbanyak dari masing-masing model
LIME	Teknik untuk menjelaskan prediksi model machine learning
Ensemble Learning	Teknik penggabungan beberapa model untuk meningkatkan akurasi
Hoaks	Informasi yang tidak benar atau menyesatkan
Sastrawi	Library stemming Bahasa Indonesia
Machine learning	Mesin belajar dari data
Lowercase	Mengubah teks ke huruf kecil
Preprocessing	Pemrosesan awal data
Confusion Matrix	Tabel evaluasi model
Library	Kumpulan kode siap pakai

INTISARI

Berita hoaks semakin marak dengan berkembangnya teknologi informasi dan media sosial, yang berdampak pada opini publik dan stabilitas sosial. Deteksi berita hoaks berbasis teks masih menghadapi tantangan dalam akurasi dan transparansi model. Oleh karena itu, penelitian ini mengintegrasikan *Explainable AI (XAI)* dalam *ensemble learning* menggunakan *Random Forest*, *AdaBoost*, dan *Ensemble Voting* untuk meningkatkan performa deteksi berita hoaks. Metode yang digunakan meliputi pengumpulan *dataset* dari sumber terpercaya, *preprocessing* teks seperti *case folding*, *stopword removal*, *tokenization*, dan *stemming*, serta ekstraksi fitur menggunakan *TF-IDF*. Model *Random Forest*, *AdaBoost*, dan *Ensemble Voting* diterapkan untuk klasifikasi, dengan evaluasi menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Untuk meningkatkan transparansi model, digunakan *LIME (Local Interpretable Model-agnostic Explanations)*. Hasil penelitian menunjukkan bahwa *AdaBoost* memiliki akurasi tertinggi sebesar 99.54%, tetapi *Ensemble Voting* lebih stabil dalam mengklasifikasikan berita hoaks dan valid. Integrasi *XAI* dengan *LIME* meningkatkan transparansi model, memungkinkan pengguna memahami faktor yang memengaruhi keputusan klasifikasi. Hasil penelitian ini dapat dimanfaatkan oleh platform media sosial, lembaga pemeriksa fakta, serta pengembang sistem deteksi berita hoaks. Penelitian selanjutnya dapat mengeksplorasi deteksi berita hoaks berbasis gambar dan video menggunakan *computer vision* dan *deep learning* untuk menangani penyebaran informasi palsu secara lebih komprehensif.

Kata kunci: Deteksi Berita Hoaks, Klasifikasi Berita, *Explainable AI (XAI)*, *Ensemble Learning*, *Machine Learning*

ABSTRACT

Fake news has become increasingly prevalent with the advancement of information technology and social media, impacting public opinion and social stability. Text-based fake news detection still faces challenges in terms of accuracy and model transparency. Therefore, this study integrates Explainable AI (XAI) into ensemble learning using Random Forest, AdaBoost, and Ensemble Voting to enhance the performance of fake news detection. The methodology includes data collection from reliable sources, text preprocessing such as case folding, stopword removal, tokenization, and stemming, and feature extraction using TF-IDF. The Random Forest, AdaBoost, and Ensemble Voting models were applied for classification, with performance evaluation based on accuracy, precision, recall, and F1-score. To improve model transparency, LIME (Local Interpretable Model-Agnostic Explanations) was employed. The results indicate that AdaBoost achieved the highest accuracy of 99.54%, while Ensemble Voting demonstrated greater stability in classifying both fake and valid news. The integration of XAI with LIME enhances model transparency, allowing users to understand the factors influencing classification decisions. The findings of this study can be utilized by social media platforms, fact-checking institutions, and developers of fake news detection systems. Future research may explore fake news detection based on images and videos using computer vision and deep learning to address the spread of misinformation more comprehensively.

Keyword: *Fake News Detection, News Classification, Explainable AI (XAI), Ensemble Learning, Machine Learning*