

**ALGORITMA *MACHINE LEARNING* YANG DITINGKATKAN
DENGAN *GRIDSEARCHCV* UNTUK KLASIFIKASI BERITA
HOAKS PADA PLATFORM BERITA INDONESIA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1-Teknik Komputer



disusun oleh

Naufal Maheswara

21.83.0663

Kepada

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

**ALGORITMA *MACHINE LEARNING* YANG DITINGKATKAN
DENGAN *GRIDSEARCHCV* UNTUK KLASIFIKASI BERITA
HOAKS PADA PLATFORM BERITA INDONESIA**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1-Teknik Komputer



disusun oleh

Naufal Maheswara

21.83.0663

Kepada

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

HALAMAN PERSETUJUAN

SKRIPSI

**ALGORITMA *MACHINE LEARNING* YANG DITINGKATKAN DENGAN
GRIDSEARCHCV UNTUK KLASIFIKASI BERITA HOAKS PADA
PLATFORM BERITA INDONESIA**

yang disusun dan diajukan oleh

Naufal Maheswara

21.83.0663

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 November 2025

Dosen Pembimbing,



Melwin Syafrizal, S.Kom., M.Eng., PhD.

NIK. 190302105

HALAMAN PENGESAHAN
SKRIPSI
ALGORITMA *MACHINE LEARNING* YANG DITINGKATKAN DENGAN
***GRIDSEARCHCV* UNTUK KLASIFIKASI BERITA HOAKS PADA**
PLATFORM BERITA INDONESIA

yang disusun dan diajukan oleh
Naufal Maheswara
21.83.0663

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 November 2025

Susunan Dewan Penguji

Nama Penguji	Tanda Tangan
<u>Dr. Dony Arivus, S.S., M.Kom.</u> NIK. 190302128	
<u>Jeki Kuswanto, S.Kom., M.Kom.</u> NIK. 190302456	
<u>Melwin Syafrizal, S.Kom., M.Eng., PhD.</u> NIK. 190302105	

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 November 2025

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Naufal Maheswara
NIM : 21.83.0663

Menyatakan bahwa Skripsi dengan judul berikut:

ALGORITMA MACHINE LEARNING YANG DITINGKATKAN DENGAN GRIDSEARCHCV UNTUK KLASIFIKASI BERITA HOAKS PADA PLATFORM BERITA INDONESIA

Dosen Pembimbing : Melwin Syafrizal, S.Kom., M.Eng., PhD.

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, 19 November 2025

Yang Menyatakan,



Naufal Maheswar

HALAMAN PERSEMBAHAN

Segala puji dan rasa syukur saya haturkan kepada Tuhan Yang Maha Esa atas limpahan rahmat, kasih, dan bimbingan-Nya, sehingga saya dapat menyelesaikan skripsi ini dengan baik. Dengan penuh rasa hormat dan kerendahan hati, karya sederhana ini saya persembahkan kepada:

1. Orang tua tercinta, atas doa, kasih, dan dukungan yang tiada henti dalam setiap langkah hidup saya.
2. Prof. Dr. M. Suyanto., M.M selaku Rektor Universitas Amikom Yogyakarta.
3. Bapak Melwin Syafrizal, S.Kom., M.Eng., Ph.D. selaku Dosen Pembimbing dari Teknik Komputer yang telah memberikan bimbingan dan bantuan kepada penulis.
4. Keluarga besar, yang senantiasa memberi semangat, perhatian, dan dorongan moral dalam perjalanan ini, serta tempat berbagi kebahagiaan atas pencapaian yang
5. Sahabat dan teman seperjuangan, atas kebersamaan, motivasi, dan dukungan yang membuat perjalanan ini lebih bermakna.

KATA PENGANTAR

Puji syukur saya panjatkan ke hadirat Allah SWT atas limpahan rahmat, taufik, serta karunia-Nya, sehingga saya dapat menuntaskan penyusunan skripsi yang berjudul “Algoritma Machine Learning yang Ditingkatkan dengan GridSearch untuk Klasifikasi Berita Hoaks pada Platform Berita Indonesia”. Karya ilmiah ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Teknik Komputer, Universitas Amikom Yogyakarta.

Dalam proses penyusunan skripsi ini, saya mendapatkan banyak dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh sebab itu, dengan penuh rasa hormat dan kerendahan hati, saya menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Orang tua tercinta, atas doa, kasih, dan dukungan yang tiada henti dalam setiap langkah hidup saya.
2. Prof. Dr. M. Suyanto., M.M selaku Rektor Universitas Amikom Yogyakarta.
3. Bapak Melwin Syafrizal, S.Kom., M.Eng., Ph.D. selaku Dosen Pembimbing dari Teknik Komputer yang telah memberikan bimbingan dan bantuan kepada penulis.

Yogyakarta, 10 November 2025

Naufal Maheswara

DAFTAR ISI

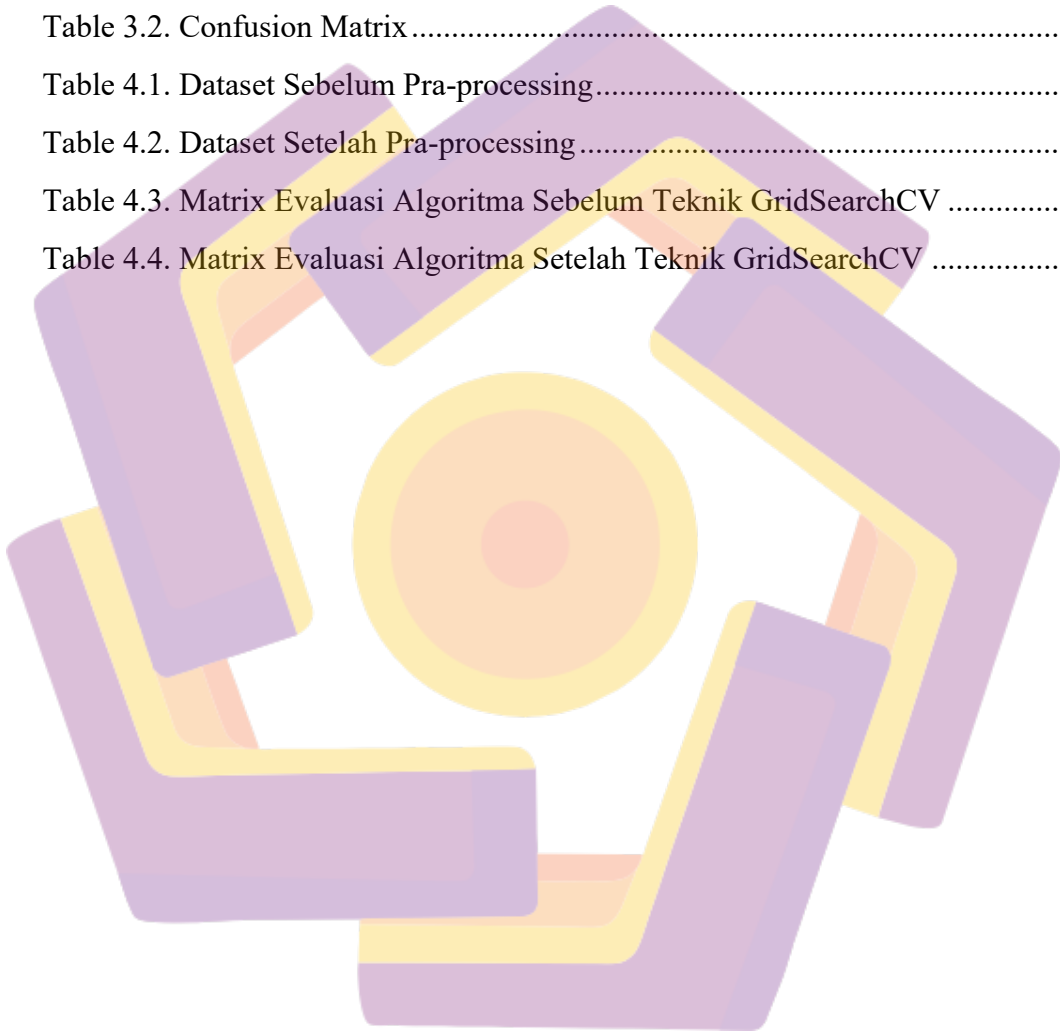
HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
INTISARI	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 Studi Literatur	7
2.2 Dasar Teori.....	13
2.2.1 Supervised Learning	13
2.2.2 <i>Machine learning</i>	13
2.2.3 Berita Hoax	14
2.2.4 <i>K-Nearest Neighbor</i>	14

2.2.5	<i>Naïve Bayes</i>	15
2.2.6	<i>Random Forest</i>	15
2.2.7	<i>Support Vector Machine</i>	17
2.2.8	<i>XGBoost</i>	17
2.2.9	Tokenizing	18
2.2.10	Stopword	18
2.2.11	<i>TF-IDF</i>	19
2.2.12	<i>Hyperparameter Tuning</i>	19
2.2.13	Evaluation Model	20
BAB III METODE PENELITIAN		22
3.1	Objek Penelitian	22
3.2	Alur Penelitian	23
3.2.1	Data Collection	24
3.2.2	Praprocessing Data	25
3.2.3	Algoritma <i>Machine learning</i>	27
3.2.4	<i>Hyperparameter Tuning</i>	30
3.2.5	Evaluation Model	31
3.2.6	Implementation of Web Apps	31
3.3	Alat dan Bahan	32
3.3.1	Bahan	32
3.3.2	Alat/Instrumen	32
BAB IV HASIL DAN PEMBAHASAN		34
4.1	Data Collection	34
4.2	Data Cleaning	35
4.3	Tokenization	38

4.4	Stopword	38
4.5	Label Encoder	39
4.6	<i>TF-IDF</i>	40
4.7	Final Dataset	43
4.8	Klasifikasi Algoritma <i>Machine learning</i> Sebelum Grid	45
4.9	Klasifikasi Algoritma <i>Machine learning</i> Setelah Grid	50
4.10	Evaluasi Perbandingan Algoritma <i>Machine learning</i>	56
4.11	Implementasi Web Apps	71
BAB V PENUTUP		78
5.1	Kesimpulan	78
5.2	Saran	78
REFERENSI		79

DAFTAR TABEL

Table 2.1. Keaslian Penelitian	10
Table 3.1. Dataset Hoax News.....	24
Table 3.2. Confusion Matrix	31
Table 4.1. Dataset Sebelum Pra-processing.....	43
Table 4.2. Dataset Setelah Pra-processing	44
Table 4.3. Matrix Evaluasi Algoritma Sebelum Teknik GridSearchCV	62
Table 4.4. Matrix Evaluasi Algoritma Setelah Teknik GridSearchCV	62



DAFTAR GAMBAR

Gambar 3.1. Alur Penelitian	23
Gambar 4.1. Dataset Data Head.....	34
Gambar 4.2. Import Dataset.....	35
Gambar 4.3. Kode Data Cleaning	35
Gambar 4.4. Hasil Setelah Data Cleaning	36
Gambar 4.5. Kode Menggabungkan Colom Judul dan Konten	36
Gambar 4.6. Hasil dari Kode Menggabungkan Colom Judul dan Konten	37
Gambar 4.7. Kata Stopword.....	38
Gambar 4.8. Kode LabelEncoder.....	39
Gambar 4.9. Hasil LabelEncoder.....	40
Gambar 4.10. Hasil dari <i>TF-IDF</i>	40
Gambar 4.11. Kode Untuk Menampilkan Hasil dari <i>TF-IDF</i>	41
Gambar 4.12. Kata Sering Muncul dari Hasil <i>TF-IDF</i>	42
Gambar 4.13. Kode Algoritma KNN Sebelum Teknik <i>GridSearchCV</i>	45
Gambar 4.14. Kode Algoritma <i>Naïve Bayes</i> Sebelum Teknik <i>GridSearchCV</i>	46
Gambar 4.15. Kode Algoritma <i>Random Forest</i> Sebelum Teknik <i>GridSearchCV</i>	47
Gambar 4.16. Kode Algoritma SVM Sebelum Teknik <i>GridSearchCV</i>	48
Gambar 4.17. Kode Algoritma <i>XGBoost</i> Sebelum Teknik <i>GridSearchCV</i>	49
Gambar 4.18. Kode Algoritma KNN Setelah Teknik <i>GridSearchCV</i>	50
Gambar 4.19. Kode Algoritma <i>Naïve Bayes</i> Setelah Teknik <i>GridSearchCV</i>	51
Gambar 4.20. Kode Algoritma <i>Random Forest</i> Setelah Teknik <i>GridSearchCV</i>	53
Gambar 4.21. Kode Algoritma SVM Setelah Teknik <i>GridSearchCV</i>	54
Gambar 4.22. Kode Algoritma <i>XGBoost</i> Setelah Teknik <i>GridSearchCV</i>	55
Gambar 4.23. Performa dari Presisi, Recal, dan F1-Score Sebelum Teknik <i>GridSearchCV</i>	56
Gambar 4.24. Performa dari Presisi, Recal, dan F1-Score Setelah Teknik <i>GridSearchCV</i>	58
Gambar 4.25. Perfoma Akurasi pada Algoritma Sebelum Teknik <i>GridSearchCV</i>	57

Gambar 4.26. Perfoma Akurasi pada Algoritma Setelah Teknik <i>GridSearchCV</i> ..	60
Gambar 4.27. Perfoma Akurasi pada Algoritma Sebelum dan Setelah Teknik <i>GridSearchCV</i>	61
Gambar 4.28. Confusion Matrix pada Algoritma Sebelum Teknik <i>GridSearchCV</i>	64
Gambar 4.29. ROC pada Algoritma Sebelum Teknik <i>GridSearchCV</i>	66
Gambar 4.30. Confusion Matrix pada Algoritma Setelah Teknik <i>GridSearchCV</i>	68
Gambar 4.31. ROC pada Algoritma Setelah Teknik <i>GridSearchCV</i>	70
Gambar 4.32. Kode Import Library dan Model <i>Machine learning</i>	71
Gambar 4.33. Kode Menyimpan Hasil Prediksi Agar Jadi Data Latih Lanjutan...	72
Gambar 4.34. Kode Route Halaman Hasil Prediksi.....	73
Gambar 4.35. Tampilan Index Deteksi Berita	74
Gambar 4.36. Tampilan Setelah Mendeteksi Berita	75
Gambar 4.37. Tampilan Riwayat Hasil Prediksi.....	75

INTISARI

Di media sosial dan platform berita, mendeteksi berita palsu menjadi krusial. Lima algoritma pembelajaran mesin, yaitu *K-Nearest Neighbors*, *Naïve Bayes*, *Random Forest*, *Support Vector Machine*, dan *XGBoost*, digunakan dalam sistem deteksi berita palsu yang diusulkan dalam studi ini untuk mengklasifikasikan berita. Prapemrosesan teks, prosedur klasifikasi, dan penyetelan hiperparameter *GridSearchCV* merupakan tiga bagian utama sistem ini. Persentase hoaks dan berita yang sah cukup tinggi dapat ditemukan dalam 3000 item berita dalam dataset, yang dibagi menjadi 80% pelatihan dan 20% pengujian. Pembersihan data, tokenisasi, eliminasi stopwords, dan komputasi *TF-IDF* untuk ekstraksi fitur merupakan langkah-langkah dalam prosedur prapemrosesan teks. Berdasarkan hasil evaluasi, model SVM memiliki akurasi terbaik (97%), diikuti oleh *Random Forest* dan *XGBoost*, yang keduanya memiliki akurasi 96%. Model-model ini berkinerja sangat baik dalam hal presisi, recall, dan skor F1, yang secara efektif membedakan antara berita yang sah dan palsu. Semua algoritma menunjukkan peningkatan kinerja yang signifikan selama penyetelan *GridSearchCV*, dengan SVM mempertahankan akurasinya yang lebih baik. Dengan temuan-temuan yang relevan dan aplikasi yang bermanfaat dalam sistem deteksi otomatis, studi ini menunjukkan bahwa penggunaan algoritma pembelajaran mesin dalam identifikasi berita hoaks dapat meningkatkan efektivitas verifikasi berita di Indonesia.

Kata kunci: Berita hoax, Deteksi hoax, Pembelajaran mesin, Algoritma klasifikasi, Klasifikasi teks.

ABSTRACT

In social media and news platforms, spotting fake news has become crucial. Five machine learning algorithms K-Nearest Neighbors, Naïve Bayes, Random Forest, Support Vector Machine, and XGBoost are used in this study's proposed fake news detection system to classify news. Text preprocessing, the classification procedure, and GridSearchCV Hyperparameter Tuning make up this system's three primary parts. A fair percentage of hoaxes and legitimate news can be found in the 3000 news items in the dataset, which is split into 80% training and 20% testing. Data cleansing, tokenization, stopword elimination, and TF-IDF computation for feature extraction are all steps in the text preprocessing procedure. According to the evaluation results, the SVM model had the best accuracy (97%), followed by Random Forest and XGBoost, both of which had 96% accuracy. These models perform exceptionally well in terms of precision, recall, and F1-score, effectively differentiating between legitimate and fake news. All algorithms demonstrated notable performance gains during GridSearchCV tuning, with SVM maintaining its better accuracy. With pertinent findings and useful applications in automated detection systems, this study shows that the employment of machine learning algorithms in hoax news identification can increase the efficacy of news verification in Indonesia.

Keyword: Hoax news, Hoax detection, Machine learning, Classification algorithms, Text classification.