TESIS

PENGARUH PENERAPAN KOMBINASI METODE *PREPROCESSING*, *VECTORIZER*, DAN ALGORITMA TERHADAP KINERJA KLASIFIKASI TEKS PERUNDUNGAN SIBER

Oleh:

Salim Athari

22206051023

PROGRAM STUDI MAGISTER INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA

YOGYAKARTA
2025

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama	: Salim Athari
NIM	: 22206051023
Jenjang	: Magister
Program Studi	: Informatika

menyatakan bahwa naskah tesis ini secara keseluruhan adalah hasil penelitian/karya saya sendiri, kecuali pada bagian-bagian yang dirujuk sumbernya.

Yogyakarta, 5 Agustus 2025

Saya yang menyatakan,

Salim Athari

NIM: 22206051023

PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan di bawah ini:

Nama	: Salim Athari
NIM	: 22206051023
Jenjang	: Magister
Program Studi	: Informatika

menyatakan bahwa naskah tesis ini secara keseluruhan benar-benar bebas dari plagiasi. Jika di kemudian hari terbukti melakukan plagiasi, maka saya siap ditindak sesuai ketentuan hukum yang berlaku.

Yogyakarta, 5 Agustus 2025

Saya yang menyatakan,

Salim Athari

NIM: 22206051023

PENGESAHAN

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1893/Un.02/DST/PP.00.9/08/2025

Tugas Akhir dengan judul : Pengaruh Penerapan Kombinasi Metode Preprocessing, Vectorizer, dan Algoritma terhadap Kinerja Klasifikasi Teks Perundungan Siber

yang dipersiapkan dan disusun oleh:

Nama	: SALIM ATHARI, S.HI., S.Kom.
Nomor Induk Mahasiswa	: 22206051023
Telah diujikan pada	: Sabtu, 16 Agustus 2025
Nilai ujian Tugas Akhir	: A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR

Ketua Sidang

Dr. Agung Fatwanto, S.Si., M.Kom.
SIGNED

Valid ID: 68a7b46013ebc

Penguji I

Prof. Dr. Ir. Shofwatul 'Uyun, S.T., M.Kom., IPM., ASEAN Eng.
SIGNED

Valid ID: 68a6d924d02e9

Penguji II

Dr. Agus Mulyanto, S.Si., M.Kom., ASEAN Eng.
SIGNED

Valid ID: 68a79a8077188

Yogyakarta, 16 Agustus 2025
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 68a7cf2057c6c

1/1 22/08/2025

PERSETUJUAN

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

BERITA ACARA UJIAN TUGAS AKHIR

Penyelenggaraan Ujian Tugas Akhir Mahasiswa

A. Waktu, Tempat dan Status Ujian Tugas Akhir:

1. Hari dan Tanggal : Sabtu, 16 Agustus 2025
2. Pukul : 08:30 s/d 09:30 WIB
3. Tempat : FST-1-101
4. Status : Utama

B. Susunan Tim Ujian Tugas Akhir:

No.	Jabatan	Nama	Tanda Tangan
1.	Ketua Sidang	Dr. Agung Fatwanto, S.Si., M.Kom.	Valid ID: 68a7b3eddf55c4
2.	Penguji I	Prof. Dr. Ir. Shofwatul 'Uyun, S.T., M.Kom., IPM., ASEAN Eng.	Valid ID: 68a5093f4161e
3.	Penguji II	Dr. Agus Mulyanto, S.Si., M.Kom., ASEAN Eng.	Valid ID: 68a79c99c799e

C. Identitas Mahasiswa yang diuji:

1. Nama : SALIM ATHARI, S.HI., S.Kom.
2. Nomor Induk Mahasiswa : 22206051023
3. Program Studi : S2 - Informatika
4. Semester : IV
5. Program : S2
6. Status Kehadiran Mahasiswa : Menghadiri Ujian

D. Judul Tugas Akhir : Pengaruh Penerapan Kombinasi Metode Preprocessing, Vectorizer, dan Algoritma terhadap Kinerja Klasifikasi Teks Perundungan Siber

E. Pembimbing/Promotor:

1. Dr. Agung Fatwanto, S.Si., M.Kom.

F. Keputusan Sidang

1. LULUS dengan Perbaikan
2. Predikat Kelulusan : 95.1 (A)
3. Konsultasi Perbaikan a. _____

b. _____

Yogyakarta, 16 Agustus 2025
Ketua Sidang/Pembimbing/Promotor,

Dr. Agung Fatwanto, S.Si., M.Kom.
SIGNED

Valid ID: 68a7b3eddf55c4

1/1 22/08/2025

NOTA DINAS PEMBIMBING

Kepada Yth.,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga
Yogyakarta

Assalamu'alaikum wr. wb.

Setelah melakukan bimbingan, arahan, dan koreksi terhadap penulisan tesis yang berjudul:

PENGARUH PENERAPAN KOMBINASI METODE *PREPROCESSING*, *VECTORIZER*, DAN ALGORITMA TERHADAP KINERJA KLASIFIKASI TEKS PERUNDUNGAN SIBER

Yang ditulis oleh:

Nama	: Salim Athari
NIM	: 22206051023
Jenjang	: Magister
Program Studi	: Informatika

Saya berpendapat bahwa tesis tersebut sudah dapat diajukan kepada Magister Informatika UIN Sunan Kalijaga untuk diujikan dalam rangka memperoleh gelar Magister Informatika.

Wassalamu'alaikum wr. wb.

Yogyakarta, 8 Agustus 2025
Pembimbing,

Dr. Agung Fatwanto, S.Si., M.Kom.
NIP: 197701032005011003

ABSTRAK

Perundungan siber merupakan ancaman digital serius yang memerlukan deteksi otomatis. *Machine Learning* (ML) dan *Natural Language Processing* (NLP) menawarkan solusi, namun kinerjanya sangat dipengaruhi oleh kombinasi metode *preprocessing*, *vectorization*, dan algoritma klasifikasi. Penelitian ini membandingkan kombinasi tersebut secara sistematis guna mengidentifikasi konfigurasi optimal karena belum adanya pendekatan yang holistik. *Preprocessing* yang digunakan yaitu *stopwords removal*, *stemming*, *lemmatization* dengan *vectorizer one-hot*, BoW, TF-IDF dan algoritma *Logistic Regression*, SVM, KNN, *Decision Tree*, *Naive Bayes* untuk klasifikasi teks perundungan siber. Tujuannya untuk menemukan kombinasi metode yang paling efektif berdasarkan metrik kinerja standar.

Pendekatan yang digunakan yaitu kuantitatif eksperimental komparatif, dengan *dataset* perundungan siber berbahasa Inggris yang telah dilabeli. *Preprocessing* data dilakukan sebelum proses *vectorization* dan kemudian dilatih menggunakan algoritma klasifikasi. Kinerja dievaluasi secara ketat menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*. Sehingga, dengan pendekatan holistik tersebut dapat memberikan hasil tertinggi dari berbagai kombinasi tersebut.

Hasil menunjukkan bahwa *preprocessing* secara signifikan meningkatkan kinerja, dengan *lemmatization* umumnya lebih unggul. *One-Hot Encoding* memberikan hasil terendah, sementara TF-IDF adalah *vectorizer* paling efektif. Di antara algoritma, SVM dan *Logistic Regression* tampil terbaik, dengan SVM menunjukkan kinerja sangat baik, terutama dengan TF-IDF. Kombinasi optimal diperoleh ketika menggunakan *stopwords removal*, TF-IDF, dan SVM yaitu mencapai *F1-score* sebesar 0.863. Sedangkan hasil kombinasi terendah sebesar 0.155 diperoleh ketika tanpa menggunakan *preprocessing*, *vectorization one-hot encoding*, dan algoritma KNN. Hal ini menekankan pentingnya *preprocessing* yang presisi, *vectorizer* yang informatif, dan algoritma yang *robust* untuk deteksi perundungan siber yang efektif.

Kata Kunci: Perundungan Siber, Klasifikasi Teks, *Preprocessing*, *Vectorizer*, Algoritma

ABSTRACT

Cyberbullying is a serious digital threat that requires automated detection. Machine Learning (ML) and Natural Language Processing (NLP) offer solutions, but their performance is significantly affected by the combination of preprocessing methods, vectorization, and classification algorithms. This study systematically compares these combinations to identify the optimal configuration due to the lack of a holistic approach. The preprocessing methods used include stopword removal, stemming, lemmatization with a one-hot vectorizer, BoW, TF-IDF, and Logistic Regression, SVM, KNN, Decision Tree, and Naive Bayes algorithms for cyberbullying text classification. The goal is to find the most effective combination of methods based on standard performance metrics.

The approach used is a comparative, quantitative, experimental approach, using a labeled English-language cyberbullying dataset. Data preprocessing was performed before vectorization and then trained using a classification algorithm. Performance was rigorously evaluated using accuracy, precision, recall, and F1-score. Therefore, this holistic approach yielded the highest results among the various combinations.

The results show that preprocessing significantly improves performance, with lemmatization generally outperforming. One-Hot Encoding yielded the lowest results, while TF-IDF was the most effective vectorizer. Among the algorithms, SVM and Logistic Regression performed best, with SVM demonstrating excellent performance, especially with TF-IDF. The optimal combination was obtained when using stopword removal, TF-IDF, and SVM, achieving an F1-score of 0.863. The lowest combination result, 0.155, was obtained when using no preprocessing, one-hot encoding vectorization, and the KNN algorithm. This emphasizes the importance of precise preprocessing, an informative vectorizer, and a robust algorithm for effective cyberbullying detection.

***Keywords**: Cyberbullying, Text Classification, Preprocessing, Vectorizer, Algorithm*

KATA PENGANTAR

Puji syukur kehadirat Allah *Subhanahu wata'ala* atas berkat dan rahmat-Nya, sehingga penulis dapat menyelesaikan penulisan tesis ini dan diajukan sebagai salah satu syarat untuk memperoleh gelar Magister pada Program Studi Magister Informatika. Penulisan tesis ini merupakan bentuk kontribusi akademis dalam upaya mengatasi isu perundungan siber yang semakin meresahkan di era digital. Penelitian ini berupaya menjawab tantangan tersebut dengan menganalisis secara sistematis dampak dari berbagai konfigurasi metode pemrosesan teks terhadap kinerja model klasifikasi.

Penulis menyadari bahwa keberhasilan dalam menyelesaikan tesis ini tidak terlepas dari dukungan, bimbingan, dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Noorhaidi, S.Ag., M.A., M.Phil., Ph.D. selaku Rektor Universitas Islam Negeri Sunan Kalijaga Yogyakarta dan seluruh jajaran pimpinan atas kesempatan yang diberikan untuk menempuh pendidikan di Universitas ini.
2. Ibu Prof. Dr. Dra. Hj. Khurul Wardati, M.Si. selaku Dekan Fakultas Sains dan Teknologi atas fasilitas yang diberikan.
3. Bapak Dr. Ir. Sumarsono, S.T., M.Kom. selaku Ketua Program Studi Magister Informatika atas dukungan dan fasilitas yang diberikan.
4. Bapak Dr. Agung Fatwanto, S.Si., M.Kom. selaku Dosen Pembimbing yang telah meluangkan waktu, memberikan arahan, bimbingan, dan motivasi yang tak henti-hentinya sejak awal hingga akhir penelitian.

5. Bapak/Ibu Dosen Penguji yang telah memberikan masukan dan saran konstruktif untuk perbaikan tesis ini.
6. Seluruh dosen Program Studi Magister Informatika yang telah membekali penulis dengan ilmu pengetahuan yang sangat berharga selama masa perkuliahan.
7. Orang tua, Istri, dan anak-anak yang telah memberikan doa, kasih sayang, serta dukungan moril dan materil yang tidak pernah putus, menjadi sumber kekuatan terbesar penulis.
8. Teman-teman Magister Informatika angkatan 2022 yang telah berjuang bersama dan saling berbagi pengalaman dan kebahagiaan.
9. Semua pihak yang tidak dapat disebutkan satu per satu yang telah memberikan kontribusi dalam penyelesaian tesis ini.

Penulis berharap tesis ini dapat memberikan manfaat, baik sebagai referensi akademis maupun sebagai inspirasi untuk penelitian-penelitian selanjutnya di bidang deteksi perundungan siber. Penulis menyadari bahwa tesis ini masih memiliki kekurangan. Oleh karena itu, segala bentuk kritik dan saran yang membangun akan diterima dengan tangan terbuka demi perbaikan di masa mendatang.

Yogyakarta, 4 Agustus 2025

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR SINGKATAN

BoW	: *Bag-of-Words*
CSV	: *Comma-Separated Values*
DT	: *Decision Tree*
FP	: *False Positive*
KNN	: *K-Nearest Neighbors*
LR	: *Logistic Regression*
ML	: *Machine Learning*
NB	: *Naive Bayes*
NLP	: *Natural Language Processing*
SVM	: *Support Vector Machines*
TF-IDF	: *Term Frequency-Inverse Document Frequency*
TN	: *True Negative*
TP	: *True Positive*

BAB I
PENDAHULUAN

A. Latar Belakang

Dalam era digital yang serba terhubung seperti saat ini, media sosial telah menjadi bagian integral dari kehidupan masyarakat modern. Platform seperti Twitter, Facebook, Instagram, dan TikTok memungkinkan pengguna untuk berbagi informasi, berinteraksi, dan mengekspresikan opini secara terbuka. Namun, kebebasan berekspresi ini juga membawa tantangan baru dalam bentuk penyebaran ujaran kebencian, pelecehan daring, dan perundungan siber (*cyberbullying*). Fenomena perundungan siber telah menjadi perhatian serius karena dampaknya yang signifikan terhadap masalah sosial dan kesehatan mental, terutama bagi kelompok usia muda. Lebih lanjut, pada penelitian Kowalski et al. (2014) menyatakan bahwa perundungan siber dapat menyebabkan depresi, kecemasan, bahkan keinginan untuk bunuh diri.

Tantangan besar dalam mengatasi perundungan siber adalah jumlah data yang sangat besar dan sifatnya yang dinamis. Konten yang merundung dapat tersembunyi dalam bentuk sarkasme, singkatan, atau bahasa yang kontekstual. Oleh karena itu, pendekatan manual untuk mendeteksi dan mengklasifikasikan konten perundungan menjadi tidak efisien dan tidak dapat diandalkan dalam skala besar (Muminovic, 2025). Dalam konteks ini, penerapan teknologi *Natural Language Processing* (NLP) dan *Machine Learning* (ML) menjadi solusi yang sangat potensial. Dengan menggunakan metode NLP, teks dapat diproses, dipahami,

dan dianalisis secara otomatis dan hasilnya dapat diklasifikasikan menggunakan algoritma tertentu untuk mendeteksi unsur perundungan (Alabdulwahab et al., 2023).

Namun, membangun sistem klasifikasi teks yang efektif tidaklah sederhana. Setiap tahap dalam *pipeline* pemrosesan data memiliki pengaruh terhadap kinerja akhir model klasifikasi. Tiga komponen utama yang sangat menentukan adalah metode *preprocessing*, teknik *vectorization*, dan algoritma klasifikasi. Masing-masing komponen ini memiliki variasi pendekatan, dan kombinasi antar metode tersebut dapat memberikan hasil yang berbeda terhadap performa klasifikasi.

Tahap *preprocessing* adalah tahap awal dan fundamental dalam pengolahan teks. Tujuannya adalah untuk menyederhanakan data teks sehingga lebih mudah diproses oleh mesin. Tiga teknik yang umum digunakan adalah *stopword removal*, *stemming*, dan *lemmatization*. *Stopword removal* menghilangkan kata-kata umum yang tidak memiliki makna signifikan seperti “dan”, “adalah”, atau “yang”. *Stemming* memotong kata hingga bentuk dasarnya tanpa memperhatikan konteks morfologis, sedangkan *lemmatization* mempertahankan bentuk dasar kata berdasarkan konteks gramatikal (Manning et al., 2008). Pemilihan teknik *preprocessing* yang sesuai dapat meningkatkan akurasi model karena membantu mengurangi dimensi dan kompleksitas data.

Setelah *preprocessing*, langkah selanjutnya adalah representasi teks dalam bentuk numerik melalui teknik *vectorization*. Tiga teknik populer adalah *one-hot encoding*, *Bag-of-Words* (BoW), dan *Term Frequency-Inverse Document Frequency* (TF-IDF). *One-hot*

encoding merepresentasikan setiap kata sebagai vektor biner yang sangat tinggi dimensinya. BoW menghitung frekuensi kemunculan kata dalam dokumen, sementara TF-IDF mempertimbangkan pentingnya kata berdasarkan distribusinya dalam seluruh korpus teks (Ramos, 2003). Setiap teknik memiliki kelebihan dan kekurangan. Misalnya, meskipun TF-IDF dapat meningkatkan akurasi, ia juga lebih kompleks secara komputasi dibandingkan BoW.

Tahap akhir dalam *pipeline* NLP adalah klasifikasi. Berbagai algoritma *supervised learning* dapat digunakan untuk memprediksi apakah sebuah teks mengandung unsur perundungan atau tidak. Beberapa algoritma populer antara lain *Logistic Regression*, *Support Vector Machines* (SVM), *K-Nearest Neighbors* (KNN), *Decision Tree*, dan *Naive Bayes*. Masing-masing algoritma ini bekerja dengan prinsip yang berbeda. *Logistic Regression* efektif untuk data linier dan interpretasi yang sederhana. SVM unggul dalam memisahkan data berdimensi tinggi dan sering digunakan dalam klasifikasi teks (Joachims, 1998). KNN bekerja berdasarkan kemiripan jarak antar data, namun sensitif terhadap dimensi tinggi. *Decision Tree* memberikan interpretasi yang baik namun rentan terhadap *overfitting*, sedangkan *Naive Bayes* efisien dalam komputasi dan cukup baik untuk data teks yang saling independen.

Permasalahan yang menarik muncul dari fakta bahwa belum banyak penelitian yang secara sistematis membandingkan kombinasi berbagai teknik *preprocessing*, *vectorization*, dan algoritma klasifikasi terhadap kinerja deteksi perundungan siber. Penelitian sebelumnya cenderung menggunakan pendekatan tunggal

atau hanya fokus pada satu atau dua komponen. Misalnya, Dinakar et al. (2012) menggunakan SVM untuk mendeteksi perundungan pada komentar YouTube, namun tidak mengeksplorasi alternatif *vectorizer* dengan membandingkan atau mengevaluasi berbagai teknik *vectorization* dan hanya melakukan *preprocessing* dasar. Hal ini menyisakan celah penelitian mengenai kombinasi metode mana yang paling efektif dalam mendeteksi perundungan pada teks dalam skenario dunia nyata.

Hubungan dan keterkaitan antara ilmu komputer, linguistik, dan psikologi sosial menjadi bahasan menarik untuk dibawa ke ruang diskusi ilmiah. Pemahaman yang lebih dalam terhadap kombinasi teknik NLP yang optimal akan memperkaya literatur ilmiah, khususnya dalam bidang *text analytics* dan pemrosesan bahasa alami. Selain itu, topik ini penting karena dapat memberikan kontribusi langsung terhadap isu sosial yang nyata yaitu perlindungan terhadap individu maupun kelompok dari kekerasan verbal di ruang digital.

Urgensi penelitian ini juga didorong oleh meningkatnya tekanan sosial dan regulasi terhadap platform digital untuk lebih bertanggung jawab dalam memoderasi konten (Kokshagina et al., 2023). Deteksi otomatis yang tidak akurat, baik dalam bentuk *false positive* maupun *false negative*, dapat menyebabkan ketidakadilan, baik terhadap korban maupun pelaku (Ricol, 2022). Oleh karena itu, penelitian ini diharapkan dapat memberikan solusi yang lebih adil, efisien, dan transparan dalam moderasi konten berbasis teks.

Manfaat praktis dari penelitian ini untuk memberikan pemahaman empiris terhadap pengaruh masing-masing teknik

terhadap akurasi dan efisiensi klasifikasi, menyediakan panduan bagi pengembang sistem deteksi perundungan berbasis NLP, dan menjadi dasar bagi pengembangan sistem moderasi konten yang lebih adaptif dan etis. Dengan demikian, penelitian ini tidak hanya memiliki nilai teoretis yang tinggi tetapi juga relevansi praktis yang besar. Penelitian ini dipandang penting untuk dilakukan seiring dengan meningkatnya fenomena perundungan siber dan kebutuhan akan sistem deteksi otomatis yang akurat, komprehensif, dan adil.

B. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan mengenai urgensi dan kompleksitas klasifikasi teks perundungan siber, penelitian ini berupaya merumuskan sejumlah pertanyaan berikut:

1. Bagaimana pengaruh penerapan metode *preprocessing* yang berbeda (*stopwords removal*, *stemming*, dan *lemmatization*), baik secara individual maupun kombinasi terhadap kinerja klasifikasi teks perundungan siber?
2. Bagaimana kinerja antara berbagai teknik *vectorization* (*one-hot encoding*, *Bag-of-Words*, dan TF-IDF) dalam merepresentasikan data teks untuk klasifikasi perundungan siber?
3. Bagaimana perbandingan kinerja algoritma klasifikasi yang berbeda (*Logistic Regression*, *Support Vector Machines*, *K-Nearest Neighbors*, *Decision Tree*, dan *Naive Bayes*) dalam mendeteksi teks perundungan siber?
4. Kombinasi metode *preprocessing*, *vectorization*, dan algoritma klasifikasi manakah yang menghasilkan kinerja klasifikasi teks

perundungan siber paling optimal berdasarkan metrik evaluasi yang relevan?

C. Batasan Masalah

Pembatasan ruang lingkup diperlukan agar penelitian ini dapat dilakukan lebih fokus dan terarah, serta untuk menghindari keluasan bahasan yang tidak relevan dengan tujuan utama penelitian. Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini hanya membahas klasifikasi teks dalam konteks *cyberbullying* atau perundungan siber. Teks yang dianalisis bersumber dari *dataset* publik yang telah tersedia secara daring, misalnya dari platform media sosial seperti Twitter atau forum diskusi daring lainnya. *Dataset* yang digunakan merupakan teks berbahasa Inggris dan telah memiliki label atau anotasi manual terkait kategori perundungan.
2. Metode *preprocessing* yang akan dibandingkan terbatas pada *stopwords*, *stemming*, dan *lemmatization*.
3. Teknik *vectorization* akan dibatasi pada tiga metode utama, yaitu *one-hot encoding*, *Bag-of-Words* (BoW), dan TF-IDF (*Term Frequency-Inverse Document Frequency*).
4. Algoritma klasifikasi yang akan dievaluasi dan dibandingkan kinerjanya terbatas pada *Logistic Regression* (LR), *Support Vector Machines* (SVM), *K-Nearest Neighbors* (KNN), *Decision Tree* (DT), dan *Naive Bayes* (NB).
5. Penilaian terhadap performa masing-masing kombinasi metode dilakukan dengan menggunakan metrik evaluasi standar dalam klasifikasi, meliputi *accuracy*, *precision*, *recall*, dan *F1-score*.

6. Eksperimen dilakukan menggunakan bahasa pemrograman Python dan pustaka *machine learning* populer seperti *Scikit-learn*, NLTK, dan *Pandas*. Proses pelatihan dan evaluasi model dilakukan dalam lingkungan komputasi lokal dan tidak mempertimbangkan penerapan sistem dalam skala produksi.

D. Tujuan Penelitian

Penelitian ini bertujuan untuk mengkaji secara komprehensif efektivitas kombinasi metode *preprocessing*, teknik *vectorization*, dan algoritma klasifikasi dalam meningkatkan akurasi klasifikasi teks yang mengandung perundungan siber. Tujuan spesifik yang hendak dicapai penelitian ini adalah sebagai berikut:

1. Mengidentifikasi dan menganalisis dampak spesifik dari penerapan metode *preprocessing* yang berbeda, yaitu *stopword*, *stemming*, dan *lemmatization*, serta kombinasi di antaranya terhadap kinerja model klasifikasi teks perundungan siber.
2. Membandingkan efektivitas metode *vectorization one-hot encoding*, *Bag-of-Words* (BoW), dan TF-IDF dalam merepresentasikan data teks perundungan siber, serta mengevaluasi bagaimana setiap metode tersebut memengaruhi kinerja klasifikasi.
3. Mengevaluasi dan membandingkan kinerja dari berbagai algoritma klasifikasi, yaitu *Logistic Regression*, *Support Vector Machines*, *K-Nearest Neighbors*, *Decision Tree*, dan *Naive Bayes*, dalam tugas klasifikasi teks perundungan siber.
4. Mengidentifikasi kombinasi metode *preprocessing*, *vectorization*, dan algoritma klasifikasi yang paling optimal dan

menghasilkan kinerja terbaik untuk deteksi teks perundungan siber, berdasarkan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*.

E. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat signifikan sebagai kontribusi baru bagi perkembangan ilmu pengetahuan di bidang *Natural Language Processing* (NLP) dan *Machine Learning* (ML), khususnya dalam konteks deteksi perundungan siber. Berikut ini beberapa manfaat penelitian tersebut:

1. Memberikan pemahaman mengenai bagaimana berbagai kombinasi metode *preprocessing*, *vectorization*, dan algoritma klasifikasi saling berinteraksi dan memengaruhi kinerja klasifikasi teks perundungan siber.
2. Mengidentifikasi kombinasi spesifik dari teknik *preprocessing*, *vectorization*, dan algoritma yang paling efektif untuk *dataset* teks perundungan siber.
3. Berkontribusi pada pengembangan kerangka kerja atau pedoman untuk memilih teknik pemrosesan teks yang paling sesuai berdasarkan karakteristik *dataset* dan domain aplikasi, tidak hanya terbatas pada perundungan siber.
4. Memberikan data empiris tentang bagaimana algoritma klasifikasi *machine learning* yang umum bekerja pada data teks perundungan siber, yang memiliki karakteristik linguistik unik.

BAB V
PENUTUP

A. Kesimpulan

Penelitian ini bertujuan untuk menguji dan membandingkan kinerja klasifikasi teks perundungan siber dengan menerapkan berbagai kombinasi metode *preprocessing*, *vectorizer*, dan algoritma pembelajaran mesin. Dari analisis mendalam terhadap hasil eksperimen, beberapa poin penting dapat disimpulkan:

1. Penggunaan metode *preprocessing* terbukti memiliki dampak signifikan terhadap kualitas data dan performa klasifikasi. Penghapusan *stopwords* secara konsisten meningkatkan kinerja model. Di antara teknik normalisasi kata, *lemmatization* menunjukkan sedikit keunggulan atau setidaknya sepadan dengan *stemming*, mengindikasikan bahwa pemulihan kata ke bentuk dasar yang linguistiknya lebih akurat lebih efektif dalam menangani nuansa teks perundungan siber.
2. Pilihan metode *vectorization* sangat menentukan representasi fitur teks. *One-Hot encoding* secara konsisten menghasilkan kinerja terendah karena menghasilkan representasi yang sangat jarang dan berdimensi tinggi tanpa menangkap hubungan antar kata. Sebaliknya, *Bag-of-Words* (BoW) menunjukkan peningkatan kinerja yang substansial. Namun, metode TF-IDF secara umum memberikan performa terbaik di antara semua *vectorizer* yang diuji, efektif dalam menyoroti kata-kata kunci yang paling diskriminatif untuk klasifikasi perundungan siber.

3. Dari algoritma yang dibandingkan, *Support Vector Machines* (SVM) dan *Logistic Regression* (LR) tampil sebagai yang paling unggul dalam mendeteksi teks perundungan siber. SVM secara khusus menunjukkan kinerja yang sangat kuat, terutama saat dipadukan dengan TF-IDF, berkat kemampuannya dalam menemukan batas pemisah optimal pada ruang fitur yang kompleks. Meskipun *Naive Bayes* juga menunjukkan hasil yang baik, algoritma seperti *Decision Tree* dan *K-Nearest Neighbors* (KNN) umumnya memiliki kinerja yang lebih rendah dibandingkan yang lain.
4. Berdasarkan seluruh evaluasi, kombinasi metode *stopwords removal* untuk *preprocessing*, TF-IDF untuk *vectorization*, dan algoritma *Support Vector Machines* (SVM) adalah konfigurasi yang paling optimal untuk tugas deteksi teks perundungan siber pada *dataset* yang digunakan. Kombinasi ini berhasil mencapai keseimbangan yang baik antara *recall* dan *precision*, menghasilkan *F1-score* tertinggi yaitu sebesar 0.863. Hal ini menegaskan bahwa pendekatan yang menyeluruh dalam pemilihan setiap tahapan pemrosesan data sangat penting untuk mengoptimalkan kinerja klasifikasi teks.

B. Saran

Berdasarkan temuan dan keterbatasan yang teridentifikasi dalam penelitian ini, beberapa saran untuk pengembangan dan eksplorasi di masa mendatang dapat dipertimbangkan:

1. Mengingat penelitian ini menggunakan data berbahasa Inggris, disarankan untuk menguji kombinasi optimal yang ditemukan

pada *dataset* perundungan siber dalam bahasa lain, seperti Bahasa Indonesia. Hal ini akan membantu menguji generalisasi dan adaptasi model terhadap karakteristik linguistik yang berbeda.

2. Meskipun *tuning hyperparameter* telah dilakukan pada tingkat dasar, eksplorasi yang lebih mendalam dengan teknik *fine-tuning* yang lebih canggih dapat dilakukan. Ini berpotensi meningkatkan kinerja model secara lebih signifikan.
3. Penelitian selanjutnya dapat membandingkan kombinasi metode *preprocessing* dan *vectorization* ini dengan model *deep learning,* misalnya jaringan saraf konvolusional, jaringan saraf berulang, atau model berbasis *transformer*. Pendekatan ini mungkin mampu menangkap pola dan hubungan yang lebih kompleks dalam teks.
4. Selain representasi tekstual, penambahan fitur-fitur lain seperti fitur linguistik, misalnya bagian dari ucapan serta struktur sintaksis, fitur sentimen yang telah ditentukan, atau bahkan *meta-data* dari postingan dapat dieksplorasi. Penggabungan fitur ini mungkin dapat memperkaya representasi data dan meningkatkan akurasi.
5. Jika *dataset* memiliki ketidakseimbangan kelas yang signifikan, teknik penanganan ketidakseimbangan, misalnya *oversampling* kelas minoritas atau *undersampling* kelas mayoritas, dapat diterapkan. Ini dapat membantu model untuk tidak bias pada kelas mayoritas dan meningkatkan kinerja, terutama pada metrik *recall* untuk kelas minoritas, misal kasus perundungan siber.

DAFTAR PUSTAKA

Aggarwal, C. C., & Zhai, C. (2012). *Mining text data*. Springer.

Alabdulwahab, A., Haq, M. A., & Alshehri, M. (2023). Cyberbullying detection using machine learning and deep learning. *International Journal of Advanced Computer Science and Applications*, 14(10).

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. O'Reilly Media, Inc.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152). ACM.

Buckland, M., & Gey, F. (1994). The relationship between recall and precision. *Journal of the American society for information science*, 45(1), 12-19.

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational intelligence magazine*, 9(2), 48-57.

Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural language engineering*, 29(3), 509-553.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education (6th ed.)*. Routledge.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21-27.

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Davidson, T., Warmsley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. arXiv preprint arXiv:1703.04009.

Dinakar, K., Reichart, R., & Lieberman, H. (2011). Modeling the detection of textual cyberbullying. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 5, No. 3, pp. 11-17).

Duda, R. O., Hart, P. E., Stork, D. G. (2001). *Pattern classification (2nd ed.)*. John Wiley & Sons.

Eronen, J., Ptaszynski, M., Masui, F., Smywiński-Pohl, A., Leliwa, G., & Wroczynski, M. (2021). Improving classifier training efficiency for automatic cyberbullying detection with feature density. *Information Processing & Management*, 58(5), 102616.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), 861-874.

Few, S. (2009). *Now you see it: Simple visualization techniques for quantitative analysis*. Analytics Press.

Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. *Journal of machine learning research*, 3(Mar), 1289-1305.

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems (2nd ed.)*. O'Reilly Media.

HaCohen-Kerner, Y., Miller, D., & Yigal, Y. (2020). The influence of preprocessing on text classification using a bag-of-words representation. *PloS one*, 15(5), e0232525.

Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3), 146-162.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758)*. New York: springer.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression (2nd ed.)*. John Wiley & Sons.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: with applications in R (Vol. 103)*. New York: springer.

Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. *Machine Learning: ECML-98*, 137–142.

Joachims, T. (1999). Transductive inference for text classification using support vector machines. In *Icml* (Vol. 99, pp. 200-209).

Joachims, T. (2002). *Learning to classify text using support vector machines: Methods, theory and algorithms*. Springer.

Jurafsky, D., & Martin, J. H. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Prentice Hall.

Kannan, S., Gurusamy, V., Vijayarani, S., Ilamathi, J., Nithya, M., Kannan, S., & Gurusamy, V. (2014). Preprocessing techniques for text mining. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.

Kelleher, J. D., Mac Namee, B., & D'Arcy, A. (2015). *Fundamentals of machine learning for predictive data analytics: Algorithms, worked examples, and case studies*. MIT Press.

Kokshagina, O., Reinecke, P. C., & Karanasios, S. (2023). To regulate or not to regulate: unravelling institutional tussles around the regulation of algorithmic control of digital platforms. *Journal of Information Technology*, 38(2), 160-179.

Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin*, 140(4), 1073.

Krovetz, R. (1993, July). Viewing morphology as an inference process. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 191-202).

Kwon, Y. M., Jun, S. H., Gal, W. M., & Lim, M. J. (2018). The performance comparison of the classifiers according to binary bow, count bow and Tf-Idf feature vectors for malware detection. *International Journal of Engineering & Technology*, *7*(3.33), 15-22.

Lewis, D. D. (1998, April). Naive (Bayes) at forty: The independence assumption in information retrieval. In *European conference on machine learning* (pp. 4-15). Berlin, Heidelberg: Springer Berlin Heidelberg.

Li, S., & Li, Z. (2025). Hate Speech Detection and Online Public Opinion Regulation Using Support Vector Machine Algorithm: Application and Impact on Social Media. *Information*, 16(5), 344.

Lovins, J. B. (1968). Development of a stemming algorithm. *Mechanical Translation and Computational Linguistics*, 11(1-2), 22-31.

Lukman, K., & Novianto, S. (2025). Komparasi Algoritma Naïve Bayes dan SVM untuk Identifikasi Cyberbullying Selebriti di Media Sosial Twitter. *Jurnal Algoritma*, 22(1), 970-981.

Mali, M., & Atique, M. (2021). The Relevance of preprocessing in text classification. In *Proceedings of Integrated Intelligence Enable Networks and Computing: IIENC 2020* (pp. 553-559). Singapore: Springer Singapore.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.

McCallum, A., & Nigam, K. (1998, July). A comparison of event models for naive bayes text classification. In *AAAI-98 workshop on learning for text categorization* (Vol. 752, No. 1, pp. 41-48).

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. arXiv preprint arXiv:1301.3781.

Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.

Muminovic, A. (2025). *Benchmarking Large Language Models for Cyberbullying Detection in Real-World YouTube Comments*. arXiv preprint arXiv:2505.18927.

Muneer, A., & Fati, S. M. (2020). A comparative analysis of machine learning techniques for cyberbullying detection on twitter. *Future Internet*, 12(11), 187.

Ng, A., & Jordan, M. (2001). *On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. Advances in neural information processing systems*, 14.

Philipo, A. G., Sarwatt, D. S., Ding, J., Daneshmand, M., & Ning, H. (2024). Assessing text classification methods for cyberbullying detection on social media platforms. *IEEE Transactions on Information Forensics and Security*.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, *14*(3), 130-137.

Powers, D. M. (2020). *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation*. arXiv preprint arXiv:2010.16061.

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.

Ramos, J. (2003, December). Using TF-IDF to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning* (Vol. 242, No. 1, pp. 29-48).

Ricol, J. (2022). *Ethical Considerations in AI-Driven Cybersecurity: Balancing Automation and Human Oversight*.

Rokach, L. & Maimon, O. Z. (2015). *Data mining with decision trees: theory and applications (Vol. 81)*. World scientific.

Romindo, R., Pangaribuan, J. J., & Barus, O. P. (2023). Implementasi algoritma TF-IDF dan Support Vector Machine terhadap analisis pendeteksi komentar cyberbullying di media sosial TikTok. *Device*, 13(1), 124-134.

Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*. McGraw-Hill.

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 34(1), 1-47.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin and Company.

Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1), 11-21.

Tharwat, A. (2021). Classification assessment methods. *Applied computing and informatics*, 17(1), 168-192.

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*, 37, 141-188.

Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer Science & Business Media.

Wibowo, A. E., Khairi, A., Humairoh, H., Fadillah, M. I., & Hartawan, M. J. D. (2023). Text Mining: Sistem Prediksi Cyberbullying pada Platform Twitter menggunakan Logistic Regression, KNN, dan Naive Bayes. J. *Rekayasa Elektro Sriwij*, 4(1), 17-23.

Witten, I. H., Frank, E., Hall, M. A., Pal, C. J. (2017). *Data Mining: Practical Machine Learning Tools and Techniques (4th ed.)*. Elsevier.