



© Hak cipta milik UIN Suska Riau

PENERAPAN SUPPORT VECTOR MACHINE DAN FASTTEXT UNTUK MENDETEKSI HATE SPEECH DAN ABUSIVE PADA TWITTER

TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

AFDHAL ZIKRI
NIM.11651103679



FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU

PEKANBARU

2023

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN

**PENERAPAN *SUPPORT VECTOR MACHINE* DAN *FASTTEXT*
UNTUK MENDETEKSI *HATE SPEECH* DAN *ABUSIVE* PADA
TWITTER**

**JURUSAN TEKNIK INFORMATIKA
UIN SUSKA RIAU**

TUGAS AKHIR

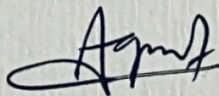
Oleh

AFDHAL ZIKRI

NIM. 11651103679

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 14 Juli 2023

Pembimbing I,



SURYA AGUSTIAN, S.T., M.Kom

NIP. 197608302011011003

LEMBAR PENGESAHAN

PENERAPAN *SUPPORT VECTOR MACHINE* DAN *FASTTEXT* UNTUK MENDETEKSI *HATE SPEECH* DAN *ABUSIVE* PADA TWITTER

JURUSAN TEKNIK INFORMATIKA

UIN SUSKA RIAU

Oleh

AFDHAL ZIKRI

NIM. 11651103679

Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 14 Juli 2023

Mengesahkan,

Ketua Jurusan,



Dr. HARTONO, M.Pd

NIP. 196403111992031003

IWAN ISKANDAR, M.T

NIP. 198212162015031003

DEWAN PENGUJI

Ketua : Muhammad Irsyad, S.T., M.T
Pembimbing I : Surya Agustian, S.T., Kom
Penguji I : Muhammad Fikry, S.T., M.Sc
Penguji II : Pizaini, S.T., M.Kom

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : Afdhal Zikri

NIM : 11651103679

Tempat/Tgl.Lahir : Sungai Pakning, 25 Oktober 1998

Fakultas/Pascasarjana : Sains dan Teknologi/S1

Prodi : Teknik Informatika

Judul Disertasi/Thesis/Skripsi/Karya Ilmiah Lainnya* :

PENERAPAN SUPPORT VECTOR MACHINE DAN FASTTEXT UNTUK MENDETEKSI HATE SPEECH DAN ABUSIVE PADA TWITTER

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Skripsi dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena Skripsi saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Skripsi saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 17 Juli 2023

Yang membuat pernyataan

 
Afdhal Zikri
NIM.11651103679

**pilih salah satu sesuai jenis karya tulis*

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kejasanaan di suatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan didalam daftar pustaka.

Pekanbaru, 17 Juli 2023

Yang membuat pernyataan,

AFDHAL ZIKRI

NIM.1651103679

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSEMBAHAN



Assalamua'alaikum Warahmatullahi Wabarakatuh

Alhamdulillahirobbil'alamiin. Puji syukur kepada Allah SWT atas seluruh rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan penelitian dan penulisan Laporan Tugas Akhir di Jurusan Teknik Informatika dengan Judul “Penerapan Support Vector Machine dan FastText untuk Mendeteksi Hate Speech dan Abusive pada Twitter”. Laporan Tugas Akhir ini merupakan prasyarat untuk memperoleh gelar sarjana strata satu di Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau (UIN SUSKA RIAU).

Selama dilaksanakannya penelitian dan penulisan Laporan Tugas Akhir ini, penulis mendapatkan banyak pengetahuan, pengalaman, bimbingan, dukungan dan juga arahan dari semua pihak yang telah membantu hingga Laporan Tugas Akhir ini dapat diselesaikan. Untuk itu pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Allah SWT dengan segala rahmat-Nya memberikan semua yang terbaik dan dengan hidayah-Nya memberikan petunjuk sehingga dalam penyusunan laporan ini berjalan lancar.
2. Bapak Prof. Dr. Hairunnas Rajab, M.Ag. selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Bapak Dr. Hartono, M.T selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Bapak Iwan Iskandar, M.T selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
5. Ibu Fitra Kurnia, S.Kom, M.T selaku Penasehat Akademik yang telah memberikan bimbingan selama penulis melakukan studi di jurusan Teknik Informatika Universitas Negeri Sultan Syarif Kasim Riau.
6. Bapak Surya Agustian, S.T, M.Kom, selaku dosen pembimbing Tugas Akhir penulis, yang telah sangat banyak berbagi waktu, ilmu dan wawasan yang

Hak Cipta Dilindungi Undang-Undang

© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dimiliki kepada penulis sehingga penelitian dan Laporan Tugas Akhir ini selesai.

7. Bapak Muhammad Fikry, S.T, M.Sc selaku Penguji I dan Bapak Pizaini, S.T, M.Kom selaku Penguji II, terimakasih telah memberikan banyak masukan dan saran dalam menyelesaikan tugas akhir ini.
8. Seluruh Bapak/Ibu dosen Teknik Informatika yang telah sabar memberikan tunjuk ajar serta ilmu yang bermanfaat kepada penulis selama masa perkuliahan.
9. Kedua orang tua penulis, Ayah dan Ibu yang tidak pernah letih memberikan semangat, nasehat, dan doa yang tulus untuk anaknya. Semoga Allah SWT senantiasa memberikan kesehatan kepada Ayah dan Ibu.
10. Teman-teman seperjuangan kelas B angkatan 2016 yang selalu berjuang bersama dan selalu saling mendukung.
11. Teman-teman Kos Awak, yang telah banyak memberikan saran, semangat, dan motivasi kepada penulis, semoga kita semua dapat mencapai apa yang diinginkan.
12. Semua pihak yang tidak dapat penulis sebutkan satu persatu. Terimakasih atas dukungan moril maupun materil dalam pengerjaan Laporan Tugas Akhir ini.

Semoga laporan ini dapat bermanfaat bagi penulis khususnya maupun pembaca pada umumnya. Serta dapat menjadi referensi dan rujukan bagi hal-hal yang bermanfaat. Penulis berharap adanya kritik dan saran guna memperbaiki atau sebagai pengembangan kedepannya. Akhir kata penulis ucapkan terima kasih dan selamat membaca.

Wassalamualaikum Wr. Wb

UIN SUSKA RIAU
Pekanbaru, 17 Juli 2023

Penulis



Penerapan Support Vector Machine dan FastText untuk Mendeteksi Hate Speech dan Abusive pada Twitter

Afhdal Zikri*, Surya Agustian

Faculty of Science and Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹*11651103678@Students.uin-suska.ac.id, ²surya.agustian@uin-suska.ac.id

Email Penulis Korespondensi: ¹11651103678@Students.uin-suska.ac.id

Abstrak—Ujaran kebencian (hate speech) dan Bahasa kasar (abusive language) adalah salah satu kecenderungan negatif yang sering muncul di media sosial. Selain itu, karena kemajuan teknologi dan penggunaan internet yang pesat, siapa pun kini dapat terlibat dalam melontarkan ujaran kebencian atau bahkan bahasa kasar, seperti pada Twitter, yang akhirnya menyebabkan perengkaman di platform media sosial tersebut. Deteksi otomatis kandungan bahasa kasar dan ujaran kebencian baik untuk melindungi khususnya di sisi aplikasi user, untuk memfilter isi tweet yang dapat merusak tatanan kehidupan sosial di dunia nyata. Tujuan penelitian ini adalah membuat model klasifikasi menggunakan Support vector machine dengan fitur word embeddings FastText dengan menggunakan suatu benchmark dataset, untuk mengidentifikasi apakah suatu tweet mengandung ujaran kebencian dan/atau bahasa kasar. Kontribusi penelitian ini adalah meningkatkan performa dari metode baseline SVM (support vector machine) dengan fitur input word embeddings FastText. Hasil eksperimen diukur berdasarkan akurasi yang akan dibandingkan dengan beberapa metode machine learning lainnya menggunakan dataset yang sama, yaitu 13.167 tweet. Hasil pengujian dengan model SVM yang paling optimal, mendapatkan akurasi rata-rata 82,65%, dengan akurasi pada kelas hate speech, abusive language dan level hate speech secara berturut-turut adalah 84,92%, 86,60%, dan 76,43%. Hasil ini lebih baik dibandingkan beberapa metode machine learning konvensional, namun belum dapat melampaui hasil yang dicapai oleh metode berbasis deep learning.

Kata Kunci: Bahasa Kasar; Ujaran Kebencian; Support Vector Machine; Fasttext; Klasifikasi Teks

Abstract—Hate speech and abusive language are negative tendencies that often appear on social media recently. In addition, due to the advancement of technology and the rapid growth of the internet, anyone can now engage in hate speech or even offensive expression such as in Twitter, which eventually leads to fights on that social media platforms. Automatic detection of offensive contents and hate speech is recommended to be applied, especially on the user application's side, to filter tweet contents which destruct social life in the real world. The purpose of this research is to create a classification model using Support Vector Machine with FastText word embeddings features, to classify if a tweet contains hate speech and/or offensive language. Our contribution in this research is an improvement in performance from the baseline SVM (support vector machine) with FastText word embeddings features input. The experiment results will also be compared with several machine learning method that have been reported using the same dataset of 13,167 tweets. The experiment using the most optimal SVM model, yields an average accuracy of 82.65%, with the accuracies of the hate speech class, abusive language class and hate speech level, are 84.92%, 86.60% and 76.43% respectively. These results are better than conventional machine learning, but cannot exceed the results achieved by deep learning.

Keywords: Abusive Language; Hate Speech; Support Vector Machine; Fasttext; Text Classification

1. PENDAHULUAN

Perkembangan teknologi informasi telah menciptakan banyak peluang baru untuk pertukaran informasi, pengetahuan, kabar, berita, meskipun dibatasi oleh jarak, waktu dan tempat. Secara informal dan tidak resmi (unofficial), pertukaran informasi dapat terjadi secara masif melalui jejaring media sosial[1]. Di samping itu, jejaring sosial membuka kesempatan tidak terbatas dalam berpendapat memungkinkan munculnya interaksi, komunikasi dan kerja sama antar penggunanya secara efektif, cepat, akurat dan relatif ekonomis.

Media sosial menawarkan banyak kemudahan dan kebebasan bagi penggunanya, yang kemudian banyak memunculkan penyalahgunaan dalam pemakaian, seperti hal-hal yang melanggar norma, kebiasaan masyarakat, dan hukum yang berlaku di negara. Penyebaran berita bohong, promosi barang dan jasa yang menyalahi aturan dan norma, menghasut terjadinya permusuhan, makin mudah ditemukan dalam berbagai postingan media sosial. Selain itu, menyebarkan ujaran kebencian dan bahasa yang menyinggung merupakan bentuk penyalahgunaan fungsi media sosial [2]. Ujaran kebencian adalah suatu tindakan yang dilakukan oleh kelompok atau individu yang bertujuan untuk memprovokasi kelompok atau individu lain. Fitnah, provokasi, dan atribusi merupakan bentuk ujaran kebencian. Ujaran kebencian di media sosial seringkali dikaitkan dengan konteks jenis kelamin, warna kulit, agama, ras dan lainnya [3].

Bahasa kasar merupakan sebuah perkataan yang memiliki makna atau ungkapan yang menyinggung atau kasar, baik yang berkaitan dengan pelecehan seksual yang vulgar, lelucon, dan menghina individu. Namun, yang sering terjadi di dalam konteks bahasa yang menyinggung seringkali menyebabkan ujaran kebencian dan dilarang di tempat umum seperti jejaring sosial [4].

Salah satu contoh penyalahgunaan media sosial dapat dilihat pada kasus Muhammad Said Didu yang dipolisikan terkait ujaran kebencian terhadap Presiden Joko Widodo dan Menteri Agama Yaqut Cholil Qoumas. Didalam kasus yang terjadi di platform Twitter Said Didu menyatakan sebuah tweet yang berisi 'Terima kasih atas penjelasan mas Qodari. Akhirnya kami tahu bahwa Bpk Presiden inginkan Menag utk "menggebuk" islam. Sekali lagi terima kasih', setelah itu pada bulan Desember 2020 Wawan selaku ketua PAC Ansor Jagakarsa melaporkan

Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Said Didu ke Bareskrim dan diterima dengan nomor LP/B/0719/XII/2020/BARESKRIM dengan dugaan tindak pidana merujuk pada golongan sara bagian ujaran kebencian sesuai dengan pasal 45A ayat (2) Jo pasal 28 ayat (2) dan/atau pasal 207 KUHP.

Berbagai penelitian telah dilakukan untuk mendeteksi secara otomatis adanya ujaran kebencian di dalam media sosial, khususnya twitter. Metode support vector machine (SVM) digunakan pada [5] untuk klasifikasi sentimen pada teks, dan untuk klasifikasi ujaran kebencian pada media sosial. Selain itu, berbagai metode machine learning lainnya juga pernah dilaporkan untuk tugas klasifikasi sentimen maupun deteksi ujaran kebencian pada twitter, beberapa dibahas di dalam menggunakan Naïve Bayes (NB)[6], Random Forest[7], Decision Tree[8], dan K-NN (k-Nearest Neighbor)[9].

SVM digunakan dalam [10] untuk mengklasifikasikan data tweet yang berisi opini, untuk mempermudah penonton untuk melihat opini tentang review film. Dataset terdiri atas 1.027 tweet dari tweet untuk film-film yang populer tahun 2016. Hasil pemeringkatan opini dibagi menjadi 3 yaitu opini positif, negatif dan netral. Terhadap akurasi klasifikasi menggunakan rata-rata 60, 70, 80, dan 90 persen data latih, secara berturut-turut adalah 76,06 persen, 76,83 persen, 81,07 persen, dan 83,3 persen.

Klasifikasi biner pada[11] membagi data menjadi dua kelas yakni kelas “positif” yang bermakna mendukung dari wacana pemindahan ibu kota baru dan juga kelas “negatif” yang memiliki makna tidak setuju dengan pemindahan ibu kota baru. Data yang digunakan dalam penelitian ini sebanyak 1.236 cuitan yang di ambil dari twitter. Hasil pengujian akurasi sebesar 96,68%, precision 95,82%, dan recall 94,04%. Ensemble Feature disusun sebagai input untuk SVM pada [5] untuk mengidentifikasi ujaran kebencian pada media sosial facebook. Fungsi ensemble (ensemble function) menggabungkan fungsi-fungsi yang tidak selaras dari teks agar memudahkan proses identifikasi. Sedangkan metode Support Vector Machine berperan dalam mengidentifikasi teks berdasarkan data fungsi yang diperoleh dari ensemble function. Akurasi yang didapatkan dalam pengujian sistem adalah 70% untuk mengidentifikasi ujaran kebencian di jejaring Facebook dengan benar.

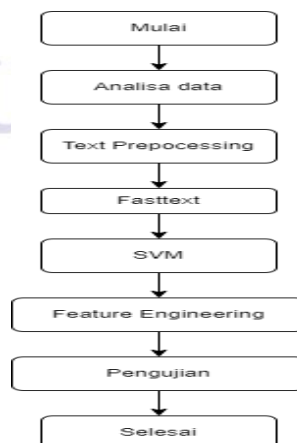
Pada penelitian klasifikasi teks, data yang akan diklasifikasikan harus berbentuk vector. Fasttext adalah salah satu metode word embedding yang ramai digunakan dan merupakan metode terbaru. Pada penelitian ini[12] yang membandingkan kinerja word embedding word2vec, glove, dan fasttext. Fasttext menunjukkan hasil kinerja yang lebih baik menggunakan f-measure dengan dataset 20 newsgroup 0,979 dan dataset 20 reuters news 0.715, untuk hasil kinerja word2vec, glove secara berturut-turut adalah 0.925, 0.958, dan 0.694, 0.688.

Penelitian tentang perbandingan metode Naïve Bayes, SVM dengan fitur bag of word TF-IDF, LSTM dan SVM dengan fitur word embedding Word2vec pada dataset yang tidak seimbang yang diusut oleh [13] mendapatkan hasil terbaik pada metode SVM dengan fitur word2vec. Metode SVM dengan word2vec mendapatkan akurasi sebesar 69% dan f1-score 65%.

Rata-rata penelitian terdahulu yang telah disebutkan di atas, hanya memeriksa atau mengklasifikasi 1 kelas saja, dengan label biner (positif dan negatif)[11], atau kadang-kadang juga menyertakan data tanpa sentimen (netral)[10], atau apakah suatu tweet ada/tidak mengandung ujaran kebencian[5]. Penelitian ini melakukan klasifikasi yang lebih mendalam terhadap suatu tweet, yaitu mendeteksi adanya ujaran kebencian, mendeteksi penggunaan bahasa kasar, dan menentukan level (strong, medium, low) dari ujaran kebencian yang terdeteksi. Metode yang diusulkan adalah Support Vector Machine (SVM), dengan input berupa vektor word embedding FastText dengan melakukan optimasi pemilihan fitur dan penyesuaian (tuning) parameter SVM untuk mendapatkan hasil akurasi yang terbaik.

2. METODOLOGI PENELITIAN

Metodologi penelitian merupakan tahapan pada penelitian, adapun metodologi penelitian dapat dilihat berikut ini:



Gambar 1. Metode Penelitian

2. Dilarang mengumumkkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

c. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.

d. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

e. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

f. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

g. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

h. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

i. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

j. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

k. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

l. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

m. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

n. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

o. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

p. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

q. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

r. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

s. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

t. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

u. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

v. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

w. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

x. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

y. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

z. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

aa. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

ab. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

ac. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

ad. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

ae. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

af. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

ag. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

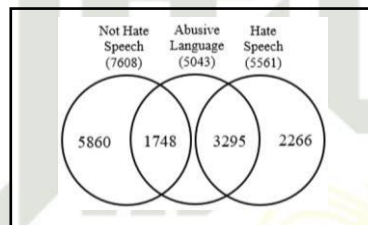
ah. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

ai. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mengutip sumber.

Penelitian ini terdapat 6 tahapan penelitian secara garis besar, yaitu rumusan masalah, seperti telah dibahas pada bab dan pendahuluan, kemudian melakukan analisa data, penerapan text preprocessing, dan dilanjutkan dengan mengembangkan model bahasa dengan Fasttext word embeddings. Setelah itu, pengembangan metode Support Vector machine (SVM) untuk tugas klasifikasi, dalam penelitian ini dilakukan terhadap 3 kelas yaitu abusive language, hate speech dan level dari hate speech. Tahap selanjutnya adalah parameter tuning, yaitu mengoptimasi model machine learning untuk mendapatkan hasil paling optimal. Pengujian dilakukan untuk memvalidasi hasil dari model yang paling optimal, untuk dibandingkan dengan penelitian dengan menggunakan data sejenis.

Analisa Data

Data yang digunakan adalah berasal dari penelitian Ibrahim dan Budi [14] dengan crawling menggunakan twitter Search API dan implementasi menggunakan Tweepy Library dengan waktu pengerjaan selama 7 bulan dari 20 Januari 2018 sampai 10 september 2018. mengenai deteksi multilabel hate speech dan abusive language pada twitter bahasa Indonesia. Data tersebut sudah di proses dalam dua tahapan anotasi data, tahapan pertama melabel data hate speech dan non hate speech dan tahapan kedua tahapan anotasi data yang bersifat abusive, yang memberikan label 13.169 tweet, dengan 7.608 tweet diantaranya bukan ujaran kebencian dan 5.561 dianggap sebagai ujaran kebencian. Statistik kelas hate speech dan abusive dapat dilihat dalam diagram venn pada gambar 2 di bawah ini. Tidak semua yang dianggap bahasa kasar adalah ujaran kebencian dan tidak semua yang disebutkan dalam ujaran kebencian merupakan bahasa kasar.



Gambar 2. Distribusi Data[14]

Penelitian ini menggunakan dataset tersebut dengan fokus kepada tiga kelas, yaitu ujaran kebencian (hate speech), bahasa kasar (abusive language), dan level dari hate speech. Dalam kelas ujaran kebencian terdapat dua kategori (bahasa kasar dan bukan bahasa kasar). Dalam kelas bahasa kasar ada dua label (bahasa kasar dan bukan bahasa kasar), dan kelas level memiliki tiga kategori (kuat, sedang, dan lemah). Salah satu contoh hate speech, abusive, dan non hate speech secara berturut-turut yaitu “Kamu transgender atau gmn anjing”, “Edan emang goollnya Fabiano”, “Kadang aku berfikir, kenapa aku tetap percaya pada Tuhan padahal aku selalu jatuh berkali-kali. Kadang aku merasa Tuhan itu meninggalkan aku sendirian. Ketika orangtuaku berencana berpisah, ketika kakakku lebih memilih jadi Kristen”.

2.2 Text Preprocessing

Text preprocessing merupakan proses awal pemilihan teks yang akan disusun melalui serangkaian tahapan agar dapat digunakan pada tahapan berikutnya. Penelitian ini meliputi beberapa tahapan seperti Case folding, Tokenizing, Punctuation dan Stopword removal.

Langkah pertama adalah Case folding. Case folding merupakan satu dari text preprocessing yang digunakan mengubah huruf yang berada di dalam dokumen dengan format huruf besar di rubah menjadi huruf kecil. Huruf dari "a" ke "z" yang hanya digunakan. Karakter selain huruf akan dihilangkan dan digunakan sebagai delimiter.

Langkah selanjutnya adalah tokenisasi, yang merupakan proses mengubah teks yang bisa berupa kalimat, paragraf, atau dokumen menjadi token atau potongan teks tertentu. Pemisah antar token biasanya terdiri dari tanda baca dan spasi.[15].

Langkah selanjutnya adalah Punctuation, yaitu proses menghilangkan tanda baca, kalimat, angka, dan karakter tambahan selain alfabet. Karena karakter yang dimaksud berfungsi sebagai pemisah kata (delimiter) dan tidak ada hubungannya dengan penyusunan teks.

Langkah terakhir adalah stopword removal, ini adalah proses menghilangkan kata-kata yang tidak perlu dan kata-kata yang tidak ada hubungannya dengan komentar atau tweet seperti preposisi, kata ganti, kata sambung dengan penelitian ini, sehingga meninggalkan kata-kata yang penting.[16].

2.3 FastText

FastText adalah library yang di kembangkan oleh Facebook yang dapat digunakan untuk membantu menghasilkan representasi kata secara efisien dan dapat membantu mengklasifikasikan kata di luar pustaka[11]. FastText merupakan salah satu metode word embedding, yang mana fasttext merupakan pengembangan dari library Word2Vec, fasttext dikenal lebih lama sebagai metode untuk proses word embedding [12].

2. Diarangi mengumumkkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hal ini bertujuan untuk melindungi hak cipta dan integritas karya tulis ini tanpa mencantumkan dan menyebutkan sumber.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

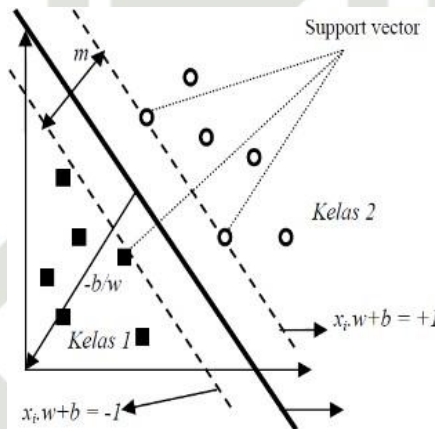
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

Berbeda dengan Word2Vec, fasttext menggunakan n-gram daripada hanya satu kata cara yang jelas dan ringkas untuk terjemahan. Contoh implementasi n-gram pada kata “bahagia” dengan trigram (n=3) berupa “bah”, “haga”, “hag”, “agi”, “gia”. Keunggulan FastText adalah waktu pemrosesannya yang relatif cepat. FastText dapat mengenali kata-kata yang belum pernah muncul di kamus, berbeda dengan Word2Vec, di mana situasi seperti ini akan mengakibatkan error.

FastText adalah metode yang digunakan untuk proses pembobotan yang mengubah kata menjadi vektor. Dataset yang besar mampu memberikan representasi struktur yang tidak muncul dalam data latih. Jika istilah nir muncul, fasttext dapat digunakan untuk memecah sebagai N-gr untuk menerima vektor. Fasttext adalah fitur bahasa yang mengubah teks menjadi vektor dengan dimensi 128.

2.4 Support Vector Machine

Support Vector Machine termasuk algoritma yang memiliki kemampuan klasifikasi dengan fungsi memisahkan data (hyperplane) dengan margin terbesar, memungkinkannya memisahkan dua set data secara optimal[18]. Support Vector Machine memiliki kemampuan untuk menganalisis hyperplane yang memaksimalkan margin antara dua kelas yang beda, tetapi juga memiliki kapasitas untuk gagal karena masalah pemilihan parameter atau fitur yang tepat. Hasil dari algoritma klasifikasi dipengaruhi secara signifikan oleh pemilihan fitur dan pengaturan parameter untuk Support Vector Machine.



Gambar 3. Margin Hyperplane[18]

Support Vector Machine pertama kali diperkenalkan tahun 1992[19] dan di kembangkan pada tahun 1995 [20] untuk klasifikasi data numerik, tetapi sekarang telah terbukti sama sukses dan cepatnya ketika digunakan untuk masalah data teks. Dimensi data teks yang tinggi, di mana ada beberapa aspek tertentu tidak penting, tetapi cenderung berkorelasi dengan satu dengan yang lain dan diklasifikasi secara linier, membuat metode Support Vector Machine (SVM) cocok untuk mengklasifikasikan data teks.

2.5 Feature Engineering

Feature engineering mengacu pada metode menemukan dan menggunakan fitur baru untuk meningkatkan fungsionalitas mesin yang membantu meningkatkan akurasi. Kemampuan ini mungkin diintegrasikan ke dalam program yang sudah ada sebelumnya atau dibuat dari awal. Langkah tersebut juga memperhitungkan pemilihan dan transformasi variabel untuk mendapatkan hasil terbaik[15].

2.5.1 Scaling MinMax

Pada penelitian ini akan dilakukan penambahan fitur Scaling, fitur Scaling ditambahkan untuk menghasilkan model yang lebih akurat untuk menunjang nilai akurasi dibandingkan dengan yang diperoleh pada tahap preprocessing. Di bagian ini, kita akan menggunakan dua Scaler MinMax dan Robust untuk menentukan fitur terbaik yang melakukan improve terhadap data.

Scaling MinMax termasuk dalam bagian tahapan untuk menormalisasikan nilai vector dengan melakukan tahapan transformasi linier terhadap vector awal dengan rentang 0-1[21]. Rumus dari scaling MinMax dapat dilihat di bawah ini

$$X(std) = \frac{X - \text{Min}(X)}{\text{Max}(X) - \text{Min}(X)} \tag{1}$$

$$X(scale) = X(std) \times (\text{Max}(F) - \text{Min}(F)) + \text{Min}(F) \tag{2}$$

Dimana X adalah nilai real dari sampel, dan F adalah nilai dari Fitur yang diinginkan. Misalnya max(F) adalah 1, dan min(F) adalah 0 (atau rentang nilai fitur adalah dari 0 sampai 1), sedangkan maksimum dari sampel mungkin saja ada yang lebih besar dari 1, atau tidak ada yang bernilai terendah 0. Teknik ini sangat cocok

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

a. Pengutipan hanya untuk kepentingan karya tulis ini tanpa mencantumkan dan menyebutkan sumber.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

c. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

d. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

e. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

f. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

g. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

h. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

i. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

j. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

k. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

l. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

m. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

n. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

o. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

p. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

q. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

r. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

s. Dilarang mengutip sebagai referensi atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

diterapkan pada hasil vektor word embeddings, di mana ada nilai komponen vektor yang lebih besar dari 1. Sedangkan SVM sangat baik bila menerima input fitur dengan nilai yang ternormalisasi di antara 0 dan 1.

2.2 Scaling Robust

Scaling robust digunakan untuk mengubah vector dengan menggunakan median dan quartiles[22]. Rumus dari scaling robust dapat dilihat di bawah ini.

$$X(\text{scale}) = \frac{X - \text{median}}{p75 - p25} \quad (3)$$

Dengan X adalah nilai real dari sample, dan median adalah nilai tengah dari sample (kuartil ke-2), sedangkan p75 dan p25 adalah berturut-turut nilai pada kuartil ke-3 (q3) atau 75% dari data nilai terurut, dan nilai pada kuartil ke-1 (q1) atau 25% dari data terurut.

2.3 Pengujian

Pada tahap ini akan dilakukan pengujian terhadap model Support vector machine. Pengujian dilakukan setelah mendapatkan kombinasi fitur dan parameter terbaik masing-masing pencarian optimal. Confusion Matrix berfungsi untuk menampilkan hasil klasifikasi dengan memberikan representasi visual dari data yang diklasifikasikan dengan benar atau salah.[15]. Confusion matrix biasanya untuk mencari akurasi, precision dan recall, namun pada penelitian ini confusion matrix yang digunakan hanya fungsi akurasinya saja. Rumus dari akurasi dapat dilihat di bawah ini.

$$\text{akurasi} = \frac{TP+TN}{\text{jumlah seluruh sampel data}} \times 100 \quad (4)$$

dengan TP adalah jumlah data kelas positif yang benar diprediksi sebagai positif (misalnya mengandung hate speech atau abusive language), dan TN adalah jumlah data kelas negatif yang benar diprediksi sebagai negatif (misalnya tidak mengandung hate speech dan abusive language).

3. HASIL DAN PEMBAHASAN

3.1 Pencarian Kombinasi Teks Preprocessing

Pada tahap ini dilakukan pencarian kombinasi fitur dari text preprocessing yaitu case folding, stopword removal, punctuation removal, filtering special character dan angka. Selain untuk menyeleksi data, kombinasi berbagai teknik text preprocessing dilakukan untuk mencari hasil akurasi terbaik. Dalam penelitian ini, baseline model SVM yang dikembangkan mengikuti teknik preprocessing pada [19, 20], yaitu menghapus angka dan karakter spesial. Selanjutnya, komposisi penerapan case folding, stopword removal dan punctuation removal dilakukan sesuai dengan skema eksperimen seperti pada Tabel 1 berikut. C1 adalah ID eksperimen baseline dari penelitian ini, yaitu tanpa melakukan pemrosesan (pre-processing) pada teks asli dari tweet.

Tabel 1. Eksperimen Preprocessing

Case Folding	Stopword	Punctuation	Id Eksperimen
Tidak	Tidak	Tidak	C1
Ya	Tidak	Tidak	C2
Tidak	Ya	Tidak	C3
Tidak	Tidak	Ya	C4
Tidak	Ya	Ya	C5
Ya	Tidak	Ya	C6
Ya	Ya	Tidak	C7
Ya	Ya	Ya	C8

3.2 Pencarian model Optimal

Untuk mendapatkan model optimal, yaitu nilai akurasi tertinggi, dilakukan pengujian menggunakan Support Vector Machine terhadap masing-masing kombinasi text preprocessing dan feature engineering, secara grid search. Parameter SVM yang digunakan dalam grid search adalah nilai C= {0.1, 1, 10}, dan gamma={1, 0.1, 0.001}. Hasil terbaik dari grid search terhadap data training 80% dan data uji 20%, untuk setiap kombinasi eksperimen, dilaporkan dalam Tabel 2 di bawah ini. ID eksperimen C1 dengan fitur Non-Scaling adalah metode baseline yang menjadi kondisi awal sistem sebelum dioptimasi.

Tabel 2. Hasil Akurasi dari Pencarian Model Optimal untuk Kelas HS, AB dan Level

Id Eksperimen	Feature	Kelas			Akurasi Rata-rata
		HS	AB	Level	
C1	Non Scaling	80.05 %	84.08 %	73.63 %	79.25 %



Id Eksperimen	Feature	Kelas			Akurasi Rata-rata
		HS	AB	Level	
C2	MinMax	81.84 %	85.19 %	72.37 %	79.80 %
	Robust	79.86 %	83.82 %	73.90 %	79.19 %
	Non Scaling	82.52 %	85.95 %	74.53 %	81.00 %
C3	MinMax	82.86 %	86.37 %	72.82 %	80.68 %
	Robust	81.65 %	85.45 %	74.26 %	80.45 %
	Non Scaling	80.27 %	84.84 %	72.73 %	79.28 %
C4	MinMax	81.99 %	85.26 %	72.28 %	79.84 %
	Robust	79.59 %	84.39 %	72.91 %	78.96 %
	Non Scaling	80.62 %	84.04 %	72.73 %	79.13 %
C5	MinMax	81.91 %	84.29 %	72.64 %	79.82 %
	Robust	80.05 %	83.59 %	72.64 %	78.76 %
	Non Scaling	80.58 %	84.88 %	72.73 %	79.40 %
C6	MinMax	81.42 %	85.34 %	72.64 %	79.80 %
	Robust	80.16 %	83.59 %	72.37 %	78.70 %
	Non Scaling	81.91 %	85.68 %	73.90 %	80.50 %
C7	MinMax	83.13 %	87.24 %	74.71 %	81.69 %
	Robust	81.26 %	85.30 %	73.63 %	80.06 %
	Non Scaling	82.44 %	86.37 %	74.17 %	80.99 %
C8	MinMax	83.63 %	87.40 %	73.90 %	81.64 %
	Robust	81.80 %	85.53 %	73.45 %	80.26 %
	Non Scaling	82.03 %	86.37 %	73.45 %	80.61 %
C8	MinMax	83.09 %	86.94 %	73.72 %	81.25 %
	Robust	81.30 %	85.64 %	73.54 %	80.16 %

Dari hasil di atas, model optimal untuk masing-masing kelas sebagai berikut:

- kelas hate speech (HS) dan abusive (AB) memiliki akurasi sebesar 83.63% dan 87.40%, dengan komposisi MinMax+C7 (menerapkan case folding dan penghapusan stopword) dengan parameter terbaik nilai $c = 10$ dan $\gamma = 1$.
- kelas level memiliki akurasi sebesar 74.71% dan memiliki akurasi rata-rata tertinggi sebesar 81.69% dengan menerapkan komposisi MinMax + C6 (menerapkan case folding dan punctuation removal, tetapi tidak membuang stopword) dengan parameter terbaik nilai $c = 10$ dan $\gamma = 1$.

Berdasarkan hasil tersebut, model C6 yang memiliki akurasi rata-rata tertinggi akan digunakan untuk pengujian akhir, yaitu pengujian dengan komposisi dataset 90% data latih dan 10% data uji. Hasil ini dibandingkan dengan baseline (kombinasi C1 non-scaling), telah menunjukkan peningkatan yang cukup baik, yaitu sekitar 3.58%, 3.32%, dan 1.08% berturut-turut untuk kelas HS, AB dan level.

3.3 Hasil Pengujian dan Pembahasan

Setelah memperoleh model final untuk masing-masing kelas, kemudian dilakukan pengujian final, agar hasilnya bisa dibandingkan dengan hasil akurasi penelitian sejenis dengan dataset dan skema pengujian yang sama. Hasil pengujian model SVM optimal terhadap 10% data test, setelah di training kembali dengan 90% data, dapat dilihat pada tabel berikut:

Tabel 3. Hasil Pengujian

Id Eksperimen	Feature	Kelas		
		HS	AB	Level
C6	Non Scaling	82.34 %	85.31 %	73.16 %
	MinMax	84.52 %	87.02 %	76.40 %
	Robust	80.92 %	84.46 %	72.31 %

Tabel 4. Hasil Perbandingan

Metode	Akurasi (%)		
	HS	AB	Level
Decision Tree[8]	71.25	77.38	63.28
KNN[9]	79.13	83.54	72.18
Random Forest[7]	77.29	80.30	68.55
Logistic Regression[23]	75.86	80.05	70.86
Naive Bayes [6]	82.99	86.48	71.87
BERT[24]	88.88	92.71	85.80
SVM (metode ini)	84.52	87.02	76.43

Hak Cipta Dilindungi Undang-Undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Dari hasil yang diperoleh, model SVM pada eksperimen C6 dengan MinMax mendapatkan akurasi paling tinggi untuk tiga kelas dibanding tanpa scaling dan Robust. Model ini juga memiliki akurasi yang paling tinggi untuk ketiga kelas yang diujikan di antara model machine learning konvensional yang lain (decision tree, random forest, k-NN, Naïve bayes, dan logistic regression), yaitu 84.52%, 87.02%, dan 76.43% untuk kelas HS, AB dan level, namun belum dapat melampaui hasil yang diperoleh dari metode deep learning menggunakan BERT.

Model deep learning masih menunjukkan hasil terbaik untuk ketiga kelas yang diklasifikasikan. Dibandingkan dengan BERT, maka SVM dapat dikatakan adalah metode yang cukup optimal untuk dataset yang kecil, sehingga tidak memerlukan resource komputasi yang tinggi, bila dibandingkan dengan kebutuhan komputasi deep learning.

4. KESIMPULAN

Algoritma SVM mendapatkan hasil terbaik dibanding metode machine learning konvensional lainnya seperti pada tabel 4 yang melakukan penelitian binary classification (seperti kasus klasifikasi HS dan AB), pada penelitian multi-label classification konvensional (kelas level dengan 3 label) juga mendapatkan hasil terbaik dibanding metode machine learning konvensional lainnya. Akurasi tertinggi pada kelas HS, AB dan level dari HS sebesar 84.52%, 87.02% dan 76,43% diperoleh dengan menerapkan MinMax scaling untuk input vektor sentence embeddings FastText. Scaling MinMax untuk seluruh klasifikasi, lebih baik hasilnya dibandingkan menggunakan scaling robust atau tanpa scaling. Hasil yang telah dicapai melalui skema optimasi metode yang telah dilakukan dalam penelitian ini, menunjukkan peningkatan yang signifikan (lebih dari 3%) untuk kelas HS dan AB (binary classification), dan peningkatan 1% untuk kelas level dari HS (multilabel classification). Metode SVM masih belum mampu melampaui hasil yang dicapai oleh metode BERT, disebabkan oleh sumber pelatihan word embeddings Fasttext hanya pada dataset yang kecil (data tweet ini saja), sedangkan BERT dilatih pada korpus teks yang sangat besar yang berasal dari dokumen dan artikel di internet. Saran yang dapat diberikan terhadap penelitian selanjutnya adalah penggunaan word embeddings yang berasal dari data latih yang lebih besar, atau menggunakan pre-trained word embeddings, baik menggunakan FastText, Glove maupun Word2vec, untuk diterapkan pada input tweet SVM untuk meningkatkan akurasi.

REFERENCES

- R. A. Kusuma, "Dampak Perkembangan Teknologi Informasi dan Komunikasi terhadap Perilaku Intoleransi dan Antisosial di Indonesia," MAWA'IZH: JURNAL DAKWAH DAN PENGEMBANGAN SOSIAL KEMANUSIAAN, vol. 10, no. 2, hlm. 273–290, Des 2019, doi: 10.32923/maw.v10i2.932.
- J. M. Teknologi dan S. Assegaff, "Evaluasi Pemanfaatan Media Sosial sebagai Sarana Knowledge Sharing," Jurnal Manajemen Teknologi, vol. 16, no. 3, hlm. 271–293, 2017, doi: 10.12695/jmt.2017.16.3.4.
- R. Bagus Triadi, "Penggunaan Makian Bahasa Indonesia pada Media Sosial (Kajian Sociolinguistik)," Jurnal Sasindo Unpam, vol. 5, no. 2, Des 2017.
- Luh Putu Ary Sri Tjahyanti, "Pendeteksian Bahasa Kasar (Abusive Language) dan Ujaran Kebencian (Hate Speech) dari Komentar di Jejaring Sosial," Daiwi Widya : Jurnal Pendidikan, vol. 07, no. 01, 2020.
- A. Kresna, B. A. Putra, M. A. Fauzi, B. D. Setiawan, dan E. Setiawati, "Identifikasi Ujaran Kebencian Pada Facebook Dengan Metode Ensemble Feature Dan Support Vector Machine," Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 2, no. 12, hlm. 2548–964, 2018, [Daring]. Available: <http://j-ptiik.ub.ac.id>
- T. Ghassani Saskia, "Klasifikasi Hate Speech dan Abusive Language pada Twitter Bahasa Indonesia dengan Metode Naive Bayes Classifier," Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, 2021.
- A. Amri, "Implementasi Algoritma Random Forest untuk Mendeteksi Hate Speech dan Abusive Language pada Twitter Bahasa Indonesia," Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, 2020.
- F. Ihsan, I. Iskandar, N. S. Harahap, dan S. Agustian, "Decision tree algorithm for multi-label hate speech and abusive language detection in Indonesian Twitter," Jurnal Teknologi dan Sistem Komputer, vol. 9, no. 4, hlm. 199–204, Okt 2021, doi: 10.14710/jtsiskom.2021.13907.
- A. Fachlah, "Penerapan Algoritma K-Nearest Neighbor untuk Mendeteksi Ujaran Kebencian dan Bahasa Kasar pada Twitter Bahasa Indonesia," Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, 2021.
- F. Rahutomo, P. Yoga Saputra, dan M. Fidyawan Agtamas, "Implementasi Twitter Sentiment Analysis untuk Review Film Menggunakan Algoritma Support Vector Machine," Jurnal Informatika Polinema, vol. 4, no. 2, hlm. 93–100, 2018.
- P. Arsy dan R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), vol. 8, no. 1, hlm. 147–156, 2021, doi: 10.25126/jtiik.202183944.
- A. Nurudin, B. Anggo, S. Aji, A. Bustamin, dan Z. Abidin, "Perbandingan Kinerja Word Embedding Word2vec, Glove, dan Fasttext Pada Klasifikasi Teks," Jurnal TEKNOKOMPAK, vol. 14, no. 2, hlm. 74, 2020.
- M. Sambuddin dan S. Agustian, "Support Vector Machine Method with Word2vec for Covid-19 Vaccine Sentiment Classification on Twitter," JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING, vol. 6, no. 1, hlm. 288–297, Jul 2022, doi: 10.31289/jite.v6i1.7534.
- M. O. Ibrohim dan I. Budi, "Multi-label Hate Speech and Abusive Language Detection in Indonesian Twitter," Proceedings of the Third Workshop on Abusive Language Online, hlm. 46–57, 2019.



- [15] F. Astuti, R. M. Candra, S. Agustian, dan D. S. Ramadhani, "Klasifikasi Sentimen Masyarakat Terhadap Pemerintah Terkait Penerapan Kebijakan New Normal Menggunakan Metode K-Nearest Neighbor," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 3, 2022.
- [16] H. Nainichah, A. Syukur, dan H. Subagyo, "Pengaruh Text Preprocessing dan Kombinasinya pada Peringkat Dokumen Otomatis Teks Berbahasa Indonesia," *Jurnal Teknologi Informasi*, vol. 15, no. 1, 2019, [Daring]. Available: <http://research.>
- [17] R. Tinjages, A. Triayudi, dan I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 3, hlm. 650, Jul 2020, doi: 10.30865/mib.v4i3.2181.
- [18] R. Achmad Rizal, I. Sanjaya Girsang, dan S. Apriyadi Prasetyo, "Klasifikasi Wajah Menggunakan Support Vector Machine (SVM)," *Riset dan E-Jurnal Manajemen Informatika Komputer*, vol. 3, no. 2, 2019.
- [19] E. H. Harahap, L. Muflikhah, dan B. Rahayudi, "Implementasi Algoritma Support Vector Machine (SVM) Untuk Penentuan Seleksi Atlet Pencak Silat," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 10, hlm. 3843–3848, 2018, [Daring]. Available: <http://j-ptiik.ub.ac.id>
- [20] F. Fauzi, M. Y. Darsyah, dan T. W. Utami, "Smooth Support Vector Machine (Ssvm) untuk Pengklasifikasian Indeks Pembangunan Manusia Kabupaten/Kota Se-Indonesia," *Statistika*, vol. 5, no. 2, 2017.
- [21] Y. E. Ardiningtyas, P. Heruningsih, dan P. Rosa, "Analisis Balancing Data untuk Meningkatkan Akurasi Dalam Klasifikasi," *Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2021*, vol. 2021, no. Prosiding SNAST 2021, hlm. A24–A28, Mar 2021.
- [22] D. Bozsims, J. Dunn, C. Pawlowski, dan Y. D. Zhuo, "Robust Classification," *INFORMS Journal on Optimization*, vol. 1, no. 1, hlm. 2–34, Jan 2019, doi: 10.1287/ijoo.2018.0001.
- [23] A. Fransiska, S. Agustian, dan Fitri Insani, dan M. Fikry, "Algoritme Logistic Regression untuk Mendeteksi Ujaran Kebencian dan Bahasa Kasar Multilabel pada Twitter Berbahasa Indonesia," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 4, 2022.
- [24] R. Saputra, "Implementasi Bidirectional Encoder Representations From Transformers (Bert) untuk Mendeteksi Hatespeech dan Abusive Language pada Twitter Bahasa Indonesia Jurusan Teknik Informatika Uin Suska Riau Tugas Akhir," *Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru*, 2022.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.