

**KLASIFIKASI CYBERBULLYING DI TWITTER
MENGUNAKAN METODE SVM (SUPPORT VECTOR
MACHINE) DAN INFORMATION GAIN**

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh :

Fauzan Aziman Zohro

NIM : 09021382025121

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024**

LEMBAR PENGESAHAN SKRIPSI

KLASIFIKASI CYBERBULLYING DI TWITTER
MENGUNAKAN METODE SVM (SUPPORT VECTOR
MACHINE) DAN INFORMATION GAIN

Oleh :

FAUZAN AZIMAN ZOHRO
NIM : 09021382025121

Palembang, 27 Juni 2024

Pembimbing I


Desty Rodiah, M.T
NIP. 198912212020122011

Pembimbing II


Hadipurnawan Satria, Ph.D.
NIP. 198004182020121001

Mengetahui,
Ketua Jurusan Teknik Informatika



Dr. M. Fachrurrozi, S.Si., M.T.
NIP.-198005222008121002

TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari ini tanggal 31 Mei 2024 telah dilaksanakan ujian komprehensif skripsi oleh jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Fauzan Aziman Zohro

NIM : 09021382025121

Judul : Klasifikasi Cyberbullying Di Twitter Menggunakan Metode Svm (Support Vector Machine) Dan Information Gain

dan dinyatakan LULUS.

1. Ketua Penguji

Rizki Kurniati, M.T

NIP. 199107122019032016

2. Penguji

Novi Yusliani, S.Kom., M.T.

NIP. 198211082012122001

3. Pembimbing I

Desty Rodiah, M.T

NIP. 198912212020122011

4. Pembimbing II

Hadipurnawan Satria, Ph.D.

NIP. 198004182020121001



Mengetahui,
Ketua Jurusan Teknik Informatika

Dr. M. Fachrurrozi, S.Si., M.T.

NIP. 198005222008121002

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Fauzan Aziman Zohro

NIM : 09021382025121

Program Studi : Teknik Informatika

Judul Skripsi : Klasifikasi Cyberbullying Di Twitter Menggunakan
Metode SVM (Support Vector Machine) Dan Information
Gain

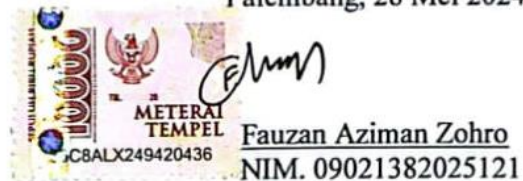
Hasil Pengecekan *Software iThenticate/Turnitin*: 14%

Menyatakan bahwa laporan proyek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan dari pihak mana pun.



Palembang, 28 Mei 2024



Fauzan Aziman Zohro
NIM. 09021382025121

MOTTO DAN PERSEMBAHAN

“Tidak ada jalan pintas menuju keberhasilan, kecuali melalui kerja keras dan ketekunan.”

"Ilmu adalah jendela dunia, dan belajar adalah kunci pembukanya."

"Tumbuh dalam Lingkungan yang Baik, Membangun Pola Pikir yang Kuat."

(Fauzan Aziman Zohro)

Kupersembahkan karya tulis ini kepada :

- **Orang Tua, Saudara/i, dan Keluargaku**
- **Dosen Pembimbing**
- **Teman-teman seperjuangan**
- **Fakultas Ilmu Komputer**
- **Universitas Sriwijaya**

ABSTRACT

Twitter, with 237.8 million daily active users, the majority aged 18-29 years, shows great potential as a communication medium. However, it also brings the problem of uncontrolled online behavior due to bullying comments. This research classifies bullying comments on Indonesian language social media using Support Vector Machine (SVM) and Information Gain. The process includes feature selection with Information Gain, calculating term weights with TF-IDF, and classification analysis with SVM. With a total of 365 bully data and 365 non-bully data. The test results show a significant increase in classification thanks to the use of Information Gain. The accuracy value of SVM without Information Gain is 0.822, while with Information Gain it is 0.830. This shows that selection of Information Gain features in SVM classification provides an increase in classification performance, which is caused by Information Gain's ability to select relevant features and the selection of C values to improve model adjustment.

Keywords: bully, Information Gain, Classification, Support Vector Machine, Twitter

ABSTRAK

Twitter, dengan 237,8 juta pengguna aktif harian, mayoritas usianya 18-29 tahun, menunjukkan potensi besar sebagai media komunikasi. Namun, juga membawa masalah perilaku online yang tidak terkendali karena komentar bully. Penelitian ini mengklasifikasikan komentar *bully* di media sosial berbahasa Indonesia menggunakan *Support Vector Machine* (SVM) dan *Information Gain*. Proses meliputi seleksi fitur dengan *Information Gain*, penghitungan bobot *term* dengan TF-IDF, dan analisis klasifikasi dengan SVM. Dengan jumlah data *bully* sebanyak 365 dan data *non-bully* sebanyak 365. Hasil pengujian menunjukkan kenaikan signifikan pada klasifikasi berkat penggunaan *Information Gain*, Nilai *accuracy* dari SVM tanpa *Information Gain* adalah 0,822, sedangkan dengan *Information Gain* adalah 0,830. Hal ini menunjukkan bahwa seleksi fitur *Information Gain* pada klasifikasi SVM memberikan kenaikan kinerja klasifikasi, yang disebabkan oleh kemampuan *Information Gain* dalam memilih fitur-fitur yang relevan dan pemilihan nilai C dalam meningkatkan penyesuaian model.

Kata Kunci: *bully*, *Information Gain*, Klasifikasi, *Support Vector Machine*, Twitter

KATA PENGANTAR

Segala Puji dan Syukur kepada Allah SWT atas rahmat, hidayah, dan karunia-Nya yang tiada terbatas sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul “Klasifikasi Cyberbullying Di Twitter Menggunakan Metode SVM (Support Vector Machine) Dan Information Gain” Tugas Akhir ini disusun dalam rangka memenuhi salah satu syarat dalam menyelesaikan pendidikan program Strata-1 Program Studi Teknik Informatika Fakultas Ilmu Komputer, Universitas Sriwijaya.

Penulis ingin mengucapkan terima kasih kepada seluruh pihak yang telah memberikan dukungan dan bantuan baik dalam segi materil maupun moril selama proses pembuatan tugas akhir ini. Adapun penulis merujuk secara khusus pihak yang telah membantu sebagai berikut:

1. Bapak Zoharmawan dan Ibu Roaidah, selaku orang tuaku tersayang. serta keluarga besarku yang selalu memberikan dukungan.
2. Bapak Prof. DR. Erwin, S.Si., M.Si., selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. M. Fachrurrozi, S.Si., M. T. selaku ketua jurusan Teknik Informatika.
4. Ibu Desty Rodiah, M.T dan Bapak Hadipurnawan Satria, Ph.D. selaku dosen pembimbing I dan dosen pembimbing II yang telah memberikan arahan, bimbingan, kritik, dan saran kepada penulis sehingga dapat menyelesaikan tugas akhir.
5. Seluruh Bapak dan Ibu Dosen Program Studi Teknik Informatika yang telah memberikan ilmu kepada penulis selama perkuliahan.
6. Staf Jurusan Teknik Informatika beserta Staf Fakultas Ilmu Komputer yang telah membantu dalam urusan administrasi dan akademik selama masa perkuliahan.
7. Anak anak LGasos sahabat terbaik yang selalu memberikan semangat, serta hiburan selama masa perkuliahan.
8. Teman-teman IF Bilingual A angkatan 2020 terutama, Faiq, Bil Hakki, Ganda, Anggi, Nazifah dan Rio yang selalu membagikan informasi, menciptakan momen kebersamaan dan selalu menemani selama perkuliahan yang tentu saja tidak akan pernah dilupakan.
9. Semua pihak lainnya yang tidak dapat saya sebutkan satu persatu yang telah berperan bagi penulis dalam menyelesaikan tugas akhir, penulis ucapkan terima kasih banyak atas semuanya.

Semoga tugas akhir ini dapat bermanfaat bagi penulis khususnya maupun pembaca pada umumnya, serta dapat menjadi referensi dan

rujukan bagi hal-hal yang bermanfaat. Penulis menyadari bahwa dalam proses penyelesaian tugas akhir ini, terdapat beberapa penjelasan yang kurang sempurna. Oleh karena itu, penulis berharap akan kritik dan saran yang membangun agar tugas akhir ini menjadi karya tulis yang sempurna agar terciptanya bekal pengetahuan yang baik bagi peneliti di masa depan.

Palembang, 7 Juli 2024

A handwritten signature in black ink, consisting of a large, stylized letter 'F' followed by several loops and a final vertical stroke.

Fauzan Aziman Zohro

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI.....	iii
HALAMAN PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xiv
BAB I PENDAHULUAN	
1.1. Pendahuluan	I-1
1.2. Latar Belakang	I-1
1.3. Rumusan Masalah	I-3
1.4. Tujuan Penelitian.....	I-3
1.5. Manfaat Penelitian.....	I-4
1.6. Batasan Masalah.....	I-4
1.7. Sistematika Penulisan.....	I-4
1.8. Kesimpulan	I-5
BAB II KAJIAN LITERATUR	
2.1. Pendahuluan	II-1
2.2. Landasan Teori.....	II-1
2.2.1. Twiter	II-1
2.2.2. <i>Bullying</i>	II-3
2.2.3. <i>Text Pre processing</i>	II-4
2.2.4. <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	II-6
2.2.5. <i>Support Vector Machine</i>	II-7
2.2.6. <i>Information Gain Feature Selection</i>	II-13
2.2.7. <i>Confusion matrix</i>	II-14
2.2.8. <i>Rational Unified Process</i>	II-15
2.2.9. Penelitian Lain yang Relevan	II-18
2.2.10. Kesimpulan	II-19

BAB III METODOLOGI PENELITIAN

3.1	Pendahuluan	III-1
3.2	Pengumpulan Data	III-1
3.2.1	Jenis dan Sumber Data	III-1
3.3	Tahap Penelitian	III-2
3.3.1.	Pembagian Data Uji Dan Data Latih	III-3
3.3.2.	Menentukan Kerangka Kerja Penelitian	III-3
3.3.3.	Menentukan Alat Bantu Penelitian	III-6
3.3.4.	Menentukan kriteria Penelitian	III-7
3.3.5.	Menentukan Format Data Penelitian	III-7
3.3.6.	Melakukan Uji Penelitain	III-8
3.3.7.	Melakukan Analisis Hasil Pengujian dan Kesimpulan Penelitian.....	III-9
3.4	Metode Pengembangan Perangkat Lunak.....	III-10
3.4.1.	<i>Inception</i>	III-10
3.4.2.	<i>Elaboration</i>	III-10
3.4.3.	<i>Construction</i>	III-10
3.4.4.	<i>Transition</i>	III-11
3.4.5.	Manajemen Proyek Penelitian	III-11
3.5	Kesimpulan	III-14

BAB IV PENGEMBANGAN PERANGKAT LUNAK

4.1.	Pendahuluan	IV-1
4.2.	Fase insepisi	IV-1
4.2.1.	Pemodelan Bisnis	IV-1
4.2.2.	Kebutuhan Sistem.....	IV-2
4.2.3.	Analisis dan Desain.....	IV-3
4.2.4.	Implementasi	IV-21
4.3.	Fase Elaborasi	IV-26
4.3.1.	Pemodelan Bisnis	IV-26
4.3.2.	Kebutuhan Sistem.....	IV-28
4.3.3.	<i>Activity Diagram</i>	IV-28
4.3.4.	<i>Sequance Diagram</i>	IV-30
4.4.	Fase Kontruksi	IV-32
4.4.1.	<i>Class Diagram</i>	IV-32
4.4.2.	Implementasi	IV-33
4.5.	Fase Transisi.....	IV-35
4.5.1.	Pemodelan Bisnis	IV-35

4.5.2.	Rencana Pengujian	IV-35
4.5.3.	Implementasi	IV-36
4.6.	Kesimpulan	IV-37
BAB V HASIL DAN ANALISIS PENELITIAN		
5.1.	Pendahuluan	V-1
5.2.	Data Hasil Penelitian	V-1
5.2.1.	Konfigurasi Percobaan	V-1
5.2.2.	Hasil Pengujian.....	V-2
5.3.	Analisis Hasil Pengujian Secara Keseluruhan	V-4
5.4.	Kesimpulan	V-9
BAB VI KESIMPULAN DAN SARAN		
6.1.	Pendahuluan	VI-1
6.2.	Kesimpulan	VI-1
6.3.	Saran	VI-2
DAFTAR PUSTAKA		viii
DAFTAR LAMPIRAN		x

DAFTAR GAMBAR

	Halaman
Gambar II- 1. Contoh <i>Hyperplane</i> Pada SVM (Nugroho et al., 2003)	II-7
Gambar II- 2. Pemetaan data ke ruang dimensi lebih tinggi	II-12
Gambar II- 3. Arsitektur RUP	II-16
Gambar III- 1 .Tahap Penelitian.....	III-3
Gambar III- 2. Kerangka Kerja Penelitian	III-4
Gambar IV- 1 Diagram <i>Use Case</i>	IV-21
Gambar IV- 2. Rancangan Input <i>File</i>	IV- 27
Gambar IV- 3. Rancangan Halaman Input dan SVM	IV- 27
Gambar IV- 4. Rancangan Halaman Input SVM dan <i>Information Gain</i>	IV- 28
Gambar IV- 5. <i>Activity Diagram</i> Input <i>File</i>	IV- 29
Gambar IV- 6. <i>Activity Diagram</i> Tabel Hasil Klasifikasi SVM.....	IV- 29
Gambar IV- 7. <i>Activity Diagram</i> Tabel Hasil Klasifikasi SVM dan <i>Information Gain</i>	IV-30
Gambar IV- 8. <i>Sequance</i> diagram input file	IV-31
Gambar IV- 9. <i>Sequance</i> diagram hasil SVM.....	IV-31
Gambar IV- 10. <i>Sequence</i> diagram hasil SVM +IG	IV-32
Gambar IV- 11. <i>Class diagram</i>	IV-33
Gambar IV- 12. <i>Interface Upload File</i>	IV-34
Gambar IV- 13. <i>Interface</i> Hasil SVM.....	IV-35
Gambar IV- 14. <i>Interface</i> Hasil SVM + <i>Information Gain</i>	IV-35
Gambar V- 1. Perbandingan <i>Accuracy</i>	V-4
Gambar V- 2. Perbandingan Rata Rata <i>Accuracy</i>	V-5
Gambar V- 3.Perbandingan <i>Precision</i>	V-6
Gambar V- 4. Rata Rata <i>Precision</i>	V-6
Gambar V- 5. Perbandingan <i>Recall</i>	V-7
Gambar V- 6.Rata Rata <i>Recall</i>	V-7
Gambar V- 7. Perbandingan <i>F-Measure</i>	V-8
Gambar V- 8. Rata Rata <i>F-Measure</i>	V-8

DAFTAR TABEL

	Halaman
Tabel II- 1 <i>Confusion Matrix</i> Dua Kelas.....	I-14
Tabel III- 1. Contoh Data Komentar	III-1
Tabel III- 2. Rancangan Tabel <i>Confusion Matrix</i> Hasil Klasifikasi	III-7
Tabel III- 3. Rancangan Tabel Hasil Pengujian	III-8
Tabel III- 4. Rancangan Tabel Hasil Analisis Klasifikasi.....	III-9
Tabel III- 5. Tabel Manajemen Proyek Perangkat Lunak.....	III-11
Tabel IV- 1.Kebutuhan Fungsional.....	IV-3
Tabel IV- 2. Kebutuhan Non-Fungsional.....	IV-3
Tabel IV- 3. Sampel Data.....	IV-4
Tabel IV- 4. Hasil <i>Casefolding</i>	IV-5
Tabel IV- 5.Hasil <i>Cleaning</i>	IV-5
Tabel IV- 6. Hasil Normalisasi	IV-6
Tabel IV- 7. Hasil <i>Stopword Removal</i>	IV-6
Tabel IV- 8. Hasil <i>Stemming</i>	IV-7
Tabel IV- 9. Hasil <i>Tokenizing</i>	IV-7
Tabel IV- 10. Perhitungan TF-IDF dari data <i>Pre-Processing</i>	IV-8
Tabel IV- 11.Perhitungan nilai TF-IDF dari Data Hasil <i>Pre-Processing</i>	IV-9
Tabel IV- 12 Perhitungan Nilai <i>Information Gain</i>	IV-13
Tabel IV- 13. Pengurutan bobot data	IV-14
Tabel IV- 14. Sampel Data.....	IV-15
Tabel IV- 15.Perhitungan TF-IDF	IV-16
Tabel IV- 16. Perhitungan nilai TF-IDF dari Data Hasil <i>Pre-Processing</i>	IV-17
Tabel IV- 17.Tabel Definisi aktor.....	IV-22
Tabel IV- 18. Tabel Definisi <i>Use Case</i>	IV-22
Tabel IV- 19. Tabel Skenario <i>Use Case</i>	IV-23
Tabel IV- 20. Perhitungan Dan Menampilkan Analisis Model SVM Tanpa <i>Information Gain</i>	IV-23
Tabel IV- 21. Perhitungan Dan Menampilkan Analisis Model SVM dengan <i>Information Gain</i>	IV-25
Tabel IV- 22.Deskripsi <i>file</i> kelas	IV-33
Tabel IV- 23.Daftar file program <i>interface</i>	IV-34
Tabel IV- 24. Rencana Pengujian <i>Use Case</i> Input File	IV-36
Tabel IV- 25. Rencana Pengujian <i>Use Case</i> Hasil Klasifikasi SVM Tanpa <i>Information Gain</i>	IV-36
Tabel IV- 26. Rencana Pengujian <i>Use Case</i> Hasil Klasifikasi SVM dengan <i>Information Gain</i>	IV-36
Tabel IV- 27. Pengujian <i>Use Case</i> Input Algoritma	IV-36
Tabel IV- 28. Pengujian <i>Use Case</i> Analisis Hasil Klasifikasi SVM	IV-37
Tabel IV- 29. Pengujian <i>Use Case</i> Analisis Hasil Klasifikasi SVM dengan <i>Information Gain</i>	IV-37
Tabel V- 1. <i>Confusion Matrix</i> SVM Tanpa <i>Information Gain</i>	V-2
Tabel V- 2. Hasil Evaluasi Klasifikasi SVM tanpa <i>Information Gain</i>	V-2
Tabel V- 3 <i>Confusion Matrix</i> SVM dengan <i>Information Gain</i>	V-3
Tabel V- 4.Hasil Evaluasi Klasifikasi dengan <i>Information Gain</i>	V-3

BAB I

PENDAHULUAN

1.1. Pendahuluan

Bab ini mendeskripsikan latar belakang, rumusan masalah, tujuan, manfaat penelitian, keterbatasan masalah, sistematika penulisan skripsi, dan kesimpulan penelitian yang merupakan gagasan utama yang menjadi dasar penelitian.

1.2. Latar Belakang

Data dari Sproutsocial menunjukkan bahwa Twitter memiliki 237,8 juta pengguna aktif setiap hari, dengan mayoritas berusia 18 hingga 29 tahun.¹ Jumlah pengguna yang lebih muda menunjukkan potensi besar *platform* ini sebagai tempat untuk berkomunikasi dan berbagi informasi. Namun demikian, hal ini juga menimbulkan masalah terkait perilaku online yang tidak terkendali.

Platform media sosial seperti Twitter memungkinkan orang di seluruh dunia berkomunikasi dan mengakses informasi, tetapi *cyberbullying* juga semakin meningkat. Kekhawatiran serius tentang dampak pelecehan online terhadap kesehatan mental dan keamanan individu muncul. Salah satu pemicu utama adalah kurangnya pemahaman tentang etika berkomunikasi di platform ini, yang menciptakan lingkungan yang mendukung perilaku berbahaya dan merugikan.

Seiring dengan kemajuan teknologi dan semakin meningkatnya interaksi *online*, sangat penting untuk melindungi pengguna *Twitter*. Akibatnya, penelitian klasifikasi *cyberbullying* di *Twitter* semakin penting. Peneliti dapat

¹ <https://sproutsocial.com/insights/new-social-media-demographics/> diakses pada 13 oktober 2023

mengembangkan solusi yang lebih baik untuk menjaga keamanan dan kesejahteraan pengguna serta mencegah efek negatif dari *cyberbullying* dengan memperoleh pemahaman yang lebih baik tentang masalah ini

Korban *cyberbullying* sering mengalami masalah kesehatan mental yang parah, seperti depresi atau bunuh diri, sebagai akibat dari penghinaan dan celaan yang mereka terima. Pelaku *cyberbullying* ini melakukan tindakan ini karena *Twitter* atau yang Sekarang dikenal sebagai "X" memungkinkan pengguna lain untuk melakukan ucapan buruk atau komentar buruk di konten seseorang. Untuk mengurangi *cyberbullying*, komentar di *Twitter* harus diklasifikasikan untuk menentukan apakah *bullying* atau *non bullying*. ada beberapa Metode yang dapat digunakan untuk melakukan klasifikasi yaitu *Naïve Bayes* (NB) dan *Support Vector Machine* (SVM) dan masih banyak yang lainnya.

Perbandingan metode klasifikasi juga pernah diteliti oleh (Chamidah & Sahawaly, 2021) yang mengklasifikasi perilaku bully pada aplikasi *Twitter* dengan judul penelitian *Comparison Support Vector Machine dan Naïve Bayes Methods for Classifying Cyberbullying in Twitter*, dengan hasil penelitian yaitu algoritma *Support Vector Machine* menghasilkan akurasi 99,6%, sedangkan *Naïve Bayes Classifier* menghasilkan akurasi 97,99%.

Support Vector Machine (SVM) merupakan metode terbaik di antara beberapa metode lainnya karena mampu mengkomputasi data dengan dimensi tinggi sehingga tingkat akurasi yang dihasilkan lebih baik (Siringoringo dan Jamaluddin, 2019).

Menurut Deng & Runger (dalam Purnamasari, et al., 2018), *Information Gain*(IG) adalah teknik pemilihan fitur untuk memperkirakan signifikansi maksimum suatu fitur. Nilai *Information Gain* suatu fitur dapat merepresentasikan seberapa penting fitur tersebut untuk mewakili suatu kelas. Perhitungan *Information Gain* dilakukan pada semua fitur data latih pada tiap kelas dan fitur yang nilainya kurang dari *threshold* atau batas akan dihapus (Pradhana, 2020).

Nilai *Information Gain* didapat melalui nilai *entropy* sebelum pemisahan dikurangi dengan nilai *entropy* setelah proses pemisahan. Nilai yang didapat akan digunakan untuk menentukan atribut mana saja yang akan digunakan atau tidak digunakan. Atribut terpilih yaitu atribut yang memenuhi kriteria pembobotan akan digunakan pada proses klasifikasi (Purnamasari et al., 2018).

Oleh karena itu Metode SVM dan *Information Gain* akan digunakan dalam penelitian klasifikasi komentar *bully* pada twitter berdasarkan hasil referensi penelitian sebelumnya.

1.3. Rumusan Masalah

Sesuai dengan latar belakang, maka rumusan masalah dari penelitian berikut :

1. Bagaimana pengaruh penggunaan *Information Gain* pada klasifikasi komentar *bully* dengan *Support Vector Machine*?
2. Bagaimana kinerja metode *support vector machine* dan *Information gain* pada Klasifikasi Komentar Bully, jika dilihat dari *accuracy*, *precison*, *recall*, dan *F-measure*?

1.4. Tujuan Penelitian

1. Untuk mengetahui pengaruh penggunaan *Information gain* pada klasifikasi

komentar *bully* dengan *Support Vector Machine*

2. Mengetahui kinerja metode *support vector machine* dan *Information gain* pada Klasifikasi Komentar Bully dengan *Support Vector Machine* jika dilihat dari *accuracy, precision, recall, dan F-measure*

1.5. Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah :

1. Hasil penelitian dapat dijadikan sebagai referensi penelitian terkait
2. Sebagai masukan bagi pengembang aplikasi *Twitter* untuk meminimalisir munculnya komentar *bully* di kolom komentar *Twitter*

1.6. Batasan Masalah

Berikut batasan penelitian :

1. Data komentar yang digunakan adalah 650 komentar berbahasa Indonesia.
2. Data yang digunakan adalah data komentar *Twitter* yang di dapat dari kaggle
3. Klasifikasi dibagi menjadi dua kategori: *bully* dan tidak-*bully*.

1.7. Sistematika Penulisan

Sistematika penyusunan skripsi adalah sebagai berikut :

Bab I. Pendahuluan

Bab ini membahas kerangka dasar penelitian, meliputi latar belakang, rumusan masalah, tujuan, manfaat penelitian, batasan masalah, dan sistematika penulisan.

Bab II. Kajian Literatur

Bab ini menjelaskan teori-teori dasar yang digunakan dalam penelitian ini, seperti definisi *Twitter*, *bullying*, klasifikasi teks, *Support Vector Machine*,

Information Gain dan beberapa tinjauan literatur dari penelitian terkait.

Bab III. Metode Penelitian

Bab ini menjelaskan tahapan penelitian, seperti pengumpulan data, analisis data, mekanisme pengujian, dan perancangan sistem.

Bab IV. Pengembangan Perangkat Lunak

Bab ini dibahas proses pengembangan perangkat lunak, dari tahap pengumpulan data hingga hasil akhir program.

Bab V. Hasil Dan Analisis Penelitian

Dalam bab ini dijelaskan terkait hasil yang didapat dalam penelitian dan analisis hasil penelitian.

Bab VI. Kesimpulan Dan Saran

Bab ini memberikan kesimpulan hasil penelitian dan saran terkait penelitian terkait.

1.8. Kesimpulan

Dalam bab ini telah membahas latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan yang akan digunakan dalam proses penelitian.

DAFTAR PUSTAKA

- Agarwal, S. D., JOHNSON, C. N., & WALKER, S. (2014). *A Model of Crowd-Enabled Organization: Theory and Methods for Understanding the Role of Twitter in the Occupy Protests*.
- Amalia, A., Zaidiah, A., & Isnainiyah, I. N. (2022). Prediksi Kualitas Udara Menggunakan Algoritma K-Nearest Neighbor. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 7(2), 496–507.
- Bara, E. A. B., Nasution, K. A., & Ginting, R. Z. (n.d.). *Penelitian tentang Twitter*.
- Chamidah, N., & Sahawaly, R. (2021). Comparison Support Vector Machine and Naive Bayes Methods for Classifying Cyberbullying in Twitter. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, 7(2), 338.
- Darma, I. M. B. S., Perdana, R. S., & Indriati. (2018). *Penerapan Sentimen Analisis Acara Televisi Pada Twitter Menggunakan Support Vector Machine dan Algoritma Genetika*.
- Hadna, N. M. S., Santosa, P. I., & Winarno, W. W. (2016). *STUDI LITERATUR TENTANG PERBANDINGAN METODE UNTUK PROSES ANALISIS SENTIMEN DI TWITTER*.
- Hidajat, M., Adam, A. R., Danaparamita, M., & Suhendrik, S. (2015). Dampak Media Sosial dalam Cyber Bullying. *ComTech: Computer, Mathematics and Engineering Applications*, 6(1), 72.
- Hidayat, A. R., Adikara, P. P., & Adinugroho, S. (2020). *Klasifikasi Hoaks Kesehatan di Media Sosial menggunakan Support Vector Machine*.

- Lei, S. (2012). A Feature Selection Method Based on Information Gain and Genetic Algorithm. *2012 International Conference on Computer Science and Electronics Engineering*, 355–358.
- Luqyana, W. A., Cholissodin, I., & Perdana, R. S. (2018). *Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine*.
- Navin, M., & Pankaja. (2016). Performance Analysis of Text Classification Algorithms using Confusion Matrix. *International Journal of Engineering and Technical Research*, 6(4).
- Noviantho, Ashianti, L., & Isa, S. M. (2017). *Cyberbullying Classification using Text Mining*.
- Nugroho, A. S., Witarto, A. B., & Handoko, D. (2003). *–Teori dan Aplikasinya dalam Bioinformatika1–*.
- Pradhana, M. O. (2020). *Analisis Sentimen Masyarakat terhadap Uji Coba LRT Jakarta Menggunakan Improved K-Nearest Neighbor dan Information Gain*.
- Purnamasari, N. M. G. D., Fauzi, I., & Dewi, L. S. (2018). Identifikasi Tweet Cyberbullying pada Aplikasi Twitter menggunakan Metode Support Vector Machine (SVM) dan Information Gain (IG) sebagai Seleksi Fitur. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 2(11).
- Riswanto, D., & Marsinun, R. (2020). Perilaku Cyberbullying Remaja di Media Sosial. *Analitika*, 12(2), 98–111.