

Tesis

Perbandingan Algoritma *Naïve Bayes Classifier* dan *Support Vector Machine* Terhadap Analisis Sentimen *Hate Speech* di Twitter

**Murni
2007048014**



**Program Studi S2 Informatika
Fakultas Teknologi Industri
Universitas Ahmad Dahlan
Yogyakarta
2023**

Tesis

Perbandingan Algoritma *Naïve Bayes Classifier* dan *Support Vector Machine* Terhadap Analisis Sentimen *Hate Speech* di Twitter

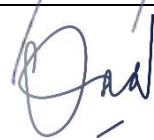
Murni
2007048014

Dipertahankan di depan Dewan Penguji
Tanggal 27 November 2023

Prof. Dr. Ir. Imam Riadi, M.Kom.
Ketua Penguji



Prof. Drs. Ir. Abdul Fadlil, M.T., Ph.D.
Penguji 1



Prof. Ir. Sunardi, S.T., M.T., Ph.D.
Penguji 2



Dr.Eng. Muhammad Kunta Biddinika
Penguji 3



Mengetahui,



Prof. Ir. Sunardi, S.T., M.T., Ph.D.
Dekan Fakultas Teknologi Industri

Pernyataan Tidak Plagiat

Saya yang bertanda tangan di bawah ini:

Nama : Murni
NIM : 2007048014
Email : murni2007048014@webmail.uad.ac.id
Program Studi : S2 Informatika
Fakultas : Teknologi Industri
Judul Tesis : Perbandingan Algoritma *Naïve Bayes Classifier* dan *Support Vector Machine* Terhadap Analisis Sentimen *Hate Speech* di Twitter

Dengan ini menyatakan bahwa:

1. Hasil karya yang saya serahkan ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar kesarjanaan baik di Universitas Ahmad Dahlan maupun di institusi pendidikan lainnya.
2. Hasil karya saya ini bukan saduran/terjemahan melainkan merupakan gagasan, rumusan, dan hasil pelaksanaan penelitian/implementasi saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing akademik dan narasumber penelitian.
3. Hasil karya saya ini merupakan hasil revisi terakhir setelah diujikan yang telah diketahui dan disetujui oleh pembimbing.
4. Dalam karya saya ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali yang digunakan sebagai acuan dalam naskah dengan menyebutkan nama pengarang dan dicantumkan dalam daftar pustaka.

Pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari terbukti ada penyimpangan dan ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya saya ini, serta sanksi lain yang sesuai dengan ketentuan yang berlaku di Universitas Ahmad Dahlan.

Yogyakarta, 27 November 2023
Yang Menyatakan



Murni

Pernyataan Persetujuan Akses

Saya yang bertanda tangan di bawah ini:

Nama : Murni
NIM : 2007048014
Email : murni2007048014@webmail.uad.ac.id
Program Studi : S2 Informatika
Fakultas : Teknologi Industri
Judul Tesis : Perbandingan Algoritma *Naïve Bayes Classifier* dan
Support Vector Machine Terhadap Analisis Sentimen
Hate Speech di Twitter

Dengan ini saya menyerahkan hak Sepenuhnya kepada Pusat Sumber Belajar Universitas Ahmad Dahlan untuk menyimpan, mengatur akses serta melakukan pengelolaan terhadap karya saya ini dengan mengacu pada ketentuan akses tesis elektronik sebagai berikut (beri tanda pada kotak):



Saya mengizinkan karya tersebut diunggah ke dalam aplikasi Repository Pusat Sumber Belajar Universitas Ahmad Dahlan.

Demikian pernyataan ini saya buat dengan sebenarnya.

Yogyakarta, 27 November 2023

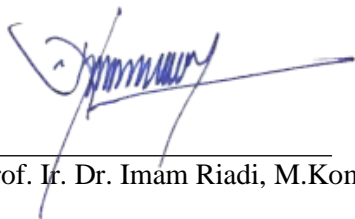
Yang Menyatakan



Murni

Mengetahui,

Pembimbing Tesis 1



Prof. Ir. Dr. Imam Riadi, M.Kom.

Pembimbing Tesis 2



Prof. Drs. Ir. Abdul Fadlil, M.T., Ph.D.

Motto dan Persembahan

"Trust in Allah plan; nothing disappoints when you have faith in Allah." -Murni

"Maka sesungguhnya bersama kesulitan ada kemudahan" (Q.S. Al Insyirah: 5)

Tesis ini kupersembahkan untuk:

1. Orang tua tersayang, Ibu Amnia dan Bapak Samahudin
Terimakasih atas kepercayaan dan kasih sayang yang diberikan selama ini.
2. Universitas Muhammadiyah Sorong yang telah memberikan kesempatan untuk dapat menempuh pendidikan di Universitas Ahmad Dahlan (UAD).
3. Dosen pembimbing akademik, Dosen pembimbing 1 dan 2, Dosen Penguji, dan Seluruh Dosen MTI UAD.
4. Keluarga, Sahabat, MTI10, MTI11, MTI12, MTI13, MTI14 yang selalu menjadi penyemangat dan penghibur selama proses penulisan Tesis.
5. Para pembaca semua.

Kata Pengantar

Bismillahirrahmaanirrahiim,

Assalaamu'alaikum warahmatullaahi wabarakaatuh

Segala puji bagi Allah SWT yang telah memberikan rahmat, hidayah, dan inayah kepada hamba-Nya sehingga penulis dapat menyelesaikan Tesis ini.

Penulis menyadari bahwa keberhasilan dalam menyelesaikan laporan ini berkat dorongan dan bimbingan berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis menghaturkan terimakasih dan penghargaan kepada Bapak Prof. Ir. Sunardi, S.T., M.T., Ph.D. Dekan Fakultas Teknologi Industri, Bapak Prof. Drs. Ir. Abdul Fadlil, M.T., Ph.D. selaku Kaprodi S2 Informatika, Bapak Prof. Dr. Ir. Imam Riadi, M.Kom. selaku Pembimbing Tesis 1 dan Bapak Prof. Drs. Ir. Abdul Fadlil, M.T., Ph.D. selaku Pembimbing Tesis 2 yang rela diganggu setiap saat selama proses penulisan tesis. Terimakasih juga penulis haturkan kepada seluruh dosen S2 Informatika yang telah memberikan banyak ilmu dan wawasan.

Selanjutnya penulis juga menyampaikan terima kasih kepada rekan-rekan angkatan 10 di S2IF UAD dan berbagai pihak yang tidak dapat disebutkan satu persatu.

Terakhir, penulis sangat mengharapkan saran dan kritik yang membangun, karena Tesis ini sangat jauh dari sempurna. Semoga Allah meridhoi langkah kita. Amin.

Wassalaamu'alaikum warahmatullaahi wabarakaatuh

Yogyakarta, 27 November 2023

Penulis

Daftar Isi

Pernyataan Tidak Plagiat.....	iii
Pernyataan Persetujuan Akses.....	iv
Motto dan Persembahan	v
Kata Pengantar	vi
Daftar Isi	vii
Daftar Gambar	ix
Daftar Tabel.....	x
Abstrak.....	xi
Abstract.....	xii
Bab 1 Pendahuluan	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah.....	2
1.3 Batasan Masalah	3
1.4 Rumusan Masalah.....	3
1.5 Tujuan Penelitian	4
1.6 Manfaat Penelitian	4
Bab 2 Tinjauan Pustaka	5
2.1 Kajian Penelitian Terdahulu	5
2.2 Landasan Teori	20
2.2.1 Twitter	20
2.2.2 <i>Preprocessing</i>	21
2.2.3 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	22
2.2.4 <i>Naïve Bayes Classifier</i>	23
2.2.5 <i>Support Vector Machine</i>	24
2.2.6 <i>Stratified K-Fold</i>	26
Bab 3 Metodologi	27

3.1 Objek Penelitian.....	27
3.2 Alat dan Bahan.....	27
3.3 Perancangan Sistem	28
3.3.1 Pengambilan Data	28
3.3.2 Analisis Sentimen	29
3.3.3 <i>Preprocessing</i>	30
3.3.4 TF-IDF	30
3.3.5 Algoritma Naïve Bayes Classifier.....	31
3.3.6 Algoritma Support Vector Machine.....	32
3.3.7 <i>Stratified K-Fold</i>	33
3.4 Pengujian Sistem.....	34
Bab 4 Hasil dan Pembahasan.....	36
4.1 Pengambilan Data	36
4.2 Analisis Sentimen	40
4.3 <i>Preprocessing</i>	41
4.3.1 <i>Case folding</i>	41
4.3.2 <i>Tokenization</i>	42
4.3.3 <i>Stopword removal</i>	43
4.3.4 <i>Stemming</i>	44
4.4 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	44
4.5 Analisis <i>Naïve Bayes Classifier</i>	49
4.6 Analisis <i>Support Vector Machine</i>	51
4.7 Perbandingan Model NBC dan SVM	55
Bab 5 Penutup	57
5.1 Kesimpulan	57
5.2 Saran	57
Daftar Acuan	58
Daftar Lampiran	61
Lampiran 1 Daftar Riwayat Hidup	62
Lampiran 2 <i>Source Code</i>	63

Daftar Gambar

Gambar 2.1 Perbandingan Penelitian Terdahulu	19
Gambar 2.2 <i>Support Vector Machine</i>	24
Gambar 3.1 Perancangan Sistem	28
Gambar 3.2 Pengambilan Data	29
Gambar 3.3 Analisis Sentimen.....	29
Gambar 3.4 <i>Preprocessing</i>	30
Gambar 3.5 TF-IDF	31
Gambar 3.6 Algoritma <i>Naïve Bayes Classifier</i>	32
Gambar 3.7 Algoritma <i>Support Vector Machine</i>	33
Gambar 3.8 Ilustrasi <i>Stratified K-Fold</i>	34
Gambar 3.9 <i>Confusion Matrix</i>	34
Gambar 4.1 <i>Auth token</i> Pengguna Twitter	36
Gambar 4.2 <i>Word Cloud Query: Hate Speech</i> Agama	38
Gambar 4.3 <i>Word Cloud Query: Hate Speech</i> Antargolongan.....	38
Gambar 4.4 <i>Word Cloud Query: Hate Speech</i> Suku atau Ras.....	39
Gambar 4.5 <i>Dataset</i>	40
Gambar 4.6 Performa <i>Classification Report</i> MNB.....	50
Gambar 4.7 Performa <i>accuracy</i> MNB	50
Gambar 4.8 <i>Confusion Matrix</i> MNB	51
Gambar 4.9 Performa <i>Classification Report</i> <i>LinearSVC</i> $C = 0,1$	53
Gambar 4.10 Performa <i>Classification Report</i> <i>LinearSVC</i> $C = 1$	54
Gambar 4.11 Performa <i>Classification Report</i> <i>LinearSVC</i> $C = 10$	54
Gambar 4.12 <i>Confusion matrix</i> <i>LinearSVC</i>	55

Daftar Tabel

Tabel 2.1 Rangkuman penelitian terdahulu	12
Tabel 3.1 Alat Penelitian.....	27
Tabel 4.1 <i>Query Hate Speech</i>	37
Tabel 4.2 Atribut Data Twitter.....	39
Tabel 4.3 Hasil Pelabelan Sentimen	40
Tabel 4.4 Tahap <i>Case folding</i>	41
Tabel 4.5 Tahap <i>Tokenization</i>	42
Tabel 4.6 Tahap <i>Stopword removal</i>	43
Tabel 4.7 Tahap <i>Stemming</i>	44
Tabel 4.8 Teks <i>Stemming</i>	45
Tabel 4.9 Perhitungan TF-IDF.....	46
Tabel 4.10 Hasil <i>Classification Naïve Bayes</i>	49
Tabel 4.11 Hasil <i>Classification</i> Naïve Bayes.....	51
Tabel 4.12 Hasil <i>Classification SVM</i>	52
Tabel 4.13 Hasil <i>Classification LinearSVC</i>	55
Tabel 4.14 Perbandingan Kinerja Terbaik <i>LinearSVC</i> dan MNB.....	56

Murni (2023). **Perbandingan Algoritma *Naïve Bayes Classifier* dan *Support Vector Machine* Terhadap Analisis Sentimen *Hate Speech* di Twitter**. Tesis, Magister Informatika, Universitas Ahmad Dahlan, Yogyakarta.

Abstrak

Twitter adalah media sosial yang menghargai kebebasan untuk berekspresi. Karakteristik Twitter yang memberikan kemudahan dalam berbagi Tweet dengan cepat dapat menyebabkan penyebaran *Hate Speech*. *Hate Speech* merujuk pada ekspresi verbal yang memiliki tujuan merendahkan individu atau kelompok tertentu berdasarkan atribut seperti suku, agama, ras, dan antargolongan (SARA). Oleh karena itu, identifikasi dan penanganan *Hate Speech* menjadi perhatian dalam upaya menjaga keamanan dan kenyamanan pengguna media sosial. Penggunaan Algoritma *Machine Learning* dalam menganalisis sentimen dibutuhkan untuk mendeteksi apakah suatu Tweet termasuk dalam kategori *Hate Speech* atau *Non-Hate Speech* di Twitter. Tujuan penelitian ini adalah melakukan analisis sentimen terhadap *Tweet* pada media sosial Twitter menggunakan algoritma *Machine Learning* guna mengidentifikasi algoritma yang tepat dalam mengenali Tweet yang termasuk dalam kategori *Hate Speech*.

Analisis sentimen merupakan bagian dari *natural language processing* untuk mendeteksi nada emosional dalam teks. Penelitian ini menggunakan algoritma *Machine Learning Naïve Bayes Classifier* dan *Support Vector Machine* untuk identifikasi *Tweet Hate Speech*. Tahapan penelitian yang dilakukan yaitu pengambilan data dari Twitter menggunakan *Twitterscraper* dan API *Twitter*, *labeling* manual, *preprocessing* (*case folding*, *tokenization*, *stopword removal*, dan *stemming*), *feature extraction* menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF), *stratified k-fold*, dan evaluasi menggunakan *Confusion Matrix*. Penelitian ini menggunakan bahasa pemrograman Python untuk proses analisis dan evaluasi.

Dataset yang dikumpulkan sebanyak 5000 data dengan sentimen *Hate Speech* sebanyak 3988 dan *Non-Hate Speech* sebanyak 1012 sentimen. Data tersebut melalui tahap *preprocessing*, yaitu *case folding*, *tokenization*, *stopword removal*, dan *stemming*. Data yang telah melalui tahap *stemming* diubah menggunakan *feature extraction* TF-IDF. Sentimen pada data yang dianalisis tidak seimbang sehingga digunakan teknik *stratified k-fold*. Model yang dianalisis yaitu *Multinomial Naïve Bayes* (MNB) dan *Linear Support Vector Machine* (*LinearSVC*). Hasil penelitian menunjukkan *LinearSVC* unggul dalam mengidentifikasi sentimen *Hate Speech* dengan *precision* sebesar 85%, sementara MNB lebih baik dalam mengenali sentimen *Hate Speech* dengan *recall* sebesar 100%. Meskipun kedua model memiliki *F1-Score* yang sama, *LinearSVC* memiliki *accuracy* yang lebih tinggi yaitu 85%, sementara MNB memiliki *accuracy* 82%, sehingga membuktikan keunggulan SVM dibandingkan MNB.

Kata kunci: *Machine Learning*, Analisis Sentimen, *Hate Speech*, Twitter.

Murni (2023). **Comparison of Naïve Bayes Classifier and Support Vector Machine Algorithms for Sentiment Analysis of Hate Speech on Twitter**. Thesis, Master Program of Informatics, Universitas Ahmad Dahlan, Yogyakarta.

Abstract

Twitter is a social media platform that values freedom of expression. The characteristics of Twitter, which facilitate the rapid sharing of Tweets, can lead to the spread of Hate Speech. Hate Speech refers to verbal expressions with the intention of demeaning individuals or specific groups based on attributes such as ethnicity, religion, race, and inter-group relations (SARA). Therefore, the identification and handling of Hate Speech have become a concern in efforts to maintain the safety and comfort of social media users. The use of Machine Learning algorithms in sentiment analysis is necessary to detect whether a Tweet falls into the category of Hate Speech or Non-Hate Speech on Twitter. The goal of this research is to conduct sentiment analysis on Tweets on the Twitter social media platform using Machine Learning algorithms to identify the appropriate algorithm for recognizing Tweets that fall into the category of Hate Speech.

Sentiment analysis is part of natural language processing to detect emotional tones in text. This research employs the Naïve Bayes Classifier and *Support Vector Machine* Machine Learning algorithms for identifying Hate Speech Tweets. The research stages include data collection from Twitter using Twitterscraper and the Twitter API, manual labeling, preprocessing (case folding, tokenization, stopword removal, and stemming), feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF), stratified k-fold, and evaluation using the Confusion Matrix. The programming language used for analysis and evaluation in this research is Python.

The dataset collected consists of 5000 entries, with 3988 labeled as Hate Speech and 1012 as Non-Hate Speech sentiments. The data undergoes preprocessing stages, including case folding, tokenization, stopword removal, and stemming. Data that has undergone stemming is transformed using TF-IDF for feature extraction. Due to imbalanced sentiment in the analyzed data, the *Stratified K-Fold* technique is employed. The models analyzed include Multinomial Naïve Bayes (MNB) and Linear *Support Vector Machine* (LinearSVC). The research results indicate that LinearSVC excels in identifying Hate Speech sentiments with a precision of 85%, while MNB is better at recognizing Hate Speech sentiments with a recall of 100%. Although both models have the same F1-Score, LinearSVC has higher accuracy at 85%, compared to MNB with an accuracy of 82%, demonstrating the superiority of SVM over MNB.

Keywords: Machine Learning, Sentiment Analysis, *Hate Speech*, Twitter