Analisis Sentimen Ujaran Kebencian pada Media Sosial Twitter Menggunakan Metode Support Vector Machine

Dwi Ahmad Dzulhijjah[1)], Muh. Fitra Rizki [2)], Mutiara Sholawati[3)]
Program Studi Teknik Informatika S1, Fakultas Teknologi Industri
Institut Teknologi Nasional Malang, Jalan Raya Karanglo km 2 Malang, Indonesia
[1)]*1818101@scholar.itn.ac.id,* [2)]*1718053@scholar.itn.ac.id,*[3)]*1818097@scholar.itn.ac.id*

ABSTRAK

Twitter merupakan salah satu media sosial yang memiliki pengguna dari berbagai kalangan masyarakat. Banyak informasi yang dapat diakses dengan mudah dan cepat. Hal ini tidak hanya memberikan dampak positif, tetapi juga dampak negatif bagi pengguna maupun non-pengguna twitter. Salah satu dampak negatif yang dapat dirasakan adalah adanya kasus ujaran kebencian yang diutarakan oleh oknum tidak bertanggung jawab kepada korban melalui perangkat teknologi, dalam hal ini adalah media sosial Twitter. Ujaran kebencian akan sangat berdampak bagi kalangan individu maupun kelompok, salah satunya timbul perpecahan dikalangan masyarakat. Hal ini mengakibatkan hubungan baik yang terjalin antar kelompok ras atau suku, agama, gender, disabilitas, politik dan berbagai kelompok lainnya menjadi renggang dan dapat mengikis rasa kesatuan dan persatuan rakyat Indonesia.Selain itu pelaku ujaran kebencian dapat dijerat pasal UU ITE Pasal 28 Ayat 2 tentang menyebarkan informasi yang dapat menimbulkan rasa kebencian atau permusuhan individu.

Pada penelitian ini dilakukan proses analisis sentimen ujaran kebencian yang bersumber dari cuitan pengguna twitter dengan mengembangkan sistem menggunakan Google Collaboratory untuk mengklasifikasikan sentimen menggunakan metode *Support Vector Machine (SVM)*. Data yang digunakan dalam penelitian ini berupa dataset tweet dengan jumlah 13169 data yang terdiri dari 7608 tweet yang bukan ujaran kebencian dan 5561 tweet ujaran kebencian. Output yang diperoleh dari sistem ini adalah berupa label prediksi "Negatif"apabila tweet tidak mengandung ujaran kebencian, dan "Positif" apabila tweet mengandung ujaran kebencian.

Dari hasil pengujian menunjukan dengan menggunakan sejumlah 13,169 data tweet, sistem mampu melakukan proses klasifikasi untuk memprediksi teks dengan sentimen positif ujaran kebencian dan negatif ujaran kebencian. Diperoleh pula hasil pengukuran evaluasi klasifikasi dengan menggunakan metode confusion matrix dengan nilai recall 82%, precision 91%, F1-Score 86% dan tingkat accuracy sebesar 83%.

***Kata Kunci** : Tweet, Support Vector machine, Analisis Sentiment, Ujaran Kebencian, Text Preprocessing, Term Frequency-Inverse Document Frequency*

PENDAHULUAN

Twitter merupakan platform media sosial yang menyediakan interaksi pengguna melalui tweet yang berisikan teks ataupun foto dan video. Twitter merupakan media sosial yang umum digunakan di Indonesia, pada bulan April 2021 pengguna aktif Twitter tercatat sebesar 15.1 juta pengguna dan terbanyak keenam di dunia [1]. Dengan pengguna aktif yang begitu banyak dan hasil interaksi antar pengguna berupa tweet maka menghasilkan berbagai macam opini atau sentimen. Sentimen yang muncul beragam mulai dari sentimen yang santun dalam berbahasa maupun sentimen yang tidak santun dalam berbahasa seperti ujaran kebencian.

Berdasarkan definisi Fortuna dan Nunes bahwa ujaran kebencian merupakan bahasa yang menyerang atau mengurangi, yang menghasut kekerasan atau kebencian terhadap suatu kelompok, berdasarkan ciri-ciri tertentu seperti penampilan fisik, agama, keturunan, asal kebangsaan atau suku, orientasi seksual, identitas gender atau lainnya, dan dapat terjadi dengan perbedaan bahasa. gaya, bahkan dalam bentuk halus atau ketika humor digunakan.[2] Ujaran kebencian merupakan bentuk dari sikap yang bertolak belakang dengan konsep kesantunan berbahasa, sama halnya dengan etika berkomunikasi. Berdasarkan UU ITE Pasal 28 Ayat 2 menyatakan bahwa setiap orang dilarang dengan sengaja dan tanpa hak menyebarkan informasi yang ditujukan untuk menimbulkan rasa kebencian atau permusuhan individu dan/atau kelompok masyarakat tertentu berdasarkan atas suku, agama, ras, dan antargolongan (SARA) [3]. Dengan adanya UU ITE Pasal 28 Ayat 2 maka selayaknya pengguna Twitter lebih memperhatikan tweetnya saat diterbitkan, apakah sudah memenuhi etika dan hukum dalam berkomunikasi di Indonesia. Dari permasalahan diatas maka terdapat banyak penelitian yang mencoba untuk mengklasifikasikan ujaran kebencian

Dari permasalahan yang diuraikan diatas maka penelitian ini mencoba untuk mengklasifikasikan apakah tweet mengandung ujaran kebencian atau tidak mengandung ujaran kebencian menggunakan algoritma *Support Vector Machine* (SVM). Sehingga dengan adanya penelitian ini diharapkan dapat membantu meningkatkan efektifitas dalam mengetahui adanya ujaran kebencian dalam suatu tweet.

Beberapa penelitian terkait analisis sentimen ujaran kebencian dilakukan oleh Yonathan Sari

Mahardhika dan Eri Zuliarso di tahun 2018 yaitu mengembangkan analisis sentimen terhadap pemerintahan joko widodo pada media sosial twitter menggunakan algoritma Naives Bayes Classifier. Pada penelitian ini, dilakukan proses analisis sentimen pengguna twitter dengan memasukkan keyword #2019gantipresiden, #2019tetapjokowi dan tweet yang berhubungan dengan pemerintahan pada saat ini yang dipimpin oleh presiden Joko Widodo dengan jumlah data tidak melebihi 400 data tweet. Hasil penelitian tersebut menunjukkan hasil tingkat akurasi yang didapatkan dengan melakukan pengujian terhadap 300 data latih dan 100 data uji dokumen tweet adalah 97% [4].

Terdapat acuan penelitian lainnya dalam mengklasifikasi ujaran kebencian seperti pada penelitian Rahman et al dengan hasil uji F-Measurement nya sebesar 53.07%. Adapun dalam penelitian ini, sistem yang akan dibuat menggunakan metode klasifikasi dari Sentiment Analysis dengan algoritma Support Vector machine (SVM). Sehingga dengan adanya penelitian ini dapat memudahkan berikutnya bagi pengguna untuk mendeteksi apakah tweetnya mengandung ujaran kebencian atau tidak[5].

ANALISIS SENTIMEN

Analisis Sentimen adalah klasifikasi dari opini dan sentimen yang diungkapkan dalam teks, yang dihasilkan oleh manusia melalui teknologi Data Mining. Analisis Sentimen menyediakan fitur ekstraksi otomatis dan kemampuan representasi dan kinerja yang lebih baik daripada teknik berbasis fitur tradisional [6]. Analisis Sentimen ditujukan untuk mencari pendapat orang lain. Ini tidak hanya berlaku untuk individu tetapi juga berlaku untuk organisasi. Contohnya saat ini, jika seseorang ingin membeli produk konsumen, tidak lagi terbatas untuk meminta pendapat teman dan keluarga seseorang karena ada banyak ulasan pengguna dan diskusi tentang produk di forum publik di website. Bagi sebuah organisasi, mungkin tidak perlu lagi melakukan survei, jajak pendapat, dan memfokuskan diri untuk mengumpulkan opini publik. Beberapa tahun terakhir, postingan pendapat di media sosial juga telah membantu membentuk bisnis, mempengaruhi sentimen publik dan emosi publik [7].

UJARAN KEBENCIAN

Ujaran kebencian didefinisikan sebagai komunikasi apa pun yang dapat menyerang seseorang maupun kelompok dalam berbagai aspek seperti ras, agama, warna kulit, etnis, kewarganegaraan, jenis kelamin, disabilitas serta orientasi seksual.Ujaran kebencian sangat berbahaya karena dapat menimbulkan konflik sosial di tengah masyarakat. Adanya ujaran kebencian membuat beberapa pihak salah satunya Twitter menginvestasikan jutaan euro untuk mengatasi ujaran kebencian di platform mereka. Namun,sebagian besar upayanya masih dilakukan secara manual oleh pengguna twitter sehingga mengakibatkan penangannya sedikit lambat. Oleh karena itu, sistem analisis sentimen ujaran kebencian diperlukan untuk mengatasi masalah tersebut agar lebih efektif dalam penangannya[8].

TERM FREQUENCY dan INVERSE DOCUMENT FREQUENCY

Algoritma TF-IDF (Term Frequency – Inverse Document Frequency) adalah salah satu algoritma yang dapat digunakan untuk menganalisa hubungan antara sebuah frase/kalimat dengan sekumpulan dokumen.

1. Term Frequency (TF)

TF (Term Frequency) adalah frekuensi dari kemunculan sebuah term dalam dokumen yang bersangkutan. Semakin besar jumlah kemunculan suatu term (TF tinggi) dalam dokumen, semakin besar pula bobotnya atau akan memberikan nilai kesesuaian yang semakin besar. Term Frequency (TF) dapat diformulasikan seperti berikut ini :

$$TF = 1 + \log log \left(F_{t,d}\right), \; F_{t,d} > 0 \qquad (2.1)$$

Dimana nilai $F_{t,d}$ adalah frekuensi term (t) pada document (d). Jadi jika suatu kata atau term terdapat dalam suatu dokumen sebanyak 5 kali maka diperoleh bobot = 1 + log (5) = 1.699. Tetapi jika term tidak terdapat dalam dokumen tersebut, bobotnya adalah nol (0).

2. Inverse Document Frequency (IDF)

IDF (Inverse Document Frequency) merupakan sebuah perhitungan dari bagaimana term didistribusikan secara luas pada koleksi dokumen yang bersangkutan. IDF menunjukkan hubungan ketersediaan sebuah term dalam seluruh dokumen. Semakin sedikit jumlah dokumen yang mengandung term yang dimaksud, maka nilai IDF semakin besar.Sedangkan untuk Inverse Document Frequency (IDF) dihitung dengan menggunakan formula sebagai berikut:

$$IDF = log(\frac{D}{df}) \quad (2.2)$$

Dimana D adalah jumlah semua dokumen dalam koleksi sedangkan dfj adalah jumlah dokumen yang mengandung term (tj).

Jenis formula TF yang biasa digunakan untuk perhitungan adalah TF murni (raw TF). Dengan demikian rumus umum untuk Term Weighting TF-IDF adalah penggabungan dari formula perhitungan raw TF dengan formula IDF dengan cara mengalikan nilai TF dengan nilai IDF [10].

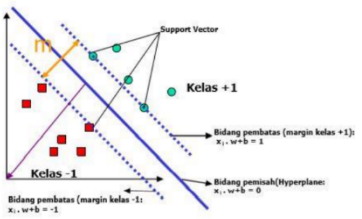
$$W = tf \; x \; idf \qquad (2.3)$$

SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) adalah suatu teknik untuk melakukan suatu prediksi, baik dalam kasus klasifikasi atau regresi. Metode SVM memiliki prinsip dasar linear classifier yaitu kasus klasifikasi yang dapat dipisahkan secara linier, namun SVM

yang dikembangkan dapat bekerja dengan problem non- linier dengan memasukkan konsep kernel pada ruang berdimensi tinggi. Pada ruang berdimensi tinggi, hyperplane yang akan dapat memaksimalkan jarak(margin) antara kelas data. Metode Support Vector Machine (SVM) berakar dari teori pembelajaran statistik yang dapat memberikan hasil yang lebih baik dari metode yang lain. [11]

SVM dapat bekerja dengan data pada non linier dengan menggunakan pendekatan kernel pada fitur awal himpunan data. Fungsi kernel yang digunakan untuk memetakan dimensi awal (dimensi yang lebih rendah) himpunan data ke dimensi baru (dimensi yang lebih tinggi). Konsep SVM adalah pencarian hyperplane terbaik yang berfungsi sebagai pemisah data dari dua kelas pada input space. Hyperplane pemisah terbaik adalah hyperplane yang terletak di tengah diantara dua set objek dari dua kelas. Hyperplane terbaik dapat dicari dengan memaksimalkan margin atau jarak dari dua set objek dari dua kelas yang berbeda. Dapat diasumsikan bahwa kedua belah kelas dapat terpisah secara sempurna oleh hyperplane(linear separable). Akan tetapi, pada umumnya dua belah kelas pada input space tidak dapat terpisah secara sempurna(non linear separable)[12]. Untuk mengatasi masalah ini SVM dirumuskan ulang dengan memperkenalkan metode margin. Seperti pada Gambar 1. [11][13]



Gambar 1 Margin Hyperplane

COFUSSION MATRIX

Untuk melakukan evaluasi hasil prediksi teks atau klasifikasi hasil analisis sentimen bisa dilakukan menggunakan beberapa nilai yaitu nilai *Recall*, *Precision*, *specificity, accuracy dan F1-Score*. Nilai-nilai tersebut didasarkan pada hasil *confusion matrix* yang didapatkan setelah melakukan pengujian model dan data testing[14][15]. *Confusion matrix* dapat dilihat pada tabel 1 dibawah ini [12]:

Tabel 1 Tabel Confusion Matrix

Predicted Class	**True Class**	
	Positif	**Negatif**
Positif	**True Positif (TP)**	**False Positif (FP)**
Negatif	**False Negatif (FN)**	**True Negatif (TN)**

Keempat parameter diatas digunakan untuk menghitung 4 metode evaluasi yakni:

1. *Recall*, yaitu perbandingan jumlah dokumen yang relevan dikenali dengan jumlah seluruh dokumen relevan. *Recall* memiliki rumusan yang ditunjukkan pada persamaan 1.

$$Recall = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

2. *Precision,* yaitu perbandingan jumlah dokumen yang relevan terkenali dengan jumlah dokumen yang terkenali. Untuk mendapatkan nilai precision dihitung menggunakan persamaan 2.

$$Precision = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

3. *Accuracy*, yaitu seberapa akurat sistem dapat mengklasifikasikan yang dihitung menggunakan persamaan 3.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \times 100\% \quad (3)$$

4. *F1-Score*, yaitu seberapa harmoni model mean klasifikasi dari precision dan recall nilai F1-score didapatkan menggunakan persamaan 4.

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \times 100\% \quad (4)$$

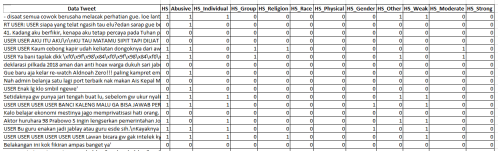
Nilai *recall, precision*, *accuracy dan F1-Score* dinyatakan dalam persen. Semakin tinggi persentase ketiga nilai tersebut menunjukan semakin baik kinerja sistem klasifikasi teks otomatis.

METODE PENELITIAN

Beberapa proses dilakukan untuk melakukan analisis sentimen data tweet yang mengandung ujaran kebencian atau tidak. Proses tersebut terdiri dari 5 proses yaitu pengumpulan data, text processing, pembobotan TF-IDF, penerapan metode Support Vector Machine (SVM) dan evaluasi metode SVM.

1. Pengumpulan Data

Data masukan yang digunakan merupakan data tweet yang bersumber dari penelitian Ibrahim dan Budi pada tahun 2019 yang bisa diakses pada https://github.com/okkyibrohim/id-multi-label-hate-speech-and-abusive-language-detection/blob/master/re_dataset.csv [16]. Data yang digunakan sebanyak 13,169 tweet yang terdiri dari 7608 tweet yang bukan ujaran kebencian dan 5561 tweet ujaran kebencian. Gambaran potongan dataset yang digunakan ditunjukkan pada Gambar 2.



Gambar 2. Potongan dataset tweet yang digunakan

2. Proses Text Preprocessing

Sebelum data tweet dilakukan klasifikasi terlebih dahulu data tweet akan melalui tahap preprocessing. Proses text preprocessing yang dilakukan diantaranya yakni case folding, cleansing, normalisasi, stopword removal, dan stemming.Berikut merupakan penjelasan dari tahap preprocessing data :

1) Pada tahapan case folding, teks dilakukan proses perubahan dari huruf besar menjadi huruf kecil dan menghilangkan seluruh tanda baca pada kalimat.

2) Pada tahapan cleansing, pembersihan terhadap data yang redundan atau ganda, inkonsisten, missing value dan outlier data.

3) Pada tahapan normalisasi, teks dilakukan proses mengubah dari kata tidak baku menjadi baku dan menghapus emoji.

4) Pada tahapan stopword removal, teks yang tidak memiliki makna dan muncul dalam jumlah yang besar akan dihapus.

5) Pada tahapan steming, teks yang merupakan kata berimbuhan akan menjadi kata dasar.

3. Proses Pembobotan TF-IDF

Pada penelitian ini pembobotan diperoleh dari frekuensi sebuah kata yang terdapat di dalam dokumen tweet atau jumlah kemunculan term dalam satu dokumen term frequency (tf) dan sebuah kata di dalam kumpulan dokumen atau jumlah kemunculan term dalam koleksi dokumen inverse document frequency (idf). Pada tahap ini menggunakan library CountVectorizer dan TfidfVectorizer.

4. Proses Metode Support Vector Machine (SVM)

Data yang akan diproses untuk diprediksi akan dibagi ke dalam dua jenis, data uji dan data ajar dan data tes, pada tahap ini kami menggunakan library train_test_split. Kemudian tahap berikutnya adalah menerapkan model SVM menggunakan fungsi SVC() dari sklearn untuk dilakukan klasifikasi SVM.

5. Proses Evaluasi Metode SVM

Setelah SVM dimodelkan dan diterapkan pada data, selanjutnya dilakukan tahap menghitung confusion matrix guna mengetahui seberapa akurat metode yang digunakan kemudian menggunakan library classification_report akan mengetahui laporan klasifikasi dengan menampilkan hasil perhitungan dari precision, recall, accuracy, F1-Score dan Support.

HASIL DAN PEMBAHASAN

Penerapan masing-masing proses yang sudah dirancang di metode penelitian kemudian diwujudkan menggunakan bahasa pemrograman python yang dilakukan pada IDE google colab untuk memudahkan tim penelitian melakukan kolaborasi, dimana hasil dari penerapan tersebut dapat diakses di https://colab.research.google.com/drive/1xaJO-Y-AERT-ernceTjLWulaadQsFwfj?usp=sharing.

Penerapan Preprocessing

Hasil dari tahap preprocessing adalah kata yang sudah menjadi kata dasar, tanpa noise seperti imbuhan dan menjadi kata baku. Berikut perbandingan antara data sebelum dan sesudah dilakukan proses preprocessing.

Data berikut adalah contoh data sebelum dilakukan tahap preprocessing

> - disaat semua cowok berusaha melacak perhatian gue. loe lantas remehkan perhatian yg gue kasih khusus ke elo. basic elo cowok bego ! ! !”

Data berikut adalah contoh data hasil proses preprocessing

> cowok usaha lacak perhati lantas remeh perhati kakak khusus basic cowok bego

Penerapan Pembobotan TF-IDF

Pada tahap TF-IDF tampak hasil dari pembobotan beberapa kata menggunakan library CountVectorizer dan TfidfVectorizer.
Berikut beberapa hasil pembobotan TF-IDF pada Gambar 3:

	TF-IDF
perhati	0.478833
cowok	0.448788
basic	0.335361
lacak	0.317111
remeh	0.300320
lantas	0.288943

Gambar 3. Hasil Perhitungan TF-IDF

Perwujudan Metode SVM

Setelah melalui tahap pembobotan TF-TDF, dataset akan dibagi menjadi data latih (*training*) dan data uji (*testing*) dengan rasio 70:30, 70% sebagai data latih dan 30% sebagai data uji. Setelah itu, membuat model klasifikasi menggunakan fungsi SVC() yang terdapat dalam library sklearn. Hasil dari model tersebut diterapkan pada data uji menghasilkan nilai *precision, recall, f1-score, dan accuracy* yang ditunjukkan pada Gambar 4 :

	precision	recall	f1-score	support
0	0.91	0.82	0.86	2553
1	0.72	0.86	0.78	1398
accuracy			0.83	3951
macro avg	0.82	0.84	0.82	3951
weighted avg	0.84	0.83	0.83	3951

Gambar 4 Akurasi Model Klasifikasi

Tingkat akurasi dari model klasifikasi analisis sentimen sebesar 83% menggunakan library dan fungsi dari SKLEARN.

Evaluasi Hasil Klasifikasi Metode SVM

Evaluasi klasifikasi bertujuan untuk mengecek kebenaran dari tingkat akurasi model yang telah dibuat. Perhitungan tingkat akurasi klasifikasi mengacu pada perhitungan tabel 1 confusion matrix. Untuk mengetahui nilai confusion matrix dapat

menggunakan fungsi confusion_matrix() yang terdapat dalam library sklearn. Sehingga diperoleh hasil seperti yang ditunjukkan pada Gambar 5:

```
array([[2084,  469],
       [ 200, 1198]])
```

Gambar 5 Confusion Matrix

Berikut Tabel 2 confusion matrix:

Tabel 2 Tabel Confusion Matrix

Predicted Class	True Class		Total
	Positif	Negatif	
Positif	2084	469	2553
Negatif	200	1198	1398
Total	2284	1667	3951

Berdasarkan tabel 2, diperoleh dokumen dengan prediksi positif dan faktanya positif sebanyak 2084 tweet, dokumen dengan prediksi positif dan faktanya negatif sebanyak 469 tweet, dokumen dengan prediksi negatif dan faktanya negatif sebanyak 1198 tweet, serta dokumen dengan prediksi negatif dan faktanya positif sebanyak 200 tweet.

Setelah melakukan perhitungan *confusion matrix* maka diperoleh nilai-nilai yang akan digunakan dalam proses perhitungan *recall* dengan persamaan (5), *precision* dengan persamaan (6) , *accuracy* dengan persamaan (7) dan *f1-score* dengan persamaan (8). Berikut perhitungan untuk masing evaluasi yang dilakukan:

$$Recall = \frac{TP}{TP+FP} \times 100\% = \frac{2084}{2084+469} \times 100\% = 82\% \quad (5)$$

$$Precision = \frac{TP}{TP+FN} \times 100\% = \frac{2084}{2084+200} \times 100\% = 91\% \quad (6)$$

$$F_1 = \frac{TP}{TP+\frac{1}{2}(FP+FN)} \times 100\% = \frac{2084}{2084+\frac{1}{2}(469+200)} \times 100\% = 86\% \quad (7)$$

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP} \times 100\% = \frac{1198+2084}{1198+2084+200+469} \times 100\% = 83\% \quad (8)$$

Sistem analisis sentimen ujaran kebencian yang dibuat telah mampu mengklasifikasikan tweet mengandung ujaran kebencian atau tweet tidak mengandung ujaran kebencian. Sistem analisis sentimen yang dirancang menggunakan metode *Support Vector Machine* (SVM) menghasilkan nilai *recall* mencapai 82%, *precision* mencapai 91% *f1-score* mencapai 86% dan *accuracy* mencapai 83%. Sehingga dapat dikatakan metode SVM dapat digunakan sebagai metode analisis sentimen yang baik.

Hasil Prediksi Akhir

Setelah dilakukan perhitungan tingkat akurasi dan memperoleh tingkat akurasi yang cukup tinggi, maka dilakukan tahap uji coba dengan menginputkan kalimat mengandung ujaran kebencian dan tidak mengandung ujaran kebencian. Sistem akan mencetak label "Positif" apabila kalimat yang diinputkan merupakan kalimat dengan ujaran kebencian (Gambar 6) dan label "Negatif" apabila kalimat tidak mengandung ujaran kebencian (Gambar 7). Berikut hasil prediksi analisis sentimen ujaran kebencian:

```
prediksi('dasar bego, gapunya otak')
```

```
'positif'
```

Gambar 6 Prediksi Positif

```
prediksi('Meski berat melangkah hatiku, hanya tak siap terluka')
```

```
'negatif'
```

Gambar 7 Prediksi Negatif

KESIMPULAN DAN SARAN

Dari evaluasi hasil metode *support vector machine* pada analisis sentimen ujaran kebencian yang telah dilakukan menggunakan 13169 data tweet menunjukkan bahwa penerapan metode menghasilkan nilai *recall* 83%, *precision* 84%, *f1-Score* 83% dan tingkat *accuracy* sebesar 84%. Namun jumlah data training dan ketepatan proses *text preprocessing* dalam menghilangkan dan membersihkan *noise* dokumen tweet mempengaruhi tingkat akurasi dan kinerja klasifikasi sentimen ujaran kebencian menggunakan metode *support vector machine* pada penelitian ini. Untuk pengembangan selanjutnya perlu ditingkatkan proses pembersihan dokumen tweet pada *text preprocessing* agar hasil lebih maksimal.

DAFTAR PUSTAKA

[1] Statista. 2021. *Leading countries based on number of Twitter users as of April 2021*. [Online] https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/.

[2] Fortuna, P. and Nunes, S. 2018. *A Survey on Automatic Detection of Hate Speech in Text. ACM Computing Surveys (CSUR)*. ACM, New York, NY, 19-33. DOI= http://doi.acm.org/10.1145/3232676.

[3] Republik Indonesia. 2016. Undang-undang ITE Pasal 28 Ayat 2 E Tentang Informasi dan Transaksi Elektronik. Jakarta : Kementerian Komunikasi dan Informatika Republik Indonesia.

[4] Yonathan Sari Mahardhika , Eri Zuliarso, 2018. *Analisis Sentimen Terhadap Pemerintahan Joko Widodo Pada Media Sosial Twitter Menggunakan Algoritma Naives Bayes Classifier*. Prosiding SINTAK 2018. ISBN: 978-602-8557-20-7. Teknik Informatika Universitas Stikubank.

[5] Umar Syahid Aulia Rahman. et al. 2020. *Implementasi Multinomial Naive Bayes Untuk Klasifikasi Ujaran Kebencian Pada Dataset Kicauan (Twitter) Bahasa Indonesia*. JATIKOM. Vol 3, No 2, 2020.

[6] Araque, O. et al. 2017. *Enhancing deep learning sentiment analysis with ensemble techniques in social applications.* Expert Systems with

Applications. Elsevier Ltd, 77, pp. 236–246. doi: 10.1016/j.eswa.2017.02.002.

[7] Zhang, L., Wang, S. and Liu, B. 2018. *Deep learning for sentiment analysis: A survey*, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), pp. 1–25. doi: 10.1002/widm.1253.

[8] Nofa Aulia, Indra Budi. 2019. *Hate Speech Detection on Indonesian Long Text Documents Using Machine Learning Approach*. ICCAI '19: Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence. Halaman 164–169. doi: https://doi.org/10.1145/3330482.3330491

[9] A. F. Hidayatullah and A. SN, "Analisis Sentimen Dan Klasifikasi Kategori Terhadap Tokoh Publik Pada Twitter," Seminar Nasional Informatika, vol.2 no.4, pp. 1–8, Okt 3,2015.

[10] https://informatikalogi.com/term-weighting-tf-idf/ , diakses tanggal 05 April 2021, pukul 13:00 WIB.

[11] Oryza Habibie Rahman, Gunawan Abdillah, & Agus Komarudin. 2021. *Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine*. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), vol.4 no.2, pp.17-23. https://doi.org/10.29207/resti.v5i1.2700

[12] Muh. Fitria Rizki, Karina Auliasari, Renaldi Primaswara Prasetya. 2020. *Analisis Sentiment Cyberbullying Pada Sosial Media Twitter Menggunakan Metode Support Vector Machine*. Jurnal JATI (Jurnal Mahasiswa Teknik Informatika), vol.4 no. 2.

[13] Achmad, R.R. Imron, S.G & Sidik, A.P. 2019.*Klasifikasi Wajah Menggunakan Support Vector Machine (SVM)*. Riset dan E-Jurnal Manajemen Informatika Komputer. vol.3 no.2.

[14] S. Ruuska, W. Hämäläinen, S. Kajava, M. Mughal, P. Matilainen, and J. Mononen, .2016. *Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle*. Behav. Processes, vol. 148, pp. 56–62.

[15] H. He and Gracia.2009.*Learning from imbalanced data*. IEEE Trans. Knowl. data Eng., vol. 21, no. 9, pp. 1263-1284.

[16] Muhammad Okky Ibrohim and Indra Budi. 2019. *Multi-label Hate Speech and Abusive Language Detection in Indonesian Twitter*. In ALW3: 3rd Workshop on Abusive Language Online, 46-57.