

IMPLEMENTASI ALGORITMA TF-IDF DAN *SUPPORT VECTOR MACHINE* TERHADAP ANALISIS PENDETEKSI KOMENTAR *CYBERBULLYING* DI MEDIA SOSIAL TIKTOK

Romindo ¹⁾, Jefri Junifer Pangaribuan ²⁾, Okky Putra Barus ³⁾

^{1), 2), 3)} Universitas Pelita Harapan

Email : romindo@uph.edu ¹⁾, jefri.pangaribuan@uph.edu ²⁾, okky.barus@uph.edu ³⁾

ABSTRAK

Cyberbullying merupakan tindakan berupa kiriman teks, gambar, atau video dengan memanfaatkan internet, *handphone* atau perangkat lain dengan tujuan menyakiti dan mempermalukan orang lain. *Cyberbullying* sering dilakukan melalui beberapa platform media sosial, salah satunya melalui komentar pada aplikasi TikTok. Menurut laporan dari *We Are Social*, TikTok memiliki 1,4 miliar pengguna aktif bulanan yang berusia 18 tahun secara *global*. Indonesia saat ini berada di posisi kedua dalam jumlah pengguna aktif TikTok di seluruh dunia. Oleh karena itu, potensi terjadinya kasus *cyberbullying* akan semakin besar apabila penggunaannya semakin banyak. Dengan menggunakan *data mining*, masyarakat dapat mengembangkan suatu sistem pendeteksi, dimana sistem tersebut dapat melakukan analisis pada komentar yang ada di aplikasi TikTok. Metode yang digunakan adalah *Term Frequency-Inverse Document Frequency* (TF-IDF) dan *Support Vector Machine* (SVM). Tahap – tahap yang dilalui adalah dengan mengumpulkan komentar-komentar yang dilabeli secara manual. Kemudian, dilakukan *text preprocessing*, *tokenizing*, dan pembobotan dengan TF-IDF. Kemudian, mengimplementasikan algoritma *Support Vector Machine* untuk mendeteksi komentar *cyberbullying*. Penelitian ini menggunakan 80% *data training* dan 20% *data testing*. Dari hasil performa algoritma *Support Vector Machine*, diperoleh 88% akurasi keseluruhan, 88% *precision*, 96% *recall*, dan 92% *f1-score* dalam mendeteksi komentar *cyberbullying* pada media sosial TikTok.

Kata Kunci : Komentar *Cyberbullying*, *Machine Learning*, TF-IDF, SVM.

ABSTRACT

Cyberbullying is the act of sending text, images, or videos using the internet, mobile phones, or other devices with the aim of hurting and shaming other people. Cyberbullying is often done through several social media platforms, one of which is through comments on the TikTok application. According to a report by We Are Social, TikTok has 1.4 billion monthly active users aged 18 and above globally. Indonesia currently ranks second in the world in terms of active TikTok users. As a result, the potential for cyberbullying instances will grow as the number of users grows. By using data mining, the public can create a detection system, which can perform analysis on comments in the TikTok application. The method used is Term Frequency-Inverse Document Frequency (TF-IDF) and Support Vector Machine (SVM). The stages passed are to collect comments that are labelled manually. Then, text preprocessing, tokenizing, and weighting were carried out with TF-IDF. Then, implement the Support Vector Machine algorithm to detect cyberbullying comments. This study uses 80% training data and 20% testing data. From the performance results of the Support Vector Machine algorithm, 88% overall accuracy, 88% precision, 96% recall, and 92% f1-score were obtained in detecting cyberbullying comments on social media TikTok.

Keywords: *Cyberbullying Comments, Machine Learning, TF-IDF, SVM.*

1. PENDAHULUAN

Saat ini segala sesuatu yang dikerjakan telah dipermudah dengan adanya internet. Penggunaan internet semakin meningkat dari tahun ke tahun. Adapun survei yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) mengenai pengguna internet Indonesia pada tahun 2021 hingga 2022, yaitu sebanyak lebih dari 210 juta jiwa penduduk Indonesia atau sekitar 77,02% dari total populasi jiwa penduduk Indonesia (Dimas Bayu, 2022). Adapun pengaruh yang diakibatkan oleh *pandemic* COVID-19 yang secara signifikan mengubah kebiasaan hidup sehari-hari manusia yang dinamakan sebagai *new normal*. *New normal* membuat manusia lebih bergantung kepada digital platform.

Sejak Maret 2020 pengguna internet di seluruh dunia meningkat 50 – 70%. Dengan peningkatan teknologi internet, manusia juga menghadapi efek negatif dari peningkatan tersebut, yang dikarenakan beberapa individu menyalahgunakan kegunaan dari media sosial dan menggunakannya sebagai media penipuan dengan bertujuan untuk mengintimidasi orang lain, yang disebut juga *cyberbullying* (Mitsu and Dawood, 2022). Menurut The National Crime Prevention Council, *cyberbullying* adalah pengiriman berupa teks, gambar, atau video baik menggunakan internet, handphone, maupun perangkat – perangkat lain dengan tujuan menyakiti dan mempermalukan orang lain (Juditha, 2021).

Dengan perkembangan internet, orang-orang menggunakan media sosial sebagai platform untuk berpartisipasi, berbagi, mengunggah ekspresi melalui teks, foto, ataupun video. Media sosial merupakan sebuah medium di internet dimana penggunaannya dapat mengekspresikan dirinya dan melakukan komunikasi, relasi dengan orang lain secara virtual (Puspitarini and Nuraeni, 2019). Salah satu media sosial yang saat ini sedang naik daun adalah Aplikasi TikTok. Berdasarkan laporan We Are Social, Aplikasi TikTok memiliki 1,4 miliar pengguna aktif bulanan yang berusia di atas 18 tahun secara global. Indonesia saat ini berada di posisi kedua dengan jumlah pengguna aktif

TikTok di seluruh dunia (Monavia Ayu Rizaty, 2022). Hubungan *cyberbullying* dan Aplikasi TikTok dapat dilihat dari komentar – komentar dari pengguna atau pembuat video. Kata atau kalimat yang mengandung unsur *cyberbullying* sangat sering dijumpai.

TF-IDF merupakan gabungan dari dua kata yaitu, Term Frequency dan Inverse Document Frequency. Term Frequency digunakan untuk mengukur seberapa sering sebuah istilah hadir dalam sebuah dataset. Sedangkan, Inverse Document Frequency akan melakukan pembagian bobot kepada istilah yang sering muncul dan jarang muncul (Qaiser and Ali, 2018). *Support Vector Machine* digunakan untuk mengubah data hasil pelatihan ke dimensi yang lebih tinggi dengan menggunakan pemetaan linear atau nonlinear (Ritonga and Purwaningsih, 2018).

Penelitian sebelumnya yang dilakukan oleh Kamal dan Rainarli pada tahun 2019 berjudul “Analisis Sentimen *Cyberbullying* pada Komentar Facebook Menggunakan Metode Klasifikasi *Support Vector Machine*” membahas tentang pengembangan model machine learning untuk mendeteksi komentar *cyberbullying* di platform media sosial Facebook. Penelitian difokuskan untuk mengklasifikasikan komentar menjadi sentimen positif (tidak mengandung unsur *cyberbullying*) dan sentimen negatif (mengandung unsur *cyberbullying*). Hasil penelitian menunjukkan bahwa penggunaan metode SVM dengan kernel RBF mencapai akurasi sebesar 96% dalam mendeteksi komentar *cyberbullying* (Kamal, 2019).

Studi lain sebelumnya yang dilakukan oleh Pardede pada tahun 2020 berjudul “Deteksi Komentar *Cyberbullying* di Media Sosial Berbahasa Inggris Menggunakan Klasifikasi *Naïve Bayes*” berfokus pada pendeteksian komentar *cyberbullying* di platform media sosial berbahasa Inggris. Penelitian diawali dengan teknik pengolahan teks, dilanjutkan dengan pemilihan fitur untuk mengubah teks menjadi matriks. Proses pengambilan keputusan menggunakan metode *Naïve Bayes*, menghasilkan akurasi sebesar

80%, presisi sebesar 81%, dan recall sebesar 80% (Pardede, Miftahuddin and Kahar, 2020).

Penelitian sebelumnya yang dilakukan oleh Nugraha Manoppo dan Hatta Fudholi pada tahun 2021 berjudul “Deteksi *Cyberbullying* Berdasarkan Unsur Pidana yang Dilanggar Menggunakan Naïve Bayes dan *Support Vector Machines*” berfokus pada pendeteksian *cyberbullying* berdasarkan unsur pidana yang dilanggar. Para peneliti menggunakan kumpulan data dari 5.000 tweet yang dikumpulkan dari Twitter. Mereka menggunakan metode N-gram dengan bobot TF-IDF untuk mengekstraksi sentimen berdasarkan penggunaan Bahasa. Hasil penelitian menunjukkan bahwa dengan memanfaatkan Naïve Bayes dan *Support Vector Machine* dengan metode SMOTE mampu memprediksi potensi *cyberbullying* yang diukur berdasarkan unsur kejahatan yang dilanggar dengan akurasi rata-rata 90% (Manoppo and Fudholi, 2021).

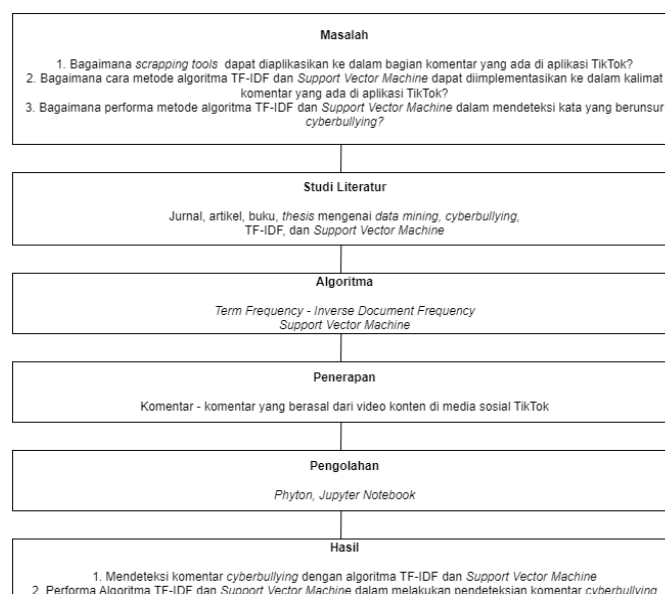
Berdasarkan paparan diatas, dapat dilakukan pengimplementasian algoritma TF-IDF dan *Support Vector Machine* untuk mendeteksi apakah komentar yang disampaikan pengguna satu kepada pengguna lainnya berunsur *cyberbullying* atau bukan. Maka penulis akan melakukan penelitian dengan judul “Implementasi Algoritma TF-IDF dan *Support Vector Machine* Terhadap Analisis Pendeteksi Komentar *Cyberbullying*

di Media Sosial TikTok” yang bertujuan untuk mencegah penyalahgunaan internet sebagai media *cyberbullying* terhadap orang lain.

2. METODE

Pengumpulan data penelitian diperoleh dari komentar pada beberapa postingan pembuat konten video di platform media sosial TikTok. Proses ekstraksi data dilakukan dengan memanfaatkan source code yang tersedia di GitHub. Setelah berhasil mengakses kode sumber, langkah selanjutnya adalah mengakses alat pengembang situs web. Kemudian, hapus semua kode sumber yang ada di konsol dan ganti dengan kode sumber yang diunduh sebelumnya dari GitHub. Setelah proses berjalan, tunggu hingga alat *scraping* berhasil mengambil komentar dari pembuat konten video TikTok.

Pada gambar 1 merupakan diagram yang menjelaskan bahwa dalam penelitian ini terdapat beberapa masalah yang ingin diselesaikan. Dengan memanfaatkan pengetahuan dari berbagai sumber dan melakukan penerapan algoritma ke dalam *dataset* yang sudah terkumpul. Kemudian, menggunakan bahasa pemrograman *Python*, *Jupyter Notebook*, algoritma TF-IDF, dan *Support Vector Machine* untuk melihat performa algoritma tersebut dalam mendeteksi komentar *cyberbullying* dalam media sosial TikTok.

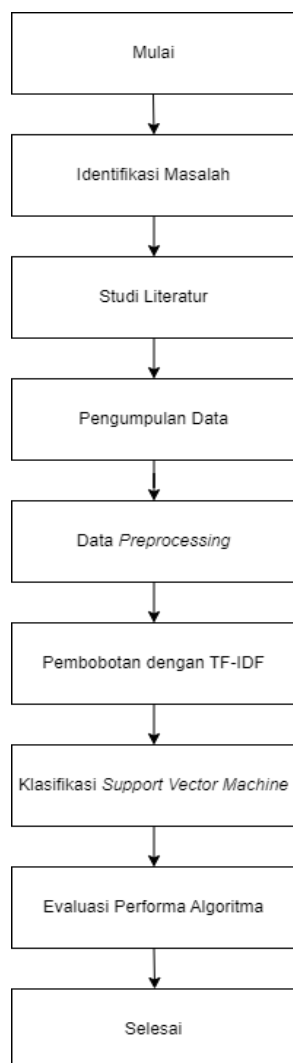


Gambar 1. Kerangka Pikir

Dalam diagram diatas dijelaskan bahwa dalam penelitian ini terdapat beberapa masalah yang ingin diselesaikan. Dengan memanfaatkan pengetahuan dari berbagai sumber dan melakukan penerapan algoritma ke dalam dataset yang sudah terkumpul. Kemudian, menggunakan bahasa pemrograman *Python*, *Jupyter Notebook*, algoritma TF-IDF, dan *Support Vector Machine* untuk melihat performa algoritma tersebut dalam mendeteksi komentar *cyberbullying* dalam media sosial TikTok.

2.1 Tahapan Penelitian

Untuk menyelesaikan permasalahan di dalam penelitian ini, diperlukan beberapa tahapan yang digambarkan dalam diagram sebagai berikut:



Gambar 2. Alur Tahapan Penelitian

Berdasarkan gambar diatas, alur tahapan penelitian ini dijelaskan sebagai berikut:

1. Identifikasi Masalah

Pada tahapan awal penelitian, diawali dengan merumuskan masalah, yaitu menganalisis pendeteksi komentar *cyberbullying* pada media sosial TikTok dengan menggunakan metode TF-IDF dan *Support Vector Machine*. Dengan tujuan untuk menganalisis tingkat akurasi metode tersebut dalam penelitian ini.

2. Studi Literatur

Pada tahap kedua, dilakukan pengumpulan materi berupa sumber referensi dan mengkaji teori yang berkaitan dengan topik penelitian, yang berfokus mengenai *cyberbullying*, klasifikasi, Pembobotan TF-IDF, Metode Support Vector Machine dari penelitian terdahulu.

3. Pengumpulan Data

Pada tahap ketiga, dilakukan pengumpulan data berupa komentar-komentar pada konten video yang diunggah pada media sosial TikTok.

4. Data labelling

Setelah tahap pengumpulan data, semua data berupa komentar – komentar akan diberikan label dengan angka 0 yang artinya negatif dan label dengan angka 1 yang artinya positif. Apabila terdapat komentar yang bersifat netral, maka komentar tersebut akan dihapus, dikarenakan pada penelitian ini penulis hanya menggunakan komentar yang bersifat negatif dan positif. Label negatif yang artinya memiliki unsur *cyberbullying* dan label positif yang artinya tidak memiliki unsur *cyberbullying*. Pada tahap ini, penulis berhasil mengumpulkan sebanyak 5610 komentar.

5. Data Preprocessing

Pada tahap kelima, dilakukan persiapan dan pengolahan data. Data yang sebelumnya sudah ada diproses dalam text preprocessing agar data tersebut dapat digunakan dan dibaca oleh sistem. Selanjutnya, preprocessing dilakukan pada Jupyter Notebook.

6. Pembobotan dengan Metode TF-IDF

Setelah melalui tahap text preprocessing, didapatkan sekumpulan data yang sudah siap untuk dilakukan uji coba klasifikasi. Data – data berupa kata akan dilakukan pembobotan dengan metode Term Frequency dan Inverse Document Frequency dengan tujuan mendapatkan nilai dari kata – kata tersebut. Proses implementasi TF-IDF menggunakan TfidfVectorizer yang ada di dalam modul scikit-learn.

Sebagai contoh dilakukan terdapat beberapa kalimat dalam dokumen yang sudah dilakukan text preprocessing sebelumnya. Berikut adalah contoh pembobotan TF-IDF dengan S_i adalah kalimat (SYUKURAN, t.t.):

Tabel 1. Term Kalimat

Kalimat	Term
S1	pemerintah malaysia libur sekolah dua hari lindung anak kabut asap asal indonesia senin
S2	indeks polusi udara wilayah negeri jiran sentuh level sangat tidak sehat
S3	otoritas paksa tutup sekolah hari selasa besok
S4	shah alam selangor indeks polusi udara turun cukup tajam
S5	angka indeks aman moderat tidak sehat sangat bahaya
S6	dapat area malaysia utama wilayah utara level tidak sehat buruk penang level
S7	kabut asap paksa berapa bandara malaysia tutup minggu
S8	acara lari maraton kuala lumpur batal

7. Klasifikasi Support Vector Machine

Tahap selanjutnya adalah implementasi algoritma Support Vector Machine untuk dilakukan klasifikasi. Proses ini terbagi menjadi 3, yaitu:

a. Membangun model Support Vector Machine

Pada proses ini dilakukan pembuatan model klasifikasi Support Vector Machine. Proses implementasi SVM

menggunakan modul SVM kernel linear yang ada di scikit-learn.

b. Data training

Setelah melalui proses klasifikasi, selanjutnya dilakukan proses data training yang menggunakan perintah .fit() dan menyertakan 2 parameter yaitu fitur (x_{train}) dan label (y_{train}).

c. Data testing

Setelah melalui proses data training, selanjutnya proses data testing akan menggunakan perintah .predict() untuk mengetahui hasil klasifikasi.

Tahap klasifikasi menggunakan metode Support Vector Machine dalam mengklasifikasikan kalimat ke dalam kelas positif dan negatif. Proses klasifikasi terbagi menjadi 2 proses, yaitu training dan classifier. Data akan dimasukan dan diubah menjadi data vektor. Format data input untuk klasifikasi SVM adalah 1 1:0.39794. Data input yang pertama 1 atau -1 menyatakan dua label awalan yang diberikan. Angka 1 pada karakter pertama menyatakan data tersebut masuk ke label positif. Angka 1 berikutnya sebelum tanda “:” menyatakan indeks dan angka 0.39794 menyatakan bobot kata. Sebagai contoh perubahan kalimat menjadi vektor ditunjukkan pada tabel 2.

Tabel 2. Pengubahan data kalimat menjadi format data vektor

Kalimat	pemerintah malaysia libur sekolah dua hari lindung anak kabut asap asal indonesia senin
Vektor	[1 1:0.903 2:0.426 3:0.903 4:0.602 5:0.903 6:0.602 7:0.903 8:0.903 9:0.602 10:0.602 11:0.903 12:0.903 13:0.903

Jika dilakukan terhadap seluruh kalimat yang ada, maka kalimat akan terwakili oleh format data vektor. Pada proses pelatihan dibutuhkan label positif dan negatif, Label positif +1 artinya kalimat tersebut berpengaruh besar terhadap hasil ringkasan, sedangkan untuk kelas negatif -1 artinya kalimat tersebut tidak

berpengaruh besar. Maka setelah melewati perhitungan pembobotan TF-IDF dan

cosine similarity sebelumnya, maka didapatkan hasil sebagai berikut:

Tabel 3. Proses SVM Training

Kalimat	Term	Label
S1	pemerintah malaysia libur sekolah dua hari lindung anak kabut asap asal indonesia senin	1
S2	indeks polusi udara wilayah negeri jiran sentuh level sangat tidak sehat	1
S3	otoritas paksa tutup sekolah hari selasa besok	1
S4	shah alam selangor indeks polusi udara turun cukup tajam	1
S5	angka indeks aman moderat tidak sehat sangat bahaya	1
S6	dapat area malaysia utama wilayah utara level tidak sehat buruk penang level	1
S7	kabut asap paksa berapa bandara malaysia tutup minggu	1
S8	acara lari maraton kuala lumpur batal	1

Setelah mendapatkan nilai label dari seluruh kalimat, selanjutnya masuk ke tahap perhitungan $x_i x_j^t$ dimana $i, j = 1, \dots, n$. Untuk x_1 adalah nilai keseluruhan dari kolom S1, $x_2 = S2$ sampai $x_8 = S8$. Sehingga setiap kalimat dapat diubah sesuai dengan hasil TF-IDF menjadi bentuk sebagai berikut, sebagai contoh bentuk pada kalimat pertama:

$$x_1 = [0.903, 0.426, 0.903, 0.602, 0.903, \\ 0.602, 0.903, 0.903, 0.602, 0.602, 0.903, \\ 0.903, 0.903, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \\ 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, \\ 0, 0, 0, 0, 0, 0, 0, 0]$$

Kemudian data vektor pada setiap kalimat sesuai dengan pembobotan tf-idf pada setiap kalimatnya. Dilakukan perhitungan $x_1 x_j^t$ pada setiap kalimat, sebagai contoh dihitung pada kalimat pertama:

[illegible]

$$= 0.816 + 0.181 + 0.816 + 0.362 + 0.816 + 0.362 + 0.816 + 0.816 + 0.362 + 0.362 + 0.816 + 0.816 + 0.816$$
$$= 8.156$$

Cara yang sama diterapkan terhadap seluruh kalimat, maka matriks yang terbentuk dari hasil perhitungan $x_1 x_j^t$ adalah sebagai berikut:

$$= \begin{bmatrix} x_1 x_1^T & 0 & 0.725 & 0 & 0 & 0.181 & 0.906 & 0 \\ 0 & 4.622 & 0 & 0.906 & 0.907 & 1.088 & 0 & 0 \\ 0.725 & 0 & 3.987 & 0 & 0 & 0 & 0.725 & 0 \\ 0 & 0.906 & 0 & 5.8 & 0.181 & 0 & 0 & 0 \\ 0 & 0.907 & 0 & 0.181 & 4.169 & 0.363 & 0 & 0 \\ 0.181 & 1.088 & 0 & 0 & 0.363 & 6.526 & 0.181 & 0 \\ 0.906 & 0 & 0.725 & 0 & 0 & 0.181 & 4.078 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4.893 \end{bmatrix}$$

Tahap selanjutnya adalah melakukan perhitungan $y_i y_j^T$ dimana $i, j = 1, \dots, n$. Untuk nilai y adalah nilai dari label yang diberikan. Sehingga dihasilkan matriks dari hasil perhitungan $y_i y_j^T$ ialah:

[illegible]

Kemudian tahap selanjutnya adalah mencari nilai a_1 , proses mendapatkan nilai a_1 diawali dengan mengubah setiap kalimat menjadi nilai

vektor. Setelah itu, kita menghitung nilai x dan y pada setiap kalimat.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan data

Data yang telah dikumpulkan melalui tahapan pengumpulan data, seluruhnya di import ke dalam notebook. Data yang diambil hanya berupa komentar pengguna dari media sosial TikTok. Data yang terkumpul dari proses ini adalah sebanyak 5610 baris dengan satu kolom atribut/fitur. Berikut merupakan tampilan data yang sudah di import ke dalam notebook.

	Comment	CS
0	A gw kira bumil anjir @j12345	0
1	Ga bisa mas, bisanya ngambang. Soalnya punya 2...	0
2	Soalnya keberatan beban depan bang,, makanya g...	0
3	woi anjing ga gitu juga🐶	0
4	Loh bukan? Kok gede perutnya!	0
...
5605	yayyyy congratsss🥳🥳	1
5606	Selamat yaaa	1
5607	diam diam mematikan ya	0
5608	Bismillah semoga bisa dpat lain kali🥳🥳🥳	1
5609	janji gak halangin rejeki saudara sendiri	0

Gambar 3. Hasil kumpulan data mentah

3.2 Hasil Data Labeling

Seluruh data yang sudah tersimpan kemudian diberikan label sebagai klasifikasi sentimen yang bersifat positif dan negatif. Disini penulis tidak mengumpulkan komentar yang bersifat netral. Pelabelan data dilakukan secara manual dengan menggunakan aplikasi Microsoft Excel. Berdasarkan hasil dari pelabelan data, didapatkan hasil berupa 4098 komentar yang bersifat negatif dan 1512 komentar yang bersifat positif.

```
0    4098
1    1512
Name: CS, dtype: int64
```

```
0 = Komentar Negatif
1 = Komentar Positif
```

Gambar 4. Hasil pelabelan komentar

3.3 Hasil Text Preprocessing

Setelah tahap pelabelan data, seluruh data berupa komentar akan melalui tahap text preprocessing yang bertujuan mengubah komentar yang sebelumnya masih bersifat data mentah menjadi data yang bisa dibaca oleh machine learning. Tahap text preprocessing sendiri terbagi menjadi beberapa tahap, hasil dari tahap-tahap text preprocessing dapat dilihat sebagai berikut:

1. Hasil Case Folding

Data akan melewati beberapa proses, proses pertama yaitu case folding. Disini merupakan proses untuk menyamakan semua huruf yang ada di dalam dataset. Seluruh huruf akan diubah menjadi huruf kecil.

```
0          a gw kira bumil anjir @j12345
1    ga bisa mas, bisanya ngambang. soalnya punya 2...
2    soalnya keberatan beban depan bang,, makanya g...
3          woi anjing ga gitu juga🐶
4          loh bukan? kok gede perutnya!

...

5605          yayyyy congratsss🥳🥳
5606          selamat yaaa
5607          diam diam mematikan ya
5608          bismillah semoga bisa dpat lain kali🥳🥳🥳
5609          janji gak halangin rejeki saudara sendiri
Name: Comment, Length: 5610, dtype: object
```

Gambar 5. Hasil case folding

Dilihat pada Gambar 5 bahwa di dalam dataset komentar masih ada variabel yang bersifat angka, emoticon, tanda baca, simbol, dan single char. Jadi, dilakukan kembali proses untuk menghilangkan semua variabel tersebut.

```
Remove Special Comments

0          a gw kira bumil anjir
1    ga bisa mas, bisanya ngambang. soalnya punya 2...
2    soalnya keberatan beban depan bang,, makanya g...
3          woi anjing ga gitu juga?
4          loh bukan? kok gede perutnya!

...

5605          yayyyy congratsss??
5606          selamat yaaa
5607          diam diam mematikan ya
5608          bismillah semoga bisa dpat lain kali ???
5609          janji gak halangin rejeki saudara sendiri
Name: Comment, Length: 5610, dtype: object
```

Gambar 6. Hasil menghilangkan special comments

Dapat dilihat pada Gambar 6 bahwa variabel special comments seperti

emoticon, link, mention, URL, tab, new line semuanya sudah terhapus dari dataset.

```
Remove number
0          a gw kira bumil anjir
1  ga bisa mas, bisanya ngambang. soalnya punya ...
2  soalnya keberatan beban depan bang,, makanya g...
3          woi anjing ga gitu juga?
4          loh bukan? kok gede perutnya!

...

5605          yayyyy congratsss??
5606          selamat yaaa
5607          diam diam mematkan ya
5608          bismillah semoga bisa dpat lain kali ???
5609          janji gak halangin rejeki saudara sendiri
Name: Comment, Length: 5610, dtype: object
```

Gambar 7. Hasil menghilangkan variabel angka

Dapat dilihat pada Gambar 7 bahwa variabel angka semuanya sudah terhapus dari dataset.

```
Remove Punctuation
0          a gw kira bumil anjir
1  ga bisa mas bisanya ngambang soalnya punya pe...
2  soalnya keberatan beban depan bang makanya gak...
3          woi anjing ga gitu juga
4          loh bukan? kok gede perutnya

...

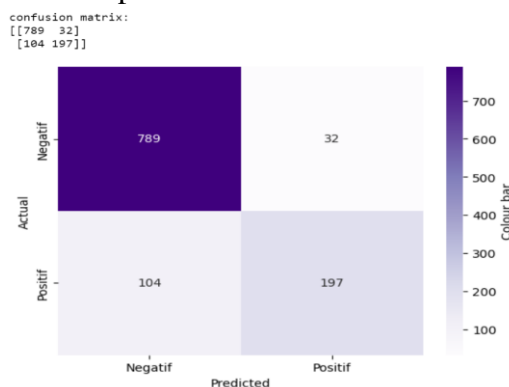
5605          yayyyy congratsss
5606          selamat yaaa
5607          diam diam mematkan ya
5608          bismillah semoga bisa dpat lain kali
5609          janji gak halangin rejeki saudara sendiri
Name: Comment, Length: 5610, dtype: object
```

Gambar 8 Hasil menghilangkan tanda baca

Dapat dilihat pada Gambar 8 bahwa variabel tanda baca seperti tanda tanya, tanda seru dll sudah terhapus dari dataset.

3.4 Hasil Penelitian

Sebelum dilakukan proses klasifikasi, data akan dibagi menjadi dua yaitu data uji dan data latih. Dengan persentase data latih sebesar 80% dan data uji sebesar 20%. Training model proses klasifikasi menggunakan modul SVM dengan kernel linear yang ada di scikit.learn. Berikut tampilan dari confusion matrix:



Gambar 9. Confusion Matrix

Pada Gambar 9, dapat dilihat bahwa confusion matrix menunjukkan hasil dari algoritma *Support Vector Machine* berhasil mengklasifikasikan 789 data sebagai negatif dengan benar, tapi terdapat 104 data misklasifikasi. Dan juga berhasil mengklasifikasikan 197 data sebagai positif dengan benar, namun terdapat 32 data misklasifikasi.

3.5 Pembahasan

Hasil yang didapatkan melalui Gambar 4.11, kemudian dilakukan evaluasi penelitian dengan menghitung nilai akurasi keseluruhan, recall, precision, dan f1-score. Berikut merupakan hasil perhitungannya:

Akurasi Keseluruhan

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{197+789}{197+789+32+104} = 0,88 \times 100 = 88\%$$

Precision, Recall, f1-score

1. Negatif (Label = 0)

$$Precision = \frac{TN}{TN+FN} = \frac{789}{789+104} = 0,88 \times 100 = 88\%$$

$$Recall = \frac{TN}{TN+FP} = \frac{789}{789+32} = 0,96 \times 100 = 96\%$$

$$F1-score = \frac{2(Precision \times Recall)}{Precision+Recall} = \frac{2(0,88 \times 0,96)}{0,88+0,96} = 0,92 \times 100 = 92\%$$

2. Positif (Label = 1)

$$Precision = \frac{TP}{TP+FP} = \frac{197}{197+32} = 0,86 \times 100 = 86\%$$

$$Recall = \frac{TP}{TP+FN} = \frac{197}{197+104} = 0,65 \times 100 = 65\%$$

$$F1-score = \frac{2(Precision \times Recall)}{Precision+Recall} = \frac{2(0,86 \times 0,65)}{0,86+0,65} = 0,74 \times 100 = 74\%$$

Tabel 4. Rangkuman perhitungan Confusion Matrix

Confusion Matrix				
Kelas	Precision	Recall	F1-score	Akurasi Keseluruhan (Accuracy)
Negatif	88%	96%	92%	88%
Positif	86%	65%	74%	

	precision	recall	f1-score	support
0	0.88	0.96	0.92	821
1	0.86	0.65	0.74	301
accuracy			0.88	1122
macro avg	0.87	0.81	0.83	1122
weighted avg	0.88	0.88	0.87	1122

Gambar 10. Hasil Perhitungan Confusion Matrix

Gambar 10 menunjukkan hasil perhitungan yang dilakukan oleh machine learning dari hasil confusion matrix sebelumnya. Didapatkan hasil dari nilai precision, recall, f1-score dan akurasi

keseluruhan dari kedua kelas adalah sama dengan hasil evaluasi performa algoritma yang dihitung secara manual dengan rumus.

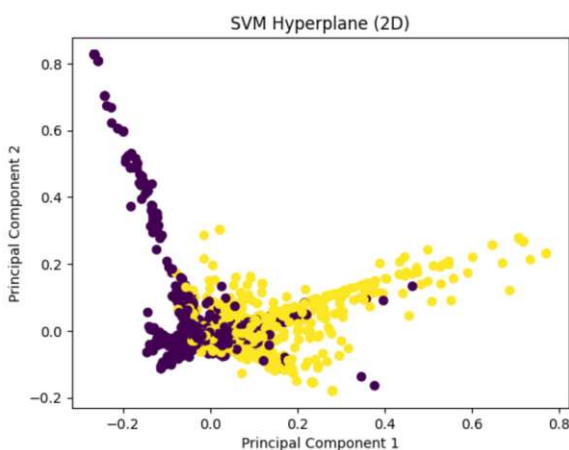
Dapat dilihat dari hasil diatas, hasil dari recall positif sebesar 65% yakni merupakan hasil yang kurang baik. Maka, penulis menerapkan metode hyperparameter search dengan menggunakan *GridSearch*. Dengan tujuan, untuk mendapatkan hasil keseluruhan yang lebih baik, dengan mencoba kernel lain seperti Linear, Sigmoid, Polynomial, dan RBF. Berikut merupakan hasil dari hyperparameter search yang dilakukan:

Parameter Terbaik : {'C': 10, 'gamma': 0.1, 'kernel': 'rbf'}
Grid Best Score : 0.8643048128342246

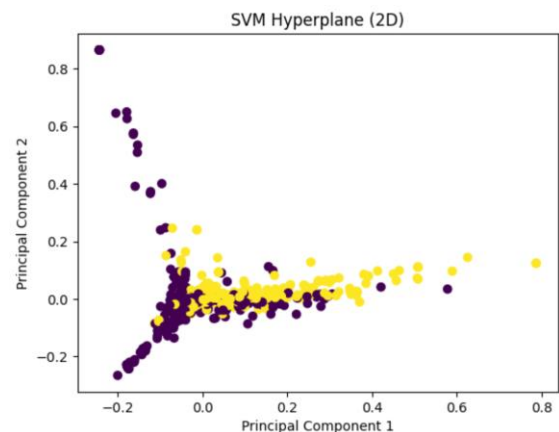
Gambar 11. Hasil Hyperparameter Search

Dari hasil hyperparameter search diatas, untuk mendapatkan hasil terbaik maka pada penelitian berikutnya disarankan dapat menggunakan nilai C sebesar 10, nilai gamma sebesar 0,1 dan kernel RBF.

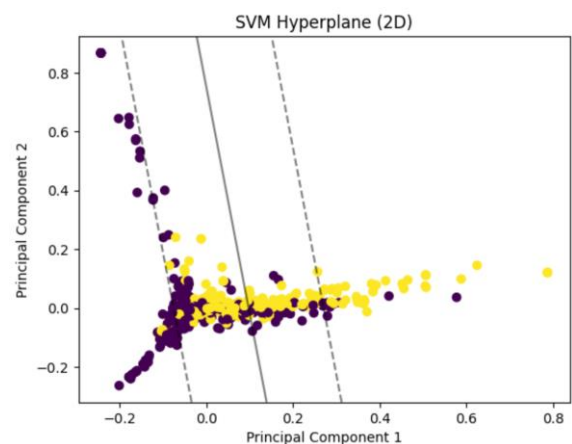
Berikut ini merupakan tampilan scatter chart dari data training dan data testing. Dapat dilihat bahwa, pada gambar terdapat dua warna yang melambangkan posisi dari kelas data positif dan data negatif.



Gambar 12. SVM Data Training



Gambar 13. SVM Data Testing



Gambar 14. Hyperplane

Pada Gambar 14, ditunjukkan sebuah garis discrimination boundaries atau hyperplane yang membatasi kedua jenis kelas. Dengan tujuan, agar dapat dilakukan klasifikasi antar kelas. Penelitian kali ini, data digambarkan ke dalam bentuk 2D dengan menggunakan kernel linear dari metode *Support Vector Machine*.

4. PENUTUP

4.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat diperoleh kesimpulan bahwa teknik *scraping tools* dapat digunakan untuk mengambil komentar – komentar yang ada pada video TikTok, dengan tujuan untuk mendeteksi komentar yang memiliki unsur *cyberbullying*. Pada penelitian ini, variabel data yang diperlukan hanya berupa komentar. Jumlah data yang berhasil didapatkan adalah sebanyak 5610 baris komentar. Metode *Term Frequency – Inverse Document Frequency* (TF-IDF) yang digunakan untuk melakukan pembobotan kata, dan *Support Vector Machine* yang digunakan sebagai metode untuk mendeteksi komentar *cyberbullying*. Dengan pembagian dataset menjadi 80% data training yang berjumlah 4488 komentar dan 20% data testing berjumlah 1122 komentar. Dapat menghasilkan nilai akurasi sebesar 88%. Hasil akurasi keseluruhan yang diperoleh algoritma *Support Vector Machine* dalam klasifikasi komentar *cyberbullying* adalah sebesar 88%. Diikuti dengan hasil *precision* sebesar 88% (Negatif) dan 86% (Positif). Lalu, hasil recall sebesar 96% (Negatif) dan 65% (Positif). Kemudian, hasil f1-score yang diperoleh adalah sebesar 92% (Negatif) dan 74% (Positif).

4.2. Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk membantu dan mendukung penelitian selanjutnya bahwa penelitian ini hanya menggunakan 2 kategori, yaitu negatif yang artinya memiliki unsur *cyberbullying* dan positif yang artinya memiliki unsur penyemangat atau support. Sehingga, pada

penelitian selanjutnya, disarankan menambahkan beberapa kategori lagi seperti netral yang artinya tidak memiliki unsur *cyberbullying* dan kritik yang artinya komentar yang bersifat kritis.

5. DAFTAR PUSTAKA

- Dimas Bayu (2022) *APJII: Pengguna Internet Indonesia Tembus 210 Juta pada 2022*, <https://dataindonesia.id/digital/detail/apjii-pengguna-internet-indonesia-tembus-210-juta-pada-2022>.
- Juditha, C. (2021) 'ANALISIS KONTEN TENTANG PERUNDUNGAN MAYA TERHADAP SELEBRITI DI INSTAGRAM ANALYSIS OF CONTENT THE CASE OF CYBERBULLYING AGAINST CELEBRITIES ON INSTAGRAM', *Jurnal Penelitian Komunikasi dan Opini Publik Vol*, 25(2), pp. 183–198.
- Kamal, R.M. (2019) *Analisis Sentimen Cyberbullying Pada Komentar Facebook Dengan Metode Klasifikasi Support Vector Machine*. Universitas Komputer Indonesia.
- Manoppo, T.N. and Fudholi, D.H. (2021) 'Deteksi Cyberbullying berdasarkan Unsur Perbuatan Pidana yang Dilanggar dengan Naive Bayes dan Support Vector Machine', *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 5(1), pp. 10–19.
- Mitsu, R. and Dawood, E. (2022) 'Cyberbullying: An Overview', *Indonesian Journal of Global Health Research*, 4(1), pp. 195–202. Available at: <https://doi.org/10.37287/ijghr.v4i1.927>.
- Monavia Ayu Rizaty (2022) *Pengguna Tiktok Indonesia Terbesar Kedua di Dunia*, <https://dataindonesia.id/digital/detail/pengguna-tiktok-indonesia-terbesar-kedua-di-dunia>.
- Pardede, J., Miftahuddin, Y. and Kahar, W. (2020) 'Deteksi Komentar Cyberbullying Pada Media Sosial Berbahasa Inggris Menggunakan Naive Bayes Classification', *JURNAL INFORMATIKA*, 7(1). Available at:

<http://ejournal.bsi.ac.id/ejurnal/index.php/ji>.

Puspitarini, D.S. and Nuraeni, R. (2019) 'Pemanfaatan media sosial sebagai media promosi', *Jurnal Common*, 3(1), pp. 71–80.

Kaiser, S. and Ali, R. (2018) 'Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents', *International Journal of Computer Applications*,

181(1), pp. 25–29. Available at: <https://doi.org/10.5120/ijca2018917395>.

Ritonga, A.S. and Purwaningsih, E.S. (2018) 'Penerapan Metode Support Vector Machine (SVM) Dalam Klasifikasi Kualitas Pengelasan Smaw (Shield Metal Arc Welding)', *Jurnal Ilmiah Educ: Pendidikan dan Informatika*, 5(1), pp. 17–25.