

Apakah *Youtuber* Indonesia Kena *Bully* Netizen?

Joviano Siahaan¹, Wella², Ririn I. Desanti³

^{1,2,3} Program Studi Sistem Informasi, Fakultas Teknik dan Informatika,
Universitas Multimedia Nusantara, Tangerang, Indonesia
² wella@umn.ac.id

Diterima 22 September 2020

Disetujui 18 November 2020

Abstract—This study will examine the cyberbullying phenomenon that was experienced by Indonesian Youtubers in their Instagram comment section. Cyberbullying is the use of electronic communication to bully a person, typically by sending messages of an intimidating or threatening nature. Youtubers are the subject of this research due to their massive following, who constantly responds to every content posted on their Instagram page. The algorithm chosen to conduct this sentiment analysis was Support Vector Machine (SVM) due to their high accuracy percentage. The data used in this analysis was retrieved from 10 Indonesian Youtuber Instagram accounts. In order to analyze this data, several steps were done including text mining, data cleansing, data modeling and applying model to test data. The result of analysis using an SVM model with an accuracy of 81.2% is 49.524% of comments on an Indonesian Youtuber comment section are considered as cyberbullying.

Index Terms—cyberbullying, Instagram, sentiment analysis, Support Vector Machine

I. PENDAHULUAN

Seiring berkembangnya zaman, semakin banyak orang yang menghabiskan harinya menggunakan media sosial. Anak berumur 12-17 memiliki aktivitas di dunia maya [1]. Media sosial dapat memberikan berbagai dampak positif seperti mempermudah komunikasi dengan orang lain, namun media sosial juga dapat memberikan berbagai dampak negatif, salah satunya adalah *cyberbullying*. *Cyberbullying* dapat didefinisikan sebagai penggunaan teknologi untuk melecehkan, mengancam, mempermalukan atau mengincar seseorang [2].

Menurut hasil riset Polling Indonesia yang bekerja sama dengan Asosiasi Penyelenggara Jasa Internet Indonesia, terdapat sekitar 49% warga Indonesia yang pernah menjadi sasaran dari bullying di media sosial [3]. *Cyberbullying* memberikan dampak yang begitu besar terhadap individu maupun kelompok. *Bullying* dapat menjadi suatu permasalahan yang mempengaruhi kesehatan mental maupun fisik dari korbannya. Efek dari *cyberbullying* adalah menurunnya kepercayaan diri, rasa gelisah, keinginan mengubah diri, balas dendam, menyakiti diri sendiri

dan keinginan untuk berhenti menggunakan media sosial[4]. Instagram adalah sebuah media sosial dimana adanya sebuah nilai yang mengukur tingkat popularitas seseorang, yaitu *followers*. Dalam sebuah akun terkenal, akun tersebut akan mendapat banyak komentar terhadap konten yang diunggah. Komentar yang ditinggalkan pada akun-akun terkenal tersebut, dapat dijadikan data untuk menganalisa seberapa sehat, atau tidak sehatnya lingkungan internet di Indonesia.

Fenomena tersebut dapat dianalisa menggunakan sentimen analisis untuk mengetahui apakah sebuah komentar mengandung unsur *cyberbullying*. Analisa sentimen merupakan masalah *text classification* yang dapat diselesaikan menggunakan algoritma-algoritma seperti *Naive Bayes* dan *Support Vector Machine* [5]. Algoritma tersebut akan mengklasifikasikan kata berdasarkan *data training* yang telah dibuat kedalam 2 *class*.

Tujuan dari analisa sentimen ini adalah mengetahui seberapa banyak *cyberbullying* yang diterima oleh *Youtuber* Indonesia di kolom komentar akun Instagram-nya. Data yang didapat dari kolom komentar tersebut akan diproses menggunakan algoritma *Support Vector Machine* yang mengklasifikasikan komentar menjadi positif atau negatif.

II. TINJAUAN PUSTAKA

A. *Cyberbullying*

'Bullying' dapat didefinisikan sebagai aksi atau perlakuan agresif yang dilakukan secara sengaja oleh seorang individu kepada seseorang yang tidak dapat membela dirinya sendiri [6]. *Cyberbullying* merupakan perluasan dari bullying dimana tindakan tersebut dilakukan menggunakan teknologi baru seperti website, text messages, social media dll. *Cyberbullying* dapat diklasifikasikan menjadi beberapa tipe, antara lain [1]:

- *Harrasment*: Mengirim pesan yang menghina atau melecehkan secara berulang-ulang.

- *Denigration*: Mendistribusikan informasi yang merendahkan orang lain melalui situs *website*, *e-mail* atau *text messaging*.
- *Flaming*: Berkelahi secara *online* dengan menggunakan huruf kapital, gambar yang tidak senonoh dan simbol-simbol untuk menambahkan emosi kedalam suatu argumen.
- *Impersonation*: Membobol *e-mail* atau akun media sosial seseorang untuk mengirim pesan atau gambar yang tidak senonoh.
- *Masquerading*: Menggunakan identitas seseorang untuk mengirim pesan negatif ke orang lain.
- *Pseudonym*: Melecehkan atau menghina dengan menggunakan alias atau *nickname* untuk merahasiakan identitas mereka.
- *Outing and Trickery*: Menyebarkan rahasia seseorang dengan cara *forwarding* atau membagikan pesan.
- *Cyber Stalking*: Merupakan bagian dari *harassment*. Melakukan tindakan berulang-ulang kali secara *online* yang membuat korban takut akan keamanannya.

B. Instagram

Instagram adalah sebuah aplikasi berbagi foto dan video yang memungkinkan pengguna mengambil foto, video, menerapkan filter digital, dan membagikannya ke berbagai layanan jejaring sosial, termasuk milik Instagram sendiri [7].

Di zaman dimana penggunaan sosial media sudah tersebar luas, mempunyai akun Instagram merupakan salah satu kewajiban bagi seorang *public figure*. Sosial media menjadi suatu sarana bagi seorang *public figure* untuk berinteraksi dengan penggemar-pengemarnya secara langsung.

C. Sentiment Analysis

Sentiment analysis memiliki tujuan untuk menganalisa opini, pendapat atau emosi seseorang terhadap suatu produk, jasa atau isu tertentu. Opini mempunyai pengaruh yang besar dalam aktivitas kehidupan manusia. Di saat kita ingin membuat keputusan, kita berharap mendapatkan opini positif dari orang lain.

Sebelum zaman teknologi, opini publik didapatkan melalui survei, jajak pendapat atau *focus groups*. Dengan perkembangan sosial media yang semakin pesat, penggunaan data-data opini dari sosial media dalam pengambilan keputusan semakin meningkat. Pengetahuan mengenai sebuah produk sudah tidak lagi sebatas opini seorang teman, melainkan dari semua orang yang pernah mengunggah opini mereka di jejaring sosial.

Dalam beberapa tahun terakhir, terdapat banyak kasus dimana opini-opini yang didapatkan dari pos sosial media memiliki pengaruh yang signifikan terhadap sistem sosial dan politik. Hal tersebut mengakibatkan meluasnya pengaruh *sentiment analysis* ke berbagai bidang, seperti produk, jasa, *healthcare* hingga *event* sosial dan pemilu politik.

Klasifikasi sentimen biasanya dirumuskan sebagai *two-class classification problem*, positif dan negatif. Klasifikasi tersebut dapat dilakukan menggunakan algoritma seperti *Naïve Bayes* dan *Support Vector Machine* (SVM) [5].

III. METODOLOGI PENELITIAN

A. Data Collection

Data yang digunakan dalam penelitian ini akan didapatkan dari komentar Instagram *YouTuber* populer menggunakan algoritma *comment scraper* [8]. Jumlah komentar dari masing-masing *YouTuber* adalah 1200 komentar (masing-masing 200 komentar dari 6 pos terakhir). Sepuluh akun Instagram dengan *subscriber* YouTube terbanyak didapatkan dari website *hypeauditor.com* pada tanggal 24 Agustus 2020 pukul 22.00 WIB [9].

Berikut adalah daftar *YouTuber* yang merupakan subjek dari analisa:

- Atta Halilintar (@atahalilintar)
- Ria Ricis (@riaricis1795)
- Jess No Limit (@jessnolimit)
- Raffi Ahmad (@raffinagita1717)
- Gen Halilintar (@genhalilintar)
- Baim Wong (@baimwong)
- Frost Diamond (@frostdiamondd)
- Naisa Alifia Yuriza (@nasiaalifiayuriza)
- Saai Halilintar (@saaihalilintar)
- MiawAug (@miawaug)

B. Preprocessing

Beberapa proses perlu dilakukan untuk mempersiapkan data untuk proses klasifikasi [10].

- *Combining Data*: Penggabungan data menjadi 1 document dengan format .csv.
- *Cleansing Data*: Proses penghapusan unsur-unsur seperti emoji, mention dan hashtag menggunakan *RegEx*.
- *Tokenize*: Komentar akan dipecahkan menjadi 1 kata berdasarkan spasi di setiap kalimat.
- *Transform Cases*: Perubahan semua huruf

menjadi *lowercase*.

- Filter *Stopwords*: Proses menghilangkan kata-kata yang tidak memiliki konotasi negatif atau positif menggunakan dataset *stopwords* [11].
- Filter *Tokens (by length)*: Proses menghilangkan token-token dengan huruf kurang dari 2.
- *Generate Bigram*: Proses menggabungkan 2 kata (*bigram*) menjadi 1 token untuk meningkatkan konteks dalam sebuah kalimat dan akurasi model [12]. Contoh kalimat “Saya tidak mau pergi” setelah melalui proses bigram menjadi. saya, saya_tidak, tidak, tidak_mau, mau, mau_pergi dan pergi.

C. Support Vector Machine Model

Sentimen analisis dilakukan dengan pembuatan model berdasarkan *labelled data* yang akan dijadikan acuan untuk memprediksi *unlabelled data*. Dalam sentimen analisis, diperlukan *labelled data* yang sesuai dengan konteks *unlabelled data*, dikarenakan dalam konteks yang berbeda, sebuah kata dapat menghasilkan makna yang berbeda.

Model SVM dibuat dari gabungan 2 dataset yaitu, *id-multi-label-hate-speech-and-abusive-language-detection* [13] dan *Indonesian-Twitter-Emotion-Dataset* [14], dimana dataset tersebut memiliki 2 kolom yaitu *text* dan *label*. Label tersebut diberikan untuk mengklasifikasikan *tweet* menjadi positif (tidak mengandung *cyberbullying*) dan negatif (mengandung *cyberbullying*). Dataset terdiri dari 15,830 *tweet* (8,390 negatif dan 7,440 positif).

Dataset tersebut akan diproses menggunakan algoritma *Support Vector Machine* (SVM) untuk menghasilkan bobot negatif atau positif dari masing-masing kata. Hasil dari algoritma SVM adalah *decision* dan *confidence*, dimana *decision* adalah jarak sebuah *text* dari *hyperplane*, sebuah garis yang dibuat untuk memisahkan kedua *class*. Sedangkan, *confidence* adalah probabilitas suatu kata diletakkan dalam *class* tertentu. Jika sebuah kata memiliki probabilitas lebih untuk masuk kedalam *class* negatif, maka model tersebut memprediksi bahwa kata tersebut mengandung unsur *cyberbullying*.

Untuk membuktikan akurasi model, dilakukan proses *cross validation* dimana data dibagi menjadi data *training* dan data *test* dengan bobot 80% dan 20%. Model tersebut menghasilkan akurasi 81.2%. Perbandingan akurasi model SVM berdasarkan metode *Cross Validation* (80% *Training* dan 20% *Test*) dan metode manual pada data komentar YouTube dengan sampel terbanyak (@jessnolimit).

Tabel 1. Perbandingan akurasi model SVM

Metode	Akurasi
Cross Validation	81.2%
Manual	71.93%

IV. HASIL DAN PEMBAHASAN

Hasil *mining* data dari berbagai akun Instagram *Youtuber* akan diproses menggunakan model SVM yang telah dibuat.

A. Hasil Analisa Sentimen

Tabel 2 merupakan hasil analisa yang didapat dari sentiment media social Instagram per *Youtuber*. Dari ± 200.000 komentar yang diambil per *Youtuber*, dilakukan pembersihan data, maka didapat ± 1.000 komentar yang dijadikan sebagai bahan analisa. Pelabelan dilakukan secara otomatis menggunakan SVM. Rincian hasil analisa dapat dilihat pada Tabel 2.

Tabel 2. Hasil analisa sentimen

User	Total Komentar	Positif	Negatif
@attahalilintar	1.001	515	486
@baimwong	974	591	383
@frostdiamondd	1.038	386	652
@genhalilintar	1.045	507	538
@jessnolimit	1.122	548	574
@miawaug	1.038	558	480
@naisaalifiayuriza	1.013	505	508
@raffinagita1717	930	492	438
@riaricis1795	1.006	587	419
@saaihalilintar	988	427	561

B. Persentase Cyberbullying dalam Komentar Instagram

Setelah berhasil melakukan pelabelan, berikut besaran persentase *cyberbullying* yang didapat oleh para *Youtuber*.

Tabel 3. Persentase *cyberbullying* dalam komentar Instagram

User	Persentase
@attahalilintar	48.55%
@baimwong	39.32%
@frostdiamondd	62.81%
@genhalilintar	51.48%
@jessnolimit	51.16%
@miawaug	46.24%
@naisaalifiayuriza	50.15%

@raffinagita1717	47.10%
@riaricis1795	41.65%
@saaihilintar	56.78%
<i>Average</i>	49.524%

Atta Halilintar mengisi akun Instagram-nya dengan konten-konten seperti kegiatan olahraga, *event* dan outfit hariannya. Komentar positif datang dari user yang memuji Atta akan aktivitas-aktivitas positif yang ia lakukan, namun terdapat juga komentar negatif yang menganggap aktivitas-aktivitas tersebut tidak penting. Terdapat juga komentar negatif dimana *user-user* menjelek-jelekkan penampilan Atta.

Sama halnya dengan Baim Wong, Baim Wong memiliki persentase *cyberbullying* terendah dikarenakan konten yang sangat positif. Konten yang diunggah adalah kegiatan sehari-hari keluarga, yang dimana menghasilkan komentar-komentar positif yang memuji keharmonisan keluarga tersebut. Terdapat banyak juga yang memuji tindakan Baim dalam melakukan *giveaway* dan kegiatan amal.

Walaupun konten frostdiamondd tergolong positif, frostdiamondd memiliki tingkat *cyberbullying* tertinggi. Komentar-komentar negatif datang dari *user* yang menjelek-jelekkan penampilan dan umur dari frostdiamondd, dengan kata-kata seperti “bocah” atau “bocil”.

Akun Instagram Gen Halilintar (@genhalilintar) diisi dengan konten sehari-hari keluarga Halilintar, beserta beberapa video yang memperlihatkan kemampuan mereka bernyanyi. Komentar positif datang dari *user-user* yang memuji penampilan keluarga tersebut.

Komentar positif dalam Instagram jessnolimit (@jessnolimit) datang dari *user-user* yang memuji keterampilan jessnolimit memainkan video *game*. Sedangkan, *cyberbullying* yang ia terima mengandung unsur-unsur SARA dan menjelek-jelekkan penampilan.

MiawAug mengisi Instagram-nya dengan konten video *game* dan *giveaway* gadget. Walau tetap tergolong *cyberbullying*, komentar negatif yang diterima MiawAug adalah kekecewaan *user-user* yang tidak memenangi *giveaway* tersebut, bukan menjelek-jelekkan MiawAug sendiri.

Konten Instagram Naisa berisi *outfit* sehari-hari, TikTok dan *endorsement* produk. Komentar negatif dan positif datang dari *user-user* yang memuji atau menjelek-jelekkan penampilan Naisa. Terdapat juga banyak komentar yang memaksa Naisa untuk melakukan “*follow back*” atau membalas komentar tersebut.

Konten-konten Instagram keluarga Raffi Ahmad mengundang 2 tipe komentar yang bertolak belakang,

komentar positif memuji Raffi dan keluarganya atas keharmonisan dan penampilannya. Namun terdapat juga *user-user* yang mendeskripsikan konten tersebut dengan kata-kata seperti “norak” atau “sombong”.

Ria Ricis menerima banyak komentar positif yang memuji penampilannya. Persentase *cyberbullying* dalam Instagram Ria menduduki peringkat kedua terendah, di angka 41.65%.

Saaih menduduki peringkat kedua dalam peringkat persentase *cyberbullying* tertinggi di angka 56.78%. Komentar negatif menjelek-jelekkan penampilan Saaih dengan kata-kata seperti “jelek”, “bodo amat” dan “sebel”.

V. SIMPULAN

Berdasarkan analisa menggunakan model *Support Vector Machine* (SVM) dengan akurasi 81.2%, bisa disimpulkan bahwa dalam sebuah kolom komentar Instagram *YouTuber* Indonesia terdapat 49.524% komentar yang mengandung unsur *cyberbullying*, yang dapat dikatakan bahwa *YouTuber* Indonesia tidak melulu dirundung oleh para masyarakat Indonesia. Hal positif yang mereka lakukan menjadi contoh bagi para masyarakat Indonesia.

Hasil penelitian ini dapat digunakan sebagai dasar penelitian lanjutan dimana ditambahkan beberapa proses-proses untuk meningkatkan kualitas. Proses tersebut yang dapat meningkatkan hasil penelitian adalah *stemming*, sinonim kata dan penggunaan *trigram*. Dapat dilakukan juga perbandingan hasil dengan algoritma klasifikasi lain seperti *Naive Bayes*.

DAFTAR PUSTAKA

- [1] Weber, N. L., & Pelfrey, W. V. (2014). *Cyberbullying: Causes, Consequences, and Coping Strategies*. LFB Scholarly Publishing LLC. Retrieved from https://books.google.co.id/books?id=6Fa_oAEACAAJ
- [2] Chadwick, S. (n.d.). Impacts of *Cyberbullying*, Resilience in and Emotional Building Social Schools.
- [3] Asosiasi Penyelenggara Jasa Internet Indonesia. (2017). Penetrasi & Profil Perilaku Pengguna Internet Indonesia. Apjii, 51. Retrieved from <https://apjii.or.id/survei2018s/download/TK5oJYBSyd8iqH A 2eCh4FsGELm3ubj>
- [4] Alisah Lusi, M. R. (2018). Memahami Pengalaman *Cyberbullying* Pada Remaja. Studi Fenomenologis Memahami, 1(1), 1–12.
- [5] Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions. Sentiment analysis: Mining Opinions, Sentiments, and Emotions*, May, 1–367. <https://doi.org/10.1017/CBO9781139084789>
- [6] Chadwick, S. (n.d.). Impacts of *Cyberbullying*, Resilience in and Emotional Building Social Schools.
- [7] Instagram. (n.d.). <https://id.wikipedia.org/wiki/Instagram>
- [8] Maulana, A. (n.d.). Instagram Comments Scraper. <https://github.com/AgiMaulana/Instagram-Comments-Scraper>
- [9] Hypeauditor. (n.d.). Top Most Subscribed YouTube Channels In Indonesia. Retrieved August 24, 2020, from

- <https://hypeauditor.com/top-youtube-all-indonesia>
- [10] Naf'an, M. Z., Bimantara, A. A., Larasati, A., Risondang, E. M., & Nugraha, N. A. S. (2019). *Sentiment analysis of Cyberbullying* on Instagram User Comments. *Journal of Data Science and Its Applications*, 2(1), 88–98. <https://doi.org/10.21108/jdsa.2019.2.20>
- [11] anpandu. (n.d.). nolimit kamus. <https://github.com/nolimitid/nolimit-kamus>
- [12] Tripathy, A., Agrawal, A., & Rath, S. K. (2016). Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications*, 57(October 2017), 117–126. <https://doi.org/10.1016/j.eswa.2016.03.028>
- [13] Meisaputri21. (2019). Indonesian-Twitter-Emotion-Dataset. <https://github.com/meisaputri21/Indonesian-Twitter-Emotion-Dataset>
- [14] Okkyibrohim. (2019). id-multi-label-hate-speech-and-abusive-language-detection. <https://github.com/okkyibrohim/id-multi-label-hate-speech-and-abusive-language-detection>.

