

# Klasterisasi Komentar Cyberbullying Masyarakat di Instagram berdasarkan K- Means Clustering

*by* Viry\_1 1201\_cek1

---

**Submission date:** 03-Aug-2023 11:25AM (UTC+0700)

**Submission ID:** 2140691329

**File name:** 846-Article\_Text-3118-1-15-20230612.docx (159.13K)

**Word count:** 2782

**Character count:** 18522

4

**J-INTECH (Journal of Information and Technology)**

Terakreditasi Kemendikbud SK No. 204/E/KPT/2022

E-ISSN: 2580-720X || P-ISSN: 2303-1425

**J-INTECH**  
Journal of Information and Technology

2

**Klasterisasi Komentar Cyberbullying Masyarakat di Instagram berdasarkan K-Means Clustering**Viry Puspaning Ramadhan<sup>1\*</sup>, Giasinta Mareskoti Namung<sup>2</sup><sup>1</sup>Universitas Merdeka Malang, Fakultas Teknologi Informasi, Sistem Informasi, Jl. Terusan Dieng No.57-59, Pisang Candi, Kec. Sukun, Kota Malang, Jawa Timur 65146, Indonesia<sup>2</sup>Universitas Merdeka Malang, Fakultas Teknologi Informasi, Sistem Informasi, Jl. Terusan Dieng No.57-59, Pisang Candi, Kec. Sukun, Kota Malang, Jawa Timur 65146, Indonesia**Informasi Artikel**Diterima: 01-01-2022  
Direvisi: 01-01-2022  
Diterbitkan: 01-01-2022**Kata Kunci**

Sentiment Analysis, K-Means Clustering, Cyberbullying, Sosial Media

**\*Email Korespondensi:**

viry.puspaning@unmer.ac.id

**Abstrak**

Cyberbullying telah menjadi masalah serius di platform media sosial seperti Instagram. Dalam usaha untuk mengatasi masalah ini, penelitian ini bertujuan untuk mengelompokkan komentar-komentar cyberbullying yang dibuat oleh pengguna Instagram. Metode yang digunakan dalam penelitian ini adalah K-Means Clustering, yang merupakan salah satu teknik pengelompokan yang umum digunakan dalam analisis data. Data komentar yang terkumpul dari Instagram kemudian dianalisis menggunakan algoritma K-Means Clustering untuk mengidentifikasi pola dan kelompok komentar yang serupa. Temuan dari penelitian ini dapat memberikan pemahaman yang lebih baik tentang jenis dan karakteristik komentar cyberbullying yang sering muncul di Instagram. Dengan mengetahui kelompok-kelompok komentar yang serupa, langkah-langkah pencegahan dan penanggulangan dapat dirancang dengan lebih efektif. Selain itu, hasil pengelompokan juga dapat membantu dalam pengembangan algoritma deteksi otomatis untuk mengidentifikasi komentar cyberbullying di platform media sosial. Berdasarkan evaluasi yang dilakukan terhadap hasil clustering dengan nilai dari silhouette score = 0,690152 yaitu pada cluster C1 yaitu cluster negatif. Maka komentar cyberbullying yang paling dominan adalah komentar negatif.

## 1. Pendahuluan

Era kemajuan teknologi saat ini, membuat akses terhadap segala bentuk teknologi semakin mudah. Terutama kehadiran teknologi informasi yang sangat memiliki pengaruh yang besar bagi masyarakat saat ini. Kemajuan teknologi tersebut yang membuat informasi-informasi yang ada dapat menyebar secara luas dengan waktu yang cukup singkat, salah satunya di social media (Khaira et al., 2020). *Social media* merupakan *platform* yang menyediakan segala bentuk informasi. Selain itu social media juga dapat digunakan sebagai media untuk komunikasi serta berkolaborasi secara luas. Oleh karena itu social media bisa dikatakan sebagai fasilitator online yang dapat membangun interaksi antar pengguna. Melalui platform *social media* ini pengguna dapat dengan mudahnya mendapatkan dan menyebar luaskan informasi yg ada pada *platform* tersebut (Jardim et al., 2022)

11 Salah satu *social media* populer saat ini yaitu Instagram. Instagram adalah salah satu platform media sosial yang populer saat ini di mana pengguna dapat memposting momen dalam bentuk foto atau video, yang sering disebut sebagai "postingan". Pengguna Instagram terdiri dari segala jenis kalangan mulai dari masyarakat biasa, pejabat, bahkan selebritas. Tetapi, kemudahan dari adanya informasi yang dapat dibagikan secara singkat melalui media sosial, membuat pengguna mungkin saja untuk menyalahgunakannya. Salah satu bentuk penyalahgunaan melalui social media adalah *cyberbullying* (Method et al., n.d.)

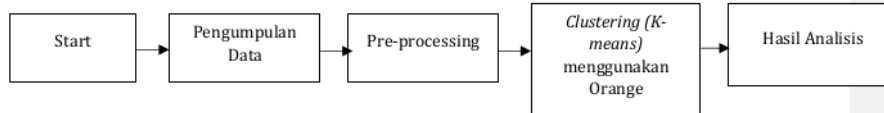
*Cyberbullying* termasuk dalam salah satu tindak kekerasan yang dilakukan oleh seorang pengguna social media terhadap korbannya di dunia maya, dimana korban dicaci, diejek, dihina, bahkan diintimidasi dan dipermalukan oleh pelaku. *Cyberbullying* sangat berdampak pada kesehatan mental seseorang, bahkan banyak korban dari *cyberbullying* yang berakhir dengan bunuh diri (Khaira et al., 2020)

6 Analisis sentimen merupakan sub bagian dari keilmuan *Natural Language Processing (NLP)*. *Natural Language Processing (NLP)* adalah bidang yang berfokus pada identifikasi perasaan yang terkandung dalam sebuah teks. Proses ini dikenal sebagai analisis sentimen, atau opinion mining, dan melibatkan ekstraksi dan pengolahan data teks secara otomatis untuk mendapatkan informasi tentang sentimen yang terkandung dalam sebuah kalimat. Dasar dari analisis sentimen adalah untuk menentukan polaritas teks dalam sebuah dokumen atau kalimat, yang dapat diklasifikasikan sebagai positif, negatif, atau netral. (Candra & Rozana, 2020)

Beberapa penelitian sebelumnya membahas tentang analisis sentiment (Syarifudin et al., 2018) opini wisata pantai di twitter menggunakan *K-Means* adalah metode yang digunakan untuk mengelompokkan opini masyarakat ke dalam kategori opini baik, buruk, dan sangat buruk. Pada penelitian tersebut penulis menggunakan metode *support vector machine* untuk pengklasifikasian opini kemudian menggunakan algoritma *k-means* untuk clusterisasi opini. Hasil dari klasifikasi tersebut mendapatkan akurasi sebesar 74,39%. Peneliti lain juga melakukan clusterisasi menggunakan *K-means* untuk (Syarifudin et al., 2018) analisis sentiment keluhan pengguna indosat sebagai masukan untuk indosat supaya memperbaiki pelayanannya.

5 Penelitian ini dilakukan untuk mengidentifikasi pola dan kelompok komentar yang serupa. Temuan dari penelitian ini dapat memberikan pemahaman yang lebih baik tentang jenis dan karakteristik komentar *cyberbullying* yang sering muncul di Instagram. Dengan mengetahui kelompok-kelompok komentar yang serupa, langkah-langkah pencegahan dan penanggulangan dapat dirancang dengan lebih efektif. Selain itu, hasil pengelompokan juga dapat membantu dalam pengembangan algoritma deteksi otomatis untuk mengidentifikasi komentar *cyberbullying* di platform media sosial.

## 2. Metode Penelitian



12

Gambar 1. Alur Penelitian

### 2.1 Pengumpulan Data

Pada tahap pengumpulan data penulis mengambil data komentar masyarakat Indonesia di Instagram Bahasa Indonesia yang ada di github (link) dimana data tersebut berisi sebanyak 400 komentar Instagram dalam bentuk csv.

### 2.2 Pre-processing

Dalam tahap pre-processing, penulis melakukan tahap awal dalam memproses data teks. Pada penelitian ini, akan dilakukan beberapa tahapan *pre-processing* seperti transformation yaitu proses mengubah data input untuk transformasi huruf kecil secara default, tokenisasi atau proses pemecahan kalimat menjadi kata per kata, *stopword* yaitu menghapus beberapa kata yang tidak penting, disini penulis menggunakan Kamus Alzami sebagai acuan penghapusan kata tidak penting.

### 2.3 Clustering menggunakan K-Means

Pada tahap klusterisasi penulis menggunakan software *orange* untuk melakukan clusterisasi dengan menentukan nilai  $k$ . Nilai  $k$  disini untuk menentukan jumlah *cluster*. Jumlah *cluster* yang digunakan adalah 2 *cluster* yaitu C1=negatif dan C2=positif. Pada *orange data mining* untuk *sentyment analysis* menggunakan algoritma *svm* dimana ada 4 *feature* yaitu *positive score*, *negative score*, *neutral score* dan *compound (combined score)*. *Compound* merepresentasikan *total sentiment* dari tweet, dimana -1 adalah paling negatif dan 1 paling positif.

Commented [AA1]: Mengapa dibagi menjadi 5 cluster? Nama clusternya apa?

3

## 3. Hasil dan Pembahasan

Dalam tulisan ini proses data yang telah diperoleh pada penelitian ini menggunakan perangkat lunak *orange data mining*, dengan langkah-langkah sebagai berikut:

### 3.1 Pengumpulan Data

Dalam penelitian ini, data yang dianalisis adalah data yang telah terkumpul komentar masyarakat Indonesia di Instagram berbahasa Indonesia. Data yang di dapatkan dari laman website Github :

<https://github.com/rizalespe/Dataset-Sentimen-Analisis-Bahasa-Indonesia>

### 3.2 Pre-Processing

Sebelum melaksanakan analisis text, untuk mengolah text dalam hal ini merupakan *preprocess text*. Text dipisahkan menjadi unit yang lebih kecil (*token*), kemudian *transformation*, *tokenization*, *normalization*, serta *filtering*. Untuk menganalisis teks menggunakan Orange Data Mining, terdapat

langkah-langkah dalam analisis yang dapat diaktifkan atau dinonaktifkan dalam widget Preprocess Text. Berikut adalah tahapan yang dilakukan dalam pra-pemrosesan teks sebelum teks dianalisis menggunakan Orange Data Mining:

### 3.2.1 Transformation

Pertama-tama, dalam pra-pemrosesan teks, langkah yang dilakukan adalah transformasi, yaitu proses mengubah data input untuk mengubah huruf menjadi kecil secara default.

### 3.2.2 Tokenization

Tokenisasi adalah metode untuk memecah teks menjadi komponen yang lebih kecil, seperti kata, kalimat, atau bigram. Ada beberapa jenis tokenisasi, termasuk Word dan Punctuation. Tokenisasi Word dan Punctuation akan membagi teks berdasarkan kata-kata dan tetap mempertahankan simbol tanda baca tanpa menghilangkannya. Sebagai contoh, "This Sample." akan dibagi menjadi (This), (Sample), (.), yang mewakili kata-kata dan tanda baca. Sementara itu, tokenisasi Tweet akan membagi teks menggunakan model pra-trained Twitter, yang mencakup hashtag, emotikon, dan simbol khusus lainnya. Contohnya, "This words. :-) #simple" akan dibagi menjadi (This), (Words), (.), (:-)), (#simple). Secara dasarnya, proses tokenisasi Word, Punctuation, dan Tweet memiliki kesamaan dalam sifat prosesnya, namun Word dan Punctuation menjadi proses utama dalam tokenisasi. Tokenisasi Word dan Punctuation juga sering digunakan dalam analisis tren.

### 3.2.3 Normalization

Langkah berikutnya adalah normalisasi, yang melibatkan stemming dan lemmatisasi teks. Teks yang telah dipisah per kata akan menghasilkan teks yang berdiri sendiri dalam sebuah kalimat. Konten status dan pendapat umumnya memiliki penulisan yang kurang sempurna, seperti kesalahan pengetikan. Dengan proses normalisasi ini, diharapkan teks tersebut dapat dikenali maknanya. Untuk melakukannya, digunakan WordNet Lemmatizer yang menggunakan jaringan sinonim kata-kata kognitif untuk menghasilkan token (kata) berdasarkan basis informasi leksikon (kamus) bahasa Indonesia yang luas dari NLTK (Natural Language Toolkit).

### 3.2.4 Filtering

Pemfilteran melibatkan penghapusan atau penyimpanan kata-kata yang dipilih. Ini mencakup *Stopwords*, yang menghapus kata-kata penutup dari teks (misalnya, menghapus 'dan', 'atau', 'di'...). Dalam penelitian ini *stopword* dilakukan dengan menggunakan Kamus Alzami. *Regex* digunakan untuk menghilangkan kata-kata yang cocok dengan ekspresi reguler tertentu, seperti `. |, |: |! |? |( ) \ | + |' |'' |'' |' |... | - | - | - $|&|>|<|/|`.

**Tabel 1.** Hasil *pre-processing*

Id	Before	After
	Geblek lo tata...cowo bgt dibela2in balikan...hadeww...ntar ditinggal lg nyalahin tuh cowo..padahal kitenya yg oon.	geblek lo tata cowo bgt dibela in balikan hadeww ntar ditinggal lg nyalahin tuh cowo padahal kitenya yg oon.
2	Kmrn termewek2 skr lengket lg duhhh kok labil bgt sih mbak kya abege ajah ato yg kmrn cari	kmrn termewek skr lengket lg duhhh kok labil bgt sih mbak kya abege ajah ato yg kmrn cari

Commented [AA2]: Buat after dan before

	sensasi biar top markotopp ertong gk berkualitas	sensasi biar top markotopp ertong gk berkualitas
3	Intinya kalau kesel dengan ATT nya, gausah ke anaknya juga. Kasian buat perkembangan psikis anak kedepannya. Itu orang bener bener tolol, skrg seandainya dia punya anak,	Intinya kalau kesel dengan att nya, gausah ke anaknya juga. Kasian buat perkembangan psikis anak kedepannya. Itu orang bener bener tolol skrg seandainya dia punya anak
4	hadewwww perempuan itu lg!!!!sakit jiwa,knp harus dia yg jd peran utama di film hantu jeruk purut,ky khabisan stok Artis ja	hadewwww perempuan itu lg sakit jiwa knp harus dia yg jd peran utama di film hantu jeruk purut ky khabisan stok artis ja
5	kebiasaan balajaer nyampah d ig para artis..suka2 yg punya ig lah mau bikin caption apa,kok balajaer yg heboh dan asik ceramahin yg punya ig.tar lama2 d bikin lagu sm teh melly loh balajaer yg berjudul"ketika balajaer bersabda"demi seonggak koin laknat nya.balajaer ngamuk berarti mengiyakan banget koin laknat nya jd penggoda d rt nya gigi.duh elah balajaer kami smua tau kale kelakuan biduan lo itu kek apa.minus segala2nya!lagi masih d bela mati2an,pd sakit jiwa	kebiasaan balajaer nyampah d ig para artis..suka yg punya ig lah mau bikin caption apa,kok balajaer yg heboh dan asik ceramahin yg punya ig tar lama d bikin lagu sm teh melly loh balajaer yg berjudul ketika balajaer bersabda demi seonggak koin laknat nya.balajaer ngamuk berarti mengiyakan banget koin laknat nya jd penggoda d rt nya gigi.duh elah balajaer kami smua tau kale kelakuan biduan lo itu kek apa.minus segala nya lagi masih d bela mati an pd sakit jiwa

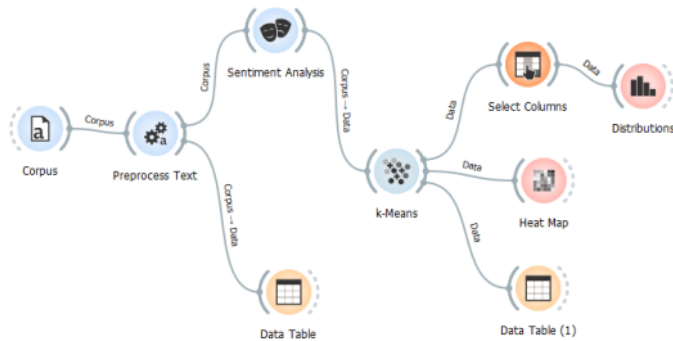
### 3.3 Sentiment Analysis

Mekanisme pemilahan kata negatif dan positif dilakukan dengan memanfaatkan fasilitas yang sudah tersedia di Orange, seperti melakukan analisis sentimen dengan menggunakan lexicon yang sudah terintegrasi di dalam software tersebut. Hal ini memudahkan penulis untuk memilah kata-kata positif dan negatif yang terkandung dalam teks yang sedang dianalisis, serta memudahkan pengguna dalam mengevaluasi kualitas teks yang dianalisis.

Pada penelitian ini menggunakan vader dari NLTK. Vader adalah lexicon- and rule-based sentiment analysis. Dalam widget Sentiment Analysis ada 4 feat. **1** score dari Vader method: positive score, negative score, neutral score dan compound (combined score). Compound merepresentasikan total sentiment dari tweet, dimana -1 merupakan jenis komentar paling negatif dan 1 merupakan komentar paling positif.

### 3.4 Clustering dengan K-Means

Proses klusterisasi menggunakan orange data mining digambarkan pada widget berikut ini:



**Gambar 2.** Design Widget Proses Clustering

Pada widget di atas proses klusterisasi dilakukan dari dataset yang terdiri dari 400 data yang belum memiliki label. Data ini berisi teks atau konten terkait cyberbullying. Sebelum melakukan klusterisasi, perlu dilakukan pre-processing pada data. *Pre-processing* mencakup langkah-langkah seperti menghilangkan tanda baca, menghapus karakter yang tidak diperlukan, mengubah teks menjadi huruf kecil, dan menghapus kata-kata umum (*stop words*). Tujuan dari *pre-processing* adalah untuk membersihkan data dan mempersiapkannya agar siap digunakan untuk analisis selanjutnya. Setelah *pre-processing* dilakukan, langkah berikutnya adalah melakukan *sentiment analysis* menggunakan algoritma *verde* yang sudah tersedia pada *orange*.

*Sentiment analysis* adalah proses untuk menentukan sentimen atau perasaan yang terkandung dalam teks, apakah positif, negatif, atau netral. Algoritma *verde* digunakan untuk menganalisis sentimen dalam teks terkait dengan cyberbullying. Setelah *sentiment analysis* dilakukan, langkah berikutnya adalah melakukan clustering menggunakan algoritma *k-means*.

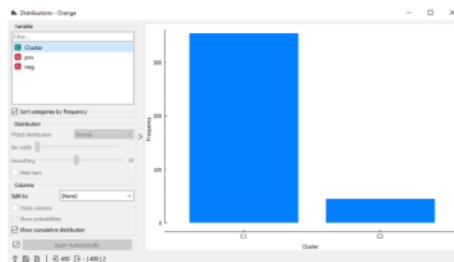
20

*K-means* adalah salah satu metode klusterisasi yang populer. Dalam konteks ini, *k-means* digunakan untuk mengelompokkan data menjadi dua kluster, yaitu kluster positif dan negatif, berdasarkan sentimen yang terdeteksi dalam langkah sebelumnya. Setelah proses klusterisasi selesai, langkah terakhir adalah melakukan pendistribusian hasil klusterisasi. Hal ini dilakukan untuk melihat kluster yang paling dominan atau mendominasi dalam dataset. Dengan melihat hasil klusterisasi, diperoleh pemahaman yang lebih baik tentang sebaran data dan dapat mengidentifikasi kluster yang paling signifikan dalam konteks cyberbullying.

Evaluasi pada proses klusterisasi ini menggunakan nilai *silhouette score* dimana cluster yang memiliki nilai *silhouette* paling mendekati angka 1 maka diartikan sebagai cluster yang paling dominan.

### 3.5 Pengujian hasil clustering dengan K-means

Dalam proses klusterisasi menggunakan metode k-means, terdapat beberapa parameter yang perlu ditentukan, yaitu jumlah kluster (k) dan jumlah iterasi maksimum. Parameter-parameter ini akan mempengaruhi penentuan centroid yang digunakan dalam perhitungan jarak antara data dan centroid menggunakan metode euclidean, serta penentuan anggota kluster. Evaluasi hasil klusterisasi dilakukan menggunakan metode koefisien *silhouette* dengan menguji jumlah kluster yaitu 2.



Gambar 3. Diagram distribusi data hasil clustering

Dari hasil klusterisasi dengan algoritma K-means menunjukkan adanya kelompok cluster negatif pada cluster C1. Cluster ini ditandai oleh komentar-komentar negatif yang dominan jika dilihat pada diagram distribusi data pada Gambar 2. Dalam penelitian ini, metode *silhouette score* digunakan untuk mengevaluasi kualitas klusterisasi. Nilai *silhouette score* sebesar 0.690152 menunjukkan bahwa cluster C1 memiliki tingkat pemisahan yang baik, dengan komentar-komentar negatif yang memiliki kesamaan yang tinggi dalam kelompok ini. Dilihat pada hasil distribusi data cluster C1 memiliki distribusi komentar negatif terbanyak.

Tabel 2. Hasil evaluasi menggunakan *silhouette score*

Cluster	Silhouette Score
C1	0.690152
C2	0.594038

Hasil evaluasi menggunakan *silhouette score* menunjukkan bahwa 0.690152 pada cluster C1 mendekati angka 1 yang memiliki arti di mana nilai yang lebih tinggi menunjukkan kualitas klusterisasi yang lebih baik. Nilai *silhouette score* pada hasil klusterisasi penelitian ini menunjukkan bahwa klusterisasi memiliki tingkat pemisahan yang baik, dengan objek-objek dalam kluster yang memiliki kesamaan yang tinggi dalam kelompok.



#### 4. Kesimpulan

Dari hasil analisis pengelompokan teks yang dilakukan pada status dan komentar mengenai cyberbullying di Instagram Indonesia, dapat disimpulkan bahwa komentar *cyberbullying* di Instagram paling dominan adalah komentar negatif. Dengan menggunakan algoritma K-means, komentar-komentar yang ditemukan dalam dataset penelitian dielaborasi menjadi kelompok-kelompok berdasarkan kesamaan karakteristik mereka. Cluster C1, yang teridentifikasi sebagai cluster negative merupakan cluster paling dominan dibuktikan dengan hasil distribusi data. Pada hasil klasterisasi juga didapatkan nilai *silhouette score* sebesar 0.690152 yaitu Cluster C1 menunjukkan tingkat pemisahan yang optimal, di mana komentar-komentar negatif dalam kelompok ini memiliki karakteristik yang serupa dan saling berkaitan. Penelitian ini memiliki implikasi penting dalam pengembangan langkah-langkah pencegahan dan penanggulangan cyberbullying di Instagram. Dengan mengetahui komentar-komentar negatif yang khas, langkah-langkah perlindungan dapat dirancang untuk melindungi pengguna dari perilaku cyberbullying, serta meningkatkan lingkungan yang lebih positif dan aman di dalam platform tersebut. Saran untuk penelitian selanjutnya adalah melakukan evaluasi score selain *silhouette coefficient* serta menguji pada beberapa *cluster* untuk mencari nilai k-terbaik. Dapat juga dilakukan proses *pre-processing* dengan menambahkan penanganan negasi untuk menghindari kesalahan interpretasi atau kesimpulan yang salah akibat adanya pernyataan-pernyataan negative sehingga dapat meningkatkan keakuratan dalam menganalisis data.

#### 5. Referensi

- Alshari, E. M. (2017). *Improvement of Sentiment Analysis based on Clustering of Word2Vec Features*. 0–3. <https://doi.org/10.1109/DEXA.2017.41>
- Candra, R. M., & Rozana, A. N. (2020). *Klasifikasi Komentar Bullying pada Instagram Menggunakan Metode K-Nearest Neighbor*. 5(1), 45–52.
- Djaballah, K. A., Boukhalifa, K., & Boussaid, O. (2019). Sentiment Analysis of Twitter Messages using Word2vec by Weighted Average. *2019 6th International Conference on Social Networks Analysis, Management and Security, SNAMS 2019*, 223–228. <https://doi.org/10.1109/SNAMS.2019.8931827>
- Hasdya, N., & Dinata, R. K. (2020). Analisis Matthew Correlation Coefficient pada K-Nearest Neighbor dalam Klasifikasi Ikan Hias. *INFORMAL: Informatics Journal*, 5(2), 57. <https://doi.org/10.19184/isj.v5i2.18907>
- Jardim, S., Mora, C., Jardim, S., & Mora, C. (2022). ScienceDirect ScienceDirect Customer reviews sentiment-based analysis and clustering for tourism services and products development or Customer reviews and clustering for positioning market-oriented tourism services and products development or positioning. *Procedia Computer Science*, 196(2021), 199–206. <https://doi.org/10.1016/j.procs.2021.12.006>
- Khaira, U., Johanda, R., Eko, P., Utomo, P., Suratno, T., & Info, A. (2020). *Sentiment Analysis of Cyberbullying on Twitter Using SentiStrength*. 3(1), 21–27.
- Liu, S. (2016). *Sentiment Clustering with Topic and Temporal Information from Large Email Dataset*. *Paclac* 30, 363–371.
- Method, M., Triyana, R., Putra, O. V., Pradhana, F. R., & Gontor, U. D. (n.d.). *Deteksi Cyberbullying Pada Tweet Berbahasa Inggris Dengan Metode Support Vector Machine*. (2017) 98–103
- Nainggolan, R., Adline, F., Tobing, T., Harianja, E. J. G., Akuntansi, K., Indonesia, U. M., Informatika, P., & Nusantara, U. M. (2023). *Analysis Sentiment in Bukalapak Comments with K-Means Clustering Method*. 9(2), 87–92.
- Ningtyas, A. M., & Herwanto, G. B. (2018). The Influence of Negation Handling on Sentiment Analysis in Bahasa

Commented [AA3]: Tidak ada tahunnya

Indonesia. *Proceedings of 2018 5th International Conference on Data and Software Engineering, ICoDSE 2018, November 2018*. <https://doi.org/10.1109/ICODSE.2018.8705802>

Patihullah, J., & Winarko, E. (2019). Hate Speech Detection for Indonesia Tweets Using Word Embedding And Gated Recurrent Unit. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 13(1), 43. <https://doi.org/10.22146/ijccs.40125>

Prastyo, P. H., Sumi, A. S., Dian, A. W., & Permanasari, A. E. (2020). Tweets Responding to the Indonesian Government's Handling of COVID-19: Sentiment Analysis Using SVM with Normalized Poly Kernel. *Journal of Information Systems Engineering*

# Klasterisasi Komentar Cyberbullying Masyarakat di Instagram berdasarkan K-Means Clustering

## ORIGINALITY REPORT

12%

SIMILARITY INDEX

12%

INTERNET SOURCES

5%

PUBLICATIONS

2%

STUDENT PAPERS

## PRIMARY SOURCES

1	<a href="https://lms.onnocenter.or.id">lms.onnocenter.or.id</a> Internet Source	2%
2	<a href="http://www.scilit.net">www.scilit.net</a> Internet Source	1%
3	<a href="https://doku.pub">doku.pub</a> Internet Source	1%
4	<a href="https://repository.uin-suska.ac.id">repository.uin-suska.ac.id</a> Internet Source	1%
5	<a href="https://www.researchgate.net">www.researchgate.net</a> Internet Source	1%
6	Eko Budi Susanto, Paminto Agung Christianto, Mohammad Reza Maulana, Satriedi Wahyu Binabar. "Analisis Kinerja Algoritma Naïve Bayes Pada Dataset Sentimen Masyarakat Aplikasi NEWSAKPOLE Samsat Jawa Tengah", Jurnal CoSciTech (Computer Science and Information Technology), 2022 Publication	1%

Submitted to Universitas Merdeka Malang

7	Student Paper	1 %
8	<a href="http://www.jurnalfti.unmer.ac.id">www.jurnalfti.unmer.ac.id</a> Internet Source	1 %
9	Submitted to Sriwijaya University Student Paper	1 %
10	<a href="http://repositor.umm.ac.id">repositor.umm.ac.id</a> Internet Source	1 %
11	<a href="http://apkcara.com">apkcara.com</a> Internet Source	<1 %
12	<a href="http://id.scribd.com">id.scribd.com</a> Internet Source	<1 %
13	Saifurrachman Chohan, Arifin Nugroho, Achmad Maezar Bayu Aji, Windu Gata. "Analisis Sentimen Pengguna Aplikasi Duolingo Menggunakan Metode Naïve Bayes dan Synthetic Minority Over Sampling Technique", Paradigma - Jurnal Komputer dan Informatika, 2020 Publication	<1 %
14	<a href="http://journal.stekom.ac.id">journal.stekom.ac.id</a> Internet Source	<1 %
15	<a href="http://www.giters.com">www.giters.com</a> Internet Source	<1 %
16	<a href="http://core.ac.uk">core.ac.uk</a> Internet Source	

<1 %

---

17 [juti.if.its.ac.id](http://juti.if.its.ac.id)  
Internet Source

<1 %

---

18 Fordjour Richmond, Nwogu Chimaraoke  
Uchechukwu, Célia M. Q. Ramos. "chapter 9  
Analyses of Visitors' Experiences in Museums  
Based on E-Word of Mouth and Tripadvisor  
Online Reviews", IGI Global, 2023  
Publication

<1 %

---

19 [media.neliti.com](http://media.neliti.com)  
Internet Source

<1 %

---

20 Rustam Rustam, Sidik Rahmatullah, Supriyato  
Supriyato, Sri Wahyuni. "PENERAPAN DATA  
MINING UNTUK PREDIKSI PENJUALAN  
PRODUK TRIPLEK PADA PT PUNCAK MENARA  
HIJAU MAS", Jurnal Informasi dan Komputer,  
2020  
Publication

<1 %

---

Exclude quotes Off

Exclude matches Off

Exclude bibliography On