

VISUALISASI ANALISIS SENTIMEN SIBERBULLYING PADA POST INSTAGRAM MENGGUNAKAN ORANGE DATA MINING

Hindun Nur Azizah¹

¹Jurusan Sistem Informasi

Universitas Alma Ata Jl. Brawijaya No.99, Jadan, Tamantirto, Kec. Kasihan, Kabupaten Bantul, Daerah Istimewa Yogyakarta 55184, Indonesia

Author Emails

^{a)} Corresponding author: azizahhindunnur@gmail.com

Abstract. Instagram is a social media that is often used in the digital era like today. Many users from various walks of life use Instagram. But of the many benefits of Instagram, the negative impact of social media is undeniable, cyberbullying is a negative impact that always lurks whenever and wherever. The dangers of cyberbullying can be troubling to many parties. Cyberbullying on Instagram often occurs in the comments column, therefore this study was conducted to analyze the sentiment of cyberbullying in Instagram comments. Sentiment analysis is a branch of science from text mining that is used to extract and process data in the form of text. To find out the results of the art, this study uses webharvy as a tool to collect data and orange data mining as a tool to analyze. This research method is qualitative because Instagram comments are research material.

Keywords : Instagram, Cyberbullying, Orange Data Mining, Sentiment Analysis

Abstraksi. Instagram merupakan sosial media yang sering digunakan pada era digital seperti sekarang ini. Banyak pengguna dari berbagai kalangan menggunakan instagram. Tetapi dari sekian manfaat instagram, dampak negatif sosial media tak dapat dipungkiri, *cyberbullying* menjadi suatu dampak negatif yang selalu mengintai kapanpun dan dimanapun. Bahaya *cyberbullying* ternyata dapat meresahkan banyak pihak. *Cyberbullying* pada instagram kerap terjadi di kolom komentar, oleh karena itu penelitian ini dilakukan untuk menganalisis sentimen *cyberbullying* pada komentar instagram. Analisis sentimen merupakan cabang ilmu dari text mining yang dipakai untuk mengekstrak dan mengolah data berbentuk teks. Untuk mengetahui hasil senimen, penelitian ini menggunakan webharvy sebagai tool untuk mengumpulkan data, dan *orange data mining* sebagai tool untuk menganalisis. Metode penelitian ini adalah metode kualitatif dikarenakan komentar instagram sebagai bahan penelitian

Kata Kunci : instagram, cyberbullying, orange data mining, analisis sentimen

PENDAHULUAN

perkembangan teknologi komunikasi di era yang serba digital kini begitu pesat bahkan sulit di kontrol, perkembangan ini merupakan realisasi dari kebutuhan manusia sebagai mahluk hidup yang menginginkan pekerjaan yang serba instan dengan penyebaran segala bentuk informasi secara cepat [1]. Perkembangan teknologi ini sejalan dengan penggunaan internet yang terus bertambah dari tahun hingga tahun. Menurut dataindonesia.id pada januari 2021 pengguna internet di indonesia mencapai 203 juta jiwa bahkan di januari 2022 mencapai 205 juta pengguna. Tak dapat dipungkiri seiring bertambahnya penggunaan internet berta[2]mbah juga penggunaan sosial media. Menurut data dari hootsuite (*we are sosial*) tahun 2019 indonesia telah mencapai 150 juta pengguna atau bertambah sebanyak 15% - 20% dari tahun 2018.

[3]. Salah satu sosial media yang populer di Indonesia adalah Instagram dengan penggunaan sebanyak 87,7% dari total masyarakat di tahun 2021 data dari Hootsuite *we are social*. Instagram adalah sebuah aplikasi yang berguna untuk membagikan foto dan video [4]. Beberapa masyarakat menggunakan Instagram sebagai tempat untuk mempromosikan produk ataupun jasa, namun beberapa masyarakat yang lain menggunakan Instagram sebagai media cyberbullying. Tentu saja hal ini menimbulkan keresahan bagi masyarakat. *Cyberbullying* merupakan perluasan dari *bullying*. *Cyberbullying* dapat diartikan sebagai kejahatan sosial media dengan mengintimidasi yang dilakukan seseorang atau sekelompok orang yang tujuannya untuk menjatuhkan, memojokkan, dan menyudutkan [5]. Dampak dari cyberbullying antara lain yaitu fisik, psikologis dan emosional, pendidikan, dan psikososial [6]. Faktor yang menyebabkan *cyberbullying* antara lain adalah pengalaman kekerasan, perbedaan gender, usia, kontrol psikologis, dan penggunaan zat-zat adiktif [2]. *Cyberbullying* pada platform Instagram biasanya terjadi di kolom komentar, hal ini menimbulkan keinginan peneliti terhadap komentar apa saja yang termasuk dalam kategori *cyberbullying*. Salah satu cara untuk mengetahui apa saja yang termasuk ke dalam kategori *cyberbullying* adalah analisis sentimen. Analisis sentimen diperlukan untuk menyaring komentar yang ada di Instagram. Analisis sentimen dilakukan untuk mengetahui komentar yang berkategori positif dan negatif [7]. Metode yang digunakan untuk memperoleh komentar adalah *scraping data* dibantu dengan tool yang bernama *webharvy*. Kemudian alat bantu untuk mendapatkan hasil analisis sentimen adalah *orange data mining*.

Orange data mining merupakan software open source untuk mengolah data analisis atau data mining. Penggunaan orange data mining dikarenakan visualisasi yang lebih unggul dari segi gambar maupun penulisan. Orange menyediakan banyak widget untuk memvisualisasikan atau menggambarkan antar proses. Dengan media seperti canvas ini, pengguna lebih mudah dalam memahami proses analisis yang berjalan [8].

TINJAUAN PUSTAKA

Pada penelitian ini, peneliti menggunakan penelitian terdahulu sebagai pembandingan dan tolak ukur dari penelitian yang akan dilaksanakan. Tujuan dari perbandingan ini adalah mempermudah peneliti dalam menyusun penelitian. Peneliti harus belajar dari penelitian lain untuk menghindari penduplikasian, pengulangan atau kesalahan yang sama seperti penelitian yang dilaksanakan sebelumnya. Penelitian sebelumnya dipakai sebagai referensi dan acuan untuk mempermudah peneliti dalam membuat penelitian ini. Peneliti telah menganalisis satu persatu dari penelitian terdahulu yang berkaitan dengan proyek penelitian ini.

Sosial Media

Media sosial merupakan salah satu tempat penyebaran *cybercrime* yang paling mudah untuk disalahgunakan. Terjadi banyak kasus *cybercrime* di beberapa sosial media, terlebih pada sosial media seperti Facebook, Instagram, ataupun Twitter. Secara garis hukum pelaku kejahatan *cybercrime* akan dikenakan Undang-Undang Nomor 11 tahun 2008 tentang Informasi dan Transaksi Elektronik (UU ITE) baik melakukan perbuatan *cybercrime* yang disengaja maupun yang tidak disengaja. Semua pengguna internet dan sosial media harus melakukan tindakan yang dapat mencegah terjadinya *cybercrime*, di antara cara mencegah terjadinya *cybercrime* adalah menjaga privasi, menghindari hoax, mengamankan akun, selalu haus terhadap informasi yang positif, mengetahui etika berinternet, serta menjaga komputer dari masuknya virus. [3]

Cyberbullying

Sebagian besar mahasiswa Universitas Pembangunan Jaya (UPJ) mengetahui *cyberbullying* dan pernah melihat tindakan *cyberbullying*. Beberapa mahasiswa juga mengakui bahwa pernah mengalami *cyberbullying* dengan berbagai motif. Dengan arti lain beberapa mahasiswa pernah menjadi korban *cyberbullying*. Penyebab terjadinya *cyberbullying* pada mahasiswa paling sering dikarenakan sakit hati, tidak bisa mengendalikan emosi (marah), dan dendam. Dengan sistem pendukung teknologi yang dapat diakses 24 jam, hal ini sangat mempermudah untuk melakukan *cyberbullying* kapanpun dengan alasan sebagai pelampiasan emosi dan sakit hati kepada orang lain. Dampak dari *cyberbullying* paling besar terjadi pada korban. Dampak *cyberbullying* akan membuat korban merasa harga dirinya rendah. Penyebab utamanya adalah korban menerima segala pesan dan perlakuan negatif dari orang lain secara terus menerus. Hal ini akan membuat hilangnya kepercayaan diri korban, sehingga korban merasa dirinya tidak berguna. [9]

Orange Data mining

Orange data mining merupakan alhoritma yang digunakan untuk pengelolaan data untuk menemukan informasi dan pengetahuan baru. Orange data mining dapat digunakan untuk menganalisis dan memvisualisasikan data. [10]

Analisis Sentimen

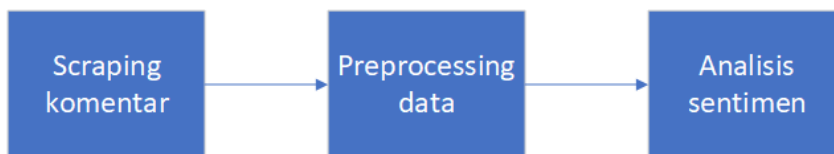
Analisis sentimen adalah ilmu data mining yang dimulai dengan cara memahami teks, mengekstrak, dan mengoah data yang berbentuk kalimat secara otomatis sehingga menjadi informasi yang dapat dipahami dengan baik. [11]

METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah metode kualitatif, dikarenakan penelitian ini berfokus pada komentar yang ada pada salah satu postingan instagram selebgram.

Desain Penelitian

Desain penelitian adalah kerangka atau proses yang dilewati peneliti guna mendapatkan hasil penelitian yang sistematis. Penelitian ini membahas tentang analisis sentimen sehingga diperoleh desain penelitian sebagai berikut:

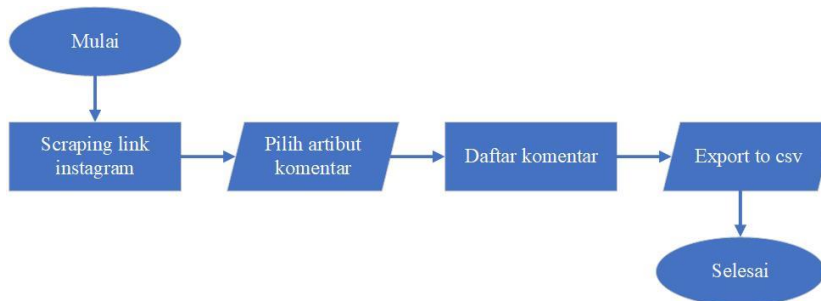


GAMBAR 1. Diagram Air Desain Penelitian

Proses pertama adalah *scraping* data (pengumpulan komentar) menggunakan *webharvy*, dilanjutkan proses kedua yaitu *preprocessing* data, dan yang terakhir adalah analisis sentimen. Setelah dilakukannya tiga proses tersebut maka akan diperoleh hasil analisis sentimen.

Metode Pengumpulan Data

langkah yang pertama diambil dalam penelitian analisis sentimen ini adalah pengumpulan komentar atau *scraping* komentar menggunakan alat bantu *webharvy*. Cara kerja *webharvy* adalah memngekstrak komentar menjadi tabel dalam bentuk file csv [12]. *Webharvy* bekerja dengan mengakses halaman web, kemudian memilih data, mengekstrak data, dan menyimpan data dalam bentuk tabel csv. Data kemudian dapat di unduh dan dapat di proses. Pada proses *scraping* ini data komentar yang diperoleh sebanyak 168 komentar. Berikut adalah diagram air dari proses *scraping* data:



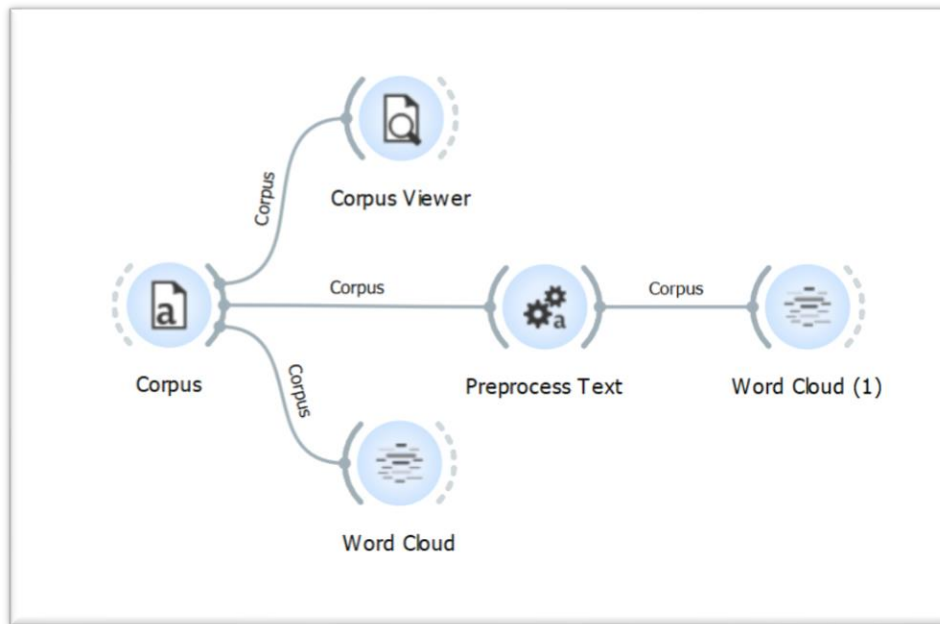
GAMBAR 2. Diagram Air Metode Pengumpulan Data

Preprocessing Data

Preprocessing data adalah membagi teks menjadi ukuran yang lebih kecil dengan melewati proses transformation, tokenization, normalization, dan filtering. Dalam setiap proses yang dilewati mempunyai fungsi masing- masing [10]. Dibawah ini merupakan tahapan *preprocessing* data menggunakan orange data mining:

- Transformation, *Transformation* merupakan proses dimana seluruh kata di ubah menjadi huruf kecil secara default.
- Tokenization, *Tokenization* merupakan tahap dimana kalimat di pecah menjadi komponen yang lebih kecil. Contoh ini buku. Menjadi (ini), (buku), (.).
- Normalization, Normalization merupakan tahap dimana pembenaran kata yang salah menjadi kata yang benar sesuai dengan KBBI (Kamus Besar Bahasa Indonesia). Proses ini diharapkan dapat membantu peneliti untuk mendapatkan data dengan ejaan yang tepat.
- Filtering, Filtering merupakan tahap penghilangan kata atau tanda yang tidak digunakan. Tahap ini dibagi menjadi 3 (tiga) bagian yaitu : stopwords (menghapus kata- kata yang telah ditetapkan dan tidak digunakan), regexp (menghapus tanda baca yang tidak penting atau tidak digunakan), dan most frequent token (menghapus kata – kata yang dipisahkan).

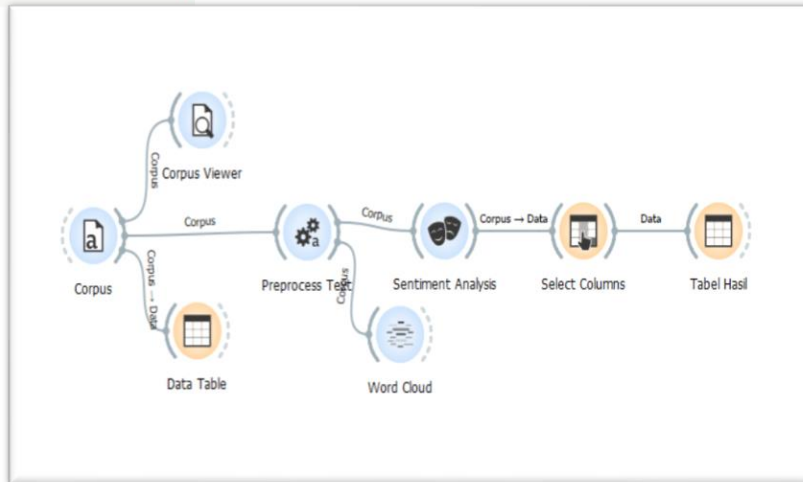
Berikut adalah widget preprocessing data dalam orange data mining :



GAMBAR 3. Widget Preprocessing Data

HASIL DAN PEMBAHASAN

Berdasarkan dari desain penelitian dimana komentar sebagai variabel yang akan diteliti, maka penggunaan orange data mining sebagai alat bantu analisis sangat berjasa penggunaannya. Berikut adalah peta konsep dari wedget yang digunakan dalam analisis sentimen ini:



GAMBAR 4. Widget Analisis Sentimen

Hasil Preprocessing Data

Word cloud merupakan visualisasi dari hasil preprocessing data pada orange data mining. *Word cloud* akan menampilkan gambar yang telah melewati tahap preprocessing text. Dari data tersebut semakin besar ukuran kata maka presentase penggunaan kata tersebut semakin banyak. Berikut adalah *Word cloud* dari proses preprocessing text:



GAMBAR 5. Word cloud Preprocessing Text

Berikut adalah 10 kata terbanyak yang dianggap sebagai cyberbullying, disajikan pada Tabel 1.

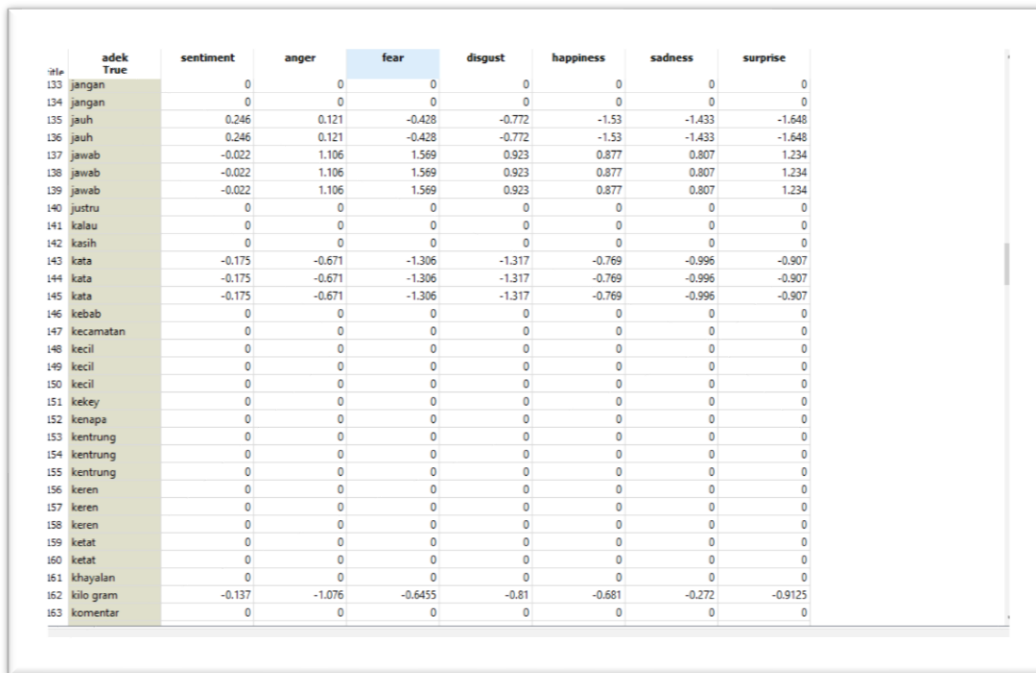
Tabel 1 Data yang dianggap cyberbullying

Count	Word
19	gitar
15	mirip
11	spanyol

10	ukulele
7	beduk
5	gendang
3	kentrung
3	serem
2	rebana
2	soneta

Analisis Sentimen

Hasil analisis sentimen diperoleh dari beberapa proses yang harus dilewati. Proses analisis sentimen di tampilkan dengan menggunakan tampilan SentiArt sebagai tabel yang akan memberikan hasil analisis berupa angka positif, negatif, dan netral. Tabel tersebut berisi tentang *sentiment* (sentimen), *anger* (kemarahan), *fear* (ketakutan), *disgust* (menjijikkan), *happiness* (kebahagiaan), *sadness* (kesedihan), dan *surprise* (kejutan) [13]. Berikut adalah hasil dari analisis sentimen menggunakan tampilan SentiArt:



ada	adek True	sentiment	anger	fear	disgust	happiness	sadness	surprise
133	jangan	0	0	0	0	0	0	0
134	jangan	0	0	0	0	0	0	0
135	jauh	0.246	0.121	-0.428	-0.772	-1.53	-1.433	-1.648
136	jauh	0.246	0.121	-0.428	-0.772	-1.53	-1.433	-1.648
137	jawab	-0.022	1.106	1.569	0.923	0.877	0.807	1.234
138	jawab	-0.022	1.106	1.569	0.923	0.877	0.807	1.234
139	jawab	-0.022	1.106	1.569	0.923	0.877	0.807	1.234
140	justru	0	0	0	0	0	0	0
141	kalau	0	0	0	0	0	0	0
142	kasih	0	0	0	0	0	0	0
143	kata	-0.175	-0.671	-1.306	-1.317	-0.769	-0.996	-0.907
144	kata	-0.175	-0.671	-1.306	-1.317	-0.769	-0.996	-0.907
145	kata	-0.175	-0.671	-1.306	-1.317	-0.769	-0.996	-0.907
146	kebab	0	0	0	0	0	0	0
147	kecamatan	0	0	0	0	0	0	0
148	kecil	0	0	0	0	0	0	0
149	kecil	0	0	0	0	0	0	0
150	kecil	0	0	0	0	0	0	0
151	kekey	0	0	0	0	0	0	0
152	kenapa	0	0	0	0	0	0	0
153	kentrung	0	0	0	0	0	0	0
154	kentrung	0	0	0	0	0	0	0
155	kentrung	0	0	0	0	0	0	0
156	keren	0	0	0	0	0	0	0
157	keren	0	0	0	0	0	0	0
158	keren	0	0	0	0	0	0	0
159	ketat	0	0	0	0	0	0	0
160	ketat	0	0	0	0	0	0	0
161	khayalan	0	0	0	0	0	0	0
162	kilo gram	-0.137	-1.076	-0.6455	-0.81	-0.681	-0.272	-0.9125
163	komentar	0	0	0	0	0	0	0

GAMBAR 6. Hasil Analisis Sentimen

KESIMPULAN

Berdasarkan hasil eksperimen tentang pola sentimen cyberbullying terdapat sepuluh kata cyberbullying terbanyak yang digunakan dalam komentar, antara lain adalah gitar, mirip, spanyol, ukulele, beduk, gendang, kentrung, serem, rebana, dan soneta. Penggunaan alat pengolahan data, yaitu Orange data mining dianggap mampu bekerja dengan baik dalam menganalisis data khususnya analisis sentimen. Tampilan SentiArt dalam orange data mining memberikan 7 kelas tabel yaitu *sentiment* (sentimen), *anger* (kemarahan), *fear* (ketakutan), *disgust* (menjijikkan), *happiness* (kebahagiaan), *sadness* (kesedihan), dan *surprise* (kejutan). Adapun penelitian yang akan datang akan membahas tentang pola clustering tentang pengguna Instagram, guna mendeteksi habit pengguna instagram tersebut.

TINJAUAN PUSTAKA

- [1] M. Rifauddin, “Fenomena Cyberbullying pada Remaja (Studi Analisis Media Sosial Facebook),” *Khazanah al-Hikmah J. Ilmu Perpustakaan, Informasi, dan Kearsipan*, vol. 4, no. 1, pp. 35–44, 2016, doi: <https://doi.org/10.24252/kah.v4i1a3>.
- [2] H. A. Dewi, Suryani, and A. Sriati, “Faktor-Faktor yang Memengaruhi Cyberbullying Pada Remaja: A Systematic Review,” *J. Nurs. Care*, vol. 3, no. 2, pp. 128–141, 2020, [Online]. Available: <http://jurnal.unpad.ac.id/jnc/article/view/24477>.
- [3] Y. Fitriani and R. Pakpahan, “Analisa Penyalahgunaan Media Sosial untuk Penyebaran Cybercrime di Dunia Maya atau Cyberspace,” *Cakrawala-Jurnal Hum.*, vol. 20, no. 1, pp. 21–27, 2020.
- [4] T. Rachman, “BAB II,” *Angew. Chemie Int. Ed.* 6(11), 951–952., pp. 10–27, 2018.
- [5] Y. Oetary and R. Hotmaulana Hutauruk, “Kajian Yuridis Terhadap Tindak Pidana Dalam Aspek Perundangan Dunia Maya (Cyberbullying): Perspektif Hukum Pidana Di Indonesia,” vol. 4, no. November, pp. 1045–1055, 2021.
- [6] A. Sukmawati and A. P. B. Kumala, “Dampak Cyberbullying Pada Remaja,” *Alauddin Sci. J. Nurs.*, vol. 1, no. 1, pp. 55–65, 2020, [Online]. Available: <http://journal.uin-alauddin.ac.id/index.php/asjn/article/download/17648/9946>.
- [7] W. Athira, I. Gholissodin, and rizal setya Perdana, “Analisis Sentimen Cyberbullying Pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 11, pp. 4704–4713, 2018, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/3051>.
- [8] S. Kasus and K. Tasikmalaya, “Jurnal saintesa,” vol. 2, 2022.
- [9] F. S. Mutma, “Deskripsi pemahaman,” pp. 165–182, [Online]. Available: <https://journal.trunojoyo.ac.id/komunikasi/article/view/5928/4250>.
- [10] R. A. raffaidy Wiguna and A. I. Rifai, “Analisis Text Clustering Masyarakat Di Twitter Mengenai Omnibus Law Menggunakan Orange Data Mining,” *J. Inf. Syst. Informatics*, vol. 3, no. 1, pp. 1–12, 2021, doi: 10.33557/journalisi.v3i1.78.
- [11] M. F. Rizki, K. Auliasari, and R. Primaswara Prasetya, “Analisis Sentiment Cyberbullying Pada Sosial Media Twitter Menggunakan Metode Support Vector Machine,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 5, no. 2, pp. 548–556, 2021, doi: 10.36040/jati.v5i2.3808.
- [12] T. S. Kasus, P. Pada, and S. Tripadvisor, “Analisis Sentimen Pelanggan Hotel di,” vol. 8106, pp. 21–29, 2021, doi: 10.20895/INISTA.V3.
- [13] A. M. Jacobs, “Sentiment Analysis for Words and Fiction Characters From the Perspective of Computational (Neuro-)Poetics,” *Front. Robot. AI*, vol. 6, no. July, pp. 1–13, 2019, doi: 10.3389/frobt.2019.00053.