

## DETEKSI CYBERBULLYING DENGAN MESIN PEMBELAJARAN KLASIFIKASI (SUPERVISED LEARNING): PELUANG DAN TANTANGAN

Yudi Setiawan<sup>\*1</sup>, Nur Ulva Maulidevi<sup>2</sup>, Kridanto Surendro<sup>3</sup>

<sup>1,2,3</sup>Institut Teknologi Bandung, Bandung

Email: <sup>1</sup>33221022@std.itb.ac.id, <sup>2</sup>ulfa@staff.stei.itb.ac.id, <sup>3</sup>endro@staff.stei.itb.ac.id

\*Penulis Korespondensi

(Naskah masuk: 05 Desember 2022, diterima untuk diterbitkan: 26 Desember 2022)

### Abstrak

Perkembangan teknologi media sosial tidak hanya memberikan kemudahan dalam berkomunikasi antar individu, akan tetapi juga dapat mengancam kehidupan sosial individu seperti tindakan cyberbullying. Bervariasinya pola dan karakteristik cyberbullying mengakibatkan sulitnya proses deteksi cyberbullying, yang dilakukan oleh pelaku cyberbullying. Penelitian deteksi pola dan karakteristik cyberbullying banyak dilakukan dengan berbagai metode, seperti dengan mengimplementasikan Machine Learning, Natural Language Processing (NLP), dan Sentiment Analysis yang memiliki variasi akurasi yang berbeda, dengan keunggulan dan kelemahan dari masing-masing metode. Implementasi Machine Learning untuk deteksi cyberbullying dapat dilakukan dengan berbagai algoritma, seperti algoritma probabilistik (Naïve Bayes) maupun supervised learning (Support Vector Machine, k-Nearest Neighbour, Decision Tree), dan metode lainnya yang hingga saat ini terus dikembangkan dengan berbagai pendekatan untuk meningkatkan akurasi deteksi cyberbullying atau non-cyberbullying. Adapun peluang dan tantangan penelitian deteksi cyberbullying seperti penerapan pada variasi domain bahasa, dan bentuk ekspresi yang dilakukan pada suatu lingkungan atau budaya, yang masih terdapat ruang untuk dikembangkan dan dijelajah secara luas. Pada artikel ini menjabarkan penelitian berikutnya berupa mengimplementasikan metode pembelajaran klasifikasi (Supervised Learning) dengan modifikasi tahapan untuk meningkatkan akurasi klasifikasi.

**Kata kunci:** *Cyberbullying, machine learning, supervised learning, media sosial*

## CYBERBULLYING DETECTION USING SUPERVISED LEARNING: OPPORTUNITIES AND CHALLENGES

### Abstract

*The development of social media technology not only provides convenience in communicating between individuals but can also threaten the social life of individuals, such as through cyberbullying. The variety of patterns and characteristics of cyberbullying makes it challenging to detect cyberbullying, which is carried out by cyberbullying perpetrators. Research on the detection of cyberbullying patterns and features has been carried out using various methods, such as implementing Machine Learning, Natural Language Processing (NLP), and Sentiment Analysis, which have different variations in accuracy, with the advantages and disadvantages of each method. The implementation of Machine Learning for cyberbullying detection can be done using various algorithms, such as probabilistic algorithms (Naïve Bayes) and supervised learning (Support Vector Machine, k-Nearest Neighbor, Decision Tree), and other methods that are currently being developed with various approaches to improve cyberbullying. Accuracy of detection of cyberbullying or non-cyberbullying. As for the opportunities and challenges of cyberbullying detection research, such as its application to a variety of language domains and forms of expression carried out in an environment or culture, there is still room to be developed and explored widely. This article describes the following research in the form of implementing a classification learning method (Supervised Learning) with step modifications to improve classification accuracy.*

**Keywords:** *Cyberbullying, machine learning, supervised learning, social media*

### 1. PENDAHULUAN

Perkembangan teknologi media sosial dan media online mempengaruhi pola komunikasi dan

interaksi antar individu. Hal ini juga mengakibatkan meningkatnya tindakan cyberbullying yang mengancam kesehatan dan mental individu (W. M.

Al-Rahmi, N. Yahaya, M. M. Alamri, N. A. Aljarboa, Y. B. Kamin, & M. S. B. Saud, 2019). Tindakan cyberbullying merupakan transformasi bentuk dari tindakan bullying tradisional. Hal ini terjadi karena terdapat dukungan perkembangan teknologi media sosial dan teknologi komunikasi elektronik yang mampu menjangkau jarak yang jauh dan masif. Tindakan cyberbullying tidak dilakukan secara fisik atau tatap muka, melainkan dilakukan dengan meninggalkan jejak digital seperti tulisan, gambar, maupun video yang berdampak besar pada kehidupan sosial korban (Sheldon dkk., 2019). Sehingga hal ini dikatakan sebagai bentuk kejahatan di dunia maya yang mampu menjatuhkan dan mengintimidasi seseorang secara digital, yang dapat dilakukan oleh siapa saja, kapan saja, dan dimana saja, umumnya melalui aplikasi media sosial (López-Vizcaíno dkk., 2021; W. M. Al-Rahmi, N. Yahaya, M. M. Alamri, N. A. Aljarboa, Y. B. Kamin, & F. A. Moafa, 2019). Cyberbullying yang merupakan tindakan verbal yang bersifat negatif, dengan faktanya adalah bahwa postingan media sosial bersifat interaktif dan bergantung pada konteks; karenanya mereka tidak boleh dianggap sebagai teks yang berdiri sendiri (Bozyigit dkk., 2021). Sehingga perlunya penelitian lanjutan untuk melakukan deteksi cyberbullying dengan mengklasifikasikan pola dan karakteristik cyberbullying.

Pada pemrosesan tekstual terdapat permasalahan untuk menentukan arti yang sebenarnya, hal ini disebabkan oleh terdapatnya variasi bahasa yang tidak konsisten dalam penulisan teks, terdapatnya variasi cara penulisan (writing style) yang tidak umum dan baku, dan penggunaan kata yang mengandung makna ambigu yang mengakibatkan kebutuhan komputasi tinggi untuk mendapatkan makna dari kontekstual teks secara tepat (Rajput, 2020). Cyberbullying pada konten tekstual memerlukan proses aturan (rules) agar pada proses klasifikasi teks, dapat dengan tepat mengklasifikasikan teks/ ujaran tersebut termasuk dalam kelas/label yang sudah dibuat (Murnion dkk., 2018). Aturan (rules) dibuat dengan membentuk aturan keputusan atas parameter-parameter yang akan dilakukan proses klasifikasi dengan metode klasifikasi tertentu.

Proses deteksi cyberbullying dapat dilakukan salah satunya dengan mengimplementasikan metode klasifikasi seperti Machine Learning, Sentiment Analysis, Fuzzy Logic maupun metode deteksi lainnya. Untuk dapat melakukan proses klasifikasi perlu adanya proses pelatihan dengan features yang tepat dan baik untuk melakukan proses klasifikasi pengalabelan cyberbullying/ non-cyberbullying, maupun klasifikasi ke dalam jenis-jenis cyberbullying.

## 2. PENELITIAN TERKAIT DAN DISKUSI

Proses deteksi cyberbullying dapat dilakukan dengan melakukan dua tahapan, yaitu; menentukan parameter sebagai fitur penentu pada proses deteksi cyberbullying dan non-cyberbullying; dan menentukan metode atau algoritma untuk melakukan proses klasifikasi maupun deteksi cyberbullying berdasarkan fitur yang dibangun atau ditentukan. Tahapan penentuan parameter atau pembentukan fitur cyberbullying dapat dilakukan dengan analisis bentuk transformasi tradisional bullying menjadi cyberbullying. Sedangkan pada penentuan metoda dapat dilakukan analisis pembeda dari masing-masing metode deteksi.

### 2.1. Bentuk Transformasi Tradisional Bullying Menjadi Cyberbullying

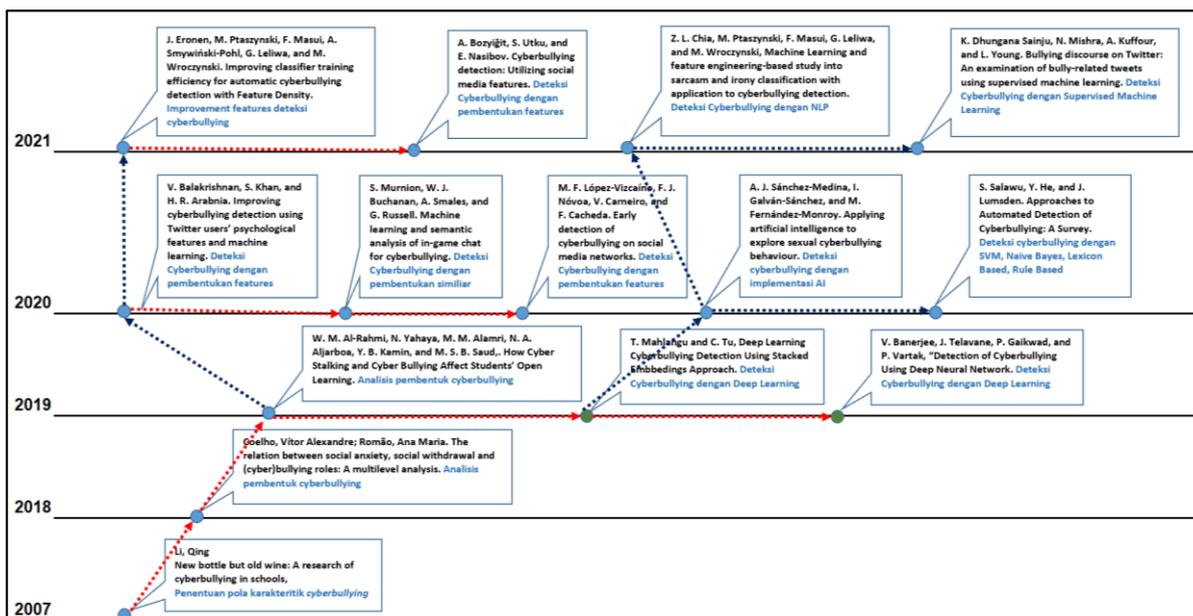
Tindakan *bullying* merupakan tindakan negatif oleh suatu individu atau kelompok kepada individu lainnya yang dilakukan secara langsung, berulang kali dengan ketidakseimbangan kekuatan antara pelaku dan korban, hal inilah yang dikatakan sebagai tradisional *bullying* (Barlett, 2019; Sheldon dkk., 2019). Tindakan *bullying* memiliki karakteristik seperti;

1. Perbedaan kekuatan (*Strength differential*)
2. Dilakukan terus-menerus (*Repeatability*)
3. Memiliki tipe serangan (*Aggression types*)

Pada *cyberbullying* terdapat karakteristik tambahan yang tidak dimiliki oleh tradisional *bullying*, yaitu; pelaku *cyberbullying* dapatlah seorang yang tidak dikenal (*Anonymity*) yang dapat terjadi di media sosial. Karakteristik-karakteristik *cyberbullying* tersebut dapat diinvestigasi sebagai *features* dari mesin pembelajaran sebagai deteksi *cyberbullying* (Chia dkk., 2021). Adapun bentuk transformasi tradisional *bullying* menjadi *cyberbullying* ditunjukkan pada Tabel 1 berikut:

Tabel 1. Perbandingan Tradisional Bullying dan Cyberbullying (Barlett, 2019; Sheldon dkk., 2019)

Tradisional <i>Bullying</i>	<i>Cyberbullying</i>
Terdapat perbedaan kekuatan ( <i>Strength differential</i> ) antara pelaku dan korban	Terdapat perbedaan Computer Skills, seperti; kemampuan membajak akun, penggunaan aplikasi komputer, pengiriman virus, dll). Perbedaan status ekonomi dan kepemilikan perangkat komputer.
Tindakan dilakukan terus menerus ( <i>Repeatability</i> )	Serangan berupa komentar yang terus berulang, atau serangan dilakukan secara masif oleh sekelompok orang yang ditujukan kepada satu korban.
Tindakan dilakukan secara tatap muka ( <i>face-to-face</i> ) Pelaku diketahui/ dikenal oleh korban	Fokus pada serangan verbal/ relasional pengucilan di media sosial. <i>Fake user, one time email</i> , menyembunyikan identitas saat online.



Gambar 1. Perkembangan Penelitian Deteksi Cyberbullying

## 2.2. Perkembangan Penelitian Deteksi Cyberbullying

Penelitian tentang *cyberbullying* terus berkembang seiring perkembangan teknologi media sosial. Hal ini disebutkan bahwa *cyberbullying* adalah masalah tindak klasik dari bentuk pengancaman kepada seseorang, akan tetapi masih dalam bentuk yang sama (*New Bottle but Old Wine*) (Li, 2007). Hal ini juga dibuktikan dengan perkembangan publikasi yang melakukan pendeteksian, pengelompokkan, dan studi *cyberbullying* yang terus berkembang terutama sejak tahun 2014 dan 2016, dan masih menjadi perhatian ilmiah yang terus-menerus dan dilakukan diseluruh disiplin ilmu. Adapun perkembangan penelitian deteksi *cyberbullying* hingga saat ini, penulis rangkum pada Gambar 1.

Pada Gambar 1 menunjukkan perkembangan penelitian deteksi *cyberbullying* baik dari perkembangan parameter atau fitur penentu proses deteksi, maupun perkembangan beberapa metode yang digunakan untuk melakukan proses deteksi.

## 2.3. Metode Deteksi Cyberbullying

Proses deteksi *cyberbullying* di media sosial ataupun platform digital lainnya, dapat dilakukan dengan cara fokus pada perbandingan sekuensial dan waktu sebagai pembelajaran cognitive yang dapat dilakukan oleh sebuah mesin (Russel & Norvig, 2021). Berbagai metode pembelajaran dan algoritma telah banyak dikembangkan dan diimplementasikan untuk melakukan proses deteksi *cyberbullying*. Proses deteksi *cyberbullying* dapat diimplementasikan dengan berbagai metode dan algoritma, ditunjukkan pada Tabel 2.

Metode deteksi *cyberbullying* dari masing-masing metode pada Tabel 2 memiliki keunggulan dalam akurasi deteksi, akan tetapi terdapat perbedaan

dari tahapan proses deteksi. Mesin pembelajaran merupakan metode saintis dimana komputer mempelajari bagaimana menyelesaikan permasalahan tanpa memprogramkan secara eksplisit (T.K. dkk., 2021).

Tabel 2. Algoritma Pada Model Pembelajaran

Model Pembelajaran	Metode/ Algoritma
Supervised Learning	Natural Language Processing (NLP) (Fortunatus dkk., 2020) Support Vector Machine (SVM) (I. Dilrukshi & K. De Zoysa, 2013) Deep Learning (Long dkk., 2021; T. Mahlangu & C. Tu, 2019)
Unsupervised Learning	Naïve Bayes (I. Dilrukshi & K. De Zoysa, 2013; M. Shirakawa dkk., 2015) K-Means (Vouros & Vasilaki, 2021)
Sentiment Analysis	Semantic-similarity (M. Shirakawa dkk., 2015; R. Devika dkk., 2021) Token-wise Sentiment (Chawla dkk., 2020)
Fuzzy Logic	Fuzzy logic (Mai & Ngo, 2018)

Mesin pembelajaran dapat melakukan proses klasifikasi atas dataset, yang kemudian dapat menghasilkan *output* berdasarkan model/ hipotesis yang dibangun. Proses klasifikasi pada mesin pembelajaran dikelompokkan sebagai pembelajaran *supervised* (Ozbay & Alatas, 2020), dimana *supervised* bekerja dengan membangun sebuah kelompok berdasarkan label pembelajaran, yang kemudian dataset diujikan dengan melakukan perbandingan dan performansi data uji dengan data latih, sehingga data tersebut menghasilkan sebuah label. Proses deteksi *cyberbullying* juga dapat dilakukan dengan pemodelan *semi-supervised*, dimana dataset yang digunakan merupakan campuran data berlabel maupun data yang tidak berlabel. *Semi-supervised* pada deteksi *cyberbullying* juga dapat dilakukan dengan mengimplementasikan algoritma Fuzzy, dimana dapat ditentukan dengan menentukan tingkat keabu-abuan data, ataupun dapat dilakukan

dengan Algoritma K-Means dengan fungsi probabilitas atau dengan *rule-based* (Tachicart & Bouzoubaa, 2021).

Natural Language Processing (NLP) juga dapat digunakan sebagai mesin pembelajaran deteksi *cyberbullying*. NLP pada proses pembelajaran melakukan pembentukan model sebagai bentuk *systematic* yang menghasilkan sebuah peta pembelajaran. (Fiok dkk., 2021).

Naïve Bayes merupakan algoritma probabilitas klasifikasi berdasarkan hipotesis. Output yang dihasilkan oleh Naïve Bayes merupakan asumsi yang memungkinkan berdasarkan hipotesis sebelumnya sebagai basis pengetahuan (S. Hassan dkk., 2011).

Berdasarkan tahapan-tahapan dari masing-masing metode atau model deteksi *cyberbullying*, terdapat peluang untuk mengembangkan metode deteksi, sehingga terdapat keterbaruan penelitian berupa tahapan yang mampu untuk meningkatkan akurasi deteksi, khususnya deteksi *cyberbullying*.

### 3. DISKUSI, PELUANG, DAN RENCANA PENELITIAN

Proses deteksi *cyberbullying* dapat dilakukan dengan melakukan persiapan dataset, melakukan analisis fitur, dan melakukan proses pembelajaran (klasifikasi *cyberbullying*).

#### 3.1. Data Collection dan Analisis Fitur

Proses pengumpulan data (*Data collection*) pada penelitian dapat dilakukan dengan mengimplementasikan Dataset yang bersumber dari ujaran/ komentar/ obrolan dari berbagai platform media sosial, seperti; Twitter, Facebook, Youtube, dll. Dimana dari dataset yang dikumpulkan diharapkan memiliki kata kunci (keyword) atau frase ucapan yang dapat digunakan untuk mengelompokkan/ mengklasifikasikan dalam proses deteksi *cyberbullying*. Proses pengumpulan data dapat dilakukan dengan melakukan penarikan data API yang tersedia dari platform media sosial, kemudian data yang dapat dilakukan proses penarikan dilanjutkan dengan proses ekstraksi untuk penentuan features dari data collection. Fitur-fitur yang diekstraksi adalah sebagai berikut:

1. Fitur teks/konten –Fitur yang meliputi jumlah karakter, karakter huruf besar, karakter huruf kecil, tagar, simbol, sebutan pengguna, URL, dan media.
2. Fitur pengguna –Ini merujuk pada fitur yang diambil dari profil masing-masing individu seperti usia akun, akun yang diverifikasi, jumlah status (yaitu, jumlah tweet (termasuk re-tweet) yang diposting oleh pengguna), jumlah daftar (mis, jumlah daftar publik tempat pengguna menjadi anggotanya), dan jumlah favorit pengguna (jumlah tweet yang disukai pengguna selama masa pakai akunnya).

3. Fitur jaringan –Fitur yang mengukur keramahan pengguna di platform masing-masing seperti jumlah pengikut dan pengikut, dan popularitas (rasio pengikut-pengikut).

Adapun data yang sudah dilakukan proses ekstraksi dilanjutkan dengan proses pembersihan (cleaning) dan proses normalisasi data (normalization) sehingga menghasilkan dataset yang akan digunakan pada proses berikutnya. Pada tahapan pre-processing dilakukan untuk mengetahui features yang terdapat pada data sehingga mampu diolah pada mesin pembelajaran.

Features dari dataset yang dihasilkan diharapkan mengandung kata-kata *bullying* (abusive words), tanda baca (punction) dan kamus leksikal (lexical dictionary) yang dapat digunakan dalam proses deteksi *cyberbullying*. Tiga metode umum yang ada dalam deteksi emosi tekstual, yaitu, berbasis kata kunci (yaitu, menggunakan sinonim dan antonim dari kamus), berbasis pembelajaran (berdasarkan pengklasifikasi yang dilatih sebelumnya) dan hibrida (kombinasi kata kunci dan metode pembelajaran).

#### 3.2. Klasifikasi Cyberbullying

Berbagai metode pembelajaran dan algoritma telah banyak dikembangkan dan diimplementasikan untuk melakukan proses deteksi *cyberbullying*. Proses deteksi *cyberbullying* dapat dilakukan dengan klasifikasi maupun dengan perangkan dari dataset. Adapun proses klasifikasi pada deteksi *cyberbullying* dapat dilakukan dengan; Naïve Bayes, K-Nearest Neighbour, Support Vector Machine (SVM), Deep Learning, maupun Natural Language Processing (NLP).

Proses klasifikasi dilakukan dengan mengimplementasikan Supervised Learning untuk melakukan proses klasifikasi dengan membangun dimensi 2 atas word cloud dan frekuensi kemunculan kata yang telah diberikan label. Deteksi *cyberbullying* dengan klasifikasi mesin pembelajaran dapat dilakukan dengan membangun fitur, yang kemudian secara spesifik dapat melakukan deteksi dari model pembelajaran yang sudah dibangun dan dilakukan perbandingan yang relevan. Disamping itu, terdapat juga model pembelajaran untuk deteksi *cyberbullying* dengan Support Vector Machine (SVM). SVM berkerja dengan melakukan model pembelajaran dengan membuat Hyperplane (Decision Boundry) sebagai pemisah kelas dari label. Dataset yang diujikan akan memiliki label berdasarkan wilayah yang dibagi oleh garis Hyperplane yang dibuat saat model pembelajaran. Deteksi *cyberbullying* dapat dilakukan dengan SVM membuat aturan diskriminasi dan dapat dilakukan pada permasalahan multidimensi yang dibangun dengan membuat kernel sebagai penentu garis Hyperplane SVM. Proses klasifikasi dan deteksi *cyberbullying* selain menggunakan metode dari machine learning, juga terdapat metode/ teknik lainnya, seperti Deep Learning yang dapat dilakukan untuk melakukan proses klasifikasi teks,

rekognisi entitas-entitas, melakukan proses parsing, sentiment analysis, dan beberapa pendekatan teknik Natural Language Processing (NLP). Proses penentuan keyphrase dengan Deep Learning juga dilakukan dengan proses pre-train terlebih dahulu untuk memberikan penanda pada keyphrase.

Komentar atau ujaran yang dituliskan pada media sosial dapat ditangkap guna mengetahui kata-kata yang tertanam pada suatu konteks. Hal ini dilakukan guna mendukung proses deteksi klasifikasi pada mesin pembelajaran. Proses deteksi *cyberbullying* tidak hanya dapat dilakukan dengan memodelkan mesin pembelajaran, akan tetapi proses deteksi juga dapat dilakukan dengan melakukan pendeteksian dengan proses Sentimen Analysis, dimana aturan model dibangung dengan pemodelan contoh/ token yang dibandingkan dengan kata-kata yang tertanam pada konteks.

Proses pengujian klasifikasi deteksi *cyberbullying* dilakukan dengan melakukan pengukuran efisiensi (accuracy), kebenaran proses latih dan uji (precision), perbandingan hasil positif dan negatif (recall), dan perbandingan efisiensi paramater dengan paramater lainnya (F Score) dari proses yang telah dilakukan.

#### 4. KESIMPULAN

Berdasarkan dari studi literatur dinyatakan bahwa; proses deteksi *cyberbullying* di media sosial masih sangat terbuka dan masih dapat dikembangkan kedepannya. Pada penentuan pola karakteristik dari *cyberbullying* masih terdapat perubahan dari waktu ke waktu, hal ini diakibatkan semakin meluasnya satu individu atau kelompok untuk dapat berinteraksi dan berkomunikasi dengan individu lainnya di dunia maya. Pola karakteristik inilah yang dapat dijadikan *features* dari *cyberbullying*. Dari sisi metode atau algoritma, masih terdapat keunggulan dan kelemahan dari masing-masing algoritma dalam penyelesaian proses deteksi *cyberbullying*.

Features *cyberbullying* dapat dikembangkan secara adaptive dan dinamis yang diharapkan dapat melakukan deteksi *cyberbullying* pada beberapa domain yang berbeda. Perbedaan domain bahasa, domain budaya atau daerah, dan domain platform digital menyebabkan sulitnya sebuah metode untuk dapat menyelesaikan permasalahan yang beragam ini. Sehingga terdapat peluang untuk membandingkan, menggabungkan, dan mengintegrasikan beberapa metode guna mendapatkan efektivitas proses pendeteksian *cyberbullying* dari berbagai sudut domain permasalahan.

Adapun tantangan yang mungkin dihadapi kedepannya yaitu; memahami kata-kata yang tertanam pada suatu konteks untuk dapat dilakukan klasifikasi dengan domain bahasa yang berbeda. Adapun tantangan lainnya berupa mempelajari sudut pandang disiplin lainnya, yang diharapkan mampu memperkaya pengetahuan dari mesin pembelajaran yang akan diterapkan nantinya.

#### DAFTAR PUSTAKA

- BARLETT, C. P. 2019. Chapter 2—Cyberbullying, Traditional Bullying, and Aggression: A Complicated Relationship. Dalam C. P. Barlett (Ed.), Predicting Cyberbullying (hlm. 11–16). Academic Press. <https://doi.org/10.1016/B978-0-12-816653-6.00002-9>
- BOZYIĞIT, A., UTKU, S., & NASIBOV, E. 2021. Cyberbullying detection: Utilizing social media features. *Expert Systems with Applications*, 179, 115001. <https://doi.org/10.1016/j.eswa.2021.115001>
- CHAWLA, P., HAZARIKA, S., & SHEN, H.-W. 2020. Token-wise sentiment decomposition for ConvNet: Visualizing a sentiment classifier. *PacificVis 2020 Workshop on Visualization Meets AI*, 4(2), 132–141. <https://doi.org/10.1016/j.visinf.2020.04.006>
- CHIA, Z. L., PTASZYNSKI, M., MASUI, F., LELIWA, G., & WROCZYNSKI, M. 2021. Machine Learning and feature engineering-based study into sarcasm and irony classification with application to cyberbullying detection. *Information Processing & Management*, 58(4), 102600. <https://doi.org/10.1016/j.ipm.2021.102600>
- FIOK, K., KARWOWSKI, W., GUTIERREZ, E., & WILAMOWSKI, M. 2021. Analysis of sentiment in tweets addressed to a single domain-specific Twitter account: Comparison of model performance and explainability of predictions. *Expert Systems with Applications*, 186, 115771. <https://doi.org/10.1016/j.eswa.2021.115771>
- FORTUNATUS, M., ANTHONY, P., & CHARTERS, S. 2020. Combining textual features to detect cyberbullying in social media posts. *Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 24th International Conference KES2020*, 176, 612–621. <https://doi.org/10.1016/j.procs.2020.08.063>
- I. DILRUKSHI & K. DE ZOYSA. 2013. Twitter news classification: Theoretical and practical comparison of SVM against Naive Bayes algorithms. *2013 International Conference on Advances in ICT for Emerging Regions (ICTer)*, 278–278. <https://doi.org/10.1109/ICTer.2013.6761192>
- LI, Q. 2007. New bottle but old wine: A research of cyberbullying in schools. *Computers in Human Behavior*, 23(4), 1777–1791. <https://doi.org/10.1016/j.chb.2005.10.005>
- LONG, S., HE, X., & YAO, C. 2021. Scene Text Detection and Recognition: The Deep

- Learning Era. *International Journal of Computer Vision*, 129(1), Art. 1. <https://doi.org/10.1007/s11263-020-01369-0>
- LÓPEZ-VIZCAÍNO, M. F., NÓVOA, F. J., CARNEIRO, V., & CACHEDA, F. 2021. Early detection of cyberbullying on social media networks. *Future Generation Computer Systems*, 118, 219–229. <https://doi.org/10.1016/j.future.2021.01.006>
- M. SHIRAKAWA, K. NAKAYAMA, T. HARA, & S. NISHIO. 2015. Wikipedia-Based Semantic Similarity Measurements for Noisy Short Texts Using Extended Naive Bayes. *IEEE Transactions on Emerging Topics in Computing*, 3(2), 205–219. <https://doi.org/10.1109/TETC.2015.2418716>
- MAI, S. D., & NGO, L. T. 2018. Multiple kernel approach to semi-supervised fuzzy clustering algorithm for land-cover classification. *Engineering Applications of Artificial Intelligence*, 68, 205–213. <https://doi.org/10.1016/j.engappai.2017.11.007>
- MURNION, S., BUCHANAN, W. J., SMALES, A., & RUSSELL, G. 2018. Machine learning and semantic analysis of in-game chat for cyberbullying. *Computers & Security*, 76, 197–213. <https://doi.org/10.1016/j.cose.2018.02.016>
- OZBAY, F. A., & ALATAS, B. 2020. Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications*, 540, 123174. <https://doi.org/10.1016/j.physa.2019.123174>
- R. DEVIKA, S. VAIRAVASUNDARAM, C. S. J. MAHENTHAR, V. VARADARAJAN, & K. KOTECHA. 2021. A Deep Learning Model Based on BERT and Sentence Transformer for Semantic Keyphrase Extraction on Big Social Data. *IEEE Access*, 9, 165252–165261. <https://doi.org/10.1109/ACCESS.2021.3133651>
- RAJPUT, A. 2020. Chapter 3—Natural Language Processing, Sentiment Analysis, and Clinical Analytics. Dalam M. D. Lytras & A. Sarirete (Ed.), *Innovation in Health Informatics* (hlm. 79–97). Academic Press. <https://doi.org/10.1016/B978-0-12-819043-2.00003-4>
- RUSSEL, S., & NORVIG, P. 2021. *Artificial Intelligence A Modern Approach*. Pearson.
- S. HASSAN, M. RAFI, & M. S. SHAIKH. 2011. Comparing SVM and naïve Bayes classifiers for text categorization with Wikitology as knowledge enrichment. 2011 IEEE 14th International Multitopic Conference, 31–34. <https://doi.org/10.1109/INMIC.2011.6151495>
- SHELDON, P., RAUSCHNABEL, P. A., & HONEYCUTT, J. M. 2019. Chapter 3—Cyberstalking and Bullying. Dalam P. Sheldon, P. A. Rauschnabel, & J. M. Honeycutt (Ed.), *The Dark Side of Social Media* (hlm. 43–58). Academic Press. <https://doi.org/10.1016/B978-0-12-815917-0.00003-4>
- T. MAHLANGU & C. TU. 2019. Deep Learning Cyberbullying Detection Using Stacked Embeddings Approach. 2019 6th International Conference on Soft Computing & Machine Intelligence (ISCMI), 45–49. <https://doi.org/10.1109/ISCMI47871.2019.9004292>
- TACHICART, R., & BOUZOUBAA, K. 2021. Moroccan Arabic vocabulary generation using a rule-based approach. *Journal of King Saud University - Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2021.02.013>
- T.K., B., ANNAVAPU, C. S. R., & BABLANI, A. 2021. Machine learning algorithms for social media analysis: A survey. *Computer Science Review*, 40, 100395. <https://doi.org/10.1016/j.cosrev.2021.100395>
- VOUROS, A., & VASILAKI, E. 2021. A semi-supervised sparse K-Means algorithm. *Pattern Recognition Letters*, 142, 65–71. <https://doi.org/10.1016/j.patrec.2020.11.015>
- W. M. AL-RAHMI, N. YAHAYA, M. M. ALAMRI, N. A. ALJARBOA, Y. B. KAMIN, & F. A. MOAFA. 2019. A Model of Factors Affecting Cyber Bullying Behaviors Among University Students. *IEEE Access*, 7, 2978–2985. <https://doi.org/10.1109/ACCESS.2018.2881292>
- W. M. AL-RAHMI, N. YAHAYA, M. M. ALAMRI, N. A. ALJARBOA, Y. B. KAMIN, & M. S. B. SAUD. 2019. How Cyber Stalking and Cyber Bullying Affect Students' Open Learning. *IEEE Access*, 7, 20199–20210. <https://doi.org/10.1109/ACCESS.2019.2891853>