

Article history

Received July 22, 2024

Accepted Nov 30, 2024

Published Nov 30, 2024

ANALISIS SENTIMEN MASYARAKAT TERHADAP KESEHATAN MENTAL PADA MEDIA SOSIAL TWITTER DENGAN MENGGUNAKAN MACHINE LEARNING

Hudatul Aulia¹⁾, Muhammad Zulfadhilah¹⁾, Septyan Eka Prastya¹⁾, Muhammad Syahid Pebriadi²⁾

¹ Program Studi Sarjana Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Sari Mulia

² Program Studi Komputerisasi Akuntansi, Jurusan Akuntansi, Politeknik Negri Banjarmasin

email: auliahuda96@gmail.com

Abstract

Mental health affects lives globally, with around 300 million people experiencing depression in 2019, including 15.6 million in Indonesia. The Covid-19 pandemic increased cases of anxiety and depression, and by 2022, WHO reported 23 million people suffering from psychiatric disorders. In Indonesia, adolescent mental health issues are also high, with excessive social media use linked to an increase in emotional disorders. Twitter, with its real-time data, is becoming an important tool for analyzing public sentiment and understanding opinions through analytics and machine learning techniques. This study aims to determine public sentiment towards mental health in Indonesia through Twitter social media and test the effectiveness of using machine learning in sentiment analysis. The results show that the Naive Bayes and Decision Tree methods are effective in analyzing sentiment, with an accuracy of 91% and 89% respectively. The average result of cross validation shows a value of 73.21% for Naive Bayes and 67.02% for Decision Tree. In this study, positive sentiment is more dominant with a percentage value of 78.7%, while negative sentiment is only 21.3%. The findings indicate that Indonesians' awareness of the importance of mental health is increasing, and they increasingly understand the importance of maintaining mental health.

Keywords: *decision tree, mental health, machine learning, naive bayes.*

Abstrak

Kesehatan mental sangat mempengaruhi kehidupan global, dengan sekitar 300 juta orang mengalami depresi pada 2019, termasuk 15,6 juta di Indonesia. Pandemi Covid-19 meningkatkan kasus kecemasan dan depresi, dan pada 2022, WHO melaporkan 23 juta orang menderita gangguan kejiwaan. Di Indonesia, masalah kesehatan mental remaja juga tinggi, dengan penggunaan media sosial berlebihan terkait peningkatan gangguan emosional. Twitter, dengan data *real-time*, menjadi alat penting untuk menganalisis sentimen publik dan memahami opini melalui teknik analisis dan *machine learning*. Penelitian ini bertujuan untuk mengetahui sentimen masyarakat terhadap kesehatan mental di Indonesia melalui media sosial Twitter serta menguji keefektifan penggunaan *machine learning* dalam analisis sentimen. Hasil penelitian menunjukkan bahwa metode *Naive Bayes* dan *Decision Tree* efektif dalam menganalisis sentimen, dengan akurasi sebesar 91% dan 89% masing-masing. Rerata hasil dari *cross validation* menunjukkan nilai 73,21% untuk *Naive Bayes* dan 67,02% untuk *Decision Tree*. Pada penelitian ini sentimen positif lebih dominan dengan persentase nilai 78,7%, sementara sentimen negatif hanya 21,3%. Temuan ini mengindikasikan bahwa kesadaran masyarakat Indonesia tentang pentingnya kesehatan mental semakin meningkat, dan mereka semakin memahami pentingnya menjaga kesehatan mental.

Kata Kunci: *decision tree, kesehatan mental, machine learning, naive bayes.*

1. INTRODUCTION

Keadaan mental yang sehat sangat penting dan berpengaruh besar pada kehidupan setiap individu [1]. Menurut data Organisasi Kesehatan Dunia (WHO) pada tahun 2019, sekitar 300 juta orang di seluruh dunia mengalami depresi, termasuk 15,6 juta penduduk Indonesia [2]. Pada tahun 2020, gangguan kecemasan meningkat sebesar 26% dan depresi sebesar 28% akibat pandemi Covid-19 [3]. Pada tahun 2022, WHO menyatakan ada sekitar 23 juta orang yang menderita gangguan kejiwaan seperti skizofrenia atau psikosis. Dari jumlah tersebut, hanya 31,3% yang menerima layanan dari spesialis jiwa [4].

Di Indonesia, masalah kesehatan mental juga tinggi di kalangan remaja. Berdasarkan data Survei Kesehatan Jiwa Remaja Nasional Indonesia tahun 2022, terdapat 15,5 juta (34,9%) remaja yang mengalami masalah kejiwaan, dengan 2,45 juta (5,5%) di antaranya menderita gangguan jiwa. Dari jumlah tersebut, hanya 2,6% yang menggunakan layanan konseling untuk masalah emosional dan perilaku [5]. Penelitian yang dilakukan merujuk pada penelitian sebelumnya yang mengaitkan media sosial dengan kesehatan mental [6]–[8].

Data tersebut tentu menjadi perbincangan di kalangan pengguna media sosial. Survei Kesehatan Jiwa Remaja Nasional Indonesia tahun 2022 memperkirakan bahwa setidaknya sepertiga anak usia 10 hingga 13 tahun mengalami masalah kesehatan jiwa. Angka serupa juga ditemukan pada remaja berusia 14 hingga 17 tahun. Jika dianalisis berdasarkan jenisnya, setidaknya 10,6% generasi milenial memiliki masalah terkait *Attention Deficit Hyperactivity Disorder* (ADHD). Semakin lama remaja menggunakan media sosial, semakin berdampak pada kesehatan mental mereka, karena mereka cenderung fokus pada diri sendiri atau dunianya sendiri dan menjadi kecanduan menggunakan media sosial [9], [10].

Twitter adalah *platform* media sosial yang banyak digunakan di seluruh dunia untuk berkomunikasi. Twitter menyediakan data dalam jumlah besar yang dapat diakses oleh masyarakat umum secara *real-time*, sehingga dapat melacak tren dan peristiwa yang sedang ramai diperbincangkan [11]. Berdasarkan demografi pengguna Twitter di Indonesia, 53% pengguna adalah laki-laki dan 47% perempuan, mayoritas berusia 16-22 tahun. Pengguna Twitter di Indonesia aktif sebesar 42% setiap hari [12].

Hal ini membuka peluang untuk melakukan analisis sentimen menggunakan data dari media sosial Twitter. Analisis sentimen adalah proses menganalisis teks digital untuk menentukan apakah teks tersebut bernilai positif atau negatif. Biasanya dilakukan dengan mengidentifikasi kata dan frasa positif atau negatif untuk menemukan pola penggunaan kata tersebut [13]. Tujuannya adalah mengetahui pendapat, perilaku, atau perasaan seseorang yang diungkapkan dalam tulisan seperti postingan media sosial, komentar, survei, atau bentuk tulisan lainnya [14]. Informasi di jejaring sosial Twitter umumnya tidak terstruktur sehingga data harus diproses ulang menggunakan pre-processing dan pelabelan untuk mengumpulkan indeks emosi atau mengklasifikasikan tingkat emosi dalam teks. Untuk melakukan analisis sentimen diperlukan algoritma yang mendukung seperti *machine learning*.

Contoh penggunaan algoritma *machine learning* terdapat pada penelitian yang dilakukan oleh Harahap, mengenai Analisis Sentimen Review Penjualan Produk UMKM di Kabupaten Nias yang menggunakan metode *Naive Bayes* (NB), *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), dan *Artificial Neural Network* (ANN) dengan akurasi mencapai 96% [15]. Penelitian yang dilakukan oleh Yuniar, mengenai Implementasi *Scraping* Data untuk Sentimen Analisis Pengguna Dompot Digital menunjukkan sentimen positif dari akun Twitter Gopay sebesar 79,6% dan LinkAja sebesar 62,1%. Sentimen negatif dari Gopay adalah 20,4% dan dari LinkAja adalah 37,9%. Hasil pengujian juga mencatat *recall* 87% untuk Gopay dan 89% untuk LinkAja [16]. Dan penelitian yang dilakukan oleh Putra & Santika, mengenai Implementasi *Machine Learning* dalam Penentuan Rekomendasi Musik dengan Metode *Content-Based Filtering* menghasilkan nilai *recall* antara 0,150 hingga 0,200 dan presisi antara 0,015 hingga 0,125, memastikan rekomendasi musik sesuai dengan preferensi pengguna [17].

Penelitian ini bertujuan untuk mengetahui sentimen masyarakat terhadap kesehatan mental di Indonesia melalui media sosial Twitter, serta menguji apakah *machine learning* dapat digunakan untuk menganalisis sentimen tersebut. Hasil analisis sentimen masyarakat terhadap kesehatan mental di Indonesia melalui Twitter akan dievaluasi menggunakan teknik *machine learning*.

2. LITERATURE REVIEW

Penelitian tentang kesehatan mental di Indonesia menunjukkan bahwa kesehatan mental yang sehat sangat penting dan berpengaruh pada kehidupan individu [1]. Terdapat sekitar 300 juta orang di dunia mengidap depresi, termasuk 15,6 juta di Indonesia [2]. Gangguan kecemasan dan depresi meningkat signifikan akibat pandemi Covid-19 [3], sementara hanya 31,3% penderita gangguan kejiwaan yang menerima layanan spesialis jiwa [4].

Kesehatan mental remaja di Indonesia juga menjadi perhatian utama. Data Survei Kesehatan Jiwa Remaja Nasional Indonesia tahun 2022 menunjukkan 15,5 juta (34,9%) remaja menderita masalah kejiwaan, dengan 2,45 juta (5,5%) di antaranya mengalami gangguan jiwa. Namun, hanya 2,6% dari mereka yang menggunakan layanan konseling untuk masalah emosional dan perilaku [5]. Persentase masalah mental tertinggi dalam rentang 2020-2022 ditemukan di Nusa Tenggara Barat (43.3%), Yogyakarta (34.7%), Jawa Tengah (33.6%), Gorontalo (33.1%), dan Kalimantan Selatan (33%) [18].

Penelitian ini didasarkan pada penelitian sebelumnya yang meneliti hubungan antara media sosial dan kesehatan mental [6]–[8]. Survei Kesehatan Jiwa Remaja Nasional Indonesia tahun 2022 memperkirakan bahwa sepertiga anak usia 10-13 tahun mengalami masalah kesehatan jiwa, dengan angka serupa pada remaja berusia 14-17 tahun. Sekitar 10,6% generasi milenial memiliki masalah terkait *Attention Deficit Hyperactivity Disorder* (ADHD). Penggunaan media sosial yang berlebihan dapat berdampak negatif pada kesehatan mental remaja, menyebabkan kecanduan dan isolasi [9], [10].

Twitter adalah *platform* media sosial yang populer dan menyediakan data dalam jumlah besar yang dapat diakses secara *real-time* [11]. Berdasarkan demografi pengguna Twitter di Indonesia, 53% pengguna adalah laki-laki dan 47% perempuan, mayoritas berusia 16-22 tahun, dengan 42% pengguna aktif setiap hari [12]. Hal ini membuka peluang untuk melakukan analisis sentimen menggunakan data dari Twitter.

Analisis sentimen adalah proses menganalisis teks digital untuk menentukan apakah teks tersebut bernilai positif atau negatif, biasanya dengan mengidentifikasi kata dan frasa yang bersifat positif atau negatif [14]. Data Twitter yang tidak terstruktur memerlukan

preprocessing dan pelabelan untuk mengumpulkan indeks emosi atau mengklasifikasikan tingkat emosi dalam teks. Algoritma machine learning mendukung analisis sentimen ini.

Ada beberapa contoh penerapan algoritma *machine learning*, seperti dalam penelitian yang dilakukan oleh Harahap [15] menggunakan algoritma *machine learning* seperti *Naive Bayes* (NB), *Support Vector Machine* (SVM), *K-Nearest Neighbor* (KNN), dan *Artificial Neural Network* (ANN) untuk analisis sentimen *review* penjualan produk UMKM, menghasilkan akurasi 96%. Penelitian yang dilakukan oleh Yuniar [16] melakukan analisis sentimen pengguna Dompot Digital menggunakan algoritma *machine learning*, menemukan sentimen positif pada akun Twitter Gopay sebesar 79,6% dan LinkAja 62,1%. Dan penelitian yang dilakukan oleh Putra & Santika [17] menggunakan *machine learning* untuk penentuan rekomendasi musik, menghasilkan *recall* antara 0,150-0,200 dan presisi antara 0,015-0,125, menunjukkan rekomendasi sesuai preferensi pengguna.

Literature review ini menyoroti pentingnya memahami kondisi kesehatan mental global dan lokal, terutama di kalangan remaja Indonesia. Penggunaan media sosial, khususnya Twitter, memberikan data berharga untuk analisis sentimen. Implementasi algoritma *machine learning* dalam analisis sentimen menunjukkan potensi besar dalam mengungkap pola dan tren dalam data media sosial, yang dapat digunakan untuk intervensi yang lebih baik dalam kesehatan mental. Penelitian lebih lanjut diperlukan untuk terus mengevaluasi dan memperbaiki metode yang ada guna mendukung kesehatan mental yang lebih baik melalui teknologi dan data digital.

3. HASIL DAN PEMBAHASAN

Pengumpulan Data

Data yang digunakan pada penelitian ini adalah *tweet* masyarakat dari media sosial twitter. Data dikumpulkan menggunakan metode *crawling* dengan memanfaatkan Google Colab dan fungsi *tool tweet-harvest*, diperlukan *auth_token* sebagai akses masuk pada halaman pencarian Twitter. Data yang berhubungan dengan kata kunci “kesehatan mental” dalam rentang waktu 1 januari 2021 sampai 29 februari 2024 di *crawling* dan disimpan dalam bentuk *file csv*. Secara rinci pembagian rentang waktu

lebih baik tentang perilaku model atau algoritma yang digunakan, memacu kemajuan dalam studi kesehatan mental. Disarankan pula untuk menguji dengan menggunakan metode lain serta memperluas penelitian dengan data dari *platform* media sosial lain. Pendekatan yang beragam ini diharapkan dapat memperdalam pemahaman tentang keandalan dan validitas hasil, serta memperluas cakupan informasi untuk mendapatkan wawasan yang lebih komprehensif mengenai perilaku, preferensi, dan tren terkait kesehatan mental.

5. REFERENSI

- [1] Kemenkes, “Pengertian Kesehatan Mental,” *kemkes.go.id*, 2018. <https://ayosehat.kemkes.go.id/pengertian-kesehatan-mental> (diakses 30 Agustus 2023).
- [2] Itsbil, “Depresi, Kesehatan Mental yang Tak Boleh Disepelekan,” *Kampus ITS*, 2023. https://www.its.ac.id/news/2023/05/22/___trashed-5/ (diakses 24 Agustus 2023).
- [3] Z. Prihatini, “WHO: Hampir 1 Miliar Orang di Dunia Alami Gangguan Kesehatan Mental,” *Kompas.com*, 2022. <https://www.kompas.com/sains/read/2022/06/20/193000823/who--hampir-1-miliar-orang-di-dunia-alami-gangguan-kesehatan-mental?page=all> (diakses 24 Agustus 2023).
- [4] P. Pandu, “Aplikasi untuk Deteksi Dini Psikosis,” *Kompas.id*, 2022. <https://www.kompas.id/baca/humaniora/2022/11/13/inovasi-iptek-aplikasi-deteksi-dini-psikosis> (diakses 24 Agustus 2023).
- [5] A. Arif, “Krisis Kesehatan Mental Melonjak di Kalangan Remaja,” *Kompas.id*, 2023. <https://www.kompas.id/baca/humaniora/2023/05/03/krisis-kesehatan-mental-melonjak-di-kalangan-remaja> (diakses 24 Agustus 2023).
- [6] S. Kaur, K. Kaur, Aprajita, R. Verma, dan Pangkaj, “Impact of Social Media on Mental Health of Adolescents,” *J. Pharm. Negat. Results*, vol. 13, no. 5, hal. 779–783, 2022, doi: 10.47750/pnr.2022.13.S05.121.
- [7] N. Z. Septiana, “Dampak Penggunaan Media Sosial Terhadap Kesehatan Mental Dan Kesejahteraan Sosial Remaja Dimasa Pandemi Covid-19,” *Nusant. Res. J. Hasil-hasil Penelit. Univ. Nusant. PGRI Kediri*, vol. 8, no. 1, hal. 1–13, 2021, doi: 10.29407/nor.v8i1.15632.
- [8] R. Al Yasin, R. R. K. A. Anjani, S. Salsabil, T. Rahmayanti, dan R. Amalia, “Pengaruh Sosial Media Terhadap Kesehatan Mental Dan Fisik Remaja: a Systematic Review,” *J. Kesehat. Tambusai*, vol. 3, no. 2, hal. 83–90, 2022, doi: 10.31004/jkt.v3i2.4402.
- [9] H. D. Jayanti *dkk.*, “2314-8101-1-Pb,” hal. 91–102, 2022.
- [10] A. Rosmalina dan T. Khaerunnisa, “Penggunaan Media Sosial dalam Kesehatan Mental Remaja,” *Prophet. Prof. Empathy, Islam. Couns. J.*, vol. 4, no. 1, hal. 49, 2021, doi: 10.24235/prophetic.v4i1.8755.
- [11] S. Sadya, “Indonesia Masuk Negara Paling Banyak Main Twitter pada Awal 2023,” *dataindonesia.id/*, 2023. <https://dataindonesia.id/internet/detail/indonesia-masuk-negara-paling-banyak-main-twitter-pada-awal-2023> (diakses 30 Agustus 2023).
- [12] Adam, “Demografi Pengguna Twitter di Indonesia Paling Banyak Pria daripada Perempuan,” *itworks.id*, 2019. <https://www.itworks.id/19408/demografi-pengguna-twitter-di-indonesia-paling-banyak-pria-daripada-perempuan.html> (diakses 30 Agustus 2023).
- [13] E. T. Pusparini, “Mengenal Apa Itu Analisis Sentimen, Tipe dan Cara Kerjanya,” *qontak.com*, 2023. <https://qontak.com/blog/analisis-sentimen-adalah/> (diakses 4 September 2023).
- [14] F. V. Sari dan A. Wibowo, “Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi,” *J. SIMETRIS*, vol. 10, no. 2, hal. 681–686, 2019.
- [15] M. Harahap, B. P. A. Sihombing, O. A. F. Laia, B. T. Saragih, F. Teknologi, dan U. P. Indonesia, “ANALISIS SENTIMEN REVIEW PENJUALAN PRODUK UMKM PADA MACHINE LEARNING Kuandi Dharma,” vol. 5, no. 2, hal. 147–154, 2021.
- [16] E. Yuniar, D. Safiroh, dan D. Wahyuningsih, “Implementasi Scraping Data Untuk Sentiment Analysis Pengguna Dompot Digital Dengan Menggunakan Algoritma

- Machine Learning Implementation of Data Scraping for Sentiment Analysis of Digital,” vol. 2, no. 1, hal. 35–42, 2022, doi: 10.25008/janitra.v2i1.145.
- [17] A. I. Putra dan R. R. Santika, “Implementasi Machine Learning dalam Penentuan Rekomendasi Musik dengan Metode Content-Based Filtering,” *Edumatic J. Pendidik. Inform.*, vol. 4, no. 1, hal. 121–130, 2020, doi: 10.29408/edumatic.v4i1.2162.
- [18] Badan Pusat Statistik, “Persentase Penduduk yang Mempunyai Keluhan Kesehatan Selama Sebulan Terakhir (Persen), 2020-2022,” *Statistik Indonesia*, 2022. <https://www.bps.go.id/indicator/30/222/1/persentase-penduduk-yang-mempunyai-keluhan-kesehatan-selama-sebulan-terakhir.html> (diakses 24 Agustus 2023).
- [19] R. Azhar, A. Surahman, dan C. Juliane, “Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes,” *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 1, hal. 267–281, 2022.
- [20] K. Verena, S. Toy, Y. A. Sari, dan I. Cholissodin, “Analisis Sentimen Twitter menggunakan Metode Naive Bayes dengan Relevance Frequency Feature Selection (Studi Kasus : Opini Masyarakat mengenai Kebijakan New Normal),” vol. 5, no. 11, hal. 5068–5074, 2021.
- [21] R. Adinugroho, *Perbandingan Rasio Split Data Training Dan Data Testing Menggunakan Metode LSTM Dalam Memprediksi Harga Indeks Saham Asia*. 2022. [Daring]. Tersedia pada: [https://repository.uinjkt.ac.id/dspace/handle/123456789/67314%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/67314/1/RAHMADHAN ADINUGROHO-FST.pdf](https://repository.uinjkt.ac.id/dspace/handle/123456789/67314%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/67314/1/RAHMADHAN_ADINUGROHO-FST.pdf)
- [22] A. Tharwat, “Classification assessment methods,” *Appl. Comput. Informatics*, vol. 17, no. 1, hal. 168–192, 2018, doi: 10.1016/j.aci.2018.08.003.