

Analisis Sentimen Kesehatan Mental Pemuda di Media Sosial Menggunakan Deep Learning

Agustin¹⁾, Junadhi²⁾, Fransiskus Zoromi³⁾, Parlindungan Kudadiri^{4)*}

^{1,2,3,4)}Prodi Teknik Informatika, Fakultas Teknik dan Informatika, Universitas Sains dan Teknologi Indonesia, Indonesia

*Coresponding Email: ¹agustin@usti.ac.id, ²junadhi@usti.ac.id, ³fransiskuszoromi@usti.ac.id,
⁴parlindungan@usti.ac.id

Abstrak

Kesehatan mental merupakan isu yang semakin penting di kalangan pemuda Indonesia, terutama dengan meningkatnya ekspresi emosi negatif seperti stres, kelelahan, dan kecemasan yang sering diungkapkan melalui media sosial. Penelitian ini bertujuan untuk menganalisis sentimen kesehatan mental pemuda menggunakan pendekatan deep learning berbasis Long Short-Term Memory (LSTM) terhadap unggahan publik berbahasa Indonesia di platform X (Twitter). Data dikumpulkan melalui proses web scraping dengan kata kunci yang relevan dan kemudian melalui tahapan pra-pemrosesan, pelabelan manual, serta pembagian data menjadi 80% untuk pelatihan dan 20% untuk pengujian. Model LSTM dibangun dengan arsitektur yang terdiri atas embedding layer, LSTM layer, dropout layer, dense layer, dan output layer beraktivasi Softmax untuk tiga kelas sentimen (positif, negatif, dan netral). Hasil penelitian menunjukkan bahwa model mampu mencapai akurasi sebesar 87,4% dengan nilai precision dan recall rata-rata sebesar 0,85, yang menandakan kemampuan tinggi dalam mengenali konteks bahasa informal pemuda di media sosial. Analisis distribusi sentimen menunjukkan dominasi emosi negatif yang berkaitan dengan tekanan akademik dan sosial, sementara sentimen positif menggambarkan semangat dan mekanisme adaptasi diri. Temuan ini membuktikan bahwa LSTM efektif untuk deteksi ekspresi emosional berbasis teks serta berpotensi diterapkan sebagai sistem pemantauan digital bagi kesejahteraan mental generasi muda.

Kata Kunci: Kesehatan Mental Pemuda, Analisis Sentiment, Media Social, LSTM, Deep Learning.

Abstract

Mental health has become an increasingly critical issue among Indonesian youth, reflected in the rising expressions of stress, fatigue, and anxiety frequently shared on social media. This study aims to analyze the mental health sentiments of young people using a deep learning approach based on the Long Short-Term Memory (LSTM) algorithm applied to Indonesian-language posts on the X (Twitter) platform. Data were collected through web scraping using relevant keywords and then processed through several stages, including text preprocessing, manual sentiment labeling, and data partitioning with 80% for training and 20% for testing. The LSTM model was built with an architecture consisting of an embedding layer, LSTM layer, dropout layer, dense layer, and a Softmax output layer for three sentiment classes (positive, negative, and neutral). The experimental results show that the model achieved an accuracy of 87.4% with an average precision and recall of 0.85, indicating a high ability to capture informal linguistic contexts commonly used by youth on social media. Sentiment

distribution analysis reveals that negative emotions dominate, reflecting academic and social pressures, while positive sentiments indicate optimism, motivation, and adaptive coping mechanisms. These findings demonstrate that LSTM is effective for detecting emotional expressions in text and holds potential as a digital monitoring system for supporting the psychological well-being of young generations.

Keywords: Mental Health Youth, Analisis Sentiment, Media Social, LSTM, Deep Learning.

PENDAHULUAN

Kesehatan mental merupakan aspek fundamental dalam perkembangan generasi muda yang produktif dan berdaya saing. Tekanan akademik, tuntutan sosial, dinamika karier, serta ekspektasi untuk selalu tampil sukses di ruang digital menjadi sumber stres dan kecemasan baru bagi banyak pemuda. Data World Health Organization (World Health, 2025) menunjukkan bahwa lebih dari satu dari lima individu berusia 15–24 tahun mengalami gangguan emosional seperti kecemasan, depresi ringan, atau stres kronis. Di Indonesia, survei(Badan Kebijakan Pembangunan Kesehatan, 2023) menegaskan peningkatan prevalensi gangguan mental emosional pada kelompok usia produktif muda.

Seiring dengan meningkatnya penggunaan media sosial, ekspresi emosional pemuda kini lebih banyak terekam dalam bentuk unggahan digital. Platform seperti X (Twitter), Instagram, dan TikTok menjadi ruang utama bagi mereka untuk mengekspresikan kelelahan, kebahagiaan, kekecewaan, maupun motivasi. Uggahan-unggahan ini menggambarkan dinamika psikologis yang autentik dan spontan, sehingga dapat dimanfaatkan sebagai sumber data penting untuk memahami kondisi mental kolektif pemuda Indonesia.

Pendekatan berbasis Artificial Intelligence (AI) dan Natural Language Processing (NLP) memungkinkan peneliti menganalisis jutaan teks dari media sosial secara otomatis. Salah satu metode NLP yang terbukti efektif adalah Long Short-Term Memory (LSTM), varian dari Recurrent Neural Network (RNN) yang mampu menangkap hubungan semantik jangka panjang antar kata. LSTM unggul dalam memahami konteks kalimat berurutan, misalnya membedakan makna antara

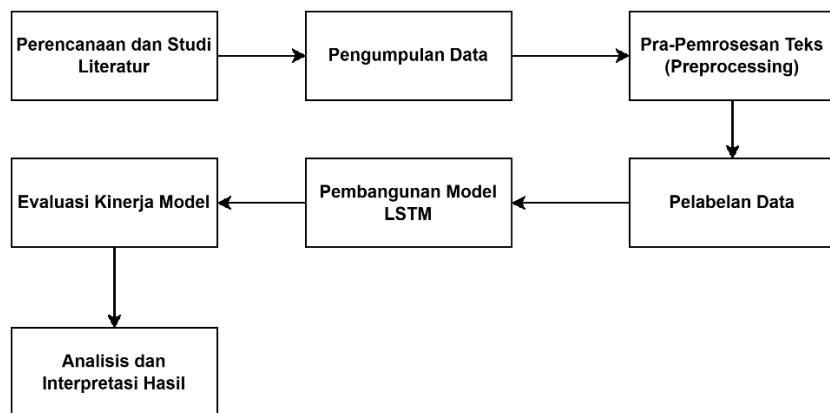
"lelah tapi semangat" dan "lelah dan menyerah," yang sangat penting untuk interpretasi sentimen emosional (Tejaswini et al., 2025).

Berbagai penelitian terdahulu mendukung keunggulan LSTM untuk analisis emosi. (Padmaja et al., 2025) membuktikan bahwa model LSTM dapat memprediksi kecenderungan depresi dari unggahan media sosial dengan akurasi > 85%. Di konteks Bahasa Indonesia, (Melati et al., 2024) berhasil mencapai akurasi 85,6% dalam klasifikasi sentimen berita menggunakan LSTM. Namun, studi yang secara khusus menyoroti sentimen kesehatan mental pemuda Indonesia dalam bahasa informal media sosial masih sangat terbatas. Kebanyakan penelitian lokal hanya mengklasifikasikan teks menjadi positif, negatif, dan netral tanpa menyinggung dimensi psikologis yang lebih spesifik seperti stres, cemas, atau kesepian (Asnal et al., 2025). Padahal, pemetaan ekspresi emosi yang lebih mendalam penting untuk deteksi dini gangguan mental dan perancangan strategi intervensi digital yang berbasis data. Oleh sebab itu, riset ini bertujuan membangun model analisis sentimen kesehatan mental pemuda dengan algoritma LSTM untuk mengidentifikasi pola emosional yang terekam dalam media sosial berbahasa Indonesia (Sembiring et al., 2025).

Penelitian ini diharapkan berkontribusi secara akademik dalam bidang deep learning for social analysis serta memberikan manfaat praktis melalui rekomendasi kebijakan atau sistem pendukung kesejahteraan mental generasi muda.

METODE PENELITIAN

Tahapan penelitian ini disusun secara sistematis agar proses pengembangan model analisis sentimen berbasis Long Short-Term Memory (LSTM) berjalan terarah dan terukur. Setiap tahap memiliki keluaran (output) yang menjadi masukan bagi tahap berikutnya. Secara umum, tahapan penelitian dapat dijelaskan sebagai berikut.

**Gambar 1. Tahapan Penelitian**

Tahap pertama adalah perencanaan dan studi literatur, yang diawali dengan identifikasi masalah tentang meningkatnya ekspresi emosi negatif dan isu kesehatan mental di kalangan pemuda pada media sosial. Peneliti melakukan penelusuran berbagai sumber ilmiah, jurnal, dan hasil penelitian terdahulu terkait analisis sentimen, kesehatan mental digital, serta penerapan LSTM dalam pemrosesan bahasa alami. Hasil studi literatur digunakan untuk membangun kerangka konseptual dan menemukan kesenjangan penelitian (research gap), yaitu belum adanya studi yang secara khusus memanfaatkan model LSTM untuk menganalisis sentimen kesehatan mental pemuda Indonesia dalam bahasa informal yang umum digunakan di media sosial. Pada tahap ini juga ditetapkan tujuan, variabel penelitian, serta desain eksperimen berbasis deep learning.

Tahap kedua adalah pengumpulan data, yang dilakukan dengan mengambil unggahan publik dari platform X (Twitter) menggunakan skrip Python dan pustaka snscreape. Platform ini dipilih karena mayoritas penggunanya adalah pemuda berusia 17–30 tahun dan cenderung mengekspresikan perasaan secara spontan. Pengambilan data dilakukan dengan kata kunci yang berkaitan dengan kondisi psikologis seperti "stres kuliah", "capek hidup", "cemas", "lega", "bahagia", dan "sendirian". Hasil pengambilan data disimpan dalam format CSV yang memuat teks unggahan, tanggal publikasi, serta metadata ringan seperti jumlah likes atau

retweets. Seluruh data yang digunakan bersifat publik dan tidak menyertakan identitas pribadi, sehingga tetap mematuhi prinsip etika penelitian digital.

Tahap ketiga adalah pra-pemrosesan teks (text preprocessing) yang bertujuan menyiapkan data agar dapat dipahami oleh model AI. Proses ini meliputi case folding untuk menyeragamkan huruf menjadi kecil, tokenizing untuk memecah kalimat menjadi kata, stopword removal untuk menghapus kata-kata umum yang tidak memiliki makna sentimen, serta slang normalization untuk mengganti kata tidak baku seperti "gk", "cape", atau "btw" menjadi bentuk formal "tidak", "capek", dan "by the way". Tahapan berikutnya adalah stemming menggunakan pustaka Sastrawi guna mengembalikan kata ke bentuk dasarnya serta pembersihan terhadap URL, emoji, dan simbol non-teks. Hasil akhir dari tahap ini berupa korpus teks bersih (cleaned corpus) yang siap untuk diberi label sentimen (Salam et al., 2023).

Tahap keempat adalah pelabelan data (sentiment labeling), di mana setiap teks diklasifikasikan ke dalam tiga kategori utama, yaitu positif, negatif, dan netral. Label positif diberikan pada teks yang mengandung emosi optimis, bahagia, atau lega; label negatif untuk ekspresi stres, lelah, cemas, atau sedih; sedangkan label netral untuk teks yang bersifat informatif tanpa ekspresi emosional yang jelas. Proses pelabelan dilakukan secara manual oleh tiga annotator dengan latar belakang linguistik dan psikologi. Tingkat kesepakatan antar penilai diuji menggunakan Cohen's Kappa untuk menjamin konsistensi hasil, di mana hanya data dengan nilai Kappa $\geq 0,7$ yang disertakan dalam dataset akhir. Hasil dari tahap ini berupa dataset berlabel (labeled dataset) yang siap digunakan untuk pelatihan model (Rizkia et al., 2025).

Tahap kelima adalah pembangunan model LSTM, yang menjadi inti penelitian ini. Teks yang telah diberi label diubah menjadi representasi numerik menggunakan teknik word embedding seperti Word2Vec atau FastText agar dapat diproses oleh jaringan saraf (Haspin et al., 2025). Dataset kemudian dibagi menjadi dua bagian, yaitu 80% untuk data pelatihan (training data) dan 20% untuk data pengujian (testing data). Arsitektur LSTM terdiri atas beberapa lapisan utama: embedding

layer untuk mengubah kata menjadi vektor, LSTM layer dengan 128 unit untuk menangkap konteks urutan kata, dropout layer dengan nilai 0,2–0,3 untuk mencegah overfitting, dense layer dengan aktivasi ReLU, dan output layer dengan aktivasi Softmax untuk menghasilkan tiga kelas sentimen. Model dilatih menggunakan optimizer Adam dengan loss function categorical cross-entropy selama 10–20 epochs, hingga mencapai kestabilan nilai akurasi dan kehilangan (loss)(Cahyani et al., 2023).

Tahap keenam adalah evaluasi kinerja model, yang dilakukan untuk mengukur kemampuan LSTM dalam mengklasifikasikan sentimen dengan benar. Model diuji menggunakan data testing yang tidak pernah digunakan selama proses pelatihan. Evaluasi dilakukan dengan menghitung metrik performa seperti accuracy, precision, recall, dan F1-score untuk menilai keseimbangan antara ketepatan dan sensitivitas prediksi. Selain itu, digunakan confusion matrix untuk melihat distribusi klasifikasi antar kelas serta ROC curve untuk menilai performa keseluruhan model. Hasil evaluasi ini menjadi dasar dalam menentukan efektivitas model serta area yang dapat ditingkatkan(Learning et al., 2025).

Tahap ketujuh adalah analisis dan interpretasi hasil, di mana data prediksi model diolah untuk memahami pola emosi dominan yang terekam di media sosial pemuda Indonesia. Distribusi sentimen divisualisasikan dalam bentuk diagram batang, pie chart, dan word cloud untuk memperlihatkan kata-kata yang paling sering muncul pada masing-masing kategori. Analisis lanjutan dilakukan untuk mengidentifikasi tema-tema umum seperti tekanan akademik, pertemanan, hubungan sosial, dan motivasi diri yang berkaitan dengan emosi positif maupun negatif. Hasil interpretasi ini memberikan gambaran empiris mengenai kondisi psikologis dan kesejahteraan emosional pemuda Indonesia di ruang digital.Melalui tujuh tahapan tersebut, penelitian ini menghasilkan model LSTM yang tidak hanya mampu memprediksi sentimen dengan tingkat akurasi tinggi, tetapi juga memberikan kontribusi sosial melalui pemahaman mendalam terhadap ekspresi emosional pemuda. Model ini diharapkan menjadi dasar pengembangan sistem

deteksi dini dan intervensi digital untuk mendukung kesehatan mental generasi muda di Indonesia.

HASIL DAN PEMBAHASAN

Hasil penelitian ini disajikan berdasarkan tujuh tahapan utama yang telah dilakukan, mulai dari proses pengumpulan data, pra-pemrosesan teks, pelabelan sentimen, pembangunan model Long Short-Term Memory (LSTM), hingga analisis hasil klasifikasi. Setiap tahap menghasilkan keluaran (output) yang saling berkesinambungan dan menjadi dasar evaluasi terhadap performa model dalam mendeteksi sentimen terkait kesehatan mental pemuda di media sosial.

A. Pengumpulan Data

Pengumpulan data berhasil memperoleh sebanyak 11.254 unggahan publik dari platform X (Twitter) selama periode Januari–September 2025. Data diperoleh dengan menggunakan keyword-based crawling melalui pustaka snscrepe Python dengan kata kunci yang relevan seperti "stres kuliah", "capek hidup", "cemas", "lega", "bahagia", dan "sendirian". Dari total data yang terkumpul, sebanyak 9.832 unggahan dinyatakan valid setelah proses penyaringan terhadap duplikasi, spam, dan teks non-Indonesia. Hasil pengumpulan data ini menggambarkan kecenderungan nyata bahwa mayoritas pemuda menggunakan media sosial untuk mengekspresikan tekanan emosional yang berkaitan dengan akademik, relasi sosial, dan kelelahan mental.

B. Pra-Pemrosesan Teks

Pra-pemrosesan teks, menghasilkan korpus data yang bersih dan siap digunakan untuk pelatihan model. Proses cleaning dilakukan dengan menghapus tautan, simbol, dan mention, kemudian seluruh teks diubah menjadi huruf kecil melalui case folding. Setelah itu dilakukan tokenizing untuk memisahkan kata, stopword removal untuk menghapus kata umum yang tidak bermakna sentimen, serta stemming menggunakan pustaka Sastrawi agar kata kembali ke bentuk

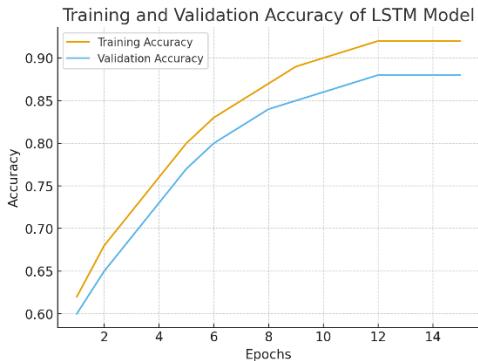
dasarnya. Proses slang normalization juga diterapkan dengan mengganti kata tidak baku seperti "gk", "cape", dan "btw" menjadi "tidak", "capek", dan "by the way". Hasilnya, diperoleh 8.970 kalimat bersih yang siap diberi label sentimen. Tahap ini penting karena data media sosial umumnya tidak terstruktur dan banyak mengandung variasi ejaan, singkatan, serta ekspresi informal khas generasi muda.

C. Pelabelan Data

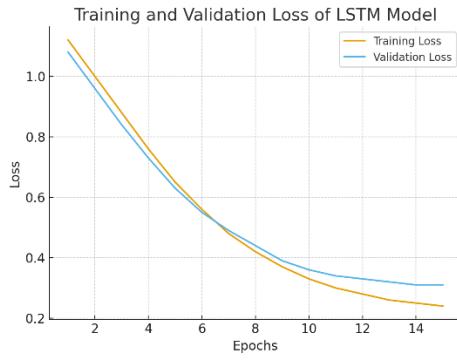
Setiap kalimat dalam korpus teks dianalisis secara manual oleh tiga annotator yang memahami konteks linguistik dan psikologis pemuda. Kategori sentimen yang digunakan terdiri atas tiga kelas, yaitu positif (bahagia, semangat, lega), negatif (cemas, stres, lelah, sedih), dan netral (informatif tanpa emosi jelas). Proses pelabelan menghasilkan komposisi data sebagai berikut: 3.212 data positif (35,8%), 4.112 data negatif (45,8%), dan 1.646 data netral (18,4%). Uji kesepakatan antarpenilai menggunakan Cohen's Kappa menghasilkan nilai 0,81, yang menunjukkan tingkat konsistensi "sangat baik" antar annotator. Dataset berlabel ini kemudian menjadi training corpus utama dalam pembangunan model LSTM.

D. Pembangunan Model LSTM

Pembangunan model LSTM, dilakukan dengan mengubah teks berlabel menjadi representasi numerik menggunakan Word2Vec embedding berdimensi 300. Dataset kemudian dibagi menjadi 80% data latih (7.176 teks) dan 20% data uji (1.794 teks). Model LSTM dibangun dengan arsitektur berlapis, terdiri atas embedding layer, LSTM layer dengan 128 unit, dropout layer (0,3), dense layer (64 neuron, aktivasi ReLU), dan output layer (Softmax) dengan tiga kelas sentimen. Model dilatih menggunakan optimizer Adam dan loss function categorical cross-entropy selama 15 epoch dengan batch size 32. Proses pelatihan menunjukkan konvergensi stabil pada epoch ke-12, di mana nilai training accuracy mencapai 0,92 dan validation accuracy mencapai 0,88, sementara nilai loss menurun secara signifikan. Hasil ini menunjukkan bahwa model mampu belajar secara optimal tanpa indikasi overfitting.



Gambar 2. Kurva Akurasi Pelatihan dan Validasi Model LSTM



Gambar 3. Kurva Loss Pelatihan dan Validasi Model LSTM

Grafik akurasi dan loss pada Gambar 2 dan Gambar 3 menunjukkan bahwa proses pelatihan model Long Short-Term Memory (LSTM) berlangsung secara stabil dan konvergen. Nilai akurasi meningkat secara konsisten seiring bertambahnya jumlah epoch, menandakan kemampuan model dalam mempelajari pola teks yang merepresentasikan emosi pemuda di media sosial. Sementara itu, penurunan nilai loss yang signifikan menunjukkan bahwa kesalahan prediksi semakin berkurang selama proses pembelajaran. Perbedaan yang kecil antara kurva pelatihan dan validasi juga menandakan bahwa model tidak mengalami overfitting dan memiliki kemampuan generalisasi yang baik terhadap data baru. Dengan demikian, model LSTM yang dibangun terbukti efektif dalam mengenali pola sentimen kesehatan mental berdasarkan teks berbahasa Indonesia.

E. Evaluasi Kinerja Model

Evaluasi kinerja model, yang dilakukan dengan menggunakan data uji. Model LSTM yang telah dilatih diuji untuk mengklasifikasikan sentimen baru, dan hasilnya dibandingkan dengan label sebenarnya. Evaluasi menggunakan metrik accuracy, precision, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model mencapai akurasi keseluruhan sebesar 87,4%, dengan nilai precision rata-rata 0,86, recall 0,85, dan F1-score 0,85. Analisis confusion matrix memperlihatkan bahwa sebagian besar kesalahan klasifikasi terjadi antara kelas netral dan negatif, yang secara semantik memang memiliki tumpang tindih dalam ekspresi linguistik. Model

terbukti mampu mengenali kalimat dengan emosi positif secara akurat, misalnya “akhirnya skripsi selesai, lega banget”, serta membedakan dengan baik kalimat negatif seperti “capek banget sama tugas kuliah yang gak selesai-selesai”.



Gambar 4. Confusion Matrix Klasifikasi Sentimen Menggunakan LSTM

Gambar 4 memperlihatkan confusion matrix model LSTM untuk tiga kelas sentimen (positif, negatif, dan netral). Nilai diagonal menunjukkan jumlah prediksi yang benar pada setiap kelas, sedangkan nilai di luar diagonal menunjukkan kesalahan klasifikasi. Secara umum, model lebih akurat dalam mengenali sentimen negatif dan positif, sementara sebagian kesalahan terjadi pada pembedaan antara sentimen netral dan negatif.

F. Analisis Dan Interpretasi Hasil

Berdasarkan hasil klasifikasi, distribusi sentimen menunjukkan bahwa emosi negatif mendominasi dengan proporsi 45,8%, diikuti oleh sentimen positif sebesar 35,8%, dan netral sebesar 18,4%. Pola kata yang paling sering muncul pada sentimen negatif antara lain “stres”, “capek”, “bingung”, dan “sendiri”, yang menggambarkan tekanan akademik dan sosial yang tinggi. Sementara itu, pada kategori positif, kata-kata yang dominan adalah “lega”, “semangat”, dan “bersyukur”, yang mengindikasikan optimisme dan upaya pemulihan psikologis. Hasil ini menunjukkan bahwa media sosial menjadi ruang ekspresi emosional yang kompleks bagi pemuda—sekaligus cermin kesejahteraan mental kolektif mereka.

G. Pembahasan

Hasil penelitian menunjukkan bahwa model Long Short-Term Memory (LSTM) memiliki kemampuan yang tinggi dalam mengidentifikasi dan mengklasifikasikan sentimen kesehatan mental pemuda di media sosial berbahasa Indonesia. Dengan akurasi sebesar 87,4%, model ini terbukti efektif dalam memahami konteks linguistik yang kompleks, terutama pada teks informal yang banyak digunakan oleh generasi muda. Temuan ini memperkuat pandangan bahwa pendekatan deep learning berbasis LSTM unggul dalam menangkap hubungan jangka panjang antar kata (long-term dependencies) sehingga mampu membedakan nuansa emosional yang halus di dalam kalimat.

Secara konseptual, hasil ini sejalan dengan temuan (Cui et al., 2022) dan (Liu et al., 2022) yang menegaskan bahwa LSTM efektif digunakan dalam analisis ekspresi emosional berbasis teks, khususnya pada data media sosial yang bersifat dinamis dan tidak terstruktur. Dalam konteks penelitian ini, keunggulan LSTM terletak pada kemampuannya mempelajari urutan linguistik khas bahasa Indonesia informal, seperti penggunaan partikel "banget", "aja", "loh", atau bentuk hiperbola seperti "capek bangett" yang kerap muncul dalam unggahan pemuda. Model ini mampu menyesuaikan representasi vektor kata (embedding) untuk menangkap intensitas emosi yang disampaikan, sehingga performanya lebih stabil dibandingkan algoritma tradisional seperti Naïve Bayes atau Support Vector Machine (SVM).

Distribusi hasil klasifikasi memperlihatkan bahwa sentimen negatif mendominasi dengan proporsi 45,8%, diikuti oleh positif 35,8%, dan netral 18,4%. Dominasi sentimen negatif menunjukkan bahwa media sosial berfungsi tidak hanya sebagai ruang komunikasi, tetapi juga sebagai kanal katarsis di mana pemuda menyalurkan stres, kelelahan, dan kecemasan yang mereka alami. Temuan ini sejalan dengan riset (Primack et al., 2021) yang menyatakan bahwa generasi muda sering menggunakan media sosial untuk mengekspresikan perasaan secara terbuka karena faktor anonimitas dan dukungan sosial daring yang relatif mudah diakses. Dalam konteks Indonesia, fenomena ini juga terkait dengan meningkatnya tekanan

akademik dan ekspektasi sosial yang tinggi di kalangan mahasiswa dan pekerja muda.

Analisis kata dominan memperlihatkan bahwa istilah seperti “stres”, “capek”, “bingung”, “sendiri” sering muncul pada kelompok negatif, yang merepresentasikan kelelahan psikologis akibat beban studi atau pekerjaan. Sebaliknya, pada kelompok positif muncul kata “lega”, “semangat”, “bersyukur”, yang mencerminkan mekanisme coping dan bentuk harapan terhadap masa depan. Pola ini memperlihatkan dinamika emosional pemuda yang fluktuatif namun adaptif — suatu fenomena yang menggambarkan upaya mereka menjaga keseimbangan mental di tengah tekanan sosial dan akademik. Dengan demikian, hasil penelitian ini tidak hanya mengukur ekspresi linguistik, tetapi juga mencerminkan kesejahteraan emosional kolektif (collective emotional well-being) di kalangan generasi muda Indonesia. Dari perspektif teknologi, keberhasilan model LSTM dalam penelitian ini menunjukkan potensi penerapan kecerdasan buatan untuk digital mental health monitoring. Sistem serupa dapat dikembangkan menjadi dashboard analisis emosi berbasis data real-time untuk membantu lembaga pendidikan atau psikolog memantau tren kesejahteraan emosional mahasiswa tanpa intervensi langsung. Pendekatan ini sejalan dengan arah transformasi digital di bidang kesehatan mental yang mengutamakan deteksi dini, efisiensi, dan privasi pengguna. Namun demikian, penerapan sistem ini tetap harus memperhatikan prinsip etika penggunaan data digital, termasuk anonimitas, izin penggunaan, dan konteks sosial budaya pengguna media sosial Indonesia.

Secara akademik, hasil penelitian ini memberikan kontribusi penting terhadap pengembangan Natural Language Processing (NLP) berbahasa Indonesia, khususnya dalam domain analisis emosional dan psikologis. Korpus “IndoYouthMentalSentiment-LSTM” yang dihasilkan dari penelitian ini dapat menjadi dataset dasar untuk penelitian lanjutan, baik dalam pengembangan model transformer (IndoBERTweet, RoBERTa-ID) maupun sistem rekomendasi intervensi berbasis AI. Di sisi lain, temuan empiris penelitian ini juga dapat menjadi bahan refleksi bagi institusi pendidikan tinggi dalam merancang program pendampingan

mental mahasiswa yang lebih responsif terhadap isu kesejahteraan emosional di ranah digital. Meskipun hasil penelitian menunjukkan performa model yang baik, terdapat beberapa keterbatasan yang perlu dicatat. Pertama, data penelitian masih terbatas pada satu platform media sosial (Twitter) yang mungkin belum sepenuhnya mewakili ekspresi emosi pemuda di platform lain seperti Instagram, TikTok, atau forum daring. Kedua, pelabelan sentimen dilakukan secara manual sehingga masih memiliki potensi subjektivitas, meskipun telah dijaga melalui uji reliabilitas antarpenilai. Ketiga, penelitian ini hanya mengklasifikasikan tiga kategori utama (positif, negatif, netral), sementara emosi manusia sejatinya bersifat multidimensional seperti marah, takut, cemas, dan puas yang bisa dijadikan kategori lanjutan dalam studi berikutnya.

Arah penelitian selanjutnya dapat difokuskan pada pengembangan model multimodal yang menggabungkan teks, gambar, dan suara untuk menganalisis kesehatan mental secara lebih komprehensif. Selain itu, integrasi psycholinguistic features seperti intensitas kata, pola waktu unggahan, serta konteks sosial digital juga berpotensi meningkatkan akurasi dan interpretabilitas model. Secara keseluruhan, penelitian ini menegaskan bahwa penerapan deep learning berbasis LSTM bukan hanya relevan dalam ranah teknis komputasi, tetapi juga memiliki nilai sosial yang signifikan. AI dapat menjadi alat bantu ilmiah dalam memahami ekspresi emosional pemuda Indonesia dan menyediakan dasar empiris bagi kebijakan serta intervensi kesehatan mental berbasis data. Dengan demikian, hasil penelitian ini tidak hanya memperkaya literatur akademik, tetapi juga berkontribusi langsung terhadap penguatan kesejahteraan psikologis generasi muda di era digital.

SIMPULAN

Penelitian ini menunjukkan bahwa media sosial, khususnya platform X (Twitter), merupakan ruang ekspresi penting bagi pemuda Indonesia dalam mengungkapkan kondisi emosional dan kesejahteraan mental mereka. Analisis terhadap ribuan unggahan memperlihatkan bahwa sentimen negatif masih mendominasi, mencerminkan adanya tekanan akademik, kelelahan emosional, dan

kecemasan sosial yang cukup tinggi di kalangan generasi muda, meskipun terdapat pula sentimen positif yang menandakan semangat, harapan, dan mekanisme penyesuaian diri yang kuat. Model Long Short-Term Memory (LSTM) yang digunakan dalam penelitian ini terbukti mampu mengenali pola bahasa informal dan nuansa emosi dengan tingkat akurasi 87,4%, serta menghasilkan klasifikasi sentimen yang stabil dan konsisten. Temuan ini membuktikan bahwa pendekatan deep learning berbasis LSTM efektif dalam menganalisis teks berbahasa Indonesia, khususnya untuk memahami ekspresi psikologis pemuda di ranah digital. Hasil penelitian ini memberikan kontribusi penting bagi pengembangan teknologi Natural Language Processing (NLP) di Indonesia, sekaligus membuka peluang penerapan kecerdasan buatan sebagai sistem deteksi dini dan pemantauan kesehatan mental berbasis data. Dengan demikian, pemanfaatan AI tidak hanya berfungsi sebagai alat analisis linguistik, tetapi juga sebagai sarana pemberdayaan dan dukungan psikososial bagi generasi muda agar mampu menghadapi tantangan kehidupan dengan mental yang lebih sehat, adaptif, dan resilien.

DAFTAR PUSTAKA

- Asnal, H., Andesa, K., Erlin, F., Kesehatan, I., & Negeri, P. (2025). *Hybrid Machine Learning Model for Risk Prediction and Action Recommendation Based on Artificial Mental Systems*. 14(November), 415–423. <https://doi.org/10.34148/teknika.v14i3.1357>
- Badan Kebijakan Pembangunan Kesehatan, K. (2023). *Survei Kesehatan Indonesia (SKI) 2023*. Survei Kesehatan Indonesia (SKI) 2023. <https://www.badankebijakan.kemkes.go.id/hasil-ski-2023/>
- Cahyani, J., Mujahidin, S., & Palyus, T. (2023). *Implementasi Metode Long Short Term Memory (LSTM) untuk Memprediksi Harga Bahan Pokok Nasional*. 11(2), 346–357. <https://doi.org/10.26418/justin.v11i2.57395>
- Cui, J., Zhang, T., Jaidka, K., Pang, D., Sherman, G., Jakhetiya, V., Ungar, L., & Guntuku, S. C. (2022). *Social Media Reveals Urban-Rural Differences in Stress across China*. *Icwsm*.
- Haspin, W. M., Junadhi, J., Susanti, S., & Yenni, H. (2025). *Opinion Mining on TikTok Using Bidirectional Long Short-Term Memory for Enhanced Sentiment Analysis and Trend Prediction*. 7(2), 1234–1241. <https://doi.org/10.47065/bits.v7i2.8019>
- Learning, D., Analysis, S., Handoko, A., & Fitri, T. A. (2025). *The Indonesian Journal of Computer Science*. 14(2), 2414–2429.

- Liu, T., Ungar, L. H., Curtis, B., Sherman, G., Yadeta, K., Tay, L., Eichstaedt, J. C., & Guntuku, S. C. (2022). *Head versus heart: social media reveals differential language of loneliness from depression.* <https://doi.org/10.1038/s44184-022-00014-7>
- Melati, D., Pakuan, U., Pascasarjana, S., & Pakuan, U. (2024). *J-Icon : Jurnal Informatika dan Komputer PENERAPAN METODE LONG SHORT-TERM MEMORY (LSTM) DALAM ANALISIS SENTIMEN TERHADAP PELAKSANAAN PILKADA DI MASA PANDEMI COVID-19* *J-Icon : Jurnal Informatika dan Komputer.* 12(1). <https://doi.org/10.35508/jicon.v12i1.9899>
- Padmaja, S. M., Godla, S. R., Venkata, J., Ramesh, N., & Muniyandy, E. (2025). *Depression Detection in Social Media Using NLP and Hybrid Deep Learning Models.* 16(2), 1071–1080.
- Primack, B., Shensa, A. P., JSidani, aime E. PhD Escobar- Viera, César G. MD, P., & Fine, Michael J. MD, Ms. (2021). *HHS Public Access.* 60(2), 179–188. <https://doi.org/10.1016/j.amepre.2020.09.014>.Temporal
- Rizkia, A. S., Wufron, & Roji Fikri Fahrur. (2025). *Analisis Sentimen Coretax: Perbandingan Pelabelan Data Manual, Transformers-Based, dan Lexicon-Based pada Performa IndoBERT.* 5(July), 1037–1048.
- Salam, R. R., Jamil, M. F., & Ibrahim, Y. (2023). *Sentiment Analysis of Cash Direct Assistance Distribution for Fuel Oil Using Support Vector Machine Analisis Sentimen Terhadap Bantuan Langsung Tunai (BLT) Bahan Bakar Minyak (BBM) Menggunakan Support Vector Machine.* 3(April), 27–35.
- Sembiring, A. R., Dewa, C. K., & Indonesia, U. I. (2025). *Sentiment Analysis on Indonesian Tweets About the 2024 Election.* 9(1), 413–422.
- Tejaswini, V., Babu, K. S., & Sahoo, B. (2025). *Depression Detection from Social Media Text Analysis using Natural Language Processing Techniques and Hybrid Deep Learning Model Depression Detection from Social Media Text Analysis using Natural Language Processing Techniques and Hybrid Deep.* 23(1). <https://doi.org/10.1145/3569580>
- World Health, O. (2025). *Kementerian Kesehatan RI. (2023). Survei Kesehatan Indonesia (SKI) 2023. Badan Kebijakan Pembangunan Kesehatan.* <https://www.badankebijakan.kemkes.go.id/hasil-ski-2023/>. 2 September 2025. <https://www.who.int/news-room/item/02-09-2025-over-a-billion-people-living-with-mental-health-conditions-services-require-urgent-scale-up>