Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

PERBANDINGAN KINERJA SUPPORT VECTOR MACHINE DAN LONG SHORT-TERM MEMORY DALAM DETEKSI CYBERBULLYING

Performance Comparison of Support Vector Machine and Long Short-Term Memory in Cyberbullying Detection

Muhammad Aditya Yudhistira[1], Rizki Yusliana Bakti[2], Muhammad Faisal[3]

[1,2,3]Universitas Muhammadiyah Makassar

[1]Email: 105841114122@student.unismuh.ac.id
[2]Email: rizkiyusliana@unismuh.ac.id
[3]Email: muhfaisal@unismuh.ac.id

Abstract
The increasingly massive phenomenon of bullying spreading on social media drives the urgency of applying computational approaches to detect cyberbullying quickly and accurately. This study aims to compare the effectiveness of the Support Vector Machine (SVM) algorithm based on TF-IDF against the Long Short-Term Memory (LSTM) architecture based on Word Embedding in classifying bullying texts. The experiment was conducted utilizing a dataset of 6,520 Indonesian comments extracted from the social media platform X (Twitter). The entire data was then divided into 80% training data and 20% testing data. Performance measurements proved that the SVM model dominated the overall prediction accuracy, achieving an Accuracy of 0.8352 (83%), Precision of 0.8356 (83%), and F1-Score of 0.8363 (83%). Conversely, the LSTM model demonstrated superiority in the sensitivity of positive class identification, scoring a Recall of 0.8552 (85%) and ROC-AUC of 0.9088 (90%). In terms of computational efficiency, the SVM model proved to be far more resource-efficient, requiring only 42.90 seconds of training time and 74.87 MB of memory, in contrast to the LSTM which consumed up to 106.67 seconds and 126.80 MB of memory. This research concludes that SVM is the most optimal method overall because it is able to balance high detection accuracy with excellent operational load efficiency.
Keywords: *Cyberbullying, Text Detection, Machine Learning, Support Vector Machine, Long Short-Term Memory*

Abstrak
Fenomena perundungan di media sosial yang semakin masif mendorong urgensi penerapan pendekatan komputasional untuk mendeteksi cyberbullying secara cepat dan akurat. Studi ini bertujuan mengkomparasikan efektivitas algoritma Support Vector Machine (SVM) berbasis TF-IDF terhadap arsitektur Long Short-Term Memory (LSTM) berbasis Word Embedding dalam mengklasifikasikan teks perundungan. Eksperimen dilakukan dengan memanfaatkan 6.520 dataset komentar berbahasa Indonesia yang didapat dari media sosial X (Twitter). Seluruh data tersebut kemudian dibagi menjadi 80% data latih dan 20% data uji. Pengukuran kinerja membuktikan bahwa model SVM mendominasi tingkat ketepatan prediksi secara umum dengan meraih Accuracy 0.8352 (83%), Precision 0.8356 (83%), dan F1-Score 0.8363 (83%). Sebaliknya, model LSTM memperlihatkan keunggulan pada sensitivitas identifikasi kelas positif dengan mencetak nilai Recall 0.8552 (85%) dan ROC-AUC 0.9088 (90%). Dari segi efisiensi komputasi, model SVM terbukti jauh lebih hemat sumber daya, hanya menghabiskan durasi pelatihan 42.90 detik dan penggunaan memori 74.87 MB, berbanding terbalik dengan

Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

LSTM yang menghabiskan waktu hingga 106.67 detik dan memori 126.80 MB. Penelitian ini menyimpulkan bahwa SVM merupakan metode paling optimal secara keseluruhan karena mampu menyeimbangkan akurasi deteksi yang tinggi dengan efisiensi beban operasional yang sangat baik.
Kata Kunci: *Cyberbullying, Deteksi Teks, Machine Learning, Support Vector Machine, Long Short-Term Memory*

PENDAHULUAN

Perkembangan teknologi informasi yang sangat pesat yang sejalan dengan tingginya penggunaan *smartphone*, telah mengubah pola interaksi sosial di tengah masyarakat (Ganiem et al., 2024; Ricoy et al., 2022). Platform media sosial seperti X (Twitter), Instagram, dan TikTok kini tidak hanya menjadi media hiburan, melainkan ruang publik digital utama untuk bertukar informasi dan berekspresi. Sayangnya, kemudahan konektivitas ini beriringan dengan munculnya risiko *online disinhibition effect* di mana pengguna merasa bebas mengabaikan norma sosial dan bertindak agresif di balik anonimitas (Amanda Putri et al., 2025; Salwa Aiza Zahra et al., 2026). Hal ini melahirkan dampak residu yang sangat merugikan, salah satunya yaitu lonjakan kasus perundungan siber atau biasa disebut sebagai *cyberbullying*. Bentuk *cyberbullying* biasanya terjadi berupa ujaran kebencian, *body shaming*, penghinaan rasial, hingga ancaman verbal di media sosial akan memicu atau memberikan dampak psikologis yang jauh lebih desktruktif bagi korban, dimulai dari penurunan *self-esteem*, depresi, kecemasan akut, hingga yang terparah yaitu risiko bunuh diri bagi para korbannya (Bansal et al., 2023; Wulandari & Sujarwo, 2024).

Di Indonesia, tingginya intensitas penggunaan media sosial menghasilkan jutaan data komentar setiap harinya (Tewu et al., 2025). Menghadapi volume data yang sangat masif sistem moderasi manual yang bergantung pada tenaga manusia terbukti tidak efisien (He et al., 2022). Ketidakefisien ini menyebabkan keterlambatan penanganan, yang pada akhirnya akan membiarkan konten negatif bertahan lebih lama dan memperparah trauma psikologis korban. Menjawab tantangan tersebut, intervensi sistem komputasi untuk deteksi cyberbullying secara otomatis menjadi sebuah urgensi. Dalam hal ini, integrasi antara *Natural Language Processing* dan *Machine Learning* hadir sebagai solusi paling mutakhir (*state of the art*) untuk menyaring dan menganalisis tumpukan data teks berskala besar secara cepat dan presisi (Kagi, 2025; Safitri et al., 2025).

Namun, dalam mengimplementasikan teknologi deteksi, tantangan terbesarnya adalah menemukan arsitektur model komputasi yang tidak hanya akurat dalam mengklasifikasikan teks, tetapi juga efisien saat diproses oleh sistem. Oleh karena itu, penelitian ini tidak hanya berfokus pada pembangunan model, melainkan juga melakukan komparasi antara algoritma *machine learning* dan pendekatan *deep learning* guna menemukan metode yang paling cocok untuk karakteristik data cyberbullying berbahasa Indonesia. Secara historis, algoritma *Machine Learning* seperti *Support Vector Machine* (SVM) merupakan salah satu pendekatan yang paling tangguh untuk klasifikasi teks. Metode berbasis supervised learning ini bekerja dengan menemukan hyperplane atau batas keputusan linier yang paling optimal untuk memisahkan antar kelas dengan memaksimalkan jarak margin (Du et al., 2024; Jabardi, 2025).

Di sisi lain, evolusi kecerdasan buatan menawarkan paradigma *Deep*

Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

Learning melalui arsitektur Long Short-Term Memory (LSTM). Sebagai pengembangan lanjut dari *Recurrent Neural Network* (RNN), LSTM dirancang secara spesifik untuk menanggulangi masalah hilangnya gradien (*vanishing gradient*) sekaligus merekam memori ketergantungan jangka panjang (Awalia Z Abidin et al., 2025; Mienye et al., 2024). Berbekal mekanisme *memory gates*, algoritma ini secara otomatis menyeleksi informasi masa lalu, sehingga handal dalam menangkap urutan temporal dan konteks semantik antar kata dalam sebuah kalimat (Al Farisi et al., 2024; Krichen & Mihoub, 2025).

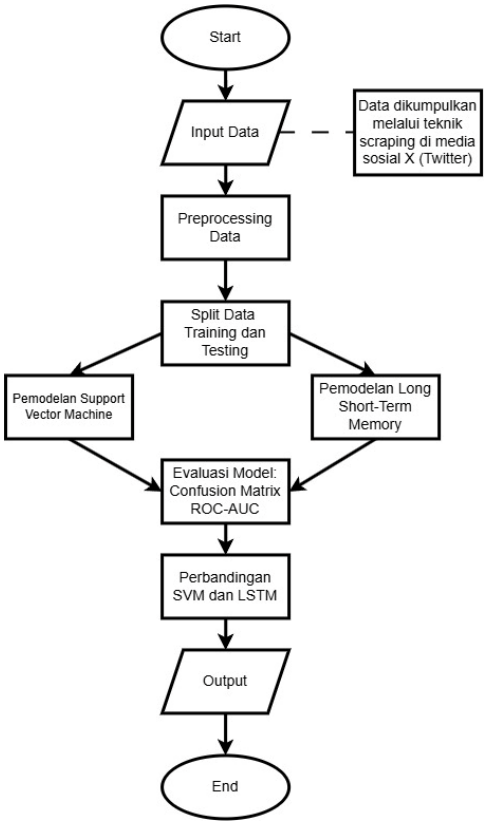
Perbedaan karakteristik antara SVM yang ringan dan LSTM yang kontekstual menciptakan urgensi untuk mengevaluasi kinerja keduanya secara objektif. Berdasarkan uraian latar belakang tersebut, rumusan masalah dalam penelitian ini adalah bagaimana implementasi model SVM dan LSTM dalam mendeteksi teks *cyberbullying*, serta bagaimana perbandingan kinerja kedua model tersebut. Sejalan dengan hal itu, tujuan penelitian ini adalah mengimplementasikan algoritma SVM dan LSTM, membandingkan kinerjanya berdasarkan metrik evaluasi *Confusion Matrix*, ROC-AUC, serta efisiensi model dan pada akhirnya menentukan metode manakah yang memberikan performa paling optimal dan akurat.

METODE

Penelitian ini dilaksanakan sepenuhnya secara daring (online) karena berfokus pada pemrosesan komputasi dan analisis terhadap data sekunder, sehingga tidak memerlukan tahapan observasi fisik di lapangan. Keseluruhan rangkaian proses penelitian, yang meliputi tahap pengumpulan dataset dari media sosial, perancangan dan pelatihan model komputasi, hingga evaluasi kinerja sistem, diselesaikan dalam kurun waktu empat bulan. Adapun pengumpulan data teks komentar yang dijadikan sebagai objek penelitian dibatasi pada periode rentang waktu di tahun 2026.

Perancangan sistem merupakan tahapan yang memodelkan alur kerja sistem deteksi cyberbullying dari awal hingga akhir. Sistem ini dirancang untuk memproses data teks mentah, mengekstraksi fiturnya, melatih dua model secara paralel, hingga mengevaluasi dan membandingkan kinerjanya. Secara visual, alur prosedur kerja dalam penelitian ini direpresentasikan melalui flowchart pada Gambar 1.

Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428



Gambar 1 Alur Kerja Sistem

Berdasarkan flowchart pada Gambar 1, alur sistem diawali dengan menerima masukan dataset cyberbullying yang sebelumnya telah dikumpulkan melalui teknik scraping dari platform media sosial X (Twitter). Dataset mentah tersebut kemudian dibersihkan dan distandarisasi melalui tahapan preprocessing data yang memuat serangkaian proses *Natural Language Processing* (NLP), yang meliputi cleaning, *case folding*, dan *tokenizing*. Setelah bersih, dataset dipecah menjadi dua bagian, yaitu data training dan data testing, yang mana pemisahan ini secara khusus dilakukan sebelum tahap ekstraksi fitur untuk mencegah terjadinya kebocoran informasi (*data leakage*) ke dalam model saat dilatih.

Selanjutnya, sistem bercabang ke dalam dua skenario pemodelan yang berjalan secara paralel. Pada skenario algoritma *Support Vector Machine* (SVM), teks ditransformasikan ke dalam matriks angka menggunakan metode pembobotan TF-IDF untuk diumpankan ke dalam model agar dapat mempelajari batas keputusan (*hyperplane*) antarkelas dan memprediksi data uji. Sementara itu, pada skenario arsitektur *Long Short-Term Memory* (LSTM), teks direpresentasikan ke dalam vektor berdimensi padat melalui teknik *Word Embedding* guna mempertahankan konteks urutan antarkata, yang kemudian diproses untuk mempelajari pola kalimat hingga mencapai titik konvergensi dan memprediksi data uji. Hasil prediksi dari masing-masing model SVM dan LSTM kemudian dihadapkan dengan label asli (*ground truth*) dari data uji untuk dihitung jumlah *True Positive*, *True Negative*, *False Positive*, dan *False Negative* ke dalam bentuk *Confusion Matrix*. Berdasarkan matriks evaluasi tersebut, sistem akan mengkalkulasikan nilai turunan berupa *Accuracy*, *Precision*, *Recall*, dan *F1-Score*, yang kinerjanya kemudian dianalisis dan dibandingkan secara kuantitatif.

Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

Pada tahap akhir, sistem menghasilkan output berupa laporan metrik kinerja serta kesimpulan mengenai model mana yang memiliki kemampuan paling optimal, efektif, dan akurat dalam mendeteksi teks cyberbullying.

Teknik analisis data dalam penelitian ini menggunakan *Confusion Matrix* untuk mengevaluasi secara objektif hasil prediksi dari kedua model komputasi yang telah dibangun. *Confusion matrix* membandingkan hasil klasifikasi sistem dengan label aktual untuk memetakan jumlah data ke dalam parameter *True Positive*, *True Negative*, *False Positive*, dan *False Negative*.

Berdasarkan keempat matriks parameter tersebut, kinerja dari model SVM dan LSTM dievaluasi menggunakan kalkulasi metrik, yang meliputi *accuracy* untuk mengukur ketepatan prediksi seacara keseluruhan, *precision* untuk memvalidasi ketepatan prediksi positif model, *recall* untuk mengukur kemampuan model dalam mendeteksi seluruh data aktual berkelas positif, serta *F1-score* yang menyintesiskan nilai *precision* dan *recall* untuk memberikan evaluasi secara menyeluruh. Lebih lanjut, kemampuan model dalam membedakan secara tepat antara kelas cyberbullying dan non-cyberbullying dianalisis menggunakan pengukuran *Receiver Operating Characteristic* (ROC) dan *Area Under Curve* (AUC). Selain menilai tingkat akurasi klasifikasi, penelitian ini juga melakukan analisis komparatif kuantitatif terkait efisiensi komputasi dari kedua model dengan membandingkan parameter waktu pelatihan, waktu prediksi, serta penggunaan memori selama model dilatih.

HASIL DAN PEMBAHASAN

Rangkaian pengujian ini mengacu pada kerangka metodologi penelitian guna mengukur akurasi, performa klasifikasi, dan efisiensi komputasi dari tiap model. Temuan eksperimen tersebut selanjutnya dijabarkan dalam beberapa bagian-bagian berikut.

Pengumpulan dan Preprocessing Data

Tahap pengumpulan data yang dilakukan melalui teknik *scraping* pada platform media sosial X (Twitter) berhasil mendapatkan total 6.520 komentar. Sebelum data tersebut dilanjutkan ke dalam arsitektur model, seluruh teks akan melalui tahapan pelabelan di mana data diklasifikasikan ke dalam kelas *Non-Cyberbullying* dan *Cyberbullying*, serta proses *preprocessing* untuk membersihkan berbagai macam noise. Rangkaian proses *preprocessing data* ini mencakup *case folding* yang menyeragamankan huruf menjadi kecil, penghapusan tautan URL, symbol dan mention, normalisasi kata tidak baku (slang), hingga *tokenizing* untuk memastikan struktur teks menjadi lebih bersih dan terstandarisasi. Sebagai gambaran representatif mengenai kondisi teks pada tahap awal sebelum dibersihkan disajikan pada Tabel 1.

Tabel 1. Data Mentah

Tweet_url	Full_text
https://x.com/ithGYJ/status/2031734100974075977	#NewProfilePic jelek banget ini makhluk apa yaallah https://t.co/w0EbrYird3
https://x.com/juhoonlaper/status/2031376728325763078	@evanndeer si gblkkkkkk najis bgt

Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

Tweet_url	Full_text
https://x.com/Lukasoswalds/status/2033412 1024011841 54	@ Noel1050 https://t.co/y4hinxB5G V Bhapp bergaul bg biar punya kawan udah bego tolol pula
https://x.com/diaaargh/status/203173773732 2393761	@lockedcarnival Asli anjirlah bodoh banget

Setelah mengidentifikasi berbagai noise pada data mentah di Tabel 1, sistem mengeksekusi tahapan preprocessing secara berurutan untuk menstandarisasi teks. Kata-kata yang disingkat, huruf kapital, serta tautan dan mention diubah menjadi bentuk baku yang lebih mudah dipahami oleh algoritma klasifikasi. Hasil transformasi teks setelah melalui keseluruhan proses preprocessing tersebut disajikan pada Tabel 2.

Tabel 2 Hasil Preprocessing Data

Teks Asli	Hasil Preprocessing
#NewProfilePic jelek banget ini makhluk apa yaallah https://t.co/w0EbrYird3	jelek banget ini makhluk apa yaallah
@evanndeer si gblkkkkkk najis bgt	goblok najis banget
@ Noel1050 https://t.co/y4hinxB5G V Bhapp bergaul bg biar punya kawan udah bego tolol pula	bergaul biar punya kawan udah bego tolol pula
@lockedcarnival Asli anjirlah bodoh banget	asli anjir bodoh banget

Berdasarkan tahapan standarisasi tersebut, keseluruhan dataset kemudian dipartisi secara proporsional ke dalam dua kelompok utama untuk keperluan eksperimen, yaitu sebanyak 80% *data training* (5.216 data) dan 20% *data testing* (1.304 data) yang siap diproses ke tahap pemodelan.

Implementasi SVM dan LSTM

Pada penelitian ini, kedua model dieksekusi dengan pengaturan parameter yang disesuaikan untuk menangani karakteristik data teks komentar. Rincian komponen, parameter, dan konfigurasi arsitektur yang digunakan pada pemodelan SVM maupun LSTM dirangkum pada Tabel 3.

Tabel 3 Konfiurasi Model

Model	Parameter	Konfigurasi atau Nilai
SVM	*Feature extraction*	*Term Frequency-Inverse Document Frequency* (TF-IDF)
	Kernel	*Linear*
	Penanganan kelas	*class_weight="balanced"*
LSTM	Representasi teks	Tokenizer
	Padding sequence	Maksimal 40 token
	Word embedding	32
	Regularisasi	*SpatialDropout1D*, *Dropout*, *Recurrent Dropout*
	Layer LSTM	32 unit LSTM
	Hidden layer	Dense layer(16 neuron, aktivasi Relu)
	Output layer	1 neuron, aktivasi sigmoid
	Optimizer dan *loss function*	Adam (learning rate 0.0003) dan *Binary Crossentropy*
	Parameter pelatihan	50 *epoch*, *Batch size* 32

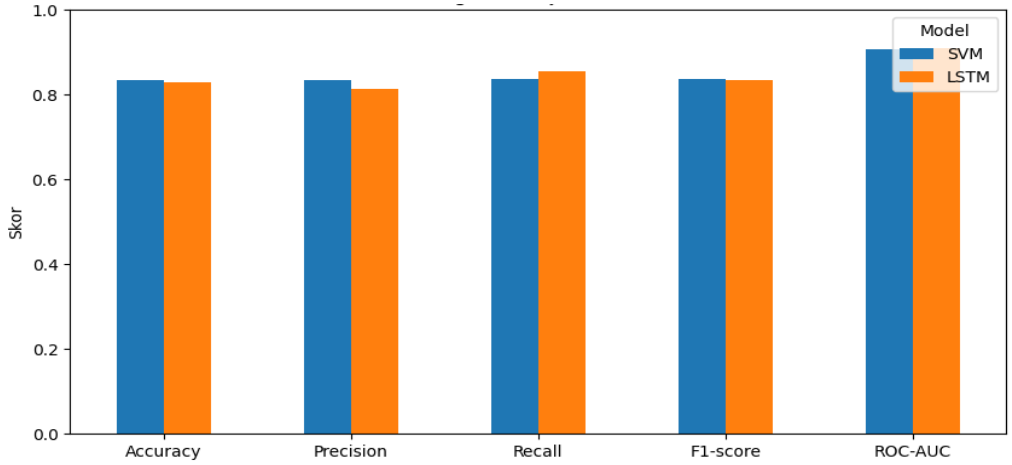
Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

Model	Parameter	Konfigurasi atau Nilai
	Callback	*EarlyStopping, ReduceRonPlateu*

Pada skenario pendekatan *machine learning*, model dibangun menggunakan algoritma SVM dengan konfigurasi kernel linear yang dinilai optimal untuk menangani data teks berdimensi tinggi hasil pembobotan TF-IDF. Penggunaan parameter penanganan kelas yang seimbang diterapkan agar algoritma tetap memperhatikan proporsi data *cyberbullying* dan *non-cyberbullying* selama pelatihan. Sementara itu, pada skenario *deep learning*, teks diproses menjadi urutan numerik dan diseragamkan panjangnya sebelum diproyeksikan ke dalam ruang dimensi vektor melalui layer *Word Embedding*. Fitur sekuensial tersebut kemudian dipelajari oleh layer LSTM dan dilewatkan ke *dense layer* untuk menghasilkan prediksi probabilistik klasifikasi biner pada *output layer*. Model LSTM dilatih secara iteratif dengan menerapkan *callback EarlyStopping* dan *ReduceLROnPlateau* yang berfungsi memantau pergerakan metrik validasi AUC guna mencegah *overfitting* sekaligus mengoptimalkan *learning rate*.

Hasil Evaluasi Model

Kinerja dari kedua model komputasi dievaluasi menggunakan *data testing* sebanyak 1.304 komentar. Evaluasi ini tidak hanya mengukur tingkat ketepatan prediksi secara keseluruhan (*accuracy*), tetapi juga mengukur kemampuan spesifik model melalui metrik *precision*, *recall*, *F1-score*, dan nilai ROC-AUC. Visualisasi perbandingan skor metrik evaluasi antara model SVM dan LSTM disajikan pada Gambar 2, dengan rincian persentase numeriknya dirangkum pada Tabel 4.



Gambar 2 Perbandingan Hasil Evaluasi Model

Tabel 4 Rincian Hasil Metrik Evaluasi Model

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
SVM	0.8352 (83%)	0.8356 (83%)	0.8369 (83%)	0.8363 (83%)	0.9076 (90%)
LSTM	0.8284 (82%)	0.8130 (81%)	0.8552 (85%)	0.8336 (83%)	0.9088 (90%)

Gambar 2 dan Tabel 4 menunjukkan bahwa model SVM dengan pendekatan *fitur extraction* TF-IDF memiliki tingkat ketepatan prediksi yang lebih unggul secara dengan capaian *accuracy* sebesar 0.8352, *precision* 0.8356, dan *F1-score* 0.8363. Tingginya nilai presisi pada SVM mengindikasikan bahwa model ini

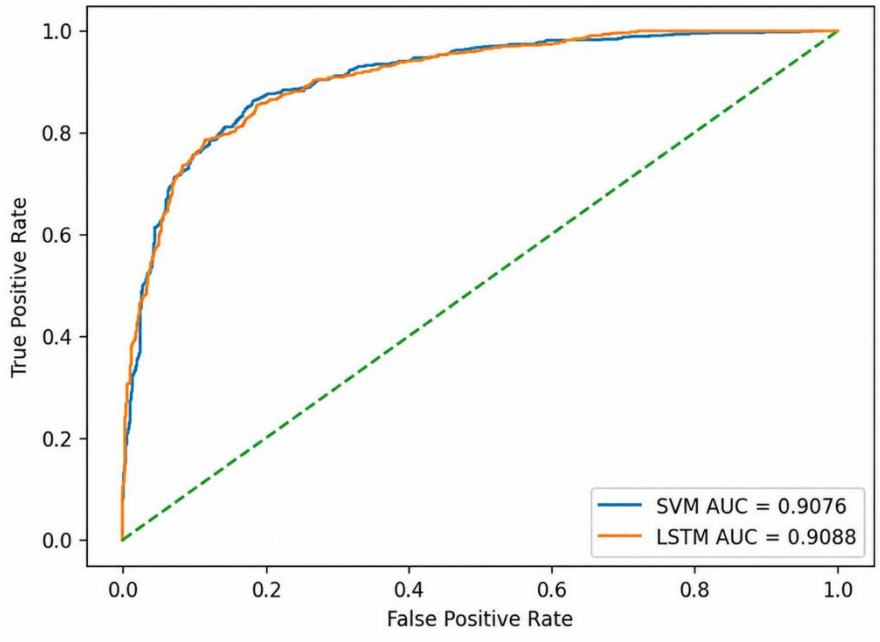
Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

menghasilkan lebih sedikit kesalahan dalam mengklasifikasikan komentar *non-cyberbullying* sebagai *cyberbullying*. Di sisi lain, model LSTM menunjukkan keunggulannya pada tingkat sensitivitas pengenalan kelas target, yang dibuktikan dengan perolehan *recall* sebesar 0.8552.

Kemampuan model dalam membedakan secara tepat antara kelas cyberbullying dan non-cyberbullying pada berbagai tingkat ambang batas divisualisasikan melalui kurva *Receiver Operating Characteristic* (ROC) pada Gambar 3.

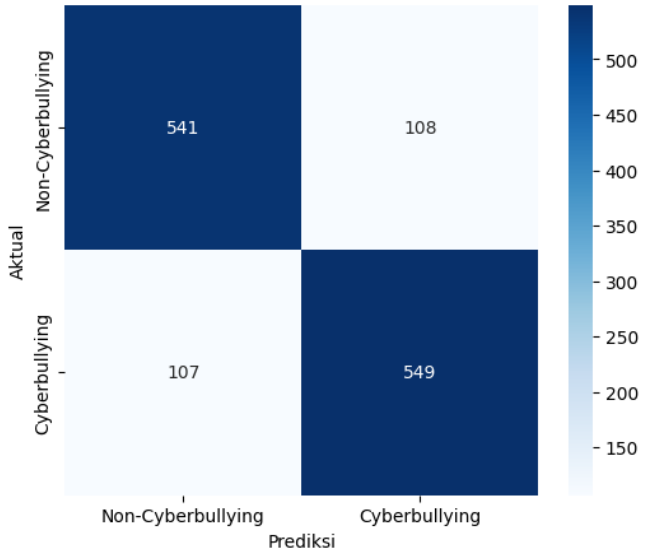
Kurva ROC pada Gambar 3 memperlihatkan probabilitas pemisahan kelas oleh kedua model. Model LSTM sedikit lebih unggul dalam membedakan kelas target dengan perolehan nilai luasan *Area Under Curve* (AUC) sebesar 0.9088, dibandingkan model SVM yang memperoleh nilai 0.9076. Meskipun perbedaannya sangat tipis, capaian nilai AUC di atas 0.90 membuktikan bahwa kedua arsitektur tersebut memiliki klasifikasi prediktif yang sangat baik dan andal pada data yang digunakan.

Untuk mendapatkan pemahaman yang lebih mendalam mengenai distribusi prediksi yang benar dan yang salah, evaluasi dilanjutkan dengan menganalisis *Confusion Matrix* dari masing-masing model di visualisasikan pada Gambar 4 dan 5.



Gambar 3 Kurva ROC-AUC

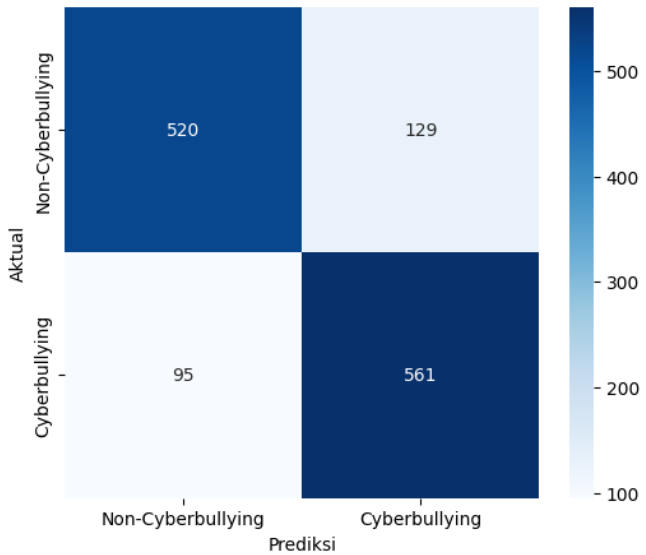
Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428



Gambar 4 *Confusion Matrix* SVM

Visualisasi Gambar 4 memperlihatkan bahwa model SVM menghasilkan total 1.090 prediksi benar, yang terdiri dari 541 data *non-cyberbullying* dan 549 data *cyberbullying* yang diklasifikasikan secara tepat. Sementara itu, SVM menghasilkan 215 prediksi salah, yang mencakup 108 *false positive* dan 107 *false negative*.

Sebagaimana ditampilkan pada Gambar 5, model LSTM menghasilkan total 1.081 prediksi benar, yang berasal dari 520 data *non-cyberbullying* dan 561 data *cyberbullying* yang berhasil diklasifikasikan dengan tepat. Di sisi lain, model ini menghasilkan 224 prediksi salah, yang terdistribusi ke dalam 129 *false positive* dan 95 *false negative*.



Gambar 5 *Confusion Matrix* LSTM

Tabel 5 Perbandingan Nilai *Confusion Matrix*

Model	True Positive	True Negative	False Positive	False Negative

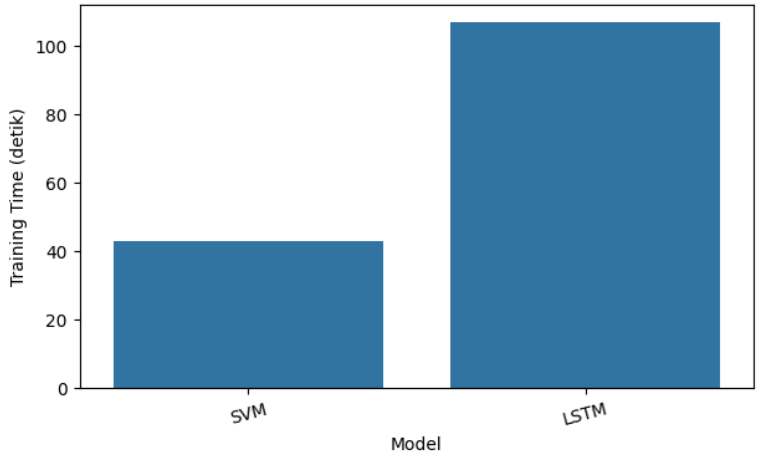
Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

SVM	549	541	108	107
LSTM	561	520	129	95

Perbandingan pada Tabel 5 membuktikan bahwa model SVM lebih tangguh dalam mengidentifikasi kelas *non-cyberbullying* secara benar, dengan perolehan *True Negative* (TN) yang lebih tinggi dibandingkan LSTM (541 berbanding 520). Sebaliknya, model LSTM berkinerja lebih baik dalam mendeteksi komentar yang memuat unsur perundungan, yang ditunjukkan oleh tingginya angka *True Positive* (TP) sebanyak 561 data. Keunggulan LSTM juga sangat jelas terlihat pada rendahnya angka *False Negative* (FN) (95 kesalahan) dibandingkan SVM (107 kesalahan). Kondisi ini menegaskan bahwa arsitektur sekuensial LSTM jauh lebih mumpuni dalam meminimalkan jumlah komentar cyberbullying yang gagal terdeteksi oleh sistem.

Hasil Pengukuran Efisiensi Model

Selain mengevaluasi ketepatan klasifikasi, penelitian ini juga mengukur aspek efisiensi komputasi dari kedua model. Efisiensi ini ditinjau berdasarkan tiga indikator utama, yaitu waktu pelatihan, waktu prediksi, dan penggunaan memori selama pelatihan.

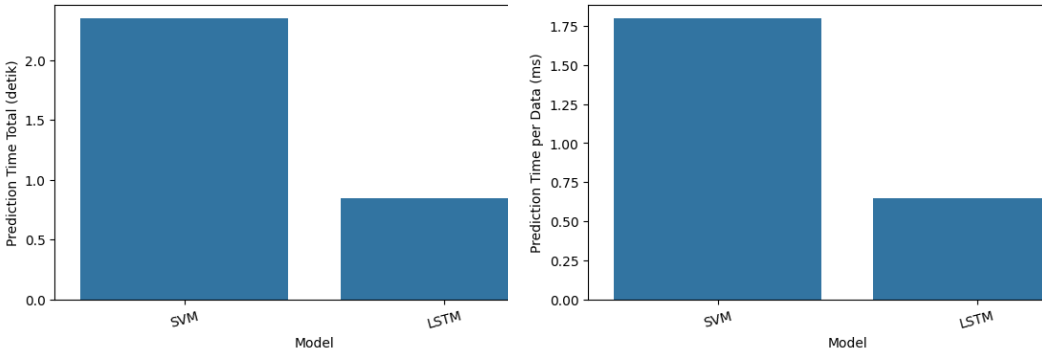


Gambar 6 Perbandingan Waktu Pelatihan

Berdasarkan Gambar 6, model SVM terbukti jauh lebih efisien dalam proses pelatihan dengan waktu hanya sebesar 42.90 detik, dibandingkan model LSTM yang membutuhkan waktu 106.67 detik. Kecepatan SVM didukung oleh penggunaan *fitur extraction* TF-IDF, di mana model hanya perlu mempelajari pola pemisahan kelas dari vektor numerik tersebut tanpa melalui proses pelatihan berulang. Sebaliknya, tingginya waktu pelatihan pada LSTM disebabkan oleh kompleksitas komputasi arsitektur *deep learning*, di mana model harus mempelajari representasi kata melalui embedding layer, memahami urutan kata, serta memperbarui bobot secara berulang pada setiap *epoch* melalui proses *backpropagation*.

Tahap evaluasi selanjutnya difokuskan pada kecepatan respons model saat melakukan klasifikasi terhadap data baru. Perbandingan total waktu prediksi dan waktu prediksi per data divisualisasikan pada Gambar 7.

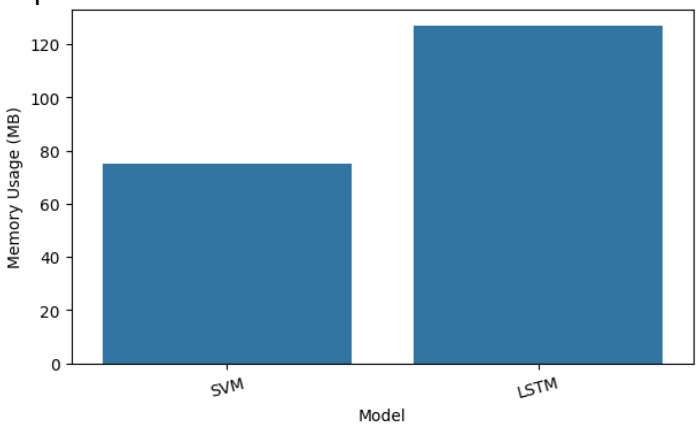
Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428



Gambar 7 Perbandingan Waktu Prediksi

Berbanding terbalik dengan hasil pada waktu pelatihan, Gambar 7 menunjukkan bahwa model LSTM jauh lebih cepat pada tahap prediksi. Model LSTM hanya membutuhkan waktu prediksi total sebesar 0.8476 detik (rata-rata 0.6493 ms per data), sedangkan SVM membutuhkan waktu 2.3466 detik (rata-rata 1.7981 ms per data). Waktu prediksi SVM menjadi lebih lama karena model harus memproses data teks yang telah direpresentasikan ke dalam vektor TF-IDF berdimensi sangat tinggi, ditambah dengan beban kalkulasi untuk menghasilkan probabilitas prediksi klasifikasi. Sementara itu, model LSTM memproses input yang telah distandarisasi ukurannya melalui tahapan *padding sequence*. Dengan input berupa urutan angka berukuran seragam, model LSTM dapat mengeksekusi perhitungan secara *batch* melalui operasi *forward pass* pada layer yang bobotnya sudah tetap, sehingga respons prediksinya menjadi jauh lebih cepat.

Sebagai bagian dari analisis yang menyeluruh terhadap sumber daya sistem, penelitian ini juga mengevaluasi alokasi penggunaan memori, sebagaimana divisualisasikan pada Gambar 8.



Gambar 8 Perbandingan Penggunaan Memori

Berdasarkan pengujian pada Gambar 8, model SVM lebih hemat dengan hanya mengalokasikan memori sebesar 74.87 MB, berbanding cukup jauh dengan model LSTM yang menyerap memori hingga 126.80 MB. SVM menggunakan memori yang lebih sedikit karena hasil pembobotan TF-IDF cenderung berbentuk *matrix sparse* (sebagian besar vektor bernilai nol). Di sisi lain, besarnya konsumsi memori pada LSTM dipicu oleh kompleksitas lapisan neural, di mana sistem harus menyimpan dan memperbarui banyak parameter, seperti representasi *word embedding*, status *cell state*, *gradien*, hingga parameter *optimizer* selama berjalannya *epoch*.

Nusantara Hasana Journal
Volume 6 No. 3 (Agustus 2026), Page: 42-55
E-ISSN: 2798-1428

Selanjutnya keseluruhan hasil pengukuran efisiensi komputasi dari kedua model dirangkum pada Tabel 6.

Tabel 6 Rincian Perbandingan Efisiensi Model

Model	Training Time	Prediction Time Total	Prediction Time Per Data	Memory Usage
SVM	42.90 detik	2.3466 detik	1.7981 ms	74.87 MB
LSTM	106.67 detik	0.8476 detik	0.6493 ms	126.80 MB

PEMBAHASAN

Berdasarkan serangkaian pengujian yang telah dilakukan, studi ini mengonfirmasi adanya *trade-off* antara pemanfaatan pendekatan *machine learning* dengan arsitektur *deep learning* dalam menangani klasifikasi teks *cyberbullying*.Apabila ditinjau dari performa klasifikasi, model SVM memperlihatkan hasil metrik yang lebih unggul dan stabil secara akumulatif. Hal ini dibuktikan melalui perolehan nilai *accuracy*, *precision*, dan *F1-score* yang lebih tinggi dibandingkan LSTM. Keunggulan tersebut sangat dipengaruhi oleh efektivitas metode ekstraksi fitur TF-IDF dalam menyeleksi sekaligus merepresentasikan kata-kata kunci ke dalam bentuk numerik. Dalam skenario analisis komentar negatif, keberadaan kosakata spesifik seperti kata kasar atau umpatan eksplisit, kerap menjadi indikator diskriminatif yang kuat untuk menentukan kelas *cyberbullying*. Oleh karena itu, integrasi antara pemetaan matriks TF-IDF dan pencarian *hyperplane* pada algoritma SVM terbukti bekerja secara sangat efektif pada dataset yang digunakan.

Pada perspektif yang berbeda, arsitektur LSTM mendemonstrasikan keunggulannya pada tingkat sensitivitas deteksi, yang direpresentasikan oleh capaian *recall* dan ROC-AUC yang lebih memadai, serta kuantitas *False Negative* yang lebih minim. Kapasitas LSTM dalam mempelajari representasi teks melalui *word embedding* dan memori sekuensial memungkinkannya untuk memahami konteks relasi antarkata, alih-alih sekadar menganalisis frasa secara terpisah. Karakteristik ini menjadikan LSTM sangat mumpuni dalam menangkap pola kalimat perundungan. Kendati demikian, tingginya tingkat sensitivitas tersebut memicu sifat reaktif pada model, sehingga beberapa komentar non-cyberbullying rentan diklasifikasikan secara keliru sebagai *cyberbullying*. Dampaknya, nilai presisi yang dihasilkan oleh LSTM berada di bawah performa SVM.

Beralih pada evaluasi efisiensi model, SVM terbukti jauh lebih ringan dari segi alokasi waktu pelatihan maupun konsumsi memori, lantaran model ini hanya memproses pemisahan fitur secara linier tanpa memerlukan iterasi berbasis epoch. Sebaliknya, LSTM menuntut sumber daya memori dan durasi pelatihan yang jauh lebih masif karena kompleksitas struktur neuralnya saat memperbarui bobot model secara dinamis. Walaupun begitu, LSTM justru memamerkan keunggulan signifikan pada fase implementasi operasional, waktu prediksi terhadap data baru berjalan jauh lebih singkat dibandingkan SVM. Fenomena ini terjadi karena LSTM memproses matriks input berukuran seragam secara *batch*, sementara SVM menuntut komputasi ekstra untuk mengalkulasi transformasi fitur TF-IDF berdimensi tinggi sekaligus menghitung probabilitas klasifikasinya.

Analisis komparatif ini pada akhirnya menunjukkan bahwa pemilihan model yang ideal sangat bergantung pada orientasi implementasi sistem yang akan dibangun. Apabila prioritas pengembangan dititikberatkan pada urgensi untuk

mencegah lolosnya komentar perundungan melalui optimasi nilai *recall*, serta tuntutan kecepatan respons klasifikasi pada aplikasi real-time, arsitektur LSTM menawarkan kapabilitas operasional yang sangat menjanjikan. Di sisi lain, jika tolok ukur utamanya adalah keseimbangan performa yakni kapabilitas dalam mempertahankan tingkat akurasi prediksi yang tinggi sembari mengendalikan beban komputasi dan konsumsi memori seefisien mungkin pendekatan SVM terbukti memberikan nilai *trade-off* yang paling menguntungkan secara menyeluruh pada studi ini.

KESIMPULAN

Penelitian ini telah berhasil mengimplementasikan dan membandingkan kinerja algoritma SVM dan LSTM dalam mendeteksi teks *cyberbullying* pada platform media sosial X (Twitter). Berdasarkan hasil evaluasi, model SVM dengan ekstraksi fitur TF-IDF menunjukkan keunggulan pada tingkat ketepatan prediksi secara holistik, dengan capaian *accuracy* sebesar 0.8352, *precision* 0.8356, dan *F1-score* 0.8363. Sementara itu, model LSTM dengan pendekatan *word embedding* unggul dalam aspek sensitivitas deteksi, yang dibuktikan dengan nilai *recall* sebesar 0.8552 dan ROC-AUC 0.9088, serta keberhasilannya dalam menekan angka kesalahan gagal deteksi atau *False Negative* menjadi lebih rendah, yakni 95 kesalahan berbanding 107 kesalahan pada SVM. Ditinjau dari efisiensi komputasi, model SVM terbukti jauh lebih ringan dengan waktu pelatihan selama 42.90 detik dan konsumsi memori sebesar 74.87 MB, dibandingkan dengan model LSTM yang menuntut beban komputasi lebih besar dengan waktu pelatihan 106.67 detik dan memori 126.80 MB. Meskipun LSTM mencatatkan waktu prediksi yang lebih cepat saat menguji data baru, keseimbangan antara performa akurasi klasifikasi yang tinggi dan efisiensi penggunaan sumber daya menjadikan SVM lebih unggul secara akumulatif. Oleh karena itu, penelitian ini menyimpulkan bahwa pemodelan SVM merupakan metode yang paling optimal secara keseluruhan untuk diterapkan dalam klasifikasi deteksi cyberbullying pada studi ini.

DAFTAR PUSTAKA

Al Farisi, F. A., Perdana, R. S., & Adikara, P. P. (2024). Klasifikasi Intensi dengan Metode Ling Short-Term Memory pada Chatbot Bahasa Indonesia. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11 (5), 1051-1058.

Amanda Putri, T., Gasa Nova, J., Amilia Putri, T., Anita Puriani, R., & Novirson, R. (2025). Tren Penelitian Dampak Kecanduan Media Sosial Bagi Remaja.

Awalia Z Abidin, N. S., Andira Adrianti Sofyan, A., Hasmin, E., & Arifin, A. (2025). Analisis Sentimen Cyberbullying pada "Tiktok" Menggunakan Metode Long Short Term Memory.

Bansal, S., Garg, N., Singh, J., & Van Der Walt, F. (2023). Cyberbullying and mental health: past, present and future. *Frontiers in Psychology*, 14.

Du, K. L., Jiang, B., Lu, J., Hua, J., & Swamy, M. N. S. (2024). Exploring Kernel Machines and Support Vector Machines: Principles, Techniques, and Future Directions. *Mathematics*, 12 (24).

Ganiem, L. M., Setiawati, R., Suardi, S., Nurhayai, N., & Ramdhani, R. (2024). Society in the Digital Era: Adaptation, Change, and Response to

Communication Technology. *Journal International Dakwah and Communication*, 4 (1), 123-135.

He, Q., Hong, Y., & Raghu, T. S. (2022). *The Effects of Machine-powered Content Modereation: An Empirical Study on Reddit*.

Jabardi, M. (2025). Support Vector Machines: Theory, Algorithms, and Applications. *Infocommunications Journal*, 17 (1), 66-75.

Kagi, shivkumar. (2025). *Cyberbullying Detection Using Machine Learning*. www.jsrtjournal.com

Krichen, M., & Mihoub, A. (2025). Long Short-Term Memory Networks: A Comprehensive Survey. *AI (Switzerland)*, 6 (9).

Mienye, I. D., Swart, T. G., & Obaido, G. (2024). Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications. *Information (Switzerland)*, 15 (9).

Ricoy, M. C., Martínez-Carrera, S., & Martínez-Carrera, I. (2022). Social Overview of Smartphone Use by Teenagers. *International Journal of Environmental Research and Public Health*, 19 (22).

Safitri, M., Alhaura Zafira, N., Dyahrestu Putri, A., Putri Pamungkas, S., Azahrah, F., Lestari, A., Kusdiyah, D., & Wayan Parwati, N. (2025). Deteksi Cyberbullying Tweet Menggunakan Machine Learning. *Seminar Nasional Riset dan Inovasi Teknologi (SEMNAS RISTEK)*.

Salwa Aiza Zahra, Betie Febriana, & Wahyu Endang. (2026). Hubungan Kecanduan Media Sosial terhadap Kejadian Cyberbullying pada Remaja di Kota Semarang. *Jurnal Riset Ilmu Kesehatan Umum Dan Farmasi (JRIKUF)*, 4 (1), 182-201.

Tewu, D., Destine, D., & Gunawan, I. (2025). Analysis of Social Media User Growth and Its Implications for Digital Marketing Strategies in Indonesia 2024. *International Journal of Management Studies and Social Science Research*, 07 (03), 236-245.

Wulandari, T. P., & Sujarwo, S. (2024). Psychological and Social Dynamics of Cyberbullying Victims: An Analysis of the Impact on Middle School Students Dinamika Psikologis dan Sosial Korban Cyberbullying: Analisis Dampak pada Peserta Didik Sekolah Menengah Pertama. *Jurnal Imiah Psikologi*, 12, 434-443.