Vol. 5 No. 1 November 2025, hal. 1-5.

ISSN 2808-8468 (Online)
sensasi.upnjatim.ac.id

IMPLEMENTASI MACHINE LEARNING UNTUK MENINGKATKAN KESADARAN MASYARAKAT DALAM DETEKSI SPAM SMS MELALUI PROGRAM MBKM MAGANG MANDIRI DI HACKTIV8 INDONESIA

Raja Valentino Kristananda[1*], Amri Muhaimin[2]
*Email : 21083010068@student.upnjatim.ac.id

Universitas Pembangunan Nasional "Veteran" Jawa Timur[1-2]

Abstrak. Spam adalah pesan-pesan tidak diinginkan yang dikirim secara massal tanpa izin dari penerima, dalam konteks pesan elektronik seperti SMS. Penelitian ini bertujuan untuk mengembangkan model deteksi spam SMS menggunakan pendekatan *deep learning* berbasis *LSTM (Long Short-Term Memory)*. Data yang digunakan adalah dataset SMS yang telah diklasifikasikan sebagai "ham" (tidak spam) atau "spam". Dataset dibagi menjadi set pelatihan dan uji, kemudian dilakukan tokenisasi teks dan padding urutan untuk persiapan data. Model LSTM dibangun dengan lapisan *embedding* dan *dropout* spasial untuk mengatasi *overfitting*, diikuti oleh lapisan LSTM dan lapisan dense dengan fungsi aktivasi sigmoid untuk klasifikasi biner. Hasil eksperimen menunjukkan bahwa model mencapai akurasi sekitar 86.55% pada data uji, dengan nilai kehilangan sebesar 0.395. Studi ini mengonfirmasi bahwa pendekatan *deep learning* dengan LSTM dapat efektif dalam mendeteksi spam SMS, meskipun penting untuk mempertimbangkan kelas tidak seimbang dalam dataset yang digunakan.

Kata kunci: Deteksi Spam SMS, Deep Learning, LSTM, Pengolahan Bahasa Alami.

Abstract. *Spam is unwanted messages sent in bulk without permission from the recipient, particularly in electronic messaging like SMS. This research aims to develop a spam SMS detection model using a deep learning approach based on Long Short-Term Memory (LSTM). The data used consists of SMS datasets classified as "ham" (not spam) or "spam". The dataset is split into training and test sets, followed by text tokenization and sequence padding for data preparation. The LSTM model is constructed with embedding layers, spatial dropout to mitigate overfitting, an LSTM layer, and a dense layer with sigmoid activation for binary classification. Experimental results show the model achieves approximately 86.55% accuracy on the test data, with a loss value of 0.395. This study confirms that the LSTM-based deep learning approach is effective in detecting spam SMS, though careful consideration of class imbalance in the dataset is crucial.*

Keywords: *SMS Spam Detection, Deep Learning, LSTM, Natural Language Processing*.

✉Corresponding author : Raja Valentino Kristananda
Email: 21083010068@student.upnjatim.ac.id

Vol. 5 No. 1 November 2025, hal. 1-5.
ISSN 2808-8468 (Online)
sensasi.upnjatim.ac.id

Pendahuluan

Short Message Service (SMS) atau layanan pesan singkat merupakan salah satu media komunikasi jarak jauh yang masih banyak digunakan pada era sekarang ini untuk mengirim pesan singkat (Dwiyansaputra, 2021). Penggunaan SMS sebagai sarana komunikasi telah meluas secara signifikan dalam berbagai aspek kehidupan, dari komunikasi pribadi hingga transaksi bisnis. Namun, peningkatan pesan spam yang tidak diinginkan telah mengganggu efisiensi dan privasi pengguna. Fenomena ini menjadi perhatian penting karena tidak hanya mengganggu penggunaan layanan SMS secara efektif, tetapi juga mengancam keamanan data pribadi pengguna. Spam atau *stupid pointless annoying messages* merupakan serangan pesan yang dikirimkan ke sejumlah pengguna layanan pesan yang tidak secara khusus meminta pesan tersebut (Hayuningtyas, 2017). Spam juga dapat didefinisikan sebagai pengiriman pesan secara berulang-ulang.

Penggunaan teknologi machine learning, khususnya metode *deep learning* seperti *Long Short-Term Memory* (LSTM). *Long-Short Term Memory* (LSTM) adalah salah satu varian dari algoritma *Recurrent Neural Network* (RNN) yang melakukan penambahan memory cell untuk dapat menyimpan informasi pada jangka waktu yang lama (Cholissodin et al., 2021). LSTM menawarkan solusi potensial dalam meningkatkan deteksi spam SMS. Metode ini memungkinkan pengenalan pola yang lebih kompleks dan adaptif dalam teks pesan, membedakan dengan lebih baik antara pesan yang sah dan spam. Dengan menerapkan model ini, diharapkan dapat meningkatkan akurasi deteksi dan mengurangi jumlah pesan spam yang mencapai pengguna akhir.

Penelitian ini mengangkat pertanyaan tentang bagaimana implementasi model LSTM dapat meningkatkan akurasi dalam deteksi spam SMS dibandingkan dengan pendekatan konvensional. Dengan fokus pada penggunaan dataset yang relevan dan teknik *preprocessing* yang tepat, penelitian ini bertujuan untuk menguji efektivitas model LSTM dalam konteks deteksi spam SMS.

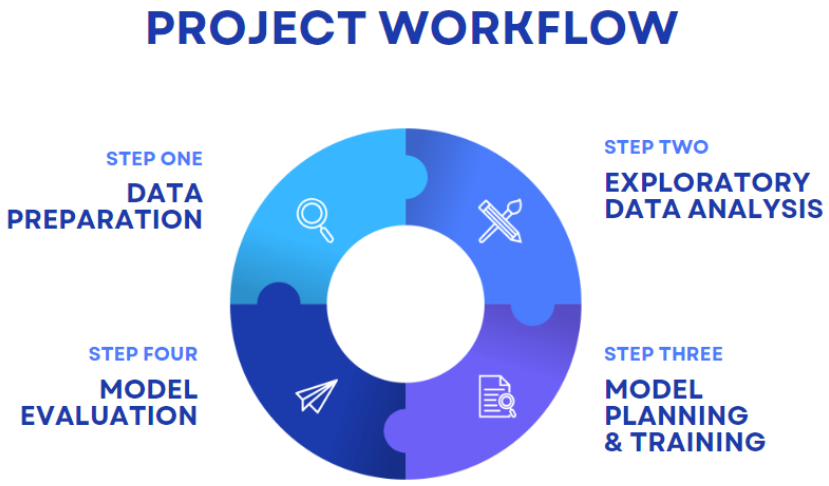
Tujuan penelitian ini adalah mengembangkan model LSTM yang dapat menghasilkan hasil deteksi spam yang lebih akurat dan dapat diandalkan, dengan harapan dapat memberikan panduan praktis bagi pengembangan sistem deteksi spam yang lebih efektif di masa depan. Dengan demikian, penelitian ini tidak hanya berkontribusi pada pengembangan teknologi deteksi spam, tetapi juga menghadirkan solusi yang lebih baik dalam menjaga keamanan dan privasi pengguna layanan SMS.

Metode Pelaksanaan

Penelitian ini menggunakan pendekatan eksperimental untuk menguji efektivitas model LSTM dalam deteksi spam SMS. Data yang digunakan berasal dari dataset publik yang tersedia di Kaggle, yang terdiri dari sampel SMS yang sudah diklasifikasikan menjadi spam dan *non-spam* (ham), dengan variabel label untuk klasifikasi dan *text* untuk konten SMS.

Gambar 1. *Project Workflow*

Vol. 5 No. 1 November 2025, hal. 1-5.
ISSN 2808-8468 (Online)
sensasi.upnjatim.ac.id



Sumber: Dokumen Pribadi (2024)

Proses pengumpulan data melibatkan pengambilan sampel acak dari dataset Kaggle, yang telah diakui keandalannya dalam konteks penelitian ini. Data tersebut kemudian dipersiapkan dengan tahapan *preprocessing* yang meliputi pembersihan teks dan tokenisasi menggunakan alat bantu *Tokenizer* dari Keras. Langkah selanjutnya adalah mengubah teks yang telah diolah menjadi urutan bilangan bulat (*sequences*) menggunakan *Tokenizer*, dan mengaplikasikan *padding* untuk memastikan semua teks memiliki panjang yang seragam.

Implementasi model LSTM dilakukan menggunakan *framework* Keras di atas TensorFlow. Model ini terdiri dari beberapa layer, termasuk layer *embedding* untuk mengubah teks menjadi vektor numerik, serta layer LSTM dengan *dropout* untuk mengurangi *overfitting*. Proses *training* model dilakukan dengan menyesuaikan parameter melalui proses iteratif menggunakan data *training*, sambil memantau performa menggunakan data validasi.

Analisis hasil dilakukan dengan membandingkan metrik performa seperti akurasi, presisi, dan *recall* antara model LSTM yang dihasilkan dengan baseline atau model lain yang tersedia. Evaluasi ini dilakukan secara kualitatif dan kuantitatif untuk menilai efektivitas dan keunggulan model LSTM dalam deteksi spam SMS.

Metode pelaksanaan ini dirancang untuk memberikan evaluasi yang komprehensif terhadap kemampuan model LSTM dalam mengatasi tantangan deteksi spam SMS, dengan harapan memberikan kontribusi signifikan dalam pengembangan teknologi keamanan komunikasi melalui SMS di masa depan.

Hasil dan Pembahasan

Hasil eksperimen ini menunjukkan bahwa model LSTM yang dikembangkan berhasil mencapai akurasi sebesar 86.55% dalam mengklasifikasikan SMS sebagai spam atau non-spam (ham). Evaluasi model juga menunjukkan performa yang baik dalam mengenali SMS spam, dengan nilai *recall* yang cukup tinggi untuk kelas spam, yang menunjukkan kemampuan model dalam mengidentifikasi sebagian besar SMS yang seharusnya ditandai sebagai spam dalam dataset.

Terdapat beberapa faktor yang mempengaruhi kinerja model, termasuk ukuran dataset, teknik penanganan ketidakseimbangan kelas, serta penyetelan parameter seperti dropout rate pada lapisan LSTM. Analisis ini memberikan pemahaman mendalam tentang tantangan dalam deteksi spam SMS menggunakan pendekatan machine learning, dan menyoroti pentingnya strategi yang tepat dalam pengembangan model untuk meningkatkan akurasi dan efisiensi.

Vol. 5 No. 1 November 2025, hal. 1-5.
ISSN 2808-8468 (Online)
sensasi.upnjatim.ac.id

Secara keseluruhan, penggunaan model LSTM dalam deteksi spam SMS menjanjikan untuk diterapkan dalam meningkatkan keamanan komunikasi digital. Temuan ini tidak hanya bermanfaat secara praktis untuk aplikasi keamanan data pribadi pengguna, tetapi juga memberikan landasan untuk penelitian lanjutan dalam pengembangan sistem deteksi spam yang lebih kompleks dan responsif terhadap variasi pola spam yang terus berkembang.

Simpulan

Berdasarkan hasil penelitian ini, penggunaan model LSTM untuk deteksi spam SMS mencapai tingkat akurasi yang signifikan, yaitu 86.55% dalam uji coba. Hal ini menegaskan bahwa pendekatan machine learning dengan teknologi *deep learning* mampu efektif dalam mengidentifikasi dan memblokir pesan spam secara tepat waktu. Implikasi praktis dari penelitian ini adalah potensi penggunaan model ini untuk melindungi pengguna dari gangguan komunikasi yang tidak diinginkan dan berpotensi merugikan. Penelitian ini juga menggarisbawahi perlunya penyesuaian parameter dan manajemen data yang cermat untuk meningkatkan performa model dalam mendeteksi spam SMS di masa mendatang, yang dapat berkontribusi dalam menjaga keamanan dan integritas komunikasi digital.

Ucapan Terimakasih

Saya ingin mengucapkan terima kasih yang sebesar-besarnya atas dukungan dan kontribusi dari berbagai pihak dalam penelitian ini. Terima kasih kepada institusi yang telah memberikan akses data dan sumber daya yang diperlukan. Serta terima kasih kepada semua pihak yang telah memberikan bimbingan, saran, dan dukungan moral selama proses penelitian ini berlangsung. Dukungan ini sangat berarti bagi kelancaran dan kesuksesan penelitian ini.

Daftar Pustaka

Cholissodin, I., Pinasthika, K., Prastiwi, J.H. and Dewantara, R. (2021). *"Social Computing to Create Government Public Policy Document Blueprint Draft Based on Social Media Data About Covid-19 Using LSTM and MMR Hybrid Algorithms. Proceedings of the International Conference on Green Technology",* 11(1), p.6. Diakses dari: https://doi.org/10.18860/icgt.v11i1.1394.

Dwiyansaputra, R., Nugraha, G., Bimantoro, F. dan Aranta, A. (2021). *"Deteksi SMS Spam Berbahasa Indonesia Menggunakan TF-IDF dan Stochastic Gradient Descent Classifier".* JTIKA. Diakses dari: https://jtika.if.unram.ac.id/index.php/JTIKA/article/view/145

R. Y. Hayuningtyas, (2017). "*Aplikasi Filtering of Spam Email Menggunakan Naïve Bayes*" IJCIT (Indonesian J. Comput. Inf. Technol., vol. 2, no. 1, hal. 53–60. Diakses dari: https://journal.uii.ac.id/AUTOMATA/article/view/19514/11556

Chrismanto, A. dan Lukito, Y. (2017). "*Klasifikasi Komentar SPAM pada Instagram Berbahasa Indonesia Menggunakan K-NN*". SNATIK. Diakses dari: https://www.researchgate.net/profile/Antonius-Rachmat/publication/323257693_KLASIFIKASI_KOMENTAR_SPAM_PADA_INSTAGRAM_BERBAHASA_INDONESIA_MENGGUNAKAN_K-NN/links/5a8a6974458515b8af94d649/KLASIFIKASI-KOMENTAR-SPAM-PADA-INSTAGRAM-BERBAHASA-INDONESIA-MENGGUNAKAN-K-NN.pdf