# BAB I PENDAHULUAN

## Latar Belakang

Analisis sentimen telah menjadi salah satu aspek penting dalam Natural Language Processing (NLP) karena kemampuannya untuk memahami opini dan emosi yang terkandung dalam teks. Analisis ini memiliki peran strategis dalam berbagai bidang, seperti layanan pelanggan untuk menganalisis kepuasan pengguna, kesehatan mental untuk mendeteksi potensi masalah emosional, serta media sosial untuk memahami dinamika emosi masyarakat secara real-time (Singgalen, 2024; Madan & Kumar, 2024). Selain itu, dalam konteks pendidikan, analisis emosi dapat digunakan untuk meningkatkan kualitas pembelajaran dengan memahami kebutuhan emosional siswa (Sorour et al., 2024).

Dalam pengembangannya, analisis sentimen telah berkembang dari sekadar klasifikasi polaritas (positif, negatif, netral) menjadi analisis emosi mendalam, seperti deteksi emosi marah, senang, atau sedih (Singh et al., 2024). Namun, pengembangan model NLP berbasis emosi dalam Bahasa Indonesia masih menghadapi sejumlah tantangan, termasuk keterbatasan dataset yang representatif, kompleksitas struktur bahasa, dan kurangnya model yang dirancang khusus untuk Bahasa Indonesia (Phan et al., 2024).

Salah satu pendekatan yang menjanjikan dalam analisis sentimen emosi adalah penggunaan algoritma hybrid CNN-LSTM. Algoritma ini memadukan kekuatan Convolutional Neural Networks (CNN) untuk menangkap fitur lokal dalam teks dan Long Short-Term Memory (LSTM) untuk memahami hubungan temporal dan konteks kata. Pendekatan ini telah terbukti efektif untuk meningkatkan akurasi dalam berbagai aplikasi analisis sentimen (Tran et al., 2024; Parra-Gallego & Arias-Vergara, 2025). Namun, penerapan algoritma hybrid ini untuk teks dalam Bahasa Indonesia masih memerlukan eksplorasi lebih lanjut untuk memastikan efektivitasnya (Das & Pradhan, 2024; Kukreja & Kundra, 2024).

Penelitian ini bertujuan untuk mengembangkan model NLP berbasis emosi untuk analisis sentimen teks Bahasa Indonesia menggunakan algoritma hybrid CNN-LSTM. Melalui penelitian ini, diharapkan dapat dihasilkan model yang tidak hanya mampu meningkatkan akurasi analisis sentimen emosi, tetapi juga memberikan kontribusi pada pengembangan model NLP yang lebih baik untuk Bahasa Indonesia. Selain itu, hasil penelitian ini juga diharapkan dapat menjadi acuan dalam pengembangan aplikasi berbasis NLP untuk berbagai bidang seperti pendidikan, layanan pelanggan, dan kesehatan mental (Chowdhury & Huq, 2024).

## Rumusan Masalah

Analisis sentimen merupakan salah satu cabang penting dalam Natural Language Processing (NLP) yang mampu mengungkap opini dan emosi dari teks. Dalam beberapa tahun terakhir, analisis ini telah berkembang dari sekadar pengklasifikasian polaritas (positif, negatif, netral) menjadi analisis emosi mendalam yang mencakup berbagai kategori emosi seperti marah, sedih, dan bahagia. Namun, pengembangan model NLP berbasis emosi untuk Bahasa Indonesia menghadapi berbagai kendala, seperti keterbatasan dataset representatif yang sesuai dengan struktur unik bahasa ini, kompleksitas pola linguistik, dan minimnya model yang dirancang khusus untuk Bahasa Indonesia. Hal ini menjadikan penerapan analisis sentimen berbasis emosi dalam Bahasa Indonesia belum optimal jika dibandingkan dengan bahasa global seperti Inggris.

Di sisi lain, algoritma hybrid CNN-LSTM menawarkan solusi yang menjanjikan dengan memadukan kemampuan Convolutional Neural Networks (CNN) untuk menangkap pola lokal dalam teks dan Long Short-Term Memory (LSTM) untuk memahami hubungan temporal antar kata. Pendekatan ini telah terbukti meningkatkan akurasi analisis sentimen dalam berbagai bahasa dan domain. Namun, penerapannya untuk teks Bahasa Indonesia masih kurang tereksplorasi. Pertanyaan yang muncul adalah bagaimana pendekatan ini dapat diadaptasi secara efektif untuk Bahasa Indonesia agar dapat mengatasi keterbatasan yang ada.

Berdasarkan permasalahan tersebut, diperlukan penelitian yang mampu menjawab tantangan ini dengan mengembangkan model analisis sentimen berbasis emosi menggunakan algoritma hybrid CNN-LSTM. Penelitian ini juga perlu memastikan bahwa model yang dihasilkan mampu memberikan kinerja yang lebih akurat, relevan, dan aplikatif di berbagai bidang seperti layanan pelanggan, pendidikan, dan kesehatan mental.

## Tujuan Penelitian

Penelitian ini bertujuan untuk mengembangkan model analisis sentimen berbasis emosi untuk teks dalam Bahasa Indonesia menggunakan algoritma hybrid CNN-LSTM. Tujuan utama dari penelitian ini adalah untuk meningkatkan akurasi analisis sentimen emosi dengan memanfaatkan kemampuan CNN dalam mengenali pola-pola lokal dalam teks dan LSTM dalam memahami hubungan temporal dan konteks antar kata.

Secara lebih spesifik, penelitian ini memiliki beberapa tujuan sebagai berikut:

1. Mengeksplorasi kemampuan algoritma hybrid CNN-LSTM dalam mengatasi tantangan unik struktur bahasa Indonesia, seperti kekayaan morfologi, sintaksis, dan semantik.
2. Mengidentifikasi dan memanfaatkan dataset yang representatif untuk Bahasa Indonesia, sekaligus mengatasi keterbatasan data dengan teknik augmentasi atau metode lain yang relevan.
3. Menguji dan mengevaluasi performa model yang diusulkan pada berbagai kategori emosi, sehingga model ini dapat digunakan untuk berbagai aplikasi, mulai dari layanan pelanggan hingga analisis sosial dan pendidikan.
4. Memberikan kontribusi pada pengembangan teknologi pemrosesan bahasa alami (NLP) yang lebih adaptif terhadap konteks dan kebutuhan lokal.

Melalui pencapaian tujuan-tujuan tersebut, penelitian ini diharapkan mampu menjadi solusi yang efektif dalam mendorong inovasi NLP untuk Bahasa Indonesia.

## Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat signifikan baik dari segi teoretis maupun praktis, yang dapat dirinci sebagai berikut:

1. **Manfaat Teoretis:**

Penelitian ini berkontribusi pada literatur ilmiah mengenai analisis sentimen berbasis emosi, terutama dalam pengembangan model NLP yang dirancang khusus untuk Bahasa Indonesia. Temuan dari penelitian ini akan memperkaya wawasan mengenai efektivitas pendekatan hybrid CNN-LSTM dalam menangani bahasa dengan struktur kompleks seperti Bahasa Indonesia. Model yang dihasilkan dapat menjadi referensi untuk penelitian lanjutan di bidang NLP, khususnya dalam pengembangan metode analisis sentimen berbasis emosi yang lebih presisi dan aplikatif.

1. **Manfaat Praktis:**

Penelitian ini juga memiliki manfaat praktis yang luas dalam berbagai bidang:

* 1. **Layanan Pelanggan:** Model yang diusulkan dapat digunakan untuk menganalisis kepuasan pelanggan secara otomatis, sehingga perusahaan dapat mengambil keputusan strategis berdasarkan opini pelanggan.
  2. **Pendidikan:** Dalam dunia pendidikan, model ini dapat membantu pengajar memahami kondisi emosional siswa melalui analisis respon atau umpan balik, sehingga proses pembelajaran dapat lebih personal dan efektif.
  3. **Kesehatan Mental:** Dalam bidang kesehatan mental, model ini dapat diterapkan untuk mendeteksi emosi negatif seperti stres atau depresi dari teks yang dihasilkan oleh pengguna, misalnya di media sosial atau survei.
  4. **Media Sosial dan Analisis Sosial:** Model ini dapat digunakan untuk memahami dinamika emosi masyarakat secara real-time, yang bermanfaat dalam penelitian sosial, pengelolaan krisis, atau bahkan pengembangan kebijakan publik.

1. **Manfaat Lokal:**

Dengan berfokus pada Bahasa Indonesia, penelitian ini akan memberikan dampak langsung pada pengembangan teknologi yang lebih inklusif, relevan, dan sesuai dengan kebutuhan masyarakat Indonesia. Hal ini diharapkan dapat mempercepat adopsi teknologi NLP dalam berbagai sektor di Indonesia.

Secara keseluruhan, penelitian ini tidak hanya diharapkan meningkatkan akurasi analisis sentimen berbasis emosi, tetapi juga menjadi pijakan penting dalam pengembangan aplikasi berbasis NLP yang mendukung berbagai inovasi di Indonesia.

**Daftar Referensi**

1. Singgalen, Y. A. (2024). *A Hybrid CNN-LSTM Model with SMOTE for Enhanced Sentiment Analysis of Hotel Reviews*. ResearchGate. [Link PDF](https://www.researchgate.net/publication/386378606_A_Hybrid_CNN-LSTM_Model_with_SMOTE_for_Enhanced_Sentiment_Analysis_of_Hotel_Reviews)
2. Madan, A., & Kumar, D. (2024). *Real-time topic-based sentiment analysis for movie tweets using hybrid approach*. Knowledge and Information Systems, Springer. [DOI Link](https://link.springer.com/article/10.1007/s10115-024-02298-x)
3. Sorour, S. E., Alojail, A., El-Shora, A., & Amin, A. E. (2024). *A Hybrid Deep Learning Approach for Enhanced Sentiment Classification and Consistency Analysis in Customer Reviews*. ProQuest. [Full Text](https://search.proquest.com/openview/e894b27e123a40e845721ef3787e1eef/1)
4. Singh, A. K., Bhushan, A., & Dwivedi, D. (2024). *Analyzing Sentiments on Twitter Using Deep Learning Techniques*. International Journal of Modern Education and Computer Science. [PDF Full Text](https://www.mecs-press.org/ijmecs/ijmecs-v16-n6/IJMECS-V16-N6-2.pdf)
5. Phan, H. T., Seo, Y. S., & Nguyen, N. T. (2024). *Fuzzy Hybrid CNN-LSTM Model for Sentence-Level Sentiment Analysis*. SSRN. [PDF Full Text](https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=4994871)
6. Tran, X. T., Nguyen, D. M., & Nguyen, M. T. (2024). *A Comprehensive Framework for Sentiment Analysis and Cold-Start Recommendations in Vietnam Hospitality Sector*. Springer. [DOI Link](https://link.springer.com/chapter/10.1007/978-981-96-0178-3_19)
7. Parra-Gallego, L. F., & Arias-Vergara, T. (2025). *Multimodal Evaluation of Customer Satisfaction from Voicemails Using Speech and Language Representations*. Digital Signal Processing, Elsevier. [DOI Link](https://www.sciencedirect.com/science/article/pii/S1051200424004457)
8. Das, S., & Pradhan, B. (2024). *Advanced Deep Learning Models for Automatic Detection of Driver's Facial Expressions, Movements, and Alertness in Varied Lighting Conditions: A Comparative Analysis*. Springer. [DOI Link](https://link.springer.com/article/10.1007/s11042-024-20428-z)
9. Kukreja, V., & Kundra, D. (2024). *Deep Learning for Sentiment Analysis: Combining CNN and LSTM for Enhanced Performance on Facebook Data*. IEEE Xplore. [DOI Link](https://ieeexplore.ieee.org/abstract/document/10760800)
10. Chowdhury, T. A., & Huq, M. R. (2024). *An Optimized Bangla Speech Emotion Recognition System Leveraging CNN-LSTM and Boosting Classifiers*. IEEE Xplore. [DOI Link](https://ieeexplore.ieee.org/abstract/document/10756428)