



Pendekatan Komputasi dalam *Sentiment Analysis* Produk pada Media Sosial

Achmad Solechan

Program Studi Bisnis Digital, Universitas Telogorejo Semarang, Indonesia

Penulis Korespondensi: achmad_solechan@universitastelogorejo.ac.id

Abstract. *The rapid growth of social media has generated large amounts of user-generated content containing opinions about products and services. This condition encourages the use of sentiment analysis to extract useful insights for business decision-making. Along with advances in Artificial Intelligence and Natural Language Processing, computational approaches to sentiment analysis have evolved from lexicon-based methods, machine learning, deep learning, transformer-based models, to Large Language Models (LLMs). This study aims to analyze the development of computational approaches in product sentiment analysis on social media using a Narrative Literature Review method. The review synthesizes 25 relevant scientific articles from reputable national and international journals. The findings show that transformer-based models, especially BERT and its variants, generally provide better classification performance than traditional machine learning and deep learning methods because they can capture contextual meaning more effectively. In addition, LLMs such as GPT show strong potential for analyzing complex and limited data. This study provides an overview of the evolution, strengths, limitations, and future research opportunities of computational approaches in sentiment analysis.*

Keywords: *Large Models; Machine Learning; Natural Language; Sentiment Analysis; Social Media.*

Abstrak. Perkembangan media sosial menghasilkan banyak *user-generated content* yang berisi opini pengguna terhadap produk dan layanan. Kondisi ini mendorong penggunaan *sentiment analysis* untuk mengekstraksi informasi yang dapat mendukung pengambilan keputusan bisnis. Seiring kemajuan *Artificial Intelligence* dan *Natural Language Processing*, pendekatan komputasi dalam *sentiment analysis* berkembang dari metode berbasis kamus, *machine learning*, *deep learning*, *transformer*, hingga *Large Language Models* (LLMs). Penelitian ini bertujuan menganalisis perkembangan pendekatan komputasi dalam *sentiment analysis* produk pada media sosial menggunakan metode *Narrative Literature Review*. Kajian dilakukan dengan mensintesis 25 artikel ilmiah yang relevan dari jurnal nasional dan internasional bereputasi. Hasil kajian menunjukkan bahwa model berbasis *transformer*, khususnya BERT dan variannya, secara umum memberikan performa klasifikasi lebih baik dibandingkan metode *machine learning* dan *deep learning* konvensional karena mampu memahami konteks bahasa secara lebih efektif. Selain itu, LLM seperti GPT memiliki potensi besar dalam menganalisis data yang kompleks maupun terbatas. Penelitian ini memberikan gambaran mengenai evolusi, keunggulan, keterbatasan, serta peluang penelitian masa depan dari pendekatan komputasi dalam *sentiment analysis*.

Kata kunci: *Large Models; Machine Learning; Media Sosial; Natural Language; Sentiment Analysis.*

1. LATAR BELAKANG

Perkembangan teknologi informasi dan internet telah mendorong pertumbuhan media sosial sebagai salah satu sumber utama informasi yang dihasilkan oleh pengguna (*user-generated content*). Platform seperti Twitter (X), Facebook, Instagram, TikTok, YouTube, dan berbagai platform e-commerce memungkinkan pengguna untuk menyampaikan pengalaman, opini, kritik, maupun rekomendasi terhadap suatu produk atau layanan secara real time. Informasi tersebut menjadi aset yang sangat berharga bagi perusahaan karena mampu menggambarkan persepsi konsumen secara langsung dan berpengaruh terhadap citra merek, loyalitas pelanggan, serta keputusan pembelian. Oleh karena itu, perusahaan tidak lagi hanya mengandalkan survei konvensional, tetapi mulai memanfaatkan data media sosial sebagai

sumber informasi dalam mendukung strategi pemasaran dan pengambilan keputusan berbasis data (Kaur & Sharma, 2023); (Rahardjo et al., 2024); (Hakim & Azizli, 2024).

Besarnya volume data yang dihasilkan setiap hari menyebabkan proses analisis secara manual menjadi tidak efisien. Selain jumlahnya yang sangat besar, data pada media sosial memiliki karakteristik tidak terstruktur, mengandung bahasa informal, singkatan, emoji, *slang*, hingga penggunaan bahasa campuran yang menyulitkan proses analisis. Kondisi tersebut mendorong berkembangnya sentiment analysis sebagai salah satu cabang *Natural Language Processing* (NLP) yang bertujuan mengidentifikasi dan mengklasifikasikan opini pengguna ke dalam kategori positif, negatif, maupun netral. Sentiment analysis telah dimanfaatkan secara luas pada berbagai bidang, seperti analisis ulasan produk, evaluasi kepuasan pelanggan, prediksi perilaku konsumen, pengembangan produk, pemantauan reputasi merek, hingga pengambilan keputusan bisnis (Elmitwalli & Mehegan, 2024); (Rajesh & Hiwarkar, 2023); (Septiani et al., 2024).

Seiring meningkatnya kompleksitas data media sosial, pendekatan komputasi yang digunakan dalam sentiment analysis juga mengalami perkembangan yang sangat pesat. Pada tahap awal, penelitian lebih banyak menggunakan lexicon-based approach, yaitu metode yang mengandalkan kamus sentimen untuk menentukan polaritas suatu teks. Pendekatan ini relatif sederhana dan mudah diimplementasikan, namun memiliki keterbatasan dalam memahami konteks kalimat, ironi, sarkasme, maupun kata-kata baru yang terus berkembang di media sosial. Selanjutnya, pendekatan machine learning mulai banyak diterapkan dengan memanfaatkan algoritma seperti Naïve Bayes, Support Vector Machine (SVM), Decision Tree, Random Forest, dan Logistic Regression. Pendekatan ini mampu meningkatkan akurasi klasifikasi karena mempelajari pola dari data berlabel, tetapi masih bergantung pada proses *feature engineering* yang cukup kompleks (Li, 2022); (Elmitwalli & Mehegan, 2024); (Rahardjo et al., 2024).

Perkembangan berikutnya ditandai dengan munculnya pendekatan deep learning yang mampu mempelajari representasi data secara otomatis tanpa memerlukan rekayasa fitur secara manual. Arsitektur seperti Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), serta berbagai model hibrida CNN-LSTM terbukti mampu meningkatkan performa sentiment analysis, khususnya dalam menangkap hubungan kontekstual antar kata. Beberapa penelitian menunjukkan bahwa model hybrid menghasilkan akurasi yang lebih tinggi dibandingkan pendekatan machine learning konvensional karena mampu mengombinasikan keunggulan

ekstraksi fitur lokal dan hubungan semantik jangka panjang (Kaur & Sharma, 2023); (Rajesh & Hiwarkar, 2023); (Jain et al., 2025); (Septiani et al., 2024).

Dalam beberapa tahun terakhir, perkembangan Transformer telah mengubah paradigma sentiment analysis secara signifikan. Model seperti Bidirectional Encoder Representations from Transformers (BERT), RoBERTa, DistilBERT, hingga berbagai model berbasis transformer lainnya mampu memahami konteks kalimat secara lebih mendalam dibandingkan model deep learning sebelumnya. Penelitian terbaru bahkan menunjukkan bahwa model Large Language Models (LLMs) seperti GPT-3 dan GPT-4 memiliki kemampuan generalisasi yang sangat baik, terutama pada data dengan anotasi terbatas. (Elmitwalli & Mehegan, 2024) menunjukkan bahwa BERT memperoleh nilai *F1-score* tertinggi pada dataset standar, sedangkan GPT-3 memberikan performa terbaik pada dataset spesifik dengan jumlah data berlabel yang terbatas. Penelitian lain juga menunjukkan bahwa model BERT yang telah dioptimasi mampu mencapai tingkat akurasi sangat tinggi pada analisis komentar media sosial sehingga semakin memperkuat dominasi transformer dalam penelitian sentiment analysis modern (Elmitwalli & Mehegan, 2024); (Das et al., 2025); (Jaber et al., 2025).

Meskipun berbagai pendekatan komputasi telah dikembangkan dan menunjukkan peningkatan performa yang signifikan, penelitian yang ada umumnya masih berfokus pada perbandingan algoritma tertentu, penggunaan dataset tunggal, atau domain aplikasi yang spesifik, seperti Amazon Reviews, Twitter, maupun platform e-commerce tertentu. Selain itu, sebagian besar penelitian hanya mengevaluasi performa model berdasarkan akurasi klasifikasi tanpa memberikan sintesis yang komprehensif mengenai perkembangan pendekatan komputasi, karakteristik masing-masing metode, kelebihan, keterbatasan, serta arah perkembangan teknologi sentiment analysis secara menyeluruh. Kondisi tersebut menyebabkan peneliti maupun praktisi masih mengalami kesulitan dalam menentukan pendekatan komputasi yang paling sesuai dengan karakteristik data dan tujuan analisis yang berbeda-beda (Kaur & Sharma, 2023); (Jaber et al., 2025); (Hakim & Azizli, 2024).

Berdasarkan kondisi tersebut, diperlukan suatu kajian literatur yang mampu menyajikan sintesis secara komprehensif mengenai perkembangan pendekatan komputasi dalam sentiment analysis produk pada media sosial. Penelitian ini menggunakan metode *Narrative Literature Review* (NLR) untuk mengidentifikasi, mengelompokkan, membandingkan, dan menganalisis berbagai pendekatan komputasi mulai dari *lexicon-based approach*, *machine learning*, *deep learning*, *transformer*, hingga *Large Language Models* (LLMs). Hasil penelitian diharapkan dapat memberikan gambaran mengenai tren perkembangan penelitian, keunggulan dan keterbatasan setiap pendekatan, serta mengidentifikasi peluang penelitian di masa mendatang.

Selain memberikan kontribusi akademik sebagai referensi bagi peneliti, penelitian ini juga diharapkan dapat menjadi dasar pertimbangan bagi praktisi dalam memilih pendekatan komputasi yang paling sesuai untuk pengembangan sistem sentiment analysis pada berbagai platform media sosial.

2. KAJIAN TEORITIS

Sentiment Analysis

Sentiment analysis atau *opinion mining* merupakan bagian dari *Natural Language Processing* (NLP) yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini, emosi, serta sikap pengguna ke dalam sentimen positif, negatif, atau netral. Teknik ini banyak digunakan untuk menganalisis ulasan produk, komentar media sosial, kepuasan pelanggan, reputasi merek, dan perilaku konsumen (Rajesh & Hiwarkar, 2023; Septiani et al., 2024).

Dalam konteks bisnis digital, sentiment analysis membantu perusahaan memahami persepsi konsumen terhadap produk secara lebih cepat dan berbasis data. Hasil analisis dapat digunakan untuk mendukung strategi pemasaran, pengembangan produk, peningkatan layanan, serta prediksi keberhasilan produk di pasar (Kaur & Sharma, 2023; Hakim & Azizli, 2024).

Media Sosial sebagai Sumber Data

Media sosial menjadi sumber data penting karena memuat *user-generated content* berupa komentar, ulasan, kritik, dan pengalaman konsumen terhadap suatu produk. Platform seperti Twitter/X, Instagram, Facebook, TikTok, YouTube, dan marketplace memungkinkan perusahaan memperoleh umpan balik konsumen secara real time (Rahardjo et al., 2024; Hakim & Azizli, 2024).

Namun, data media sosial memiliki karakteristik tidak terstruktur, mengandung bahasa informal, singkatan, emoji, *slang*, sarkasme, serta campuran bahasa. Kondisi ini menyebabkan analisis sentimen menjadi lebih kompleks dan membutuhkan pendekatan komputasi yang mampu memahami konteks bahasa secara lebih baik (Das et al., 2025; Jaber et al., 2025).

Pendekatan Komputasi dalam Sentiment Analysis

Pendekatan komputasi dalam sentiment analysis berkembang dari metode sederhana menuju model yang lebih kompleks. Pendekatan lexicon-based menggunakan kamus sentimen untuk menentukan polaritas teks. Metode ini mudah diterapkan, tetapi kurang mampu memahami konteks, sarkasme, dan variasi bahasa media sosial (Li, 2022); (Elmitwalli & Mehegan, 2024).

Pendekatan machine learning menggunakan algoritma seperti Naïve Bayes, Support Vector Machine, Logistic Regression, Decision Tree, dan Random Forest. Pendekatan ini

mampu meningkatkan akurasi karena belajar dari data berlabel, tetapi masih bergantung pada proses *feature engineering* seperti TF-IDF, BoW, dan *n-gram* (Rahardjo et al., 2024); (Jaber et al., 2025).

Selanjutnya, pendekatan deep learning seperti CNN, RNN, LSTM, Bi-LSTM, dan CNN-LSTM mampu mempelajari fitur teks secara otomatis. Model ini lebih baik dalam menangkap hubungan semantik dan konteks kalimat dibandingkan machine learning konvensional (Kaur & Sharma, 2023); (Rajesh & Hiwarkar, 2023); (Septiani et al., 2024).

Perkembangan terbaru ditandai dengan penggunaan Transformer dan Large Language Models (LLMs) seperti BERT, RoBERTa, GPT-3, dan GPT-4. Model ini mampu memahami konteks bahasa secara lebih mendalam dan menunjukkan performa tinggi pada berbagai tugas sentiment analysis, terutama pada data yang kompleks dan bersifat kontekstual (Elmitwalli & Mehegan, 2024); (Das et al., 2025).

3. METODE PENELITIAN

Penelitian ini menggunakan metode Narrative Literature Review (NLR) untuk mengkaji pendekatan komputasi dalam *sentiment analysis* produk pada media sosial. NLR dipilih karena bersifat fleksibel dan tematik, sehingga sesuai untuk menyintesis perkembangan metode, tren penelitian, kelebihan, keterbatasan, serta peluang penelitian terdahulu terkait *sentiment analysis* berbasis produk, ulasan konsumen, media sosial, dan e-commerce. Kajian pustaka merupakan proses penting dalam penelitian karena membantu peneliti memahami teori, konsep, hasil penelitian sebelumnya, serta menemukan celah penelitian yang masih dapat dikembangkan (Sugiyono, 2019). Selain itu, pendekatan kualitatif dalam kajian literatur memungkinkan peneliti melakukan interpretasi mendalam terhadap berbagai sumber ilmiah yang berkaitan dengan fenomena yang diteliti (Moleong, 2017).

Pencarian artikel dilakukan menggunakan kata kunci: product sentiment analysis social media, sentiment analysis product reviews machine learning, social media sentiment analysis deep learning, BERT product review sentiment analysis, aspect-based sentiment analysis product reviews, dan consumer sentiment social media analytics. Kata kunci tersebut digunakan untuk menemukan artikel yang membahas analisis sentimen pada ulasan produk, media sosial, e-commerce, serta pendekatan komputasi seperti *machine learning*, *deep learning*, BERT, Transformer, dan LLM (Ghatora et al., 2024; Ashbaugh & Zhang, 2024; Chen et al., 2025).

Kriteria inklusi dalam penelitian ini adalah: artikel terbit pada periode 2015–2026, membahas *sentiment analysis* atau *opinion mining*, menggunakan data produk, ulasan

konsumen, media sosial, atau e-commerce, menerapkan pendekatan komputasi seperti lexicon-based, machine learning, deep learning, BERT/Transformer, atau Large Language Models, serta tersedia dalam bentuk *full-text* atau *open access* bila memungkinkan. Artikel yang tidak relevan dengan analisis sentimen produk, tidak menjelaskan metode komputasi, duplikat, atau tidak tersedia naskah lengkap dikeluarkan dari analisis.

Berdasarkan proses seleksi tersebut, diperoleh 25 artikel yang dianalisis dan disintesis secara naratif. Artikel-artikel tersebut kemudian dikelompokkan berdasarkan jenis pendekatan komputasi, yaitu *lexicon-based*, *machine learning*, *deep learning*, *hybrid model*, Transformer/BERT, dan LLM. Sintesis dilakukan dengan membandingkan fokus penelitian, dataset, algoritma, hasil utama, kelebihan, dan keterbatasan tiap artikel. Pendekatan ini memungkinkan penelitian menghasilkan gambaran menyeluruh mengenai perkembangan pendekatan komputasi dalam *sentiment analysis* produk pada media sosial.

4. HASIL DAN PEMBAHASAN

Ringkasan Penelitian Terdahulu

Tabel 1 menyajikan sintesis penelitian terdahulu yang menjadi dasar dalam mengidentifikasi perkembangan pendekatan komputasi pada sentiment analysis produk di media sosial. Sintesis dilakukan terhadap 25 artikel yang dipublikasikan pada jurnal internasional maupun nasional bereputasi dengan fokus pada metode komputasi, sumber data, temuan utama, serta kesenjangan penelitian.

Tabel 1. Ringkasan Riset Terdahulu.

No.	Penulis dan Tahun	Objek Penelitian	Pendekatan Komputasi	Dataset/Media	Hasil Penelitian
1	(Elmitwalli & Mehegan, 2024)	Amazon Product Review	Machine Learning, Deep Learning, Transformer, GPT-4	Amazon Reviews	Transformer memberikan akurasi tertinggi pada dataset standar.
2	(Kaur & Sharma, 2023)	Consumer Review	Hybrid Feature + LSTM	Consumer Reviews	Hybrid feature extraction meningkatkan precision dan F1-score.
3	(Rajesh & Hiwarkar, 2023)	Text Sentiment Analysis	CNN + Bi-LSTM + Attention	IMDB, Review Dataset	Attention meningkatkan akurasi hingga lebih dari 94%.
4	(Jain et al., 2025)	Multilingual Sentiment Analysis	CNN-LSTM-GWO	Multilingual Dataset	Optimasi GWO meningkatkan performa CNN-LSTM.

No.	Penulis dan Tahun	Objek Penelitian	Pendekatan Komputasi	Dataset/Media	Hasil Penelitian
5	(Jaber et al., 2025)	Arabic Sentiment Analysis	Ensemble Machine Learning	Arabic Twitter	Ensemble memberikan performa lebih stabil dibanding model tunggal.
6	(Das et al., 2025)	Social Media Comment	Improved BERT	Social Media	Fine-tuning BERT meningkatkan klasifikasi sentimen.
7	(Rahardjo et al., 2024)	Consumer Behavior	Random Forest, Naïve Bayes, Logistic Regression	Twitter, Facebook, Instagram	Random Forest memperoleh akurasi terbaik.
8	(Hakim & Azizli, 2024)	Product Launch	SVM, Random Forest	Twitter, Instagram	Sentimen positif berkorelasi dengan keberhasilan produk.
9	(Septiani et al., 2024)	Social Media	CNN, LSTM, CNN-LSTM	Twitter	CNN-LSTM menghasilkan performa terbaik.
10	(Tabssam et al., 2025)	Duolingo & Babbel Reviews	Machine Learning + Network Analysis	App Reviews	Sentiment analysis mampu mengidentifikasi pengalaman pengguna.
11	(Duru & Sunar, 2025)	Amazon Reviews	ML, LSTM, Transformer, embeddings	Amazon Magazine & Gift Card Reviews	DistilBERT mencapai akurasi tertinggi 92% dan loss terendah; FastText lebih baik daripada GloVe pada model tradisional dan LSTM.
12	(Ghatora et al., 2024)	Product Reviews	RF, NB, SVM, GPT-4	Product review text	SVM efektif pada teks pendek, sedangkan GPT-4 lebih unggul pada teks panjang dan kaya konteks.
13	Ashbaugh & Zhang (2024)	Customer Reviews	CNN, RNN, LR, RF, NB	Amazon Product Reviews	ML dan DL memiliki kelebihan masing-masing dalam klasifikasi sentimen pelanggan.

No.	Penulis dan Tahun	Objek Penelitian	Pendekatan Komputasi	Dataset/Media	Hasil Penelitian
14	(Sasikala & Sheela, 2020)	Online Product Reviews	DLMNN dan IANFIS	Online product review	DLMNN digunakan untuk klasifikasi sentimen, sedangkan IANFIS untuk prediksi produk masa depan.
15	(Fang & Zhan, 2015)	Product Review Data	Sentiment polarity categorization	Amazon Reviews	Mengusulkan proses umum kategorisasi polaritas pada level kalimat dan review.
16	(Liu et al., 2020)	E-commerce Product Reviews	BERT-BiGRU-Softmax	>500 ribu review e-commerce	Model BERT-BiGRU-Softmax mencapai akurasi lebih dari 95,5%.
17	(Cao et al., 2022)	Agricultural Product Reviews	Improved BERT	Online agricultural product reviews	Improved BERT memperoleh F1-score 89,86%, meningkat dari BERT dasar.
18	(Bellar et al., 2024)	E-commerce Recommendations	CNN, RNN, Bi-LSTM, BERT, FastText, Word2Vec	Online women's clothing review	Deep learning dan embedding transformer mendukung prediksi rating/sentimen produk.
19	(Vadla et al., 2024)	Product Design	BERT dan T5	Amazon eco-friendly product reviews	BERT memperoleh akurasi 92% dan lebih efisien dibanding T5 untuk deteksi aspek.
20	(Gooljar et al., 2024)	Online Purchase	Predictive model dan sentiment analysis	Review dan data pembelian online	Sentiment-based predictive model membantu memahami preferensi pelanggan dalam Marketing 5.0.
21	(Ramadhan et al., 2023)	E-commerce Product Reviews	Naïve Bayes dan SVM	Tokopedia wireless earphone reviews	Naïve Bayes dengan tuning menghasilkan akurasi 77,03%, sedangkan SVM unggul pada presisi.

No.	Penulis dan Tahun	Objek Penelitian	Pendekatan Komputasi	Dataset/Media	Hasil Penelitian
22	(Chen et al., 2025)	ABSA Research	Bibliometric mapping dan topic modeling	1.325 artikel ABSA	ABSA berkembang pesat dan menjadi pendekatan penting untuk analisis sentimen berbasis aspek.
23	(Susanto et al., 2025)	Social Media Sentiment for SMEs	Dictionary-based, NB, SVM, hybrid framework	Facebook dan Instagram UMKM	ML lebih baik daripada dictionary-based untuk bahasa informal; hybrid framework potensial untuk UMKM.
24	(Daraghmi & Zyadeh, 2026)	Amazon Product Reviews	RF, LR, SVM, NB, GBM, CNN, LSTM, RoBERTa	Fine Food dan Unlocked Mobile Reviews	Oversampling meningkatkan akurasi; RF unggul pada ML, RoBERTa unggul pada DL/Transformer.
25	(Li, 2022)	Brand Potential User Identification	Sentiment dictionary, topic model, resampling	Social media text corpus	Domain sentiment lexicon dan topic-based model membantu identifikasi calon pelanggan potensial.

Tren Perkembangan Pendekatan Komputasi

Hasil sintesis terhadap 25 artikel menunjukkan bahwa penelitian mengenai sentiment analysis produk pada media sosial mengalami perkembangan yang sangat pesat selama lima tahun terakhir. Perkembangan tersebut tidak hanya ditunjukkan oleh peningkatan jumlah publikasi, tetapi juga oleh evolusi pendekatan komputasi yang digunakan. Pada periode awal, penelitian lebih banyak mengandalkan pendekatan berbasis kamus (*lexicon-based approach*) yang memanfaatkan daftar kata positif dan negatif sebagai dasar klasifikasi sentimen. Pendekatan ini dipilih karena implementasinya sederhana dan tidak memerlukan data berlabel. Namun demikian, pendekatan tersebut memiliki keterbatasan dalam memahami konteks kalimat, ironi, sarkasme, maupun penggunaan bahasa tidak baku yang umum ditemukan pada media sosial. Keterbatasan tersebut mendorong peneliti mengembangkan pendekatan berbasis pembelajaran mesin yang mampu mempelajari pola dari data secara otomatis.

Machine Learning menjadi pendekatan yang paling dominan pada penelitian periode 2018–2022. Algoritma seperti *Naïve Bayes*, *Support Vector Machine (SVM)*, *Decision Tree*,

Logistic Regression, dan *Random Forest* banyak digunakan karena memiliki kompleksitas komputasi yang relatif rendah serta mampu menghasilkan akurasi yang baik pada dataset berukuran kecil hingga menengah. Rahardjo et al. (2024) menunjukkan bahwa *Random Forest* memperoleh performa terbaik dibandingkan *Naïve Bayes* dan *Logistic Regression* dalam mengidentifikasi sentimen konsumen pada data Twitter, Facebook, dan Instagram. Hasil tersebut memperlihatkan bahwa metode ensemble mampu meningkatkan stabilitas klasifikasi dibandingkan algoritma tunggal.

Meskipun demikian, pendekatan Machine Learning masih memiliki keterbatasan karena memerlukan proses *feature engineering* yang cukup kompleks. Pemilihan fitur seperti TF-IDF, Bag of Words, maupun Word2Vec sangat memengaruhi performa model. Selain itu, hubungan semantik antar kata belum dapat dipahami secara optimal sehingga performanya menurun ketika dihadapkan pada data media sosial yang mengandung bahasa informal, emoji, dan singkatan. Kondisi ini mendorong berkembangnya pendekatan *Deep Learning* yang mampu melakukan ekstraksi fitur secara otomatis tanpa memerlukan rekayasa fitur manual.

Penelitian-penelitian terbaru menunjukkan bahwa arsitektur CNN, LSTM, Bi-LSTM, serta kombinasi CNN-LSTM mampu meningkatkan kemampuan model dalam memahami hubungan antar kata. CNN efektif mengekstraksi pola lokal dari teks, sedangkan LSTM mampu menangkap ketergantungan jangka panjang sehingga kombinasi keduanya menghasilkan performa yang lebih baik dibandingkan penggunaan masing-masing model secara terpisah. (Kaur & Sharma, 2023) melaporkan bahwa pendekatan hybrid berbasis LSTM menghasilkan nilai F1-score yang lebih tinggi dibandingkan model konvensional, sedangkan (Rajesh & Hiwarkar, 2023) menunjukkan bahwa penambahan mekanisme *attention* mampu meningkatkan akurasi klasifikasi hingga lebih dari 94%.

Perkembangan paling signifikan terjadi dengan hadirnya model Transformer, terutama BERT, RoBERTa, DistilBERT, dan berbagai variannya. Model-model tersebut memanfaatkan mekanisme *self-attention* sehingga mampu memahami hubungan kontekstual antar kata secara lebih baik dibandingkan CNN maupun LSTM. (Das et al., 2025) menunjukkan bahwa model Improved BERT menghasilkan performa klasifikasi yang lebih tinggi pada data komentar media sosial, sedangkan Elmitwalli dan Mehegan (2024) menunjukkan bahwa BERT menjadi model dengan performa terbaik pada sebagian besar dataset standar sentiment analysis.

Framework Konseptual Perkembangan Pendekatan Komputasi dalam Sentiment Analysis

Berdasarkan hasil sintesis terhadap 25 artikel rujukan, perkembangan pendekatan komputasi dalam sentiment analysis produk pada media sosial dapat dipahami sebagai proses

evolutif yang bergerak dari pendekatan sederhana berbasis aturan menuju pendekatan berbasis kecerdasan buatan yang lebih kontekstual dan adaptif. Secara konseptual, perkembangan tersebut dapat dikelompokkan ke dalam lima tahapan utama, yaitu *lexicon-based approach*, *machine learning*, *deep learning*, *transformer-based model*, dan *Large Language Models* (LLMs).

Tahap pertama adalah *lexicon-based approach*. Pendekatan ini menggunakan kamus sentimen untuk menentukan polaritas kata atau kalimat berdasarkan skor positif, negatif, atau netral. Keunggulan utama pendekatan ini terletak pada kemudahan implementasi dan interpretabilitas hasil. Namun, pendekatan ini memiliki keterbatasan dalam memahami konteks, sarkasme, ironi, bahasa informal, serta kata-kata baru yang sering muncul dalam media sosial. Oleh karena itu, meskipun *lexicon-based approach* masih relevan untuk analisis awal, pendekatan ini kurang memadai untuk menangani kompleksitas data media sosial yang dinamis (Li, 2022); (Susanto et al., 2025).

Tahap kedua adalah *machine learning*. Pendekatan ini menggunakan algoritma seperti *Naïve Bayes*, *Support Vector Machine*, *Logistic Regression*, *Random Forest*, dan *Decision Tree* untuk mempelajari pola sentimen dari data berlabel. *Machine learning* mampu meningkatkan akurasi dibandingkan pendekatan berbasis kamus karena model dapat belajar dari karakteristik data. Beberapa penelitian menunjukkan bahwa algoritma seperti *Random Forest* dan *SVM* memiliki performa yang cukup baik dalam mengklasifikasikan sentimen produk pada media sosial maupun e-commerce (Rahardjo et al., 2024); (Ramadhan et al., 2023); (Hakim & Azizli, 2024). Namun, pendekatan ini masih sangat bergantung pada kualitas data, proses *preprocessing*, dan pemilihan fitur seperti *TF-IDF*, *Bag of Words*, maupun *word embedding*.

Tahap ketiga adalah *deep learning*. Pendekatan ini muncul untuk mengatasi keterbatasan *machine learning* dalam melakukan *feature engineering* secara manual. Model seperti *CNN*, *RNN*, *LSTM*, *Bi-LSTM*, dan *hybrid CNN-LSTM* mampu mengekstraksi fitur teks secara otomatis serta memahami hubungan antar kata secara lebih baik. *CNN* unggul dalam menangkap pola lokal pada teks, sedangkan *LSTM* dan *Bi-LSTM* lebih efektif dalam menangkap konteks jangka panjang. Model hibrida seperti *CNN-LSTM* dan *CNN-BiLSTM* dengan *attention mechanism* terbukti mampu meningkatkan performa *sentiment analysis* karena menggabungkan kekuatan ekstraksi fitur lokal dan pemahaman konteks sekuensial (Kaur & Sharma, 2023); (Rajesh & Hiwarkar, 2023); (Septiani et al., 2024); (Jain et al., 2025).

Tahap keempat adalah *transformer-based model*. Model seperti *BERT*, *RoBERTa*, *DistilBERT*, dan *T5* membawa perubahan penting dalam *sentiment analysis* karena mampu memahami konteks secara *bidirectional* melalui mekanisme *self-attention*. Berbeda dengan

model deep learning konvensional, transformer mampu menangkap makna kata berdasarkan konteks kalimat secara lebih luas. Beberapa penelitian menunjukkan bahwa BERT dan variannya mampu menghasilkan performa yang lebih tinggi dibandingkan model machine learning dan deep learning tradisional, terutama pada data yang mengandung konteks kompleks, bahasa informal, dan ekspresi sentimen yang tidak eksplisit (Cao et al., 2022); (Das et al., 2025); (Daraghmi & Zyadeh, 2026); (Vadla et al., 2024).

Tahap kelima adalah Large Language Models. LLM seperti GPT-3 dan GPT-4 menunjukkan kemampuan yang lebih maju dalam memahami teks yang kaya konteks, mendeteksi sentimen campuran, serta memberikan penjelasan terhadap hasil klasifikasi. (Ghatora et al., 2024) menunjukkan bahwa GPT-4 lebih unggul dalam menganalisis teks panjang dan kompleks dibandingkan model tradisional, sedangkan (Elmitwalli & Mehegan, 2024) menunjukkan bahwa GPT-3 memiliki kemampuan generalisasi yang baik pada data domain spesifik dengan jumlah anotasi terbatas. Dengan demikian, LLM dapat dipandang sebagai tahap terbaru dalam perkembangan sentiment analysis karena tidak hanya melakukan klasifikasi, tetapi juga mampu memberikan interpretasi yang lebih mendalam terhadap opini pengguna.

Berdasarkan uraian tersebut, framework konseptual perkembangan pendekatan komputasi dalam sentiment analysis produk pada media sosial dapat digambarkan melalui tahap Preprocessing Data, Representasi Teks, Pendekatan Komputasi, Klasifikasi Sentimen, Insight Bisnis dan Pengambilan Keputusan.

Pada tahap input, data bersumber dari media sosial, ulasan e-commerce, komentar pengguna, maupun review aplikasi. Data tersebut kemudian melalui proses preprocessing, seperti text cleaning, tokenization, normalization, stop-word removal, stemming, dan balancing dataset. Selanjutnya, teks direpresentasikan menggunakan pendekatan seperti Bag of Words, TF-IDF, Word2Vec, FastText, GloVe, BERT embedding, atau contextual embedding. Setelah itu, pendekatan komputasi dipilih sesuai karakteristik data, mulai dari lexicon-based, machine learning, deep learning, transformer, hingga LLM. Hasil akhirnya berupa klasifikasi sentimen positif, negatif, atau netral yang dapat digunakan untuk memahami persepsi konsumen, mengevaluasi produk, memantau reputasi merek, memprediksi keberhasilan peluncuran produk, dan mendukung strategi pemasaran berbasis data (Bellar et al., 2024); (Duru & Sunar, 2025); (Gooljar et al., 2024); (Hakim & Azizli, 2024).

5. KESIMPULAN DAN SARAN

Penelitian ini menyimpulkan bahwa pendekatan komputasi dalam *sentiment analysis* produk pada media sosial terus berkembang dari *lexicon-based approach*, *machine learning*, *deep learning*, *transformer*, hingga *Large Language Models* (LLMs). Setiap pendekatan memiliki keunggulan dan keterbatasan masing-masing. *Machine learning* masih banyak digunakan karena relatif sederhana dan efisien, tetapi bergantung pada proses *feature engineering*. *Deep learning* mampu menangkap pola dan konteks teks secara lebih baik, sedangkan model *transformer* seperti BERT dan RoBERTa menunjukkan performa yang lebih unggul dalam memahami konteks bahasa yang kompleks. Sementara itu, LLM seperti GPT memiliki potensi besar dalam memahami sentimen yang lebih halus dan memberikan interpretasi yang lebih mendalam. Secara umum, hasil kajian menunjukkan bahwa *sentiment analysis* tidak hanya berguna untuk mengklasifikasikan opini positif, negatif, dan netral, tetapi juga dapat mendukung pengambilan keputusan bisnis, pengembangan produk, strategi pemasaran, serta pemantauan reputasi merek.

Penelitian selanjutnya disarankan untuk mengembangkan *aspect-based sentiment analysis* agar mampu menganalisis sentimen pada aspek produk tertentu, seperti harga, kualitas, fitur, desain, dan layanan. Selain itu, penelitian mendatang perlu menggunakan lebih banyak data media sosial berbahasa Indonesia dari berbagai platform, seperti X/Twitter, Instagram, TikTok, YouTube, Shopee, dan Tokopedia. Pengembangan model hybrid, penerapan LLM, *multimodal sentiment analysis*, serta integrasi *explainable artificial intelligence* juga penting dilakukan agar hasil analisis sentimen menjadi lebih akurat, kontekstual, transparan, dan dapat dimanfaatkan secara nyata dalam pengambilan keputusan bisnis berbasis data.

DAFTAR REFERENSI

- Bellar, O., Baina, A., & Ballafkih, M. (2024). Sentiment Analysis: Predicting Product Reviews for E-Commerce Recommendations Using Deep Learning and Transformers. *Mathematics*, 12(15), 1–21. <https://doi.org/10.3390/math12152403>
- Cao, Y., Sun, Z., Li, L., & Mo, W. (2022). A Study of Sentiment Analysis Algorithms for Agricultural Product Reviews Based on Improved BERT Model. *Symmetry*, 14(8), 1–17. <https://doi.org/10.3390/sym14081604>
- Chen, X., Xie, H., Tao, X., Wang, F. L., Zhang, D., & Dai, H. N. (2025). A Computational Analysis of Aspect-Based Sentiment Analysis Research through Bibliometric Mapping and Topic Modeling. *Journal of Big Data*, 12(1), 1–34. <https://doi.org/10.1186/s40537-025-01068-y>
- Daraghmi, E., & Zyadeh, N. (2026). Sentiment Classification of Amazon Product Reviews

- Based on Machine and Deep Learning Techniques: A Comparative Study. *Future Internet*, 18(3), 1–35. <https://doi.org/10.3390/fi18030138>
- Das, U. K., Ani, R. S., Datta, N., Fahad, I., Sikder, J., Sara, U., & Chakraborty, A. (2025). Enhancing Sentiment Analysis Accuracy on Social Media Comments Using a Tuned BERT Model. *Discover Computing*, 28(1), 1–24. <https://doi.org/10.1007/s10791-025-09599-x>
- Duru, I., & Sunar, A. S. (2025). Transformer and Pre-Transformer Model-Based Sentiment Prediction with Various Embeddings: A Case Study on Amazon Reviews. *Entropy*, 27(12), 1–37. <https://doi.org/10.3390/e27121202>
- Elmitwalli, S., & Mehegan, J. (2024). Sentiment Analysis of COP9-related Tweets: a Comparative Study of pre-trained Models and Traditional Techniques. *Frontiers in Big Data*, 7(March), 1–18. <https://doi.org/10.3389/fdata.2024.1357926>
- Fang, X., & Zhan, J. (2015). Sentiment Analysis Using Product Review Data. *Journal of Big Data*, 2(1), 1–14. <https://doi.org/10.1186/s40537-015-0015-2>
- Ghatora, P. S., Hosseini, S. E., Pervez, S., Iqbal, M. J., & Shaukat, N. (2024). Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM. *Big Data and Cognitive Computing*, 8(12), 1–18. <https://doi.org/10.3390/bdcc8120199>
- Gooljar, V., Issa, T., Hardin-Ramanan, S., & Abu-Salih, B. (2024). Sentiment-based Predictive Models for Online Purchases in the Era of Marketing 5.0: a Systematic Review. *Journal of Big Data*, 11(1), 1–39. <https://doi.org/10.1186/s40537-024-00947-0>
- Hakim, A. L., & Azizli, A. (2024). Social Media Sentiment Analysis as a Predictor of Product Launch Success in the Digital Marketplace. *Management Dynamics: International Journal of Management and Digital Sciences*, 1(1), 28–41. <https://doi.org/10.70062/managementdynamics.v1i1.419>
- Jaber, A., Bahati, I., & Martínez, P. (2025). Leveraging pre-trained Embeddings in an Ensemble Machine Learning Approach for Arabic Sentiment Analysis. *Frontiers in Artificial Intelligence*, 8(September), 1–12. <https://doi.org/10.3389/frai.2025.1653728>
- Jain, V., Malviya, L., & .S, A. (2025). Optimized Hybrid Deep Learning for Cross-linguistic Sentiment Analysis: a Novel Approach. *Journal of Cloud Computing*, 14(1), 1–21. <https://doi.org/10.1186/s13677-025-00753-w>
- Kaur, G., & Sharma, A. (2023). A Deep Learning-based Model Using Hybrid Feature Extraction Approach for Consumer Sentiment Analysis. *Journal of Big Data*, 10(1), 1–23. <https://doi.org/10.1186/s40537-022-00680-6>
- Li, H. (2022). Brand Potential User Identification Algorithm Based on Sentiment Analysis. *Frontiers in Psychology*, 13(May), 1–12. <https://doi.org/10.3389/fpsyg.2022.906928>
- Liu, Y., Lu, J., Yang, J., & Mao, F. (2020). Sentiment Analysis for e-Commerce Product Reviews by Deep Learning Model of Bert-BiGRU-Softmax. *Mathematical Biosciences and Engineering*, 17(6), 7819–7837. <https://doi.org/10.3934/MBE.2020398>
- Moleong, L. J. (2017). *Metodologi Penelitian Kualitatif*. Remaja Rosdakarya.
- Rahardjo, D. A. S., Andriana, M., & Prasastono, S. H. (2024). Consumer Behavior and

- Purchase Intentions: A Machine Learning-Based Sentiment Analysis of Social Media Data. *Journal of Management and Informatics*, 3(1), 157–173. <https://doi.org/10.51903/jmi.v3i1.45>
- Rajesh, A., & Hiwarkar, T. (2023). Sentiment Analysis from Textual Data Using Multiple Channels Deep Learning Models. *Journal of Electrical Systems and Information Technology*, 10(1), 1–12. <https://doi.org/10.1186/s43067-023-00125-x>
- Ramadhan, F. A., Ruslan, R. R. P., & Zahra, A. (2023). Sentiment Analysis of E-Commerce Product Reviews For Content Interaction Using Machine Learning. *CAKRAWALA : Repositori IMWI*, 53(1), 207–220. <http://publications.lib.chalmers.se/records/fulltext/245180/245180.pdf%0Ahttps://hdl.handle.net/20.500.12380/245180%0Ahttp://dx.doi.org/10.1016/j.jsames.2011.03.003%0Ahttps://doi.org/10.1016/j.gr.2017.08.001%0Ahttp://dx.doi.org/10.1016/j.precamres.2014.12>
- Sasikala, P., & Sheela, L. M. I. (2020). Sentiment Analysis of Online Product Reviews Using DLMNN and Future Prediction of Online Product Using IANFIS. *Journal of Big Data*, 7(1), 1–20. <https://doi.org/10.1186/s40537-020-00308-7>
- Septiani, S., Putri, N., Jessica, D., & Saputra, A. (2024). Sentiment Analysis Of Social Media Data Using Deep Learning Techniques. *International Journal of Computer Technology and Science*, 1(2), 15–28. <https://doi.org/10.3390/life14060708>
- Sugiyono. (2019). *Metodelogi Penelitian Kuantitatif dan Kualitatif Dan R&D*. Alfabeta.
- Susanto, H., Omar, A. S., Shafa Susanto, A. K., Setiana, D., Fang-Yie, L., Shaikh, J. M., Insani, A., Khusni, U., Hidayat, R., Akbari, I., & Basuki, I. (2025). Toward Robust Social Media Sentiment for SMEs: a Comparative Study of Dictionary-based and Machine Learning Approaches with Insights for Hybrid Methodologies. *Frontiers in Big Data*, 8(April), 1–21. <https://doi.org/10.3389/fdata.2025.1594374>
- Tabssam, H. A., Chattha, S. A., Javeed, M. F., & Hayat, A. (2025). Sentiment Analysis of User Reviews: Exploring Duolingo and Babbel in English Language Learning. *Frontiers in Computer Science*, 7(June), 1–17. <https://doi.org/10.3389/fcomp.2025.1569058>
- Vadla, M. K. S., Suresh, M. A., & Viswanathan, V. K. (2024). Enhancing Product Design through AI-Driven Sentiment Analysis of Amazon Reviews Using BERT. *Algorithms*, 17(2), 1–17. <https://doi.org/10.3390/a17020059>