

1. Pendahuluan

Latar Belakang

Kemajuan teknologi informasi dan komunikasi telah membawa perubahan besar dalam cara manusia berinteraksi serta mengakses informasi. Berbagai platform daring memungkinkan pengguna untuk mengekspresikan pendapat, berbagi pengalaman, dan memberikan ulasan mengenai berbagai topik, salah satunya seperti film. Seiring perkembangan tersebut, jumlah data teks yang dihasilkan oleh pengguna internet kian meningkat pesat. Di bidang perfilman, ulasan dan pendapat penonton berperan penting dalam mengukur respons masyarakat terhadap sebuah film, yang dapat membantu rumah produksi, kritikus, maupun calon penonton untuk memahami persepsi publik secara lebih objektif.

Analisis sentimen merupakan salah satu metode yang kerap digunakan untuk mengurai informasi sentimen pada sebuah teks. Dengan analisis sentimen, teks ulasan dapat diklasifikasikan ke dalam kategori emosi atau polaritas tertentu, seperti positif, negatif, atau netral [1] dengan memanfaatkan *machine learning* (ML). Namun, performa model ML sangat bergantung pada kualitas representasi teks yang digunakan. Tantangan utamanya adalah bagaimana mengubah teks menjadi representasi numerik yang dapat menangkap makna secara efektif.

Meskipun kinerja model *deep learning* (DL) yang dapat mencapai hasil *state-of-the-art* dalam penelitian analisis sentimen [2], model ML masih relevan, terutama dalam skenario ketika sumber daya komputasi dan dataset yang terbatas [3, 4]. Selain itu, kinerja model ML dapat ditingkatkan secara signifikan dengan penggunaan teknik representasi teks yang efektif [5].

Salah satu teknik representasi teks adalah dengan *embedding* statis, seperti Word2Vec, di mana setiap kata dipetakan ke satu vektor numerik, tetapi tanpa memperhitungkan konteks kalimatnya. Keterbatasan ini mendorong pengembangan *embedding* kontekstual, salah satunya adalah Text-embedding-3-large dari OpenAI, yang dirilis pada awal tahun 2024 [6]. Keunggulan utamanya terletak pada kemampuannya menangkap nuansa semantik yang kompleks dan hubungan kontekstual dalam teks, yang sangat penting untuk tugas seperti analisis sentimen [7], dan telah terbukti unggul dalam berbagai tugas *Natural Language Processing* (NLP) [8].

Kecanggihan *embedding* Text-embedding-3-large ini mendukung dua pendekatan klasifikasi yang akan dieksplorasi dalam penelitian ini. Pertama, pendekatan *supervised learning*, di mana model ML tradisional seperti *Support Vector Machine* (SVM) dan *Logistic Regression* (LR) dilatih menggunakan data berlabel. Model-model ini dipilih karena kesederhanaan dan efisiensi komputasinya [9]. Kedua, pendekatan klasifikasi *zero-shot*. Pendekatan ini memungkinkan klasifikasi sentimen pada data baru tanpa memerlukan data latih berlabel spesifik untuk tugas tersebut; yang dimungkinkan berkat pemahaman semantik yang mendalam dari *embedding* OpenAI.

Penelitian ini akan mengeksplorasi penggunaan model Text-embedding-3-large untuk klasifikasi sentimen, baik secara *supervised learning* dan juga melalui klasifikasi *zero-shot* pada dataset ulasan film *Movie Reviews* (MR) [10]. Meskipun sudah ada sejumlah penelitian yang menguji berbagai model dan *embedding* teks pada dataset MR [11, 12, 13, 14, 15], masih terdapat kesenjangan penelitian terkait bagaimana performa model *embedding* teks kontekstual dari OpenAI pada dataset tersebut ketika diintegrasikan dengan beragam pendekatan klasifikasi.

Topik dan Batasannya

Berdasarkan latar belakang yang telah diuraikan, kami merumuskan beberapa masalah utama yang akan dikaji. Pertama, penelitian ini akan mengevaluasi bagaimana performa model *embedding* kontekstual Text-embedding-3-large untuk klasifikasi sentimen pada dataset MR secara *supervised learning* menggunakan model ML. Selanjutnya, akan diukur sejauh mana peningkatan performa model Text-embedding-3-large jika dibandingkan dengan *embedding* statis Word2Vec. Penelitian ini juga akan mengidentifikasi model ML mana yang paling sesuai untuk mengklasifikasikan sentimen pada dataset MR menggunakan model Text-embedding-3-large. Terakhir, penelitian ini akan menguji bagaimana performa model Text-embedding-3-large dengan pendekatan klasifikasi *zero-shot* pada dataset MR.

Untuk memastikan penelitian dapat diselesaikan dalam lingkup waktu yang tersedia, serta dengan mempertimbangkan sumber daya yang ada, maka ditetapkan beberapa batasan masalah. Pertama, penelitian ini hanya menggunakan dataset MR yang berbahasa Inggris sebagai objek studi. Pengujian tidak dilakukan pada dataset lain yang memiliki format berbeda (misalnya, ulasan berbahasa Indonesia atau dari domain selain film).

Kedua, sebagai pembandingan untuk model Text-embedding-3-large yang berbasis kontekstual, penelitian ini menggunakan metode *embedding* statis, yaitu Word2Vec. Model *embedding* kontekstual lain seperti BERT (*Bidirectional Encoder Representations from Transformers*) tidak diikutsertakan dalam perbandingan. Batasan ini dikarenakan adanya keterbatasan sumber daya komputasi perangkat keras untuk menjalankan model tersebut secara efisien.

Dan yang terakhir, pendekatan *supervised learning* dalam penelitian ini terbatas pada penggunaan algoritma ML tradisional, yaitu LR dan SVM. Penelitian ini tidak melibatkan implementasi dan perbandingan dengan arsitektur model DL.

Tujuan

Penelitian ini bertujuan untuk memberikan kontribusi berupa informasi mengenai bagaimana model *embedding* Text-embedding-3-large dari OpenAI dapat diimplementasikan secara optimal pada kedua pendekatan klasifikasi untuk klasifikasi sentimen, dan seberapa besar peningkatan kinerja yang dapat dicapai. Diharapkan, penerapan model Text-embedding-3-large akan memberikan representasi semantik yang mampu menghasilkan akurasi klasifikasi sentimen yang tinggi.

Organisasi Tulisan

Setelah bagian Pendahuluan ini, jurnal dilanjutkan dengan Bab 2 yang membahas landasan teori dan tinjauan pustaka mengenai penelitian-penelitian sebelumnya yang berkaitan dengan analisis sentimen, *embedding* teks, dan model *embedding* OpenAI. Selanjutnya, Bab 3 akan menjelaskan secara rinci metodologi yang diimplementasikan, mulai dari dataset, prapemrosesan, *generate* vektor *embedding*, hingga implementasi model klasifikasi *supervised learning* dan *zero-shot*. Bab 4 akan memaparkan hasil eksperimen yang disertai dengan analisis terhadap performa setiap konfigurasi model yang diuji. Terakhir, jurnal ini ditutup dengan kesimpulan di Bab 5 yang merangkum temuan-temuan utama, menjawab tujuan penelitian, dan memberikan saran untuk pengembangan di masa mendatang.