

Analisis Sentimen Multi-Aspek Ulasan Film Menggunakan BiLSTM pada Data Platform X

Astria MP Sinaga ^{1,*}, Yuliant Sibaroni ², Sri Suryani Prasetyowati ³

^{1,2,3} School of Computing, Informatika, Universitas Telkom, Bandung, Indonesia

Email: ^{1,*} triameilynnp@student.telkomuniversity.ac.id, ² yuliant@telkomuniversity.ac.id,

³ srisuryani@telkomuniversity.ac.id

Penulis Korespondensi Email: triameilynnp@student.telkomuniversity.ac.id

Abstrak

Industri film menghasilkan banyak ulasan film setiap tahunnya, yang mencerminkan opini penonton terhadap berbagai aspek film seperti cerita, musik, penampilan, dan sebagainya. Ulasan-ulasan tersebut merupakan sumber yang baik untuk menganalisis opini publik secara otomatis. Analisis sentimen dan aspek ulasan film berdasarkan model klasifikasi multitugas yang berakar pada struktur Bidirectional Long Short-Term Memory (BiLSTM) menjadi tema penelitian ini. Tujuan penelitian ini adalah untuk mengembangkan dan mengevaluasi model multitugas berbasis BiLSTM yang mampu mengklasifikasikan polaritas sentimen dan aspek ulasan film secara bersamaan untuk meningkatkan penggalan opini yang terperinci. Data dikumpulkan dari Platform X melalui perayapan web dan menjalani berbagai langkah praproses teks sebelum memasukkannya ke dalam model. Tidak seperti pendekatan tradisional yang memperlakukan sentimen dan klasifikasi aspek sebagai operasi independen, metode yang diusulkan dalam karya ini melakukan keduanya secara bersamaan—prediksi sentimen (positif, netral, negatif) dan kategori aspek (alur cerita, musik, aktor, lainnya). Model tersebut dibandingkan antara tiga ukuran lapisan BiLSTM yang berbeda—32, 64, dan 128 unit—untuk menyelidiki pengaruh kapasitas model terhadap kinerja. Skema validasi silang 10 kali lipat juga diterapkan untuk mengonfirmasi keandalan dan ketahanan hasil. Temuan eksperimen mengungkapkan bahwa model BiLSTM 128 unit mengungguli model lain secara keseluruhan, khususnya dalam mengambil hubungan kontekstual yang halus, untuk mencapai skor akurasi tertinggi dalam kedua tugas. Meskipun model ini secara signifikan lebih lama untuk dilatih, generalisasi yang ditingkatkan—terutama untuk pasangan aspek sentimen yang sulit seperti kategori netral atau sumber daya rendah—memvalidasi trade-off. Temuan tersebut memvalidasi efektivitas pembelajaran multitugas berbasis BiLSTM untuk analisis ulasan film yang komprehensif, yang menunjukkan pentingnya kapasitas model dalam menangani pola bahasa yang kompleks dan identifikasi opini yang terperinci.

Kata kunci: Analisis Sentimen, Multi Aspek, BiLSTM, Multitugas

Abstract

The film industry generates scores of movie reviews annually, reflecting viewer opinion towards various aspects of movies such as story, music, performances, and so on. They are a good source to publicly analyze opinion automatically. Aspect-based and sentiment analysis of movie reviews based on a multitask classification model rooted in the Bidirectional Long Short-Term Memory (BiLSTM) structure is the theme of this study. Data was collected from Platform X through web crawling and subjected to various text preprocessing steps before feeding them into the model. Unlike traditional approaches that treat sentiment and aspect classification as independent operations, the method proposed in this work is performing both simultaneously—sentiment prediction (positive, neutral, negative) and aspect categories (plot, music, actors, others). The model was compared between three different sizes of BiLSTM layers—32, 64, and 128 units—to investigate the influence

of model capacity on performance. A 10-fold cross-validation scheme also implemented to confirm the reliability and robustness of results. Experiment findings reveal that the 128-unit BiLSTM model outperformed other models across the board, particularly at picking up subtle contextual relationships, to achieve the highest accuracy score in both tasks. Although this model significantly longer to train, its improved generalization—most notably for difficult sentiment- aspect pairs such as neutral or low-resource categories—validated the trade-off. The findings validate the effectiveness of BiLSTM-based multitask learning for comprehensive movie review analysis, demonstrating the importance of model capacity in tackling complex language patterns and fine-grained opinion identification.

Keywords: Sentiment Analysis, Multi-Aspect, BiLSTM, Multitask.

Pendahuluan

Pembuatan film kini telah menjadi bagian dari kehidupan modern dan memainkan peran penting dalam pengalaman pribadi maupun lingkungan sosial. Industri film memproduksi jutaan film setiap tahunnya, yang sebagian besar didorong oleh greed for money atau keserakahan akan keuntungan finansial. Sayangnya, kualitas film sering kali tidak sebanding dengan gambar-gembornya. Cuplikan film (film trailers) terkadang menyedapkan penonton, sehingga penonton merasa kecewa dan tidak puas saat isi film tidak sesuai dengan ekspektasi mereka[1].

Dengan adanya akses internet di era digital, kini lebih mudah dari sebelumnya bagi penonton untuk memperoleh dan bertukar informasi. Sebagian besar individu mengecek ulasan secara daring sebelum menonton film, guna mengetahui persepsi publik. Situs media sosial, khususnya Platform X, telah menjadi platform utama di mana pengguna memberikan opini dalam bentuk ulasan, komentar, dan tweet[2]. Teks-teks yang dihasilkan pengguna ini memberikan wawasan berharga tentang sentimen umum masyarakat dan dapat dimanfaatkan melalui metode sentiment analysis[3].

Sentiment analysis adalah subdisiplin dari *natural language processing (NLP)* yang berfokus pada identifikasi emosi dalam teks, baik positif, negatif, maupun netral. *Sentiment analysis* menjadi alat penting untuk merekam sentimen publik terhadap suatu produk, layanan, atau peristiwa. Namun, mengamati sentimen keseluruhan dari sebuah ulasan sering kali tidak cukup informatif, terutama jika ulasan tersebut mengandung opini campuran terhadap berbagai aspek. Misalnya, sebuah ulasan film bisa saja menyatakan suka terhadap alur cerita tetapi tidak suka terhadap musik atau akting, seperti pada pernyataan: "I liked the plot but hated the music"[4], [5]. Contoh seperti ini menyoroti relevansi dari *aspect-based sentiment analysis (ABSA)*, yang bertujuan untuk mengklasifikasikan sentimen berdasarkan fitur atau elemen individual.

Upaya awal dalam *sentiment analysis* menggunakan pendekatan *machine learning* tradisional seperti *Naïve Bayes*. Salah satu eksperimen yang menggunakan *Naïve Bayes* pada dataset *tweet* hanya mencapai akurasi 65%, yang menunjukkan kelemahannya dalam memproses teks yang kompleks dan kontekstual[6]. Eksperimen lain mencapai akurasi lebih tinggi sebesar 78,96%, tetapi tetap mengalami kesulitan dalam mengenali ekspresi sentimen yang halus, yang menunjukkan bahwa pendekatan ini mungkin kurang cocok untuk dataset dengan kompleksitas tinggi[7]. *Random Forest*, sebagai klasifikator populer lainnya, mencatat akurasi 86%, namun diketahui kurang mampu melakukan generalisasi terhadap data baru[8]. Sementara itu, *K-Nearest Neighbor (KNN)*—meskipun sederhana dan mencatat akurasi 87%—rentan terhadap *overfitting* jika parameter K terlalu kecil dan sensitif terhadap pemilihan nilai K[9].

Deep learning telah merevolusi kinerja dalam *sentiment analysis*. *Long Short-Term Memory (LSTM)*, sebagai pengembangan dari *Recurrent Neural Networks (RNNs)*, mampu menangkap ketergantungan jangka pendek dan panjang dalam teks. Misalnya, *LSTM* berhasil mencapai akurasi sebesar 88,46% dalam klasifikasi ulasan film, dan 89,9% pada dataset IMDb yang disusun oleh Andrew Maas[10], [11]. Meskipun model *LSTM* efektif, model ini membutuhkan sumber daya komputasi yang besar dan rentan terhadap

overfitting jika dilatih pada data yang kecil atau bising. Di samping itu, pendekatan ini belum selalu memberikan peningkatan akurasi yang signifikan dibandingkan metode lain. Selain itu, keterbatasannya pada kemampuan multibahasa dan ketergantungannya pada dataset berbahasa Inggris membatasi tingkat generalisasinya. Penelitian selanjutnya mengusulkan penggabungan metode seleksi fitur untuk meningkatkan performa model secara keseluruhan[12].

Studi komparatif juga menunjukkan bahwa model seperti *Bag of Words* yang dipadukan dengan *Random Forest*, *Support Vector Machines (SVM)*, atau *Logistic Regression* dapat mencapai akurasi hingga 75,59%[13]. Namun, pendekatan ini umumnya masih kurang dalam menangkap konteks kaya yang diperlukan dalam tugas *sentiment analysis* tingkat lanjut, yang kembali menekankan pentingnya pemilihan model dan penyetelan *hyperparameter*.

Meskipun banyak studi telah mengeksplorasi berbagai metode *sentiment analysis*, sebagian besar hanya berfokus pada sentimen keseluruhan atau pada klasifikasi sentimen berbasis aspek, namun belum menggabungkan keduanya dalam satu kerangka kerja yang terpadu. Selain itu, banyak penelitian yang hanya menggunakan dataset berbahasa Inggris dan belum cukup adaptif terhadap ulasan pengguna yang lebih kompleks dan kontekstual di dunia nyata. Keterbatasan ini menunjukkan perlunya pengembangan model yang mampu menangkap baik polaritas sentimen maupun aspek-aspek spesifik dalam ulasan secara bersamaan, serta dalam konteks multibahasa. Untuk menjawab kebutuhan ini, kami mengusulkan implementasi sistem *multi-aspect sentiment analysis* berbasis *BiLSTM* terhadap ulasan film yang dikumpulkan dari *Platform X*.

Untuk mengatasi berbagai keterbatasan tersebut, *Bidirectional Long Short-Term Memory (BiLSTM)* telah diusulkan sebagai solusi yang lebih efisien. Dengan memproses data secara maju dan mundur, *BiLSTM* mempertahankan konteks yang lebih kaya sehingga menghasilkan akurasi klasifikasi sentimen yang lebih baik. Dalam sebuah studi yang menggunakan *BiLSTM* untuk analisis ulasan produk di media sosial, diperoleh akurasi luar biasa sebesar 96,58%, yang menunjukkan potensi besar metode ini dalam *multi-aspect sentiment analysis*[4]. Kemampuan *BiLSTM* untuk mengakses konteks teks sebelum dan sesudah suatu kata memungkinkan pemahaman yang lebih menyeluruh terhadap emosi pengguna.

Lebih lanjut, upaya telah dilakukan untuk menggabungkan *BiLSTM* dengan model bahasa canggih seperti *BERT (Bidirectional Encoder Representations from Transformers)*. Ma et al. (2021) merancang *Multi-task Learning Double BiLSTM Model (MLDBM)* yang mengombinasikan *BiLSTM* dan *BERT* untuk melakukan *aspect-based sentiment analysis* pada ulasan klinis, dengan akurasi mencapai 75,59%[14]. Pendekatan hibrida ini memiliki potensi besar untuk klasifikasi sentimen di bidang khusus. Studi lain yang menggunakan kombinasi *BiLSTM* dan *BERT* bahkan mencapai akurasi lebih tinggi, yakni hingga 81%, terutama dalam menangani kompleksitas temporal dan kontekstual dalam data teks[15].

Model yang diusulkan dirancang menggunakan kerangka multi-task learning untuk mengklasifikasikan baik polaritas sentimen (positif, negatif, netral) maupun aspek-aspek (alur cerita, musik, akting, dan lainnya). Dengan memanfaatkan kekuatan *BiLSTM* dalam menangkap konteks dua arah, penelitian ini bertujuan memberikan wawasan yang lebih mendalam tentang sentimen pengguna dan berkontribusi pada pemahaman yang lebih rinci dalam analisis ulasan film.