Jumistik
E-ISSN: 2964-3953 P-ISSN: 2964-3988
Jurnal Manajemen Informatika, Sistem Informasi dan Teknologi Komputer

Vol.4 No.2 Desember 2025
Hal 587-598
Https://ojs.amiklps.ac.id

IMPLEMENTASI BERT UNTUK ANALISIS *EXCITEMENT* AUDIENS FILM MENDATANG

Royfansyah Muhammad Razavi[1*], Petrus Sokibi[2], Putri Rizqiyah[3]
[1,2,3]Teknik Informatika Fakultas Teknologi Informasi, Universitas Catur Insan Cendekia
E-mail: royfansyahae@gmail.com[1*], petrus.sokibi@cic.ac.id[2], putri.rizqiyah@cic.ac.id[3]

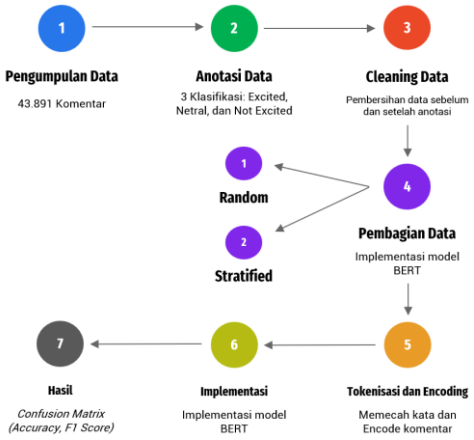
INFO ARTIKEL

Sejarah Artikel
Diterima : 31/08/2025
Direvisi : 07/10/2025
Diterbitkan : 01/12/2025

***Corresponding author**
royfansyahae@gmail.com

DOI: 10.70247/jumistik.v4i2.198
Https://ojs.amiklps.ac.id

GRAPHICAL ABSTRACT



ABSTRACT

Social media comments about upcoming films often contain diverse and informal expressions of audience excitement, making manual classification challenging. This study implements a BERT-based classification model to categorize English comments into three classes: Excited, Neutral, and Not Excited. Data were collected from social media, manually annotated, and processed through cleaning, tokenization, and encoding. The dataset was divided using random and stratified splitting before training and evaluation. Experimental results show that BERT achieved strong performance, with stratified splitting providing higher stability and accuracy compared to random splitting. The model was further integrated into a simple web application to demonstrate practical implementation. These findings highlight the effectiveness of BERT for analyzing audience excitement in social media comments.

Keywords: *BERT, Sentiment Analysis, Social Media, Audience Excitement, NLP*

ABSTRAK

Komentar media sosial tentang film yang akan datang sering kali mengekspresikan tingkat antusiasme penonton secara tidak terstruktur, sehingga sulit diklasifikasikan secara manual. Penelitian ini mengimplementasikan model klasifikasi berbasis BERT untuk mengelompokkan komentar berbahasa Inggris ke dalam tiga kategori: *Excited*, Netral, dan *Not Excited*. Data dikumpulkan dari media sosial, dianotasi secara manual, lalu melalui tahap pembersihan, *tokenisasi*, dan *encoding*. Selanjutnya, data dibagi menggunakan *random split* dan *stratified split* sebelum proses pelatihan dan evaluasi model. Hasil eksperimen menunjukkan bahwa BERT mampu memberikan performa klasifikasi yang baik, dengan *stratified split* menghasilkan akurasi yang lebih tinggi dan stabil dibandingkan *random split*. Model juga diintegrasikan ke dalam aplikasi web sederhana sebagai implementasi praktis. Hasil penelitian ini membuktikan efektivitas BERT dalam menganalisis *excitement* audiens terhadap film mendatang melalui komentar media sosial.

Kata kunci: BERT, Analisis Sentimen, Komentar Media Sosial, *Excitement* Audiens, NLP

PENDAHULUAN

Era digitalisasi telah menghadirkan transformasi signifikan dalam ekosistem industri perfilman global. Kemajuan teknologi digital yang pesat, disertai dengan penetrasi media sosial yang masif, telah mengubah lanskap pemasaran dan interaksi antara produser konten dengan audiensnya. Transformasi ini tidak hanya berdampak pada cara film diproduksi dan didistribusikan, tetapi juga pada bagaimana audiens mengakses, berinteraksi, dan mengekspresikan antusiasme mereka terhadap konten sinematik.

Data terkini dari *Motion Picture Association* menunjukkan bahwa sektor home/mobile entertainment global telah mencatatkan pendapatan sebesar USD 78,5 miliar, dengan peningkatan 14% dibandingkan periode sebelumnya. Lonjakan ini terutama dipicu oleh pertumbuhan *eksponensial* langganan *platform streaming*, yang mencapai 1,3 miliar pengguna dengan pertumbuhan 14% dari tahun 2020 [1]. Fenomena ini menandakan pergeseran paradigma konsumsi konten film dari model tradisional menuju ekosistem digital yang lebih dinamis dan interaktif.

Kawasan Asia Pasifik telah muncul sebagai kekuatan baru dalam industri perfilman global. Pertumbuhan yang signifikan dalam jumlah bioskop dan audiens di wilayah ini, dengan China yang berhasil melampaui Amerika Serikat sebagai pasar bioskop terbesar dunia [1], mencerminkan dinamika perubahan geopolitik ekonomi kreatif global. Transformasi ini mengindikasikan adanya evolusi fundamental dalam cara audiens mengakses dan berpartisipasi dalam pengalaman sinematik.

Proyeksi industri menunjukkan optimisme yang tinggi terhadap pemulihan pasca pandemi. Data dari Statista memproyeksikan bahwa pendapatan pasar sinema internasional akan mencapai *USD* 86,10 miliar pada tahun 2025, dengan proyeksi peningkatan hingga *USD* 104,35 miliar pada tahun 2029, mencerminkan *compound annual growth rate* sebesar 4,92% [2]. Jumlah penonton bioskop global diperkirakan akan mencapai 1,9 miliar pengguna pada tahun 2029, dengan tingkat penetrasi meningkat dari 22,4% menjadi 24,1%. Rata-rata pendapatan per penonton diestimasi sebesar *USD* 49,13, dengan Amerika Serikat tetap memimpin sebagai pasar dengan pendapatan tertinggi, yakni USD 23,52 miliar pada tahun 2025 [2].

Studio-studio besar telah mengantisipasi momentum ini dengan merilis berbagai film *blockbuster* sepanjang tahun 2025. *Marvel Studios* menghadirkan trilogi film unggulan termasuk *Captain America: Brave New World*, *Thunderbolts*, dan *The Fantastic Four: First Steps*. *Warner Bros.* turut berkontribusi dengan *Final Destination: Bloodlines*, sementara *Disney* merilis film yang dinantikan seperti *Lilo & Stitch*. Kehadiran film-film dari studio major ini tidak hanya meningkatkan ekspektasi audiens, tetapi juga memperkuat posisi media sosial sebagai platform strategis dalam membangun antisipasi publik sebelum perilisan.

Platform media sosial seperti Facebook, YouTube, Twitter, TikTok, dan Reddit telah bertransformasi menjadi instrumen vital dalam membangun *excitement* dan ekspektasi audiens terhadap film yang akan datang. Keunggulan media sosial terletak pada aksesibilitas informasi yang tinggi, sifat interaktif, dan kemampuan memicu diskusi publik yang lebih luas dibandingkan metode pemasaran konvensional. Strategi pemasaran berbasis media sosial terbukti efektif dalam meningkatkan kepuasan penonton dan mempengaruhi intensi pembelian terhadap produk film [3].

Aktivitas tinggi pengguna media sosial dalam mendiskusikan film yang akan datang telah menciptakan peluang baru dalam memahami dinamika antusiasme audiens. Diskusi yang berkembang sebelum perilisan, mencakup reaksi terhadap *trailer*, poster, pengumuman *cast*, dan informasi produksi lainnya, seringkali menjadi refleksi dari ekspektasi dan *hype* yang berkembang di masyarakat. Media sosial telah menjadi ruang ekspresi publik yang memfasilitasi pengguna untuk menyampaikan opini, prediksi, dan respons emosional terhadap konten yang belum dirilis.

Komentar dan diskusi yang muncul di platform digital ini menggambarkan proses pembentukan dan penyebaran persepsi awal terhadap film. Oleh karena itu, analisis sentimen dari percakapan yang terjadi di media sosial dapat memberikan *insight* yang berharga bagi stakeholder industri film, baik internasional maupun lokal, dalam merancang strategi komunikasi, promosi, dan distribusi yang lebih efektif. *Excitement* yang tercermin dari ekspresi audiens dapat berfungsi sebagai indikator prediktif terhadap potensi daya tarik dan tingkat penerimaan film di kalangan publik.

Namun demikian, tantangan signifikan muncul dalam upaya menganalisis *excitement* audiens secara akurat dan otomatis. Komentar di media sosial memiliki karakteristik *real-time* dengan volume data yang masif serta variasi ekspresi bahasa yang kompleks, termasuk penggunaan ironi, sarkasme, dan manifestasi emosional yang tidak eksplisit [4]. Pendekatan konvensional dalam analisis sentimen, seperti metode berbasis leksikon atau sistem berbasis aturan, seringkali menghadapi limitasi dalam memahami nuansa emosional yang kompleks, khususnya dalam mengidentifikasi *excitement* audiens dengan tingkat akurasi yang memadai.

Metode tradisional menunjukkan keterbatasan dalam memahami dinamika bahasa yang digunakan audiens dalam mengekspresikan antusiasme mereka, sehingga menghasilkan analisis yang kurang akurat dan tidak kontekstual [5]. Kompleksitas bahasa natural, terutama dalam konteks media sosial yang informal dan dinamis, memerlukan pendekatan yang lebih canggih untuk dapat

menangkap makna dan konteks yang terkandung dalam ekspresi textual.

Untuk mengatasi tantangan tersebut, pendekatan berbasis *artificial intelligence* telah mulai diaplikasikan secara luas dalam berbagai domain, termasuk analisis sentimen berbasis teks. Dalam konteks ini, *Bidirectional Encoder Representations from Transformers* (BERT) telah muncul sebagai model yang menonjol karena kemampuannya dalam memahami konteks dan semantik kata dalam struktur kalimat secara komprehensif. BERT memiliki keunggulan dalam menangkap relasi semantik dan sintaksis dalam bahasa natural, yang sangat krusial untuk mengidentifikasi ekspresi emosional yang kompleks dalam konten media sosial.

Model BERT telah terbukti memberikan performa tinggi dalam berbagai tugas klasifikasi teks, termasuk analisis sentimen, dengan pencapaian skor *accuracy*, *precision*, recall, dan *F1-score* di atas 0.8 pada berbagai *dataset* validasi. Keunggulan tersebut menjadikan BERT sebagai solusi yang tepat untuk mengatasi tantangan dalam analisis *excitement* audiens terhadap film yang akan datang.

Penelitian ini diharapkan dapat memberikan kontribusi signifikan bagi industri film, khususnya bagi distributor film internasional di Indonesia, agensi digital marketing, dan pengelola konten media sosial, dalam memahami pola *excitement* audiens sebelum perilisan film. Penelitian ini bertujuan untuk mengembangkan sistem berbasis BERT yang mampu menganalisis komentar dari media sosial secara otomatis dan akurat, guna mendukung pengambilan keputusan terkait strategi promosi dan distribusi film internasional di pasar Indonesia.

Meskipun data komentar bersumber dari pengguna global, fokus utama penelitian ini adalah pada pemanfaatan insight *excitement* tersebut untuk merancang strategi pemasaran yang relevan dalam konteks lokal. Penelitian ini diposisikan sebagai solusi berbasis teknologi informasi yang mengintegrasikan artificial intelligence dan analisis media sosial untuk menjawab tantangan dalam memahami perilaku audiens di era digital.

Dengan demikian, penelitian ini bertujuan untuk mengimplementasikan *Bidirectional Encoder Representations from Transformers* (BERT) dalam menganalisis *excitement* audiens terhadap film yang akan datang, yang membedakannya dari penelitian sebelumnya karena tidak hanya menerapkan BERT dalam analisis *excitement* audiens, tetapi juga bertujuan menghasilkan model analisis sentimen yang lebih akurat dan kontekstual untuk kebutuhan industri perfilman baik internasional maupun lokal. Harapan dari penelitian ini adalah terciptanya sistem analisis yang dapat memberikan insight prediktif yang akurat mengenai potensi penerimaan film di pasar, sehingga dapat mendukung optimalisasi strategi pemasaran dan distribusi dalam industri perfilman modern.

TINJAUAN PUSTAKA

Pengembangan teknologi *Natural Language Processing* berbasis Transformer telah membuka peluang baru dalam analisis sentimen tekstual, khususnya untuk konten media sosial yang bersifat dinamis dan kompleks. Model *Bidirectional Encoder Representations from Transformers* (BERT) dan variannya telah menjadi fokus utama dalam berbagai penelitian analisis sentimen film selama lima tahun terakhir, menunjukkan superioritas dibandingkan pendekatan konvensional dalam memahami nuansa emosional dari teks.

Nkhata et al. [5] mengusulkan integrasi BERT dengan *Bidirectional Long Short-Term Memory* (BiLSTM) untuk analisis sentimen ulasan film, dengan hasil yang menunjukkan pencapaian akurasi hingga 98,76%. Penelitian mereka mengidentifikasi bahwa kombinasi arsitektur *Transformer* dengan jaringan rekuren mampu menangkap dependensi temporal dan kontekstual secara lebih efektif. Metode yang digunakan melibatkan *preprocessing* data ulasan dari berbagai platform, *tokenisasi* menggunakan BERT *tokenizer*, dan *fine-tuning* model *hybrid* BERT-BiLSTM. Hasil eksperimen menunjukkan bahwa pendekatan ini unggul dalam klasifikasi sentimen positif, negatif, dan netral. Namun, penelitian tersebut masih terbatas pada analisis ulasan film yang telah dirilis, tanpa mempertimbangkan aspek prediktif terhadap konten yang belum tayang.

Zhang et al. [6] memfokuskan penelitiannya pada kombinasi BERT dengan *Convolutional Neural Network* (CNN) untuk deteksi sentimen negatif dalam ulasan film. Pendekatan mereka melibatkan ekstraksi fitur menggunakan layer konvolusi setelah representasi BERT, dengan tujuan meningkatkan kemampuan deteksi pola lokal dalam teks. Metodologi penelitian mencakup *augmentasi* data, optimisasi *hyperparameter*, dan evaluasi menggunakan *dataset* IMDb. Hasil penelitian menunjukkan bahwa model BERT-CNN mampu mendeteksi sentimen negatif dengan presisi yang lebih tinggi dibandingkan model baseline. Meskipun demikian, fokus penelitian tetap berada pada ulasan film yang sudah tayang, bukan pada analisis ekspektasi audiens terhadap film mendatang.

Ning et al. [7] mengembangkan pendekatan yang mengintegrasikan BERT dengan CNN dan mekanisme attention untuk meningkatkan performa klasifikasi sentimen. Penelitian mereka menghasilkan peningkatan *F1-score* menjadi 0,92 dengan akurasi 91,12%. Metodologi yang diterapkan melibatkan *preprocessing* teks multi-tahap, *feature engineering* menggunakan BERT *embeddings*, dan implementasi *attention mechanism* untuk memberikan bobot yang berbeda pada bagian teks yang relevan. Hasil eksperimen menunjukkan bahwa mekanisme *attention* mampu meningkatkan kemampuan model dalam memahami konteks penting dalam ulasan. Namun, penelitian ini masih berorientasi pada analisis retrospektif terhadap film yang telah dirilis, bukan pada prediksi *excitement* untuk konten yang akan datang.

He et al. [8] mengusulkan arsitektur *hybrid* BERT-CNN-BiLSTM-*Attention* yang dirancang khusus untuk menangani karakteristik komentar pendek dan kompleks dalam media sosial. Penelitian mereka mengidentifikasi bahwa kombinasi *multiple architectures* dapat mengatasi keterbatasan individual dari setiap model. Metodologi penelitian melibatkan *preprocessing* data dari *multiple social media platforms*, *tokenisasi* adaptif, dan *ensemble learning*. Hasil penelitian menunjukkan bahwa model *hybrid* tersebut unggul dalam menangani variasi bahasa informal dan ekspresi emosional yang ambigu. Meskipun signifikan dalam konteks teknologi, penelitian ini belum mengeksplorasi aplikasi spesifik untuk analisis ekspektasi audiens terhadap film yang belum dirilis.

Khan et al. [9] melakukan studi komparatif antara berbagai model *deep learning* termasuk BERT, XLNet, dan varian lainnya untuk analisis sentimen film. Penelitian mereka menunjukkan bahwa XLNet mengungguli BERT dalam hal akurasi dengan pencapaian 87,68% dibandingkan 82,24% untuk BERT. Metodologi penelitian mencakup evaluasi *cross-validation*, analisis kompleksitas komputasi, dan pengujian *robustness* model terhadap variasi data. Temuan penelitian mengindikasikan bahwa model berbasis *Transformer* generasi terbaru memiliki keunggulan dalam memahami dependensi jangka panjang dalam teks. Namun, evaluasi mereka tetap terfokus pada *dataset* ulasan konvensional, bukan pada data ekspektasi atau antusiasme *pre-release*.

Jin et al. [10] mengembangkan pendekatan BUGE (BERT + *Graph*) yang mengintegrasikan representasi BERT dengan *graph neural networks* untuk menghindari masalah *over-smoothing* pada *dataset* teks besar. Penelitian mereka mengidentifikasi bahwa pendekatan berbasis graf dapat mempertahankan informasi lokal sambil menangkap relasi global dalam *dataset*. Metodologi yang digunakan melibatkan konstruksi graf berdasarkan similarity semantik, implementasi *graph convolution layers*, dan optimisasi *end-to-end*. Hasil eksperimen menunjukkan efektivitas pendekatan ini dalam menangani *dataset* skala besar dengan variasi tinggi. Meskipun inovatif dalam aspek teknis, penelitian ini belum mengaplikasikan pendekatan tersebut untuk analisis *excitement* spesifik terhadap film yang akan datang.

Sayeed et al. [11] melakukan evaluasi komprehensif terhadap kemampuan BERT dalam menangani berbagai jenis sentimen, termasuk analisis terhadap kemampuan model dalam memahami konteks dan menangani review panjang. Penelitian mereka mengidentifikasi bahwa BERT unggul dalam memahami konteks semantik yang kompleks, namun mengalami kesulitan dalam mengklasifikasikan sentimen netral dengan akurasi yang konsisten. Metodologi penelitian melibatkan *fine-tuning* BERT pada *multiple domains*, analisis *confusion matrix* untuk identifikasi pattern error, dan evaluasi *robustness*. Temuan penelitian menunjukkan bahwa BERT memerlukan data training yang lebih beragam untuk menangani sentimen netral secara efektif. Fokus penelitian mereka masih berada pada analisis umum, bukan pada aspek prediktif *excitement*.

Farasalsabila et al. [12] mengimplementasikan pendekatan klasik menggunakan *Support Vector Machine* (SVM) dengan *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk analisis sentimen ulasan film, mencapai akurasi 91,27%. Penelitian mereka menunjukkan bahwa meskipun pendekatan tradisional masih dapat memberikan hasil yang baik, namun tidak memiliki kemampuan memahami konteks semantik yang kompleks seperti yang dimiliki oleh model berbasis *Transformer*.

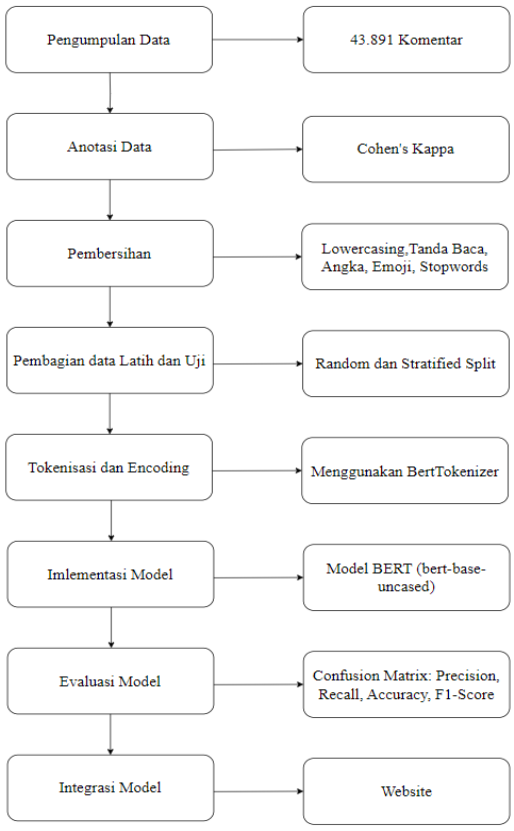
Berdasarkan tinjauan pustaka yang telah dipaparkan, teridentifikasi bahwa meskipun model BERT dan turunannya telah terbukti efektif dalam analisis sentimen ulasan film dengan performa tinggi, terdapat kesenjangan penelitian yang signifikan dalam hal analisis *excitement* audiens terhadap film yang belum dirilis. Mayoritas penelitian terdahulu berfokus pada analisis retrospektif terhadap ulasan film yang telah tayang, dengan pendekatan evaluatif terhadap konten yang sudah dikonsumsi audiens.

Penelitian ini memiliki kebaruannya pada penerapan BERT untuk analisis prediktif *excitement* audiens terhadap film yang akan datang berdasarkan data media sosial. Kebaruan penelitian ini mencakup pengembangan metodologi analisis sentimen yang spesifik untuk mengidentifikasi dan mengkuantifikasi tingkat antusiasme audiens sebelum perilisan film. Penelitian ini mengisi gap yang ada dengan mengadaptasi kemampuan BERT dalam memahami konteks emosional untuk keperluan prediksi *excitement*, yang berbeda secara fundamental dari analisis sentimen retrospektif pada penelitian sebelumnya. Kontribusi utama penelitian ini adalah pengembangan model yang dapat memberikan insight prediktif untuk mendukung strategi pemasaran dan distribusi film dalam konteks industri perfilman modern.

METODOLOGI PENELITIAN

Tahapan Penelitian

Penelitian ini dirancang dengan menggunakan pendekatan kuantitatif eksperimental yang melibatkan implementasi model *machine learning* untuk analisis sentimen teks. Kerangka metodologi penelitian disusun secara sistematis untuk memastikan validitas dan reliabilitas hasil yang diperoleh. Tahapan penelitian ini dapat digambarkan dalam diagram alir sebagai berikut:



Gambar 1. Alur Tahap Penelitian
Sumber: Hasil Olah Data

Berdasarkan Gambar 1, tahapan dalam alur penelitian dapat dijelaskan sebagai berikut:

a. Input
Sumber data primer penelitian ini berupa komentar audiens terhadap film yang akan datang, yang diperoleh dari platform media sosial Facebook dan Instagram. Pemilihan komentar difokuskan pada respons terhadap konten promosi film seperti *trailer*, poster, pengumuman *cast*, dan material pemasaran lainnya. *Platform* Facebook dan Instagram dipilih berdasarkan tingginya penetrasi penggunaan di Indonesia serta aktifitas diskusi yang signifikan terkait konten hiburan dan film.

b. Proses
Tahapan proses penelitian terdiri dari dua komponen utama yaitu *preprocessing* data dan implementasi model BERT.
Tahap *Preprocessing*, pada fase ini, data mentah komentar melalui serangkaian proses persiapan. Tahap awal melibatkan anotasi manual untuk memberikan label kategori emosi *excitement* pada setiap komentar berdasarkan tingkat antusiasme yang teridentifikasi. Proses pembersihan data (*data cleaning*) dilakukan untuk mengeliminasi elemen-elemen yang tidak relevan seperti karakter khusus, *hyperlink*, *hashtag*, dan mention. Normalisasi teks diterapkan untuk mengkonversi seluruh karakter menjadi *lowercase* dan mengstandardisasi format penulisan. Data yang telah dibersihkan kemudian dipartisi menjadi *training set* dan *testing set* dengan proporsi yang telah ditentukan. Tahap terakhir adalah *tokenisasi*, yaitu konversi teks menjadi format token yang kompatibel dengan arsitektur BERT.
Tahap Implementasi model BERT dilakukan melalui proses *fine-tuning* menggunakan data yang telah dipreparasi. Komentar yang telah ditokenisasi menjadi input untuk model BERT yang akan melakukan klasifikasi berdasarkan tiga kategori: *excited*, *not excited*, dan netral. Proses training model menggunakan teknik *supervised learning* dengan optimisasi *hyperparameter* untuk mencapai performa optimal. Evaluasi model dilakukan menggunakan *confusion matrix* dengan perhitungan metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

c. Output
Hasil akhir penelitian berupa model terlatih yang mampu mengklasifikasikan komentar berdasarkan tingkat *excitement*. Evaluasi performa model divisualisasikan melalui metrik klasifikasi yang mencakup *accuracy*, *precision*, *recall*, dan *F1-score*. Model yang telah divalidasi diintegrasikan ke dalam sistem berbasis *web application*, memungkinkan klasifikasi otomatis komentar baru secara *real-time*. Output sistem diharapkan dapat memberikan insight strategis untuk mendukung pengambilan keputusan dalam pemasaran film.

Pengumpulan Data

Metode pengumpulan data dalam penelitian ini menggunakan pendekatan web scraping melalui ekstensi *browser Google Chrome* yang dirancang khusus untuk ekstraksi komentar dari platform media sosial. Pemilihan metode ini didasarkan pada efisiensi dan kemudahan implementasi tanpa memerlukan akses *Application Programming Interface* (API) atau pengembangan *script* khusus. Data primer yang dikumpulkan berupa komentar publik pada postingan resmi yang berkaitan dengan film yang akan dirilis, mencakup respons terhadap *trailer* film, poster promosi, pengumuman *casting*, dan konten pemasaran lainnya. Periode pengumpulan data dilaksanakan selama bulan Maret hingga April 2025, yang bertepatan dengan fase promosi intensif berbagai film *blockbuster* yang dijadwalkan rilis tahun 2025.

Kriteria seleksi komentar meliputi: (1) komentar dalam bahasa Inggris, (2) berkaitan langsung dengan film yang akan datang, (3) mengandung ekspresi opini atau emosi, dan (4) bukan spam atau konten promosi. Data yang berhasil diekstraksi disimpan dalam format *Comma Separated Values* (CSV) untuk memfasilitasi proses analisis dan manipulasi data pada tahap selanjutnya.

Pengumpulan data dilakukan pada periode 1–30 April 2025, dengan mempertimbangkan fase promosi aktif dari delapan film yang menjadi objek penelitian, yaitu *A Minecraft Movie*, *Thunderbolts*, *Final Destination: Bloodlines*, *Mission Impossible: The Final*

Reckoning, *Lilo & Stitch*, *How to Train Your Dragon*, *Jurassic World Rebirth*, dan *The Fantastic Four: First Step*. Karena bersumber dari promosi film internasional, penelitian ini hanya menggunakan komentar berbahasa Inggris agar sesuai dengan kebutuhan model.

Anotasi Data

Tahapan anotasi dilakukan untuk memberikan label pada setiap komentar sesuai tingkat antusiasme audiens. Sebelum proses anotasi, dilakukan perekrutan anotator dengan memberikan panduan berisi ketentuan klasifikasi dan contoh komentar sebagai acuan. Setiap komentar diberi label oleh annotator berdasarkan pedoman yang telah disusun agar hasilnya konsisten dan tidak bias [13]. Data komentar kemudian dibagi rata kepada anotator untuk mempermudah proses penandaan.

Proses anotasi menghasilkan tiga kelas utama, yaitu *Excited*, *Not Excited*, dan Netral. Untuk menjaga kualitas dan konsistensi, dilakukan evaluasi kesepakatan antar anotator menggunakan *Cohen's Kappa*. Setelah itu, komentar yang dianggap tidak relevan dihapus dari *dataset*. Hasil akhir berupa kumpulan data berlabel yang siap diproses lebih lanjut.

Pembersihan Data

Data yang telah dianotasi selanjutnya melalui tahap *cleaning* untuk memastikan kualitas teks yang digunakan. Tahapan *cleaning* meliputi penghapusan karakter asing, simbol, emoji, spasi ganda, hingga baris kosong. Selain itu, seluruh teks diubah menjadi huruf kecil untuk menyamakan format penulisan. Komentar dengan kata-kata informal atau slang dinormalisasi menjadi bentuk baku agar dapat dipahami dengan lebih baik oleh *tokenizer* BERT.

Pembagian Data Latih dan Uji

Setelah data bersih diperoleh, tahap berikutnya adalah pembagian data menjadi data latih dan data uji dengan proporsi 80:20. Dua pendekatan digunakan dalam proses ini, yaitu *random split* dan *stratified split*.

Pada *random split*, data dibagi secara acak tanpa mempertimbangkan distribusi label. Pada *stratified split*, distribusi label dipertahankan agar seimbang di data uji, sehingga evaluasi performa model dapat mencerminkan kinerja yang lebih adil pada setiap kelas.

Tokenisasi dan Encoding

Tahapan berikutnya adalah mengubah data teks ke dalam format yang dapat diproses oleh model BERT. Proses ini dilakukan melalui dua langkah, yaitu *tokenisasi* dan *encoding*. *Tokenisasi* bertujuan memecah teks menjadi token berbasis sub-kata, serta menambahkan token khusus seperti [CLS] dan [SEP]. Setelah itu, setiap token dikonversi ke dalam representasi numerik menggunakan *BertTokenizer* dari pustaka *Huggingface*. Dengan proses ini, komentar yang semula berupa teks dapat dipahami oleh model dalam bentuk vektor numerik.

Implementasi Model BERT

Penelitian ini menggunakan *BERT-base uncased*, sebuah model bahasa pra-pelatihan (*Pre-trained Language Model*) yang telah melalui proses pelatihan awal menggunakan kumpulan data berskala besar dengan berbagai tujuan seperti *Masked-Language Modeling* dan *Causal Language Modeling* [14].

Secara arsitektural, BERT merupakan salah satu penerapan langsung dari *deep learning* dalam bidang pemrosesan bahasa alami. *Deep learning* sangat efektif dalam menangani data berukuran besar, tidak terstruktur, dan mengandung pola yang kompleks seperti gambar, suara, dan teks [15]. Model ini dibangun di atas arsitektur *Transformer encoder* yang terdiri dari banyak lapisan jaringan saraf tiruan, sehingga memiliki kedalaman (*depth*) dan kompleksitas yang tinggi [14]. Sebagai model *deep learning*, BERT tidak hanya mempelajari pola sederhana dalam teks, tetapi juga menangkap hubungan antar kata melalui mekanisme *self-attention*.

Dalam penelitian ini, *BERT-base uncased* dilatih menggunakan data latih hasil *tokenisasi*. Proses pelatihan dilakukan melalui beberapa eksperimen dengan parameter yang berbeda untuk menemukan konfigurasi terbaik dalam mengklasifikasikan komentar. Evaluasi performa model dilakukan menggunakan *confusion matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil pengujian dari berbagai eksperimen kemudian dibandingkan untuk menentukan performa terbaik dari model.

Evaluasi Model

Evaluasi model merupakan tahapan penting yang bertujuan untuk menilai sejauh mana sistem klasifikasi yang dibangun mampu memberikan hasil prediksi yang sesuai dengan label sebenarnya. Pada penelitian ini, evaluasi difokuskan pada kemampuan model *Bidirectional Encoder Representations from Transformers* (BERT) dalam mengklasifikasikan komentar ke dalam tiga kategori utama, yaitu *Excited*, Netral, dan *Not Excited*. Metode utama yang digunakan dalam evaluasi adalah *confusion matrix*. Melalui *confusion matrix*, dapat diketahui distribusi prediksi model terhadap label aktual sehingga terlihat dengan jelas jumlah komentar yang diprediksi benar maupun salah pada masing-masing kelas. Informasi ini menjadi dasar untuk menghitung sejumlah metrik evaluasi yang lebih komprehensif, yaitu *accuracy*, *precision*, *recall*, dan *F1-score*. Dengan kombinasi metrik tersebut, evaluasi tidak hanya terbatas pada tingkat akurasi secara umum, tetapi juga memperhatikan performa model pada setiap kelas. Hal ini penting mengingat adanya kemungkinan ketidakseimbangan distribusi data pada tiap label, sehingga model perlu dinilai secara lebih adil.

Integrasi Model

Tahap akhir dari penelitian ini adalah melakukan integrasi model yang telah dilatih dan dievaluasi ke dalam sebuah aplikasi web sederhana. Integrasi ini berfungsi sebagai sarana visualisasi hasil klasifikasi sekaligus memberikan fasilitas pengujian interaktif terhadap komentar baru yang dimasukkan pengguna. Aplikasi web dirancang dengan memanfaatkan bahasa pemrograman *HTML* untuk antarmuka serta *Python* beserta pustaka pendukung untuk menghubungkan model BERT yang telah dilatih sebelumnya. Dengan adanya integrasi ini, sistem tidak hanya berhenti pada tahap evaluasi model, tetapi juga dapat menunjukkan penerapannya secara langsung dalam bentuk aplikasi yang mampu memberikan prediksi kategori *excitement* terhadap komentar film.

HASIL DAN PEMBAHASAN

Persiapan Data

Dataset pada penelitian ini diperoleh melalui proses *web scraping* menggunakan *Comment Exporter*, sebuah ekstensi pada *Web Browser* yang memungkinkan pengeksporan komentar dari media sosial. Sumber data yang dikumpulkan berasal dari platform Facebook dan Instagram, dengan fokus pada komentar pada unggahan resmi terkait film yang akan datang, seperti *trailer*, poster, dan materi promosi lainnya.

Proses pengumpulan dilakukan dalam rentang waktu 1–30 April 2025, bertepatan dengan fase promosi aktif dari delapan film yang diteliti. Pemilihan periode tersebut bertujuan agar data yang diperoleh dapat merepresentasikan respons audiens sebelum film ditayangkan di bioskop.

Secara keseluruhan, terkumpul 43.891 komentar berbahasa Inggris yang kemudian dijadikan dasar dalam tahap *preprocessing* dan anotasi. Distribusi jumlah komentar dari masing-masing film ditampilkan pada Tabel 1.

Tabel 1. Jumlah Data untuk setiap Film

No	Judul Film	Jumlah Komentar
1	*A Minecraft Movie*	5015
2	*Thunderbolts*	5608
3	*Final Destination: Bloodlines*	5661
4	*Mission Impossible: The Final Reckoning*	5102
5	*Lilo & Stitch*	5490
6	*How to Train Your Dragon*	5836
7	*Jurassic World Rebirth*	5621
8	*The Fantastic Four: First Step*	5558
	Total Komentar	43.891

Setelah tahap akuisisi data, dilakukan proses pembersihan untuk memastikan kualitas komentar yang akan dianalisis. *Cleaning* ini mencakup penghapusan komentar yang bersifat duplikat, tautan URL, simbol asing, komentar kosong, serta elemen lain seperti emotikon atau stiker yang tidak memiliki makna kontekstual. Dari total 43.891 komentar awal, proses *cleaning* menghasilkan 27.266 komentar, sehingga sekitar 62% data tetap digunakan. Komentar dibagi ke dalam 10 file berisi sekitar 2.700 komentar per file (termasuk *overlap*). Setiap file memuat campuran komentar dari satu hingga tiga film yang dipilih secara acak untuk memudahkan distribusi anotasi dan menjaga keragaman konteks. Persentase jumlah berkurangnya komentar dapat dilihat pada Tabel 2.

Tabel 2. Jumlah Pembagian Komentar untuk Anotasi

No	Judul Film	Jumlah Komentar
1	*A Minecraft Movie*	2702
2	*A Minecraft Movie, Thunderbolts*	2701
3	*Thunderbolts, Final Destination: Bloodlines*	2699
4	*Final Destination: Bloodlines, Mission Impossible: The Final Reckoning*	2702
5	*Mission Impossible: The Final Reckoning, Lilo & Stitch*	2701
6	*Lilo & Stitch, How to Train Your Dragon*	2702
7	*How to Train Your Dragon*	2703
8	*How to Train Your Dragon, Jurassic World Rebirth*	2702
9	*Jurassic World Rebirth, The Fantastic Four: First Step*	2702
10	*The Fantastic Four: First Step*	2702
	Total Komentar	27.266

Proses Anotasi

Proses anotasi dilakukan untuk memberikan label klasifikasi pada komentar yang telah terkumpul. Anotasi ini dilakukan secara manual oleh 10 orang anotator yang telah melalui tahap perekrutan dan diberikan panduan klasifikasi berupa dokumen berisi ketentuan pelabelan serta contoh komentar. Dengan adanya panduan ini, diharapkan seluruh anotator memiliki persepsi yang seragam dalam menentukan label.

Pada tahap pembagian data, setiap anotator memperoleh satu file Excel yang berisi rata-rata 2.700 komentar. Selain itu, pada masing-masing file disisipkan 50 komentar identik (*overlap* data) yang didistribusikan kepada dua anotator berbeda. Komentar tambahan ini digunakan untuk mengukur tingkat kesepakatan antar-anotator menggunakan *Cohen's Kappa*. Form anotasi yang digunakan berisi tiga kolom utama, yakni Title (judul film), Content (isi komentar), dan Klasifikasi. Anotator memilih salah satu label yang telah disediakan, yaitu *Excited*, Netral, *Not Excited*, dan Tidak Relevan.

Hasil evaluasi kesepakatan menunjukkan bahwa nilai *Cohen Kappa* berada pada rentang 0,77 hingga 0,90 untuk seluruh pasangan anotator. Rentang nilai tersebut menunjukkan tingkat kesepakatan yang kuat (*substantial agreement*), sehingga hasil anotasi dianggap reliabel untuk dijadikan *dataset* pelatihan. Rincian hasil evaluasi ditampilkan pada Tabel 3.

Tabel 3. Nilai *Cohen Kappa* pasang Anotator

No	Anotator	Nilai *Cohen Kappa*
1	Pasangan Anotator 1 dan 2	0.90
2	Pasangan Anotator 3 dan 4	0.80
3	Pasangan Anotator 5 dan 6	0.85
4	Pasangan Anotator 7 dan 8	0.77
5	Pasangan Anotator 9 dan 10	0.80

Setelah seluruh file selesai dianotasi, dilakukan proses pembersihan untuk menghapus komentar yang diberi label "Tidak Relevan", seperti komentar spam, komentar dalam bahasa non-Inggris, serta komentar yang tidak memiliki makna jelas. Dari total 27.266 komentar hasil anotasi, sebanyak 4.549 komentar (16,7%) dihapus. Dengan demikian, jumlah data yang siap digunakan pada tahap selanjutnya adalah 22.717 komentar atau setara dengan 51,8% dari total data awal (43.891 komentar). Hasil dari proses pembersihan ini akan dijabarkan lebih lanjut pada tabel berikut.

Tabel 4. Distribusi Klasifikasi Berdasarkan Judul

No	Judul Film	Excited	Netral	Not Excited	Jumlah
1	*A Minecraft Movie*	1019	773	880	2672
2	*Thunderbolts*	1653	1290	312	3255
3	*Final Destination: Bloodlines*	2317	269	220	2806
4	*Mission Impossible: The Final Reckoning*	1798	653	39	2490
5	*Lilo & Stitch*	1676	881	201	2758
6	*How to Train Your Dragon*	2149	1003	151	3303
7	*Jurassic World Rebirth*	1253	722	679	2654
8	*The Fantastic Four: First Step*	1080	969	730	2779
	Total	12945	6560	3212	22717

Pembersihan

Setelah proses anotasi dilakukan, data yang memiliki label Tidak Relevan dikeluarkan dari *dataset*. Dari total komentar yang terkumpul, tersisa sebanyak 22.717 komentar yang kemudian melalui tahap pembersihan untuk memastikan kualitas data sebelum diproses pada tahap tokenisasi dan pelatihan model. Proses pembersihan dilakukan melalui beberapa langkah. Pertama, karakter asing, simbol yang tidak memiliki makna, emotikon yang tidak diperlukan, serta spasi ganda dihapus dari teks. Kedua, seluruh komentar diubah menjadi huruf kecil (*lowercase*) untuk menyamakan bentuk penulisan. Komentar kosong atau baris yang tidak berisi teks juga dikeluarkan dari *dataset*. Selanjutnya, dilakukan normalisasi terhadap kata-kata gaul (slang) atau bentuk tulisan informal agar menjadi kata baku dalam bahasa Inggris.

Sebagai contoh, komentar "*omg this looks sooo dopeee!!*" setelah pembersihan berubah menjadi "*oh my god this looks so dope!!*". Pada contoh tersebut, kata singkatan "*omg*" dinormalisasi menjadi "*oh my god*", sedangkan kata dengan huruf berlebih seperti "*sooo*" dan "*dopeee*" disederhanakan menjadi "*so*" dan "*dope*". Dengan langkah ini, data menjadi lebih konsisten namun tetap mempertahankan nuansa emosional yang terkandung, sehingga memudahkan model dalam memahami konteks *excitement* yang diekspresikan oleh pengguna.

Pembagian Data

Total data yang digunakan pada tahap ini adalah sebanyak 22.717 komentar. Pembagian data bertujuan untuk memisahkan sebagian besar data sebagai data latih (*training* data), sementara sebagian lainnya digunakan sebagai data uji (testing data) untuk mengukur kinerja model terhadap data yang belum pernah dilihat sebelumnya.

Proporsi pembagian ditetapkan sebesar 80:20, yaitu 80% untuk data latih dan 20% untuk data uji. Dalam penelitian ini digunakan dua pendekatan berbeda, yaitu:

a. *Random Split*
   Pada metode ini, data dibagi secara acak tanpa mempertimbangkan distribusi label pada setiap kelas. Dengan pendekatan ini, meskipun jumlah data yang digunakan sesuai proporsi, ada kemungkinan distribusi antarlabel menjadi tidak seimbang, khususnya pada data uji.
b. *Stratified Split*
   Berbeda dengan *random split*, metode ini mempertahankan distribusi label yang seimbang pada data uji. Dengan demikian, proporsi komentar pada kategori *Excited*, Netral, dan *Not Excited* tetap konsisten antara data latih dan uji. Hal ini diharapkan dapat memberikan hasil evaluasi yang lebih representatif, terutama dalam menilai performa model pada masing-masing kelas. Distribusi jumlah data pada masing-masing metode ditunjukkan dalam tabel berikut:

Tabel 5. Distribusi Jumlah Data Latih dan Uji (*Random*)

Klasifikasi	Data Uji	Data Latih
Excited	10353	2592
Netral	5231	1329
Not Excited	2589	623

Selanjutnya, adalah distribusi untuk jumlah data latih dan uji dengan pendekatan *Stratified*.

Tabel 6. Distribusi Jumlah Data Latih dan Uji (*Stratified*)

Klasifikasi	Data Uji	Data Latih
Excited	12289	642

Netral	5945	642
Not Excited	2557	642

Tokenisasi dan Encoding

Data latih dan uji hasil pembagian dengan metode *random split* dan *stratified split* kemudian diproses melalui tahap *tokenisasi* dan *encoding* menggunakan *BertTokenizer* dari pustaka *Huggingface*. Proses ini mengubah setiap komentar menjadi token berbasis sub-kata, menambahkan token khusus [CLS] dan [SEP], lalu mengonversinya ke dalam bentuk indeks numerik sesuai vocabulary BERT.

Hasil dari tahap ini disimpan dalam format tensor .pt agar dapat langsung digunakan pada proses pelatihan. Empat berkas yang dihasilkan adalah train_encoded_random.pt, test_encoded_random.pt, train_encoded_balanced.pt, dan test_encoded_balanced.pt, masing-masing merepresentasikan data latih dan uji untuk kedua pendekatan pembagian (*Random* dan *Stratified*).

Pelatihan

Pada tahap ini, model yang digunakan untuk mengklasifikasikan komentar adalah *Bidirectional Encoder Representations from Transformers* (BERT) dengan memanfaatkan arsitektur *BertForSequenceClassification*. Model ini dipilih karena telah terbukti efektif dalam memahami konteks kalimat dan relasi antar kata, sehingga sesuai untuk tugas klasifikasi komentar.

Variasi yang digunakan adalah *bert-base-uncased*, yaitu model *pra*-latih pada korpus bahasa Inggris yang mengabaikan perbedaan huruf kapital. Pemilihan varian ini didasarkan pada alasan utama yaitu bahasa komentar yang digunakan adalah bahasa Inggris informal yang tidak selalu konsisten dalam penggunaan kapitalisasi.

Pelatihan dilakukan dengan beberapa skenario eksperimen, yang masing-masing memanfaatkan kombinasi *epoch* dan *learning rate* berbeda. Tujuannya adalah untuk melihat bagaimana pengaturan parameter memengaruhi performa model dalam melakukan klasifikasi. Rincian parameter dari tiap eksperimen ditunjukkan pada Tabel 7 berikut.

Tabel 7. Parameter Eksperimen Pelatihan Model

No	Eksperimen	*Random* dan *Stratified*	
		Epoch	***Learning Rate***
1	A	3	2e-5
2	B	4	2e-5
3	C	3	3e-5
4	D	4	3e-5
5	E	6	2e-5
6	F	6	3e-5

Seluruh proses pelatihan dijalankan pada *Google Colab* dengan dukungan GPU. Pemanfaatan GPU ini sangat penting mengingat arsitektur BERT memiliki jumlah parameter yang besar dan memerlukan komputasi intensif. Selain itu, implementasi model dilakukan dengan bantuan *library PyTorch* dan *Huggingface Transformers*, yang menyediakan fungsi siap pakai untuk *tokenisasi*, pengelolaan *dataset*, pelatihan, hingga penyimpanan model terlatih.

Dari eksperimen yang dijalankan, dapat diamati bahwa variasi jumlah *epoch* dan *learning rate* memberikan dampak yang berbeda terhadap performa model. Misalnya, penggunaan *epoch* yang lebih banyak memungkinkan model mempelajari pola dengan lebih baik, namun juga berpotensi menimbulkan *overfitting*. Sebaliknya, penyesuaian *learning rate* yang terlalu tinggi dapat membuat model sulit mencapai konvergensi optimal.

Evaluasi Model

Metode evaluasi yang digunakan adalah *confusion matrix*. *Confusion Matrix* menjadi dasar perhitungan seluruh evaluasi model klasifikasi dan sangat membantu dalam mengidentifikasi pola kesalahan model, seperti kecenderungan memprediksi kelas tertentu secara berlebihan atau kekurangan [16]. Dasar perhitungan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. *Confusion matrix* menyajikan informasi mengenai jumlah prediksi yang benar maupun salah berdasarkan label aktual dan prediksi model, yang mencakup komponen *True Positive*, *True Negative*, *False Positive*, dan *False Negative*.

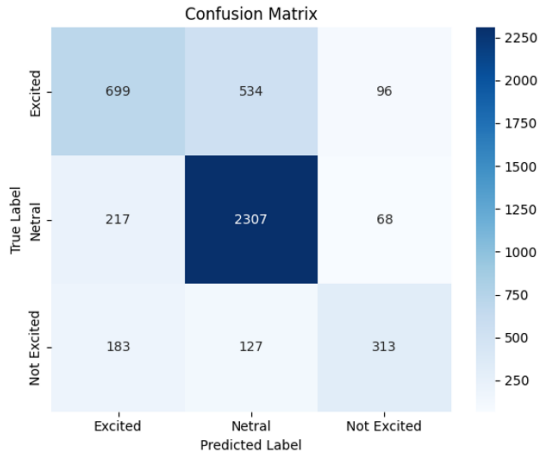
Berdasarkan hasil eksperimen dengan pembagian data secara acak (*random split*), model menghasilkan performa yang cukup baik dengan nilai akurasi berada pada kisaran 72% hingga 73%. Namun, performa model cenderung kurang stabil dalam mengklasifikasikan komentar dari kelas yang cenderung ambigu atau beririsan secara konteks, khususnya antara kelas Netral dan *Not Excited*. Rangkuman hasil evaluasi dari seluruh eksperimen A hingga F pada *random split* disajikan pada tabel berikut.

Tabel 8. Hasil Evaluasi Model pada Data *Random*

No	Eksperimen	*Precision* (%)	*Recall* (%)	*Accuracy* (%)	*F1* (%)
1	A	71.70	72.82	72.82	71.77
2	B	71.86	72.62	72.62	72.11
3	C	71.89	72.82	72.82	72.11
4	D	71.90	72.93	72.93	71.97
5	E	71.94	73.04	73.04	71.98
6	F	71.63	72.47	72.47	71.83

Rata - rata	71.82	72.78	72.78	71.96

Berdasarkan Tabel diatas, model dengan pembagian data secara acak menghasilkan rata-rata *F1-score* sebesar 71.96%, yang menunjukkan performa yang cukup baik namun masih dapat ditingkatkan. Nilai *F1-score* ini mengindikasikan bahwa model mampu mencapai keseimbangan yang wajar antara *precision* (71.82%) dan *recall* (72.78%). Meskipun *accuracy* mencapai 72.78%, *F1-score* memberikan gambaran yang lebih akurat tentang performa model mengingat adanya ketidakseimbangan kelas dalam *dataset*.



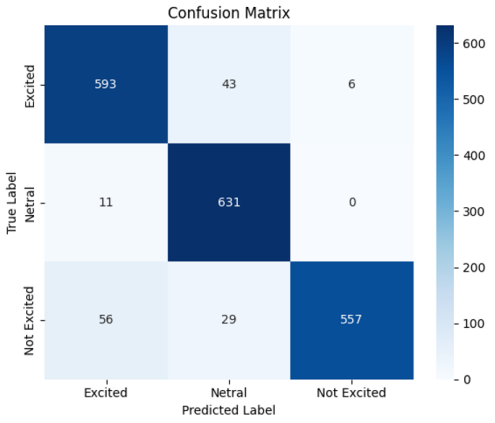
Gambar 2. Matrix Terbaik dengan Data *Random*
Sumber: Hasil Pelatihan Model

Selanjutnya, penggunaan pembagian data stratified memberikan hasil yang lebih stabil dan optimal dibandingkan *random split*. Nilai akurasi yang diperoleh berada pada rentang 84% hingga 92%. Selain itu, *precision*, *recall*, dan *F1-score* untuk masing-masing kelas juga menunjukkan peningkatan yang signifikan. Hal ini menunjukkan bahwa keseimbangan distribusi label pada data latih dan uji memberikan dampak positif terhadap performa model. Hasil lengkap evaluasi pada *stratified split* dapat dilihat pada tabel berikut ini.

Tabel 9. Hasil Evaluasi Model pada Data *Stratified*

No	Eksperimen	*Precision* (%)	*Recall* (%)	*Accuracy* (%)	*F1* (%)
1	A	85.04	84.32	84.32	84.31
2	B	89.86	89.35	89.35	89.33
3	C	87.63	87.12	87.12	87.08
4	D	91.39	90.91	90.13	90.89
5	E	91.96	91.58	91.58	91.57
6	F	92.84	92.47	92.47	92.45
Rata - rata		89.45	89.29	89.16	89.27

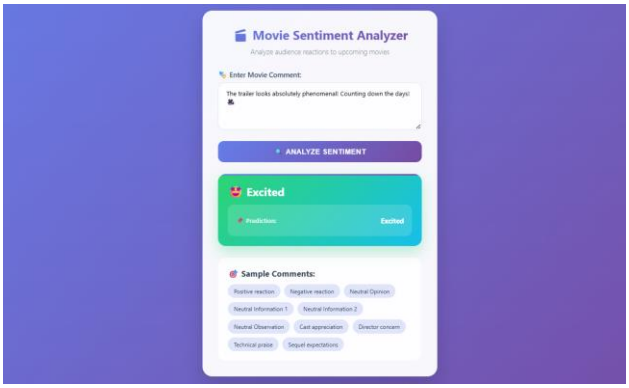
Berdasarkan tabel di atas, terlihat bahwa penggunaan pembagian data secara *stratified* memberikan hasil rata – rata yang lebih stabil dan optimal dibandingkan dengan pembagian data acak. *F1-score* meningkat signifikan menjadi 89.27%, menunjukkan peningkatan 17.31% dari hasil *random* split. Peningkatan *F1-score* mengindikasikan bahwa model tidak hanya akurat secara keseluruhan, tetapi juga mampu mengenali ketiga kelas sentimen dengan lebih seimbang. Nilai *precision* (89.45%) dan *recall* (89.29%) yang hampir setimbang mengkonfirmasi bahwa *F1-score* memberikan representasi yang tepat tentang performa model. Hal ini membuktikan bahwa keseimbangan distribusi label pada data latih dan uji berdampak positif terhadap performa model.



Gambar 3. Matrix Terbaik dengan Data *Stratified*
Sumber: Hasil Pelatihan Model

Integrasi Model

Tahap akhir penelitian ini adalah mengintegrasikan model yang telah dilatih dan dievaluasi ke dalam sebuah aplikasi web sederhana. Integrasi ini bertujuan sebagai sarana visualisasi sekaligus pengujian hasil prediksi model secara interaktif.



Gambar 4. Antarmuka Integrasi dengan Model
Sumber: Hasil Integrasi Model

Seperti terlihat pada Gambar 4, antarmuka web dirancang dengan tampilan minimalis untuk mempertahankan fokus pada fungsi utama, yaitu menampilkan hasil klasifikasi komentar. Pengguna dapat memasukkan teks komentar pada kolom input

yang tersedia, kemudian menekan tombol submit untuk memproses masukan tersebut. Model yang telah dilatih sebelumnya akan mengolah teks tersebut dan menampilkan hasil prediksi pada halaman yang sama. Selain itu, disediakan pula beberapa contoh komentar di bagian bawah antarmuka sebagai data uji cepat (*quick test*) yang dapat digunakan untuk menguji sistem secara langsung.

KESIMPULAN DAN SARAN

Penelitian ini dilakukan untuk menjawab rumusan masalah mengenai efektivitas model BERT dalam mengklasifikasikan *excitement* audiens terhadap film yang akan datang berdasarkan komentar di media sosial. Berdasarkan hasil evaluasi, model BERT mampu memberikan performa yang cukup baik dengan akurasi tertinggi mencapai 92,47% pada pembagian data *stratified*. Selain itu, nilai *precision*, *recall*, dan *F1-score* juga menunjukkan konsistensi yang memadai pada ketiga kelas (*Excited*, Netral, *Not Excited*), yang berarti model mampu mengenali pola bahasa secara kontekstual. Proses *preprocessing* yang meliputi pembersihan teks, *tokenisasi*, *encoding*, serta penerapan pembagian data *random* dan *stratified* terbukti berkontribusi pada peningkatan kinerja model. Dengan demikian, dapat disimpulkan bahwa BERT merupakan model yang efektif dan sesuai digunakan untuk analisis *excitement* audiens di ranah ulasan film.

Meskipun begitu, penelitian ini memiliki keterbatasan, khususnya pada cakupan data yang hanya berasal dari komentar berbahasa Inggris. Selain itu, jumlah data yang relatif terbatas serta adanya potensi ketidakseimbangan antar kelas juga dapat memengaruhi stabilitas model, sangat disarankan agar penelitian selanjutnya mengikutsertakan komentar dalam berbagai bahasa, termasuk Bahasa Indonesia, guna mencerminkan persepsi audiens yang lebih beragam. Selain itu, perluasan sumber data dari berbagai *genre* film maupun *platform* media sosial lain seperti Twitter, YouTube, atau TikTok juga dapat memberikan konteks yang lebih kaya.

Dari sisi model, meskipun BERT sudah menunjukkan performa yang baik, pemanfaatan varian *transformer* lain seperti RoBERTa, DistilBERT, atau IndoBERT bisa menjadi opsi yang menarik untuk dibandingkan, sehingga dapat diperoleh gambaran yang lebih komprehensif mengenai efektivitas tiap arsitektur dalam klasifikasi komentar. Penelitian berikutnya juga dapat mengeksplorasi pengaturan parameter secara lebih mendalam, termasuk variasi *epoch*, *learning rate*, dan *batch size*, atau memanfaatkan teknik pelatihan tambahan untuk meningkatkan kinerja model.

Selain itu, arah penelitian dapat difokuskan lebih spesifik pada tingkat excitement terhadap tiap film. Dengan pendekatan ini, distribusi sentimen *Excited*, Netral, dan *Not Excited* dapat dipetakan secara terpisah untuk masing-masing film. Hasil analisis yang lebih detail seperti ini akan memungkinkan evaluasi menggunakan metrik *precision*, *recall*, dan *F1-score* per film, sehingga memberikan wawasan yang lebih mendalam mengenai pola respons audiens terhadap film tertentu dibandingkan film lainnya.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada seluruh pihak yang telah memberikan dukungan, baik berupa bimbingan, fasilitas, maupun motivasi, sehingga penelitian ini dapat terselesaikan dengan baik. Ucapan terima kasih juga ditujukan kepada pembimbing dan rekan-rekan yang berkontribusi dalam proses anotasi data, serta semua pihak lain yang secara langsung maupun tidak langsung membantu kelancaran penelitian ini.

DAFTAR PUSTAKA

[1] M. Picture Association, "Theme Report 2021: The State of Theatrical and Home Entertainment Market."

[2] www.statista.com, "Cinema - Worldwide," https://www.statista.com/outlook/amo/media/cinema/worldwide.

[3] M. Fadheel Djamaly, R. Astini, dan D. Asih, "LITERATURE REVIEW: PERAN MEDIA SOSIAL DALAM PEMASARAN FILM INDONESIA: ANALISIS KEPUASAN DAN NIAT BELI PENONTON," 2023.

[4] B. Kurniawan, K. Rofiki, J. Raya Telang, K. Kamal, K. Bangkalan, dan J. Timur, "PT. Media Akademik Publisher PERAN MEDIA SOSIAL DALAM STRATEGI PEMASARAN FILM: PENGARUHNYA TERHADAP KEPUASAN DAN ANTUSIASME PENONTON FILM," *JMA)*, vol. 2, no. 7, hlm. 3031–5220, 2024, doi: 10.62281.

[5] G. Nkhata, U. Anjum, dan J. Zhan, "Sentiment Analysis of Movie Reviews Using BERT," Feb 2025, [Daring]. Tersedia pada: http://arxiv.org/abs/2502.18841

[6] B. Zhang, "A BERT-CNN Based Approach on Movie Review Sentiment Analysis," *SHS Web of Conferences*, vol. 163, hlm. 04007, 2023, doi: 10.1051/shsconf/202316304007.

[7] W. Ning, F. Wang, W. Wang, H. Wu, Q. Zhao, dan T. Zhang, "Research on movie rating based on BERT-base model," *Sci Rep*, vol. 15, no. 1, Des 2025, doi: 10.1038/s41598-025-92430-w.

[8] A. He dan M. Abisado, "Text Sentiment Analysis of Douban Film Short Comments Based on BERT-CNN-BiLSTM-Att Model," *IEEE Access*, vol. 12, hlm. 45229–45237, 2024, doi: 10.1109/ACCESS.2024.3381515.

[9] S. S. Khan dan Y. Alharbi, "Sentiment analysis of movie review classifications using deep learning approaches," 1 Agustus 2024, *Institute of Advanced Science Extension (IASE)*. doi: 10.21833/ijaas.2024.08.016.

[10] Y. Jin dan A. Zhao, "Bert-based graph unlinked embedding for sentiment analysis," *Complex and Intelligent Systems*, vol. 10, no. 2, hlm. 2627–2638, Apr 2024, doi: 10.1007/s40747-023-01289-9.

[11] M. S. Sayeed, V. Mohan, dan K. S. Muthu, "BERT: A Review of Applications in Sentiment Analysis," 1 Juni 2023, *Ital Publication*. doi: 10.28991/HIJ-2023-04-02-015.

[12] D. D. Nur Cahyo *dkk.*, "Sentiment Analysis for IMDb Movie Review Using Support Vector Machine (SVM) Method," *Inform : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 8, no. 2, hlm. 90–95, Mar 2023, doi: 10.25139/inform.v8i2.5700.

[13] M. Adnanul Islam, M. Saddam Hossain Mukta, P. Olivier, dan M. M. Rahman, "Comprehensive guidelines for emotion annotation," dalam *IVA 2022 - Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*, Association for Computing Machinery, Inc, Sep 2022. doi: 10.1145/3514197.3549640.

[14] A. Vaswani *dkk.*, "Attention Is All You Need," 2023.

[15] G. Nkhata, S. Gauch, U. Anjum, dan J. Zhan, "Fine-tuning BERT with Bidirectional LSTM for Fine-grained Movie Reviews Sentiment Analysis," Feb 2025, [Daring]. Tersedia pada: http://arxiv.org/abs/2502.20682

[16] M. C. Hinojosa Lee, J. Braet, dan J. Springael, "Performance Metrics for Multilabel Emotion Classification: Comparing Micro, Macro, and Weighted F1-Scores," *Applied Sciences (Switzerland)*, vol. 14, no. 21, Nov 2024, doi: 10.3390/app14219863.