

LAPORAN PROJECT MACHINE LEARNING



Analyzing YouTube Gaming Comments: Like Count Regression and Sentiment-Emotion Classification

Kelas LM01 (Kelompok 07) :

Alfito Faiz Rizqi - 2702316724

Bryan Cristhoper Chandra Putra - 2702272306

Reuben Anselmus Adel - 2702314901

Mata Kuliah Machine Learning
Semester Genap 2024/2025

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital telah mengubah cara manusia berinteraksi, termasuk melalui media sosial dan platform berbagi video seperti YouTube. YouTube tidak hanya menjadi sarana hiburan, tetapi juga ruang diskusi dan komunitas daring yang aktif, terutama dalam konten bertema gaming yang mengalami lonjakan popularitas. Komentar pada video gaming mencerminkan berbagai interaksi dan emosi pengguna, namun sebagian besar data teks tersebut belum dimanfaatkan secara optimal. Salah satu indikator penting dalam komentar adalah jumlah like yang mencerminkan tingkat apresiasi komunitas.

Melalui pendekatan regresi, komentar dapat dianalisis untuk memprediksi jumlah like, sedangkan klasifikasi sentimen dan emosi dapat mengelompokkan komentar berdasarkan nuansa perasaan yang terkandung. Untuk mendukung analisis ini, digunakan teknik Natural Language Processing (NLP) dan machine learning yang memungkinkan pemrosesan teks secara otomatis dan pembangunan model prediktif. Penelitian ini berfokus pada dua hal: memprediksi jumlah like dari komentar YouTube gaming dan mengklasifikasikan sentimen serta emosi pengguna. Hasil analisis diharapkan memberikan kontribusi terhadap pemahaman interaksi pengguna, serta menjadi dasar pengembangan sistem rekomendasi dan moderasi konten.

1.2 Rumusan Masalah

Beberapa permasalahan yang ingin diselesaikan melalui proyek ini adalah:

1. Apa saja faktor dalam komentar YouTube bertema gaming yang mempengaruhi jumlah like yang diterima?
2. Bagaimana model regresi dapat digunakan untuk memprediksi jumlah like berdasarkan isi komentar?
3. Bagaimana klasifikasi sentimen dan emosi dapat diterapkan untuk memahami isi komentar pengguna secara lebih mendalam?
4. Sejauh mana hasil klasifikasi sentimen dan emosi dapat memberikan gambaran terhadap suasana interaksi dalam komunitas gaming di YouTube?

1.3 Tujuan Penelitian

Tujuan penelitian ini adalah :

1. Menganalisis hubungan antara isi komentar dengan jumlah like yang diterima pada video YouTube bertema gaming.
2. Membangun model regresi yang mampu memprediksi jumlah like berdasarkan fitur teks dari komentar. Melakukan klasifikasi sentimen (positif, negatif, netral) dan emosi (senang, marah, sedih, dll.) terhadap komentar pengguna.

3. Memberikan pemahaman lebih lanjut mengenai pola interaksi dan emosi pengguna dalam komunitas YouTube gaming melalui pendekatan NLP dan machine learning.

1.4 Manfaat

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut :

1. Manfaat Akademik: Penelitian ini dapat menjadi latihan praktis dalam menerapkan teknik pemrosesan bahasa alami dan machine learning sederhana terhadap data komentar media sosial. Selain itu, penelitian ini memperkuat pemahaman tentang hubungan antara konten teks dan analisis sentimen.
2. Manfaat Praktis: Hasil penelitian ini dapat menjadi referensi awal bagi mahasiswa atau pengembang pemula yang ingin membangun model sederhana untuk menganalisis komentar daring, misalnya dalam sistem rekomendasi atau penyaringan komentar otomatis.
3. Manfaat Sosial: Penelitian ini dapat memberikan gambaran umum tentang suasana komunikasi dan interaksi dalam komunitas gaming online, serta membantu memahami bagaimana emosi dan opini pengguna berkembang di platform digital.

BAB II

METODOLOGI

2.1 Pengumpulan dan Pra-pemrosesan Data

Data yang digunakan di penelitian ini diambil dari komentar pengguna di video YouTube milik beberapa YouTuber gaming populer. Total terdapat 22 kanal YouTube yang dipilih berdasarkan popularitas dan tingkat keterlibatan pengguna (engagement). Beberapa di antaranya termasuk channel @PewDiePie, @jacksepticeye, @Dream, @markiplier, dan lain-lain.

Pengambilan data dilakukan menggunakan YouTube Data API v3 untuk mengakses komentar publik dari video-video milik YouTuber tersebut. Untuk masing-masing channel, dipilih beberapa video populer dengan jumlah komentar yang tinggi. Dari setiap video, dikumpulkan komentar-komentar pengguna beserta metadata penting seperti:

- Isi komentar
- Jumlah Like pada komentar

Data dikumpulkan hingga mencapai kuota maksimum 4500 komentar per channel, dengan target pembagian sebagai berikut:

- 1500 komentar kategori **Low**
- 1500 komentar kategori **Medium**
- 1500 komentar kategori **High**

Proses pra-pemrosesan melibatkan beberapa langkah utama, yaitu:

- Pembersihan Teks

Komentar dibersihkan dari berbagai elemen yang dapat mengganggu analisis teks, antara lain:

- Karakter non-alfabetik
- Tautan (link)
- Emoji dan simbol khusus
- Tag atau entitas HTML

Pembersihan ini bertujuan menyederhanakan representasi teks dan mengurangi noise dalam data.

- Tokenisasi dan normalisasi

Setelah dibersihkan, komentar kemudian :

- Ditokenisasi menjadi kata-kata individual.
- Diubah menjadi huruf kecil (lower casing).

- Dihilangkan stopword-nya, yaitu kata-kata umum yang tidak membawa makna penting (seperti “yang”, “dan”, “itu”).

Langkah ini bertujuan menstandarkan isi komentar agar siap digunakan dalam proses analisis atau pelatihan model berbasis teks.

- **Pengelompokan ke dalam Bin**

Komentar dikelompokkan berdasarkan jumlah like menggunakan metode quantile-based discretization dengan `pd.qcut()` dari pustaka `pandas`. Komentar dibagi ke dalam tiga kategori (bin):

- **Low**: 33% komentar dengan jumlah like terendah.
- **Medium**: 33% komentar dengan jumlah like sedang.
- **High**: 33% komentar dengan jumlah like tertinggi.

Pembagian ini memudahkan klasifikasi komentar berdasarkan tingkat popularitas relatif.

- **Sampling Seimbang**

Untuk menjaga keseimbangan distribusi data antar kategori, dilakukan proses **undersampling**. Setiap kategori hanya menyimpan maksimal **1.500 komentar**. Jika salah satu bin memiliki komentar kurang dari 1.500, maka bin lain akan **disamakan mengikuti jumlah bin terendah**. Misalnya, jika kategori *High* hanya memiliki 1.022 komentar, maka *Low* dan *Medium* juga akan disamakan menjadi 1.022 komentar.

Langkah ini dilakukan agar model atau analisis yang menggunakan data tersebut tidak bias terhadap kelas mayoritas.

- **Penyimpanan dan Struktur Data**

Data akhir disimpan dalam format CSV dengan struktur sebagai berikut:

- `youtuber_name` – Nama channel YouTube sumber komentar
- `comment` – Isi komentar
- `like_count` – Jumlah like pada komentar
- `bin_category` – Kategori bin (low, medium, high)
- `sentiment` – Label sentimen (misalnya: positive, negative, neutral)
- `emotion` – Label emosi (misalnya: joy, surprise, anger, dll.)

Data inilah yang digunakan untuk proses analisis dan pelatihan model pada tahap berikutnya.

2.2 Implementasi Model

Pada tahap ini, dua jenis pendekatan machine learning digunakan untuk menyelesaikan dua tugas berbeda: (1) klasifikasi sentimen dan emosi dari komentar YouTube, serta (2) prediksi jumlah likes berdasarkan konten komentar. Seluruh implementasi dibangun dalam bentuk pipeline yang modular dan dapat direproduksi.

Arsitektur dan Alur Pemodelan

Langkah awal dalam pipeline adalah klasifikasi sentimen dan emosi menggunakan model berbasis RoBERTaForSequenceClassification. Komentar dari YouTube terlebih dahulu di tokenisasi dengan RobertaTokenizerFast, kemudian dimasukkan ke model RoBERTa yang telah difine-tuning untuk masing-masing tugas klasifikasi. Label yang digunakan pada tugas sentimen terdiri dari: positive, neutral, dan negative, sementara label untuk tugas emosi mencakup: joy, sadness, anger, fear, disgust, surprise, dan neutral.

Setelah memperoleh prediksi sentimen dan emosi, informasi tersebut digabungkan dengan teks asli komentar dalam format:

{komentar} [SEP] {label_emosi} [SEP] {label_sentimen}

Gabungan teks ini selanjutnya diubah menjadi representasi numerik menggunakan model BERT (bert-base-uncased). Vektor representasi yang diambil berasal dari token [CLS] pada output layer terakhir dari BERT, yang merepresentasikan konteks keseluruhan kalimat.

Embedding vektor hasil dari BERT kemudian digunakan sebagai input untuk model regresi jumlah likes. Model regresi yang digunakan adalah XGBRegressor, yang telah dilatih sebelumnya dan disimpan dalam format .pkl. Prediksi jumlah likes dilakukan menggunakan model.predict(embedding), kemudian hasilnya dibulatkan dan dijaga agar tidak negatif.

Seluruh proses dilengkapi dengan pengecekan perangkat (CPU atau GPU), dan setiap model dijalankan dalam mode evaluasi (model.eval()) untuk memastikan efisiensi dan konsistensi hasil. Model dan tokenizer dimuat ulang dari direktori Google Drive sebelum digunakan dalam proses inferensi.

Dengan pendekatan ini, pipeline inferensi lengkap tersusun sebagai berikut:

- Input komentar teks
- Prediksi sentimen dan emosi dengan RoBERTa
- Gabungkan komentar dan hasil klasifikasi
- Ekstraksi embedding menggunakan BERT
- Prediksi jumlah likes menggunakan XGBRegressor

2.3 Evaluasi Model

Performa model dievaluasi menggunakan dataset uji (*test set*) yang tidak pernah dilihat oleh model selama proses pelatihan maupun validasi, guna menilai kemampuan generalisasi terhadap data baru. Evaluasi dilakukan terhadap dua jenis tugas: klasifikasi (sentimen dan emosi) dan regresi (jumlah likes). Masing-masing tugas memiliki metrik evaluasi yang disesuaikan.

Untuk model klasifikasi sentimen dan emosi berbasis RoBERTa, digunakan beberapa metrik utama sebagai berikut:

- Accuracy: Proporsi prediksi yang benar terhadap total prediksi, mencerminkan kemampuan model dalam mengklasifikasikan sentimen atau emosi secara keseluruhan.
- Precision: Mengukur ketepatan prediksi untuk setiap label; menunjukkan seberapa banyak dari prediksi suatu kelas yang benar-benar sesuai dengan label tersebut.
- Recall: Menunjukkan seberapa banyak contoh dari suatu kelas yang berhasil dikenali oleh model.
- F1-Score: Rata-rata harmonis dari precision dan recall, memberikan penilaian yang seimbang terhadap performa model, terutama penting dalam kondisi distribusi label yang tidak seimbang

Sementara itu, untuk tugas regresi jumlah likes yang dilakukan dengan model LightGBM Regressor, digunakan metrik sebagai berikut:

- Mean Squared Error (MSE): Mengukur rata-rata kuadrat selisih antara nilai prediksi dan nilai aktual.
- Root Mean Squared Error (RMSE): Akar dari MSE yang mengembalikan kesalahan dalam skala asli (jumlah likes).
- R² Score (Koefisien Determinasi): Mengukur seberapa besar variansi data yang dapat dijelaskan oleh model.

Evaluasi dilakukan terhadap model yang telah dilatih menggunakan kombinasi input berupa teks komentar, label prediksi emosi, dan sentimen yang digabung dan diubah menjadi embedding melalui model BERT, lalu digunakan sebagai input untuk model regresi jumlah likes. Hasil evaluasi menunjukkan bahwa model klasifikasi cukup andal dalam mengenali emosi dominan seperti *joy* dan *sadness*, meskipun masih terdapat tantangan pada kelas minoritas seperti *disgust* dan *surprise*. Sementara itu, model regresi mampu menghasilkan prediksi jumlah likes yang cukup berkorelasi dengan nilai aktual, terutama pada komentar bernada positif dan antusias. Prediksi kemudian dibulatkan dan disesuaikan agar tetap valid secara semantik, seperti menghindari nilai negatif. Secara keseluruhan, pipeline model menunjukkan performa yang baik dalam memahami konteks komentar dan memetakan informasi tersebut ke dalam klasifikasi semantik serta estimasi popularitas.

2.4 Visualisasi Hasil

Untuk memahami karakteristik data komentar YouTube bertema gaming, dilakukan beberapa visualisasi awal. Visualisasi distribusi channel (Gambar 2.1) dan total like per channel (Gambar 2.2) memberikan gambaran awal mengenai tingkat popularitas dan volume data per kreator. Selain itu, distribusi sentimen dan emosi per channel (Gambar 2.3 dan Gambar 2.4) menunjukkan dominasi sentimen positif dan emosi joy pada sebagian besar channel. Analisis juga menunjukkan bahwa terdapat banyak komentar dengan jumlah like nol (Gambar 2.5), yang dapat memengaruhi distribusi target regresi. Visualisasi klasifikasi jumlah like (low, medium, high) pada tiap channel (Gambar 2.6) menunjukkan pola distribusi engagement, sementara scatter plot hubungan antara panjang komentar dan jumlah like (Gambar 2.7) menunjukkan bahwa komentar yang lebih panjang tidak selalu mendapatkan like lebih banyak, tetapi outlier tertentu bisa sangat populer.

LIST NAMA YOUTUBE CHANEL (Gambar 2.1)

	youtuber_name
1.	Dream
2.	DanTDM
3.	Technoblade
4.	PewDiePie
5.	Ludwig
6.	MrBeast
7.	jacksepticeye
8.	Markiplier
9.	TommyInnit
10.	Asmongold TV
11.	penguinz0

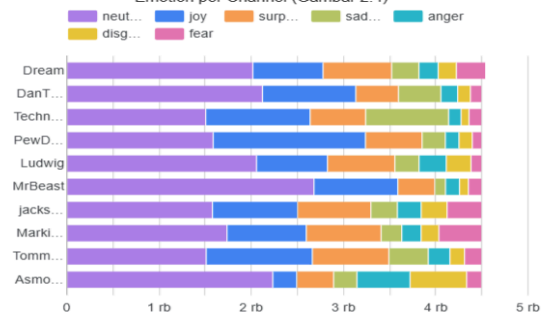
Sentiment per Channel (Gambar 2.3)

youtuber_name	sentiment	comment
Dream	positive	2.782
	negative	1.623
	neutral	137
PewDiePie	positive	3.037
	negative	1.357
	neutral	106
TommyInnit	positive	2.774
	negative	1.586
	neutral	140

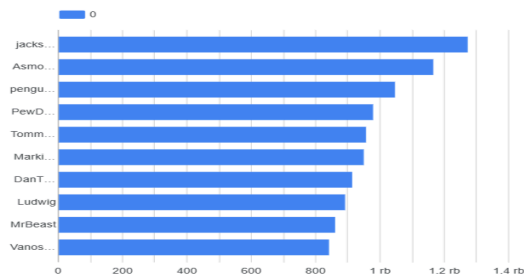
Total Like per Channel (Gambar 2.2)

	youtuber_name	like_co...
1.	stampylonghead	1.001.947
2.	penguinz0	5.648.944
3.	ohnepixel	1.229.194
4.	jacksepticeye	6.149.007
5.	VanossGaming	3.142.506
6.	TommyInnit	6.946.752
7.	TenZ	12.189
8.	Technoblade	6.899.792
9.	Shroud	2.993.258

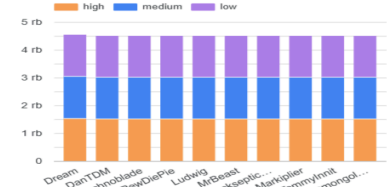
Emotion per Channel (Gambar 2.4)



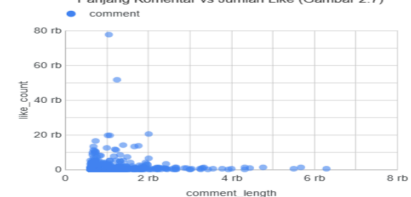
Jumlah Data dengan like_count = 0 (Gambar 2.5)



Like Count (Low, Medium, High) per Channel (Gambar 2.6)



Panjang Komentar vs Jumlah Like (Gambar 2.7)



BAB III

IMPLEMENTASI SISTEM

3.1 Frontend System

Frontend dari aplikasi YouTube Comment Analyzer dibangun menggunakan **Next.js 14** (React framework) dengan **TypeScript**, serta didukung oleh **Tailwind CSS** dan **Shadcn/ui** untuk antarmuka yang responsif dan modern. Aplikasi ini menyediakan tiga tab utama: **Analyze**, **History**, dan **Statistics**. Di tab **Analyze**, pengguna dapat memasukkan komentar melalui text area yang dilengkapi fitur validasi dan real-time feedback. Hasil klasifikasi emosi, sentimen, dan prediksi jumlah like ditampilkan secara interaktif dalam bentuk kartu animasi yang disertai emoji dan warna sesuai konteks. Riwayat analisis tersimpan otomatis di tab **History** dengan fitur manajemen seperti penghapusan data dan timestamp. Sementara itu, tab **Statistics** menyajikan dasbor visualisasi dalam bentuk grafik distribusi emosi, sentimen, dan like, serta rekap jumlah prediksi. UI dilengkapi dengan status koneksi model, animasi loading, dan error handling yang ramah pengguna, menciptakan pengalaman penggunaan yang intuitif dan menyenangkan.

3.2 Backend System

Backend dibangun menggunakan Flask sebagai server utama dan mengelola seluruh proses inferensi machine learning. Komentar dari frontend dikirim ke backend melalui REST API (`/predict`) yang memicu pipeline prediksi. Pipeline ini melibatkan dua model RoBERTa untuk klasifikasi sentimen dan emosi, serta model BERT untuk menghasilkan embedding. Embedding tersebut kemudian diproses oleh model XGBoost untuk memprediksi jumlah like. Semua model dimuat menggunakan pustaka transformers, PyTorch, dan joblib. Backend juga menyediakan endpoint tambahan seperti `/history`, `/stats`, dan `/debug` untuk mendukung fitur riwayat dan statistik. Arsitektur aplikasi mengadopsi pendekatan microservices, di mana frontend dan backend berjalan pada port terpisah (3000 dan 5000) namun terhubung secara asinkron melalui API. Struktur ini memungkinkan fleksibilitas dalam pengembangan dan skalabilitas sistem kedepannya.

3.3 Pelatihan Model dan Tuning Hyperparameter

Pelatihan model dilakukan dalam dua tahap utama, masing-masing dengan pendekatan tuning hyperparameter yang berbeda untuk tugas klasifikasi (emosi dan sentimen) serta regresi jumlah likes. Untuk tugas klasifikasi, digunakan model `RoBERTaForSequenceClassification` dari pustaka Huggingface Transformers. Pelatihan dilakukan menggunakan kelas `Trainer` dan `TrainingArguments`, dengan beberapa hyperparameter penting yang disetel secara manual berdasarkan praktik fine-tuning model BERT. Pengaturan pelatihan meliputi

- `num_train_epochs = 2` — jumlah epoch disesuaikan agar cukup untuk konvergensi tanpa overfitting
- `learning_rate = 2e-5` — nilai learning rate standar untuk fine-tuning RoBERTa
- `per_device_train_batch_size = 32` — menyesuaikan kapasitas GPU yang tersedia
- `eval_strategy = "epoch"` dan `load_best_model_at_end = True` — untuk memilih model terbaik berdasarkan metrik validasi
- `fp16 = True` — pelatihan dengan presisi campuran (jika menggunakan GPU)

Sementara itu, untuk tugas regresi jumlah like, digunakan model `XGBRegressor` yang dituning menggunakan metode `RandomizedSearchCV` dari scikit-learn. Proses ini dilakukan pada embedding yang telah dihasilkan sebelumnya dari model BERT, dan dilakukan dengan pencarian kombinasi hyperparameter berikut: embedding BERT tersebut digunakan sebagai

representasi fitur utama dari komentar karena kemampuannya menangkap konteks semantik secara mendalam, dan fitur ini kemudian diperkaya dengan hasil prediksi emotion serta sentiment yang diperoleh dari model RoBERTa, sehingga model XGBRegressor dapat memanfaatkan informasi konteks teks sekaligus nuansa emosional dan sentimen untuk meningkatkan akurasi prediksi jumlah like, dan dilakukan dengan pencarian kombinasi hyperparameter berikut:

- `n_estimators`, `learning_rate`, `num_leaves`, `max_depth`
- `reg_alpha`, `reg_lambda` (regularisasi L1 dan L2)
- `colsample_bytree` dan `subsample` (untuk mengontrol fraksi fitur dan sampel)

Pencarian dilakukan dengan:

- `cv = 5` (lima lipatan cross-validation)
- `n_iter = 50` (jumlah kombinasi hyperparameter yang diuji)
- Metrik evaluasi `neg_mean_absolute_error` sebagai acuan optimasi

Model terbaik kemudian disimpan sebagai file `xgboost_BERT_embeddings.pkl` menggunakan `joblib`. Dengan pendekatan ini, sistem memiliki dua model terlatih yang masing-masing telah disesuaikan untuk performa optimal baik dalam klasifikasi maupun prediksi numerik.

BAB IV

EVALUASI DAN HASIL

4.1 Akurasi Model

Model klasifikasi komentar yang dibangun menggunakan arsitektur RoBERTa menunjukkan performa yang sangat baik dalam mengelompokkan komentar YouTube bertema gaming berdasarkan sentimen dan emosi. Pada pengujian terhadap data uji, model klasifikasi sentimen berhasil mencapai:

- **Akurasi: 93.12%**
- **Precision: 92.0%**
- **Recall: 91.0%**
- **F1-Score: 92.0%**

Sementara itu, model klasifikasi emosi yang menangani tujuh kelas berbeda juga menunjukkan kinerja yang kompetitif dengan hasil evaluasi sebagai berikut:

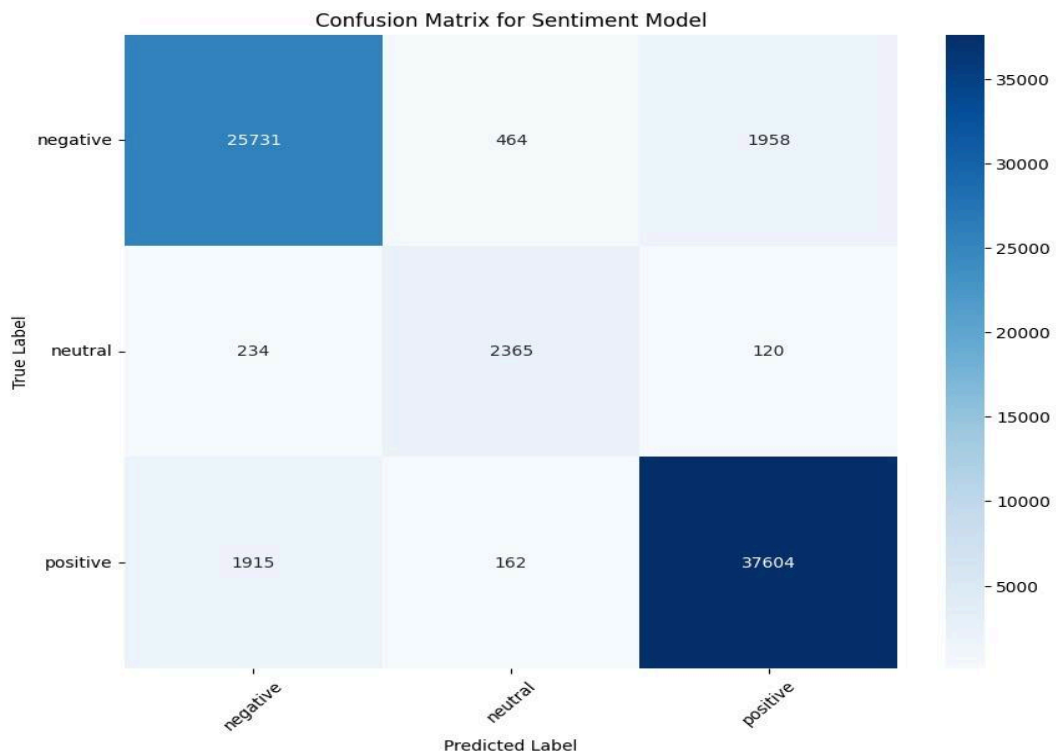
- **Akurasi: 87.58%**
- **Precision: 83.0%**
- **Recall: 85.0%**
- **F1-Score: 84.0%**

Hasil ini mengindikasikan bahwa model sentimen mampu membedakan komentar positif, netral, dan negatif dengan sangat akurat, sedangkan model emosi cukup andal dalam mengenali variasi ekspresi seperti joy, sadness, dan anger. Nilai F1-score yang tinggi pada kedua model menunjukkan keseimbangan yang baik antara presisi dan sensitivitas, menjadikannya efektif untuk memahami persepsi dan reaksi emosional pengguna secara otomatis dalam komunitas gaming online.

4.2 Confusion Matrix

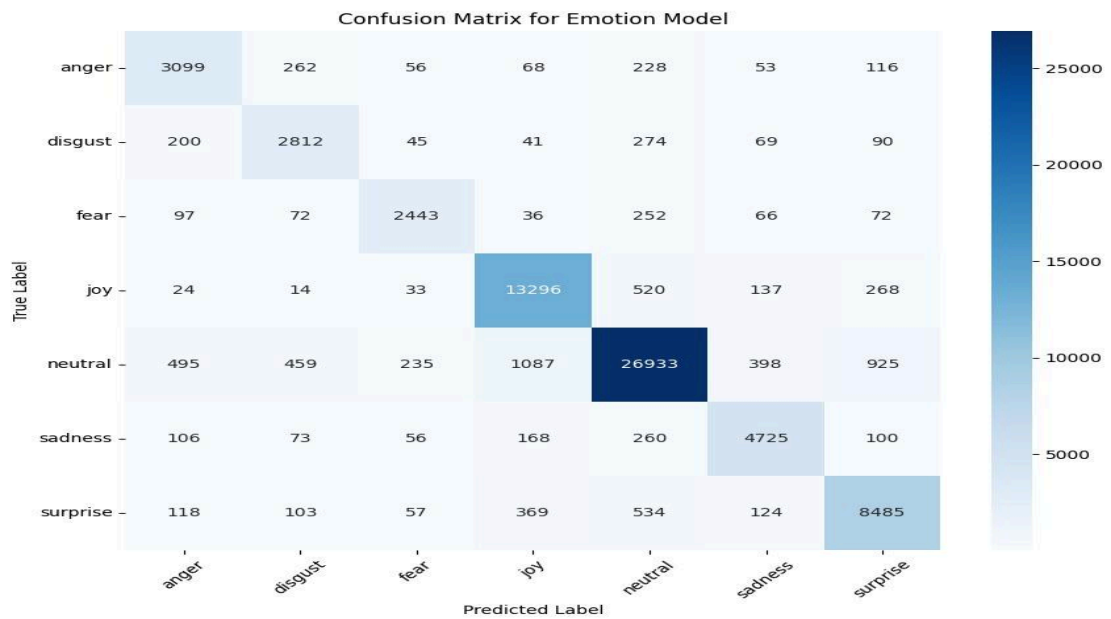
Evaluasi menggunakan confusion matrix pada data uji menunjukkan bahwa model mampu mengklasifikasikan komentar dengan akurasi yang baik, sebagaimana terlihat pada visualisasi heatmap berikut :

Pada **model sentimen** (Gambar 4.1), mayoritas komentar berhasil diklasifikasikan dengan benar, terutama pada kelas *positive* (37.604 benar dari 39.681). Kesalahan paling umum terjadi antara kelas *positive* dan *negative*, namun secara keseluruhan distribusi prediksi cukup akurat.



Gambar 4.1

Sementara itu, **model emosi** (Gambar 4.2) menunjukkan hasil yang baik pada kelas *joy* dan *neutral*, masing-masing dengan lebih dari 13.000 dan 26.000 prediksi benar. Beberapa kesalahan antar kelas, seperti antara *anger*, *disgust*, dan *fear*, masih terjadi tetapi berada dalam batas wajar mengingat kemiripan ekspresi emosi dalam teks.



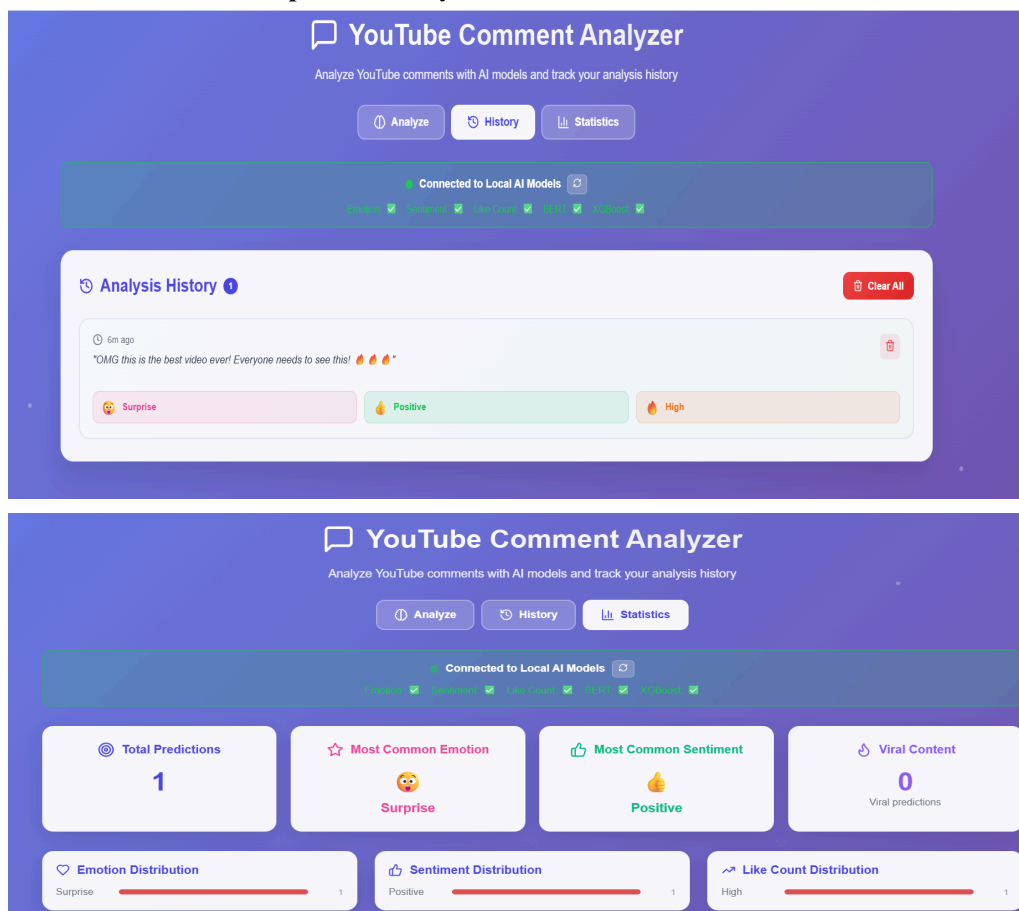
Gambar 4.2

4.3 Pengujian Aplikasi

Tampilan saat user ingin input komentar setelah itu ada 3 kategori yang keluar yaitu Emotion, Sentiment dan Like Count

The screenshot displays the 'YouTube Comment Analyzer' application interface. At the top, the title 'YouTube Comment Analyzer' is shown with the subtitle 'Analyze YouTube comments with AI models and track your analysis history'. Below this are three tabs: 'Analyze', 'History', and 'Statistics'. A status bar indicates 'Connected to Local AI Models' with checkmarks for Emotion, Sentiment, Like Count, BERT, and XGBERT. The 'Comment Input' section contains a text area with the comment 'OMG this is the best video ever! Everyone needs to see this!' and a character count '67/500'. Below the text area is an 'Analyze Comment' button and a 'Reset' button. The analysis results are displayed in three cards: 'Emotion' showing 'Surprise' (Detected emotion), 'Sentiment' showing 'Positive' (Detected sentiment), and 'Like Count' showing 'High' (101-1000 likes, Predicted engagement level).

Berikut ini adalah tampilan History dan Statistics



BAB V

PEMBAGIAN TUGAS

No	Nama (NIM)	Program Studi	Bidang Ilmu	Uraian Tugas
1	Alfito Faiz Rizqi 2702316724	Computer Science	Machine Learning Deployment	Mengumpulkan Data, mengerjakan frontend, integrasi API backend, pengujian sistem, penyusunan laporan dan ppt, dan visualisasi interaktif.
2	BRYAN CRISTHOPER CHANDRA PUTRA 2702272306	Computer Science	Machine Learning	Melatih dan mengevaluasi model, pelatihan model menggunakan data yang tersedia, validasi hasil menggunakan metrik evaluasi seperti akurasi, F1-score, atau AUC, serta melakukan tuning parameter untuk meningkatkan performa model
3	REUBEN ANSELMUS ADEL 2702314901	Computer Science	Machine Learning	Mengumpulkan data dan mem preprocess data, serta melakukan quality control

Tabel 5.1 Pembagian Tugas Kelompok

LAMPIRAN

Canva :

https://www.canva.com/design/DAGpl1_AUKQ/Fy6R76zzE0vKjQ8pxEM79Q/edit?utm_content=DAGpl1_AUKQ&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton

predict:

<https://colab.research.google.com/drive/1sR-IJZ-mKl59VF2JyXGVJb2hgclL1lOl?usp=sharing>

trainingnya :

<https://colab.research.google.com/drive/1PVCftQQqPKmQrrOS9my69fYKTHHZiWZz?usp=sharing>

Mengumpulkan Dataset :

<https://colab.research.google.com/drive/1uokj6OBbHPN9YJH7z79J6Ie4yZ1bNwf1>

Looker : <https://lookerstudio.google.com/reporting/83253f36-1ebc-43e4-8fd7-5d403262c643>