

**METODE KORELASI PEARSON UNTUK
SENTIMEN KOMENTAR TERHADAP PERFORMA
YOUTUBE JKT48 MENGGUNAKAN METODE *LSTM***

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ALIF IKHWAN HABIBI

21.11.3990

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**METODE KORELASI PEARSON UNTUK
SENTIMEN KOMENTAR TERHADAP PERFORMA
YOUTUBE JKT48 MENGGUNAKAN METODE *LSTM***

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

ALIF IKHWAN HABIBI

21.11.3990

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

**METODE KORELASI PEARSON UNTUK
SENTIMEN KOMENTAR TERHADAP PERFORMA
YOUTUBE JKT48 MENGGUNAKAN METODE *LSTM***

yang disusun dan diajukan oleh

Alif Ikhwan Habibi

21.11.3990

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 30 Juli 2025

Dosen Pembimbing,



Melany Mustika Dewi, S.Kom., M.Kom
NIK. 19030455

HALAMAN PENGESAHAN
SKRIPSI
METODE KORELASI PEARSON UNTUK
SENTIMEN KOMENTAR TERHADAP PERFORMA
YOUTUBE JKT48 MENGGUNAKAN METODE *LSTM*

yang disusun dan diajukan oleh

Alif Ikhwan Habibi

21.11.3990

Telah dipertahankan di depan Dewan Penguji
pada tanggal 30 Juli 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Anna Baita, S.Kom., M.Kom
NIK. 190302290

Rizqi Sukma Kharisma, S.Kom., M.Kom
NIK. 190302215

Melany Mustika Dewi, S.Kom., M.Kom
NIK. 190302455



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 30 Juli 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Alif Ikhwan Habibi
NIM : 21.11.3990

Menyatakan bahwa Skripsi dengan judul berikut:

"Metode Korelasi Pearson Untuk Sentimen Komentar Terhadap Performa Youtube Jkt48 Menggunakan Metode *Lstm*"

Dosen Pembimbing: Melany Mustika Dewi, S.Kom, M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 30 Juli 2025

Yang Menyatakan,



Alif Ikhwan Habibi

HALAMAN PERSEMBAHAN

Dengan rasa syukur yang mendalam kepada Tuhan Yang Maha Esa, saya mempersembahkan skripsi ini kepada pihak-pihak yang telah berkontribusi dalam perjalanan saya selama ini. Persembahan ini sebagai wujud rasa terima kasih atas dukungan, doa, serta semangat yang telah diberikan kepada saya. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Kepada Allah SWT, yang telah memberikan rahmat dan karunia-Nya sehingga saya dapat menyelesaikan tugas akhir ini.
2. Kepada Ayah, Ibu, dan Adik, yang selalu memberikan doa, dukungan, semangat, serta kasih sayang yang tiada henti dalam setiap langkah perjalanan hidupku.
3. Kepada Dosen pembimbing, yang telah sabar dalam memberikan bimbingan, arahan, motivasi dan ilmu yang sangat berarti selama proses penyusunan skripsi ini.
4. Kepada Teman-teman seperjuangan WDM Andri, Dimas, Dika, Rian, Rio, Hizqil, dan teman-teman kelas IF 03 yang senantiasa memberikan semangat dan bantuan, selama menjalani proses perkuliahan.
5. Diriku sendiri, yang telah berjuang dan pantang menyerah dalam menyelesaikan skripsi ini.

Semoga karya ini dapat memberikan manfaat bagi langkah besar berikutnya dan menjadi bagian kecil dari proses pembelajaran saya selama menempuh pendidikan.

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas limpahan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Metode Korelasi Pearson Untuk Sentimen Komentar Terhadap Performa Youtube Jkt48 Menggunakan Metode *Lstm*” sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta.

Penyusunan skripsi ini tidak lepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Melany Mustika Dewi, S.Kom, M.Kom., selaku dosen pembimbing yang telah memberikan arahan, masukan, dan dukungan selama proses penyusunan skripsi ini.
2. Tim dosen penguji yang telah meluangkan waktu, memberikan kritik, saran, dan penilaian yang membangun demi penyempurnaan skripsi ini.
3. Orang tua dan keluarga penulis yang selalu memberikan doa, semangat, dan dukungan moral maupun material kepada penulis dalam menyelesaikan pendidikan hingga tahap ini.
4. Pihak-pihak terkait yang selalu memberikan dukungan dan bantuan selama penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan demi penyempurnaan karya ini di masa yang akan datang. Semoga skripsi ini dapat memberikan manfaat bagi semua pihak yang membacanya.

Yogyakarta, 30 Juli 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
<i>ABSTRACT</i>	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.2 Batasan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6

2.2	Dasar Teori	17
2.2.1	Analisis Sentimen	17
2.2.2	<i>Natural Language Processing (NLP)</i>	17
2.2.3	<i>Youtube API</i>	18
2.2.4	<i>Word Embedding</i>	18
2.2.5	<i>SMOTE (Synthetic Minority Over-sampling Technique)</i>	18
2.2.6	<i>LSTM</i>	19
2.2.7	<i>Confusion Matrix</i>	21
2.2.8	<i>Pearson Correlation</i>	22
BAB III	METODE PENELITIAN	24
3.1	Objek Penelitian	24
3.2	Alur Penelitian	25
3.2.1	<i>Scraping Data</i>	25
3.2.2	<i>Pre-Processing</i>	26
3.2.3	<i>Labeling</i>	27
3.2.4	<i>EDA</i>	27
3.2.5	<i>Pre-Modeling</i>	28
3.2.6	<i>Modeling LSTM</i>	29
3.2.7	<i>Model Evaluation</i>	29
3.2.8	<i>Correlation</i>	30
3.3	Alat dan Bahan	30
3.3.1	Bahan Penelitian	30
3.3.2	Alat Penelitian	30
BAB IV	HASIL DAN PEMBAHASAN	32
4.1	<i>Scraping Data Youtube</i>	32
4.2	<i>Preprocessing Data</i>	33
4.2.1	<i>Text Cleaning</i>	33
4.2.2	<i>Case Folding</i>	34
4.2.3	<i>Normalized</i>	34

4.2.4	<i>Stopword Removal</i>	35
4.2.5	<i>Tokenized</i>	35
4.2.6	<i>Stemming</i>	36
4.3	<i>Labeling</i>	36
4.4	<i>EDA</i>	37
4.4.1	Distribusi Sentimen	37
4.4.2	<i>Word cloud</i> Keseluruhan Data	38
4.4.3	<i>Word Cloud</i> Sentimen Positif dan Negatif	39
4.5	<i>Premodeling</i>	40
4.5.1	<i>Splitting Dataset</i>	41
4.5.2	<i>Tokenizing & Padding</i>	41
4.5.3	<i>Word Embedding</i>	43
4.5.4	<i>SMOTE</i>	44
4.6	<i>Modeling LSTM</i>	46
4.7	<i>Model Evaluation</i>	47
4.8	<i>Correlation</i>	49
4.8.1	Perhitungan Persentase Sentimen Per Video	49
4.8.2	Korelasi pearson	50
BAB V PENUTUP		52
5.1	Kesimpulan	52
5.2	Saran	53
REFERENSI		54
LAMPIRAN		57

DAFTAR TABEL

Tabel 2. 1 Keaslian Penelitian	9
Tabel 2. 2 <i>Confusion Matrix</i>	21
Tabel 3. 1 Bahan Penelitian	31
Tabel 4. 1 informasi video Youtube	32
Tabel 4. 2 Hasil <i>Scraping</i> Dataset Youtube	33
Tabel 4. 3 Data Sebelum <i>Text_Cleaning</i>	33
Tabel 4. 4 Data Sesudah <i>Text_Cleaning</i>	33
Tabel 4. 5 Hasil <i>Case Folding</i>	34
Tabel 4. 6 Hasil <i>Normalized</i>	34
Tabel 4. 7 Hasil <i>Stopword Removal</i>	35
Tabel 4. 8 Hasil <i>Tokenized</i>	35
Tabel 4. 9 Hasil <i>Stemming</i>	36
Tabel 4. 10 Hasil <i>Labeling</i>	36
Tabel 4. 11 Hasil <i>Tokenizing & Padding</i>	42
Tabel 4. 12 Hasil <i>Word Embedding</i>	43
Tabel 4. 13 Hasil <i>Training & Validation LSTM</i>	46
Tabel 4. 14 Evaluasi model <i>LSTM</i>	48
Tabel 4. 15 perhitungan presentase per video	50
Tabel 4. 16 Evaluasi korelasi pearson	51

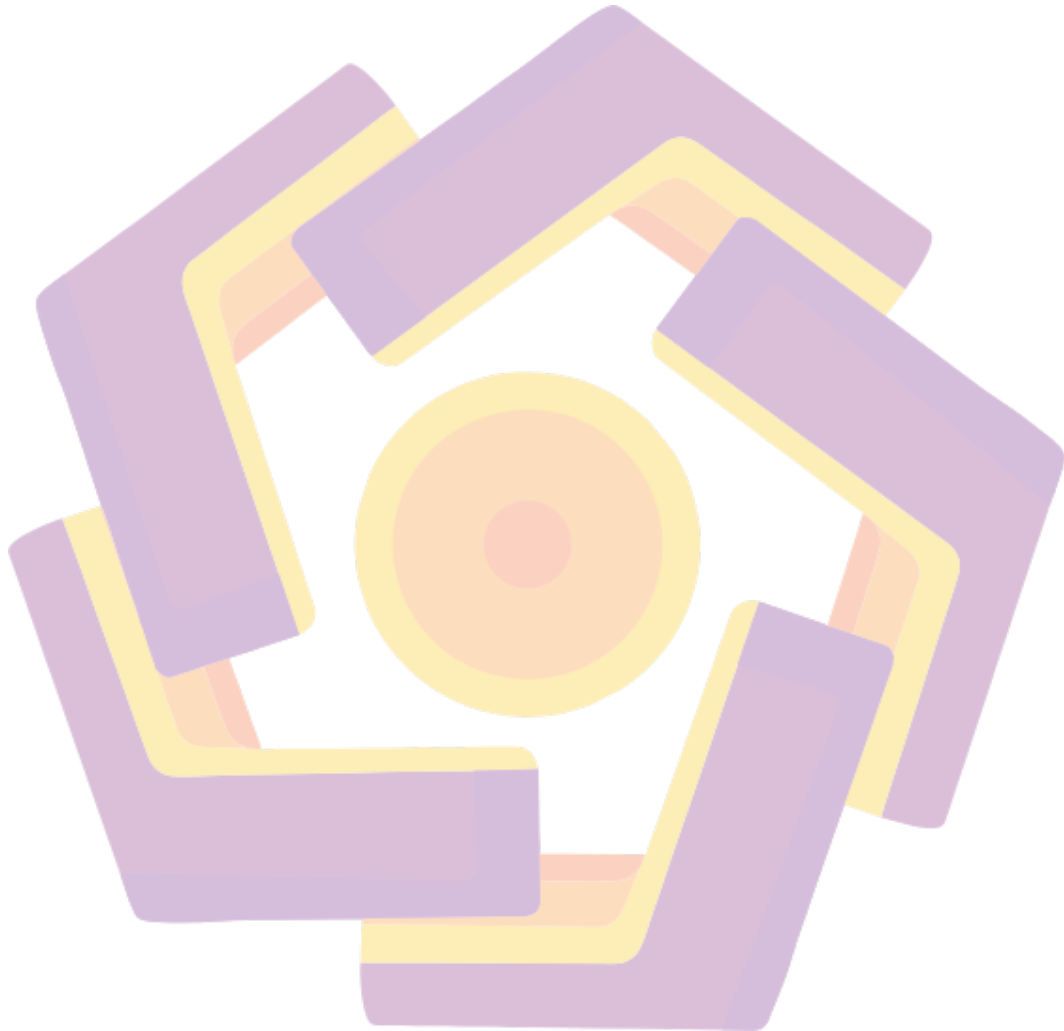
DAFTAR GAMBAR

Gambar 2. 1 <i>LSTM Architecture</i>	19
Gambar 3. 1 Alur Penelitian	25
Gambar 4. 1 Gambar Distribusi Sentimen	Error! Bookmark not defined.
Gambar 4. 2 <i>Word Cloud</i> Keseluruhan Data	39
Gambar 4. 3 <i>Word Cloud</i> Sentimen Positif	40
Gambar 4. 4 <i>Word Cloud</i> Sentimen Negatif	40
Gambar 4. 5 Split Data <i>Train & Test</i>	41
Gambar 4. 6 Distribusi data sebelum <i>SMOTE</i>	45
Gambar 4. 7 Distribusi data setelah <i>SMOTE</i>	45
Gambar 4. 8 Tingkat Akurasi dari <i>Epoch LSTM</i>	47
Gambar 4. 9 Tingkat <i>Loss</i> dari <i>Epoch LSTM</i>	47
Gambar 4. 10 <i>Confusion Matrix LSTM</i>	49



DAFTAR LAMPIRAN

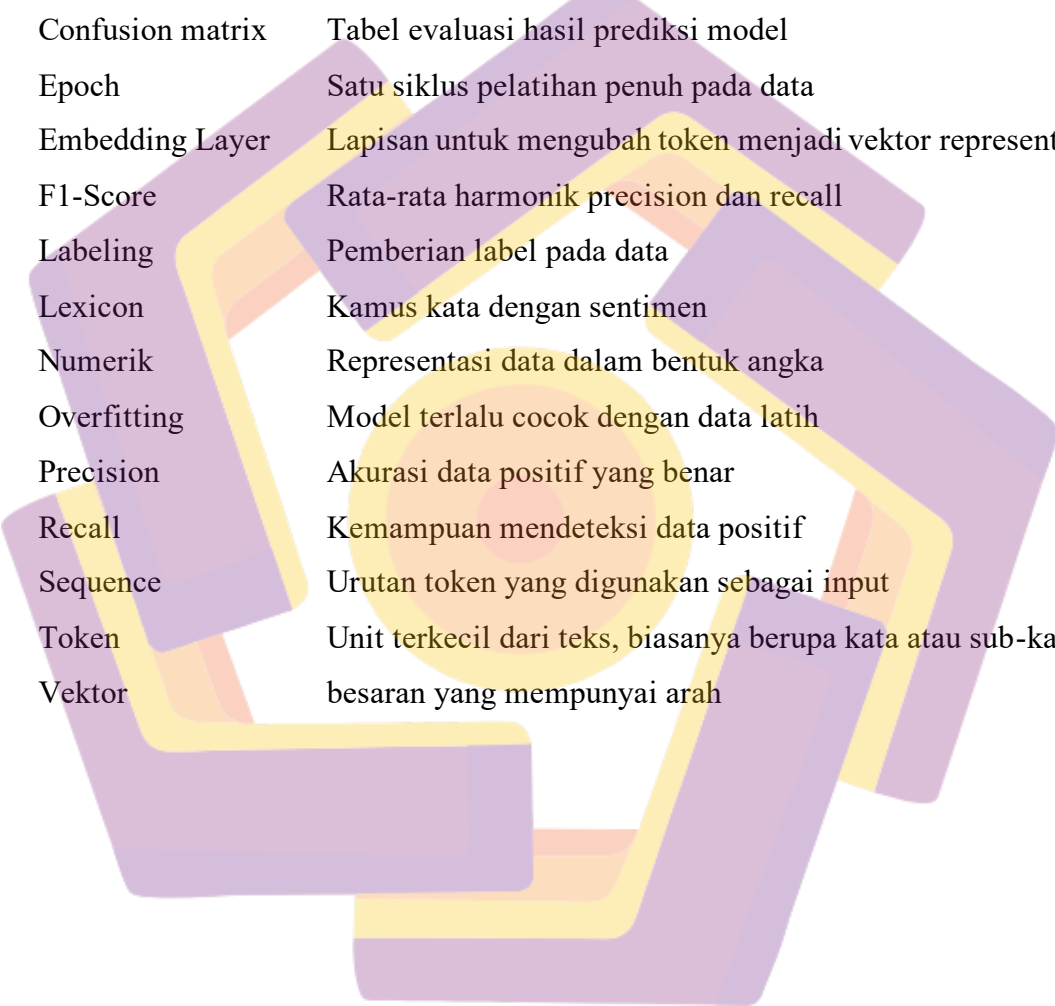
Lampiran 1 Kode <i>Scraping</i> Data	57
Lampiran 2 Kode <i>Pre-Processing</i>	59
Lampiran 3 Kode <i>Labeling</i>	60
Lampiran 4 Kode <i>EDA</i>	62
Lampiran 5 Kode <i>Premodeling</i>	65
Lampiran 6 Kode <i>Modeling LSTM</i>	70
Lampiran 7 Kode <i>Correlation</i>	71



DAFTAR LAMBANG DAN SINGKATAN

σ	Fungsi aktivasi sigmoid
Tanh	Fungsi aktivasi hyperbolic tangent
f_t	Nilai forget gate pada waktu ke-t
i_t	Nilai input gate pada waktu ke-t
o_t	Nilai output gate pada waktu ke-t
$\tilde{c}_t$	Informasi baru untuk cell state pada waktu ke-t
c_t	Cell state pada waktu ke-t
h_t	Hidden state pada waktu ke-t
$h_{\{t-1\}}$	Hidden state dari waktu sebelumnya
X_t	Input pada waktu ke-t
W_f, W_i, W_o	Bobot untuk forget gate, input gate, output gate
b_f, b_i, b_o	Bias untuk forget gate, input gate, output gate
AI	Artificial Intelligence
API	Application Programming Interface
CSV	Comma-Separated Values
EDA	Exploratory Data Analysis
FN	False Negative
FP	False Positive
LSTM	Long Short-Term Memory
NLP	Natural Language Processing
OOV	Out of Vocabulary
RNN	Recurrent Neural Network
SMOTE	Synthetic Minority Oversampling Technique
TP	True Positive
TN	True Negative
TF-IDF	Term Frequency–Inverse Document Frequency

DAFTAR ISTILAH



Accuracy	Persentase prediksi yang benar
Automatic Labeling	Pelabelan otomatis tanpa manusia.
Bias	Nilai tambahan dalam jaringan saraf
Batch Size	Jumlah data yang diproses dalam satu iterasi pelatihan
Confusion matrix	Tabel evaluasi hasil prediksi model
Epoch	Satu siklus pelatihan penuh pada data
Embedding Layer	Lapisan untuk mengubah token menjadi vektor representasi
F1-Score	Rata-rata harmonik precision dan recall
Labeling	Pemberian label pada data
Lexicon	Kamus kata dengan sentimen
Numerik	Representasi data dalam bentuk angka
Overfitting	Model terlalu cocok dengan data latih
Precision	Akurasi data positif yang benar
Recall	Kemampuan mendeteksi data positif
Sequence	Urutan token yang digunakan sebagai input
Token	Unit terkecil dari teks, biasanya berupa kata atau sub-kata
Vektor	besaran yang mempunyai arah

INTISARI

Penelitian ini bertujuan untuk menganalisis korelasi antara sentimen komentar pengguna terhadap video Youtube JKT48 dan performa video tersebut, yang diukur melalui jumlah *like* dan *view*. Sentimen komentar dianalisis menggunakan model Long Short-Term Memory (LSTM), dengan data pelatihan yang diperoleh dari hasil pelabelan otomatis dan teknik oversampling SMOTE untuk mengatasi ketidakseimbangan kelas. Model LSTM yang dibangun menunjukkan performa tinggi dengan akurasi sebesar 94%, precision dan recall di atas 93% untuk kedua kelas sentimen (positif dan negatif), serta f1-score rata-rata sebesar 0.94. Hasil klasifikasi menunjukkan bahwa komentar negatif lebih dominan (55,3%) dibanding komentar positif (44,7%). Analisis korelasi Pearson menunjukkan hubungan yang sangat kuat antara sentimen komentar dan performa video. Nilai korelasi antara komentar positif dan jumlah *likes* sebesar $r = 0.981$, serta terhadap jumlah *views* sebesar $r = 0.846$. Sementara itu, komentar negatif memiliki korelasi sebesar $r = 0.995$ terhadap jumlah *likes* dan $r = 0.876$ terhadap jumlah *views*. Hal ini menunjukkan bahwa baik komentar positif maupun negatif berkontribusi signifikan terhadap peningkatan engagement video. Penelitian ini membuktikan bahwa interaksi melalui komentar berperan penting dalam meningkatkan popularitas video. Temuan ini dapat menjadi pertimbangan dalam pengembangan strategi konten digital, khususnya dalam mengelola interaksi pengguna di platform media sosial seperti Youtube.

Kata kunci: Analisis Sentimen, Youtube, JKT48, LSTM, Korelasi Pearson, Komentar, Engagement

ABSTRACT

This study aims to analyze the correlation between user comment sentiment on JKT48's Youtube videos and the videos' performance, as measured by the number of likes and views. Sentiment analysis was conducted using a Long Short-Term Memory (LSTM) model, trained on labeled data with class imbalance addressed using the SMOTE oversampling technique. The resulting LSTM model demonstrated strong performance, achieving 94% accuracy, with both precision and recall exceeding 93% for positive and negative sentiment classes, and an average f1-score of 0.94. The sentiment classification revealed that negative comments were more dominant (55.3%) compared to positive ones (44.7%). Pearson correlation analysis showed a very strong relationship between sentiment and video performance. The correlation between positive comments and likes was $r = 0.981$, and with views was $r = 0.846$. Meanwhile, negative comments showed a correlation of $r = 0.995$ with likes and $r = 0.876$ with views. These results indicate that both positive and negative sentiments significantly contribute to increasing user engagement. This research demonstrates that user interaction through comments plays a crucial role in enhancing video popularity. These insights can inform digital content strategies, particularly in managing user engagement on social media platforms like Youtube.

Keyword: *Sentiment Analysis, Youtube, JKT48, LSTM, Pearson Correlation, Comments, Engagement*