

Full Paper

## Song Lyric Meaning Generator Using Transformer (Case Study: Drake's Song Lyrics)

**Tubagus Alwasi<sup>1\*</sup>**,Universitas Islam Sultan Agung, Kota Semarang,  
Indonesia,[tubagusalesi@gmail.com](mailto:tubagusalesi@gmail.com)**Sam Farisa Chaerul Haviana<sup>2</sup>**,Universitas Islam Sultan Agung, Kota Semarang,  
Indonesia,[sam@unissula.ac.id](mailto:sam@unissula.ac.id)

\*Corresponding Author

### ABSTRACT


The limited availability of structured and credible sources that interpret Drake's song lyrics makes it difficult for listeners to fully grasp the meaning and emotional depth within his music. This study aims to develop a web-based lyric meaning generator capable of automatically interpreting Drake's lyrics using the Transformer architecture. The system employs the LLaMA 3 model, which is fine-tuned through Low-Rank Adaptation (LoRA) to improve training efficiency while maintaining high semantic accuracy. The curated dataset consists of Drake's song lyrics, their corresponding interpretations, and metadata such as album and reference sources. Data preprocessing and supervised fine-tuning were conducted using the Hugging Face framework in Google Colab. A gradio-based web application was implemented with a Retrieval-Augmented Generation (RAG) mechanism to enhance contextual relevance during inference. Evaluation metrics, including Semantic Similarity and ROUGE-L, were applied to measure the model's ability to produce coherent and contextually aligned interpretations. The results demonstrate that the fine-tuned LLaMA 3 model effectively generates accurate lyric meanings while reducing computational cost. Overall, this study highlights the potential of Transformer-based models to bridge the gap between music and natural language understanding, particularly in analyzing metaphorical and emotion-rich song lyrics.

### KEYWORDS

Song lyric meaning generator; Transformer; LLaMA 3; Drake; NLP; LoRA; Fine-tuning; RAG

Alwasi<sup>1</sup>, T. Haviana. S. F. C. (2025). Song Lyric Meaning Generator Using Transformer (Case Study: Drake's Song Lyrics). *jasmed*, 3(1), pp. 30-40. <https://doi.org/10.20895/jasmed.v3i1.10132>

Article Submitted 14/11/2025. Revision uploaded 23/12/2025. Accepted 23/12/2025.

© 2025 by the authors of this article. Published under CC-BY 

## 1. INTRODUCTION

Perkembangan teknologi kecerdasan buatan (Artificial Intelligence atau AI) telah membawa transformasi besar dalam berbagai bidang, termasuk pemrosesan bahasa alami (Natural Language Processing atau NLP). NLP berfokus pada bagaimana komputer dapat memahami, menginterpretasikan, dan menghasilkan bahasa manusia secara bermakna. Salah satu tonggak penting dalam perkembangan NLP adalah munculnya arsitektur Transformer, yang diperkenalkan oleh Vaswani et al. [1]. Model ini mengandalkan mekanisme self-attention untuk mempelajari hubungan antar-kata dalam konteks yang luas, menggantikan pendekatan berbasis urutan seperti RNN dan LSTM. Keunggulan Transformer terletak pada kemampuannya untuk memproses teks secara paralel, meningkatkan efisiensi dan akurasi dalam berbagai tugas NLP seperti text generation, translation, dan summarization. Perkembangan ini memicu munculnya berbagai model bahasa besar seperti GPT, BERT, T5, dan LLaMA yang menunjukkan peningkatan signifikan dalam kemampuan pemahaman dan generasi bahasa alami [2], [3], [4], [5], [6], [7], [8].

Dalam konteks musik, lirik lagu merupakan salah satu bentuk ekspresi linguistik yang kompleks dan sarat makna. Lirik tidak hanya menyampaikan pesan eksplisit, tetapi juga makna tersirat melalui metafora, simbolisme, serta konteks budaya tertentu. Namun, interpretasi makna lirik sering kali bersifat subjektif, tergantung pada pengalaman dan pemahaman individu terhadap konteks lagu. Salah satu artis yang dikenal dengan kompleksitas liriknya adalah Drake, seorang musisi asal Kanada yang banyak menggabungkan unsur rap, R&B, dan pop. Lirik-liriknya sering kali berisi refleksi emosional, kritik sosial, serta narasi personal yang memerlukan interpretasi mendalam untuk dipahami secara utuh [3], [9], [10].

Permasalahan yang dihadapi penggemar maupun peneliti musik adalah keterbatasan sumber yang menyediakan penjelasan atau makna lirik secara sistematis. Sumber daring seperti Genius.com memang menyediakan interpretasi, tetapi sebagian besar ditulis oleh pengguna dan tidak selalu konsisten secara semantik. Kondisi ini menimbulkan kebutuhan akan sistem otomatis yang mampu menghasilkan makna lirik lagu secara kontekstual dan koheren. Di sinilah peran model bahasa besar (Large Language Model atau LLM) menjadi relevan. Model seperti GPT, BERT, T5, dan LLaMA telah menunjukkan kemampuan luar biasa dalam memahami hubungan semantik antar-kata serta menghasilkan teks dengan struktur yang logis [3], [6], [7], [11]. Selain itu, pendekatan Retrieval-Augmented Generation (RAG) terbukti meningkatkan akurasi generasi teks karena memungkinkan model mengakses pengetahuan eksternal sebelum menghasilkan output [12].

Dalam penelitian ini, dikembangkan sebuah sistem Generator Makna Lirik Lagu berbasis Transformer dengan studi kasus pada lirik lagu Drake. Sistem ini menggunakan model LLaMA 3, sebuah model bahasa besar yang dikembangkan oleh Meta, yang kemudian di-fine-tune menggunakan teknik Low-Rank Adaptation (LoRA) untuk menyesuaikan parameter model terhadap dataset khusus tanpa memerlukan pelatihan ulang penuh. LoRA dikenal efisien karena hanya memperbarui sebagian kecil bobot pada lapisan attention, sehingga menghemat sumber daya komputasi hingga 90% dibanding pelatihan konvensional [13], [14]. Pendekatan ini sangat ideal untuk penelitian akademik dengan keterbatasan perangkat keras seperti GPU di Google Colab.

Selain itu, sistem ini dikembangkan menjadi aplikasi berbasis web menggunakan Streamlit, agar hasil penelitian dapat diakses secara interaktif oleh pengguna. Arsitektur sistem mengikuti pendekatan Retrieval-Augmented Generation (RAG) yang menggabungkan proses pencarian konteks (retrieval) dengan generasi teks (generation). Dengan demikian, sistem tidak hanya menghasilkan teks berdasarkan model, tetapi juga mempertimbangkan referensi relevan dari dataset saat memproses input lirik pengguna [12].

Beberapa penelitian terdahulu telah membahas pemanfaatan Transformer dalam konteks musik, namun fokusnya lebih pada klasifikasi emosi atau pembuatan lirik baru. Ventura dan Toker (2022) memperkenalkan TRBLLmaker, model Transformer yang digunakan untuk menghasilkan lirik secara otomatis, bukan untuk interpretasi makna [15]. Penelitian lain oleh Agrawal et al. (2021) menyoroti penggunaan Transformer dalam mengenali emosi dari lirik lagu, tetapi tidak membahas aspek semantik yang lebih dalam seperti analisis makna atau pesan di balik teks [16]. Selain itu, beberapa penelitian mengenai analisis semantik lirik menunjukkan bahwa lirik musik memuat struktur metaforis dan simbolis yang kompleks sehingga memerlukan pendekatan NLP berbasis deep learning untuk dianalisis secara akurat [9], [17]. Oleh karena itu, penelitian ini memiliki nilai kebaruan (novelty) karena berfokus pada analisis makna lirik, bukan sekadar generasi teks atau klasifikasi emosi.

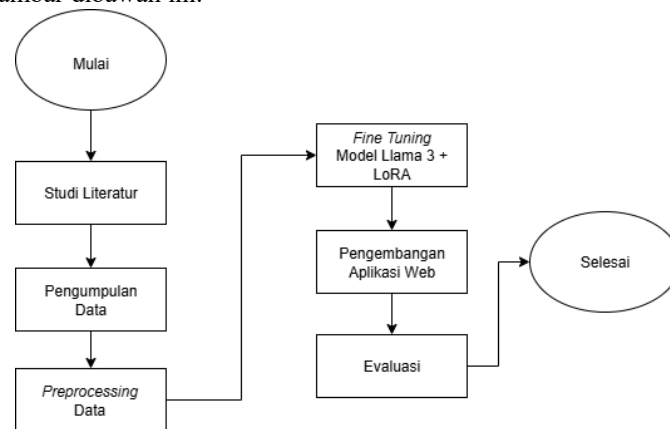
Secara ilmiah, penelitian ini berkontribusi dalam memperluas penerapan model Transformer, khususnya LLaMA 3, dalam bidang analisis teks non-formal seperti lirik lagu. Penelitian ini juga menunjukkan bahwa dengan strategi fine-tuning menggunakan LoRA, performa model tetap tinggi meskipun dijalankan dengan sumber daya

terbatas. Secara praktis, sistem yang dikembangkan dapat digunakan oleh pendengar musik, kritikus, atau peneliti linguistik untuk memahami makna lirik secara lebih objektif dan konsisten.

Berdasarkan uraian di atas, dapat disimpulkan bahwa penelitian ini dilakukan untuk mengembangkan sistem Generator Makna Lirik Lagu Drake berbasis Transformer (LLaMA 3) yang di-fine-tune menggunakan LoRA dan diimplementasikan melalui aplikasi web Streamlit. Tujuan akhirnya adalah menghasilkan model yang mampu memberikan interpretasi makna lirik secara otomatis, kontekstual, dan dapat diakses publik melalui antarmuka web yang sederhana.

## 2. METHODS

Metode penelitian ini menggunakan pendekatan eksperimental kuantitatif dengan fokus pada pengembangan dan pengujian model bahasa berbasis *Transformer* untuk menghasilkan makna lirik lagu secara otomatis. Alur tahapan penelitian dapat dilihat pada gambar dibawah ini.



**Gambar 1** Flowchart Penelitian.

### A. Studi Literatur

Tahap pertama yaitu studi literatur, dilakukan untuk memahami konsep dasar *Natural Language Processing* (NLP), arsitektur *Transformer*, dan metode *fine-tuning* efisien seperti *Low-Rank Adaptation* (LoRA). Referensi utama yang digunakan meliputi penelitian Vaswani et al. [1] tentang *Transformer architecture*, Zhang et al. [3] terkait LLaMA-Adapter, dan Wu et al. [4] mengenai pengembangan LLaMA-Pro. Studi ini juga meninjau penelitian terdahulu tentang pemrosesan teks musik dan analisis lirik lagu, seperti yang dilakukan oleh Agrawal et al. [2] dan Ventura & Toker [5]. Hasil tinjauan menunjukkan adanya celah penelitian pada aspek interpretasi makna lirik, bukan sekadar klasifikasi emosi atau generasi teks.

### B. Pengumpulan Data

*Dataset* yang digunakan dalam penelitian ini dikumpulkan dari sumber daring terpercaya seperti Genius.com dan Pitchfork, yang memuat lirik lagu-lagu Drake beserta makna interpretatifnya. *Dataset* terdiri atas 514 entri yang mencakup judul lagu, album, potongan lirik, makna lirik, dan sumber referensi. Data disusun dalam format JSONL (*JSON Lines*) untuk memudahkan pemrosesan dan kompatibilitas dengan pustaka *Hugging Face Datasets*. Tujuan pengumpulan data ini adalah membangun korpus teks yang representatif dan berkualitas untuk pelatihan model bahasa.

### C. Pra-pemrosesan Data

Tahap ini melibatkan proses pembersihan data (*data cleaning*) untuk memastikan kualitas *input* yang optimal sebelum digunakan dalam pelatihan model. Proses ini mencakup penghapusan karakter-karakter khusus seperti tanda baca dan emoji yang tidak memiliki makna semantik, normalisasi teks ke dalam bentuk huruf kecil (*lowercasing*) agar konsisten, serta penghapusan data ganda atau duplikasi entri yang dapat memengaruhi akurasi pelatihan. Selain itu, dilakukan juga penyesuaian format agar setiap baris data hanya memuat pasangan *input-output*, di mana bagian lirik lagu digunakan sebagai *input* dan maknanya berfungsi sebagai target *output*. Setelah seluruh data melalui tahap pembersihan dan normalisasi, dilakukan proses tokenisasi menggunakan *AutoTokenizer* dari model *meta-llama/Llama-3.2-1B-Instruct*. Tokenisasi ini berperan penting dalam mengubah teks menjadi representasi numerik agar dapat diproses oleh model *Transformer* secara efektif, serta menjaga keterhubungan semantik antar kata di dalam setiap konteks kalimat.

#### D. Fine-tuning Model LLaMa 3 + LoRA

Model LLaMA 3 yang telah dilatih sebelumnya (*pre-trained*) disesuaikan (*fine-tuned*) menggunakan teknik LoRA (*Low-Rank Adaptation*) agar efisien dalam penggunaan sumber daya. Pelatihan dilakukan menggunakan dataset yang telah diproses untuk menghubungkan lirik lagu dengan maknanya.

#### E. Pengembangan Aplikasi Web Menggunakan Gradio di Google Colab

Tahap pengembangan aplikasi web difokuskan pada implementasi antarmuka yang responsif dan efisien. Aplikasi ini dikembangkan menggunakan *framework* Gradio, yang menyediakan antarmuka pengguna (UI) yang sederhana dan mendukung integrasi cepat. Untuk mengatasi kendala komputasi dan memastikan performa inferensi yang optimal, aplikasi ini dijalankan di lingkungan *cloud* Google Colab yang menyediakan akses akselerasi GPU. Pendekatan ini berhasil mereduksi waktu pemrosesan makna lirik menjadi kurang dari 60 detik. Sistem ini dirancang dengan arsitektur *Retrieval-Augmented Generation* (RAG). Alur kerjanya dimulai saat pengguna memasukkan lirik; sistem kemudian menjalankan fase *retrieval* untuk mengambil konteks lagu yang relevan, yang kemudian digunakan oleh model LLaMA 3 yang telah di-*fine-tune* untuk menghasilkan interpretasi makna secara otomatis.

#### F. Evaluasi Sistem

Evaluasi dilakukan untuk mengukur sejauh mana model LLaMA 3 yang telah di-*fine-tune* mampu menghasilkan interpretasi makna lirik lagu yang memiliki kemiripan dengan makna referensi yang tersedia dalam *dataset*. Karena sistem ini bersifat generatif dan *output* yang dihasilkan berupa teks naratif bebas struktur (tidak berbentuk label klasifikasi), maka metrik evaluasi yang digunakan harus mampu menilai kesamaan makna, bukan hanya kemiripan struktur kata.

Metrik pertama berfokus pada metode evaluasi berbasis *semantic similarity*, yaitu pengukuran kesamaan makna antar dua teks dengan pendekatan vektor. Evaluasi dilakukan menggunakan model *pre-trained* dari *library sentence-transformers*, yaitu all-MiniLM-L6-v2, yang mampu mengubah teks (kalimat atau paragraf) menjadi representasi vektor berdimensi tinggi (*embedding*). Dua teks yang memiliki makna yang sama atau mirip akan menghasilkan vektor yang saling mendekati dalam ruang vektor.

Perhitungan *similarity* dilakukan dengan rumus *Cosine Similarity*, yang secara matematis dituliskan sebagai berikut:

$$\text{Similarity}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} \quad (1)$$

Keterangan:

$A$  dan  $B$  adalah vektor *embedding* dari dua teks yang dibandingkan, yaitu makna hasil prediksi model dan makna referensi dari dataset.  $A \cdot B$  adalah dot *product* antara kedua vektor.  $\|A\|$  dan  $\|B\|$  adalah panjang (norma) dari masing-masing vektor. Nilai *cosine similarity* berada dalam rentang 0 hingga 1, dimana: Nilai 0 menunjukkan bahwa dua teks tidak memiliki kesamaan makna. Nilai 1 menunjukkan bahwa dua teks sangat mirip atau identik secara semantik.

Sebagai pelengkap, metrik kedua yaitu ROUGE-L (*Longest Common Subsequence*) digunakan untuk mengevaluasi kesamaan struktural dan leksikal. ROUGE-L menilai kualitas generasi teks dengan mengukur panjang urutan kata terpanjang (LCS) yang sama antara teks yang dihasilkan (kandidat) dan teks referensi. Skor ini membantu memastikan bahwa model tidak hanya menangkap makna yang benar, tetapi juga mampu menyusun kalimat dengan pilihan kata dan struktur yang mirip dengan referensi.

Dengan demikian, penggunaan kombinasi *semantic similarity* dan ROUGE-L ini akan mencerminkan seberapa baik model dapat memahami, menafsirkan, dan menyajikan kembali lirik lagu dengan konteks yang mendekati makna aslinya.

## 3. RESULT

### A. Pengumpulan Data

Pada tahap pengumpulan data, penelitian ini menggunakan data lirik lagu dari artis hip-hop Drake, yang mencakup judul lagu, nama album, lirik lagu, interpretasi makna lirik, dan sumber referensi makna. Data dikumpulkan dari berbagai situs terpercaya seperti Genius.com, Pitchfork, wawancara resmi dengan Drake, serta media musik lainnya. Dataset ini mencakup lagu-lagu dari album Drake, mulai dari *More Life* hingga *Dark Lane Demo Tapes*. Total data yang terkumpul adalah 514 baris data, dengan setiap baris merepresentasikan bait lirik beserta maknanya. Berikut data dari *dataset* yang terdapat pada Tabel 1.

**Tabel 1** Dataset Lirik Lagu

album	song_title	lyric_bait	meaning_bait	Source
Dark Lane Demo Tapes	Deep Pockets	Dressed in fatigues, I rep the East with my trustees Smokin' on crushed leave ...	lirik ini menggambarkan perjalanan Drake dari masa kecil yang penuh tantangan hingga menjadi sosok ...	LyricsMeanings.com
Dark Lane Demo Tapes	When To Say When	Yeah Somebody, um Somebody hit me today and they were like "Nah, just make sure, ...	Bait ini mencerminkan sikap dermawan dan keterbukaan Drake terhadap orang-orang di sekitarnya. Ia menekankan bahwa berkat yang ...	Genius.com

Pada Tabel 1 merupakan tampilan dari dataset yang telah terkumpul diorganisir dalam format tabel. Tabel ini menampilkan struktur *dataset* yang digunakan, yang memuat informasi mengenai nama album (*album*), judul lagu (*song\_title*), potongan lirik atau frasa kunci yang akan diinterpretasikan maknanya (*lyric\_bait*), interpretasi makna dari potongan lirik tersebut (*meaning\_bait*), dan sumber asli dari lirik atau interpretasi makna (*source*). Sebagai contoh, pada baris pertama terlihat lirik "*Dressed in fatigues, I rep the East with my trust*" dari album "*Dark Lane Demo Tapes*" dan lagu "*Deep Pockets*", yang memiliki interpretasi makna "Drake ingin menggambarkan perjalanan" dengan sumber dari "LyricsMeanings.com". Contoh lain menunjukkan lirik dari lagu "*When To Say When*" dan "*Chicago Freestyle (Ft. Giveon)*", yang juga disertai dengan interpretasi makna dan sumbernya, seperti Genius.com. Total lirik/frasa yang ditampilkan dalam contoh dataset ini adalah 5 baris, namun *dataset* tersebut sebenarnya berisi 514 baris.

## B. Pra-pemrosesan Data

Tahap pra-pemrosesan dilakukan untuk memastikan bahwa data siap digunakan dalam pelatihan model LLaMA 3. *Dataset* berisi potongan lirik lagu, makna, judul lagu, album, dan sumber referensi dibaca menggunakan pustaka *Pandas*, kemudian diformat ulang menjadi pasangan *prompt-response* sesuai format *instruction fine-tuning*. Proses pembersihan mencakup penghapusan spasi berlebih, karakter khusus (*punctuation removal*), serta perubahan seluruh teks menjadi huruf kecil (*lowercase*) agar konsisten dan mudah diproses oleh model. Tabel 2 berikut memperlihatkan contoh hasil pembersihan teks pada tahap *data cleaning*.

**Tabel 2** Tampilan Sebelum dan Sesudah *Cleaning*

Sebelum	Sesudah
<i>First time in a long time</i>	<i>first time in a long time</i>
<i>Hurtin' deeply inside</i>	<i>hurtin deeply inside</i>
<i>First time in a long time</i>	<i>first time in a long time</i>
<i>Hurtin' deeply</i>	<i>hurtin deeply</i>

Tabel 2 menunjukkan perbedaan antara teks mentah dan hasil pembersihan, di mana karakter khusus dihapus dan huruf kapital diubah menjadi huruf kecil untuk menjaga konsistensi teks. Setelah tahap pembersihan, *dataset* disimpan dalam format JSONL (*JavaScript Object Notation Lines*) yang mendukung efisiensi pembacaan oleh pustaka *Hugging Face Datasets*. Format ini memungkinkan setiap baris data berisi pasangan lengkap antara lirik, makna, judul lagu, album, dan sumber.

**Tabel 3** Tampilan Sebelum dan Sesudah Diubah ke Format JSONL

Sebelum (CSV)	Sesudah (JSONL)
album: <i>Dark Lane Demo Tapes</i> title: <i>Not You Too (Ft. Chris Brown)</i> lyric_bait: <i>First time in a long time...</i> meaning_bait: <i>Bait ini menggambarkan perasaan sakit hati...</i>	{"lyrics": "First time in a long time\nHurtin' deeply inside", "meaning": "Bait ini menggambarkan perasaan sakit hati yang mendalam...", "title": "Not You Too (Ft. Chris Brown)", "album": "Dark Lane Demo Tapes", "source": "genius.com"}

Konversi ini menghasilkan 513 entri data yang telah terformat secara konsisten. Setelah itu dilakukan proses tokenisasi menggunakan *AutoTokenizer* dari model *meta-llama/Llama-3.2-1B-Instruct*. Tokenisasi mengubah teks menjadi representasi numerik dengan pengaturan *truncation=True* dan *padding="max\_length"* agar seluruh *input*

memiliki panjang tetap (maksimal 512 token). *Dataset* hasil tokenisasi ini menjadi *input* akhir yang siap digunakan pada tahap *fine-tuning* model LLaMA 3.

### C. Pengembangan dan *Fine-tuning* Model LLaMA 3 + LoRA

Model utama yang digunakan dalam penelitian ini adalah LLaMA-3.2-1B-*Instruct* yang dikembangkan oleh Meta. Model ini merupakan varian dari arsitektur *Transformer* dengan kemampuan *instruction-following* yang unggul dalam memahami konteks dan menghasilkan teks secara relevan. Pemilihan model ini didasarkan pada kemampuannya dalam menangani tugas berbasis bahasa alami yang kompleks, seperti interpretasi makna lirik lagu. Untuk mengoptimalkan efisiensi sumber daya, model dimuat menggunakan teknik kuantisasi 8-bit melalui modul *BitsAndBytesConfig*, yang memungkinkan pelatihan dilakukan pada GPU dengan kapasitas memori terbatas seperti NVIDIA T4.

Proses *fine-tuning* dilakukan dengan pengaturan parameter yang telah dioptimalkan agar model dapat menyesuaikan diri terhadap *dataset* lirik tanpa kehilangan kemampuan generalisasinya. Konfigurasi lengkap ditunjukkan pada Tabel 4 berikut.

**Tabel 4** Konfigurasi *Fine-tuning*

<i>Hyperparameter</i>	Nilai ( <i>Value</i> )	Deskripsi Singkat
<code>num_train_epochs</code>	3	Jumlah total siklus pelatihan pada dataset.
<code>per_device_train_batch_size</code>	1	Jumlah data yang diproses per perangkat dalam satu langkah.
<code>gradient_accumulation_steps</code>	2	Jumlah langkah untuk mengakumulasi gradien (simulasi <i>batch size</i> 2).
<code>optim</code>	<code>paged_adamw_8bit</code>	Optimizer yang hemat memori.
<code>learning_rate</code>	5e-5	Tingkat laju pembelajaran untuk optimizer.
<code>bf16</code>	True	Mengaktifkan <i>mixed-precision training</i> untuk efisiensi.
<code>weight_decay</code>	0.01	Nilai regularisasi untuk mencegah <i>overfitting</i> .

Konfigurasi di atas dirancang agar proses pelatihan berjalan stabil meskipun pada lingkungan dengan sumber daya terbatas. *Learning rate* 5e-5 digunakan agar adaptasi pengetahuan berlangsung secara hati-hati, sementara *weight decay* 0.01 membantu menghindari *overfitting*. Nilai *batch size* kecil (1) dikompensasi dengan *gradient accumulation steps* sebanyak 2 agar simulasi pelatihan tetap stabil. Selain itu, *optimizer paged\_adamw\_8bit* digunakan untuk mengurangi beban memori, dan pengaktifan *mixed precision (bf16)* mempercepat proses komputasi tanpa mengurangi akurasi.

Selanjutnya, efisiensi pelatihan ditingkatkan menggunakan *Low-Rank Adaptation (LoRA)*, yaitu teknik *parameter-efficient fine-tuning* yang hanya melatih sebagian kecil parameter model. Konfigurasi LoRA ditampilkan pada Tabel 5.

**Tabel 5** Konfigurasi LoRA

<i>Hyperparameter</i>	Nilai ( <i>Value</i> )	Deskripsi Singkat
<code>R</code>	32	Ukuran (rank) dari matriks adaptasi LoRA yang bisa dilatih.
<code>lora_alpha</code>	32	Faktor penskalaan untuk bobot LoRA.
<code>target_modules</code>	["q_proj", "k_proj", "v_proj", "o_proj"]	Lapisan spesifik pada model yang menjadi target adaptasi LoRA.
<code>lora_dropout</code>	0.15	Nilai <i>dropout</i> pada lapisan LoRA untuk mencegah <i>overfitting</i> .

Pendekatan LoRA bekerja dengan “membekukan” sebagian besar parameter model dan hanya melatih matriks kecil (adapter) pada lapisan self-attention. Penempatan adapter pada modul *query* ( $q\_proj$ ), *key* ( $k\_proj$ ), *value* ( $v\_proj$ ), dan *output* ( $o\_proj$ ) memungkinkan model menyesuaikan hubungan semantik antarkata secara efektif tanpa harus memperbarui seluruh bobot *Transformer*. Nilai *rank* dan *alpha* yang seimbang (keduanya 32) memastikan kapasitas pembelajaran cukup tanpa mengganggu stabilitas model.

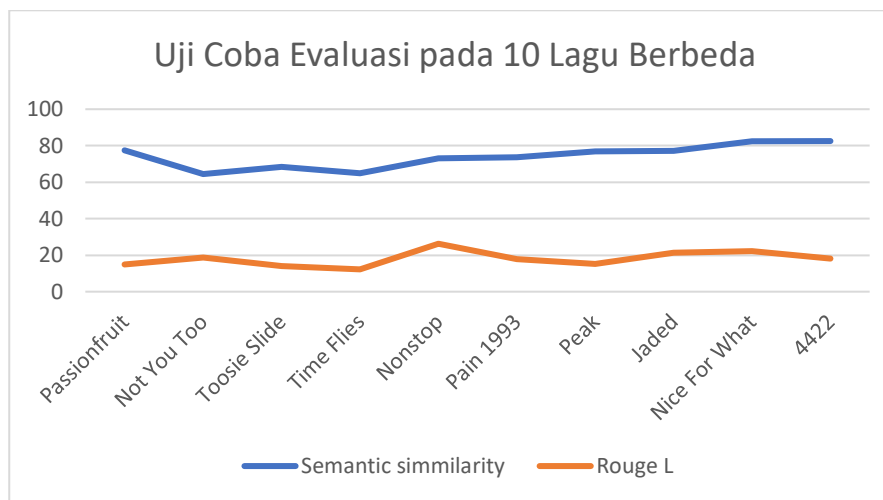
Selain itu, *lora\_dropout* sebesar 0.15 digunakan sebagai mekanisme regularisasi untuk mengurangi ketergantungan model pada jalur pembelajaran tertentu. Proses pelatihan ini diatur untuk jenis tugas *Causal Language Modeling* (CAUSAL\_LM), di mana model memprediksi token berikutnya berdasarkan konteks sebelumnya. Kombinasi antara konfigurasi *hyperparameter* dan strategi LoRA ini menghasilkan proses *fine-tuning* yang efisien, stabil, serta mampu mempertahankan kualitas semantik tinggi dalam interpretasi makna lirik lagu.

#### D. Evaluasi Model

Evaluasi dilakukan untuk mengukur sejauh mana model LLaMA 3 yang telah di-*fine-tune* mampu menghasilkan interpretasi makna lirik lagu Drake secara akurat dan relevan. Dua metrik utama digunakan, yaitu *Semantic Similarity* dan ROUGE-L (*Longest Common Subsequence*), untuk menilai kualitas semantik dan kesamaan struktur teks.

Pada pendekatan *Semantic Similarity*, digunakan model *SentenceTransformer* (all-MiniLM-L6-v2) untuk mengubah teks hasil keluaran dan teks referensi menjadi vektor numerik. Nilai kemiripan dihitung dengan *cosine similarity*, di mana skor mendekati 1 menunjukkan kesamaan makna yang tinggi. Metrik ini memastikan bahwa model memahami isi lirik secara kontekstual, bukan sekadar menyalin teks.

Sebagai pelengkap, metrik ROUGE-L digunakan untuk mengukur kesamaan urutan kata antara hasil model dan referensi. Skor ROUGE-L yang tinggi menunjukkan struktur kalimat dan pilihan kata yang sejalan dengan teks asli. Kedua metrik ini dihitung otomatis setiap kali sistem melakukan analisis, dan hasilnya ditampilkan bersamaan di antarmuka aplikasi untuk memudahkan perbandingan.



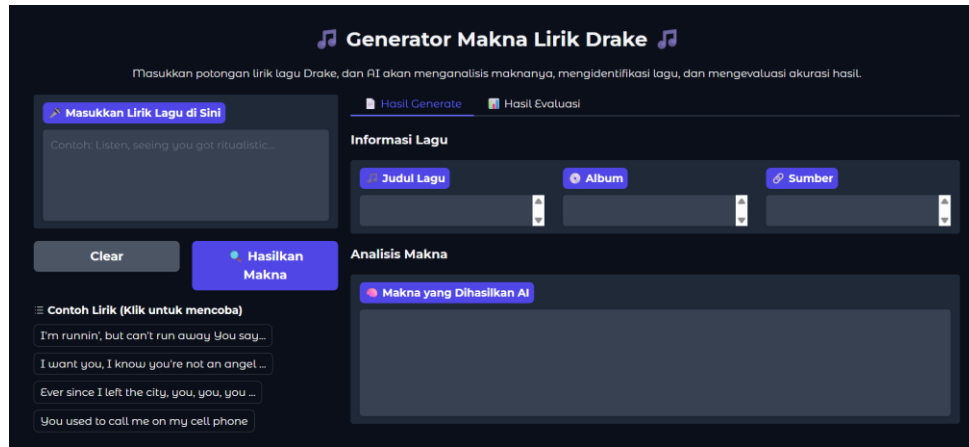
**Gambar 2** Chart Uji Coba Evaluasi.

Gambar 2 memperlihatkan hasil dari sepuluh uji coba model. Garis biru menunjukkan nilai *Cosine Similarity*, sedangkan garis oranye menampilkan nilai ROUGE-L. Terlihat adanya fluktuasi skor antar uji coba akibat perbedaan kompleksitas lirik — lirik dengan banyak metafora dan ambiguitas menghasilkan variasi skor yang lebih besar.

Secara umum, hasil evaluasi menunjukkan bahwa skor *Cosine Similarity* relatif stabil dan tinggi, menandakan kemampuan model memahami makna dengan baik. Variasi pada skor ROUGE-L menunjukkan bahwa model mampu memparafrasekan teks dengan struktur berbeda tanpa kehilangan makna utama. Dengan demikian, kombinasi dua metrik ini memberikan evaluasi yang menyeluruh terhadap kemampuan model dalam memahami dan mengekspresikan makna lirik lagu secara semantik maupun leksikal.

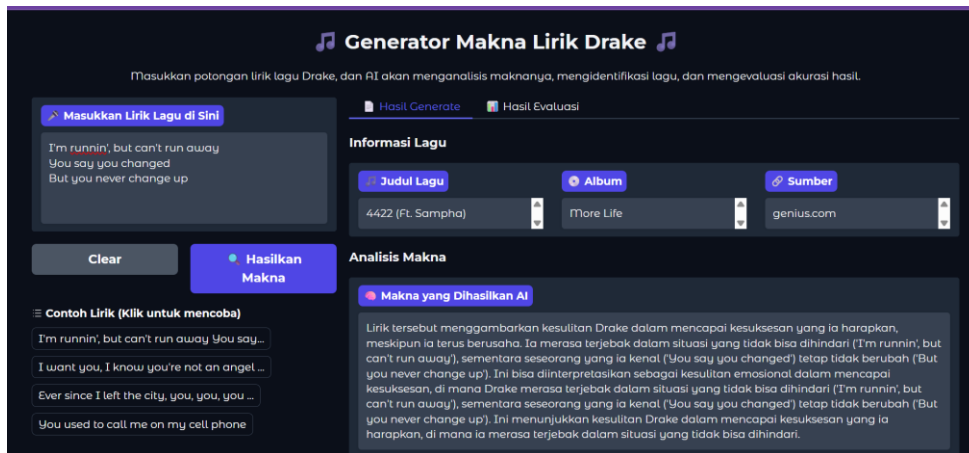
### E. Implementasi Antar Muka Menggunakan Gradio di Google Colab

Antarmuka aplikasi dirancang sederhana dan intuitif agar pengguna dapat berinteraksi langsung dengan sistem tanpa hambatan. Implementasi ini menunjukkan bahwa model hasil *fine-tuning* LLaMA 3 dapat diterapkan secara praktis dalam bentuk aplikasi web interaktif.



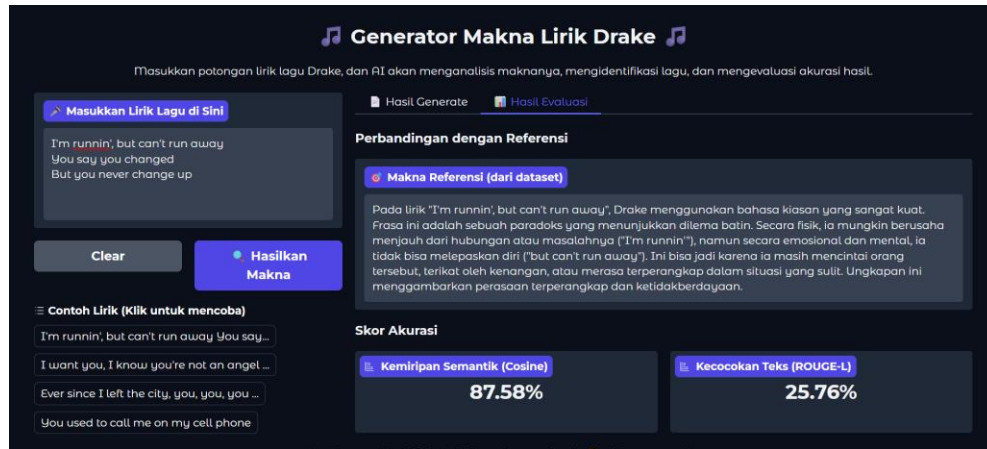
Gambar 3 Antarmuka Halaman Input Aplikasi Gradio.

Gambar 3 menampilkan antarmuka utama aplikasi Generator Makna Lirik Drake. Desainnya menggunakan *gr.Blocks* dengan tata letak dua kolom: sisi kiri berfungsi sebagai panel *input* dan sisi kanan sebagai panel *output*. Panel *input* berisi area teks utama “Masukkan Lirik Lagu di Sini”, tombol “Hasilkan Makna”, tombol “Clear”, serta fitur “Contoh Lirik” (*gr.Examples*) yang memudahkan pengujian cepat tanpa memasukkan data manual. Sementara panel *output* menampilkan dua tab, yaitu “Hasil Generate” dan “Hasil Evaluasi”, yang memungkinkan pengguna melihat hasil interpretasi dan nilai evaluasi secara terpisah.



Gambar 4 Antarmuka Halaman Hasil Generate.

Gambar 4 memperlihatkan alur kerja sistem berbasis *Retrieval-Augmented Generation* (RAG). Setelah pengguna memasukkan lirik, sistem menjalankan tahap *retrieval* menggunakan metode TF-IDF untuk menemukan lirik serupa dalam *dataset*. Informasi lagu yang sesuai, seperti judul, album, dan sumber referensi (misalnya *genius.com*), kemudian ditampilkan sebagai bukti bahwa sistem menemukan konteks yang relevan. Berdasarkan hasil pencarian ini, model LLaMA 3 menghasilkan “Makna yang Dihasilkan AI” — penjelasan semantik dari lirik yang diberikan. Proses ini membuktikan kemampuan model dalam mengubah bahasa puitis menjadi makna yang lebih eksplisit dan logis tanpa kehilangan konteks emosional.



**Gambar 5** Halaman Hasil Evaluasi dan Skor Kuantitatif *Real-time*.

Selanjutnya, Gambar 5 menunjukkan tab “Hasil Evaluasi”, yang menampilkan hasil validasi kuantitatif secara *real-time*. Sistem membandingkan hasil keluaran model dengan makna referensi dari *dataset*, kemudian menampilkan dua skor evaluasi utama: *Cosine Similarity* sebesar 87.58% dan ROUGE-L sebesar 25.76%. Nilai ini menandakan bahwa model memahami makna inti lirik dengan baik, meskipun menggunakan struktur kalimat yang berbeda — mencerminkan kemampuan parafrase dan abstraksi semantik yang kuat. Fitur ini tidak hanya memberikan transparansi terhadap kinerja model, tetapi juga memperlihatkan keseimbangan antara pemahaman makna dan kebebasan linguistik dalam proses generasi teks.

#### 4. DISCUSSION

Hasil penelitian ini menunjukkan bahwa model LLaMA 3, setelah melalui proses *fine-tuning* menggunakan metode *Low-Rank Adaptation* (LoRA) dan diintegrasikan ke dalam kerangka *Retrieval-Augmented Generation* (RAG), mampu menghasilkan interpretasi makna lirik lagu Drake secara koheren dan kontekstual. Temuan ini memperkuat dasar teoretis yang telah dijelaskan pada bagian pendahuluan, bahwa arsitektur berbasis *Transformer* dapat diadaptasi secara efektif untuk tugas interpretasi semantik dalam pemrosesan bahasa alami, khususnya pada teks berbentuk lirik musik.

Dibandingkan dengan penelitian sebelumnya, studi ini memberikan kontribusi yang berbeda dan signifikan. Penelitian Ventura dan Toker [7] berfokus pada pembuatan lirik (*lyric generation*), sedangkan Agrawal et al. [8] menerapkan model *Transformer* untuk klasifikasi emosi pada musik. Penelitian ini, sebaliknya, menitikberatkan pada interpretasi semantik, yaitu proses menerjemahkan ekspresi puitis dalam lirik menjadi makna yang dapat dipahami tanpa kehilangan kedalaman emosional dan nilai artistiknya. Dengan demikian, penelitian ini memperluas penerapan model *Transformer* dari sekadar generasi dan klasifikasi menuju pemahaman semantik yang lebih mendalam.

Dari sisi teknis, kombinasi antara *fine-tuning* menggunakan LoRA dan teknik kuantisasi berhasil menurunkan jumlah parameter yang dilatih hingga sekitar 0,13% dari total parameter asli. Efisiensi ini memungkinkan proses adaptasi model dilakukan pada infrastruktur dengan sumber daya terbatas seperti Google Colab, tanpa penurunan kinerja yang berarti. Temuan ini mendukung hasil penelitian Hu et al. (2021) yang menunjukkan bahwa LoRA efektif meningkatkan efisiensi pelatihan model bahasa berskala besar.

Implementasi sistem berbasis Gradio juga membuktikan kepraktisan pendekatan ini. Antarmuka satu halaman yang interaktif memungkinkan pengguna untuk memasukkan lirik, melihat hasil interpretasi, serta mengakses metrik evaluasi seperti *Cosine Similarity* dan ROUGE-L secara *real-time*. Nilai evaluasi sebesar 87,58% untuk kesamaan semantik dan 25,76% untuk ROUGE-L menunjukkan bahwa model mampu menangkap makna inti dengan baik meskipun menggunakan struktur kalimat yang berbeda. Hal ini membuktikan bahwa model memiliki kemampuan parafrase semantik, bukan sekadar meniru teks referensi.

Secara teoretis dan praktis, penelitian ini memberikan kontribusi nyata. Dari sisi teoretis, penelitian ini memperluas cakupan penerapan model bahasa besar dalam ranah budaya dan seni, khususnya pemahaman semantik lirik lagu. Dari sisi praktis, penelitian ini menunjukkan bahwa AI berbasis *Transformer* dapat dimanfaatkan sebagai alat bantu analisis musik, pendidikan, maupun penelitian budaya untuk menafsirkan makna lirik secara otomatis dan konsisten.

Namun demikian, terdapat beberapa keterbatasan yang perlu diakui. Kinerja model masih sangat bergantung pada keragaman *dataset*. Model cenderung menghasilkan makna yang lebih umum ketika diberikan lirik yang terlalu pendek atau minim konteks. Selain itu, kemampuan model dalam memahami metafora kompleks dan referensi budaya spesifik masih terbatas. Oleh karena itu, penelitian lanjutan disarankan untuk memperluas *dataset* agar mencakup lebih banyak artis dan menambahkan analisis *multimodal* (teks dan audio) guna memperdalam kemampuan interpretasi model.

Secara keseluruhan, penelitian ini berhasil membuktikan bahwa kombinasi antara arsitektur RAG dan metode *fine-tuning* yang efisien merupakan fondasi yang kuat untuk membangun sistem kecerdasan buatan yang mampu melakukan interpretasi semantik lirik lagu secara efektif dan inovatif.

## 5. CONCLUSION

Penelitian ini berhasil mencapai tujuannya untuk membangun sebuah aplikasi web generator makna lirik yang dapat menginterpretasikan lirik lagu Drake secara otomatis. Keberhasilan ini dicapai melalui tahapan terstruktur, mulai dari pengumpulan data dan melakukan *fine-tuning* pada model LLaMA 3 dengan pendekatan *Low-Rank Adaptation* (LoRA). Pada tahap implementasi akhir, dilakukan revisi signifikan dengan mengganti framework antarmuka dari Gradio di Google Colab (yang dijalankan secara lokal) menjadi Gradio yang dieksekusi di lingkungan Google Colab. Revisi ini terbukti krusial untuk mengatasi kendala performa, dengan mereduksi waktu inferensi dari  $\pm 5$  menit menjadi kurang dari 60.

Aplikasi yang dihasilkan secara efektif mengimplementasikan arsitektur *Retrieval-Augmented Generation* (RAG), yang memungkinkan sistem mengambil konteks lagu relevan sebelum menghasilkan interpretasi. Kinerja sistem divalidasi secara kuantitatif melalui evaluasi ganda yang juga diintegrasikan secara transparan ke dalam antarmuka aplikasi. Metrik kesamaan semantik (*semantic similarity*) menunjukkan kesamaan makna yang tinggi antara hasil model dengan data referensi. Sebagai pelengkap, metrik ROUGE-L digunakan untuk mengukur kesamaan struktural dan leksikal. Secara keseluruhan, penelitian ini menyajikan prototipe fungsional yang membuktikan kelayakan integrasi model bahasa canggih dengan teknik *fine-tuning* efisien untuk aplikasi analisis teks yang responsif dan tervalidasi.

## 6. REFERENCE

- [1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention Is All You Need," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [2] F. Rahman and S. Ratna, "Perancangan E-Learning Berbasis Web Menggunakan Framework Codeigniter," *Technol. J. Ilm.*, vol. 9, no. 2, p. 95, 2018, doi: 10.31602/tji.v9i2.1370.
- [3] Y. Agrawal, R. G. R. Shanker, and V. Alluri, "Transformer-Based Approach Towards Music Emotion Recognition from Lyrics," in *Advances in Information Retrieval*, D. Hiemstra, M.-F. Moens, J. Mothe, R. Perego, M. Potthast, and F. Sebastiani, Eds., Lecture Notes in Computer Science, vol. 12657. Cham, Switzerland: Springer, 2021, pp. 160–174, doi: 10.1007/978-3-030-72240-1\_12.
- [4] W. Duan, Z. Zhang, Y. Yu, and K. Oyama, "Interpretable Melody Generation from Lyrics with Discrete-Valued Adversarial Training," in *Proceedings of the 30th ACM International Conference on Multimedia (MM '22)*, Lisbon, Portugal, 2022, pp. 6973–6975, doi: 10.1145/3503161.3547742.
- [5] C. Wu *et al.*, "LLaMA Pro: Progressive LLaMA with Block Expansion," Jan. 2024, [Online]. Available: <http://arxiv.org/abs/2401.02415>
- [6] C. Raffel *et al.*, "Exploring the limits of transfer learning with a unified text-to-text transformer," *J. Mach.*

- Learn. Res.*, vol. 21, pp. 1–67, 2020.
- [7] Y. Mao *et al.*, “A survey on LoRA of large language models,” *Front. Comput. Sci.*, vol. 19, no. 7, pp. 1–144, 2025, doi: 10.1007/s11704-024-40663-9.
- [8] C. M. V. de Andrade, W. Cunha, D. Reis, A. S. Pagano, L. Rocha, and M. A. Gonçalves, “A Strategy to Combine 1stGen Transformers and Open LLMs for Automatic Text Classification,” Aug. 2024, [Online]. Available: <http://arxiv.org/abs/2408.09629>
- [9] S. Naseri, S. Reddy, J. Correia, J. Karlgren, and R. Jones, “The Contribution of Lyrics and Acoustics to Collaborative Understanding of Mood,” May 2022, [Online]. Available: <http://arxiv.org/abs/2207.05680>
- [10] A. Han and R. Park, “Deep-Hop Rapper: Using LSTM and Transformer for Rap Lyric Generation Stanford CS224N {Custom, Default} Project.” [Online]. Available: <https://deepbeat.org/>.
- [11] R. Zhang *et al.*, “LLaMA-Adapter: Efficient Fine-tuning of Language Models with Zero-init Attention,” Mar. 2023, [Online]. Available: <http://arxiv.org/abs/2303.16199>
- [12] N. Weir, A. Poliak, and B. Van Durme, “Probing Neural Language Models for Human Tacit Assumptions,” *Proc. 42nd Annu. Meet. Cogn. Sci. Soc. Dev. a Mind Learn. Humans, Anim. Mach. CogSci 2020*, pp. 377–383, 2020.
- [13] E. Hu *et al.*, “Lora: Low-Rank Adaptation of Large Language Models,” *ICLR 2022 - 10th Int. Conf. Learn. Represent.*, pp. 1–26, 2022.
- [14] F. Dennis Heraldi and F. Zakhralativa Ruskanda, “EasyChair Preprint Effective Intended Sarcasm Detection Using Fine-Tuned Llama 2 Large Language Models Effective Intended Sarcasm Detection Using Fine-tuned Llama 2 Large Language Models,” 2024.
- [15] M. Ventura and M. Toker, “TRBLLmaker -- Transformer Reads Between Lyrics Lines maker,” Dec. 2022, [Online]. Available: <http://arxiv.org/abs/2212.04917>
- [16] H. Touvron *et al.*, “Llama 2: Open Foundation and Fine-Tuned Chat Models,” 2023, [Online]. Available: <http://arxiv.org/abs/2307.09288>
- [17] E. M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell, “On the dangers of stochastic parrots: Can language models be too big?,” *FAccT 2021 - Proc. 2021 ACM Conf. Fairness, Accountability, Transpar.*, pp. 610–623, 2021, doi: 10.1145/3442188.3445922.

## 7. AUTHORS

Tubagus Alwasi'i merupakan mahasiswa Program Studi Teknik Informatika di Universitas Islam Sultan Agung, Semarang. Minat penelitiannya meliputi bidang Kecerdasan Buatan, Pembelajaran Mesin, dan Pengembangan Aplikasi Android. Ia memiliki pengalaman dalam pengembangan perangkat lunak serta pernah mengikuti program Studi Independen Bersertifikat dengan fokus pada Android Developer. Email: [tubagusalesi@gmail.com](mailto:tubagusalesi@gmail.com).

Sam Farisa Chaerul Haviana merupakan dosen di Universitas Islam Sultan Agung, Semarang. Fokus penelitiannya meliputi *Machine Learning*, *Data Mining*, dan *Natural Language Processing (NLP)*, dengan ketertarikan pada pengembangan dan penerapan metode kecerdasan buatan untuk pemecahan masalah berbasis data. Email: [sam@unissula.ac.id](mailto:sam@unissula.ac.id).