



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



KLASIFIKASI LIRIK LAGU MELALUI *SPEECH-TO-TEXT* DENGAN INDOBERT *MULTICLASS CLASSIFICATION* DAN LLAMA3:70B

SKRIPSI

MICHAEL NATANAEL 2107411002

**POLITEKNIK
NEGERI
JAKARTA**

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA**

2025



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



KLASIFIKASI LIRIK LAGU MELALUI *SPEECH-TO-TEXT* DENGAN INDOBERT *MULTICLASS CLASSIFICATION* DAN LLAMA3:70B

SKRIPSI

**Dibuat untuk Melengkapi Syarat-Syarat yang Diperlukan untuk
Memperoleh Diploma Empat Politeknik**

**POLITEKNIK
NEGERI
JAKARTA**

MICHAEL NATANAEL

2107411002

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER
POLITEKNIK NEGERI JAKARTA**

2025



SURAT PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan di bawah ini:

Nama : Michael Natanael
NIM : 2107411002
Jurusan/Program Studi : T. Informatika dan Komputer/Teknik Informatika
Judul Skripsi : Klasifikasi Lirik Lagu melalui *Speech-to-Text*
dengan *IndoBERT Multiclass Classification* dan
LLAMA3:70B

Menyatakan dengan sebenarnya bahwa skripsi ini benar-benar merupakan hasil karya saya sendiri, bebas dari peniruan terhadap karya dari orang lain. Kutipan pendapat dan tulisan orang lain ditunjuk sesuai dengan cara-cara penulisan karya ilmiah yang berlaku.

Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa dalam skripsi ini terkandung ciri-ciri plagiat dan bentuk-bentuk peniruan lain yang dianggap melanggar peraturan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Depok, 30 Juni 2025

Yang membuat pernyataan,



Michael Natanael

NIM 2107411002

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

LEMBAR PENGESAHAN

Skripsi diajukan oleh:

Nama : Michael Natanael
NIM : 2107411002
Program Studi : Teknik Informatika
Judul Skripsi : Klasifikasi Lirik Lagu melalui *Speech-to-Text* dengan *IndoBERT Multiclass Classification* dan *LLAMA3:70B*

Telah diuji oleh tim penguji dalam Sidang Skripsi pada hari Kamis, Tanggal 26, Bulan Juni, Tahun 2025 dan dinyatakan **LULUS**.

Disahkan oleh

Pembimbing I : Rizki Elisa Nalawati, S.T., M.T. (.....)

Penguji I : Euis Oktavianti, S.Si., M.T.I. (.....)

Penguji II : Hata Maulana, S.Si., M.T.I. (.....)

Penguji III : Zahra Azizah, S.Kom., M.I.S. (.....)

Mengetahui:

Jurusan Teknik Informatika dan Komputer
Ketua



Dr. Anita Hidayati, S.Kom., M.Kom.
NIP. 197908032003122003



KATA PENGANTAR

Puji dan syukur senantiasa penulis panjatkan kepada Tuhan Yang Maha Esa atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Lirik Lagu melalui *Speech-to-Text* dengan *IndoBERT Multiclass Classification* dan *LLAMA3:70B*” sebagai salah satu persyaratan untuk mencapai gelar Sarjana Terapan Program Studi Teknik Informatika di Politeknik Negeri Jakarta.

Penulis menyadari bahwa banyak pihak yang telah memberikan dukungan dan bantuan dalam proses penyelesaian studi dan skripsi ini. Dengan demikian, sudah sepantasnya penulis mengungkapkan rasa terima kasih yang tulus dan berdoa agar Allah memberikan balasan yang terbaik. Secara khusus penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu Dr. Anita Hidayati, S.Kom., M.Kom., selaku Ketua Jurusan Teknik Informatika dan Komputer.
2. Ibu Euis Oktavianti, S.Si., M.T.I., selaku Kepala Program Studi Teknik Informatika.
3. Ibu Rizki Elisa Nalawati, S.T., M.T., selaku dosen pembimbing yang telah dengar sabar membimbing, mengarahkan, membantu, serta menyemangati dalam proses penyelesaian skripsi ini.
4. Ibu AD, selaku ahli psikolog anak dan remaja yang telah bersedia menjadi narasumber wawancara serta memberikan validasi terhadap hasil klasifikasi sistem.
5. Seluruh responden yang telah bersedia meluangkan waktu untuk memberikan data dan masukan dalam pengujian sistem.
6. Seluruh Bapak/Ibu dosen yang telah mendidik penulis sehingga menjadi pribadi yang lebih baik.
7. Orang tua serta keluarga penulis yang telah memberikan segala dukungan baik berupa moral dan material.
8. Rekan-rekan dan sahabat penulis yang senantiasa memberikan semangat.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritis atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

Akhir kata, semoga skripsi penulis ini dapat bermanfaat bagi semua pihak, baik dari segi pendidikan maupun penelitian. Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna, penulis mengucapkan permohonan maaf atas kekurangan dan keterbatasan dalam skripsi ini. Oleh karena itu, penulis mengharapkan segala bentuk kritik, saran, dan masukan yang dapat membantu memperbaiki serta menyempurnakan skripsi ini.

Depok, 30 Juni 2025

Penulis,

Michael Natanael



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK
KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Politeknik Negeri Jakarta, saya bertanda tangan dibawah ini :

Nama : Michael Natanael
NIM : 2107411002
Jurusan/Program Studi : T. Informatika dan Komputer/Teknik Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Politeknik Negeri Jakarta Hak Bebas Royalti Non-Eksklusif atas karya ilmiah saya yang berjudul :

Klasifikasi Lirik Lagu melalui *Speech-to-Text* dengan *IndoBERT Multiclass Classification* dan *LLAMA3:70B*

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksklusif ini Politeknik Negeri Jakarta Berhak menyimpan, mengalihmediakan/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Depok, 30 Juni 2025

Yang menyatakan,



Michael Natanael

NIM 2107411002



KLASIFIKASI LIRIK LAGU MELALUI *SPEECH-TO-TEXT* DENGAN INDOBERT *MULTICLASS CLASSIFICATION* DAN LLAMA3:70B

ABSTRAK

Popularitas lagu-lagu Indonesia di media sosial dan layanan streaming telah menimbulkan kekhawatiran terkait lirik yang tidak sesuai usia, terutama dampaknya terhadap perilaku dan emosi anak-anak dan remaja. Oleh karena itu, diperlukan pengembangan model yang mampu mengklasifikasikan lirik lagu Indonesia ke dalam empat kategori usia pendengar: anak-anak, remaja, dewasa, dan semua usia. Penelitian terdahulu telah melakukan klasifikasi lirik lagu Indonesia berdasarkan kategori usia menggunakan algoritma *Naïve Bayes Classifier (NBC)*, menghasilkan hasil yang kurang memuaskan dengan akurasi 65%. Penelitian ini mengusulkan metode baru yang menggabungkan *Large Language Model (LLM)*, khususnya LLAMA3:70B untuk meningkatkan kinerja klasifikasi dan memperluas jumlah dataset hingga 1000%. Tiga strategi augmentasi dataset diterapkan dalam metode ini, yaitu *zero-shot*, *sample-based*, dan *translation-based*. Performa klasifikasi berhasil ditingkatkan secara signifikan melalui model terbaik hasil *fine-tune IndoBERT multiclass classification* dengan pendekatan *sample-based*, yang mencapai *F1 macro score* sebesar 74%. Model ini kemudian diimplementasikan dalam bentuk aplikasi web yang dilengkapi dengan fitur *speech-to-text* dengan *Whisper-Large-V3*. Hasil pengujian menunjukkan bahwa akurasi klasifikasi mencapai 75% pada benchmark dataset. Hasil evaluasi usability menunjukkan nilai *System Usability Scale (SUS)* sebesar 78,21 dan *Net Promoter Score (NPS)* sebesar 45%.

Kata kunci: *IndoBERT multiclass classification*, Kategori usia pendengar, Klasifikasi lirik lagu Indonesia, LLAMA3, *Whisper*

POLITEKNIK
NEGERI
JAKARTA

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



DAFTAR ISI

SURAT PERNYATAAN BEBAS PLAGIARISME.....	iii
LEMBAR PENGESAHAN	iv
KATA PENGANTAR	v
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	vii
ABSTRAK	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah.....	3
1.3. Batasan Masalah.....	4
1.4. Tujuan dan Manfaat.....	4
1.4.1. Tujuan	4
1.4.2. Manfaat	4
1.5. Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. <i>Bidirectional Encoder Representations from Transformers</i> (BERT)	7
2.2. IndoBERT	7
2.3. <i>Multiclass Classification</i>	8
2.4. Kategorisasi Usia.....	9
2.4.1. Anak-anak (0-12 tahun).....	9
2.4.2. Remaja (13-18 tahun)	9
2.4.3. Dewasa (19 tahun ke atas)	9
2.4.4. Semua Usia	9
2.5. <i>Large Language Model</i> (LLM)	10
2.6. LLAMA	10
2.7. LLAMA3:70B	10
2.8. Whisper	11
2.9. BERTScore.....	11

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

2.10. <i>Explainable Artificial Intelligence</i> (XAI).....	12
2.10.1. <i>Local Interpretable Model-agnostic Explanations</i> (LIME).....	12
2.10.2. <i>SHapley Additive exPlanations</i> (SHAP).....	12
2.11. <i>Hallucination</i> dalam LLM.....	13
2.11.1. <i>Benchmark Dataset</i> untuk Evaluasi <i>Hallucination</i>	13
2.11.2. <i>Fine-Tuning</i> dengan <i>Benchmark Dataset</i>	13
2.12. Flask	14
2.13. PostgreSQL	14
2.14. Tinjauan Pustaka	14
BAB III METODE PENELITIAN.....	18
3.1. Rancangan Penelitian	18
3.2. Tahapan Penelitian	18
3.2.1. Identifikasi Kebutuhan.....	19
3.2.2. Pengembangan Model.....	19
3.2.3. Pengujian Model.....	20
3.2.4. Pengembangan Web	20
3.2.5. Pengujian Web.....	20
3.2.6. Penulisan Laporan.....	20
3.3. Objek Penelitian	21
BAB IV HASIL DAN PEMBAHASAN	22
4.1. Analisis Kebutuhan	22
4.1.1. Kebutuhan Pengembangan Model.....	22
4.1.2. Kebutuhan Pengembangan Web	24
4.1.3. Kebutuhan Perangkat Keras.....	24
4.1.4. Kebutuhan Perangkat Lunak.....	25
4.2. Perancangan Sistem.....	27
4.2.1. Perancangan Model.....	27
4.2.2. Perancangan Web	34
4.3. Implementasi Sistem	43
4.3.1. Implementasi Model	43
4.3.2. Implementasi Web	54
4.4. Pengujian Sistem	59
4.4.1. Deskripsi Pengujian	60
4.4.2. Prosedur Pengujian	63



© Hak Cipta milik Politeknik Negeri Jakarta

4.4.3. Data Hasil Pengujian	63
4.4.4. Analisis Data Pengujian.....	74
BAB V PENUTUP.....	86
5.1. Simpulan.....	86
5.2. Saran.....	87
DAFTAR PUSTAKA	88
DAFTAR RIWAYAT HIDUP PENULIS	91
LAMPIRAN.....	92



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



DAFTAR TABEL

Tabel 2.1. <i>Literature Review</i>	14
Tabel 4.1. <i>Kebutuhan Dataset</i>	22
Tabel 4.2. <i>Kebutuhan Model</i>	23
Tabel 4.3. <i>Kebutuhan Fungsional</i>	24
Tabel 4.4. <i>Kebutuhan Non-Fungsional</i>	24
Tabel 4.5. <i>Kebutuhan Perangkat Keras Pengembangan Model</i>	25
Tabel 4.6. <i>Kebutuhan Perangkat Keras Pengembangan Web</i>	25
Tabel 4.7. <i>Kebutuhan Perangkat Lunak Pengembangan Model</i>	25
Tabel 4.8. <i>Kebutuhan Perangkat Lunak Pengembangan Web</i>	26
Tabel 4.9. <i>Jumlah Lirik Sintetis Hasil Augmentasi Dataset</i>	44
Tabel 4.10. <i>Generate Indonesian Lyrics Prompt</i>	44
Tabel 4.11. <i>Generate Indonesian Lyrics Using Example Prompt</i>	45
Tabel 4.12. <i>Generate English Lyrics Prompt</i>	46
Tabel 4.13. <i>Translate English Lyrics to Indonesian Lyrics Prompt</i>	47
Tabel 4.14. <i>Hasil Uji Similarity Score dengan BERTScore</i>	47
Tabel 4.15. <i>Hyperparameter Pelatihan Model</i>	49
Tabel 4.16. <i>Hasil Pre-training IndoBERT</i>	49
Tabel 4.17. <i>Hasil Fine-tuning IndoBERT</i>	51
Tabel 4.18. <i>Hyperparameter Whisper</i>	53
Tabel 4.19. <i>Skenario Black Box Testing</i>	60
Tabel 4.20. <i>Pertanyaan SUS</i>	62
Tabel 4.21. <i>Pertanyaan NPS</i>	62
Tabel 4.22. <i>Kategori dan Kelompok Sampel</i>	63
Tabel 4.23. <i>Hasil Black Box Testing</i>	64
Tabel 4.24. <i>Hasil SUS</i>	68
Tabel 4.25. <i>Nilai SUS</i>	70
Tabel 4.26. <i>Hasil NPS</i>	72
Tabel 4.27. <i>Perhitungan Nilai NPS</i>	73

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



DAFTAR GAMBAR

Gambar 2.1. Arsitektur IndoBERT	8
Gambar 3.1. Tahapan Penelitian	18
Gambar 3.2. Tahapan Metode Klasifikasi Lirik Berdasarkan Usia	19
Gambar 4.1. <i>Zero-Shot Lyric Augmentation</i>	27
Gambar 4.2. <i>Sample-Based Lyric Augmentation</i>	28
Gambar 4.3. <i>Translation-Based Lyric Augmentation</i>	30
Gambar 4.4. <i>Use Case Diagram</i>	34
Gambar 4.5. <i>Activity Diagram</i> Klasifikasi Lirik Lagu.....	36
Gambar 4.6. <i>Activity Diagram Login</i>	38
Gambar 4.7. <i>Activity Diagram</i> Registrasi Akun	39
Gambar 4.8. <i>Activity Diagram</i> Melihat Riwayat Klasifikasi Lirik Lagu.....	41
Gambar 4.9. Struktur <i>Database PostgreSQL</i>	42
Gambar 4.10. Halaman Utama Klasifikasi Lirik Lagu	54
Gambar 4.11. Halaman Hasil Prediksi Usia.....	55
Gambar 4.12. Halaman Registrasi Akun.....	56
Gambar 4.13. Halaman <i>Login</i> Pengguna	57
Gambar 4.14. Halaman Riwayat Prediksi Bagian 1.....	58
Gambar 4.15. Halaman Riwayat Prediksi Bagian 2.....	58
Gambar 4.16. Halaman Riwayat Prediksi Bagian 3.....	59
Gambar 4.17. <i>Confusion Matrix</i> IndoBERT <i>Multiclass Classification</i>	65
Gambar 4.18. Jenis Kelamin Responden Kuesioner Uji Coba Web.....	67
Gambar 4.19. Kategori Usia Responden Kuesioner Uji Coba Web	67
Gambar 4.20. Pekerjaan Responden Kuesioner Uji Coba Web.....	68
Gambar 4.21. <i>LIME Explanation</i> Ganteng Ganteng SWAG.....	75
Gambar 4.22. <i>SHAP Explanation</i> Ganteng Ganteng SWAG	75
Gambar 4.23. <i>LIME Explanation</i> Bengawan Solo	76
Gambar 4.24. <i>SHAP Explanation</i> Bengawan Solo.....	77
Gambar 4.25. <i>LIME Explanation</i> Garam & Madu	78
Gambar 4.26. <i>SHAP Explanation</i> Garam & Madu.....	78
Gambar 4.27. <i>LIME Explanation</i> Satu Satu Aku Sayang Ibu	80

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritikan atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

Gambar 4.28. SHAP <i>Explanation</i> Satu Satu Aku Sayang Ibu.....	80
Gambar 4.29. Diagram Nilai SUS Berdasarkan Jenis Kelamin.....	81
Gambar 4.30. Diagram Nilai SUS Berdasarkan Kategori Usia	82
Gambar 4.31. Diagram Nilai SUS Berdasarkan Pekerjaan.....	83
Gambar 4.32. Diagram Nilai NPS Berdasarkan Jenis Kelamin.....	84
Gambar 4.33. Diagram Nilai NPS Berdasarkan Kategori Usia	84
Gambar 4.34. Diagram Nilai NPS Berdasarkan Pekerjaan.....	85



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

DAFTAR LAMPIRAN

Lampiran 1 Transkrip Wawancara Latar Belakang	92
Lampiran 2 Sampel <i>Dataset</i> Lirik Lagu Indonesia (<i>Human-Annotated</i>).....	95
Lampiran 3 Sampel <i>Dataset</i> Lirik Lagu Sintetis (<i>Zero-Shot</i>).....	96
Lampiran 4 Sampel <i>Dataset</i> Lirik Lagu Sintetis (<i>Sample-Based</i>).....	98
Lampiran 5 Sampel <i>Dataset</i> Lirik Lagu Sintetis (<i>Translation-Based</i>).....	100





BAB I PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi memberikan pengaruh besar pada berbagai industri kreatif, khususnya pada industri musik Indonesia karena tingginya penetrasi media sosial dan layanan *streaming*. Perubahan ini tidak hanya terjadi pada cara orang menikmati musik, tetapi juga pada kebiasaan mendengarkan musik yang telah beralih ke platform digital. Media seperti stasiun radio juga bermigrasi ke ranah digital dengan menyediakan aplikasi yang memungkinkan pendengar mengakses siaran melalui ponsel pintar secara *streaming* (Minarni and Dewi, 2023).

Berbagai survei telah dilakukan untuk memahami sejauh mana penerimaan dan pola perilaku masyarakat Indonesia terhadap layanan *streaming* musik. DailySocial melaporkan bahwa 88% dari 1.955 responden di Indonesia secara rutin mendengarkan musik melalui layanan *streaming* dalam enam bulan terakhir, dimana 51% di antaranya menghabiskan 1 hingga 14 jam per minggu untuk mendengarkan musik secara *online* (Fauzan, Cleverin and Lumbaa, 2023). Selain itu, data dari IDN Research Institute mengungkap penggunaan platform *streaming* musik di tiga kelompok usia, yaitu 21-26 tahun, 16-20 tahun, dan 11-15 tahun. Spotify menempati peringkat teratas di kalangan pendengar berusia 11-15 tahun dengan tingkat penggunaan 94%, diikuti oleh kelompok usia 16-20 tahun sebanyak 66% dan 21-26 tahun sebanyak 61%. Platform seperti YouTube, JOOX, dan Apple Music juga turut menjadi pilihan meskipun dengan persentase yang lebih rendah (IDN Research Institute, 2024).

Kemudahan akses terhadap jutaan lagu melalui berbagai platform digital memberikan dampak positif bagi perkembangan industri musik, namun juga menimbulkan tantangan baru. Distribusi musik yang masif ini tidak diimbangi dengan sistem klasifikasi lirik yang memadai untuk mengelompokkan lagu berdasarkan kategori usia pendengar. Hal ini mengakibatkan anak-anak dan remaja dapat dengan mudah mengakses lagu-lagu dengan lirik yang mengandung bahasa

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

kasar, diskriminatif, atau referensi kekerasan dan seksual tanpa pengawasan orang dewasa (Rospocher and Eksir, 2023). Hal ini sangat kontras dengan industri film yang telah menerapkan label klasifikasi usia secara jelas dalam memilih tontonan yang sesuai (Shadiqin Thirafi and Rahutomo, 2018).

Berbagai penelitian telah mengeksplorasi dampak jenis lagu tertentu terhadap perilaku dan emosi anak-anak dan remaja. Sebuah studi dalam jurnal medis "*Family Medicine and Community Health*" mengungkapkan bahwa 90,8% dari 3.633 referensi lirik lagu-lagu teratas Billboard 1999–2018 mengandung tema dewasa, meliputi lirik sugestif seksual, bahasa kasar, konsumsi alkohol, dan penggunaan senjata. Paparan lirik bertema dewasa ini dapat meningkatkan kecenderungan perilaku berbahaya, seperti agresi atau kekerasan, bunuh diri, penggunaan tembakau, alkohol, dan obat-obatan terlarang, serta praktik seksual yang tidak aman (Kury *et al.*, 2020). Pada tahun 2016, YouTuber Indonesia Awkarin dan Young Lex memicu kontroversi dengan merilis video musik yang mengandung bahasa vulgar dan konten seksual eksplisit. Konten seperti ini dapat memberikan pengaruh buruk jika dikonsumsi tanpa filter yang sesuai (Rachmawaty, 2021).

Berdasarkan hasil wawancara dengan Ibu AD selaku psikolog, terungkap bahwa lirik lagu dengan tema kekerasan, bahasa kasar, atau seksual sangat memengaruhi perilaku anak-anak dan remaja (selengkapnya dapat dilihat pada Lampiran 1). Beliau menyaksikan sendiri bagaimana murid-murid di beberapa sekolah dasar bisa mengatakan hal-hal dewasa yang tidak sesuai dengan usia mereka. Menurut beliau, anak-anak cenderung menggunakan kata-kata dari lirik lagu yang mereka dengar dalam kehidupan sehari-hari tanpa mengetahui makna sebenarnya dari lirik lagu tersebut. Sementara itu, remaja cenderung dapat memahami makna sebenarnya dari lirik lagu yang didengar. Menurut beliau, fitur penyaringan konten usia pada platform *streaming* musik sangat efektif dalam mencegah dampak negatif ini.

Saat ini, telah ada penelitian sejenis yang mengembangkan klasifikasi lirik lagu Indonesia menggunakan metode Naive Bayes *Classifier*. Namun, penelitian tersebut hanya mencapai akurasi sebesar 65% dengan rasio 80% data latih dan 20% data uji (Shadiqin Thirafi and Rahutomo, 2018). Rendahnya akurasi ini disebabkan

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



oleh keterbatasan metode *Naive Bayes Classifier* yang mengasumsikan kelompok usia lirik dengan menganalisis frekuensi kata dan memperlakukan setiap kata sebagai fitur independen sehingga gagal memahami hubungan semantik dan kontekstual dalam lirik lagu, serta kurangnya jumlah *dataset* yang tersedia untuk pelatihan model (Kelly and Johnson, 2021).

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan sistem klasifikasi lirik lagu Indonesia berdasarkan kelompok usia dengan mengintegrasikan teknologi *speech-to-text* dan *IndoBERT Pre-trained Language Model (PLM)*. *IndoBERT* mampu memprediksi kelompok usia lirik melalui mekanisme *self-attention* yang telah dilatih pada korpus berukuran besar untuk memahami hubungan kontekstual antar kata, daripada memperlakukan setiap kata sebagai fitur independen. Dalam penelitian ini, *LLAMA3:70B* digunakan untuk menghasilkan *dataset* augmentasi lirik yang mengadopsi mekanisme generasi lirik dan metode pengambilan sampel (Jonason *et al.*, 2023; Liu *et al.*, 2024). Kemudian *IndoBERT multiclass classification* digunakan untuk mengklasifikasikan lagu berdasarkan kelompok usia, meliputi anak-anak, remaja, dewasa, dan semua usia. Setelah itu, model yang telah dikembangkan diimplementasikan ke dalam bentuk aplikasi web yang terdapat fitur *speech-to-text* untuk mengubah audio menjadi lirik lagu yang dapat diolah oleh model. Dengan demikian, penelitian ini diharapkan dapat menghadirkan sistem klasifikasi yang lebih akurat dalam memilah lirik lagu sesuai dengan kategori usia.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka perumusan masalah dalam penelitian ini adalah:

1. Bagaimana mengembangkan teknologi *speech-to-text* menggunakan *Whisper* untuk mendeteksi lirik lagu dari audio?
2. Bagaimana mengklasifikasikan lirik lagu Indonesia berdasarkan kategori usia menggunakan *IndoBERT multiclass classification*?
3. Bagaimana melakukan augmentasi *dataset* menggunakan mekanisme generasi lirik dan metode pengambilan sampel dengan *LLAMA3:70B*?

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

4. Bagaimana mengevaluasi performa model klasifikasi menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan Kappa?

1.3. Batasan Masalah

Dalam penelitian ini, terdapat beberapa batasan masalah yang perlu diperhatikan agar fokus dan ruang lingkup penelitian tetap terjaga:

1. Penelitian hanya berfokus pada lirik lagu berbahasa Indonesia.
2. Klasifikasi lirik lagu menggunakan IndoBERT *multiclass classification* dibatasi pada empat kategori usia pendengar, yaitu anak-anak, remaja, dewasa, dan semua usia.
3. Augmentasi *dataset* dilakukan menggunakan LLAMA3:70B melalui mekanisme generasi lirik dan pengambilan sampel.
4. Evaluasi performa model menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan Kappa.

1.4. Tujuan dan Manfaat

1.4.1. Tujuan

Tujuan dari penelitian ini meliputi:

1. Mengembangkan teknologi *speech-to-text* menggunakan Whisper untuk mendeteksi lirik lagu dari audio.
2. Mengklasifikasikan lirik lagu Indonesia berdasarkan kategori usia menggunakan IndoBERT *multiclass classification*.
3. Melakukan augmentasi *dataset* menggunakan mekanisme generasi lirik dan metode pengambilan sampel dengan LLAMA3:70B.
4. Mengevaluasi performa model klasifikasi menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan Kappa.

1.4.2. Manfaat

Penelitian ini diharapkan memberikan beberapa manfaat, yaitu:

1. Manfaat bagi industri musik:
 - Membantu dalam segmentasi pasar musik berdasarkan kategori usia.
2. Manfaat bagi pengguna platform musik:



Membantu menyaring dan menyesuaikan lirik lagu berdasarkan kategori usia pengguna.

3. Manfaat bagi peneliti *Natural Language Processing* (NLP):
Menjadi referensi dalam pengembangan penelitian klasifikasi lirik lagu berbahasa Indonesia dengan metode augmentasi *dataset* LLM.

1.5. Sistematika Penulisan

Dalam penulisan laporan ini, sistematika penulisan dibuat berdasarkan beberapa bab yang dimana masing-masing bab dapat diuraikan sebagai berikut:

BAB I PENDAHULUAN

Berisi tentang latar belakang penulisan laporan tugas akhir, tujuan penulisan laporan tugas akhir, manfaat penulisan tugas akhir, dan sistematika penulisan laporan tugas akhir.

BAB II TINJAUAN PUSTAKA

Berisi tentang kumpulan teori dasar dan pendapat para ahli yang mendukung masalah yang diteliti.

BAB III METODE PENELITIAN

Berisi penjelasan tentang metode yang akan digunakan, meliputi rancangan penelitian, tahapan penelitian, serta objek penelitian.

BAB IV HASIL DAN PEMBAHASAN

Berisi pembahasan menjelaskan pemaparan dan analisis hasil pengembangan model dan web klasifikasi lirik lagu Indonesia berdasarkan kategori usia, yang meliputi analisis kebutuhan, perancangan, implementasi, pengujian serta hasil analisis pengujian.

BAB V PENUTUP

Berisi penjelasan mengenai kesimpulan akhir dari penelitian dilengkapi dengan saran untuk penelitian selanjutnya.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

BAB V PENUTUP

5.1. Simpulan

Penelitian mengenai “Klasifikasi Lirik Lagu melalui *Speech-to-Text* dengan IndoBERT *Multiclass Classification* dan LLAMA3:70B” telah berhasil diselesaikan. Hasil penelitian ini mencakup augmentasi *dataset* dengan LLAMA3:70B yang telah teruji kelayakannya melalui BERTScore, serta pengembangan model klasifikasi usia berbasis IndoBERT dengan pendekatan *sample-based* yang menghasilkan F1 *Macro* sebesar 74%. Selain itu, penelitian ini juga berhasil menghasilkan aplikasi web klasifikasi lirik lagu berbahasa Indonesia berdasarkan kategori usia yang mengimplementasikan dua model utama, yaitu IndoBERT untuk klasifikasi usia dan Whisper-Large-V3 untuk transkripsi otomatis melalui *speech-to-text*. Fitur yang tersedia mencakup klasifikasi lirik, hasil prediksi usia, riwayat prediksi, serta sistem autentikasi pengguna melalui registrasi dan *login*. Sistem mampu mengelompokkan lirik lagu berbahasa Indonesia ke dalam empat kategori usia, yaitu anak-anak, remaja, dewasa, dan semua usia.

Sistem yang telah dikembangkan ini telah menjalani serangkaian pengujian, baik secara mandiri, wawancara dengan ahli psikolog, maupun melalui kuesioner uji coba web kepada pengguna akhir. Hasil pengujian menunjukkan bahwa akurasi klasifikasi sistem mencapai 75% pada *benchmark dataset evaluation*. Selain itu, pengujian kualitas pengalaman pengguna juga menunjukkan hasil positif dengan nilai *System Usability Scale* (SUS) sebesar 78,21 dan *Net Promoter Score* (NPS) sebesar 45%. Analisis lebih lanjut menggunakan pendekatan *Explainable AI* (XAI) dengan *Local Interpretable Model-agnostic Explanations* (LIME) dan *SHapley Additive exPlanations* (SHAP) menunjukkan bahwa kesalahan klasifikasi umumnya disebabkan oleh keterbatasan representasi data latih, khususnya untuk lirik yang mengandung bahasa kasar, eksplisit, atau bertema kompleks. Hal ini menegaskan pentingnya peningkatan keragaman *dataset* agar sistem mampu memahami nuansa tematik yang lebih luas.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



5.2. Saran

Berdasarkan hasil penelitian dan pengujian aplikasi web klasifikasi lirik lagu berdasarkan kategori usia, berikut beberapa saran untuk pengembangan lebih lanjut:

1. Peningkatan Kualitas Model: Untuk meningkatkan akurasi klasifikasi dan kemampuan adaptasi model terhadap lirik lagu modern yang cenderung dinamis, disarankan untuk mengimplementasikan pendekatan *Reinforcement Learning from Human Feedback* (RLHF). Dalam hal ini, pakar seperti psikolog dapat memberikan label korektif terhadap prediksi yang tidak sesuai, yang kemudian digunakan sebagai *reward signal* bagi sistem. Pendekatan ini memungkinkan model belajar secara berkelanjutan dan adaptif terhadap konteks baru yang lebih kompleks.
2. Diversifikasi Fitur: Sistem dapat ditingkatkan dengan menambahkan fitur identifikasi lagu dari potongan lirik atau input audio. Dengan demikian, sistem tidak hanya memberikan klasifikasi usia, tetapi juga dapat menampilkan informasi terkait lagu seperti judul, penyanyi, tahun rilis, dan genre. Fitur ini akan memperluas fungsi sistem sebagai alat bantu edukatif dan informatif bagi para pengguna.
3. Pengujian Lebih Luas: Untuk memperkuat generalisasi model, perlu dilakukan pengujian terhadap populasi pengguna yang lebih luas dan beragam, termasuk dari berbagai kelompok usia, budaya, dan bahasa daerah. Hal ini juga dapat memperkuat akurasi dan *fairness* model dalam memproses lirik dari berbagai konteks dan gaya bahasa dalam lirik lagu Indonesia.

Dengan implementasi saran-saran tersebut, diharapkan aplikasi web klasifikasi lirik lagu berdasarkan kategori usia dapat terus ditingkatkan kualitas dan fungsionalitasnya, sehingga dapat memberikan manfaat yang lebih besar bagi pengguna di bidang edukasi, penelitian, dan industri musik digital di Indonesia.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



DAFTAR PUSTAKA

Cui, Y., Yang, Z. and Yao, X. (2024) 'Efficient and Effective Text Encoding for Chinese LLaMA and Alpaca'. arXiv. Available at: <http://arxiv.org/abs/2304.08177> (Accessed: 20 August 2024).

Devlin, J. *et al.* (2019) 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding'. arXiv. Available at: <http://arxiv.org/abs/1810.04805> (Accessed: 29 August 2024).

Fauzan, A., Cleverin, R. and Lumbaa, S.B.P. (2023) 'THE EFFECTS OF DIGITAL PLATFORMS IN THE MUSIC INDUSTRY TO R&B MUSIC IN INDONESIA', 12(2).

Gohel, P., Singh, P. and Mohanty, M. (2021) 'Explainable AI: current status and future directions'. arXiv. Available at: <https://doi.org/10.48550/arXiv.2107.07045>.

Grattafiori, A. *et al.* (2024) 'The Llama 3 Herd of Models'. arXiv. Available at: <https://doi.org/10.48550/arXiv.2407.21783>.

Grinberg, M. (2018) *Flask web development: developing web applications with python*. O'Reilly Media, Inc.

Hakim, L.N. (2020) 'Urgensi Revisi Undang-Undang tentang Kesejahteraan Lanjut Usia', *Aspirasi: Jurnal Masalah-masalah Sosial*, 11(1), pp. 43–55. Available at: <https://doi.org/10.46807/aspirasi.v11i1.1589>.

IDN Research Institute (2024) *INDONESIA GEN Z REPORT 2024 (Understanding and Uncovering the Behavior, Challenges, and Opportunities)*. IDN Media. Available at: <https://cdn.idntimes.com/content-documents/indonesia-gen-z-report-2024.pdf> (Accessed: 9 January 2025).

Jonason, N. *et al.* (2023) 'Retrieval Augmented Generation of Symbolic Music with LLMs'. arXiv. Available at: <http://arxiv.org/abs/2311.10384> (Accessed: 18 August 2024).

Kelly, A. and Johnson, M.A. (2021) 'Investigating the Statistical Assumptions of Naïve Bayes Classifiers', in *2021 55th Annual Conference on Information Sciences and Systems (CISS)*. *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, USA: IEEE, pp. 1–6. Available at: <https://doi.org/10.1109/CISS50987.2021.9400215>.

Koto, F. *et al.* (2020) 'IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP', in *Proceedings of the 28th International Conference on Computational Linguistics. Proceedings of the 28th International Conference on Computational Linguistics*, Barcelona, Spain

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

(Online): International Committee on Computational Linguistics, pp. 757–770. Available at: <https://doi.org/10.18653/v1/2020.coling-main.66>.

Kury, Elise *et al.* (2020) ‘Lyrical Content of Contemporary Popular Music (1999–2018) and the Role of Healthcare Providers in Media Education of Children and Adolescents’, *Cureus* [Preprint]. Available at: <https://doi.org/10.7759/cureus.11214>.

Li, J. *et al.* (2022) ‘A Survey on Deep Learning for Named Entity Recognition’, *IEEE Transactions on Knowledge and Data Engineering*, 34(1), pp. 50–70. Available at: <https://doi.org/10.1109/TKDE.2020.2981314>.

Liu, S. *et al.* (2024) ‘MS²UGen: Multi-modal Music Understanding and Generation with the Power of Large Language Models’. arXiv. Available at: <http://arxiv.org/abs/2311.11255> (Accessed: 20 August 2024).

Lundberg, S. and Lee, S.-I. (2017) ‘A Unified Approach to Interpreting Model Predictions’. arXiv. Available at: <https://doi.org/10.48550/arXiv.1705.07874>.

Mardi Suryanto, T.L., Faroqi, A. and Simarmata, W.N. (2022) ‘SYSTEM USABILITY SCALE (SUS) SEBAGAI METODE PENGUJIAN KEGUNAAN PADA SITUS PROGRAM STUDI’, *Prosiding Seminar Nasional Teknologi dan Sistem Informasi*, 2(1), pp. 285–294. Available at: <https://doi.org/10.33005/sitasi.v2i1.314>.

Minarni, S. and Dewi, K.S. (2023) ‘PENGARUH WORD of MOUTH, DIGITAL MARKETING, BRAND AWARENESS TERHADAP COSTUMER LOYALTY PADA ARTIS BARU SEVEN MUSIC INDONESIA’, *Jurnal Manajemen dan Bisnis*, 3(1), pp. 1–17. Available at: <https://doi.org/10.32509/jmb.v3i1.2683>.

Rachmawaty, M. (2021) ‘The Necessity and Feasibility of Imposing Conventional Media Content Regulation on Digital Media Content in Indonesia (Case Study: The Controversy of Awkarin and Young Lex’s Music Video)’, 10(05).

Radford, A. *et al.* (2022) ‘Robust Speech Recognition via Large-Scale Weak Supervision’. arXiv. Available at: <https://doi.org/10.48550/arXiv.2212.04356>.

Rhamdani, N.I. (2021) ‘Net Promoter Score sebagai Tolok Ukur Ketercapaian Customer Loyalty Peserta Pekerja Penerima Upah Badan Usaha’, *Jurnal Jaminan Kesehatan Nasional*, 1(2), pp. 123–135. Available at: <https://doi.org/10.53756/jjkn.v1i2.34>.

Ribeiro, M.T., Singh, S. and Guestrin, C. (2016) ‘“Why Should I Trust You?”: Explaining the Predictions of Any Classifier’. arXiv. Available at: <https://doi.org/10.48550/arXiv.1602.04938>.



© Hak Cipta milik Politeknik Negeri Jakarta

Rospocher, M. and Eksir, S. (2023) ‘Assessing Fine-Grained Explicitness of Song Lyrics’, *Information*, 14(3), p. 159. Available at: <https://doi.org/10.3390/info14030159>.

Shadiqin Thirafi, Moch.F. and Rahutomo, F. (2018) ‘Implementation of Naïve Bayes Classifier Algorithm to Categorize Indonesian Song Lyrics Based on Age’, in *2018 International Conference on Sustainable Information Engineering and Technology (SIET)*. *2018 International Conference on Sustainable Information Engineering and Technology (SIET)*, Malang, Indonesia: IEEE, pp. 106–109. Available at: <https://doi.org/10.1109/SIET.2018.8693201>.

Sholeh, M. *et al.* (2021) ‘Black Box Testing on ukmbantul.com Page with Boundary Value Analysis and Equivalence Partitioning Methods’, *Journal of Physics: Conference Series*, 1823(1), p. 012029. Available at: <https://doi.org/10.1088/1742-6596/1823/1/012029>.

The PostgreSQL Global Development (2024) ‘PostgreSQL’, <https://www.postgresql.org/>. Available at: <https://www.postgresql.org/>.

Xu, Z., Jain, S. and Kankanhalli, M. (2025) ‘Hallucination is Inevitable: An Innate Limitation of Large Language Models’. arXiv. Available at: <https://doi.org/10.48550/arXiv.2401.11817>.

Zhang, T. *et al.* (2020) ‘BERTScore: Evaluating Text Generation with BERT’. arXiv. Available at: <https://doi.org/10.48550/arXiv.1904.09675>.

**POLITEKNIK
NEGERI
JAKARTA**

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



DAFTAR RIWAYAT HIDUP PENULIS



Michael Natanael

Lahir di Jakarta, 30 Desember 2002. Anak kedua dari dua bersaudara. Lulus dari SDK Kanaan Jakarta pada tahun 2014, SMPK 3 Penabur Jakarta pada tahun 2017, dan SMKN 26 Jakarta pada tahun 2021. Saat ini sedang menempuh Pendidikan Sarjana Terapan pada Program Studi Teknik Informatika di Politeknik Negeri Jakarta. Tertarik pada bidang *machine learning* dan *software development*.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta





Hak Cipta :
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penerbitan karya ilmiah, penerbitan laporan, penerbitan kritikan atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

LAMPIRAN

Lampiran 1 Transkrip Wawancara Latar Belakang

Transkrip Wawancara

Wawancara dengan : Ahli Psikologi Anak dan Remaja

Hari/Jam : Selasa, 7 Januari 2025/ Pukul 19.00 - 19.45 WIB

Daftar Pertanyaan:

1. Bisakah Anda menjelaskan bagaimana anak-anak dan remaja memproses pesan dari lirik lagu yang mereka dengar?

Jawab:

Berdasarkan latar belakang saya di bidang psikologi dan pengalaman sebagai guru SD, saya melihat bahwa anak-anak cenderung meniru lagu yang mereka dengar, tanpa benar-benar memahami isi liriknya. Mereka menyanyikannya bersama teman-teman hanya karena terdengar menarik. Lagu-lagu yang populer sering kali mengandung konten dewasa yang sebenarnya belum sesuai dengan usia mereka. Mereka menirukan lirik-lirik tersebut karena daya tarik sosial dan lingkungan, bukan karena memahami maknanya.

2. Menurut pengalaman Anda, sejauh mana lirik lagu dengan tema kekerasan, bahasa kasar, atau seksual memengaruhi perilaku anak-anak dan remaja?

Jawab:

Saya pernah membaca dan melihat video terkait kesehatan mental yang menjelaskan bahwa jika anak-anak terlalu sering mendengar lirik lagu seperti itu, mereka cenderung menormalisasi konten tersebut. Mereka menyerap hal-hal dari lingkungan sekitarnya secara cepat. Hal ini berbahaya karena mereka bisa mulai meniru atau melakukan hal-hal yang sebenarnya tidak pantas untuk usianya, hanya karena terbiasa mendengarnya.

3. Apakah ada perbedaan pengaruh antara anak-anak dan remaja dalam merespons konten lirik lagu yang tidak sesuai usia?

Jawab:

Anak-anak dan remaja sama-sama rentan, tapi dengan cara yang berbeda. Anak-anak lebih sering meniru tanpa tahu makna, hanya karena seru. Sedangkan remaja biasanya sudah mulai memahami isi lirik dan justru lebih berbahaya karena bisa menormalisasi isi lagu yang tidak seharusnya. Mereka bisa merasa bahwa lirik itu hal yang wajar karena dinyanyikan artis terkenal, atau karena ingin ikut tren (FOMO).

4. Dalam pandangan Anda, apakah anak-anak dan remaja dapat membedakan antara hiburan semata dan pesan yang ada dalam lirik lagu?

Jawab:

Anak-anak kecil belum bisa membedakan. Mereka hanya menyerap tanpa filter. Remaja mungkin sudah bisa, tapi karena faktor psikologis seperti pubertas dan tekanan sosial, mereka tetap bisa terpengaruh dan menganggap lirik yang negatif itu keren. Maka dari itu, tetap perlu bimbingan dari orang dewasa.

5. Apakah Anda pernah menangani kasus di mana dampak negatif lirik lagu terlihat jelas pada perilaku anak atau remaja? Bisa diceritakan secara umum?

Jawab:

Secara langsung menangani belum pernah. Tapi saya pernah kaget mendengar anak SD menyanyikan lagu-lagu bernuansa politik atau dewasa yang mereka dapat dari handphone kakaknya. Dari situ saya sadar, anak-anak bisa menyerap konten dewasa dari media sosial tanpa benar-benar tahu konsekuensinya. Ini menandakan bahwa mereka sudah terpapar konten yang tidak sesuai dari media sosial atau lingkungan sekitar.

6. Menurut Anda, siapa yang paling bertanggung jawab untuk memastikan anak-anak dan remaja terhindar dari lirik lagu yang tidak sesuai? Orang tua, sekolah, atau platform streaming musik?

Jawab:

Orang tua. Karena waktu anak-anak banyak dihabiskan bersama orang tua. Tapi kalau orang tua tidak punya waktu atau edukasi cukup, maka pendekatan holistik diperlukan. Sekolah, guru, masyarakat, dan bahkan platform streaming musik juga harus ikut andil dalam memberikan batasan dan edukasi.

7. Apakah ada tanda-tanda tertentu yang bisa diamati oleh orang tua atau guru jika anak terpengaruh oleh lirik lagu yang tidak sesuai usianya?

Jawab:

Bisa saja anak mengalami konflik batin, terutama jika di rumah dia melihat atau mendengar konten dewasa, sementara di sekolah diajarkan nilai-nilai moral atau agama. Anak bisa bingung menentukan mana yang benar, karena tidak ada pedoman yang jelas.

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

(lanjutan)

8. Dari sudut pandang psikologi, seberapa efektif fitur penyaringan konten usia pada platform streaming musik dalam mencegah dampak negatif ini?

Jawab:

Sangat penting. Setidaknya ada filter usia dasar terlebih dahulu. Tidak semua orang tua punya edukasi untuk memahami dampak lirik lagu. Jadi jika platform bisa menyediakan penyaringan, itu akan sangat membantu dalam membatasi akses anak-anak terhadap konten yang tidak sesuai usia. Jika dari skala satu sampai sepuluh, saya beri nilai sembilan karena fitur ini sangat penting.

9. Menurut Anda, apakah penting bagi industri musik untuk memiliki regulasi yang lebih ketat terkait klasifikasi usia pada lagu?

Jawab:

Penting, tapi jangan terlalu ketat. Kalau terlalu ketat justru bisa membuat anak-anak semakin penasaran. Yang penting ada sistem klasifikasi dan edukasi yang seimbang, serta tetap melibatkan peran orang tua dalam pengawasan.

10. Apakah ada saran yang bisa Anda berikan kepada para orang tua dan pendidik untuk meminimalkan dampak negatif dari lirik lagu dewasa terhadap anak-anak dan remaja?

Jawab:

Kalau dari orang tua, saya tahu sekarang hampir semua anak sudah pegang gadget, tapi sebisa mungkin ketika diberikan gadget, didampingi penggunaannya. Terutama ketika anak-anak itu masih belum paham betul apa yang mereka lihat di media sosialnya.

Kalau dari pihak sekolah, tentunya sekolah bisa memberikan pedoman-pedoman moral, agama, atau etika terkait penggunaan media sosial, mendengarkan lirik lagu juga termasuk di dalamnya.

Kalau dari sisi platform, menurut saya akan baik jika lagu-lagu anak dibuatkan wadah tersendiri. Contohnya, setahu saya Netflix punya fitur *Netflix Kids*. Jadi orang tua bisa memilih tayangan khusus untuk anak. YouTube juga punya *YouTube Kids*, seperti itu. Menurut saya, itu sudah menjadi salah satu opsi yang membantu orang tua dalam memberikan batasan serta perlindungan terhadap anak-anak.

- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Lampiran 2 Sampel *Dataset* Lirik Lagu Indonesia (*Human-Annotated*)

Judul	Lirik Lagu	Label Usia
Lirik Lagu Pok Ame Ame - Lagu Anak Indonesia	Pok ame ameBelalang kupu kupuDitepuk biar ramaiDiupah air susuKalau nangis jangan bertingkahPada waktunya minum susuKalau nanti suka bertingkahBisa kurus seperi batu	Anak
Lirik Lagu Aku Anak Gembala - Tasya	Aku adalah anak gembalaSelalu riang serta gembiraKarena aku senang bekerjaTak pernah malas ataupun lengahTralala la la laTralala la la la la la la laSetiap hariku bawa ternak ke padang rumput,di kaki bukit rumputnya hijau,subur dan banyakTernakku makan tak pernah sedikit Tralala.....	Anak
Lirik Lagu Cinta Mati - Aura Kasih	Selamanya kau telah pergi tinggalkanKu sisakan pedih di hatiku iniDan kini ku sendiri mengenangmuTak ada lagi yang selalu ku nantiDan aku takkan pernah percayaDan aku takkan mampu percayaKau telah tiadaKaulah cinta matikuKini kau pergi tinggalkanku sendiriTanpa cintamu lagiKaulah cinta matikuDan ku ingin kau tuk kembaliJangan pergiKu tak mau begini	Dewasa
Lirik Lagu Cinta Satu Malam - Melinda	Walau cinta kita sementaraAku merasa bahagiaKalau kau kecup mesra di keningkuKu rasa bagai di SurgaCinta satu malamOh indahnyaCinta satu malamBuatku melayangWalau satu malamAkan selalu ku kenangDalam hidupkuCinta satu malamOh indahnyaCinta satu malamBuatku melayangWalau satu malamAkan selalu ku kenangSelama-lamanyaSentuhanmu membuatku terlenaAku telah terbuai mesraYang ku rasa hangat indahnya cintaHasratku kian membara	Dewasa
Lirik Lagu Kekasih Terhebat - Anji	Aku manusia yang paling butuh kamuMembutuhkanmuHatiku memalingkan pandang hanya padamuPada hatimuJangan lelah menghadapikuBiarkan aku jadi yang terhebatJadilah kamu kekasih yang kuatGenggam tanganku bernyanyi bersamaKarena kamu kekasih terhebatAku manusia yang paling butuh kamuMembutuhkanmuJangan lelah menghadapiku	Remaja
Lirik Lagu Sempurna - Andra & The Backbone	Kau begitu sempurnaDi mataku kau begitu indahkau membuat diriku akan selalu memujimuDi setiap langkahkuKukan selalu memikirkan dirimuTak bisa kubayangkan hidupku tanpa cintamuJanganlah kau tinggalkan dirikuTakkan mampu menghadapi semuaHanya bersamamu ku akan bisa Kau adalah darahkuKau adalah jantungkuKau adalah hidupkuLengkapi dirikuOh sayangku, kau begituSempurna.. Sempurna..Kau genggam tangankuSaat diriku lemah dan terjatuhKau bisikkan kata dan hapus semua sesalku	Remaja
Lirik Lagu Buat Apa Susah - Koes Plus	Buat apa susah...Buat apa susah...Lebih baik kita bergembiraBuat apa susah...Buat apa susah...Lebih baik kita bergembiraKekasihku apa yang kau risaukanKerjamu hanya melamun sajaTak berguna kau bersedih hatiTertawalah... sayang... Kekasihku apa yang kau pikirkanHidup ini hanya sementaraTak berguna kau bersedih hatiPercayalah... sayang...	Semua usia
Lirik Lagu Tanah Airku - Kotak	Tanah air ku tidak ku lupakanKan terkenang selama hidupkuBiarpun saya pergi jauhTidak kan hilang dari kalbuTanah ku yang ku cintaiEngkau ku hormaiWalaupun banyak negeri ku jalaniYang masyhur permai di kata orangTetapi kampung dan rumahkuDi sanalah ku rasa senangTanah ku tak ku lupakanEngkau ku banggakanTanah ku yang ku cintaiEngkau ku hormaiKu hormai	Semua usia

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Lampiran 3 Sampel *Dataset* Lirik Lagu Sintetis (*Zero-Shot*)

Judul	Lirik Lagu Sintetis (Buatan LLM)	Label Usia
Menggapai Mimpi	Mimpi itu, kita gapai dengan tangan Dengan percaya diri dan tekad yang kuat Kita tidak boleh menyerah Dalam perjalanan panjang Kita belajar dari kegagalan Agar kita dapat bangkit kembali Kita tidak boleh putus asa Dalam menghadapi rintangan Kita harus percaya diri Bahwa kita pasti dapat mencapai Dengan kita berjuang dan tidak menyerah Menggapai mimpi, kita tidak boleh Menyerah dalam perjalanan Kita pasti dapat mencapai Dengan kita percaya diri Kita pasti dapat menyelesaikan	Semua usia
Ketulusan Hati	Ketulusan hati, kita jaga dengan baik Dalam setiap langkah yang kita ambil Kita tidak boleh berpura-pura Dalam setiap kata yang kita ucapkan Kita harus mengetahui batas Dalam setiap tindakan yang kita ambil Kita harus memahami Bahwa ketulusan hati sangat berharga Kita harus menjaga dengan baik Dalam setiap langkah yang kita ambil Kita tidak boleh berpura-pura Kita harus mengetahui batas Dalam setiap tindakan yang kita ambil Ketulusan hati, kita jaga dengan baik Dalam setiap langkah yang kita ambil	Semua usia
Gembira Itu Sederhana	Gembira itu sederhana saja Tidak perlu sesuatu yang banyak Cukup bermain dan tertawa Itu sudah membuatku gembira Gembira itu ada di sekitar kita Mereka selalu membuat kita gembira Cukup berbagi dan bersama Itu sudah membuatku merasa nyaman Gembira itu bukan sesuatu yang mahal Cukup bermain dan berteriak Itu sudah membuatku gembira Gembira itu sederhana saja Tidak perlu sesuatu yang banyak Cukup bermain dan tertawa Itu sudah membuatku gembira	Anak
Lagu Untuk Ibu	Ibu ku yang tercinta Selalu ku sayangi Mengasuh dan menjaga ku Menjadi teman yang setia Ku ingin membuat ibu ku Bangga dan bahagia Dengan lagu ini ku cipta Untuk ibu ku yang tercinta Ibu ku adalah anugerah Bagi ku yang kecil Ku ingin membuat ibu ku Senang dan bahagia Dengan lagu ini ku cipta Untuk ibu ku yang tercinta Ibu ku takkan pernah Kehilangan kasih sayang ku Ku ingin membuat ibu ku	Anak
Menyongsong Masa Depan	Kami anak muda, penuh cita-cita Menggapai bintang, yang tak pernah pudar Menjadi penerus, yang menggapai masa Membangun negeri, dengan hati yang mulia Kami tak pernah,-menyerah pada kegagalan Kami berjuang, untuk meraih keyakinan Membela kebenaran, yang kita pegang Menjadi pemimpin, yang membawa negeri Kami remaja, yang penuh semangat Membangun masa depan, yang lebih baik Kami tak kan, meninggalkan jejak Mengikuti kebenaran, yang kita yakini Membela negara, dengan semua kita Menjadi generasi, yang penuh harapan Membangun negeri, yang lebih sejahtera Menggapai cita, yang kita impikan	Remaja
Kami Anak Negeri	Kami anak negeri, yang penuh harapan Menggapai cita-cita, yang kita gapai Menjadi pribadi, yang lebih baik Menggapai bintang, yang tak pernah pudar Kami tak pernah, menyerah pada kegagalan Kami berjuang, untuk meraih keyakinan Membela kebenaran, yang kita pegang Menjadi pemimpin, yang membawa negeri Kami anak negeri, yang penuh semangat Membangun masa depan, yang lebih baik Kami tak kan, meninggalkan jejak Mengikuti kebenaran, yang kita yakini Membela negara, dengan semua kita Menjadi generasi, yang penuh harapan Membangun negeri, yang lebih sejahtera Kami anak negeri, yang penuh makna	Remaja

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkannya dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



(lanjutan)

Ku Yakin Kau Bisa	Ku yakin kau bisa Melakukan apa yang kau inginkan Ku yakin kau kuat Menghadapi apa yang kau hadapi Ku yakin kau bijak Mengambil keputusan yang tepat Ku yakin kau pandai Melakukan apa yang kau cita-citakan Ku yakin kau hebat Menghadapi apa yang kau takuti Ku yakin kau bisa Melakukan apa yang kau mimpi Ku yakin kau kuat Menghadapi apa yang kau hadapi Ku yakin kau bijak Mengambil keputusan yang tepat	Dewasa
Jarak yang Kita Buat	Kita berjalan beriringan Namun masih ada jarak di antara Ku tak tahu apa yang terjadi Kita berjauhan dan tak bersama Ku melihatmu dari kejauhan Namun tak dapat kugapai Kita membiarkan jarak itu Menghilangkan komunikasi yang sempat Ku tak tahu bagaimana memulai Karena kita sudah terlalu jauh Kita membiarkan cinta memudar Menghilangkan kasih sayang yang sempat Ku tak tahu harus apa Karena kita sudah berbeda	Dewasa



- Hak Cipta :**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Lampiran 4 Sampel *Dataset* Lirik Lagu Sintetis (*Sample-Based*)

Judul	Lirik Lagu Sintetis (Buatan LLM)	Label Usia
Jalan Kesuksesan	Kita berjalan untuk kesuksesan Jalan yang tidak pernah kita tinggalkan Kita tidak pernah menyerah Karena kita percaya diri Kita akan menulis sejarah Bersama kita akan berjaya Kita tidak pernah berhenti Berjalan menuju kesuksesan Kita akan selalu berjuang Menaikkan kepercayaan diri kita Kita akan menang atas kesulitan Kita akan menang atas kesusahan Kita tidak pernah sendiri Karena kita memiliki saudara Kita saling bantu dan membantu Menaikkan kepercayaan diri kita	Semua usia
Tanah Air Kita	Di atas bumi, yang kita injak Terbentang lahan, yang subur dan hijau Di dalam udara, yang kita hirup Terdapat kesegaran, yang tak pernah mati Di dalam air, yang kita minum Terbentang lahan, yang subur dan hijau Di dalam api, yang kita gunakan Terbentang lahan, yang subur dan hijau Kita pernah menemukan, makna dari hidup Di atas bumi, yang kita sayang Kita pernah menemukan, jalan dari kesulitan Di atas bumi, yang kita jalani Kita pernah menemukan, kesegaran dari alam Di atas bumi, yang kita pijak Kita pernah menemukan, keindahan dari bumi Di atas bumi, yang kita tuju	Semua usia
Kolam Ikan	Kolam ikan di depan mata Ikan-ikan berenang kesana-kemari Kita tunggu dengan sabar Ikan-ikan sangat lucu Kita beri makan ikan-ikan Ikan-ikan sangat bahagia Kolam ikan sangat indah Ikan-ikan adalah sahabat Kita belajar tentang ikan Ikan-ikan sangat penting Kolam ikan sangat sejuk Ikan-ikan sangat berwarna Kita sangat menyukai ikan Ikan-ikan sangat indah Kolam ikan sangat bersih Ikan-ikan sangat seimbang	Anak
Si Kecil Tukang Kayu	Si kecil tukang kayu, bekerja dengan keras Membuat mainan, untuk anak-anak Si kecil tukang kayu, sangat berbakat Membuat rumah, untuk burung beristirahat Si kecil tukang kayu, bekerja dengan senang Membuat meja, untuk anak-anak belajar Si kecil tukang kayu, sangat pandai Membuat kursi, untuk orang tua duduk Si kecil tukang kayu, bekerja dengan rajin Membuat bangku, untuk anak-anak duduk Si kecil tukang kayu, sangat bersemangat Membuat lemari, untuk anak-anak menyimpan Si kecil tukang kayu, bekerja dengan gembira Membuat pintu, untuk anak-anak masuk Si kecil tukang kayu, sangat berbahagia Membuat jendela, untuk anak-anak melihat	Anak
Seperti Kekasih	Kau adalah kekasihku Yang membuatku bahagia Kau adalah kekasihku Yang membuatku bahagia Ku tak pernah lupa Bahwa kau adalah kekasihku Ku tak pernah lupa Bahwa kau adalah kekasihku Ku ingin kita bersama Dan ku jaga hatimu Ku tak pernah lupa Bahwa kau adalah kekasihku Ku tak pernah lupa Bahwa kau adalah kekasihku Ku ingin kita bersama Dan ku jaga hatimu	Remaja
Saat Kau Di Sini	Saat kau di sini, seperti langit cerah Saat kau di sini, seperti bumi yang stabil Ku rasakan hati ini bergetar kembali Saat kau di sini, seperti pelangi setelah hujan Ku tahu kau ada, seperti ku tahu kau ada Saat kau di sini, seperti denyut nadi Ku bersama kau, kita bersama Saat kau di sini, seperti cahaya yang terang Ku tahu kau ada, seperti ku tahu kau ada Saat kau di sini, seperti kesempurnaan Ku rasakan hati ini bergetar kembali Saat kau di sini, seperti pelangi setelah hujan	Remaja

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



© Hak Cipta milik Politeknik Negeri Jakarta

(lanjutan)

Cinta yang Terluka	Cinta kita dahulu sangat indah Tapi kini hanya tinggal luka Aku ingat waktu kita bersama Dan cinta kita yang dahulu sangat kuat Tapi sekarang semua telah luka Dan aku hanya tinggal dengan kesedihan Aku mencari cinta yang terluka Tapi takkan pernah kembali lagi Aku ingin kembali ke waktu yang lalu Dan Cinta kita yang dahulu sangat indah Tapi sekarang hanya tinggal luka Aku mencari cinta yang terluka Tapi takkan pernah kembali lagi	Dewasa
Cinta di Waktu yang Salah	Ku takkan pernah lupakan Masa-masa yang kita lalui Ku ingat kembali Waktu yang kita kenal Ku ingin kau tahu Ku cinta kamu lebih dari apapun Ku takkan pernah berpaling Karena kau adalah cinta yang sejati Ku ingin kau mengerti Ku takkan pernah melepaskan Karena kau adalah bagian dari jiwaku Ku ingin kau tahu Ku takkan pernah mengingkari Karena kau adalah cinta yang pertama Ku ingin kau mengingat Kita akan selalu bersama Dalam cinta yang kita miliki	Dewasa

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta

Lampiran 5 Sampel *Dataset* Lirik Lagu Sintetis (*Translation-Based*)

Judul	Lirik Lagu Sintetis (Buatan LLM)	Label Usia
Warna-Warna Pelangi	Merah, oranye, kuning, hijau dan biru Ini adalah warna yang saya cintai untuk dilakukan Menggambar dunia dengan nuansa yang cerah Karya masterpiece, hanya saya dan kamu Di pelangi, kita menemukan kebanggaan kita Di mana mimpi dan sihir tidak pernah surut Setiap warna memiliki cerita yang spesial Sebuah cerita untuk dibagi, tanpa kegagalan Ketika matahari bersinar, kita menari dan bermain Di bawah pelangi, pada hari yang lebih cerah Imajinasi kita tidak terbatas Di mana cinta dan kegembiraan bergema Di setiap warna, kita menemukan suara kita Menciptakan harmoni, membuat pilihan yang tepat Di warna-warna pelangi, kita bersatu Bersama, hati kita akan terbang	Semua usia
Detak Jantung Cinta	Di dalam irama cinta, hati kita berkelana Ketika irama cinta, menjadi cinta yang ilahi Dunia luar, mungkin mencoba memisahkan Tapi cinta kita, berdiri kuat dengan hati yang terbuka Cinta yang kita bagikan, adalah cinta yang kita berikan Tanpa syarat, atau harapan untuk hidup Di dalam irama cinta, hati kita terbang Ketika irama cinta, menjadi cinta yang begitu cerah Dunia mungkin menilai, tapi cinta kita akan berdiri Kita memilih untuk mencintai, dengan tangan yang terbuka Di dalam irama cinta, hati kita bebas Ketika irama cinta, menjadi cinta yang sejati	Semua usia
Sihr Pelangi	Warna-warna pelangi, bersinar dengan cerah Sihr di udara, pada malam yang ajaib Simfoni warna, dalam harmoni yang sempurna Sihr pelangi, keajaiban yang dapat dilihat Merah, oranye, kuning, hijau, dan biru Sihr pelangi, hanya menunggu Anda Ungu dan pink, dan semua sisanya Sihr pelangi, pada yang terbaik Jembatan warna, yang menghubungkan bumi dan langit Sihr pelangi, yang lewat Mimpi-mimpi dan keinginan, pada angin sepoi-sepoi Sihr pelangi, membawa kegembiraan dengan mudah Di dunia penuh keajaiban, di mana sihir adalah nyata Sihr pelangi, harta yang dapat dirasakan Di mana cinta dan sihir, selalu bersatu Sihr pelangi, harta yang sangat ilahi	Anak
Awan Ringan	Awan ringan bergerak, di langit yang biru Menonton mereka berubah, menjadi bentuk yang baru Lembut dan ringan, mereka bergerak perlahan Awan ringan berbisik, dengan nada yang lembut Mereka membawa mimpi, dan rahasia dari atas Awan ringan berbisik, dengan cinta yang lembut Mereka seperti binatang, atau mungkin hanya wajah Awan ringan bercerita, tentang ruang yang luas Bayang-bayang mereka menari, di atas rumput yang hijau Awan ringan berbisik, dalam adegan yang damai Mereka akan kembali, ke tempat mereka berada Di mana matahari bersinar, dan cinta akan bersinar	Anak

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta



Ke Dalam Jurang	Kegelapan menutupi saya, seperti kain yang dibungkus Sebuah perasaan putus asa, yang menghimpit hati saya Dunia luar sangat kejam, dengan suara-suara yang dingin Sebuah rasa pahit kenyataan, yang sulit diungkapkan Saya hilang dalam labirin, pikiran saya sendiri Sebuah maze emosi yang rumit, di mana bayangan memiliki kuasa Jurang memanggil saya, dengan bisikan yang manis Sebuah nyanyian sirene kesunyian, yang bergema sepanjang tahun Saya berjalan di atas garis tipis, antara cahaya dan kegelapan Sebuah benang harapan yang rapuh, yang memegang sebentar Ke dalam jurang saya jatuh, di mana kesunyian adalah sahabat saya Sebuah perjalanan kesendirian, di mana hati saya tidak pernah berakhir	Remaja
Halaman yang Tak Tertulis	Bab ini belum selesai, cerita ini belum ditulis Halaman ini masih kosong, kata-kata ini belum diucapkan Saya mencari pena, saya mencari suara Saya mencoba mengisi halaman, saya mencoba menulis nama saya Saya bukan karakter, saya bukan pion Saya adalah penulis takdir saya, saya adalah yang menggambar Saya menulis skenario saya sendiri, saya menulis skor saya Saya mengambil kendali, saya keluar dari pintu Saya bukan korban, saya bukan bodoh Saya adalah pahlawan cerita saya, saya adalah yang keren Saya menulis takdir saya sendiri, saya menulis nama saya Saya mengambil kembali kekuatan saya, saya mematahkan rantai Saya mengisi halaman saya sendiri, saya menulis jalan saya	Remaja
Serenade Senja	Matahari terbenam, di atas bukit yang bergulir Malam yang damai, dengan cinta yang sejati Bintang-bintang bersinar, di langit malam Ketika saya memetik gitar, dan menyanyikan lagu ini Musik berbisik lembut, tentang cinta yang sejati Ketika saya menyanyikan lagu ini, untuk kamu Dunia ini berhenti, dan cinta kita akan tumbuh Di dalam kesunyian, cinta kita akan bersinar Saya akan menyanyikan lagu ini, untuk kamu Dan dalam kesunyian, cinta kita akan bersinar Di malam yang damai, kita akan menari Di bawah bintang yang bersinar, kita akan bersinar cerah	Dewasa
Nyala Cinta Abadi	Dalam api cinta kita, nyala akan tetap menyala Cinta yang tidak akan pernah hilang, cinta yang akan tetap tahun Sisa-sisa hati kita, sebuah percikan yang akan tetap menyala Cinta yang akan tetap tinggal, dalam diriku Aku akan memegang cinta, yang kita pernah miliki Dan menghargai momen, yang kita pernah bagikan Dalam nyala cinta abadi, aku akan menemukan jalanku Ke tempat cinta akan menyembuhkan, dan menunjukkan jalanku ke cahaya Nyala cinta abadi, akan tetap tinggal Sebuah cinta yang akan tetap tinggal, dalam diriku Nyala cinta abadi, akan menunjukkan jalanku Ke tempat cinta akan menyembuhkan, dan menunjukkan jalanku ke cahaya	Dewasa

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Politeknik Negeri Jakarta