

REKOMENDASI LAGU BERDASARKAN JENIS MOOD MENGGUNAKAN TF-IDF DAN COSINE SIMILARITY

Callista Virginia ¹⁾, Christy ²⁾, Hafiz Irsyad ^{3)*}

^{1,2,3} Informatika, Ilmu Komputer dan Rekayasa, Universitas Multi Data Palembang
Jl. Rajawali No.14, 9 Ilir, Kec. Ilir Tim. II, Kota Palembang, Sumatera Selatan 30113
¹callistavirginia_2226250118@mhs.mdp.ac.id, ²christy_2226250005@mhs.mdp.ac.id,
^{3*}hafizirsyad@mdp.ac.id

ABSTRAK

Sistem rekomendasi lagu saat ini semakin berkembang seiring dengan meningkatnya kebutuhan pengguna akan konten yang sesuai dengan suasana hati (*mood*). Penelitian ini bertujuan untuk membangun sistem rekomendasi lagu berdasarkan jenis *mood* pengguna dengan memanfaatkan teknik pengolahan teks, yaitu *Term Frequency-Inverse Document Frequency* (TF-IDF) dan *Cosine Similarity*. Metodologi yang digunakan melibatkan tahap *preprocessing* teks lirik lagu, ekstraksi fitur menggunakan TF-IDF, serta penghitungan tingkat kemiripan antara input deskripsi *mood* dari pengguna dan koleksi lirik lagu menggunakan *Cosine Similarity*. Dataset yang digunakan berupa kumpulan lirik lagu yang telah dikategorikan ke dalam beberapa label *mood*, seperti *sadness*, *joy*, *anger*, dan *love*. Hasil penelitian menunjukkan bahwa sistem dapat merekomendasikan lagu yang relevan dengan deskripsi *mood* pengguna secara cukup akurat, dengan tingkat akurasi mencapai 90%, nilai *precision* mencapai 91,8%, nilai *recall* mencapai 90% dan nilai *F1-Score* mencapai 86,9% berdasarkan evaluasi yang telah dilakukan. Simpulan dari penelitian ini menyatakan bahwa kombinasi TF-IDF dan *Cosine Similarity* efektif untuk membangun sistem rekomendasi lagu berbasis teks yang sederhana namun fungsional.

Kata kunci : *Cosine Similarity*, Lirik Lagu, *Mood*, Rekomendasi Lagu, Sistem Cerdas, TF-IDF.

ABSTRACT

The music recommendation system has been evolving in response to the increasing demand for content that aligns with users' moods. This study aims to develop a music recommendation system based on users' mood types by utilizing text processing techniques, specifically Term Frequency-Inverse Document Frequency (TF-IDF) and Cosine Similarity. The methodology involves preprocessing song lyrics, feature extraction using TF-IDF, and calculating the similarity between users' mood descriptions and the song lyric collection using Cosine Similarity. The dataset consists of song lyrics categorized into various mood labels, such as sadness, joy, anger, and love. The results show that the system can accurately recommend songs that match users' mood descriptions, achieving an accuracy of 90%, precision of 91.8%, recall of 90%, and an F1-Score of 86.9% based on the evaluation. The conclusion of this research states that the combination of TF-IDF and Cosine Similarity is effective in building a simple yet functional text-based music recommendation system.

Keywords: *Cosine Similarity, Intelligent System, Mood, Song Lyrics, Song Recommendation, TF-IDF.*

PENDAHULUAN

Dalam kehidupan manusia, musik tidak hanya berperan sebagai sarana hiburan, tetapi juga sebagai sarana untuk mengekspresikan emosi dan pikiran. Di zaman modern ini, lagu dan musik telah menjadi salah satu media hiburan yang populer dan banyak digemari oleh masyarakat, mulai dari anak-anak, remaja, hingga orang dewasa. Oleh karena itu, musik dan lagu menjadi salah satu media yang sering digunakan dan diputar dalam berbagai situasi, misalnya saat berkendara, mengerjakan tugas, atau bahkan saat hendak tidur.

Di era digital saat ini, musik terus berkembang pesat mengikuti perkembangan teknologi yang ada. Musik yang sebelumnya hanya terdiri dari beberapa genre, kini telah menjadi berbagai genre, mulai dari *jazz*, *pop*, *hip hop*, klasik, dan masih banyak lagi. Selain itu, sistem teknologi rekomendasi lagu juga sudah mulai dikembangkan sehingga memudahkan kita untuk menemukan jenis lagu yang kita inginkan. “Jenis genre musik menjadi lambang dari suasana hati manusia, kalau lagi sedih yang dengar musiknya yang sedih-sedih, sedangkan kalau lagi bahagia dengar musiknya yang sesuai dengan isi hatinya.” (Wildani et al., 2023)

Menurut Ramadhan et al. (2024), “sistem rekomendasi musik adalah bidang penelitian yang berkembang pesat, terutama dengan meningkatnya popularitas layanan streaming musik.” “Layanan streaming musik menjadi metode paling populer bagi konsumen untuk mendengarkan musik. Layanan streaming menawarkan konsumen akses tak terbatas ke katalog musik besar.” (Noviani et al., 2020).

Streaming musik merupakan salah satu layanan yang disediakan oleh sistem jaringan atau aplikasi musik. Layanan ini memungkinkan kita untuk memutar berbagai jenis lagu secara langsung tanpa harus mengunduhnya terlebih dahulu. Akan tetapi, saat ini masih banyak orang yang belum menemukan lagu yang tepat berdasarkan suasana hati atau emosinya. Kebanyakan orang juga hanya mengingat atau menemukan satu hingga tiga lagu berdasarkan suasana hati tertentu atau mudah bosan dengan lagu yang itu-itu saja. Oleh karena itu, sistem rekomendasi lagu berdasarkan suasana hati menjadi salah satu solusi yang dapat

dikembangkan agar kita dapat menemukan berbagai ide pilihan lagu sesuai dengan suasana hati atau emosi tertentu.

Metode *Term Frequency-Inverse Document Frequency* (TF-IDF) merupakan salah satu teknik yang umum digunakan dalam analisis teks. “Algoritma *Term Frequency-Inverse Document Frequency* (TF-IDF) merupakan algoritma yang digunakan untuk menghitung bobot kata dengan cara mempertimbangkan ada berapa banyak kata muncul (frekuensi kata) dan berapa banyak kata tersebut ditemukan dalam dokumen.” (Widaningrum et al., 2022). Dalam konteks sistem rekomendasi lagu, TF-IDF akan mengubah lirik lagu dalam bentuk teks menjadi representasi numerik (vektor) yang mencerminkan pentingnya kata-kata tertentu dalam lirik tersebut, seperti kata-kata yang mencerminkan kebahagiaan, kesedihan, kemarahan atau semangat.

Setelah itu, *Cosine Similarity* digunakan untuk menghitung tingkat kesamaan antara *mood* yang diinputkan pengguna dengan lirik lagu dalam dataset untuk menemukan lagu yang paling relevan. “*Cosine Similarity* adalah metode yang digunakan untuk mengukur kesamaan antara dua vektor teks berdasarkan sudut *cosinus* yang terbentuk di antara keduanya. Teknik ini sangat efektif dalam berbagai aplikasi, seperti pencarian informasi, sistem rekomendasi, dan pengelompokan dokumen” (Putra et al., 2024). Dengan pendekatan ini, sistem diharapkan dapat mengidentifikasi lagu-lagu dengan *mood* yang mirip berdasarkan kesamaan struktur liriknya.

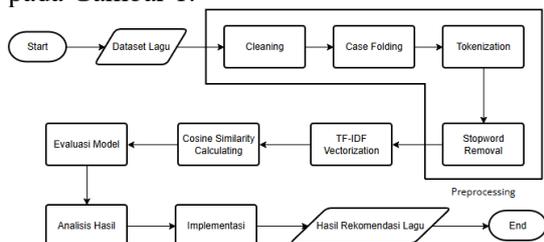
Berdasarkan paparan diatas, peneliti ingin mengangkat topik berjudul “Rekomendasi Lagu Berdasarkan Jenis Mood Menggunakan TF-IDF dan *Cosine Similarity*”. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi lagu berdasarkan *mood* menggunakan metode TF-IDF dan *Cosine Similarity*. Sistem ini juga diharapkan tidak hanya membantu pengguna menemukan lagu-lagu yang sesuai dengan *mood* mereka tetapi juga memberikan berbagai ide lagu baru yang mungkin belum pernah mereka temukan.

METODE PENELITIAN

Penelitian ini menggunakan metode eksperimental untuk mengembangkan dan

mengevaluasi sistem rekomendasi lagu. “Penelitian eksperimen merupakan salah satu metode penelitian yang bertujuan untuk menguji hubungan sebab akibat antara variabel-variabel yang diteliti. Dalam penelitian ini, peneliti secara sengaja memberikan perlakuan (*treatment*) pada subjek penelitian dan kemudian mengukur pengaruhnya terhadap variabel lain.” (Yulianti et al., 2024). Metode eksperimental digunakan karena melibatkan serangkaian langkah sistematis untuk merancang, mengimplementasikan, dan menguji sistem. Fokusnya adalah pada pengujian efektivitas algoritma (TF-IDF dan *Cosine Similarity*) dalam mencocokkan *mood* pengguna dengan lirik lagu.

Penelitian ini terbagi menjadi beberapa tahapan yaitu, pengumpulan data, *pre-processing* data, implementasi algoritma TF-IDF dan *Cosine Similarity*, evaluasi sistem, dan pengembangan antarmuka. Tahapan perancangan sistem dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Metode Ilmiah

Pengumpulan data dilakukan dari situs *website* Kaggle menggunakan dataset Spotify yang berjumlah 498.052 data dengan 39 jenis fitur yang mencakup atribut seperti *artist, song, text, length, emotion, genre, album, release date, key, tempo, loudness, time signature, explicit, popularity, energy, danceability, positiveness, speechiness, liveness, acousticness, instrumentalness*, dan lainnya yang dapat dilihat pada Tabel 1.

Tabel 1. Sampel Data

Kolom	Type Data	Kolom	Type Data
Artist(s)	object	Instrumentalness	float64
song	object	Good for Party	float64
text	object	Good for Work/Study	float64
Length	object	Good for Relaxation/Meditation	float64
emotion	object	Good for Exercise	float64
Genre	object	Good for	float64

		Running	
Album	object	Good for Yoga/Stretching	float64
Release Date	object	Good for Driving	float64
Key	object	Good for Social Gatherings	float64
Tempo	float64	Good for Morning Routine	float64
Loudness (db)	float64	Similar Artist 1	object
Time Signature	object	Similar Song 1	object
Explicit	object	Similarity Score 1	float64
Popularity	int64	Similar Artist 2	object
Energy	int64	Similar Song 2	object
Danceability	int64	Similarity Score 2	float64
Positiveness	int64	Similar Artist 3	object
Speechiness	int64	Similar Song 3	object
Liveness	int64	Similarity Score 3	object
Acousticness	float64		

Setelah dataset dikumpulkan, tahapan *preprocessing* yang dilakukan berupa *case folding, stopword removal, dan tokenization* untuk membersihkan dan menyiapkan lagu sebelum proses representasi dan klasifikasi *mood* lagu. Kemudian, teks yang telah diproses akan direpresentasikan dalam bentuk vektor dengan menggunakan metode TF-IDF. Algoritma TF-IDF akan menghitung frekuensi kemunculan setiap kata dalam teks relatif terhadap keseluruhan dataset. TF-IDF adalah sistem pembobotan yang memberikan nilai bobot pada setiap kata dalam dokumen berdasarkan frekuensi kemunculan kata (TF) dan frekuensi resiprokal dokumen (IDF). Kata-kata yang memiliki skor bobot lebih tinggi dianggap lebih penting. “TF-IDF memberikan bobot lebih tinggi pada kata-kata yang sering muncul dalam satu dokumen tetapi jarang muncul di dokumen lain, karena kata-kata seperti itu lebih signifikan dalam membedakan dokumen satu dengan yang lainnya.” (Putra et al., 2024)

Setelah *preprocessing* dan representasi teks dalam bentuk TF-IDF *matrix*, dilakukan penghitungan kemiripan antara *mood* dan lagu menggunakan *Cosine Similarity* berdasarkan vektor TF-IDF yang telah dihitung sebelumnya. Hasilnya akan digunakan dalam fungsi *recommend_song* untuk memberikan rekomendasi lagu berdasarkan input *mood* pengguna, dengan mencocokkan rata-rata vektor TF-IDF dari lagu-lagu dengan *mood* yang sama.

$$tf - idf(t, d) = tf(t, d) * idf(t) \quad (1)$$

$$idf(t) = \log \left(\frac{N}{df(t)} \right) \quad (2)$$

Lagu-lagu yang memiliki nilai *similarity* tertinggi (berdasarkan skor *Cosine Similarity*) akan direkomendasikan kepada pengguna. Lima lagu teratas dengan skor *similarity* tertinggi dipilih untuk ditampilkan dalam output. Semakin kecil sudut antara dua vektor, semakin besar nilai *Cosine Similarity* berarti kedua vektor tersebut lebih mirip satu sama lain. Jika nilai *Cosine Similarity* mendekati 1 artinya kedua vektor sangat mirip atau bahkan identik dalam hal orientasi atau arah. Sebaliknya, jika nilai *Cosine Similarity* mendekati 0 berarti kedua vektor tersebut sangat berbeda.

$$\cos \theta = \frac{A \cdot B}{\|A\| \|B\|} \quad (3)$$

Tahapan selanjutnya adalah evaluasi model untuk mengukur kinerja sistem rekomendasi dengan menggunakan metrik evaluasi berupa *accuracy*, *precision*, *recall*, dan *F1-score*. *Mood* yang telah diprediksi oleh sistem akan dibandingkan dengan label *mood* yang asli dari dataset. Berdasarkan perbandingan tersebut, akan dihitung nilai *accuracy*, *precision*, *recall*, dan *F1-score* menggunakan fungsi dari scikit-learn pada Google Collab.

Kemudian, digunakan *Confusion Matrix* untuk memvisualisasikan perbandingan distribusi antara prediksi yang benar dan salah yang membantu dalam identifikasi pola kesalahan yang ada. “*Confusion matrix* adalah tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah.” (Normawati et al., 2021).

Tabel 2. Confusion Matrix

		Prediksi	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

Selain visualisasi dengan *Confusion Matrix*, hasil evaluasi juga ditampilkan dalam bentuk kurva *precision* dan *recall* untuk memberikan informasi tambahan mengenai kinerja model dalam sistem.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \quad (6)$$

Pemanfaatan kurva *precision* dan *recall* dipilih karena sistem ini

menitikberatkan pada ketepatan hasil rekomendasi, bukan hanya banyaknya prediksi yang benar. Saat menghadapi dataset yang distribusi mood-nya tidak merata, tingkat akurasi menjadi kurang representatif.

Precision menunjukkan tingkat relevansi hasil rekomendasi, sedangkan *recall* mencerminkan kemampuan sistem dalam menampilkan semua lagu yang sesuai.

F1-score merupakan metrik evaluasi yang menggabungkan *precision* dan *recall* dalam satu nilai harmonis. Ini berguna saat data tidak seimbang karena memberikan gambaran seimbang antara ketepatan dan kelengkapan hasil rekomendasi. Nilai *F1-score* tinggi menunjukkan bahwa sistem tidak hanya akurat dalam merekomendasikan lagu yang relevan, tetapi juga mampu menangkap sebagian besar lagu yang sesuai dengan mood pengguna.

HASIL DAN PEMBAHASAN

Proses pengolahan data diawali dengan tahap *data cleaning* yaitu menggabungkan beberapa file menjadi satu kesatuan dan penanganan nilai yang hilang (*missing value*) yang bertujuan untuk mengoptimalkan data. Pada tahap ini, dilakukan pembersihan nama kolom dari karakter yang tidak terlihat atau *Byte Order Mark* (BOM) dan spasi yang berlebihan. Selain itu juga, kolom dengan nilai hilang (*missing value*) diisi sesuai dengan nilai default yang sesuai contohnya mengisi kolom *emotion* dengan ‘Unknown’, kolom *song* dengan ‘Unknown Song’, kolom *Artist(s)* dengan ‘Unknown Artist’ dan kolom *text* dengan string kosong. Hal ini bertujuan agar model tidak terganggu oleh data kosong.

Setelah data bersih diperoleh, selanjutnya adalah tahap ekstraksi fitur dengan menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Proses ini digunakan untuk mengubah teks menjadi representasi numerik dalam bentuk matriks. Penggunaan metode ini penting diterapkan dikarenakan dapat diukur seberapa pentingnya suatu kata dalam sebuah lirik lagu dalam konteks keseluruhan koleksi dataset. Menurut Chihab (2021), dalam teori *Information Retrieval* (IR), metode TF-IDF berperan penting dalam meningkatkan performa klasifikasi teks karena mampu

merepresentasikan kata-kata yang paling informatif dari suatu dokumen. Teknik ini secara efektif menonjolkan istilah-istilah yang memiliki nilai pembeda tinggi antar dokumen. Hasil dari transformasi TF-IDF berupa *sparse matrix* (matriks jarang), yakni bentuk penyimpanan data yang efisien di mana sebagian besar elemennya bernilai nol. Matriks ini hanya menyimpan nilai-nilai yang memiliki bobot, disertai dengan koordinat baris (dokumen) dan kolom (kata) yang sesuai. Sebagai contoh, nilai (0, 235) 0.07443871089562687 mengindikasikan bahwa dalam dokumen ke-0, kata dengan indeks ke-235 memiliki bobot TF-IDF sebesar 0.0744, menunjukkan tingkat kepentingan kata tersebut dalam dokumen tersebut.

Secara keseluruhan, terdapat 829 nilai non-nol dalam matriks berukuran 16 baris dan 588 kolom, yang berarti terdapat 16 dokumen (lirik lagu) dan 588 fitur (kata unik) yang diekstraksi dari korpus. Kondisi ini menunjukkan bahwa setiap dokumen hanya memuat sebagian kecil dari keseluruhan kosakata, sehingga sebagian besar sel bernilai nol. Oleh karena itu, penggunaan *sparse matrix* menjadi sangat tepat dan efisien dalam merepresentasikan data teks pada tahap ini. Setelah vektorisasi data, *resampling* data digunakan untuk menyeimbangkan distribusi kelas misalnya jika beberapa *mood* seperti *joy* atau *anger* lebih sering muncul dibandingkan dengan jenis *mood* lainnya, maka akan menyebabkan distribusi kelas yang tidak seimbang dan bisa menyebabkan model menjadi lebih condong untuk memprediksi *mood* yang lebih sering muncul.

Proses tokenisasi data juga dilakukan untuk mengubah *text* menjadi bentuk yang lebih kecil seperti kata atau frasa yang lebih mudah diproses oleh model. Melalui proses ini, *text* akan dihitung dan dianalisis untuk menentukan kesamaan antar lagu misalnya jika *text* sebuah lagu "I feel happy today," maka setelah tokenisasi, kata-kata yang dihasilkan adalah: ["I", "feel", "happy", "today"].

Sistem ini menggunakan metode *Cosine Similarity* untuk mengukur kesamaan antara vektor lirik lagu dengan vektor *mood* yang di *input* oleh pengguna. Chihab (2021) juga menjelaskan bahwa algoritma *Cosine Similarity* digunakan untuk mengukur tingkat kemiripan antara dua vektor dalam hal ini,

antara vektor representasi lirik lagu dan vektor *mood* pengguna. Nilai *similarity* yang lebih tinggi menunjukkan tingkat relevansi emosional yang lebih kuat antara lagu dan suasana hati pengguna. Sebagai contohnya nilai 0.19185489 antara dua lagu menunjukkan bahwa kedua lirik memiliki kemiripan emosional yang cukup tinggi. Sebaliknya, nilai 0.0 menunjukkan tidak ada kesamaan konten yang berarti antar lirik tersebut.

Setelah data selesai diproses, sistem rekomendasi lagu dirancang dengan menggunakan antarmuka Gradio yang telah tersedia di Google Collaboratory yang memungkinkan pengguna untuk memilih *mood* mereka dan menerima hasil rekomendasi lagu berdasarkan *mood* yang dipilih. Sistem ini menggunakan fungsi *recommend_song* yang menyaring lagu berdasarkan *mood* dan menghitung *Cosine Similarity* antara representasi dari metode TF-IDF setiap lagu dengan vektor representasi *mood* yang dipilih.

Sebagai contoh hasil, untuk *mood* "sadness", sistem mengembalikan lima lagu dengan tingkat kesamaan tertinggi, di antaranya seperti lagu *The Box* dari (Hed) P.E. dengan tingkat kemiripan 0.4606 dan lagu *Bringin' Da Noise* dari *NSYNC dengan kemiripan 0.4520. Sementara untuk *mood* "joy", direkomendasikan lagu *Mashallah* dari *\$kinny* dengan kemiripan 0.5231 serta *FREE* dari #TheSet dengan nilai 0.4762.

Untuk beberapa *mood* seperti "anger" dan "love", sistem mendeteksi satu lagu yang paling relevan, masing-masing dengan tingkat kemiripan sempurna (1.0000), yaitu lagu *Initials* oleh *Hair Tribe* untuk *mood* "anger" dan *Champagne and Cocaine* dari *12 Dirty Bullets* untuk *mood* "love".

Implementasi algoritma TF-IDF dan *Cosine Similarity* dalam menyarankan lagu berdasarkan suasana hati tidak terlepas dari konsep psikologi musik, terutama teori *Mood Congruency Effect*. Teori ini menjelaskan bahwa seseorang cenderung memilih musik yang sejalan dengan kondisi emosional yang sedang dialaminya. Hauck dan Hecht (2023) mengemukakan bahwa keselarasan antara mood pada musik dan teks dapat memperkuat keterlibatan emosional pendengar. Demikian pula, hasil studi Zhang dan Asano (2024) memperlihatkan bahwa kesesuaian antara

aspek emosional seperti valensi dan tingkat rangsangan dari audio dan visual mampu meningkatkan resonansi afektif dalam pengalaman pengguna. Dengan demikian, sistem ini secara tidak langsung memanfaatkan kekuatan emosi dalam lirik untuk menghadirkan rekomendasi yang lebih personal dan menyentuh aspek afektif pendengar.

Kemudian, dilakukan analisis terhadap hasil rekomendasi lagu yang diberikan berdasarkan input *mood* dari pengguna.

Rekomendasi Lagu untuk mood 'sadness':
 1. The Box - (Hed) P.E. (Similarity: 0.4606)
 2. Bringin' Da Noise - *NSYNC (Similarity: 0.4520)
 3. I Drive Myself Crazy - *NSYNC (Similarity: 0.4383)
 4. Forsaken - 100 Demons (Similarity: 0.4258)
 5. LEGO - 1 800 PAIN (Similarity: 0.4161)

Gambar 2. Rekomendasi Lagu berdasarkan Mood Sadness

Rekomendasi Lagu untuk mood 'joy':
 1. Mashallah - \$kinny (Similarity: 0.5231)
 2. FREE - #TheSet (Similarity: 0.4762)
 3. Daddys Girl - 1 Girl Nation (Similarity: 0.4550)
 4. CAKE - 1 800 PAIN (Similarity: 0.4454)
 5. SHELF - 1 800 PAIN (Similarity: 0.4288)

Gambar 3. Rekomendasi Lagu berdasarkan Mood Joy

Dari kedua perbandingan hasil sistem rekomendasi lagu yang ditambahkan dalam Gambar 2 dan Gambar 3 terdapat beberapa hal yang dapat dianalisis. Pertama, dapat dilihat hasil rekomendasi untuk setiap *mood* yang diinginkan menunjukkan variasi yang cukup signifikan. Gambar 2 merekomendasikan untuk *mood sadness* dan Gambar 3 merekomendasikan untuk *mood joy*. Hal ini menunjukkan bahwa algoritma TF-IDF sudah bisa membedakan dan merespon input yang berbeda dengan hasil rekomendasi yang relevan sesuai dengan *mood* yang diinginkan oleh penggunanya.

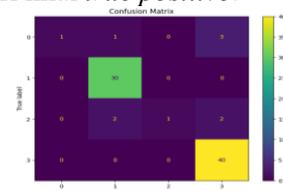
Kedua, *similarity score* yang diberikan untuk setiap rekomendasi lagu menunjukkan seberapa relevan lagu tersebut dengan *mood* yang diinput. Dalam rekomendasi lagu untuk *mood sadness*, lagu 'The Box' oleh (Hed) P.E. memiliki *similarity score* tertinggi sebesar 0.4606, sedangkan untuk lagu 'Mashallah' yang direkomendasikan untuk *mood joy* memiliki *similarity score* tertinggi dari *mood* tersebut sebesar 0,5231. Perbedaan *score* ini mengindikasikan bahwa algoritma TF-IDF lebih efektif dalam merekomendasikan lagu untuk

mood joy daripada *sadness*.

Ketiga, dapat dilihat bahwa algoritma TF-IDF cukup sensitif dalam mengidentifikasi variasi dari lagu yang sama. Hal ini dapat dilihat seperti untuk *mood sadness* terdapat lagu 'I Drive Myself Crazy' yang dipopulerkan oleh NSYNC dengan *similarity score* 0.4383 yang menunjukkan bahwa variasi lagu yang relevan meskipun ada perbedaan dalam versi atau *artist* yang berbeda pula. Hal ini menunjukkan bahwa metode TF-IDF memang dapat mengenali kemiripan lirik lagu meskipun ada perbedaan dalam versi lagu.

Evaluasi sistem juga dilakukan dengan menggunakan metrik-metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score* yang dihitung berdasarkan perbandingan antara *mood* yang sebenarnya (*true labels*) dan *mood* yang diprediksi oleh sistem. Akurasi 0.90 menunjukkan bahwa 90% dari rekomendasi yang diberikan oleh sistem relevan dengan *mood* yang diminta. *Precision* 0.918 berarti 91,8% dari lagu yang direkomendasikan benar-benar relevan dengan *mood* yang diminta, sementara 8,2% tidak sesuai. *Recall* 0.9 menunjukkan bahwa 90% lagu relevan yang ada dalam dataset berhasil ditemukan oleh sistem untuk *mood* tertentu. Sementara, *F1-Score* dengan 0.8694 menunjukkan keseimbangan antara *precision* dan *recall*, menunjukkan bahwa sistem cukup baik dalam memberikan rekomendasi yang relevan dan menemukan lagu-lagu yang sesuai dengan *mood* yang diminta.

Confusion matrix memberikan bagaimana model melakukan prediksi untuk setiap kelas *mood*, hasilnya menunjukkan bahwa model memiliki kesusahan dalam mengelompokkan beberapa *mood* seperti *sadness* dan *anger*. Walaupun *mood joy* dengan 30 prediksi benar dan *love* dengan 40 prediksi benar, menunjukkan performa yang lebih baik tetapi masih mempunyai kesalahan klasifikasi misalnya terdapat kesalahan prediksi untuk *mood joy* dan *sadness* yang terlihat dari nilai *true positive*.



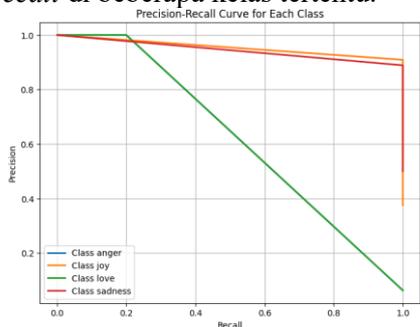
Gambar 4. Confusion Matrix

Hasil rekomendasi sebanyak lima lagu teratas yang memiliki kesamaan paling tinggi dengan *mood* yang dipilih oleh pengguna. Penggunaan *Cosine Similarity* memungkinkan sistem untuk memberikan rekomendasi lagu yang lebih relevan berdasarkan liriknya.



Gambar 5. Interface Rekomendasi User

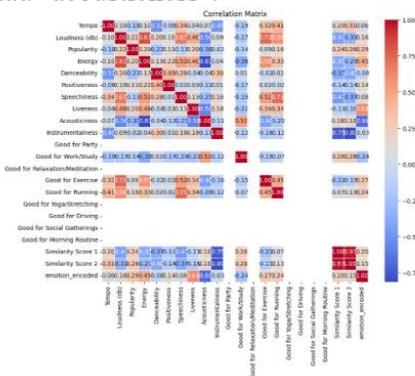
Kurva *Precision-Recall* memberikan gambaran tentang *trade-off* antara *precision* dan *recall* pada setiap *mood*. *Precision* dimanfaatkan untuk mengukur tingkat kebenaran dari rekomendasi yang diberikan oleh sistem, sementara *recall* digunakan untuk mengukur seberapa banyak identifikasi model memberikan rekomendasi yang relevan. Berdasarkan Gambar 6, dapat dilihat bahwa model memiliki performa lebih baik pada *mood joy* dan *love* dengan *precision* yang lebih tinggi. Penurunan *precision* yang tajam seiring dengan peningkatan *recall* menunjukkan bahwa model terlalu fokus pada *recall* di beberapa kelas tertentu.



Gambar 6. Kurva Precision-Recall

Correlation matrix menunjukkan hubungan antara berbagai fitur dalam dataset yang mencakup atribut kolom numerik seperti *loudness*, *energy*, *tempo*, *energy* dan fitur terkait lainnya. Dari hasil *correlation matrix* pada Gambar 7, adanya hubungan yang signifikan antara atribut seperti "good for party" dan "energy" atau "danceability" yang saling berkorelasi tinggi. Hal ini menunjukkan bahwa atribut ini saling mendukung dalam menentukan *mood* seperti *joy* atau *love* yang lebih energik. Sebaliknya, jika *mood* yang lebih tenang seperti *sadness* lebih terkait dengan atribut seperti "good for relaxation"

atau "acousticness".



Gambar 7. Correlation Matrix

PENUTUP

Penelitian ini menunjukkan bahwa sistem rekomendasi lagu berdasarkan *mood* menggunakan metode TF-IDF dan *Cosine Similarity* telah berhasil memberikan rekomendasi lagu yang sesuai dengan *mood* pengguna. Sistem ini dapat mengidentifikasi kesamaan antara *mood* yang diisi oleh pengguna dengan lirik lagu dalam dataset tersebut sehingga bisa menghasilkan rekomendasi yang lebih tepat sesuai dengan *emotion* atau suasana hati yang diinginkan. Keberhasilan sistem ini ditunjukkan melalui hasil evaluasi kinerja model dengan metrik kuantitatif yang mencakup akurasi sebesar 0.90, *precision* sebesar 0.918, *recall* sebesar 0.90, dan *F1-score* sebesar 0.8694. Hasil ini menunjukkan keseimbangan yang baik dalam menemukan rekomendasi lagu sesuai dengan *mood* pengguna.

Penerapan algoritma TF-IDF juga terbukti efektif dalam menghitung kata dan mengenali kesamaan struktur lirik lagu, yang terlihat dari hasil transformasi berupa 829 nilai non-nol dari total 9.408 elemen matriks (16 dokumen × 588 fitur). Nilai TF-IDF tertinggi tercatat sebesar 0.2975, menunjukkan bahwa kata tersebut memiliki pengaruh signifikan dalam representasi dokumen, sedangkan sebagian besar kata memiliki bobot lebih kecil, seperti 0.0744, yang mencerminkan tingkat kepentingannya dalam konteks lirik tertentu. Lagu dengan skor *similarity* 0.5231 memiliki kemiripan yang lebih tinggi terhadap *mood* tertentu dibandingkan lagu dengan skor 0.4288. Implementasi algoritma TF-IDF dan *Cosine Similarity* ini sangat relevan untuk layanan *streaming* musik dimana sistem ini dapat

memuaskan pengguna dengan rekomendasi lagu yang sesuai dengan *mood* mereka.

Namun, penelitian ini masih membuka peluang untuk pengembangan lebih lanjut. Penelitian selanjutnya dapat mengarah pada integrasi data pengguna yang lebih lengkap, seperti riwayat aktivitas mendengarkan musik dan penerapan algoritma berbasis *deep learning* untuk meningkatkan akurasi dan personalisasi dalam memberikan rekomendasi lagu. Hal ini akan memungkinkan sistem untuk terus berkembang, menjadi lebih efektif dan memberikan manfaat yang lebih besar bagi pengguna layanan *streaming* musik.

Selain itu, penggunaan jumlah data yang lebih banyak juga dapat dipertimbangkan agar model dapat memberikan rekomendasi yang lebih akurat dan bervariasi. Karena dalam penelitian ini kami hanya menggunakan 1.605 data dari 498.052 dataset. Hal tersebut dilakukan karena jika menggunakan semua data, maka beban komputasi menjadi berat karena pengelolaan data yang besar serta keterbatasan memori kapasitas pemrosesan perangkat yang digunakan. Jika menggunakan keseluruhan dataset, maka dapat berpotensi menyebabkan keterlambatan waktu yang signifikan dari berbagai proses analisis baik dari segi waktu eksekusi maupun penggunaan sumber daya perangkat keras.

DAFTAR PUSTAKA

Cespedes-Guevara, J. dan Dibben, N. 2022. *Music and Emotion: The Role of Visual Imagery and Bodily Imagery in Emotional Experience with Music. Musicae Scientiae*. Vol 26 No 1. Hal 75–92.

Chihab, Y. 2021. *Netflix Recommendation System Based on TF-IDF and Cosine Similarity Algorithms. Prosiding 2nd International Conference on Big Data, Modelling and Machine Learning (BML 2021)*. Hal 15–20.

<https://www.scitepress.org/Papers/2021/107275/107275.pdf>. Diakses tanggal 14 Mei 2025.

Hauck, P. dan Hecht, H. 2023. *Mood Congruency in Music and Text: How Emotion in Music and Text Interact in Songs. Psychology of Music*. Vol 51 No 2. Hal 427–443.

Normawati, D. dan Prayogi, S. A. 2021. *Implementasi Naive Bayes Classifier dan*

Confusion Matrix pada Analisis Sentimen berbasis Teks pada Twitter. Jurnal Sains Komputer & Informatika (J-SAKTI). Vol 5 No 2. Hal 697-711.

Noviani, D., Pratiwi, R., Silvianadewi, S., Alexandri, M. B., dan Hakim, M. A. 2020. *Pengaruh Streaming Musik terhadap Industri Musik di Indonesia. Jurnal Bisnis STRATEGI*. Vol 9 No 1. Hal 14–25.

Putra, R. R., Putri, N. A., dan Putra, A. D. 2024. *Teknik Cosine Similarity dan TF-IDF dalam Analisis Data. Serasi Media Teknologi*. Sumatera Barat.

Ramadhan, M. A., Najiyah, I., Abillutfi, R. M., Musaropah, R., dan Pramanik, N. D. 2024. *Implementasi dan Evaluasi Sistem Rekomendasi Musik Berbasis Lirik dengan Algoritma Term Frequency-Inverse Document Frequency (TF-IDF). Seminar Teknologi Majalengka (STIMA)*. Vol 8. Hal 1–6.

Völker, D. 2022. *Measuring BRECVEM: Development and Validation of an Inventory for the Assessment of Music-Evoked Emotions Mechanisms. Frontiers in Psychology*. Vol 13. Hal 1–15.

Widaningrum, I., Mustikasari, D., Arifin, R., Tsaqila, S. L., dan Fatmawati, D. 2022. *Algoritma Term Frequency-Inverse Document Frequency (TF-IDF) dan K-Means Clustering untuk Menentukan Kategori Dokumen. Prosiding SISFOTEK*. Vol 6 No 1. Hal 145–149.

Wildani, N., Sari, S., dan Narti, S. 2023. *Musik Hip-Hop sebagai Identitas Diri Komunitas Planktone.Ind. Jurnal ISO: Jurnal Ilmu Sosial, Politik, dan Humaniora*. Vol 3 No 2. Hal 163-168.

Yulianti, R., Syukrilla, W. A., Effendi, Febriyanti, T. L., Rahayu, D. S., Agung, B. H., et al. 2024. *Metode Penelitian Eksperimen: Konsep, Implementasi, dan Studi Kasus. PT. Mifandi Mandiri Digital*. Sumatera Utara.

Zhang, Z. dan Asano, Y. M. 2024. *Affect Congruence and Crossmodal Processing in Multimedia Experiences. Journal of Media Psychology*. Vol 36 No 1. Hal 10–21.