

ANALISIS PENGARUH FITUR EMOSI MUSIK TERHADAP POPULARITAS LAGU SPOTIFY MENGGUNAKAN RANDOM FOREST DAN SHAP

Giatika Chrisnawati¹⁾, Muhammad Taqy Hafizh²⁾, Hisar Maini Siregar³⁾, Faizah Nabilah⁴⁾

^{1,2,3,4}Prodi Teknologi Informasi, Teknik dan Informatika, Universitas Bina Sarana Informatika

Correspondence author: G. Chrisnawati, giatika.gcw@bsi.ac.id, Jakarta, Indonesia

Abstract

The rapidly growing digital music industry, enabled by the Spotify platform, enables in-depth analysis of the emotional factors that influence a song's popularity. This study aims to identify significant emotional features, analyze their influence on the number of streams, and develop a data-driven model for predicting song popularity. The research method is an experimental artificial intelligence (machine learning) approach that applies the Random Forest Regression algorithm combined with the Shapley Additive Explanations (SHAP) interpretability method. The research stages include collecting datasets from Kaggle, data preprocessing, model training, and evaluation using the Coefficient of Determination (R^2), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). The results show that valence, danceability, acousticness, and energy have the greatest influence on the number of streams, while instrumentalness has the least. Based on the Shapley Additive Explanations analysis, high emotional values, valence, and danceability contribute positively to song popularity. This study concludes that combining the Random Forest algorithm with the Shapley Additive Explanations method is effective for predicting and understanding the emotional factors that determine song success on Spotify.

Keywords : *prediction, music popularity, emotional features, Spotify, machine learning*

Abstrak

Industri musik digital yang berkembang pesat melalui platform Spotify memungkinkan analisis mendalam mengenai faktor emosional yang memengaruhi popularitas sebuah lagu. Penelitian ini bertujuan untuk mengidentifikasi fitur emosional yang signifikan, menganalisis pengaruhnya terhadap jumlah *stream*, serta membangun model prediksi popularitas lagu menggunakan pendekatan berbasis data. Metode penelitian yang digunakan adalah pendekatan eksperimental kecerdasan buatan (*machine learning*) dengan menerapkan algoritma *Random Forest Regression* yang dipadukan dengan metode interpretabilitas *SHapley Additive Explanations* (SHAP). Tahapan penelitian meliputi pengumpulan kumpulan data (*dataset*) dari Kaggle, pra-pemrosesan data, pelatihan model, serta evaluasi menggunakan metrik *Coefficient of Determination* (R^2), *Mean Absolute Error* (MAE), dan *Root Mean Square Error* (RMSE). Hasil penelitian menunjukkan bahwa fitur *valence*, *danceability*, *acousticness*, dan *energy* memiliki pengaruh paling signifikan terhadap jumlah *stream*, sementara fitur

instrumentalness tercatat memiliki pengaruh paling kecil. Berdasarkan analisis *SHapley Additive Explanations*, nilai emosional yang tinggi pada fitur *valence* dan *danceability* berkontribusi positif terhadap popularitas lagu. Penelitian ini menyimpulkan bahwa kombinasi algoritma *Random Forest* dan metode *SHapley Additive Explanations* efektif dalam memprediksi serta memahami faktor-faktor emosional yang menentukan kesuksesan lagu di platform Spotify.

Kata Kunci : prediksi, popularitas musik, fitur emosi, spotify, *machine learning*

A. PENDAHULUAN

Perkembangan teknologi digital telah membawa perubahan besar dalam industri musik. Kehadiran platform streaming seperti Spotify tidak hanya memungkinkan pengguna mengakses jutaan lagu secara instan, tetapi juga menyediakan data yang kaya mengenai karakteristik audio dan perilaku pendengar (I. K. D. Setiawan et al., 2026). Musik, selain sebagai bentuk hiburan dan ekspresi seni, merupakan media komunikasi emosional yang kompleks. Berbagai penelitian menunjukkan bahwa musik mampu membangkitkan, merepresentasikan, dan memodulasi emosi manusia secara signifikan, baik secara sadar maupun tidak sadar (Feng & Wang, 2025).

Musik dalam kehidupan sehari-hari berfungsi sebagai media untuk mengatur suasana hati, mendukung berbagai aktivitas, serta merepresentasikan identitas pendengar. Penelitian oleh (Loureiro et al., 2024), menunjukkan bahwa pemilihan musik secara personal berperan dalam proses regulasi emosi dan pembentukan identitas individu, di mana pendengar memanfaatkan musik untuk menyesuaikan mood, merefleksikan pengalaman personal, serta membangun keterikatan emosional dengan diri mereka sendiri. Popularitas musik kini menjadi indikator penting dalam industri kreatif karena berpengaruh pada strategi pemasaran, sistem rekomendasi, dan keputusan bisnis (Simalango et al., 2025). Penelitian empiris menunjukkan bahwa faktor-faktor seperti perilaku

pengguna, algoritma rekomendasi, dan konteks sosial budaya memiliki peran signifikan dalam menentukan tingkat popularitas musik, selain pengaruh promosi dan reputasi artis (Ye et al., 2024).

Seiring tersedianya data musik dalam jumlah besar, pendekatan analisis berbasis data menjadi sangat relevan untuk memahami dan memprediksi tren popularitas di era digital (Tannady et al., 2024). Dalam penelitian ini, *Random Forest Regressor* dipilih karena kemampuannya untuk mengenali pola hubungan yang tidak linier dan fitur-fitur yang saling berinteraksi secara kompleks dalam data. Dengan menyatukan hasil dari sejumlah pohon keputusan yang membentuk *ensemble*, metode ini dapat memberikan prediksi yang lebih konsisten dan akurat, terutama ketika pola data tidak dapat ditangani oleh model regresi linier sederhana. (A. Setiawan et al., 2024). Model ini juga menunjukkan performa yang baik pada data musik karena dapat mengolah variabel-variabel audio yang saling berhubungan.

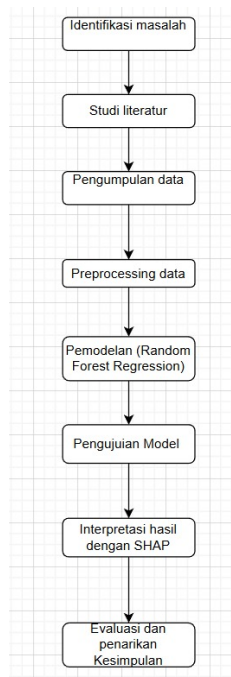
Penelitian ini memfokuskan analisis pada fitur-fitur emosi musik seperti *danceability*, *energy*, *valence*, *acousticness*, *speechiness*, *liveness*, dan *instrumentalness* untuk mengidentifikasi faktor yang paling berpengaruh terhadap jumlah *stream* lagu di Spotify. Dengan memanfaatkan *Random Forest* dan metode interpretabilitas *SHAP* (*SHapley Additive Explanations*), penelitian ini bertujuan untuk memberikan pemahaman yang lebih mendalam

mengenai kontribusi masing-masing fitur emosional terhadap popularitas musik.

Hasil penelitian diharapkan dapat memberikan kontribusi dalam pengembangan sistem rekomendasi berbasis emosi, mendukung strategi industri musik, serta memperkaya kajian ilmiah terkait analisis data musik menggunakan pendekatan *machine learning*.

B. METODE PENELITIAN

Penelitian ini dilakukan melalui serangkaian tahapan sistematis yang dimulai dari pengumpulan data hingga interpretasi hasil model. Fokus utama penelitian ini terletak pada penerapan algoritma *Random Forest* sebagai metode pemodelan prediktif dan *SHapley Additive exPlanations* (SHAP) sebagai metode interpretasi model. Alur kerja penelitian secara lengkap digambarkan pada Gambar 1.



Gambar 1. Tahapan Penelitian

Adapun tahapan dalam penelitian ini adalah sebagai berikut:

Identifikasi Masalah

Tahapan penelitian ini diawali dengan identifikasi masalah yang berfokus pada tantangan dalam proses pemodelan prediktif menggunakan dataset yang tersedia. Permasalahan yang dikaji tidak hanya berfokus pada pencapaian tingkat akurasi prediksi yang tinggi, namun juga pada aspek interpretabilitas model yakni kemampuan model dalam memberikan penjelasan komprehensif terkait pengaruh setiap variabel input yang digunakan. Oleh karena itu, perhatian utama penelitian ini diarahkan pada optimasi algoritma *Random Forest* untuk menghasilkan prediksi yang andal, serta integrasi metode *SHapley Additive exPlanations* (SHAP) untuk menginterpretasikan kontribusi masing-masing fitur terhadap hasil prediksi secara transparan.

Studi Literatur

Tahap ini dilakukan melalui penelaahan berbagai referensi ilmiah terkait *Random Forest* dan *SHapley Additive exPlanations* (SHAP). Tujuannya adalah untuk memahami landasan teoretis, prinsip kerja, serta meninjau implementasi metode tersebut pada penelitian terdahulu guna memperkuat kerangka berpikir dalam penelitian ini.

Pengumpulan Dataset

Data diperoleh dari sumber publik Kaggle.com yang berisi dataset lagu Spotify dengan atribut audio numerik seperti *danceability*, *energy*, *valence*, *tempo*, dan *loudness* (Branchris et al., 2026). Variabel yang relevan dipilih kemudian melalui proses pengolahan menyeluruh, meliputi pembersihan data (*data cleaning*) dan normalisasi fitur. Variabel-variabel utama yang menjadi fokus penelitian ini disajikan pada Tabel 1.

Tabel 1. Fitur Emosi Musik

Atribut	Defenisi
<i>Danceability</i>	seberapa enak lagi itu unutm menari
<i>Energy</i>	seberapa kuat atau intens musiknya
<i>Valence</i>	apakah lagu itu terdengar bahagia atau sedih

Atribut	Defenisi
<i>Acousticness</i>	seberapa akustik atau alami bunyinya.
<i>Speechiness</i>	seberapa banyak unsur percakapan / vokal
<i>Liveness</i>	apakah lagu terdengar seperti rekaman live.
<i>Instrumental</i>	seberapa banyak bagian instrumental tanpa vokal.

Pra-pemrosesan Data (Pre-processing)

Data yang telah dikumpulkan selanjutnya melalui tahapan pra-pemrosesan secara sistematis untuk meningkatkan kualitas *dataset*. Tahap pembersihan data dilakukan dengan mengidentifikasi serta menghapus anomali, duplikasi, dan nilai yang hilang (*missing values*) guna menjamin validitas informasi (Gori et al., 2024). Selain itu, dilakukan normalisasi untuk menyelaraskan skala antarvariabel guna mencegah bias dan memastikan konsistensi data. Rangkaian proses ini bertujuan menghasilkan dataset yang bersih dan terstruktur, sehingga siap digunakan dalam tahap pelatihan model *Random Forest*.

Pembangunan Model Random Forest

Setelah data siap, algoritma *Random Forest* digunakan untuk membangun model prediksi. *Random Forest* merupakan algoritma berbasis ensemble learning yang terdiri dari banyak pohon keputusan (*decision tree*). Model ini bekerja dengan menggabungkan hasil prediksi dari masing-masing pohon untuk menghasilkan prediksi yang lebih stabil dan akurat.

Pengujian Model

Model yang telah dilatih dievaluasi menggunakan data uji (*testing data*) untuk menilai kinerja dan kemampuan generalisasinya terhadap data baru. Evaluasi ini krusial untuk memastikan model tidak hanya optimal pada data latih, tetapi juga andal saat menghadapi pola data yang bersifat umum. Kinerja model diukur menggunakan tiga metrik regresi utama:

1. R^2 (*Coefficient of Determination*): Mengukur sejauh mana variasi data dapat dijelaskan oleh model.

2. MAE (*Mean Absolute Error*): Menghitung rata-rata kesalahan absolut antara nilai prediksi dan nilai aktual.
3. RMSE (*Root Mean Square Error*): Ukuran kesalahan yang lebih sensitif terhadap penyimpangan prediksi yang besar.

Kombinasi nilai yang tinggi serta nilai MAE dan RMSE yang rendah menunjukkan bahwa model *Random Forest* yang dibangun memiliki tingkat akurasi dan keandalan yang optimal dalam memprediksi popularitas musik.

Interpretasi Model Menggunakan SHAP

Setelah tahap pelatihan model *Random Forest*, dilakukan interpretasi prediksi menggunakan metode *SHapley Additive exPlanations* (SHAP). Sebagai bagian dari pendekatan *Explainable Artificial Intelligence* (XAI), SHAP menguraikan kontribusi setiap fitur terhadap hasil prediksi berdasarkan konsep nilai Shapley dari teori permainan (Hapsari et al., 2026). Metode ini memungkinkan atribusi pengaruh fitur dilakukan secara aditif dan konsisten, sehingga memberikan transparansi terhadap keputusan model yang bersifat *black-box* (Rosyid & Pramaditya, 2025).

Hasil interpretasi SHAP disajikan melalui berbagai visualisasi, seperti *summary plot* dan *dependence plot*, untuk menggambarkan pengaruh relatif setiap fitur secara kuantitatif maupun individual. Pendekatan ini bertujuan untuk mengidentifikasi variabel yang paling signifikan memengaruhi *output* model, sehingga keputusan yang dihasilkan dapat dipertanggungjawabkan secara ilmiah dan lebih mudah dipahami (Airlangga, 2024).

Evaluasi dan Penarikan Kesimpulan

Tahap akhir penelitian ini melibatkan evaluasi menyeluruh terhadap kinerja model *Random Forest* menggunakan metrik penilaian regresi yang relevan guna mengukur tingkat keandalan dan konsistensi prediksi. Selain aspek teknis, hasil interpretasi dari metode SHAP

dianalisis untuk memahami pola kontribusi fitur terhadap *output* model secara mendalam.

Berdasarkan integrasi hasil evaluasi kinerja dan analisis interpretabilitas tersebut, disusunlah sebuah kesimpulan komprehensif. Kesimpulan ini mencakup tingkat akurasi model, identifikasi fitur-fitur yang paling signifikan memengaruhi popularitas lagu, serta sejauh mana model yang dikembangkan menjawab rumusan masalah penelitian. Melalui pendekatan ini, hasil penelitian tidak hanya memberikan luaran teknis tetapi juga penjelasan ilmiah yang dapat dipertanggungjawabkan.

C. HASIL DAN PEMBAHASAN

Penelitian ini menganalisis pengaruh fitur emosional musik terhadap popularitas lagu di Spotify menggunakan algoritma *Random Forest Regression* dan metode interpretabilitas SHAP. Dataset yang bersumber dari Spotify Web API mencakup fitur emosional seperti *danceability*, *energy*, *valence*, *acousticness*, *speechiness*, *liveness*, dan *instrumentalness* dengan variabel target berupa jumlah *stream*.

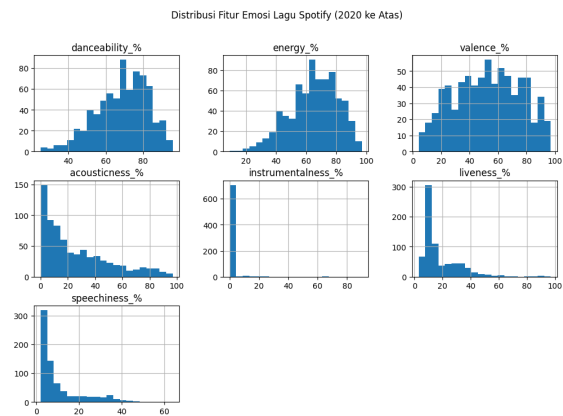
Data dibagi dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian. Hasil evaluasi menunjukkan model *Random Forest* memiliki performa yang cukup baik dengan nilai prediksi yang mendekati nilai aktual. Meskipun terdapat kecenderungan *underestimate* (prediksi sedikit di bawah nilai asli) pada jumlah *stream* tertentu, model ini tetap dinilai andal dalam menangkap tren popularitas berdasarkan karakteristik audio yang digunakan.

Distribusi Fitur Emosi Lagu

Hasil analisis statistik deskriptif terhadap dataset lagu populer Spotify sejak tahun 2020 menunjukkan karakteristik emosional yang spesifik. Fitur *danceability* mencatat rata-rata tertinggi ($\approx 68,88$), diikuti oleh *energy* ($\approx 64,88$). Hal ini mengindikasikan bahwa audiens cenderung menyukai lagu dengan ritme yang kuat,

terstruktur, dan memiliki intensitas energi yang tinggi.

Sementara itu, fitur *valence* berada pada angka moderat ($\approx 52,49$), menunjukkan keseimbangan antara nuansa emosi positif dan negatif. Rendahnya nilai *instrumentalness* ($\approx 1,13$) menegaskan dominasi unsur vokal pada lagu-lagu populer. Secara keseluruhan, temuan ini menyimpulkan bahwa tren lagu populer pada periode tersebut memiliki ciri khas dominasi vokal, berenergi tinggi, serta bersifat ritmis.



Gambar 2. Distribusi Fitur Emosi Lagu

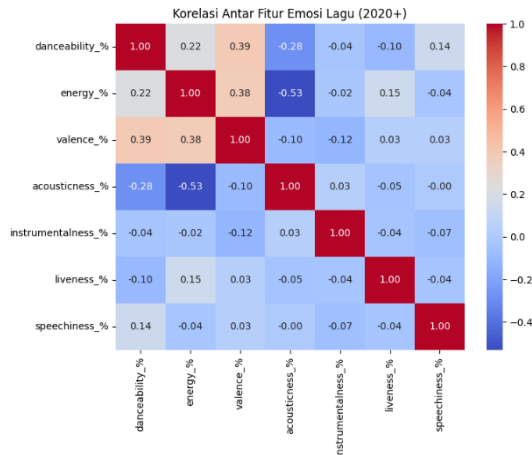
Korelasi Antar Fitur Emosi Lagu

Hubungan antar fitur audio secara umum berada pada tingkat lemah hingga sedang. Hal ini menunjukkan bahwa setiap fitur merepresentasikan dimensi karakteristik audio yang berbeda. Temuan utama dalam analisis korelasi ini adalah:

1. Korelasi Positif: Fitur *danceability*, *energy*, dan *valence* saling berkorelasi positif, menunjukkan bahwa lagu yang energik dan ritmis cenderung membawa nuansa emosional yang positif.
2. Korelasi Negatif: *Acousticness* memiliki korelasi negatif yang cukup kuat dengan *energy*, mengindikasikan bahwa lagu akustik cenderung memiliki intensitas energi yang lebih rendah.
3. Independensi Fitur: Rendahnya korelasi pada fitur *instrumentalness*, *liveness*, dan *speechiness* terhadap variabel lain

menunjukkan pengaruh yang bersifat kontekstual dan non-linear.

Rendahnya risiko multikolinearitas (tidak adanya korelasi yang sangat tinggi antar-fitur) memperkuat justifikasi penggunaan algoritma *Random Forest Regression* dalam penelitian ini, karena setiap fitur dapat memberikan kontribusi informasi yang unik terhadap model.



Gambar 3. Korelasi Antar Fitur Emosi Lagu

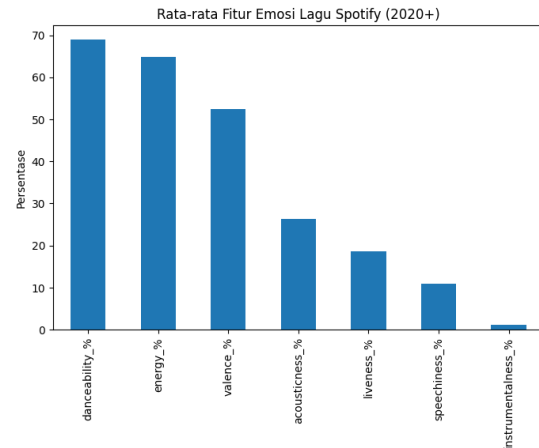
Rata-Rata Fitur Emosi Lagu

Analisis terhadap fitur emosional lagu populer di Spotify periode 2020 ke atas menunjukkan pola karakteristik yang konsisten:

1. *Danceability* rata-rata tertinggi ($\approx 68,88$), menandakan ritme yang mudah diikuti dan mendukung gerakan tari.
2. *Energy* relatif tinggi ($\approx 64,88$), mencerminkan intensitas akustik kuat, tempo cepat, dan nuansa produksi yang energik.
3. *Valence* berada di tingkat menengah ($\approx 52,49$), menunjukkan keseimbangan antara ekspresi emosional positif dan negatif.
4. *Acousticness* sedang ($\approx 26,34$), mengindikasikan dominasi elemen digital dibandingkan akustik murni.
5. *Liveness* ($\approx 18,58$) dan *Speechiness* ($\approx 10,99$) rendah-menengah, menandakan

minimnya format live maupun spoken word.

6. *Instrumentalness* paling rendah ($\approx 1,13$), menegaskan dominasi vokal dibandingkan komposisi instrumental.

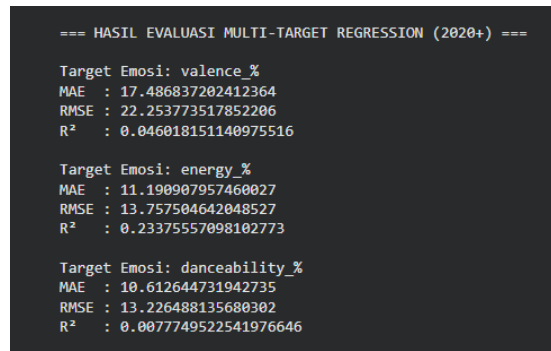


Gambar 4. Rata-Rata Fitur Emosi Lagu

Hasil Evaluasi Multi-Target Regression

Hasil evaluasi terhadap fitur emosi lagu menunjukkan bahwa:

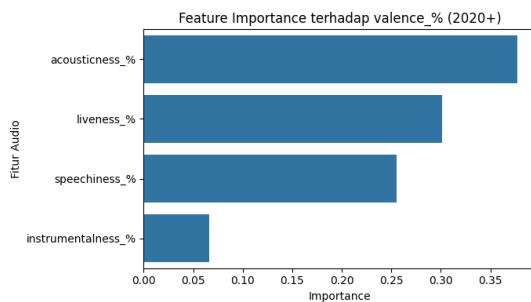
1. *Energy_ %* memiliki performa prediksi terbaik dengan nilai $R^2 = 0,23$, yang berarti model mampu menjelaskan sekitar 23% variasi tingkat energi lagu. Hal ini menegaskan bahwa energi lagu cukup erat terkait dengan karakteristik fisik audio (intensitas, tempo, struktur suara), sehingga lebih mudah ditangkap oleh model.
2. *Valence_ %* dan *Danceability_ %* menunjukkan nilai R^2 rendah, menandakan keterbatasan model dalam memprediksi emosi positif maupun tingkat kemudahan lagu untuk ditarikan hanya dari fitur audio numerik.
3. Kedua aspek tersebut lebih bersifat subjektif dan kontekstual, sehingga dipengaruhi oleh faktor eksternal seperti lirik, genre, serta preferensi pendengar.



Gambar 5. Hasil evaluasi Multi-Target

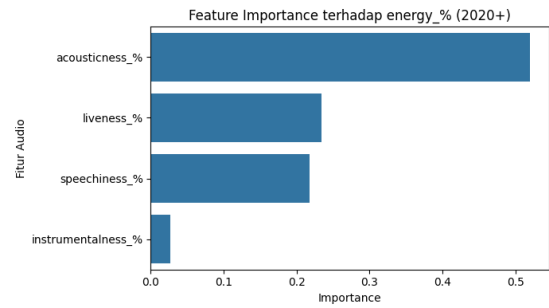
Pembahasan *Feature Importance*

Analisis *feature importance* memperlihatkan bahwa setiap dimensi emosi lagu dipengaruhi oleh kombinasi fitur audio dengan kontribusi berbeda.



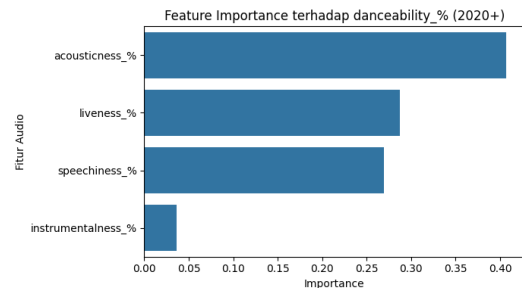
Gambar 6. *Feature Importance* terhadap *valence*

1. *Acousticness* ($\approx 0,38$) → faktor paling dominan, menunjukkan bahwa karakter akustik berperan besar dalam membentuk persepsi emosional positif/negatif.
2. *Liveness* ($\approx 0,30$) → kontribusi signifikan, menandakan suasana live turut memengaruhi valensi emosional.
3. *Speechiness* ($\approx 0,25$) → keberadaan elemen spoken word juga memberi pengaruh cukup kuat terhadap persepsi emosi.
4. *Instrumentalness* ($\approx 0,07$) → kontribusi kecil, menegaskan bahwa instrumen murni tidak terlalu menentukan valensi emosional.



Gambar 7. *Feature Importance* terhadap energi

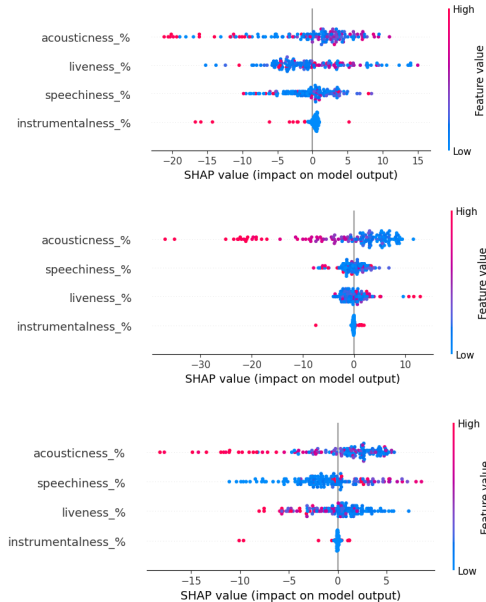
Energy_% *Acousticness* kembali menjadi faktor dominan ($\approx 0,52$), menunjukkan bahwa tingkat energi lagu sangat terkait dengan sifat akustik atau digital dari produksi musik. *Liveness* ($\approx 0,23$) dan *speechiness* ($\approx 0,22$) juga berperan, meskipun dengan pengaruh lebih rendah. *Instrumentalness* memiliki kontribusi minimal ($\approx 0,03$), sehingga dapat disimpulkan bahwa energi lebih ditentukan oleh intensitas akustik dan dinamika vokal dibandingkan komposisi *instrumental*.



Gambar 8. *Feature Importance* terhadap *danceability*

Danceability_% *Acousticness* tetap menjadi variabel utama ($\approx 0,41$), diikuti oleh *liveness* ($\approx 0,29$) dan *speechiness* ($\approx 0,27$). Hal ini menunjukkan bahwa kemudahan lagu untuk ditarikan tidak hanya bergantung pada ritme, tetapi juga dipengaruhi oleh kualitas akustik, nuansa live, serta elemen vokal/spoken word. *Instrumentalness* kembali memiliki pengaruh yang sangat kecil ($\approx 0,04$).

SHAP Value Analysis



Gambar 9. SHAP Value Analysis

Visualisasi SHAP digunakan untuk menginterpretasikan pengaruh fitur audio terhadap output model. Hasil analisis menunjukkan:

1. *Acousticness_%* kontribusi paling besar, ditunjukkan oleh sebaran nilai SHAP yang paling luas. Nilai tinggi meningkatkan output model, sedangkan nilai rendah memberi pengaruh negatif.
2. *Liveness_%* dan *Speechiness_%* berpengaruh cukup berarti, namun dengan rentang SHAP lebih sempit. Keduanya berperan sebagai faktor pendukung dalam prediksi.
3. *Instrumentalness_%* kontribusi paling rendah, dengan nilai SHAP mayoritas di sekitar nol. Perubahan nilai fitur ini tidak signifikan terhadap hasil prediksi.

Analisis Hasil

Hasil pemodelan dan pengujian menunjukkan bahwa fitur emosional musik memiliki pengaruh berbeda terhadap popularitas lagu di Spotify. Model *Random Forest Regression* berhasil menangkap hubungan non-linear antar fitur audio, sehingga prediksi jumlah stream lagu cukup representatif.

Statistik deskriptif memperlihatkan bahwa lagu populer cenderung memiliki *danceability* dan *energy* tinggi, serta *instrumentalness* sangat rendah. Hal ini menegaskan preferensi pendengar terhadap lagu dengan ritme mudah diikuti, energi kuat, dan dominasi vokal.

Feature importance & SHAP mengungkapkan bahwa:

1. *Acousticness* menjadi fitur paling berpengaruh, dengan nilai tinggi maupun rendah berdampak signifikan pada prediksi.
2. *Liveness* dan *Speechiness* berperan sebagai faktor pendukung dengan kontribusi moderat.
3. *Instrumentalness* konsisten menunjukkan pengaruh paling rendah.

D. PENUTUP

Penelitian ini berhasil mengimplementasikan algoritma *Random Forest Regression* untuk memprediksi popularitas lagu di Spotify berdasarkan fitur emosi musik dengan hasil yang cukup representatif. Berdasarkan analisis deskriptif, lagu-lagu populer pasca-2020 didominasi oleh karakteristik *danceability* dan *energy* yang tinggi, yang mencerminkan preferensi audiens terhadap musik ritmis dan enerjik. Hasil interpretasi menggunakan metode SHAP menunjukkan bahwa fitur *acousticness* dan *valence* memiliki pengaruh paling signifikan terhadap jumlah *stream*. Meskipun model memiliki kecenderungan sedikit merendahkan estimasi (*underestimate*) pada angka pemutaran yang sangat tinggi, secara keseluruhan hasil penelitian ini mampu memberikan transparansi ilmiah mengenai faktor-faktor emosional yang mendorong keberhasilan sebuah karya musik di platform digital.

E. DAFTAR PUSTAKA

Airlangga, G. (2024). Comparing BDD and

- TDD: Machine Learning Analysis of Software Quality with SHAP Interpretability. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, 8(4), 2615–2625. <https://doi.org/10.33395/sinkron.v8i4.14201>
- Branchris, Yech, K. A., & Wijaya, A. (2026). Klasterisasi Lagu pada Dataset Spotify Berdasarkan Fitur Audio Menggunakan Algoritma K-Means. *Sinergi: Jurnal Ilmiah Multidisiplin*, 2(1), 106–117. <https://publikasi.ahlalkamal.com/index.php/sinergi/article/view/251>
- Feng, Y., & Wang, M. (2025). Effect of music therapy on emotional resilience, well-being, and employability: a quantitative investigation of mediation and moderation. *BMC Psychology*, 13, 47. <https://doi.org/10.1186/s40359-024-02336-x>
- Gori, T., Sunyoto, A., & Fatta, H. Al. (2024). Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(1), 215–224. <https://doi.org/10.25126/jtiik.20241118074>
- Hapsari, L. N., Fannani, I., Rahmawati, Y., & Muhariya, A. (2026). Explainable Machine Learning untuk Prediksi Risiko Penyakit Jantung Menggunakan Random Forest dan Analisis SHAP. *REMIK: Riset Dan E-Jurnal Manajemen Informatika Komputer*, 10(1), 190–199. <https://doi.org/10.33395/remik.v10i1.15766>
- Loureiro, C., van der Meulen, K., & del Barrio, C. (2024). Why I listen to music: Emotion regulation and identity construction through music in mid-adolescence. *Empiria*, 60. <https://doi.org/10.5944/empiria.60.2024.39285>
- Rosyid, I. F., & Pramaditya, H. (2025). Visual Interpretation of Machine Learning Models (Random Forest) for Lung Cancer Risk Classification Using Explainable Artificial Intelligence (SHAP & LIME). *JUTIF: Jurnal Teknik Informatika*, 6(4), 2187–2206. <https://doi.org/10.52436/1.jutif.2025.6.44925>
- Setiawan, A., Utami, E., & Ariatmanto, D. (2024). Cattle Weight Estimation Using Linear Regression and Random Forest Regressor. *Jurnal RESTI*, 8(1), 72–79. <https://doi.org/https://doi.org/10.29207/resti.v8i1.5494>
- Setiawan, I. K. D., Sudiarsa, I. W., Naratama, I. P. D., & Nope, A. T. (2026). Analisis Klasifikasi Perilaku Pengguna Terhadap Customer Churn Pada Layanan Musik Spotify Menggunakan Metode Random Forest. *JMA : Jurnal Media Akademik*, 4(1), 1–15. <https://doi.org/10.62281/qcrest95>
- Simalango, H. M., Amelia, E., Marsyanda, R., Ananda, D., Putra, M. R. S., Setiawan, A. Y., & Chanthay, C. (2025). Music, Algorithms, and Popularity: How TikTok Trends Act as Catalysts in Shaping Contemporary Culture. *Musikolastika: Jurnal Pertunjukan Dan Pendidikan Musik*, 7(2), 140–153. <https://doi.org/10.24036/musikolastika.v7i2.239>
- Tannady, H., Andry, J. F., & Honni. (2024). Analisis Big Data Spotify dengan Metode Data Mining. *JBASE - Journal of Business and Audit Information Systems*, 7(2), 52–59. <https://doi.org/10.30813/jbase.v7i2.6261>
- Ye, G., Zhao, Y., Qiu, T., & Wang, H. (2024). Analysis of Factors Influencing Music Popularity and Trend Prediction. *Frontiers in Science and Engineering*, 4(12). <https://doi.org/10.54691/sftbz407>