

**KLASIFIKASI PREFERENSI MUSIK GENERASI Z  
BERDASARKAN FITUR AUDIO SPOTIFY  
MENGUNAKAN ALGORITMA  
RANDOM FOREST**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Sistem Informasi



disusun oleh

**NI MADE SUPARTINI**

**22.12.2479**

Kepada

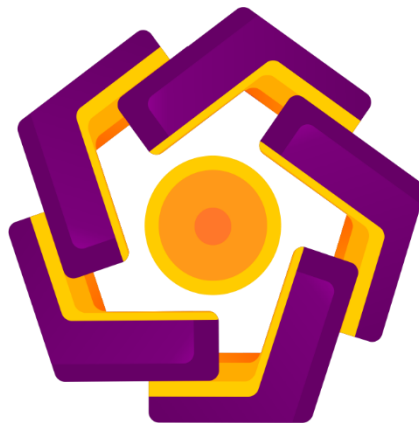
**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

**KLASIFIKASI PREFERENSI MUSIK GENERASI Z  
BERDASARKAN FITUR AUDIO SPOTIFY  
MENGUNAKAN ALGORITMA  
RANDOM FOREST**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Sistem Informasi



disusun oleh

**NI MADE SUPARTINI**

**22.12.2479**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**KLASIFIKASI PREFERENSI MUSIK GENERASI Z BERDASARKAN  
FITUR AUDIO SPOTIFY MENGGUNAKAN ALGORITMA RANDOM  
FOREST**

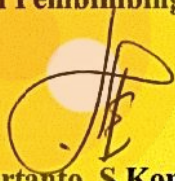
yang disusun dan diajukan oleh

**NI MADE SUPARTINI**

**22.12.2479**

telah disetujui oleh Dosen Pembimbing Skripsi  
25 Februari 2026

**Dosen Pembimbing,**



**Anggit Dwi Hartanto, S.Kom., M.Kom.**

**NIK. 190302163**

**HALAMAN PENGESAHAN**  
**SKRIPSI**  
**KLASIFIKASI PREFERENSI MUSIK GENERASI Z BERDASARKAN**  
**FITUR AUDIO SPOTIFY MENGGUNAKAN ALGORITMA RANDOM**  
**FOREST**

yang disusun dan diajukan oleh

**NI MADE SUPARTINI**

**22.12.2479**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 25 Februari 2026

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Atik Nurmasani, S.kom., M.Kom.**  
**NIK. 190302354**

**Arif Nur Rohman, M.Kom.**  
**NIK. 190302684**

**Anggit Dwi Hartanto, S.Kom., M.Kom.**  
**NIK. 190302163**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 25 Februari 2026

**DEKAN FAKULTAS ILMU KOMPUTER**



**Prof. Dr. Kusrini, M.Kom.**  
**NIK. 190302106**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : NI MADE SUPARTINI  
NIM : 22.12.479

Menyatakan bahwa Skripsi dengan judul berikut:

### KLASIFIKASI PREFERENSI MUSIK GENERASI Z BERDASARKAN FITUR AUDIO SPOTIFY MENGGUNAKAN ALGORITMA RANDOM FOREST

Dosen Pembimbing : Anggit Dwi Hartanto, S.Kom., M.Kom.

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

Yogyakarta, 25 Februari 2026

Yang Menyatakan,



NI MADE SUPARTINI

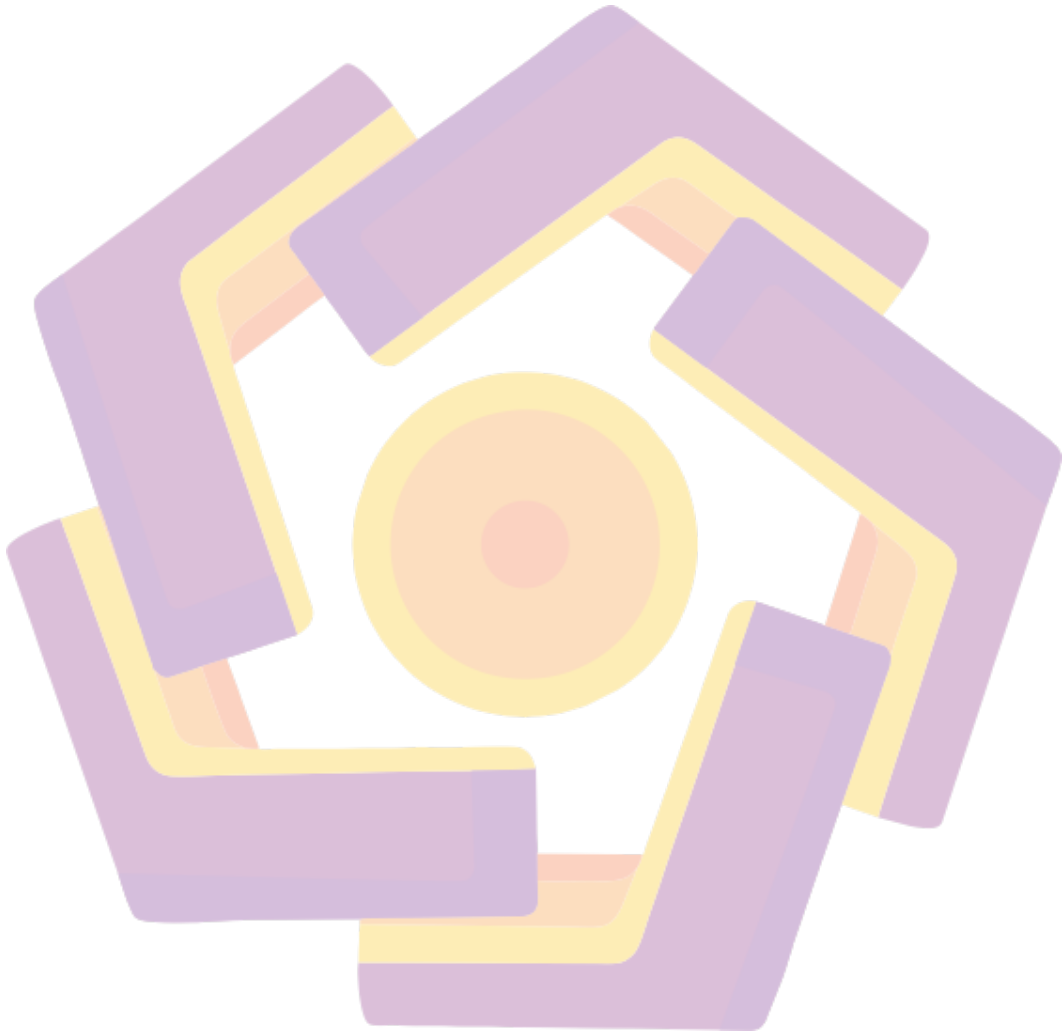
## HALAMAN PERSEMBAHAN

Puji syukur kepada Ida Sang Hyang Widhi Wasa atas asung kerta wara anugerah-Nya kepada penulis beserta keluarga, dosen pembimbing, dan rekan seperjuangan dapat menyelesaikan skripsi tepat waktu. Skripsi ini penulis persembahkan sebagai wujud yadnya (korban suci tulus ikhlas):

1. Kedua orang tercinta, Bapak dan Mama, serta adikku Ketut, yang selalu menjadi sumber doa, kasih sayang, dan kekuatan dalam setiap langkah penulis. Terima kasih atas pengorbanan, kesabaran, dan dukungan tanpa henti sehingga penulis mampu menyelesaikan hingga tahap ini.
2. Saudara perempuan kembar saya, Putu, yang telah berjuang bersama dalam menyelesaikan skripsi ini. Terima kasih karena telah melalui setiap proses secara berdampingan mengerjakan skripsi bersama, berbagi suka dan duka, saling menguatkan saat lelah, bahkan menangis dan bangkit bersama. Kebersamaan ini menjadi kenangan yang sangat berharga karena kita dapat sampai di titik menyelesaikan skripsi secara bersama-sama.
3. Keluarga besar Pakyan, Mekah, Pakmang, Mekdek, Ajik dan Biang ayu, Ninik, dadong senantiasa memberikan doa, perhatian serta semangat sehingga penulis mampu menyelesaikan skripsi ini dengan baik.
4. Ayuk Savitri, Ayuk sayu, Kakak dan Adi Saputra yang selalu memberikan motivasi dan dukungan dalam setiap proses perjalanan ini.
5. Mbok Saras, dengan tulus membantu sejak awal penyusunan hingga ini dapat diselesaikan.
6. Pakyan bos bagio, yang telah meyakinkan bapak dan memberikan keyakinan bahwa penulis mampu melanjutkan pendidikan ke jenjang kuliah.
7. Sahabat terdekat, Egi Apreza dan Gita Regina Savitri, yang selalu hadir dalam suka dan duka, berbagi semangat, dan menjadi teman bertumbuh selama masa perkuliahan.
8. Dosen pembimbing, Bapak Anggit yang bersabar memberikan arahan, masukan, dan motivasi hingga skripsi ini tersusun dengan baik.

9. Seluruh teman Angkatan 2022, yang telah menjadi bagian dari perjalanan dan perjuangan bersama hingga mencapai tahap akhir perkuliahan.

Semoga persembahan ini menjadi bagian dari yadnya sebagai ungkapan bakti dan terima kasih atas setiap doa, dukungan, serta kebaikan yang telah diberikan.



## KATA PENGANTAR

Puji syukur kepada Ida Sang Hyang Widhi Wasa atas asung kerta wara anugra-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Preferensi Musik Berdasarkan Fitur Audio Spotify Menggunakan Algoritma Random Forest” dengan baik dan tepat waktu. Skripsi ini disusun sebagai salah satu syarat penyelesaian Program Studi Sistem Informasi, Universitas Amikom Yogyakarta.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa penelitian ini tidak terselesaikan tanpa dukungan, bimbingan, serta arahan dari berbagai pihak. Oleh karena itu, penulis menyampaikan rasa hormat dan terima kasih kepada:

1. Prof. Dr. M. Suyanto, M.M. (Rektor Universitas AMIKOM Yogyakarta)
2. Prof. Dr. Kusrini, M.Kom. (Dekan Fakultas Ilmu Komputer)
3. Dr. Andi Sunyoto, M.Kom. (Wakil Dekan I Fakultas Ilmu Komputer)
4. Melwin Syafrizal, M.Eng., Ph.D. (Wakil Dekan II Fakultas Ilmu Komputer)
5. Anggit Dwi Hartanto, S.Kom., M.Kom. (Kaprodi Sistem Informasi dan Dosen Pembimbing)
6. Seluruh dosen Program Studi S1 Universitas Sistem Informasi Amikom Yogyakarta.
7. Orang-orang terdekat yang tidak bisa disebutkan satu per satu yang telah memberikan semangat dan dukungan sampai saat ini.

Yogyakarta, 23 Februari 2026

NI MADE SUPARTINI

## DAFTAR ISI

|                                           |       |
|-------------------------------------------|-------|
| HALAMAN JUDUL .....                       | i     |
| HALAMAN PERSETUJUAN.....                  | ii    |
| HALAMAN PENGESAHAN .....                  | iii   |
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI ..... | iv    |
| HALAMAN PERSEMBAHAN .....                 | v     |
| KATA PENGANTAR .....                      | vii   |
| DAFTAR ISI.....                           | viii  |
| DAFTAR TABEL.....                         | xii   |
| DAFTAR GAMBAR.....                        | xiii  |
| DAFTAR LAMBANG DAN SINGKATAN .....        | xv    |
| DAFTAR ISTILAH.....                       | xvi   |
| INTISARI .....                            | xvii  |
| <i>ABSTRACT</i> .....                     | xviii |
| BAB I PENDAHULUAN.....                    | 1     |
| 1.1    Latar Belakang.....                | 1     |
| 1.2    Rumusan Masalah.....               | 2     |
| 1.3    Batasan Masalah .....              | 3     |
| 1.4    Tujuan Penelitian .....            | 3     |
| 1.5    Manfaat Penelitian .....           | 3     |
| 1.6    Sistematika Penulisan .....        | 4     |
| BAB II TINJAUAN PUSTAKA .....             | 6     |

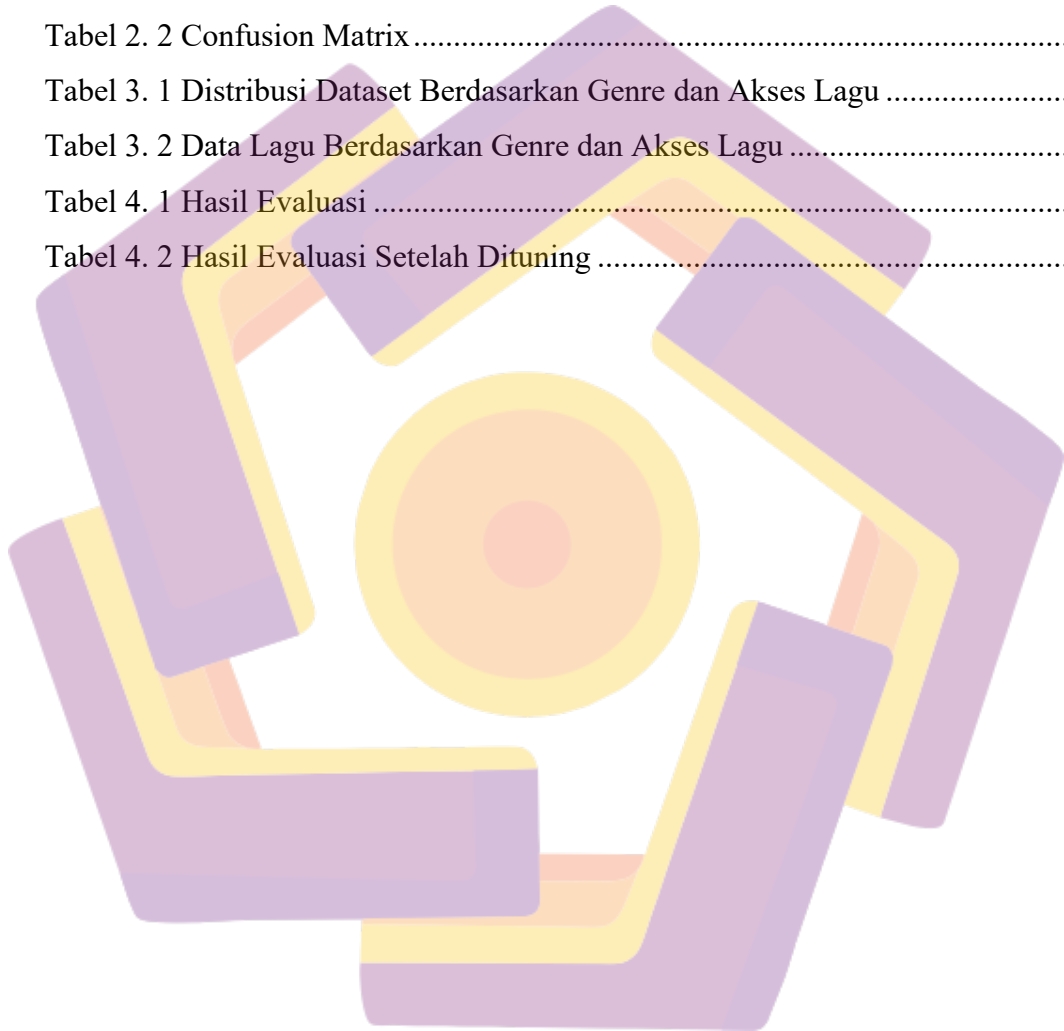
|                                 |                                   |    |
|---------------------------------|-----------------------------------|----|
| 2.1                             | Studi Literatur .....             | 6  |
| 2.2                             | Dasar Teori.....                  | 18 |
| 2.2.1                           | Klasifikasi .....                 | 18 |
| 2.2.2                           | Preferensi Musik Generasi Z ..... | 18 |
| 2.2.3                           | Spotify.....                      | 18 |
| 2.2.4                           | Fitur audio Spotify .....         | 19 |
| 2.2.5                           | Genre Musik.....                  | 20 |
| 2.2.6                           | Preprocessing .....               | 21 |
| 2.2.7                           | Machine Learning .....            | 22 |
| 2.2.8                           | Random Forest.....                | 22 |
| 2.2.9                           | <i>Hyperarameter</i> Tuning ..... | 23 |
| 2.2.10                          | <i>Random Search</i> .....        | 24 |
| 2.2.11                          | <i>Confusion Matrix</i> .....     | 24 |
| 2.2.12                          | <i>Accuracy</i> .....             | 25 |
| 2.2.13                          | <i>Precision</i> .....            | 26 |
| 2.2.14                          | <i>Recall</i> .....               | 26 |
| 2.2.15                          | F1-Score.....                     | 28 |
| 2.2.16                          | <i>Google Coloboratory</i> .....  | 29 |
| 2.2.17                          | Python .....                      | 29 |
| BAB III METODE PENELITIAN ..... |                                   | 30 |
| 3.1                             | Objek Penelitian.....             | 30 |
| 3.2                             | Alur Penelitian .....             | 32 |

|                                   |                                                                                   |    |
|-----------------------------------|-----------------------------------------------------------------------------------|----|
| 1.                                | Pengumpulan data .....                                                            | 33 |
| 2.                                | EDA .....                                                                         | 33 |
| 3.                                | Pra-pemrosesan .....                                                              | 35 |
| 4.                                | Split data .....                                                                  | 36 |
| 5.                                | Pemodelan <i>Random Forest</i> .....                                              | 36 |
| 6.                                | <i>Hyperparameter</i> Tuning .....                                                | 37 |
| 7.                                | Evaluasi Model .....                                                              | 37 |
| 3.3                               | Alat dan Bahan.....                                                               | 37 |
| 3.3.1                             | Data Penelitian .....                                                             | 37 |
| 3.3.2                             | Alat Instrumen .....                                                              | 38 |
| BAB IV HASIL DAN PEMBAHASAN ..... |                                                                                   | 40 |
| 4.1                               | Pengumpulan data .....                                                            | 40 |
| 4.2                               | Exploratory Data Analysis (EDA).....                                              | 40 |
| 4.2.1                             | Analisis Distribusi Genre .....                                                   | 41 |
| 4.2.2                             | Genre Musik Paling Populer Per tahun.....                                         | 42 |
| 4.2.3                             | Analisis Korelasi Fitur Audio (Heatmap) .....                                     | 43 |
| 4.2.4                             | Rata-rata <i>Danceability</i> dan <i>Valence</i> berdasarkan kategori musik ..... | 45 |
| 4.2.5                             | Top 10 Artis berdasarkan rata-rata popularitas.....                               | 46 |
| 4.2.6                             | Analisis Top 10 lagu yang diminati .....                                          | 47 |
| 4.2.7                             | Analisis trend genre musik populer berdasarkan tahun .....                        | 48 |
| 4.3                               | Pra-pemrosesan data .....                                                         | 49 |
| 4.3.1                             | Data Cleaning.....                                                                | 49 |

|           |                                                  |    |
|-----------|--------------------------------------------------|----|
| 4.3.2     | Normalisasi .....                                | 50 |
| 4.3.3     | Encoding Label .....                             | 50 |
| 4.4       | Split data .....                                 | 51 |
| 4.5       | Pemodelan Random Forest .....                    | 52 |
| 4.6       | Hyperparameter Tuning .....                      | 53 |
| 4.7       | Evaluasi Model .....                             | 54 |
| 4.7.1     | Confusion Matrix Random Forest Baseline .....    | 54 |
| 4.7.2     | Hasil Evaluasi .....                             | 56 |
| 4.7.3     | Confusion Matrix Model Tuning .....              | 57 |
| 4.7.4     | Hasil Evaluasi Setelah di Tuning .....           | 59 |
| 4.8       | Analisis Top 10 <i>Feature Importances</i> ..... | 60 |
| 4.9       | Deploy .....                                     | 62 |
| 4.9.1     | Backend aplikasi .....                           | 62 |
| 4.10      | Antarmuka pengguna .....                         | 67 |
| 4.11      | Pemanfaatan model .....                          | 73 |
| BAB V     | PENUTUP .....                                    | 74 |
| 5.1       | Kesimpulan .....                                 | 74 |
| 5.2       | Saran .....                                      | 75 |
| REFERENSI | .....                                            | 76 |

## DAFTAR TABEL

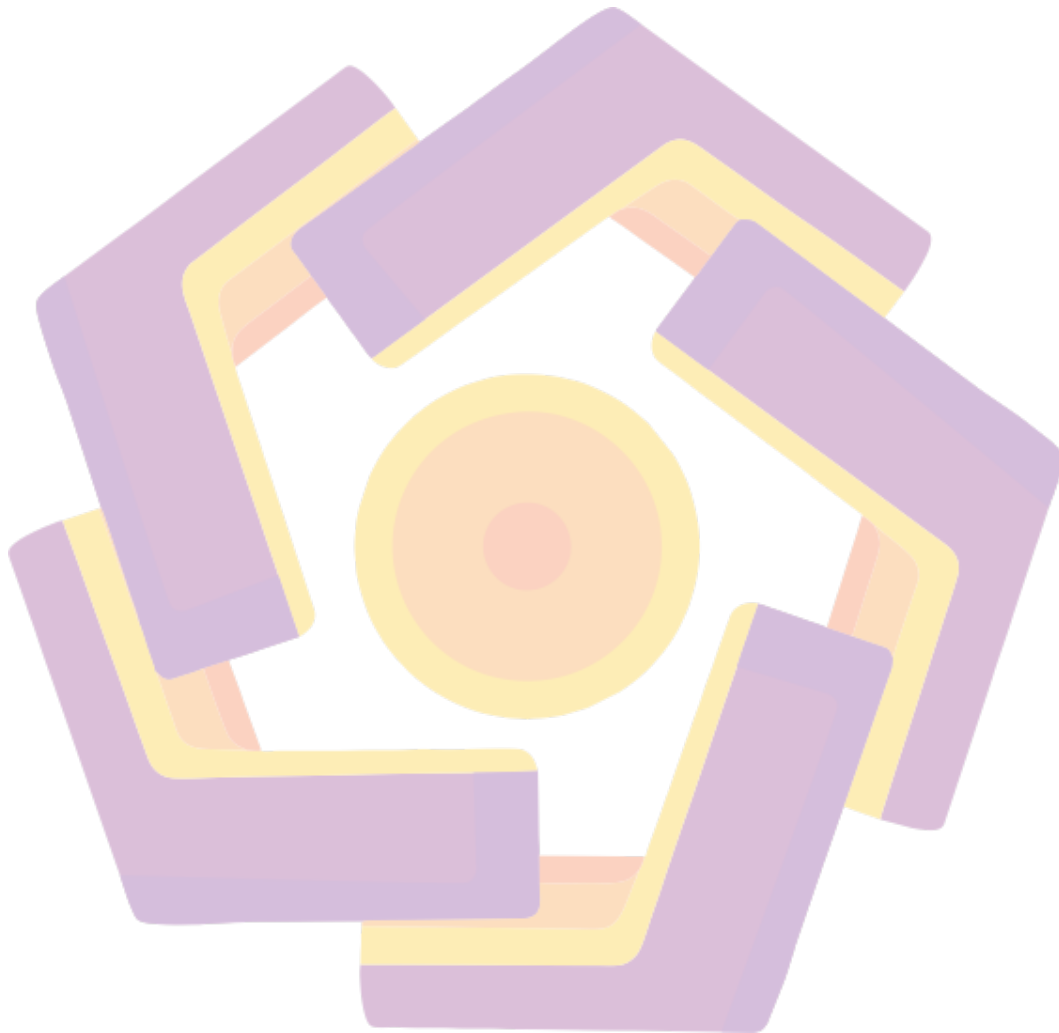
|                                                                      |    |
|----------------------------------------------------------------------|----|
| Tabel 2. 1 Keaslian Penelitian .....                                 | 8  |
| Tabel 2. 2 Confusion Matrix .....                                    | 24 |
| Tabel 3. 1 Distribusi Dataset Berdasarkan Genre dan Akses Lagu ..... | 30 |
| Tabel 3. 2 Data Lagu Berdasarkan Genre dan Akses Lagu .....          | 31 |
| Tabel 4. 1 Hasil Evaluasi .....                                      | 56 |
| Tabel 4. 2 Hasil Evaluasi Setelah Dituning .....                     | 59 |



## DAFTAR GAMBAR

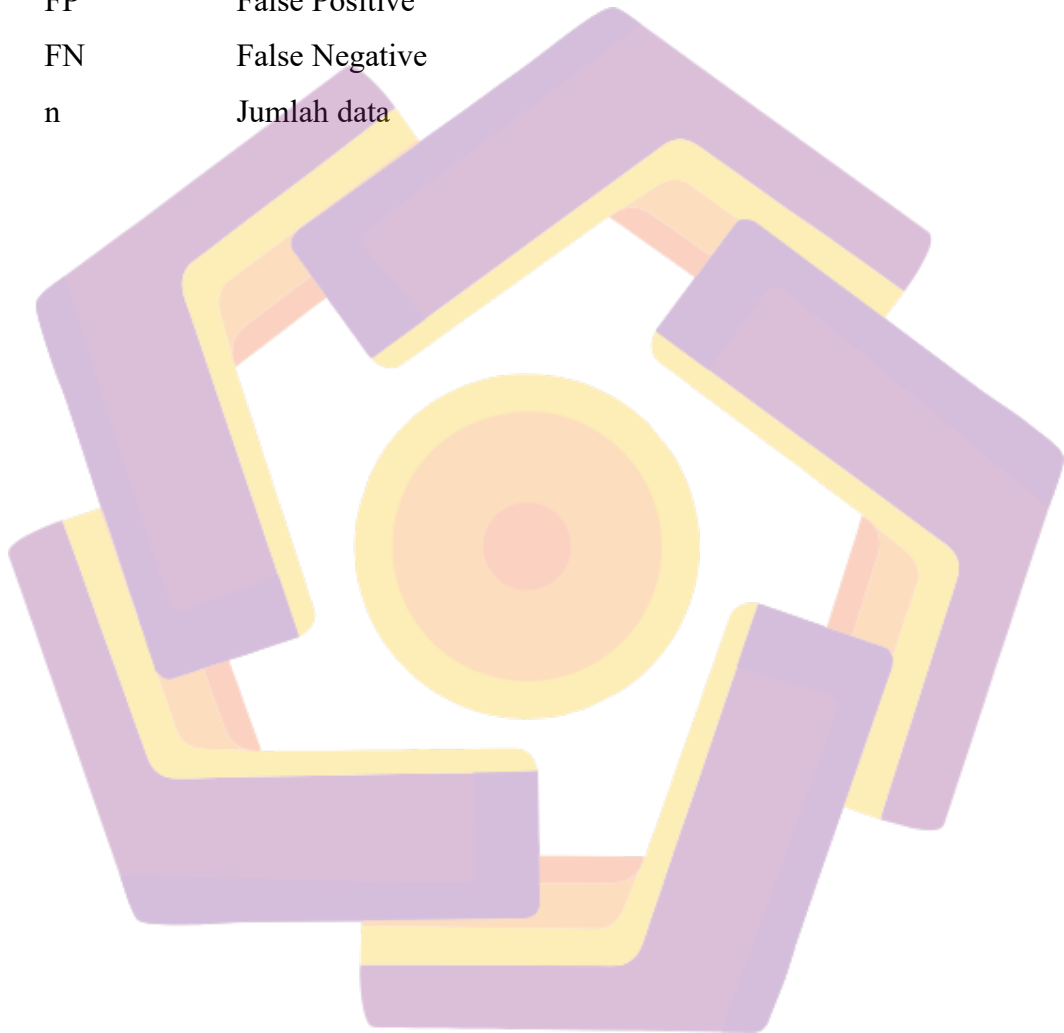
|                                                                              |    |
|------------------------------------------------------------------------------|----|
| Gambar 3. 1 Alur Penelitian .....                                            | 32 |
| Gambar 4. 1 Load Data .....                                                  | 40 |
| Gambar 4. 2 Visualisasi Genre Paling Populer Didataset.....                  | 41 |
| Gambar 4. 3 Visualisasi Genre Musik Paling Populer Per tahun .....           | 42 |
| Gambar 4. 4 Heatmap .....                                                    | 44 |
| Gambar 4. 5 Visualisasi Perbandingan Rata-Rata Danceability dan Valence..... | 45 |
| Gambar 4. 6 Visualisasi Top 10 Artis Berdasarkan Popularitas.....            | 46 |
| Gambar 4. 7 Visualisasi Top 10 Lagu Paling Diminati .....                    | 47 |
| Gambar 4. 8 Trend Musik Paling Populer Berdasarkan Tahun .....               | 48 |
| Gambar 4. 9 Cleaning .....                                                   | 49 |
| Gambar 4. 10 Normalisasi .....                                               | 50 |
| Gambar 4. 11 Encoding Label .....                                            | 50 |
| Gambar 4. 12 Split Data.....                                                 | 51 |
| Gambar 4. 13 Hasil Split Data .....                                          | 51 |
| Gambar 4. 14 Data Tabel Yang Setelah Dipreprocessing .....                   | 52 |
| Gambar 4. 15 Permodelan Random Forest .....                                  | 52 |
| Gambar 4. 16 Hyperparameter Tuning .....                                     | 53 |
| Gambar 4. 17 Random Search .....                                             | 54 |
| Gambar 4. 18 Confusion Matrix Model Baseline.....                            | 55 |
| Gambar 4. 19 Confusion Matrix Setelah Dituning .....                         | 58 |
| Gambar 4. 20 Feature Importances .....                                       | 61 |
| Gambar 4. 21 Import Library .....                                            | 62 |
| Gambar 4. 22 Kode Program Python .....                                       | 63 |
| Gambar 4. 23 Kode program Python .....                                       | 63 |
| Gambar 4. 24 Kode Program Python .....                                       | 63 |
| Gambar 4. 25 Kode Program Python .....                                       | 64 |
| Gambar 4. 26 Kode program Python .....                                       | 65 |
| Gambar 4. 27 Kode Program Python .....                                       | 66 |
| Gambar 4. 28 Tampilan Halaman Klasifikasi.....                               | 68 |

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| Gambar 4. 29 Tampilan Hasil Klasifikasi.....                          | 69 |
| Gambar 4. 30 Tampilan Informasi Fitur Audio .....                     | 70 |
| Gambar 4. 31 Tampilan Visualisasi Distribusi Lagu Bulan & Tahun ..... | 71 |
| Gambar 4. 32 Tampilan Visualisasi Distribusi Top 10 Lagu .....        | 72 |



## DAFTAR LAMBANG DAN SINGKATAN

|    |                |
|----|----------------|
| RF | Random Forest  |
| TP | True Positive  |
| TN | True Negative  |
| FP | False Positive |
| FN | False Negative |
| n  | Jumlah data    |



## DAFTAR ISTILAH

|               |                                                                                                                                                                                                  |
|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Genre musik   | Kategori atau pengelompokan musik berdasarkan karakteristik tertentu.                                                                                                                            |
| Fitur audio   | Representasi numerik dari karakteristik musik <i>danceability</i> , <i>energy</i> , <i>tempo</i> , <i>valence</i> , <i>speechiness</i> yang digunakan sebagai variabel dalam proses klasifikasi. |
| Random Forest | Algoritma klasifikasi menggunakan ensemble learning yang membangun sejumlah pohon keputusan dan menggabungkan hasil prediksi model dengan data aktual.                                           |
| Overfitting   | Kondisi ketika terlalu menyesuaikan dengan pelatihan sehingga performanya menurun saat diuji pada data baru.                                                                                     |

## INTISARI

Transformasi digital mengubah masyarakat dalam mengakses dan menikmati musik melalui platform streaming seperti Spotify yang menyediakan berbagai fitur audio sebagai representasi karakteristik lagu. Ketersediaan data berbasis fitur audio seperti *danceability*, *tempo*, *energy*, dan *speechiness* mendukung proses analisis musik dilakukan secara lebih sistematis dan objektif. Pada Generasi Z, preferensi musik tidak hanya dipengaruhi oleh label genre, tetapi juga oleh karakteristik audio suasana dan pola konsumsi musik yang dinamis.

Penelitian ini bertujuan untuk mengklasifikasi preferensi musik Generasi Z berdasarkan fitur audio menggunakan algoritma *Random Forest*. Dataset yang digunakan merupakan data lagu yang diperoleh dari *Spotify Track Dataset* dan *Spotify Web API* dengan rentang tahun 2015 hingga 2025 dan difokuskan pada empat genre utama, yaitu Pop, K-pop, Hip-hop dan R&B. Proses penelitian meliputi tahapan penelitian pengumpulan data, Exploratory Data Analysis (EDA), pra-pemrosesan, analisis data pembangunan model, serta evaluasi model. Algoritma *Random Forest* dipilih karena kemampuannya dalam mengolah data dengan jumlah variabel yang cukup besar serta menghasilkan model klasifikasi yang stabil.

Hasil penelitian ini menunjukkan bahwa model *Random Forest* memiliki performa yang sangat baik dalam mengklasifikasi preferensi musik Generasi Z. Model baseline memperoleh akurasi sebesar 94% dan meningkatkan menjadi 95% setelah dilakukan proses *hyperparameter tuning*. Nilai *precision*, *recall*, dan *F1-score* yang tinggi menunjukkan bahwa model memiliki kemampuan klasifikasi dengan tingkat kekeliruan antar kelas rendah. Dengan demikian, model yang dihasilkan dapat digunakan secara efektif dalam mengklasifikasi preferensi musik berbasis fitur audio.

**Kata kunci:** Random Forest, Klasifikasi, Fitur Audio, Spotify, Generasi Z.

## **ABSTRACT**

*Digital transformation has changed the way people access and enjoy music through streaming platforms such as Spotify, which provide various audio features as representations of song characteristics. The availability of audio-based features such as danceability, tempo, energy, and speechiness supports music analysis in a more systematic and objective way. Among Generation Z, music preferences are influenced not only by genre labels but also by audio characteristics, mood, and dynamic listening patterns.*

*This study aims to classify Generation Z music preferences based on audio features using the Random Forest algorithm. The dataset used consists of song data obtained from the Spotify Track Dataset and Spotify Web API, covering the period from 2015 to 2025 and focusing on four main genres Pop, K-pop, Hip-hop, and R&B. The research process includes data collection, Exploratory Data Analysis (EDA), preprocessing, model development, and evaluation. The Random Forest algorithm was selected because of its ability to process data with a relatively large number of variables and to produce a stable classification model.*

*The results of this study show that the Random Forest model performs very well in classifying Generation Z music preferences. The baseline model achieved an accuracy of 94%, which increased to 95% after the hyperparameter tuning process. High precision, recall, and F1-score values indicate that the model has strong classification capability with low inter-class misclassification. Therefore, the resulting model can be effectively used to classify music preferences based on audio features.*

**Keywords:** Random Forest, Classification, Audio Feature, Spotify, Generation Z.