

BAB I

PENDAHULUAN

1.1 Latar Belakang

Musik sudah menjadi hal yang lumrah bagi semua kalangan dan senantiasa hadir dalam berbagai macam kesempatan, seperti acara kenegaraan, berbagai jenis pesta perayaan, dan ritual keagamaan yang sarat akan kesakralannya. Musik telah lahir sejak zaman manusia purba, lalu berlanjut dan terus berkembang seiring berlalunya zaman. Dengan demikian, musik terus hadir dalam tiap kehidupan manusia dan akan terus melahirkan musik baru.

Seiring berjalannya waktu, musik terus menerus berevolusi yang membentuk berbagai jenis musik. Terdapat berbagai cara dalam mengelompokkan musik. Salah satu dari jenis pengelompokan musik adalah pengelompokan berdasarkan genre. Genre musik adalah kategori yang muncul melalui interaksi kompleks antara budaya, artis, dan kekuatan pasar untuk mengumpulkan kesamaan antara musisi atau komposisi dan untuk mengatur koleksi musik (Scaringella dkk., 2006).

Genre musik juga bisa diartikan sebagai klasifikasi instrumental ke dalam kategori secara umum yang membantu dalam identifikasi musik menurut praktik budaya yang saling melengkapi atau menurut acuan budaya yang berlaku (Mehta dkk., 2021). Mengelompokkan musik ke dalam genre tertentu terbilang cukup sulit karena batasan antargenre masih belum jelas, membuat klasifikasi genre musik secara otomatis menjadi tugas yang tidak mudah (Scaringella dkk., 2006). Klasifikasi

berdasarkan genre merupakan masalah yang kompleks dan luas yang bukan hanya sekadar representasi dari bentuk matematis dari suatu data audio (Nguyen dkk., 2017).

Klasifikasi genre musik adalah pengelompokan musik sesuai label tertentu setelah proses ekstraksi data audio berupa musik dan menganalisis beberapa informasi penting dari data audio tersebut (Du dkk., 2016). Klasifikasi genre musik umumnya memiliki dua tahapan: ekstraksi data audio dan klasifikasi dengan metode tertentu. Untuk sistem pengambilan informasi musik (*Music Information Retrieval*, MIR) standar, salah satu langkah dasarnya adalah mengumpulkan sebanyak mungkin fitur musik untuk mengklasifikasi intisari musik dengan tepat dan mencari satu musik lebih efisien.

Klasifikasi genre musik dapat digunakan di berbagai bidang dalam industri musik, seperti membantu produser musik dalam mengidentifikasi genre musik mana yang lebih populer (Mogonediwa, 2024). Selain itu, klasifikasi genre musik diperlukan karena berperan penting dalam layanan pemutaran musik berbasis aplikasi digital agar mampu memberikan rekomendasi dan pencarian musik bagi pengguna (Lau, 2020). Dengan demikian, proses klasifikasi genre musik menjadi fitur tambahan guna meningkatkan pengalaman pengguna (Mogonediwa, 2024).

Topik klasifikasi genre musik pertama kali diteliti oleh Tzanetakis dan Cook (2002). Keduanya melakukan penelitian terhadap klasifikasi genre musik otomatis, yang bersumber dari data audio bersumber dari koleksi audio *real-world*, yang diklasifikasi ke dalam sepuluh genre musik, yaitu *classical*, *country*, *disco*, *hiphop*, *jazz*, *rock*, *blues*, *reggae*, *pop*, dan *metal*. Penelitian tersebut menggunakan parameter ekstraksi fitur (*feature extraction*), yang di antaranya *timbral texture features*, *timbral texture features vector*, *rhythmic content features*, dan *pitch content features*. Fitur-

fitur tersebut diklasifikasi menggunakan *statistical pattern recognition* (SPR). Adapun akurasi dari klasifikasi tersebut dihitung berdasarkan kinerja fitur yang digunakan saat diterapkan dalam pengoreksian koleksi audio. Hasil perhitungan akurasi diperoleh bahwa metode yang digunakan memiliki akurasi klasifikasi sebesar 61% secara *nonreal time* dan 44% secara *real time*.

Setelah penelitian dari Tzakinesis dan Cook dilakukan, data yang digunakan dalam penelitian tersebut digunakan untuk pengenalan genre musik di tahun-tahun elahnya. Data tersebut dinamakan GTZAN yang berasal dari nama George Tzakinesis. Menurut Sturm (2013), dari 467 artikel yang diterbitkan mengenai pengenalan genre musik, 100 artikel menggunakan data GTZAN sebagai data yang digunakan sehingga menjadi data yang paling banyak digunakan. Adapun data GTZAN terdiri dari 1.000 data musik berdurasi 30 detik yang terdiri dari 10 genre, yaitu *blues*, *classical*, *country*, *disco*, *hip-hop*, *jazz*, *metal*, *pop*, *reggae*, dan *rock*.

Seiring dengan banyaknya penggunaan data GTZAN dalam topik pengenalan genre musik dan klasifikasi genre musik, banyak terbentuk data lain yang dapat digunakan dalam topik pengenalan genre musik dan klasifikasi genre musik. Salah satunya adalah data *Free Music Archive* (FMA). Data FMA adalah data publik yang mudah diakses yang dapat digunakan untuk pengambilan informasi musik atau *Music Information Retrieval* (MIR), yaitu sebuah bidang yang berkaitan dengan penelusuran, pencarian, dan pengorganisasian koleksi musik dalam jumlah besar (Defferard dkk., 2016). data FMA memiliki 4 sub data, yaitu:

1. *Full*: versi lengkap dari data FMA, terdiri dari 106.574 data audio yang belum dipangkas dengan 161 genre.

2. *Large*: data lengkap dengan audio yang dibatasi hanya berdurasi 30 detik untuk masing-masing data audio. Sub data *large* memiliki jumlah data audio yang sama seperti sub *full*.
3. *Medium*: sub data FMA yang terdiri dari 25.000 data audio berdurasi 30 detik dengan 16 genre.
4. *Small*: sub data yang terdiri dari 8.000 data audio berdurasi 30 detik dengan 8 genre. Dapat dikatakan bahwa sub data ini serupa dengan data GTZAN dengan manfaat tambahan berupa metadata yang berisi fitur-fitur audio selengkapnya yang ada di data GTZAN.

Tahun demi tahun, berbagai cara tercipta yang digunakan dalam mengklasifikasi genre musik. Zhang dkk (2019) menggunakan metode *Convolutional Neural Network* (CNN) dan model arsitektur VVG19 dalam pengenalan data audio dari Freesound Data yang telah diubah ke dalam citra spektogram. Dari penelitian tersebut, skor *label-weighted label-ranking average precision* dan akurasi yang didapat masing-masing sebesar 81,3% dan 88,9%.

Mehta dkk. (2021) melakukan klasifikasi genre musik dengan pendekatan data audio dari *data* GTZAN diubah ke dalam bentuk spektogram. Hasil berupa gambar tersebut diproses dengan *transfer learning*. Model arsitektur dalam *transfer learning* yang digunakan dalam penelitian tersebut di antaranya Resnet34, Resnet50, VGG16, dan Alexnet. Dari semua arsitektur *transfer learning* yang digunakan, akurasi dengan arsitektur Resnet34 tertinggi dibandingkan arsitektur yang lain, yakni sebesar 79%.

Penelitian yang dilakukan Simarmata dan Supriana (2023) mengklasifikasi genre musik menggunakan dua model, yaitu *random forest* dan *XGB classification* dengan sumber data dari *data* GTZAN. Dengan menggunakan dua model tersebut,

didapatkan hasil bahwa akurasi klasifikasi dengan *random forest* sebesar 72% dan akurasi klasifikasi dengan *XGB classification* sebesar 73%.

Chaudury dkk. (2022) membandingkan hasil klasifikasi genre musik dengan beberapa pengklasifikasi yang digunakan, yaitu *naïve Bayes*, *decision tree*, regresi logistik, dan *random forest*; data yang digunakan dalam penelitian tersebut menggunakan data dari GTZAN. Dari semua pengklasifikasi, *random forest* unggul dari semua pengklasifikasi dengan mendapatkan akurasi sebesar 90%.

Dalam banyak penelitian yang membahas klasifikasi genre musik dengan berbagai metode klasifikasi, spektrogram menjadi aspek penting karena kombinasi antara fitur gambar auditori dengan fitur spektral, yang keduanya bersumber dari spektrogram, menghasilkan akurasi yang lebih tinggi (Khodzhaev, 2024).

Dalam penelitian ini, dilakukan klasifikasi genre musik dengan menggunakan data suara yang akan dialih bentuk ke dalam spektrogram. Fitur-fitur di dalam spektrogram diekstraksi dengan bantuan bahasa pemrograman Python. elah fitur-fitur audio terekstraksi, klasifikasi dengan *random forest* menggunakan fungsi dari *package* di dalam Python. *Random forest* memiliki beberapa keunggulan, yaitu relatif mudah disesuaikan dengan kebutuhan klasifikasi, efektif digunakan untuk data besar yang melibatkan fitur berdimensi tinggi seperti spektrogram, dan mudah dalam interpretasi komponen-komponen yang berperan penting dalam klasifikasi (Raschka dan Mirjalili, 2017).

Sebagai pembeda dari penelitian yang dilakukan oleh Simarmata dan Supriana (2023) dan Chaudury dkk. (2022), model *random forest* dibuat ke dalam dua jenis model, yaitu model *random forest* tanpa *hyperparameter tuning* dan model *random forest* dengan *hyperparameter tuning*. Menurut Arnold dkk. (2024), *hyperparameter*

dapat memengaruhi seberapa baik model pembelajaran mesin (*machine learning*) untuk data yang berada di luar sampel (*out-of-sample data*). Selain itu, pembeda lainnya adalah penambahan data FMA sub *small* (atau dituliskan sebagai FMA *small*) dalam penelitian ini sehingga kedua model dari data yang berbeda dapat dibandingkan, baik secara akurasi maupun dari metrik evaluasi model lainnya.

Penelitian ini dilakukan sebagai bentuk kontribusi dalam pelatihan data dan meningkatkan akurasi dalam mengklasifikasi data suara. Lebih jauh, peranan penelitian ini tidak hanya sebatas digunakan dalam bidang musik dan akustik saja, melainkan dapat digunakan dalam konsentrasi lain, seperti bidang Biologi, Kedokteran, Psikologi, dan dapat digunakan untuk pengembangan kecerdasan buatan (*artificial intelligent*) yang terfokus dalam bidang suara.

1.2 Rumusan Masalah

Berdasarkan penjelasan yang telah dipaparkan, rumusan masalah dalam penelitian ini adalah bagaimana proses klasifikasi genre musik menggunakan *random forest* berdasarkan fitur-fitur audio dalam spektogram. Kemudian ingin dilihat pula tingkat ketepatan model klasifikasi genre musik yang dilakukan dengan *random forest* berdasarkan fitur-fitur audio yang diekstraksi dari spektogram, dan fitur-fitur audio apa saja yang berpengaruh terhadap klasifikasi genre musik menggunakan *random forest* berdasarkan citra audio dari spektogram.

1.3 Batasan Masalah

Penelitian ini hanya berfokus kepada membandingkan hasil akurasi model klasifikasi, baik akurasi model klasifikasi dari data *Free Music Archive* (FMA) sub data *small* maupun akurasi model klasifikasi dari data GTZAN. Selain itu, model yang digunakan dalam klasifikasi menggunakan dua data ini dibagi menjadi dua jenis, yakni

model klasifikasi dengan *hyperparameter tuning* dan model klasifikasi tanpa *hyperparameter tuning*. Model klasifikasi tersebut dibuat dengan bantuan *package* Scikit-learn dari Python.

1.4 Tujuan Penelitian

Tujuan yang hendak dicapai dalam penelitian ini, yaitu menerapkan *random forest* dalam mengklasifikasi genre musik dan mengukur ketepatan model dengan *hyperparameter tuning* dan model tanpa *hyperparameter tuning* dalam memprediksi genre musik. Setelah itu, model dengan *hyperparameter tuning* dan model tanpa *hyperparameter tuning* yang didapat dari *random forest* dan spektogram dievaluasi menggunakan metrik-metrik tertentu. Tujuan lainnya adalah untuk mengidentifikasi fitur-fitur audio yang dominan berpengaruh dalam klasifikasi genre musik menggunakan *random forest* berdasarkan citra audio dari spektogram.

