

Team ID: CSD-109

Nama Anggota:

1. M299R5310 - Muhammad Aldi Surya Putra
2. M014V6060 - Ayu Kadek Nadya Oktaviana

Tema yang dipilih: User Generated Content

Judul Proyek: Model Machine Learning Untuk Klasifikasi dan Rekomendasi Lagu Berdasarkan Genre

Executive Summary:


Mendengarkan lagu telah menjadi fenomena konstan di masyarakat. Menurut artikel penelitian (Sanfilippo, K.R.M. *et al*, 2020) menyebutkan orang dewasa mendengarkan musik atau lagu rata-rata selama 18 jam dalam seminggu. Artikel tersebut juga mengatakan bahwa ketika sebuah lagu dipilih oleh pendengar, respons emosional terhadap lagu tersebut lebih kuat. Oleh karena itu, kami berencana untuk membuat model *machine learning* yang dapat mengklasifikasikan dan merekomendasikan lagu berdasarkan genre. Genre yang digunakan dalam proyek kami terbatas sehingga beberapa lagu dengan genre tertentu tidak bisa digunakan dalam model sistem rekomendasi. Kami berharap model yang kami kerjakan dapat membantu memberikan rekomendasi lagu yang benar berdasarkan preferensi pengguna.

Project Scope & Deliverables:

Secara umum, kami akan membuat sebuah model *machine learning* yang mengklasifikasikan dan merekomendasikan lagu berdasarkan genre. Klasifikasi lagu yang kami kerjakan terbatas hanya beberapa genre saja sehingga pada sistem rekomendasi yang kami kerjakan tidak bisa melakukan rekomendasi lagu dengan genre-genre lagu tertentu.

Proyek ini dimulai setelah dinyatakan aman dalam artian tidak ada revisi dari tim Panitia Dicoding. Dalam waktu 4-5 minggu yang diberikan, kami berharap bisa menyelesaikan kedua model *machine learning* ini. Pada minggu pertama, kami mulai membuat model klasifikasi lagu. Minggu kedua kami melakukan evaluasi terhadap model agar hasil klasifikasi menjadi lebih baik dari sebelumnya. Dari minggu ketiga, kami mulai mengembangkan model sistem rekomendasi. Pada minggu keempat, kami mengerjakan dokumentasi yang diperlukan seperti video, slide, dan lain-lain. Rincian kegiatan yang akan dilakukan dapat ditemukan di bagian *Project Schedule*.

Project Schedule:

 Project Schedule Capstone

Project Resources:

Dalam pembuatan dan pengembangan *project* ini, terdapat beberapa *resource* yang akan digunakan sebagai bahan maupun referensi agar perkembangan *project* ini dapat terlaksana dengan baik dan tepat waktu, antara lain :

1. Bahasa Pemrograman :

- Python 3.7.12
2. Framework / Library :
    - Tensorflow 2.7.0
    - Pandas 1.1.5
    - Numpy 1.19.5
    - Matplotlib 3.2.2
    - Scikit-Learn 1.0.1
    - Seaborn 0.11.2
  3. Dataset :

Karena kami akan membuat 2 model yaitu model klasifikasi dan juga model rekomendasi, jadi dalam membangun kedua model tersebut kami menggunakan 2 dataset yang berasal dari kaggle antara lain,

    1. GTZAN Dataset

Dataset ini berisi audio dari tiap kategori genre, dalam dataset ini terdapat 10 genre yaitu, blues, classical, country, disco, hiphop, jazz, metal, rock, pop, dan reggae. Jumlah audio dalam dataset ini berjumlah 1000 audio dengan rincian 100 audio pada setiap genre. Dataset ini digunakan sebagai bahan dalam training model pertama yaitu klasifikasi musik/lagu
    2. Top Spotify Songs from 2010 - 2019

Dataset ini berisi data lagu pada aplikasi *spotify*, jumlah lagu yang ada pada dataset ini berjumlah 603 lagu. Dataset ini digunakan sebagai output pada model kedua yaitu sistem rekomendasi.
  4. Artikel

Terdapat beberapa artikel yang digunakan sebagai referensi dalam *project* ini antara lain,

    - [Music Genre Classification using python](#)
    - [How to build a simple song recommender](#)

#### Risk and Issue Management Plan:

- Dalam pengembangan model klasifikasi, data yang digunakan untuk training model berupa audio sehingga kami akan cukup kesulitan dalam proses implementasinya karena kami belum cukup ahli dalam pembuatan *Machine Learning Model*
- Karena saat ini kita masih berada dalam situasi pandemi, kami cukup kesulitan dalam berkomunikasi satu sama lain karena hanya bisa mengandalkan *platform online meeting* dan *Whatsapp Messenger*. Serta kami memiliki perbedaan waktu wilayah yaitu WIB dan WITA sehingga kami merasa cukup kesulitan dalam menentukan waktu yang tepat untuk mendiskusikan *progress* kedepannya.
- Dalam rencana kami, *project* ini tidak akan sampai tahap deployment melainkan hanya sampai tahap evaluasi model mengingat terbatasnya waktu dalam pengerjaan *project* ini.