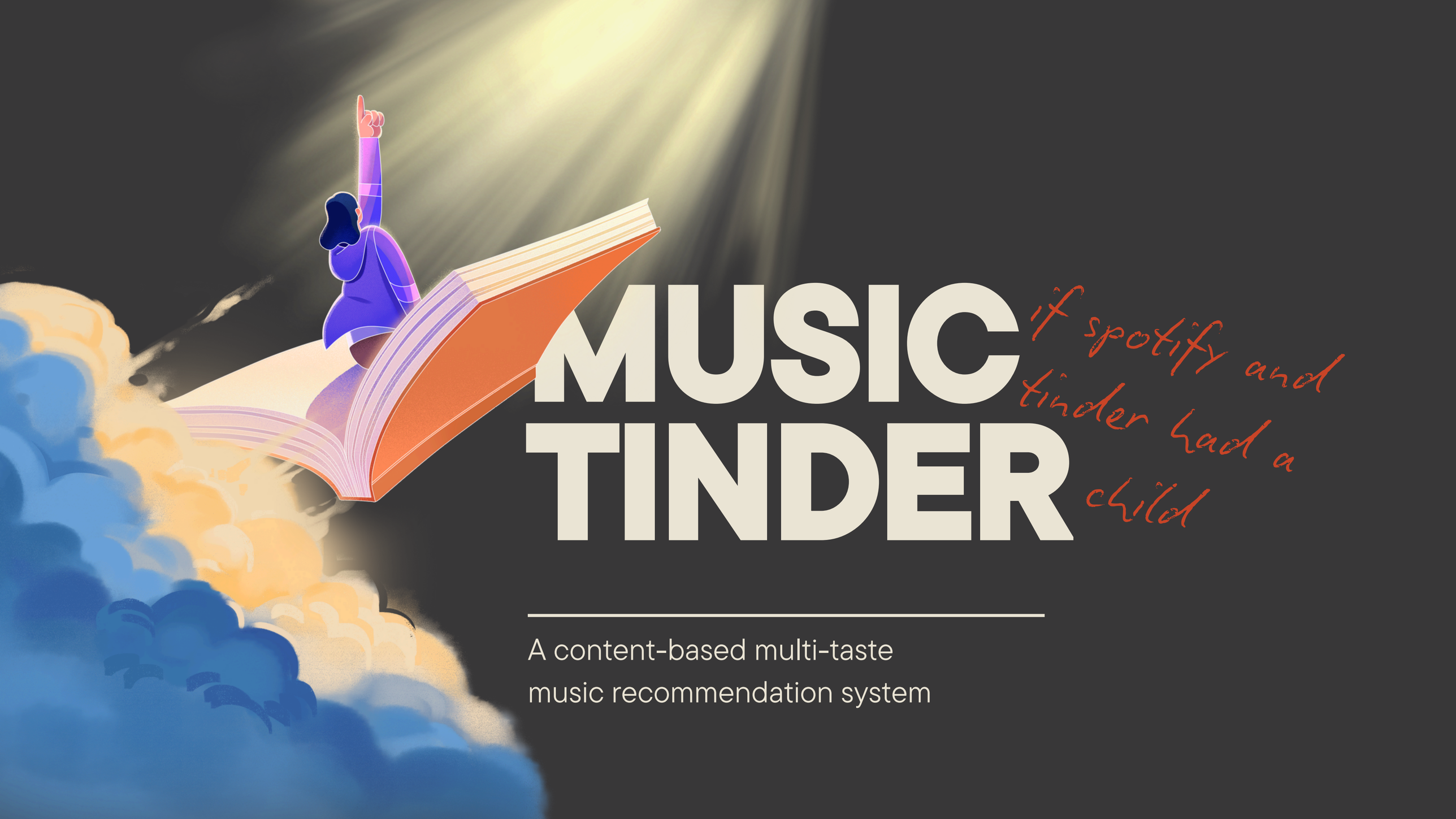


# MUSIC TINDER

*if spotify and tinder had a child*

A content-based multi-taste
music recommendation system

# MEMBER LIST

**Mario**
Roggi

**Luca**
Sartorelli

**Gianmario**
Palleschi

# LIST OF CONTENT

# INTRODUCTION

# INTRODUCTION

Music recommendation systems are a central component of modern streaming platforms. Their goal is to **reduce information overload** by selecting a **small subset** of tracks that are likely to match a user's preferences.

This project explores how recommendation systems behave when **social data**, **user registration**, and **large-scale online signals** are **completely removed**, focusing instead on **pure content-based personalization**.

# BACKGROUND

Commercial platforms such as Spotify heavily rely on **collaborative filtering**, **large user bases**, and **continuous data collection**. These systems benefit from millions of users, shared playlists, listening history across devices, and real-time feedback.

In contrast, this project investigates a **local**, **offline-capable** recommender system, designed for a **single user**, where all **preferences are learned** exclusively through explicit feedback on songs and their audio features.

# PROBLEM FORMULATION

Traditional content-based recommenders often represent a user with **a single preference vector**, implicitly assuming musical taste is **linear and coherent**.

In reality, users frequently enjoy **multiple, disconnected genres** (for example classical music and heavy metal), which cannot be accurately captured by a **single centroid**.

The problem addressed is how to model **non-linear**, **multi-taste preferences** without using collaborative or social data.

# RESEARCH

# Research OBJECTIVES

To design a recommender system that works without user accounts or shared data

To model multiple independent musical tastes per user avoiding a single centroid vector

To compare simple similarity methods with supervised machine learning approaches

To evaluate how recommendation quality evolves with increasing feedback

# Research METHODS

Songs are represented as numerical vectors derived from audio features such as energy, valence, danceability, tempo, and loudness.

Both similarity-based ranking and machine-learning models are employed.

User feedback (like/dislike) is simulated or collected explicitly and used to:

1. update taste profiles
2. train supervised models when sufficient data is available
3. dynamically switch between recommendation strategies

# Research BENEFIT

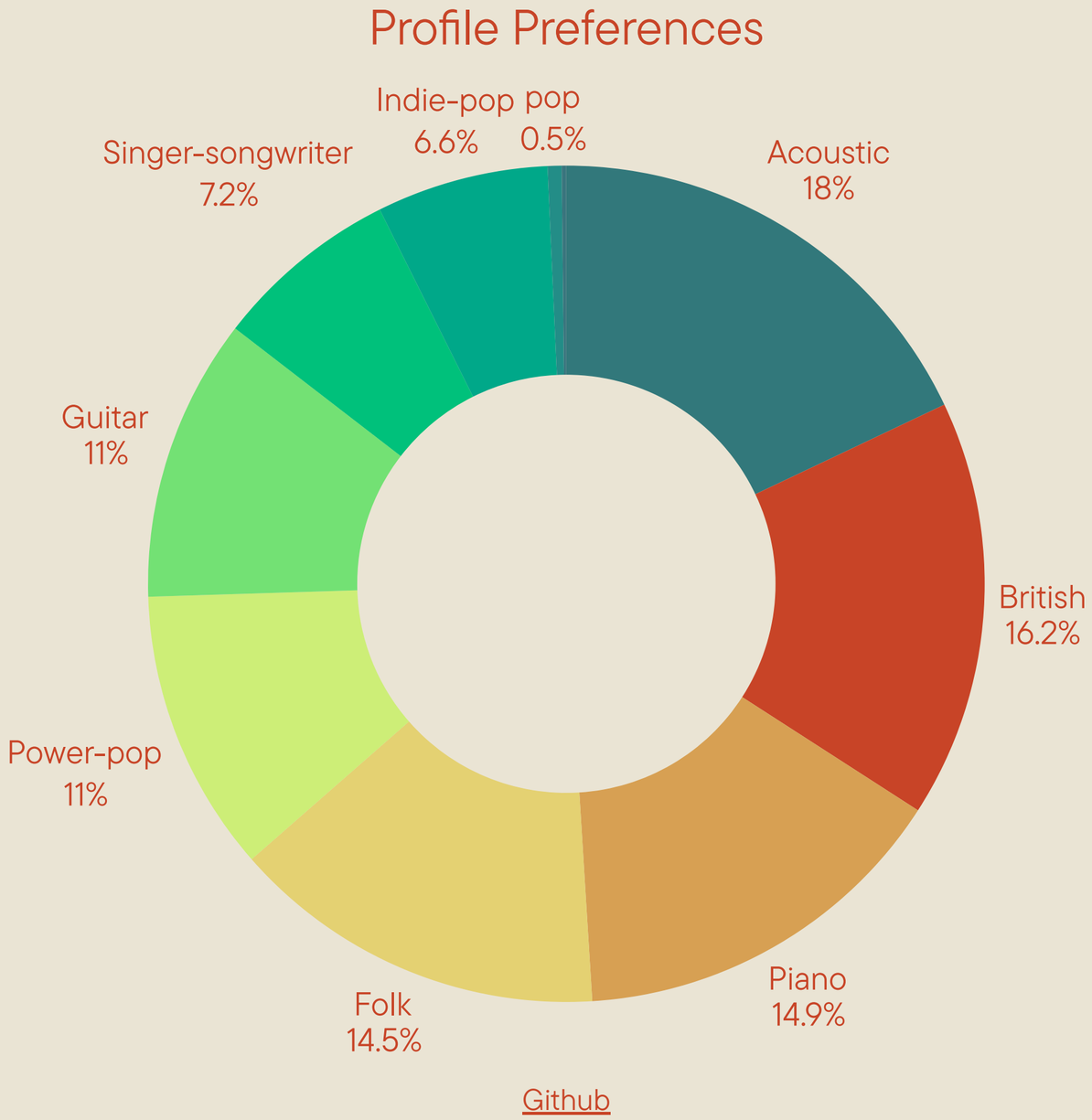
From an **educational perspective**, this project provides a clear view of how recommendation systems function internally, without abstraction layers imposed by large platforms.

It highlights the **strengths** and **weaknesses** of content-based filtering and demonstrates why large-scale systems rely on hybrid and **collaborative approaches**.

# Research FLOW

1. Initial cold-start recommendations using genre diversity
2. Collection of explicit user feedback
3. Clustering of preferences into independent taste profiles
4. Similarity-based recommendation within each taste
5. Transition to supervised learning models as data grows

This flow allows gradual improvement without requiring large datasets upfront.




Github

# Research RESULT

The system successfully adapts to users with **multiple distinct tastes**, producing more coherent recommendations than a **single-vector** approach.

However, results also show **instability** when **feedback is sparse** and highlight sensitivity to feature selection and noise.

**Performance improvements** are noticeable once **supervised models** are introduced, but remains limited by **dataset size**.

# SUMMARY

# CONCLUTION & RECOMMENDATION

This model does **not match Spotify-level** performance because it **lacks**: collaborative filtering, cross-user behavioral patterns, large-scale interaction data and real-time online learning

Some users prefer large platforms for accuracy and discovery, while others may prefer simple, private, offline systems for transparency and control.
Future work should explore **hybrid models** and **neural approaches**.

# QUESTION & ANSWER

**QUESTIONS ARE WELCOME**

- Why those algorithms?
- What are the limitations?
- How does it compare with commercial systems?
- what are your plans on improvements?

# THANK YOU

+39 338 7535149

sartorelli.luca@outlook.com

Borgo Vittoria, torino, TO