

Melodia

Arthur Santos Afonso Ferreira¹, Ian Theodoro Campanha¹, Kenny Jun Takahashi¹, Yuri Nichimura Alves¹

¹Curso de Ciência da Computação Universidade Presbiteriana Mackenzie (UPM) – São Paulo, SP – Brasil

{10381332, 10381723, 10401701, 10396373}@mackenzista.com.br

Abstract. *This project proposes the development of a music recommendation system using Machine Learning (ML) and the Spotify API, focusing on analyzing user preferences and audio features (e.g., rhythm, energy, valence). The initiative stems from the current challenges in delivering personalized music suggestions that truly align with individual tastes. By leveraging real user data and Spotify's audio analysis capabilities, the system aims to enhance recommendation accuracy through advanced ML techniques. The expected outcome is an improved musical experience through more precise and relevant recommendations, thereby contributing to the advancement of recommendation systems in digital entertainment.*

Resumo. *Este projeto propõe o desenvolvimento de um sistema de recomendação musical utilizando Machine Learning (ML) e a API do Spotify, com foco na análise de preferências dos usuários e características das músicas (ex.: ritmo, energia, valência). A iniciativa surge dos atuais desafios em fornecer sugestões musicais personalizadas que realmente correspondam aos gostos individuais. Ao utilizar dados reais de usuários e os recursos de análise de áudio do Spotify, o sistema visa melhorar a precisão das recomendações por meio de técnicas avançadas de ML. Espera-se que o projeto desenvolvido melhore a experiência musical através de recomendações mais precisas e relevantes, contribuindo assim para o avanço das técnicas de recomendação no âmbito do entretenimento digital.*

Introdução

Contextualização

Com o crescimento exponencial de plataformas de streaming musical, como o Spotify, a quantidade de músicas disponíveis tornou-se praticamente ilimitada. No entanto, essa abundância de escolhas trouxe um novo desafio: como ajudar os usuários a descobrir músicas que realmente se alinhem com seus gostos pessoais?

Os sistemas de recomendação atuais, embora úteis, frequentemente falham em capturar a complexidade das preferências individuais, gerando sugestões genéricas ou repetitivas. Nesse cenário, a inteligência artificial (IA) e o Machine Learning (ML) surgem como ferramentas promissoras para analisar padrões complexos em grandes volumes de dados, como histórico de reprodução e características musicais (ex.: ritmo, energia, tom). A API do Spotify oferece acesso a esses dados, permitindo a criação de modelos mais precisos e personalizados.

Justificativa

A motivação deste projeto baseia-se em três pilares principais:

- Limitações dos sistemas atuais: Algoritmos tradicionais de recomendação muitas vezes priorizam popularidade em vez de relevância individual.
- Oportunidade tecnológica: A API do Spotify fornece dados ricos sobre músicas e usuários, ainda subutilizados em projetos independentes.
- Impacto na experiência do usuário: Recomendações precisas podem aumentar o engajamento e a satisfação em plataformas de streaming.

Além disso, o projeto contribui para o campo acadêmico, explorando técnicas avançadas de ML aplicadas ao entretenimento digital – uma área em crescimento, mas ainda com muitas lacunas de pesquisa.

Objetivo

Este projeto tem como objetivo principal desenvolver um sistema de recomendação musical inteligente que:

- Integre dados da API do Spotify para acessar:
 - Características técnicas das músicas (bpm, energia, dançabilidade);
 - Histórico de reprodução e preferências dos usuários;
- Implemente algoritmos de Machine Learning para:

Analisar padrões complexos de comportamento musical;

Identificar relações não óbvias entre diferentes gêneros e estilos;

Personalizar recomendações baseadas no perfil individual;

- Superar as limitações dos sistemas atuais por meio de:

Melhor diversidade nas recomendações;

Adaptação dinâmica às mudanças de preferência;

Opção do Projeto

A escolha por focar no Spotify deve-se a:

-Acessibilidade: Sua API é bem documentada e oferece dados granulares sobre músicas e usuários (com permissão);

-Relevância: É a plataforma de streaming mais popular globalmente, com milhões de playlists públicas para análise;

-Inovação: Possibilita explorar features avançadas não disponíveis em outras APIs;

Descrição do Problema

Os sistemas atuais de recomendação musical enfrentam desafios significativos na era do streaming, onde a abundância de opções dificulta a descoberta de conteúdo relevante. As principais limitações incluem: personalização superficial baseada apenas em popularidade e histórico básico, recomendações repetitivas que criam "bolhas musicais", e dificuldade em adaptar-se às mudanças de preferência dos usuários. Dessa maneira, os problemas se agravam pela alta complexidade dos dados musicais e pela subjetividade na avaliação de qualidade.

Os sistemas existentes falham em integrar adequadamente três dimensões cruciais: comportamento do usuário, características das músicas e contexto de escuta. Além disso, carecem de adaptatividade para evoluir com novas interações e usam métricas que privilegiam engajamento em detrimento da satisfação real.

Estas limitações resultam em frustração por parte de usuários, menor visibilidade para artistas menos conhecidos e estagnação tecnológica no setor. Nosso projeto propõe superar esses obstáculos através de uma abordagem que combina a análise avançada de dados da API do Spotify com modelos de Machine Learning, focando não apenas no "o que é popular", mas no "porquê" determinadas músicas ressoam com cada ouvinte.

Aspectos Éticos e Responsabilidade no Desenvolvimento

Os sistemas de recomendação musical com IA enfrentam desafios éticos complexos que exigem soluções equilibradas. Na privacidade, a conformidade com leis como LGPD e GDPR é fundamental, pois os padrões de escuta revelam aspectos pessoais sensíveis - exigindo anonimização rigorosa e consentimento transparente. O viés algorítmico representa outro obstáculo, com riscos de reforçar artistas populares e marginalizar nichos culturais, problema que demanda técnicas avançadas de Machine Learning justo combinadas com curadoria humana especializada.

A responsabilidade contínua completa este quadro, com a necessidade de testes abrangentes, canais de feedback eficazes e atualizações periódicas para corrigir distorções. A solução ideal integra expertise técnica, ética aplicada e profundo entendimento musical, visando sistemas que enriqueçam a experiência do usuário enquanto respeitam diversidade cultural e privacidade individual. Este equilíbrio delicado entre personalização e responsabilidade social define o futuro ético da recomendação musical automatizada.

Descrição do Dataset

Para o desenvolvimento do sistema de recomendação de músicas, foi utilizado um conjunto de dados extraído da conta pessoal do usuário por meio da API do Spotify. Os dados foram coletados de forma anonimizada, respeitando os princípios éticos de privacidade e segurança das informações.

A coleta dos dados foi realizada utilizando o script `coleta_dados.py`, o qual realiza a autenticação na API do Spotify por meio do protocolo OAuth2. Após a autenticação, o script acessa as músicas salvas na biblioteca do usuário, extraindo características musicais disponibilizadas pela API, tais como:

- danceability: nível de dançabilidade da música;
- energy: intensidade percebida da faixa;
- speechiness: presença de palavras faladas;
- acousticness: probabilidade de a faixa ser acústica;
- instrumentalness: ausência de vocais;
- liveness: probabilidade de a faixa ter sido gravada ao vivo;
- valence: positividade emocional transmitida pela música;
- tempo: tempo (BPM) da faixa;
- duration_ms: duração total da música em milissegundos.

Adicionalmente, informações como `track_id`, `track_name`, `artist` e um identificador anônimo do usuário (`user_anon_id`) também foram armazenadas. Os dados extraídos foram salvos no arquivo `spotify_anonimizado.csv`, localizado na pasta `dataset/`.

A análise exploratória foi realizada no script `analise_dados.py`. Este script tem como finalidade explorar estatisticamente e graficamente os dados obtidos, com o objetivo de compreender melhor sua estrutura e identificar possíveis padrões que serão úteis nas etapas posteriores do projeto.

A análise contemplou:

- Exibição das primeiras entradas do conjunto de dados (`head()`);
- Geração de estatísticas descritivas (média, desvio padrão, valores máximos e mínimos);
- Visualização gráfica da distribuição das características musicais por meio de histogramas;
- Análise de correlação entre as variáveis numéricas por meio de uma matriz de calor (`heatmap`);
- Normalização dos dados utilizando o `StandardScaler`, do `scikit-learn`, com o intuito de padronizar os valores das variáveis para posterior aplicação de algoritmos de Machine Learning.

Essa etapa permitiu obter uma visão geral do perfil musical do usuário e preparar os dados de forma adequada para alimentar os modelos de recomendação que serão desenvolvidos na fase seguinte do projeto.

Metodologia e Resultados Esperados

O projeto foi dividido em duas etapas principais: coleta e análise de dados. Na primeira, utilizamos a API do Spotify para extrair informações de músicas salvas por um usuário, como nome, artista e características sonoras (dançabilidade, energia, etc.). Na segunda etapa, foi feita uma análise exploratória, com gráficos, estatísticas e correlações entre as variáveis. Os dados também foram normalizados para uso posterior em modelos de recomendação.

Como resultado esperado, pretende-se que, em uma próxima etapa, o sistema seja capaz de recomendar músicas semelhantes às preferidas pelo usuário, utilizando os dados tratados e organizados. A expectativa é mostrar que, com o uso de dados reais e algumas técnicas de Inteligência Artificial, é possível oferecer sugestões personalizadas de forma eficaz.

link GitHub: <https://github.com/junntk/Projeto-IA---Melodia>

Bibliografia

Oliveira, M. (2024) “Como Usar a API Web do Spotify | Integre Facilmente com o Spotify “. Disponível em: <https://apidog.com/pt/blog/spotify-web-api/>.

Alkhovik, R. (2024) “O que é a API da Web do Spotify: automatizando análises de música e experiências do usuário “. Disponível em: <https://latenode.com/pt-br/blog/what-is-the-spotify-web-api-automating-music-analytics-and-user-experiences>.

Loyola, N., Macohin, A., Barbosa, I., Fagundes, V., Graça, F. (2023) “Ética na IA”. Disponível em: <https://www.serpro.gov.br/menu/noticias/noticias-2023/etica-na-ia> .

UNICEP (2024) “Inteligência Artificial e Ética: Conheça o impacto ético da IA na sociedade.”. Disponível em: <https://www.unicep.edu.br/post/inteligencia-artificial-e-etica-conheca-o-impacto-etico-da-ia-na-sociedade>.

Nardon, B. (2023) “Inteligência Artificial: oportunidades e desafios éticos para uma sociedade que emerge”. Disponível em: <https://mittechreview.com.br/inteligencia-artificial-oportunidades-e-desafios-eticos-para-uma-sociedade-que-emerge/>.

Taraves, J. (2024) “Quais são os desafios éticos da Inteligência Artificial?”. Disponível em: <https://www.mindtek.com.br/2024/12/quais-sao-os-desafios-eticos-da-inteligencia-artificial/>.

Moraes, H., Szentmiklósy, J. (2024) “Desafios éticos do uso de IA para o tratamento de dados pessoais”. Disponível em: <https://www.migalhas.com.br/depeso/405003/desafios-eticos-do-uso-de-ia-para-o-tratamento-de-dados-pessoais>.