

**UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO**

Bruno Antônio Vieira

**QUAIS CARACTERÍSTICAS SÃO MAIS RELEVANTES
PARA A POPULARIDADE DE UMA MÚSICA EM DIFERENTES PAÍSES**

Proposta de projeto científico de
Monografia do bacharelado em Sistemas
de Informação da Universidade Federal
de Minas Gerais.

Orientador:
Prof. Dr. Flavio Vinicius Diniz de Figueiredo

**Belo Horizonte
2020**

INTRODUÇÃO

O Spotify é um serviço de streaming de música e, mais recentemente, podcasts que dá acesso à milhões de conteúdos por meio de um plano gratuito ou por opções pagas. Dentro dessa grande quantidade de conteúdo disponível, a plataforma disponibiliza meios de sugerir músicas mais relevantes para seus usuários utilizando sistemas de recomendação e também fornecendo *playlists* com as paradas de sucesso atuais.

Conhecer os critérios e atributos que fazem uma música se tornar mais popular é um meio interessante, como mencionado por Al-Beitawi et al. [3], já que este conhecimento pode fomentar melhorias na experiência e na satisfação do usuário final, garantindo melhores recomendações. O próprio Spotify disponibiliza por meio de uma API (Interface de Programação de Aplicações) alguns atributos referentes à cada música, caracterizando-as nos quesitos: *acousticness*, *danceability*, *energy*, *instrumentalness*, *liveness*, *loudness*, *speechness*, *valence* e *tempo*. Estes atributos, apesar de representarem qualitativamente características de cada música, são expressos em escalas quantitativas que serão melhor exemplificadas ao decorrer deste trabalho.

Com isso, pretende-se explorar, utilizando técnicas de aprendizado supervisionado e não supervisionado, a relação entre um conjunto de músicas disponíveis em paradas musicais de diferentes regiões e a sua popularidade nos respectivos países. A partir dessa exploração, busca-se entender as diferenças culturais entre estes locais em relação à musicalidade e, possivelmente, prever o padrão de cada nação com base nas características musicais fornecidas pelo Spotify.

TRABALHOS RELACIONADOS

Alguns trabalhos sobre a popularidade de músicas no Spotify e a análise de seus atributos já foram conduzidos, seguindo diferentes perspectivas. No trabalho de Suh [2], por exemplo, foram utilizados os atributos quantitativos já mencionados anteriormente (*acousticness*, *danceability*, *energy*, *instrumentalness*, *liveness*, *loudness*, *speechness*, *valence* e *tempo*) para determinar o sucesso de faixas na parada musical das 200 melhores músicas do Spotify, utilizando regressões e levando em consideração as divergências em cinco países diferentes, tendo como fator de popularidade a estabilidade da faixa em uma determinada posição na parada musical.

Já no trabalho de Al-Beitawi et al. [3], apesar do uso dos mesmos atributos quantitativos, a abordagem escolhida para análise foi a separação das faixas em clusters e a utilização do algoritmo K-means. Os resultados mostraram que os clusters que possuem músicas com o quesito *danceability* mais alto tendem a estar no topo de paradas musicais.

Um dos papéis deste trabalho é analisar as diferenças culturais na esfera musical em diferentes países e, com isso, é relevante entender como a música popular evolui ao longo do tempo e se transforma, assim como a cultura de um povo no geral. Os trabalhos de Mauch et al. [11] e de Vigliensoni e Fujinaga [5] investigam essa evolução por meio de uma análise da parada musical Billboard Hot 100 entre os anos de 1960 e 2010 e pela criação de um dataset unificado utilizando dados extraídos do Last.fm, respectivamente. Em ambos os trabalhos temos um retrato da realidade musical, sendo o primeiro mais focado nos pontos-chaves de mudanças, como as "revoluções" do pop ocorridas em 1964, 1983 e 1991 e o segundo focando em características mais demográficas, produzindo um dataset útil para eventuais análises.

Assim como este trabalho, três estudos levaram em consideração a relação entre o gosto musical e a cultura de um país. Em [8], Schedl et al. utilizaram a correlação de Spearman e QAP (Quadratic Assignment procedure) para medir a correlação dos fatores socioeconômicos de um país com o seu gosto musical, gerando resultados que podem melhorar sistemas de recomendação baseados em

cultura. O artigo [9] seguiu o mesmo ideal, gerando indicadores curiosos, como a falta de relação significativa entre fatores geoeconômicos e as preferências por artistas, álbuns e gêneros. Por fim, o artigo [10] utiliza as mesmas características musicais já mencionadas neste texto em conjunto com fatores culturais e socioeconômicos para aumentar a acurácia de um algoritmo de recomendação também baseado em cultura. O modelo foi desenvolvido usando dados sobre reprodução de música de cerca de 55 mil usuários de 36 países e, em seguida, foi aplicado em um sistema de recomendação já existente.

Em conclusão, os artigos [7] e [6], apesar de não analisarem dados sobre preferências musicais, identificam outros padrões humanos relacionados à cultura e fatores socioeconômicos. Os algoritmos utilizados para mensuração de hábitos alimentares, assim como, padrões de comportamento no Twitter são úteis para influenciar este trabalho, além de nortear formas de lidar com os aspectos culturais presentes na pesquisa.

METODOLOGIA

Primeiramente, foram selecionados alguns artigos relacionados com o tema proposto, buscando um bom embasamento para desenvolvimento deste trabalho. Junto com a seleção desses artigos foi selecionado um dataset contendo cerca de 2 milhões de linhas com o ranking diário das 200 músicas mais ouvidas em 53 países nos anos de 2017 e 2018.

A partir deste dataset, serão coletados dados mais refinados de cada música, contendo seus atributos acústicos, por meio do uso da API do Spotify. Com estes atributos será realizada uma análise exploratória a fim de investigar e caracterizar os dados existentes.

O terceiro passo será o uso de técnicas de aprendizado supervisionado e não supervisionado para classificar as músicas presentes no conjunto de dados e entender, com base em suas características, como elas divergem entre países diferentes.

Por fim, será desenvolvido um modelo que tente prever em qual país uma música tende a ser mais popular com base em suas características. Pretende-se fazer uso do algoritmo SVM forest e, possivelmente, comparar seu desempenho com outras soluções.

RESULTADOS ESPERADOS

Ao fim do semestre, espera-se responder a questão levantada no título deste artigo sobre quais características são mais relevantes para a popularidade de uma música em cada país. Essa resposta deverá ser contemplada por meio de informações estatísticas e gráficos que demonstrem a padronização ou não dos dados diante das variáveis utilizadas.

Além disso, por meio da verificação de padrões em regiões diferentes do mundo é possível visualizar discrepâncias culturais. A partir destes dados distintos de cada região, diferentes estratégias publicitárias para atender o público podem ser tomadas pelo Spotify para melhorar o seu serviço, incluindo a melhoria de sistemas de recomendação.

ETAPAS E CRONOGRAMA

| Atividades | Ago | Set | Out | Nov |
|-------------------------------------|-----|-----|-----|-----|
| Leitura de artigos relacionados | | | | |
| Montagem do dataset | | | | |
| Análise exploratória | | | | |
| Aplicação de modelos de aprendizado | | | | |
| Revisão e entrega do trabalho | | | | |

REFERÊNCIAS

- [1] O que é o Spotify?. Disponível em:
<https://support.spotify.com/br/using_spotify/getting_started/what-is-spotify/>
Acesso em: 19 Ago. 2020
- [2] SUH, Brendan Joseph. International Music Preferences: An Analysis of the Determinants of Song Popularity on Spotify for the US, Norway, Taiwan, Ecuador, and Costa Rica. 2019.
- [3] AL-BEITAWI, Zayd; SALEHAN, Mohammad; ZHANG, Sonya. Cluster Analysis of Musical Attributes for Top Trending Songs. In: **Proceedings of the 53rd Hawaii International Conference on System Sciences**. 2020.
- [4] ARAUJO, Carlos Soares; CRISTO, Marco; GIUSTI, Rafael. Predicting Music Popularity on Streaming Platforms. In: **Anais do XVII Simpósio Brasileiro de Computação Musical**. SBC, 2019. p. 141-148.
- [5] VIGLIENSONI, Gabriel; FUJINAGA, Ichiro. The Music Listening Histories Dataset. In: **ISMIR**. 2017. p. 96-102.
- [6] GARCIA-GAVILANES, R. O. G. G.; QUERCIA, Daniele; JAIMES, Alejandro. Cultural dimensions in twitter: Time, individualism and power. 2013.
- [7] SILVA, Thiago H. et al. You are what you eat (and drink): Identifying cultural boundaries by analyzing food & drink habits in foursquare. **arXiv preprint arXiv:1404.1009**, 2014.
- [8] SCHEDL, Markus et al. Indicators of country similarity in terms of music taste, cultural, and socio-economic factors. In: **2017 IEEE International Symposium on Multimedia (ISM)**. IEEE, 2017. p. 308-311.
- [9] LIU, Meijun; HU, Xiao; SCHEDL, Markus. The relation of culture, socio-economics, and friendship to music preferences: A large-scale, cross-country study. **PloS one**, v. 13, n. 12, p. e0208186, 2018.

- [10] ZANGERLE, Eva; PICHL, Martin; SCHEDL, Markus. User Models for Culture-Aware Music Recommendation: Fusing Acoustic and Cultural Cues. **Transactions of the International Society for Music Information Retrieval**, v. 3, n. 1, 2020.
- [11] MAUCH, Matthias et al. The evolution of popular music: USA 1960–2010. **Royal Society open science**, v. 2, n. 5, p. 150081, 2015.