

PIBIFSP	PROJETO DE PESQUISA
---------	---------------------

TÍTULO DO PROJETO: Sistema de recomendação de músicas baseado na correlação entre áudios e arquivos digitais.

Área do Conhecimento (Tabela do CNPq):	1	.	0	3	.	0	4	.	0	2	-	9
--	---	---	---	---	---	---	---	---	---	---	---	---

1. RESUMO

Nos últimos anos, o consumo de música *online* aumentou consideravelmente e, com isso, muitos sistemas de *stream* de música passaram a não só disponibilizar as músicas para serem ouvidas, mas também recomendá-las de acordo com o perfil dos usuários. A recomendação de músicas podem ser feitas baseada na opinião de outros usuários, no conteúdo das músicas ou em abordagens híbridas. Este trabalho tem o objetivo de analisar a relação entre o áudio das músicas e o padrão binário dos arquivos digitais. Caso haja essa correlação, será possível criar um sistema de recomendação baseado somente nessa informação.

2. FUNDAMENTAÇÃO TEÓRICA

Com o aumento do consumo de música *online*, surgem também os sistemas de recomendação destas. Várias são as técnicas de recomendação utilizadas para música, sendo as mais simples aquelas que agrupam as músicas em relação ao seu gênero (*pop*, *rock*, etc). Entretanto, existem técnicas que exploram outras formas de se correlacionar músicas (ou seus arquivos digitais), como apresentado em [1].

Algumas dessas técnicas são:

- Filtragem colaborativa (*Collaborative Filtering* - CF): uma técnica comum em vários contextos, não só para recomendação de música. Basicamente é baseada na recomendação de usuários, cruzando essas recomendações para usuários com perfis similares [2]. É uma estratégia que não leva em conta o conteúdo daquilo que é recomendado, e sim o perfil de quem gostou ou não gostou;
- Baseada em conteúdo (*Content-Based approach* - CB): se baseia no conteúdo da música, ou seja, som, gênero, etc. Essa abordagem é menos utilizada, pois é mais complexa que a CF. Essa técnica é derivada de técnicas de recuperação de informação [3];
- Abordagem Híbrida: trata-se de um híbrido em relação às duas técnicas anteriores.

Em relação à abordagem baseada em conteúdo, que é foco deste trabalho, como descrito em [1], há alguns sistemas que a utilizam como, por exemplo, o Pandora¹ [4]. Esse sistema recomenda músicas para os usuários baseando-se em

¹ Pandora (<http://www.pandora.com>) é um sistema de *stream* de música, com recomendação baseada em conteúdo. Não está disponível para o Brasil. Acessado em 15/05/2020.

características como tom da música, *riffs* de guitarra, arranjo instrumental, etc. Essas características são comparadas com as de músicas indicadas pelo usuário como aquelas que ele ou ela gosta e então é feita a recomendação. Para esse tipo de abordagem, o Pandora utiliza a base de dados da Music Genome Project².

Outro sistema que utiliza essa abordagem é o Musicoverly³, que classifica as músicas em um plano cartesiano de duas dimensões, sendo um eixo representado por “calmo – energético” e “negativo – positivo” (tradução do autor para “*calm – energetic*” e “*dark – positive*”). As músicas são classificadas nesse plano de acordo com seu conteúdo sonoro.

Este trabalho tem como objetivo ir um pouco mais a fundo no conteúdo do arquivo digital que contém a música, investigando o padrão binário do arquivo e relacionando esse padrão com a sua sonoridade. Dessa forma, caso essa relação existir, será possível identificar se duas músicas são similares (ou dissimilares) e o grau dessa similaridade, somente analisando o padrão binário dos arquivos digitais.

O primeiro passo dessa pesquisa será a investigação dos formatos digitais de música mais comuns. Por exemplo, é sabido que um dos formatos mais populares, o MP3, executa uma compressão do áudio, para que os arquivos diminuam em tamanho. Dessa forma, é provável que esse formato, devido à compressão, elimine essa relação entre a representação binária e o áudio que representa. Entretanto, outros formatos também utilizados, como o WAVE, preservam melhor essa relação.

Uma vez identificado quais formatos possuem essa relação, seria possível, por meio de funções de similaridade, identificar o grau de similaridade entre dois arquivos digitais de música e, dessa forma, baseado nas músicas que um usuário aprecie, recomendá-la ou não.

Funções de similaridade são funções que tem a seguinte forma:

$$\text{sim}(a, b) = [0, 1]$$

sendo que se $\text{sim}(a, b) = 1$, a e b são totalmente similares e $\text{sim}(a, b) = 0$, a e b são totalmente dissimilares. Existem diversas funções de similaridade para valores binários como, por exemplo, a similaridade do cosseno, que mede o cosseno do ângulo entre vetores binários (que varia de 0 a 1) e a similaridade de Hamming, que, também para sequências binárias, calcula a razão entre o número de bits iguais nas mesmas posições pelo número total de bits (também normalizado em uma escala de 0 a 1).

Este trabalho tem como objetivo analisar os principais formatos digitais de áudio e verificar se há relação entre o áudio contido no arquivo e seu padrão binário. Caso haja essa relação, utilizar funções de similaridade para mensurar a similaridade entre arquivos de áudio e comparar a qualidade dessa abordagem com outras. Caso a correlação não se comprove, será feita estudada uma estratégia de se criar

² Music Genome Project - <https://www.pandora.com/about/mgp> . Acessado em 15/05/2020.

³ Musicoverly - <http://b2b.musicoverly.com/> . Acessado em 15/05/2020.

assinaturas digitais das músicas, de forma a representar seu conteúdo sonoro em *strings* binárias e usar essas assinaturas para a comparação de similaridade.

3. OBJETIVOS

Objetivo geral: Verificar se é possível a recomendação de músicas baseado somente nas características binárias de arquivos de áudio.

Objetivos específicos:

- compreender os principais formatos de áudio utilizados atualmente;
- verificar se há correlação entre o padrão binário dos arquivos de áudio e o som que os mesmos reproduzem;
- verificar quais as funções de similaridade para valores binários são as mais adequadas para esse contexto;
- testar o resultado da aplicação dessas funções de similaridade em arquivos de áudio;
- comparar essa abordagem com outras abordagens baseada em conteúdo existentes na literatura;
- analisar a qualidade dos resultados obtidos, do ponto de vista da percepção das pessoas em relação à similaridade entre os áudios.

4. MATERIAIS E MÉTODOS

O aluno de iniciação científica utilizará computador e acesso a Internet para executar o projeto.

O aluno fará o levantamento bibliográfico, com supervisão do orientador, sobre o estado da arte em relação a representação de áudio em forma digital e sobre funções de similaridade que possam ser utilizadas nesse projeto. Nessa etapa já será feita uma análise sobre a correlação entre áudio e padrão binário

Após essa etapa, será feito um levantamento sobre bases de arquivos de áudio gratuitas e abertas que possam ser usadas para testes. Em [15] são apresentadas bases que podem ser usadas.

Escolhida a base, os testes iniciarão nos arquivos de áudio da base escolhida e os resultados serão colhidos e analisados.

5. PLANO DE TRABALHO

Na Tabela 5.1 são apresentadas as tarefas determinadas para o decorrer do projeto. Na Tabela 5.2 é apresentado o cronograma para as tarefas estabelecidas na Tabela 5.1.

Tabela 5.1 Metas estabelecidas para a pesquisa.

METAS	DESCRIÇÃO
1	Levantamento bibliográfico e pesquisa sobre o estado da arte sobre representação de áudio de forma digital.
2	Levantamento bibliográfico e pesquisa sobre o estado da arte sobre funções de similaridade relacionadas com o problema proposto.
3	Definição de base de dados a serem testados.
4	Relatório Parcial entrega até 06/10/20
5	Implementação do ambiente de testes.
6	Execução dos testes.
7	Análise dos resultados.
8	Relatório Final entrega até 28/02/2021

Tabela 5.2 Cronograma proposta para cumprimento das metas.

METAS	MESES								
	JUN	JUL	AGO	SET	OUT	NOV	DEZ	JAN	FEV
1	x	x							
2		x	x						
3			x	x					
4					x				
5						x	x		
6								x	
7									x
8									x

6. VIABILIDADE DE EXECUÇÃO

Para a execução do projeto são necessários os seguintes recursos:

- computador para o aluno realizar as pesquisas e testes;
- acesso à Internet para pesquisa sobre a literatura pertinente e busca por base de dados de áudio.

Dessa forma, não há a necessidade de adquirir equipamentos extras para a execução do projeto.

7. RESULTADOS ESPERADOS E DISSEMINAÇÃO

É esperado que como resultado desse projeto, essa forma de se comparar a similaridade entre áudio e padrão binário de arquivos digitais seja estudada, testada e sistematizada. Isso poderá gerar um método que poderá ser publicado em revistas e congressos científicos sobre o assunto.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] Marius Kaminskas, Francesco Ricci, Contextual music information retrieval and recommendation: State of the art and challenges, Computer Science Review, Volume 6, Issues 2–3, 2012, Pages 89-119, ISSN 1574-0137

[2] U. Shardanand, P. Maes, Social information filtering: algorithms for automating “word of mouth”, CHI’95: Proceedings of the SIGCHI Conference on Human Factors

in Computing Systems, ACM Press/Addison-Wesley Publishing Co., New York, NY, USA (1995), pp. 210-217

[3] P. Lops, M. de Gemmis, G. Semeraro, Content-based recommender systems: state of the art and trends F. Ricci, L. Rokach, B. Shapira, P. Kantor (Eds.), Recommender Systems Handbook, Springer Verlag (2011), pp. 73-105

[4] Joyce, John. (2006). Pandora and the music genome project: Song structure analysis tools facilitate new music discovery. 23. 14+40-41.

[5] Rei Morikawa, 15 Best Audio Datasets for Machine Learning obtido em <https://lionbridge.ai/datasets/12-best-audio-datasets-for-machine-learning/> - acessado em 15/05/2020.