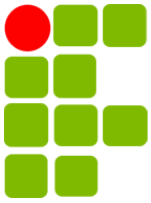


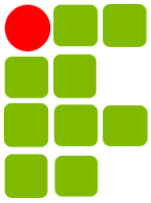
Instituto Federal de Educação, Ciência e Tecnologia da Paraíba
Campus Campina Grande
Coordenação do Curso Superior de Engenharia de Computação

Aplicação do algoritmo K-means para análise de similaridade entre gêneros musicais brasileiros

JOSENILDO SIMÃO DA SILVA
RUBEM RIBEIRO DE BARROS

Orientador: Igor Barbosa da Costa, D.Sc.

Campina Grande, Junho de 2023



Instituto Federal de Educação, Ciência e Tecnologia da Paraíba
Campus Campina Grande
Coordenação do Curso Superior de Engenharia de Computação

Aplicação do algoritmo K-means para análise de similaridade entre gêneros musicais brasileiros

JOSENILDO SIMÃO DA SILVA
RUBEM RIBEIRO DE BARROS

Trabalho de Conclusão de Curso apresentado ao Curso Engenharia de Computação, do Instituto Federal da Paraíba – Campus Campina Grande, em cumprimento às exigências parciais para a obtenção do título de Bacharel em Engenharia de Computação.

Orientador: Igor Barbosa da Costa, D.Sc.

Campina Grande, Junho de 2023

S586a Silva, Josenildo Simão da.
Aplicação do algoritmo K-means para análise de similaridade entre gêneros musicais brasileiros / Josenildo Simão da Silva, Rubem Ribeiro de Barros. - Campina Grande, 2023.
25 f.: il.

Trabalho de Conclusão de Curso (Graduação em Engenharia de computação) - Instituto Federal da Paraíba, 2023.
Orientador: Prof. Dr. Igor Barbosa da Costa.

1. Algoritmo K-means 2. Gêneros musicais 3. Coeficiente máximo de agrupamento. I.Barros, Rubem Ribeiro de II.Costa, Igor Barbosa da. III Título.

CDU 004.021

Aplicação do algoritmo K-means para análise de similaridade entre gêneros musicais brasileiros

Trabalho de Conclusão de Curso apresentado ao Curso Engenharia de Computação, do Instituto Federal da Paraíba – Campus Campina Grande, em cumprimento às exigências parciais para a obtenção do título de Bacharel em Engenharia de Computação.

Aprovada em 31/05/2023

Igor Barbosa da Costa, D.Sc.
Orientador

Henrique do Nascimento Cunha, Sc. M
Membro da Banca

Paulo Ribeiro Lins Júnior, D.Sc.
Membro da Banca

Campina Grande, Paraíba, Brasil
Junho/2023

"Nada resiste ao trabalho."
Autor desconhecido

À Deus. À meus pais, familiares e amigos, por todo apoio e carinho!

Agradecimentos

Por Josenildo: Gostaria de expressar meus sinceros agradecimentos a Deus por ter me protegido das adversidades que enfrentei durante a graduação. Aos meus queridos pais, Ivone e Josinaldo, que, mesmo não tendo tido a oportunidade de serem alfabetizados, sempre compreenderam que o melhor caminho é buscar conhecimento. Sou imensamente grato por todo amor e suporte que recebi. À minha amada Vó Francisca, que sempre me aconselhou e apoiou incondicionalmente. Aos meus tios João, Fabiano, Ivonete, Adriana e Adeilton, que sempre estiveram ao meu lado, oferecendo seu apoio.

Gostaria de estender meus agradecimentos especiais aos meus padrinhos Beto e Edjane, que me acolheram como parte de sua família. Vocês se tornaram minha segunda família. Aos meus amigos Matheus, Emerson e Wanderson, que considero como irmãos, gostaria de expressar minha profunda gratidão. Vocês estiveram ao meu lado durante os momentos em que estive longe da minha família, e sua amizade foi inestimável. À Yasmin, que se tornou mais do que uma amiga, uma verdadeira companheira ao longo da minha jornada acadêmica.

Desejo agradecer a todos os meus professores, que me guiaram com seu vasto conhecimento. Em especial, gostaria de mencionar e agradecer aos professores Henrique Cunha, Ana Cristina e Paulo Ribeiro. Por fim, meu orientador Igor Barbosa, que me apoiou de forma incansável e me ajudou a tornar possível essa etapa final rumo à minha formação.

A todos mencionados, minha gratidão eterna.

Por Rubem: Aos meus pais, irmão, família e aos grandes mentores desta Odisseia educacional, minha gratidão sincera. Esta jornada repleta de desafios me proporcionou aprendizados inestimáveis. Seguindo as palavras de Newton, reconheço que alcancei horizontes mais amplos ao estar sobre os ombros de gigantes. Sou profundamente grato por cada um de vocês.

Sumário

Lista de Abreviaturas

API	*Application Programming Interface*
t-SNE	*t-Distributed Stochastic Neighbor Embedding*
CMA	*Coeficiente Máximo de Agrupamento*

Lista de Figuras

Lista de Tabelas

Aplicação do algoritmo *K-means* para análise de similaridade entre gêneros musicais brasileiros

Silva S. Josendildo[1], **Barros R. Rubem**[1], **Barbosa C. Igor**[1]

[1]Instituto Federal de Ciência e Tecnologia– IFPB - Campus Campina Grande (IFPB)
CEP 58432-300– Campina Grande - PB– Brazil

{silva.josenildo, rubem.barros,igor.costa}@academico.ifpb.edu.br

Resumo. *A música desempenha um papel essencial na sociedade humana e, no Brasil, reflete a diversidade cultural do país ao longo de sua rica trajetória musical. A interação e a influência mútua entre os artistas desempenham um papel importante na evolução dos gêneros musicais brasileiros. Compreender a interrelação e a influência entre esses gêneros é fundamental para decifrar a identidade musical do Brasil e orientar o desenvolvimento da indústria musical. Este estudo recorre a dados obtidos através da API do Spotify com o objetivo de analisar as conexões e similaridades entre os gêneros musicais brasileiros, dando especial ênfase às características sonoras. A análise, que envolveu dados de 1896 artistas brasileiros, utilizou técnicas de aprendizado não supervisionado, especificamente o algoritmo K-means, para identificar padrões de similaridade. Os experimentos revelaram notável homogeneidade entre os artistas de música gospel e uma correlação expressiva entre os artistas de forró e pagode. Adicionalmente, este trabalho oferece um conjunto de dados estruturado que poderá ser usado em futuras pesquisas para um entendimento mais aprofundado dos gêneros musicais brasileiros e suas interligações.*

Abstract. *Music plays an essential role in human society and, in Brazil, it reflects the country's cultural diversity throughout its rich musical journey. The interaction and mutual influence among artists play a significant role in the evolution of Brazilian musical genres. Understanding the interrelation and influence between these genres is crucial to deciphering Brazil's musical identity and guiding the development of the music industry. This study utilizes data obtained from Spotify's API to analyze the connections and similarities between Brazilian musical genres, with a special emphasis on sound characteristics. The analysis, which involved data from 1896 Brazilian artists, employed unsupervised learning techniques, specifically the K-means algorithm, to identify similarity patterns. The experiments revealed remarkable homogeneity among gospel music artists and a striking correlation between forró and pagode artists. Additionally, this work provides a structured dataset that could be used in future research for a deeper understanding of Brazilian musical genres and their interconnections.*

1. Introdução

A música desempenha um papel fundamental na sociedade humana desde tempos remotos e representa um elemento essencial do patrimônio cultural. No caso do Brasil, a música possui uma trajetória rica e diversificada, refletindo a enorme riqueza cultural do país.

Desde os primeiros ritmos indígenas até as influências africanas, europeias e americanas, a música brasileira se desenvolveu de forma singular, resultando em uma ampla variedade de gêneros musicais ao longo dos séculos [McCann 2004].

A interação entre os artistas e a influência mútua que exercem uns sobre os outros desempenham um papel importante na evolução da música brasileira. A partir de eventos sociais, políticos, culturais e experiências pessoais, os artistas brasileiros têm moldado e transformado os gêneros musicais ao longo do tempo. Essa dinâmica resultou em um rico patrimônio musical, que abrange desde o samba e a bossa nova até o funk, o axé e o sertanejo, entre muitos outros gêneros [Perrone and Dunn 2002].

Compreender a interrelação e a influência mútua entre os gêneros musicais é essencial para decifrar a identidade musical brasileira. Além disso, a análise da evolução dos gêneros musicais brasileiros pode orientar o desenvolvimento da indústria musical, fornecendo aos artistas e produtores uma melhor compreensão das tendências e influências que moldam o panorama musical. Com isso, podem ajudar os produtores à auxiliar artistas para produzir músicas que terão uma maior probabilidade de tornar-se um sucesso entre os ouvintes. Já que poderão comparar a música produzida com os sucesso atuais.

Nesse contexto, este estudo tem como objetivo analisar a cadeia de influências entre os gêneros musicais brasileiros, investigando suas semelhanças e diferenças, especialmente em termos de sonoridade. Para alcançar esses objetivos, serão abordadas duas questões principais:

1. Qual o grau de similaridade entre os artistas de um determinado gênero?
2. Quais gêneros apresentam maior similaridade entre si?

Para responder a essas perguntas, foi compilado um conjunto de dados utilizando a API do Spotify [Spotify 2023]. Esse conjunto de dados inclui informações processadas de 1896 artistas brasileiros, selecionados com base em critérios que contemplam a diversidade de gêneros musicais. A análise dos dados foi conduzida por meio de técnicas de aprendizagem não supervisionada, como o algoritmo *K-means*, a fim de identificar padrões de similaridade entre os gêneros musicais.

A avaliação desses dados proporcionou percepções valiosas sobre a interconexão dos gêneros musicais brasileiros, como a notável semelhança entre os artistas do gospel e a forte correlação entre artistas do forró e do pagode. Além disso, este estudo também contribui disponibilizando um conjunto de dados estruturado com informações de artistas brasileiros, que pode servir como base para futuras pesquisas nessa área.

2. Fundamentação Teórica

Esta seção tem como objetivo apresentar os conceitos importantes para uma compreensão mais aprofundada deste trabalho.

2.0.1. Características de Áudio

O Spotify é conhecido por seu algoritmo de recomendação que personaliza as listas de reprodução com base nas preferências musicais de cada usuário. Para fornecer essa personalização, a plataforma coleta dados de músicas e playlists, utilizando informações

sobre as características de áudio de cada faixa para criar um modelo de preferência do usuário. Essas características de áudio estão disponíveis para os desenvolvedores por meio da API do Spotify [Spotify 2023].

As *características de áudio* da API do Spotify se referem a informações sobre os aspectos sonoros de uma faixa musical, como dançabilidade, energia, volume, valência, entre outras. Essas características permitem que os desenvolvedores criem aplicativos e serviços que ofereçam recursos personalizados para seus usuários, como recomendações musicais com base nas preferências de áudio ou a criação de listas de reprodução adaptadas a diferentes atividades e momentos do dia.

Aqui estão algumas das principais características de áudio fornecidas pela API do Spotify:

- **Dançabilidade (Danceability):** indica o quão adequada uma faixa é para dançar, levando em consideração fatores como andamento, estabilidade rítmica, força da batida e regularidade. O valor varia de 0,0 (menos dançável) a 1,0 (mais dançável).
- **Energia (Energy):** mede a intensidade e atividade percebidas em uma faixa, com valores variando de 0,0 a 1,0. Faixas com alta energia geralmente têm uma sonoridade rápida, alta e barulhenta, enquanto faixas com baixa energia podem ser mais suaves e calmas.
- **Tom (Key):** indica a tonalidade da faixa, representada por um número inteiro mapeado para notas musicais (por exemplo, 0 = C, 1 = C#/Db, 2 = D, e assim por diante). Um valor de -1 indica que a tonalidade não foi detectada.
- **Volume (Loudness):** representa o volume geral da faixa em decibéis (dB). Os valores de volume são calculados em média em toda a faixa e podem ser úteis para comparar o volume relativo entre diferentes faixas.
- **Modo (Mode):** indica a modalidade (maior ou menor) da faixa. O valor 1 representa o modo maior, enquanto o valor 0 representa o modo menor.
- **Fala (Speechiness):** detecta a presença de palavras faladas em uma faixa. Valores próximos a 1,0 indicam que a faixa é predominantemente composta por palavras faladas, como programas de rádio ou audiolivros, enquanto valores abaixo de 0,33 geralmente representam faixas instrumentais.
- **Acústica (Acousticness):** medida que indica a probabilidade de a faixa ser acústica, variando de 0,0 (menor probabilidade) a 1,0 (maior probabilidade).
- **Instrumentalidade (Instrumentalness):** prevê se uma faixa não contém vocais. Valores próximos de 1,0 indicam uma alta probabilidade de a faixa ser instrumental.
- **Vivacidade (Liveness):** detecta a presença de uma plateia na gravação da faixa. Valores mais altos indicam uma maior probabilidade de a faixa ter sido gravada ao vivo.
- **Valência (Valence):** mede a positividade musical transmitida por uma faixa, variando de 0,0 (mais negativa) a 1,0 (mais positiva). Faixas com alta valência têm uma sonoridade mais alegre e animada, enquanto faixas com baixa valência podem transmitir tristeza ou melancolia.
- **Tempo (Tempo):** indica o andamento estimado da faixa em batidas por minuto (BPM). É uma medida da velocidade ou ritmo da música.
- **Duração (Duration-ms):** representa a duração da faixa em milissegundos.

- **Assinatura temporal (Time Signature):** fornece uma estimativa da fórmula de compasso da faixa. A fórmula de compasso especifica quantas batidas existem em cada compasso, variando de 3 a 7.

Essas características de áudio fornecem informações detalhadas sobre o conteúdo sonoro de uma faixa, permitindo que os desenvolvedores criem aplicativos e serviços que ofereçam recursos personalizados para os usuários do Spotify [Spotify 2023].

2.1. t-SNE

O *t-SNE* (*t-Distributed Stochastic Neighbor Embedding*) é um algoritmo de redução de dimensionalidade não linear amplamente utilizado para visualizar dados de alta dimensionalidade em um espaço de menor dimensão [Van der Maaten and Hinton 2008].

O *t-SNE* é especialmente útil para a visualização de dados em aprendizado de máquina e análise não supervisionada. Ele permite explorar padrões e agrupamentos em conjuntos de dados de alta dimensão, como análise de agrupamento de músicas ou reconhecimento de padrões de fala.

O algoritmo funciona mapeando cada ponto de dados em um espaço bidimensional ou tridimensional, ao mesmo tempo em que tenta preservar as distâncias entre os pontos no espaço de alta dimensão. Isso significa que pontos semelhantes no espaço de alta dimensão serão mapeados próximos uns dos outros no espaço de baixa dimensão, enquanto pontos diferentes no espaço de alta dimensão serão mapeados distantes uns dos outros no espaço de baixa dimensão.

O *t-SNE* é implementado em várias bibliotecas de aprendizado de máquina, como o *scikit-learn* em Python. Essa técnica é frequentemente utilizada em pesquisas científicas e análises de dados de empresas, como o Spotify, que utiliza o *t-SNE* para agrupar músicas em seus recursos de recomendação.

Ao aplicar o *t-SNE* a conjuntos de dados de música, é possível visualizar padrões, similaridades e relações entre as faixas em um espaço de baixa dimensão. Isso pode ser usado para entender melhor as preferências dos usuários, identificar agrupamentos musicais ou até mesmo criar visualizações interativas para explorar o catálogo musical.

2.2. K-means

O algoritmo *K-means* é amplamente utilizado em várias áreas como uma técnica de aprendizado não supervisionado, sendo aplicado especialmente na análise de dados de mercado e na segmentação de clientes [MacQueen 1967]. O *K-means* é conhecido por sua simplicidade e eficiência na tarefa de agrupar dados em K *clusters*, onde K é um valor predefinido. Sua abordagem segue uma sequência de passos básicos:

1. **Inicialização**: São selecionados K centróides iniciais aleatoriamente ou de forma estratégica.
2. **Atribuição**: Cada ponto de dados é atribuído ao centróide mais próximo, formando *clusters* preliminares.
3. **Atualização**: Os centróides são atualizados recalculando-se suas posições com base nos pontos de dados atribuídos a eles.
4. **Repetição**: Os passos 2 e 3 são repetidos até que os centróides não se movam significativamente ou um critério de parada seja alcançado.

Uma das limitações do *K-means* é que ele assume que os *clusters* têm formas esféricas e possuem a mesma variância, o que pode não ser verdadeiro em alguns conjuntos de dados. Além disso, o resultado do *K-means* pode depender da inicialização dos centróides, podendo levar a diferentes soluções.

No entanto, o *K-means* continua sendo uma técnica popular e eficiente para o agrupamento de dados. É amplamente utilizado devido à sua simplicidade, escalabilidade e interpretabilidade. O *K-means* está implementado em várias bibliotecas de aprendizado de máquina, como o *scikit-learn* em Python, o que facilita sua aplicação e uso em projetos.

2.3. Correlação de Pearson

A correlação de Pearson é uma medida de associação entre duas variáveis quantitativas. É amplamente utilizada para avaliar o grau de relação linear entre as variáveis e pode ser calculada a partir da covariância e do desvio padrão de cada variável. A correlação de Pearson varia entre -1 e 1, onde valores próximos a -1 indicam uma correlação negativa perfeita, valores próximos a 1 indicam uma correlação positiva perfeita, e valores próximos a 0 indicam ausência de correlação [Myers and Myers 2010].

A correlação de Pearson é comumente empregada em análises estatísticas, tais como análise de dados de mercado e pesquisas científicas. No entanto, é importante ressaltar que ela mede apenas a relação linear entre as variáveis, não sendo apropriada para avaliar relações não lineares.

O cálculo da correlação de Pearson pode ser realizado facilmente utilizando diversas ferramentas de análise de dados, como o Excel, e linguagens de programação como Python e R. Além disso, é possível interpretar a correlação de Pearson por meio de um coeficiente que indica a intensidade e a direção da relação entre as variáveis, permitindo insights sobre o comportamento dessas variáveis em conjunto.

2.4. Coeficiente Máximo de Agrupamento (CMA)

O Coeficiente Máximo de Agrupamento (CMA) é uma medida utilizada na análise de redes complexas para avaliar a densidade de agrupamentos de nós altamente conectados em uma rede. Ele é definido como o maior coeficiente de agrupamento em uma rede, ou seja, o coeficiente de agrupamento máximo que pode ser alcançado em qualquer nó da rede.

O CMA é uma medida útil para avaliar a eficácia de algoritmos de agrupamento em redes complexas, como em análises de redes sociais ou biológicas. Uma rede com um alto valor de CMA indica a presença de vários grupos de nós altamente conectados na rede, o que pode ser útil para identificar comunidades ou subgrupos de interesse.

O cálculo do CMA pode ser realizado utilizando algoritmos de agrupamento, como o algoritmo de Louvain, que foi utilizado em estudos que aplicaram o CMA para análise de redes complexas [Opsahl and Panzarasa 2009].

2.5. Pipeline ETL

O *Pipeline ETL* (*Extract, Transform, Load*) é uma estrutura de processamento de dados amplamente utilizada em projetos de análise de dados e inteligência de negócios. Ele consiste em três etapas principais: extração, transformação e carga de dados [Kimball and Ross 2011].

Na etapa de extração, os dados são obtidos de diversas fontes, como bancos de dados, arquivos, APIs e sistemas externos. A extração pode envolver consultas a bancos de dados, coleta de dados em tempo real ou importação de arquivos estruturados ou não estruturados

Em seguida, na etapa de transformação, os dados são limpos, organizados e formatados para que possam ser integrados em um único conjunto de dados e analisados de forma eficiente. Essa etapa pode envolver a limpeza de dados inconsistentes ou duplicados, a padronização de formatos, a agregação de dados e a aplicação de regras de negócio.

Por fim, na etapa de carga, os dados transformados são inseridos em um repositório de dados ou em um sistema de análise para serem utilizados posteriormente. Essa etapa pode envolver a criação de estruturas de armazenamento, como *data warehouses* ou *data lakes*, e a carga dos dados transformados nesses ambientes.

O *Pipeline ETL* é uma parte fundamental do processo de análise de dados, pois permite que os dados sejam processados de forma sistemática e automatizada, aumentando a eficiência e a precisão da análise. Ele é amplamente utilizado em projetos de *big data* e inteligência de negócios, nos quais grandes quantidades de dados precisam ser coletadas, organizadas e analisadas.

3. Revisão de Literatura

Esta seção tem como objetivo apresentar trabalhos relacionados ao tema deste artigo, buscando expor as diferentes abordagens utilizadas para detecção de similaridade entre gêneros musicais e entre musicas.

Na leitura já existe trabalho relacionado a similaridade entre gêneros musicais. [Gabriela 2014] investigou as subestruturas dentro da música popular (rock, hip hop, pop, etc.). Durante a pesquisa criou subgêneros musicais utilizando aprendizagem de máquina não-supervisionado em características de áudio que capturam timbre, ritmo e tempo e observam as relações entre eles. Utilizando o algoritmo k-means, observou que as músicas do mesmo cluster são mais semelhantes do que as músicas de outros clusters. Além disso, observou que certos sub-gêneros gerados pelos k-means são extremamente diferentes de outros sub-gêneros também gerados pelo k-means em termos de . Por exemplo, o “Fast-hardalternative-punk rock” é muito diferente do “Melancholic rock pop”, mas muito semelhante do “Energy pop hard rock”.

Encontrar conjuntos de dados para pesquisa científica relacionado a música, artista ou gêneros brasileiros é um desafio. Pensando nesse problema [Paulo and Denise 2021] desenvolveram uma base de dados destinada à classificação automática de gêneros musicais brasileiros. Para tanto, utilizou o Spotify para identificar músicas relacionadas aos gêneros Axé, Bossa Nova, Brega, Choro, Forró, Frevo, Funk Carioca, Maracatu, Música Sertaneja, Pagode e Samba. A base de dados ficou com um total de 1907 registros com 684 artistas únicos. O conjunto de dados contém nome do artista, gênero e características das áudio como Acousticness, Danceability, Duration_ms, Instrumentalness, Key, Liveness, Loudness, Mode, Speechiness, Tempo e Valence.

Para encontrar a similaridade musical, além da abordagem baseada em conteúdo da música como características sonoras, há também a abordagem baseada em contexto.

Essa abordagem contextual refere-se a todas as informações relevantes para a música que não estão incluídas no próprio sinal de áudio. [Peter and Markus 2013] desenvolveram uma pesquisa sobre os tipos de similaridade musical baseada em dados de contexto. Os autores identificaram três principais tipos de abordagens de similaridade baseadas em contexto: abordagens baseadas em recuperação de texto (baseadas em textos da web, tags ou letras), abordagens baseadas em coocorrência (baseadas em listas de reprodução, contagens de páginas, microblogs ou redes ponto a ponto) e abordagens baseadas em avaliações do usuário ou hábitos de escuta.

Abordagens baseadas em recuperação de texto

Para abordagens baseadas em texto é utilizado mecânicos de pesquisa na web, filtrando a consulta de termos que pode incluir nome do artista seguidas de palavras chaves relacionadas ao tema, como crítica musical ou estilo de gênero musical. Para isso, são utilizadas diversas técnicas como recuperação de informação (IR), processamento de linguagem natural (NLP), Análise Semântica Latente (LSA) e marcação de parte da fala (PoS).

Abordagens baseadas em coocorrência

Dentro das abordagens baseadas em ocorrência, a similaridade é estimada a partir da ocorrência de peças musicais ou artistas dentro do mesmo contexto. Repetições ou contagens de palavras referindo-se a dois artistas dentro de páginas web, microblogs (Twitter), listas de reprodução ou redes ponto-a-ponto (P2P) que pode indicar algum tipo de semelhança sobre os artistas.

Abordagens baseadas em avaliações do usuário

A abordagem baseada em feedback do usuário ou abordagem colaborativa, explora dois tipos de relações de similaridade por rastreamento dos hábitos dos usuários: similaridade item a item e similaridade usuário para usuário.

Para similaridade item a item são computadas matrizes de similaridade entre todos os usuários e itens/artista musical. As colunas da matriz representam os usuários e as linhas apresentam o item ou artista musical. Para encontrar a similaridade nesse caso, será comparado os vetores de linha da matrix. Portanto, considera-se a similaridade, itens/artista musical que obtiveram avaliações similares de grupos de usuários.

Já quando falamos de similaridade de usuário para usuário estamos nos referindo a uma matriz de similaridade entre todos os usuários. Utilizando como exemplo a matriz do exemplo anterior, para obter a similaridade é comparado todos os vetores de colunas dessa matriz. Então, considera-se a similaridade, usuários que deram avaliações semelhantes a um conjunto comum de itens.

A utilização de técnica de Deep Learning pode ser alternativa para obter a similaridade musical para classificação de estilos musicais. Nesse contexto, [Xiuli 2023] desenvolveu uma pesquisa para detectar similaridade musical utilizando convolutional neural network (CNN). Em sua pesquisa, utilizou o algoritmo Harmony and Percussive Source Separation (HPSS) para separar os espectrograma do sinal musical original e decompor em dois componentes: harmônicos característicos de tempo e frequência choques característicos. Esses dois elementos foram inseridos na CNN junto com os dados no espectrograma original para processamento. O autores concluíram que esse método é

superior em relação aos métodos clássicos de detecção de similaridade musical.

4. Conjunto de Dados

Esta seção tem como objetivo apresentar a metodologia utilizada para a construção do conjunto de dados utilizado neste trabalho, fornecendo informações relevantes sobre o conteúdo, qualidade, tamanho e fonte dos dados. Serão também discutidas eventuais limitações e pré-processamentos realizados, com o objetivo de fornecer uma visão completa e transparente sobre a origem e a preparação dos dados utilizados na pesquisa. Dessa forma, busca-se garantir a qualidade e a confiabilidade dos dados utilizados nas análises realizadas neste trabalho.

O desenvolvimento do conjunto de dados seguiu um pipeline de extração, transformação e carregamento dos dados a partir de consultas à API do Spotify. A sequência de passos foi a seguinte:

1. Extração da lista de todos os gêneros musicais presentes na plataforma.
2. Identificação manual dos gêneros tipicamente brasileiros, como Gospel, Forró, Sertanejo, Bossa Nova, Samba, MPB, Pagode, Rap, Funk e Rock Brasileiro.
3. Para cada gênero, foram buscados dados dos artistas pertencentes ao gênero, incluindo nome do artista, popularidade, gêneros relacionados, ID do artista e URI.
4. Para cada artista, foram extraídos os álbuns lançados.
5. Para cada álbum, foram extraídas as faixas musicais correspondentes.
6. Para cada faixa musical, foram extraídas as características de áudio (como danceability, energy, key, loudness, mode, speechiness, acousticness, instrumentalness, liveness, valence, tempo, duration_ms e time_signature).
7. Para cada artista, foi calculada uma média para cada característica individual de áudio utilizando as faixas musicais extraída da API.

Após a etapa de extração dos dados, foram realizados alguns pré-processamentos para garantir a consistência e a integridade dos dados. Isso incluiu a verificação e a remoção de registros duplicados, a correção de eventuais erros ou inconsistências nos metadados dos artistas, álbuns e faixas, além da padronização e normalização das características de áudio.

No geral, o conjunto de dados resultante é composto por informações sobre artistas brasileiros de diversos gêneros musicais, seus álbuns e faixas correspondentes, bem como as médias calculadas para cada característica de áudio. Esses dados fornecem uma base sólida para a realização de análises e investigações relacionadas à música brasileira e às suas características artísticas.

Na etapa final do pipeline, os dados tratados anteriormente foram consolidados em um único conjunto de dados. Cada registro do conjunto de dados contém as seguintes informações:

- Nome do artista ou banda musical.
- Médias individuais de cada característica de áudio, como danceability, energy, key, loudness, mode, speechiness, acousticness, instrumentalness, liveness, valence, tempo, duration_ms e time_signature.
- Número de faixas utilizadas para calcular as médias.
- Gênero musical utilizado como critério de pesquisa para encontrar o artista.

O conjunto final de dados tem informações de 1896 artistas. A Figura 1 ilustra a distribuição de artistas por gênero musical no dataset.

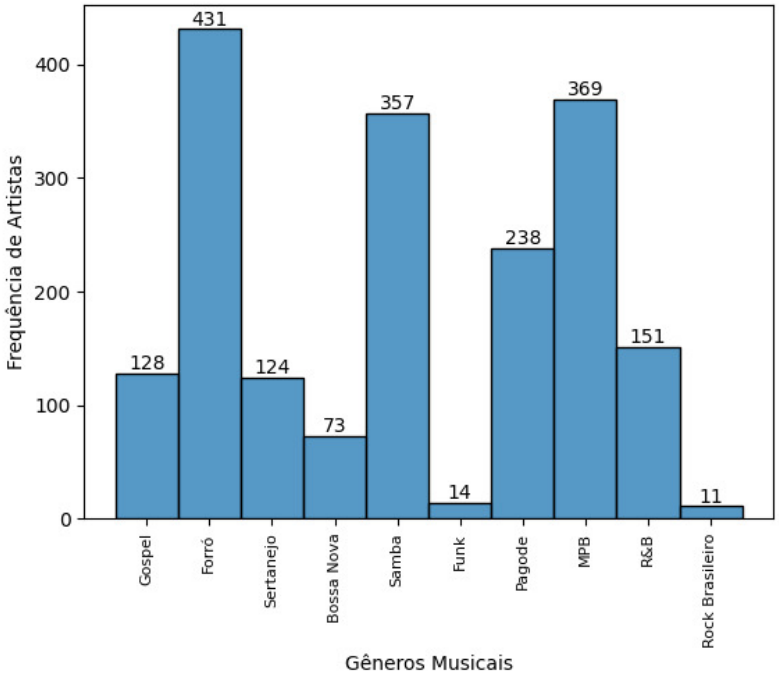


Figura 1. Distribuição de artistas por gênero musical.

5. Resultados

Nesta seção, serão apresentados os resultados obtidos usando técnicas de aprendizagem de máquina e estatística para responder às perguntas de pesquisa.

5.1. Qual o grau de similaridade entre os artistas de um determinado gênero?

Para avaliar a similaridade entre os artistas de um determinado gênero, foi utilizado o algoritmo K-Means para realizar o agrupamento [Dabbura 2018]. Esse algoritmo emprega o método *Expectation-Maximization* para atribuir os artistas a grupos com base em suas características musicais.

Antes de realizar o agrupamento, definiu-se previamente o número de grupos como sendo 10 (dez), k = 10, correspondendo à quantidade de gêneros presentes no conjunto de dados. O parâmentro init foi configurado como sendo 50, sendo o melhor a convergencia. Por fim, foi utilizando o max_iter = 1000. Observou que o resultado para essa configuração estava estável. Além disso, aplicou-se a técnica de redução de dimensionalidade t-SNE para transformar as características de 14 (quatorze) dimensões em 2 (duas) dimensões. Isso permitiu visualizar os grupos em um gráfico bidimensional. O melhor algoritmo de redução de dimensionalidade que possibilitou uma vizualização mais evidentes dos grupos formados pelo k-means com o conjunto de dados desenvolvido, foi o t-SNE.

A Figura 2 ilustra essa representação, na qual cada ponto de dados representa um artista e as cores indicam os diferentes grupos (clusters) gerados pelo algoritmo K-Means.

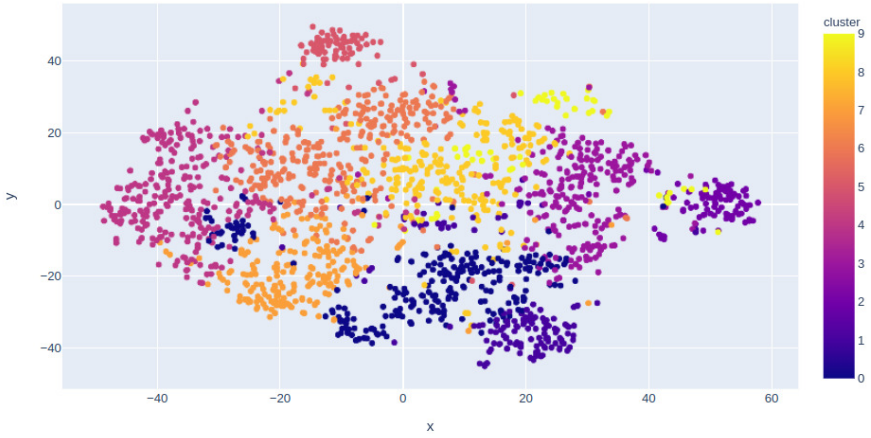


Figura 2. Grupos de artistas de acordo com o algoritmo K-Means

Para analisar se o algoritmo de agrupamento foi capaz de separar com sucesso diferentes gêneros com base em suas características, utilizou-se o Coeficiente Máximo de Agrupamento (CMA). O CMA é calculado como a proporção entre o número máximo de artistas de um gênero em um grupo (cluster) e o número total de artistas desse gênero.

O raciocínio por trás disso é o seguinte: em um gênero no qual os artistas são mais similares entre si, haverá uma tendência de que mais artistas estejam agrupados no mesmo grupo. Por outro lado, em um gênero no qual os artistas são mais distintos, haverá uma distribuição maior entre diferentes grupos. Nesse sentido, o CMA pode ser utilizado como uma métrica para determinar o grau de similaridade entre os artistas de um determinado gênero. Quanto mais próximo de 1 for o valor do CMA, maior será a similaridade entre os artistas desse gênero. A Tabela 1 apresenta os resultados obtidos, ordenados pelo CMA.

Tabela 1. Coeficiente Máximo de Agrupamento (CMA) para os gêneros musicais

Gênero	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	CMA
Gospel	11.0	93.0	8.0	4.0	0.0	2.0	0.0	1.0	9.0	0.0	0.73
Sertanejo	47.0	2.0	0.0	6.0	2.0	0.0	0.0	66.0	1.0	0.0	0.53
Funk	2.0	0.0	0.0	0.0	1.0	1.0	0.0	6.0	3.0	1.0	0.43
Forró	56.0	4.0	5.0	8.0	179.0	3.0	111.0	39.0	19.0	7.0	0.42
MPB	37.0	20.0	41.0	144.0	2.0	3.0	9.0	4.0	104.0	5.0	0.39
Rock Brasileiro	4.0	2.0	0.0	0.0	1.0	0.0	0.0	1.0	3.0	0.0	0.36
R&B	12.0	14.0	6.0	7.0	8.0	54.0	3.0	2.0	30.0	15.0	0.36
Samba	56.0	5.0	10.0	38.0	11.0	15.0	119.0	28.0	45.0	30.0	0.33
Pagode	18.0	3.0	1.0	6.0	69.0	6.0	59.0	58.0	14.0	4.0	0.29
Bossa Nova	14.0	1.0	18.0	21.0	0.0	0.0	4.0	1.0	8.0	6.0	0.29

Ao analisar os resultados, é evidente que os artistas do gênero gospel obtiveram o maior valor de CMA (aproximadamente 0,73), indicando que cerca de 73% desses artistas foram agrupados no mesmo grupo (cluster 2). Isso revela que os artistas desse gênero possuem uma alta similaridade entre si. Por outro lado, os gêneros de pagode e bossa nova apresentaram os menores valores de CMA, sugerindo que seus artistas são menos similares entre si e estão distribuídos em diferentes grupos. Isso implica que os

artistas desses gêneros possuem características musicais distintas e um menor grau de similaridade em relação aos demais artistas do mesmo gênero.

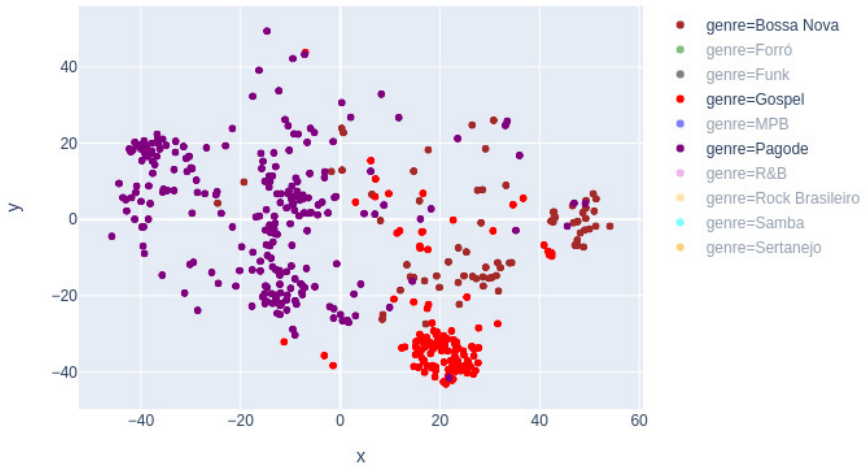


Figura 3. Resultado do agrupamento entre os gêneros musicais bossa nova, gospel e pagode

Para avaliar se esses resultados são estatisticamente significativos, foi realizado um teste A/B. A Figura 4 ilustra o resultado do teste A/B sobre o CMA. O CMA real (em azul) é plotado em comparação com o CMA gerado randomicamente pelo teste A/B (em vermelho). As linhas de intervalo em preto representam o intervalo de confiança de 95%.

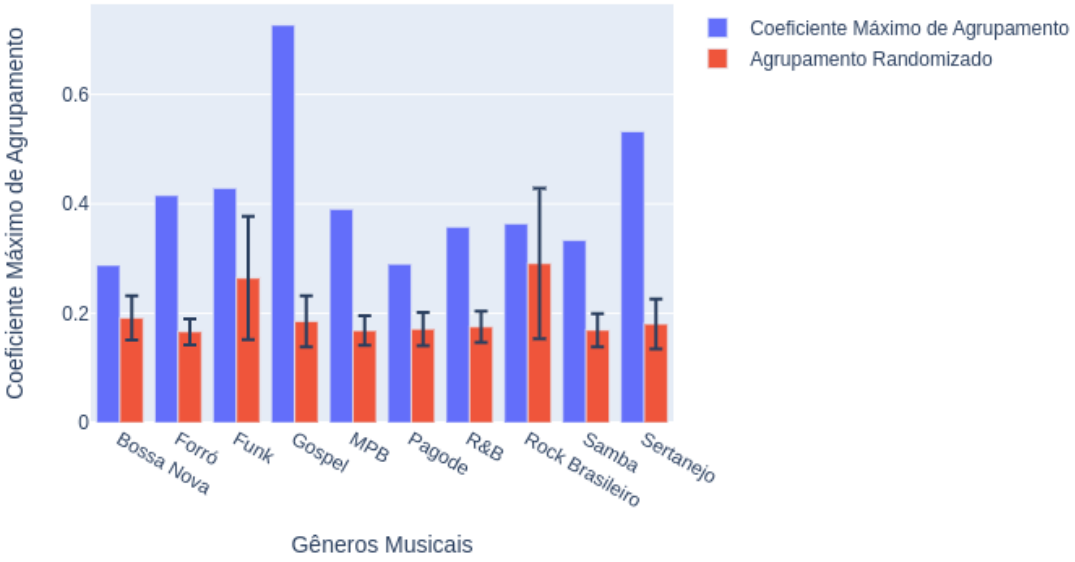


Figura 4. Resultado do teste A/B sobre o Coeficiente Máximo de Agrupamento (CMA)

Observa-se que os valores reais de CMA estão fora da margem do intervalo de confiança, indicando que os resultados da clusterização são estatisticamente significativos (valor $p < 0.05$), com exceção do gênero "rock brasileiro". Isso sugere que o algoritmo K-Means foi capaz de separar de forma significativa os diferentes gêneros com base em suas características musicais, proporcionando insights reais sobre o grau de similaridade entre os artistas de cada gênero.

5.2. Quais gêneros apresentam maior similaridade entre si?

Para entender a similaridade entre os gêneros, foi utilizada uma medida de correlação chamada correlação de Pearson. A Figura 5 ilustra a matriz de confusão com a correlação de Pearson entre cada par de gêneros.

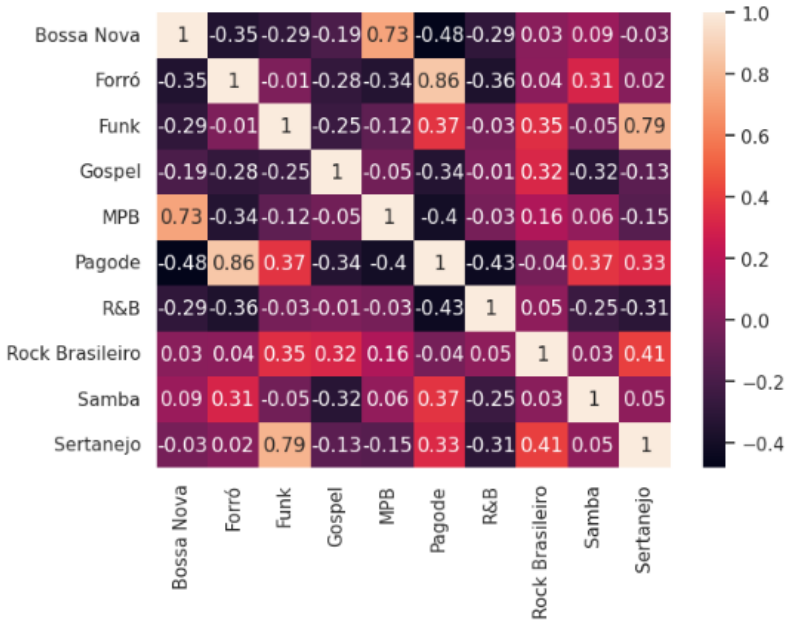


Figura 5. Matriz de confusão com a correlação de Pearson para cada gênero.

Pode-se observar que diversos pares de gêneros apresentam correlação significativa. Por exemplo, os gêneros Forró e Pagode possuem uma correlação de $0,86$, indicando uma forte associação entre esses dois gêneros. Além disso, os gêneros Sertaenjo e Bossa Nova, como também Funk e Sertanejo apresentam uma correlação positiva, sugerindo uma similaridade entre esses dois gêneros musicais. Por outro lado, gêneros como Pagode e Bossa Nova apresentam uma correlação negativa de $-0,48$, indicando uma dissimilaridade entre esses gêneros. A Figura 6 ilustra os três pares de gêneros com a maior correlação de acordo com a medida de correlação de Pearson.

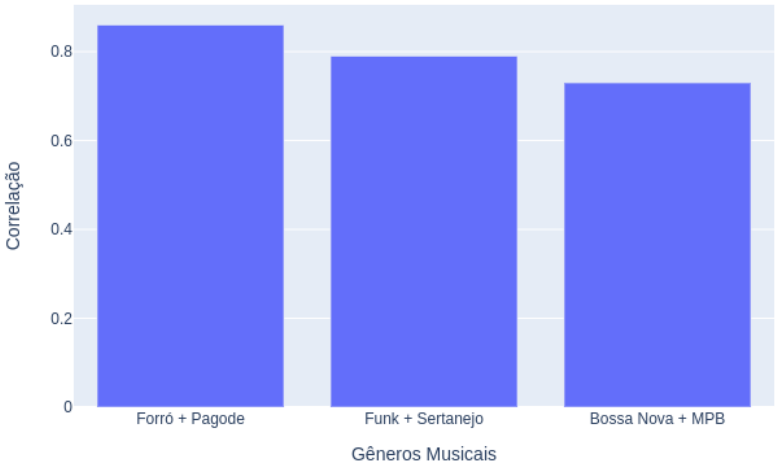


Figura 6. Pares de gêneros com maior correlação de acordo com a medida de correlação de Pearson.

Esses resultados indicam que existem associações e similaridades entre os gêneros musicais estudados. Alguns gêneros apresentam uma correlação positiva, indicando uma maior similidade entre reles, enquanto outros apresentam uma correlação negativa sugerindo diferenças musicais mais distintas.

A figura 7 ilustra a correlação entre o gênero Forró e o gênero Pagode. Cada ponto correspode a um artista. Podemos perceber que os artitas de ambos os gêneros, em sua maioria, estão posicionados próximo de si. Confirmando que os gêneros tem forte correlação entre si.

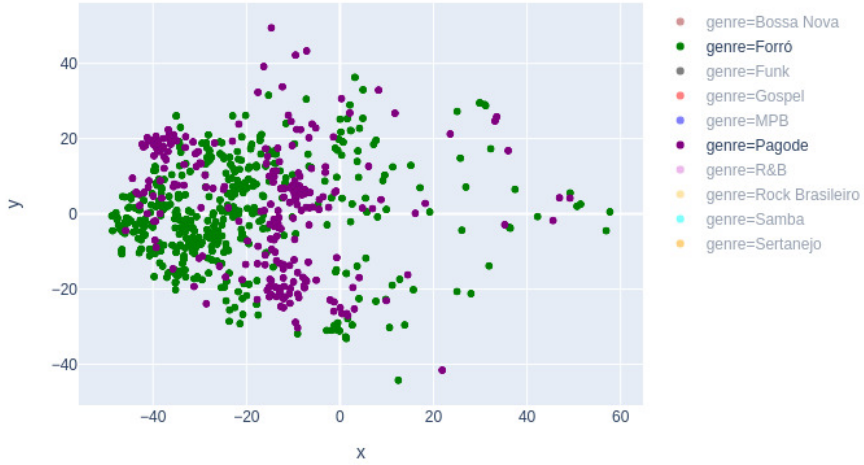


Figura 7. Cluster do par de gênero com maior correlação de acordo com a medida de correlação de Pearson.

6. Conclusão

Com base nas informações analisadas, foi possível identificar que artistas pertencentes a gêneros musicais distintos podem apresentar similaridade entre si. Observou-se também que existem gêneros musicais dentro do escopo deste estudo em que os artistas possuem maior similaridade. A utilização das características dos artistas e dos recursos de áudio permitiu a identificação de agrupamentos, tanto entre os artistas quanto entre os gêneros musicais analisados. Esses resultados indicam que este estudo pode servir como base para implementações relacionadas a sistemas de recomendação musical, visando recomendar produtos a grupos de pessoas com preferência por artistas ou gêneros musicais similares.

Além disso, destaca-se a relevância deste trabalho para a pesquisa brasileira, uma vez que foi construído um dataset contendo artistas e gêneros musicais tipicamente brasileiros. Essa base de dados poderá ser utilizada como referência para futuras investigações nessa área.

6.1. Limitações

É fundamental reconhecer as limitações inerentes a este estudo, a fim de garantir uma interpretação completa e realista dos resultados obtidos. A seguir, são apresentadas as limitações identificadas:

1. **Limitações dos dados:** As análises foram baseadas em um conjunto de dados obtido por meio da API do Spotify, o que implica que estão sujeitos a possíveis erros ou incompletude.
2. **Representatividade dos gêneros:** A seleção dos gêneros musicais analisados foi baseada em critérios prévios e, portanto, pode não abranger todos os gêneros presentes na música brasileira. Outros gêneros menos populares ou emergentes podem não ter sido incluídos na análise, o que pode afetar a representatividade dos resultados.esc
3. **Generalização dos resultados:** Os resultados obtidos neste estudo são específicos para a amostra de artistas e gêneros musicais considerados. Eles podem não ser generalizáveis para o cenário musical como um todo ou para outras populações de artistas e gêneros. Portanto, é importante ter cautela ao extrapolar os resultados para contextos diferentes.
4. **Limitações das técnicas utilizadas:** As técnicas de aprendizagem de máquina e estatística aplicadas neste estudo têm suas próprias limitações. As escolhas dos algoritmos, parâmetros e métodos de análise podem influenciar os resultados, e existem outras abordagens que poderiam ser exploradas para obter perspectivas adicionais.
5. **Viés dos dados:** É possível que os dados utilizados apresentem algum tipo de viés, seja em relação à popularidade dos artistas, preferências musicais da plataforma Spotify ou outros fatores que possam impactar a representatividade dos resultados.

Essas limitações devem ser consideradas ao interpretar e discutir os resultados deste trabalho, bem como fornecer oportunidades para futuras pesquisas que possam abordar essas lacunas e aprimorar a compreensão da interação e influência entre os gêneros musicais brasileiros.

6.2. Trabalhos Futuros

Com base nas descobertas e nas limitações deste estudo, várias oportunidades de trabalho futuro surgem para aprofundar a compreensão da interação e influência entre os gêneros musicais brasileiros. Algumas sugestões são:

Ampliação do conjunto de dados: É recomendável ampliar o conjunto de dados utilizado, buscando incluir uma maior diversidade de artistas e gêneros musicais brasileiros. Isso pode ser feito por meio de parcerias com outras fontes de dados musicais, além da API do Spotify, ou por meio da coleta manual de informações.

Análise temporal: Realizar análises longitudinais para investigar a evolução dos gêneros musicais ao longo do tempo. Isso permitirá compreender as tendências e transformações na música brasileira, bem como identificar possíveis influências históricas e culturais.

Inclusão de dados qualitativos: Complementar as análises quantitativas com dados qualitativos, como letras de músicas ou entrevistas com especialistas da indústria musical. Isso pode fornecer perspectivas adicionais sobre as relações entre os gêneros musicais e as dinâmicas artísticas envolvidas.

Exploração de técnicas avançadas de aprendizado de máquina: Investigar o uso de técnicas avançadas de aprendizado de máquina, como redes neurais profundas,

para identificar padrões mais sutis e complexos na música brasileira. Isso pode ajudar a descobrir relações ainda mais precisas e detalhadas entre os gêneros musicais.

Análise de redes de artistas: Construir redes de artistas com base em colaborações musicais e medir a influência de cada artista na rede. Isso permitirá compreender a estrutura e as conexões da indústria musical brasileira de maneira mais abrangente.

Aplicações práticas: Explorar as aplicações práticas dos resultados deste estudo, como o desenvolvimento de sistemas de recomendação musical personalizados, análise de tendências musicais e planejamento estratégico para a indústria musical.

Ao abordar esses trabalhos futuros, será possível expandir o conhecimento sobre a música brasileira, sua diversidade e sua evolução ao longo do tempo. Essas pesquisas adicionais também podem contribuir para áreas aplicadas, beneficiando a indústria musical e o público em geral.

Referências

Dabbura, I. (2018). K-means clustering: Algorithm, applications, evaluation methods, and drawbacks — by imad dabbura — towards data science. `https://l1nq.com/m0dE7`. (Acessado em 05/15/2023).

Gabriela, G. (2014). Understanding music genre similarity. *Stanford Edu*, 31(2):5.

Kimball, R. and Ross, M. (2011). *The data warehouse toolkit: the complete guide to dimensional modeling*. John Wiley & Sons.

MacQueen, J. (1967). Classification and analysis of multivariate observations. In *5th Berkeley Symp. Math. Statist. Probability*, pages 281–297. University of California Los Angeles LA USA.

McCann, B. (2004). *Hello, hello Brazil: popular music in the making of modern Brazil*. Duke University Press.

Myers, R. H. and Myers, S. L. (2010). Correlação de pearson— research design and statistical analysis. wiley. `https://www.wiley.com/en-us/Research+Design+and+Statistical+Analysis\%2C+3rd+Edition-p-9780471746847`. (Acessado em 05/15/2023).

Opsahl, T. and Panzarasa, P. (2009). Clustering in weighted networks. *Social networks*, 31(2):155–163.

Paulo, M. and Denise, T. (2021). Proposta de base de dados para classificação automática de gêneros musicais brasileiros. *Revista Brasileira de Educação em Ciência da Informação*, 8(2):234.

Perrone, C. A. and Dunn, C. (2002). Brazilian popular music and globalization. *Journal of Popular Music Studies*, 14(2):163–165.

Peter, K. and Markus, S. (2013). A survey of music similarity and recommendation from music context data. *Association for Computing MachineryNew YorkNYUnited States*, 10(2):1–21.

Spotify (2023). Web API — Spotify for Developers — developer.spotify.com. https://developer.spotify.com/documentation/web-api. [Accessado em 13/05/2023].

Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(11).

Xiuli, W. (2023). Music similarity detection guided by deep learning model. *Computational Intelligence and Neuroscience*, 2023(2):10.

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DA PARAÍBA
Campus Campina Grande
R. Tranqüilino Coelho Lemos, 671, Dinamérica, CEP 58432-300, Campina Grande (PB)
CNPJ: 10.783.898/0003-37 - Telefone: (83) 2102.6200

Documento Digitalizado Ostensivo (Público)

Versão Final do TCC

Assunto: Versão Final do TCC
Assinado por: Josenildo Simao
Tipo do Documento: Dissertação
Situação: Finalizado
Nível de Acesso: Ostensivo (Público)
Tipo do Conferência: Cópia Simples

Documento assinado eletronicamente por:

- **Josenildo Simao da Silva, ALUNO (201811250032) DE BACHARELADO EM ENGENHARIA DE COMPUTAÇÃO - CAMPINA GRANDE**, em 29/06/2023 19:03:03.

Este documento foi armazenado no SUAP em 29/06/2023. Para comprovar sua integridade, faça a leitura do QRCode ao lado ou acesse https://suap.ifpb.edu.br/verificar-documento-externo/ e forneça os dados abaixo:

Código Verificador: 866061
Código de Autenticação: 18131e2d58