

LUCAS BONILHA HENRIQUE

**APLICAÇÃO DE MODELOS DE APRENDIZAGEM DE MÁQUINA NA
CLASSIFICAÇÃO DE EMOÇÕES MUSICAIS COM BASE NO
MODELO CIRCUMPLEXO DE RUSSELL**

**Monografia apresentada ao Programa de
Educação Continuada da Escola
Politécnica da Universidade de São Paulo,
para obtenção do título de Especialista,
pelo Programa de Pós-Graduação em
Engenharia de Dados e Big Data.**

SÃO PAULO

2024

LUCAS BONILHA HENRIQUE

**APLICAÇÃO DE MODELOS DE APRENDIZAGEM DE MÁQUINA NA
CLASSIFICAÇÃO DE EMOÇÕES MUSICAIS COM BASE NO
MODELO CIRCUMPLEXO DE RUSSELL**

Monografia apresentada ao Programa de Educação Continuada da Escola Politécnica da Universidade de São Paulo, para obtenção do título de Especialista, pelo Programa de Pós-Graduação em Engenharia de Dados e Big Data.

Área de concentração: Tecnologia da Informação – Engenharia/ Tecnologia/ Gestão

Orientador: Prof.^a Dr^a Marcia Ito

SÃO PAULO

2024

FICHA CATALOGRÁFICA

B. Henrique, Lucas

APLICAÇÃO DE MODELOS DE APRENDIZAGEM DE MÁQUINA NA
CLASSIFICAÇÃO DE EMOÇÕES MUSICAIS COM BASE NO MODELO
CIRCUMPLEXO DE RUSSELL / L. B. Henrique -- São Paulo, 2024.
63 p.

Monografia (Especialização em Engenharia de Dados e Big Data) -
Escola
Politécnica da Universidade de São Paulo. PECE – Programa de Educação
Continuada em Engenharia.

1.Reconhecimento de Emoções Musicais 2. Aprendizado de Máquina
3.Modelo Circumplexo de Russell I.Universidade de São Paulo. Escola
Politécnica. PECE – Programa de Educação Continuada em Engenharia
II.t.

AGRADECIMENTOS

Ao chegar ao final desta jornada acadêmica, reflito sobre todos aqueles que estiveram ao meu lado, oferecendo apoio incondicional, incentivo e sabedoria. É com grande gratidão que dedico estas palavras a cada um deles.

Primeiramente, Deus que por sua bondade e misericórdia tem nos guiado todos os dias, sendo um refúgio e um alicerce. À minha esposa, cujo amor, compreensão e paciência foram fundamentais em cada etapa deste caminho. Suas palavras de incentivo e seu apoio incansável foram a força que me impulsionou nos momentos mais desafiadores. A sua presença constante e amorosa foi a luz que guiou minha jornada.

À minha família, que sempre acreditou em mim e nas minhas capacidades, mesmo nos momentos em que eu duvidava de mim mesmo. O apoio e incentivo de vocês foram os pilares que me sustentaram, proporcionando o ambiente necessário para que eu pudesse crescer e evoluir tanto pessoal quanto profissionalmente. Cada um de vocês desempenhou um papel essencial em minha jornada, enchendo minha vida de motivação e alegria.

À minha equipe no Itaú, que se tornou uma fonte constante de inspiração e motivação. Trabalhar ao lado de profissionais tão dedicados e apaixonados pelo que fazem me incentivou a buscar sempre o meu melhor. A camaradagem, o espírito de equipe e o ambiente de suporte mútuo foram fundamentais para o meu desenvolvimento profissional e pessoal.

Aos meus gestores, cuja liderança e orientação foram essenciais em minha trajetória. Agradeço a confiança depositada em mim e pelas oportunidades de crescimento que me foram oferecidas. O aprendizado e as experiências compartilhadas sob sua supervisão foram inestimáveis.

Por fim, mas certamente não menos importante, à minha orientadora, que me encorajou nos momentos em que duvidei da possibilidade de alcançar este objetivo. Sua sabedoria, paciência e orientação incisiva foram fundamentais para superar os desafios e alcançar este marco em minha vida acadêmica.

CURSO ENGENHARIA DE BIG DATA

Coord.: Prof. Solange N. Alves de Souza

Vice-Coord.: Prof. Pedro Luiz Pizzigatti Corrêa

Perspectivas profissionais alcançadas com o curso:

Consegui expandir consideravelmente meu escopo de atuação na empresa. Além de assumir projetos mais complexos e desafiadores, passei a desempenhar um papel fundamental no treinamento e desenvolvimento de profissionais mais jovens, compartilhando com eles as competências e ideias adquiridas durante a especialização.

Essa experiência de orientar novos talentos não apenas contribuíram para o crescimento da equipe, mas também reforçou meu próprio desenvolvimento profissional. A capacidade de ensinar e orientar outros profissionais reafirma e solidifica o conhecimento, além de desenvolver habilidades de liderança e comunicação. Comecei a receber convites para participar de comunidades profissionais de Engenharia de Dados, uma oportunidade excelente para networking, troca de experiências e permanência atualizada com as tendências e inovações do setor. Além disso, outras empresas começaram a demonstrar interesse no meu perfil profissional.

A especialização não apenas aprimorou minhas habilidades técnicas, mas também abriu novas portas e me proporcionou uma oportunidade para influenciar positivamente a área e contribuir para o desenvolvimento de outros profissionais.

RESUMO

Este trabalho aborda o desafio de classificar emoções em músicas utilizando as características de áudio disponibilizadas pela API do Spotify. Inserido no contexto da crescente influência das plataformas de streaming na indústria musical e no consumo de mídia, o estudo foca na exploração do campo do Reconhecimento de Emoções Musicais (MER). O objetivo central é avaliar a eficácia das características de áudio do Spotify, como ritmo, timbre, harmonia, entre outras, na classificação das emoções musicais, baseando-se nos quadrantes de Russell. A metodologia adotada envolveu a construção e análise de um conjunto de dados com 900 faixas musicais, a preparação e normalização desses dados, e a aplicação de diferentes modelos de aprendizado de máquina, incluindo *K-Nearest Neighbors*, *Gaussian Naive Bayes*, *Support Vector Machine*, *Random Forest*, *Gradient Boosting* e *XGBoost*. Os resultados evidenciaram variações significativas no desempenho dos modelos, com alguns exibindo maior precisão em classes emocionais específicas. Conclui-se que as características de áudio do Spotify possuem potencial para a classificação de emoções em músicas, porém a complexidade e variações das emoções musicais exigem estratégias mais integradas ou híbridas, combinando diferentes modelos para melhor precisão. As descobertas deste trabalho abrem novos caminhos para a personalização de listas de reprodução e aplicações em musicoterapia, destacando a relevância da interação entre ciência de dados, engenharia de áudio e psicologia musical.

Palavras-chave: Reconhecimento de Emoções Musicais, Aprendizado de Máquina, Modelo Circumplexo de Russell.

ABSTRACT

This study addresses the challenge of classifying emotions in music using the audio features provided by the Spotify API. Set against the backdrop of the growing influence of streaming platforms in the music industry and media consumption, the research focuses on exploring the field of Musical Emotion Recognition (MER). The primary goal is to assess the effectiveness of Spotify's audio features, such as rhythm, timbre, harmony, and others, in classifying musical emotions, based on Russell's quadrants. The methodology involved constructing and analyzing a dataset of 900 musical tracks, preparing and normalizing this data, and applying various machine learning models, including K-Nearest Neighbors, Gaussian Naive Bayes, Support Vector Machine, Random Forest, Gradient Boosting, and XGBoost. The results showed significant variations in the performance of the models, with some demonstrating higher accuracy in specific emotional classes. It concludes that Spotify's audio features have potential for classifying emotions in music, but the complexity and nuances of musical emotions require more integrated or hybrid strategies, combining different models for improved accuracy. The findings of this work pave new avenues for playlist personalization and applications in music therapy, highlighting the significance of the interplay between data science, audio engineering, and musical psychology.

Keywords: Musical Emotion Recognition, Machine Learning, Emotion Classification, Russell's circumplex model

LISTA DE FIGURAS

Figura 1 – Modelo Circumplexo de Afeto (adaptado de Russel (1980) (RUSSELL, 1980)).....	19
Figura 2 – Espectrografia comparativa do trompete, bateria e contrabaixo.	21
Figura 3 - Gráfico comparativo da receita do mercado musical brasileiro de 2018 a 2022 (adaptado de (PRÓ-MÚSICA, 2022)).....	24
Figura 4 - Linha do Conjunto de Dados Original	30
Figura 5 - Evidência da inexistência de valores ausentes no Dataset.....	33
Figura 6 - Evidência de tipos de dados adequados.....	33
Figura 7 - Histograma da variável <i>energy</i>	35
Figura 8 - Correlação das variáveis numéricas	36
Figura 9 - Distribuição da variável <i>valence</i> por Quadrante de Russell.....	36
Figura 10 - Relação de dispersão entre <i>energy</i> e <i>valence</i>	37

LISTA DE TABELAS

Tabela 1 - Comparação dos Modelos de Aprendizagem de Máquina.....49

SUMÁRIO

1 INTRODUÇÃO	11
1.1 Objetivo.....	11
1.2 Justificativa	13
1.3 Metodologia	13
1.4 Organização do Trabalho	14
2 FUNDAMENTAÇÃO TEÓRICA	16
2.1 Reconhecimento de Emoções Musicais (MER).....	16
2.1.1 Métodos Analíticos e Computacionais.....	17
2.1.2 Modelo Circumplexo de Russel	18
2.2 Características de Áudio	19
2.2.1 Ritmo	19
2.2.2 Timbre	20
2.2.3 Harmonia	22
2.2.4 Melodia	22
2.2.5 Dinâmica.....	23
2.3 Plataformas de Streaming de Música e Influência nas Emoções	23
2.3.1 Impacto no consumo de Música	23
2.3.2 Influência das Plataformas de <i>Streaming</i> nas Emoções.....	25
2.3.3 A API do Spotify e suas características.....	25
2.4 Modelos de Aprendizagem de Máquina	26
2.4.1 Modelos utilizados na tarefa de classificação do MER	27
3 DESENVOLVIMENTO.....	29
3.1 Descrição do Conjunto de Dados	29
3.1.1 Processamento e Geração do Conjunto de dados	31

3.1.2	Características do Conjunto de Dados e sua Relevância para o MER..	32
3.1.3	Tamanho e Diversidade do Conjunto de Dados	32
3.2	Pré-processamento e Limpeza dos Dados	32
3.2.1	Verificação e Tratamento de Valores Ausentes.....	32
3.2.2	Normalização dos Dados.....	34
3.3	Análise Exploratória de Dados (EDA).....	34
3.4	Modelagem e Algoritmos de Aprendizado de Máquina com Preparação de Dados e Otimização	37
3.4.1	Preparação dos Dados	37
3.4.2	Modelagem e Otimização de Hiperparâmetros	38
3.5	Avaliação dos Modelos.....	40
3.5.1	<i>K-Nearest Neighbors (KNN)</i>	40
3.5.2	Gaussian Naive Bayes	41
3.5.3	<i>Support Vector Machine (SVM)</i>	42
3.5.4	Random Forest.....	43
3.5.5	Gradient Boosting	44
3.5.6	XGBoost	44
3.6	Classificação de Novas Músicas Utilizando Modelos de Aprendizado de Máquina.....	45
3.6.1	Detalhes Técnicos da Rotina de Classificação	46
3.6.2	Exemplo Prático de Classificação.....	46
3.6.3	Implicações e Potenciais Aplicações	47
4	ANÁLISE DOS RESULTADOS.....	49
4.1	Comparação dos Modelos de Aprendizagem de Máquina	49
4.2	Discussão dos Resultados.....	50
4.3	Conclusão sobre os resultados	50
5	CONCLUSÃO E TRABALHOS FUTUROS.....	51

5.1 Conclusão.....	51
5.2 Direções para Pesquisas Futuras.....	51
5.2.1 Exploração de Modelos Híbridos:.....	51
5.2.2 Uso de Técnicas Avançadas de Aprendizado Profundo:.....	52
5.2.3 Aplicações Práticas e Comerciais:.....	52
5.3 Reflexão Final.....	53
REFERÊNCIAS BIBLIOGRÁFICA	54

1 INTRODUÇÃO

A evolução do consumo de música na era digital tem sido notável, particularmente com o surgimento de serviços de streaming como o Spotify. Esta plataforma, que hoje oferece mais de 70 milhões de faixas, tem transformado a maneira como as pessoas acessam e interagem com a música. Paralelamente, o campo do Reconhecimento de Emoções Musicais (MER) vem ganhando destaque ao explorar as complexas relações entre música e emoções. MER investiga como a música evoca emoções nos ouvintes e busca quantificar essas experiências emocionais através da análise computacional das características de áudio das faixas musicais. (GRIFFITHS et al., 2021). Portanto, emoção musical pode ser entendida como o conjunto de respostas emocionais que a música pode suscitar em quem a ouve, sendo essas emoções influenciadas por aspectos intrínsecos à música, tais como sua estrutura, elementos e performance.

Assim, partindo da crescente influência das plataformas de streaming na indústria musical e o impacto direto no consumo de mídia, destacando o Spotify devido à sua popularidade, observou-se não apenas acesso a uma vasta biblioteca de músicas, mas também a disponibilidade de dados de áudio que oferecem oportunidades para análise e pesquisa em MER. No entanto, apesar dessa disponibilidade, existe uma lacuna na compreensão de como as características de áudio fornecidas pelo Spotify podem ser empregadas como métricas representativas no campo do MER. Embora o foco do Spotify tenha sido construir personalizações baseadas principalmente no histórico de escuta dos usuários, com menos ênfase no conteúdo de áudio devido à complexidade da tarefa computacional, a análise dessas características de áudio pode fornecer insights valiosos para a compreensão da Emoção Musical (PANDA et al., 2021).

1.1 Objetivo

O objetivo geral deste trabalho é fazer uma análise comparativa dos modelos de aprendizagem de máquina para a classificação das emoções com base nas características de áudio fornecidas pela API do Spotify. Um conjunto de dados

baseado nos quadrantes de Russell (RUSSELL, 1980) é utilizado, após a análise comparativa o modelo mais eficaz para esta tarefa é identificado.

Os objetivos específicos para atingir a meta proposta são:

Objetivos Específicos

- **Construção e Análise do Conjunto de Dados:** Construir um conjunto de dados apropriado que utilize os quadrantes de Russell para a classificação das emoções em músicas e analisar as características de áudio fornecidas pela API do Spotify;
- **Preparação dos Dados:** Descrever as etapas de preparação dos dados, incluindo a limpeza, normalização e transformação dos dados de áudio para garantir que sejam adequados para o treinamento dos modelos de aprendizagem de máquina;
- **Treinamento e Validação de Modelos de Aprendizagem de Máquina:** *K-Nearest Neighbors (KNN):* Treinar um modelo KNN, explicar a seleção do número de vizinhos e avaliar a eficácia do modelo para a classificação de emoções musicais. *Naive Bayes:* Aplicar o modelo *Naive Bayes*, discutir sua simplicidade e eficiência e avaliar seu desempenho no conjunto de dados. *Gradient Boosting:* Implementar o modelo de *Gradient Boosting*, detalhar a escolha de parâmetros como a taxa de aprendizagem e o número de árvores, e avaliar sua precisão e eficiência. *Random Forest:* Treinar e validar um modelo de *Random Forest*, abordar o papel do número de árvores na decisão e a importância de features aleatórias, e testar sua eficácia. *Support Vector Machine (SVM):* Implementar um modelo SVM, explorar a escolha do kernel e parâmetros como margem e C, e testar sua capacidade de classificar emoções em músicas.
- **Análise Comparativa dos Modelos:** Comparar os modelos de aprendizagem de máquina com base em métricas como precisão, *recall*, F1-score e tempo de treinamento, destacando os pontos positivos e negativos de cada modelo no contexto da classificação das emoções musicais.
- **Seleção do Melhor Modelo:** Identificar o modelo mais eficaz para a classificação de emoções musicais com base nas características de áudio do Spotify, considerando os resultados da análise comparativa.

1.2 Justificativa

A popularidade no uso das plataformas de *streaming* de música, como o Spotify, no cotidiano das pessoas aumenta a relevância de compreender como as características de áudio influenciam as emoções, uma questão central no campo do Reconhecimento de Emoções Musicais (MER).

Um aspecto fundamental desta pesquisa é a potencial aplicação de MER na musicoterapia, especialmente no tratamento de pacientes psiquiátricos. O estudo de (LAZAROV; PINE; BAR-HAIM, 2017), demonstra como intervenções terapêuticas que integram música podem ser eficazes no tratamento de transtornos emocionais, como o transtorno de ansiedade social. Este trabalho sugere que a aplicação direcionada de música em contextos terapêuticos pode ter efeitos benéficos significativos no bem-estar emocional dos pacientes (LAZAROV; PINE; BAR-HAIM, 2017).

Além disso, as características de áudio podem ser usadas de maneira confiável para inferir emoções em músicas, o que é relevante para a seleção de músicas na musicoterapia, permitindo intervenções mais personalizadas e potencialmente mais eficazes (GRIFFITHS et al., 2021)

Este trabalho visa contribuir para o campo do MER e fornecer ideias de soluções práticas que pode ser aplicado no tratamento de condições psiquiátricas. Avaliar a eficácia das características de áudio do Spotify pode levar a avanços significativos na personalização da musicoterapia, oferecendo benefícios tangíveis para os pacientes.

1.3 Metodologia

O método de pesquisa adotado para esta trabalho é um estudo exploratório e aplicado, que utiliza a abordagem de mineração de dados para analisar os modelos de aprendizagem de máquina na classificação de emoções em músicas por meio das características de áudio do Spotify. A pesquisa foi realizada em várias etapas sequenciais que envolvem a coleta de dados, pré-processamento, aplicação de modelos de aprendizagem de máquina e análise estatística dos resultados.

Na execução do experimento, seleciona-se um conjunto de dados preenchido com as características de áudio de músicas e suas respectivas classificações emocionais baseadas nos quadrantes de Russell. É feito um pré processamento no conjunto de dados, para assegurar a qualidade e a uniformidade dos dados. Aplicam-se diversos modelos de aprendizagem de máquina - como *KNN*, *Naive Bayes*, *Gradient Boosting*, *Random Forest*, e *SVM* - para realizar a análise comparativa e ao final identificar o modelo que tem a melhor eficiência na classificação das emoções segundo o quadrante de Russell.

Cada modelo é avaliado de acordo com um protocolo estabelecido, que envolve a divisão dos dados em conjuntos de treinamento e teste, a otimização de parâmetros por meio da validação cruzada e o ajuste fino baseado em métricas de desempenho como precisão, recall e F1-score. Discutem-se os pontos positivos e negativos de cada modelo, considerando o desempenho e a complexidade computacional.

Os resultados desta análise são detalhados no capítulo do experimento, onde se apresenta uma comparação entre os modelos, resultando na seleção do modelo mais eficiente para a tarefa proposta.

1.4 Organização do Trabalho

O presente trabalho está estruturado em capítulos de uma forma que visem facilitar a compreensão do estudo e dos resultados obtidos.

Capítulo 1: Introdução – Define-se o contexto da pesquisa, apresentando-se a relevância do estudo no campo do Reconhecimento de Emoções Musicais (MER) e a utilização das características de áudio do Spotify. Discutem-se também a motivação e os objetivos do trabalho.

Capítulo 2: Referencial Teórico – Aborda-se a fundamentação teórica, revisando-se os conceitos relacionados ao MER, às características de áudio relevantes, às plataformas de streaming de música e aos modelos de aprendizagem de máquina. Também são descritos os modelos de aprendizagem de máquina pertinentes a este estudo, proporcionando um entendimento dos métodos e tecnologias envolvidas.

Capítulo 3: Desenvolvimento – Detalha-se a seleção e preparação do conjunto de dados e implementação dos modelos, a configuração dos experimentos, o processo de treinamento e validação e a análise dos resultados.

Capítulo 4: Análise dos Resultados – Compara-se os modelos de aprendizagem de máquina quanto à sua eficácia na classificação das emoções musicais, discutindo-se os pontos positivos e negativos de cada um.

Capítulo 5: Conclusão e Trabalhos Futuros – Conclui-se sobre a eficácia das características de áudio do Spotify e a adequação dos modelos de aprendizagem de máquina, sugerindo-se direções para pesquisas futuras com base nas descobertas do estudo.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são abordada a fundamentação teórica para o estudo do Reconhecimento de Emoções Musicais (MER), enfatizando a importância de características de áudio específicas e a influência das plataformas de *streaming* de música, como o Spotify, no contexto da música e emoções. Adicionalmente, são explorados os modelos de aprendizagem de máquina pertinentes a este estudo, proporcionando um entendimento dos métodos e tecnologias envolvidas.

2.1 Reconhecimento de Emoções Musicais (MER)

O estudo do Reconhecimento de Emoções Musicais (MER) explora a relação entre música e emoção. Esta área combina análise musical e ciência de dados para entender como a música influencia os sentimentos. O MER é importante para descobrir como diferentes músicas evocam reações emocionais variadas em ouvintes, usando métodos analíticos e computacionais para essa finalidade.

As raízes históricas do (MER) parte da interseção da música, psicologia e ciência da computação. Uma compreensão sólida dos fundamentos históricos do MER torna-se relevante para contextualizar o desenvolvimento posterior desta área multidisciplinar. Um dos marcos mais importantes na história do MER remonta à antiga Grécia, onde filósofos como Aristóteles e Platão exploraram as conexões entre música e emoção. Eles acreditavam que a música poderia influenciar o estado emocional das pessoas e desempenhar um papel na catarse e na terapia emocional (XU et al., 2023).

Durante os séculos seguintes, a música continuou a ser considerada uma poderosa ferramenta para expressar e evocar emoções. No século XIX, compositores românticos como Beethoven e Wagner criaram obras que visavam provocar fortes respostas emocionais em seus ouvintes (SENA MOORE, 2016).

A influência do MER na música e na psicologia cresceu à medida que novas tecnologias, como a análise de áudio computacional, permitiram a quantificação das características musicais associadas às emoções. Hoje, o MER combina análise musical, teorias psicológicas e algoritmos de aprendizagem de máquina para entender

como a música evoca emoções humanas. A pesquisa nessa área tem implicações significativas, desde a criação de lista de reprodução emocionais personalizadas até aplicações na terapia musical (XU et al., 2023).

2.1.1 Métodos Analíticos e Computacionais

A análise de características musicais desempenha um papel fundamental, e representa uma etapa relevante que envolve a extração e a quantificação de atributos específicos das músicas, como ritmo, timbre, harmonia, melodia e dinâmica. Essas características são essenciais para compreender como a música evoca emoções nos ouvintes. Através da análise detalhada desses elementos, é possível identificar padrões que estão relacionados às emoções específicas transmitidas pela música (GOMEZ-CANON et al., 2021).

O uso de algoritmos de aprendizado de máquina e reconhecimento de padrões é uma parte relevante do MER. Esses algoritmos são treinados com base em conjuntos de dados que associam características musicais a etiquetas emocionais, permitindo que os modelos aprendam a reconhecer padrões emocionais nas músicas. Abordagens como Support Vector Machines (SVM), Redes Neurais Artificiais (RNA) e Random Forest são comumente empregadas para esse fim. Essas técnicas são capazes de generalizar a relação entre características musicais e emoções, tornando possível prever as emoções evocadas por músicas não vistas anteriormente (Han et al., 2022).

É importante mencionar o desenvolvimento de conjuntos de dados *benchmark* no MER. Esses conjuntos de dados contêm faixas musicais rotuladas com emoções específicas, permitindo uma avaliação objetiva do desempenho dos algoritmos de reconhecimento de emoções. Além disso, o desenvolvimento de *benchmarks* para análise emocional da música é essencial para promover a consistência e a comparação entre diferentes abordagens de MER. Ter acesso a conjuntos de dados de alta qualidade e devidamente rotulados é um pré-requisito para avançar na pesquisa nessa área (ALJANAKI; YANG; SOLEYMANI, 2017).

Além da análise de áudio, as abordagens multimodais estão se tornando cada vez mais relevantes no MER. Isso envolve a combinação de informações de áudio com outros tipos de dados, como informações visuais, contextuais ou até mesmo dados

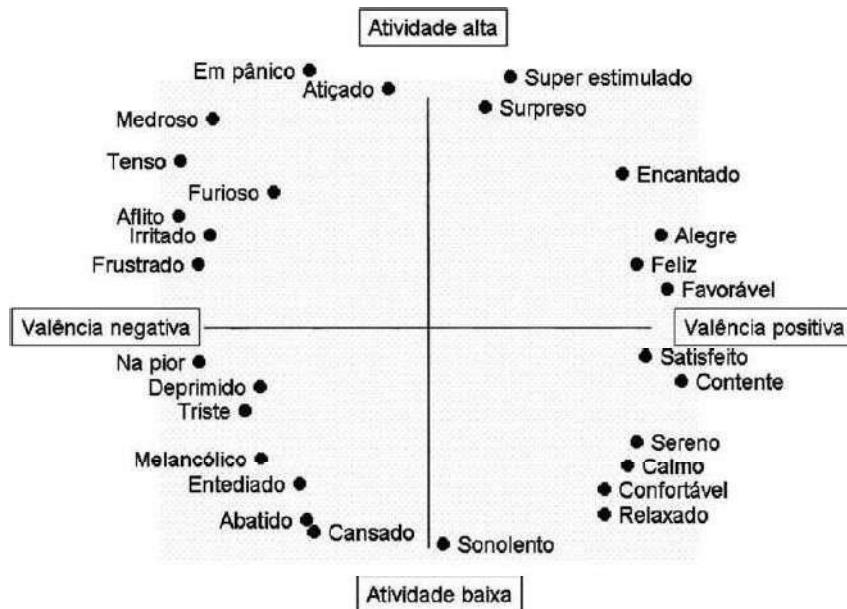
fisiológicos dos ouvintes. A inclusão de várias fontes de informação pode enriquecer a compreensão das emoções evocadas pela música. Por exemplo, a combinação de áudio com informações visuais de videoclipes pode proporcionar uma representação mais completa das emoções transmitidas pela música. Essa abordagem multidimensional está se tornando um tópico de pesquisa importante no campo do MER (HAN et al., 2022).

2.1.2 Modelo Circumplexo de Russel

Um dos conceitos fundamentais no estudo das emoções em música é o modelo circumplexo de afeto de Russell. Este modelo propõe que as emoções podem ser mapeadas em um espaço bidimensional com dois eixos principais: ativação e valência. A ativação refere-se à intensidade da emoção, enquanto a valência descreve a qualidade positiva ou negativa da emoção. O modelo é dividido em quatro quadrantes, cada um representando diferentes combinações de ativação e valência (RUSSELL, 1980).

No primeiro quadrante, com alta ativação e valência positiva, encontram-se emoções como 'excitação' e 'felicidade'. Já o segundo quadrante, que também tem alta ativação, mas com valência negativa, inclui emoções como 'raiva' e 'ansiedade'. No terceiro quadrante, com baixa ativação e valência positiva, estão emoções como 'relaxamento' e 'calma'. Por fim, o quarto quadrante, com baixa ativação e valência negativa, engloba emoções como 'tristeza' e 'depressão'. A Figura 1 apresenta o Modelo Circumplexo de Afeto (RUSSELL, 1980).

Figura 1 – Modelo Circumplexo de Afeto (adaptado de Russel (1980) (RUSSELL, 1980)).



Fonte: (SANTANA et al., 2021)

Esta abordagem circular e bidimensional das emoções oferece uma maneira de entender como as pessoas categorizam suas experiências afetivas, desviando-se de visões anteriores que viam as emoções como fenômenos independentes e isolados (SCHLOSBERG, 1952). Em vez disso, o modelo circumplexo sugere uma continuidade e inter-relação entre diferentes estados afetivos, fornecendo uma perspectiva mais integrada e holística das emoções humanas.

2.2 Características de Áudio

As características de áudio são elementos intrínsecos à música, como ritmo, timbre, harmonia, melodia e dinâmica, que desempenham um papel fundamental na evocação de emoções nos ouvintes. Explorar as principais características de áudio utilizadas no Reconhecimento de Emoções Musicais (MER) e como elas estão relacionadas às reações emocionais dos ouvintes torna-se relevante.

2.2.1 Ritmo

Ritmo, um elemento fundamental da música, é a organização temporal de sons e silêncios, criando padrões que estruturam a música. Em termos de Reconhecimento

de Emoções Musicais (MER), o ritmo é essencial, pois atua como um dos principais indicadores das emoções transmitidas por uma peça musical. O ritmo influencia a percepção de energia e movimento na música, impactando diretamente a resposta emocional dos ouvintes (THAUT; MCINTOSH; HOEMBERG, 2015).

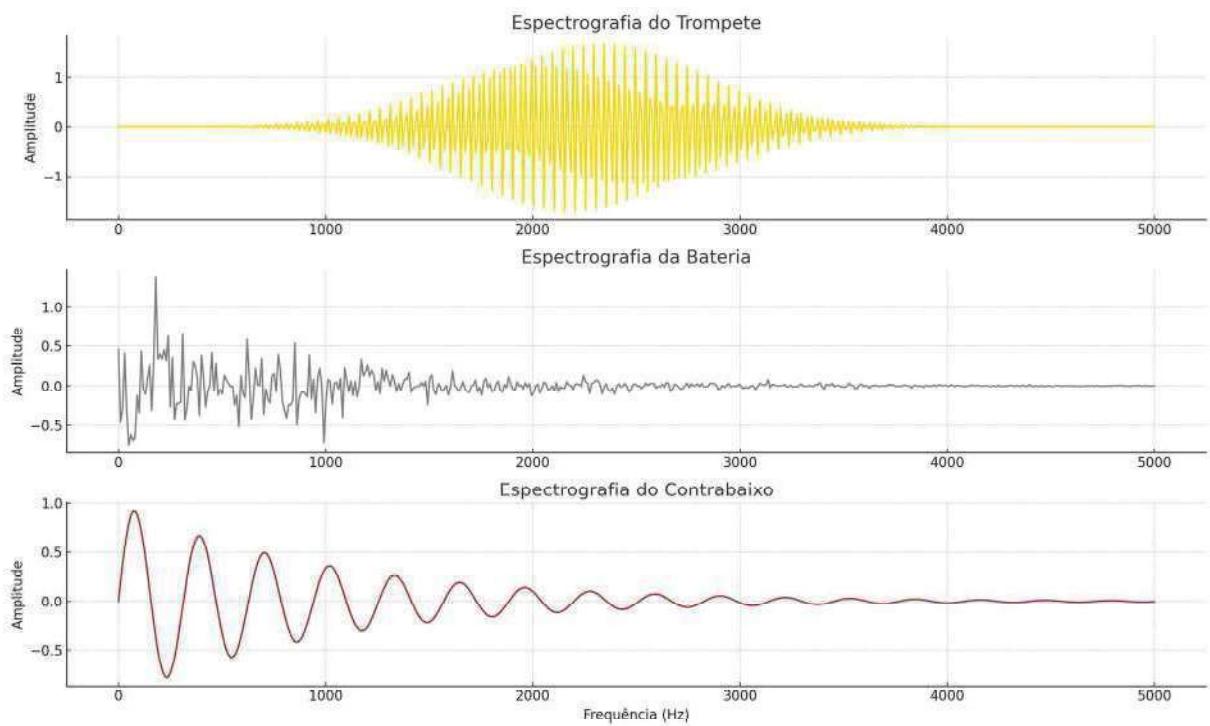
Diferentes padrões rítmicos possuem a capacidade de evocar emoções variadas. Por exemplo, ritmos rápidos e regulares tendem a ser associados a sentimentos de alegria ou excitação, enquanto ritmos lentos e irregulares podem evocar tristeza ou calma. Essa relação é explicada pela maneira como o cérebro humano processa padrões rítmicos e os associa a estados emocionais específicos. A sincronização do ritmo com as batidas cardíacas ou respiração pode também influenciar a resposta emocional (JUSLIN; VÄSTFJÄLL, 2008).

2.2.2 Timbre

O timbre é a qualidade do som que permite distinguir diferentes fontes sonoras, mesmo quando elas produzem notas na mesma frequência. Ele é determinado por aspectos como a forma da onda sonora e a sua espectrografia. No contexto do MER, o timbre é relevante, pois diferentes timbres podem alterar significativamente a percepção e a emoção evocada por uma peça musical (IVERSON; KRUMHANSL, 1993).

Para exemplificar a diferença do timbre, a Figura 2 representam um conjunto de gráficos com a espectrografia de três instrumentos distintos:trompete, bateria e contrabaixo.

Figura 2 – Espectrografia comparativa do trompete, bateria e contrabaixo.



Fonte: Autor (2023)

- **Trompete :** O espectro do trompete mostra picos agudos em frequências mais altas, refletindo seu som brilhante e penetrante. O trompete tem uma presença marcante em harmônicos elevados, o que contribui para sua qualidade sonora distintiva. A energia concentrada em frequências altas é típica dos instrumentos de metal.
- **Bateria :** A bateria, sendo um conjunto de instrumentos de percussão, exibe um espectro mais irregular, com variações amplas em amplitude ao longo de todo o espectro de frequência. Isso reflete a natureza percussiva e a diversidade de sons que a bateria pode produzir, desde graves profundos de tambores até sons agudos de pratos.
- **Contrabaixo:** O espectro do contrabaixo é dominado por frequências mais baixas, com um decaimento gradual em frequências mais altas. Isso ilustra o som profundo e ressonante do contrabaixo, um instrumento fundamental na produção de linhas de baixo e na adição de profundidade harmônica à música.

O timbre tem um impacto profundo na emoção transmitida pela música. Variações sutis no timbre podem mudar a atmosfera de uma música, influenciando como as emoções são percebidas e processadas pelo ouvinte. Por exemplo, um timbre mais

escuro e mais profundo pode evocar sentimentos de melancolia ou serenidade, enquanto timbres mais claros e nítidos podem ser associados a sentimentos de alegria ou excitação (IVERSON; KRUMHANSL, 1993).

2.2.3 Harmonia

Harmonia, em música, refere-se à forma como as notas são combinadas simultaneamente, criando acordes, e sequencialmente, formando progressões harmônicas. É um aspecto que influencia a estrutura, textura e cor emocional de uma peça musical. No contexto do MER (Reconhecimento de Emoções Musicais), a harmonia desempenha um papel significativo na indução de emoções, pois diferentes progressões harmônicas podem evocar sentimentos distintos (KRUMHANSL, 1997).

A relação entre harmonia e emoção é complexa e profundamente enraizada na psicologia humana. Acordes maiores são frequentemente associados a emoções positivas, como alegria e felicidade, enquanto acordes menores tendem a ser relacionados a sentimentos de tristeza ou melancolia. A tensão e resolução criadas pelas progressões harmônicas também têm um impacto significativo na experiência emocional do ouvinte (HURON, 2006).

2.2.4 Melodia

Melodia é a sequência linear de notas musicais que é percebida como uma única entidade. Ela guia o ouvinte através da música, oferecendo o tema central ao redor do qual a harmonia e o ritmo são construídos. No contexto do MER (Reconhecimento de Emoções Musicais), a melodia é relevante, pois carrega a identidade emocional da música, transmitindo sentimentos e estados de ânimo (MEYER, 1956).

A melodia tem um impacto direto na emoção percebida em uma peça musical. Melodias ascendentes frequentemente são associadas a sentimentos de otimismo e alegria, enquanto melodias descendentes podem evocar tristeza ou nostalgia. A variação melódica, incluindo o alcance e a velocidade das notas, também desempenha um papel significativo na expressão emocional (JUSLIN; LAUKKA, 2003).

2.2.5 Dinâmica

A dinâmica na música refere-se ao grau de intensidade ou volume com que uma nota, frase ou peça é tocada. Ela é um elemento que contribui para a expressão emocional na música, oferecendo variações de suavidade e força. No MER (Reconhecimento de Emoções Musicais), a dinâmica representa diferentes níveis de intensidade e podem transmitir diferentes emoções, desde a sutileza de um pianíssimo até a força de um fortíssimo (SCHUBERT, 2004).

A dinâmica afeta profundamente a experiência emocional da música. Passagens tocadas suavemente podem evocar sentimentos de calma ou tristeza, enquanto passagens mais fortes podem expressar alegria ou raiva. A variação dinâmica dentro de uma peça também pode criar um senso de tensão e liberação, contribuindo para a narrativa emocional da música (JUSLIN, 1993).

2.3 Plataformas de Streaming de Música e Influência nas Emoções

O impacto cultural e social exercidos pelas plataformas de streaming de música promove a necessidade de se observar o consumo musical e suas influências nas emoções humanas. Em particular, destaca-se o papel do Spotify, que, por meio de sua API, oferece uma janela para a análise detalhada das características de áudio e é tratada aqui como representante das demais plataformas enquanto agentes de mudanças na relação entre música, ouvinte e sociedade.

2.3.1 Impacto no consumo de Música

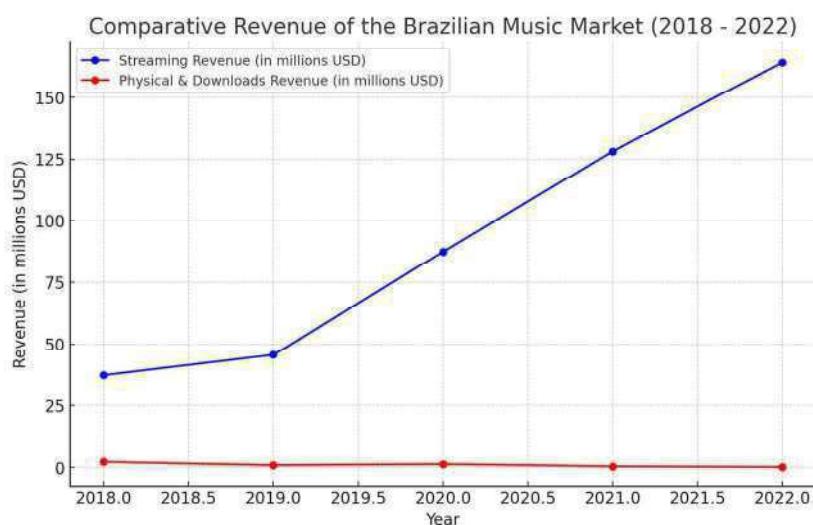
As plataformas de streaming de música redefiniram a indústria musical, alterando o comportamento do consumidor e a economia da música. A facilidade de acesso e a personalização oferecidas por esses serviços mudaram a maneira como as pessoas descobrem e interagem com a música. Este paradigma de acesso, em oposição à posse, sugere uma nova era na qual o valor está menos na propriedade do conteúdo e mais na experiência proporcionada pelo serviço (ALMEIDA SANTOS; FREIRE, 2013). Além disso, a presença da música por meio do streaming influencia o ambiente

social e cultural, promovendo a diversificação da experiência musical e oferecendo novas oportunidades para artistas emergentes e estabelecidos.

A curadoria, tanto algorítmica quanto humana, desempenha um papel relevante na orientação dos ouvintes através de vastas coleções de música, permitindo a descoberta personalizada e a criação de comunidades em torno de gostos musicais compartilhados (MOSCHETTA; VIEIRA, 2018). Este novo modelo também levanta questões sobre a sustentabilidade econômica para os artistas, uma vez que os mecanismos de remuneração baseados em streaming diferem significativamente das vendas tradicionais de álbuns e faixas individuais.

A Figura 3 apresenta um comparativo sobre o mercado fonográfico brasileiro entre 2018 e 2022.

Figura 3 - Gráfico comparativo da receita do mercado musical brasileiro de 2018 a 2022 (adaptado de (PRÓ-MÚSICA, 2022)).



Fonte: (PRÓ-MÚSICA, 2022)

As receitas de *streaming* (em azul) mostram um crescimento significativo ao longo dos anos, refletindo a crescente popularidade e adoção deste formato de consumo de música no Brasil. Por outro lado, as receitas de vendas físicas e *downloads* (em vermelho) apresentam uma tendência de declínio durante o mesmo período, indicando uma mudança clara na preferência dos consumidores em direção ao *streaming*.

Em resumo, as plataformas de *streaming* têm efeitos profundos não apenas na economia da música, mas também na cultura musical global. Elas democratizam o acesso à música, mas também criam desafios e dinâmicas tanto para consumidores quanto para criadores de música(SINCLAIR; TINSON, 2017).

2.3.2 Influência das Plataformas de *Streaming* nas Emoções

As plataformas de streaming de música têm um papel significativo na influência das emoções dos ouvintes. Estudos indicam que o uso dessas plataformas pode facilitar a regulação emocional, permitindo que os usuários selezionem músicas que correspondam ao seu estado emocional desejado ou atual (WADLEY et al., 2019). A capacidade de escolher entre uma variedade quase ilimitada de músicas e a criação de lista de reprodução personalizadas contribuem para que o ouvinte encontre conforto, alegria ou até mesmo companhia em momentos de solidão.

A personalização e o acesso imediato à música em *streaming* também podem intensificar a experiência emocional da música. A plataforma de streaming atua como um intermediário que oferece não apenas entretenimento, mas também um espaço para experiências emocionais profundas e significativas (MOSCHETTA; VIEIRA, 2018). Esta interação contínua com a música através do streaming pode ter um impacto duradouro no bem-estar emocional dos indivíduos, influenciando o humor, a produtividade e até a saúde mental.

As plataformas de *streaming* são uma ferramenta poderosa para a regulação emocional e para a criação de uma atmosfera personalizada que pode elevar ou acalmar o espírito do ouvinte, desempenhando um papel essencial na vida diária e na saúde emocional (HAGEN, 2015).

2.3.3 A API do Spotify e suas características.

A API (Application Programming Interface) do Spotify desempenha um papel relevante na análise de características de áudio das faixas disponíveis em seu catálogo. Através dela, é possível obter uma análise detalhada de baixo nível da estrutura musical e do conteúdo de uma faixa, abrangendo elementos como ritmo, *pitch* (afinação) e timbre,

elementos essenciais para compreender a influência da música nas emoções do ouvinte (VOLOKHIN; AGICHTEIN, 2018).

A API do Spotify fornece uma análise detalhada das características de áudio das faixas (SPOTIFY, 2024). Estas características incluem:

- *Acousticness*: Medida de quão acústica é uma faixa.
- *Danceability*: Quão adequada uma faixa é para dançar.
- *Duration*: Duração da faixa em milissegundos.
- *Energy*: Medida de intensidade e atividade de uma faixa.
- *Instrumentalness*: Previsão de se uma faixa contém vocais.
- *Liveness*: Presença de público na gravação.
- *Loudness*: Volume geral da faixa em decibéis.
- *Speechiness*: Presença de palavras faladas em uma faixa.
- *Valence*: Positividade musical transmitida por uma faixa.
- *Tempo*: Tempo estimado da faixa em batidas por minuto.
- *Key*: Tonalidade em que a faixa está.
- *Mode*: Modalidade (maior ou menor) da faixa.

Essas características permitem análises sobre como diferentes aspectos da música podem influenciar as emoções dos ouvintes, fornecendo um recurso para pesquisa e desenvolvimento de aplicações relacionadas à música e emoção.

2.4 Modelos de Aprendizagem de Máquina

A aprendizagem de máquina, uma vertente fundamental da inteligência artificial, tem desempenhado um papel revolucionário em diversos campos, desde o reconhecimento de padrões até a tomada de decisões complexas (M. BISHOP, 2006). Esta área da ciência da computação é dedicada ao desenvolvimento de algoritmos que permitem que as máquinas aprendam e melhorem suas performances baseadas em experiências passadas, sem serem explicitamente programadas para cada tarefa específica(GOODFELLOW; BENGIO; COURVILLE, 2016).

Existem diferentes tipos de aprendizagem de máquina, cada um com características e aplicações específicas:

Aprendizagem Supervisionada: Neste tipo, o algoritmo é treinado em um conjunto de dados rotulado, aprendendo a mapear entradas para saídas. Este é o tipo mais comum de aprendizagem de máquina e é amplamente utilizado em aplicações como classificação de e-mails em spam e não spam (JAMES et al., 2013).

Aprendizagem Não Supervisionada: Aqui, o algoritmo é treinado usando um conjunto de dados que não possui rótulos, explorando a estrutura dos dados para extrair padrões ou insights (M. BISHOP, 2006).

Aprendizagem por Reforço: Neste caso, o algoritmo aprende a tomar decisões sequenciais, descobrindo através de tentativa e erro quais ações maximizam uma recompensa ou minimizam uma penalidade ao longo do tempo (SUTTON; BARTO, 2018).

Esses métodos possibilitam a análise de dados complexos, fornecendo ferramentas para descobrir padrões ocultos, fazer previsões precisas e melhorar a tomada de decisões baseadas em dados (RUSSELL; PETER, 2020).

2.4.1 Modelos utilizados na tarefa de classificação do MER

K-Nearest Neighbors (KNN): Este é um algoritmo simples e eficaz, utilizado para problemas de classificação e regressão. No contexto da monografia, foi empregado para classificar emoções em músicas. Matematicamente, a classificação pelo KNN é baseada na distância entre os pontos de dados. Seja x um ponto a ser classificado, e X o conjunto de dados de treinamento com rótulos conhecidos. O KNN identifica os k pontos mais próximos de x em X e atribui a x a classe mais frequente entre esses k pontos (COVER; HART, 1967).

Support Vector Machine (SVM): O SVM é um modelo robusto e versátil, adequado para a classificação linear e não linear. Funciona construindo um hiperplano ou conjunto de hiperplanos em um espaço de alta dimensão, que pode ser usado para classificação ou regressão. Matematicamente, para um conjunto de dados de

treinamento $(x_1, y_1), \dots, (x_n, y_n)$ onde y_i é a classe de x_i , o SVM procura o hiperplano que maximiza a margem entre as classes. A solução envolve resolver um problema de otimização quadrática (CORTES; VAPNIK; SAITTA, 1995).

Naive Bayes: Este modelo é baseado no teorema de Bayes e é particularmente eficiente para grandes volumes de dados. Na monografia, foi utilizado para classificar emoções musicais. Matematicamente, para uma dada variável de classe y e um vetor de atributo x_1, \dots, x_n , o teorema de Bayes é usado para calcular a probabilidade $(P(y|x_1, \dots, x_n))$, que é proporcional a $P(x_1, \dots, x_n | y)P(y)$. Na prática, o modelo assume que as variáveis de entrada são independentes entre si, o que simplifica o cálculo $P(x_1, \dots, x_n | y)$. (RISH; RISH, 2001).

Random Forest: Este modelo é um ensemble de árvores de decisão. Funciona criando múltiplas árvores de decisão durante o treinamento e produzindo a classe que é a moda das classes (classificação) ou a média das previsões (regressão) das árvores individuais. Matematicamente, cada árvore é construída a partir de uma amostra aleatória dos dados de treinamento. A decisão final é uma média ou uma votação majoritária das decisões de cada árvore (BREIMAN, 2001).

Gradient Boosting: Este modelo constrói um ensemble de árvores de decisão de maneira sequencial, onde cada árvore subsequente tenta corrigir os erros da árvore anterior. É formulado como um problema de otimização, onde a função de perda a ser minimizada é definida com base na diferença entre as previsões atuais e os valores reais. A cada iteração, um modelo é ajustado ao gradiente negativo da função de perda (FRIEDMAN, 2001).

XGBoost: Uma extensão do Gradient Boosting, o XGBoost é conhecido por sua eficiência e desempenho. Ele se diferencia pela sua capacidade de fazer ajustes rápidos e precisos. Matematicamente, o XGBoost adiciona regularização ao modelo de Gradient Boosting, o que ajuda a prevenir o overfitting. O modelo é otimizado para ser altamente eficiente e escalável (CHEN; GUESTRIN, 2016).

3 DESENVOLVIMENTO

Nessa etapa, a teoria e a prática se unem no contexto da Engenharia de Dados e Big Data. Aqui, é feita a ponte entre os conceitos fundamentais discutidos anteriormente e a sua aplicação prática, necessária para a compreensão e o manejo eficaz dos dados no estudo das emoções na música.

O experimento começa com uma exploração detalhada do conjunto de dados. Esta seção delinea a origem e as características dos dados escolhidos, e explica a razão deles serem essenciais para este estudo. A preparação dos dados, incluindo etapas de limpeza e normalização é abordado, garantindo que a base para as análises subsequentes seja sólida e confiável.

A Análise Exploratória de Dados (EDA) é o primeiro passo do entendimento dos dados, onde padrões e tendências são identificados, proporcionando uma compreensão melhor dos dados. É feita a modelagem de dados e os modelos de aprendizado de máquina selecionados são cuidadosamente descritos, justificando a escolha de cada um e detalhando sua implementação.

O processo de treinamento e validação desses modelos é feito e os resultados analisados. Esta seção analisa o desempenho dos modelos, das técnicas e abordagens adotadas. O capítulo, apresenta os desafios enfrentados e as limitações encontradas, tem-se os resultados obtidos e contribui com uma compreensão mais detalhada do campo em estudo.

Com este capítulo, aspira-se não só a atender aos requisitos acadêmicos, mas também a fazer uma contribuição significativa para a área de MER, refletindo as habilidades adquiridas no curso de especialização e demonstrando a aplicabilidade prática dos conceitos de Engenharia de Dados e Big Data.

3.1 Descrição do Conjunto de Dados

Os dados selecionados para o experimento, é fruto do trabalho realizado por (PANDA et al., 2021) e fornece uma análise detalhada da API do Spotify, enfatizando sua aplicabilidade no campo do Reconhecimento de Emoções Musicais (MER). O conjunto

de dados gerado nesse trabalho foi composto por 900 faixas de áudio com a marcação dos quadrantes de Russell e características de áudio obtidas pelo “*The AllMusic*” de forma balanceada, tanto em termos de quadrantes emocionais quanto de gêneros musicais.

Descreve-se ainda em (PANDA et al., 2021) que a correspondência entre as tags de humor do AllMusic e as normas de valência e excitação para palavras em inglês de Warriner (PANDA; MALHEIRO; PAIVA, 2020) foi usada para transformar as tags em anotações de quadrantes, seguindo uma validação manual feita por voluntários. Esta abordagem assegurou que as anotações refletissem as emoções expressas pelas faixas de maneira equilibrada e representativa. A construção balanceada deste conjunto de dados visavam garantir a validade e a aplicabilidade dos resultados da pesquisa na identificação automática de emoções em música.

A Figura 4 apresenta a uma linha de exemplo do conjunto de dados original

Figura 4 - Linha do Conjunto de Dados Original

Song	MT0000004637
Artist	Charlie Poole
Title	Bulldog Down in Sunny Tennessee
Quadrant	Q3
PQuad	0.666667
MoodsTotal	3
Moods	3
MoodsFoundStr	circular; greasy; messy
MoodsStr	Circular; Greasy; Messy
MoodsStrSplit	Circular; Greasy; Messy
Genres	2
GenresStr	Country; International
Sample	1
SampleURL	http://rovimusic.rovicorp.com/playback.mp3?c=l...

Fonte: Autor (2023)

3.1.1 Processamento e Geração do Conjunto de dados

Para a preparação específica do conjunto de dados a ser utilizado neste projeto, foi desenvolvida uma rotina em Python. Esta rotina foi responsável por gerar o conjunto de dados da principal fonte de dados para o experimento. Este procedimento permite personalizar e otimizar os dados para as necessidades específicas deste estudo.

A rotina em Python executa um procedimento de extração e análise de dados musicais através da API do Spotify, utilizando as bibliotecas Python *requests* e *pandas* e segue a seguinte ordem:

1. Autenticação na API do Spotify para obtenção de um token de acesso, utilizando credenciais de cliente armazenadas de maneira segura.
2. Definição de uma função para a recuperação de características de áudio de faixas específicas, baseando-se nos nomes das faixas e dos artistas.
3. Extração de informações de um conjunto de dados carregado de um arquivo CSV, incluindo título da faixa, nome do artista, gênero, entre outros.
4. Busca das características de áudio de cada faixa na API do Spotify e armazenamento dos resultados em um DataFrame dos pandas.
5. Tratamento de casos em que as faixas não são encontradas, com armazenamento desses casos em um DataFrame separado.
6. Salvamento dos DataFrames - um com as características das faixas encontradas e outro com as faixas não encontradas para análises posteriores ou uso em aplicações de dados.

Os detalhes estão presentes no ANEXO A.

3.1.2 Características do Conjunto de Dados e sua Relevância para o MER

O conjunto de dados gerado, contém variáveis necessárias para a análise de emoções em músicas. Estas variáveis incluem, mas não se limitam a, características de áudio como ritmo, timbre, tonalidade, entre outras, que são fornecidas pela API do Spotify.

A relevância deste conjunto de dados para o estudo de emoções musicais reside no fato de que ele oferece informações sobre as propriedades sonoras das faixas musicais. Essas propriedades são fundamentais para o entendimento de como diferentes aspectos da música podem influenciar as emoções dos ouvintes. Além disso, a possibilidade de comparar esses dados com benchmarks de MER de estado-da-arte permite uma análise mais contextualizada dentro do campo do Reconhecimento de Emoções Musicais.

3.1.3 Tamanho e Diversidade do Conjunto de Dados

O conjunto de dados é composto por 706 registros, com uma diversidade de gêneros musicais e estilos. Essa diversidade permite garantir a abrangência e a robustez das análises, permitindo que os resultados sejam generalizáveis para um amplo espectro musical.

3.2 Pré-processamento e Limpeza dos Dados

O pré-processamento e a limpeza dos dados é necessária principalmente na área do Reconhecimento de Emoções Musicais (MER), pois a precisão e a qualidade dos dados são importantes para um resultado mais assertivo.

3.2.1 Verificação e Tratamento de Valores Ausentes

Inicialmente, foi realizado uma exploração dos dados para identificar valores ausentes ou nulos. Valores ausentes podem distorcer os resultados analíticos e, portanto, devem ser tratados adequadamente.

Figura 5 - Evidência da inexistência de valores ausentes no Dataset

Valores Ausentes:	
Nome da Faixa	0
Cantor ou Compositor	0
Genero	0
Quadrante	0
acousticness	0
danceability	0
energy	0
instrumentalness	0
key	0
liveness	0
loudness	0
mode	0
speechiness	0
tempo	0
time_signature	0
valence	0

Fonte: Autor (2023)

Como pode ser observado na Figura 5, não foram encontrados valores ausentes para o conjunto de dados, além dos tipos de dados estarem adequados para as próximas etapas conforme Figura 6.

Figura 6 - Evidência de tipos de dados adequados

Tipos de Dados:	
Nome da Faixa	object
Cantor ou Compositor	object
Genero	object
Quadrante	object
acousticness	float64
danceability	float64
energy	float64
instrumentalness	float64
key	int64
liveness	float64
loudness	float64
mode	int64
speechiness	float64
tempo	float64
time_signature	int64
valence	float64
dtype:	object

Fonte: Autor (2023)

3.2.2 Normalização dos Dados.

A normalização dos dados foi aplicada para assegurar que as características de áudio estivessem em uma escala comum. Para isso foi utilizada a classe “*StandardScaler*” da biblioteca do “*Sklearn*”.

Para cada característica do conjunto de dados, é calculado a média e o desvio padrão. Cada valor de característica x é transformado usando a fórmula:

$$z = \frac{x - \mu}{\sigma}$$

Asas x é o valor original, μ é a média e σ é o desvio padrão. Essa normalização é conhecida como Z-score.

Após a transformação, cada característica tem uma média aproximada de 0 e um desvio padrão de 1 e significa que os valores da característica são comparáveis em uma escala comum.

Isso é particularmente importante quando diferentes variáveis possuem escalas distintas, o que pode levar a um desequilíbrio na influência dessas variáveis nos modelos de aprendizado de máquina.

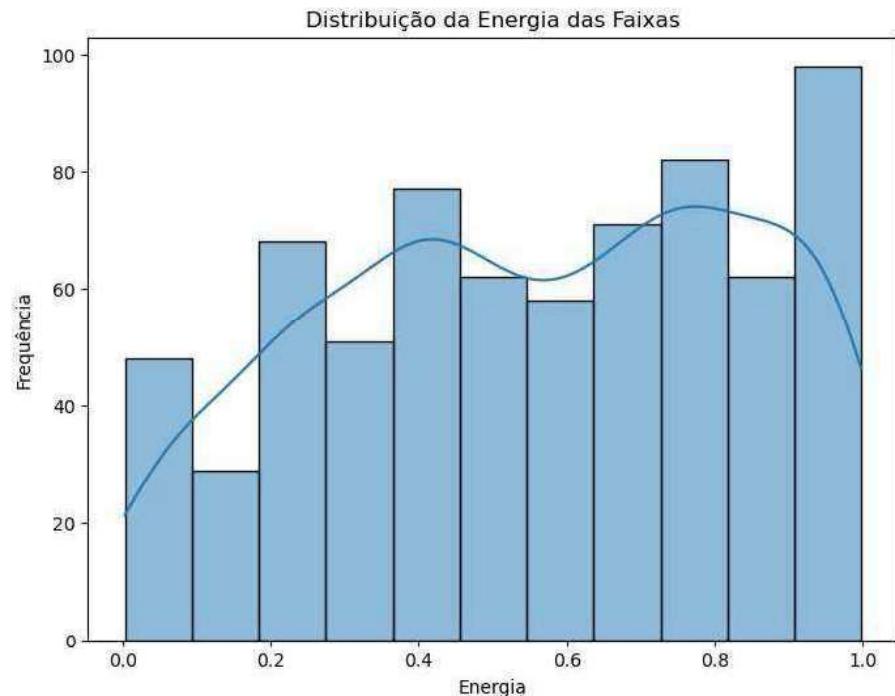
3.3 Análise Exploratória de Dados (EDA)

A Análise Exploratória de Dados (EDA) permite a compreensão do conjunto de dados, permitindo identificar padrões, tendências e possíveis anomalias que podem influenciar no resultado da análise final. Esta fase é caracterizada pelo uso de técnicas estatísticas e visualizações de dados para explorar as propriedades e relações intrínsecas nos dados.

Os detalhes dessa etapa, pode ser consultado no Apêndice A. Para exemplificar, há uma amostragem dos seguintes elementos:

- Histograma para a variável ‘energy’ na Figura 7

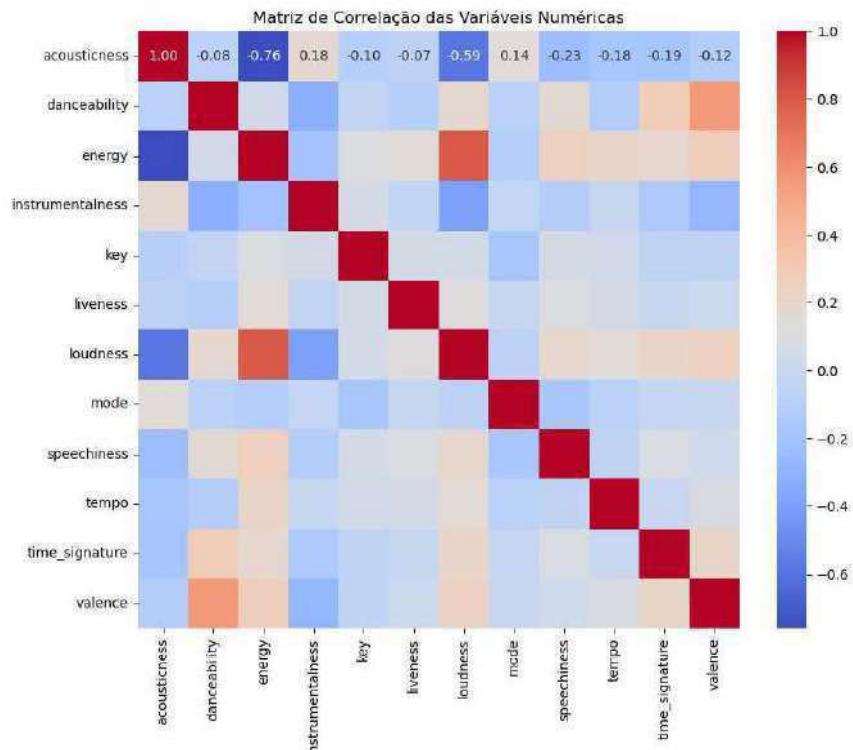
Figura 7 - Histograma da variável energy



Fonte: Autor (2023)

- Matriz de correlação das variáveis numéricas na Figura 8

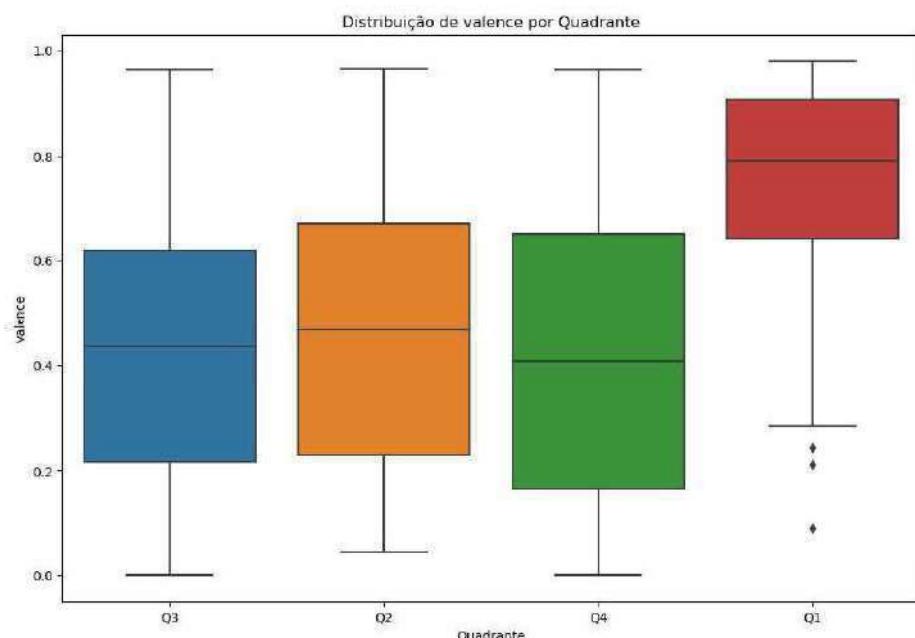
Figura 8 - Correlação das variáveis numéricas



Fonte: Autor (2023)

- Distribuição da variável ‘valence’ por “Quadrante” conforme observado na Figura 9

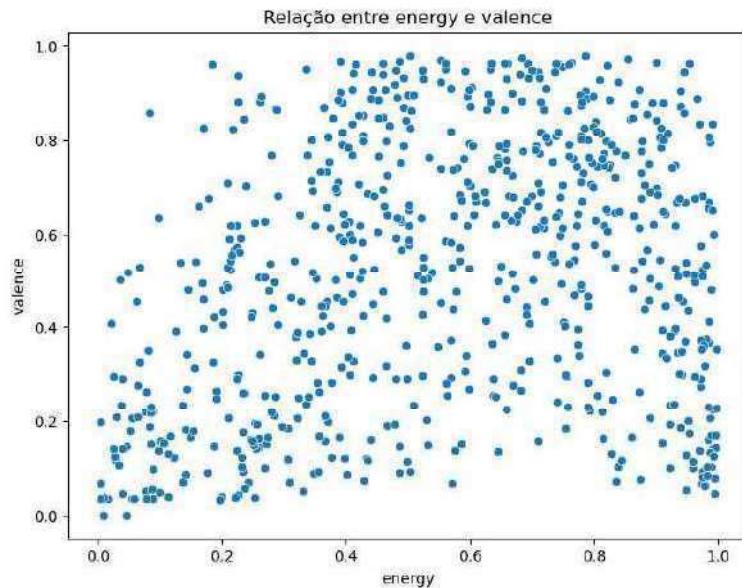
Figura 9 - Distribuição da variável valence por Quadrante de Russell



Fonte: Autor (2023)

- Gráfico de Dispersão para explorar relações entre o par de variáveis ‘energy’ e ‘valence’ na Figura 10

Figura 10 - Relação de dispersão entre *energy* e *valence*



Fonte: Autor (2023)

Esta etapa ajudou a identificar padrões de distribuição, como normalidade ou assimetria, e destacar quaisquer *outliers* ou anomalias.

3.4 Modelagem e Algoritmos de Aprendizado de Máquina com Preparação de Dados e Otimização

A escolha dos modelos foi baseada em sua adequação para o objetivo da monografia, que é analisar e prever emoções musicais com base em características de áudio extraídas de faixas musicais.

3.4.1 Preparação dos Dados

Antes da aplicação dos modelos de aprendizado de máquina, foi realizada uma preparação dos dados. A rotina em Python desenvolvida para este fim incluiu:

1. Carregamento do conjunto de dados.

2. Separação dos dados em características de áudio (*feature*) e *target* (variável 'Quadrante').
3. Codificação da variável alvo categórica utilizando *LabelEncoder*.
4. Divisão dos dados em conjuntos de treinamento e teste.
5. Pré-processamento com *Pipeline* incluindo imputação de valores ausentes e normalização das variáveis numéricas.
6. Exportação do pré-processador e do *LabelEncoder* para uso futuro.

Esses passos garantiram que os dados estivessem limpos, normalizados e prontos para serem aplicados aos modelos de aprendizado de máquina e estão disponíveis em detalhe no ANEXO B

3.4.2 Modelagem e Otimização de Hiperparâmetros

Para cada um dos modelos selecionados, foi estabelecida uma grade de hiper parâmetros para ser utilizada no *GridSearchCV*. Este método permite uma busca exaustiva sobre os parâmetros especificados, utilizando validação cruzada para determinar a melhor combinação de hiper parâmetros. Os modelos explorados incluíram:

1. *K-Nearest Neighbors (KNN)*:

- **Vantagens:** Simplicidade e eficácia para classificação baseada em proximidade.
- **Limitações:** Sensível a dados não normalizados e de alta dimensionalidade; desempenho pode ser afetado em grandes conjuntos de dados.

2. *Support Vector Machine (SVM)*:

- **Vantagens:** Eficiente em espaços de alta dimensão; eficaz na separação de classes.
- **Limitações:** Seleção de kernel e parâmetros pode ser complexa; não ideal para conjunto de dados muito grandes.

3. ***Naive Bayes:***

- **Vantagens:** Simples, rápido e eficaz para conjunto de dados de tamanho moderado.
- **Limitações:** Suposição de independência entre variáveis pode ser um ponto fraco; performance pode ser limitada em comparação com outros modelos mais complexos.

4. ***Random Forest:***

- **Vantagens:** Bom para lidar com muitas variáveis; robusto contra *overfitting*.
- **Limitações:** Pode ser lento para treinar; interpretabilidade dos resultados pode ser desafiadora.

5. ***Gradient Boosting:***

- **Vantagens:** Alta precisão; bom para lidar com diferentes tipos de dados.
- **Limitações:** Mais suscetível a *overfitting*; o tempo de treinamento pode ser longo.

6. ***XGBoost:***

- **Vantagens:** Rápido e eficiente; frequentemente supera outros modelos.
- **Limitações:** Pode ser desafiador ajustar; *overfitting* é uma preocupação se não for bem regulado.

A escolha e a aplicação destes modelos foram realizadas com uma compreensão das suas vantagens e desvantagens. A otimização de hiper parâmetros e a validação cruzada reforçaram a adequação dos modelos ao contexto específico do estudo, com foco em classificar emoções musicais com base em características de áudio.

3.5 Avaliação dos Modelos

Os resultados de desempenho de cada modelo, incluindo métricas são apresentadas para que se possa fazer a análise comparativa entre eles e identificar qual obteve melhor desempenho no contexto do estudo.

3.5.1 K-Nearest Neighbors (KNN)

O modelo KNN foi otimizado através do ***GridSearchCV***, considerando um conjunto de hiperparâmetros, incluindo o número de vizinhos (***n_neighbors***), o método de ponderação (***weights***) e a métrica de distância (***metric***). Os melhores parâmetros identificados foram '***metric***: 'manhattan', '***n_neighbors***': 11, e '***weights***': 'uniform', com uma pontuação de validação cruzada (acurácia) de aproximadamente 0.608.

Os resultados no conjunto de teste foram os seguintes:

Precisão e Recall por Classe: Classe 0: Teve a maior taxa de recall (0.76) mas com precisão relativamente baixa (0.46), indicando uma tendência do modelo em identificar corretamente esta classe, embora com um número significativo de falsos positivos. Classe 1: Apresentou um equilíbrio melhor entre precisão (0.74) e recall (0.68). Classe 2 e 3: Tiveram tanto precisão quanto recall mais baixos, sugerindo uma dificuldade do modelo em identificar corretamente estas emoções.

F1-Score: As variações nos valores de F1-Score refletem a capacidade inconsistente do modelo em equilibrar precisão e recall entre as classes, com melhor desempenho na Classe 1 (0.71) e pior na Classe 2 (0.41).

Acurácia Geral: A acurácia total de 0.542 mostra um desempenho moderado, indicando que o modelo teve sucesso em algumas áreas, mas também apresentou limitações.

O modelo KNN, com os parâmetros otimizados, demonstrou ser eficaz em algumas áreas, mas enfrentou desafios na classificação consistente de todas as emoções musicais. A escolha da métrica de Manhattan e o número de vizinhos indicam um enfoque na proximidade das características, o que pode ter beneficiado algumas

classes em detrimento de outras. A variação nas métricas de desempenho entre as classes sugere que o KNN, apesar de ser um modelo valioso, pode não ser suficiente por si só para capturar a complexidade total das emoções musicais.

3.5.2 Gaussian Naive Bayes

O modelo Gaussian Naive Bayes foi otimizado utilizando `GridSearchCV` para explorar um conjunto maiores de valores para o parâmetro `var_smoothing`. Este parâmetro auxilia no ajuste da suavização das probabilidades no modelo. O melhor `var_smoothing` encontrado foi **1e-09**, com uma pontuação de validação cruzada em acurácia de cerca de 0.583.

O modelo foi avaliado no conjunto de teste, resultando nas seguintes métricas:

Precisão e Recall por Classe: Classe 0: Precisão de 0.43 e recall de 0.62, indicando uma tendência do modelo em identificar corretamente esta classe, mas com uma taxa considerável de falsos positivos. Classe 1: Melhor equilíbrio entre precisão (0.67) e recall (0.76). Classe 2 e 3: Precisão e recall moderados, sugerindo desafios na classificação precisa destas emoções.

F1-Score: Os valores de F1-Score variam entre as classes, com o melhor desempenho na Classe 1 (0.71) e desempenhos mais baixos nas Classes 2 e 3.

Acurácia Geral: A acurácia total de 0.535 mostra um desempenho geral moderado.

O *Gaussian Naive Bayes*, apesar de sua simplicidade e eficiência computacional, mostrou limitações na classificação das emoções musicais. A variação nas métricas de precisão e *recall* por classe sugere que o modelo teve dificuldades em capturar a complexidade e as diferenças sutis na classificação das emoções. A escolha do `var_smoothing` de **1e-09** indica a necessidade de uma suavização mínima, refletindo as características dos dados.

3.5.3 *Support Vector Machine (SVM)*

O modelo SVM foi otimizado usando **GridSearchCV** para avaliar uma variedade de hiperparâmetros, incluindo 'C' (parâmetro de regularização), 'gamma' (coeficiente para os kernels 'rbf', 'poly' e 'sigmoid'), 'kernel' (tipo de kernel usado no algoritmo) e 'degree' (grau do polinômio para o kernel 'poly'). O melhor conjunto de parâmetros encontrado foi 'C': 0.1, 'degree': 2, 'gamma': 'scale', 'kernel': 'linear', com uma pontuação de validação cruzada (acurácia) de aproximadamente 0.612.

A avaliação do modelo no conjunto de teste produziu as seguintes métricas:

Precisão e Recall por Classe: Classe 0: Precisão de 0.51 e recall de 0.65, mostrando uma capacidade moderada do modelo em identificar corretamente esta classe. Classe 1: Precisão de 0.72 e recall de 0.82, indicando um bom desempenho na classificação desta classe. Classe 2 e 3: Precisões e recalls mais baixos em comparação, sugerindo dificuldades na identificação correta destas emoções.

F1-Score: F1-Scores variados, com a Classe 1 apresentando o melhor desempenho (0.77) e a Classe 3 o mais baixo (0.39).

Acurácia Geral: Uma acurácia de 0.563 reflete um desempenho moderadamente bom, com o modelo exibindo capacidade razoável em classificar as emoções musicais.

O modelo SVM com kernel linear e os parâmetros otimizados mostrou ser eficaz em algumas classes, mas enfrentou desafios em outras. A escolha do kernel linear sugere que uma fronteira de decisão linear foi adequada para algumas das classes de emoção, enquanto outras poderiam se beneficiar de abordagens mais complexas. As diferenças nas métricas de desempenho entre as classes sugerem que o SVM, apesar de ser uma ferramenta poderosa, pode não ser totalmente adequado para capturar a complexidade de todas as emoções musicais.

3.5.4 Random Forest

O modelo *Random Forest* foi otimizado com o uso do **GridSearchCV**, considerando uma variedade de hiperparâmetros como 'n_estimators' (número de árvores), 'max_depth' (profundidade máxima da árvore) e 'min_samples_split' (número mínimo de amostras necessárias para dividir um nó). Os melhores parâmetros identificados foram '**max_depth**: 10, '**min_samples_split**: 6, e '**n_estimators**: 400, alcançando uma pontuação de validação cruzada em acurácia de aproximadamente 0.619.

A avaliação no conjunto de teste revelou as seguintes métricas:

Precisão e Recall por Classe: Classe 0: Precisão de 0.47 e recall de 0.62, sugerindo uma capacidade razoável do modelo em identificar esta classe com uma taxa de falsos positivos moderada. Classe 1: A melhor precisão (0.75) e recall (0.71), indicando um bom desempenho nesta classe. Classe 2 e 3: Precisões e recalls mais baixos, apontando para desafios na identificação correta destas emoções.

F1-Score: As variações de F1-Score refletem a capacidade variada do modelo em equilibrar precisão e recall entre as diferentes classes.

Acurácia Geral: Uma acurácia de 0.556 mostra um desempenho geral moderado.

O modelo Random Forest, com os parâmetros otimizados, exibiu um desempenho considerável, especialmente na Classe 1. A escolha de uma profundidade máxima de 10 e um mínimo de 6 amostras para divisão de nós, juntamente com um número robusto de 400 árvores, sugere um modelo equilibrado que evita *overfitting* enquanto mantém uma boa capacidade de generalização. No entanto, as métricas de desempenho por classe indicam que o modelo ainda enfrenta desafios em capturar com precisão a complexidade das emoções musicais em algumas classes. Este resultado sugere que o *Random Forest*, apesar de suas qualidades, pode beneficiar-se de abordagens complementares ou da integração com outros modelos para uma análise mais abrangente.

3.5.5 Gradient Boosting

O modelo Gradient Boosting foi otimizado usando **GridSearchCV**, explorando um conjunto de hiperparâmetros como 'n_estimators' (número de árvores), 'learning_rate' (taxa de aprendizado) e 'max_depth' (profundidade máxima da árvore). Os melhores parâmetros identificados foram '**learning_rate**: 0.01, **'max_depth'**: 4, e '**n_estimators**: 50, com uma pontuação de validação cruzada (acurácia) de aproximadamente 0.606.

A avaliação do modelo no conjunto de teste produziu as seguintes métricas:

Precisão e Recall por Classe: Classe 0: Precisão de 0.45 e recall de 0.65, indicando uma capacidade razoável do modelo em identificar corretamente esta classe. Classe 1: A melhor precisão (0.77) e recall (0.68), sugerindo um desempenho mais forte na classificação desta classe. Classe 2 e 3: Precisões e recalls mais baixos, apontando para desafios na identificação precisa destas emoções.

F1-Score: Variações nos F1-Scores entre as classes, com o melhor desempenho na Classe 1 (0.72) e o mais baixo na Classe 2 (0.32).

Acurácia Geral: Uma acurácia de 0.521 mostra um desempenho geral moderado.

O modelo *Gradient Boosting*, com uma taxa de aprendizado baixa e um número relativamente pequeno de árvores, mostrou ser eficiente em algumas classes, mas enfrentou dificuldades em outras. A escolha de parâmetros sugere uma abordagem cautelosa para evitar *overfitting*, mas pode ter limitado o potencial do modelo em capturar a complexidade total das emoções musicais. As métricas de desempenho variadas por classe indicam que o Gradient Boosting, embora seja uma ferramenta poderosa, pode precisar de ajustes adicionais ou de ser complementado com outros modelos para uma análise mais abrangente das emoções musicais.

3.5.6 XGBoost

O modelo XGBoost foi otimizado usando **GridSearchCV**, avaliando uma variedade de hiperparâmetros, incluindo 'n_estimators' (número de árvores), 'learning_rate' (taxa de

aprendizado) e 'max_depth' (profundidade máxima da árvore). Os melhores parâmetros identificados foram '**learning_rate**': 0.01, '**max_depthn_estimators**': 50, alcançando uma pontuação de validação cruzada em acurácia de aproximadamente 0.622.

A avaliação do modelo no conjunto de teste produziu as seguintes métricas:

Precisão e Recall por Classe: Classe 0: Precisão de 0.44 e recall de 0.59, indicando uma capacidade moderada do modelo em identificar corretamente esta classe. Classe 1: A melhor precisão (0.76) e recall (0.74), sugerindo um desempenho forte na classificação desta classe. Classe 2 e 3: Precisões e recalls mais baixos, apontando para desafios na identificação precisa destas emoções.

F1-Score: Variações nos F1-Scores entre as classes, com o melhor desempenho na Classe 1 (0.75) e o mais baixo na Classe 2 (0.34).

Acurácia Geral: Uma acurácia de 0.507 reflete um desempenho geral moderado.

O modelo XGBoost, com uma taxa de aprendizado baixa e um número limitado de árvores, mostrou ser eficaz em algumas classes, mas enfrentou dificuldades em outras. A escolha dos parâmetros sugere uma abordagem cautelosa para evitar *overfitting*, mas isso pode ter restringido o potencial do modelo em capturar a complexidade das emoções musicais. As métricas de desempenho variadas por classe indicam que o XGBoost, apesar de ser uma ferramenta avançada de aprendizado de máquina, pode precisar de ajustes adicionais ou combinação com outros métodos para abordar todas as diferenças sutis das emoções musicais.

3.6 Classificação de Novas Músicas Utilizando Modelos de Aprendizado de Máquina

Neste segmento do capítulo, destaca-se a aplicação prática dos modelos de aprendizado de máquina desenvolvidos. Para ilustrar a funcionalidade dos modelos, foi implementada uma rotina em Python capaz de classificar novas músicas com base em suas características de áudio obtidas através da API do Spotify. Esta rotina permite a aplicação dos modelos em cenários reais, como a classificação automática de

músicas em plataformas de streaming. Os detalhes podem ser verificados no Apêndice A.

3.6.1 Detalhes Técnicos da Rotina de Classificação

A rotina desenvolvida realiza os seguintes passos:

- **Autenticação e Acesso à API do Spotify:** Utiliza-se as credenciais do Spotify para obter um token de acesso, permitindo a interação com a API.
- **Busca de Informações da Música:** Através da API, busca-se informações detalhadas de uma faixa específica, incluindo o nome da música, artista, álbum e outras informações relevantes.
- **Extração de Características de Áudio:** Obtém-se as características de áudio da música, como danceabilidade, energia, valência e outras, essenciais para o processo de classificação.
- **Preparação e Normalização dos Dados:** Os dados são preparados e normalizados de acordo com o mesmo procedimento aplicado ao conjunto de dados original, garantindo a consistência na entrada dos modelos.
- **Carregamento dos Modelos Treinados e Classificação:** Carregam-se os modelos treinados (KNN, SVM, Naive Bayes, Random Forest, Gradient Boosting e XGBoost) e utilizam-se para prever o quadrante emocional da música.
- **Resultados da Classificação:** Exibem-se os quadrantes emocionais previstos por cada modelo, oferecendo uma visão comparativa das previsões.

3.6.2 Exemplo Prático de Classificação

A rotina foi aplicada à música "Halo" de Beyoncé. Cada modelo de aprendizado de máquina forneceu uma previsão para o quadrante emocional da música. Os resultados foram:

- KNN: Q2 - Baixa Valência, Alta Ativação (Tensa, Nervosa)
- SVM: Q3 - Baixa Valência, Baixa Ativação (Triste, Melancólica)

- Naive Bayes: Q1 - Alta Valência, Alta Ativação (Alegre, Energética)
- Random Forest: Q3 - Baixa Valência, Baixa Ativação (Triste, Melancólica)
- Gradient Boosting: Q4 - Alta Valência, Baixa Ativação (Calma, Relaxante)
- XGBoost: Q3 - Baixa Valência, Baixa Ativação (Triste, Melancólica)

Esses resultados destacam as variações nas previsões dos diferentes modelos e a complexidade inerente à tarefa de classificação de emoções em músicas. A variabilidade nas previsões sublinha a necessidade de uma análise mais aprofundada e talvez de uma abordagem híbrida ou enriquecida para melhorar a precisão.

3.6.3 Implicações e Potenciais Aplicações

A implementação desta rotina de classificação demonstra o potencial prático dos modelos de aprendizado de máquina no campo do Reconhecimento de Emoções Musicais (MER). As aplicações incluem:

- Recomendação Personalizada de Músicas: Plataformas de streaming podem utilizar estes modelos para recomendar músicas que correspondam ao estado emocional desejado dos ouvintes.
- Análise de Tendências Musicais: A indústria da música pode se beneficiar da análise automatizada para entender as preferências emocionais dos ouvintes em diferentes contextos.
- Terapia Musical: A rotina pode ser adaptada para selecionar músicas que auxiliem em tratamentos terapêuticos, proporcionando uma experiência personalizada e eficaz.

A inclusão desta rotina realça a aplicabilidade prática dos conceitos e técnicas estudados. Demonstra-se assim, não apenas a viabilidade teórica, mas também a utilidade prática dos modelos de aprendizado de máquina em contextos reais relacionados à música e às emoções humanas.

Este avanço abre portas para futuras pesquisas e desenvolvimentos, visando aprimorar ainda mais a precisão e a eficácia dos modelos na classificação de emoções

em músicas e na aplicação prática em diversas áreas relacionadas à música e bem-estar humano.

4 ANÁLISE DOS RESULTADOS

4.1 Comparação dos Modelos de Aprendizagem de Máquina

Neste capítulo, será comparado os modelos de aprendizado de máquina utilizados para classificar emoções musicais com base nas características de áudio do Spotify. Os modelos *incluíram K-Nearest Neighbors (KNN), Naive Bayes (Gaussian), Support Vector Machine (SVM), Random Forest, Gradient Boosting e XGBoost.*

Tabela 1 - Comparação dos Modelos de Aprendizagem de Máquina

Modelo	Melhor Acurácia de Validação Cruzada	Acurácia no Conjunto de Teste	Melhores Parâmetros
KNN	0.6082	0.5423	Metric: manhattan, n_neighbors: 11, weights: uniform
Naive Bayes	0.5834	0.5352	var_smoothing: 1e-09
SVM	0.6118	0.5634	C: 0.1, degree: 2, gamma: scale, kernel: linear
Random Forest	0.6188	0.5563	max_depth: 10, min_samples_split: 6, n_estimators: 400
Gradient Boosting	0.6064	0.5211	learning_rate: 0.01, max_depth: 4, n_estimators: 50
XGBoost	0.6224	0.5070	learning_rate: 0.01, max_depth: 4, n_estimators: 50

Fonte: Autor (2023)

4.2 Discussão dos Resultados

A comparação dos modelos revela que cada um tem seus pontos fortes e fracos na classificação das emoções musicais. Enquanto alguns modelos, como o SVM e o *Random Forest*, mostraram eficácia em classes específicas, outros enfrentaram dificuldades em abranger um conjunto mais completo de emoções. Isso sugere que a complexidade das emoções musicais pode exigir abordagens mais integradas ou híbridas, combinando diferentes modelos para melhor capturar a variedade e diferenças sutis das emoções musicais.

Os resultados indicam a importância de uma seleção cuidadosa de hiperparâmetros e estratégias de modelagem. Por exemplo, o equilíbrio entre *overfitting* e *underfitting* foram importantes em modelos como *Gradient Boosting* e *XGBoost*, onde a escolha da taxa de aprendizado e o número de estimadores tiveram um impacto no desempenho.

4.3 Conclusão sobre os resultados

A análise dos modelos de aprendizado de máquina demonstra que, embora cada modelo tenha suas qualidades, nenhum deles é superior em todos os aspectos para a classificação das emoções musicais. Isso destaca a necessidade de estratégias de modelagem mais complexas e combinação de diferentes abordagens para melhorar a precisão na classificação das emoções musicais. A escolha do modelo ideal pode depender das especificidades das emoções que estão sendo analisadas, bem como das características particulares dos dados de áudio disponíveis.

5 CONCLUSÃO E TRABALHOS FUTUROS

5.1 Conclusão

Este estudo analisou a eficácia das características de áudio do Spotify na classificação de emoções musicais, utilizando uma variedade de modelos de aprendizado de máquina. A análise revelou que cada modelo possui pontos fortes e limitações distintas ao abordar a complexidade das emoções musicais. Modelos como o SVM e Random Forest mostraram eficiência em classes específicas, enquanto o Gradient Boosting e XGBoost, apesar de avançados, apresentaram desafios na generalização para todas as emoções.

As características de áudio do Spotify fornecem uma base sólida para o Reconhecimento de Emoções Musicais (MER), mas a eficácia na classificação depende significativamente do modelo de aprendizado de máquina escolhido e da configuração de seus parâmetros. A complexidade inerente às emoções musicais sugere que abordagens mais integradas ou híbridas podem ser necessárias para capturar com precisão toda a gama de emoções.

5.2 Direções para Pesquisas Futuras

Com base nas descobertas deste estudo, propõem-se as seguintes direções para pesquisas futuras:

5.2.1 Exploração de Modelos Híbridos:

Investigar a combinação de diferentes modelos de aprendizado de máquina para aproveitar as forças de cada um e melhorar a classificação geral das emoções musicais.

5.2.2 Uso de Técnicas Avançadas de Aprendizado Profundo:

Explorar redes neurais profundas e convolucionais, que têm mostrado resultados promissores em outras áreas de processamento de áudio e podem oferecer melhorias na classificação de emoções musicais.

5.2.3 Aplicações Práticas e Comerciais:

Personalização de Experiências Musicais: Explorar como as características de áudio relacionadas a emoções musicais podem ser utilizadas para criar sistemas de recomendação mais sofisticados. Além de considerar o histórico de escutas dos usuários, as plataformas de streaming podem incorporar análises emocionais para sugerir músicas que se alinhem com o estado emocional atual do ouvinte ou objetivos emocionais, como levantar o ânimo ou acalmar.

Terapia Musical Assistida: Desenvolver ferramentas baseadas em inteligência artificial que utilizem a classificação de emoções musicais para auxiliar profissionais de saúde mental na seleção de músicas para terapia musical. Tais ferramentas poderiam personalizar sessões de terapia para atender às necessidades emocionais específicas dos pacientes, potencializando os benefícios terapêuticos da música.

Análise de Tendências e Marketing Musical: Utilizar insights sobre emoções musicais para entender melhor as preferências dos ouvintes e as tendências de consumo musical. Isso poderia ajudar artistas e gravadoras a criar músicas que atendam a demandas emocionais específicas do mercado ou a desenvolver campanhas de marketing que ressoem melhor com o estado emocional do público-alvo.

Educação Musical e Engajamento: Empregar análises de emoções musicais para desenvolver ferramentas educacionais que ensinem sobre a conexão entre música, emoção e expressão artística. Isso pode aumentar o engajamento dos estudantes com a música, fornecendo uma compreensão mais profunda de como diferentes aspectos da música afetam as emoções.

5.3 Reflexão Final

Este estudo contribui para o crescente campo do Reconhecimento de Emoções Musicais, oferecendo ideias sobre a interseção entre a ciência de dados, engenharia de áudio e psicologia musical. As descobertas deste estudo não apenas avançam o conhecimento acadêmico, mas têm o potencial de influenciar a indústria da música e a experiência musical do ouvinte no futuro.

Ao expandir o entendimento e a aplicação de emoções musicais, este estudo abre portas para uma gama diversificada de aplicações que transcendem a experiência de escuta tradicional. Essas aplicações não apenas têm o potencial de revolucionar a indústria da música, mas também de enriquecer a vida das pessoas, oferecendo novas maneiras de interagir com a música de forma mais pessoal e significativa. Ao aproveitar a ciência por trás das emoções musicais, as plataformas de streaming podem liderar a caminho para uma era de consumo musical mais conectada e emocionalmente inteligente.

REFERÊNCIAS BIBLIOGRÁFICA

- ALJANAKI, A.; YANG, Y. H.; SOLEYMANI, M. Developing a benchmark for emotional analysis of music. **PLoS ONE**, v. 12, n. 3, 1 mar. 2017.
- ALMEIDA SANTOS, E. B.; FREIRE, O. B. D. L. THE INFLUENCE OF MUSIC ON CONSUMER BEHAVIOR. **Independent Journal of Management & Production**, v. 4, n. 2, 1 set. 2013.
- BREIMAN, L. **Random Forests**. [s.l]: s.n.].
- CHEN, T.; GUESTRIN, C. **XGBoost**. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. **Anais...** New York, NY, USA: ACM, 13 ago. 2016.
- CORTES, C.; VAPNIK, V.; SAITTA, L. **Support-Vector Networks EditorMachine Learning**. [s.l.] Kluwer Academic Publishers, 1995.
- COVER, T.; HART, P. Nearest neighbor pattern classification. **IEEE Transactions on Information Theory**, v. 13, n. 1, p. 21–27, jan. 1967.
- FRIEDMAN, J. H. Greedy function approximation: A gradient boosting machine. **The Annals of Statistics**, v. 29, n. 5, 1 out. 2001.
- GOMEZ-CANON, J. S. et al. Music Emotion Recognition: Toward new, robust standards in personalized and context-sensitive applications. **IEEE Signal Processing Magazine**, v. 38, n. 6, p. 106–114, 1 nov. 2021.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [s.l.] The MIT Press, 2016.
- GRIFFITHS, D. et al. A multi-genre model for music emotion recognition using linear regressors. **Journal of New Music Research**, v. 50, n. 4, p. 355–372, 8 ago. 2021.
- HAGEN, A. N. **The Playlist Experience: Personal Playlists in Music Streaming Services. Popular Music and Society** Routledge, , 20 out. 2015.

- HAN, D. et al. **A survey of music emotion recognition.** *Frontiers of Computer Science* Higher Education Press Limited Company, , 1 dez. 2022.
- HURON, D. **Sweet Anticipation.** [s.l.] The MIT Press, 2006.
- IVERSON, P.; KRUMHANSL, C. L. Isolating the dynamic attributes of musical timbre). **The Journal of the Acoustical Society of America**, v. 94, n. 5, p. 2595–2603, 1 nov. 1993.
- JAMES, G. et al. **An Introduction to Statistical Learning.** New York, NY: Springer New York, 2013. v. 103
- JUSLIN, P. N. **Handbook of Music and Emotion: Theory, Research, Applications.** [s.l.] Oxford University Press, 1993.
- JUSLIN, P. N.; LAUKKA, P. **Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?** *Psychological Bulletin*, set. 2003.
- JUSLIN, P. N.; VÄSTFJÄLL, D. Emotional responses to music: The need to consider underlying mechanisms. **Behavioral and Brain Sciences**, v. 31, n. 5, 2008.
- KRUMHANSL, C. L. An exploratory study of musical emotions and psychophysiology. **Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale**, v. 51, n. 4, p. 336–353, dez. 1997.
- LAZAROV, A.; PINE, D. S.; BAR-HAIM, Y. Gaze-Contingent Music Reward Therapy for Social Anxiety Disorder: A Randomized Controlled Trial. **The American journal of psychiatry**, v. 174, n. 7, p. 649–656, 1 jul. 2017.
- M. BISHOP, C. **Pattern Recognition and Machine Learning.** 1. ed. Nova York: Springer, 2006.
- MEYER, L. B. **Emotion and Meaning in Music.** Chicago: [s.n.].
- MOSCHETTA, P. H.; VIEIRA, J. Music in the streaming era: Music curatorship and discovery in Spotify. **Sociologias**, v. 20, n. 49, p. 258–292, 1 set. 2018.

PANDA, R. et al. How Does the Spotify API Compare to the Music Emotion Recognition State-of-the-Art? WELCOME-Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Comorbidities View project LINK-Linking Excellence in Biomedical knowledge and Computational Intelligence Research for personalized management of CVD within PHC View project HOW DOES THE SPOTIFY API COMPARE TO THE MUSIC EMOTION RECOGNITION STATE-OF-THE-ART? **18th Sound and Music Computing Conference (SMC 2021)**, p. 238–245, 2021.

PANDA, R.; MALHEIRO, R.; PAIVA, R. P. Novel Audio Features for Music Emotion Recognition. **IEEE Transactions on Affective Computing**, v. 11, n. 4, p. 614–626, 1 out. 2020.

PRÓ-MÚSICA. **Mercado Fonográfico Brasileiro**. Disponível em: <<https://pro-musicabr.org.br/wp-content/uploads/2023/03/2023-03-20-Mercado-Brasileiros-em-2023.pdf>>. Acesso em: 14 jan. 2024.

RISH, I.; RISH, I. **An Empirical Study of the Naïve Bayes Classifier An empirical study of the naive Bayes classifier**. [s.l.: s.n.]. Disponível em: <<https://www.researchgate.net/publication/228845263>>.

RUSSELL, J. A. A circumplex model of affect. **Journal of Personality and Social Psychology**, v. 39, n. 6, p. 1161–1178, dez. 1980.

RUSSELL, S.; PETER, N. **Stuart Russell, Peter Norvig - Artificial Intelligence_ A Modern Approach (Pearson Series in Artificial Intelligence)-Pearson (2020)**. [s.l.: s.n.].

SANTANA, M. A. et al. **Reconhecimento Automático de Emoções a partir de Sinais Multimodais e Inteligência Artificial**. [s.l.: s.n.]. Disponível em: <<https://www.researchgate.net/publication/348560154>>.

SCHLOSBERG, H. The description of facial expressions in terms of two dimensions. **Journal of Experimental Psychology**, v. 44, n. 4, p. 229–237, 1952.

SCHUBERT, E. Modeling Perceived Emotion With Continuous Musical Features. **Music Perception**, v. 21, n. 4, p. 561–585, 1 jun. 2004.

SENA MOORE, K. Understanding the Influence of Music on Emotions: A Historical Review. **Music Therapy Perspectives**, p. miw026, 11 nov. 2016.

SINCLAIR, G.; TINSON, J. Psychological ownership and music streaming consumption. **Journal of Business Research**, v. 71, p. 1–9, 1 fev. 2017.

SPOTIFY. **Spotify Documentation Web API**. Disponível em: <<https://developer.spotify.com/documentation/web-api>>. Acesso em: 14 jan. 2024.

SUTTON, R. S.; BARTO, A. G. **Reinforcement Learning: An Introduction**. 2. ed. [s.l.] The MIT Press, 2018.

THAUT, M. H.; MCINTOSH, G. C.; HOEMBERG, V. **Neurobiological foundations of neurologic music therapy: rhythmic entrainment and the motor system**. *Frontiers in Psychology* Frontiers Media S.A., , 18 fev. 2015.

VOLOKHIN, S.; AGICHTEIN, E. **Towards intent-aware contextual music recommendation: Initial experiments**. 41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2018. *Anais*...Association for Computing Machinery, Inc, 27 jun. 2018.

WADLEY, G. et al. **Use of music streaming platforms for emotion regulation by international students**. ACM International Conference Proceeding Series. *Anais*...Association for Computing Machinery, 2 dez. 2019.

XU, L. et al. How have music emotions been described in Google books? Historical trends and corpus differences. **Humanities and Social Sciences Communications**, v. 10, n. 1, 1 dez. 2023.

APÊNDICE A - REPOSITÓRIO DE DADOS

Os dados e notebooks utilizados para a construção das rotinas em Python estão disponíveis no endereço: <https://github.com/lubohen/eEDB-018>.

A.1. Códigos Fontes

Os códigos fontes nos dois próximos anexos complementa a etapa de Construção do Conjunto de Dados e Pré-processamento explicada no capítulo de desenvolvimento.

ANEXO A - GERAÇÃO DO CONJUNTO DE DADOS

```

import requests
import pandas as pd
from config import SPOT_CLIENT_SECRET, SPOT_CLIENT_ID

# Função para obter o token de acesso ao Spotify
def get_spotify_token(client_id, client_secret):
    """
    Obtém um token de acesso ao Spotify.

    Parâmetros:
    client_id (str): ID do cliente Spotify
    client_secret (str): Segredo do cliente Spotify

    Retorna:
    str: Token de acesso
    """
    auth_url = "https://accounts.spotify.com/api/token"
    auth_header =
    base64.b64encode(f"{client_id}:{client_secret}".encode('utf-
8')).decode('utf-8')
    headers = {"Authorization": f"Basic {auth_header}"}
    data = {"grant_type": "client_credentials"}

    response = requests.post(auth_url, headers=headers, data=data)
    if response.status_code != 200:
        raise Exception(f"Erro ao obter token. Código
{response.status_code}: {response.text}")

    return response.json()['access_token']

# Obter o token de acesso
CLIENT_ID = SPOT_CLIENT_ID
CLIENT_SECRET = SPOT_CLIENT_SECRET
TOKEN = get_spotify_token(CLIENT_ID, CLIENT_SECRET)

# Cabeçalhos para solicitações à API do Spotify
headers = {"Authorization": "Bearer {TOKEN}"}

# Função para buscar características de áudio de uma faixa específica
def get_audio_features(track_name, artist_name):
    """
    Busca características de áudio de uma faixa específica usando a API do
    Spotify.

    Parâmetros:
    track_name (str): Nome da faixa
    artist_name (str): Nome do artista

    Retorna:
    dict: Características de áudio da faixa ou None se não encontrada
    """
    response =
    requests.get(f"{BASE_URL}search?q=track:{track_name}+artist:{artist_name}&t
ype=track", headers=headers)

```

```

response_json = response.json()

if 'tracks' not in response_json:
    print(f"Erro na resposta para: {track_name} de {artist_name}.")
    Resposta: {response_json}")
    not_found_tracks.append([track_name, artist_name])
    return None

items = response_json['tracks']['items']
if not items:
    print(f"Não foi possível encontrar: {track_name} de {artist_name}")
    not_found_tracks.append([track_name, artist_name])
    return None

track_id = items[0]['id']
track_response = requests.get(f"{BASE_URL}audio-features/{track_id}",
headers=headers)
return track_response.json()

# Carregar dados do arquivo CSV
basedir = r"C:\Users\luboh\OneDrive\Área de
Trabalho\USP\EngDadosBigData\eEDB-018_REPO\DEVELOPER\eEDB-018\prj\DATA"
data_file = basedir + r'\panda_dataset_taffc_metadata.csv'
data_source = pd.read_csv(data_file)

# Processamento de faixas
tracks = data_source[['Title', 'Artist', 'GenresStr', 'Quadrant']].values
not_found_tracks = []
data = []

for track_name, artist_name, genre, quadrant in tracks:
    audio_features = get_audio_features(track_name, artist_name)

    if audio_features is None:
        continue

    data.append([
        track_name, artist_name, genre, quadrant,
        audio_features['acousticness'], audio_features['danceability'],
        audio_features['energy'], audio_features['instrumentalness'],
        audio_features['key'], audio_features['liveness'],
        audio_features['loudness'], audio_features['mode'],
        audio_features['speechiness'], audio_features['tempo'],
        audio_features['time_signature'], audio_features['valence']
    ])

# Criação dos DataFrames
df = pd.DataFrame(data, columns=[
    'Nome da Faixa', 'Cantor ou Compositor', 'Genero', 'Quadrante',
    'acousticness', 'danceability', 'energy', 'instrumentalness',
    'key', 'liveness', 'loudness', 'mode', 'speechiness', 'tempo',
    'time_signature', 'valence'
])
df_not_found = pd.DataFrame(not_found_tracks, columns=['Nome da Faixa',
'Cantor ou Compositor'])

# Salvando os DataFrames

```

```
df.to_csv('spotify_dataset.csv', index=False)
df_not_found.to_csv('not_found_spotify.csv', index=False)
```

ANEXO B - PRÉ-PROCESSAMENTO

```

import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
from sklearn.impute import SimpleImputer
from sklearn.preprocessing import LabelEncoder
import joblib

# Função para carregar e preparar os dados
def prepare_data(file_path, target_column, numeric_features, test_size=0.2,
random_state=42):
    # Carregar dados
    df = pd.read_csv(file_path)

    # Separar em features e target
    X = df[numeric_features] # Supõe-se que todas as features sejam
    numéricas
    y = df[target_column]

    # Codificar a variável alvo categórica
    label_encoder = LabelEncoder()
    y_encoded = label_encoder.fit_transform(y)

    # Dividir os dados em conjuntos de treinamento e teste
    X_train, X_test, y_train, y_test = train_test_split(X, y_encoded,
    test_size=test_size, random_state=random_state)

    # Construir o pré-processador para features numéricas
    numeric_transformer = Pipeline(steps=[
        ('imputer', SimpleImputer(strategy='median')), # Preencher valores
        ausentes
        ('scaler', StandardScaler()) # Normalizar dados
    ])

    # Coluna Transformer para aplicar as transformações nas colunas
    numéricas
    preprocessor = ColumnTransformer(
        transformers=[
            ('num', numeric_transformer, numeric_features)
        ])
    ])

    # Pré-processar os dados de treinamento e teste
    X_train_processed = preprocessor.fit_transform(X_train)
    X_test_processed = preprocessor.transform(X_test)

    # Salvar o pré-processador e o label encoder
    joblib.dump(preprocessor, scaler_path)
    joblib.dump(label_encoder, label_encoder_path)

```

```
# Retorna os conjuntos de dados processados e o pré-processador para uso futuro
return X_train_processed, X_test_processed, y_train, y_test,
preprocessor, label_encoder

# Uso da função
file_path = 'spotify_dataset.csv'
target_column = 'Quadrante'
numeric_features = ['acousticness', 'danceability', 'energy',
'instrumentalness', 'key', 'liveness', 'loudness', 'speechiness', 'tempo',
'velence', 'mode', 'time_signature']
scaler_path = 'preprocessor.joblib'
label_encoder_path = 'label_encoder.joblib'

# Prepare os dados
X_train, X_test, y_train, y_test, preprocessor, label_encoder =
prepare_data(file_path, target_column, numeric_features)

# Agora X_train, X_test, y_train e y_test estão prontos para serem usados em qualquer modelo de ML.
```