



ESTUDO SOBRE A APLICAÇÃO DAS TECNOLOGIAS DE INTELIGÊNCIA ARTIFICIAL E MACHINE LEARNING PARA MAPEAMENTO DE PROVÁVEIS HITS NA INDÚSTRIA MUSICAL

Marina Haru Marcoulakis (IC) e Eurico Luiz Prospero Ruivo(Orientador)

Apoio: PIBIC MackPesquisa.

RESUMO

Este estudo foca na utilização de modelos MLP de rede neural preditivos de Inteligência Artificial (IA) para prever o sucesso musical, especialmente em termos de aumento da popularidade de músicas e posições nas paradas. A metodologia integra uma ampla gama de features musicais para prever a taxa de sucesso de uma música. Foi utilizado um referencial de sucesso musical com taxas de sucesso baseadas nas classificações da Billboard e comparamos isso com um modelo orientado por características musicais. Essas descobertas têm o potencial de oferecer insights para melhorar composições musicais e otimizar estratégias de lançamento, indo além dos contextos tradicionais de classificação.

Palavras-chave: Inteligência Artificial. Indústria Musical. Hit Song Science.

ABSTRACT

Abstract - This study focuses on the utilization of predictive Artificial Intelligence (AI) models to forecast musical success, particularly in terms of enhancing song popularity and chart rankings. Our methodology integrates various musical features to predict a song's success rate. We establish a benchmark for musical success by using success rates based on Billboard rankings and compare them with a model driven by musical features. These findings have the potential to offer valuable insights for improving musical compositions and optimizing release strategies, extending beyond traditional ranking contexts.

Keywords: Artificial Intelligence. Music Industry. Hit Song Science.

1. Introdução

A evolução tecnológica, especialmente no campo da Inteligência Artificial (IA) e do Machine Learning (ML), tem transformado diversas indústrias, incluindo a música. No cenário musical, prever o sucesso de uma música antes de seu lançamento tem sido um desafio constante para produtores, artistas e gravadoras. Tradicionalmente, essa tarefa envolvia uma combinação de experiência intuitiva, marketing e análise de tendências passadas.



Contudo, a crescente complexidade do mercado musical, impulsionada por mudanças nos padrões de consumo, plataformas digitais e a globalização, exige métodos mais precisos e baseados em dados para garantir o sucesso de novos lançamentos.

Neste contexto, a aplicação de técnicas de IA, como redes neurais artificiais, para prever hits musicais representa uma inovação significativa. Estas tecnologias permitem a análise de grandes volumes de dados, identificando padrões e correlações que podem escapar à análise humana tradicional. A pesquisa sobre predição de hits, conhecida como "Hit Song Science" (HSS), tem ganhado destaque, fornecendo ferramentas que podem influenciar diretamente as decisões estratégicas na indústria da música, como a escolha de singles, o timing de lançamento e o direcionamento de campanhas de marketing.

O presente estudo busca explorar essa intersecção entre IA e música, investigando como modelos preditivos podem ser utilizados para antecipar o sucesso de faixas musicais. Utilizando um conjunto de dados robusto, que inclui músicas presentes na Billboard Year-End Chart de 1958 a 2022, integramos diversas características musicais extraídas da API do Spotify. Essas características, como ritmo, energia e dança, são utilizadas para treinar um modelo de rede neural, especificamente um MLPRegressor, para prever a posição futura de uma música nas paradas de sucesso.

Ao longo deste estudo, comparamos o desempenho do modelo de IA com um referencial de sucesso musical tradicional, representado pelas classificações da Billboard, buscando validar a eficácia das técnicas aplicadas. Além disso, exploramos as implicações dos achados, tanto em termos práticos quanto teóricos, para a indústria musical. Por exemplo, ao identificar padrões que contribuem para o sucesso de uma música, podemos oferecer insights valiosos para compositores, produtores e executivos, auxiliando na criação de músicas com maior potencial de sucesso e na otimização das estratégias de lançamento.

Este trabalho, portanto, não apenas contribui para o campo emergente da ciência de predição de hits, mas também destaca o papel crescente da IA na reconfiguração dos processos criativos e comerciais na indústria musical. Ao proporcionar uma análise baseada em dados, esperamos fornecer uma ferramenta que possa aumentar a precisão na previsão de hits, reduzindo os riscos associados a novos lançamentos e maximizando o retorno sobre o investimento para gravadoras e artistas.

2. Referencial Teórico

Como explorado por Park e Jun (2023), é evidente a drástica alteração da indústria cultural como um todo nos últimos anos devido à digitalização e à maior influência da



internet na vida cotidiana. Com isso, há uma constante reformulação na maneira que mercados culturais, como o da música, se comportam, o que consequentemente impactou nas estratégias de venda para as massas, conforme demonstrado por Avlonitis (2023). Portanto, diante das novas demandas do mercado fonográfico, surge a ideia de desenvolver uma ferramenta que, baseada em dados estatísticos, musicais, demográficos e de "sucessos" anteriores no meio, aumente a acurácia na projeção da receita de novos lançamentos para as gravadoras, conforme as tendências populares da indústria.

Por conta dessa constante e emergente situação de renovação de interesses artístico-culturais, faz-se necessária a utilização de novos métodos para a predição dos resultados de grandes lançamentos. Desse modo, a aplicação de inteligências artificiais para avaliar como as músicas podem ter um impacto em determinado público mostra-se uma proposta interessante para determinar lançamentos mais eficientes, com maiores retornos e menores riscos para os produtores. Neste contexto, o trabalho analisa a possibilidade de prever o sucesso musical antes do lançamento das faixas, utilizando técnicas de aprendizado de máquina e considerando tanto as propriedades técnicas da música quanto o conteúdo de sua letra.

“No entanto, da perspectiva de um produtor musical, artista e estúdio, seria interessante analisar o sucesso de uma música a priori. Um sistema que prevê o sucesso de uma música com base em sua técnica e propriedades seria de grande benefício para os produtores musicais. Poder prever a resposta esperada para uma música do ponto de vista do ouvinte, por não considerar apenas os recursos de áudio bem como o conteúdo da letra” (RAZA.; NANATH, 2020, *tradução nossa*).

Com a fluidez da sociedade atual, uma ferramenta capaz de calcular como as audiências reagiriam às músicas seria uma verdadeira revolução. Com isso, seria possível entregar um conteúdo personalizado para cada consumidor de acordo com seus gostos musicais, economizando tempo e dinheiro¹ ao evitar investimentos em músicas que não seriam bem-sucedidas. Além disso, seria possível economizar com pesquisas de marketing custosas, uma vez que essa metodologia seria mais precisa e barata.

Graças ao avanço da inteligência artificial, surgiram novas possibilidades em diversas áreas, incluindo a indústria musical. A capacidade de calcular a probabilidade de sucesso de uma música antes de seu lançamento é uma das novas oportunidades para pesquisa e aprimoramento nessa área. Estudos como estes são úteis tanto para as

¹ "Em 2013, as gravadoras investiram um total de USD 4,3 bilhões em novos talentos ao redor do mundo" Investing in music. *RIAA*, Washington, D.C., 11 nov. 2014.



produtoras, que podem maximizar seus lucros, quanto para os artistas, que podem obter melhores conclusões sobre o futuro apelo de suas novas faixas.

De acordo com (RAZA; NANATH, 2020), sistemas de inteligência artificial podem detectar músicas de sucesso com alto nível de precisão, permitindo aos usuários obter uma estimativa do apelo de uma faixa antes de seu lançamento. Essa análise pode ajudar os produtores a maximizar seus lucros e a reduzir os riscos associados à produção musical.

Com isso, o uso de aprendizado de máquina possibilitará a análise e identificação de padrões utilizando uma grande quantidade de dados musicais. Baseando-se na linha de pesquisa que vem sendo feita com sistemas de composição automática², o programa pode levar a conclusões valiosas para produtores durante a criação de novas músicas, aumentando as chances na criação de *hits*.

Além disso, a análise de dados também pode ajudar a personalizar recomendações para gravadoras e fornecer uma vantagem competitiva na indústria musical. Pesquisas anteriores já obtiveram sucesso ao aplicar tecnologias similares para um gênero musical único; por isso, surge a ideia de se aplicar para vários gêneros e se obter um panorama mais amplo, visando a possibilidade de desenvolver uma ferramenta que seja capaz de indicar possíveis mudanças a serem feitas antes do lançamento, para que faixas mais bem-sucedidas sejam obtidas.

Além disso, essa aplicação de inteligências artificiais e tecnologias de *machine learning* para esse tipo de pesquisa já foi comprovada e julgada promissora, como mostra o artigo *Dance hit song prediction*: Esta pesquisa prova que a popularidade de músicas eletrônicas pode ser aprendida a partir da análise dos padrões musicais" (MARTENS; HERREMANS; SÖRENSEN, 2014, tradução nossa).

3. Metodologia

Neste estudo, conduzimos um processo sistemático de coleta de dados visando adquirir um conjunto de composto por músicas apresentadas na Billboard Year-End Chart entre 1958 e 2022. O objetivo principal foi analisar a relação entre as características das músicas e suas posições nas paradas.

a. Dados

² MARTENS, David; HERREMANS, Dorien; SÖRENSEN, Kenneth. *Classification and generation of composer-specific music using global feature models and variable neighborhood search*, set. 2015.



O processo de coleta de dados começou utilizando a biblioteca "billboard-top-100" do Node.js, que forneceu acesso aos dados da Billboard Year-End Chart. Esta biblioteca facilitou a extração de uma lista completa de músicas que apareceram na parada durante o período especificado.

Para enriquecer o conjunto de dados, integramos a API do Spotify à pipeline de coleta de dados, o que, segundo pesquisas anteriores, proporciona resultados positivos. Essa integração permitiu obter identificadores de músicas, como ID do Artista, ID do Álbum e ID da Faixa. Esses identificadores foram fundamentais para estabelecer conexões entre as músicas e seus metadados, possibilitando uma análise detalhada dos atributos das músicas.

Foram utilizadas 14 características principais da API do Spotify em abril de 2023. Essas 14 características foram escolhidas por sua capacidade de capturar aspectos essenciais das músicas, como ritmo, energia, dança, e outros elementos que influenciam diretamente a popularidade e a aceitação de faixas musicais pelo público. A seleção dessas características nos permitiu realizar uma análise mais focada e precisa no contexto da predição de hits musicais.

Essas características incluíam características importantes, como *acousticness*, *danceability*, *duração em milissegundos*, *energy*, *gênero*, *instrumentalness*, *key*, *liveness*, *loudness*, *mode*, *speechiness*, *tempo*, *time signature* e *valence*. Essas métricas formaram a base para as análises subsequentes, fornecendo percepções sobre as características das músicas nos conjuntos de dados.

O processo de compilação do conjunto de dados envolveu a seleção de músicas com base em suas aparições nas paradas do Billboard Top 100, como foi feito com sucesso em pesquisas anteriores (CIBILS, 2023). Essa seleção garantiu a inclusão de músicas que tiveram um impacto notável na indústria da música e ganharam atenção significativa do público. O conjunto de dados incorporou uma ampla gama de informações, incluindo posições nas paradas, detalhes dos artistas, informações sobre álbuns, características das faixas e classificações de gêneros.

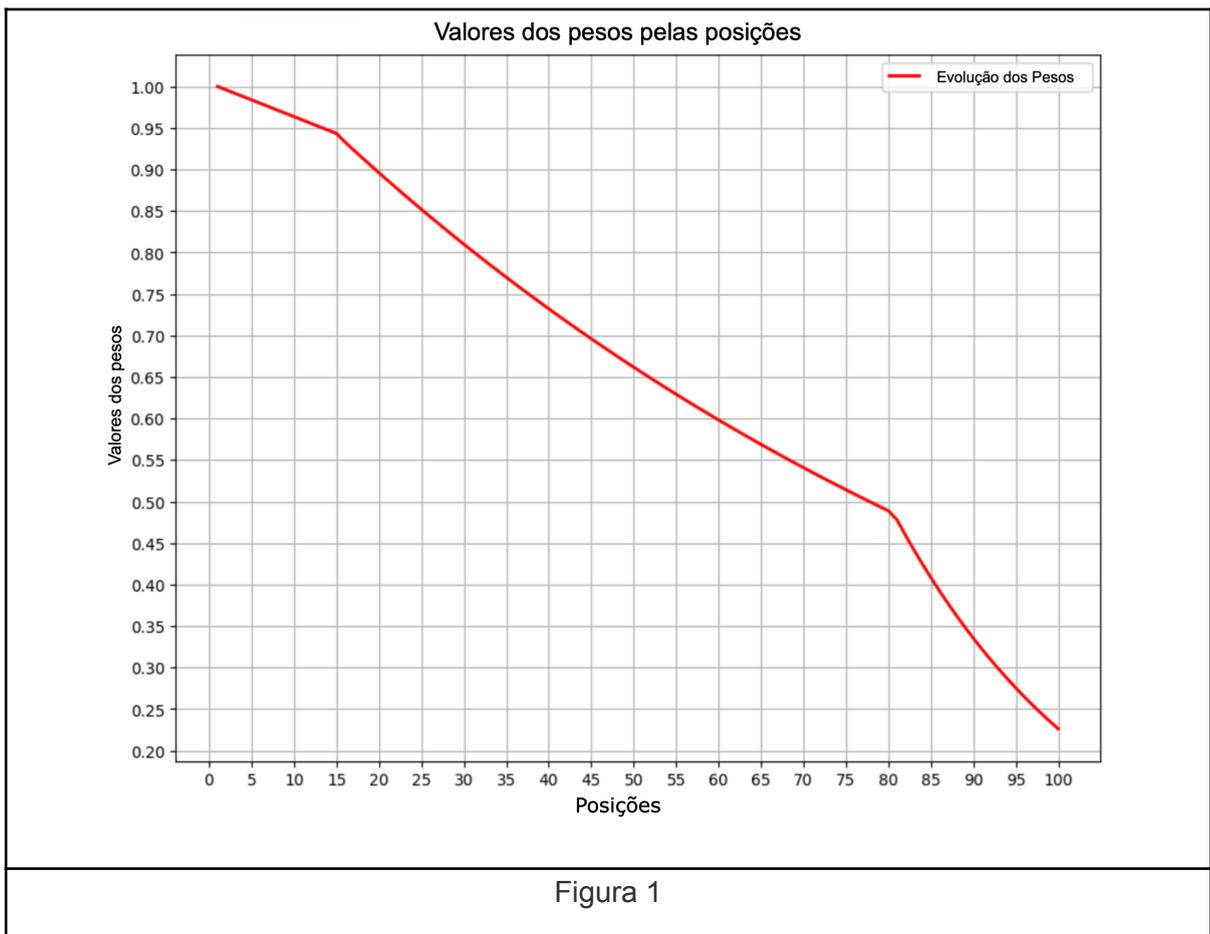
Para aprimorar a análise, introduzimos a métrica de "taxa de sucesso", um método para quantificar o desempenho de uma música nas paradas. Esta métrica envolveu a atribuição de pesos específicos às posições nas paradas, conforme descrito na Figura 1. Esses valores ponderados foram então utilizados em uma equação de média ponderada, para calcular uma pontuação geral de posição para cada música dentro das classificações da Billboard Year-End Chart demonstrado na fórmula abaixo:



$$Taxa de Sucesso = \frac{\sum(\text{valores não nulos} \times \text{valor} \times \text{peso})}{\text{Total de semanas não nulas nos rankings}}$$

Essas classificações forneceram uma visão abrangente do desempenho de cada música ao longo do tempo, indicando com que frequência uma música manteve cada posição em um determinado ano de parada.

No entanto, durante a análise, encontramos um desafio significativo ao perceber que alguns dos dados coletados da API do Spotify estavam incorretos devido a imprecisões no processo de automação de coleta de dados. Essa discrepância exigiu uma revisão manual e minuciosa de todos os dados coletados. Tivemos que verificar e validar cada música para garantir a confiabilidade da análise, levando em consideração a presença de diversas músicas não correspondentes com seus identificadores. Esse processo exigiu a reconstrução manual de banco de dados, cruzando várias fontes para verificar a exatidão das informações coletadas.



b. Pré-Processamento



Organizamos o conjunto de dados utilizando os identificadores e dados obtidos da Billboard Year-End Chart. A abordagem foi centrada no tempo, onde dividimos o conjunto de dados em semanas. A posição de cada música na parada para uma semana específica foi indexada nessa estrutura. Esse método garantiu um registro histórico claro e organizado do desempenho das músicas ao longo do tempo.

Além dos dados fundamentais das músicas, coletamos informações suplementares. Isso inclui gêneros musicais associados ao artista de cada música e outras características relevantes que contribuíram para um perfil musical abrangente.

Para melhorar a eficiência da recuperação de dados e minimizar as solicitações à API, refinamos o conjunto de dados. O transformamos em uma coleção de objetos JSON, cada um identificado pelo nome da música. Esses objetos continham dois componentes: características detalhadas da música e uma array de 100 elementos. Esse array documentava as posições de classificação da música, incluindo a frequência e as datas correspondentes.

Após a criação dos objetos JSON, os convertimos em um formato de array. Esse array serviu como a estrutura de dados fundamental para as etapas de análise subsequentes. Geramos quatro arquivos CSV distintos a partir desse array. O primeiro arquivo CSV incluía os nomes das músicas e suas posições nas paradas, simplificando o acesso aos dados de desempenho nas paradas. O segundo arquivo CSV fornecia uma lista abrangente de classificações, identificadas como "rank_1", "rank_2" e assim por diante, junto com os nomes das músicas. Os dois arquivos CSV restantes seguiam uma estrutura semelhante, mas também incluíam características das músicas para uma análise abrangente.

Para tornar o conjunto de dados mais eficiente, utilizamos a API do Spotify para integrar características de análise musical. Essa etapa forneceu insights valiosos sobre características das músicas, como tempo, key, e danceability.

c. Modelo

Na busca por modelagem de regressão precisa, empregamos a biblioteca scikit-learn em Python, fazendo uso do MLPRegressor — um modelo de aprendizado de máquina explicitamente voltado para tarefas de regressão. O MLPRegressor, sustentado por uma arquitetura de perceptron multicamadas (rede neural), demonstra notáveis capacidades de previsão de valores numéricos com base em características de entrada complexas. A flexibilidade inerente dessa arquitetura permite capturar adequadamente relações não-lineares complexas dentro dos dados.



O desempenho inicial do modelo foi avaliado utilizando dados do ano de 2020. O conjunto de dados foi dividido em 80% para treinamento e 20% para teste, garantindo uma avaliação justa e alinhada com boas práticas de aprendizado de máquina. Optamos pelo MLPRegressor devido à sua capacidade de capturar relações não-lineares complexas, características da indústria musical.

Através de uma exploração dos parâmetros do modelo, identificamos a combinação ideal, que culminou em uma pontuação baixa de Erro Médio Absoluto (MAE), girando em torno de 0,2. O MAE, uma métrica essencial na avaliação do desempenho de modelos de regressão, quantifica a discrepância absoluta média entre os valores previstos e os valores reais, fornecendo uma medida do erro típico de previsão do modelo nas mesmas unidades que a variável alvo. Notavelmente, a capacidade do modelo de manter um MAE baixo ressalta sua superior precisão preditiva.

Impulsionados por esses resultados iniciais, iniciamos uma avaliação abrangente do funcionamento do modelo, implantando um conjunto de dados mais extenso, abrangendo 29.532 músicas. No decorrer dessa avaliação, examinamos vários esquemas de particionamento de dados, incluindo divisões percentuais de dados de teste/treinamento de 10%, 15%, 20%, 25%, 30% e 40%.

No entanto, a precisão do modelo foi de 17.53%, o que indicou a necessidade de um processo de validação manual para lidar com a imprecisão dos dados utilizados. Reconhecendo esse desafio, tivemos que recolher os dados e conduzir um processo de validação manual em todos os dados coletados para garantir sua confiabilidade. Observamos que havia 3.987 discrepâncias entre as músicas em banco de dados e as esperadas. Essas discrepâncias destacaram a extensão das imprecisões em conjunto de dados inicial.

Prosseguimos para analisar os dados ano a ano usando o modelo para validar seu desempenho. Essa abordagem minuciosa nos permitiu identificar discrepâncias e inconsistências. Durante essa análise, também percebemos que o modelo poderia se beneficiar da incorporação de um fator temporal. Observamos que músicas mais recentes frequentemente exibiam tendências de desempenho diferentes em comparação com músicas mais antigas. Portanto, percebemos a necessidade de atribuir mais valor aos pontos de dados recentes. Ao adicionar esse fator temporal ao modelo, visamos aumentar sua precisão preditiva, garantindo que ele pudesse capturar efetivamente a natureza dinâmica da indústria da música.

4. Resultados e discussão



A primeira análise dividiu o conjunto de dados em períodos de 10 anos, conforme mostrado no gráfico da Figura 3. A precisão média alcançada pelo modelo foi de 62,19%, variando significativamente entre os períodos analisados (figura 3), com a maior precisão (em torno de 67%) sendo alcançada quando o modelo foi treinado com dados de 1998-2007 e testado com dados de 2008-2017, e vice-versa. Isso sugere que o modelo apresenta bom desempenho quando há uma sobreposição temporal entre os dados de treinamento e teste, refletindo tendências consistentes na indústria da música nesses períodos.

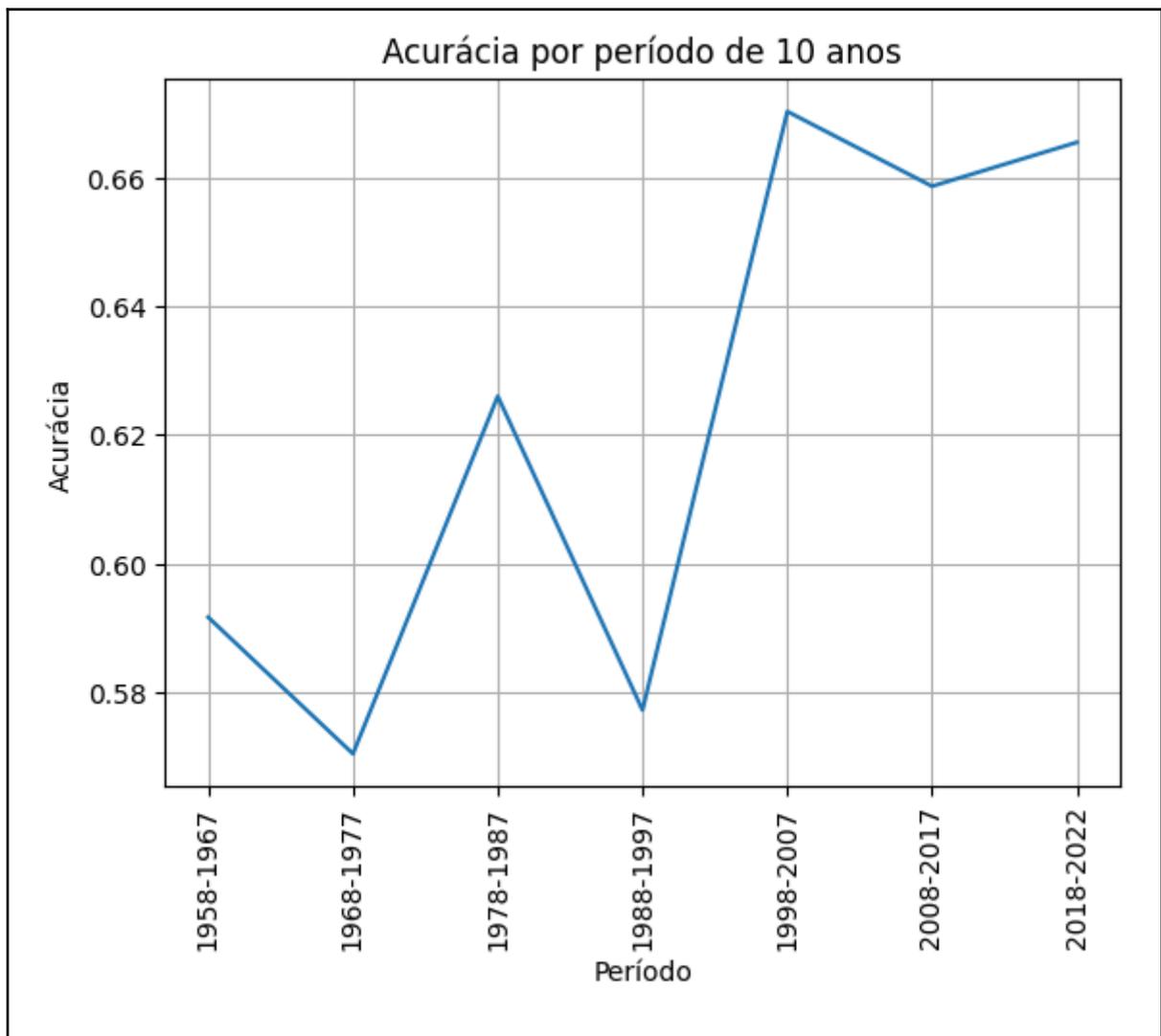
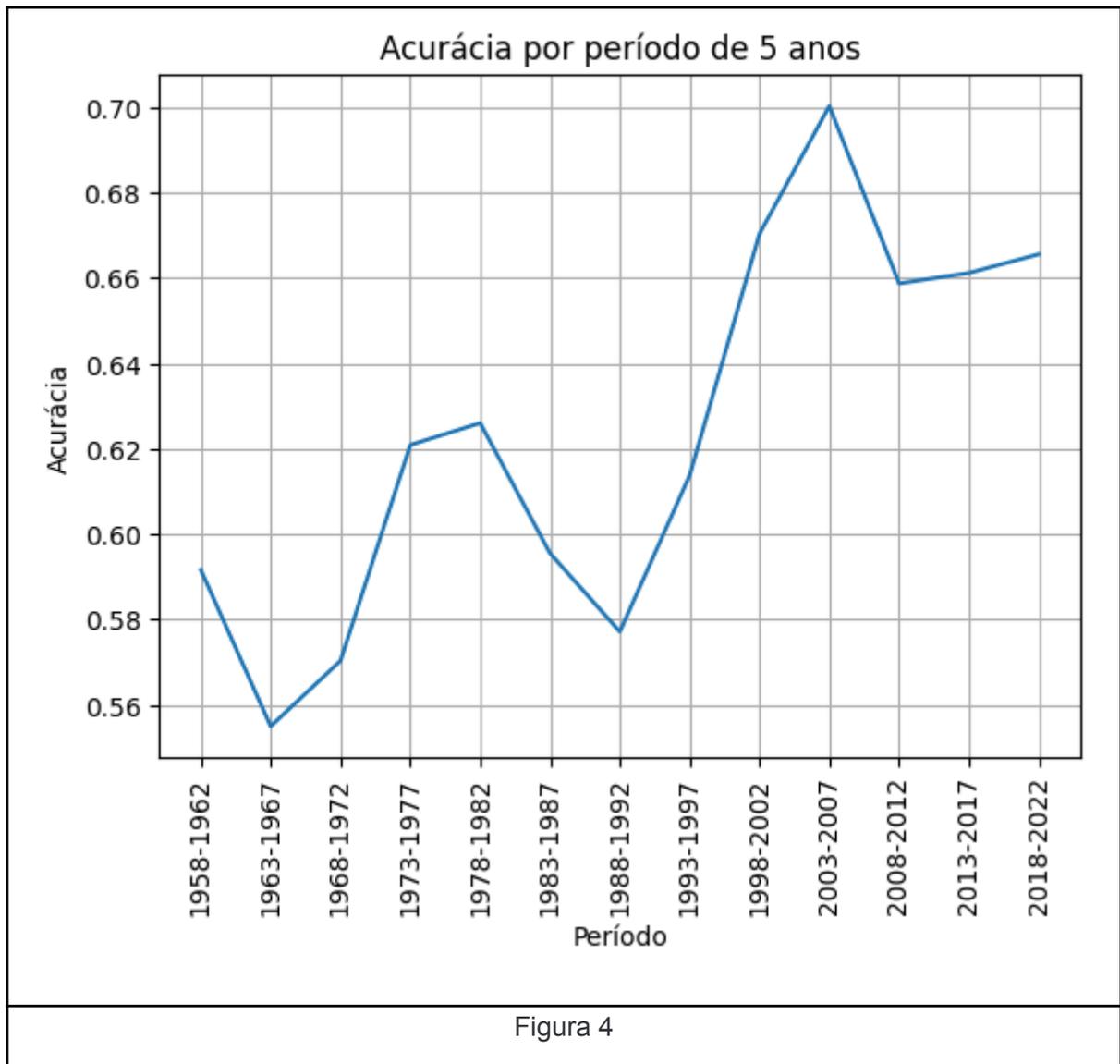


Figura 3

Para obter uma compreensão mais detalhada, dividimos ainda mais o conjunto de dados em períodos de 5 anos, conforme mostrado no gráfico intitulado "Acurácia por período de 10 anos". A precisão média nesta análise foi ligeiramente maior, em aproximadamente 0,6232. Os resultados novamente mostraram variabilidade na precisão entre os diferentes períodos, com a maior precisão (em torno de 0,70) ocorrendo quando o



modelo foi treinado com dados de 2008-2012 e testado com dados de 2003-2007, assim como o inverso. Isso indica que divisões temporais ainda mais detalhadas podem capturar tendências mais específicas, levando a um desempenho melhorado do modelo.



No entanto, é importante ressaltar que originalmente o modelo foi concebido para retornar um valor contínuo, indicando uma "taxa de sucesso", o que caracteriza um problema de regressão. Nesse contexto, a métrica apropriada para avaliação seria o Erro Absoluto Médio (MAE). Entretanto, as referências à "acurácia" sugerem uma abordagem de classificação, onde a acurácia é uma métrica válida.

Se o modelo foi ajustado para uma tarefa de classificação (por exemplo, prever se uma música seria um "hit" ou "não-hit"), é necessário esclarecer como essa transformação



foi realizada e como a acurácia foi calculada. Para um problema de regressão, o uso de acurácia seria inadequado, e as métricas de erro (como MAE) deveriam ser usadas para avaliar o desempenho

As análises revelaram uma clara influência temporal na precisão do modelo, com períodos que exibem maior proximidade temporal entre os conjuntos de dados de treinamento e teste geralmente apresentando maior precisão. Isso destaca a importância dos dados recentes na previsão das tendências atuais. A notável variabilidade na precisão entre os diferentes períodos reflete ainda a natureza dinâmica e em evolução da indústria da música, sugerindo que tendências e padrões na popularidade musical podem mudar significativamente ao longo do tempo.

Ao incorporar esses fatores temporais, o modelo pode ser ajustado para priorizar dados recentes, aumentando assim sua precisão preditiva e tornando-o mais responsivo às tendências atuais na indústria da música. Essa abordagem permite um modelo mais robusto e adaptável, capaz de capturar melhor as nuances de um cenário em constante mudança, levando, em última análise, a previsões e conclusões mais confiáveis.

O sucesso do modelo pode ser atribuído a um conjunto de parâmetros cuidadosamente ajustados, que colaboram para sua precisão na previsão de valores numéricos. Um dos componentes fundamentais é a função de ativação escolhida, a ReLU (Rectified Linear Unit), amplamente utilizada em redes neurais devido à sua capacidade de introduzir não-linearidade de forma eficiente. Essa função, definida como $f(x) = \max(0, x)$, permite ao modelo capturar relações complexas nos dados, o que melhora a convergência durante o treinamento.

Outro aspecto crucial é o tamanho do lote (batch size), configurado em 32. Esse valor representa o número de amostras processadas antes que os pesos do modelo sejam atualizados. Optar por um tamanho de lote pequeno, mas equilibrado, como 32, garantiu atualizações frequentes dos parâmetros, resultando em maior adaptabilidade ao conjunto de dados.

A arquitetura do modelo também desempenha um papel significativo no desempenho. As camadas ocultas foram projetadas com tamanhos de 10, 20 e 10 neurônios, respectivamente, formando uma estrutura hierárquica que permite ao modelo capturar padrões intrincados e relações não-lineares presentes nos dados.

Essas escolhas, aliadas a outros parâmetros, como uma taxa de aprendizado inicial de 0,001, a utilização do solver 'adam' e a desativação do baralhamento de dados e do *early stopping*, reforçam a robustez do modelo, além de destacar a capacidade do modelo de não

apenas capturar, mas também representar fielmente relações intrincadas e não-lineares dentro do conjunto de dados, reafirmando seu mérito como uma ferramenta inestimável no campo da análise de regressão.

5. Considerações Finais

Esta pesquisa demonstra a importância do uso de modelos preditivos de Inteligência Artificial (IA) na indústria da música. Especificamente, aplicamos modelos de IA para prever a popularidade de músicas e posições nas paradas, aproveitando um conjunto diversificado de características musicais. A investigação desse trabalho se baseou na avaliação objetiva das capacidades preditivas da IA, usando as classificações da Billboard HOT 100 como referência.

Este estudo demonstrou que os modelos de IA, especificamente o MLPRegressor, apresentam uma precisão considerável na predição de atributos musicais e popularidade. A inclusão de fatores temporais e a validação manual dos dados foram determinantes para o sucesso das análises.

Por meio de uma rigorosa coleta e análise de dados, estabelecemos a precisão tangível da IA na projeção do sucesso de músicas. Este estudo quantifica as dimensões intrincadas das composições musicais, demonstrando o potencial transformador da IA na indústria da música.

Em termos práticos, esta pesquisa representa um avanço substancial nos procedimentos preditivos e de aprimoramento da indústria da música. Ela fornece conclusões respaldadas por dados que podem orientar profissionais da indústria e pesquisadores.

A notável precisão preditiva do modelo, caracterizada por consistentemente baixos escores de Erro Absoluto Médio (MAE), destaca sua adequação para tarefas de regressão. Este sucesso é uma consequência direta do ajuste meticuloso de parâmetros e da adaptabilidade das arquiteturas de redes neurais.

Este trabalho demonstra as significativas capacidades preditivas dos modelos de IA quando aplicados às características musicais. A análise revelou várias conclusões importantes, como a importância de considerar o fator temporal, a variabilidade e a capacidade de melhoria do modelo.

A análise detalhada de períodos de 10 anos e 5 anos destacou a importância dos fatores temporais na previsão de tendências musicais. Esses resultados mostraram que o modelo apresenta bom desempenho quando há uma sobreposição temporal entre os dados



de treinamento e de teste, capturando tendências consistentes na indústria da música dentro desses períodos.

Em conclusão, este estudo apresenta um framework concreto para o potencial da IA em revolucionar a previsão de sucesso de músicas e estratégias de otimização, indo além dos paradigmas convencionais de classificação de hits. À medida que avançamos, a integração de modelos de IA promete elevar a precisão e a eficácia da indústria da música, fornecendo ferramentas úteis para artistas, produtores e executivos que navegam na indústria.

Agradecimentos

Ao Instituto Presbiteriano Mackenzie, pelo apoio financeiro e pela concessão da bolsa PIBIC, que foi essencial para a realização desta iniciação científica.

Ao Prof. Dr. Eurico Luiz Prospero Ruivo, pela orientação e suporte inestimáveis durante o desenvolvimento desta iniciação científica. Sua expertise, conselhos e dedicação foram fundamentais para o sucesso deste trabalho.

À Sophia de Oliveira Pimentel Souza Lima, pela sua participação ativa na elaboração e pesquisa dos artigos de submissão e científico.

Ao Prof. Dr. Leandro Carlos Fernandes, pelas dicas e valiosos comentários em conversas casuais.

Ao Prof. Dr. Orlando Bisacchi Coelho, pelas sugestões, ajuda e apoio fornecido em conversas informais.

Dedicatória

Por fim, dedico este trabalho aos meus queridos irmãos, André e Tiago Markoulakis, que sempre me incentivaram com sua admiração e confiança na realização de minha jornada. À minha mãe, que esteve ao meu lado em minha trajetória profissional, acadêmica e pessoal, oferecendo seu apoio e amor incondicional. E ao Professor Doutor Orlando Bisacchi Coelho, cuja paixão pela ciência e educação sempre me inspirou e que, infelizmente, nos deixou no começo deste ano.

6. Referências

AVLONITIS, George; **KARAYANNI**. The Impact of Internet Use on Business-to-Business Marketing: Examples from American and European Companies. Set. de 2000. Disponível



em:

https://www.sciencedirect.com/science/article/pii/S0019850199000711?casa_token=EesfNQWM9UwAAAAA:3QCogreht0Gd1A254Tg6nUdoxImWTxeQmxZ4hreBAAdly7L_uncBOSManbwKLSJWVvorl8LfBPpwk. Acesso em: 28 mar. 2023.

CIBILS, Cristian; et al. Predicting a Song's Path through the Billboard Hot 100. Stanford.edu, 12 dez. 2015. Disponível em: https://cs229.stanford.edu/proj2015/012_report.pdf. Acesso em: 16 set. 2023.

HERREMANS, Dorien; et al. Dance Hit Song Prediction. Journal of New Music Research, vol. 43, no. 3, 3 jul. 2014, pp. 291–302. Disponível em: <https://doi.org/10.1080/09298215.2014.881888>. Acesso em: 16 set. 2023.

DIMOLITSAS, Ioannis; et al. SpotHitPy: A Study for ML-Based Song Hit Prediction Using Spotify. ArXiv (Cornell University), 19 jan. 2023. Disponível em: <https://doi.org/10.48550/arxiv.2301.07978>. Acesso em: 16 set. 2023.

JUN, Jong-Kun; PARK, Cheol. A Cross-Cultural Comparison of Internet Buying Behavior: Effects of Internet Usage, Perceived Risks, and Innovativeness. Out. 2003. Disponível em: https://www.emerald.com/insight/content/doi/10.1108/02651330310498771/full/html?casa_token=B7yJl4rhJ1EAAAAA:izlkHHe4G3myhOmmrKwz99H_uZx5MzuFA7fipfKeUH9LcHWUbygYMm2cmhf52xcmHNx8OnzHeUkRfhzeDp5lydMnqWR2INRY_BezDDupbeOGy1RWPxeX. Acesso em: 28 mar. 2023.

MARTENS, David; **HERREMANS**, Dorien; **SÖRENSEN**, Kenneth. Classification and Generation of Composer-Specific Music Using Global Feature Models and Variable Neighborhood Search. Set. 2015. Disponível em: <https://ieeexplore.ieee.org/document/7258427>. Acesso em: 28 mar. 2023.

MARTENS, David; **HERREMANS**, Dorien; **SÖRENSEN**, Kenneth. Dance Hit Song Prediction. Set. 2014. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/09298215.2014.881888>. Acesso em: 28 mar. 2023.

MIDDLEBROOK, Kai; **SHEIK**, Kian. Song Hit Prediction: Predicting Billboard Hits Using Spotify Data. ArXiv (Cornell University), 22 ago. 2019. Acesso em: 16 set. 2023.

RAJYASHREE, R.; et al. Predicting Hit Music Using MIDI Features and Machine Learning. IEEE Xplore, 1 out. 2018. Disponível em: <https://ieeexplore.ieee.org/abstract/document/8724001>. Acesso em: 10 nov. 2020.



RAZA, Agha Haider; **NANATH**, Krishnadas. Predicting a Hit Song with Machine Learning: Is There an A Priori Secret Formula? Jul. 2020. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9190613>. Acesso em: 28 mar. 2023.

RIAA. Investing in Music. Washington, D.C., 2013. Nov. 2014. Disponível em: <https://www.riaa.com/reports/investing-in-music/>. Acesso em: 28 mar. 2023.

YANG, Li-Chia; et al. Revisiting the Problem of Audio-Based Hit Song Prediction Using Convolutional Neural Networks. ArXiv (Cornell University), 5 abr. 2017. Disponível em: <https://doi.org/10.48550/arxiv.1704.01280>. Acesso em: 16 set. 2023.

ZHAO, Mengyisong; et al. An Analysis of Classification Approaches for Hit Song Prediction Using Engineered Metadata Features with Lyrics and Audio Features. 10 mar. 2023.

Contatos: contato.marcoulakis@gmail.com e eurico.ruivo@mackenzie.br