

UNIVERSIDADE PRESBITERIANA MACKENZIE
FACULDADE DE COMPUTAÇÃO E INFORMÁTICA
SISTEMAS DE INFORMAÇÃO

07K

Inteligência Artificial

Relatório do Projeto

1) Aplicação de MIR (Music Information Retrieval) e IA no auxílio à composição de peças musicais

2) Integrantes

- Guilherme Augusto de Lima Gavazzi, 10403271@mackenzista.com.br
- Guilherme Mello Florio, 10400766@mackenzista.com.br
- Matheus Yudi Akizuki, 10400810@mackenzista.com.br

3) Resumo

A composição de uma peça musical pode ser, por muitas razões, um processo trabalhoso e devagar, comumente desmotivando aspirantes à música a continuarem suas produções. Em momentos de desesperança, muitos dos envolvidos acabam por buscar ferramentas que os auxiliem a progredir, seja por meio de feedback digital ou análise e sugestões. Com o avanço da Inteligência Artificial, uma série de tecnologias pertinentes a essa situação surgem e clamam por uma nova abordagem à composição musical assistida por algoritmos. É nesse cenário que propomos uma ferramenta capaz de, através de input MIDI (Musical Instrument Digital Interface) do usuário, sugerir progressões à melodia ou aos acordes, analisando e seguindo o estilo imputado.

4) Introdução

a. Contextualização

A música faz parte da história humana há pelo menos 40 mil anos (WASSILIWIZKY; MENNINGHAUS, 2021), de onde se data o primeiro instrumento musical encontrado. Durante esses milênios de história da música a mesma teve que se reinventar diversas vezes, gerando padrões e gêneros, e, posteriormente, algoritmos para guiar a composição, conhecidos como a teoria musical. A teoria musical ocidental é estruturada de formas específicas, como harmonia (SCHUBERT; STEVENS, 2006), escalas, claves, acordes e suas progressões. Um dos principais instrumentos musicais que é utilizado para visualizar e estudar essa é o piano ou o teclado, uma vez que é mais fácil visualizar a divisão de tons de uma maneira intuitiva.

Criar e compor música é um processo muito humano, entretanto, por falta de conhecimentos estruturais de teoria musical e bloqueios criativos, isso pode atrasar ou desanimar um compositor menos experiente (NEWMAN et al, 2023). Dessa forma seria conveniente para este compositor ter alguém para complementar seu conhecimento e dar reconhecimento do gênero musical para o que está tentando compor.

b. Justificativa

Nestes casos, coisas simples como uma nova ideia apresentada por terceiros ou uma nova perspectiva no mesmo trabalho servem como uma ferramenta para superar este bloqueio (UoA, 2022). Tendo isso em mente, o conceito de um assistente digital capaz de auxiliar seu usuário sobre maneiras de como prosseguir em seu esforço criativo se apresenta como uma solução interessante para este problema. A ideia de juntar um modelo de Inteligência Artificial (IA) com o ato de compor música é uma ideia antiga, com pesquisas e artigos relacionados ao tópico sendo publicados desde os anos 60 (BERZ; BOWMAN, 1995). Ademais, com o crescente interesse em IA presente na atualidade desde o lançamento público e popular de plataformas como ChatGPT, era só questão de tempo até que surgissem tentativas de utilizar IA generativa com o intuito de criar música, com instruções sendo providas por textos, áudios ou através da análise de imagens providas pelos usuários. Entretanto, ao pesquisar sobre aplicativos já existentes e artigos e publicações relevantes ao assunto, nota-se a preferência para a geração de peças criadas completamente por IA.

c. Objetivo

Este projeto busca iterar sobre ideia de um programa capaz ajudar compositores a escaparem seus bloqueios mentais, utilizando pequenas melodias que já escreveram através o uso de inputs MIDI (Musical Instrument Digital Interface) e

recebendo informações relevantes a esta composição, o programa será capaz de analisar sua estrutura e então dar feedback a seu usuário através do uso de IA e recuperação de informação musical (MIR, do inglês Music Information Retrieval), discutindo possíveis caminhos para continuar sua peça e provendo insights sobre aquilo já escrito, uma solução que remonta todo o contexto do compositor e da sua necessidade.

d. Opção do projeto

A opção escolhida foi a integração de uma API de um grande modelo de linguagem no código, para auxiliar melhor o processo de análise musical.

5) Descrição do problema

A música faz parte do dia a dia da grande maioria dos Brasileiros (ABRAMUS, 2024) e da população do mundo moderno, seja de modo passivo, ao andar pela rua ou jogando algum jogo, ou de modo ativo, via aplicativos de streaming e rádio, sendo claro que esta linguagem universal que nos acompanha desde a era pré-histórica (MORLEY, 2013) se mantém altamente relevante na atualidade. Faz sentido então que parte desta população busque interagir de modo mais direto com este campo da arte através da criação de suas próprias peças musicais. Contudo, mesmo entre aqueles educados em teoria musical, a tarefa de compor uma melodia completa, quem dirá uma música com múltiplos instrumentos e até vocais, tende a ser bem mais complexa na prática do que na teoria, visto que até mesmo grandes mestres clássicos como Rachmaninoff (WINSTON, 2010) e sucessos modernos como Adele (CLASH NEWS, 2008) já enfrentaram momentos nos quais se viam impossibilitados de realizar seus desejos de compor. Seja devido a seu estado mental, ambiente físico e social ou até mesmo por falta de conhecimento. Ser incapaz de concluir uma atividade, especialmente algo tão pessoal e querido a tantos indivíduos como música, se apresenta como frustrante e desencorajador.

6) Discutir a respeito dos aspectos Éticos do uso da IA e a sua Responsabilidade no desenvolvimento da solução;

O uso de IA, no âmbito musical, é bastante controverso e levanta uma série de questões em relação à ética, sendo um dos principais pontos de debate do nosso projeto. A IA é majoritariamente usada atualmente na música por duas funções: Criar uma música a partir de um prompt, ou cantar uma música na voz de outra pessoa, que, apesar de serem propostas interessantes, acabam acarretando controvérsias tanto éticas quanto legais (DROTT, 2020). Nosso projeto busca uma abordagem diferente, no que ele apoia o compositor com sugestões de acordes e melodias nas quais ele pode

se inspirar e criar algo novo com base no que foi fornecido. Mas ainda assim, existem algumas questões éticas a serem tratadas quando se fala de geração de melodias e acordes. Como exemplo: Um possível viés ou problema de overfitting pode causar a cópia acidental uma melodia exata de uma música real, levantando questões legais que indicam uma responsabilidade nossa na hora de montar datasets a serem utilizados pelo nosso sistema.

7) Dataset, se for o caso (anonimizados quando necessário), descrição detalhada do seu conteúdo/origem, análise exploratória e preparação dos dados em Python;

<https://www.kaggle.com/datasets/soumikrakshit/classical-music-midi/data>

<https://www.kaggle.com/datasets/kritanjali/jain/maestropianomidi/code>

<https://www.kaggle.com/datasets/jackvial/themaestrodatasetv2>

<https://www.kaggle.com/datasets/bhavesjain/midi-data>

<https://www.kaggle.com/datasets/gautamgc75/classical-piano-midi-music>

<https://www.kaggle.com/datasets/charliecodex/classical-music>

Os Datasets acima foram encontrados no Kaggle e contém uma série de arquivos MIDI de peças de valsa erudita (e pseudo midi, em csv) que serão utilizados em testes do algoritmo de IA, para comparação e análise.

Análise exploratória disponível no Github: <https://github.com/guilherme-mello-florio/silksong-composition-helper>

8) Metodologia e Resultados Esperados: apresentar a abordagem que pretende empregar na resolução do problema e quais são os resultados esperados.

Levando em conta a exploração e análise de dados em um ambiente computacional com variáveis controláveis, a pesquisa cobrirá a área das Ciências Exatas e da Terra, de natureza aplicada e propósito descritivo. Será realizada em laboratório, em um ambiente experimental, com dados de caráter quantitativo. Será realizada com métodos mistos, fazendo uso de ground-theory e experimentos com o modelo gerado (GIL, 2022).

Inicialmente, foram realizadas pesquisas sobre trabalhos similares e relevantes nas bases de dados apontadas (SCOPUS, Web of Science, ISMIR). Tendo adquirido um ponto norteador para o projeto, foram levantados os métodos (IBM, 2021) (IBM; Stryker, 2024), aplicações (Thickstun; Harchaoui; Kakade, 2016) (Roberts et al, 2019) e

bibliotecas (Sun; Cuthbert, 2018) (Raffel; Ellis, 2014) (Weiß; Peeters 2022) a serem utilizadas, com estas informações sendo revisadas e, então, consolidadas com o intuito de gerar uma estimativa para o processo de desenvolvimento do projeto.

A primeira etapa estimada se refere a criação ou adaptação de algum sistema com o intuito de receber inputs MIDI para a aplicação. Foi escolhido um possível candidato: uma junção das bibliotecas para Python Music21 (Sun; Cuthbert, 2018) e pretty-midi (Raffel; Ellis, 2014), que, juntas, seriam capazes de manipular e analisar dados MIDI. Se aproximando da conclusão desta etapa, será iniciado o tratamento da base de dados escolhida com o intuito de formatá-la de modo com que esta seja compatível ao formato estabelecido pela etapa anterior.

Em paralelo, conforme os dados forem sendo formatados, estes serão imediatamente utilizados no modelo do agente escolhido e utilizando bases de dados provenientes da plataforma Kaggle (Adegoke, 2023), devido ao vasto número de bases referentes a piano clássico, visto que o foco desta prova de conceito será a valsa erudita (Buja, 2022), com outras possíveis fontes disponíveis na MusicNet (Thickstun; Harchaoui; Kakade, 2016), o que necessitará do uso da biblioteca Python Librosa (Weiß; Peeters 2022).

Neste ponto, que corresponde à etapa cinco, o foco será o uso de uma LLM. Ainda pesquisaremos um agente ideal para a interpretação e análise de arquivos MIDI.

Simultaneamente a esta etapa, será criada uma interface para o programa com o intuito de simplificar o processo de geração, visando melhoria na experiência de uso e uma identidade visual para posterior versão pública do programa.

Finalmente, há de se realizar a união da aplicação concluída com a interface gerada e, então, a revisão do trabalho realizado, focando na coerência do material escrito, no debugging e na eficácia do sistema em sua proposta.

Ao final de todo o processo, pretendemos chegar a uma aplicação pronta capaz de receber inputs MIDI do usuário, processá-los através de uma ferramenta de IA, e devolver feedback que envolva sugestões e faíscas para despertar a criatividade do próprio usuário, auxiliando-o a superar seu bloqueio no processo de criação.

9) Referências: citadas dentro do texto do projeto / 10) Bibliografia: na qual o grupo se baseará para criar e desenvolver o projeto e/ou Referências: citadas dentro do texto do projeto

ABADI, Martín et al. TensorFlow: A System for Large-Scale Machine Learning. In: USENIX SYMPOSIUM ON OPERATING SYSTEMS DESIGN AND IMPLEMENTATION (, 12., 2016, Savannah, GA, USA. Proceedings [...]. [S. l.: s. n.], 2016. Disponível em:

<https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf>. Acesso em: 13 nov. 2024.

ADEGOKE, Joshua. A Beginner's Guide to Kaggle for Data Science. [S. l.], 17 abr. 2023. Disponível em: <https://www.makeuseof.com/beginners-guide-to-kaggle/>. Acesso em: 13 nov. 2024.

Consumo de Música no Brasil. Abramus - Associação Brasileira de Música e Artes, Disponível em: <https://www.abramus.org.br/noticias/16444/consumo-de-musica-no-brasil/>. Acesso em: 20 set. 2024.

BERZ, William L.; BOWMAN, Judith. An Historical Perspective on Research Cycles in Music Computer-Based Technology. Bulletin of the Council for Research in Music Education, Illinois, n. 126, p. 15-28, Outono 1995.

BUJA, Maureen. Dance, Dance, Dance: The Waltz. [S. l.], 4 nov. 2022. Disponível em: <https://interlude.hk/dance-dance-dance-the-waltz/>. Acesso em: 13 nov. 2024.

DROTT, Eric. Copyright, compensation, and commons in the music AI industry. Creative Industries Journal, [s. l.], v. 14, ed. 2, p. 190-207, 29 out. 2020. DOI 10.1080/17510694.2020.1839702. Disponível em: <https://doi.org/10.1080/17510694.2020.1839702>. Acesso em: 20 set. 2024.

GIL, Antonio C. Como Elaborar Projetos de Pesquisa. 7th ed. Rio de Janeiro: Atlas, 2022. E-book. p.1. ISBN 9786559771653. Disponível em: <https://app.minhabiblioteca.com.br/reader/books/9786559771653/>. Acesso em: 17 nov. 2024.

IBM. What is machine learning (ML)?. [S. l.], 22 set. 2021. Disponível em: <https://www.ibm.com/topics/machine-learning>. Acesso em: 13 nov. 2024.

IBM; STRYKER, Cole. What is a recurrent neural network (RNN)?. [S. l.], 4 ago. 2024. Disponível em: <https://www.ibm.com/topics/recurrent-neural-networks>. Acesso em: 13 nov. 2024.

MORLEY, Iain. *The Prehistory of Music: Human Evolution, Archaeology, and the Origins of Musicality*. Oxford: Oxford University Press, 2013. ISBN 9780199234080.

NEWMAN, Michele; MORRIS, Lidia; LEE, Jin Ha. Human-AI Music Creation: Understanding the Perceptions and Experiences of Music Creators for Ethical and Productive Collaboration. *Proceedings of the 24th International Society for Music Information Retrieval Conference*, p. 80-88, 2023. Disponível em: <https://doi.org/10.5281/zenodo.10265227>. Acesso em: 20 set. 2024.

NICODEMUS, Jonas. PINNs-based MPC: First release. 2021. Disponível em: <https://doi.org/10.5281/zenodo.5520662>. Acesso em: 20 set. 2024.

RAFFEL, Colin; ELLIS, Daniel P. W. INTUITIVE ANALYSIS, CREATION AND MANIPULATION OF MIDI DATA WITH pretty_midi. [S. l.], 2014. Disponível em: <https://colinraffel.com/publications/ismir2014intuitive.pdf>. Acesso em: 13 nov. 2024.

ROBERTS, J.; ENGEL, J.; MANN, Y.; GILLICK, J.; KAYACIK, C.; NÜRLY, S.; DINCULESCU, M.; RADEBAUGH, C.; HAWTHORNE, C.; ECK, D. Magenta Studio: Augmenting creativity with deep learning in Ableton Live. In: *Proc. International Workshop on Musical Metacreation (MUME)*, 2019. Disponível em: <https://doi.org/10.5281/zenodo.428526>. Acesso em: 20 set. 2024.

SCHUBERT, Emery; STEVENS, Catherine. The effect of implied harmony, contour and musical expertise on judgments of similarity of familiar melodies. *Journal of New Music Research*, v. 35, n. 2, p. 161-174, 2006. Disponível em: <https://doi.org/10.1080/09298210600835000>. Acesso em: 20 set. 2024.

SUN, S. H.; CUTHBERT, M. S. Emotion Painting: Lyric, Affect, and Musical Relationships in a Large Lead-Sheet Corpus. *Empirical Musicology Review*, v. 12, n. 3-4, p. 327-348,

2018. Disponível em: <https://doi.org/10.18061/emr.v12i3-4.5889>. Acesso em: 20 set. 2024.

THICKSTUN, John; HARCHAOUI, Zaid; KAKADE, Sham M. MusicNet. [S. l.], 30 nov. 2016. Disponível em: <https://zenodo.org/records/5120004#.Yhxr0-jMJBA>. Acesso em: 13 nov. 2024.

WASSILIWIZKY, Eugen; MENNINGHAUS, Winfried. Why and How Should Cognitive Science Care about Aesthetics? Trends in Cognitive Sciences, v. 25, n. 6, p. 437-449, 2021. ISSN 1364-6613. Disponível em: <https://doi.org/10.1016/j.tics.2021.03.008>. Acesso em: 20 set. 2024.

WEIß, C.; PEETERS, G. Comparing Deep Models and Evaluation Strategies for Multi-Pitch Estimation in Music Recordings. IEEE/ACM Transactions on Audio, Speech, and Language Processing, v. 30, p. 2814-2827, 2022. DOI: 10.1109/TASLP.2022.3200547.

Winston, Robert. How great artists have fought creative block. BBC News, 27 jul. 2010. Disponível em: <https://www.bbc.com/news/magazine-10766308>. Acesso em: 20 set. 2024.

ZHANG, Weiwei. Automatic Synthesis Technology of Music Teaching Melodies Based on Recurrent Neural Network. Scientific Programming, v. 2021, p. 1-10, 2021. Disponível em: <https://doi.org/10.1155/2021/1704995>. Acesso em: 20 set. 2024.

Adele On Writers Block. Clash News, 13 nov. 2008. Disponível em: <https://www.clashmusic.com/news/adele-on-writers-block/>. Acesso em: 20 set. 2024.

Overcoming writer's block. University of Amsterdam, 21 jul. 2022. Disponível em: <https://student.uva.nl/en/topics/overcoming-writer-s-block>. Acesso em: 20 set. 2024.