

**HarmonIA: um software experimental para visualização e manipulação
interativa de sucessões de acordes geradas por um modelo preditivo
conexionista**

Eixo temático: Interdisciplinaridades teórico-analíticas

Palavras-chave: Harmonia tonal; análise musical assistida por computador;
inteligência artificial.

Resumo: Relato nesta comunicação o desenvolvimento de um software experimental para manipulação e visualização interativa de sucessões de acordes geradas por um modelo preditivo connexionista baseado em Redes Neurais Recorrentes com mecanismo LSTM (*Long Short-Term Memory*), solução frequentemente empregada para previsão e geração textual. O desenvolvimento partiu de um *fork* (bifurcação) da solução de código-aberto disponibilizada por KIM (2019), na qual esta combinação de algoritmos atua na geração textual - com tokenização por palavras - a partir de um modelo preditivo treinado em um *dataset* da obra poética de Shakespeare. Após adaptações pontuais, o algoritmo foi adaptado para geração de sequências textuais de cifras de acordes, e diversos modelos preditivos foram treinados a partir de *datasets* obtidos por da mineração de dados - com posterior filtragem e tratamento - de um popular site brasileiro de cifras. Foram criados modelos para diferentes tonalidades e artistas diversos, visando uma exploração livre dos resultados em um esforço de musicologia empírica (vide COOK, 2014).

Modelos preditivos connexionistas são sistemas frequentemente descritos como *black box* (caixa-preta), cujo funcionamento (objetivamente, os *pesos* e conexões entre as camadas de uma rede neural) é pouco transparentes e difícil de ser interpretado. Cada modelo preditivo carrega em si predileções, vieses (*biases*) e potenciais erros difíceis de serem detectados, o que torna a avaliação dos modelos e reconhecimento de resultados anômalos uma tarefa especialmente complicada, demandando conhecimentos aprofundados no domínio específico dos dados utilizados. Outra metáfora - mais lúdica - utilizada para descrever as dificuldades de avaliação de modelos computacionais em música tem sido o “cavalo”, em alusão ao cavalo “Hans Esperto” que supostamente sabia fazer contas e responder a perguntas (STURM, 2014). Em face dessa demanda, o conhecimento do teorista-analista, acostumado em lidar com diferentes tipos de padrões musicais, certamente ocupa um espaço privilegiado na tarefa de avaliar modelos

computacionais para geração e modelagem musical. Alguns pesquisadores tem validado seus modelos gerativos a partir de variações do *Teste de Turing* (veja HADJERES, PACHET e NIELSEN, 2016), que, supostamente, afere a indistinguibilidade dos produtos gerados em relação a uma produção humana, porém tais métodos dificilmente aferem o potencial real destas ferramentas para um usuário final (o músico). Cientes dessa inadequação, alguns pesquisadores tem experimentado a formação de equipes interdisciplinares para avaliar os resultados em contextos musicais mais concretos (STURM et. al, 2019).

Em direção a este objetivo exploratório de avaliar concretamente os resultados da geração musical por modelos conexionistas a partir da perspectiva de um usuário final (o músico), buscou-se desenvolver no software HarmonIA uma interface para visualização e audição em tempo real das hipóteses e probabilidades geradas pelos modelos preditivos treinados. Nesta interface (Figura 1), as continuidades hipotéticas para uma sequência de acordes são dispostas em um *squaremap* proporcional, distinguindo visualmente os eventos seguintes mais prováveis dos eventos seguintes menos prováveis. Desta forma, o propósito do software desenvolvido se direciona para manipulação livre das hipóteses sugeridas - *down the rabbit hole* – em busca de uma avaliação (intuitiva e subjetiva) do potencial da geração musical assistida por algoritmos de inteligência artificial, seja para compreensão teórica de regras estilísticas ou criação musical assistida. Para auxiliar este propósito avaliativo, foram acrescentadas outras ferramentas analíticas, tais como processamento de linguagem natural para agrupamento de sucessões de acordes mais comuns (Figura 2). Com estas ferramentas, busca-se visualizar padrões nos modelos treinados e possíveis predileções estilísticas. Foi criada também uma interface para comparação simultânea entre modelos preditivos (Figura 3), que se provou bastante útil para a investigação dos resultados obtidos a partir do ajuste dos hiperparâmetros (GOODFELLOW, BENGIO, COURVILLE, 2014) de treinamentos da rede neural, porém menos proveitoso na comparação entre modelos devido às diferenças de vocabulário (itens *tokenizados*). Para tornar a manipulação das sucessões de acordes mais aprazível, foram também mineradas as posições dos acordes no violão do mesmo site de cifras, as quais foram utilizadas para gerar resultados auditivos mais próximo dos contextos musicais modelados. As posições dos acordes podem ser alternadas em uma partitura interativa (Figura 4), porém os resultados devem ser considerados muito mais como abstrações do que modelagens efetivas de estilos. Foi implementada também uma

ferramenta de exportação para o software de notação musical aberto MuseScore, possibilitando manipulações mais livres dos resultados obtidos.

Um dos maiores desafios consistiu na construção de um *dataset* consistente. Através da mineração de dados de um popular site colaborativo brasileiro de cifras, foram obtidas, após diversas filtrações, 53270 músicas para construção dos modelos preditivos. Foram construídos duas categorias de modelos preditivos: a) tonalidades maiores e menores, usando como ponto de partida a classificação fornecida pelo próprio site e descartando de forma automatizada tonalidades erroneamente classificadas; b) sucessões de acordes nas músicas de artistas específicos, propósito que se demonstrou mais problemático em virtude da quantidade de dados necessárias para generalizações sem *overfitting* estatístico. O tratamento dos dados consistiu na limpeza por filtração (exclusão automatizada de músicas por diversos mecanismos) e uniformização (das cifras), buscando a redução do vocabulário (quantidade de cifras *tokenizadas*) dos modelos preditivos para otimização dos resultados. Foram experimentadas duas estratégias de normalização dos dados: transposição para as 12 tonalidades (como em HADJERES, PACHET e NIELSEN, *op. cit*) e normalização para Dó maior, com o propósito de viabilizar a comparação entre modelos com vocabulários distintos. No entanto, tais estratégias aparentemente esbarraram em outros vieses estatísticos, em especial a quantidade inconsistente de repetições das sequências de acordes nas diferentes músicas, o que sugere a necessidade de estratégias mais sofisticadas para a construção de *datasets* musicais para modelagem.

Os resultados obtidos até o momento apontam para uma necessidade de maior aprofundamento na avaliação de modelos conexionistas em contextos musicais concretos e corroboram a grande complexidade da construção de *datasets* apropriados em quantidade e qualidade dos dados bem como o manejo de hiperparâmetros de treinamento, com influências significativas no resultado dos modelos preditivos. Apontam também para a importância de ferramentas de controle para avaliação mais adequada dos modelos. Por fim, o pesquisador disponibiliza os *datasets* brutos e o código-fonte completo da aplicação na plataforma GitHub, sob a licença de uso permissivo MIT.