

Inteligência Artificial no ensino e aprendizagem de bateria e percussão: ferramentas atuais e perspectivas futuras

Daniel Gohn

Universidade Federal de São Carlos
orcid.org/0000-0002-3940-1229
dgohn@ufscar.br

GOHN, Daniel. Inteligência Artificial no ensino e aprendizagem de bateria e percussão: ferramentas atuais e perspectivas futuras. *Revista da Abem*, [s. l.], v. 33, n. 1, e33113, 2025.





Inteligência Artificial no ensino e aprendizagem de bateria e percussão: ferramentas atuais e perspectivas futuras

Resumo: Este artigo investiga as implicações da inteligência artificial (IA) no ensino e na aprendizagem de bateria e percussão. Inicialmente, contextualiza o surgimento e a evolução da IA, com ênfase nos modelos de linguagem de grande escala e seu impacto na Educação Musical. Em seguida, discute o potencial da IA generativa para criar materiais didáticos, apoiar o aprendizado teórico e oferecer feedback, ao mesmo tempo em que se problematizam limitações como alucinações e preconceitos. A partir de uma análise empírica de interações com o ChatGPT, avaliam-se suas capacidades de fornecer feedback sobre áudio e vídeo de performances de bateria. Os resultados indicam que, embora a IA demonstre sofisticação ao emular discursos avaliativos musicais, ainda carece de precisão contextual e sensibilidade estética para oferecer retornos pedagogicamente relevantes. Conclui-se apontando para um futuro da Educação Musical permeado pela inteligência artificial, destacando a importância do letramento em IA para que educadores e estudantes possam utilizá-la de forma crítica, consciente e criativa na construção de práticas pedagógicas inovadoras.

Palavras-chave: Inteligência artificial, Bateria e percussão, Modelos de linguagem de grande escala, ChatGPT.

Artificial Intelligence in drum and percussion education: current tools and future perspectives

Abstract: This article investigates the implications of artificial intelligence (AI) in the teaching and learning of drums and percussion. It begins by contextualizing the emergence and evolution of AI, with an emphasis on large language models and their impact on music education. It then discusses the potential of generative AI to create educational materials, support theoretical learning, and provide feedback, while also addressing limitations such as hallucinations and biases. Based on an empirical analysis of interactions with ChatGPT, the study evaluates its ability to offer feedback on audio and video recordings of drum performances. The findings indicate that, although AI demonstrates sophistication in emulating evaluative musical discourse, it still lacks contextual accuracy and aesthetic sensitivity to provide pedagogically meaningful responses. The article concludes by pointing to a future of music education shaped by artificial intelligence, highlighting the importance of AI literacy so that educators and students can engage critically, consciously, and creatively with these tools in the development of innovative pedagogical practices.

Keywords: Artificial intelligence, Drums and percussion, Large language models, ChatGPT.

Inteligencia Artificial en la enseñanza y el aprendizaje de batería y percusión: herramientas actuales y perspectivas futuras

Resumen: Este texto investiga las implicaciones de la inteligencia artificial (IA) en la enseñanza y el aprendizaje de batería y percusión. En primer lugar, contextualiza el surgimiento y la evolución de la IA, con énfasis en los modelos de lenguaje de gran escala y su impacto en la educación musical. A continuación, se analiza el potencial de la IA generativa para crear materiales didácticos, apoyar el aprendizaje teórico y ofrecer retroalimentación, al mismo tiempo que se abordan limitaciones como las alucinaciones y los sesgos. A partir de un análisis empírico de interacciones con ChatGPT, se evalúa su capacidad para ofrecer retroalimentación sobre grabaciones de audio y video de interpretaciones de batería. Los resultados indican que, aunque la IA demuestra sofisticación al emular discursos evaluativos musicales, todavía carece de precisión contextual y sensibilidad estética para brindar respuestas pedagógicamente significativas. El artículo concluye señalando un futuro de la educación musical permeado por la inteligencia artificial, destacando la importancia del alfabetismo en IA para que educadores y estudiantes puedan utilizar estas herramientas de manera crítica, consciente y creativa en la construcción de prácticas pedagógicas innovadoras.

Palabras clave: Inteligencia artificial, Batería y percusión, Modelos de lenguaje de gran escala, ChatGPT.





Introdução

O campo de pesquisa acadêmica sobre inteligência artificial (IA) surgiu na década de 1950, quando Alan Turing propôs o “Teste de Turing”, uma forma de medir a capacidade das máquinas para se passar como humanos em trocas de textos escritos (Turing, 1950). Desde então, as tecnologias denominadas como IA avançaram significativamente em determinados momentos, como no surgimento do reconhecimento facial digital, mas houve longos períodos de estagnação, chamados por pesquisadores de “invernos de IA” (Molick, 2024).

Nas duas primeiras décadas dos anos 2000, tornaram-se comuns aplicativos como Google Maps e Waze, com informações de trânsito em tempo real, tendo como base a inteligência artificial preditiva. Estes são sistemas que utilizam dados históricos e algoritmos de aprendizado de máquina (machine learning) para fazer previsões sobre eventos futuros ou identificar padrões que ainda não ocorreram (Hancock; Ren, 2024). No mesmo período, também foram difundidos sistemas de recomendação em serviços de streaming como Spotify e YouTube, utilizando outro tipo de inteligência artificial, chamada discriminativa, na qual o aprendizado de máquina se concentra em identificar fronteiras entre classes ou categorias, classificando dados a partir de critérios pré-estabelecidos (Maheshwari, 2023; Convington; Adams; Sargin, 2016). Embora essas tecnologias tenham alcançado amplo reconhecimento na sociedade, o termo “inteligência artificial” ainda não era facilmente encontrado fora da ficção científica ou de círculos acadêmicos.

Tal cenário foi drasticamente modificado com a chegada dos modelos de linguagem de grande escala (large language models), trazendo à discussão a inteligência artificial generativa. Esta é voltada para a geração de novos conteúdos (texto, imagem, áudio) a partir de dados de treinamento (Feuerriegel *et al.*, 2024). Em novembro de 2022, a empresa OpenAI tornou público o ChatGPT 3.5 e em março de 2023 houve uma atualização para o ChatGPT 4. A grande diferença entre as duas versões criou a expectativa de um desenvolvimento rápido para novas ferramentas, tornando urgente a discussão sobre um letramento em inteligência artificial.

Diversos estudos começaram a investigar os impactos das novas ferramentas de IA na educação (Grassini, 2023; Chiu, 2023). No campo da música, a





pesquisa sobre IA cresce rapidamente, com diversos temas ainda em aberto para explorações. Entre esses temas estão as implicações que a inteligência artificial traz para os processos de ensino e aprendizagem de instrumentos musicais.

Objetivando contribuir para a discussão, o presente texto concentra-se principalmente – porém não exclusivamente – na IA generativa, apresentando o olhar de um professor da Universidade Federal de São Carlos (UFSCar), a partir da observação de seus alunos em conjunto com a recente literatura da área, assim como algumas conjecturas que poderão ser realidade em um futuro breve. Como apoio para esse exercício, antes das considerações finais, são apresentadas trocas de ideias com o ChatGPT, que se mostram significativas para a compreensão do atual momento tecnológico.

Modelos de Linguagem de grande escala

O ChatGPT, assim como Claude, Gemini e LLaMa, entre outros, são modelos de linguagem de grande escala, sistemas de inteligência artificial generativa baseados em redes neurais profundas que aprendem a compreender e gerar textos através do processamento de enormes quantidades de dados durante um treinamento. Esses modelos utilizam uma arquitetura chamada *Transformer*, que analisa as relações entre palavras em uma frase, permitindo a compreensão do contexto e do significado de cada palavra com base em sua posição em relação a outras palavras.

Durante o treinamento, são aprendidos padrões estatísticos na linguagem, prevendo a próxima palavra em bilhões de frases, desenvolvendo gradualmente uma compreensão implícita de gramática, semântica, conhecimento factual e até mesmo raciocínio. O resultado é um sistema “inteligente”, capaz de colocar questões em contextos apropriados e gerar texto coerente, funcionando essencialmente como um sofisticado sistema de preenchimento automático, assim como já existia nos aparelhos celulares e em formulários on-line.

Além de produzir conteúdo textual, a IA generativa também existe em sistemas multimodais capazes de gerar imagens e vídeos. A partir de instruções textuais (prompts), podem gerar pixels e frames, produzindo cenas ultrarrealistas. Por último, há ferramentas que aplicam mecanismos e arquiteturas derivadas dos





modelos de linguagem para aprender padrões e estruturas musicais, permitindo a geração de músicas. Esses sistemas geralmente utilizam arquiteturas de rede neural, como *Transformers*, mas são adaptadas para lidar com sequências musicais em vez de linguagem textual. Por meio desse conjunto de tecnologias, chegamos a uma situação na qual o território da produção artística, antes dominado somente por autores humanos, passa a ser ocupado também por máquinas.

O treinamento com dados da internet e documentos diversos implica em eventuais preconceitos e erros fatuais, que tornam tais modelos de linguagem perigosos em alguns momentos. Muitas vezes os erros são decorrentes das alucinações – um termo técnico usado para descrever respostas incorretas ou inventadas que foram geradas pelos modelos de linguagem, quando dados fictícios são misturados à realidade. Em certas situações, essa capacidade dos modelos de linguagem de gerar informações não-fatuais pode ser útil, pois se assemelha à qualidade humana que chamamos de “imaginação” (Hoffman, 2023).

Por exemplo, ao solicitar que um modelo de linguagem ofereça diferentes ideias para o parágrafo seguinte de um texto que estamos escrevendo, ampliamos nossa capacidade humana de pensar criativamente e conceitualmente. Neste contexto, uma alucinação pode ser bem-vinda, pois há potencial para caminhos inovadores e inusitados. Podemos pensar nas implicações para roteiros de filmes, design de jogos e anúncios publicitários.

No entanto, as alucinações geram questionamentos quanto à confiabilidade de respostas da IA generativa e suas possíveis consequências. Casos absurdos são facilmente detectáveis, como a sugestão da IA do Google para grudar queijo na massa da pizza com cola não-tóxica ou para todo humano comer uma pedra por dia (McMahon; Kleinman, 2024). Por outro lado, informações plausíveis, mas incorretas, são problemáticas, porque são difíceis de detectar, já que os modelos de linguagem as apresentam com autoridade convincente – por exemplo, citações de professores fictícios advindos de universidades que não existem, o que têm sido uma experiência comum para pesquisadores acadêmicos que utilizando o ChatGPT.

Preconceitos emergem como reflexo da ampla variedade de fontes públicas de texto com as quais os modelos de linguagem são alimentados, como livros, artigos, sites e outras bases de conhecimento, perpetuando visões estreitas em





relação a questões de gênero e raça. Por exemplo, a profissão de enfermeiro foi relacionada ao sexo feminino em 98% de cenários hipotéticos construídos por três diferentes modelos de linguagem: GPT-4, Gemini-pro e Llama-2-70B (Menz *et al.*, 2024). Essas visões restritivas também são perceptíveis no campo da música, no qual plataformas de IA, como Suno (www.suno.com) e Udio (www.udio.com), possibilitam a geração de músicas a partir de especificações textuais.

Esses modelos têm como base as estruturas rítmicas e tonais da música ocidental e utilizam tais padrões mesmo quando a solicitação envolve gêneros não-ocidentais, como música indiana ou do Oriente Médio (Mehta *et al.*, 2025). Evidência anedótica também revela que a música brasileira é subrepresentada no treinamento dessas plataformas. Por exemplo, no Suno, a solicitação de uma música de baião resulta atualmente em ritmos e sonoridades bastante distantes dos elementos característicos desse gênero.

Visando o uso de inteligência artificial na Educação, esforços para minimizar problemas como alucinações e preconceitos têm sido realizados, como a construção de modelos personalizados, nos quais há limites de segurança e “redes de proteção” – mecanismos que orientam e regulam o comportamento do sistema (Khan, 2024). Por exemplo, o Khanmigo (www.khanmigo.ai) é um sistema criado por Sal Khan, responsável também pela Khan Academy, projeto de grande reconhecimento no âmbito das tecnologias educacionais. Em sua perspectiva, a IA pode democratizar o acesso à educação de qualidade sendo um tutor individual, como Aristóteles foi para Alexandre o Grande.

Como seria se cada estudante no planeta tivesse acesso a um tutor pessoal com inteligência artificial: uma IA capaz de escrever junto com o estudante; uma IA com a qual os estudantes pudessem debater qualquer tópico; uma IA que afinasse os pontos fortes inerentes do estudante e preenchesse quaisquer lacunas de aprendizado? (Khan, 2024, p. 8).

Dentro de um pensamento socrático, o Khanmigo foi elaborado para não passar respostas diretas para as perguntas recebidas, mas sempre conduzir os alunos até que encontrem respostas por si mesmos. Da mesma forma, serviços como Magic School (www.magicschool.ai), School AI (<https://schoolai.com>) e Diffit (<https://web.diffit.me>) têm como base os modelos de linguagem para criar ambientes mais seguros em que o potencial da inteligência artificial seja explorado





por professores e estudantes. Todos possibilitam a elaboração de materiais educacionais, que podem ser facilmente adaptados de acordo com a idade dos alunos. O School AI oferece também a oportunidade de “dialogar” com personagens famosos da história mundial, como Charles Darwin ou Charlie Chaplin, por meio de treinamentos da inteligência artificial com todas as informações disponíveis sobre esses notáveis.

Interações desse tipo no School AI são controladas e minimizam riscos, ao contrário de websites que não adotam nenhum limite de segurança, como o CharacterAI (<https://character.ai>). Essa é uma plataforma que disponibiliza chatbots treinados para simular conversas com personagens diversos, sem mecanismos claros de moderação ou controle de conteúdo. Relatos indicam que a ausência desses filtros pode contribuir para consequências graves, como o caso de um jovem que cometeu suicídio após interações prolongadas com a ferramenta (Montgomery, 2024).

É fundamental a compreensão de que, ainda que possam oferecer certas proteções aos seus usuários, websites educacionais de inteligência artificial estão inseridos em um contexto de plataformização (Van Dijck; Poell; De Waal, 2018) e levantam questões sobre valores públicos, governança e responsabilidade social. Essas plataformas pertencem a corporações tecnológicas privadas que operam em uma zona regulatória cinzenta, sem garantias de que sua orientação educacional não seja enviesada por valores defendidos pelas próprias empresas. Além disso, ao interagir com tutores de IA, estudantes fornecem dados sobre seus processos de aprendizagem, frequentemente sem transparência sobre como estão alimentando bancos de dados corporativos (Alves, 2023). Portanto, há um dilema entre o potencial de democratização do acesso educacional e os interesses comerciais desse modelo de negócio, o que pode comprometer valores educacionais básicos.

Embora seja possível desenvolver nesses websites qualquer assunto em planos de aulas, desenvolvimento de rubricas e na elaboração de materiais didáticos, suas ferramentas têm como base a palavra escrita e não há funcionalidades para gerar conteúdo musical sonoro. Em outro caminho, serviços como os já mencionados Suno e Udio surgiram com o propósito de gerar música, transformando prompts em produções completas, com arranjos para diferentes





formações instrumentais. Com facilidade são geradas letras, melodias e linhas vocais, formas e harmonias, além de títulos e elementos visuais, tudo determinado pela escolha de palavras do usuário. Além das questões sobre a delimitação dos dados usados no treinamento desse tipo de IA, também são levantadas discussões sobre a democratização da produção musical, em um contexto no qual qualquer pessoa potencialmente poderá atuar como agente criador profissional de música. A situação não usual na autoria das obras criadas por IA gera dúvidas sobre como os vários agentes irão perceber a autenticidade das músicas (Tan, 2024).

Portanto, a partir das possibilidades abertas pela inteligência artificial generativa surgem diversos questionamentos, sendo o mais imediato o impacto que tais recursos terão nas profissões ligadas à música. Empresas em busca de músicas comerciais continuarão a contratar músicos de carne e osso, se podem gerar com a IA inúmeros jingles até chegar ao resultado desejado? O interesse pelo estudo de instrumentos musicais irá diminuir, se qualquer pessoa está capacitada a transformar suas ideias em música? Se esses serviços promovem uma democratização na produção musical, quais são as implicações na educação musical? Como atribuir direitos autorais para “novas” composições feitas por IA, se são geradas a partir do treinamento com obras já existentes? Qual é o processo desse treinamento e quais obras são usadas como referência?

Tais questionamentos deverão gerar profundas reflexões no universo musical. Além disso, no sentido oposto aos serviços que geram música a partir de textos, podemos ter expectativas de crescimento para o caminho inverso, no qual a IA recebe músicas, analisa o conteúdo e gera descrições, comentários e sugestões. Como um professor atento, os sistemas poderão guiar cada vez com mais precisão os passos de estudantes de música, seja em suas criações artísticas ou no aprendizado de instrumentos musicais. Um aprendiz enviará seus vídeos tocando e receberá detalhadas análises de suas performances, com a observação de pontos a melhorar e exercícios práticos para corrigir erros.

Embora situações como essa já estejam no horizonte, como iremos discutir mais adiante, a capacidade dos modelos de linguagem para lidar com a música ainda é incipiente. Em grande parte, isso é explicado pela enorme complexidade das artes musicais, que contribuem para um desenvolvimento de IA mais lento no





campo musical do que aqueles que têm como foco textos, imagens e vídeos (Tan, 2024). Assim, na trilha aberta pelos sistemas multimodais, a inteligência artificial tende a provocar transformações intensas nos processos de ensino e aprendizagem de instrumentos musicais.

IA, bateria e percussão

Após esta breve contextualização sobre os modelos de linguagem de grande escala e seu impacto educacional, nos dirigimos agora para um campo mais específico: o ensino e aprendizagem de bateria e instrumentos de percussão. Vamos começar pensando na origem desses instrumentos. Há um consenso entre os estudiosos da história da Música de que os instrumentos de percussão figuram entre os primeiros instrumentos musicais da humanidade (Blades, 1992). Primeiramente com a percussão de diferentes partes do corpo humano, com palmas e batidas no peito, e depois com o uso de objetos como pedras e pedaços de madeira, nossos ancestrais exploraram os sons em múltiplas situações e provavelmente se surpreendiam com muitos dos resultados – sons que representam o que há de mais antigo na música, em oposição às inteligências artificiais, que atualmente estão nas fronteiras das tecnologias digitais.

Embora tenham sido desenvolvidos e lapidados ao longo do tempo, os instrumentos de percussão mantêm até hoje esse elemento instintivo e visceral, contrastando com a tecnologia elaborada que envolve diversas áreas da música na atualidade. Há tecnologias avançadas na fabricação dos atuais instrumentos de percussão, visando melhores sonoridades e ergonomia, mas permanece o fato de que o som é gerado a partir de movimentos que percutem, raspam, agitam ou friccionam seus elementos. Ou seja, para aprender a tocá-los (assim como todos os outros instrumentos musicais, em alguma medida), nos remetemos a movimentos corporais e a procedimentos específicos.

Dentro de tal quadro, considerando a fisicalidade envolvida, de que formas os processos de ensino e aprendizagem de instrumentos de percussão serão modificados pelo desenvolvimento da inteligência artificial generativa? Considerando também que, como observado previamente, a IA da atualidade é





mais desenvolvida com textos, imagens e vídeos do que com a música, quais serão as repercussões de curto e longo prazo?

Nas disciplinas de bateria e percussão na UFSCar, o que pode ser constatado no presente momento é o uso pelos alunos de softwares como o Moises, um exemplo de inteligência artificial discriminativa. Programas desse tipo analisam e transformam conteúdos já existentes, sendo usados para separar faixas de áudio, removendo vocais ou quaisquer instrumentos de uma gravação (Pereira *et al.*, 2023). Tecnologias similares foram utilizadas para produzir uma música inédita dos Beatles (Savage, 2023), tornando-as bastante populares. Com esse recurso, os alunos retiram a percussão das músicas e depois sobrepõem sua interpretação, acompanhados dos outros instrumentos da gravação original. Muitas vezes, assim são produzidas as performances apresentadas no final do semestre e vídeos que servem como avaliação nas aulas.

Além disso, em anos recentes, não tem sido percebido o uso de inteligências artificiais generativas durante as aulas de bateria e percussão na UFSCar. Esse fato pode ser parcialmente explicado pela natureza prática das disciplinas, voltadas para o conhecimento procedimental com a percussão. Conforme afirma Sloboda (1985), para a aquisição de habilidades musicais é necessário desenvolver hábitos que se tornam automáticos, ou seja, há uma demanda por exercícios que são repetidos até que determinados movimentos sejam dominados. Por exemplo, para aprender a tocar samba no pandeiro é praticada uma sequência de toques, buscando a produção de boa sonoridade com pele e platinelas e um resultado com o ritmo característico, sem tensões musculares. O caminho para desenvolver essa habilidade é a prática constante ou, como se diz no jargão popular, a disposição para conseguir suas “horas de voo”.

Nesse sentido, exercícios têm sido demonstrados pelo professor e praticados pelos alunos, e o impacto que a IA generativa terá nas aulas ainda permanece como interrogação. Historicamente, as demandas na performance musical têm sido modificadas pelo desenvolvimento tecnológico, como exemplificado pelas novas técnicas de canto ocasionadas pelo microfone ou pela exigência de maior precisão temporal suscitada com o uso do metrônomo, o que nas práticas de bateria foi chamado por Zagorski-Thomas (2010) de “culto à consistência”. Possíveis novas





demandas geradas pela inteligência artificial deverão ser observadas e devidamente incorporadas em planejamentos curriculares e pedagógicos no futuro.

Além do conceito de hábito, Sloboda (1985) também afirma que o desenvolvimento de habilidades musicais envolve a passagem do conhecimento fatural (saber o que fazer) para o conhecimento procedimental (saber como fazer), e neste processo as ferramentas de inteligência artificial generativa já estão impactando o estudo da percussão. Modelos de linguagem podem ser usados para criar planos de estudo, compilar materiais didáticos, gerar ideias de práticas musicais, entre várias outras situações em que o conhecimento fatural está envolvido. Dessa forma, pode-se passar da mera repetição de algo para uma prática deliberada, guiada por desafios progressivamente mais complexos. Molick (2024) exemplifica a questão com uma comparação hipotética entre duas estudantes de piano, em que uma repete sempre os mesmos exercícios e outra é guiada por um professor, e sugere que ferramentas de IA podem participar de processos de mentoria dessa natureza.

Mais do que isso, em adição às perspectivas das ferramentas generativas, também a inteligência artificial discriminativa deve causar grandes transformações. À medida que compreenderem melhor sons e imagens, poderão proporcionar maior facilidade e eficiência nos sistemas de buscas por vídeos de percussão. Os serviços de busca usuais até o momento, como o disponível no YouTube, se mostram limitados na compreensão do que um percussionista possa estar procurando. Por exemplo, um estudante de jazz pode buscar pelo baterista Art Blakey e diversos vídeos irão surgir, mas uma pesquisa por “Art Blakey tocando um rulo na caixa” somente terá resultados significativos se um vídeo contiver essas palavras em seu título ou na descrição.

Com o desenvolvimento futuro da IA, pode-se esperar que o sistema compreenda a solicitação, identifique o músico em questão e busque por trechos de vídeos em que o elemento musical específico seja demonstrado. Assim, a mesma pesquisa poderá indicar os pontos exatos em que Art Blakey estiver tocando um rulo na caixa, facilitando o estudo desse rudimento e a passagem de uma explicação teórica para sua realização sonora.





Por último, para o desenvolvimento de habilidades musicais, Sloboda (1985, p. 210) menciona o feedback como sendo “essencial para que somente os procedimentos bem-sucedidos sejam aprendidos; qualquer procedimento que leve à repetição de fracassos é descartado”. Neste sentido, surge a interrogação da medida em que a inteligência artificial poderá contribuir na análise de alunos tocando instrumentos de percussão, para em seguida avaliar as performances e dar recomendações de como melhorar. Softwares populares como SmartMusic (atualmente chamado MakeMusic Cloud), Tonestro e Yousician já apresentavam sistemas de feedback no estudo de instrumentos musicais, mas de maneira bastante superficial, checando a acuidade de alturas e ritmos a partir de partituras determinadas. Segundo Aquilino e Scavone (2022), entre as limitações desses sistemas estão: a ideia de que para ser um bom músico basta tocar as notas certas no ritmo certo, mesmo que sem nenhuma expressividade; o objetivo que passa a ser a obtenção da maior avaliação possível, ainda que isso resulte em frases quebradas e sonoridade robótica; e o modelo estatístico que mede o progresso do estudante, que não proporciona mais elementos pedagogicamente relevantes para o crescimento artístico.

Certamente, os softwares citados se mostram como poderosos estimulantes para o estudo musical, principalmente pela oportunidade de tocar com material pré-gravado e simular a experiência de participar de uma banda. Entretanto, até o presente momento, nenhum dos exemplos mencionados oferece opções para instrumentos de percussão. Existem aplicativos como BeatMirror e liveBPM, que detectam o tempo musical e monitoram o bpm (batimentos por minuto) captado pelo microfone de smartphones, mas sem a capacidade de fornecer feedback mais significativo para percussionistas. Esse quadro deverá ser drasticamente modificado com o desenvolvimento das inteligências artificiais.

Reconhecer as diferentes sonoridades e a vasta gama de ritmos envolvidos na percussão é algo bastante intrincado, principalmente porque há diversas formas de tocar “certo”. Em diversos ritmos musicais, como no caso de ritmos brasileiros, diferentes interpretações de uma mesma notação rítmica são aceitas como válidas (Gohn, 2021). Portanto, a expectativa futura é de que a compreensão que a IA possui dessa complexidade musical aumente exponencialmente e, em vez de avaliar uma performance comparando-a com a matemática fixa de uma partitura,



oferecendo simples sinalizações de certo ou errado, possa gerar feedback a partir de amplas bases de conhecimento. Ao recomendar caminhos para aprimorar o estudo da percussão, a IA poderá analisar não apenas o som, mas também a qualidade dos movimentos realizados pelo percussionista, identificando tensões musculares, desequilíbrios e inconsistências indesejadas.

Conversas com o ChatGPT: uma análise empírica

Para compreender as potencialidades e limitações atuais da inteligência artificial generativa no contexto do ensino e aprendizagem de percussão, foram realizadas interações estruturadas com o ChatGPT (versão 4-turbo) em abril de 2025. A metodologia consistiu em três etapas: (1) solicitação de informações gerais sobre o uso de IA na educação musical, (2) envio de um arquivo de áudio de bateria para obter feedback, e (3) submissão de um vídeo da execução para avaliação da performance. Face ao prompt “liste três maneiras para professores de música utilizarem modelos de linguagem nas aulas de instrumentos musicais”, o texto elaborado pelo ChatGPT indicou o seguinte:

1. Geração de material didático personalizado – Criar partituras simplificadas, exercícios técnicos específicos para os alunos, resumos teóricos sobre harmonia, ritmo e técnica instrumental, além de elaborar roteiros para estudo e prática.
2. Apoio no aprendizado teórico e histórico – Explicar conceitos musicais de forma acessível, fornecer resumos sobre a história de determinado instrumento, analisar o contexto de uma peça musical e até criar quizzes interativos para reforçar o aprendizado.
3. Feedback e planejamento de estudos – Auxiliar na criação de planos de estudo personalizados com metas e cronogramas, sugerir repertório baseado no nível do aluno e até analisar gravações (por meio de transcrição e feedback textual) para identificar pontos de melhoria na execução (OPENAI, 2025).

Dentro dessa resposta geral, o terceiro item especificamente remete ao auxílio da IA com feedback no estudo de instrumentos, trazendo possíveis contribuições com a análise e comentários sobre gravações musicais. Nessa situação, são envolvidos sistemas multimodais de inteligência artificial, com elementos discriminativos, para compreender e classificar performances, e generativos, para gerar críticas e sugestões. Para testar essa possibilidade, foi enviado ao ChatGPT um arquivo de áudio em formato wav contendo uma gravação de bateria, produzida para uma música de 4 minutos e 40 segundos de duração,



com a solicitação de “analisar a qualidade da performance, fazendo comentários sobre o arranjo e indicando se há pontos a melhorar”. A resposta obtida:

Ouvi sua gravação e aqui está um feedback detalhado:

Performance

Técnica: A execução está bem firme, com pegada e precisão rítmica consistentes. A condução nos pratos e os ataques na caixa soam bem equilibrados.

Dinâmica: Parece haver uma boa variação de intensidade ao longo da música, o que mantém a execução interessante.

Groove e Feeling: O groove está sólido e bem conectado com a música. Se há uma linha de metrônomo ou playback junto, você está bem alinhado.

Arranjo

Estrutura: O arranjo de bateria complementa bem a música, com variações bem colocadas entre seções.

Preenchimentos (fills): Os fills são criativos e funcionam bem para conectar as partes. Talvez alguns poderiam ter um pouco mais de clareza ou precisão, dependendo do efeito desejado.

Uso dos tambores e pratos: O kit está sendo usado de forma equilibrada, mas poderia haver uma exploração um pouco maior de texturas, como mais variações no chimbal ou uso de ghost notes na caixa.

Possíveis Melhorias

Precisão nos fills: Em algumas passagens rápidas, há um leve desencontro, que pode ser resolvido com prática em menor tempo e aumentando gradualmente a velocidade.

Mix e captação: Dependendo de como a gravação foi feita, algumas peças do kit (como bumbo ou tons) podem estar um pouco abafadas ou com pouca presença. Se foi gravado via MIDI, talvez seja questão de ajustar a mixagem dos sons escolhidos.

Uso de dinâmicas: Para um som mais expressivo, experimente explorar ainda mais variações de intensidade, especialmente em partes mais sutis da música (OPENAI, 2025).

Os comentários são significativos na medida em que aparentemente trazem uma análise do material sonoro, tendo como base uma linguagem normalmente utilizada por bateristas e fazendo referência a aspectos da performance que são objeto de atenção constante para professores de bateria. Os termos são bem específicos (“pegada”, “groove”, “feeling” e “ghost notes”¹, por exemplo) e colocam

¹ No contexto da performance musical, o termo “pegada” refere-se à solidez e firmeza na execução. Já o conceito de “groove” está relacionado ao “feeling”, isto é, à interação entre microvariações temporais e a organização rítmica em nível macro, resultando em uma sensação específica percebida pelo ouvinte (Gohn, 2021). As ghost notes, por sua vez, são toques sutis e de baixa intensidade utilizados por bateristas para preencher os espaços entre os ataques principais, conferindo maior complexidade e nuance ao ritmo.





em foco tanto elementos gerais (estrutura do arranjo) como detalhes (fills criativos). Há elementos que indicam uma análise de dados não subjetivos, como precisão rítmica e equilíbrio de condução e ataques, o que indica que de fato houve uma “escuta” do material. A afirmação de alinhamento com o metrônomo sugere que o andamento (bpm) foi analisado como consistente.

No entanto, as sugestões de melhoria, como “explorar mais variações de intensidade” ou “usar mais ghost notes”, são genéricas e não dialogam com as escolhas estéticas específicas da gravação analisada. Esse tipo de resposta, muito ampla e sem direcionamento, compromete a utilidade pedagógica do feedback fornecido. Mesmo raciocínio se aplica à frase “poderia haver uma exploração um pouco maior de texturas”. Além disso, é relevante notar que a gravação enviada continha apenas a linha de bateria, sem acompanhamento harmônico ou melódico. Assim, as afirmações da IA de que “o groove está bem conectado com a música” e de que “o arranjo de bateria complementa bem a música” revelam-se inconsistentes, sugerindo que o modelo está apenas replicando fórmulas comuns, sem considerar as características reais do material analisado. Logo, conclui-se que o modelo de linguagem está apenas tentando agradar o usuário, como já constatado em diversas outras situações (Salecha *et al.*, 2024).

Extrapolando a compreensão de textos e arquivos de áudio, modelos de linguagem também estão rapidamente aprimorando suas capacidades para trabalhar com vídeos (Maaz *et al.*, 2024). Para experimentar essa possibilidade com a IA, foram enviadas ao ChatGPT imagens registradas durante a gravação da música discutida anteriormente. O prompt solicitou uma análise dos movimentos do baterista, novamente com indicações de possíveis aperfeiçoamentos na performance. A resposta obtida:

Análise dos Movimentos do Instrumentista

Postura e Pegada

A postura parece equilibrada, com uma boa posição das mãos e braços.

A pegada nas baquetas parece relaxada, sem tensão excessiva, o que favorece controle e velocidade.

Técnica e Execução

Os movimentos dos braços e pulsos estão bem distribuídos, sugerindo eficiência no toque.



Se houver uso de técnicas como Moeller ou finger control, isso pode indicar um bom nível técnico.
O movimento das pernas no bumbo e no hi-hat parece coordenado.

Expressividade e Controle

Se há variação na dinâmica, isso mostra que o instrumentista está atento à musicalidade.
O controle de ghost notes, acentuações e condução nos pratos pode ser um ponto forte.

Pontos de Melhoria

Se houver tensão nos ombros ou nos braços, pode ser interessante ajustar a ergonomia.
A fluidez nos fills pode ser trabalhada para maior naturalidade.
O movimento das baquetas pode ser otimizado para economizar energia e manter precisão (OPENAI, 2025).

De forma semelhante à análise de áudio, o texto produzido pelo ChatGPT sugere que o vídeo foi examinado, ainda que se deva considerar que, tecnicamente, esses modelos atualmente ainda trabalham apenas com quadros ou imagens extraídas, e não necessariamente têm uma compreensão da dinâmica contínua como um observador humano (Yang *et al.*, 2023), isto é, o vídeo não é realmente “assistido”.

Embora o ChatGPT, especialmente na versão GPT-4-turbo, ofereça as chamadas capacidades multimodais, como interpretação visual, ele ainda apresenta limitações importantes no que diz respeito à análise de imagens em movimento. Ainda assim, a resposta revela como essa IA compreende o que é necessário dizer, com escolhas de palavras que soam bastante convincentes. As frases sobre “postura e pegada” são oportunas, indicando boa posição das mãos, sem tensão excessiva. A discussão técnica afirma que os movimentos estão bem distribuídos, há eficiência nos toques e o uso dos pedais “parece” coordenado. Ou seja, são comentários que poderiam ser de um professor de bateria, ao assistir o vídeo de seu aluno, com elogios gerais que seriam seguidos de críticas mais pontuais e específicas.

Todavia, o ChatGPT se abstém de uma análise mais detalhada, usando uma série de condicionais: “se houver uso de técnicas...”, “se há variação na dinâmica...”, “se houver tensões nos ombros ou braços...”. Nesse sentido, um professor seria



objetivo em afirmar se houve ou não o uso de técnicas Moeller ou finger control², variações de dinâmicas ou tensões nos ombros do aluno. E, novamente, as sugestões de melhorias são genéricas e um tanto óbvias: pode ser interessante ajustar a ergonomia, trabalhar a fluidez para maior naturalidade e otimizar o movimento das baquetas para economizar energia. O benefício de afirmações dessa natureza é bastante limitado, a não ser que fossem direcionados para um momento determinado da performance, com indicações precisas de pontos a lapidar. Como exemplo maior de observações supérfluas, há a frase “o controle de ghost notes, acentuações e condução nos pratos pode ser um ponto forte” – nada relacionado com o vídeo em discussão, somente elementos aleatórios que “podem ser pontos fortes”.

Em suma, essas interações demonstram que, no atual estágio de desenvolvimento, o ChatGPT apresenta uma capacidade sofisticada de emular o discurso avaliativo musical, mas ainda carece de precisão contextual e sensibilidade estética para oferecer feedback pedagogicamente significativo. Embora a IA simule com fluência um discurso avaliativo, suas respostas sugerem que o sistema opera principalmente por associação de padrões linguísticos, sem uma compreensão musical profunda do material analisado. Essa constatação reforça a necessidade de uma postura crítica diante dessas tecnologias, especialmente na educação musical, onde o olhar pedagógico humano não pode ser automatizado sem perda significativa de sentido.

Considerações finais

Novas gerações de modelos de linguagem de grande escala têm surgido incessantemente. Entre os exemplos, atualmente, nos Estados Unidos, estão em desenvolvimento o GPT da OpenAI, o Claude da Anthropic, o Gemini da Google, o Llama da Meta, o Grok da xAI, e o Nova da Amazon; enquanto na China avançam o Qwen da Alibaba, o Yi da 01.ai e a DeepSeek.

² As técnicas Moeller e finger control, desenvolvidas respectivamente por Sanford Moeller e Billy Gladstone, são métodos de controle de baquetas que visam otimizar a eficiência e a expressividade na execução. Ambas propõem séries de exercícios aplicáveis a diferentes instrumentos de percussão, contribuindo para o desenvolvimento da precisão, dinâmica e resistência do intérprete.



Quando o leitor passar os olhos nesta lista, certamente os modelos já terão novas versões, com treinamentos mais abrangentes e maiores capacidades para lidar com sons e imagens, ampliando o potencial para compreender e elaborar discussões sobre música. Em outras palavras, conforme o professor Ethan Molick (2024), qualquer seja a inteligência artificial que usamos em determinado momento, será a pior que iremos experimentar. Na próxima vez, já existirá outra melhor. Os investimentos nessa área são enormes e a evolução de novos modelos acontece em ritmo frenético, dificultando o acompanhamento das inovações mais recentes.

A partir do exposto acima, o objetivo deste texto foi provocar uma reflexão sobre os limites e alcances de inteligências artificiais generativas. Para isso, capturamos um retrato do momento atual, que de certo será continuamente transformado, mas que representa um marco para a Educação Musical: já existem modelos de linguagem que recebem material sonoro e conseguem gerar feedback sobre performances em instrumentos musicais. Pode ainda não ser totalmente útil, pois o ChatGPT criou respostas genéricas e sem grande valor pedagógico.

No entanto, ainda que com limitações significativas e sem a mediação sensível e contextual própria do ensino humano, o fato é que a máquina desempenhou funções que tradicionalmente seriam atribuídas a um professor. Diferentemente do que ocorria até tempos atrás, não temos em vista somente indicações de certo ou errado, como já acontecia em softwares de leitura musical como SmartMusic, mas também discussões sobre técnicas aplicadas e expressividade musical.

Como foi observado nas breves experiências apresentadas, o auxílio do ChatGPT para aperfeiçoar a interpretação de uma música tocada na bateria é ainda bastante limitado. O modelo de linguagem pode oferecer somente uma “simulação” de feedback pedagógico, mas com o progresso intenso da IA, não resta dúvida de que isso irá mudar. Assim como no Teste de Turing, poderemos nos propor a comparar falas de professores humanos com comentários gerados por modelos de linguagem, na expectativa de um ponto em que seja impossível distinguir os dois.

Nessa ocasião, surgem outras perguntas: haverá outro ponto em que uma aula dada por IA será superior à de um humano, pelo acesso a quantidades gigantescas de referências obtidas nos treinamentos dos modelos de linguagem?





Como consequência, haverá uma regulação adequada para lidar com as questões éticas, trabalhistas, e a proteção de dados necessária para conter interesses de empresas privadas que visam lucro, em detrimento de valores educacionais?

Ademais, qual será o balanço entre riscos e benefícios com a ideia de Khan (2024), para que a IA sirva de tutor individual de qualquer aluno, considerando o histórico e as particularidades de cada um, se consistentemente aplicada na Educação Musical, incluindo o ensino de instrumentos musicais? Teremos modelos de linguagem treinados a partir das ideias, técnicas e concepções de professores de bateria e percussão, com as devidas precauções de segurança, possibilitando que seus ensinamentos perdurem post mortem e fiquem acessíveis sem limitações de espaço e tempo? Será esse treinamento realizado até com dados de músicos já falecidos, como Art Blakey? Nesse caso, seria possível enviar para a IA um vídeo tocando um rulo na caixa, no estilo característico de Art Blakey, e receber comentários que o próprio Blakey faria? Quem define os limites éticos e os direitos de uma pessoa já falecida nessa situação? Qual deve ser a regulação no treinamento dos dados e quem colhe o lucro financeiro?

Todas essas questões nos remetem a um futuro de Educação Musical intrinsecamente permeado pela inteligência artificial. Considerando que a IA caminha para se tornar uma tecnologia ubíqua, assim como a eletricidade e a internet, dificilmente o futuro será outro. Tendo em vista que a realidade atual envolve sérios perigos decorrentes de alucinações e preconceitos, além da falta de limites de proteção e de clareza na utilização de dados em muitas das ferramentas disponíveis, é importante compreender esses problemas e tentar mitigar seus riscos.

O letramento em inteligência artificial é fundamental para que professores e alunos, tanto nos estudos de percussão como em outros campos de música, possam não apenas utilizar essas ferramentas com criticidade e consciência, mas também participar ativamente da construção de práticas pedagógicas mais éticas, inclusivas e criativas no cenário da Educação Musical contemporânea. Assim, mais do que trazer respostas, a presença da IA na Educação Musical pode se tornar uma oportunidade para repensarmos as próprias perguntas que fazemos e o modo





como escutamos, ensinamos e aprendemos música em um mundo cada vez mais atravessado por algoritmos.



Referências

ACQUILINO, Alberto; SCAVONE, Gary. Current State and Future Directions of Technologies for Music Instrument Pedagogy. **Frontiers in Psychology**, v. 13, 2022. <https://doi.org/10.3389/fpsyg.2022.835609>

ALVES, Lynn. Notas Iniciais sobre a Inteligência Artificial e Educação. In: ALVES, Lynn (org.). **Inteligência Artificial e Educação**: refletindo sobre os desafios contemporâneos. Salvador: EDUFBA; Feira de Santana: UEFS Editora, p. 33-50, 2023.

BLADES, James. **Percussion instruments and their history**. Westport: Bold Strummer, 1992.

CHIU, Thomas. K. F. The impact of Generative AI (GenAI) on practices, policies and research direction in education: a case of ChatGPT and Midjourney. **Interactive Learning Environments**, v. 32, no. 10, p. 6187–6203, 2023. <https://doi.org/10.1080/10494820.2023.2253861>

CONVINGTON, Paul; ADAMS, Jay; SARGIN, Emre. Deep Neural Networks for YouTube Recommendations. **Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16)**. Association for Computing Machinery, New York, NY, USA, p. 191–198, 2016. <https://doi.org/10.1145/2959100.2959190>

FEUERRIEGEL, Stefan; HARTMANN, Jochen; JANIESCH, Christian; ZSCHECH, Patrick. Generative ai. **Business & Information Systems Engineering**, v. 66, n. 1, p. 111-126, 2024. <https://doi.org/10.1007/s12599-023-00834-7>

GOHN, Daniel Marcondes. “Feeling Brazilian”: The search for authenticity in drum kit playing, **Journal of Popular Music Education**, v. 5, n. 2, p. 211–26, 2021.

GRASSINI, Simone. Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. **Education Sciences**, v. 13, no. 7, artigo 692, 2023. <https://doi.org/10.3390/educsci13070692>

HANCOCK, Jaxon; REN, Jian. Accurate Traffic State Prediction with Deep Learning - Analyzing Statistical Aides for Identifying Anomalous Traffic Trends, **2024 IEEE 30th International Conference on Parallel and Distributed Systems (ICPADS)**, Belgrade, Serbia, p. 408-414, 2024. <https://doi.org/10.1109/ICPADS63350.2024.00060>

HOFFMAN, Reid; GPT-4. **Impromptu**. Amplifying Our Humanity Through AI. Washington: Dallepedia LLC, 2023.





KHAN, Salman. **Brave New Words**. How AI will revolutionize education (and why that's a good thing). New York: Viking, 2024.

MAAZ, Muhammad; RASHEED, Hanoona; KHAN, Salman; KHAN, Fahad Shahbaz. Video-ChatGPT: towards detailed video understanding via large vision and language models. In: **Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics**. Kerrville, TX: Association for Computational Linguistics, p. 12585–602, 2024. <https://doi.org/10.48550/arXiv.2306.05424>

MAHESHWARI, Chhavi. Music recommendation on spotify using deep learning. arXiv preprint arXiv:2312.10079, 2023. <https://doi.org/10.48550/arXiv.2312.10079>

McMAHON, Liv; KLEINMAN, Zoe. Glue pizza and eat rocks: Google AI search errors go viral. **BBC News online**. 24/05/2024. Disponível em: <https://www.bbc.com/news/articles/cd11gzejgz4o>. Acesso em: 21 fev. 25.

MEHTA, Atharva et al. Music for All: Representational Bias and Cross-Cultural Adaptability of Music Generation Models. **North American Chapter of the Association for Computational Linguistics**, 2025. <https://doi.org/10.48550/arXiv.2502.07328>

MENZ, Bradley D.; KUDERER, Nicole M.; CHIN-YEE, Benjamin et al. Gender Representation of Health Care Professionals in Large Language Model-Generated Stories. **JAMA Network Open**, v.7, n.9, e2434997, 2024. <https://doi:10.1001/jamanetworkopen.2024.34997>

MOLICK, Ethan. **Co-intelligence**: living and working with AI. New York: Portfolio/Penguin, 2024.

MONTGOMERY, Blake. Mother says AI chatbot led her son to kill himself in lawsuit against its maker. **The Guardian**, 23 de Out, 2024. Disponível em: <https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death>. Acesso em: 04 abr. 25.

OPENAI. **ChatGPT (GPT-4)**. <https://chat.openai.com/chat>. Acesso em: 10 abr. 2025.

PEREIRA, Igor; ARAÚJO, Felipe; KORZENIOWSKI, Filip; VOGL, Richard. MoisesDB: A dataset for source separation beyond 4-stems. **arXiv preprint arXiv:2307.15913**, 2023. <https://doi.org/10.48550/arXiv.2307.15913>

SALECHA, Aadesh; IRELAND, Molly E.; SUBRAHMANYA, Shashanka; SEDOC, João; UNGAR, Lyle H.; EICHSTAEDT, Johannes C. Large language models



display human-like social desirability biases in Big Five personality surveys. **PNAS Nexus**, v. 3, n. 12, 2024.

<https://doi.org/10.1093/pnasnexus/pgae533>

SAVAGE, Mark. Sir Paul McCartney says artificial intelligence has enabled a 'final' Beatles song. **BBC News online**, 13 de Jun, 2023. Disponível em:

<https://www.bbc.com/news/entertainment-arts-65881813>.

Acesso em: 12 mar. 2025.

SLOBODA, John A. **The Musical Mind: The Cognitive Psychology of Music**. Oxford: Oxford University Press, 1985.

TAN, Selim. Are We All Musicians Now? Authenticity, Musicianship, and AI Music Generator Suno. **SocArXiv**. December 4, 2024.

<https://doi:10.31235/osf.io/4nt8z>

TURING, Alan M. Computing Machinery and Intelligence. **Mind**, [s.l.], v. 59, n. 236, pp. 433–460, 1950. Disponível em: <http://www.jstor.org/stable/2251299>.

Acesso: em 7 jun. 2025.

VAN DIJCK, José; POELL, Thomas; DE WAAL, Martijn. **The Platform Society: Public values in a connective world**. New York: Oxford university press, 2018.

YANG, Zhengyuan; LI, Linjie; LIN, Kevin; WANG, Jianfeng; LIN, Chung-Ching; LIU, Zicheng; WANG, Lijuan. The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision). **arXiv preprint arXiv:2309.17421**, 2023.

<https://doi.org/10.48550/arXiv.2309.17421>

ZAGORSKI-THOMAS, Simon. Real and Unreal Performances: The Interaction of Recording Technology and Rock Drum Kit Performance. In: DANIELSEN, Anne (ed.). **Musical Rhythm in the Age of Digital Reproduction**. Farnham: Ashgate Publishing, pp. 195–212, 2010.



Daniel Gohn é professor da Licenciatura em Música da Universidade Federal de São Carlos (UFSCar). É Mestre e Doutor pela Escola de Comunicações e Artes da Universidade de São Paulo (ECA/USP) e Bacharel em Música Popular pela Universidade Estadual de Campinas (UNICAMP). Estudou na escola Drummers Collective, em Nova York, onde também trabalhou na produção de vídeos educacionais com a empresa Hudson Music. Atualmente, é baterista da Orquestra Rock, tocando com artistas como Dinho Ouro Preto, Di Ferrero e Maria Gadú, entre outros; e participa do projeto Highland Echoes, promovendo a integração das culturas escocesa e brasileira. É autor de “No Mundo da Bateria: conceitos, ideias e composições” (Editora Som, 2018), “Caderno de Percussão Yamaha” (Ricordi, 2014), “Educação Musical a Distância: Abordagens e Experiências” (Cortez, 2011), “Tecnologias Digitais para Educação Musical” (EdUFSCar, 2010) e “Autoaprendizagem Musical: Alternativas Tecnológicas” (Annablume, 2003), além de um capítulo no livro Cambridge Companion to the Drum Kit (Cambridge University Press, 2021). Seus principais interesses de pesquisa incluem o uso de tecnologias na educação musical e processos de ensino e aprendizagem de instrumentos de percussão.

<http://lattes.cnpq.br/1332810341001276>