

Projet Étudiant Master Data Science

Génération de tablatures de guitare basse pour l'accompagnement de guitare en musique populaire occidentale

Alexandre D'Hooge – Doctorant en 3^è année, équipe Algomus

`alexandre.dhooge@univ-lille.fr`

6 septembre 2024

En musique populaire occidentale (e.g. rock, pop, blues, funk...), la **guitare basse** joue généralement un rôle important dans la **perception de la progression harmonique d'un morceau**. Conjointement avec la batterie, la basse participe également grandement à l'aspect rythmique. Lors de la composition de morceaux, les musiciens n'ont pas toujours une maîtrise de la guitare basse et ne sont donc pas en mesure de concevoir une partie de basse accompagnant au mieux des parties de guitares préalablement écrites. Un outil permettant de leur **proposer automatiquement une tablature de basse pertinente** pourrait donc les aider dans cette tâche.

Objectifs

Nous proposons dans ce projet de nous intéresser à la génération de tablatures de basse conditionnée par une tablature de guitare, composée préalablement. Un tel outil aurait le potentiel d'assister les guitaristes non-bassistes dans la composition de musique populaire. Pour cela, nous proposons d'appliquer les méthodes du **Traitement Automatique du Langage Naturel** via des modèles comme les *transformers*, qui ont prouvé leur efficacité au traitement de la musique écrite [1-3]. En particulier, nous suggérons l'exploration de tâches d'entraînement spécialisées comme le Masked Language Modeling (MLM) sur des modèles type BERT [4] ou BART [5]. L'étudiant-e devra implémenter une tokenization des chansons de la **base de données de tablatures DadaGP** [6] (dont le code de tokenization original est déjà disponible publiquement). Ces données seront ensuite utilisées afin d'entraîner un réseau de neurones avec modules d'attention sur la prédiction d'une partie de basse. Suite à ça, une phase plus exploratoire pourra être entamée, en envisageant **plusieurs méthodes de contrôle et conditionnement** telles qu'elles pourraient être désirées par une personne faisant usage de cet outil en composant.

Encadrement

Le projet sera réalisé au sein de l'équipe Algomus¹ du laboratoire CRISAL. L'encadrement sera principalement assuré par Alexandre D'Hooge, doctorant en troisième année. La supervision sera également assurée par Ken Déguernel, Chargé de Recherche CNRS.

1. www.algomus.fr

Références

- [1] Drew Edwards et al. "MIDI-to-Tab : Guitar Tablature Inference via Masked Language Modeling". In : *Proc. of the 25th Int. Society for Music Information Retrieval Conf. (ISMIR)*. 9 août 2024. arXiv : 2408.05024 [cs]. url : <http://arxiv.org/abs/2408.05024>.
- [2] Ziyu Wang et Gus Xia. "Musebert : Pre-Training of Music Representation for Music Understanding and Controllable Generation". In : *Proc. of the 22nd Int. Society for Music Information Retrieval Conf. (ISMIR)*. 2021.
- [3] Mingliang Zeng et al. "MusicBERT : Symbolic Music Understanding with Large-Scale Pre-Training". In : *Findings of the Association for Computational Linguistics : ACL-IJCNLP 2021*. Findings of the Association for Computational Linguistics : ACL-IJCNLP 2021. Online, 2021. doi : 10.18653/v1/2021.findings-acl.70.
- [4] Jacob Devlin et al. "BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding". In : *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 24 mai 2019. arXiv : 1810.04805 [cs]. url : <http://arxiv.org/abs/1810.04805>.
- [5] Mike Lewis et al. "BART : Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension". In : *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Juill. 2020. doi : 10.48550/arXiv.1910.13461. arXiv : 1910.13461 [cs, stat].
- [6] Pedro Sarmiento et al. "DadaGP : A Dataset of Tokenized GuitarPro Songs for Sequence Models". In : *Proc. of the 22nd Int. Society for Music Information Retrieval Conf. (ISMIR)*. Online, 2021.