

Kyle Zarzana & Liam Leece

CS 4395.001

Dr. Karen Mazidi

11/13/2022

### Symbolic Music Understanding

The paper we chose titled, “MusicBERT: Symbolic Music Understanding with Large-Scale Pre-Training”, covers the development of MusicBERT, a pre-trained model for symbolic music understanding. The model uses a corpus of more than one million songs. Applying pre-training techniques that work on text-based data is insufficient as the structure of music data is more complex. Each song can have a different tempo, use different instruments, and have different time signatures. Previous models have been able to extract large amounts of data from songs, but the amount of data that makes up an entire song becomes too large for a pre-trained model to work on. Thus, they were unable to perform tasks that require a song-level amount of information. The model was evaluated based on its success during four tasks. Those four tasks are melody completion, accompaniment suggestion, genre classification, and style classification.

The four tasks that MusicBERT was tested on truly show the success of the authors’ research. The melody completion and accompaniment suggestion tasks focus mainly on a particular section of a song while the genre and style classification tasks make decisions about the whole. The results of these tests show that MusicBERT outperformed all other models designed for these problems, often doing so on a small pre-trained data set, which evidences how better results will follow with a larger more complete data set. They tested MusicBERT on each task against other models designed for the same things, using the same data set. In each case,

MusicBERT outperformed all other models tested. While these results are truly amazing, we also want to take note of the accomplishment that is making an effective and universal music encoding system. Their creation and use of OctupleMIDI is what allowed them to make such a large corpus of music to train MusicBERT but still have the data be complete and readable. Most other encodings lose bounds of information about a song when it is encoded but the integrity of the pieces were vital to the pre-training of their model. We believe that this encoding is their most valuable unique contribution to the field of nlp. The creation of MusicBERT is certainly amazing, but the OctupleMIDI is what paved the way and will allow others interested in this domain to take things even further.

This research had many different contributors that aided in the research and publishing of the paper. For one, Mingliang Zeng, it is their first published research paper and has been cited a total of 21 times. On the opposite end Zeng's mentor, Tie-Yan Liu, has over 40,000 citations on google scholar. For the authors in between and their citations we have; Zeqian Ju with 117 citations, Rui Wang with 705, Xu Tan with 4,883, and lastly Tao Qin with 16,335. Apart from Zeng, the other authors are all more established in their research. These impressive minds have worked on many other nlp topics in the past such as making text-to-speech faster and higher quality, achieving human-level accuracy and speed in Chinese-to-English speech translation and creating multilingual neural machine translators.

We first took interest in this research paper because of our mutual love for music. We believe this work is important because music identification has proven to be a fast-growing topic of interest and the MusicBERT has shown that there is so much more that models like these can achieve. It is our hope that this research will inspire people to learn about nlp in music and help pave the way for others in the field to take things even further.

## Citations

Zeng, Mingliang, et al. "Musicbert: Symbolic music understanding with large-scale pre-training." *arXiv preprint arXiv:2106.05630* (2021).