Tutorial 4

Music Classification: Beyond Supervised Learning, Towards Real-world Applications

Keunwoo Choi, Minz Won and Janne Spijkervet

Abstract

Music classification is a music information retrieval (MIR) task to classify music items to labels such as genre, mood, and instruments. It is also closely related to other concepts such as music similarity and musical preference. In this tutorial, we put our focus on two directions - the recent training schemes beyond supervised learning and the successful application of music classification models.

The target audience for this session is researchers and practitioners who are interested in state-of-the-art music classification research and building real-world applications. We assume the audience is familiar with the basic machine learning concepts. For those who are not, we kindly refer to [1, 2] to be prepared for this session.

We plan to present three lectures as follows:
1. Music classification overview: Task definition, applications, existing approaches, datasets
2. Beyond supervised learning: Semi- and self-supervised learning for music classification
3. Towards real-world applications: Less-discussed, yet important research issues in practice

We provide an accompanying code repository and Jupyter notebooks that can be used along with the video presentation. With the material, attendees can easily train semi- and self-supervised models with their own audio data.

[1] A Tutorial on Deep Learning for Music Information Retrieval, Keunwoo Choi et al., 2017 (Concepts in deep learning) https://arxiv.org/abs/1709.04396

[2] An Introduction to Statistical Learning, Daniela Witten et al., 2013 (Chapter 2-4 for ML fundamentals) https://www.statlearning.com/

Biographies of Presenters

Keunwoo Choi (website) is a research scientist at ByteDance, developing machine learning products for music recommendation and discovery. He received a Ph.D degree from Queen Mary University of London (c4dm) in 2018. As a researcher, he also has been working at Spotify (2018 - 2020) and several other music companies as well as open-source projects such as Kapre, librosa, and torchaudio. He argues that he writes some good music.

Minz Won (website) is a Ph.D candidate at the Music Technology Group (MTG) of Universitat Pompeu Fabra in Barcelona, Spain. His research focus is music representation learning. Along with his academic career, he has put his knowledge into practice with industry internships at Kakao Corp., Naver Corp., Pandora, Adobe, and he recently joined ByteDance as a research scientist. He contributed to the winning entry in the WWW 2018 Challenge: Learning to Recognize Musical Genre.

Janne Spijkervet (website) graduated from the University of Amsterdam in 2021 with her Master's thesis titled "Contrastive Learning of Musical Representations". The paper with the same title was published in 2020 on self-supervised learning on raw audio in music tagging. She has started at ByteDance as a research scientist (2020 - present), developing generative models for music creation. She is also a songwriter and music producer, and explores the design and use of machine learning technology in her music.