

# Tutorial 5

## Programming MIR Baselines from Scratch: Three Case Studies

Rachel Bittner, Mark Cartwright and Ethan Manilow

### Abstract

This tutorial will walk through the creation of MIR baselines programmed live, including pitch tracking, instrument identification, and drum transcription. Each case study will start with building a system and finish with evaluations and visualization/sonification, each using different tools and styles of programming. This tutorial is both beginner and experienced programmer-friendly and will start from the basics but will move quickly. While the tutorial is not interactive, all code will be made available afterwards.

### Biographies of Presenters

**Rachel Bittner** is a Senior Research Scientist at Spotify in Paris. She received her Ph.D. in Music Technology in 2018 from the Music and Audio Research Lab at New York University under Dr. Juan P. Bello, with a research focus on deep learning and machine learning applied to fundamental frequency estimation. She has a Master's degree in mathematics from New York University's Courant Institute, as well as two Bachelor's degrees in Music Performance and in Mathematics from the University of California, Irvine. In 2014-15, she was a research fellow at Telecom ParisTech in France after being awarded the Chateaubriand Research Fellowship. From 2011-13, she was a member of the Human Factors division of NASA Ames Research Center, working with Dr. Durand Begault. Her research interests are at the intersection of audio signal processing and machine learning, applied to musical audio. She is an active contributor to the open-source community, including being the primary developer of the pysox and mirdata Python libraries.

**Mark Cartwright** is an Assistant Professor at New Jersey Institute of Technology in the Department of Informatics. He completed his PhD in computer science at Northwestern University as a member of the Interactive Audio Lab, and he holds a Master of Arts from Stanford University (CCRMA) and a Bachelor of Music from Northwestern University. Before his current position, he spent four years as a researcher in the Music and Audio Research Lab (MARL) and the Center for Urban Science and Progress (CUSP) at New York University (NYU). His research lies at the intersection of human-computer interaction, machine learning, and audio signal processing. Specifically, he researches human-centered machine listening and audio processing tools for creative expression with sound and understanding the acoustic world.

**Ethan Manilow** is a PhD candidate in Computer Science at Northwestern University under advisor Prof. Bryan Pardo. His research lies in the intersection of signal processing and machine learning, with a focus on source separation, automatic music transcription, and open source datasets and applications. Previously he was an intern at Mitsubishi Electric Research Labs (MERL) and at Google Magenta. He is one of the lead developers of nussl, an open source audio separation library. He lives in Chicago, where he spends his free time playing his guitar and smiling at dogs he passes on the sidewalk.