IS/CS 698 Machine Listening

Fall 2023

Instructor: Mark Cartwright (mark.cartwright@njit.edu)
Teaching Assistant: Sripathi Sridhar (ss645@njit.edu)
Format: In-person
In-person Location: Mechanical Engineering Bldg 221
Class Hours: Friday 1:00pm - 3:50pm
Office Hours: Tuesday 1–3pm in GITC 3902 or by appointment

Course Description

Machine listening is the study of algorithms for the machine understanding of audio. It's the auditory sibling of computer vision. It has diverse applications such as automatic music transcription, music recommendation, audio search, smart home acoustic sensing, machine condition monitoring, audio captioning, urban noise monitoring, and wildlife monitoring. This course will provide an overview of topics in the field, focusing on the machine understanding of music and environmental sound. Lectures will cover the fundamental problems in the field and both classical and modern approaches to solving these problems. Students will complete Python-based assignments, present machine-listening research papers to the class, and complete a semester-long project of their choosing.

Topics

- Musical sound understanding (e.g. algorithms for estimating pitch, chords, beat, instruments, rhythm, structure)
- Environmental sound understanding (e.g. algorithms for detecting and localizing sound events)

Prerequisites

- Basic Python programming knowledge (Required since language instruction will not be covered in class)
- Basic Machine learning knowledge (Recommended, but not required. Basics will be covered)

Course Goals

The goal of this course is to provide students with:

1. Basic understanding of background theory for machine listening, including digital signals, human auditory perception, and machine learning
1. Broad knowledge of common topics in machine listening, including their definitions, challenges, and evaluation metrics.
2. Understanding of the current state-of-the-art and remaining challenges for each topic
3. The ability to critically analyze the behavior of machine listening systems
4. Practical skills for working with open-source machine listening resources in Python

Textbook and readings

1. Müller, Meinard. “Fundamentals of Music Processing” (Available online through the NJIT Library)
2. Additional readings posted on Canvas

Grading policy

The class will be graded on a 100-point scale.

A	90 and up
B+	85 to 89
B	80 to 84
C+	75 to 79
C	70 to 74
D	65 to 69
F	Less than 65

Point distribution:

- 5 pts. class participation
- 25 pts. homework
- 10 pts. Midterm 1
- 10 pts. Midterm 2
- 20 pts. paper discussion roles
- 30 pts. final project

Homework and reading assignments are solo assignments and must be original work.

Final projects may be group assignments and all members of a group will share a grade for all parts of the assignment.

Participation Policy

You are expected to attend class and actively participate in both class and online discussion, e.g. responding to online / in-person polls, asking questions, volunteering answers, sharing tips/resources, etc.

Homework

We will use Canvas for distributing and submitting any written assignments, and Github Classroom for distributing and submitting any programming-based assignments. No late assignments will be accepted, but you can drop your lowest homework grade.

Software

All coding assignments will be in Python and will utilize the SciPy Stack and other audio and machine-listening-specific packages, e.g. `librosa`, `mir_data`, `mir_eval`, `sed_eval`, etc. In the first homework assignment (HW0), you will be tasked with setting up your programming environment and familiarizing yourself with the tools we will be using throughout the course.

Paper Presentations

We will read research papers throughout the class to understand the current state of the field. The whole class will participate in the discussion of these papers. Students will be assigned roles such as “scientific peer reviewer” and “archaeologist” (see https://colinraffel.com/blog/role-playing-seminar.html list of roles). Multiple students may be assigned to the same role. Each role group will present a short presentation to the class.

Final Project

Students will work on a final project. The final project can be done in groups of 1-3 and will have code, written, and presentation components. Project proposals will be due roughly a month into the semester (exact date TBD). I will meet with each project group 4 times. Once to give feedback on an initial pitch. A second time to give feedback on a concrete project proposal and plan. And two more times to give interim feedback during the project.

Class environment

Please do your part by seeking to promote the success of others, and by treating each other in ways that respect and celebrate the diversity of talent that is drawn to this field. The classroom is an open forum for discussion, and I encourage all students to feel free to ask questions in class. Please do not be afraid to ask any question, no matter how basic it may seem. What is basic to some of the class may be completely new to the rest.

Accessibility

If you are in need of accommodations due to a disability please contact Chantonette Lyles, Associate Director of the Office of Accessibility Resources & Services (OARS), Fenster Hall Room 260 to discuss your specific needs. A Letter of Accommodation Eligibility from the OARS authorizing your accommodations will be required.

Academic Integrity

Academic integrity is the cornerstone of higher education and is central to the ideals of this course and the university. Cheating is strictly prohibited and devalues the degree that you are working on. As a member of the NJIT community, it is your responsibility to protect your educational investment by knowing and following the academic code of integrity policy that is found at: http://www5.njit.edu/ policies/sites/policies/files/academic-integrity-code.pdf.

Please note that it is my professional obligation and responsibility to report any academic misconduct to the Dean of Students Oce. Any student found in violation of the code by cheating, plagiarizing or using any online software inappropriately will result in disciplinary action. This may include a failing grade of F, and/or suspension or dismissal from the university. If you have any questions about the code of Academic Integrity, please contact the Dean of Students Oce at dos@njit.edu.

AI Policy

The use of generative AI tools (e.g., ChatGPT, CoPilot, etc.) is permitted in this course and is recommended to speed up the development of plotting and data-wrangling code. However, it is recommended to only use such tools as an assistant when coding, since you will be expected to fully understand any code you submit and will be assessed on exams accordingly with that expectation.

Feedback

I will solicit (anonymous) feedback from students throughout the course through anonymous surveys in Canvas, but if you have pressing or specific issues, please do not hesitate to let me know if any aspect of our course or class community can be improved.

Calendar

WEEK	TOPIC
1	Introduction / Music, Acoustics, Signals, and Perception
2	Digital signal theory overview
3	Machine learning overview
4	Classic Music Transcription (Chords, Melody, Beat/Tempo)
5	Modern Music Transcription / Midterm 1
6	Classic Timbre Recognition
7	Modern Timbre Recognition
8	Classic Source Separation
9	Modern Source Separation
10	Classic Sound Event Detection
11	Modern Sound Event Detection
12	Additional Topics 1 / Midterm 2
13	Additional Topics 2
14	Additional Topics 3