Fall 2023

IS 485 - 001: Machine Listening

Mark Cartwright

IS/CS 485 Machine Listening

Fall 2023

Instructor: Mark Cartwright (mark.cartwright@njit.edu)
Teaching Assistant: Sripathi Sridhar (ss645@njit.edu)
Format: In-person
In-person Location: Faculty Memorial Hall 321
Class Hours: Tuesday and Friday 10am - 11:20am
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 and classical approaches of the field, as well as the discussion of the field's recent advances. Students will complete weekly Python-based assignments and either present machine-listening research papers to the class or complete a semester-long project of their choosing.

Topics

- Musical sound understanding (e.g. algorithms for estimating pitch, chords, beat, instruments, rhythm, and 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
2. Broad knowledge of common topics in machine listening, including their definitions, challenges, and evaluation metrics
3. Awareness of the current state-of-the-art and remaining challenges for each topic
4. The ability to critically analyze the behavior of machine listening systems
5. 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) ISBN: 978-3-319-21945-5
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:

- 10 pts. class participation
- 35 pts. homework
- 15 pts. Midterm 1
- 15 pts. Midterm 2
- 25 pts. paper presentation or course project

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

Students may choose to either complete a final project or read and present research papers of interest to the class. Final projects and paper presentations will be group assignments. Grades for these components will be a combination of the instructor's assessment as well as peer evaluation of each group member's contribution.

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 (Option 1)

In the final weeks of the semester, we will read research papers of interest to the class. 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 will be assigned to the same role. Each role group will present a short presentation to the class. Students will be expected to be able to answer in-class questions about their presentation topics.

Final Project (Option 2)

Instead of end-of-semester paper presentations, students may opt to work on a final project. The final project can be done in groups of 2-3 students and will have code, written, and presentation components. Project proposals will be due a month into the semester. 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	DATE	TOPIC
1	Sep-5	Intro
1	Sep-8	Music, Acoustics, Signals, and Perception
2	Sep-12	Music, Acoustics, Signals, and Perception
2	Sep-15	Digital signal theory overview
3	Sep-19	Digital signal theory overview *(Prof. Cartwright at Conference)*
3	Sep-22	Digital signal theory overview *(Prof. Cartwright at Conference)*
4	Sep-26	Machine learning overview
4	Sep-29	Machine learning overview
5	Oct-3	Machine learning overview
5	Oct-6	Midterm 1
6	Oct-10	Chords
6	Oct-13	Chords
7	Oct-17	Melody
7	Oct-20	Melody
8	Oct-24	Beat and tempo (*Prof. Cartwright at Conference*)
8	Oct-27	Beat and tempo
9	Oct-31	Timbre
9	Nov-3	Timbre / Representation learning
10	Nov-7	Identification (Fingerprinting, Structure, Cover Song ID)
10	Nov-10	Source separation
11	Nov-14	Source separation

11	Nov-17	Environmental Sound Scene Description
12	Nov-22	Midterm 2
12	Nov-24	THANKSGIVING BREAK
13	Nov-28	Paper discussion
13	Dec-2	Paper discussion
14	Dec-5	Paper discussion
14	Dec-8	Paper discussion
15	Dec-12	Paper discussion
	Exam time	Project presentation