

MLSP Class Project Guidelines

Cem Subakan

September 2025

1 Defining the expectations

The goal of this document is to specify the expectations and guidelines on your MLSP class projects. In most general terms, we expect the students to carry out an applied machine learning project. We would like the projects,

- *To be realistic:* You should be able to implement the project within a reasonable amount of time, with limited computational resources (unless you have access to a larger compute cluster). One resource you can use is google colab, where you can access to a free GPU. You can also get the paid version for an affordable amount which can make running experiments easier. We will also give you access to the Valeria cluster.
- *To have a well defined dataset:* Especially if you are going to train a supervised model, you should have a well-defined dataset with labels. For deep learning models typically a large dataset is required. However, bear in mind that it's currently very popular to use large pretrained models (e.g. foundation models such as Whisper, CLIP, GPT 2/3/4, ...). Even if you will use pretrained models (which I recommend you to do. You can for instance browse on <https://huggingface.co> to see if there can be pretrained models you can take advantage of.
- *To clearly define what will be the input, and what will be the output:* In your project proposals and your final presentations, it would really help if you could clearly define what you will give to the model, and what will be the prediction of the model. It would also help a lot if you can specify what dataset you will use to train the model to obtain the output you specify. For example, if you are planning to solve a music source separation task, you need to specify if you will have access to the isolated instrument tracks to be able to train the model. Or, if you are planning to automatic speech recognition on a low-resource language, you need to say if you will have access to training data that corresponds to your inputs and outputs. This is of course related to the previous item, but we are trying to think about the dataset in the context of input and output of the model.
- *To be not trivial, or even be somewhat ambitious or interesting:* It is of course easy to come up with a project proposal that would fit the criteria we have listed above, if you pick a project that is very simple. For instance if you say, you want to classify handwritten digits, then that is something that can be done within our lab sessions. So I expect a bit more creativity / challenge. The creativity can come from different sources. It can be a modeling idea you want to explore on a known problem, or it could be that you want to attack an interesting applied problem. For teams that is composed of only undergrads, it is acceptable to the team to implement/interpret a recent paper. Otherwise, I expect you to teach me something I don't know. :) – If you reach this expectation we can potentially publish together also.
- *To be within the machine learning and/or signal processing domain:* As I have tried to explain, the area of signals is extremely wide. As you also know people are applying machine learning on almost any field these days. So, basically almost any machine learning project would cut it. However, you

definitely need to include some machine learning or signal processing techniques in your projects. We can not for instance accept a project where you only collect data, and present ideas on how to solve the problem. We need to see some concrete results where a model obtains predictions. If you want to go for a project that really lies within the core of the MLSP field, you can take a look into ICASSP 2023 proceedings <https://ieeexplore.ieee.org/xpl/conhome/10094559/proceeding>.

- *To be well situated within the existing literature:* I expect the project reports to contain a literature review on existing works in your subject. If there are no papers that overlap with the problem you are attacking, still try to include some similar papers, or papers that could be useful for the methodology.

2 Format for the project proposals

I would expect you to follow this section structure in your proposals (I want you to have the following sections in your proposal):

- Abstract – 300 words max. You should briefly describe what you will do in the project, and what expected experimental results you foresee obtaining.
- Introduction – Here I expect you to verbally define your project more in detail, and introduce the solution technique you would want to employ.
- Literature review – Here I expect you to compare and contrast your solution with the existing techniques in the literature.
- Input-Output-Dataset – Here I expect you to clearly define (with a diagram) what will be your input, what will your model predict. I also expect you to specify what data you will use in training and testing. Do you have sufficient labeled/supervision data for training? Are you going to use pretrained models?
- Performance Metrics – What metrics will you use to measure how well you do, and potentially to compare with others?
- Computational Resources – What computational resources do you have, and how much do you think you will need for your project?
- Checklist – Please include a checklist that comments on how you plan to satisfy the expectations I listed in the first section. Include a bullet list, and comment with 1-3 sentences on how your project proposal satisfies the criteria.

3 Example Project Ideas:

Do not take these too literally, I am just improvising here.

- Controllable Symbolic Music Generation
- Fine Tuning a Large Language Model for a Specific Task
- Fine Tuning a Pretrained Generative Model on your Specific Dataset that generates data that respects certain constraints
- Text-Guided Speech / Audio Source Separation for Speaker/Source Extraction
- Zero-Shot Classification using CLAP/CLIP, Understanding the limits of generalization
- Interpretable Classification of Music Genres
- Music Transcription for New Instruments (Train on a set instruments that you haven't seen before)
- Fine tuning pretrained image / audio captioning models.

4 Places you can look for ideas:

- UCI repository: <http://archive.ics.uci.edu/>
- Huggingface
- Kaggle: <https://www.kaggle.com/>
- ICASSP/IEEE MLSP/IEEE WASPAA/ Interspeech proceedings (your UL account would give you access)
- NeurIPS / ICML/ ICLR proceedings.
- Try to read papers..