

ML Project Planning Checklist

For proof-of-concept and academic research

1. Project Motivation

- What is the problem you want to solve?
- Which strategic goal is it linked to?

2. Problem Definition

- What specific **output** do you want to predict?
- What **input** data do you have for the algorithm?

For each dataset describe:

- Number of examples (roughly)
- Hours of audio (roughly)
- Location (where is it saved & how can it be accessed?)
- How is the dataset generated & how frequently is it updated?
- Are there any big changes / **biases** in the data?
- What are the most **relevant factors** for predicting your specific output?
- How many **training examples** can you provide?

3. Performance Measurement

- Do you have a **simple benchmark** to compare against?
If yes:
 - How does it perform?
 - Do you have documentation on how it works?
- How will you measure the **performance** of the predictions?

4. Expected Results

- What is the minimum level of accuracy you expect?
- What would a perfect solution look like?
- Are there reference solutions (e.g. research papers)?

5. Milestones

- Are there any deadlines to be aware of?
- When do you need to see the first results?
- When do you want to have a finished solution?

6. Contacts

- Who is responsible for the project (PM)?
- Who can grant access to the datasets?
- Who can help understand the current process and / or the simple benchmark (domain expert)?

7. Collaboration

- Set up a weekly update with your team!
- Who should be involved?
- What should they learn?
- Define where code & issues are located and how to access them.
- How much compute power do you need and where can you get it?