

Machine Learning

Street fighting ML

Emanuele Rodolà
rodola@di.uniroma1.it



SAPIENZA
UNIVERSITÀ DI ROMA

2nd semester a.y. 2023/2024 · May 21, 2024

STREET FIGHTING ML



1.

Pitch your idea.

If no one's around, pitch it to ChatGPT.

2.

Have an interactive notebook ready.
With pre-installed libraries.

3.

Quickly prototype with baselines.
(decision trees, logistic regression)

4.

Create small demos of your own methods.
Use them for quick sharing and re-use.

5.

Ensure reproducibility.

6.

Think twice before training.

7.

Think twice before training from scratch.

8.

Look at the data.

9.

Look for structure.
Look for clusters.

10.

Do dimensionality reduction.

 Use K-Means.

 Use PCA.

11.

Can you do it with nearest neighbors?

12.
Try overfitting first.

13.

Don't aim for one big, elegant model.

14.

Favor simplicity.

15.

Familiarize with the libraries.

16.

Don't get rusty.

The rolling stone doesn't gather moss.

17.

Use more data if you can afford it.

18.

Don't use code you don't understand.

19.

Always do cross-validation.

20.

Avoid data leaking.

21.

Do data augmentation.

22.
Add sparsity.

23.

Beware of numerical problems.
Add epsilon.

24.

Look at the errors (e.g. misclassifications).
Do they have a structure?

25.

Do hyperparameter tuning.
Grid search, random search.

26.
Scale and normalize.

27.

Look for success stories (Kaggle).

28.

Stay curious!

A 5-min read each day keeps you up to snuff.

29.
Share your tricks!

30.

Start small, then scale up.