

Response to
Student Feedback
for
CS-EJ3211
Edition: Jan-Mar 2021

Round 1 – Three Components of ML



Predictor Function



4.2

Hypothesis Space



4.2

Loss Function



4.3

Labels



4.6

Features



4.6

Round 2 – Regression

1	2	3	4	5
---	---	---	---	---

Linear Predictor Function



4.4

Weight Vector



4.2

Squared Error Loss



4.4

Huber Loss



4.2

Training Error



4.3


Round 3 – Model Validation and Selection


	1	2	3	4	5	
Regularization				■		3.7
Model Selection				■		3.9
K-fold Cross-Validation				■		4.0
Validation Error				■		4.1
Splitting into Training and Validation Set				■		4.4
Training Error				■		4.2


Round 4 – Classification

1	2	3	4	5
---	---	---	---	---


Binary Classification  4.2

Logistic Loss  4.0

Logistic Regression  4.1

Confusion Matrix  4.1

Multiclass Classification  4.1

Decision Trees  4.0

Round 5 – Clustering



DBSCAN		4.0
Gaussian Mixture Model		3.8
K-means		4.2
Density based clustering		4.0
Hard Clustering		4.3
Soft Clustering		4.2

Round 6 – Feature Learning



Basic Idea of Feature Learning



3.9

Benefits of Feature Learning



3.9

Principal Component Analysis (PCA)



3.7

Combining PCA with Linear Regression



3.6

S: The theoretical formulas could be explained more in details with Calculus mathematics and providing examples for formulas.

T: We have revised the notebooks to emphasize the basic ideas and intuition behind machine learning methods. We try to limit the use of mathematical formulas to the absolute necessity.

S: Quizzes with only few question and one try have way too big impact on grade on those even one or two wrong answer can really impact for the grade of course which lasts like three months and that doesn't really seem balanced to me.

T: We have reduced weight of some quiz questions for the grading. Moreover, we have made the autograding for coding assignments more fine-grained so that partial solutions are also taken into account.

S: Lack of real lectures. I had to watch the available YouTube made the of Alex Jung. Some of the methods did not really open just by completing the tasks of online material.

T: We will now offer some pre-recorded “real lectures” that provide more background on the machine learning methods implemented in the Python notebooks.

S: I think that the only thing are the instructions in the final project, which sometimes were not clear to me and I had to correct my code couple of times just because of that.

T: We will provide more detailed instructions about the project requirements in the beginning of the course.

S: ...the coding evaluation is fully automated and lacks of human sense evaluation. This has caused a lot of zero points while the effort was pretty much and the approach has dismissed the semi-skilled coding which is close to the correct answer but it has evaluated as same as a person who just skipped the question with zero effort

T: We have revised the autograded assignments to ensure a more fine-grained grading. Moreover, we will prepare a list of common programming mistakes. We will also highlight the opportunity for students to ask course staff to review the autograding results.

S: The slack forums was helpful but not so student friendly, meaning the questions are answered in a way that still took significant time to find the correct answers for coding assignments.

T: We will make the role of the discussion forum more clear. The forum is meant to help with basic questions related to Python programming and not to provide partial or full solutions to student tasks.

S: I may suggest more basic recommendations the first time. I didn't have code experience before. It's hard for me to start at that time.

T: We have revised the Python quickstart notebook (“Round 0”) to help students to avoid common pitfalls.

S: I think the rounds should have opened straight away from the start, as they consisted of automatic tests. They were quite short and I had a lot of fun with them so I would've liked to do them wholly during the previous period when I had more time.

T: Our course is offered several times per year and we continuously revise the notebooks based on student feedback. To have more time for the revision we prefer to release the rounds sequentially.

Q: Some student tasks were definitely more difficult than others. Where one assignment could be completed by applying code used in an example and advancing it further, other would demand I read through scikit-learn's (frankly, a bit overwhelming) documentation in order to find what parameters or functions to use on an object for it to even work.

T: We will revise the weighting of different assignments to better reflect to varying level of difficulty.

S: feedback on assignments could be much better. It is not very clear at all.

T: We will revise the reference solutions to make them clearer.

S: The term "validating" was used in a confusing way. Both model selection and final evaluation were called "validating" despite they are very different in nature although computational methods are same.

T: Validation is conceptually very similar to testing. Strictly speaking, the difference between validation and testing is only in how the results are used. Validation errors are used for model selection. Test errors are only used for the final performance evaluation but not to further improve the method. We try to make the distinction between validation and testing clearer in the course materials.

S: Please cover the topic how to establish baseline accuracy for machine learning models. There were not too much discussion about how to interpret model performance numbers. Without prior experience, it is difficult to get idea what score values are "good", or what values indicate a useless model.

T: We will now discuss how to obtain baseline or reference levels for the performance of ML methods in some of the notebooks and lectures. In particular, we will detail how a probabilistic models allow to derive baselines.

S: Would have been nice to know the right answers to the assignments and that we learned and worked on...

Maybe more feedbacks after we finished the assignment, such as releasing some correct answers for reference

T: We have revised the instructions for how to find the reference solutions on jupyterhub after the assignments have been closed.