

## Peer-Review Questions for ML Project

Note, that {0,1,2} are grading levels, not points. Final points for submission are computed as described [here](#).

**Q1.1 -1.5 are the same as in Stage 1.**

**Q2.1 -2.5 are the same as in Stage 2.**

### **Stage 3. ML problem formulation – MODEL and LOSS**

Q3.1. Does the report clearly explain the models (hypothesis spaces) underlying all ML methods that are used in the project? Chapter 3 of [mlbook.cs.aalto.fi](http://mlbook.cs.aalto.fi) discusses the models used by some well-known ML methods.

- 0 – Models are not discussed or there is/are major mistake(s) in the discussion.
- 1 – Models are explained only partially or poorly.
- 2 – Models are explained clearly.

Q3.2. Does the report clearly explain why certain the models (hypothesis spaces) were chosen (justification). It can include discussion on the assumption of the features-label relationship (linear, non-linear), computational efficiency, model's complexity, and interpretability etc.

- 0 – Justification is not provided or contains major mistake(s).
- 1 – Justification is provided, but partially or poorly.
- 2 – Justification is provided.

Q3.3. Does the report clearly specify the loss function(s) used to evaluate the quality of a hypothesis? Note that it might be useful to use a different loss function for learning a hypothesis (e.g. logistic loss) than for computing the validation error (e.g., metrics “accuracy” as the average 0/1 loss). If using several different loss functions, both should be described.

- 0 – The loss functions used for training and validating the ML methods are not defined or defined incorrectly.
- 1 - The loss functions defined but described partially or poorly.
- 2 - The loss function is explicitly defined.

Q3.4 Does the report explicitly discuss why these loss functions and metrics were chosen? For example, “The Huber loss is used as it is robust towards outliers.” or “The accuracy (1/0 loss) is used as it is easier interpret in contrast to logistic loss used for parameters fitting.”.

- 0 – Justification is not provided or contains major mistake(s).
- 1 – Justification is provided, but partially or poorly.
- 2 – Justification is provided.

Q3.5 Does the report explicitly discuss how the validation set is constructed, e.g., using a single split into training and validation set or k-fold CV?

0 – The construction of training and validation sets are not discussed at all or discussed incorrectly.

1 – The construction of training and validation sets are discussed superficially.

2 – The construction of training and validation sets are discussed very clearly. I would be able to reproduce this construction on my own.