Question 1

Suppose we solve a binary classification task and our solution is scores with logloss. What predictions are more preferable in terms of logloss if true labels are y_true = [0, 0, 0, 0].

Correct answers:

- y_pred = [0.5, 0.5, 0.5, 0.5].

Incorrect answers:

- y_pred = [0, 0, 0, 1]. Incorrect! Technically, the loss is infinite in this case, while is it is not for other options, so it cannot be the right answer.

- y_pred = [0.4, 0.5, 0.5, 0.6]. Incorrect! What is better to predict [0.5, 0.5] or [0.4, 0.6]? To answer this question think how $\log$ function behaves. If you plot it you will clearly see that $\log(6) - \log(5) < \log(5) - \log(4)$, thus $\log(4) + \log(6) < \log(5) + \log(5)$. In fact it follows from Jensen's inequality.

Question 2

Suppose we solve a regression task and we optimize MSE error. If we managed to lower down MSE loss on either train set or test set, how did we change Pearson correlation coefficient between target vector and the predictions on the same set?

Correct answers:

- Any behavior is possible. Yes! We cannot monotonically relate MSE and Pearson correlation similarly to how e.g. R-squared monotonically related MSE,

Incorrect answers:

- The correlation did not change. Try to come up with a counterexample when correlation is zero, but MSE can be lowered down.

- The correlation was also lowered. No! Try to come up with a counterexample when correlation is zero, but MSE can be lowered down.

- The correlation became larger.

Question 3

What would be a best constant prediction for a following multiclass classification task with 4 classes? The number of objects of each class in train set is: 18, 3, 15, 24. Enter four comma separated values. Round each to two decimal places.

Correct answer:

- 0.3,0.05,0.25,0.4. To compute the best constant prediction for multi-logloss we need to divide number of objects in each class by the total size of the dataset. E.g. 18/(18 + 3 + 5 + 24) = 0.3 .

Question 4

What is the best constant predictor for R-squared metric?

Correct answers:

- Target mean. Yes! As it is up to a constant is equal to MSE metric.

Incorrect answers:

- One minus target mean. No! Well, maybe when the target mean is 0.5 this will work out :)

- Target mean divided by target variance. No! It's hard to justify this answer.

- (Log of target mean) + 1. No, this option would be true if the question asked about RMSLE metric.

- 0.5. Does not depend on data at all?..

Question 5

Select the correct statements.

Correct answers:

- Optimization loss can be the same as target metric. Yes! Sometimes we can use target metric as optimization loss. For example if our target metric is MSE.

- Optimization loss can different to target metric. Yes! Sometimes we cannot use target metric as optimization loss. For example if our target metric is accuracy.

Incorrect answers:

- Optimization loss is always different to target metric.

- Optimization loss is always the same as target metric.

Question 6

Suppose the target metric is **M1**, and optimization loss is **M2**. We train a model and monitor its quality on a *holdout set* using metrics **M1** and **M2**. Select the correct statements.

Correct answers:

- There is no definite relation between the best iterations for M1 score and M2 score.

Incorrect answers:

- If the best M1 score is attained at iteration N, then the best M2 score is always attained after N-th iteration.
- If the best M1 score is attained at iteration N, then the best M2 score is always attained before N-th iteration.
- If the best M1 score is attained at iteration N, then the best M2 score is always attained also at the iteration N. No! It is not true in general. There are exceptions though. For example if M1 is MSE and M2 is R-squared.

✓ Complete   Go to next item