McGill | *my*Courses   Winter 2019 - COMP-551-001 - Applied ...

Content   Discussions   Quizzes   Assignments   Groups   Grades   Course Evaluations   Lecture Recordings

Discussions List ⟩ View Topic ⟩ View Thread   ⚙ Settings   ❓ Help   🔍 Search Model Pre...

Model Prediction Metrics

Matthew Lesko-Krleza posted Mar 12, 2019 7:48 PM   ☆ Subscribe

Hi,

Are we allowed to only report on our models' predictions on the test set in regards to the results section in our report? We trained our models with the entirety of the training set, and made predictions on the test set. We're wondering if it would be ok to only present metrics on the test set, as opposed to also adding its score on a validation set.

Thanks,

Matt

Reply to Thread

Filter by: All Posts ⌄ | Clear filters   Show: Threaded

Will Hamilton
March 15 at 2:51 PM

Hi,

Ideally, you should have been doing model development with a validation set, rather than relying only on Kaggle. It is certainly fine to run on the entire training set before submitting to Kaggle once you have chosen a strong model through validation experiments, but in general hyperparameter optimization should be done on a validation set.

I suppose the short answer is that you would be "allowed" to do this, but a high-quality report would include some model exploration on a validation set (i.e., it would be tough to get 50/50 on the report if you don't have any validation set). Basically, we are asking that you treat the Kaggle competition like a test set in an academic paper, and it would be viewed as very poor practice to do hyperparameter optimization on the test set in an academic paper.

Hope that helps,

Will

↩ Reply

Matthew Lesko-Krleza
March 15 at 6:55 PM

That helps, thanks!

↩ Reply

◁ 1 / 1 ▷

Reply to Thread