12-748-612

Date: 12.05.2022

Name: Nicola Crimi

Keywords: error convergence, improvement test coverage, training service, wiki documentation

Overall Progress: 4

Activities completed:

- US8 - Training Service Documentation: Started to document the whole service, from the service itself over the test files etc. Defined as a group a standard which we will follow (Style Guide Google).

- Fixed code smells in training service

- Remodelling of the training service after some discussions in the group. Mainly, we wanted to implement a parameter which decides upon the continuation of the training process. Thus, we implemented the *error convergence* term which checks for the delta of losses between the current and the last sample (returned by the model.prediction() function) if the delta is below the *error convergence* we terminate the training process. If the error convergence is not met, we iterate over the defined number of epochs. However, after having implemented this new process, we do encounter some problems with pushing the accuracy in the region of 0.95-0.99. This will be analysed over the weekend and fixed on Monday (16.05.2022).

- Furthermore, we ensured that the accuracy of a model is validated against a global validation set. This was implemented to make a more reliable statement about the accuracy of a model. So whenever a new set of data is used for training a model, the validation set will be concatenated with the previous validation sets.

Current activities:

- Analysing training process, since we do not reach anymore accuracies of around 0.95-0.99. Which is probably the case because of the comparison of the loss between the current and the last sample. This implementation is probably wrong and we do have to come up with a more robust way to exit the training process.

- Update wiki.

- Enhancing overall performance

Goals or Activities for next deliverable:

- Fixing and finalising training service

- Add some additional fields on the db which are important when switching the model. We would like to know on how many files a model has been trained.

- Prepare the presentation

- Individual report

Immediate attention:

- -