12-748-612

Date: 07.04.2022

Name: Nicola Crimi

Keywords: set up train-service**,** train- & test-pipeline, jupyter notebooks, code-fetch, change to flask, discard Azure, Mongo DB, Docker

Overall Progress: 3

Activities completed:

- US3a - Train Service: Train service implemented such that the accuracy of the current model (the model which the end user is interacting) is fetched from the MongoDB. After this, the improve_model() is called and the current model is trained. After the train process, the new accuracy of the model is compared to the current accuracy. If new_acc > curr_acc, then there will be a switched initiated (future user story).
- US3b - Train & Test splitting: Setting up a train and test splitting such that an overfit of the model can be avoided. Research on optimal batch-size (see Deliverable 1) resulted in B=100. This is the most common use case.
- Code fetcher: We fetched data from the official Github API to populate our Mongo DB.
- Set up proper Github Wiki

Current activities:

- Research about how to appropriately switch the model.
- Improve dev set up with a “hot load”-possibility for docker
- Update Wiki
- First thoughts about how to set up a smooth demo.
- Research about how to guarantee less than 99% down time, since this is an architectural issue which could potentially cause troubles in the last two sprints.

Goals or Activities for next deliverable:

- US4 - Update Model: The switch of the newly trained model with a higher accuracy must be initiated and implemented in seamless way.
- US12 Base Training Dataset: The goal of this activity is to trigger the train service manually such that the train process can be observed and be tweaked in case of a) any unforeseen circumstances b) demonstration reasons.

Immediate attention: -

General remark: Massive struggles with dockerization of our Flask API´s. No previous Docker knowledge in the team. We pivoted away from our Azure cloud functions because Azure can’t be dockerized. Due to this circumstances, we decided to switch to the relatively easy Flask framework. Additionally, the new M1 chips of Apple are causing troubles with Docker.