Master thesis

University of Tartu

April 5, 2023

1 Meeting notes 8

1. **IMPORTANT:** It is not a competition(in terms of results) since we are not inventing any methods. Thus we don't care about good results.
2. Start putting everything into one document for Yova to make suggestions and later analyze results. From this moment no more report, but instead I will be inserting everything in my thesis report already.
3. Continue training models until the converge.
4. No need to do bootstrapping test since as researchers we did everything(sent them emails, used their codebased, and followed their instruction from [Neeman et al., 2022]).
5. Retrain on of the fine-tuned model with val-loss instead of val-accuracy. Choose the best one and compare with test results on val-accuracy.
6. Put threshold to 100 epochs. We can stop at any point when we recognize that the model converged. Early stopping might be used too, but for now I think I could analyze convergence manually.
7. There is clear difference in terms of results between my and [Neeman et al., 2022] codebases. We will be using whichever gives best results.

Todos:

1. Retrain not fully converged models
2. Write about dataset and results sections
3. Finish all the hyperparameters search

4. Prepare slides for my next presentation
5. Read HPC documentation, to understand allocation of two machines.

Desirable outcome:

1. Write-up for experiment section.
2. Finish hyperparameter search
3. Improved presentation of my experiments

References

[Neeman et al., 2022] Neeman, E., Aharoni, R., Honovich, O., Choshen, L., Szpektor, I., and Abend, O. (2022). Disentqa: Disentangling parametric and contextual knowledge with counterfactual question answering.