Section	Details

Progress Overview (inkl. Reasons for Deviations of the Plan)	✔ Problem Statement
	✎ Finalizing documentation of Literature Review
	✎ Reproducing PASSION results
	    − had to fix issues in current evaluation package
	    − to reproduce an experiment on GPUHub takes 32.5h (6.5h x 5 folds) - and there are 4 experiments
	✘ evaluate baseline with fairness metrics
	    − Reproduction is not yet done

Accomplishments	✔ Categorization of the methods ✔ Got the evaluation running ✔ Progress on the report

Risks and Measures	Issues with interpreting the results
	  • Clarifying the open questions
	The fine-tuning-experiments takes to long
	  • Improve checkpoint handling
	  • Use another machine (afraid to lose information and canceled runs due to getting disconnected from GPUHub)
	Focus loss bc of tendency to want to fix / improve systems
	  • Setting clear boundaries for myself: only do what must be fixed to enable running the mitigation experiments efficiently

Next Steps	• Get checkpoints for each experiment • Evaluate the fairness for each subgroup on the experiment checkpoints
Discussion Points / Questions	• Have I understood the experiment result csv correctly? *(- Details in presentation)* • Do you have any tips on how to work efficiently on the GPU Hub? • How is the correct naming for the attributes in the PASSION dataset? Maybe I misunderstood the terms Features and Labels
Additional Notes	I was sick for roughly 4 days.
Attachments	• experiment_standard_split_conditions_passion.csv

Protocol	Topic
	  •
	Topic
	  •