PS9. Reproducible AI & Experiment Tracking using MLOps

Organised by Viktor Stenby Johanson (Technical University of Denmark), Joakim Bruslund Haurum (Aalborg University) and Kenneth Borup (Aarhus University)

Session overview

Being able to replicate an experiment and obtain the same results is a crucial part of the scientific method, though when it comes to machine learning applications, machine learning practitioners might often find themselves with a plethora of confusing Jupyter notebooks, where executing the cells in a very specific order is the best hope of reproducibility.

In this session, Weights and Biases will tell about their MLOps platform for experiment tracking, dataset versioning and model management. Then, all attendees will invited to be putting their ML skills to use in a Kaggle InClass competition. Finally, Danish startup Alvenir will tell about how they manage their experimental budget when training large models.

Speakers

- Morgan McGuire, Growth Machine Learning Engineer at Weights & Biases
- Martin Carsten Nielsen, Co-founder and CEO at Alvenir

Programme

10:15 – 11:00 **Weights & Biases**
Morgan McGuire from Weights & Biases will walk through ML tools to show how to track, compare and visualise ML experiments in practice.

11:00 – 11:45 **Kaggle InClass Classification Competition featuring W&B**
Weights & Biases will run a classification competition to give everyone a taste of how to use W&B tracking. A starter Colab is provided, and submissions are made via Kaggle InClass competitions.

11:45 – 12:15 **Managing Expectations and Learning from Experience when Experimenting on a Budget, Alvenir**
As the size of our models increases, warranting for long training runs and massive compute requirements, the experimental budget feels tighter and tighter. Subsequently this requires us to be smarter about our training choices and hypotheses to ensure we gain tangible value from each run, this is where the MLOps cycle plays a crucial role.