

The Mutual Effects of Example Importance and Hyperparameter Optimization

For the hyperparameter optimization step in a machine learning or deep learning pipeline, the dataset is usually randomly split into a number of validation folds, and the hyperparameter configurations will be evaluated on individual folds. Given the difference between the importance of examples in the dataset, it might be beneficial to construct the validation folds using heuristics that consider notions of example importance. One goal of this thesis is to investigate if using such heuristics can lead to finding better hyperparameters.

On the other hand, it is also interesting to investigate the training dynamics of examples in the validation fold corresponding to the winning hyperparameter trial when we train the final model on the full training dataset.

Studying these two questions can lead to more insights into how important examples are influenced by - and influence - the hyperparameter optimization step, and can lead to faster training of better models.

Contact: **Sina Sheikholeslami** <sinash@kth.se>

Distributed Computing Group, SCS Division, KTH

Examiner: **Prof. Vladimir Vlassov** <vladv@kth.se>

Distributed Computing Group, SCS Division, KTH

Technical Requirements:

- Familiarity with deep neural networks and hyperparameter optimization techniques
- Proficiency in developing libraries using PyTorch (preferred) or TensorFlow

References:

- Ariafar et al. "**Weighting is worth the wait: Bayesian optimization with importance sampling.**" arXiv preprint arXiv:2002.09927. 2020.
- Toneva et al. "**An Empirical Study of Example Forgetting during Deep Neural Network Learning.**" In *International Conference on Learning Representations*. 2018.
- Katharopoulos and Fleuret. "**Not all samples are created equal: Deep learning with importance sampling.**" In *International conference on machine learning*, pp. 2525-2534. PMLR, 2018.

