

Paper Assignment DeepBayes Summer School 2019 (deepbayes.ru)

Е. Koreko

8 апреля 2019 г.

Sida I. Wang and Christopher D. Manning. Fast Dropout Training. ICML 2013.

- **How would you summarize the main idea of this paper in one sentence?**

How to achieve the benefit of dropout training without actually sampling, thereby using all the data efficiently.

- **In which cases the fast dropout approximation is exact? Why is this the case?**

Fast dropout is directly applicable to dropping out the final hidden layer of neural networks. Because of use Gaussian approximating that is justified by the central limit theorem and empirical evidence.

- **Where do the time savings come from?**

It comes from drawing samples, because drawing samples of the approximating Gaussian S of $Y(z)$, a constant time operation and we speeding up process by order of magnitude.

- **Can fast dropout be used to optimize w.r.t. dropout rates p in a meaningful way? If so, how? If not, why?**

Yes, it can. In the article provided deterministic and easy-to-compute objective function approximately equivalent to that of real dropout training.