Seminar Review 4.20-4.26

Jiaxin Hu

04/20/2020

IFDS 4.20

Title: Bad global minima can be reached by overparameterized model with SGD

Author: Dimitris Papailiopoulos

This talk explains why overparameterized model trained with SGD can work well in many cases. Contrast to the consensus that "there is no bad global minima", the overparameterized SGD is attracted by the bad global minimia. The bad global minima can be considered as a complex boundary in classification and it can lead vanilla methods to get 100% training accuracy while poor performance in tests. The initial point can be set randomly with unlabeled data, which leads the SGD converge to the bad minima quickly. To eliminate the effect from the random initialization, a regularization term can be added to let SGD escape and have a good test performance finally.

Questions:

1. As presenter shows, the vanilla SGD with random initialization and SOTA(the overparameterized SGD method) with adversarial initialization point have similar performance in practice. Which one should we use? Is there any pre-judgement steps?

Possible Answer: The overparameterized SGD has smaller computational complexity. The data which has poor performance with traditional methods may suitable for this method.

SILO 4.22

Title: Biologically interpretable machine learning modeling for understanding functional genomics - with an application to the human brain

Author: Daifeng Wang

This talk firstly explains the complexity of the "rules" from human genotype to phenotype(diseases). The "rules" can not be discovered by GWAS because the genes express distinctively among different cells. A deep learning network, DSPN, is proposed to find the mechanism from SNPs to phenotype. Every layer in DSPN is interpreted by the biological meaning such as tissues, mRNA numbers, non-coding genes. The trained pathway in the network is the mechanism we desired and is used for brain diseases predication. The experiments show the improved accuracy of DSPN over logistic regression.

Questions:

1. Is there any validation for each step? If you only do validation for the phenotype, how can you guarantee the mechanism is almost correct?
2. How to decide the number of layers and the number of nodes for each layer? Is there any risk of overfitting?

Possible Answer: Step-wise validation is needed and we can integrate the prior molecular knowledge or scientific conclusion to verify the mechanism. Maybe the specific parameters for the network are adjusted to get a higher accuracy. Also, this model is very complex, expensive and consuming and people should discuss whether it is worth to use such a model.