DS-GA 1006 Fall 2017
Capstone Project Status Update I

Voyagers:
Yurui Mu, Sudhir Nallam
Xiaoyu Wang, Wenqing Zhu

Nov 16, 2017

In past week, we have scrutinized [1] and understood the mechanics of generative adversarial network. To further explain, we have understood that the process is to train the classifier f with an adversarial generator r, such that

$$p(f(X;\theta_f) = s|z) = p(f(X;\theta_f) = s|z') \tag{1}$$

In words, we are looking for a prediction function f which is pivotal quantity with respect to the nuisance parameter. This implies that $f(X;\theta_f)$ and Z are independent random variables. To train such generative adversarial networks, we then consider the value function

$$E(\theta_f, \theta_r) = L_f(\theta_f) - L_r(\theta_f, \theta_r) \tag{2}$$

that we optimize by finding the minimax solution

$$\hat{\theta_f}, \hat{\theta_r} = arg \min \max E(\theta_f, \theta_r) \tag{3}$$

Meanwhile, to complete our roadmap in model experiments, we have submitted tasks on Prince to train Gated Recurrent Neural Network model. We then can compare the AUC score and ROC curve with our previous results. More specifically, we are going to compare the background efficiency when signal efficiency is constrained to be 0.5 with previous results from Simple Recursive Neural Network and Simple Recursive Neural Network with 1x1x1 Convolution as activation layer.

Moreover, as we have observed performance improvement by using three layers of 1x1x1 convolution, we want to parametrize the activation function through controlling the number of convolution layers to use, and optimize for best result.

References

[1] Gilles Louppe, Michael Kagan, and Kyle Cranmer. Learning to Pivot with Adversarial Networks. 2016.