project_proposal

October 1, 2019

Deep Neural networks has become a promising method for solving many real word problems. It has potential wide applications in different fields, including image recognition, face recognition and object detection (Yedida, 2019). The accuracy and efficiency of training neural networks are greatly impacted by choosing the activation function and hyper-parameters. To optimize the deep neural networks, manual adjustment of several hyper-parameters, including learning rate, weight decay and dropout rate, is needed. Among them, the learning rate is generally considered to be the most important hyper-parameters. Learning rate controls the speed of adjusting the weights of the neural networks with respect to the loss gradient. When the learning rate is too small, it could take a long time to converge. When the learning rate is too large, the local minima could be missed and cause divergence problems. Therefore, optimizing the learning rate is crucial for improving the performance of the neural networks. Typically, the learning rate is configured randomly or set by the user according to intuition or past experiences. However, this method is not only time-consuming but also hard for getting the proper learning rate. In recent work, researchers proposed a non-monotonic learning rate scheduling system and agreed that it offers faster convergence compared to a fixed learning rate value (Seong, 2018).

In this project, we would like to study the existing algorithms of adaptive learning rate, evaluate and explore the performance of these algorithms by computational experiments, and propose a new method to mathematically identify an optimal learning rate and improve the performance of existing algorithms. This project is worth trying since the improvement of optimizing learning rate could largely boost the efficiency and accuracy of training neural networks. Other than the potential value of this project, we believe this project is feasible for a three month period for mainly three reasons. First, many recent works provide theoretical framework for the project. There are lots of recent works propose potential novel ways to adaptively change the learning rate and could provide intuition to our project. Second, the project group members have experiences in research and training neural networks models. The complementary of group member's knowledge backgrounds ensures the overall practice and theoretical basis of success of the project. Third, the project is well designed and will follow a proper timeline. We plan to use one month to study the existing algorithms, one month to conduct the computational experiments and explore the potential improvement and one month to validate our idea and compare results. Since we want to validate the reliability and the performance of the algorithms in different tasks, multiple datasets will be used.

1. Rahul Yedida, Snehanshu Saha, A novel adaptive learning rate scheduler for deep neural networks, arXiv:1902.07399, 2019.
2. Sihyeon Seong, Yekang Lee, Youngwook Kee, Dongyoon Han, and Junmo Kim. Towards atter loss surface via nonmonotonic learning rate scheduling. In UAI2018 Conference on Uncertainty in Arti_cial Intelligence. Association for Uncertainty in Arti_cial Intelligence (AUAI), 2018.