project_proposal

October 1, 2019

Deep Neural networks has become a promising method for solving many real-world problems. It is widely applied in different fields, including image recognition, face recognition and object detection (Yedida, 2019). The accuracy and efficiency of trained neural networks are greatly affected by the choice of the activation function and hyper-parameters. To optimize the deep neural networks, several hyper-parameters, including learning rate, weight decay and dropout rate, are manually adjusted. Among these hyper-parameters, the learning rate is generally considered to be the most important. 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 resulting in divergence problem. Therefore, optimizing 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 have 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 existing algorithms for 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 significant since the improvement of optimizing learning rate could largely boost the efficiency and accuracy of training neural networks. Also, the experience we will garner from this project will help deepen our understanding of many of the concepts learned in the course (ORIE 4741 - Learning with Big Messy Data) that emphasized the theory and significance of learning rate. Other than the potential value of this project, we believe this project is feasible for a three-month period for three reasons. First, there are lots of recent work that have proposed potential novel ways to adaptively change the learning rate. These recent works will provide the theoretical framework for the project. Second, the project group members have prior experiences in research in general and neural network models in particular that will be employed in this project. Third, the project is well designed and will follow a structured and realistic timeline. We plan to use one month to study the existing algorithms, one month to conduct computational experiments and explore the potential improvement and one month to validate our idea and compare results. Finally, 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.