The Name of the Title is Hope

Your Name
University of Passau
your.name@uni-passau.de

ABSTRACT

Here goes your abstract.

1 INTRODUCTION

Adaptive learning algorithms like AdaGrad and Adam are being extensively used thanks to there shorter learning time, However the benefit of using such algorithms on datasets where the number of model parameters to train on is larger than the available data, is to be investigated. In this paper we try to reproduce the experiment done by Wilson & al. [WRS+17] in order to investigate the results of using adaptive learning algorithms compared to there non-adaptive counter parts in the case where the available number of points in the data is smaller than the number of model parameters to train.

2 BACKGROUND

In this section we are going to introduce the algorithms used in the experiment. We are going to introduce them using an informal and a mathematical description. To visualize how these algorithms function differently we suggest that you take a look at Lili Jiang's tool Gradient Descent Viz [Jan20].
In the next sections we define x to be a vector or a scalar, t is the iteration number, L is the loss function, W is the weights matrix, α is the learning rate and β_i are decay rates.

2.1 Non-adaptive algorithms

Stochastic Gradient descent [KW52] is the most basic Learning algorithm. To minimize the Loss function it calculates the function's gradient at a particular point and updates the point coordinates with the negative value of that gradient. Formally, in the iteration t SGD calculates the next point coordinate using the following formula:

$$x_{t+1} = x_t - \alpha * \nabla L(x_t)$$

One problem with SGD is that its learning speed is very slow [DAN+18] and can get caught in a local minimum easily [BJ19].

2.2 Part Two

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

3 PART THREE

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

4 RELATED WORK

5 CONCLUSION

6 MODIFICATIONS

REFERENCES

[BJ19] A. Bouillard and P. Jacquet. Quasi black hole effect of gradient descent in large dimension: Consequence on neural network learning. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8365–8369, 2019.

[DAN+18] E. M. Dogo, O. J. Afolabi, N. I. Nwulu, B. Twala, and C. O. Aigbavboa. A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. In *2018 International Conference on*

Computational Techniques, Electronics and Mechanical Systems (CTEMS), pages 92–99, 2018.

[Jan20] LiLi Jang. Gradient descent viz. https://github.com/lilipads/gradient_descent_viz, Online, accessed 11 June 2020.

[KW52] J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.*, 23(3):462–466, 09 1952.

[WRS+17] Ashia C. Wilson, Rebecca Roelofs, Mitchell Stern, Nathan Srebro, and Benjamin Recht. The marginal value of adaptive gradient methods in machine learning, 2017.