Devoir Maison

Student name: *Durand Enzo 21107517*

Course: *Algorithmes sur les arbres et les graphes en bio-informatique (AAGB)*
Due date: *December 10th, 2021*

Sujet : analyse d'une base de donnée de bio-informatique

Gradient descent is one of the most popular algorithms to perform optimization and by far the most common way to optimize neural networks. At the same time, every state-of-the-art Deep Learning library contains implementations of various algorithms to optimize gradient descent (e.g. lasagne's, caffe's, and keras' documentation). These algorithms, however, are often used as black-box optimizers, as practical explanations of their strengths and weaknesses are hard to come by.

This blog post aims at providing you with intuitions towards the behaviour of different algorithms for optimizing gradient descent that will help you put them to use. We are first going to look at the different variants of gradient descent. We will then briefly summarize challenges during training. Subsequently, we will introduce the most common optimization algorithms by showing their motivation to resolve these challenges and how this leads to the derivation of their update rules. We will also take a short look at algorithms and architectures to optimize gradient descent in a parallel and distributed setting. Finally, we will consider additional strategies that are helpful for optimizing gradient descent.