

A comprehensive study on SGD variants using Transfer Learning

Sai Krishna Karthikeya Challa

The gradient descent is an important part of building a neural network which reduces the losses. It has many variants including momentum, RMSProp, ADAM & Adagrad. Each of these optimizers updates the weights and biases in different ways. Transfer learning utilizes the already proven models like mobilenet, resnet50, VGG16 etc. and change them to fit our particular application.

So, understanding the behavior of the optimizers for loss and accuracy calculations on variety of datasets would be useful. When compiled using various models from transfer learning would help in understanding which optimizer helps us fit the model better.

A thorough understanding of neural networks, gradient descent and models is needed as a background for this project which is obtained through coursework and tutorials.

The data for this project will be obtained from Tensorflow Datasets which includes: oxford_flowers102 and cassava leaf.

In this project, the mobilenet, inception and resnet50 models will be utilized to fit the application of image classification on above mentioned datasets. Each of these models will be compiled using gradient descent variants RMSProp, ADAM, ADAGRAD, ADAMAX and SGD. The data will be divided into testing, training and validation sets and some methodologies to avoid overfitting like dropout and early stopping will also be implemented to make the model better.

The image classification accuracy metric will be used to do a comparative study on transfer learning models and optimizer variants. All the results will be provided in the form of plots and figures for a visual understanding of the study.