https://github.com/TheSarang/CIFAR10-Image-Classification

video tutorial : <https://www.youtube.com/playlist?list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv>

Stochastic Gradient Descent, epoch, and batch

1. Stochastic Gradient Descent:
   1. An optimization algorithm to train ML models (ANN in deep learning etc.)
   2. Goal: to find a set of internal model parameters

Performance measure such as logarithmic loss or mean squared error

* 1. Optimization means searching process
  2. Iterative: means each discrete step, including calculating, comparing, and updating
  3. Update procedure: depends on different algorithms. For ANN is backpropagation update algorithm

1. Sample:
   1. Contains inputs and outputs
   2. Other names: an instance, an observation, an input vector, or a feature vector.
2. Batch:
   1. The batch size is a hyperparameter that defines the number of samples to work through before updating the internal model parameters.
   2. “for loop”
   3. A training dataset can be divided into one or more batches.
   4. batch gradient descent VS stochastic gradient descent
3. epoch:
   1. The number of epochs is a hyperparameter that defines the number times that the learning algorithm will work through the entire training dataset.
   2. a for-loop over the number of epochs where each loop proceeds over the training dataset. Within this for-loop is another nested for-loop that iterates over each batch of samples, where one batch has the specified “batch size” number of samples.
4. Epoch VS batch:
   1. The batch size is a number of samples processed before the model is updated.
   2. The number of epochs is the number of complete passes through the training dataset.