

Script for presentation

Slide 1: Today I am going to present the work of mariya toneva and co-authors titled as an empirical study of example forgetting during deep neural networks learning. This was presented in ICLR 2019 and had opened a new line of work for other researchers to work on catastrophic forgetting. In this presentation our focus will remain narrowed towards the implementation part of the paper. Authors made their work public on github. Hence I cloned the repository and have tried to reproduce the results for the same. In further slides I will show two dataset because the approach to solve the problem of example forgetting for CIFAR has done differently than MNIST.

Slide 2: The first step is to setup the arguments and fix the parameters. In other words, the arguments which are passed in the python command line are channelized via python dictionary. Here we are keeping all arguments and parameters to their default values to get the first output.

Slide 3: Load the splitted dataset to the trainset and testset respectively.

Slide 4: in this step you train the learning algorithm with the MNIST dataset for 200 epochs. After every epoch you want to save the paramters from the learning for forgetting statistics. Save the data to dictionary and save that to pkl file.

Slide 5: Meanwhile in the training process you want to save the indices of the real examples in the batch. that is within `tain()` do permute samples on each mini-batch and get

indices of each samples. If provided transformation flag in the parameter then transform the dataset appropriately. Get indices for the mini-batch which is a subset of permuted indices.

Slide 6: Once done, do the forward pass and compute the losses. Also, get predicted labels. Cross entropy loss is used. The code below get the index of the max in data which is a tensor.

Slide 7: get accuracy for each sample in teh mini batch and compute statistics. list of T and F where predicted and targets are equal. iterate over indices in the mini-batch. Get index in original dataset (not sorted by forgetting). Compute miss-classification margin. get the probability outputs for the correct expected class.

Slide 8: Combine the stats and save a pkl file.

Slide 9: For CIFAR dataset CIFAR10 was first trained with no sorting, no sample removal, no data augmentation and no cutout. For each event, calculate loss and predict output. Updata accuracy, loss stats of each event and save it in example_stats

Slide 10: Sort dataset based on forgettable events Step 1: Use accuracy value after training the CIFAR data to compute whether event is learned or unlearned or forgettable - Learned event: event with accuracy of 1 - Unlearned event: event with accuracy of 0 - Forgettable event: event with accuracy dropping from 1 to 0 # Forgetting event is a transition in accuracy from 1 to 0. # Find all presentations when forgetting occurs # Find number of presentations needed to learn example, #

e.g. last presentation when acc is 0 # Find the misclassification margin for each presentation of the example # Find the presentation at which the example was first learned, # e.g. first presentation when acc is 1

Slide 11: Sort dataset based on forgettable events Step 2: Sort the example_stats to rank the sample from the highest forgetting count to the lowest forgetting count Step 3: Save the sorted file with a stat of sample ID and sample values

Slide 12: Process 3: Train the learning algorithm again and repeat process 1 and 2 with random data removal of samples / examples

Slide 13: Process 4: Train the learning algorithm again and repeat process 1 and 2 with sorted removal of samples / examples