

APS360 Fundamentals of AI

Jake Snell

Lecture 4; May 16, 2019

Agenda

Last Class:

- ▶ Batching
- ▶ Learning Rate
- ▶ Training Curve (Learning Curve)

Today:

- ▶ Hyperparameter Tuning
- ▶ Validation Set
- ▶ **Lab 2:** Hyperparameter Tuning

A bit about lab 2

You'll be implementing a neural network to differentiate between images of cats and images of dogs.

The neural network models are built for you. Your job is to train the networks using different learning rates, batch sizes, and other settings.

Hyperparameters

Hyperparameters

Hyperparameters are neural network settings that cannot be optimized using an optimizer

Hyperparameter examples

- ▶ Size of network
 - ▶ Number of layers
 - ▶ Number of neurons in each layer
- ▶ Choice of Activation Function
- ▶ Learning Rate
- ▶ Batch Size

Q: How do we tune hyperparameters?

Hyperparameter examples

- ▶ Size of network
 - ▶ Number of layers
 - ▶ Number of neurons in each layer
- ▶ Choice of Activation Function
- ▶ Learning Rate
- ▶ Batch Size

Q: How do we tune hyperparameters?

By trying various values, see which one trains “best”

Training Curve

- ▶ **x-axis:** epochs or iterations
- ▶ **y-axis:** loss, error, or accuracy



Code!

- ▶ Go through “Neural Network Training”, code snippets [8-15]

Tracking Training Curve

Last week, we proposed to set aside a **test set** to help determine how our model will perform on a new data set.

Should we track the **test accuracy** (accuracy over test set) in our training curve?

The problem with tracking test accuracy

If we track test loss/accuracy in our training curve, then:

- ▶ We may make decisions about neural network architecture using the test accuracy!
- ▶ The final test accuracy will not be a realistic estimate of how our model will perform on a new data set!

Validation Set

We still want to track the loss/accuracy on a data set not used for training

Idea: set aside a separate data set, called the **validation set**

- ▶ Track validation accuracy in the training curve
- ▶ Make decisions about hyperparameters using the validation set

More Code!!

- ▶ Go through “Neural Network Training”, code snippets [16-17]
- ▶ Note: Try setting the DataLoader parameter **shuffle=True** and see its effect on the training curve. What do you think this setting does?

Still More Code!!

- ▶ Go through “Neural Network Training”, code snippets [18-]