Author: Benjamin Smidt
Created: September 22, 2022
Last Updated: September 30, 2022

Deep Learning Resources

Contents

1 General **3**

2 Data Preprocessing **3**
2.1 Weight Initialization 3
2.2 Batch Normalization 3

3 Loss Functions **3**

4 Architectures **4**

5 Hyperparameters **4**
5.1 Update Rules 4
5.2 Regularization 4
5.3 Dropout 4

6 Hardware and Software **5**

1 General

1. "Training Neural Networks I." CS231N Convolutional Neural Networks for Visual Recognition, Stanford University, 2022. [*Modeling a Neuron, Activation Functions, ReLU*]

2 Data Preprocessing

1. "Training Neural Networks II." CS231N Convolutional Neural Networks for Visual Recognition, Stanford University, 2022. [*Mean Subtraction, Normalization, PCA and Whitening*]

2.1 Weight Initialization

1. "Training Neural Networks II." CS231N Convolutional Neural Networks for Visual Recognition, Stanford University, 2022. [*Pitfalls All Zero, Small Random Numbers, Calibrating Variances, Batch Normalization*]

2. He et al, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification", ArXiv 2015 [*Original Kaiming Initialization Paper*].

2.2 Batch Normalization

1. Sergey Ioffe and Christian Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift" , ICML 2015. [*Batch Normalization Paper*]

3 Loss Functions

1. "Training Neural Networks I." CS231N Convolutional Neural Networks for Visual Recognition, Stanford University, 2022. [*Problem of Large Number of Classes, Attribute Classification, Regression v. Classification*]

4 Architectures

1. "Training Neural Networks I." CS231N Convolutional Neural Networks for Visual Recognition, Stanford University, 2022. [*Layer-Wise Organization, Naming Conventions, Ex. Feed-forward, Representational Power, Capacity*]

5 Hyperparameters

5.1 Update Rules

1. "Training Neural Networks III." CS231N Convolutional Neural Networks for Visual Recognition, Stanford University, 2022. [*SGD, Momentum, Nesterov Momentum, Adagrad, RMSprop, Adam*]

2. Diederik Kingma and Jimmy Ba, "Adam: A Method for Stochastic Optimization", ICLR 2015 [*Original Adam Paper*]

5.2 Regularization

1. "Training Neural Networks II." CS231N Convolutional Neural Networks for Visual Recognition, Stanford University, 2022. [*L2, L1, Max Norm Constraints*]

5.3 Dropout

1. "Training Neural Networks II." CS231N Convolutional Neural Networks for Visual Recognition, Stanford University, 2022. [*Dropout, Inverted Dropout, Code Implementation*]

2. Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting." University of Toronto, 2014

3. Hinton, Geoffrey E. et al. "Improving neural networks by preventing co-adaptation of feature detectors." ArXiv, 2012. [*Dropout*]

6 Hardware and Software

1. Chadha, Amani. "CS231N Deep Learning Hardware and Software." Aman's AI Journal, 2020.