Deeplearning.ai III Deep Neural Networks

Train/dev/test sets

Iterative process

Mismatched train/test distribution

Bias and variance

Basic recipe for machine learning

Regularizing neural networks

Regularization

Lambda – regularization parameter

Logistic regression

L1 regularization

L2 regularization

Frobenius Norm

Weight Decay

Regularization reduces overfitting

Dropout Regularization

Shrink Weight

Keep\_prob

Regularization Methods

Data Augmentation using symmetry

Data Augmentation using rotation

Early Stopping

Orthogonalization

Normalizing inputs

Normalizing training sets

Vanishing/Exploding gradients

Xavier initialization

Weight Initialization for deep networks

Numerical approximation of gradients

Check derivative computation

Approximation error

Gradient checking

Gradient checking implementation notices

Mini-batch Gradient Descent

Batch vs mini-batch gradient descent

Vectorization (compute efficiently)

Mini-batch gradient descent algorithm

Training with mini-batch gradient descent

Choosing the mini-batch size

Exponentially weighted averages

Implementing exponentially weighted average

Bias correction in exponentially weighted averages

Bias Correction

Gradient descent with momentum

Implementation details

RMSprop

Adam Optimization Algorithm

Hyperparameter choice

Learning rate decay

Learning rate decay methods

Discrete Staircase

Manual Decay

Problem of Local Optima

Local Optima in Neural Networks

Saddle Point

Problem of plateaus

Hyperparameter tuning

Tuning process

Hyperparameters

Grid search should not be used

Coarse to fine

Using an appropriate scale to pick hyperparameters

Picking hyperparameters at random

Appropriate scale for hyperparameters

Tuning using log-scale

Hyperparameters for exponentially weighted averages

Re-test hyperparameters occasionally

Babysitting one model

Training many models in parallel

Panda vs caviar

Batch optimization

Normalizing activations in a network

Normalizing inputs to speed up learning

Implementing batch norm

Fitting batch norm into a neural network

Adding batch norm to a network

Learning on shifting input distribution

Batch norm as regularization

Batch norm at test time

Multi-class classification

Softmax regression

Softmax layer

Training a softmax classifier

Deep learning frameworks

Tensorflow