Stanford CS230 Notes

1 Introduction to Deep Learning

1.1 What is a Neural Network?

neuron, links.

1.2 Supervised Learning with Neural Networks

Supervised Learning

Examples: Standard NN, Convolutional NN, Recurrent NN

Structured Data: tabular data; Unstructured Data: audio/image/text

1.3 Why is Deep Learning taking off?

Amount of labeled data.

2 Basics of Neural Network Programming

2.1 Binary Classification

2.1.1 Binary Classification

image $\rightarrow$ 1 (cat) vs 0 (non cat)

2.1.2 Notation

m: number of examples

n_x: input size

n_y: output size

x: input, column vector

y: output, 0/1

X: input matrix, shape = (n_x, m)

Y: output matrix, shape = $(1, m)$

2.2 Logistic Regression

Given: $x \in \mathbb{R}^{n_x}$, $0 \leq \hat{y} \leq 1$

Parameters: $w \in \mathbb{R}^{n_x}$, $b \in \mathbb{R}$

Output:

$$z = w^T x + b \tag{1}$$

$$\hat{y} = \sigma(z) \tag{2}$$

$$\sigma(z) \approx \frac{1}{1 + e^{-z}} \tag{3}$$

$$z \approx \infty, \sigma(z) \approx \frac{1}{1 + 0} = 1 \tag{4}$$

$$z \approx -\infty, \sigma(z) \approx \frac{1}{1 + \infty} = 0 \tag{5}$$

Simplified Parameters: $x_0 = 1$, $x \in \mathbb{R}^{n_x + 1}$

$$\theta_0 = b, \theta_1 ... \theta_{n_x} = w \tag{6}$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ ... \\ \theta_{n_x} \end{bmatrix} \tag{7}$$

$$\hat{y} = \sigma(\theta^T x) \tag{8}$$

2.3 Logistic Regression cost function

Given $\{(x^{(1)}, y^{(1)}, ..., (x^{(m)}, y^{(m)}\}$, want $\hat{y}^{(i)} \approx y^{(i)}$.

Loss (error) function: