Neural Network Overview

미래연구소 12기 3주차

0. Week 2에서 배운 내용

Logistic Regression(로지스틱 회귀) - Binary Cross Entropy

Gradient Descent(경사하강법)

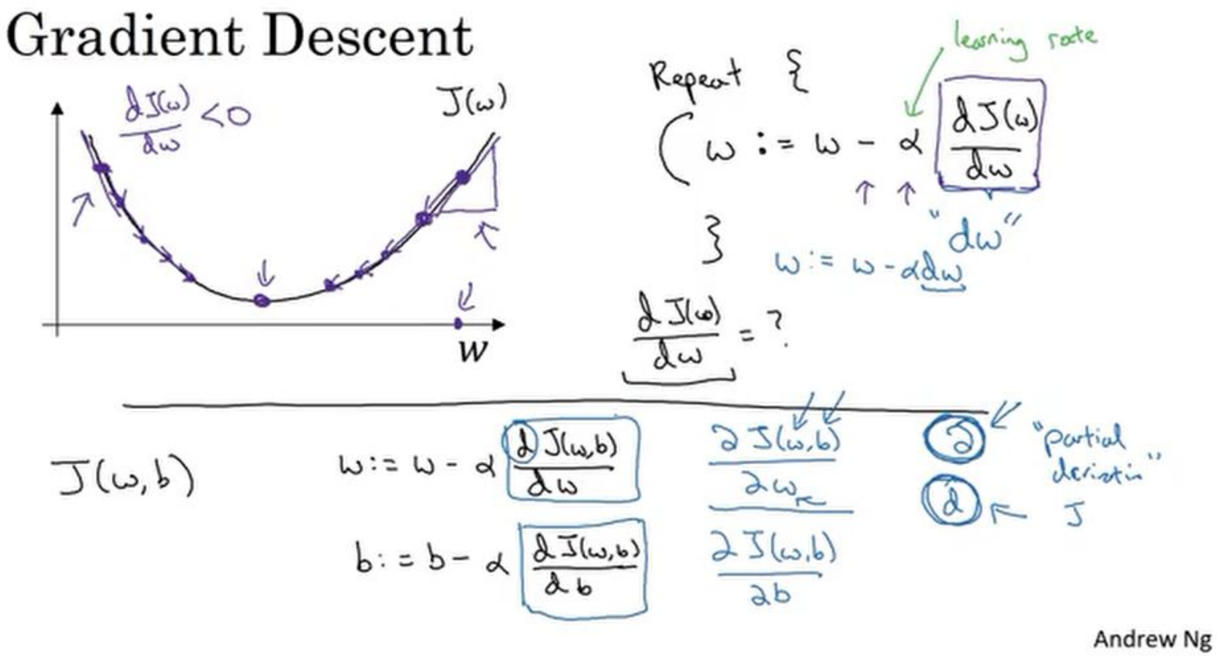
Backward Propagation(오차역전파)

Vectorization(벡터화)

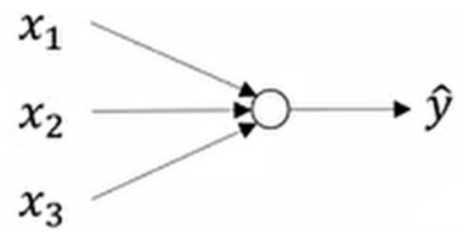
Binary Cross Entropy

$$BCE = -\frac{1}{N}\sum_{i=0}^{N} y_i \cdot log(\hat{y}_i) + (1 - y_i) \cdot log(1 - \hat{y}_i)$$

Gradient Descent



한 번의 parameter update

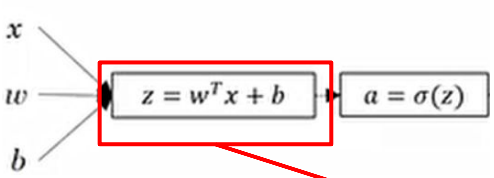
<table>
<tr><td></td><td></td><td>계산 과정</td></tr>
<tr><td>모델</td><td>
</td><td rowspan="2">(x = (x1, x2, x3), y)가 m개 주어진다. = (X, Y)

1) w1, w2, w3, b initalize

2) np.dot(w.T, X) + b => Z
sigmoid(Z) => A = Y^

3) compute J(w, b)

4) backpropagation => dw, db

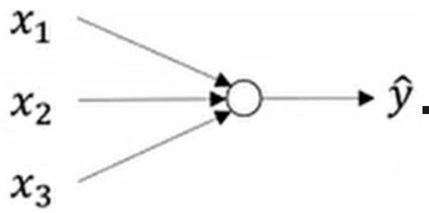
5) w = w - αdw, b = b - αdb</td></tr>
<tr><td>computation graph</td><td>
</td></tr>
<tr><td></td><td colspan="2">$$np.dot(w.T, X) + b = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1^{(1)} & x_1^{(2)} \\ x_2^{(1)} & x_2^{(2)} \\ x_3^{(1)} & x_3^{(2)} \end{bmatrix} + b$$</td></tr>
</table>

실제 행렬 계산

<새로운 표현법>

w 앞의 숫자: unit index = 현재 layer index

w 뒤의 숫자: feature index = 이전 layer index



$$np.dot(w.T, X) + b = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1{}^{(1)} & x_1{}^{(2)} \\ x_2{}^{(1)} & x_2{}^{(2)} \\ x_3{}^{(1)} & x_3{}^{(2)} \end{bmatrix} + b$$

$$np.dot(w.T, X) + b = \begin{bmatrix} -w.T- \end{bmatrix} \begin{bmatrix} x_1{}^{(1)} & x_1{}^{(2)} \\ x_2{}^{(1)} & x_2{}^{(2)} \\ x_3{}^{(1)} & x_3{}^{(2)} \end{bmatrix} + b$$

$$np.dot(w.T, X) + b = \begin{bmatrix} w_{1\ 1} & w_{1\ 2} & w_{1\ 3} \end{bmatrix} \begin{bmatrix} x_1{}^{(1)} & x_1{}^{(2)} \\ x_2{}^{(1)} & x_2{}^{(2)} \\ x_3{}^{(1)} & x_3{}^{(2)} \end{bmatrix} + b$$

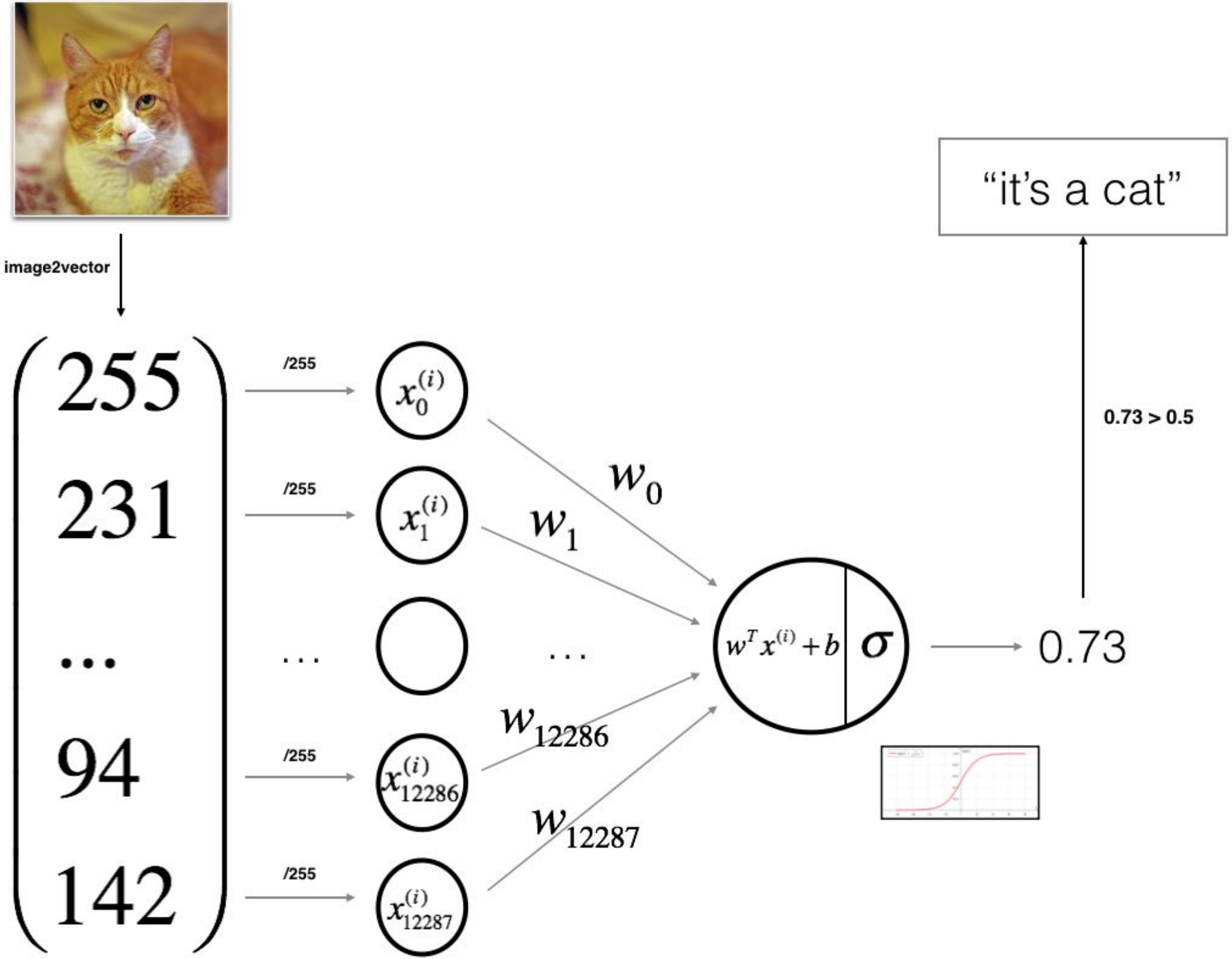
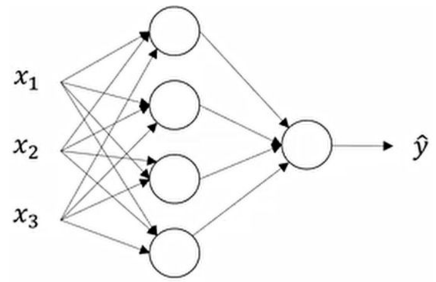


image2vector
255
231
...
94
142
/255
/255
/255
/255
$x_0^{(i)}$
$x_1^{(i)}$
$x_{12286}^{(i)}$
$x_{12287}^{(i)}$
w_0
w_1
w_{12286}
w_{12287}
$w^T x^{(i)} + b$
σ
0.73
0.73 > 0.5
"it's a cat"

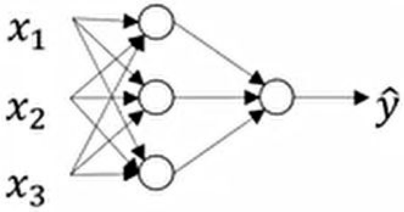
1. Week 3에서 배울 내용

<table>
<tr><td>모델</td><td>
</td><td rowspan="3">week 2 모델 연산에서

forward propagation
backward propagation

부분만 2배로 증가</td></tr>
<tr><td>computation graph</td><td>$x, W^{[1]}, b^{[1]} \rightarrow \boxed{z^{[1]} = W^{[1]}x + b^{[1]}} \rightarrow a^{[1]} = \sigma(z^{[1]}) \rightarrow z^{[2]} = W^{[2]}a^{[1]} + b^{[2]} \rightarrow a^{[2]} = \sigma(z^{[2]})$
$W^{[2]}, b^{[2]}$</td></tr>
<tr><td></td><td>$$np.dot(W^{[1]}, X) + b = \begin{bmatrix} w_{1\,1} & w_{1\,2} & w_{1\,3} \\ w_{2\,1} & w_{2\,2} & w_{2\,3} \\ w_{3\,1} & w_{3\,2} & w_{3\,3} \\ w_{4\,1} & w_{4\,2} & w_{4\,3} \end{bmatrix} \begin{bmatrix} x_1^{(1)} & x_1^{(2)} \\ x_2^{(1)} & x_2^{(2)} \\ x_3^{(1)} & x_3^{(2)} \end{bmatrix} + b$$</td></tr>
</table>

1. Week 3에서 배울 내용



$$np.dot(W^{[1]}, X) + b = \begin{bmatrix} -w_1^T- \\ -w_2^T- \\ -w_3^T- \\ -w_4^T- \end{bmatrix} \begin{bmatrix} x_1^{(1)} & x_1^{(2)} \\ x_2^{(1)} & x_2^{(2)} \\ x_3^{(1)} & x_3^{(2)} \end{bmatrix} + b$$

$$np.dot(W^{[1]}, X) + b = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{bmatrix} \begin{bmatrix} x_1^{(1)} & x_1^{(2)} \\ x_2^{(1)} & x_2^{(2)} \\ x_3^{(1)} & x_3^{(2)} \end{bmatrix} + b$$

2. Week 3에서 주로 다뤄질 내용

1) hidden layer의 등장: lec 2
2) hidden unit(node)가 여러 개인 상황: lec 3 ~ lec 5
3) activation function: lec 6 ~ 8
4) hidden layer가 있을 때의 5가지 step 진행: lec 9 ~ lec 10
5) initialization (unit이 여러 개인 상황): lec 11