

시즌 1 - 딥러닝의 기본 - Lecture 06

노트북: 모두를 위한 머신러닝
만든 날짜: 2018-12-31 오후 4:13
작성자: gpfl104@gmail.com
태그: #모두를 위한, .Lecture

수정한 날짜: 2018-12-31 오후 5:22

Lecture 6

= Logistic regression



x 입력 받음

w를 가지고 계산을 한 다음 나오는 값이 z 값이 됨

z값을 sigmoid 함수에 넣고 나오는 값은 0~1값을 가짐 (y-hat) 우리는 이게 H(x) 임

= Multinomial classification

$$\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [w_1 x_1 + w_2 x_2 + w_3 x_3]$$

$$\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = [w_1 x_1 + w_2 x_2 + w_3 x_2]$$

$$\begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = [w_1 x_1 + w_2 x_2 + w_3 x_3]$$

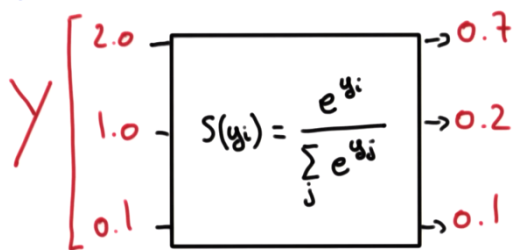
세 개 연산을 따로 계산하면 복잡하고 길어지니까, w를 합쳐보자.

$$\begin{bmatrix} w_{A1} & w_{A2} & w_{A3} \\ w_{B1} & w_{B2} & w_{B3} \\ w_{C1} & w_{C2} & w_{C3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} w_{A1}x_1 + w_{A2}x_2 + w_{A3}x_3 \\ w_{B1}x_1 + w_{B2}x_2 + w_{B3}x_3 \\ w_{C1}x_1 + w_{C2}x_2 + w_{C3}x_3 \end{bmatrix} = \begin{bmatrix} \bar{y}_A \\ \bar{y}_B \\ \bar{y}_C \end{bmatrix}$$

$H_A(x)$, $H_B(x)$, $H_C(x)$

y-hatA, y-hatB, y-hatC가 각각 $H_A(x)$, $H_B(x)$, $H_C(x)$ 가 된다.

= softmax



SCORES \longrightarrow PROBABILITIES

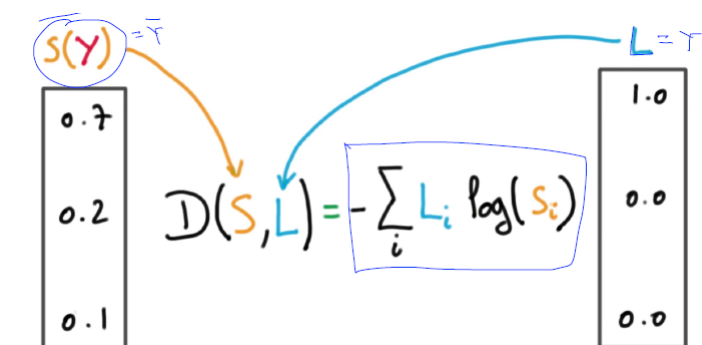
\hat{y} -hat 벡터 값을 모두 더하면 1이 되도록 하자. \Rightarrow softmax

이 값은 A, B, C가 나올 확률이라고 할 수 있다.

그 중 하나만 고르기 위해 가장 큰 값을 1.0으로 하고 나머지는 0.0으로 하는 것이 **ONE-HOT ENCODING**

argmax

= cross-entropy cost function



L 이 label, 정답, 실제 값

$S(y)$ 는 softmax에 넣어서 예측한 값

두 개의 차이가 얼마나 나는지 구하는 식

맞게 예측한 경우에는 0, 잘못 예측한 경우는 ∞ 가 나옴

= Logistic cost VS cross entropy

$$C(H(x), y) = y \log(H(x)) - (1 - y) \log(1 - H(x))$$

$$D(S, L) = -\sum_i L_i \log(S_i)$$

두 식은 왜 같을까?

참고 - <http://mazdah.tistory.com/791>

<https://www.youtube.com/watch?v=jMU9G5WEtBc&feature=youtu.be>