

Type something...

01) Backprop and Neural Networks

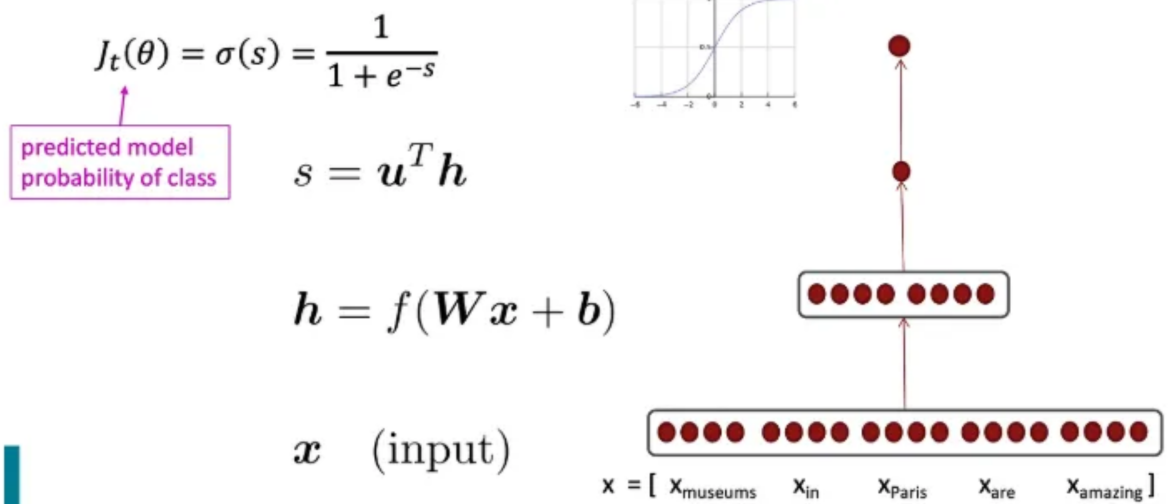
01_Named entity recognition(NER)

01) Backprop and Neural Networks

01_Named entity recognition(NER)

- context window의 이웃 단어들을 이용하여 각 단어들을 분류하는 방식이다.
- 예를 들어 "the museums in Paris are amazing to see" 라는 문장에서 "Paris"라는 단어에 대해 NER을 해본다고 가정해보자.
- 이때 window size가 2라면 다음과 같은 벡터가 만들어진다.

$$X_{\text{window}} = [x_{\text{museums}} \quad x_{\text{in}} \quad x_{\text{Paris}} \quad x_{\text{are}} \quad x_{\text{amazing}}]^T$$



- 각 단어에 대해 5차원 벡터를 만들고 이를 neural network layer에 넣어준다.
- Linear를 통과시킨 후 activation function을 이용하여 nonlinear 속성을 추가해준다.
- extra vector u 를 통해 낮은 차원의 벡터로 변환해주고 logistic 변환을 통해서 단어가 특정 클래스에 속할 확률을 구해준다. (sigmoid 사용)

$$\theta^{new} = \theta^{old} - \alpha \nabla_{\theta} J(\theta)$$

α = step size or learning rate

- SGD를 이용해서 backpropagation을 진행할 것이다.
 - 여기서는 행렬을 backpropagation하므로 자코비안 행렬을 이용해준다.
- backpropagation은 cs231n에서 다루었기 때문에 생략한다.