

CS224N

Lecture 3: Word Window Classification, Neural Networks, and Matrix Calculus

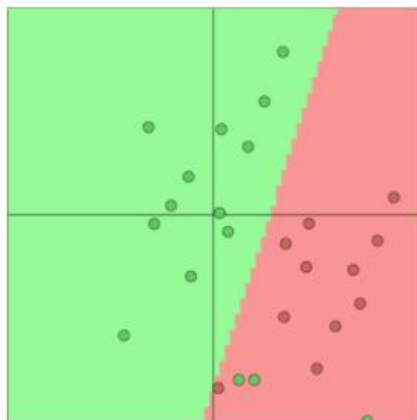
1. Classification

-

$$\{x_i, y_i\}_{i=1}^N$$

training dataset을 1부터 N까지 inputs과 outputs에 대해 training dataset을 가지고 있다. NLP에서 input은 단어나, 문장, 파일을 의미하며, output은 words나 다른 것일 수도 있다.

전통적인 Classification은 softmax나 logistic regression을 이용해서 class를 구분하는 linear boundary를 결정했다.



Visualizations with ConvNetJS by Karpathy!

<http://cs.stanford.edu/people/karpathy/convnetjs/demo/classify2d.html>

<Softmax classifier>

$$p(y|x) = \frac{\exp(f_y)}{\sum_{c=1}^C \exp(f_c)} = \text{softmax}(f_y)$$

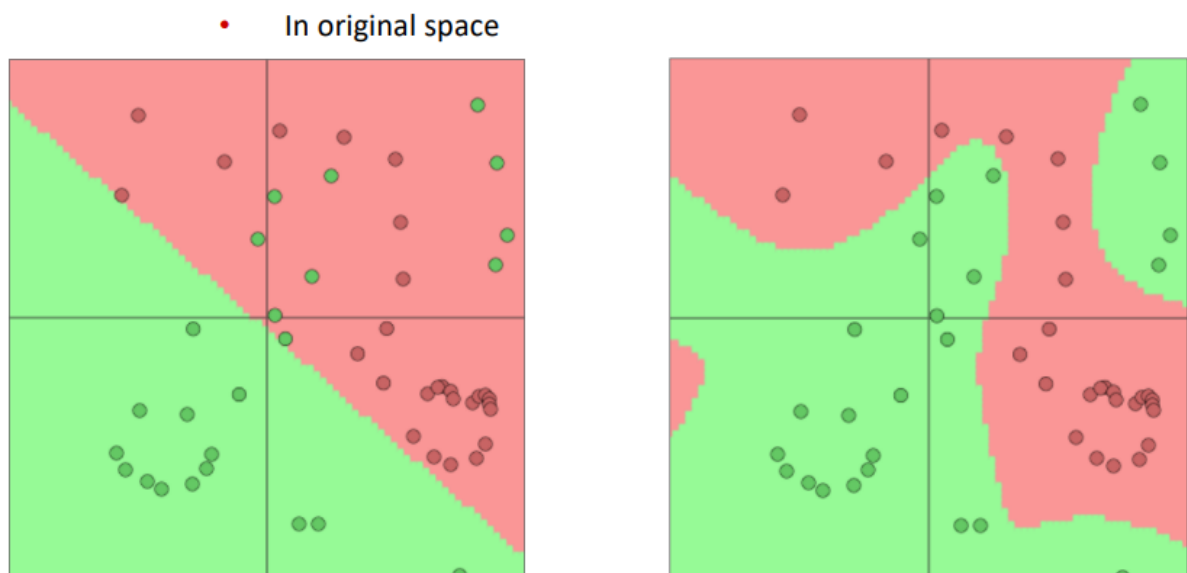
Training with softmax and cross-entropy loss

$$-\log p(y|x) = -\log \left(\frac{\exp(f_y)}{\sum_{c=1}^C \exp(f_c)} \right)$$

값을 학습할 때, 올바른 y 값을 예측하도록 확률을 극대화하거나, negative log값을 최소화하도록 학습을 하게 된다.

2. Neural networks

- decision boundary



non-linear boundary(오른쪽 그림)를 설정함으로써 분류의 정확도를 더 높일 수 있다.

