

시즌 1 - 딥러닝의 기본 - Lecture 07

노트북: 모두를 위한 머신러닝
만든 날짜: 2019-01-02 오후 3:24
작성자: gpfl104@gmail.com
태그: #모두를 위한, .Lecture

수정한 날짜: 2019-01-02 오후 4:27

Lecture 7

= learning rate

- Large learning rate: overshooting

learning rate를 너무 크게 잡아서 코스트가 줄어들지 않고 큰 값이 찍히게 되고 학습이 안 됨

- Small learning rate: takes too long, stops at local minimum

너무 조금씩 줄어들어서 시간이 오래 걸림

보통은 0.01 로 시작

= Data preprocessing

- data (X) preprocessing for gradient descent

데이터 값에 큰 차이가 있을 경우, 노말라이즈 해야 함

zero-centered data: 데이터의 중심이 0으로 갈 수 있도록 바꿔줌

normalized data: 어떤 범위 안에 항상 들어가도록 바꿔줌

- Standardization

```
X_std[:,0] = (X[:,0] - X[:,0].mean()) / X[:,0].std()
```

= Overfitting

너무 가지고 있는 데이터에만 정확한 모델

실제 환경에서 적용하기 어려움

오버피팅 줄이는 방법

1. 트레이닝 셋 데이터 늘리기
2. 티치 개수 줄이기 중복 줄이기
3. Regularization (일반화 시키기)

- Regularization

모델을 구부리지 말고 펴자!

W을 제공하여 합한 값을 최소화

앞에 상수를 줄 수도 있다. -> regularization strength

0: 사용하지 않겠다

1: 굉장히 중요하게 생각한다

0.001: 사용하긴 하지만 그렇게 중요하게 생각하진 않는다

```
l2reg = 0.001 * tf.reduce_sum(tf.square(W))
```

= Training / Testing 데이터 셋

training set 으로 학습시키고 test set으로 얼마나 학습되었는지 확인

training set에서도 training과 validation으로 나눌 수 있음

training으로 학습시키고 validation으로 튜닝함 (모의 시험)

완벽하게 되면 testing set으로 확인 (실전)

- Online learning

100만 개를 잘라서 10만 개씩 학습

데이터를 추가적으로 학습할 수 있기 때문에 좋은 모델

- Accuracy (정확성)