

모델 정규화

- 기존의 모델을 보면 차원수가 많아 'overfitting' 이 발생했다.
- 정확화 향을 붙이는 방법으로 해결가능하다.
- 대표적으로 Ridge, Lasso, Elastic Net 존재

Ridge

- 각 계수의 제곱을 더한 값을 식에 포함하여 계수의 크도 함께 최소화하도록 만듦
- α 값이 크면 계수를 많이 줄이고, α 값이 적으면 계수를 조금 줄인다.

underfitting

overfitting

- 릿지가 기본으로 사용하고 실제로 사용하는 feature 가 있다면 라쏘와 엘라스틱 사용 (여유는 아래에서)
- Ridge의 가중치는 '0' 이 되지는 않는다.

- (식) \rightarrow (모든 계수) 합을 의미

$$J(\theta) = \text{MSE}(\theta) + \alpha \frac{1}{2} \sum_{i=1}^N \theta_i^2$$

Lasso

- Ridge와와는 조금 다르게 제곱대신 절대값 사용한다.
- Lasso는 Ridge와 다르게 가중치가 '0' 이 될수 있다.

- 덜 중요한 features의 가중치를 '0' 으로 만들어 변수선택 효과가 있다. (엄청 중요한 feature 가 훈련데이터보다 많으면 문제발생해서 엘라사용)

- (식)

$$J(\theta) = \text{MSE}(\theta) + \alpha \sum_{i=1}^N |\theta_i|$$

Elastic-Net

- 릿지, 라쏘를 절충한 모델이다.
- 두 모델을 혼합 비율을 사용해 조절 한다.
- 극단적으로 변수의 수가 훈련샘플의 수 보다도 많고

변수 몇개가 강하게 연관되어 있을 경우에는 라쏘보다 엘라스틱넷 선호

- (식)

$$J(\theta) = \text{MSE}(\theta) + \gamma \cdot \alpha \sum_{i=1}^N |\theta_i| + \frac{1-\gamma}{2} \alpha \sum_{i=1}^N \theta_i^2$$