

모델

머신러닝은 간단히 '데이터를 이용한 모델링 기법'이라고 표현할 수 있습니다. 다시 말하면, 데이터들 중에서 '모델'을 뽑아내는 기법입니다. 모델은 머신러닝을 통해 얻을 수 있는 최종 결과물을 의미합니다. 예를 들어, 스팸을 분리해내는 시스템을 만들었다면 해당 시스템이 '모델'에 해당합니다.

모델은 '가설(Hypothesis)'이라고도 부릅니다.

- **오버피팅**

- 오버피팅(과대적합, Overfitting) : 머신러닝에서 과대적합(Overfitting)은 큰 문제 중에 하나이다. 과대적합이란 머신러닝(여기서는 신경망) 모델이 너무 과한 학습으로 인해 훈련 데이터에만 지나치게 적응되어 그 이외의 시험 데이터에 대해서는 제대로 대응하지 못하는 것을 말한다.

- 머신러닝에서는 다양한 상태에 대하여 일반적인 성능을 내는 모델을 지향하므로, 훈련 데이터에 없는 상황이 주어지더라도 올바른 결과를 내는 모델이 더 좋은 모델이다. 따라서 복잡하고 표현력이 좋은 모델을 만들면서도, 과대적합을 억제하는 기술이 중요해지는 것이다.