Overfitting and Regularization

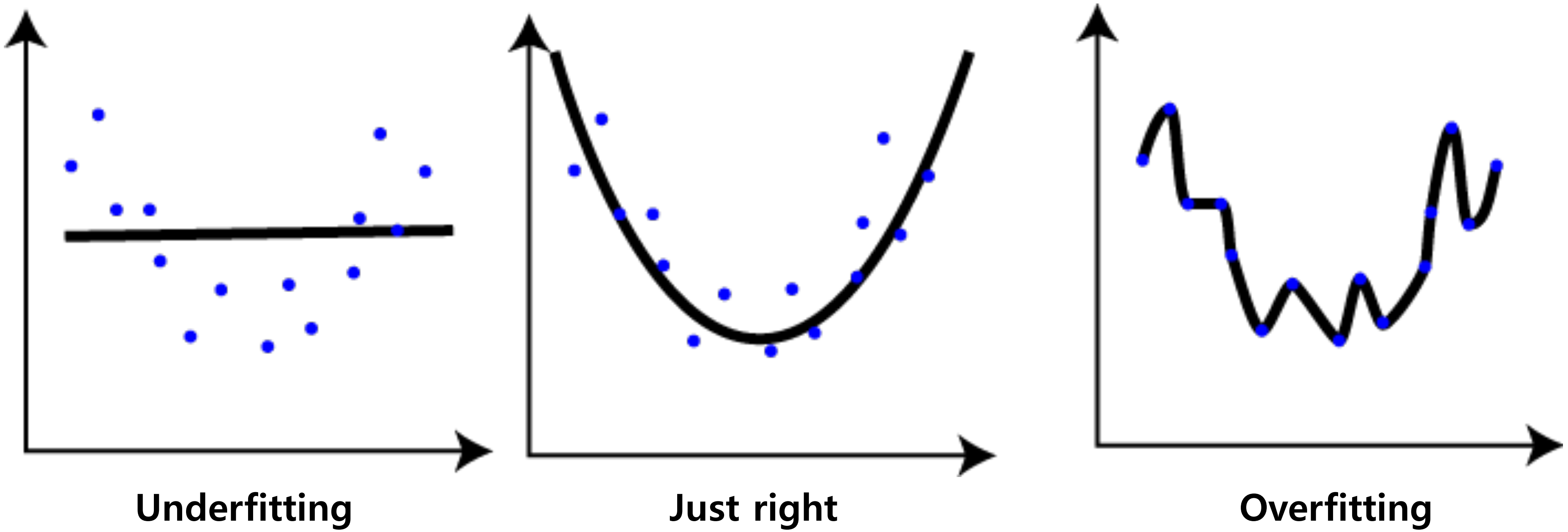
Linear Regression

Director of TEAMLAB
Sungchul Choi

Overfitting

Overfitting

학습데이터 과다 최적화 → 새로운 데이터의 예측 ↓



https://goo.gl/aP8iFa

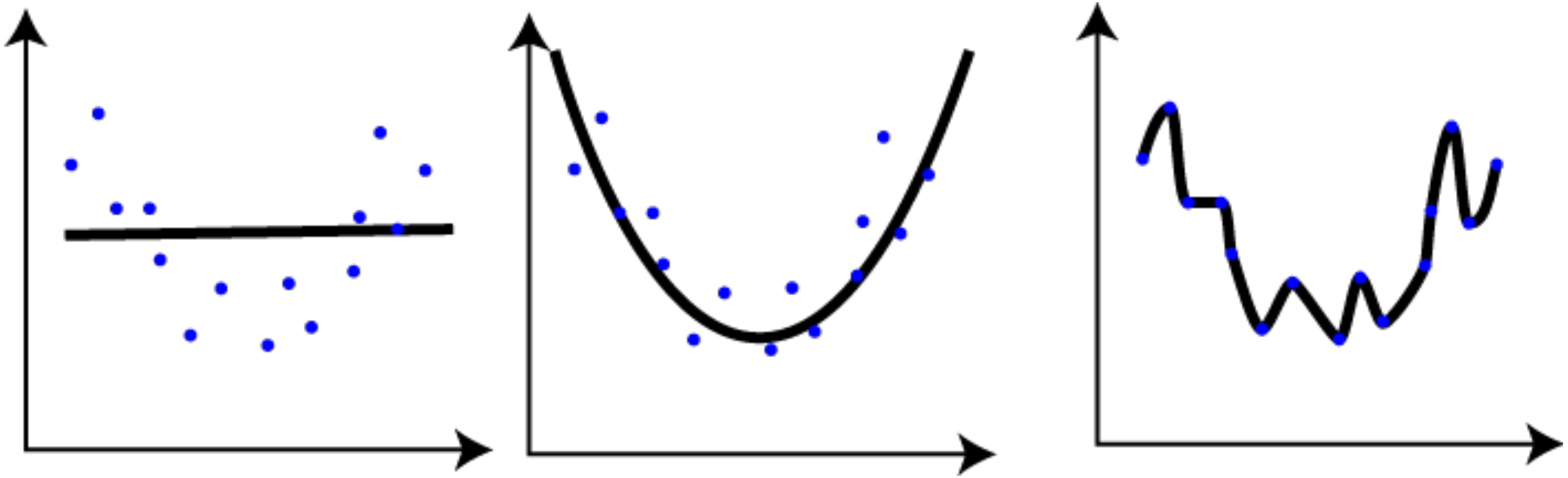
보다 적은 수의 논리로 설명이 가능한 경우, ₩
많은 수의 논리를 세우지 말라

Occam’s razor

https://ko.wikipedia.org/wiki/%EC%98%A4%EC%BB%B4%EC%9D%98_%EB%A9%B4%EB%8F%84%EB%82%A0

Bias – Variance tradeoff

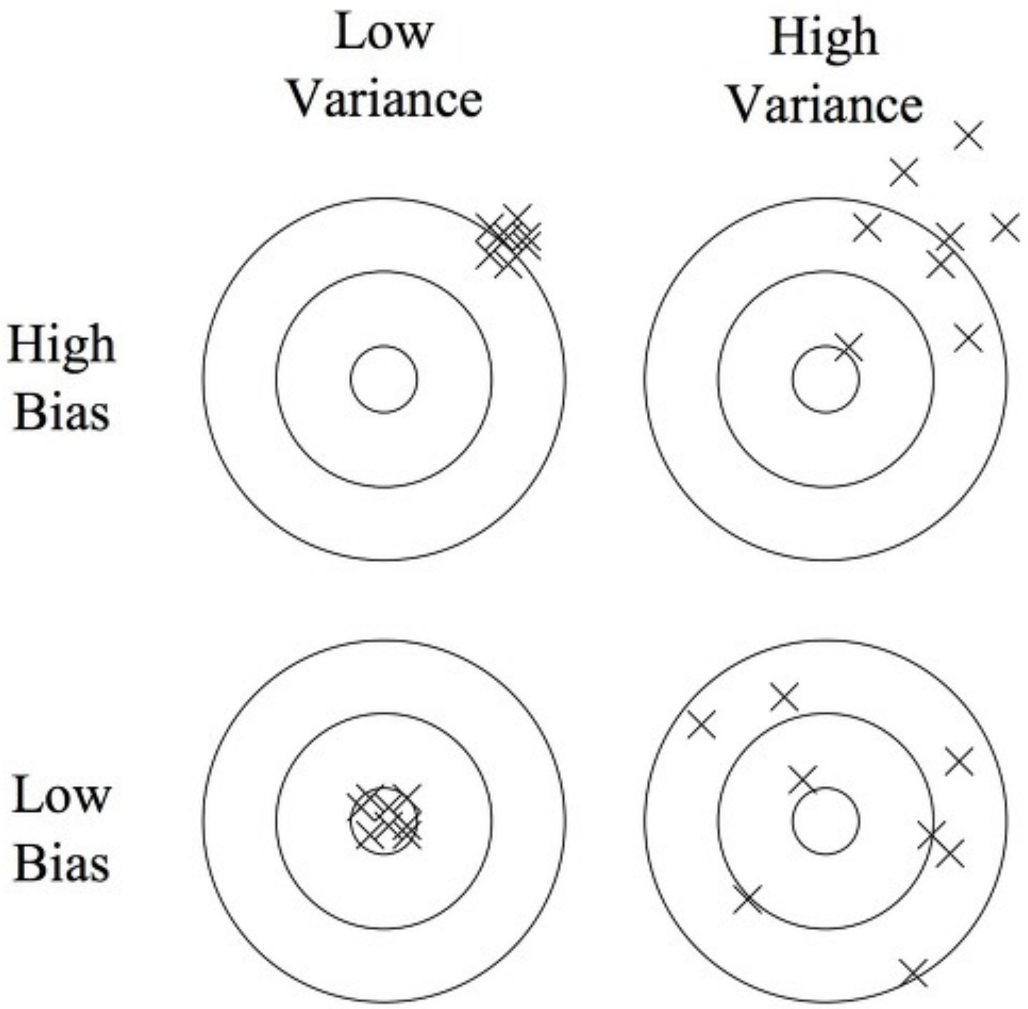
학습데이터 과다 최적화 → 새로운 데이터의 예측 ↓



High bias　　Just right　　High variance

https://goo.gl/aP8iFa

Bias – Variance tradeoff



High bias

원래 모델에 많이 떨어짐

잘못된 데이터만 계속 학습함

→ 잘못된 Weight만 Update

High variance

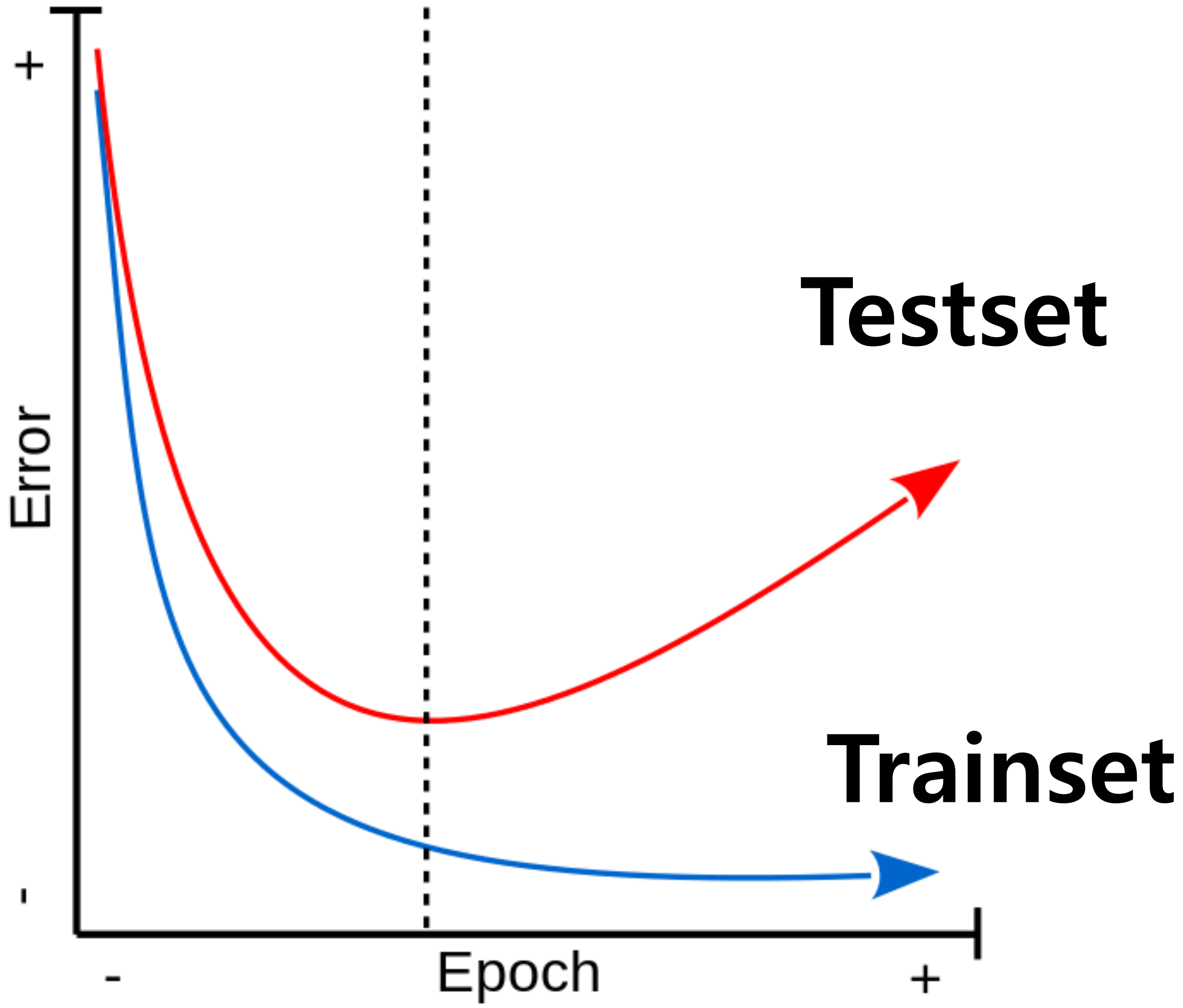
모든 데이터에 민감하게 학습

Error를 고려하지 않음

→ 모든 Weight가 Update

http://scott.fortmann-roe.com/docs/BiasVariance.html

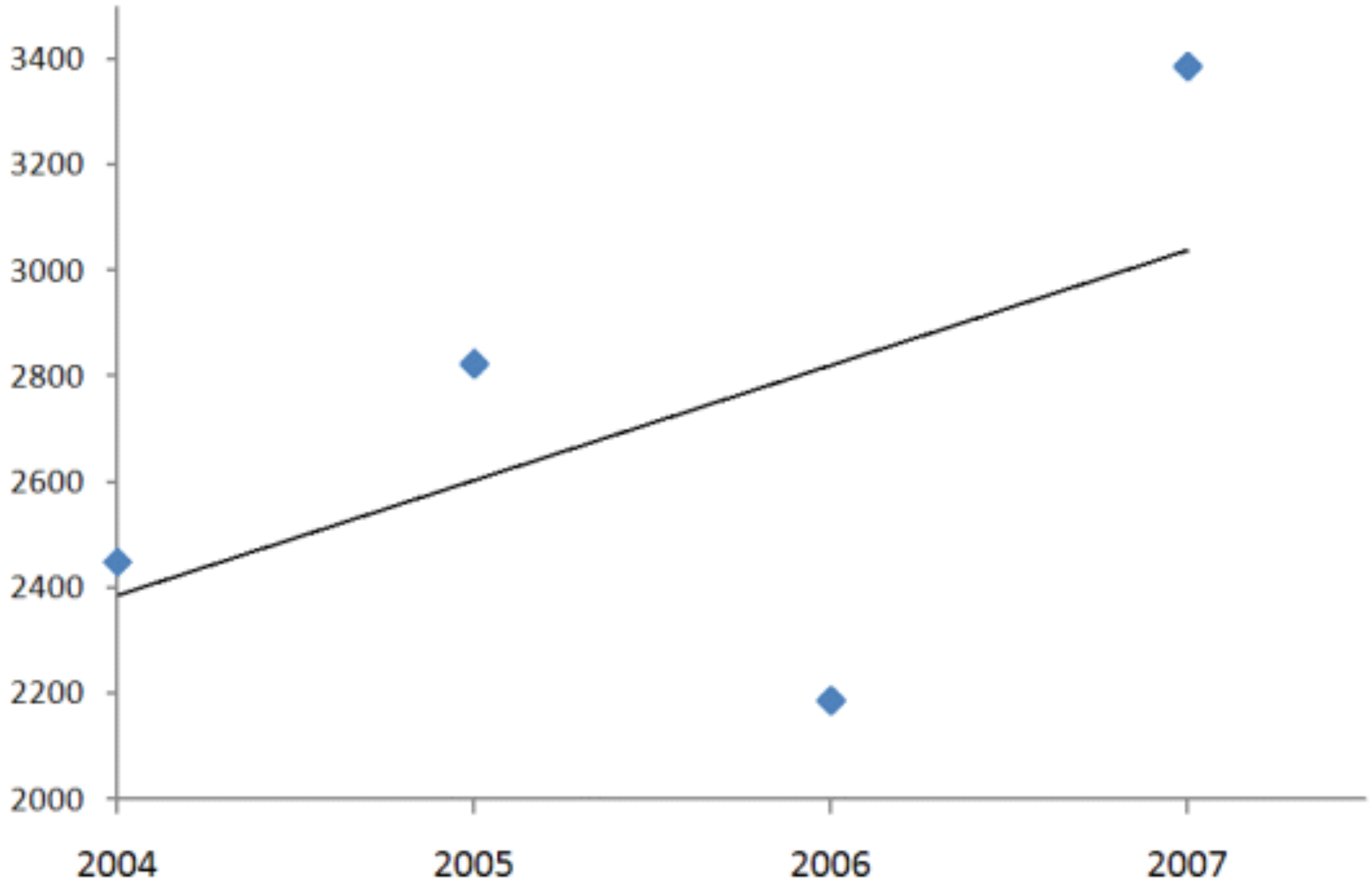
Train-Test Error



Overcoming Overfitting

- 더 많은 데이터를 활용한다.
- Feature의 개수를 줄인다
- 적절히 Parameter를 선정한다
- Regularization

Regularization



Regularization

$$J(w_0, w_1) = \frac{1}{2m} \sum_{i=1}^{m} (w_1 x^{(i)} + w_0 - y^{(i)})^2$$

$$\frac{\partial J}{\partial w_0} = \frac{1}{m} \sum_{i=1}^{m} (w_1 x^{(i)} + w_0 - y^{(i)})$$

$$\frac{\partial J}{\partial w_1} = \frac{1}{m} \sum_{i=1}^{m} (w_1 x^{(i)} + w_0 - y^{(i)}) x^{(i)}$$

TEAMLAB

Human knowledge belongs to the world.