

• 3/4 Regression (회귀) : 연속형의 규모값 예측

Classification (분류) : 0 or 1 예측 + 성능평가 (Metric)

• confusion Matrix

		실제값 (Label)	
		True	False
예측값	True	True positive	False positive
	False	False Negative	True Negative

• Precision (정밀도)

(우리 Model이 True인
보유한 것 중真正 True인 것의 비율)

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

• Recall (재현율)

hit rate
실제 True인 것 중에서
우리의 Model이 True라고 예측한
비율

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

• Accuracy (정확도)

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}}$$



"bias"

(사건CT)

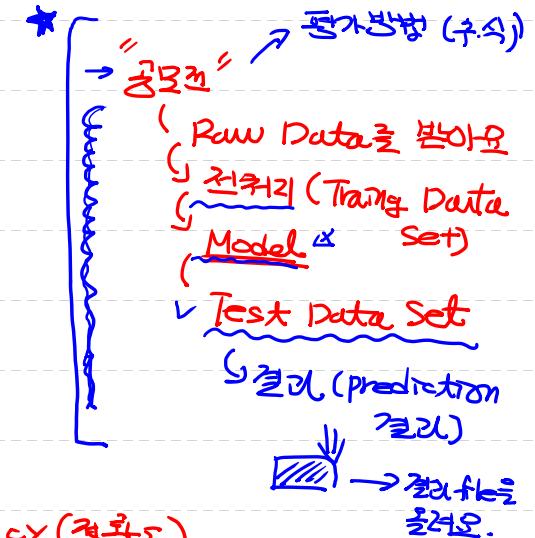
우리 Model이 확률값을 줌.

$$\frac{1}{1000000} 확률$$

Model → 무조건 확률값이

"아니다"

→ 99.9%



• Precision, Recall, Accuracy

해결하려는 문제(Data)에 따라

종종 이것을 사용하기가 결정

일반적으로
Precision Vs Recall
비례

Precision 를

Recall의 조화평균

$$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

\Rightarrow F1 Score

• Fall-out

$$\text{Fall-out} = \frac{FP}{TN+FP}$$

실제값은 False인가
우리 Model이 True를
잘못 예측한 비율.

• Recall & Fall-out \Rightarrow 그래프로 표현

ROC curve

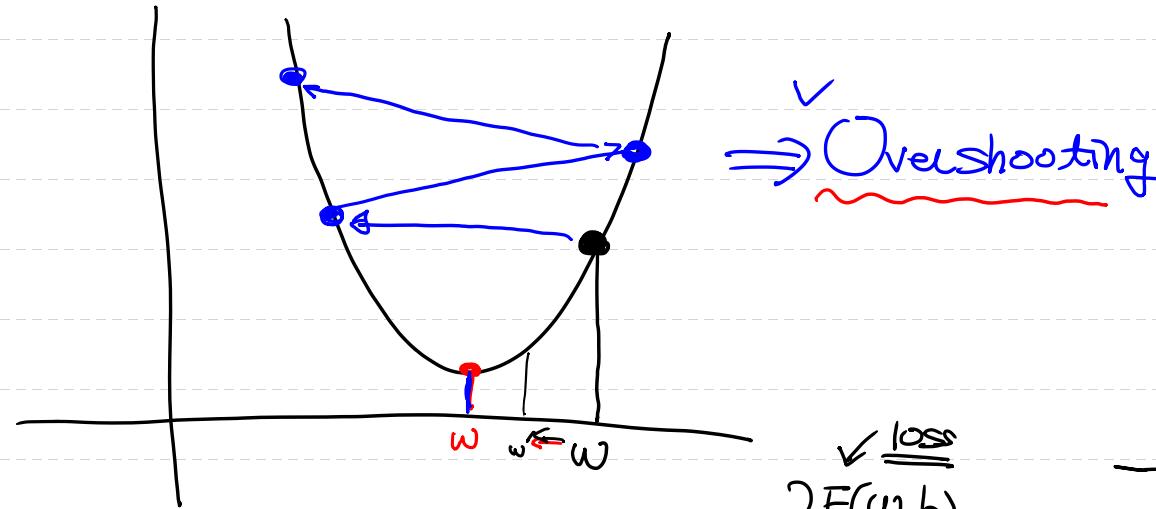
F

주제적으로 알아두어야 할 내용!!

(Learning Rate를 틀는) 이전 주제의 본문이 끝나요 ?

- Learning Rate \rightarrow w 와 b 를 갱신할 때
사용되는 Customizable 수치, ($1e^{-4}$)

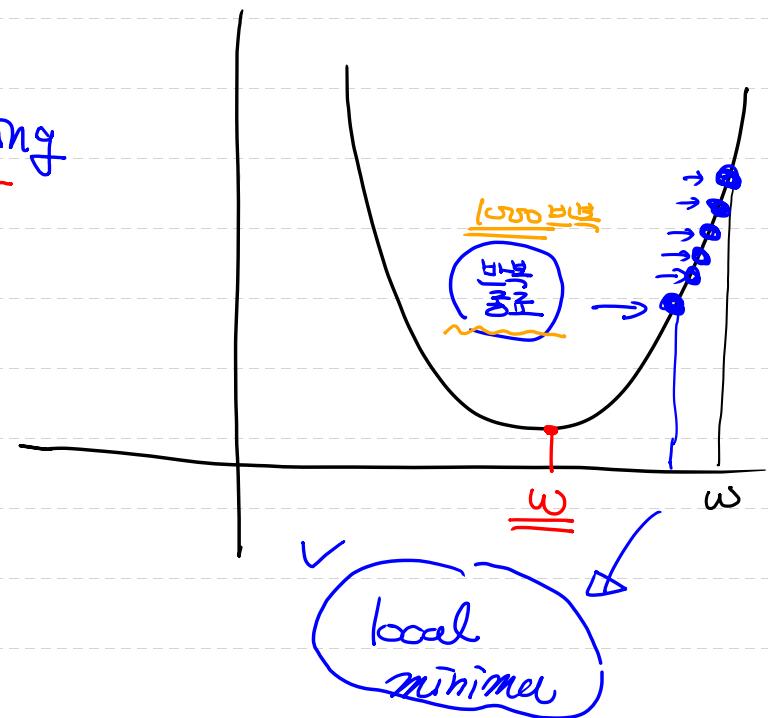
▼ 만약 Learning Rate를 훨씬 이상으로 크게 설정



$$w' = w - \alpha \cdot \frac{\partial E(w, b)}{\partial w}$$

↑
Learning rate ↑

▼ 만약 Learning Rate를 훨씬 이상으로
작게 설정



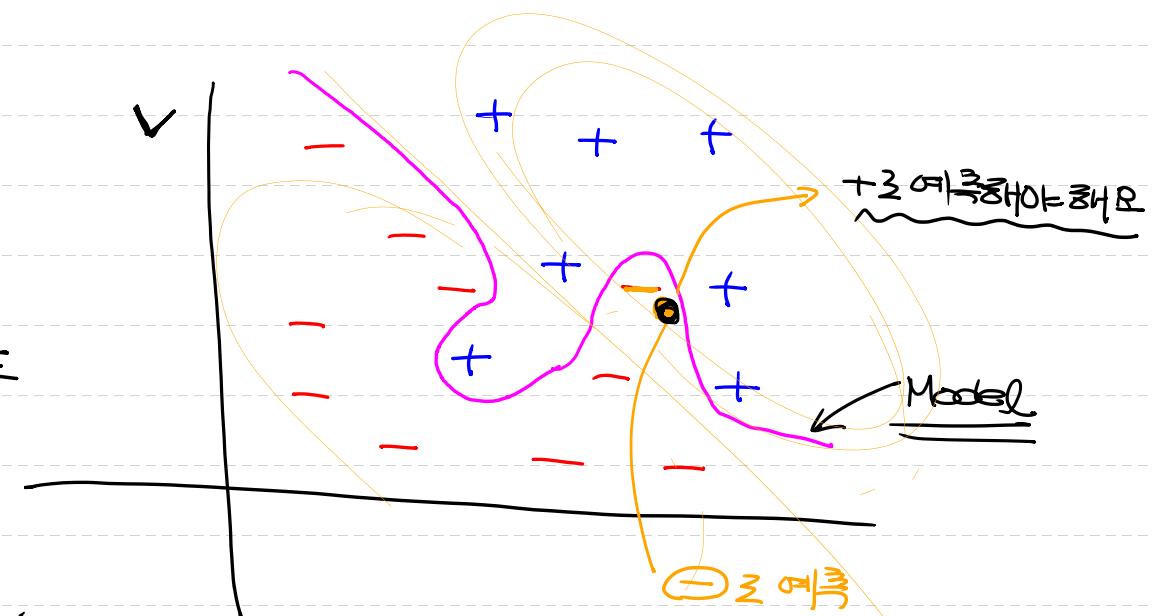
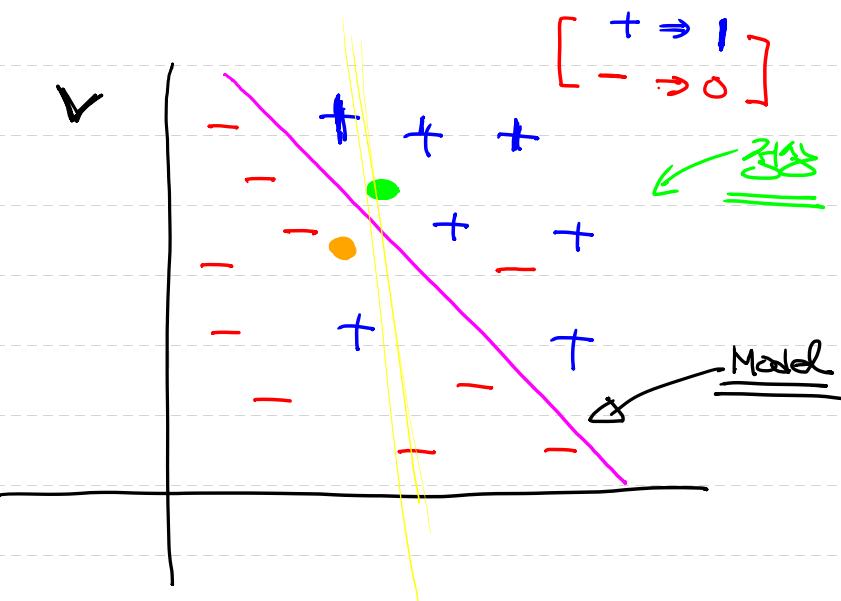
● Overfitting (과대적합) 과적합 ✅ Underfitting (과소적합) ✅

총부분 학습률 하지 않아서

Model 가족이 비쁘다

✗

Data를 이용해서 학습이 끝난 후 우리 model의 성과를 기록하는 용어!!



Overfitting

epoch (주기)

• Data가 많아 (많은 Data들이 필요) → 가짜 data는 필요

없는 경우에는 어떤가요??

Feature의 개수를 줄여요!! (중복되는 feature는 제외, 필요 없는 feature는 제거)

우리 Model이 Training Data Set에

너무 잘 맞아서 Training Data Set을 딱 들어맞는
형태로 만들어 줘야요!!

Evaluation (성능평가)

Training Data Set으로 성능평가를

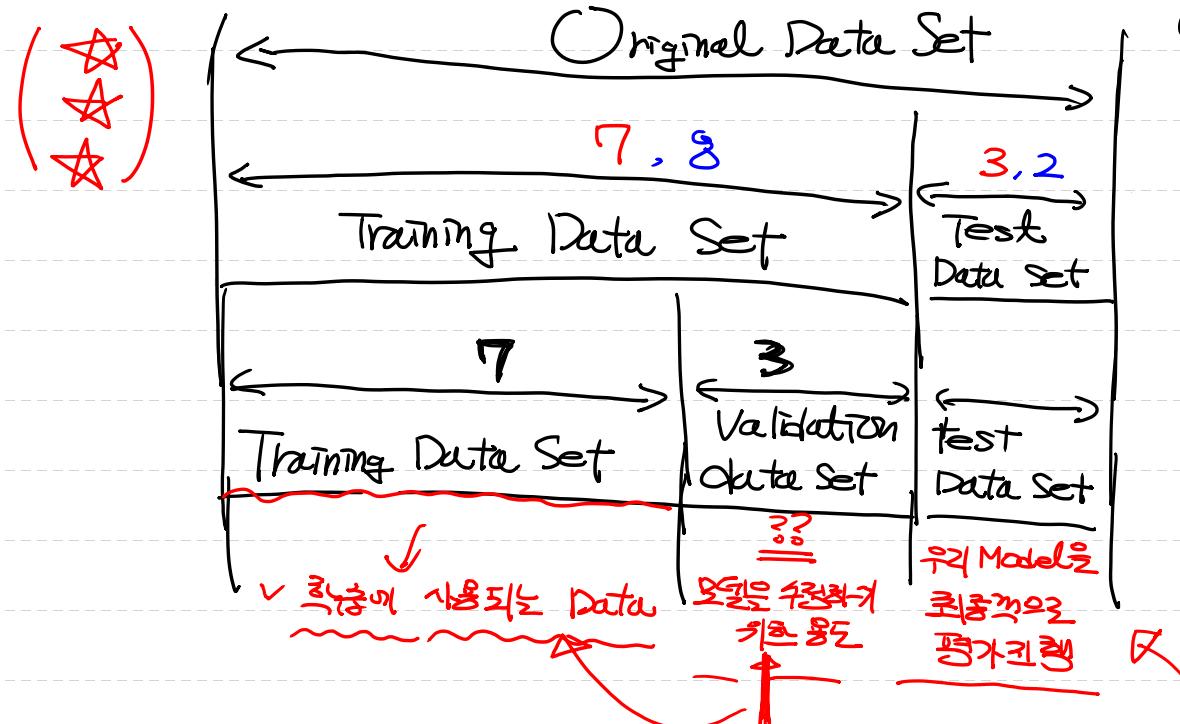
기준과면 X

거의 100%의 accuracy를 도출함

CV k-fold Cross Validation

총 데이터의 구가 각을 걸

전체 Data의 양이 많아야 해요



Epoch 당 Model이 갱신.

Training Data Set은 평가

Validation Data Set은 평가

최적 epoch

- 드라이버 조작자이 Logistic Regression의 기본 예제

⇒ Titanic 예제

Data는 kaggle에서 블아서 사용

[Training Data Set]
[Test Data Set]

"Tensorflow로
구현해 보아요 !! "

"Titanic"

