

4/4/2019

모델 평가와 성능 향상

장서영

CONTENTS

1. 회귀모델의 평가지표
2. 분류모델의 평가지표
3. 교차검증
4. 그리드 서치를 통한
파라미터 튜닝

모델 평가의 목적

- 모델이 미래의 (처음 본) 데이터에 대한 예측 성능인 **일반화** 정확도를 추정
- 학습 알고리즘을 튜닝하고 주어진 가설 공간 안에서 가장 성능 좋은 모델을 골라 **예측 성능** 향상
- 최대 성능을 내는 알고리즘을 선택하기 위해 여러 알고리즘들을 **비교**

모델 평가 지표 - Regression

- MSE

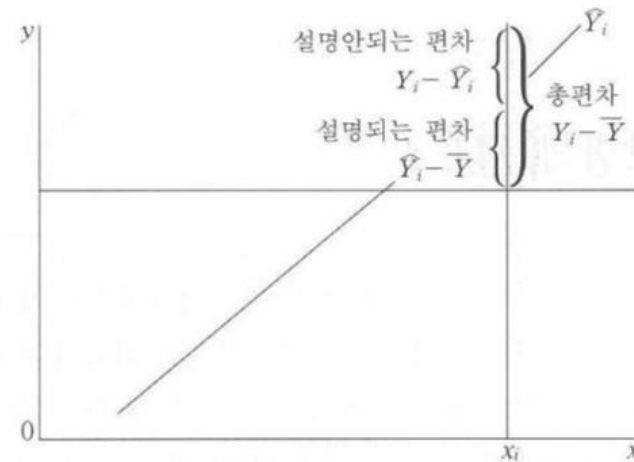
$$MSE = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}$$

- RMSE

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

- R^2

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$



$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})}{\sum(y_i - \bar{y})^2} = \frac{\text{회귀선에 의해 설명되는 변동}}{\text{전체 변동}}$$

$$\text{adjusted } R^2 = 1 - \frac{n - 1}{(n - p - 1)(1 - R^2)}$$

모델 평가 지표 - Classification

Confusion Matrix

		Predicted Data		Total
		Predicted Condition POSITIVE	Predicted Condition NEGATIVE	
Actual Data	Condition TRUE	TP True Positive	FN False Negative	P
	Condition FALSE	FP False Positive	TN True Negative	N
Total		P [^]	N [^]	P+N

- True Positive (TP): 양성을 양성으로 제대로 검출
- True Negative (TN): 음성을 음성으로 제대로 검출
- False Negative (FN): 양성을 음성으로 잘못 검출
- False Positive (FP): 음성을 양성으로 잘못 검출

모델 평가 지표

1. 정확도 Accuracy

실제값과 예측값이 일치한 비율

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

2. 정밀도 Precision

양성으로 예측한 값 중 실제로 양성인 값인 비율

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

		Predicted Data		Total
		Predicted Condition POSITIVE	Predicted Condition NEGATIVE	
Actual Data	Condition TRUE	TP True Positive	FN False Negative	P
	Condition FALSE	FP False Positive	TN True Negative	N
Total		P`	N`	P+N

모델 평가 지표

3. 재현율 Recall = 민감도 Sensitivity

실제 양성인 값 중 양성으로 정확히 예측한 값의 비율

$$\text{sensitivity} = \frac{TP}{P} = \frac{TP}{TP + FN}$$

$$\text{specificity} = \frac{TN}{N} = \frac{TN}{FP + TN}$$

• 특이도:

실제 음성인 값 중 음성으로 정확히 예측한 값의 비율

4. F1점수

정밀도와 재현율의 조화평균
(정밀도 재현율 반비례 관계)

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

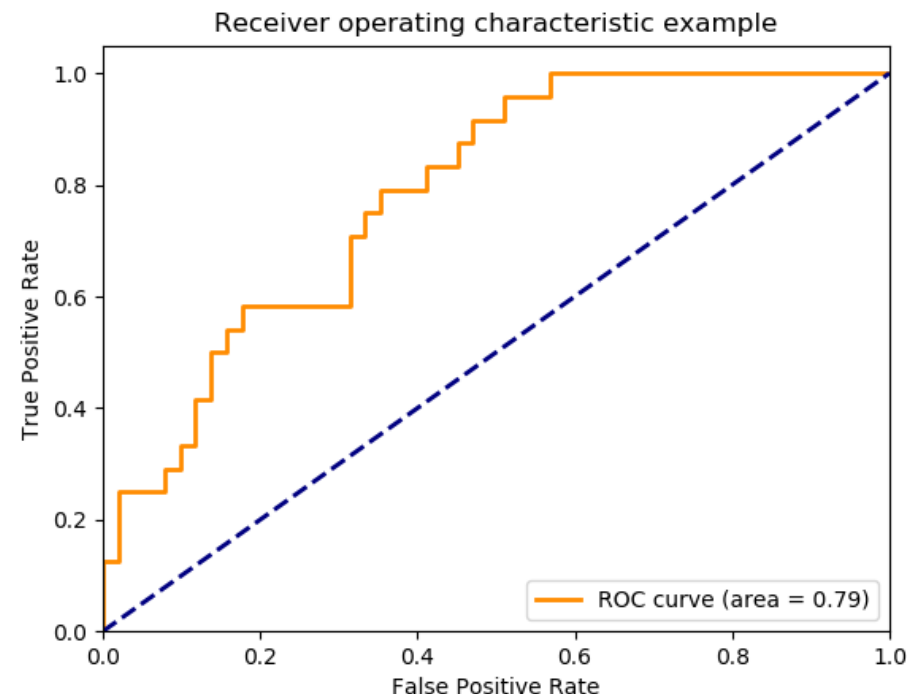
$$F_{\beta} = \frac{(1 + \beta^2) \times \text{precision} \times \text{recall}}{\beta^2 \times \text{precision} + \text{recall}}$$

• F_{β} 점수:

정밀도의 β 배만큼 재현율에 가중치

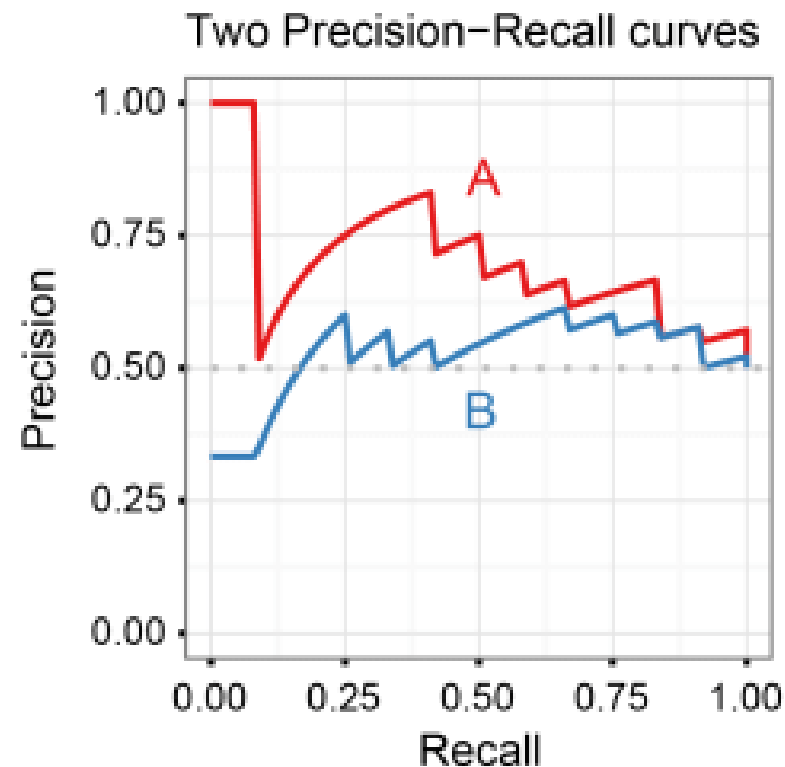
ROC (Receiver Operating Characteristics)

- ROC 곡선: Specificity를 x축, Sensitivity를 y축으로 나타낸 그래프
- 그래프가 왼쪽 위 코너로 붙을수록 좋은 모델, $y=x$ 보다 위에 있어야 확률이 0.5를 넘겨서 쓸모가 있는 모델
- AUC(Area Under Curve)값으로 평가
 - ROC그래프의 면적
 - 1에 가까울 수록 좋음



PR (Precision-Recall)

- PR 곡선: Recall을 x축, Precision을 y축으로 나타낸 그래프
- 데이터 라벨의 분포가 불균등할 때 사용
 - ex) 정상거래 98%, 비정상거래 2%
- 그래프가 오른쪽 위 코너로 붙을수록 좋은 모델,
Base line = $P/(P+N)$ 보다 위에 있어야 쓸모 있는 모델 (N: 전체데이터)
- AUC값으로 모델의 정확도 평가



교차검증 (Cross-Validation)

통상적 정확도 검정 방식

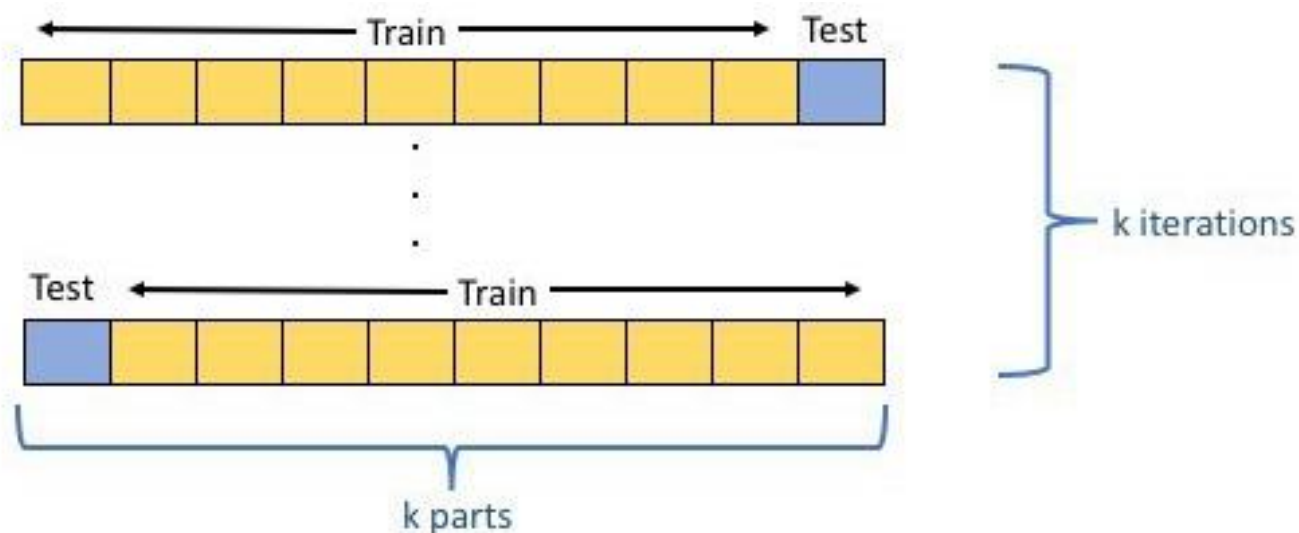
Train set과 test set을 나누어서 train set에 기반해 모델을 만들고 test set에 대한 예측력 및 정확도를 도출

교차검증

- 데이터를 어떻게 분리하느냐에 따라 검증 성능이 조금씩 달라질 수 있으므로 여러가지 방식으로 데이터를 분리하여 검증을 실시하는 방식
- **평균성능(mean performance)과 성능분산(performance variance)**를 모두 구함
- 데이터의 수가 적은 경우 검증 성능의 신뢰도가 떨어지는 문제 해결

K-fold Cross Validation

1. 데이터를 K개의 부분 집합(폴드)로 분리
2. 첫 번째 폴드를 test set으로 두고 나머지 K-1개의 폴드를 training set으로 이용하여 K개의 모델을 추정
3. 테스트 폴드를 바꾸어 가며 모든 폴드가 사용될 때까지 반복



그리드 서치를 통한 파라미터 튜닝

- 매개변수를 튜닝하여 일반화 성능을 높이는 것이 목표

- **Grid Search:**

관심있는 매개변수들을 대상으로 가능한 모든 조합을 시도

- `sklearn.model_selection.GridSearchCV()`

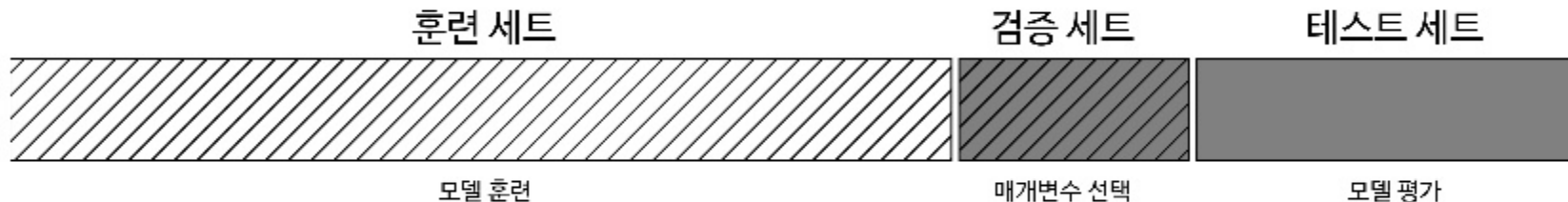
어떤 파라미터를 실험하고 싶은지, 어떤 값들을 시험해 보아야 하는지를 알려주면 교차검증을 사용해 파라미터 값들의 모든 가능한 조합들을 평가

Validation Set

여러가지 파라미터 조합에 대해 평가

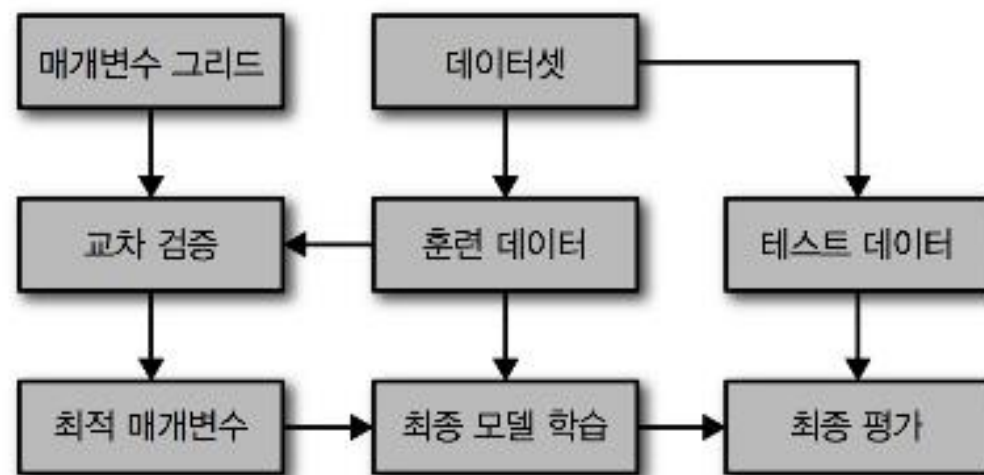
-> 최종 평가를 위해서는 독립된 데이터 세트가 필요!

학습용 데이터를 학습용과 검증용으로 분리해 학습용 데이터로 학습한 후
검증용 데이터로 검증



모델 평가 과정

매개변수 탐색의 전체 과정



실습 코드

Session09 model evaluation.ipynb로 실습해 봅시다 😊

Quest

Sklearn의 breast cancer 데이터의 mean radius, mean texture 변수를 사용해 종양을 분류하는 모델에 대한 평가와 성능 향상

1. 두 개 이상의 분류 모델의 AUC를 비교해 가장 좋은 모델 한 가지를 선택해 주세요. (ROC, PR 곡선 중 데이터에 더 적합한 기준 선택)
2. 선택된 모델에 대해 5겹 교차검증을 통한 그리드 서치를 시행해 최적의 파라미터를 도출해 주세요.
(1번에서 선택한 모델보다 좋은 성능을 내는 파라미터 찾기)
3. 최종적으로 최적의 파라미터로 튜닝된 모델로 score를 산출해 주세요!

참고자료

- <https://bcho.tistory.com/1206>
- <https://codedragon.tistory.com/6943>
- <https://datascienceschool.net/view-notebook/ff4b5d491cc34f94aea04baca86fbef8/>
- <https://www.slideshare.net/RickyPark3/5model-evaluation-and-improvement>
- <https://datascienceschool.net/view-notebook/266d699d748847b3a3aa7b9805b846ae/>
- <https://cyan91.tistory.com/18?category=230402>
- 김현세 (GH 3기), 교차검증 세션 자료