

1. 회귀 모형의 평가지표

SSE 오차 제곱합

SST 전체 제곱합

SSR 회귀 제곱합

AE 평균 오차

MAE 평균 절대 오차

RMSE 평균 제곱근 오차

MAPE 평균 절대 백분율 오차

MPE 평균 백분율 오차

2. ROC 곡선

가로축을 거짓 긍정률, 세로축을 민감도로 두어 시각화한 그래프

AUC의 판단 기준

AUC는 진단 정확도 측정할 때 사용하는 것으로, 곡선 아래 면적을
평가지표로 삼음

0.9~1.0 Excellent

0.8~0.9 Good

0.7~0.8 Fair

0.6~0.7 Poor

0.5~0.6 Fail

3. 모수 유의성 검정

모집단 평균에 대한 유의성 검정

Z-검정

귀무가설에서 검정 통계량의 분포를 정규분포로 근사할 수 있는 통계 검정

T-검정

검정하는 통계량이 귀무가설 하에서 T-분포를 따르는 통계적 가설검정

분산 분석 (ANOVA)

2개 이상의 집단 간 비교를 수행할 때 집단 내 분산, 총 평균과 각 집단의 평균 차이에 의해 생긴 집단 분산 비교로 얻은 F-분포를 이용하여 가설검정을 수행하는 방법

모집단 분산에 대한 유의성 검정

카이제곱 검정

관찰된 빈도가 유의미하게 다른지를 검정하기 위해 사용되며, 카이제곱 분포에 기초한 통계적 검정 방법

F-검정

두 표본의 분산에 대한 차이가 통계적으로 유의미한가를 판별하는 검정 기법

4. 과대 적합

제한된 학습 데이터 세트에 너무 특화되어 새로운 데이터에 대한 오차가

매우 커지는 현상

과대 전파 방지

데이터 증강

데이터의 양이 적을 경우, 데이터를 변형해서 늘릴 수 있음

모델의 복잡도 감소

인공신경망의 은닉층의 수를 감소시키거나 모델의 수용력을 낮추어 복잡도를 줄임

가중치 규제 적용

개별 가중치 값을 제한하여 모델을 간단하게 하는 방법

드롭 아웃

학습 과정에서 신경망 일부를 사용하지 않는 방법

