

Chapter 22

배운걸 적용해 보자: 실전 모델 개발 프로세스

|실전 그리드 서치

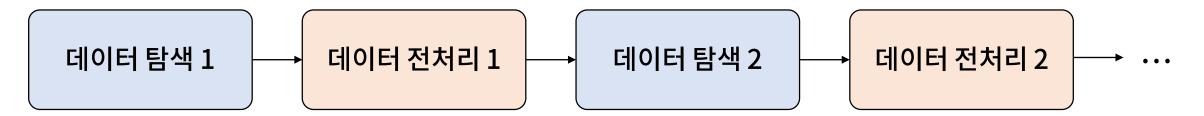
FAST CAMPUS ONLINE 데이터 탐색과 전처리 I

강사. 안길승

l큰 그림을 보자

FAST CAMPUS

실제 프로세스에서는 데이터 탐색과 전처리 사이에 피드백 루프가 존재함



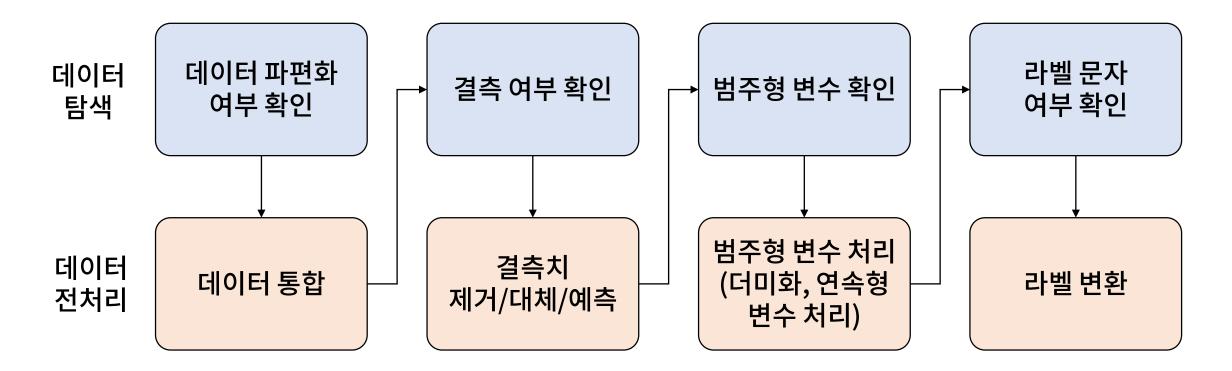
- 즉, 탐색을 한 번에 다하고 전처리를 한 번에 다하는 것이 아님
- 전처리에도 다양한 하이퍼 파라미터가 있기에 모델의 하이퍼 파라미터와 같이 튜닝해야 함

ONLINE 안길승 강사.



Ⅰ현실 지도학습 모델 개발 프로세스 (1) 필수 전처리

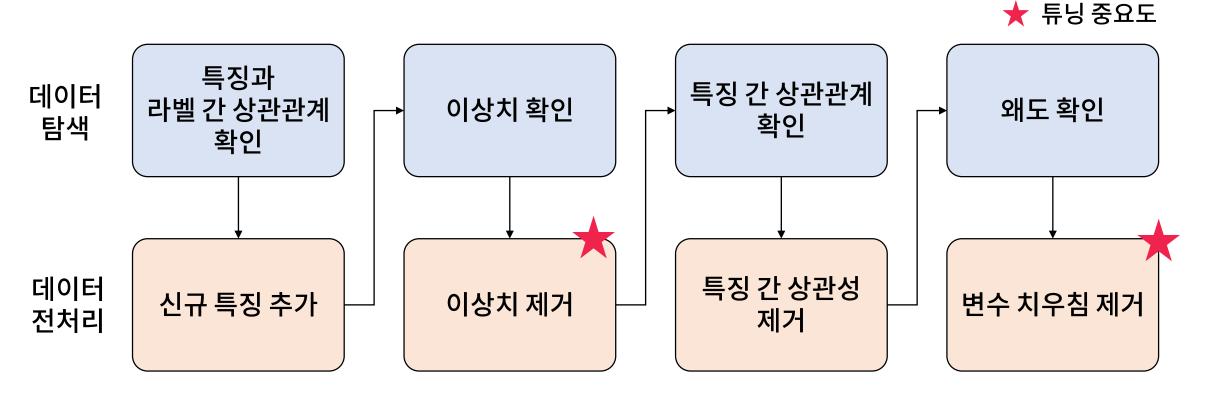
• 일반적으로 필수 전처리는 특별한 튜닝이 필요하지 않으므로 순차적으로 진행





l 현실 지도학습 모델 개발 프로세스 (2) 성능 향상을 위한 전처리

• 일반적으로 성능 향상을 위한 전처리는 튜닝도 같이 수행해야 함



 특징 선택 단계에서 차원을 축소하므로 신규 특징 추가, 변수 치우침 제거 등에서 특징을 다수 추가하는 것이 이전 스텝으로 돌아가는 피드백 루프를 줄일 수 있음

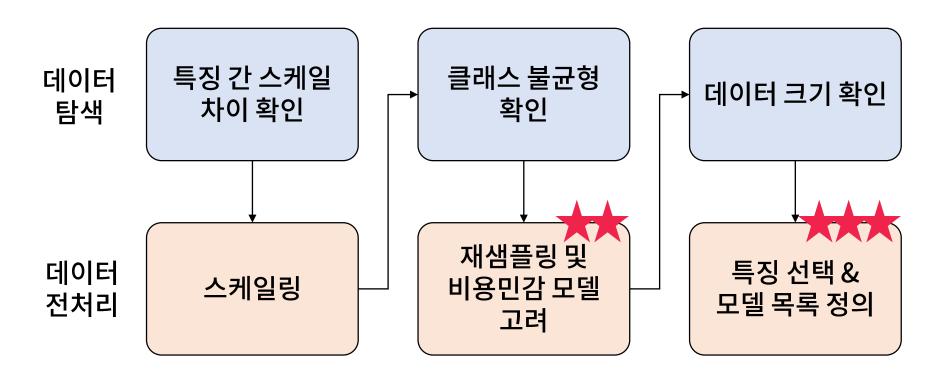
FAST CAMPUS ONLINE 안길승 강사.



l 현실 지도학습 모델 개발 프로세스 (2) 성능 향상을 위한 전처리

• 일반적으로 성능 향상을 위한 전처리는 튜닝도 같이 수행해야 함

★ 튜닝 중요도



fast campus

I 파라미터 그리드 설계

탐색을 제대로 하지 않으면 파라미터 그리드에 포함된 파라미터 개수가 수십 ~ 수백만개도 됨

이상치 제거 재샘플링 모델 특징 선택 모델 및 파라미터 파라미터 및 파라미터 파라미터 ▶ 특징 개수 Linear regression parameter IQR 기준치 SMOTE parameter **DBSCAN** Lasso / Ridge parameter **NearMiss** parameter parameter Neural network parameter Other **SVM** parameter resamplers **Decision tree parameter** Random forest parameter

FAST CAMPUS ONLINE 안길승 강사.





Chapter.

배운걸 적용해 보자: 실전 모델 개발 프로세스

감사합니다

FAST CAMPUS ONLINE 데이터 탐색과 전처리 I

강사. 안길승