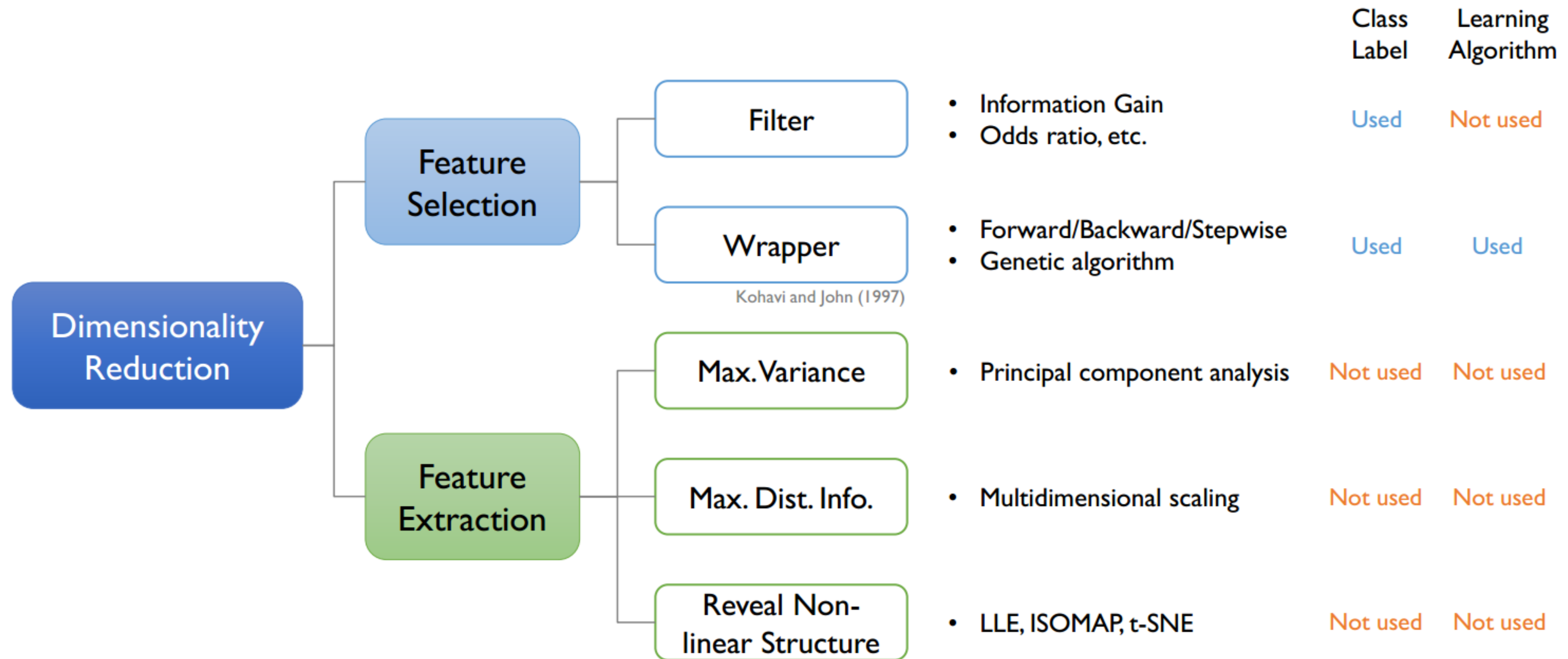


Feature_Selection(3)

240512_8기 데이터 분석 적용반



Mlxtend(Machine Learning Extensions)



Mlxtend (machine learning extensions) is a Python library of useful tools for the day-to-day data science tasks.

SequentialFeatureSelector

- **전진 선택 (forward=True):** 특성이 없는 상태에서 시작하여 모델을 개선하는 특성을 탐욕적으로 추가합니다. 원하는 특성 수에 도달할 때까지 각 반복에서 모델을 가장 많이 개선하는 특성을 추가합니다.
- **후진 제거 (forward=False):** 모든 특성을 포함한 상태에서 시작하여 각 반복에서 가장 중요도가 낮은 특성을 탐욕적으로 제거합니다.

ExhaustiveFeatureSelector

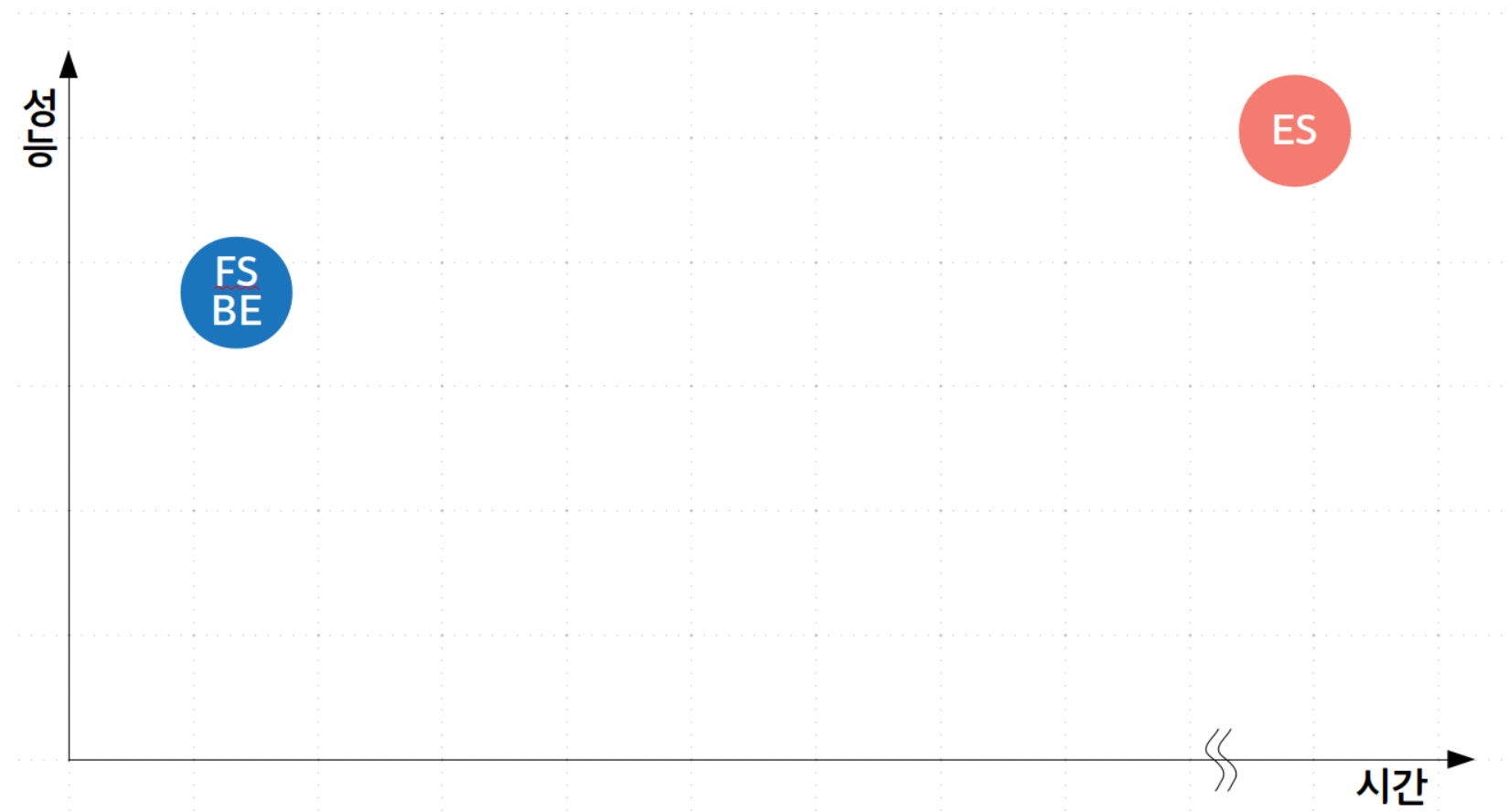
가능한 모든 특성 조합을 평가하여 지정된 성능 지표에 따라 최적의 세트를 결정하는 완전 탐색 특성 선택을 수행

래퍼(Wrapper)

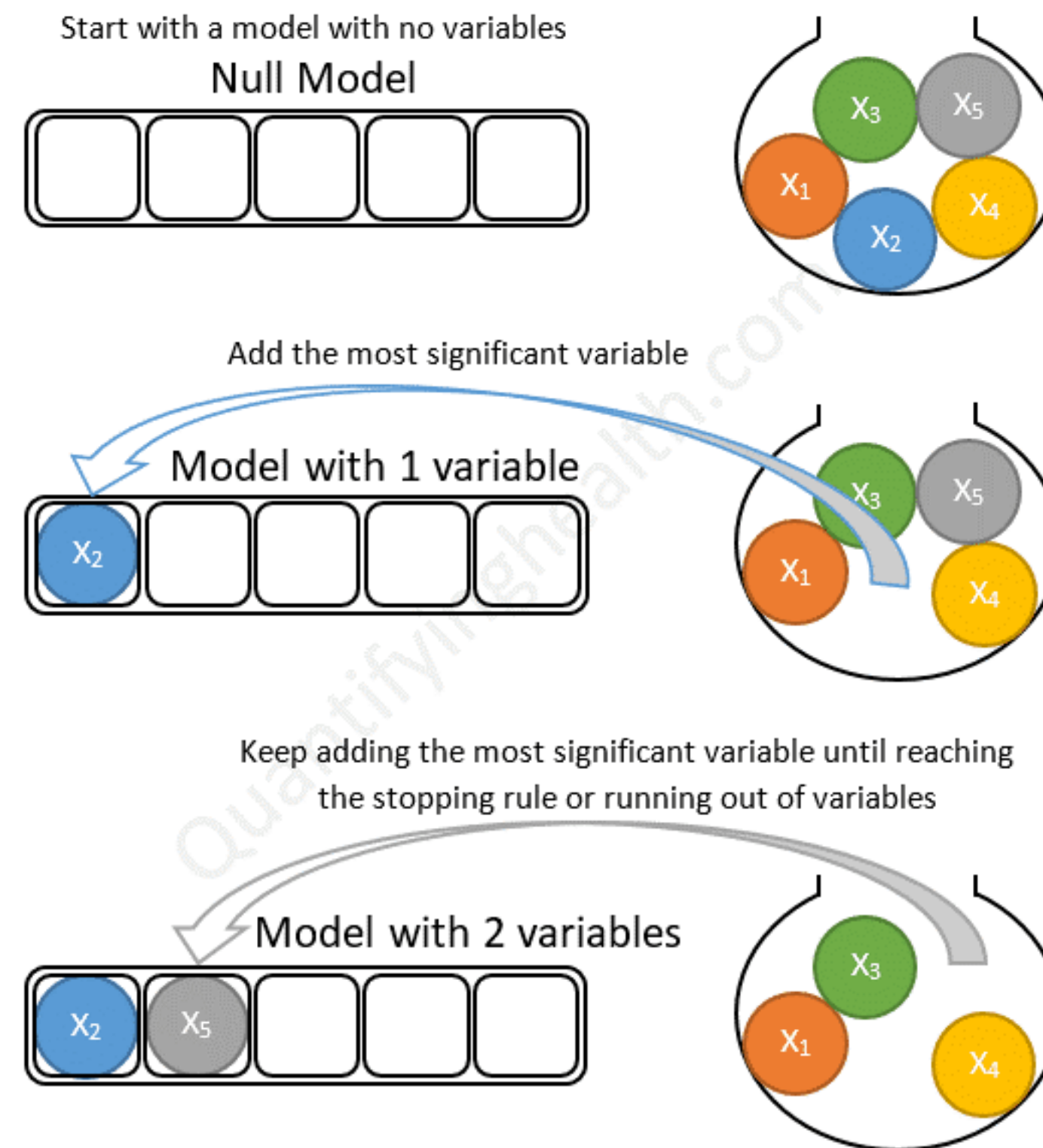
특성 선택(Feature selection)에 속하는 방법 중 하나로, 반복되는 알고리즘을 사용하는 지도 학습 기반의 차원 축소법입니다. 래퍼 방식에는 전진 선택(Forward selection), 후진 제거(Backward elimination), Stepwise selection 방식 뿐만아니라 유전 알고리즘(Genetic algorithm) 방식도 사용



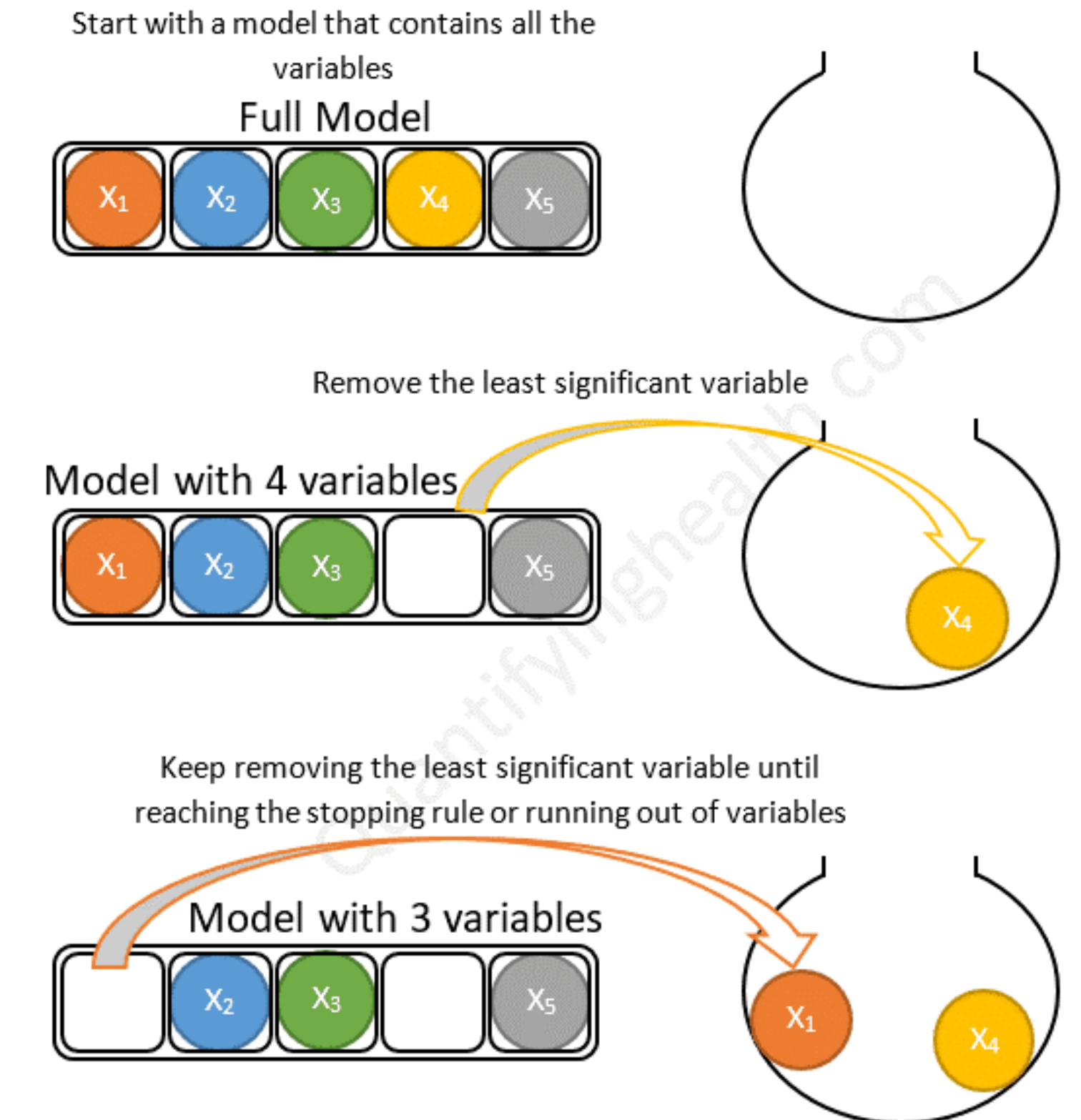
Forward Selection / Backward Elimination



Forward stepwise selection example with 5 variables:

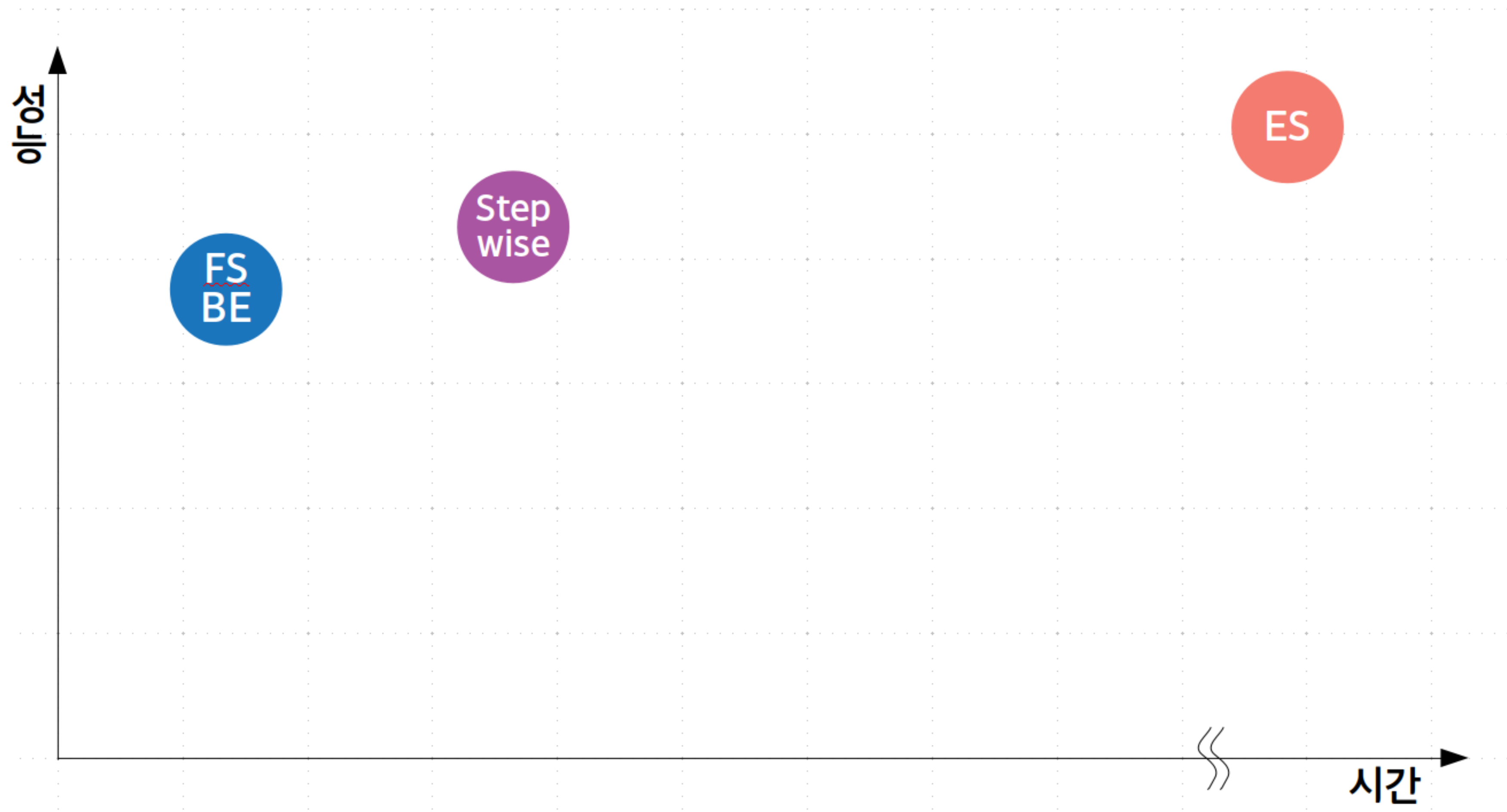


Backward stepwise selection example with 5 variables:



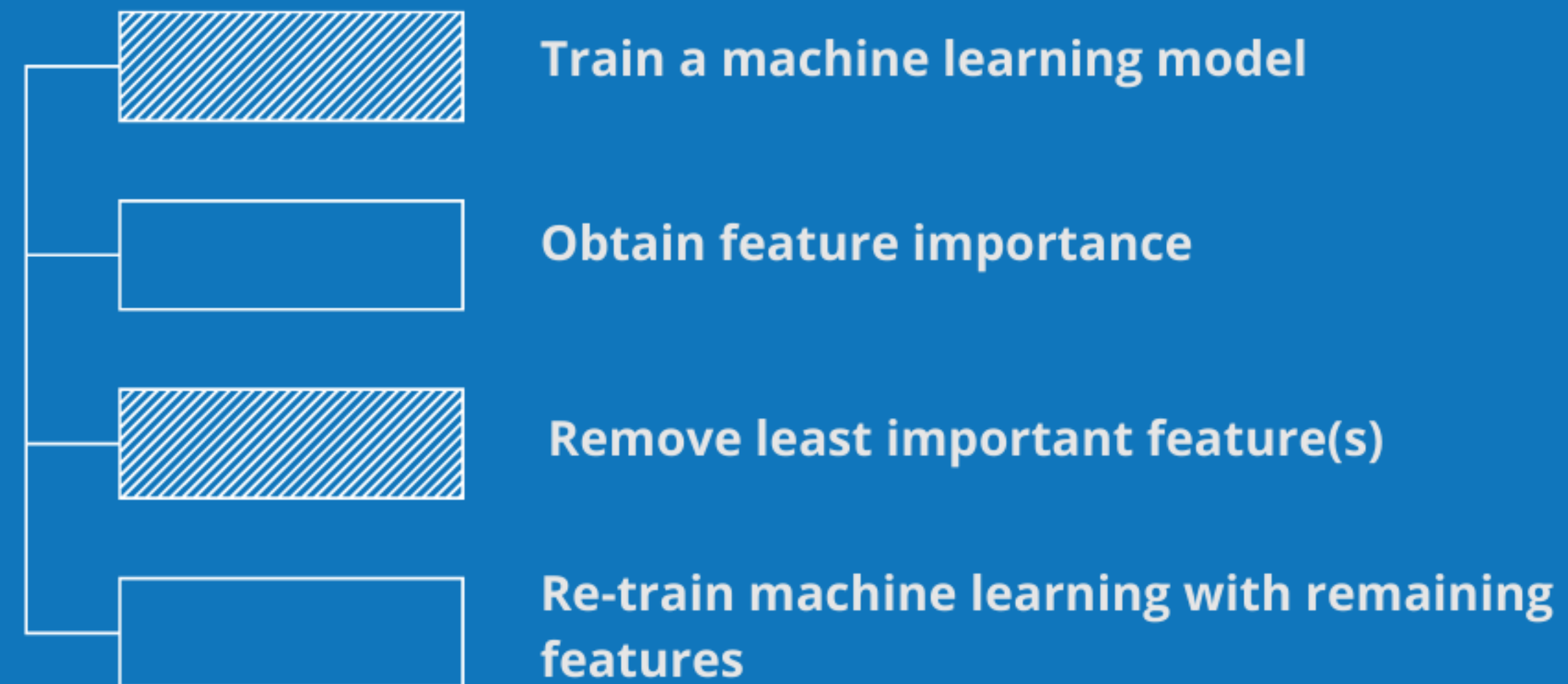
Stepwise Selection

전진 선택에서는 한 번 선택된 특성은 제거되지 않고, 후진 제거에서는 한 번 제거된 특성은 다시 선택되지 않습니다. 그렇기 때문에 두 방법 모두 더 많은 특성 조합에 대해 모델을 평가할 수 없다는 단점을 가지고 있습니다. Stepwise selection은 전진 선택과 후진 제거 방식을 매 단계마다 반복하여 적용하는 방식입니다. 이전 두 방법보다는 더 오래 걸리지만 최적의 변수 조합을 찾을 확률이 높습니다.

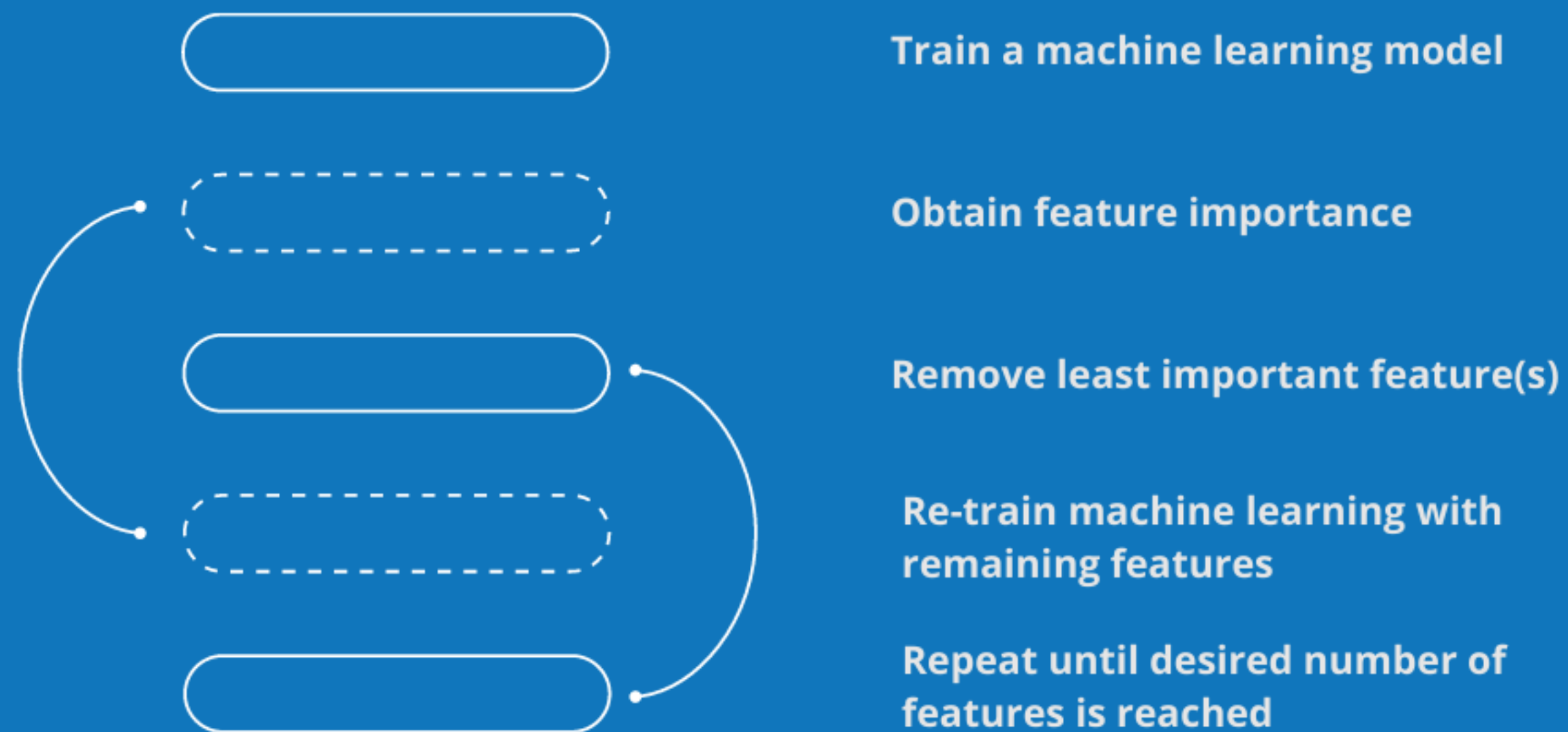


Recursive Feature Elimination

RFE - initial steps



RFE - Scikit-learn



Recursive Feature Elimination with Cross-Validation

Recursive Feature Elimination with Cross-Validation

