

사이킷런으로 수행하는 타이타닉 생존자 예측

[타이타닉 생존자 ML 예측 구현]

- 데이터 전처리
 - Null 처리
 - 불필요한 속성 제거
 - 인코딩 수행
- 모델 학습 및 검증/예측/평가
 - 결정트리, 랜덤포레스트, 로지스틱 회귀 학습 비교
 - K 폴드 교차 검증
 - `cross_val_score()`와 `GridSearchCV()` 수행

[2장 Summary]

머신러닝 지도 학습 프로세스

1. 데이터 전처리
 - 데이터 클리닝
 - 결손값 처리(Null/NaN 처리)
 - 데이터 인코딩(레이블, 원-핫 인코딩)
 - 데이터 스케일링
 - 이상치 제거
 - Feature 선택, 추출 및 가공
2. 데이터 세트 분리
 - 학습 데이터 / 테스트 데이터 분리
3. 모델 학습 및 검증 평가
 - 알고리즘 학습
 - 교차검증 | `cross_val_score()` | `GridSearchCV`
4. 예측 수행
 - 테스트 데이터로 예측 수행
5. 평가 – 예측 평가