

머신러닝 프로젝트 진행 절차

머신러닝 프로젝트 진행 절차

1. 문제 정의 : 비즈니스 문제 정의
2. 데이터 수집 : 데이터 세트를 식별하고 수집
3. 데이터 준비 : 데이터로부터 통찰을 얻기 위해 탐색하고 시각화
 - 결측치 대치
 - 중복 레코드 제거
 - 값 정규화
 - 다른 유형의 정리나 매핑 수행
 - 완전한 특성 추출
 - 상관 특성 제거
4. 데이터 분리 : 데이터를 훈련 세트, 검증 세트, 테스트 세트로 분할
5. 모델 학습 : 훈련 데이터 세트를 사용해 기계 모델은 훈련
6. 후보 모델 평가 : 모델 정확도를 결정하기 위해 테스트용 데이터와 검증용 데이터를 사용해 모델 성능 측정
7. 모델 배포 : 모델이 선택되면 추론을 위해 제품으로 배포
8. 성능 모니터링 :
 - 모델 성능을 지속적으로 모니터링하고 그에 따라 재교육 및 보정.
 - 새로운 데이터를 수집해 계속해서 모델 개선
 - 오래된 데이터를 사용하지 않도록 주의