

딥러닝

1. 인공지능 데이터 분석의 절차

1. 문제 정의 (Problem Definition)

분석의 목적과 목표를 명확히 설정

해결하려는 비즈니스 문제나 연구 질문 정의

필요한 데이터의 종류와 분석 방법 구상

2. 데이터 수집 (Data Collection)

내부 데이터: 데이터베이스, 로그 파일, CRM 등

외부 데이터: 공공 데이터, API, 웹 크롤링 등

정형 데이터(테이블 형태)와 비정형 데이터(텍스트, 이미지, 영상 등) 수집

3. 데이터 전처리 (Data Preprocessing)

결측치 처리 (NaN 값 제거, 대체)

이상치 탐지 및 제거

데이터 정규화 및 변환 (스케일링, 원-핫 인코딩 등)

데이터 통합 및 정리

4. 탐색적 데이터 분석 (Exploratory Data Analysis, EDA)

데이터 분포 및 기본 통계 확인 (평균, 표준편차, 분산 등)

시각화를 통한 패턴 탐색 (히스토그램, 박스플롯, 산점도 등)

변수 간 관계 분석 (상관 분석, 피어슨 상관계수 등)

5. 모델링 (Modeling)

지도 학습: 회귀 분석, 분류 모델 (예: 선형 회귀, 랜덤 포레스트, 딥러닝)

비지도 학습: 군집 분석, 차원 축소 (예: K-means, PCA)

시계열 분석: ARIMA, LSTM 등

최적의 모델 선택 및 하이퍼파라미터 튜닝

6. 모델 평가 (Model Evaluation)

성능 지표 확인 (정확도, F1-score, RMSE 등)

과적합 여부 확인 (교차 검증, 정규화)

모델 개선 (피처 엔지니어링, 데이터 추가, 알고리즘 변경 등)

7. 결과 해석 및 시각화 (Interpretation & Visualization)

주요 인사이트 도출

그래프 및 대시보드를 활용한 결과 시각화

모델이 제공하는 의미 있는 정보 정리

8. 배포 및 유지보수 (Deployment & Maintenance)

웹 애플리케이션, 대시보드, API로 배포

모델의 지속적인 모니터링 및 성능 개선

데이터 업데이트 및 재학습 수행

2. 인공지능 분류

구분	정의	특징
인공지능 (AI)	인간의 지능을 모방하는 기술. 다양한 문제 해결을 목표로 한다.	넓은 범위의 기술 포함. 머신러닝, 딥러닝도 AI의 일부다.
머신러닝 (ML)	데이터를 학습해 스스로 규칙을 찾아내는 AI의 하위 분야.	데이터 기반 학습. 알고리즘이 패턴을 인식하고 예측한다.
딥러닝 (DL)	머신러닝의 하위 분야로, 신경망을 사용해 복잡한 데이터를 처리한다.	이미지, 음성, 자연어 처리 등에서 뛰어난 성능을 보인다.

3. 머신 러닝 알고리즘의 정의

머신 러닝 종류	학습 방식	대표 사례
지도 학습	정답(레이블) 있음	이메일 스팸 필터링, 얼굴 인식
비지도 학습	정답 없음, 패턴 찾기	고객 세분화, 이상 탐지

강화 학습	보상 기반 학습	알파고, 자율 주행
자기 지도 학습	일부 데이터 기반 예측 (생성형 AI)	ChatGPT, 이미지 생성 AI

4. 지도 학습의 label(종속변수 : 답) 분류 (Classification)

분류 유형	예시	주요 알고리즘
이진 분류 (Binary)	합격 vs 불합격	로지스틱 회귀, SVM, 랜덤 포레스트
다중 분류 (Multi-class)	개 vs 고양이 vs 새	소프트맥스 회귀, 랜덤 포레스트, 신경망(CNN)
선형 분류 (Linear)	운동 선수 연봉 예측	퍼셉트론, LDA, 선형 SVM

5 활성화 함수

활성화 함수	수식	특징
시그모이드 (Sigmoid)		출력 범위 (0~1), 이진 분류
렐루 (ReLU)		연산 간단, 딥러닝에서 가장 많이 사용
소프트맥스 (Softmax)		다중 분류에서 확률값 출력

CNN(Convolutional neural network)

Convolutional neural network는 수십 또는 수백 개의 계층을 가질 수 있으며, 각 계층은 영상의 서로 다른 특징을 검출한다. 각 훈련 영상에 서로 다른 해상도의 필터가 적용되고, 컨벌루션된 각 영상은 다음 계층의 입력으로 사용된다. 필터는 밝기, 경계와 같이 매우 간단한 특징으로 시작하여 객체를 고유하게 정의하는 특징으로 복잡도를 늘려갈 수 있다.