

머신러닝 : 회귀분석부터 앙상블모델까지

1. 강의 개요

교육 목표: 이 강의는 머신러닝의 핵심 개념을 이해하고, 데이터 기반 모델을 구현 및 평가하는 능력을 기르는 것을 목표로 합니다. 지도학습(회귀/분류), 비지도학습(군집화/차원축소), 앙상블 학습을 배우며 실습을 통해 머신러닝 모델을 효과적으로 활용하는 방법을 익힙니다.

교육 대상: KAIST 학생 및 교직원, 일반인, Python 기초 문법을 알고 있고, 머신러닝을 배워보고 싶은 분

선수 학습: Python 기본 문법 (변수, 리스트, 함수 등), Pandas 및 Numpy 기본 사용법, 기초적인 선형대수 및 확률/통계 개념이 있으면 좋음

2. 강의 내용

1일차 - 지도학습(회귀 & 분류)

주제	시간	교육 내용
머신러닝 개요 및 파이프라인 이해	1	머신러닝의 정의, 지도/비지도 학습 개념, 모델 학습 과정
회귀 분석 (1) - 단순 선형 회귀	1	회귀의 개념, 가설 함수, 손실 함수(MSE), 최소 제곱법
회귀 분석 (2) - 다중 선형 회귀	1	다중 변수 활용, 벡터 표현, 성능 평가 (R^2 , MSE)
마무리	1	질의응답 및 실습 복습
과적합과 규제	1	과적합 문제, 정규화 기법 ($L1$, $L2$), 모델 복잡도 조절
분류 개념 및 KNN 모델	1	분류 vs 회귀 차이, KNN 원리, K 값 조정 실습
로지스틱 회귀 & SVM	1	로지스틱 회귀 기본 원리, SVM 개념 및 실습
성능 평가 지표 및 교차 검증	1	분류 문제 평가 (정확도, 정밀도, 재현율, F1-

마무리		score), 교차 검증 질의응답 및 실습 복습
-----	--	-------------------------------

2일차

주제	시간	교육 내용
비지도 학습 개요	1	지도학습과 차이점, 비지도학습의 필요성
K-means 군집화 실습 밀도 기반 및 계층적 군집화	1	K-means 개념, 최적의 K 선택법 (Elbow Method), 군집 시각화
	1	DBSCAN 개념 및 실습, 계층적 군집화 원리
마무리	1	질의응답 및 실습 복습
차원 축소 및 시각화 의사결정나무 랜덤포레스트 & 앙상블 학습	1	PCA, t-SNE 개념 및 2D 시각화
	1	결정 트리 구조, Gini 지수 및 엔트로피 개념, 과적합 문제 해결
	1	랜덤포레스트 개념, 배깅 & 부스팅 기법, 실습
마무리	1	질의응답 및 실습 복습

강의 진행 및 상황에 따라 내용이 변경될 수 있습니다.

3. 강의 세부 사항

필수 강의 내용: 머신러닝의 핵심 개념(회귀, 분류, 군집화, 차원 축소)부터 성능 평가 및 앙상블 모델까지 실습을 통해 배우는 실전 중심 강의

- 머신러닝의 기본 개념과 파이프라인
- 회귀 분석(단순 회귀, 다중 회귀)
- 과적합 문제 및 규제 기법 (L1, L2 정규화)
- 분류 모델 (KNN, SVM, 로지스틱 회귀)
- 성능 평가 지표 및 교차 검증
- 비지도학습 (군집화, 차원축소)
- 의사결정나무와 랜덤포레스트

교육 환경: >= Python 3.6, Scikit-learn, Pandas, Numpy

추천 참고 책: 파이썬 머신러닝 완벽 가이드