

머신러닝과 데이터분석 A-Z

1. 강의 소개

본 강의는 머신러닝 입문자를 대상으로, 파이썬 기초부터 머신러닝까지 개념을 이해한 다음 머신러닝을 활용한 데이터분석까지 실습을 통해 진행하는 과정입니다.

2. 파트 구성

- 01. Python Programming 기초

파이썬을 아직 접해보지 않은 분들을 위한 파트입니다. 파이썬의 기초적인 문법부터 클래스와 인스턴스, 정규표현식까지 꼭 배워야하는 문법들을 학습합니다.

- 02. 데이터 수집을 위한 Python (Crawling)

파이썬은 웹 크롤링을 위한 대표적인 언어입니다. 반복되는 작업을 효율적으로 줄일 수 있는 크롤링을 위한 간단한 문법과 실제 예제를 통해 크롤링을 배워봅니다.

- 03. 데이터 처리를 위한 Python (Numpy)

파이썬은 수치계산과 벡터연산에 최적화된 라이브러리 numpy를 제공하고 있습니다. numpy를 통해 데이터 처리의 기초를 학습해봅니다.

- 04. 데이터 분석을 위한 Python (Pandas)

데이터는 '엑셀'과 같이 행과 열로 이루어져 있습니다. 파이썬은 pandas를 통해 행렬로 이루어진 데이터프레임을 다룰 수 있습니다. pandas에 대해 알아봅니다.

- 05. Machine learning의 개념과 종류

본격적으로 머신러닝을 학습하기 전에 머신러닝과 딥러닝 전반에 대한 개념을 숙지합니다.

- 06. 회귀분석

머신러닝의 가장 기본 '회귀분석'을 알아봅니다. 회귀분석을 통해 지도학습의 개념을 이해하고 나아가 차원축소를 위한 PCA도 간단하게 다뤄봅니다.

- 07. 기본적인 Machine learning 모형

대표적으로 사용되는 머신러닝 모형은 나이브베이즈, KNN, LDA, SVM, 의사결정나무, 신경망 총 6개 정도로 추려볼 수 있습니다. 수많은 알고리즘들이 더 있지만 기본적으로 알아야 하는 6개 알고리즘을 먼저 알아보고 머신러닝의 개념을 더 깊게 이해합니다.

- 08. 앙상블 기법의 종류와 원리

앞에서 배운 알고리즘들을 이리저리 섞어서 더 좋은 정확도를 내는 방법으로 앙상블 기법이 사용됩니다. 앙상블의 원리와 개념을 실습을 통해 알아봅니다.

- 09. 클러스터링

K-means 클러스터링은 비지도학습의 가장 기본적인 모형입니다. 이를 먼저 이해하고 더 나아가 Hierarchical 클러스터링과 DBSCAN 클러스터링까지 알아봅니다.

- 10. 불균형 데이터

실제 데이터를 마주하면 항상 일정한 비중의 데이터가 주어지지 않습니다. 불균형 데이터를 받았을 때 어떤 방식으로 처리해야 좋은 성능을 얻어낼 수 있을지 알아봅니다.

- 11. 빅콘테스트 Review

강사님이 참여한 빅콘테스트에 대해 개발 과정과 인사이트를 공유합니다.

- 12. 딥러닝

딥러닝은 신경망 알고리즘으로 구성된 머신러닝의 한 기법입니다. 신경망의 원리와 손실함수, 최적화 알고리즘 등을 통해 딥러닝을 이해하고 대표적인 딥러닝 모델인 CNN, RNN에 대해 알아봅니다.

- 13~17. [Project]

지금까지 배운 내용을 활용해 단계적으로 실습 프로젝트에 도전해봅니다.