

파이썬 데이터분석 강의 소개

소개

최근 몇 년 사이에 머신러닝/딥러닝 기술이 획기적으로 발전하면서 데이터분석 및 인공지능 관련 연구의 중요성이 사회, 경제, 산업의 거의 모든 분야에 지대한 영향을 미치고 있으며, 앞으로 한 동안 그런 경향이 더욱 강화될 것으로 기대된다. 이에 본 강의에서는 데이터분석의 기본 아이디어와 다양한 활용법을 실전 예제와 함께 전달하여 학습자의 문제해결능력을 향상시키고자 한다. 데이터분석의 다양한 개념과 기법에 대한 이해는 이어지는 머신러닝/딥러닝 강좌 수강을 위한 필수조건이기도 하다.

실습에 사용되는 언어는 파이썬(Python)이며, 현재 데이터분석 및 머신러닝 분야에서 가장 많이 사용되는 언어이다. 파이썬 프로그래밍 지식은 강의에 필요한 만큼만 학기 초반에 소개한다. 다만 짧은 시간에 상대적으로 많은 내용을 다루기에 파이썬 기초지식을 학기 시작 이전에 학습해두어야 한다. 미리 공부해야 할 파이썬 기초 내용은 아래 **예비학습 안내**에서 소개한다.

데이터분석 알고리즘의 작동원리를 제대로 이해하려면 선형대수, 확률과통계, 미적분학에 대한 기본 지식이 있어야 한다. 하지만 이론은 최소로 필요한 정도만 다룰 것이기에 따로 예비학습이 요구되지는 않는다.

참고: 인공지능, 데이터과학, 머신러닝, 딥러닝은 아래의 관계를 갖는다.

- 인공지능: 사고나 학습 등 인간이 가진 지적능력을 컴퓨터를 통해 구현하는 기술 또는 해당 연구 분야
- 데이터과학: 수학과 통계 지식을 활용하여 데이터로부터 특정 정보를 추출하는 과학
- 머신러닝: 컴퓨터가 데이터로부터 스스로 정보를 추출하는 기법. 데이터과학, 인공지능 등의 분야에서 가장 중요한 문제해결 기법 또는 해당 연구 분야
- 딥러닝: 인공신경망 이론을 기반으로 복잡한 비선형 문제를 해결하는 머신러닝 기법 또는 해당 연구 분야



그림 출처: [교보문고: 에이지 오브 머신러닝 \(https://bit.ly/2YmwB7U\)](https://bit.ly/2YmwB7U)

참여 전제조건

- 선형대수, 확률과통계, 미적분학 과목 이수
- 아래 예비학습 안내에서 소개한 내용을 학기 시작 이전에 미리 공부하기

주요 학습내용

- 파이썬 언어의 기본 소개
- 내장 자료구조, 함수, 파일
- numpy 기본: 배열과 벡터 연산
- pandas 시작하기
- 데이터 로딩과 저장, 파일 형식
- 데이터 정제 및 준비
- 데이터 준비하기: 조인, 병합, 변형
- 그래프와 시각화
- 데이터 집계와 그룹 연산
- 시계열
- 고급 pandas
- 파이썬 모델링 라이브러리
- 데이터분석 예제

교재

아래 세 권의 책의 내용을 적절히 조합해서 다룰 것이며 온라인으로 강의노트를 제공한다.

- 파이썬 라이브러리를 활용한 데이터 분석, 웨스 맥키니, 한빛미디어

머신러닝(응용수학전공) 강의 예비학습 안내

예비학습 주요 내용과 목적

학기가 시작하기 전에 아래 내용을 미리 공부해야 하며, 이유와 목적은 다음과 같다.

- 파이썬 프로그래밍 왕기초 학습
- 학기 중에 학습할 내용 일부 익혀두기
- 학기 초에 아래 내용을 반복하겠지만 보다 많은 내용을 보다 빠른 속도로 학습 예정
- 파이썬 기초지식이 부족하면 학습에 많은 어려움 발생

예비학습 1

- 구글 코랩(Google Colab) 활용법 익혀두기
- 구글 코랩만 이용하여 강의 진행 예정
- 참조: [구글 코랩 기초 사용법 안내 동영상 \(https://bit.ly/2L1NAcE\)](https://bit.ly/2L1NAcE)

예비학습 2

- 파이썬 데이터과학 입문 강좌에서 소개하는 파이썬 프로그래밍 왕기초 학습
 - 참조: [파이썬 데이터과학 입문 \(https://formal.hknu.ac.kr/Gongsu-DataSci/\)](https://formal.hknu.ac.kr/Gongsu-DataSci/)
- 실습환경: [구글 코랩\(Google Colab\) \(https://colab.research.google.com/\)](https://colab.research.google.com/)

예비학습 3

- 예비학습2 과정을 마친 후 아래 동영상 시리즈 시청 추천
 - 참조: [Data analysis in Python with Pandas \(http://bit.ly/3ov40rn\)](http://bit.ly/3ov40rn)
 - 내용: 다양한 판다스 실전 활용법 소개
- 자동번역 한글자막 수준이 매우 우수함.
 - 설정 => 자막 => 자동 번역 => 한국어.
- 참조: [강의노트 \(http://bit.ly/3osVCJ7\)](http://bit.ly/3osVCJ7)

학습법

- 학습법: 강의노트에서 소개하고 설명하는 내용을 구글 코랩에서 직접 타이핑 하면서 따라해볼 것.
- 주의사항: 주어진 구글 코랩 강의노트를 단순히 실행만 하면 제대로 학습되지 않을 것임.