

교육 과정 소개서.

파이썬을 활용한 머신러닝 CAMP



코스요약

코스명	파이썬을 활용한 머신러닝 CAMP
기간	2020.07.04 - 2020.09.05 (8월 15일 휴강)
일정	매주 토요일 10:00-13:00
장소	패스트캠퍼스 강남강의장
준비물	개인노트북
담당자	02-501-9396 / help-ds@fastcampus.co.kr
수강료	1,500,000
상세페이지 url	fastcampus.co.kr/data_camp_pmb/

코스목표

본 강의는 이론 뿐만 아니라 실습을 통해 머신러닝 알고리즘 적용의 결과를 해석하고 이를 시각화 할 수 있고, 강의 내에서 제공되는 코드를 통해 실제의 내 데이터에 적용하여, 데이터 분석 기반을 다질 수 있습니다.

코스정보

본 강의에서는 머신러닝을 처음 접하시는 분들을 위하여 머신러닝이 무엇인지, 무엇을 할 수 있고 무엇을 할 수 없는지를 소개하며 머신러닝에 들어가는 수학, 통계적인 내용을 이해하기 쉽게 알려드립니다. 더불어 앞서 배운 이론을 어떻게 파이썬 코드로 구현하는지에 대하여 학습합니다.



코스특징

사례를 통해 재미있게 배우는 머신러닝

어렵게만 보였던 머신러닝, 지루하기까지 했나요? ‘타이타닉 탑승객의 분류’, ‘농구선수의 슛 성공/실패 분류’, ‘목소리 신호 데이터로 성별 분류’와 같이 모두에게 흥미로운 예제를 통해 머신러닝을 보다 재미있게 이해해봅니다.

이해하기 쉽도록 반복되는 학습과 피드백

개념과 수식 이해, 실습 등이 유기적으로 이뤄져야 머신러닝 알고리즘을 제대로 이해할 수 있습니다. 개념설명 → 수식설명 → 파이썬 실습을 반복하며 알고리즘의 이해를 돕고, 이를 온전히 가져갈 수 있도록 퀴즈 및 피드백을 제공합니다.

필요한 만큼의 수식 설명으로 수학에 대한 부담감 해소

머신러닝 공부의 가장 큰 장벽 수학! 하지만 모든 수식에 대해 알아야 할 필요는 없습니다. 알고리즘 이해를 위한 최소한의 수학적 내용만을 압축해서 전달하여, 수학에 대한 부담감이 대폭 줄어듭니다.



커리큘럼

1주차 ● 머신러닝 개요 및 파이썬 기초 (Introduction)

- 머신러닝의 활용사례와 이슈
- 실습을 위한 파이썬 환경 구성
- Anaconda 를 이용한 실습 환경 설정
- Github 을 이용한 실습 코드 및 데이터 공유
- Python, Numpy, Pandas, Bokeh 소개

2주차 ● 선형 회귀 모델 (lineaar regression)

- Least Squares 와 Maximum Likelihood Estimation 를 이용한 선형 회귀 모델 풀이
- 미분을 이용한 최적화 탐색
- 다중 선형 회귀 모델 (Multiple linear regression)
- 다항 회귀 모델 (Polynomial regression)
- 모델 적합도와 검정

3주차 ● 로지스틱 회귀를 이용한 분류 (logistic regression)

- 로지스틱 회귀의 소프트맥스 회귀로의 확장
- 과적합과 L2/L1 regularization
- Logit 에 대한 선형 회귀 모델로의 해석
- 문서 분류 예시를 통한 모델 계수의 의미 해석

4주차 ● -1 변수 엔지니어링과 데이터 전처리

- 데이터의 벡터 표현 및 정규화 (normalization)
- 변수 추출법인 Singular Vector Decomposition
- 고차원 벡터의 차원 축소를 통한 시각화 (t-SNE, UMAP)

4주차 ● -2 모델의 성능 검증 및 데이터 분리

- 판별 성능 척도 (precision, recall, F1, accuracy, confusion matrix)
- 회귀 성능 척도 (MSE, MAP)
- 교차 검증 (cross validation)
- 하이퍼 패러미터 탐색 방법



커리큘럼

5주차 • 인공신경망을 이용한 비선형 분류 및 회귀 (feed forward neural network)

- Linear inseparable 문제 (XOR)
- 인공신경망의 구성 요소 (hidden layer, activation function)
- 인공신경망의 기하학적 해석 (representation learning)
- optimizers (SGD, Adam)

6주차 • 지지 벡터 머신 (Support Vector Machine)

- Linear kernel SVM
- Radial Basis Function (RBF) kernel SVM
- Support Vector Regression

7주차 • -1 의사결정 나무 (Decision Tree)

- 의사결정 나무를 이용한 분류와 회귀

7주차 • -2 앙상블 기법

- Bagging / Boosting
- Random Forest 와 Extreme Forest
- XGBoost
- Isolation Forest 를 이용한 이상치 (outlier) 탐색

8주차 • 최인접이웃 기반 모델 (k-Nearest Neighbor models)

- k-NN 기반 분류/회귀 모델
- Ball Tree 를 이용한 빠른 최인접이웃 검색

9주차 • 군집화 (Clustering)

- k-means
- hierarchical clustering
- DBSCAN
- UMAP 을 이용한 군집화 학습 결과의 시각화



강사소개



김현중

- (2016~현재) 패스트캠퍼스에서 파이썬, 머신러닝 관련 강의 다수 진행
- 서울대학교 산업공학과 데이터마이닝 연구실 박사과정
- 한국어 텍스트 분석을 위한 soynlp 라이브러리 개발
- PYCON KOREA 2017 '노가다 없는 텍스트 분석을 위한 한국어 NLP' 발표



수강환경

강남강의장



❖ 강의에 따라 강의장이 변경될 수 있습니다.