

XML 이론(데이터 분석) 및 실습

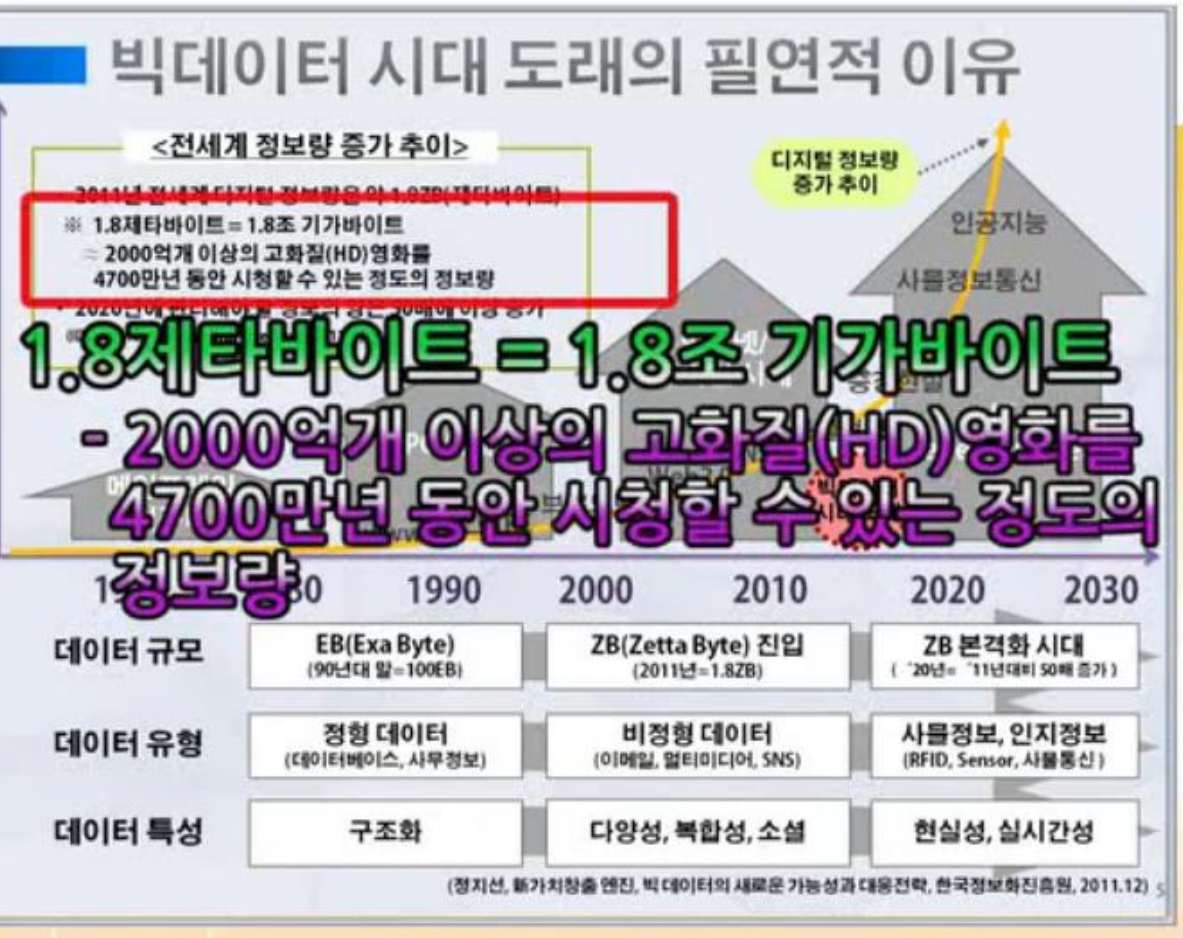
강윤희

yhkang@bu.ac.kr, yunh.kang@gmail.com

차례

- 빅데이터시대 준비하기
- 수업목표
- 교재 및 부교재
- 주차별 수업내용
- 수강 이유
- 수업 진행 및 평가

빅데이터 시대 준비하기



하드웨어 → 소프트웨어 → 데이터 기반 사회로 이동



빅데이터 시대 준비하기

빅데이터의 사회경제적인 가치 전망

구분	기관명	주요 내용
산업 경제성	Economist (2010)	☑ 데이터는 자본이나 노동력과 거의 동등한 레벨의 경제적 투입 자본, 비즈니스의 새로운 원자재 역할
	Gartner (2011)	☑ 데이터는 21세기 원유 , 데이터가 미래 경쟁 우위를 좌우 ☑ 기업은 다가올 ' 데이터 경제 시대 ' 를 이해하고 정보 고립(Information Silo)을 경계해야 성공 가능
	McKinse (2011)	
국가 경쟁력	美 대통령 과학기술자문위	☑ 미국 정부기관들이 데이터를 지식으로, 지식을 행동으로 변환하는 전략에 집중해야 함을 주장
	싱가포르	☑ 데이터를 기반으로 싱가포르를 위협하는 리스크 에 대한 평가와 환경변화를 탐지

데이터는 21세기 원유



KERIS 수요포럼
**빅데이터시대
전망과 대응전략**



김 현 곧기획단장

한국정보화진흥원

한국교육학술정보원

빅데이터 시대 준비하기

4차 산업혁명 시대에서 소음 진동 분야의 전망과 역할은 무엇입니까?

최근 소음 진동 신호를 기반으로 제품의 이상 유무를 판단하는 “기계 진단” 분야에서 4차 산업혁명의 핵심 기술인 인공지능이 활용되고 있습니다. 현대자동차의 AI 카닥터와 에스엠인스트루먼트사의 Doberman이 대표적인 예입니다. 해당 제품은 **측정한 데이터로부터 고장 유무와 관련 있는 특징을 추출하고 이를 기반으로 인공지능을 학습시키는 과정**으로 이루어집니다. 이 기술은 인공지능을 활용함으로써 데이터 측정·분석·판단과 같은 소음·진동 분야의 전통적인 진단 과정을 간소화하는 동시에 정확도 또한 향상시키는 효과가 있습니다. 이처럼 인공지능 기술의 도래가 기계 진단의 패러다임 변화를 야기한 것과 같이 4차 산업혁명 기술, 학문 간 융합이 소음 진동 분야 전반의 Process 변화를 초래할 수 있습니다.

다만 소음·진동 분야에 인공지능 기술을 접목하기 위해서는 기본적으로 신호의 특징을 잘 이해하고, 이를 가공하여 적절한 기능을 구현하도록 인공지능을 학습시키는 과정이 필요합니다. 이를 위해서는 소음 진동 발생의 근본적인 물리적 현상에 대한 이해가 반드시 필요합니다. 따라서 소음 진동 분야에 종사하는 엔지니어들이 그간의 전통적인 이론을 바탕으로, 새롭게 다가오는 4차 산업혁명 기술과의 연결점을 찾기 위한 많은 노력을 기울여야 할 것으로 생각합니다.

수업목표

- JSON/XML 형식을 포함한 다양한 정보 표현 자료형식을 이해한다
 - 웹 (HTML) 및 공공데이터(JSON/XML)를 수집하여 데이터 분석을 위한 데이터 준비과정을 이해한다
 - 정제, 변환을 포함한 데이터 준비 과정인 데이터 파이프라인(data pipeline)을 프로그래밍한다
 - 수집되어진 자료에서 필요한 자료를 추출한 후 추출된 자료 저장을 프로그래밍한다.
 - 데이터 분석처리 결과를 위해 시각화(visualization) 방법을 이해하고 프로그래밍한다
- ❖ 자신의 수강목표와 수업목표와 일치 여부를 생각함

교재 및 부교재

- **Text:** 수업시간에 제공하는 자료 (미출판, 빅데이터 분석)

- 부교재


- (데이터수집자동화)한입에 웹 크롤링, 김경록,서경덕
- (패키지)모두의 데이터 분석 with 파이썬, 송석리, 이현아

- KOCW

- 빅데이터 프로그래밍, 건국대학교 이정환

1.  [컴퓨터 프로그램의 이해](#)

2.  [파이썬 프로그래밍 -01](#)

3.  [파이썬 프로그래밍 -02](#)

4.  [파이썬 프로그래밍 -03](#)

5.  [파이썬 프로그래밍 -04](#)



교재 및 부교재

#1

fastest growing
programming language

SOURCE: TechRepublic

28%

increase in data science
jobs by 2020

SOURCE: IBM

65%

developers use & want to
use Python

SOURCE: The Economist

1. 파이썬 프로그래밍 소개 (1)

- 파이썬 프로그램 파일
 - ✓ 파이썬 코드는 ".py"로 끝나는 파일과 함께 텍스트 파일에 저장 됨
 - ✓ 모든 행은 파이썬 문 또는 그 일부라고 가정함
- 모듈
 - ✓ 파이썬의 대부분의 기능은 modules에 의해 제공됨
 - ✓ 모듈을 import하면 dir 함수를 사용하여 모듈안에 포함된 심볼/함수 확인가능
- 변수 및 유형
 - ✓ 심볼 이름 : 영숫자 a-z, A-Z, 0-9 및 _ 와 같은 특수문자 포함
 - ✓ 할당 : 대입 연산자 '='

교재 및 부교재

2. 파이썬 프로그래밍 소개 (2)

- 복합 데이터 유형
 - ✓ 문자열 : 텍스트 메시지 저장
 - ✓ 리스트 : 각 요소가 모든 유형이 될 수 있다는 점을 제외하고는 문자열과 유사함
 - ✓ 튜플 : 생성되면 수정이 불가능함
 - ✓ 사전 : 각 요소는 키-값(key-value) 쌍으로 구성됨
- 흐름 제어
 - ✓ 조건문 : if, elif, else
- 루프
 - ✓ for loops
 - ✓ while loops

교재 및 부교재

3. 파이썬 프로그래밍 소개 (3)

인쇄하기

■ 함수

✓ 기본 인수 및 키워드 인수

- 함수가 취하는 인수에 기본값을 줄 수 있음
- 함수 myfunc를 호출 할 때 debug 인수의 값을 제공하지 않으면 함수정의에 제공된 값이 기본값으로 사용하게 됨

✓ 이름없는 함수(lambda 함수)

- lambda 키워드를 사용함

교재 및 부교재

3. 파이썬 프로그래밍 소개 (3)

인쇄하기

■ 클래스

- 객체와 객체에서 수행할 수 있는 작업을 나타냄
- 속성(변수)과 메소드(함수)를 포함

■ 모듈

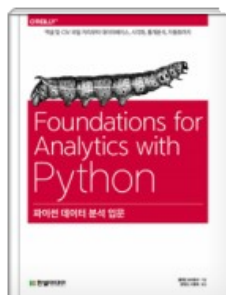
- 모듈에서 관련된 변수, 함수 및 클래스를 수집할 수 있는 고레벨(higher-level) 모듈식 프로그래밍 구조
- 파이썬 파일 (확장자 .py)로 정의됨

■ 예외처리

- 예외 생성을 위해 raise 문 사용가능

Python 추천 입문서

파이썬 데이터 분석 입문 엑셀 및 CSV 파일 처리부터 데이터베이스, 시각화, 통계분석, 자동화까지



★★★★★ 0.0 | 네티즌리뷰 2건

저자 클린턴 브라운리 공역 | 역자 한창진, 이병욱 | 한빛미디어 | 2017. 10. 25

원제 Foundations for Analytics with Python

페이지 368 | ISBN 9791162240144 | 판형 규격외 변형

도서 24,300원 27,000원 -10%

e북 17,010원 18,900원 -10%

구매해택 상세보기 >

39



바로구매

예스24	N Pay 1%	24,300원	구매
인터넷 교보문고	N Pay 1%	24,300원	구매
알라딘	N Pay 1%	24,300원	구매
인터파크 도서	N Pay 6%	24,300원	구매
강컴닷컴	N Pay 1%	24,300원	구매
반디앤루니스	N Pay 6%	24,300원	구매
영풍문고	N Pay 6%	24,300원	구매
도서11번가		24,300원	구매
커넥트북	N Pay 6%	24,300원	구매

e북 예스24	N Pay 1%	17,010원	구매
e북 인터파크 도서	N Pay 6%	17,010원	구매
e북 인터넷 교보문고	N Pay 1%	17,010원	구매
e북 리디북스	N Pay 1%	18,900원	구매

책정보

출판사 서평

네티즌 리뷰

가격정보

책소개

엑셀의 한계를 넘어 파이썬을 이용한 데이터 분석으로

엑셀은 쉽고 직관적이지만 수백 개의 파일을 다루기는 무리다. 파이썬을 이용하면 엑셀의 한계를 넘어 다양한 형식의 데이터를 가공하고 수많은 파일을 분석할 수 있다. 이 책은 일반인을 위한 파이썬 기초부터 CSV, 엑셀, 데이터베이스의 데이터를 분석하는 파이썬 코드 작성법까지 친절하게 알려준다. 파일 파싱, 그룹화, 통계 산출, 시각화에 필요한 각종 파이썬 라이브러리도 함께 알아본다. 데이터 분석 고수가 되겠다는 의지가 있다면 프로그래밍 경험은 필요 없다.

[인터넷 교보문고 제공]

오디오북 까칠한 아이
딱 7일간 90% 할인



출판사 추천도서



엄마의 말하기 연습
박재연

Python 추천 입문서

책 홈 > 컴퓨터/IT > IT 전문서 > 프로그래밍언어

파이썬 프로그래밍 데이터 과학을 위한



★★★★★ 9.0 | 네티즌리뷰 1건

저자 최성철 | 한빛아카데미 | 2019.01.01

페이지 536 | ISBN 9791156644361 | 판형 규격외 변형

도서 23,500원 25,000원 -6%

구매해택 상세보기 >

5



바로구매

에스24	N Pay 1%	25,000원	구매
인터넷 교보문고	N Pay 1%	25,000원	구매
알라딘	N Pay 1%	25,000원	구매
인터파크 도서	N Pay 6%	24,250원	구매
강컴닷컴	N Pay 1%	24,500원	구매
반디앤루니스	N Pay 6%	23,500원	구매
영풍문고	N Pay 6%	24,000원	구매
도서11번가		25,000원	구매
커넥츠북	N Pay 6%	25,000원	구매

책정보

네티즌 리뷰

가격정보

책소개

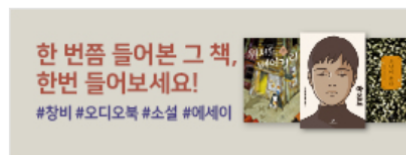
데이터 과학을 기반으로 쉽게 배우는 파이썬 프로그래밍

《데이터 과학을 위한 파이썬 프로그래밍》은 K-MOOC 공개강좌인 최성철 교수의 ‘데이터 과학을 위한 파이썬’ 강의를 기반으로 집필하였습니다.

이 책은 파이썬으로 프로그래밍을 시작하는 입문자가 쉽게 이해할 수 있도록 기본 개념을 상세하게 설명하며, 다양한 예제를 제시합니다. 또한 프로그래밍의 기초 원리를 이해하면서 파이썬으로 데이터를 처리하는 기법도 배웁니다. 파이썬을 혼자 공부하기 어려워하는 독자를 위해 저자 직강의 동영상 강의를 유튜브를 통해 무료로 제공하며, 이를 통해 학습의 이해도를 더욱 높일 수 있습니다.

※ 본 도서는 대학 강의용 교재로 개발되었으므로 연습문제 해답은 제공하지 않습니다

[인터넷 교보문고 제공]



같은 분야의 인기책



모두의 데이터 분석 with 파이썬
송석리 외 1명
★★★★★ 0



Do it! 안드로이드 앱 프로그래밍
정재곤
★★★★★ 0

Python 추천 입문서

점프 투 파이썬



지은이 : 박응용

최종 편집일시 : 2019년 2월 12일 9:22 오후

저작권 : (CC) BY-NC-ND

e-book 판매가 : 5,000원 ([구매하기](#))

👍 2,432 명이 추천

점프 투 파이썬 오프라인 책 출간 !! (2016.03)

- [책 구입 안내](#)

이 책은 파이썬이란 언어를 처음 접해보는 독자들과 프로그래밍을 한 번도 해 본적이 없는 사람들을 대상으로 한다. 프로그래밍을 할 때 사용되는 전문적인 용어들을 알기 쉽게 풀어서 쓰려고 노력하였으며, 파이썬이란 언어의 개별적인 특성만을 강조하지 않고 프로그래밍 전반에 관한 사항을 파이썬이란 언어를 통해 알 수 있도록 알기 쉽게 설명하였다.

파이썬에 대한 기본적인 지식을 알고 있는 사람이라도 이 책은 파이썬 프로그래밍에 대한 흥미를 가질 수 있는 좋은 안내서가 될 것이다. 이 책의 목표는 독자가 파이썬을 통해 프로그래밍에 대한 전반적인 이해를 갖게하는 것이며, 또 파이썬이라는 도구를 이용하여 원하는 프로그램을 쉽고 재미있게 만들 수 있게 하는 것이다.