

딥러닝 이론을 몰라도 개발하는 AI 프로그램

2021.01.09



AI Programming

이론 없이, 바로 지금 출발합니다!



이론 없이 출발한다고 말은 했지만...
이론은 실습하면서 자연스럽게
몸으로 배우게 될 것 같습니다. ^-^)/

agenda

- Python 기초
- NLP 기초
- 뉴스 수집
- 뉴스 분석
- 관련기사 찾기, 분류



개발환경
google colab
`colab.research.google.com`





Colaboratory에 오신 것을 환영합니다

파일 수정 보기 삽입 런타임 도구 도움말

공유



목차



+ 코드

+ 텍스트

드라이브로 복사

연결

수정 가능



시작하기

데이터 과학



머신러닝



추가 리소스

머신러닝 예시

섹션



Colaboratory란?

줄여서 'Colab'이라고도 하는 Colaboratory를 사용하면 브라우저에서 Python을 작성하고 실행할 수 있습니다. Colab은 다음과 같은 이점을 자랑합니다.

- 구성이 필요하지 않음
- GPU 무료 액세스
- 간편한 공유

학생이든, 데이터 과학자든, AI 연구원이든 Colab으로 업무를 더욱 간편하게 처리할 수 있습니다. [Colab 소개 영상](#)에서 자세한 내용을 확인하거나 아래에서 시작해 보세요.

시작하기

지금 읽고 계신 문서는 정적 웹페이지가 아니라 코드를 작성하고 실행할 수 있는 대화형 환경인 **Colab** 메모장입니다. 예를 들어 다음은 값을 계산하여 변수로 저장하고 결과를 출력하는 간단한 Python 스크립트가 포함된 코드 셀입니다.

```
[ ] seconds_in_a_day = 24 * 60 * 60
    seconds_in_a_day
```

86400

Class 1 Python 기초



Class 1. python 기초

- 변수
- 배열(list, numpy, pandas)
- tuple
- dictionary
- 제어문(if, while, for)

Class 2 NLP 기초



Class 2. NLP 기초

- Word embedding
 - Sparse representation
 - Dense representation
- String tokenizer
 - Space tokenizer
 - Word tokenizer
 - 형태소 분석기
- 명사 추출 분석
- 단어 관계 시각화

Class 3

뉴스 수집



Class 3. 뉴스 수집

- 검색 api 신청
- Sample code 실행
- 검색 옵션 변경(display, start, sort)
- Pandas를 이용하여 읽기
- 뉴스 수집 함수 작성
- 뉴스 수집

요청 변수	타입	필수 여부	기본값	설명
query	string	Y	-	검색을 원하는 문자열로서 UTF-8로 인코딩한다.
display	integer	N	10(기본값), 100(최대)	검색 결과 출력 건수 지정
start	integer	N	1(기본값), 1000(최대)	검색 시작 위치로 최대 1000까지 가능
sort	string	N	sim, date(기본값)	정렬 옵션: sim (유사도순), date (날짜순)

Class 4

뉴스 분석



Class 4. 뉴스 분석

- 뉴스 불러오기
- 형태소(명사) 추출
- word2vec 모델 만들기
- word2vec 모델 테스트
- wordcloud를 통한 시각화

Class 5

관련기사 찾기, 분류



Class 5. 관련기사 찾기, 분류

- Doc2vec 모델 만들기
- Doc2vec 모델을 이용한 관련기사 찾기
- 휴대폰관련뉴스 분류기 만들기
 - Train data pre processing
 - Test data pre processing
- 휴대폰분류기 활용

End of Document

다음주에 만나요...

