

기말 프로젝트

- 데이터 공유 사이트 만들기
 - 조건에 맞도록 사이트 구성
 - HTML, CSS, JS 를 이용하여 기능 구현
 - 수업시간 내용 및 개인 공부한 내용 추가 허용
 - 차주 수업까지 팀원 등록한 팀에 한하여 팀프로젝트 허용

I. 데이터 공유 사이트 목표

■ 교육의 場

- 학생들에게 **실습** 가능한 데이터셋을 제공
- 오픈 SW를 통해 자신의 방법과 다른 사람의 방법과 **경쟁**하며 성장

■ 취업의 場

- 산업 현장에서 얻은 데이터셋을 이용한 **실무 경험** 제공
- OSS Resume와 연계한 **산학 연계** 과제 수행 혹은 **인턴쉽** 진행

■ 진학의 場

- **대학원**에서 연구하는 데이터셋을 공유하여 관심있는 학생을 모집
- 최근 **연구적**으로 이슈가 되고 있는 데이터셋을 체험하는 공간 제공



II. 데이터 공유 사이트 구성

상명 Data Wiki

- 현재의 Data Wiki
 - 데이터셋 확보를 위한 기본 틀과 대표적인 60개의 데이터셋 보유

항목	내용
ID(name)	Dataset명
Description	Dataset 설명
Link	URL
Category	Dataset 카테고리
Field	Dataset 분야
Task	Dataset의 목적
Instance	Dataset내의 data 수
Num	Index
Tutorial	Dataset 관련 사이트 (일부)

- 추진 계획
 - 기존 데이터 공유 사이트와의 차별성 확보를 위한 콘텐츠 확보
 - 데이터셋을 이용한 학습이 가능한 사이트
 - 데이터셋의 목적성을 제공하는 사이트
 - 데이터셋에 대해 친절한 설명이 가능한 사이트

II. 데이터 공유 사이트 구성

상명 Data Wiki



- Data Share Hub
 - 데이터 분석을 위한 기초 자료를 제공
 - 기존 교육자료를 활용한 교육 콘텐츠
 - TensorFlow
 - Google 사의 대표 ML 라이브러리
 - 라이브러리 활용을 위한 예제 코드 제공
 - Pytorch
 - Facebook에서 제공하는 ML 오픈SW
 - 라이브러리 활용을 위한 콘텐츠 제공
- 각 데이터셋 별 Open SW 연계
- 설명이 적은 데이터셋 Data Competition 진행 및 공유 자료 확보



초보자용

사용자 친화적 Keras Sequential API로 시작하는 것이 가장 좋습니다. 구성요소를 함께 연결하여 모델을 빌드해 보세요. 이러한 튜토리얼을 완료한 후 [Keras 가이드](#)를 읽어보세요.

초급자용 빠른 시작

이 'Hello, World!' 메모장에는 Keras Sequential API 및 `model.fit`가 표시됩니다.

Keras 기본사항

이 메모장 컬렉션은 Keras를 사용한 기본적인 머신러닝 작업을 보여줍니다.

데이터 로드

이러한 튜토리얼에서는 `tf.data`를 사용하여 다양한 데이터 형식을 로드하고 입력 파이프라인을 빌드합니다.



파이토치(PyTorch) 레시피 [+]

파이토치(PyTorch) 시작하기 [-]

파이토치(PyTorch) 기본 익히기

빠른 시작(Quickstart)

텐서(Tensor)

Dataset과 DataLoader

변형(Transform)

신경망 모델 구성하기

`torch.autograd`를 사용한 자동 미분

모델 매개변수 최적화하기

모델 저장하고 불러오기

II. 데이터 공유 사이트 구성

상명 Data Wiki

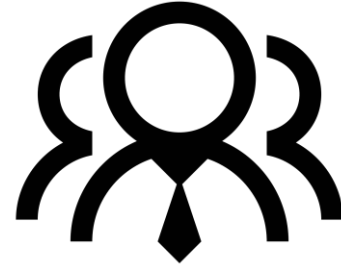


- Data Colosseum
 - 교내 경진대회를 통한 데이터셋 활용
 - 연 2회 Data Wiki 내 데이터셋을 이용한 경진대회 개최
 - 활용사례가 적은 최신 데이터셋에 대한 오픈 SW 자료 확보
 - 경진대회 수상자는 상금과 6개월 간 Data Manager 의무 수행
 - 수업 과제로 Data Wiki 활용
 - 데이터 및 인공지능 수업에서 활용
 - Data Wiki를 활용한 교수법 연구를 통한 활용방법 공유

II. 데이터 공유 사이트 구성

상명 Data Wiki

- Data Manager
 - Data Competition의 우승자에게 부여되는 권한
 - 우승한 데이터셋에 대한 설명 및 수정권한을 부여 받음
 - Data Wiki를 통하여 수상자와 질의 응답이 가능한 통로를 만듦



III. 사이트 구성

■ 상명 Data Wiki

- 사이트 구성



Home

Data Wiki

Share Hub

Colosseum

Forum



■ 사이트 관련 사진 출력 (5장, 5초마다 슬라이드 됨)

■ 공지사항

- 추가 데이터셋
- 경진대회

■ 인기 많은 글

- Data Wiki 중 상위 3개
- Colosseum 중 상위 3개
- Forum 중 상위 3개



III. 사이트 구성

■ 상명 Data Wiki

- 사이트 구성



Home

Data Wiki

Share Hub

Colosseum

Forum



■ Data Wiki

- 연구용 데이터셋 및 공공데이터 검색 및 링크 제공
- 이름순, Category순, task순, tutorial 유/무로 정렬 가능
- 데이터셋은 다음과 연계됨
 - Data Manager (1인)
 - 데이터셋 설명 (관리자 및 Data Manager가 작성 및 수정)
 - 해당 데이터셋을 활용한 Open SW
 - 데이터셋을 설명한 Blog 링크
 - 데이터셋 Tutorial Code

III. 사이트 구성

■ 상명 Data Wiki

- 사이트 구성



Home

Data Wiki

Share Hub

Colosseum

Forum



정렬기준

데이터 리스트

데이터 이름

- 데이터 분류
- 데이터의 수 (instances)
- 데이터의 목적 (task)
- 데이터 설명

데이터 링크 (iframe 연결) / 튜토리얼 링크 (iframe 연결)

-1

-2

-3

-4

-5

-.

-.

-.

-60

III. 사이트 구성

■ 상명 Data Wiki

- 사이트 구성



Home

Data Wiki

Share Hub

Colosseum

Forum



■ Share Hub

- 데이터를 다운 받은 사람이 활용하는 사이트
- 데이터셋의 유형에 따른 예제코드 정리
 - 자연어 (csv, text) / 이미지 (jpg, png) / 동영상 (avi, mp3) / 음성 (wav)
 - <https://www.tensorflow.org/tutorials?hl=ko>
 - <https://tutorials.pytorch.kr/>
 - 위 사이트와 연계
- 데이터 분석에 필요한 배경지식 & 실습자료 제공

III. 사이트 구성

■ 상명 Data Wiki

- 사이트 구성



Home

Data Wiki

Share Hub

Colosseum

Forum



■ Forum

- 사이트 이용자들의 정보 공유 공간
- 데나무숲 : Data Wiki의 Data Manager 질의 응답 게시판(자유롭게)