

1 번



Trusted

Python 3 O

Edit View Insert Cell Kernel Widgets Help
[File] [New] [Cell] [Run] [Kernel] [Widgets] [Help]

Tabular 문제 - 당뇨병 진행률 예측(20점)

문제 개요

다음에 제시되는 당뇨병 데이터를 분석하여 1년 후 당뇨병 진행률을 예측하는 모델을 만드세요.
(Professional 자격인증 문제는 총 3개입니다. 이번 문제는 3개 중 1번 문제입니다.)

제공 자원

문제당 CPU 2CORE, Memory 16G, GPU 1 개
제공 GPU 사양

- Tensor 코어 : 320
- CUDA 코어 : 2,560
- VRAM : 16GB GDDR6
- 대역폭 : 320GB/s
[상세 사양 참고](#)

SW 사양

- Python 3.6.9
- Tensorflow 2.4.2
- OpenJDK 11

달성 목표

test set의 데이터로 RMSE 70이하가 나오도록 모델을 만드세요.
test의 label 값은 별도로 알려드리지 않습니다.

데이터 설명

데이터 위치 및 폴더 구조

- 문제 파일과 동일한 곳에 위치
- 01_train.csv: 훈련 데이터셋입니다. 훈련 데이터셋으로 모델을 만드세요.
- 01_test_x.csv: 테스트 데이터셋입니다. 테스트 데이터셋에는 라벨(Y)값이 제외되어 있습니다.

Label 정보

jupyter 01073001524_1 Last Checkpoint: 1분 전 (autosaved)



Trusted | Python 3 O

Edit View Insert Cell Kernel Widgets Help
File New Cell Kernel Help
Run Cell Markdown Cell

• OpenJDK 11

달성 목표

test set의 데이터로 RMSE 70이하가 나오도록 모델을 만드세요.
test의 label 값은 별도로 알려드리지 않습니다.

데이터 설명

데이터 위치 및 폴더 구조

- 문제 파일과 동일한 곳에 위치
- 01_train.csv: 훈련 데이터셋입니다. 훈련 데이터셋으로 모델을 만드세요.
- 01_test_x.csv: 테스트 데이터셋입니다. 테스트 데이터셋에는 라벨(Y)값이 제외되어 있습니다.

Label 정보

- Y: 1년 후 당뇨병 진행률

Data 정보

당뇨병 환자 442명에 대한 데이터입니다.

- AGE: 숫자로 된 나이 정보
- SEX: 1 또는 2로 된 성별 정보
- BMI: 숫자로 된 체질량 지수
- BP: 숫자로 된 평균 혈압
- S1~S6: 6가지의 혈청 측정 정보

데이터는 csv 파일 2개가 제공됩니다.

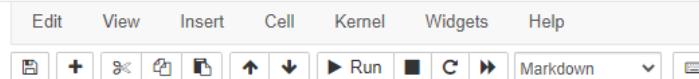
위 문제 해결 가이드를 참고하여 훈련 데이터셋으로 모델을 만들고 테스트 데이터셋으로 라벨(Y)값을 예측하여 csv파일로 저장하세요

- 01_train.csv는 훈련 데이터셋입니다.
- 01_test_x.csv는 테스트 데이터셋입니다. 테스트 데이터셋에는 라벨(Y)값이 제외되어 있습니다.

데이터 예시

해당 예시 데이터는, 예시용 데이터로, 학습용 데이터셋에 포함되어 있지 않습니다.

jupyter 01073001524_1 Last Checkpoint: 1분 전 (autosaved)



Trusted | Python 3

데이터 예시

해당 예시 데이터는, 예시용 데이터로, 학습용 데이터셋에 포함되어 있지 않습니다.

AGE	SEX	BMI	BP	S1	S2	S3	S4	S5	S6	Y
64	1	25.3	83	225	156.6	46	5	4.7185	84	49
63	1	23	73	161	97.8	54	3	3.8286	91	200
61	1	19.9	75	175	108.6	54	3	4.1271	72	99
61	1	22.6	89	139	64.8	61	2	4.1897	68	182
60	1	25.3	84	198	131.4	40	5	4.8903	89	199

문제 해결 가이드

(1) 훈련 데이터셋을 이용하여 1년 후 당뇨병 진행률 예측하는 모델을 만드세요.

- 필요 시 추가로 함수를 만들어거나 cell을 만들어도 됩니다.
- 성능이 가장 좋은 모델을 저장하세요.
- 파일명: 본인핸드폰번호_1.h5 (또는 다른 확장자의 모델 파일)

(2) 라벨(Y)값이 제외된 테스트 데이터셋 파일을 읽어들여 예측값에 대한 컬럼을 추가하여 csv로 저장하세요.

- 파일명: 본인핸드폰번호_1.csv
- 테스트 데이터 파일을 읽고 마지막 컬럼 뒤에 Y 컬럼 생성 후, 모델 예측하여 예측값을 채워주세요.
- 예측값을 채운 테스트 데이터 형상과 훈련 데이터 형상과 동일하게 됩니다.(똑같은 컬럼명 보이게 됨)
- 예측값을 채운 테스트 데이터를 '본인핸드폰번호_1.csv' 파일로 저장하세요.

(3) 본 문제지를 답안코드 작성된 상태로 저장하세요.

- 파일명: 본인핸드폰번호_1.ipynb (본 문제지 상단의 저장버튼을 눌러 저장)

제출할 파일 목록

- 본인핸드폰번호_1.h5 (또는 다른 확장자의 모델 파일)
- 본인핸드폰번호_1.csv
- 본인핸드폰번호_1.ipynb (본 문제지 상단의 저장버튼을 눌러 저장)

※ 예시: 01012345678_1.h5

jupyter 01073001524_1 Last Checkpoint: 1분 전 (autosaved)



Edit View Insert Cell Kernel Widgets Help

Trusted Python 3 O

File + % Cell Run C Markdown

- 본인핸드폰번호_1.ipynb (본 문제지 상단의 저장버튼을 눌러 저장)
※ 예시: 01012345678_1.h5

채점 진행 방식 안내

1번 문항 배점 : 20 점

- 본인핸드폰번호_1.ipynb 실행 가능 여부 확인
- 생성된 모델과 코드의 동일성 여부 확인
- 본인핸드폰번호_1.csv로 제출된 예측 데이터를 정답지와 대조하여, RMSE 70 이하 달성

안내사항

- 답안 작성 시 TensorFlow, Pytorch 등 프레임워크 제한 없이 사용 가능합니다.
- 필요한 라이브 라이브러리를 설치하여 사용하세요.
- 모델 저장시 본인이 사용하는 프레임워크의 모델 저장 확장자를 사용하셔도 됩니다.(h5, pth, pt,.pkl, pickle 등)

문제 풀기 전 아래 코드 실행

GPU 할당을 위해 아래 코드를 실행하세요.

```
In [2]: import tensorflow as tf
gpus = tf.config.list_physical_devices('GPU')
if gpus:
    try:
        for gpu in gpus:
            tf.config.experimental.set_memory_growth(gpu, True)
        logical_gpus = tf.config.experimental.list_logical_devices('GPU')
        print(len(gpus), 'Physical GPUs,', len(logical_gpus), 'Logical GPUs')
    except RuntimeError as e:
        print(e)
```

1 Physical GPUs, 1 Logical GPUs

답안 작성

필요한 cell을 추가하여 작성하세요.

2번



Edit View Insert Cell Kernel Widgets Help
[File] [New] [Cell] [Run] [Kernel] [Widgets] [Help] Trusted Python 3 O

Text 문제 - 기가지니 발화 데이터 카테고리 분류(40점)

문제 개요

기가지니 발화 데이터에 대해 어느 카테고리에 속하는지 분류하는 모델을 만드세요.
(Professional 문제는 총 3개입니다. 이번 문제는 3개 중 2번 문제입니다.)

제공 자원

문제당 CPU 2CORE, Memory 16G, GPU 1 개

제공 GPU 사양

- Tensor 코어 : 320
- CUDA 코어 : 2,560
- VRAM : 16GB GDDR6
- 대역폭 : 320GB/s

[상세 사양 참고](#)

SW 사양

- Python 3.6.9
- Tensorflow 2.4.2
- OpenJDK 11

달성 목표

test set의 데이터로 정확도가 70% 이상을 달성하도록 모델을 만드세요.
test의 label 값은 별도로 알려드리지 않습니다.

데이터 설명

데이터 위치 및 폴더 구조

- 문제 파일과 동일한 곳에 위치
- 02_train.csv: 훈련 데이터셋입니다. 훈련 데이터셋으로 모델을 만드세요.
- 02_test_x.csv: 테스트 데이터셋입니다. 테스트 데이터셋에는 라벨(Y)값이 제외되어 있습니다.



jupyter 01073001524_2 Last Checkpoint: 3시간 전 (unsaved changes)



Trusted

Python 3

데이터 설명

데이터 위치 및 폴더 구조

- 문제 파일과 동일한 곳에 위치
- 02_train.csv: 훈련 데이터셋입니다. 훈련 데이터셋으로 모델을 만드세요.
- 02_test_x.csv: 테스트 데이터셋입니다. 테스트 데이터셋에는 라벨(Y)값이 제외되어 있습니다.

Label 정보 : act_category

- 제어
- 미디어
- 채팅
- 지능형대화
- 생활정보

Data 정보

기가지니에서 수집된 발화 데이터입니다.

- reqmsg : 수집된 음성 메시지

데이터는 csv 파일 2개가 제공됩니다.

위 문제 해결 가이드를 참고하여 훈련 데이터셋으로 모델을 만들고 테스트 데이터셋으로 라벨(Y)값을 예측하여 csv파일로 저장하세요

- 02_train.csv는 훈련 데이터셋입니다.
- 02_test_x.csv는 테스트 데이터셋입니다. 테스트 데이터셋에는 라벨(Y)값이 제외되어 있습니다.

데이터 예시

해당 예시 데이터는, 예시용 데이터로, 학습용 데이터셋에 포함되어 있지 않습니다.

reqmsg	act_category
영화 콜드 우먼 보여줘	미디어
tv 켜줘	제어
그냥 한번 불러봤어	채팅
영어 상어 tv youtube 틀어줘	지능형대화
미세먼지	생활정보



Edit View Insert Cell Kernel Widgets Help
[File] [New] [%] [Cell] [Run] [Kernel] [Widgets] [Help] [Markdown]

Trusted | Python 3 O

문제 해결 가이드

(1) 훈련 데이터셋을 이용하여 발화 데이터 카테고리 분류하는 모델을 만드세요.

- 필요 시 추가로 함수를 만들거나 cell을 만들어도 됩니다.
- 성능이 가장 좋은 모델을 저장하세요.
- 파일명: 본인핸드폰번호_2.h5 (또는 다른 확장자의 모델 파일)

(2) 라벨(Y)값이 제외된 테스트 데이터 파일을 읽어들여 예측값에 대한 컬럼을 추가하여 csv로 저장하세요.

- 파일명: 본인핸드폰번호_2.csv
- 테스트 데이터 파일을 읽고 마지막 컬럼 뒤에 act_category 컬럼 생성 후, 모델 예측하여 예측값을 채워주세요.
- 예측값을 채운 테스트 데이터 형상과 훈련 데이터 형상과 동일하게 됩니다.(똑같은 컬럼명 보이게 됨)
- 예측값을 채운 테스트 데이터를 '본인핸드폰번호_2.csv' 파일로 저장하세요.

(3) 본 문제지를 답안코드 작성된 상태로 저장하세요.

- 파일명: 본인핸드폰번호_2.ipynb (본 문제지 상단의 저장버튼을 눌러 저장)

제출할 파일 목록

- 본인핸드폰번호_2.h5 (또는 다른 확장자의 모델 파일)
- 본인핸드폰번호_2.csv
- 본인핸드폰번호_2.ipynb (본 문제지 상단의 저장버튼을 눌러 저장)
※ 예시: 01012345678_2.h5

채점 진행 방식 안내

2번 문항 배점 : 40 점

- 본인핸드폰번호_2.ipynb 실행 가능 여부 확인
- 생성된 모델과 코드의 동일성 여부 확인
- 본인핸드폰번호_2.csv로 제출된 예측데이터를 정답지와 대조하여, 70% 이상 정확도 달성

안내사항

- 답안 작성 시 TensorFlow, Pytorch 등 프레임워크 제한 없이 사용 가능합니다.
- 필요한 라이브러리는 설치하여 사용하세요.
- 모델 저장시 본인이 사용하는 프레임워크의 모델 저장 확장자를 사용하셔도 됩니다.(h5, pth, pt,.pkl, pickle 등)

3번

jupyter 01073001524_3 Last Checkpoint: 7분 전 (autosaved)

No kernel Trusted Python 3

[KT] AICE Professional_12.1(목) 15:00-18:00

시험 시간: 15시 00분 - 18시 00분 (180분)
남은 시간: 0분 0초

임... 최...

시험이 종료되었습니다.

AICE Professional 자격인증 문제-3

Image 문제 - 업무용 데스크 점검(40점)

문제 개요

이미지 데이터를 분석하여 업무용 데스크 보안점검을 위한 모델을 만드세요.
업무용 데스크 이미지를 분석하여 서류/도서 등 문구류가 없이 깨끗이 정리되어 있으면 clean, 그렇지 않으면 not clean으로 분류합니다.
(Professional 자격인증 문제는 총 3개입니다. 이번 문제는 3개 중 3번 문제입니다.)

제공 자원

문제당 CPU 2CORE, Memory 16G, GPU 1 개
제공 GPU 사양

- Tensor 코어 : 320
- CUDA 코어 : 2,560
- VRAM : 16GB GDDR6
- 대역폭 : 320GB/s

[상세 사양 참고](#)

SW 사양

- Python 3.6.9
- Tensorflow 2.4.2
- OpenJDK 11

달성 목표

jupyter 01073001524_3 Last Checkpoint: 몇 초 전 (autosaved)



xywh

1/10

Edit View Insert Cell Kernel Widgets Help

Trusted Python 3

- TensorFlow 2.4.2
- OpenJDK 11

달성 목표

test set의 데이터로 정확도 80% 이상을 달성하도록 모델을 만드세요.
test의 label 값은 별도로 알려드리지 않습니다.

데이터 설명

데이터 위치 및 폴더 구조

- 문제 파일과 동일한 곳에 위치
- 아래 data 정보에 기입된 파일구조와 같이 하위 폴더로 구성되어있음

Label 정보

- clean
- not_clean

Data 정보

- train용 clean desk image 2341장 : "03_clean_desk/train/clean"
- train용 not clean desk image 2339장 : "03_clean_desk/train/not_clean"
- test용 desk image 630장 : "03_clean_desk/test"

문제 해결 가이드

(1) train 폴더의 이미지를 이용하여 clean과 not clean 데스크 분류하는 모델을 만드세요.]

- 필요 시 추가로 함수를 만들거나 cell을 만들어도 됩니다.
- 성능이 가장 좋은 모델을 저장하세요.
- 파일명: 본인핸드폰번호_3.h5 (또는 다른 확장자의 모델 파일)

(2) csv 파일을 생성하고 test image 파일명과 예측값에 대해 아래와 같이 csv로 저장하세요.

- 파일명: 본인핸드폰번호_3.csv
- 1번째 column을 filename으로 하여, test image 데이터셋 안에 있는 파일명들을 기입해주세요.
- 2번째 column을 label으로 하여, 'clean' / 'not_clean'의 값에 대한 테스트 이미지 파일의 예측값을 기입해주세요. (아래 예시 참고)

filename	label
image_0_999684536126.jpg	clean



jupyter 01073001524_3 Last Checkpoint: 몇 초 전 (autosaved)



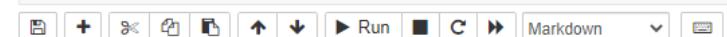
xywh

1/10



Edit View Insert Cell Kernel Widgets Help

Trusted Python 3 O



- 필요 시 추가로 함수를 만들거나 cell을 만들어도 됩니다.
- 성능이 가장 좋은 모델을 저장하세요.
- 파일명: 본인핸드폰번호_3.h5 (또는 다른 확장자의 모델 파일)

(2) csv 파일을 생성하고 test image 파일명과 예측값에 대해 아래와 같이 csv로 저장하세요.

- 파일명: 본인핸드폰번호_3.csv
- 1번째 column을 filename으로 하여, test image 데이터셋 안에 있는 파일명들을 기입해주세요.
- 2번째 column을 label으로 하여, 'clean' / 'not_clean'의 값에 대한 테스트 이미지 파일의 예측값을 기입해주세요. (아래 예시 참고)

filename	label
image_0_999684536126.jpg	clean

(3) 본 문제지를 답안코드 작성된 상태로 저장하세요.

- 파일명: 본인핸드폰번호_3.ipynb (본 문제지 상단의 저장버튼을 눌러 저장)

제출할 파일 목록

- 본인핸드폰번호_3.h5 (또는 다른 확장자의 모델 파일)
- 본인핸드폰번호_3.csv
- 본인핸드폰번호_3.ipynb (본 문제지 상단의 저장버튼을 눌러 저장)
※ 예시: 01012345678_3.h5

채점 진행 방식 안내

3번 문항 배점 : 40 점

- 본인핸드폰번호_3.ipynb 실행 가능 여부 확인
- 생성된 모델의 동일성 여부 확인
- 본인핸드폰번호_3.csv로 제출된 label 데이터를 정답지와 대조하여, 80% 이상 정확도 달성

안내사항

- 답안 작성 시 TensorFlow, Pytorch 등 프레임워크 제한 없이 사용 가능합니다.
- 필요한 라이브러리는 설치하여 사용하세요.
- 모델 저장시 본인이 사용하는 프레임워크의 모델 저장 확장자를 사용하셔도 됩니다.(h5, pth, pt,.pkl, pickle 등)