

The background is a dark blue space filled with glowing binary code (0s and 1s) that appears to be streaming or orbiting around a central, bright, glowing sphere. The sphere has a lens flare effect. The overall aesthetic is high-tech and digital.

Workshop

2021

팀별 수행(9:30~16:10)

컨퍼런스홀 입장 및 팀별 발표(16:30~)

우수팀 수상(대상/금상/은상/동상) (17:15~)

총 4팀

- 1조: 신예성, 양영준, 유정근
- 2조: 김도영, 하진아, 홍영표
- 3조: 윤일중, 윤진선, 이현규
- 4조: 김정아, 이이삭, 최진현

공통 문제 2개 (필수) +
선택 문제(최소 1문제 이상)

공통 문제 1

- IoT 장비의 네트워크 로그(출처: Kaggle)
- **Target: normality**
 - 0- normal
 - 1- wrong setup
 - 2- ddos
 - 3- Data type probing (ultrasonic sensor was used so in data type probing mostly string values are sent to the server)
 - 4 - scan attack
 - 5 - man in the middle
- **Preprocessed_data.csv를 사용해서 normality를 Target으로 예측하도록 모델링**

- **CIFAR-10 데이터셋으로 이미지 분류**

- 60000장의 32X32 픽셀 이미지: 50000개 Training, 10000개 Testing
- 3개 채널의 컬러 이미지
- Target: 10개의 범주, 0: airplane, 1: automobile, 2: bird, 3: cat, 4: deer, 5: dog, 6: frog, 7: horse, 8: ship, 9: truck

- **데이터 다운로드**

- `from keras.datasets import cifar10`
- `(x_train, y_train), (x_test, y_test) = cifar10.load_data()`

공통 문제 2

- CIFAR-10 데이터셋으로 이미지 분류

- 이미지 확인

- `from google.colab.patches import cv2_imshow`
 - `for i in range(10):`
 - `cv2_imshow(x_train[i])`

- Fashion_mnist 코드를 참고하며, 대신, 흑백 이미지 대신 컬러이미지인 것은

`input_shape`에 (32,32,3)으로 입력

– 예:

```
model.add(Conv2D(filters=32, kernel_size=(3, 3),  
padding='same', input_shape=(32,32,3), activation  
='relu'))
```

선택 문제 1

- **한글 영화 리뷰의 감성 분석**
 - **수업 예제를 활용하여, 영화 리뷰에 대해 긍부정 분류 모델을 만드세요**
 - Training에서 최소 10000개 이상의 리뷰로 모델링
 - Testing에서 최소 2000개 이상의 리뷰로 성능을 측정
 - **Tensorflow를 활용해서 딥러닝 모델을 적용하고, sklearn의 분류 모델 중 하나 (MLPClassifier 제외)를 선택해서 적용한 후, 두 모델의 성능을 비교**
- **데이터셋**
 - Training: rating_train.txt
 - Testing: rating_test.txt

선택 문제 2

- 항공기 엔진 예측 정비 (참고: MS Azure Gallery)
 - *This predictive maintenance template focuses on the techniques used to predict when an in-service machine will fail, so that maintenance can be planned in advance.*
 - 엔진이 특정 일수(예를 들어 30일) 내에 고장이 발생하는지를 예측
 - 가능한 모형을 사용하여 모델링하세요
 - 데이터셋
 - Training: pm_train.csv
 - Testing: pm_test.csv

선택 문제 2

- 장비 예측 정비 (참고: MS Azure Gallery)

Index	Data fields	Type	Descriptions
1	id	Integer	aircraft engine identifier, range [1, 100]
2	cycle	Integer	time, in cycles
3	setting1	Double	operational setting 1
4	setting2	Double	operational setting 2
5	setting3	Double	operational setting 3
6	s1	Double	sensor measurement 1
7	s2	Double	sensor measurement 2
...	...		
26	s21	Double	sensor measurement 21

- ttf 변수: Time To Failure, ttf 변수를 사용하여, 30일 이하인 경우 1(고장 발생), 그 외의 경우 0(정상 작동)으로 처리하여 Target으로 사용하세요
 - 예: `df_train['ttf'].apply(lambda x: 1 if x <= 30 else 0)`

선택 문제 3

- 부도 예측
 - 신용점수, 부동산, 근무기간, 대출금액 등을 바탕으로 부도 발생 여부를 기록한 **mortdefault2009.csv** 파일을 읽어서 아래의 작업을 하세요
 - Data partitioning- 7:3으로 training, testing
 - Data scaling- MinMax 또는 Standard Scaling
 - 분류 모델링, Testing셋으로 정분류율 평가
 - Target: default
 - Scaling된 creditScore, houseAge, yearsEmploy, ccDebt로 군집분석 수행

- ✓ 16:30 부터 컨퍼런스홀(게더타운 內)로 이동해서 발표 시작
- ✓ 팀별 발표 10~15분 + 피드백 및 Q&A 5분
 - ✓ 발표 내용은 주어진 템플릿을 참고 / 다른 양식을 사용하셔도 무방
- ✓ 발표 후 시상: 대상 / 금상 / 은상 / 동상

Q&A