

# Project : Character classification Challenge

## I. 프로젝트 개요

- English Character & Number with various fonts **분류 문제**
  - 영문 및 숫자 인쇄 데이터 분류하기
  - 총 52개의 class
  - 각 class는 같은 character, 서로 다른 font 의 활자들로 구성됨.
- 자유롭게 모델 구성
  - Accuracy 를 높게 하는 것 목표
  - Dense, CNN, RNN , Dropout, Batch\_norm 등 모두 적용 가능
- 제한
  - Colab GPU 환경 기준 **10분 이내 학습 완료**
  - 구성된 **model 과 학습기법에 대한 근거를 보고서에 설명**
- 주어진 dataset class를 활용하여 dataloader 구성



# Project : Character classification Challenge

## II. 데이터셋

### • 영문/숫자

- Label 0~9: 숫자(digit)
- Label 10~52: 영문자 (letter)
  - 대/소문자 구별이 어려운 ["c", "k", "l", "O", "p", "s", "v", "w", "x", "z"] class 는 제외
- Torchvision 에서 제공하지 않으며, 직접 다운 받아 사용해야 함. ([train](#) / [valid](#))

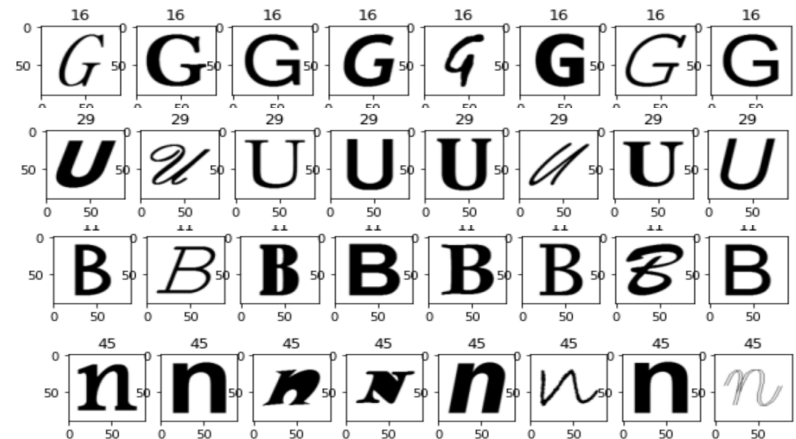
### • Data shape

- 1\*90\*90 (C\*H\*W)

### • 데이터

- Train: 37,232
- Valid : 7,800
- Test : 7,800 제공 X
- Totally balanced (각 class 별 data 개수 모두 같음.)

```
label_dict = {'0': 0, '1': 1, '2': 2, '3': 3, '4': 4, '5': 5, '6': 6, '7': 7, '8': 8, '9': 9,
'A': 10, 'B': 11, 'C': 12, 'D': 13, 'E': 14, 'F': 15, 'G': 16, 'H': 17,
'I': 18, 'J': 19, 'K': 20, 'L': 21, 'M': 22, 'N': 23, 'P': 24, 'Q': 25,
'R': 26, 'S': 27, 'T': 28, 'U': 29, 'V': 30, 'W': 31, 'X': 32, 'Y': 33,
'Z': 34, 'a': 35, 'b': 36, 'd': 37, 'e': 38, 'f': 39, 'g': 40, 'h': 41,
'i': 42, 'j': 43, 'm': 44, 'n': 45, 'o': 46, 'q': 47, 'r': 48, 't': 49, 'u': 50, 'y': 51}
```



# Project : EMNIST Challenge

## III. 프로젝트 일정 / 제한 사항

### • 제출 기한

- 12월 14일 (화) 자정(23:59) / 지각 제출에 대한 감점 있음
- 지각 제출 기한 : 12월 19일 (일) 자정 (23:59) / 이후 제출 불가

### • 제한 사항

- Train 시간 Google Colab GPU 기준 10분 이내에 완료
  - Train 시간 10분 초과 소요에 대한 감점 있음
- Pre-trained model 사용 불가
- Python 3.5+, Pytorch 1.2+ 이상
- 라이브러리 제한
  - Python 표준 라이브러리, Pytorch 제공 라이브러리, Numpy, Matplotlib 등 학습을 위한 라이브러리 외 학습성능 향상을 위한 라이브러리 사용시 조교에게 문의



# Project : EMNIST Challenge

## IV. 제출 형태 / 채점 기준

### • 제출 형태: PJ\_학번.zip

1. Train 에 사용된 파일 : train\_학번.py
2. Train 에서 저장한 model : 학번.pth
3. Test 파일 : Test\_학번.py

train 을 통해 저장한 model.pth 를 불러와서 동작하도록 작성 (Validation set 을 test 하도록 작성)

### 4. 보고서 : PJ\_학번.pdf

1. 과제 목표
2. 배경 이론
3. 과제 수행 방법
4. 결과 및 토의
5. 참고 문헌

### • 채점기준

- Model Accuracy (순위) : 50%
- 프로젝트 보고서 : 50%

