

# week04\_ 정회수

## 실험 결과

	CNN1	CNN2	VGG
Epoch	46	46	29
Best Score	0.4041	0.4020	0.7884

## 가장 좋은 모델은?

과제 설명에도 적혀있듯이 VGG는 이미지넷으로 1000개의 클래스로 사전 훈련을 거친 상태이다. 반면에 우리의 커스텀 데이터셋은 배치 사이즈를 32개로 설정 했을때 152개의 배치가 나오므로 대략 4864개임을 알 수 있고, 따라서 학습 데이터셋의 크기부터 현저히 차이가 난다는 것을 알 수 있다. 훈련을 돌려본 결과 첫 에포크부터 0.5 이상의 validation accuracy를 보여주면서 우리 커스텀 모델의 성능을 압도했다. (기울어진 운동장)

## CNN1과 CNN2의 차별점?

CNN2는 기존의 CNN1의 FCN 레이어 부분에 ReLU 및 Dropout을 추가하여 데이터에 비선형성을 추가해준 것인데, 사실 이론적으로는 input 값들과 가중치의 선형 결합으로 이루어지는 초평면에 곡선 형태가 추가되면서 목표값들을 더 잘 표현할 수 있어야하는게 맞다. 근데 CNN의 태생적 한계 때문인지 실험 결과로는 두드러지는 차이를 볼 수가 없었고, 개인적으로 BatchNorm2D도 추가해서 실험을 돌려보았지만 결과는 더 나빴다.

## Pytorch를 사용해야 하는 이유는?

일단 내부적으로 알아서 미분값을 계산해주는 것이 너무 편리하다. 또한, 레이어를 생성만 해주면 나중에 모델의 forward 함수를 설정하는 과정에서 레고 블록 쌓듯이 레이어를 쌓아주는 과정이 너무 편했다.