



8/19(목) 회의록

🕒 작성일시	@2021년 8월 19일 오전 9:06
👤 작성자	하람 이하람
👥 참석자	
🕒 최종 편집일시	@2021년 8월 19일 오후 8:46
📌 회의 유형	일일 회의

👉 학습 내용 공유

1. 강의 내용 중 질문하기 🙋

Q1. 모델을 가져올때 pretrained=True를 주고 가져오는건 특정 task에 맞춰서 학습된건가요?

가령 MNIST를 이용해서 분류문제를 풀도록 pretrained된 vgg모델을 불러와서 개와 고양이 사진으로 돌린다고 하면 마지막에 nn.Linear만 바꿔주면 분류를 알아서 하고 결과도 잘 나오는건가요?

A1. 훨씬 방대한 여러 이미지로 분류를 잘 하게끔 미리 학습시켜서 마지막 레이어만 잘 수정 해주면 결과가 잘 나오게 된다.

Q2. Dataloader의 num_workers개수가 늘어나도 소요시간이 줄지 않는 이유?

A2. core갯수가 같고(colab의 경우 1개) subprocess가 늘어나도 일을 쪼개서 할 뿐 전체 연산 시간은 같습니다.(concurrency만 증가)

과제질문:

encoder에 함수 encode를 배정하는건지 ?

그렇다면 아래에 ['hello'] 하는 식으로 들어오는건 어떻게 처리하는지?

encode('hello') 라면 되겠지만

encode['hello']는 안될텐데?

그리고 encode를 encoder에 할당한다면, 기본 아규먼트는 무엇으로 쓰는지?

2. CNN 논문 리뷰 🤖

EfficientNet

발표자료 :

https://github.com/Barleysack/BoostCampPaperStudy/tree/main/Day7_EfficientNet

모델사이즈가 커지는 것에도 한계가 있다.

한정된 범위 내에서 d, r의 밸런스를 맞추는게 중요하다. (d:depth, r:resolution)

모델의 크기만을 바꾸기 때문에 바탕이 되는 모델을 적절히 선택하는게 중요!

MBConv(Mobile inverted Bottleneck)에 SE(Squeeze and Excitation)를 더한 레이어가 주 구성요소.

3. 마스터클래스 🕶️

앞으로 알아야할 것 들

- MLOPs 도구들
- 당연히 데이터베이스 SQL
- Cloud - AWS, GCP, Azure
- Spark (+ Hadoop)
- Linux + Docker + 쿠버네티스
- 스케줄링 도구들 (쿠브플로우, MLFlow, AirFlow)

오르는 집값에 쫓지말고 주식 공부 시작하기

4. 멘토링 질문

진규

1. 모델을 공부하다 보니까, 어떤 문제를 풀 때 어떤 모델을 사용하는 것이 적합한 지는 어떻게 알까요?? 그냥 최신의 모델만 사용하나요?? 아님 다 써보고 정확도가 가장 높은 것을 쓰나요??

주로 EfficientNet을 가져다 쓰는 경향이 있다. ImageNet SOTA모델도 자주 가져다 쓰고 의료분야에서는 유넷??을 쓰기도 한다.

태호

1. 공부하다가 지칠때 리프레시 노하우..

도메인 선택은 무엇을 기준으로 결정하면 좋을까요?