

Handong Honor Code

- You are responsible for understanding and complying with Handong Honor Code.
- If you copy someone else's coding or homework assignments or use someone else's creative work without any indication, you will get an F grade. Anyone who shares the code with you will also get an F grade.

Copyright Notice

You may not make copies of this and use or distribute it for any purpose.

Jaeyoung Chun | School of Applied Artificial Intelligence | Handong Global University

Final Project : 나의 해방일지

<목표> 딥러닝 수업으로부터 해방하기 위한 해방일지를 작성해주세요 :-)

(처음부터 끝까지 문제를 "먼저" 다 읽어보시고, 문제의 큰 그림을 판단하신 후 시작하는 것이 좋을 듯 합니다)

1 --

1.1 Stage 1

수업 시간에 배운 MNIST 데이터셋을 이용하여 손글씨 숫자 판별을 가능케하는 딥러닝 모델을 구축합니다. 아주 기본적인 형태의 Fully Connected Network 을 구성하고 이를 base model 이라 부르겠습니다.

Training set 과 validation set 을 사용하여 훈련 및 검증을 하고 loss 및 accuracy 그래프를 그립니다.

Validation set 으로 검증을 하면서 accuracy 를 측정하고 95% 이상의 accuracy 가 측정 되면, 만들어진 모델을 파일로 저장하고 (*.pth), 다음 stage 로 진행합니다.

1.2 Stage 2

- Test set 을 불러와서 최종 테스트를 합니다.
- Test set 을 사용해서 accuracy 가 어느 정도 나오는지 측정합니다.
- 만약 만족할 만한 accuracy 가 나오지 않는다면, 다시 Stage 1 로 가서 모델을 수정해도 좋습니다.

1.3 Stage 3

- 손글씨로 숫자 0-9 까지를 종이에 써서 스캐너, 사진기 또는 핸드폰 카메라를 이용하여 손글씨를 디지털화합니다 (.jpg or .png).
- 디지털화된 손글씨 이미지를, 앞서 구성한 딥러닝 모델에 입력하여 inference 을 수행하고, 자신의 모델이 정확한 답을 예측했는지 아닌지를 확인합니다. 예를 들어, 종이에 손글씨로 5 라고 쓴 것을 정말로 5 라고 판단하는지를 수행해보고, 만약 정확한 예측을 내렸다면 몇 퍼센트의 확률로 예측하는지를 확인해봅니다.
- 모델에 입력하기 전/후의 손글씨를 matplotlib/Pillow/OpenCV 와 같은 패키지를 사용해서 화면에 표시하고, 모델이 판단한 예측값과 퍼센트 확률도 같이 화면에 표시해주세요.

1.3.1 Preprocessing

디지털화 된 손글씨 이미지를 자신이 만든 모델이 요구하는 입력 포맷으로 어떻게 변환시켰는지 “각각” 의 모든 전처리 스텝 (preprocessing steps) 들을 상세히 해방일지에 기록하세요. 예를 들면, 해상도를 낮추기 위해서 어떤 프로그램 또는 어떤 코드를 사용했는지, 만약 프로그램 (e.g. PhotoShop) 을 사용했다면 프로그램 안에 있는 어떤 기능을 어떤 순서로 어떻게 사용했는지 구체적으로 기록하세요.

- Screenshot 을 사용하여 설명하는 것을 적극 추천합니다.
- 동영상을 제공해도 괜찮으나 이 경우 반드시 .mp4 또는 .mov 파일로 제출하셔야 합니다.
- 문서는 markdown 을 (.md) 추천하며, Microsoft Word 나 Google Docs 를 사용할 경우 원본 파일을 제출하되 꼭 PDF 로 변환해서 PDF 파일도 같이 제출하셔야 합니다.

만약 어떤 딥러닝 프로젝트를 수행하기 위해서 직접 데이터를 수집하고 정제하는 작업을 해야한다면, 만약 이 작업을 다른 사람에게 시키거나 외주를 (outsourcing) 줘야한다고 한다면, 그 사람들에게 어떻게 잘 설명할 수 있을지, 메뉴얼을 만든다는 기분으로 작업하시면 될 것 같습니다.

여러번의 시도가 필요할 것으로 생각합니다.

성공과 실패의 여부와 상관없이 모든 시도한 내용을 체계적으로 하나하나 꼼꼼히 해방일지에 기록해야합니다. 그리고 각 실험에 어떤 변화를 주었는지도 자세하게 빠짐없이 해방일지에 기록해주세요. 각 실험에 사용된 원본이미지와 모델에 입력하기 바로 직전의 이미지 모두를 제출해야합니다.

1.3.2 데이터 수집

필수

- 자신의 손글씨 원본 이미지
 - "최소" 두 세트 (0 두 개, 1 두 개, ..., 9 두 개)
 - 인식에 실패한 손글씨 이미지도 모두 포함해서 제출합니다. 단, 인식 성공과 실패를 명확히 구분해주세요.
- 한동대학교 친구 손글씨 0-9 까지 ("최소" 3명)

선택사항

- 한동대학교 교수님 손글씨 0-9 까지
 - 선택사항이지만, 만약 손글씨 이미지를 확보하면 최종 점수에 무조건 5점 가산 (인식 성공/실패 상관없이)
 - 최대 3명까지. 세 분의 교수님 손글씨를 받으면 최종 점수에 무조건 15점 가산
 - 교수님께 부탁드립니다 경우는 공손히 자기소개 및 프로젝트 이름을 소개 하신 후 양해를 구하고 진행 해주시길 부탁드립니다.
 - 한동대학교 외국인 교수님의 경우는 최종 점수에 3점 가산 (MNIST 가 외국인 손글씨를 모은 데이터셋이기에 아무래도 외국인 손글씨 인식이 상대적으로 조금 더 쉬움. 예를 들면 2 나 7 은 한국인이 잘 사용 안하는 손글씨 모양이 존재)
- 총장님 (또는 부총장님) 손글씨 0-9 까지
 - 10점 추가 보너스

유의사항

- 섭외한 분의 이름, 소속학부를 기록해주세요.
- 섭외한 사람의 손글씨라는 것을 증명할 수 있는 인증 사진을 반드시 제출해야 합니다.
- 손글씨는 반드시 첨부된 PDF 파일을 A4 용지에 출력하신 후, 그 위에 손글씨를 받으셔야 하며, 손글씨를 받은 종이가 같이 최종 제출해야합니다. 그러므로, 꾸겨지거나 찢어지지 않도록 보관을 잘 부탁드립니다.
- 이미 다른 학우에게 손글씨를 부탁 받은 친구/교수님에게는 다시 부탁드립니다 안됩니다.
- 제출하는 압축 파일에 여러개의 폴더와 파일들이 포함될텐데, 각 폴더와 파일을 설명하는 문서를 작성해서 같이 제출하세요.

1.3.3 기타

최적의 성능을 내기 위하여 원본이미지 데이터를 어떻게 수집하고 어떻게 전처리하면 되는지 그 구체적인 전략을 수립하고 배워가는 시행착오를 해방일지에 적어나가시면 됩니다.

결과와 상관없이 실험을 어떻게 디자인하고 진행했는지 그 과정을 더 중요하게 보고 거기에 맞는 점수를 드립니다. 과학계 말하면 인식률이 저조해도 100점에 준하는 점수를 받을 수 있다는 이야기입니다

어떻게 데이터를 수집해야하고 가공하는지는 매우 중요한 문제입니다. "*Garbage In Garbage Out*" - 쓰레기 데이터를 넣으면 아무리 좋은 모델이어도 쓰레기를 뱉아 냅니다.

만약 성능을 높이기 위해 최초 수립한 **base model** 을 후반부에 수정했다면 무엇을 어떻게 수정했는지, 어떤 절차로 바꾸었는지 자세하게 해방일지에 기록해주세요.

해방일지에 각 작업을 한 날짜를 반드시 적어주세요 (예: 11월 8일, 11월 15일, 12월 2일, ...)

2 -

준비되는데로 다시 알려드립니다 (TBA)