

**딥러닝 hw2 과제**

**과목명 딥러닝**

**담당교수 장익범교수님**

**제출일 20231015**

**전공 컴퓨터전자시스템**

**학번 201904458**

**이름 이준용**

1. Learning curve (y축: Train\_loss & validation\_loss curve, x축: Epochs)

스크린샷, 텍스트, 라인, 그래프이(가) 표시된 사진

자동 생성된 설명

2. Hyperparameter

에폭(epoch) 수 => 10

: 데이터셋을 반복하는 횟수

Epoch은 batch size와 learning rate을 각각 수치를 조정해가면서 loss curve를 확인했는데 그래프가 epoch 수가 5이상부터는 일정하게 감소하는 형태를 띄어서 10으로 계속 고정하여 훈련시켰습니다.

배치 크기(batch size) => 32

: 매개변수가 갱신되기 전 신경망을 통해 전파된 데이터 샘플의 수

Batch size를 32와 64를 둘 다 훈련시켰을 때 32인 경우에 더 훈련이 잘되어 32로선택했습니다.

학습률(learning rate) =>0.00001

: 각 batch/epoch에서 모델의 매개변수를 조절하는 비율. 값이 작을수록 학습 속도가 느려지고, 값이 크면 학습 중 예측할 수 없는 동작이 발생할 수 있습니다.

처음 훈련시켰을 때 0.001로 훈련시켰는데 예측할 수 없는 loss로 “nan”이 출력되었습니다.

그래서 0.1씩 곱해가면서 learning rate을 줄여갔습니다. 200\*200\*3= 120000 정도의 feature인데 learning rate도 1e-5정도 되어야 fitting이 잘 될 것이라고 생각했습니다.

3. 예측에 가장 중요한 feature를 순서대로 1,000가지 나열하기.

Feature = (image) 200\*200\*3 = 120000개

Feature 1: 눈 주위의 주름

Feature 2: 미간 주름

Feature 3: 입 주위의 주름

Feature 4: 주름 깊이

Feature 5: 머리카락 색상

Feature 6: 피부 톤

Feature 7: 머리카락 스타일

Feature 8: 머리카락 밀도

Feature 9: 눈동자 크기

Feature 10: 피부 상태

Feature 11: 피부 변화

Feature 12: 표정

Feature 13: 미소 여부

Feature 14: 머리카락 모양

Feature 15: 알레르기 반응

Feature 16: 머리카락 상태

Feature 17: 주름의 형태

Feature 18: 머리카락 길이

Feature 19: 입모양

Feature 20: 코 형태

Feature 21: 피부 잡티

Feature 22: 머리카락 굵기

Feature 23: 입술 모양

Feature 24: 머리카락 꼬임 정도

Feature 25: 눈 색상

Feature 26: 입술 색상

Feature 27: 얼굴 형태

Feature 28: 코 길이

Feature 29: 미간 깊이

Feature 30: 얼굴 비율

...

Feature 1000: 배경 조명

Feature Importance: weight

4. 위의 Feature Importance Map으로부터 깨달은 사실 혹은 배운 점에 대해 서술

Feature Importance Map은 머신러닝 모델이 예측을 수행하는 데 있어 각 피처가 얼마나 중요한 역할을 하는지를 시각화한 도구입니다. 이를 통해, 다음과 같은 사실들을 깨달을 수 있습니다:

피처의 상대적 중요성: Feature Importance Map은 각 피처가 모델의 예측에 어느 정도 기여하는지 보여줍니다. 이를 통해 가장 중요한 피처들이 무엇인지, 그리고 그것들이 어떻게 타겟 변수와 관련되어 있는지 이해할 수 있습니다.

모델의 해석 가능성: 복잡한 머신러닝 모델, 특히 앙상블 모델과 신경망 등은 종종 "블랙 박스"로 비유됩니다. Feature Importance Map은 이러한 복잡한 모델의 내부 작동 방식에 대해 좀 더 많은 인사이트를 제공하며, 모델이 왜 특정 예측을 만드는지 설명하는데 도움이 됩니다.

특징 선택: Feature Importance Map으로부터 가장 영향력 있는 변수들을 파악하고, 이 정보를 바탕으로 feature selection 혹은 feature engineering을 진행할 수 있습니다. 즉, 중요도가 낮거나 영향력 없는 변수들을 제거함으로써 계산 비용을 줄일 수 있고, 과적합(overfitting) 문제를 완화할 수 있습니다.

데이터 문제 식별: 만약 예상치 못하게 어떤 변수가 너무 큰 영향력을 가진다면 데이터에 문제가 있는 것일 가능성도 확인해볼 필요가 있습니다. 예를 들어 데이터 유출(leakage)나 전처리 과정에서 발생한 오류 등입니다.

따라서 Feature Importance Map는 데이터 과학자와 분석가들에게 매우 유용한 도구로 작용합니다.

5. 과제를 진행하며 어려웠던 점 서술.

얼굴 사진을 통한 나이 예측 모델 학습 관련 경험이 없다 보니 Hyperparameter 값을 이렇게 저렇게 해보면서 시간을 소모했습니다.

6. 과제를 진행하며 배운 점과 느낀 점을 서술

이번 과제를 차근차근 해보면서 다음 딥러닝 관련 주제로 스스로 훈련 할 때는 이 과제가 밑거름이 되어 잘 수행할 수 있을