

우울증 종합 진단 플랫폼

딥러닝팀

김예찬

윤지영

채소연

한지원

홍지우



1. 주제 소개



1. 주제 소개



딥러닝을 이용한 우울증 진단 모델





2-1. 한국어 데이터 EDA

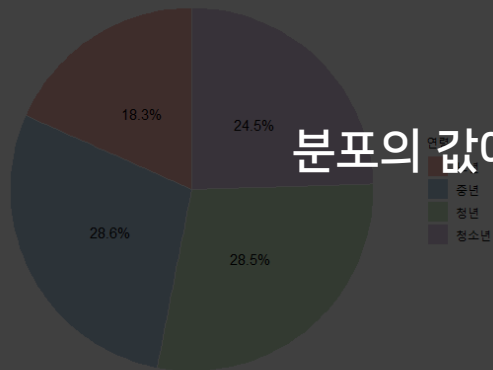


2-1. 한국어 데이터 EDA



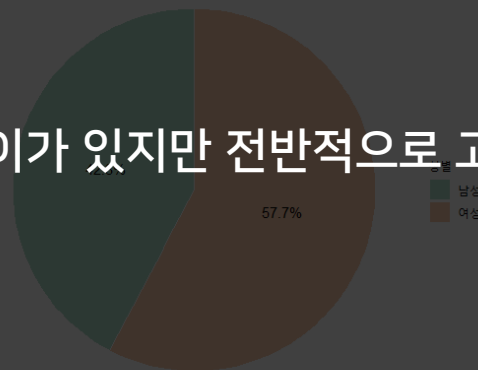
Training Data

연령



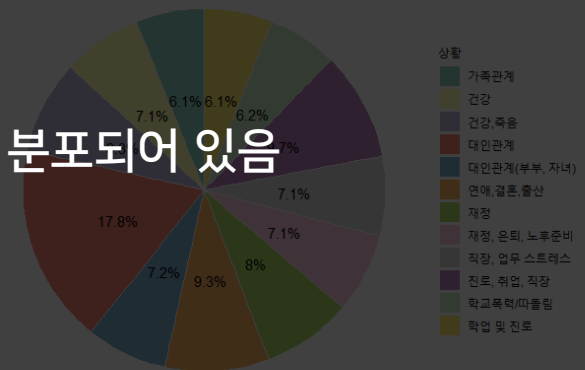
청년이 가장 많고
노년이 가장 적음

성별



여성이 남성보다
더 자료가 많음

상황별



대인관계 항목이 가장 자료가 많고
그 외에는 분포의 차이가 크지 않음

분포의 값에는 차이가 있지만 전반적으로 고르게 분포되어 있음





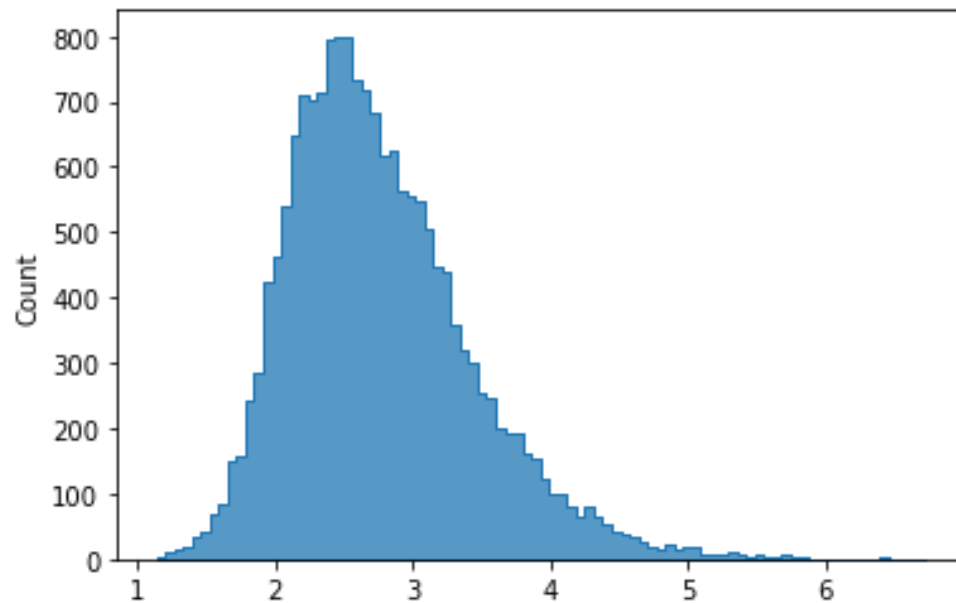
2-2. 영어 데이터 EDA



2-2. 영어 데이터 EDA



음성 데이터 길이



음성 데이터의 길이가 대체로
2초~3초 정도 사이인 것을
확인할 수 있음

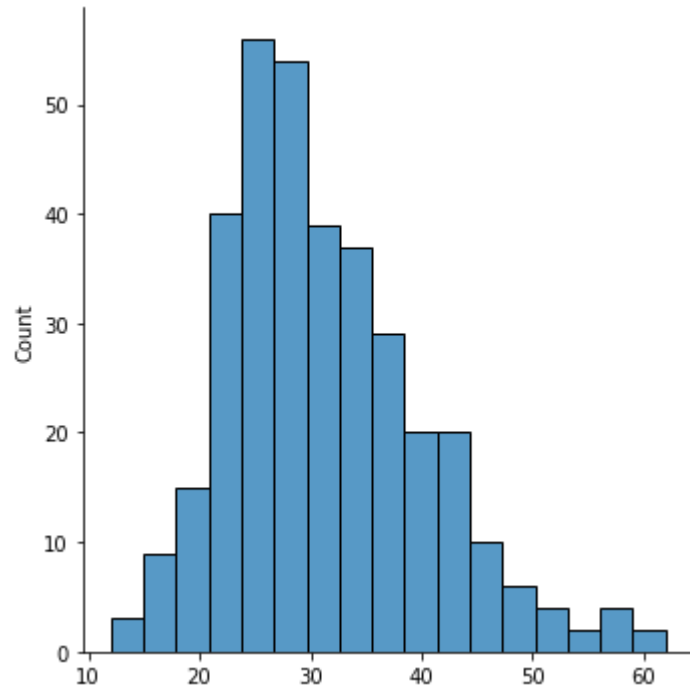




2-2. 영어 데이터 EDA



텍스트 데이터 길이



텍스트의 길이를 전처리한 결과
대략 30개의 알파벳으로 구성되어
있는 것을 확인할 수 있음





4-1. 음성 감정 분류



4-1. 음성 감정 분류



Data input

```
#DataFlair - Split the dataset
x_train,x_test,y_train,y_test=load_data(test_size=0.25)

#DataFlair - Get the shape of the training and testing datasets
print((x_train.shape, x_test.shape))

#DataFlair - Get the number of features extracted
print(f'Features extracted: {x_train.shape[1]}')

#DataFlair - Initialize the Multi Layer Perceptron Classifier
model=MLPClassifier(alpha=0.01, batch_size=256, epsilon=1e-08, hidden_layer_sizes=(300,), learning_rate='adaptive', max_iter=1500,early_stopping=False, warm_start=True)

#DataFlair - Train the model
model.fit(x_train,y_train)

import matplotlib.pyplot as plt

from sklearn.neural_network import MLPClassifier
from sklearn.preprocessing import MinMaxScaler
from sklearn import datasets
from sklearn.exceptions import ConvergenceWarning

#DataFlair - Predict for the test set
y_pred=model.predict(x_test)

#DataFlair - Calculate the accuracy of our model
accuracy=accuracy_score(y_true=y_test, y_pred=y_pred)

#DataFlair - Print the accuracy
print("Accuracy: {:.2f}%".format(accuracy*100))
```

MLP Classifier를 활용해서 모델 학습



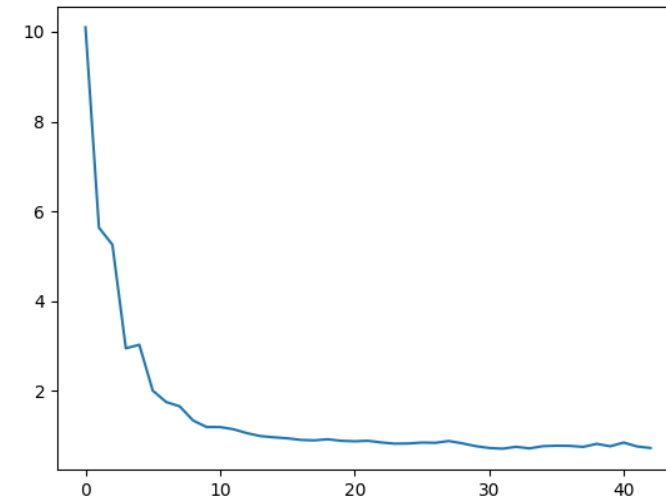
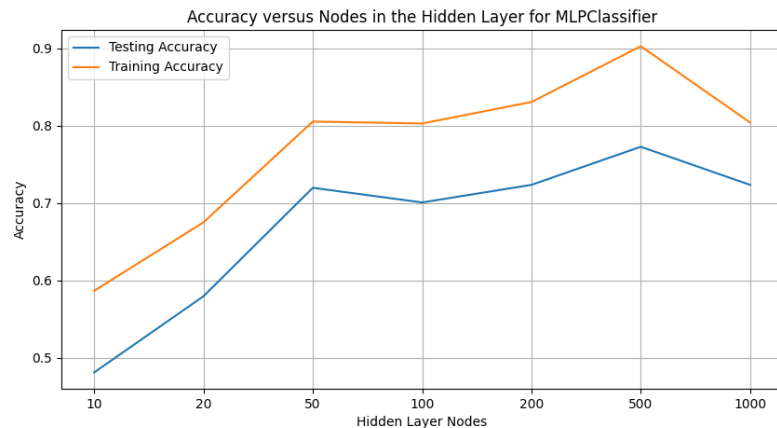


4-1. 음성 감정 분류



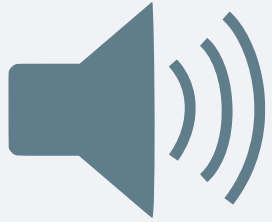
Accuracy/Loss plot

Accuracy: 69.47%
((7279, 188), (2427, ...))
Features extracted: 188
Accuracy: 69.47%



최종적으로 수렴하는 것을 확인할 수 있음





4-2. 텍스트 감정분류



4-2. 텍스트 감정분류



모델링 진행

| | | | | | | |
|-----|----|-----|-----|----|----|-----|
| 101 | 54 | 200 | 821 | 85 | 94 | 102 |
|-----|----|-----|-----|----|----|-----|



KoBERT Classifier



'기쁨'으로 분류

| | | | | | |
|---------|--------|--------|--------|--------|--------|
| 기쁨(86%) | 불안(3%) | 당황(3%) | 분노(3%) | 슬픔(3%) | 상처(2%) |
|---------|--------|--------|--------|--------|--------|



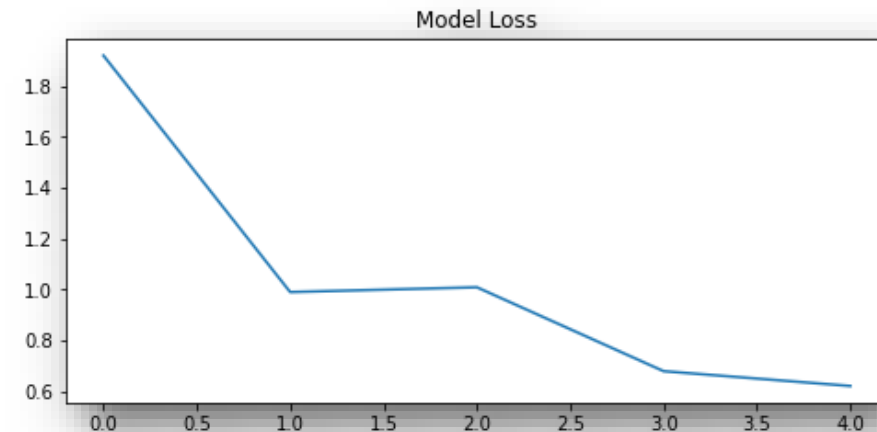
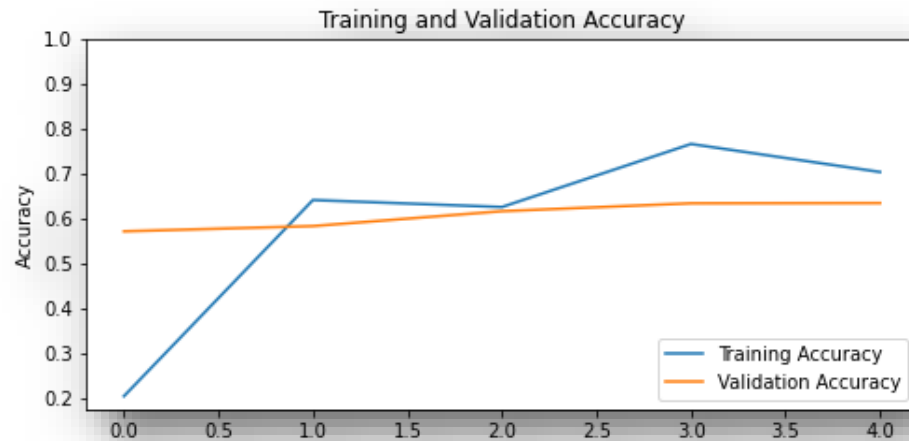


4-2. 텍스트 감정분류



Accuracy plot/Loss plot

train accuracy: 0.809, validation accuracy: 0.635



Loss는 최종적으로 수렴하나, 정확도는 epoch이 증가할 수록 과적합 현상을 보임





5. Multi-Modal 감정 분류 모델

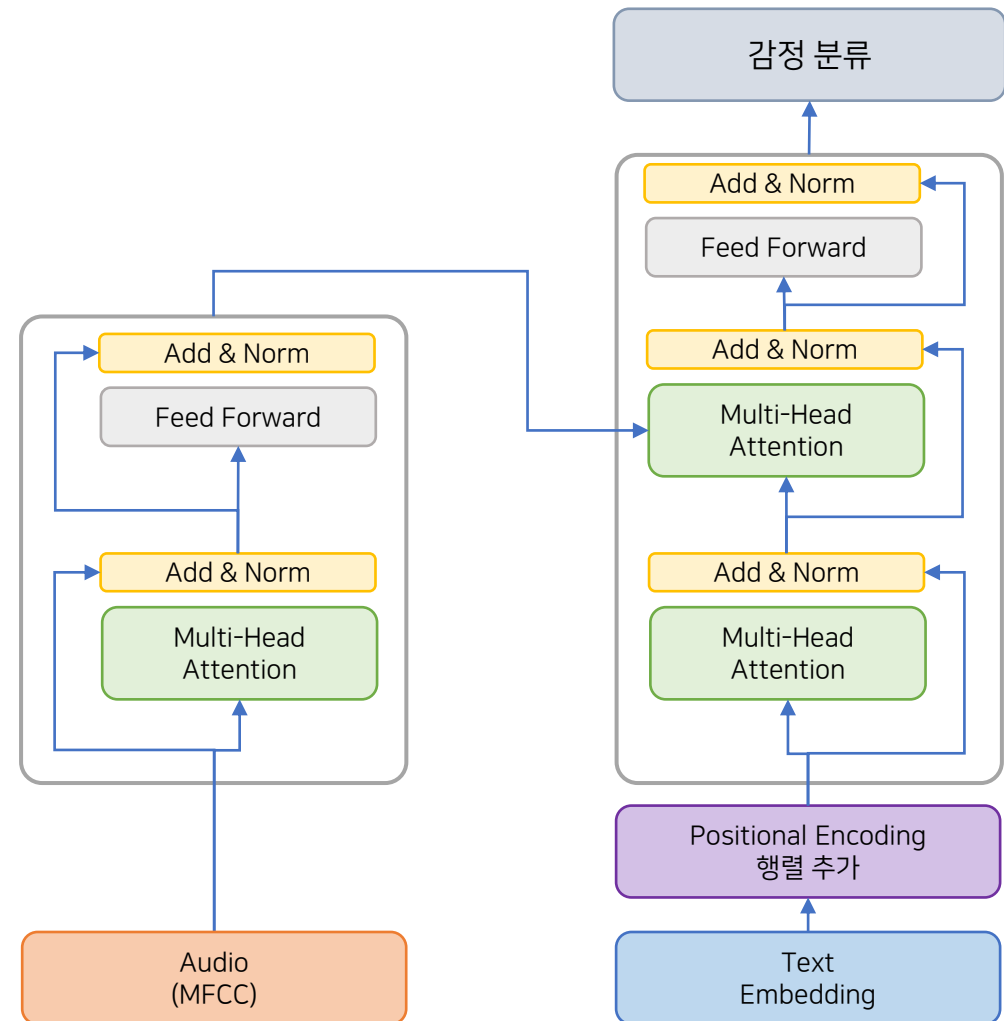


5. Multi-Modal 감정 분류 모델



Multi-modal Transformer

| 속성 | 내용 |
|-----------|---|
| 파라미터 수 | 약 350만 개 |
| 감정분류 Cell | FC Layer 활용 예정 |
| 특이사항 | <ul style="list-style-type: none">음성의 특성을 텍스트에 반영최종 분류시 텍스트 Feature만 사용텍스트 임베딩: 모델 내에서 진행 |



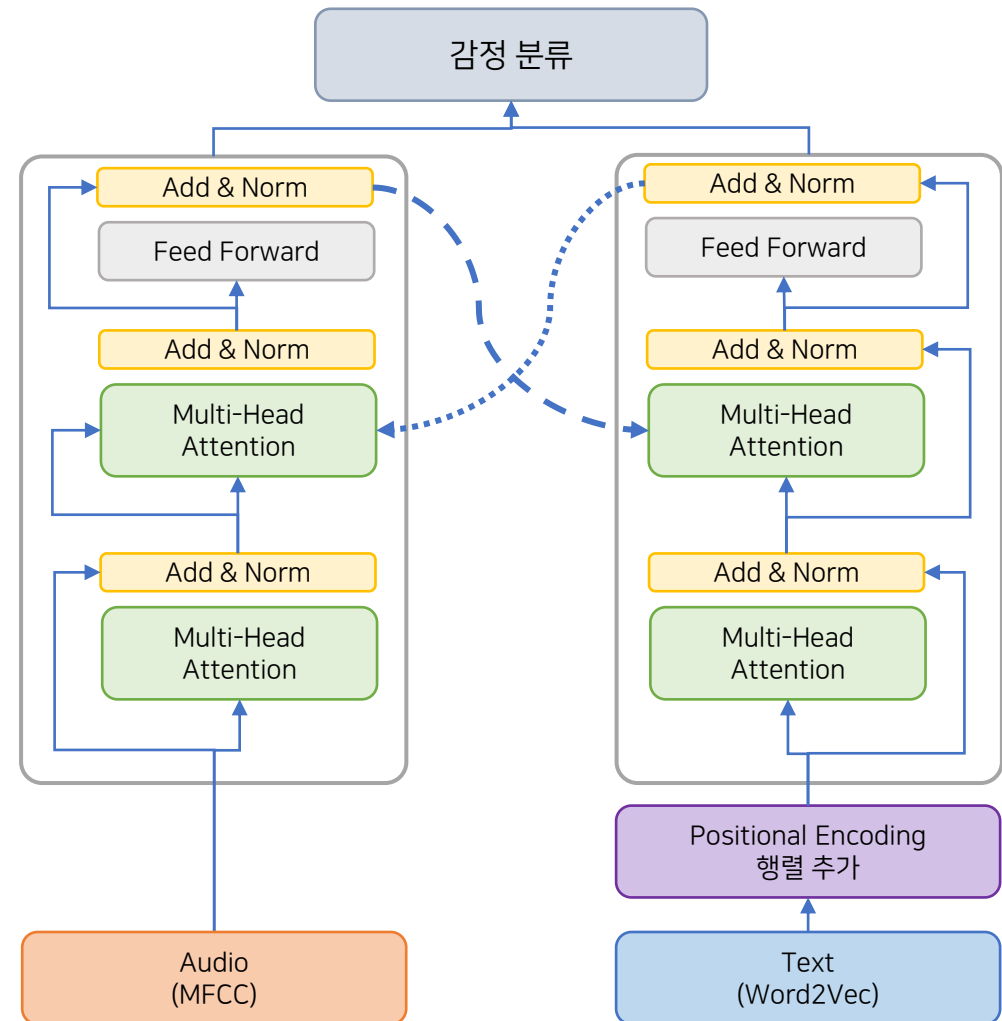


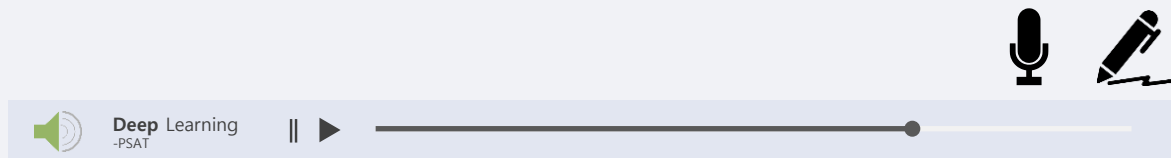
5. Multi-Modal 감정 분류 모델



Cross Attention Transformer

| 속성 | 내용 |
|-----------|---|
| 파라미터 수 | 약 400만 개 |
| 감정분류 Cell | Convolutional Layer 사용 예정 |
| 특이사항 | <ul style="list-style-type: none">서로의 특성을 상호 반영최종 분류시 양쪽 Feature Stack텍스트 임베딩: Word2Vec |





멀티모달 감정인식 모델

딥러닝팀

김예찬

윤지영

채소연

한지원

홍지우



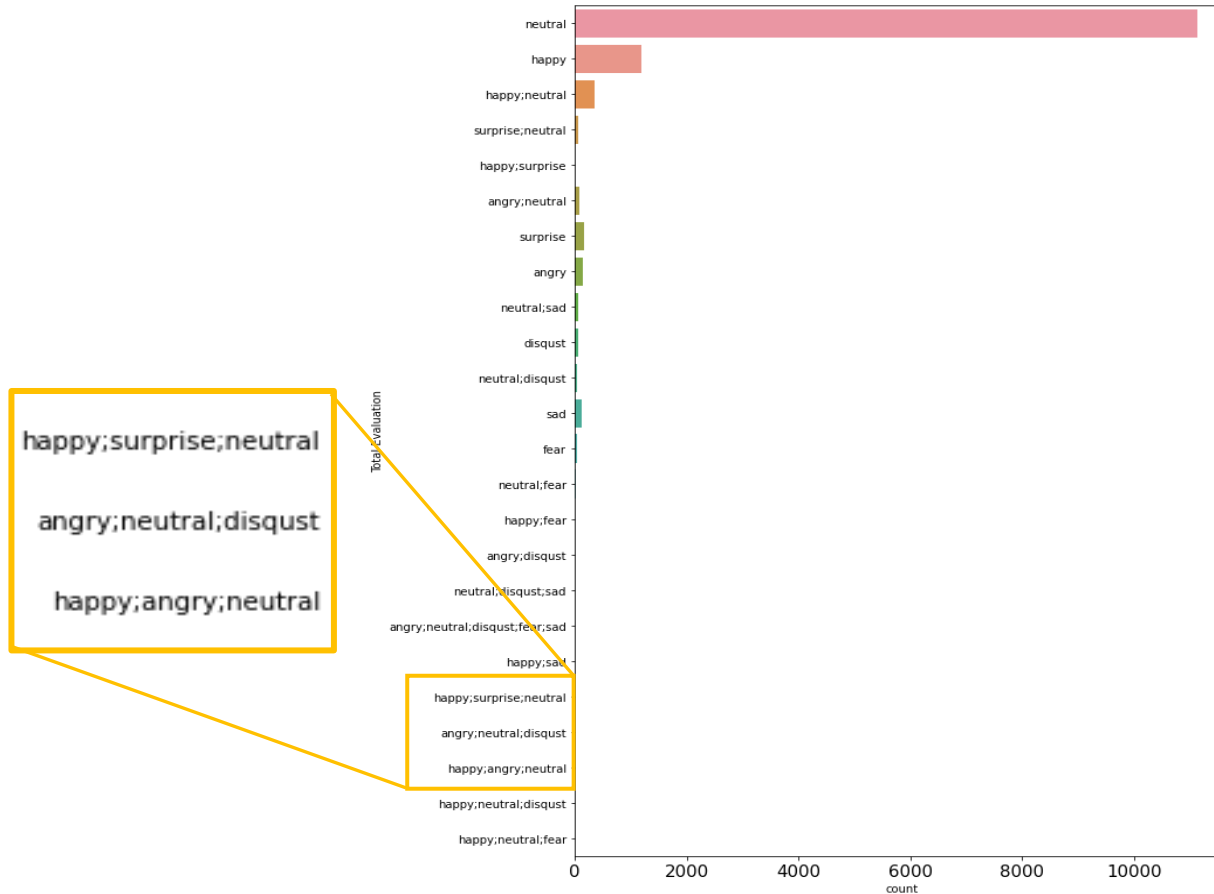
1. 데이터 소개 및 EDA



1. 데이터 소개 및 EDA



감정분포



← Neutral

Neutral(중립) 감정의 비율이

불균형적으로 큼

다수의 감정이 라벨링 된 데이터 존재





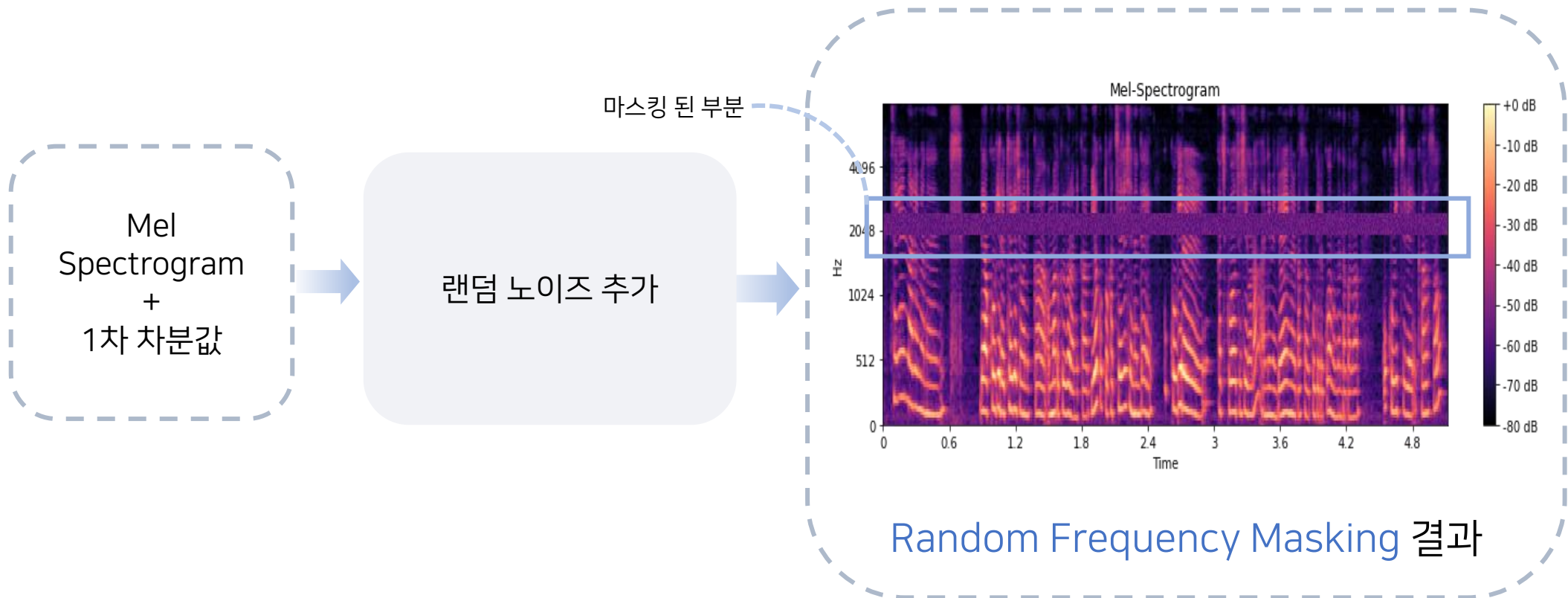
3. 음성 데이터 전처리



3. 음성 데이터 전처리



데이터 증강(augmentation) – Random Frequency Masking





4. 텍스트 데이터 전처리

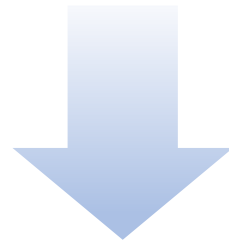


4. 텍스트 데이터 전처리



텍스트 데이터 토큰화 (koBERT)

너는 대선 때 투표할 수 있어?



사전 훈련된 KoBERT 토크나이저

| | | | | | | | | | |
|----|---|-----|-----|----|-----|---|----|-----|---|
| _너 | 는 | _내년 | _대선 | _때 | _투표 | 할 | _수 | _있어 | ? |
|----|---|-----|-----|----|-----|---|----|-----|---|





4. 텍스트 데이터 전처리

불균형한 텍스트 데이터

텍스트 데이터 증강방법

| 방법 | 문장 |
|---------------|-------------------------------------|
| 원래 문장 | 중심을 잃고 목소리도 잃고 비난받고 사람들과 멀어지는 착각 속에 |
| SR (동의어 대체) | 핵심을 잃고 목소리도 잃고 비판받고 인간들과 멀어지는 혼란 속에 |
| RI (단어 임의 대체) | 돈을 잃고 목소리도 잃고 비난받고 사람들과 멀어지는 착각 속에 |
| RS (단어 위치 변경) | 중심을 잃고 사람들과 잃고 비난받고 목소리도 멀어지는 착각 속에 |
| RD (단어 임의 삭제) | 중심도 목소리도 사람들과 멀어지는 착각 |

의미 왜곡 위험이 적은 RS, RD 선택





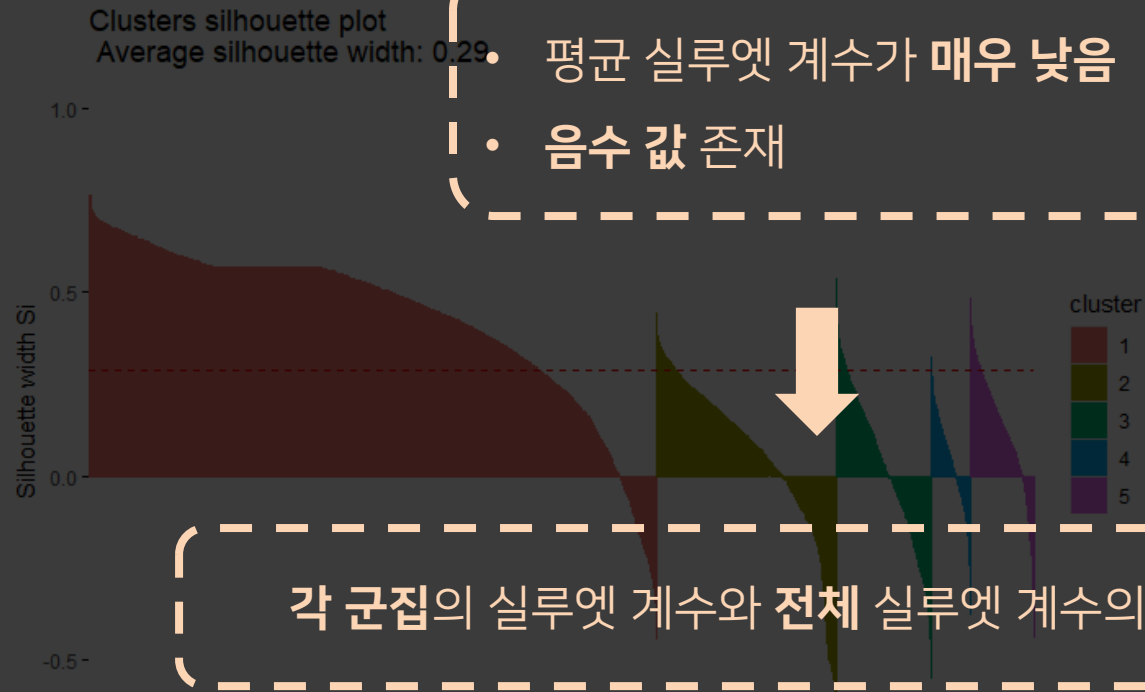
5. 바이오 데이터 분석



5. 바이오 데이터 분석



바이오 데이터 - Silhouette



색깔 너비

=

군집에 속한 데이터 수





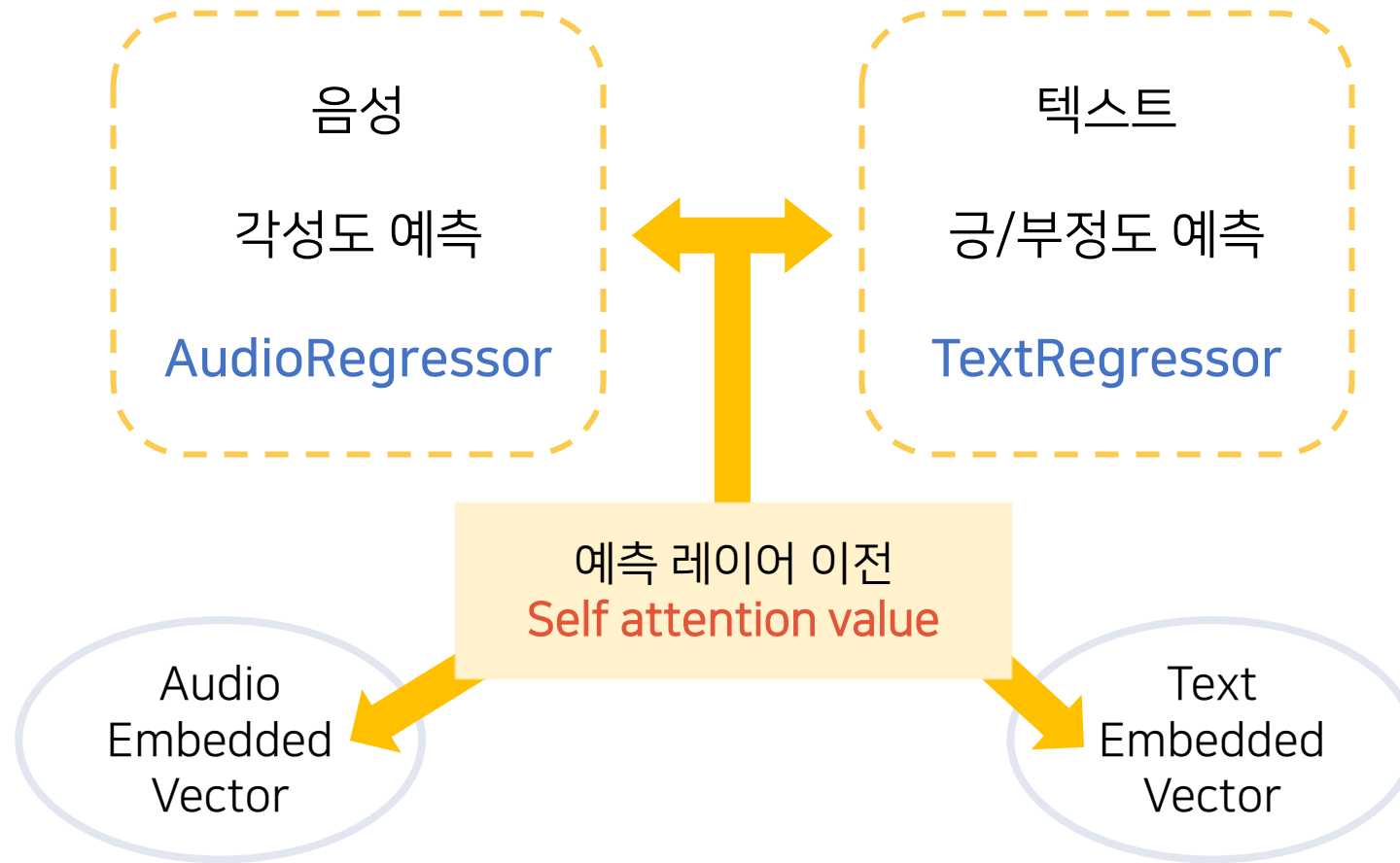
6. 모델 설명



6. 모델 설명



Multimodal Emotion Classifier





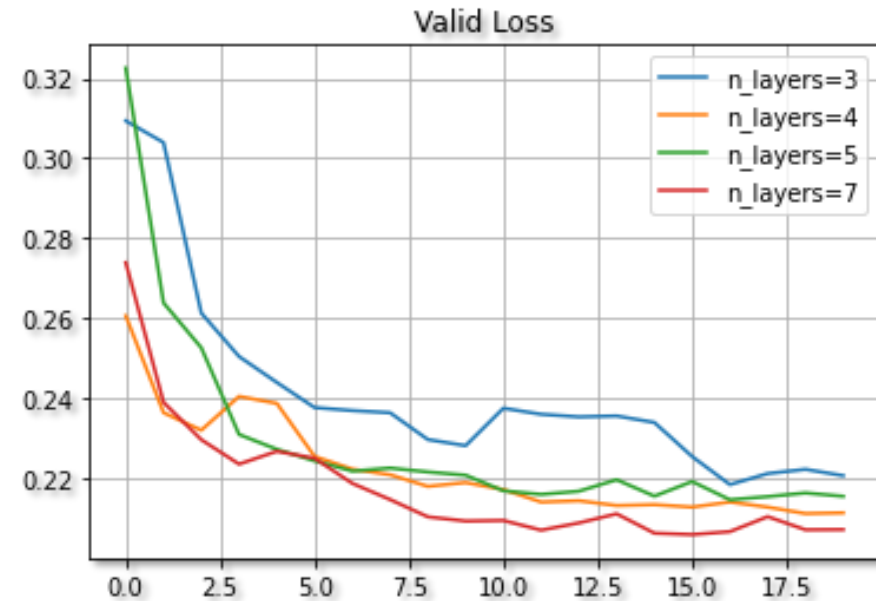
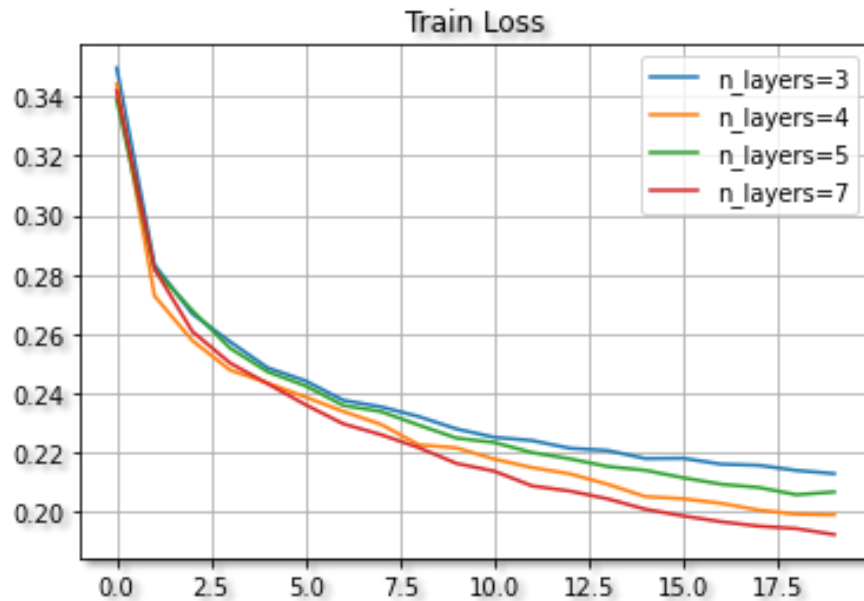
7. 학습 결과



7. 학습 결과



음성 싱글 모달리티- Arousal 예측 결과



Multihead Attention Layer의 반복 횟수에 따른 RMSE변화 → 7회일 때 최저

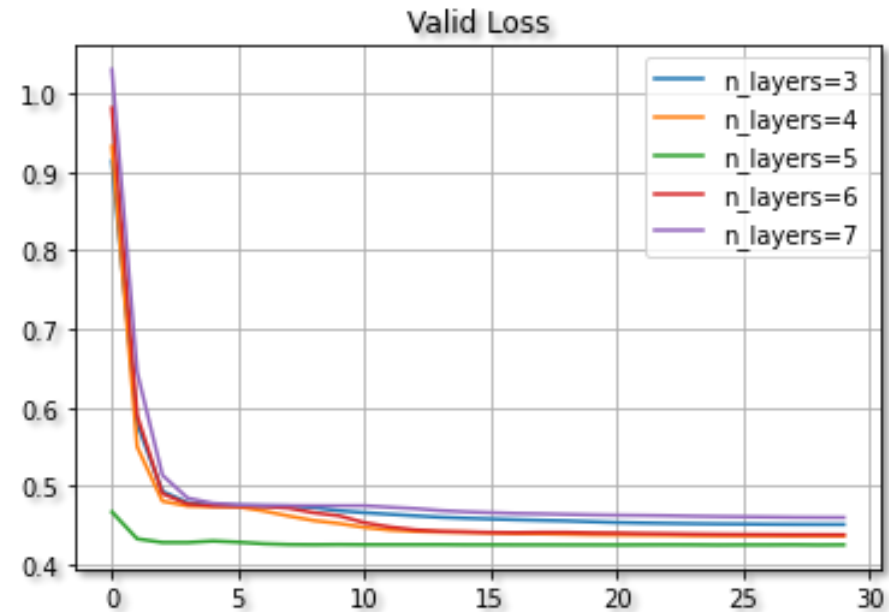
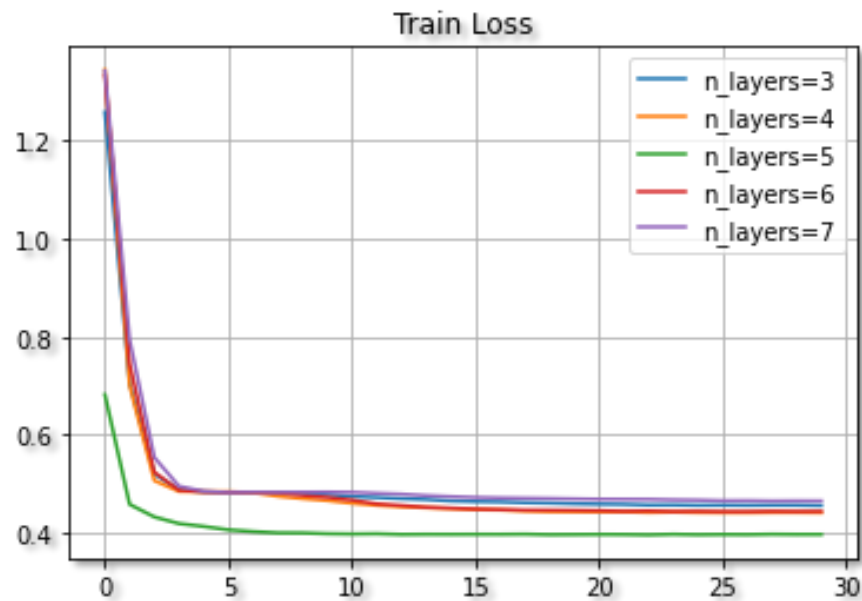




7. 학습 결과



텍스트 싱글 모달리티- Valence 예측 결과



Multihead Attention Layer의 반복 횟수에 따른 RMSE변화 → 5회일 때 최저





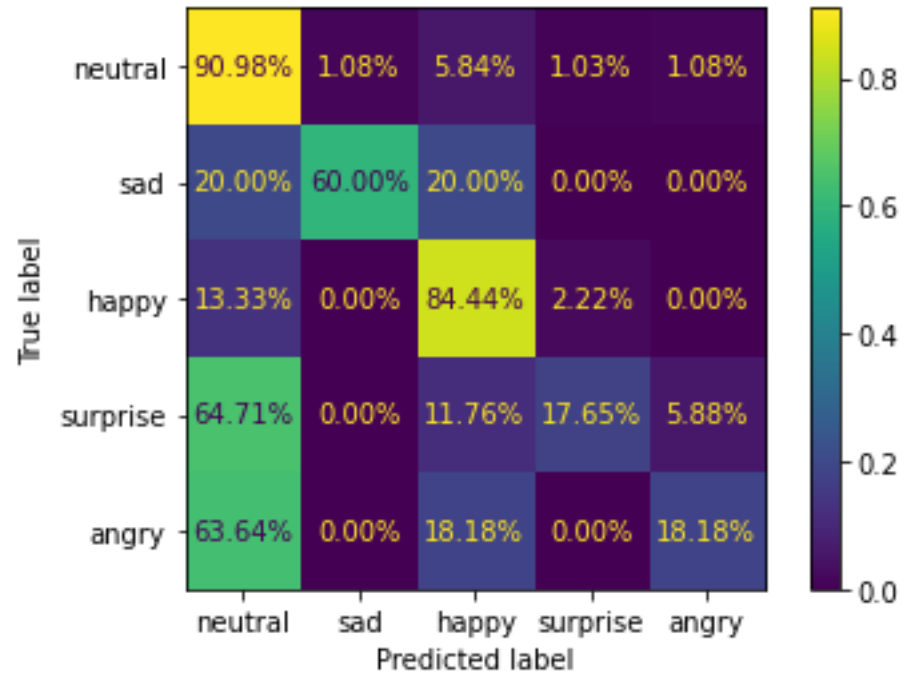
8. 성과 & 한계



8. 성과 & 한계



성과



불균형한 데이터임에도
불구하고 데이터 수가
적은 라벨들도 예측 성공

