
각성도 및 긍/부정도의 싱글모달 사전학습 예측 모델 기반 멀티모달 감정인식 모델

Multimodal Emotional Recognition Model based on
Singlemodal Pretrained Prediction Model of Valence and Arousal

홍지우 김예찬 윤지영 채소연 한지원

Department of Statistics, Sungkyunkwan University

Contents

01 제안 배경

- A. 멀티모달 감정 인식 모델 선행연구
- B. 감정의 전달

02 문제 인식

- A. 다중 분류 문제
- B. 데이터 불균형 문제

03 데이터 구조 파악

04 데이터 선정(모달리티)

- A. 생체신호 데이터 분석
- B. 각성도/공부정도 데이터 분석

05 데이터 전처리

- A. 음성 데이터 전처리
- B. 텍스트 데이터 전처리

06 모델 구성

- A. 베이스라인 모델 선정
- B. Transformer
- C. 음성-각성도 예측 모델
- D. 텍스트-공/부정도 예측 모델
- E. 음성/텍스트 특징 추출 블록
- F. 감정 인식 블록

07 모델 평가

- A. F1 Score
- B. Confusion Matrix
- C. 모델 비교
- D. 최종 모델 평가
- E. 의의

01

제안 배경

1. 제안 배경

A. 멀티모달 감정 인식 모델 선행연구

Emotions from Text: Machine Learning for Text-based Emotion Prediction

Cecilia Ovesdotter Alm, D. Roth, R. Sproat • Published in HLT 6 October 2013

In addition to information, text contains attitudinal, and m
addition, we present plans for a more cognitively sound s
larger set of basic emotions. [Expand](#)

Understanding Emotions in Text Using Deep Learning and Big Data

Deep learning approach to text analysis for human emotion detection from big data

^aadhakrishnan Srikanth ^b, Michel Galley ^c,

Jia Guo

From the journal [Journal of Intelligent Systems](#)

Emotion analysis of Arabic tweets using deep learning approach

[Massa Baali](#) & [Nada Ghneim](#) 

언어적 요소만을 기반으로 하는 싱글모달 / 멀티모달 감정인식 모델 多

1. 제안 배경

B. 감정의 전달

언어적 요소

① 발화

② 텍스트

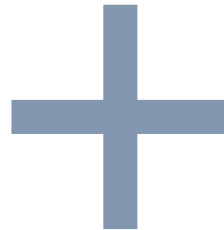
1. 제안 배경

B. 감정의 전달

언어적 요소

① 발화

② 텍스트



비언어적 요소

① 표정

② 몸짓

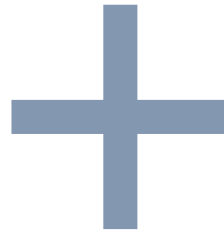
③ 손짓

1. 제안 배경

B. 감정의 전달

언어적 요소

- ① 음성
- ② 텍스트



비언어적 요소

- ① 생체신호
- ② 각성도
- ③ 긍부정도

1. 제안 배경

B. 감정의 전달



언어적 요소 기반의 감정인식모델 + 비언어적 요소를 반영

정확도 ↑

02

문제 인식

2. 문제 인식

해결해야 할 문제

1

다중 분류 문제

2

데이터 불균형 문제

2. 문제 인식

A. 다중 분류 문제

감정 레이블
1. 기쁨
2. 놀람
3. 분노
4. 중립
5. 혐오
6. 공포
7. 슬픔

언어적 요소만으로는 7가지 감정을 정확하게 분류 불가

2. 문제 인식

A. 다중 분류 문제

언어적 요소

① 음성

② 텍스트



비언어적 요소

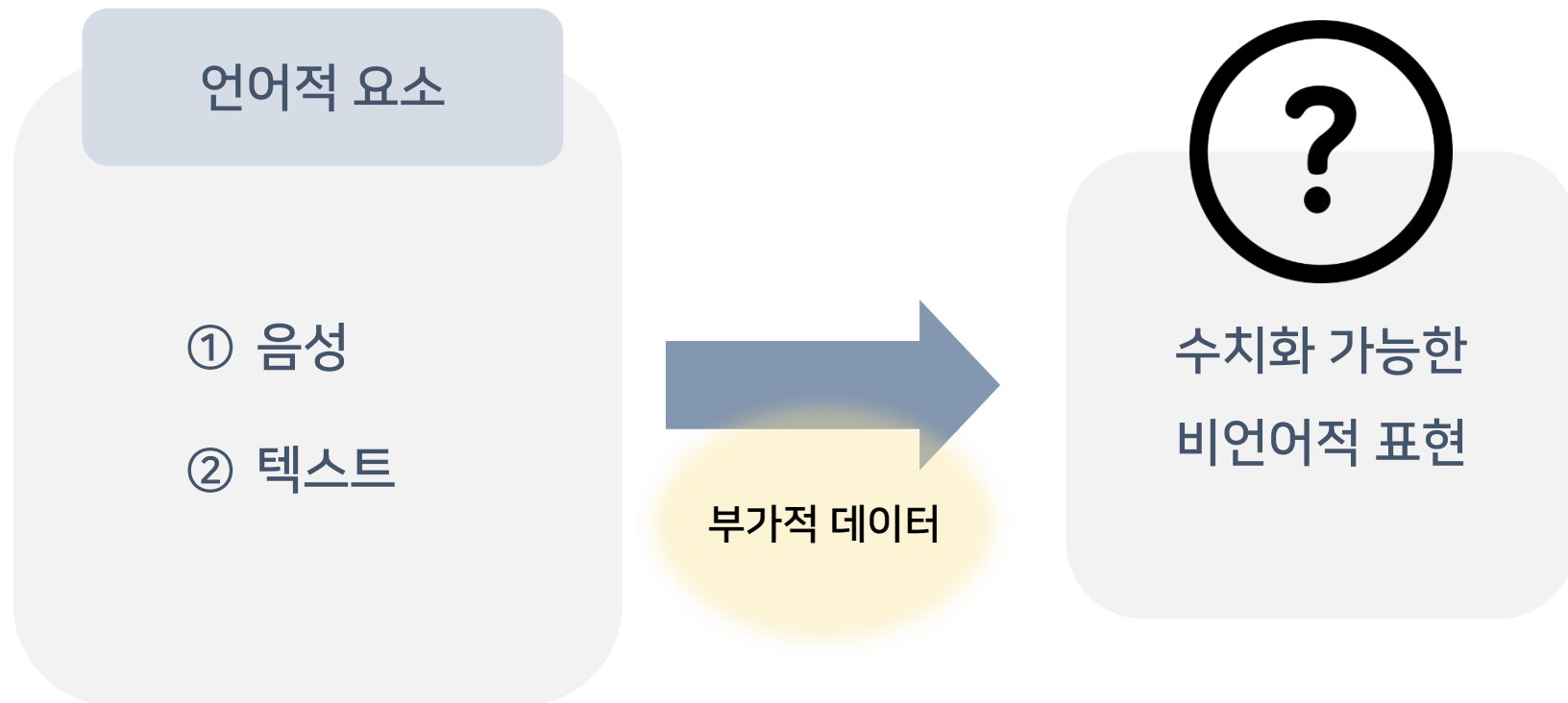
① 생체신호

② 각성도

③ 긍부정도

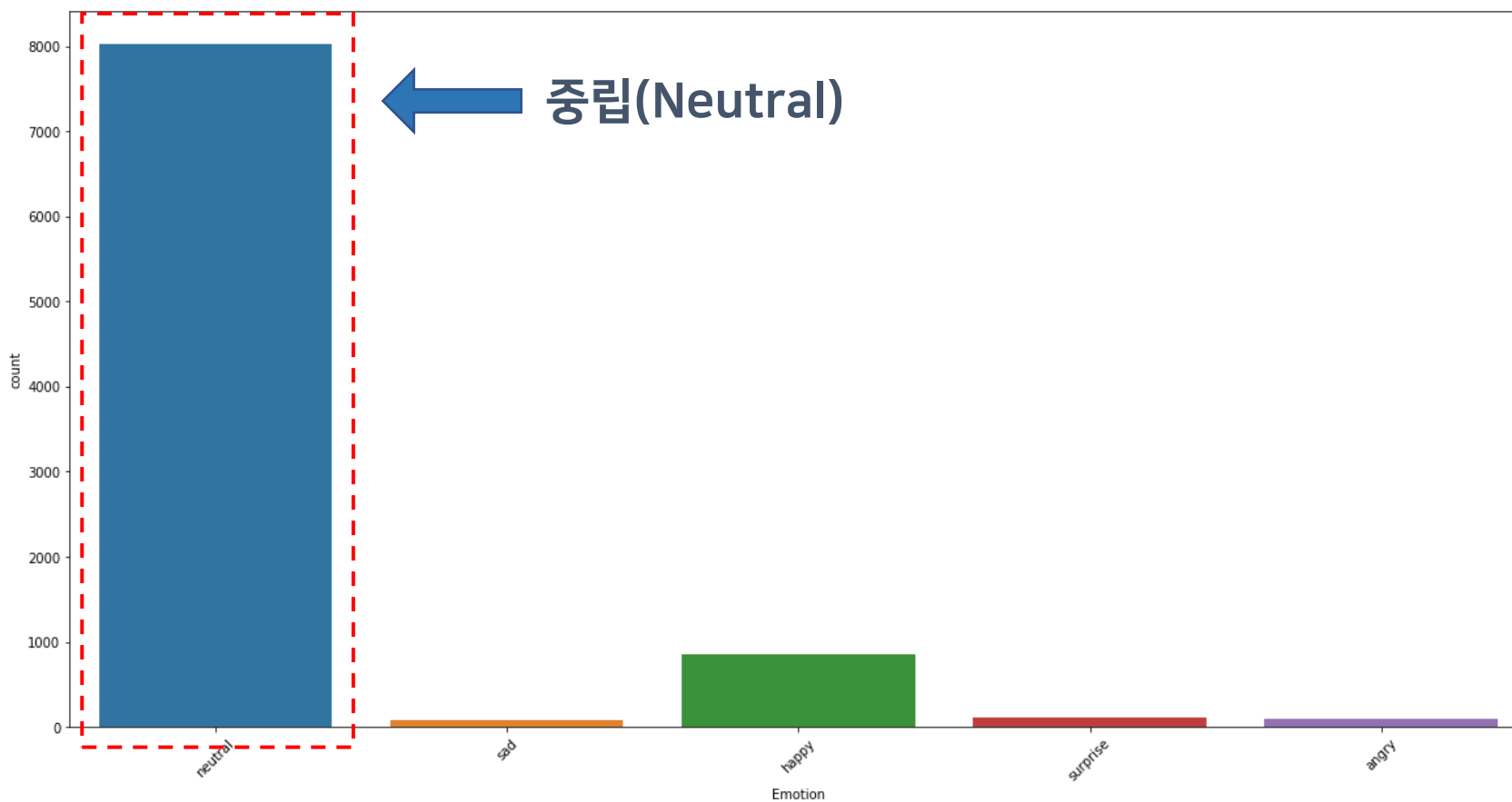
2. 문제 인식

A. 다중 분류 문제



2. 문제 인식

B. 데이터 불균형 문제



[그림] 학습 데이터 감정라벨별 분포

2. 문제 인식

B. 데이터 불균형 문제

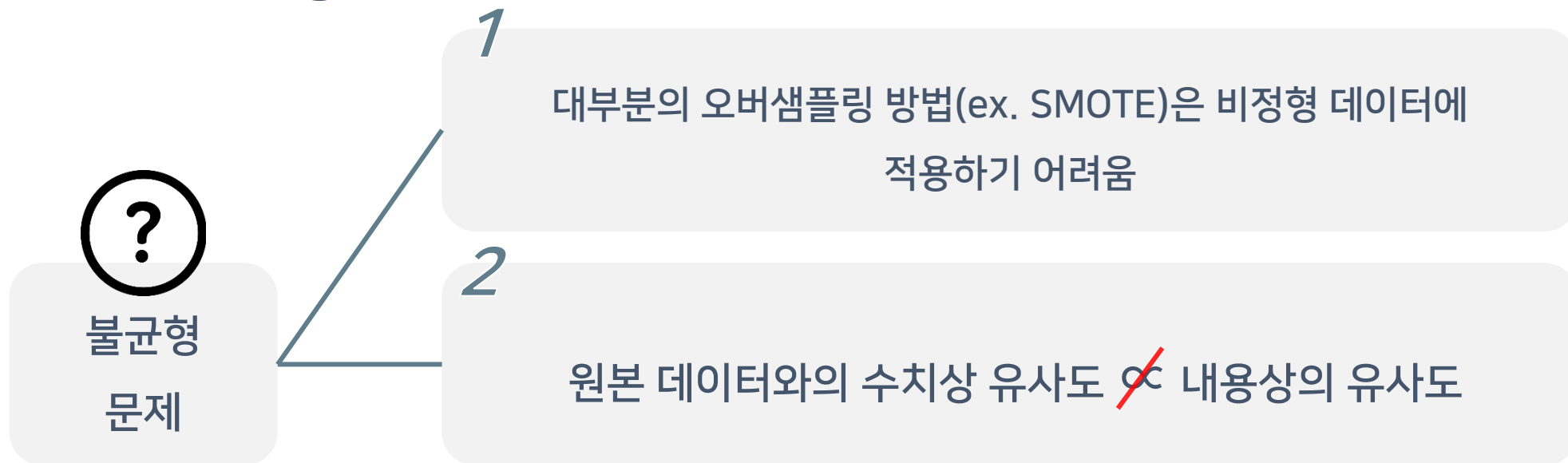


1

대부분의 오버샘플링 방법(ex. SMOTE)은 비정형 데이터에
적용하기 어려움

2. 문제 인식

B. 데이터 불균형 문제



2. 문제 인식

B. 데이터 불균형 문제



03

데이터 구조 파악

3. 데이터 구조 파악

한국어 멀티모달 감정 데이터셋 2020(KEMDy20)



3. 데이터 구조 파악

한국어 멀티모달 감정 데이터셋 2020(KEMDy20)

디렉토리	구성	설명
annotation	.csv	발화에 대한 관찰자의 감정 레이블 평가
wav	.wav / .txt	자유대화 발화 음성/텍스트 파일
EDA	.csv	E4 장치로 수집한 피부전도도 데이터
IBI	.csv	E4 장치로 수집한 심장박동수 데이터
TEMP	.csv	E4 장치로 수집한 체온 데이터

3. 데이터 구조 파악

한국어 멀티모달 감정 데이터셋 2020(KEMDy20)

디렉토리	구성	감정 레이블	설명
annotation	.CSV	1. 기쁨	참여자 감정 레이블 평가
wav	.wav / .txt	2. 놀람	대화 음성/텍스트 파일
EDA	.CSV	3. 분노	수집한 피부전도도 데이터
IBI	.CSV	4. 중립	수집한 심장박동수 데이터
TEMP	.CSV	5. 혐오	
		6. 공포	수집한 체온 데이터
		7. 슬픔	

3. 데이터 구조 파악

한국어 멀티모달 감정 데이터셋 2020(KEMDy20)

디렉토리	구성	감정 레이블	설명
annotation	.CSV	1. 기쁨	참여자 감정 레이블 평가
wav	.wav / .txt	2. 놀람	대화 음성/텍스트 파일
EDA	.CSV	3. 분노	수집한 피부전도도 데이터
IBI	.CSV	4. 중립	수집한 심장박동수 데이터
TEMP	.CSV	5. 혐오	수집한 체온 데이터
		6. 공포	
		7. 슬픔	

데이터 수
100개 미만

3. 데이터 구조 파악

한국어 멀티모달 감정 데이터셋 2020(KEMDy20)

디렉토리	구성	감정 레이블	설명
annotation	.CSV	1. 기쁨	참가자의 감정 레이블 평가
wav	.wav / .txt	2. 놀람	대화 음성/텍스트 파일
EDA	.CSV	3. 분노	수집한 피부전도도 데이터
IBI	.CSV	4. 중립	수집한 심장박동수 데이터
TEMP	.CSV	5. 혐오	수집한 체온 데이터
		6. 공포	
		7. 슬픔	

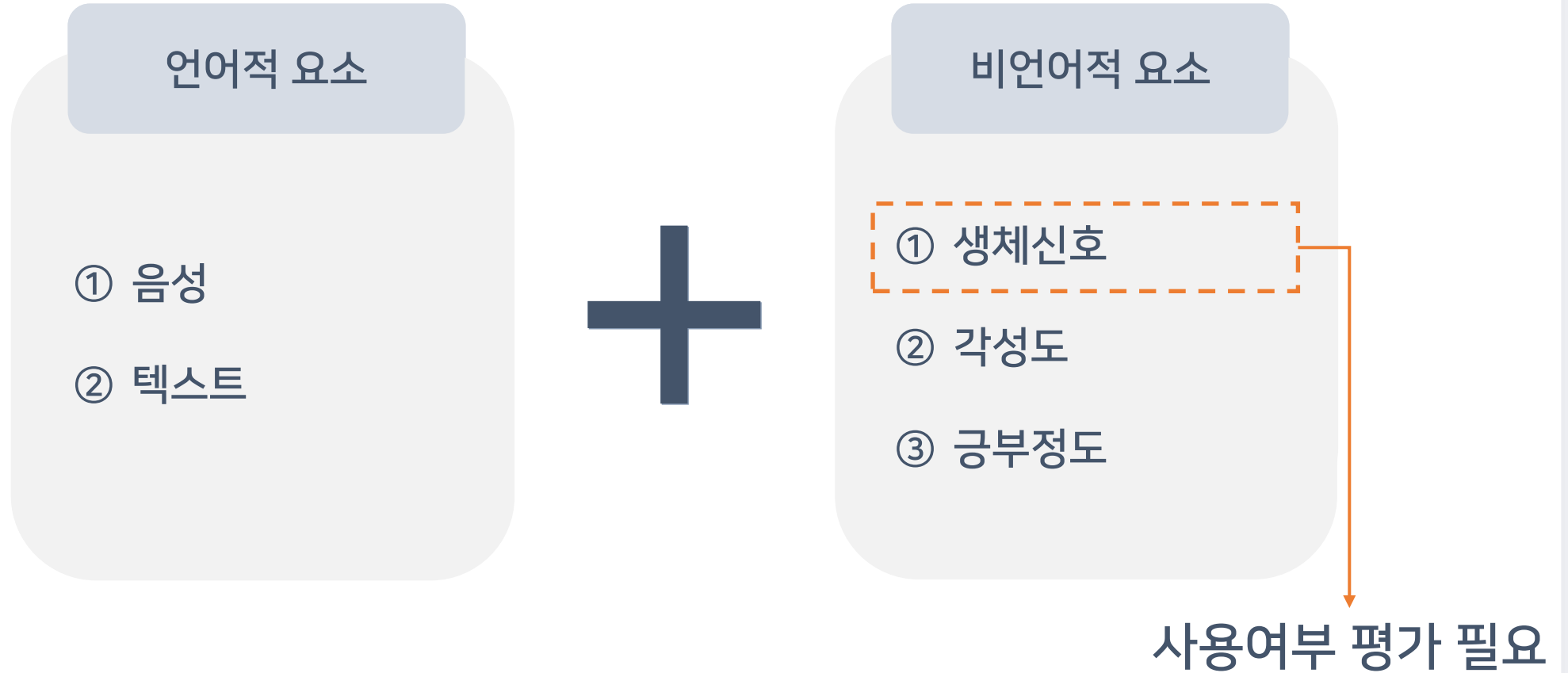
5가지
감정 선택

04

데이터 선정(모달리티)

4. 데이터 선정(모달리티)

비언어적 요소 사용여부 평가



4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

목적

- 생체신호의 변화가 발화자의 감정 분류에 유의한 영향을 주는지에 대한 분석

→ 생체신호 데이터의 사용여부 결정

4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

목적

- 생체신호의 변화가 발화자의 감정 분류에 유의한 영향을 주는지에 대한 분석

→ 생체신호 데이터의 사용여부 결정

과정

1

EDA, IBI, TEMP의 변화량과
감정의 변화량 그래프를 통해
대략적인 상관관계 확인

4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

목적

- 생체신호의 변화가 발화자의 감정 분류에 유의한 영향을 주는지에 대한 분석

→ 생체신호 데이터의 사용여부 결정

과정



4. 데이터 선정(모달리티)

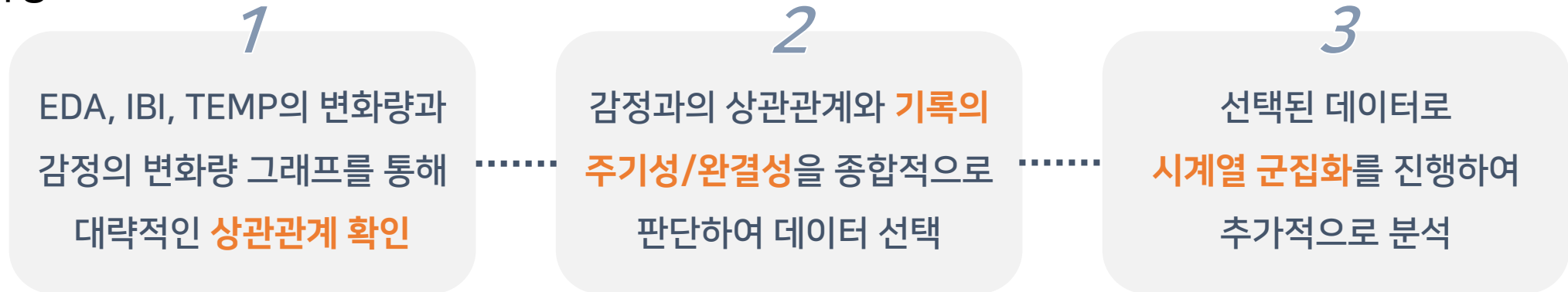
A. 생체신호 데이터 분석

목적

- 생체신호의 변화가 발화자의 감정 분류에 유의한 영향을 주는지에 대한 분석

→ 생체신호 데이터의 사용여부 결정

과정



4. 데이터 선정(모달리티)

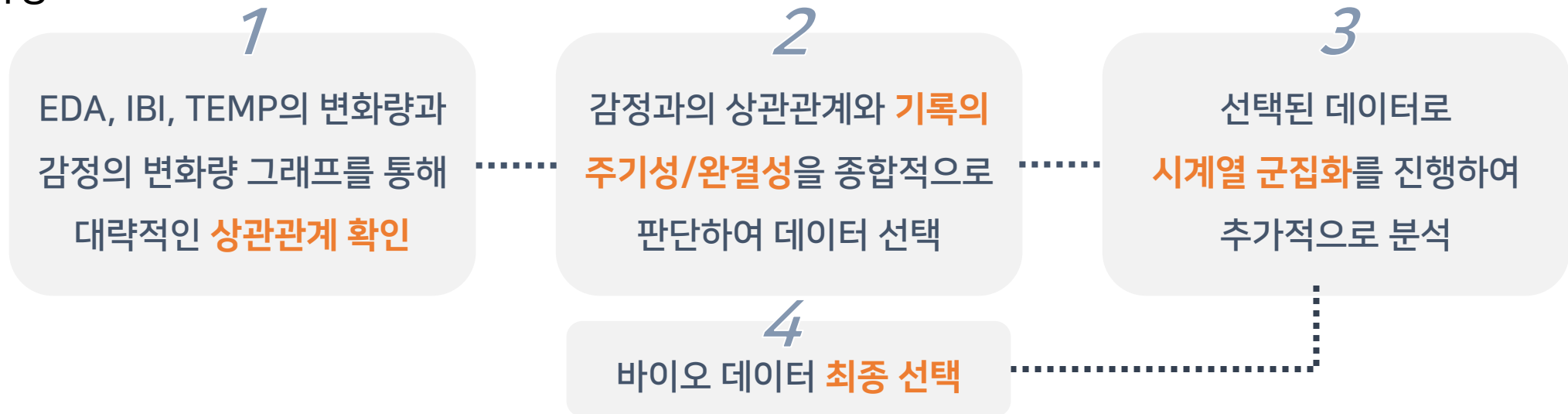
A. 생체신호 데이터 분석

목적

- 생체신호의 변화가 발화자의 감정 분류에 유의한 영향을 주는지에 대한 분석

→ 생체신호 데이터의 사용여부 결정

과정

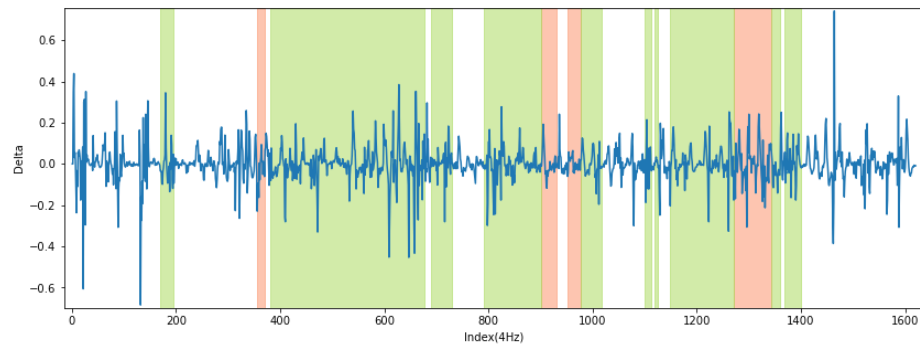


4. 데이터 선정(모달리티)

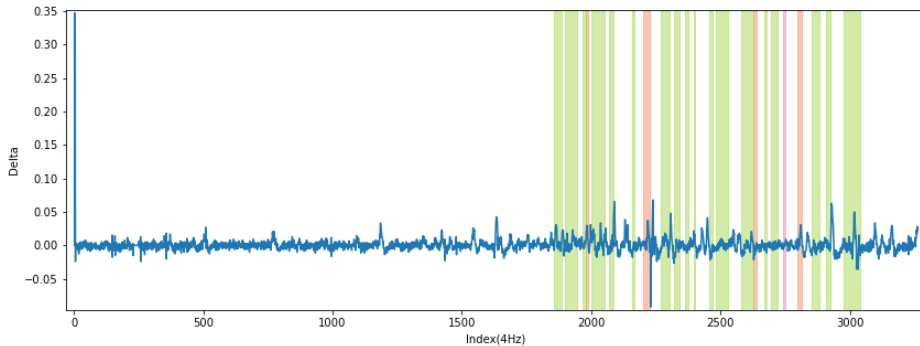
A. 생체신호 데이터와 감정의 상관관계 확인

(1) 시간에 따른 생체신호의 변화와 감정 라벨링 시각화

EDA of Session 01 - Script 05 - User 002M

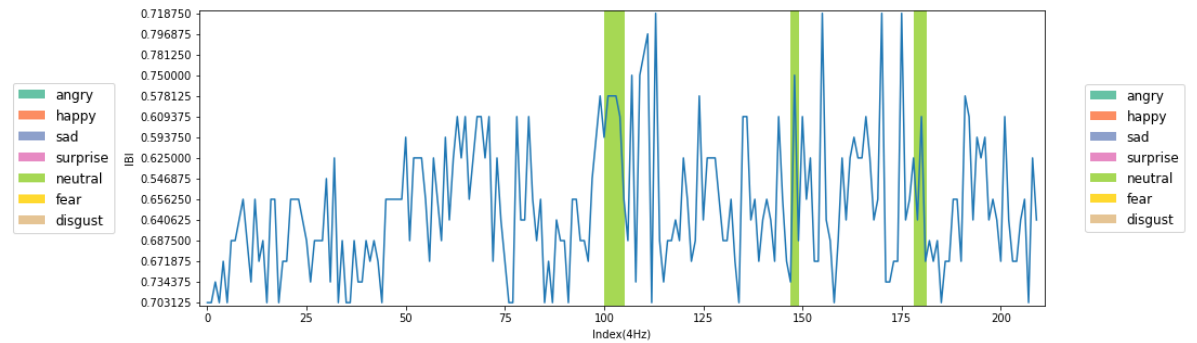


EDA of Session 29 - Script 01 - User 057F

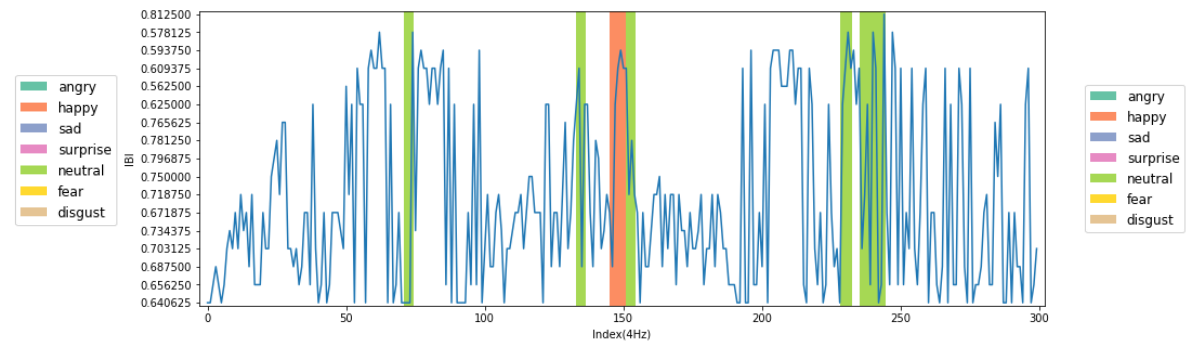


[그림] EDA와 감정 그래프

Session01 - s2_u1_f



Session01 - s1_u1_f



[그림] IBI와 감정 그래프

4. 데이터 선정(모달리티)

A. 생체신호 데이터 주기성/완결성

(2) EDA, IBI 데이터의 일부 Session에 빈 구간 존재

Arousal	Valence	Segment	TEMP	IBI	EDA
3.2	3.4	Sess11_script06_User022M_038	[33.71 33.7]	[0.609375	[9.212907 9.2065
2.7	2.4	Sess11_script06_User022M_039	[33.71 33.7]	[0.671875	[9.121938 9.121938
3.6	3.7	Sess12_script01_User081F_001	[32.75 32.7]	[]	
3.6	4	Sess12_script01_User081F_002	[32.77 32.7]	[]	
3.1	3.4	Sess12_script01_User081F_003	[32.79 32.7]	[]	
3.5	3.9	Sess12_script01_User081F_004	[32.77 32.7]	[]	
3.3	3.6	Sess12_script01_User081F_005	[32.77 32.7]	[]	
3.6	4	Sess12_script01_User081F_006	[32.79 32.7]	[]	
3.8	4	Sess12_script01_User081F_007	[32.79 32.7]	[]	
3.4	3.7	Sess12_script01_User081F_008	[32.81 32.8]	[]	
3	3.4	Sess12_script01_User081F_009	[32.83 32.8]	[]	
2.9	3.5	Sess12_script01_User081F_010	[32.84 32.8]	[]	
3.6	4.1	Sess12_script01_User081F_011	[32.84 32.8]	[]	
3.6	4.1	Sess12_script01_User081F_012	[32.84 32.8]	[]	
3.9	4.1	Sess12_script01_User081F_013	[32.84 32.8]	[]	
3.8	4	Sess12_script01_User081F_014	[32.87 32.8]	[]	
3.7	4.1	Sess12_script01_User081F_015	[32.89 32.8]	[]	
3.5	3.2	Sess12_script01_User081F_016	[32.91 32.8]	[]	
3.6	3.1	Sess12_script01_User081F_017	[32.91 32.8]	[]	
3.6	3.7	Sess12_script01_User081F_018	[32.91 32.8]	[]	
3	3.3	Sess12_script01_User081F_019	[32.91 32.8]	[]	
3.1	3	Sess12_script01_User081F_020	[32.91 32.8]	[]	
2.9	3.2	Sess12_script01_User081F_021	[32.89 32.8]	[]	

A. 생체신호 데이터 분석

결과

- ## ② 기록의 주기성/데이터의 완결성 고려

- EDA, IBI 데이터의 일부 Session에 빈 구간 존재

결과

- ① EDA, IBI의 경우 각각의 변화량과 감정과의 **상관관계가 유의하지 않다고 판단**
- ② 기록의 **주기성/데이터의 완결성** 고려
 - EDA, IBI 데이터의 일부 Session에 빈 구간 존재

4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

(2) EDA, IBI 데이터의 일부 Session에 빈 구간 존재

결과

Arousal	Valence	Segment	TEMP	IBI	EDA
3.6	3.7	Sess12_script01_User081F_001	[32.75 32.7]	0	
3.6	4	Sess12_script01_User081F_002	[32.77 32.7]	0	
3.1	3.4	Sess12_script01_User081F_003	[32.79 32.7]	0	
3.9	3.9	Sess12_script01_User081F_004	[32.77 32.7]	0	
3.3	3.6	Sess12_script01_User081F_005	[32.77 32.7]	0	
3.6	4	Sess12_script01_User081F_006	[32.79 32.7]	0	
3.8	4	Sess12_script01_User081F_007	[32.79 32.7]	0	
3.4	3.7	Sess12_script01_User081F_008	[32.81 32.7]	0	
3	3.4	Sess12_script01_User081F_009	[32.83 32.7]	0	
2.9	3.5	Sess12_script01_User081F_010	[32.84 32.7]	0	
3.6	3.6	Sess12_script01_User081F_011	[32.84 32.7]	0	
3.6	3.6	Sess12_script01_User081F_012	[32.84 32.7]	0	
3.9	4.1	Sess12_script01_User081F_013	[32.84 32.7]	0	
3.8	4	Sess12_script01_User081F_014	[32.87 32.7]	0	
3.7	4.1	Sess12_script01_User081F_015	[32.89 32.7]	0	
3.5	3.2	Sess12_script01_User081F_016	[32.91 32.7]	0	
3.6	3.1	Sess12_script01_User081F_017	[32.91 32.7]	0	
3	3.3	Sess12_script01_User081F_018	[32.91 32.7]	0	
3.1	3	Sess12_script01_User081F_020	[32.91 32.7]	0	
2.9	3.2	Sess12_script01_User081F_021	[32.89 32.7]	0	

EDA 미사용

Arousal	Valence	Segment	TEMP	IBI	EDA
3.1	3	Sess01_script01_User001F_013	[34.59 34.5]	0	
3.4	3.7	Sess01_script01_User001F_015	[34.59 34.5]	0	
2.9	3.4	Sess01_script01_User001M_001	[34.81 34.5]	0	
2.9	3.1	Sess01_script01_User002M_002	[34.79 34.7]	0	
3	3.1	Sess01_script01_User002M_003	[34.75 34.7]	0	
3.1	3.7	Sess01_script01_User002M_004	[34.77 34.7]	0	
3.2	3	Sess01_script01_User002M_005	[34.75 34.7]	0	
3.3	3.3	Sess01_script01_User002M_006	[34.75 34.7]	0	
3.6	3.9	Sess01_script01_User002M_008	[34.75 34.7]	0	
3.7	4.1	Sess01_script01_User002M_009	[34.75 34.7]	0	
3.4	3.7	Sess01_script01_User002M_010	[34.77 34.7]	0	
3.6	4.2	Sess01_script01_User002M_011	[34.77 34.7]	0	
3	3.1	Sess01_script01_User002M_012	[34.79 34.7]	0	
2.9	3	Sess01_script01_User002M_013	[34.79 34.7]	0	
3	3.3	Sess01_script01_User002M_015	[34.83 34.5]	0	
2.7	2.9	Sess01_script01_User002M_016	[34.84 34.5]	0	
2.8	3.2	Sess01_script01_User002M_017	[34.83 34.5]	0	
2.9	2.9	Sess01_script01_User002M_018	[34.89 34.5]	0	
3	3.1	Sess01_script01_User002M_019	[34.87 34.5]	0	

IBI 미사용

Arousal	Valence	Segment	TEMP	IBI	EDA
7	2.2	1851	7.260261	7.275625	7.29355
8	5.9	7441	8.802428	8.839558	8.833156
2	3.5	6493	2.788578	2.678377	2.652749
2	5.4	7035	2.653442	2.631658	2.614999
2	3.9	6217	2.975664	3.02692	3.071769
3	5.8	9019	3.692863	3.659546	3.409672
3	4.4	2096	3.49937	3.532687	3.53525
3	4.6	9006	3.50934	3.463209	3.457084
3	4.2	6492	2.676422	2.627642	2.626998
3	3.3	0224	3.33329	3.358415	3.337154
3	5.4	5451	3.823566	3.986305	4.090631
3	5.0	8822	3.375605	3.255903	3.157234
3	5.0	834	3.51731	3.641607	3.744119
4	2.0	0224	4.150249	4.092586	4.111806
4	3.7	3214	4.380902	4.386028	4.396279
4	5.3	5627	5.443366	5.470275	5.401079
5	5.7	0175	5.754748	5.771406	5.740652
5	5.1	653	5.140953	5.080727	5.206305
5	5.1	3218	5.92475	5.612086	5.284046
6	1.0	2866	6.105429	5.952941	5.800454

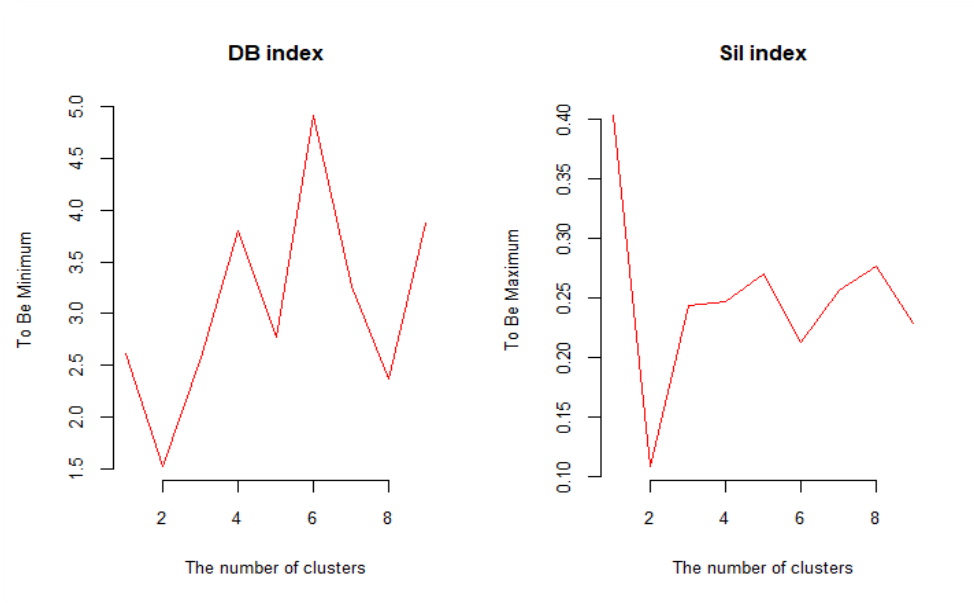
TEMP 사용

4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

(3) TEMP 데이터의 DTW 시계열 군집화

- ① 발화 데이터의 특성상 데이터의 길이가 모두 상이 → DTW(Dynamic Time Warping) 적용
- ② 클러스터 개수 선택을 위해 Davis-Bouldin Index / Silhouette Index 비교분석



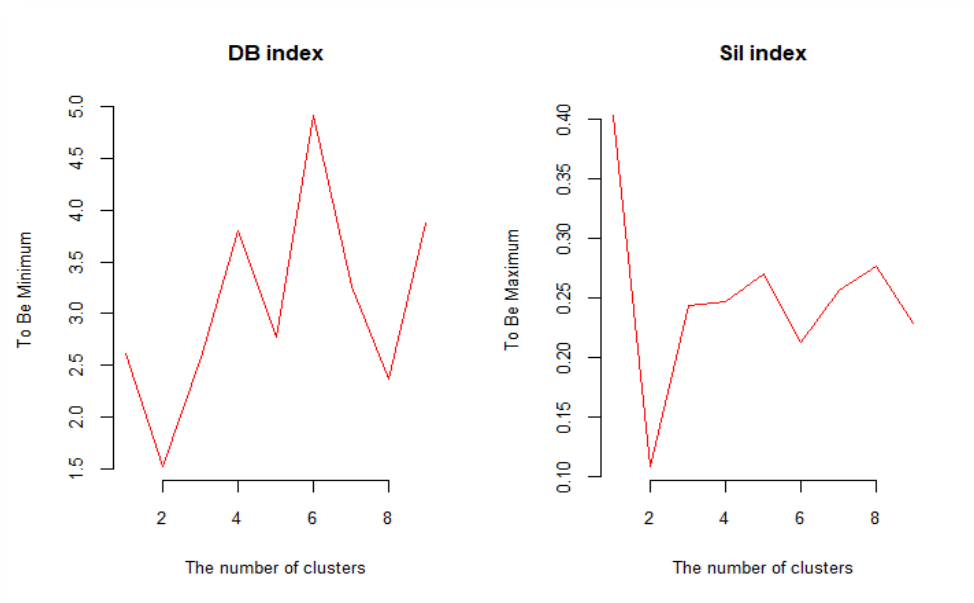
[그림] Davis-Bouldin Index/Silhouette Index

4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

(3) TEMP 데이터의 DTW 시계열 군집화

- ① 발화 데이터의 특성상 데이터의 길이가 모두 상이 → DTW(Dynamic Time Warping) 적용
- ② 클러스터 개수 선택을 위해 Davis-Bouldin Index / Silhouette Index 비교분석



[그림] Davis-Bouldin Index/Silhouette Index

각 Index 비교분석 결과

+

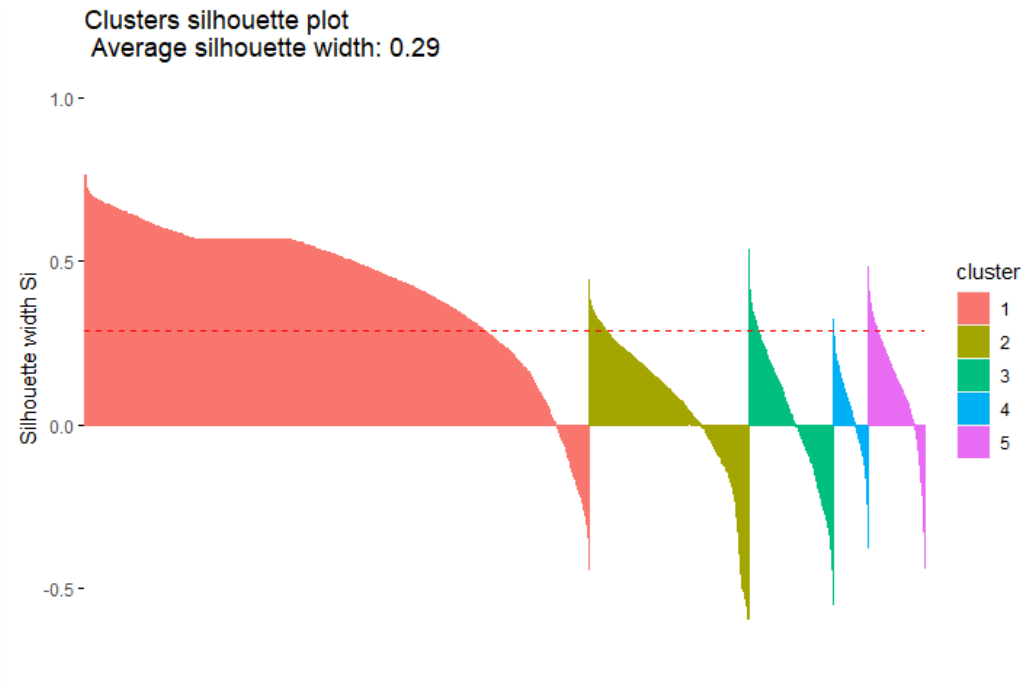
감정 라벨 5가지

클러스터 개수: 5개

4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

(3) TEMP 데이터의 DTW 시계열 군집화



[그림] Silhouette Analysis 결과

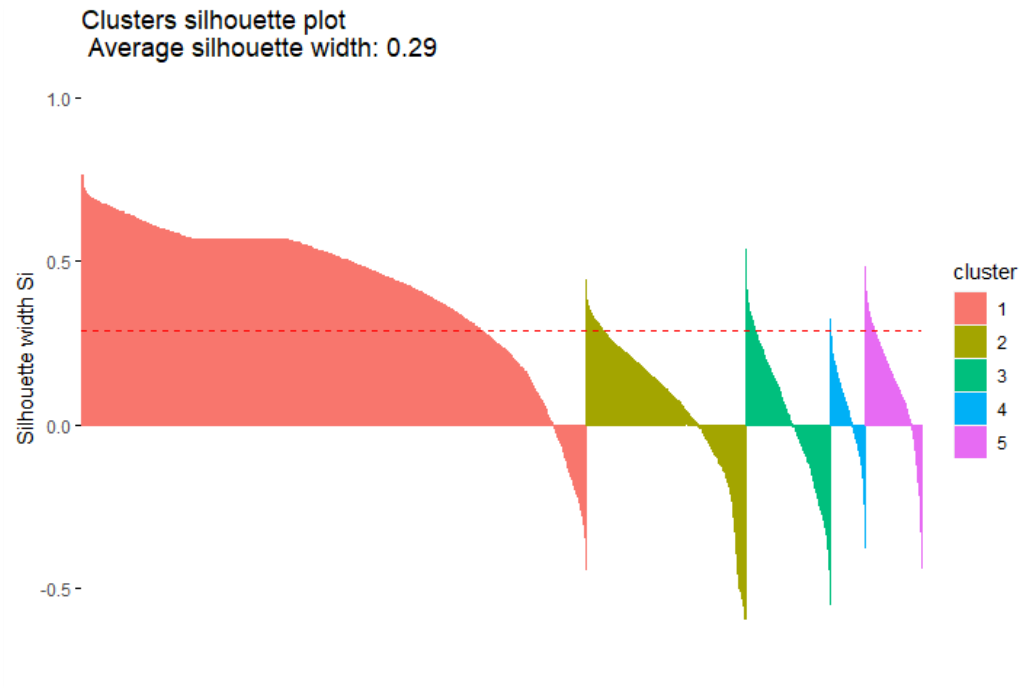
① 실루엣 계수가 전체적으로 낮은 값을 보임

② 음수인 값 존재

4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

(3) TEMP 데이터의 DTW 시계열 군집화



[그림] Silhouette Analysis 결과

① 실루엣 계수가 전체적으로 낮은 값을 보임

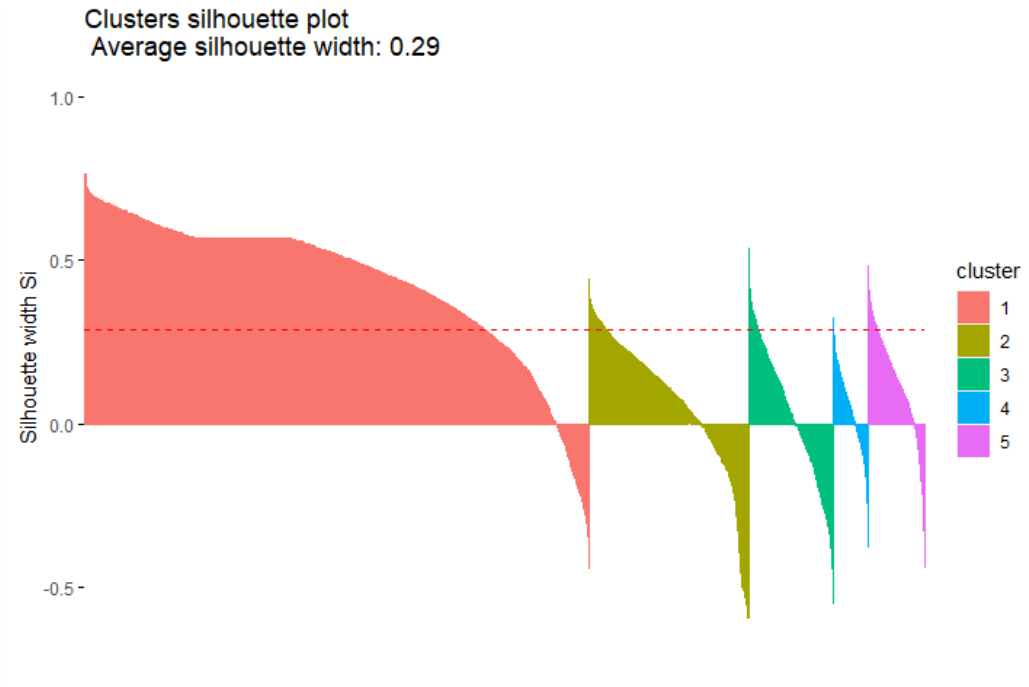
② 음수인 값 존재

피부 온도는 감정과 유의한 관계를 가지지 않는다

4. 데이터 선정(모달리티)

A. 생체신호 데이터 분석

(3) TEMP 데이터의 DTW 시계열 군집화



[그림] Silhouette Analysis 결과

① 실루엣 계수가 전체적으로 낮은 값을 보임

② 음수인 값 존재

피부 온도는 감정과 유의한 관계를 가지지 않는다

생체신호 데이터 **미사용**

4. 데이터 선정(모달리티)

비언어적 요소 사용여부 평가



4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

목적

- 5개의 감정 라벨 그룹에 따라 각성도 및 공부정도 데이터가 유의한 차이를 보이는지 검정

→ **각성도/공부정도 데이터의 사용여부** 결정 과정

4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

목적

- 5개의 감정 라벨 그룹에 따라 각성도 및 공부정도 데이터가 유의한 차이를 보이는지 검정

→ **각성도/공부정도 데이터의 사용여부** 결정 과정

과정

1

정규성·등분산성 검정

4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

목적

- 5개의 감정 라벨 그룹에 따라 각성도 및 공부정도 데이터가 유의한 차이를 보이는지 검정

→ **각성도/공부정도 데이터의 사용여부** 결정 과정

과정



4. 데이터 선정(모달리티)

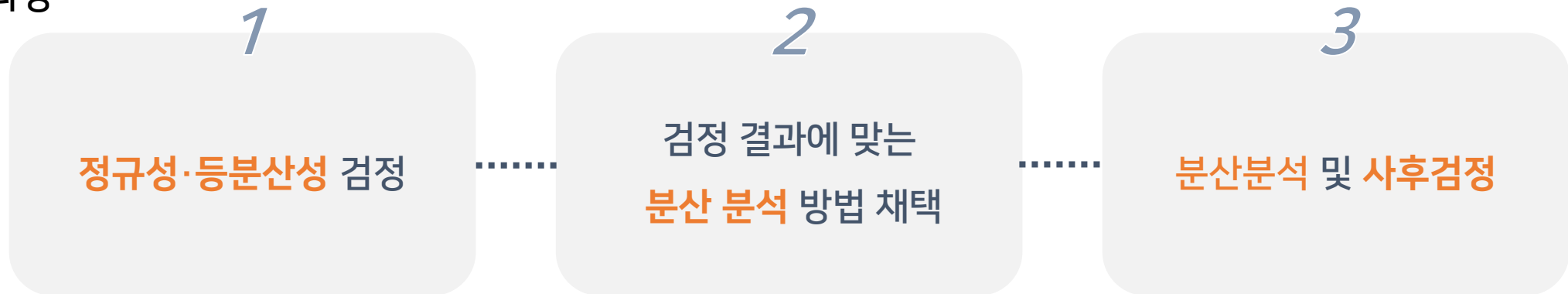
B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

목적

- 5개의 감정 라벨 그룹에 따라 각성도 및 공부정도 데이터가 유의한 차이를 보이는지 검정

→ **각성도/공부정도 데이터의 사용여부** 결정 과정

과정



4. 데이터 선정(모달리티)

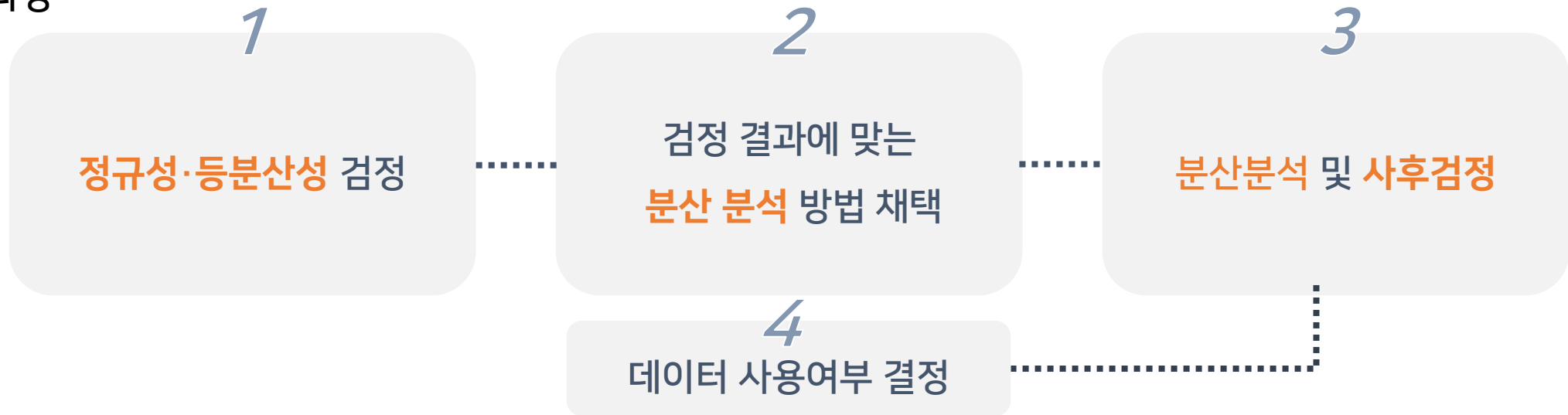
B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

목적

- 5개의 감정 라벨 그룹에 따라 각성도 및 공부정도 데이터가 유의한 차이를 보이는지 검정

→ **각성도/공부정도 데이터의 사용여부** 결정 과정

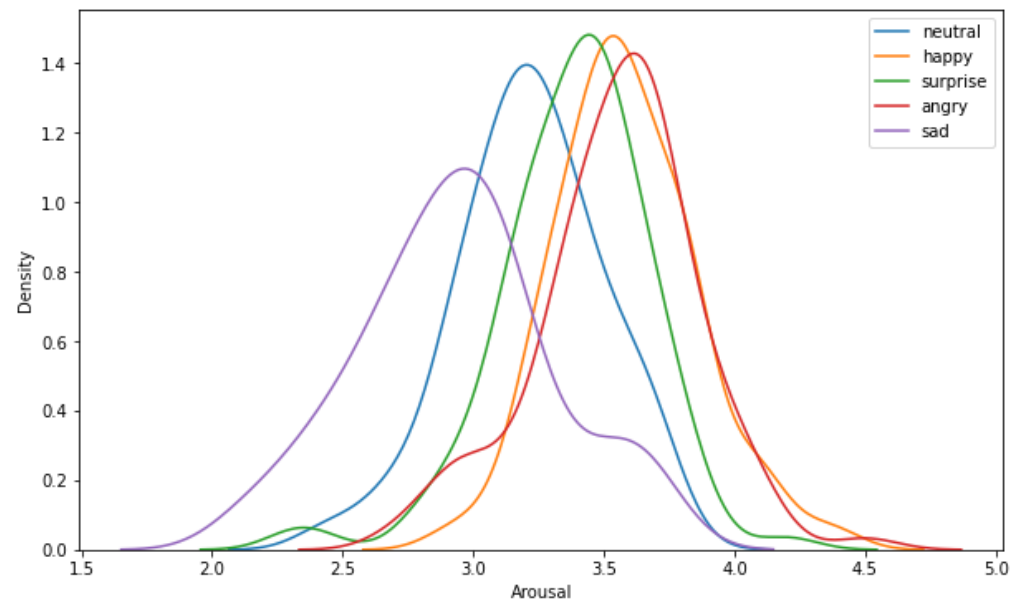
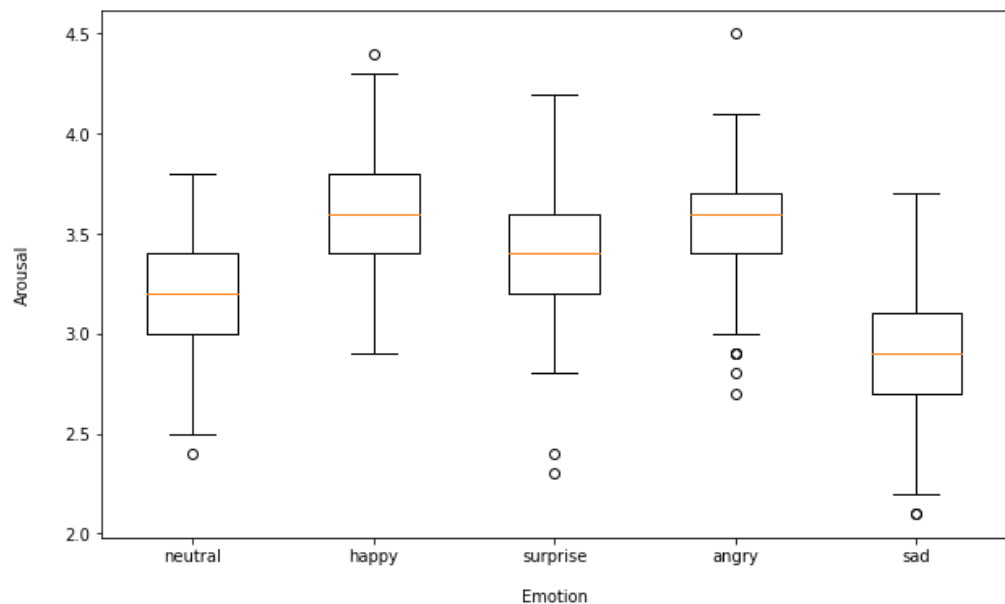
과정



4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

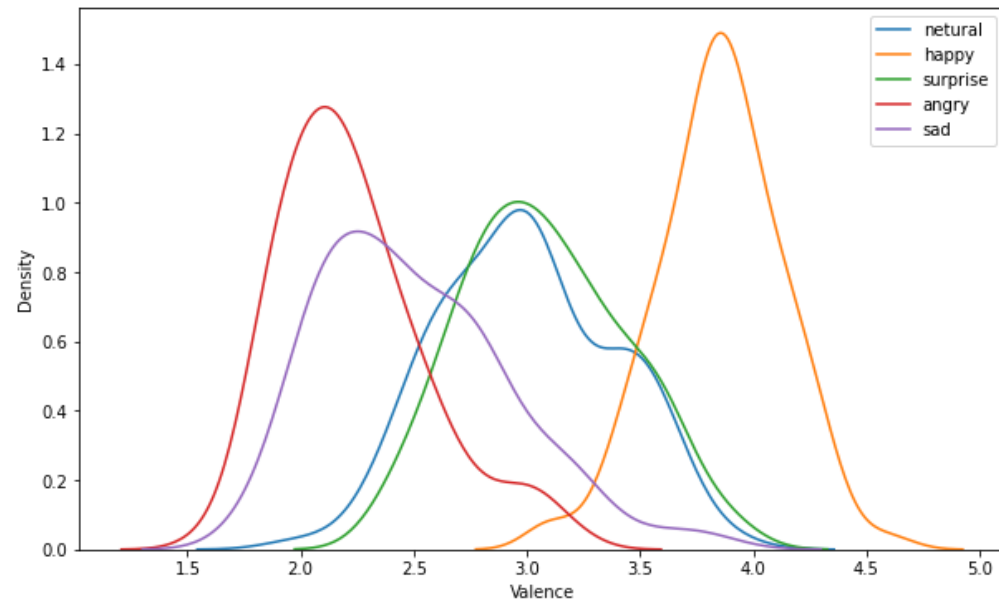
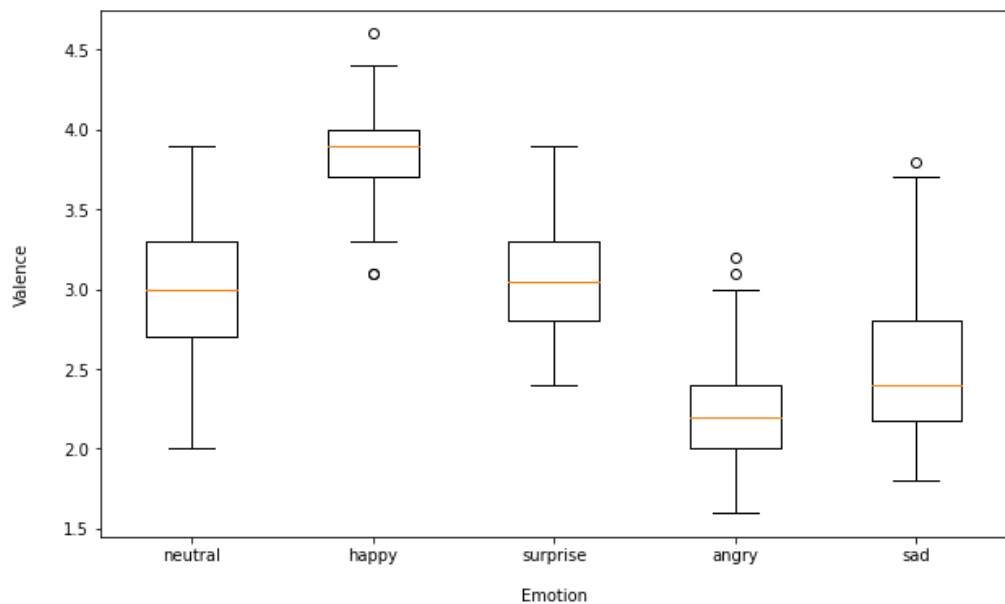
(0) 시각화 - 5개 감정 그룹별 각성도(Arousal)



4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

(0) 시각화 - 5개 감정 그룹별 공부정도(Valence)



4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

(1) 정규성·등분산성 검정 - 정규성 검정

정규성 만족 여부 (0 / X)	<div>$n < 10$$10 \leq n \leq 30$$30 < n$</div>		
정규성 검정 0	X	0	0
정규성 검정 X	X	X	0

4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

(1) 정규성·등분산성 검정 - 정규성 검정

정규성 만족 여부 (0 / X)	$n < 10$	$10 \leq n \leq 30$	$30 < n$
정규성 검정 0	X	0	0
정규성 검정 X	X	X	0

감정 라벨별 표본의 수가 100개 이상이므로, 정규성 가정 만족

4. 데이터 선정(모달리티)

B. 각성도/금부정도 데이터 분석 - 분산 분석 (ANOVA)

(1) 정규성·등분산성 검정 - 등분산성 검정

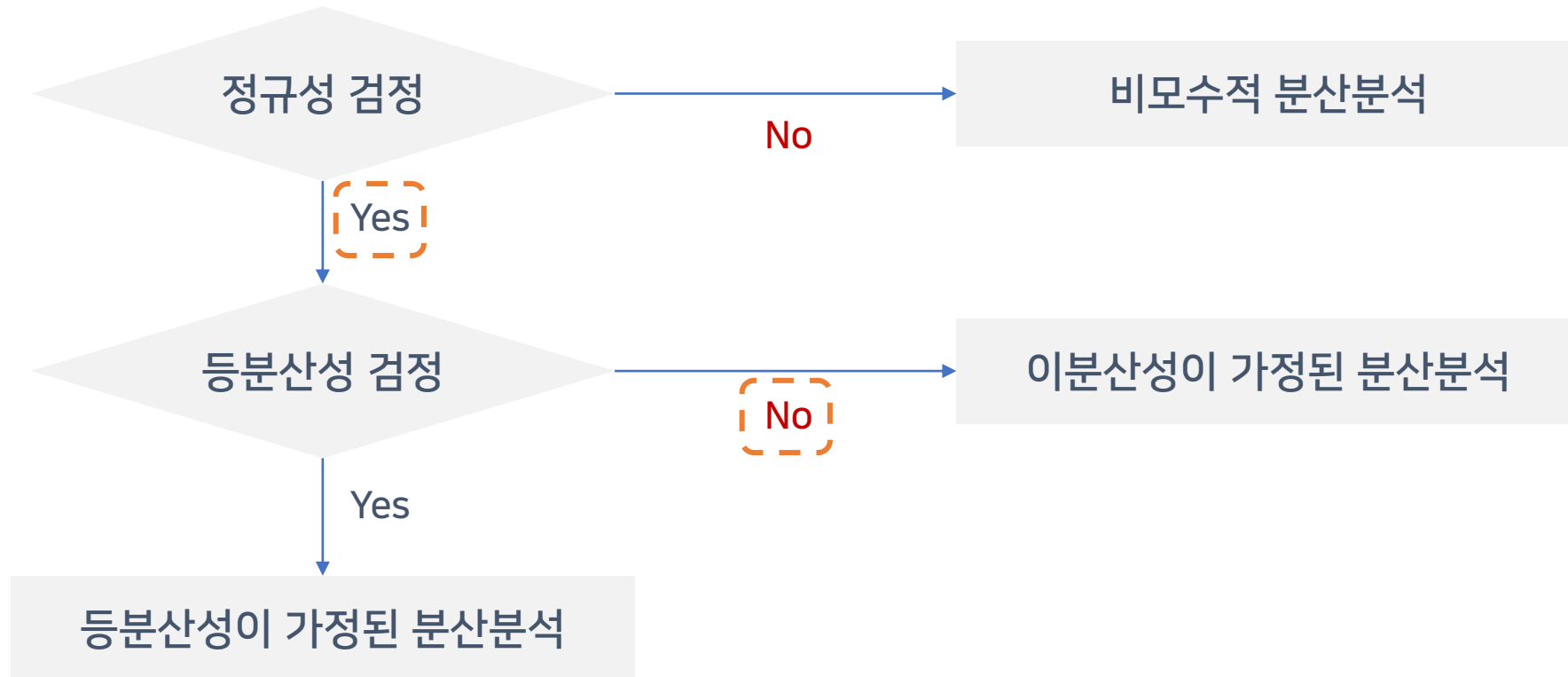
등분산성 검정 (P-value)	Barlett 검정 (정규성 만족시)		Levene 검정 (정규성 만족 무관)	
각성도	1.7815e-05		1.3512e-05	
금부정도	9.0970e-29		3.8527e-23	

모든 항목에서 $P\text{-value} < 0.05$ 이므로, 등분산성 만족 X

4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

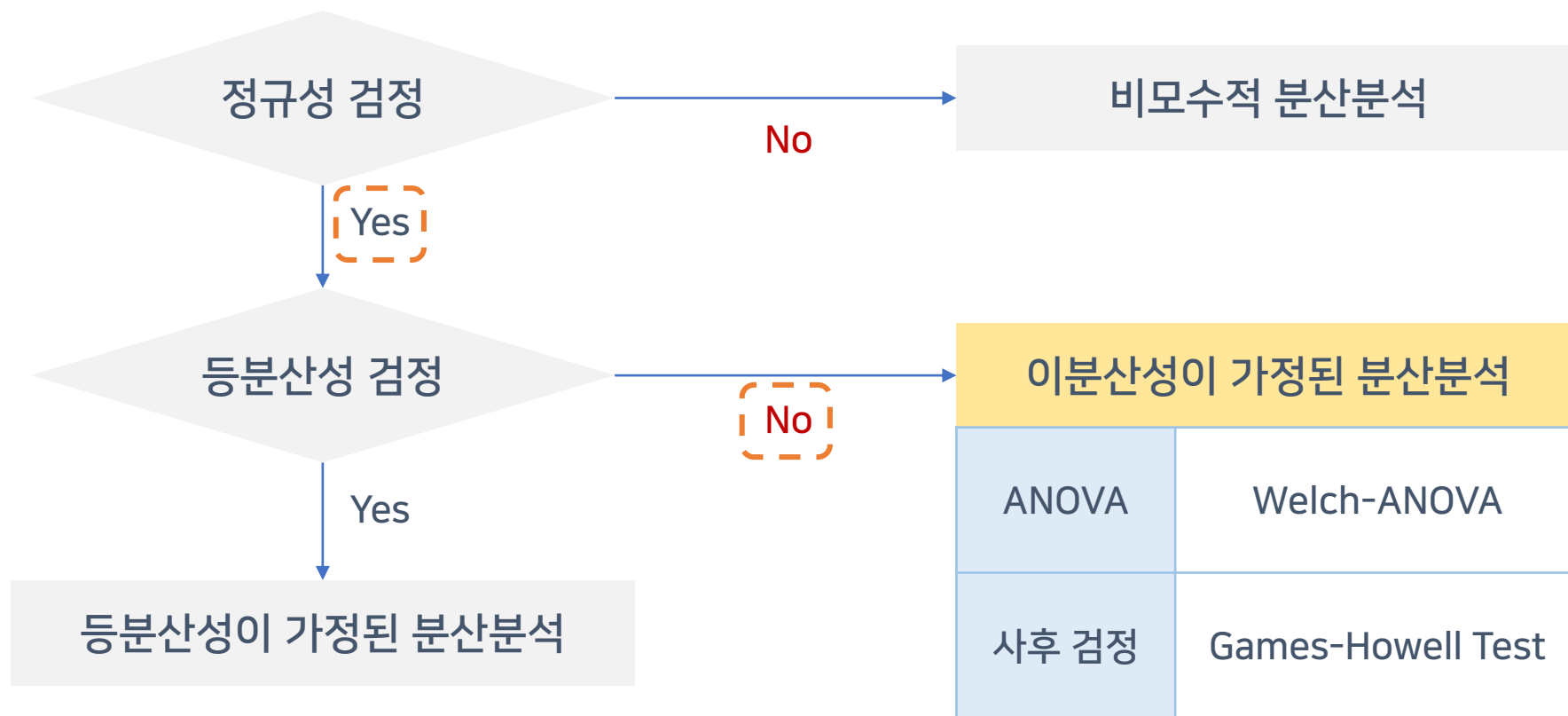
(2) 분산분석 방법 채택



4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

(2) 분산분석 방법 채택



4. 데이터 선정(모달리티)

B. 각성도/긍부정도 데이터 분석 - 분산 분석 (ANOVA)

(3) 분산분석 및 사후 검정 - 감정 **각성도** 분산분석 (Welch-ANOVA)

Source	ddof2	ddof2	F-value	p-value	np2
감정 라벨	4	246.83	63.12	1.1026e-36	0.379

p-value < 0.05

다섯 집단(neutral, happy, surprise, sad, angry)의 **각성도 평균에 유의한 차이가 존재**

4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

(3) 분산분석 및 사후 검정 - 감정 **공부정도** 분산분석 (Welch-ANOVA)

Source	ddof1	ddof2	F-value	p-value	np2
감정 라벨	4	245.7521	422.26	8.3110e-109	0.702

p-value < 0.05

다섯 집단(neutral, happy, surprise, sad, angry)의 **공/부정도 평균에 유의한 차이**가 존재

4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

(3) 분산분석 및 사후 검정 - 감정 **각성도** 사후 검정 (Games-Howell Test)

p-value < 0.05

A	B	mean(A)	mean(B)	diff	P-value
angry	happy	3.560	3.583	-0.023	9.782316e-01
angry	neutral	3.560	3.213	0.347	0.000000e+00
angry	sad	3.560	2.939	0.621	8.693046e-14
angry	surprise	3.560	3.380	0.180	9.345341e-05
⋮	⋮	⋮	⋮	⋮	⋮
neutral	sad	3.213	2.939	0.274	6.231712e-07
sad	surprise	2.939	3.380	-0.441	0.000000e+00

모든 경우에서 **P-value < 0.05**이므로, 모든 감정 라벨 간 각성도의 평균에 차이가 존재

4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산 분석 (ANOVA)

(3) 분산분석 및 사후 검정 - 감정 공부정도 사후 검정 (Games-Howell Test)

p-value < 0.05

A	B	mean(A)	mean(B)	diff	P-value
angry	happy	2.212	3.824	-1.612	3.663736e-15
angry	neutral	2.212	2.995	-0.783	2.253753e-14
angry	sad	2.212	2.508	-0.296	3.431279e-07
angry	surprise	2.212	3.093	-0.881	0.000000e+00
⋮	⋮	⋮	⋮	⋮	⋮
neutral	sad	2.995	2.508	0.487	2.287059e-14
sad	surprise	2.508	3.093	-0.585	0.000000e+00

모든 경우에서 P-value < 0.05이므로, 모든 감정 라벨 간 공부정도의 평균에 차이가 존재

4. 데이터 선정(모달리티)

B. 각성도/공부정도 데이터 분석 - 분산분석 (ANOVA)

모든 감정 라벨 간 각성도, 공/부정도에는

(3) 분산분석 및 사후 검정 - 감정 공부정도 사후 검정 (Games-Howell Test)

p-value < 0.05

유의한 차이가 존재

A	B	mean	df	t	P-value
angry	happy	3.56	3.583	-0.023	9.782316e-01
angry	neutral	3.56	3.213	0.347	0.000000e+00
angry	sad	3.56	2.939	0.621	8.693046e-14
angry	surprise	3.56	3.380	0.180	9.345341e-05
⋮	⋮	⋮	⋮	⋮	⋮
neutral	sad	3.21	3.38	3.380	5.156983e-04
sad	surprise	2.13	3.38	3.380	0.000000e+00

각성도, 공/부정도 데이터 채택

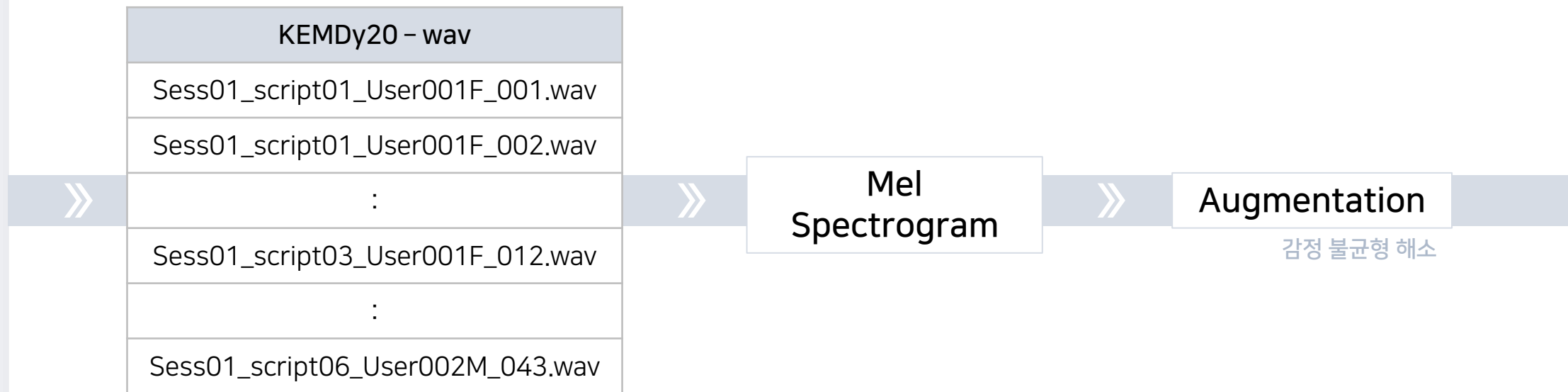
모든 경우에서 P-value < 0.05이므로, 모든 감정 라벨 간 공부정도의 평균에 차이가 존재

05

데이터 전처리

5. 데이터 전처리

음성 데이터 전처리 흐름



5. 데이터 전처리

A. 음성 데이터 전처리

1

음성파일(.wav)을 **Mel Spectrogram**으로 처리

- ① 멜(mel) 필터 수 = 128
- ② 최소 주파수 = 100
- ③ 최대 주파수 = 8000

5. 데이터 전처리

A. 음성 데이터 전처리

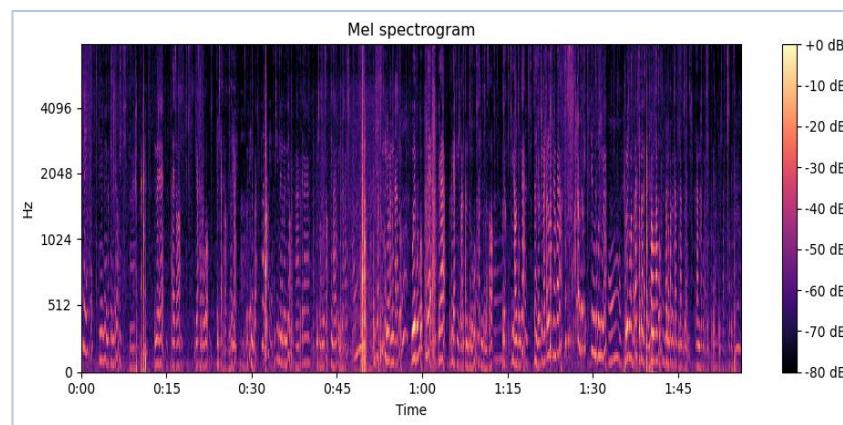
1

음성파일(.wav)을 **Mel Spectrogram**으로 처리

- ① 멜(mel) 필터 수 = 128
- ② 최소 주파수 = 100
- ③ 최대 주파수 = 8000

2

음성의 변화 강조를 위해 Mel Spectrogram에
대한 **1차 차분값** 함께 활용

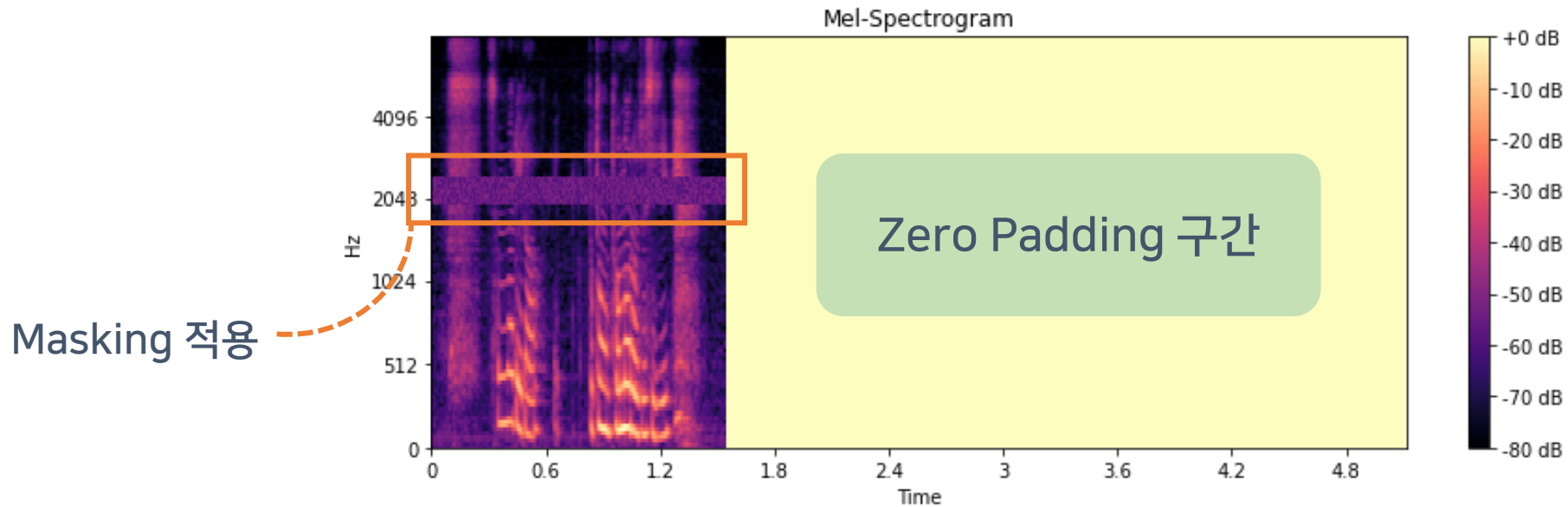


5. 데이터 전처리

B. 음성 데이터 증강

과정

- ① Mel Spectrogram 변환 후 일정 주파수 영역을 가리는 **Frequency Masking 기법** 적용

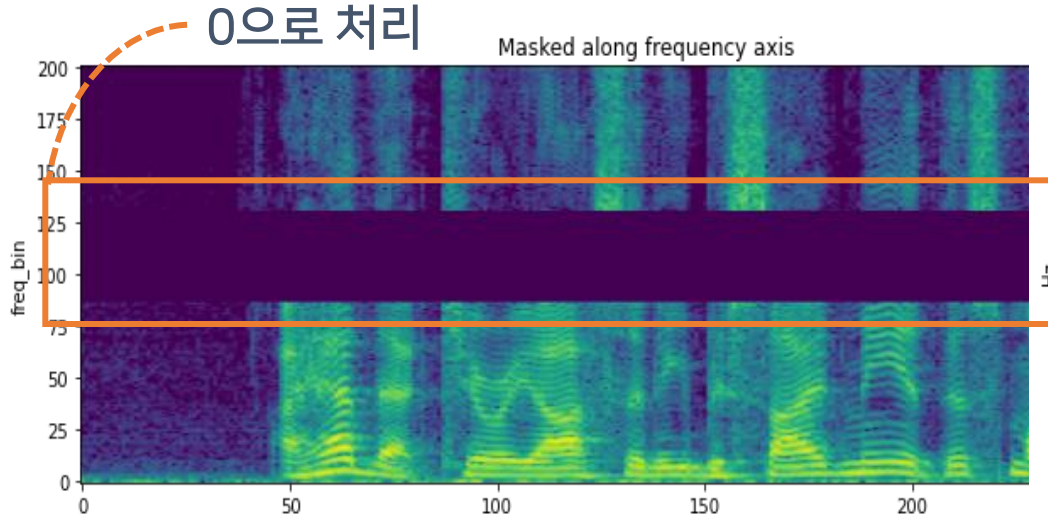


5. 데이터 전처리

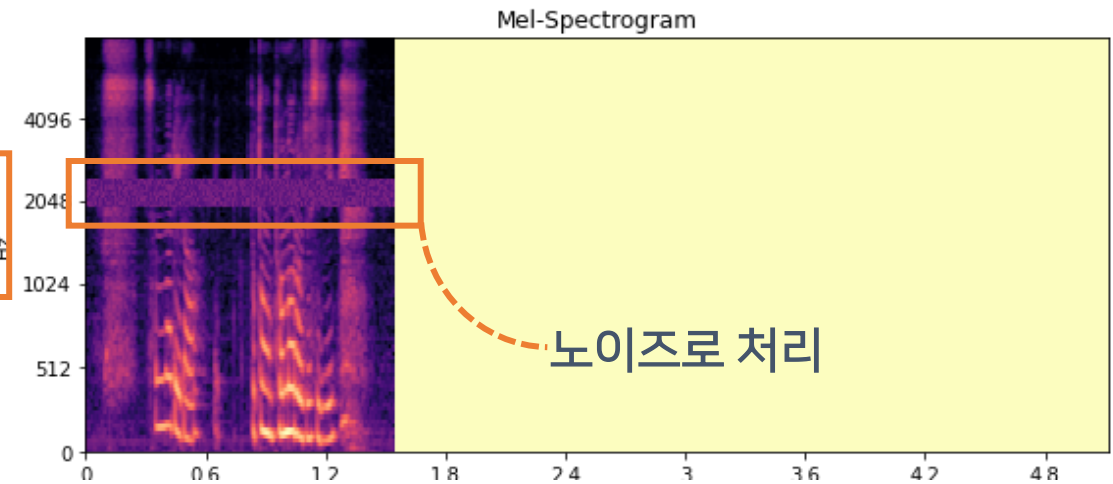
B. 음성 데이터 증강

과정

- ① Mel Spectrogram 변환 후 일정 주파수 영역을 가리는 **Frequency Masking 기법** 적용
- ② 랜덤 노이즈를 추가하는 **Random Frequency Masking** 도입 → 일반화 성능 향상 시도



[그림] 기존 Frequency Masking 결과



[그림] Random Frequency Masking 결과

5. 데이터 전처리

텍스트 데이터 전처리 흐름

./wav/~/.txt

: 발화세그먼트의 음성 이외 사운드상황에 대한 태깅

- c/ : 휴지구간이 확보되지 않은 연속발성(0.3초 미만)
- n/ : 발성 이외의 단발적인 소음
- N/ : 음성 구간의 50% 이상 잡음이 포함된 상황
- u/ : 단어의 내용을 알아 들을 수 없는 상황
- l/ : 발성중 음음 소리가 포함된 상황 (small 'L')
- b/ : 발성 중 숨소리, 김침 소리가 포함된 상황
- * : 단어 중 일부만 알아 듣거나 알아들었으나 애매한 상황
- + : 발성 중 말을 반복적으로 더듬는 상황
- / : 간투사

토큰화

Augmentation

감정 불균형 해소

5. 데이터 전처리

C. 텍스트 데이터 전처리

1

텍스트 분석에 도움이 되지 않는 음성 관련
효과음들에 대한 표시들을 모두 **불용어**로 지정

불용어 리스트

C/	휴지구간이 확보되지 않은 연속발성
n/	발성 이외의 단발적 소음
N/	음성 구간의 50% 이상 잡음이 포함
u/	단어의 내용을 알아들을 수 없음
l/	발성 중 음음 소리가 포함
b/	발성 중 숨소리, 기침 소리가 포함
*	단어 중 일부만 알아듣거나 애매한 상황
+	발성 중 말을 반복적으로 더듬는 상황
/	간투사

5. 데이터 전처리

C. 텍스트 데이터 전처리

1

텍스트 분석에 도움이 되지 않는 음성 관련
효과음들에 대한 표시들을 모두 **불용어**로 지정

불용어 리스트

불용어 리스트	
C/	휴지구간이 확보되지 않은 연속발성
n/	발성 이외의 단발적 소음
N/	음성 구간의 50% 이상 잡음이 포함
u/	단어의 내용을 알아들을 수 없음
l/	발성 중 음음 소리가 포함
b/	발성 중 숨소리, 기침 소리가 포함
*	단어 중 일부만 알아듣거나 애매한 상황
+	발성 중 말을 반복적으로 더듬는 상황
/	간투사

2

불용어 **제거**

Text

Text
n/ 아 친구들도? l/
l/ 나는 생일?
생일날이면은 b/ 내가 고기를 되...
b/ 선물이라 이 보통 돈으로 주...
:
u/ 그걸 점점 발전을 시키잖아, 사 람처럼 느껴지게.
뭐 진짜 막 어떤 욕심이 가득 찬 사 람은 내가...



Text

Text
아 친구들도?
나는 생일?
생일날이면은 내가 고기를 되...
선물이라 이 보통 돈으로 주...
:
그걸 점점 발전을 시키잖아, 사 람처럼 느껴지게.
뭐 진짜 막 어떤 욕심이 가득 찬 사 람은 내가...

5. 데이터 전처리

C. 텍스트 데이터 전처리

3

텍스트 데이터 **토큰화(KoBERT tokenizer)**

너는 대선 때 투표할 수 있어?



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

5. 데이터 전처리

C. 텍스트 데이터 전처리

3

텍스트 데이터 **토큰화(KoBERT tokenizer)**

너는 대선 때 투표할 수 있어?



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

4

최대 길이를 **64**로 지정 / **Zero Padding** 적용

Text(최종 텍스트 데이터)
[_어, _나는, _워낙, _많아, _가지고, _고, 등, _, 고등학교, _3, ...
[_근, 데, _이제, _진, 지, _이, 거, _진짜, 예, 요, _진짜, _꾸,...
[_그리고, _그, 걸, _또, _그, 걸, _또, _빌려, 달라, 고, _하는, ...
[_아, _뭐, 었, 지, _그, .]
:
[_그때, 도, _아, 마, _, 밌, 었, 었, 을, _때, 었, 으니, 까, _...
[_그냥, _오히려, _군, 대, _둘, _다, _, 갔다, _, 왔, 잤, 아, ...

5. 데이터 전처리

D. 텍스트 데이터 증강

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

5. 데이터 전처리

D. 텍스트 데이터 증강

텍스트 데이터 증강방법

방법	문장
원래 문장	중심을 잃고 목소리도 잃고 비난받고 사람들과 멀어지는 착각 속에
SR (동사의 대체)	백심을 잃고 목소리도 잃고 비난받고 사람들과 멀어지는 착각 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

SR, RI와 같은 단어 대체 방법은
문장의 의미를 왜곡할 수 있다고 판단!

5. 데이터 전처리

D. 텍스트 데이터 증강

텍스트 데이터 증강방법

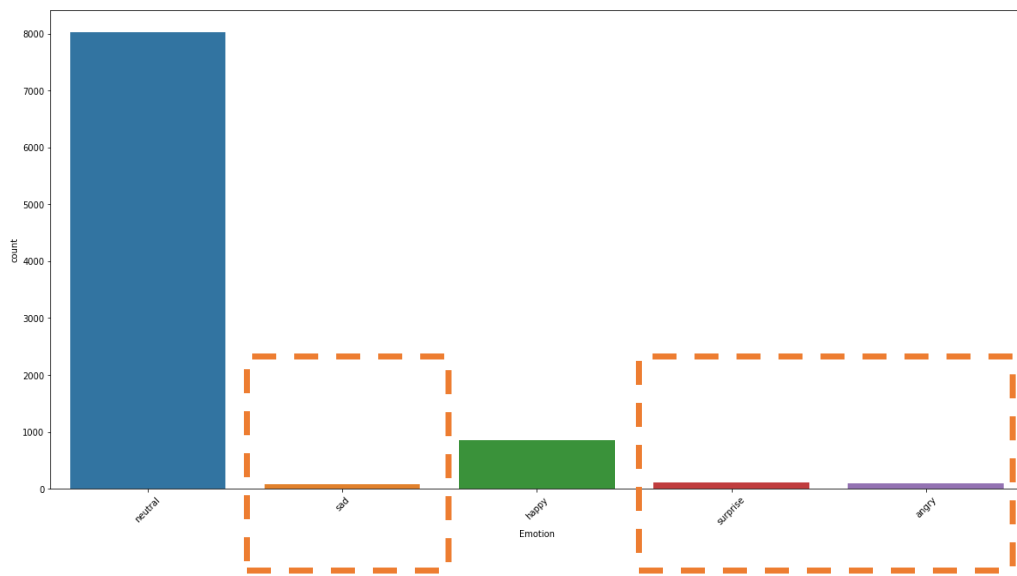
방법	문장
원래 문장	중심을 잃고 목소리도 잃고 비난받고 사람들과 멀어지는 착각 속에
SR (동의어 대체)	핵심을 잃고 목소리도 잃고 비판받고 인간들과 멀어지는 혼란 속에
RI (단어 임의 대체)	돈을 잃고 목소리도 잃고 비난받고 사람들과 멀어지는 착각 속에
RS (단어 위치 변경)	중심을 잃고 사람들과 잃고 비난받고 목소리도 멀어지는 착각 속에
RD (단어 임의 삭제)	중심도 목소리도 사람들과 멀어지는 착각

의미 왜곡 위험이 적은 RS, RD 선택

5. 데이터 전처리

D. 텍스트 데이터 증강

텍스트 데이터 증강 결과

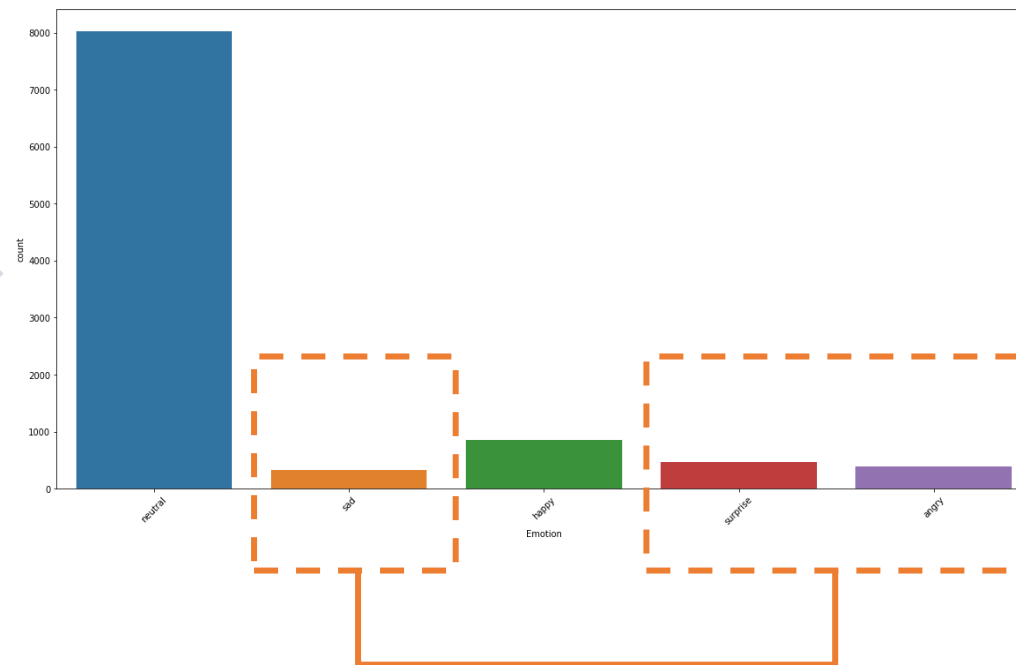
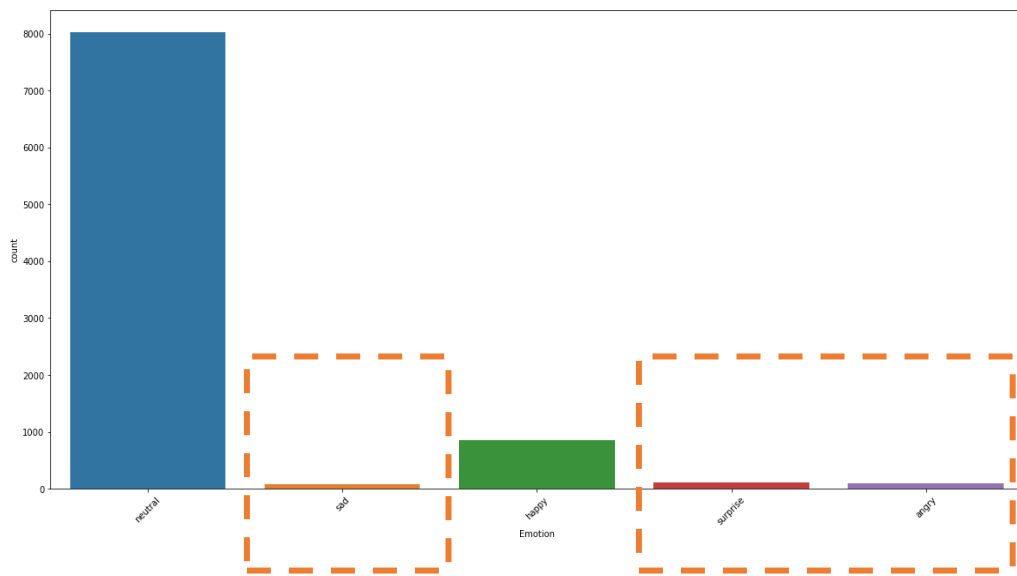


- ① 음성 데이터 증강 50%
- ② 텍스트 데이터 증강 50%
- ③ 음성 + 텍스트 데이터 증강 50%

5. 데이터 전처리

D. 텍스트 데이터 증강

텍스트 데이터 증강 결과



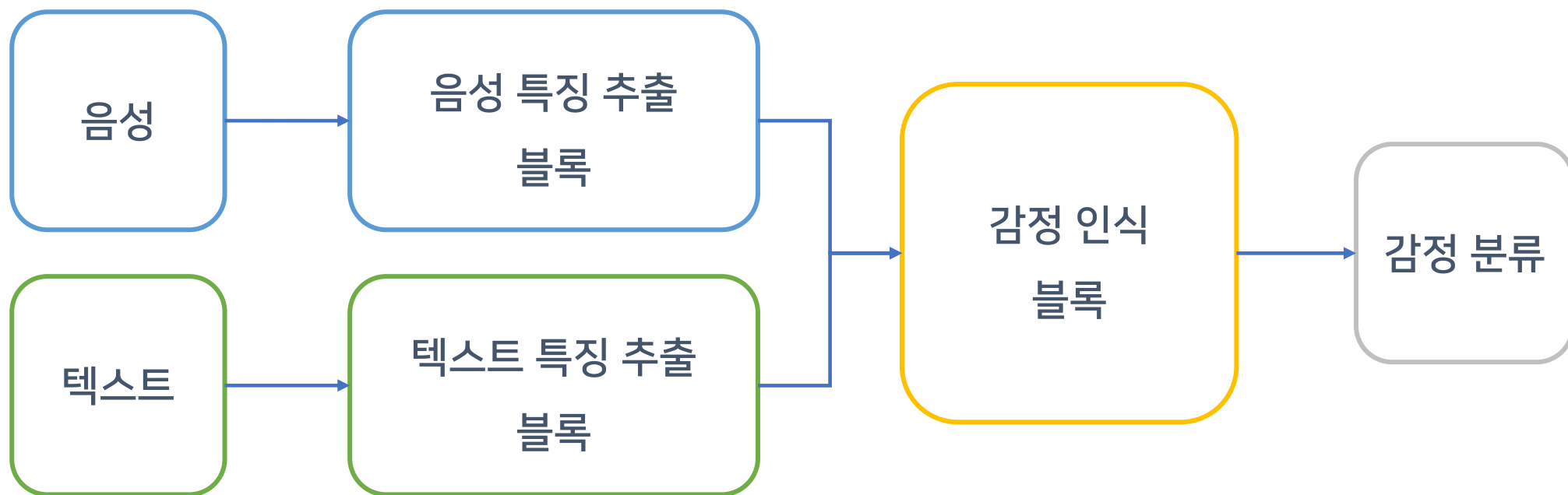
라벨별 250% 증강

06

모델 구성

6. 모델 구성

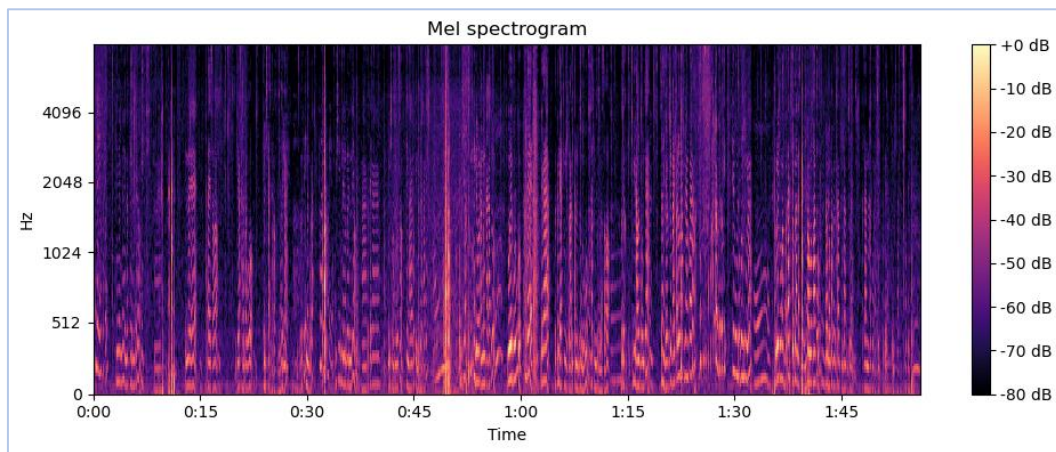
A. 베이스라인 모델 선정



6. 모델 구성

A. 베이스라인 모델 선정

음성(Audio)



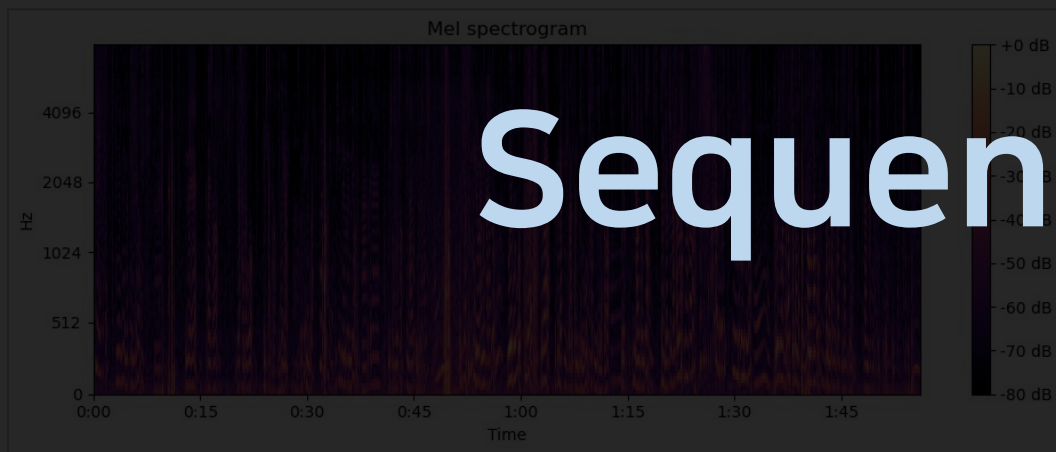
텍스트(Text)

생일날이면은 미역국도 같이 해주시곤 했지.

6. 모델 구성

A. 베이스라인 모델 선정

음성(Audio)



텍스트(Text)

싱글날이건은 바역국도 같이 해주시곤 했지.

Sequential Data

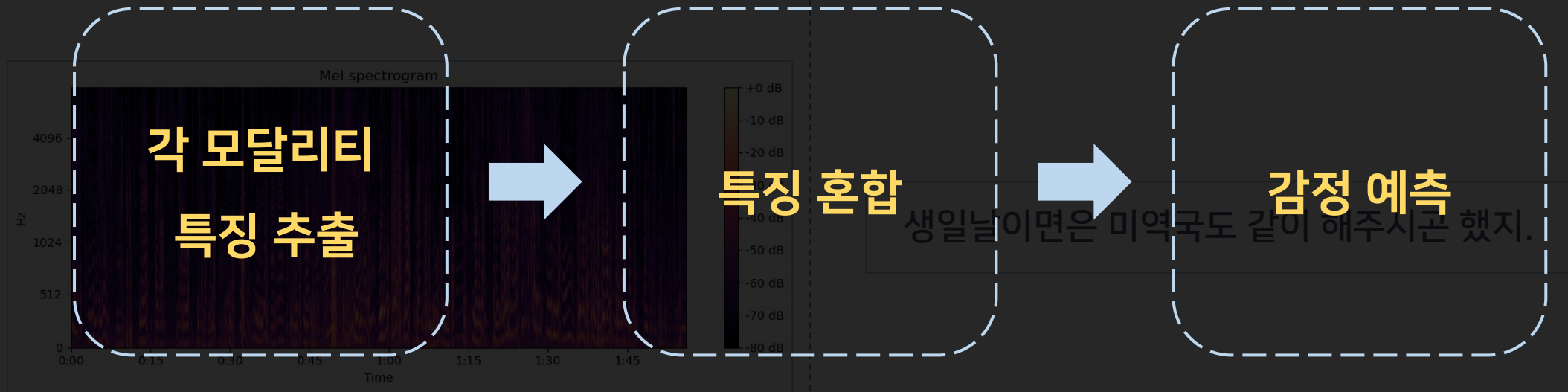
6. 모델 구성

A. 베이스라인 모델 선정

Sequential Data

음성(Audio)

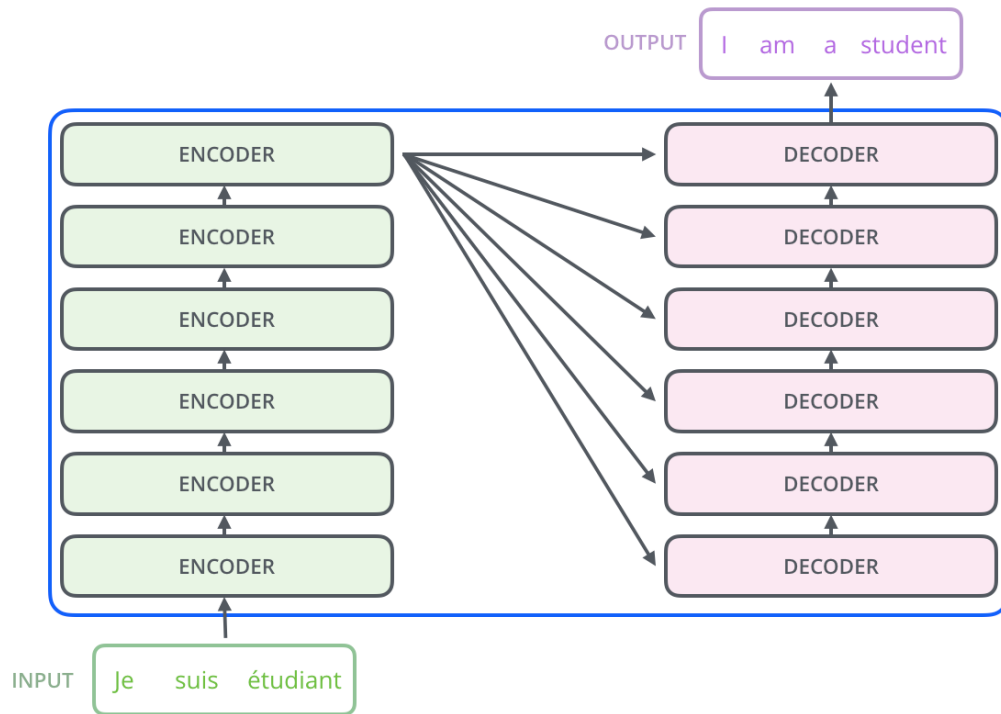
텍스트(Text)



순서 정보 보존

6. 모델 구성

B. Transformer (Vaswani et al., 2017)



Transformer의 특징

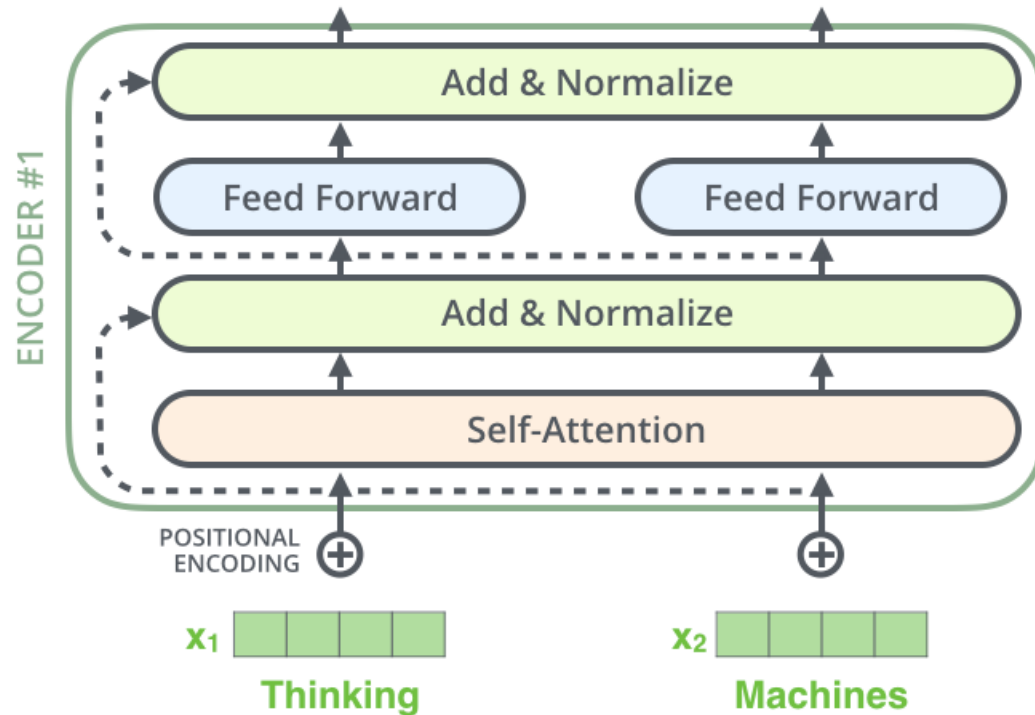
- ① Seq2Seq 구조의 데이터 압축 과정 배제
- ② Self Attention: 입력 데이터 간 관계 학습
- ③ 자연어처리(NLP)에서 우수한 성능

Encoder의 특징

- ① 입력 데이터의 Size 및 순서 정보 보존
- ② Residual Learning: 반복적인 Self Attention 학습

6. 모델 구성

B. Transformer (Vaswani et al., 2017)



Transformer의 특징

- ① Seq2Seq 구조의 데이터 압축 과정 배제
- ② Self Attention: 입력 데이터 간 관계 학습
- ③ 자연어처리(NLP)에서 우수한 성능

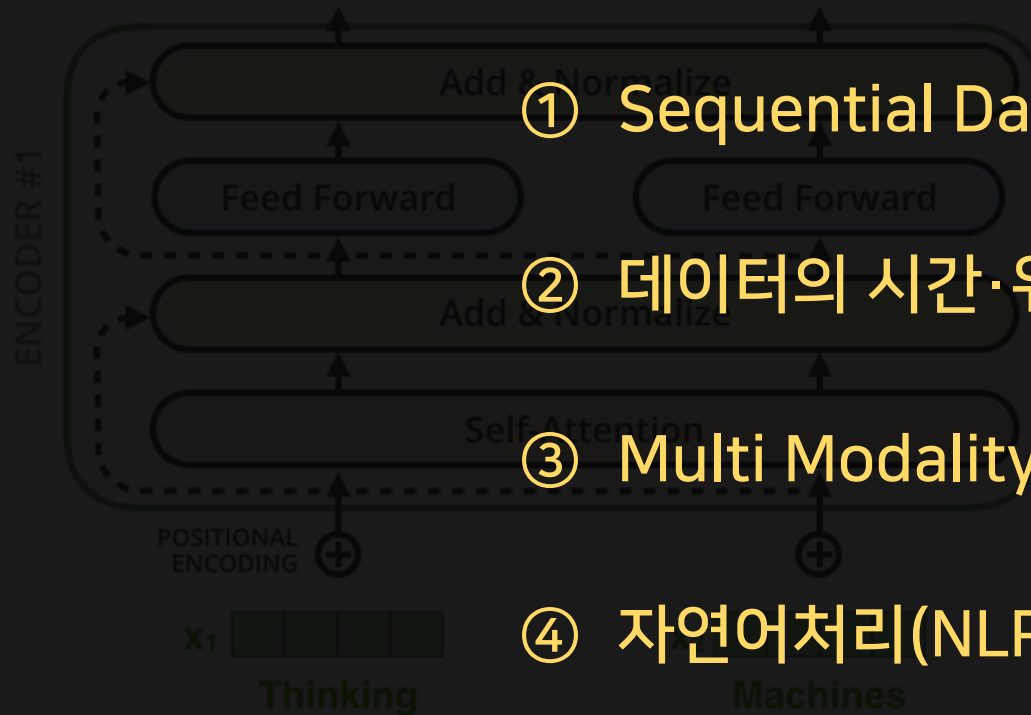
Encoder의 특징

- ① 입력 데이터의 Size 및 순서 정보 보존
- ② Residual Learning: 반복적인 Self Attention 학습

6. 모델 구성

B. Transformer (Vaswani et al., 2017)

최종 선정 배경



① Sequential Data의 특징 추출 용이성

② 데이터의 시간·위치 정보 보존 가능

③ Multi Modality의 합성 용이성

④ 자연어처리(NLP)에서 뛰어난 효과

Transformer의 특징

① Self-Attention: 입력 데이터 압축 과정 배제

② Self-Attention: 입력 데이터 간 관계 학습

③ Self-Attention: Transformer에서 우수한 성능

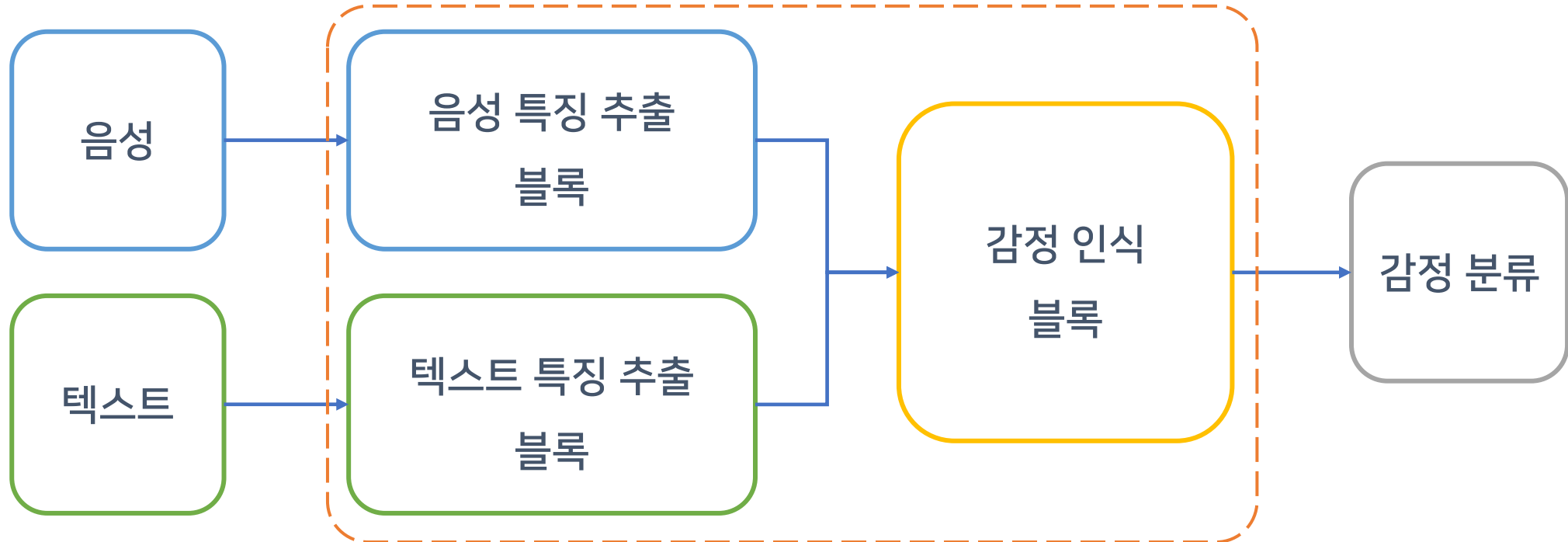
① 입력 데이터의 Size 및 순서 정보 보존

② Residual Learning: 반복적인 Self Attention 학습

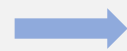
1) https://nlpinkorean.github.io/images/transformer/transformer_residual_layer_norm.png

6. 모델 구성

B. Transformer (Vaswani et al., 2017)



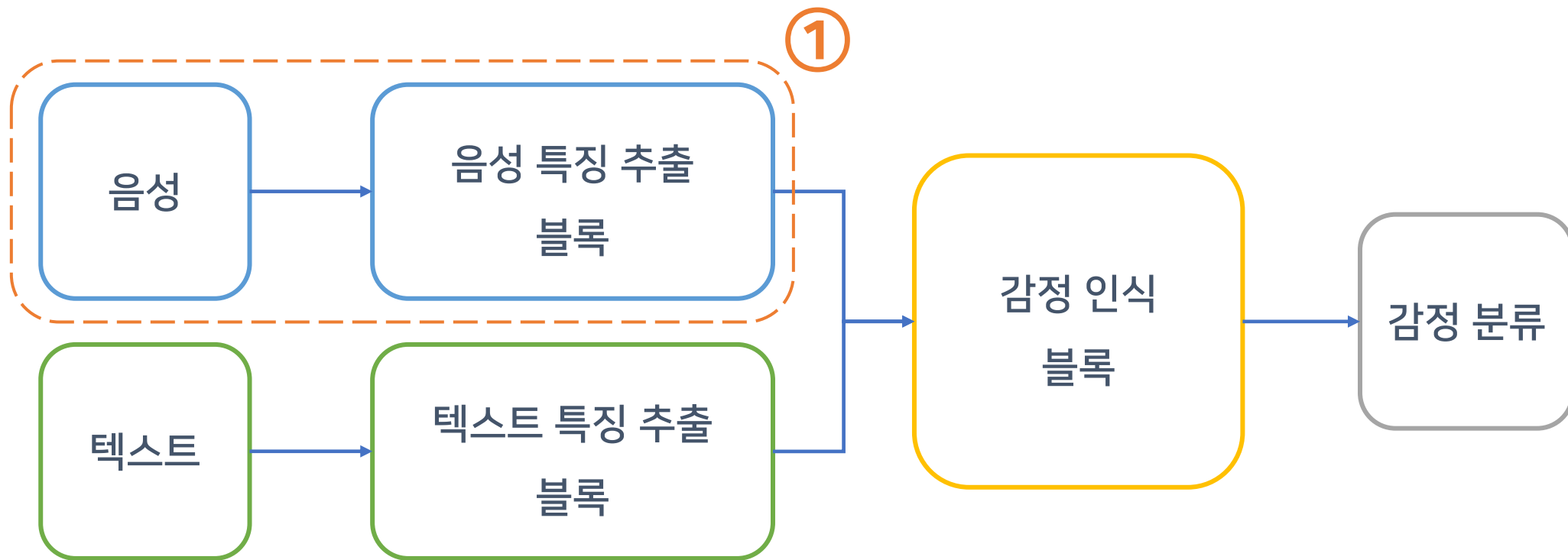
Multihead Attention 기반 Encoder 구조 응용



데이터 구조 보존 + 예측 성능 ↑

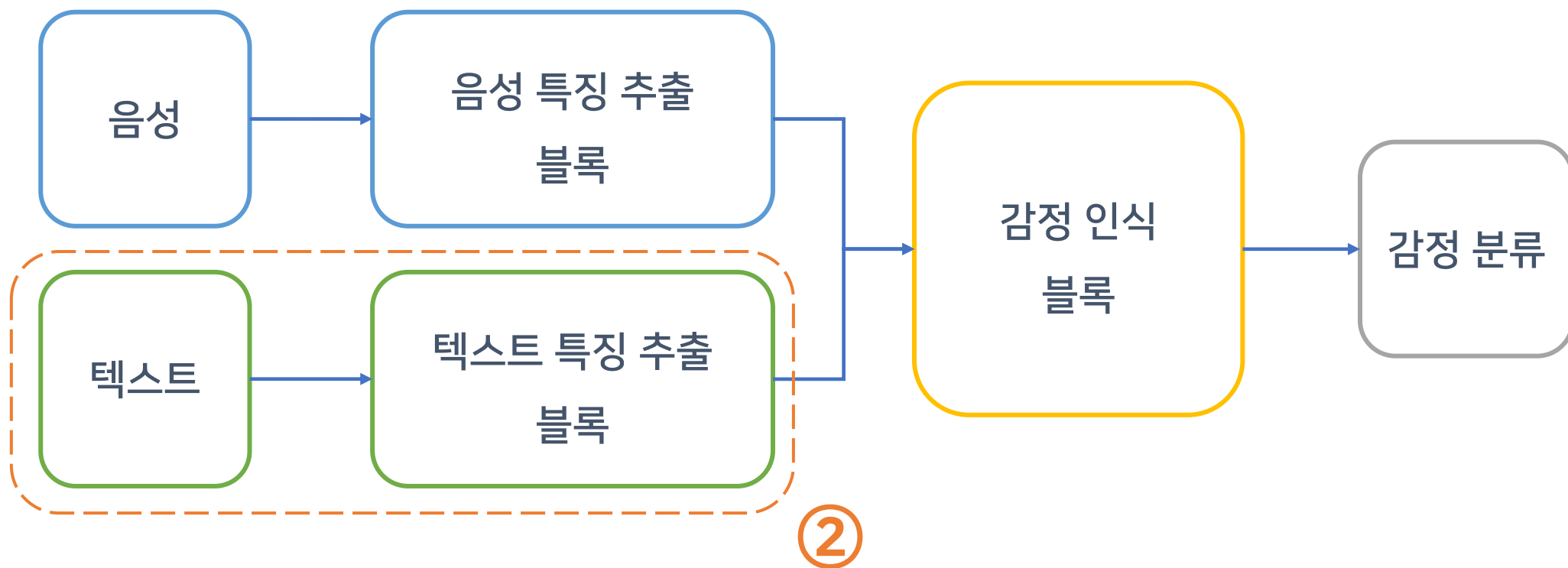
6. 모델 구성

B. Transformer (Vaswani et al., 2017)



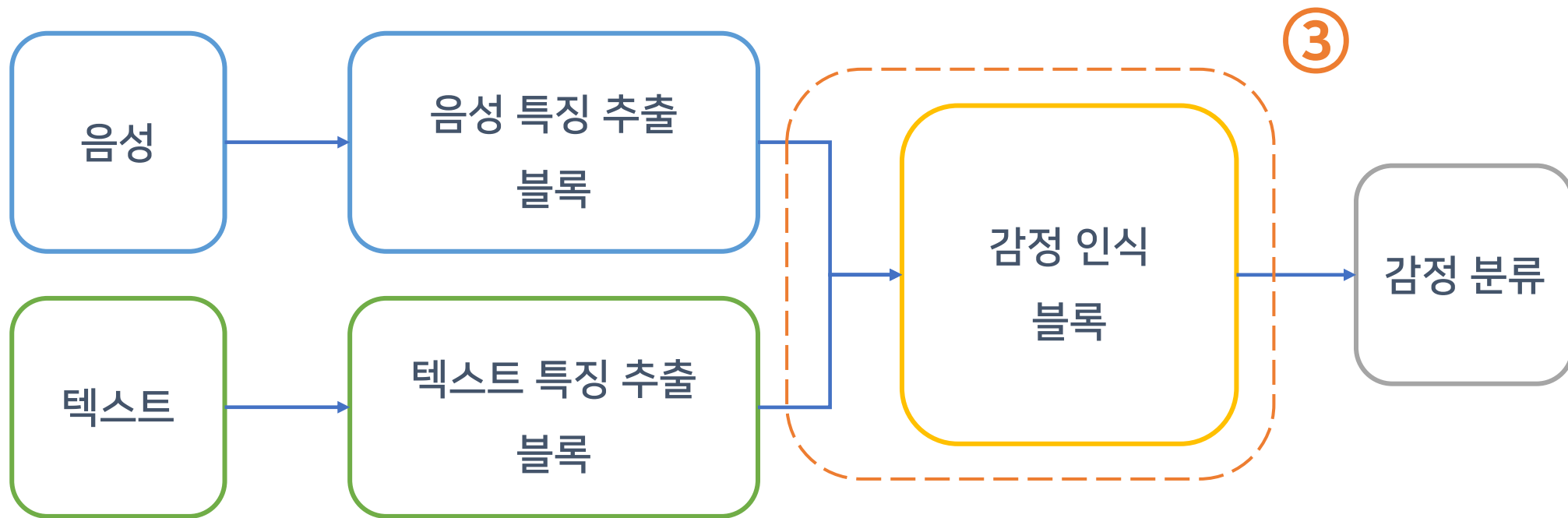
6. 모델 구성

B. Transformer (Vaswani et al., 2017)



6. 모델 구성

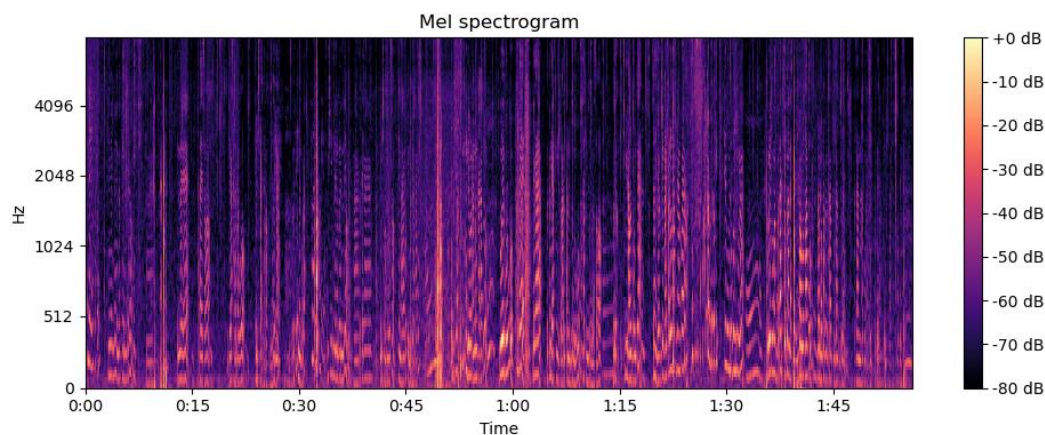
B. Transformer (Vaswani et al., 2017)



6. 모델 구성

각성도와 공부정도

음성(Audio)와 텍스트(Text)



생일날이면은 미역국도 같이 해주시곤 했지.

각성도(Arousal)와 공부정도(Valence)

3.5

2.7

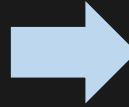
4.2

비정형 데이터를 사용하는 딥러닝 모델에 하나의 **실수값**을 어떻게 사용할 수 있을까?

6. 모델 구성

각성도와 공부정도
음성

음성(Audio)과 텍스트(Text)



각성도
예측

각성도(Arousal)와 공부정도(Valence)

3.5

공부정도
예측

2.7

4.2

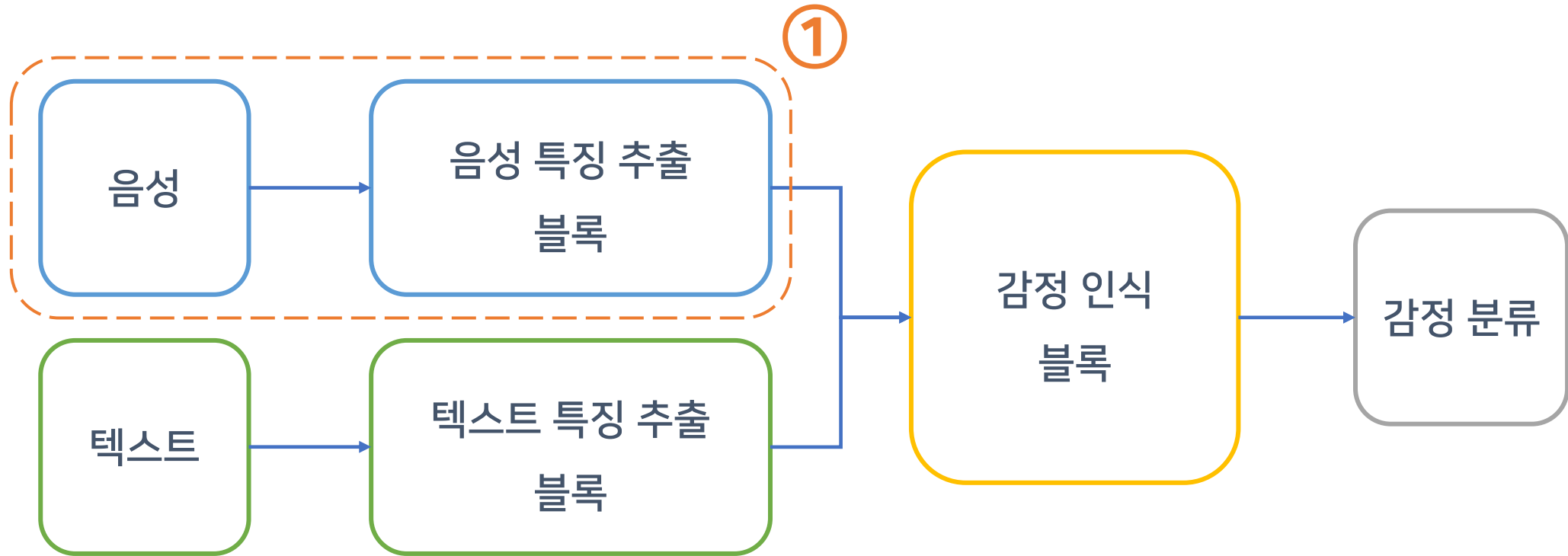
싱글모달 예측 모델
사전학습



특징 추출 블록에
파라미터 사용!

6. 모델 구성

C. 음성 - 각성도 예측 모델

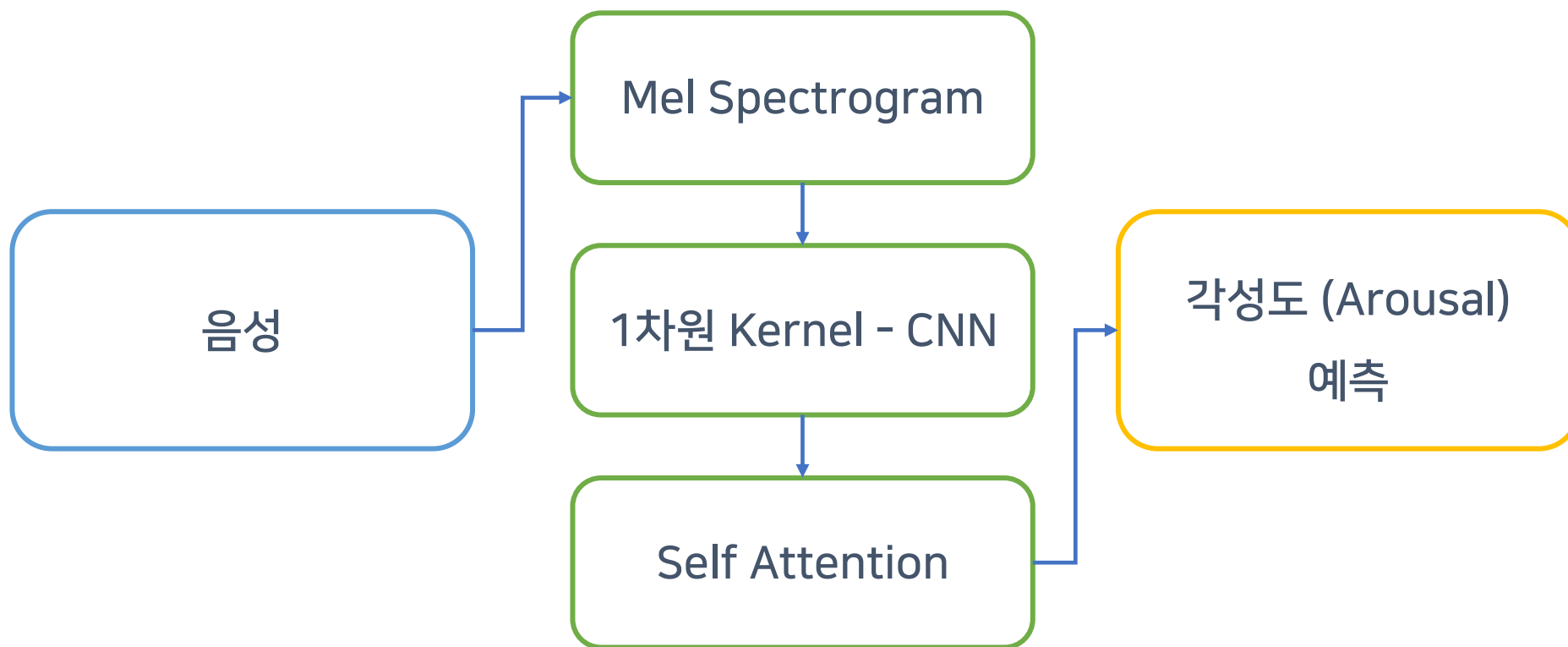


음성 - 각성도 예측 모델의 파라미터를 음성 특징 추출 블록에 차용 ➡ 발화자의 각성도 반영

6. 모델 구성

C. 음성 - 각성도 예측 모델

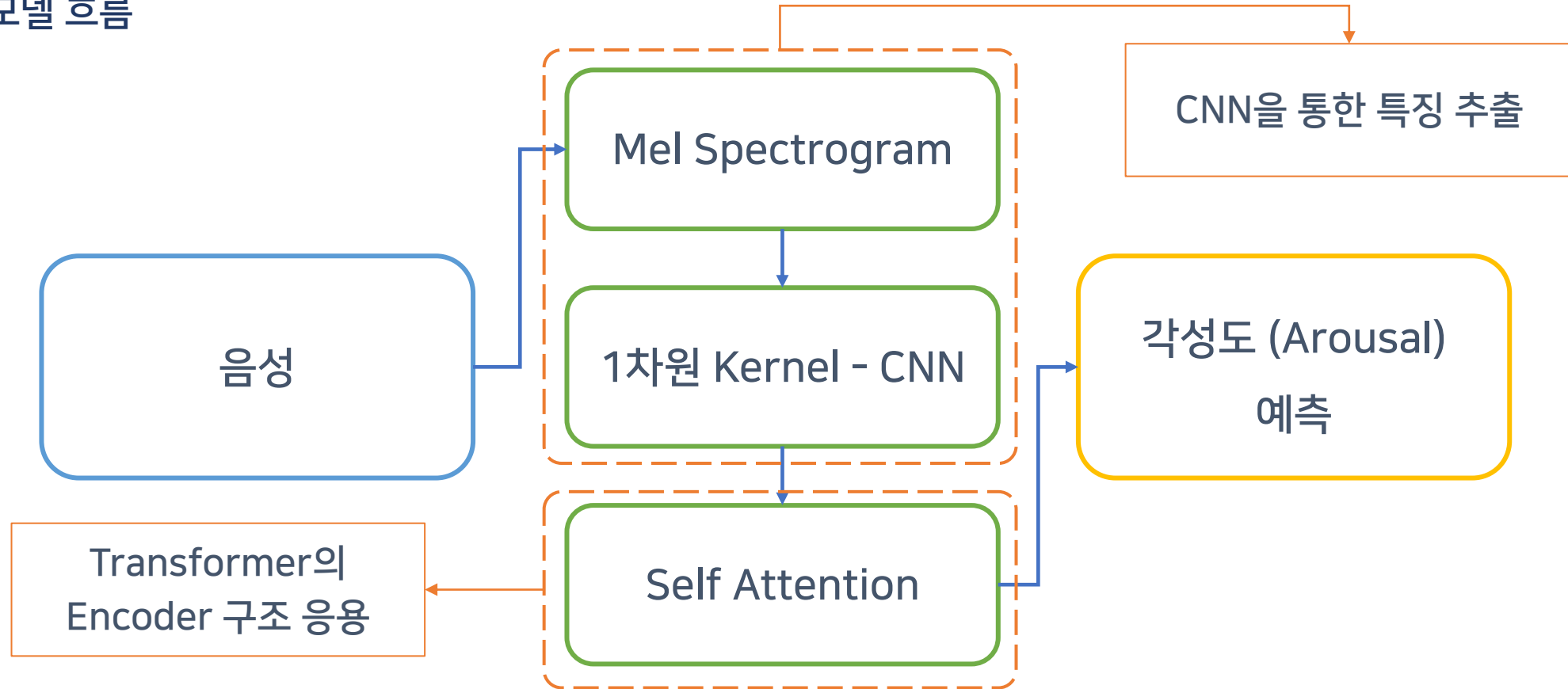
모델 흐름



6. 모델 구성

C. 음성 - 각성도 예측 모델

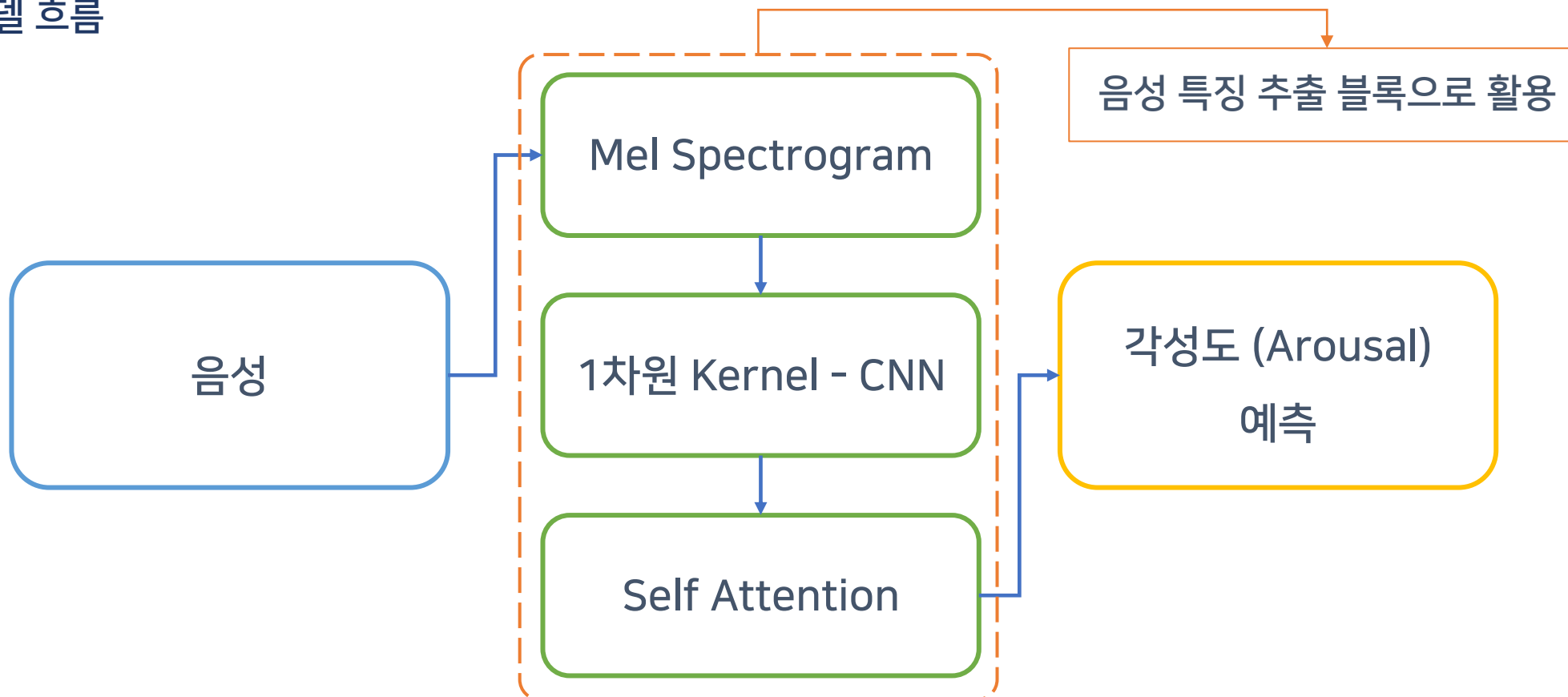
모델 흐름



6. 모델 구성

C. 음성 - 각성도 예측 모델

모델 흐름

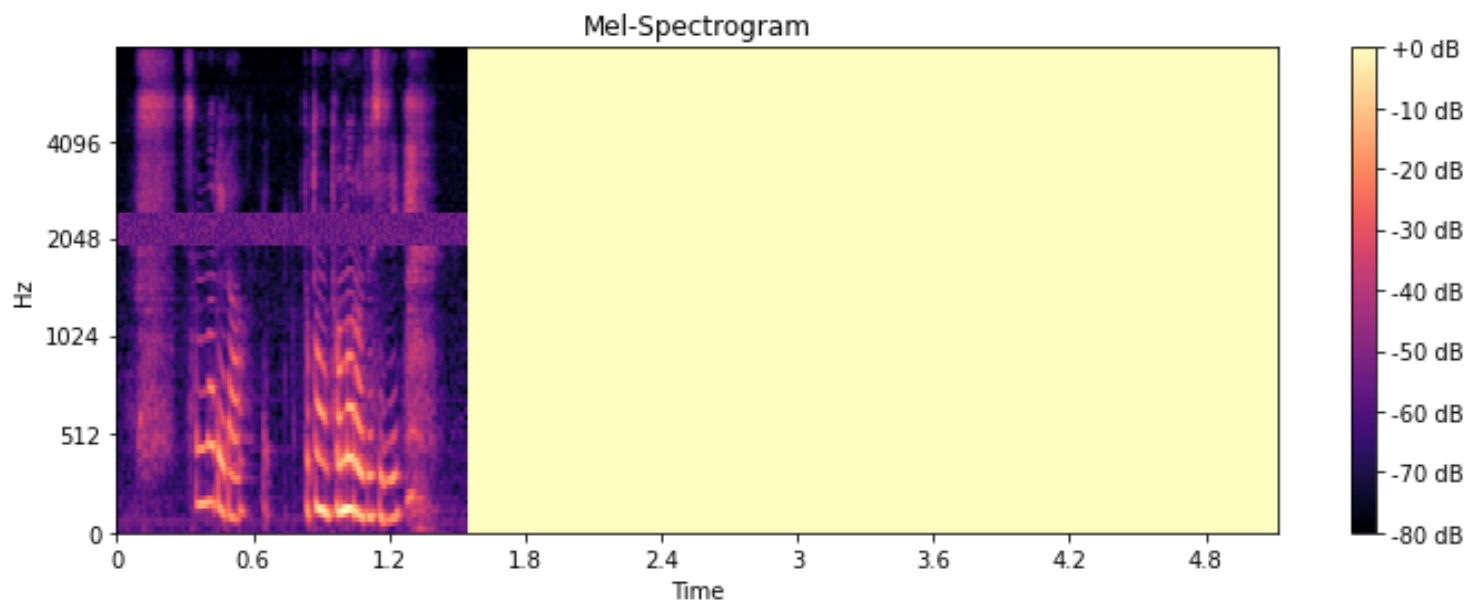


6. 모델 구성

C. 음성 - 각성도 예측 모델

(1) Mel-Spectrogram

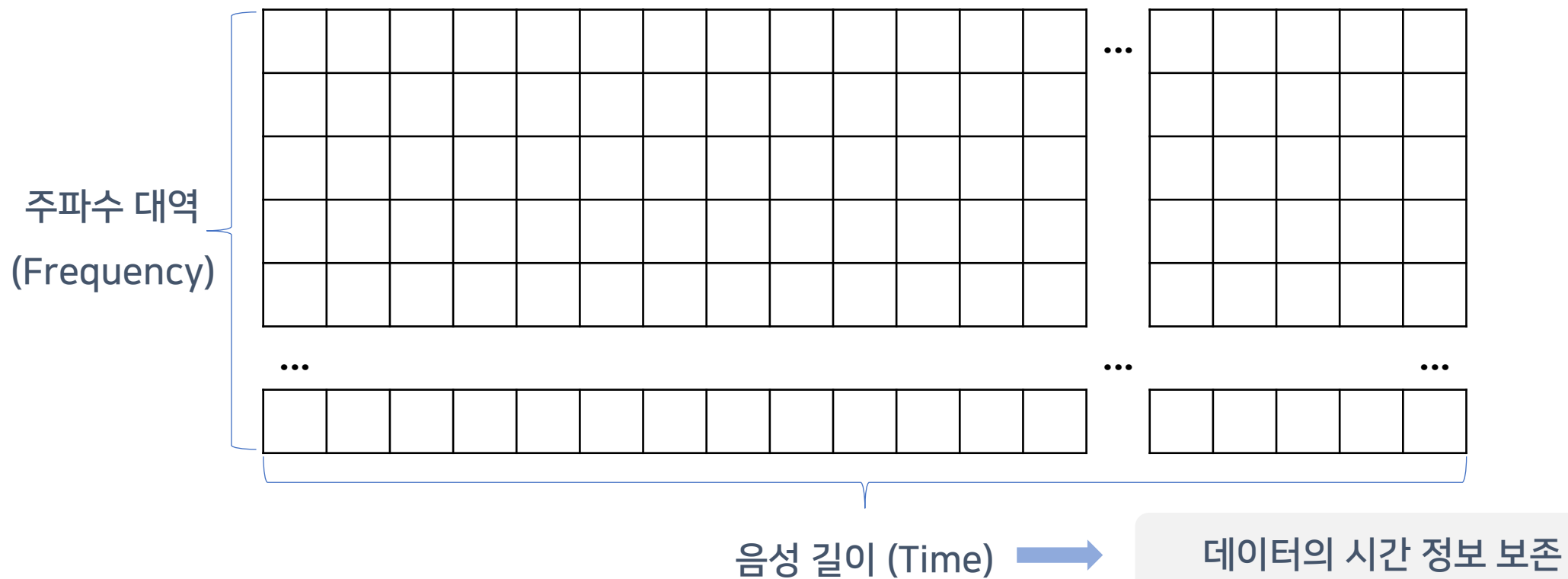
- ① $n_mels=128$, 최소 주파수 = 100, 최대 주파수 = 8000
- ② Zero-Padding & Random Frequency Masking 적용



6. 모델 구성

C. 음성 - 각성도 예측 모델

(2) 1차원 Kernel - CNN



6. 모델 구성

C. 음성 - 각성도 예측 모델

(2) 1차원 Kernel - CNN



6. 모델 구성

C. 음성 - 각성도 예측 모델

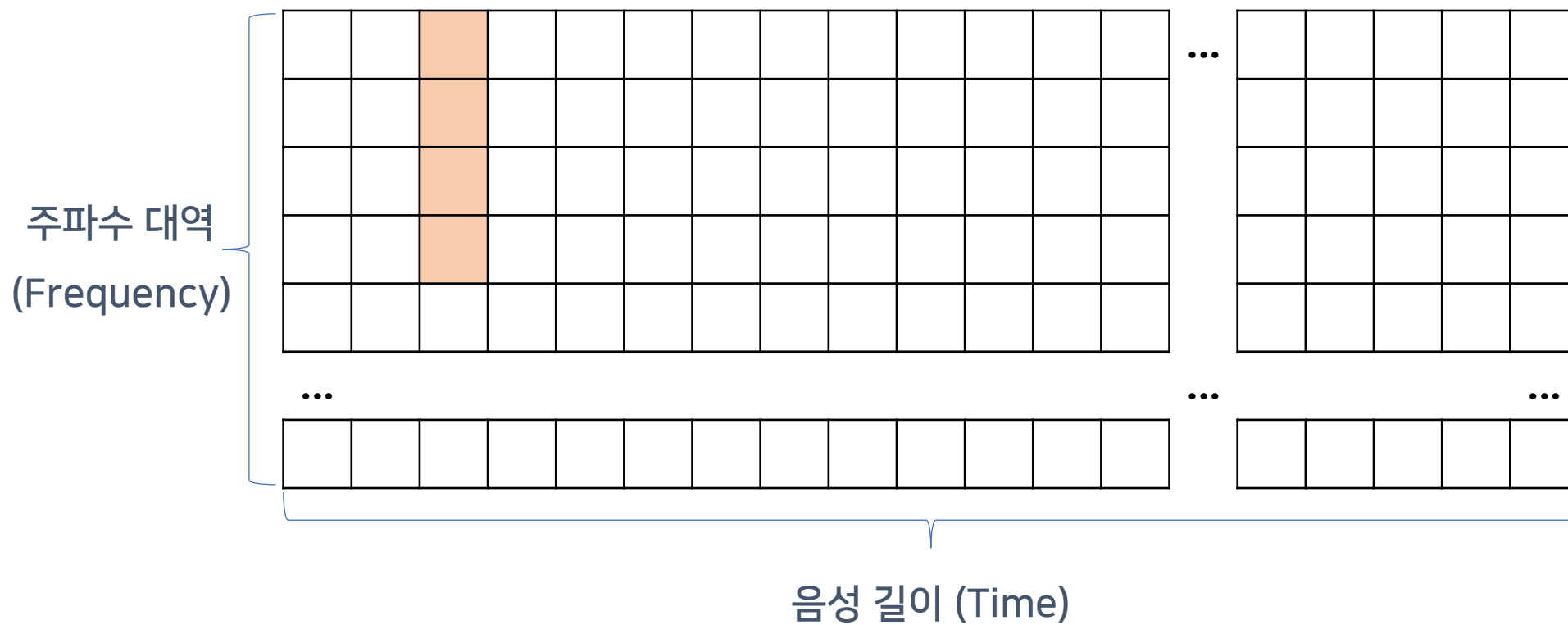
(2) 1차원 Kernel - CNN



6. 모델 구성

C. 음성 - 각성도 예측 모델

(2) 1차원 Kernel - CNN



6. 모델 구성

C. 음성 - 각성도 예측 모델

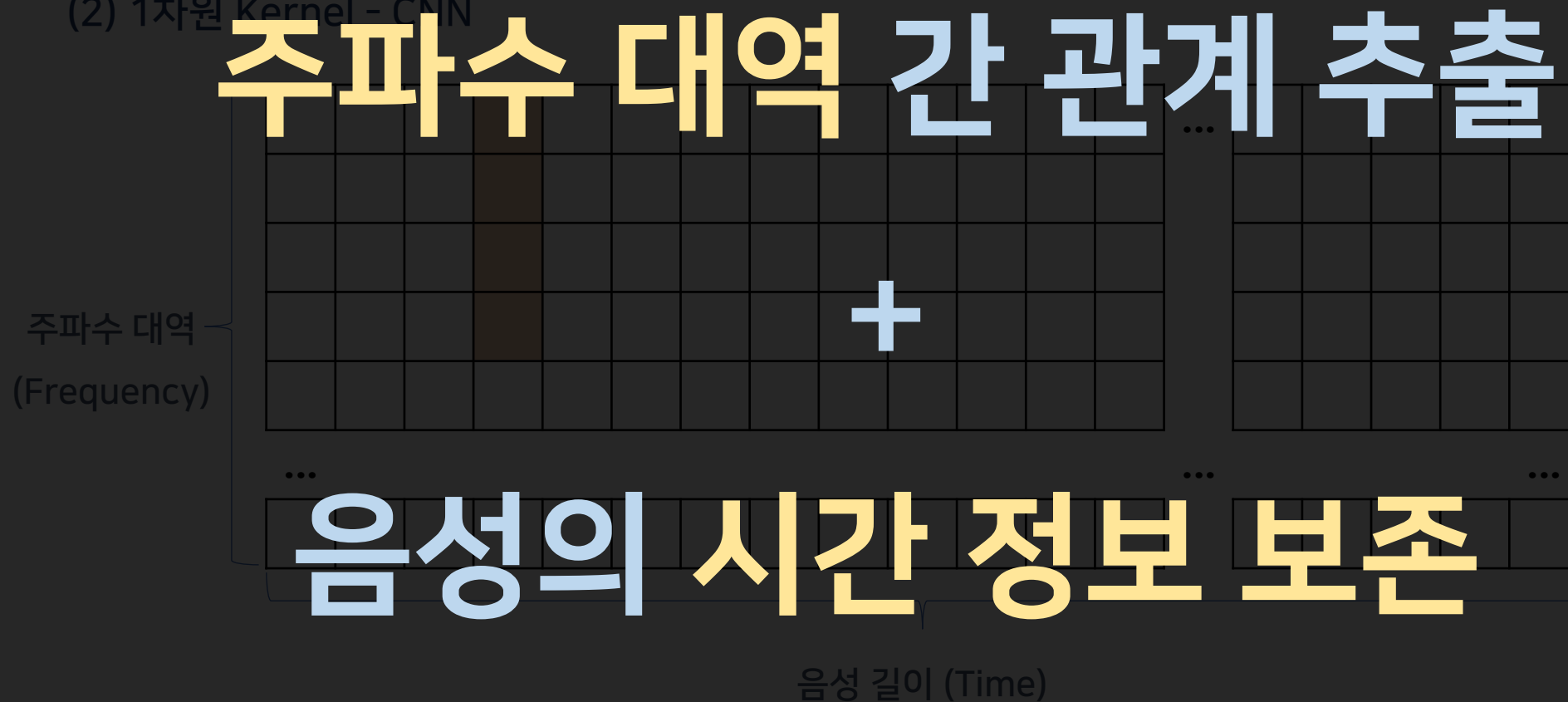
(2) 1차원 Kernel - CNN



6. 모델 구성

C. 음성 - 각성도 예측 모델

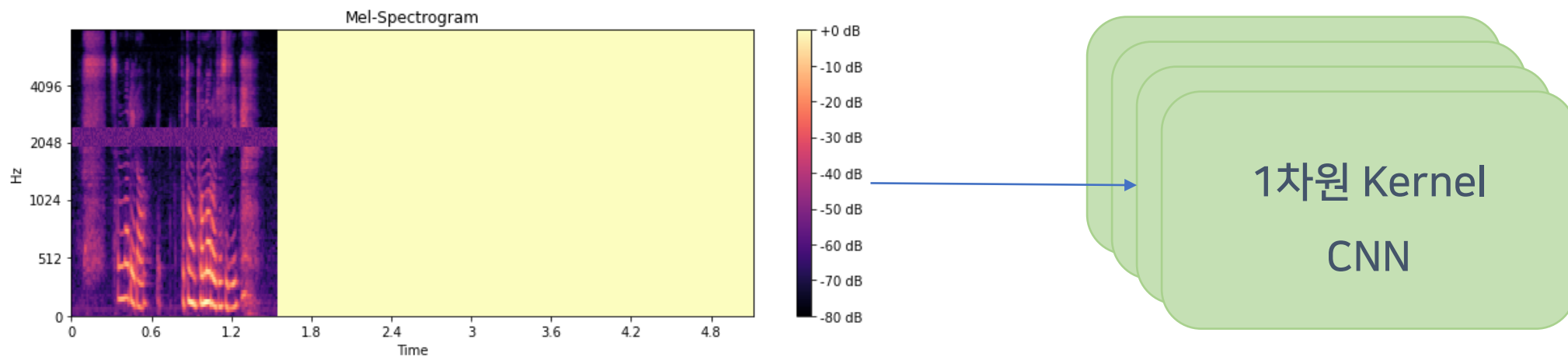
(2) 1차원 Kernel - CNN



6. 모델 구성

C. 음성 - 각성도 예측 모델

(2) 1차원 Kernel - CNN

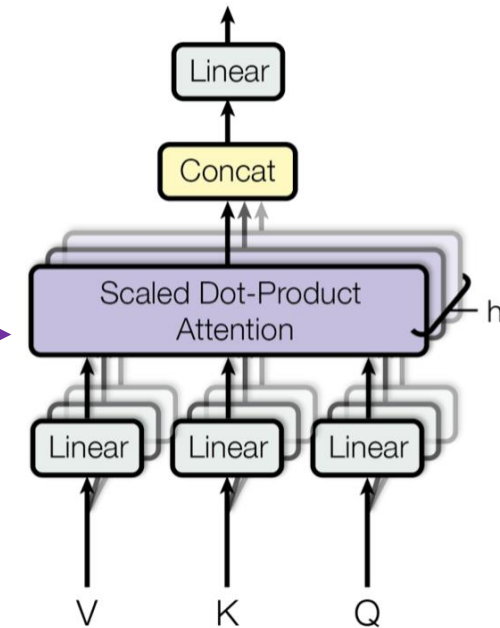
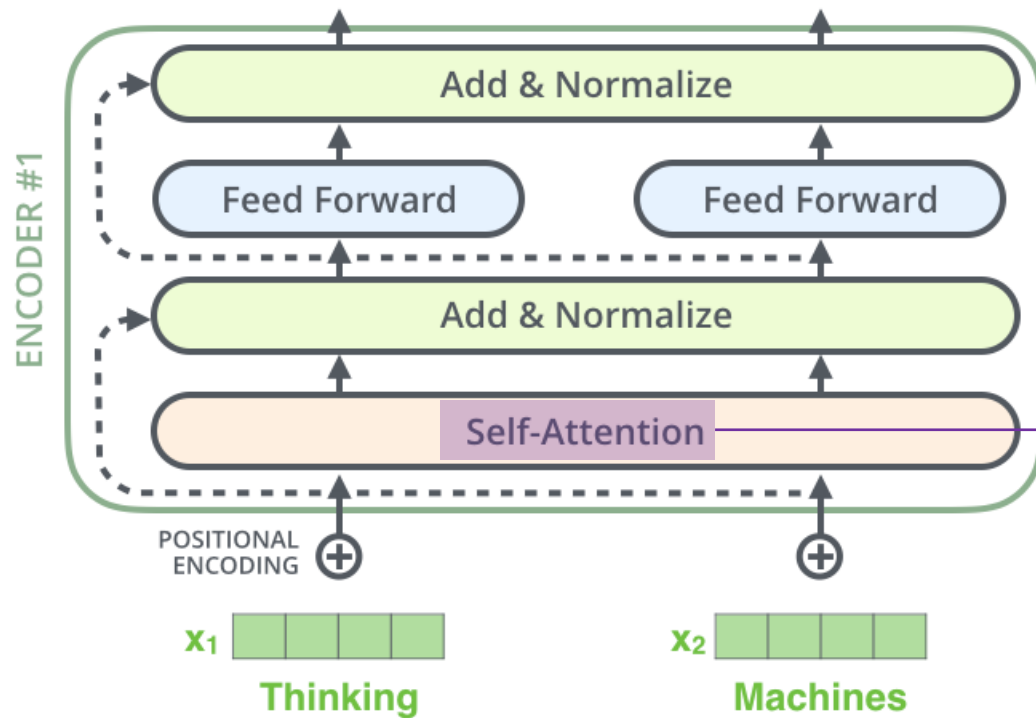


1차원 Kernel CNN Layer → 4회 반복

6. 모델 구성

C. 음성 - 각성도 예측 모델

(3) Self Attention



Multihead Attention

6. 모델 구성

C. 음성 - 각성도 예측 모델

(4) 모델 학습



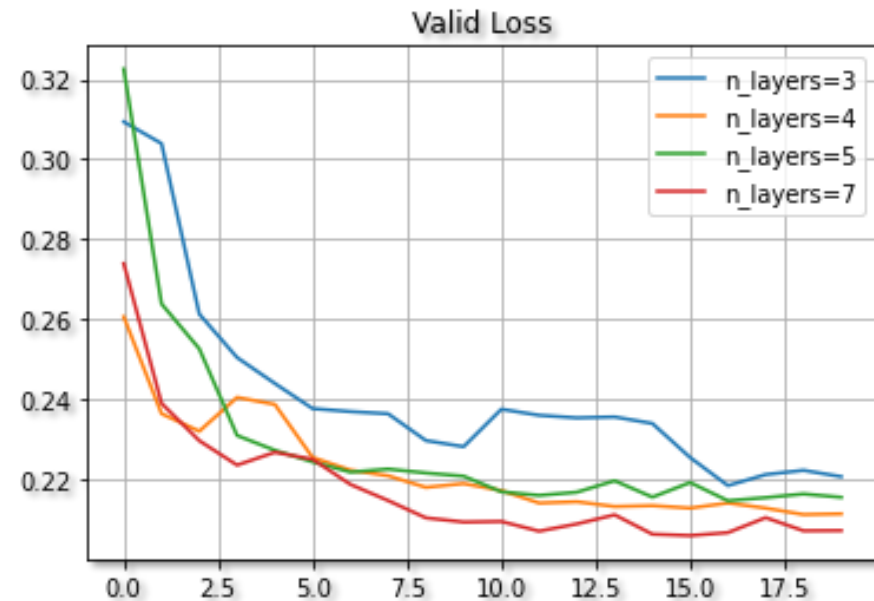
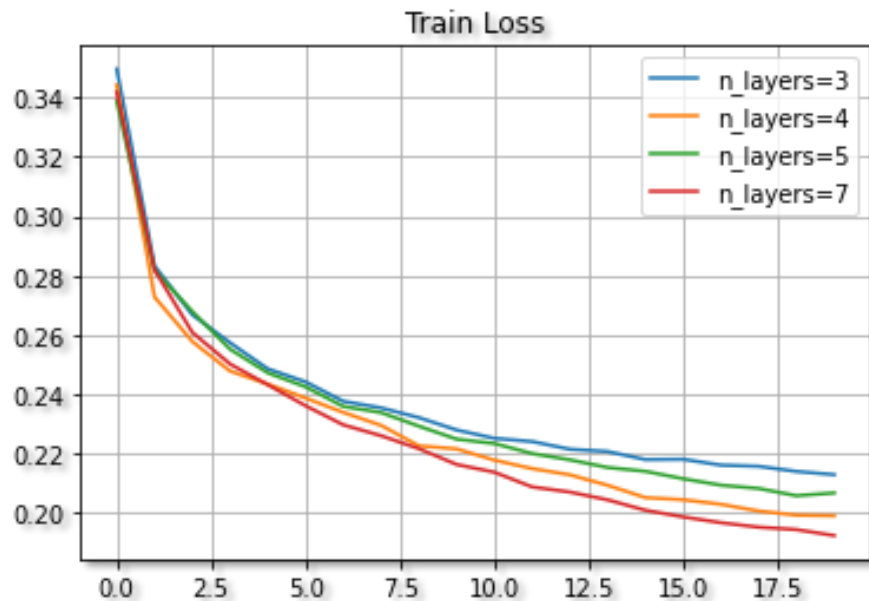
Self Attention 학습 횟수에 따른 예측 성능(RMSE) 비교

6. 모델 구성

C. 음성 - 각성도 예측 모델

(4) 모델 학습

n_layers = Self Attention 반복 횟수



Self Attention 반복 횟수에 따른 RMSE



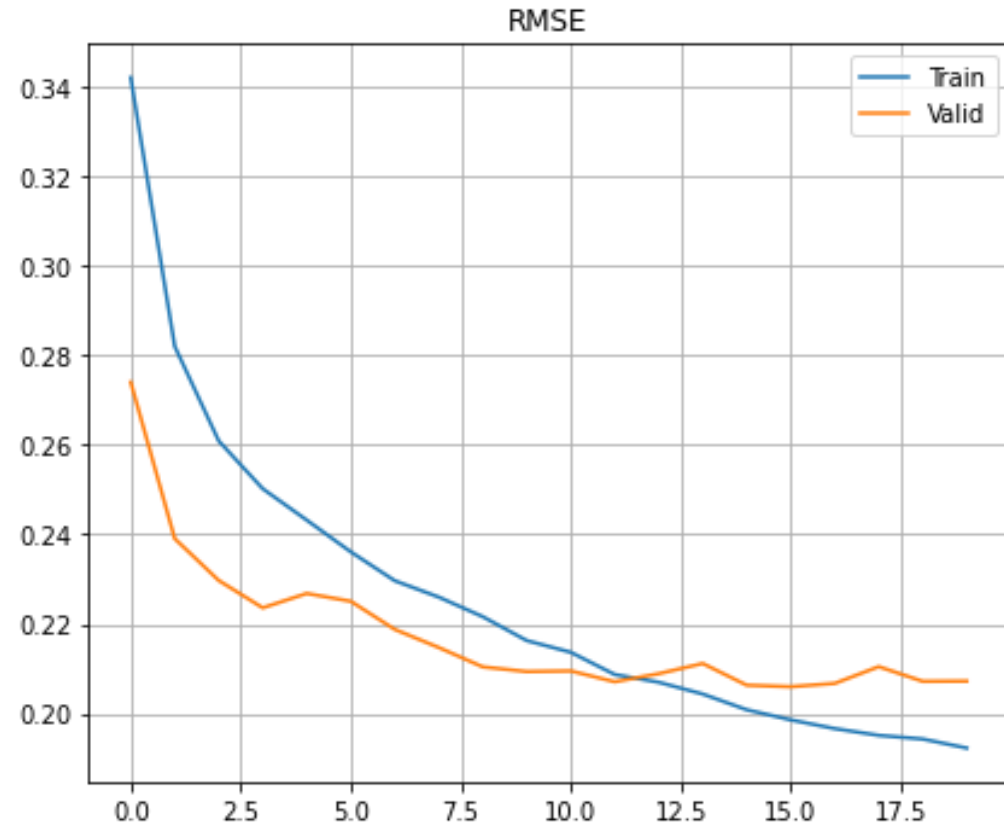
7회 반복시 최저 RMSE

6. 모델 구성

C. 음성 - 각성도 예측 모델

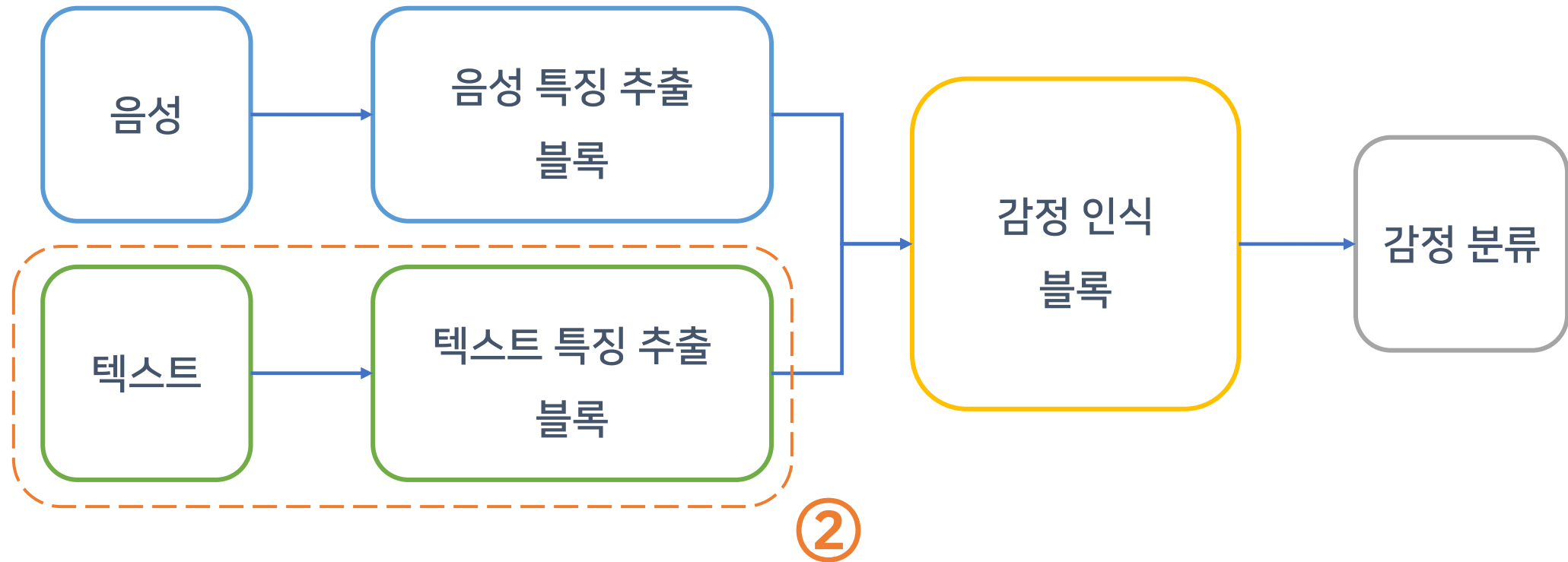
(5) 모델 학습 결과 (n_layers=7)

Train	Valid
0.341884649	0.273954606
0.281943787	0.239045343
0.260837908	0.229711024
0.250174330	0.223605140
.	.
.	.
.	.
0.196662220	0.206742686
0.195157791	0.210485889
0.194336303	0.207206603
0.192330844	0.206007988



6. 모델 구성

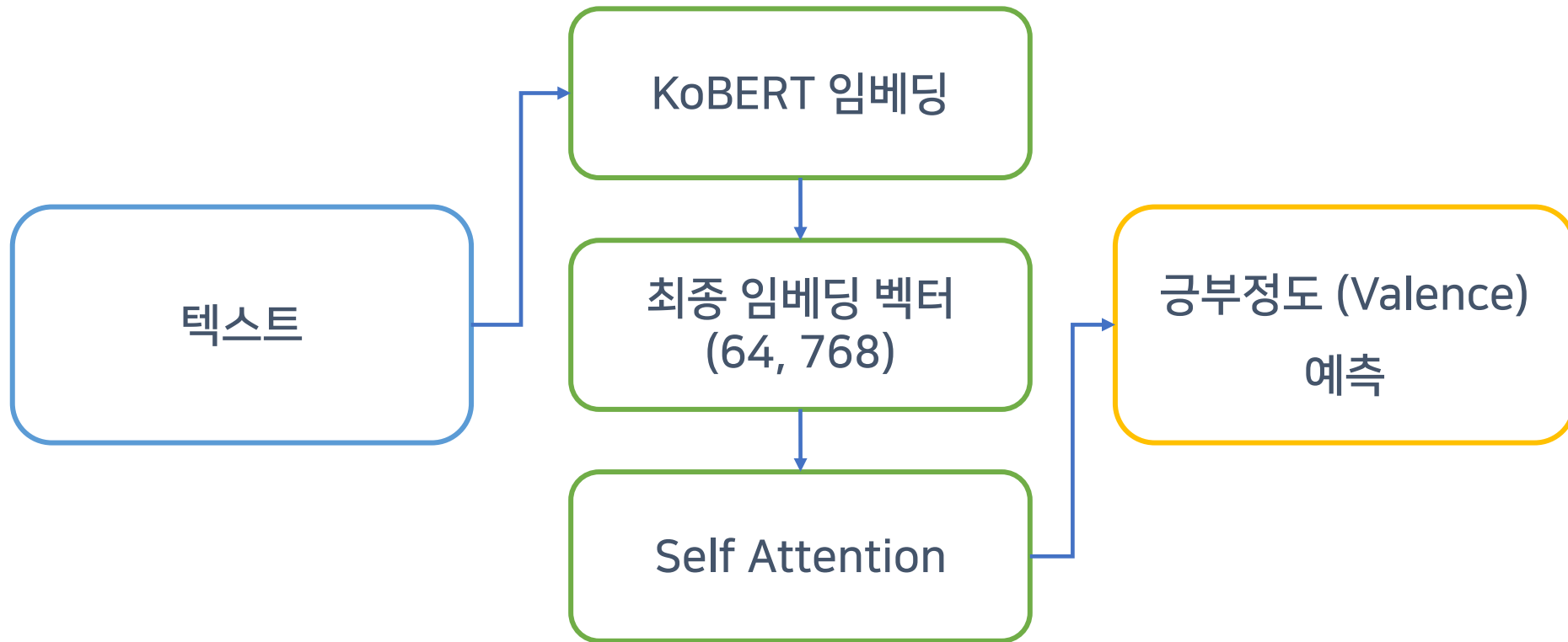
D. 텍스트 - 긍/부정도 예측 모델



6. 모델 구성

D. 텍스트 - 긍/부정도 예측 모델

모델 흐름

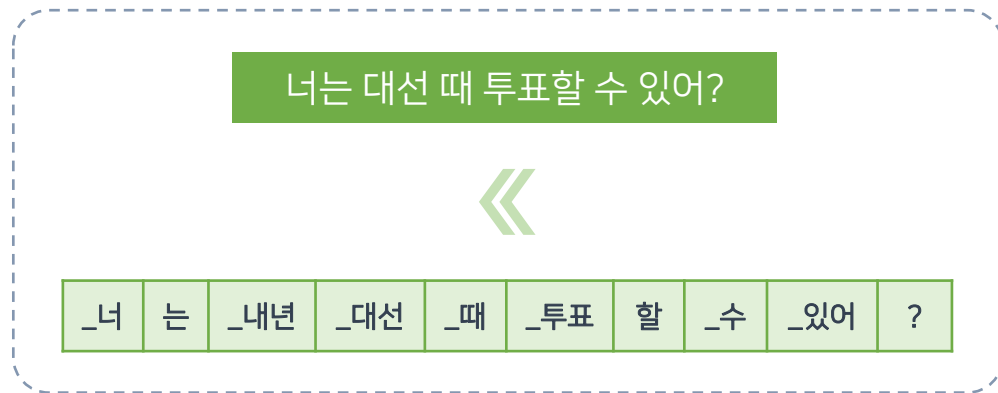


6. 모델 구성

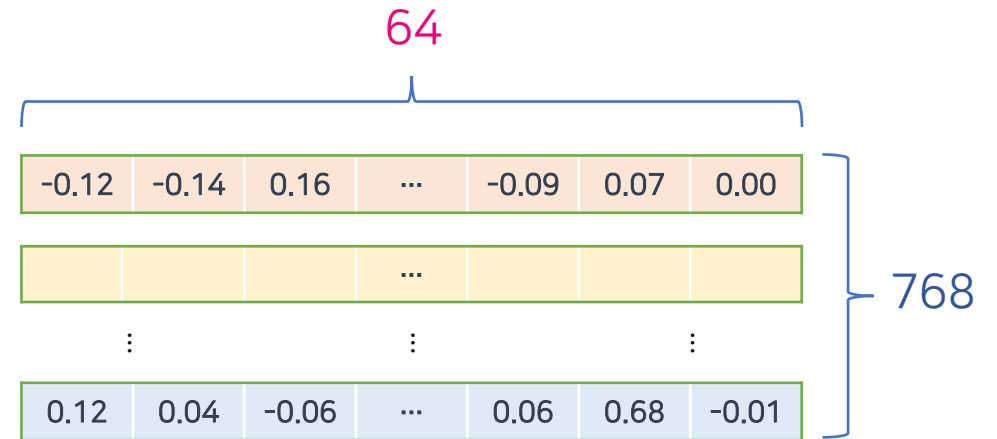
D. 텍스트 - 긍/부정도 예측 모델

(1) KoBERT 임베딩

- 문맥 고려측면에서 KoBERT 임베딩 선택



토큰화된 문장

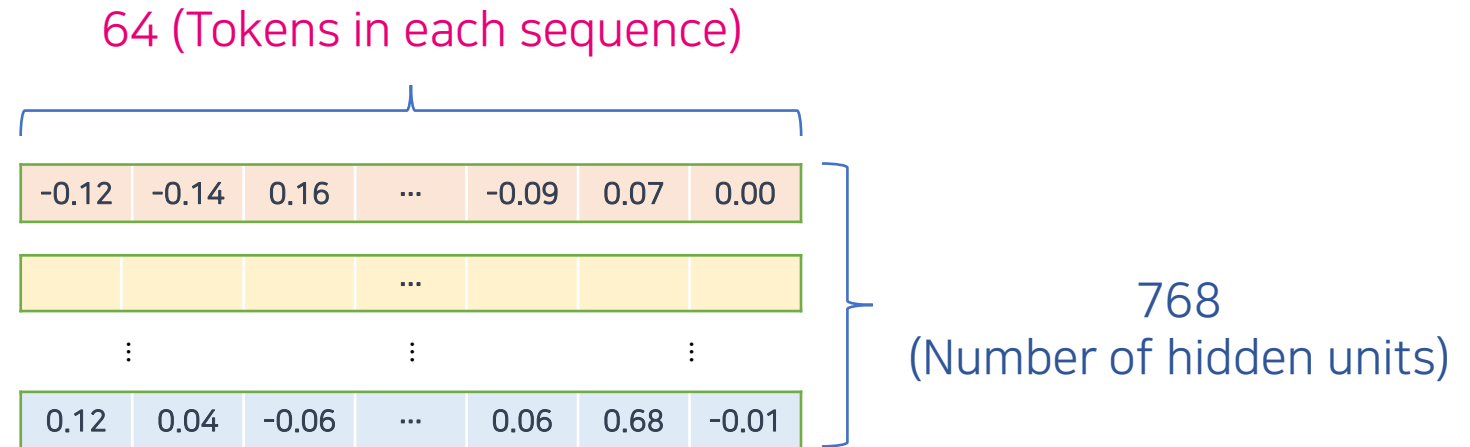


[64, 768]

6. 모델 구성

D. 텍스트 - 긍/부정도 예측 모델

(1) KoBERT 임베딩



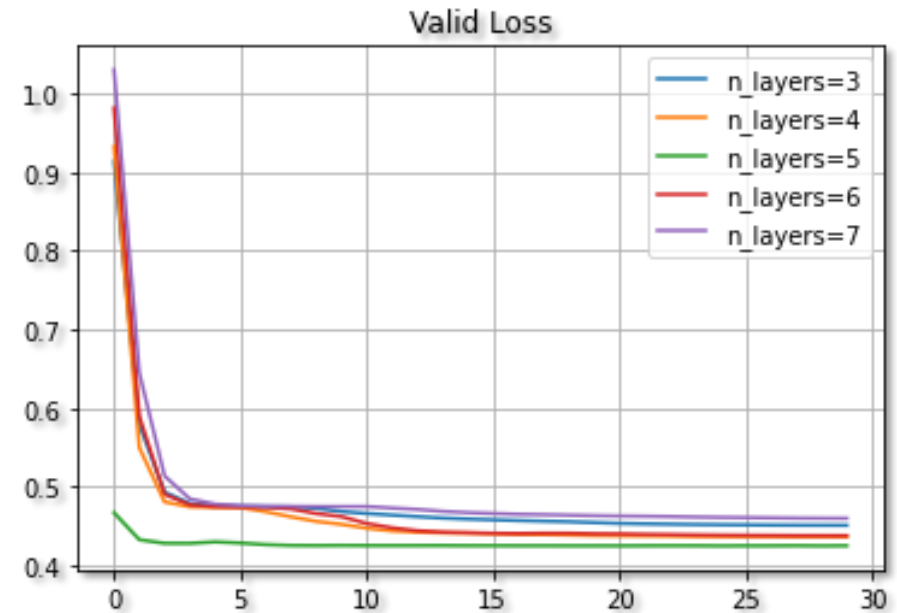
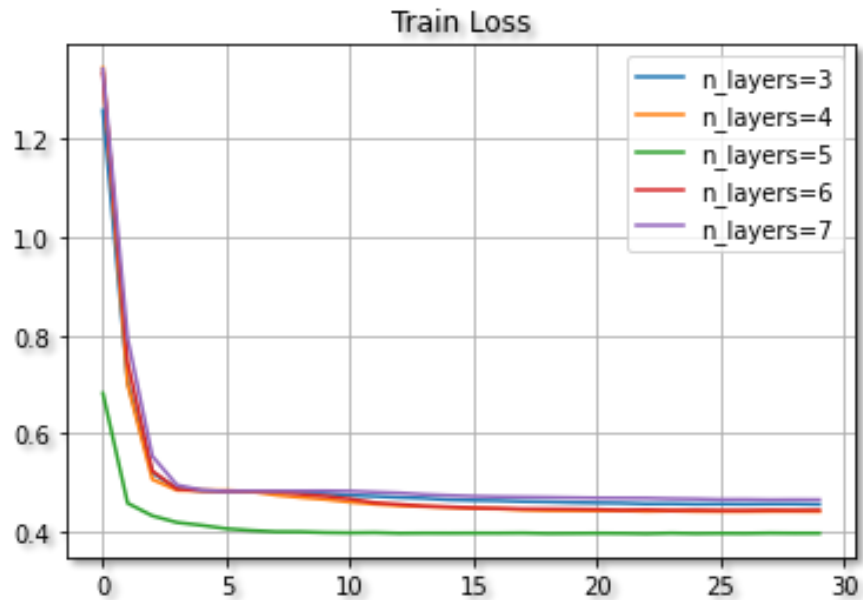
GRU 모델로 예측

6. 모델 구성

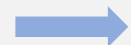
C. 텍스트 - 긍/부정도 예측 모델

(2) 모델 학습

n_layers = Self Attention 반복 횟수



Self Attention 반복 횟수에 따른 RMSE



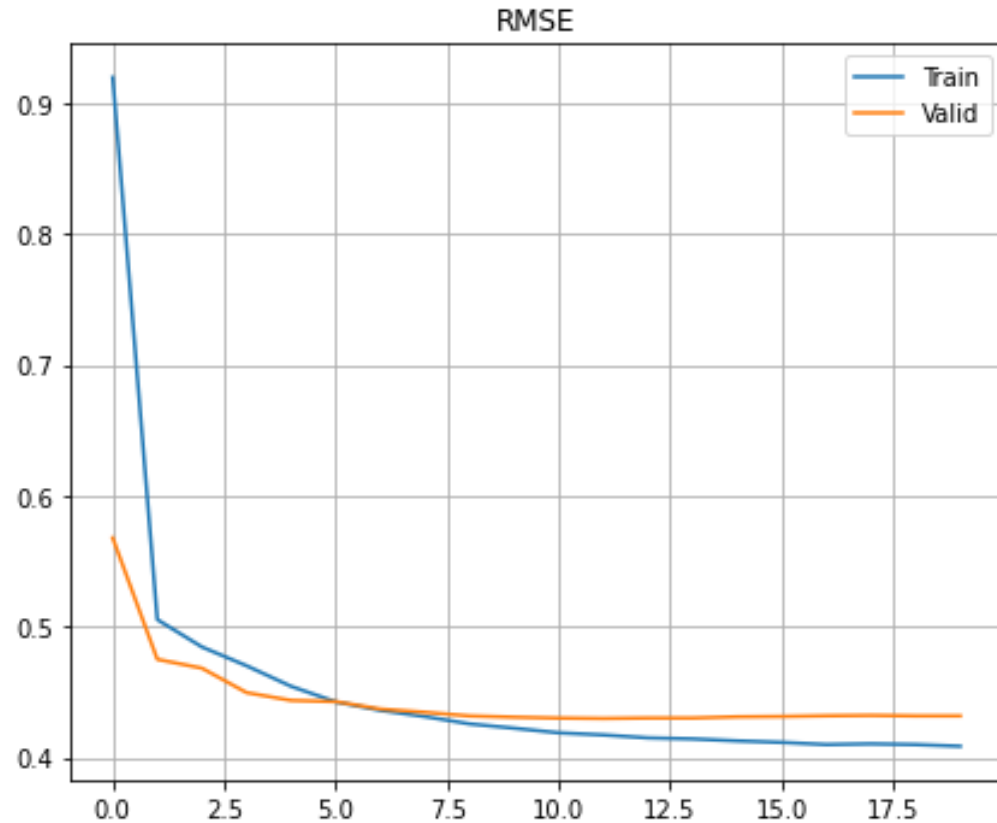
5회 반복시 최저 RMSE

6. 모델 구성

C. 텍스트 - 긍/부정도 예측 모델

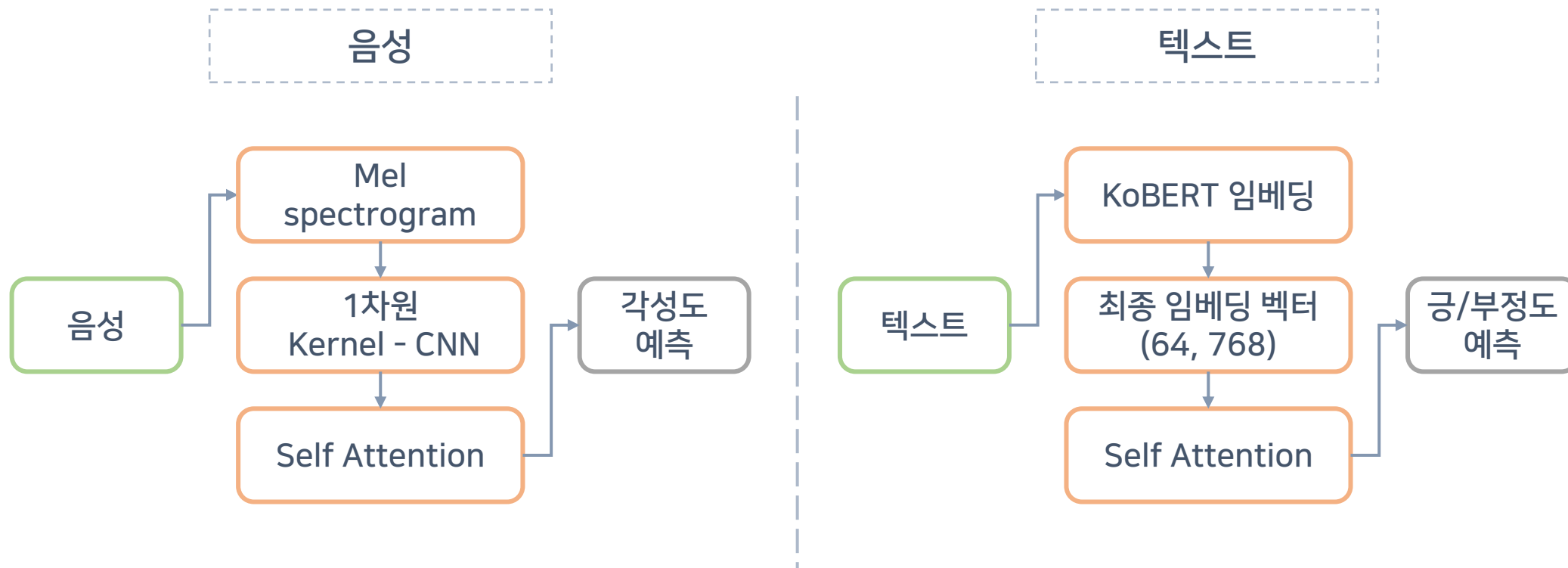
(3) 모델 학습 결과 (n_layers=5)

Train	Valid
0.91981	0.475089
0.505456	0.468349
0.484568	0.449671
0.470203	0.443702
.	.
.	.
.	.
0.409979	0.431955
0.410524	0.432256
0.409952	0.431876
0.408682	0.431883



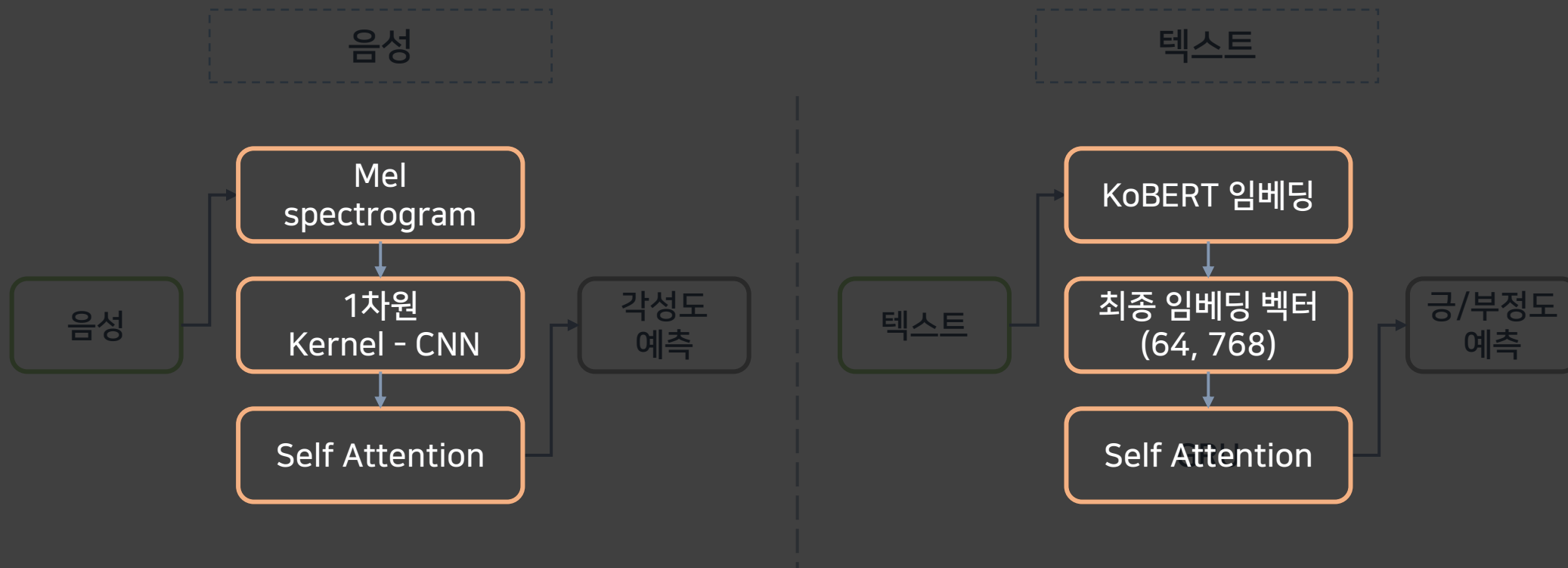
6. 모델 구성

E. 음성/텍스트 특징 추출 블록



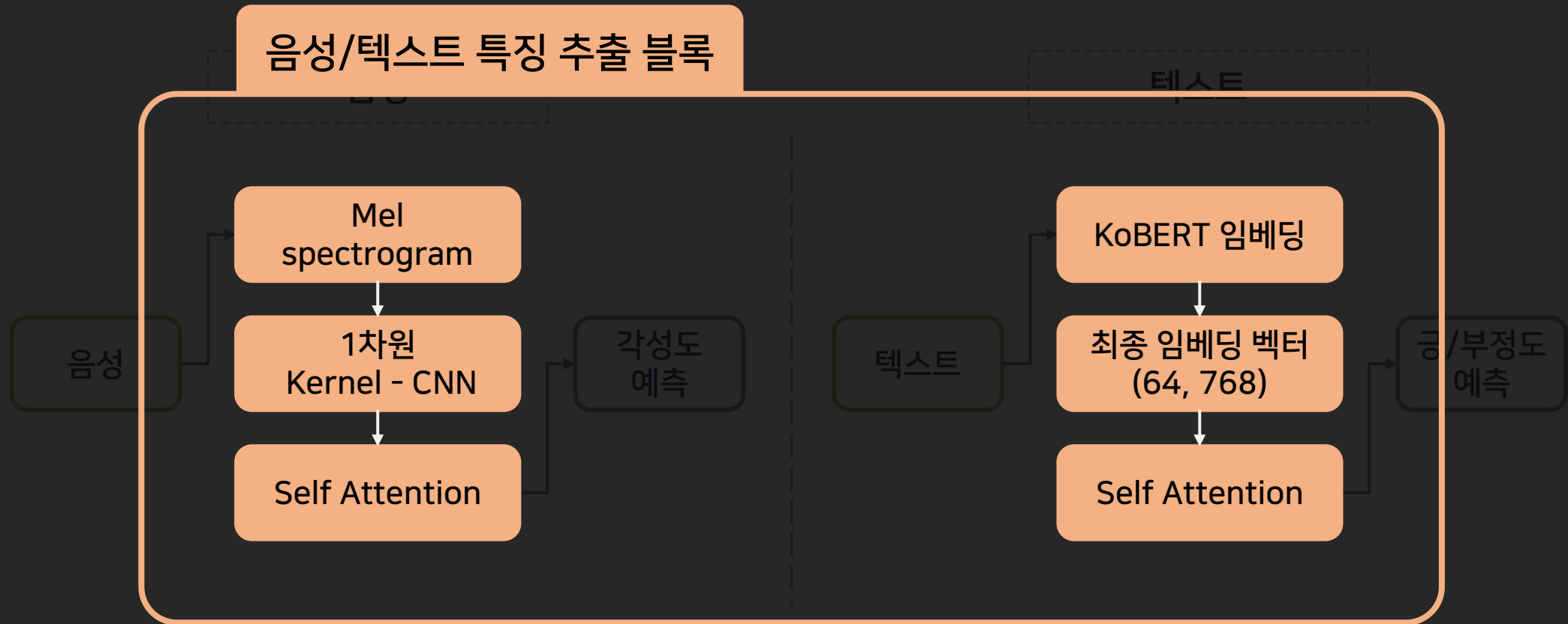
6. 모델 구성

E. 음성/텍스트 특징 추출 블록



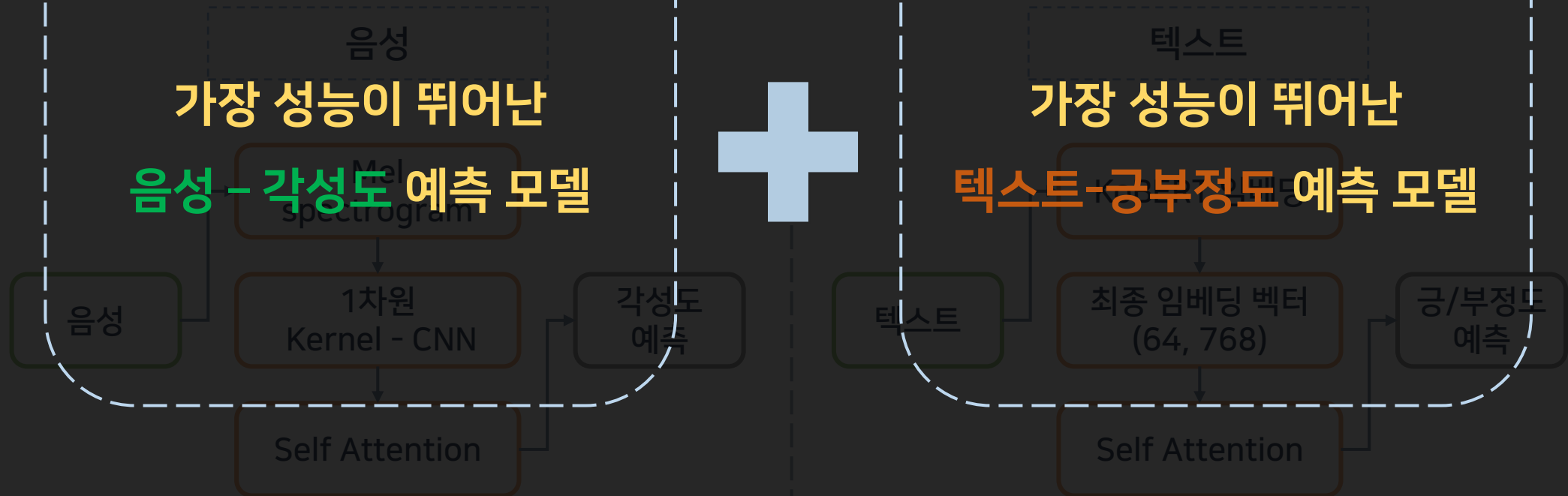
6. 모델 구성

E. 음성/텍스트 특징 추출 블록



6. 모델 구성

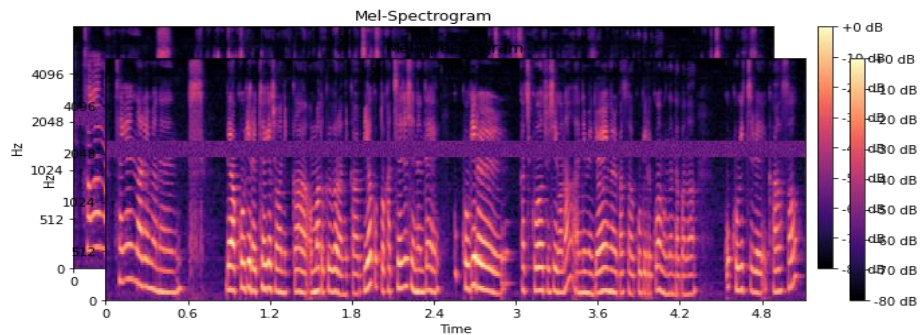
E. 음성/텍스트 특징 추출 블록



감정 분류 성능↑

6. 모델 구성

E. 음성 특징 추출 블록

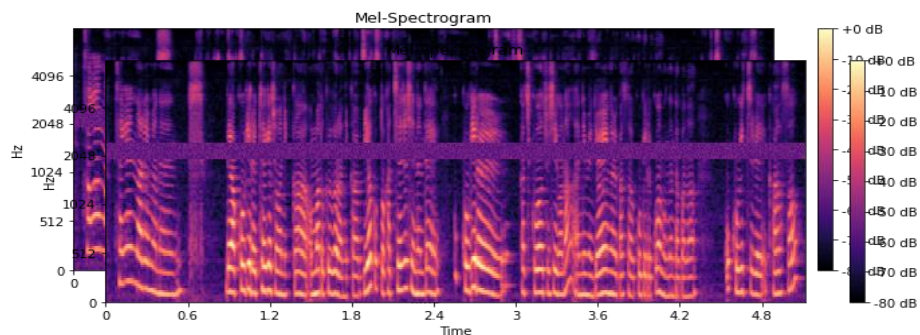


[2, 512, 128]

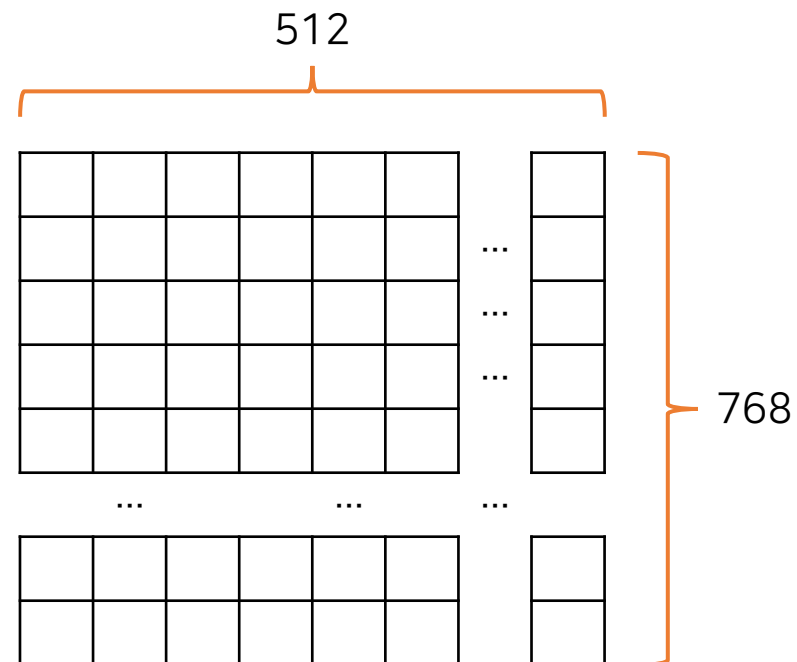
Mel Spectrogram

6. 모델 구성

E. 음성 특징 추출 블록



[2, 512, 128]

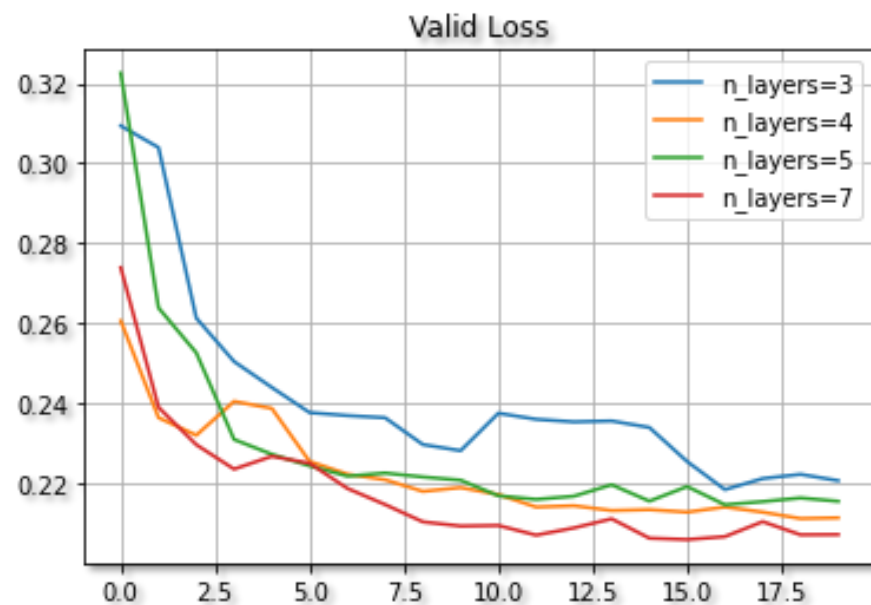
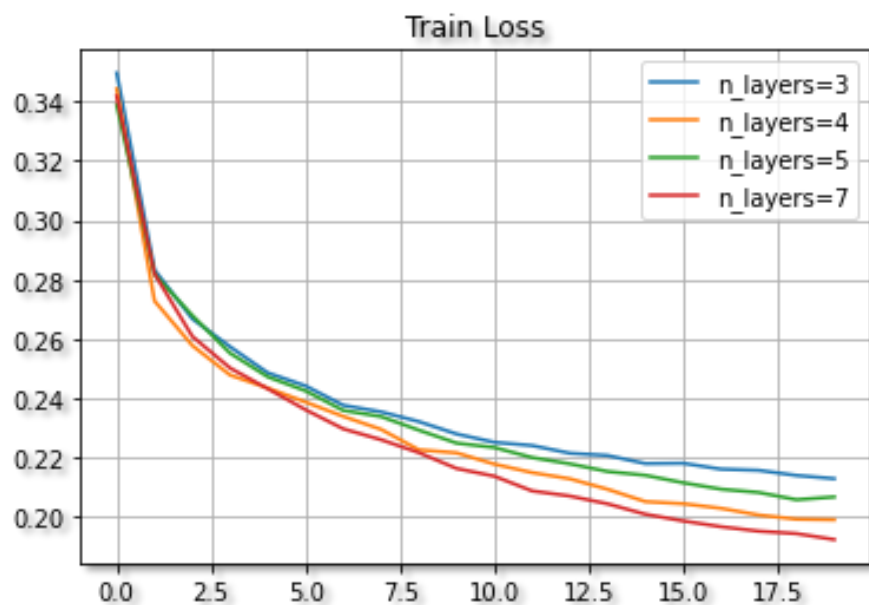


Mel Spectrogram

128차원 → 768차원

6. 모델 구성

E. 음성 특징 추출 블록



Mel Spectrogram

128차원 → 768차원

Self Attention - 7회

6. 모델 구성

E. 텍스트 특징 추출 블록

너는 대선 때 투표할 수 있어?

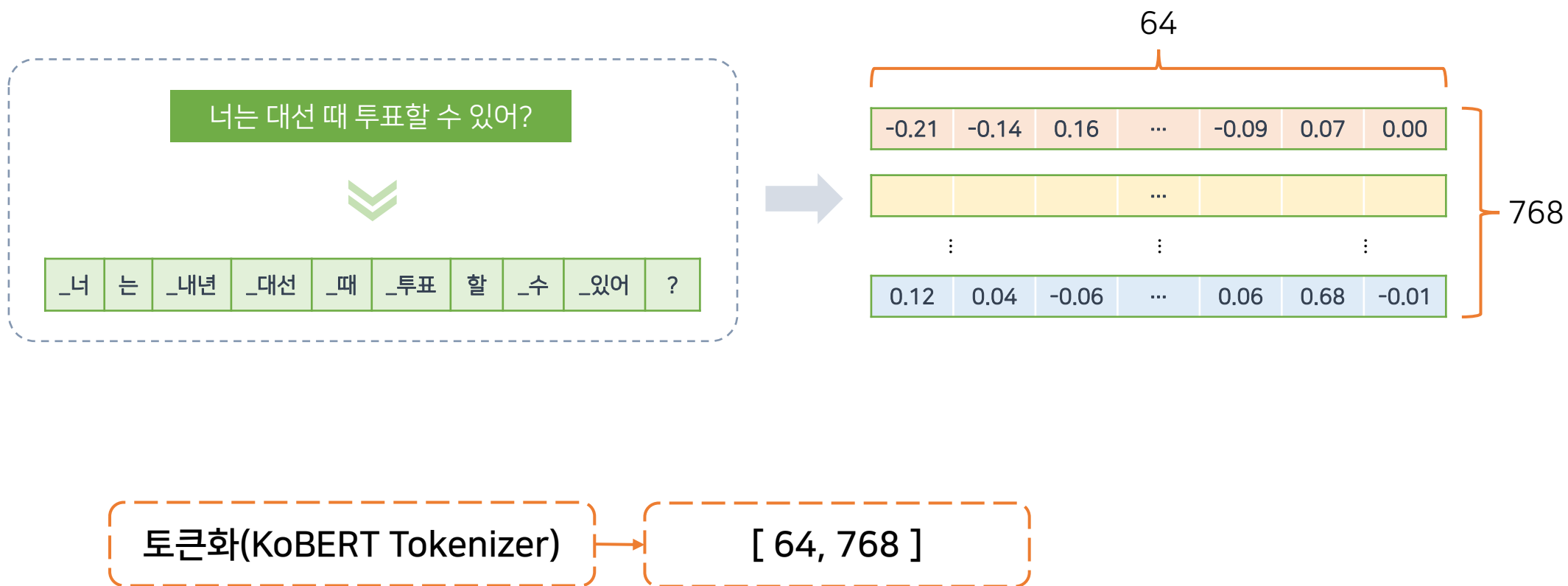


_너	는	_내년	_대선	_때	_투표	할	_수	_있어	?
----	---	-----	-----	----	-----	---	----	-----	---

토큰화(KoBERT Tokenizer)

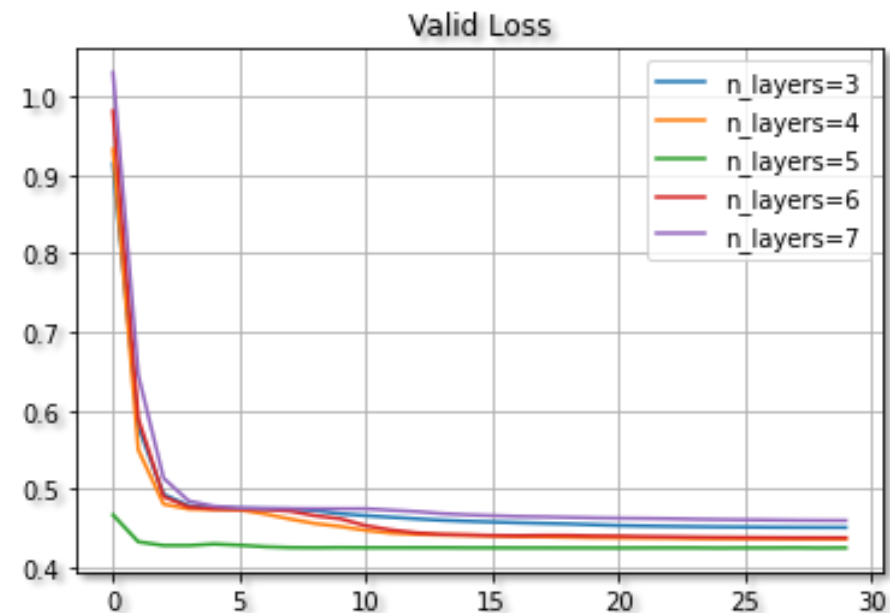
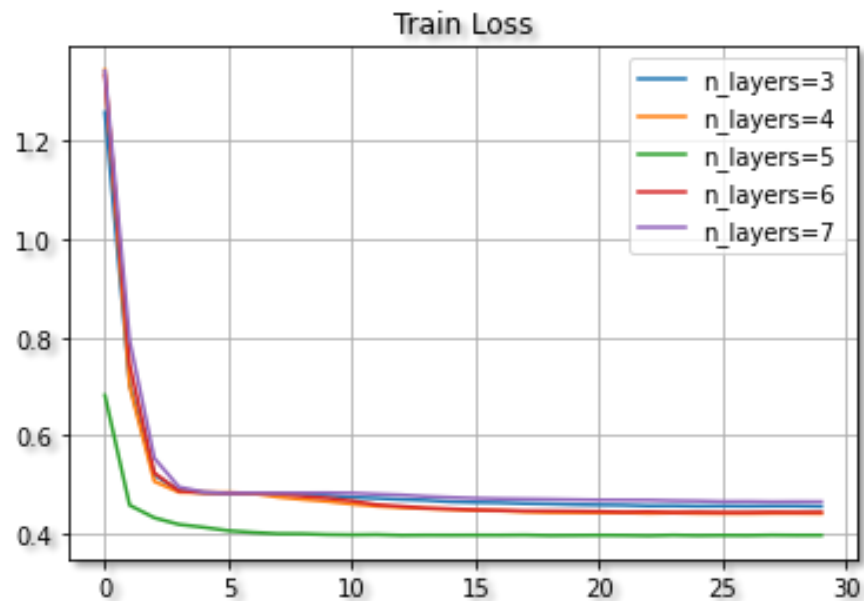
6. 모델 구성

E. 텍스트 특징 추출 블록



6. 모델 구성

E. 텍스트 특징 추출 블록



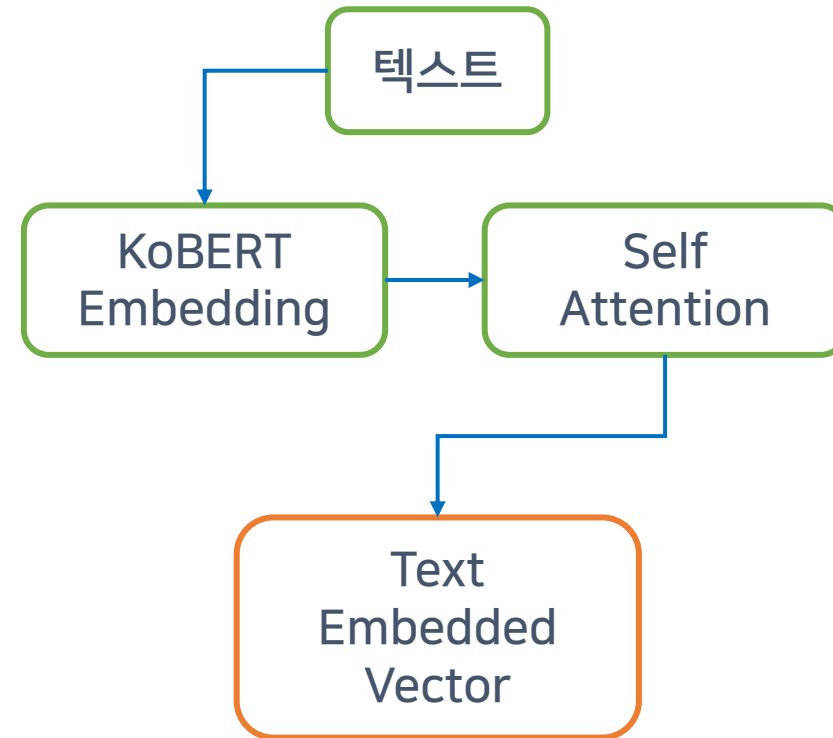
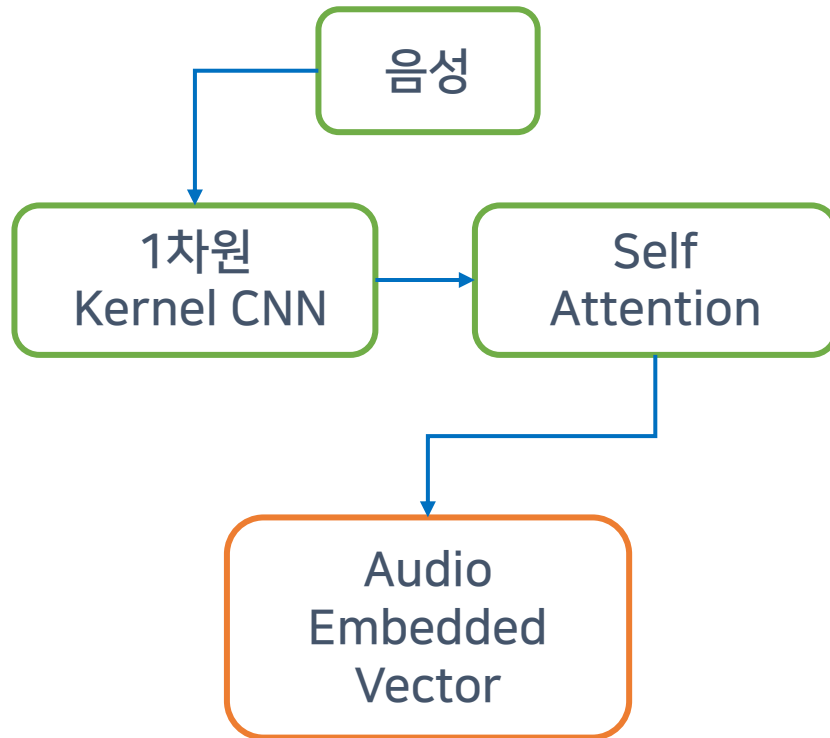
토큰화(KoBERT Tokenizer)

[64, 768]

Self Attention - 5회

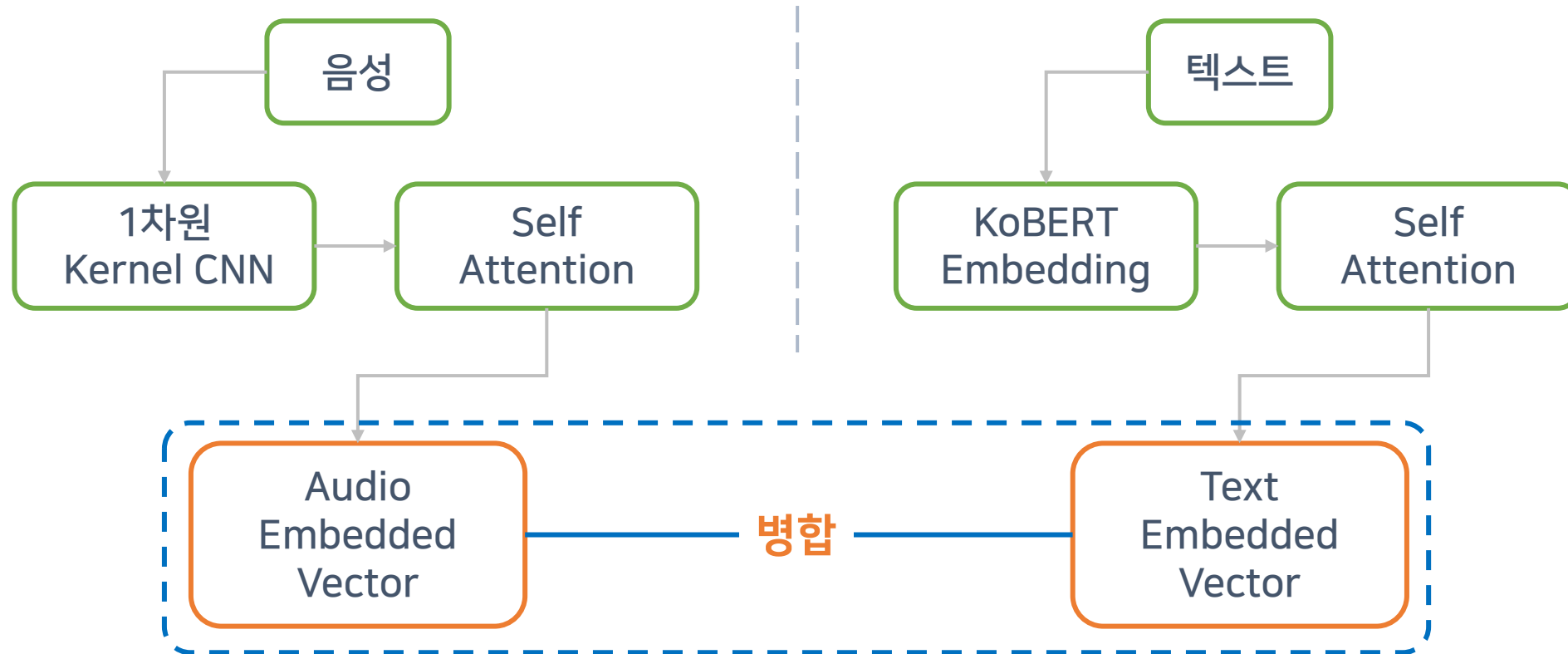
6. 모델 구성

F. 감정 인식 블록



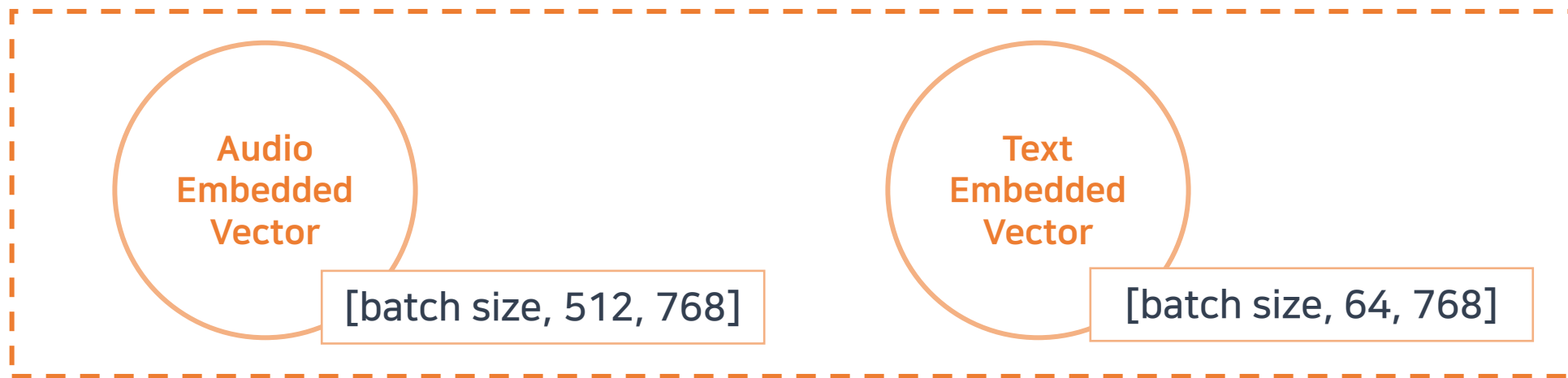
6. 모델 구성

F. 감정 인식 블록



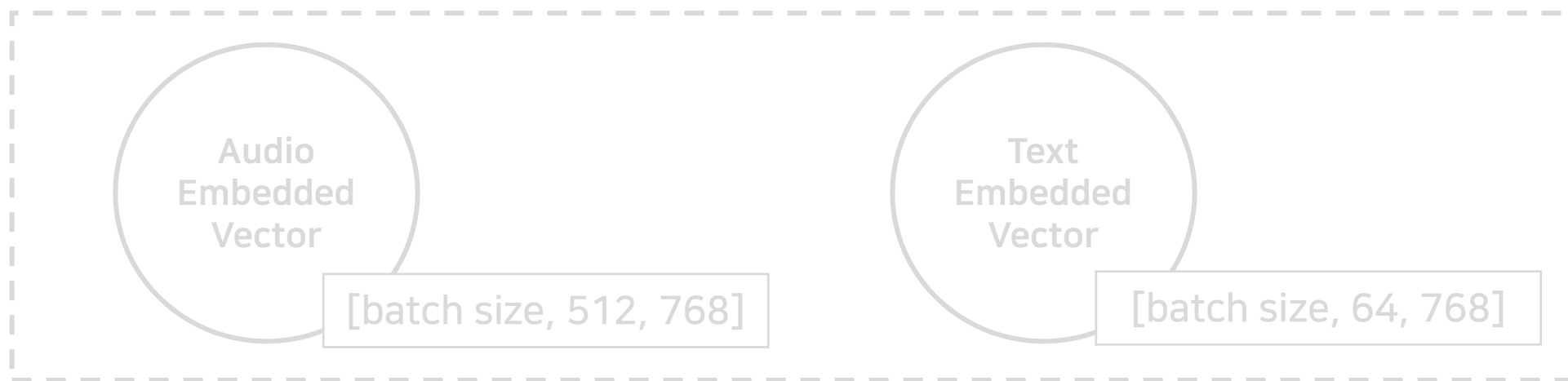
6. 모델 구성

F. 감정 인식 블록



6. 모델 구성

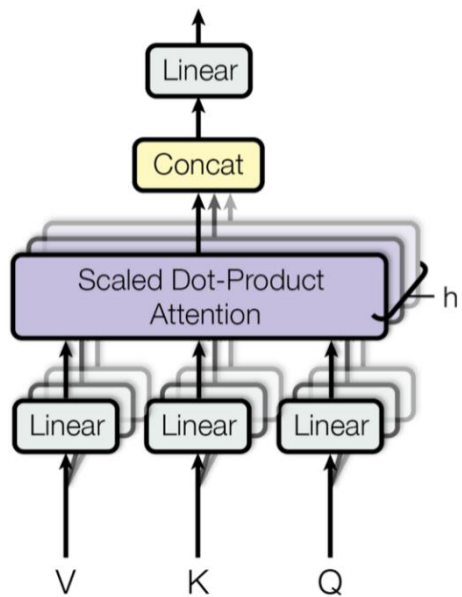
F. 감정 인식 블록



[batch size, **576**, 768] **Tensor 생성**

6. 모델 구성

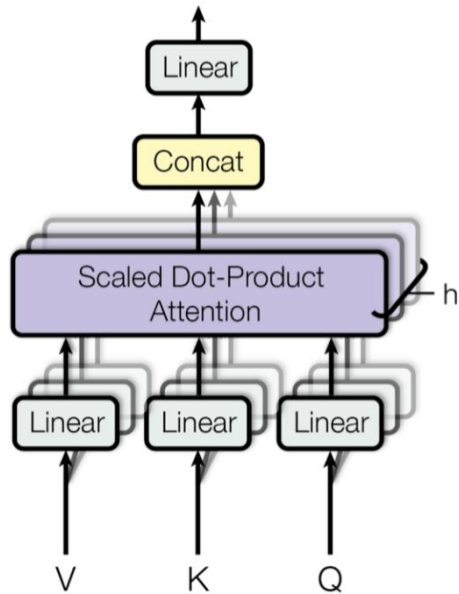
F. 감정 인식 블록



Multihead Attention
Self attention 학습

6. 모델 구성

F. 감정 인식 블록



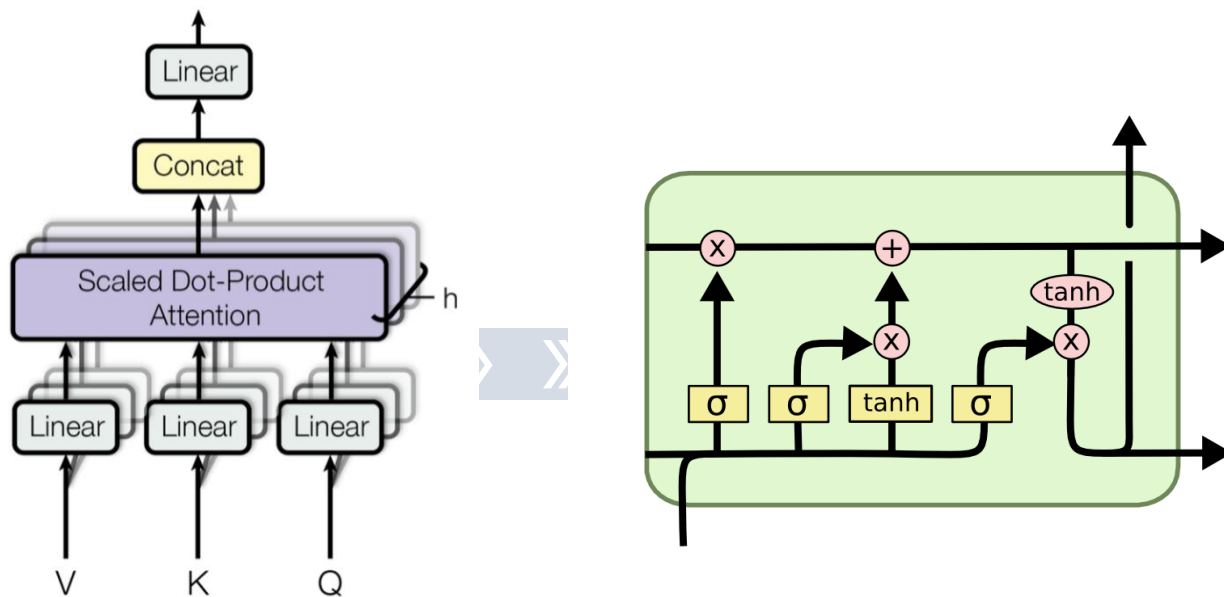
Multihead Attention
Self attention 학습

Attention
Value

데이터의 시간·위치 정보가 모두 보존됨

6. 모델 구성

F. 감정 인식 블록



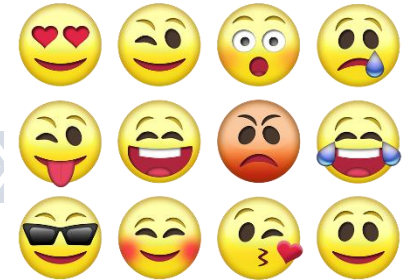
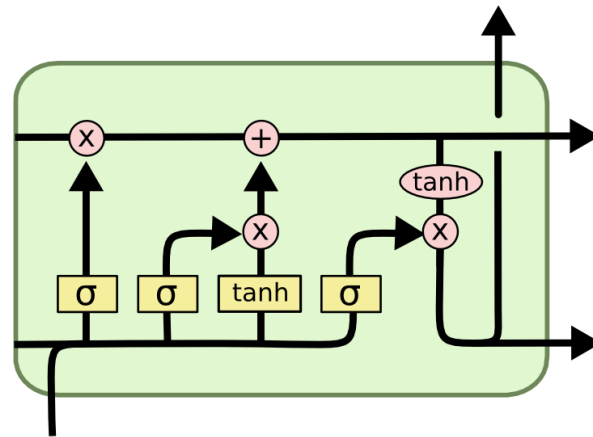
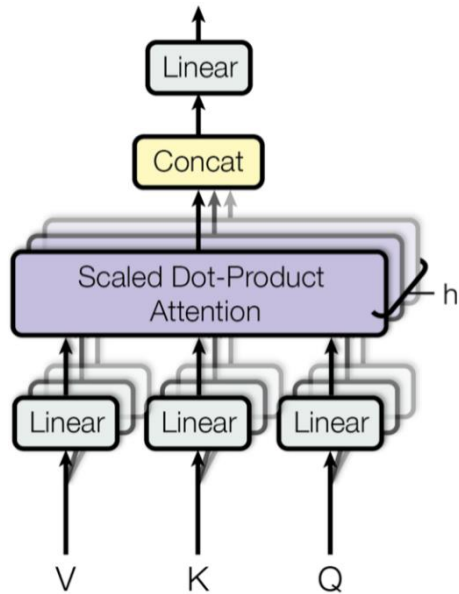
Multihead Attention
Self attention 학습

Attention
Value

LSTM

6. 모델 구성

F. 감정 인식 블록



Multihead Attention
Self attention 학습

Attention Value

LSTM

감정분류

07

모델 평가

7. 모델 평가

A. F1 Score

F1 Score란?

- : 주로 분류 클래스 간 데이터가 심각한 불균형을 이루는 경우 사용
- 정밀도와 재현율의 평균

$$F_1 = 2 \cdot \frac{1}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

정밀도(precision) : 예측한 것 중에 정답의 비율 재현율(recall) : 찾아야 할 것 중에 실제로 찾은 비율 Scikit learn 패키지의 average = weighted 사용

7. 모델 평가

모델 성능 실험

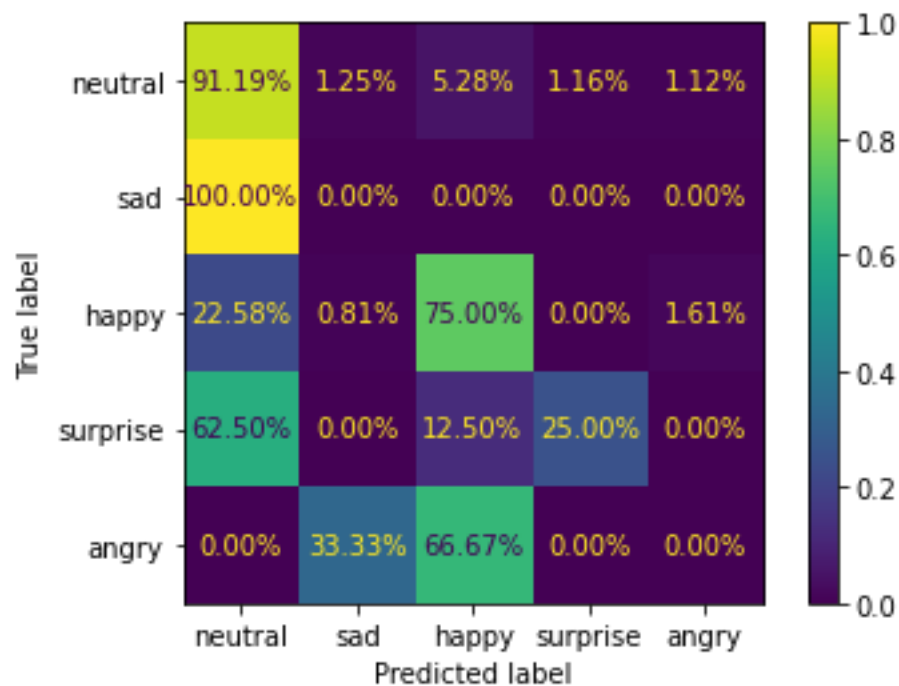
가장 뛰어난 성능을 보인 Self Attention 반복 횟수

	1번 조합	2번 조합	3번 조합
음성 블록 (n_layers / 순위)	6 / 2순위	7 / 1순위	7 / 1순위
텍스트 블록 (n_layers / 순위)	6 / 2순위	6 / 2순위	5 / 1순위

음성 - 각성도 모델과 텍스트 - 긍부정도 모델에서 Self Attention 반복 횟수에 따라 조합

7. 모델 평가

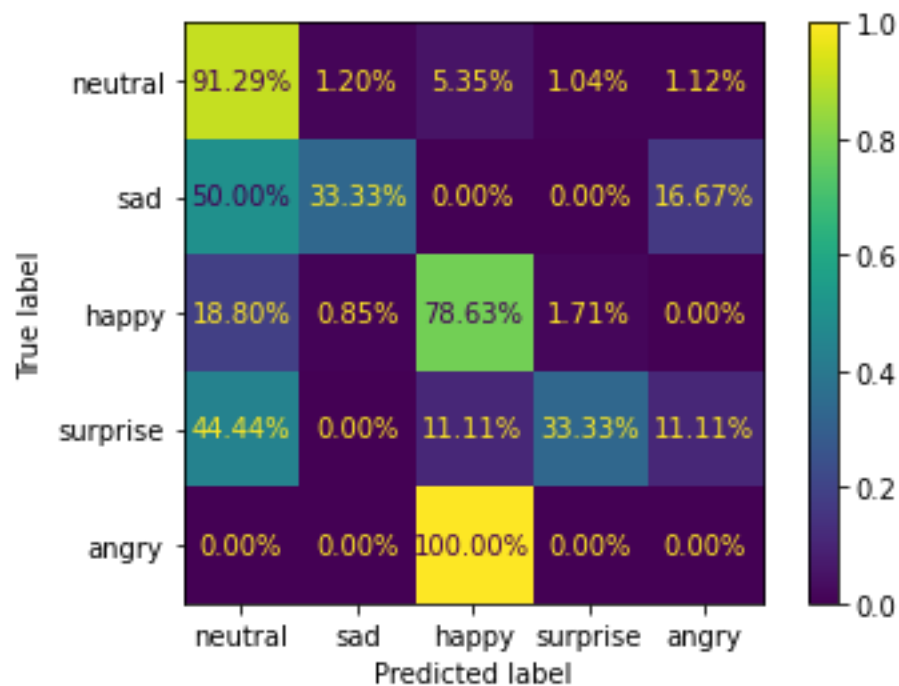
B. Confusion matrix - 1번 조합



- **F1 Score : 87.54%**
- 음성 n_layers = 6(2nd)
- 텍스트 n_layers = 6(2nd)

7. 모델 평가

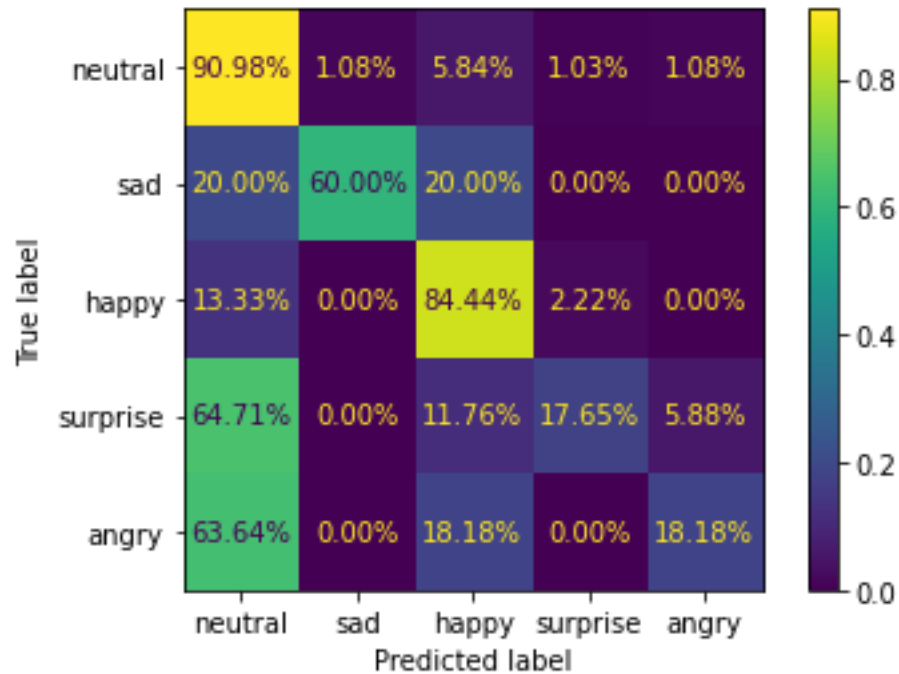
B. Confusion matrix - 2번 조합



- **F1 Score : 87.60%**
- 음성 n_layers = 7(1st)
- 텍스트 n_layers = 6(2nd)

7. 모델 평가

B. Confusion matrix - 3번 조합

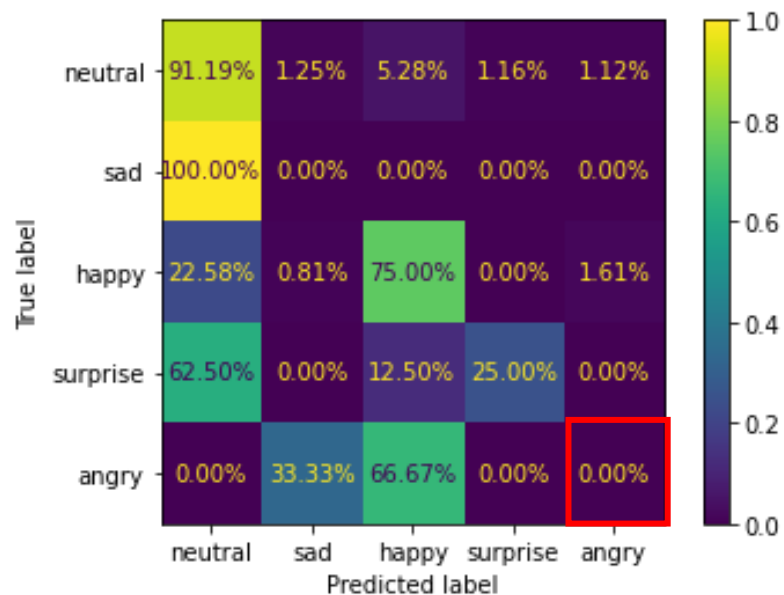


- **F1 Score : 86.93%**
- 음성 n_layers = 7(1st)
- 텍스트 n_layers = 5(1st)

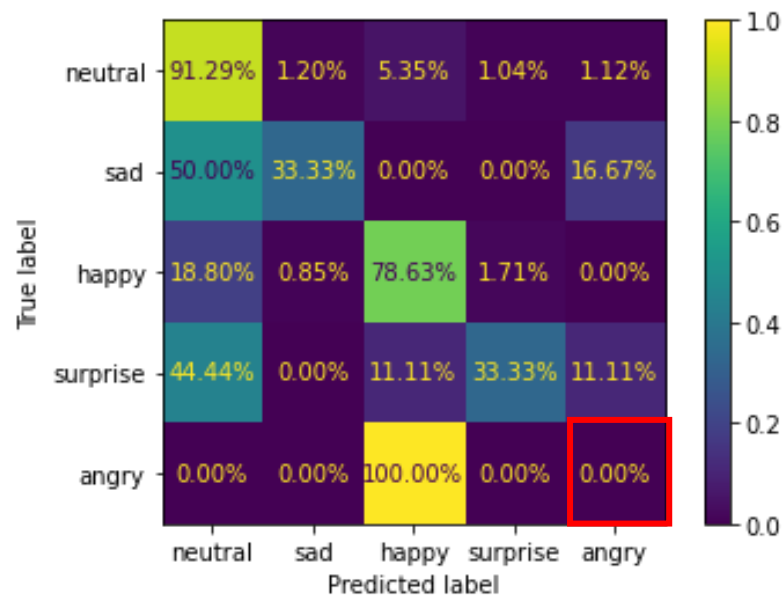
7. 모델 평가

C. 모델 비교 - 1번 & 2번

1번 조합



2번 조합

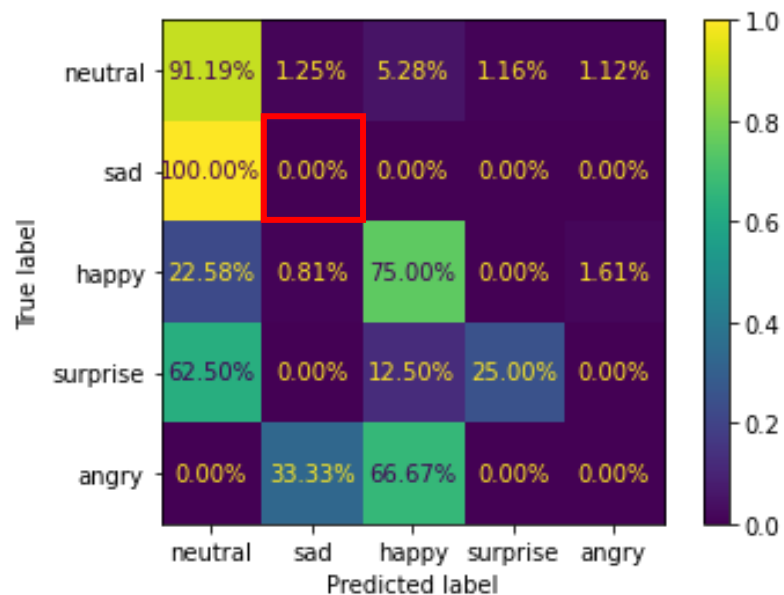


분노(Angry) 예측 정확도 0%

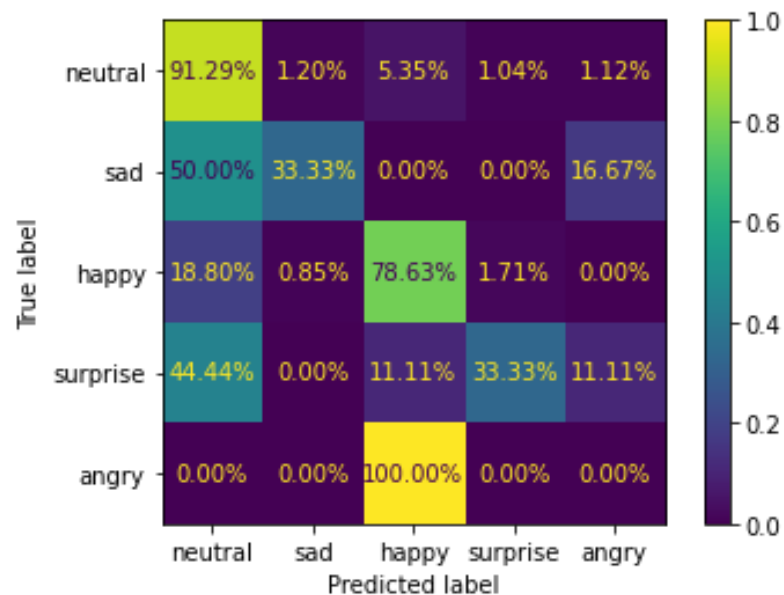
7. 모델 평가

C. 모델 비교 - 1번 & 2번

1번 조합



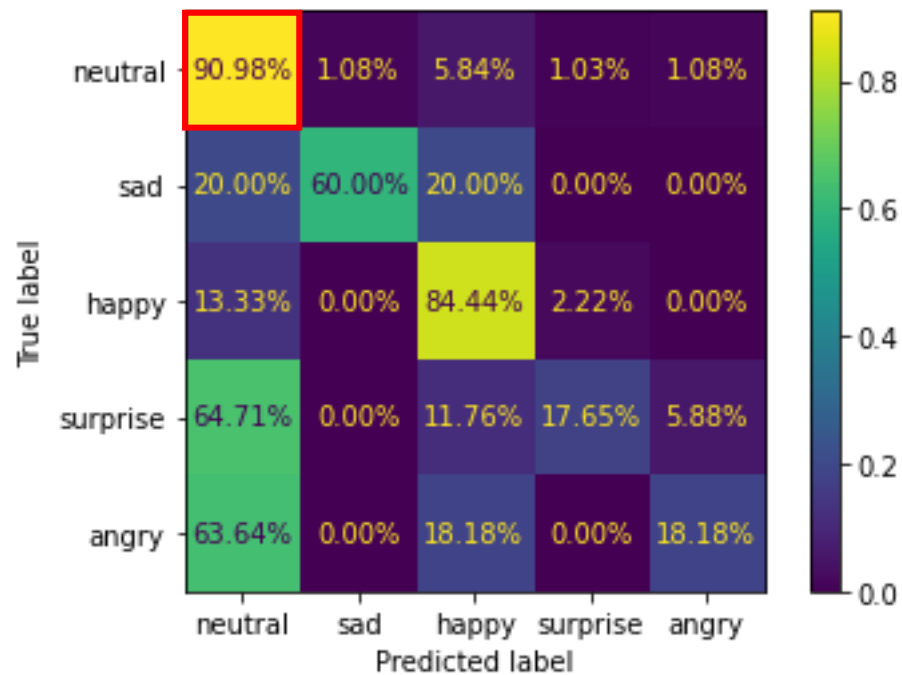
2번 조합



1번 조합 슬픔(Sad) 예측 정확도 0%

7. 모델 평가

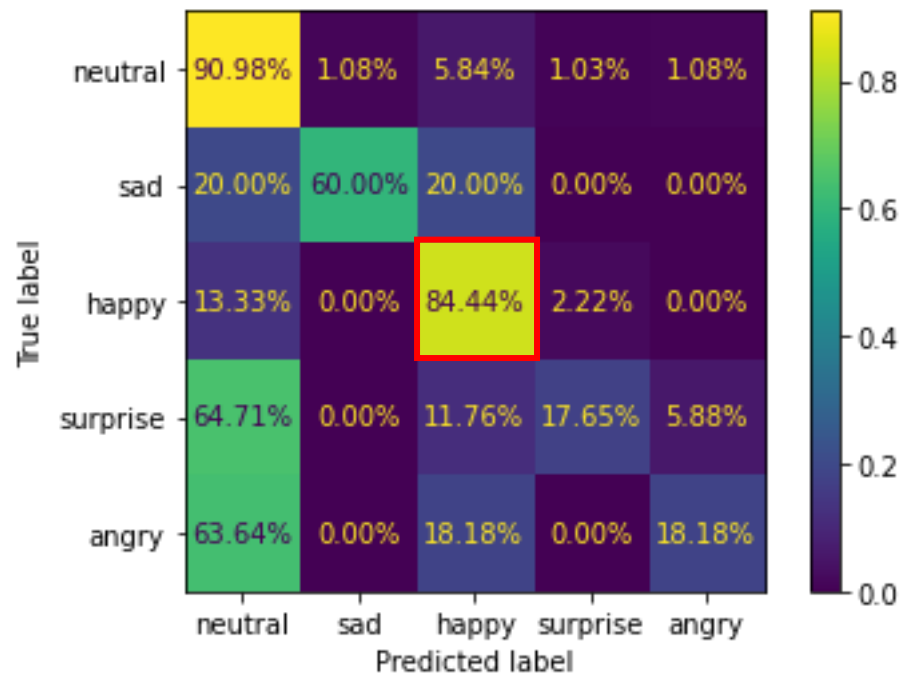
D. 최종 모델 평가 - 3번 조합



- ① 가장 데이터가 많았던 neutral 라벨
90.98%로 예측 성공
- ② Happy 라벨 또한 84.44%로 예측 성공
- ③ 데이터가 적은 세가지 라벨 모두 예측 성공 사례 존재
- ④ 특히 sad 라벨은 60%로 높은 예측 정확도

7. 모델 평가

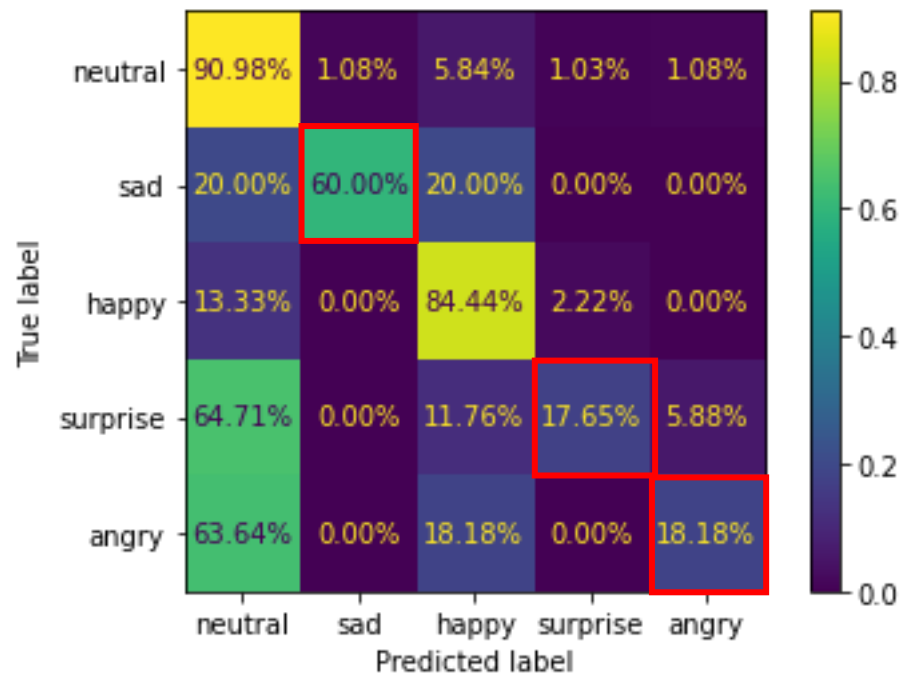
D. 최종 모델 평가 - 3번 조합



- ① 가장 데이터가 많았던 neutral 라벨
90.98%로 예측 성공
- ② Happy 라벨 또한 84.44%로 예측 성공
- ③ 데이터가 적은 세가지 라벨 모두 예측 성공 사례 존재
- ④ 특히 sad 라벨은 60%로 높은 예측 정확도

7. 모델 평가

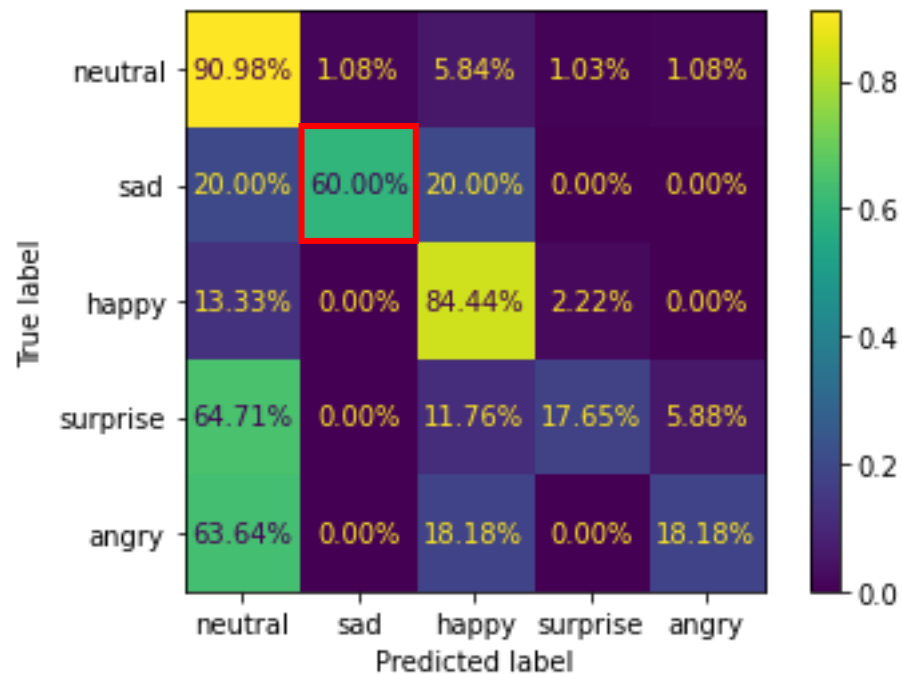
D. 최종 모델 평가 - 3번 조합



- ① 가장 데이터가 많았던 neutral 라벨
90.98%로 예측 성공
- ② Happy 라벨 또한 84.44%로 예측 성공
- ③ 데이터가 적은 세가지 라벨 모두 예측 성공 사례 존재
- ④ 특히 sad 라벨은 60%로 높은 예측 정확도

7. 모델 평가

D. 최종 모델 평가 - 3번 조합



- ① 가장 데이터가 많았던 neutral 라벨
90.98%로 예측 성공
- ② Happy 라벨 또한 84.44%로 예측 성공
- ③ 데이터가 적은 세가지 라벨 모두 예측 성공 사례 존재
- ④ 특히 sad 라벨은 60%로 높은 예측 정확도

7. 모델 평가

D. 최종 모델 평가 - 3번 조합



상대적으로

데이터 불균형에 강건!

① 가장 데이터가 많았던 neutral 라벨

90.98%로 예측 성공

② Happy 라벨 또한 84.44%로 예측 성공

데이터가 적은 surprise 라벨도 모든 예측 성공 사례 존재

④ 특히 sad 라벨은 60%로 높은 예측 정확도

7. 모델 평가

E. 의의

- ① 언어적 요소와 비언어적 요소를 복합적으로 활용하는 멀티모달 감정인식 모델 제안
- ② 심한 데이터 불균형을 일정 부분 극복할 수 있는 가능성 제시
- ③ 멀티모달 딥러닝 모델에서 데이터 증강의 효과 확인
- ④ 사전학습 기반 멀티모달 모델의 예측 성능 확인

감사합니다

Q & A