

# 2023

군장병 부문

# 국방 AI 경진대회

# 결과물 발표평가



팀명 | ROKAFNET



대한민국 국방부  
Ministry of National Defense



과학기술정보통신부  
Ministry of Science and ICT



**김진수**  
공군사관학교  
정보통신대대



**이민준**  
제3미사일방어여단  
정보통신대

# ROKAFNET.



**박성준**  
공군사관학교  
정보통신대대



**박진영**  
공군사관학교  
정보통신대대

# II 발표 내용\_ [평가 항목] 문제 이해도

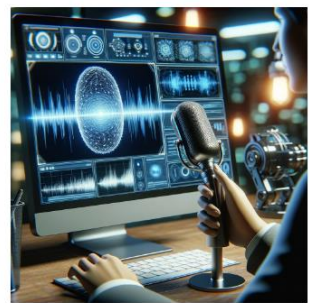


“ 음성인식 기술을 활용한 효과적인 지휘관의 전술 명령 전달 ”

## Automatic Speech Recognition (자동 음성 인식)

### 전략1. 노이즈 제거

음성 인식 모델 개발 경진대회



AI 음성 인식은 인공지능 기술을 활용하여 사람의 음성을 이해하고 그 의미를 파악하는 기술입니다. 이 기술은 음성 명령을 컴퓨터가 이해하고 처리할 수 있도록 해주며, 음성으로 입력된 정보를 텍스트나 명령어 등으로 변환합니다.

군사 작전에서 지휘관의 명령은 매우 중요합니다. 특히 소음이 많은 환경에서도 명료하고 정확한 음성 명령은 작전 수행의 핵심 요소입니다. 그러나 소음이나 다른 환경적 요인으로 인해 음성 명령의 전달이 어려울 수 있습니다.

전투 상황에서 군장병들은 지휘관의 명령을 정확하게 이해하고 신속하게 실행해야 합니다. 따라서 음성 통신의 효과성은 군사 작전의 성공에 매우 중요한 역할을 합니다.

전투나 군사 작전 시 소음이 발생하는 환경에서 음성 명령의 전달은 어려울 수 있습니다. 이러한 환경에서 효과적으로 음성을 전달하고 전사하는 모델의 필요성이 대두되고 있습니다.

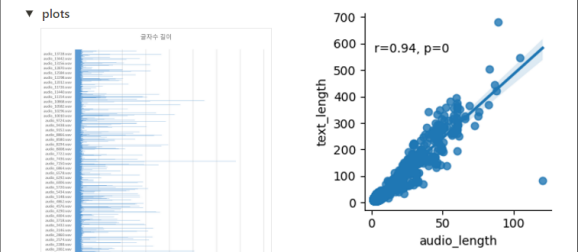
소음이 많은 환경에서도 명료하고 정확한 음성 명령은 작전 수행의 핵심 요소입니다. 그러나 소음이나 다른 환경적 요인으로 인해 음성 명령의 전달이 어려울 수 있습니다.

군에 실 적용 시, 노이즈 제거가 가장 중요하다.

### 전략2. 텍스트 EDA

#### EDA 결과

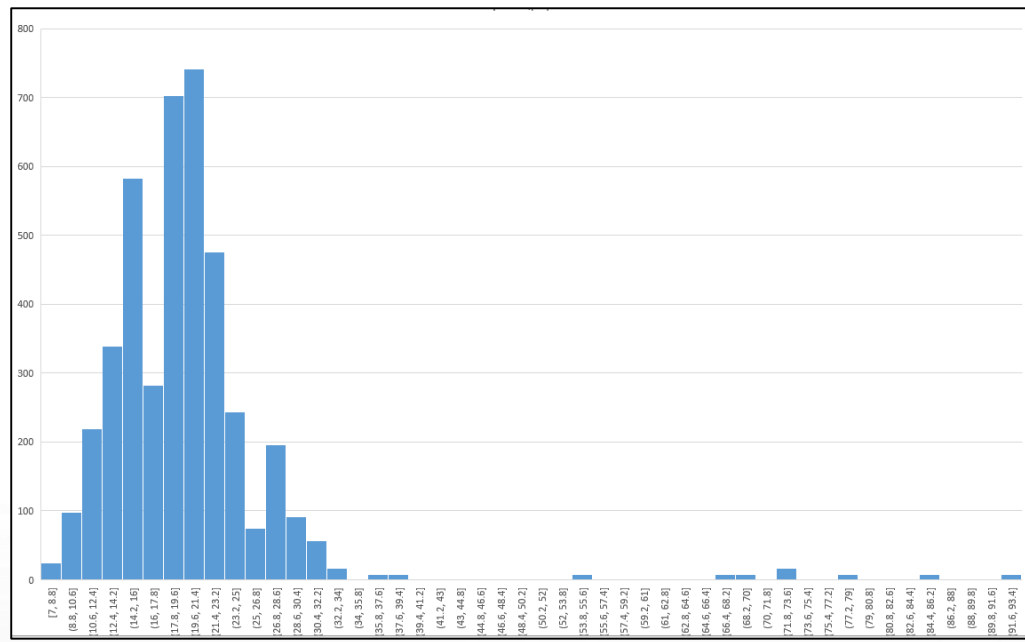
- Train Data
  - 14000개, 16.3 시간
  - 4447개의 다른 text로 이루어져 있음 → 같은 텍스트가 많지만, 서로 다른 사람의 목소리임.
  - 7.5초 이상의 데이터에는 라벨이 걸치는 데이터가 없음. 데이터의 성격도 다름 → 다른 접근이 필요할 가능성 높음.
  - 길이 매우 다양(2초~1분) → 분포 고려해야 함
  - 잡음(주로 백색소음) 매우 심함 + 목소리 불륨 작아서 잘 안 들림
    - Denoising + ASR 결합 구조도 염두에 둘 것
  - sampling rate: 16000 → resampling 필요 X
  - punctuation 없애기
  - Text의 대부분의 영어와 숫자는 한국어로 표현
    - 예외) audio\_91.wav의 경우 'tsb'
    - (발음 단양 주의: ex) audio\_51.wav에서는 '티브이', audio\_26.wav에서는 '티브이', 7296.wav에서는 'tv')
  - 공백으로 끝나는 텍스트도 있음(trimming 필요할 것으로 보임)
  - 전체 서로 다른 텍스트는 4449개
  - "부탁해" → 108개 (4%), "알려줘" → 76개 (1~2%) "가까운" → 24개



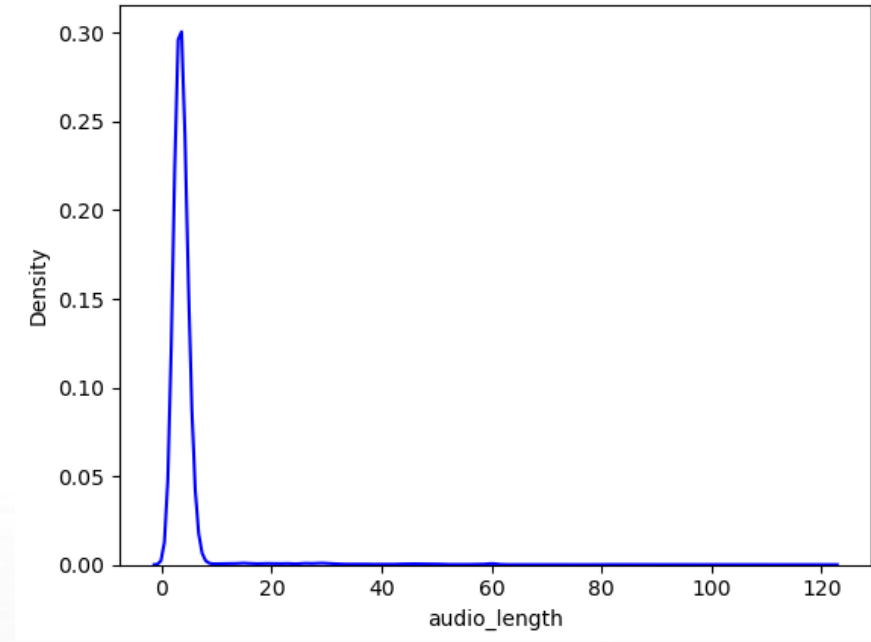
“ 텍스트 자체를 보면서 텍스트의 특징을 분석해야 한다.”



## 텍스트 EDA



텍스트 수



오디오 길이

**길이 분포가 편향적**



## 텍스트 EDA

### TEXT 길이 < 41

audio_10.wav	속소에 휴식 시설 있으면 안내해 줄래?	21 ?
audio_100.wa	다섯시 드라마 놓치지 않게 예약 좀 해주게.	24 .
audio_1000.w	잠 다 깬으니까 깨우지 마.	15 .
audio_10000.	내과 접수처에 접수를 해야 하는데 어디서 해?	25 ?
audio_10001.	가정통신문에 적힌 거 다 보여 줘.	19 ?
audio_10002.	공항 부근에 지금 예약 가능해?	17 ?
audio_10003.	근처 숙박 가능한 방 알려 주게.	18 .
audio_10004.	속소에 주차한 기간 동안 발생한 주차 요금 정산해 줘.	30 .
audio_10005.	내과에 홍길동 선생님은 언제 진료 가능해?	23 ?
audio_10007.	휴식 주변에 가 볼 만한 곳 있어?	19 ?
audio_10008.	여기 망고 들어있니?	11 ?
audio_10009.	네, 세탁기 다 돌아갔습니다.	16 .
audio_1001.w	일회용 교통카드 보증금으로 바꿔 주시요.	22 .
audio_10010.	영화관에서 가장 가까운 오락실은 어디로 가야 하는지 말해줘.	33 .
audio_10012.	채소 합인 시작은 언제 해?	19 ?
audio_10014.	가장 싼 호텔 알려줘.	12 .
audio_10015.	외래 진료 예약 시간 변경?	15 .
audio_10016.	서울대병원 시설 지도 띄워주세요.	22 .
audio_10017.	노래 어떤 게 있는지 검색해 드릴게요.	21 .
audio_10018.	내일 숙제 다 한거 맞아?	14 ?
audio_10019.	티비가 화면 꺼져 있나?	13 ?
audio_1002.w	기저귀 교환대 설치되어 있는 방으로 예약 부탁해.	27 .
audio_10020.	근처에 고객 실터가 마련되어 있습니다.	21 .
audio_10022.	자주 듣는 이 노래 틀어 줄 수 있어?	21 ?
audio_10023.	일급 시 되면 또 부르게 즐겨찾기 해.	21 .
audio_10024.	호텔에서 서울역까지 가는 데 몇 분 걸리니?	24 ?
audio_10025.	은 게 없어서 슬퍼요.	12 .
audio_10026.	유제품은 어디가 가장 싸게 팔아요?	19 ?
audio_10027.	언제부터 하는지 확인해.	13 .
audio_10028.	예약할 수 있는 좌석은 어떤 게 있어?	21 ?
audio_10029.	몇 시간 뒤에 목적지까지 도착하는지 알려줘!	24 !
audio_1003.w	내가 티비를 켜고 싶은데 켜 놓여?	19 ?
audio_10030.	간병사 불러 주면 좋겠어.	14 .

**가장 싼 호텔 알려줘.**

### 명령문, 안내 위주의 발화

### TEXT 길이 >= 41

audio_10011.	현재 대학교 1학년 전공 과정을 준비하면서 아직은 반도체 전공정이 무엇인지 제대로 배우지 못했습니다. 하지만 북한	143 .
audio_10013.	신선하고 맛있는 커피를 즐기기 위해서 제가 커피를 먹고 싶을 때 사서 먹는 게 그게 제일 즐길 수 있고 맛있게 먹을 수	77 .
audio_10021.	오늘 꼭 마무리해야 하는 업무가 있다면 우선 저에게 맡은 바를 충실히 다 이행할 후 남은 부분은 같은 경우는 각각의	222 .
audio_10039.	디지털이 보편화되면서 이를 제대로 활용하는 게 중은 지식이 늘어나고 소득도 증가하는 반면에 디지털을 이용하지 못	117 .
audio_1004.w	저는 저나 타인을 모두 포함해서 자존감이 조금 낮았을 때 제 의견에 대한 확신이 없었던 경험이 있습니다. 자존감이 높	358 고
audio_10047.	소통이란 서로의 입장을 이해하고 합의점을 찾는 것이라고 생각합니다. 특히 대학 행정의 경우 여러 부서 간의 협업을	251 .
audio_10080.	협상이라고 하는 것은 일반적인 회사에서뿐만 아니라 사람과의 관계에서 반드시 필요한 것이라고 생각을 합니다. 협상	145 .
audio_10089.	제가 생각하는 amd 직무에서 필요한 역량은 마켓 분석 능력입니다. amd 직무와 같은 경우 상품 기획을 하거나 채고와	422 .
audio_10153.	우선 쿠키는 클라이언트 쪽에 로컬에 저장되는 키 값이 들어있는 작은 데이터 파일이고 이것은 이제 어디에 저장하느니	178 개
audio_10196.	저의 경쟁력이 여러 가지가 있는데요. 남들보다 뛰어나다고 생각하고 있는 부분에 대해서 질 수 없다고 생각을 합니다.	177 개
audio_10199.	취업 준비를 하면서 스스로 두 가지 기준을 정했었습니다. 첫째는 화장품이었고 둘째는 전략이었습니다. 감사하게도 전	343 .
audio_10267.	자신의 일을 중요하게 느끼며 성취감을 느끼는 것이 가장 중요하다고 생각합니다.	43 .
audio_10345.	시장 조사를 할 때 구글 서칭을 많이 하고 요구한 회사나 그래서 많이 조사를 하면서 시장 조사를 하고 있습니다.	62 .
audio_10373.	저는 10점 만점에 8점 정도 된다고 생각합니다. 이 점은 경희의료원에 입사하면 채워질 것으로 생각되는데요. 8점인 이	141 .
<b>자신의 일을 중요하게 느끼며 성취감을 느끼는 것이 가장 중요하다고 생각합니다.</b>		
audio_10431.	저는 주로 협상할 때 먼저 들어주는 먼저 그 어떤 일인지 물어줍니다. 그리고 그다음 제 생각을 얘기한 뒤 최종안을 하	76 .
audio_1056.w	인간관계에서 가장 중요한 것은 대화와 소통이 중요하 것 같습니다. 각각의 인간들은 성격이 너무나 다르고 가치관이 나	131 .
audio_10586.	저를 만들어 준 사람을 부모님, 자랑 이야기해 주는 사람을 친구로 생각하고 있어요	45 요
audio_1064.w	대변과 차변 차변은 왼쪽이고 대변은 오른쪽입니다. 차변에는 자산과 관련된 항목들이 위치하고 대변에는 부채와 자본	161 .
audio_10682.	제가 인간관계에서 가장 중요하게 생각하는 것은 관심이라고 생각합니다. 모든 사람의 인간관계는 관심에서 시작된다.	446 .
audio_10683.	고등학교 때 제가 원하는 동아리를 들고 싶어서 직접 동아리를 설립하고 동아리 회장이 되었던 적이 있습니다. 저한테는	179 .
audio_10786.	sns를 크게 활용하고 있던 않지만 기본적으로 sns 관련된 여러 가지 정보들은 알고 있습니다.	53 .
audio_10851.	저는 그 한 사람을 이해하는 것이 가장 중요하다고 생각합니다. 사람들은 모두 다른 성격과 특성을 가지고 있으며 그 한	165 .
audio_10888.	우선 승무원의 직급은 주니어와 시니어 그리고 사무장이 있다고 알고 있습니다. 각 직급별로 담당하는 구역도 다른데	318 .
audio_109.wa	만약에 저는 먼저 보고를 할 것 같습니다. 왜냐하면 저 혼자 윤리적인 윤리적인지 않냐 윤리적인 판단을 하면 분명히	202 .
audio_10928.	첫째로는 진단평가에 인공지능 적용이 시급하다고 생각합니다. 왜냐하면 아이들의 시작점 분석을 통해 1년의 학습 시	548 .
audio_1097.w	제 삶에 가장 영향력을 미친 사람은 아버지입니다. 아버지는 지금까지 살아오시면서 여러 가지 다양한 분야에 몸을 담	295 .
audio_10982.	10년 뒤에 저는 전문성을 가진 공공기관 직원으로서 성장해 있는 모습을 기대할 수 있는데요. 국립박물관문화재단에서	137 .
audio_11016.	저의 약점은 한 문제에 대해 집착을 한다는 것입니다. 해결되지 않은 문제에 대해서 해결하기 위해 시간과 노력을 매우	262 .
audio_11086.	우선 손님을 모시기 전에 안전 먼저 체크하는 것이 우선이라고 생각합니다. 서비스가 먼저가 아닌 안전요원에 중점을	130 .
audio_11095.	준비물을 꼼꼼하게 체크하는 것이 중요합니다. 무엇보다도 그전에 여행사 보험 교통편 숙박에 대한 일정들을 꼼꼼히	89 .

### 자기소개, 발표 위주의 발화

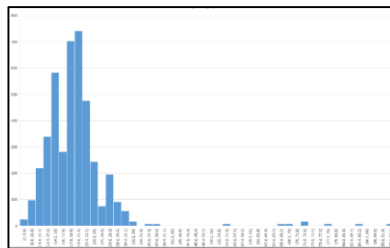
## 1. LOOK FOR

명령문, 안내 위주의 발화 => ' Short Text '

자기소개, 발표 위주의 발화 => ' Long Text '

## 2. INSIGHT

Short Text 중심으로  
학습하는 모델 설계



Long Text는 325개  
Short Text는 13675개  
Short Text의 비율 전체의  
**97.7%**

## 3. PRIORITY

한정된 시간을 효율적으로  
사용하기 위한 우선순위

길이 기반 Data 선별



Training



Hyper-parameter 조정



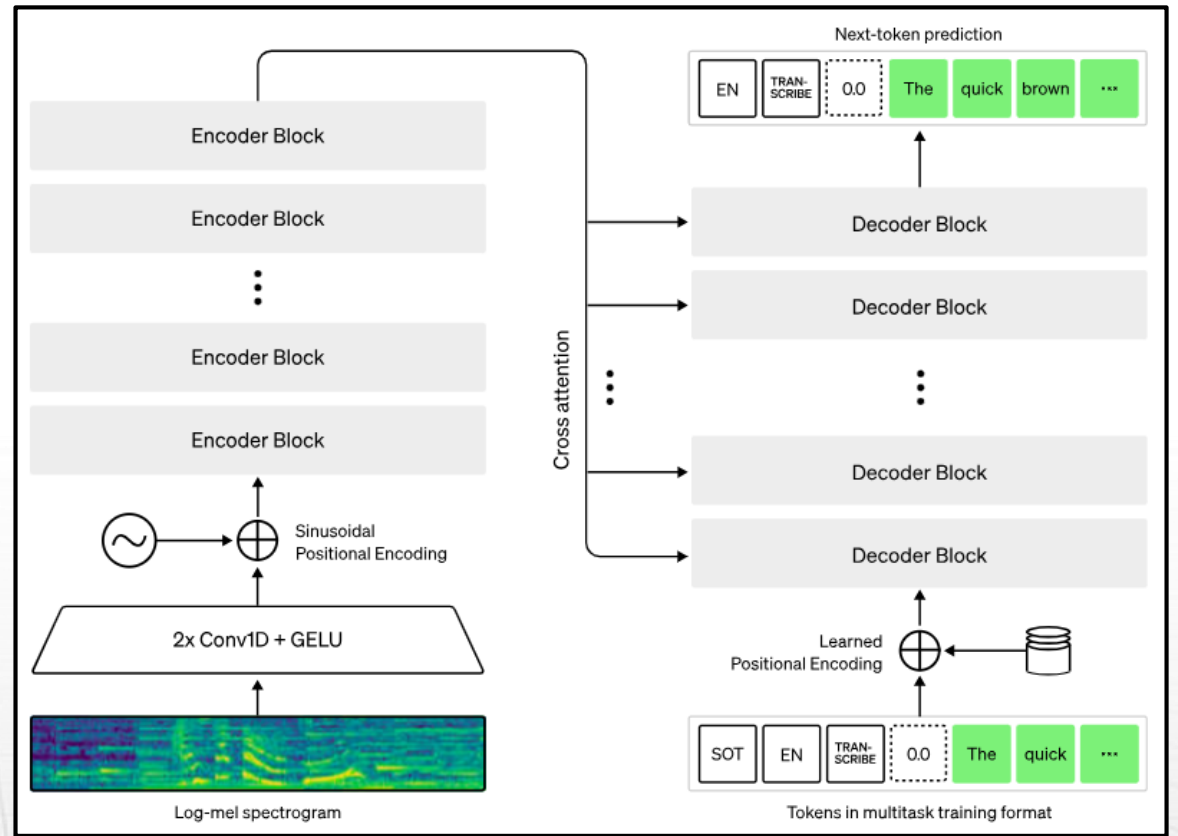
## Model

### Whisper

- ASR 분야 SOTA 모델, **소음**, 억양 등에 **Robust**
- 한정된 자원 고려 'small' model 선정

[seastar105/whisper-small-ko-zeroth](#) like 2

Size	Parameters
tiny	39 M
base	74 M
<b>small</b>	<b>244 M</b>
medium	769 M

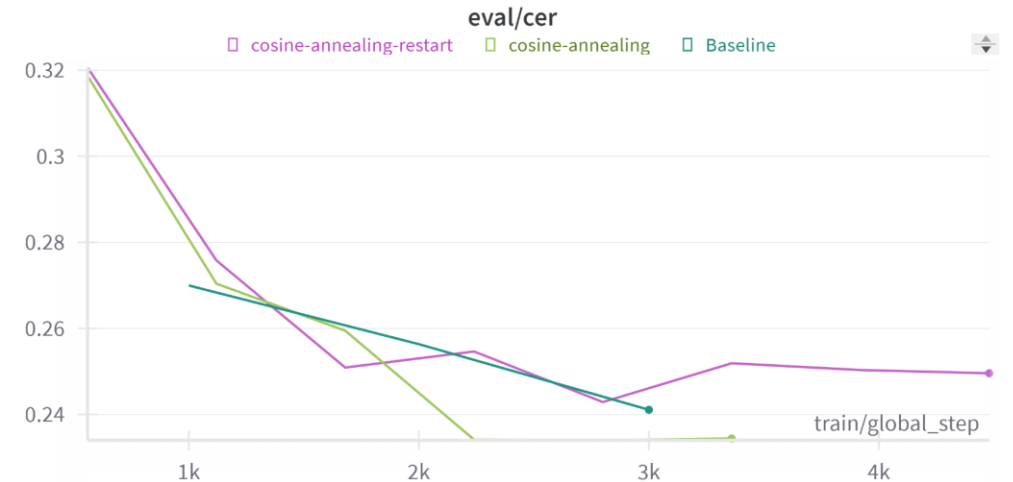


## Model

### Optimizer & Hyperparameters

- Wandb + Notion 이용 실험 프로세스 관리
- Scheduler: Val CER 이용 결정

실험 관리		
Aa branch name	☰ baseline model	# Public Score
📄 <a href="#">cosannealing + postprocess(0.7)</a>	seastar105/whisper-small-ko-zeroth · Hugging Face	0.86572373494 55758
📄 <a href="#">cosannealing + postprocess(0.71)</a>	seastar105/whisper-small-ko-zeroth · Hugging Face	0.8649
📄 <a href="#">cosannealing + postprocess(0.72)</a>	seastar105/whisper-small-ko-zeroth · Hugging Face	0.86478116823 3087



Scheduler	Validation CER
Linear	0.7589
<b>Cosine Annealing</b>	<b>0.766</b>
Cosine Annealing + Restart	0.7571



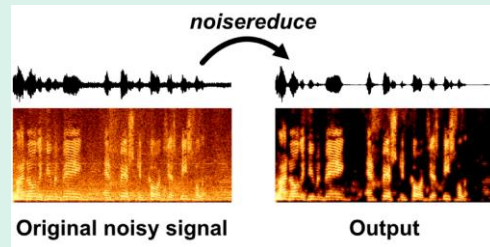
## Pre-process

**Input  
Audio**



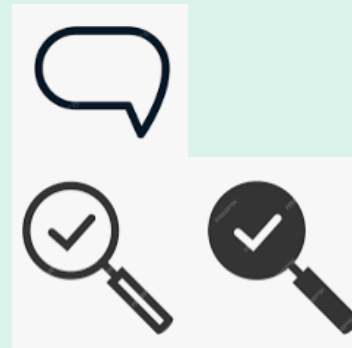
### 1. 노이즈 제거

(5479.wav)에  
1분 가량의 Pure Noise가 있다.  
→ Noise Crop 후 통계량 이용  
Denoising 수행



### 2. 데이터 정제

공백, 문장 부호 제거



**Dataset**

## Post-process

Model



### 1. 반복 문자 제거

근처 구구산단센터 근처 근처 근처 근처 근처 근처  
예약 시간 예약 시간 예약 시간 예약 시간 예약 시간  
숙소에서 할인 할인 할인 할인 할인 할인 할인 할인

결과 Submission에서 1~2단어  
주기로 반복되는 패턴이 발견됨.  
반복 패턴들을 제거해주는  
후처리 알고리즘을 구현함.

### 2. 철자 오류 수정

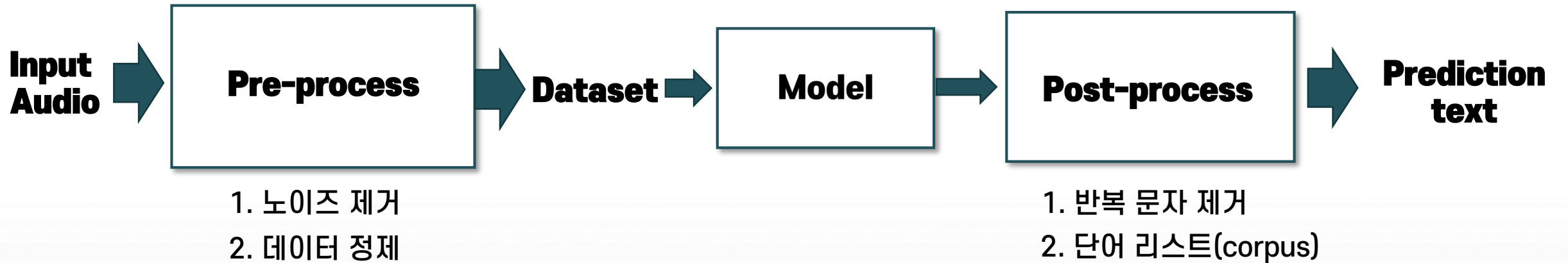


- 1) Train Data의 Label을 이용해  
주로 나오는 단어 리스트(Corpus)를 만들고
- 2) 유사도가 높을 시  
Corpus에서 단어를 가져와 사용하는  
방식을 구현함.



Prediction  
text

# II 발표 내용\_ [평가 항목] 수행 프로세스





## 2023 국방 AI 경진대회



# Where AI & Defense Connect

소감 한마디

훌륭한 팀원들과 함께한 천방지축 어리둥절 빙글빙글 돌아가는 2박3일 이였지만, 수많은 고난을 함께 이겨내가며 진정한 전우애를 느낄 수 있었습니다.