

**SK네트웍스 Family AI 과정 14기  
 데이터 전처리 학습된 인공지능 모델**



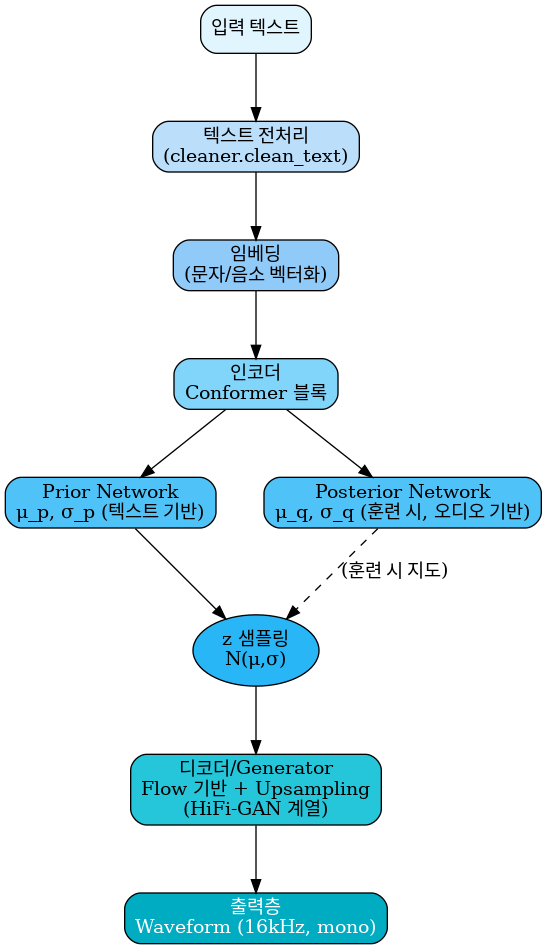
| **산출물 단계** | 데이터 전처리 |
| --- | --- |
| **평가 산출물** | 학습된 인공지능 모델 |
| **제출 일자** | 2025.08.29 |
| **깃허브 경로** | [4팀\_깃허브](https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN14-FINAL-4Team) |
| **작성 팀원** | 공지환 |

1. **모델 목적** : 특정 인플루언서의 음성을 학습하여 프로젝트의 한국어 TTS 모델 구축
2. **모델 아키텍처 설계**

* 선정 모델 : MeloTTS
* 아키텍처 개요 :

| **계층** | **구성 요소** | **역할** |
| --- | --- | --- |
| 입력층 | cleaner.clean\_text() | 텍스트 전처리 |
| 인코더 | Variational Inference + Attention | 텍스트 특징 인코딩 |
| 디코더 | Flow-based Decoder + GAN | 오디오 파형 합성 |
| 출력층 | Waveform | 24kHz 음성 출력 |

* 아키텍처 시각화 : 입력 텍스트 → 임베딩 → 인코더블록(Conformer) → Variational Latent Space → 출력 오디오



* 모델 사양 요구 사항 :  
  + 프레임워크 : PyTorch 2.8.0
  + GPU/CPU 호환 여부 : CUDA 지원(권장), CPU도 가능
* 환경 설정 파일 : requirements.txt, config\_finetune\_my.json
* 설계 근거
  + MeloTTS는 vits2는 end-to-end 음성 합성에 강점을 가짐. -> 다른 모델에 비해 자연스러움.
  + 한국어 전용 전처리를 지원함. -> TTS 시 안정성 확보

1. **모델 학습 요약**

* 학습 데이터 수: 41건
* 평가 데이터 수: 4건
* 성능 평가 결과: 학습 실패
  + 실패 이유:
    - 모델 train시 순환참조 문제 해결 실패
    - 마지막 에러:
      * MemoryError
      * 메모리 에러 임의 해결 시 RuntimeError 발생
      * Checkpoint 로딩 코드의 무한 재귀가 문제로 추정

1. **향후 계획**

* 다른 모델 탐색 및 학습 재시도
* conqui-ai/TTS
* Spark-TTS
* MeloTTS 학습실패 원인 파악 후 재시도