

[ 실증적 SW 개발 프로젝트 ]

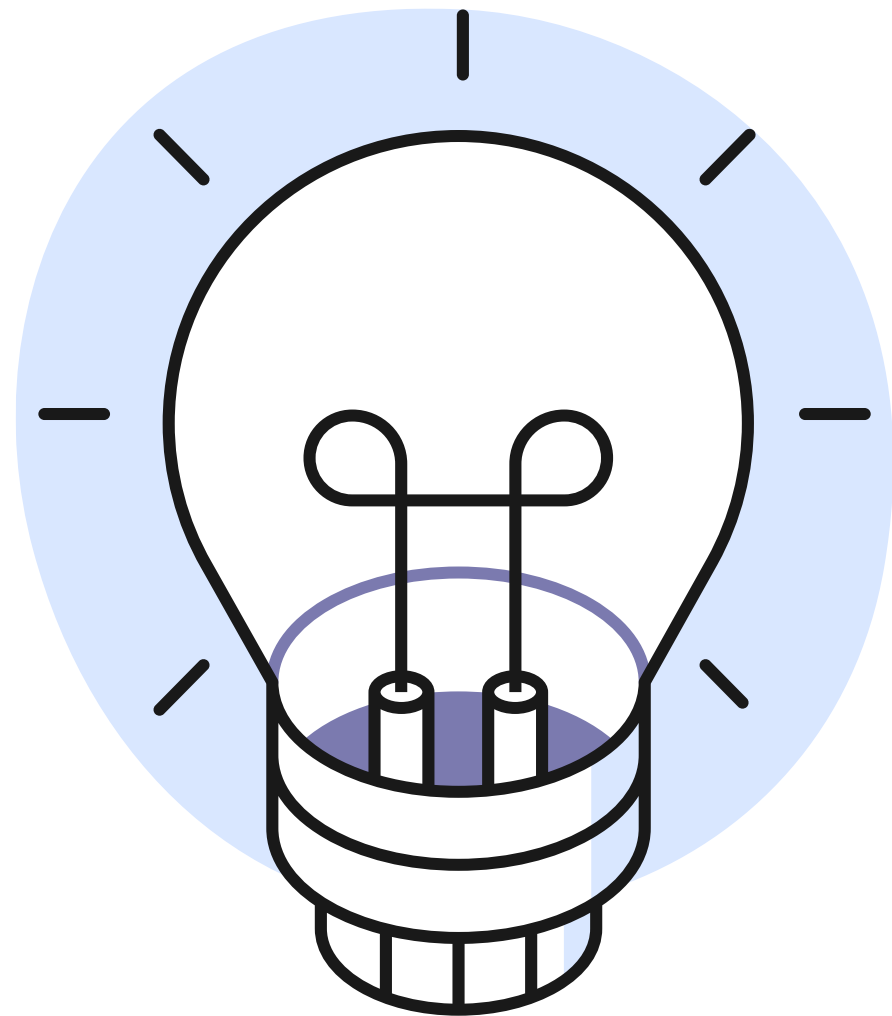
# 디지털화된 오디오를 네트워크로 전송시 에러율 확인 소프트웨어 ◦

D e t e c t i o n   E r r o r   I n  
C o m m u n i c a t i o n   S y s t e m

컴퓨터공학과

1924351 정성운 1924362 박유진  
1923846 강태원 1934285 전민재

# [ 서론 ]



## 배경 소개 및 문제 정의

네트워크 통신 환경에서는 **효율적인 통신**이 중요한데,  
배경 소음과 같은 노이즈로 인해 통신 품질이 떨어지면  
**정확한 정보 전달**에 어려움이 있음

## 문제의 중요성 및 필요성

정확하고 안정적인 통신 시스템이 이루어지지 않으면 **생명과 안전**이 위협받을 수 있으며, 통신 오류는 곧 **심각한 사고**로 이어질 수 있음

# [ 서론 ]

“

에러율 검출은  
안전과 신뢰성을  
위한 핵심 요소



## 네트워크 이상 감지

네트워크 내에서 발생하는 노이즈 및 패킷 이슈 등과 같은 이상 상황을 감지함



## 문제 분석 및 에러율 계산

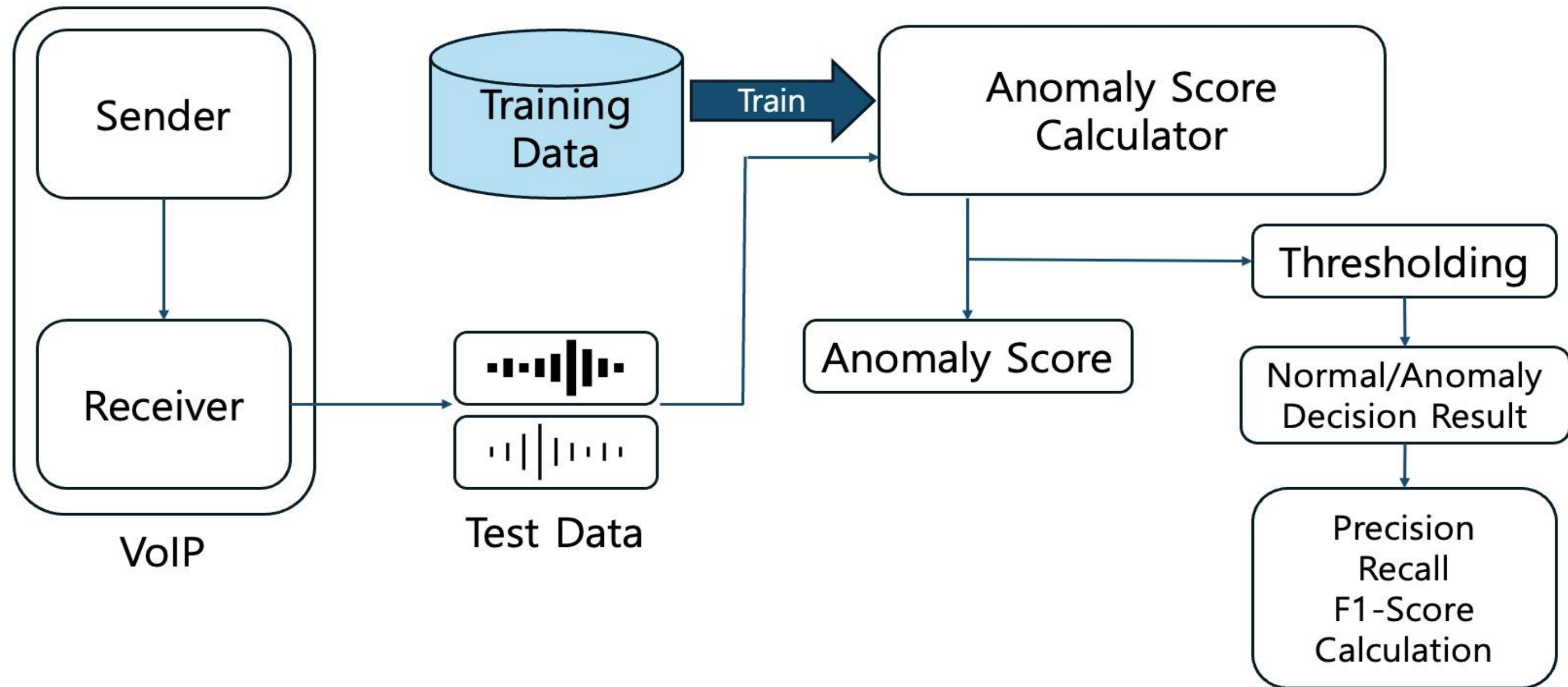
감지된 이상 데이터에 대한 분석을 통해 문제의 원인과 심각도를 파악하고 에러율을 계산함



## 사전 조치 및 예방

계산된 에러율을 기반으로 시스템을 조정하거나 예방 조치를 시행하여 네트워크의 안정성과 신뢰성을 높임

## [ 시스템 구조 ]



# [ 프로토콜 ]

송신자

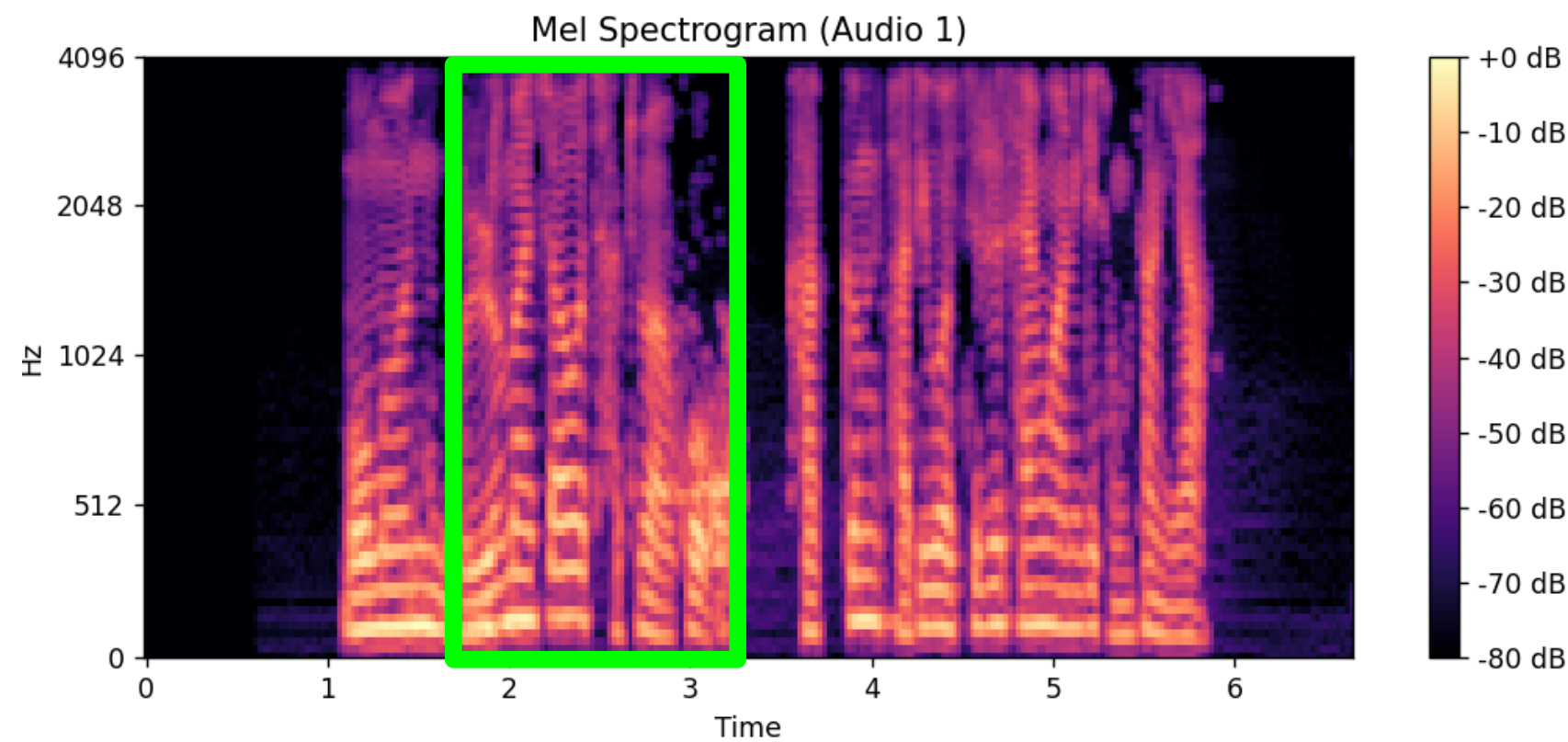


수신자

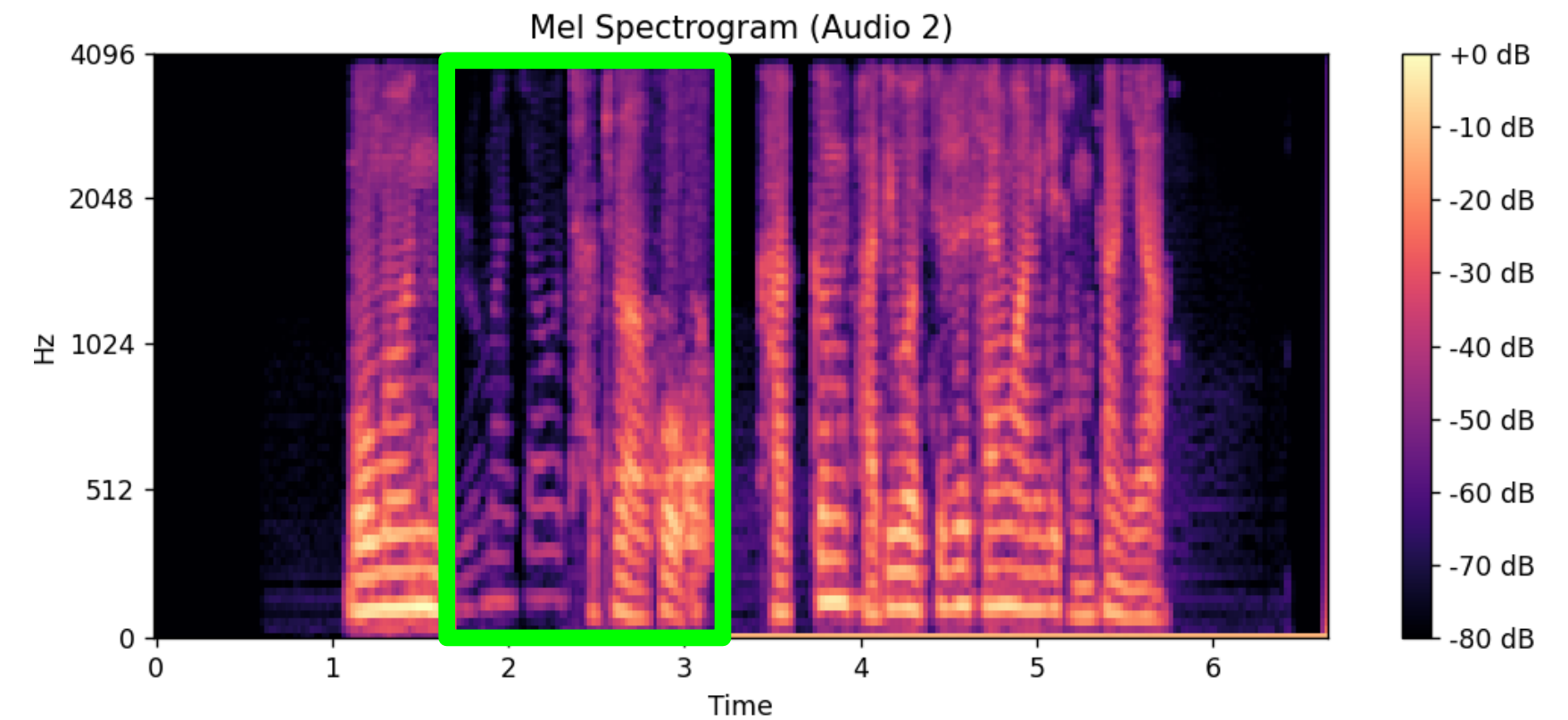


# [ 스펙트로그램 ]

## Normal Data



## Abnormal Data

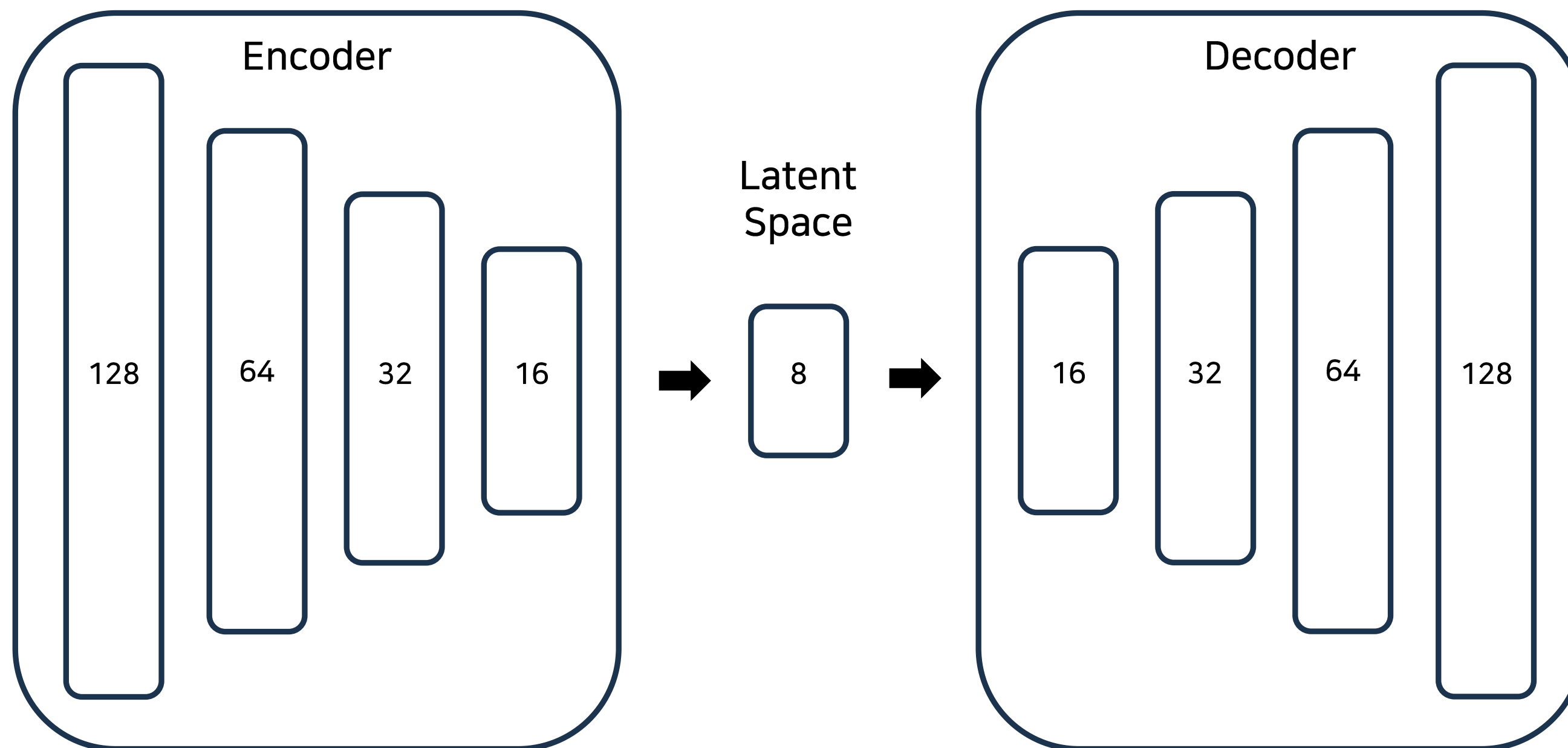


- 음성 인터넷 프로토콜 환경에서 패킷 이슈를 인위적으로 발생시킴
- 이러한 과정을 반복해 테스트를 위한 정상 및 비정상 음성 샘플을 확보함

# [ 데이터 수집 ]

데이터 종류	이상 여부	데이터 특성	비고
러닝 데이터	정상	- 오토인코더 모델의 학습을 위해 확보 - 노이즈가 없는 정상 음성 데이터	13,500개
테스트 데이터	정상	- 오토인코더 모델의 테스트를 위해 확보 및 생성 - 노이즈가 없는 정상 음성 데이터	481개
	비정상	- 오토인코더 모델의 테스트를 위해 확보 및 생성 - 인위적으로 노이즈를 삽입한 비정상 음성 데이터	461개

# [ 오토인코더 ]





# [ 오토인코더 ]

	259	260	261	262	263
94	-28.1397	-41.0995	-42.4378	-38.7723	-49.0917
95	-30.5573	-42.4889	-53.1547	-51.1967	-52.1647
96	-46.8496	-50.1426	-53.3506	-50.7871	-52.1269
97	-43.7436	-48.2642	-51.5776	-50.4783	-51.8175
98	-49.9245	-50.1166	-56.1945	-56.645	-55.4987

정상 데이터 텐서

	259	260	261	262	263
94	-45.2891	-47.1479	-53.6491	-51.7262	-53.5158
95	-76.2029	-71.9809	-74.1143	-76.9652	-71.432
96	-85.2256	-81.5356	-84.6857	-88.6854	-91.4685
97	-106.8742	-108.6237	-110.9142	-107.9351	-102.9452
98	-14.8582	-16.7877	-18.4515	-19.7964	-20.9194

비정상 데이터 텐서

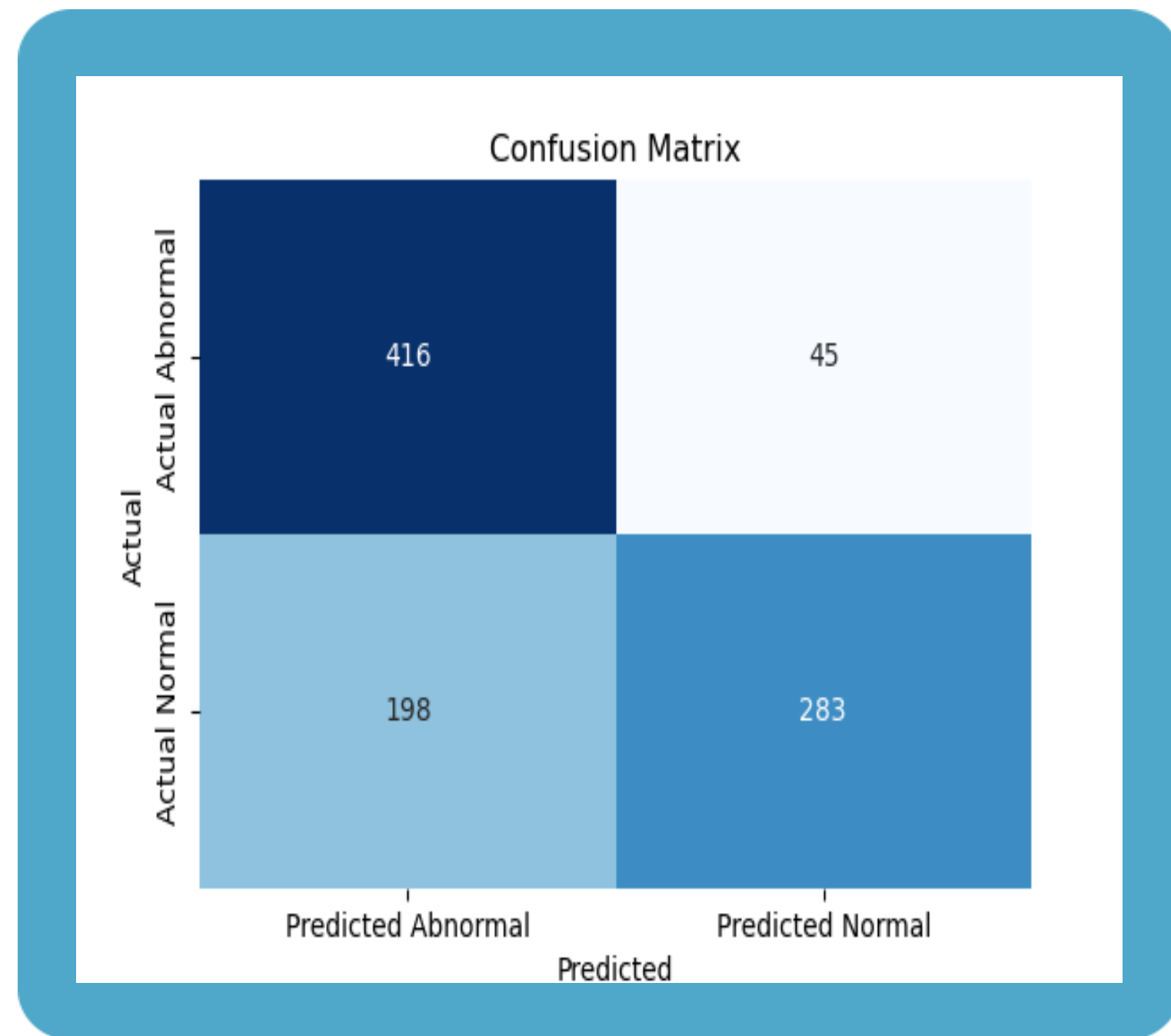
	259	260	261	262	263
94	-29.4682	-43.5187	-44.9348	-40.1571	-52.2297
95	-34.1957	-44.5652	-55.3247	-53.6478	-56.6824
96	-58.0675	-59.1742	-61.9803	-60.9834	-57.0217
97	-59.8095	-60.8019	-62.1974	-61.7112	-58.8848
98	-52.0164	-54.5057	-59.377	-60.2851	-58.4032

비정상 데이터를 재구성한 텐서

## 오토인코더 입출력 데이터 텐서

- 입력 데이터의 샘플 수는 97개이고, 각 샘플의 특징 벡터 차원은 320임
- 패킷 이슈가 포함된 입력 데이터를 재구성하며, 일부 이슈는 완벽히 복원하진 못함
- 출력 데이터와의 재구성 오차를 통해 에러율을 측정하여 이상 여부를 판별함

# [ 초기 모델 ]



## 01 정밀도

비정상으로 예측한 샘플 중 실제로 비정상인 샘플 비율

$$Precision = \frac{TP}{TP + FP} = \frac{416}{416 + 198} \approx 0.678$$

## 02 재현율

실제 비정상인 샘플 중 모델이 올바르게 예측한 비율

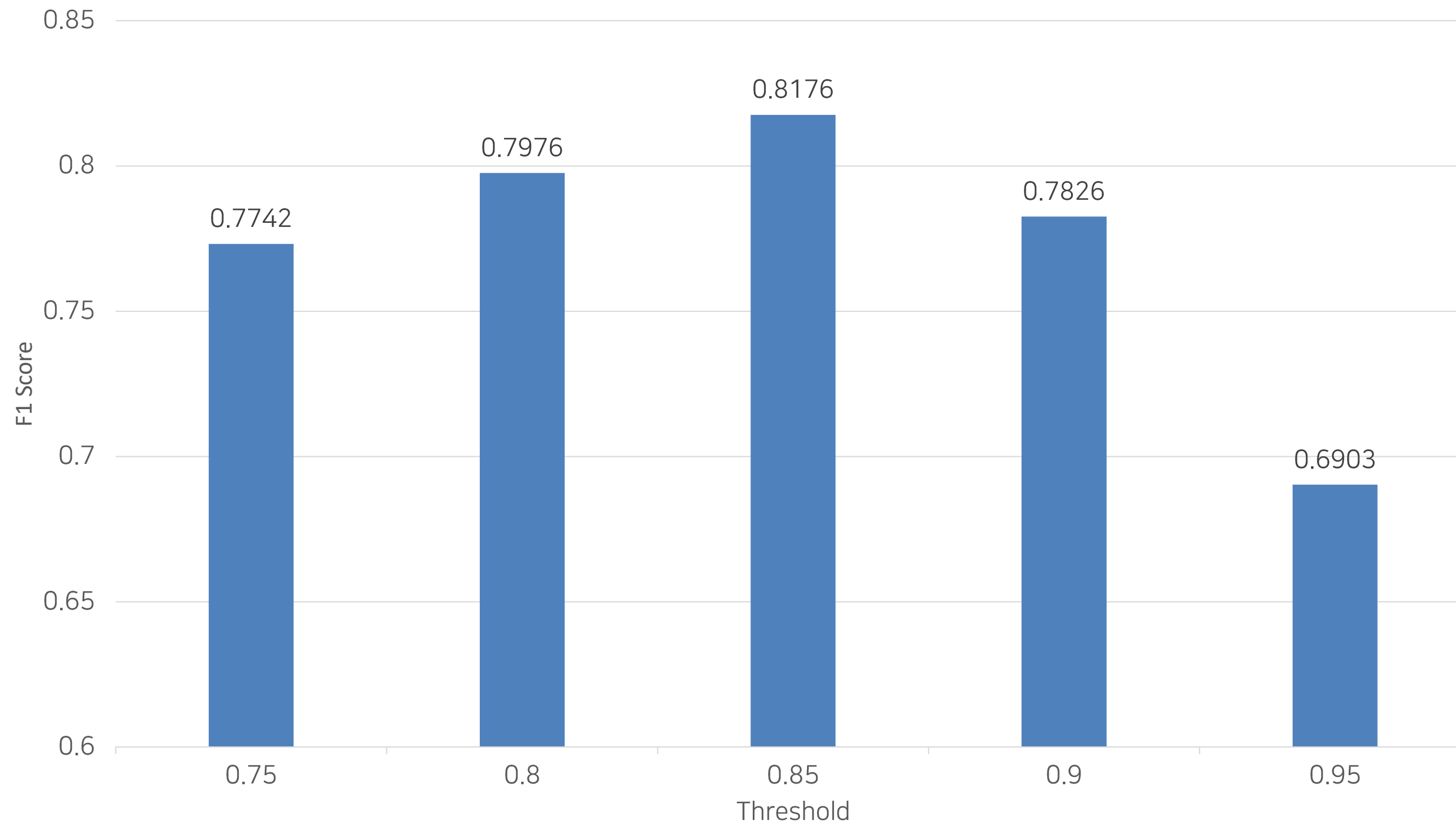
$$Recall = \frac{TP}{TP + FN} = \frac{416}{416 + 45} \approx 0.903$$

## 03 F1 Score

모델의 예측 성능을 종합적으로 평가하는 지표

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.678 \times 0.903}{0.678 + 0.903} \approx 0.774$$

# [ 성능 개선 ]



## 그래프 구성

X축: 임계값  
Y축: F1 Score

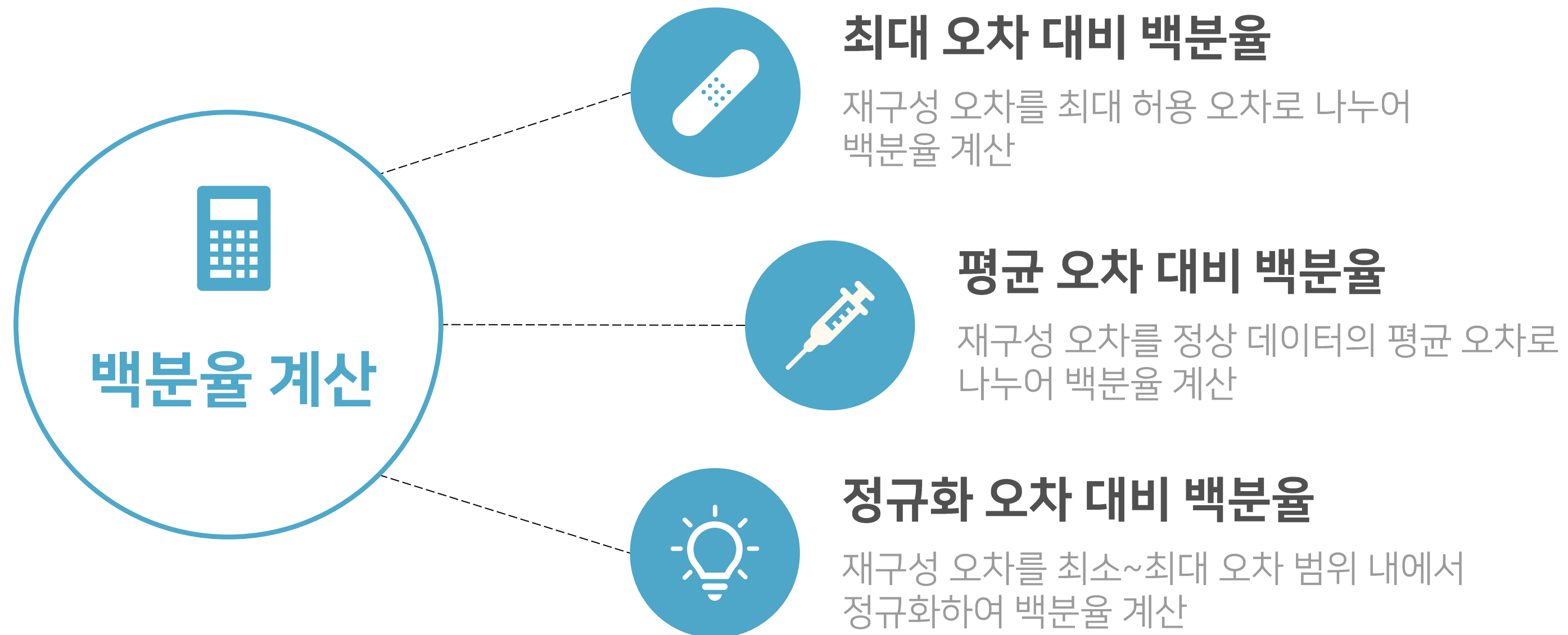


## 그래프 해석

임계값이 0.85일 때,  
가장 높은 F1 Score가 나타남

# [ 에러율 검출 ]

Q. 에러율을 백분율로 표현하려면 ?



# [ 에러율 검출 ]

## 평균 오차 대비 백분율

$$Error\ Rate = \frac{Current\ Reconstruction\ Error}{Average\ Normal\ Error} \times 100 = \frac{0.2}{0.05} \times 100 = 400\%$$

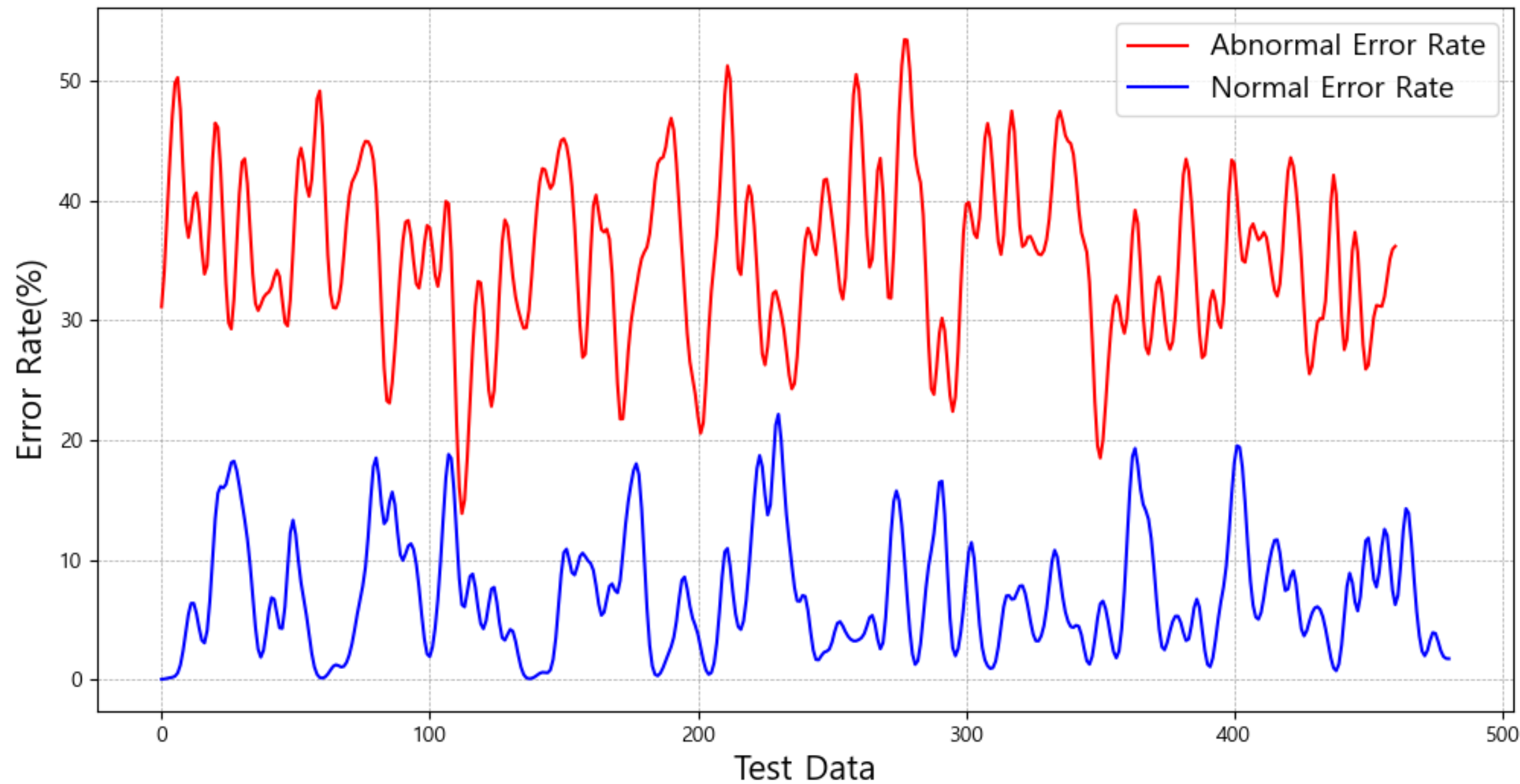


오차율을 상대적으로 표현하여 원본과의 차이를 명확하게 드러내기 위해 수정함

$$Error\ Rate = \left(1 - \frac{Average\ Normal\ Error}{Current\ Reconstruction\ Error}\right) \times 100 = \left(1 - \frac{0.05}{0.2}\right) \times 100 = 75\%$$

# [ 에러율 검출 ]

Test Data Error Rate Distribution



## 그래프 구성

X축: 테스트 데이터

Y축: 에러율

빨간선: 비정상 에러율

파란선: 정상 에러율

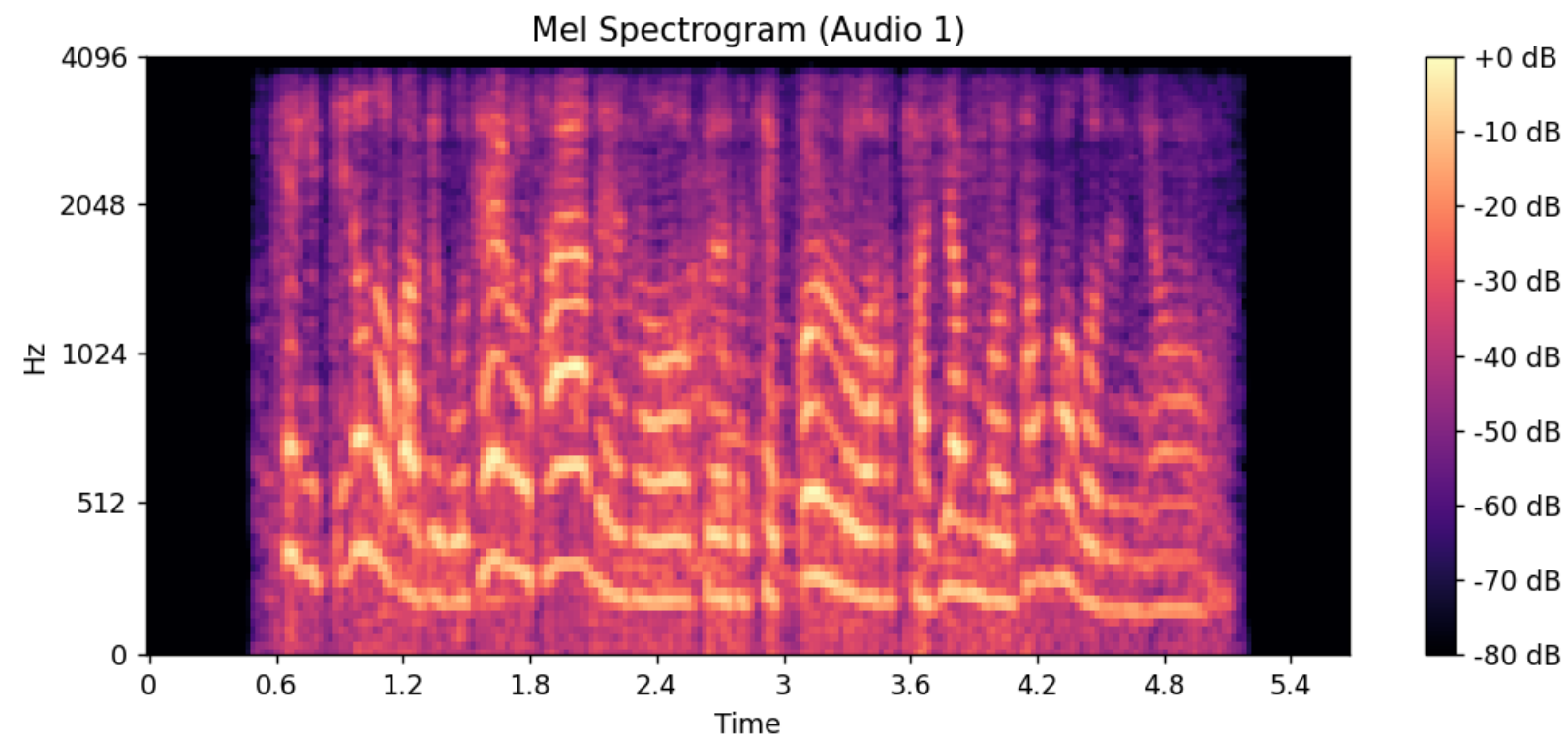


## 그래프 해석

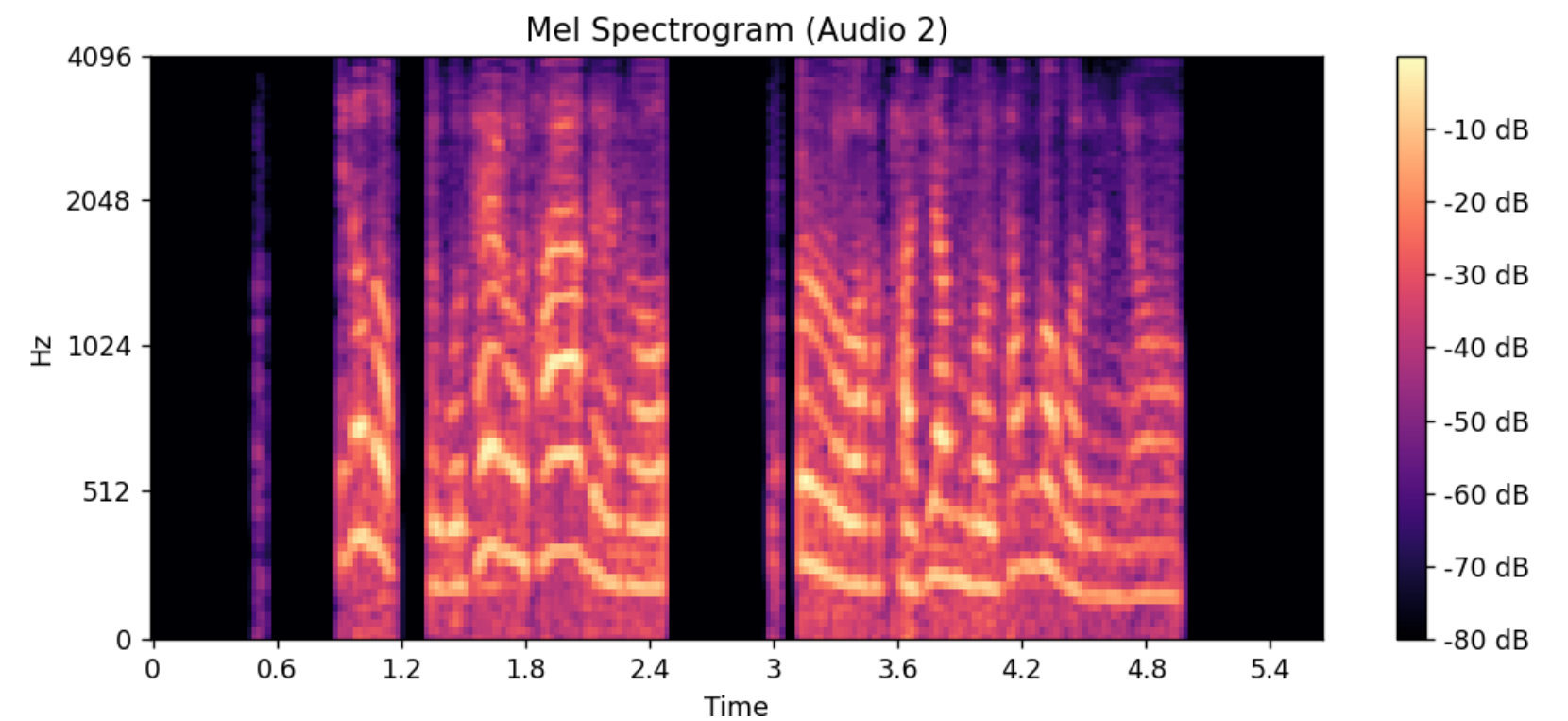
정상 에러율 범위: 0 ~ 20%

비정상 에러율 범위: 20 ~ 50%

# [ 구현 결과 ]



Normal.wav Reconstruction Error = 39.3998



Abnormal.wav Reconstruction Error = 66.2315

Average Normal Error: 39.1084

Current Reconstruction Error: 66.2315



$$Error\ Rate = \left( 1 - \frac{39.1084}{66.2315} \right) \times 100 = \mathbf{40.9519\%}$$

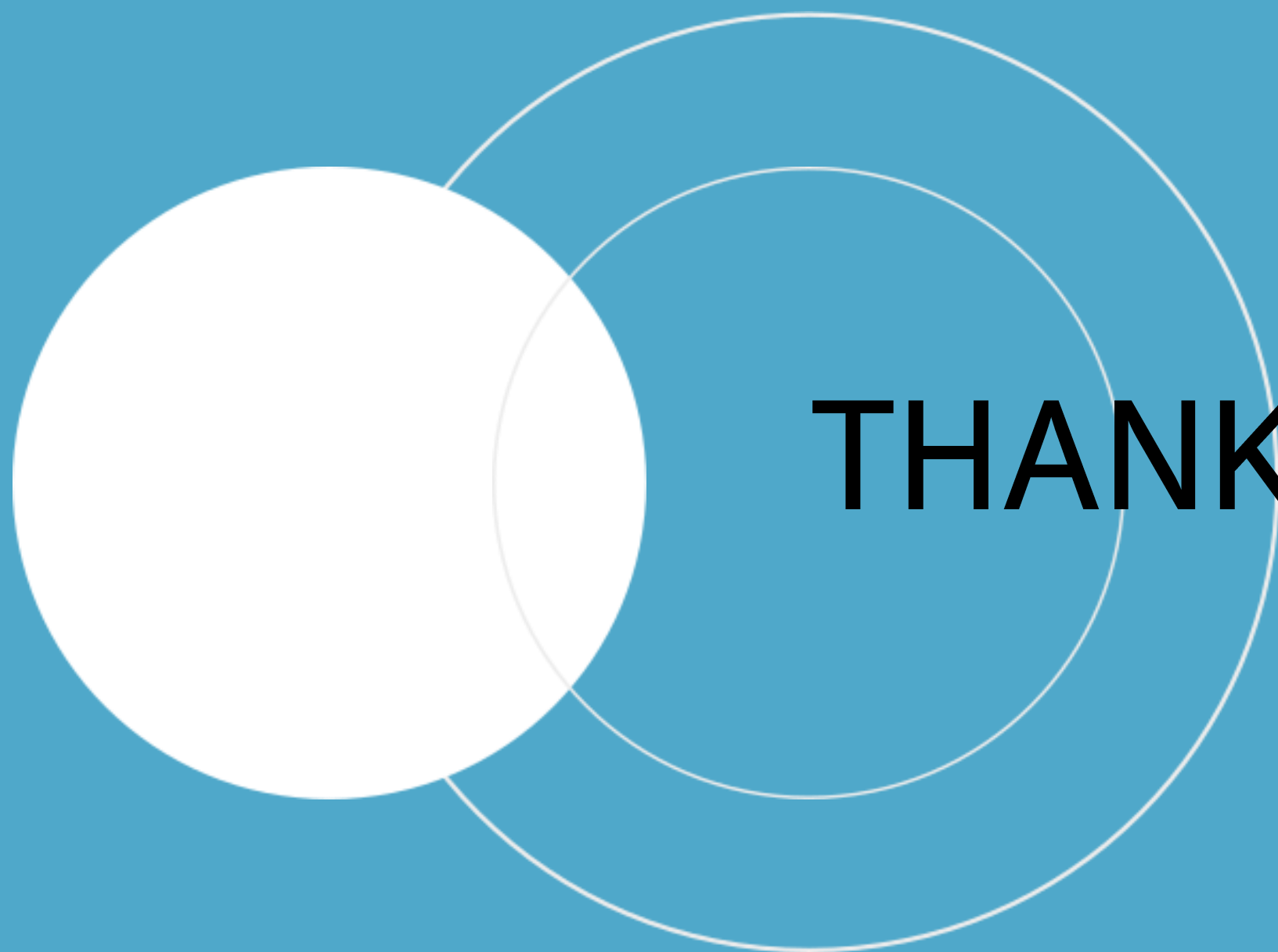
## [ 결론 ]

오토인코더 모델을 사용하여 정상 데이터는 높은 정확도로 재구성하는 데 성공,  
비정상 데이터의 경우 재구성 오차를 통해 차이를 측정할 수 있음을 확인함.

예시 비정상 데이터에 대해 40.95%의 에러율을 나타내어 비정상 데이터를 일정 부분  
탐지할 수 있는 가능성을 보여줌.

- ☑ 비정상 데이터 탐지 성능을 더욱 개선하기 위해 모델 구조의 최적화와 데이터셋 확장이 필요
- ☑ 향후, 노이즈 제거 오토인코더를 추가하여 네트워크 안정성과 신뢰성을 확보할 계획





**THANK YOU!**

