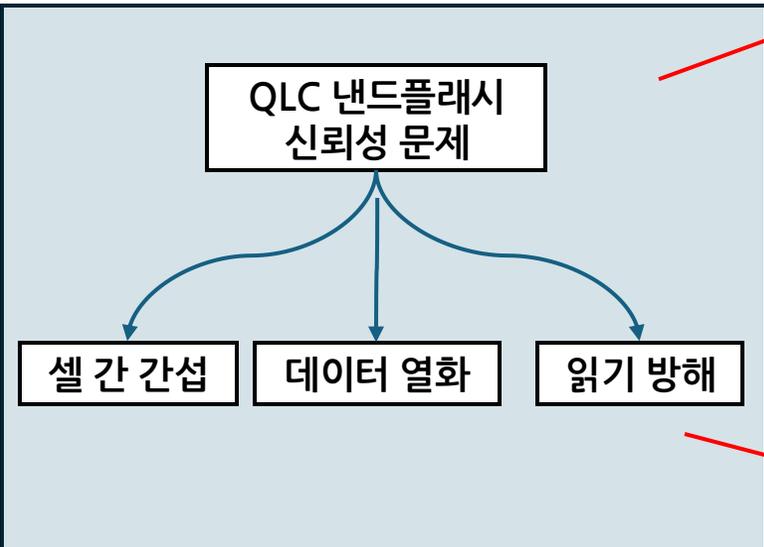


랜덤화 기법 및 해석 가능한 인공지능을 활용한 QLC 낸드플래시 데이터 신뢰성 개선

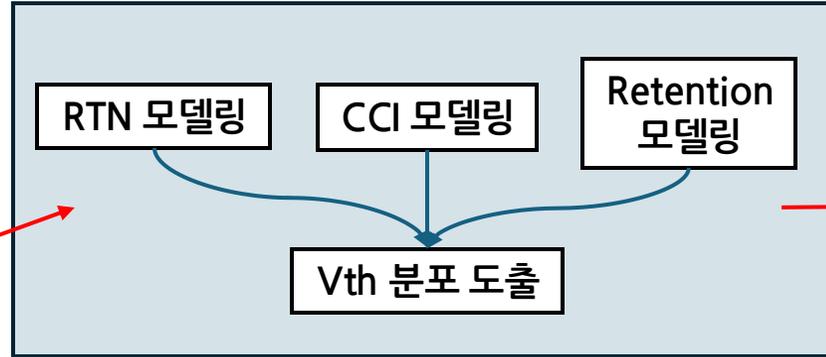


- 선택 문제 : 1번
- 팀 명 : 믿음을주세요
- 팀원 : 변지환(조장), 김세훈, 김영진, 나희재

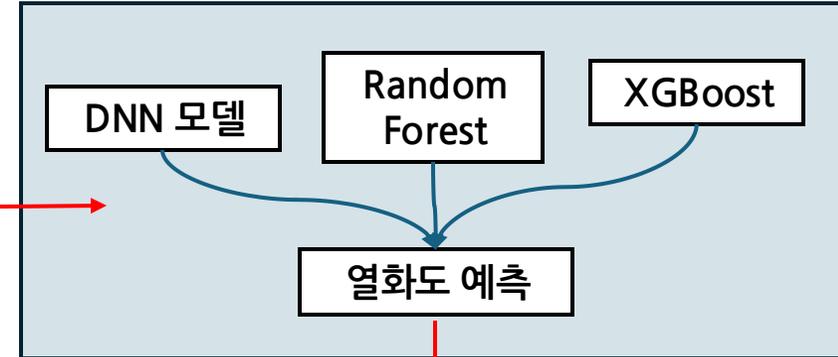
Executive summary



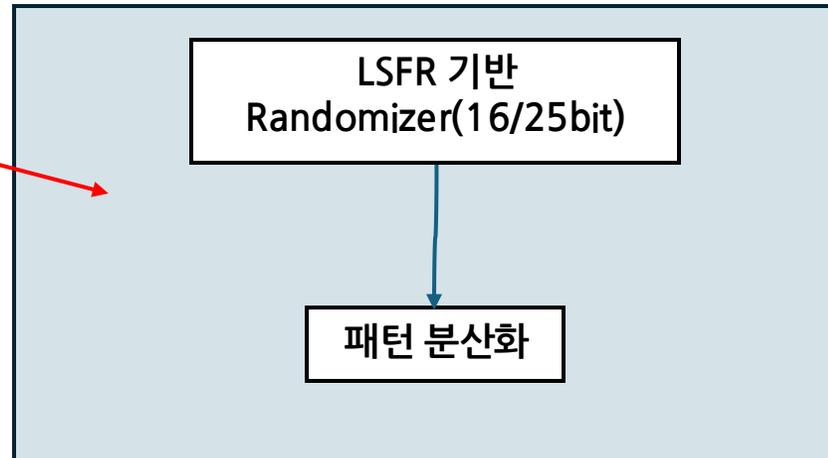
Task 2.



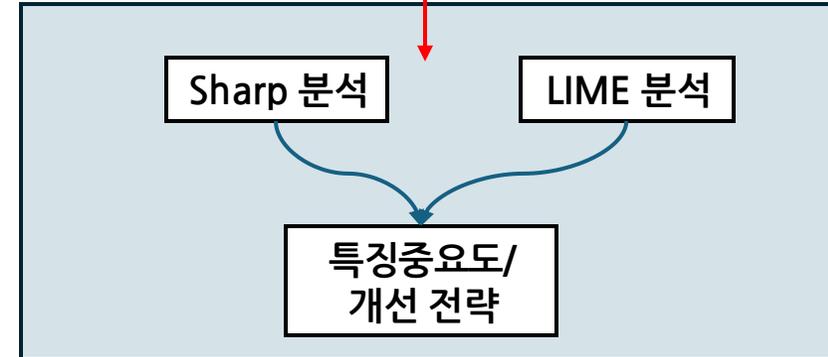
Task 3.



Task 1.



Task 3.



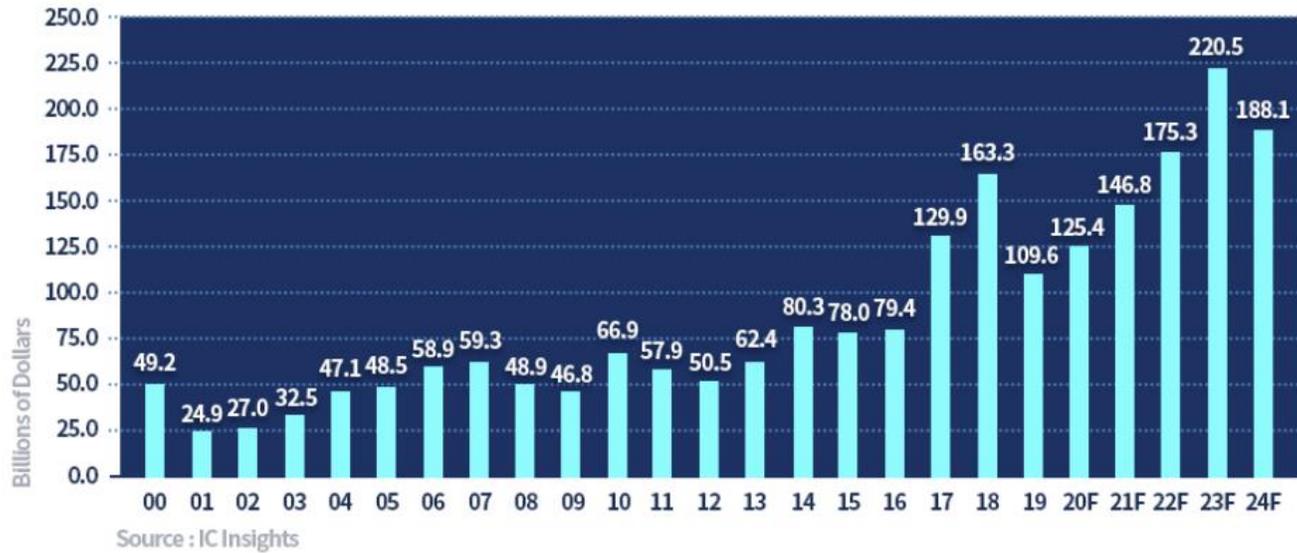
데이터 패턴 의존성 해결

열화도 해석가능에 따른 개선 전략 제시

연구 배경 설명 - 메모리 반도체 시장 현황



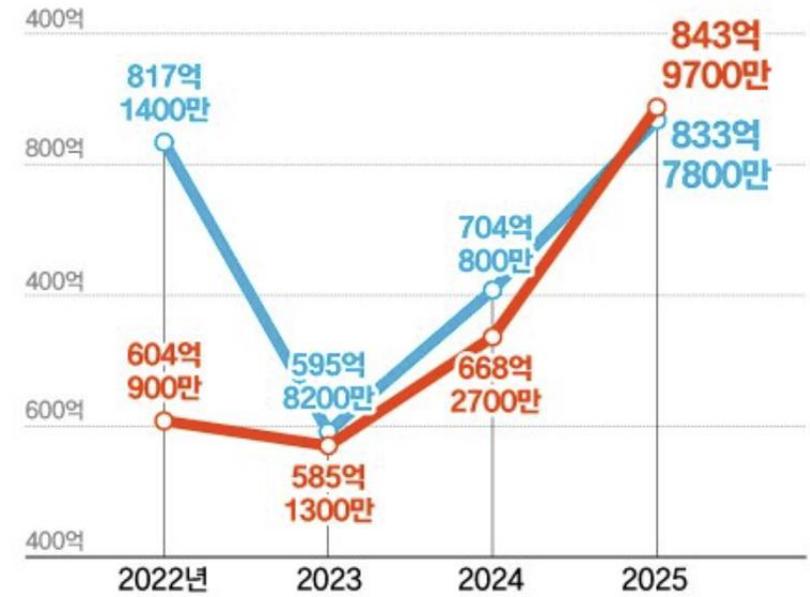
TOTAL MEMORY IC MARKET (\$B)



D램보다더 커질 낸드 시장규모

단위: 달러

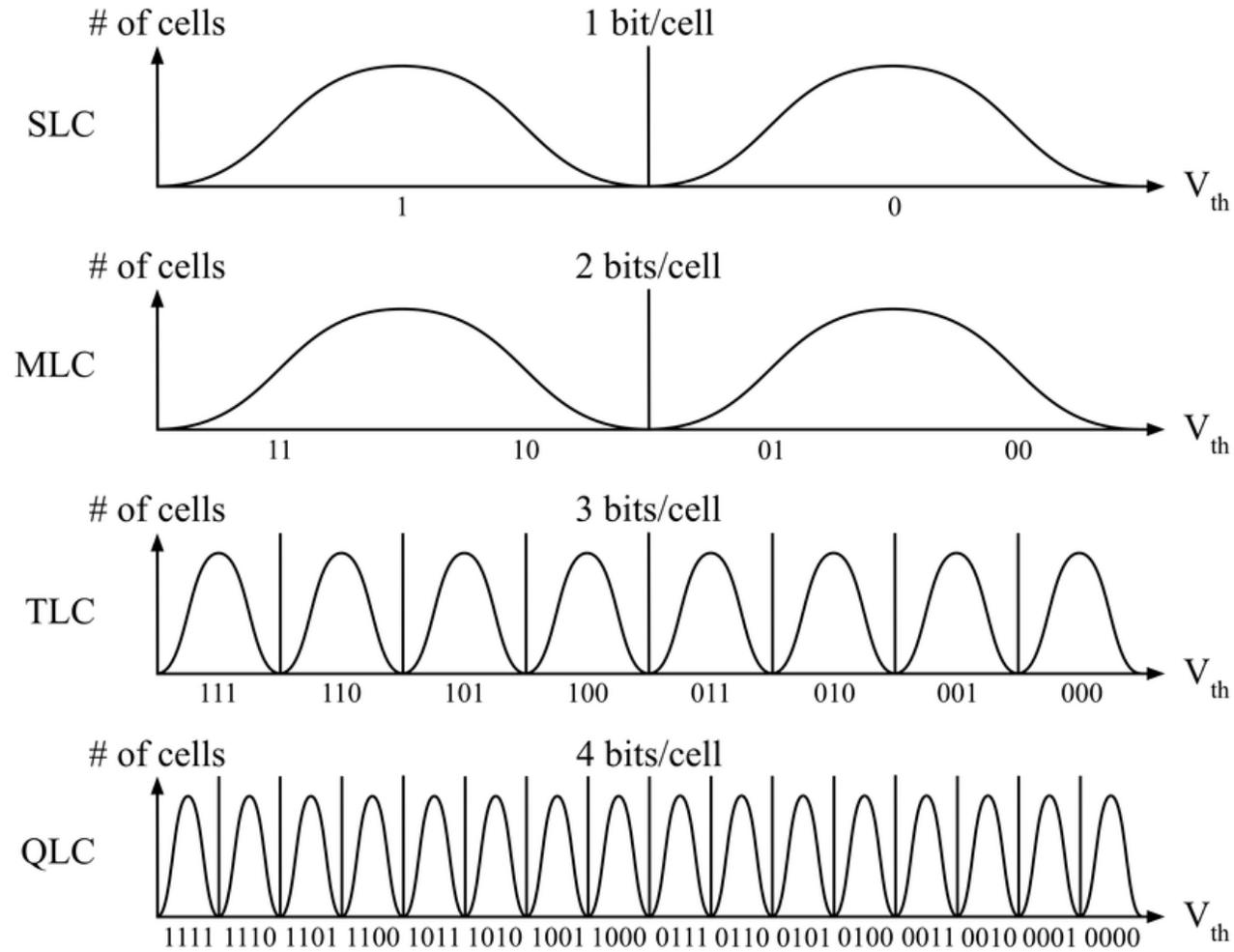
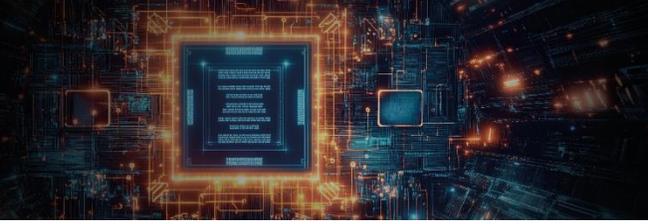
— D램 — 낸드 플래시



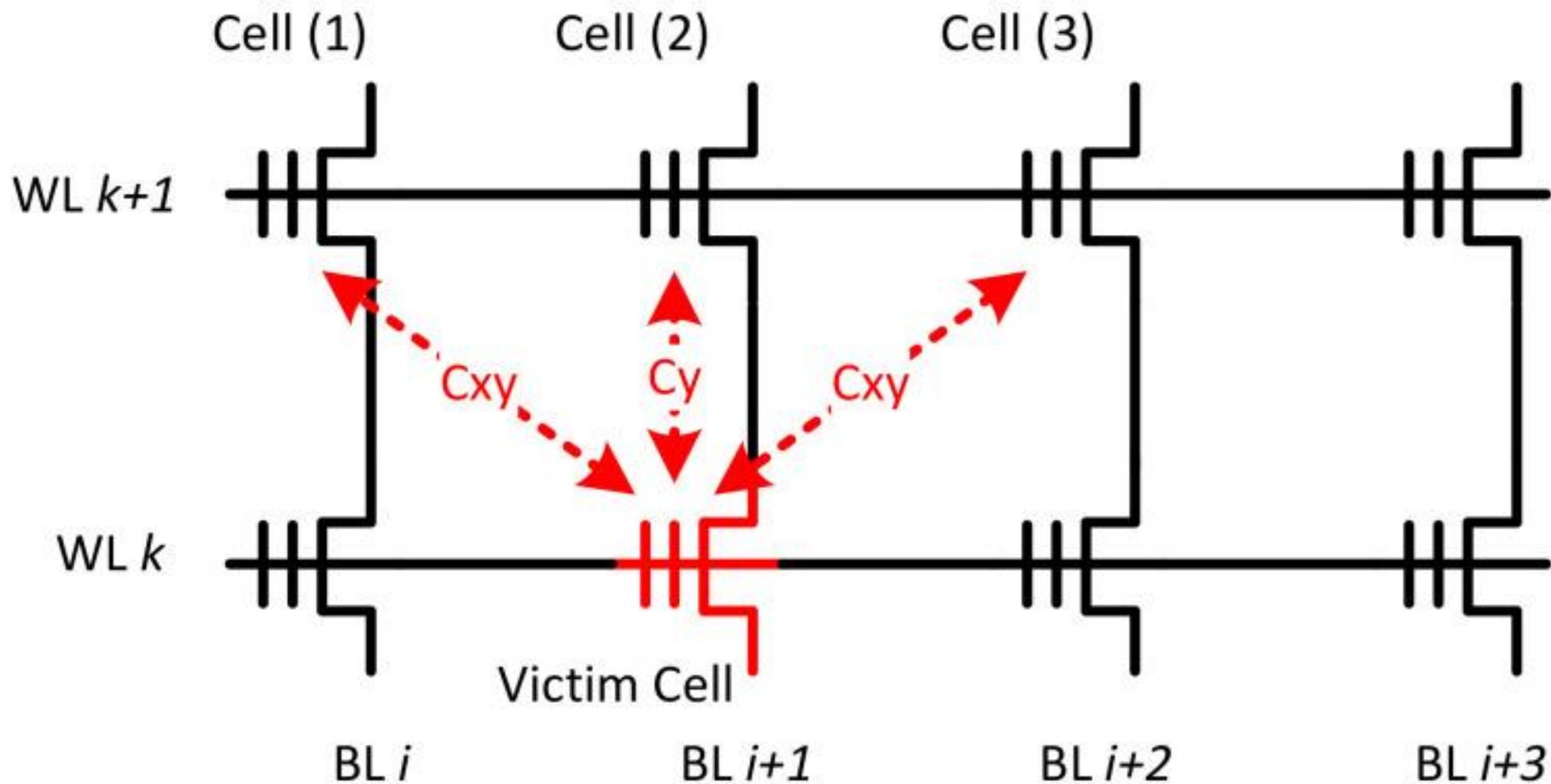
자료: 옴디아

The JoongAng

연구 배경 설명 - QLC 낸드플래시의 발전



연구 배경 설명 - 데이터 신뢰성 문제





데이터 열화 요인:

1. Data retention

2. P/E 사이클 증가

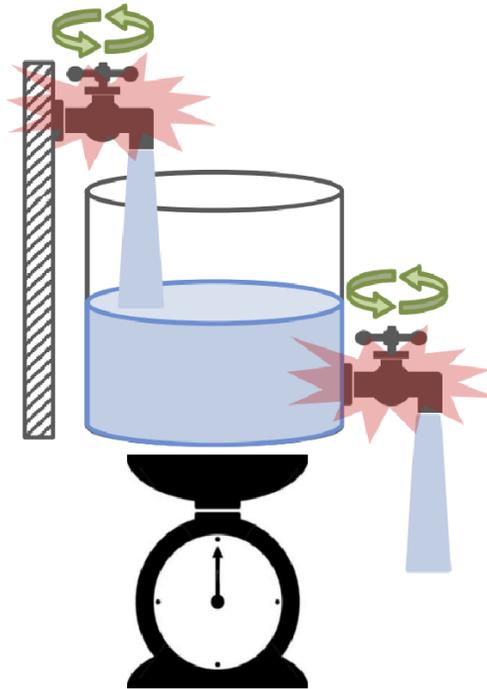
3. Read disturb

연구 배경 설명 - 데이터 열화 요인



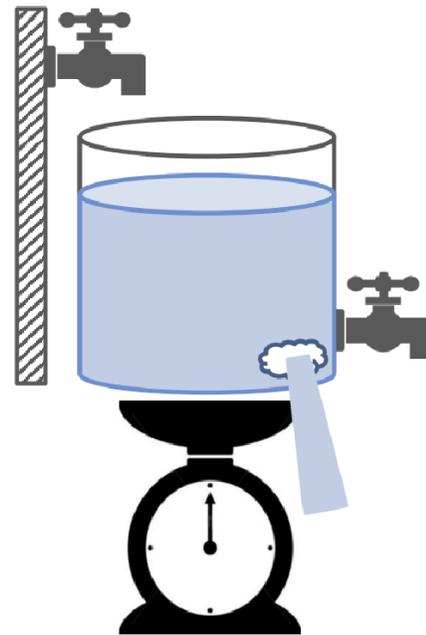
Endurance

A limited number of program/erase operations



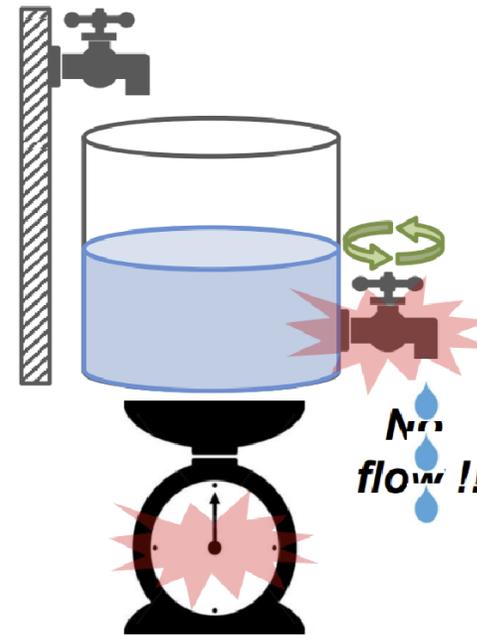
Retention

Data loss by leakage



Read Disturb

A limited number of read operations



목차

Task 1. 랜덤화 기법

Task 2. 열화 현상 모델링

Task 3. 해석가능 AI 프레임워크

목차

Task 1. 랜덤화 기법

Task 2. 열화 현상 모델링

Task 3. 해석가능 AI 프레임워크

Task 1. 랜덤화 기법



Task 1. 어떤 데이터 패턴에서 열화?



특정 데이터 패턴

Ex)

000000000000

111111111111

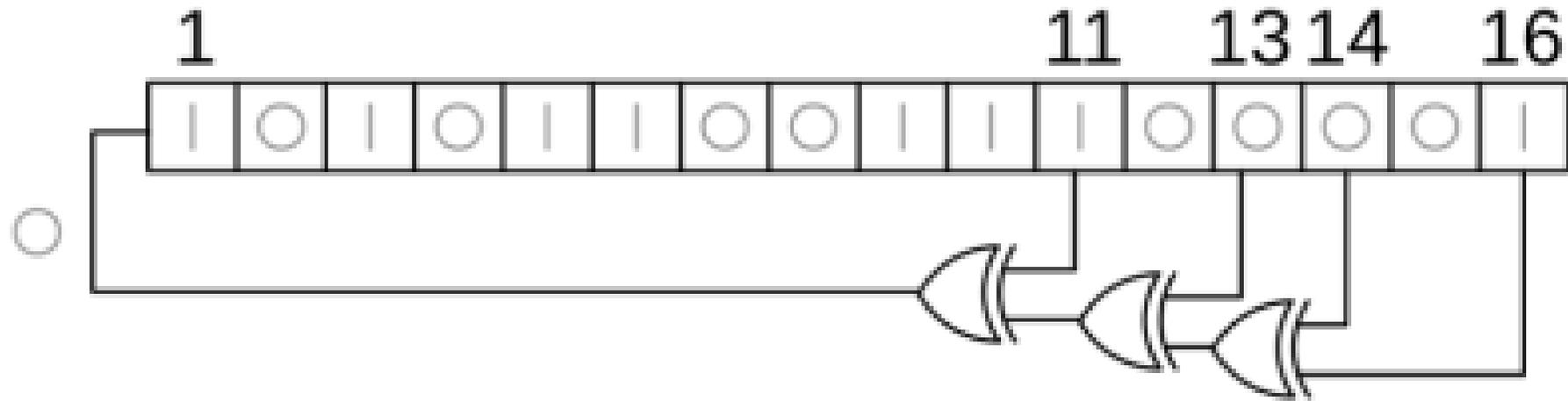
특정 부분의 셀 마모가 빨라지고 전압 변동이 커질 수 있다.

=> 데이터를 랜덤화하여 패턴을 없애보자!

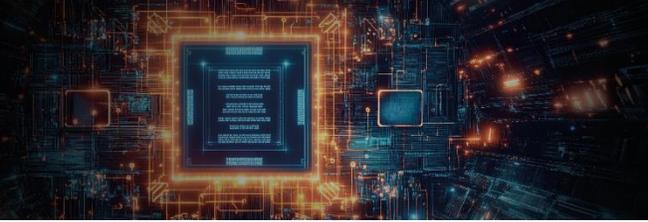
Task 1. 랜덤화 기법



16bit LFSR 구조



Task 1. 랜덤화 기법



항목	16bits LFSR	24bits LFSR
다항식	$x^{16} + x^{14} + x^{13} + x^{11} + 1$	$x^{24} + x^{22} + x^{21} + x^{19} + 1$
시드 값 크기	16비트 (0x0000 ~ 0xFFFF)	24비트 (0x000000 ~ 0xFFFFFFFF)
랜덤화 강도	충돌 가능성 높음, 보안 수준 낮음	충돌 가능성 낮음, 보안 수준 높음
성능	빠른 연산 속도, 적은 메모리 사용	느린 연산 속도, 더 많은 메모리 사용
적용 분야	간단한 랜덤화가 필요한 경우	높은 보안과 무작위성이 필요한 경우

Task 1. 랜덤화 기법



랜덤성이 낮은 데이터의 실험 결과:

구분	원본	16bits LFSR	24bits LFSR
총 비트 수	71,544	71,544	71,544
최대 연속 길이	9	15	14
평균 연속 길이	2.75	2.01	2.01
총 시퀀스 수	25,993	35,678	35,666
0 시퀀스 수	12,997	17,839	17,833
1 시퀀스 수	12,996	17,839	17,833

Task 1. 랜덤화 기법

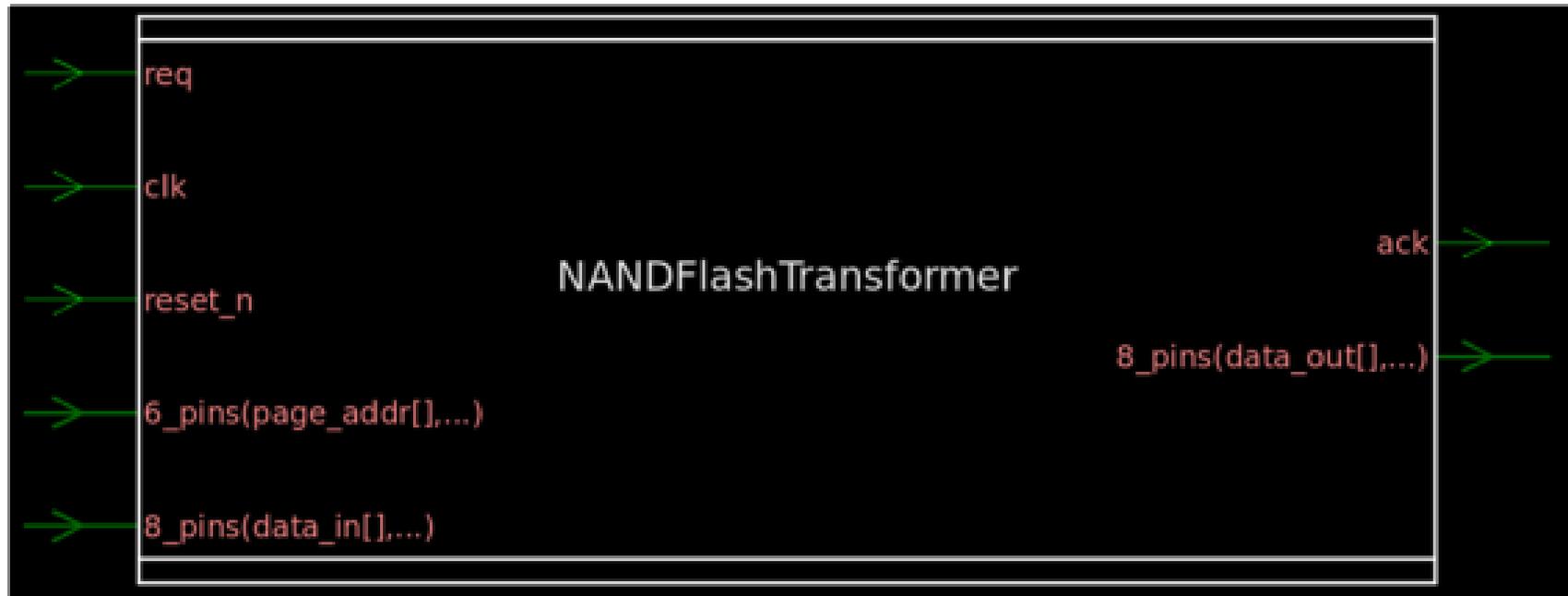


일반적인 영어 문장에 대한 실험 결과:

구분	원본	16bits LFSR	24bits LFSR
총 비트 수	14,800	14,800	14,800
최대 연속 길이	9	12	14
평균 연속 길이	1.96	2.01	2.01
총 시퀀스 수	7,557	7,347	7,346
0 시퀀스 수	3,779	3,673	3,673
1 시퀀스 수	3,778	3,674	3,673

Task 1. 랜덤화 기법

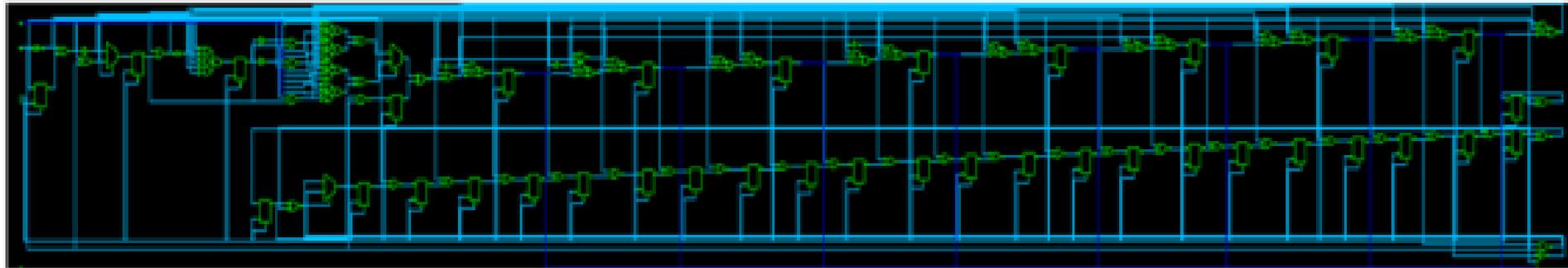
Top모듈 (NANDFlashTransformer)



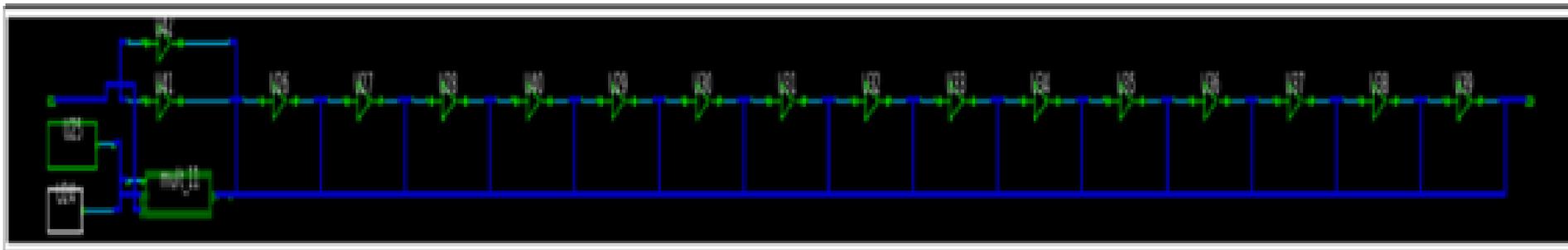
Task 1. 랜덤화 기법



데이터 변환 회로 (DataTransformer)



Seed 기반 ReferenceData 생성 회로 (SeedGenerator)



Task 1. 실 구현 가능성



Verilog를 통한 하드웨어 시뮬레이션
- 약 2938.85 μm^2 의 면적
- 약 0.90 ns의 Critical Path Delay

매우 높은 속도로도 충분히 동작 가능한 구조
실제 구현 시에도 큰 오버헤드 없이 수행 가능

목차

Task 1. 랜덤화 기법

Task 2. 열화 현상 모델링

Task 3. 해석가능 AI 프레임워크

목차

Task 1. 랜덤화 기법

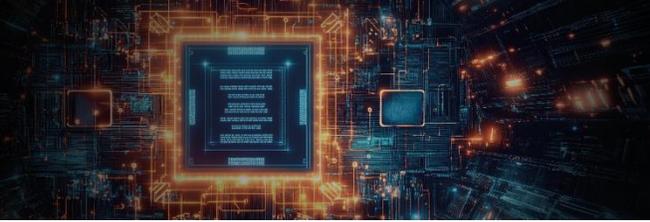
Task 2. 열화 현상 모델링

Task 3. 해석가능 AI 프레임워크

Task 2. 열화 현상 모델링



Task 2. 열화 현상 모델링



RTN(Random Telegraph Noise)

He,R.;Hu,H.;Xiong,C.; Han, G. Artificial Neural Network Assisted Error Correction for MLC NAND Flash Memory. *Micromachines* 2021, 12, 879.

$$p_r(x) = \frac{1}{\sigma_r \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_r^2}}, \quad \sigma_r = 0.00027 \times \text{PE}^{0.62}$$

CCI(Cell-to-Cell Interference)

$$\Delta V_{\text{victim}} = \sum_n (\Delta V_t^{(n)} \cdot \gamma^{(n)}) \quad \gamma_y = 0.08s, \gamma_{xy} = 0.006s$$

Retention Noise

$$p_t(x) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(x-\mu_t)^2}{2\sigma_t^2}} \quad \mu_t = \Delta V_t [A_t(\text{PE})^{\alpha_i} + B_t(\text{PE})^{\alpha_o}] \log(1 + T), \quad \sigma_t = 0.3|\mu_t|$$

Task 2. 열화 현상 모델링

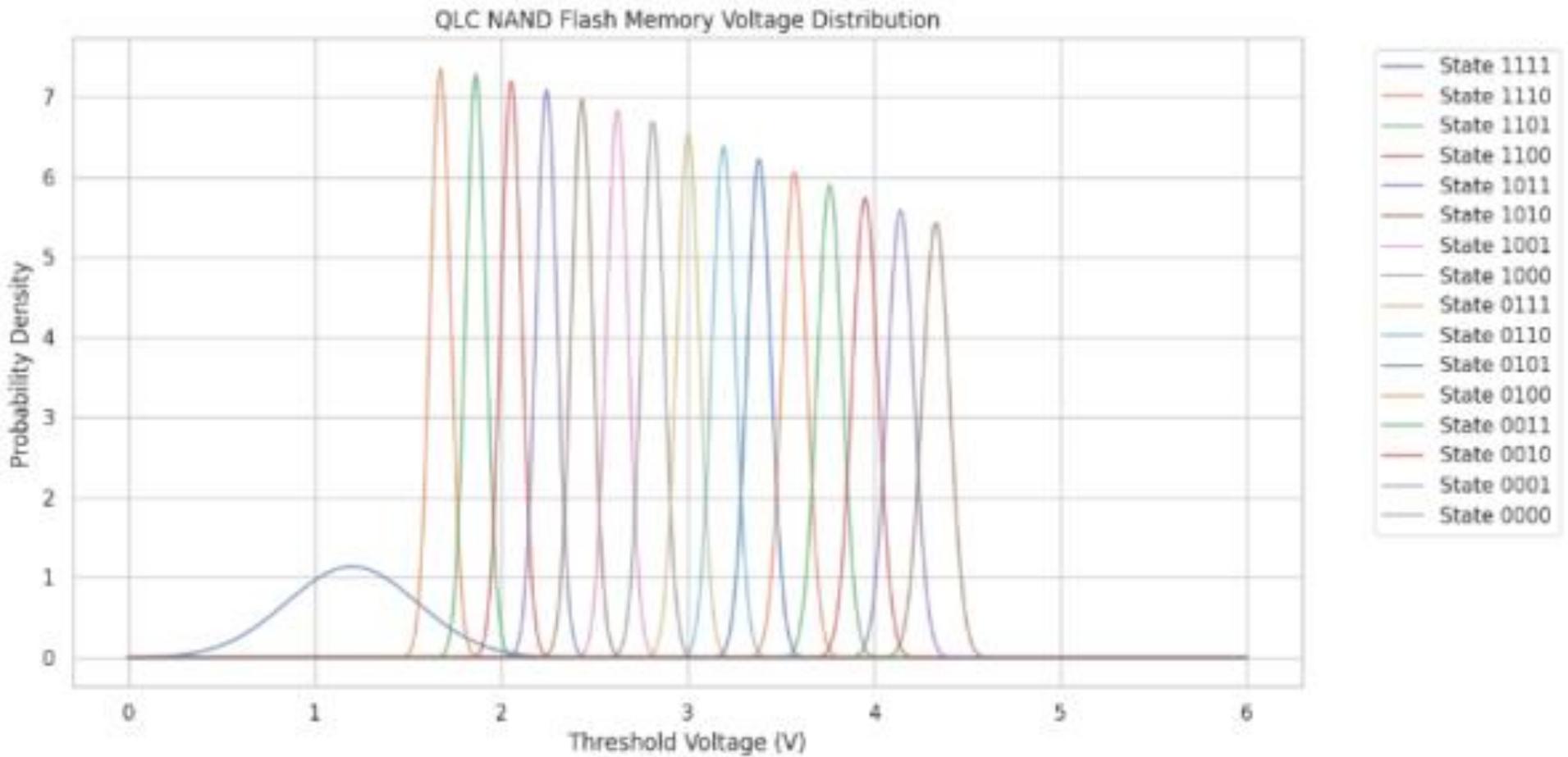


He,R.;Hu,H.;Xiong,C.; Han, G. Artificial Neural Network Assisted Error Correction for MLC NAND Flash Memory. *Micromachines* 2021, 12, 879.

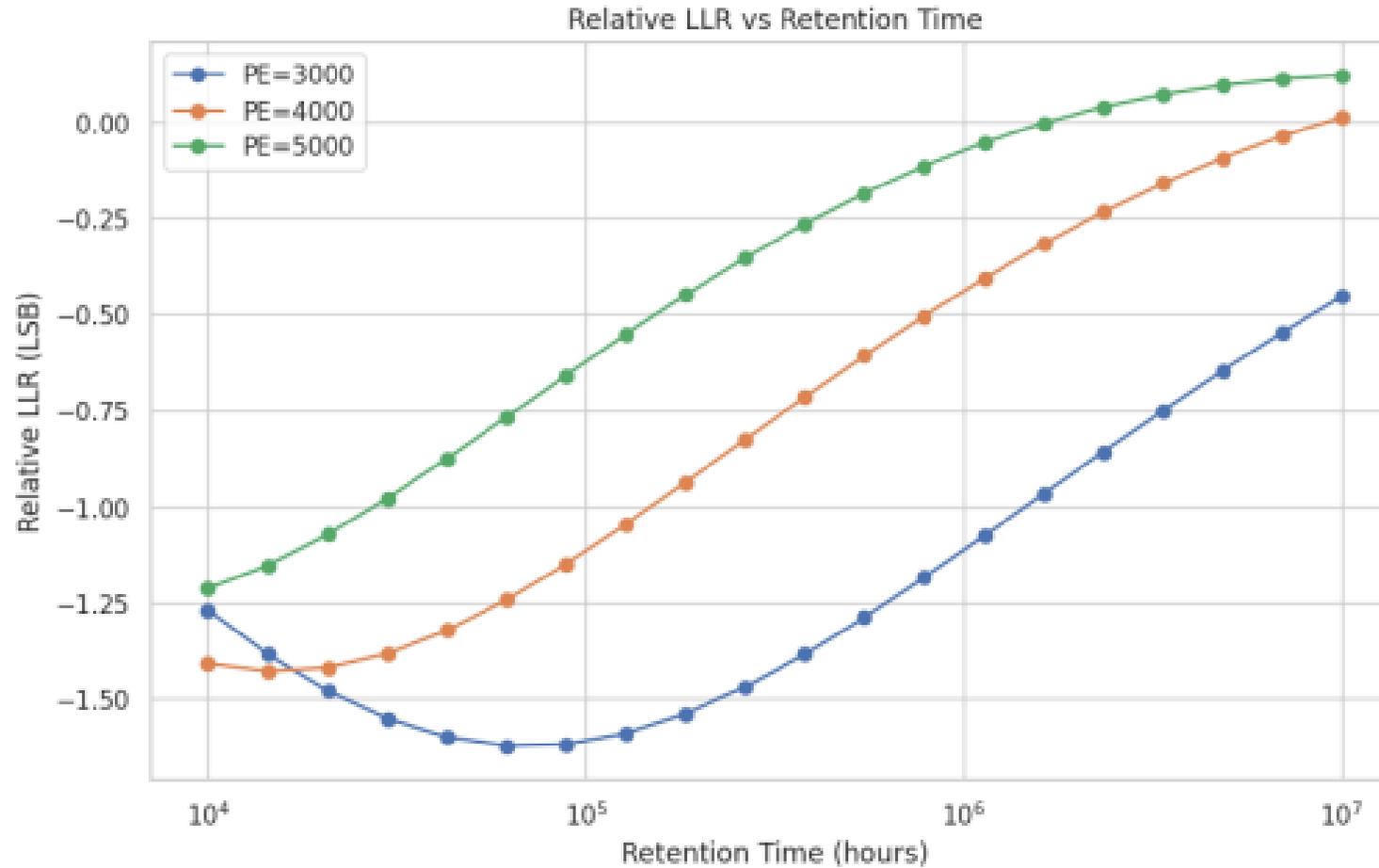
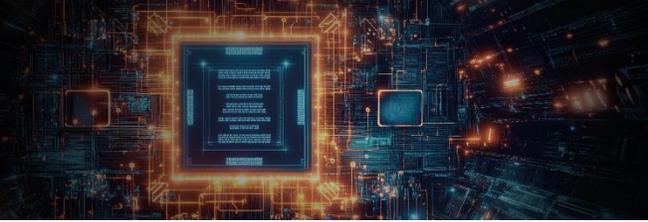
모든 개별적인 노이즈 요소를 결합한
NAND 플래시 메모리의 임계 전압:

$$V_{th} = V + n_{RTN} + \Delta V_{CCI} - n_{retention}$$

Task 2. 열화 현상 모델링 - 전압 분포 모델링



Task 2. 열화 현상 모델링 - LLR과 리텐션 Time



목차

Task 1. 랜덤화 기법

Task 2. 열화 현상 모델링

Task 3. 해석가능 AI 프레임워크

목차

Task 1. 랜덤화 기법

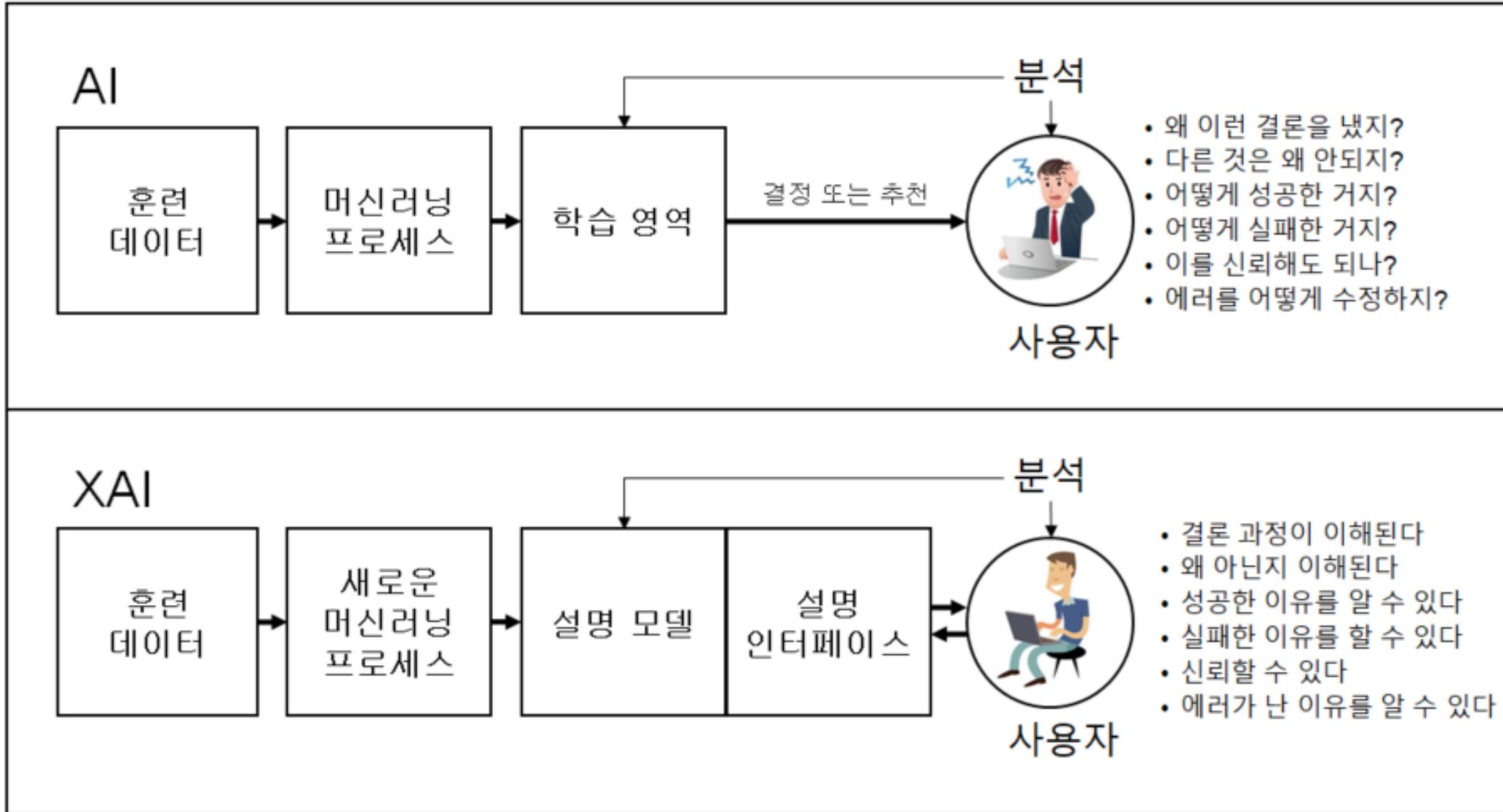
Task 2. 열화 현상 모델링

Task 3. 해석가능 AI 프레임워크

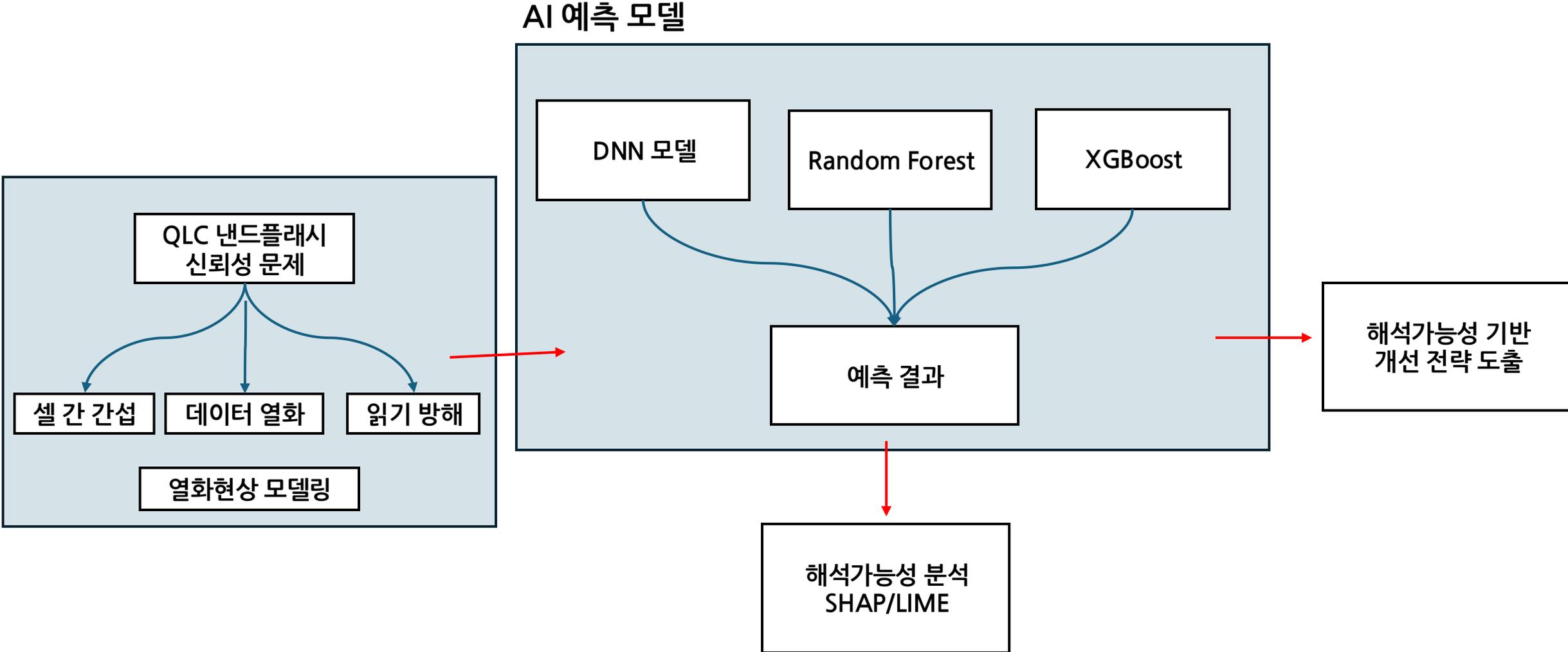
Task 3. 해석가능 AI 프레임워크



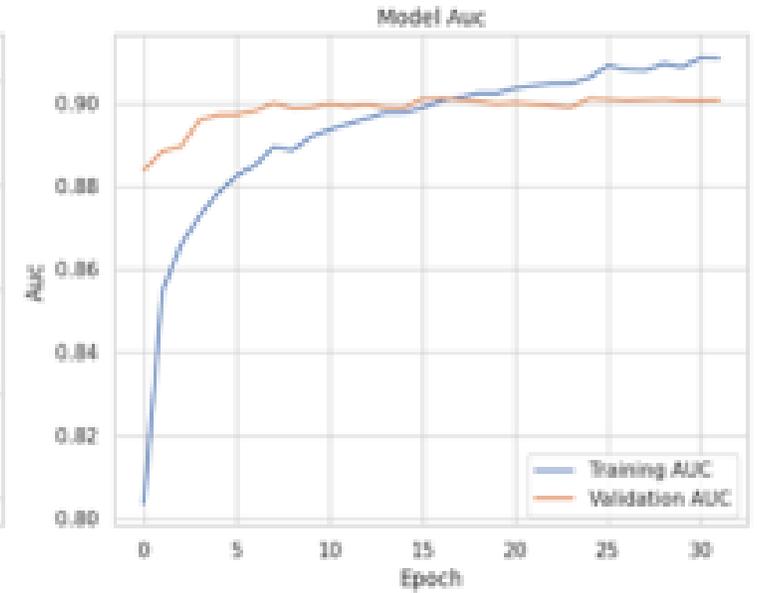
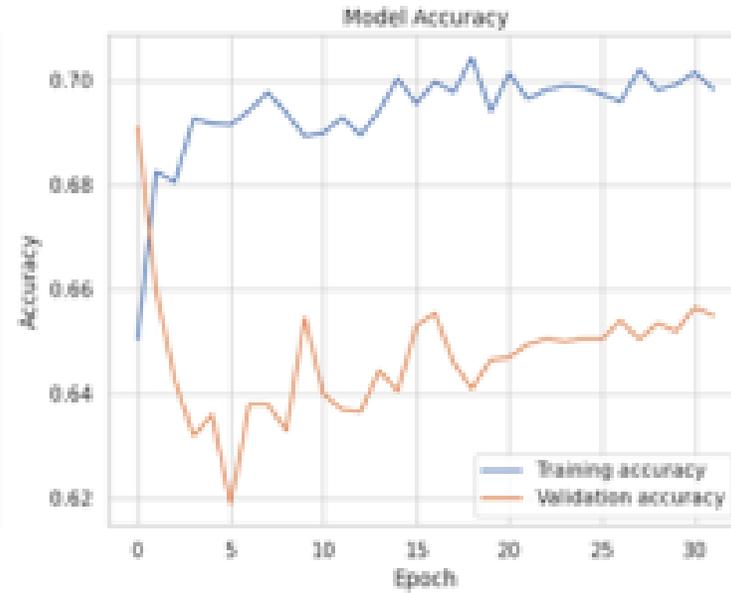
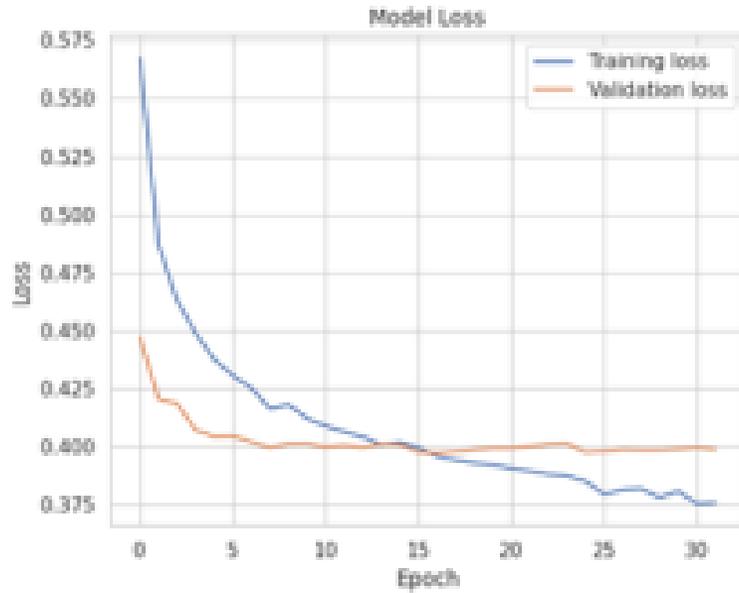
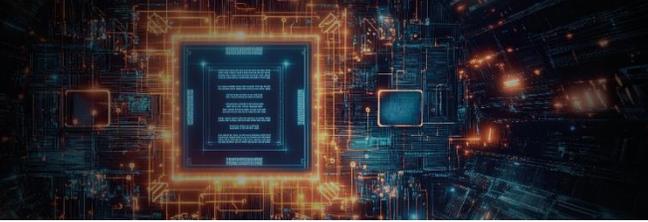
Task 3. 해석가능 AI란?



Task 3. 해석가능 인공지능 RAIN 프레임워크 (Reliability Assessment and Improvement framework for NAND flash)

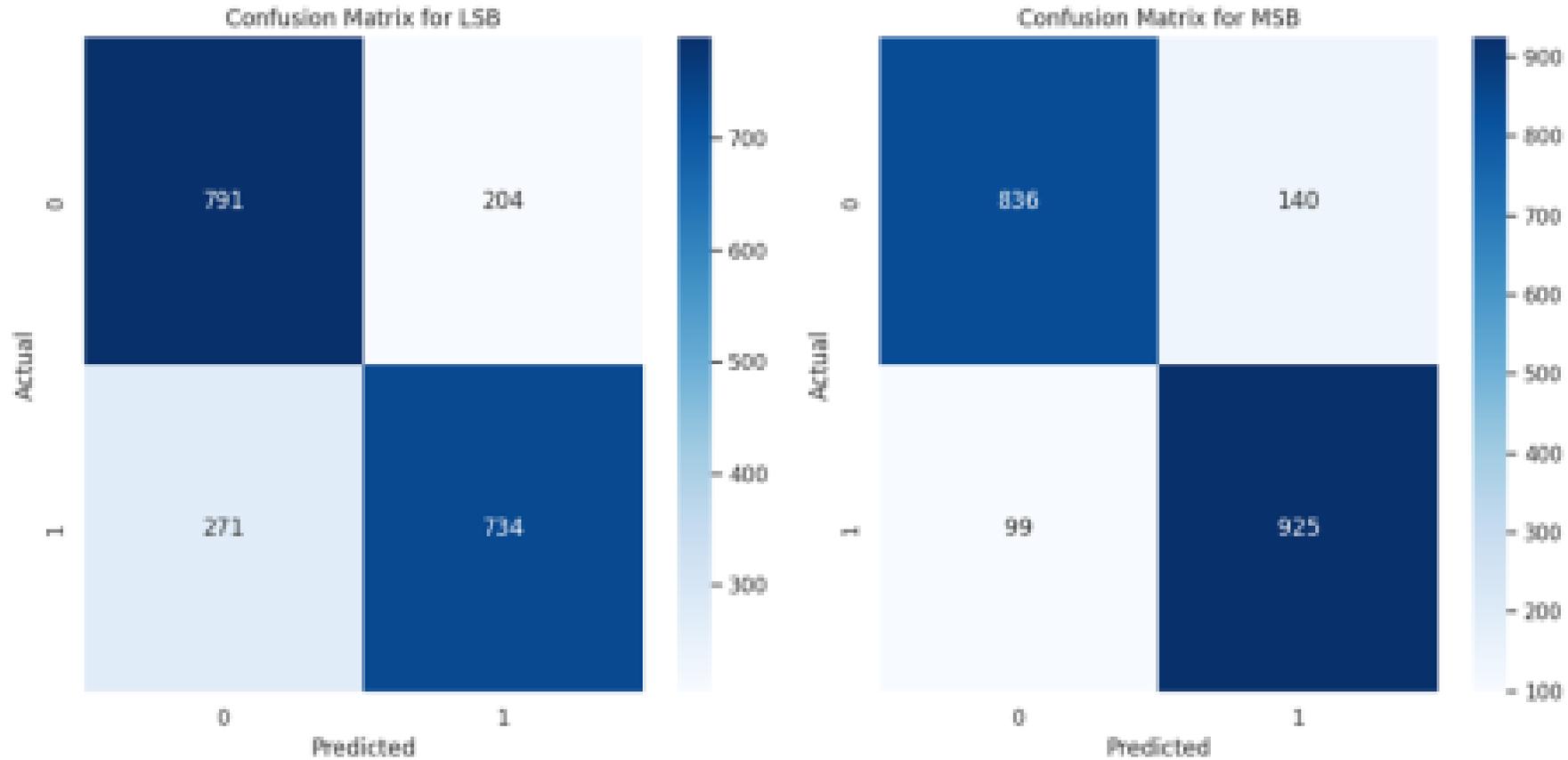


Task 3. 해석가능 AI 프레임워크



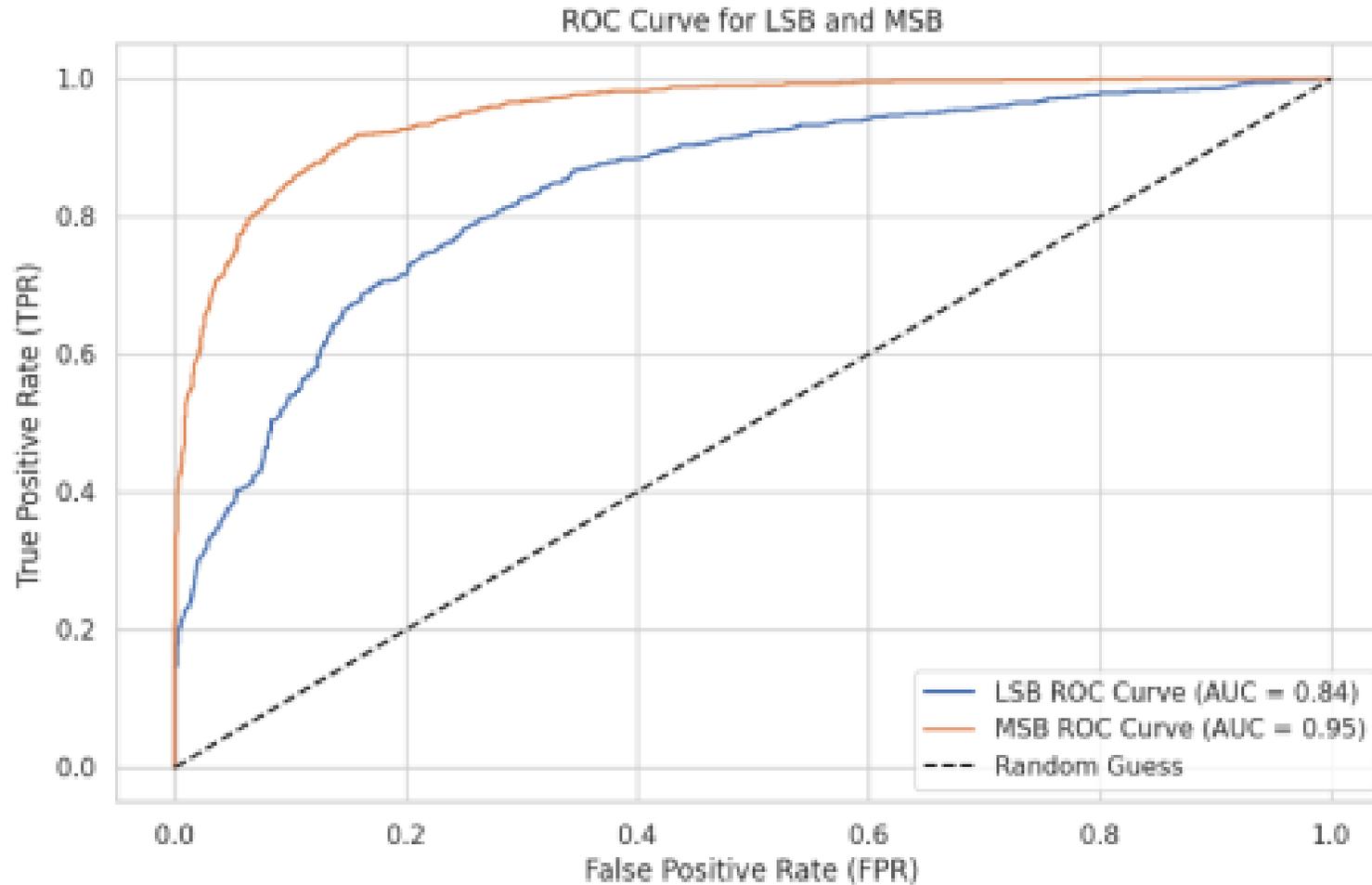
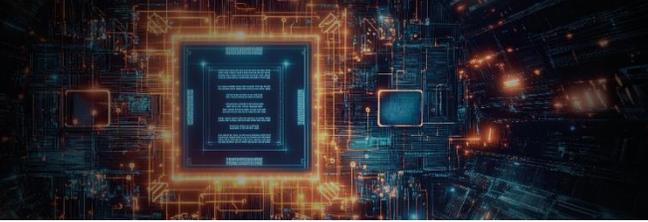
모델이 안정적으로 잘 학습되었음을 시각화를 통해 확인 가능

Task 3. 해석가능 AI 프레임워크



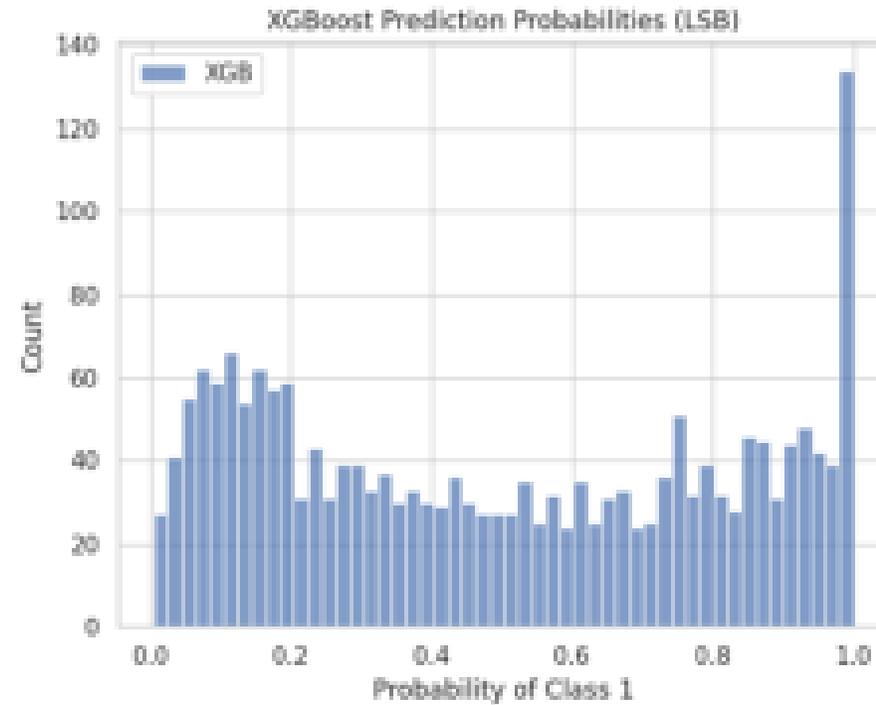
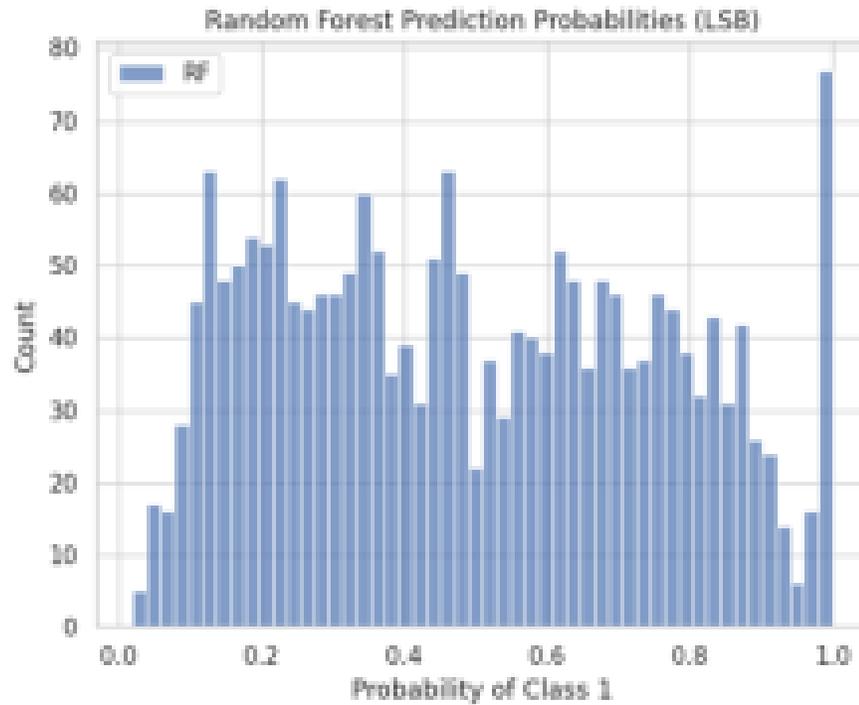
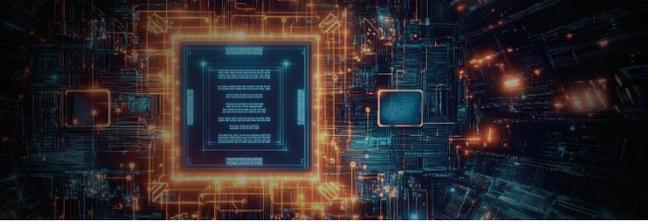
혼동 행렬을 통한 모델 파악

Task 3. 해석가능 AI 프레임워크



ROC커브를 통한 모델의 성능 파악

Task 3. 해석가능 AI 프레임워크

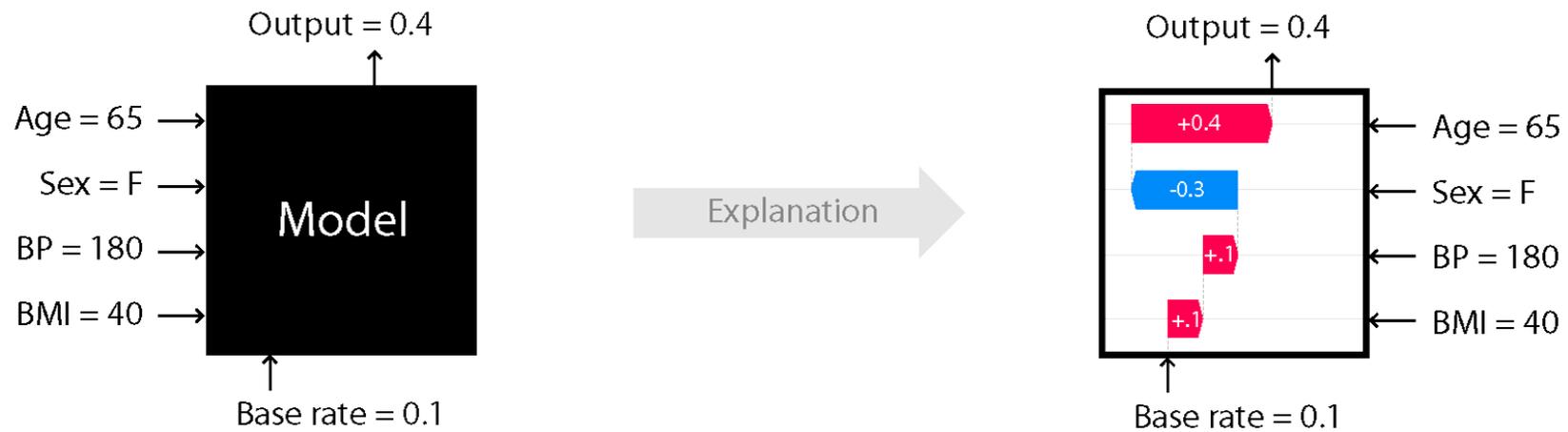


모델의 예측 경향성 확인

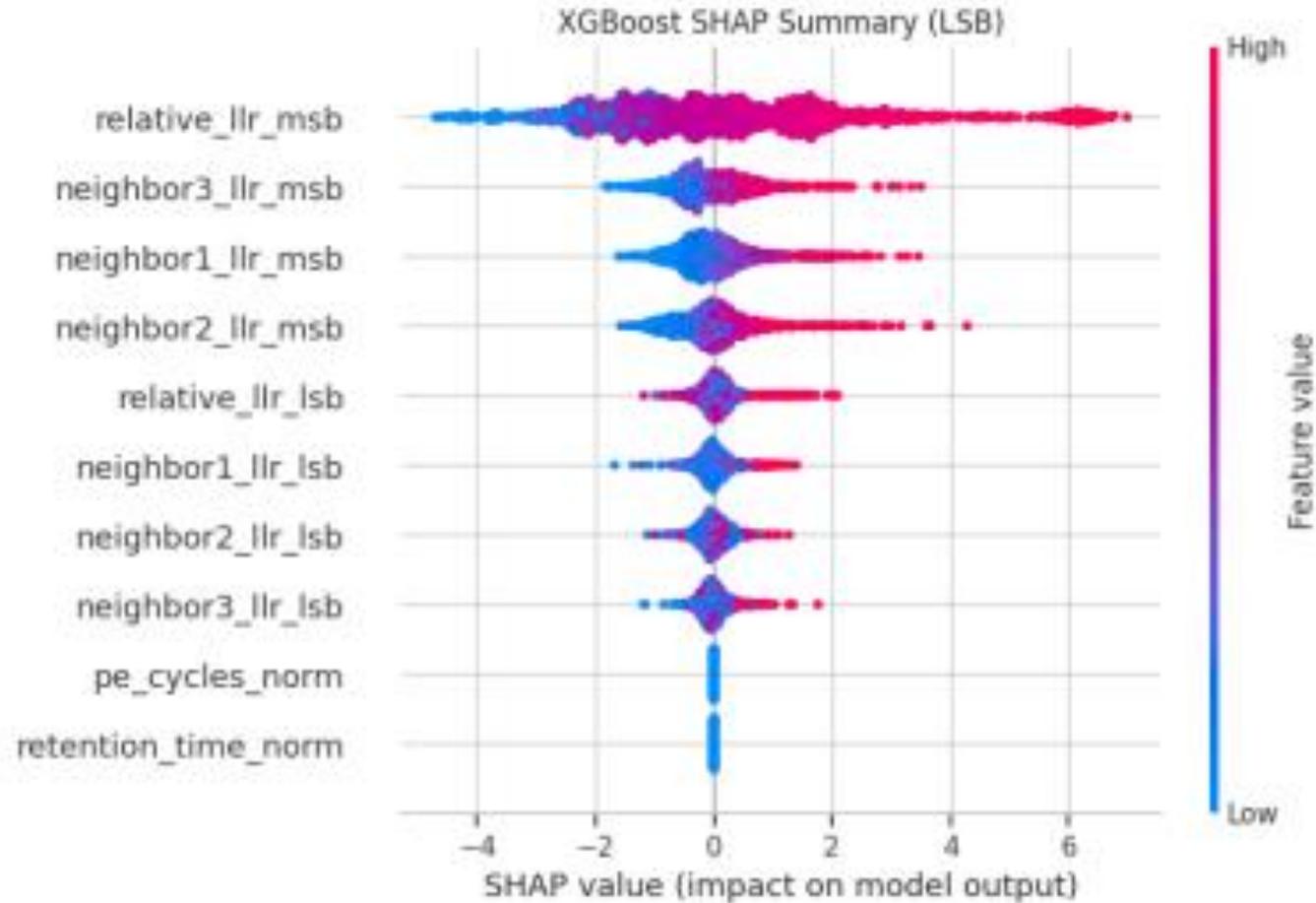
Task 3. SHAP란?



SHAP



Task 3. 해석가능 AI 프레임워크

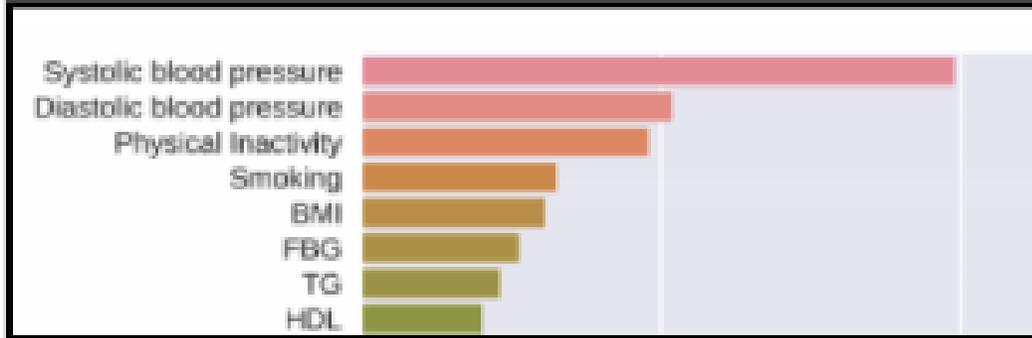


relative_llr_msb(가장 중요한 비트인 MSB의 LLR 값)가 약 31~37% 정도로 가장 큰 영향력

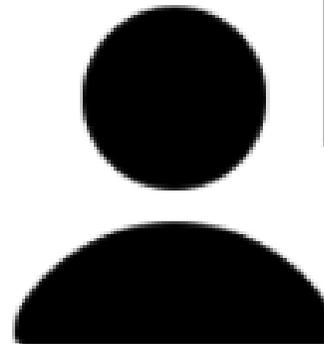
Task 3. LIME이란?



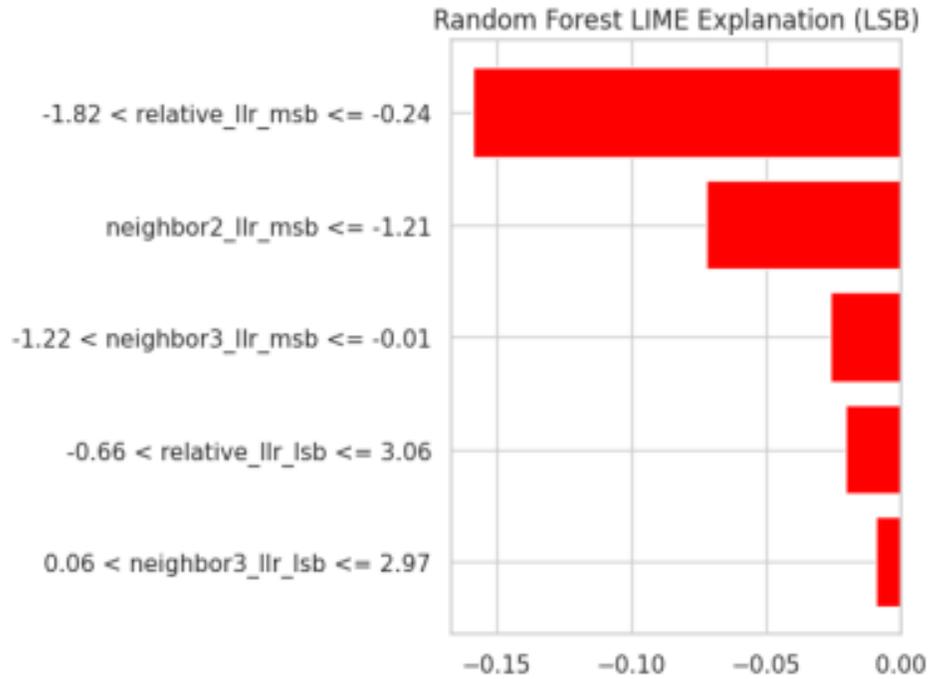
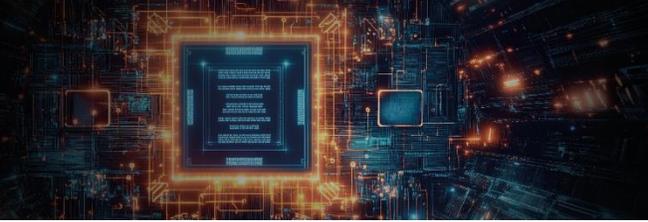
살을 좀 빼셔야..



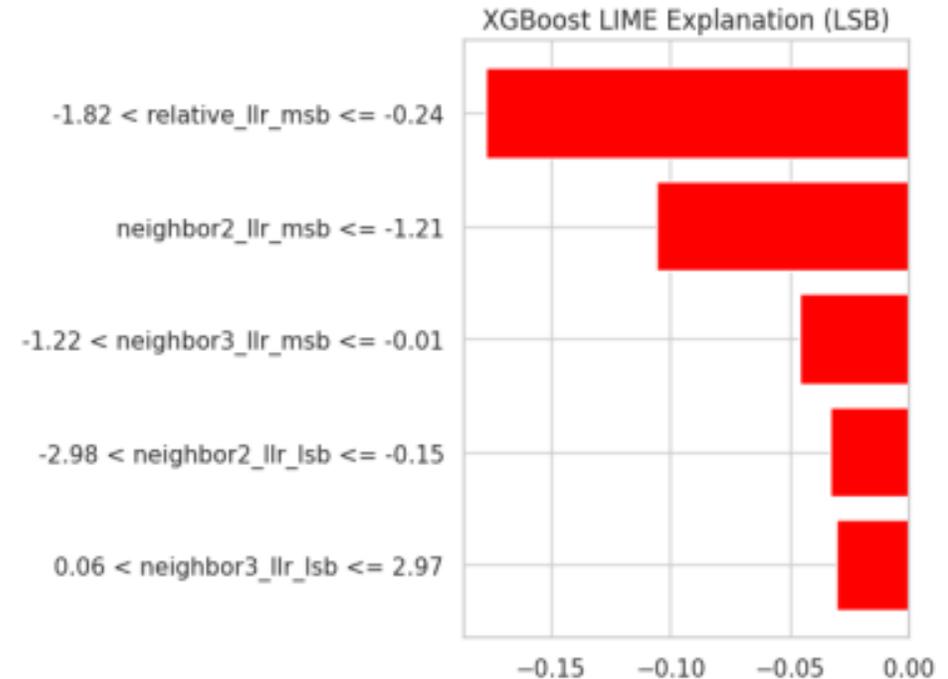
π



Task 3. 해석가능 AI 프레임워크



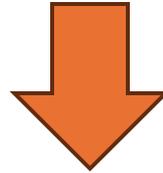
〈Random Forest 모델의 개별 예측 LIME 해석〉



〈XGBosst 모델의 개별 예측 LIME 해석〉

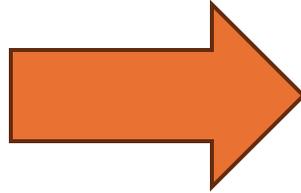
특정 개별 예측 사례에 대해 "이 셀의 MSB LLR이 이 정도 값일 때 오류 확률이 증가한다"는 식으로 지역적(로컬) 설명 가능

QLC 낸드플래시에서 발생하는
데이터 신뢰성 문제



"랜덤화 기법"
+
"해석 가능한 인공지능 Framework"

“오류가 난다”
X



“왜 오류가 나는지”
O

MSB 관리 강화, 인접 셀 간섭 완화, ECC 자원 효율적 할
당 등 구체적인 전략 수립 가능



Q & A