

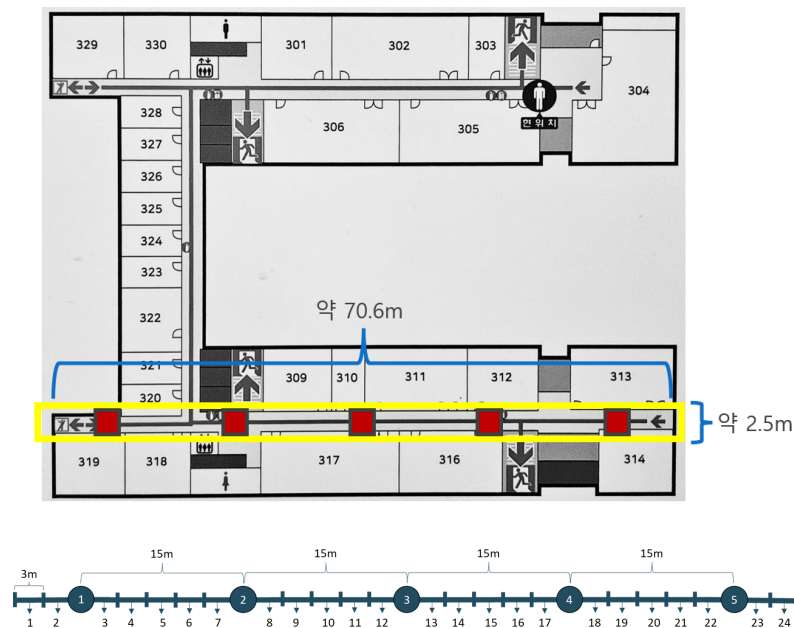
10월 2주차(2024.10.07~2024.10.13)

Model

- 비콘을 활용한 위치 추정에서 가장 좋은 성능을 보인 방법은 연속된 5개의 데이터를 평균 처리한 후 Random Forest를 적용한 모델이다.
- 해당 모델은 평균 거리 오차가 가장 낮았으며, 구역 내 예측이 비교적 정확하게 이루어졌다. 그러나 정확도를 중시할 경우, XGBoost 모델이 더 나은 결과를 보였다.
- 전반적으로 RSSI 값의 변동성과 데이터 부족이 성능 저하의 원인이며, 추가 데이터 수집과 전처리 개선이 필요하다.

구조

- 면적
 - 폭: 약 2.5m
 - 길이: 약 70.6m
 - 높이: 약 3.3m
- 비콘 개수: 5개
- 비콘 간격: 15m
- Zone 개수: 24개
- Zone 간격: 3m



데이터셋

학습 및 검증 데이터는 연속된 5개의 수신 데이터를 평균 처리하였고, 테스트 데이터는 실시간으로 수신된 데이터를 사용해 진행했다.

[데이터 전처리 방법 1.]

- 학습 데이터: 8,871개
- 검증 데이터: 2,219개
- 테스트 데이터: 실시간으로 수신되는 원본 데이터를 기반으로 구성

[데이터 전처리 방법 2.]

- 학습 데이터: 7,475개
- 검증 데이터: 1,870개
- 테스트 데이터: 7,464개(각 구역 311개/실시간으로 수신되는 원본 데이터를 기반으로 구성)

데이터 전처리

1. 연속된 5개 데이터의 평균
2. 각 비콘의 평균 RSSI 값에서 오차가 5를 초과하는 데이터를 제거한 후, 유효한 데이터를 기준으로 5개 데이터의 평균

모델 성능 비교

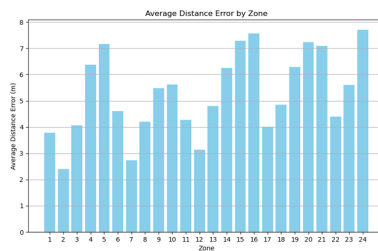
[데이터 전처리 1. 연속된 5개 데이터의 평균]

- Random Forest로 진행
- 각 구역의 검증 정확도

Zone	1	2	3	4	5	6	7
정확도	76%	71%	62%	59%	67%	59%	62%
Zone	13	14	15	16	17	18	19
정확도	69%	55%	59%	60%	84%	71%	59%

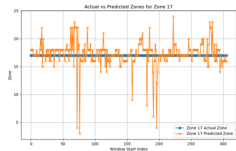
- 테스트 결과 각 구역의 평균 거리 오차

Zone	1	2	3	4	5	6	7
거리오차	3.77m	2.40m	4.06m	6.38m	7.17m	4.60m	2.73
Zone	13	14	15	16	17	18	19
거리오차	4.80m	6.25m	7.29m	7.57m	4.01m	4.85m	6.28

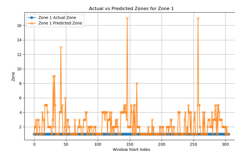


- 테스트 결과
 - 파란색 그래프: 실제 Zone
 - 주황색 그래프: 예측 Zone

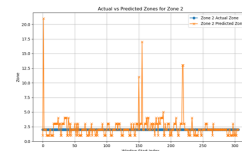
Zone 17: 30.39%(가장 높은 검증 정확도: 84%)



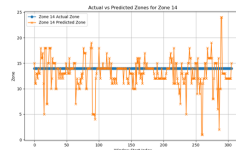
Zone 1: 47.39%(가장 높은 테스트 정확도)



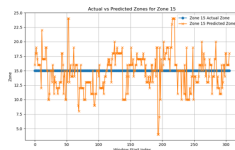
Zone 2: 45.75%(가장 작은 거리 오차: 2.40m)



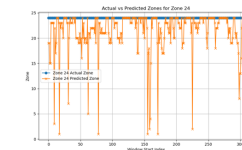
Zone 14: 22.88%(가장 낮은 검증 정확도: 55%)



Zone 15: 12.75%(가장 낮은 테스트 정확도)



Zone 24: 40.20%(가장 큰 거리 오차: 7.71m)



검증 정확도에 비해 테스트 정확도가 매우 낮음.. 데이터의 전처리 문제로 예상

[데이터 전처리 2. 각 비콘의 평균 RSSI 값에서 오차가 5를 초과하는 데이터를 제거한 후, 유효한 데이터를 기준으로 5개 데이터의 평균]

- Random Forest와 XGBoost로 진행
1. 앱으로 들어오는 각 비콘 RSSI 값의 평균 계산
 2. 각 비콘의 RSSI 값과 평균 비교 후 오차 5 초과 데이터 제거
 3. 유효한 데이터를 활용하여 연속된 5개의 데이터 평균 활용

RSSI 평균	60:98:66:32:98:58(B1)	60:98:66:32:8E:28(B2)	60:98:66:32:BC:AC(B3)	60:98:66:30:A9:6E(B4)	60:98:66:30:8E:28(B5)
Zone 1	-63.17	-77.57	-81.32	-81.22	-82.07
Zone 5	-71.91	-69.02	-76.30	-78.25	-82.15
Zone 11	-80.70	-70.22	-63.02	-75.21	-77.99
Zone 17	-83.54	-78.78	-73.57	-59.34	-71.38
Zone 19	-83.53	-79.79	-77.63	-67.41	-67.67
Zone 23	-82.73	-80.93	-82.36	-75.58	-55.12

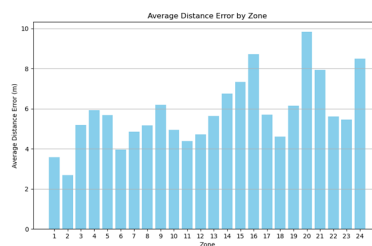
<Random Forest>

- 각 구역의 검증 정확도

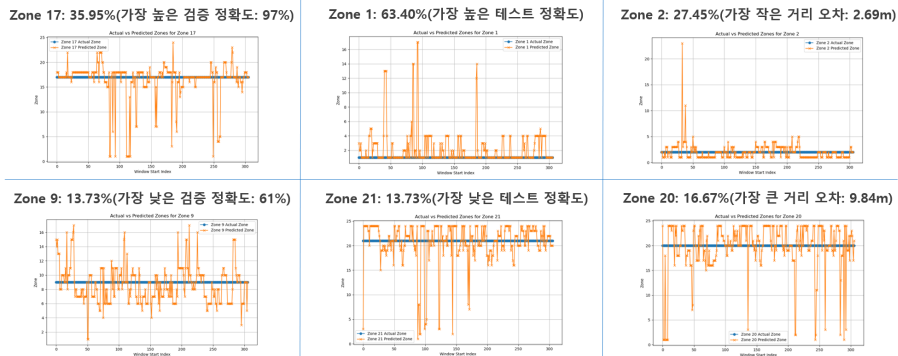
Zone	1	2	3	4	5	6	7
정확도	88%	70%	77%	73%	87%	81%	78%
Zone	13	14	15	16	17	18	19
정확도	88%	77%	83%	80%	97%	73%	75%

- 테스트 결과 각 구역의 평균 거리 오차

Zone	1	2	3	4	5	6	7
거리오차	3.58m	2.69m	5.20m	5.93m	5.69m	3.97m	4.86m
Zone	13	14	15	16	17	18	19
거리오차	5.63m	6.75m	7.34m	8.72m	5.70m	4.60m	6.15m



- 테스트 결과



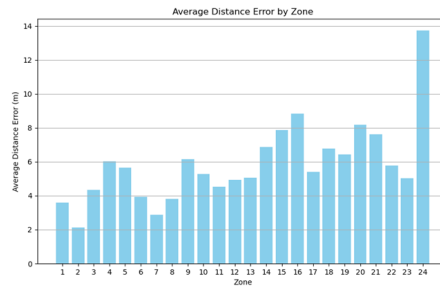
<XGBoost>

- 각 구역의 검증 정확도

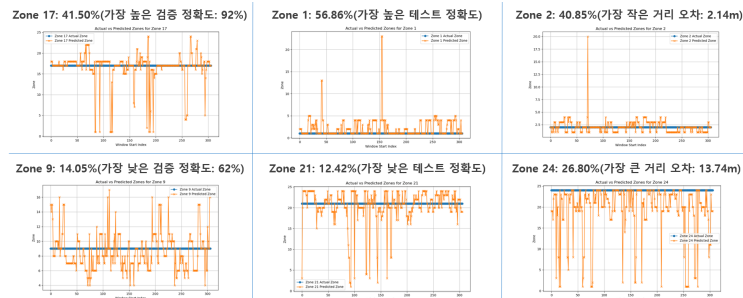
Zone	1	2	3	4	5	6	7
정확도	86%	75%	72%	81%	87%	81%	81%
Zone	13	14	15	16	17	18	19
정확도	82%	79%	89%	81%	92%	82%	74%

- 테스트 결과 각 구역의 평균 거리 오차

Zone	1	2	3	4	5	6	7
거리오차	3.60m	2.14m	4.34m	6.04m	5.64m	3.92m	2.88
Zone	13	14	15	16	17	18	19
거리오차	5.06m	6.85m	7.87m	8.84m	5.39m	6.78m	6.42



- 테스트 결과



종합 비교

1. 연속된 5개 데이터 평균(Random Forest)
2. 각 비콘 RSSI 값의 평균 오차 제거(Random Forest)
3. 각 비콘 RSSI 값의 오차 제거(XGBoost)

- 평균 거리 오차: $1 < 2 < 3$
- 정확도: $1 < 2 < 3$

- 1을 선택하는 경우: 평균 거리 오차 중요
- 2를 선택하는 경우: 거리 오차와 정확도의 균형 중요
- 3을 선택하는 경우: 정확도 중요

향후 계획

- **개선 부분 1.** 데이터셋의 부족(현재, 한 구역 당 약 350개)
 - 추가 데이터 수집 진행 완료(10/10)
- **개선 부분 2.** 데이터 전처리의 중요도 상승
 - 가중 최소 제곱법, 이동 평균 필터 등 고려
- **고려 사항**
 - 일자 형태의 구조에서 구역을 세분화하는 것은 매우 어려운 문제
 - 작업자의 위치 등의 거리 또는 좌표 고려
 - T자 형태 및 벽이 존재한다면 RSSI 값의 변화가 확실하여 비교적 쉬움
 - 비콘 수 \geq 구역 수, 구역의 명확한 구분