**오픈소스전문프로젝트**

**Report#06**

**담당 교수님: 박수창 교수님**

**담당 조교님: 김소연 조교님**

**컴퓨터공학과**

**8조**

**강주형, 이찬진, 정주형, 최준호**

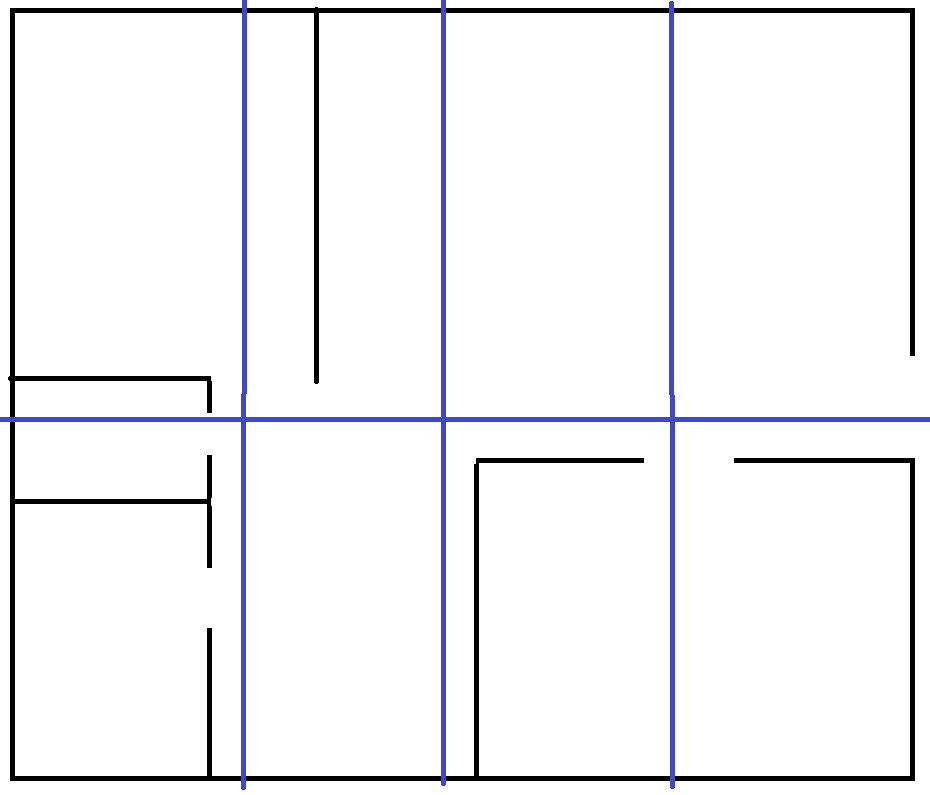
**목 차**

1. Wi-Fi 데이터 수집
2. 다층 퍼셉트론 (Keras)
3. 데이터 분석

**1. Wi-Fi 데이터 수집**

실내의 여러 곳에서 측정하는 Wi-Fi의 신호(RSSI)는 모두 다를 것이다. 이러한 점은 실내를 구역별로 구분할 수 있는 기준이 될 수 있다. 이를 이용해 각 구역에서 Wi-Fi들의 RSSI 값을 수집 및 측정하고 이를 통해 수집한 데이터들을 실내의 구역별로 나눈다. 각 구역별 RSSI 데이터들을 인공지능에 학습시킨 후 차후의 RSSI 수집값을 학습된 값을 기반으로 분석해 사용자가 현재 어느 구역에 있는지 표시할 수 있을 것이다.

RSSI 값 수집을 위해 계획중인 실내의 구역을 8개로 나누었다. 이때, 사진에 비해 실제 실내는 더욱 크기 때문에 한 곳에 서서 데이터를 수집하는 것은 그저 한 지점의 RSSI 값을 수집하는 것이기에 각 구역에서 다섯 지점에서 데이터를 수집하여 구역의 전체적인 RSSI 값을 수집하였다.



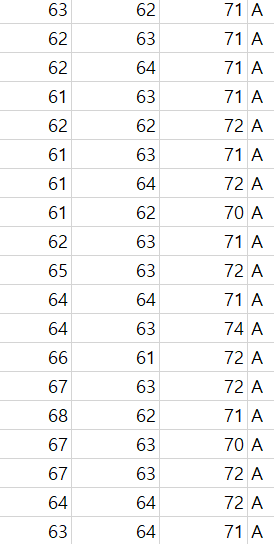
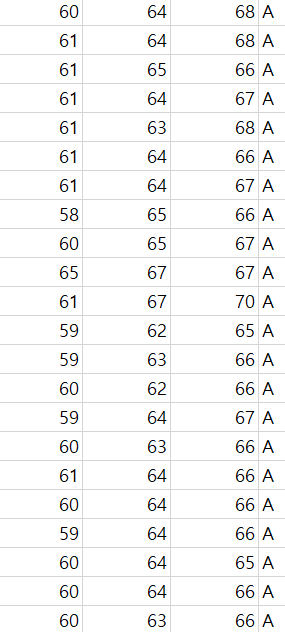
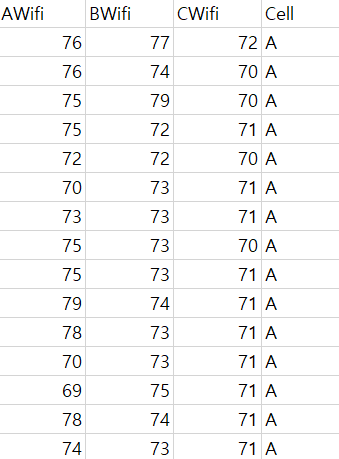
<왼쪽 위부터 오른쪽까지의 구역 이름: A,B,C,D,E,F,G,H>

8구역에서 각 구역마다 5지점에서 수집되는 각 Wi-Fi에 대한 RSSI 값을 100개씩 수집하였다.

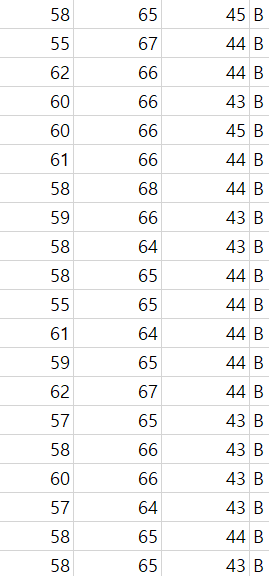
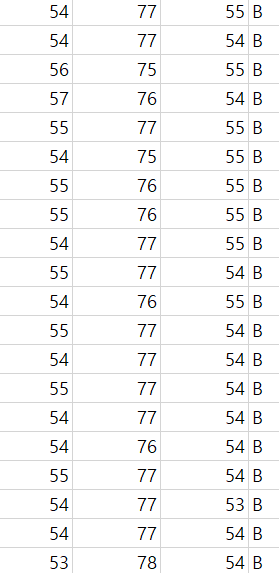
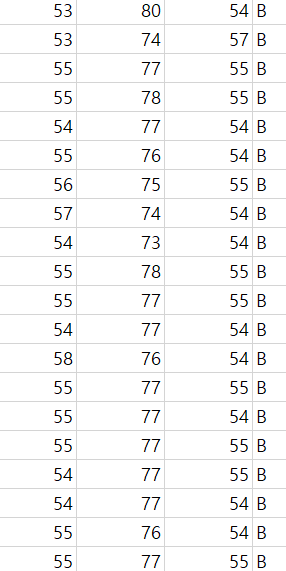
(구역당 500개씩 약 4천개)

많은 RSSI 값 중 공통적으로 수집되는 Wi-Fi는 radioactivity, U+NetA990, olleh\_WiFi\_075B였기에 이들의 RSSI 값을 추출 후 양수로 변환, csv파일에 저장하였다.

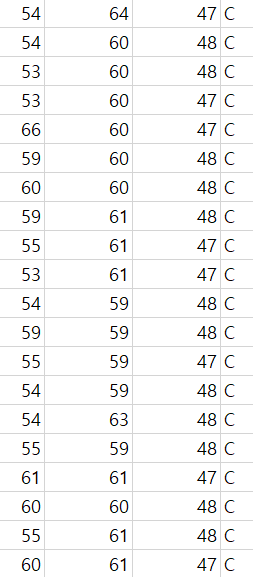
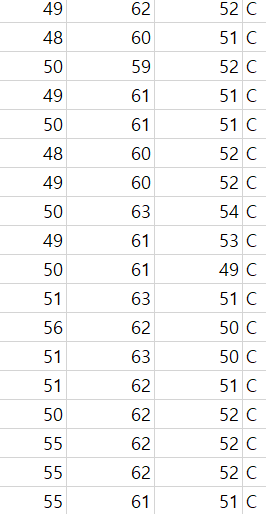
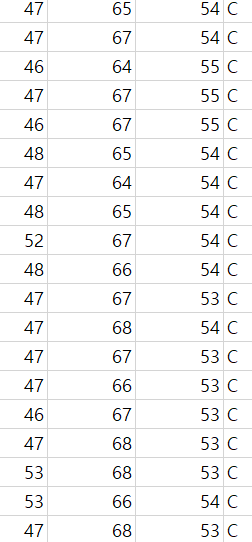
(여기서 Awifi=radioactivity, Bwifi=U+NetA990, Cwifi=olleh\_WiFi\_075B이다.)



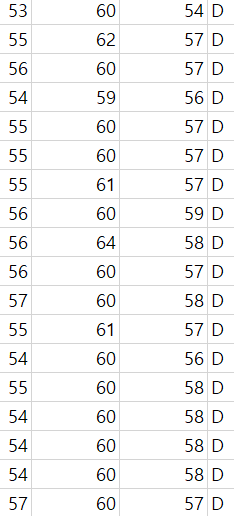
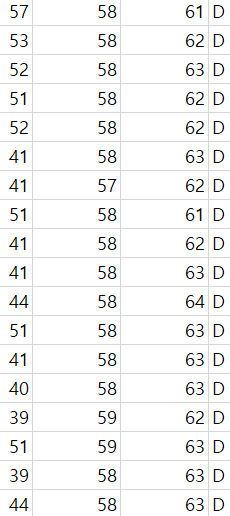
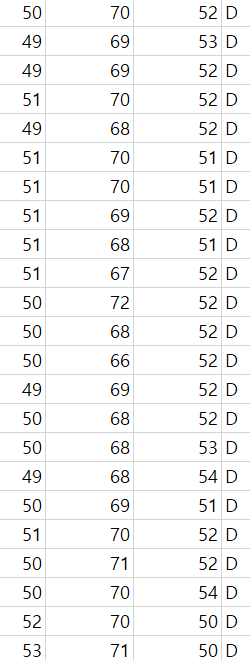
<A 구역의 RSSI 값>



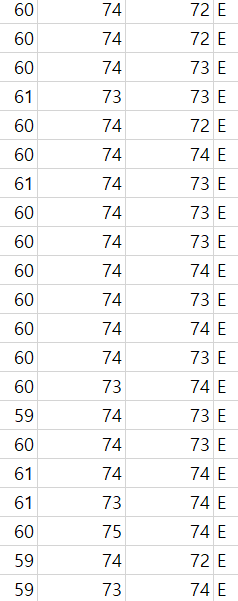
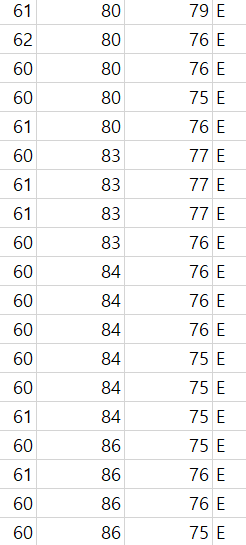
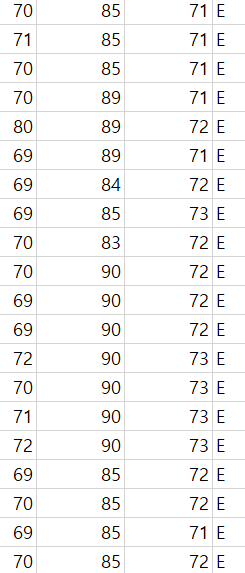
<B 구역의 RSSI 값>



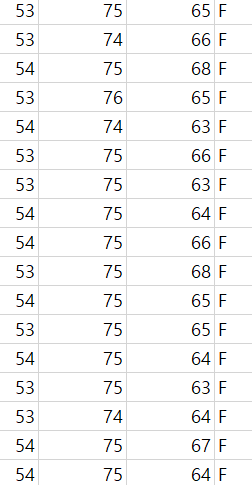
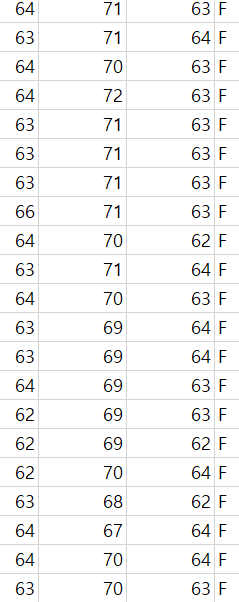
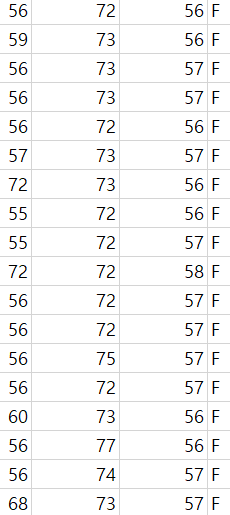
<C 구역의 RSSI 값>



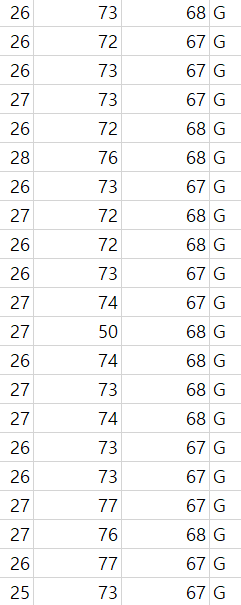
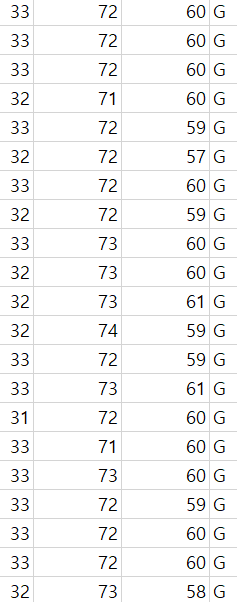
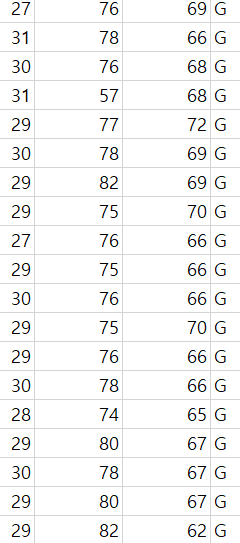
<D 구역의 RSSI 값>

‘

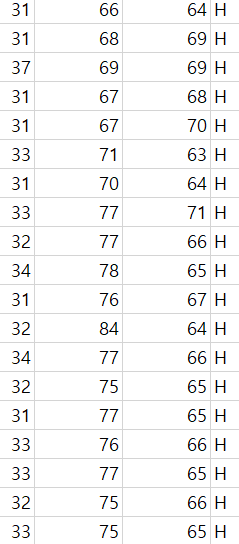
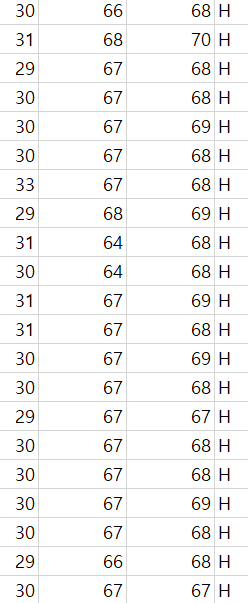
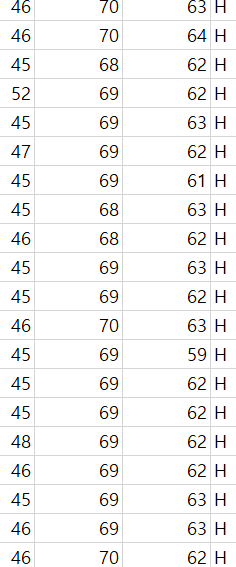
<E 구역의 RSSI 값>



<F 구역의 RSSI 값>



<G 구역의 RSSI 값>



<H 구역의 RSSI 값>

**2. 다층 퍼셉트론(Keras)**

작업 환경 : 구글 Colab

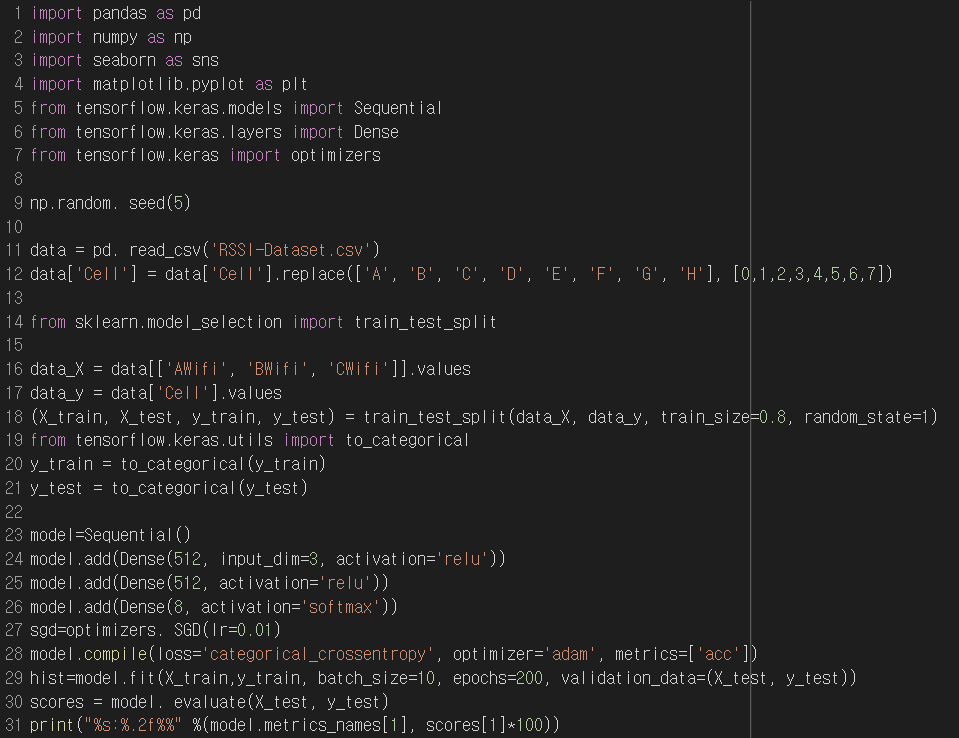
구글에서 오픈소스로 공개한 기계학습 라이브러리이다. 케라스를 이용해 수집한 RSSI 값들을 주고 학습시킴으로써 사용자가 어플 사용 시 수집되는 Wi-Fi들의 RSSI 값을 읽어 스스로 어느 구역인지 분석시키도록 한다. 이번 실험에선 데이터를 전달하고 자료의 80%만 학습시키고 나머지 20%를 어떻게 분석하는지 알아보도록 한다.

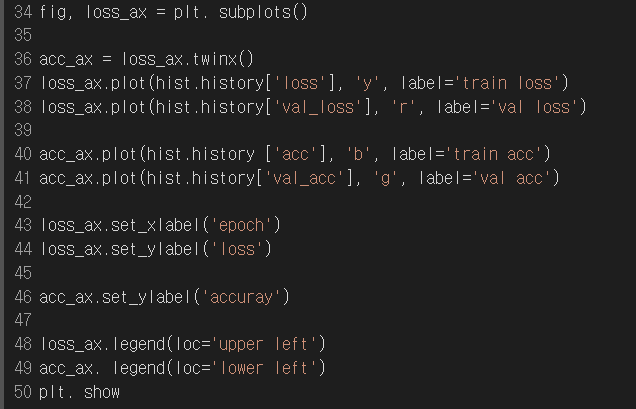
1. 세 개 이상의 선택지를 고르는 것이므로 다중 클래스 분류를 한다.

2. 8개 구역의 모든 확률의 합은 1이여야 하므로 출력층의 활성화 함수는 softmax를 이용한다.

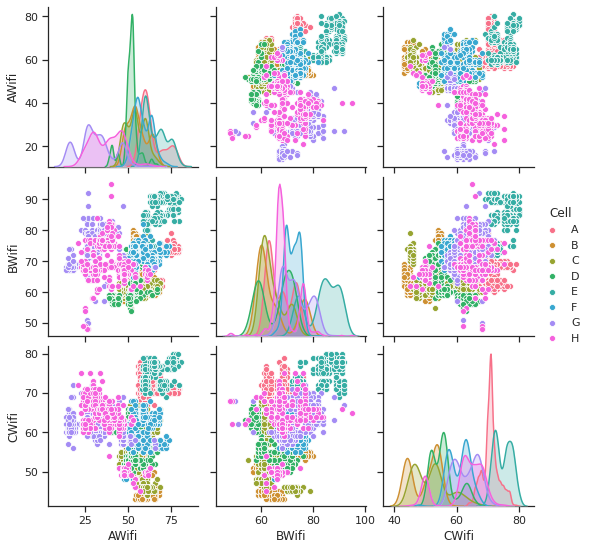
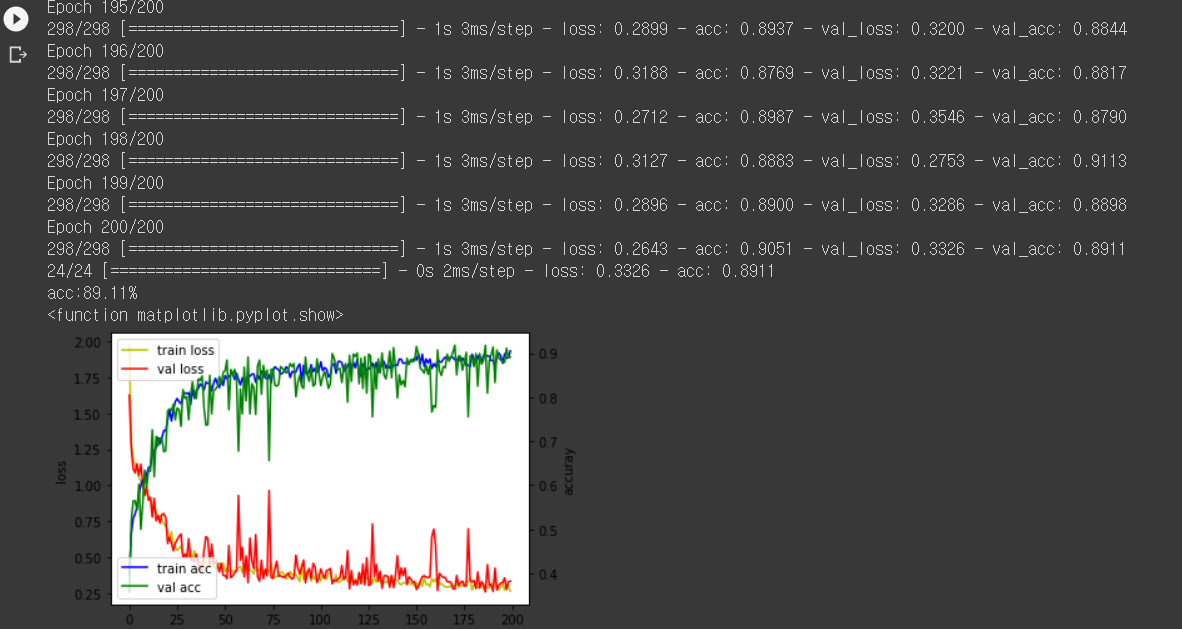
3. 각 클래스 간의 관계가 균등하므로 출력되는 값을 원-핫 벡터로 표현한다.

1)Source Code





2)결과



**3. 데이터 분석**

은닉층의 노드는 512개, 최적화 함수로는 adam이, 손실 함수로는 categorical\_crossentropy가 정확도가 우수했다.

구역별로 여러 Wi-Fi가 수집되었지만 그 중 모든 구역에서 공통으로 수집된 Wi-Fi가 있고, 이들은 구역별로 다른 값을 가지기에 구역을 나눌 수 있을 것이라 생각하였고, Tensorflow를 통해 이를 확인하였다.

Tensorflow를 통한 분석에서, 정확도가 약 89%로, 정확한 분석을 하기에는 한계점이 있다.

이는 데이터 수의 부족의 이유도 있을 것이고 또는 실제 실내에서 RSSI 측정 시 위의 실내지도보다 실제 실내가 더 크기 때문에 학습된 값보다 실제 측정될 RSSI의 값에 차이가 생길 수 있음을 보여준다. 따라서 RSSI 측정 시 셀을 더욱 나누거나 여러 군데에서 측정하여 구역의 RSSI값의 범위를 더 정밀하게 조정할 필요가 있다.