

딥러닝을 활용한 위험 지역 예측 기법

차주형 · 김성희 · 우영운

목적

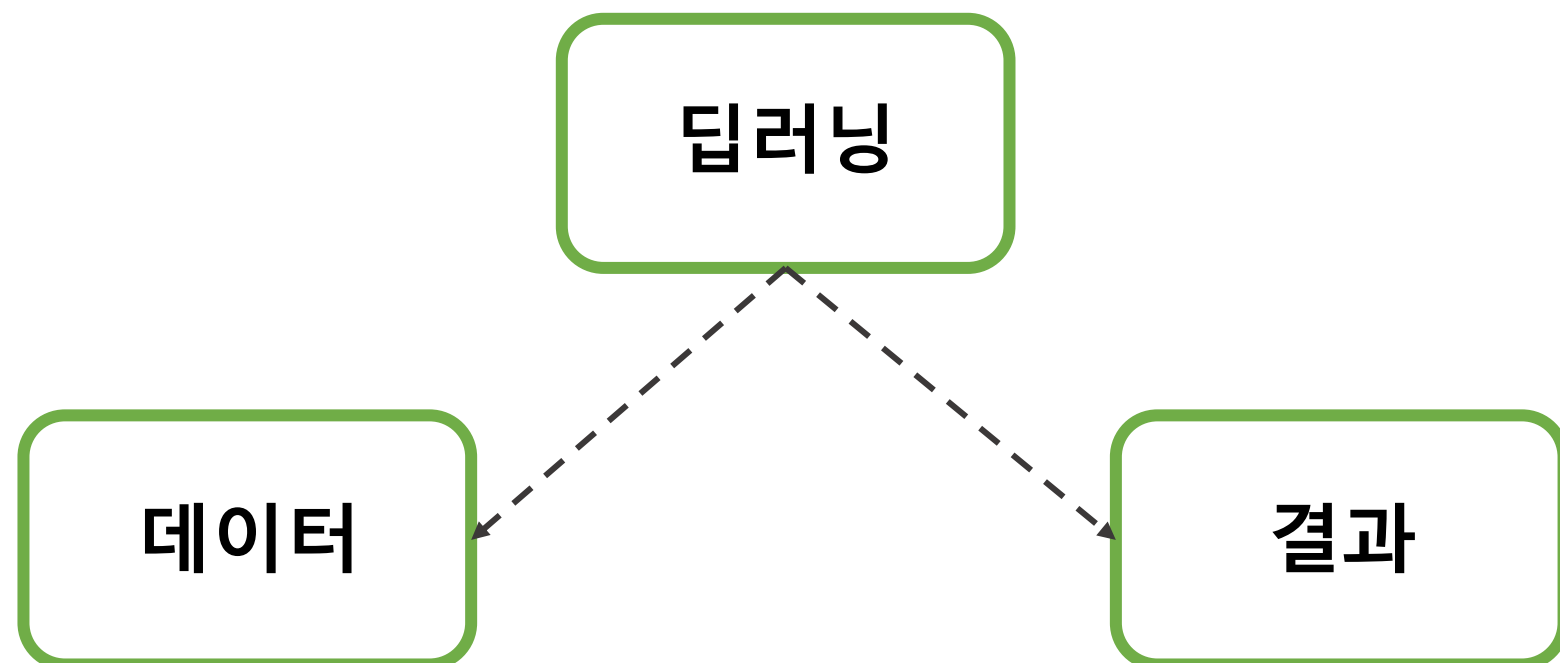
요약

이 논문에서는 유해-불건전 장소, 청소년의 안전 사고 데이터를 수집을 하여 각 위치에 위험 지역 정보를 지닌 태그를 지정 하여, '특정 구역이 어떤 공통적인 태그 정보를 가지는가'에 분석한다. 이때, 분석하는 기법으로 딥러닝을 통하여 분석한다. 딥러닝이 효율적인 예측 모델이 가능한지 확인 하기 위한 방법은 다음과 같다. 위치 데이터와 위험 요소를 임의적으로 만든다. 이때 특정 위치가 어떠한 위험 요소를 가질 수 있는지 결과를 예측하여 알려 주는 모델을 만들어 결과를 보여준다. 추후 개선 방안으로 시간 데이터나 다른 데이터와 같이 사용하여 개선 방향을 제시한다.

접근

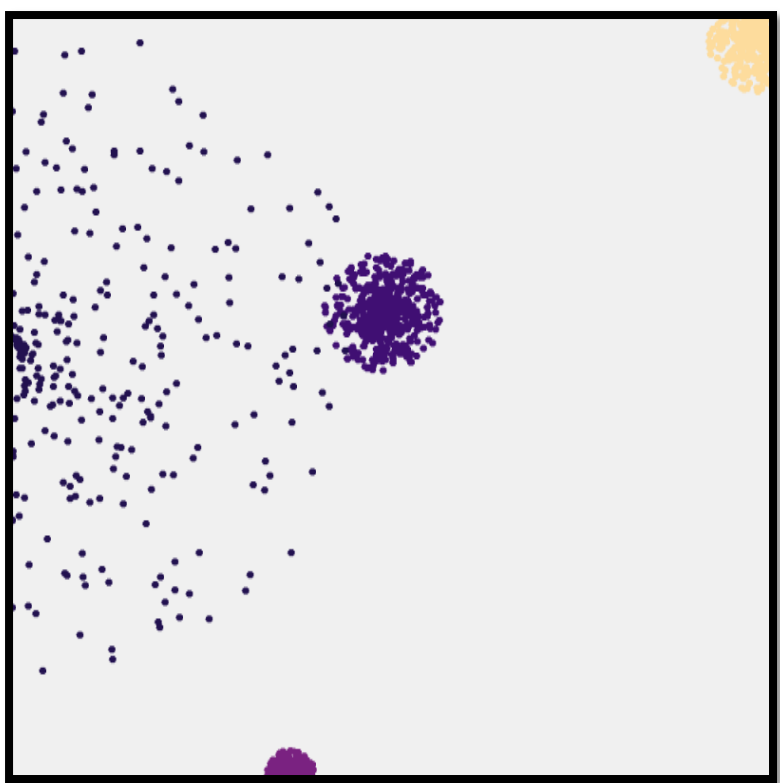
이 논문에서는 사용자가 위치와 위험 요소 데이터를 입력했을 때, 해당 데이터를 활용하기 위해서 데이터를 군집 분석을 통하여 해야한다. 그래서 데이터를 형태를 분석한 결과 입력과 결과 데이터가 따로 존재 하기 때문에, 딥러닝을 통하여 데이터를 군집화 하는 방향으로 정하였다. 이때, 출력 노드에서 소프트 맥스를 이용 하여 하나의 위험 요소를 출력하여 결과를 확인 할 수 있게 하여, 클러스터링 결과를 한눈에 볼 수 있는 이점이 있다.

결과

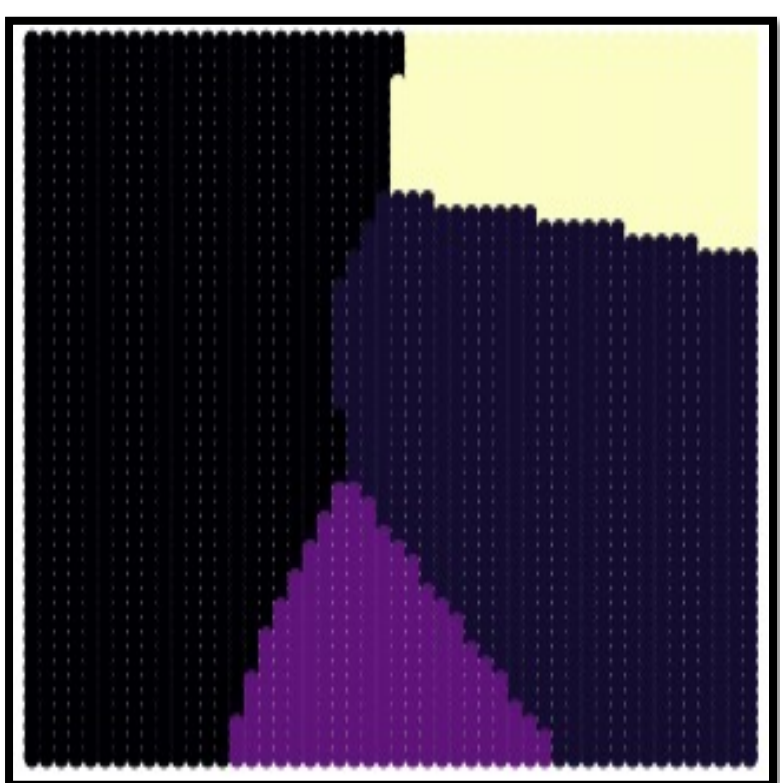


< 예측은 위와 같은 시퀀스를 지닌다. 시퀀스는 위치 데이터를 입력 받으며, 위험 요소 정보를 반환 한다. >

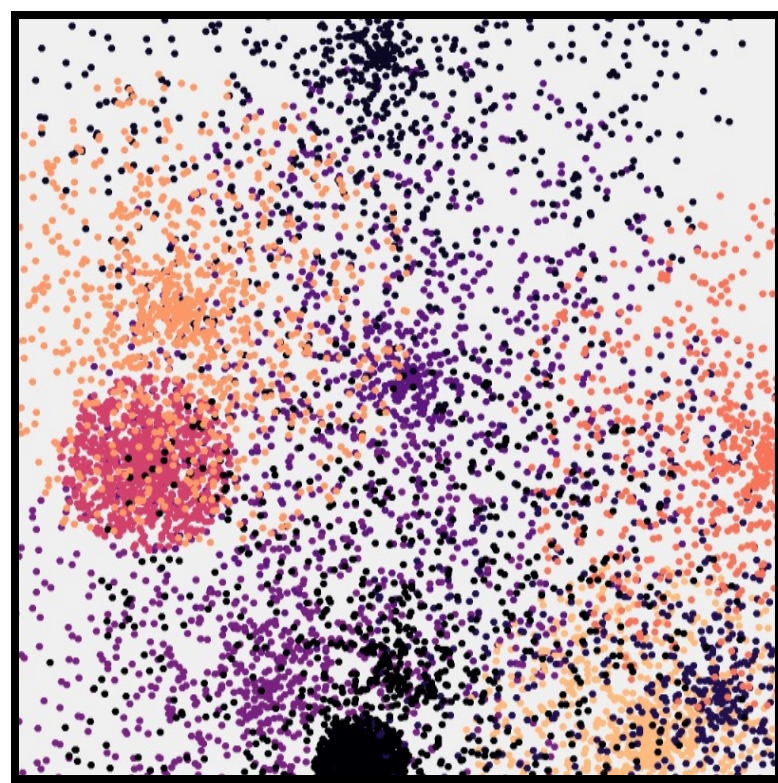
4개 태그 정보



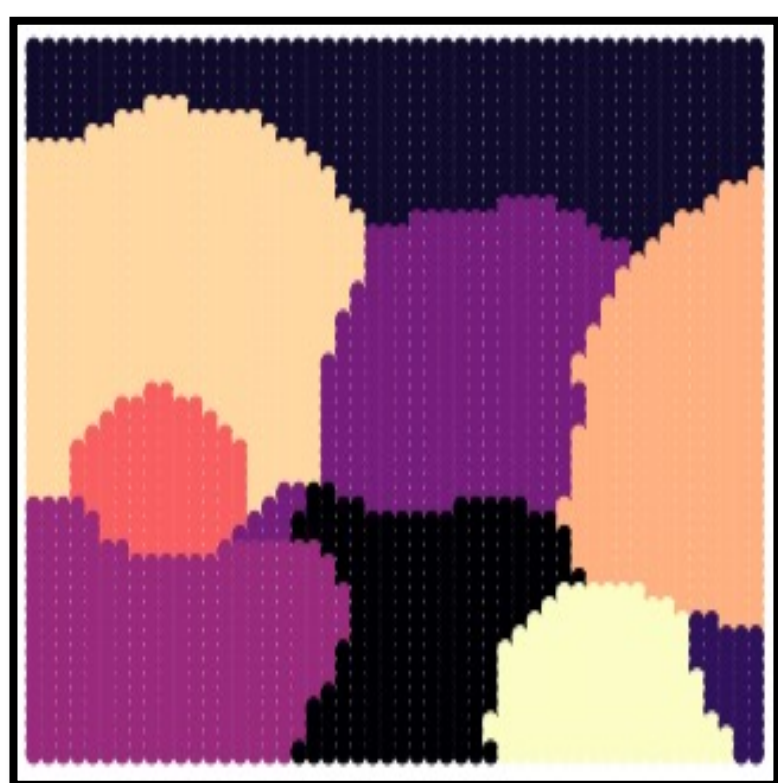
4개 태그 학습 결과



9개 태그 정보

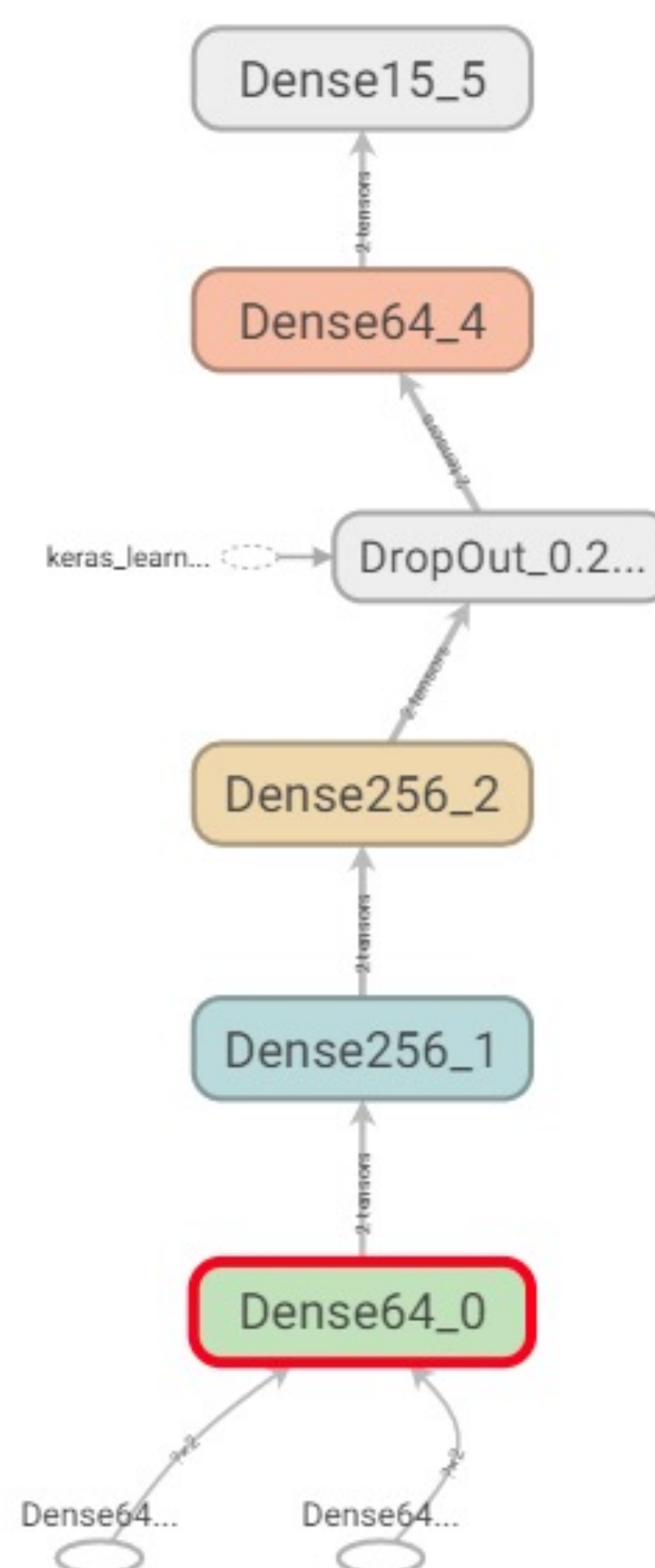


9개 태그 학습 결과



< 제시한 딥러닝 모델을 구현하여 표현을 한 형식 중 하나. 위 그림은 위치 데이터와 위험 요소 데이터와 함께 학습하여 지도 형식으로 결과를 출력한 결과이다. >

접근



< 제안한 딥러닝 모델. 아래가 입력이며, 위쪽이 출력 노드 이다. 2개의 데이터를 입력 받으며, 소프트맥스를 통하여 15개의 위험 요소를 반환 한다. 딥러닝을 통하여 추후 데이터 추가 및 노드 수정이 쉽게 하였다. >

결론

태그 기반으로 해당 영역이 어느 정도의 위험도를 가지는지에 대한 알고리즘을 딥러닝 기법으로 작성하였다. 입력의 개수가 2개밖에 되지 않으므로 이 개수를 확장할 필요성이 있다.

그래서 위치 데이터를 위도, 경도로 받는 것이 아닌 약 1km² 단위 영역을 분할 하여 데이터를 세분화 한다. 추가로 영역별 관계 및 영향을 거리별로 가중치를 조절을 하는 방식으로 학습을 한다면 더 좋은 모델이 가능할 것 이다. 또한 시간데이터를 추가하여 과거 데이터와 현재 데이터 함께 학습이 가능할 것이다.