

박서진 기말 프로젝트 발표 시작하겠습니다.

저는 최근 10년간 한반도 지진 발생 데이터를 분석하였습니다.

올해 2024년 6월 12일 오전 8시 26분경 전라북도 부안군에서 규모 4.8의 지진이 발생하여 많은 국민들을 놀라게 하였습니다. 이번 부안 지진은 올해 발생한 최대 규모의 지진으로 지진이 많지 않은 호남 내륙에선 역대 최대 규모를 기록했다고 합니다. 그동안 한반도에서 호남권은 상대적으로 지진이 발생하지 않는 안전지대로 여겨졌습니다. 하지만 이번 부안 지진으로 인해 전문가들은 한반도 전역에서 언제든지 강한 지진이 발생할 수 있다고 입을 모았습니다. 이와 같이 우리나라의 지진 데이터를 분석하여 지진을 예방하는 것은 필수가 되었습니다. 지진은 주기를 가지며 특정한 단층에서 일어나는 재해이므로 자주 발생하는 지역에서는 대비가 필요합니다. 따라서 저는 한반도에서 발생하는 지진의 패턴을 분석하고, 지진이 자주 일어나는 지역과 지진의 규모를 정리하여 미래의 재해를 대비하는데 도움을 주고자 주제를 선정했습니다.

지진 발생 데이터는 2015년 1월부터 2024년 6월까지의 데이터를 모았고 출처는 기상청 기상자료개발포털에 있습니다. 지진데이터 csv파일의 데이터 칼럼은 총 6개로 시간, 규모, 진앙, 위도, 경도, 위치로 구성되어 있고, 데이터의 총 개수는 1071개입니다. 데이터를 상세하게 알려드리면 NULL값은 없고, 규모와 위치는 오브젝트 타입, 진앙은 인트 타입 그리고 나머지는 플롯 타입입니다.

저는 파이썬을 이용하여 데이터 분석을 진행했고, 툴은 스파이더를 사용했습니다. 사용한 모듈과 라이브러리는 다음과 같습니다.

저는 우선 지진 데이터를 분석해 보았는데요. 연도별로 규모 3.0이상의 지진 발생 횟수는 왼쪽의 표와 같습니다. 과거 경주에서 발생했던 지진이 기억 나시나요? 그 시기였던 2016년에 3.0 이상의 지진이 가장 많았던 것으로 나왔습니다. 또한 10년간 지진 발생 횟수를 보았을 때 경북 지역이 월등히 높은 것으로 나왔습니다. 따라서 경북 지역에서는 지진 대비와 재난 매뉴얼 체계가 확실히 되어야 할 것으로 보입니다. 마지막으로 2016년 경북 지역 월별 최대 지진 규모를 관측했을 때 지진 발생이 1년간 꾸준히 일어난 것으로 관측이 되고, 그 중 3.0 이상의 규모가 5번이나 발생했습니다. 이는 우리나라 지진 위험성을 잘 보여주는 결과이고 이에

따라 경북 지역의 대비가 필요한 것으로 보여집니다.

다음으로 지진 데이터 예측 분석 결과를 보여드리겠습니다.

첫 번째로 선형 회기 분석을 사용하여 지진의 규모를 예측했습니다. 독립변수는 지진 발생연도, 월, 진앙위치, 위도, 경도이고, 종속 변수는 규모로 선정했습니다. 그리고 전체 데이터의 30%를 테스트 데이터로 사용하였습니다. 결과는 외쪽 사진처럼 꽤 비슷하게 규모를 예측하였고, mae는 0.32, mse는 0.24가 나왔습니다.

두 번째로 랜덤포레스트를 이용하여 지진의 지역을 예측했습니다. 위치 라벨을 생성한 후 육지를 0, 해역을 1로 두고 어떤 지역에서 지진이 발생했는지를 예측하는 모델입니다. 독립변수는 지진 발생연도, 월, 규모, 위도, 경도이고, 종속 변수는 위치 라벨 즉, 육지와 해역입니다. 이 또한 전체 데이터의 30%를 테스트 데이터로 사용했고 좌측 사진처럼 결과를 예측했습니다. 오차행렬을 이용하여 결과를 평가하면 정확도는 0.97로 꽤나 높은 점수를 기록했습니다.

마지막으로 프로젝트 결과를 정리하겠습니다.

한반도에서 일어난 규모 3.0 이상의 지진 횟수는 총 114회입니다. 강도가 큰 지진이 적지 않게 일어났다고 볼 수 있고, 특히 경남 지역에서 지진 발생률이 높기 때문에 특별히 주의해야 한다고 볼 수 있습니다.

또한, 선형 회기 분석을 통해 다양한 독립 변수를 이용하여 종속 변수인 지진의 규모를 예측해 규모가 큰 지진을 대비해 볼 수 있습니다.

한편으로는, 랜덤포레스트 분류기를 사용해 지진이 육지에서 일어날지 해변에서 일어날지 판단하고 해변 혹은 육지에 사는 주민들에게 차별적으로 예측된 지진을 알려주어 지진을 대피할 수 있습니다.

제가 준비한 발표는 여기까지입니다. 감사합니다.