|  |
| --- |
| **1. 주제**  스포츠 경기 우천 취소 예측 앱  **분반, 팀, 학번, 이름**  (가)반, 20221813, 정연재 |
| **2. 요약**  이 앱의 목표는 야외 스포츠 경기 당일 날 우천 취소 가능성을 예측하여 알려주는 것이다. 앱의 핵심 내용은 기상청 날씨 누리의 초단기 강수 예측 데이터를 불러와 스포츠 경기가 있는 곳의 6시간 이내의 강수 가능성을 예측해서 우천 취소 가능성을 제공해주는 것이다. 기상청 데이터와 지난 우천 취소 경기의 기상 데이터를 이용할 예정이다. 이 앱으로 우천 취소 예측으로 스포츠 경기를 즐기는 이들에게 편의성을 제공하고 경기장 사용에 있어 보다 효율적인 관리를 기대할 수 있다. |
| **3. 대표 그림**  - 개발 배경 : 스포츠 경기를 볼 때, 경기장에 가보지 전까지 우천 취소가 될지 안될지 모르는 불편한 상황이 발생  - 예상 결과 : 스포츠 경기 당일 날, 보러 갈 예정인 스포츠 경기를 클릭하면 우천 취소 가능성을 제공함  . |

|  |
| --- |
| **4. 서론**  **- 배경 설명, 사례 분석**  야구나 다른 스포츠 경기를 즐겨 보는 사람들이라면 경기장까지 갔다가 강수로 인해 콜드게임이 되거나 우천취소가 되는 경험을 분명 해보았을 것이다. 관중들은 왜 우천취소가 될 것이라고 예측하지 못했을까? 날씨는 기상청에서 제공하는 정보만으로 파악하기가 어렵다. 국내에서 제공하는 날씨에서 비 구름 정보 같은 경우 강수량을 제대로 알 수 없으며, 변수도 많고 정확하지 않을 때가 많기 때문이다. 이는 단순히 관중에게만 국한되는 문제가 아니다. 선수단과 경기장에도 피해가 있다. 아래는 엑스포츠 뉴스 기사의 몇 부분을 발췌한 내용이다.  *‘문제는 날씨다. 이날 수원 지역에는 오전부터 빗방울이 흩날렸다. … 홈 팀 KT는 일찌감치 내야를 방수포로 덮고 … 일기 예보상으로 이날 저녁까지 비구름이 수원 지역을 지나간다. NC는 만약 이날 kt전이 우천취소 될 경우 이틀 더 수원에 머물러야 해 여러 가지로 어려움을 겪을 수 밖에 없다.’* 이처럼 예측하기 힘든 강수 상황은 선수단에게도 컨디션 난조, 일정 문제, 시설 문제 등의 다양한 크고 작은 문제를 발생시킨다.  **- 문제 정의**  위와 같이 강수 등의 특수한 기상 상황으로 인해 경기 상황을 예측할 수 없어 피해를 보는 경우가 많다. 보통 기상청에서 제공하는 날씨는 강수량의 퍼센트 정도만 알 수 있어 정확한 예측이 힘들다. 야외에서 진행되며, 날씨의 영향을 많이 받는 스포츠 경기의 경우 우천으로 취소 되는 경우가 굉장히 많다. 특히 야구는 한 시즌에 10경기 이상이 기상 상황으로 인해 우천 취소된다. 하지만 경기를 관람하러 가는 관중들은 직접 경기장에 가보고, 경기가 시작되기 전까지는 그 가능성을 예측 할 수 없지만, 모든 사람들이 이를 경기 관람에 있어서 당연하게 겪을 수 밖에 없는 상황으로 여기고 있다. 하지만 나는 이러한 스포츠 시스템을 조금 더 개선할 수 있을 것이라고 생각했다. 그래서 스포츠 경기 운영에 있어서 강수 상황을 예측할 수 없다는 문제점을 개선하기 위해 이 프로젝트를 구상하게 되었다.  **- 극복 방안**  스포츠 경기를 즐기는 이들이 날씨 정보만 가지고는 쉽게 경기 상황을 예측할 수 없다는 기존의 한계점과 ‘예측 불가능한 기상 상황’ 은 보통 기상청에서 제공하는 날씨 상황(사용자들이 보편적으로 알고 있고 사용하는 기상 정보)보다 더욱 정확성이 높은 ‘초단기 강수 예측’ 데이터를 사용하여 해결 할 수 있다. 이 ‘초단기 강수 예측’ 데이터와 지난 우천 취소 경기 당시의 강수량 데이터, 스포츠 경기 우천 취소 기준 등을 종합하여 우천 취소 가능성을 예측하는 빅데이터 알고리즘을 만들 것이며, 이를 이용해 당일 경기 6시간 전에 우천 취소 가능성을 사용자에게 노출 시킬 것이다. |

|  |
| --- |
| **5. 본론 (1장 이내)**  필요한 기술 요소 중 가장 핵심적인 부분은 강수 상황에 대한 데이터이다. 강수 상황에 대한 데이터는 기상청 날씨 누리에서 제공하는 ‘초단기 강수 예측 서비스’를 이용할 예정이다. 초단기 강수 예측 서비스란 레이더 초단기 예측 자료와 수치모델 자료로 현재 실황부터 10분 간격의 강수 영역을 예측하여 시간당 강수량을 보여주는 서비스이다. 이러한 초단기 강수 정보는 색깔로 강수량을 구분할 수 있는 ‘비정형 데이터’이다. 비정형 데이터 마이닝을 이용해 데이터를 알고리즘에서 사용할 수 있게 한다.  두번째로 필요한 기술 요소는 우천 예측 알고리즘이다. 우천 예측 알고리즘 데이터베이스에는 우천 취소가 되었던 경기 당시의 강수 데이터와 야외 스포츠들의 우천 취소 기준이 저장 되어있다. 사용자가 알고 싶어 하는 경기를 입력하면 우천 예측 알고리즘이 실시간으로 ‘초단기 강수 예측’ 데이터를 불러온다. 의사결정 나무 기법의 랜덤포레스트를 이용하여 수집한 기상 데이터를 전처리 과정에서 라인 별로 분리 한 뒤 랜덤포레스트, Gradient Boosting 예측모형을 구축하여 우천 취소 가능성을 예측하고 우천 취소 선언 예상 시간을 도출한다. 사용자가 입력한 경기의 우천 취소 확률을 계산하고, 사용자에게 보여준다.  왼쪽은 간단하게 앱이 어떻게 작동되는지 시스템 개요를 나타낸 것이다. 여기서 반드시 필요한 기술 요소는 기상청 Data 와 우천예측 Algorithm 이다.  <시스템 개요>  스포츠 경기를 즐기는 모든 이들에게 경기를 보기 전, 예측할 수 없는 기상 상황으로 인한 불안함을 느끼지 않도록 하는 것, 보다 즐거운 경기 관람을 위한 편의성을 제공하는 것이 이 앱의 목표이다. 이 앱의 개발 방향은 앱의 예측 오류를 점차 줄여나가는 것이다. 사용자는 앱의 예측 값에 따라 경기를 보러 갈지 말지 결정할 것이기 때문에 오류를 최대한 0에 수렴하게 만들어야 한다. 따라서 부스팅 알고리즘의 Gradient Boosting Classifier를 도입하여 잘못 예측된 데이터에 대하여 가중치를 부여하여 오류를 개선해 나갈 것이다. |

|  |
| --- |
| **6. 결론**  이 앱은 스포츠 경기를 즐기는 이들을 위해 개발될 예정이다. 사용자가 원하는 경기를 입력하면 해당 경기 위치의 초단기 강수 예측 데이터를 불러와 이전 데이터와 종합하여 우천 취소 가능성을 예측하여 사용자에게 노출 시키는 것이 앱의 주요 내용이다. 이 앱은 관중들과 선수들에게 편의성 제공과 경기장을 효율적으로 관리하는 효과를 기대할 수 있다.  향후 할 일을 정리하자면 앞서 본론에서 언급했던 우천취소예측 알고리즘에 필요한 데이터를 모두 수집한다. 그 후 실시간으로 데이터를 불러와 기존 데이터와 비교 분석하는 빅데이터 알고리즘을 구현한다. 그리고 앱 UI를 디자인하고, 사용자가 사용할 button 을 간단명료하게 정리하고 꾸민다. 위 알고리즘을 앱에서 실행 구현하고, 앱을 시연하고 테스팅할 예정이다. |

**7. 출처**

[1] "고척돔 오늘 하루만 빌려서 야구하면 안 되나요?" [현장:톡], 엑스포츠 뉴스, 2022.10.10

[2] 네이버 지식백과 ‘콜드게임’,

<https://terms.naver.com/entry.naver?docId=1215762&cid=40942&categoryId=31946> (2022.10.13)

[3] 기상청 날씨누리 – 초단기강수예측, <https://www.weather.go.kr/w/image/vshrt/rain.do>

(2022.10.13)

[4] 오민지 외 3명, ‘게임데이터를 이용한 승패예측 및 세분화된 변수 중요도 도출 기법’, 한국빅데이터학회지 제 5권 제 2호, 2020,pp. 231-240

**8. 부록**

콜드게임 : 야구에서 양팀 모두 5회 이상 공격을 동일하게 진행한 뒤 해가 지거나 폭우 등의 이유로 경기를 진행할 수 없게 된 때

랜덤포레스트 : 지도 머신러닝 알고리즘. 의사결정나무 기법의 한 종류로 개별 의사결정나무의 결과를 결합하여 모델 생성.

Gradient Boosting Classfier : 약한 예측 모델들을 결합해 강한 예측 모델을 만드는 알고리즘으로 회귀와 분류 모두에 사용될 수 있음. 개별 의사결정나무의 예측 오차가 높다 하더라도 여러 개의 의사결정나무를 결합한 최종 예측 모델은 예측 오차가 낮음.