**Seoul FP-Weather Script**

Code States AI bootcamp 15th

정 경 재

안녕하세요. AI부트캠프 15기 정경재입니다.  
서울시 날씨 정보로 식중독 환자 수를 예측하는 프로그램인 서울 FP 웨더를 소개하고자 합니다.

우선 서론에서 프로그램에 대한 간단한 소개를 한 뒤에  
개발 과정을 다루고, 시연해보는 순서로 발표를 진행하고자 합니다.

서론입니다.

식중독은 식품 섭취에 의해 발생하는 질환이고 세균성 식중독이 가장 흔한 형태입니다.

이러한 식중독균은 35~36도 부근에서 가장 빠르게 증식하기 때문에 고온 다습한 여름철에 가장 위험도가 높습니다.

정부기관에서 발표하는 식중독 지수 역시 온도와 습도를 고려하여 계산하기 때문에 식중독은 날씨와 연관이 깊습니다.

따라서 저는 이러한 날씨 정보를 바탕으로 식중독 환자 수를 인공지능으로 예측해보는 프로그램을 기획하고 개발하게 되었습니다.

본 프로젝트의 전체적인 파이프라인은 다음과 같습니다.

식중독 환자수 데이터와 기상 데이터를 받아오고, 클라우드 데이터베이스를 구축하여 저장하였습니다.

그리고 그 데이터베이스를 기반으로 모델링을 진행하고, 웹페이지로 구현하여 배포하였습니다.

조금 더 세부적인 과정들을 살펴보겠습니다.

우선 데이터 수집 단계입니다.

식중독 환자 데이터는 식품의약품안전처 Open API를 통해 수집하였고,

서울시 월별 기상자료는 기상청 및 서울시 공공데이터를 다운받아 가공하였습니다.

기상 정보는 기상청에서 정의하는 8개의 주요요소 데이터를 수집하였습니다.

그 다음으로 아마존 RDS와 PostgreSQL을 이용하여 클라우드 관계형 데이터베이스를 구축하였고, 데이터 구조는 아래 그림과 같습니다.

이렇게 저장한 데이터들을 조합하고 추출하여 모델링 및 대시보드 제작에 이용하였습니다.

기상정보를 통한 환자수를 예측하기 위한 회귀분석을 실시하였고,   
속도와 성능이 뛰어나다고 알려진 LightGBM으로 모델링을 진행하였습니다.

만족스럽지는 않은 모델 성능이긴 하지만, 진행에는 무리가 없을 것이라 판단하고 진행하였습니다.

그 다음으로 대시보드와 웹페이지를 구현하였습니다.

트렌드를 확인하기 위한 대시보드는 구글 루커 스튜디오와 태블로로 구현하였고,

예측에 대한 대시보드는 루커 스튜디오로 구현하였습니다.

그 후에 웹페이지를 제작하고 플라스크와 헤로쿠를 통해 배포하였습니다.

이제 프로그램을 직접 시연해보도록 하겠습니다.

해당 링크를 통해 웹페이지에 접속해보도록 하겠습니다.

(클릭)

홈페이지는 홈, 트렌드, 예측, 정보로 구성 되어 있습니다.

<트렌드>

우선 트렌드 대시보드에서는

연도를 선택해서 보실 수도 있고,

특정 월을 선택해 확인 해 보실 수도 있습니다.

다음 대시보드에서는 계절별로 기상요인과 환자수의 관계를 살펴 보실 수 있습니다.

<예측>

그럼 이제 예측을 진행해보도록 하겠습니다.

이번달이 11월이니 11월 평년값 데이터를 입력해서 예측해보도록 하겠습니다.

이번달은 68명의 식중독 환자가 생길 것으로 보이네요.

예측은 이렇게 진행해 보실 수 있습니다.

<예측 대시보드>

이런 방식으로 평년 값을 통해 모델이 예측 한 결과 값을 대시보드로 정리하였습니다.

대체적으로 여름이 압도적으로 높게 나타나는 패턴이 나타나는군요.

좀 더 자세한 기상 정보를 얻으시려면 아래 링크에서 확인하실 수 있습니다.

<ppt 복귀>

성공적인 시연이었습니다.

이제 슬슬 결론을 지어보도록 하겠습니다.

제가 생각하는 프로그램 활용방안입니다.

트렌드 대시보드를 통해 기간별, 기상요인별 트렌드를 파악해볼 수 있고,

예측을 통해 월별 식중독 환자 수를 예측 해볼 수도 있습니다.

또한 평년값을 통해 앞으로의 예상되는 환자수를 파악하는 데에도 활용해볼 수 있습니다.

제가 생각하는 한계점들을 뽑아보자면, 적은 데이터 수와 한정된 기상 요인을 꼽을 수 있습니다.

이러한 한계점들은 수도권이나 전국으로 분석범위를 확대하거나, 다른 기상요인들을 추가 분석하는 것으로 극복해 볼 수 있을 것이라 생각합니다.

본 프로젝트의 핵심은 서울시의 월별 기상요인을 통해 식중독 발생 환자 수를 예측하는 프로그램을 구현하는 것이고,

이를 위해 데이터 수집부터 웹페이지 구현 및 모바일 최적화 작업까지 다양한 작업들을 수행하였습니다.

프로젝트를 진행하면서 많은 것들을 배울 수 있어 뿌듯했고,

다만, 만족할 만한 모델 성능이 나오진 않아 아쉽다고 느꼈습니다.

그럼 이상으로 발표를 마치도록 하겠습니다. 감사합니다.