

예방해요! 유행성질병 -국민건강알림서비스

국민건강보험공단



분석개요

- **추진배경 및 필요성**
 - 주요질병의 위험도 동향과 알림 제공 필요
- **추진목적**
 - 예측모델의 정확도를 향상시키고 사용자화면을 개선하여 정확하고 친절한 국민건강알람 서비스를 구축 하여 국민건강증진
- **성과**
 - 7개권역별로 동일한 기상변수 및 환경변수를 사용하는 예측모델을 개발하여 질병 별 변수 확인
 - 위험도 모델링을 통해 질병 별 위험도 구간을 세분화하여 위험도 구간 설정
- **활용방안**
 - 5개질병에 대한 실시간 전국 국민 건강알람서비스 제공

활용데이터

- **활용데이터**

데이터 종류	수집 채널	수집 방법	수집 간격
정형	국민건강보험공간	관리도구 FTP	일별
	식약처	관리도구	수시
	기상청	API	시간별, 일별
	국립환경과학원	API	시간별, 일별
비정형	트위터	API	일별
	블로그	API	일별
	뉴스	API	일별
	검색 트렌드	API	주별

분석모델링

• 분석 프로세스

정형데이터와 비정형데이터를 이용, 질병의 진료건수를 예측하는 모형으로 적합한 음이항회귀모형을 선택

-고도화알람서비스

진료건수에유의하게영향을미치는변수들(기상변수,환경변수, 소셜데이터)의 데이터를 지연없이 결합한 예측모형

○ 감기 진료발생건수 예측모델

$$y = \exp(8.890e+00 - 2.447e-03 \cdot \text{최저기온} + 5.481e-03 \cdot \text{일교차} + 6.264e-04 \cdot \text{습도} + 8.171e-04 \cdot \text{뉴스} + 5.290e-06 \cdot \text{트위터} + \text{의료급여} + \text{지역가중치})$$

• y = 예측진료건수

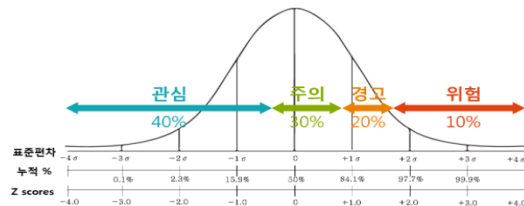
• e : 십의 자승을 의미합니다. (예: $1.0E+03 = 1000 / 1.0E-02 = 0.01$)

○ 감기 위험도 예측

예측된 질병 발생 건수를 과거 2년 평균 질병 발생 건수와 비교하여, 평균으로부터 벗어나 있는 정도를 기준으로 해당값의 분포도에 따라 위험도를 관심, 주의, 경고, 위험 4단계로 구분합니다.

단계	구간	구간 확률	누적 확률
● 관심	$-\infty \leq Z \leq -0.11$	40.00%	40.00%
● 주의	$-0.11 < Z \leq 0.58$	30.00%	70.00%
● 경고	$0.58 < Z \leq 1.14$	20.00%	90.00%
● 위험	$1.14 < Z \leq \infty$	10.00%	100%

• $Z = (X - \mu) / \sigma$ where μ = 평균, σ = 표준편차



분석결과

• 분석결과

- 전국 예측모델과 지역적유사성을 가지고 있는 7개권역별로 권역별로 동일한 기상변수 및 환경변수를 사용하는 예측모델을 개발
- 예측모델 구동결과 실제값과 예측값 사이에 0.9이상의 높은상관도 존재

전국기준

구분	감기	눈병	식중독	피부염	천식
상관도	0.912	0.97	0.9157	0.971	0.933

• 분석결과 활용 측면

- AS-IS : 질병특성 고려, 질병 별 3단계 위험도 구간 알림 서비스/알람 대상 질병 (9개)
- TO-BE : 정확하고 친절한 알람 서비스 구축
 1. 유관기관 협업
 2. 정형/비정형 데이터 보강
 3. 사용자 화면 개선
 4. 대상질병 확대 및 예측 정확도 향상

분석결과 및 시각화 예시

- 국민건강 알람서비스 홈페이지 화면



국민건강 알람서비스