|  |
| --- |
| COVID-19 Risk Analytics and Safe Activity Assistant with Clustering and Classification Algorithms  ABSTRACT  (abstract)  Keywords:  클러스터링 및 분류 알고리즘을 이용한 COVID-19 위험 분석 및 안전 활동 보조  요 약  요약 내용  키워드: |

# Introduction

# Related Work

관련 연구 내용

# Design of COVID-19 Safety Risk Assessment

COVID-19 안전 위험 평가는 COVID-19 데이터를 이용하여 감염자, 접촉자(감염자와 접촉한 사람) 혹은 장소에 대한 위험성을 계산하고 안전성에 대해 평가하는 것을 의미한다. 이 평가의 목적은 사회에서 완화되어가는 COVID-19 전염병의 위험도를 수치화하여 보여줌으로써 개인이 사회의 일원으로서 놓치고 있었던 막연한 경각심을 일깨워 방어적인 행동을 더 유발하게 해 공익에 좀 더 기여할 수 있는데 있다.

## Factors for Safety Risk Assessment

…

..

Age

Date

Disease

…

…

## Metrics for Safety Risk

COVID-19 안전 위험 평가는 두 가지 관점에서 접근할 수 있다. ISR(Individual Safety Risk, 개인 안전 위험)과 GSR(Group Safety Risk, 집단 안전 위험)이다. ISR은 개인이 현재 얼마나 위험한 상태인지 보여주는 척도다. 결과값의 범위는 [0, 1]이며 0이면 무결한 상태를 의미하고 1이면 매우 위험한 상태를 의미한다. GSR은 시, 도처럼 넓은 지역을 거시적으로 접근하거나, 건물 혹은 편의시설처럼 미시적으로 접근하여 그 구역이 얼마나 위험한 상태인지 보여주는 척도다. 결과값의 범위는 [0, 1]이며 0이면 무결한 장소를 의미하고 1이면 감염성이 매우 높은 상태를 의미한다.

이러한 COVID-19 안전 위험 평가를 계산하기 위한 요소는 크게 ISR을 위한 요소와 GSR을 위한 요소로 나뉜다. ISR을 위한 요소는 다음과 같다.

*Incurred\_date*는 접촉하거나 감염된 날을 의미한다. *Ageuser*는 개인의 현재 나이를 의미한다. *Ageuser*의 범주는 [0, 100]이며 100세 이상은 같은 계산식을 적용한다. *Diseaseuser*는 지병을 의미하며 범주는 [0, 1]이며 0은 COVID-19와 밀접한 관련이 없는 지병을 가지고 있거나 건강한 상태를 의미하며 1은 COVID-19에 감염될 경우 매우 치명적일 수 있는 지병을 가지고 있음을 의미한다.

Group에 대한 설명

Group의 종륲

*Severitypeople*은 해당 지역, 건물 혹은 편의시설에 있었던 사람의 위험도를 나타낸다. *Severitypeople*의 범주는 [0, 1]이며 0이면 건강한 상태를 의미하며 1에 가까울수록 그 사람이 얼마나 위독한 상태임을 의미한다. *Covid\_Statuspeople*은 해당 지역, 건물 혹은 편의시설에 있던 사람이 COVID-19와 관련된 상태인지를 나타낸다. 범주는 무결함, 접촉함 또는 감염됨 세가지가 있다. *Density*는 해당 지역, 건물 혹은 편의시설의 사람들의 밀집도를 나타낸다. 범주는 [0, 1]이며 0에 가까울수록 공간이 넓거나 사람이 적음을 의미하고 1에 가까울수록 비좁은 공간에 사람이 많다는 것을 의미한다.

마지막으로 COVID-19 안전 위험 평가에 사용될 요소들에 적용할 가중치가 있다. 클러스터링 알고리즘을 사용할 때 거리 함수(Distance Function)를 사용하게 된다. 이때 거리 함수에 각 요소별로 가중치를 다르게 줘서 COVID-19 안전 위험 평가에 요소가 미치는 영향력을 조절할 수 있다. 또한 각 가중치의 합을 1로 고정함으로써 요소가 미치는 영향력의 비율을 조절할 수 있다.

Weight 값들 고려한 계산 공식 추가

# Clustering of People and Groups

## Clustering People using Factors

클러스터링이란 거리 함수를 이용하여 주어진 데이터들 간의 거리를 계산하고 여러 개의 클러스터를 생성하여 데이터를 클러스터별로 분류하는 것을 의미한다. 즉, ISR을 이용한 클러스터링이란 ISR Factor Set을 이용하여 클러스터링하는 것을 의미한다.

People Clustering 필요성

Clustering Alogoirhm

Clustering Results

Clustering 결과의 활용

## Clustering Groups using Factors

Group Clustering 필요성

Clustering Alogoirhm

Clustering Results

Clustering 결과의 활용

# Design of Safe Activity Assistant

Query processing

….

….

# Experiments and Assessment

# Conclusion

References

1. …
2. …
3. …