

<https://doi.org/10.7236/JIIBC.2024.24.5.177>
JIIBC 2024-5-25

노인 고독사 방지를 위한 빅데이터 기반 고독사 고위험 지역 탐지 연구

A Big Data Analysis to Prevent Elderly Solitary Deaths by High-risk Area Clusterization

김소연*, 김수형**, 이봉규***

Soyon Kim*, Soo Hyung Kim**, Bong Gyou Lee***

요약 본 연구는 서울시 노인 고독사 방지를 위해 고독사 고위험 지역을 탐지하기 위한 빅데이터 기반 분석 방법을 제안한다. 본 연구에서는 고독사 위험 요인을 인구학적, 건강학적, 경제적, 사회환경적 요인으로 분류하여 분석하였다. 서울시 열린데이터 광장과 공공데이터 포털에서 수집된 데이터를 이용하여 변인을 생성하고 K-means 클러스터링을 사용하여 산포도를 작성한 후 지도를 통해 시각적으로 구현하였다. 분석 결과 중랑구, 강북구, 노원구, 은평구, 강서구, 관악구가 가장 고위험 지역으로 나타났다. 본 연구에는 기존 설문조사 기반 연구의 한계를 빅데이터 분석을 통해 보완하였다는 함의가 존재한다. 연구결과가 고독사 예방 사업의 효율성을 높이고 자치구별로 적절한 예산 배정을 위한 의사결정 기반 자료로 활용되기를 기대한다.

Abstract This study proposes a big data-based analytical method to detect high-risk areas for solitary deaths among the elderly in Seoul. The study categorizes and analyzes the risk factors of solitary deaths into demographic, health, economic, and socio-environmental factors. Using data collected from the Seoul Open Data Plaza and Public Data Portal, variables were generated and scatter plots were created using K-means clustering, followed by visual implementation through map creation. The analysis identified Jungnang-gu, Gangbuk-gu, Nowon-gu, Eunpyeong-gu, Gangseo-gu, and Gwanak-gu as the highest-risk areas. This study addresses the limitations of previous survey-based research through big data analysis. The findings are expected to enhance the efficiency of solitary death prevention programs and serve as a basis for informed decision-making in budget allocation across districts.

Key Words : Elderly Solitary Death, High-Risk Area Detection, Big Data Analysis, Seoul

*정회원, 연세대학교 정보대학원

**준회원, 연세대학교 정보대학원

***정회원, 연세대학교 정보대학원(교신저자)

접수일자 2024년 7월 12일, 수정완료 2024년 8월 3일
게재확정일자 2024년 10월 4일

Received: 12 July, 2024 / Revised: 3 August, 2024 /

Accepted: 4 October, 2024

***Corresponding Author: bglee@yonsei.ac.kr

Graduated School of Information, Yonsei University, Korea

I. 서 론

보건복지부에서 2022년에 조사한 고독사 실태조사에 따르면 고독사 사망자 수는 2017년부터 최근 5년간 매년 증가 추세에 있으며 2021년에는 3,378명에 이를 정도로 중요한 도시문제로 제기되고 있다. 이 중 서울은 해당 기간 고독사가 가장 많이 발생한 지역 2위이며, 2021년 서울시 고독사 사망자 중 60대 이상이 절반 가량의 비율을 차지할 정도로 노인 인구에 대한 예방 대책이 시급한 것으로 확인되었다^[1]. 고연령 독거노인의 경우 다양한 신체, 건강, 사회, 경제, 심리적 제약이 존재하여 다차원적으로 취약한 환경에 놓여있다^[2]. 해당 문제에 대응하기 위해 정부에서는 공모를 통해 고독사 예방·관리 시범사업을 추진 중이다. 이는 지역 내 고독사 위험자를 발굴한 후 안부 확인, 생활 지원, 정신 및 심리 지원, 사후관리 등의 서비스를 선택적으로 지원하는 사업을 말한다^[3]. 서울시에서는 사회적 관계망이 취약한 고위험 1인가구에게 IoT 스마트 플러그를 설치하여 안전과 건강을 확인하는 서비스를 실시하고 있다. 스마트 플러그 서비스란 플러그를 통해 생활 활동을 감지 및 분석하여 일정 기간 변화가 없을 경우 동 주민센터 복지플래너에게 위험 신호를 전달하는 서비스를 말한다. 서울시에서는 독거노인의 건강 및 안전 관리 지원을 위해 서비스를 지속적으로 확대하고자 계획하고 있다^[4]. 정부 예산과 인력은 한정적이기 때문에 모든 지역에 동시에 예방 서비스를 시행할 수 없으므로 고독사 고위험 지역을 우선적으로 선별하여 해당 지역부터 순차적으로 서비스를 확대할 필요가 있다. 현재 서울시에서는 현장에서 지역 내 독거노인을 대상으로 수요조사를 실시하거나 담당자의 추천을 통해 대상자를 선정한다^[4]. 그러나 이와 같은 방식으로는 해당 서비스가 반드시 필요한 지역을 객관적으로 파악하기가 어렵다는 한계가 존재한다.

본 연구는 이와 같은 노인 고독사 방지 대책에 도움이 되기 위한 분석 방법을 제안하는 것을 목적으로 한다. 본 연구에서는 서비스를 우선적으로 실시하기 위한 고독사 고위험 지역을 선정하기 위해 서울시에서 제공하는 빅데이터를 기반으로 고독사 고위험 지도를 작성하고 위험도가 높은 지역을 탐지하는 방법을 제안한다. 이를 통해 서비스 우선 실시 및 예산 배분이 필요한 지역을 객관적으로 탐지할 수 있도록 하여 한정된 예산과 인력을 효율적으로 사용하는 데에 기여할 수 있을 것으로 기대된다.

II. 이론적 배경

고독사 예측을 위해서는 먼저 고독사 위험 요인 발굴이 선행되어야 한다. 선행연구에서는 고독사 위험 요인을 도출하기 위하여 여러 변수를 고려하여 위험군의 유형화를 수행하고 있다. 송문선은 수집을 통해 얻은 비정형 데이터에 기반하여 독거노인 고독사 요인을 다섯 가지의 영역별 요인으로 범주화한다. 이는 인구사회학적 요인, 경제적 요인, 건강적 요인, 사회적 요인, 개인적 요인의 다섯 가지로 분류된다. 인구학적 요인에는 연령과 독거 유형, 복지급여 형태와 액수가 포함된다. 건강학적 요인은 신체적 건강과 정신적 건강으로 분류된다. 경제적 요인은 자산 및 부채, 주택, 직업 등의 요인이 포함된다. 사회적 요인에는 사회관계와 주거, 문화 및 정보 접근 등의 요인이 해당한다. 마지막으로 개인적 요인은 개인의 성향을 의미한다. 분석 결과 한국의 경우 인구학적 요인에서는 성별은 남성, 연령대는 65세 이상, 독거 형태로 살아온 기간이 길수록 고독사의 위험이 높았다. 건강학적으로는 만성질환 수가 높고 질병이 있을 경우, 경제적으로는 절대적 빈곤을 경험할 경우 고독사 위험군에 해당한다. 사회적 요인으로는 사회적으로 고립되어 있으며, 소규모 주택이나 임대 아파트 및 월세가 낮고 폭염 및 한파 취약지구에서 위험군의 가능성이 높았다. 마지막으로 개인적 요인으로는 사회적 관계가 무력화되었을 경우가 해당했다^[5].

정순돌 외에 따르면 대표적으로 고독사 위험 수준에 따라 분류하는 지자체로 서울시와 부산시가 존재한다^[6]. 고독사 위험군 분류는 일반적으로 경제적 요인과 가구 취약성, 사회적 고립의 특성에 기반하여 유형화된다^[7, 8]. 정순돌 외는 이와 같은 해당 유형을 전국적으로 확대하고자 사회적 고립 및 가구 취약성, 사회인구학적 특성을 영향 요인으로 선정하고 설문조사 결과를 다항로지분분석하여 최종적으로 위험군에 영향을 미치는 요인을 도출해냈다. 결과적으로 사회인구학적 특성 중에는 연령(60~70대), 성별(남성), 주관적 경제 수준, 주관적 건강 수준, 일상생활에 제한이 될 정도의 건강, 배우자 없음, 1인 가구의 요인이 유의하게 나타났다^[6]. 이러한 위험 요인은 보건복지부의 고독사 실태조사에서도 동일하게 포함되었다^[3].

선행연구에서는 데이터에 기반하여 위험 요인을 도출해내었으나 분석을 위해 인간의 주관성이 개입될 수 있는 설문조사 방식만을 사용했다는 한계가 존재한다. 또한 설문조사는 예산과 시간이 한정되어 있어 대규모의

지역을 대상으로 분석하기 어렵다. 본 연구에서는 선행 연구에서 도출된 요인을 객관적인 빅데이터에 적용함으로써 서울시 전체를 대상으로 과거의 위험 지역을 탐지하고 앞으로의 위험 지역을 분류하는 방식을 제안하고자 한다.

III. 연구방법

본 연구에서는 선행연구에서 분석된 고독사 위험 요인을 참고하여 인구학적, 건강학적, 경제적, 사회·환경적 요인으로 분류하여 취약 요인을 분석하고 해당하는 데이터를 서울시 열린데이터 광장과 공공데이터 포털에서 수집하였다^{9, 10}. 서울시 열린데이터 광장과 공공데이터 포털의 경우 공공 데이터 개방을 통해 시민의 적극적 활용과 빅데이터 가치 창출을 모색하고 있다¹¹.

먼저 인구학적 요인으로는 65세 이상의 고령 인구나 남성이 고독사에 더 취약함이 확인되었다. 고독사 사망자 수는 매해 여성 대비 남성이 4배 이상 많으며, 60세 이상 고령층이 서울시 고독사 인구의 절반을 차지한다^{11, 5, 7}. 경제적 요인으로는 저소득 고령인구와 국민기초생활수급권자의 경우 높은 고독사율을 보이는 것으로 나타났다. 절대적 빈곤 경험자의 경우 장기적 경제 스트레스로 이어져 고독사로 연결되며, 복지급여 영역의 기초생활보장 수급자의 경우 기존의 연구에서 높은 고독사율이 확인되었다⁵. 건강학적 요인으로는 의료 서비스에 멀어질수록 만성질환, 의료적 관리가 어려워 고독사율이 높을 것으로 보았다^{3, 5, 6, 12}. 사회적 요인으로는 단독, 다세대, 연립, 빌라를 포함한 주택에서의 고독사 발생률이 절반 이상을 차지함에 따라 다세대 주택 밀집 지역의 경우 고독사 위험군으로 판단되었다³. 환경적 요인으로는 폭염취약지구와 난방이 부족한 한파취약지구의 경우 위험군으로서의 가능성이 높았다⁵. 해당 지구를 도출하기 위해 서울시에서 제공하는 S·DoT 빅데이터에서 추출한 환경정보 중 기온·풍속 데이터를 기초로 체감온도를 계산하였다⁷. 기상청 기상특보 발표기준에 따라 폭염(일 최고 체감온도가 33℃ 이상인 날) 및 한파(최저기온이 -12℃ 이하인 날)의 빈도가 높을수록 취약하다고 정의하였다.

이렇게 정의된 요인을 기초로 고독사 고위험 점수 변인을 생성하였다. 수집 기간은 모두 2021년 1월부터 12월까지이며 수집된 데이터를 정리하면 표 1과 같다.

표 1. 고독사 취약 요인 데이터

Table 1. Data on Factors Contributing to Solitary Deaths

요인	데이터명	설명	수집처
인구학적	서울시 독거노인 연령별 현황 통계	집계구별 65세 이상 독거노인 거주 건수 통계	서울 열린 데이터 광장
	서울시 1인 가구 성별 현황 통계	집계구별 남성 노년 1인 가구 거주 건수 통계	
경제적	서울시 독거노인 저소득 현황 통계	집계구별 저소득 독거노인 거주 건수 통계	서울 열린 데이터 광장
	서울시 국민기초생활보장 수급자 통계	자치구별 국민기초생활수급자 거주 가구수 통계	
사회 환경적	서울시 주택종류별 주택 통계	집계구별 단독, 다세대, 연립, 빌라 위치 현황 통계	공공 데이터 포털
	S·DoT 환경정보	서울시 행정동별 기온·풍속 데이터	
건강학적	전국 병의원 현황 데이터	시군구별 병·의원·약국 위치 데이터	공공 데이터 포털

고독사 취약 요인 데이터의 경우 S·DoT 환경정보와 병·의원 위치정보에만 전처리가 필요하였다. 환경정보를 통해 폭염 및 한파일수를 계산하였으며, 위치정보의 기존 행정동 단위를 자치구 단위로 변환하여 전처리를 수행하였다. 이렇게 처리된 고독사 취약 요인 데이터를 기반으로 고독사 고위험 점수 변인을 생성하였으며 변인 생성 시 가중치를 주기 위해 선행연구에 기반하여 인구학적, 경제적, 사회적, 건강학적, 환경적 순으로 점수를 부여하였다.

다음으로 실제 노인 고독사 발생률 데이터를 수집하여 자치구별 실제 노인 고독사율 변인을 생성하였다. 수집된 데이터는 표 2와 같다.

표 2. 실제 노인 고독사 발생률 데이터

Table 2. Actual Elderly Solitary Death Rate Data

데이터명	설명	수집처
서울시 자치구별 사망자 수 통계	자치구별 전체 사망자 수 합계	서울 열린 데이터 광장
서울시 노인 고독사 발생 현황 통계	자치구별 65세 이상 고령 인구 고독사 발생 건수	서울시 복지재단
서울시 고독사 발생 현황 통계	자치구별 고독사 사망자 수 통계	

수집된 데이터를 바탕으로 고령인 고독사 수를 전체 사망자 수로 나누어 노인 고독사율 변인을 생성하였다. 서울시 고독사 발생 현황 통계의 경우 결과 평가를 위해 활용하였으며, 21년도 데이터가 공개되어 있지 않아 20년도 1월부터 12월까지의 데이터를 수집하였다⁸.

분석 절차는 다음의 순서로 진행되었다. 전처리가 완료된 데이터를 이용하여 요인별 단계 구분도를 생성한 후, 요인별 가중치에 따른 고위험 점수를 부여하여 고위험 점수 변인을 생성한다. 다음으로 K-means 클러스터링의 군집분석을 수행하여 자치구별 실제 고독사율 변인 대비 고위험 점수 변인의 산포도를 작성한다. 분석된 각 군집에 HH에서 LL까지의 위험순위를 부여한다. 분석 결과를 이용하여 고독사 고위험 지도를 구현한 후 최종적으로 위험 지역을 탐지하여 제시하였다. 분석 절차를 도식으로 정리하면 그림 1과 같다.

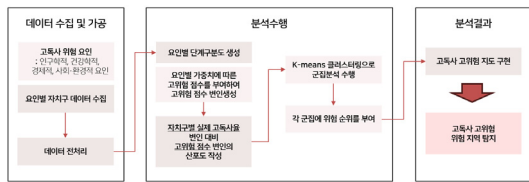


그림 1. 분석 절차
Fig. 1. Analysis Procedure

IV. 실험 및 결과

K-means 클러스터링을 수행하기에 앞서 최적의 K 값을 선정하기 위해 엘보우 메소드(Elbow method)와 실루엣 계수(Silhouette coefficient)를 사용하였다. 엘보우 메소드는 클러스터 개수별로 각 클러스터의 가장 가운데에서부터 각 포인트의 거리를 제곱합해서 나온 수치를 비교하여 적절한 K의 수를 찾는 방법이다^[14]. 실루엣 계수는 각 군집 간의 거리를 측정하고 분리된 거리의 효율성을 계산하여 가장 적절한 K의 수를 선정한다^[15]. 선정된 K의 수는 4개이므로 고독사 고위험 점수 변인과 노인 고독사율 변인에 따라 4개로 군집분류한 K-means 클러스터링 결과는 그림 2와 같이 나타났다.

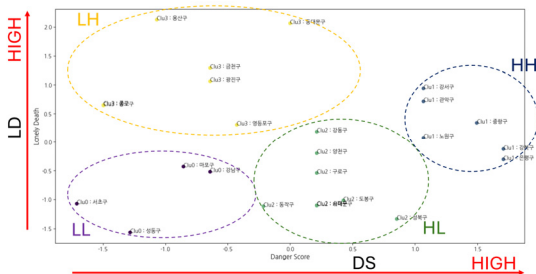


그림 2. K-means 클러스터링 결과
Fig. 2. K-means Clustering Results

가장 고위험지구인 HH에는 중랑구, 강북구, 노원구, 은평구, 강서구, 관악구가 포함되는 것으로 나타났다. 결과 평가를 위하여 2020년도 자치구별 실제 고독사 수의 순위를 매긴 결과를 고독사 고위험 지도와 비교해보았다. 결과적으로 상위 6개의 자치구 중 동대문구를 제외하고 HH의 5개 구가 포함됨을 확인하였다. 클러스터링 결과를 지도로 시각화하여 나타내면 그림 3과 같다.

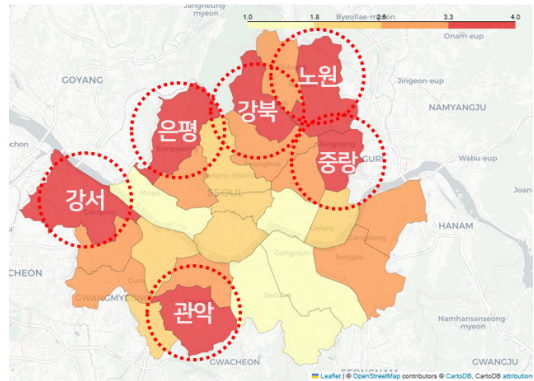


그림 3. 고독사 고위험 서울시 지도
Fig. 3. High-Risk Solitary Death Map of Seoul

V. 결론

본 연구에서는 노인 고독사 방지를 위하여 빅데이터를 이용하여 서울시 고독사 고위험 지역을 탐지하는 방법을 제안하였다. 대표적인 고독사 예방사업인 스마트플러그 사업은 현재 실태조사와 주민 신고 등을 통해 발굴된 고독사 위험 1인 가구를 대상으로 시행되고 있다. 그러나 본 연구에서 사용된 분석을 이용하면 자치구별로 적절한 설치 예산을 배정하기 위한 의사결정의 기반 자료로써 활용 가능하다. 서울시에서는 2022년도 당시 스마트플러그의 2천 개 추가 설치를 위해 2억 원의 예산을 배정하였으며 해당 예산은 스마트플러그 확대 운영과 관련된 25개 자치구를 대상으로 수요조사를 실시하여 분배되었다. 그러나 본 방법을 활용하여 참여의사를 밝힌 16개의 자치구 중 HH로 분류된 5개 자치구에 우선적으로 예산을 배분하고, 사업 참여를 검토 중이었던 9개의 자치구 중 HH로 분류된 1개의 자치구에 참여를 독려함으로써 예산과 인력을 보다 효율적으로 분배할 수 있을 것이다. 이와 같은 방법을 통해 수요조사를 위한 예산을 절감할 수 있으며, 객관적인 고독사 고위험 지도에 근거하여 독거노인 고독사 예방을 위한 빅데이터 기반 체계적 관리

방안을 수립할 수 있다는 연구의 실무적 함의가 존재한다.

본 연구의 학술적 함의는 다음과 같다. 설문조사만을 사용하여 탐구하였던 기존 연구의 한계를 빅데이터 기반 분석 방법을 활용하여 보완할 수 있었다. 이를 통해 노인 고독사 방지를 위한 연구의 외연을 보다 객관적인 방법을 통해 확장하였다는 의의가 존재한다.

본 연구에는 분석 대상을 자치구 단위에 한정하였다는 한계가 존재한다. 향후에는 행정동 단위의 지도를 만들어 구에서 활용할 수 있는 기반 자료를 생성할 수 있도록 추가 연구를 수행하고자 한다. 또한 2021년도 기준 데이터만을 분석에 활용하였다는 한계가 존재하는데 향후 지속적으로 추가 연도 데이터셋을 확보하여 고독사 고위험 지역 예측 분류 모델을 구현하는 연구로 확장하고자 한다.

References

- [1] Ministry of Health and Welfare, "Announcement of 2022 Solitary Death Survey Results," Ministry of Health and Welfare.
https://www.mohw.go.kr/react/al/sal0301vw.jsp?PAR_MENU_ID=04&MENU_ID=0403&page=1&CONT_SEQ=374084.
- [2] S. W. Kim, H. O. Park, and S. Yoon, "A Study on Changes in Daily Life and Satisfaction of the Elderly Living Alone Using IoT Sensor-based Care Service," *Journal of Korea Academia-Industrial cooperation Society*, vol. 24, no. 5, pp. 155-163, 2023.
DOI: <https://doi.org/10.5762/KAIS.2023.24.9.155>.
- [3] Ministry of Health and Welfare, "Pilot Project for Solitary Death Prevention and Management Starts in August," Ministry of Health and Welfare.
https://www.mohw.go.kr/board.es?mid=a10503010100&bid=0027&act=view&list_no=372433&tag=&nPage=1.
- [4] Seoul Metropolitan Government, "Deployment of 2,000 'Smart Plugs' to Prevent Solitary Deaths in Households with No Electricity Usage," Seoul Metropolitan Government.
<https://news.seoul.go.kr/welfare/archives/542803>.
- [5] S. M. Sun, "Establishment of Data-base Prediction Indicators for Risk Group of Elderly Living Alone," *Journal of the Korean Gerontological Society*, vol. 42, no. 2, pp. 273-303, 2022.
DOI: <https://doi.org/10.31888/JKGS.2022.42.2.273>.
- [6] S. Chung, J. Yoo, M. Jo, and J. Jung, "Characteristics of Vulnerable Groups at Risk of Lonely Deaths in South Korea," *Korean Journal of Social Welfare*, vol. 75, no. 3, pp. 397-420, 2023.
DOI: <https://doi.org/10.20970/kasw.2023.75.3.014>.
- [7] The Seoul Institute, "Diagnosis of Heatwave Vulnerability in Residential Areas of Seoul Using S-DoT Big Data," The Seoul Institute, Data Insight Report 06,
<https://data.si.re.kr/data-insight-report/65208>
- [8] I. J. Song and E. J. Mo, "A Study on the Current Status of Risk Groups for Solitary Deaths in Seoul," Seoul Welfare Foundation, Research Report.
<https://wish.welfare.seoul.kr/swflmsfront/board/boar dr.do?bmnno=10015&opno=10005&bno=94004>.
- [9] D. I. Lee and J. Chun, "A Study on the Current Status and Countermeasures of Deaths in Busan City," *Hangdo Busan*, vol. 41, pp. 531-565, 2021.
DOI: <https://doi.org/10.19169/hd.2021.2.41.531>.
- [10] Seoul Open Data Plaza, "Public data," Seoul Open Data Plaza.
<https://data.seoul.go.kr/dataList/datasetList.do?datasetKind=1&searchFlag=M>.
- [11] Open Data Portal, "Data list," Open Data Portal.
<https://www.data.go.kr/tcs/dss/selectDataSetList.do>.
- [12] H. K. Noh, S. Y. Park, S. Y. Hwang, D. J. Shin, Y. S. Lee, J. J. Kim, and K. Park, "Big Data Analytic System Based on Public Data," *Journal of the Institute of Internet, Broadcasting and Communication*, vol. 20, no. 5, pp. 195-205, 2020.
DOI: <https://doi.org/10.7236/IIBC.2020.20.5.195>.
- [13] H. Kim, K. Kim, and J. Y. Kang, "A Study on Impact of Medical Facility Accessibility on the Elderly Population Growth: A Comparative Analysis using Regression Analysis and Machine Learning Techniques," *Journal of Korean Institute of Information Technology*, vol. 20, no. 5, pp. 113-121, 2022.
DOI: <https://doi.org/10.14801/jkiit.2022.20.5.113>.
- [14] J. Y. Jeon, J. S. Choi, and H. G. Byun, "Implementation of Elbow Method to Improve the Gases Classification Performance Based on the RBFN-NSG Algorithm," *Journal of Sensor Science and Technology*, vol. 25, no. 6, pp. 431-434, 2016.
DOI: <https://doi.org/10.7236/1225-5475/eISSN2093-7563>.
- [15] J. S. Ahn, J. W. Kim, J. Yoo, and D. J. Choi, "Prediction of Public Bicycle Usage in Changwon City Using Cluster Characteristics and Deep Neural Networks," *The Journal of the Korea Contents Association*, vol. 23, no. 4, pp. 159-170, 2023.
DOI: <https://doi.org/10.5392/JKCA.2023.23.04.159>.

저 자 소 개

김 소 연(정회원)



- 2014년 8월 : 성균관대학교 정치외교학과 (정치학 학사)
- 2022년 8월 : 성균관대학교 대학원 정치외교학과 (정치학 석사)
- 2022년 9월 ~ 현재 : 연세대학교 정보대학원 박사과정
- 주관심분야 : 생성형 AI, LLM, 빅데이터 분석, 마이데이터

김 수 형(준회원)



- 2018년 9월 : University of California, Davis Dept. of Computer Science (BS)
- 2024년 3월 ~ 현재 : 연세대학교 정보대학원 석사과정
- 주관심분야 : 빅데이터 분석, 딥러닝, Transformer, Forecasting

이 봉 규(정회원)



- 1988년 : 연세대학교 상경대학 (경제학 학사)
- 1992년 : Cornell University, Dept. of CRP (MS)
- 1994년 : Cornell University, Dept. of CRP (Ph.D)
- 2005년 ~ 현재 : 연세대학교 정보대학원 교수
- 주관심분야 : 플랫폼 비즈니스 전략, ICT, Digital Transformation, 빅데이터