

국토정책 Brief

국토연구원에서 수행한 주요 연구과제의 핵심 내용과 정책제안 등을 압축해 국민께 알려드리고자 하는 발간물입니다.

2023. 1. 16.
No. 900



발행처 국토연구원
발행인 강현수
www.krihs.re.kr

장요한 국토연구원 국토데이터랩 팀장
이영민 국토연구원 부연구위원
임시영 국토연구원 부연구위원
박서현 서울시립대학교
도시빅데이터융합학과 연구원
김현경 국토연구원 연구원

감염병 대응을 위한 딥러닝 기반 목적별 통행수요 예측모형 개발 연구

주요 내용

- ① 실시간 유동정보(관측교통량)로부터 기종점별·목적별 통행수요 예측모형 개발
- 전국도로에 매설된 차량검지기 관측교통량 정보로부터 출발지와 도착지를 추정하는 수리모형과 다양한 머신러닝 및 딥러닝을 이용한 AI 활용모형 개발
- ② 감염병 확산 전후로 확인되는 시공간(일별·지역별) 데이터 패턴 등을 딥러닝 모형과 목적별 통행수요 예측모형을 결합하여 고도화
- 한국기업 데이터, 가계신용 데이터, 전국 사업자 주소록, 인구, GIS 네트워크, 정부의 감염병 대응 정책, 유동인구 데이터 등 다양한 공공 및 민간 빅데이터의 융합을 통한 공간 단위 분석과 딥러닝 기반의 목적별 통행수요를 예측
- ③ 통행과 통행목적 간 인과관계 이해와 포스트 감염병 시기 변화된 통행목적의 파악을 통해 포스트 감염병에 대응하기 위한 정책의 근거 자료 제시
- 다양한 머신러닝 및 딥러닝 모형을 정책연구에 적용한 국정운영의 과학화
- 모형의 민감도 분석을 통한 감염병 확산 이후 바뀐 지역별·목적별 통행수요를 파악

활용 방안

- ① 실시간 목적별 통행수요 변화와 감염병 방역 정책 전개에 따른 목적별 통행수요 영향 모니터링을 위한 과학적 정책지원 분석틀 마련
- ② 다양한 공공 및 민간 빅데이터를 시간과 공간의 축을 중심으로 융복합하여 다목적으로 활용할 수 있는 데이터 활용 사례 제시
- ③ 공공 인프라를 이용한 국토모니터링 가능성 제시 및 기타 범주에서의 추가 확장성 및 효용성 제고
- ④ 과학적 정책지원 도구를 이용한 감염병 확산의 영향을 공간적으로 파악하고, 이를 이용한 지역 맞춤형 방역 정책지원의 근거 자료로 활용

01. 포스트 감염병 대응을 위한 목적별 통행수요 예측모형 개발 의의

통행의 인과관계는 고려하지 않고, 공간 밀집도만을 조절하는 일원화된 방역 정책은 부작용을 초래

감염병 발생 초기, 정부 방역 조치가 즉각적인 호흡기 감염병 확산 방지에는 다소 효과가 있었던 것으로 추정되나, 지역 특성이 고려되지 않은 단일화된 방역 조치는 지역경제 및 내수시장 악화로 이어지는 등 부작용이 발생

- 신종 호흡기 감염병의 높은 전파력은 접촉을 통해 빠르게 감염되고 확산한다는 특징이 있어, 정부는 사람들이 모이는 공간 중 개인 방역 취약 시설을 중심으로 감염병 확산 방지 정책을 전개함
- 이동에 관한 인과관계 고민보다 이동 후 접촉 최소화를 위한 활동공간(밀도) 억제에만 집중한 정부 방역 정책은 일괄적으로 추진되고 장기화됨에 따라 방역 피로감과 풍선효과 등 부작용이 뒤따름

공간 밀집도를 강제적으로 조절하기보다는 ‘왜 그 공간으로 이동하여야 했는지’에 대한 이해가 필요

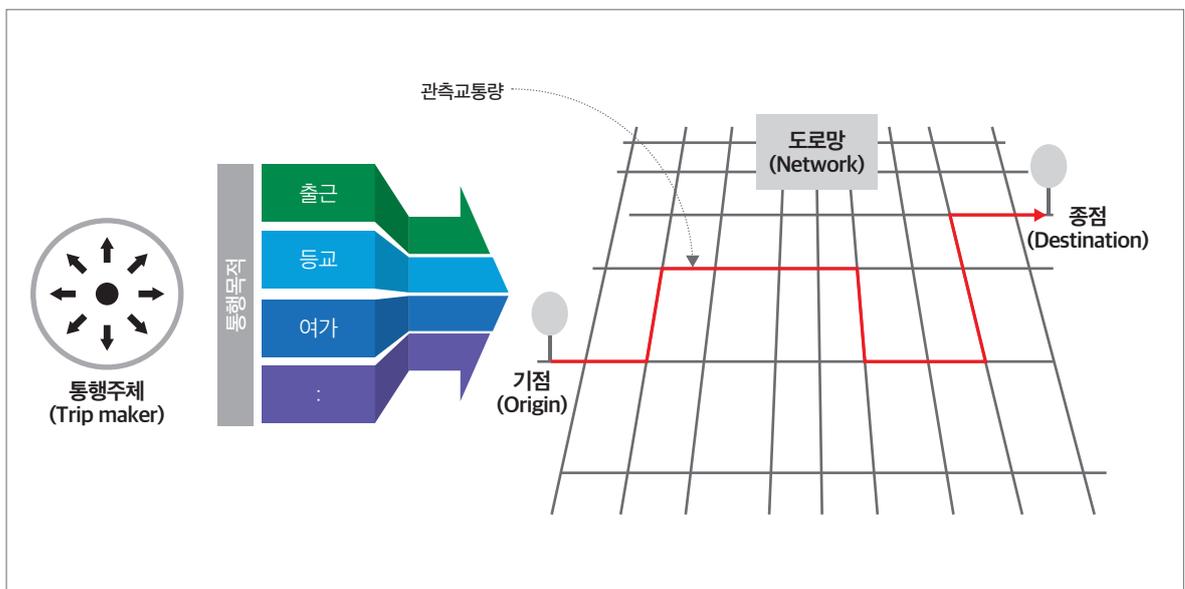
통행목적에 대한 이해는 현재 통행뿐만 아니라 장래 통행패턴 예측에 핵심 정보가 될 수 있다는 점에서 매우 중요

- 통행주체(Trip maker)는 업무, 등교, 쇼핑, 친교 등 다양한 통행목적에 따라 통행을 발생
- 전통적인 통행수요 파악 방법은 가구통행실태조사 등이 가장 대표적이거나, 5년에 한 번씩 이루어지기 때문에 감염병 등 시의성 높은 현안에 즉각 적용하기에는 한계가 있음
- 감염병 확산 이후, 구글(Google)과 애플(Apple)에서는 주기적 이동성 보고서를 통해 통행의 종점 특성에 따른 이동성 모니터링을 제공 중

공간 밀집도를 강제적으로 조절하는 방법보다는 공간 간 이동이 이루어진 통행의 목적에 대한 원인을 파악하는 방역 추진 방식은 지역 맞춤형 정책을 전개하고 방역 정책의 부작용도 완화할 것으로 기대

- 사회적 거리두기 정책 이후 달라진 통행패턴 파악 등 통행량이 증감한 주요 지역과 해당 지역의 공간적인 특성을 연계하여 살피는 등 데이터 기반의 접근 방법이 필요
- 관측교통량 데이터 등 세밀하고 실질적인 이동에 근거한 데이터와 이에 상응하는 분석 방식이 지역별 분석과 선제적 감염병 대응 정책 마련에 현실적일 것으로 기대

그림 1 통행주체의 통행목적별 기중점 통행과 도로의 관측교통량 등의 개념 도식화



주: 통행주체는 다양한 통행목적에 따라 통행을 결정하고, 기중점 간 통행이 드러나는 도로망 상의 관측교통량은 통행주체의 이동 이유를 추론해 볼 수 있는 실마리가 됨.

자료: 저자 작성.

02. 딥러닝 기반 목적별 통행수요 예측모형 개발

다양한 공공 및 민간 빅데이터를 수집, 가공

통행주체가 어떤 공간으로 이동한 이유를 통행수요 유발요인으로 간주, 이를 추론하기 위한 작업으로서 사회, 경제, 인구 등 경제활동 정보에 해당하는 다양한 데이터를 선별하고 지표로 활용하여 지역의 매력도를 추론

- 인구 관련 정보는 통계청에서 확인되는 지역별·성별·연령별 인구정보를 활용
- 지역의 특징 파악을 위해 디비리아(DBRIA)에서 제공하는 전국 422만 개의 업종별 지역 사업자 분포정보를 활용
- 지역의 기업 관계 분석을 위해 한국기업데이터(Korea Enterprise Data, KED)에서 제공하는 650만 건의 전국 사업체 및 거래관계망 정보를 활용
- 지역의 경제활동 분석을 위한 자료로, 가계신용 빅데이터인 코리아크레딧뷰로(Korea Credit Bureau, KCB)에서 제공하는 전국 소득·소비·부채와 관련된 데이터를 활용
- 지역 간 목적별 통행 정보와 교통네트워크 분석을 위해, 국가교통데이터베이스(Korea Transport DataBase, KTDB)의 기종점통행량 여객자료와 교통네트워크 분석용 도로망 자료를 활용
 - 국가교통데이터베이스에서는 시군구 및 읍면동을 교통분석존(Traffic Analysis Zone, TAZ)으로 구별하여 출근, 등교, 업무, 쇼핑, 귀가, 여가/오락/친지방문, 기타 등 7가지 통행목적별 수요를 구분하여 제공
- 지역 간 목적별 통행 정보를 위해, SDC 통계청의 통신 모바일 인구이동량 통계 데이터를 이용하여 목적별 통행량의 참값으로 같음하여 활용
 - 특정 지역으로 이동한 유동인구를 10세 단위 연령대별 및 일 단위로 집계한 데이터로, 해당 통계에서 고려하는 입지유형은 상업지역, 관광지, 대형아웃렛(쇼핑), 사무지역, 레저/스포츠/여가, 주거지역/기타 등으로 구별하여 제공
- 교통량정보제공시스템(Traffic Monitoring System, TMS)에서 전국 3,800여 곳의 상시 및 수시 교통량 조사지점과 1시간 단위의 방향별 교통량정보를 활용
- 감염병 확산 및 사회적거리두기 단계 정보는 손재선 외(2021)에서 제시한 데이터 취합방식을 이용하여 활용

표 1 모형에서 활용된 데이터와 항목 및 출처

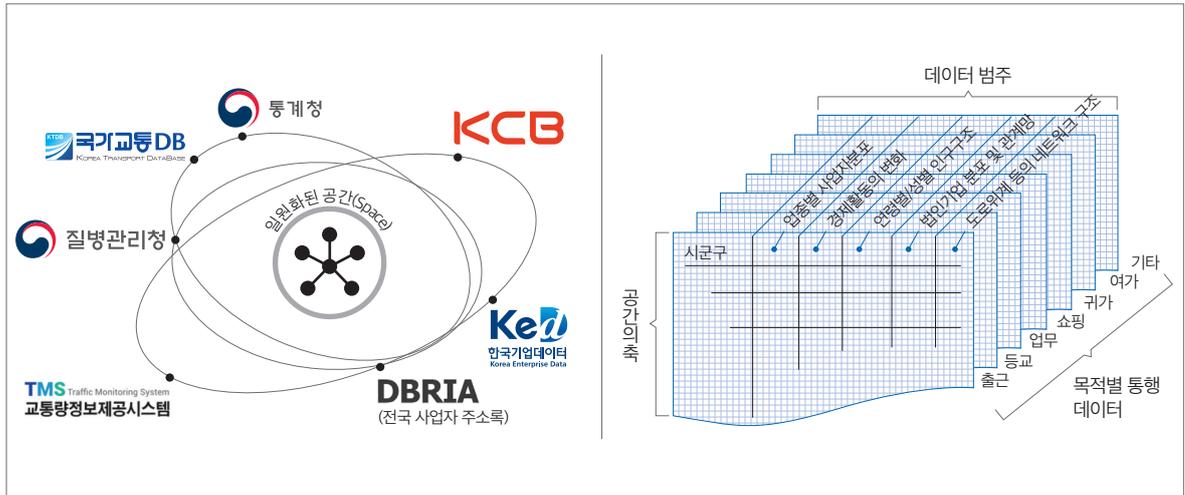
구분	데이터 항목	목적	출처	공공/민간
정적(static) 데이터	목적별 통행수요	목적별 수요 분석	국가교통데이터베이스(KTDB)	공공
	GIS 도로망	교통망		
	수시 교통량	공간의 수요 유발요인 분석		
	성별·연령별·지역별 인구	공간 요인 및 특성 분석	통계청	민간
	전국 사업자주소 데이터		디비리아(DBRIA)	
	기업 거래관계 데이터		한국기업데이터(KED)	
	소득·소비·부채 등 가계신용 데이터		코리아크레딧뷰로(KCB)	
동적(dynamic) 데이터	휴대전화 기반 유동인구 데이터	목적별 수요 분석	SDC 통계청(SKT 제공)	공공 및 민간
	상시 교통량	동적교통량 정보 개발 및 정산	교통량정보제공시스템(TMS)	공공
	지역별 코로나-19 확진자 기록	감염병 영향 분석	질병관리청	
	사회적거리두기 변화	정책대응 영향 분석	보건복지부	

자료: 저자 작성.

취합된 다양한 공공 및 민간 빅데이터를 시군구 공간 단위로 취합

업종별 사업자분포, 경제활동의 변화, 연령별/성별 인구구조, 기업 관계망 정보, 도로망 정보 등을 시군구 단위로 취합하여 머신러닝 훈련을 위해 통행목적별로 데이터셋을 구성

그림 2 본 연구에서 활용한 공공 및 민간 빅데이터의 공간 단위 취합 개념



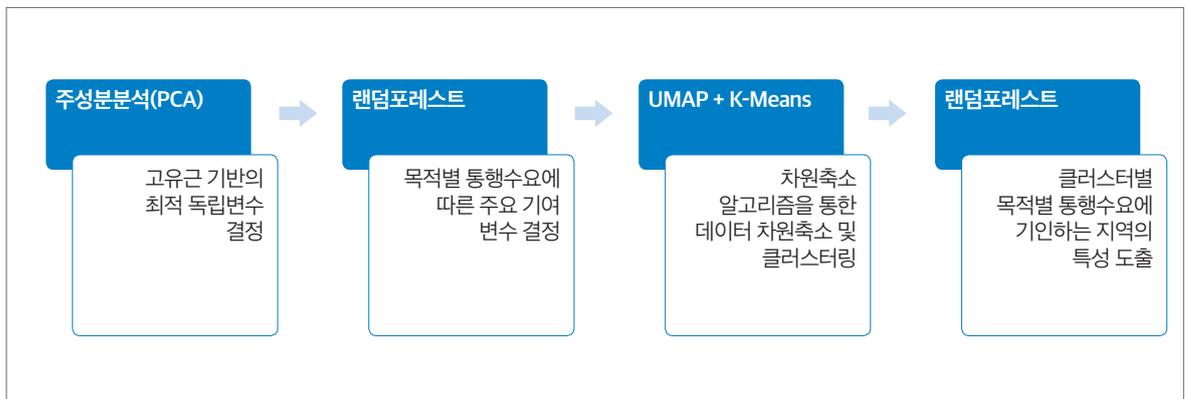
자료: (좌)통계청, 국가교통DB, 질병관리청, TMS교통량정보제공시스템, 디비리아, KCB, KED 이미지 등을 취합하여 저자 작성; (우)저자 작성.

머신러닝 및 딥러닝 등 다양한 AI 모형을 응용하여 예측모형 구축: 공간의 통행수요 유발요인 추정

공간의 통행수요 유발요인 추정을 위해 주성분분석(Principal Component Analysis, PCA), 랜덤포레스트(Random forest), UMAP, K-Means 클러스터링 등의 순차적인 분석을 진행

- 주성분분석(PCA)¹⁾은 데이터 분포로부터 데이터 변동성이 최대가 되는 축을 찾아 주성분을 결정하는 데 활용
- 랜덤포레스트는 목적별 통행에 영향을 주는 것으로 보이는 독립변수와 독립변수의 기여도를 파악하는 데 활용
- 데이터 차원축소 알고리즘 중 하나인 UMAP은, 다차원으로 구성된 입력 데이터의 특성은 유지하면서 데이터의 차원을 축소하는 역할로 활용
 - K-Means는 UMAP으로 차원축소된 데이터 분포로부터 군집 분석을 통해 지역을 선별해내기 위해 활용
- UMAP과 K-Means로 군집된 지역들의 통행수요 주요 기여 요인을 랜덤포레스트를 이용하여 추론

그림 3 공간의 통행수요 유발요인 추정과정 개념 도식화



자료: 저자 작성.

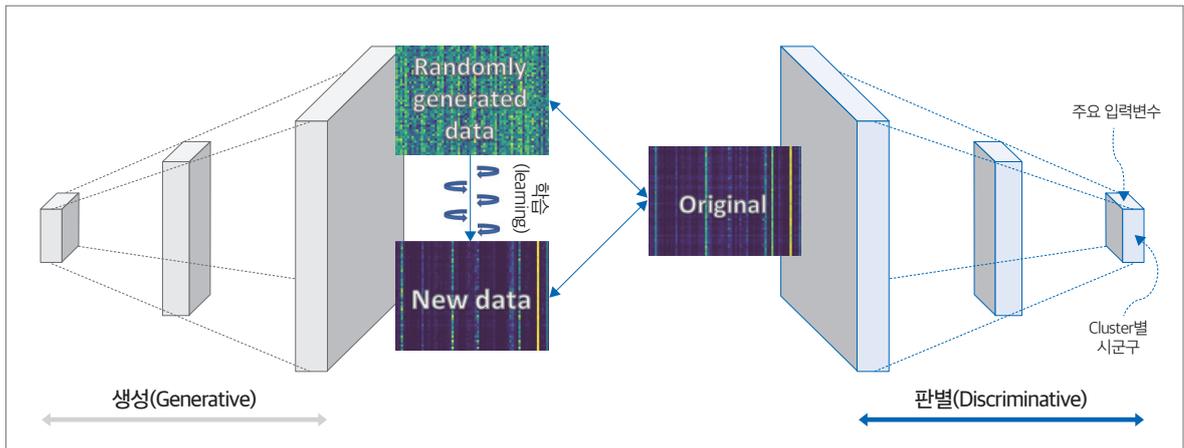
1) 주성분분석 방법은 분산이 큰 벡터를 기준으로 계산되기 때문에 주어진 입력변수로부터 영향을 주는 변수의 수 확인에는 용이하나, 데이터가 방대해질 경우 발생하는 비선형 데이터구조에서는 특성추출 효율성이 떨어짐.

머신러닝 및 딥러닝 등 다양한 AI 모형을 응용하여 예측모형 구축: 공간의 통행수요 유발요인 부각화

머신러닝 구현과정 중 구축된 학습데이터의 교통량 정보에 대한 훈련도 제고를 위해, 적대적생성신경망(Generative Adversarial Network, GAN) 딥러닝 모형을 이용하여 정적(static) 데이터인 공간과 관련된 통행수요 유발요인 등을 추가로 생성

- 상대적으로 한정된 패턴으로 제공되는 정적 데이터군과 역동적으로 변화하는 관측교통량과 유동인구 등 동적 데이터군 간의 고도화된 학습을 통해, 교통량 변화 등에 따른 목적별 통행행태 변화를 미세하게 감지하고 예측하기 위한 접근 방법

그림 4 적대적생성신경망(GAN)을 이용한 정적 데이터의 추가생성 모식도



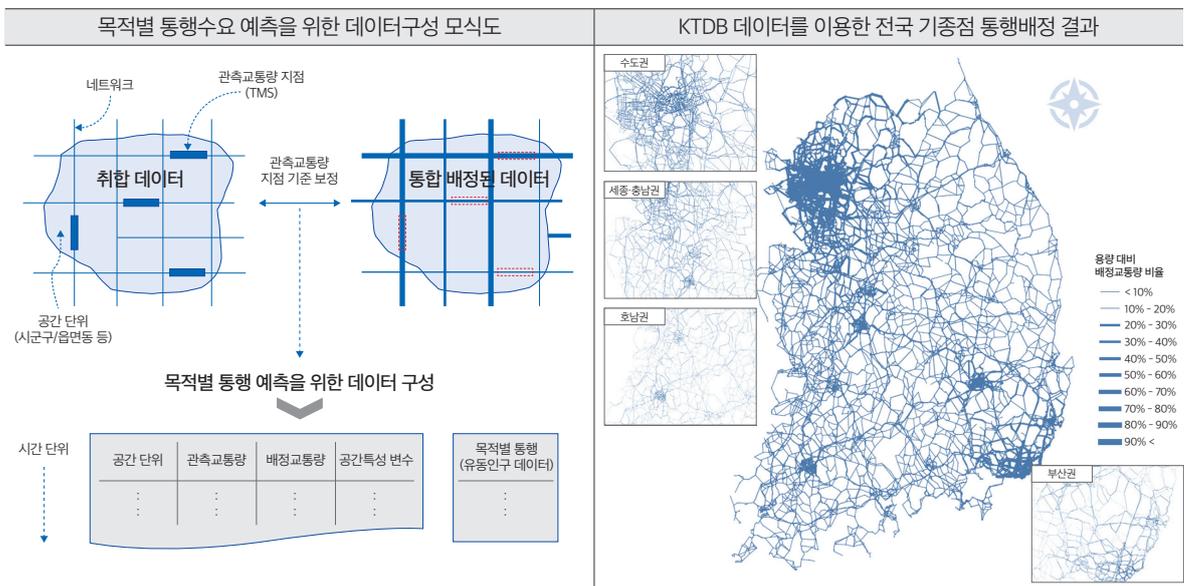
주: 클러스터별·통행목적별로 적대적생성신경망 모형을 생성하여 분석에 활용.
자료: 저자 작성.

머신러닝 및 딥러닝 등 다양한 AI 모형을 응용하여 예측모형 구축: 통행목적별 감염병 전후 변화예측

국가교통데이터베이스(KTDB)에서 기초사된 기종점통행정보(Origin-Destination trip matrix, O/D), 네트워크 데이터를 이용하여 Frank-Wolfe 기반의 통행배정(Traffic Assignment) 시뮬레이션을 진행

- 교통량정보시스템(TMS)에서 구득한 교통량정보 데이터는 전체 도로망의 정보가 아닌 지점별 교통량이기 때문에, 통행배정 시뮬레이션을 통해 교통량 보정 및 전체 링크의 교통량을 추정하는 단계가 필요
- 교통분석존(TAZ)과 해당 교통분석존 내 배정된 모든 링크의 교통량, 관측교통량, 그리고 관측된 목적별 통행수요를 클러스터별·통행목적별 주요 기여 요인들과 결합하여 AI 모형에 학습할 수 있도록 구성

그림 5 목적별 통행수요 예측을 위한 통행배정과 데이터 취합



자료: 저자 작성.

03. 코로나19 전후의 목적별 통행수요 변화 추정모형

AI 모형이 2019년 관측교통량 패턴으로 확인한 2020년 관측교통량의 목적별 통행수요 변화

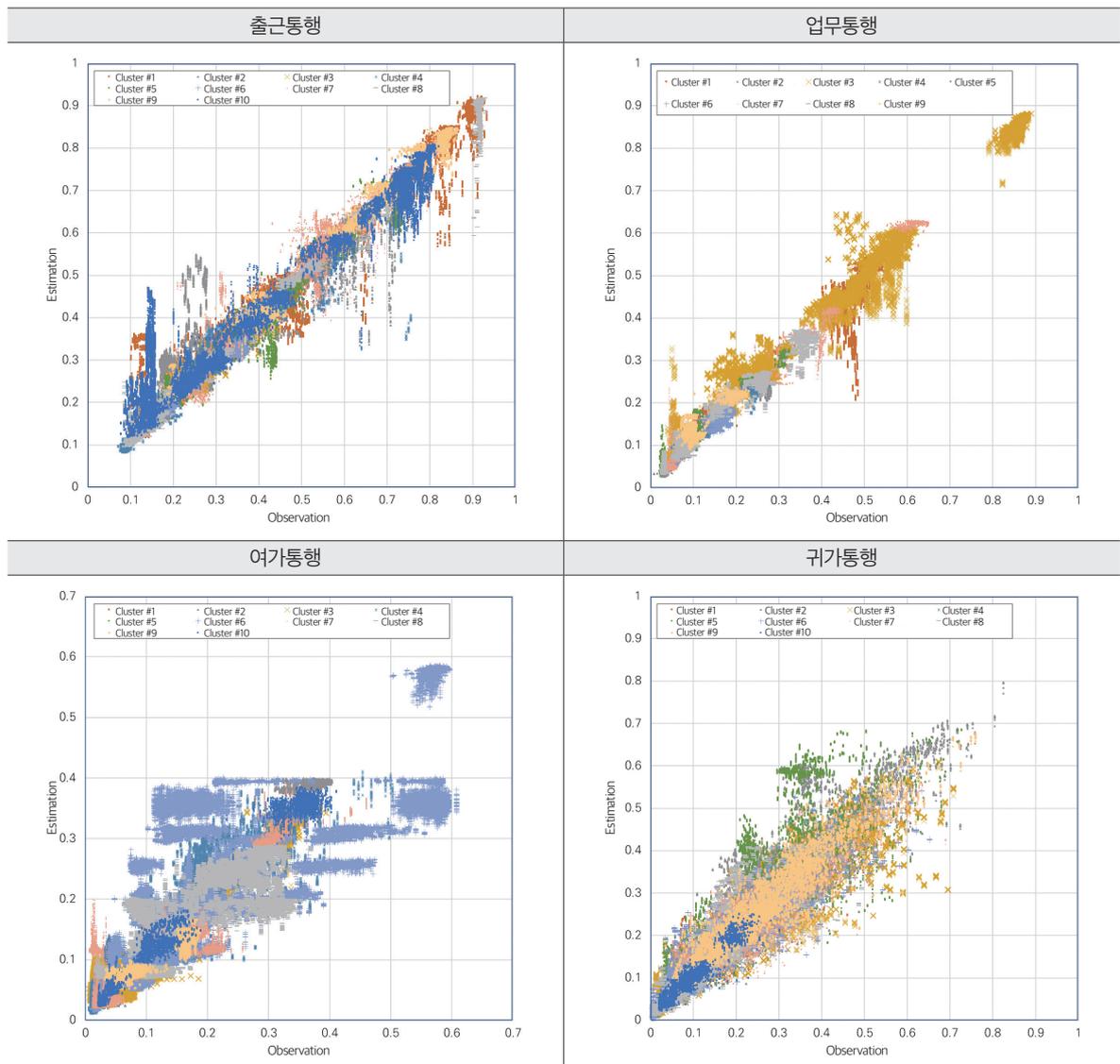
코로나19 발생 이전(2019년)의 통행목적별, 클러스터별 교통량 패턴에 따른 목적별 통행수요의 비율변화를 머신러닝 모형에 학습하도록 한 뒤, 감염병 발생 이후(2020년) 교통량 패턴에 따른 목적별 통행수요를 어떻게 예측하는지 분석함

- 감염병 발생 전 관측교통량과 목적별 통행수요 비율변화를 학습한 뒤, 감염병 발생 이후 동일한 지역 및 클러스터 등에서 수집된 관측교통량 통행패턴에 따라 목적별 통행수요를 어떻게 예측하는지 확인
 - 예측된 목적별 통행비율이 실제 관측된 목적별 통행수요 비율보다 적다면, 감염병 발생 이후 해당 지역에 해당 목적의 통행이 감염병 이후 증가한 경우이고, 반대의 상황에는 감소한 경우로 해석해 볼 수 있음

필수통행이라 할 수 있는 출근통행과 업무통행은 증감이 확인되긴 하지만 대체로 대각선 형태에서 크게 흩어지지 않는 패턴인 반면, 비필수통행이라 할 수 있는 여가, 쇼핑 등의 목적별 통행은 산포된 패턴을 보임

- 특히 여가통행의 경우, 관측값 대비 급격한 증가와 감소 모두가 확인되어 통행의 빈익빈 부익부가 확인되는 등 감염병 확산 이후 지역적 여파가 짐작됨

그림 6 통행목적별 감염병 영향 추정 결과 예시(학습된 통행패턴 대비 실제 관측통행패턴)



주: 각 그래프의 x축은 실제 2020년 관측 통행패턴을, y축은 2019년 데이터로 학습된 모형을 이용하여 2020년 관측교통량으로부터 도출된 통행패턴 예측 결과임.
자료: 저자 작성.

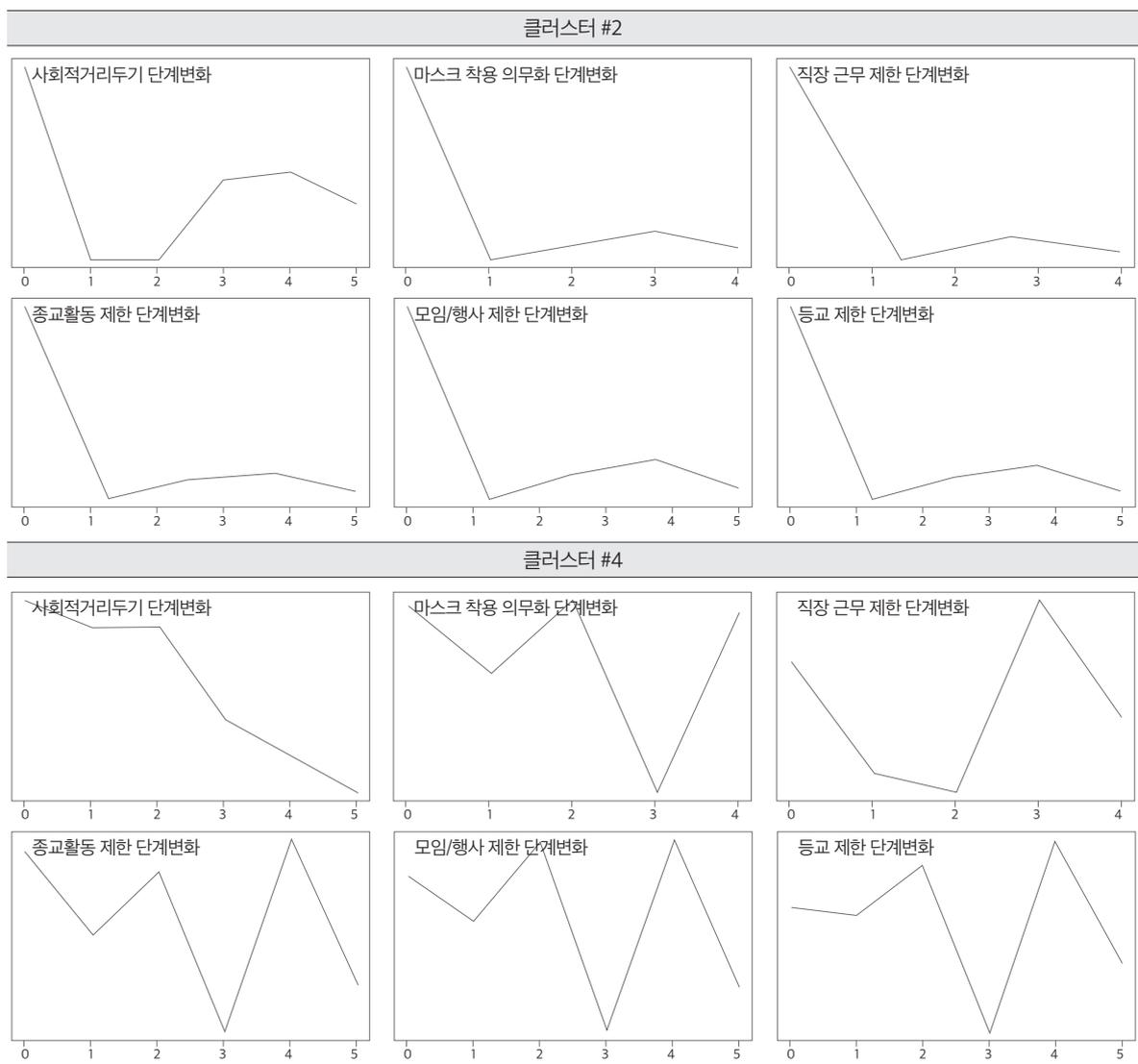
04. 포스트 감염병 대응을 위한 목적별 통행수요 예측모형 활용방안

감염병 대응 정책변화에 따른 통행패턴 변화 민감도 분석

본 연구에서 구축한 데이터구조에서 사회적거리두기 단계별 데이터를 학습하여 시공간(시기별·시군구별) 특징에 따른 감염병 대응 정책변화와 목적별 통행수요 변화를 확인함

- 감염병 발생 이후, 사회적거리두기 단계변화와 시군구별 확진자 정보 등을 취합하여 교통량 패턴과 목적별 통행수요 패턴을 학습하여 진행하였고, 동일한 통행목적에도 불구하고 클러스터 분류 지역마다 다양한 감염병 대응 정책변화에 따른 민감도 변화를 보임
 - 출근통행의 경우 클러스터 #2에서는 사회적거리두기 단계변화에 급격한 통행비율 감소가 확인되는 데 반해, 클러스터 #4에서는 완만한 통행비율 감소가 확인되며 기타 정책변화에 급격한 증감패턴을 보임
 - 클러스터 #2는 경상북도 일대 지역으로 건설업 및 에너지와 관련된 기업체 관계가 주요 지역의 통행수요 유발요인이었던 것으로 확인되는 데 반해, 클러스터 #4는 서울 지역으로 시설 임대업과 협회 및 개인 서비스업 등이 클러스터 지역의 유발요인으로 확인되어 감염병으로 인한 지역별 영향도 편차가 확인됨

그림 7 통행목적별 감염병 대응 정책변화에 따른 민감도 분석 결과 출근통행 예시 (위 - 클러스터 #2, 아래 - 클러스터 #4)



주: 각 그래프의 x축은 사회적거리두기 항목별 단계를 나타내고, y축은 통행목적별 통행비율 증감률을 나타냄.
자료: 저자 작성.

05. 결론 및 향후 연구 과제

최근 고도화된 알고리즘과 컴퓨팅 자원, 데이터 댐 등으로 축적된 다양한 공공 및 민간 빅데이터 융합을 통한 딥러닝 기반의 시의성 높은 목적별 통행수요 예측모형을 시범적으로 개발

- 사람의 학습을 모방하여 정교하고 예측력 높은 딥러닝 모형을 다양한 분석 및 데이터 융복합 단계에서 활용함으로써 향후 알고리즘 확장과 연계방안을 함께 고려
- 감염병 확산과 정부의 사회적거리두기 대응이라는 첨예하게 대조적인 상관관계에 다각도로 모형을 이용할 수 있다는 장점이 있어 향후 감염병 및 유사 사회적거리두기 정책 전개에도 유용하게 활용될 것으로 기대

모형에서 제시하는 클러스터 및 해당 지역을 고려한 일련의 방역정책 전개가 감염 확산에 어떠한 영향을 미칠지 모니터링 연계 필요

- 다양한 목적별 통행수요 예측모형의 전개와 지역의 특성을 고려한 맞춤형 방역정책 추진이 필요

모형에서 산출한 목적별 통행수요 예측치와 실제 목적별 통행량 간의 이격을 확인하여 현실감 있는 정책지원 도구를 지향

- 본 연구에서 제시된 모형이 예측의 범주에서도 그 효용성이 있으므로, 감염병 확산과 사회적거리두기 정책의 전개 그리고 통행 간의 현상 등을 함께 모니터링할 수 있는 추가연구가 필요
- 분석과정의 복잡함은 본 연구의 한계로, 향후 연구과제에서는 학습 방법 고도화와 함께 분석 절차의 간소화 등을 향후 과제로 추진

참고문헌

- 교통량정보제공시스템(TMS). <http://www.road.re.kr/main/main.asp> (2021년 7월 7일 검색).
- 국가교통DB. <https://www.ktdb.go.kr/www/contents.do?key=23> (2021년 7월 6일 검색).
- 손재선, 장요한, 황명화, 임은선, 이경주. 2021. 감염병의 시공간 확산모형을 활용한 정책시뮬레이션 방법 개발 연구. 세종: 국토연구원.
- 조성철, 장요한, 장은교, 김석윤. 2020. 지역산업 클러스터의 경쟁력과 발전방안 연구. 세종: 국토연구원.
- 코리아크레딧뷰로(KCB). <http://www.koreacb.com> (2021년 11월 11일 검색).
- 한국기업데이터(KED). <http://www.kedkorea.com> (2021년 11월 11일 검색).
- Mclnnes, L., Healy, J. and Melville, J. 2018. Umap: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv preprint arXiv:1802.03426.
- SDC 통계청. 통신 모바일 인구가동량 통계. <https://data.kostat.go.kr/social/mobilePopMoveInfoPage.do> (2022년 2월 10일 검색).
- Sheffi, Y. 1985. *Urban Transportation Networks: Equilibrium Analysis With Mathematical Programming Methods*. MIT press.
- ※ 이 브리프는 “장요한·이영민·임시영·박서현·김현경. 2022. 포스트 감염병 대응을 위한 딥러닝 기반 목적별 통행수요 예측모형 개발 연구. 세종: 국토연구원” 보고서를 요약·정리한 것임.

- **장요한** 국토연구원 국토데이터랩 팀장(ycanns@krihs.re.kr, 044-960-0406)
- **이영민** 국토연구원 국토데이터랩 부연구위원(ymlee@krihs.re.kr, 044-960-0420)
- **임시영** 국토연구원 공간정보사회연구본부 부연구위원(limsy@krihs.re.kr, 044-960-0565)
- **박서현** 서울시립대학교 도시빅데이터융합학과 연구원(psh0196@uos.ac.kr, 02-6490-5684)
- **김현경** 국토연구원 국토모니터링연구센터 연구원(withkim@krihs.re.kr, 044-960-0316)

