Check for updates
pISSN: 1226-7147, eISSN: 2383-9171
https://doi.org/10.17208/jkpa.2025.08.60.4.188
Journal of Korea Planning Association Vol.60, No.4 (2025)
대한국토·도시계획학회지 「국토계획」 제60권 제4호 pp.188-203

근린환경 특성과 도시활력의 비선형 관계 분석
: 해석 가능성 기반 기계학습 모형의 적용*

Analysis of the Nonlinear Relationships between Neighborhood Environmental Characteristics and Urban Vitality
: Applications of Interpretability-based Machine Learning Model

조월** · 김선재*** · 이수기****
Cao, Yue · Kim, Sunjae · Lee, Sugie

Abstract

Creating a vibrant neighborhood environment is a key component of sustainable urban development. Urban theorist Jane Jacobs explains that urban vitality occurs through the interactions of human activities with neighborhood environments. Drawing on the recent development of big data and machine learning technologies, this study analyzes the impact of neighborhood environmental factors on urban vitality. This study utilizes big data such as De Facto Population, Points-Of-Interest (POI), and Street View images for the city of Seoul and employs a machine learning model to understand urban vitality. It derives key variables that affect urban vitality and checks the nonlinear relationships between variables by utilizing explainable machine learning model. The main analysis results are as follows. It also indicates that land use characteristics and POI show strong associations with urban vitality. Specifically, SHapley Additive exPlanations (SHAP) analysis results confirm that the independent variables largely show nonlinear relationships with urban vitality. Moreover, the study identified critical thresholds for variables such as residential area density and distance to subway stations, beyond which their impact on urban vitality becomes constant. This study is significant because it provides a clearer understanding of the key neighborhood environmental factors that affect urban vitality. Furthermore, this study offers planning and policy implications that promoting urban vitality and social interaction.

주제어 도시활력, 생활인구, 해석가능한 인공지능, 관심시설, Shapley Index
Keywords Urban Vitality, De Facto Population, eXplainable Artificial Intelligence, Points-Of-Interest, Shapley Index

Ⅰ. 서 론

1. 연구의 배경 및 목적

도시활력(Urban Vitality)은 도시의 질을 평가하는 중요한 요소 중 하나이다. 활력이 있는 도시 공간은 사회적 교류와 지속가능한 도시를 만드는 데 도움이 된다(Sulis et al., 2018). 도시 이론가인 Jacobs(1961)에 따르면, 도시활력은 사람의 활동과 근린환경의 상호작용을 통해서 일어난다고 설명하였다. Lynch(1981)는 주민의 다양한 수요를 충족시킬 수 있는 안전한 도시환경과 형태가 도시활력에 영향을 미칠 수 있다고 하였다. 또한, Long and Zhou(2016)는 가로의 활력은 사회적 활력을 의미하며, 가로의

* 이 논문은 2021년 4월 대한국토·도시계획학회 춘계산학학술대회에서 우수논문상을 수상하였으며, 조월의 한양대학교 석사학위 논문을 수정·보완하여 작성하였음. 또한, 이 논문은 한양대학교 인문사회예체능 전문학술논문지원사업의 지원을 받아 수행된 연구임(HYU-202300000003596).
** Master's Degree, Department of Urban Planning & Engineering, Hanyang University (First Author: tracy0703@hanyang.ac.kr)
*** Ph.D. Student, Department of Urban Planning & Engineering, Hanyang University (sunjaekim@hanyang.ac.kr)
**** Professor, Department of Urban Planning & Engineering, Hanyang University (Corresponding Author: sugielee@hanyang.ac.kr)

물리적 공간 자체는 활력을 만들어내지 못하지만, 사람들이 활동하는 사회적 공간이 될 때 가로활력이 생긴다고 하였다.

이에 따라, 물리적 환경과 도시활력과의 관계에 관한 연구는 그동안 여러 연구자에 의해 진행되었다. 그러나 대부분의 연구에서 분석자료의 한계로 도시활력과 관련된 연구는 소규모 관찰연구로 진행된 한계점이 있다. 특히, 현장조사와 관찰조사는 샘플 수도 제한적이고 많은 시간이 소요되기 때문에 도시 빅데이터를 적극적으로 활용하여 도시활력을 정량화하는 연구가 증가하고 있다(He et al., 2018; Yue et al., 2019; Liu et al., 2020).

또한, 기존연구는 도시 근린환경 특성을 정량화하여 도시활력에 대한 근린환경 특성의 상대적 중요성을 평가하고 있다(Sung and Lee, 2015; Sung et al., 2015; Ye et al., 2018). 그러나 대부분의 연구는 도시활력에 영향을 미치는 요인과의 선형관계 분석에 초점을 맞추고 있어 비선형관계를 설명하는 데 한계를 가지고 있다. 최근 비선형관계 분석에 대한 전통적인 회귀분석의 한계를 극복하기 위해 기계학습(Machine Learning) 방법론이 활용되고 있다(Scepanovic et al., 2021; Yang et al., 2021).

따라서 본 연구는 도시 공간 빅데이터와 설명 가능한 인공지능(eXplainable Artificial Intelligence, XAI) 기법을 활용하여 도시활력에 영향을 미치는 근린환경 특성을 도출하고자 한다. 특히, 도시활력에 영향을 미치는 환경 요인의 비선형적 임곗값을 규명함으로써, 특정 임곗값을 기준으로 도시활력 증진을 위한 정책적 시사점을 제시할 수 있다. 이를 통해 본 연구는 도시활력 제고를 위한 실효성 있는 도시계획 및 정책 수립을 위한 과학적 근거를 제공하고, 도시활력 증진과 사회적 교류 활성화를 위한 공간적 전략을 도출하는 데 기여하고자 한다.

II. 선행연구 고찰

1. 도시활력의 개념과 정의

도시활력은 사회적인 개념으로 다양한 의미를 가지고 있다. 도시활력은 "사회, 경제활동을 충족시킬 수 있는 도시의 중요한 요소"(Jacobs, 1961), "상업, 유흥시설이 많고 보행인구가 밀접한 지역의 특성"(Maas, 1984), "다양한 만남과 활동이 일어나는 장소의 특성"(Montgomery, 1998) 등으로 다양하게 정의되고 있다. 대부분의 도시활력 관련 연구를 살펴보면 유동인구나 생활인구의 집중으로 사람들의 활동이 활발하게 일어나는 정도를 도시활력으로 정의하고 있음을 알 수 있다.

이와 같은 다양한 도시활력의 개념을 정량적으로 측정하기 위해서 다양한 요인들이 활용되고 있다. 구체적으로 도시활력을 측정하기 위한 지표로 음식점 리뷰 개수(Yue et al., 2019), 카드매출(Kim, 2018), 야간 조명(Zhou et al., 2023) 등이 활용되었다.

그리고 최근 스마트폰 기반 빅데이터를 가지고 도시활력을 대표하는 논문들이 증가하였다(Jacobs-Crisioni et al., 2014; Long and Zhou, 2016; Tang et al., 2018; Li et al., 2020). Jacobs-Crisioni et al.(2014)의 연구는 스마트폰 데이터를 기반으로 토지이용의 혼합 정도가 도시활력에 미치는 영향을 분석하였고, Long and Zhou(2016)의 연구는 스마트폰 데이터를 가지고 가로의 활력에 영향을 미치는 요인을 분석하였다. 특히, Tang et al.(2018)의 연구는 스마트폰 기반 빅데이터가 도시활력에 있어 두 가지 장점을 가지고 있다고 언급하였다. 첫 번째 장점은 지리 위치로 통계된 스마트폰 데이터는 미시적인 시간 및 다양한 공간 척도에서 인구의 분포를 반영할 수 있다는 것이다. 두 번째 장점은 스마트폰 데이터가 전체 사용자를 포함하기 때문에 넓은 공간을 고려할 수 있고 도시활력의 시간적, 공간적 변화를 반영할 수 있다고 하였다.

한편, 국내에서도 스마트폰 기반 생활인구를 활용하여 도시활력을 측정하는 연구가 증가하고 있는 추세이다(이지혜·김형중, 2019; 조월 외, 2021). 구체적으로 조월 외(2021)는 가로활력 대리변수로 생활인구를 사용하여 토지이용 현황과 건조환경 요인이 가로활력에 미치는 영향을 서울시 집계구 단위에서 분석하였다. 또한, 정재훈·남진(2019)은 서울시 2017년 생활인구 빅데이터를 활용하여 서울시 행정동 단위를 기준으로 도시공간 내 생활인구의 시계열적 패턴을 유형화하고, 각 유형의 지역 특성을 파악하였다. 분석 결과, 서울의 생활인구는 공간 범위와 시간대에 따라 다른 패턴을 갖고 있으며, 이를 바탕으로 인구의 활동을 측정하기 위해 생활인구 빅데이터가 유용하게 활용될 수 있다는 시사점을 제시하였다. 이처럼, 생활인구 데이터는 도시활력을 측정하는 데 있어 적합한 자료인 것을 알 수 있다.

2. 도시활력 영향요인

도시활력 영향요인에 대한 대표적인 이론은 Jacobs(1961)에 의해 제시되었다. Jacobs는 가로활력을 제고하기 위한 네 가지 조건으로 적절한 밀도, 작은 블록 구조, 토지이용의 혼합 정도, 오래된 건물의 중요성 등을 제시하였다. 이러한 관점은 이후 많은 연구에서 도시활력에 미치는 물리적 환경 특성의 중요성을 강조하는 근거가 되었다(Montgomery, 1998).

최근에 이루어진 도시활력 관련 실증연구도 물리적 환경에 초점을 두고 있으며, 주로 토지이용 특성(임하나 외, 2016; Wu et al., 2018b), 교통환경 특성(Wu et al., 2018a; 곽호찬 외, 2018), 가로형태(Yue et al., 2019), 도시형태 특성(Buchanan, 1988), 공공시설의 접근성 등 다양한 관점에서 연구가 이루어졌다. 특히, 토지이용 특성변수는 대부분의 연구에서 도시활력과 밀접한 관련이 있음을 보이고 있다. 그러나 넓은 범위에서 주거, 상업, 업무 등의 토지이용 연면적 지표에 집중하고 있어 구체적으로 어떤 토지이용 시설이 도시활력과 밀접한 관련이 있는지 제시하는

데 한계가 있다. 이러한 맥락에서 최근 구체적인 토지이용 유형을 대변하는 Point-of-Interest(POI) 자료를 분석에 활용하고 있다(He et al., 2018; Liu et al., 2020).

나아가 최근 가로경관 이미지 빅데이터와 기계학습 딥러닝(Deep Learning) 기술의 발전으로 정량화된 가로경관 이미지 자료가 보행환경 관련 실증연구에 활용되고 있다(Wang and Vermeulen, 2020; Ma et al., 2021). 예를 들면, Wang and Vermeulen(2020)은 가로경관 이미지를 활용하여 도로 비율, 표지판 비율이 높은 곳이 보행보다는 자동차 통행이 많기 때문에 도시활력이 낮게 나타나는 것으로 설명하였다.

3. 도시활력 분석 방법론

한편, 도시활력을 분석하기 위해 다양한 연구방법론이 이루어졌다. 도시의 근린환경 요소가 도시활력에 미치는 영향에 대해 수행된 대부분의 기존 연구방법은 보편적인 방법으로 최소제곱법(Ordinary Least Square: OLS) 회귀모형을 사용하고 있다. 그러나 실제 근린환경과 도시활력과의 관계는 비선형관계(Non-linear Relationship)가 많이 도출되고 있다(Ding et al., 2018; Tao et al., 2020; Yang et al., 2021). 전통적인 회귀분석 방법론은 이차 함수 이상의 비선형적 관계를 분석하는 데 한계를 가지고 있다. 도시활력에 영향을 미치는 독립변수의 비선형 관계를 고려하지 못하면 도시활력 증진을 위한 구체적인 정책개발 방향을 제시하는 데 한계가 있을 수 있다(Wu et al., 2019). 최근, 도시활력을 분석한 Yang et al.(2021)의 연구는 도시활력에 영향을 미치는 근린환경 요소가 비선형적인 임곗값을 가지고 있다고 주장하였다. 이는 일부 변수의 경우 임곗값을 초과할 때만 도시활력과 연관이 있고, 특정 변수는 임곗값 범위 내에 있을 때만 연관성을 가질 수 있음을 의미한다. 이러한 측면에서 도시활력에 영향을 미치는 요인을 분석하는 데 있어 선형관계보다 변수 간의 비선형관계 분석이 중요하다고 볼 수 있다.

도시활력과 설명변수 간의 비선형관계를 반영하기 위하여 인공지능의 주요 분야 중 하나로 기계학습(Machine Learning) 방법이 유용하게 활용될 수 있다. 최근 기계학습으로 독립변수와 종속변수 간의 비선형관계를 분석하고자 하는 연구는 다양한 분야에서 진행되었다(Wang et al., 2014; Gan et al., 2020; Yan et al., 2020; Yang et al., 2021). 전통적인 OLS 회귀분석 방법론은 분석 변수나 오차항의 분포, 변수 간의 다중공선성 등과 관련하여 엄격한 가정을 전제로 하고 있다. 반면, 기계학습 방법론은 이러한 가정을 전제로 하고 있지 않아 다양한 변수를 모형에 고려할 수 있으며 비선형관계 분석에도 장점을 가지고 있다.

4. 연구의 차별성

선행연구 고찰을 통해 기존 도시활력에 대해 연구의 한계점은 다음과 같이 정리할 수 있다. 첫째, 도시활력의 연구 방법에 있어 대다수의 선행연구는 전통적인 선형회귀 모형을 사용했다. 선형회귀 모형은 종속변수와 독립변수의 선형적인 관계를 가질 것으로 가정한다. 하지만, 실제 도시활력에 영향을 미치는 도시환경 요인은 선형적으로 설명하는 데 한계가 있을 수 있으며, 선형적 관계를 가정할 때 분석 결과가 왜곡될 수 있다. 따라서 본 연구는 기계학습을 활용하여 도시활력과 근린 환경 요인 간의 비선형 관계를 분석하고, 특정 변수의 임곗값을 규명함으로써 도시활력 증진을 위한 도시계획적 시사점을 도출하고자 한다.

둘째, 기존에 기계학습으로 도시활력에 영향을 미치는 요인을 분석한 연구는 기계학습의 블랙박스(Black Box) 구조 때문에 어떤 근린환경 요인이 영향을 미쳤는지 확인할 수 없는 한계가 있다. 따라서 본 연구는 다양한 모형과 비교뿐만 아니라 기계학습 모형을 해석할 수 있는 SHapley Additive exPlanations(SHAP)로 기계학습 모형의 결과를 해석하고자 한다. 이를 통해 도시활력에 영향을 미치는 요인의 상대적 중요도를 평가하고, 그 영향이 어떻게 변화하는지에 대한 설명 가능성을 높이고자 한다. 이런 방법은 근린 환경 설계 및 정책 수립에 대한 실질적 가이드라인을 제시할 수 있을 것으로 기대된다.

마지막으로 비선형 관계를 고려한 기존 연구들도 존재하지만, 대부분은 변수 간 관계를 확인하는 데 초점을 맞췄으며, 특정 요인의 임곗값을 규명하는 데 구체적 분석이 부족했다. 그러나 정확한 임곗값을 도출하는 것은 도시활력 증진을 위한 정책개발 측면에서 매우 중요하다. 따라서 본 연구는 SHAP 기법을 활용하여 도시의 근린환경이 도시활력에 미치는 영향을 분석하고 비선형적 변화 패턴을 분석하고, 특정 임곗값을 규명함으로써 정책적인 시사점을 제시하고자 한다.

Ⅲ. 연구 방법론

1. 연구의 범위

본 연구는 서울 열린데이터광장에서 제공하는 서울시 생활인구 데이터를 도시활력 지표로 활용하였다. 도시활력 분석모형은 전통적인 회귀분석 모형과 기계학습 모형으로 구축하고 모형 간의 비교를 통해 가장 설명력이 높은 모형을 도출하였다. 도시활력 설명변수의 중요도와 비선형 관계를 파악하기 위해, 최종 모형에 SHAP 방법론을 활용하여 분석하였다.

본 연구의 공간적인 범위는 〈그림 1〉과 같다. 구체적으로는 서울특별시 25개 자치구에 분포된 17,196개[1] 집계구를 기본적인 공간분석 단위로 사용하였다. 또한, 연구의 시간적 범위는

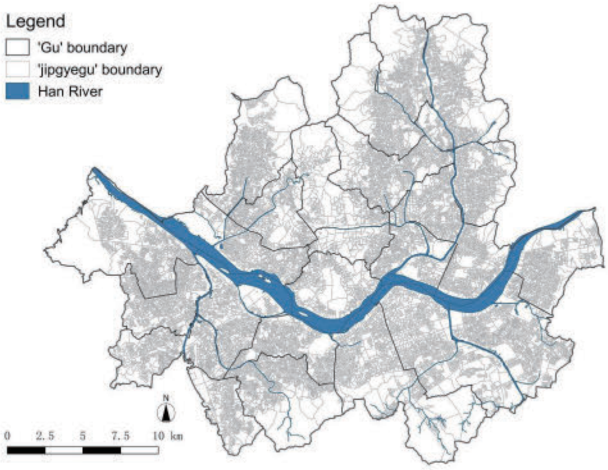


그림 1. 서울시 자치구 및 집계구 경계
Figure 1. The boundaries of administrative 'gu' and 'jipgyegu'

COVID-19가 발생하기 이전인 2019년 1월 1일 9시부터 2019년 12월 31일 18시까지 1년간의 생활인구 데이터를 사용하였다.[2] 이에 따라, 집계구 수준의 설명변수에 관한 자료의 시점은 대부분 2019년 자료를 활용하였지만, 일부 물리적 환경변수는 1년 사이 큰 변화가 없었다는 전제하에 2020년 자료를 분석에 활용하였다.

2. 분석변수와 자료 출처

본 연구에서 사용한 분석변수의 구체적인 설명과 데이터 출처는 〈표 1〉과 같다. 연구의 종속변수는 서울 열린데이터광장에서 제공되는 생활인구 데이터를 이용하여 집계구별 낮 시간대의 일평균 생활인구 수를 산출하였다. 독립변수로는 서울 열린데이터광장, 도로명주소 DB, 카카오맵과 Naver Street View(NSV) 자료를 활용하여 토지이용 특성, POI시설 특성, 교통 특성 그리고 시각적으로 보이는 보행환경 특성을 선정하였다.

3. 변수 선정

도시활력은 사람들의 다양한 활동으로 생겨나는 개념이기 때문에, 이러한 개념적 특성을 바탕으로 실시간 스마트폰 기반으로 집계된 생활인구 수를 도시활력 대리변수로 사용하였다. 종속변수인 '도시활력'의 조작적 정의는 서울시 생활시간조사 데이터와 낮 시간대의 활동 인구를 고려하여 집계구 단위로 낮 시간대(9시-18시)의 주중, 주말 생활인구의 시간대별 평균으로 정의하였다. 구체적으로 서울시 열린데이터광장에서 제공하고 있는 2019년 1월 1일부터 2019년 12월 31일까지 1년간 서울시 생활인구 데이터 자료를 활용하였다. 서울시의 생활인구 추계 매뉴얼에 따르면, 생활인구는 "조사 시점 현재 서울에 머무르고 있는 현재인구로 서울에 주소지를 둔 사람은 물론 업무, 관광, 의료 및 교육 등 일시적으로 서울을 찾아 행정수요를 유발하는 인구를 총칭하는 인구를 말한다"로 정의되고 있다(서울특별시, 2018, p.6).

본 연구에서 활용한 독립변수는 크게 토지이용 특성, POI시설 특성, 교통 특성, 그리고 보행환경 특성으로 나누어 구성하였다(〈표 2〉 참조). 독립변수 중 토지이용 특성 변수의 경우 인구 분포와 보행활동에 영향을 미칠 수 있는 중요한 변수이다(Jacobs, 1961). 토지이용 특성 변수는 도로명주소 DB에서 제공되고 있는 건축물 데이터를 활용하여 구축하였다. 선행연구를 참조하여 집계구별 주거시설 연면적 밀도, 상업시설 연면적 밀도, 업무시설 연면적 밀도로 계산하여 구축하였다(이수기 외, 2016; 임하나 외, 2016; 조월 외, 2021). 또한, 토지이용 혼합도는 엔트로피 지표의 개념을 활용하였다. 토지이용 혼합도는 0에서 1 사이의 값을 가지며, 1에 가까울수록 해당 집계구에서 토지이용의 혼합 정도가 높음을 의미하고, 서울의 경우 평균 0.30의 수치로 나타났다.

다음으로 POI 데이터를 활용하여 POI시설 변수를 구축하였다. POI시설 변수는 선행연구에서 주로 사용했던 주거, 상업, 업무 등 대분류 토지이용 변수의 한계를 보완할 수 있다. 특히 사람들의 활동과 밀접한 관련을 가지는 토지이용시설의 공간적인 집중을 대변할 수 있는 장점이 있다. POI시설 변수는 카카오맵을 통해 구득하였으며, 서울시 전체 약 428,668개의 POI시설 데이터를 분석에 활용하였다. POI시설 데이터는 지도에 미시적으로 표시하는 지점 단위의 공간 데이터로, 시설물, 지하철, 호텔, 음식점 등의 종류를 확인할 수 있고, 시설의 경위도, 시설 종류, 시설 명칭 등 정보를 포함하고 있다. 한편, POI시설 데이터는 한 건물에 하나의 주 용도를 대표하는 전통적인 토지이용 데이터보다 건물 단위에서 다양한 용도의 시설을 더 세분화하여 측정할 수 있는 장점을 가지고 있다(Wu et al., 2018b).

표 1. 변수설명
Table 1. Description of variables

Variable		Description	Data source
종속변수 Dependent variable	주중 일평균 생활인구(9-18시) A weekday average daily de facto population	주중 낮 시간대의 일평균 생활인구 수 The average daily de facto population in the daytime hours of the weekday	서울 열린데이터광장(2019) Seoul Open Data Plaza (2019)
	주말 일평균 생활인구(9-18시) A weekend average daily de facto population	주말 낮 시간대의 일평균 생활인구 수 The average daily de facto population in the daytime hours of the weekend	

다음 페이지에 계속(Continue on next page)

Variable			Description	Data source
독립변수 Independent variable	토지이용 특성 Land use characteristics	주거용도 연면적 밀도(m^2/m^2) Density of total floor area (residential)	주거용도 총 연면적 / 집계구 면적 Total floor area of residential facility/ Jipgyegu total area	도로명주소 DB(2019) Road name address DB (2019)
		상업용도 연면적 밀도(m^2/m^2) Density of total floor area (commercial)	상업용도 총 연면적 / 집계구 면적 Total floor area of commercial facility/ Jipgyegu total area	
		업무용도 연면적 밀도(m^2/m^2) Density of total floor area (office)	업무용도 총 연면적 / 집계구 면적 Total floor area of office facility/Jipgyegu total area	
		토지이용 혼합도 Land Use Mix	집계구별 토지이용의 혼합 정도 (주거, 상업, 업무, 녹지, 기타) Land-use mix of jipgyegu (Residential+Commercial+Office+Green+Other)	
	POI시설 특성 POI facilities characteristics	공공기관시설 밀도(개/km^2) Density of public institutions facilities (Num./km^2)	집계구별 공공기관시설 총개수 / 집계구 면적 Total number of public institutions facilities by jipgyegu / Jipgyegu total area	카카오맵(2020) Kakaomap (2020)
		교육시설 밀도(개/km^2) Density of educational facilities (Num./km^2)	집계구별 교육시설 총개수 / 집계구 면적 Total number of educational facilities by jipgyegu / Jipgyegu total area	
		편의시설 밀도(개/km^2) Density of convenience facilities (Num./km^2)	집계구별 편의시설 총개수 / 집계구 면적 Total number of convenience facilities by jipgyegu / Jipgyegu total area	
		금융시설 밀도(개/km^2) Density of finance facilities (Num./km^2)	집계구별 금융시설 총개수 / 집계구 면적 Total number of finance facilities by jipgyegu / Jipgyegu total area	
		레저시설 밀도(개/km^2) Density of leisure facilities (Num./km^2)	집계구별 레저시설 총개수 / 집계구 면적 Total number of leisure facilities by jipgyegu / Jipgyegu total area	
		문화시설 밀도(개/km^2) Density of culture facilities (Num./km^2)	집계구별 문화시설 총개수 / 집계구 면적 Total number of culture facilities by jipgyegu / Jipgyegu total area	
		병원시설 밀도(개/km^2) Density of hospital facilities (Num./km^2)	집계구별 병원시설 총개수 / 집계구 면적 Total number of hospital facilities by jipgyegu / Jipgyegu total area	
		산업시설 밀도(개/km^2) Density of industrial facilities (Num./km^2)	집계구별 산업시설 총개수 / 집계구 면적 Total number of industrial facilities by jipgyegu / Jipgyegu total area	
		여행시설 밀도(개/km^2) Density of trip facilities (Num./km^2)	집계구별 여행시설 총개수 / 집계구 면적 Total number of trip facilities by jipgyegu / Jipgyegu total area	
		음식점시설 밀도(개/km^2) Density of dining facilities (Num./km^2)	집계구별 음식점 총개수 / 집계구 면적 Total number of dining facilities by jipgyegu / Jipgyegu total area	
	교통 특성 Transportation characteristics	교차로 밀도(개/km^2) Density of intersections(Num./km^2)	집계구별 교차로 총개수 / 집계구 면적 Number of intersections / Jipgyegu total area	서울 열린 데이터 광장 (2019) Seoul Open Data Plaza (2019)
		버스정류장 밀도(개/km^2) Density of bus station (Num./km^2)	집계구별 버스정류장 총개수 / 집계구 면적 Number of bus station / Jipgyegu total area	
		지하철역까지의 거리(m) Distance to the subway station(m)	집계구별 중심점에서 가장 가까운 지하철역까지 거리 The nearest distance from each jipgyegu to the subway station	
	보행환경 특성 Walking environment characteristics	보행로 비율 Sidewalk ratio	집계구 평균 보행로 비율 Average sidewalk ratio	네이버 지도(2020) Naver Map (2020)
		녹시율 Green view index	집계구 평균 녹시율 Average green view index	

표 2. 기초 통계분석
Table 2. Descriptive statistics

	Variable	Obs.	Mean	S.D.	Min.	Max.	VIF
종속변수 Dependent variable	주중 생활인구(9-18시) A weekday average daily de facto population (9:00-18:00)	17,196	4064.26	10203.76	8.39	406839.44	-
	주말 생활인구(9-18시) A weekend average daily de facto population (9:00-18:00)	17,196	1515.83	2958.01	4.32	81108.04	-
토지이용 특성 Land use characteristics	주거용도 연면적 밀도(m^2/m^2) Density of total floor area (residential)	17,196	1.50	1.41	0.00	20.73	1.59
	상업용도 연면적 밀도(m^2/m^2) Density of total floor area (commercial)	17,196	0.16	0.34	0.00	13.86	1.65
	업무용도 연면적 밀도(m^2/m^2) Density of total floor area (office)	17,196	0.09	0.43	0.00	17.99	1.25
	토지이용 혼합도 Land use mix	17,196	0.30	0.30	0.00	0.99	2.13
POI시설 특성 POI facilities characteristics	공공기관시설 밀도(개/km^2) Density of public institutions facilities (Num./km^2)	17,196	13.39	45.75	0.00	1856.12	1.05
	교육시설 밀도(개/km^2) Density of educational facilities (Num./km^2)	17,196	93.71	160.57	0.00	3597.35	1.21
	편의시설 밀도(개/km^2) Density of convenience facilities (Num./km^2)	17,196	342.25	495.48	0.00	13854.17	2.21
	금융시설 밀도(개/km^2) Density of finance facilities (Num./km^2)	17,196	14.43	63.01	0.00	2168.60	1.40
	레저시설 밀도(개/km^2) Density of leisure facilities (Num./km^2)	17,196	25.96	62.49	0.00	1266.92	1.48
	문화시설 밀도(개/km^2) Density of culture facilities (Num./km^2)	17,196	27.81	134.21	0.00	14731.29	1.25
	병원시설 밀도(개/km^2) Density of hospital facilities (Num./km^2)	17,196	46.96	134.09	0.00	2461.28	1.59
	산업시설 밀도(개/km^2) Density of industrial facilities (Num./km^2)	17,196	210.51	364.61	0.00	9032.73	1.64
	여행시설 밀도(개/km^2) Density of trip facilities (Num./km^2)	17,196	4.69	26.73	0.00	2476.68	1.03
	음식점시설 밀도(개/km^2) Density of dining facilities (Num./km^2)	17,196	206.61	373.16	0.00	11062.17	2.08
교통 특성 Transportation characteristics	교차로 밀도(개/km^2) Density of intersections (Num./km^2)	17,196	17.27	42.25	0.00	1161.31	1.09
	버스정류장 밀도(개/km^2) Density of bus station (Num./km^2)	17,196	28.13	55.95	0.00	1161.31	1.10
	지하철역까지의 거리(m) Distance to the subway station (m)	17,196	622.27	407.22	5.81	4071.93	1.06
보행환경 특성 Walking environment characteristics	보행로 비율 Sidewalk ratio	17,196	3.11	1.55	0.00	17.79	1.19
	녹시율(GVI) Green view index	17,196	15.93	13.10	0.00	82.85	1.44

〈표 3〉은 POI시설의 종류와 종류별 차지하는 POI 현황을 포함하였다. 기본적으로 카카오맵에서 제공하는 분류를 활용하였으며 이 중에 도시활력과 관련이 없는 POI는 분석에서 제외하였다. 분석대상 지역의 POI시설 현황을 살펴보면 편의, 산업, 음식점과 교육 시설이 차지하는 비율이 상대적으로 높게 나타났다. 또한, 데이터를 구득한 시점은 2020년 6월부터 8월까지이며, 각 시설은

표 3. POI 빅데이터의 종류
Table 3. Types of POI big data

Types	Detailed facilities	Num. of facilities (%)
Public institutions	Community service center, police agency, court, ministry of education, etc.	6,498 (1.52)
Education	Private institutions, library, school, etc.	34,476 (8.04)
Convenience	Beauty salon, furniture sales, clothes, mart, etc.	131,470 (30.67)
Finance	Banks, insurance, finance, etc.	9,677 (2.26)
Leisure	Basket ball club, football field, golf club, hiking, etc.	11,162 (2.60)
Culture	Theater, performing art, museums, etc.	14,077 (3.28)
Hospital	Pharmacy, dental hospital, health clinic, etc.	20,819 (4.86)
Industry	Factory, telecommunication, construction, architecture, etc.	104,943 (24.48)
Trip	Tourist attractions, guest house, etc.	3,873 (0.90)
Dining	Cafe, western food, Korean food, etc.	91,673 (21.39)

Source: Types of POI are based on existing studies (Huang et al., 2020; Li et al., 2020; Tang et al., 2018; Wu et al., 2018b).

집계구별 밀도로 계산하였다. 〈그림 2〉는 서울시 성동구 왕십리 지역을 예시로 카카오맵에서 POI시설의 종류와 공간적 분포를 나타낸 그림이다. 그림을 보면 한 건물에서 다양한 시설을 확인할 수 있고, 이를 통해 POI시설 데이터는 건물의 세부적인 이용을 대표할 수 있음을 알 수 있다.

교통 특성 변수는 집계구별 지하철역까지의 거리, 교차로 수와

그림 2. POI 종류와 공간적 분포(서울 왕십리 지역 예시)
Figure 2. POI types and spatial distribution (Example of Wangsimni area in Seoul)

버스정류장 수를 고려하였다. 우선 서울열린데이터광장에서 제공되는 교차로 데이터와 지하철역 위치자료 그리고 버스정류소 위치 데이터를 활용하였다. 각 집계구 중심점에서 가장 가까운 지하철역까지의 거리를 계산하였고, 집계구의 면적을 고려해 교차로 수의 밀도와 버스정류장 수의 밀도를 계산하였다.

마지막으로 본 연구에서는 보행환경 요소를 산출하기 위해 가로경관 빅데이터인 NSV 파노라마 이미지를 네이버 지도에서 API를 이용하여 구득하였다. 우선, NSV 파노라마 이미지를 구득하기 위해 Open Street Map(OSM) 도로망 데이터를 따라 10m 간격으로 지점(Point)을 생성하였으며, 해당 지점들을 대상으로 경위도 좌표(WGS84)를 산출하였다. 이후 산출된 경위도 좌표와 API를 활용하여 개별 지점들에 대한 629,467장 NSV 파노라마 이미지를 구득하였다. 또한, 본 연구가 활용하고자 하는 NSV 이미지의 경우, 계절 요인과 다른 변수와의 구축 시점을 고려해 2020년 4, 5, 6월 이미지를 사용하였다.

구득한 NSV 이미지를 가지고 딥러닝 의미론적 분할 기법(Semantic Segmentation)을 통해 보행환경의 경관 요소를 도출하였다. 의미론적 분할은 픽셀 단위로 각 픽셀의 이미지 구성요소가 어떤 요소인지 구별할 수 있는 기법이다. 의미론적 분할 기법은 다양한 모형이 존재하지만 본 논문은 벤치마크(Benchmark)에서 높게 나타나는 모형(OCRNet+HRNet48)과 ADE20K 데이터 세트를 사전 학습한 모형을 활용하여 분석하였다. 의미론적 분할 기법으로 도출된 NSV 파노라마 이미지의 예시 사진은 〈그림 3〉과 같다. 본 논문에 사용되는 녹시율(Green View Index: GVI) 변수는 식 (1)과 식 (2)와 같이 이미지 전체 픽셀 수 대비 녹지로 분류된 이미지 픽셀 수로 나타낼 수 있으며, 3차원 공간에서 인지되는 녹지의 양을 의미한다(Li et al., 2015). 본 분석에서 사용한 ADE20K 데이터 세트의 Tree, Grass, Plant 비율의 합계를 활용하여 녹시율로 구축하였다. 즉, 이미지 내 녹시율로 분류된 픽셀은 0에서 100 사이의 값을 가지며, 그 값이 클수록 해당 지점에서 시각적으로 보이는 녹지의 비율은 높다는 것을 의미한다(Ki and Lee, 2021). 또한, 동일한 방법으로 가로경관 이미지에 포함된 보행로 비율을 측정하여 분석에 활용하였다.

$$\text{녹지픽셀수} = tree\text{픽셀수} + grass\text{픽셀수} + plant\text{픽셀수} \tag{1}$$

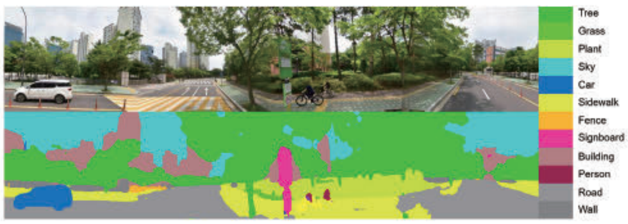


그림 3. 네이버 파노라마 이미지와 딥러닝 의미론적 분할 결과
Figure 3. Naver panorama image and deep learning semantic segmentation result

$$GVI = \frac{\text{녹지로 분류된 픽셀 수}}{\text{전체 픽셀 수}} \times 100 \qquad (2)$$

4. 분석과정 및 분석방법

1) 기계학습 방법론

본 연구는 도시활력에 대한 예측을 위해 전통적인 OLS 회귀분석 모형과 다양한 유형의 기계학습 모형을 검토하고 그중에서 설명력이 가장 높은 모형을 제시하고자 한다. 구체적으로, Random Forest(RF)는 앙상블 방법으로 배깅(Bagging)을 사용하고 결정나무(Decision Tree)가 모여서 랜덤 포레스트(Random Forest)를 구성한다. RF는 결정 나무를 독립적으로 훈련하고 훈련 세트에서 무작위 집합을 사용하기 때문에 과적합(Overfitting) 문제를 방지할 수 있다는 장점이 있다(Liaw and Wiener, 2002).

다음으로 Gradient Boosting Regression(GBR) 모형은 한 번에 하나의 트리를 구축하며 각각의 새로운 트리는 모든 이전의 트리가 도출된 잔차(Residual)를 조정함으로써 최종 모형의 예측 성능을 높이는 방법이다(Friedman, 2001). GBR 모형은 잔차를 줄일 수 있지만 과적합이 일어날 수도 있다는 단점이 있다.

반면 Extreme Gradient Boosting(XGBoost) 모형은 GBR에 비해 데이터 세트의 훈련 속도가 더 빠르고 정규화(Regularization) 방식으로 과적합 문제를 피할 수 있는 장점을 가지고 있다(Chen and Guestrin, 2016). 따라서 XGBoost 모형은 다양한 분야에서 높은 정확도를 보이는 기계학습 모형 중 하나이다. 마지막으로 심층신경망(Deep Neural Network, DNN) 모형은 앙상블 학습법과 달리 입력층(Input Layer)과 출력층(Output Layer) 사이에 여러 개의 은닉층(Hidden Layer)으로 이뤄진 인공신경망(Artificial Neural Network, ANN)을 활용하는 방법이다(Schmidhuber, 2015).

이와 같이 다양한 유형의 기계학습 방법론이 활용되고 있으며, 각 방법론마다 장단점을 가지고 있다. 따라서 본 연구는 서울시 집계구 단위의 도시활력을 분석하기 위해서 대표적인 기계학습 방법론인 RF, GBR, XGBoost, 딥러닝 모형을 비교분석하고 가장 설명력이 높은 기계학습 모형을 도출하고자 한다.

2) 해석가능한 기계학습 방법론

기계학습 모형은 설명력이나 예측의 정확도를 향상시킬 수 있지만, 모형 내부 메커니즘의 복잡성이 증가함에 따라 과정을 이해하는 데 어려움이 있다. 이러한 문제를 해결하기 위해 기계학습의 분석 결과를 설명할 수 있는 SHAP 방법론이 다양한 연구에서 활용되고 있다(Lundberg and Lee, 2017; Kim and Lee, 2023).

도시활력에 영향을 미치는 변수의 기여도를 도출하고 변수 간의 구체적인 비선형 관계를 확인하기 위해 SHAP 방법론을 활용하였다. SHAP 방법론을 통해 도출되는 Shapley Value는 가능한 모든 변수의 조합에 대해서 개별 변수의 기여도를 종합적으로 평가한 값이며, 이는 해당 독립변수가 종속변수를 예측할 때 얼마나 기여할 수 있는지에 대해 정량화할 수 있는 지표이다(Lundberg and Lee, 2017).

Ⅳ. 분석결과

1. 도시활력 분석 모형 평가 결과

도시활력을 대표하는 생활인구의 분석모형은 Python 기계학습 관련 라이브러리(Library)인 Scikit-learn을 활용하였다. 모형 간의 예측력을 객관적으로 비교하기 위해 수집된 전체 표본 17,196개는 학습 표본과 평가 표본으로 분할하였다. 이때 학습 표본은 전체의 80%인 13,756개, 평가 표본은 20%인 3,440개로 할당하였다. 구체적인 분석 순서는 먼저 학습 표본을 통해 각 모형을 학습하였으며 학습한 모형들에 대해 평가 표본들의 속성값(독립변수)을 통해서 생활인구 수를 예측하여 정확도를 평가하였다.

일반 OLS 모형과 4가지 기계학습 모형을 비교 분석하기 위해 선행연구에서 주로 사용하고 있는 결정계수(R^2), 평균 제곱근 오차(RMSE), 그리고 평균 절대 오차(MAE)를 평가지표를 활용하였다(Chen et al., 2020). 〈표 4〉의 결과에 따르면 주중 생활인구의 결정계수(R^2) 값은 0.120~0.565로 나타났고, RMSE는 6.375~8.362, MAE는 2.025~3.082로 나타났다. 한편, 주말 생활인구의 결정계수(R^2) 값은 0.226~0.564이고, RMSE는 1.811~2.424, MAE는 0.726~1.060의 값으로 주중 생활인구보

표 4. 모델 평가
Table 4. Model evaluation

Dependent variable	Model	MAE	RMSE	R^2
주중 생활인구 (Weekday de facto population)	Linear regression	3.082	8.362	0.120
	LightGBM	**2.025**	**6.375**	**0.565**
	Random forest	2.076	6.443	0.514
	Gradient boosting	2.047	6.375	0.528
	XGBoost	2.038	6.377	0.528
주말 생활인구 (Weekend de facto population)	Linear regression	1.060	2.424	0.226
	LightGBM	**0.726**	**1.811**	**0.564**
	Random forest	0.773	1.928	0.502
	Gradient boosting	0.761	1.856	0.541
	XGBoost	0.765	1.903	0.328

다 오차가 적은 것으로 나타났다.

분석결과 네 가지 기계학습 모형의 평가결과는 모두 일반 선형 회귀 모형에 비해 좋은 성능을 보여주는 것으로 나타났다. 특히 다섯 가지 모형 중 가장 설명력이 높은 모형은 LightGBM으로 나타났다. LightGBM 모형은 균형 트리 분할(Leaf Wise Tree) 방식으로 오차가 큰 노드를 중심으로 나누면서 예측에 대한 오차를 최소화하는 방법으로 학습시간이 빠른 장점을 가지고 있다. 따라서 LightGBM 모형을 최종모형으로 활용하였으며, 해석가능한 기계학습 방법론을 활용하여 도시활력에 영향을 미치는 설명변수의 비선형관계를 분석하였다.

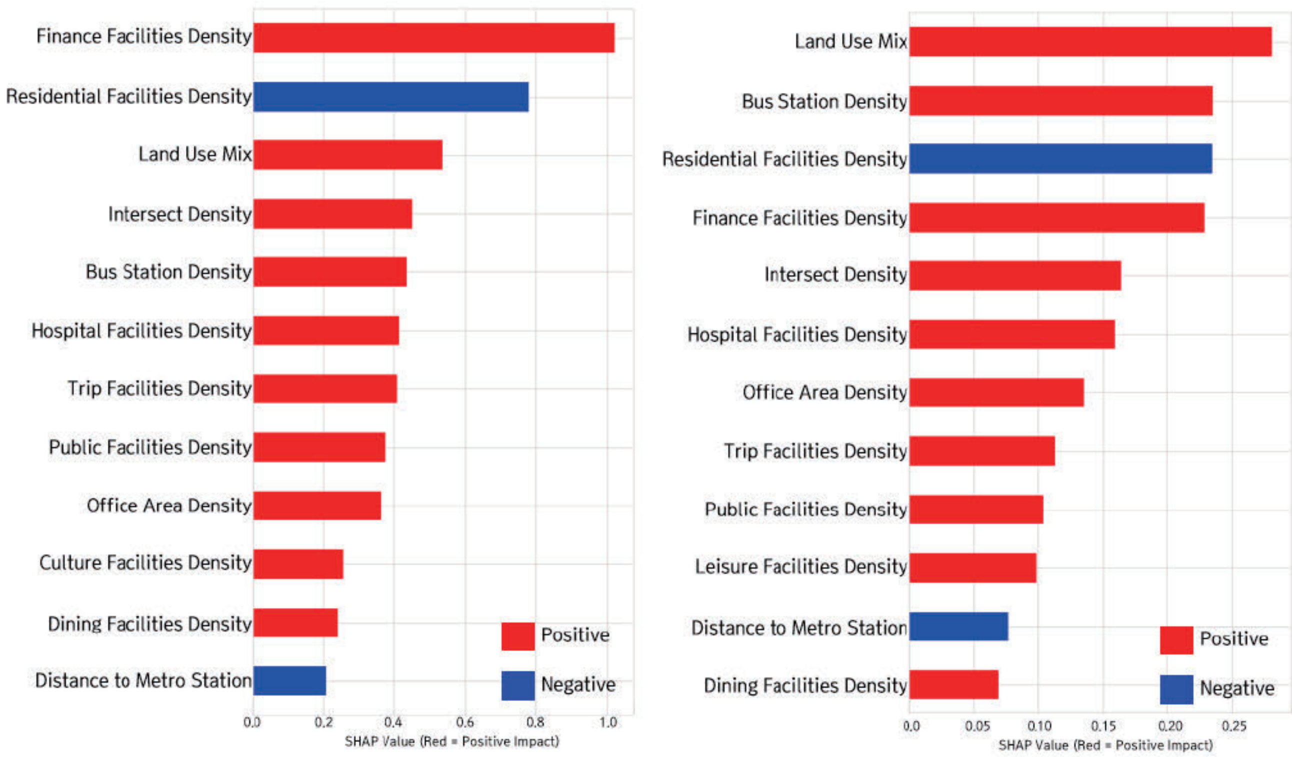
2. SHAP 변수의 기여도 분석 결과

본 연구는 각 변수의 중요성을 확인하기 위해 훈련된 LightGBM 모형을 기반으로 학습하고, 기계학습을 해석할 수 있는 SHAP 방법론으로 Shapley Value(기여도)를 도출하였다. Shapley Value는 전체 데이터에서 각 변수에 대한 절대적인 기여도를 설명할 수 있으며, 각 변수의 중요도를 나타낸 지표이다. 이에 따라 각 변수의 기여도를 확인하면 〈그림 4〉와 같다.

〈그림 4(a)〉는 주중 생활인구에 미치는 변수 기여도이고, 〈그림 4(b)〉는 주말 생활인구에 미치는 변수 기여도를 의미한다. X축은 각 변수의 Shapley Value로 볼 수 있고, Y축은 집계구의 단위로 해당 변수의 기여도를 표시하였다. 변수의 순서는 총기여도의 값에 따라 정렬하였다. 그리고 특정 변수의 색상이 빨간색의 막대 그래프일 때 생활인구와 양(+)의 관계를 의미하고, 파랑의 경우 생활인구와 음(-)의 관계를 의미한다.

구체적으로 주중 생활인구(그림 4(a))의 경우 금융시설 밀도, 주거시설 밀도, 토지이용 혼합도, 교차로 밀도, 버스정류장 밀도 순으로 기여도가 높은 것으로 나타났다. 이는 금융시설 밀도가 높고, 토지이용이 혼합되거나, 교차로의 밀도가 높을수록 도시활력과 양(+)의 관계를 보이는 것을 확인할 수 있다. 그리고 음(-)의 관계를 보이는 변수로 주거시설 밀도가 높고, 지하철역까지의 거리가 길수록 주중활력이 낮아지는 관계로 나타났다.

한편, 주말의 경우 토지이용 혼합도가 가장 중요한 요인이고, 버스정류장 밀도, 주거시설 밀도, 금융시설 밀도, 교차로 밀도 순으로 도시활력에 미치는 영향력이 크게 나타났다. 음(-)의 관계 변수는 주중과 같은 주거시설 밀도, 지하철역까지의 거리로 나타났다. 주중과 주말의 생활인구를 비교했을 때, 주중은 사람들이 금융시설과 같은 업무공간에서 근무하기 때문에 주로 금융시설이 높은 집계구에서 도시활력이 높은 것으로 나타났다. 반면에 주말은 주로 주거, 업무, 상업의 토지이용이 혼합된 집계구일수록 도시활력이 높은 것으로 나타났다. 이는 선행연구에서 토지이용이 혼합된 도시일수록 도시의 활력과 양(+)의 관계를 보인다는 연구와 동일한 결과로 볼 수 있다(조월 외, 2021).



(a) 주중 생활인구 변수 기여도
The contribution of weekday de facto population variable

(b) 주말 생활인구 변수 기여도
The contribute of weekend de facto population variable

그림 4. SHAP로 도출된 각 변수의 기여도(Global 값)
Figure 4. Contribution of each variable derived by SHAP (Global value)

3. 주중 · 주말 비선형 관계 분석 결과

앞서 사용한 SHAP로 도출한 전역적 변수 기여도(Global Feature Importance)는 도시 환경변수와 도시활력 간의 복잡한 비선형 관계를 구체적으로 설명할 수 없으므로 추가적으로 국지적 변수 기여도(Local Feature Importance)를 통해서 독립변수가 도시활력과 어떤 연관성이 있는지 분석하였다. 구체적으로 본 논문은 LightGBM 모형을 기반으로 독립변수와 주중, 주말 도시활력 사이의 관계를 설명하고자 하였다. X축은 해당 독립변수의 값이고, Y축의 각 독립변수는 도시활력을 예측할 때 상대적인 낮 시간대의 일평균 생활인구 수에 미치는 변화량을 나타낸다. 〈그림 5〉는 SHAP의 기여도를 바탕으로 도시환경 변수들이 주중 낮 시간대 일평균 생활인구와의 관계를 나타내며, 〈그림 6〉은 도시환경 변수와 주말 낮 시간대 일평균 생활인구 관계를 설명한다. 특히, 변수의 변화가 도시활력에 미치는 영향이 어떻게 변화하는지 확인하기 위해서 Cleveland(1979)의 Locally Weighted Scatterplot Smoothing(LOWESS) 기법을 활용하여, 전체적인 경향을 확인하고, 이러한 추이를 통해서 비선형 임곗값을 확인하였다.

대부분의 근린환경 변수와 도시활력 사이에는 비선형적인 관계가 나타났다. 우선 〈그림 5(a)〉와 〈그림 6(a)〉에서 주거용도 건물의 연면적 밀도는 주거 연면적 밀도가 약 3(m^2/m^2)까지는 도시활력이 증가하고, 이후 증가폭은 감소하는 것으로 나타났다. 주거용도 건물 연면적은 거주지에 있는 건물 연면적이 증가할수록 사람이 거주할 가능성은 커지지만, 일정 수준을 초과하면 도시활력에 미치는 상관관계는 크지 않다는 것을 의미한다.

한편, 토지이용 혼합도(그림 5(b), 그림 6(b))에서 도시활력 간의 관계는 약 0.4를 기점으로 도시의 활력이 증가하는 경향을 보였다. 특히, 주말에서 증가폭은 더 크게 나타났다. 이러한 결과는 높은 토지이용 혼합 정도를 가지고 있는 지역은 사람들이 서로 다른 시간대에 활동할 수 있으며(Jacobs, 1961), 토지이용이 혼합될 때 사람들이 다양한 사회활동과 경제활동을 할 수 있다. 결과적으로 토지이용 혼합도가 지역의 도시활력에 긍정적인 영향을 미친다는 선행연구와 맥락을 같이한다(조월 외, 2021). 또한, 임곗값을 중점적으로 보면 서울시의 토지이용 혼합 구간이 0.4전까지 사람들의 일상 수요를 충분히 충족시키지 못하고 있으며, 토지이용 혼합 엔트로피 지수가 0.4를 넘어야 더 많은 사람이 해당 공간에서 다양한 활동이 가능해진다고 볼 수 있다. 따라서 도시활력을 높이기 위해서는 단일 용도보다는 다른 용도의 시설과 적절한 토지이용 혼합이 필요함을 시사한다.

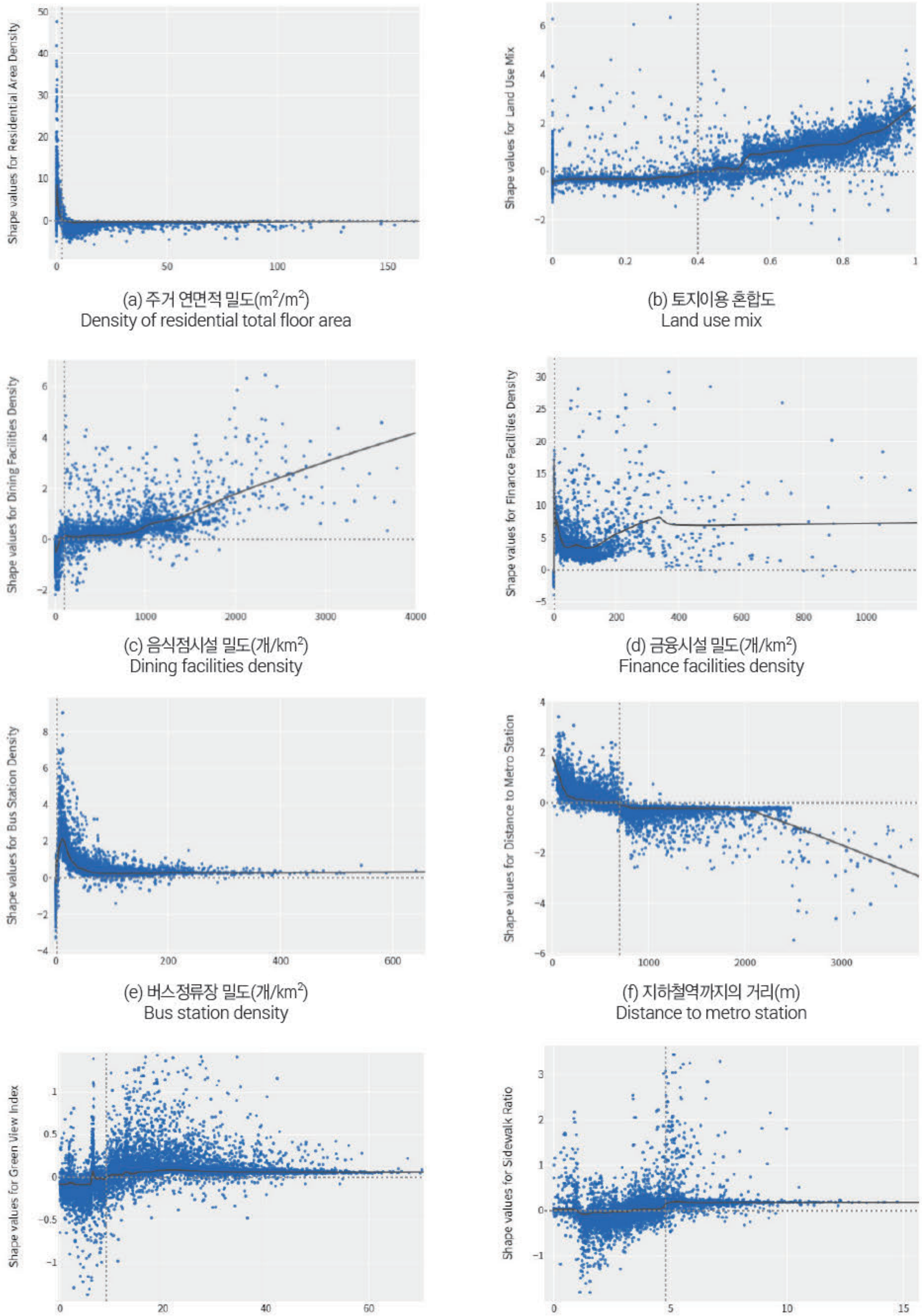
다음으로, 〈그림 5(c)〉와 〈그림 6(c)〉에서 도시활력과 음식점 시설을 비선형 관계를 비교했을 때 음식점시설의 밀도가 높을수록 도시활력의 기여도가 증가하는 것으로 나타났다. 이는 음식점 시설의 밀도가 높을수록 주중 낮 시간대의 생활인구 밀도가 높아지는 것으로 나타났다. 특히, 주중에서는 음식점 밀도가 100을 중심으로 증가하고, 800 이상인 지역에서 증가폭은 커지는 경향으로 나타났다. 음식점과 같은 업체의 생존은 보행자가 밀집된 도시 활동에 달려 있고(Ye et al., 2018), 음식점시설의 밀도가 높을수록 집중 효과를 일으키기 때문에 다른 지역의 사람들을 끌어올 수 있는 시설로 해석할 수 있다. 이러한 결과는 서울시 대상으로 가로활력에 영향을 미치는 상업업종 중 음식업이 가장 중요하다고 했던 연구의 결과와 맥락을 같이한다(임하나 외, 2017).

금융시설밀도와 도시활력과의 관계(그림 5(d), 그림 6(d))를 보면 금융시설의 밀도가 변화하는 임곗값은 그래프에서 확인할 수 없으나, 금융시설 밀도가 높을수록 생활인구가 많은 것으로 나타났다. 특히 금융시설 밀도가 주중 생활인구에 미치는 기여도가 15인 반면, 주말 생활인구는 3.8로 나타났다. 이는 금융시설의 경우 주로 업무 용도이기 때문에 주말과 비교했을 때, 주중 도시활력에 미치는 영향력이 큰 것을 의미한다.

한편, 〈그림 5(e)〉와 〈그림 6(e)〉에서 대중교통의 접근성은 모두 비선형적인 관계이고, 임곗값을 가지는 것으로 나타났다. 구체적으로 버스정류장 밀도는 2개/km^2 정도까지 주중 도시활력이 급격하게 증가하고, 17개/km^2 정도부터 증가폭이 감소하다가 60개/km^2 정도부터는 증가폭이 일정해지는 경향이 있는 것으로 나타났다. 이는 대중교통의 밀도가 60개/km^2을 넘어선 이후 주중 도시활력의 변화가 일정한 것은 사람들의 이동 수요가 보편적으로 충족되면 활력을 촉진하기 위해 대중교통의 밀도를 더 높이는 것이 효과가 크지 않은 것을 볼 수 있다.

그리고 지하철역까지의 거리와 도시활력 간의 관계가 음(-)의 방향으로 나타났다(그림 5(f), 그림 6(f)). 이는 지하철역과 거리가 멀어질수록 도시활력이 감소한다는 것을 의미한다. 특히 지하철역까지의 거리가 650~700m 이내일 경우 지속적으로 도시활력에 긍정적인 영향을 미칠 수 있으며, 이는 해당 지역이 지하철역과 가까워지면서 더 큰 교통 편의를 제공함으로써 시민들이 해당 지역에 방문 및 다양한 일상활동을 장려할 수 있다. 이러한 결과는 Yang et al.(2021)의 연구와 비슷한 맥락으로 나타났고, 서울시는 도시활력에 있어 역세권 범위는 650~700m로 볼 수 있음을 시사한다.

보행환경의 경관요소의 녹시율(그림 5(g), 그림 6(g))은 도시활력과 양(+)의 관계를 가지는 것으로 나타났으며, 녹시율은 10%를 기점으로 도시활력에 긍정적인 영향을 주며, 23% 이상이면 도시활력에 미치는 영향력이 일정해지는 경향이 나타났다. 이에 대한 이유로, 가로변에 존재하는 녹지가 잘 갖춰진 지역의 경우, 쾌적한 걷기 환경으로 다양한 목적의 일상활동이 보행을 통해 이루어질 수 있고 공원의 이용도를 증진시킴으로써(Ki and Lee, 2021) 도시활력의 증가에 영향을 미치는 데 도움이될 수 있는 것으로 해석할 수 있다. 하지만 녹지의 비율이 23% 정도 넘어선 이후 도시활력에 미치는 영향력의 증가폭이 일정하게 나타났다. 이러한 결



(a) 주거 연면적 밀도(m^2/m^2)
Density of residential total floor area

(b) 토지이용 혼합도
Land use mix

(c) 음식점시설 밀도(개/km^2)
Dining facilities density

(d) 금융시설 밀도(개/km^2)
Finance facilities density

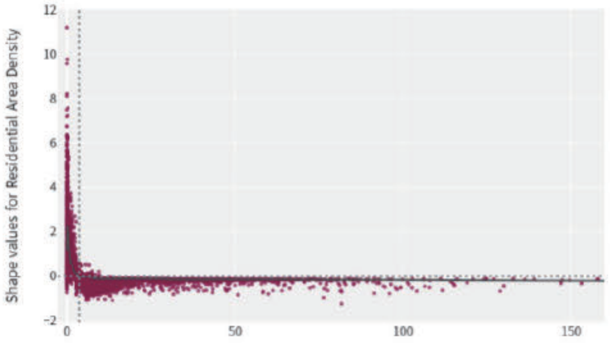
(e) 버스정류장 밀도(개/km^2)
Bus station density

(f) 지하철역까지의 거리(m)
Distance to metro station

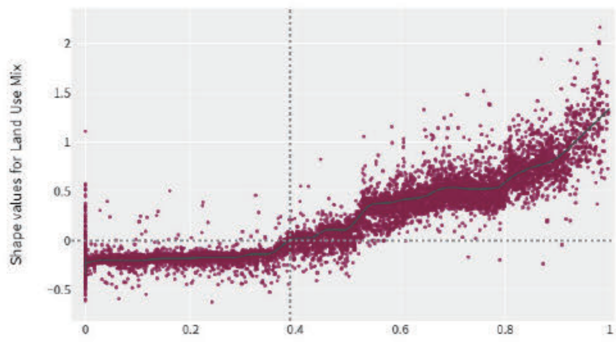
(g) 녹시율
Green view index

(h) 보행로 비율
Sidewalk ratio

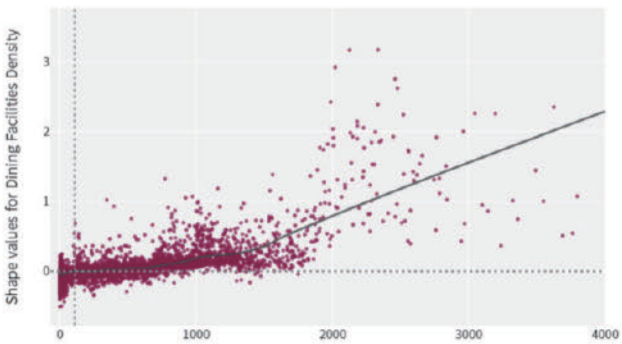
그림 5. 주중 도시활력과 근린환경 변수의 비선형 관계
Figure 5. Nonlinear relationship between weekday urban vitality and neighborhood variables



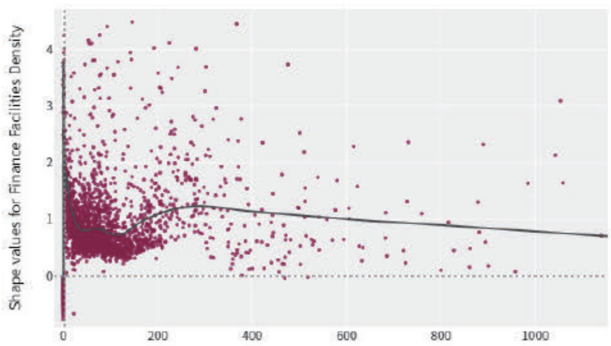
(a) 주거 연면적 밀도(m^2/m^2)
Density of residential total floor area



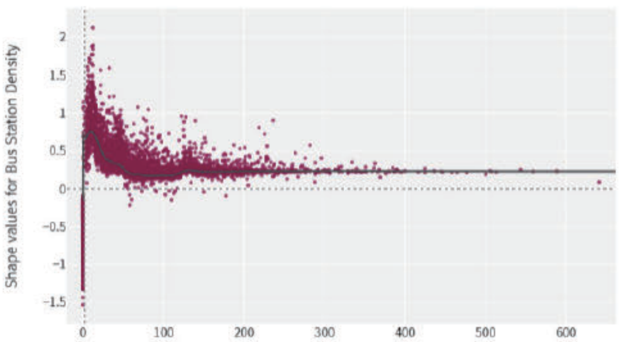
(b) 토지이용 혼합도
Land use mix



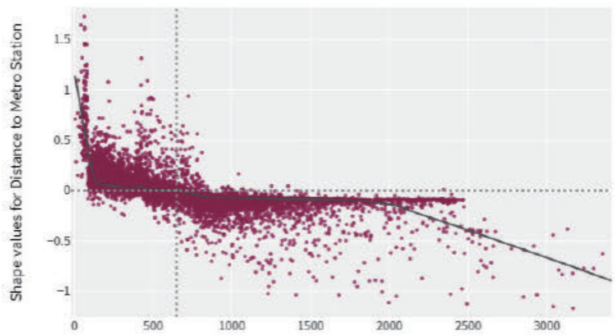
(c) 음식점시설 밀도(개/km^2)
Dining facilities density



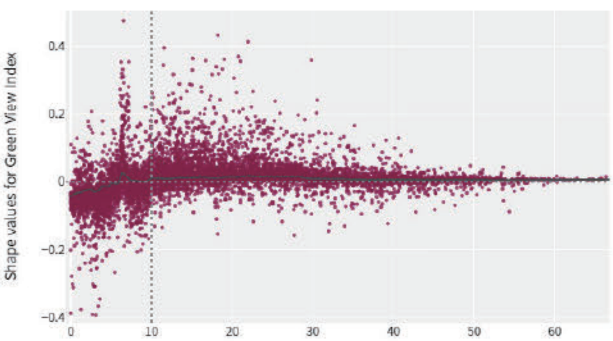
(d) 금융시설 밀도(개/km^2)
Finance facilities density



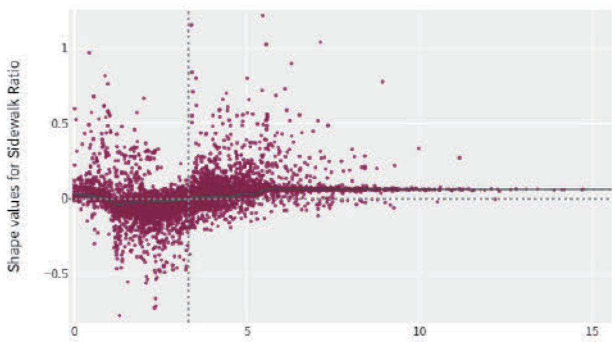
(e) 버스정류장 밀도(개/km^2)
Bus station density



(f) 지하철역까지의 거리(m)
Distance to metro station



(g) 녹시율
Green view index



(h) 보행로 비율
Sidewalk ratio

그림 6. 주말 도시활력과 근린환경 변수의 비선형 관계
Figure 6. Nonlinear relationship between weekend urban vitality and neighborhood variables

과는 Yang et al.(2021)의 연구에서 도출된 녹시율이 24%보다 작아야 보행경향(Walking Propensity)과 양(+)의 관계가 있으며, 녹시율이 24%를 초과하면 보행경향에 미치는 영향은 제한적이라는 결과와 같은 맥락이다. 가로 녹지의 비율이 높은 지역의 경우 자연상태의 숲이나 사람들이 접근할 수 없는 녹지이기 때문에 가로활력과 유의미한 관계가 나오지 않은 것으로 판단된다.

마지막으로, 보행로 비율(그림 5(h), 그림 6(h))은 4% 정도 이상에서 도시활력이 증가하고 이후 영향력의 변동이 크지 않게 나타났다. 가로활력과 관계가 있는 보행로 비율이 일정수준 이상으로 증가할 경우 보행로의 비율을 더 높이는 것은 가로활력과는 큰 관계가 없다는 것을 확인할 수 있다. 녹지와 보행로 변수의 결과를 종합해 보면 도시공간에서 시각적으로 볼 수 있는 녹지와 보행로의 비율이 높을수록 도시활력이 높을 수 있지만, 일정 비율 이상이 되면 낮 시간대의 생활인구 수와 유의미한 관계가 없다는 것을 알 수 있다.

V. 결 론

본 연구는 서울시를 대상으로 2019년 생활인구 자료를 토대로 주중과 주말의 낮 시간대 일평균 생활인구 수를 종속변수로 활용하였다. 생활인구에 미치는 근린환경을 찾기 위해 독립변수는 도시 빅데이터를 활용하여 토지이용 특성, POI 시설 특성, 교통 특성, 보행환경 특성 총 4가지 변수 유형을 고려하였다. 나아가, 기계학습 방법론으로 도시활력을 정량적으로 예측하는 모형을 구축하였고, 해석가능한 방법론을 활용하여 도시활력에 영향을 미치는 변수의 기여도를 도출하고 변수 간의 비선형 관계를 확인하였다. 연구의 분석결과와 정책적 시사점은 다음과 같다.

첫째, 도시활력과 근린환경 요인 간의 관계에서 비선형적 패턴이 존재함을 확인하였다. 기존 연구에서는 선형 관계를 가정하는 경우가 많았으나, 본 연구는 도시의 다양한 환경 요인과 도시활력은 특정 임곗값을 기준으로 변화하는 비선형관계가 있음을 보였다. POI시설의 경우, 각 시설 특성에 따라 도시활력 간의 관계가 다르게 나타났다. 특히, 음식점시설은 사람들이 자주 이용하는 시설로, 밀도가 높아질수록 도시활력이 증가하는 경향을 확인할 수 있다. 반면, 금융시설은 주로 주중과 연관있는 업무시설이기 때문에 주중 도시활력과 연관 있는 것으로 나타났다. 이는 특정 POI 시설의 밀도가 주중, 주말의 도시활력에 다른 영향을 미치고, 각 시설의 특성에 따라 도시활력과 관련된 밀도를 조절할 필요가 있음을 시사한다.

또한, 토지이용과 관련된 변수나 대중교통의 접근성은 도시활력과 비선형 관계를 가지는 것을 확인하였다. 구체적으로, 주거지역 연면적의 밀도와 지하철역까지의 거리가 증가할수록 도시의 활력이 감소하지만, 임곗값이 일정 수준(주거지역 연면적의 밀도 약 3(m^2/m^2); 지하철까지의 거리 약 650~700m)을 넘어가면 도시활력과의 관계는 일정한 것으로 나타났다. 이러한 결과는 Yue et al.(2019)의 논문 결과와 유사하고, 서울시 도시활력의 측면에서 보면 지하철역의 영향권이 650~700m에 이르며, 이는 서울시 역세권 개발 정책을 설정할 때 실증적 근거로 고려할 수 있음을 시사한다.

그리고 토지이용 혼합도의 비선형 관계를 확인해 본 결과 음(-)의 관계에서 양(+)의 관계로 변화하는 임곗값이 도출되었다. 구체적으로 토지이용 혼합 정도는 0.4 이상 단위에서 도시활력이 양(+)의 관계로 변화하는 것을 볼 수 있다. 이와 같은 결과는 토지이용의 측면에서 사람들의 통근·서비스·생활 등이 혼합되어 있을 때 더 활력이 있고 안전한 도시가 될 수 있다는 Jacobs(1961) 주장에 대한 실증적 근거가 될 수 있다.

그리고 버스정류장 밀도의 경우, 전체적으로 도시활력과 비선형적인 관계가 있는 것으로 확인되었다. 버스정류장의 밀도가 증가함에 따라 도시활력에는 양(+)의 관계가 강해졌으나, 임곗값(60개/km^2)을 넘어선 이후 도시활력에 미치는 영향력은 일정한 추세를 보였다. 이러한 결과는 버스정류장의 밀도가 임곗값 이전까지는 자동차보다 대중교통을 통해 사람들의 활동을 장려할 수 있음을 시사한다.

도시환경에서 시각적으로 보이는 녹시율과 보행로 비율의 경우, 일정 수준 이상(녹시율 약 24%; 보행로 비율 약 4%)일 때 도시활력과 유의미한 관계가 나타나지 않았다. 이는 녹시율과 보행로 비율은 낮 시간대 생활인구의 활동이 주로 이루어지는 장소와 관련이 크다는 것을 시사한다. 따라서 도시의 활력을 높이기 위한 전략은 단순히 녹지 공간을 확장하는 것이 아니라, 보행 환경의 질을 높이고, 보행 중심의 공공 공간을 조성하는 방안이 중요하다고 할 수 있다.

둘째, 도시활력과 관계가 있는 변수의 영향 순위를 확인하기 위해 SHAP 방법으로 각 변수의 기여도를 도출하였다. 구체적으로 주중은 금융시설 밀도, 주말은 토지이용 혼합도가 도시활력에 가장 큰 영향을 미치는 요인으로 나타났다. 주중과 주말 도시활력에 영향을 미치는 주요 요인이 다르게 나타났으므로, 주중과 주말 유동인구 특성을 반영한 도시활력 제고 방안이 필요하다. 추가적으로 해당 분석 결과를 바탕으로 Large Language Model(LLM)을 같이 활용한다면, 다양한 정량적인 수치뿐만 아니라 정성적인 요인을 바탕으로 도시활력을 측정할 수 있을 것으로 판단된다.

셋째, 본 연구는 5가지 모형을 비교함으로써 어떤 모형이 서울시 도시활력을 예측하는 데 더 적합한지 검증하였다. 트리(Tree) 구조의 4가지 기계학습 모형과 일반 선형 모형의 3가지 평가지표를 통해 각 모형의 성능을 평가하였다. 비교한 결과, 트리 구조의 기계학습 모형 모두 일반 선형회귀 모형보다 좋은 성능을 가지고 있는 것으로 나타났다. 특히, 제시된 4가지 기계학습 알고리즘 중에서는 LightGBM 모형의 예측 정확도가 가장 높은 것으로 나타났다.

반면, 본 연구는 다음과 같이 몇 가지 한계점을 가진다. 첫째, 서울시 POI시설 데이터는 각 시설의 개수로 구축되어 있기 때문에 해당 시설의 영향 효과와 건물의 연면적을 고려할 수 없는 한계점이 있다. 향후 연구에서 POI시설에 가중치를 추가하거나 토지이용 연면적 데이터와 조합해 도시환경을 더 잘 대표할 수 있는 변수를 구축하여 활용할 필요가 있다. 둘째, SHAP 방법론을 활용하여 도시활력에 영향을 미치는 변수별 기여도와 영향 관계를 도출하였지만, SHAP 방법론은 독립변수가 종속변수에 미치는 영향에 대한 인과관계는 해석할 수 없다는 한계점이 있다. 셋째, 도시환경 요소로 4가지 유형만 고려하였지만, 낮 시간대 도시활력에 영향을 미치는 다양한 유형의 변수에 대한 추가적인 고려가 필요할 것으로 판단된다. 해당 지역의 도시활력 정도는 주민의 인구, 사회경제적 특성, 생활만족도 등에 따라 차이가 있을 수 있다. 따라서 다양한 정성적 요인을 포함하여 도시활력 분석의 범위를 확장할 필요가 있다. 마지막으로, 연구는 데이터 구득의 한계로 최신 생활인구 데이터를 반영하지 못한 한계가 있다. 다만, COVID-19 팬데믹 이전의 정상적인 도시활동 패턴을 반영하기 위해 2019년 생활인구 데이터를 활용하였으며, 팬데믹 이후의 변동성을 배제하여 도시활력과 근린환경 간의 구조적 관계를 분석하고자 하였다.

그럼에도 불구하고, 본 연구는 도시 빅데이터와 기계학습 방법론으로 도시활력에 영향을 미치는 근린환경 요인을 분석했다는 점에서 의의가 있다. 또한, 기존의 선행연구에서 한계점으로 지적된 비선형관계를 주요 설명변수로 확인하였고, 구체적인 정책적 시사점을 제공했다는 점에서 의의가 크다고 판단된다. 도시활력을 제고하기 위한 정책개발에 있어, 도시활력에 영향을 미치는 설명변수의 비선형관계에 대한 정확한 이해는 효과적인 정책개발에 유용하게 활용될 수 있을 것으로 판단된다.

주1. 2019년 생활인구 데이터는 2016년의 집계구(19,163개) 기준으로 집계되었다. 하지만, 독립변수인 네이버 이미지로 도출된 보행환경 요소는 도로가 없거나 서비스를 제공하지 않아서 1,957개 집계구의 데이터를 확인할 수 없다. 따라서 본 연구는 보행환경 요소 데이터가 없는 집계구를 제외하고 최종적으로 총 17,196개 집계구를 기본적인 분석단위로 사용하였다.

주2. 2019년 생활인구 데이터는 10월 15일부터 27일까지 총 13일의 데이터가 누락이 되었다. 따라서 본 연구는 총 352일의 생활인구 데이터를 활용하여 낮 시간대의 일평균 생활인구 수를 계산하였다.

인용문헌
References

1. 곽호찬·송지영·엄진기·김경태, 2018. "이동통신 자료를 활용한 대도시 유동인구 영향요인 분석", 「한국도시철도학회논문집」, 6(4): 373-381.
Kwak, H.C., Song, J.Y., Eom, J.K., and Kim, K.T., 2018. "A Study on Factors Influencing Floating Population Using Mobile Phone Data in Urban Area", *Journal of The Korean Society For Urban Railway*, 6(4): 373-381.

2. 서울특별시, 2018. 「서울 생활인구 추계 매뉴얼」, 서울.
Seoul Metropolitan Government, 2018. *Seoul Living Population Estimation Manual*, Seoul.

3. 이수기·고준호·이기훈, 2016. "근린환경특성이 보행만족도에 미치는 영향 분석: 서울서베이 2013년 자료를 중심으로", 「국토계획」, 51(1): 169-187.
Lee, S., Koh, J.H., and Lee, G.H., 2016. "An Analysis of Neighborhood Environment Affecting Walking Satisfaction – Focused on the 'Seoul Survey' 2013", *Journal of Korea Planning Association*, 51(1): 169-187.

4. 이지혜·김형중, 2019. "생활인구 데이터를 활용한 노인인구 공간적 분포 및 군집분석: 서울시를 중심으로", 「한국디지털콘텐츠학회논문지」, 20(7): 1365-1371.
Lee, J.H. and Kim, H.J., 2019. "Identification of Spatial Distribution of an Aged Population and Analysis on Characterization of the Cluster: Focusing on Seoul Metropolitan Area", *Journal of the Korea Digital Contents Association*, 20(7): 1365-1371.

5. 임하나·성은영·최창규, 2017. "상업시설의 다양성과 가로활력과의 관련성 실증분석 – 서울시 상업지역과 주거지역을 구분하여", 「도시설계」, 18(6): 37-49.
Im, H.N., Seong, E.Y., and Choi, C.G., 2017. "Relationship Between Diversity of Commercial Store and Street Vitality – By District Types in Seoul", *Journal of the Urban Design Institute of Korea Urban Design*, 18(6): 37-49.

6. 임하나·이수기·최창규, 2016. "서울시 토지이용 혼합유형과 보행량의 연관성 실증분석", 「국토계획」, 51(7): 21-38.
Im, H.N., Lee, S., and Choi, C.G., 2016. "Empirical Analysis of the Relationship Between Land Use Mix and Pedestrian Volume in Seoul, Korea", *Journal of Korea Planning Association*, 51(7): 21-38.

7. 정재훈·남진, 2019. "위치기반 빅데이터를 활용한 서울시 활동인구 유형 및 유형별 지역 특성 분석", 「국토계획」, 54(3): 75-90.
Jung, J.H. and Nam, J., 2019. "Types and Characteristics Analysis of Human Dynamics in Seoul Using Location-Based Big Data", *Journal of Korea Planning Association*, 54(3): 75-90.

8. 조월·하재현·이수기, 2021. "서울시 생활인구의 시간대별 혼합수준에 영향을 미치는 요인 분석", 「국토계획」, 56(1): 22-38.
Cao, Y., Ha, J.H., and Lee, S., 2021. "Analysis on the Determinants of Hourly-based Mixed Level of De Facto Population in Seoul, Korea", *Journal of Korea Planning Association*, 56(1): 22-38.

9. Buchanan, P., 1988. "What City? A Plea for Place in the Public Realm", *The Architectural Review*, 184(1101): 31-41.

10. Chen, L., Yao, X., Liu, Y., Zhu, Y., Chen, W., Zhao, X., and Chi, T., 2020. "Measuring Impacts of Urban Environmental Elements on Housing Prices Based on Multisource Data—A Case Study of Shanghai, China", *ISPRS International Journal of Geo-Information*, 9(2): 1-23.

11. Chen, T. and Guestrin, C., 2016. "Xgboost: A Scalable Tree

Boosting System", *Paper presented at the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining,* 16: 785-794, San Francisco California USA.

12. Cleveland, W.S., 1979. "Robust Locally Weighted Regression and Smoothing Scatterplots", *Journal of the American Statistical Association*, 74(368): 829-836.
13. Ding, C., Cao, X., and Næss, P., 2018. "Applying Gradient Boosting Decision Trees to Examine Non-linear Effects of the Built Environment on Driving Distance in Oslo", *Transportation Research Part A: Policy and Practice*, 110: 107-117.
14. Friedman, J.H., 2001. "Greedy Function Approximation: A Gradient Boosting Machine", *The Annals of Statistics*, 29(5): 1189-1232.
15. Gan, Z., Yang, M., Feng, T., and Timmermans, H.J., 2020. "Examining the Relationship Between Built Environment and Metro Ridership at Station-to-station Level", *Transportation Research Part D: Transport and Environment,* 82: 102332.
16. He, Q., He, W., Song, Y., Wu, J., Yin, C., and Mou, Y., 2018. "The Impact of Urban Growth Patterns on Urban Vitality in Newly Built-up Areas Based on An Association Rules Analysis Using Geographical 'Big Data'", *Land Use Policy*, 78: 726-738.
17. Huang, B., Zhou, Y., Li, Z., Song, Y., Cai, J., and Tu, W., 2020. "Evaluating and Characterizing Urban Vibrancy Using Spatial Big Data: Shanghai as a Case Study", *Environment and Planning B: Urban Analytics and City Science*, 47(9): 1543-1559.
18. Jacobs, J., 1961. *The Death and Life of Great American Cities*, New York: Vintage Books.
19. Jacobs-Crisioni, C., Rietveld, P., Koomen, E., and Tranos, E., 2014. "Evaluating the Impact of Land-use Density and Mix on Spatiotemporal Urban Activity Patterns: An Exploratory Study Using Mobile Phone Data", *Environment and Planning A*, 46(11): 2769-2785.
20. Ki, D. and Lee, S., 2021. "Analyzing the Effects of Green View Index of Neighborhood Streets on Walking Time using Google Street View and Deep Learning", *Landscape and Urban Planning*, 205: 1-11.
21. Kim, Y.L., 2018. "Seoul's Wi-Fi Hotspots: Wi-Fi Access Points as an Indicator of Urban Vitality", *Computers, Environment and Urban Systems*, 72: 13-24.
22. Kim, S. and Lee, S., 2023. "Nonlinear Relationships and Interaction Effects of an Urban Environment on Crime Incidence: Application of Urban Big Data and an Interpretable Machine Learning Method", *Sustainable Cities and Society*, 91: 104419.
23. Li, S.J., Wu, C., Lin, Y., Li, Z.Y., and Du, Q.Y., 2020. "Urban Morphology Promotes Urban Vibrancy from the Spatiotemporal and Synergetic Perspectives: A Case Study Using Multi-Source Data in ShenZhen, China", *Sustainability*, 12(12): 1-24.
24. Li, X., Zhang, C., Li, W., Kuzovkina, Y.A., and Weiner, D., 2015. "Who Lives in Greener Neighborhoods? The Distribution of Street Greenery and Its Association with Residents' Socioeconomic Conditions in Hartford, Connecticut, USA", *Urban Forestry & Urban Greening*, 14(4): 751-759.
25. Liaw, A. and Wiener, M., 2002. "Classification and Regression By RandomForest", *R News*, 2(3): 18-22.
26. Liu, S., Zhang, L., Long, Y., Long, Y., and Xu, M., 2020. "A New Urban Vitality Analysis and Evaluation Framework Based on Human Activity Modeling Using Multi-Source Big Data", *ISPRS International Journal of Geo-Information*, 9(11): 617.
27. Long, Y. and Zhou, Y., 2016. "Quantitative Evaluation on Street Vibrancy and Its Impact Factors: A Case Study of Chengdu", *New Archit*, 1: 52-57.
28. Lundberg, S. and Lee, S.I., 2017. "A Unified Approach to Interpreting Model Predictions", *ArXiv Preprint ArXiv*, 1705.07874: 1-10.
29. Lynch, K., 1981. *A Theory of Good City Form,* Cambridge, MA: MIT Press Books.
30. Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., Zhang, C., and Ren, F., 2021. "Measuring Human Perceptions of Streetscapes to Better Inform Urban Renewal: A Perspective of Scene Semantic Parsing", *Cities*, 110: 103086.
31. Maas, P.R., 1984. "Towards a Theory of Urban Vitality", Master's Thesis, University of British Columbia.
32. Montgomery, J., 1998. "Making a City: Urbanity, Vitality and Urban Design", *Journal of Urban Design*, 3(1): 93-116.
33. Scepanovic, S., Joglekar, S., Law, S., and Quercia, D., 2021. "Jane Jacobs in the Sky: Predicting Urban Vitality with Open Satellite Data", *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1): 1-25.
34. Schmidhuber, J., 2015. "Deep Learning in Neural Networks: An OverView", *Neural Networks*, 61: 85-117.
35. Sulis, P., Manley, E., Zhong, C., and Batty, M., 2018. "Using Mobility Data as Proxy for Measuring Urban Vitality", *Journal of Spatial Information Science*, 16(2018): 137-162.
36. Sung, H.G. and Lee, S., 2015. "Residential Built Environment and Walking Activity: Empirical Evidence of Jane Jacobs' Urban Vitality", *Transportation Research Part D: Transport and Environment*, 41: 318-329.
37. Sung, H.G., Lee, S., and Cheon, S.H., 2015. "Operationalizing Jane Jacobs's Urban Design Theory: Empirical Verification From the Great City of Seoul, Korea", *Journal of Planning Education and Research*, 35(2): 117-130.
38. Tang, L.J., Lin, Y., Li, S.J., Li, S., Li, J.Y., Ren, F., and Wu, C., 2018. "Exploring the Influence of Urban Form on Urban Vibrancy in Shenzhen Based on Mobile Phone Data", *Sustainability*, 10(12): 1-21.
39. Tao, T., Wu, X., Cao, J., Fan, Y., Das, K., and Ramaswami, A., 2020. "Exploring the Non-linear Relationship Between the Built Environment and Active Travel in the Twin Cities", *Journal of Planning Education and Research*, 43(3): 637-652.
40. Wang, M. and Vermeulen, F., 2020. "Life Between Buildings From a Street View Image: What do Big Data Analytics Reveal About Neighbourhood Organisational Vitality?", *Urban Studies*, 58(15): 3118-3139.
41. Wang, X., Wen, J., Zhang, Y., and Wang, Y., 2014. "Real Estate

Price Forecasting Based on SVM Optimized By PSO", *Optik*, 125(3): 1439-1443.

42. Wu, J.Y., Ta, N., Song, Y., Lin, J., and Chai, Y.W., 2018a. "Urban Form Breeds Neighborhood Vibrancy: A Case Study Using a GPS-Based Activity Survey in Suburban Beijing", *Cities*, 74: 100-108.

43. Wu, C., Ye, X., Ren, F., and Du, Q., 2018b. "Check-in Behaviour and Spatio-temporal Vibrancy: An Exploratory Analysis in Shenzhen, China", *Cities*, 77: 104-116.

44. Wu, X., Tao, T., Cao, J., Fan, Y., and Ramaswami, A., 2019. "Examining Threshold Effects of Built Environment Elements on Travel-Related Carbon-Dioxide Emissions", *Transportation Research Part D: Transport and Environment*, 75: 1-12.

45. Xiao, L., Lo, S., Liu, J., Zhou, J., and Li, Q., 2021. "Nonlinear and Synergistic Effects of TOD on Urban Vibrancy: Applying Local Explanations for Gradient Boosting Decision Tree", *Sustainable Cities and Society*, 72: 103063.

46. Yan, X., Liu, X., and Zhao, X., 2020. "Using Machine Learning for Direct Demand Modeling of Ridesourcing Services in Chicago", *Journal of Transport Geography*, 83: 102661.

47. Yang, J., Cao, J., and Zhou, Y., 2021. "Elaborating Non-linear Associations and Synergies of Subway Access and Land Uses with Urban Vitality in Shenzhen", *Transportation Research Part A: Policy and Practice*, 144: 74-88.

48. Ye, Y., Li, D., and Liu, X., 2018. "How Block Density and Typology Affect Urban Vitality: An Exploratory Analysis in Shenzhen, China", *Urban Geography*, 39(4): 631-652.

49. Yue, W.Z., Yang, C., Zhang, Q., and Liu, Y., 2019. "Spatial Explicit Assessment of Urban Vitality Using Multi-source Data: A Case of Shanghai, China", *Sustainability*, 11(3): 1-20.

50. Zhou, Y., He, X., and Zikirya, B., 2023. "Boba Shop, Coffee Shop, and Urban Vitality and Development—A Spatial Association and Temporal Analysis of Major Cities in China from the Standpoint of Nighttime Light", *Remote Sensing*, 15(4): 904.

Date Received	2024-09-08
Reviewed(1st)	2024-11-08
Date Revised	2025-03-29
Reviewed(2nd)	2025-04-21
Date Accepted	2025-04-21
Final Received	2025-07-09