



모바일폰 위치기반 생활이동 빅데이터를 활용한 통행목적별 도시활력 영향요인 분석

: PageRank 알고리즘과 SHAP 기계학습을 활용하여*

Analysis of Determining Factors of Urban Vitality with Mobile Phone Location-Based Origin-Destination Bigdata by Travel Purpose

: Using the PageRank Algorithm and SHAP Machine Learning

박준상** · 김선재*** · 이수기****

Park, Junsang · Kim, Sunjae · Lee, Sugie

Abstract

Urban vitality is an important indicator for evaluating a city's sustainability. Urban vitality increases when a social space emerges where people can interact with each other in a city. Although many studies have tried to measure urban vitality and its determining factors, few studies have measured it using mobile phone location-based, origin-destination (OD) big data. The aim of this study is to analyze the determining factors of urban vitality with mobile phone big data using the PageRank algorithm and interpretable machine learning techniques. The focus is the nonlinear relationships between urban vitality and its determining factors. The main results of the analysis are as follows. First, urban vitality according to mobile phone location-based, OD big data by travel purpose has different determining factors. For instance, while the perception of street scenery had a considerable influence on the urban vitality of non-commuting travel, it had no impact on the urban vitality of commuting travel. Second, restaurant Point of Interest (POI) density and subway station exit density had positive associations with urban vitality for both leisure and utility travel purposes. Third, street safety was a significant variable for urban vitality, regardless of travel purposes of the population. This finding indicates that the safety of the street environment encourages urban vitality. Finally, the interpretable machine learning analysis indicated that the relationships between urban vitality and its determining factors were nonlinear. Overall, the study findings demonstrate the useful application of mobile phone, location-based, OD big data to examine urban vitality and provide specific policy implications for promoting it.

주제어 도시활력, 생활인구 이동, PageRank, 다층 퍼셉트론, 해석가능한 기계학습

Keywords Urban Vitality, Mobility of Living Population, PageRank, Multi-layer Perceptron, Interpretable Machine Learning

* 이 논문은 한양대학교 교내연구지원사업(HY-2022)으로 연구되었으며, 2022년도 4월 대한국토·도시계획학회 춘계산학협력대회에서 발표하여 우수논문상 수상 논문을 수정·보완한 것임.

** Master's Student, Department of Urban Planning & Engineering, Hanyang University (First Author: june1264@hanyang.ac.kr)

*** Master's Degree, Department of Urban Planning & Engineering, Hanyang University (Co-author: sunjaekim@hanyang.ac.kr)

**** Professor, Department of Urban Planning & Engineering, Hanyang University (Corresponding Author: sugielee@hanyang.ac.kr)

I. 서론

1. 연구의 배경 및 목적

도시의 활력은 도시에 있어서 중요한 요소 중 하나일 뿐만 아니라 도시의 지속가능성을 평가하는 지표 중 하나이다(He et al., 2018; 조월·이수기, 2021). 높은 수준의 활력을 가진 도시는 사회적 상호작용, 근린 생활, 활동 및 장소에 대한 애착에 반영되는 강한 공동체 의식을 촉진하는 경향이 있기 때문이다(Krier, 2009; Jia et al., 2019). 대표적인 도시 이론가인 Jacobs(1961)는 도시의 활력은 거리에서의 보행자 활동으로 구성되며, 도시의 활력은 이웃을 인식할 수 있는 거리, 광장, 공원 등에서의 활동과 높은 관련이 있다고 주장하였다. 따라서 도시활력을 증대시키기 위해서는 사람들이 자주 마주칠 수 있도록 작은 블록, 높은 밀도, 적절하게 높은 토지이용 혼합, 오래된 건물 등이 필요하다고 주장하였다(Jacobs, 1961). 이와 더불어 Ying and Yin(2016)은 도시활력의 사회적 차원에 집중하여 가로의 물리적 환경이 사람들이 활동하는 사회적 공간이 될 때 도시의 활력이 높아진다고 주장하였다.

이후, 도시의 활력을 정량화하기 위해 다양한 방법의 연구가 진행되었다(Azmi and Karim, 2012; Filion and Hammond, 2003; Lu et al., 2019; Yue et al., 2019). 대부분의 선행연구에서는 도시활력을 정량화하기 위해 현장 조사 또는 설문조사를 활용하였으며, 사람들의 주관적인 응답을 통해 분석을 진행하였다. 한편, 유동 인구 또는 지역 상권의 매출 데이터를 활용하여 도시 활력을 분석한 연구도 있다(조월·이수기, 2021; Ye et al., 2018). 다만, 유동인구나 소매업 매출 자료는 비공개되어 있는 자료가 많고 도시활력을 반영하기에 한계가 존재한다. 하지만 최근 들어 통신 기술의 발전과 휴대폰과 같은 GPS 내장 전자기기의 발전으로 개개인의 이동 궤적 데이터와 같은 방대한 자료를 수집하고 분석할 수 있게 되었다. 이에 따라 서울시에서는 서울시 보유 공공 빅데이터와 KT의 휴대폰 LTE+5G 시그널 데이터를 활용하여 서울 생활이동 빅데이터를 제공하고 있다. 서울 생활이동 데이터는 도시민의 출발지, 도착지 및 시간대, 이동 인원, 이동자의 성별, 연령, 이동 소요 시간 및 거리 정보를 담고 있어 도시민의 활동성을 반영하여 도시의 활력을 정량화하는 것이 가능하다. 따라서 본 연구에서는 서울시 생활이동 데이터를 활용하여 도시활력을 정량화하고자 한다.

한편, 인간 활동의 물리적, 사회적 공간으로서 도시 경관은 이론적으로 도시의 활력과 밀접한 관련이 있다(Zhang et al., 2021). 환경이나 경관에 대한 인식은 사용자의 행동에 영향을 미치며 사용자의 행동에 대한 동기, 선호도 및 태도를 반영할 수 있기 때문이다(Nasar, 2008; Yuen and Hien, 2005). 따라서 도시 경관에 대한 사람들의 인식을 반영하는 것은 도시활력을 측정하는 데 있어 중요하다고 볼 수 있다. 하지만 가로경관에 대한 사람

들의 인식을 반영하여 도시활력을 측정하는 연구는 부족한 실정이다. 최근 들어, 빅데이터와 머신러닝, 딥러닝 등 인공지능 기술의 발전으로 가로경관에 대한 사람들의 인식을 정량화하는 것이 가능해졌다(Ordóñez and Berg, 2014; Naik et al., 2017; Yao et al., 2019). 이를 통해 도시활력 영향요인을 분석하는 데 있어 선행연구에서 반영하였던 물리적 환경뿐 아니라 도시의 경관에 대한 사람들의 주관적 인식을 반영하는 것도 가능하다.

이러한 배경에서 본 연구는 서울시 생활이동 데이터와 PageRank 알고리즘을 활용하여 도시활력이라는 개념에 대해 통행의 목적에 따라 통근 통행과 비통근 통행으로 구분하여 조작적으로 정의하였다. 이는 통근 통행과 비통근 통행의 성격이 서로 다르기 때문이다(민병학 외, 2016). 통행의 목적별로 통근 통행은 출퇴근, 등하교 등과 같이 뚜렷한 목적을 가지고 있는 통행이며, 비(非)통근 통행은 통근 통행을 제외한 대부분의 통행을 포함한다. 이러한 비통근 통행은 일상생활 활동이 일어나는 근린 지역과 밀접한 연관을 가지며, 근린 환경 변수에 영향을 많이 받는다(박영준·박소현, 2019; Krizek, 2003; 이남휘·최창규, 2020). 또한, 도시활력에 영향을 미치는 요인을 분석하기 위해 주요 토지이용을 대변할 수 있는 POI 관심시설, 건축물, 교통환경, 가로경관 등을 조작화하여 분석에 활용하였다.

분석 방법론으로는 각 변수가 도시활력과 가지는 선형 또는 비선형 관계를 도출하기 위해 해석가능한 기계학습을 활용하였다. 나아가 분석 결과를 통해 통행목적별 도시활력을 증진시키기 위한 정책적 시사점을 도출하였다.

II. 선행연구 고찰

1. 도시활력의 정의 및 측정 방법

도시활력의 개념은 '서로 다른 시간과 장소에서 특정 도시공간이 얼마나 바쁜가'라는 관점에서 도시 연구자들에 의해 정의되었다(Ravenscroft, 2000; 김영룡, 2020). 이후 도시활력의 개념은 사회적 차원과 경제적 차원으로 구분되어 정의되었다(Bromley and Thomas, 2002). 사회적 활력이란 도시의 여러 공간에서 발생하는 유동 인구로 인한 북적거림(buzz)을 의미하며, 경제적 활력은 소비, 거래 투자 등의 활동을 의미한다(김영룡, 2020). 본 연구에서는 도시민의 활동성을 반영할 수 있는 서울시 생활이동 데이터를 활용하여 도시활력의 두 가지 차원의 개념 중 사회적 차원의 도시활력에 대해 조작적으로 정의하고자 한다.

최근 들어 다양한 유형의 도시 빅데이터가 생성되고 축적됨에 따라 이를 이용하여 도시활력을 객관적으로 측정하기 위한 다양한 연구들이 진행되었다. 구체적으로 Yue et al.(2019)는 음식점 시설의 리뷰 개수를 활용하여 도시활력을 측정하였으며, 이 밖에도 야간 조명, 와이파이 핫스팟 및 신용카드 매출 자료, 핸드폰 위

치 데이터 등을 활용하여 도시활력을 측정하는 연구들이 진행되었다(Jin et al., 2017; Kim, 2018; Li et al., 2020). 이것으로 미루어 볼 때, 도시활력은 다양한 종류의 빅데이터를 활용하여 정의될 수 있는 차원을 가지고 있다는 것으로 이해할 수 있다(정시운·전병운, 2020; 조월·이수기, 2021).

특히, 최근 스마트폰의 광범위한 사용과 GPS 기반 위치정보 축적 기술의 발전으로 이를 활용한 도시활력 측정 연구가 증가하고 있다(Ying and Yin, 2016; 조월·이수기, 2021). 대부분의 선행연구에서 도시활력을 측정하기 위해 사용한 위치 데이터 및 생활인구 데이터는 특정 시점에 특정 위치에 존재하는 데이터를 활용하였다. 따라서 실제 외부 활동이 일어나지 않은 데이터도 도시활력을 높이는 데에 포함되어 연구가 진행되었다. 하지만, 앞서 언급했다시피 도시의 활력은 도시의 여러 공간에서 발생하는 유동 인구에 의해 정의되었다. 따라서 실제 이동이 발생한 데이터만을 활용하여 도시활력을 측정하는 것이 필요하다고 볼 수 있다.

2. 도시활력 영향요인

Jane Jacobs(1961)는 도시의 활력에 있어 주요한 네 가지 필수 조건을 언급하였다. 구체적으로, 도시의 다양성과 활력을 유지하기 위한 토지이용 혼합, 작은 크기의 도시 블록, 오래된 건물, 높은 밀도를 도시의 활력에 영향을 미치는 주요 요인으로 제시하였다. 이러한 요인은 도시공간에서 도시민이 모일 수 있도록 만드는 요소로 볼 수 있고, 따라서 도시활력은 도시민의 활동을 통해서 모여서 만들어지는 것으로 이해할 수 있다(조월·이수기, 2021). Jacobs는 이러한 도시의 활력을 「미국 대도시의 죽음과 삶」에서 사회, 경제활동을 충족시키는 도시에서의 중요한 요소라고 주장하였다. 이러한 물리적 환경 요인은 특정 공간에 도시민이 많이 모이도록 하는 요소로 볼 수 있다. 도시의 활력은 도시민의 활동으로 인한 상호작용이 일어날 때 생기는 것으로 이해할 수 있기 때문이다(조월·이수기, 2021).

이후, 도시활력에 영향을 미치는 요인을 분석하기 위해 Jacobs(1961)의 이론에 기반한 요인을 포함하여 객관적인 물리적 환경 요인을 중심으로 분석이 진행되었다(Sung and Lee, 2015; Sung et al., 2015; Wu et al., 2018; 임하나 외, 2016; 조월·이수기, 2021). Katz(1994)는 도시의 조밀한 정도, 보행의 규모, 토지이용의 기능적 혼합 및 적절한 수준의 건물 밀도가 도시의 활력에 영향을 미치는 중요한 요소로 주장하였다. 이와 더불어 조월·이수기(2021)는 낮 시간대 생활인구 수에 영향을 미치는 요인을 분석하였으며, 음식점, 교육·학문 시설, 금융·보험시설 등이 도시 활력에 긍정적인 영향을 미치고 있음을 주장하였다. 또한, 지역별로 도시활력에 영향을 미치는 POI 요인이 다른 것으로 나타났다. 이러한 선행연구를 바탕으로 도시의 물리적 환경은 도시활력과 중요한 연관이 있음을 알 수 있다.

다음으로, 이러한 물리적 환경을 통해 도시민의 활동을 향상시키기 위해서는 활동에 영향을 미치는 도시 물리적 환경의 속성을 측정해야 하며, 이는 인식의 문제가 포함된다(Nasar, 2008). 물리적 환경에 대한 직접적인 측정은 사람들의 인식으로 직접적으로 전환되지 않으며, 일부 인식은 신체 활동과 독립적인 연관성을 가질 수 있기 때문이다(Michael et al., 2006; Nasar, 2008). 이에 따라 도시의 물리적 환경에 대한 인식과 신체 활동 간의 연관성을 분석한 연구가 다수 진행되었다. 이와 관련하여 Miles and Panton(2006)은 도시환경에 대한 인식과 보행 활동에 대해 연구를 진행하였으며, 사람들은 안전하거나 쾌적하다고 인식하는 지역에서 더 많이 걷는다는 분석 결과가 도출되었다. 즉, 가로경관 인식에 대한 이해는 도시민의 활동을 장려하는 장소에 대한 더 나은 계획을 가능하게 할 수 있음을 알 수 있다.

3. 해석가능한 기계학습

기계학습은 머신러닝(Machine Learning)이라고도 불리며, 컴퓨터에 입력된 변수 간의 관계를 파악하고 입력된 데이터를 바탕으로 패턴을 평가 및 예측하는 알고리즘 기법을 의미한다(권철민, 2019; Bzdok et al., 2017). 따라서 다양한 상황에 대해 유연한 적용이 가능하여 복잡한 데이터의 분석에 적합한 다양한 알고리즘을 제공한다. 또한, 기존의 전통적인 회귀분석은 분석 변수에 대한 선형가정, 결측값 등 데이터에 대한 다양한 제약사항이 존재하였다. 하지만 기계학습의 알고리즘은 이러한 변수 간의 복잡한 상호작용 문제와 관련하여 뛰어난 유연성을 제공한다(Caraviello et al., 2006). 추가적으로 독립변수와 종속변수 간의 선형 또는 비선형 관계를 분석하고자 하는 여러 분야에서 유용하게 활용될 수 있다(Gan et al., 2020; Yan et al., 2020).

그러나 훈련된 기계학습 알고리즘 모형의 내부는 블랙박스로 간주되어 각 독립변수가 종속변수의 예측에 기여한 정도를 알 수 없다는 한계점이 있었다. 하지만 해석가능한 기계학습(Interpretable Machine Learning)의 발전으로 블랙박스를 해석할 수 있게 되면서 도시 분야를 포함한 다양한 분야에서 활용되기 시작하였다. 블랙박스를 해석하는 대표적인 방법론에는 Skater, Local Interpretable Model-Agnostic Explanation (LIME), SHapley Additive exPlanations (SHAP), Partial Dependence Plot(PDP) 등이 있다. 그중 SHAP은 1953년 게임 이론을 바탕으로 Shapley(1953)에 의해 처음 제안되었으며, 모델의 예측력에 있어 설명변수의 기여도에 대해 통찰력 있는 측정값을 제공한다. 이후, Lundberg and Lee(2017)에 의해 다양한 기계학습 모델에 대한 SHAP를 계산할 수 있는 Python 패키지가 개발되면서 블랙박스를 해석하는 데 있어 대중적으로 사용되고 있다(Stojic et al., 2019).

앞선 선행연구를 정리하면 기계학습 방법론은 전통 선형 회귀

분석 방법론에 비해 데이터에 대한 제약이 없어 유연하고 자유로움을 알 수 있다. 또한, 독립변수와 종속변수가 가지는 선형 또는 비선형 관계를 확인할 수 있고, 전통 회귀모형에 비해 예측력이 높다는 장점이 존재한다. 따라서 본 연구의 방법론으로 기계학습 방법론을 활용하고, 모형의 블랙박스를 해석하여 도시의 활력에 영향을 미치는 요인을 분석하고자 한다.

4. 연구의 차별성

선행연구 고찰 결과에 따른 한계점은 다음과 같다. 첫째, 기존의 도시활력을 측정하는 방법은 현장 조사, 설문조사 등을 활용하거나, 핸드폰 위치 데이터, 생활인구 빅데이터 등을 활용하였다. 하지만 이러한 데이터는 시간과 비용이 많이 들고 대상지가 한정되어 있거나, 도시민의 활동성을 반영하기에는 한계가 존재하며 도시민의 실제 이동 데이터를 활용한 도시활력 측정 연구는 부족한 실정이다.

둘째, 대다수의 선행연구에서 사용한 토지이용 변수의 경우, 해당 지역의 건축물 데이터를 기반으로 활용하고 있다. 건축물 데이터의 경우 연면적을 통해 해당 토지이용의 수직적인 면적도 반영할 수 있다는 장점이 있다. 하지만, 주 용도를 기준으로 구축되어 있는 건축물 데이터의 용도는 복합용도 건축물의 경우 부정확한 자료로 볼 수 있다(조월·이수기, 2021). 따라서 이러한 한계점을 보완할 수 있는 다양한 종류의 건물 용도를 정밀하게 반영하는 것이 필요하다.

셋째, 도시의 경관에 대한 인식은 도시민의 행동에 대한 동기, 선호도 등을 결정하여 긍정적인 가로경관은 도시활력을 증진시키는 데에 도움이 될 수 있다. 하지만, 데이터 구축과 기술의 한계로 인해 가로경관에 대한 사람들의 인식을 정량화할 수 없다는 한계점이 존재했다. 따라서 가로경관에 대한 인식이 도시활력에 미치는 영향에 대해 분석한 연구는 부족한 실정이다. 마지막으로 기계학습 모형은 전통 회귀분석에 비해 입력 데이터에 대한 제약이 자유로우며, 높은 예측력을 가지고 있고, 변수 간의 선형 또는 비선형 관계를 분석할 수 있다. 이러한 장점에도 불구하고 이를 활용하여 도시활력과 영향요인 간의 관계를 분석한 연구는 부족한 실정이다.

이에 따른 본 연구의 차별성은 다음과 같다. 첫째, 기존 연구에서 도시활력 지표로 많이 사용되었던 집계된 인구 자료의 한계를 극복하고 도시민의 실제 활동을 반영하기 위하여 서울시 생활이동 데이터와 네트워크 중심성을 분석할 수 있는 PageRank 알고리즘을 활용하여 도시의 활력을 정량화하였다. 둘째, 기존 건축물의 주 용도를 기반으로 구축된 토지이용 변수의 한계점을 극복하기 위해 건축물의 용도를 다양하게 분류하여 용도별 밀도를 도출할 수 있는 POI 빅데이터를 활용하였다. 셋째, 도시활력에 있어 가로경관에 대한 도시민의 인식을 반영하기 위하여 딥러닝 알

고리즘을 활용하여 서울시의 가로경관 이미지에 대한 도시민의 인식을 객관적으로 정량화하였다. 마지막으로 해석가능한 기계학습을 통해 Jacobs(1961)의 이론을 바탕으로 구성된 도시 물리적 환경과 가로경관 인식 변수가 도시활력에 미치는 선형 또는 비선형 관계를 분석하여 정책적 시사점을 도출하였다.

III. 분석자료 및 분석방법

1. 분석자료

본 연구는 도시 물리적 환경과 가로경관에 대한 도시민의 인식이 도시활력과 가지는 관계를 분석하는 데에 초점을 두고 있다. 따라서 도시의 활력을 정량화하기 위해 서울특별시에서 제공하는 서울 생활이동 데이터를 활용하였으며, 따라서 연구의 공간적 범위는 서울특별시이고 분석 단위는 행정동이다(그림 1). 도시민이 통행을 하는데 있어 특히 비통근 통행 시에는 북한산, 관악산 등 주요 산지와 한강 및 지류들이 통행의 목적지로 활용될 수 있기 때문에 행정동 면적 계산 시 행정동의 전체 면적을 활용하였다. 연구의 시간적 범위는 COVID-19로 인한 도시민의 이동 패턴 변화를 제어하기 위하여 2021년 1년 동안의 데이터를 사용하였다. 2020년부터 2년 가까이 지속된 COVID-19로 인해 익숙해진 비대면 생활로 통행의 패턴이 어느 정도 정착이 되어 있기 때문에 포스트 코로나 시대에는 COVID-19가 발생하기 이전의 모습 그대로 완전히 되돌아가지는 않을 확률이 높기 때문이다(정혜선, 2021).

다음 <표 1>은 본 연구에서 사용한 종속변수와 독립변수에 대한 정의 및 출처이다. 종속변수인 서울 생활이동 데이터는 KT 기지국에 집계되는 LTE와 5G 신호를 바탕으로 출발지와 목적지별 이동 인원수를 가공하여 개방한 데이터이다. 도시민의 체류 패턴 분석을 통해 새벽 5시 정각을 기준으로 야간 상주지와 주간 상주지가 추정되었다. 구체적으로, 해당 월 기준 특정 지역에서 새벽

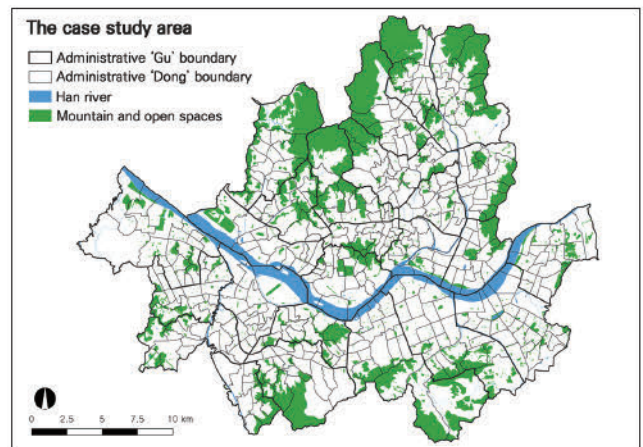


그림 1. 연구의 공간적 범위 및 분석 단위
Figure 1. The case study area and unit of analysis

표 1. 변수의 정의 Table 1. Description of variables

Variables	Description of variables		Sources		
종속 변수 Dependent variable	통근 통행 목적 도시활력 CCommuting driven urban vitality	Urban vitality index for commuting traffic using PageRank algorithm	Seoul open data (2021)		
	비통근 통행 목적 도시활력 Non-commuting driven urban vitality	Urban vitality index for non-commuting traffic using PageRank algorithm			
POI 특성 POI characteristics	주거시설 POI 밀도(개/km ²) POI density of residential facilities (no./km ²)	Number of POIs in residential facilities by administrative dong/ Area of administrative dong	Kakao Map (2020)		
	근린상업시설 POI 밀도(개/km ²) POI density of neighborhood commercial facilities (no./km ²)	Number of POIs for neighborhood commercial facilities by administrative dong/ Area of administrative dong			
	제조업시설 POI 밀도(개/km ²) POI density of manufacturing facilities (no./km ²)	Number of POIs for manufacturing facilities by administrative dong/ Area of administrative dong			
	병원 POI 밀도(개/km ²) POI density of hospitals (no./km ²)	Number of hospital POIs by administrative dong/ Area of administrative dong			
	스포츠·레저시설 POI 밀도(개/km ²) POI density of sports and leisure facilities (no./km ²)	Number of POIs for sports and leisure facilities by administrative dong/ Area of administrative dong			
	음식점 POI 밀도(개/km ²) POI density of restaurants (no./km ²)	Number of POIs for restaurants by administrative dong/ Area of administrative dong			
	POI 혼합도 POI facility mix	Entropy index of POI facilities by administrative dong			
	독립 변수 Independent variable	건축물 특성 Building characteristics		건물 밀도(km ² /km ²) Building density (km ² /km ²)	New address DB (2021)
		평균 건물 높이(m) Average building height (m)		Average building height by administrative dong	
		평균 건물 나이 Average building age		Average building age by administrative dong	
건물 나이 다양성 Building age diversity		Standard deviation of average building age by administrative dong			
교통 환경 특성 Traffic environment characteristics	3지 교차로 밀도(개/km ²) Three-way intersection density (no./km ²)	Number of three-way intersections by administrative dong/ Area of administrative dong	Open Street Map (2021)		
	4지 교차로 밀도(개/km ²) Four-way intersection density (no./km ²)	Number of four-way intersections by administrative dong/ Area of administrative dong			
	평균 도로 폭(m) Average road width (m)	Average road width by administrative dong	National transportation DB (2021)		
	평균 도로 길이(km) Average road length (km)	Average road length by administrative dong			
	버스정류장 밀도(개/km ²) Bus stop density (no./km ²)	Number of bus stops by administrative dong/ Area of administrative dong	Seoul open data (2021)		
지하철 출구 밀도(개/km ²) Subway exit density (no./km ²)	Number of subway exits by administrative dong/ Area of administrative dong	New address DB (2021)			
가로경관 인식 특성 Street view perception characteristics	활발함 Lively	가로경관의 활발함 인식 지수 Street view 'lively' perception index	Naver Street View (2020)		
	안전함 Safety	가로경관의 안전함 인식 지수 Street view 'safety' perception index			

5시를 포함하여 체류한 누적 시간이 많은 곳을 거주지(H), 당일 기준 특정 지역에서 새벽 5시 이후에 체류를 시작하여 이튿날 새벽 5시 이전에 체류를 끝낸 누적 시간이 가장 많은 곳을 근무지(W), 그 이외의 지역을 기타 장소(E)로 구분하여 이동의 성격이 추정된 자료이다. 따라서 특정 지역 간에 반복적으로 이동하는 생활이동 인구와 비정기적 이동을 구분하는 것이 가능하다. 이를 활용하여 통행의 목적별로 도시활력에 미치는 영향을 구분하기 위해 통근 통행과 비통근 통행으로 구분하여 분석하였다. 본 연구에서 활용된 통근 통행의 데이터로는 주중의 거주지에서 근무지로의 통행(H-W), 근무지에서 거주지로의 통행(W-H) 데이터가 포함되었으며, 비통근 통행 데이터로는 주말의 통근 통행을 제외한 데이터와 업무 관련 통행인 근무지에서 근무지로의 통행(W-W)을 제외한 데이터를 활용하였다.

본 연구에서 사용된 독립변수로는 크게 POI 특성, 건축물 특성, 교통환경 특성, 가로경관 인식 특성으로 구성하였다. 먼저, POI 특성 변수는 각각의 시설이 위치기반 점 형태로 구성되어 있는 공간 데이터이며, 건축물 데이터 기반의 토지이용 변수보다 용도가 더 세분화되어 있다는 장점이 있다(Yue et al., 2017). 이러한 유형별 POI 시설은 생활이동의 분포에 영향을 미치며, 도시에서의 활동과 관련이 있는 중요한 변수이다(조윌·이수기, 2021). 따라서 POI 시설의 특성을 주거시설, 근린 상업시설, 제조업 시설, 병원, 스포츠·레저시설, 음식점, POI 혼합도로 구분하여 구축하였다. 다만, 주거시설 POI는 아파트의 경우 동 1개당 1개의 POI가 생성되며, 단독주택의 경우에도 건물 1개당 1개의 POI가 생성된다. 주거시설 이외의 시설은 같은 건물 안에 있더라도 해당 시설점포 1개당 1개의 POI가 생성된다. POI 혼합도는 유형별 POI 시설 개수에 대한 엔트로피 지수를 활용하여 구축하였다. 구체적으로, POI 혼합도는 0부터 1 사이의 값을 가지며, 1에 가까울수록 해당 행정동의 POI 혼합유형이 고르게 이루어지고 있음을 의미한다.

건축물 특성 변수의 경우에는 Jacobs가 언급한 도시활력에 영향을 미치는 주요 물리적 환경 요인을 반영하고자 하였다. 앞서 언급하였듯이 Jacobs는 높은 건물 밀도, 작은 블록, 혼합적 토지 이용, 다양한 연령의 건물을 주요 요인으로 주장하였다. 따라서 높은 건물 밀도, 평균 건물 높이, 평균 건물 나이, 건물 나이 다양성으로 구분하여 구축하였다. 교통환경 특성의 경우, 도로망의 구조를 막다른 골목과 격자형 도로 구조를 설명하여 블록의 크기를 대변할 수 있는 3지 교차로, 4지 교차로의 밀도를 고려하였다. 이외에도 평균 도로 폭, 평균 도로 길이, 대중교통으로의 접근성을 나타내는 버스 정류장 밀도, 지하철 출구 밀도를 구축하였다.

마지막으로, 가로경관 인식 특성 변수의 경우에는 '활발함(Lively)', '안전함(Safety)' 인식을 활용하여 구축하였다. 이를 위해 딥러닝 모델과 머신러닝 모델을 활용하여 서울시의 가로경관

에 대한 인식을 정량화하였다. 구체적으로, 딥러닝 모델은 MIT에서 제공한 Place Pulse 2.0 데이터를 기반으로 딥러닝 의미론적 분할(Semantic Segmentation)을 진행하여 가로경관 이미지 구성요소를 비율로서 정량화하였다. 이후, 정량화된 가로 이미지 데이터를 머신러닝 모델을 활용하여 학습시킨 후, 서울시 가로경관 빅데이터 셋인 Naver Street View(NSV) 파노라마 이미지를 활용하여 가로경관 인식 지표를 도출하였다.

Place Pulse 2.0 데이터는 전 세계 28개 국가의 56개 도시에서 총 110,988개의 Street View 이미지를 총 162개 국가의 81,630명이 평가한 데이터로 이루어져 있다. 평가에 활용된 여섯 가지 지표는 Beautiful, Boring, Wealthy, Lively, Safety, Depressing이며 각 질문별로 두 장 이미지의 우열을 비교한 데이터이다. 이러한 승패로 구성된 데이터를 정량화하기 위해 Microsoft Trueskill 알고리즘을 활용하여 이미지 간 서열을 점수화하였다(Herbrich et al., 2007; Zhang et al., 2018).

최종적으로 가로경관에 대한 인식 지표를 예측하는 데 사용된 서울시의 Street View 데이터는 NSV 파노라마 이미지를 활용하였다. NSV 파노라마 이미지는 2020년 도로망을 기준으로 10m 간격으로 지점을 생성하고, 각 지점에 해당하는 파노라마 이미지를 구축하여 사용하였다. 다만, 본 연구에서는 변수의 의미를 국문으로 번역하였을 때 직관적으로 이해하고 해석하기 난해하거나, 상관관계가 높은 변수는 제거하였다. 따라서 최종적으로 '활발함(Lively)', '안전함(Safety)'에 대한 인식을 정량화하여 가로경관 인식 변수로 선정하였다. 가로경관에 대한 인식을 정량화한 분석 프레임워크는 다음 <그림 2>와 같다.

2. PageRank 알고리즘

본 연구에서는 도시활력을 정량화하기 위해 서울 생활이동 데이터와 PageRank 알고리즘을 활용하였다. PageRank 알고리즘은 Google 검색에서 웹사이트 간의 순위를 지정하는 데 사용되는 알고리즘으로, 페이지에 대한 링크의 수와 품질을 계산하여 웹사이트가 얼마나 중요한지 추정하는 데 사용된다(Page et al., 1999). PageRank 알고리즘은 위세 중심성의 한계인 특정 노드와 연결된 모든 링크의 값이 동일하게 계산되는 한계를 극복하고 웹 페이지 간의 링크 관계를 통해 구글 검색 알고리즘을 만들려는 목적에서 개발되었다. PageRank 알고리즘을 활용하여 도시의 활력을 계산할 수 있으며, 전통적인 인원 집계 방식보다 인간의 활동을 더 밀접하게 나타낼 수 있다(Jia et al., 2019).

PageRank 알고리즘 네트워크에서 가장 중요한 링크의 속성은 인바운드(In-bound)와 아웃바운드(Out-bound)의 수이다. PageRank 알고리즘에서는 모든 노드의 인바운드 링크를 동일하게 계산하는 것이 아니며 아웃바운드된 노드의 중요도와 수에 의해 결정된다. PageRank에 사용된 공식은 다음과 같으며, $PR(T_i)$

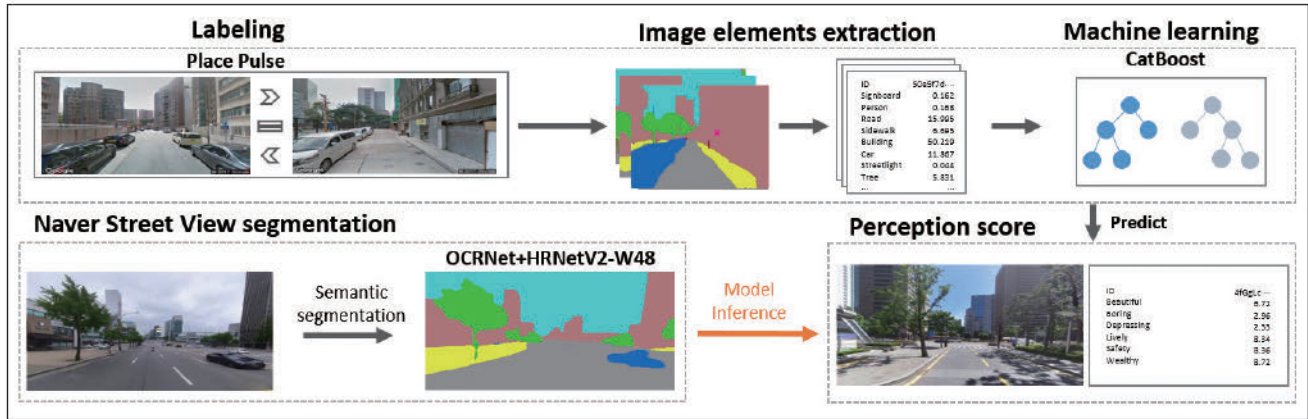


그림 2. 머신러닝을 활용한 가로경관 인식 정량화 프레임워크
 Figure 2. A framework for quantifying street view perception using machine learning

는 A와 연결된 페이지의 PageRank이고 $C(T_i)$ 는 T_i 가 가지고 있는 링크의 총 개수를 의미한다. d 는 0과 1의 범위에서 설정하는 댄핑 계수이며 다른 노드로 이동할 확률로 정의된다. 즉, 도시민이 A 행정동에서 B 행정동으로 이동한 후, B 행정동에 계속 머물지 않고 또 다른 행정동으로 이동할 확률이라고 볼 수 있다. 다양한 연구에서 댄핑 계수에 대한 다양한 값을 테스트했지만, 일반적으로 댄핑 계수는 0.85로 설정해야 한다고 가정하였다(Brin and Page, 1998; Jia et al., 2019). 이러한 선행연구를 참고하여 본 연구에서 설정한 댄핑 계수의 값은 0.85이다.

$$PR(A) = (1-d) + d \times \sum_i \frac{PR(T_i)}{C(T_i)} \quad (1)$$

본 연구에서는 생활이동 데이터를 사용하였으므로 각각의 노드는 행정동이며 인바운드 링크는 통행 유입, 아웃바운드 링크는 통행유출로 간주하였다. 이후, 각각의 이동에 대한 인구수는 가중치로 설정하여 계산하였다.

3. 해석가능한 기계학습

본 연구의 목적은 도시활력에 영향을 미치는 요인들 간의 관계를 분석하는 것이다. 이에 따라 독립변수를 토지이용 특성, 건축물 특성, 교통환경 특성, 가로경관 인식 특성으로 구축하였다. 이를 분석하기 위한 가장 보편적인 방법은 최소제곱법(OLS) 회귀 모형이 있다. 하지만 이러한 전통적인 회귀분석에 비해 모형의 예측력이 높고 결측값, 선형가정 등 데이터에 대한 제약이 적어 변수 간의 선형 또는 비선형 관계를 분석할 수 있는 기계학습(Machine Learning) 방법론이 유용하다. 따라서 본 연구에서는 데이터에 대해 뛰어난 유연성과 높은 예측력을 제공하는 기계학습 방법론을 사용하였다. 이후 최종적으로 해석가능한 기계학습을 활용하여 모형의 블랙박스를 해석하였다.

기계학습은 인공지능의 한 분야로, 통계적인 수학 기법을 활용하여 컴퓨터가 학습한 뒤 이를 통해 향상된 성능의 알고리즘을 구현하는 것이다(Tom, 1997). 본 연구에서는 기계학습 방법론 중 Scikit-learn에서 제공하는 Multi-layer Perceptron(MLP) Regressor 모형을 사용하였다. MLP Regressor는 입력, 출력 및 은닉 계층으로 구성된 여러 뉴런에 의해 학습되는 알고리즘이다. 해당 모형의 예측력을 평가하기 위한 지표로 다수의 연구에서 사용하고 있는 결정계수(R^2), 평균절대오차(Mean Absolute Error), 평균제곱오차(Mean Squared Error)를 활용하였다.

이후, 모형의 블랙박스를 해석하기 위해 해석가능한 기계학습 방법론인 SHapley Additive exPlanations(SHAP)를 활용하였다. SHAP는 모든 기계학습 모형의 기여도를 설명하기 위해 게임이론을 바탕으로 만들어진 방법이다(Lundberg and Lee, 2017). Shapley Value(기여도)는 전역적 지수(Global Shapley Value)와 국지적 지수(Local Shapley Value)의 두 가지로 나누어진다. Global Shapley Value는 영향력의 크기가 큰 순서대로 위쪽부터 나열되며, 전체적인 변수의 영향력이 양(+)의 관계일 경우에는 붉은색, 음(-)의 관계일 경우에는 푸른색으로 나타난다.

Local Shapley Value는 밀도 산점도를 통해 변수의 기여도를 자세하게 나타낸다. Local Shapley Value의 점 하나당 데이터 하나를 의미하며, 본 연구의 분석단위는 424개의 서울시 행정동이므로 각 변수당 424개의 점이 나타난다. 점의 색깔은 해당 변수의 실제 값을 의미하며, 변수의 값이 높을수록 붉은색, 낮을수록 푸른색으로 표시된다. 즉, X축의 0을 기준으로 변수의 값이 증가할수록 오른쪽에 표시된다면 해당 변수는 도시활력과 양(+)의 관계를 가지고 있으며, 왼쪽에 표시된다면 음(-)의 관계를 가지고 있다는 것으로 해석할 수 있다.

IV. 분석 결과

1. 기술 통계 분석

본 연구에서 사용된 변수에 대한 기술 통계는 다음 <표 2>와 같다. 종속변수는 통근 통행 도시활력, 비통근 통행 도시활력의 두 가지이며, 독립변수는 토지이용 특성, 건축물 특성, 교통환경 특성, 가로경관 인식 특성으로 구성되어 있다. 기술 통계량은 각 변수에 대한 평균, 표준편차, 최솟값, 최댓값으로 계산하였다. 먼저, 종속변수인 통근 통행 도시활력과 비통근 통행 도시활력의 경우 평균값은 3.07로 동일하게 나타났다. 이는 PageRank 이론 특성상 각각의 노드들이 가진 점수를 주고받는 방식이기 때문으로 판단된다. 통근 통행의 경우 편차가 비통근 통행 시 도시활력보다 높아 특정 행동동에 활력이 더 집중되어 있음을 알 수 있다.

POI 특성의 경우, POI 밀도는 근린 상업시설, 음식점, 주거시설 등의 순으로 높은 밀도를 가지는 것으로 나타났다. 또한, 행정동별 POI 혼합도의 평균값은 0.70로 나타났다. 밀도를 나타내는 건물 밀도의 경우 평균 1.06으로 나타났다. 오래된 건물을 의미하는 건물 연령 다양성의 경우, 평균 13.93으로 도출되었다. 마지막으로 서울시의 행정동별 건물의 평균 연령은 30년 정도인 것으로 나타났으며, 평균 연령이 가장 적은 행정동은 4.45년, 가장 많은 행정동은 54.18년으로 나타났다. 마지막으로 건축물의 평균 높이는 약 12.65m로 나타났고, 평균 높이가 가장 낮은 행정동은 4.21m, 가장 높은 동은 67.73m로 나타났다.

교통 환경 변수 중 교차로 밀도는 막다른 길과 그렇지 않은 교차로를 구분하기 위하여 3지 교차로의 밀도와 4지 교차로의 밀도를 활용하였으며, 3지 교차로의 평균 밀도는 113.99, 4지 교차로의 평균 밀도는 35.56으로 나타났다. 도시 형태 이론에서 막다른 길의 밀도가 높은 지역은 자연발생적으로 성장한 곳으로 볼 수 있으며, 4지 교차로의 밀도가 높은 지역은 계획적으로 조성되었다고 판단할 수 있다(이수기 외, 2016). 또한, 행정동별 평균 도로 폭은 7.01m, 평균 도로 길이는 1.68km로 나타났다.

추가적으로 대중교통 접근성 변수로서 버스 정류장 밀도와 지하철역 출구의 밀도를 계산하였다. 버스 정류장 밀도의 평균값은 21.38, 지하철 출구의 평균 밀도는 3.90으로 나타났다. 마지막 가로경관 인식 특성의 경우, 서울시의 가로경관에 대해 활발함(Lively), 안전함(Safety) 인식을 정량화하여 구축하였다. 활발한 인식에 대한 평균값은 0.51이며, 안전한 인식에 대한 평균값은 0.43으로 나타났다.

2. PageRank를 통한 도시활력 측정

본 연구에서는 서울시 생활이동 데이터와 PageRank 알고리즘을 활용하여 행정동별 도시활력을 정량화하였다. 추가적으로, 통

행의 유형에 따라 영향을 미치는 요인이 다를 것으로 판단하고 통행의 유형을 통근 통행과 비통근 통행으로 구분하였다. 먼저, 서울 생활이동 데이터 내부의 체류지 구분은 다음과 같이 정의되었다. 해당 월 기준 특정 지역에서 새벽 5시를 포함하여 체류한 누적 시간이 가장 많은 지역을 야간 상주지(거주지), 당일 기준 특정 지역에서 새벽 5시 이후에 체류를 시작하여 이튿날 새벽 5시 이전에 체류를 끝낸 누적 시간이 가장 많은 지역을 주간 상주지(근무지), 이외의 지역을 기타 지역으로 구분하였다.

따라서 통근 통행 유형에 포함된 데이터는 평일 기준 거주지에서 근무지로 이동한 데이터(출근 혹은 등교 이동), 근무지에서 거주지로 이동한 데이터(퇴근 혹은 하교 이동)가 포함되었다. 비통근 통행의 경우에는 주말 기준 거주지에서 기타지역, 기타지역에서 근무지, 기타지역에서 거주지, 기타지역에서 기타지역, 거주지에서 거주지, 근무지에서 기타지역으로 이동한 데이터가 포함되었다. 이외의 통행은 통근 통행, 비통근 통행으로 명확히 정의하기에는 한계가 있다고 판단하여 분석 데이터에서 제외되었다.

앞서 정의한 통근 통행, 비통근 통행 데이터를 활용하여 PageRank 알고리즘을 활용한 행정동별 도시활력 지수를 도출하였다(그림 3). 먼저 통근 통행의 경우, 강남, 여의도, 종로의 서울 3도심에 집중적으로 높은 활력 지수가 나타났다. 이는 강남, 여의도, 종로에 업무시설, 상업시설 등 통근 통행 유형에서 통행 유입을 유발하는 시설들이 다수 존재하기 때문으로 판단된다. 반면, 평창동, 내곡동, 시흥동, 신원동 등 서울 외곽 지역의 경우에는 도시활력이 가장 낮은 수준으로 나타났다. 전체적으로 보았을 때, 통근 통행의 경우에는 도시활력의 수준이 서울 3도심과 그 외 지역의 편차가 큰 것으로 도출되었다.

비통근 통행의 경우에도 마찬가지로 강남, 종로, 여의도의 도시활력 수준이 높은 것으로 나타났다. 하지만, 통근 통행의 경우와는 다르게 특히 강남에서의 도시활력 수준이 줄어들고 서울시의 외곽 지역 등 전체적인 도시의 활력 수준이 높아진 것으로 나타났다. 비통근 통행의 경우에는, 통근 통행과 달리 다양한 종류의 통행 목적이 존재하고, 이에 따라 통행 유입을 유발하는 시설이 통근 통행에 비해 다양해졌기 때문으로 판단된다. 전체적인 결과값을 보았을 때, 생활이동 데이터와 PageRank를 활용한 도시활력의 측정은 합리적인 결과를 도출하였다고 판단할 수 있다.

3. 가로경관 인식 정량화 기계학습 모형

본 연구에서는 먼저, Place Pulse 2.0 데이터셋에서 평가된 가로경관 이미지에 대해 Microsoft Trueskill 알고리즘을 활용하여 인식 지표별 가로 이미지를 점수화하였다. 인식 지표의 경우 개인의 주관적인 판단이므로 개인 특성에 따라 가로경관에 대한 인식이 차이가 있을 수 있다. 이러한 주관적인 평가를 예측할 경우에는 이진 분류 모형이 회귀 모형보다 더 정확하다(Datta et

표 2. 기술 통계 분석표 Table 2. Descriptive statistical analysis

Variables		Obs.	Mean	S.D	Min.	Max.	VIF
종속변수 Dependent variable	통근 통행 도시활력 Commuting driven urban vitality	424	3.07	1.78	0.75	17.57	-
	비통근 통행 도시활력 Non-commuting driven urban vitality	424	3.07	1.62	0.65	14.38	-
POI 특성 POI characteristics	주거시설 POI 밀도 POI density of residential facilities	424	97.80	60.97	0.00	395.51	2.00
	근린상업시설 POI 밀도 POI density of neighborhood commercial facilities	424	315.11	262.85	3.29	2,238.70	4.31
	제조업시설 POI 밀도 POI density of manufacturing facilities	424	16.29	28.96	0.00	359.75	1.66
	병원 POI 밀도 POI density of hospitals	424	48.49	36.05	0.00	240.97	2.87
	스포츠·레저시설 POI 밀도 POI density of sports and leisure facilities	424	24.92	16.68	0.00	100.09	3.70
	음식점 POI 밀도 POI density of restaurants	424	206.33	164.65	2.19	1,124.01	4.72
	POI 혼합도 POI facility mix	424	0.70	0.16	0.00	0.88	1.21
	건축물 특성 Building characteristics	건물 밀도 Building density	424	1.06	0.52	0.00	3.01
건물 평균 높이(m) Average building height		424	12.65	7.35	4.21	67.73	3.54
건물 평균 연령 Average building age		424	30.26	5.69	4.45	54.18	1.75
건물 연령 다양성 Building age diversity		424	13.93	13.59	2.66	176.89	1.08
교통환경 특성 Traffic environment characteristics	3지 교차로 밀도 Three-way intersection density	424	113.99	78.60	5.56	424.21	2.63
	4지 교차로 밀도 Four-way intersection density	424	35.56	28.08	0.00	185.61	3.18
	평균 도로 폭(m) Average road width	424	7.01	3.52	2.71	25.92	3.58
	평균 도로 길이(km) Average road length	424	1.68	2.05	0.10	15.63	2.36
	버스 정류장 밀도 Bus stop density	424	21.38	13.46	0.00	87.66	1.22
	지하철 출구 밀도 Subway exit density	424	3.90	4.72	0.00	32.75	1.74
가로경관 인식 특성 Street view perception characteristics	활발함 Lively	424	0.51	0.06	0.32	0.84	3.41
	안전함 Safety	424	0.43	0.06	0.33	0.79	4.97

al., 2006; Ordonez and Berg, 2014). 본 연구에서는 점수화된 어 순위가 나열된 가로 이미지 데이터의 상위 10%, 하위 10%를 활용하여 학습을 진행하였다.

이분화된 가로경관 인식 지표를 학습하는 데 있어 대표적 몇 가지 모형을 비교하고 성능이 가장 좋은 모형을 사용하였다(표 3). 비교 결과 최종적으로 사용한 기계학습 모형은 CATBoost이다.

CATBoost는 기존 경사 하강법 기반의 모형에 범주형 피처 처리를 위한 알고리즘을 도입하여 예측 성능이 우수한 모형이다. 해당 모형의 정확도는 ‘활발함’이 75%, ‘안전함’이 80%로 나타나 두 가지 지표 모두에서 75% 이상의 정확도를 나타내고 있는 것으로 나타났다. 또한, 모형의 전반적인 성능을 의미하는 Area Under the ROC Curve(AUC)의 경우에도 정확도와 유사한 결과를 나

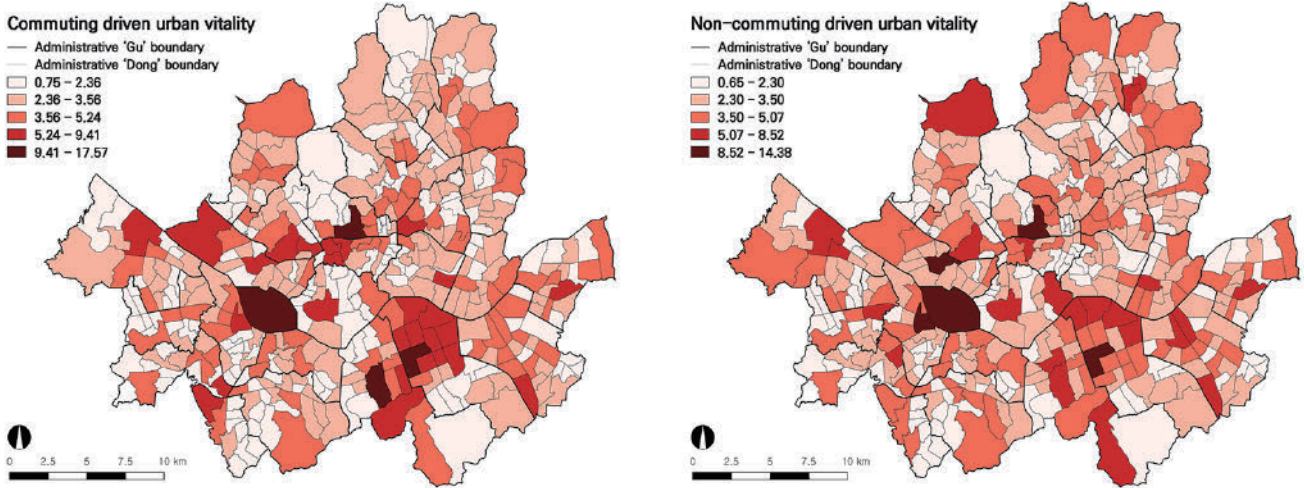


그림 3. 통행 목적별 도시활력 지수 (좌: 통근 통행, 우: 비통근 통행)
 Figure 3. Urban vitality index by trip purpose (left: commute, right: non-commute)

표 3. 가로경관 인식 정량화 모형 평가 결과

Table 3. Evaluation results of street view perception quantification model

Model	Lively			Safety		
	Accuracy	AUC	F1-Score	Accuracy	AUC	F1-Score
CATBoost Classifier	0.75	0.818	0.754	0.80	0.871	0.804
Random Forest Classifier	0.74	0.813	0.741	0.75	0.825	0.749
Gradient Boosting Classifier	0.74	0.815	0.747	0.75	0.827	0.752

타내고 있음을 알 수 있다.

이후, 앞서 언급했다시피 최종적으로 학습된 기계학습 모형에 서울시의 NSV 파노라마 이미지를 활용하여 가로경관에 대한 인식의 예측값을 도출하였다. <그림 4>는 본 연구에서 활용한 가로경관 인식 지표인 활발함(Lively) 인식 점수가 높은 지역과 낮은 지역에 대한 예시 이미지이다. 대체적으로 활발함(Lively) 인식

이 높은 지역은 간판과 자동차 등이 많은 상업지역 인근의 이미지가 많이 도출된 것을 알 수 있다. 그리고 안전함(Safety) 인식이 높은 지역은 도로가 넓은 주거지역 인근의 가로경관의 이미지가 많이 도출되었음을 알 수 있다. 즉, 교통사고로부터의 안전, 범죄로부터의 안전, 보행 안전 등 가로경관의 전반적인 안전 인식은 잘 정돈된 가로경관의 안전 지표가 높게 나타남을 알 수 있다.



그림 4. 서울시 가로경관에 대한 활발함과 안전함 인식 예시
 Figure 4. Example of perception of lively and safety streetscape in Seoul

4. 도시활력 예측 모형 평가

본 연구에서는 도시활력에 영향을 미치는 요인을 도출하고, 변수의 선형 또는 비선형 관계를 도출하기 위해 기계학습을 활용하였다. 대표적으로 사용되고 있는 몇 가지 기계학습 모형의 성능을 매개변수 조정 후 비교하고, 가장 성능이 좋은 모형을 선택하였다(표 4). 최종적으로 본 연구에 활용된 기계학습 모형은 MLP Regressor이다. 모형의 설명력을 평가하기 위해 80%의 학습 데이터와 20%의 평가 데이터로 분할하였다. 즉, 학습 데이터는 전체 424개의 데이터 중 339개, 평가 데이터는 85개의 데이터로 할당되었다. 이후, 모형의 과적합을 방지하고 예측력을 높일 수 있도록 모형의 학습에 영향을 미치는 주요 매개변수를 조정하였다. 최종적으로 학습된 모형에 대해 평가 데이터의 독립변수를 활용하여 평가 데이터의 종속변수인 도시활력을 예측하도록 하였다. 이후 전통 회귀분석 결과와 비교하였으며, 기계학습 모형인 MLP Regressor의 결정계수(R^2)는 통근 통행에서 0.34, 비통근 통행에서 0.26으로 전통 회귀모형의 결정계수인 0.20, 0.17보다 높게 나와 도시활력 영향요인을 분석하는 데 있어 기계학습 모형이 더 적절함을 알 수 있다.

〈표 4〉는 통근 통행 도시활력과 비통근 통행 도시활력에 대한 MLP Regressor의 모형 평가 결과 및 조정된 매개변수를 나타낸 표이다. 조정된 매개변수로는 먼저 은닉층 레이어의 개수와 뉴런의 수를 조정하는 'hidden_layer_sizes', 뉴런이 학습을 진행할 때의 학습률을 조정하는 'learning_rate', L2 페널티를 활용하여 모델의 복잡도를 제어하는 'alpha', 모형의 최대 학습 횟수를 설정하는 'max_iter', 한 번의 연산에 들어가는 데이터의 크기를 설정하는 'batch_size'이다. 또한 뉴런의 활성화 함수는 'tanh', 가중치 최적화를 위한 함수는 'adam'을 사용하였다. 최적 파라미터를 찾는 방법으로는 그리드를 설정하여 최적의 매개변수를 찾는 GridSearchCV 검증을 활용하였다.

5. 해석가능한 기계학습 방법론을 활용한 도시활력 영향요인 분석

〈그림 5〉와 〈그림 6〉은 본 연구에서 사용한 MLP Regressor 모형에 대해 해석가능한 기계학습인 SHAP를 활용하여 통행 목적별로 변수의 중요도를 도출한 결과이다. 먼저, 통근 통행의 경우, 음식점 POI 밀도, 주거시설 POI 밀도, 평균 건물 나이 등의 순으로 도시의 활력에 영향을 미치는 것으로 나타났다. 구체적으로 가장 영향력이 높은 음식점의 경우, Local Shapley Value 값이 높은 붉은 점이 Y축을 기준으로 양(+)의 방향에 존재하여 도시활력과는 긍정적인 관계를 가지고 있다고 볼 수 있다. 이와 반대로, 주거시설 POI 밀도, 평균 건물 나이의 경우에는 Local Shapley Value의 값이 높은 붉은 점이 X축의 0을 기준으로 음(-)의 방향에 존재하여 도시활력과는 부정적인 관계를 가지고 있는 것으로 나타났다.

비통근 통행의 경우, 주거시설 POI 밀도, 지하철 출구 밀도, 음식점 POI 밀도 등의 순으로 도시의 활력에 영향을 미치는 것으로 나타났다. 주거시설 POI 밀도는 통근 통행과 비통근 통행 모두에서 도시활력에 높은 부정적인 관계를 가지고 있는 것으로 나타났다. 이는 주거시설의 밀도가 높을수록 출발지로서의 역할을 많이 하게 되기 때문으로 판단된다. 이와 반대로 음식점 POI의 경우, 통근 통행과 비통근 통행 모두에서 도시활력에 높은 양(+)의 관계를 가지는 것으로 나타났다. 이는 음식점 경우 통근 통행의 목적지인 주요 도심에 많이 분포하고 있고 비통근 통행의 주요 목적지로 기능을 하고 있기 때문으로 판단된다.

POI 특성을 살펴보면, 통근 통행 도시활력과 양(+)의 관계가 있는 요인은 음식점, POI 혼합도, 제조업시설 등으로 나타났다. 이와 반대로 주거시설, 스포츠·레저시설 등의 시설은 통근 통행 시 도시활력과 음(-)의 관계를 가지는 것으로 나타났다. PageRank 값은 해당 노드의 아웃바운드 링크(통행 유출)보다는 인바운

표 4. 도시활력 영향요인 분석모형 평가 결과 및 조정 매개변수

Table 4. Assessment results of analysis models for influencing factors of urban vitality and adjustment parameters

Model	Commuting driven urban vitality			Non-commuting driven urban vitality		
	R^2	MSE	MAE	R^2	MSE	MAE
Linear Regression	0.20	5.313	1.150	0.17	2.933	1.166
Gradient Boosting Regressor	0.26	4.924	1.095	0.14	3.038	1.130
Random Forest Regressor	0.10	6.005	1.282	0.17	3.161	1.128
MLP Regressor	0.34	4.419	1.048	0.26	2.611	1.038
Parameters	Hidden layer sizes	250,10,100		200,100,200		
	Max iter	1,000		1,000		
	Alpha	0.9		0.7		
	Batch size	64		64		
	Learning rate	adaptive		adaptive		

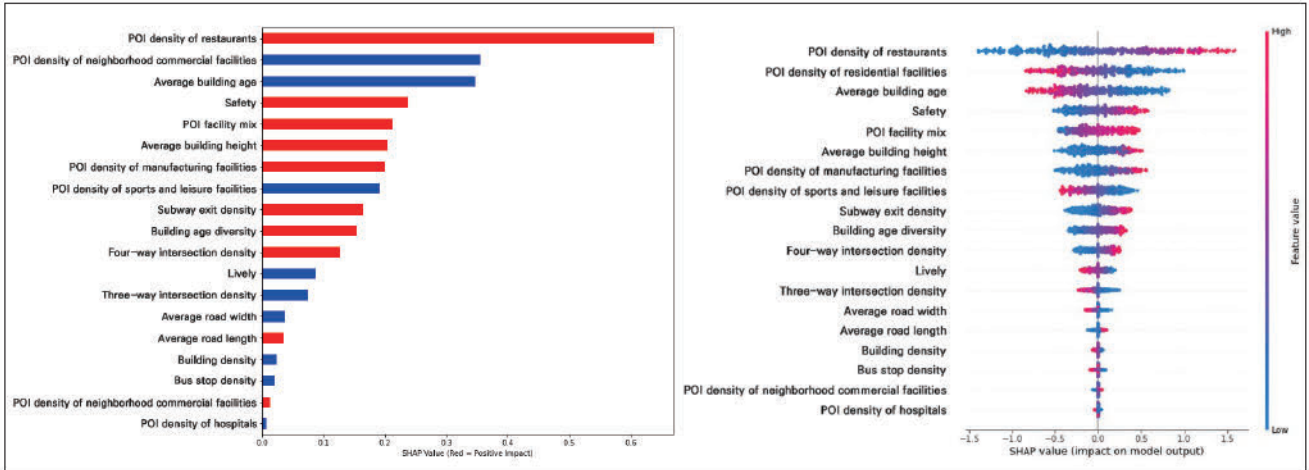


그림 5. 통근 통행 도시활력에 영향을 미치는 SHAP 변수 중요도

Figure 5. Importance of SHAP variables influencing commuting driven urban vitality

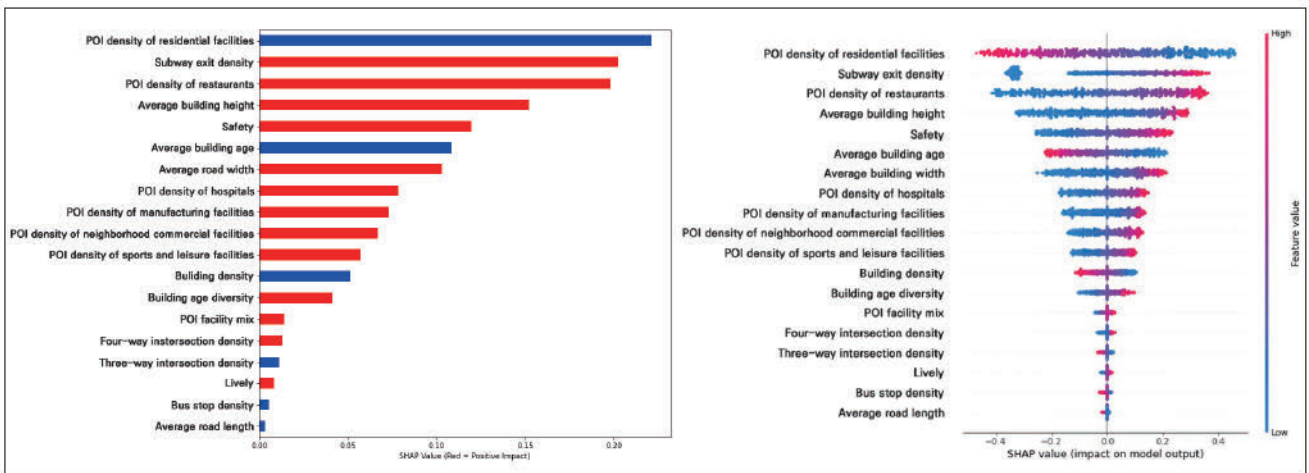


그림 6. 비통근 통행 도시활력에 영향을 미치는 SHAP 변수 중요도

Figure 6. Importance of SHAP variables influencing non-commuting driven urban vitality

드 링크(통행 유입)가 높을 때 점수가 올라간다. 따라서 주거시설 및 스포츠·레저시설 등의 경우에는 통근 통행의 목적지로 활용될 확률이 낮아 음(-)의 관계가 도출된 것으로 판단된다. 이를 보았을 때, 통근 통행의 목적지로 활용될 수 있는 POI 시설이 많이 존재할수록, 출발지, 또는 비통근 통행 시의 목적지로 활용될 수 있는 POI 시설이 적게 존재할수록 통근 통행 도시활력이 증가함을 알 수 있다.

이와 반대로, 비통근 통행의 경우에는 음식점, 병원, 스포츠·레저시설 등 비통근 통행에서의 목적지로 활용될 수 있는 시설이 많이 존재할수록 도시의 활력이 증가함을 알 수 있다. 병원 시설 POI의 경우 도시활력과 양(+)의 관계로 나타났는데, 이는 한의원의 수가 고령인구의 비통근 통행에 긍정적인 영향을 미친다는 고승욱·이승일(2017)의 연구와 비슷한 맥락의 결과로 판단된다. POI 혼합도의 경우, 통근 통행과 비통근 통행 모두에서 도시활력에 긍정적인 영향을 미치는 것으로 나타났다. 이는 통행의 목적에 상관없이 다양한 목적의 POI가 존재하는 곳일수록 다양한 목

적의 통행이 발생하기 때문으로 판단할 수 있다. 이는 비목적 여가통행의 경우 건축물 주 용도 수가 많을수록 보행 만족도에 긍정적인 영향을 미친다는 김창국(2015)의 연구와 유사한 맥락으로 해석할 수 있다.

건축물 특성의 경우 높은 건물이 많을수록 통근 통행과 비통근 통행 모두에서 도시의 활력에 긍정적인 영향을 미치는 것으로 나타났다. 건물 나이의 경우에는 건물 나이가 적을수록, 건물 나이의 다양성이 높을수록 도시활력에 긍정적인 영향을 미치는 것으로 나타났다. 건물 나이 다양성의 경우에는 Jacobs(1961)가 언급했던 도시활력에 영향을 미치는 주요 요인과 일치하는 결과가 도출되었다. 이는 시간이 흘러감에 따라 다양한 요소들이 복잡하고 유기적으로 발달하여 도시의 활력에 긍정적인 역할을 함을 의미한다. 다만, 건물 밀도의 경우에는 두 경우 모두에서 영향력이 낮은 것으로 나타났으며, 음(-)의 관계를 가지고 있는 것으로 나타났다. 이는 POI 변수와 함께 생각해보았을 때, 도시활력을 높일 수 있는 시설의 효율적인 입지가 도시의 활력에 긍정적인 역할을

함을 의미한다.

교통환경 특성의 경우, 3지 교차로와 4지 교차로의 밀도는 서로 반대의 결과가 도출되었다. 그중 블록의 크기를 대변할 수 있는 4지 교차로의 경우, 통근 통행과 비통근 통행 모두에서 양(+)의 영향을 미치는 것으로 나타났다. 이는 블록의 크기가 작을수록 다양한 목적의 통행이 형성되어 도시의 활력에 긍정적인 역할을 한다는 Jacobs(1961)의 주장과 일치하는 결과로 볼 수 있다. 크기가 작은 블록은 경로 선택에 대한 다양성이 높다. 이동의 목적지가 같은 경우에도 똑같은 길로 갈 필요가 없으며, 여러 길을 선택하여 이동하는 것이 가능하게 하기 때문이다. 따라서 작은 크기의 블록은 가로 이용의 유동성과 다양한 경로의 혼잡을 야기시켜 도시의 활력에 긍정적인 역할을 한다고 볼 수 있다. 대중교통으로의 접근성을 나타내는 변수 중 하나인 지하철 출구 밀도의 경우, 비통근 통행에서 도시의 활력에 긍정적인 영향을 미치며 높은 영향력을 가지는 것으로 나타났다. 이는 대중교통의 활성화 측면과 지하철 역사 인근 상권의 활성화 측면에서 비통근 통행을 많이 발생시킨 것으로 해석할 수 있다. 다만, 버스 정류장의 경우 두 경우 모두에서 도시의 활력에 낮은 영향력을 미치는 것으로 나타났다.

마지막 가로경관 인식 특성의 경우, 안전함(Safety) 인식은 통근 통행과 비통근 통행 모두에서 높은 영향력을 가지며, 도시활력을 높이는 데 기여하고 있음이 나타났다. 필연적으로 발생하는 통근 통행과 그렇지 않은 비통근 통행 두 경우 모두에 긍정적인 역할을 하는 것으로 보아, 안전한 가로경관은 도시민의 통행을 증가시키는 데 기여를 하고 있음을 알 수 있다. 반면에, 활발함(Lively)의 경우에는 필연적으로 발생하는 통근 통행의 경우에는 음(-)의 영향을 미치는 것으로 나타났으며, 그렇지 않은 비통근 통행의 경우에는 양(+)의 영향을 미치는 것으로 나타났다. 따라서, 활발한 가로경관은 도시민의 비통근 통행을 증가시켜 도시활력에 긍정적인 영향을 미치는 것으로 볼 수 있다.

본 연구는 국가적 현상인 저출산의 지역 간 격차에 주목하고, 이를 유발하는 주요 요인으로서의 지역 주택가격에 초점을 맞추었다. 지역의 주택가격은 결혼과 출산을 위한 주거공간을 마련하는 데 가장 중요하게 고려되는 사항이기 때문이다.

6. 도시활력 비선형관계 분석

〈그림 7〉은 Shapley Value가 통행의 목적별로 차이가 크거나, 반대로 도출된 주요 변수에 대해 단일 변수의 밀도 산점도를 나타낸 결과이다. 먼저, 근린 상업시설은 Shapley Value가 0인 구간이 있는 것을 통해서 통근 통행에 영향을 미치지 않는 행정동이 나타났다. 하지만, 비통근 통행은 밀도가 약 250 이상인 경우 도시활력에 긍정적인 영향을 미치는 것으로 나타났다. 그리고 근린 상업시설의 밀도가 약 650 이상인 구간에서는 비통근 통행 도시

활력 증가에 미치는 영향이 일정해지는 비선형 관계로 나타났다. 이는 비통근 통행에 의한 도시활력에 있어 적절한 근린 상업시설의 밀도가 존재함을 시사한다.

스포츠·레저시설 POI의 경우에는 통행의 목적별로 반대의 결과가 도출되었다. 특히, 비통근 통행의 경우 도시활력과 양(+)의 관계를 갖는 것으로 나타났다. 이는 스포츠·레저시설의 밀도가 통근 통행보다 비통근 통행에 밀접한 연관이 있는 것으로 볼 수 있다. 다만, 두 경우 모두에서 약 50 정도의 밀도보다 높아질 경우 도시의 활력에 미치는 영향력이 일정해지는 비선형 관계가 도출되었다. 이는 비통근 통행 시 도시활력 증진을 위한 스포츠·레저 시설의 추가 설치에 레저시설 이용에 대한 일정 밀도 이상으로 충족되면 도시활력에 미치는 영향력은 일정해진다는 것으로 이해할 수 있다.

다음으로 교차로 변수 중 4지 교차로는 통근 통행과 비통근 통행 모두에서 양(+)의 관계를 가지는 것으로 나타났지만, 일정 수준 이상에서 완만해지는 비선형 관계가 있는 것으로 나타났다. 4지 교차로의 밀도가 높을수록 블록의 크기가 작아 도시민의 이동에 있어 경로 선택의 다양성을 증대시키고, 도시활력에 긍정적인 역할을 한다는 것으로 이해할 수 있다. 다만, 4지 교차로의 밀도가 통근 통행에서는 28, 비통근 통행에서는 25를 기점으로 SHAP Value가 음(-)에서 양(+)으로 바뀌는 것으로 나타났다. 이는 너무 큰 크기의 블록(4지 교차로 밀도 25 미만)은 도시의 활력에 오히려 부정적인 영향을 미칠 수 있음을 시사한다. 통근 통행과 비통근 통행의 경우 4지 교차로의 밀도가 적정 수준 이상일 때 도시활력에 긍정적인 영향을 미칠 수 있음을 의미한다. 그러나 4지 교차로의 밀도가 증가한다고 해도 통근 통행의 경우 100 이상일 때, 비통근 통행의 경우 65 이상일 때 도시활력은 더 이상 증가하지 않는 비선형 관계가 있는 것으로 나타났다.

마지막으로 가로경관 인식 변수 중 안전함(Safety)은 통근 통행과 비통근 통행 모두에서 도시활력과 양(+)의 관계를 가지고 있음을 알 수 있다. 비통근 통행을 중심으로 보면, 안전함 지수가 0.42 이상일 때 도시활력에 양의 영향을 가지고 있지만 약 0.55 이상일 때 효과가 일정해지는 비선형 관계가 도출되었다. 이는 도시민이 가로경관을 통해 느끼는 안전함 인식이 일정 수준까지 도시의 활력을 증대시키는 데 중요함을 의미한다. 따라서 도시민들에게 안전한 인식을 줄 수 있는 보안등, 비상벨, CCTV와 같은 안전시설을 확충하여 도시활력을 증대시키는 것이 필요함을 시사한다.

반면, 활발함(Lively)은 통근 통행과 비통근 통행에서 반대의 패턴을 가지는 것으로 나타났다. 통근 통행에서는 음(-)의 관계로 나타났으며, 비통근 통행에서는 양(+)의 관계를 가지는 것으로 나타났다. 가로경관의 활발함이 높은 지역은 통근 통행의 목적지인 업무 중심지이기보다는 여가, 쇼핑 등 상업 활동이 많이 일어나는 지역으로 볼 수 있다. 따라서 비통근 통행의 경우 가로경관

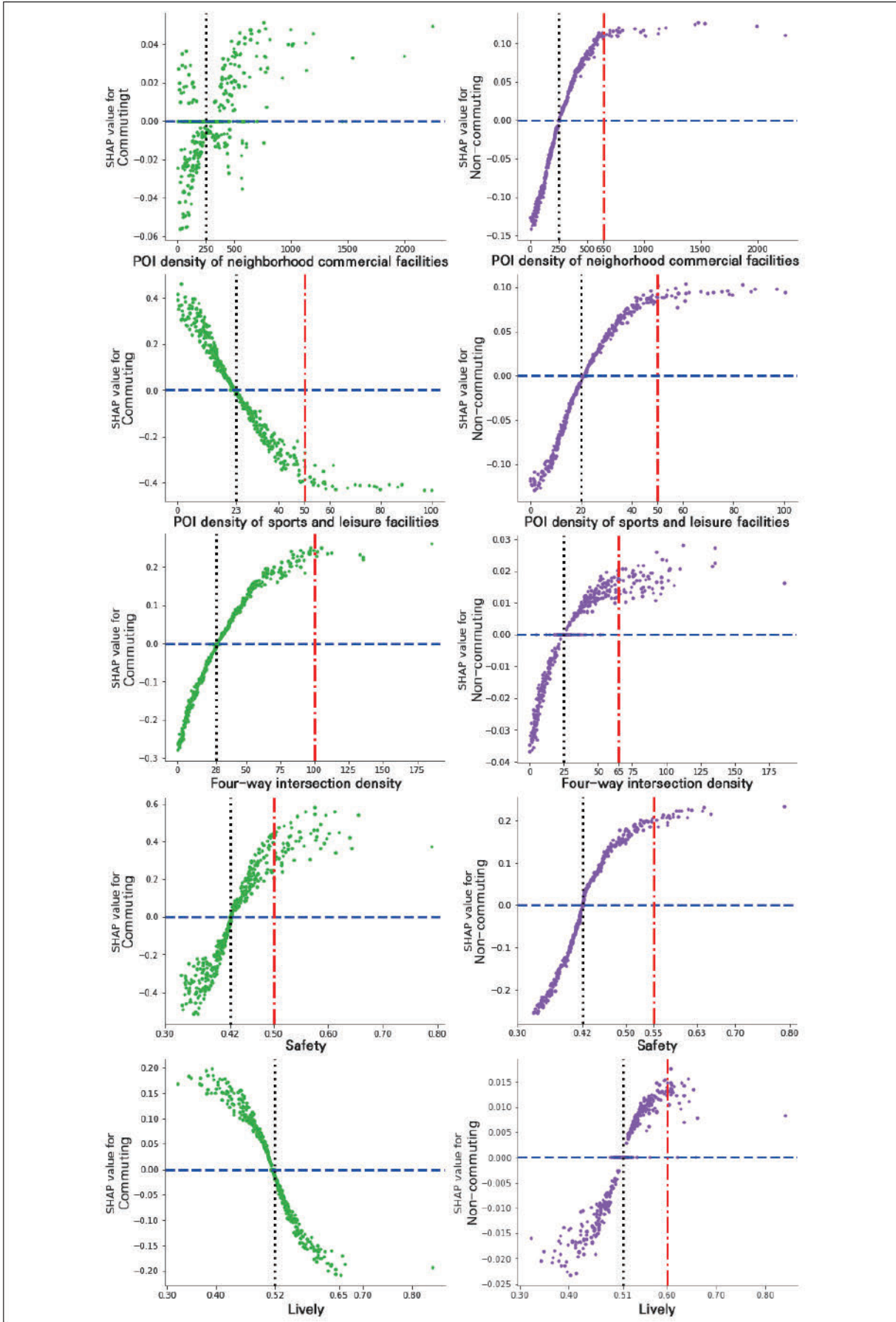


그림 7. 동행목적별 도시활력에 영향을 미치는 단일 변수에 대한 산점도 (왼쪽: 통근 통행, 오른쪽: 비통근 통행)
 Figure 7. Scatterplot of single variable affecting urban vitality by trip purpose (left: commute, right: non-commute)

의 활발함 지수가 0.52 이상일 때 가로 활력이 양의 영향을 가지는 것으로 나타났으며 0.6 이상일 때 효과가 감소하는 비선형 관계로 나타났다.

V. 결론

본 연구에서는 2021년 1년 동안의 서울시 생활이동 데이터와 PageRank 알고리즘을 활용하여 행정동별 도시활력을 정량화하였다. 추가적으로, 가로경관에 대한 사람들의 인식을 정량화하여 도시활력에 영향을 미치는 영향요인을 분석하였다. 나아가 분석 방법론으로 해석가능한 기계학습을 활용하여 도시의 활력에 영향을 미치는 변수별 기여도의 선형 또는 비선형 관계를 분석하였다. 분석 결과에 따른 종합적인 결론은 다음과 같다. 먼저, 도시의 활력을 측정하는 데 있어 실제 이동이 발생한 데이터를 활용하여 전통적으로 사용되었던 인원 집계방식의 한계점을 보완하였다. 또한, 분석 방법론으로 해석가능한 기계학습을 활용하여 도시활력에 영향을 미치는 요인을 분석하고, 각 변수의 기여도를 도출하였다. 나아가 통행목적별 통근 통행과 비통근 통행으로 나누어 도시활력 영향요인의 비선형관계를 도출하였다.

POI 특성 중 주거시설 밀도와 음식점 밀도는 도시의 활력에 가장 큰 영향을 미치며, 주거시설 밀도가 낮을수록, 음식점 밀도가 높을수록 도시의 활력이 증가하는 것으로 나타났다. 이와 더불어, POI 혼합도는 높을수록 도시활력에 긍정적인 역할을 하는 것으로 나타났다. 이는 POI 시설의 다양성이 높을수록 다양한 목적의 통행 수요를 유발하며, 이에 따라 도시의 활력이 증가함을 의미한다. 특히, 통근 통행에서는 음식점, 제조업 시설이 중요한 역할을 하며, 비통근 통행에서는 병원, 근린 상업시설, 스포츠·레저시설 등이 중요한 역할을 하는 것으로 나타났다. 통행의 목적에 따른 POI 시설의 높은 혼합이 도시의 활력에 긍정적인 영향을 미칠 수 있음을 시사한다. 다양한 용도가 효과적으로 혼합된 시설들이 각기 다른 목적을 가진 사람을 각기 다른 시간에 이동시키기 때문이다. 또한, 대부분의 도시활력 영향요인은 선형관계보다는 특정한 임계값 이상을 넘어갈 때 효과가 감소하는 비선형관계가 있음을 확인하였다.

건축물 특성의 경우 건물의 나이가 적고 다양할수록 도시의 활력이 증가한다는 결과가 나타났다. 다만, Jacobs(1961)의 주장과는 다르게 건물 나이의 평균이 낮을수록 도시활력이 높아진다는 결과가 함께 나타났으며, 건물 나이의 다양성보다 변수 중요도가 높은 것으로 나타났다. 이는 서울시의 경우 행정동 단위에서 건물의 나이가 다양하고 새로운 건축물이 많을수록 도시활력이 높은 것을 의미한다. 이러한 결과는 오랜된 건물이 많고 건축물 연령의 다양성이 높은 근린에서 도시활력이 높다고 주장한 Jacobs(1961)의 이론과는 약간의 차이가 있는 결과이다. 또한, Jacobs의 도시활력 이론을 실증분석한 Sung and Lee(2015)의

연구결과와도 차이가 있다. 이러한 차이는 도시활력의 정의(definition)와 측정방법에 따라 나타날 수 있는 문제로 볼 수 있으며, 향후 추가적인 연구가 필요할 것으로 판단된다.

다음으로, 교통환경 특성 중 4지 교차로 밀도의 경우 통근 통행과 비통근 통행 모두에서 도시활력과 양(+)의 관계를 가지는 것으로 나타났다. 블록의 크기가 작을수록 다양한 통행 수요를 만드는 데 긍정적인 역할을 하며 도시의 활력을 증가시키는 데에 기여할 수 있음을 의미한다. 하지만, 4지 교차로 밀도는 도시활력과 비선형 관계를 가지고 있음을 확인하였다. 이는 4지 교차로 밀도가 일정 수준 이상일 때 도시활력에 더 이상 큰 영향을 미치지 않음을 의미한다. 이러한 분석결과는 도시활력을 증진시키기 위해서는 적정 수준의 4지 교차로 밀도가 존재함을 시사한다.

마지막으로, 가로경관에 대한 인식은 통행 목적별로 도시의 활력에 유의미한 영향을 미치고 있는 것으로 나타났다. 도시의 활력에 가장 큰 영향을 미친 안전함(Safety)의 경우, 통근 통행과 비통근 통행 모두에서 도시의 활력과 양(+)의 비선형 관계를 가지고 있는 것으로 나타났다. 이는 통행의 목적에 상관없이 안전한 가로경관은 도시민의 활동을 증가시키는 데에 일정 수준 긍정적인 역할을 한다는 것으로 볼 수 있다. 따라서 도시활력을 높이기 위해서는 보행자 교통사고나 범죄 등으로부터 안전한 가로환경 조성이 매우 중요함을 시사한다. 이에 반해, 활발함(Lively)의 경우 통근 통행에서는 음(-)의 관계를 가지는 것으로 나타났으며, 비통근 통행에서는 양(+)의 관계를 가지고 있는 것으로 나타났다. 활발한 가로경관이 비통근 통행에 긍정적인 역할을 한다는 것은 가로경관에 대한 인식이 도시민의 여가 활동에 선호, 동기를 부여할 수 있음을 시사한다. 따라서 활력 있는 가로경관의 조성은 도시활력을 높이는 데 긍정적인 영향을 미칠 수 있을 것으로 판단된다.

한편, 본 연구는 다음과 같이 몇 가지의 한계점을 가진다. 먼저, 연구의 시간적 범위인 2021년의 데이터는 2020년부터 지속되어 온 COVID-19 감염병으로 인해 익숙해진 비대면 생활로 통행의 패턴이 어느 정도 정착되어 있을 수 있으나, COVID-19의 영향을 완전히 제어하지는 못했다는 한계가 존재한다. 또한, 본 연구에서 가로경관 인식을 정량화하기 위해 학습 데이터로 사용하였던 Place Pulse 2.0 데이터는 세계 여러 나라의 가로경관 이미지를 활용하여 구축되었다. 현대 도시의 가로경관은 상당히 유사하지만 실제 우리나라 가로경관에 대해 느끼는 도시민의 인식과는 어느 정도 차이가 있을 수 있다. 마지막으로 본 연구에서 활용한 POI 데이터의 경우 상업이나 업무 용도의 경우 한 건물 안에 다양한 유형의 비즈니스를 고려할 수 있는 장점을 가지고 있다. 그러나 주거시설에서 아파트 단지의 경우 세대별 1개의 POI 자료가 아닌 건물 동별로 1개의 POI가 구축되기 때문에 아파트 단지의 정확한 규모를 고려하지 못하는 한계를 가지고 있다.

그럼에도 불구하고 본 연구는 서울 생활이동 빅데이터와 Pag-

eRank 알고리즘을 활용하여 도시활력을 정량화하고 도시활력의 유형별 영향요인을 도출하였다. 또한, 가로경관에 대한 인식 지표를 포함하여 도시의 활력에 영향을 미치는 요인을 해석가능한 기계학습 방법론을 활용하여 분석했다는 점에서 의의가 있다. 또한, 가로경관에 대한 인식이 실제 도시의 활력에 영향을 미친다는 점을 확인하였고, 도시활력의 유형을 구분하여 시사점을 제공했다는 데에 의의가 있다. 도시활력에 대한 정확한 측정과 이에 영향을 미치는 요인에 대한 이해는 도시활력을 증진시키는 데 유용하게 활용될 수 있을 것으로 판단된다.

인용문헌 References

- 고승욱·이승일, 2017. “통행목적지로서 서울 행정동의 특성이 연령인구 연령대별 비통근 통행에 미치는 영향 분석”, 『한국지역개발학회지』, 29(1): 79-98.
Go, S.W. and Lee, S.I., 2017. “A Study on Impact of Characteristics Destination Districts of Seoul on Non-commuting Travel of Elderly Population by Age Groups”, *Journal of the Korean Regional Development Association*, 29(1): 79-98.
- 권철민, 2019. 『파이썬 머신러닝 완벽 가이드』, 파주: 위키북스.
Kwon, C.M., 2019. *Python Machine Learning Perfect Guide*, Paju: Wikibooks.
- 김영룡, 2020. 「빅데이터를 활용한 도시활력 측정과 도시공간 유형 분류」, 수원: 경기연구원.
Kim, Y.L., 2020. *Measuring Urban Vitality and Typology of Urban Space using Big Data*, Suwon: Gyeonggi Research Institute.
- 김창국, 2015. “물리적 가로환경 구성요소가 보행자의 지점방문 목적별 보행만족도에 미치는 영향 분석”, 한양대학교 대학원 석사학위논문.
Kim, C.G., 2015. “Impact of Street’s Physical Environment on the Satisfaction of Pedestrian Regarding with the Different Walking Purposes in Seoul”, Master’s Dissertation, Hanyang University.
- 민병학·이건원·김세용, 2016. “토지이용특성이 통행수단 및 통행목적별 통행패턴에 미치는 효과: 서울시 행정동을 대상으로”, 『대한건축학회논문집-계획계』, 32(4): 77-87.
Min, B.H., Lee, G.W., and Kim, S.Y., 2016. “The Effects of Land-use Characteristics on Trip Patterns by Trip Modes and purposes: Focused on Seoul metropolitan administrative Division”, *Journal of the Architectural Institute of Korea Planning & Design*, 32(4): 77-87.
- 박영준·박소현, 2019. “근린지역을 중심으로 한 전업주부들의 자가용 - 비통근 통행특성”, 『서울도시연구』, 20(1): 45-59.
Park, Y.J. and Park, S.H., 2019. “Car-use and Non-work Travel of Housewives in Residential Neighborhood of Seoul”, *Seoul Studies*, 20(1): 45-59.
- 이남휘·최창규, 2020. “유형별 주중 여가통행의 영향요인에 관한 연구”, 『국토계획』, 55(2): 91-100.
- Lee, N.H. and Choi, C.G., 2020. “A Study on the Influence Factors of Weekday Leisure Travel”, *Journal of Korea Planning Association*, 55(2): 91-100.
- 이수기·고준호·이기훈, 2016. “근린환경특성이 보행만족도에 미치는 영향 분석: 서울서베이 2013년 자료를 중심으로”, 『국토계획』, 51(1): 169-187.
Lee, S., Ko, 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.
- 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.
- 정시윤·전병운, 2020. “공간 빅데이터와 야간 위성영상을 활용한 도시 활력 평가: 대구시를 사례로”, 『한국지리정보학회지』, 23(4): 217-233.
Jeong, S.Y. and Jun, B.W., 2020. “Urban Vitality Assessment using Spatial Big Data and Nighttime Light Satellite Image: A Case Study of Daegu”, *Journal of the Korean Association of Geographic Information Studies*, 23(4): 217-233.
- 정혜선, 2021. “코로나-19 4차 대유행을 넘어 워드 코로나의 시대로”, 『월간 공공정책』, 193: 61-63.
Jeong, H.S., 2021. “Beyond the COVID-19 4th Pandemic, to the Era of with COVID”, *Monthly Public Policy*, 193: 61-63.
- 조월·이수기, 2021. “서울시 POI 빅데이터를 활용한 도시활력과 영향요인 분석”, 『국토계획』, 56(7): 87-102.
Cao, Y. and Lee, S., 2021. “Analysis of Urban Vitality and Its Determinant Factors using POI Bigdata in Seoul, Korea”, *Journal of Korea Planning Association*, 56(7): 87-102.
- Azmi, D.I. and Karim, H.A., 2012. “Implications of Walkability towards Promoting Sustainable Urban Neighborhood”, *Procedia - Social & Behavioral Sciences*, 50: 204-213.
- Brin, S. and Page, L., 1998. “The Anatomy of a Large-scale Hypertextual Web Search Engine”, *Computer Networks and ISDN Systems*, 30(1998): 107-117.
- Bromley, R.D. and Thomas, C.J., 2002. “Food Shopping and Town Centre Vitality: Exploring the link”, *The International Review of Retail, Distribution and Consumer Research*, 12(2): 109-130.
- Bzdok, D., Krzywinski, M., and Altman, N., 2017. “Machine Learning: A Primer”, *Nature Methods*, 14(12): 1119-1120.
- Caraviello, D.Z., Weigel, K.A., Craven, M., Gianola, D., Cook, N.B., Nordlund, K.V., Fricke, P.M., and Wiltbank, M.C., 2006. “Analysis of Reproductive Performance of Lactating Cows on Large Dairy Farms using Machine Learning Algorithms”, *Journal of Dairy Science*, 89(12): 4703-4722.
- Datta, R., Joshi, D., Li, J., and Wang, J.Z., 2006. “Studying Aesthetics in Photographic Images Using A Computational Approach”, *Computer Vision*, 3953: 288-301.

19. Filion, P. and Hammond, K., 2003. "Neighbourhood Land Use and Performance: The Evolution of Neighbourhood Morphology over the 20th Century", *Environment and Planning B: Urban Analytics and City Science*, 30(2): 271-296.
20. 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.
21. 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.
22. Herbrich, R., Minka, T., and Graepel, T., 2007. "TrueSkill™: A Bayesian Skill Rating System", *Advances in Neural Information Processing Systems*, 19: 569-576.
23. Jacobs, J., 1961. *The Death and Life of Great American Cities*, New York: Random House.
24. Jia, C., Du, Y., Wang, S., Bai, T., and Fei, T., 2019. "Measuring the Vibrancy of Urban Neighborhoods using Mobile Phone Data with an Improved PageRank algorithm", *Transactions in GIS*, 23(2): 241-258.
25. Jin, X., Long, Y., Sun, W., Lu, Y., Yang, X., and Tang, J., 2017. "Evaluating Cities' Vitality and Identifying Ghost Cities in China with Emerging Geographical Data", *Cities*, 63: 98-109.
26. Katz, P., 1994. *The New Urbanism: Toward an Architecture of Community*, New York: McGraw-Hill.
27. Kim, Y.L., 2018. "Seoul's Wi-Fi Hotspots: Wi-Fi Access Points as and Indicator of Urban Vitality", *Computers, Environment and Urban Systems*, 72: 13-24.
28. Krier, L., 2009. *The Architecture of Community*, Washington, D.C.: Island Press.
29. Krizek, K.J., 2003. "Neighborhood Services, Trip Purpose, and Tour-Based Travel", *Transportation*, 30(4): 387-410.
30. Li, S., Wu, C., Lin, Y., Li, Z., and Du, Q., 2020. "Urban Morphology Promotes Urban Vibrancy from the Spatiotemporal and Synergetic Perspectives: A Case Study using Multisource Data in Shenzhen, China", *Sustainability*, 12(12): 1-24.
31. Lu, S.W., Huang, Y.P., Shi, C.Y., and Yang, X.P., 2019. "Exploring the Associations between Urban Form and Neighborhood Vibrancy: A Case Study of Chengdu, China", *ISPRS International Journal of Geo-Information*, 8(4): 1-15.
32. Lundberg, S.M. and Lee, S.I., 2017. "A Unified Approach to Interpreting Model Predictions", *Advances in Neural Information Processing Systems*, 4765-4774.
33. Michael, Y.L., Beard, T., Choi, D., Farquhar, S., and Carlson, N., 2006. "Measuring the Influence of Built Neighborhood Environments on Walking in Older Adults", *Journal of Aging and Physical Activity*, 14(3): 302-312.
34. Miles, R. and Panton, L., 2006. "The Influence of the Perceived Quality of Community Environments on Low-income Women's Efforts to Walk More", *Journal of Community Health*, 31(5): 379-392.
35. Naik, N., Kominers, S.D., Raskar, R., Glaeser, E.L., and Hidalgo, C.A., 2017. "Computer Vision Uncovers Predictors of Physical Urban Change", *Proceedings of the National Academy of Sciences of the USA*, 114(29): 7571-7576.
36. Nasar, J.L., 2008. "Assessing Perceptions of Environments for Active Living", *American Journal of Preventive Medicine*, 34(4): 357-363.
37. Odonez, V. and Berg, T.L., 2014. "Learning High-level Judgments of Urban Perception", paper presented at European Conference on Computer Vision—ECCV 2014, Switzerland: Zurich, 494-510.
38. Page, L., Brin, S., Motwani, R., and Winograd, T., 1999. *The PageRank Citation Ranking: Bringing Order to the Web*, Stanford Digital Library Technologies Project.
39. Ravenscroft, N., 2000. "The Vitality and Viability of Town Centres", *Urban Studies*, 37(13): 2533-2549.
40. Shapley, L., 1953. "A Value for n-Person Games", in *Contributions to the Theory of Games II*, edited by Kuhn, H. and Tucker, A. 307-317, Princeton: Princeton University Press.
41. Stojić, A., Stanić, N., Vuković, G., Stanišić, S., Perišić, M., Šoštarić, A., and Lazić, L., 2019. "Explainable Extreme Gradient Boosting Tree-Based Prediction of Toluene, Ethylbenzene and Xylene Wet Deposition", *Science of The Total Environment*, 653: 140-147.
42. Sung, H. 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.
43. Sung, H., 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.
44. Tom, M., 1997. *Machine Learning*, Chicago: McGraw Hill Books.
45. Wu, C., Ye, X., Ren, F., and Du, Q., 2018. "Check-in Behaviour and Spatio-temporal Vibrancy: An Exploratory Analysis in Shenzhen, China", *Cities*, 77: 104-116.
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. Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., Wang, R., and Guan, Q., 2019. "A Human-machine Adversarial Scoring Framework for Urban Perception Assessment using Street-view Images", *International Journal of Geographical Information Science*, 33(12): 2363-2384.
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. Ying, L. and Yin, Z., 2016. "Quantitative Evaluation on Street Vibrancy and Its Impact Factors: A Case Study of Chengdu", *New Architecture*, 2016(1): 52-57.

50. Yue, W.Z., Chen, Y., Zhang, Q., and Liu, Y., 2019. "Spatial Explicit Assessment of Urban Multi-source Data: A Case of Shanghai, China", *Sustainability*, 11(3): 638.
51. Yue, Y., Zhuang, Y., Yeh, A.G., Xie, J.Y., Ma, C.L., and Li, Q.Q., 2017. "Measurements of POI-based Mixed Use and Their Relationships with Neighbourhood Vibrancy", *International Journal of Geographical Information Science*, 31(4): 658-675.
52. Yuen, B. and Hien, W.N., 2005. "Resident Perceptions and Expectations of Rooftop Gardens in Singapore", *Landscape and Urban Planning*, 73(4): 263-276.
53. Zhang, A., Li, W., Wu, J., Lin, J., Chu, J., and Xia, C., 2021. "How Can the Urban Landscape Affect Urban Vitality at the Street Block Level? A Case Study of 15 Metropolises in China", *Environment and Planning B: Urban Analytics and City Science*, 48(5): 1245-1262.
54. Zhang, F., Zhou, B, Lie, L., Liu, Y., Fung, H.H., Lin, H., and Ratti, C., 2018. "Measuring Human Perceptions of a Large-scale Urban Region using Machine Learning", *Landscape and Urban Planning*, 180: 148-160.

Date Received 2022-05-26
 Reviewed(1st) 2022-07-17
 Date Revised 2022-08-12
 Reviewed(2nd) 2022-08-28
 Date Accepted 2022-08-28
 Final Received 2022-09-19