해석가능한 기계학습을 활용한 보행목적별 보행만족도 영향 요인 분석

Analysis of Influencing Factors of Walking Satisfaction by Purpose Using Interpretable Machine Learning

박준상**·이수기***
Park, Junsang·Lee, Sugie

Abstract
The uncoordinated development and expansion of cities along with industrialization and urbanization have caused various urban issues, such as traffic congestion, heavy energy consumption, and environmental pollution. To alleviate the negative impacts of the automobile-centered urban environment, pedestrian-oriented urban planning and design practices have been proposed and implemented for several decades in Seoul, Korea. However, the factors influencing the walking satisfaction by walking purpose have not been sufficiently investigated in urban planning and design literature. In addition, the majority of previous studies include attempts to diagnose only the linear relationships between the walking satisfaction and the built environment. With the recent development of an interpretable machine learning (IML) model, the nonlinear relationships between the walking satisfaction and the built environment is investigated in this study. Further, the interaction effects of the built environmental variables on the walking satisfaction are identified. The results indicate that the machine learning model shows a significantly higher explanatory power compared with the conventional model. In addition, it is confirmed that IML is a useful tool to understand the nonlinear relationships between the walking satisfaction and the built environment. The analysis results on the aforementioned relationship suggest that they can be used as important data to promote a pedestrian-friendly urban environment. Additionally, the policy implications of promoting walking satisfaction by purpose are presented in this study.

Keywords
Walking Satisfaction, Interpretable Machine Learning, Physical Environment, Nonlinear Relationship, Interaction Effect

1. 서론
산업화와 도시화에 따른 도시의 무분별한 개발과 외연적 방향은 도시 내 공간 구조를 자동차 중심적으로 변화시켰으며 이로 인한 교통체증, 에너지 소비, 환경오염은 주요 도시문제로 부상하였다(조혜빈·이수기, 2016). 이와 같은 자동차 중심의 도시구조가 미치는 부정적인 영향을 완화하기 위해 그동안 보행자 중심의 도시계획 방안과 공공정책이 제시되어 왔다. 보행자 중심의 보행 친화적 도시환경은 교통, 환경, 에너지 소비 배관 변화에 영향을 미치기적으로 논의되고 있는 도시문제를 해결할 수 있을 것이다.
로 기대되기 때문이다.


다른 한편으로, 최근 도시계획 분야에 활용되기 시작한 기계 학습 모형은 높은 성능과 인식적 관계 분석의 장점에도 불구하고 적합성이 미흡한 실정이다. 기계학습 방법론은 일반적으로 기존의 전통적인 회귀모형보다 높은 설명력을 가지고 있다. 그러나 알고리즘을 통해 훈련된 모형은 복잡성을 간주하여 각 설명 변수가 종속변수의 예측에 얼마나 기여하였는지 알 수 없다는 단점이 있다. 하지만 최근 들어 담당 일정을 포함한 해석 가능

한 기계학습(interpretable machine learning) 모형의 발전으로 보행행보를 어느 정도 해석할 수 있게 되면서 도시 및 교통계획 분야에서 활용되기 시작하였다(은운호, 이수기, 2019; 조희민, 2021). 기계학습 모형의 복잡성을 해석하는 방법에는 여러 가지가 있다. 그 중 Partial Dependence Plot(PDP)는 모형에 포함된 모든 변수의 평균적인 영향을 고려하여 종속변수에 대한 설명변수의 한계효과(marginal effect)를 그래프로 설명하는 방법이다(Chung, 2013; Friedman, 2001). PDP 방법인 설명변수와 종속변수 사이의 선형관계 또는 비선형 관계를 분석할 수 있는 장점이 있다.

따라서 본 연구는 전통적인 회귀모형과 비교하여 보행만족도 분석 모형으로 기계학습 모형의 우수성을 살펴보고, 해석 가능한 기계학습을 활용하여 보행만족도와 영향을 미치는 도시의 물리적 환경 요인을 보행특성을 도출한다. 나아가 보행만족도에 영향을 미치는 변수의 상관작용 효과를 분석하고 보행만족도 측정을 위한 정책적 시사점을 도출한다.

II. 선형연구 고찰

1. 보행만족도 영향요인 연구


보고자학적 보행만족도에 영향을 미치는 근거 건조환경에 대한 분석의 강령국 외(2016)는 높은 흔적적 목적의 보행만족도가 그 목적의 보행만족도보다 높다는 결과를 도출하였으며, 도시민이 만족하는 보행환경의 특성이 보행의 목적에 따라 다르다는 것을 확인하였다. 이와 비슷한 맥락에서 보행자의 지점 우선적 특성 보행만족도에 영향을 미치는 구체적 구성요소에 대해 분석한 강령국(2015)은 보행만족도에 영향을 미치는 요인이 보행의 목적
에 따라 통계적으로 유의한 차이가 있다고 주장하였다. 추가적으 로, 보행의 유의미한 모형이 통합적인 목적의 모형보다 설명력이 높게 나타나므로 보행의 성격이 무엇보다 보행만족도에 영향을 미치는 변수들로 더 명확하게 나타났다고 주장하였다.

또한, 딥러닝과 구글 Street View API를 활용하여 가로 이지자의 보행만족도를 분석한 박근익(2021)의 연구에서는 위치적 분석을 활용한 도로 이용에서 나타난 구성요소들을 도출하고, 각 요소들 중 보행만족도와 이와 같은 관계를 보이는지 분석하였다. 분석 결과는 가로 이지지에서 하는, 나무, 식물 등의 이용은 보행만족도와 양(+)의 상관관계가 나타났으며, 반대로 건물, 벽, 차량 등은 보행만족도와 음(-)의 상관관계가 나타났다.

마지막으로, 머신러닝 및 컴퓨터 비전 기법을 활용하여 보행만족도에 영향을 미치는 가로의 물리적 특성과 분석한 이승호(2021)의 연구에서는 보행만족도에 영향을 많이 미치는 가로관련의 시각적 요소들이 도로 밀도, 비율, 보도면적 비율, 가로시내의 면적 비율 등이 나타났다. 따라서 보행만족도가 높은 가로를 조성하기 위해서는 보행자의 눈높이에서 경험하는 시각적 물리적 환경의 중요성을 알 수 있다.

2. 기계학습 방법론

기계학습은 머신러닝(machine learning)이라고도 불리며, 인간지능의 한 분야로 데이터의 패턴을 분석하여 알고리즘의 예측 오류를 최소화하는 수학 기법을 적용하여, 신뢰도 있게 결과를 예측하는 기법이다(권철민, 2019).


이처럼 머신러닝 방법론은 비선형 관계에 대한 데이터를 효과적으로 해석하는 것이 가능하다. 또한, 최근 들어 해석 가능한 기계학습 방법론을 활용하여 비선형적인 문제를 해석하는 것으로 연구의 영역이 넓어지고 있다. 본 연구의 모형은 다양한 종류의 머신러닝 모형을 비교하고, R-Squared 값과 표준화오차(MAE), 표준편차요인(MSE)을 비교하여 가장 우수한 머신러닝 모형을 선정하여 분석에 활용하고 있다. 나아가 해석 가능한 기계학습을 활용하여 최종적으로 선정된 모형의 특별적 특성을 해석함으로써 모형의 예측력에 영향을 미치는 요소들을 도출하고 있다. 따라서 해석 가능한 기계학습 방법론은 도시의 물리적 환경을 분석하는 데 있어 활용 가능성이 높다는 것을 알 수 있다.

3. 연구의 차별성

성행연구 결과에 따른 학문적 결과는 다음과 같다. 먼저, 기존의 전통적인 분석 방법은 기계학습에 비해 설명력이 낮음에도 불구하고 모형의 특별적 특성을 해석할 수 있다는 한계점으로 인해 설명력이 낮은 전통적인 회귀분석을 활용한 연구가 주를 이루었다. 둘째, 보행 활동은 물리적 환경과 성행의 관계를 가정하기에는 더욱 복잡한 비선형 관계가 있음을 볼 때 보행 연구의 대부분이 보행만족도에 미치는 영향요인을 다중 회귀분석을 활용하여 성행관계를 진단하는 상관에서 분석이 진행되었다. 셋째, 생활 만족도에 영향을 미치는 요인을 보유하려고 구축하지 않고 동일한 일반화 모형만으로는 미치는 연구가 주를 이루었으며, 보행특성을 구분하여 진행한 연구는 부족한 실정이다. 마지막으로 대부분의 연구에서는 하나의 독립변수가 보행만족도에 미치는 영향을 분석한 연구가 대부분이며, 독립변수의 상호작용 효과를 고려하여 분석한 연구는 부족한 실정이다.

이에 따른 본 연구의 차별성은 다음과 같다.첫째, 기존의 전통 적인 분석 방법에 비해 설명력이 높은 해석 가능한 기계학습을 활용하여 분석을 진행하였다. 둘째, 성행 관계에 근거하여 있는 대 부분 성행연구의 한계점을 해결하기 위해 기계학습 모형의 특별적 특성을 해석하여 설명변수와 보행만족도 간의 비선형적인 관계를 분석하였다. 셋째, 보행의 특성에 따라 도시의 물리적 환경이 미치는 영향력이 다르기 때문에 보행특성의 일부를 보행특성과 운동특 성으로 구분하여 보행만족도에 영향을 미치는 물리적 환경변수를 분석하였다. 마지막으로 통계적 결과가 보행 만족도에 미치는 영향을 분석하여 보행특성별로 보행만족도에 미치는 시각적 특성을 도출하였다.
III. 방법론

1. 분석자료


그림 1은 연구의 공간적 범위인 서울특별시와 설문조사 응답자의 거주지 구분을 나타낸다. 설문조사 응답자는 일상생활만족도의 경우 "귀하께서는 주거지역 인근에서 일상으로 전 기가 얼마나 만족스러우실까요?"라는 질문에서 1점의 경우 "매우 불만족", 7점의 경우 "매우 만족"으로 구분하여 응답하였고, 육중생활만족도의 경우에는 "귀하께서는 주거지역 인근에서 운동으로 걷기가 얼마나 만족스러우실까요?"라는 질문의 7점 최도로 응답하였다. 

본 연구의 목적에 따라 도시 물리적 환경 분석을 위한 반경이 500m 이내의 범위를 본 연구에서는 "가로환경요인으로 분류하되(이수기, 2010; 최창규 외, 2013). 따라서 설문조사 응답자의 거주지 위치로부터 반경 500m를 기준으로 한 경도의 지역에 대한 주거환경의 최도를 기반으로 설정하였다. 

Figure 1. The case study area and respondents’ address locations

그림 2. 2016년 NSV 이미지 의미론적 분할 결과 예시

Figure 2. An example of the 2016 NSV image semantic segmentation results

가로환경요인은 가로환경의 시각적 특성과 보행만족도와의 관계를 분석한 박근석의 연구를 참고하여 도시

행만족도와 밀접한 관련이 있는 것으로 도출된 몇 가지 요인 중
가로환경의 건물 비율과 녹지 비율을 구체화하며 활용하였다. 또한, 가장 많은 비율이 보행만족도에 영향을 미칠 수 있을 것으로 판단하여 분석 변수로 추가하였다. 가로환경의 요소별 비율을 산출하기 위해 서울시 가로환경 박해티 세트인 2016년 Naver Street View(NSV) 파노라마 이미지(Panoramic Image) 41,140개
을 구축하였다. NSV 이미지 구축은 설문조사 년도인 서울시의 2016년 도로망을 기준으로 50m 간격으로 지정을 생성하고, 각 지정별 파노라마 이미지를 구축하여 분석에 활용하였다.

이후, 이미지 내의 가로환경 요소를 추출하기 위해 컴퓨터 비전 기법의 하나인 랜더링 의미론적 분할 기법을 활용하였다. 의미론적 분할 기법은 건물, 녹지, 공공, 소방 등 이미지 내에 존재하는 요소를 랜더링 위주로 구분하지는 기법이다. 이러한 의미론적 분할

포함으로는 Deeplabv3+, U-Net, PSPNet, FCN, FCN+에 있으며, 본 연구에서는 2016년 구글의 개발 이후 현재까지 지속적인 발전을 통해 우수한 성능을 나타내는 Deeplabv3+ 모형을 활용하였다(Chen et al., 2018). Deeplabv3+ 모형을 활용하여 의미론적 분할을 진행한 결과는 그림 2와 같다.

해당 모형에서는 차량, 보행자, 허리, 녹지, 핸드백, 벽, 앞으로 등 19개의 항목을 분류하였지만, 최종적으로는 앞서 설명한 바와 같이 건물, 녹지, 허리의 비율을 분석에 활용하였다. 분할이 진행된 이미지를 대상으로 이미지 전체 박해 수 대비 해당 가로환경요인 이미지 박해 수로 계산하여 변수를 구체하였으며, 활용된 수식은 다음과 같다.

가로환경요인 비율 = 가로환경요인으로 분류된 박해 수 

전체 이미지 박해 수 × 100

즉, 해당 가로환경요인 비율은 분류된 박해가 존재하지 않는 경우 0, 모든 박해가 해당 가로환경요인으로 분류될 경우 최대 100의 값으로 계산되며 값이 클수록 해당하는 지점에서 시각적으로 보이는 요소의 비율이 높다는 것으로 해석할 수 있다(기호화 의, 2021).

토지이용 특성 변수는 주거시설 연면적, 상업시설 연면적, 업
<table>
<thead>
<tr>
<th>Variables</th>
<th>Description of variables</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>일상보행만족도&lt;br&gt;Daily walking satisfaction</td>
<td>Daily walking satisfaction (1-7 points scale)</td>
<td></td>
</tr>
<tr>
<td>운동보행만족도&lt;br&gt;Exercise walking satisfaction</td>
<td>Exercise walking satisfaction (1-7 points scale)</td>
<td>Survey (2016)</td>
</tr>
<tr>
<td><strong>Personal characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>일상보행시간&lt;br&gt;Daily walking time (min.)</td>
<td>Average daily walking time (minute) per day × Average number of daily walking days per week</td>
<td></td>
</tr>
<tr>
<td>운동보행시간&lt;br&gt;Exercise Walking time (min.)</td>
<td>Average exercise walking time (minute) per day × Average number of exercise walking days per week</td>
<td></td>
</tr>
<tr>
<td>성별&lt;br&gt;Gender</td>
<td>Male (0), Female (1)</td>
<td></td>
</tr>
<tr>
<td>나이&lt;br&gt;Age</td>
<td>Age of survey respondents</td>
<td>Deep Learning semantic segmentation using Naver street view images (2016)</td>
</tr>
<tr>
<td><strong>Street environment characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>건물비율&lt;br&gt;Percentage of buildings</td>
<td>Percentage (% of buildings within a 500 m radius from a survey location)</td>
<td></td>
</tr>
<tr>
<td>하늘비율&lt;br&gt;Percentage of sky</td>
<td>Percentage (% of sky within a 500 m radius from a survey location)</td>
<td></td>
</tr>
<tr>
<td>녹지비율&lt;br&gt;Percentage of green area</td>
<td>Percentage (% of green area within a 500 m radius from a survey location)</td>
<td></td>
</tr>
<tr>
<td><strong>Landuse characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>토지이용혼합도&lt;br&gt;Land use mix</td>
<td>Entropy index of residential, commercial, and office facilities within a 500 m radius from a survey location</td>
<td></td>
</tr>
<tr>
<td>주거시설 면적(k㎡)&lt;br&gt;Total floor area of residential facilities (k㎡)</td>
<td>Total floor area (k㎡) of residential facilities within a 500 m radius from a survey location</td>
<td>National spatial information portal (2017)</td>
</tr>
<tr>
<td>상업시설 면적(k㎡)&lt;br&gt;Total floor area of commercial facilities (k㎡)</td>
<td>Total floor area (k㎡) of commercial facilities within a 500 m radius from a survey location</td>
<td></td>
</tr>
<tr>
<td>오피스시설 면적(k㎡)&lt;br&gt;Total floor area of office facilities (k㎡)</td>
<td>Total floor area (k㎡) of office facilities within a 500 m radius from a survey location</td>
<td></td>
</tr>
<tr>
<td>교지로 개수&lt;br&gt;Number of intersections</td>
<td>Number of intersections within a 500 m radius from a survey location</td>
<td>National transportation DB (2016)</td>
</tr>
<tr>
<td>건물면적 다양성&lt;br&gt;Diversity of building age</td>
<td>Standard deviation of building age within a 500 m radius from a survey location</td>
<td>National spatial information portal (2017)</td>
</tr>
<tr>
<td>평균 건물 연령&lt;br&gt;Average building age</td>
<td>Average building age within a 500 m radius from a survey location</td>
<td></td>
</tr>
<tr>
<td><strong>Accessibility characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>버스정류장 접근성&lt;br&gt;Accessibility to bus stop</td>
<td>Distance (m) from the survey location to the nearest bus stop</td>
<td>Seoul open data (2018)</td>
</tr>
<tr>
<td>지하철역 접근성&lt;br&gt;Accessibility to subway station</td>
<td>Distance (m) from the survey location to the nearest subway station</td>
<td>New address DB (2016)</td>
</tr>
<tr>
<td><strong>Border vacuum characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>대규모 건물 면적(k㎡)&lt;br&gt;Total floor area of large buildings (k㎡)</td>
<td>Total floor area (k㎡) of large buildings (top 15% large buildings in Seoul) within a 500 m radius of a survey location</td>
<td>National spatial information portal (2017)</td>
</tr>
<tr>
<td>대규모 공원 유무&lt;br&gt;Existence of large parks</td>
<td>Existence of large parks (top 15% of park size in Seoul) within a 500 m radius of a survey location</td>
<td></td>
</tr>
<tr>
<td>하천(강) 유무&lt;br&gt;Existence of rivers</td>
<td>Existence of rivers within a 500 m radius of survey location</td>
<td>National spatial information portal (2016)</td>
</tr>
<tr>
<td>고속도로 유무&lt;br&gt;Existence of highways</td>
<td>Existence of highways within a 500 m radius of survey location</td>
<td>New address DB (2016)</td>
</tr>
</tbody>
</table>
무시설 연면적을 포함하였다. 이와 더불어 주거시설, 상업시설, 업무시설의 혼입도 또한, 보행만족도에 영향을 미칠 수 있기 때문에 세 종류의 혼입도를 엘리트모에서 지수로 계산하여 토지이용 혼합도 변수를 구축하였다. 추가적으로 Jacobss가 도시 환경에 영향을 미치는 토지이용 요구로 계산한 블록 크기가 보행만족도의 영향을 미칠 수 있기 때문에 블록 크기를 나타낼 수 있는 교차로의 개수를 변수로 구축하여 추가하였다. 마찬가지로 다양한 연령의 건물 또한, Jane Jacobs 정의 도시 환경에 영향을 주는 밀도적인 요소 중 하나였기 때문에 건물 연령의 표준편차와 평균값을 계산하여 변수에 추가하였다.

다음으로 일상 및 운동보행만족도에 있어 대중교통과의 접근성도 영향을 미칠 수 있다. 이수기 (2014)의 연구는 설문조작가의 거리 지점으로부터 거리로 위치한 다양한 환경공간은 보행만족도에 영향을 미칠 수 있음을 시사하였다. 따라서 대표적인 대중교통통로 비속정량화와 지하철 접근성 변수를 분석에 포함하였다. 이때 접근성은 설문조사자의 거리 지점으로부터 가장 가까운 대중교통수단의 거리(m)를 측정하였다. 대표적인 대중교통수단비속정량화, 지하철 변수를 통해 도시 공간 내 주요교통시설과의 접근성을 분석하는 데 목적이 있다.

마지막으로 경계 공백 지대(border vacuums) 특성을 설명하기로 해서 사용된 4가지 변수는 다음과 같은 이유로 선정하였다. 우선 대규모 건물(건물 면적 상위 15%) 유무와 대규모 공원(공원 면적 상위 15%) 유무 변수는 Jacobss가 지정한 도시 내 대규모 단업용 도시이용을 측정하는 데 목적이 있다. Jacobss에 따르면 대규모 단업용 건물, 대규모 공원은 해당 지역뿐만 아니라 인접한 공간들의 보행 특성을 단일로 확장하기 어려움이 높기 때문에 도시에서 이 같은 경계 공백 지대를 지양해야 함을 주장하였다.

경계 공백 지대가 소재매매에 미치는 영향에 대해 분석한 이동업. 성현권 (2020)의 연구를 참고하여 서울시 전체 건물들의 연면적 중 상위 15% 건물들이 실문조사 지점으로부터 500m 반경에 있는지, 서울시 전체 공원 중 상위 15% 공원들이 설문조사 지점으로부터 500m 반경에 있는지 여부를 미리 변수로 구축하여 활용하였다.

또한, 경계 공백 지대 특성을 분석하는 데 있어 하천 또는 강은 보행자의 통행에 어려운 반면 4가지 이상의 도로, 고속도로, 지상철도와 함께 보행을 단절시키는 요인으로 작용할 수 있다(이동업. 성현권, 2020). 따라서 설문조사 지점 반경 500m 이내 하천이 존재하지만의 여부를 미리 변수로 구축하여 분석에 활용하였다. 이와 마찬가지로 고속도로 접근성의 경우에도 보행의 단절시키는 요인으로 작용할 수 있다. 2016년 도로망을 기준으로 고속도로, 도시 고속도로, 고속도로 연결 랜프 도로를 주요하여 설문조사 지점으로부터 반경 500m 내 유무를 미리 변수로 구축하여 활용하였다.

2. 분석 방법론


본 연구에서는 기계러닝의 여러 모형 중 보행만족도를 가장 잘 예측할 수 있는 모형을 선정하기 위해 현재 가장 많이 사용되고 있는 몇 가지 모형을 선정해 각각의 성능을 토니 회귀모형과 비교하였다. 선정된 모형은 Random Forest, Light Gradient Boost, CAT-Boost, Extreme Gradient Boost (XGBoost), Gradient Boost 모형이다. 선정된 모형들을 평가하기 위한 방법으로서 선형연구를 참고하여 주로 사용하고 있는 평가지표인 결정계수(R²), 평균제곱오차(Mean Squared Error), 평균절대오차(Mean Absolute Error)의 3가지를 활용하여 모형을 평가하였다(Chen et al., 2020; 조일외, 2021).

이후, 각 변수의 중요성을 확인하기 위해 SHAP 방법론을 활용하여 Shapley Value(기여도)를 도출하였다. 변수의 기여도는 Global Shapley Value와 Local Shapley Value로 나누어진다. 먼저, Global Shapley Value의 경우 변수의 점값으로 크기가 큰 순서대로 차례로 배치되어 있다. 각 변수별로 Shapley 값이 양수인 경우에는 붉은색, 음수인 경우에는 파란색으로 표시된다.

Local Shapley Value의 경우에는 Global Shapley Value를 보완하여 모형의 예측에 대해 변수의 영향 정도를 방향 및 크기로써 표현한다(Lundberg et al., 2018). 변수의 순서는 Global Shapley Value와 마찬가지로 모형에 영향을 미치는 크기를 순서대로 배치되며, 변수의 값이 높을수록 붉은색, 변수의 값이 낮을수록 파란색의 점으로 나타난다. 이 후 X축의 0값 기준으로 오른쪽은 양(+)의 자계, 왼쪽은 음(-)의 자계 방향으로 모형의 성능 평가에 기여한다.

하지만 SHAP 방법론으로는 물리적 환경변수가 가지는 복잡한 선형관계는 구체적으로 설명할 수 없는 한계가 존재한다. 따라서 PDP를 활용하여 각각의 독립변수가 보행만족도와 어떠한 연관성을 가지고 있는지 분석하였다. 따라서 기계학습모형을 기반으로 변수의 PDP 그래프를 생성하고 해당 독립변수와 보행만족도 사이의 관계를 분석하였다.

Journal of Korea Planning Association Vol.57, No.1 (2022) 31
IV. 분석결과

1. 기술 통계 분석

본 연구에서 사용된 변수들의 기술 통계량은 (표 2)와 같다. 종속변수는 일상보행만족도, 운동보행만족도이며 독립변수는 개인 특성, 거주환경 특성, 토지이용 특성, 접근성 특성이 해당한다. 기초통계량은 각 변수에 대한 평균, 표준편차, 최소값, 최대값으로 구성되어 있다. 종속변수는 일상보행만족도와 운동보행만족도이며 설문조사 항목의 1점부터 5점까지의 척도를 의미한다. 일상보행만족도의 평균은 4.71점, 운동보행만족도는 4.78점으로 두 변수 간 평균점수는 차이가 크지 않은 것으로 나타났다.

개인 특성 중 일상보행시간과 운동 시 보행시간은 보행목적별로 일주일 동안 보행한 시간(분)을 의미하며 일상생활 시 보행에 소요하는 평균 시간은 약 250분으로 운동 시 평균적으로 소요하는 평균 시간은 약 181분보다 높은 것으로 나타났다. 이는 운동시 보행하는 시간보다 일상생활에서 보행하는 시간이 평균적으로 높다는 것으로 해석된다. 거주환경 특성 중 거주 밀도는 비율이 평균적으로 가장 높게 나타났으며 논지 비율이 가장 낮은 것으로 나타났다. 이는 본 연구의 공간적 범위가 메트로시티 서울이라다는 점이 영향된 것으로 보인다.

토지이용 특성 중 토지이용 혼합도는 주거, 상업, 업무지역의 토지이용 혼합도를 측정한 것으로 기술통계 분석표에 따르면 토지이용 혼합도의 평균값은 약 0.40로 나타났으며 주거지역 업무지역의 평균값의 상업과 업무지역의 영역과는 높은 것으로 나타났다. 접근성 특성은 설문조사 자료로부터 가장 근접한 교통시

<table>
<thead>
<tr>
<th>Variables</th>
<th>Obs.</th>
<th>Mean, %</th>
<th>S.D.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>종속변수</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>일상보행만족도</td>
<td>2,501</td>
<td>4.71</td>
<td>1.17</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>운동보행만족도</td>
<td>2,501</td>
<td>4.78</td>
<td>1.22</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>개인특성</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>일상보행시간</td>
<td>2,501</td>
<td>260.54</td>
<td>290.69</td>
<td>0</td>
<td>4,200</td>
</tr>
<tr>
<td>운동보행시간</td>
<td>2,501</td>
<td>180.99</td>
<td>288.93</td>
<td>105</td>
<td>6,000</td>
</tr>
<tr>
<td>성별 Gender</td>
<td>2,501</td>
<td>Female (1)</td>
<td>59.78%</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>나이 Age</td>
<td>2,501</td>
<td>44.33</td>
<td>10.20</td>
<td>20</td>
<td>64</td>
</tr>
<tr>
<td>거주환경 특성</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>건물 비율 Percentage of buildings</td>
<td>2,501</td>
<td>27.66</td>
<td>17.55</td>
<td>0.00</td>
<td>73.22</td>
</tr>
<tr>
<td>하늘 비율 Percentage of sky</td>
<td>2,501</td>
<td>9.61</td>
<td>8.29</td>
<td>0.00</td>
<td>45.07</td>
</tr>
<tr>
<td>녹지 비율 Percentage of green area</td>
<td>2,501</td>
<td>7.62</td>
<td>6.03</td>
<td>0.00</td>
<td>35.61</td>
</tr>
<tr>
<td>토지이용 특성</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>토지이용 혼합도 Land use mix</td>
<td>2,501</td>
<td>0.49</td>
<td>0.23</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>주거시설 면적(km²) Total floor area of residential facilities(km²)</td>
<td>2,501</td>
<td>0.67</td>
<td>0.29</td>
<td>0.00</td>
<td>1.80</td>
</tr>
<tr>
<td>상업시설 면적(km²) Total floor area of commercial facilities(km²)</td>
<td>2,501</td>
<td>0.12</td>
<td>0.10</td>
<td>0.00</td>
<td>1.10</td>
</tr>
<tr>
<td>오피스시설 면적(km²) Total floor area of office facilities(km²)</td>
<td>2,501</td>
<td>0.07</td>
<td>0.15</td>
<td>0.00</td>
<td>1.37</td>
</tr>
<tr>
<td>교차로 가수 Number of intersections</td>
<td>2,501</td>
<td>429.37</td>
<td>294.62</td>
<td>2</td>
<td>1,504</td>
</tr>
<tr>
<td>건물연령 다양성 Diversity of building age</td>
<td>2,501</td>
<td>11.31</td>
<td>3.18</td>
<td>1.03</td>
<td>23.14</td>
</tr>
<tr>
<td>평균 건물 연령 Average building age</td>
<td>2,501</td>
<td>25.68</td>
<td>4.91</td>
<td>2.20</td>
<td>43.32</td>
</tr>
</tbody>
</table>

(Continue on next page)
2. 기계학습 모형

기계학습 모형 간의 설명력을 객관적으로 평가하기 위해 수집된 데이터 2,510개 중 이어가치에 해당하는 일부인 평균 보행시간이 6,000분을 초과하는 데이터 6개와 가로 간격 비율이 약 94.5%인 데터 3개를 제거한 2,501개의 데이터를 학습 데이터와 평가 데이터로 분할하였고, 이 때, 학습 표본은 전체의 80%인 2,001개, 평가 데이터는 20%인 500개로 할당하였다. 이후, 과적합을 방지하기 위해 각각 모형의 학습에 영향을 미치는 주요 예측변수를 조정하였다. 최종적으로 학습 데이터를 통해 각 모형을 추정하도록 하였으며, 추정된 모형들에 대한 평가 데이터의 유효도를 판단한 후 종속변수인 보행만족도를 예측하도록 하였다. <표 3>은 앞서 언급한 5가지 기계학습 모형과 결정트리 모형을 평가한 결과이다.

<표 3>의 결과에 따르면 기계학습 모형의 평가 결과는 일반적인 선형회귀 모형에 비해 좋은 성능을 보여주고 있는 것으로 나타났다. 또한, 5가지 기계학습 모형 중 Gradient Boost 모형은 일반 회귀 모형의 예측력인 10%, 15%에 비해 43%, 52%로 모형 설명력이 향상된 것을 알 수 있다. 따라서 Gradient Boost 모형이 도시 물리적 환경으로 도시민의 보행만족도를 예측할 때 가장 적절한 것으로 나타났다. Gradient Boost 모형은 정확 하방 법을 활용한 모형에서 손실함수를 최소화하는 역할을 한다. 이러한 손실함수를 파라미터로 미분하여 기울기를 산출하고, 이 기울기 값이 작아지는 방향으로 파라미터를 움직이게 하는 방법이다. 즉, Gradient Boost 모형은 단계에서 학습된 모형

<table>
<thead>
<tr>
<th>모형</th>
<th>일상보행만족도</th>
<th>운동보행만족도</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily walking satisfaction</td>
<td>Exercise walking satisfaction</td>
</tr>
<tr>
<td>R²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear Regression</td>
<td>0.10</td>
<td>0.15</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.23</td>
<td>0.28</td>
</tr>
<tr>
<td>Light Gradient Boost</td>
<td>0.28</td>
<td>0.30</td>
</tr>
<tr>
<td>CATBoost</td>
<td>0.38</td>
<td>0.44</td>
</tr>
<tr>
<td>XGBoost</td>
<td>0.40</td>
<td>0.45</td>
</tr>
<tr>
<td>Gradient Boost</td>
<td>0.43</td>
<td>0.52</td>
</tr>
</tbody>
</table>
의 장차를 다음 모형에서 학습하여 보완을 하는 방식의 모형이라(Friedman, 2001) 파라미터 조정에서는 Gradient Boost 모형이 가장 적절한 모형이라고 판단하고 최종 분석에 활용하였다. Gradient Boost 모형의 주요 매개변수로는 모형의 최대 깊이를 설정하는 max_depth, weak learner의 개수를 조정하는 n_estimators, 모형이 학습을 진행할 때마다 학습률 조정을 위해 적용하는 learning_rate, weak learner에 따른 학습에 사용하는 데이터 샘플링의 비율을 조정하는 subsample 등이 있다. 본 연구에서 최적화된 매개변수의 값을 찾기 위해 Grid Search의 방식을 활용하였다. 조정한 매개변수의 값은 일상 및 운동 보행만족도 모두 max_depth = 15, n_estimators = 1000, learning_rate = 0.01, subsample = 0.6으로 설정하였다.

3. 변수의 기여도 분석

그림 3과 4는 본 연구에서 활용한 Gradient Boost 모형에 대한 Global Shapley Value와 Local Shapley Value에 대한 결과이다. Global Shapley Value는 높은 값으로 표현된 변수일 경우에는 보행만족도와 양(+)의 관계를 가지고 있음을 나타내며, 반대로 높은 값으로 표현된 변수일 경우에는 보행만족도와 음(-)의 관계를 가지고 있음을 나타낸다. 그림 3, 4는 본 연구에서 활용한 기여도 모형인 Gradient Boost 모형을 기반으로 모형을 학습시키고, 플롯을 구성하는 점들은 각각 하나의 설문조사 응답 데이터 및 독립변수 데이터, 즉 입력 데이터 샘플 값이며, 여러 개의 점들이 모여 오른쪽이나 왼쪽으로 만드는 긴 포리 모양은 특정

![그림 3. 일상보행만족도에 대해 'SHAP'로 도출된 각 변수의 중요도](image-url)

**Figure 3. Importance of each variable derived by 'SHAP' on the daily walking satisfaction**

![그림 4. 운동보행만족도에 대해 'SHAP'로 도출된 각 변수의 중요도](image-url)

**Figure 4. Importance of each variable derived by 'SHAP' on exercise walking satisfaction**
개인에게 극단적인 측정값이 중요할 수 있음을 뜻한다(Lundberg et al., 2018; 이지윤, 2021).

본 석 결과 일상생활행태모드의 운동생활행태도 두 경우 모두에서 브로젝트 및 노동의 물리적 환경에 따라 모형의 예측력에 큰 영향을 미치고 있음을 알 수 있었다. 이후, 물리적 환경의 경우에는 일상생활행태모드의 경우 교차로 개수, 건물 연면 면적, 토지이용 허용가능 인가 수준 안전 환경을 미치고 있었으며, 운동생활행태모드의 경우에는 교차로 개수, 경로 접근성, 건물 연면 면적 등과의 순으로 영향을 미치고 있었다. 특히 교차로 개수의 경우에는 Local Shapley Value 값을 보았다. 높은 값을 가진다는 점은 원의 중심과 음(+)의 반향으로 등 계 나타났다. 이는 교차로 개수가 많을수록 즉, 축소의 크기가 작을수록 보행행태도에 매우 부정적인 영향을 미친다는 것을 알 수 있다. 이는 가로광활 면에서 분석하여 Jacobs(1961)가 언급했던 작 은 크기의 블록화는 단면의 결과이며, 토익의 크기가 작을수록 교차로의 횡단 등으로 인한 보행의 연속성을 저하시키기 때문이라고 판단된다. 또한, 건물 연면 해체율과 건물 연면 면적도 일상생활행태모드 운동생활행태도 두 경우 모두에서 음(-)의 관계가 나타났다. 건물 연면의 영향이 낮음수록 보행행태도가 높아지며, 건물 연면의 영향력이 보행행 태도에 미치는 정도가 크게 나타났다.

토지이용 허용도를 살펴보면 Global Shapley Value의 경우에 일상생활행태모드와 운동생활행태도 두 가지 경우 모두에서 음 (-)의 관계가 있는 것으로 도출되었다. 그러나 Local Shapley Value의 경우에는 높은 징과 푸른색 점이 혼합된 정도가 높은 것으로 보아 비선형 관계가 있음을 시사한다. 이와 관련하여 상업시설 연면적과 업무시설 연면적의 경우에도 운동생활행태도 일상생활행태도 모두 음(-)의 관계로 나타났다. 특히, 상업시설 연면적의 경우에도 Local Shapley Value의 높은 점과 둥근색 점이 혼합되어 나타나는 것으로 보아 비선형 관계가 있음을 알 수 있다.

경제 공백 지대 번수의 경우에는 대부분 번수가 보행행태도에 미치는 영향이 적은 것으로 나타난다. 먼저 일상생활행태도의 경우 대규모 건물 면적, 고도수 유무, 하천 유무, 대규모 공원 유무의 순으로 변수의 영향력이 큰 것으로 도출되었다. 대규모 공원 유무로 제한된 대규모 건물 면적, 고도수 유무, 하천 유무 등의 경우에는 일상생활행태도의 음(+)의 관계를 가지는 것으로 나타났다. 이는 고도수로, 대규모 공원, 하천이 있을 경우 보행행태도가 높다는 것을 뜻한다. 이러한 결과는, 가로광활 면에서 Jacob의 주요성에 상응하는 부분이 있으면, 그래프에서 높은 점과 둥근색 점과의 혼합도가 높은 것을 보아 비선형 관계를 가지고 있는 것으로 판단된다.

운동생활행태도의 경우에는 대규모 건물 면적, 하천 유무, 고도수 유무, 대규모 공원 유무 순으로 변수의 영향력이 높은 것으로 도출되었다. 하천 유무와 고도수 유무의 경우에는 양 (+)의 관계를 가지는 것으로 나타났으며, 대규모 건물 면적이 짧은 경우 공원 유무의 경우에는 음(-)의 관계를 가지는 것으로 나타났다. 하천과 고도수로도 전부 양 경우 보행행태도가 높아지며, 대규모 건물 면적이 짧은 경우 모르 보행행태도가 높아지게 의미한다. 대규모 건물 면적도의 경우 운동생활행태도에서 변수가 아이디어 상을 차지하고 있어 우선적인 제어가 필요할 것으로 판단된다. 하지만, 미나리 경계 공백 지대 번수의 경우에도 일상생활행태도와 마찬가지로 변수의 중요도가 높아 향후 추가적인 연구가 필요할 것으로 예상된다.

4. PDP을 활용한 비선형 관계 분석

(그림 5)는 Gradient Boost 모형을 기반으로 변수의 Partial Dependence Plot(PDP) 그래프로 생성한 결과이다. X축은 해당 독립변수의 값이며, Y축은 보행행태도를 나타낸다. 전체적으로 대부분의 물리적 환경변수와 보행행태도 사이는 비선형적 인 연관성이 있는 것으로 도출되었다. 먼저, 교차로 개수와 토지이용 허용도의 경우에는 일상생활행태모드 운동생활행태모드에서 비선형한 형태의 결과가 도출되었다. PDP 그래프는 앞서 분석한 Local Shapley Value에서 높은 점과 푸른 점의 혼합도가 높아 비선형적 관계가 있을 것으로 판단되는 변수를 중심으로 결과를 도출하였다.

앞선 Shapley Value에서 언급했던 것처럼 교차로의 개수와 많은수록 보행행태도는 낮아지는 것으로 나타났다. 토지이용 허용도의 경우에는 일정한 수준을 유지하기가 약 0.4~0.6 정도로 되어 최적값을 나타내고 전적 보행행태도 감소하는 비선형 관계가 나타났다. 이러한 결과는 일상생활행태도 운동생활행태도에서 유사하게 나타났다. 이는 적당한 수준의 토지이용 허용도 상이 및 운동생활행태도 중에 균일한 영향을 미치고 있음을 시사한다. 건물 연면 면적의 경우에는 다양성이 높음을 수록 보행행태도는 감소하는 결과가 도출되어 노후 건물의 재정비 및 외관개선은 도시민의 보행행태도에 긍정적인 효과가 있을 것으로 판단된다.

대규모 건물 면적도의 경우에는 일상생활행태모드에 유의한 영향을 미치지 않았지만, 운동생활행태모드의 경우에는 500m2 이상인 대규모 건물 면적이 0.6km2 하일 때 보행행태도가 가장 높게 나타났다. 주거시설 면적도의 경우 보행행태에서 차이가 있는 것으로 나타났다. 주거시설 면적도의 경우 0.6km2 이상일 때 일상생활행태도는 높게 나타났지만 운동생활행태도는 구준히 감소하는 형태로 나타났다. 이는 동일 목적의 보행은 주거 시설의 연면적 전체 수록 보행행태도가 높다는 것을 의미한다. 상업시설 면적도의 경우에는 일상생활행태도 운동생활행태도 사이에 반대의 결과가 도출되었다. 먼저 일상생활행태도의 경우
그래프 5. 보행만족도 Partial Dependence Plot (PDP) 분석결과

Figure 5. Partial Dependence Plot (PDP) analysis results of walking satisfaction

에는 상업시설 연면적이 높아짐에 따라 보행만족도가 높아지는 형태를 나타내며 약 0.2km²을 기점으로 서서히 낮아지는 결과가 도출되었다. 반면, 운동생활만족도의 경우에는 상업시설 연면적이 높아짐에 따라 보행만족도는 점점 감소하는 형태를 보여 운동 시에는 상업시설의 연면적이 보행만족도에 부정적인 영향을 미친다는 점을 시사한다.

가로 간을 비율의 경우에는 일상생활만족도의 경우 일정한 수준을 보이나 약 30%를 넘어서면서부터 보행만족도가 급격하게 증가하는 형태를 보이고 있다. 하지만 운동생활만족도의 경우에는 간을의 비율이 약 30% 정도일 때 보행만족도가 가장 높다는
결과가 나타났으며, 적당한 정도의 건물 비율이 운동 시 보행만족도에 긍정적인 영향을 미칠 수 있음을 시사한다. 마찬가지로, 논지 분만의 경우에는 일상생활방면의 경우 어느 정도 일정한 수준을 유지하다가 논지 비율이 약 15%정도 보행만족도가 가장 높을 수 있다. 하지만 운동보행만족도의 경우에는 논지 비율이 높을수록 보행만족도는 높아지는 경향을 보였다. 논지 비율이 약 7.5%를 넘어서부터는 보행만족도가 급격히 감소함을 알 수 있다. 따라서 이는 일상생활의 경우는 논지 비율이 높을수록 보행만족도는 높아지는 경향을 보였다. 논지 비율이 약 7.5%를 넘어서부터는 보행만족도가 급격히 감소하는 놀라운 사실이다. 

근데(表2)의 연구와 논리의 비율은 보행의 특성에 따라 보행만족도와 비선형적인 관계가 있음을 시사한다.

5. 상호작용 효과 분석

다음 〈그림 6〉은 독립변수의 상호작용 효과에 대한 보행만족도와 관계를 그래프로 나타낸 결과이다. 상호작용 효과는 토지이용 혼합도와 상업시설 연면적, 대규모 건물 연면적과 주거시설 연면적, 가로 논지를 가로 건물 비율의 세 가지 조합에 집중하여 분석하였다. 우선 토지이용 혼합도와 상업시설 연면적 조합의 경우 상업시설의 연면적에 따른 토지이용 혼합도가 일정 및 운동보행만족도에 미치는 영향을 분석하기 위해서 선정하였다. 두 번째, 대규모 건물 연면적과 주거시설 연면적 조합의 경우 주거연면적 수준에 따른 대규모 건물 연면적이 경제적 공백의 역할을 하여 일정 및 운동보행만족도에 미치는 영향을 분석하기 위해서 선정하였다. 마지막으로 가로 건물 비율과 논지 비율의 상호작용은 일상생활방면과 운동보행만족도에 서로 다른 영향을 미칠 것으로 판단되어 분석하였다.

먼저, 〈그림 6〉의 (a)와 (c)는 토지이용 혼합도와 상업시설 연면적과의 관계가 보행만족도에 미치는 영향을 나타낸 결과이다. 토지이용 혼합도의 경우는 앞선 PDP 결과와 같이 일상생활방면과 운동보행만족도 모두 약 0.4~0.6 범위의 토지이용혼합도 수준일 때 보행만족도가 높게 나타났다. 하지만 비슷한 수준의 토지이용혼합도에서도 상업시설 연면적이 높을수록 일상생활방면도 높아지는 경향이 있었으며, 이와 반대로 운동보행만족도는 상업시설 연면적이 작을수록 높아지는 경향을 나타내고 있다. 이는 일상생활에서도 상업시설 연면적이 높은 수준의 적당한 토지이용 혼합도에서 보행만족도가 높으며, 운동 시는 그 반대의 경우일 때 보행만족도가 높아진다는 토지이용 혼합도라도 보행방면에 따라 혼합도의 변화를 다르게 해주며 종합을 시사한다.

두 번째, 〈그림 6〉의 (b)와 (c)는 주거시설 연면적과 대규모 건물 연면적이 보행만족도에 미치는 영향에 대한 그래프이다. 먼저, 일상생활방면의 경우에는 대규모 건물과 주거시설 연면적의 높음수록 보행만족도는 높다는 결과가 도출되었다. 이는 대규모 건물과 주거시설 연면적 모두 주기적인 보행만족도는 높다는 결과가 도출되었다. 이는 일상생활에 있어서는 대규모 주거시설이 많이 분포한 수록 보행만족도는 높아지고, 반대로 운동 시에는 주거시설 연면적이 높은 지역은 보행만족도가 감소하는 것을 의미한다. 마찬가지로, 〈그림 6〉의 (d)와 (f)는 가로 논지 비율과 가로 건물 비율이 보행만족도를 미치는 영향에 대한 그래프이다. 먼저 일상생활방면의 경우에는 가로 논지 비율이 약 15%, 가로 건물 비율이 높을수록 보행만족도는 높다는 결과가 도출되었다. 이는 일상생활에 있어서는 수준의 논지와 보행방면의 다양성을 높일 수 있는 건물의 비율이 높을수록 보행만족도가 높을 것으로 시사한다. 운동보행만족도의 경우에는 가로 논지 비율이 약 7.5%, 가로 건물 비율이 약 30%의 적당한 비율이 더 보행만족도가 가장 높은 결과가 도출되었다. 이는 두가지로도 많은 양의 논지 또는 건물의 비율이 높을수록 건물의 특성에 따라 높이 부정적인 영향을 줄 수 있으며, 적정 수준의 비율이 보행만족도에 긍정적인 영향을 미친다는 것을 시사한다.

V. 결론

본 연구에서는 서울시 거주민 대상으로 실시한 2016년 설문 조사 자료를 바탕으로 일상생활보행만족도와 운동보행만족도 영향요인을 분석하였다. 독립변수로는 설문조사에서 대상개인 특성변수와 둘러싼 미비적으로 분할 기법을 활용한 가로환경 특성변수, 토지이용 변수, 경제적공간 지배 변수의 유형을 구분하여 구체화하였다. 그리고, 분석방법론으로는 독립적인 기계학습 방법론을 활용하여 보행만족도에 영향을 미치는 변수의 기여도를 분석하고, 변수가 보행만족도와 가지는 비선형적 관계와 상호작용 효과를 분석하였다.

분석 결과를 종합해보면, 도시의 물리적 환경요인과 보행만족도 사이에는 비선형적인 관계가 강하게 나타나는 것을 알 수 있다. 또한, 이러한 비선형적 관계는 보행방면별로 다르게 나타났으며, 물리적 환경변수 사이의 상호작용 효과도 보행만족도에 영향을 미치는 것으로 나타났다. 이러한 분석 결과는 보행방면에 따른 보행만족도 개선을 위해 가능성을 제시할 필요성을 시사한다. 구체적인 분석 결과와 정책적 시사점은 다음과 같다.

첫째, 해석 가능한 기계학습을 활용하여 전통적인 회귀분석 방법론의 모형보다 활성화 높은 설명력을 도출하였다. 일상생활방면의 경우에는 전통 회귀모형의 설명력이 10%에서 기계학습 모형의 경우 설명력이 43%로 향상되었다. 이는 유사하게 운동보행만족도의 경우 전통 회귀모형의 설명력이 15%에서 기계학습 모형의 설명력이 52%로 매우 높게 나타났다. 따라서 기계학습 방법론
의 활용은 기존의 전통적인 환경분석 방법론과 비교하여 도형의
설명력 향상이 매우 높은 것을 알 수 있다. 이러한 도형 설명력의
향상은 환경 가로의 도형적 특성 및 도형의 개발에
 크게 기여할 수 있다. 이러한 도형적 특성의 도형의 개발은 도형
계획도 기반의 도형 경로 등을 제공하여 서비스 이용자에게 적절
한 경로를 제시할 수 있다는 점이 있다(이재송·이현희, 2019).

Figure 6. Interaction effect of built environmental variables on walking satisfaction by purpose
둘째, 일상보행만족도와 운동보행만족도 두 경우 모두에서 교차로의 개수가 적을수록 보행만족도는 높아지는 결과가 도출되었다. 이러한 결과는 분석의 규모가 작아지기 때문인 보행 반도 및 다양성은 증가할 수 있었다고, 조사연구 및 통계등으로 보행 활동의 연속성을 저해시키기 때문에 보행만족도에는 부정적인 영향을 미치는 것으로 판단되었다. 따라서 도심인의 보행만족도가 높아질수록 보행 활동에 끌어들일 수 있도록 보행의 연속성을 강화할 수 있는 접근이 보행만족도 커뮤니티에 매우 유용한 것으로 볼 수 있다.

셋째, 토지이용 혼합도의 경우에는 역 U자 모양의 표준선으로 진행적인 비선형 관계를 가진 것으로 나타났다. 토지이용 혼합도 지수가 높아질수록 보행만족도가 증가하지만, 토지이용 혼합도가 약 0.4~0.6 이상일 때 보행만족도는 빠르게 감소하는 것으로 나타났다. 또한, 상호작용 효과가 추가적으로 보였을 때, 비슷한 수준의 토지이용 혼합도 중에서도 운동보행만족도의 경우에는 상대적 연변에선 작용만족도 탑승하는 양이 증가하였다. 그리고 이러한 결과는 특정 토지이용 밀도와 토지이용 혼합도가 보행 목적에 따라 보행만족도에 미치는 영향을 다스리기를 시사한다.

넷째, 건물 연면의 경우 건물 연면 평균값이 약 22년일 때 일상 보행만족도가 가장 높은 것으로 나타났다. 운동보행만족도의 경우에는 24~26년 정도 수준이 되면 보행만족도가 높은 것으로 나타났다. 건물 연면 다양성의 경우에는 일상보행만족도와 운동보행 만족도 두 가지 모두에서 낮음수록 보행만족도는 높다는 결과가 도출되었다. 이런 결과를 종합하여 볼 때 건물 연면의 평균이 20~24년 정도의 통치에 지어진 건물일수록 보행만족도가 높다는 결론이 도출되었다. 따라서 도심인의 보행만족도 증진을 위해서는 역 따른 노출 건축물에 대한 꾸준한 정비가 필요함을 시사한다.

다섯째, 운동보행만족도의 경우 가로 파노라마 이미지를 기준으로 가로의 녹지 비율은 약 7.5%, 건물 비율은 약 30% 수준일 때 보행만족도가 가장 높은 것으로 나타났다. 또한, 상호작용 분석 결과를 같이 고려할 때, 고품질 녹지 비율의 증가는 오히려 보행 만족도에 부정적인 영향을 미치는 것으로 나타났다. 이러한 결과는 보행목적에 따라 가로환경에서의 건물 및 녹지와의 조화가 이루어질 때 보행만족도가 증가할 수 있음을 시사한다.

마지막으로 경계 공공 지대 반도의 경우에는 대규모 건물 연면적을 제한한 경우, 고품질, 대규모 공원 등은 보행 만족도에 미치는 중요도가 크지 않은 것으로 나타났다. 이들 경계 공공 지대 반도는 대규모 반도에 구축되기에 때문에 생태적인 영향력을 보기에 위해서는 추가적인 연구가 필요할 것으로 판단된다. 반면 대규모 건물 연면적은 주거시설 연면적에 상호작용 효과를 보였을 때 일상보행만족도의 경우 상당히 복잡한 비선형적이 관계가 나타난 반면, 운동보행만족도에서 대규모 주거시설 연면적이 높수록 보행만족도는 감소하는 경향이 두드러져 나타났다. 따라서 대규모 아파트 단지와 같은 주거시설의 증가는 경계

공백으로 작용하여 보행만족도에 부정적인 영향을 미칠 수 있음을 시사한다.

한편, 본 연구는 다음과 같이 몇 가지의 한계점을 가진다. 우선, 본 연구에서 사용한 NSV 이미지의 경우 자동차를 기준으로 촬영된 이미지기 때문에 보행자가 인식하는 가로환경과는 차이가 있을 수 있다. 또한, 설문조사 자료에서 구축한 개인 특성변수의 경우 보행만족도에 대한 영향력이 높음에도 불구하고 이에 대한 해석이 부족하다는 한계점이 존재한다. 따라서 앞으로 최근의 도시환경을 반영하여 개인적인 특성인 보행만족도에 어떠한 영향을 미치지 살펴볼 필요가 있을 것으로 판단된다.

그럼에도 불구하고 본 연구는 해석 가능한 기계학습 방법론을 활용하여 도시인의 보행만족도에 영향을 미치는 도시 물리적 환경요인을 분석했다는 점에서 의미가 있다. 또한, 설문연구에서도 한계점으로 언급된 수도 간의 비선형 관계를 확인하였고, 보행특성별로 다른 정책적 시가지를 제공하려는 점에서 의의가 있다.

이러한 기준을 따르면 본 연구는 보행만족도 증진 목적의 도시환경에 대한 정책적 이론은 보행 환경적인 도시를 조성하고 이에 따른 사용자의 만족감을 높이는 데 중요한 자료로서 활용될 수 있을 것으로 시사한다.

인용문헌

References

1. 권칠민, 2019. 「미래에 미신이라는 백과', 2세기, 웹북스.
5. 박근석, 기동현, 이수기, 2021. "서울시 가로환경 조사의 시각적 특성 별 보행여건과 미치는 영향 분석: 구글 가로환경 미리보기 앱 이용

6. 성현근, 2014. "주거지 건조환경과 보행활동과의 비선형 연관성


8. 유정은, 2017. "기계학습을 통한 TIMSS 2011 중학생의 실생활

9. 윤준호, 이수기, 2019. "기계학습과 LIME 방법론을 활용한 서울


11. 이동섭, 성현근, 2020. "재난 제의자료의 경제적책정책과 소재의

12. 이수기, 고홍주, 이기훈, 2016. "근린환경특성에 보행여건에

13. 이수기, 이은성, 염정란, 2014. "보행자 언발매달 보행여건에


15. 이송휘, 이혜린, 2019. "기계학습 알고리즘을 이용한 보행여건

16. 이지윤, 2021. "보행여건에 영향을 미치는 가로환경의 물리적

17. 조원, 김성재, 이수기, 2021. "근린환경이 도시생활에 미치는 영

18. 조현민, 이수기, 2016. "보행여건의 보행활동시간에 영향을 미치

19. 조현민, 이수기, 2017. "근린환경특성에 사회적 자본의 수준에 미

20. 최창규, 성현근, 이수기, 김태현, 고준호, 윤정란, 2013. "지속가능

Elements on Housing Prices Based on Multisource Data—A Case Study of Shanghai, China”, ISPRS International Journal of Geo-Information, 9(2): 106.


