

# 기계학습을 이용한 스마트폰 이용이 보행속도에 미치는 영향 분석

## Effects of Smartphone Usage on Walking Speed using Machine Learning Method

진혜련\* · 도명식\*\*

\* 주저자 : 한밭대학교 SOC 자산관리센터 연구원

\*\* 교신저자 : 한밭대학교 도시공학과 교수

Hye ryun Jin\* · Myung sik Do\*\*

\* Center of Infrastructure Asset Management, Hanbat National University

\*\* Dept. of Urban Eng., Hanbat National University

† Corresponding author : Myung sik Do, msdo@hanbat.ac.kr

Vol.18 No.2(2019)

April, 2019

pp.93~103

pISSN 1738-0774

eISSN 2384-1729

<https://doi.org/10.12815/kits.2019.18.2.93>

2019.18.2.93

### 요약

본 연구에서는 대전광역시 내의 보행로 2개소를 대상으로 보행 중 스마트폰 사용이 보행속도에 미치는 영향을 분석하였다. 분석을 위해 스마트폰 사용실태에 대한 영상자료를 취득하고 도로용량편람에서 제시한 보행자 서비스수준의 보행밀도를 기준으로 보행속도를 산정하였다. 보행속도에 미치는 영향을 분석하기 위한 방법으로는 기계학습을 통한 다중회귀분석과 의사결정나무(Decision tree)를 활용하였으며, 설명변수로는 성별, 스마트폰 미사용, 청각을 이용한 스마트폰 사용, 시각을 이용한 스마트폰 사용, 서비스수준A, 서비스수준B, 서비스수준C가 선정되었다. 분석결과 서비스수준C가 보행속도 변화에 가장 높은 영향을 미치는 것으로 나타났으며, 서비스수준C에서 시각을 활용해 스마트폰을 이용한 여성그룹이 가장 낮은 속도로 보행하는 것으로 나타났다. 특히 음악 감상이나 통화와 같은 청각을 이용한 유형보다 시각을 이용하는 경우 대부분의 경우에서 보행속도가 크게 저하되며 통계적으로 유의한 차이가 있음을 확인하였다.

핵심어 : 기계학습, 의사결정나무, 다중회귀분석, 스마트폰, 보행속도

### ABSTRACT

This study analyzed the impact of smartphone usage on walking speed during walking on two pedestrian walkways in Daejeon Metropolitan City. For the analysis, the video data about the actual use of smartphone was acquired and the walking speed was calculated based on the walking density of the pedestrian Level Of Service(LOS) presented in the Road Capacity Manual. Multiple regression analysis and decision tree using machine learning were used to analyze the impact of smartphone usage on walking speed, and as the explanatory variables, gender, disable smartphone, use of smartphone using auditory function, use of smartphone using visual function, LOS A, LOS B, LOS C were adopted. The result showed that LOS C had the highest impact on walking speed change and the women's group using their visual function was founded to have the slowest walking speed in LOS C. In particular, the author found that walking speed significantly decreased in the case of use of visual function rather than listening to music or the hearing on the phone.

Key words : Machine Learning, Decision tree, Multiple regression, Smartphone, Walking speed

Received 4 January 2019

Revised 31 January 2019

Accepted 4 March 2019

© 2019. The Korea Institute of Intelligent Transport Systems. All rights reserved.

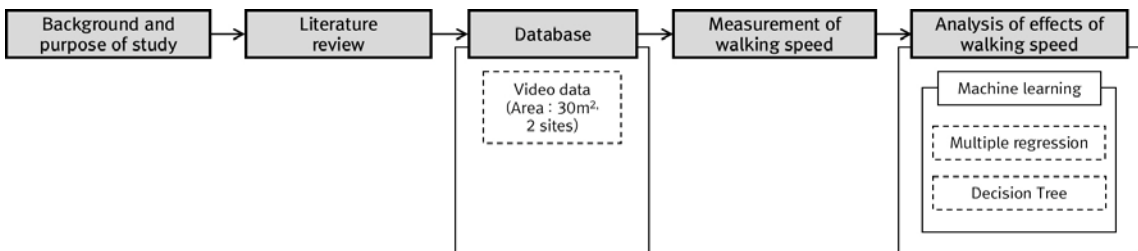
# I. 서론

## 1. 개요

최근 스마트폰의 보급이 증가하여 2016년 기준으로 국내 만 6세 이상 인구의 88.5%가 스마트기기를 보유하고 있으며, 스마트폰 가입자와 인터넷 이용자가 4천만 명을 넘어섰다. 스마트폰은 이동 중에도 사용이 편리하다는 장점을 가지고 있다. 그러나 보행 중에 스마트폰을 이용할 경우 걷는 행위와 스마트폰을 이용하는 행위 간에 서로 간섭을 일으켜 인지능력과 사고력, 판단력을 저하시키는 이중과제간섭(dual task interference)이 발생하게 되어 보행자가 두 행위에 대한 수행력이 감소하게 된다(Pellecchia, 2005; Shumway-Cook and Woollacott, 2000; Frenkel-Toledo et al., 2005). 이 경우 보행자가 보행환경으로부터 주의가 분산되어 보도에 설치된 시설물이나 다른 보행자와의 충돌이 발생할 가능성을 가지고 있으며, 특히 보행자와 차량이 함께 통행하는 횡단보도에서는 보행자가 차량과 충돌할 가능성이 높다.

2014년부터 2016년까지 최근 3년간 차대사람 사고 중 보행 중 주의분산 행동과 관련된 사고 중 휴대전화 사용에 의한 교통사고는 주의분산행동사고의 61.7%인 1,085건이 발생하였으며, 1,105명의 사상자수가 발생한 것으로 나타났다(Park, 2017). 이와 같이 보행 중 스마트폰 사용으로 인한 교통사고는 국내뿐만 아니라 전 세계적으로 큰 문제가 되고 있는 실정이다.

따라서 본 연구에서는 보행 중 스마트폰 사용이 보행속도에 미치는 영향력을 분석하기 위해 예측 성능이 우수하고 효율적으로 모델을 구축할 수 있는 기계학습법을 활용하였다. 이를 위해 현장조사를 실시하여 습득한 보행자 통행현황에 대한 영상자료를 스마트폰 사용방법으로 구분하여 보행속도를 측정하였다. 데이터는 다중회귀분석을 실시하여 스마트폰 사용에 관련된 변수들의 보행속도에 대한 주 효과를 선별하였으며, 기계학습의 의사결정나무(Decision tree) 방법을 통해 보행속도에 영향을 미치는 변수들의 중요도와 복합적 영향을 분석하는 것을 목적으로 한다.



<Fig. 1> Flow-chart for Study

## 2. 선행연구 고찰

보행 중 스마트폰 이용에 따른 위험성 관련 선행연구를 살펴보면 다음과 같다. 국내의 경우, Lee and Yang(2015)은 보행 중 스마트폰 이용이 시각과 청각 자극에 영향을 미쳐 보행속도를 감소시킴을 확인하였으며, Park and Park(2016)은 보행 중 스마트폰을 이용했을 때가 그렇지 않았을 때보다 보행속도(gait velocity), 보폭(step length), 보장(stride length)이 감소한다고 분석하였다. 또한 Kang et al.(2016)은 보행 중 스마트기기 이용에 대한 사용실태와 행동패턴에 대한 분석을 통해 스마트 기기 이용이 보행자에게 주위 환경에 대한 주의를 분산시켜 교통사고로 이어질 수 있음을 밝혔으며, You and Kwon(2017)은 20~30대 성인을 대상으로 스

마트폰 이용 동기 및 이용 실태가 보행안전도에 미치는 영향은 27.6%에 달하며, 스마트폰을 이용한 인터넷 검색이 보행안전도에 가장 큰 영향을 끼치는 것으로 분석하였다.

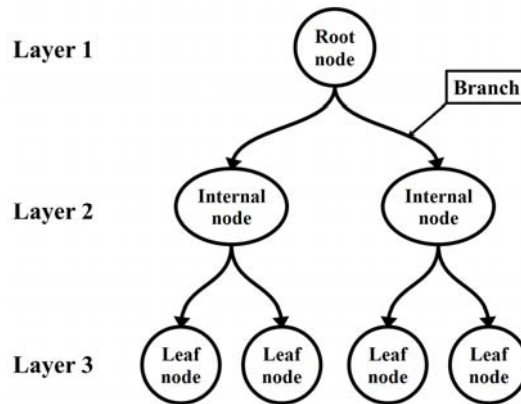
한편, 해외의 경우, Thompson et al.(2012)는 통화, 핸드프리통화, 문자발송이 일반 보행자에 비해 각각 0.75초, 1.29초, 1.87초 정도 보행시간이 더 소요되는 것을 밝혔다. Hyman et al.(2010)는 보행자 통행량이 많은 곳을 대상으로 일정 구간에 대해 보행 중 휴대전화나 음향기기 사용 등에 의한 주의분산행동을 관찰한 결과 일반 보행자의 통행시간은 74.8초로, 음향기기 이용 보행자는 73.7초, 통화하는 보행자는 82.5초, 대화하는 보행자는 86.2초로 나타나 일반 보행자에 비해 보행시간이 더 소요됨을 밝혔다. 이는 대화의 경우 다른 경우에 비해 방향 전환이 잦고 비틀거리는 경우가 많기 때문인 것으로 분석하였다. 또한 Schwebel et al.(2012)는 보행 중 스마트폰 이용자가 그렇지 않은 사람에 비해 정신이 분산되어 주행 차량에 대한 위험도가 높음을 밝혔다.

선행연구를 살펴본 결과 스마트폰 이용실태나 스마트폰 이용에 따른 위험성에 대한 의식조사, 신체반응조사, 보행속도 변화에 대한 연구가 주를 이루고 있으며 그 위험성을 서비스수준(Level Of Service; LOS)과 연계하여 정량화한 연구는 거의 없는 것으로 나타났다. 따라서 본 연구에서는 현장조사를 통해 스마트폰 이용이 보행속도에 미치는 영향을 기계학습을 이용하여 정량적으로 살펴보고 이를 서비스수준과 연계해 정책적 의미를 살펴보고자 하였다.

## II. 분석방법론

### 1. 기계학습과 의사결정나무

의사결정나무는 기계학습 알고리즘 중의 하나로, 여러 의사결정규칙을 도표화하여 관심대상이 되는 집단을 몇 개의 소집단으로 분류(classification)하거나 예측(prediction)을 수행하는 방법이다. 의사결정나무는 데이터의 분리기준(splitting criterion)과 데이터의 분리가 더 이상 일어나지 않도록 하는 정지규칙(stopping rule) 등에 따라 CHAID(Chi-squared Automatic Interaction Detection), CART(Classification And Regression Trees), C5.0(successor of ID4), C4.5(successor of ID3), ID3(Iterative Dichotomiser 3) 등으로 구분되며, 모델의 기본형태는 <Fig. 2>와 같다(Rokach and Maimon, 2015).



<Fig. 2> Decision tree structure

이러한 의사결정나무는 나무구조의 형태로 표현되어 있어 분석자가 모형을 이해하기 쉽다. 또한 선형성(linearity)이나 정규성(normality) 또는 등분산성(homoscedasticity) 등의 가정을 필요로 하지 않는 비모수적 방법으로, 순위만 분석에 영향을 미치지 때문에 이상치(outlier)에 민감하지 않다는 장점이 있어 기존 통계모델보다 결과 예측에 유리하다.

따라서 본 연구에서 종속변수는 연속형임에 따라 회귀나무를 형성하는 CART 알고리즘을 활용하였다. CART 알고리즘은 이진분리(binary split)를 수행하며 서로 다른 분리기준을 가지고 있으며, 종속변수가 이산형인 분리나무(classification tree)와 연속형인 회귀나무(regression tree)를 형성한다. 그 중 회귀나무는 식 (1)과 같은 분산을 분리기준으로 이용하여 나무를 구성하게 된다.

$$V = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \tag{1}$$

여기서  $V$ 는 분산,  $n$ 은 해당 노드에 포함된 목표변수의 개수,  $y_i$ 는 해당 노드의  $i$ 번째 목표변수,  $\bar{y}$ 는 노드의 목표변수의 평균을 나타낸다. 이를 노드의 목표변수의 평균을 해당 노드에 포함되는 모든 개체에 대한 예측값으로 사용된다면 예측오차를 최소화 하는 것과 같은 맥락이라 할 수 있어, 회귀나무는 식 (2)와 같은 분산의 감소를 분리기준으로 이용하여 나무를 형성하게 된다. 여기서  $n$ 은 분리된 노드의 상위노드의 데이터 개수,  $n_L$ 과  $n_R$ 은 분리된 노드의 데이터 수,  $V_L$ 과  $V_R$ 은 분리된 노드의 분산을 말하며,  $L$ 과  $R$ 은 분리된 노드의 왼쪽 노드와 오른쪽 노드를 나타낸다.(Choi et al., 1998).

$$\Delta V = V - \frac{n_L}{n} V_L - \frac{n_R}{n} V_R \tag{2}$$

본 연구에는 Python 프로그래밍 언어에 대한 라이브러리 중 Scikit-Learn을 활용하였다. 오픈 소스인 Python은 범용 프로그래밍 언어에 대한 장점에 더해 Matlab이나 R과 같은 일정 분야에 특화된 스크립팅 언어의 편리함을 갖춰 데이터의 시각화, 통계분석, 언어처리 등에 필요한 라이브러리를 포함하고 있다. Python의 여러 라이브러리 중 Scikit-learn은 알고리즘 사용에 있어 사용자 친화적으로 효율적이고 생산적인 인터페이스를 제공한다는 장점을 가지고 있다. 이러한 Scikit-learn은 Python의 배포판인 Anaconda, Enthought Canopy, Python(x,y) 등을 통해 활용되며, 특히 Anaconda의 경우 대용량 데이터의 처리 분석, 예측 분석에 유용하다(Müller and Guido, 2016; Raschka, 2016). 따라서 본 연구에서는 Python의 여러 배포판 중 Anaconda를 활용하여 회귀나무를 형성하는 의사결정나무 분석을 수행하였다.

## 2. 다중회귀분석

다중회귀분석은 하나의 종속변수와 여러 개의 독립변수 간의 선형관계가 성립함을 가정하고, 주어진 학습 데이터를 통해 각 독립변수에 대한 영향력인 회귀계수를 추정하는 모형이다. 이러한 다중회귀분석의 경우 의사결정나무와 달리 선형성이나 정규성 등의 가정을 필요로 하는 모수적 방법으로, 변수들의 주효과를 분석할 수 있다(Lee and Noh, 2012). 분석에는 기존의 전통적 방법보다 설명력의 예측이 우수하고 효율적인 모델 구축을 할 수 있는 기계학습 알고리즘을 활용하였다. 알고리즘은 Anaconda를 통해 활용하였으며 이를 통해 스마트폰이 보행속도에 미치는 영향을 정량적으로 분석하였다(Müller and Guido, 2016).

### Ⅲ. 스마트폰 사용여부에 따른 보행속도 측정

#### 1. 조사개요

본 연구에서는 대전광역시 내의 보행로 2개소를 선정하여 구간길이 10m, 보도폭 3m로 구성된 30㎡의 범위를 조사대상지로 설정하여 보행자의 스마트폰 사용여부와 보행속도를 영상촬영을 통해 조사하였다. 데이터는 취득의 용이성을 위해 유동인구가 많은 유성구 궁동에 소재한 충남대학교 인근의 보행로(A)와 서구 둔산동에 소재한 갤러리아 타임월드 백화점 인근의 보행로(B)를 선정하였다.

촬영한 영상은 MOLIT(2013)의 도로용량편람에서 제시한 보행자 서비스수준의 보행밀도를 활용하여 보행속도를 산정하였다. 보행속도는 서비스수준에 따라 성별, 스마트폰 사용유형을 구분하였다. 스마트폰 사용유형에는 1) 스마트폰을 사용하지 않은 유형과 2) 문자메시지 보내기나 동영상 시청과 같은 시각을 사용한 유형, 3) 음악 감상이나 통화와 같은 청각을 사용한 유형의 세 가지로 분류하였다.

<Table 1> LOS on Pedestrian Roads (MOLIT, 2013)

LOS	Pedestrian traffic flow rate (person/min/m)	Occupied space (m <sup>2</sup> /person)	Density (person/30m <sup>2</sup> )	Speed (m/sec)
A	≤20	≥0.11	≤9	≥1.3
B	≤32	≥0.07	≤15	≥1.2
C	≤46	≥0.05	≤21	≥1.2
D	≤70	≥0.03	≤33	≥1.0
E	≤106	≥0.01	≤78	≥0.7
F	-	<0.01	>78	<0.7

#### 2. 보행속도산정

보행속도를 산정함에 앞서, 스마트폰을 사용하지 않는 보행자 중 음식을 섭취하는 보행자는 그로 인해 보행속도가 저하될 가능성이 있어 제외하였다.

보행속도는 보행자 1인이 10m의 구간길이를 갖는 임의의 A지점과 B지점을 통과하는 시간을 영상촬영자료를 기반으로 취득한 뒤, 구간길이와 구간통과시간을 이용하여 산정하였다. 또한, 보행자의 서비스수준은 보행자의 구간통과시간동안 측정된 해당 면적 내의 총 보행인원을 기준으로 <Table 1>에 근거하여 산정하였다. 본 연구는 스마트폰의 사용여부, 사용유형 등에 따른 보행자의 보행속도를 측정하기 위해 제한된 환경이 아닌 실제 보행자도로에서 다수의 불특정난 피실험자를 대상으로 데이터를 취득하였다. 따라서, 영상을 기반으로 판별이 어려운 보행자의 연령 등과 같은 데이터는 분석변수에서 제외하였다.

산정된 보행속도조사 결과는 <Table 2>와 같다. 조사대상자는 636명으로 남성 341명, 여성 295명으로 나타났다. 스마트폰을 사용한 보행자는 299명, 스마트폰을 사용하지 않은 보행자는 337명으로 서비스수준은 C까지 관찰되었으며 계단 등에서 나타나는 극심한 혼잡인 서비스수준 D이하의 나타나지 않았다.

스마트폰 사용유형별 보행속도를 비교하기 위해 스마트폰 사용유형별로 서비스수준을 분류하고 이를 분산분석을 통해 통계적 유의성을 검정하였다. 정규성 검정을 위해 Levene test를 실시하여 정규성을 만족하는

경우 F검정을, 만족하지 않는 경우  $\chi^2$ 검정을 하였다. 또한 LOS 간 다중비교는 정규성을 만족하는 경우 Scheffe 방법을, 그 외는 Dunnett T3 방법을 활용하였다(Lee and Noh, 2012; Sim et al., 2016).

분산분석 결과 분류한 데이터들의 집합은 유의확률이 0.05보다 작아 스마트폰 사용유형과 서비스수준에 따른 보행속도는 통계적으로 유의한 것으로 나타났다. 스마트폰을 사용했을 때의 평균보행속도는 그렇지 않았을 때보다 남성의 경우 0.06~0.38m/s, 여성의 경우 0.06~0.35m/s 느린 것으로 나타났으며 서비스수준이 낮아질수록 평균보행속도가 낮아지는 것으로 나타났다. 또한, 남성보다 여성이 보행속도가 더 느린 것으로 나타났으며, 남성보다 여성이 스마트폰을 이용할 때 보행속도가 상대적으로 느린 것으로 나타났다.

다중비교 분석 결과, 서비스수준 간의 차이가 모두 유의수준 0.05 이하로 통계적으로 유의한 것으로 나타났다. 분석된 집단 모두 서비스수준 A와 서비스수준 C 간의 평균보행속도 차이가 가장 큰 것으로 분석되었다. 스마트폰을 사용하지 않은 경우는 남녀 모두 서비스수준 A와 서비스수준 C의 보행속도 차이, 서비스수준 B와 서비스수준 C의 보행속도 차이 순으로 높게 나타났고, 스마트폰을 사용한 경우는 서비스수준 A와 서비스수준 C의 보행속도 차이, 서비스수준 A와 서비스수준 B의 보행속도 차이 순으로 높게 나타났다

<Table 2> Result of walking speed

Gender	Usage type	LOS	N	AWS (m/s)	Levene test Sig.	ANOVA			Multiple comparison		
						F	$\chi^2$	Sig.	LOS 1(A)	LOS 2(B)	(A)-(B)
Male	Non-use	A	71	1.50	0.000	-	81.220	0.000*	A	B	0.109*
		B	59	1.39						C	0.519*
		C	31	0.98						B	C
	Auditory-use	A	13	1.44	0.001	-	40.016	0.000*	A	B	0.330*
		B	14	1.11						C	0.607*
		C	24	0.83						B	C
	Visual-use	A	46	1.44	0.000	-	87.903	0.000*	A	B	0.434*
		B	52	1.01						C	0.623*
		C	31	0.82						B	C
Fe-male	Non-use	A	74	1.43	0.000	-	80.242	0.000*	A	B	0.090*
		B	68	1.34						C	0.475*
		C	34	0.96						B	C
	Auditory-use	A	12	1.37	0.336	42.473	-	0.000*	A	B	0.364*
		B	17	1.00						C	0.540*
		C	8	0.83						B	C
	Visual-use	A	31	1.24	0.355	51.082	-	0.000*	A	B	0.254*
		B	39	0.99						C	0.453*
		C	12	0.79						B	C

AWS: Average walking speed

\* p-value<0.05

## IV. 스마트폰 이용이 보행속도에 미치는 특성 분석

### 1. 변수설정

보행자의 스마트폰 이용에 따른 보행속도 특성을 분석하기 위해 성별, 스마트폰 사용유형, 서비스수준을 독립변수로 설정하고, 보행속도를 종속변수로 설정하였다(Table 3). 변수는 영상자료에서 취득할 수 있는 자료를 활용하여 설정하였으며, 독립변수는 개인속성과 집단속성으로 구분하여 개인속성에는 성별과 스마트폰 사용유형이 포함되었고, 집단속성에는 서비스수준이 포함되었다. 의사결정나무를 활용한 분석은 각 변수들의 중요도와 복잡한 영향력을 고려할 수 있는 장점이 있어 다양한 서비스수준과 스마트폰 사용유형에 따른 복잡한 보행환경이 보행속도에 미치는 영향을 살펴보기 위해 <Table 3>과 같이 변수를 설정하였다.

<Table 3> Variables applied in the empirical analysis

Variable			Description	Unit
Explanatory variable	Personal property	Gender	Male and Female	-
		Usage type	Non-use, Auditory-use, Visual-use	-
	Group Property	Level Of Service(LOS)	LOS A, LOS B, LOS C	-
Target variable	Pedestrian walking speed			m/sec

### 2. 기계학습을 통한 보행특성 분석

본 연구에서 활용된 기계학습의 경우 입력된 데이터를 바탕으로 목표(종속)변수를 예측하는 지도학습 방법으로, 훈련 세트와 테스트 세트를 구분하여 예측하는 것이다. 따라서 본 연구에서는 일반적으로 사용되는 75%의 훈련 세트, 25%의 테스트 세트로 나눠 분석하였다(Table 4).

의사결정나무는 일반화 되지 않은 데이터의 훈련 세트를 이용할 경우, 복잡한 나무가 형성되는 과적합(overfitting)의 문제가 발생하게 된다. 이러한 문제를 해결하기 위해, 나무의 최대 깊이와 잎 노드의 최소 샘플수에 대한 파라미터를 설정하여 나무의 형태가 복잡해지기 전에 알고리즘을 종료하는 가지치기(pruning) 기법을 분석에 적용하였다. 가지치기 기법은 시행착오법을 기반으로, 1~20개의 최대 깊이와 최소 샘플수를 고려하여 수행되었다. 여기서, 최소 샘플 수는 평균오차제곱, 최대 깊이는 결정계수의 비율차이가 최소가 되는 경우를 최적 파라미터로 판단하게 된다. 실험 결과, 나무의 최대 깊이는 5, 잎 노드의 최소 샘플수는 5로 나타났으며 이를 적용하여 분석을 수행하였다.

다중회귀분석 결과 훈련 세트와 테스트 세트의 설명력은 0.626, 0.581로 나타났다. 분석에 사용된 변수는 보행속도에 모두 음(-)의 영향을 미치는 것으로 나타났으며 그 중 서비스수준 C가 기준 변수(LOS A)에 비해 가장 높은 음의 영향을 가지며, 여성이 남성에 비해 보행속도가 낮은 것으로 분석되었다.

의사결정나무의 경우 훈련 세트와 테스트 세트의 설명력은 0.647, 0.721로 다중회귀분석의 설명력보다 다소 높게 나타났다. 의사결정나무는 비모수적 분석으로 회귀계수가 도출되는 것이 아닌 목표변수에 대한 중요도(importance)가 도출이 된다. 분석결과 서비스수준 C가 목표변수인 보행속도에 대해 가장 큰 영향을 미치는 것을 확인하였으며, 가장 낮은 영향을 미치는 변수는 여성으로 분석되어 다중회귀모형과 큰 차이가 없음을 알 수 있다.

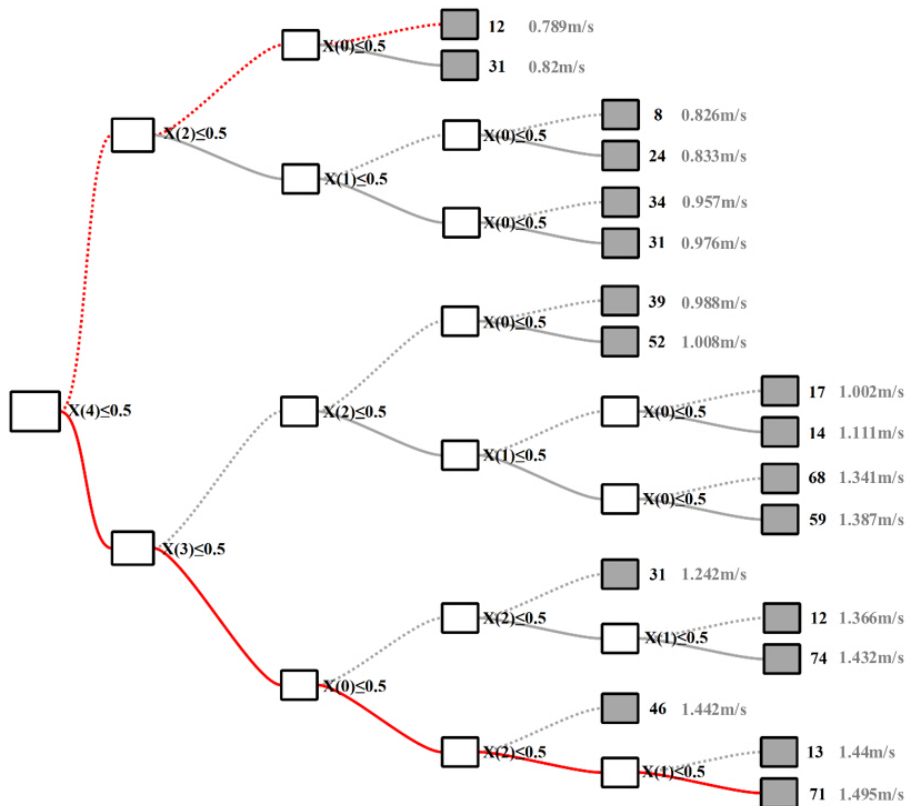
의사결정나무는 나무의 형태로 분석 모델이 나타나는 알고리즘으로 <Fig. 3>에서 알 수 있는 바와 같이

보행속도에 영향을 미치는 변수들의 복합적 영향을 분석할 수 있다. 그림에서 회색의 직사각형은 18개의 노드로 구성된 잎 노드(leaf node)를 나타내며, 중간 노드(internal node)는 17개의 흰색 직사각형을 사용하여 표현하였다. 그 중 가장 왼쪽에 위치한 흰색 직사각형은 뿌리 노드(root node)로, 나무구조가 시작되는 마디로써 학습된 데이터의 분류를 추적할 수 있다.

<Table 4> Results of the empirical analysis using machine learning models

Variable		Multiple Regression		Decision Tree	
		$\beta$	Sig.	Importance	Rank
(Constant)		1.522	0.000*	-	-
Personal property	X(0) Gender : Female	-0.049	0.004*	0.027	5
	X(1) Usage type : Auditory-use	-0.174	0.000*	0.073	4
	X(2) Usage type : Visual-use	-0.224	0.000*	0.178	3
Group property	X(3) LOS : B	-0.199	0.000*	0.187	2
	X(4) LOS : C	-0.553	0.000*	0.535	1
Explanatory Power		Training $R^2$ : 0.626 Test $R^2$ : 0.581		Training $R^2$ : 0.647 Test $R^2$ : 0.721	

\*p-value<0.05

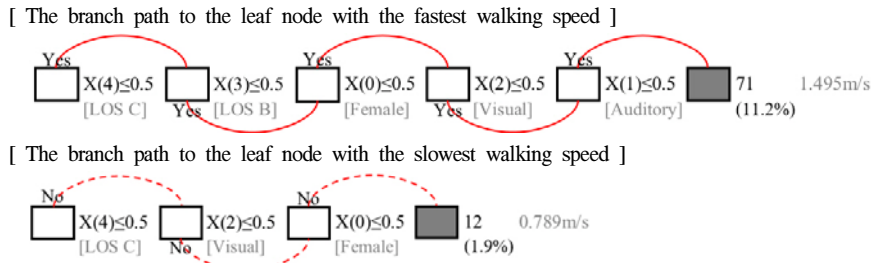


<Fig. 3> Decision tree for Pedestrian walking speed considering smartphone use



모델의 구조에서 실선은 중간 노드의 조건에 부합되면 하위 항목으로 이동함을 의미하고 파선은 조건에 만족하지 않았을 때 상위 항목으로 이동함을 의미한다. 노드의 오른쪽에 표시된 방정식은 각 노드당 배정된 샘플을 분할하는 조건을 의미하며, X(n)은 <Table 4>의 독립변수를 의미한다. 잎 노드의 오른쪽에 위치한 첫 번째 숫자와 두 번째 숫자는 잎 노드에 할당된 샘플수와 샘플의 평균이다. 의사결정나무 모델에서 전체 구조를 분석할 수 있으며, 본 절에서는 그림에서 빨간 선으로 표현된 평균보행속도가 가장 높을 때와 가장 낮을 때의 경우를 대상으로 해석해 보기로 한다. 여기서, 모델의 구조는 통계적인 기준이 아닌 의사결정나무 알고리즘기반의 기계학습에 의한 결과물로 별도의 통계적인 검증은 요구되지 않는다.

먼저 보행속도가 가장 높을 때의 경우는 1.495m/s의 속도로 보행하며 전체의 11.2%인 71명의 보행자가 해당된다. 해당 잎 노드에 대하여 경로선정에 영향을 주는 요인은 서비스수준 B와 C에 해당하는 X(3)와 X(4), 성별인 X(0), 스마트폰의 시각적 활용인 X(2)와 청각적 활용인 X(1)이다. 이 경로는 서비스수준A에서의 스마트폰을 사용하지 않는 남성 그룹에 대한 경우이다.



<Fig. 4> Detail branch paths to the leaf nodes with the fastest and the slowest walking speed

반면 보행속도가 가장 낮을 때의 경우를 살펴보면, 0.789m/s의 속도로 보행하며 이 경우에는 12명(1.9%)의 보행자가 여기에 해당된다. 경로선정에 영향을 미치는 요인은 서비스수준 C에 해당하는 X(4), 스마트폰의 시각적 활용인 X(2)와 성별을 구분하는 X(0)이다. 이는 서비스수준C에서의 시각을 이용하여 스마트폰을 이용한 여성 그룹에 대한 결과라 해석할 수 있다.

## V. 결론

본 연구에서는 보행 중의 스마트폰 이용이 보행속도에 미치는 영향을 기계학습을 통한 다중회귀분석과 의사결정나무를 통해 분석하고자 하였다. 이 때, 대상지역은 대전광역시 내의 보행로 2개소를 선정하였으며 현장조사를 통해 보행자의 스마트폰 이용행태에 대한 영상자료를 취득하여 도로용량편람에서 제시한 보행자 서비스수준의 보행밀도를 기준으로 보행속도의 특징을 살펴보았다.

보행속도 조사결과 스마트폰을 사용했을 때의 평균보행속도는 그렇지 않았을 때보다 남성의 경우 0.06~0.38m/s, 여성의 경우 0.06~0.35m/s 느린 것으로 나타났다. 또한 서비스수준이 낮아질수록 보행속도가 느려지며, 남성보다 여성의 경우가 보행속도가 더 느린 것으로 나타났다. 보행자의 스마트폰 이용에 따른 위험성을 분석하기 위해 성별, 스마트폰 사용유형, 서비스수준을 독립변수로 설정하고, 보행속도를 종속변수로 설정하여 분석에 활용하였다.

기계학습을 통한 다중회귀분석 결과 분석에 사용된 변수는 모두 음(-)의 영향을 미치며 서비스수준 C가

기준 변수(LOS A)에 비해 보행속도에 가장 높은 영향을 미치는 것으로 나타났다. 의사결정나무의 경우 서비스수준C의 경우가 보행속도에 대해 가장 중요도가 높은 것으로 나타났으며, 훈련 세트와 테스트 세트의 설명력은 0.647, 0.721로 다중회귀분석에 비해 다소 높게 나타난 것으로 분석되었다. 의사결정나무의 모델 분석 결과 서비스수준A에서의 스마트폰을 사용하지 않은 남성 그룹이 1.495m/s의 보행속도로 가장 높게 나타났으며, 서비스수준C에서의 시각을 이용하여 스마트폰을 이용한 여성 그룹이 0.789m/s의 가장 낮은 속도로 보행하는 것으로 나타났다. 의사결정나무 방법을 통해 스마트폰의 사용에 따른 보행속도에 영향을 미치는 변수를 살펴본 결과 보행환경(밀도, 인/m<sup>2</sup>)이 개인속성(성별, 사용유형) 보다 보행속도 저하에 큰 영향을 미치는 것을 확인하였다. 특히 음악 감상이나 통화와 같은 청각을 이용한 유형보다 시각을 이용하는 경우 대부분의 경우에서 보행속도가 크게 저하되며 통계적으로 유의한 차이가 있음을 확인하였다. 이는 청각을 이용하는 것과 스마트폰을 보면서 보행하는 경우 주위의 위험성을 판단하기 힘들어 보행자간의 충돌, 구조물 및 차량과의 충돌로 이어질 가능성이 높아짐을 알 수 있다.

마지막으로 본 연구에서는 자료 취득의 한계로 인해 향후에는 본 연구에서 다루지 못한 서비스수준C 미만(서비스수준D, E, F)에 대한 데이터와 다양한 토지이용에서의 스마트폰 이용, 보행자의 연령대, 직업, 다양한 보행패턴에 대한 조사도 이루어져야 할 것으로 판단된다. 본 연구의 분석결과를 기반으로 보행안전 기본 계획 또는 보행안전 종합대책의 수립 시, 스마트폰 사용자가 많은 보행환경에서의 보행특성을 고려한 대책 수립을 위한 기초자료로 활용이 가능할 것으로 기대된다.

## REFERENCES

- Choi J. H., Han S. T., Kang H. C. and Kim E. S.(1998), *Data Mining Decision Tree Analysis Using Answer Tree*, SPSS academy.
- Frenkel-Toledo N., Giladi C. and Peretz T.(2005), "Effect of gait speed on gait rhythmicity in Parkinson's disease: variability of stride time and swing time respond differently," *J. of Neuroeng Rehabil.*, vol. 2, no. 23, pp.1-7.
- Hyman Jr I. E., Boss S. M., Wise B. M., McKenzie K. E. and Caggiano J. M.(2010), "Pedestrians, vehicles, and cell phones," *Accident Analysis and Prevention*, vol. 42, pp.589-594.
- Kang S. C., Lee S. W. and Shim J. I.(2016), "A study on patterns and distraction of Smart Devices Usages While Walking," *J. of Transport Research*, vol. 23, no. 2, pp.27-39.
- Lee H. Y. and Noh S. C.(2012), *Advanced Statistical Analysis*, Moon woo sa.
- Lee J. S. and Yang S. H.(2015), "The Effect of Using a Smartphone about Response Rate of the Body to the Visual and Auditory Stimuli During Walking," *J. of the Korea Entertainment Industry Association*, vol. 9, no. 3, pp.299-305.
- Ministry of Land, Infrastructure and Transport(2013), *Korea Highway Capacity Manual*.
- Müller C. A. and Guido S.(2016), *Introduction Machine Learning with Python*, Hanbit Media, Inc.
- Park G. Y.(2017), *Risk Assessment and Prevention Plans on Drunk and Distracted Pedestrian Behavior*, Samsung Traffic Safety Research Institute.
- Park H. W. and Park S. K.(2016), "Effects of Smart-phone Use on Gait Biomechanics and Auditory Cognition," *J. of Sport and Leisure Studies*, vol. 64, pp.783-795.
- Pellecchia G.(2005), "Dual-task Training reduces impact of cognitive task on postural sway," *J. of*

*Motor Behavior; Washington*, vol. 37, no. 3, pp.239-246.

Raschka S.(2016), *Python Machine Learning*, Packt Publishing.

Rokach L. and Maimon O.(2015), *Data Mining with Decision Trees*, World Scientific.

Schwebel D. C., Stavrinou D., Byington K. W., Davis T., O'Neal E. E. and de Jong D.(2012), "Distraction and Pedestrian Safety: How Talking in the Phone, Texting, and Listening to Music Impact Crossing the Street," *Accident Analysis & Prevention*, vol. 45, pp.266-271.

Shumway-Cook A. and Woollacott M.(2000), "Attentional demands and postural control: the effect of sensory context," *J. of Gerontology-Biological Sciences and Medical Sciences*, vol. 55, no. 1, pp. 10-16.

Sim H. J., Do M. S. and Chong K. S.(2016), "Characteristic Analysis on Driver's Glance Durations with Different Running Speeds on the Expressway," *J. of Korea Institute of Intelligent Transportation System*, vol. 15, no. 1, pp.77-86.

Thompson L. L., Rivara F. P., Ayyagari R. C. and Ebel B. E.(2012), "Impact of social and technological distraction on pedestrian crossing behavior; and observational study," *Injury Prevention*, vol. 19, pp.76-82.

You S. H. and Kwon C. H.(2017), "Effects of Smartphone Usage Patterns on Human Factors of Pedestrian Safety: Focused on Users in Their 20's and 30's," *J. of Digital Convergence*, vol. 15, no. 9, pp.79-85.