Australasian Transport Research Forum 2023 Proceedings
29 November – 1 December, Perth, Australia
Publication website: http://www.atrf.info

Evaluating the pedestrian level of service of sidewalks using a machine learning model

Deborah Paul[1], Sara Moridpour[2]

[1]Civil and Infrastructure engineering

[2] Civil and Infrastructure engineering

Email for correspondence (presenting author): s3764996@student.rmit.edu.au

Abstract

The service quality of pedestrian facilities varies from place to place depending upon the comfort experienced by the pedestrians. The feeling of comfort experienced by the pedestrians on sidewalks in the central cities differs from the comfort experienced by the pedestrians in residential areas mainly due to the differences in the volume of pedestrians, land use and land use mix, and urban structures. Pedestrian comfort also differs from person to person depending on their gender, age, trip purpose, etc. The Pedestrian Level of Service (PLOS) on sidewalks in central cities is an issue to consider because of the essential amenities such as universities, tourist attractions, big shopping centers, and city businesses. Pedestrian mobility and comfort level on the sidewalks must be assessed from time to time. This study determines the factors influencing the PLOS of sidewalks in central cities by collecting data from pedestrians walking in the Melbourne Central Business District (CBD). The data were collected using questionnaire surveys in Melbourne CBD. The pedestrian responses and density recorded by sensors were used in developing a PLOS model using Machine Learning (ML) techniques. The variables used in model development have been selected using the Chi-square test, which assists in choosing the influencing variables on PLOS. The Naïve Bayes algorithm is used to develop a model to predict the PLOS of sidewalks using the six significant variables from the Chi-square test. The accuracy of PLOS prediction using the Naïve Bayes classifier model is 60%. The Naïve Bayes model was analysed using LIME, which stands for Local Interpretable Model-agnostic Explanations that help to find the variables influencing prediction of PLOS for any instance using the ML model.

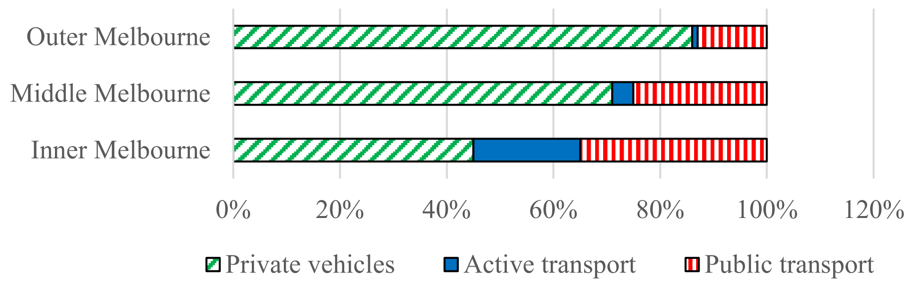
Keywords: Pedestrian; Comfort; Level of Service; Sidewalks.

1. Introduction

1.1. Background

Walking is an essential mode of transport that has been considered a form of active transportation in transport infrastructure. Effective Job density Walk score relates to the ease with which a person can access many jobs by walking (SGS Economics and Planning, 2014). If public transport is involved, then it produces a high value of that score. From the Victorian Integrated Survey of Travel and Activity, in the 2018 survey results, we find that 55% of the people traveling in the city to their work use public transport or bicycle or walk. As we move far from the city, in the middle and outer Melbourne, this percentage decreases and is replaced by cars and private vehicles, as shown in figure 1 (VISTA, 2018).

Figure 1: The proportion of journeys to work by region and mode in Melbourne in 2018 (VISTA, 2018)



As the population of cities increases every year, more space needs to be allocated for pedestrians. Otherwise, the comfort level of pedestrians would drop, people would feel discouraged from walking, and it will cost the economy. Alfonzo has referred to walkable communities as those planned and designed with built environmental variables that influence the walking activity of people and show relatively high numbers of incidental walking by adults for leisure or transport (Alfonzo, 2005). The neighbourhood's Walkability is strongly related to the physical activity level of the people living in that area by linking their Body mass indices (BMI) and the cause of chronic diseases (Smith et al., 2011). By 2036, Melbourne expects about 1.4 million people to walk around its city. To maintain a competitive environment for people to live, work and do business, the transport department wants to provide a safe, comfortable, and efficient public space and streets to move around and enjoy the city (City of Melbourne, 2020). In central cities, the service level of pedestrians must be measured at timely intervals to keep up a safe and healthy environment for walking.

Walkability and Pedestrian Level of Service (PLOS) are two commonly used terms that refer to the quality of service that a pedestrian facility can offer to pedestrians. Although the early methods have used pedestrian density and walkway width to calculate the PLOS (Fruin, 1971, HCM, 2000), the current plans involve pedestrians' perception of the level of service, in addition to features such as walkway width, speed, and density of pedestrians, etc. (Kim et al., 2014, Sahani et al., 2017, Ye et al., 2015). Pedestrians' comfort changes with the time of day, day of the week, location, and trip purpose (Kadali and Vedagiri, 2015). It is necessary to find a suitable model that will help evaluate the PLOS by identifying the factors that affect the LOS after the occurrence of events that affect people's physical, social, and mental aspects.

1.2. Objectives and scope

Several studies have formulated a model to find the PLOS of sidewalks, crosswalks, intersections, and other pedestrian facilities. This study mainly focuses on determining the factors contributing to the PLOS of sidewalks using Chi-square test and formulating a machine learning model for predicting the pedestrian level of service using categorical variables such as Pedestrian crowd, High-speed traffic, non-slippery footpath, Personal space, Vehicle Volume, and Construction works for modelling.

This paper is made of the following sections. In the next section, the previous studies on PLOS and the methods used have been discussed. In section 3, the data used in the research and the data collection methods have been explained. Section 4 represents the methodology and the analysis of the significant variables used in the model. Section 5 describes the model development and its performance. Finally, section 6 summarises the research and suggests further research work.

2. Literature review

The pedestrian level of service has been widely used to refer to the operational quality of the pedestrian facilities such as sidewalks, footpaths, and crosswalks, and it has been derived from the Level of Service concept in traffic studies, which was first used in vehicles and then later developed for pedestrian facilities (Cepolina et al., 2018). The first method to measure the performance of sidewalk capacity was expanded by Fruin, known as Fruin Scale, which gives the standards based on speed, ability to overtake slow-moving pedestrians, and bidirectional flow movements (Fruin, 1971). The measure of pedestrians' perception of safety and quality was first developed by field measurements by Landis, using factors such as speed, travel time, traffic interruption, freedom to manoeuvre, comfort, and convenience of using the facility (Landis, 2001). The Highway Capacity Manual gives a measure to evaluate LOS based on the effective width of the walkway, pedestrian flow rate, and pedestrian density, which forms a base for the development of other models that evolved further at later stages (HCM, 2010). The behaviour of pedestrians is quite different compared to the conduct of vehicles while assessing the level of service. Hence, factors such as personal space and evasive movements have been considered apart from pedestrian density and flow used by traditional methods such as HCM2000 (Kim et al., 2014, HCM, 2000). Studies have shown that land use planning needs to be considered to ensure safe and comfortable pedestrian access, as different land use conditions influence pedestrian behaviour and hence reduce pedestrian safety in many developing countries (Dumbaugh and Li, 2010). Network-related factors are considered to evaluate the pedestrian quality attribute of the walkway while calculating the cost of route choices. Individual characteristics also affect routing decisions and network factors based on time constraints and personal abilities (Czogalla, 2012). Thus, we find that the Pedestrian Level of Service measures the service quality of the pedestrian facilities by using variables at the micro level compared to those involved in assessing Walkability.

Cepolina et al. (2018), has assessed the level of service of a pedestrian facility as a function of aggregation of individual comfort levels. At every given time and position, each pedestrian will experience a personal comfort level which can be evaluated from the available space, called open space, and the required space. If the available space is less than needed, an individual will feel reduced comfort. In this methodology, the concept of the Voronoi area and Voronoi link has been taken from Xiao et al.'s (2016) description of the Voronoi diagram. The PLOS has been assessed from the aggregation of the loss of required space for each pedestrian in each segment, and the values are compared against the HCM values and it was found that each pedestrian feels a different level of comfort based on their position (Cepolina et al., 2018). This method requires following the pedestrian trajectory; applying it to larger areas will be difficult. Cepolina et al. (2017) also analysed the impact of social groups on the level of service of the people walking in groups. In a corridor in the Institute of Massachusetts Institute of Technology in the US, Microsoft kineticV1 and a video camera were used to collect data such as pedestrian flow, speed, and time taken to cross the section by focusing on a rectangular area of 4.9 m^2. HCM 2000 method explains the pedestrian LOS in terms of the space available for pedestrians. The speed and density of each pedestrian have been calculated, and the values have been plotted on a graph (speed, density) by distinguishing the points of people walking alone and those of people walking in groups. From the chart, the LOS has been derived for each pedestrian using the HCM 2000 method, and it was found that 36% of points cannot be assessed under LOS as they had very different speeds and densities. This irregularity in speed and density has been claimed because of voluntary groups, as the speed of a person walking in groups is less than the speed of a person walking alone (Cepolina et al., 2017).

Kim et al. (2014) have considered factors such as personal space and evasive pedestrian movements in assessing PLOS for sidewalks. Subjective data was collected using a questionnaire survey carried out in Seoul, South Korea, at 28 different sites; 569 participants were asked about the comfort level of the sidewalk they had walked through and the top two factors which they considered important among pedestrian facilities, such as aesthetics, the influence of surrounding facilities, maintenance of sidewalk pavements, etc. Video recordings produced 468 sample recordings of 5 mins each, showing the evasive movements of pedestrians during peak and off-peak hours. The objective data, such as the effective width of the sidewalk and pavements and the number and size of fixed objects on the pedestrian facilities, were also recorded from videos, and pedestrian volume at those sites' residential, commercial, and recreational areas were recorded also collected. The pedestrian LOS calculated by KHCM's method using Multi regression analysis has been compared with the perceived LOS by 216 pedestrians at 12 facilities. The main factor used in KHCM to describe LOS has been checked against the top two primary factors rated by most pedestrians about the facilities: walking speed and the number of conflicts (Kim et al., 2014).

Ye et al. (2015) have considered the effect of delays of pedestrians at signalized intersections under mixed traffic conditions in countries like China, where there are non-uniform arrival rates and non-compliant behavior. In the Highway capacity manual, Webster's model has been mentioned where the pedestrian arrival rate is uniform and they comply with traffic signals (HCM, 2000), but further delay models have been proposed for different scenarios by Braun and Roddin and Li-et al. (Braun, 1978, Li et al., 2005). The data has been collected for this study at ten crosswalks from five intersections with a reasonable flow of pedestrians, a traffic signal, and a crossing facility and sidewalk are available. Three hours of video recording data were collected, which helped to gain information about the arrival rate of pedestrians, walking time of pedestrians during green and non-phase, number of pedestrians who ran while crossing the road, delay, number of pedestrians starting to cross during non-green phase and time taken by a randomly chosen pedestrian who didn't encounter any conflicts. The proposed model was derived from Webster's model with some correction factors included for non-uniform arrival rates and signal compliance and is calculated as – the sum of delay during the green and non-green phases. The factor for the non-uniform arrival rate is taken from Li et al.'s model (Li et al., 2005). A total of 1257 responses were collected from pedestrians in that area when they were asked to rate the LOS of the pedestrian facility from A to F. Cumulative logistic regression was used to develop a LOS model which will consider pedestrian delay as one of the factors that influence the LOS (Ye et al., 2015).

A model for the pedestrian level of service at signalized intersections in developing countries has been developed considering pedestrians' safety, convenience, and efficiency using fuzzy linear regression analysis. Data has been collected from nine signalized intersections in India, using surveys, regarding the efficiency, convenience, and safety level perceived by crosswalk pedestrians. Simultaneously a video camera has been used to record the factors such as walking speed, delay of pedestrians, and traffic volume. The Pearson correlation test was performed to find the significant variables contributing to the combined score of efficiency, safety, and convenience. Four variables that significantly affect the PLOS, such as vehicle volume, pedestrian delay, probability of pedestrian interaction with vehicles, and pedestrian facilities, have been used in the Fuzzy linear regression model with the help of MATLAB R2014 to derive the mathematical model for finding PLOS score from A-F (Marisamynathan and Vedagiri, 2019). Zhao et al. (2016) quantified the effects of environmental factors, road facilities, and traffic conditions by image characteristics extraction and edge detection method. It was found that factors such as pedestrian flow rate, motorized and non-motorized flow rate, road crossing section, the effective width of sidewalks, frequency of obstructions on sidewalks, segregated facility between pedestrians and vehicles, on-street parking facility, greenery, the orderliness of shops away from sidewalks affect the pedestrian satisfaction on streets. The pedestrian perspective's sense of safety and comfort was collected through the intercept survey method in residential, commercial, and transportation hubs on different kinds of roads such as expressway, arterial, minor arterial, and branch roads. As in this study, the PLOS was affected by multifactor; a fuzzy neural network model was used to do multiple inputs and a single output. The pedestrian satisfaction ratings from 1-10 were converted into six grades of the level of service using the Fuzzy clustering method. This method, when compared against linear regression, and the results were found to agree. However, this method requires quantifying environmental variables by image recognition, which is rarely used in PLOS calculations (Zhao et al., 2016).

Asadi-Shekari et al. (2014) have suggested an analytical point system for finding PLOS by comparing the current condition of various features in a university campus in Malaysia to a standard. It will help identify the existing streets' problems and can give information about improving those pedestrian facilities. This process has been done in three stages. In the first stage, the 27 indicators that affect the pedestrian facility has been selected, such as slower traffic speed, buffers and barriers, fewer traffic lanes, mid-block crossing, landscape and trees, fire hydrants, furniture, footpath pavement, marking, corner island, the sidewalk on both sides, advance stop bar, width of the footpath, driveway, lighting, signing, bollard, slope, curb ramp, tactile pavement guiding, tactile pavement warning, ramp, grade, signal, bench, and seating area, wheelchair accessible fountain and drinking fountain. These indicators have been assigned a coefficient as they have a different effect on the pedestrian level of service. Standards of guidelines for these indicators are used to give points or weights to the indicators as well as to compare the present condition against the required standard. In this study, PLOS is derived as a percentage of existing PLOS. The ideal PLOS and various PLOS ranges are calculated for the university campus (Asadi-Shekari et al., 2014). In India, one of the developing countries, the level of service at midblock crosswalks at unsignalized places has been analysed. Such scenarios are prevalent when pedestrians cross the roads due to easy access to nearby land-use activities such as shopping, educational institutions, or business areas. Eight locations with unprotected mid-block crosswalks were selected, in Mumbai, India, for conducting a questionnaire and graphic video survey, and pedestrians were asked to rate their perception of the LOS in terms of vehicle flow, pedestrian comfort, and waiting time as their difficulty factor. The video camera collected data about vehicle encounters, pedestrian speed, waiting time, and crossing time. The ordered probit model has been used to find the effect of the variables such as perceived crossing difficulty, safety, number of lanes, median width, and number of vehicles on pedestrian-perceived LOS as it was found that land-use types used in this study, such as Industrial, Business, Mixed, Residential, and Shopping, has a good correlation with median width, crossing time, number of lanes, and pedestrian wait time which in turn influence the pedestrian perception of LOS. The model has been validated using a questionnaire and video survey from 102 respondents in a residential area and has produced a successful prediction rate of 67.64% for perceived PLOS (Kadali and Vedagiri, 2015).

The PLOS of midblock in a Greek city was evaluated by conducting questionnaire surveys and collecting data about the sociodemographic features such as gender and age of pedestrians and by considering their perceived comfort and street characteristics. An ordinal regression model was developed based on the accumulated questionnaire data comprising the qualitative data. This method can help find the perceived level of service of pedestrians of different age groups and genders (Georgiou et al., 2021). Pandemic pedestrian level of service has been calculated using the data obtained from pedestrian sensors, Apple and Google based on social distancing to relieve congestion in areas where there is more pedestrian flow in Madrid (Talavera-Garcia and Perez-Campana, 2021). PLOS of sidewalks in Dhaka metropolitan area has been evaluated by using 15 variables related to sidewalks safety and security collected through questionnaire surveys. These variables were then grouped under three latent variables using factor analysis to find the effect of these latent variables on the LOS by applying Structural equation modelling (Jahan et al., 2020).

Table 1: Pedestrian level of service models in literature

Models	References	Limitations
Mathematical models (Voronoi area based, Average value and points system)	Cepolina et al., 2018, Cepolina et al., 2017, Asadi-Shekari et al., 2014, HCM, 2010, Talavera-Garcia and Soria-Lara, 2015.	-sometimes involves complex calculations -The results cannot be validated
Statistical models (Regression, structural equation modelling and probit model)	Kim et al., 2014, Ye et al., 2015, Kadali and Vedagiri, 2015, Georgiou et al., 2021, Marisamynathan and Vedagiri, 2019.	-The accuracy of these models is not very high - The results cannot be validated using test data

Thus, we find that PLOS has been evaluated using various methods such as multilinear regression, fuzzy linear regression, cumulative logistic regression, structural equation modelling, ordinal regression, fuzzy neural network, and ordered probit model. A comparison of some major methods has been listed in Table 1. These methods have used different set of variables to evaluate the PLOS which is specific to the location and country where the study has been conducted. Various locations considered for the study includes crosswalks, signalized intersections, and sidewalks. The purpose of this research is to focus only on sidewalks and to find the factors influencing the PLOS and the model to predict it for planning purposes.

3. Data

Questionnaire surveys were conducted in Melbourne CBD to identify pedestrians' viewpoints on the pedestrian facilities and their comfort while walking on sidewalks in Melbourne (Moran et al., 2018, Shatu et al., 2019, Jena et al., 2017, Bai et al., 2017). Seven locations in the CBD have been selected to conduct surveys. These locations include RMIT University City Campus building 14 and 80, which a representative of education centres where pedestrians are mainly students, and Flinders Street train station and Southern Cross train station, where the pedestrians are mostly frequent commuters that go to work, and Bourke Street, Lygon street east and Lygon street west where the pedestrians mostly go for recreational purposes. These seven locations also have installed pedestrian sensors, which provide the count of pedestrians passing through those walkways every hour.

In this research, walkway features, such as the surface of the footpath, the width of the pathway, continuous footpath, street furniture, lighting, buffers, road verge, street benches, street vendors, slow-moving pedestrians, and volume of pedestrians in the opposite direction, and road traffic factors such as noise of traffic, detours, and on-street parking have been considered as qualitative variables to find if they influence the pedestrians' perceived level of service. Pedestrian volume changes at different times of the day, days of the week, and at various locations in Melbourne CBD. Thus, the pedestrian flow rate is taken as the quantitative variable to determine the pedestrian level of service. The pedestrians' feelings about less crowded streets and covid safe distance or social distancing have also been considered as qualitative variables, as shown in Table 2. The pedestrians rated these questions on a scale of 1 to 5, where one stands for strongly disagree and 5 for strongly agree. The questionnaire has two questions on how the pedestrian feels about the pedestrian crowd around them and vehicle traffic volume at that time on a scale of 1 to 5, where one is very uncomfortable, to 5 stands for very comfortable.

The final question is about the overall comfort experienced by the pedestrians while walking on that sidewalk which is rated from 1 to 5, which denotes the pedestrian perceived level of service from A to E, where A (rating 5) stands for excellent, and E (rating 1) stands for the worst condition of level of service. Traditional methods have used a scale of 1-6 to identify the different levels of service (HCM, 2010). But for midblock pedestrian walkways, it is hard for pedestrians to feel comfortable on a scale of 1-6. Hence following the range of A to E used by previous research studies, PLOS has been categorized into five levels (Kadali and Vedagiri, 2015, Bai et al., 2017, Li et al., 2012, Kang and Lee, 2012).

Table 2: Categorical explanatory variables

Variable name	Description	Values	Frequency
Nominal Variables			
Location	Survey location	1. Flinders Street station 2. Southern Cross station 3. RMIT 14 4. RMIT 80 5. Bourke Street 6. Lygon Street East 7. Lygon Street West	25% (170) 19% (132) 4% (29) 27% (183) 6% (40) 14% (96) 5% (34)
Gender	Gender of respondents	0. Male 1. Female	58% (394) 42% (288)
Social Distancing	Pedestrians follow social distancing on the sidewalk	0. No 1. Yes	62% (423) 38% (261)
Ordinal Variables			
Age Group	Age of respondents	0. 18-35 1. 36-50 2. 51-65 3. >65	68% (465) 21% (141) 6% (43) 5% (35)
Comfort Distance	Comfort distance of a pedestrian from other pedestrians on the sidewalk	0. 0-1 1. 1-1.5 2. 1.5-2 3. >2	41% (282) 43% (291) 13% (90) 3% (21)
Pedestrian Crowd	Pedestrian's feeling of comfort concerning other pedestrians around them	1. Very uncomfortable 2.UnComfortable 3. Neutral 4. Comfortable 5. Very comfortable	0% (2) 2% (14) 12% (84) 35% (240) 50% (344)
Vehicle Volume	Pedestrian's feeling of comfort concerning cars and vehicles on the road next to them	1. Very uncomfortable 2.UnComfortable 3. Neutral 4. Comfortable 5. Very comfortable	0% (2) 2% (14) 14% (95) 33% (229) 50% (344)
Pedestrian path and Flow characteristics			

Continuous Footpath	Footpaths are continuous on both sides of the road	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	0% (0) 1% (10) 9% (60) 31% (214) 59% (400)
Wide Footpath	Footpaths are wide enough to walk	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	1% (5) 6% (40) 16% (111) 36% (244) 42% (284)
Street Furniture	Minimal street furniture gets in the way while walking	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	1% (5) 7% (51) 19% (130) 36% (247) 37% (251)
Footpath Surface	The footpath surface is safe and in good condition	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	1% (6) 8% (52) 20% (137) 36% (250) 35% (239)
Lighting	Street and footpath lighting is always good	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	0% (1) 3% (18) 12% (80) 33% (227) 52% (358)
Buffers	Buffers are present between pedestrians and the road	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	2% (12) 9% (59) 22% (153) 35% (236) 32% (224)
Non-slippery	Footpaths are not slippery even after rain	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	3% (24) 14% (96) 34% (231) 33% (224) 16% (109)
Landscaping	Road verge is present on footpaths	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	6% (38) 10% (70) 26% (180) 30% (203) 28% (193)
Street Benches	Street benches are widely available	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	2% (15) 11% (79) 24% (162) 30% (208) 32% (220)
Street Vendors	Street vendors and outdoor seating are not disturbing to walk	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	2% (12) 5% (34) 17% (118) 34% (232) 42% (288)

On-street Parking	On-street parking does not affect visibility	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	2% (12) 4% (28) 15% (102) 31% (209) 48% (333)
Slow moving Pedestrians	Slow-moving pedestrians rarely block the footpath	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	4% (30) 15% (99) 24% (165) 35% (239) 22% (151)
Opposite direction Flow	Pedestrians in the opposite direction rarely get in the way	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	3% (18) 14% (98) 27% (185) 34% (232) 22% (151)
Personal Space	A comfortable personal space can always be maintained.	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	4% (25) 13% (87) 28% (193) 29% (198) 26% (181)
Highspeed Traffic	Footpaths next to high-speed or high-volume traffic are not disturbing	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	1% (10) 5% (37) 22% (154) 34% (224) 38% (259)
Detours	Detours on footpaths due to roadworks are minimal	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	1% (10) 5% (36) 20% (132) 34% (233) 40% (273)
Construction Sites	Safe passage is available when construction sites occupy footpaths	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	1% (9) 4% (29) 17% (116) 30% (204) 48% (326)
Covid Safe Distance	It is always possible to maintain COVID-safe social distancing (1.5m)	1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree	13% (92) 18% (125) 33% (217) 22% (153) 14% (97)
PLOS	The overall comfort of walking on the footpath	1. Very Poor (E) 2. Poor (D) 3. Average (C) 4. Good (B) 5. Very Good (A)	3% (24) 12% (84) 32% (217) 42% (286) 11% (73)

After analyzing the pedestrian volume data from sensors, the peak hours of the pedestrian crowd near the university location were found to be from 10:00 am to 12:00 pm and from 2:00 to 4:00 pm, whereas, for train stations, the pedestrian volume is high during the morning peak from 7:00 to 9:00 am, lunchtime from 12:00 to 2:00 pm and afternoon peak from 3:00 to 5:00

pm. The survey was conducted between 7:30 am to 5:00 pm at all seven locations to identify the pedestrians' comfort and the perceived level of service during the peak and off-peak periods. The total number of responses collected from the survey was 682, out of which 58% of respondents were male, and 42% were female respondents. Among the male and female survey respondents, 68% were in the age group of 18-35 years due to universities in the city. 21% of pedestrians were in the age group of 36-50, and their purpose for the trip was mostly work. 11% of participants belonged to >50 years, and their purpose of the trip was mainly recreation, medical purpose, and other.

The pedestrians were questioned if they followed social distancing while walking on that sidewalk. About 38% of the pedestrians who took part in the survey answered 'yes,' and only 62% of the pedestrians replied 'no' for social distancing. Hence it shows people mostly don't care about social distancing, and the effects of Covid has been removed from people's mind.

4. Methodology

4.1. Chi-square test

The machine learning model can use the chi-square test for feature selection. It is used to test the independence of two variables. It is an important method to find the relationship between two categorical variables in a contingency table. Chi-squared statistics and P-values are used to assess the relationship between two variables. In this case, the Chi-square statistic is a number that reflects how much difference exists between observed and expected values that are collected. The higher the Chi-square value, the feature or variable is more dependent on the response variable and can be included in training the model. The formula for the Chi-square statistic is given by:

$$\chi_{e}^{2} = \Sigma \frac{(O-E)^{2}}{E} \qquad \text{---------------------- (1)}$$

Where e is the degrees of freedom, "O" is the observed value, and "E" is the expected value. The number of categories minus 1 gives degrees of freedom.
Degrees of freedom and "α" (alpha) values are required to find the p-value and chi-square statistic. α denotes the significance level used to state the association between two variables if they are associated or not. The value of α is usually taken as 0.01 or 0.05. If p- value ≤ α then we reject the null hypothesis and assume that there is significant association between the two variables. If p- value ≥ α then we cannot assume there is significant association between the two variables. Thus, smaller the p-value, the difference is significant or small, and the variable can be selected for model training.

4.2. Naïve Bayes Classifier

It is a probabilistic machine-learning model that is used for classification tasks. Based on the Bayes theorem, the probability of event A happening can be found, given that event B has already occurred. It assumes that all features/predictors are independent and contribute equally to the outcome. Hence it is called Naïve. Bayes theorem could be stated mathematically as,

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)} \qquad \text{-----------------------------(2)}$$

To the dataset, Bayes theorem can be applied in the following way:

$$P(y \mid X) = \frac{P(X \mid y)\, P(y)}{P(X)} \qquad \text{-------------------------------(3)}$$

Where y is the class or response variable, and X is a dependent feature vector of size n,
Where X = (x_1, x_2, x_3...... x_n)
There are three types of Naïve Bayes Classifiers: Gaussian, Multinomial, and Bernoulli, based on whether the feature vectors consist of continuous, discrete, or Boolean values. It performs better than many other models if independent predictor variables are assumed. It needs a small amount of training data to predict test data and hence less time consumed. It is easy to implement. The demerits to be considered while using this model are the assumption that all variables are independent does not occur in real life. If test data has a category not observed in training data, it assigns probability as zero. A smoothing technique is needed to overcome this issue.

4.3. Performance metrics and LIME

Classification models are used to find the target class of the dataset, where the target variable, such as PLOS, is categorical. It is necessary to check the performance of classification models before they can be applied to real-world situations. Hence, various measures are available to fit the model's performance. The most used metric for evaluating the performance of machine learning models is accuracy. Accuracy is an appropriate measure for assessing performance only if the dataset has a balanced class. Precision, recall, and f-score must be considered for imbalanced classification problems. These metrics can be understood with the help of a confusion matrix for binary classification problems, as shown in **Table 3**.

Table 3: A confusion matrix representing a binary classification problem

		True Class		
Predicted	Class	Positive Class	Negative Class	Σ
	Positive Class	True Positive (TP)	False Positive (FP)	TP+FP
	Negative Class	False Negative (FN)	True Negative (TN)	FN+TN
	Σ	TP+FN	FP+TN	TP+FP+FN+TN

True Positive (TP): True positive is the value of correct predictions of positives out of actual positive cases.

False Positive (FP): False positive denotes the value of incorrect positive predictions

True Negative (TN): True Negative represents the number of correct predictions of negatives out of actual negative cases.

False Negative (FN): False negative represents the number of incorrect negative predictions.

Precision: The precision score represents the classification model's ability to correctly predict the positives out of its total positive predictions. It is helpful to evaluate the performance when the classes are imbalanced.

Recall: Recall score represents the ability of the model to correctly predict the positives out of the actual positives. It is different from the precision score, which gives the value of how many predictions are positive out of all positive predictions.
F1-score: The F1-score metric is a function of both Precision and Recall. It gives equal weight to them without using the total number of observations and is used to evaluate the performance of imbalanced classification models.
Accuracy: Accuracy is the value that informs the number of times the model predicts correctly out of the total number of predictions.

4.3.1. LIME

Lime stands for Local Interpretable Model-agnostic Explanations, meaning it explains model-based local values. It helps to find which features are essential in predicting the outcome and what values are assigned to each feature to give the result. LIME will be used to understand how the model will predict the PLOS rating for each instance.

4.4. Comparison of machine learning models

Classification is the method of allocating objects under different classes. Most often used machine learning algorithms for classification problem are Decision tree, Naïve Bayes, k-Nearest Neighbor (kNN) and Support Vector Machine (SVM). The performance of these models has been tested by Sheth et al (2022), where five well-known datasets such as placement dataset, wine quality dataset, heart disease dataset, glass quality dataset and classification of jobs has been taken for working out the performance metrices which are precision, recall, accuracy, and f1-score using the popular machine learning algorithms. The value of metrics calculated using decision tree, k-NN, Naïve Bayes and SVM shows that Naïve Bayes performs well in terms of accuracy, precision, recall and f1-score compared to rest of the other algorithms with SVM placed second followed by k-NN and Decision tree models (Sheth et al, 2022). The energy performance of buildings that can be classified using 13 parameters has been modelled using three machine learning languages which are decision tree, k-NN and Naïve Bayes to compare the performance of the three models. It has been found that the average F-measure of decision tree was 0.676, k-NN was 0.543 and Naïve Bayes was 0.780. Thus proving that the precision and recall values of Naïve Bayes is also better than the decision tree and k-NN (Ashari et al, 2013). Hence Naïve Bayes algorithm has been used in this study to predict the PLOS of pedestrians on sidewalks which needs a classification model.

4.5. Significant variables selection and analysis

The pedestrian questionnaire survey had questions on pedestrian footpath features such as footpath surface, continuity of footpath, street furniture, on-street parking, the noise of vehicles on the roads next to them, and pedestrian flow attributes such as slow-moving pedestrians, opposite direction flow of pedestrians, covid safe social distance and comfort experienced at the presence of pedestrian crowd around them. The final question about the footpath's overall comfort level and volume compared to an ideal pathway is taken as the PLOS value in a range of 1 to 5. Initially, a Chi-square test is performed to find the most significant variables that influence the PLOS value and could be used in the model. The variables with higher chi-square statistics and p-value less than 0.05 are the most important variables influencing the PLOS.

Table 4: Significant variables

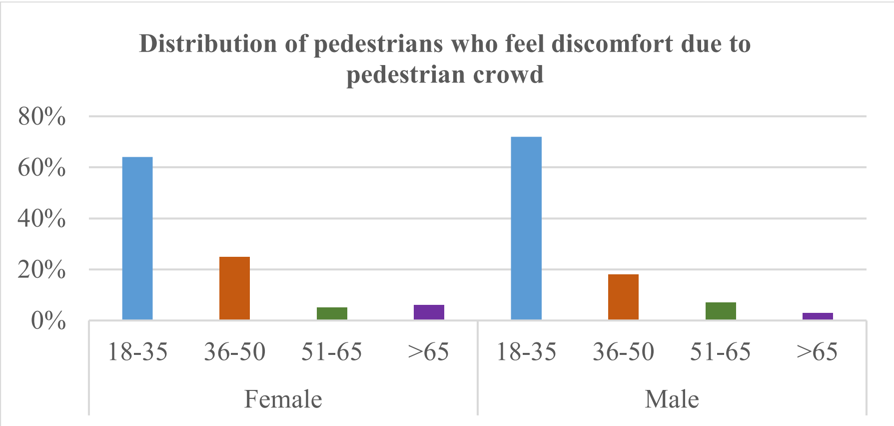
No.	Features	Chi-square statistic	P-Value
1	Pedestrian Crowd	248.17	1.22e-43
2	Highspeed Traffic	247.75	1.49e-43
3	Non-slippery	240.26	5.11e-42
4	Vehicle Volume	192.43	2.67e-32
5	Construction sites	189.70	9.47e-32
6	Personal Space	180.39	7.01e-30

Table 4 shows the variables Pedestrian Crowd, Highspeed Traffic, Non-slippery, Vehicle Volume, Construction Sites, and Personal Space that have P-values less than 0.05 and higher Chi-square statistic values compared to the rest of the variables and hence would be used for modeling. The only quantitative variable, the pedestrian flow rate per hour, has yet to appear significant and will not be included in the model.

4.5.1. Pedestrian crowd and personal space

Pedestrian Crowd refers to the comfort experienced by pedestrians concerning the volume of pedestrians around them. Personal space shows the comfort zone each person experiences around them, which the pedestrian crowd affects. The rating for these factors can change from peak to off-peak hours when pedestrians' volume drops. It can also vary based on age group and gender. Previous studies have considered pedestrian volume and personal space as essential factors in finding the PLOS for aggregated and disaggregated methods (Sahani and Bhuyan, 2014, Kim et al., 2014, Jia et al., 2022). Figure 2 shows the distribution of Pedestrian discomfort due to the pedestrian crowd around them.

Figure 2: Effect of the pedestrian crowd on pedestrians by gender and age group

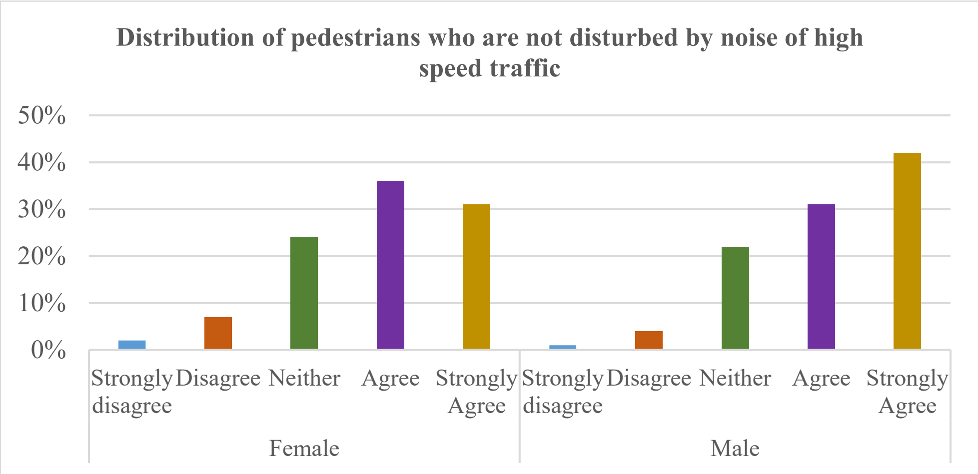


The graph shows that a similar level of discomfort has been felt by both male and female pedestrians across all age groups. Elderly females are affected more compared to elderly males. Personal space is also a very important factor, and loss of personal space has been calculated for disaggregated methods to find the PLOS using the HCM method (Cepolina et al., 2018).

4.5.2 Vehicle volume and noise of traffic

Vehicle volume refers to pedestrian comfort concerning the number of vehicles on the road while walking. The variable high-speed traffic has been used to include the noise disturbance caused by cars on the road next to them (Vallejo-Borda et al., 2020). These are essential factors influencing the PLOS and have been included in finding the LOS in the literature (Zhao et al., 2016, Bivina and Parida, 2019, Hasan et al., 2015). Figure 3 shows the distribution of pedestrians by gender who are not affected by the noise of cars and heavy vehicles on the nearby road.

Figure 3: Effect of the noise of traffic on pedestrians by gender



It shows that the percentage of female pedestrians who disagree is more than that of male pedestrians, and males quite strongly agree that traffic noise doesn't affect them compared to females.

4.5.3 Non-slippery footpath surface:

The footpath surface and the non-slippery surface have been analyzed as two separate variables because the former refers to the cleanliness and the surface free from trip hazards (Bivina and Parida, 2019), and the latter refers to the footpath surface not being slippery after rain. Melbourne often has light to heavy showers, and the chances of dirt and trash causing the surface to be oiled after rain are high. Only 16 % of the respondents rated that they strongly agree that the surface is not slippery, whereas 17% of pedestrians disagreed that the footpath surface is safe after rain.

4.5.4 Construction sites/works:

After Covid, construction works in the city were resumed, and the number of work sites has considerably increased. Although this is not a long-term factor at a particular location in the city, construction work continues as renovations happen at various locations (Basbas). Construction sites have been a significant variable in assessing the level of service because more than 75% of the pedestrians have agreed that safe passage is available at places where construction work happens.

5. Model development and results

The collected pedestrian survey data has been split into two, 80% of the data was used for training the model, and the remaining 20% was used for testing the model. Classes D and E were combined as there needed to be more responses in class E. As a result, the PLOS were graded from A to D, like the previous research in the literature (Kadali and Vedagiri, 2015, Bai et al., 2017, Li et al., 2012, Kang and Lee, 2012). The Naïve Bayes Classifier algorithm was coded using Python in Jupyter Notebook 6.4.8, and the model was generated with 60% accuracy in predicting the pedestrian level of service. The corresponding values of precision-recall and f1 score for the four categories of PLOS are shown below in Table 5.

Table 5: Model performance metrics

PLOS category	Precision	Recall	F1-score	support
A	0.59	0.71	0.65	14
B	0.60	0.67	0.63	55
C	0.59	0.37	0.46	43
D	0.59	0.76	0.67	21
Macro average	0.59	0.63	0.60	133

The dataset used in this study with four classes of PLOS as dependent variable is imbalanced data. Hence the f1-score of the model is considered for evaluating if the model could be accepted. The overall f1-score of 0.60 for the test data shows that the model has good predictive capacity. For PLOS classes A, B, and D, the recall, precision, and f1-score values are high, with an average value of 0.65 which shows that nearly 65% of the prediction belonging to classes A, B, and D are the same as their actual observed classes. The recall and f1-score for PLOS category C are slightly lower than other classes, which could be understood using the confusion matrix shown in Figure 4. The second row of the matrix corresponds to the prediction of class C, where there is a misclassification rate of 56% which is high compared to other classes.

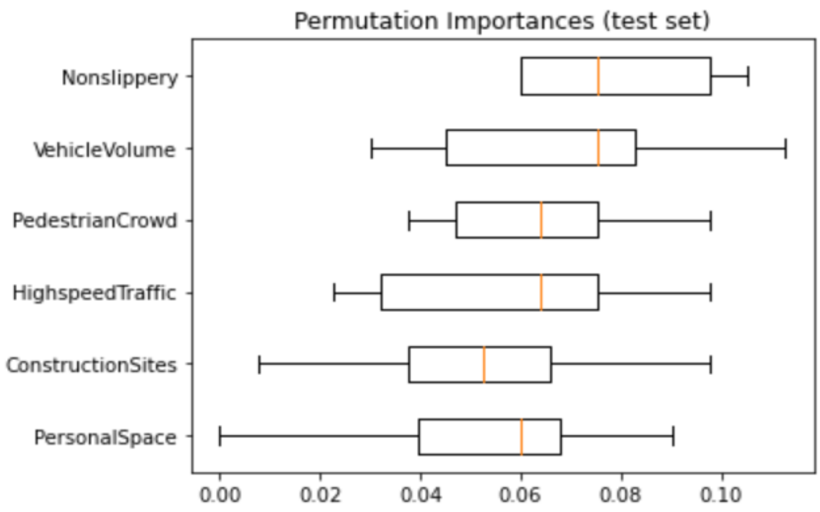
Figure 4. Confusion matrix and Important metrics of the Naïve Bayes model

```
Baseline Metrics for dataset: postcovid is as follows
accuracy: 0.5939849624060151
confusion matrix:
 [[16  3  2  0]
 [ 6 16 19  2]
 [ 5  8 37  5]
 [ 0  0  4 10]]
              precision    recall  f1-score   support

           0       0.59      0.76      0.67        21
           1       0.59      0.37      0.46        43
           2       0.60      0.67      0.63        55
           3       0.59      0.71      0.65        14

    accuracy                           0.59       133
   macro avg       0.59      0.63      0.60       133
weighted avg       0.59      0.59      0.58       133
```

The support column in Figure 4 indicates that out of the 133 test data samples, nearly 85% of the pedestrian responses were related to PLOS class A, B, or C. The answers received for the poor condition of sidewalks, or PLOS D, were less than other classes. The misclassification in class C is not the limitation of the Naïve Bayes model but rather due to the imbalanced number of classes in the data available to build the model.

The importance of features in predicting the pedestrian level of service in this model can be explained using the permutation importance function in machine learning. It is a model-agnostic technique, representing the decrease in the model scores when a single feature is randomly shuffled. It mainly depends on the dataset used for developing the model, and it can be calculated based on the train data or test dataset.

Figure 5: Feature importance using the Permutation Importance technique on test data



From Figure 5, we can see that the features of Non-slippery footpath surface, Vehicle Volume, Pedestrian crowd, and High-speed traffic are the features that mainly influence the PLOS rating using the Naïve Bayes classifier model.

In machine learning models, the global interpretation can lead to the features being overlooked when understanding the feature contribution for a particular instance or an individual. Hence, local interpretation is required, and in this study, Local Interpretable Model Agnostic Explanations (LIME) have been used to explain the prediction for a particular instance.

Figure 6: LIME output to explain local feature contribution.

The above figure 6 shows the LIME outcome for record 5 in the dataset. The response variables are numbered 2 to 5 because rating 1 and 2 have been combined due to insufficient data for class 1. The model has a higher probability (57%) of predicting the level of service rating as C (3 in the figure) and the Pedestrian Crowd (0.15), Vehicle Volume (0.10), non-slippery footpath surface (0.07) and Construction sites (0.04) are the top four variables contributing to the rating

3. The features such as Pedestrian crowd, Construction sites, and High-speed traffic says the PLOS rating cannot be 3, whereas the ratings for features such as Vehicle Volume, Non-slippery, and Personal space say the PLOS rating is 3. This technique gives a clear picture of the contribution of different factors in determining the PLOS rating and which aspect needs to be improved to improve the comfort of pedestrians.

6. Conclusion and future research direction

This study used a machine learning algorithm to develop a model to predict the pedestrian level of service of sidewalks in cities. Different techniques such as multilinear regression, structural equation modelling, fuzzy linear regression, fuzzy neural network, logistic regression, and ordered probit model have been used to develop a PLOS model for pedestrian walking locations such as signalized intersections, crosswalks, and sidewalks. A novel Naïve Bayes approach for modelling the pedestrian level of service helps to understand if the pedestrian's comfort factors have changed since the social disturbances occurred. The variables used for modelling were the pedestrian crowd, personal space, footpath surface non-slippery, construction works, vehicle volume, and noise of high-speed traffic. These were all qualitative variables that had the rating of pedestrians.

The accuracy of predicting the PLOS using the Naïve Bayes Classifier was 60%. As the dataset had imbalanced classes, measures such as precision, recall, and f-score values have been calculated to check the model's performance. It was found that PLOS class A, B, and D have f1-score values around 0.65, while for class C it was slightly small due to a smaller number of observations corresponding to those groups and not due to the limitation of the model. Enough data must be collected for each category of LOS so that categories can be classified from A to E.

The Naïve Bayes classifier model has been explained using the Permutation Importance function, which ranked the variables in the order in which they influenced the PLOS. The surface of the footpath and Vehicle volume are the top influencing variables of the PLOS. The other variables include pedestrian crowds, the noise of high-speed traffic, construction sites, and personal space. The LIME technique has been used to study the local interpretation of the model in predicting the PLOS, and it is found that for each instance of prediction, the top influencing variables affecting the PLOS could change. This study could be extended by modelling using other machine learning algorithms and comparing their performance with the Naïve Bayes model. City planners and councils can use this technique to find the significant affecting factors of PLOS in a locality or region, which may vary from time to time based on the changes occurring in that area. Hence there is scope for further research to use more advanced techniques and features to model the pedestrian level of service in cities' sidewalks.

7. References

ALFONZO, M. A. 2005. To Walk or Not to Walk? The Hierarchy of Walking Needs. *Environment and Behavior,* 37**,** 808-836.

ASADI-SHEKARI, Z., MOEINADDINI, M. & ZALY SHAH, M. 2014. A pedestrian level of service method for evaluating and promoting walking facilities on campus streets. *Land Use Policy,* 38**,** 175-193.

ASHARI ET AL 2013. Performance Comparison between Naïve Bayes, Decision Tree and k-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool.

International Journal of Advanced Computer Science and Applications, Vol. 4, No. 11, 2013.

BAI, L., LIU, P., CHAN, C.-Y. & LI, Z. 2017. Estimating level of service of mid-block bicycle lanes considering mixed traffic flow. *Transportation Research Part A: Policy and Practice,* 101**,** 203-217.

BASBAS, S. E. A. Pedestrian Level of Service Assessment in an Area Close to an Under-Construction Metro Line in Thessaloniki, Greece. *Transportation Research Procedia 45 (2020): 95–102*.

BIVINA, G. R. & PARIDA, M. 2019. Modelling Perceived Pedestrian Level of Service of Sidewalks: A Structural Equation Approach. *Transport,* 34**,** 339-350.

BRAUN, R. R. A. M. F. R. 1978. Quantifying the benefits of separating pedesstrians and vehicles. *TRB, National Research Council, Washington D.C***,** 88-91.

CEPOLINA, E. M., MENICHINI, F. & GONZALEZ ROJAS, P. 2017. Pedestrian Level of Service: the Impact of Social Groups on Pedestrian Flow Characteristics. *International Journal of Sustainable Development and Planning,* 12**,** 839-848.

CEPOLINA, E. M., MENICHINI, F. & GONZALEZ ROJAS, P. 2018. Level of service of pedestrian facilities: Modelling human comfort perception in the evaluation of pedestrian behaviour patterns. *Transportation Research Part F: Traffic Psychology and Behaviour,* 58**,** 365-381.

CITY OF MELBOURNE. 2020. *Transport Strategy 2030, City of Melbourne* [Online]. Available: https://www.melbourne.vic.gov.au/SiteCollectionDocuments/transport-strategy-2030-city-of-melbourne.pdf [Accessed].

CZOGALLA, O. 2012. A Model for Pedestrian Routing in Walkable Networks. *IFAC Proceedings Volumes,* 45**,** 303-308.

DUMBAUGH, E. & LI, W. 2010. Designing for the Safety of Pedestrians, Cyclists, and Motorists in Urban Environments. *Journal of the American Planning Association,* 77**,** 69-88.

FRUIN, J. 1971. *Pedestrian Planning and Design*

GEORGIOU, A., SKOUFAS, A. & BASBAS, S. 2021. Perceived pedestrian level of service in an urban central network: The case of a medium size Greek city. *Case Studies on Transport Policy,* 9**,** 889-905.

HASAN, T., SIDDIQUE, A., HADIUZZAMAN, M. & MUSABBIR, S. R. 2015. Determining the Most Suitable Pedestrian Level of Service Method for Dhaka City, Bangladesh, through a Synthesis of Measurements. *Transportation Research Record: Journal of the Transportation Research Board,* 2519**,** 104-115.

HCM 2000. Highway Capacity Manual. *Transportation Research Board,* Washington, D.C.

HCM 2010. Highway Capacity Manual. *Transportation Research Board,* Washington, D.C.

JAHAN, M. I., MAZUMDAR, A. A. B., HADIUZZAMAN, M., MASHRUR, S. M. & MURSHED, M. N. 2020. Analyzing Service Quality of Pedestrian Sidewalks under Mixed Traffic Condition Considering Latent Variables. *Journal of Urban Planning and Development,* 146.

JENA, S., ATMAKURI, P. & BHUYAN, P. K. 2017. Evaluating Service Criteria of Urban Streets in Developing Countries Based on Road Users' Perception. *Transportation in Developing Economies,* 4.

JIA, X., FELICIANI, C., MURAKAMI, H., NAGAHAMA, A., YANAGISAWA, D. & NISHINARI, K. 2022. Revisiting the level-of-service framework for pedestrian comfortability: Velocity depicts more accurate perceived congestion than local density. *Transportation Research Part F: Traffic Psychology and Behaviour,* 87**,** 403-425.

KADALI, B. R. & VEDAGIRI, P. 2015. Evaluation of pedestrian crosswalk level of service (LOS) in perspective of type of land-use. *Transportation Research Part A: Policy and Practice,* 73**,** 113-124.

KANG, K. & LEE, K. 2012. Development of a bicycle level of service model from the user's perspective %J KSCE journal of civil engineering. 16**,** 1032-1039.

KIM, S., CHOI, J., KIM, S. & TAY, R. 2014. Personal space, evasive movement and pedestrian level of service. *Journal of Advanced Transportation,* 48**,** 673-684.

LANDIS, B. W., VATTIKUTI, V. R., OTTENBERG, R. M., MCLEOD, D. S., & GUTTENPLAN, M 2001. Modelling the roadside walking environment: Pedestrian level of service. *Transportation Research Record: Journal of the Transportation Research Board*, 82-88.

LI, Q., WANG, Z., YANG, J. & WANG, J. 2005. Pedestrian delay estimation at signalized intersections in developing cities. *Transportation Research Part A: Policy and Practice,* 39, 61-73.

LI, Z., WANG, W., LIU, P. & RAGLAND, D. R. 2012. Physical environments influencing bicyclists' perception of comfort on separated and on-street bicycle facilities. *Transportation Research Part D: Transport and Environment,* 17, 256-261.

MARISAMYNATHAN, S. & VEDAGIRI, P. 2019. Pedestrian Perception-based Level-of-service Model at Signalized Intersection Crosswalks. *Journal of Modern Transportation*, 1-16.

MORAN, M. R., RODRÍGUEZ, D. A. & CORBURN, J. 2018. Examining the role of trip destination and neighborhood attributes in shaping environmental influences on children's route choice. *Transportation Research Part D: Transport and Environment,* 65, 63-81.

SAHANI, R. & BHUYAN, P. K. 2014. Pedestrian Level of Service Criteria for Urban Off-Street Facilities in Mid-Sized Cities. *Transport,* 32, 221-232.

SAHANI, R., OJHA, A. & BHUYAN, P. K. 2017. Service levels of sidewalks for pedestrians under mixed traffic environment using Genetic Programming clustering. *KSCE Journal of Civil Engineering,* 21, 2879-2887.

SGS ECONOMICS AND PLANNING. 2014. *CBD Pedestrian Analysis, Technical report, City of Melbourne* [Online]. Available: https://www.sgsep.com.au/assets/main/SGS-Economics-and-Planning-CBD-Pedestrian-Analysis.pdf [Accessed].

SHATU, F., YIGITCANLAR, T. & BUNKER, J. 2019. Objective vs. subjective measures of street environments in pedestrian route choice behaviour: Discrepancy and correlates of non-concordance. *Transportation Research Part A: Policy and Practice,* 126, 1-23.

SHETH ET AL 2022. A Comparative analysis of machine learning algorithms for classification purpose. *The 4th International Conference on Innovative Data Communication Technology and Application.*

SMITH, K. R., ZICK, C. D., KOWALESKI-JONES, L., BROWN, B. B., FAN, J. X. & YAMADA, I. 2011. Effects of neighborhood walkability on healthy weight: assessing selection and causal influences. *Soc Sci Res,* 40, 1445-55.

TALAVERA-GARCIA, R. & PEREZ-CAMPANA, R. 2021. Applying a Pedestrian Level of Service in the Context of Social Distancing: The Case of the City of Madrid. *Int J Environ Res Public Health,* 18.

VALLEJO-BORDA, J. A., CANTILLO, V. & RODRIGUEZ-VALENCIA, A. 2020. A perception-based cognitive map of the pedestrian perceived quality of service on urban sidewalks. *Transportation Research Part F: Traffic Psychology and Behaviour,* 73, 107-118.

VISTA. 2018. *Victorian Integrated Survey of Travel and Activity, Department of Transport* [Online]. Available: https://public.tableau.com/profile/vista#!/vizhome/VISTA-JourneytoeducationAccess/JTE-methodoftravel [Accessed].

YE, X., CHEN, J., JIANG, G. & YAN, X. 2015. Modeling Pedestrian Level of Service at Signalized Intersection Crosswalks under Mixed Traffic Conditions. *Transportation Research Record: Journal of the Transportation Research Board,* 2512, 46-55.

ZHAO, L., BIAN, Y., RONG, J., LIU, X. & SHU, S. 2016. Evaluation method for pedestrian level of service on sidewalks based on fuzzy neural network model. *Journal of Intelligent & Fuzzy Systems,* 30, 2905-2913.