

Article

A Comprehensive Analysis of Clustering Public Utility Bus Passenger's Behavior during the COVID-19 Pandemic: Utilization of Machine Learning with Metaheuristic Algorithm

Maela Madel L. Cahigas ^{1,2,*} , Ferani E. Zulvia ¹, Ardvin Kester S. Ong ¹  and Yogi Tri Prasetyo ^{3,4} 

¹ School of Industrial Engineering and Engineering Management, Mapúa University, 658 Muralla St., Intramuros, Manila 1002, Philippines

² School of Graduate Studies, Mapúa University, 658 Muralla St., Intramuros, Manila 1002, Philippines

³ International Bachelor Program in Engineering, Yuan Ze University, 135 Yuan-Tung Rd., Chung-Li 32003, Taiwan

⁴ Department of Industrial Engineering and Management, Yuan Ze University, 135 Yuan-Tung Rd., Chung-Li 32003, Taiwan

* Correspondence: mmlcahigas@mapua.edu.ph; Tel.: +63-2-8247-5000 (ext. 6202)

Abstract: Public utility bus (PUB) systems and passenger behaviors drastically changed during the COVID-19 pandemic. This study assessed the clustered behavior of 505 PUB passengers using feature selection, K-means clustering, and particle swarm optimization (PSO). The wrapper method was seen to be the best among the six feature selection techniques through recursive feature selection with a 90% training set and a 10% testing set. It was revealed that this technique produced 26 optimal feature subsets. These features were then fed into K-means clustering and PSO to find PUB passengers' clusters. The algorithm was tested using 12 different parameter settings to find the best outcome. As a result, the optimal parameter combination produced 23 clusters. Utilizing the Pareto analysis, the study only considered the vital clusters. Specifically, five vital clusters were found to have comprehensive similarities in demographics and feature responses. The PUB stakeholders could use the cluster findings as a benchmark to improve the current system.

Keywords: public utility bus (PUB); passenger; feature selection; k-means clustering; particle swarm optimization (PSO)



Citation: Cahigas, M.M.L.; Zulvia, F.E.; Ong, A.K.S.; Prasetyo, Y.T. A Comprehensive Analysis of Clustering Public Utility Bus Passenger's Behavior during the COVID-19 Pandemic: Utilization of Machine Learning with Metaheuristic Algorithm. *Sustainability* **2023**, *15*, 7410. <https://doi.org/10.3390/su15097410>

Academic Editor: Mariano Gallo

Received: 28 March 2023

Revised: 22 April 2023

Accepted: 26 April 2023

Published: 29 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Public utility bus (PUB) transports passengers from one city to another. It also provides employment and resources that benefit everyone worldwide. Due to rapid urban growth, the demand for public transportation has increased. Hence, Mayo and Taboada [1] recommended focusing on the public transportation system's safety, affordability, accessibility, and sustainability. In the Philippines, commuting passengers frequently use PUB for public transportation [1,2]. Additionally, PUB is used to transport goods and products, which positively influences economic development [1]. According to Dela Peña [3], approximately 8% of the Metro Manila population (one of the largest urban areas in the Philippines) uses PUB daily. Metro Manila comprises 12.4% of the total population in the Philippines [4]. Hence, 87.6% of Filipinos reside outside Metro Manila. They live in rural and urban-rural mixed areas and residents of these cities use PUBs frequently [1]. Thus, the Philippines needs consistent and effective PUB operations regardless of the population density.

Since the coronavirus disease (COVID-19) started in the Philippines, public transportation restrictions have been implemented. By the end of 2021, PUB was only allowed to operate at 75% capacity [5]. Before entering the vehicle, PUB passengers were required to undergo a temperature check and alcohol disinfection. PUB passengers were also mandated to wear face masks and shields. Moreover, standing passengers were not allowed. Since

several factors affected the Philippines' PUB system during the COVID-19 pandemic, Cahigas et al. [6] identified the most relevant features, the researchers evaluated the following variables with underlying features: accessibility, safety, economic benefit, crisis management, trust, attitude, subjective norm, perceived behavioral control, and intention to use. The corresponding features are fed into machine learning and metaheuristic algorithms.

One of the machine learning techniques utilized in the study is known as feature selection or feature engineering. It distinguishes the most important features and eliminates unimportant features [7,8]. The important features are utilized to increase data prediction accuracy [9]. Through feature selection, data are transformed into a more logical set of information. This current research study used feature selection to find the most suitable features affecting the behavior of PUB passengers. Since feature selection consists of several extraction techniques, this study focused on six feature selection techniques. Specifically, the feature selection techniques were (1) filter-correlation, (2) filter-univariate selection, (3) wrapper-backward elimination, (4) recursive feature elimination (RFE), (5) embedded-LASSO, and (6) stepwise regression. The current study used these feature selection techniques because of their high prediction/accuracy rates [9–11]. After feature selection, the researchers utilized another machine learning technique known as K-means clustering.

K-means clustering is a well-known clustering algorithm that groups data sets into different clusters by applying ordinary sample mean with asymptotic behavior [12–14]. It is an iterative process that generates different clusters for every initialization [12,13]. Researchers commonly use K-means clustering because of its simplicity and efficiency [13,14]. However, the foundation of K-means clustering is weak because it starts with a random initial centroid. Since the current study utilizes actual survey responses with several features/attributes, random initial centroid might negatively affect the succeeding centroids of features. Therefore, this study applied a metaheuristic algorithm to find the optimal initial centroid of K-means.

Particle swarm optimization (PSO) is a metaheuristic approach that analyzes the movement of particles [15,16]. Each particle generates and keeps new locations until the parameters are met [15]. The current study used PSO because it has memory and keeps different sets of solution and fitness values until the optimal solution is met, which other metaheuristic approaches lack.

Despite available studies, researchers have not yet explored the combination of multiple-feature selection techniques, K-means clustering, and PSO in PUB passenger behavior. Although some studies compared at least three feature selection techniques, the results were limited to distinguishing the advantages and disadvantages of feature selection techniques [10,17,18]. These feature selection studies failed to comprehensively analyze machine learning and metaheuristic algorithms. The studies of Anderson [19] and Fotouhi and Montazeri-Gh [20] applied K-means clustering to improve the transport system but failed to discuss the passengers' behavior. Furthermore, researchers commonly used PSO to optimize the number of passengers and develop transport routes [19–24]. However, these PSO-related studies excluded the importance of analyzing passengers' behavior which could also affect ridership and movement of public transportation modes. Therefore, there is a lack of studies utilizing feature selection, K-means clustering, and PSO to learn and properly assess the PUB passengers' behavior.

Following the research gap, the subsequent research questions are developed: (1) How can feature selection, K-means, and PSO generate the optimal clusters? (2) What are the optimal PUB passengers clusters? and (3) What are the characteristics of the optimal clusters?

The researchers aimed to find the optimal PUB passenger clusters based on their COVID-19 pandemic behavior by integrating feature selection, K-means clustering, and PSO. This study proposed a novel method by eliminating traditional statistical procedures. It employed six feature selection techniques (filter-correlation, filter-univariate selection, wrapper-backward elimination, recursive feature elimination, embedded-LASSO, and stepwise regression) to find the most important features. In addition, K-means clustering and PSO were combined for clustering purposes. Therefore, the purpose of integrating feature

selection, K-means clustering, and PSO stemmed from the identification of significant features to focus on essential PUB passenger-related factors during the era of COVID-19. Determining critical factors allows stakeholders to maximize their resources upon the materialization of practical implications. The researchers also combined the methods to analyze Filipino passengers' demographic features. Since the study would receive real-time survey responses, the K-means clustering and PSO could compartmentalize PUB passenger profiles. Overall, the integrated methods reflected the actual scenarios affecting Philippine public transportation, the economy, and PUB passengers' behavior.

The study's contributions are (1) determining the appropriate PUB passengers' clusters during the COVID-19 pandemic. It is necessary to group the passengers according to their unique characteristics to easily find their similarities. (2) The government, PUB companies, and PUB drivers can adapt the optimal clusters in developing a more efficient PUB system. Every region caters to different demographic characteristics and PUB passengers' needs, and the results are anticipated to help the PUB stakeholders create comprehensive protocols. Since quarantine restrictions limit the mobility of passengers, passengers have a high demand for quality and efficient public transport services. (3) The application of feature selection, K-means clustering, and PSO to the public transportation and academic sector. To the best of the researchers' knowledge, researchers have not published this combined approach yet. Thus, presenting a study that employs feature selection, K-means cluster, and PSO reveals the possibility of exploring the three methods' different uses.

2. Literature Review

This section presents the reviews of relevant studies over the years. The first section tackles nine features affecting PUB passengers' behaviors. The second section reveals the six feature selection techniques. The third subsection discusses the principles and past studies about K-means clustering. Lastly, particle swarm optimization (PSO) is elaborated by showing its theories and applications in transportation and related studies.

2.1. Features Affecting the PUB Passengers' Behaviors

Accessibility pertains to the ease of accessing a public transportation mode, specifically a public utility bus (PUB) [25,26]. PUB is necessary for Filipinos because it is one of the Philippines' primary public transportation modes [27]. It is imperative to have accessible PUB as it transports passengers and goods. Thus, the government needs to continuously improve PUB stop locations, routes, and ticket purchase systems since these factors affect passenger satisfaction. Chen et al. [28] added that the number of PUBs and drivers must be carefully analyzed alongside the passenger demands. Tiglao et al. [27] emphasized that accessibility in the Philippines significantly affects passengers' perceptions. In Vietnam, PUB passenger ridership decreased because other public transportation modes provided convenience and innovative systems [29]. Furthermore, inconvenience and inaccessibility result in poor public transportation modes [30]. Hence, accessibility is a way to determine areas of improvement in public transport planning [31].

Safety has a direct relationship with passengers, drivers, and the environment. It deals with passenger security and riding comfort [25,27]. The driver's attitude and driving capability are also considered [27,32]. Moreover, the environment describes cleanliness, onboarding and offboarding systems, low crime rate, and accident-free ride [27,32,33]. All these factors contribute to public safety. One study stated that some designated stops in the Philippines are poorly maintained due to the lack of road lights and signs [27]. It is recommended to prioritize road safety because most passengers feel satisfied when their safety is guaranteed [33]. Thus, a past study ran multiple simulations to ensure appropriate bus boarding and alighting design [34]. They proposed four strategies and found two significant strategies. First, they recommended the transition from disembarkation to embarkation process to improve bus disembarkation efficiency [34]. Second, the redesigning of lower and rear doors increased boarding and alighting efficiency by 40% to 50% [34]. These

heightened approaches toward public transportation safety would increase the country's economic growth.

Economic benefit describes the person's ability to afford a PUB ride for essential needs (work, healthcare, school, etc.) [31]. PUB rides should be affordable and efficient to support the country's economy [26]. A developed public transportation system means positive social benefits and urban living [22]. Public transportation system development is associated with an employment increase, and many citizens benefit from employment. In return, cash flow circulates perpetually. Rasoolimanesh et al. [35] concluded that economic benefit positively affects passengers' perceptions and infrastructure development. In addition, Atombo and Wemegah [26] also noted that passengers must be allowed to access various public transport modes because this affects passenger satisfaction. An economic benefit is not only measured through affordability but also by the saving of time [31].

Crisis management refers to the COVID-19 preventive measures implemented by the government and healthcare organizations [36]. The past study stated that crisis management plays a primary role in recovering from passengers' perceived risks of using public transportation modes. In the Philippines, passengers, drivers, and conductors must follow COVID-19 preventive measures [37]. Before boarding PUBs, all individuals must wear face masks, pass the temperature check, and disinfect their hands. PUBs can only operate at 75% capacity, and standing passengers are not allowed [5]. Similar to Indonesia's situation, people preferred following COVID-19 protocols such as avoiding large crowds, using face masks, and participating in vaccination programs [38]. In Australia, citizens have negative perceptions of using public transport modes [39]. Although Australia had minimal COVID-19 cases in 2021, the study of Thomas et al. [39] revealed that the number of COVID-19 cases did not positively affect the passenger's public transportation behavior. Instead, the presence of COVID-19 makes people fear and be more vigilant. Bus passengers were inclined to avoid crowds and limit human interaction [34,40]. People feel threatened by sharing public space and prefer using private transport modes [37,40].

Trust describes the reliability of a specific action or object [41]. It may apply to behavioral relationships between people and dependency on a particular object. Trust is measured by knowing a passenger's confidence in riding PUB during the COVID-19 pandemic [42]. Moreover, PUB's physical condition matters (cleanliness, mechanical parts, etc.) [25]. Another study confirmed that trust strongly influenced passengers' attitudes toward using public transportation modes [41]. Trust is also a component of the PUB's service atmosphere. It coincides with the PUB's layout and mechanics, which were found the strongest predictors influencing passengers' usage of buses [29]. Overall, trust is a long-term insight, resulting in a highly positive or negative influence [43].

Attitude defines the person's behavior to perform a behavior [44]. It involves positive and negative emotions affecting a person's judgment [45]. Borhan et al. [41] connected attitude in choosing a transportation mode in the transportation sector. People have different preferences and may prefer a certain transportation mode over other options. Nowadays, passengers also consider the effects of COVID-19 when riding any public transportation mode. Lee et al. [46] discovered a significant relationship between passengers' attitudes and willingness to use public transportation during the COVID-19 pandemic. They confirmed that the presence of COVID-19 made people feel reluctant to ride public transportation modes. Similarly, non-regular users of public transportation in Australia have a negative travel attitude due to COVID-19 [39]. Since public transportation is a shared public space, people are more willing to explore other travel modes, such as cycling and walking [47]. Thus, attitude helps understand behavioral changes, which may result in either favorable or unfavorable evaluations [44].

In addition, factors affecting individual behavior under the theory of planned behavior affect overall intentions. The subjective norm is the resulting behavior of an individual based on social pressure [44,47]. Social pressure occurs due to the perceptions of family members, friends, and acquaintances [47]. People share experiences and principles and tend to follow the group's social norms for consistency [48]. They tend to follow the

majority to mitigate conflicts. As a resulting behavior, people are inclined to seek social approval because they dislike being punished and abandoned [49]. Meanwhile, others use social pressure to contribute to goodness and encourage positivity [50].

Perceived behavioral control is the ability of a person to perform the intended behavior [44]. In this study, perceived behavioral control pertains to the ease or difficulty of riding a PUB. Perceived behavioral control is affected by several factors, including the environment, available resources, and psychological interests [44]. Thus, it directly impacts the behavioral changes of individuals [47]. A study was conducted in Qatar to determine the behavior of citizens in riding public transportation [51]. Past research disclosed that the citizens' perceived behavioral control is important in considering public transport for work purposes. However, there is a difference between the perceptions of the two genders. Although perceived behavioral control is important, women have less perceived behavioral control than men. This instance is associated with the culture and norms adopted by the citizens of Qatar. Gao et al. [52] also evaluated the perceptions of children and parents about using public transport. Perceived behavioral control directly impacts students' commuting behavior, but it has a lesser effect because most parents drop off and pick up students for safety purposes. These past studies supported that perceived behavioral control is essential for determining passenger behavior.

2.2. Feature Selection

Feature selection is a data filtering process that eliminates unimportant features and retains the important features [7–9]. It has several benefits aiming to improve the overall data. The accuracy rate increases and computation time decreases by reducing the number of features [7,53]. Additionally, the data complexity is reduced, making the data easier to understand [53]. By removing the noise, overfitting is reduced to ensure data consistency [53]. Generally, feature selection improves data structure by increasing its prediction performance.

Previous studies used feature selection in the transportation sector for different purposes. Many researchers applied feature selection in optimizing transportation routes. Xiong et al. [54] optimized the emergency vehicle flow for elderly patients. These researchers integrated feature selection into the decision tree model. They tested the combination of several features by setting appropriate training and test set parameters. In addition, Liu et al. [55] used two feature selection techniques to improve the multi-modal transportation system in China. First, they proposed a new feature selection technique by considering the demographic profiles, travel modes, geographic locations, and time. Second, an embedded method was combined with a bipartite graph to visualize transportation networks.

Furthermore, Soares et al. [56] applied feature engineering to the transportation and technology industries. The aforementioned study enhanced the location accuracy of travel modes and reduced unnecessary transportation costs through feature selection and other automated machine learning. Additionally, Rodríguez-Sanz et al. [57] predicted the queuing behavior of airport passengers by applying machine learning algorithms. They enhanced feature selection results by employing a random forest. Another study predicted the passengers' transportation mode choice among car, bus, train, or walking [58]. The past study compared six feature selection techniques and applied benchmarking and compressive sensing-based feature selection algorithms to find the most optimal transport mode choice.

Researchers have developed many feature selection techniques to find the most significant features in data. Some researchers adapt the existing techniques, while others generate a more enhanced approach. This current study used feature selection techniques with a supported accuracy rate [9–11]. The researchers also filtered the suitable feature selection techniques by considering the study's data structure. Specifically, this study used six feature selection techniques (filter-correlation, filter-univariate selection, wrapper-backward elimination, recursive feature elimination, embedded-LASSO, and stepwise regression).

2.2.1. Filter Method—Correlation

The filter method is the most straightforward feature selection technique commonly used by many researchers [10,59]. It has many variations, and correlation is among the standard approaches. This approach measures the importance of features by calculating the correlation scores between features and dependent variables [17,18]. Afterward, a ranking criterion is applied, and the high scorers are considered the most important features. The filter method involves a statistical-based threshold to eliminate unimportant features [17,53]. Thus, it reduces overfitting, which helps produce optimal solutions [53]. Moreover, the filter method using correlation is unbiased with any classifiers because it evaluates features individually [60]. Although the method performs individual assessment, the computation is fast and efficient [59,61]

2.2.2. Filter Method—Univariate Selection

Univariate selection is another filter approach proposed by researchers that calculates the scores of important features. It employs analysis of variance (ANOVA) to predict the important features based on the target variables [17,18]. ANOVA uses an F-test with a corresponding p -value that tests each feature's significance [18]. Moreover, this method uses a specific parameter to separate important and unimportant features [61]. The univariate selection ensures that the features fit the proposed model based on ANOVA principles.

2.2.3. Wrapper Method—Backward Elimination

The wrapper method finds the optimal solution among the multiple subsets containing different features [61,62]. It can be processed through forward selection, backward elimination, and recursive feature elimination (RFE). In forward selection, features are continuously added until the performance model becomes constant [62]. Backward elimination is the opposite process of forward selection. All features are automatic inputs of the model and are removed one by one until the model achieves optimal performance [62]. Meanwhile, RFE is discussed in the next subsection. In this study, the forward selection is excluded from the analysis because it produces a lesser accuracy than backward elimination and RFE.

Through backward elimination, the order of features within each subset is automatically arranged according to the respective importance [10,18]. This approach uses classifiers to train the combination of all subsets [59,61]. All subsets are compared, and the subset with the most significant p -value is considered the best result. Many researchers have applied the wrapper method because of its high accuracy performance [18,53]. Since the wrapper method is more extensive than the filter method, it results in greater processing time and heavier computation.

2.2.4. Wrapper Method—Recursive Feature Elimination

Recursive feature elimination (RFE) trains and compares different sets of feature combinations [18]. It is a continuous process of eliminating weak features until the RFE score hits the highest prediction rate [53,55]. RFE scores range from 0 to 1 to find the most important features. The RFE score is determined using a machine algorithm of Jupyter Notebook, considering the combination of features and the number of evaluated features [53]. Furthermore, Liu et al. [55] recommended RFE in big data structure and industrial problems, which applies to the current study.

2.2.5. Embedded Method—LASSO

The embedded method identifies feature subsets that undergo a learning system to evaluate feature importance [59]. This approach requires a predictive model to train the data [17]. In this study, the Least Absolute Shrinkage and Selection Operator (LASSO) model is the learning algorithm employed in the predictive model. In the LASSO model, coefficients of unimportant feature combinations are set to zero, and non-zero values are retained [63]. LASSO applies linear regression principles with L1 regularization [18]. L1

regularization is also known as the subset selection through least squares regression [63]. The embedded method is almost similar to the wrapper method, but the embedded trains features without iteration while the wrapper considers the iteration process.

2.2.6. Stepwise Regression

Stepwise regression is a feature selection method that continually adds significant features and removes insignificant features to/from the subset [9]. It is a statistical approach used to predict the relationship between variables of linear models. Moreover, it evaluates the effect of each feature/variable on the model; features are eliminated or added depending on the statistical significance values [64]. Since stepwise regression reduces dimensionality, all features undergo refinement. Hence, it can easily find the best predictor in every subset [7]. Overall, stepwise regression increases the correlation between dependent and independent variables.

2.3. K-Means Clustering

K-means is a partition-based algorithm that assigns data points to clusters [12,13]. Afterward, data points in similar clusters generate centroids [12–14]. The centroid is also known as the mean of the data points within a cluster, but the initial centroid in K-means is randomly generated. Overall, K-means aims to group data points into the optimal number of clusters and minimize the sum of squared error (SSE). SSE should be minimized because it measures the error between data points in a cluster and the nearest cluster [13]. If the SSE has a lower value, clusters are more compact. It also signifies that data points are appropriately grouped or clustered. Since K-means is an iterative process, old and new centroids are constantly compared according to tolerance limits [13]. This approach makes K-means a well-built clustering algorithm because it automatically detects all clusters with appropriate centroids. In this study, the important features from each feature selection technique undergo K-means clustering.

Many researchers used K-means clustering in the transportation sector. Fotouhi and Montazeri-Gh [20] developed driving patterns in Tehran, Iran. They analyzed the driver's average speed and idle time through K-means. As a result, they generated clusters with corresponding traffic conditions. Anderson [19] identified clusters describing road accidents in London, England, through attributes and types of road collisions. The past study's attributes are relevant to roads, transportation, and infrastructure.

Meanwhile, road collisions pertain to the characteristics of accidents. Aside from road characteristics, some studies also evaluate passenger behavior through K-means clustering. Eltvéd et al. [65] identified the impact of a 3-month rail line closure in Greater Copenhagen, Denmark. The researchers applied the K-means algorithm to identify passengers' behavior before and after the rail station closure. Li et al. [66] combined K-means clustering and operation research approaches to determine the effectiveness of passenger satisfaction assessment on rail transits in Shanghai, China. The researchers used a modified K-means method to find appropriate clusters of passengers dependent on the rail transit lines. Additionally, Shen et al. [67] used K-means to cluster passengers according to travel time and distance. These researchers customized a bus boarding system to determine the appropriate bus stop and destination points.

2.4. Particle Swarm Optimization

Particle swarm optimization (PSO) is derived from the swarming behavior of organisms, such as bird flocking and fish schooling [15,16]. Particles or organisms move to a specific location without instructions, enhancing the particles' abilities to find the best fitness points [68]. Fitness pertains to the optimal solution based on the PSO algorithm's objective. In this study, PSO reflects the movement of data points within the clusters. The movement consists of changes in location and velocity, inclined towards the particle's best experience (*pbest*) and global best experience (*gbest*). *Pbest* is the best fitness

achieved by a particle, while *gbest* is the best fitness achieved by any particle from the entire population [14,15].

Past studies utilized PSO to improve public transportation routes. Specifically, Zhong et al. [22] proposed an improved PSO algorithm to optimize bus transit routes in Dalian, China. The study aimed to serve more bus passengers to increase bus efficiency and meet passenger demands. Kechagiopoulos and Beligiannis [21] focused on the road problems affecting public and private transportation users. The PSO inputs are travel time, demand, and road network structure. Interestingly, Peng et al. [69] applied the multiple-objective PSO-crowding distance approach to the structural design of railroad vehicles. They reduced the crash impacts of rail accidents by integrating the proposed PSO into other optimization methods.

Furthermore, Li et al. [24] evaluated China's three primary high-speed rail networks. The optimal high-speed rail routes and passenger assignments for each rail section are identified through PSO. Moreover, Xiao et al. [70] integrated PSO and neural network to increase travel mode detection accuracy in Global Positioning System (GPS). They supported that PSO enhanced the neural network's classification performance.

Table 1 outlines the aforementioned studies from Sections 2.2–2.4. All these studies are relevant to the public transportation system and passenger behavior. Most importantly, the researchers utilized machine learning and metaheuristic algorithms.

Table 1. Summary of relevant studies.

Author(s)	Year	Country	Public Transport Mode	Purpose of the Study	Methodology
Xiong et al. [54]	2021	USA	Emergency vehicle	Optimization of emergency vehicle flow for elderly patients	Feature selection (general) and decision tree model
Liu et al. [55]	2021	China	Bus, train, car, taxi, and bicycle	Improvement of the multi-modal transportation system in China	Feature selection (proposed and embedded method), bipartite graph, and post-processing algorithm (proposed)
Soares et al. [56]	2019	USA	Car, train, and bus	Enhancement of the travel modes' location accuracy and reduction of unnecessary costs	Feature engineering (principal component analysis) and automated machine learning (AutoSklearn)
Rodríguez-Sanz et al. [57]	2021	Spain	Aviation	Prediction of airport passengers' queuing behavior at check-in desks and security controls	Random forest, feature analysis (proposed), and machine learning (simulation)
Yang and Ma [58]	2019	Sydney, Australia	Car, train, and bus	The identification of the most optimal public transport mode choice	Feature selection (Laplacian score, ReliefF, SimbaLinear, mutual information quotient, Genetic Programming, Dynamic Relevance, and Joint Mutual Information Maximization approach) and compressive sensing-based feature selection algorithm
Fotouhi and Montazeri-Gh [20]	2013	Tehran, Iran	Car	The enhancement of traffic conditions by evaluating driving patterns	Feature extraction (proposed) and K-means clustering
Anderson [19]	2009	London, England	General	Mitigation of road accidents to promote safety and security	Kernel Density Estimation and K-means clustering
Eltved et al. [65]	2021	Greater Copenhagen, Denmark	Train	Investigation of passenger behavior before and after the train station closure	K-means algorithm
Li et al. [66]	2021	Train	Shanghai, China	The effectiveness of passenger satisfaction assessment on train transit lines	K-means clustering and operation research (double hierarchy hesitant linguistic term sets (DHHLTSs) and alternative queuing method (AQM))
Shen et al. [67]	2021	China	Bus	Customization of the bus boarding system to determine the appropriate bus stop and destination points	K-means clustering
Zhong et al. [22]	2016	Dalian, China	Bus	Proposal of new bus transit routes to increase bus efficiency and meet passenger demands	Particle swarm optimization
Kechagiopoulos and Beligiannis [21]	2014	Switzerland	Bus	Centered on the road problems affecting users of public and private transportation modes	Particle swarm optimization

Table 1. Cont.

Author(s)	Year	Country	Public Transport Mode	Purpose of the Study	Methodology
Peng et al. [69]	2023	China	Train	Proposal of hybrid optimization decision systems in assessing the structural design of railroad vehicles	VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR), multiple objective particle swarm optimization-crowding distance (MOPSO-CD), and evolutionary algorithm and repetitive VIKOR
Li et al. [24]	2020	China	Train	Evaluation of three primary high-speed rail networks and passenger assignments for each rail section	Bi-level multi-objective mixed integer nonlinear programming model and particle swarm optimization
Xiao et al. [70]	2015	Shanghai, China	Bike, bus, and car	Increasing the travel mode detection accuracy in Global Positioning System	Particle swarm optimization and neural network

3. Methodology

The methodology section is divided into three parts: (1) data collection, (2) feature selection, and (3) a combination of K-means clustering and PSO algorithm. First, the data collection introduced the sampling technique and demographic profile of participants. Second, feature selection discussed the application of six feature selection techniques. The K-means clustering and PSO algorithm utilized the optimal output from feature selection. Finally, the K-means and PSO algorithms were implemented. Figure 1 demonstrates the proposed processes.

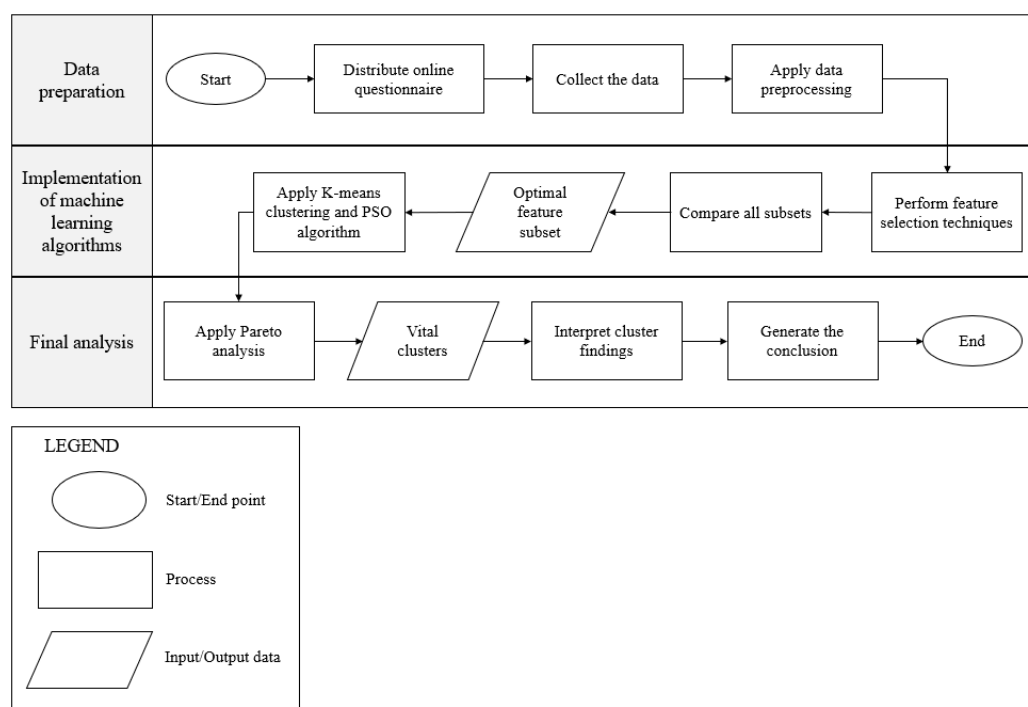


Figure 1. Methodology framework.

3.1. Data Collection and Preparation

This study employed a purposive sampling technique to collect data from the targeted participants. Based on the Yamane Taro Formula and optimal 5% sampling error, at least 399 participants should be collected [71]. This study exceeded the minimum required participants by gathering 505 PUB passengers residing in different cities of the Philippines. All the participants voluntarily participated in an online questionnaire hosted through Google Forms. The study's objectives and real-life contributions were posted on social media platforms (LinkedIn, Facebook, and Instagram), enticing the targeted participants to partake.

The researchers adopted the questionnaire from Cahigas et al. [6]. The questionnaire was utilized because it focused on PUB passenger behavior during the COVID-19 pandemic in the Philippines. The questionnaire contained 9 variables affecting PUB passenger behavior during the pandemic. These variables were accessibility, safety, economic benefit, crisis management, trust, attitude, subjective norm, perceived behavioral control, and intention to use—corresponding to 58 features. The indicator numbers in the study of Cahigas et al. [6] were identified as feature numbers (e.g., AC1, AC2, SA1, SA2, EB1, EB2, etc.) in the current study.

Table S1 in Supplementary Materials shows 58 features evaluated in the study. It was previously discussed that the features were collated by Cahigas et al. [6] in investigating PUB passengers' perceptions. The present study adopted the same set of features because all 58 features were supported at a 0.05 significance level. In addition, all features passed the 0.50 minimum factor loading value. Significance level and factor loading were important aspects of selecting the features because they validate logical relevance and consistency to the overall model.

Furthermore, the participants' demographic profiles were gathered from the questionnaire. Only a few male PUB passengers (17.43%) participated because they were more open to using an informal public transport mode. More than half of the participants were female (82.57%) because they preferred PUB, an economical formal type of public transportation in the Philippines. The common ages of participants ranged from 18 to 24 (52.87%) and 25 to 34 (29.11%) because these ages were less susceptible to severe COVID-19 effects. They could use PUB with fewer worries than other age groups. Another study testified that young passengers use a bus more frequently than older generations [28]. Most of the participants were students (35.05%), followed by unemployed individuals (29.70%), full-time employees (23.96%), self-employed individuals (6.73%), and part-time employees (4.56%). The employment status coincides with the education level. College students ranked first (32.08%), bachelor's degree holders ranked second (31.09%), and associate's degree holders ranked third (9.31%). Furthermore, at least 80% of the participants chose the smallest amount for allowances and expenses. Most of the participants allotted at most PHP 500 for daily allowance and at most PHP 200 for daily PUB expenses. This result suggested that individuals prioritize inexpensive public transport modes because the COVID-19 pandemic impacted the global economy and individual resources. Passengers were inclined to save money since they realized the importance of needs over wants. Lastly, only 22.57% of the participants possessed a private car at home, while 77.43% did not own any vehicles. Thus, more than half of the participants were forced to ride a public transport mode to perform their intended activities.

3.2. Feature Selection

This study assessed 58 features from 9 latent variables. Since several features affected the PUB passenger behavior during COVID-19, it was essential to filter the features through feature selection. This technique aimed to retain significant features producing a higher accuracy performance. Hence, insignificant features were eliminated from the model. Furthermore, the data set was classified as supervised learning because the class (perceived passenger behavior) was known. Figure 2 illustrates the universal process flow of the feature selection technique. Jupyter Notebook and SPSS 22 were utilized to perform this process.

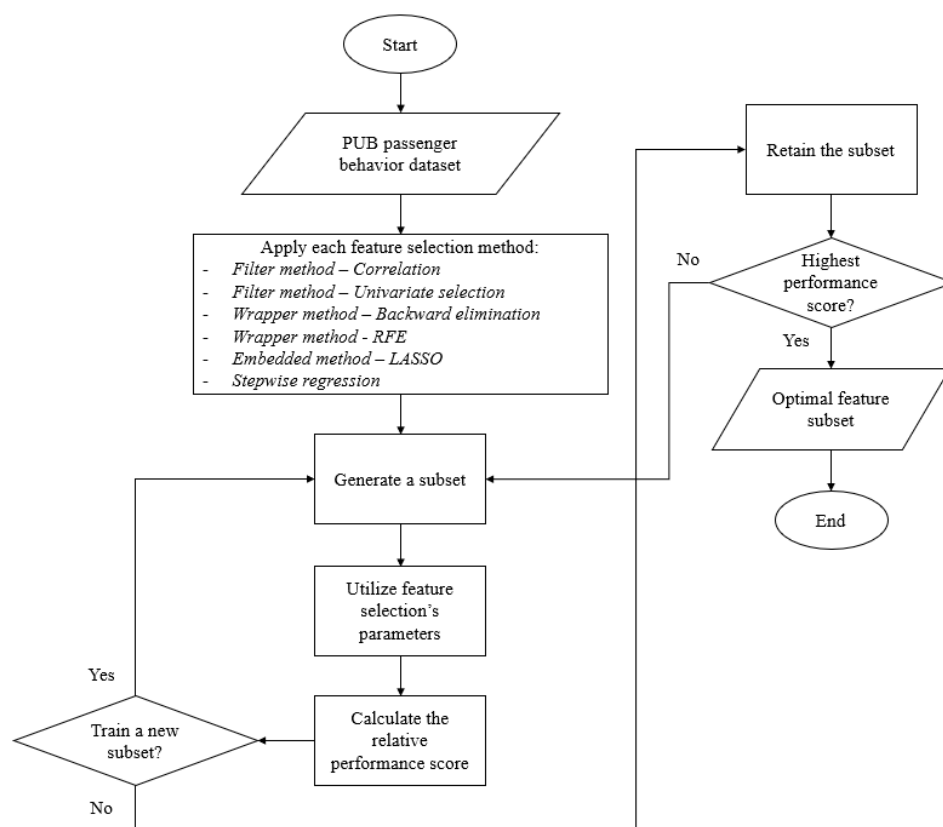


Figure 2. Feature selection method.

3.2.1. Filter Method—Correlation

This study assessed 58 features from 9 latent variables. The relationship between the 58 features and 1 dependent variable (perceived passenger behavior) was evaluated in the filter method using Pearson correlation. Features with a correlation value ranging from -1.00 to -0.50 and 0.50 to 1.00 were deemed acceptable [53]. A correlation value that is closer to 1 indicates a strong positive correlation, and values closer to -1 define a strong negative correlation. The equation can be found in Supplementary Materials (Feature Selection Equations).

A total of 23 features were generated in this step. Afterward, multicollinearity was verified by applying correlation to the pre-identified 23 features. Multicollinearity occurs when multiple independent variables are highly correlated [72]. All features paired and correlated with each other using Equation (S1) and underwent the second correlation step. Features with less than 0.50 correlation values passed the multicollinearity test. Meanwhile, features with at least 0.50 correlation values in the second correlation step were to be reevaluated. In the reevaluation process, the paired features' first-step correlation values were evaluated instead of the second value. A feature with a higher original correlation value would remain in the optimal feature subset list and the one with a lesser value would be eliminated.

3.2.2. Filter Method—Univariate Selection

The univariate selection was performed by combining SelectKBest and chi2 packages. This combination helps find the features with the highest scores. The researchers found the best features by eliminating features that failed to meet the minimum p -value significance level of 0.05. Its equation is presented in Supplementary Materials (Feature Selection Equations).

3.2.3. Wrapper Method—Backward Elimination

All the features were considered at the beginning of the wrapper method's backward elimination. This study utilized the Ordinary Least Squares (OLS) model to perform backward elimination. OLS predicts the regression between multiple features and targeted variables [51]. Through OLS, the model's performance was validated by removing the worst-performing feature one by one. The iteration process ended when all the p -values were below 0.05. Backward elimination's formula representation is located in Supplementary Materials (Feature Selection Equations).

3.2.4. Wrapper Method—Recursive Feature Elimination

Through the wrapper method's Recursive Feature Elimination (RFE), all 58 features were fed into the model. Unimportant features were removed, and the model was trained recursively according to the same training and testing sets (random state = 0). The researchers tried test sizes from 0.10 to 0.90 with an increment of 0.10 to find the highest RFE accuracy score. For each training and testing set, the optimal number of features was determined. Finally, the optimal number was keyed into the model to generate the optimal feature subset.

3.2.5. Wrapper Method—Embedded Method—LASSO

An embedded method applies LASSO regression to shrink the coefficient values [63]. LASSO regression penalized coefficients of unimportant features by setting them to 0. Thus, features with 0 coefficient values were eliminated from the model. Meanwhile, non-zero values were retained as they positively or negatively impacted the model. This overall process is expressed as a formula (Supplementary Materials: Feature Selection Equations).

3.2.6. Stepwise Regression

Stepwise regression was performed by using SPSS. SPSS tool's stepwise regression function tests the model's accuracy by continuously adding or removing features [9,64]. First, the data was imported into the SPSS. The dependent variable pertains to the data's class, and the 58 features were considered independent variables. Finally, the SPSS generated the features' corresponding coefficients and p -values. This study applied a 0.05 minimum significance level.

3.3. K-Means Clustering and PSO Algorithm

This study combined K-means clustering and the PSO algorithm to generate appropriate PUB passenger clusters. K-means clustering and the PSO algorithm were combined to improve the clustering of data points. Since K-means clustering had a weak initial centroid approach, PSO strengthened the generation solution. The data set was generated from the feature selection's optimal feature subsets. Afterward, the subsets underwent iteration until the optimal centroid was developed. Figure 3 displays the solution representation of the combined methods, where C is the generated centroid, D is the dimension from the optimal subset, and t is the number of iterations.

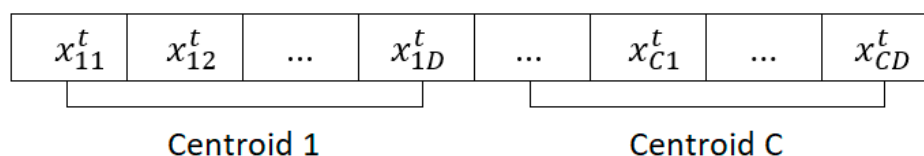


Figure 3. Solution representation.

Table 2 defines the parameters utilized in the model. Since PSO is a metaheuristic algorithm, predefined parameters were necessary. The number of particles (N), inertia weight (w), first acceleration coefficient ($c1$), and second acceleration coefficient ($c2$) underwent simulation as they produced different fitness values. Afterward, the calculated

cluster (M) value was rounded up to the nearest whole number (Supplementary Materials: PSO Initialization Equations). The number of iterations (T) was determined through the elbow method. Using the existing parameters, the iteration that produced stable results was chosen. This parameter was also supported by Kuo et al. [68]. Lastly, the optimal number of runs (r) was supported by Ryan et al. [73].

Table 2. Parameter settings of the model.

Parameter	Value	Note
N	20, 30, 40	References: Xu et al. [74]; Piotrowski et al. [75]
w	0.6, 0.9	References: Xu et al. [74]; Piotrowski et al. [75]
$c1$	1, 2	References: Xu et al. [74]; Piotrowski et al. [75]
$c2$	1, 2	References: Xu et al. [74]; Piotrowski et al. [75]
M	23	Supplementary Materials: PSO Initialization Equations
T	200	Measured through simulations and supported by Kuo et al. [68]
r	10	Reference: Ryan et al. [73]

After the parameters were established, data normalization followed. Data normalization organized the data set by cleaning and filtering unstructured data from the optimal feature selection subsets. The normalization process ensured that the data set was standardized. Data redundancy and errors were removed since the study employed a massive data size. Aside from data modification, the processing time was also reduced. The corresponding equations are displayed in Supplementary Materials (PSO Initialization Equations).

Once the data set was normalized, particles, velocities, and centroids were randomly generated. Additional conditions were incorporated into the values of particles because random particles ranged from 0 to 1 in decimals. To convert the particles to binary, random particles with less than 0.5 values were set to 0; otherwise, they were set to 1. Next, the sum of particles was calculated to generate the updated clusters and initial centroids. The cluster validity approaches used to find the optimal number of clusters were the Sum of Squares Error (SSE) or Sum of Squares Within (SSW), Sum of Squares Between (SSB), and Total Sum of Squares (TSS). The cluster validity formulas are as follows:

$$SSE = \sum_{j=1}^k \sum_{i=1}^{n_j} \left(X_{ij} - \bar{X}_j \right)^2 \quad (1)$$

$$SSB = \sum_{j=1}^k \left(\bar{X}_j - \bar{\bar{X}} \right)^2 \quad (2)$$

$$TSS = SSE + \left(\frac{1}{SSB} \right) \quad (3)$$

where i is the row, j is the column, and k and n are the respective data dependent on the row or column.

The fitness value pertains to the minimum value generated from the sum of TSS for each cluster. Fitness was needed to generate $pbest$ and $gbest$. The $pbest$ is the best fitness of each particle and the $gbest$ is the best fitness of the entire population. The particle's velocity is updated by applying Equation (4):

$$v_{ij}^t = wv_{ij}^{t-1} + r_1c_1^t \left(pbest_{ij}^t - x_{ij}^{t-1} \right) + r_2c_2^t \left(gbest_{ij}^t - x_{ij}^{t-1} \right) \quad (4)$$

where w is the inertia weight that controls the search space of velocity. It maintains the convergence and diversity of the algorithm. The $r1$ and $r2$ are random numbers from 0 to 1. Meanwhile, the $c1$ and $c2$ are fixed acceleration coefficients that influence velocity to

improve $pbest$ and $gbest$. After the particle's velocity was updated, the particle's location was updated through Equation (5):

$$x_{ij}^t = x_{ij}^{t-1} + v_{ij}^t \quad (5)$$

where x_{ij}^t is the *particle* (i, j) at iteration t . Meanwhile, v_{ij}^t is the updated particle's velocity.

Once the particle's velocity was updated, the process was repeated by adding the particles' conditions, then updating the number of clusters, updating the centroids, and applying cluster validity to find the new fitness value. The particle's location and velocity were continuously updated until the final $gbest$ met the stopping criteria. Updating process was terminated once the new $pbest$ fitness had a value greater than the old $pbest$ fitness. The algorithm aims to produce a lower fitness value than an increasing value. In the end, the final $gbest$ was used as the optimal centroid. Cluster validity through SSE was applied using the final $gbest$. As a result, the optimal number of clusters and their corresponding characteristics were produced. MATLAB R2021a was utilized to perform the entire algorithm combination.

4. Results and Discussion

Section 4 comprises four subsections. First, feature selection results were discussed. Second, the findings using K-means clustering and PSO algorithms were elaborated. Third, the summarized cluster findings were presented. Fourth, the researchers proposed managerial implications.

4.1. Feature Selection

The results of the filter method using correlation are presented in Table 3. Out of 58 features, 7 features were found to be the most significant. If the features were ranked from highest to lowest, the features were arranged as follows: IU1, IU3, PBC2, AT2, PBC5, TR3, and SN1. Nevertheless, their ranking did not matter in feature selection as the goal was to identify the most optimal number of features. Considering the optimal features, the features pertain to internal variables affecting passengers' behaviors (trust, attitude, subjective norm, perceived behavioral control, and intention to use). This suggested that the filter method through correlation found more importance on internal features than external features (accessibility, safety, economic benefit, and crisis management).

Table 3. Filter method through correlation.

Number	Feature	Correlation (R) Value
1	TR3	0.5373
2	AT2	0.5674
3	SN1	0.5040
4	PBC2	0.5885
5	PBC5	0.5564
6	IU1	0.6588
7	IU3	0.6403

In this study, only 12.07% of the total features were significant. Compared to the results generated by Granados-López et al. [53]; they found 72.09% significant features using the filter method's correlation. This drastic difference occurred because the current study evaluated multicollinearity, while the past study did not assess the data's multicollinearity. It is vital to eliminate multicollinearity because of misinterpretation due to the extreme influence of one feature [72]. Furthermore, the correlation values of the seven significant features (ranging from 0.50 to 0.66) were considered moderate (0.50 to 0.70), which coincides with the acceptable multicollinearity principle.

Meanwhile, the univariate selection produced fewer significant features than the correlation method. According to one study, the univariate selection was one of the poor-

performing feature selections unless *limma* was added to the filtering process [17]. This current study had similar results to Bommert et al. [17] because the univariate selection was the sole feature selection method that generated only one significant variable (intention to use). Specifically, it was found that 3 out of 58 features were significant. Nonetheless, the *p*-values of IU2, IU5, and IU6 were less than 0.05. The three features' corresponding chi-square values and *p*-values are shown in Table 4.

Table 4. Filter method through univariate selection.

Number	Feature	Chi-Square	<i>p</i> -Value
1	IU2	88.9751	0.0043
2	IU5	88.6177	0.0046
3	IU6	81.3142	0.0189

Redundancy is a disadvantage of univariate selection because this filter method does not consider dependency between features [61]. However, the current study aimed to evaluate features individually. Thus, the results were deemed acceptable and features that did not produce a substantial impact on the target variable/class (perceived passenger behavior) were eliminated. Similar to the study of Matharaarachchi et al. [18], features were ordered based on the most significant *p*-values relaying informative features.

By using the wrapper method's backward elimination, 13 out of 58 features (22.41%) were considered to be the strongest performing features (Table 5). All 13 features met the minimum 0.05 significance level. Meanwhile, 7 out of 9 variables had at least 1 feature representative. Specifically, the seven variables were accessibility (AC), economic benefit (EB), crisis management (CM), trust (TR), subjective norm (SN), perceived behavioral control (PBC), and intention to use (IU). Thus, attitude (ATT) and safety (SA) were insignificant for the wrapper method's backward elimination. It could be seen that backward elimination yielded a better result than the two filter selection techniques. A past study also noted that backward elimination performed slightly better than other methods because it could ignore feature arrangement [62]. However, the past study overlooked that RFE was a better wrapper method than backward elimination.

Table 5. Wrapper method through backward elimination.

Number	Feature	<i>p</i> -Value
1	AC3	≤0.05
2	AC7	≤0.05
3	EB2	≤0.05
4	CM6	≤0.05
5	TR4	≤0.05
6	SN5	≤0.05
7	PBC1	≤0.05
8	PBC3	≤0.05
9	PBC5	≤0.05
10	IU2	≤0.05
11	IU3	≤0.05
12	IU5	≤0.05
13	IU6	≤0.05

In RFE, several features were removed recursively until the most optimal combination was generated. Moreover, the technique entails training and testing sets. Table 6 summarizes the highest RFE accuracy with the corresponding optimal feature combinations. As seen in the table, RFE accuracy ranges from 0.4817 to 0.7100. There was no minimum cut-off for RFE accuracy, but a score closest to 1.0000 was the most promising.

Table 6. Wrapper method through recursive feature elimination.

Training Size	Test Size	Optimal Number	Features	RFE Accuracy
90%	10%	26	AC3, AC7, SA3, SA6, EB1, EB2, EB4, CM3, CM6, TR3, TR4, TR5, AT5, AT6, SN3, SN5, PBC1, PBC2, PBC3, PBC4, PBC5, IU1, IU2, IU3, IU5, IU6	0.7100
80%	20%	7	CM6, TR4, PBC3, PBC5, IU2, IU5, IU6	0.6436
70%	30%	10	AC3, CM6, TR4, SN5, PBC3, PBC5, IU2, IU3, IU5, IU6	0.6375
60%	40%	13	AC3, AC7, CM6, TR4, AT6, SN5, PBC1, PBC3, PBC5, IU2, IU3, IU5, IU6	0.6560
50%	50%	8	CM6, TR4, PBC3, PBC5, IU2, IU3, IU5, IU6	0.6095
40%	60%	9	CM6, TR4, SN5, PBC3, PBC5, IU2, IU3, IU5, IU6	0.6004
30%	70%	8	CM6, TR4, PBC3, PBC5, IU2, IU3, IU5, IU6	0.5259
20%	80%	8	CM6, TR4, PBC3, PBC5, IU2, IU3, IU5, IU6	0.4817
10%	90%	3	TR4, PBC3, IU6	0.5299

The highest RFE accuracy belongs to 90% training and 10% testing, and this RFE split was considered to be the optimal solution among all the subsets. At least one feature in nine variables was considered in the highest score. Overall, 26 features were found important for the 90:10 split. It is noted that a higher training percentage was better to ensure data testing reliability. Additionally, 60% training and 40% testing had the second-highest RFE accuracy. This combination resulted in thirteen optimal features. These thirteen features were derived from seven variables. The third-highest RFE pertains to 80% training and 20% testing. In the 80:20 split, the seven optimal features were considered significant based on four variables. Past studies disclosed that 80% training and 20% testing was the best combination for feature selection [10,53]. However, the current study disagreed with the past study since the 80:20 split only had the third-highest RFE accuracy. A past study revealed 70% RFE accuracy was found to be the highest with a corresponding 70:30 training:testing size [58]. It resulted in 7 optimal features while comparing 8 data sets with varying 34 to 279 features. As the present study argued, 7 features were considered too low for a data set consisting of 279 features, but the number was sufficient for 34 features. The current findings also added that the 70:30 split only ranked fourth compared to other training:testing sizes. Meanwhile, one study noted that RFE was best applied in industrial production [55]. However, the present study argued that wrapper-RFE could also be applied to service problems, such as the public transportation system.

Figure 4 illustrates features' coefficients through the embedded method LASSO regression. Out of 58 features, 36 features contained zero coefficients and were eliminated from the model. The remaining 22 features had non-zero coefficients and were retained in the model. These 22 features were considered the most important features yielded by the LASSO method. Since the three features (SA3, AC7, and EB2) had negative correlations, they produced a negative relationship with the perceived passenger behavior. Meanwhile, 19 features (IU6, TR4, IU2, PBC5, PBC3, IU5, CM6, AC3, PBC4, IU3, SN5, AT6, IU1, TR3, PBC1, SN3, PBC2, IU4, and EB1) had a positive relationship with the perceived passenger behavior. Regardless of the relationship's direction, all 22 features were considered the optimal subset. Comparing the LASSO's subsets to others, LASSO was the second-best technique next to RFE. A past study noted that LASSO was a notable competitor because it cross-validated small, moderate, and large feature numbers [63]. Based on the findings, the LASSO model's alpha value was 0.0218 and the regression was 0.7134. These results were acceptable because alpha must be close to zero while regression should be close to one [63]. In a similar study, the LASSO model was utilized to reduce the model complexity by training the model iteratively [59].

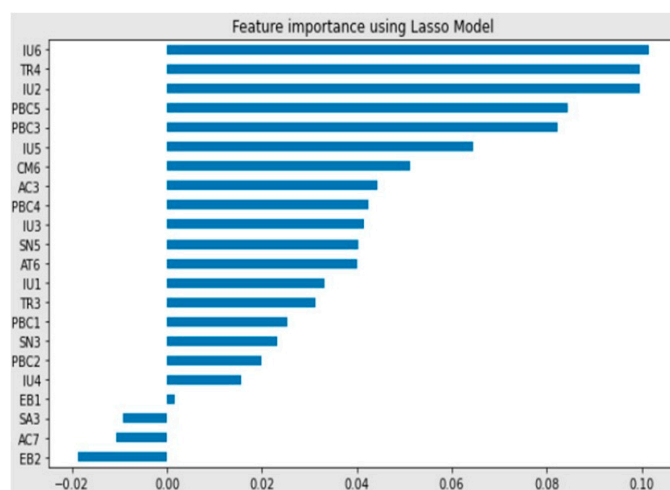


Figure 4. Embedded method using LASSO regression.

Lastly, stepwise regression generated nine important features with a p -value ≤ 0.05 as seen in Table 7. Since the original model had nine variables, stepwise regression supported six significant variables. These variables were accessibility (AC), economic benefit (EB), crisis management (CM), trust (TR), perceived behavioral control (PBC), and intention to use (IU). The model's correlation R-value was 0.855 and the regression R-squared value was 0.732. These results were acceptable since they were close to 1.00. Moreover, it posited that the model had an acceptable accuracy value and features had a positive relationship with the perceived passenger behavior. Likewise, Żogała-Siudem and Jaroszewicz [64] supported the promising performance of stepwise regression. Another study extracted the most number of features through stepwise regression [7]. In this study, stepwise regression only ranked third for the highest optimal features among all feature selection techniques. Nevertheless, the final model's correlation and regression values were adequate. Moreover, a past study found significant stepwise regression results, and the subset was considered the primary input of K-means clustering [9]. However, the present study did not utilize stepwise regression subsets in the succeeding K-means clustering algorithm because it was outperformed by other feature selection techniques.

Table 7. Stepwise regression.

Number	Feature	p -Value
1	AC3	0.0007
2	EB1	0.0117
3	EB2	0.0413
4	CM6	0.0367
5	TR4	0.0031
6	PBC3	0.0165
7	PBC5	0.0001
8	IU2	0.0005
9	IU6	0.0018

Table 8 demonstrates the summarized results of all feature selection methods. The optimal number and most important features for each method are presented in the table. Moreover, the determining factor was included to support the inclusion of important features in each subset.

Only two feature selection techniques (wrapper-RFE and embedded-LASSO) comprised features from all nine variables. The remaining six feature selection techniques did not entail all nine variables from the original model. Thus, this result reflects the superiority of wrapper-RFE and embedded LASSO. RFE contained a total of 26 optimal

features compared to the LASSO method's 22 features. Since feature selection aims to find the highest number of features with the highest accuracy determiner, the wrapper method's RFE with a 90% training set and 10% testing set was considered the optimal solution among all feature selection techniques. Likewise, RFE was the best performer because it could maintain a sufficient number of features without sacrificing accuracy, as well as eliminating overfitting and underfitting [53]. In another study, RFE appeared to have an average result when compared to a hybrid feature selection technique [18]. However, when compared to other basic techniques similar to the current study's approach, RFE was still superior [18].

Table 8. Feature selection summary results.

Feature Selection Method	Optimal Features	Determining Factor
Filter Method—Correlation	7—TR3, AT2, SN1, PBC2, PBC5, IU1, IU3	R cutoff = 0.5
Filter Method—Univariate Selection	3—IU2, IU5, IU6	<i>p</i> -value cutoff = 0.05
Wrapper Method—Backward Elimination	13—AC3, AC7, EB2, CM6, TR4, SN5, PBC1, PBC3, PBC5, IU2, IU3, IU5, IU6	<i>p</i> -value cutoff = 0.05
Wrapper Method—Recursive Feature Elimination (90:10)	26—AC3, AC7, SA3, SA6, EB1, EB2, EB4, CM3, CM6, TR3, TR4, TR5, AT5, AT6, SN3, SN5, PBC1, PBC2, PBC3, PBC4, PBC5, IU1, IU2, IU3, IU5, IU6	RFE accuracy = 0.7100
Embedded Method—LASSO	22—IU6, TR4, IU2, PBC5, PBC3, IU5, CM6, AC3, PBC4, IU3, SN5, AT6, IU1, TR3, PBC1, SN3, PBC2, IU4, EB1, SA3, AC7, EB2	Lasso alpha = 0.0218 Lasso regression score = 0.7134
Stepwise Regression	9—AC3, EB1, EB2, CM6, TR4, PBC3, PBC5, IU2, IU6	<i>p</i> -value cutoff = 0.05

Considering these findings, the researchers identified the strengths and weaknesses of each feature selection (Table 9). This table discussed several indicators, such as computation efficiency, data structure, significant features, determining factor's effectivity, underfitting and overfitting issues, bias concerns, and parameter settings. It was previously discussed that wrapper-RFE and embedded LASSO yielded promising figures. Their numerical findings coincided with the descriptive advantages and disadvantages. Both wrapper-RFE and embedded LASSO outperformed other feature selection techniques based on quantitative and descriptive analysis, but between the two, wrapper-RFE performed better than embedded LASSO.

Table 9. Advantages and disadvantages of feature selection techniques.

Feature Selection	Advantage	Disadvantage
Filter method's correlation	Easy computational approach Applied multicollinearity concept Non-complex data could be maximized better Found significant internal features (TR, AT, SN, PBC, IU)	Resulted in premature convergence Produced low R-values Lacked distinct parameters to support the significance of external features (AC, SA, EB, CM)
Filter method's univariate selection	Simple mathematical process Better used for problems requiring only one significant factor influencing the target variable	Prone to overfitting Could not identify subfeatures redundancy, resulting in a highly concentrated result Highly dependent on the target variable Produced the least number of features with unpromising <i>p</i> -values
Wrapper method's backward elimination	More appropriate for grouped features instead of individual features	Failed to identify <i>p</i> -values of the features individually Attitude and safety features were not found significant Extensive computation

Table 9. Cont.

Feature Selection	Advantage	Disadvantage
Wrapper method's recursive feature elimination	Applied multiple recursion until the stopping criteria were met Flexible parameters Eliminated weak features first, ensuring that the most important features were retained Produced a balanced feature subset since all features had corresponding subfeatures	Needed a higher training size to generate an optimal number of features coinciding with a good accuracy score Extensive computation
Embedded method's LASSO regression	Could process complex model Determined the relationships between features and target variables in two directions (positive and negative) Produced a balanced feature subset since all features had corresponding subfeatures	Too dependent on the target variable Lacked a fixed alpha value Prone to bias due to unstable parameters
Stepwise regression	Uncomplicated computation Enforced both adding and removing of features into/from the model	Too focused on the first optimal model as it could not assess multiple optimal solutions with varying features Lacked distinct parameters to support the significance of SA, AT, and SN

4.2. K-Means and PSO Algorithms

The parameter settings from Table 2 were applied to generate the needed results for the combined K-means and PSO algorithms. For a simpler illustration, Table 10 was created to classify parameter values for each parameter number. A total of 12 parameter combinations were simulated through MATLAB. Furthermore, K-means clustering and PSO algorithms utilized the data set from the most optimal feature selection method, which is the wrapper method's RFE 90:10 split. A total of 26 features or dimensions were fed into each parameter.

Table 10. K-means and PSO parameter combinations.

Parameter	<i>N</i>	<i>w</i>	<i>c1</i>	<i>c2</i>
1	20	0.6	1	1
2	30	0.6	1	1
3	40	0.6	1	1
4	20	0.9	1	1
5	30	0.9	1	1
6	40	0.9	1	1
7	20	0.6	2	2
8	30	0.6	2	2
9	40	0.6	2	2
10	20	0.9	2	2
11	30	0.9	2	2
12	40	0.9	2	2

Figure 5 portrays *gbest* fitness outliers and normality by executing 12 parameters in 10 runs. Ten out of twelve parameter combinations (parameters 1, 2, 3, 5, 6, 7, 8, 9, 11, and 12) had similarities, while two parameter combinations (parameters 4, 10) generated unstable *Gbest* fitness values. Hence, the 10 parameters were seen to have a close relationship with the optimal solution while the remaining 2 parameters were considered outliers. *Gbest* fitness or SSE is the cluster validity of the K-means and PSO algorithms. These values should be minimized to mitigate model error within the clustered groups. Hence, lower and more consistent values meant that the PUB passengers were clustered appropriately. The *gbest* statistics of all parameters are further elaborated in Supplementary Materials (Table S2).

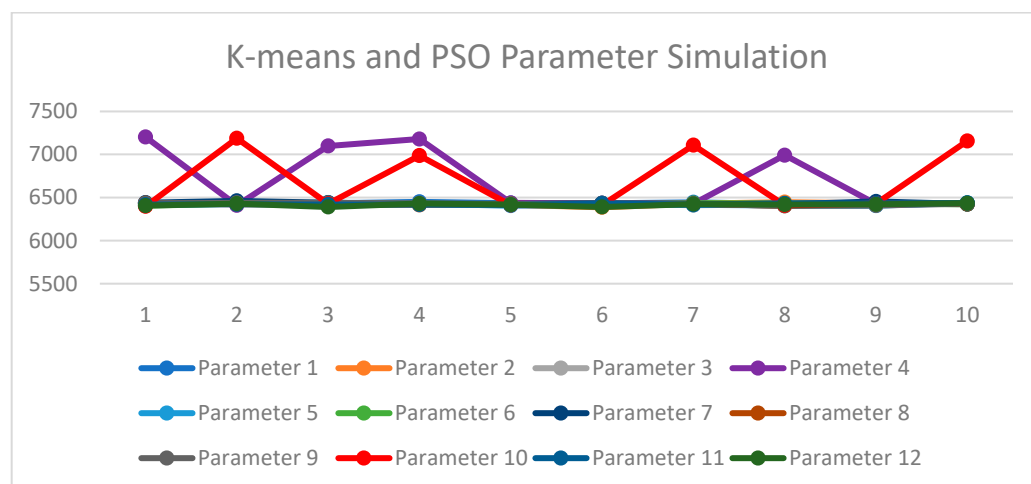


Figure 5. *Gbest* fitness simulation of 12 parameter combinations in 10 runs.

As a summary, Table 11 demonstrates the best and worst *gbest* fitness solutions. Parameter 12 produced the lowest mean (6415.86) while parameter 4 had the highest mean (6702.46). Since the study aims to minimize the error, parameter 12 was deemed the optimal solution as it had the lowest *gbest* fitness mean among all parameter combinations. Meanwhile, parameter 4 was the worst because it had the highest *gbest* fitness mean (6702.46) and the second-highest standard deviation (361.79). Parameter 10 produced the second-highest *gbest* fitness mean (6692.55) and produced the highest standard deviation (362.59). Hence, parameters 4 and 10 were identified as outliers due to a very high standard deviation, resulting in greater errors within the clusters. Other parameters' *gbest* fitness means were <6435 with a standard deviation of < 16.00. Although parameter 12 did not produce the least standard deviation among all combinations, 15.89 is fairly low considering that it also produced the lowest mean. Among all individual runs, parameter 2 (6387.30) had the minimum *gbest* fitness, and parameter 4 (6387.30) had the maximum *gbest* fitness. While parameter 2 had the best *gbest* fitness among individual runs, it did not necessarily pertain to the optimal solution since mean values were considered more reliable. Nevertheless, parameter 2's *gbest* fitness mean (6424.26) was close to parameter 12's *gbest* fitness mean (6415.86), and parameter 4 had the worst individual *gbest* fitness run and this result coincided with its outlier characteristics.

Table 11. Summary of all parameter combinations.

<i>Gbest</i> Descriptive Statistics	Value	Parameter
Lowest mean	6415.86	12
Highest mean	6702.46	4
Lowest standard deviation	7.82	11
Highest standard deviation	362.59	10
Minimum among all runs	6387.30	2
Maximum among all runs	7202.90	4

Therefore, the optimal K-means and PSO parameter combination is parameter number 12 with the following parameters: $N = 40$, $w = 0.9$, $c1 = c2 = 2$. Table 12 shows the corresponding *gbest* fitness for each run. The optimal *gbest* fitness learning curve was also provided (Supplementary Materials: Figure S1). *Gbest* fitness ranged from 6389.70 to 6434.20 with a standard deviation of 15.89. Piotrowski et al. [75] referred to one of the past studies in which 40 N was the best number of particles for the 30-dimensional setting problem. In this case, 26 dimensions were used based on the optimal feature selection and could be categorized close to the 30-dimensional setting when rounded up. Past studies tested various inertia weights (w), including 0.9 [74,75]. However, none of these past

studies disclosed the best inertia weight. On the other hand, another study tested 0.6, 0.7, and 0.8 wherein they found that 0.6 yielded the best w parameter [68]. Unlike the past results, the current study tested 0.6 and 0.9 and found that 0.9 generated a better solution. Like Xu et al. [74], acceleration coefficients ($c1$ and $c2$) with a value of two produced the optimal solution.

Table 12. Optimal parameter among all combinations.

Parameter 12	Value
N	40
w	0.9
c1	2
c2	2
No. of Runs	Gbest Fitness
1	6406.10
2	6429.10
3	6389.80
4	6430.40
5	6417.00
6	6389.70
7	6423.80
8	6417.80
9	6420.70
10	6434.20
Mean	6415.86
Standard Deviation	15.89
Minimum	6389.70
Maximum	6434.20

Cluster groups were generated by utilizing the optimal parameter #12. A total of 23 clusters were considered the optimal groupings of PUB passengers. These passengers were grouped according to their similarities and the corresponding formation was illustrated in Figure 6. The step-by-step plotting is demonstrated in Supplementary Materials (Figure S2). Other studies produced a lesser number of clusters formed by different tuning parameters [14,76]. For example, a study only discovered three optimal clusters because they evaluated the seven pre-determined data sets [14]. The current study sought actual responses from PUB passengers instead of the data sets available online. Hence, the current study's customized approach was more comprehensive in identifying the optimal cluster. Another instance was when a study concluded that 17 clusters produced consistent *gbest*, which entailed parameters $w = 0.8$, $c1 = 1$, $c2 = 2$ [76]. The present study did not consider this parameter combination because 0.9 was a better performer than 0.8 inertia weight (w). Moreover, the researchers prioritized similar $c1$ and $c2$ values to maintain model equilibrium. It was seen that $c1$ and $c2$ with unequal values led to premature convergence at either *pbest* or *gbest*.

Table 13 presents the comprehensive benefits and drawbacks of utilizing machine learning algorithms. K-means clustering and PSO were discussed individually to determine their unique characteristics. It is noted that the combination of K-means clustering and PSO yielded immense advantages compared to the individual algorithm.

However, not all 23 clustered groups gave importance as some only had a few participants (<30). Participants' specific details were provided in Supplementary Materials (Table S3). Thus, a Pareto chart was utilized to focus on the vital few similar to the approach of Peng et al. [69]. As reflected in Figure 7, five out of twenty-three clusters were considered vital—clusters 10, 14, 15, 16, and 21. These five clusters were presented in green-colored vertical graphs while non-vital clusters were demonstrated in blue bar graphs. The figure is also accompanied by a supporting procedure (Supplementary Materials: Table S4). How-

ever, Pareto analysis should be used strategically by ensuring that clusters were optimal. Unlike the study of Bommert et al. [17], they concluded that feature combinations had no ranking as they failed to incorporate appropriate clustering methods after feature selection.

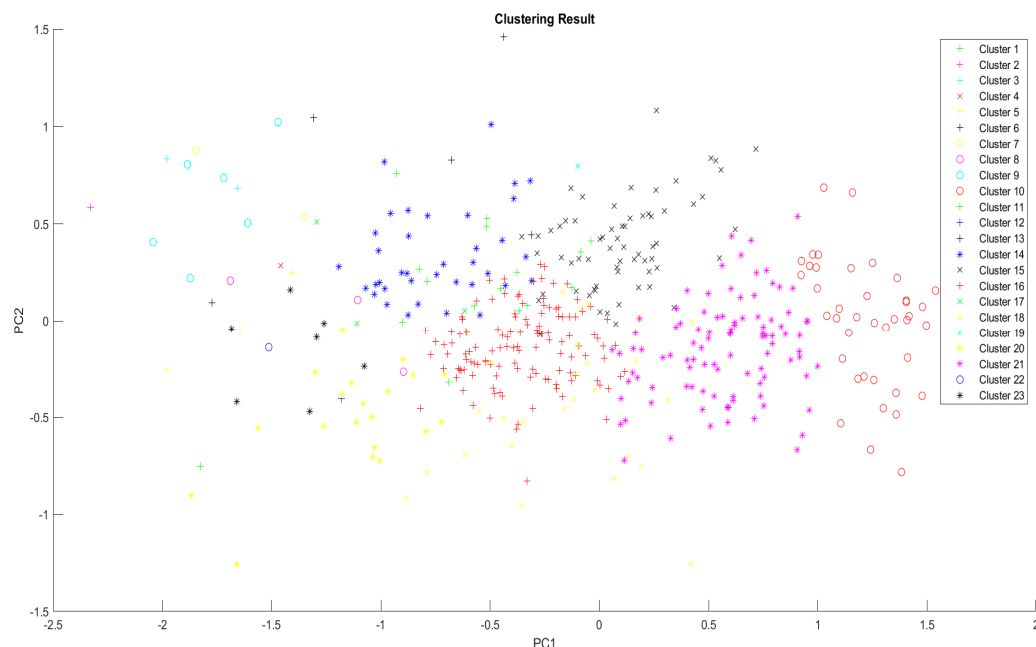


Figure 6. K-means and PSO clustering results.

Table 13. Advantages and disadvantages of machine learning algorithms.

Algorithm	Advantage	Disadvantage
K-means clustering	The simplest machine learning algorithm Could be used for either supervised or non-supervised learning	Weak fundamental due to the presence of random initial centroid Could only process 3 parameters based on the standard model settings Generated very poor SSE mean (7049.09) and standard deviation (126.22)
Particle Swarm Optimization (PSO)	Consisted of multiple iterations and would only stop if the termination condition was satisfied Stored multiple optimal solutions with varying <i>gbest fitness</i> values	Lacked initial clustered data, which would prompt randomization Time-consuming trial and error method to find the best parameters Needed human intervention to ensure that premature convergence at <i>pbest</i> was eliminated
Combined K-means clustering and PSO	Processed 12 parameters with varying <i>N</i> , <i>w</i> , <i>c1</i> , and <i>c2</i> A total of 10 parameters yielded consistent <i>gbest</i> results Produced a significantly lower <i>gbest</i> (6415.86) and standard deviation (7.82) compared to the results of K-means Generated 23 cluster groups with feature and demographic similarities among PUB passengers	Computationally extensive due to multiple parameter combinations The combination needed higher <i>N</i> , <i>w</i> , <i>c1</i> , and <i>c2</i> values

Figure 8 portrays the K-means clustering and PSO results for the five vital clusters. These clusters have unique legends in the plot. Cluster 16 is equivalent to red “+”, cluster 21 is magenta “*”, cluster 10 is red “o”, cluster 15 is black “x”, and cluster 14 is blue “*.”

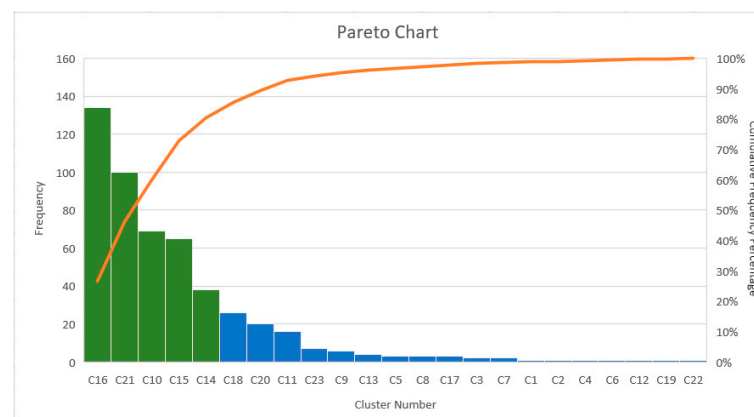


Figure 7. Pareto chart.

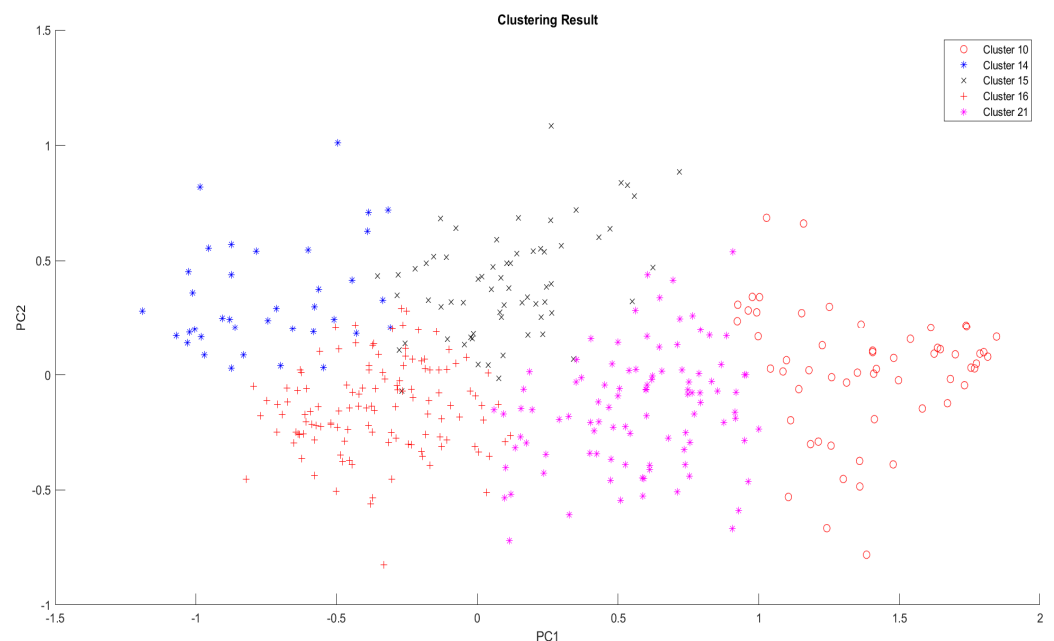


Figure 8. Vital clusters' plot results.

The succeeding results reflect the interpretation of vital few clusters (Clusters 10, 14, 15, 16, and 21). They were assessed to focus on important groupings. Figure S3 in Supplementary Materials shows the demographic profile of participants according to their clusters. Since the overall gender demographics showed that female passengers dominated the questionnaire, it also produced the same results across all clusters. Similarly, most PUB passengers from the overall demographics mentioned that they did not own a car; thus, results showed a higher PUB passenger frequency for those who did not own a vehicle than the ones who owned a car. However, there were interpretation differences for age, employment status, education level, allowance, and expenses (Supplementary Materials: Vital Clusters' Demographic Profiles). On the other hand, Figure S4 in Supplementary Materials illustrates the comparison of clusters dependent on PUB passengers' 5-point Likert responses to the 26 optimal features.

4.3. Clustering Summary

For accessibility, only two features (AC3 and AC7) were considered; wherein AC3 refers to PUBs' accessibility during daytime working hours and AC7 pertains to reasonable PUB stop locations and distances. Other studies noted that these accessibility factors produced positive PUB passenger satisfaction [26,27]. Similarities occurred because PUB

passengers yearn for convenience. Additionally, most participants from cluster 16 found AC3 and AC7 neutral. For cluster 21, most participants gave four out of five ratings, agreeing with AC3 and AC7. For clusters 10 and 15, a lot of participants extremely agreed (five out of five scores) with AC3 and AC7. Meanwhile, cluster 14 had varying results for AC3 and AC7 since they agreed with AC3 but were neutral for AC7. On one hand, a study concluded that AC3 was not a significant feature, rather AC5, AC6, and AC7 were the significant ones [6]. Through machine learning and statistical techniques, AC7 was considered a safe bet in all types of methods.

Safety's optimal features were SA3 and SA6. SA3 ensured that crimes were not frequent in PUBs, while SA6 guaranteed that passengers felt safe when riding PUBs. Most participants from clusters 16 and 14 answered neutrally to both features. Cluster 21 agreed with SA3 and responded neutrally to SA6. Cluster 10 extremely agreed with both features. Finally, cluster 15 was neutral with SA3 and agreed with SA6. Regardless of the agreeableness level, only SA6 was found to be the most common significant factor in comparison to the past study [6]. Inconsistent results occurred because the past study considered exogenous and endogenous relationships, while the present study considered SA features as an individual variable. Another piece of research evaluated SA independently and had a similar result to the current findings, whereas safety measures were mostly overlooked by PUB stakeholders [32].

Economic benefit had three optimal features: EB1, EB2, and EB4. EB1 supported the fairness of PUB fares, EB2 stated that PUB passengers allotted a small percentage of their income for travel expenses, and EB4 validated that passengers ride PUB due to its affordability. Cluster 16 mostly answered neutrally to all three economic benefit features. Cluster 21 agreed with EB1 and EB4 and was neutral with EB2. Cluster 10 extremely agreed with the three economic benefit features. Cluster 15 extremely agreed with EB1, was neutral with EB2, and agreed with EB4. Lastly, cluster 14 was neutral for EB1 and agreed with EB2 and EB4. These varying cluster responses coincided with EB1 and EB4 results presented by Cahigas et al. [6]. Two studies tried to prove EB2 but failed to find its significance [6,35]. Nevertheless, the current findings demonstrated the importance of EB2 because most PUB passengers were students and fresh graduates who had the least amount of budget and income.

Crisis management had two optimal features (CM3 and CM6), whereby CM3 revealed that PUB passengers appreciated the COVID-19 precautions mandated by the PUB system and government and CM6 stated that PUB passengers acknowledged the effectiveness of COVID-19 precautions followed in PUBs. On a similar note, CM3 was given importance by the passengers of Thomas et al. [39]. However, the past study preferred the removal of COVID-19 precautions, unlike the current study, which requested strict implementation. The remarkable difference was due to lesser active COVID-19 cases in Australia and New Zealand compared to the Philippines. Furthermore, clusters 16 and 14 had neutral responses to CM3 and CM6. Cluster 21 agreed with CM3 and CM6. Both clusters 10 and 15 extremely agreed with CM3 and CM6. For crisis management's two features, participants that belonged to the same clusters had similar types of responses.

Trust's optimal features were TR3, TR4, and TR5. According to TR3, PUB passengers thought that PUBs were reliable during the pandemic. In another study, TR3 was a significant feature but it was less reliable than TR5 [6]. For TR4, they considered PUBs to be essential. A past study agreed that instilling trust in PUB passengers was necessary, especially since PUB is one of the main public transportation modes [42]. For TR5, they felt comfortable riding PUBs during the pandemic. Likewise, TR5 was the dominant feature among these three optimal features [6]. This implied that PUB passengers prioritized comfort in sharing public space and using PUBs daily. Moreover, most participants from cluster 16 responded neutrally to all three trust features. Then, most participants under cluster 21 agreed with the three features. Most participants from cluster 10 extremely agreed with all three features. For cluster 15, most participants agreed with TR3 and TR4 and were neutral for TR5. Lastly, cluster 14 was neutral for TR3 and TR5 and agreed with TR4.

Attitude produced two optimal features, AT5 and AT6. PUB passengers felt satisfied with the current PUB system for AT5 and they found it highly acceptable to use PUB despite the pandemic for AT6. In a similar study, AT5 was a significant feature but AT6 was considered insignificant [6]. This inferred that the statistical approach of the past study failed to support passenger attitude's approval of PUB, while the present study sustained a positive attitude through the combined machine learning algorithms. Moreover, clusters 16, 15, and 14 dominated neutral answers to AT5 and AT6. Cluster 21 dominated agree or 4 out of 5 ratings for both features. Cluster 10 dominated extremely agree or 5 out of 5 ratings for the two attitude features. Thus, most participants in the same clusters dominated one type of response for AT5 and AT6.

Subjective norms also generated two optimal features, which were SN3 and SN5. For SN3, PUB passengers chose to ride PUB during the pandemic when they were surrounded by people who do the same thing. For SN5, PUB passengers' friends and family anticipated their loved one's usage of PUBs during the pandemic. Between S3 and S5, SN3 held a greater value based on the findings of Cahigas et al. [6]. In the current study, both features had equal importance because feature and cluster results disproved ranking. This conveyed that SN3 and SN5 best suited the overall model accuracy of predicting PUB passengers' behaviors. Additionally, most participants from clusters 16 and 15 answered neutrally to SN3 and SN5. Cluster 21 agreed and Cluster 10 extremely agreed with SN3 and SN5. Surprisingly, cluster 14 answered neutral to SN3 and disagreed with SN5. Only participants from cluster 14 mostly disagreed with one subjective norm feature.

Five out of six perceived behavioral control features were considered significant by the feature selection. PBC1, PBC2, and PBC3 indicated that riding PUB was a sole decision, easy, and acceptable. Moreover, PBC4 and PBC5 disclosed that PUB passengers felt confident with their COVID-19 knowledge and the PUB precautions. On the contrary, only three out of six PBC features were supported by a study [6]. Since the past study utilized multivariate analysis, the researchers concluded that machine learning algorithms could further prove the significance of features. Most participants under cluster 16 answered neutral for the five optimal PBC features. Cluster 21 participants mostly agreed with the five features. Cluster 10 participants mostly extremely agreed with the five features. For Cluster 15, participants extremely agreed with PBC1 and PBC4, responded neutral to PBC2, and PBC3, and agreed with PBC5. Finally, cluster 14 participants extremely agreed with PBC1, disagreed with PBC2, responded neutral to PBC3, and agreed with PBC4 and PBC5. Interestingly, most clusters had neutral and positive insights into PBC except for cluster 14, which disagreed with PBC2.

Intention to use PUB originally had six features, and the five optimal features were IU1, IU2, IU3, IU5, and IU6. Based on these five optimal features, PUB passengers will make an effort to use PUB, ride PUB during rush hour, ride PUB for essential purposes, increase PUB travel frequency, and have a strong intention to use PUB despite the pandemic. Most participants from clusters 16 and 15 answered neutral for the five IU features. Cluster 21 agreed with the five IU features. Cluster 10 extremely agreed with the five IU features. Meanwhile, cluster 14 responded neutral to IU1 and disagreed with IU2, IU3, IU5, and IU6. Cluster 14 had varying results compared to other clusters. Consistently, these IU optimal features produced an almost similar result to a past study [6]. The current and past studies only differed in using PUBs for leisure purposes. The present findings noted that PUB passengers disapproved of using PUBs for leisure and should only be used for essentials (e.g., grocery shopping, medical, school, and work). While hesitations were present, passengers made an effort to ride PUBs due to the lack of private cars [46]. Based on the present study's actual survey responses, 77% of the participants were not equipped with personal vehicles. Thus, it was a complementary condition that made passengers use PUBs.

4.4. Managerial Implications

The researchers present managerial implications in Table 14. These strategies aim to help PUB stakeholders improve the current PUB system during the COVID-19 pandemic. Since PUB passengers were clustered according to their demographics' and feature responses' similarities, the PUB stakeholders could utilize the clusters as a benchmark for each city. For example, cities mostly comprised of universities/schools (e.g., Manila City) should focus on implications present under clusters 16, 21, and 10 because students dominated these clusters. Meanwhile, if the PUB stakeholders aim to evaluate the PUB passenger behavior of employed individuals, they should focus on clusters 14 and 15. These clusters could be used as a benchmark to assess city demographics. Supplementary Materials (Five Clusters' Comprehensive Details and Table S5) gradually discuss the interpretation of PUB passengers' demographics and features clustered findings.

Table 14. Recommended strategies to improve the PUB system.

Cluster	Clustered Demographic	Strategies	Corresponding Features Requiring Improvement
16	• All age ranges (≤ 17 to ≥ 55 years old) • Dominated by high school levels, college students, and unemployed individuals • Most PUB passengers preferred the least PUB allowances and expenses	The PUB stakeholders must enhance the PUB system services according to all 26 features' characteristics. There should be adequate numbers of PUB on the road, especially in the daytime. The bus stops should have a reasonable distance. Crimes must not be feasible inside the bus; whereby, drivers/operators shall inspect the belongings of passengers or the PUB companies can install metal detectors. PUB fares should be affordable. The drivers and operators shall ensure that COVID-19-mandated protocols are implemented and strictly followed. Furthermore, the PUB stakeholders should improve the public-sharing comfortability experience by ensuring that PUBs are reliable; thus, regular PUB maintenance shall be checked by the PUB companies, drivers, and operators.	AC3, AC7, SA3, SA6, EB1, EB2, EB4, CM3, CM6, TR3, TR4, TR5, AT5, AT6, SN3, SN5, PBC1, PBC2, PBC3, PBC4, PBC5, IU1, IU2, IU3, IU5, and IU6
21	• ≤ 17 to 54 years old • Dominated by college students, bachelor's degree holders, and unemployed individuals • Most PUB passengers preferred the least PUB allowances and expenses	The PUB stakeholders must improve the public-sharing comfortability experience and ensure that the PUB fare increase is mitigated.	SA6 and EB2
10	• ≤ 17 to 44 years old and ≥ 55 years old • Dominated by high school levels, college students, and unemployed individuals • Most PUB passengers preferred the least PUB allowances and expenses	Since most passengers extremely agreed with the services provided by 26 features, there are no improvements needed. Instead, the PUB stakeholders must maintain the current PUB system.	N/A
15	• ≤ 17 years old to 44 years old • Dominated by college students, high school students, and full-time employees • Most PUB passengers preferred the least PUB allowances and expenses	The PUB stakeholders must ensure that crimes are not feasible inside the bus. PUB fares must be maintained at the lowest cost. Passengers should feel comfort and convenience when riding the PUB despite the pandemic.	SA3, EB2, TR5, AT5, AT6, SN3, SN5, PBC2, PBC3, IU1, IU2, IU3, IU5, and IU6
14	• ≤ 17 years old to 44 years old • Dominated by bachelor's degree holders, college students, and full-time employees • Most PUB passengers preferred the least, high, and mid (in chronological order) PUB allowances	The PUB stakeholders must ensure that bus stops have reasonable distance, crimes are not feasible inside the bus, and PUB fares are affordable. They should also improve the public-sharing comfortability experience by ensuring that PUBs are reliable. Moreover, the drivers and operators shall ensure that COVID-19-mandated protocols are implemented and strictly followed.	AC7, SA3, SA6, EB1, CM3, CM6, TR3, TR5, AT5, AT6, SN3, SN5, PBC2, PBC3, IU1, IU2, IU3, IU5, and IU6

Cluster 14 should be prioritized by the PUB stakeholders out of the five vital clusters. Cluster 14's corresponding features should be given importance since cluster 14 was the only cluster that expressed disagreement with some features. Suppose the government could identify the cities that had similarities with cluster 14. In that case, the strategies shall

be applied to the respective PUB city routes to ensure positive PUB passenger behavior during the pandemic.

For PUB passengers under Cluster 15, safety, affordability, comfort, and convenience were considered the most important factors. At every PUB stop, checkpoints should be overseen by military officers regularly. These officers were also advised to perform random inspections at a random PUB stop and time. PUB companies must discourage tinted windows on the passenger side to diminish the chances of holdups and hijacking. Meanwhile, PUB fare affordability could be maintained by increasing PUB subsidies under the government. With proper budget allocation, the government could benefit from PUBs economically due to the increasing demand for public transportation modes.

Furthermore, it was noted that PUB designs are dissimilar because PUBs are operated by different private companies. To improve the PUB passenger experience of clusters 15 and 21, the PUB companies were encouraged to follow ergonomically designed chairs, doors, and interiors. The comfort of riding a public transportation mode must be on par with the ergonomic standards to entice PUB passengers to use this mode frequently.

Meanwhile, cluster 16 entailed a more comprehensive approach since this cluster was comprised of inclusive backgrounds and young generations. Kids and teenagers were more susceptible to COVID-19 transmission but they were not frequent users of PUBs. Although cluster 16 was not a priority, PUB stakeholders were encouraged to follow the presented strategies in Table 14 for cities consisting of kids and teenagers.

5. Conclusions

PUB is a transportation mode greatly affected during the COVID-19 pandemic in the Philippines. PUB stakeholders applied system and protocol changes to prevent the spread of the virus. Following these changes, the researchers utilized feature selection, K-means clustering, and PSO to identify PUB passenger behaviors and similarities.

The three research questions were answered as follows: (1) Feature selection generated the most important features or dimensions, whereby the wrapper method's RFE was found the most optimal technique. Moreover, the combined K-means and PSO algorithm produced the optimal clusters by testing multiple parameters. (2) There were 23 optimal PUB passengers' clusters, and 5 clusters were deemed the most vital ones (Clusters 10, 14, 15, 16, and 21). (3) The clusters' demographics vary and were discussed in Table 14. Cluster 10 comprised PUB passengers who extremely agreed and felt very content with all features. Cluster 14 had the most diverse responses from disagree to strongly agree. It is the only cluster that disagrees with the current PUB system. Next, cluster 15 generated different results from neutral to strongly agree across all features. The majority of PUB passengers under cluster 16 had neutral responses. Finally, cluster 21 PUB passengers agreed on almost all features except for responding neutral to two features and extremely agree to another two features.

Following the related literature, none of them utilized feature selection, K-means clustering, and PSO in the context of PUB passengers' resulting behavior during the COVID-19 pandemic. This circumstance added significance to the novel approach of the present study. For instance, Shen et al. [67] improved bus layout, onboarding behavior, and off-boarding system but only focused on K-means clustering. Moreover, Zhong et al. [22] and Kechagiopoulos and Beligiannis [21] enhanced bus efficiency and road patterns through PSO. While past studies maximized the PUB and road efficiency, their findings lacked a comprehensive approach as their method was limited to data optimization. The present study assessed PUB and road problems alongside passengers' insights through the application of feature selection, K-means clustering, and PSO. Meanwhile, PUB passenger factor similarities were discovered in a past study [22]. However, the past study focused on primary factors instead of the underlying features of each factor. In line with this, the current findings covered significant features and demographic characteristics. Participants were grouped according to their appropriate clusters. These clusters were essential to PUB

stakeholders (operators, drivers, private transportation groups, and the government) in implementing case-by-case transportation issues.

Although the study had promising contributions, the researchers recognized the study's limitations that future scholars could explore. First, the study could be extended by evaluating the passengers' behavior according to the respective Philippine cities since each city's PUB system and route vary. As the results could be compared with city demographics, future researchers should include participants' geographical locations in the questionnaire. The results were still acceptable as the primary benchmark process could be dependent on the Philippines' overall city demographics. Other researchers may also consider their local cities' public transportation modes or perform a comparison of international public transportation modes. An additional recommendation was to investigate the perceptions of non-regular commuters in the current public transportation system since they could bring innovative insights. Moreover, feature selection techniques could be combined because the researchers only assessed each feature selection technique individually. For example, future scholars could merge filter correlation and univariate filter selection. Another suggestion was the integration of filter-wrapper, filter-embedded, and wrapper-embedded. Nevertheless, the researchers assessed the techniques individually to produce a standard due to the lack of relevant studies in the PUB and COVID-19 contexts. Future studies could also opt for the combination of feature selection techniques and other machine learning algorithms, such as random forest classifier, support vector machine, and Naïve Bayes. In further studies centered on the impacts of COVID-19 on public transportation, scholars should eliminate parameter 4 ($N = 20, w = 0.9, c1 = c2 = 1$) and parameter 10 ($N = 20, w = 0.9, c1 = c2 = 2$) from the K-means and PSO initialization. These two conditions were considered the worst-performing parameters. Lastly, future researchers could explore other K-means and PSO parameter settings ($N, w, c1, c2$) to find more solution sets. The present study tested 12 parameter settings and future researchers could test at least 20 combinations. Nonetheless, the current study used the best parameter settings according to past studies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/su15097410/s1>, Table S1: Corresponding meaning of each feature; Table S2: Summary of gbest fitness descriptive statistics; Table S3: Descriptive analysis of PUB passengers' clusters; Table S4: Pareto analysis; Table S5: Summarized analysis of the vital clusters; Figure S1: The learning curve of the optimal solution; Figure S2: Cluster groups of participants using the optimal parameter settings; Figure S3: Cluster comparison according to demographics; Figure S4: Cluster comparison according to features; Feature Selection Equations; PSO Initialization Equations; Vital Clusters' Demographic Profiles; Vital Clusters' Comprehensive Details.

Author Contributions: Conceptualization, M.M.L.C. and F.E.Z.; methodology, M.M.L.C., F.E.Z., A.K.S.O. and Y.T.P.; software, M.M.L.C., F.E.Z. and A.K.S.O.; validation, M.M.L.C., F.E.Z., A.K.S.O. and Y.T.P.; formal analysis, M.M.L.C., F.E.Z. and A.K.S.O.; investigation, M.M.L.C., F.E.Z. and A.K.S.O.; resources, M.M.L.C., F.E.Z., A.K.S.O. and Y.T.P.; data curation, M.M.L.C., F.E.Z., A.K.S.O. and Y.T.P.; writing—original draft preparation, M.M.L.C., F.E.Z., A.K.S.O. and Y.T.P.; writing—review and editing, M.M.L.C., F.E.Z., A.K.S.O. and Y.T.P.; visualization, M.M.L.C. and F.E.Z.; supervision, F.E.Z., A.K.S.O. and Y.T.P.; project administration, M.M.L.C., F.E.Z., A.K.S.O. and Y.T.P.; funding acquisition, A.K.S.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Mapúa University Directed Research for Innovation and Value Enhancement (DRIVE).

Institutional Review Board Statement: This study was approved by the Mapua University Research Ethics Committees (FM-RC-23-01-06).

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study (FM-RC-23-02-06).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The researchers would like to extend their deepest gratitude to the respondents of this study despite the current COVID-19 inflation rate.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mayo, F.L.; Taboada, E.B. Ranking Factors Affecting Public Transport Mode Choice of Commuters in an Urban City of a Developing Country Using Analytic Hierarchy Process: The Case of Metro Cebu, Philippines. *Transp. Res. Interdiscip. Perspect.* **2020**, *4*, 100078. [CrossRef]
2. Guillen, M.D.; Ishida, H.; Okamoto, N. Is the Use of Informal Public Transport Modes in Developing Countries Habitual? an Empirical Study in Davao City, Philippines. *Transp. Policy* **2013**, *26*, 31–42. [CrossRef]
3. Dela Peña, K. Sensible Public Transport: A Post-Pandemic Dream. Available online: <https://newsinfo.inquirer.net/1507740/sensible-public-transport-a-post-pandemic-dream> (accessed on 10 January 2022).
4. Philippine Statistics Authority. Highlights of the Philippine Population 2020 Census of Population and Housing (2020 CPH). Available online: <https://psa.gov.ph/content/highlights-philippine-population-2020-census-population-and-housing-2020-cph> (accessed on 10 January 2022).
5. Abadilla, E.V. DOTR Issues Level 3 Directive to Buses & Puvs. Manila Bulletin. Available online: <https://mb.com.ph/2022/01/03/dotr-level3-directive-to-buses-puvs/> (accessed on 10 January 2022).
6. Cahigas, M.M.; Prasetyo, Y.T.; Persada, S.F.; Ong, A.K.; Nadlifatin, R. Understanding the Perceived Behavior of Public Utility Bus Passengers during the Era of COVID-19 Pandemic in the Philippines: Application of Social Exchange Theory and Theory of Planned Behavior. *Res. Transp. Bus. Manag.* **2022**, *45*, 100840. [CrossRef]
7. Tsai, C.-F. Feature Selection in Bankruptcy Prediction. *Knowl. Based Syst.* **2009**, *22*, 120–127. [CrossRef]
8. Budak, A.; Sarvari, P.A. Profit Margin Prediction in Sustainable Road Freight Transportation Using Machine Learning. *J. Clean. Prod.* **2021**, *314*, 127990. [CrossRef]
9. Chou, S.-Y.; Dewabharata, A.; Bayu, Y.C.; Cheng, R.-G.; Zulvia, F.E. An Automatic Energy Saving Strategy for a Water Dispenser Based on User Behavior. *Adv. Eng. Inform.* **2022**, *51*, 101503. [CrossRef]
10. Rose, S.; Nickolas, S.; Sangeetha, S. A Recursive Ensemble-Based Feature Selection for Multi-Output Models to Discover Patterns among the Soil Nutrients. *Chemom. Intell. Lab. Syst.* **2021**, *208*, 104221. [CrossRef]
11. Joshi, C.; Ranjan, R.K.; Bharti, V. A Fuzzy Logic Based Feature Engineering Approach for Botnet Detection Using Ann. *J. King Saud Univ. Comput. Inf. Sci.* **2022**, *34*, 6872–6882. [CrossRef]
12. MacQueen, J. *Some Methods for Classification and Analysis of Multivariate Observations*; Project Euclid: Durham, NC, USA, 1966.
13. Jain, A.K. Data Clustering: 50 Years beyond K-Means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [CrossRef]
14. Kuo, R.J.; Rizki, M.; Zulvia, F.E.; Khasanah, A.U. Integration of Growing Self-Organizing Map and Bee Colony Optimization Algorithm for Part Clustering. *Comput. Ind. Eng.* **2018**, *120*, 251–265. [CrossRef]
15. Eberhart, R.; Kennedy, J. A New Optimizer Using Particle Swarm Theory. In Proceedings of the Sixth International Symposium on Micro Machine and Human Science (MHS'95), Nagoya, Japan, 4–6 October 1995; pp. 39–43.
16. Sahab, M.G.; Toropov, V.V.; Gandomi, A.H. A Review on Traditional and Modern Structural Optimization. In *Metaheuristic Applications in Structures and Infrastructures*; Elsevier: London, UK; Waltham, MA, USA, 2013; pp. 25–47.
17. Bommert, A.; Welchowski, T.; Schmid, M.; Rahnenführer, J. Benchmark of Filter Methods for Feature Selection in High-Dimensional Gene Expression Survival Data. *Brief. Bioinform.* **2021**, *23*, bbab354. [CrossRef] [PubMed]
18. Matharaarachchi, S.; Domaratzki, M.; Muthukumarana, S. Assessing Feature Selection Method Performance with Class Imbalance Data. *Mach. Learn. Appl.* **2021**, *6*, 100170. [CrossRef]
19. Anderson, T.K. Kernel Density Estimation and K-Means Clustering to Profile Road Accident Hotspots. *Accid. Anal. Prev.* **2009**, *41*, 359–364. [CrossRef] [PubMed]
20. Fotouhi, A.; Montazeri-Gh, M. Tehran Driving Cycle Development Using the K-Means Clustering Method. *Tehran Driv. Cycle Dev. Using K Means Clust. Method. Sci. Iran.* **2013**, *20*, 286–293.
21. Kechagiopoulos, P.N.; Beligiannis, G.N. Solving the Urban Transit Routing Problem Using a Particle Swarm Optimization Based Algorithm. *Appl. Soft Comput.* **2014**, *21*, 654–676. [CrossRef]
22. Zhong, S.; Zhou, L.; Ma, S.; Jia, N.; Zhang, L.; Yao, B. The Optimization of Bus Rapid Transit Route Based on an Improved Particle Swarm Optimization. *Transp. Lett.* **2016**, *10*, 257–268. [CrossRef]
23. Li, N.; Yang, L.; Li, X.; Li, X.; Tu, J.; Cheung, S.C.P. Multi-Objective Optimization for Designing of High-Speed Train Cabin Ventilation System Using Particle Swarm Optimization and Multi-Fidelity Kriging. *Build. Environ.* **2019**, *155*, 161–174. [CrossRef]
24. Li, X.; Li, D.; Hu, X.; Yan, Z.; Wang, Y. Optimizing Train Frequencies and Train Routing with Simultaneous Passenger Assignment in High-Speed Railway Network. *Comput. Ind. Eng.* **2020**, *148*, 106650. [CrossRef]
25. Ittamalla, R.; Srinivas Kumar, D.V. Determinants of Holistic Passenger Experience in Public Transportation: Scale Development and Validation. *J. Retail. Consum. Serv.* **2021**, *61*, 102564. [CrossRef]
26. Atombo, C.; Dzibordi Wemegah, T. Indicators for Commuter's Satisfaction and Usage of High Occupancy Public Bus Transport Service in Ghana. *Transp. Res. Interdiscip. Perspect.* **2021**, *11*, 100458. [CrossRef]

27. Tiglaio, N.C.; De Veyra, J.M.; Tolentino, N.J.; Tacderas, M.A. The Perception of Service Quality among Paratransit Users in Metro Manila Using Structural Equations Modelling (SEM) Approach. *Res. Transp. Econ.* **2020**, *83*, 100955. [\[CrossRef\]](#)
28. Chen, M.-C.; Hsu, C.-L.; Huang, C.-H. Applying the Kano Model to Investigate the Quality of Transportation Services at Mega Events. *J. Retail. Consum. Serv.* **2021**, *60*, 102442. [\[CrossRef\]](#)
29. Quy Nguyen-Phuoc, D.; Ngoc Su, D.; Nguyen, T.; Vo, N.S.; Thi Phuong Tran, A.; Johnson, L.W. The Roles of Physical and Social Environments on the Behavioural Intention of Passengers to Reuse and Recommend Bus Systems. *Travel Behav. Soc.* **2022**, *27*, 162–172. [\[CrossRef\]](#)
30. Wang, L.; Zhang, S.; Sun, W.; Chen, C.-L. Exploring the Physical and Mental Health of High-Speed Rail Commuters: Suzhou-Shanghai Inter-City Commuting. *J. Transp. Health* **2020**, *18*, 100902. [\[CrossRef\]](#)
31. Wang, Y.; Cao, M.; Liu, Y.; Ye, R.; Gao, X.; Ma, L. Public Transport Equity in Shenyang: Using Structural Equation Modelling. *Res. Transp. Bus. Manag.* **2022**, *42*, 100555. [\[CrossRef\]](#)
32. Deveci, M.; Öner, S.C.; Canitez, F.; Öner, M. Evaluation of Service Quality in Public Bus Transportation Using Interval-Valued Intuitionistic Fuzzy QFD Methodology. *Res. Transp. Bus. Manag.* **2019**, *33*, 100387. [\[CrossRef\]](#)
33. Shen, W.; Xiao, W.; Wang, X. Passenger Satisfaction Evaluation Model for Urban Rail Transit: A Structural Equation Modeling Based on Partial Least Squares. *Transp. Policy* **2016**, *46*, 20–31. [\[CrossRef\]](#)
34. Xue, Y.; Zhong, M.; Xue, L.; Zhang, B.; Tu, H.; Tan, C.; Kong, Q.; Guan, H. Simulation Analysis of Bus Passenger Boarding and Alighting Behavior Based on Cellular Automata. *Sustainability* **2022**, *14*, 2429. [\[CrossRef\]](#)
35. Rasoolimanesh, S.M.; Jaafar, M.; Kock, N.; Ramayah, T. A Revised Framework of Social Exchange Theory to Investigate the Factors Influencing Residents' Perceptions. *Tour. Manag. Perspect.* **2015**, *16*, 335–345. [\[CrossRef\]](#)
36. Turner, M.; Kwon, S.-H.; O'Donnell, M. State Effectiveness and Crises in East and Southeast Asia: The Case of COVID-19. *Sustainability* **2022**, *14*, 7216. [\[CrossRef\]](#)
37. Chuenyindee, T.; Ong, A.K.; Ramos, J.P.; Prasetyo, Y.T.; Nadlifatin, R.; Kurata, Y.B.; Sittiwatethanasiri, T. Public Utility Vehicle Service Quality and Customer Satisfaction in the Philippines during the COVID-19 Pandemic. *Util. Policy* **2022**, *75*, 101336. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Cahigas, M.M.; Prasetyo, Y.T.; Alexander, J.; Sutapa, P.L.; Wiratama, S.; Arvin, V.; Nadlifatin, R.; Persada, S.F. Factors Affecting Visiting Behavior to Bali during the Covid-19 Pandemic: An Extended Theory of Planned Behavior Approach. *Sustainability* **2022**, *14*, 10424. [\[CrossRef\]](#)
39. Thomas, F.M.F.; Charlton, S.G.; Lewis, I.; Nandavar, S. Commuting before and after COVID-19. *Transp. Res. Interdiscip. Perspect.* **2021**, *11*, 100423. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Strielkowski, W.; Zenchenko, S.; Tarasova, A.; Radyukova, Y. Management of Smart and Sustainable Cities in the Post-COVID-19 Era: Lessons and Implications. *Sustainability* **2022**, *14*, 7267. [\[CrossRef\]](#)
41. Borhan, M.N.; Ibrahim, A.N.; Miskeen, M.A. Extending the Theory of Planned Behaviour to Predict the Intention to Take the New High-Speed Rail for Intercity Travel in Libya: Assessment of the Influence of Novelty Seeking, Trust and External Influence. *Transp. Res. Part A Policy Pract.* **2019**, *130*, 373–384. [\[CrossRef\]](#)
42. Fallah Zavareh, M.; Mehdizadeh, M.; Nordfjærn, T. Demand for Mitigating the Risk of COVID-19 Infection in Public Transport: The Role of Social Trust and Fatalistic Beliefs. *Transp. Res. Part F Traffic Psychol. Behav.* **2022**, *84*, 348–362. [\[CrossRef\]](#)
43. Restuputri, D.P.; Indriani, T.R.; Masudin, I. The Effect of Logistic Service Quality on Customer Satisfaction and Loyalty Using Kansei Engineering during the COVID-19 Pandemic. *Cogent Bus. Manag.* **2021**, *8*, 1906492. [\[CrossRef\]](#)
44. Ajzen, I. The Theory of Planned Behavior. *Organ. Behav. Hum. Decis. Process.* **1991**, *50*, 179–211. [\[CrossRef\]](#)
45. Cahigas, M.M.; Prasetyo, Y.T.; Persada, S.F.; Nadlifatin, R. Examining Filipinos' Intention to Revisit Siargao after Super Typhoon Rai 2021 (Odette): An Extension of the Theory of Planned Behavior Approach. *Int. J. Disaster Risk Reduct.* **2023**, *84*, 103455. [\[CrossRef\]](#)
46. Lee, J.; Baig, F.; Pervez, A. Impacts of COVID-19 on Individuals' Mobility Behavior in Pakistan Based on Self-Reported Responses. *J. Transp. Health* **2021**, *22*, 101228. [\[CrossRef\]](#)
47. van Wee, B.; Witlox, F. COVID-19 and Its Long-Term Effects on Activity Participation and Travel Behaviour: A Multiperspective View. *J. Transp. Geogr.* **2021**, *95*, 103144. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Wu, D.; Gu, H.; Gu, S.; You, H. Individual Motivation and Social Influence: A Study of Telemedicine Adoption in China Based on Social Cognitive Theory. *Health Policy Technol.* **2021**, *10*, 100525. [\[CrossRef\]](#)
49. Krettenauer, T.; Lefebvre, J.P. Beyond Subjective and Personal: Endorsing pro-Environmental Norms as Moral Norms. *J. Environ. Psychol.* **2021**, *76*, 101644. [\[CrossRef\]](#)
50. Cahigas, M.M.; Prasetyo, Y.T.; Persada, S.F.; Nadlifatin, R. Filipinos' Intention to Participate in 2022 Leyte Landslide Response Volunteer Opportunities: The Role of Understanding the 2022 Leyte Landslide, Social Capital, Altruistic Concern, and Theory of Planned Behavior. *Int. J. Disaster Risk Reduct.* **2023**, *84*, 103485. [\[CrossRef\]](#)
51. Shaaban, K.; Maher, A. Using the Theory of Planned Behavior to Predict the Use of an Upcoming Public Transportation Service in Qatar. *Case Stud. Transp. Policy* **2020**, *8*, 484–491. [\[CrossRef\]](#)
52. Gao, Y.; Chen, X.; Shan, X.; Fu, Z. Active Commuting among Junior High School Students in a Chinese Medium-Sized City: Application of the Theory of Planned Behavior. *Transp. Res. Part F Traffic Psychol. Behav.* **2018**, *56*, 46–53. [\[CrossRef\]](#)
53. Granados-López, D.; Suárez-García, A.; Díez-Mediavilla, M.; Alonso-Tristán, C. Feature selection for CIE Standard Sky Classification. *Sol. Energy* **2021**, *218*, 95–107. [\[CrossRef\]](#)

54. Xiong, C.; Yang, M.; Kozar, R.; Zhang, L. Integrating transportation data with Emergency Medical Service Records to improve triage decision of high-risk trauma patients. *J. Transp. Health* **2021**, *22*, 101106. [\[CrossRef\]](#)
55. Liu, Y.; Lyu, C.; Liu, Z.; Cao, J. Exploring a large-scale multi-modal transportation recommendation system. *Transp. Res. Part C Emerg. Technol.* **2021**, *126*, 103070. [\[CrossRef\]](#)
56. Soares, E.F.d.S.; Campos, C.A.V.; Lucena, S.C.d. Online Travel Mode Detection Method Using Automated Machine Learning and Feature Engineering. *Future Gener. Comput. Syst.* **2019**, *101*, 1201–1212. [\[CrossRef\]](#)
57. Rodríguez-Sanz, Á.; Fernández de Marcos, A.; Pérez-Castán, J.A.; Comendador, F.G.; Arnaldo Valdés, R.; París Loreiro, Á. Queue Behavioural Patterns for Passengers at Airport Terminals: A Machine Learning Approach. *J. Air Transp. Manag.* **2021**, *90*, 101940. [\[CrossRef\]](#)
58. Yang, J.; Ma, J. Compressive Sensing-Enhanced Feature Selection and Its Application in Travel Mode Choice Prediction. *Appl. Soft Comput.* **2019**, *75*, 537–547. [\[CrossRef\]](#)
59. Thabtah, F.; Kamalov, F.; Hammoud, S.; Shahamiri, S.R. Least Loss: A Simplified Filter Method for Feature Selection. *Inf. Sci.* **2020**, *534*, 1–15. [\[CrossRef\]](#)
60. Cekik, R.; Uysal, A.K. A Novel Filter Feature Selection Method Using Rough Set for Short Text Data. *Expert Syst. Appl.* **2020**, *160*, 113691. [\[CrossRef\]](#)
61. Labani, M.; Moradi, P.; Ahmadizar, F.; Jalili, M. A Novel Multivariate Filter Method for Feature Selection in Text Classification Problems. *Eng. Appl. Artif. Intell.* **2018**, *70*, 25–37. [\[CrossRef\]](#)
62. Kohavi, R.; John, G.H. Wrappers for Feature Subset Selection. *Artif. Intell.* **1997**, *97*, 273–324. [\[CrossRef\]](#)
63. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 267–288. [\[CrossRef\]](#)
64. Żogała-Siudem, B.; Jaroszewicz, S. Fast Stepwise Regression Based on Multidimensional Indexes. *Inf. Sci.* **2021**, *549*, 288–309. [\[CrossRef\]](#)
65. Eltvéd, M.; Breyer, N.; Ingvardson, J.B.; Nielsen, O.A. Impacts of Long-Term Service Disruptions on Passenger Travel Behaviour: A Smart Card Analysis from the Greater Copenhagen Area. *Transp. Res. Part C Emerg. Technol.* **2021**, *131*, 103198. [\[CrossRef\]](#)
66. Li, Q.; Liu, R.; Zhao, J.; Liu, H.-C. Passenger Satisfaction Evaluation of Public Transport Using Alternative Queuing Method under Hesitant Linguistic Environment. *J. Intell. Transp. Syst.* **2021**, *26*, 330–342. [\[CrossRef\]](#)
67. Shen, C.; Sun, Y.; Bai, Z.; Cui, H. Real-Time Customized Bus Routes Design with Optimal Passenger and Vehicle Matching Based on Column Generation Algorithm. *Phys. A Stat. Mech. Its Appl.* **2021**, *571*, 125836. [\[CrossRef\]](#)
68. Kuo, R.J.; Nugroho, Y.; Zulvia, F.E. Application of Particle Swarm Optimization Algorithm for Adjusting Project Contingencies and Response Strategies under Budgetary Constraints. *Comput. Ind. Eng.* **2019**, *135*, 254–264. [\[CrossRef\]](#)
69. Peng, Y.; Li, T.; Bao, C.; Zhang, J.; Xie, G.; Zhang, H. Performance Analysis and Multi-Objective Optimization of Bionic Dendritic Furcal Energy-Absorbing Structures for Trains. *Int. J. Mech. Sci.* **2023**, *246*, 108145. [\[CrossRef\]](#)
70. Xiao, G.; Juan, Z.; Gao, J. Travel Mode Detection Based on Neural Networks and Particle Swarm Optimization. *Information* **2015**, *6*, 522–535. [\[CrossRef\]](#)
71. German, J.D.; Redi, A.A.; Ong, A.K.; Prasetyo, Y.T.; Sumera, V.L. Predicting Factors Affecting Preparedness of Volcanic Eruption for a Sustainable Community: A Case Study in the Philippines. *Sustainability* **2022**, *14*, 11329. [\[CrossRef\]](#)
72. Voss, D.S. Multicollinearity. In *Encyclopedia of Social Measurement*; Elsevier: Amsterdam, The Netherlands, 2005; pp. 759–770.
73. Ryan, L.; Kuhn, S.; Colreavy-Donnelly, S.; Caraffini, F. Particle Swarm Optimisation in Practice: Multiple Applications in a Digital Microscope System. *Appl. Sci.* **2022**, *12*, 7827. [\[CrossRef\]](#)
74. Xu, G.; Cui, Q.; Shi, X.; Ge, H.; Zhan, Z.-H.; Lee, H.P.; Liang, Y.; Tai, R.; Wu, C. Particle Swarm Optimization Based on Dimensional Learning Strategy. *Swarm Evol. Comput.* **2019**, *45*, 33–51. [\[CrossRef\]](#)
75. Piotrowski, A.P.; Napiorkowski, J.J.; Piotrowska, A.E. Population Size in Particle Swarm Optimization. *Swarm Evol. Comput.* **2020**, *58*, 100718. [\[CrossRef\]](#)
76. Kuo, R.J.; Potti, Y.; Zulvia, F.E. Application of metaheuristic based Fuzzy K-modes algorithm to supplier clustering. *Comput. Ind. Eng.* **2018**, *120*, 298–307. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.