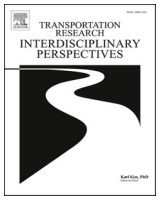
Transportation Research Interdisciplinary Perspectives 27 (2024) 101238

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives



Toward real-time deterrence against fare evasion risk in public transport

Benedetto Barabino [a], Massimo Di Francesco [b,*], Roberto Ventura [a], Simone Zanda [b]

[a] Department of Civil, Environmental, Architectural Engineering and Mathematics (DICATAM), University of Brescia, Brescia, Italy
[b] Department of Mathematics and Computer Science, University of Cagliari, 09123 Cagliari, Italy

ARTICLE INFO

Keywords:
Fare Evasion
Risk Prediction
Risk Management
Artificial Neural Network

ABSTRACT

Fare evasion is a critical threat for Transit Agencies (TAs) and/or Public Transport Companies (PTCs) everywhere, especially in Proof-of-Payment Transit Systems (POP-TSs). The research on fare evasion risk is limited and based on econometric models restricting time characterization to a single period. This paper aims to enhance the use of fare evasion risk over several periods for possible real-time deterrence against fare evasion. The paper moves from an existing framework, identifying the factors of fare evasion and risk exposure in terms of frequency (or probability) and severity (or vulnerability), and adopts Artificial Neural Networks (ANNs) to shed light on the intricate nexus between these components, estimating the fare evasion risk for every (segment of a) route. Next, the risk index is evaluated for each time period of interest. The predictions are ranked and represented by time-dependent dashboards to recognize routes with high-risk evasion that require deterrence strategies. Some real-time strategies are simulated from fare inspection logs, passenger surveys, and probability distributions on data collected in three years. In conclusion, this research provides actionable insights for TAs/PTCs in dealing with fare compliance and can be integrated into any bus transit management system.

1. Introduction

Fare evasion is a primary threat to Transit Authorities/Public Transport Companies (TAs/PTCs) worldwide, owing to economic, social, and psychological implications. Although no public transportation system is immune from fare evasion, Proof-of-Payment Transit Systems (POP-TSs) without barriers have the highest evasion rate (e.g., Barabino et al., 2020; Barabino et al., 2024). Indeed, passengers must buy and validate the ticket (or pass) before using the service, but they may astutely choose to evade the fare because tickets are not automatically checked.

Addressing fare evasion is tricky due to policy, deterrence, enforcement, operational costs, and equipment facets (e.g., Wolfgram et al., 2022). To date, a unified strategy against fare evasion does not exist. TAs/PTCs have two main "weapons" to fight fare evasion: ticket (or pass) inspections and fine charging for evaders (e.g., Sasaki, 2014; Guarda et al., 2016; Dai et al., 2018; Alhassan et al., 2022; Barabino et al., 2022b; Wolfgram et al., 2022; Celse and Grolleau, 2023). The employment of inspectors is known in the literature as the "conventional" strategy (e.g., Delbosc and Currie, 2019).

1.1. State of the art

Inspections protect revenue and reduce aggressive behavior and vandalism (e.g., Barabino and Salis, 2020). Optimal inspection levels reduce fare evasion (e.g., Barabino et al., 2015; Cools et al., 2018; Porath and Galilea, 2020), whereas suboptimal levels are less effective (e.g., Guzman et al., 2021). The research on deterrence shows that offenders are more dissuaded by the certainty of being caught than by the severity of the fine (e.g., Smith and Clarke, 2020). Thus, proper inspection levels could decrease crime and increase passenger security (e.g., Killias et al., 2009). Several studies investigated the deterrence using simulated data (e.g., Boyd et al., 1989; Boyd, 2020) and real data (e.g., Barabino et al., 2013, Barabino et al., 2014; Barabino and Salis, 2019), as well as optimization methods for scheduling inspections (e.g., Yin et al., 2012; Correa et al., 2017; Brotcorne et al., 2021; Escalona et al., 2024).

In many TAs/PTCs worldwide, the metrics of fare inspection effectiveness are based on the ratio E/P, where E usually refers to fare evaders (or a proxy) and P to passengers. The most common metric is the fare evasion ratio (FER), where E is measured by fines, and P is the number of inspected passengers and is measured by the number of tickets checked in a time window (e.g., Wolfgram et al., 2022). The higher the indicator, the higher the fare evasion. Even though this statistic is frequently used to quantify fare evaders, TAs and PTCs often disagree on the definition of "evader." Moreover, determining the number of checked passengers is a complex task (e.g., Dauby and Zoltan, 2007; Wolfgram et al., 2022). In

* Corresponding author.
E-mail address: mdifrance@unica.it (M. Di Francesco).

https://doi.org/10.1016/j.trip.2024.101238
Received 16 December 2023; Received in revised form 17 July 2024; Accepted 21 September 2024
Available online 4 October 2024

more advanced PTCs, this metric can be replaced by the ratio between the number of validated smart cards and the number of passengers boarded, which is measured by automatic passenger counting for each location (e.g., Pourmonet et al., 2015). The lower the indicator, the higher the fare evasion. However, this *modus operandi* has flaws as well: in many transit systems worldwide, it is not mandatory for pass holders to tap in/out their tickets, which are often paper-based. Therefore, a low value of this ratio could provide biased information on the concentration of evaders. Moreover, because smart card data and passenger data were automatically collected, they also provide biased measures if they are not correctly handled. To summarize, both metrics could be inaccurate.

The conventional inspection strategy can be enhanced by the introduction of a risk metric, as done in many fields of engineering and economics (e.g., Aven, 2015; Ventura et al., 2023). Indeed, TAs/PTCs can identify high-risk evasion areas on the transit network, implement targeted strategies to reduce revenue losses and develop policies and actions promoting fairness and equity. Moreover, the determination of high-risk evasion routes (areas) can lead TAs/PTCs to an effective deployment of personnel, devices, and technology. Furthermore, TAs/PTCs can identify patterns and areas where measures need to be strengthened to ensure the safety and security of passengers and staff, thereby fostering a comfortable traveling experience. Since fare inspection can lead to delays on buses and can negatively impact passenger satisfaction, the implementation of alternative measures against fare evasion risk may result in improved operational efficiency, smoother boarding processes, and enhanced passenger experiences. Finally, the evaluation of fare evasion risk involves analyzing data on inspection policy, demographics, and behavioral factors. This process promotes data-driven decision-making within TAs/PTCs to understand the dynamics of fare evasion and devise suitable targeted strategies. This evidence-based decision-making fosters efficiency, transparency, and accountability in managing transit networks.

The risk concept integrates the frequency and/or probability, severity and/or consequences, and exposure terms. They must be integrated because they depend on several explanatory factors (e.g., attributes, determinants, predictors, or variables). Very few studies on fare evasion adopted the risk concept and separately investigated the frequency and/or probability and severity in South America, Australia, and Europe by descriptive statistics and inferential modeling.

Specific studies in Santiago (Chile) revealed a positive correlation between the frequency of fare evasion and the number of passengers boarding (and alighting), the number of doors on buses, the number of passengers boarding through the back door, high occupancy rates, and long headway (e.g., Guarda et al., 2016). Furthermore, fare evasion is more common in the afternoon and the evening for young men, crowded buses without turnstiles, and bus stop far from metro stations without devices for off-board payment and ticket sales (e.g., Cantillo et al., 2022).

Other studies have investigated how potential fare evaders' characteristics, behaviors, motivations, and attitudes can affect the probability of fare evasion. These facets were investigated according to a one-size-fits-all, *a priori*, or *a posteriori* segmentation.

One-size-fits-all segments showed that the probability of fare evasion could be influenced in different ways by sociodemographic factors, travel behavior, and situational variables (e.g., Eddy, 2010 Bucciol et al., 2013; Barabino et al., 2015; Dai et al., 2017; Cools et al., 2018; Allen et al., 2019).

As for the *a priori* segmentation, Currie and Delbosc (2017) showed that honesty and evasion tolerance explained both intentional and unintentional evasion. On the contrary, ticket competency and perceived ease of evasion explained intentional and unintentional evasion, respectively. According to Barabino and Salis (2020), male gender and a history of fare evasion are common characteristics that increase the likelihood of fare evasion among students, workers, and jobless passengers. Conversely, specific predictors characterize each segment.

As for the *a posteriori* segmentation, Delbosc and Currie (2016) identified three segments of passengers: deliberate evaders, unintentional evaders, and evaders with no history of evasion, each one having the specific attitude to (not) evade the fare. They were largely clustered according to the structural and operational elements of the transit system. González and Codocedo (2019) identified "proud," "empathetic," and "circumstantial" paying passengers and "radical," "strategic," "ambivalent," and "accidental" non-paying passengers, mainly clustered according to several personality factors (e.g., values and attitudes). According to Guzman et al. (2021), age and evasion records are the strongest determinants of evasion. Additionally, the likelihood of fare evasion decreases as the level of satisfaction increases. However, personality traits can limit how satisfaction affects fare evasion. Barabino and Salis (2023) showed that the probability of evading the fare depends on the employment level and discovered a medium segment of captive and chronic unemployed evader passengers, many captive and frequent evader students, and a minor portion of rare evader choice workers.

Less attention was devoted to the severity, which included several issues, such as lost income and/or violence aboard. A few works were based on circumstances that had already occurred, and they usually estimated the lost revenues from fare evasion (e.g., Cosby, 1985; Isreal and Strathman, 2002; Prokosch and Gartsman, 2017).

Barabino et al. (2023) first applied the risk concept in this field. They proposed a systematic framework for assessing the risk of fare evasion along each route or segment of a transport network. This framework proposed generalized regression models for frequency, logistic regression for severity, and a risk function linking these components. The method yielded the fare evasion risk index for every (segment of a) route, which is then rated and shown on user-friendly dashboards to plan and implement deterrence strategies.

1.2. Literature gaps

All previous works provided valuable contributions to our understanding of some components of fare evasion. However, some gaps persist.

First, although Barabino et al. (2023) introduced an approach for the average measure of risk over a single period of interest, none has investigated the change of fare evasion risk over several periods to plan and implement real-time deterrence strategies accordingly. This is particularly useful against the so-called calculators, who are always on guard and ready to defraud the ticketing system intentionally (e.g., Delbosc and Currie, 2016; Salis et al., 2017). For example, they often share information on inspections in real-time by virtual communities of fare evaders, i.e., maps of the potential places/routes of inspection activities (e.g., Assaf and Van den Broeck, 2022). Therefore, it is worth capturing the variation of fare evasion risk over time to target the inspection along routes.

Second, all prediction methods are based on econometric models (e.g., negative binomial regression for the frequency, binomial logit for the severity). They have the great ability to recognize the effect of each factor on fare evasion, which is an essential prerequisite to understanding the phenomena. Moreover, the research on factors influencing fare evasion considers generalized linear models. No work has been made on Artificial Intelligence models such as Artificial Neural Networks (ANNs), which are expanding in many engineering and economic fields and can be leveraged to estimate the risk components. Unlike econometric models, ANNs are functional relationships between system inputs and system outputs even if the effect of each predictor on the response variable may not be easy-to-understand (e.g., Zhang, 2010). Moreover, leveraging ANNs is intriguing in capturing the intricate (often nonlinear) relationship between the frequency and severity of fare evasion and its influencing factors to provide recommendations for fare evasion risk planning and management.

1.3. Objective

This paper aims to cover the previous gaps with a new framework for modeling, forecasting, ranking, and managing risk in a different way.

Building on Barabino et al. (2023), it leverages an ANN model to capture the complex relationship among the frequency, the severity of fare evasion, and the related factors to build a risk function for each (segment of a) route. Next, the risk function is evaluated in short time windows, ranked for real-time utilization, and represented by time-dependent dashboards to identify routes requiring major attention. Next, some real-time deterrence strategies are simulated. A real-world experiment is effectively done on data gathered over three years by passenger surveys, inspection logs, and probability distributions.

This experiment contributes both theoretically and practically. To our knowledge, no theoretical research has been done toward real-time estimation and management of fare evasion risk by ANN. In terms of application, this framework provides an advanced applicative tool for TAs/PTCs to assess the fare evasion risk of routes. It can serve as a decision support tool to enhance fare payment on specific routes and quickly warn TAs/PTCs of high-risk fare evasion on (segment of a) route.

1.4. Paper outline

The remainder of the paper follows. Section 2 shows the framework for real-time estimating and managing the risk of fare evasion by ANN. Section 3 presents the application of this framework in a real case study. Section 4 reports the results and some discussions. Section 5 draws conclusions and future research.

2. Methodology

This section describes the methodology for evaluating and managing the fare evasion risk on a bus transit network, which is summarized in Fig. 1.

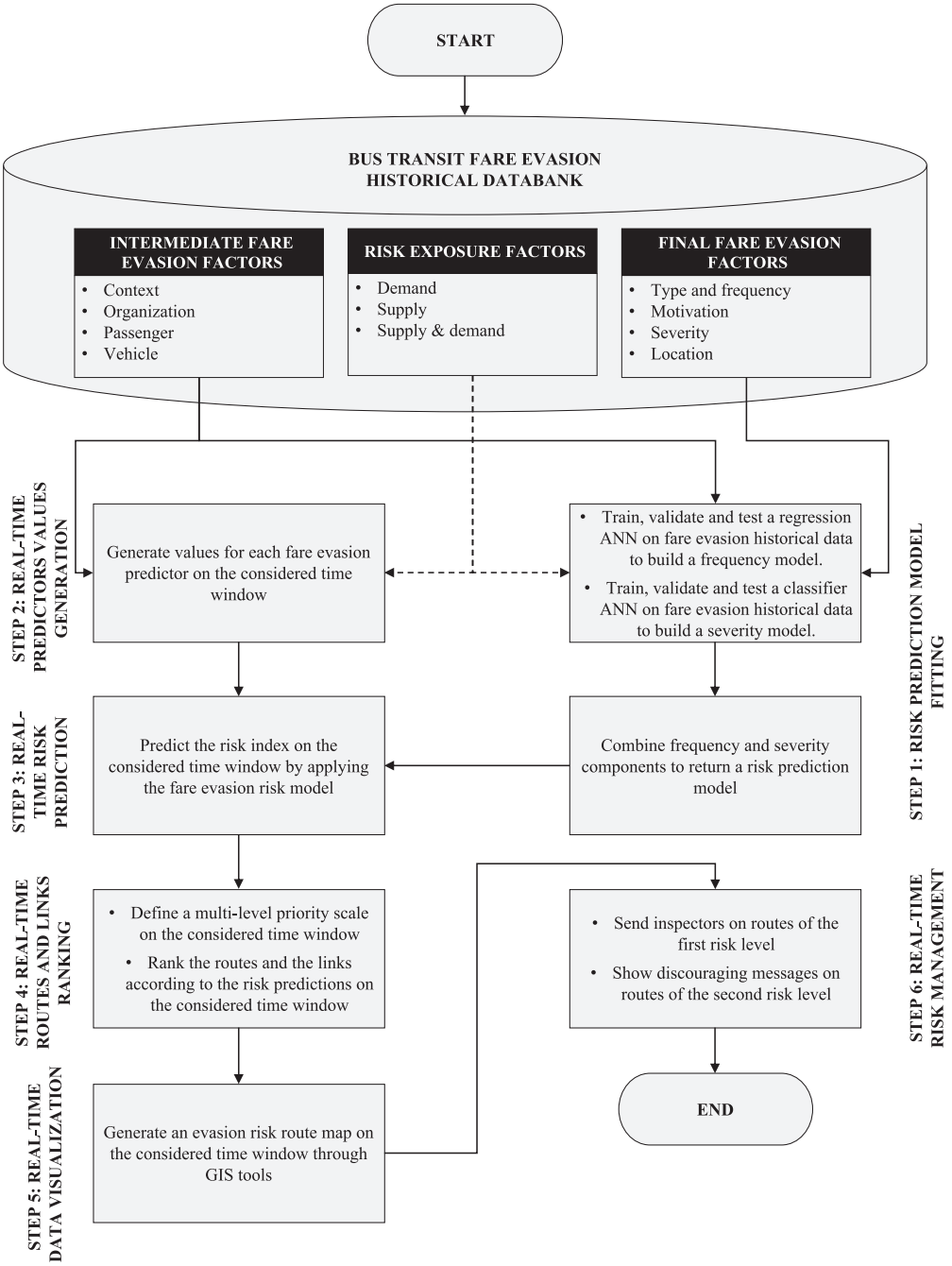


Fig. 1. Flow chart of the proposed methodology for the evaluation and real-time management of the fare evasion risk on a bus transit network.

It adopts the symbols recommended by the American National Standards Institute (ANSI) and the International Organization of Standardization (International Organization of Standardization, 1985).

2.1. Bus transit fare evasion historical dataset

A fare evasion event is generated by the interaction between *intermediate factors* and *risk exposure factors*. *Intermediate factors* may generate the conditions of a potential risk event. They depend on the context, organization, passengers, and vehicles. *Risk exposure factors* reflect the actual use of the transit network in terms of the number of individuals or services interested in fare evasion occurrence. They include supply-oriented, demand-oriented, and both supply- and demand-oriented factors. Both intermediate and risk exposure factors can affect the frequency (or the probability) and severity of fare evasion. Their interplay affects the occurrence and severity of the fare evasion event and results in *final fare evasion factors*. These factors include the type and frequency of fare evasion, related motivations, severity, and location. The experimental dataset of TAs/PTCs can be exploited to collect all these factors. More details are reported in Barabino et al. (2023).

2.2. STEP 1 – Risk prediction model fitting

According to STEP 1, an index to evaluate the risk of fare evasion on a bus transit network is defined according to this notation. Let:

- L be the set of routes.
- $S(l)$ be the set of homogeneous segments of route $l \in L$. They are legs with the same road configuration or the same urban environment or legs between two consecutive bus stops.
- T be the set of time windows discretizing the measure of fare evasion risk.
- FEE_{slt} be the Fare Evasion Exposure on homogeneous segment $s \in S(l)$ in time window $t \in T$. It represents the rate of occurrence of a potential fare evader travelling on the considered segment.
- FEP_{slt} be the Fare Evasion Probability on homogeneous segment $s \in S(l)$ in time window $t \in T$. It represents the likelihood that, if a potential fare evader pops-up, the fare evasion event takes place.
- FEC_{slt} be the Fare Evasion Consequences on homogeneous segment $s \in S(l)$ in time window $t \in T$; it represents the most likely result of a potential fare evasion event (e.g., additional costs for TAs/PTCs).
- FEF_{slt} be the Fare Evasion Frequency on homogeneous segment $s \in S(l)$ in time window $t \in T$; it is a driver of the probability FEP_{slt}.
- FES_{slt} be the Fare Evasion Severity on homogeneous segment $s \in S(l)$ in time window $t \in T$; it is a driver of the fare evasion consequences FEC_{slt}.
- E be the set of exposure factors and $e_{slt} \in E$ the generic exposure factor observed on homogeneous segment $s \in S(l)$ of route $l \in L$ in time window $t \in T$.
- X be the set of frequency predictors and $x_{slt} \in X$ the generic frequency predictor on homogeneous segment $s \in S(l)$ of route $l \in l$ in time window $t \in t$.
- Y be the set of severity predictors and $y_{slt} \in Y$ the generic severity predictor on homogeneous segment $s \in S(l)$ of route $l \in L$ in time window $t \in T$.

Hence, for each route $l \in L$ and time window $t \in T$, the risk function R_{lt} is defined by summing the risk R_{slt} on each segment $s \in S(l)$ of route $l \in L$ in time window $t \in T$. According to Fine (1971), the risk function is defined as follows:

$$R_{lt} \overset{\text{def}}{=} \sum_{s \in S(l)} R_{slt} \overset{\text{def}}{=} \sum_{s \in S(l)} FEP_{slt} * FEE_{slt} * FEC_{slt};\ \forall l \in L; \forall t \in T \tag{1}$$

Each term of (1) needs to be estimated according to the historical dataset of fare evasion. A straightforward method for this task is to build a complete bivariate risk model based on frequency and severity functions and include the exposure factor in the frequency model (Barabino et al., 2023). The latter is a logical choice that naturally follows from the definitions of probability, exposure, and frequency factors. It can be straightforwardly demonstrated that the frequency of fare evasion events at time $t \in T$ (i.e., FEF_{slt}) is obtained by multiplying the following components: 1) the likelihood that, if a possible fare evader pops-up, the fare evasion occurs (i.e., FEP_{slt}); 2) the rate of repetition of the fare evaders (i.e., FEE_{slt}). Hence, the frequency and the severity of fare evasion are assumed to be response variables, and they are predicted as a function of the selected explanatory factors. Specifically, the frequency is assessed as a function of the risk exposure factors and a set of predictors mined from intermediate safety outcome factors. Similarly, the severity is forecasted as a function of a set of predictors taken from intermediate safety outcome factors. Finally, the risk index is obtained by multiplying frequency and severity outcomes.

$$R_{lt} \overset{\text{def}}{=} \sum_{s \in S(l)} R_{slt} \overset{\text{def}}{=} \sum_{s \in S(l)} FEF_{slt}(\{e_{slt} \in E\}, \{x_{slt} \in X\}) * FES_{slt}(\{y_{slt} \in Y\}); \forall l \in L; \forall t \in T \tag{2}$$

Therefore, each component of eqn. (2) should be modeled and estimated. In this study, both frequency and severity components are modeled by ANNs. A regression ANN is fitted for the frequency component, which is modelled by an unbounded discrete variable. Conversely, a classifier ANN is chosen for the severity component since it is represented by a binary variable. More formally, let:

- J be the set of frequency or severity records in the historical dataset.
- $TR \subset J$, $VA \subset J$ and $TE \subset J$ be the training, validation, and test subsets, respectively.
- F be the set of frequency or severity explanatory factors, i.e., E and X, or Y, respectively, and $f_j \in F$ the generic factor in record $j \in J$.
- $\overline{\overline{INP}} \in \mathbb{R}^{|J| \bullet |F|}$ be the input matrix for the ANN fitting process, i.e., the matrix of each explanatory factor $f \in F$ related to each record $j \in J$.
- $\overline{\overline{TAR}} \in \mathbb{R}^{|J|}$ be the target vector for the ANN fitting process, i.e., the vector of the Fare Evasion Frequency or the Fare Evasion Severity with entries tar_j for each $j \in J$.
- ω be a function linking the input matrix $\overline{\overline{INP}}$ to the target vector $\overline{\overline{TAR}}$.
- $\widetilde{\omega}$ be an approximation of ω.
- P be the set of ANN hidden layer perceptrons[1] and $p \in P$ the generic perceptron.
- inp_{pj} be the input signal of the hidden layer perceptron $p \in P$ for $j \in J$.
- out_{pj} be the output signal of the hidden layer perceptron $p \in P$ for $j \in J$.
- $\overline{inp_j}$ be the input signal of the output layer perceptron for $j \in J$.
- $\overline{out_j}$ be the output signal of the output layer perceptron for $j \in J$.
- b_p be the bias for the input signal of hidden layer perceptron $p \in P$.
- $\overline{b}$ be the bias for the input signal of the output layer perceptron.
- w_{pf} be the weight factor for the input signal of the hidden layer perceptron $p \in P$ for $f \in F$.
- $\overline{w_p}$ be the weight factor for output signal of the hidden layer perceptron $p \in P$.
- ϕ be the activation function of the hidden layer perceptrons.
- $\overline{\phi}$ be the activation function of the output layer perceptron.
- ϱ be the input variable of the activation functions.
- ς be the slope parameter of the linear activation function.
- $\overline{\overline{\theta}} \in \mathbb{R}^{|P|(|F|+1)+1}$ be the generic vector with the parameters of the ANN model and $\overline{\overline{\theta}}_0 \in \mathbb{R}^{|P|(|F|+1)+1}$ be the parameter vector obtained through the learning phase.

[1] A perceptron is the node of the network, the ANN.

- ε_j be the generic residual value, i.e., the difference between the observed and predicted frequency or severity for $j \in J$, and $\overline{\overline{\varepsilon}} \in \mathbb{R}^{|J|}$ be the vector of the residual values.
- *RSS* be the Residual Sum of Squares, i.e., the sum of the squared residual values.
- *CF* be the Cost Function quantifying the prediction error of the ANN model.

Then, ω links the input matrix $\overline{\overline{INP}}$ to the target vector $\overline{\overline{TAR}}$, such that $\overline{\overline{TAR}} = \omega(\overline{\overline{INP}})$. This function can be interpreted as a computation model linking the causes ($\overline{\overline{INP}}$) with their observed outcomes ($\overline{\overline{TAR}}$). Generally speaking, ANNs define a mapping $\overline{\overline{TAR}} = \widetilde{\omega}(\overline{\overline{INP}}, \overline{\overline{\theta}}) + \overline{\overline{\varepsilon}}$ and discover the values of the parameter vector $\overline{\overline{\theta}}$ that leads the best approximation (Ian et al., 2016), i.e., that minimizes a properly defined Cost Function (*CF*).

A two-layer feed-forward network is chosen to perform this mapping (e.g., Schmidhuber, 2015). In this architecture, the information proceeds forward from the input nodes, through the hidden nodes, and towards the output nodes. As shown in Fig. 2, this network is formed by two layers of perceptrons: a hidden layer of $|P|$ perceptrons and an output layer on one perceptron.

As for the hidden layer, the input signal inp_{pj} of each perceptron $p \in P$ is obtained by multiplying the factors values f_j with their weight factors w_{pf} and then adding the bias factor b_p, according to eqn. (3):

$$inp_{pj} = b_p + \sum_{f \in F} w_{pf} * f_s; \forall p \in P; \forall j \in J \tag{3}$$

Then, the output signal out_{pj} of each perceptron in the hidden layer is determined by an activation function ϕ on the input signal. A hyperbolic tangent sigmoid activation function is selected for the hidden layer (e.g., Anastassiou, 2011; Romero Reyes et al., 2013; Sramka et al., 2019). Therefore:

$$\phi(\varrho) = tanh(\varrho) = \frac{e^{\varrho} - e^{-\varrho}}{e^{\varrho} + e^{-\varrho}}; \tag{4}$$

$$out_{pj} = \phi\left(inps_{pj}\right); \forall p \in P; \forall j \in J \tag{5}$$

Similarly, as for the output layer, the input signal $\overline{inp_j}$ of perceptron is obtained by multiplying the output signals out_{pj} originating from the hidden layer perceptrons by the weight factors $\overline{w_p}$ and then adding the bias factor $\overline{b}$, according to eqn. (6):

$$\overline{inp_j} = \overline{b} + \sum\nolimits_{p \in P} \overline{w_p} * out_{pj}; \forall j \in J \tag{6}$$

Then, the output signal ($\overline{out_j}$) of the perceptron in the output layer is determined by an activation function $\overline{\phi}$ to the input signal. The expression of $\overline{\phi}$ depends on the typology of the response variable. Particularly, since the frequency is an unbounded response variable, a linear activation function is selected. Conversely, a sigmoidal shape is preferred for the severity considering the binary nature of this response variable. Thus, a regression or a classification ANN is obtained for the frequency or the severity component, respectively. Hence:

$$\overline{\phi}(\varrho) = \begin{cases} \varsigma\varrho & \textit{for the frequency model;} \\ \dfrac{e^{\varrho}}{e^{\varrho}+1} & \textit{for the severity model;} \end{cases} \tag{7}$$

$$\overline{out_j} = \overline{\phi}\left(\overline{inp_j}\right); \forall j \in J \tag{8}$$

Hence, the predicted value of the response variable (denoted as $\widetilde{tar_j}$) equals the perceptron output signal. More formally:

$$\widetilde{tar_j} = \overline{out_j}; \forall j \in J \tag{9}$$

Before fitting, the records (i.e.,*J*) are randomly partitioned into three subsets: training, validation, and test. The training set (i.e., *TR*) is provided to the network during training to directly estimate weights and biases. The validation set (i.e., *VA*) is employed to assess the network generalization and check if overtraining begins to occur. The testing set (i.e., *TE*) has no effect on training and offers an unbiased metric of network performance.

During the training-validation phase, the training data *TR* is employed to tune the weights and the biases to minimize the cost function *CF* calculated on the validation subset *TR* (Brownlee, 2019). In this paper, two forms of *CF* are selected:

- The Mean Squared Error (denoted as *MSE*) for the frequency model, being the most popular for regression problems where a quantity is predicted (Reed and Marks, 1999).
- The Cross Entropy (denoted as *CE*) for the severity model, as it is the most appropriate for classification problems, where outputs are interpreted as likelihoods of belonging to a specified class (Reed and Marks, 1999; MathWorks, 2022a).

More formally:

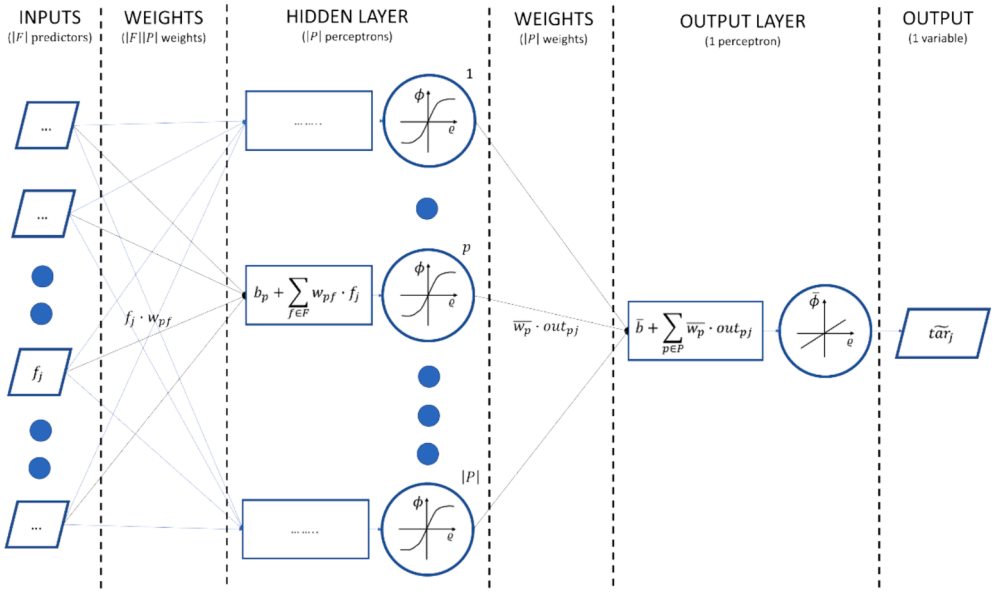


Fig. 2. Schematic structure of the two-layer feed-forward ANN selected for the frequency and severity prediction models.

$$CF: \begin{cases} MSE = \frac{\sum_{j\in VA} \varepsilon_j{}^2}{|VA|} & \text{for the frequency model;} \\ CE = -\frac{1}{|VA|}\sum_{j\in VA} tar_j \log(\widetilde{tar_j}) + (1 - tar_j)\log(1 - \widetilde{tar_j}) & \text{for the severity model;} \end{cases} \tag{10}$$

Hence, the training-validation procedure is formalized as follows:

$$\bar{\bar{\theta}}_0 = \underset{\bar{\bar{\theta}}}{\text{argmin}}\ CF(\bar{\bar{\theta}}) \tag{11}$$

A back-propagation algorithm is chosen, being one of the most popular for training–validating feedforward neural networks (Ian et al., 2016). Three back-propagation algorithms are adopted to solve (11): Levenberg-Marquardt, Bayesian Regularization, and Scaled Conjugate Gradient. The first typically requires more memory but less training time, while the second is generally more time-consuming, but it can lead to a better generalization for complex, small, or noisy datasets. The third requires less memory (MathWorks, 2022b). In all algorithms, the training stops when there is no further improvement in generalization, which is indicated by an increase in the cost function on the validation sample.

It may be noted that training–validating several times will produce different results due to dissimilar initial conditions and random splitting. Hence, multiple instances are issued, and the best solution is chosen according to the lowest CF. At the end of the training-validation phase, $\theta\grave{\iota}_{\acute{c}0}$ is selected, and the functional form of the ANN model is obtained:

$$\widetilde{tar_j} = \tilde{\omega}\left(\left\{f_j \in F\right\}, \bar{\bar{\theta}}_0\right); \forall j \in J \tag{12}$$

After the training-validation phase, the ANN model is tested to evaluate its performance in the fitting of training and validation data (i.e., *TR* and *VA*), as well as in predicting unseen records (i.e., *TE*). To execute the testing procedure, different metrics are considered depending on the response variable at hand, as summarised in Table 1.

More precisely, as for the frequency:

- The Regression Plot (*RP*) is graphed for *TR*, *VA* and *TE* by plotting the couples $(tar_j, \widetilde{tar_j})$ on a Cartesian plane; hence, the closer the points to the first quadrant bisector, the greater the fitting goodness.
- The Pearson correlation coefficient between tar_j and $\widetilde{tar_j}$ (denoted as *R*) is computed for *TR*, *VA* and *TE*; it is a positively oriented score, i.e., the greater the value, the better the outcomes.
- The Mean Absolute Error (denoted as *MAE*), the Root Mean Squared Error (denoted as *RMSE*), and the Coefficient of Variation (denoted as *CoV*) are computed for *TR*, *VA* and *TE*; hence, the results are negative oriented scores, i.e., the smaller the values, the better outcomes.

As for the severity:

- The Confusion Matrix (*CM*) is computed for *TR*, *VA* and *TE*; in the rows it reports the true class of severity and in the columns the predicted class. To clarify, *CM* is a 2x2 matrix with True Negative (denoted as *TN*) and True Positive (denoted as *TP*) values along the main diagonal and False Positive (denotes as *FP*) and False Negative (denoted as *FN*) values along the antidiagonal. The closer to 0 the values along the antidiagonal, the greater the fitting goodness, since null numbers for *FP* and *FN* indicate a perfect prediction.
- The True Positive Rate (denoted as *TPR*), the True Negative Rate (denoted as *TNR*) and the accuracy (denoted as *ACC*) are computed for *TR*, *VA* and *TE*; hence, the results are positively oriented scores, i.e., the greater the values, the better the results.

Next, the relevance of each explanatory factor is evaluated by computing the Permutation Feature Importance (denoted as PFI_f), that it is recognized as a reliable indicator, particularly for non-linear or opaque estimators (Strobl et al., 2007). The PFI_f is stated as a reduction in a

Table 1
Performance metrics for frequency and severity models, defined on the *TE* subset. As the for *TR* and *VA* subsets, performance metrics are computed according to the same definitions, where *TE* is replaced with *TR* or *VA*, respectively.

Response variable	Metric	Symbol	Definition
Frequency	Regression Plot	*RP*	Cartesian graph of $(tar_j, \widetilde{tar_j})$ couples $(\forall j \in TE)$
	Pearson correlation coefficient	*R*	$\frac{\sum_{j\in TE}(tar_j - mean(tar_j)) * (\widetilde{tar_j} - mean(\widetilde{tar_j}))}{\sqrt{\sum_{j\in TE}(tar_j - mean(tar_j))^2} * \sqrt{\sum_{j\in TR}(\widetilde{tar_j} - mean(\widetilde{tar_j}))^2}}$
	Mean Absolute Error	*MAE*	$\frac{\sum_{j\in TE}\lvert\varepsilon_j\rvert}{\lvert TE\rvert}$
	Root Mean Squared Error	*RMSE*	$\sqrt{\frac{\sum_{j\in TE}\varepsilon_j{}^2}{\lvert TE\rvert}}$
	Coefficient of Variation	*CoV*	$\frac{RMSE}{\frac{\sum_{j\in TE}\lvert tar_j\rvert}{\lvert TE\rvert}}$
	Ratio among the total number of predicted and observed fare evasion events	π	$\frac{\sum_{j\in TE}\widetilde{tar_j}}{\sum_{j\in TE} tar_j}$
Severity	Confusion Matrix	*CM*	A 2x2 matrix with $\mathscr{TN}$ and $\mathscr{TP}$ values along the main diagonal, and $\mathscr{FP}$ and $\mathscr{FN}$ values along the antidiagonal
	True Positive Rate (or Sensitivity)	*TPR*	$\frac{\mathscr{TP}}{\mathscr{TP} + \mathscr{FN}}$
	True Negative Rate (or Specificity)	*TNR*	$\frac{\mathscr{TN}}{\mathscr{TN} + \mathscr{FP}}$
	Accuracy	*ACC*	$\frac{\mathscr{TP} + \mathscr{TN}}{\mathscr{TP} + \mathscr{FN} + \mathscr{TN} + \mathscr{FP}}$

model score after a single factor value is casually rearranged (Breiman, 2001). This approach destroys the link between the factor and the target, and the model's score drops, thus revealing how strongly the model relies on the factor. This indicator is calculated on *TE* to stress the importance of features in out-of-sample prediction. More precisely, let:

- N be the set of arbitrary permutation instances.
- $\overline{\overline{INP}_{n,f}} \in \mathbb{R}^{|J|,|F|}$ be the corrupted version of the input matrix at the permutation $n \in N$, that is generated by casually rearranging the column related to $f \in F$ of the uncorrupted input matrix $\overline{\overline{INP}}$.
- $CF(\theta_0)_{TE}$ be the *CF* of the trained ANN model, calculated on *TE* by the uncorrupted input matrix $\overline{\overline{INP}}$.
- $CF(\theta_0)_{TEn,f}$ be the *CF* of the trained ANN model, calculated on *TE* by the corrupted input matrix $\overline{\overline{INP}_{n,f}}$.
- $s = -CF(\theta_0)_{TE}$ be the reference score of the trained ANN model on the uncorrupted input matrix $\overline{\overline{INP}}$.
- $s_{n,f} = -CE(\theta_0)_{TEn,f}$ be the score of the trained ANN model on the corrupted input matrix $\overline{\overline{INP}_{n,f}}$.

Hence, several permutation instances ($|N|$) are executed for each $f \in F$ and the related PFI_f are calculated by taking the average of the scores according to eqn. (13).

$$PFI_f = s - \frac{1}{|N|}\sum_{n \in N} s_{n,f}; \forall f \in F \tag{13}$$

The higher PFI_f, the higher the importance of $f \in F$.

Although the effect of each predictor on the response variable is not evident, a deeper analysis can be made by calculating the associated partial derivative. More formally, let:

- $\frac{\partial \widetilde{\omega}}{\partial \overline{f}_j}$ be the partial derivative of the function $\widetilde{\omega}\left(\left\{f_j \in F\right\}, \overline{\overline{\theta}}_0\right)$ with respect to the specific factor $\overline{f} \in F$ computed in record $j \in J$.
- $\Delta\overline{f}$ be a "small" increase in the value taken by the factor $\overline{f} \in F$.

Thus, $\frac{\partial \widetilde{\omega}}{\partial \overline{f}_j}$ is numerically estimated by computing the associated incremental ratio as follows:

$$\frac{\partial \widetilde{\omega}}{\partial \overline{f}_j} \cong \frac{\widetilde{\omega}\left(\overline{f}_j + \Delta\overline{f}; \left\{f_j \in F : f \neq \overline{f}\right\}, \overline{\overline{\theta}}_0\right) - \widetilde{\omega}\left(\left\{f_j \in F\right\}, \overline{\overline{\theta}}_0\right)}{\Delta\overline{f}; \forall \overline{f} \in F; \forall j \in J} \tag{14}$$

By analyzing the statistical distribution of $\frac{\partial \widetilde{\omega}}{\partial \overline{f}_j}$, some interesting insights emerge. For example, a positive mean value of $\frac{\partial \widetilde{\omega}}{\partial \overline{f}_j}$ (computed across the dataset) indicates that an increase in $\overline{f} \in F$ leads to an average increase in the predicted value of the target variable tar_j. Conversely, a negative mean value of $\frac{\partial \widetilde{\omega}}{\partial \overline{f}_j}$ indicates that an increase in $\overline{f} \in F$ could lead to an average decrease in the predicted value of the target variable tar_j.

Finally, once the ANN frequency and severity models have been fitted, they are combined to return a risk prediction model according to eqn. (15).

$$R_{lt} = \sum_{s \in S(l)} R_{slt} = \sum_{s \in S(l)} \widetilde{\omega}((\{e_{slt} \in E\}, \{x_{slt} \in X\}), \overline{\overline{\theta}}_0) * \omega(\{y_{slt} \in Y\}, \overline{\overline{\theta}}_0); \forall l \in L; \forall t \in T \tag{15}$$

2.3. STEP 2 – Real-time explanatory generation of factors' values

Once the risk prediction model has been fitted to the historical dataset, it could be leveraged by TAs/PTCs to forecast in real-time the fare evasion risk on the bus transit network during upcoming time windows to implement management actions. Note that the method is general, as one can make forecasts either for the following periods in time (e.g., the risk in the soft period of the morning from the peak periods in the early morning) and the period with specific characteristics (e.g., the risk in the soft period of the morning from the morning soft periods in the dataset).

Thus, a procedure to generate predictor values could be defined in several ways. On the one hand, if adequate technology is available, frequency and exposure data may be collected daily by inspection logs and automatic passenger counting; conversely, severity data can be collected by on-board digital video cameras while respecting privacy concerns. On the other hand, if this technology is not available, the former data can be generated from historical data by Monte Carlo methods, as described in the following. More formally, let:

- $\check{E}$ be the set of generated exposure factors and $\check{e}_{slt} \in \check{E}$ the value of the specific predictor generated for the homogeneous segment $s \in S(l)$ in time window $t \in T$.
- $\check{X}$ be the set of generated frequency predictors and $\check{x}_{slt} \in \check{X}$ the value of the specific predictor generated for the homogeneous segment $s \in S(l)$ in time window $t \in T$.
- $\check{Y}$ be the set of generated severity predictors and $\check{y}_{slt} \in \check{Y}$ the value of the specific predictor generated for the homogeneous segment $s \in S(l)$ in time window $t \in T$.
- $\psi_{\check{e},s,l,t}$, $\psi_{\check{x},s,l,t}$ and $\psi_{\check{y},s,l,t}$ be the probability density functions of the predictors $\check{e}_{slt} \in \check{E}$, $\check{x}_{slt} \in \check{X}$, and $\check{y}_{slt} \in \check{Y}$, respectively, for the homogeneous segment $s \in S(l)$ and time window $t \in T$.
- $\Psi_{\check{e},s,l,t} \overset{\text{def}}{=} \int \psi_{\check{e},s,l,t} d\check{e}_{slt}$, $\Psi_{\check{x},s,l,t} \overset{\text{def}}{=} \int \psi_{\check{x},s,l,t} d\check{x}_{slt}$ and $\Psi_{\check{y},s,l,t} \overset{\text{def}}{=} \int \psi_{\check{y},s,l,t} d\check{y}_{slt}$ be the cumulative distribution functions of the predictors $\check{e}_{slt} \in \check{E}$, $\check{x}_{slt} \in \check{X}$, and $\check{y}_{slt} \in \check{Y}$, respectively, for the homogeneous segment $s \in S(l)$ and time window $t \in T$.
- $n_{\check{e},s,l,t} \in [0;1]$, $n_{\check{x},s,l,t} \in [0;1]$ and $n_{\check{y},s,l,t} \in [0;1]$, be the random numbers generated for the prediction of $\check{e}_{slt} \in \check{E}$, $\check{x}_{slt} \in \check{X}$, and $\check{y}_{slt} \in \check{Y}$, respectively, for the homogeneous segment $s \in S(l)$ in time window $t \in T$.

In this procedure, each predictor is interpreted as a random variable, and then an appropriate probability function is built from the historical dataset. Each distribution function depends not only on the time window $t \in T$, but also on the homogeneous segment $s \in S(l)$ of route $l \in L$. For example, the user's age could be a significant severity predictor of fare evasion. On the one hand, the distribution of age varies among different time windows (e.g., morning peak hour, midday peak hour, evening peak hour, off-peak hour), because different user categories use the service (e.g., young students and older workers might be prevalent in the morning and the evening peak hours, respectively). On the other hand, the age distribution also varies among the different routes and segments along the bus network (e.g., young students might prevail on routes that pass near schools, while older workers might be more frequent on routes that cross industrial or commercial areas).

For each predictor, bus route, homogeneous segment and time window, a random number obeying a uniform distribution, is generated in the interval $[0;1]$ (i.e., $n_{\check{e},s,l,t}$, $n_{\check{x},s,l,t}$ or $n_{\check{y},s,l,t}$). Then, the value of the predictor (i.e., $\check{e}_{slt}$, $\check{x}_{slt}$ or $\check{y}_{slt}$) is determined by inverting the associated cumulative probability function (i.e., $\Psi_{\check{e},s,l,t}$, $\Psi_{\check{x},s,l,t}$ or $\Psi_{\check{y},s,l,t}$), according to Eqns. (16), (17) and (18), respectively.

$$\check{e}_{slt} = \Psi^{-1}{}_{\check{e},s,l,t}\left(n_{\check{e},s,l,t}\right);\ \forall \check{e}_{slt} \in \check{E}; \forall s \in S(l); \forall l \in L; \forall t \in T; \tag{16}$$

$$\check{x}_{slt} = \Psi^{-1}{}_{\check{x},s,l,t}\left(n_{\check{x},s,l,t}\right);\ \forall \check{x}_{slt} \in \check{X}; \forall s \in S(l); \forall l \in L; \forall t \in T; \tag{17}$$

$$\check{y}_{slt} = \Psi^{-1}{}_{\check{y},s,l,t}\left(n_{\check{y},s,l,t}\right);\ \forall \check{y}_{slt} \in \check{Y}; \forall s \in S(l); \forall l \in L; \forall t \in T; \tag{18}$$

2.4. STEP 3 – Real-time risk prediction

Once the predictor values have been generated for upcoming $t \in T$, for each route $l \in L$, the risk index (denoted as $\check{R}_{lt}$) is predicted in real-time through the model fitted in STEP 1, by summing the risk $\check{R}_{slt}$ on each segment $s \in S(l)$ and route $l \in L$:

$$\check{R}_{lt} = \sum_{s \in S(l)} \check{R}_{slt} = \sum_{s \in S(l)} \tilde{\omega}\left(\left(\left\{\check{e}_{slt} \in \check{E}\right\}, \left\{\check{x}_{slt} \in \check{X}\right\}\right), \bar{\bar{\theta}}_0\right) * \tilde{\omega}\left(\left\{\check{y}_{slt} \in \check{Y}\right\}, \bar{\bar{\theta}}_0\right); \forall l \in L; \forall t \in T; \quad (19)$$

2.5. STEP 4 – Real-time routes and links ranking

Next, the routes and links are assessed using a five-level risk scale, which accounts for the quartile distribution of the fare evasion risk index predicted for the upcoming $t \in T$.

More precisely, the route risk value and the link risk value (i.e., $\check{R}_{lt}$ and $\check{R}_{slt}$) are arranged in a growing order. Next, the $Q1_t$, $Q2_t$, $Q3_t$ quartiles (i.e., the 25th, 50th and 75th percentiles, respectively), along with the interquartile range (IQR_t), i.e., the difference between $Q3_t$ and $Q1_t$, are computed. Thus, the routes and the links are ranked accordingly, but there is no need to adopt these specific ranges.

Table 2 illustrates the five-level ranking scale and specifies the lower and upper boundaries for each risk level. It is noteworthy that this scale is temporal dependent, as the lower and the upper bounds may be different for each upcoming time window $t \in T$.

2.6. STEP 5 – Real-time data visualization

Once the risk ranking has been established, charts and maps are built to offer lucid representations of the outcomes for the upcoming time window $t \in T$. Charts serve as a straightforward and comprehensible dashboard for the ranking of routes based on their predicted fare evasion risk for each $t \in T$. Consequently, managers can gain a clear overview of how fare evasion risk will impact the entire network during the upcoming time windows. On the other hand, maps illustrate the ranking of fare evasion risk for individual segments of the route; each section is color-coded to reflect the anticipated risk level for fare evasion during the upcoming time window $t \in T$. These maps are generated using GIS and are easy to understand, to support managers in identifying the areas of the route that will demand special attention during the upcoming $t \in T$.

2.7. STEP 6 – Real-time deterrence strategies

Next, the maps created in the previous step are exploited to plan in real-time deterrence strategies. Among these strategies, inspection and messaging activities during the upcoming time window $t \in T$ are planned for those segments $s \in S(l)$, which are predicted to fall into the two highest levels, i.e., R_1 and R_2, respectively. More precisely, as for segments in R_1, inspectors will be sent by managers to check passenger tickets. These measures are anticipated to have a substantial impact on curbing fare evasion, because they increase users' perceptions on the certainty of getting caught (e.g., Smith and Clarke, 2020). Additionally, as noted by Barabino et al. (2023), a higher perception of inspection frequency corresponds to a reduced inclination to engage in fare evasion.

Since the number of inspectors is not sufficient to visit all high-risk segments, as for segments in R_2, a less resource-consuming strategy based on messaging actions is proposed. This strategy is based on the evidence that beliefs about the occurrence of fare evasion (e.g., the social norm) are among the main determinants of fare evasion (e.g., Celse and Grolleau, 2023; Ayal et al., 2021). Particularly, Ayal et al. (2021) noted that passengers were less likely to evade the fare when they were informed about the low rate of evasion through social norm messaging while simultaneously exposed to a control eye-cues poster. Drawing from this experience, this framework proposes communication actions that display messages on-board. More precisely, when a vehicle travels along a R_2 route segment, a message stating, e.g., "*Along this bus route, more than pg% of travellers own a regular ticket or pass*" will be displayed on a Variable Message Panel (VMP), together with watching eye cues. Furthermore, a second message stating, "*This bus is a video surveillance area,*" will be displayed simultaneously on the VMPs to increase users' feeling of "being observed" by others.

It is noteworthy that the stated percentage of passengers with a regular ticket (denoted as *pg%)* might be greater than the real one because the latter is likely to be too low on R_2 routes to discourage unethical behaviour. However, the adoption of a misleading *pg%* does not raise concerns if passengers are not supposed to know the real percentage of regular travelers. Nevertheless, to avoid passengers doubting the truthfulness of the messages, the *pg%* value should not be constant in time and space. Thus, for each segment $s \in S(l)$ and time window $t \in T$, an effective strategy might be to set the *pg%* value by randomly sorting a number within appropriate ranges.

Table 2
Definition of the risk ranges of fare evasion during each upcoming time window $t \in T$.

Level	Range values		Colour
	Lower bound	Upper bound	
R_1	$(Q3_t + 1.5\ IQR_t)$	MAX_t	
R_2	$Q3_t$	$(Q3_t + 1.5\ IQR_t)$	
R_3	$Q2_t$	$Q3_t$	
R_4	$Q1_t$	$Q2_t$	
R_5	$(Q1_t\text{-}1.5\ IQR_t)$	$Q1_t$	

3. Experiment in an Italian metropolitan area

Research context

The comprehensive methodology was implemented in MATLAB and tested in the metropolitan area of Cagliari, Italy, which boasts a population of approximately 0.4 million residents. The study focused on the transit network of CTM, the largest PTC in this area. CTM manages its transit services with a fleet of 271 between buses and trolley, providing 35.8 M+trips annually (CTM, 2023). We mined the most pertinent data for the model from PTC records gathered in 2013, 2014, and 2015.[2] It is worth noting that these data are easily manageable, provided that PTCs routinely collect fare evasion monitoring data and consistently gauge passenger perceptions regarding ticket inspections.

3.2. Data and sources

Data on final, intermediate, and risk-exposure factors were mined from the dataset presented in Barabino et al. (2023). More specifically, various data sources were considered.

Final factors on the fare evasion **frequency** were mined from the fare inspection logs, commonly conducted on-board to prevent service disruptions. Inspection logs report the number of passengers inspected and fined for each route and time window. The frequency was modeled as the total number of fined, eluded, and escaped passengers in a time window. **Final factors** on fare evasion **severity** were gathered by administering onboard surveys to passengers by asking for their self-reported fare evasion. The measure of severity was defined according to the revenue loss. Specifically, individuals with lacking or unvalidated tickets were classified as severe fare evaders, while minor irregularities (e.g., just expired tickets) were not categorized as severe offenders. In total, data for evaluating the frequency and severity of fare evasion were gathered over a span of at least two weeks, encompassing both weekdays and weekends, from 07:00 to 19:00, which amounted to more than 19,000 records.

Data on **intermediate factors** related to **context**, **organization,** and **vehicle** were gathered from the inspection log files. They mainly include temporal elements, system planning and design attributes, inspection patterns, and vehicular characteristics. Data on **intermediate factors** on **passengers** were gathered by administering a four-section questionnaire, including sociodemographic characteristics, travel behavior characteristics, situational factors, and personality traits.

Tables A1 and A2 in the Appendix provide descriptive statistics on the intermediate and final factors of the fare evasion datasets adopted to set up the frequency and severity model proposed in this research.

Data on **risk exposure factors** were gathered from the PTC's dataset. These factors encompass passenger volumes, which are determined by ride checks on a route-by-route basis, the length of the inspection path (defined as the distance between the bus stops where passengers alight and board during each inspection run), and the level of occupancy (defined as the ratio of passengers to the vehicle's capacity).

Finally, in this study, the segmentation between consecutive stops was adopted because of:

- Exposure Term Collection: Passenger numbers are gathered at consecutive bus stops, allowing for a better representation of exposure heterogeneity.
- Industry Practice: PTCs widely use segmentation between consecutive stops as the standard unit of analysis, and data and performance metrics are often reported in this manner (e.g., Barabino, et al., 2016; Ceder, 2016; Pili et al., 2019). This choice enhances the comprehensibility of results within the PTCs' community.
- Experiment Replicability: Stating segments as the "leg between two consecutive bus stops" eliminates any vagueness that might derive from segmentation based on different features.

[2] The analysis of more recent data is not feasible because of CTM's confidentiality policy.

4. Results and discussion

4.1. Fare evasion frequency model

According to STEP 1, the frequency model was fitted by adopting an ML technique based on a regression ANN, as indicated in eqns. from (3) to (12). To identify training (i.e., *TR*), validation (i.e., *VA*) and test (i.e., *TE*) subsets, respectively, a splitting ratio of 70 %, 15 % and 15 % was used (Flach, 2012). By changing the training procedure and the hidden layer's number of neurons, many calibrations were made to enhance the fitting. After tuning, a network with 10 perceptrons in the hidden layer, trained using the Levenberg-Marquardt algorithm, is selected as it provides the greatest data fitting. The training-validation process was stopped at epoch 13, which returned the value of the minimum cost function (i.e., *MSE*) on the validation subset (i.e., *VA*) as shown in Fig. 3. Weights and biases of the hidden layer perceptrons (i.e., w_{pf} and b_p) and for the output layer perceptron (i.e., $\overline{w_p}$ and $\overline{b}$) are provided in Tables A3 and A4 in the Appendix, respectively.

The fitting and prediction performances of the model were tested through the metrics in Table 1 for the frequency response variable. Fig. 4 shows that the model adequately describes frequency data, since (tar_j, $\widetilde{tar_j}$) points are located close to the regression lines in all subsets. The other metrics (Table 3) corroborate the goodness of fit and prediction of the ANN model, which accounts for more than 75 % of the variability in the data (as shown by $R > 0.75$). Furthermore, the results of ANN look conservative because it slightly overestimates the number of fare evasion events (as shown by $\pi \gtrsim 1$).

Next, the Permutation Feature Importance and the partial derivative of each factor in the frequency model (i.e., PFI_f and $\widetilde{\frac{\partial \omega}{\partial f_j}}$) were computed according to eqns. (13) and (14). The outcomes are reported in Table 4 and Fig. 5.

To summarize, an exposure factor (i.e., passenger volume) and two organization factors (i.e., inspection types B and A) showed the strongest contributions in explaining the fare evasion frequency, being the values of PFI_f significantly higher than those associated with the other factors (Fig. 5).

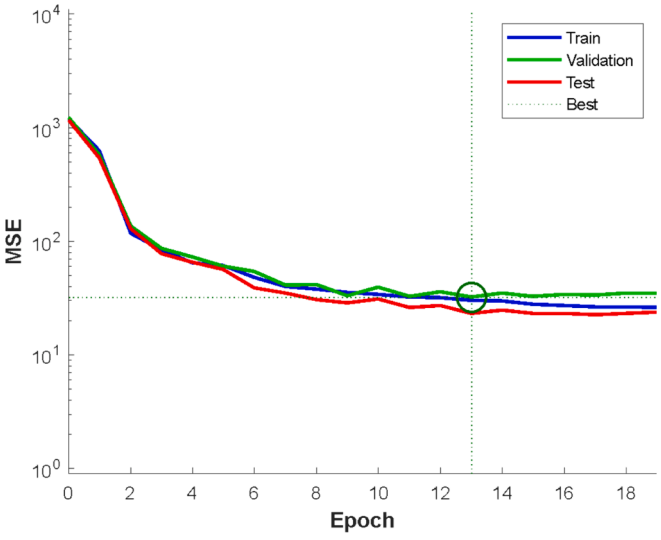


Fig. 3. Cost function for the frequency model (i.e., *MSE*) as a function of training epoch. In epoch 13, the best validation performance was attained as indicated by the green circle. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

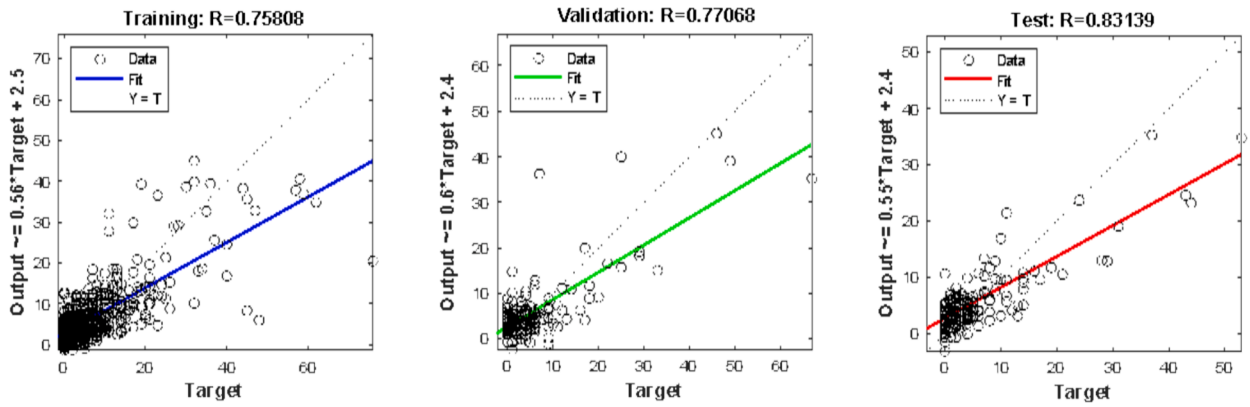


Fig. 4. Regression Plots (RPs) for the frequency model.

Table 3
Performance metrics for the frequency model.

Metric	Value on *TR*	Value on *VA*	Value on *TE*
R	0.76	0.77	0.83
MAE	3.39	3.78	3.33
RMSE	5.49	5.67	4.80
CoV	0.994	0.977	0.892
π	1.011	1.032	1.000

Focusing on each factor separately, the volume of passengers carried during a specific time window (i.e., *Pass_vol*) was the most important predictor, having a PFI_f greater than 70 and a positive partial derivative. This finding corroborates Barabino et al. (2023), who proved that the larger the exposure in terms of passenger volume, the more frequent the fare evasion. Moreover, the level of occupancy looks less important, unlike that observed by some authors (i.e., Lee, 2011; Guarda et al., 2016; Cantillo et al., 2022). Next, all the factors related to targeted inspection type followed, with categories B and A ranking second and third, respectively, and PFI_f values at least twice as large as C and D, which ranked fourth and fifth, respectively. These findings are consistent with Barabino et al. (2023), in which the presence of checkers (A and B) results in a higher rate of fare evasion compared to the presence of inspectors alone (C and D). Although these results look unexpected, they might be explained by crafty evaders who are constantly on guard to elude inspectors and could travel along alternative routes where the inspection is not yet targeted (e.g., Assaf and Van den Broeck, 2022). Moreover, the positive partial derivatives of all targeted inspection types (i.e., from A to D) seem to contrast with the assumption that highly targeted inspections would reduce fare evasion (e.g., Boyd, 2020). Hence, random inspections look more effective than those following targeted paths in lowering fare evasion, endorsing real-world results in Barabino et al. (2023) and laboratory experiments in Dai et al. (2017). The last experiment showed that targeted inspections reduce fare evasion rates by around 50 % in the first four days and result in a burst of fraud during subsequent periods with no inspection. Next, a standard patrol of two inspectors (i.e., *Standard*) rated sixth, and a negative value was found for its partial derivatives. The relatively high importance of this predictor and its lowering effect on fare evasion frequency are expected results. Indeed, previous studies showed that a two-inspector patrol is more effective than three or more inspectors in dealing with fare evasion (e.g., Barabino et al., 2023; Barabino and Salis, 2019). Finally, the time period during which the inspection was performed (i.e., 7:30–9:30 and 12:30–14:30) placed seventh and eighth, respectively, and both presented negative partial derivatives. Interestingly, 7:30–9:30 was found to have the highest PFI_f among the time period-related

Table 4
Permutation feature importance and ranking of each factor for the ANN frequency model.

Factor		Abbreviation	PFI	PFI Ranking	Mean partial derivative*	Probability that $\frac{\partial\tilde{\omega}}{\partial\tilde{f}_j} \geq 0$ **
Exposure						
Passenger volume		*Pass_vol*	70.281	1	0.001	100.00 %
Level of occupancy		*Occ_Lev*	0.731	11	4.216	100.00 %
Context						
Time period	7:30–9:30	*Rush_Morn*	1.506	7	−1.354	35.62 %
	12:30–14:30	*Rush_Half*	1.296	8	−0.592	32.54 %
	17:30–19:30	*Rush_Even*	0.978	10	1.162	80.07 %
Organization						
Inspection Type	A	*A*	16.062	3	4.036	68.93 %
	B	*B*	24.214	2	6.302	98.75 %
	C	*C*	5.772	4	4.350	85.70 %
	D	*D*	2.672	5	5.199	93.85 %
Level of enforcement	Standard	*Standard*	2.131	6	−1.424	16.19 %
Vehicle						
Capacity		*Capacity*	0.650	12	0.030	89.70 %
Length	Medium	*Medium*	0.615	13	2.489	89.17 %
	Long	*Long*	1.023	9	0.524	67.83 %

*Computed as the mean of $\frac{\partial\tilde{\omega}}{\partial\tilde{f}_j}$ values in the dataset. ** Computed as the ratio between the number of records in which $\frac{\partial\tilde{\omega}}{\partial\tilde{f}_j} \geq 0$ and the total number of records in the dataset.

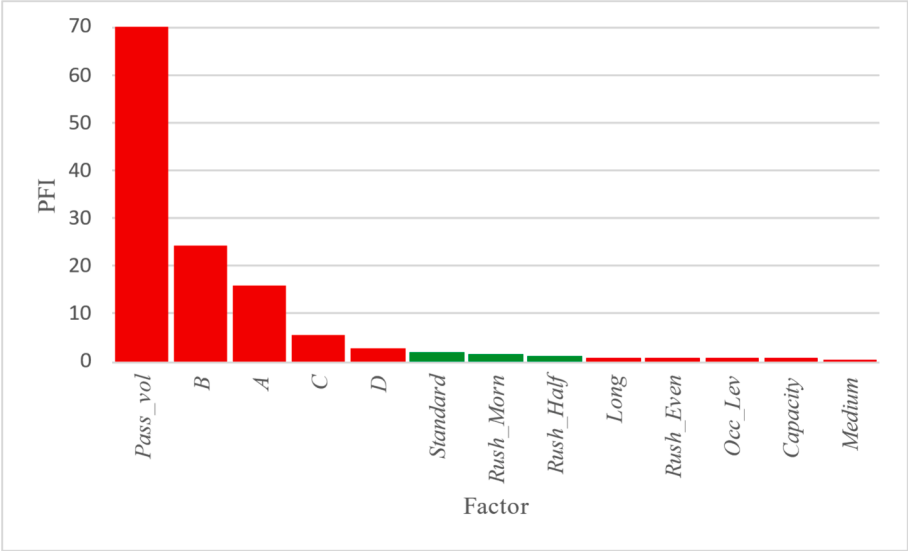


Fig. 5. Bar chart representing the PFI_f of each factor for the ANN frequency model. Factors are sorted from most important to least important. Red and green bars indicate that an increase in the factor "on average" leads to an increase or decrease in fare evasion frequency, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

factors. This is an expected outcome because Barabino et al. (2023) found that rush morning was the time period-related factor associated with the greater coefficient in terms of absolute value. Indeed, they showed that the lowest frequency of fare evasion occurred during the morning rush hours. Although these findings were gathered in a midsized European city, they are consistent with Lee (2011) and Reddy et al. (2011) in big cities of San Francisco and New York (USA) as well as in Guarda et al. (2016) and Cantillo et al. (2022) in the big city of Santiago (Chile). The other variables look less important.

4.2. Fare evasion severity model

According to STEP 1, the severity model was set up by fitting a classifier ANN, as specified in eqns. from (3) to (12). Like the frequency model, a splitting ratio of 70 %, 15 % and 15 % was used for *TR*, *VA* and *TE* subsets, respectively. After several attempts at improvement, a network with 40 perceptrons in the hidden layer and trained by the Levenberg-Marquardt algorithm was chosen. The training-validation process was terminated at epoch 29, corresponding to the minimum *CE* on *VA* (Fig. 5). Weights and biases of the hidden layer perceptrons (i.e., w_{pf} and b_p) and the output layer perceptron (i.e., $\overline{w_p}$ and $\overline{b}$) are provided in Tables A5 and A6 in the Appendix, respectively.

Next, the metrics in Table 1 for the severity response variable were considered to test the fitting and prediction performances of the model. On the one hand, the *CMs* (Fig. 6) and the *TPR* metric (Table 5) indicated that the model slightly underestimates the number of severe fare evasion events, probably due to the strong skewing of the dataset toward no severe events. All in all, the model was found to be sufficiently accurate, as 90 %+ of the observed data were correctly predicted. Indeed, *ACC* values greater than 0.9 were attained not only for the *TR* and *VA* datasets, but also for the *TE* dataset (Table 5), which had no influence on training and, therefore, provides an unbiased measure of model performance. See (Fig. 7).

Afterwards, the Permutation Feature Importance and the mean partial derivative of each factor in the frequency model (i.e., PFI_f and $\frac{\widetilde{\partial \omega}}{\partial f_j}$) were computed according to eqns. (13) and (14). Results are summarized in Table 6 and in Fig. 8.

Generally speaking, a personality factor (i.e., honesty), a situational

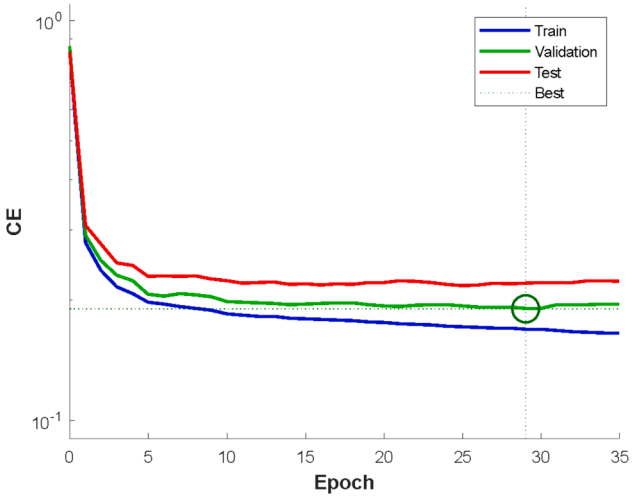


Fig. 6. Cost function for the severity model (i.e., *CE*) as a function of training epoch. In epoch 29, the best validation performance was attained as indicated by the green circle. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 5
Performance metrics for the severity model.

Metric	Value on *TR*	Value on *VA*	Value on *TE*
TPR	0.370	0.339	0.323
TNR	0.984	0.980	0.983
ACC	0.932	0.922	0.917

factor (i.e., fines in the past), two sociodemographic characteristics (i.e., two different educational qualification levels), and a travel behavior characteristic (i.e., trip purpose) demonstrated the highest impacts in explaining the fare evasion severity (Fig. 8).

Focusing on each factor, honest users (i.e., *Hon_Y*) turned out to be the most important predictor, with a PFI_f of about 0.1 (significantly higher than the PFI_f all other factors). Moreover, a negative partial derivative indicated a decreased severity, which is an expected outcome.

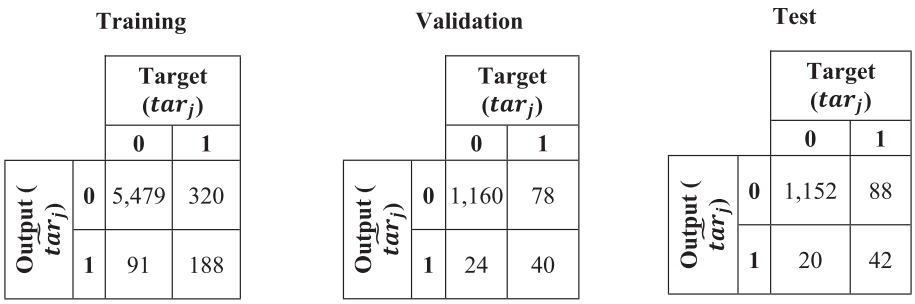


Fig. 7. Confusion Matrixes (CMs) for the severity model.

Table 6
Permutation feature importance and ranking of each factor for the ANN severity model.

Factor		Abbreviation	PFI	Rank	Mean partial derivative*	Probability that $\frac{\partial\widetilde{\omega}}{\partial\widetilde{f}_j} \geq 0$**
Sociodemographic characteristics						
Gender	Male	*Gen_M*	0.006	8	0.029	90.29 %
Age class	Above 65	*Above_65*	0.001	15	−0.027	20.99 %
	51–65	*51–65*	0.006	9	−0.046	2.63 %
	36–50	*36–50*	0.001	16	−0.027	7.32 %
	26–35	*26–35*	<0.001	19	−0.026	16.62 %
	18–25	*18–25*	0.007	7	0.028	78.38 %
Educational qualification	Upper school graduate	*Upper_Sc*	0.020	4	−0.059	1.31 %
	Middle school graduate	*Middle_Sc*	0.021	3	−0.036	22.15 %
	Middle school not graduated	*Middle_Sc_N*	0.005	10	−0.002	49.37 %
Employment	Worker	*Worker*	<0.001	17	0.023	87.55 %
Reason to use the bus	Other	*Other_Use_Bus*	0.003	12	−0.022	13.33 %
Travel behaviour characteristics						
Trip purpose	Work	*Work_Trips*	0.002	13	−0.011	33.43 %
	Study	*Study_Trips*	0.020	5	−0.043	26.53 %
In-vehicle time	More than 15 min	*Time_More_15*	0.005	11	−0.038	5.99 %
Quality rating	Satisfied	*Statisf_Y*	0.002	14	−0.020	37.37 %
Situational factors						
Perceived inspection frequency	From 6 to 10	*Insp_Freq_6_10*	<0.001	20	−0.020	12.21 %
	From 1 to 5	*Insp_Freq_1_5*	<0.001	18	−0.009	48.47 %
Know the amount of the fine	Yes	*Know_Fine_Y*	0.007	6	0.035	95.94 %
Fined in the past	Yes	*Fine_Past_Y*	0.032	2	0.081	99.91 %
Personality factors						
Honest	Yes	*Hon_Y*	0.100	1	−0.157	0.00 %

*Computed as the mean of $\frac{\partial\widetilde{\omega}}{\partial\widetilde{f}_j}$ values in the dataset. ** Computed as the ratio between the number of records in which $\frac{\partial\widetilde{\omega}}{\partial\widetilde{f}_j} \geq 0$ and the total number of records in the dataset.

Specifically, honesty presented one of the highest negative coefficient estimates (in terms of absolute value), and being honest was proved to reduce the likelihood of a severe event by about 9 times compared to being dishonest. Moreover, this result confirms previous *post hoc* segmentation studies (e.g., Barabino and Salis, 2023; Currie and Delbosc, 2017; Delbosc and Currie, 2016).

Next, the history of prior fines (i.e., *Fine_Past_Y*) emerged as the second-ranking factor, and a positive value was found for its partial derivative. This confirmed previous research (e.g., Barabino et al., 2023; Barabino and Salis, 2023; Barabino et al., 2022a; Barabino and Salis, 2020). In contrast, the findings of Bucciol et al. (2013) indicated that prior fines were insignificant in influencing fare evasion behavior. Furthermore, unlike the results presented by Dai et al. (2017), our research shows that passengers previously fined may attempt to offset their losses by evading fares once again.

Next, the educational qualification of the users followed with the categories “middle school graduate” (i.e., *Middle_Sc*) and “upper school graduate” (i.e., *Upper_Sc*) ranking third and fourth, respectively. Both factors presented a lower impact on fare evasion severity. These results confirm the conclusions of some previous works while contrasting others. Indeed, Barabino et al. (2023) found that passengers with a lower educational level significantly boost the severity of fare evasion, and Bucciol et al. (2013) concluded that this variable has a negligible effect. Hence, more research is needed to better understand the influence of this predictor. Finally, the trip purpose (i.e., *Study_Trips*) rated fifth and presented a negative partial derivative. This is an expected finding, as it has been demonstrated that severe fare evasion is more likely to occur for other trip motivations, e.g., when passengers are making occasional trips (e.g., Barabino et al., 2023). Interestingly, the perception of the inspections does not strongly influence the severity model. This fact could partially explain why past research did not observe clear trends. Indeed, while some studies showed how this perception affects fare evasion (e.g., Barabino et al., 2015; Cools et al., 2018; Porath and Galilea, 2020) and agree with the research on deterrence (e.g., Clarke et al., 2010), others disagree (e.g., Bucciol et al., 2013; Barabino and Salis, 2023).

4.3. Real-time risk prediction and ranking and data visualization

Next, according to STEP 2, in this case study, a Monte Carlo method

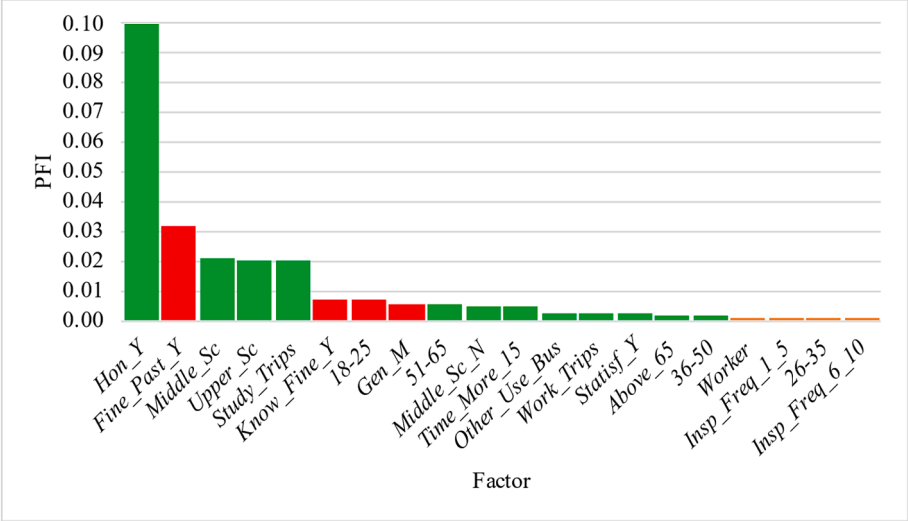


Fig. 8. Bar chart representing the permutation feature importance of each factor for the ANN severity model. Factors are sorted from most to least important. Red and green bars indicate that an increase in the factor "on average" leads to an increase or decrease in fare evasion severity, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

generated the explanatory factors for each upcoming temporal window ($t \in T$). The duration of the time windows was chosen to ensure uniformity in the typology of the period within each window (i.e., soft hours or rush hours). As a result, the duration of each window was different according to the type of the period (i.e., two hours for rush and 18 h for soft). To run the Monte Carlo simulation, the probability density functions associated with each predictor were preliminary defined according to historical data. More details about these functions are provided in Appendix (Section 3).

Next, according to STEP 3, for each route, the risk index was real-time predicted for each window by eqn. (19), where the explanatory factors generated in STEP 2 were assumed as input. Next, according to STEP 4, the risk predictions were ranked for each time window by the five-level risk scale based on the distribution quartiles. In what follows, results are presented considering segments rather than routes to provide greater granularity. The case study time windows are chosen to cover all time periods and type of day combinations. The outcomes are reported in Table 7.

On the one hand, in the case of *Weekday – Other intervals* and *Weekend – Other intervals*, the upper bounds of the maximum risk level are comparable. On the other hand, during *Weekday – 7:30–9:30*, the upper bound of the maximum risk level was significantly higher than *Weekend – 7:30–9:30*. This outcome might depend on the higher values of exposure during the former rather than the latter, due to the overlap of work and study activities. Next, in *Weekend – 12:30–14:30* and the *Weekend – 17:30–19:30*, the upper bounds of the maximum risk level were found to be notably higher than those of the *Weekday – 12:30–14:30* and the *Weekday – 17:30–19:30*. This is probably due to the joint effect of the larger number of leisure/tourism trips with the lower rate of inspections performed during the weekend.

Next, according to STEP 5, maps were plotted to clearly represent risk predictions. They help visualize the ranking of routes at a disaggregated level and, thus, better appreciate the variability of risk along each segment of the entire transit network (Fig. 9).

Fig. 9 shows that the fare evasion risk has a strong spatial and temporal dependence. Consider as an illustrative example, *Weekday –*

Table 7
Ranking scale for link risk ranges during the eight different time windows taken as illustrative examples.

Level	*Weekday - 7:30-9:30*		*Weekday – 12:30-14:30*		*Weekday – 17:30-19:30*		*Weekday – Other intervals*		*Colour*
	Low.	*Upp.*	*Low.*	*Upp.*	*Low.*	*Upp.*	*Low.*	*Upp.*	
R_1	0.38	1.25	0.75	1.43	1.34	3.16	0.87	1.57	
R_2	0.20	0.38	0.39	0.75	0.74	1.34	0.47	0.87	
R_3	0.13	0.20	0.24	0.39	0.48	0.74	0.29	0.47	
R_4	0.08	0.13	0.15	0.24	0.33	0.48	0.21	0.29	
R_5	0.00	0.08	0.00	0.15	0.00	0.33	0.00	0.21	
Level	*Weekend - 7:30-9:30*		*Weekend – 12:30-14:30*		*Weekend – 17:30-19:30*		*Weekend – Other intervals*		*Colour*
	Low.	*Upp.*	*Low.*	*Upp.*	*Low.*	*Upp.*	*Low.*	*Upp.*	
R_1	0.31	0.94	0.69	1.81	1.17	3.49	0.77	1.55	
R_2	0.15	0.31	0.33	0.69	0.63	1.17	0.41	0.77	
R_3	0.08	0.15	0.19	0.33	0.43	0.63	0.25	0.41	
R_4	0.03	0.08	0.10	0.19	0.27	0.43	0.16	0.25	
R_5	0.00	0.03	0.00	0.10	0.00	0.27	0.00	0.16	

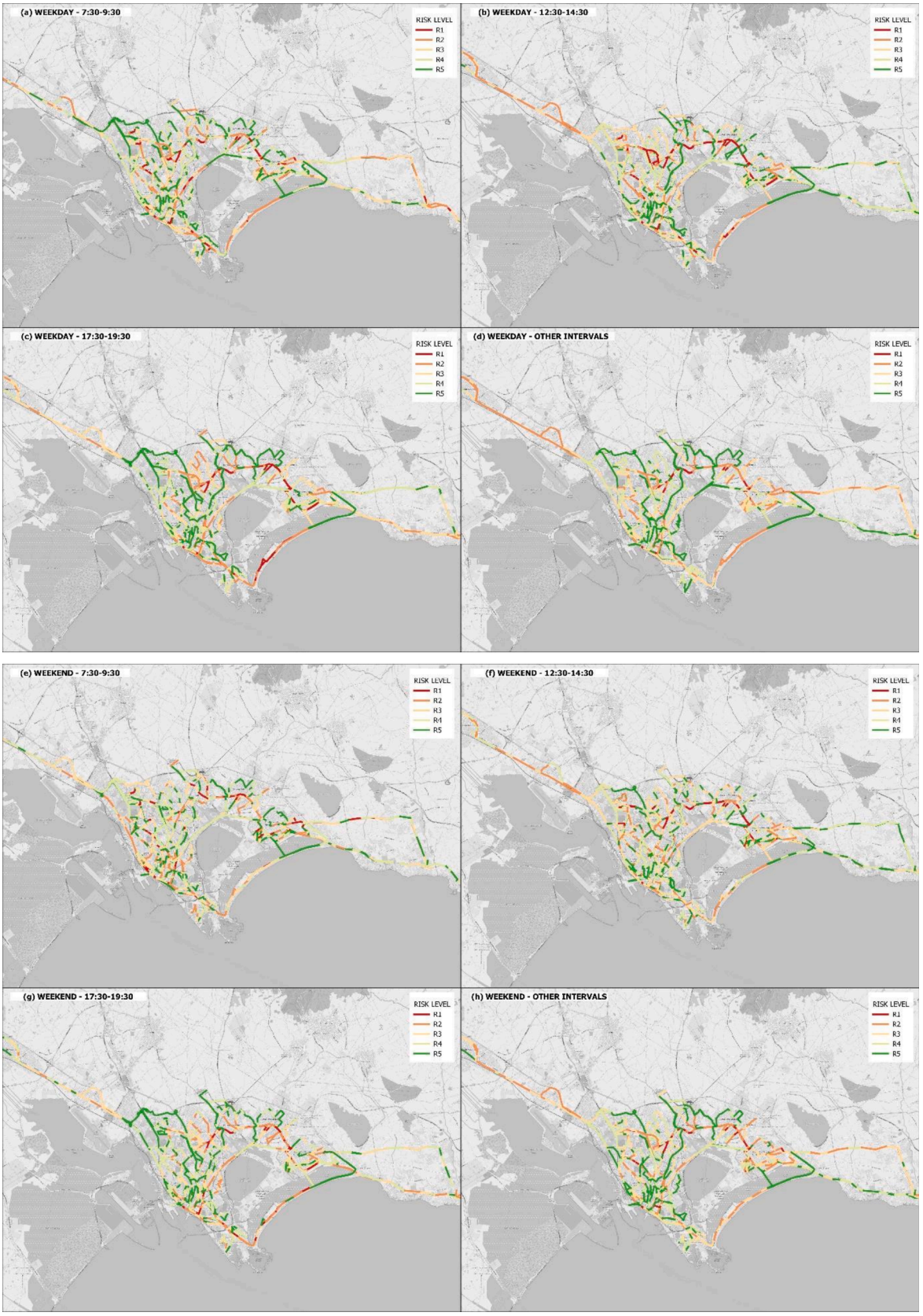


Fig. 9. Fare evasion risk maps on the overall network during the eight different time windows taken as illustrative examples: (a) Weekday – 7:30–9:30; (b) Weekday – 12:30–14:30; (c) Weekday – 17:30–19:30; (d) Weekday – Other intervals; (e) Weekend – 7:30–9:30; (f) Weekend – 12:30–14:30; (g) Weekend – 17:30–19:30; (h) Weekend – Other intervals.

Table 8
Simulated deterrence strategies. For each scenario, the percentage of segments that would be affected by the different strategies is shown.

Strategy	***Weekday – 7:30–9:30***	***Weekday – 12:30–14:30***	***Weekday – 17:30–19:30***	***Weekday – Other intervals***
Inspections	5.46 %	7.84 %	6.48 %	3.12 %
Messaging	19.54 %	17.16 %	18.52 %	21.88 %
Do nothing	75.00 %	75.00 %	75.00 %	75.00 %
Strategy	***Weekend – 7:30–9:30***	***Weekend – 12:30–14:30***	***Weekend – 17:30–19:30***	***Weekend – Other Intervals***
Inspections	6.15 %	5.64 %	5.89 %	3.51 %
Messaging	18.85 %	19.36 %	19.11 %	21.49 %
Do nothing	75.00 %	75.00 %	75.00 %	75.00 %

12:30–14:30 (Fig. 9 (b)) and *Weekday – 17:30–19:30* (Fig. 9 (c)) maps. In *Weekday – 12:30–14:30*, the area with the higher density of segments and the maximum risk level (i.e., R1) is the northern part of the city, where there are many activities. Conversely, in *Weekday – 17:30–19:30*, the riskiest segments are along the southern coastline, where leisure attractions prevail. Moreover, the time windows affect the number of segments with a given risk level. Furthermore, in the cases of *Weekday – 12:30–14:30* (Fig. 9 (b)) and *Weekend – 12:30–14:30* (Fig. 9 (f)), the number of R1 segments is sensibly greater in the former than in the latter, probably due to a different distribution of transit user's activities within the city.

4.4. Simulation of real-time risk deterrence strategies

According to STEP 6, real-time risk deterrence strategies are simulated for eight scenarios, one for each time window, to provide an illustrative example of the proposed methodology (Table 8).

Noteworthy, whilst in *Weekday – Other intervals* and *Weekend – Other intervals,* the fractions of segments involved by the more severe strategies are similar, in Weekend – 7:30–9:30 the fraction of inspected segments is quite higher (about 0.7 %) than in Weekday – 7:30–9:30. Conversely, in *Weekday – 12:30–14:30* and *Weekday – 12:30–14:30*, the percentages of inspected segments turned out to be substantially higher than in *Weekend – 12:30–14:30* and *Weekend – 12:30–14:30.* Specifically, this difference exceeded 2 % for the 12:30–14:30 period. At first glance, these results could appear to be incompatible with the risk level bounds. Indeed, there is a low number of segments to be inspected in the time windows when a high risk of fare evasion is predicted and vice versa. Nevertheless, the two outcomes are not contradictory because they are just two sides of the same phenomenon. More specifically, on the one hand (e.g., *Weekend – 12:30–14:30*), there are comparatively few segments with a maximum risk level of fare evasion (i.e., R_1), but the upper bound associated with this level is relatively high. Strictly speaking, there are few but very risky segments. On the other hand (e.g., *Weekday – 12:30–14:30*), the number of segments with the maximum risk level (i.e., R_1) is comparatively larger, but the upper bound with this level is relatively lower (i.e., there are many but less risky segments). Moreover, this quantitative result confirms what is qualitatively highlighted from the risk maps (e.g., Fig. 9 (b) and Fig. 9 (f)). Finally, it is interesting to note that the percentage of segments that would be affected by inspection activities in all time windows is quite comparable to the optimum inspection rate that maximizes the PTC's profit, i.e., 3.4 %-4.0 % (Barabino et al., 2014; Barabino and Salis, 2019).

4.5. Summary of the key findings and fare evasion deterrence strategies

Finally, this section highlights the key findings and translates them into policy actions to reduce fare evasion. These are reported in Table 9, which is self-explicative and structured into four columns: "Risk Component", disaggregating the fare evasion risk in frequency and severity components; "Factor(s)", listing the main variables influencing fare evasion risk according to the PFI values; "Observed Effect", summarizing how each factor impacts fare evasion risk according to the mean partial derivative values; and "Suggested Deterrence Strategy," recommending strategies to address each main risk factor, often involving dynamic real-time methods. By focusing on these refined actions, TAs/PTCs can effectively reduce both the frequency and severity of fare evasion, ensuring a more compliant and fair public transportation system.

Table 9
Summary of the key findings and corresponding fare evasion deterrence strategies.

Risk Component	**Factor(s)**	**Observed Effect**	**Suggested Deterrence Strategy**
Frequency	Passenger Volume	Higher passenger volumes result in more frequent fare evasion events	Implement strategies to manage passenger volumes during peak times and ensure better monitoring and control of the busiest routes. This includes, e.g., deploying inspection teams based on real-time risk predictions.
	Inspection Types	Targeted inspection with checkers (types A and B) generates more evasion than target inspections with inspectors only (C and D). Moreover, random inspections are more effective than targeted ones in lowering fare evasion frequency.	Avoid inspection Types A and B in segments with high fare evasion risk. Vary inspection times and locations (i.e., implement random inspections) to prevent evaders from predicting inspection patterns, e.g., by deploying inspection teams based on real-time risk predictions.
Severity	Honesty	Honest individuals are much less likely to commit severe fare evasion.	Develop programs encouraging honesty and integrity among passengers. This includes community engagement initiatives that promote fare compliance as a social norm. These campaigns might be based on real-time risk predictions and include, e.g., information dissemination by VMPs strategically located in vehicles and at bus stops.
	Educational Level	Higher educational levels (middle and upper school graduates) correlate with lower fare evasion severity.	Launch educational campaigns targeting populations with lower educational levels to raise awareness about the importance of fare compliance. These campaigns might be based on real-time risk predictions, as are done in the case of honesty.
	Trip Purpose	The trip purpose significantly influences the severity of fare evasion. Occasional trips are associated with a greater fare evasion severity than study and work trips.	Intensify inspections during periods when the fraction of occasional trips is larger (e.g., off-peak hours, weekends) by deploying inspection teams based on real-time risk predictions.

Interestingly, many TAs/PTCs worldwide are addressing fare evasion, according to the guidelines in Table 9.

For instance, managing passenger volumes during peak times is crucial. Effective strategies include dynamic deployment of inspection teams, as in New York City (MTA, 2019). Moreover, the city is moving from a "citations first" approach to a "warnings first" approach, i.e., to educate passengers. This choice shows that first-time fare evaders should receive an official warning recorded, prioritizing this measure over the immediate issuance of citations or arrests (Blue-Ribbon Panel – BRP, 2023).

Random inspections have proven more effective than targeted ones. For instance, the towns of Adelaide and San Francisco in California (USA) address fare evasion by unpredictable inspection patterns in terms of times and locations (e.g., Milnes, 2012; SFMTA, 2020).

Promoting honesty among passengers through community engagement initiatives has also been effective. For example, San Francisco has implemented public awareness campaigns to establish fare compliance as a social norm (e.g., SFMTA, 2020). Likewise, Sydney, Australia's efforts in public education on fare compliance serve as a useful model (e.g., Cassidy, 2018). Implementing similar campaigns informed by real-time risk data can effectively raise awareness and promote lawful behavior.

5. Conclusions and research perspective

Nowadays, fare evasion is a major threat to TAs/PTCs. It is an emerging topic for academic research to recommend possible deterrence strategies. The current approach models the fare evasion risk from all factors that could be gathered daily in the activities of fare inspectors and surveying passengers (if any). This study contributes to the literature as follows:

- It investigates fare evasion risk by machine learning methods. Specifically, two ANN models are trained, validated, and tested to assess the frequency and severity of fare evasion to account for the intricate (potentially nonlinear) nexus connecting these components.
- It provides a real-time estimation of risk value for each (segment of a) route and time window of interest, which is ranked accordingly.
- It recommended real-time deterrence strategies based on inspections and messaging actions to lessen fare evasion along the most critical (segments of a) route.
- It shows how to plan accurate fare compliance strategies as opposed to static daily measures in the literature. The performance of fare evasion on (segment of a) route can be diagnosed using simple-to-read control dashes based on charts and maps, which help prioritize interventions on the high-risk evasion routes according to inspection and communications strategies.

To the best of our knowledge, this is the first research using ANNs to predict fare evasion risk in transit networks in real-time. Moreover, this research can foster the assessment and management of the risk of fare evasion because the framework works with any PTC. If new data becomes available, this framework can be employed in any urban context to investigate the effects of various factors on the risk of fare evasion. For instance, TAs/PTCs can schedule accurate randomized inspection paths along the network.

This study indicates some research developments. Because original data were manually collected, each factor and route segment were simulated in any time window. Nevertheless, in the era of strong advancements in technology, all these data could always be available because they can be automatically collected by several tools and/or sensors on board the bus. For instance, frequently, fare evasion data are collected by inspection logs. Although these logs are available daily, in many TAs/PTCs, they are still gathered by paper and pencil, which presents some critical issues (e.g., digital data transfer). However, the opportunity for palmtop devices to support inspection activities is strongly recommended. These devices could reduce the timing of ticket controls, automate the citation process, manage reiterated passenger misbehavior, and quickly return data available for processing (e.g., Egu and Bonnel, 2020). Risk exposure data (i.e., passenger data) could easily be collected by automated passenger counting systems quantifying boarding and alighting passengers. Thus, these data could improve the estimation of the exposure factors, as well as improve the passenger experience by facilitating efficient boarding, enabling more accurate occupancy information, and enhancing the overall quality of public transport services. The severity data used in this study were based on surveying passengers and identifying specific factors. This option could be quite an expensive activity. Conversely, recent computer vision development has adopted pattern recognition in images. Therefore, most factors could be inferred by digital video cameras that are frequently available on vehicles (e.g., Huang et al., 2022). This technology may contribute to automatically inferring passenger factors, thus refining the measure of the severity of fare evasion. Nevertheless, according to the European General Data Protection Regulations (GDPR), collecting such digital characteristics must be implemented with appropriate measures in accordance with data privacy. Finally, as these data are collected by different devices and sensors, their integration should be investigated to develop more refined real-time risk prediction models that should include multi-source data fusion algorithms.

Second, future research should precisely measure the type of evasion and analyze its severity to produce a more accurate evaluation of risk. Furthermore, the exposure factors could be explicitly modeled instead of being part of the frequency model. Moreover, adding the accurate locations of each passenger's boarding bus stop might raise the quality of the fare evasion datasets.

Third, future research should evaluate the predictive capabilities of ANN models compared to more traditional econometric models, specifically for the frequency and severity components of fare evasion risk. Such a comparison could uncover scenarios in which ANN models excel over traditional econometric approaches and vice versa, enhancing the understanding of fare evasion dynamics. Additionally, this method could highlight potential areas for model improvement and innovation.

Finally, exploring new techniques for segmenting transit networks could provide further insights. Current segmentation methods might not fully capture the complex patterns of fare evasion. Studies can better identify high-risk areas and tailor interventions by developing and applying innovative segmentation strategies. This could lead to more precise and targeted policies, ultimately improving the efficiency of fare evasion prevention measures.

In summary, this study shows the time-dependent degree of fare evasion throughout the network using the fare evasion risk estimation as a driver. Research challenges will arise from the development of optimization approaches for inspection planning, scheduling, and randomized patrols on route segments.

CRediT authorship contribution statement

Benedetto Barabino: Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Massimo Di Francesco:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization. **Roberto Ventura:** Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Data curation, Conceptualization. **Simone Zanda:** Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

This paper was initiated, and a preliminary draft was completed while the first author was employed at CTM SpA. Therefore, we would like to express our gratitude to the General Management of CTM SpA for their support in this research and for granting us access to their authentic data for experimentation.

Furthermore, this research received partial funding from the Fondazione di Sardegna, Progetto biennale bando 2021, as part of the project "Computational Methods and Networks in Civil Engineering (COMANCHE)," as well as from the Department of Civil, Environment, Land and Architecture Engineering and Mathematics (DICATAM), University of Brescia, through the research grant "Valuation of the risk of fare evasion in an urban public transport network," CUP: F73C22001140007.

It is essential to note that any viewpoints, findings, conclusions, or recommendations presented in this material are solely those of the authors and do not necessarily represent the perspectives of CTM and MUR. Any remaining errors or inaccuracies are the sole responsibility of the authors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trip.2024.101238.

References

Alhassan, I., Matthews, B., Toner, J., Susilo, Y., 2022. Seamless public transport ticket inspection: exploring users' reaction to next-generation ticket inspection. J. Public Transp. 24, 100004. https://doi.org/10.1016/j.jpubtr.2022.100004.

Allen, J., Muñoz, J.C., Ortúzar, J.de.D., 2019. On evasion behaviour in public transport: dissatisfaction or contagion? Transp. Res. A Policy Pract. 130, 626–651. https://doi.org/10.1016/j.tra.2019.10.005.

Anastassiou, G.A., 2011. Multivariate hyperbolic tangent neural network approximation. Comput. Math. Appl. 61 (4), 809–821. https://doi.org/10.1016/j.camwa.2010.12.029.

Assaf, C., Van den Broeck, P., 2022. The quest for a fairer formula. How re-institutionalisation begins with neo-illegal transport communities at the fringe of social innovation in Brussels. European Journal of Spatial Development. https://doi.org/10.5281/zenodo.7215256.

Aven, T., 2015. Risk analysis. Wiley. https://doi.org/10.1002/9781119057819.

Ayal, S., Celse, J., Hochman, G., 2021. Crafting messages to fight dishonesty: a field investigation of the effects of social norms and watching eye cues on fare evasion. Organ. Behav. Hum. Decis. Process. 166, 9–19. https://doi.org/10.1016/j.obhdp.2019.10.003.

Barabino, B., Di Francesco, M., Ventura, R., 2023. Evaluating fare evasion risk in bus transit networks. Transportation Research Interdisciplinary Perspectives. https://doi.org/10.1016/j.trip.2023.100854.

Barabino, B., Carra, M., Currie, G., 2024. Fare inspection in proof-of-payment transit networks: A review. J. Public Transp. 26, 100101. https://doi.org/10.1016/j.jpubtr.2024.100101.

Barabino, B., Salis, S., Useli, B., 2013. A modified model to curb fare evasion and enforce compliance: empirical evidence and implications. Transp. Res. A Policy Pract. 58, 29–39. https://doi.org/10.1016/j.tra.2013.10.007.

Barabino, B., Salis, S., Useli, B., 2014. Fare evasion in proof-of-payment transit systems: deriving the optimum inspection level. Transp. Res. B Methodol. 70, 1–17. https://doi.org/10.1016/j.trb.2014.08.001.

Barabino, B., Salis, S., 2019. Moving towards a more accurate level of inspection against fare evasion in proof-of-payment transit systems. Netw. Spat. Econ. 19 (4), 1319–1346. https://doi.org/10.1007/s11067-019-09468-3.

Barabino, B., Salis, S., 2020. Do students, workers, and unemployed passengers respond differently to the intention to evade fares? An empirical research. Transportation Research Interdisciplinary Perspectives 7, 100215. https://doi.org/10.1016/j.trip.2020.100215.

Barabino, B., Salis, S., Useli, B., 2022b. Assessing the intention to evade fares for demographic segments of passengers: empirical research in italy for building smart (er) cities. Journal of Urban Planning and Development 148 (1). https://doi.org/10.1061/(ASCE)UP.1943-5444.0000804.

Barabino, B., Salis, S., 2023. Segmenting fare-evaders by tandem clustering and logistic regression models. Public Transport 15 (1), 61–96. https://doi.org/10.1007/s12469-022-00297-1.

Barabino, B., Salis, S., Useli, B., 2015. What are the determinants in making people free riders in proof-of-payment transit systems? Evidence from Italy. Transp. Res. A Policy Pract. 80, 184–196. https://doi.org/10.1016/j.tra.2015.07.017.

Barabino, B., Lai, C., Casari, C., Demontis, R., Mozzoni, S., 2016. Rethinking transit time reliability by integrating automated vehicle location data, passenger patterns, and web tools. IEEE Trans. Intell. Transp. Syst. 18 (4), 756–766.

Barabino, B., Lai, C., Olivo, A., 2020. Fare evasion in public transport systems: a review of the literature. Public Transport 12 (1), 27–88. https://doi.org/10.1007/s12469-019-00225-w.

Barabino, B., Di Francesco, M., Maternini, G., Mozzoni, S., 2022a. An offline framework for the diagnosis of transfer reliability using automatic vehicle location data. IEEE Intell. Transp. Syst. Mag. 14 (4), 163–182. https://doi.org/10.1109/MITS.2021.3051977.

Blue-Ribbon Panel – BRP, 2023. Technical Report on MTA Fare and Toll Evasion. Available from https://new.mta.info/document/111531 accessed on 15/11/2023.

Boyd, C., 2020. Revisiting the foundations of fare evasion research. Transp. Res. A Policy Pract. 137, 313–324. https://doi.org/10.1016/j.tra.2020.05.004.

Boyd, C., Martini, C., Rickard, J., Russel, A., 1989. Fare evasion and non-compliance: a simple model. JTEP 23 (2), 189–197.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Brotcorne, L., Escalona, P., Fortz, B., Labbé, M., 2021. Fare inspection patrols scheduling in transit systems using a Stackelberg game approach. Transp. Res. B Methodol. 154, 1–20. https://doi.org/10.1016/j.trb.2021.10.001.

Brownlee, J., 2019. *Loss and Loss Functions for Training Deep Learning Neural Networks.* https://machinelearningmastery.com/loss-and-loss-functions-for-training-deep-learning-neural-networks/.

Bucciol, A., Landini, F., Piovesan, M., 2013. Unethical behavior in the field: demographic characteristics and beliefs of the cheater. J. Econ. Behav. Organ. 93, 248–257. https://doi.org/10.1016/j.jebo.2013.03.018.

Cantillo, A., Raveau, S., Muñoz, J.C., 2022. Fare evasion on public transport: who, when, where and how? Transp. Res. A Policy Pract. 156, 285–295. https://doi.org/10.1016/j.tra.2021.11.027.

Cassidy, 2018. New Frontiers for BI: Beyond Scaling. Site: https://www.artd.com.au/news/new-frontiers-for-bi-beyond-scaling/.

Ceder, A., 2016. Public transit planning and operation. CRC Press. https://doi.org/10.1201/b18689.

Celse, J., Grolleau, G., 2023. Fare evasion and information provision: what information should be provided to reduce fare-evasion? Transp. Policy 138, 119–128. https://doi.org/10.1016/j.tranpol.2023.05.008.

Clarke, R.V., Contre, S., Petrossian, G., 2010. Deterrence and fare evasion: results of a natural experiment. Secur. J. 23 (1), 5–17. https://doi.org/10.1057/sj.2009.15.

Cools, M., Fabbro, Y., Bellemans, T., 2018. Identification of the determinants of fare evasion. Case Studies on Transport Policy 6 (3), 348–352. https://doi.org/10.1016/j.cstp.2017.10.007.

Correa, J., Harks, T., Kreuzen, V.J.C., Matuschke, J., 2017. Fare evasion in transit networks. Oper. Res. 65 (1), 165–183. https://doi.org/10.1287/opre.2016.1560.

Cosby, S., 1985. A method for measuring the revenue loss due to fraud within a public transport undertaking. Traffic Engineering & Control.

CTM, 2023. Carta mobilità 2023-2024. Available at: https://ctmcagliari.portaletrasparenza.net/it/trasparenza/servizi-erogati/carta-dei-servizi-e-standard-di-qualita.html. Accessed on 16.07.2024.

Currie, G., Delbosc, A., 2017. An empirical model for the psychology of deliberate and unintentional fare evasion. Transp. Policy 54, 21–29. https://doi.org/10.1016/j.tranpol.2016.11.002.

Dai, Z., Galeotti, F., Villeval, M.C., 2017. The efficiency of crackdowns: a lab-in-the-field experiment in public transportations. Theor. Decis. 82 (2), 249–271. https://doi.org/10.1007/s11238-016-9561-0.

Dai, Z., Galeotti, F., Villeval, M.C., 2018. Cheating in the lab predicts fraud in the field: an experiment in public transportation. Manag. Sci. 64 (3), 1081–1100. https://doi.org/10.1287/mnsc.2016.2616.

Dauby, L., Zoltan, K., 2007. Fare evasion in light rail systems. Transportation Research Circular 6–8.

Delbosc, A., Currie, G., 2016. Cluster analysis of fare evasion behaviours in Melbourne, Australia. Transp. Policy 50, 29–36. https://doi.org/10.1016/j.tranpol.2016.05.015.

Delbosc, A., Currie, G., 2019. Why do people fare evade? A global shift in fare evasion research. Transp. Rev. 39 (3), 376–391. https://doi.org/10.1080/01441647.2018.1482382.

Eddy, D., 2010. Fare Evasion is it a youth issue? Transit Australia-Australia's Urban Passenger Transport Journal 65 (12), 1–7.

Egu, O., Bonnel, P., 2020. Can we estimate accurately fare evasion without a survey? Results from a data comparison approach in Lyon using fare collection data, fare inspection data and counting data. Public Transport 12 (1), 1–26. https://doi.org/10.1007/s12469-019-00224-x.

Escalona, P., Brotcorne, L., Fortz, B., Ramirez, M., 2024. Fare inspection patrolling under in-station selective inspection policy. Ann. Oper. Res. 332 (1), 191–212.

Fine, W.T., 1971. Mathematical Evaluation for Controlling Hazards. J. Safety Res. 3, 157–166.

Flach, P., 2012. Machine Learning: The Art and Science of Algorithms That Make Sense of Data. Cambridge University Press.

González, B., Codocedo, 2019. Fare evasion in public transport: grouping transantiago users' behavior. Sustainability 11 (23), 6543. https://doi.org/10.3390/su11236543.

Guarda, P., Galilea, P., Paget-Seekins, L., Ortúzar, J.de.D., 2016. What is behind fare evasion in urban bus systems? An econometric approach. Transp. Res. A Policy Pract. 84, 55–71. https://doi.org/10.1016/j.tra.2015.10.008.

Guzman, L.A., Arellana, J., Camargo, J.P., 2021. A hybrid discrete choice model to understand the effect of public policy on fare evasion discouragement in Bogotá's Bus Rapid Transit. Transp. Res. A Policy Pract. 151, 140–153. https://doi.org/10.1016/j.tra.2021.07.009.

Huang, S., Liu, X., Chen, W., Song, G., Zhang, Z., Yang, L., Zhang, B., 2022. A detection method of individual fare evasion behaviours on metros based on skeleton sequence and time series. Inf. Sci. 589, 62–79. https://doi.org/10.1016/j.ins.2021.12.088.

Ian, G., Yoshua, B., Courville, A., 2016. *Deep Learning* (The MIT Press, Ed.).

International Organization of Standardization, 1985. *ISO 5807:1985- Information processing - Documentation symbols and conventions for data, program and system flowcharts, program network charts and system resources charts*.

Isreal, S., Strathman, J.G., 2002. *Analysis of Transit Fare Evasion in the Rose Quarter.*.

Killias, M., Scheidegger, D., Nordenson, P., 2009. The effects of increasing the certainty of punishment. Eur. J. Criminol. 6 (5), 387–400. https://doi.org/10.1177/1477370809337881.

Lee, J., 2011. Uncovering San Francisco, California, Muni's Proof-of-payment patterns to help reduce fare evasion. Transportation Research Record: Journal of the Transportation Research Board 2216 (1), 75–84. https://doi.org/10.3141/2216-09.

MathWorks, 2022a. *crossentropy - Neural network performance*. https://it.mathworks.com/help/deeplearning/ref/crossentropy.html.

MathWorks, 2022b. *Neural Net Fitting - Solve fitting problem using two-layer feed-forward networks*. https://it.mathworks.com/help/deeplearning/ref/neuralnetfitting-app.html.

Milnes, M., 2012. Random checks to stem fare evasion on public transport. The Advertiser. Site: https://www.adelaidenow.com.au/news/south-australia/random-checks-to-stem-fare-evasion-on-public-transport/news-story/b06428c7322c42dd3c9bc566af4d1a14.

MTA, 2019. MTA Announces Surge in Police Presence to Deter Fare Evasion. Site: https://new.mta.info/press-release/archive.

Pili, F., Olivo, A., Barabino, B., 2019. Evaluating alternative methods to estimate bus running times by archived automatic vehicle location data. IET Intel. Transport Syst. 13 (3), 523–530. https://doi.org/10.1049/iet-its.2018.5339.

Porath, K., Galilea, P., 2020. Temporal analysis of fare evasion in Transantiago: a socio-political view. Res. Transp. Econ. 83, 100958. https://doi.org/10.1016/J.RETREC.2020.100958.

Pourmonet, H., Bassetto, S., Trepanier, M., 2015. Vers la maîtrise de l'évasion tarifaire dans un réseau de transport collectif. *1e Congrès International De Génie Industriel (Québec, Canada)*.

Prokosch, A., Gartsman, A., 2017. All-Door Boarding Without Proof-of-Payment: Revenue Impacts and Operational Implications. In: TRB 96th Annual Meeting Compendium of Papers, pp. 1–14.

Reddy, A.V., Kuhls, J., Lu, A., 2011. Measuring and controlling subway fare evasion. Transportation Research Record: Journal of the Transportation Research Board 2216 (1), 85–99. https://doi.org/10.3141/2216-10.

Reed, R., Marks II, R.J., 1999. Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks. MIT Press.

Salis, S., Barabino, B., Useli, B., 2017. *Segmenting fare evader groups by factor and cluster analysis*. 503–515. doi: 10.2495/UT170431.

Sasaki, Y., 2014. Optimal choices of fare collection systems for public transportations: Barrier versus barrier-free. Transp. Res. B Methodol. 60, 107–114. https://doi.org/10.1016/j.trb.2013.12.005.

Schmidhuber, J., 2015. Deep Learning in neural networks: an overview. Neural Netw. 61, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003.

SFMTA, 2020. Fare inspection reimagined. Site. https://www.sfmta.com/blog/fare-inspection-reimagined.

Smith, M.J., Clarke, R.V., 2020. Crime and public transport. Crime and Justice. A Review of Research 169–233.

Sramka, M., Slovak, M., Tuckova, J., Stodulka, P., 2019. Improving clinical refractive results of cataract surgery by machine learning. PeerJ 2019 (7), 1–23. https://doi.org/10.7717/peerj.7202.

Strobl, C., Boulesteix, A.L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinf. 8. https://doi.org/10.1186/1471-2105-8-25.

Ventura, R., Maternini, G., Barabino, B., 2024. Traffic hazards on main road's bridges: Real-time estimating and managing the overload risk. IEEE Trans. Intell. Transp. Syst. 25 (9). https://doi.org/10.1109/TITS.2024.3371265.

Wolfgram, L., Pollan, C., Hostetter, K., Martin, A., Spencer, T., Rodda, S., Amey, A., 2022. Measuring and managing fare evasion. Transp. Res. Board. https://doi.org/10.17226/26514.

Yin, Z., Jiang, A.X., Tambe, M., Kiekintveld, C., Leyton-Brown, K., Sandholm, T., Sullivan, J.P., 2012. TRUSTS: scheduling randomized patrols for fare inspection in transit systems using game theory. AI Mag. 33 (4), 59–72. https://doi.org/10.1609/aimag.v33i4.2432.

Zhang, P., 2010. Industrial control system simulation routines. In Advanced Industrial Control Technology (pp. 781–810). Elsevier. doi: 10.1016/B978-1-4377-7807-6.10019-1.