



Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com



Integrating Inverse Data Envelopment Analysis and Machine Learning for Enhanced Road Transport Safety in Iran

Seyyed Mojtaba Tabatabai^{1*}, Fatemeh Al-Sadat Tabatabai²

1. Faculty of Engineering, Islamic Azad University Birjand Branch, Iran

Corresponding author: seyyedmojtatabatabai@gmail.com

 <https://doi.org/10.22115/SCCE.2023.399661.1657>

ARTICLE INFO

Article history:

Received: 29 May 2023

Revised: 24 June 2023

Accepted: 08 July 2023

Keywords:

DEA;

Safety status;

Inefficiency;

Risk scores;

ANN;

Sensitivity analysis.

ABSTRACT

The purpose of this research is to present a new method for considering accidents according to the environmental, traffic and geometrical conditions of the road, which considers accidents according to the interaction of the components that lead to them. In order to enter the physical characteristics, this approach divides the road into units or parts with homogeneous physical characteristics, and as a result, the decision about the safety status of the road is made for a length of road with specific characteristics instead of a single point. This approach has been carried out using the Data Envelopment Analysis (DEA) method, which, unlike regression methods, does not require obtaining the distribution function and considering hypotheses about it. This method gives scores (inefficiencies) that allow road segments to be appropriately ranked and prioritized in terms of accident proneness. In the current research, a case study was conducted on routes with a length of 144.4 kilometers, which resulted in the identification of 154 road sections with different relative risk scores, thus the accident sections were identified and prioritized with the proposed method, which in terms of the definition of entry indicators and the output based on the data coverage analysis method is considered as a new experience for the priority of road sections. Furthermore, this study focuses on the application of artificial neural networks (ANNs) in analyzing road safety. An idealized ANN model is developed using a database of various input parameters related to road segments, and the weighted index of accidents as the target variable. The results reveal the relative importance of different parameters on the weighted index, with the Ratio of curvature, Length of the segment, and Condition of the pavement identified as the most influential factors. These findings highlight the significance of road curvature, segment length, and pavement condition in determining accident severity. The study underscores the potential of ANNs for assessing road safety and informs targeted interventions to mitigate accidents.

How to cite this article: Tabatabai SM, Tabatabai FA. Integrating inverse data envelopment analysis and machine learning for enhanced road transport safety in Iran. *J Soft Comput Civ Eng* 2024;8(1):141-160. <https://doi.org/10.22115/scce.2023.399661.1657>

2588-2872/ © 2023 The Authors. Published by Pouyan Press.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



1. Introduction

Transportation is a concept that mankind has been involved with since the beginning of its creation and this concept has played a key role in human life since long ago. After the invention of the wheel, transportation underwent a fundamental transformation, and more advanced and complex devices were used in this field day by day, so that today the issues raised in the field of transportation are among the most complex issues in different dimensions. Despite the many benefits of new technologies used in this field, unfortunately, this industry is associated with financial and life risks for humans. Every year, many people die in traffic accidents in the world [1]. Iran is also one of the countries whose statistics of accidents and casualties in the field of transportation, especially in its road type, are very high compared to world standards [2]. Therefore, dealing with the issue of safety in Iran's road transportation can help improve road safety and help the authorities in better planning in this direction. With the prevalence of road transportation, traffic accidents have become an important factor threatening social safety. According to WHO statistics [3], traffic accidents in the world's industrial cities account for about one third of the total number of accidents that lead to death, which are considered to be one of the most important urban problems along with traffic. Also, the number of accidents in third world countries is several times higher than that of industrialized countries [4]. An increase in traffic violations and accidents can cause an increase in disorder and social chaos; If these violations are not dealt with, it will cause the rights of other citizens to be violated, and this issue will cause the phenomenon of cultural backwardness to spread. During the past years, about 26,000 of our compatriots have lost their lives in accidents and about ten times that number have been injured [5]. The global average death rate due to traffic accidents is between 14 and 15 people per 100,000 people, but this amount in Iran is twice the world average, that is, about 30 people per 100,000 people [6]. Therefore, the importance and necessity of the current research can be taken at the ideal level to take preliminary steps to reduce accidents. By identifying the most important influencing factors of the road in accidents, useful information can be provided to policy makers for preventive measures in order to eliminate them with the least amount of time and cost, and as a result, the severity of financial and life losses can be dramatically reduced. In general, according to what has been said, it seems necessary to carry out a research under the title of identifying and ranking the effective road factors on road safety. In recent years, a substantial body of research has highlighted the significance of artificial intelligence in the field of engineering [7–13].

In this article, a new approach has been introduced to identify the parts of the accident. One of the advantages of the new approach compared to the previous researches is the study of the accident of road parts instead of road points. Since the interaction of groups of factors leads to the occurrence of an accident in a piece of road, therefore, it is more logical to consider the interruptions with specific length and characteristics instead of defining the accident point in the past, whose exact range is not defined. This approach has been carried out using data envelopment analysis method. Data envelopment analysis does not require obtaining the distribution function and considering hypotheses about it. This method evaluates the potential of converting inputs such as geometric characteristics and road side factors to output compared to the best performance of each part compared to other parts.

Nowadays, Due to the great importance of safety in transportation, various studies have been conducted in this field in the past. The improvement of efficiency and safety in road transportation has emerged as a prominent focus in safety science and transportation science [14]. Among the methods used to assess safety efficiency by considering multiple inputs and outputs of decision-making units (DMUs), Data Envelopment Analysis (DEA) stands out as one of the most popular techniques [15]. In the context of transportation safety evaluation, DEA has been applied in various scenarios. For instance, in [8], a non-radial DEA model was introduced to evaluate the railway efficiency of European countries, specifically regarding safety at railway level crossings. Additionally, Nahangi et al. [16] utilized DEA to evaluate the safety efficiency of construction sites, identifying a correlation between safety efficiency and climate conditions. Furthermore, a double frontier cross efficiency method with an evidential reasoning approach was proposed in [17] to assess the safety efficiency of road transportation. Omrani et al. [18], DEA was combined with the group best-worst method to evaluate the safety efficiency of Iran's road transportation. However, while DEA offers a scientific approach to evaluate safety efficiency, it does not directly address the challenge of effectively achieving safety objectives. This is where the concept of inverse DEA comes into play. Inverse DEA, proposed by Wei et al. [19], provides a powerful tool for determining the optimal path to achieve specific safety objectives given a certain efficiency level [20]. It is particularly useful for solving two types of problems: firstly, determining the amount of additional outputs a particular DMU can produce with given additional inputs while maintaining its current efficiency relative to others; secondly, establishing the additional inputs required by a DMU to generate given additional outputs, while maintaining the same efficiency relative to others. The concept of inverse DEA has gained significant attention and found practical applications in various domains. For instance, Yan et al. [21] introduced an extended inverse DEA model with preference cone constraints to incorporate decision makers' preferences into resource reallocation decisions. Addressing the issue of variable returns to scale (VRS), Lertworasirikul et al. [22] proposed an inverse DEA model that preserves the relative efficiencies of all DMUs. Moreover, Lim [23] used the inverse DEA method to establish product targets by considering frontier changes, while Amin et al. [24] applied goal programming to inverse DEA for devising inputs-outputs plans in the banking industry.

2. Theoretical foundations of DEA

Data envelopment analysis was created by Charnes et al. [25], Cooper and Rhodes as a tool to test the relative efficiency of production units or decision-making units based on the information of produced outputs and consumed inputs. Using this system, the relative efficiency score of the units is calculated and efficient and ineffective units are determined. So far, the data coverage analysis method has been used in traffic safety topics only in identifying intersections with high accidents in the city and comparing the safety status of countries. They presented a model that has the ability to measure efficiency with multiple inputs and multiple outputs. The model known as Data Envelopment Analysis (DEA) was first introduced by Charnes et al. [25]. It is often referred to as the CCR model, derived from the initials of these three individuals. The primary objective of the CCR model is to assess and compare the relative efficiency of decision-making

units, such as schools, hospitals, bank branches, and other similar cases, which involve multiple comparable inputs and outputs. To evaluate the efficiency of a unit under review using the CCR model, the ratio of the weighted sum of outputs to the weighted sum of inputs is used as a scale for efficiency measurement. In cases where each unit has m inputs to produce s outputs, the fractional form of data envelopment analysis for evaluating efficiency will be as follows [26]:

$$MAX EF_j = \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \quad (1)$$

In which:

$$\frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1. u_r, v_i \geq 0 \quad (2)$$

In this non-linear and non-convex problem, EF_j is the efficiency of the unit DMU $_j$ and the other variables are as follows:

x_{ij} : i -th input amount for j -th unit ($i=1,2,\dots,m$).

y_{rj} : RAM output amount for unit j ($r=1,2,\dots,s$).

u_r : RAM output weight,

v_i : i -th input weight.

The problem in this problem is that this model has infinite solutions. Because if the optimal values of the variables are u^* and v^* , then αu^* and αv^* will also be the optimal solution of this model. To solve this problem, after changing the linear shape variable twice, the classical data coverage analysis model is presented in the following form [19]:

$$MAX EF_j = \sum_{r=1}^s \mu_r y_{rj} \quad (3)$$

So that,

$$\sum_{r=1}^s \mu_r y_{rj} - \sum_{i=1}^m w_i x_{ij} \leq 0. u_r, v_i \geq 0 \quad (4)$$

$$\sum_{i=1}^m w_i x_{ij} = 1 \quad (5)$$

$$\mu_r, w_i \geq 0 \quad (6)$$

In this regard, the change of variables is as follows:

$$\mu_r = \frac{u_r}{\sum_{i=1}^m v_i x_{ij}} \cdot w_i = \frac{v_i}{\sum_{i=1}^m v_i x_{ij}} \quad (7)$$

The DEA method can successfully divide decision making units into two groups of efficient units (efficiency values equal to one) and ineffective units (efficiency values less than one). Using this method, inefficient units are ranked, but it is not possible to rank efficient units with this method [27]. In order to solve this problem, researchers have proposed different methods. They presented the (AP) method for ranking efficient units, which makes it possible to determine the most efficient unit, by using this method, the score of efficient units can become a number greater than one, as a result, an overall ranking is provided for efficient and ineffective units. In this article, the Anderson-Peterson (AP) method is used, for this purpose, it is enough to consider the units whose efficiency value is equal to one in the first-order solution of the CCR model, and

by removing the limitation related to that unit from the total limitations of the first-order solution of the CCR model, it is solved again for that unit. By performing this operation, the ranking of all effective units will be achieved.

3. Methodology

This article aims to compare the road sections in terms of accident proneness through the method of data envelopment analysis. Data envelopment analysis performs the relative efficiency of decision-making units based on the amount of produced outputs and consumed inputs. Efficiency is the ratio of outputs to inputs. Here, the output refers to the number of accidents on the road section in question, and the input refers to the factors affecting accidents that exist on the road section under investigation. In order to use the data envelopment analysis method, it is necessary to have decision-making units with similar functions. Considering that the type of road specifications as input and accidents as output for each segment are similar to other segments, each segment of the road is considered as a decision unit. In this article, road segmentation approach is considered to create units with similar functions. The purpose of ranking road sections is ultimately to help in the optimal allocation of resources and appropriate policies to improve safety. Therefore, since the DEA model examines risk indicators (the number and severity of accidents) with the involvement of road characteristics, it can provide a more comprehensive approach in policy making and planning.

In section 3-1, the factors affecting accidents are reviewed and in section 2-3, the case study and the method of data collection are explained. In the following, the methodology of road segmentation and identification of homogeneous parts, factors affecting accidents (model input) and accident criteria (model output) and the use of data coverage analysis to identify and compare accident-prone parts are explained.

3.1. Road accident factors

Identifying accident-prone parts requires knowing the factors affecting the occurrence of accidents. Of course, it should be noted that the factors in this discussion are factors that are dependent on the location and hence factors such as special weather conditions, the condition of the driver and the type of vehicle we are considering. Based on previous researches [28,29], the characteristics that can be considered to evaluate the safety performance of the route are: average daily traffic (ADT), curvature (length and radius), straight route length, cross section characteristics (lane width, shoulder width), density of accessible routes, roadside hazards, sight distance, road slope, pavement condition, speed limit.

3.2. Data collection

In general, the required information includes road specifications, traffic and accident information. The study was carried out on a sample of 144.4 kilometers of Khorasan-Razavi province roads, including two-lane two-way sections of the Mashhad-Kalat and Mashhad-Freeman axes. The details of the route plan were obtained from the Mashhad Road and Transport Department. Unfortunately, due to the oldness of the mentioned routes, there have been changes over the years in these routes that have not been recorded, for this reason, the visit was carried

out using the movement mode of the GPS device to collect horizontal location information and compare it with the maps.

These investigations were done by driving on the extreme right side of the lane at an average speed of 50 km/h. Fortunately, in these roads, the changes in the horizontal plan in the sections of the two-way lanes were limited only in some points to the line deviation, which was also observed in the field in a separate visit. Roadside risk, number of approaches, sections with speed limit and pavement quality index were taken by field visits by road experts. The information related to the vertical curves and their interference with the horizontal curve was omitted due to the lack of access to the longitudinal profile maps of the routes and the functional speed factor of the vehicles at different times due to the lack of speed recording equipment.

3.3. Route segmentation

Until now, several researchers have attempted to estimate accident models using the road segmentation approach [26,28]. However, many of these studies have solely focused on road sections with fixed lengths or those between two primary intersections. In contrast, Ref. [30] took a different approach by modeling the road with homogeneous characteristics concerning traffic flow and geometric conditions, such as horizontal curvature degree, width of shoulder and middle island, lane width, and other relevant factors related to accidents. Additionally, Cafiso et al. [28] introduced a comprehensive segmentation method that combines exposure to risk, geometric conditions, compatibility, and conceptual variables associated with safety performance to model accidents.

In this article, cutting the path and identifying homogeneous parts is done based on the factors affecting the accidents. Some of the accident factors that can be used for this purpose are:

- Average daily traffic (ADT).
- The width of the movement lines and height
- Speed limit
- Curvature change rate
- Pavement condition

The amount of ADT of each route is known, and the beginning and end of sections with a change in width or speed limit can be determined by field visit and its amount can be measured. The curvature change rate is determined from the specifications of the horizontal plan, which can be defined for each section as follows [31]:

$$CCR_{sec} = \frac{\sum_{i=1}^n |\gamma_i|}{L} \quad (8)$$

In which γ_i is the angle of deviation for i th arc in the length of L . To obtain sections with homogeneous CCR, cumulative deviation angles Y are plotted in terms of kilometers and then smooth trend lines are fitted. The CCR value for each given segment is equal to the slope of the drawn line. This definition is shown in Figure 1 based on a sample of information collected for the present study.

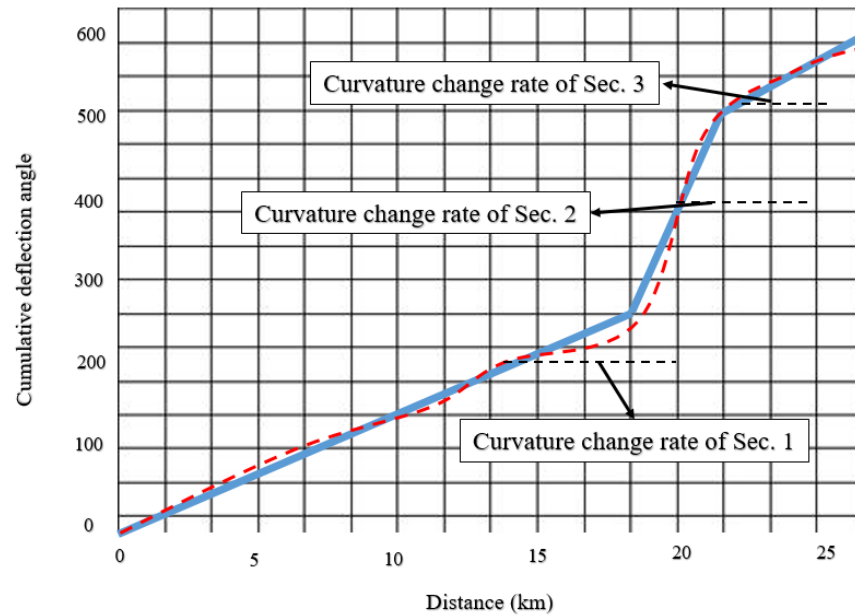


Fig. 1. Slope of the curvature line.

In (Anastasopoulos et al., 2008) [32] is shown that the condition of the pavement in terms of driving quality and slippage is effective on the rate of accidents. In this research, the road was divided into homogeneous parts based on the level of surveillance level of 7. The scoring method is based on the AASHTO method [33]. The inspectors score the condition of the road pavement between zero (very poor) and five (very good). First, sections with a fixed length of 500 meters are considered, and sections with similar scores are combined with others, and larger sections are obtained. It should be noted that in places where the quality of the pavement has changed significantly, the length of 500 meters has not been observed, and sections with a shorter length have also been considered. Also, due to the lack of access to friction measurement devices, This factor has been omitted. Based on the change of each of the above factors, homogeneous road sections can be defined in such a way that a homogeneous section is a section where the mentioned factors do not change in that section.

3.4. DEA model inputs

The inputs of the data envelopment analysis model are the characteristics of the decision-making unit and effective on the output, which include the variables used for segmentation and other characteristics that are calculated for each segment separately. These features include:

- The length of the piece
- curvature ratio
- Direct path ratio
- Roadside hazard index
- Access density
- Proportion of prohibited overtaking areas

The ratio of the distance from the population centers at the beginning and end of the route

The segment length is obtained at each segmentation stage. According to the horizontal geometric plan of the track, the curvature ratio (CR) and the straight track ratio (TR) are calculated as follows:

$$CR = \frac{\sum_{j=1}^m L_{cj}}{L_{HS}} \quad (9)$$

$$TR = \frac{MAX_{e=1}^N (L_{Te})}{L_{HS}} \quad (10)$$

L_{HS} is the total length of the homogeneous piece, L_{cj} is the j th arc length in a homogeneous piece with M arc, L_{Te} is the length of e th straight path in a homogeneous section where there are N straight paths.

Cafiso et al. [28] have presented a Roadside Hazard Index (RHS) for use in 200-meter sections of road. In this index, a score (0 = absent, 1 = low risk, 2 = high risk) is assigned to 5 roadside hazards (embankments, bridges, entrance nose and the transfer area of guardrails, trees and other rigid obstacles) to the right and left side separately. Then the weighted average of 5 factors can be calculated as follows:

$$RSH_i = \frac{\sum_{k=1}^2 \max(score_{ijk} \times weight_j)}{2} \quad (11)$$

so that k is the direction of visit (1=right, 2=left) and the score of items i and j in the i th unit of visit is in the direction of k . The relative weight of the first item on the side of the road is based on AASHTO accident severity indicators, which are:

3 for embankments, 5 for bridges, 4 for inlet nose and guardrail transition area, 2 for trees and other rigid obstacles, and 1 for culverts.

This order of roadside risk is evaluated by safety inspectors for 200-meter segments using the designed zero-level check, and then for each of the homogeneous segments, its average value is considered as the cross-sectional risk index.

The density of accesses and the proportion of prohibited overtaking areas are respectively obtained by dividing the number of access roads and the total length of prohibited overtaking areas by the length of the entire plot. Due to the presence of concentrated and various industrial and recreational uses on the side of the roads near the cities, there are more factors of distraction, flow volume and traffic chaos, therefore, Ayati [34] researches show that the rate of accidents has an inverse ratio with the logarithm of the distance from the city. Therefore, in order to give effect to the different importance of the beginning and the end of the route, the index related to the ratio of the distance from the population centers of the beginning and the end of the route is defined as follows:

$$DCI = \frac{P_a \times \log D_b + P_b \times \log D_a}{(\log(D_b + D_a)) \times (P_a + P_b)} \quad (12)$$

where in DCI is ratio of the distance from population centers at the beginning and end of the route D_a and D_b are the distance between the center of the piece and the beginning and end of the route (cities a and b). P_a and P_b are population of cities a and b in a population case study Each city is obtained from the results of the population and housing census of 2015.

3.5. DEA model outputs

For each of the homogeneous components, the output in the DEA model is the number of accidents. However, in addition to the frequency of accidents, one of the factors that are effective in identifying a place as a high accident spot is the severity of the accidents that occurred in that place. Different researchers have mentioned different coefficients in their studies for the ratio of importance or severity of damage, injury and death accidents. For example, the Belgian Ministry of Transport uses ratios of 1, 3, and 5 for damage, injury, and death accidents, and the Portuguese Road Administration uses ratios of 100, 10, and 1 feet, severe injury, and minor injury. Road and Transport Department of Khorasan Province uses coefficients 1, 3, and 5 for damage, injury, and death accidents in order to identify accident-prone points, and any point that gets a total score of more than 30 is considered as an accident-prone point. Ref. [35] by examining various relationships according to the current conditions of accident reporting and safety culture in Iran, has emphasized the use of the same ratios of 1, 3, and 5 for damage, injury, and death accidents. In this research, using these coefficients, the weighted index of accidents has been calculated as an output in the DEA model for each part.

4. Scoring and prioritizing road sections

A part of Mashhad-Kalat and Mashhad-Fariman two-lane road with a length of 144.4 kilometers was considered as a case study. In general, 154 homogeneous parts of 144.4 kilometers were obtained in the selected routes, the longest part is 5 kilometers long and the shortest part is 0.15 kilometers. In the current research, it is assumed that each of the 154 road sections obtained by the method defined in section 3-3 are decision making units. In these decision-making units, the input variables are the length of the segment (X1), the condition of the pavement (X4), the ratio of the straight path (X3), the ratio of curvature (X2), the speed limit (X6), the width of the movement lines and shoulders (X5), the ratio of overtaking areas (X8), access density (X7), curve change rate (X10), roadside risk index (X9), distance index from population centers (X11) and average daily traffic volume (X12) and output variable, weighted index of accidents (Y1) , (1, 3 and 5 for damage, injury and death accidents respectively) and an example of these values can be seen in Table 1. For example, segment No. 1 of the Mashhad-Kalat axis has a length of 1000 meters, curvature ratio, 0.208, straight line ratio, 0.427 degree, current level of pavement (PSR) 3.8, width of movement lines and shoulders 7.3 meters, speed limit 80 hr/km, density of accesses, 0.003 ratio of overtaking areas, 0.101 road side danger index, 4 curvature change rate, 0.009 distance index from population centers 0.998 and average daily traffic volume 1500 vehicle per day and the output variable, the weighted index of accidents is 32.

Since the occurrence of an accident is an undesirable factor, therefore, an inefficiency index is defined for each part instead of an efficiency index. The inefficiency (score) of each unit is calculated using the CCR model for each of the road parts (DMUs) and in the next step, the rank of the units with inefficiency equal to one is also calculated using the AP method. Based on the results, the sections with the highest level of inefficiency are considered as the most accident-

prone units on the road, and on this basis, the prioritization of the road sections becomes possible.

CCR model was programmed using Excel spreadsheet software and was used to calculate the inefficiency of different road parts. In Figure (2), the results of calculating the inefficiency of 154 road sections with this method are presented. Then, the sections with an inefficiency of 1 were ranked using the AP method, and the results for these sections can be seen in Table No. (1). For example, the level of inefficiency or risk-creating potential of part 1 in comparison with other parts was 63.2, which according to this score is placed in the fifth priority for safety or improvement. It is worth mentioning that according to the level of budget allocated for the safety of routes, high-accident points can be selected in order of priority.

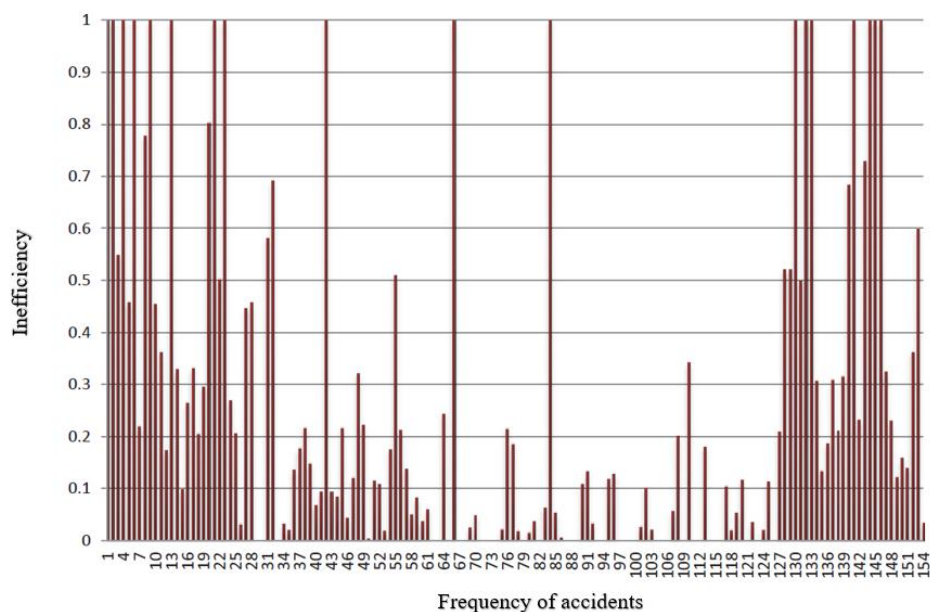


Fig. 2. The level of inefficiency of road parts.

In Table 1, road sections with an inefficiency rate greater than one are shown, among these 18 sections, 7 sections are related to the Freeman-Mashhad route and 11 sections are related to the Mashhad-Kalat route. Also, in order to compare the two roads Mashhad-Kalat and Mashhad-Fariman, the mean of their inefficiency was calculated, which was 1.10 for the Fariman-Mashhad road and 0.47 for the Mashhad-Kalat road, which indicates that the Freeman route is more critical than the Kalat. Figure (3) shows the level of ineffectiveness of road parts against the weighted index of accident frequency (method of road and transportation administration). Although in the method used by the road administration, the length of the accident-prone point is not specified, but in order to compare with the proposed methods, The rate of weighted accidents of each piece was calculated. In the road administration method, the points that have an accident frequency index higher than 30 are introduced as accident-prone points, while in the proposed method, a higher level of inefficiency indicates that the part is accident-prone. According to the figure, a large number of parts simultaneously have a weight index of less than 30 and an inefficiency value of less than one.

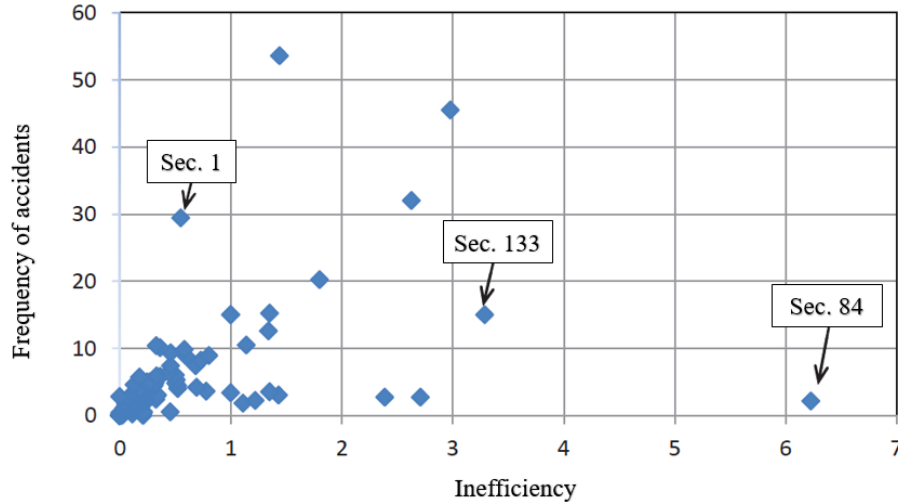


Fig. 1. The inefficiency values of parts by AP method against the weighted index of accident frequency.

Although a number of parts are classified as accident-prone parts in both methods, but a part like part No. 3, despite the high accident weighted index, has a low inefficiency, which shows that compared to other parts, it had a good performance according to its specifications. . On the contrary, in parts such as 84 and 133, which have a low accident weight index and high inefficiency, their relative performance is not suitable according to their specifications, and the number and severity of accidents was expected from them. In the previous method, points like plot 3 may be secured and introduced again in the following years as an accident-prone point, and points like plot 84 may never be introduced as an accident-prone point, While with the proposed method, this point is taken into consideration and the amount of accidents may be completely eliminated by spending a small amount of money.

The accuracy and reliability of this method in identifying accident-prone places can be checked by the economic evaluation of the benefits of the points improved by this method with another method. Using this method in evaluating road safety in combination with other methods can provide more appropriate analyses.

Table 1.
Prioritization of accident-prone parts using AP method.

Sec Number	Route	Mileage from the beginning	End mileage	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	Y ₁	Inefficiency	Rank
1	Kalat	0	1000	1000	0.208	0.109	3.8	7.3	80	0.003	0.101	4	0.009	0.958	1500	32	2.63	3
2	Kalat	1000	4400	3400	0.325	0.254	3.6	7.3	80	0.009	0.835	6.2	0.009	0.980	1500	2.7	1.44	6
4	Kalat	6000	6900	900	0.145	0.186	3.6	7.5	80	0	0	5.2	0.009	0.980	1500	45.5	2.39	4
6	Kalat	7900	12900	5000	0.254	0.352	3.8	7.2	80	0.004	0.201	7.1	0.011	0.960	1500	20.22	2.98	2
9	Kalat	14091	13400	691	0.325	0.452	4.2	7.3	95	0	0.786	6	0.009	0.962	1500	3.4	1.8	5
13	Kalat	18800	19800	1000	0.178	1	3.8	9.2	80	0.009	1	6	0.015	0.970	1500	10.55	1	13
21	Kalat	25300	25800	500	0	1	3.6	12.3	80	0	0	6	0.009	0.980	1500	2.2	1.14	9
23	Fariman	400	1400	1000	0.388	0.345	3	16.5	95	0.003	0.4	5	0.001	0.965	11821	12.6	1.34	8
42	Fariman	1400	1800	400	0	1	4	9.7	50	0.025	0.2	6	0.001	0.980	11821	15	1	13
66	Fariman	1800	2600	600	0	1	4	9.7	45	0.004	0.75	6	0.001	0.950	11821	15	3.29	1
84	Fariman	8900	9900	1000	0	1	4.2	9.7	50	0.010	0.5	6	0.001	0.980	11821	2.7	2.71	3
130	Fariman	12400	1000	1000	0	1	3.8	9.7	55	0.005	0.7	4.5	0.009	0.980	11821	15.2	1.35	7
132	Fariman	13400	1000	1000	0	1	4.2	9.7	45	0.001	0.45	6	0.005	0.940	11821	15	1	13

5. Machine learning based model

Machine learning based models, with a focus on artificial neural networks, are computational models inspired by the structure and functioning of the human brain. These models are designed to learn patterns and relationships from data and make predictions or decisions without being explicitly programmed. Artificial neural networks (ANNs) are a specific type of machine learning model commonly used in various applications. They are composed of interconnected nodes, called artificial neurons or simply "neurons." These neurons are organized in layers: an input layer, one or more hidden layers, and an output layer. Each neuron receives input signals, applies a mathematical transformation to those signals, and produces an output signal that is transmitted to the next layer [36–40].

Training an artificial neural network involves two key phases: forward propagation and backpropagation. During forward propagation, the input data is fed through the network, and computations are performed layer by layer, leading to the generation of an output. The output is then compared to the expected output, and the difference is quantified using a loss function, which measures the network's performance. Backpropagation is the process of updating the network's parameters (weights and biases) based on the calculated loss. It involves propagating the error backwards through the network, adjusting the parameters in a way that reduces the loss. This iterative process is performed over multiple training examples until the network converges to a state where the predictions align closely with the desired outputs. The strength of artificial neural networks lies in their ability to learn complex, non-linear relationships from large amounts of data. They excel at tasks such as image and speech recognition, natural language processing, and recommendation systems. The hidden layers in ANNs enable them to capture and represent hierarchical features in the data, allowing for sophisticated pattern recognition. However, artificial neural networks also come with challenges. They require substantial amounts of labeled training data to generalize well and avoid overfitting. Additionally, training deep neural networks can be computationally expensive and may require specialized hardware, such as graphics processing units (GPUs) or tensor processing units (TPUs), to speed up the calculations.

Table 2 presents a statistical summary of the dataset utilized for training an artificial neural network (ANN) model. The dataset consists of several input parameters and a target variable, the Weighted index of accidents. The input parameters include the Length of the segment, Ratio of curvature, Ratio of the straight path, Condition of the pavement, Width of the movement lines and shoulders, Speed limit, Access density, Ratio of overtaking areas, Roadside risk index, Curve change rate, Distance index from population centers, and Average daily traffic volume. These parameters provide information about various characteristics of the road segments. The statistical summary in Table 2 showcases key statistics such as minimum, maximum, mean, and standard deviation for each parameter, giving an overview of their range and variability within the dataset. Analyzing this summary can provide insights into the distribution and central tendencies of the input variables, aiding in understanding their potential influence on the target variable, the weighted index of accidents.

Table 2
Statistical summary of the data utilized for ANN modeling.

Parameter	Min	Max	Mean	Standard Deviation
Length of the segment	400.00	5000.00	1463.23	1517.81
Ratio of curvature	0.00	0.39	0.18	0.14
Ratio of the straight path	0.11	1.00	0.53	0.27
Condition of the pavement	3.00	4.20	3.72	0.31
Width of the movement lines and shoulders	7.20	16.50	10.22	2.73
Speed limit	45.00	95.00	66.40	16.44
Access density	0.00	0.10	0.01	0.01
Ratio of overtaking areas	0.00	1.00	0.45	0.32
Roadside risk index	4.00	7.10	5.62	0.79
Curve change rate	0.00	0.02	0.01	0.01
Distance index from population centers	0.95	0.98	0.97	0.01
Average daily traffic volume	1500.00	11821.00	6727.52	5176.90
Weighted index of accidents	2.20	45.50	13.72	12.27

Fig. 3 displays the regression values for the training data of an idealized network, along with an error histogram. The regression values represent the predicted output or target variable values generated by the network for the training data. These values are likely compared to the actual target values from the training data to assess the performance of the network. The regression values indicate how well the network is able to approximate or predict the target variable based on the given input data. A good network would produce regression values that closely match the actual target values, indicating a strong predictive capability. The error histogram provides a visual representation of the discrepancies or errors between the predicted regression values and the actual target values. It shows the frequency or distribution of these errors, allowing for an assessment of the network's overall accuracy. A well-performing network would exhibit a histogram with small errors, indicating that the predicted values closely align with the true values.

By examining both the regression values and the error histogram, insights can be gained into the performance and accuracy of the idealized network on the training data. This information helps in evaluating the effectiveness of the network's predictions and identifying areas where improvements might be necessary.

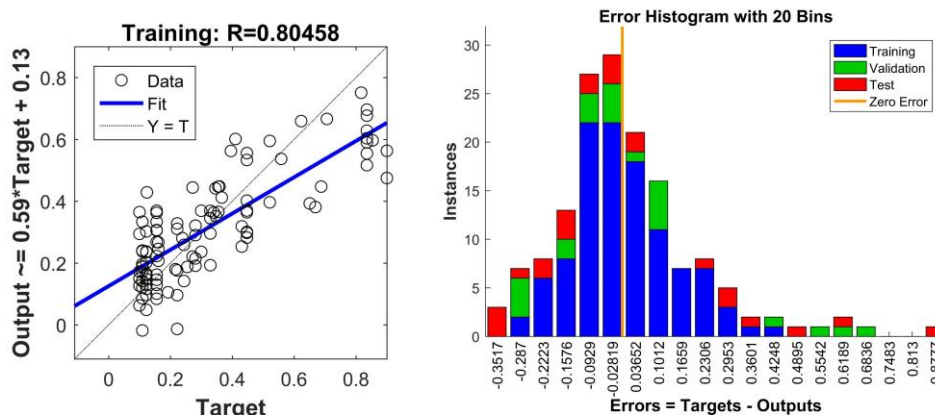


Fig. 3. Regression values for the training data of the idealized network and error histogram.

Fig. 4 displays the performance of the idealized network. It indicates that the network has achieved good results. The term "performance" typically refers to how well the network is performing in terms of its predictive capabilities and accuracy. In this context, it suggests that the idealized network has demonstrated positive outcomes in terms of its ability to make accurate predictions or decisions based on the given input data. The term "good results" implies that the network's performance is satisfactory or meets the desired criteria. This could mean that the network's predictions closely match the actual values of the target variable, indicating a high level of accuracy. It is also possible that the network has achieved a low error rate, indicating minimal discrepancies between the predicted values and the true values. In summary, Fig. 4 signifies that the idealized network has achieved favorable performance, suggesting that it is capable of producing accurate predictions or decisions based on the provided data. The specific details of the figure, such as metrics, visualizations, or numerical values, would provide more precise insights into the nature of the network's performance and the extent of its success.

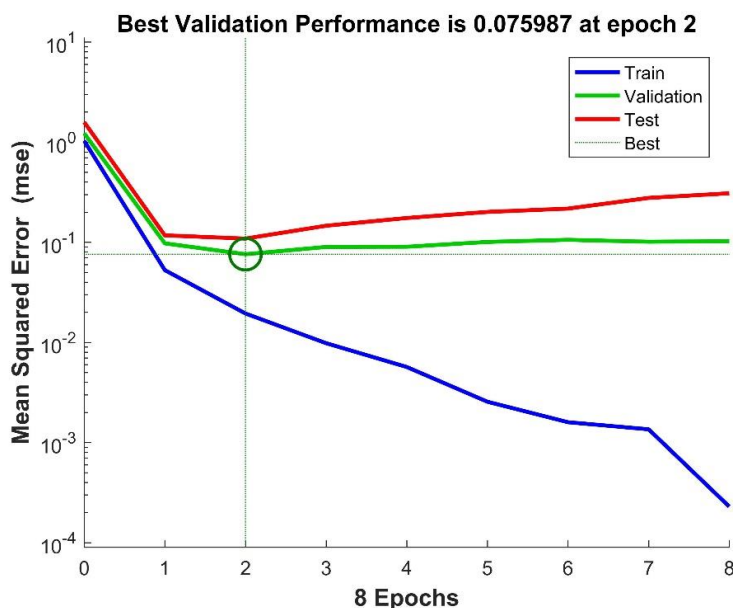


Fig. 4. Performance of the idealized network.

Fig. 5 depicts the training procedure of the idealized artificial neural network (ANN). The training procedure refers to the process of optimizing the parameters of the ANN through iterative steps to improve its performance. This optimization is achieved by exposing the network to a set of training data, consisting of input samples with corresponding target values. The specific steps involved in the training procedure can vary, but they generally follow a similar pattern. Here is a typical overview of the training procedure for an ANN:

1. Initialization: The network's parameters, such as weights and biases, are initialized with random values or predefined initializations.
2. Forward Propagation: The training data is fed into the network, and the input signals are propagated forward through the layers. The network computes outputs based on the current parameter values.

3. **Loss Calculation:** The computed outputs are compared to the corresponding target values from the training data, and a loss function is used to quantify the discrepancy between the predicted and actual values.
4. **Backpropagation:** The error or loss from the previous step is propagated backward through the network. The network calculates gradients, which represent the sensitivity of the loss with respect to the parameters.
5. **Parameter Update:** The network's parameters (weights and biases) are updated based on the calculated gradients. Various optimization algorithms, such as gradient descent, are commonly used to determine the appropriate updates.
6. **Iteration:** Steps 2-5 are repeated for multiple iterations or epochs. The network continues to refine its parameters, gradually minimizing the loss and improving its predictive capability.

The training procedure aims to find the optimal set of parameters that minimize the discrepancy between the predicted outputs of the network and the target values. This iterative process allows the network to learn patterns, relationships, and representations in the training data, enabling it to make accurate predictions on new, unseen data. The details and specific visualizations within Fig. 5 would provide a more comprehensive understanding of the training procedure of the idealized ANN, including information about convergence, learning curves, or other metrics used to assess the network's progress and performance throughout the training process.

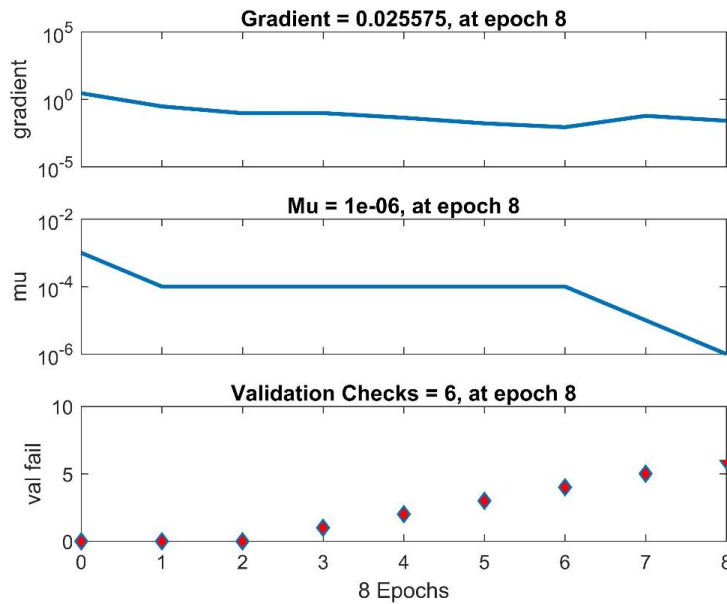


Fig. 5. Training procedure of the idealized ANN.

Table 3 presents the weights of the idealized artificial neural network (ANN) for all hidden layers. The weights play a crucial role in determining the strength of connections between neurons within the network. They represent the numerical values assigned to these connections, indicating their influence on the network's computations. By analyzing the weights in Table 3, valuable insights can be gained regarding the network's architecture and information flow within the hidden layers. The specific values of the weights provide an understanding of the network's

decision-making process and the importance of each neuron in the overall computations. These weights are typically learned during the training process, where the network adjusts them to minimize the discrepancy between predicted outputs and target values. The optimized weights enable the network to capture complex relationships within the input data and make accurate predictions or classifications. Overall, Table 3 provides a glimpse into the inner workings of the idealized ANN, shedding light on the significance of the weights in its functioning.

Table 3

Weights of the idealized ANN for all hidden layers.

W1												W2
-0.263	-0.185	-1.044	1.116	0.264	-0.076	0.114	0.226	0.020	0.146	-0.441	-1.130	-0.444
-0.335	-0.174	-0.125	0.313	-0.093	-0.216	0.655	0.122	0.797	-0.870	1.016	0.956	-0.352
0.022	-0.003	-0.635	0.191	1.158	-0.882	-0.543	0.191	-0.073	-0.231	-0.470	0.651	-0.105
0.814	0.116	-0.302	-0.632	0.569	-0.827	0.421	-0.112	-0.515	0.216	0.710	-0.292	-0.993
-0.209	0.042	-0.629	-0.257	0.880	-0.683	0.648	0.634	0.652	-0.544	0.709	0.083	-0.501
-0.314	1.138	-1.148	-0.116	-0.200	-0.792	0.303	0.452	0.279	-0.514	0.438	0.804	-0.720
0.537	-0.629	0.740	0.169	0.292	-0.559	-0.039	-0.485	-0.813	-0.317	-0.653	-0.184	-0.364
-0.313	0.696	-0.133	0.585	-0.087	0.167	-0.587	0.461	0.806	-0.586	-0.552	0.567	0.167
0.542	0.415	0.109	-0.539	-0.362	0.074	-1.018	0.707	0.137	0.566	-0.694	-0.874	-0.389
0.586	0.541	-0.075	-0.216	0.285	0.988	-0.184	0.624	-0.207	-0.883	0.405	0.454	-0.270
0.101	0.503	0.002	-0.612	-0.653	-0.785	0.131	0.353	0.370	-0.151	-1.186	0.530	0.477
0.572	-0.173	0.786	-0.581	-0.639	0.873	-0.985	0.668	-0.436	-0.238	0.028	-0.628	0.899
-0.688	1.022	-0.146	0.820	0.087	0.070	-0.160	-1.083	-0.366	-0.572	-0.511	-0.526	-0.540
0.185	-0.465	0.568	-0.190	-0.869	0.600	0.224	-0.398	0.466	-0.604	-0.751	-0.035	-1.010
-0.160	-0.753	-0.404	0.156	-0.163	-0.698	0.731	-0.459	-0.640	0.815	-0.001	-0.112	-0.060
-0.043	-0.641	0.071	-0.221	1.118	-0.864	0.880	0.372	-0.061	0.292	0.038	0.163	0.092
-0.667	-0.623	-0.086	-0.603	0.534	0.401	0.492	0.760	0.569	-0.590	0.346	-0.572	-0.319
0.488	0.228	-0.586	-0.278	-0.330	0.243	0.819	0.100	0.920	-0.952	0.872	-0.328	0.646
-0.392	-0.006	-0.635	0.278	0.328	-0.038	0.691	-0.541	-1.429	-0.745	0.391	-0.872	0.821
-0.132	0.166	0.057	-0.040	-0.788	-0.659	-0.776	0.825	0.742	-0.358	-0.471	-0.891	-0.375
0.362	-0.785	-0.366	0.434	-0.626	0.164	0.523	-0.035	0.651	-0.633	-0.822	0.329	-0.032
-0.541	0.671	-0.303	0.248	-0.003	0.592	-0.792	-0.038	0.587	-0.676	0.777	-0.790	-0.653
-0.191	0.885	-0.472	0.070	0.245	0.793	0.571	-0.783	0.815	-0.157	0.013	0.214	0.343
0.654	-0.354	-0.844	0.037	-0.152	0.014	-0.229	-0.500	-0.892	0.979	0.697	-0.281	-0.956
0.457	-0.098	-0.109	-1.010	-0.954	-0.081	-0.318	-0.058	0.762	0.228	-0.074	0.227	-0.216
-0.357	0.528	-0.924	1.163	0.773	-0.185	0.222	-0.067	-0.238	0.542	0.507	-0.015	-0.183
0.402	-0.592	0.837	-1.131	-0.116	-0.280	-0.581	0.150	-0.205	0.104	-0.556	0.930	-0.415
1.050	0.248	0.338	-0.313	0.126	-0.294	0.697	0.468	-0.770	1.099	-0.062	0.211	-0.511
1.184	0.765	0.188	-0.027	-1.071	0.195	-0.211	-0.675	-0.474	-0.606	0.377	-0.061	-0.491
0.489	-0.575	0.622	-0.115	-0.872	0.469	0.278	-0.377	-0.048	0.589	-0.176	-0.496	0.255

Fig. 6 reveals insightful information regarding the relative importance of input parameters on the weighted index of accidents. Among the factors considered, namely the Ratio of curvature,

Length of the segment, and Condition of the pavement, these parameters emerge as the most influential in determining the severity and occurrence of accidents. The Ratio of curvature denotes the degree of road curvature in proportion to its length, where a higher ratio indicates an increased accident risk. Meanwhile, the Length of the segment represents the distance covered by a specific road segment, and variations in its characteristics can impact accident likelihood. Equally significant is the Condition of the pavement, which highlights the crucial role of well-maintained road surfaces in promoting safety. By recognizing these parameters as the most vital, the findings from Fig. 6 emphasize the importance of proactive measures in improving road design, considering appropriate segment lengths, and prioritizing pavement maintenance. Addressing these influential factors has the potential to significantly reduce accidents and enhance overall road safety standards.

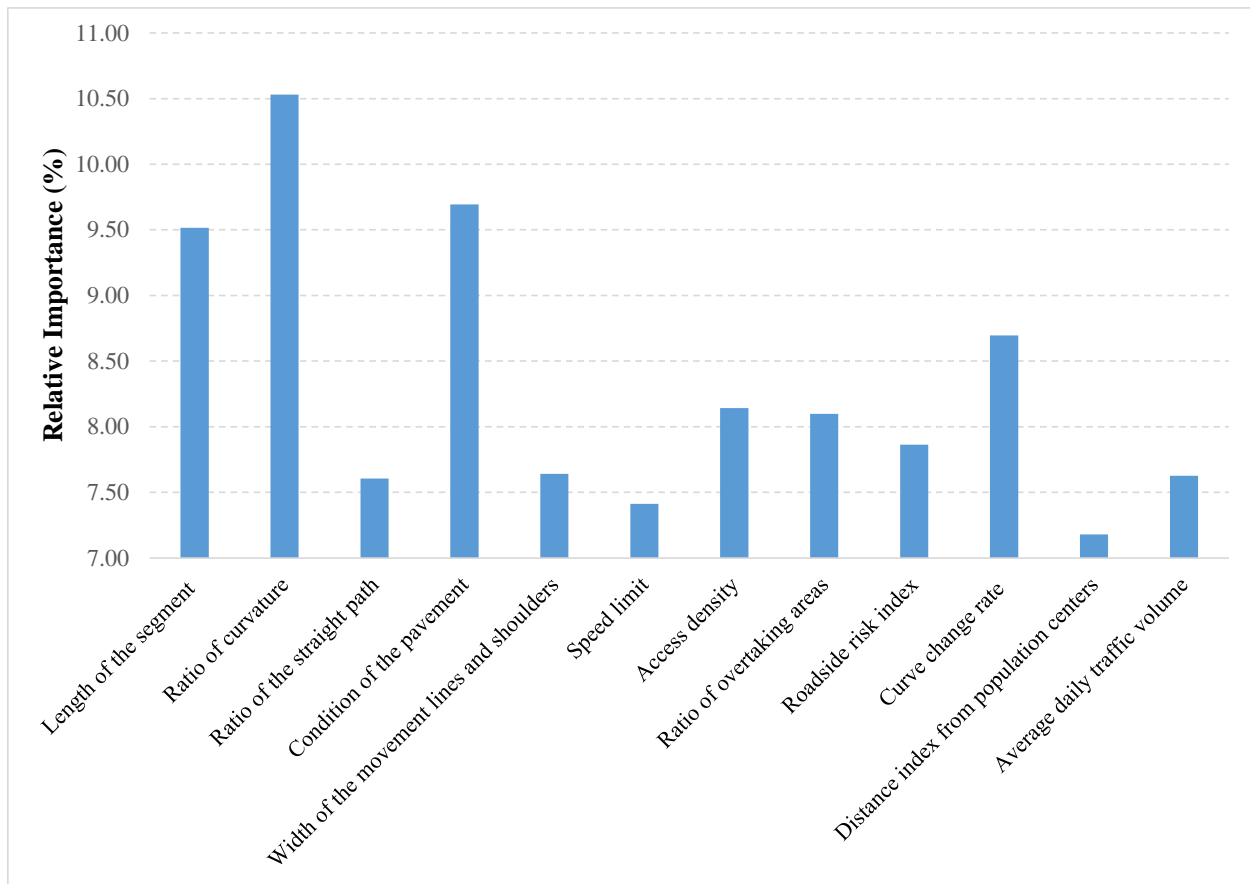


Fig. 6. Relative importance of the input data on weighted index of accidents.

6. Conclusions

Quantifying qualitative factors using modeling is one of the methods used in managerial decision making. The good results obtained from these methods in planning affairs have caused more tendency to use them. Based on this, in this article, a new method for entering the environmental, traffic and geometric characteristics of the road was presented in order to identify accident-prone parts, in such a way that the road is divided into units or parts with homogeneous physical characteristics, and as a result, a decision about the safety situation is made on the units with

specific characteristics. . This method considers accidents according to the interaction of the components that lead to it. Also, instead of points, it introduces a length of the route with known characteristics that can be improved in this specific interval.

The application of linear programming within the framework of data envelopment analysis enables a method for comparing road segments, intersections, areas, or entire routes in the domain of road safety organizations concerning the safety of other roads. In this current research, a novel approach was employed to determine the relative inefficiency of 154 road sections, introducing a new perspective in terms of defining input and output indicators based on the data envelopment analysis method for prioritizing road sections. The case study demonstrated that the existing method overlooked several sections despite their inadequate efficiency, whereas more favorable outcomes could be achieved with a relatively small investment. Consequently, this method provides scores (inefficiency) that facilitate appropriate ranking and prioritization of road segments, leading to better allocation of resources for enhancing road safety.

In addition to the data envelopment analysis, the study also incorporates artificial neural networks (ANNs) for further analysis of road safety. An idealized ANN model is developed, utilizing a database of various input parameters related to road segments, with the weighted index of accidents as the target variable. The results obtained from the ANN modeling reveal the relative importance of different parameters in determining accident severity. Factors such as the Ratio of curvature, Length of the segment, and Condition of the pavement are identified as the most influential. These findings provide valuable insights into the significant contributors to accidents and inform targeted interventions to mitigate risks.

By combining the data envelopment analysis and ANN modeling, this study presents a comprehensive approach to quantifying qualitative factors and analyzing road safety. The integration of these methods allows for a more holistic understanding of accident-prone areas and provides valuable guidance for decision-makers in resource allocation and risk mitigation strategies. Further research can explore the synergies between these techniques and other methodologies to enhance road safety practices and outcomes.

References

- [1] Saffarzadeh M, Sahebi S, Fallah Z. M. Classification of Recent Approaches in Road Safety Research. *Traffic Eng* 2013;52:5–10.
- [2] Taheri N, Ahmadi S, Malekiha M. The effect of pavement texture on road safety and reducing road accidents. *Second Int. Conf. Res. Find. Civ. Eng. Archit. Urban Plan.*, Tehran: 2016, p. 1–20.
- [3] World Health Organizatio. Global launch: decade of action for road safety 2011-2020. No. WHO/NMH/VIP11. 08. World Health Organization; 2011.
- [4] Elvik R, Veisten K. Barriers to the use of efficiency assessment tools in road safety policy. No. 785/2005. Institute of Transport Economics; 2005.
- [5] Ansari E, Mohamadi A, Saeidi S. Studying Social And Cultural Factors Affecting Inner-City Traffic Accidents (Case of Study: The Province of Kohgiluyeh and Boyer-Ahmad). *J Sociol Urban Stud* 2013;3:81–102.
- [6] Barandak F. Assessing Road Safety Performance in Iranian Provinces. *Majl Rahbord* 2022;29:207–41. <https://doi.org/10.22034/mr.2021.4473.4384>.

- [7] Naderpour H, Mirrashid M. Proposed soft computing models for moment capacity prediction of reinforced concrete columns. *Soft Comput* 2020;24:11715–29. <https://doi.org/10.1007/s00500-019-04634-8>.
- [8] Naderpour H, Mirrashid M. Moment capacity estimation of spirally reinforced concrete columns using ANFIS. *Complex Intell Syst* 2019. <https://doi.org/10.1007/s40747-019-00118-2>.
- [9] Khademi A, Behfarnia K, Kalman Šipoš T, Miličević I. The Use of Machine Learning Models in Estimating the Compressive Strength of Recycled Brick Aggregate Concrete. *Comput Eng Phys Model* 2021;4:1–25. <https://doi.org/10.22115/cepm.2021.297016.1181>.
- [10] Naderpour H, Khatami SM, Barros RC. Prediction of Critical Distance Between Two MDOF Systems Subjected to Seismic Excitation in Terms of Artificial Neural Networks. *Period Polytech Civ Eng* 2017. <https://doi.org/10.3311/PPci.9618>.
- [11] Rezaei Ranjbar P, Naderpour H. Probabilistic evaluation of seismic resilience for typical vital buildings in terms of vulnerability curves. *Structures* 2020;23:314–23. <https://doi.org/10.1016/j.istruc.2019.10.017>.
- [12] Sharbatdar MK, Vaez SRH, Amiri GG, Naderpour H. Seismic Response of Base-Isolated Structures with LRB and FPS under near Fault Ground Motions. *Procedia Eng* 2011;14:3245–51. <https://doi.org/10.1016/j.proeng.2011.07.410>.
- [13] Babaei A, Parker J, Moshaver P. Adaptive Neuro-Fuzzy Inference System (ANFIS) Integrated with Genetic Algorithm to Optimize Piezoelectric Cantilever-Oscillator-Spring Energy Harvester: Verification with Closed-Form Solution. *Comput Eng Phys Model* 2022;5:1–22. <https://doi.org/10.22115/cepm.2023.375302.1227>.
- [14] Bergland H, Pedersen PA. Efficiency and traffic safety with pay for performance in road transportation. *Transp Res Part B Methodol* 2019;130:21–35. <https://doi.org/10.1016/j.trb.2019.10.005>.
- [15] Blagojević A, Kasalica S, Stević Ž, Tričković G, Pavelkić V. Evaluation of Safety Degree at Railway Crossings in Order to Achieve Sustainable Traffic Management: A Novel Integrated Fuzzy MCDM Model. *Sustainability* 2021;13:832. <https://doi.org/10.3390/su13020832>.
- [16] Nahangi M, Chen Y, McCabe B. Safety-based efficiency evaluation of construction sites using data envelopment analysis (DEA). *Saf Sci* 2019;113:382–8. <https://doi.org/10.1016/j.ssci.2018.12.005>.
- [17] Seyedalizadeh Ganji SR, Rassafi A, Xu D-L. A double frontier DEA cross efficiency method aggregated by evidential reasoning approach for measuring road safety performance. *Measurement* 2019;136:668–88. <https://doi.org/10.1016/j.measurement.2018.12.098>.
- [18] Omrani H, Amini M, Alizadeh A. An integrated group best-worst method – Data envelopment analysis approach for evaluating road safety: A case of Iran. *Measurement* 2020;152:107330. <https://doi.org/10.1016/j.measurement.2019.107330>.
- [19] Wei Q, Zhang J, Zhang X. An inverse DEA model for inputs/outputs estimate. *Eur J Oper Res* 2000;121:151–63. [https://doi.org/10.1016/S0377-2217\(99\)00007-7](https://doi.org/10.1016/S0377-2217(99)00007-7).
- [20] Zhang G, Cui J. A general inverse DEA model for non-radial DEA. *Comput Ind Eng* 2020;142:106368. <https://doi.org/10.1016/j.cie.2020.106368>.
- [21] Yan H, Wei Q, Hao G. DEA models for resource reallocation and production input/output estimation. *Eur J Oper Res* 2002;136:19–31. [https://doi.org/10.1016/S0377-2217\(01\)00046-7](https://doi.org/10.1016/S0377-2217(01)00046-7).
- [22] Lertworasirikul S, Charnsethikul P, Fang S-C. Inverse data envelopment analysis model to preserve relative efficiency values: The case of variable returns to scale. *Comput Ind Eng* 2011;61:1017–23. <https://doi.org/10.1016/j.cie.2011.06.014>.
- [23] Lim D-J. Inverse DEA with frontier changes for new product target setting. *Eur J Oper Res* 2016;254. <https://doi.org/10.1016/j.ejor.2016.03.059>.
- [24] Amin GR, Al-Muharrami S, Toloo M. A combined goal programming and inverse DEA method for target setting in mergers. *Expert Syst Appl* 2018;115. <https://doi.org/10.1016/j.eswa.2018.08.018>.

- [25] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *Eur J Oper Res* 1978;2:429–44. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8).
- [26] Mayora JMP, Rubio RL. Relevant variables for crash rate prediction in Spain's two lane rural roads. *Proc. 82nd Transp. Res. Board Annu. Meet. Washingt. DC, USA, 2003*.
- [27] Jahanshahloo GR, Hosseinzadeh Lotfi F, Khanmohammadi M, Kazemimanesh M, Rezaie V. Ranking of units by positive ideal DMU with common weights. *Expert Syst Appl* 2010;37:7483–8. <https://doi.org/10.1016/j.eswa.2010.04.011>.
- [28] Cafiso S, Di Graziano A, Di Silvestro G, La Cava G, Persaud B. Development of comprehensive accident models for two-lane rural highways using exposure, geometry, consistency and context variables. *Accid Anal Prev* 2010;42:1072–9. <https://doi.org/10.1016/j.aap.2009.12.015>.
- [29] Zhang C, Ivan JN. Effects of Geometric Characteristics on Head-On Crash Incidence on Two-Lane Roads in Connecticut. *Transp Res Rec* 2005;1908:159–64. <https://doi.org/10.1177/0361198105190800119>.
- [30] Abdel-Aty MA, Radwan AE. Modeling traffic accident occurrence and involvement. *Accid Anal Prev* 2000;32:633–42. [https://doi.org/10.1016/S0001-4575\(99\)00094-9](https://doi.org/10.1016/S0001-4575(99)00094-9).
- [31] Rainey D, Parenteau MA, Kales SN. Sleep and Transportation Safety: Role of the Employer. *Sleep Med Clin* 2019;14:499–508. <https://doi.org/10.1016/j.jsmc.2019.08.007>.
- [32] Anastasopoulos P, Tarko A, Mannering F. Tobit Analysis of Vehicle Accident Rates on Interstate Highways. *Accid Anal Prev* 2008;40:768–75. <https://doi.org/10.1016/j.aap.2007.09.006>.
- [33] AASHTO. *Roadside Design Guide*, Washington: 1996.
- [34] Ayati I. *Irans road accidents (analysis, comparison and cost calculation)*. Ferdowsi Mashhad University Publications; 1992.
- [35] Yazdani H. *Investigation of identification and prioritization methods of Fakhiz intersection points and their effectiveness*. 2010.
- [36] Naderpour H, Rafiean AH, Fakharian P. Compressive strength prediction of environmentally friendly concrete using artificial neural networks. *J Build Eng* 2018;16:213–9. <https://doi.org/10.1016/j.jobbe.2018.01.007>.
- [37] Naderpour H, Sharei M, Fakharian P, Heravi MA. Shear Strength Prediction of Reinforced Concrete Shear Wall Using ANN, GMDH-NN and GEP. *J Soft Comput Civ Eng* 2022;6:66–87. <https://doi.org/10.22115/SCCE.2022.283486.1308>.
- [38] Rezazadeh Eidgahee D, Jahangir H, Solatifar N, Fakharian P, Rezaeemanesh M. Data-driven estimation models of asphalt mixtures dynamic modulus using ANN, GP and combinatorial GMDH approaches. *Neural Comput Appl* 2022;34:17289–314. <https://doi.org/10.1007/s00521-022-07382-3>.
- [39] Ghanizadeh AR, Ghanizadeh A, Asteris PG, Fakharian P, Armaghani DJ. Developing bearing capacity model for geogrid-reinforced stone columns improved soft clay utilizing MARS-EBS hybrid method. *Transp Geotech* 2023;38:100906. <https://doi.org/10.1016/j.trgeo.2022.100906>.
- [40] Fakharian P, Rezazadeh Eidgahee D, Akbari M, Jahangir H, Ali Taeb A. Compressive strength prediction of hollow concrete masonry blocks using artificial intelligence algorithms. *Structures* 2023;47:1790–802. <https://doi.org/10.1016/j.istruc.2022.12.007>.