Machine learning established by using crowdsourced investigation vehicle data for forecast of expressway crash risk

Lee Chen, Don Chen, Chang Li, Bing Pan, Lixuan Zhang, Zheng Xiang
Faculty of Computer Science and Information System, Universiti Teknologi MARA (UiTM), Malaysia

ABSTRACT

Real-time prediction of crash risk can support traffic incident management by generating critical information for practitioners to allocate resources for responding to anticipated traffic crashes proactively. Unlike previous studies using archived traffic data covering a limited highway environment such as a segment or corridor, this study uses a statewide live traffic database from HERE to develop real-time traffic crash prediction models. This database pro- vides crowdsourced probe vehicle data that are high-resolution real-time traffic speed for the entire freeway network (nearly 2,000 miles) in Alabama. This study aims to use machine learning models to predict crash risk on freeways according to pre-crash traffic dynamics (e.g., mean speed, speed reduction) along with static freeway attributes. Traffic speed char- acteristics were extracted from the HERE database for both pre-crash and crash-free traffic conditions. Random Forest (RF), Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) were developed and compared. Separate models were estimated for three major crash types: single-vehicle, rear-end, and sideswipe crashes. The model predic- tion accuracy indicated that the RF models outperform other models. Models for rear-end crashes are found to have greater accuracy than other models, which implies that rear-end crashes have a significant relationship with pre-crash traffic dynamics and are more predict- able. The traffic speed factors that are ranked high in terms of feature importance are the speed variance and speed reduction prior to crashes. According to partial dependence plots, the rear-end crash risk is positively related to the speed variance and speed reductions. More results are discussed in the paper.

KEYWORDS: Crowd sourced probe vehicle data; machine learning; pre- crash traffic dynamics; real- time crash risk prediction; rear-end crash

1.0 INTRODUCTION

Traffic crashes are a leading cause of non-natural deaths in the United States for people under 54 years old and often lead to traffic congestions that waste a lot of time of travelers on road, especially during peak hours. The crash-led congestions pose a risk of secondary crashes which could further deteriorate the traffic conditions and increase travel delays. To mitigate the impacts of traffic crashes, Traffic Incident Management (TIM) plays an essential role in detecting and responding to crashes and restoring the traffic [1-5]. Most of current TIM practices are a reactive process that traffic crashes or other incidents are responded after a crash or incident has occurred and this process heavily relies on incident reporting and detection. Significant research efforts have been dedicated to developing or improving the methods of detecting traffic crashes or incident detection. In addition to the traditional traffic surveillance methods such as closed circuit television (CCTV) systems, 911 services, and highway patrols, researchers have also developed methods using loop detector data, probe vehicle data, and other emerging sensor data [6-10]. However, this reactive process is inherently associated with response delays between the occurrence of the incident/crash and the detection of an incident/crash. The consequences of response delays are longer traffic queues, increased risk of secondary crashes, and reduced likelihood of crash survivals (FHWA, 2018). A practical question arises as to whether the TIM responses can be upgraded from a reactive to proactive process [11-17]. Theoretically, it is possible and one strategy to reduce or completely avoid the response delay is to predict the probability of having a crash based on traffic dynamics prior to the event of a crash so that traffic incident managers can proactively allocate resources for responding to the sites which have high risk of crashes. Many studies have developed a variety of models to explain the relation- ships between traffic characteristics and crash risk, and some studies focused on improving the accuracy of crash risk pre- diction models. However, a common feature of these studies is that the data used in these studies are usually from a limited environment such as a road segment or a corridor which may not be able to capture the diverse road environments [18-23]. With the advances of information and communication technologies, an increasing amount of crowdsourced traffic data are being generated by mobile devices and electronic sensors and pre- sent an opportunity for researchers to scrutinize the traffic dynamics in greater depth and width than ever before. The objective of this study is to develop real-time crash risk prediction models by exploiting crowdsourced probe vehicle data that are not limited to a specific environment but cover diverse road environments [23-31]. Specifically, the crowdsourced data used in this study are from a live HERE database that reports the real-time traffic conditions of the entire freeway net- work in the State of Alabama. The models developed in this study are expected to predict the crash risk on freeways according to pre-crash traffic dynamics along with static freeway attributes. The pre-crash traffic dynamics are traffic speed characteristics prior to the event of crashes extracted from the HERE database for both pre-crash and crash-free traffic conditions. The modeling methods include both regression-based and machine learning methods, including random forest (RF), extreme gradient boosting (XGBoost), and support vector machine (SVM). Considering the different causes of crashes of different types, this study estimates separate models for major crash types including single-vehicle, rear-end, and sideswipe crashes. Further, unlike previous studies that focused on the increasing model performance of machine learning models, this study explores the methods to interpret the machine learning modeling results and attempts to translate the modeling results into action- able items to support proactive TIM responses [32-40].

2.0 LITERATURE REVIEW

Topics related to traffic incident management (TIM) as an important component of the Transportation System Management and Operations (TSMO) have been extensively researched. According to the tem- poral phrases of incident management, relevant research can be grouped into three categories: 1) Before the occurrence of an incident or proactive traffic incident management 2) During the occurrence of an incident or the incident detection research and 3) After the occurrence of an incident or incident impact analysis (e.g., queue length and incident duration). The first two categories may use the same traffic data to develop models that relate the traffic flow dynamics to crashes. The critical difference between these two categories is whether their models use the traffic information prior to or during/after the event of a crash. Literally, the models that use traffic information during or after the occurrence of a crash are for crash detection and not for crash prediction [1-7]. Given the objective of this study, the literature review is limited to studies that developed crash risk prediction models that use the pre- crash traffic information. Studies that develop models to detect crashes are not discussed in the literature. The studies that focused on crash risk prediction have generated significant insights to support pro- active traffic incident management. Some earlier studies relied on the static data (e.g., average daily traffic, land use type, road geometry characteristics) to identify sites with high crash risks at an aggregate level, e.g., the crash count at a segment with limited sight distance. In recent years, researchers take advantage of traffic data generated by electronic sensors such as loop detectors and mobile devices and have developed a variety of models to predict the crash risk in a real- time or near-real-time manner [8-13]. Table 1 summarizes selected studies that used pre-crash traffic dynamic information to develop crash risk prediction models. A large portion of relevant studies relied on the traffic loop detector data. Loop detectors play an essential role in traffic management and have been widely installed at intersections and on freeways to monitor the traffic. The loop detector data can be used to generate common traffic flow characteristics such as volume, speed, and density. The loop detector data are readily available for researchers to explore the relationships between crash risk and traffic flow characteristics. However, such data are limited to sites with loop detectors and do not cover the entire road network. Some studies used sensor data from Microwave Vehicle Detection System (MVDS) Automated Vehicle Identification (AVI), and Bluetooth Detectors [14-19]. Similar to loop detectors, these sensors are fixed on the road or roadside, and the data from these sensors have the same coverage limitation. Probe vehicle data are not limited to specific road segments or areas. Researchers started to use probe vehicles to understand the relationship between crash risk and traffic flow dynamics. The focus of the studies using probe vehicle data is primarily on the risk of secondary crashes instead the initial or primary crashes. From the model perspective, researchers have developed a variety of regression-based and also machine learning models. The most commonly used regression modeling approach is the logistic model, including binary logistic regression, sequential logistic model, and random parameters logistic regression. Among the machine learning models, the Support Vector Machine (SVM) has been frequently used by researchers, and the SVM models produced decent accuracies [20-27]. Recently, more sophisticated modeling approaches were introduced for crash risk prediction. For example, used a convolutional neural network model to predict crash risk based on a deep convolutional generative adversarial network (DCGAN) and achieved good performance. In most of these studies, the crash types were not discussed, and it is likely that their models were to predict the risk of all crashes. Assuming that crashes of different types could have different relationships with the traffic dynamics, some studies built separate models for specific crash types such as rear-end crashes and sideswipe crashes [28-31].

Table 1. Selected studies on crash risk prediction.

Authors	Year	Study area	Data source	Model	Key independent variables or features (traffic dynamics)	Dependent variable or target features
Hossain and Muromachi (2012)	2012	11.9 km Shibuya 3 and 13.5 km Shinjuku 4 expressways, Tokyo	Loop detector	Bayesian network (model)	Congestion index, speed and detector occupancy	Crash risk
Qu et al. (2012)	2012	9.3-mile I-894, Milwaukee, Wisconsin	Loop detector	Support vector machine	Speed, occupancy, volume	Rear-end crash risk
Xu et al. (2013)	2013	29-mile I-880, San Francisco, California	Loop detector	Sequential logit model	Vehicle count, speed, detector occupancy	Crash risk at different severity levels
Yu and Abdel-Aty (2013)	2013	15-mile I-70, Colorado	Remote Traffic Microwave Sensor	Support vector machine	Speed, occupancy, volume	Crash risk
Qu et al. (2013)	2013	9.3-mile I – 894 Milwaukee, Wisconsin	Loop detector	Support vector machine	Traffic state and variances between adjacent lanes	Side-wipe crash risk
Wang et al. (2015)	2015	22-mile SR 408, Central Florida	Microwave Vehicle Detection System (MVDS) detector	Multilevel Bayesian logistic regression model	Speed, occupancy, volume	Crash risk for weaving segments
Park and Haghani (2016)	2016	51-mile I-695	Probe vehicle data	Neural network model	Speed	Secondary incident occurrences risk
Wu et al. (2018a)	2018	7-mile I-75, 11-mile I-4, Tampa and 11-mile SR-408, Florida	Loop and radar detectors; Microwave Vehicle Detection System (MVDS) sensors	Binary logistic regression	Speed, occupancy, volume	Crash risk
Wu et al. (2018b)	2018	Segment of I-4 in Florida	Remote Traffic Microwave Sensor (RTMS)	Random parameters logistic regression	Speed, volume	Rear-end crash risk
Cai et al. (2020)	2020	24-mile SR 408 in Central Florida	Microwave Vehicle Detection System (MVDS) detector	Convolutional Neural Network (CNN)	Speed, occupancy, volume	Crash risk
Huang et al. (2020)	2020	13.78-mile I-235, Des Moines	Roadside radar sensors	Support vector machine	Speed, occupancy, volume	Crash risk

To the best of the authors' knowledge, crowd- sourced probe vehicle data have not been extensively explored by researchers regarding crash risk prediction. Using such data could overcome the limitation of sensor data that cover only road segments where fixed detection units are installed. Besides, models using crowd sourced data could uncover the relation- ships between traffic dynamics and crash risk over a diverse road or land use environment, and such models may have greater applicability. Using statewide crowd sourced probe vehicle data, this study attempts to develop machine learning models to predict crash risk for the entire freeway network in Alabama. Further, previous studies focused on improving the model performance, especially when the machine learning models are adopted. To improve the transportation systems by reducing the crashes, it is essential to interpret the modeling results and translated models into actionable items such as countermeasures to address the issues at some locations where high crash risks are identified [32-40].

3.0 RESEARCH METHODOLOGY

The Random Forest (RF) is one of the most popular ensemble methods in machine learning for data classification. The RF is a combination of multiple decision trees, and each tree is a class prediction model. The result of RF is the average pre- diction of all decision trees. The general idea behind the random forest is shown. Compared with the individual tree or class prediction model, the RF is expected to have lower variance and over- come the overfitting problem. In this study, the RF was adopted to model the complex nonlinear rela- tionships between the crash risk and associated factors based on its flexible modeling structure. In this study, two hyperparameters, including the number of decision trees in the forest and the number of fea- tures, were considered for each decision tree when splitting a node, and the three-fold-cross validation was performed to optimize the model performance. To determine the number of features at each split, two commonly used methods, including the square root of the total number of variables and base two logarithm of the total number of variables, were tested. For the number of trees, values from 50 to 550 at an interval of 50 were tested. Overall, 20 possible combinations were examined for gaining the optimized parameters.The crowdsourced probe vehicle data used in this study are from the HERE traffic database, which pro- vides the live traffic information for the entire freeway network in Alabama (HERE, 2021). Figure 1 shows the freeway network in Alabama (I-459, I-59, I-65, I- 85, I-10, I-20, I-20/I-59, I-22, I-565). The HERE traffic database provides the live speed information updated every minute for each traffic management channel (TMC). A TMC is a pre-defined section of the road for traffic data reporting (Esri, 2020). As shown in Figure 2, the size of TMC can range from under 0.1 mile to a few miles. Table 2 shows the descriptive statistics of 635 TMCs for the Alabama freeway net- work. The total length of these TMCs is 1,976 miles, and the average length is 3.1 miles. Note, the TMCs are specified for each direction. One TMC may be split into several shorter dynamic units to capture the speed information in a higher resolution when the traffic flow speed changes quickly within a TMC.

This study paired the crowdsourced traffic data with the traffic crash data to extract the pre-crash traffic flow dynamics. The traffic crash data were obtained from the CARE crash database maintained by the Alabama Department of Transportation (ALDOT) (FHWA, 2020). The crashes occurred in 2019 on the targeted freeways (I-459, I-59, I-65, I-85, I-10, I-20, I-20/I-59, I-22, I-565) within the state of Alabama were extracted from the CARE crash data- base. After data cleaning (removing crash records that miss the time or location information), the remaining 9,997 crash records were used to extract the pre-crash traffic information from the HERE database. Table 3 shows the descriptive statistics of sampled crashes regarding crash types. In total, the top three crash types - single-vehicle crash, rear-end crash and side- swipe crash account for 35.6%, 34.7%, and 18.3% of the total crashes, respectively. Spatially, approximately 40% of the crashes occurred in the I-65, followed by 14.4% of the crashes that occurred in I-20/I-59. The spatial distribution of interstate freeway crash frequency and crash density is shown in Figure 3(a) and (b), respectively.

Figure 1. Selected Interstate freeways.

To reveal the impact of the diverse road environments on crash risk, this study pulled the road infra- structure data from the Highway Performance Monitoring System (HPMS) database. HPMS database provides road environment variables, including annual average daily traffic (AADT), number of lanes, distance to the closest upstream ramp, and land use. As shown in Figure 4, the spatial distribution of freeway traffic volume (AADT) is consistent with the spatial distribution of crash density.

4.0 RESULT

Data used in this study are from three sources, including the HERE traffic database, the CARE crash database, and the HPMS highway infrastructure data- base. This study undertook a significant effort to link the data features and created variables for crash risk prediction. Figure 5 shows the framework of data linkage and variable creation. First, the TMCs in the HERE database were linked to the crashes from the CARE database according to spatial locations of TMCs and crashes. To ensure enough spatial coverage for extracting the traffic data, this study linked crashes to TMCs within 5 miles upstream and downstream of the crash location. As a result, each crash was paired with a 10-mile segment for traffic data extraction from linked HERE TMCs. Second, from paired HERE TMCs this study extracted the traffic speed records before the crash occurrence time. This study considered three different pre-crash time points (10, 15, and 20 minutes before the crash) to extract the traffic data. To ensure the modeling results are comparable, this study extracted the traffic data from the same 20-min intervals. Note, other time intervals were also attempted, and the 20-min intervals were selected given the model performance. The different pre-crash time points start from the middle of the 20-min intervals. For instance, if a crash occurred at 8:00 am, this study extracted the traffic data between 7:40 am, and 8:00 am for the 10-minute pre-crash time point, between 7:35 am and 7:55 am for the 15-min time point, and between 7:30 am and 7:50 am for the 20- min pre-crash time point. Then, for each crash, a spatial-temporal speed matrix was created to document the speed records extracted from the HERE TMCs. Figure 6 shows a schematic diagram of part of the spatial-temporal speed matrix. The speed matrix has a resolution of 0.1 miles by 1minute. Last, the road environment attributes from the HPMS database were matched to crashes and TMCs. Data used in this study are from three sources, including the HERE traffic database, the CARE crash database, and the HPMS highway infrastructure data- base. This study undertook a significant effort to link the data features and created variables for crash risk prediction. Figure 5 shows the framework of data linkage and variable creation. First, the TMCs in the HERE database were linked to the crashes from the CARE database according to spatial locations of TMCs and crashes. To ensure enough spatial coverage for extracting the traffic data, this study linked crashes to TMCs within 5 miles upstream and downstream of the crash location. As a result, each crash was paired with a 10-mile segment for traffic data extraction from linked HERE TMCs. Second, from paired HERE TMCs this study extracted the traffic speed records before the crash occurrence time. This study considered three different pre-crash time points (10, 15, and 20 minutes before the crash) to extract the traffic data. To ensure the modeling results are comparable, this study extracted the traffic data from the same 20-min intervals. Note, other time intervals were also attempted, and the 20-min intervals were selected given the model performance. The different pre-crash time points start from the middle of the 20-min intervals. For instance, if a crash occurred at 8:00 am, this study extracted the traffic data between 7:40 am, and 8:00 am for the 10-minute pre-crash time point, between 7:35 am and 7:55 am for the 15-min time point, and between 7:30 am and 7:50 am for the 20- min pre-crash time point. Then, for each crash, a spatialtemporal speed matrix was created to document the speed records extracted from the HERE TMCs. Figure 6 shows a schematic diagram of part of the spatial-temporal speed matrix. The speed matrix has a resolution of 0.1 miles by 1minute. Last, the road environment attributes from the HPMS database were matched to crashes and TMCs. Data used in this study are from three sources, including the HERE traffic database, the CARE crash database, and the HPMS highway infrastructure data- base. This study undertook a significant effort to link the data features and created variables for crash risk prediction. Figure 5 shows the framework of data linkage and variable creation. First, the TMCs in the HERE database were linked to the crashes from the CARE database according to spatial locations of TMCs and crashes. To ensure enough spatial coverage for extracting the traffic data, this study linked crashes to TMCs within 5 miles upstream and downstream of the crash location. As a result, each crash was paired with a 10-mile segment for traffic data extraction from linked HERE TMCs. Second, from paired HERE TMCs this study extracted the traffic speed records before the crash occurrence time. This study considered three different pre-crash time points (10, 15, and 20 minutes before the crash) to extract the traffic data. To ensure the modeling results are comparable, this study extracted the traffic data from the same 20-min intervals. Note, other time intervals were also attempted, and the 20-min intervals were selected given the model performance. The different pre-crash time points start from the middle of the 20-min intervals. For instance, if a crash occurred at 8:00 am, this study extracted the traffic data between 7:40 am, and 8:00 am for the 10-minute pre-crash time point, between 7:35 am and 7:55 am for the 15-min time point, and between 7:30 am and 7:50 am for the 20- min pre-crash time point. Then, for each crash, a spatialtemporal speed matrix was created to document the speed records extracted from the HERE TMCs. Figure 6 shows a schematic diagram of part of the spatial-temporal speed matrix. The speed matrix has a resolution of 0.1 miles by 1minute. Last, the road environment attributes from the HPMS database were matched to crashes and TMCs. of dynamic sub-TMCs within a 20-min time interval at different pre-crash time points.

Table 2. Descriptive statistics for TMC length (Unit: mile).

Primary Road	Direction	Number of TMC	Total Length of TMC	Min	Max	Mean	Var
I-459	NS	28	65.401	0.082	4.581	2.336	1.850
I-59	NS	54	222.200	0.026	17.046	4.115	18.057
I-65	NS	214	731.130	0.029	13.872	3.416	9.833
I-85	NS	64	159.891	0.064	9.242	2.498	5.924
I-10	WE	43	132.749	0.044	13.426	3.087	10.075
I-20	WE	56	168.769	0.042	7.539	3.014	4.109
I-20/I-59	WE	88	259.822	0.027	9.340	2.953	7.285
I-22	WE	56	191.724	0.034	7.284	3.424	2.517
I-565	WE	32	44.009	0.238	4.485	1.375	1.019
Total		635	1,975.695	0.026	17.046	3.111	8.119

Besides the traffic dynamics prior to crashes, this study also extracted the traffic speed information for crash-free conditions paired with the crash conditions. Specifically, for each crash record, one crash-free record was generated at the same location and the same time of day but on a different day when there was no crash. For the crash-free records, the same sets of variables were created through the same methods of data linkage and variable creation. Note that crash-free observations can be created for any times and locations where no crashes occurred.

5.0 CONCLUSION

Real-time crash risk is expected to support proactive traffic incident management by generating critical information for traffic managers to allocate incident response resources to high-risk sites before the occur- rence of crashes. This study developed crash risk pre- diction models by taking advantage of the HERE crowdsourced probe vehicle data from a live database that reports and archives minute-by-minute real-time traffic speeds for freeways. The data are not limited to a specific road segment and cover the entire freeway

network in Alabama. Based on the HERE data, this study created a variety of variables to capture the pre- crash traffic dynamics, which are traffic speed charac- teristics prior to the event of crashes, measured by mean speed, speed variance, and speed reduction. In addition to the pre-crash traffic dynamics, this study also extracted the traffic dynamics for crash-free con- ditions. With the data processed for pre-crash and crash-free traffic dynamics, this study developed logis- tic regression and machine learning models to predict the crash risk on freeways according to traffic dynam- ics along with static freeway attributes. Three machine learning approaches, including random forest (RF), support vector machine (SVM), and extreme gradient boosting (XGBoost) were tested and compared. Separate models were developed for all crashes, single- vehicle crashes, rear-end crashes, and sideswipe crashes. Modeling results were interpreted by using permutation feature importance and partial depend- ence plots.

The results indicated that the traffic dynamics closer to the event of a crash are more predictive of the crash risk. Models for rear-end crashes are found to have a greater accuracy than other models, which implies that rear-end crashes have a significant rela- tionship with pre-crash traffic dynamics (especially speed) and are more predictable than other crashes. For rear-end crashes, in comparison with the logistic regression model, XGBoost model, and SVM model, the RF model had a slightly improved prediction per- formance according to the AUC value (68.4%) and the accuracy (ACC) (65.1%). For rear-end crash mod- els, this study calculated the permutation feature importance for each variable and computed the partial dependence for pre-crash speed-related top-ranked variables according to the feature importance. Though the feature importance rankings are different across models, the pre-crash speed-related variables are gen- erally ranked high in all three models. According to the estimated partial dependence, the rear-end crash risk is positively related to the speed variance and speed reductions. A higher rear-end crash risk is asso- ciated with a more significant speed variance upstream, and the risk for a rear-end crash increases when the traffic speed at this location decreases significantly.

This study contributes by using crowdsourced probe vehicle data to develop real-time models to pre- dict crash risk on freeways. Such data have been increasingly used by agencies for traffic monitoring and management at a reasonable cost without instal- ling and maintaining electronic sensors on the road or roadside. Agencies with such data could potentially implement the crash risk prediction models to facili- tate their traffic incident responses. This study also identified the issues using such data for developing crash risk models. Continuing efforts are needed to examine the accuracy of the crowdsourced probe vehicle data, which is likely to vary across the geo- graphic areas and times. Future research will expand the modeling efforts to develop models that account for the data issues such as unobserved heterogeneity.

REFERENCES

[1] Dimitrijevic, Branislav, et al. Segment-Level Crash Risk Analysis for New Jersey Highways Using Advanced Data Modeling. No. CAIT-UTC-NC62. Rutgers University. Center for Advanced Infrastructure and Transportation, 2020.
[2] Li, Chang, et al. "Machine learning for text mining based on prediction of occupational accidents and safety risk calculation." Australian Journal of Engineering and Applied Science 13.6 (2020): 11-17.
[3] Bahrami, Javad, Viet B. Dang, Abubakr Abdulgadir, Khaled N. Khasawneh, Jens-Peter Kaps, and Kris Gaj. "Lightweight implementation of the lowmc block cipher protected against side-channel attacks." In *Proceedings of the 4th ACM Workshop on Attacks and Solutions in Hardware Security*, pp. 45-56. 2020.
[4] Chen, Lee, et al. "Machine learning established by using crowdsourced investigation vehicle data for forecast of expressway crash risk ." International Journal of Applied Science and Information Science 11.8 (2020): 356-363.
[5] Ahmadinejad, Farzad, Javad Bahrami, Mohammd Bagher Menhaj, and Saeed Shiry Ghidary. "Autonomous Flight of Quadcopters in the Presence of Ground Effect." In *2018 4th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*, pp. 217-223. IEEE, 2018.
[6] Zhang, Lixuan, et al. "Machine Learning Models established toward the Car Smash Injury Difficulty." European Journal of Applied Engineering and Basic Sciences 19.17 (2020): 4678-4685.
[7] Bozorgasl, Zavareh, and Mohammad J. Dehghani. "2-D DOA estimation in wireless location system via sparse representation." In *2014 4th International Conference on Computer and Knowledge Engineering (ICCKE)*, pp. 86-89. IEEE, 2014.
[8] Xiang, Zheng, et al. "Machine Learning process for injury severity prediction and Traffic accidents classification." International Journal of Management System and Applied Science 23.12 (2020): 997-1003.
[9] Amini, Mahyar, and Aryati Bakri. "Cloud computing adoption by SMEs in the Malaysia: A multi-perspective framework based on DOI theory and TOE framework." Journal of Information Technology & Information Systems Research (JITISR) 9.2 (2015): 121-135.
[10] Silva, Philippe Barbosa, Michelle Andrade, and Sara Ferreira. "Machine learning applied to road safety modeling: A systematic literature review." Journal of traffic and transportation engineering (English edition) 7.6 (2020): 775-790.
[11] Amini, Mahyar. "The factors that influence on adoption of cloud computing for small and medium enterprises." (2014).
[12] AlMamlook, Rabia Emhamed, et al. "Comparison of machine learning algorithms for predicting traffic accident severity." 2019 IEEE Jordan international joint conference on electrical engineering and information technology (JEEIT). IEEE, 2019.
[13] Amini, Mahyar, et al. "Development of an instrument for assessing the impact of environmental context on adoption of cloud computing for small and medium enterprises." Australian Journal of Basic and Applied Sciences (AJBAS) 8.10 (2014): 129-135.
[14] Rezapour, Mahdi, Amirarsalan Mehrara Molan, and Khaled Ksaibati. "Analyzing injury severity of motorcycle at-fault crashes using machine learning techniques, decision tree and logistic regression models." International journal of transportation science and technology 9.2 (2020): 89-99.
[15] Amini, Mahyar, et al. "The role of top manager behaviours on adoption of cloud computing for small and medium enterprises." Australian Journal of Basic and Applied Sciences (AJBAS) 8.1 (2014): 490-498.
[16] Rahim, Md Adilur, and Hany M. Hassan. "A deep learning based traffic crash severity prediction framework." Accident Analysis & Prevention 154 (2021): 106090.
[17] Amini, Mahyar, and Nazli Sadat Safavi. "Critical success factors for ERP implementation." International Journal of Information Technology & Information Systems 5.15 (2013): 1-23.
[18] Siebert, Felix Wilhelm, and Hanhe Lin. "Detecting motorcycle helmet use with deep learning." Accident Analysis & Prevention 134 (2020): 105319.
[19] Amini, Mahyar, et al. "Agricultural development in IRAN base on cloud computing theory." International Journal of Engineering Research & Technology (IJERT) 2.6 (2013): 796-801.
[20] Yang, Yang, et al. "Identification of dynamic traffic crash risk for cross-area freeways based on statistical and machine learning methods." Physica A: Statistical Mechanics and Its Applications 595 (2022): 127083.
[21] Amini, Mahyar, et al. "Types of cloud computing (public and private) that transform the organization more effectively." International Journal of Engineering Research & Technology (IJERT) 2.5 (2013): 1263-1269.
[22] Fu, Yuchuan, et al. "A decision-making strategy for vehicle autonomous braking in emergency via deep reinforcement learning." IEEE transactions on vehicular technology 69.6 (2020): 5876-5888.
[23] Amini, Mahyar, and Nazli Sadat Safavi. "Cloud Computing Transform the Way of IT Delivers Services to the Organizations." International Journal of Innovation & Management Science Research 1.61 (2013): 1-5.
[24] Alkinani, Monagi H., Wazir Zada Khan, and Quratulain Arshad. "Detecting human driver inattentive and

aggressive driving behavior using deep learning: Recent advances, requirements and open challenges." Ieee Access 8 (2020): 105008-105030.
[25] Amini, Mahyar, and Nazli Sadat Safavi. "A Dynamic SLA Aware Heuristic Solution For IaaS Cloud Placement Problem Without Migration." International Journal of Computer Science and Information Technologies 6.11 (2014): 25-30.
[26] Wahab, Lukuman, and Haobin Jiang. "Severity prediction of motorcycle crashes with machine learning methods." International journal of crashworthiness 25.5 (2020): 485-492.
[27] Amini, Mahyar, and Nazli Sadat Safavi. "A Dynamic SLA Aware Solution For IaaS Cloud Placement Problem Using Simulated Annealing." International Journal of Computer Science and Information Technologies 6.11 (2014): 52-57.
[28] Muhammad, Khan, et al. "Deep learning for safe autonomous driving: Current challenges and future directions." IEEE Transactions on Intelligent Transportation Systems 22.7 (2020): 4316-4336.
[29] Sadat Safavi, Nazli, et al. "An effective model for evaluating organizational risk and cost in ERP implementation by SME." IOSR Journal of Business and Management (IOSR-JBM) 10.6 (2013): 70-75.
[30] Cai, Qing, et al. "Applying a deep learning approach for transportation safety planning by using high-resolution transportation and land use data." Transportation research part A: policy and practice 127 (2019): 71-85.
[31] Sadat Safavi, Nazli, Nor Hidayati Zakaria, and Mahyar Amini. "The risk analysis of system selection and business process re-engineering towards the success of enterprise resource planning project for small and medium enterprise." World Applied Sciences Journal (WASJ) 31.9 (2014): 1669-1676.
[32] Wahab, Lukuman, and Haobin Jiang. "A comparative study on machine learning based algorithms for prediction of motorcycle crash severity." PLoS one 14.4 (2019): e0214966.
[33] Sadat Safavi, Nazli, Mahyar Amini, and Seyyed AmirAli Javadinia. "The determinant of adoption of enterprise resource planning for small and medium enterprises in Iran." International Journal of Advanced Research in IT and Engineering (IJARIE) 3.1 (2014): 1-8.
[34] Ziakopoulos, Apostolos, and George Yannis. "A review of spatial approaches in road safety." Accident Analysis & Prevention 135 (2020): 105323.
[35] Safavi, Nazli Sadat, et al. "An effective model for evaluating organizational risk and cost in ERP implementation by SME." IOSR Journal of Business and Management (IOSR-JBM) 10.6 (2013): 61-66.
[36] Mannering, Fred, et al. "Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis." Analytic methods in accident research 25 (2020): 100113.
[37] Khoshraftar, Alireza, et al. "Improving The CRM System In Healthcare Organization." International Journal of Computer Engineering & Sciences (IJCES) 1.2 (2011): 28-35.
[38] Mokhtarimousavi, Seyedmirsajad. "A time of day analysis of pedestrian-involved crashes in California: Investigation of injury severity, a logistic regression and machine learning approach using HSIS data." Institute of Transportation Engineers. ITE Journal 89.10 (2019): 25-33.
[39] Abdollahzadegan, A., Che Hussin, A. R., Moshfegh Gohary, M., & Amini, M. (2013). The organizational critical success factors for adopting cloud computing in SMEs. Journal of Information Systems Research and Innovation (JISRI), 4(1), 67-74.
[40] Silva, Philippe Barbosa, Michelle Andrade, and Sara Ferreira. "Machine learning applied to road safety modeling: A systematic literature review." Journal of traffic and transportation engineering (English edition) 7.6 (2020): 775-790.