IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 10, No. 2, June 2021, pp. 365~373
ISSN: 2252-8938, DOI: 10.11591/ijai.v10.i2.pp365-373

Spatial analysis model for traffic accident-prone roads classification: a proposed framework

Anik Vega Vitianingsih[1], Nanna Suryana[2], Zahriah Othman[3]
[1]Department of Informatics, Universitas Dr. Soetomo, Surabaya, Indonesia
[1,2,3]Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

Article Info

Article history:

Received Feb 25, 2020
Revised Dec 10, 2020
Accepted Apr 2, 2021

Keywords:

GIS software engineering
Hybrid estimation model-based
MCDM-ANNs
Proposed framework
Spatial analysis model
Traffic accident-prone roads

ABSTRACT

The classification method in the spatial analysis modeling based on the multi-criteria parameter is currently widely used to manage geographic information systems (GIS) software engineering. The accuracy of the proposed model will play an essential role in the successful software development of GIS. This is related to the nature of GIS used for mapping through spatial analysis. This paper aims to propose a framework of spatial analysis using a hybrid estimation model-based on a combination of multi-criteria decision-making (MCDM) and artificial neural networks (ANNs) (MCDM-ANNs) classification. The proposed framework is based on the comparison of existing frameworks through the concept of a literature review. The model in the proposed framework will be used for future work on the traffic accident-prone road classification through testing with a private or public spatial dataset. Model validation testing on the proposed framework uses metaheuristic optimization techniques. Policymakers can use the results of the model on the proposed framework for initial planning developing GIS software engineering through spatial analysis models.

Corresponding Author:

Anik Vega Vitianingsih
Department of Informatics
Universitas Dr. Soetomo
Jalan Semolowaru 84 Surabaya, 60118, Surabaya, Indonesia
Email: vega@unitomo.ac.id

1. INTRODUCTION

Model accuracy prediction in the development of frameworks on GIS software is the first step in efforts to improve the quality of GIS software developed and is part of quality control and quality assurance [1]. Quality control will determine the method of spatial analysis to test quality standards [1]. A spatial analysis modeling is a process to build an artificial intelligence (AI) model that is combined with trials on spatial datasets [2], gathering spatial knowledge through spatial datasets and providing knowledge of models in the framework through AI methods from various sources. The purpose of the spatial analysis model is to make a description of the GIS software that will be developed, conduct simulations to test spatial datasets through models on the AI method used on the proposed framework that has already been described. Spatial datasets in GIS relate to how primary and secondary data are obtained through the collection process, and then how the data is processed through spatial analysis to be information in the decision support system [3]. Visualization of spatial data can be done with cloud-terminal integration GIS to provide convenience in the process of spatial analysis on a large number of spatial datasets [4], aggregation-based spatial datasets information retrieval system [5]. Spatial datasets as the key to the value of big data in spatial data mining (SDM) that refers to the description of attribute data requirements, how the data is obtained, and what AI method is used to perform spatial analysis of the data [6], [4]. Spatial datasets become the basic structure in GIS for the process of spatial analysis algorithms, analyzing algorithm principles, or adapting existing algorithms [7]. The classification model in machine learning is prevalent [8] to be used research in the field of spatial analysis of GIS. However, there is no concrete statement regarding which classification algorithm is best to use with certainty because the accuracy, precision, and recall (APR) tests in each study use different sample data. It is also based on the field of study, which is always other on the object of research conducted.

Previous research proposed a framework using the CART model (classification and regression trees), which reported a 10-fold increase in the best value for crash severity prediction [9]. However, the CART model has a weakness in the number of training data samples because changes in training and testing data samples affect the results of spatial analysis [10]. Spatial analysis model using data mining decision tree (J48, ID3, and CART) and naïve bayes classifiers [11] States that the accuracy value of 96.30% on the J48 method is higher than ID3, CART, and naïve bayes, where the naïve bayes have better performance even though the accuracy value is small. Different studies suggest that the accuracy of prediction of classification models with the decision tree approach to reach 84.1% [12]. Also, indicate that the enhanced empirical bayesian (EB) method is a spatial analysis approach that is preferred for prediction of the number of accidents in road segments [13]. Maximizes the accuracy value of the model for Geo-spatial data using the adaptive k-nearest neighbor (kNN) classifier, i.e., by dynamically selecting k for each instance, the value being classified reaches a ROC AUC score of 0,9. The fuzzy deep-learning approach model is used to reduce the uncertainty of data in the prediction of traffic flows that affect road traffic accident rates [14]. Convolutional long short-term memory (ConvLSTM) neural network model [15] states that the proposed framework is sufficiently accurate and significant to improve accuracy in traffic accident prediction for heterogeneous data. The road accident classification model using random forests and boosted trees works equally well with an average value of 80% accuracy and a sensitivity value of 50% [16].

The discussion in this paper emphasizes the comparison in modeling spatial analysis using classification methods for hybrid models through the proposed framework. The general contribution of this proposed framework will be used for future work is integrated through the GIS-platform for the safe management and risk assessment [17], [18] of traffic accident-prone roads classification, to analyze multi-criteria parameters that influence the results on the traffic accident-prone road classification, to purpose new parameters of spatial datasets, to enhance a framework of spatial analysis using a hybrid estimation model-based on a combination of MCDM-ANNs, and to evaluate the enhancement of the new model through the hybrid. Model evaluation needs to be done to provide best practices for the resulting model [19]. Model performance assessment is influenced by balanced data to describe the quality of the resulting model, so as not to lead to misleading conclusions [16]. The proposed framework of classification models with MCDM-ANNs hybrid to the implementation of prone-roads traffic accident classification and its differences with existing frameworks are presented of classification models. The selection of a model-based hybrid estimation on a combination of MCDM-ANNs classification in this proposed framework study is based on a literature review. The collection of dataset multi-criteria parameter for prone-roads traffic accident classification which has been used in the paper articles obtained to evaluate the proposed framework of classification models, explains also the validation and evaluation techniques of the proposed model. Modeling of group analytic hierarchy process (GAHP) technique to develop weighting technique on multi-parameter criteria applied to MCDM Methods which still use are a human assumption in weighting, proving through the sensitivity and stability test of GAHP technique modeling to MCDM methods by comparing the weight was given the human by manual assumption.

Multi-criteria decision making (MCDM) methods are used in this study to process the determinant parameter data in the classification of accident-prone areas that include road conditions, traffic volume, accident rate [20], [21], assign weighting values to each factor based on literature and surveys to expert sources [22]. From the classification of the accident-prone areas, it becomes crucial to provide recommendations to the road auditor to conduct a traffic safety audit to obtain assessment criteria, implementation expenses, the number of involved traffic participants, the effect of road safety, protective effect, and social factors presenting difficulties [23]. The traffic safety audit is carried out by the administration of the road auditor by conducting a feasibility study of the network of accident-prone road categories [24]. MCDM methods have been used for analysis with simple additive weight (SAW), analytical hierarchy process (AHP), and fuzzy AHP method, used for road safety analysis (RSA) that can help decisions process in n determining the priority of road management and provide mitigating actions against the most vulnerable to accidents [25]. The MCDM method with technique for order preference by similarity to ideal solution (TOPSIS) method is used in the management of road safety, and road safety is one of the factors to reduce the number of traffic accidents by knowing the position of a road safety study in Bushehr province Bushehr-Borazjan roads and Borazjan-Genaveh based on various quantitative and qualitative criteria [26]. The MCDM model is one of the right approach models to deal with the problem of accident-prone road section (APRS) because it uses several road and environmental criteria, both quantitative or qualitative; MCDM is related to the results of decision making for planning that involves stakeholders [27]. A framework to be proposed through the process of a literature review from several studies that have been done before. This proses to evaluate the benefits of research that has been done, to know the limitations of the method used, to identify research gaps that have been conducted, and to advise development for further research to get the right framework in the research the new [28]. The research questions in research are intended to focus on the subject area of the study by identifying and classifying the spatial analysis framework for accident-prone traffic roads to be done [29].

2. RESEARCH METHOD

The spatial analysis model using MCDM is a multi-criteria spatial decision support system (MC-SDSS) developed in GIS technology by integrating MCDM as a method to determine the best alternative from the many choices available based on the spatial datasets described [30]. ANNs classification is a data mining technique in machine learning, mapping various attributes as input layer in a node, adding the hidden layer, which is then used to get the threshold to the non-linear output layer [31]. The proposed framework with the steps in Figure 1.

The initial stage a proposed framework in Figure 1 is to plan topics and research trends with identifying in research needs for the literature review process through state-of-the-art frameworks, methods, datasets requirements, and gap analysis of existing methods and frameworks. Action adapting, improving, and hybrid implementation to model accuracy prediction in the development of frameworks. The state-of-the-art from the literature review within the primary study is displayed in Table 1.

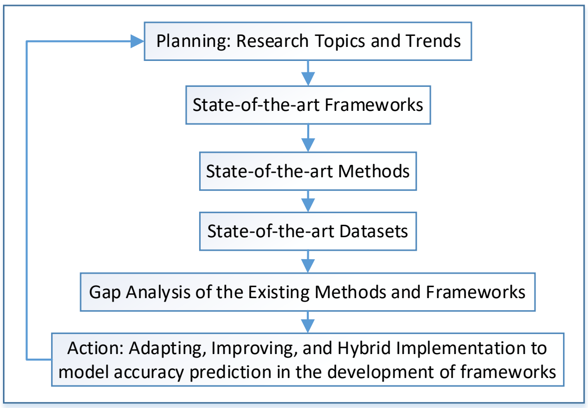

Figure 1. Research method steps

The literature review is in Table 1. The research [32] not shown the comparison of the accuracy and consistency of each method used with the confusion matrix. The meaning of empirical bayes has the best accuracy and consistency value that is not really visible. The standard deviation of the data distribution value in the sample data is only used to calculate the disaster-prone traffic accident rate, and there is no proof of the truth of the model used [33]. Discussion [34] is still limited to the use of an existing method, and knowledge combination has not been done as a hybrid model approach. The results of the comparison of the two methods are stated to be more accurate, but no precise accuracy value is given based on the value of the confusion matrix [35]. On research [36] have not considered the type of road type design, for example, arterial roads, collector roads, or roads based on their nature (geometric road), there are no studies on adaptive models that can expand machine learning through a combination of online learning and deep learning [37]. Paper discussion [38] is still limited to the use of an existing method; knowledge combination has not been done as a hybrid model approach. The DTR model in conducting the prediction accuracy in this study is still a macro-level crash count [39]. The model has weaknesses in terms of data simulation because it requires accident data at the beginning of the calculation [27]. Mathematical modeling in the comparison algorithm does not exist, so the comparison of results is difficult [40]; there is no evaluation of the models offered because the test data collected does not have a long-time span [16]. MLP is more accurate for available spatial datasets but becomes very vulnerable when there is data noise that can cause errors in predictions [34]. PNN has probabilistic outputs with multilayer perceptron networks, producing fairly accurate predictions [34]. RBF is very weak in making predictions [34]. VKT parameters proved to be the most influential in road traffic accidents, then the V/C variable and driver speed based on the RReliefF algorithm calculation method [34]. The evaluation to perform the technique, The site consistency test (SCT), The method consistency test (MCT), The total rank differences test (TRDT), and The total score test (TST) [38].

Table 1. Literature reviews a framework comparison

Framework	Model and method	Spatial datasets	Results
[32]	Model-based spatial statistical methods: Poisson regression, Negative Binomial regression, Empirical Bayesian.	The accidents, injuries, and deaths by years	In this study comparing all methods used, where Empirical Bayes has the best accuracy and consistency, recommended by the Highway Safety Manual (HSM) and the European Union Acquis
[33]	Model-based spatial statistical methods: Kernel density analysis, Nearest neighbor, K-function	Intercity accidents, accidents leading to injury, accidents leading to death, and accidents leading to damages	The observed value curve on the spatial analysis process, the value of spatial datasets is above the 5% confidence interval
[36]	Spatial analysis techniques: Nearest Neighborhood Hierarchical (NNH) Clustering, Spatial-Temporal Clustering Analysis (STAC)	Road accidents involving all types of vehicles	The results of the spatial analysis vary according to the parameter values in the spatial datasets, where is STAC has a 461,57 higher Prediction Accuracy Index (PAI) compared to NNH 163,69.
[34]	ANNs techniques: Extreme learning machine (ELM), Probabilistic neural network (PNN), Radial basis function (RBF), and Multilayer perceptron (MLP).	V/C, speed, vehicle kilometer traveled (VKT), roadway width, the existence of median, and allowable/not-allowable parking	Evaluation method using Nash – Sutcliffe (NS), mean absolute error (MAE), and root means square error (RMSE). ELM, as a feed-forward neural network, becomes the algorithm that has the best performance and the most accurate prediction results (RMSE =3,576; NS =0,81; MAE =2,5062) by randomly selecting hidden nodes using random weights.
[35]	Hot spot analysis (Getis-Ord Gi*): Network spatial weights, Kernel Density method	The traffic accident)	Hotspot analysis gives better results because it is done by considering the weight of spatial datasets
[37]	The support vector machine combines the techniques of statistical learning, machine learning, the neural networks based: Support vector machine, Deep neural network	Accident, person, vehicle, road, and environment data	They proposed a real-time online deep learning framework Based on traffic accident black spots. SVM algorithm in machine learning has 63% precision and a 61% recall rate in analyzing the black spots of traffic accidents. If the training data period is added, the SVM and deep neural network values increase by 95% and 89% accuracy, 69%, and 79% recall rates.
[38]	Black spot identification (BSID) method and Segmentation method: Empirical Bayesian (EB), Excess Empirical Bayesian (EEB), Accident Frequency (AF), Accident Rate (AR).	The traffic accident	AF method has the best performance with a consistency of 93.1% compared to EB 92.2%, and EEB 77.4%. The performance of the EEB and AR methods is the weakest in the case of segmentation in most cases of segmentation.
[39]	Machine learning techniques to prediction model: Decision tree regression (DTR) methods Regression tree framework, Ensemble techniques. Model assessment: Average Squared Error (ASE), Standard Deviation of Error (SDE)	Statewide Traffic Analysis Zone (STAZ)	The DTR model to prediction accuracy works better than the spatial DTR model. To improve prediction accuracy using ensemble techniques (bagging, random forest, and gradient boosting) with slightly better results, depending on the amount of training data.
[27]	Multicriteria decision making (MCDM) model: Weighted linear combination (WLC) method	The traffic accident Reports	The model was developed to determine the criteria weights that have been determined by experts with interest in subjective results.
[40]	Prediction model: Deep neural network model, Gene expression programming (GEP), Random effect negative binomial (RENB) models, Regular negative binomial model (FENB)	The road geometry, traffic, and road environment)	The DNN model experienced an increase in road prediction with 0.914 (RMSE =7.474) by GEP, and 0.891 (RMSE =8.862). GEP works better than RENB to measure the ranking of variables that influence accidents.
[16]	Random effects negative binomial model: Hierarchical cluster method	Real time-frequency of accident data and contributing factors	The model developed can provide information on the main causes of accidents at road intersections

3. RESULTS AND DISCUSSION

The proposed framework is based on a literature study by comparing the existing framework to determine the performance of the spatial analysis model offered for traffic accident prone roads in Table 1 and the MCDM-based framework [27], [41]-[44]. The framework being compared includes the method used for primary study (PS) spatial analysis for accident-prone traffic roads, the spatial analysis model used, spatial datasets used to test the model through method selection, and the value of the measurement results through the assessment.

The framework that has been developed by previous researchers will be described in this section. The framework model [44], was developed to create the Maycock and Hall's accident prediction model. This model provides sensitivity analysis on modeling results using multi-objective optimization (MOO) using multi-criteria decision making for the analytical hierarchy process (AHP model). The needs primary spatial data sets in the road geometry category, the necessity secondary spatial, i.e., the numbers and types of traffic accidents, traffic and demand for structural flow, visual distance and vehicle speed, road signs, and equipment, lighting, driver behavior. The value of the multi-criteria parameters obtained will be done mathematic spatial data modeling to produce the sensitivity of spatial analysis, the results of multi-criteria optimization in the form of traffic efficiency (TS), and traffic safety (TS) to the predicted traffic accident. MOO model is measured using a consistency index (CI) and consistency ratio (CR), the model is proven to have a good structure with a value of CR≤10, or the CR value is 0.00298; this shows that the MOO model with MCDM on the AHP model has a consistent value the good one.

The [41] framework was developed by the PROMETREE-RS MCDM model. MCDM is used because it can use more than one parameter to get the best results from the alternatives produced. This model was developed to evaluate the DEA and TOPSIS methods in road safety to reduce risk the number of accidents on the road through the road safety index. The model is tested by using the Robustness of the composite index. The average correlation value, the average rank value, and the cluster variation average values will be entered into the MCDM PROMETHEE-RS to test the resulting model. Multi-criteria parameters tested in this model, i.e., the Police Department data, fatalities, serious injuries, number of inhabitants, number of registered vehicles, traffic risk, and public risk. This parameter will be used to mathematic spatial data modeling through DEA and TOPSIS, the produce optimal composite index through the value of final risk efficiency. DEA-WR provides the best ranking results compared to the DEA-based composite indicator model (DEA-CI).

The [42], [43] framework is a model built using MCDM. The purpose of this model is to create a knowledge data mining rule decision tree through FP-growth and apache spark framework. A trial model on road accident analysis, where the results have a high degree of accuracy and work well to improve road safety. The multi-parameter criteria used are the road accident data to death and injuries attribute. The testing model for the relevant association rule is done by testing and validation by measuring quality measurement. MCDM model involves many criteria, so it is suitable to overcome the problem of accident-prone road section (APRS) on the type of horizontal alignment, vertical alignment, intersections, significant places, and shoulder widths with an accuracy value of 0.8830 for threshold values 1 [27].

The proposed framework in previous research will be used by the author as a reference in developing further activities of the framework that will be proposed. The framework of the research proposed in Figure 2 has the main differences from the existing framework. Prepare data requirements for spatial datasets as primary and secondary spatial datasets in determining the road categories to be studied (using a private or public spatial dataset type). Perform a literature study relating to multi-criteria parameters used on each road category. Mathematics modeling for spatial analysis to the proposed framework for hybrid estimation model-based on a combination of MCDM-ANNs multi-class classification. In this case, the pre-processing data process will run for the classification analysis process. The range of classification will be performed through mathematical modeling using the Guttman method. The results of the multi-class classification will be validated with SCT, MCT, and ARC validation. Focuses on the propose a classification of roads prone to accidents using multiple criteria parameters (data series), make modeling of road prone to accidents by calculating the value of traffic accident by type of events and the index of the accidents, the density that of roads traffic accident happened to each zone and the amount of data in each year, risk factors based on the severity of the accidents, severity of roads traffic accident events, crash prediction models using data series, and the value of the societal cost of each type the accident. The ANN strategy has the most noteworthy rating of techniques that are regularly utilized in the literature review in essential considers. The empirical Bayes method and decision tree in data mining are also broadly used within the clustering category in spatial information modeling of accident-prone zones. This considers a proposed framework of classification used a hybrid estimation model based on a combination of MCDM-ANN classification. Test the consistency of the method from the model produced with the MCT, SCT, and the value of ARC model evaluations. ANNs classification methods are the most popular data mining techniques in the field spatial analysis of accident-prone roads and the factors that affect the accident rate, among others (neural networks, extreme learning machines, k-nearest neighbor, naive bayes, decision trees) [45], [31].

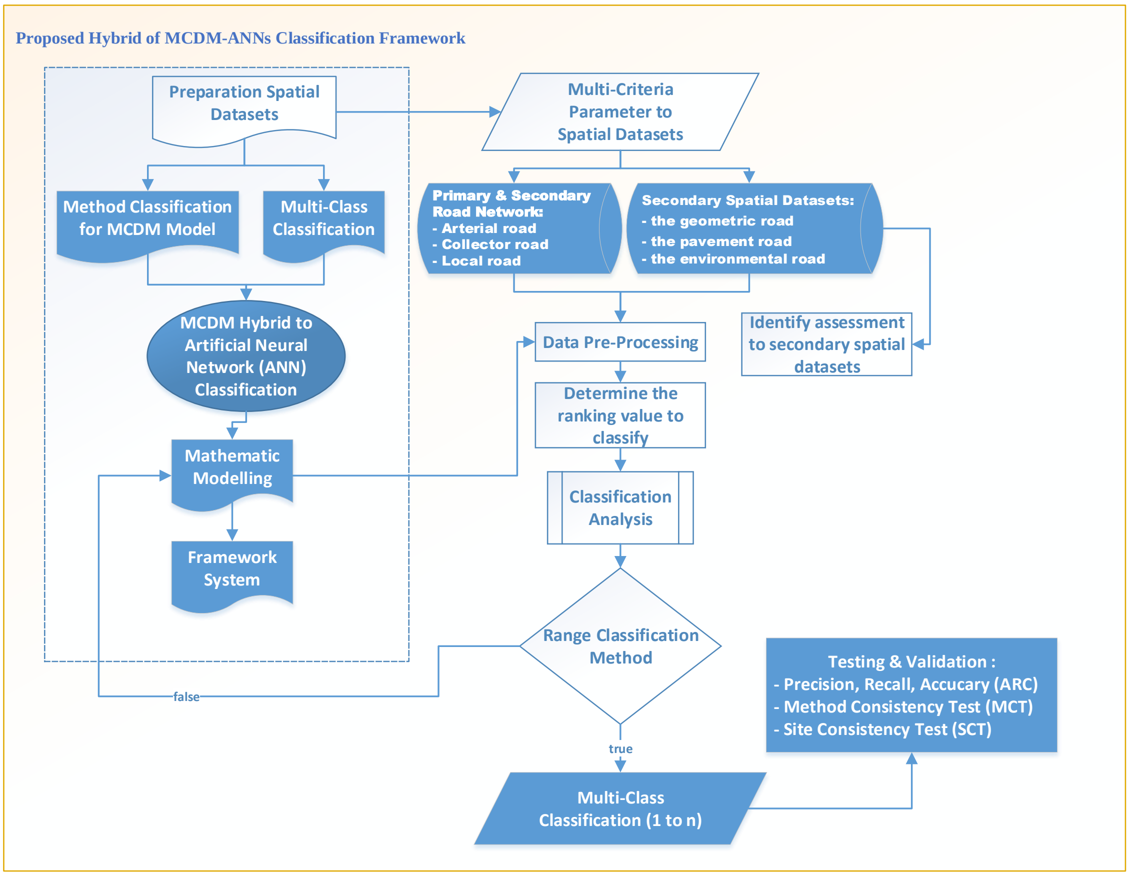

Figure 2. Proposed hybrid of MCDM-ANNs classification framework to evaluate and rank spatial analysis model traffic accident prone roads

4. CONCLUSION

The proposed framework in this study will act as a model-based hybrid estimation approach on a combination of MCDM-ANNs classification to strengthen data mining techniques in spatial multi-criteria analysis in multi-class classification decision making. In the literature review on the primary study, there are no research topics that discuss on the traffic accident-prone roads classification on the arterial road, collector road, and type of road based on its nature (pavement, geometry, and local road) categories. The spatial analysis model using MCDM among others, analytic hierarchy process (AHP), analytical network process (ANP), weighted sum model (WSM), weighted product (WP), weight product model (WPM), simple additive weighting (SAW), technique for order preference by similarity to ideal solution (TOPSIS), preference ranking organization method for enrichment of evaluations (PROMETHEE), multi-attribute utility theory (MAUT), elimination and choice expressing reality (ELECTRE), and vlsekriterijuska optimizacija i komoromisno resenje (VIKOR). The results of the best methods through APR measurement will be a reference in decision making in road management. Existing research is still limited to one type of road used as an object (specific region), and 96% is used private spatial datasets. In this study, it was using an Inductive qualitative approach in the modeling of road prone to accidents to identify the findings of science that is done during the research process. The proposed a classification of roads prone to accidents using multiple criteria parameters, make a modeling of road prone to accidents calculating by the value of traffic accident by type of events and the index of the accidents, the value of the density that of roads traffic accident happened to each zone and the amount of data in each year, the value of risk factors based on the severity of the accidents, the value of severity of roads traffic accident events, the value of crash prediction models, the value of the societal cost of each type the accident, and the test result is using the method the SCT, the MCT, and APR.

ACKNOWLEDGEMENTS

This research is supported by Universiti Teknikal Malaysia Melaka, Malaysia, and Universitas Dr. Soetomo, Indonesia. The development of results of a study funded by The Directorate General of Strengthening Research and Development of Research, Technology, and Higher Education Ministry-Indonesia in 2015-2016.

REFERENCES

[1] J. Albrecht, “GIS Project Management,” in *Comprehensive Geographic Information Systems*, Elsevier Inc., 2018, pp. 446–477.

[2] A. Banerjee and S. Ray, “Spatial models and geographic information systems,” *Encyclopedia of Ecology, 2nd Edition*. Elsevier Inc., pp. 1–10, 2018.

[3] K. E. Brassel and R. Weibel, “A review and conceptual framework of automated map generalization,” *Int. J. Geogr. Inf. Syst.*, vol. 2, no. 3, pp. 229–244, 1988.

[4] S. Wang, Y. Zhong, and E. Wang, “An integrated GIS platform architecture for spatiotemporal big data,” *Futur. Gener. Comput. Syst.*, vol. 94, no. May, pp. 160–172, 2019.

[5] J. Lacasta, F. J. Lopez-Pellicer, B. Espejo-García, J. Nogueras-Iso, and F. J. Zarazaga-Soria, “Aggregation-based information retrieval system for geospatial data catalogs,” *Int. J. Geogr. Inf. Sci.*, vol. 31, no. 8, pp. 1583–1605, 2017.

[6] D. Li, S. Wang, H. Yuan, and D. Li, “Software and applications of spatial data mining,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 6, no. 3, pp. 84–114, 2016.

[7] L. Zhao, L. Chen, R. Ranjan, K. K. R. Choo, and J. He, “Geographical information system parallelization for spatial big data processing: a review,” *Cluster Comput.*, vol. 19, no. 1, pp. 139–152, 2016.

[8] N. F. Hordri, A. Samar, S. S. Yuhaniz, and S. M. Shamsuddin, “A systematic literature review on features of deep learning in big data analytics,” *Int. J. Adv. Soft Comput. its Appl.*, vol. 9, no. 1, pp. 32–49, 2017.

[9] M. Effati and A. Sadeghi-Niaraki, “A semantic-based classification and regression tree approach for modelling complex spatial rules in motor vehicle crashes domain,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 5, no. 4, pp. 181–194, 2015.

[10] M. A. Raihan, M. Hossain, and T. Hasan, “Data mining in road crash analysis: the context of developing countries,” *Int. J. Inj. Contr. Saf. Promot.*, vol. 25, no. 1, pp. 41–52, 2018.

[11] T. K. Bahiru, D. Kumar Singh, and E. A. Tessfaw, “Comparative Study on Data Mining Classification Algorithms for Predicting Road Traffic Accident Severity,” in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2018, pp. 1655–1660.

[12] Z. Zheng, P. Lu, and D. Tolliver, “Decision Tree Approach to Accident Prediction for Highway–Rail Grade Crossings,” *Transp. Res. Rec. J. Transp. Res. Board*, vol. 2545, no. 1, pp. 115–122, 2016.

[13] A. S. Lee, W. H. Lin, G. S. Gill, and W. Cheng, “An enhanced empirical bayesian method for identifying road hotspots and predicting number of crashes,” *Journal of Transportation Safety and Security*, vol. 0, no. 0, Taylor & Francis, pp. 1–17, 2018.

[14] W. Chen *et al.*, “A novel fuzzy deep-learning approach to traffic flow prediction with uncertain spatial–temporal data features,” *Futur. Gener. Comput. Syst.*, vol. 89, no. June, pp. 78–88, 2018.

[15] M. Kibanov, M. Becker, M. Atzmueller, and A. Hotho, “Adaptive kNN Using Expected Accuracy for Classification of Geo-Spatial Data,” in *Proceedings of the 33rd Annual ACM Symposium on Applied Computing Pages 857-865*, 2018, pp. 857–865.

[16] M. Schlögl, “A multivariate analysis of environmental effects on road accident occurrence using a balanced bagging approach,” *Accid. Anal. Prev.*, vol. 136, no. March, pp. 1–12, 2020.

[17] W. Li and S. Wang, “PolarGlobe: A web-wide virtual globe system for visualizing multidimensional, time-varying, big climate data,” *Int. J. Geogr. Inf. Sci.*, vol. 31, no. 8, pp. 1562–1582, 2017.

[18] M. P. Repetto, M. Burlando, G. Solari, P. De Gaetano, M. Pizzo, and M. Tizzi, “A web-based GIS platform for the safe management and risk assessment of complex structural and infrastructural systems exposed to wind,” *Adv. Eng. Softw.*, vol. 117, pp. 29–45, 2018.

[19] H. Gong, F. Wang, B. Brenda, and S. Dent, “Application of random effects negative binomial model with clustered dataset for vehicle crash frequency analysis,” *Int. J. Transp. Sci. Technol.*, no. April, pp. 1–12, 2020.

[20] T. Sipos, “Spatial statistical analysis of the traffic accidents,” *Period. Polytech. Transp. Eng.*, vol. 45, no. 2, pp. 101–105, 2017.

[21] A. V. Vitianingsih and D. Cahyono, “Geographical Information System for Mapping Road Using Multi-Attribute Utility Method,” in *International Conference on Science and Technology-Computer (ICST)*, 2016, pp. 0–4.

[22] R. Al-Ruzouq, K. Hamad, S. Abu Dabous, W. Zeiada, M. A. Khalil, and T. Voigt, “Weighted Multi-attribute Framework to Identify Freeway Incident Hot Spots in a Spatiotemporal Context,” *Arab. J. Sci. Eng.*, vol. 44, no. 10, pp. 8205–8223, 2019.

[23] Á. Török, “Statistical Analysis of a Multi-Criteria Assessment of Intelligent Traffic Systems for the Improvement of Road Safety,” *J. Financ. Econ.*, vol. 4, no. 5, pp. 127–135, 2016.

[24] Y. Huvarinen, E. Svatkova, E. Oleshchenko, and S. Pushchina, “Road Safety Audit,” in *Transportation Research Procedia*, 2017, vol. 20, no. September 2016, pp. 236–241.

[25] S. Kanuganti, R. Agarwala, B. Dutta, P. N. Bhanegaonkar, A. P. Singh, and A. K. Sarkar, “Road safety analysis using multi criteria approach: A case study in India,” in *Transportation Research Procedia*, 2017, vol. 25, pp.

4653–5665.
[26] M. S. Fatemeh Haghighat, "Application of a Multi-Criteria Approach To Road Safety Evaluation in the Bushehr Province , Iran," *Traffic Plan. Prelim. Commun.*, vol. 23, no. 5, pp. 341–352, 2011.
[27] F. Yakar, "A multicriteria decision making–based methodology to identify accident-prone road sections," *J. Transp. Saf. Secur.*, pp. 1–15, 2019.
[28] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering," *Engineering*, vol. 2, p. 1051, 2007.
[29] M. Vierhauser, R. Rabiser, and P. Grünbacher, "Requirements monitoring frameworks: A systematic review," *Inf. Softw. Technol.*, vol. 80, pp. 89–109, 2016.
[30] S. M. Ghavami, "Multi-criteria spatial decision support system for identifying strategic roads in disaster situations," *Int. J. Crit. Infrastruct. Prot.*, vol. 24, pp. 23–36, 2019.
[31] N. Sang and M. Aitkenhead, "Data Mining, Machine Learning and Spatial Data Infrastructures for Scenario Modelling," in *Modelling Nature-based Solutions*, 2020, pp. 276–304.
[32] M. A. Dereli and S. Erdogan, "A new model for determining the traffic accident black spots using GIS-aided spatial statistical methods," *Transp. Res. Part A Policy Pract.*, vol. 103, no. September, pp. 106–117, 2017.
[33] G. A. Shafabakhsh, A. Famili, and M. S. Bahadori, "GIS-based spatial analysis of urban traffic accidents: Case study in Mashhad, Iran," *J. Traffic Transp. Eng.*, vol. 4, no. 3, pp. 290–299, 2017.
[34] H. Behbahani and A. Mohamadian, "Forecasting accident frequency of an urban road network : A comparison of four artificial neural network techniques," *J. Forecast.*, vol. 37, no. 7, pp. 767–780, 2018.
[35] H. E. Colak, T. Memisoglu, Y. S. Erbas, and S. Bediroglu, "Hot spot analysis based on network spatial weights to determine spatial statistics of traffic accidents in Rize, Turkey," *Arab. J. Geosci.*, vol. 11, no. 151, pp. 1–11, 2018.
[36] S. S. R. Shariff, H. A. Maad, N. N. A. Halim, and Z. Derasit, "Determining hotspots of road accidents using spatial analysis," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 9, no. 1, pp. 146–151, 2018.
[37] Z. Fan, C. Liu, D. Cai, and S. Yue, "Research on black spot identification of safety in urban traffic accidents based on machine learning method," *Saf. Sci.*, vol. 118, no. April, pp. 607–616, 2019.
[38] M. Ghadi and Á. Török, "A comparative analysis of black spot identification methods and road accident segmentation methods," *Accid. Anal. Prev.*, vol. 128, no. February, pp. 1–7, 2019.
[39] M. S. Rahman, M. Abdel-Aty, S. Hasan, and Q. Cai, "Applying machine learning approaches to analyze the vulnerable road-users' crashes at statewide traffic analysis zones," *J. Safety Res.*, vol. 70, no. September, pp. 275–288, 2019.
[40] G. Singh, M. Pal, Y. Yadav, and T. Singla, "Deep neural network-based predictive modeling of road accidents," *Neural Comput. Appl.*, vol. 32, no. 16, pp. 12417–12426, 2020.
[41] M. Rosić, D. Pešić, D. Kukić, B. Antić, and M. Božović, "Method for selection of optimal road safety composite index with examples from DEA and TOPSIS method," *Accid. Anal. Prev.*, vol. 98, no. January, pp. 277–286, 2017.
[42] A. Ait-Mlouk, F. Gharnati, and T. Agouti, "An improved approach for association rule mining using a multi-criteria decision support system: a case study in road safety," *Eur. Transp. Res. Rev.*, vol. 9:40, no. September, pp. 1–13, 2017.
[43] A. Ait-Mlouk, T. Agouti, and F. Gharnati, "Mining and prioritization of association rules for big data: multi-criteria decision analysis approach," *J. Big Data*, vol. 4, no. 1, pp. 1–21, 2017.
[44] H. Pilko, S. Mandžuka, and D. Barić, "Urban single-lane roundabouts: A new analytical approach using multi-criteria and simultaneous multi-objective optimization of geometry design, efficiency and safety," *Transp. Res. Part C Emerg. Technol.*, vol. 80, no. July, pp. 257–271, 2017.
[45] V. Rovšek, M. Batista, and B. Bogunović, "Identifying the key risk factors of traffic accident injury severity on Slovenian roads using a non-parametric classification tree," *Transport*, vol. 32, no. 3, pp. 272–281, 2017.

BIOGRAPHIES OF AUTHORS

Anik Vega Vitianingsih. A bachelor's degree in Informatics Engineering in 2004 and a Master's Degree in Game Tech was obtained in 2011. The author is a Permanent Lecturer in the Informatics Department, editor in chief of the International Journal of Artificial Intelligence and Robotics Universitas Dr. Soetomo, and students in the Ph.D. at the Faculty of Information and Communication Technology (FTMK), Universiti Teknikal Malaysia Melaka, Malaysia. The field of interest in Spatial Analysis, and Spatial Data Modeling, Artificial Intelligence in Geographical Information Systems. Experiences in writing papers according to their fields in the Scopus Journal include 2019-International Journal of Intelligent Engineering and Systems, 2019-Data in Brief, 2018-International Journal of Engineering and Technology (UAE), 2018-Journal of Telecommunication, Electronic and Computer Engineering. The author has been a reviewer for Taylor and Francis Ltd publishers, including the Journal of Transportation Safety and Security, IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India), and International Journal of Injury Control and Safety Promotion, as well as a reviewer on the International Social Science Journal Publisher Wiley-Blackwell Publishing Ltd.

Prof. Dr. Nanna Suryana Herman. Professor at the Faculty of Information and Communication Technology (FTMK), Universiti Teknikal Malaysia Melaka, Malaysia. Bachelor's degree in Soil and Water Engineering from Padjadjaran University, Bandung, Indonesia. Master's degree in Computer Assisted Regional Planning at the International Institute for Geoinformatics and Earth Observation (ITC), Enschede, The Netherlands. Doctoral Degree in Department of Remote sensing and GIS, Research University of Wageningen, Holland. His research interests are spatial data analytics, image processing, and spatial modeling, and remote sensing. Active as the Editorial Board of International Journals, member of The ASEAN European Academic University Network (ASEA-UNINET), and EURAS-Eurasian Universities Union.

Zahriah Othman. Lecturer at the Software Engineering department, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka. Bachelor's degree in Information Technology from Universiti Utara Malaysia in 2001. In 2003 he received a master of science degree in Software Engineering from the School of Informatics, Department of Computing, and University of Bradford, United Kingdom. Ph.D. in Computer Science from Universiti Teknikal Malaysia Melaka, Malaysia. His areas of research interest are Software Engineering, Artificial Intelligence, and information retrieval, specifically on terminology disagreement in retrieving geospatial data.