Advances of vehicular ad hoc network using machine learning approach

See Thian Meng[1], Sumendra Yogarayan[1], Siti Fatimah Abdul Razak[1], Subarmaniam Kannan[1], Afizan Azman[2]
[1]Faculty of Information Science and Technology, Multimedia University, Melaka, Malaysia
[2]School of Computing, Faculty of Information and Technology, Taylors University, Subang Jaya, Malaysia

Article Info

Article history:

Received Jun 21, 2023
Revised Jul 22, 2023
Accepted Aug 7, 2023

Keywords:

Algorithm
Intelligent transportation systems
Machine learning
Technique
Vehicular ad hoc network

ABSTRACT

Vehicular ad hoc networks (VANETs) play a crucial role in intelligent transportation systems (ITS), enabling seamless communication between vehicles and other entities. VANETs provide a wide range of services, allowing vehicles to communicate with each other and with roadside infrastructure. With the increasing amount of data generated by VANETs, machine learning approaches have emerged as valuable tools to address complex challenges in this domain. This paper presents a comprehensive literature review on the application of machine learning in VANETs. The paper discusses the potential challenges and future research directions in the field, emphasizing the need for more accessible machine learning solutions for VANETs. This review emphasizes the significant role of machine learning approach in advancing the capabilities of VANETs and shaping the future of intelligent transportation systems.

Corresponding Author:

Sumendra Yogarayan
Faculty of Information Science and Technology, Multimedia University
Ayer Keroh Lama Street, 75450 Melaka, Malaysia
Email: sumendra@mmu.edu.my

1. INTRODUCTION

Vehicular ad hoc networks (VANETs) serve as a fundamental component of intelligent transportation systems (ITS), enabling seamless communication between vehicles and other entities [1], [2]. When compared to traditional ITS, the VANETs provides a wider range of services. This is because VANET can communicate not just with one another, but also with base stations located along the road. Network-based VANET’s communication capabilities between vehicle to vehicle (V2V), roadside infrastructure and other vehicles (V2I) are its main advantages. Utilising the communication platforms and protocols offered by the supporting infrastructure makes this possible [3]–[5]. VANETs have the potential to enhance safety, comfort, and overall road efficiency. VANETs consist of interconnected vehicles that can form networks even without prior knowledge of each other [6]. The contemporary systems for traffic control and road safety make extensive use of these networks. Through the exchange of basic safety messages (BSMs), for instance, crucial information like the present location, speed, and acceleration can be transmitted between automobiles, improving eyesight and awareness of the situation for the near surroundings [7], [8].

Furthermore, VANETs enable vehicles to communicate safety and traffic-related information with each other while on the road. VANETs are distinguished by numerous nodes (vehicles) and constant movement. The security of VANETs is essential for establishing a reliable ITS that ensures safety. According to World Health Organisation (WHO), road accidents are ranked as the eighth major contributor to global fatalities [9], [10]. Next, VANETs are understood to be essential tools for preventing traffic accidents and fatalities. Vehicles can transmit and receive data packets from roadside units (RSUs) within these networks in addition to exchanging data packets among themselves. Because VANETs generate and communicate such a significant amount of data, applying machine learning techniques is advantageous because they may be applied to a variety of goals [11]–[13]. Artificial intelligence (AI) refers to the ability of a computer or robot to perform tasks that are typically associated with intelligent beings, guided by computer programming [14]. The primary aim often involves creating systems that possess intellectual capabilities like those demonstrated by humans. These capabilities encompass reasoning, comprehension, generalization, and the ability to derive meaning or learn from prior experiences [15]–[17]. Many facets of society could be transforms by AI, from bettering healthcare outcomes to enabling more effective transportation systems such as VANET. Besides, different methods, such as rule-based systems, machine learning, deep learning, and evolutionary computation, can be used to create them [18], [19].

Recently, machine learning has been used in wireless networks to offer a data-driven method to solving formerly difficult challenges [20]. This has led to the development of some solutions to enhance the performance of autonomous vehicles (AV) and VANET. Machine learning can be used for VANET testing and training by applying machine learning algorithms. then determining which machine learning algorithm is best for VANET [21], [22]. During the training stage of machine learning, a model is developed by learning from the provided training data. This involves extracting patterns and relationships from the data to create a predictive model. In the subsequent testing stage, the trained model is utilized to make predictions or generate outputs based on new, unseen data. This two-step process allows us to harness the power of machine learning to address complex problems that were traditionally difficult to solve [23]–[27].

2. LITERATURE REVIEW

Machine learning techniques can be categorized into three main divisions: reinforcement learning, unsupervised learning, and supervised learning [28], [29]. Among these, supervised learning is commonly used in practical applications and relies on labelled datasets. Traditional methods in this category include K-nearest neighbours (KNN), decision trees (DT), support vector machines (SVM), neural networks (NN), and random forests (RF) [30]–[33]. Unsupervised learning involves analysing data without prior knowledge or labels, focusing on either the specific characteristics of individual samples or the similarities between samples to group them together. Classic clustering techniques, such as k-means, hierarchical clustering, spectrum clustering, and dirichlet process, are commonly used in this context [34], [35]. Reinforcement learning aims to maximize the rewards obtained from interacting with an environment. It relies on the markov decision process (MDP) model, which revolves around taking actions and receiving corresponding rewards.

Rashid *et al.* [36] presented the issue of detecting malicious nodes in VANET is addressed in this study, where a machine learning-based solution is developed for real-time identification of dangerous nodes. Furthermore, a distributed multi-layer classifier is proposed, and its performance is evaluated using objective modular network testbed in C++ (OMNeT++) and simulation of urban mobility (SUMO), along with machine learning models such as gradient boosting trees (GBT), logistic regression (LR), multilayer perceptron classifier (MLPC), RF, and SVM. The proposed model is designed to be applied to a dataset containing information about attacking and regular vehicles. To compare the performance of different machine learning models, GBT, LR, MLPC, RF, and SVM are employed. RF is particularly recommended for enhanced detection of misbehaviours. The results accurately forecast how effectively the suggested architecture manages the research within the VANET environment.

Kezia and Anusuya [37] outlined to predict the local density of vehicles and the channel busy ratio, with the exclusion of beacon considerations. To achieve this, various machine learning techniques such as LR, KNN, Naive Bayes (NB), DT, and RF algorithms are utilized. The most accurate predictions are then utilized in a congestion control system based on prediction and adaptation. The results reveal that the RF classification technique achieves a remarkable 99.3% accuracy in predicting the number of cars. Additionally, the DT method effectively determines the channel busy ratio with a root mean square error (RMSE) of 0.018 for the predicted number of cars. By incorporating these predictions into the system, the adaptation of transmission parameters is improved, leveraging the calculations of local density and channel busy ratio (CBR). Consequently, this enhancement significantly enhances the overall network performance.

Khan *et al.* [38] highlighted that enabling vehicles to make autonomous decisions is a fundamental requirement in VANET systems. Vehicles select the best path using cognitive memory, which saves all the past routes. Communication is key in networks. With cellular-based vehicle-to-everything (CV2X) communication, each vehicle broadcasts a cooperative awareness message (CAM) that is used to exchange critical information. Finding that results from the proposed SVM-CV2X-M4 system are precise. The vehicles can therefore choose their best path on their own. The proposed system is effective and simple to use thanks to a variety of strategies and algorithms. The planned SVM-CV2X-M4 system's goals are to assist users and make commuting easier.

Ajay and Kumaraswamy [39] demonstrated that advancements in vehicle safety measures have contributed to a reduction in traffic accidents. During congested periods, predicting traffic patterns becomes crucial in selecting optimal routes, resulting in time and fuel savings. This article focuses on traffic forecasting and utilizes various machine learning techniques to gather and analyse data. By employing a specialized machine learning model, the collected dataset significantly enhances the accuracy of traffic prediction compared to previous methods. The RF algorithm achieves an impressive 97.82 % of accuracy rate in predicting traffic data [34].

Alhaidari and Alrehan [40] introduced a state-of-the-art simulation technique to generate a reliable distributed denial of service dataset specifically tailored for automotive ad hoc networks. The architecture, traffic volume, attack power, and node mobility of the VANET are all considered when analysing the distributed denial of service attack traffic in the VANET dataset. Popular simulation programme including SUMO, OMNeT++, Veins, and internet networking (INET). Finding that apart from SVM, which displays an accuracy of 97%, the classifiers reported by VANET distributed denial of service dataset (VDDD) obtained excellent detection accuracies that were typically better than 99%. When compared to other applicable machine learning classifiers, the RF classifier has the greatest accuracy (99.7%).

Alsarhan *et al.* [41] incorporated SVM algorithm to detect intrusions in VANET. Numerous computational benefits of the SVM structure include direction at a finite sample and independence of algorithm complexity to sample dimension. Finding that based on the network security laboratory–knowledge discovery in databases (NSL-KDD) data, this article simulates the interactions between the driver and the VANET, as well as the communications between the vehicles and RSUs of the VANET. The KDD99 version of NSL-KDD data has been cleaned up and refined. It fixes a few of the KDD99's fundamental issues. The entire KDD dataset is contained in NSL-KDD. Internet traffic that was handled by an actual intrusion detection system is contained in the records.

Gad *et al.* [42] highlighted the extensive utilization of the NSL-KDD and KDD-CUP99 datasets in various studies. Nevertheless, these datasets suffer from a limitation of not encompassing up-to-date attack information. In our research, we utilized a realistic dataset called telemetry data, operating data, network data from internet of things (ToN-IoT), which was obtained from a large-scale, diverse internet of things (IoT) network. We experimented with various machine learning techniques to address both binary and multi-class classification challenges. The findings demonstrate that employing the Chi2 feature selection method and synthetic minority over-sampling technique (SMOTE) leads to meaningful results when using KNN algorithm. Furthermore, the outcomes for all applied machine learning techniques using the Chi2 feature selection and SMOTE techniques were evaluated.

Kadam and Sekhar [43] proposed a new secure framework for identifying distributed denial of service (DDoS) using a novel hybrid KSVM scheme, which combines KNN and SVM algorithms. attacks is established The research considers multiple machine learning approaches and evaluates the proposed hybrid kernel support vector machine (KSVM) algorithm in terms of accuracy, sensitivity, precision, recall, and error for detecting malicious DDoS activity. The results demonstrate that the hybrid KSVM algorithm outperforms other machine learning algorithms, making it a superior choice for DDoS attack detection.

Sonker and Gupta [44] focused on detecting five main types of attacks in VANETs: constant attack, constant offset attack, random attack, random offset attack, and eventual attack. Various machine learning techniques are employed to identify these attacks. Initially, binary classification methods are utilized to detect each individual attack. Subsequently, a novel approach is developed to perform multi-classification of the attacks, using the best-performing machine learning algorithm identified from the previous step. The results show that NB, DT, and RF achieve 100% accuracy in binary classification for detecting all attacks. In terms of multi-classification, the RF technique achieves the highest accuracy of 97.62%. The vehicular reference misbehaviour (VeReMi) dataset, which is a publicly available repository for detecting malicious nodes in VANETs, is utilized in this research.

Wang *et al.* [45] addressed the escalating issue of severe traffic congestion resulting from the growing number of vehicles, leading to increased travel time and financial burden on transportation users. To some extent, forecasting future traffic patterns can help to solve this issue. In this study, it employs three machine learning algorithms, namely LR, DT, and SVM, to simulate traffic flow. Prior to conducting the simulation, it pre-processes the data using Selenium, open-source software (OSS), and a message queue. next, it examines Beijing's actual traffic information as displayed on the Baidu map. This study demonstrates that RF has the highest accuracy of the three approaches, at 0.719. The second is SVM and LR.

Anand and Sankhe [46] discussed the various applications of traffic flow forecasting in effectively managing regional and urban traffic. Traffic control in large cities is quite challenging. However, it was discovered that considering certain physical factors like the surroundings and weather made the forecasts more accurate. In this work, a traffic flow forecasting system model is created to anticipate traffic data at intervals of between one hour and twenty-four hours. In the past, research using predictive algorithms was also carried out, but it was discovered that not many systems that make it simple to forecast traffic flow and allow access to general consumers. The system is made to address issues with time series and history. Datasets of previous traffic were gathered from public sources and cleansed upon request. A system that collects data from the road using vehicle detecting sensors and storing it in a database for future predictions was developed using machine learning algorithms. To retrieve weather information, it additionally incorporated the weather service's application programming interface (API). This traffic flow forecasting model was developed using artificial neural networks (ANN) and SVM, two of the most popular machine learning forecasting approaches currently in use.

Le and Maple [47] describe three novel elements, discovered through research on *n* consecutive locations, that play a crucial role in determining a vehicle's behavior. Two machine learning algorithms which is KNN and SVM. Those machine learning algorithm use these attributes to detect assaults in the VeReMi dataset. It demonstrates that the overall precision rates may reach. Besides, this paper has four different algorithms were examined, and it was discovered that KNN and RF classifiers produce the best outcomes. Then, using the same VeReMi dataset, we compared the outcomes of our strategy (which makes use of KNN) with those of other current strategies that are listed in the literature. According to the findings, the suggested strategy significantly surpasses the currently used methods for overall attack detection in terms of both precision and recall.

Mythili and Magendran [48] discussed the limitations associated with unscrupulous individuals who attempt to manipulate information for their own purposes, emphasizing that such information exchange is not always reliable or accurate. However, how to identify numerous attacks in the VANET and how to increase security while maintaining the level of service quality in the VANET becomes a very challenging problem. Finding that the outcome of the experiment demonstrated how effective the suggested algorithm was in detecting malicious nodes. The accuracy of the SVM algorithm is 97.46%, and the accuracy of the enhanced SVM algorithm is 98.01%. However, for multiple attack detection in VANET, powerful optimisation-based clustering method and ensemble SVM classifier beat previous techniques by offering 98.99% accuracy.

Shi and Wu [49] proposed a machine learning algorithm-based model for detecting road functionality based on data aggregation in VANETs. Numerous simulations are used to assess the scheme's performance. The simulation findings demonstrate that the method of combining the data from moving cars may precisely identify the state of a dysfunctional road and pinpoint a specific malfunctioning place, especially for large dimensions. Finding that to identify the aberrant traffic state, a road functionality detection model based on VANET, and machine learning has been built. In the simulation, SVM and deep neural network (DNN) are compared to determine the dysfunctional road state, and DNN is then used to estimate where the defect will be exactly. The evaluation's results indicate that it is possible to accurately determine the present traffic situation by merging various data about the moving items on the road.

So *et al.* [50] introduced a system framework that employs plausibility checks as a feature vector for machine learning models in order to detect and classify misbehavior. The results indicate that by incorporating KNN and SVM, the overall precision of detecting misbehaviour using plausibility checks in the feature vectors can be increased by more than 20%, while maintaining a recall rate of less than 5%. The research findings demonstrate that our proof-of-concept misbehaviour detection system, which combines machine learning and plausibility checks, achieves a 20% improvement in precision while keeping the recall within a 5% range of the recall achieved by plausibility checks alone. This framework provides a system that not only identifies inappropriate behaviours but also categorizes them, enabling more targeted corrective actions. Table 1 (see Appendix) shows the existing works of machine learning approach on VANET deployment.

3. RESEARCH GAPS AND DISCUSSION

3.1. Datasets

One of the crucial research gap exists in the realm of VANET research due to the scarcity of standardized and comprehensive benchmark datasets. While studies utilize various datasets such as VeReMi, NSL-KDD, and ToN-IoT, there is a need for standardized datasets that encompass a wide range of VANET scenarios and capture real-world traffic patterns, mobility, and attacks. The availability of high-quality datasets would enable researchers to compare and validate their proposed solutions effectively. Collaboration among researchers, industry partners, and governmental agencies is crucial for establishing standardized benchmark datasets for VANET research. The datasets should capture various real-world scenarios, including different traffic patterns, mobility patterns, and attack scenarios. Creating a repository for sharing these benchmark datasets would facilitate progress in the field and ensure fair and accurate comparisons among different research endeavors.

3.2. Performance

A notable research gap in VANET applications lies in the absence of comprehensive comparative evaluations of machine learning algorithms. Although studies employ different algorithms such as GBT, LR, MLPC, RF, and SVM, there is a need for rigorous benchmarking to determine their accuracy, efficiency, and scalability. Comparative evaluations would provide valuable insights into the strengths and weaknesses of different approaches and help identify the most suitable algorithms for specific VANET scenarios. To address this research gap, it is important to emphasize the need for standardized datasets, evaluation criteria, and factors such as real-time performance in order to ensure fair and meaningful comparisons. By conducting comprehensive benchmarking studies, researchers can contribute to advancing the field of VANETs and facilitate the development of more effective and efficient machine learning solutions.

3.2. Security

Another research gap pertains to the limited consideration of diverse attack scenarios in VANETs. While some studies focus on detecting specific types of attacks, there is a need for a comprehensive exploration of various attack scenarios to develop robust security mechanisms. Understanding and addressing different attack types, including DDoS, misbehaving nodes, and intrusion detection, are crucial for ensuring the security and integrity of VANETs. Developing a taxonomy specific to VANETs and systematically investigating each attack category can provide valuable insights for the development of effective security measures. Furthermore, collaboration with industry partners and governmental agencies is essential to gather real-world attack data, enabling more accurate analysis and testing of security mechanisms. By addressing this research gap, VANETs can be better safeguarded against a wide range of attacks, ensuring their integrity and reliability.

3.4. Scalability and resource constraints

The scalability and resource constraints of VANETs present a significant research gap. As the number of vehicles in VANETs increases, issues such as network congestion, energy consumption, and limited bandwidth become critical factors that need to be addressed. Existing research often lacks comprehensive solutions that can handle the scaling challenges of vehicular networks while optimizing resource utilization and maintaining system performance and security. Thus, approaches such as distributed computing, edge computing, or cloud-assisted solutions should be incorporated to handle the scalability challenges of vehicular networks. These approaches can leverage distributed resources, optimize energy consumption, and enhance overall system performance. Additionally, energy-efficient algorithms could be employed to optimize resource utilization without compromising system performance or security.

4. CONCLUSION

In conclusion, machine learning techniques have emerged as valuable tools in addressing the challenges of VANETs. These techniques have been applied to improve security, predict traffic patterns, enhance network performance, and detect malicious nodes in VANETs. The literature review highlights the effectiveness of algorithms such as SVM, RF, DT, and ensemble classifiers in achieving accurate results for different VANET applications. As VANETs continue to play a crucial role in ITS, integrating machine learning techniques will further enhance vehicular networks safety, efficiency, and overall performance. Future research can explore machine learning algorithms, hybrid models, and optimization techniques to tackle complex challenges in VANETs and improve their functionality. By harnessing the capability of machine learning, VANETs have the potential to revolutionize the transportation industry and pave the way for safer, smarter, and more efficient road networks.

APPENDIX

Table 1. Existing work of machine learning approach on VANET deployment

Study	Features	Algorithm	Training set	Accuracy
Rashid *et al.* [36]	DDoS	SVM	OMNeT++, SUMO++ and Veins	97.00%
		RF		98.00%
Kezia and Anusuya [37]	Traffic	KNN	CBR	75.18%
		RF		99.30%
		DT		86.22%
Khan *et al.* [38]	Traffic	SVM	SVM-CV2X-M4 system	99.60%
Ajay and Kumaraswamy [39]	Traffic	RF	Traffic condition	82.33%
		SVM		97.82%
				90.67%
Alhaidari and Alrehan [40]	DDoS	SVM	VDDD	97.36%
		RF		99.75%
		KNN		99.73%

Table 1. Existing work of machine learning approach on VANET deployment *(continue)*

Study	Features	Algorithm	Training set	Accuracy
Alsarhan *et al.* [41]	Misbehavior detection	GA-SVM PSO-SVM ACO-SVM	NSL-KDD	98.00% 94.00% 89.00%
Gad *et al.* [42]	Intrusion detection system	KNN SVM RF	ToN-IoT (binary classification)	86.80% 98.80% 97.90%
Kadam and Sekhar [43]	DDoS	KNN SVM	Hybrid KSVM	86.00% 92.00%
Sonker Gupta [44]	Misbehavior detection	KNN RF	VeReMi (type 1 attack)	99.10% 100%
Wang *et al.* [45]	Traffic	SVM RF	Traffic condition	71.90% 65.70%
Anand and Sankhe [46]	Traffic	ANN SVM	Performance measure system (PEMS)	99.00% 99.00%
Le and Maple [47]	Misbehavior detection	KNN SVM	VeReMi	-
Mythili and Magendran [48]	Attack detection	SVM	-	97.46%
Shi and Wu [49]	Traffic	SVM DNN	Planung transport Verkehr-verkehr in stadten simulations model (PTV-VISSIM)	84.00% 87.10%
So *et al.* [50]	Misbehavior detection	KNN SVM	-	-

REFERENCES

[1] F. Gonçalves, J. Macedo, and A. Santos, "Evaluation of VANET datasets in context of an intrusion detection system," in *29th International Conference on Software, Telecommunications and Computer Networks, SoftCOM*, Sep. 2021, pp. 1–6, doi: 10.23919/softcom52868.2021.9559058.

[2] Z. Xia *et al.*, "A comprehensive survey of the key technologies and challenges surrounding vehicular ad hoc networks," *ACM Transactions on Intelligent Systems and Technology*, vol. 12, no. 4, pp. 1–30, Jun. 2021, doi: 10.1145/3451984.

[3] A. A. Hussein and D. A. Mahmood, "Connectivity analysis in vehicular ad-hoc network based on VDTN," *Journal of Communications Software and Systems*, vol. 19, no. 2, pp. 147–157, 2023, doi: 10.24138/jcomss-2022-0166.

[4] S. Yogarayan, S. F. A. Razak, A. Azman, M. F. A. Abdullah, S. Z. Ibrahim, and K. J. Raman, "A review of routing protocols for vehicular ad-hoc networks (VANETs)," in *8th International Conference on Information and Communication Technology, ICoICT*, Jun. 2020, pp. 1–7, doi: 10.1109/ICoICT49345.2020.9166174.

[5] P. Ajmani, N. Singh, and P. Verma, "Internet of vehicles taxonomy and evaluation: architectures, protocols, and issues," in *5th International Conference on Contemporary Computing and Informatics, IC3I*, Dec. 2022, pp. 1163–1170, doi: 10.1109/IC3I56241.2022.10072434.

[6] H. Shahwani, S. A. Shah, M. Ashraf, M. Akram, J. (Paul) Jeong, and J. Shin, "A comprehensive survey on data dissemination in vehicular ad hoc networks," *Vehicular Communications*, vol. 34, p. 100420, Apr. 2022, doi: 10.1016/j.vehcom.2021.100420.

[7] L. Liang, H. Ye, and G. Y. Li, "Toward intelligent vehicular networks: a machine learning framework," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 124–135, Feb. 2019, doi: 10.1109/JIOT.2018.2872122.

[8] N. H. Hussein, C. T. Yaw, S. P. Koh, S. K. Tiong, and K. H. Chong, "A comprehensive survey on vehicular networking: communications, applications, challenges, and upcoming research directions," *IEEE Access*, vol. 10, pp. 86127–86180, 2022, doi: 10.1109/ACCESS.2022.3198656.

[9] S. Sharma and B. Kaushik, "A survey on internet of vehicles: applications, security issues & solutions," *Vehicular Communications*, vol. 20, p. 100182, Dec. 2019, doi: 10.1016/j.vehcom.2019.100182.

[10] K. Kiela *et al.*, "Review of V2X-IoT standards and frameworks for ITS applications," *Applied Sciences (Switzerland)*, vol. 10, no. 12, p. 4314, Jun. 2020, doi: 10.3390/app10124314.

[11] L. Sleem, H. N. Noura, and R. Couturier, "Towards a secure ITS: overview, challenges and solutions," *Journal of Information Security and Applications*, vol. 55, p. 102637, Dec. 2020, doi: 10.1016/j.jisa.2020.102637.

[12] M. Sangare, S. Banerjee, P. Muhlethaler, and S. Bouzefrane, "Predicting transmission success with support vector machine in VANETs," in *IFIP/IEEE International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks, PEMWN*, Sep. 2018, pp. 1–8, doi: 10.23919/PEMWN.2018.8548826.

[13] H. Gao, C. Liu, Y. Li, and X. Yang, "V2VR: reliable hybrid-network-oriented V2V data transmission and routing considering RSUs and connectivity probability," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3533–3546, Jun. 2021, doi: 10.1109/TITS.2020.2983835.

[14] M. C. Buiten, "Towards intelligent regulation of artificial intelligence," *European Journal of Risk Regulation*, vol. 10, no. 1, pp. 41–59, Mar. 2019, doi: 10.1017/err.2019.8.

[15] H. Ye, L. Liang, G. Y. Li, J. Kim, L. Lu, and M. Wu, "Machine learning for vehicular networks: recent advances and application examples," *IEEE Vehicular Technology Magazine*, vol. 13, no. 2, pp. 94–101, Jun. 2018, doi: 10.1109/MVT.2018.2811185.

[16] P. A. D. Amiri and S. Pierre, "An ensemble-based machine learning model for forecasting network traffic in VANET," *IEEE Access*, vol. 11, pp. 22855–22870, 2023, doi: 10.1109/ACCESS.2023.3253625.

[17] J. Wu and S. Shang, "Managing uncertainty in AI-enabled decision making and achieving sustainability," *Sustainability (Switzerland)*, vol. 12, no. 21, pp. 1–17, Oct. 2020, doi: 10.3390/su12218758.

[18] X. Di and R. Shi, "A survey on autonomous vehicle control in the era of mixed-autonomy: From physics-based to AI-guided driving policy learning," *Transportation Research Part C: Emerging Technologies*, vol. 125, p. 103008, Apr. 2021, doi: 10.1016/j.trc.2021.103008.

[19] S. S. Band *et al.*, "When smart cities get smarter via machine learning: an in-depth literature review," *IEEE Access*, vol. 10, pp. 60985–61015, 2022, doi: 10.1109/ACCESS.2022.3181718.

[20] A. Talpur and M. Gurusamy, “Machine learning for security in vehicular networks: a comprehensive survey,” *IEEE Communications Surveys and Tutorials*, vol. 24, no. 1, pp. 346–379, 2022, doi: 10.1109/COMST.2021.3129079.

[21] M. C. Lavanya, M. Akshatha, P. Radhakrishnan, K. V. S. Bai, S. Vijayalakshmi, and J. Ranga, “Machine learning based enhanced autonomous driving for autonomous vehicles,” in *International Conference on Inventive Computation Technologies (ICICT)*, Apr. 2023, pp. 78–82, doi: 10.1109/ICICT57646.2023.10134201.

[22] K. Kandali, L. Bennis, H. Halaq, and H. Bennis, “A novel k-means powered algorithm for an efficient clustering in vehicular ad-hoc networks,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 3, pp. 3140–3148, Jun. 2023, doi: 10.11591/ijece.v13i3.pp3140-3148.

[23] S. Sulthana and B. N. R. M. Reddy, “Machine learning algorithms for privacy preserving in vehicular ad hoc network,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 2, pp. 1021–1028, May 2023, doi: 10.11591/ijeecs.v30.i2.pp1021-1028.

[24] E. S. Ali *et al.*, “Machine learning technologies for secure vehicular communication in internet of vehicles: recent advances and applications,” *Security and Communication Networks*, vol. 2021, pp. 1–23, Mar. 2021, doi: 10.1155/2021/8868355.

[25] M. Saleem, S. Abbas, T. M. Ghazal, M. A. Khan, N. Sahawneh, and M. Ahmad, “Smart cities: fusion-based intelligent traffic congestion control system for vehicular networks using machine learning techniques,” *Egyptian Informatics Journal*, vol. 23, no. 3, pp. 417–426, Sep. 2022, doi: 10.1016/j.eij.2022.03.003.

[26] J. Posner, L. Tseng, M. Aloqaily, and Y. Jararweh, “Federated learning in vehicular networks: opportunities and solutions,” *IEEE Network*, vol. 35, no. 2, pp. 152–159, Mar. 2021, doi: 10.1109/MNET.011.2000430.

[27] A. Handa, A. Sharma, and S. K. Shukla, “Machine learning in cybersecurity: a review,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 4, p. e1306, Feb. 2019, doi: 10.1002/widm.1306.

[28] Y. Kumar, K. Kaur, and G. Singh, “Machine learning aspects and its applications towards different research areas,” in *International Conference on Computation, Automation and Knowledge Management, ICCAKM*, Jan. 2020, pp. 150–156, doi: 10.1109/ICCAKM46823.2020.9051502.

[29] J. G. Nahr, H. Nozari, and M. E. Sadeghi, “Artificial intelligence and machine learning for real-world problems (a survey),” *International Journal of Innovation in Engineering*, vol. 1, no. 3, pp. 38–47, Oct. 2021, doi: 10.59615/ijie.1.3.38.

[30] J. M. L. Domínguez, F. Al-Tam, T. de J. M. Sanguino, and N. Correia, “Analysis of machine learning techniques applied to sensory detection of vehicles in intelligent crosswalks,” *Sensors (Switzerland)*, vol. 20, no. 21, p. 6019, Oct. 2020, doi: 10.3390/s20216019.

[31] M. Bansal, A. Goyal, and A. Choudhary, “A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning,” *Decision Analytics Journal*, vol. 3, p. 100071, Jun. 2022, doi: 10.1016/j.dajour.2022.100071.

[32] J. D. C. Julio-Rodríguez, C. A. Rojas-Ruiz, A. Santana-Díaz, M. R. Bustamante-Bello, and R. A. Ramirez-Mendoza, “Environment classification using machine learning methods for eco-driving strategies in intelligent vehicles,” *Applied Sciences (Switzerland)*, vol. 12, no. 11, p. 5578, May 2022, doi: 10.3390/app12115578.

[33] D. Hindarto and H. Santoso, “Performance comparison of supervised learning using non-neural network and neural network,” *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, vol. 11, no. 1, pp. 49–62, Apr. 2022, doi: 10.23887/janapati.v11i1.40768.

[34] K. S. Meena and S. Suriya, “A survey on supervised and unsupervised learning techniques,” in *International Conference on Artificial Intelligence, Smart Grid and Smart City Applications*, Springer International Publishing, 2020, pp. 627–644.

[35] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, “A survey on machine learning for data fusion,” *Information Fusion*, vol. 57, pp. 115–129, May 2020, doi: 10.1016/j.inffus.2019.12.001.

[36] K. Rashid, Y. Saeed, A. Ali, F. Jamil, R. Alkanhel, and A. Muthanna, “An adaptive real-time malicious node detection framework using machine learning in vehicular ad-hoc networks (VANETs),” *Sensors*, vol. 23, no. 5, p. 2594, Feb. 2023, doi: 10.3390/s23052594.

[37] M. Kezia and K. V. Anusuya, “A comparative study on machine learning algorithms for congestion control in VANET,” in *International Conference on Intelligent Innovations in Engineering and Technology, ICIIET*, Sep. 2022, pp. 38–44, doi: 10.1109/ICIIET55458.2022.9967614.

[38] M. A. Khan *et al.*, “Support-vector-machine-based adaptive scheduling in mode 4 communication,” *Computers, Materials and Continua*, vol. 73, no. 2, pp. 3319–3331, 2022, doi: 10.32604/cmc.2022.023392.

[39] C. N. Ajay and H. V. Kumaraswamy, “Traffic prediction using random forest machine learning algorithms,” *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, vol. 11, no. 4, 2022.

[40] F. A. Alhaidari and A. M. Alrehan, “A simulation work for generating a novel dataset to detect distributed denial of service attacks on vehicular ad hoc network systems,” *International Journal of Distributed Sensor Networks*, vol. 17, no. 3, p. 155014772110002, Mar. 2021, doi: 10.1177/15501477211000287.

[41] A. Alsarhan, M. Alauthman, E. Alshdaifat, A. R. Al-Ghuwairi, and A. Al-Dubai, “Machine learning-driven optimization for SVM-based intrusion detection system in vehicular ad hoc networks,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 5, pp. 6113–6122, Feb. 2023, doi: 10.1007/s12652-021-02963-x.

[42] A. R. Gad, A. A. Nashat, and T. M. Barkat, “Intrusion detection system Using machine learning for vehicular ad hoc networks based on ToN-IoT dataset,” *IEEE Access*, vol. 9, pp. 142206–142217, 2021, doi: 10.1109/ACCESS.2021.3120626.

[43] N. Kadam and K. R. Sekhar, “Machine learning approach of hybrid KSVN algorithm to detect DDoS attack in VANET,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, pp. 718–722, 2021, doi: 10.14569/IJACSA.2021.0120782.

[44] A. Sonker and R. K. Gupta, “A new procedure for misbehavior detection in vehicular ad-hoc networks using machine learning,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 3, pp. 2535–2547, Jun. 2021, doi: 10.11591/ijece.v11i3.pp2535-2547.

[45] H. Wang, X. Wei, J. Yao, and Y. Zhang, “Traffic flow prediction using machine learning methods,” in *3rd International Conference on Machine Learning, Big Data and Business Intelligence, MLBDBI*, Dec. 2021, pp. 30–35, doi: 10.1109/MLBDBI54094.2021.00014.

[46] R. Anand and S. Sankhe, “Traffic prediction for intelligent transportation systems using machine learning,” *International Journal for Modern Trends in Science and Technology*, vol. 8, no. 7, pp. 276–280, Jul. 2022, doi: 10.46501/ijmtst0807041.

[47] A. Le and C. Maple, “Shadows don’t lie: N-sequence trajectory inspection for misbehaviour detection and classification in VANETs,” in *IEEE Vehicular Technology Conference*, Sep. 2019, vol. 2019-Septe, pp. 1–6, doi: 10.1109/VTCFall.2019.8891137.

[48] A. Mythili and S. K. Magendran, “Clustering algorithm and ensemble SVM for attack detection in VANET,” *International Journal of Pure and Applied Mathematics*, vol. 119, no. 12, pp. 15407–15419, 2018.

[49] Y. R. Shi and H. Wu, “Data aggregation for road functionality detection based on machine leaning and vanet,” *Journal of Computers (Taiwan)*, vol. 29, no. 2, pp. 161–173, 2018, doi: 10.3966/199115992018042902016.

[50] S. So, P. Sharma, and J. Petit, “Integrating plausibility checks and machine learning for misbehavior detection in VANET,” in *17th IEEE International Conference on Machine Learning and Applications, ICMLA*, Dec. 2019, pp. 564–571, doi: 10.1109/ICMLA.2018.00091.

BIOGRAPHIES OF AUTHORS

See Thian Meng is a Bachelor of Information Technology (Hons) student in Faculty of Information Science and Technology, Multimedia University (MMU), Melaka, Malaysia, majoring in Data Communication and Networking. His research interests include internet of things, vehicular ad hoc network, machine learning and embedded device. He can be contacted at email: 1201302671@student.mmu.edu.my.

Sumendra Yogarayan is a lecturer at the Faculty of Information Science and Technology, Multimedia University (MMU), Melaka, Malaysia. He is an active member of the Centre for Intelligent Cloud Computing (CICC), Multimedia University (MMU). He graduated from Multimedia University (MMU) with a Master of Science (Information Technology) in 2019 and a Bachelor of Information Technology (Security Technology) in 2015. He completed his Doctor of Philosophy (Ph.D.) in Information Technology at Multimedia University (MMU) in 2022. His research interests include intelligent transportation systems, vehicular ad hoc networks, wireless communication, and mesh networks. He can be contacted at email: sumendra@mmu.edu.my.

Siti Fatimah Abdul Razak is a senior lecturer at the Faculty of Information Science and Technology, Multimedia University, since 2005. She graduated from Multimedia University (MMU) with a Doctor of Philosophy (Ph.D.) in Information Technology in 2018 and a Master of Information Technology (Science and System Management) in 2004. She is also an active member of the Centre for Intelligent Cloud Computing. Her research interest includes vehicle safety applications, the internet of things, rule mining, information systems development, and educational technology. She can be contacted at email: fatimah.razak@mmu.edu.my.

Subarmaniam Kannan is a senior lecturer in Faculty of Information Science and Technology, Multimedia University since 2000. He has a Ph.D. in Semantic Learning (Knowledge Engineering) from Multimedia University. He is also a Certified Information Systems Auditor (CISA) and Certified Cisco Networking Associate (CCNA) registrar and instructor for MMU-Melaka Local Networking Academy. His research area includes semantic web technology and knowledge management, automatic speech recognition for Bahasa Malaysia, information system audit, internet of things, and edge computing. He can be contacted at email: subar.kannan@mmu.edu.my.

Afizan Azman is an Associate Professor at Taylors Univerity. Previously he was the Vice Chancellor of Research and Innovation at Kolej Universiti Islam Melaka (KUIM). He graduated from Loughborough University with a Doctor of Philosophy (Ph.D.) in Computer Science in 2013 and from University College of London with a Master of Information Technology in Computer Science in 2005. His research interests include human-computer interaction, artificial intelligence, vehicle systems and technology, data analytics, and the internet of things. He can be contacted at email: afizan.azman@taylors.edu.my.