

Research Article

An Efficient Traffic Incident Detection and Classification Framework by Leveraging the Efficacy of Model Stacking

Zafar Iqbal ¹, Majid I. Khan,¹ Shahid Hussain,^{1,2} and Asad Habib ³

¹Department of Computer Science, COMSATS University Islamabad, Islamabad 45550, Pakistan

²Department of Computer Science and Information Science, University of Oregon, OR 97403, USA

³Institute of Computing, Kohat University of Science and Technology, Kohat, Pakistan

Correspondence should be addressed to Zafar Iqbal; zafariqbal0703@gmail.com

Received 26 January 2021; Revised 21 February 2021; Accepted 29 March 2021; Published 10 April 2021

Academic Editor: Átila Bueno

Copyright © 2021 Zafar Iqbal et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Automatic incident detection (AID) plays a vital role among all the safety-critical applications under the parasol of Intelligent Transportation Systems (ITSs) to provide timely information to passengers and other stakeholders (hospitals and rescue, police, and insurance departments) in smart cities. Moreover, accurate classification of these incidents with respect to type and severity assists the Traffic Incident Management Systems (TIMSs) and stakeholders in devising better plans for incident site management and avoiding secondary incidents. Most of the AID systems presented in the literature are incident type-specific, i.e., either they are designed for the detection of accident or congestion. While traveling along the road, one may come across different types of traffic incidents, such as accidents, congestion, and reckless driving. This necessitates that the AID system detects and classifies not only all the popular traffic incident types, but severity as well that is associated with these incidents. Therefore, this study aims to propose an efficient incident detection and classification (E-IDC) framework for smart cities, by incorporating the efficacy of model stacking, to classify the incidents with respect to their types and severity levels. The experimental results showed that the proposed E-IDC framework achieved performance gains of 5%–56% in terms of incident severity classification and 1%–14% in terms of incident type classification when applied with different classifiers. We have also applied the Wilcoxon test to benchmark the performance of our proposed framework that reflects the significance of our approach over existing individual incident predictors in terms of severity and type classification. Moreover, it has been observed that the proposed E-IDC framework outperforms the existing ensemble technique, such as XGBoost used for the classification of incidents.

1. Introduction

The worldwide urban and highway transportation comes across several traffic incidents in every single moment with the increase in the number of on-road vehicles, and a single irregularity may disturb daily transportation and logistics operations [1]. The Road Traffic Incidents (RTIs), referred to as incidents, are defined as the unexpected scenarios that occur on the road network, often in potentially dangerous situations, and disrupt the normal traffic flow. The interrupted flow of road traffic can be majorly categorized into three types, i.e., (1) stoppage/hurdle on the road, (2) recurrent/nonrecurrent congestion, and (3) traffic violations. Derived from these categories, accidents, congestion, and reckless driving are the most common types of road

incidents [2] and are highly interlinked with each other, as stated by the U.S. Department of Transportation (DOT) in [3]. According to the World Health Organization (WHO) report published in [4], road accidents are the major category of traffic incidents and are considered as the ninth leading cause of death globally. Moreover, it was also highlighted that the major cause of accidents is reckless driving that is often defined as a mental state in which the driver displays a wanton disregard for the rules of the road; the drivers misjudge common driving procedures, e.g., speeding, driving under the influence of drugs, and nonuse of helmets and seat belts. According to report [5], reckless driving causes 33% of all deaths involving major car accidents, which are more than 13,000 each year. Another challenging type of on-road traffic incident is

congestion that relates to an excess of vehicles on a portion of a roadway at a time resulting in slower speeds, sometimes much slower than normal or “free flow” speeds, the Federal Highway Administration (FHWA) [6]. Congestion can be categorized into two types: recurrent which happened due to insufficient road resources and nonrecurrent which occurred due to another primary incident. According to a report by the Texas Transportation Institute (TTI), [7] on urban transportation, congestion is a reason for 8.8 billion extra hours in travel time and wastage of 3.3 billion gallons of fuel, and on average, the autocommutter has to spend extra 54 hours traveling compared to normal traveling time [8]. Authors in [9] worked to identify the correlation between the co-occurrence of colocated traffic incidents and eventually categorized them into primary and secondary incidents. It is also observed that 20%–50% of traffic incidents occur due to primary incidents and 50% of secondary incidents occur within 10 minutes of the primary incident. Similarly, an additional minute collapse in the clearance of the incident site increases the chances of near future incidents by 2.8% [10]. Moreover, the interdependency of these RTIs makes the boundary of incident classification blurred, which increases false alarm rate (FAR) and misclassification of the incident type (IT), and makes it hard to guess the risk factor (RF) also known as incident severity (IS) associated with the RTIs. The risk of secondary incidents can be reduced by deploying an incident management system (IMS) to timely and accurately detect the primary incident type (accident, congestion, and reckless driving) and its associated severity (low, medium, and high) and reporting the RTI with its details to appropriate rescue departments for timely clearance.

Existing IMS for smart cities can be classified into preincident management systems and postincident management systems [11]. Preincident management involves the study of incident prediction and prevention. These systems are designed to generate alerts for vehicles before they get involved in any incident. The International Transport Forum (ITF) [12] reports the assumptions for these systems in achieving “zero-incident.” A zero-incident situation is a symbol of perfection that is hard to achieve as it requires each component of the system to be perfect and working all the time in a synchronized manner. On the other hand, postincident management systems include an entire Traffic Incident Management System (TIMS) [13]. It comprises three major modules including incident detection, reporting, and clearance. The incident detection module includes the systems that evolved from nonautomatic incident detection systems via mobile telephone calls by eye-witnesses to automatic incident detection (AID) systems [14] that utilize parametric data gathered from on-road and in-vehicle sensing devices. The data are then analyzed using comparative, statistical, time-series, and ML algorithms to extract traffic patterns that are used to classify incidents [15]. The reporting module reports the detected incident to the appropriate stakeholders, e.g., hospitals, insurance agencies, traffic control room (TCR), law enforcement agencies, and neighboring vehicles. In the clearance module, actionable activities are arranged to clear the incident sight, e.g.,

ambulances, response teams, and changeable message signs (CMSs) [16, 17].

This paper focuses on ML-based AID systems for smart cities that can be further classified into two categories with respect to (a) incident context and (b) classification: a context-aware AID system is designed to classify a single type of incident only, such as accident or congestion. The context-free systems are more generic and can classify among different types of traffic incidents, such as accident, congestion, and reckless driving. For example, a context-free AID system is proposed [18] in which accidents, road debris, and road works are considered as incidents and are detected using restricted Boltzmann machine (RBM). However, the system lacks in providing real-time support due to very high detection time. On contrary, most of the ML-based AID systems available in the literature, such as [19, 20], are context-aware. These systems have a very high probability of incident misclassification because these are designed to classify only a specific type of incident, e.g., an AID deployed to classify accident from traffic data may classify a stalled vehicle due to congestion to an accident. In addition, the ML-based AID system can be characterized as a single-feature classification (SFC) which classifies the single feature or one aspect of the underlying scenario while multifeature classification (MFC) can classify or predict multiple features or multiple aspects about the underlying problem. Therefore, a context-free AID system equipped with MFC can extract multiple aspects of the incident, such as type and severity.

Hence, the main aim of this research is to design an efficient incident detection and classification (E-IDC) framework for smart cities by leveraging the efficacy of model stacking to classify the incidents accurately with respect to incident type (accident, congestion, and reckless driving) and severity level (low, medium, and high). Following research questions (RQs) are listed to evaluate the performance of the proposed E-IDC framework:

RQ-1: is the proposed framework effective to classify the incidents at their severity levels?

RQ-2: is the proposed framework effective to classify the incidents into appropriate categories/contexts?

RQ-3: does the proposed framework outperform the existing state-of-the-art ensemble learning approach used for incident prediction?

The rest of the paper is structured as follows: Section 2 discusses the related work, Section 3 presents the proposed E-IDC framework, Section 4 describes the dataset creation mechanism and the evaluation model, Section 5 discusses the experiments and results, Section 6 describes the threat to validity, and Section 7 concludes the discussion of the article.

2. Related Work

ML-based AID systems can be categorized into two classes: (1) individual algorithm, a standalone algorithm usually based on a single classifier having classification and detection logic, and (2) fusion algorithms, integration of multiple

algorithms through ensemble learning to improve and enhance the performance of individual classifiers. Multiple systems consider incident classification as a multiclass problem in which the target/dependent variable has more than two outcomes; that is, if the severity of the accident is a target class/label then fatal, severe, slight, and injurious are the possible outcomes that may lead towards misclassification. The following subsections discuss briefly the individual and fusion-based AID systems.

2.1. Individual Algorithm-Based AID Systems. The native ML-based AID systems were based on individual algorithms that take incident classification as a binary or multiclass problem. The binary classification represents two possible outcomes in the target variable, i.e., “1” if an incident occurs and “0” otherwise, while the multiclass also selects the target variable with multiple labels but selects one label at a time. In this section, existing ML-based incident detection and classification systems are discussed with respect to the context of the investigated incidents, parameters used in the incident predictor, and machine learning technique along with their support for multifeature classification or single-feature classification.

2.1.1. Accident Detection Systems. ML-based accident detection systems are deployed to detect accidents and their relevant details, i.e., crash severity, injury severity, driver behavior, weather condition, road conditions, type of vehicles involved in accidents, location of the vehicle, peak hours, and nonpeak hours. The study [18] proposed a spatiotemporal pattern network (STPN) and restricted Boltzmann machine (RBM)-based system that can find the presence of incident along with its root-cause analysis, but due to lack of data to train the model, the case study in this paper does not differentiate secondary incidents from all incidents and it focuses on accident only. Authors in [21] proposed an accident injury classification by using hierarchical clustering and artificial neural networks. The classification is restricted to a limited number of resources, large overhead, and lack of precision for a large number of clusters. The literature [22] proposes a traffic accident detection system using the random forest in which vehicles share their microscopic vehicular parameters with neighbor vehicles. Incidents are considered as outliers, which limits it to take binary classification decisions. A copula-based approach is employed in [23] with a multivariate temporal ordered probit model to simultaneously predict injury and damage severity. Though it provides MFC support in classification, its functionality is limited within the same context. Authors in [24] find the weightage of attributes by using rough set theory and support vector machine (SVM) to predict the severity of accidents. In the literature, a semisupervised model is proposed that detects vehicle incident trajectories coming from the cameras using a deep-learning-based You Look Only Once (YOLOv3) classifier was projected [25]. In another study, [26] presented a system to investigate accident likelihood and severity by manipulating real-time traffic and weather data. The results

presented in this paper are based on data from two urban arterials. In [27], a parametric sensitivity analysis method is proposed to shape an algorithm accident detection using SVM for rural roads. The development of automatic severity classification of accidents using advanced machine learning algorithms was carried out in the traffic accident literature [28]. The discussion on accident detection and classification systems is summarized in Table 1.

2.1.2. Congestion Detection Systems. In the literature, researchers have focused on another important issue of traffic congestion that costs billions of dollars around the globe. The SFC-based ML approaches are currently being used to detect the presence, severity, and causal of congestion. The study [29] used the speed performance index to measure present congestion levels on urban roads. A procedure based on a piecewise switched linear (PWSL) macroscopic traffic model and an exponentially weighted moving average (EWMA) monitoring chart for the detection of traffic congestion was carried out in [30] which suffers from a lack of precision. In the literature [31], congestion levels are estimated by using clustering and find out the road area per capita along with the ownership of vehicles per thousand people, which affects the ratio of congestion. In [32], a novel traffic congestion detection algorithm is designed from two aspects. One is the offline traffic data processing, and the other is congestion mode judgment by online monitoring, but the system has very low support for real-time implementation. The study [33] presented an integrated and computable definition of congestion in which a congestion dataset that contains multiple traffic scenes is constructed, and an algorithm is developed based on inverse perspective mapping (IPM). The deep-learning method does not work well, and a stronger label including the position of vehicles and a carefully designed deep network shall be exploited. In another study [34], a congestion detection algorithm is proposed using multilayer perceptron (MLP) and convolutional neural networks (CNNs), but there is a lack of incident type and severity. The study [35] presented a system for automated congestion detection on a road segment based on lane-changing properties. The literature [36] also proposed a Tree-Augmented Naïve Bayesian (TAN) classifier-based algorithm to estimate congestion. Due to lack of suitable datasets and transferability of information, the TAN-based algorithms are not evaluated in this paper. Table 2 summarizes the congestion detection systems.

2.1.3. Reckless Driving Detection Systems. Reckless driving is an obvious reason for traffic accidents that also leads to congestion as well. In the literature [37], the authors presented the detection of centerline crossing in abnormal driving using a deep convolutional neural network, but the system lacks in providing precision and effectiveness. A study [38] proposed a five-layer architecture with the context-aware property that can collect the driving situation and detect bad driving patterns. A system to detect bad driving by training and testing a model with real-world vehicle data was also proposed [39]. The abnormal driving

TABLE 1: Summary of accident detection systems.

Ref #	ML technique	Parameters relevant to						MFC	
		Accident	Driver	Vehicle	Road	Weather	Detector		Speed
[18]	Unsupervised	✗	✗	✓	✗	✗	✗	✓	✓
[21]	Supervised	✓	✓	✗	✓	✓	✗	✗	✗
[22]	Supervised	✗	✗	✓	✗	✗	✗	✗	✗
[23]	Supervised	✓	✓	✓	✗	✓	✗	✗	✓
[24]	Supervised	✗	✗	✓	✗	✗	✗	✗	✗
[25]	Semisupervised	✓	✗	✗	✗	✗	✗	✗	✗
[26]	Supervised	✗	✗	✗	✗	✗	✓	✗	✗
[27]	Supervised	✗	✓	✓	✗	✗	✗	✓	✗
[28]	Supervised	✓	✗	✗	✓	✗	✗	✗	✗

TABLE 2: Summary of congestion detection systems.

Ref #	ML technique	Independent variables	MFC
[29]	Unsupervised	Speed	✗
[30]	Unsupervised	Density	✗
[31]	Unsupervised	Road area, car ownership, and vehicle miles traveled	✗
[32]	Unsupervised	Average speed, density, traffic flow	✗
[33]	Supervised	Spatial-temporal activity, image pixel for the speed	✗
[34]	Supervised	Loop data	✗
[35]	Unsupervised	ST properties	✗
[36]	Supervised	Volume, speed, occupancy	✗

pattern of vehicles using multiple view geometry in space and time, e.g., transverse motion, acceleration/deceleration, and meandering, is proposed [40]. Another work [41] proposed a real-time GPS-based abnormal driving detection for intelligent vehicles, but there is a lack of precision due to insufficient training data. The authors in [42] proposed an innovative driver behavior monitoring system in real-time, because of vehicle dynamics, road design values, and passenger comfort. To identify and detect the abnormal driving behavior, e.g., lane changing, weaving, and sudden braking by using pattern matching techniques were carried out in [43]. The study did not consider various road types, vehicles and placement positions of the phone, and variable sampling rates, and their effect on the pattern. In the literature [44], an anomaly detection system is proposed, namely, D&R Sense to enrich smart transportation and making road journeys safe and comfortable. Table 3 summarizes various aspects of reckless driving detection systems.

2.2. Fusion Algorithm-Based AID Systems. Fusion algorithms are commonly implemented through ensemble learning technique which is one type of ML technique in which multiple classifiers are combined to improve the performance and accuracy of the classification. Stacking is an ensemble learning that is used to explore a space of different models for the same problem. The logic is to take a learning problem with different types of models that are capable to learn some part of the problem, but not the whole space of the problem. Hence, multiple different learners could be built and further used to build an intermediate prediction, one prediction for each learned model. Then, a new model

could be added which learns from the intermediate predictions for the same target. Traffic incident detection systems use the ensemble learning approach to achieve accuracy and minimize the false alarm rate.

Authors in the literature [45] proposed a model for injury severity analysis by using two-layer stacking but the effectiveness of the model is achieved by setting the values of various parameters as fixed (e.g., values in the ideal environment). The literature [46] proposed a freeway accident detection by using a two-stage approach that optimally sums up the data generated from sensors using Bayesian quickest change detection framework, but lack of MFC support and intertype classification was found. An ensemble learning technique is used in which SVM and KNN models are used for traffic incident detection [19], but there is no support for incident type classification. An incident detection and management system are proposed for traffic signals comprising of urban traffic networks [47], but the system requires extra hardware for proper functioning, hence providing a costly solution to the incident detection. There is a methodology in which traffic signals are extracted from video and converted as spatial-temporal (ST) signals to find the normal and abnormal driving patterns by using SVM and adaptive boosting on these ST signals [48]. In [49], it is studied that the causes of traffic congestion using collaboration between connected vehicles and used voting procedure (VP) and belief function (BF) methods for congestion estimation, but it has very low support for real-time implementation due to lack of validity. A framework was designed for real-time distributed congestion classification on a diverse urban road network using a decision tree and probabilistic Naïve Bayes classifier [50]. In a study [20], an

TABLE 3: Summary of reckless driving detection systems.

Ref #	ML technique	Independent variables	MFC
[37]	Supervised	Position, size, orientation, speed, reflect factor, and texture	✗
[38]	Supervised	Speed, position, acceleration, driver's eye, and intoxication	✗
[39]	Supervised	Engine speed, battery voltage, clutch engagement, and TP	✗
[40]	Supervised	Image-based parameters, i.e., speed, acceleration, and angles	✗
[41]	Unsupervised	Position and velocity	✗
[42]	Supervised	Speed/acceleration	✗
[43]	Unsupervised	Speed and acceleration	✗
[44]	Supervised	Latitude, longitude, timestamp, and accelerations along the axis	✗

AID algorithm is developed based on standard SVM with the bagging method but the system lacks providing real-time implementation support, and boosting may outperform bagging. In the study [51] (XGBoost), the ML technique is used to detect the occurrence of accidents using a set of real-time data comprised of traffic, network, demographic, land use, and weather features but the performance of this system is limited to highway-based accident detection. The literature on fusion algorithms shows that multiexpert and multistage systems are currently being used to improve the performance of individual algorithm-based incident detection and classification systems. However, all these systems are context-aware in nature and do not support multitarget classification. Table 4 summarizes various aspects of fusion algorithm-based AID systems.

To summarize the entire discussion on existing AID systems, it could be said that existing systems support univariate or multivariate target variable but within the same incident context (either it detects the presence of incident or it classifies its severity) and these are not compatible for real-time deployment due to the imprecise training based on offline features (collected from the police department or hospitals after the incident is rescued). So, in the context of TIMS, it requires a full understanding of several aspects of the incident (incident type and severity, type of vehicles collapsed, peak or off-peak hours and location, etc.) that help in devising more efficient incident clearance plans that may help in minimizing the occurrence of secondary incidents in smart city paradigm. Thus, an effective incident detection and classification framework is needed to reduce the clearance delay in any circumstances (accident, congestion, reckless driving, and other environmental conditions such as weather and time). In each dimension of TIMS, the quick and valid response based on maximum information relevant to the incident could be more beneficial for all stakeholders in devising appropriate decision making as well as to avoid the near future incidents.

3. Efficient Incident Detection and Classification (E-IDC) Framework

We propose an efficient incident detection and classification framework based on enhanced TIMS architecture as well as incident detection and classification algorithm to effectively detect and classify the incident type and severity for smart cities.

TABLE 4: Summary of fusion algorithm-based incident detection systems.

Ref #	ML technique	ML algorithm	MFC	Incident context
[45]	Multiexpert	Stacking	✗	Accident
[46]	Multistage	—	✗	Accident
[19]	Multiexpert	Voting	✗	Accident
[47]	Multistage	—	✗	Congestion
[48]	Multistage	Boosting	✗	Congestion
[49]	Multiexpert	Voting	✗	Congestion
[50]	Multistage	Boosting	✗	Congestion
[20]	Multistage	Bagging	✗	Congestion
[51]	Multistage	Boosting	✗	Accident

3.1. Enhanced TIMS. Similar to the existing TIMS architecture, the proposed enhanced TIMS architecture also comprises three major components: (1) on-road networks, (2) intelligent roadside units, and (3) traffic control room. However, the enhancements have been proposed in all the abovementioned components to improve the operations of the TIMS as depicted in Figure 1 and discussed as follows:

- (1) On-road network: it comprises various sensing resources (e.g., loop detectors and radar sensors) that can be used to collect the desired on-road parameters of interest from a smart city environment, such as traffic flow information, weather information, and vehicle specific information.
- (2) Intelligent roadside unit (I-RSU): traditionally, the RSUs were used as access points (APs) to collect the data from the road network and forward it to the traffic control room. We propose to enhance the capabilities of RSUs by deploying a pretrained incident detection and classification algorithm on them. The resulting RSU is named intelligent roadside units (I-RSUs), as they are now capable of processing the data received from the on-road network and classify incidents based on type (accident, congestion, or reckless driving) and severity. Figure 1 depicts the I-RSU and its two components, (1) transceiver, which receives real-time data from on-road sensors and forwards it to TCR along with processed incident-specific information, and (2) incident detection and classification algorithm.
- (3) Traffic control room (TCR): TCR is a centralized traffic control room dedicated to managing the traffic flow of a city. In this study, TCRs are used for offline

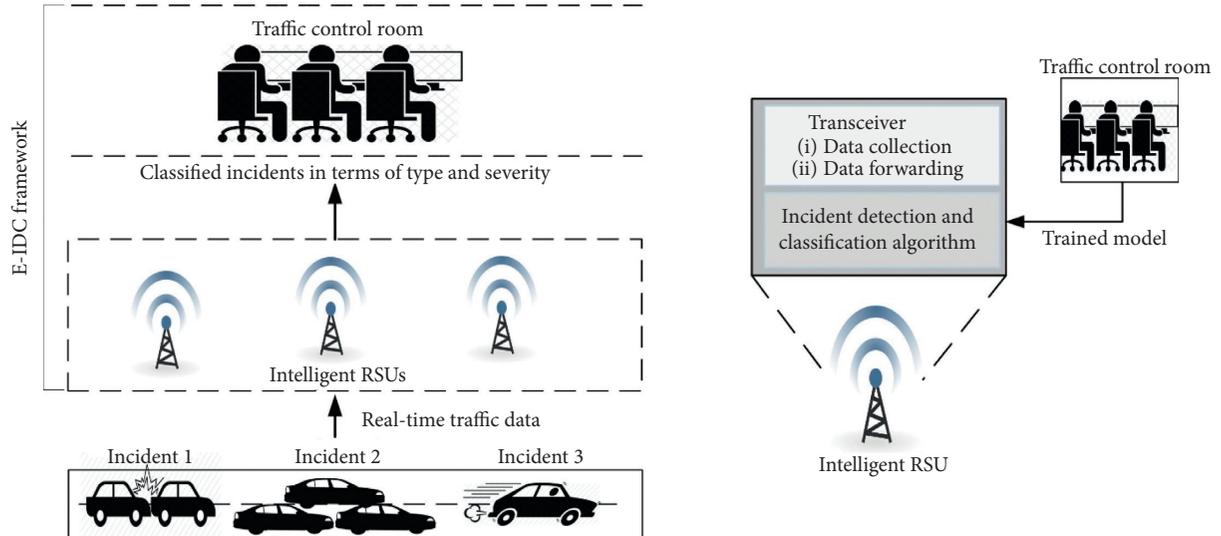


FIGURE 1: E-IDC framework (enhanced TIMS).

training of the incident detection and classification algorithm that is deployed on the I-RSU.

3.2. Incident Detection and Classification Algorithm. The proposed E-IDC framework utilizes model stacking to detect and classify road traffic incidents based on type and severity. Though the model stacking has been used in various studies [52] to provide a solution for the classification problems, its performance is restricted to SFC as it considers only one target variable for decision making. In this study, we have enhanced the capability of the model stacking approach to provide MFC and make it context-free. The model stacking comprises of two levels: the “level 1 ($L1$)” classifies the incident severity and “level 2 ($L2$)” is to classify the incident type. As a result, the classification of the incident (type) with its associated impact (severity) can be classified in a single iteration. Moreover, since the incident type and severity classification are interdependent problems, therefore, incident severity prediction at $L1$ can help to improve the classification (accuracy and performance) of incident type at $L2$. The entire classification procedure of the proposed E-IDC framework is logically presented in Figure 2 as a block diagram.

We have considered three incident types, i.e., accident (A), congestion (C), and reckless driving (R) with their associated severity in three levels, i.e., low (L), medium (M), and high (H), which results in a total of nine (type and severity) pairs. At $L1$, nine models ($M1$ to $M9$) are trained using raw data X (obtained from the on-road network), to classify one of the nine (type and severity) pairs, e.g., $M1$ is trained to classify (type (accident) and severity (low)). Each model will give its output as either 0 or 1 according to the belonging of the (type and severity) pair. The output vector X' of $L1$ (comprising of nine outputs from the $L1$ models) is then passed as input to the

$L2$ that has three models ($N1$ to $N3$) that are trained to classify the three incident types (A , C , and R). Figure 3 further elaborates the details of the proposed E-IDC framework in terms of the structure and the procedure for training and testing. Firstly, the raw dataset (X) is taken having the order “ $i \times j$,” where “ i ” represents the number of instances (incident records) and “ j ” represents the features (traffic attributes) with severity as a target variable “ T_{sev} ”. The dataset is then split into nine data segments (A_L , A_M , and so on) with respect to the (type and severity) pairs. These data segments are used to train the nine $L1$ models, i.e., $M1 (A, L)$, $M2 (A, M)$, $M3 (A, H)$, \dots , $M9 (R, H)$ for (type and severity) pair classification. After training of $L1$ models, the severity predictions for “ i ” records are taken using k -fold cross-validation, e.g., all the nine models predict the severity label as 1 or 0; that is, “1” will be predicted if the instance belongs to the model and “0” will be predicted if it does not belong. Afterward, the predictions from nine $L1$ models “ $PM1, PM2, \dots, PM9$ ” for “ i ” records are taken as features for creating the training dataset (termed as X') for $L2$. The target variable “ T_{type} ” is assigned the ground truth values/labels that are known during the training procedure. X' is then split into three data segments each for a specific incident type classification (A , C , and R). These data segments are used to train three $L2$ models, such as $N1$ (accident), $N2$ (congestion), and $N3$ (reckless driving) using k -fold cross-validation; that is, all the models predict the incident type label as 0 or 1. After training the model stacking-based E-IDC algorithm, the unlabeled testing data (T) having the order “ $i \times j$ ” without target variable “ T_{sev} ” (i.e., “?” in Figure 3 shows the emptiness of the target variable) are passed to the nine trained models of $L1$. All nine models will predict the incident severity label for “ i ” instances of “ T ” and create a dataset (T') from these predictions. The dataset T' with empty target variable “ T_{type} ” is passed into

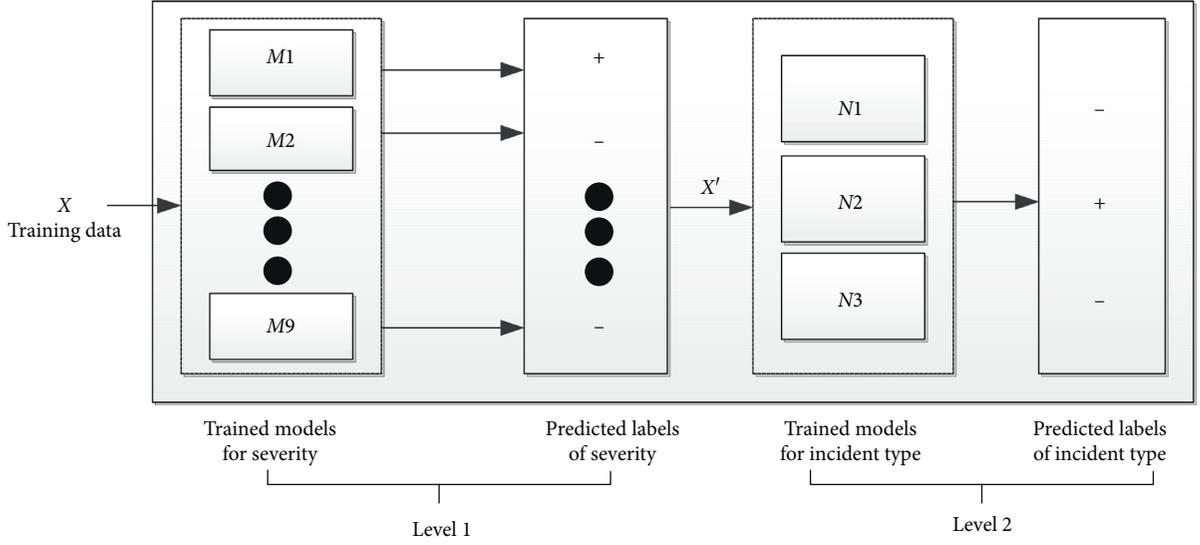


FIGURE 2: Block diagram of proposed classification algorithm.

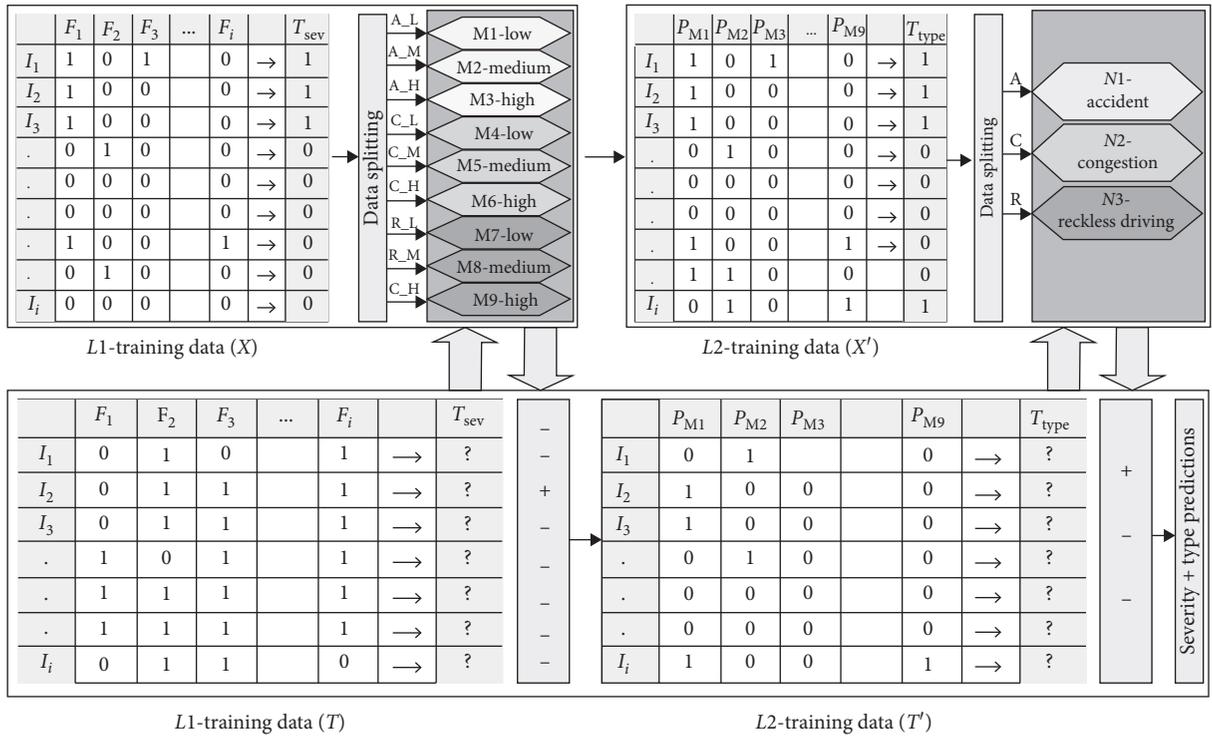


FIGURE 3: Elaborated view of the proposed classification algorithm.

three trained models of $L2$ that will predict the incident type labels for all the test instances. Hence, the predictions of both levels give the final incident severity and type predictions and make the “E-IDC framework” compatible with the MFC support.

4. Dataset Description and Evaluation Model

The following section describes the dataset creation procedure along with the evaluation criterion set for the proposed E-IDC framework:

4.1. Dataset Description. The dataset description presented in this section comprises three activities including dataset extraction, preprocessing, and sensitivity analysis.

4.1.1. Dataset Extraction. There is no single public dataset that collectively contains information about the three major road traffic incidents (accident, congestion, and reckless driving). Therefore, the accident dataset (D1) was retrieved from the link (smoosavi.org/datasets/us_accidents) which was spans over 49 states of the US from February 2016 to onwards. The D1 is used in the literature [53] for accident risk prediction. The dataset on congestion and reckless driving was collected from the United States (US) Government dataset repository (data.gov). The congestion dataset (D2) contains the past data on congestion of 1270 Chicago's arterial roads from August 2011. The data at Chicago Traffic Tracker are continuously collected through GPS traces received from Chicago Transit Authority (CTA) buses. Thirdly, the dataset for reckless driving (D3) is collected which spans the historic data of the motorists who exceed the safe speed limit of Chicago City by at least 12 miles per hour (mph), and based on these data, citations/penalties could be generated. Since the datasets on congestion and reckless driving were relevant to only one state, i.e., Chicago; therefore, we have only selected the data of Chicago out of 49 states of D1, and the timespan for D1 was in-line with the time frame of D2 and D3 to maintain the uniformity. The retrieved datasets independently provide labeled data for a specific type (i.e., accident or congestion) along with their severity.

4.1.2. Data Preprocessing. In order to create a single dataset containing incidents of three types (accident, congestion, and reckless driving) with their three associated severity levels (low, medium, and high), the collected datasets (D1, D2, and D3) are integrated to create a single dataset. The integration of the data is done based on the common demographic (location) and temporal (time and date) attributes from all three datasets to align in a single file. Eleven attributes (start latitude, start longitude, road segment number, temperature (F), humidity (%), pressure (in), visibility (mi), wind speed (mph), weather condition, traffic signal, and sunrise/sunset) are selected from the D1 dataset, and five attributes (start latitude, start longitude, road segment number, date, and time) were selected from D2 and D3 datasets.

The weather conditions can also play a vital role in transportation behavior, and it was missing in D2 and D3 datasets which were extracted from the online weather data sources by using date, time, latitude, and longitude. By merging these attributes, a total of eleven attributes were combined in a single file, and labels were assigned to each instance as given in the original dataset. The created dataset was then analyzed, and it was observed that the selected attributes were not enough to classify multiple incident types along with their severity. The reason for this deficiency was the lack of vehicle-related attributes that would better describe the traffic condition to differentiate the type of

incidents that occur on the road. Therefore, we have used the Simulation of Urban Mobility (SUMO) simulator to impersonate the incidents that are already present in the historical datasets by using the location (latitude and longitude) and environmental situations. A road traffic scenario was generated in SUMO in which we export the Chicago City maps from OpenStreetMap (OSM) website. Random trips were generated on various areas of the map, and various detectors were applied to multiple places on the map. As SUMO is a collision-free simulator, we have forcefully simulated accidents by implementing stops in vehicles at random places and random lanes. Congestions are simulated at various lanes, and speed is increased three times to simulate drivers' aggressive driving behavior. The final dataset after integration holds thirty-two attributes of which thirty (historical and simulator-based attributes) are serving as independent and two as dependent variables. The description of dependent variables, i.e., incident severity and incident type, is given in Table 5.

To improve the effectiveness of the created dataset in terms of accuracy in incident classification, the under-sampling technique is applied on the integrated data in which all the instances with missing values were excluded to remove the ambiguities, and all the classes are balanced with an equal number of instances.

4.1.3. Dataset Sensitivity Analysis. To identify the validity and contribution of each feature of the created dataset to the classification results of the E-IDC framework, we have extracted the weights corresponding to each feature by applying information gain attribute evaluation [54]. Figures 4 and 5 show the contribution of each factor in the classification of incidents at its types and severity levels, respectively. All the features are classified into two categories, i.e., historical data and simulated data. The features of the simulated data have high significance over most of the variables from the historical data category which shows that selected simulated attributes are imported to be available for real-time incident detection and classification. Moreover, it is clarified from the sensitivity analysis that the accuracy of incident classification results majorly depends on the presence of road, vehicular, and traffic flow-related attributes contribute about 70%–73% in the case of incident type and severity classification.

4.2. Evaluation Model for Framework. The performance evaluation of the proposed framework is done in two steps. In the first step, we have extracted various model evaluation metrics for the implemented E-IDC framework. As our proposed approach falls under the classification problem, therefore, we considered five performance evaluation metrics to evaluate our algorithm. These five metrics include true positive rate (TP), false positive rate (FP), precision (P), recall (R), and F -measure. Though we have calculated all the values of evaluation metrics, to simplify the presentation of results and considering the significance of F -measure, we have used it as a primary investigating factor.

TABLE 5: Description of dependent variables.

Var. name	Classes	Description
Incident severity	High	The incident either accident, congestion, or reckless driving that takes more than 10 minutes before clearance/identification
	Medium	The incident either accident, congestion, or reckless driving that takes from 5 to 10 minutes before clearance/identification
	Low	The incident either accident, congestion, or reckless driving that takes up to 5 minutes before clearance/identification
Incident type	Accident	The vehicular and weather-related variables show the classified incident is an accident
	Congestion	The vehicular and weather-related variables show congestion
	Reckless driving	The vehicular and weather-related variables show that the incident is reckless driving

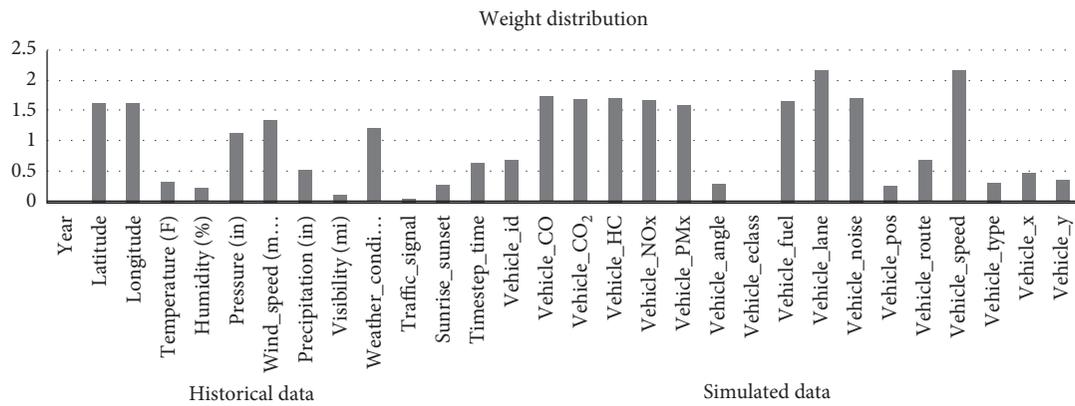


FIGURE 4: Feature weight distribution in case of incident severity classification.

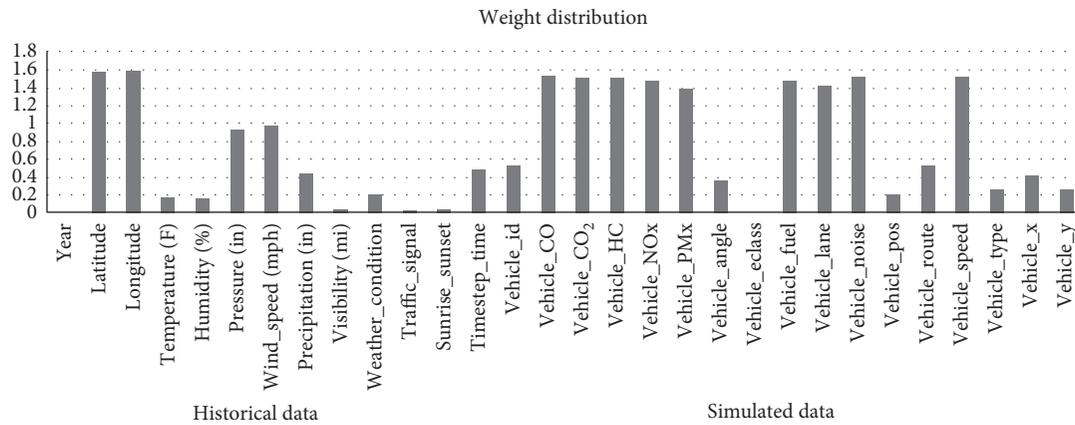


FIGURE 5: Feature weight distribution in case of incident type classification.

In the second step, the Wilcoxon test is conducted, and boxplots are drawn to benchmark the performance of the proposed E-IDC framework. Though there are several statistical tests (e.g., McNemar's, Friedman's, and Nemenyi tests) to benchmark the significance of one algorithm over the other but based on the nature and domain of our problem, we have used the Wilcoxon test. The Wilcoxon test is conducted to determine if two or more sets of pairs are different from one another in a statistically significant manner [55]. In this study, we have two groups of algorithms to be judged, i.e., individual classifiers and classifier integrated with the proposed E-IDC framework over a single

dataset. The results of the Wilcoxon test are interpreted in terms of significance values and are discussed in the next section. Afterward, the boxplots are drawn to visualize and compare the F-measures of existing (individual) and proposed (E-IDC) incident classification framework. The F-measures of both groups (individual and E-IDC framework) are taken as sample populations. The boxplots are interpreted using the lower, median, and upper quartile notations, i.e., Q_1 , Q_2 , and Q_3 of the sample data, respectively. These notations help us to determine the size of the boxes (i.e., interquartile range (IQR)), the position of median, and whisker in the boxplot. Finally, through these

parameters, we have interpreted the difference between the two groups. Following rules are considered for the interpretation of boxplots [56]:

Rule 1. F-measure comparison between two groups is considered different when boxplots do not overlap each other. Consequently, the difference between the two sample distributions can be endorsed if the following equation holds:

$$Q_1(\text{Individual}) > Q_3(\text{E-IDC}) \vee Q_1(\text{E-IDC}) > Q_3(\text{Individual}). \quad (1)$$

Rule 2. F-measure comparison between two groups will be the same when their boxplots overlap each other via medians and the following equation holds:

$$Q_2(\text{Individual}) > Q_2(\text{E-IDC}) \wedge Q_2(\text{Individual}) < Q_2(\text{E-IDC}) \vee Q_2(\text{Individual}) [Q_2(\text{E-IDC}) \wedge Q_2(\text{Individual})] Q_1(\text{Individual}). \quad (2)$$

Rule 3. F-measure comparison between two groups will be the considered uncertain when their boxplots overlap each other without medians. In this case, the subrules are derived as follows.

Subrule 3A. F-measure comparison between two groups of classes will be the same when their IQR is the same, defined using

$$\text{IQR}(\text{Individual}) = \text{IQR}(\text{E-IDC}). \quad (3)$$

However, it could not define that the underlying F-measure distribution is symmetric.

Subrule 3B. The underlying F-measure distribution will be symmetric if the difference of Q_1 and Q_3 with Q_2 (median) is the same, shown in equation (4). The same assumption can be done between whiskers and Q_2 :

$$\text{diff}(Q_1, Q_2) = \text{diff}(Q_3, Q_2). \quad (4)$$

Furthermore, we have considered the distance between medians (DBMs) and overall visible spread (OVS) measures to significantly distinguish among the distributions of [57]. According to the fractional percentage (i.e., $\text{DBM}/\text{OVS} \times 100$), the significant difference between groups can be defined in terms of sample size. For example, when the sample size is up to 30, then the fractional percentage $>33\%$ indicates the significant difference between groups. Subsequently, the fractional percentages of 20% and 10% are recommended when the sample size is ≥ 100 and 1000 observations, respectively. The sample size of the system under study is fewer than 30; therefore, the results are benchmarked according to a fractional percentage of 33%.

5. Experiments and Results Analysis

In this section, we have presented an experimental procedure along with the algorithm followed by the results. The main aim is to describe the procedure that how the proposed E-IDC framework can be employed to (1) incident severity classification, (2) incident type classification, and (3) comparison of the proposed technique with state-of-the-art ensemble technique in terms of incident type and severity classification.

5.1. Experimental Procedure. We have performed several experiments and grouped them into three phases. In the

proposed E-IDC technique training of the incident classification module is periodically improved based on the incident detection and classification received from the I-RSUs. Therefore, the classifiers with low training overhead are preferred. Moreover, due to real-time incident classification requirements, the proposed incident classification module should have a minimum response time. After conducting the experiments with a total of 14 classifiers, we have selected nine classifiers that fulfill the above requirements to be fused with the proposed E-IDC technique. While in the second phase experiments, we employed the proposed E-IDC framework described in Algorithm 1 by incorporating all the nine classifiers of the first phase (such as NB) and analyzed the improvement in the classification decision.

In the algorithm, steps 1-2 describe level 1 and predict the incident severity labels, whereas steps 3-4 are responsible for incident type classification and hence describe level 2 of model stacking. To refer to experiments of the second phase, we have used the notation “E-IDC + NB” for the NB to analyze in [58] and compare the improvement in the performance of individual incident predictors (i.e., NB) with the proposed framework (i.e., E-IDC + NB). The same notation is used for the rest of the individual incident predictors. The experiment of the third phase is performed to compare the performance of the proposed E-IDC with the existing state-of-the-art ensemble technique (i.e., XGBoost). The experiments of all phases are performed with k-fold (i.e., $k = 10$) cross-validation to assess the effectiveness of incident predictors. Besides, we performed all experiments using WEKA and R tool. Finally, to benchmark the performance of the proposed E-IDC compared to the existing incident predictors, we have applied various nonparametric tests. The first and second phase experiments aim to respond to the RQ-1 and RQ-2, respectively, while experiments of the third phase aim to respond to the RQ-3.

5.1.1. Response to RQ-1. The aim of RQ-1 is to assess the effectiveness of E-IDC as compared to individual incident predictors in terms of identification of severity level of incidents. Consequently, in this regard, we performed experiments of the first and second phases by calling Algorithm 1.

The effectiveness of individual incident predictors and their incorporation with the proposed ensemble framework

<p>Input:</p> <ol style="list-style-type: none"> 1. “X” = $S[i][j]$ is a <i>matrix</i> of training dataset with “i” instances and “j” independent variables 2. “C” = {Set of individual classifiers selected under study i.e., <i>Naïve Bayes, Decision Table. ... J48</i>} <p>Procedure Begin</p> <ol style="list-style-type: none"> 1. Split “X” to “m” datasets <ol style="list-style-type: none"> 1.1. Assign <i>binary labels</i> for each <i>severity class</i> in “X_m” 2. FOR “m” = 1 to 9 <ol style="list-style-type: none"> 2.1. Apply <i>classification algorithm</i> from “C” on “X_m” <ol style="list-style-type: none"> 2.1.1. Predict <i>severity labels</i> for “i” rows of “X” 2.1.2. Write predicted <i>severity labels</i> in “X'” 3. Split “X'” to “l” datasets <ol style="list-style-type: none"> 3.1. Assign <i>binary labels</i> for each <i>incident type class</i> in “X_l” 4. FOR $l=1$ to 3 <ol style="list-style-type: none"> 4.1. Apply <i>classification algorithm</i> from “C” on “X_l” <ol style="list-style-type: none"> 4.1.1. Predict <i>type labels</i> for “i” rows of “X'” 4.1.2. Apply <i>evaluation criteria</i> to assess the <i>F-measure</i> of the classifier <p>Procedure End</p> <p>Output: The outperform incident predictor with best <i>F-measure along with severity and type labels</i></p>

ALGORITHM 1: Model stacking [E-IDC (X, C)].

(i.e., E-IDC) in terms of F-measure is shown in Figure 6. The major findings of experimental results are as follows:

- (i) A significant increase (5%–56%) in the performance of individual incident predictors when exploiting through E-IDC except for NB
- (ii) 48% improvement in the effectiveness of individual incident predictor, namely, H-Tree when it is incorporated with E-IDC (i.e., E-IDC + H-Tree)
- (iii) Improvement in the performance of NB (1%), RF (5%), and J48 (6%) when incorporated with E-IDC as compared to their performance at the individual level
- (iv) E-IDC + J48 (F -measure = 99%) and E-IDC + LMT (F -measure = 99%) outperformed other incident predictors for correct estimation of the severity levels of an incident

From Figure 6, it can be observed that the improvements for the individual incident predictors may vary.

Consequently, there is a need to benchmark the significance of improvement for individual incident predictors when incorporated with E-IDC. In this regard, we applied a nonparametric test, namely, a two-tailed Wilcoxon signed-rank test to rank the significant improvement in classification decision of incident predictors when incorporated with E-IDC. The F-measure values of incident predictors with and without E-IDC are considered as two sample populations. The null hypothesis (H_0 : the performance of individual incident predictors with and without incorporation with E-IDC is the same for the prediction of incident severity levels). The observed value of the Wilcoxon signed-rank test (i.e., 0.004) with p value < 0.05 indicates the rejection of null hypothesis. Consequently, we can conclude

that the performance of individual predictors is not the same when incorporated with E-IDC.

Moreover, to evaluate the boxplot drawn for incident severity classification shown in Figure 7, we apply the rules defined earlier. It can be observed that the IQR of both groups overlaps without medians; hence, first and second rules are rejected. In the case of individual and E-IDC incident predictor groups for incident severity classification presented in Figure 7, the IQR of both incident predictors slightly overlaps and it can be observed that the Q_2 and Q_3 of E-IDC incident predictors are greater than Q_2 of individual incident predictors. Moreover, it can be concluded from the values of whiskers and medians that the performance of E-IDC incident predictors is better than individual predictors, which makes rule 3 vague. Hence, we will apply the DBM and OVS fractional percentage test to benchmark the performance of our approach. The Q_2 value of the individual incident predictor (0.871) is less than the value of the E-IDC incident predictor (0.971). Hence, the value of DBM for this case is 0.1, and the value of OVS is 0.27. Therefore, their fractional percentage calculated through the formula is 37% that is greater than 33% (significant value for this study) significance of E-IDC over individual incident predictors for incident severity classification.

5.1.2. Response to RQ-2. RQ-2 aims to assess the effectiveness of E-IDC as compared to individual incident predictors in terms of type identification of traffic incidents. Consequently, in this regard, we performed experiments of the first and second phases, but this time, we consider the incident type rather than incident severity levels by using Algorithm 1. The effectiveness of individual incident predictors and their incorporation with

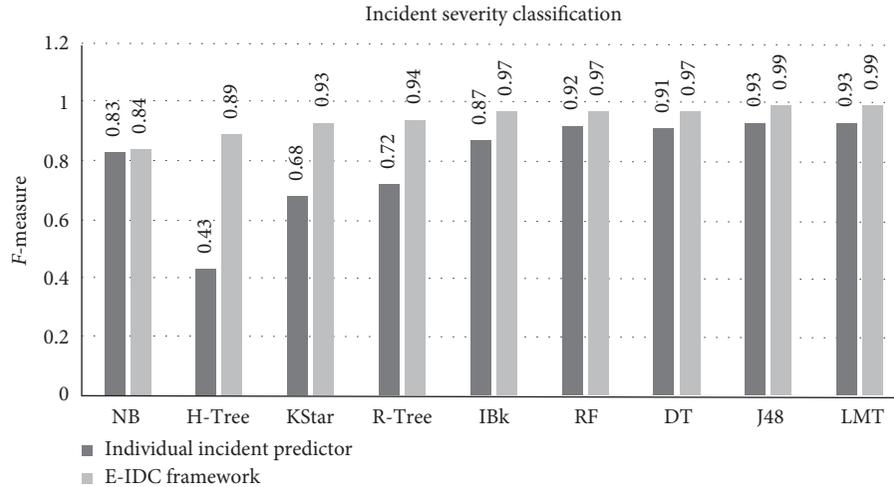


FIGURE 6: F-measures of individual and ensemble E-IDC predictors.

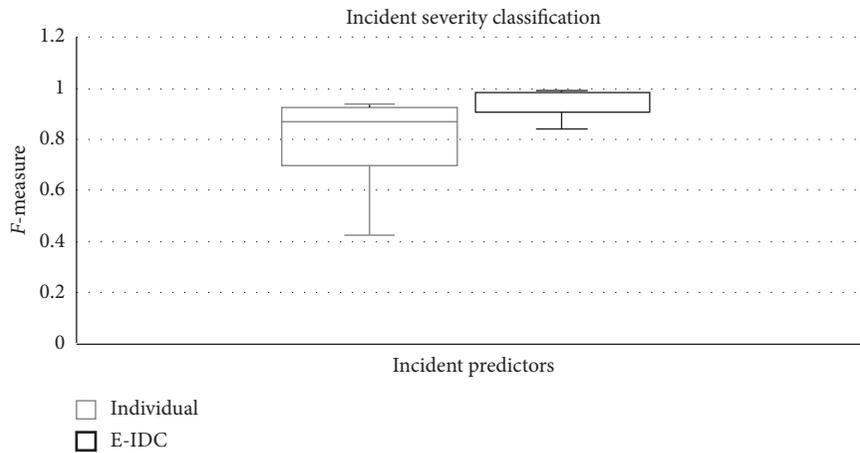


FIGURE 7: Boxplot of incident severity classification.

the proposed technique (E-IDC) in terms of F-measure is shown in Figure 8. Major findings of experimental results are as follows:

- (i) A significant increase (1%–14%) is observed in the performance of individual incident predictors when exploiting through E-IDC except for DT and J48
- (ii) 14% improvement in individual incident predictor H-Tree is observed when incorporated in E-IDC
- (iii) Minor improvement in the performance of RF, NB, IBk, and KStar is observed when incorporated with E-IDC as compared to their performance at the individual level
- (iv) We observed E-IDC + R-Tree (F -measure = 100%), E-IDC + IBk (F -measure = 100%), E-IDC + RF (F -measure = 100%), E-IDC + KStar (F -Measure = 100%), and E-IDC + LMT (F -measure = 100%) as best performer among considered predictors for the identification of types of incidents

From Figure 8, we can observe that the significant improvement in individual incident predictors varies. Consequently, there is a need to benchmark the significant improvement of individual incident predictors when incorporated with E-IDC.

In this regard, like severity level prediction, we also applied a nonparametric test, namely, the two-tailed Wilcoxon signed-rank test to rank the significant improvement in classification decision (i.e., identification of incident types) of incident predictors when incorporated with E-IDC. The null hypothesis (H_0 : the performance of individual incident predictors with and without incorporation with E-IDC is the same for the prediction of incident types).

The observed value of the Wilcoxon signed-rank test (i.e., 0.03) with p value < 0.05 indicates the rejection of null hypothesis. Consequently, we can conclude that the performance of individual incident predictors is not the same when incorporated with E-IDC. Afterward, we draw a boxplot-based comparison in Figure 9 to benchmark the

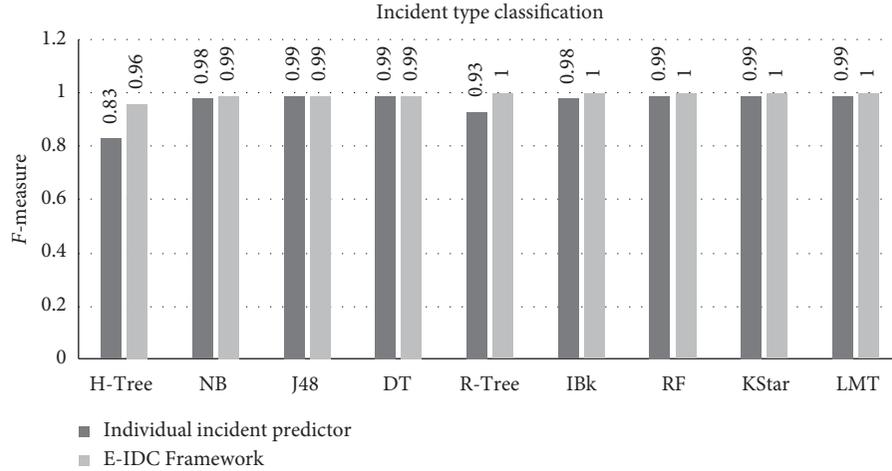


FIGURE 8: F-measures of individual and ensemble E-IDC predictors.

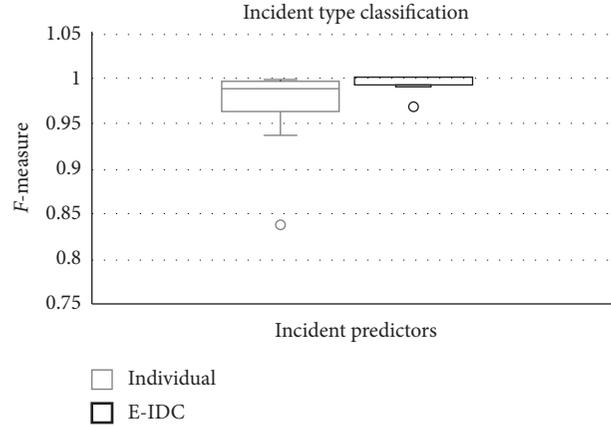


FIGURE 9: Boxplot of incident type classification.

significant improvement in the performance of the proposed incident predictor for the identification of incident types. The IQRs of both groups slightly overlap (without Q_2) from the bottom of E-IDC with the top of individual incident predictors which rejects rules 1 and 2.

Moreover, it can be observed that Q_2 and Q_3 of E-IDC are greater than Q_2 of individual incident predictors which implies E-IDC incident predictors are better as compared to individual incident predictors in terms of incident type classification. Moreover, it can also be concluded that the groups are not symmetric and rule 3 is weak to make any clear decision about the groups. Hence, to benchmark the performance, we will calculate the percentage of DBM and OVS. The Q_2 value of the individual incident predictor (0.987) is less than the value of the E-IDC incident predictor (1). Hence, the value of DBM for this case is 0.013 and the value of OVS is 0.037. Therefore, their fractional percentage calculated through the formula is 35% (>33% significant

value) which shows the significance of E-IDC over individual incident predictors.

5.1.3. Response to RQ-3. In the context of incident detection, recently, [51] has reported the implication of an ensemble technique XGBoost by leveraging the capabilities of a decision tree classifier. In order to respond to the RQ-3, we performed a comparative analysis of outperformed proposed incident type predictors (such as E-IDC + LMT) with XGBoost. The results are shown in terms of F-measure in Figure 10. In the case of incident severity level prediction, the proposed ensemble framework has improved the classification decision up to 8% of the proposed E-IDC incident predictor (i.e., E-IDC + LMT) as compared to XGBoost. Similarly, in the case of incident type prediction, the proposed E-IDC framework has improved the classification decision up to 1% of the individual incident predictor (i.e.,

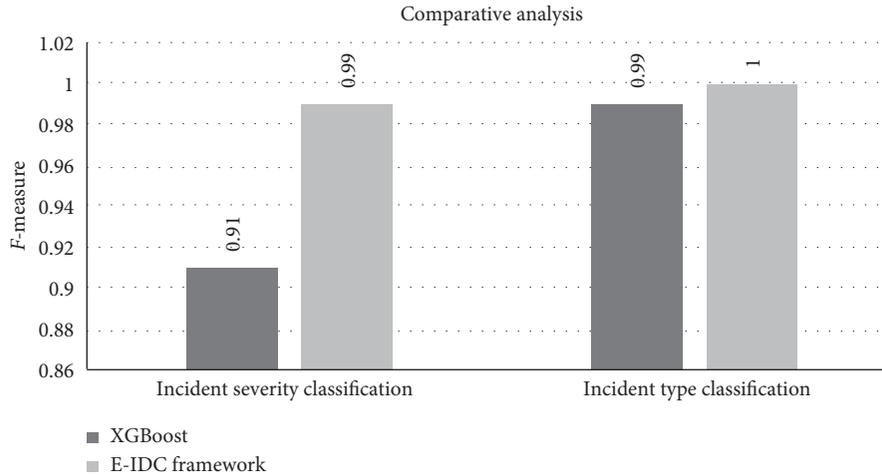


FIGURE 10: Comparison with state-of-the-art ensemble framework.

E-IDC + LMT) as compared to XGBoost. Though the low improvement in the classification decision of LMT through E-IDC has been observed, it will be true for the rest of the cases when XGBoost is ensembled with another individual classifier.

5.2. Results Evaluation Summary. The following discussion evaluates the experimental results:

- (1) The performance of an individual incident predictor can be enhanced through the proposed ensemble framework, namely, E-IDC in terms of F -measure and accuracy. In the case of severity level prediction, we observe a 5%–56% improvement for prediction of severity level while a 1%–14% improvement for the type of prediction of incidents through E-IDC.
- (2) The experimental results indicate the effectiveness of E-IDC to identify the type and its associated severity level of an incident simultaneously in a single iteration rather than identification of a specific incident type or severity level done by existing incident predictors.
- (3) We observe that E-IDC performs better compared to existing ensemble techniques, such as XGBoost used for the prediction of incidents.
- (4) The proposed model stacking approach shows improvements in both severity classification based on specialized incident severity predictors at level 1 and incident type classification based on intermediary predictions at level 2 that narrow down the domain of incident type. Level 2, thus, predicts incident type more accurately with a very low misclassification rate.
- (5) It could be observed from the experiments that the model stacking approach could be utilized for MFC-based applications in which multiple aspects of the same problem/event are to be classified. As traffic incident type and severity classification is a multivariate problem and both the target variables are highly interdependent, therefore, both the target

variables could be classified simultaneously by stacking them at various levels of model stacking.

- (6) Existing MFC-based approaches [18, 23] train the classifiers using the entire training dataset for each target variable independently and thus involve high training overhead whereas the proposed model stacking approach reduces the training overhead by exploiting the predictions of intermediate levels to obtain the final predictions.
- (7) The results of the proposed approach could be enhanced near to reality by deploying various real vehicular features in the datasets extracted from the smart vehicle, smart road infrastructure, and smart city environment as the traffic parameters show a high dynamicity in real-time.

6. Threats to Validity

In this study, we have identified the following four threats to validity:

- I. The first threat is related to the generalization of data preprocessing techniques. We have performed experiments and concluded the results with balanced incidental instances/records. An unbalanced and continuous stream of real-time data could present the modified results.
- II. The second threat is related to the use of an integrated dataset and the selection of independent parameters specific to vehicles and weather conditions. The reported results could be altered if more features are added and tuning is done.
- III. The third threat is related to the consideration of few incident predictors to assess the efficacy of the E-IDC framework. The benchmark result could be varied either if the number of incident predictors is increased or incident predictors of other families are considered.
- IV. The fourth threat is relevant to the use of selected performance measures to assess the effectiveness of the proposed approach. We used only F -measure, while

other performance metrics could be considered to assess class-level accuracy of the E-IDC framework.

7. Conclusion

E-IDC framework is based on incident classification via model stacking, an ensemble learning approach aid to provide a promising base for classifying various incident types and their severity levels simultaneously. We conclude several important consequences with respect to the experimental results. First, no individual algorithm could be specified to be stacked in the model stacking approach and varies across MFC-supported incident classification problems. All the individual algorithms provide SFC support, but their performance capabilities are enhanced by integrating them with our proposed E-IDC framework. The performance of the proposed E-IDC framework is enhanced by 5%–56% in terms of incident severity classification and 1%–14% improvement is observed in incident type classification. Second, the performance of a classifier degrades with the increase in the number of classes in the target variable, e.g., the F-measure for incident type classification is far more pleasant than severity classification as it has more classes in the target variable. Moreover, in order to determine more than two aspects of the incident simultaneously, more levels in stacking should be added as we have considered only the type and severity of the incident; therefore, we have defined only two levels in the model stacking approach. Third, to benchmark the performance of our proposed approach, we have used the Wilcoxon nonparametric test for incident severity (i.e., 0.004) and type classification (i.e. 0.03) which shows lesser values than the significant value (i.e. 0.05) and hence rejects the null hypothesis. Moreover, we observe that the E-IDC framework outperforms (8% for the case of incident severity and 1% in case of type classification) the existing ensemble framework such as XGBoost used for the prediction of incidents.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

The authors would like to thank the United States (US) Government and Chicago Transit Authority (CTA) for providing the road incidental data.

References

- [1] N. Owens, A. Armstrong, P. Sullivan, C. Mitchell, and D. Newton, *Traffic Incident Management Handbook*, Science Applications International Corporation (SAIC), Greensboro Drive McLean, VA, USA, 2010, https://ops.fhwa.dot.gov/eto_tim_pse/publications/timhandbook/tim_handbook.pdf.
- [2] H. S. Song, Y. G. Sin, and G. C. Gang, "A study of drivers' behaviors using a driving simulator with special reference of driving anger and traffic congestion," *Journal of Korean Society of Transportation*, vol. 23, no. 2, pp. 61–74, 2005.
- [3] M. E. Hallenbeck, J. Ishimaru, and J. Nee, *Measurement of Recurring versus Non-recurring Congestion*, Department of Transportation, Washington, DC, USA, 2003, <http://depts.washington.edu/trac/bulkdisk/pdf/568.2.pdf>.
- [4] World Health Organization, *Global Launch: Decade of Action for Road Safety 2011–2020*, World Health Organization, Washington, DC, USA, 2011, <http://www.who.int/roadsafety/decadeofaction/plan/planenglish.pdf>.
- [5] R. J. Fenstersheib, *An Overview of Reckless Driving Accidents*, Attorneys Hallandale, Tallahassee, FL, USA, 2020, <https://www.tellrobert.com/accidents/reckless-driving-accidents/an-overview-of-reckless-driving-accidents>.
- [6] C. Systematics, *Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation*, https://ops.fhwa.dot.gov/congestion_report/congestion_report_05.pdf, United States: Federal Highway Administration, Sacramento, CA, USA, 2005, https://ops.fhwa.dot.gov/congestion_report/congestion_report_05.pdf.
- [7] D. Schrank, B. Eisele, and T. Lomax, *2019 Urban Mobility Report*, <https://static.tti.tamu.edu/tti.tamu.edu/documents/mobility-report-2019.pdf>, Texas A&M University System and Texas Transportation Institute, San Antonio, TX, USA, 2019, <https://static.tti.tamu.edu/tti.tamu.edu/documents/mobility-report-2019.pdf>.
- [8] D. Schrank, B. Eisele, and T. Lomax, *Urban Mobility Report: Powered by Inrix Traffic Data*, Texas A&M University System and Texas Transportation Institute, San Antonio, TX, USA, 161302-1.pdf (<http://tamu.edu>), 2014.
- [9] H. J. Haule, T. Sando, R. Lentz, C.-H. Chuan, and P. Alluri, "Evaluating the impact and clearance duration of freeway incidents," *International Journal of Transportation Science and Technology*, vol. 8, no. 1, pp. 13–24, 2019.
- [10] R. Elvik, P. Christensen, and A. H. Amundsen, *Speed and Road Accidents: An Evaluation of the Power Model*, Transportøkonomisk Institutt, Oslo, Norway, 2004, https://vbn.aau.dk/ws/files/316398068/Speed_and_road_accidents_PhD.pdf.
- [11] M. G. Karlaftis, S. P. Latoski, N. J. Richards, and K. C. Sinha, "ITS impacts on safety and traffic management: an investigation of secondary crash causes," *ITS Journal - Intelligent Transportation Systems Journal*, vol. 5, no. 1, pp. 39–52, 1999.
- [12] P. Tyagi and D. Dembla, "Investigating the security threats in vehicular ad hoc networks (VANETs): towards security engineering for safer on-road transportation," in *Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 2084–2090, Delhi, India, September 2014.
- [13] K. Ozbay and P. Kachroo, *Incident Management in Intelligent Transportation Systems* Artech House Publishers, Norwood, MA, USA, 1999, https://digitalscholarship.unlv.edu/ece_fac_articles/103.
- [14] Z. Iqbal and M. I. Khan, "Automatic incident detection in smart city using multiple traffic flow parameters via V2X communication," *International Journal of Distributed Sensor Networks*, vol. 14, no. 11, 2018.
- [15] E. Parkany and C. Xie, *A Complete Review of Incident Detection Algorithms and Their Deployment: What Works and What Doesn't*, New England Transportation Consortium, 2005, <http://docs.trb.org/00988875.pdf>.
- [16] P. B. Farradyne, *Traffic Incident Management Handbook*, United States: Prepared for Federal Highway Administration,

- Office of Travel Management, Traffic Incident Management Handbook (<http://hcm2010.org>), 2000.
- [17] L. E. Dougald, V. Ramkumar, and J. G. Noah, *Traffic Incident Management Quick Clearance Guidance and Implications*, http://www.virginiadot.org/vtrc/main/online_reports/pdf/16-r9.pdf, Virginia Transportation Research Council, Charlottesville, VA, USA, 2016, http://www.virginiadot.org/vtrc/main/online_reports/pdf/16-r9.pdf.
 - [18] C. Liu, M. Zhao, A. Sharma, and S. Sarkar, "Traffic dynamics exploration and incident detection using spatiotemporal graphical modeling," *Journal of Big Data Analytics in Transportation*, vol. 1, no. 1, pp. 37–55, 2019.
 - [19] J. Xiao, "SVM and KNN ensemble learning for traffic incident detection," *Physica A: Statistical Mechanics and Its Applications*, vol. 517, pp. 29–35, 2019.
 - [20] L. Li, S. He, J. Zhang, and F. Yang, "Bagging-SVMs algorithm-based traffic incident detection," in *Proceedings of the 16th COTA International Conference of Transportation Professionals*, pp. 1461–1469, Shanghai, China, 2016.
 - [21] M. Taamneh, S. Taamneh, and S. Alkheder, "Clustering-based classification of road traffic accidents using hierarchical clustering and artificial neural networks," *International Journal of Injury Control and Safety Promotion*, vol. 24, no. 3, pp. 388–395, 2017.
 - [22] N. Dogru and A. Subasi, "Traffic accident detection using random forest classification system," in *Proceedings of the IEEE 15th Learning and Technology Conference (LandT)*, pp. 40–45, Jeddah, Saudi Arabia, February 2018.
 - [23] K. Wang, T. Bhowmik, S. Yasmin, S. Zhao, N. Eluru, and E. Jackson, "Multivariate copula temporal modeling of inter-section crash consequence metrics: a joint estimation of injury severity, crash type, vehicle damage and driver error," *Accident Analysis & Prevention*, vol. 125, pp. 188–197, 2019.
 - [24] X. Jianfeng, G. Hongyu, T. Jian, L. Liu, and L. Haizhu, "A classification and recognition model for the severity of road traffic accident," *Advances in Mechanical Engineering*, vol. 11, no. 5, 2019.
 - [25] P. Chakraborty, A. Sharma, and C. Hegde, "Freeway traffic incident detection from cameras: a semi-supervised learning approach," in *Proceedings of the IEEE 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1840–1845, Maui, HI, USA, November 2018.
 - [26] A. Theofilatos, "Incorporating real-time traffic and weather data to explore road accident likelihood and severity in urban arterials," *Journal of Safety Research*, vol. 61, pp. 9–21, 2017.
 - [27] X. Liu, J. Xu, M. Li, and J. Peng, "Sensitivity analysis based SVM application on automatic incident detection of rural road in China," *Mathematical Problems in Engineering*, vol. 2018, Article ID 9583285, 9 pages, 2018.
 - [28] H. Nguyen, C. Cai, and F. Chen, "Automatic classification of traffic incident's severity using machine learning approaches," *IET Intelligent Transport Systems*, vol. 11, no. 10, pp. 615–623, 2017.
 - [29] F. He, X. Yan, Y. Liu, and L. Ma, "A traffic congestion assessment method for urban road networks based on speed performance index," *Procedia Engineering*, vol. 137, pp. 425–433, 2016.
 - [30] A. Zeroual, F. Harrou, Y. Sun, and N. Messai, "Monitoring road traffic congestion using a macroscopic traffic model and a statistical monitoring scheme," *Sustainable Cities and Society*, vol. 35, pp. 494–510, 2017.
 - [31] C. Bian, C. Yuan, C. W. Kuang, and D. Wu, "Evaluation, classification, and influential factors analysis of traffic congestion in Chinese cities using the online map data," *Mathematical Problems in Engineering*, vol. 2016, Article ID 1693729, 10 pages, 2016.
 - [32] X. Xu, X. Gao, X. Zhao, and Z. X. H. Chang, "A novel algorithm for urban traffic congestion detection based on GPS data compression," in *Proceedings of the IEEE International Conference on (SOLI)*, pp. 107–112, Beijing, China, 2016.
 - [33] J. Wan and Y. Y. Q. Wang, "Traffic congestion analysis: a new perspective," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1398–1402, New Orleans, LA, USA, 2017.
 - [34] L. Zhu, F. Guo, R. Krishnanand, and J. W. Polak, "A deep learning approach for traffic incident detection in urban networks," in *Proceedings of the IEEE 21st International Conference on ITSC*, pp. 1011–1016, Maui, HI, USA, 2018.
 - [35] J. Ren, Y. Chen, L. Xin, and J. S. H. Mahama, "Detecting and locating of traffic incidents in a road segment based on lane-changing characteristics," *Transportmetrica A: Transport Science*, vol. 13, no. 10, pp. 853–873, 2018.
 - [36] D. Li, X. Hu, C. J. Jin, and J. Zhou, "Learning to detect traffic incidents from data based on tree augmented naive bayesian classification system," *Discrete Dynamics in Nature and Society*, vol. 2017, Article ID 8523495, 9 pages, 2017.
 - [37] M. Kim and S. Chi, "Detection of centerline crossing in abnormal driving using CapsNet," *The Journal of Supercomputing*, vol. 75, no. 1, pp. 189–196, 2019.
 - [38] S. Al-Sultan, A. H. Al-Bayatti, and H. Zedan, "Context-aware driver behavior detection system in intelligent transportation systems," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4264–4275, 2013.
 - [39] J. Hu, L. Xu, X. He, and W. Meng, "Abnormal driving detection based on normalized driving behavior," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 6645–6652, 2017.
 - [40] K. Saruwatari, F. Sakae, and J. Sato, "Detection of abnormal driving using multiple view geometry in space-time," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 1102–1107, Alcalá de Henares, Spain, 2012.
 - [41] I. Mohamad, M. A. Ali, and M. Ismail, "Abnormal driving detection using real time global positioning system data," in *Proceedings of the IEEE IconSpace*, pp. 1–6, Penang, Malaysia, 2011.
 - [42] A. S. Zeeman and M. J. Booyesen, "Combining speed and acceleration to detect reckless driving in the informal public transport industry," in *Proceedings of the 16th International IEEE Conference on ITSC*, pp. 756–776, The Hague, Netherlands, 2013.
 - [43] K. Sasidhar and A. Upasini, "Two wheeler rash drive detection using smartphone," in *Proceedings of the 11th International COMSNETS*, pp. 754–758, Bengaluru, India, 2019.
 - [44] B. Bose, J. Dutta, S. Ghosh, P. Pramanick, and S. Roy, "Smartphone based system for real-time aggressive driving detection and marking rash driving-prone areas," in *Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and Networking*, Varanasi, India, 2018.
 - [45] J. Tang, J. Liang, C. Han, Z. Li, and H. Huang, "Crash injury severity analysis using a two-layer Stacking framework," *Accident Analysis & Prevention*, vol. 122, pp. 226–238, 2019.
 - [46] Y. W. Liyanage, D. S. Zois, and C. Chelmiss, "Robust freeway accident detection: a two-stage approach," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3387–3391, Brighton, UK, 2019.
 - [47] Y. E. Hawas, M. Sherif, and M. D. Alam, "Optimized multistage Fuzzy-based model for incident detection and management on urban streets," *Fuzzy Sets and Systems*, vol. 381, pp. 78–104, 2020.

- [48] L.-L. Wang, H. Y. T. Ngan, and N. H. C. Yung, "Automatic incident classification for large-scale traffic data by adaptive boosting SVM," *Information Sciences*, vol. 467, pp. 59–73, 2018.
- [49] R. Al Mallah, A. Quintero, and B. Farooq, "Cooperative evaluation of the cause of urban traffic congestion via connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 59–67, 2019.
- [50] A. M. Ranwa, A. Quintero, and B. Farooq, "Distributed classification of urban congestion using VANET," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2435–2442, 2017.
- [51] A. B. Parsa, A. Movahedi, H. Taghipour, S. Derrible, and A. K. Mohammadian, "Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis," *Accident Analysis & Prevention*, vol. 136, Article ID 105405, 2020.
- [52] C. Cheng, P. F. Xu, H. Cheng, Y. Ding, and J. Zheng, "Ensemble learning approach based on stacking for unmanned surface vehicle's dynamics," *Ocean Engineering*, vol. 207, Article ID 107388, 2020.
- [53] S. Moosavi, M. H. Samavatian, S. Parthasarathy, R. Teodorescu, and R. Ramnath, "Accident risk prediction based on heterogeneous sparse data: new dataset and insights," in *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 33–42, Chicago, IL, USA, 2019.
- [54] A. G. Karegowda, A. S. Manjunath, and M. A. Jayaram, "Comparative study of attribute selection using gain ratio and correlation based feature selection," *International Journal of Information Technology and Knowledge Management*, vol. 2, no. 2, pp. 271–277, 2020.
- [55] R. L. M. C. Martinez and J. D. Naranjo, "A pretest for choosing between logrank and wilcoxon tests in the two-sample problem," *Metron*, vol. 68, no. 2, pp. 111–125, 2010.
- [56] S. Hussain, J. Keung, M. K. Sohail et al., "Methodology for the quantification of the effect of patterns and anti-patterns association on the software quality," *IET Software*, vol. 13, no. 5, pp. 414–422, 2019.
- [57] C. J. Wild, M. Pfannkuch, M. Regan, and N. J. Horton, "Towards more accessible conceptions of statistical inference," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 174, no. 2, pp. 247–295, 2011.
- [58] S. Hussain, J. Keung, A. A. Khan et al., "Implications of deep learning for the automation of design patterns organization," *Journal of Parallel and Distributed Computing*, vol. 117, pp. 256–266, 2018.