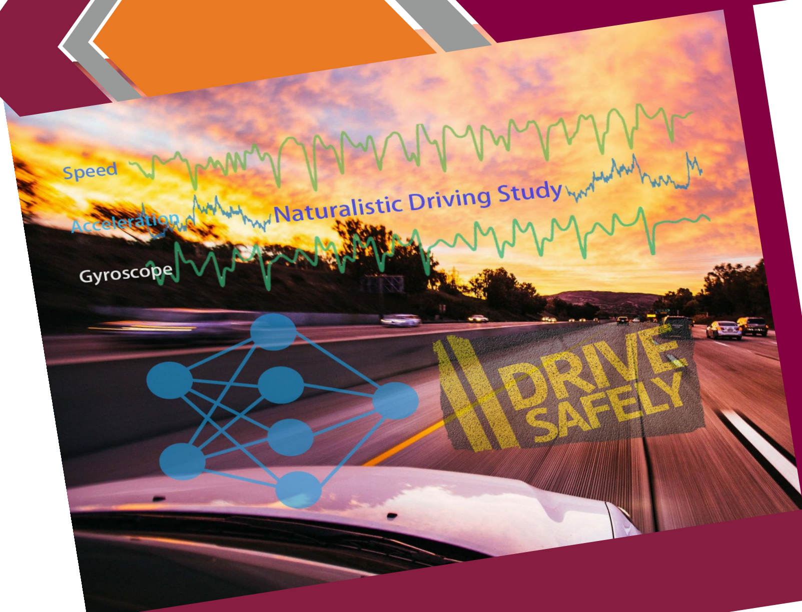
Driving Risk Assessment Based on High-frequency, High-resolution Telematics Data

April 2022 | Final Report



SAFE-D
SAFETY THROUGH DISRUPTION

VT TRANSPORTATION INSTITUTE
VIRGINIA TECH.

Texas A&M Transportation Institute

San Diego State University
Leadership Starts Here

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. VTTI-00-028	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Driving Risk Assessment Based on High-frequency, High-resolution Telematics Data		5. Report Date: April 2022 6. Performing Organization Code:
7. Author(s) Feng Guo (VTTI/VT)* Chen Qian (Student-VTTI/VT) Liang Shi (Student-VTTI/VT)		8. Performing Organization Report No. VTTI-00-028
9. Performing Organization Name and Address: Safe-D National UTC Virginia Polytechnic Institute and State University Virginia Tech Transportation Institute 3500 Transportation Research Plaza Blacksburg, Virginia 24061 USA		10. Work Unit No. 11. Contract or Grant No. 69A3551747115/ VTTI-00-028
12. Sponsoring Agency Name and Address Office of the Assistant Secretary for Research and Technology University Transportation Centers Program Department of Transportation Washington, DC 20590 United States		13. Type of Report and Period Final Research Report 14. Sponsoring Agency Code
15. Supplementary Notes This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program.		

16. Abstract

The emerging connected vehicle and Automated Driving System (ADS), the widely available advanced in-vehicle telematics data collection/transmitting systems, as well as smartphone apps produce gigantic amount of high-frequency, high-resolution driving data. These telematics data provide comprehensive information on driving style, driving environment, road condition, and vehicle conditions. The high frequency telematics data has been used for several safety areas such as insurance pricing, teenage driving risk evaluation, and fleet safety management. This report study advances traffic safety analysis in the follow aspects: 1) characterize the high-frequency kinematic signatures for safety critical events compared to normal operations; and 2) develop models to distinguish and predict crashes from normal driving scenarios based on the high frequency data. Two deep learning models were developed. The first one combines the strength of convolutional neural network (CNN), gated recurrent unit (GRU) network and extreme gradient boosting (XGBoost). The second approach is based on a novel variational inference for extremes (VIE) to address the rarity of crashes. The models proposed in this project can benefit a variety of traffic research and applications including connected vehicles and ADS real-time safety monitoring, NDS data analysis, ride-hailing safety prediction, as well as fleet and driver safety management programs.

17. Key Words High frequency driving data; kinematics driving data; crash prediction; convolutional neural network; gated recurrent unit network; Variational Inference; Connected vehicles.		18. Distribution Statement No restrictions. This document is available to the public through the Safe-D National UTC website, as well as the following repositories: VTechWorks, The National Transportation Library, The Transportation Library, Volpe National Transportation Systems Center, Federal Highway Administration Research Library, and the National Technical Reports Library.	
19. Security Classif. (of this report) Unclassified	20. Security Classify. (of this page) Unclassified	21. No. of Pages 21	22. Price $0

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Abstract

Emerging connected vehicle and Automated Driving Systems (ADS), widely available advanced in-vehicle telematics data collection and transmitting systems, as well as smartphone apps produce gigantic amount of high-frequency, high-resolution driving data. This telematics data provides comprehensive information on driving style, driving environment, road condition, and vehicle conditions. High frequency telematics data has been used in several safety areas such as insurance pricing, teenage driving risk evaluation, and fleet safety management. This report study advances traffic safety analysis in the follow aspects: 1) characterize high-frequency kinematic signatures for safety critical events compared to normal operations; and 2) develop models to distinguish and predict crashes from normal driving scenarios based on high frequency data. Two deep learning models were developed. The first one combines the strength of convolutional neural network (CNN), gated recurrent unit (GRU) network and extreme gradient boosting (XGBoost). The second approach is based on a novel variational inference for extremes (VIE) to address the rarity of crashes. The models proposed in this project can benefit a variety of traffic research and applications including connected vehicles and ADS real-time safety monitoring, naturalistic driving study (NDS) data analysis, hail-driving driver safety prediction, as well as fleet and driver safety management programs.

Acknowledgements

This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program. We thank Jeremy Sudweeks for providing valuable subject matter expert comments and Eric Glenn for support throughout the project.

Table of Contents

List of Figures

List of Tables

Introduction and Background

Emerging transportation technologies, such as connected vehicles (CVs) and automated driving systems (ADS), as well as widely available advanced in-vehicle telematics data collection systems, produce massive amount of high-frequency, high-resolution driving data. For example, Tesla collects a large amount of data on vehicle operation, system status, kinematics, and service history. The OnStar® system by General Motors also includes sensors for an automatic crash response. At a much larger scale, modern smartphones are typically equipped with inertial measurement units (IMUs), which can support apps that also measure vehicle kinematics. Such ubiquitous, high-resolution kinematic driving data provides a unique opportunity for driving risk assessment, crash prediction, and safety countermeasure development.

Crashes have been a gold standard in crash risk evaluation. Although the overall number of crashes is large, around 35,000 per year in the US, crashes are rare events for an individual driver. On average, a driver would experience a fatal crash every 5,000 years and one police-reportable crash every 30 years. The rarity of crashes is especially challenging for continuous driving data as crashes typically only last a few seconds, which is extremely short compared to normal driving. For example, there are only around 2,000 crashes from the Second Strategic Highway Research Plan (SHRP 2) Naturalistic Driving Study (NDS) with more than one million hours of continuous driving data [1].

Although occurring at a shallow frequency, rare events can have far-reaching impacts on society and new technology. For example, fatal crashes are rarely observed for ADS, but even a few fatalities could profoundly reduce public confidence in the new technology and impede its development. Identifying, assessing, and predicting rare events and determining the underlying risk factors can benefit a broad range of applications, such as fraud detection, traffic safety evaluation, and other rare event scenarios. Highly efficient, scalable models are required to identify crashes from the vast amount of normal driving data.

Kinematic driving data, such as speed, three-dimensional acceleration data and yaw rate, reflect critical characteristics of driving behavior and the driving environment. High-frequency, high-resolution telematics data provide crucial information on both long-term driver behavior as well as instantaneous driving conditions. Telematics driving data have been used for individual driver risk evaluation, teenage driving behavior intervention, use-based insurance pricing, and identifying safety-critical events (SCEs) from NDS [2].

One of the most significant developments in the travel service provider sector in the last decade is the surge of ride-hailing services such as Uber and Lyft [3]. Ride-hailing services provide a novel alternative travel mode by connecting travelers and ride-hailing drivers via smartphone apps. Leveraging the rapid development in wireless and smartphone technology, the ride-hailing platform can effectively match the travel demand and supply via sophisticated computational algorithms. The ride-hailing app can conveniently collect kinematic information from sensors on smartphones. The constantly connected app with the cloud server allows such telematics information to be collected from millions of drivers, which could lead to Big Data analysis challenges. The mega-scale telematics data with substantial noise imposes tremendous challenges in the data processing and analysis. A distributed cloud data server, extensive parallel computing system, and advanced analytics methods are needed to implement a smartphone-based safety management system successfully.

Telematics information has been used to evaluate and predict crashes at the driver level, for example, hard braking and swerving events [35]. The premise is that aggressive or unsafe driving behavior could lead to a high number of situations that require an evasive maneuver to prevent crashes from happening. Thus, if a driver can drive conservatively and cautiously to avoid such hazardous situations, fewer cases will require high g-force maneuvers, defined as a maneuver that leads to acceleration or yaw rates higher than a pre-specified threshold value, to avoid crashes. Studies have shown that the rate of high g-force events can be used to predict crash risk. For example, Simons-Morton et al. [4] showed that high g-force events predict crashes for teenage drivers. A previous Safe-D UTC project confirms that kinematic driving signature does affect individual crash risk [3]. In addition, the threshold values used to define high g-force events can vary depending on the objective of a study [35].

Modern CVs and ADS are equipped with full suite of sensors and constantly connected to the network [3]. The high-definition driving data from these systems open up opportunities for diverse applications, including real-time vehicle status monitoring, automated driving algorithms, and provide real-time services for drivers. High-frequency kinematics information is essential for such applications.

However, existing research typically only uses simple features, such as maximum deceleration, for risk prediction and crash detection. Though proven to be associated with crash risk at the driver level, the extreme value-based features ignore a large amount of information from the high-frequency data. Comprehensive feature extraction can help utilize the rich information collected through NDS or CVs. Advanced statistical and machine learning models are required to capture the full information from the rich source data, as well as address methodology challenges such as the rarity of crash events.

This study leverages the rich information available through NDS data collected in the last decade and state-of-the-art machine learning and artificial intelligence (AI) models to bridge the gap between rich data and limitations in application due to relatively simple analytics methods. The study's overall goal is to provide a comprehensive portrait of the telematics characteristics of SCEs and apply them to risk prediction and evaluation. The study's outputs can be used in driving risk assessment, event classification, and real-time risk prediction and detection.

Crashes and SCEs involve impact with other objects or evasive maneuvers, which will be reflected in vehicle kinematics, such as excessive deceleration rate. Using such kinematics information to predict crashes can lead to instantaneous crash prediction and benefit emergency response, large-scale NDS processing, and active safety system development. However, the actual driving data are typically noisy, and the key signatures are not easy to extract. In addition, the duration of SCEs only accounts for an infinitesimal portion of the total driving time. To address these challenges, we propose two alternative approaches. The first approach focuses on feature representation of high-dimensional, high-frequency kinematics data through deep learning models. The second approach focuses on the rarity of crashes.

This project offers two main contributions. First, it proposes a hybrid deep learning approach that consists of feature engineering based on convolutional neural network (CNN) and gated recurrent unit (GRU), as well as classification based on extreme gradient boosting (XGBoost) to predict

SCEs, including crashes and near-crashes (CNCs), from normal braking behavior. A novel time-series classification approach is used for imbalanced data based on kinematic signatures. Second, it formulates a variational representation learning scheme that models disentangled extremes and develops a robust, powerful prediction approach that joins the strength of a generalized additive model and an isotonic neural net.

The rest of the report is organized as follows. Section 2 presents the instantaneous crash risk prediction using deep learning models; Section 3 introduces the methodology for risk prediction using rare event models. Section 4 provides a summary and discussion. Section 5 introduces the web tools for application and dissemination.

Instantaneous Crash Risk Prediction Using Deep Learning Models

Substantial research effort has been conducted on identifying crashes in real time, with environment-based approaches [5-7] and kinematics signal-based approaches [8-11] being the two common alternatives. Environment-based approaches monitor accident-prone areas by predicting risks on specific road sections and providing warning when the risk is high. These methods are typically constrained to specific locations, and complicated, often site-specific, environmental conditions bring challenges in model generalization. The environment-based approaches are less informative for in-vehicle monitoring and applications.

The kinematics signal-based approaches utilize data from the vehicle network or sensors installed on vehicles, e.g., data from NDS and CV technology. By learning from the high-frequency kinematic data from individual vehicles, traffic crashes can be detected or predicted. This approach does not rely on environmental information and is not limited to specific locations.

Although encouraging results have been reported for crash and non-crash detection and prediction [8-10], it is challenging to generalize them to crash, near-crash, normal driving detection [8]. One reason is that crashes are not always associated with robust and recognizable features under the extreme imbalanced condition, as traffic crashes are rare for individual vehicles. To accurately model the key characteristics of crashes, we utilize a weighted categorical cross-entropy loss-function-based CNN-GRU to enhance representation power for traffic crash identification. This data-driven approach can identify key features that were ignored by traditional domain-knowledge-based feature extraction methods. While the weighted categorical cross-entropy loss-function can address the class imbalance issue, it could lead to high class-weight risk, i.e., lower precision and F1-score to each class [12]. To avoid the high class-weight risk, we adopt an XGBoost classifier to predict CNCs based on features extracted by the aforementioned CNN-GRU model. Compared with the neural network, the XGBoost has the advantages of processing the discrete data as input, tuning the hyperparameter easily, and not depending on a large amount of training data [13]. The combination of CNN-GRU and XGBoost led to a powerful crash prediction model with high accuracy and computation efficiency. The deep learning approach with weighted loss-function extracts critical representative features of driving kinematics, especially of crashes. The XGBoost classifier builds an efficient and precise classification model while mitigating the class-weight risk in the deep learning model. The structure of the model is illustrated in Figure 1.

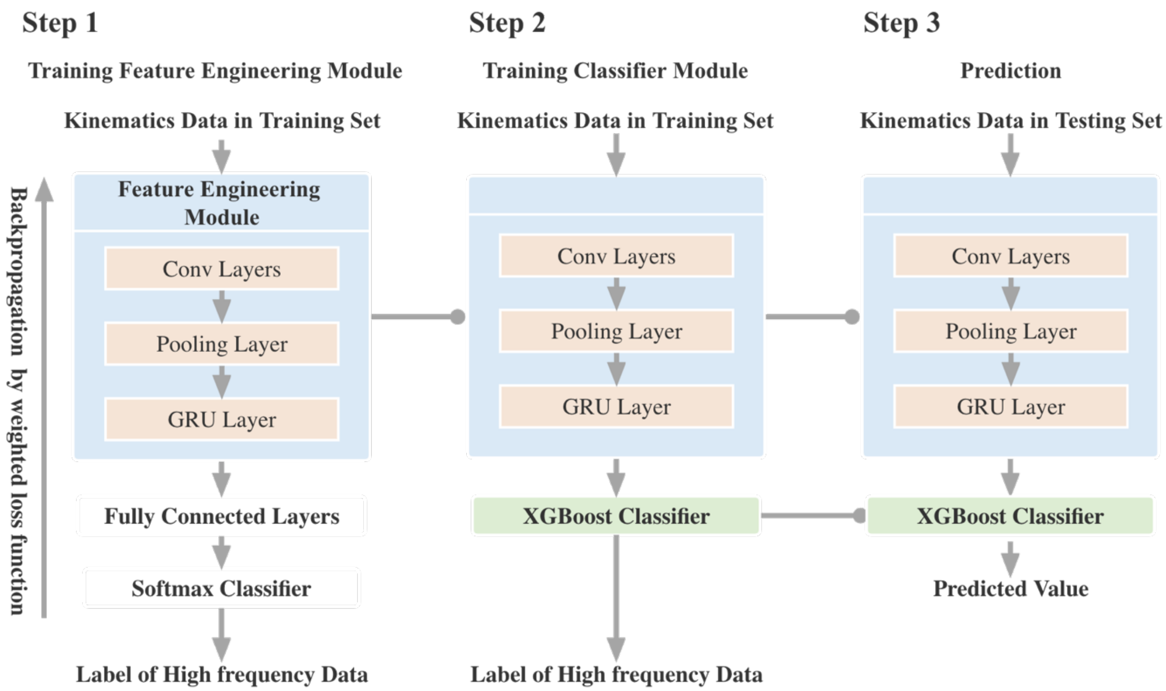


Figure 1. Illustration. Model architecture for instantaneous crash prediction model.

The algorithm includes three steps:

- Step 1 is the feature engineering module, which includes three components, a convolutional network, pooling layers, and a GRU layer. The convolutional network captures local characteristics. The maximum pooling layer integrates adjacent aspects for the convolutional layers. The GRU layer captures temporal patterns of the driving data. The outputs of the CNN-GRU are the features that will be fed to the classification component, which is a fully connected neural network. Due to an imbalance issue, class weight is selected for the cross-entropy loss. The model is fitted by the backpropagation method. The outputs of the GRU layer represent the key characteristics of driving segments and are used for the subsequent classification task.
- Step 2 trains an XGBoost classifier based on outputs of the feature engineering module trained from Step 1. The results show that the classification performance improves significantly over the full connected neural network classifier in Step 1.
- Step 3 builds the final prediction model by integrating the CNN-GRU feature engineering module trained in Step 1 and the XGBoost classifier trained in Step 2.

CNN-GRU Models for Feature Engineering

The feature extraction for the high-frequency driving kinematic data uses the CNN-GRU approach. CNN's ability to characterize spatial information and GRU's efficiency in capturing the temporal structure of local features extracted by CNNs has been extensively demonstrated in the literature [14, 15]. In combination, the CNN-GRU model can extract represented patterns while building prediction models for high-frequency and noisy data [15, 16]. The first set of layers is the primary convolution layers, which are used to identify the local information, such as abrupt changes in driving. The output of the CNN is input into a GRU structure to refine the inherited dependency structure for time-series kinematic information. Figure 2 shows the overall structure of the feature extraction approach. The input results from data processing, a tensor shaped (N, 51, 3). For feature extraction, two 1-dimension convolutional layers are used to catch the relationship between each kernel. A maximum pooling follows the convolutional layers to catch extreme driving behavior changes. The output of this layer is (N, 18, 96). Several GRU layers follow in processing the time-series dependence. The output of GRU (N, 96) is the features, which are fed into the fully connected layers for classification. The fully connected layers are comprised of four dense layers. Batch normalization and dropout methods follow each dense layer to reduce random noise and improve the model's generalization. The last layer is a SoftMax classifier to determine whether the input data are crash, near-crash, or normal driving segments. From grid search, the class weight of crashes, near-crashes, and baselines is set as 30: 7: 1.

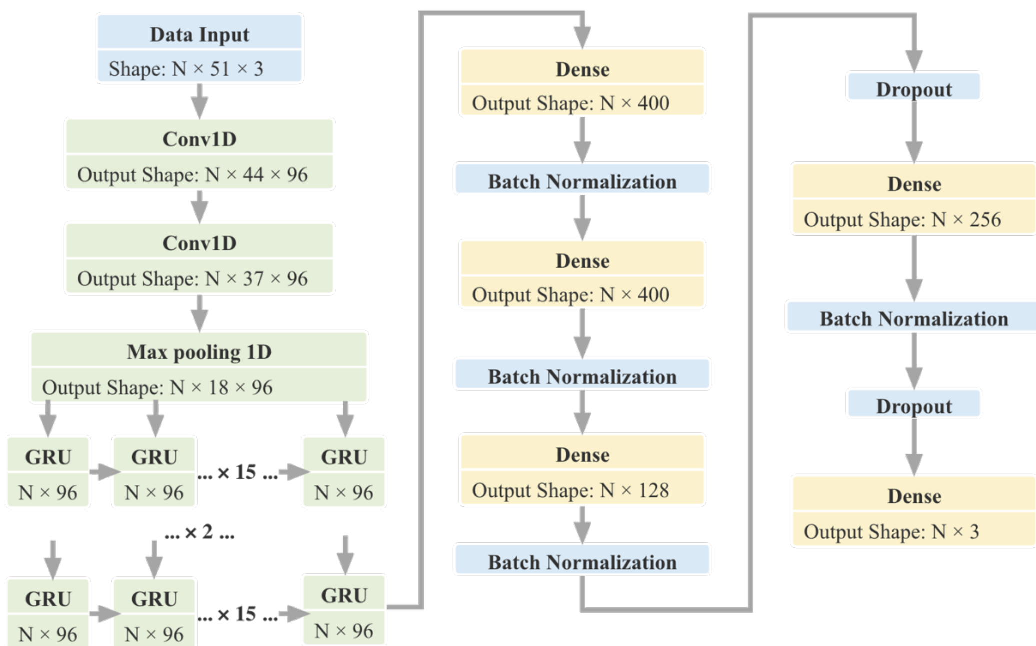


Figure 2. Illustration. Deep learning models for feature engineering.

The feature extraction module described above is trained through a supervised learning process. The outputs are fed to four fully connected neural network layers. A SoftMax classifier is utilized to link with labels (C, CN, and ND).

To address the severe imbalance issue, i.e., small number of crashes compared to normal driving segments, we used the weighted cross-entropy loss function with higher class weights for CNCs. Higher weights improve the model's ability to capture the characteristics of these underrepresented classes. One drawback of the weighted loss function is the associated class-weighted risk reflected in the low precision and F1-scores [12], which is addressed in Step 2.

XGBoost for Prediction

The output features of the CNN-GRU are used for an XGBoost classifier, a supervised machine learning system based on a tree boosting approach [17]. The XGBoost follows the Gradient Boosting framework and has been shown to provide better performance and efficiency than other boosting methods [17]. One of the most important parts of the XGBoost is a regularization term that avoids overfitting. In the proposed model, the XGBoost replaces the fully connected layers in Step 1 to alleviate the class-weighted risk [12] and preserve high precision and F1-scores.

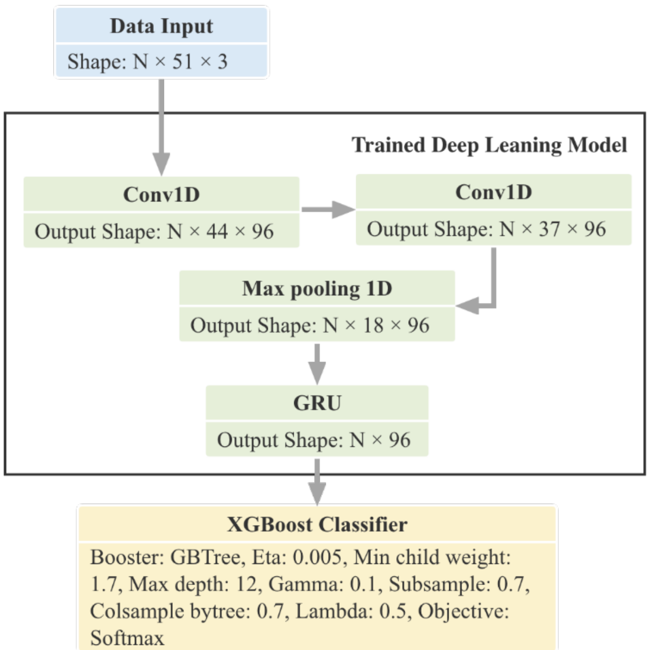


Figure 3. Diagram. XGBoost model for instantaneous crash prediction model.

Figure 3 illustrates the XGBoost module in the overall model structure, as well as its key parameters. The same training data are fed into the trained deep learning model. The output after the GRU layer is the features, a matrix with shape (N, 96). This matrix is fed into an XGBoost classifier; that is, the booster as gradient boosted trees at a 0.005 learning rate; the minimum child weight is 1.7; the maximum depth is 12; the $L1$ regularization parameter (γ) is 0.1; the subsample ratio is 0.7; the subsample ratio of columns when constructing each tree is 0.7; the $L2$ regularization on leaf weights (λ) is 0.5; and the objective function is SoftMax. The outputs of the XGBoost classifier predict whether the inputted driving segments are crash, near-crash, or normal driving segments.

Application of Deep Learning Crash Detection Models

The time-series datasets from the SHRP 2 NDS [26] are used to evaluate the proposed deep learning crash detection approach. The SHRP 2 NDS database contains more than 1,000,000 hours of NDS kinematic time-series data driven by 3,221 drivers over a 5-year period. The dataset includes 1,821 crashes and 6,848 near-crashes. Crashes are divided into four severity levels: most severe crash (Level 1), police-reportable crash (Level 2), minor crash (Level 3), and low-risk tire strike (Level 4).

To prepare the data to be suitable for real applications, the following two aspects are considered. Firstly, speed is a lurking variable between acceleration and the classification result; that is, it is very likely to have a zero speed or substantially lower speed after CNCs. In contrast, if normal driving segments are randomly selected, the beginning and ending speed are likely to be remarkably similar due to the short segment length. To make the model more suitable for real applications, normal driving segments should be selected with similar beginning and ending speeds as SCEs. The speed-matching step will ensure the model's capability to distinguish CNCs from normal stopping or slowing maneuvers. For each CNC, we search the entire driving data for the *same* driver to identify matched normal driving segments. More than 2 million matched normal driving segments have been identified, and the application used 60,000 randomly selected matched normal driving segments.

The length of driving segments in the SHRP 2 NDS dataset varies considerably; e.g., some of the segments last for less than 10 seconds; some last for more than 30 seconds. In comparison, the driving segments containing a CNC are 15.4 seconds on average. To avoid the nuanced information not related to the CNC, proper truncation and alignment to the segments should be applied to capture the most volatile parts.

The turbulence selection module is needed to capture the part with maximum volatility, which corresponds to the time when the CNC occurred. Figure 4 shows the process of turbulence selection. The first step calculates a longitudinal and lateral acceleration resultant, horizontal acceleration. The second step sets a moving standard deviation of the horizontal acceleration and assumes the maximum moving standard deviation is the center of turbulence. Each segment consists of 25 time-series data points (2.5 seconds) before the center of turbulence and 25 data points after. The duration of selected turbulence is 5.1 seconds, mostly representing the volatility after comparing other durations (10.1 and 2.1 seconds).

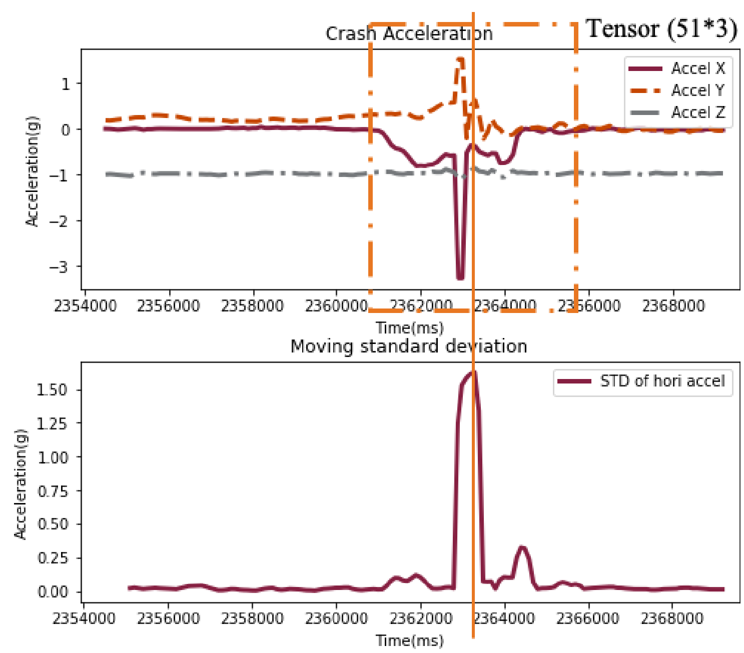


Figure 4. Graphs. Turbulence selection.

The triaxial acceleration (the horizontal direction of the vehicle is positive x) is used as the input to determine whether a specific NDS time-series signal is a crash, near-crash, or normal driving behavior. For N events, the resulting input dataset is a tensor with shape (N, 51, 3) and is the input for the model.

Feature Representation

The performance of the model depends heavily on the research representation, i.e., the outputs of the CNN-GRU model, of the kinematics of driving segments. A good representation should provide clear separation among different types of events. We compared our CNN-GRU automatic feature extraction method with three state-of-the-art domain-knowledge-based feature extraction methods in processing kinematics data, including Taccari et al. [8], Winlaw et al. [9], and Osman et al.[10].

For comparing the representation power of different feature extraction methods, we utilized t-Distributed Stochastic Neighbor Embedding (t-SNE) [27]. The t-SNE method reduces the original multidimensional data to a lower dimension to visualize the representation power of different feature extraction methods. The t-SNE of features extracted by the CNN-GRU method is shown in Figure 5(a), and the t-SNE of features extracted by the methods in Taccari et al. [8], Winlaw et al. [9], and Osman et al. [10] are shown in Figure 5(b),(c), and (d). Different types of events are represented by different colors. CNCs do differ from normal driving behaviors, as shown in each figure. However, many crashes mix up with near-crashes in the benchmarks (Figure 5(b),(c), and (d)), while the CNN-GRU feature extraction method can distinguish most of them (Figure. 5(a)), indicating a superior performance of the proposed model.

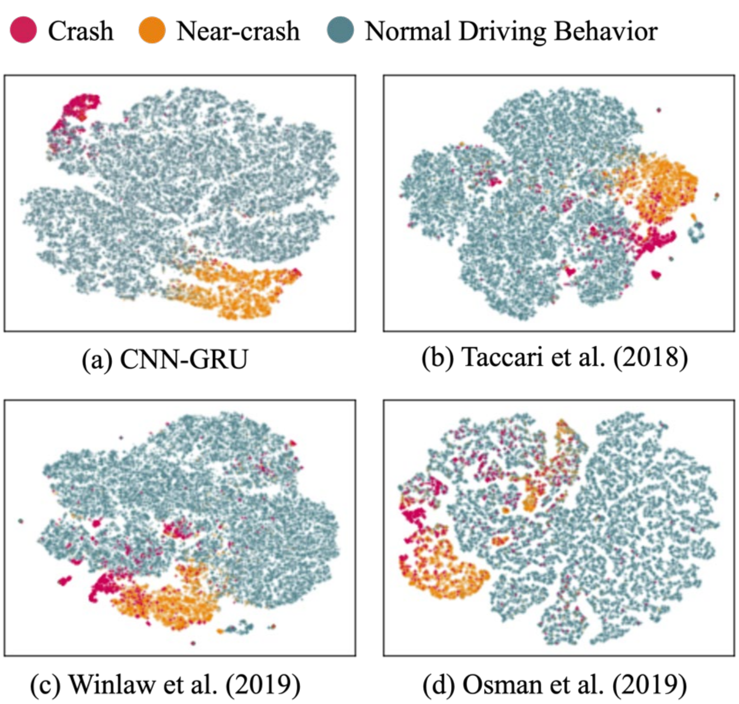


Figure 5. Graphs. t-SNE visualization for features.

Performance Classification

To quantitatively evaluate the model, we use four standard evaluation metrics: accuracy of the model, precision of crashes, recall of crashes, and the F1-score of crashes.

Table 1. Confusion Matrix for 3-class Problem for Crash Detection Model

	Predicted: Crash	**Predicted: Near-crash**	**Predicted: Non-crash**
Actual: Crash	655	108	156
Actual: Near-crash	57	3,069	285
Actual: Non-crash	61	177	29,765

Table 1 shows the confusion matrix of the prediction result on the *test set,* which accounts for about half of the data, containing 773 crashes, 3,354 near-crashes, and 30,206 normal driving epochs. The global accuracy is 97.54%, which means the overwhelming majority of the prediction of events is correct. The recall of crash and near-crash is 71.27% and 89.97% respectively, which indicates that the model can detect the majority of the SCEs.

The precision of crashes and near-crashes is 84.73% and 91.50% respectively, which suggests strong power to distinguish crashes, near-crashes, and normal driving behavior. These evaluation metrics show that even if under a highly imbalanced sample, the model achieves high performance in crash identification.

We compared the model performance with various benchmarks, and the results are shown in Table 2. There is no major difference in the accuracy of the overall model in each method because normal driving behaviors constitute the vast majority of the data. However, compared with the benchmark models, the recall of crashes in the proposed method increases by 5.80%, 15.77%, 40.85%, and 8.57%, respectively, while the F1-score of crashes in the proposed method increases by 5.97%, 7.33%, 23.32%, and 24.01%, respectively, which suggests that the proposed model provides a better overall performance compared to the benchmark models.

Table 2. Performance Classification Comparison for Crashes

Method	Feature Extraction	Classification	Accuracy	Recall	Precision	F1-score
Proposed method	CNN-GRU	XGBoost	**97.54%**	**71.27%**	84.73%	**77.42%**
Arvin et al. [11]	CNN-LSTM	DNN	97.27%	67.36%	81.13%	73.60%
Taccari et al. [8]	Statistical Features	Random Forest	97.37%	61.56%	**87.08%**	72.13%
Winlaw et al. [9]	Statistic Features	Logistic Regression	96.61%	50.96%	81.74%	62.78%
Osman et al. [10]	Statistic Features	Adaboost	96.08%	65.95%	59.27%	62.43%

The severity of crashes varies, and the kinematic signatures of different types of crashes can vary substantially. We examined the model performance by crash severity. As mentioned previously, the SHRP 2 NDS data separates crashes into four groups: most severe, police-reportable, minor, and low-risk tire strike. There are altogether 1,821 crashes, including 101 most severe crashes, 182 police-reportable crashes, 769 minor crashes, and 769 low-risk tire strikes. Table 3 lists the recall rate by crash severity. The recall for most severe crashes is 98.02%, while the recall for minor crashes and low-risk tire strikes is 72.43% and 78.28%, respectively, suggesting considerably better performance for the most severe crashes.

Table 3. Recall by crash severity

Severity	Total crashes	Predicted crashes	Recall rate
Most severe	101	99	98.02%
Police-reportable	182	162	89.01%
Minor crashes	769	557	72.43%
Low-risk tire strike	769	602	78.28%

In summary, the proposed CNN-GRN-XGBoost deep learning approach achieves a good prediction power in distinguishing CNCs from normal driving segments. The model outperforms benchmark models in precision, recall, and F1-score overall. The model can recall most of the severe crashes with 98.02% recall rate.

Rare Event Prediction Models

Severe event class imbalance poses a significant challenge to traffic collision prediction in traffic safety [18] and autonomous vehicle testing [19]. Characterized by severe event class imbalance and the lack of minority labels, rare event modeling falls outside the comfort zone of standard statistical approaches [20]. Without explicit statistical adjustments, the imbalance drives a learning agent to bias towards the majority class; at the same time, the absence of adequate minority examples causes unprotected models to capture spurious features that do not generalize well.

The existing solutions mainly focus on differential sampling or sample reweighting approaches to alleviate the imbalance issue. However, these statistical sample adjustments, such as resampling and reweighting, do not handle inputs outside the norm well. Resampling typically involves either over-sampling or under-sampling [21]. The former is often subjected to the loss of estimation efficiency [20], and the latter is often associated with compromised generalization [22]. The reweighting schemes are often criticized for their numerical instability [23]. Recent works by Cao et al. [22] and Lin et al. [24] modify the hinge loss and entropy loss to capitalize the minority class prediction [23], mentioning that resampling and reweighting may cause overfitting issues due to training bias and label noise, which leads to inferior performance in generalization or fit on unseen datasets.

To address the rare event issue, we take a novel alternative view on promoting generalization. Our proposal is formulated under the generative Bayesian framework, positing that predictors are the stochastic proxies of latent causes, whose exceedance leads to extreme events. Our solution adopts a generalized Pareto distribution as prior and is modeled with variational inference to capture the extended tail accurately. We devised a disentangled additive monotonic neural architecture to predict the risk by the assertion that exceedance leads to extremes. Our model acknowledges representation uncertainties while at the same time embracing improved interpretability, generalization, and robustness. The proposed model was applied to the SHRP 2 NDS for predicting crashes, near-crashes, and normal driving segments.

Variational Inference of Extremals for Rare Event models

The key idea of the variation inference of extremals (VIE) is to amortize the difficulty of direct prediction of rare events to the representation learning stage. The premise is that extreme latent representations lead to extreme events. The approach partitions the representation space into normal and extreme regions, where the prevalence of target events in the latter far exceeds those in the former. This alleviates the majority bias issue that afflicts conventional schemes, as the event distribution is more balanced for the extreme region, which lends better statistical efficiency when learning from the predictors. The VIE model is composed of two parts, an encoder and decoder.

For supervised learning tasks with rare event modeling, the core of the generative Bayesian model is to reconstruct the conditional probability of data. Optimizing over the conditional probability is challenging because of the computational intractability of the integration, as well as for high dimensional predictors. To address this issue, the conditional Variational Auto Encoder based on the Multilayer Perceptron (MLP) utilizes variational inference to transform the intractable integration in equation into optimization over the evidence lower bound (ELBO). Since we are primarily interested in the modeling of events instead of reconstructing input features, the following conditional ELBO is the main training objective function

$$p_\theta(y \mid x) \geq E_{q_\phi(z|x)}[\log p_\theta(y \mid z)] - \mathcal{D}_{\mathrm{KL}}\left[q_\phi(z \mid x) \,||\, p_\theta^{\mathrm{EP}}(z)\right] \triangleq \ell_{\mathrm{VIE}}(\theta, \phi),$$

As there is no closed form solution for the Kullback-Leibler (KL) divergence, the Monte Carlo method is used to estimate the empirical KL divergence with:

$$\mathcal{D}_{\mathrm{KL}}\left[q_\phi(z \mid x) \,||\, p_\theta(z)\right] = E_{q_\phi(z|x)}\left[\log \frac{q_\phi(\boldsymbol{z}|\boldsymbol{x})}{p_\theta(\boldsymbol{z})}\right] \approx \frac{1}{L}\sum_{l=1}^{L} \log\left(\frac{q_\phi\left(z^{(l)}|\boldsymbol{x}\right)}{p_\theta\left(z^{(l)}\right)}\right),$$

For predictions with an input vector **x**, the marginal likelihood can be approximated by the following Monte Carlo estimator:

$$p_\theta(y \mid \boldsymbol{x}) = \int q_\phi(\boldsymbol{z} \mid \boldsymbol{x}) p_\theta(y \mid \boldsymbol{z}) d\,\boldsymbol{z} = \frac{1}{L}\sum_{l=1}^{L} p_\theta\left(y \mid z^{(l)}\right) \quad \text{where} \quad z^{(l)} \sim q_\phi(\boldsymbol{z} \mid \boldsymbol{x}),$$

To enable the modeling of extreme representations, we extend standard variational inference via incorporating a generalized Pareto-based prior to accommodate heavy tails. For a scalar latent variable z, let μ be the cutting point for being "extreme." We define the extreme prior (EP) as a weighted mixture of a generalized Pareto distribution and standard Gaussian distribution to guarantee its continuity:

$$p_\theta^{\mathrm{EP}}(z) = \frac{w}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) + \mathcal{I}(z \geq \mu) \times \frac{1-w}{\sigma}\left(1 + \xi\frac{z-\mu}{\sigma}\right)^{\left(-\frac{1}{\xi}-1\right)},$$

where w is the accumulated mass of the normal prior at the threshold. To induce more flexibility into the model, all the parameters in the EP are learnable. The purpose of this prior distribution is to insert a heavy-tailed part to represent how those rare events deviate from the bulk. If the tailed behavior does not actually lead to "rarity," the flexibility of the prior will fall back to a standard Gaussian distribution. Generalization to multi-dimensional latent components can be achieved by applying the above setup to each individual dimension, although the parameters need not be shared.

Next, to facilitate interpretability and generalization for rare event modeling, it is beneficial to impose constraints on the decoder part. In this paper, we combine two prominent techniques: the generalized additive model and isotonic regression. Let y come from a Bernoulli distribution and we focus on modelling the conditional event probability. We model this under the following generalized additive models:

$$\psi_\theta(y \mid z) = \log\left(\frac{p}{1-p}\right) = \sum_i f_j(z_j), \quad \text{s.t.} \quad f_j(z_j) = \alpha_j + \beta_j \int_{s_0}^{z_j} \exp\left(h_j(t;\theta)\right) dt,$$

The generalized additive model uses the linear combination of functions for every latent variable to model the generation of y. Each function f_j takes input from the j-th dimension of the latent, and this additive decomposition encourages disentangled representation. It is easy to see that all f_j are monotonic, with the direction dictated by the sign of β_j. We model $h_j\,(t{:}\,\theta)$ above using deep neural networks, and compute $f_j(z_j)$ via Clenshaw-Curtis numerical integration.

To sum up, the proposed VIE framework brings together the strength of variational inference, extreme value theory, and isotonic regression to advance rare event modeling. The model provides enhanced representation of rare events and improve the capability to detect rare events. The overall VIE model structure is shown in Figure 6.

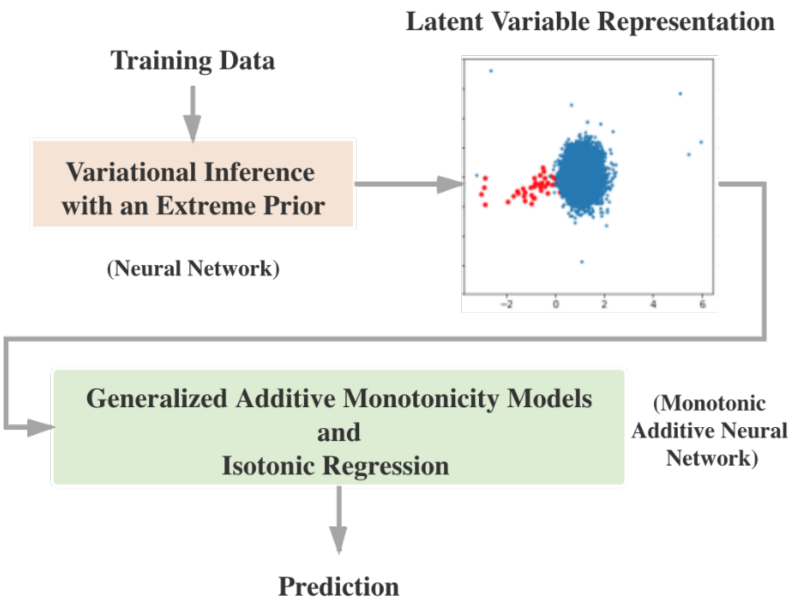


Figure 6. Diagram. VIE for rare event modeling model structure.

Application of VIE Models

The VIE model was also applied to the SHRP 2 NDS. To fully capture discriminative characteristics for normal driving segments and traffic crashes, we employed seven commonly used features in transportation safety research for *x*-, *y*-, and *z*-dimensional acceleration separately, which led to a total of 21 features to represent the acceleration dynamics within every driving segment. The utilized features for prediction models include:

- SD: standard deviation of the acceleration, representing the driving variability [28];
- RV: realized volatility [29] of the acceleration;
- MAX: maximum of the acceleration, typically referring to hard braking or acceleration in driving; known as the elevated gravitational-force event in Morton et al. [4];
- KT for kurtosis of the acceleration, SK for skewness, AC for autocorrelation, and UQ for 75-percent quantile. For a more detailed discussion, please refer to Shi et al. [30].

For comparison, the following baseline deep learning approaches are used: (1) Sampling MLP: a resampling scheme based on multi-layer perceptron over-samples from the less-represented class [21]; (2) LDAM-HG: a reweighting-based loss that adapts the hinge loss accounting for the imbalance[22]; (3) MAML: a few-shot learning scheme [31]; and (4) Deep-SVDD: a one-class classification-based anomaly detection approach [32]. In addition, we also considered several widely available statistical approaches: (5) GBDT: a tree-based classifier [33] where we set the maximum depth to be 10, preventing it from over-fitting; and (6) Elastic Net: logistic regression that alters the class weight to adapt to an imbalanced ratio.

A 5/5 split was adopted for training and testing, as well as a fixed training-validation split ratio of 8/2 for all state-of-the-art benchmarks' hyperparameter selections. As indicated by existing empirical studies [34-36], analytical modeling is sensitive to kinematic noise (such as signal connection issues or different driving environments), where model performance can largely deteriorate as noise increases [28]. To develop a robust model, we created an artificially noisy setting where Gaussian noise N (0; 2) was injected into the input features. A model trained using this added-noise data will be more resistant to noise.

To validate the effectiveness of the proposed VIE framework, we compared the VIE against popular baselines on various real-world applications. The applications include identifying traffic crashes based on kinematic IMU acceleration, identifying fraudulent credit card transactions, and news detection. We conducted the experiments via Pytorch on an NVIDIA V100 GPU cluster. To ensure fair comparisons across different models, we set bottleneck dimensions to 2; i.e., the number of latent, in all experiments, with 1000*1000*2 fully connected layers for encoder and 1000*1000*2 for the unconstrained monotonic neural network (UMNN)in recognizing the decoder part. We used Leaky ReLU for the activation function throughout.

The interpretations for the analytical models were also evaluated. Figure 7 shows the learned representations for different models, where the VIE model shows a remarkable distribution tail and a clear discriminative boundary between "normality" and "events." Elastic Net logistic regression and logistic generalized additive models have no latent variables to consider, so we used the linear principle component analysis (PCA) to approximate its two-dimensional latent plot.

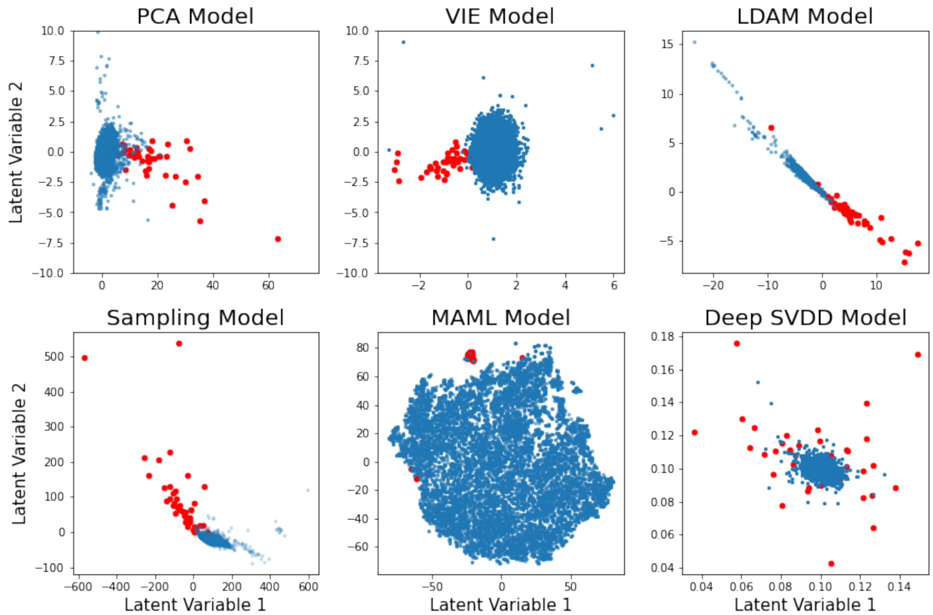


Figure 7. Graphs. SHRP 2 learned representations in testing set for different models – red represents crash events and blue represents non-crash events.

For fair and comprehensive comparison among different models, we used several classification criteria to evaluate model performance. As reliability is essential for decision making, we used negative log likelihood (NLL for short) and expected calibration error (ECE) to measure uncertainty generalization under an imbalanced testing setup, which describes the reliability of decision making based on the established predictive models. The ECE evaluates the relationship between model predicted probabilities and empirical accuracy. Another group of evaluation metrics measured the discriminative power. We adopted the area under the curve of the receiver operating characteristic curve (ROC-AUC; ROC for short) and precision recall AUC (PR-AUC for short). We also used the F1-score to evaluate the models' precision and recall performance. The cutoff points for decision are discussed in each experiment.

The model performance metrics are presented in Table 4 and Table 5. The proposed VIE outperforms alternative benchmark models based on the classification evaluation metrics on all metrics. Importantly, the VIE shows robust discriminant power under noisy testing conditions.

Table 4. Traffic Crash Identification Model Comparison – Original

Method	NLL	ECE	ROC	PR-AUC	F1-Score
GBDT	1.98×10^{-2}	1.45×10^{-3}	0.91	0.80	0.80
MAML	1.26×10^{-2}	1.09×10^{-2}	0.94	0.86	0.85
ELASTIC NET	1.98×10^{-2}	4.02×10^{-3}	0.99	0.83	0.67
LDAM-HG	1.74×10^{-1}	1.25×10^{-2}	0.92	0.83	0.86
DEEP-SVDD	1.82×10^{-2}	3.00×10^{-3}	0.93	0.43	0.08
SAMPLING MLP	2.49×10^{-2}	6.65×10^{-1}	0.91	0.90	0.86
VIE	3.44×10^{-3}	1.30×10^{-3}	**0.99**	**0.93**	**0.89**

Table 5. Traffic Crash Identification Model Comparison – Noise-infused Data

Method	NLL	ECE	ROC	PR-AUC	F1-Score
GBDT	4.30×10^{-1}	3.65×10^{-2}	0.88	0.43	0.13
MAML	5.05×10^{-2}	1.28×10^{-1}	0.92	0.50	0.57
ELASTIC NET	5.93×10^{-1}	1.04×10^{-1}	0.91	0.56	0.04
LDAM-HG	3.50×10^{-2}	4.72×10^{-3}	0.89	0.54	0.40
DEEP-SVDD	1.95×10^{-2}	1.62×10^{-3}	0.81	0.24	0.08
SAMPLING MLP	7.54×10^{-2}	5.27×10^{-1}	0.78	0.56	0.56
VIE	3.20×10^{-2}	2.23×10^{-3}	**0.92**	**0.71**	**0.63**

Figure 8 plots the marginal relationship between the latent variable and associated marginal risk. We can see that Latent Variable 1 contributes significantly and positively to event risk prediction. This indicates that the major components of Latent Variable 1 are x_{RV}, x_{KT}, y_{RV}, z_{MAX}, x_{UQ}, and z_{KT}, where (x, y, z) represents the three-dimensional acceleration. This finding suggests that x_{RV}, x_{KT}, y_{RV}, z_{MAX}, x_{UQ}, and z_{KT} contribute the most for distinguishing between traffic crashes and normal driving segments based on IMU acceleration information. This finding is consistent with various empirical results in the traffic safety research domain [4, 10, 30].

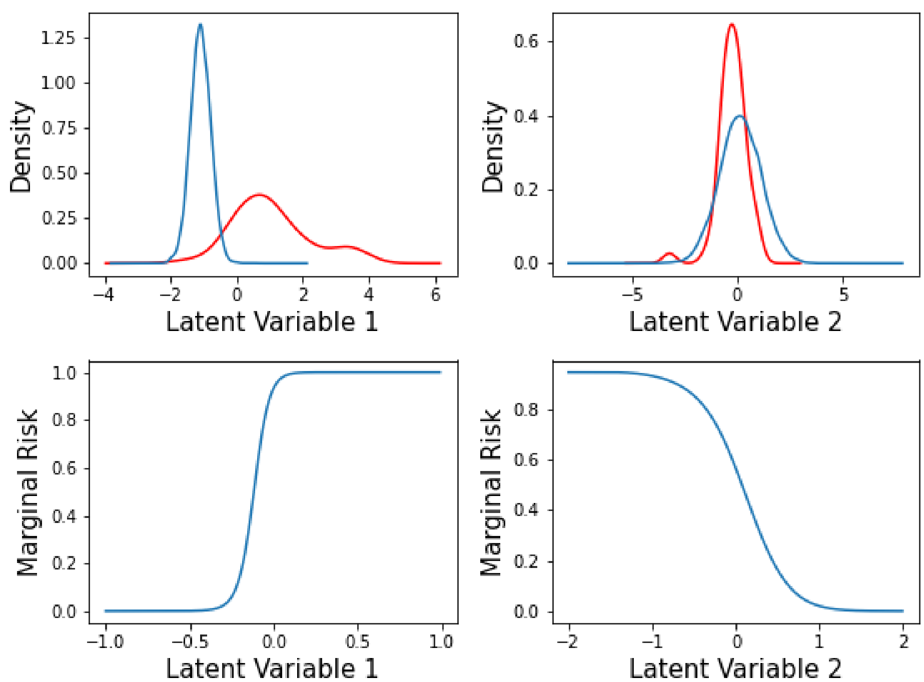


Figure 8. Graphs. SHRP 2 marginal relationship of latent variable and risk – red represents crash.

Summary and Discussions

This project focuses on instantaneous crash risk prediction for a short time window with high-frequency, high-resolution kinematic driving data. The instantaneous risk will provide critical information for current and future transportation systems for the development of safety countermeasures. This is especially useful when CVs, ADS, and smartphones make kinematics data almost universally available. There are, however, numerous challenges in utilizing such data for risk prediction and assessment, among them the Big Data, large noise, and rarity of crashes. There is an imperative need for innovative methodologies to meet these challenges.

This project developed two approaches based on state-of-the-art deep learning and statistical methods. The first approach uses a combination of multiple deep learning methods, including CNN, GRU, and XGBoost, to fit the unique demand of driving risk prediction. The CNN-GRU component concerns how to extract distinct features that represent the key characteristics of a driving segment, which is essential for predicting crashes from normal driving. This powerful combination can incorporate the crucial property of high-frequency kinematic driving data, i.e., local and time series correlation. To address the rarity of crashes, we adopted a weighted cross-entropy loss function for feature engineering in combination with an XGBoost classifier to relieve the class-weighted risk. This sophisticated modeling framework is tailored for complex, noisy driving data while maintaining high computational efficiency.

We applied the proposed deep learning risk prediction model to the SHRP 2 NDS, the largest naturalistic driving study to date with more than one million hours of continuous driving data. The results showed a clear separation of crashes and normal driving segments (with matched initial and ending speed) in the latent representation space. Quantitative metrics show promising results with high precision, recall and overall F1-score. For both latent representation and predictive assessment metrics, the proposed model outperforms the popular benchmark models.

The second approach starts from the rarity of crashes. We propose a VIE approach for rare event prediction. The VIE model induces uncertainty to representation learning through a generalized Pareto distribution, which can accommodate rare events through its long distribution tails. Another common issue with rare event models is overfitting due to the small number of events. We address this issue through generalized additive monotonicity, which leads to a more robust model for rare events. We provide theoretical properties on the efficacy of the proposed model. Extensive empirical experiments were conducted to confirm the model performance over various application scenarios with promising results, especially with respect to model generalization and interpretability.

The VIE rare event model was also applied to the SHRP 2 NDS data. The proposed model universally outperforms several state-of-the-art benchmark models. The approach proposed in this paper can positively impact the research field and society in several regards. The VIE framework dramatically improves the generalizability and interpretability of rare event modeling, two challenging issues associated with limited events. The application of VIE could accurately depict the risk associated with rare events and allow the general public and decision-makers to set realistic expectations for rare events. Identifying adverse events at an early stage may allow mitigation of the damage and loss associated with the events. The features identified through the model are crucial for researchers and practitioners to identify the causes of rare events and take proper countermeasures to prevent and reduce the occurrence of future adverse events.

The two novel modeling frameworks proposed in this study allow prediction of instantaneous crash risk based on high-frequency kinematic driving data, which are widely available through smartphones, CV technology, and ADS. There is a wide range of applications in safety countermeasures, emergency response, naturalistic driving data processing, and insurance.

Education and Workforce Development and Technology Transfer

Education and Workforce Development and Technology Transfer is an essential part of the project. For the education and workforce development, we supported two Ph.D. students, and the research outputs and methodology were presented in graduate- and undergraduate-level courses. For technology transfer and dissemination, we completed two manuscripts for submission to peer-reviewed journals for publication. We presented the results of the project at a public webinar. A webtool was developed for dissemination and technology transfer.

The products created as part of this project have been or will be located on the project page of the Safe-D website.[1] The available datasets resulting from the final project have been or will be located in the Safe-D Collection of the VTTI Dataverse.

Education and Workforce Development Products

- The project supported two Ph.D. students, Liang Shi and Chen Qian.
- The project contents have been used in multiple courses:
    - Stat Epi. & Obs. Studies (STAT 5374), Spring 2020
    - Applied Multivariate Analysis (STAT5504G/STAT4504), Spring 2021
- The project was presented in a webinar, "Driving Risk Assessment Based on High-frequency, High-resolution Telematics Data," on August 26, 2021.

Technology Transfer Products

We have drafted two papers based on the results of the project.

- "C'est La VIE: Variational Inference of Extremal for Rare Event Modeling," submitted to *Annals of Applied Statistics*
- "Real-Time Traffic Safety Critical Events Prediction using Deep Learning Models"; draft manuscript completed.

The methods developed in this study require a high level of technical knowledge to apply appropriately. To facilitate the dissemination and technology transfer, we developed a publicly available web tool to allow users to observe the input data and outputs of the deep learning models. The web tool also allows users to upload their own data for evaluation.

The web tool was developed using the MySQL + Node.js. This system enables users to find the high-risk driving behavior and the corresponding probability. Users can upload their own data to fit our model and get the result or use our sample data. This can be the basis for risk management, which can improve traffic safety. The web tool can be accessed from the following link: https://utc-deep-prediction.cloud.vtti.vt.edu/

[1] https://safed.vtti.vt.edu/projects/driving-risk-assessment-based-on-high-frequency-high-resolution-telematics-data/

Figure 9 shows an interface of the web tool with the following main components:

1. The true label of the event.
2. The result of current sensor data detection.
3. Menu for selecting a sample event. The web tool includes five crashes, five near-crashes, and five normal driving behavior segments from the SHRP 2 NDS.
4. Adjust the vertical acceleration (to remove one gravity unit).
5. Visualization of the results, including
    a. 3-dimensional acceleration represented by lines;
    b. Probability of crash represented by red area;
    c. Probability of near-crash represented by yellow area.

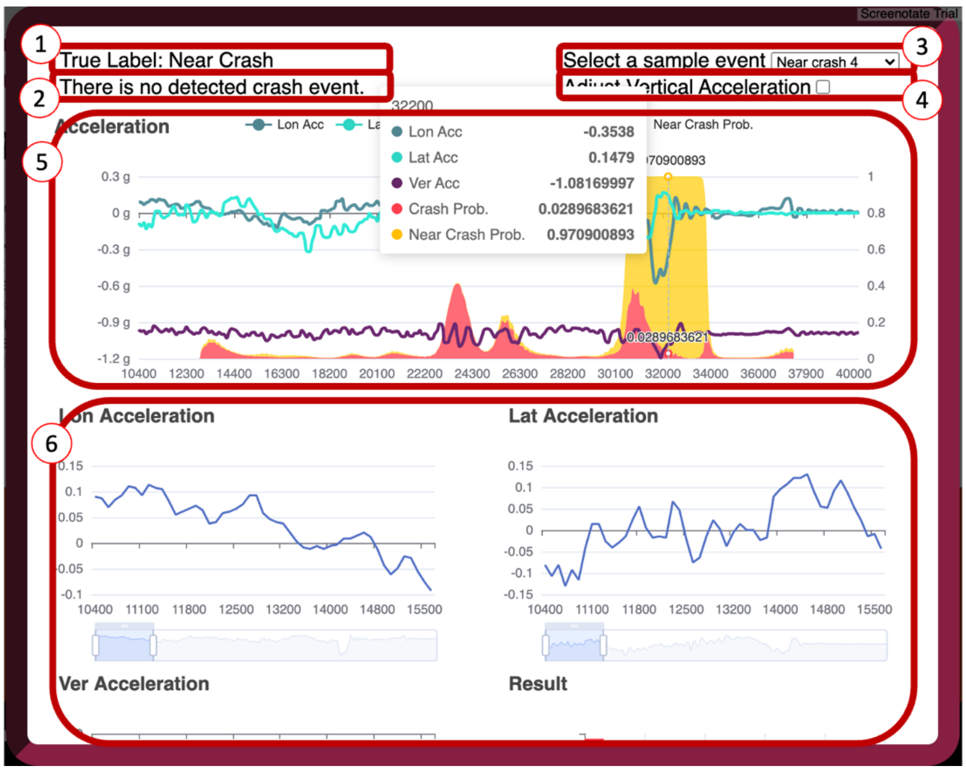


Figure 9. Screen capture. Web tool for risk prediction.

Data Products

The SHRP 2 dataset was used to evaluate the performance of the proposed time-series data classification approach. This database contains more than 1,000,000 hours of continuous naturalistic driving data driven by 3,221 drivers under daily normal driving conditions without specific instructions. A sophisticated data acquisition system was installed in each participant's own vehicle, which continuously collected key driving data, including four camera views, GPS, three-dimensional acceleration, yaw rates, etc.

The research team extracted tens of thousands of normal driving conditions with matching initial and ending speed as the crashes and near-crashes. These data will be saved in data repositories for future use.

References

1. Antin, J.F., et al., *Second strategic highway research program naturalistic driving study methods.* Safety Science, 2019. **119**: p. 2-10.
2. Mao, H., et al., *Decision-adjusted driver risk predictive models using kinematics information.* Accident Analysis & Prevention, 2021. **156**: p. 106088.
3. Mao, H., et al., *Driving safety assessment for ride-hailing drivers.* Accident Analysis & Prevention, 2021. **149**: p. 105574.
4. Simons-Morton, B.G., et al., *Do elevated gravitational-force events while driving predict crashes and near crashes?* American journal of epidemiology, 2012. **175**(10): p. 1075-1079.
5. Yuan, J., et al., *Real-time crash risk prediction using long short-term memory recurrent neural network.* Transportation research record, 2019. **2673**(4): p. 314-326.
6. Huang, X., et al., *Intelligent intersection: Two-stream convolutional networks for real-time near-accident detection in traffic video.* ACM Transactions on Spatial Algorithms and Systems (TSAS), 2020. **6**(2): p. 1-28.
7. Li, P., M. Abdel-Aty, and J. Yuan, *Real-time crash risk prediction on arterials based on LSTM-CNN.* Accident Analysis & Prevention, 2020. **135**: p. 105371.
8. Taccari, L., et al. *Classification of crash and near-crash events from dashcam videos and telematics*. in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 2018. IEEE.
9. Winlaw, M., et al., *Using telematics data to find risky driver behaviour.* Accident Analysis & Prevention, 2019. **131**: p. 131-136.
10. Osman, O.A., et al., *Prediction of near-crashes from observed vehicle kinematics using machine learning.* Transportation Research Record, 2019. **2673**(12): p. 463-473.
11. Arvin, R., A.J. Khattak, and H. Qi, *Safety critical event prediction through unified analysis of driver and vehicle volatilities: application of deep learning methods.* Accident Analysis & Prevention, 2021. **151**: p. 105949.
12. Xu, Z., et al. *Class-weighted classification: Trade-offs and robust approaches*. in *International Conference on Machine Learning*. 2020. PMLR.
13. Sun, Y., et al., *Surrogate-assisted evolutionary deep learning using an end-to-end random forest-based performance predictor.* IEEE Transactions on Evolutionary Computation, 2019. **24**(2): p. 350-364.
14. Krizhevsky, A., I. Sutskever, and G.E. Hinton, *Imagenet classification with deep convolutional neural networks.* Advances in neural information processing systems, 2012. **25**.
15. Bengio, Y., A. Courville, and P. Vincent, *Representation learning: A review and new perspectives.* IEEE transactions on pattern analysis and machine intelligence, 2013. **35**(8): p. 1798-1828.
16. Tian, C., et al., *Enhanced CNN for image denoising.* CAAI Transactions on Intelligence Technology, 2019. **4**(1): p. 17-23.
17. Chen, T. and C. Guestrin. *Xgboost: A scalable tree boosting system*. in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016.

18. Dingus, T.A., et al., *Driver crash risk factors and prevalence evaluation using naturalistic driving data.* Proceedings of the National Academy of Sciences, 2016. **113**(10): p. 2636-2641.
19. O'Kelly, M., et al., *Scalable end-to-end autonomous vehicle testing via rare-event simulation.* Advances in neural information processing systems, 2018. **31**.
20. Wang, H. *Logistic regression for massive data with rare events*. in *International Conference on Machine Learning*. 2020. PMLR.
21. Byrd, J. and Z. Lipton. *What is the effect of importance weighting in deep learning?* in *International Conference on Machine Learning*. 2019. PMLR.
22. Cao, K., et al., *Learning imbalanced datasets with label-distribution-aware margin loss.* Advances in neural information processing systems, 2019. **32**.
23. Ren, M., et al. *Learning to reweight examples for robust deep learning*. in *International conference on machine learning*. 2018. PMLR.
24. Lin, T.-Y., et al. *Focal loss for dense object detection*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
25. Hastie, T.J. and R.J. Tibshirani, *Generalized additive models*. 2017: Routledge.
26. Hankey, J.M., M.A. Perez, and J.A. McClafferty, *Description of the SHRP 2 naturalistic database and the crash, near-crash, and baseline data sets*. 2016, Virginia Tech Transportation Institute.
27. Belkina, A.C., et al., *Automated optimized parameters for T-distributed stochastic neighbor embedding improve visualization and analysis of large datasets.* Nature communications, 2019. **10**(1): p. 1-12.
28. Guo, F., *Statistical methods for naturalistic driving studies.* Annual review of statistics and its application, 2019. **6**: p. 309-328.
29. Andersen, T.G., et al., *Modeling and forecasting realized volatility.* Econometrica, 2003. **71**(2): p. 579-625.
30. Shi, X., et al., *A feature learning approach based on XGBoost for driving assessment and risk prediction.* Accident Analysis & Prevention, 2019. **129**: p. 170-179.
31. Finn, C., P. Abbeel, and S. Levine. *Model-agnostic meta-learning for fast adaptation of deep networks*. in *International conference on machine learning*. 2017. PMLR.
32. Ruff, L., et al. *Deep one-class classification*. in *International conference on machine learning*. 2018. PMLR.
33. Friedman, J.H., *Greedy function approximation: a gradient boosting machine.* Annals of statistics, 2001: p. 1189-1232.
34. Perez, M.A., et al., *Performance of basic kinematic thresholds in the identification of crash and near-crash events within naturalistic driving data.* Accident Analysis & Prevention, 2017. **103**: p. 10-19.
35. Kluger, R., et al., *Identification of safety-critical events using kinematic vehicle data and the discrete fourier transform.* Accident Analysis & Prevention, 2016. **96**: p. 162-168.
36. Johnson, T., R. Sherony, and H.C. Gabler. *Driver lane keeping behavior in normal driving using 100-car naturalistic driving study data*. in *2016 IEEE Intelligent Vehicles Symposium (IV)*. 2016. IEEE.
37. Wali, B. and A.J. Khattak, *Harnessing ambient sensing & naturalistic driving systems to understand links between driving volatility and crash propensity in school zones–A generalized hierarchical mixed logit framework.* Transportation research part C: emerging technologies, 2020. **114**: p. 405-424.