Received: 11 March 2021 | Accepted: 5 July 2022

DOI: 10.1111/poms.13816

POMS

ORIGINAL ARTICLE

Is it all about you or your driving? Designing IoT-enabled risk assessments

Yi-Jen (Ian) Ho[1] | Siyuan Liu[1] | Jingchuan Pu[2] | Dian Zhang[3]

[1]Smeal College of Business, Pennsylvania State University, University Park, Pennsylvania, USA

[2]Warrington College of Business, University of Florida, Gainesville, Florida, USA

[3]College of Computer Science and Software Engineering, Shenzhen University, Guangdong, China

Correspondence
Yi-Jen (Ian) Ho, Smeal College of Business, Pennsylvania State University, University Park, PA, USA.
Email: ian.ho@psu.edu

Handling Editor: Vijay Mookerjee

Abstract

Technological applications disrupt the way to assess risks in the auto-insurance business. Contrasted with the common practice based on static demographics, usage-based insurance predicts risks using driving data collected from Internet-of-things–enabled telematics. This study proposes a novel solution leveraging the synergy between big data and hierarchical modeling. We specifically consider two aspects of mobility, namely, *trait* and *trajectory*, monitored by global positioning system (GPS), on-board diagnostics, and in-vehicle cameras in real time. *Traits* here refer to drivers' distinctive driving behaviors (styles), whereas *trajectories* consist of the vehicle motion sequences and the contextual factors on trips. We operationalize semantic features of the two to assess risks at both trip and driver levels. Using fine-granular driving data and crash reports, we find that behavioral traits play a significant role in predicting crashes, given individual heterogeneity and temporal dynamics. In a series of empirical validations, the proposed solution outperforms the current practice and alternative predictive models considered by prior literature. We show that the mobility-based models are superior to the demographic-based ones. Moreover, our model achieves the comparable performance of neural networks, improving the recall of class-weighted logistic regression, nested support vector machine, and cost-sensitive random forests by 44.23%, 29.18%, and 24.59%, respectively. Last, our approach is robust, data independent, and computationally efficient for skewed and small samples. This study provides several managerial implications and a blueprint for the auto-insurance industry to operationalize IoT-enabled risk assessments in the era of 5G communication.

KEYWORDS

driving risk, Internet of things, in-vehicle camera, on-board diagnostics, usage-based insurance

1 | INTRODUCTION

Risk assessments—identifying, evaluating, and prioritizing the effect of uncertainty—have been considered the first step of various operational risk management, including but not limited to production (Cole et al., 2017), inventory (Michalski, 2009), and procurement (Nagali et al., 2008). Risk analysis is critical but challenging due to the lack of either accurate information (Ale, 2016) or effective assessment models (Goldstein et al., 2017). The Internet of things (IoT) and artificial intelligence (AI) capture and model subtle clues and hidden uncertainties to minimize operational risks (Kumar et al., 2018). Here are two cases illustrating the strategic use of IoT and AI to manage risks. HSBC exemplifies AI-enabled risk detection with its anti–money-laundering solution (Mejia, 2020). Virgin Atlantic makes a fleet of Boeing 787 s and cargo equipment connected via IoT (Accelya, 2016), leading to a 20% reduction in delay risks. The AI identifies more money-laundering patterns than the traditional rule-based method, reducing the risk of false-positive misclassification by 20%.

Technology-enabled risk assessments shall not only improve general operation efficiencies but disrupt specific business practices, especially in the auto-insurance industry. Both incumbents (e.g., State Farm) and insurgents (e.g., Metromile) are experimenting with usage-based insurance (UBI) that charges premiums based on actual driving.

Accepted by Vijay Mookerjee, after two revisions.

Growing at a compound annual growth rate of 23%, UBI is cannibalizing market shares from the traditional practice that prices insurance using the typical bevy of demographics (ReportLinker, 2020). Such a disruptive shift in the auto-insurance industry starts to draw the attention of academia. The related studies focus on either associating crashes with contextual factors or predicting crash risks using personal driving information. For example, Jun et al. (2011) find a positive correlation between speed and accidents. Guo and Fang (2013) apply k-means to cluster drivers by their personality types and predict crash risks. While prior studies take an initial step to predict crash risks using limited data and standard predictive models, the development of innovative crash-risk assessments should be accelerated along with state-of-the-art technologies. There thus exists an urgent need for an operationally efficient solution. Specifically, data collection and risk prediction are fully integrated by IoT and AI to evaluate risks in real time, as Chen and Jiang (2019) believe that leveraging driving monitoring technologies generate efficiency gain in the auto-insurance industry.

A fundamental question naturally arises: *How could an insurance provider leverage IoT telematics and AI predictive analytics to assess crash risks in a disruptive way?* Though we witness the emergence of UBI, it remains unclear what the full potential of UBI would be. In this paper, we propose a generalizable UBI solution, namely, *mobility-based risk assessment* (MRA), which analyzes the real-time data using a flexible modeling approach. Our IoT-enabled solution is designed to accelerate the transformation of the industry by improving the predictive performances of the current demographic- and usage-based predictive practices significantly. Before coming out with the desired design, we face the following challenges: (i) *What data and information would be useful and shall be collected via IoT?* (ii) *Which factors are the most significant in risk predictions?* (iii) *What are the specific requirements a robust predictive model needs to fulfill in the auto-insurance business?*

First, it is generally true that businesses could benefit from fine-granular data for efficient decision making. Due to the lack of proper technologies, auto-insurance providers are used to collecting personal, limited information (e.g., age and self-reported driving tenure) to assess crash risks, leading to undesirable outcomes, such as substantial revenue losses and discrimination. Indeed, the House passes the Prohibit Auto Insurance Discrimination Act in 2019 to prevent statistical discrimination. Prior literature has explored new data for profiling drivers, ranging from naturalistic driving (e.g., Guo & Fang, 2013) to in-vehicle recording (e.g., Boyce & Geller, 2002; WÅhlberg, 2017). They jointly confirm that high-resolution data improve model predictions. A contemporary system thus shall take all kinds of relevant information into account, given the faster IoT data transmission nowadays. Contrasted with a common practice based on demographic-based analytics, we propose a mobility-based solution that incorporates comprehensive behaviors collected via the global positioning system (GPS), on-board diagnostics (OBD), and in-vehicle cameras (IVCs).

Second, it is critical but challenging to determine the most significant crash-contributing factors among enormous semantic features, although more information seems to enhance predictive power in general. Extant literature tends to utilize all of the available attributes to calibrate models (e.g., Guo & Fang, 2013; Wu et al., 2014). Such a kitchen-sink-regression approach is computationally feasible only if the number of features is manageable. However, it often results in overfitting in predictions as more and more predictors are considered. This is particularly troublesome when 5G data communication provides big data in a fast and economical fashion. One may suggest that humans help select the right variables, but human cherry-picking could introduce another source of biases. To have efficient, automatic operations, our robust model embeds feature selection into learning, wherein the feature selection process is specifically modeled as one of the criteria to optimize the predictive engine. Moreover, prior studies usually ignore the cost of misclassifications, especially in the case of false negatives (i.e., a crash-involving trip predicted as a safe one). To fit our model to the auto-insurance industry better, we further integrate a cost-sensitive objective function into the model calibration by penalizing false negatives.

Third, predicting risks is naturally framed as a classification problem from the modeling perceptive. A desirable design ought to fulfill the unique needs in the auto-insurance context, given various models available to exercise predictions. Specifically, it is necessary to balance predictive performances (i.e., precision and recall) and computational efficiency (i.e., time) from an operational perspective. After all, predicting crash risks in real time is quite different from static predictive exercises, since driving behaviors and road conditions may evolve over time. The solution shall be agile in data processing and model training to handle such driving dynamics. Additionally, it would be ideal if the model is data independent but data-skewness-handling. Also, the system should consider multilevel profiling (i.e., trajectory and driver levels), which better accounts for driver heterogeneity and reduce statistical biases. The current UBI practice aggregates data at the driver level to predict crash risks based on standard logistic regressions (LRs).[1] Our design, by contrast, starts with determining trajectory-level risks and synthesizes them as driver-level risks. The modeling approach enhances model performances and enables new prediction capabilities (e.g., trajectory-centric predictions). In this research, we take the three needs as the design principles for an efficient model that predicts multilevel risks in real time.[2]

Besides the above technical considerations, the context of auto insurance by itself could serve as an iconic showcase for researchers and practitioners to understand how technologies help redefine the best practice. The auto-insurance industry at heart runs a service-oriented business, and the core competence of an auto-insurance provider would directly come from precise risk assessments. Yet, insurance providers bear unnecessary revenue losses due to the improper profiling of drivers. For example, State Farm lost 7 billion dollars on its auto-insurance underwriting in 2016 (Simpson, 2017). With its nature of information-intensive decision making, the auto-insurance industry shall benefit more from big data and advanced predictive models than other service sectors. IT-enabled product innovations continue disrupting the current business models, as UBI is cannibalizing significant market shares (Global Market Insights, 2019). We thus are motivated to propose a novel IT solution that enhances the core operational efficiency in the auto-insurance business.

To explore the abovementioned potentials of MRAs, we start to consider fine-granular driving behaviors of a focal driver observed by GPS trackers, OBD, and IVCs. GPS tracks every trajectory of hers and summarizes her usage. OBD is connected to the sensors of her vehicle, recording safety alerts and engine conditions. IVCs with facial recognition monitor her in-vehicle distraction and fatigue. Following prior literature (e.g., Xie et al., 2017) and domain experts, we operationalize a comprehensive set of semantic features extracted from the three IoT-connected devices. We specifically consider two aspects of driving mobility, namely, *trait* and *trajectory*. Traits refer to drivers' distinctive driving behaviors (styles), while a trajectory consists of the vehicle motion sequences along with the contextual factors on a trip.

We face several design challenges, though the development seems a straightforward procedure. The solution needs to deal with high-dimensional data and conduct multilevel analyses. We adopt an eXtreme Gradient Boosting (XGBoost)-based model as the prediction engine that considers risks at both trip and driver levels using classification and regression trees (CARTs). The choice of gradient boosting is attributed to its seamless integration with feature selection, time efficiency in big-data handling (i.e., parallel processing), and robustness to skewed and sparse data. The model also proactively incorporates the cost sensitivity of misclassification into optimization. We train and evaluate our MRA solution using real-world driving data. We partner with one of the Fortune 500 insurance companies. The insurance provider has rolled out the installation program of an IoT-based monitoring system on its clients' vehicles in a major city in China since 2015. We access the 30-week driving trajectories of 398 drivers randomly selected from the entire driver pool. The 184,209 driver-trajectory observations consist of comprehensive driving details, ranging from trip mileage to fatigue level. We have the associated crashes in the police and company reports as ground truths. The driver demographics are also collected to replicate traditional demographic-based profiling. We conduct 10-fold cross-validation to learn the model parameters and obtain several robust results from the calibrated model itself, internal validations, and external benchmarking.

Our interpretable model brings a vital message that behavioral traits are more relevant to crashes than trajectory contexts. The automatic feature selection process lists the top crash-contributing factors, including accelerations, phone using, sharp turns, speeding, and yawning, consistent with prior literature. We also find the salient driver heterogeneity of risks using our hierarchical model approach, which cannot be captured by the existing demographic-based models. When zooming in, we further notice that driver risks evolve. Such dynamic patterns depict diverse adaptive behavioral traits (i.e., self-enforcement to be either safer or more dangerous) across drivers that need to be carefully managed from the modeling perspective. To ensure internal validity, we consider various model specifications and data granularity. The proposed model incorporates driving traits at the individual level and contextual factors at the trip level, outperforming driver- and trip-centric specifications. Also, we confirm that high data resolution improves predictive performance significantly, highlighting the power of real-time IoT-enabled data collection.

In a series of external benchmarking, we first consider a 2×2 evaluation framework regarding input features and modeling choices. The proposed solution (i.e., an XGBoost-based model learned from mobility) is the most preferred. On the one side, mobility-based solutions are consistently superior to demographic-based ones (i.e., the traditional practice), highlighting the value of mobility data. On the other side, our MRA model improves predictive performance metrics, compared with other algorithms. Quantitatively, the proposed model brings a higher recall (precision) than class-weighted LR, nested support vector machine (SVM), and cost-sensitive random forest (RF) by 44.23% (46.68%), 29.18% (28.56%), and 24.59% (23.13%), respectively.[3] Our model also achieves comparable predictive performances to neural networks (NNs) with much lower computational requirements. Second, to test the model robustness to unusual use cases (i.e., skewed and sparse data), we reconstruct the training data by manipulating the distributions of crashes and the data sizes. The results show that our design remains the most robust option. Third, we confirm the computational efficiency of the proposed framework that takes a smaller training set and shorter computing time while achieving decent performance, especially compared with NN. Last, we show our model is more data independent than the four benchmarks.

This study makes several contributions to the literature. We rationalize the business value of IoT from an operational perspective. Prior research has documented that IoT-like technologies, such as radio-frequency identification (RFID), interconnect products, and people to have operational synergy (Rawat et al., 2014). We further argue that the intelligence for real-time decision making complements IoT to be an innovative catalyst, as Olsen and Tomlin (2020) envision. We thus contribute to the literature by developing a novel IoT-based solution that gains operational efficiency in risk assessments by overcoming technical challenges. The proposed solution distinguishes itself from the related literature regarding data uniqueness, modeling novelty, and business applications. We are among the first to comprehensively extract in-vehicle activities to predict crashes. Second, we frame crash predictions as a multilevel classification problem, whereas prior studies mainly conduct single-level analyses (e.g., Paefgen et al., 2014; Xie et al., 2017). Our model also accounts for the risk evolution that is commonly ignored in the crash-analysis literature. Last, we introduce learning analytics to manage the bias–variance trade-off for both needs of association and prediction inferences.

The findings also lead to executable strategies for practitioners. First, our solution showcases how an auto-insurance provider leverages IoT to enhance its core competence. Comprehensive mobility data provide the highest data resolution to improve predictions significantly. Real-time risk assessments are increasingly desirable, as faster data transmission and higher analytical capabilities are cost-efficient in the era of the 5G network. Second, practitioners need to recognize the notable evolution of risks. The ignorance of such dynamic changes results in operations inefficiency. However, insurancers are used to modeling a driver's risk score as a constant on a half-year basis. For example, while Progressive embraces the concept of UBI, its Snapshot program still calculates half-year premiums using driving data in a fixed, 30-day period. Third, our solution leads to new insurance product designs. With GPS tracking, providers can profile risks at the trip level. This helps reinvent trip-level products for taxi services, car rental, and car sharing. Last, auto manufacturers could learn from the insights of this study. The results indicate that distraction and fatigue (besides abnormal accelerations, sharp turns, and speeding) play a considerable role in crashes. To further protect drivers and passengers, auto manufacturers may integrate IoT-enabled devices with face recognition, such as IVC we study here, into online on-board monitoring. The system is proactive in warning drivers once risky driving behaviors are detected. Such a smart system leads to an all-win situation for manufacturers, drivers, and insurance providers.

The rest of the paper proceeds as follows. We first develop the theoretical foundation by reviewing the literature in Section 2. Following the theoretical guidance, we define the problem of risk assessments in Section 3 and propose a novel, IoT-enabled framework to predict risks in Section 4. Section 5 discusses data, model calibration, and internal validations. Our proposed models are benchmarked with alternative classification algorithms in Section 6. We conclude this study with theoretical contributions and managerial implications in Section 7.

2 | LITERATURE

We start reviewing prior work on disruptive technologies in the literature on operations management and information systems. These studies help understand how leveraging advanced technologies enhances value creation and cost reduction in various industries. We then focus on the related research of transportation operations that assess drivers' risks using various archival data.

2.1 | Connectivity-based operations disruption

Academia consistently makes significant efforts to justify the value of IT in business operations due to the exponential growth of IT investments over a few decades. Mithas et al. (2012) consider that IT generates substantial business values regarding revenue enhancement. The revenue-generating capability is enhanced through *productivity gain* (e.g., Hitt & Brynjolfsson, 1996) and *cost reduction* in supply chain coordination (e.g., Dong et al., 2009; Rai & Tang, 2014; Rai et al., 2006). This notion continues along with technological advancements. IoT and the 5G data communication are thus expected to make a trillion-dollar impact on business innovation and operations (McKinsey, 2021). Olsen and Tomlin (2020) even envision that IoT—a vast array of interconnected, intelligent sensors on people, machines, and products—will make more efficient decisions in real time.

IoT itself is a fashionable term, while applications based on a similar concept have been adopted. The contributions of IoT-like technologies to business values are ubiquitous. RFID exemplifies such applications. RFID redefines the process of inventory management in practice, motivating researchers to investigate what could be its optimal use case (e.g., Heese, 2007; Lee & Özer, 2007). Stylized analytical models are developed to consider the interaction between RFID and contextual factors. Delen et al. (2007) show that RFID monitors key performance metrics (e.g., lead time at a retail store) and enables the coordination between a distribution center and a retail store. In the same direction, mobile technologies and location-based services extend interconnections from machine-to-machine connectivity to behavioral human-to-human and human-to-machine interactions. These behavioral big data empower businesses to exercise demand forecasting (e.g., Ban & Rudin, 2019; Feng & Shanthikumar, 2018) and take necessary, proactive actions in advance. Besides, empiricists investigate revenue enhancement in various marketing strategies using mobility data. For example, Brynjolfsson et al. (2013) and Caro and Sadr (2018) study that mobile and augmented reality enhance offline shopping experiences by providing necessary information commonly provided in online shopping contexts. Ho et al. (2020) show that geofencing ads can effectively entice consumers to visit local businesses and make subsequent purchases. Ghose et al. (2019) use trajectory-based mobile coupons to promote sales in a shopping mall. In short, there always exist new types of efficiency gained from the technology-enabled interconnections among people, machines, and products.

IoT is unique due to its ubiquitous interconnectivity and intelligent potential. In other words, IoT enhances or creates business value added using (i) the connected sensors and devices that collect and transfer big data and (ii) AI and machine learning that seamlessly process the collected data and react timely. Ideally, an IoT solution is expected to gather and process high-dimensional data and provide analytical results in a real-time manner (Kumar et al., 2018; Olsen & Tomlin, 2020). As Shacklett (2017) illustrates, a useful IoT-enabled application for a food supply chain could start with monitoring overall plantation environmental factors, such as temperature and humidity. It then makes the corresponding decisions of food replenishment, storage, and delivery by linking smart sensors to an AI-supported enterprise system. While the concept seems straightforward, developing an IoT-based solution is never an easy task. For example, in the theme of retailing, Ghose et al. (2019) reveal that it is technically challenging to implement a location-tracking system to predict consumer trajectories in a real-time manner.

Accordingly, we contribute to the emerging topic of the business value of IoT by designing a novel application in transportation operations. Specifically, we focus on crash-risk assessments and predictions due to their relevance to all auto-insurance stakeholders, as Chen and Jiang (2019) show that leveraging driving behaviors improves risk assessments and operational efficiency significantly in an ecosystem. Given a strong motivation to refine the current UBI practice and disrupt the traditional demographics-based model, we develop a model-independent framework integrating two system components, namely, *IoT-enabled data collection* and *AI-based predictions*. By doing so, we take an initial step to inspire academia to unlock the full potential of IoT in operations management.

2.2 | IoT-enabled risk assessment and crash prediction

The importance of risk assessments and management has been recognized in the literature on operations management. From a broad perspective, prior literature studies risks in various domains (Sodhi et al., 2012), including but not limited to services (e.g., Cohen, 2018), inventory (e.g., Kouvelis & Li, 2019), supply chain coordination (e.g., Gao et al., 2017), manufacturing (e.g., Ding et al., 2015; Matsui & Hillier, 2008), and transportation (e.g., Chung et al., 2017). No matter what reactive strategies are implemented to mitigate risks, risk management always starts with assessing uncertainty, making accessing the right information a prerequisite. As Choi et al. (2016) discuss, IoT provides the highest resolution of big data required to take the first step of risk assessments in general.

This study focuses on transportation. Compared with other operational activities, transportation operations management, especially traffic control and crash assessments, involves complicated decision making due to driver behavioral dynamics. The associated behavioral uncertainty shall be better identified and evaluated when more and more data are collected. We thus believe that related businesses could benefit from advanced technologies in assessing traffic-related risks (e.g., crash risks). In this regard, the objective of this paper is to showcase how IoT and AI are integrated to classify driver risks effectively. The proposed solution contributes to the literature in the aspects of data uniqueness, modeling novelty, and business innovations. Table 1 selects the related work focusing on the crash-risk analysis regarding the above three.

First, we obtain a real-world data set with exclusive characteristics and consider the multifaceted driving behaviors. Prior literature attributes crashes to a set of generic factors collected through offline GPS trackers and OBD (e.g., Jun et al., 2007; Xie et al., 2017), ranging from driving speed to time of day. To characterize drivers thoroughly, we consider semantic driving features and various in-vehicle activities that are recognized by IVCs with facial recognition capabilities. Partnering with an insurance provider in the Fortune 500, we access these features at the driver-trip level in real time. The fine-granular data set enables more in-depth analyses than extant studies that rely on either aggregated data at the driver level (e.g., Guo & Fang, 2013; Wu et al., 2014) or cross-sectional data at the trip level.

Second, this study distinguishes itself from prior literature from a modeling perspective. The existing research generally conducts analyses at a single level (i.e., trip-centric or driver-centric). Ignoring the nature of multilevel data could lead to negative statistical consequences and a poor understanding of heterogeneity if any (Osborne, 2000). A single-level model also hinders individual profiling to some extent due to the mixture of driving traits and contextual effects. Following the theoretical guidance (e.g., Machin & Sankey, 2008; Musicant et al., 2010; Toledo et al., 2008), we split mobility into two semantic components (i.e., behavioral traits and trajectory characteristics) in a multilevel hierarchy. Our model not only profiles each driver's risk by aggregating her behavioral risks at the trip level but incorporates environmental factors across trips. This flexible modeling approach better captures the accident-causing factors, compared with prior work (e.g., Paefgen et al., 2014; Wu et al., 2014). As a result, our model provides an advantage in risk predictions. Deviating from static prediction tasks (e.g., Guo & Fang, 2013; Xie et al., 2017), predicting crash risks naturally involves behavioral evolutions and contextual dynamics over time. The multilevel model can assess both trip- and driver-level risks while controlling for potential evolutions accurately. Our predictions are useful to reprice the existing products (e.g., regular half-year premium) in a shorter period and design new products (e.g., trip-based insurance).

Last, we propose a fully integrated system that demonstrates the synergy between mobile data and novel modeling. In other words, the best system performance comes from the integration of the two, which leads to great association and prediction exercises. Analyzing associations requires a relatively high-*bias* model to infer the data-generating process backward; by contrast, exercising predictions needs a relatively high-*variance* model to choose the correct identifier from a set of outcomes forward (Dietterich & Kong, 1995). The conflict between the two exercises refers to the bias–variance dilemma in machine learning. While prior studies adopt high-*bias* models, such as regressions (e.g., Paefgen et al., 2014) and clustering (e.g., Guo & Fang, 2013), we develop a balanced learning model. Our model

TABLE 1 Selected studies of crash risk analysis

	Data			Modeling			Application		
Reference	GPS	OBD	IVC	Trip	Driver	Time[a]	Association	Prediction	Research focus
This study	√	√	√	√	√	√	√	√	Developing a novel IoT-based solution to predict crash risks
Xie et al. (2019)			√[b]	√				√	Predicting crash risks using trajectories data
Xie et al. (2017)	√			√				√	Predicting crash risks at the geographical-grid level
Paefgen et al. (2014)	√	√			√		√		Associating crashes with trajectory data
Wu et al. (2014)		√	√		√		√		Associating traffic events with demographics
Guo and Fang (2013)		√	√		√		√	√	Associating crashes with demographics and predicting crash risks using *k*-means
Jun et al. (2011)	√				√		√		Associating crashes and driving speed
Musicant et al. (2010)	√	√		√	√		√		Associating crashes with driving time
Jun et al. (2007)	√	√		√	√		√		Associating crashes with driving patterns

[a]Time refers to the consideration of risk evolution over time.
[b]Xie et al. (2019) use traffic videos to construct vehicle trajectories.

manages the trade-off using feature selection to optimize the multilevel learning discussed earlier. In short, we contribute to the literature by introducing a new learning technique along with unique data to shed light on risk assessments and management.

3 | SYSTEM

In this section, we formulate the MRA as a classification problem and decompose the problem into three interrelated modules. We then introduce our design framework encompassing the three components accordingly.

3.1 | Problem definition

Let $U = \{u_i\}$, $i = 1, 2,\ldots, n$, be a set of drivers, and $S = \{s_{ij}\}$, $j = 1, 2,\ldots, m$, be a set of m trips driven by n drivers. Driver u_i involves Risk r_{ij} on Trip s_{ij}, wherein a crash is considered associated with her driving behaviors (e.g., Toledo et al., 2008) and trip-related contexts (e.g., Jun et al., 2011; Musicant et al., 2010). Thus, r_{ij} consists of two semantic components, behavioral and contextual risks.

The majority of prior works tend to pool the two together and conduct risk analyses at either the driver or trip level (e.g., Guo & Fang, 2013). Such a cross-section modeling approach, however, does not consider driver heterogeneity and could lead to undesirable statistical biases (Osborne, 2000). For example, a route that is risky to the general public may not be risky to a skillful driver. It is ideal to balance the average risk at the population level with the specific risk at the individual level, motivating our multilevel modeling. Moreover, behavioral and contextual dynamics update at different paces. For instance, drivers' habits may evolve slowly, whereas route conditions could change dramatically. Put together, we analyze crash risks in a multilevel manner (i.e., population means along with individual heterogeneity), given the hierarchical nature from both behavioral and data-modeling perspectives. On the one hand, a driver's *behavioral traits* refer to her intrinsic driving style that is not volatile across her trips but distinctive among the other drivers'. On the other hand, *trajectory contexts* capture environmental conditions that may vary across trips. Accordingly, we specify the Risk r_{ij} of Driver i on Trip j as

$$r_{ij} = r_i^B + r_j^T + e_{ij}, \tag{1}$$

where r_i^B and r_j^T represent the risks from behavioral traits and trajectory contexts, respectively. IoT-enabled devices, such as OBD, collect rich mobility information at the trip level. The mobility data, M, are operationalized and divided into two corresponding sets of features. Behavioral features, $B = \{\mathbf{b}_{ij}\}$, is the set of associated behavioral states at the trip level, while trajectory features, $T = \{\mathbf{t}_{ij}\}$, is a vector of trip-level characteristics.

MRA problem: *Given a set of drivers U, their existing trips S with features B and T, predict crash risk r_{ij} of Driver i on Trip j by minimizing the errors, where the risk is defined in Equation (1).*

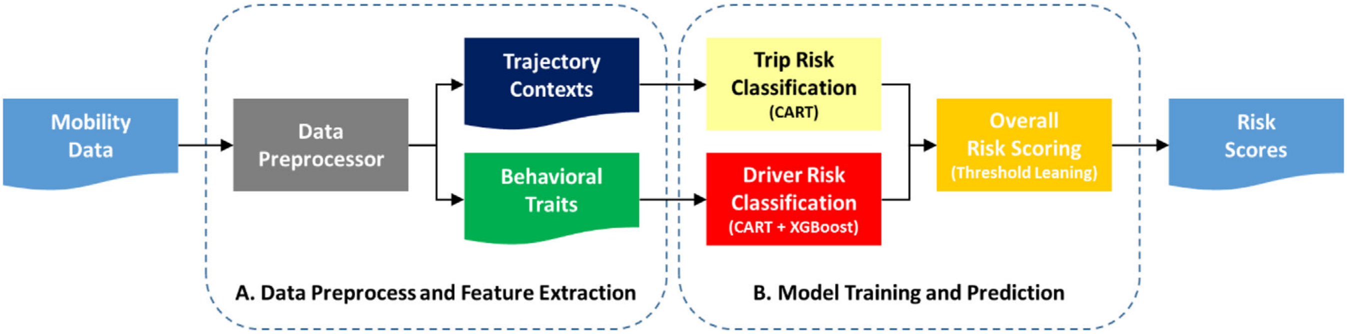

FIGURE 1 Overview of the proposed framework [Color figure can be viewed at wileyonlinelibrary.com]

We have to accomplish three specific tasks, given the MRA framed as a predictive exercise, including (i) trip risk classification (TRC), (ii) driver risk classification (DRC), and (iii) overall risk scoring (ORS). We start by assessing the trip-level risks. Given the crash labels of S, the TRC module is a classifier learned from trajectory features T. In the end, the module predicts and labels each trip of S in probability as the output. Second, the driver risk classifier profiles each driver's risk by aggregating the risks of her trips learned from behavior feature B instead. The DRC module outputs the predicted probability labels at the driver level. Last, given the labels, the ORS module calibrates the parameters and predicts risk scores, R, by weighting the outputs of TRC and DRC.

3.2 | Framework overview

The proposed framework is illustrated in Figure 1. Part A and B focus on data preprocessing and model training, respectively. In Part A, the data preprocessor categorizes the IoT-collected mobility data into trajectory contexts and behavioral traits. It is worth noting that behavior features could be skewed, sparse, and varying across drivers. Turning to Part B, we train the trip risk classifier using the trajectory features correspondingly, while the driver risk classifier learns from behavior features. ORC finally predicts the overall risk ratings.

The proposed framework is data- and model independent to accommodate additional data (e.g., point-of-interest features) and implement the modules with other learning methods. This flexible framework can be applied to not only specific transportation settings, such as the flight delay assessment, but general operational contexts as long as trajectory patterns play a critical role in utility, such as consumer offline preference profiling and local criminal detection.

4 | ALGORITHM

To predict in real time, we design the mobility-based solution to assess crash risks for effective predictive performance and efficient computing time. A data preprocessor is developed to extract the predefined features from IoT-collected data. Then, we model the MRA algorithm as a classifier using CART, XGBoost, and threshold learning.

4.1 | Data preprocess

Suppose IoT-enabled devices continuously upload data to the MRA system. The preprocessor transforms timestamped mobility data (e.g., geo-coordinates and facial images) into a set of features at the trip level. A trip (i.e., trajectory) is defined as follows.

Definition. *Trip s_{ij} of Driver i is defined as a temporally ordered set of tuples $s_{ij} = \{(l_j^1, ts_j^1), \ldots, (l_j^p, ts_j^p)\}$, where $l_j^p = (lat_j^p, long_j^p)$ are the geo-coordinates of Location p, and ts_j^p is the corresponding timestamp.*

Per the domain knowledge, we operationalize S by satisfying three additional conditions, as such (i) the trip duration lasts at least 4 min long ($ts_j^p - ts_j^1 \geq 4$ min), (ii) the speeds of ts_j^1 and ts_j^p are 0, and (iii) any temporary stop within a trip has to be shorter than 4 min. We, thereafter, can calculate the mileage and duration of each trip.[4]

The preprocessor maps the semantic features from OBD and IVCs using timestamps given S. While the OBD data are aggregated at the trip level, the preprocessor applies a pretrained facial recognition model to analyze the videos collected by IVC. At the end of data processing, Trip s_{ij} has a set of mobility features, $\mathbf{m}_{ij} = \{\mathbf{b}_{ij}, \mathbf{t}_{ij}\}$, where $\mathbf{b}_{ij}$ and $\mathbf{t}_{ij}$ refer to behavioral and trajectory features, respectively. We then start to assess crash risks at the trip level while taking the police and company crash reports, $Y = \{y_{ij}\}$, as the ground truth, where $y_{ij} = 1$ if Trip s_{ij} involves in a crash and $y_{ij} = 0$, otherwise.

4.2 | Trip risk classification

We are motivated by prior literature to consider trajectory features, T, and the ground truth, Y, to classify trip risks. We employ a decision-tree–based model to predict the label of S. Specifically, we implement CART (Breiman et al., 2017) with Gini impurity calculation to have risk probabilities that can be readily integrated into the ORC model. We train the classifier by calculating the Gini index of each feature in T, $\text{Gini}(S, T^k)$, where T^k is the kth element of T. The tree-building process is searching for the feature with the lowest value of the index iteratively as

$$\text{Gini}(S, T^k) = \frac{|S_1|}{|S|}\,\text{Gini}(S_1) + \frac{|S_2|}{|S|}\text{Gini}(S_2), \quad (2)$$

where S_1 and S_2 are the subsets of S, and $|S_1|$, $|S_2|$, and, S, refer to the number of trips in Set S_1, S_2, and S, respectively. Intuitively, T^k divides S into S_1, S_2, and the Gini impurity of each subset is formulated as

$$\text{Gini}(S_*) = 1 - \text{Pr}_{S_*}(Y_* = 1)^2 - \text{Pr}_{S_*}(Y_* = 0)^2. \quad (3)$$

If T^k is a binary variable, the tree splits S based on its value; if not, the subsets are divided by searching for the threshold that minimizes $\text{Gini}(S, T^k)$. The CART classifier outputs the risks, r_j^T, at the trip level using trajectory features, T.

4.3 | Driver risk classification

We have to accomplish two tasks to profile drivers' behavioral traits. Specifically, we (i) classify trip risks using behavioral features, B, and (ii) aggregate those classified risks at the trip level to obtain the behavioral traits at the driver level. First, we replicate the exercise of TRC with the consideration of B instead to have r_j^B. It is straightforward but naïve to average out all of a driver's trip risks to assess her risk. This approach is never preferred due to the ignorance of the driver heterogeneity, lowering the predictive power severely. In other words, the impact of a semantic behavioral feature could vary across drivers.[5] In this regard, we follow Chen and Guestrin (2016) to develop a stylized XGBoost algorithm to manage the variance–bias trade-off. Our XGBoost-based model incorporates a cost function to improve predictions by alleviating the variance across drivers. Besides, it is worth noting that the DRC model manages skewed and sparse data efficiently since it inherits parallel tree learning from XGBoost.

Suppose the XGBoost-based model predicts Trip s_{ij} as a crash-involving trip with Risk $x_{s_{ij}}(y = 1)$. We transform the predicted risk into the corresponding probability $\text{Pr}_{S_{ij}}(y = 1)$ with SoftMax function:

$$\text{Pr}_{s_{ij}}(y = 1) = \frac{e^{x_{s_{ij}}(y=1)}}{e^{x_{s_{ij}}(y=1)} + e^{x_{s_{ij}}(y=0)}}. \quad (4)$$

To achieve a cost-sensitive purpose (He & Garcia, 2019), we design and minimize the following cross-entropy function (Shore & Johnson, 1980) to learn the model parameters as

$$\Psi = \lambda - \sum_{i=1}^{n}\sum_{j=1}^{m_i} \text{I}_{s_{ij}}(y = 1)\ln\left(\text{Pr}_{s_{ij}}(y = 1)\right) + \text{I}_{s_{ij}}(y = 0)\ln\left(\text{Pr}_{s_{ij}}(y = 0)\right), \quad (5)$$

where λ is the regularized parameter for all tree classifiers. $\text{I}_{s_{ij}}(y_{ij} = 1)$ is an indicating function wherein

$$y_{ij} = \begin{cases} 1, & \text{I}_{s_{ij}}(y = 1) = 1 \text{ and } \text{I}_{s_{ij}}(y = 0) = 0 \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Last, we conduct the iterative gradient tree boosting processing to optimize Ψ using the first- and second-order approximations (i.e., gradient $\mathbf{gred}_S$ and Hessian $\mathbf{Hess}_S$) specified as

$$\begin{aligned} \mathbf{gred}_S &= \frac{\partial\Psi}{\partial x_{s_{ij}}(y = 1)} + \frac{\partial\Psi}{\partial x_{s_{ij}}(y = 0)} \\ &= \left[\text{Pr}_{s_{ij}}(y = 1) - \text{I}_{s_{ij}}(y = 1)\right] \\ &\quad + [\text{Pr}_{s_{ij}}(y = 0) - \text{I}_{s_{ij}}(y = 0)] \end{aligned} \quad (7)$$

and

$$\begin{aligned} \mathbf{Hess}_S &= \frac{\partial^2\Psi}{\partial^2 x_{s_{ij}}(y = 1)} + \frac{\partial^2\Psi}{\partial^2 x_{s_{ij}}(y = 0)} \\ &= 2\text{Pr}_{s_{ij}}(y = 1)\cdot\text{Pr}_{s_{ij}}(y = 0), \end{aligned} \quad (8)$$

respectively. We have the XGBoost-based classifier outputting the corresponding risks, r_i^B, at the driver level using behavioral features, B.

4.4 | Overall risk scoring

We take the outputs of the TRC and DRC modules (i.e., r_j^T and r_i^B) to predict risk scores. It is reasonable to consider that crash risks may evolve due to unobserved trends, motivating us to evaluate risks on a weekly basis to better capture the trends. In this regard, we obtain behavior risk r_{iw}^B of Drive i and trajectory Risk r_{jw}^T on Trip j in Week w. Given a total of W weeks, we employ a linear weighting vector as

$$\omega(|\omega| = W, \omega_q = q), \quad (9)$$

to smooth out Driver i's time-varying risk evolution over time. A time-discounting weighting is conducted to highlight the importance of more recent patterns, followed by a significant weighting focusing on more representative weeks.

(a) On-Board Diagnostics with GPS-tracking **(b) In-Vehicle Camera**

FIGURE 2 IoT-enabled devices [Color figure can be viewed at wileyonlinelibrary.com]

By doing so, we expect the weighted behavioral traits and trajectory contexts to be more robust.

We finally synthesize r_i^B and r_j^T to predict the risk score of Driver i on Trip j by learning a linear specification as

$$r_{ij} = \alpha r_i^B + \beta r_j^T, \tag{10}$$

where r_i^B and r_j^T are normalized, and α and β are parameters to indicate the significance of the two types of risks. We use a Bayesian optimization method (Snoek et al., 2012) to estimate α and β by anchoring θ, a threshold for labeling crashes.

5 | ANALYSIS

We conduct empirical analyses to examine our method using real-world data, which are collected via IoT-enabled devices. We start this section by detailing the context and data. Then, we apply the fine granular trajectorial data to calibrate our model parameters and obtain interpretable machine learning insights. To further ensure internal validity, we consider different model specifications and data granularity at the end.

5.1 | Data

We partner with a Fortune 500 auto-insurance provider to access fine-granular data.[6] To be innovative, the insurance provider attempts to understand its clients' behaviors by embracing IoT-based technologies and rolling out the monitoring program in a major city in China in 2015. Commercial and residential drivers are required and incentivized by the local government policy and premium reduction to join the program, respectively. To construct the samples, we randomly select 398 drivers from the enrolled pool and collect their representative driving records through IoT-enabled GPS trackers, OBD, and IVCs between August 1, 2017, and March 31, 2018.[7] The final samples consist of 184,209 trips. Recall that the ultimate goal of this study is to develop a novel algorithm, rather than quantifying monitoring-driven behavioral changes. Thus, self-selection and monitoring are less likely to jeopardize the algorithm development while potentially affecting the enrolled drivers' risks.[8] Yet, for the sake of caution, it is mindful of these sampling issues when we interpret the model parameters.

As Figure 2 illustrates, the installed devices become online and start to monitor the mobility of a driver's vehicle when she starts the engine. The GPS tracker uploads the geo-coordinates every 15 s, helping infer the semantic features. We follow prior literature and domain expertise to operationalize risk factors, including speed (Paefgen et al., 2014) and speeding (Cestac et al., 2011), lane changes (Xu et al., 2012), sharp turns (Guo et al., 2021), and trip mileage as well as duration (Jun et al., 2007). OBD capture behavioral states in real time, including accelerations, hard breaks, and front-collision alerts. As Guidotti and Nanni (2020) note, these unusual actions are highly associated with crush accidents. In addition, IVCs with face detection recognize in-vehicle activities, including distraction and fatigue.[9] We thus specifically consider phone use (Razi-Ardakani et al., 2019), smoking (Hutchens et al., 2008), yawning (Vural et al., 2007), and long eye-closing. Last, we map the above GPS-coded trajectories to the geographic information system of the company to extract additional contextual factors, including speed limits (Lee et al., 2006) and weather conditions (Jung et al., 2010). For each trip, timestamps help capture the time of the day (Jun et al., 2011) and the day of the week (Musicant et al., 2010).

Table 2 groups these semantic driving states into behavioral and trajectory features. We use police and company crash filing reports to construct the ground truth. Figure 3 helps illustrate a driver's multiple trajectories on a specific day, wherein the red segment refers to a crash-involving trip. Also, demographics are acquired to replicate the traditional practice (i.e., demographic-based profiling), including age, gender, income, education, driving tenure, and crash history.

We summarize the statistics of the mobility features at the trip level in Table 3.[10] *Crash*, the outcome of predictions, refers to the crash involving Driver i's Trip j in Week t. *Crash* has a mean of 0.0021, indicating that the overall crash rate is 0.21% for the entire data set. Drivers on average drive at

TABLE 2 Descriptions of mobility features

Variable	Description
Outcome of interest	
Crash	Binary indicator of crash involvement
Behavioral features	
Speed	Average speed of the entire trip (km/h)
Speeding	Binary indicator of speeding
Accelerations	# of accelerations ≥ 1.8 m/s^2
HardBreaks	# of decelerations ≥ 1.8 m/s^2
SharpTurns	# of turns where angular velocity ≥ 30 radian/s
LaneChanges	# of turns where 10 radian/s < angular velocity < 30 radian/s
FrontAlerts	# of front-collision alerts
Yawning	# of yawns
EyeClosing	# of eye-closing ≥ 1 s
PhoneUsing	# of phone-related behaviors
Smoking	# of smoking behaviors
Trajectory features	
Mileage	Total distance of a trip (km)
Duration	Total travel time of a trip (min)
SpeedLimit	Average road speed limit of a trip (km)
Daytime	Binary indicator of a daytime trip
Weekday	Binary indicator of a weekday trip
Raining	Binary indicator of raining

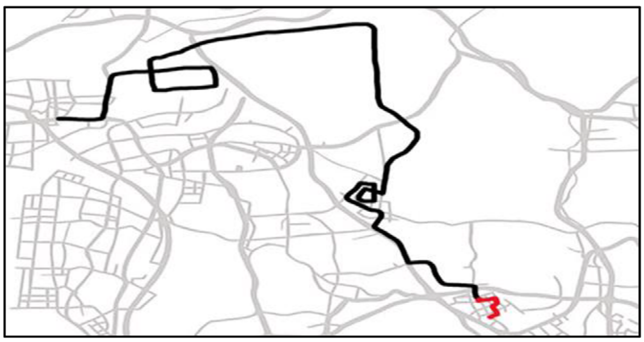

FIGURE 3 Illustration of a driver's trajectories [Color figure can be viewed at wileyonlinelibrary.com]

a speed of 23.79 (km/h) with a range from 0 to 141; 8.1% of trips involve speeding. On a trip, the average number of abnormal accelerations, sharp turns, hard breaks, and lane changes made are 1.51, 3.02, 3.12, and 6.02, respectively. The corresponding ranges and standard deviations of the first three vary significantly. A driver receives 2.02 alerts from the sensor to keep a safe distance. *Yawning* and *EycClosing* indicate the fatigue level; drivers make 1.90 yawns and 0.31 long eye-closing, on average. We also notice that a driver is distracted by phones and smoking 2.98 and 0.12 times, respectively. Regarding trajectory characteristics, a typical

TABLE 3 Summary statistics (N = 184,209)

Variable	Mean	Std. dev.	Min.	Max.
Outcome of interest				
Crash	0.0021	0.0042	0	1
Mobility				
Speed	23.79	48.48	0	141
Speeding	0.081	0.189	0	1
Accelerations	1.51	3.67	1	28
HardBreaks	3.02	8.76	1	25
SharpTurns	3.12	4.90	1	23
LaneChanges	6.02	5.98	3	48
FrontAlerts	2.02	2.45	0	16
Yawning	1.90	4.22	0	19
EyeClosing	0.31	2.98	0	8
PhoneUsing	2.98	5.67	0	32
Smoking	0.12	1.90	0	22
Mileage	10.02	8.46	2.50	60.00
Duration	28.50	26.75	6	131
SpeedLimit	35.02	7.50	15	120
Daytime	0.73	0.26	0	1
Weekday	0.90	0.55	0	1
Raining	0.23	0.04	0	1
Demographics				
Gender	0.65	0.15	0	1
Age	40.32	7.75	25	63
Income	5772.34	1325.70	4872	10,921
Education	12.90	2.41	9	18
Tenure	16.72	7.56	1	40
CrashHistory	3.56	1.65	0	12

trajectory is 10.02 km and 28.50 min long. The average speed limit across all trips is 35.02 (km/h). Seventy-three percent of trips occur in the daytime, while 90% of them are taken on weekdays. About one-quarter of trips are taken place in rainy road conditions. Besides, the summary of demographics indicates that 65% of drivers are male. The average age of the drivers is around 40, while the average driving tenure is over 16 years.

5.2 | Calibration and interpretation

We calibrate the parameters of our MRA algorithm in the following process. Tenfold cross-validation is conducted to learn the model parameters that minimize the errors.[11] Before predicting risk scores, parameters α, β, and θ are estimated with an initial value of 0.5. The choice of 0.5 for α and β can be seen as an equal weight on the information contributed by the behavioral and contextual risks, respectively. θ also starts at 0.5 to avoid undesirable bias. Following Snoek et al.

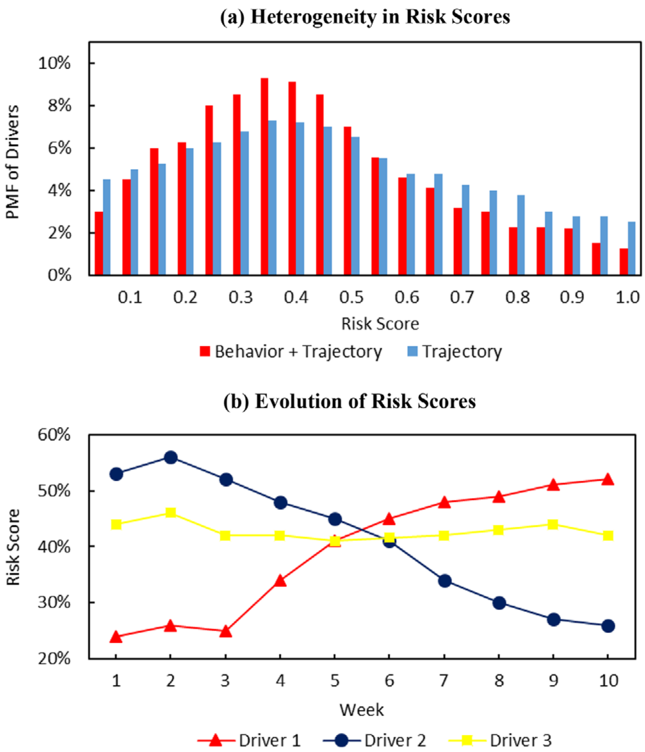

FIGURE 4 Risks of drivers [Color figure can be viewed at wileyonlinelibrary.com]

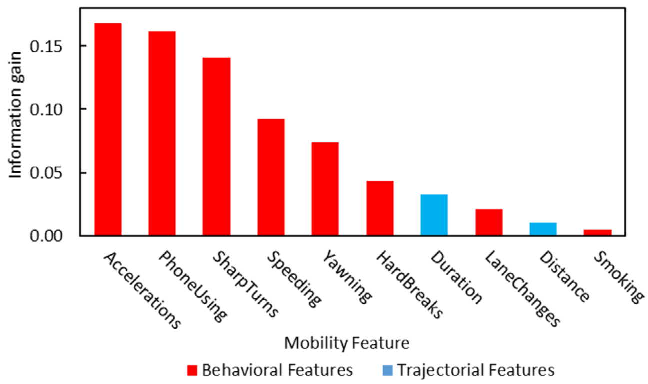

FIGURE 5 Feature importance [Color figure can be viewed at wileyonlinelibrary.com]

(2012), we calibrate the three using the Bayesian optimization method. Intuitively, the parameters are fine-tuned, given the feedback on model performance.

In the process of model calibration, we obtain several interpretable insights into driving behaviors. First, drivers' risk scores are heterogeneous when their driving states are considered. Panel a of Figure 4 compares the cases with and without accounting for behavioral traits, suggesting that they play a critical role in characterizing the hidden heterogeneity across drivers. We show that behavioral features improve predictive performance substantially. Second, we confirm that the propensity to involve in a crash evolves over time. We plot three drivers' distinctive evolution patterns in panel b of Figure 4. Driver 1 initially has a low-risk score but becomes riskier, whereas Driver 2 demonstrates the opposite behavior. Such unexpected patterns illustrate diverse adaptive driving traits (i.e., self-enforcing) across drivers. We smooth these dynamically aggressive (defensive) risks in Equation (9). Last, we can conclude the most significant crash-contributing factors by looking at feature importance obtained from the feature selection process.

As Figure 5 illustrates, behavioral features in red are more significant than trajectory features in blue. Aggressive driving behaviors (i.e., abnormal accelerations, sharp turns, and speeding) increase crash probabilities substantially. Moreover, it is worth noting that distractions (i.e., phone use and smoking) and fatigue (i.e., yawning) contribute to fatal accidents as well. As for the trajectorial features, we confirm that longer trips are associated with higher accident rates. While rainy conditions may result in more crashes, we do not find the evidence. One possible explanation could be that drivers avoid driving on rainy days and use additional caution.

5.3 | Validation

We so far already calibrated our MRA model and interpreted the feature significance. Yet, it is necessary to ensure that the proposed model is the most suitable across alternative model specifications within the same concept. Recall Section 4.3, in which we follow theoretical guidance to consider (i) behavioral factors for profiling driving traits at the driver level and (ii) contextual factors for profiling trajectory risks at the trip level. We synthesize (i) and (ii) to have the proposed model using Equation (10). To show the value of our novel modeling approach, we specify two alternatives, wherein behavioral and contextual risks are assessed at either the driver level or the trip level. We rewrite Equation (10) for the driver-centric and trip-centric models as

$$r_i = \alpha r_i^B + \beta r_i^T \tag{11}$$

and

$$r_j = \alpha r_j^B + \beta r_j^T, \tag{12}$$

respectively. As a result, the driver-centric model may not capture the significance of contextual factors, whereas the trip-centric model could ignore driver heterogeneity. We compare the three variants regarding precision and recall since we frame risk predictions as a classification problem. Table 4 summarizes the comparisons. The proposed model stands out and results in the best performance, followed by the driver-centric one. Quantitatively, the former boosts the precision and recall of the latter by 9.69% and 10.74%. The evidence illustrates the flexibility of our unique hierarchical modeling as the heart of the solution.

Besides, it is worth experimenting with the impact of data granularity on model performance, given the big trajectorial

TABLE 4 Alternative specifications

	Main model $Behavior_i$ + $Trajectory_j$	Driver-centric $Behavior_i$ + $Trajectory_i$	Trip-centric $Behavior_j$ + $Trajectory_j$
Precision	89.76%	81.83%	74.89%
Recall	89.02%	80.39%	73.29%

TABLE 5 Data granularity

	15 s	60 s	300 s
Precision	89.76%	85.36%	80.04%
Recall	89.02%	85.03%	78.21%

Note: The data are sampled every 15, 60, and 300 s, respectively.

data available. We thus reconstruct the data by extending the data sample window from 15 to 60 and 300 s. This investigation shall help show the data-model sensitivity. Table 5 reports the corresponding model performances. The sensitivity analysis suggests that the finer the data granularity is, the better the model performs, highlighting the value of IoT-enabled real-time data collection. In sum, the two internal validations confirm our hierarchical modeling approach and suggest finer data granularity, respectively. Next, we conduct external benchmarking by considering various combinations of alternative predictive models and data specifications.

6 BENCHMARKING

We benchmark the proposed solution against alternative model specifications in the following empirical experiments. The primary goal of this research is to investigate whether the IoT-enabled risk assessment is more efficient to disrupt the current best practice (i.e., demographic-based profiling). While using precision and recall as the performance metrics, we would pay additional attention to recall in the following analyses due to costly false-negative cases in the insurance industry. Following the literature on machine learning, we specifically consider four advanced, representative classifiers as the benchmark methods, namely, class-weighted LR (King & Zeng, 2001), nested SVM (Lee & Scott, 2010), cost-sensitive RFs (Devi et al., 2019), and artificial NN (Kim et al., 2005). We follow the literature to fine-tune model parameters for each benchmark, such that we shall have decent comparisons. In the following benchmarking, we comprehensively consider predictive performance, robustness to data skewness and sparsity, time efficiency, and data independence.

We start with a 2×2 evaluation framework regarding input data and modeling choices. As Figure 6 visualizes, the vertical dimension represents the comparison between mobility- and demographic-trained modes. Meanwhile, the proposed model is compared with other classifiers in the horizontal dimension for the same training feature set.

We summarize the model performances in Table 6. The results highlight the advantage of mobility-based profiling.

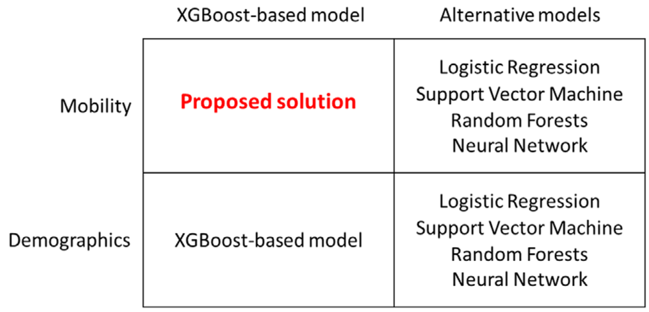

FIGURE 6 2×2 evaluation framework [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 6 Benchmarking—Predictive performance

(a) Precision

	(1) Main model	(2) LR	(3) SVM	(4) RF	(5) NN
(i) Mobility	89.76%	61.12%	69.82%	72.90%	90.02%
(ii) Demographics	63.12%	52.32%	56.31%	56.51%	63.50%

(b) Recall

	(1) Main model	(2) LR	(3) SVM	(4) RF	(5) NN
(i) Mobility	89.02%	61.72%	68.91%	71.45%	89.58%
(ii) Demographics	61.45%	50.12%	55.34%	55.40%	62.08%

The models learned from mobility significantly improve the metrics (i.e., precision and recall), compared with those based on demographics. Specifically, the proposed design generates a much higher recall than a same-specification model trained by demographic by 44.86% (see Model 1 in Panel b). Moreover, when we compare mobility-data–trained models, our multilevel modeling approach outperforms class-weighted LR, nested SVM, and cost-sensitive RF, lifting recall by 44.23%, 29.18%, and 24.59%, respectively (see Row (ii) in Panel b).[12] Thus, our model should be even more accurate than the current UBI practice based on standard LRs. While providing comparable predictive performance to NNs, our model is more computationally efficient and qualitatively interpretable. These similar patterns can be found for precision in Panel a. The above empirics suggest that the best performance occurs only if IoT-collected mobility data are applied to an AI-based predictive engine, as the proposed MRA solution does.

Second, we are interested in testing whether the MRA solution is robust to unusual use cases. It is quite challenging for machines to learn from noninformative (i.e., skewed or small) training data. We thus reconstruct the training set by manipulating the distribution of positive labels (i.e., crash-involving trip) in the training set. In this regard, an extreme case would be that only 10% of positive cases are used in training. We plot the model performance across the five models regarding

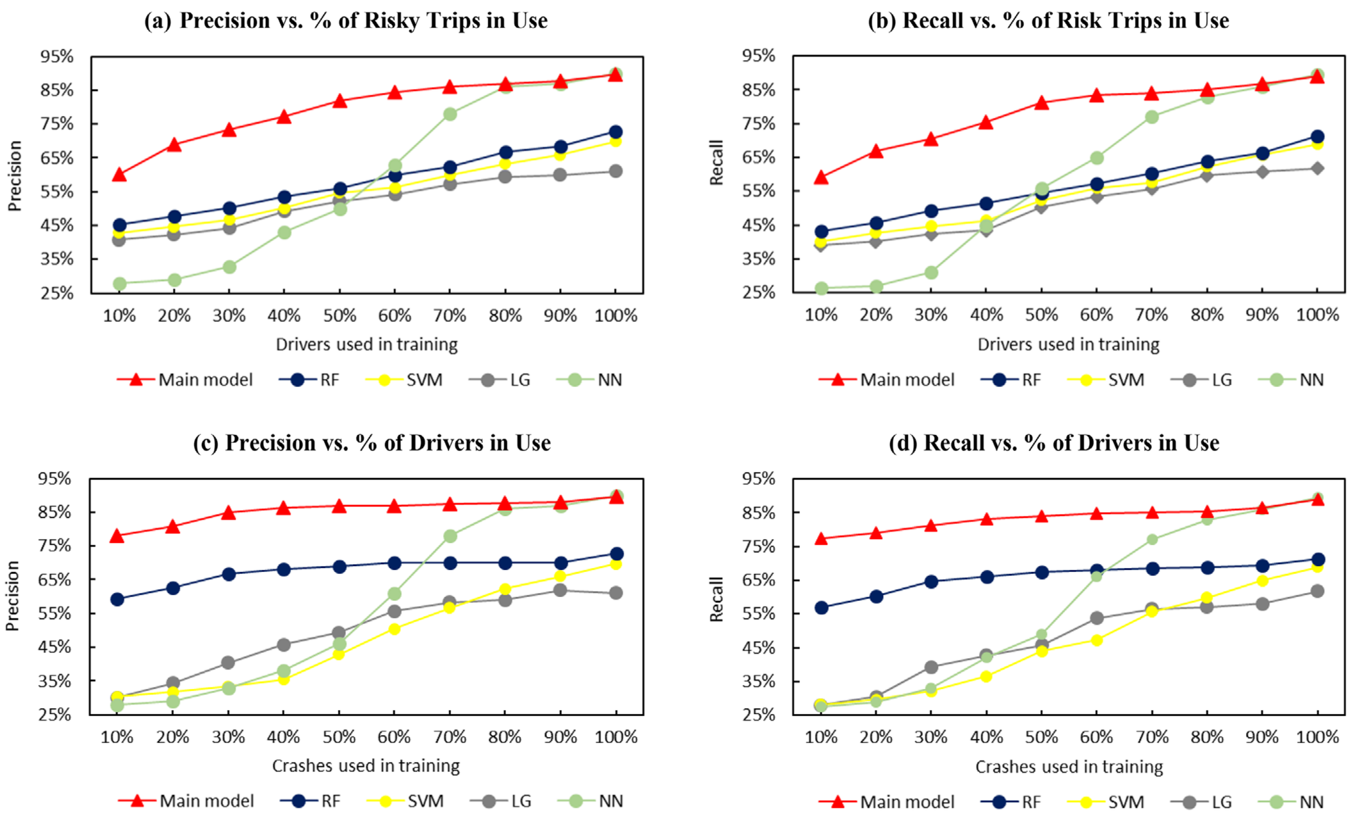

FIGURE 7 Benchmarking—Data skewness and sparsity [Color figure can be viewed at wileyonlinelibrary.com]

the manipulated distribution in Figure 7. The proposed solution is the most robust to skewed data. Panels a and b show that the XGBoost-based model learns efficiently. It has decent precision and recall (i.e., over 80%) using less than 50% of positive label samples, whereas NN requires more than 70% of data. RF, though inferior to our model, demonstrates its better learning efficiency than SVM and LR. Besides, we are curious about how the MRA algorithm manages the issue of small training data. We train models using smaller data sets that consist of fewer drivers, starting from 10% of the full training samples. The proposed model dominates each benchmark algorithm in precision and recall, as shown in Panels c and d, respectively. The model has a performance of 75% initially, improves precision and recall to 80% promptly, and reaches the saturation point of 89% using 10%, 20%, and 100% of the full training set, respectively. These two sets of results jointly conclude that the proposed model is robust to both skewed and small training sets.

Third, we report the processing time of each model in Table 7. Though logistics regression is the fastest among the four due to its simplicity, it trades predictive power for faster processing speed. This reminds us of the trade-off between predictive performance and computational time. The proposed model is preferred since it can deliver the same outperformance in a much shorter time. In other words, the proposed model could save processing time using a relatively smaller training set than the benchmarks in order to achieve the same performance. As Figure 8 visualizes, it needs only

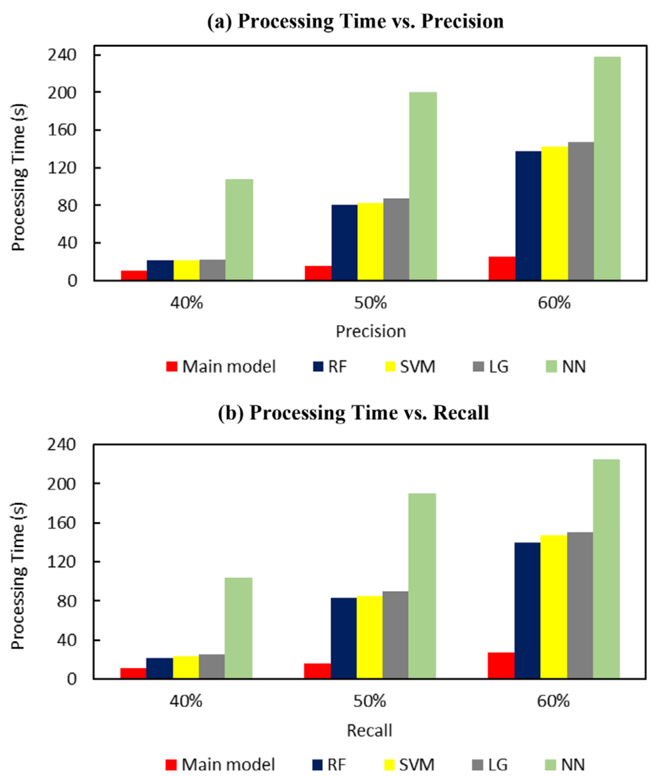

FIGURE 8 Model comparison for varying performance [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 7 Benchmarking—Processing time

Weeks of data	1	5	10	15	20	25	30
Main model	6.48	35.40	67.67	101.93	130.33	180.36	201.74
Logistic regression	6.71	41.23	87.47	135.62	181.24	222.17	272.91
Random forest	5.57	27.51	59.83	089.89	121.43	151.23	178.68
Support vector machine	5.38	27.83	61.43	091.46	122.31	153.87	181.11
Neural network	8.78	58.92	109.20	200.13	278.39	300.33	401.27

Note. The processing time is measured in the computing environment: CPU of 24cores Intel i7-6850K @ 3.60GHZ, Memory of 64GB DDR4, and Hard drives of 1 TB SSD and 4 TB HDD.

TABLE 8 Benchmarking—Data independence

	(1) Main model	(2) LR	(3) SVM	(4) RF	(5) NN
(a) Precision					
(i) In-sample	89.76%	61.12%	69.82%	72.90%	90.02%
(ii) Out-of-sample	87.18%	59.22%	65.18%	70.12%	80.12%
(b) Recall					
(i) In-sample	89.02%	61.72%	68.91%	71.45%	89.58%
(ii) Out-of-sample	87.18%	59.22%	65.18%	70.12%	80.12%

27 s to have 60% precision and recall quantitatively, whereas the other three take more than 140 s to achieve the same performance, not to mention the case of NN.

Last, we conclude model benchmarking with data independence. This investigation shows whether model performances are data specific. A preferred model should be less data specific. To conduct external validity, we collect the second set of drivers from another major city in China between August 1, 2017, and March 31, 2018, resulting in 176,823 individual-trip observations. The model parameters learned from the first data set are applied to predict risks using the second data set. Table 8 reports the cases regarding in-sample and out-of-sample validations. Model predictive performances drop consistently. The proposed model is the most data independent among the five, whereas NN is the most data sensitive. It is also worth noting that our model outperforms NN in the out-of-sample validation. To sum up, the benchmarking jointly highlights the synergy between the model and trajectorial data regarding predictive performance, robustness to skewed and sparse data, computational efficiency, and adaptability.

7 | CONCLUSION AND FUTURE RESEARCH

We have proposed a novel design to predict crash risks using the latest technologies. Different from the traditional demographic-based practice, the proposed framework leverages mobility-based modeling. Specifically, we collect the fine-granular data from various IoT-enabled devices, including GPS trackers, OBD, and IVCs. To classify drivers' risks, we develop a robust predictive engine that considers crash risks regarding driver-centric behaviors along with trajectory-centric contexts. We implement an XGBoost-based model that can predict crashes at both trajectory and driver levels flexibly. Applying the proposed solution to the real-world data, we obtain several robust results from the calibrated model itself and the model evaluations. The interpretable results of the model indicate that driving traits play a significant role in accidents. The top impactful features include abnormal accelerations, sharp turns, and distractions. We notice the salient heterogeneity of driving behaviors that could not be captured by static demographics previously. The zoomed-in weekly analyses further show that drivers' risks evolve dynamically. Such unexpected patterns illustrate diverse adaptive driving traits (self-enforcing) across drivers. In other words, drivers could become more aggressive (defensive) over time. Before moving forward to external benchmarking, we conduct internal validations by comparing modeling approaches and data granularity to ensure proper model and data specifications.

We benchmark the proposed solution against the alternatives in a series of empirical validations. We first consider a 2×2 evaluation framework regarding input features and modeling choices. The results jointly demonstrate that the proposed solution outperforms the other approaches. The models learned from mobility improve precision and recall significantly, compared with the models learned from demographics. Specifically, the MRA solution results in higher recall (precision) than the XGBoost-based model based on demographics by 44.86% (42.21%). Given the mobility features, our approach has comparable performance to NN, improving the recalls of LR, SVM, and RF by 44.23%, 29.18%, and 24.59%, respectively. Second, we are interested in testing the model robustness to unusual use cases (i.e., skewed and small training data). We reconstruct the training set by experimenting with the various distribution of crashes and the different number of drivers included. The empirical evaluations show that the proposed solution is the most robust. Third, the within-solution comparison highlights the value of the multilevel modeling when we compare the full model with the parsimonious alternatives (i.e., trajectory-centric and driver-centric models). Last, the proposed model is more computationally efficient regarding a smaller training set requirement with superior performances. Last, our model is the most data independent, whereas NN is the most data sensitive. It is also worth noting that the proposed model significantly outperforms NN in external validations.

This study contributes to the literature by developing a novel solution wherein IoT and AI improve operational efficiency in assessing crash risks. While technically challenging, the proposed solution shows its uniqueness from prior research in three aspects: data, modeling, and applications. We first utilize IoT-enabled sensors (e.g., IVC) to gather proprietary data that monitor behavioral details in real time, whereas prior studies rely on archival data through OBD (e.g., Paefgen et al., 2014). These high-resolution data further enable us to profile risks in a multilevel manner. Unlike single-level analyses (e.g., Xie et al., 2017), our driver-trip profiling separates behavioral risks from contextual dangers (e.g., speed limits) to calibrate drivers' heterogeneity more precisely. We factor in the existing evolution of risk classification that is commonly ignored (e.g., Wu et al., 2014). Last, we account for the bias–variance dilemma to meet the needs for association and prediction inferences. We introduce an efficient learning technique to improve the performances of naïve models significantly (e.g., Guo & Fang, 2013).

These results are relevant to the auto-insurance–related business and provide several new directions for practitioners. First, it is essential to incorporate mobility data into risk assessments while major players in the United States still hesitate. Mobility-based modeling improves the performance of risk predictions significantly, confirmed in the empirical validations. Insurance providers are encouraged to embrace this new best practice proactively to fuel core competencies. The investment in the mobility-based solution can be justified by the efficiency gain (i.e., cost reduction in underwriting) as the costs of IT catalysts (e.g., IoT and AI) diminish over time. As expected, this paradigm can be even more potent to predict in real time with the faster data communications of the 5G network. Second, practitioners must pay attention to the evolution of risks. The ignorance of dynamic changes underestimates or overestimates crash risks, resulting in operations inefficiency. Insurance providers, however, are used to modeling a driver's score as a constant on a half-year basis, no matter using static demographics or limited mobility data. For example, Progressive's *Snapshot* program evaluates the pooled data of a fixed, 30-day driving period for half-year premiums. This half-year profiling causes such derivative issues as moral hazards. We thus suggest that practitioners keep monitoring drivers' behaviors and model the corresponding risks in a dynamic manner. Third, our solution can lead to new product designs for the auto-insurance industry. With GPS tracking, insurance providers can profile risks at the trip level, as does the trajectory-centric component in the MRA solution. This new capability helps reinvent trip-level products perfect for taxi services, car rental, and car-sharing (e.g., Zipcar). A compelling use case would be that car-sharing companies do not charge a predefined flat rate upfront but a customized trip-level premium afterward.

Besides, auto manufacturers could learn from the results to further protect drivers and passengers. Distraction and fatigue are equally dangerous, though crashes are commonly ascribed to improper actions, such as abnormal accelerations and sharp turns. Ignoring distraction and fatigue becomes fatal when no further actions are taken. This study showcases a solution to prevent drivers from such a danger. With facial recognition, IVCs monitor drivers' concentration levels by detecting the behaviors of distraction (e.g., phone use and smoking) and signs of fatigue (e.g., yawning). It will be ideal for auto manufacturers to integrate IoT-enabled IVCs into cloud-based monitoring in real time. The system can be proactive in taking necessary interventions, such as warming and vibrating steering, once the undesired behaviors are detected. We believe that such a smart system shall lead to an all-win situation for auto manufacturers, drivers, and insurance providers.

This study faces the following limitations. First, we suffer from the lack of comprehensive trajectory contexts while accessing high-resolution mobility data. We acknowledge that utilizing additional trajectory features is one way to improve model performances, as Jia et al. (2018) and Paefgen et al. (2014) study the impacts of points of interest and road conditions, respectively. We alternatively leverage driving behaviors to predict risks. The two shall complement each other, focusing on extrinsic (i.e., environmental) characteristics and intrinsic (i.e., behavioral) factors. In this regard, our solution is purposely designed to incorporate every new environment feature whenever available (e.g., city and rural settings) to amplify the predictive power. In other words, our model would be relatively data independent. Second, there will be more powerful algorithms available at pace with technological advancements. One could expect that the predictive performance will be further improved, while XGBoost is considered one of the most preferred classifiers at this moment. We thus propose a generalizable, model-independent framework so practitioners would be less constrained to choose backend learning methods to accomplish context-specific tasks. As acknowledged, our solution cannot be the all-time best in a specific context but shall be generally applicable to various industries for predictions. Third, there exist privacy concerns in the data collection process. In the beginning, drivers may be reluctant to share their driving data, especially when the notion of UBI has not been widely accepted. Yet, we shall expect that more and more drivers would try and enroll in new insurance programs in a similar adoption process of credits and online shopping. With the emergence of UBI, privacy issues seem gradually moderated and shall be less concerned in the near future. The last limitation is related to the profit-efficiency qualification. It is very operationally efficient to fine-tune product premiums using the drivers' risks profiled by the MRA solution. Yet, we are restricted from manipulating pricing due to the scope of research collaboration. Our system serves as the first stepping stone to efficiency gain in the subsequent operations; after all, the core of the insurance businesses is hedging, whose initial, critical step is all about identifying risks. As a future direction, we pursue maximizing an insurance provider's profits by pricing the existing products and designing new trip-level products exquisitely on the top of the already-precise predictions.

Despite these limitations, this study is one of the first to develop an IoT-enabled solution for robust risk assessments. The generalizable framework provides useful managerial implications on how the auto-insurance industry utilizes IoT and AI to redefine the best practice, reinvent insurance products, and disrupt the existing business model in the era of the 5G network.

ACKNOWLEDGMENTS

The authors thank Professor Vijay Mookerjee, the senior editor, and anonymous reviewers for constructive suggestions. The authors also thank the participants at the 2020 Symposium on Statistical Challenges in Electronic Commerce and Research as well as the research seminars at University of Florida and National Dong Hua University, Taiwan, for helpful comments.

ORCID

Yi-Jen (Ian) Ho https://orcid.org/0000-0002-2480-480X

ENDNOTES

[1] We interview two UBI providers, including the one we collaborate with. Due to the nondisclosure agreement, the providers remain anonymous.

[2] The real-time system is implemented in a distributed setting to achieve real-time processing by following Chen and Gurestrin (2016). We have eight virtual cores, 30 GB of RAM, and two 80 GB SSD local disks in each cluster machine. We integrate the proposed solution with Apache Spark's MLLIB framework using XGBoost4J-Spark.

[3] Recall is chosen as the primary metric due to the higher cost of false-negative cases than false-positive ones.

[4] We calculate the distances of all segments within a trip using Haversine formula and sum them out.

[5] For example, Driver A is an experienced driver and managers to drive fast, whereas Driver B is an amateur driver but tends to drive fast. While fast driving is dangerous in general, the effect of this behavioral state would be different for the two.

[6] Due to the efficiency of UBI (i.e., claim cost reduction), major insurance companies are proactively offering UBI programs and retiring the traditional products, while insurgent startups often seem more innovative.

[7] We randomly select 0.1% of drivers from the driver pool due to the computation limitations. The samples represent the entire pool, given the two's statistically identical distributions of demographics. Commercial and residential drivers account for 48.2% and 51.8% of the final sample. Driving trajectories covers various city and rural areas.

[8] We acknowledge that the data may consist of fewer risky driving actions (e.g., sharp turns), making it harder to train our model with lower data variation. In such a less ideal scenario, our model still outperforms the traditional method. We shall expect that the predictive performance could be further improved as more drivers join the program.

[9] We use the facial recognition algorithm pretrained by the data provider. Due to the nondisclosure policy, we cannot reveal many details of the facial recognition algorithm. It is built upon a Facial Action Coding system (Vural et al. 2007) using the proprietary expert-labeled data. The trained model results in 96% and 97% of precision and recall, respectively.

[10] The correlations among variables are reported in Table A1 (see the Supporting Information).

[11] We alternatively conduct a holdout training strategy, wherein 90% of the data are used to train the hyperparameters. The models based on the two training methods have indistinguishable performances summarized in Table A2 in the Supporting Information.

[12] We also provide additional benchmarks of standard LR, SVMs, and RFs in Table A3 in the Supporting Information.

REFERENCES

Ale, B. (2016). Risk analysis and big data. *Safety and Reliability*, *36*(3), 153–165. https://doi.org/10.1080/09617353.2016.1252080

Accelya. (2016). *5 great ways airlines are using the Internet of Things*. https://w3.accelya.com/blog/5-great-ways-airlines-are-using-the-internet-of-things

Ban, G. Y., & Rudin, C. (2019). The big data newsvendor: Practical insights from machine learning. *Operations Research*, *67*(1), 90–108. https://doi.org/10.1287/opre.2018.1757

Boyce, T. E., & Geller, E. S. (2002). An instrumented vehicle assessment of problem behavior and driving style: Do younger males really take more risks? *Accident Analysis & Prevention*, *34*(1), 51–64.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). *Classification and regression trees*. Routledge.

Brynjolfsson, E., Hu, Y., & Rahman, M. S. (2013). Competing in the age of omnichannel retailing. *MIT Sloan Management Review*, *54*(4), 23–29.

Caro, F., & Sadr, R. (2018). Internet of things (IoT) in retail: Bridging supply and demand. *Business Horizons*, *62*(1), 47–54. https://doi.org/10.1016/j.bushor.2018.08.002

Cestac, J., Paran, F., & Delhomme, P. (2011). Young drivers' sensation seeking, subjective norms, and perceived behavioral control and their roles in predicting speeding intention: How risk-taking motivations evolve with gender and driving experience. *Safety Science*, *49*(3), 424–432. https://doi.org/10.1016/j.ssci.2010.10.007

Chen, Y. -. H., & Jiang, B. (2019). Effects of monitoring technology on the insurance market. *Production and Operations Management*, *28*(8), 1957–1971. https://doi.org/10.1111/poms.13023

Chen, T., & Guestrin, C. (2016). XGboost: A scalable tree boosting system. In Balaji Krishnapuram, Mohak Shah, Alexander J. Smola, Charu Aggarwal, Dou Shen, & Rajeev Rastogi (Eds.), *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. Association for Computing Machinery (pp. 785–794).

Choi, T. M., Chan, H. K., & Yue, X. (2016). Recent development in big data analytics for business operations and risk management. *IEEE Transactions on Cybernetics*, *47*(1), 81–92. https://doi.org/10.1109/TCYB.2015.2507599

Chung, S. H., Ma, H. L., & Chan, H. K. (2017). Cascading delay risk of airline workforce deployments with crew pairing and schedule optimization. *Risk Analysis*, *37*(8), 1443–1458. https://doi.org/10.1111/risa.12746

Cohen, M. C. (2018). Big data and service operations. *Production and Operations Management*, *27*(9), 1709–1723. https://doi.org/10.1111/poms.12832

Cole, S., Giné, X., & Vickery, J. (2017). How does risk management influence production decisions? Evidence from a field experiment. *The Review of Financial Studies*, *30*(6), 1935–1970. https://doi.org/10.1093/rfs/hhw080

Delen, D., Hardgrave, B. C., & Sharda, R. (2007). RFID for better supply-chain management through enhanced information visibility. *Production and Operations Management*, *16*(5), 613–624. https://doi.org/10.1111/j.1937-5956.2007.tb00284.x

Devi, D., Biswas, S. K., & Purkayastha, B. (2019). A cost-sensitive weighted random forest technique for credit card fraud detection. In Prakash Duraisamy (Ed.), *Proceedings of 10th international conference on computing, communication and networking technologies*. Institute of Electrical and Electronics Engineers (pp. 1–6).

Dietterich, T. G., & Kong, E. B. (1995). *Machine learning bias, statistical bias, and statistical variance of decision tree algorithms*. Technical report, Oregon State University Corvallis, OR.

Ding, X., Tian, Y., & Yu, Y. (2015). A real-time big data gathering algorithm based on indoor wireless sensor networks for risk analysis of industrial operations. *IEEE Transactions on Industrial Informatics*, *12*(3), 1232–1242. https://doi.org/10.1109/TII.2015.2436337

Dong, S., Xu, S. X., & Zhu, K. X. (2009). Research note—Information technology in supply chains: The value of IT-enabled resources under competition. *Information Systems Research*, *20*(1), 18–32. https://doi.org/10.1287/isre.1080.0195

Feng, Q., & Shanthikumar, J. G. (2018). How research in production and operations management may evolve in the era of big data. *Production and Operations Management*, *27*(9), 1670–1684. https://doi.org/10.1111/poms.12836

Gao, L., Yang, N., Zhang, R., & Luo, T. (2017). Dynamic supply risk management with signal-based forecast, multi-sourcing, and discretionary selling. *Production and Operations Management*, *26*(7), 1399–1415. https://doi.org/10.1111/poms.12695

Ghose, A., Li, B., & Liu, S. (2019). Mobile targeting using customer trajectory patterns. *Management Science*, *65*(11), 5027–5049. https://doi.org/10.1287/mnsc.2018.3188

Global Market Insights. (2019). *Usage-based insurance market revenue to hit USD 115 bn by 2026*. https://www.prnewswire.com/news-releases/usage-based-insurance-market-revenue-to-hit-usd-115-bn-by-2026-global-market-insights-inc-300970226.html

Goldstein, B. A., Navar, A. M., Pencina, M. J., & Ioannidis, J. (2017). Opportunities and challenges in developing risk prediction models with electronic health records data: A systematic review. *Journal of the American Medical Informatics Association*, *24*(1), 198–208. https://doi.org/10.1093/jamia/ocw042

Guidotti, R., & Nanni, M. (2020). Crash prediction and risk assessment with individual mobility networks. In Karine Zeitouni, Cristian Borcea, Flora Salim, Dimitrios Skoutas, & Shaoyi Yin (Eds.), *Proceedings of 2020 21st IEEE international conference on mobile data management*. Institute of Electrical and Electronics Engineers (pp. 89–98).

Guo, F., & Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. *Accident Analysis & Prevention*, *61*, 3–9.

Guo, M., Zhao, X., Yao, Y., Yan, P., Su, Y., Bi, C., & Wu, D. (2021). A study of freeway crash risk prediction and interpretation based on risky driving behavior and traffic flow data. *Accident Analysis & Prevention*, *160*, 106328.

He, H., & Garcia, E. A. (2019). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284.

Heese, H. S. (2007). Inventory record inaccuracy, double marginalization, and RFID adoption. *Production Operations Management*, *16*(5), 542–553. https://doi.org/10.1111/j.1937-5956.2007.tb00279.x

Hitt, L. M., & Brynjolfsson, E. (1996). Productivity, business profitability, and consumer surplus: Three different measures of information technology value. *MIS Quarterly*, *20*(2), 121–142. https://doi.org/10.2307/249475

Ho, Y.-J., Dewan, S., & Ho, Y.-C. (2020). Distance and local competition in mobile geofencing. *Information Systems Research*, *31*(4), 1421–1442. https://doi.org/10.1287/isre.2020.0953

Hutchens, L., Senserrick, T. M., Jamieson, P. E., Romer, D., & Winston, F. K. (2008). Teen driver crash risk and associations with smoking and drowsy driving. *Accident Analysis & Prevention*, *40*(3), 869–876.

Jung, S., Qin, X., & Noyce, D. A. (2010). Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accident Analysis & Prevention*, *42*(1), 213–224.

Jia, R., Khadka, A., & Kim, I. (2018). Traffic crash analysis with point-of-interest spatial clustering. *Accident Analysis & Prevention*, *121*, 223–230.

Jun, J., Ogle, J., & Guensler, R. (2007). Relationships between crash involvement and temporal-spatial driving behavior activity patterns: Use of data for vehicles with global positioning systems. *Transportation Research Record*, *2019*(1), 246–255. https://doi.org/10.3141/2019-29

Jun, J., Guensler, R., & Ogle, J. (2011). Differences in observed speed patterns between crash-involved and crash-not-involved drivers: Application of in-vehicle monitoring technology. *Transportation Research Part C: Emerging Technologies*, *19*(4), 569–578. https://doi.org/10.1016/j.trc.2010.09.005

Kim, Y., Street, N. W., Russell, G. J., & Menczer, F. (2005). Customer targeting: A neural network approach guided by genetic algorithms. *Management Science*, *51*(2), 264–276. https://doi.org/10.1287/mnsc.1040.0296

King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, *9*(2), 137–163. https://doi.org/10.1093/oxfordjournals.pan.a004868

Kouvelis, P., & Li, R. (2019). Integrated risk management for newsvendors with value-at-risk constraints. *Manufacturing & Service Operations Management*, *21*(4), 816–832.

Kumar, S., Mookerjee, V., & Shubham, A. (2018). Research in operations management and information systems interface. *Production and Operations Management*, *27*(11), 1893–1905. https://doi.org/10.1111/poms.12961

Lee, C., Hellinga, B., & Saccomanno, F. (2006). Evaluation of variable speed limits to improve traffic safety. *Transportation Research Part C: Emerging Technologies*, *14*(3), 213–228. https://doi.org/10.1016/j.trc.2006.06.002

Lee, H., & Özer, Ö. (2007). Unlocking the value of RFID. *Production Operation Management*, *16*(1), 40–64. https://doi.org/10.1111/j.1937-5956.2007.tb00165.x

Lee, G., & Scott, C. (2010). Nested support vector machines. *IEEE Transactions on Signal Processing*, *58*(3), 1648–1660.

Machin, M. A., & Sankey, K. S. (2008). Relationships between young drivers' personality characteristics, risk perceptions, and driving behaviour. *Accident Analysis & Prevention*, *40*(2), 541–547.

Matsui, M., & Hillier, F. S. (2008). *Manufacturing and service enterprise with risks*. Springer.

McKinsey. (2021). *The Internet of things: Catching up to an accelerating opportunity*. https://www.mckinsey.com/~/media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/iot%20value%20set%20to%20accelerate%20through%202030%20where%20and%20how%20to%20capture%20it/the-internet-of-things-catching-up-to-an-accelerating-opportunity-final.pdf

Mejia, N. (2020). *Artificial intelligence at HSBC—2 Use-Cases*. https://emerj.com/ai-sector-overviews/artificial-intelligence-at-hsbc

Michalski, G. (2009). Inventory management optimization as part of operational risk management. *Economic computation and economic cybernetics studies and research*, *43*(4), 213–222.

Mithas, S., Tafti, A., Bardhan, I., & Goh, J. M. (2012). Information technology and firm profitability: Mechanisms and empirical evidence. *MIS Quarterly*, *36*(1), 205–224. https://doi.org/10.2307/41410414

Musicant, O., Bar-Gera, H., & Schechtman, E. (2010). Electronic records of undesirable driving events. *Transportation Research Part F: Traffic Psychology and Behaviour*, *13*(2), 71–79. https://doi.org/10.1016/j.trf.2009.11.001

Nagali, V., Hwang, J., Sanghera, D., Gaskins, M., Pridgen, M., Thurston, T., Mackenroth, P., Branvold, D., Scholler, P., & Shoemaker, G. (2008). Procurement risk management (PRM) at Hewlett-Packard company. *Interfaces*, *38*(1), 51–60. https://doi.org/10.1287/inte.1070.0333

Olsen, T. L., & Tomlin, B. (2020). Industry 4.0: Opportunities and challenges for operations management. *Manufacturing & Service Operations Management*, *22*(1), 113–122.

Toledo, T., Musicant, O., & Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behavior. *Transportation Research Part C: Emerging Technologies*, *16*(3), 320–331. https://doi.org/10.1016/j.trc.2008.01.001

Osborne, J. W. (2000). Advantages of hierarchical linear modeling. *Practical Assessment, Research, and Evaluation*, *7*(1), 1.

Paefgen, J., Staake, T., & Fleisch, E. (2014). Multivariate exposure modeling of accident risk: Insights from Pay-as-you-drive insurance data. *Transportation Research Part A: Policy and Practice*, *61*, 27–40.

Rai, A., Patnayakuni, R., & Seth, N. (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS Quarterly*, *30*(2), 225–246. https://doi.org/10.2307/25148729

Rai, A., & Tang, X. (2014). Research commentary—Information technology-enabled business models: A conceptual framework and a coevolution perspective for future research. *Information Systems Research*, *25*(1), 1–14. https://doi.org/10.1287/isre.2013.0495

Razi-Ardakani, H., Mahmoudzadeh, A., & Kermanshah, M. (2019). What factors results in having a severe crash? A closer look on distraction-related factors. *Cogent Engineering*, *6*(1), 1708652. https://doi.org/10.1080/23311916.2019.1708652

Rawat, P., Singh, K. D., Chaouchi, H., & Bonnin, J. M. (2014). Wireless sensor networks: A survey on recent developments and potential synergies. *The Journal of Supercomputing*, *68*(1), 1–48. https://doi.org/10.1007/s11227-013-1021-9

ReportLinker. (2020). *The usage-based insurance market is projected to reach USD 125.7 billion by 2027 from an estimated USD 24.0 billion in 2019, at a CAGR of 23.0%*. https://www.globenewswire.com/news-release/2020/05/13/2032379/0/en/The-usage-based-insurance-market-is-projected-to-reach-USD-125-7-billion-by-2027-from-an-estimated-USD-24-0-billion-in-2019-at-a-CAGR-of-23-0.html

Shacklett, M. (2017). *How the Internet of things will transform food supply chains. Food logistics*. https://www.foodlogistics.com/technology/article/12382105/how-the-internet-of-things-will-transform-food-supply-chains

Shore, J., & Johnson, R. (1980). Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, *26*(1), 26–37. https://doi.org/10.1109/TIT.1980.1056144

Simpson, A. G. (2017). *State farm lost $7 billion on auto in 2016*. https://www.insurancejournal.com/news/national/2017/03/01/443224.htm

Sodhi, M. S., Son, B. G., & Tang, C. S. (2012). Researchers' perspectives on supply chain risk management. *Production and Operations Management*, *21*(1), 1–13. https://doi.org/10.1111/j.1937-5956.2011.01251.x

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. In Peter L. Bartlett, Fernando C. N. Pereira, Christopher J. C. Burges, Léon Bottou, & Kilian Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25 (NIPS 2012)*. Morgan Kaufmann Publishers 2951–2959.

Vural, E., Cetin, M., Ercil, A., Littlewort, G., Bartlett, M., & Movellan, J. (2007). Drowsy driver detection through facial movement analysis. In Michael Lew, Nicu Sebe, Thomas S. Huang, & Erwin M. Bakker (Eds.), *Proceedings of the 2007 IEEE international conference on human-computer interaction*, Springer (pp. 6–18).

Xie, K., Ozbay, K., Kurkcu, A., & Yang, H. (2017). Analysis of traffic crashes involving pedestrians using big data: Investigation of contributing factors and identification of hotspots. *Risk Analysis*, *37*(8), 1459–1476. https://doi.org/10.1111/risa.12785

Xie, K., Ozbay, K., Yang, H., & Li, C. (2019). Mining automatically extracted vehicle trajectory data for proactive safety analytics. *Transportation Research Part C: Emerging Technologies*, *106*, 61–72. https://doi.org/10.1016/j.trc.2019.07.004

Xu, G., Liu, L., Ou, Y., & Song, Z. (2012). Dynamic modeling of driver control strategy of lane-change behavior and trajectory planning for collision prediction. *IEEE Transactions on Intelligent Transportation Systems*, *13*(3), 1138–1155. https://doi.org/10.1109/TITS.2012.2187447

WÅhlberg, A. A. (2017). *Driver behaviour and accident research methodology: Unresolved problems*. CRC Press.

Wu, K. F., Aguero-Valverde, J., & Jovanis, P. P. (2014). Using naturalistic driving data to explore the association between traffic safety-related events and crash risk at driver level. *Accident Analysis & Prevention*, *72*, 210–218.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ho, Y.-J. (I)., Liu, S., Pu, J., & Zhang, D. (2022). Is it all about you or your driving? designing IoT-enabled risk assessments. *Production and Operations Management*, *31,* 4205–4222. https://doi.org/10.1111/poms.13816