Avoiding Stress Driving: Online Trip Recommendation from Driving Behavior Prediction

Rohit Verma, Bivas Mitra, Sandip Chakraborty
Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, INDIA 721302
Email: rohit.verma@iitkgp.ac.in, bivas@cse.iitkgp.ac.in, sandipc@cse.iitkgp.ac.in

***Abstract*—The growth in the market for cab companies like Uber has opened the door to high-income options for drivers. However, in order to boost their income, drivers many a time resort to accepting trips which increases their stress resulting in poor driving quality and accidents in serious cases. Every driver handles stress differently and the trip recommendation thus needs to be on a personalized level. In this paper, we explore historical trip data to compute the driving stress and its impact on various driving behavioral features, captured through vehicle-mounted GPS and inertial sensors. We utilize a Multi-task Learning based Neural Network model to learn both the common features and the personalized features from the driving data to predict the stress level of a driver. We further establish a causal relationship between the stress level of a driver and his driving behavior. Finally, we develop a trip recommendation system for cab drivers to avoid stress driving. The models have been tested over both a publicly available dataset with 6 drivers for 500 minutes of driving data and an in-house collected dataset from 8 drivers over 1700 trips for 5 months. We observe that the proposed model gives an average prediction accuracy of 94% with low false-positive rates. We also observed that the driving behavior is improved when a driver takes a recommended trip.**

***Index Terms*—Driver Trip Recommendation; Driving Stress; Driving Behavior; Multi-task Learning; Causality Analysis**

I. INTRODUCTION

The huge surge in the number of drivers joining ride-sharing companies like Uber, Ola, Lyft, etc. can be linked to the high-income opportunities these farms provide [1]. However, the drivers, in order to further boost their income, resort to taking more trips or high paying trips neglecting how this is affecting their driving quality. This trend is more prominent in developing countries where stress acts as a major factor for poor driving behavior and compromising road safety [2]. Various recent studies revealed that factors like stress-driving, road rage, driving under fatigue or poor mental conditions etc. are predominant among the ride-sharing and taxi cab drivers across low-income and middle-income countries [3]. In this context, it is essential to develop a system that could assist the drivers & the farms to decide whether to take the next trip depending on the current stress level of the driver.

Although the relationship between driver stress and driving behavior has been studied before [4], [5], the existing literature mostly relies on physiological sensors or surveys. These methods are quite obtrusive and difficult to be used for a long period. Moreover, physiological sensors are costly for mass-scale deployments over the drivers at low-income and the middle-income countries. On the contrary, spatial and inertial sensor data are widely available – cars are in-general GPS enabled; even a smartphone can provide sensor data. This opens up the opportunity to study driving data under different road and driving conditions and to correlate the stress profile of a driver with corresponding driving behavior.

However, correlating stress profile of a driver and driving behavior from historical data of driving trajectories has multiple challenges. First of all, several direct and indirect factors, such as driving schedule, road & traffic conditions contribute to the stress profile of the driver. Hence, quantifying stress as well as driving behavior, which are subjective in nature, is important. Second, personality traits have a direct impact on the stress profile of a driver [6]; therefore, the same level of workload would impact multiple drivers differently. A generic model may not be suitable for all the drivers, and behavioral patterns of the individuals need to be learned. Moreover, driving behavior may vary among drivers even for a similar driving environment; for example, a driver may engage in rash driving only under stress, whereas rash driving may be a general behavior for another driver.

In this paper, we rely on a large pool of spatio-temporal driving dataset to develop a system which predicts the impact of past driving attributes on the stress level of a driver and his driving behavior under various driving and environmental conditions (§IV). We first conduct an interview based survey to understand the effect of various parameters which results in driving stress (§III). Subsequently, we develop a model which computes driver's stress based on the driving schedule roster and his personality traits (§V). In order to include various individual-centric factors, we implement a neural network learning approach based on multi-task learning [7]. Simultaneously, we develop a mechanism to score the driver's driving behavior from the recorded sensor data (§VI-A). Finally, we develop a model to predict the driving behavior at any later stage from the calculated driver stress; this model is based on the detailed correlation as well as causality study which depicts that driver stress impacts driving behavior (§VI-D). Finally, leveraging on this relationship, we develop a recommendation system to suggest if the driver should take a trip or not (§VII). We evaluate our system over UAH-DriveSet data [8] and an in-house driving dataset collected for a period of five months (§IV) from 8 drivers ($\approx$ 1700 trips) using a custom-made device deployed on various cabs.

II. Related Work

The need to define a relationship between stress and driving behavior has been a buzz among the researchers for quite a time in the medical field. In this section, we present a brief survey of the existing literature broadly focusing on four aspects of our work – (a) stress modeling, (b) driver behavior identification, (c) relating stress and driving behavior, and (d) trip recommendation system.

Stress Modeling: As early as 1900; several works tried to use physiological sensors to calculate stress [9]. Rise in the number of accidents directed researchers to explore physiological data to quantify various stress parameters of drivers using methodologies such as correlation analysis [10], pattern recognition [11], using machine learning techniques [12]. However, their reliance on physiological sensors or personal surveys made them obtrusive.

Driver Behavior Analysis: In the later 1980s, driving behavior addressed as an extension of the problem-behavior theory [13]. Increasing number of vehicles made the problem a generalized one [14]. Recent works try to model and provide a score to the driving behavior for safe driving [15], [4]. In [16], [8], the authors have designed machine learning based techniques to identify several abnormal driving behaviors like weaving, sudden brakes, etc. and provide a score to the drivers.

Relationship between Stress and Driving Behavior: Various works [17], [18] have explored the reasons for the abnormal driving. In [19], the authors have shown that drowsy driving is one of the major factors behind road fatalities. Scott-Parker *et al.* [5] have attributed stress as a major factor behind abnormal driving. However, to the best of our knowledge, no existing works have developed a predictive model to correlate driving behavior with the stress profile considering hidden factors, such as the personality traits and the driving environments.

Trip Recommendation Systems: Existing trip recommendation systems provide preference to passenger needs and assign a trip to the driver. Some works observe which are the best matching driver/commuter pairs [20], [21], while others [22] provide algorithms to assign chain trips to drivers. [23] moves one step ahead and scores the driver after assigning a trip. However, none of these systems address if assigning the trip would adversely affect the driver. Consequently, in this paper, we devise a recommendation system which based on a driver's personality traits, decides how a trip would affect his driving behavior and recommends him to accept/reject.

III. Motivation and Challenges

The major motivation of this work comes from the need of analyzing driver stress due to workload, which is highly linked to road safety and the need to stop drivers from taking risks. In order to understand the different intricacies such as effect of workload and the type of trips a driver prefers, we performed a survey over 30 drivers in different cities of India. The survey results are given in Fig 1. The major takeaway from the survey is that most of the drivers are stressed when they have to take long distance trips, followed by long trip time and low rest time (Fig. 1(c)). This brings out the personalization aspect of driving stress as different drivers are affected by diverse factors. Fig. 1(d) shows that there are drivers who prefer to take long trips while almost similar percentage of drivers prefer to take multiple short trips instead of a single long trip. Furthermore, as Fig. 1(e) suggests, there are both classes of drivers; one who prefer to drive more at night to boost their income and other who would rather take more day trips. It is evident from Fig. 1(f) that a significant fraction of drivers accept a trip when assigned, but later regret for taking it as it was more stressful for them and affect their driving. Fig. 1(d,e,f) highlights the need of developing a system to recommend if a driver should accept a trip or skip it based on their past driving history.

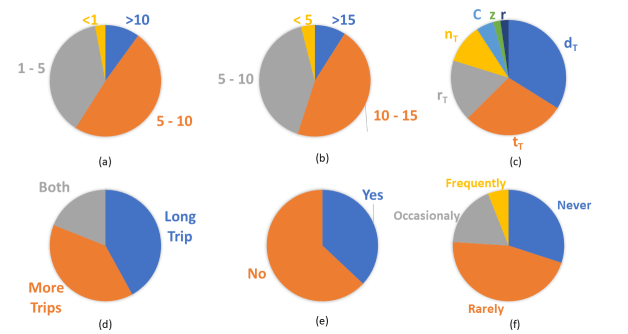


Fig. 1. Driver Survey Report (a) Years of Experience (b) Avg. daily trips (c) Features (Table II) which affect the driver's driving most (d) If a driver prefers a long trip or more short trips (e) If the driver prefers to drive more at night instead of day for extra commission? (f) Regretted taking a trip

From the aforesaid observations, we infer that different drivers are stressed by various factors and they prefer to avoid stressed trips. Therefore, we need a trip recommendation system for avoiding stress driving. However, developing such a system has a number of research challenges as follows. (1) In order to develop a non-obtrusive system, we need to avoid the physiological sensors for stress calculation. (2) Stress relies heavily on the individual's personality traits. Traditional machine learning approaches fail to account for such individual differences. (3) Although we have the stress model along with the driving score, a mere correlation between the two does not imply that a particular driving behavior is caused by the driver's stress. Therefore, a methodology needs to be designed to infer a causal relationship between the two. (4) Rejecting trips in any system incurs some penalty for a driver which requires to be addressed as a multi-objective problem.

IV. Dataset Description and System Overview

In this section, we give a brief overview of the datasets used in the study along with the overall system model.

A. Datasets

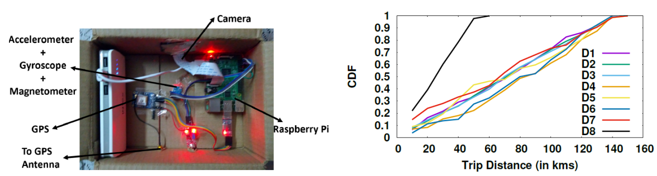
In this paper, we conduct the experiments on two datasets – (i) UAH-DriveSet data [8] and (ii) an in-house dataset

TABLE I
DRIVER AND CAR DETAILS FOR IN-HOUSE DATASET

Driver	Age Range	Vehicle Model	Fuel Type	Number of Trips
D1	20-30	Tata Indigo	Diesel	282
D2	20-30	Hindustan Ambassador	Diesel	321
D3	20-30	Hindustan Ambassador	Diesel	325
D4	30-40	Tata Sumo	Diesel	72
D5	30-40	Maruti Suzuki Dzire	Diesel	67
D6	20-30	Tata Indigo	Diesel	80
D7	30-40	Maruti Suzuki Dzire	Diesel	75
D8	20-30	Maruti Suzuki Alto	Petrol	405

collected for 8 drivers, over a period of five months for ≈ 1700 trips. The brief description of these two datasets follows.

1) UAH-DriveSet: The UAH-DriveSet data [8] is a public dataset captured by the driving monitoring app developed by the authors. The data were collected for five male drivers and one female driver in the age group $20-50$ over two different routes in Spain – (i) 25 km round-trip in a motorway type of road with 3 lanes in each direction and 120 km/h of maximum allowed speed, and (ii) 16 km round-trip in a secondary road with one lane in each direction and around 90 km/h of maximum allowed speed. The data contains both inertial sensors and GPS data along with a video clip for all the trips. The complete dataset has over 500 minutes of driving data. The data is labeled with three different driving behaviors – normal, drowsy and aggressive.



(a) (b)

Fig. 2. (a) Data collection device, (b) Route length distribution for all the drivers for the collected in-house dataset

2) In-house Dataset: Apart from the UAH-DriveSet data, we also collect a large pool of spatio-temporal driving data for 8 drivers over five months using custom-made sensing device deployed over the cars of two local cab companies (2 cars and 5 cars respectively) and one personal car. The data has been collected from a natural and uncontrolled environment, where the cab companies rent the cabs based on the requisitions from the customers. However, a single car is always driven by a single driver (details in Table I). The device developed for data collection is shown in Fig. 2(a) which contains inertial sensors like accelerometer, gyroscope and magnetometer, GPS sensor and a camera mounted on a Raspberry Pi device connected with external power supply through an 11000 mAh power bank. This custom-made device was mounted on the car dashboard and it continuously collects timestamped data while the car is on a trip. The cars were driven by 8 male drivers of age group 20–40; the details of the drivers and cars are given in Table I. We have collected data for around 1700 trips over 15 major routes and their sub-routes covering highways, urban and rural roads. The distribution of the route lengths is given in Fig. 2(b).

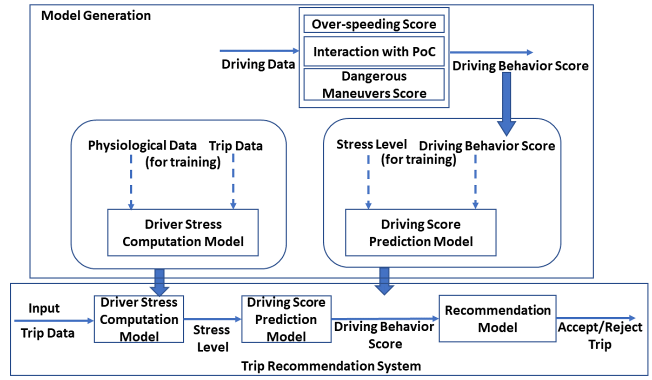


Fig. 3. Block diagram of the developed system

B. System Overview

The proposed system can be divided into four broad modules as shown in Fig. 3.

Driver Stress Model: In this module, we assert that driver's trip roster has a major impact on his stress; hence, we aim to model the stress levels based on the roster-based factors such as the number of trips, the travel time, the rest time, etc. Furthermore, to address personality traits of different drivers, we develop a personalized stress model for each driver.

Driving behavior model: We model the driving behavior based on three aspects – (a) over-speeding instances, (b) interactions with road anomalies like speed breakers and potholes and (c) frequency of dangerous maneuvers [24] such as sudden brakes, sharp turns, side-slipping, etc.

Predicting driving behavior from driver stress: We perform both correlation and causality analysis over driver stress and driving behavior to infer their relationships. Following this, we develop a machine learning based regression model to predict the driving behavior score from the driver stress model.

Trip recommendation: Finally, we use the driving behavior prediction model to predict the impact of the driving stress over the next possible trip of the driver. Whenever a new trip arrives, the trip details are fed to the stress model to compute the stress level considering the trip to be taken, and this is used to predict the driving behavior score. Based on this score, we recommend the driver to take or not to take the trip to ensure driving safety. The following sections provide a detailed description of each of these modules.

V. DRIVER STRESS MODEL

In this section, we construct a model to compute driver stress based on the historical trip information captured through various sensing data. We extract trip related features given in Table II from the collected sensor data and compute stress for the respective drivers. The model classifies driver stress into three categories, viz., *no stress*, *medium stress*, and *high stress*. Notably, the proposed model is sophisticated to capture the personality traits of the individual drivers.

TABLE II
BASE FEATURES USED FOR THE MTL-NN MODEL.

Feature	Description
No. of trips (n_T)	Number of trips the driver has covered including the current one.
Trip time covered (t_T)	Time for which the driver was driving starting from the first trip.
Trip distance covered (d_T)	Distance for which the driver was driving starting from the first trip.
Rest time (r_T)	Time for which the driver had taken rest after the last trip.
Time of the day (z)	Divided into 4 time zones. 6 AM -10 AM(0), 10 AM- 4 PM(1), 4 PM - 10 PM(2), 10 PM - 6 AM(3)
Congestion ($\mathcal{C}$)	Calculated from the trajectory data using existing models [25]
Road Type (r)	City (0), Highway (1), Rural (2). If multiple road types are on the same trip, then the score is calculated as the weighted average over the distance for which each type of road was driven on.

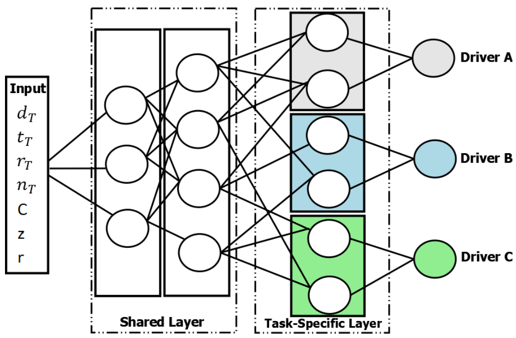


Fig. 4. MTL-NN architecture for personalized stress modeling

A. Model Development

Effectively, the model takes the trip information of the driver as input (given in Table II) and computes the driver's stress label. We develop the driver's stress model using the *Multi-task learning* (MTL) strategy [7]. Here computing the stress level of each driver is considered as *one task* in the multi-task model. The objective of the model is to conduct a robust learning by (a) *shared learning*: learning features of one driver (one task) using the related features of other drivers (related tasks), and (b) *task-specific learning*: side by side, some part of the model allows to become specialized to compute the stress of the specific driver. This part handles the personality traits by considering each driver as a separate task.

Fig. 4 depicts the multi-layered architecture of *Multi-task learning Neural Network* (MTL-NN). The input layer contains the feature vector, obtained from the trip information, which is fed into the network. The next layer is the *shared layer* which contains a set of hidden nodes; the parameters of these nodes are shared across other nodes of this layer for all the tasks. This shared layer enables inductive transfer which improves learning for one task (say, the stress profile of one driver) by using the information contained in the training signals of other related tasks (say, the stress profiles of other drivers). This improves the overall model performance since some features may be easy to learn for driver A while being difficult to learn for another driver B. This might occur since the driver B interacts with those features in a more complex way than driver A. The shared layer allows the model to eavesdrop from driver A, and learn the features for driver B. On the other hand, MTL-NN also allows few hidden nodes to become specialized for computing stress of just one driver (i.e. specialized in one task); this personalized stress computation, by capturing the characteristics of the specific driver, is carried out in the final *task-specific layer*. In this layer, stress computation of one specific driver can ignore the hidden nodes connected to other drivers, by keeping the weights connected to them small, as they do not appear useful. In this layer, the learning mechanism maps the generalized information learned at the *shared layers* to a final prediction personalized by the characteristics of the specific driver (task).

B. Model Training

We train the MTL-NN model over the HCILab dataset [26]. This dataset contains driving information of 10 drivers (3 female and 7 male), between 23 – 57 years of age, covering four trips each. The trip duration ranged from 2 minutes to 35 minutes. The collected driving data includes acceleration along the three axes, light, and GPS coordinates, speed, altitude, and bearing. Additionally, the dataset also records timestamped physiological sensor data, like electrocardiogram (ECG), skin conductance rate, body temperature, heart rate and the heart rate variability, for each driver. We extract the base features from the trip details, as summarized in Table II. Additionally, we leverage on the multiple derived features computed from statistical properties such as mean, variance, kurtosis and 80^{th} percentile calculated on those base features. Finally, a feature set constituting of a total of 27 features is built, which along with the loss function for softmax regression [27] is used to train our model. It can be noted that the physiological sensor data are used in this work just to generate the ground truth levels, and are not used for actual prediction model.

*1) **Ground Truth Stress Labels**:* We leverage the model proposed in [28] to compute the driver stress from the recorded physiological data, which is used for ground truth generation for the HCILab dataset. The model classifies driver stress into three classes, *no stress*, *medium stress*, and *high stress*. It is built upon eight standard ECG features, viz., *average QRS, RR, QQ, SS, QR, RS intervals*, *average heart beats* and *average difference beats* derived from the physiological sensors available in the PhysioNet dataset [29]. The PhysioNet dataset contains the data of different types, viz., time stamp, electro-cardiogram, EMG, foot GSR, hand GSR, heart rate and respiration rate, collected for 17 healthy drivers in Boston area. The data was taken using sensors that the drivers wore driving on prescribed routes including city streets and highways. The trips were of duration ranging from 25 – 93 minutes. Since the window size was not specified in [28], we determine it empirically based on existing literature [30] to obtain identical results. Based on the empirical results we implement a sliding window (of window size 15 seconds and shift of 500 ms) to compute the features and assign ground truth labels for each time window. Similar to [28], we implement the decision tree algorithm to classify driver stress labels from physiological data in PhysioNet dataset [29] obtaining 86% accuracy with a train-test split of 90%–10%. The 10% holdout is drawn randomly from the data and hence belong to multiple windows. We use this (validated) model to generate the ground truth stress label for each driver in the HCILab dataset. For that, we extract the features from physiological data available in the HCILab dataset for each time window and compute the respective stress label.

C. Model Evaluation

We evaluate the proposed MTL-NN model for stress calculation on the HCILab dataset. We calculate the features provided in Table II in each time window (15 seconds and 500 ms slide) and trained the model using ground truth stress labels generated using the PhysioNet dataset. The data is divided into non-overlapping set of 80% and 20% for training and testing respectively. Moreover, the training portion is divided as 60% and 20% of data for training and validation, respectively. Simultaneously, we implement a *Single Task Learning* (STL) based neural network model to train each driver in isolation, which we consider as the baseline for evaluation, where personality traits are not captured. Considering the unbalanced dataset (as in general, most of the data points are in non-stress class), we calculate the results for the *Area Under the Curve* (AUC) metric. Since we provide the multi-class result, we examine a forced binary classification method using the *one versus all* method, i.e, considering *no stress* as one class and the other two combined as the second class; similarly, for the medium-stress and the high-stress, we consider the other two combined as the second class. We then plot the ROC for all these separate scenarios and give the AUC result aggregated over the number of classes. We observe that the MTL-NN approach has an AUC of 0.931 which is a significant improvement compared to the STL approach where we get an AUC of 0.794. Fig. 5 illustrates the high prediction accuracy

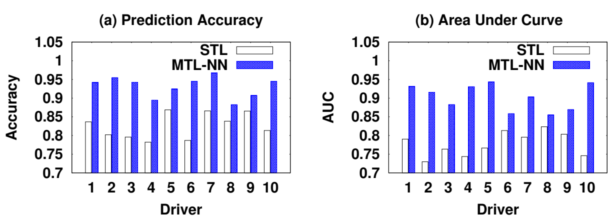


Fig. 5. Driver specific results for STL and MTL-NN

and the AUC for all the 10 drivers demonstrating the utility of the MTL-NN model in personalization.

TABLE III
COMPARISON WITH SYSTEMS WHICH RELY ON PHYSIOLOGICAL SENSORS

Model	Area Under Curve
Salai et. al. [30]	0.75
Shi et. al. [31]	0.62
Singh et. al. [32]	0.94
Our Model	0.93

In-order to check if this system fares similar to some of the existing systems which rely on physiological sensor data

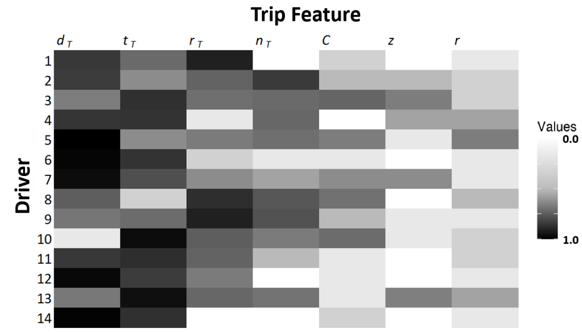


Fig. 6. Heat map showing which features affect a drivers stress computation most based on Sobol indices

only, we compare our model with three existing systems. The first work by Salai *et. al.* [30] develops a simple stress detection algorithm using only time-domain HRV features. The second work by Shi *et. al.* [31] builds a personalized stress detection algorithm using SVM. They rely on heart rate, ECG, skin conductance and respiration and temperature related features. The final work which we compare with is by Singh *et. al.* [32] and develops a neural network based model for stress computation. In order to compare the systems, we use our model to compute stress using driving data, whereas for the competing models, we use the physiological sensor data. Table III gives the AUC result,calculated same as before, for all the four models. We observe that our model fares better than [30] and [31] and gives almost similar result as [32]. This ensures that the idea to detect stress using driving data is equally effective as using physiological sensor data.

In Table IV, we rank the base features based on the classification importance by performing the sensitivity analysis using *Sobol Total Order Indices* [33]. A confidence score below 10% implies that the sample size provided is sufficient for the analysis and the measured indices are significant. Notably, the distance covered impacts driver stress most, followed by the driving time, and the rest time, as we observe from the total order indices value.

TABLE IV
SOBOL TOI FOR THE BASE FEATURES. TOI: TOTAL ORDER INDICES, TOC: TOTAL ORDER CONFIDENCE

Feature	d_T	t_T	r_T	n_T	$\mathcal{C}$	z	r
TOI	0.942	0.897	0.813	0.758	0.699	0.504	0.448
TOC (%)	3.1	3.4	2.9	3.3	3.1	2.8	2.6

The major advantage of the proposed stress model is its ability to understand a drivers personality traits. Fig. 6 shows the effect of the multiple features on the driver stress, calculated using Sobol total order indices [33]. The heat map shows considerable variation of feature importance across the drivers. For instance, (i) time-zone (z) has little importance on drivers $D6$ or $D8$ (ii) driver $D14$ is impacted by a small subset of the features (iii) whereas driver $D2$ or $D3$ are impacted significantly by all the features.

VI. Predicting Driving Behavior from Stress

Stress leads to cognitive distraction of a driver which significantly impacts his concentration and affects driving behavior. Our next objective is to predict the driving behavior from the stress level of a driver. For this purpose, we first define a quantitative measure of the driving behavior, and then develop a model to predict the driving behavior score from the stress level by establishing a causal relationship between the two.

A. Driving Behavior Score

We quantify the driving behavior of a driver from the recorded sensor data and assign a driving behavior score. We concentrate on the following three aspects while quantifying driving behavior.

Speed Profile: All the countries have statutory speed limits ($\mathcal{L}$) with an allowable tolerance ($\mathcal{T}$) defined for cars plying on different types of roads [34]. For instance, in United Kingdom, the speed limit is 48, 96 and $112kmph$ respectively for built-up areas, single-carriageways, and motorways with a tolerance level of 10%. In our scoring, we apply the country-specific speed limits and the tolerance level to categorize the vehicle speed s, obtained from the GPS sensor, into one of the three classes – *safe*, *moderate* and *dangerous*. The speed profiling score ($\mathcal{V}$) is computed as follows.

$$\mathcal{V} = \begin{cases} 0, & \text{s} < \mathcal{L} \text{ (safe)} \\ (\text{s - } \mathcal{L})/\mathcal{T}, & \mathcal{L} \leq \text{s} < (\mathcal{L}+\mathcal{T}) \text{ (moderate)} \\ 1, & \text{s} \geq (\mathcal{L}+\mathcal{T}) \text{ (dangerous)} \end{cases}$$

Interaction with PoCs: Driving behavior can be assessed by observing driver's interaction with speed breakers and potholes; sudden and sharp jerk indicates careless and abnormal driving. We detect a speed breaker from acceleration signals by implementing the standard machine learning techniques [25]. In order to detect a pothole, we implement the algorithm proposed in Wolverine [35]. Once we detect a speed breaker or a pothole, we measure jerk as the rate of change in acceleration along the z-axis, calculated as $\mathcal{J} = da(t)/d(t)$ (where $a(t)$ is the acceleration at time t along the z-axis of the sample). We designate the driving as abnormal if the calculated jerk qualifies as *critical jerk* [36]. We assign the interaction with PoC score $\mathcal{I}$ as 1 if the jerk is critical ($\mathcal{J} \leq -9.9m/s^3$) and 0 otherwise;

Dangerous Maneuvers: We extract six types of dangerous maneuvers based on [16] to determine the driving score – (a) *Weaving*: repeated lane changing at high speed, (b) *Swerving*: change direction abruptly, (c) *Side-slipping*: deviating from the proper driving direction, (d) *Fast U-turn*: suddenly taking a U-turn, (e) *Sharp turn* and (f) *Sudden brake*. The method in [16] extracts unique signatures of these maneuvers from the acceleration and orientation data and then computes a set of features. The authors have used both SVM and NN based techniques . The results of both the techniques were more or less similar, and hence, we employ the SVM based technique for our analysis. We implement an SVM (with 16 features, which are statistical and mathematical derivations of the acceleration and orientation values in different axes, as given in [16]) to detect if a dangerous maneuver has occurred. This model yields a tuple $\mathcal{M}$ of size six to store the score for each of the detected maneuvers. A score of 1 is assigned if a particular maneuver occurs in a sample window and 0 otherwise.

TABLE V
Accuracy In Assigning Scores (NA - Behavior Wasn't Observed)

Scenario	P	R	AUC	P	R	AUC
	in-house Dataset			UAH-Driveset		
Speed Profiling	0.98	1.00	0.98	0.98	1.00	0.98
PoC Interaction	0.93	0.92	0.87	0.94	0.94	0.88
Weaving	0.97	0.93	0.92	0.92	0.95	0.89
Swerving	0.95	0.97	0.88	0.93	0.93	0.90
Side-slipping	0.97	0.97	0.91	0.96	0.97	0.94
Fast U-Turn	0.91	0.91	0.89	NA	NA	NA
Sharp Turn	0.98	0.91	0.91	0.95	0.91	0.85
Sudden Brake	0.94	0.97	0.87	0.96	0.93	0.85

Overall Score: We calculate the overall driving behavior score as follows, considering that all the aforementioned aspects exhibit equal importance on the driver's driving behavior.

$$\mathbb{D} = \frac{1}{8}\sum\left(\mathcal{V}+\mathcal{I}+\sum_{i=1}^{6}\mathcal{M}_i\right) \tag{1}$$

Evidently, we assign a low score for the smooth driving and high score for the aggressive driving behavior.

We validate the overall driving behavior score $\mathbb{D}$ by individually evaluating each of the aforesaid three aspects $\mathcal{V}$, $\mathcal{I}$ and $\mathcal{M}$. For speed profiling, we evaluate whether our score properly detects the speed violations, depending on the country and the road type. In order to evaluate the interaction with PoCs (speed-breakers, potholes), we first check (a) if the landmarks are detected correctly, and subsequently validate (b) if the *critical jerks* are also correctly classified. Finally, we evaluate the correctness of identifying the dangerous maneuvers. We perform the evaluation experiments on both the datasets. The UAH-DriveSet provides files enumerating driver over-speeding, weaving, sudden brakes, etc as ground truth. For in-house dataset, we rely on the captured video clips to tag the ground truth information. Table V exhibits the evaluation results in terms of precision (P), recall (R) and the AUC metrics for all the three aspects. This is comforting for us to observe that the computed score is able to capture the diverse driving behaviors efficiently.

B. Correlating Driver Stress and Driving Behavior

We leverage on Kendall's tau coefficient [37] to conduct the correlation analysis of driver stress and driving behavior. In order to calculate the coefficient, we first compute the concordant and discordant pairs. A pair of observations $(\mathbb{S}_i, \mathbb{D}_i)$ and $(\mathbb{S}_j, \mathbb{D}_j)$ are concordant if $\mathbb{S}_i > \mathbb{S}_j$ & $\mathbb{D}_i > \mathbb{D}_j$ or $\mathbb{S}_i < \mathbb{S}_j$ & $\mathbb{D}_i < \mathbb{D}_j$. In the same vein, the pairs are said to be discordant if $\mathbb{S}_i > \mathbb{S}_j$ & $\mathbb{D}_i < \mathbb{D}_j$ or $\mathbb{S}_i < \mathbb{S}_j$ & $\mathbb{D}_i > \mathbb{D}_j$. Kendall's tau coefficient is calculated as $\tau = \frac{n_c - n_d}{n(n-1)/2}$, where, n_c and n_d are the number of concordant and discordant pairs respectively

Fig. 7. Kendall's tau coeff. between stress and driving behavior

in the dataset, and n is the total number of samples. We obtain the mean correlation coefficient τ as 0.83 between driver stress $\mathbb{S}$ and driving behavior $\mathbb{D}$ for all the drivers in both the datasets (with p-value $2.99x10^{-10}$). Moreover, Fig. 7 depicts the high coefficient τ for each individual driver (with p-values in the order of 10^{-10}). Once we demonstrate the correlation, we next aim to establish the causal relationship between the stress and the driving behavior.

C. Causality Analysis

We apply the standard causal inference techniques [38] to investigate the causal relationship between the driver stress and the driving behavior. We define driver stress $\mathbb{S}$ as the *treatment* variable (considered as intervention) and the driving behavior score $\mathbb{D}$ as the *response* to the treatment. In order to claim that $\mathbb{S}$ influences $\mathbb{D}$, (a) $\mathbb{S}$ should always precede $\mathbb{D}$ and (b) there should be no other variable $\mathbb{Z}$ which could be a possible cause for $\mathbb{D}$; $\mathbb{Z}$ is known as the *confounding variable*. Since driver stress results in the driving behavior, hence trivially $\mathbb{S}$ always precedes $\mathbb{D}$. We now concentrate on the candidate confounding variables and demonstrate their impact on $\mathbb{D}$.

*1) **Confounding Variables**:* By definition [39], confounding variables should impact both treatment variable (stress) and response variable (driving behavior). We identify four candidate confounding variables $\mathbb{Z}$ which can be measured from our dataset.

Weather condition ($\mathbb{W}$): We invoke the OpenWeatherMap Weather API [40] to get the weather information at the time of data collection. We provide scores for rainfall volume, temperature, snow and wind-speed on a three-point scale (0 as low and 2 as high). For instance, considering India, rainfall of $< 5mm$ is considered low, $> 5mm$ and $< 10mm$ is considered moderate and $> 10mm$ is considered high [41]. The final score is computed as the aggregated sum of all the four scores.

Previous driving score ($\mathbb{P}$): The last driving behavior sometimes impacts the driver's current behavior. We obtain the score from the previous $\mathbb{D}$ and consider it as a confounding variable.

Day of the week ($\mathbb{Q}$): Whether the driver is driving on a weekday or weekend might impact his driving behavior. A score of 0 is assigned for a weekday and 1 for the weekend.

Special occasion ($\mathbb{O}$): We capture if the driving was on a special occasion such as holidays, strikes, festivals etc. Any of the special occasion days are scored as 1; otherwise, we score it as 0.

TABLE VI
KENDALL'S TAU FOR CONFOUNDING VARIABLES

$\mathbb{Z}$	$\mathbb{W}$	$\mathbb{P}$	$\mathbb{Q}$	$\mathbb{O}$
τ	0.68	0.56	0.18	0.23

In Table VI, we compute Kendall's tau coefficient concerning the driving behavior $\mathbb{D}$ for all the candidate confounding variables $\mathbb{Z}$. We observe that variables such as day of the week and special occasion exhibit low correlation. Hence, we eliminate $\mathbb{Q}$ and $\mathbb{O}$ from the candidate confounding set $\mathbb{Z}$ and perform the following causality inference only on the weather condition ($\mathbb{W}$) and previous driving score ($\mathbb{P}$) along with the driver stress ($\mathbb{S}$).

*2) **Causal Inference**:* The intuition behind defining a causal relationship between the treatment (say, stress $\mathbb{S}$) and the response (say, driving behavior $\mathbb{D}$) variables is that a small change in the treatment variable would cause a significant change in the response variable, whereas no significant change in the response would be observed on changing the confounding variables $\mathbb{Z}$ [42]. This expected change in response variable is conceptualized as the average treatment effect (ATE). However, in order to observe the impact on the response variable, we need the response data both in case (a) when the treatment was performed, and (b) when the treatment was not performed; these are called as matching pairs to remove bias in observational studies [43]. In order to obtain the matching pairs, we identify two drivers (say, u and v) who have similar trip information, but one of them has high stress score (say, $\mathbb{S}(u)$) while the other one has a low score (say, $\mathbb{S}(v)$). This creates a synthetic scenario to obtain the matching pairs, which provide us the responses for both the treatment and the no-treatment cases. Once the set of matching scenarios ($\mathcal{P}$) are generated, we compute the average treatment effect [38] as,

$$ATE = \mathbb{E}\Big(\frac{\mathbb{D}(u) - \mathbb{D}(v)}{\mathbb{X}(u) - \mathbb{X}(v)}\Big)_{(u,v)\in\mathcal{P}} \tag{2}$$

where $\mathbb{E}(.)$ is the expectation, u and v are matching pairs in $\mathcal{P}$ and $X = \mathbb{S} \cup \mathbb{Z}$. The high value of ATE indicates $\mathbb{X}$ (which can be the treatment variable $\mathbb{S}$ or the confounding variables $\mathbb{Z}$) causes driving behavior $\mathbb{D}$.

TABLE VII
ATE OVER $\mathbb{D}$ W.R.T TREATMENT $\mathbb{S}$, $\mathbb{W}$, $\mathbb{P}$

Variable	$\mathbb{S}$	$\mathbb{W}$	$\mathbb{P}$
ATE	0.62	0.23	0.21

Table VII displays the average treatment effect (ATE) for the treatment variable (driver stress $\mathbb{S}$) as well as the confounding variables (weather condition $\mathbb{W}$ and previous driving score $\mathbb{P}$). Evidently, stress $\mathbb{S}$ demonstrates a significant and positive causal impact on the driving behavior $\mathbb{D}$. Notably, the low but positive ATE of $\mathbb{W}$ and $\mathbb{P}$ exhibit their mild impact on $\mathbb{D}$. However, stress exerts a major influence on the deterioration of the driving score, overshadowing the identified confounding variables.

It is also possible to assert that stress can be caused due to driving behavior, thus turning $\mathbb{D}$ as the treatment and $\mathbb{S}$ as the response. But, we cannot calculate the driving behavior unless this action has happened. However, we can calculate stress from historical information as we observed in §V. So, effectively, this analysis would help us to establish a relationship to predict the driving behavior, which would otherwise won't be possible, from the stress profile of a driver. We do not have such a motivation to perform causal analysis the other way round.

D. Model Development: Driving Score Prediction

Finally, we develop a simple machine learning based model to predict the driving behavior score from the driver stress. The model is constructed considering the driver stress level as a single feature. We implement the model following state of the art machine learning algorithms, viz. *Simple Linear Regression* (based on the least square technique [44]), *Linear Support Vector Regression (SVR)*, *SVR with Radial Bias Function (RBF) Kernel)* and *Decision Tree Regression*. We evaluate the model on the UAH-DriveSet and the in-house dataset. We obtain the ground truth driving scores of the drivers from the methodology described in §VI-A. During the evaluation, we split the dataset into the ratio of 80% (to train the model) and 20% (for testing); we perform 5-fold cross-validation and report the results in terms of Mean Square Error (MSE), Mean Absolute Error (MAE) and R^2 score.

We develop a baseline model to demonstrate the elegance of the proposed model in capturing the personality trait. It has been shown qualitatively that an increase in driver stress would deteriorate his driving behavior [5]. Following this idea, we develop a simple linear model as the baseline for driving score prediction. We consider the equation of the linear model as $y = mx$ where y is the driving behavior and x is the driver stress, assuming no intercept. We vary the angle θ, that the line makes with the x-axis, from $0°$ to $89°$ (the line should always lie in the first quadrant) and compute the value of $m = tan\theta$ for which we got the best result. On running the baseline model on the in-house dataset, the best result was obtained for $\theta = 32°$ which we use to calculate m.

Performance Evaluation: Table VIII illustrates the driving score prediction results across different machine learning algorithms for the in-house dataset. Apart from the low MSE, low MAE and high R^2 score, which show the elegance of our model, we note that simple linear regression performs best out of the four machine learning techniques. This is possibly because of the high correlation between driver stress and driving behavior score which could easily be put into the simple linear regression equation. Hence, we implement the simple linear regression to report the subsequent results. In Table IX we exhibit the model performance for the two datasets. The high accuracy of the proposed model can be attributed to the efficient personalized stress computation. Consequently, this model could efficiently learn in what range the driving score of a particular driver varies with the change in the stress level and predict it accordingly. Finally, Tables IX also shows the results of the baseline model which gives significantly poor results when compared to the proposed model. A negative R^2 score implies that the fit of the baseline model is poor than a horizontal line. This can be attributed to the fact that the impact of personality traits has not been considered in the baseline model. The assumption that an increase in stress would always deteriorate the driving performance may not hold true for all the drivers. Hence, the linear model [5] performs poorly as it cannot capture such variations. On the other hand, in our model, when the training is done on the driver data, the personality traits have already been incorporated in the driver stress score; accordingly, the correct driving behavior score can be predicted.

TABLE VIII
REGRESSION TECHNIQUES COMPARISON

Technique	Metrics		
	MSE	MAE	R^2
Simple Linear Regression	0.0103	0.0771	0.8123
Linear SVR	0.0128	0.0971	0.7184
SVR (RBF kernel)	0.0131	0.0985	0.7118
Decision Tree Regression	0.0110	0.0826	0.7861

TABLE IX
METRICS TO EVALUATE THE MODEL

Dataset	Metrics (Proposed Model)			Metrics (Linear Baseline)		
	MSE	MAE	R2	MSE	MAE	R2
in-house	0.0103	0.0771	0.8123	0.376	0.611	-0.05
UAH-DriveSet	0.0105	0.07	0.8571	0.326	0.563	-0.04

VII. ONLINE TRIP RECOMMENDATION

Finally, we leverage on the driver stress computation model as well as the driving score prediction model to develop a system which can recommend a driver whether he should take or skip the next trip. Naively, if the new trip causes high stress to a driver, then that trip shouldn't be recommended. However, it is important to note that some of the drivers are capable to handle high stress efficiently, hence the decision can not be taken only considering driver stress. Thus, we first compute the stress level of a driver using the historic trip information fed into the stress model, and next we use the prediction model (§VI-D) to estimate the driver's driving behavior score for the next trip. If the score is above a threshold, we recommend the driver to skip the trip. We keep the threshold as a configurable parameter, which can be fixed by various cab companies depending on the adopted policy for the companies and the corresponding country or city regulations. For instance, the model could be lenient for a new driver, and hence the threshold would be high. Similarly, some cab firms may require their drivers to atleast complete five trips in a day (for example, the payment boost for *Uber*[45] based on minimum trip requirements)Thus, the recommendation could start only after the fifth trip.

A. Evaluation

We have shared the trip recommendation system (an Android application based on smartphones) with the 7 drivers of the in-house cab company and recorded the results for a week. Whenever a ride was sent to a driver, the ride details were fed to the system. The recommendation system decided whether the driver should take a ride or skip and the information is flashed on the driver's phone. In our experiments, we start recommending only after the third trip and fix the driving behavior score threshold to 0.6.

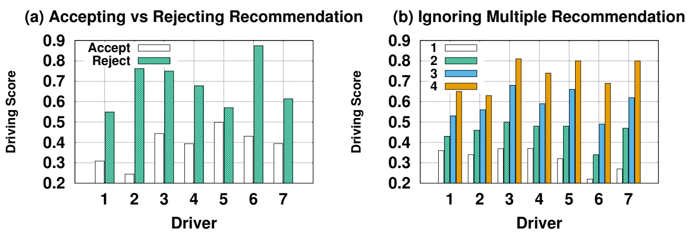


Fig. 8. (a) Average $\mathbb{D}$ on accepting or rejecting a recommendation. (b) $\mathbb{D}$ as a driver rejects one or more recommendations.

Benefit of the trip recommendation system is evident from Fig. 8(a). All drivers observe gain accepting recommendation, which gets reflected from the drop in driving scores. Notably, Driver $D2$ and $D6$ observe high gain, while it is marginal for driver $D5$. We observe that driver $D5$ inherently drives poorly, so the recommendation has marginal effect on him. Moreover, Fig. 8(b) shows that driving score increases (deteriorates) when driver starts rejecting multiple recommendations.

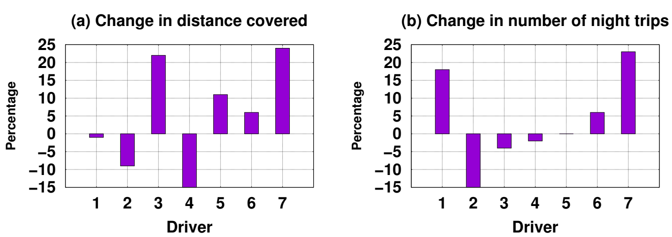


Fig. 9. Avg. % change in d_T and night trips when a recommendation is followed

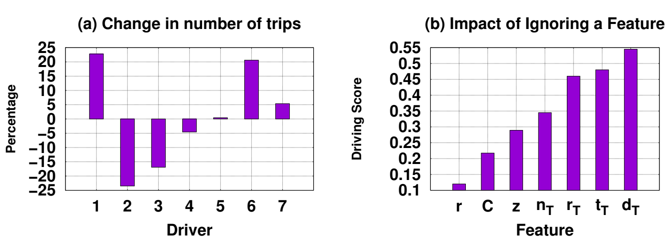


Fig. 10. (a) Avg. % change in n_T when following recommendation. (b) Avg. $\mathbb{D}$ when ignoring recommendation w.r.t. a feature.

Does the recommendation affect a driver's overall income? One primary concern of the proposed trip recommendation is that it may compromise with the driver's income based on the incentive parameters set by the cab company. We surveyed the drivers and noted that the income is primarily based on the number of trips, distance covered and number of night trips conducted by the drivers. We examine how the acceptance of a recommendation affects these aspects. Fig. 9(a) shows the change in distance covered by the drivers on following the recommendations. We observe that the change is less than 25% for all the drivers. Also most of these changes are in positive axis which rather increases the income. This increase can be linked to the decrease in trips more prominent in other features which affects his stress and hence increase in distance covered. A similar trend is observed for number of night trips (Fig. 9(b) and the total number of trips (Fig. 10(a)) also. Such small change in parameters should not adversely affect the driver's income, but maintains a fair balance between driving safety and driver's incentives. However, 20% change in the negative axis for $D2$ (Fig. 9(b) and Fig. 10(a)) and $D4$ (Fig. 9(a)) would seem a major setback for them, but this was expected as safety was the primary concern. $D2$ & $D4$ exhibit poor driving behavior, while taking more number of trips (especially night trips) and more long trips respectively. Finally, in Fig. 10(b) we observe that neglecting the recommendation linked to travel distance (d_T) heavily degrades the driving score. This result also supports the fact observed from Fig.6, where most of the drivers are most stressed by long trips. In a nutshell, we observe that the proposed recommendation system can promote safe driving by balancing the requirements of a driver, a passenger as well as the cab company.

VIII. Conclusion

With gaining popularity of cab-sharing services and the increasing risk-taking attitude of drivers to boost income, a low-cost and easy deployable solution for enforcing safe driving practice has become inevitable. In this paper, we have used historical driving data to develop a personalized model to assess driver's stress from his past driving attributes and trip history. Subsequently, we quantify driving behavior from recorded sensor data. We established a causal relationship between driver stress and driving behavior, which helped us to develop a simple model to predict driving behavior from stress. Leveraging on the stress computation and prediction model, we developed an online trip recommendation system. However, the system could have a plethora of applications. (a) The system can take into account how much driving stress a driver can handle and design a full day roster. (b) To improve driver quality, the system can award drivers who cope better in high stress scenarios. (c) Fleet managers could also use our system to recruit drivers who are on probation.

Although the model has shown high accuracy in quantifying driver stress, there are some confounding variables which are out of our hands and could result in model failure. For instance a driver, who recently faced an accident might get stressed much earlier than expected by the model. We assert that such confounding variables are usually active only for short duration; their role gets disappeared quickly with time.

REFERENCES

[1] "Uber's Ever-Renewing Workforce: One-Fourth Of Its Current U.S. Drivers Joined Last Month," 2018, https://www.forbes.com/sites/ellenhuet/2015/01/22/uber-study-workforce/109d8a5b367a.

[2] P. Stanojević, T. Lajunen, D. Jovanović, P. Sârbescu, and S. Kostadinov, "The driver behaviour questionnaire in south-east europe countries: Bulgaria, romania and serbia," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 53, pp. 24–33, 2018.

[3] Y. Wang, L. Li, and C. G. Prato, "The relation between working conditions, aberrant driving behaviour and crash propensity among taxi drivers in china," *Accident Analysis & Prevention*, 2018.

[4] H. T. Lam, "A concise summary of spatial anomalies and its application in efficient real-time driving behaviour monitoring," in *Proc. of SIGSPATIAL*. ACM, 2016, p. 30.

[5] B. Scott-Parker, C. Jones, K. Rune, and J. Tucker, "A qualitative exploration of driving stress and driving discourtesy," *Accident Analysis & Prevention*, vol. 118, pp. 38–53, 2018.

[6] Y. Ge, W. Qu, C. Jiang, F. Du, X. Sun, and K. Zhang, "The effect of stress and personality on dangerous driving behavior among chinese drivers," *Accident Analysis & Prevention*, vol. 73, pp. 34–40, 2014.

[7] R. Caruana, "Multitask learning," in *Learning to learn*. Springer, 1998, pp. 95–133.

[8] E. Romera, L. M. Bergasa, and R. Arroyo, "Need data for driver behaviour analysis? presenting the public uah-driveset," in *Proc. of ITSC*. IEEE, 2016, pp. 387–392.

[9] J. L. Deffenbacher, D. M. Deffenbacher, R. S. Lynch, and T. L. Richards, "Anger, aggression, and risky behavior: a comparison of high and low anger drivers," *Behaviour research and therapy*, vol. 41, no. 6, pp. 701–718, 2003.

[10] M. Singh and A. B. Queyam, "A novel method of stress detection using physiological measurements of automobile drivers," *International Journal of Electronics Engineering*, vol. 5, no. 2, pp. 13–20, 2013.

[11] J. Healey and R. Picard, "Smartcar: detecting driver stress," in *Proc. of Pattern Recognition*, vol. 4. IEEE, 2000, pp. 218–221.

[12] M. Muñoz-Organero and V. Corcoba-Magaña, "Predicting upcoming values of stress while driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 7, pp. 1802–1811, 2017.

[13] R. Jessor, "Risky driving and adolescent problem behavior: an extension of problem-behavior theory." *Alcohol, Drugs & Driving*, 1987.

[14] R. Bieri, T. Nef, R. M. Müri, and U. P. Mosimann, "Development of a novel driving behavior adaptations questionnaire," *International psychogeriatrics*, vol. 27, no. 6, pp. 1017–1027, 2015.

[15] W. Wang, J. Xi, and H. Chen, "Modeling and recognizing driver behavior based on driving data: A survey," *Mathematical Problems in Engineering*, vol. 2014, 2014.

[16] J. Yu, Z. Chen, Y. Zhu, Y. J. Chen, L. Kong, and M. Li, "Fine-grained abnormal driving behaviors detection and identification with smartphones," *IEEE TMC*, vol. 16, no. 8, pp. 2198–2212, 2017.

[17] Y. Wang, J. Yang, H. Liu, Y. Chen, M. Gruteser, and R. P. Martin, "Sensing vehicle dynamics for determining driver phone use," in *Proc. of MobiSys*. ACM, 2013, pp. 41–54.

[18] "HAMS: Harnessing AutoMobiles for Safety - Microsoft Research," 2018, https://www.microsoft.com/en-us/research/project/hams/.

[19] H. Park, S. Oh, and M. Hahn, "Drowsy driving detection based on human pulse wave by photoplethysmography signal processing," in *Proc. of IUCS*. ACM, 2009, pp. 89–92.

[20] J. Lehmann and D. Sommer, "Computer implemented method for allocating drivers and passengers sharing a trip," Feb. 28 2012, uS Patent 8,126,903.

[21] N. Bicocchi, M. Mamei, A. Sassi, and F. Zambonelli, "On recommending opportunistic rides," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3328–3338, 2017.

[22] C. Lambert, L. Green, F. T. Yoo, M. Haumann, D. Dryjanski, and L. Popov, "Ride chaining," Jun. 13 2017, uS Patent 9,679,489.

[23] Z. Haparnas and T. Marco, "System for navigating drivers to passengers and dynamically updating driver performance scores," Jul. 6 2017, uS Patent App. 14/984,839.

[24] L. Shamoa-Nir and M. Koslowsky, "Aggression on the road as a function of stress, coping strategies and driver style," *Psychology*, vol. 1, no. 01, p. 35, 2010.

[25] R. Verma, S. Ghosh, N. Ganguly, B. Mitra, and S. Chakraborty, "Smart-phone based spatio-temporal sensing for annotated transit map generation," in *Proc. of ACM SIGSPATIAL*. ACM, 2017, p. 16.

[26] S. Schneegass, B. Pfleging, N. Broy, F. Heinrich, and A. Schmidt, "A data set of real world driving to assess driver workload," in *Proc. of AutomotiveUI*. ACM, 2013, pp. 150–157.

[27] D. Heckerman and C. Meek, "Models and selection criteria for regression and classification," in *Proc. of UAI*. Morgan Kaufmann Publishers Inc., 1997, pp. 223–228.

[28] N. Keshan, P. Parimi, and I. Bichindaritz, "Machine learning for stress detection from ecg signals in automobile drivers," in *Proc. of Big Data*. IEEE, 2015, pp. 2661–2669.

[29] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE TITS*, vol. 6, no. 2, pp. 156–166, 2005.

[30] M. Salai, I. Vassányi, and I. Kósa, "Stress detection using low cost heart rate sensors," *Journal of healthcare engineering*, vol. 2016, 2016.

[31] Y. Shi, M. H. Nguyen, P. Blitz, B. French, S. Fisk, F. De la Torre, A. Smailagic, D. P. Siewiorek, M. alAbsi, E. Ertin *et al.*, "Personalized stress detection from physiological measurements," in *International symposium on quality of life technology*, 2010, pp. 28–29.

[32] R. R. Singh, S. Conjeti, and R. Banerjee, "A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 740–754, 2013.

[33] I. M. Sobol, "Sensitivity estimates for nonlinear mathematical models," *Mathematical modelling and computational experiments*, vol. 1, no. 4, pp. 407–414, 1993.

[34] W. H. Organization, *Global status report on road safety 2015*. World Health Organization, 2015.

[35] R. Bhoraskar, N. Vankadhara, B. Raman, and P. Kulkarni, "Wolverine: Traffic and road condition estimation using smartphone sensors," in *Proc. of COMSNETS*. IEEE, 2012, pp. 1–6.

[36] M. Nygård *et al.*, "A method for analysing traffic safety with help of speed profiles," 1999.

[37] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.

[38] B. Lu, E. Zanutto, R. Hornik, and P. R. Rosenbaum, "Matching with doses in an observational study of a media campaign against drug abuse," *Journal of the American Statistical Association*, vol. 96, no. 456, pp. 1245–1253, 2001.

[39] A. Mehrotra, F. Tsapeli, R. Hendley, and M. Musolesi, "Mytraces: investigating correlation and causation between users emotional states and mobile phone interaction," *IMWUT*, vol. 1, no. 3, p. 83, 2017.

[40] "Weather API - OpenWeatherMap," 2018, https://openweathermap.org/api.

[41] N. Deshpande, A. Kulkarni, and K. Krishna Kumar, "Characteristic features of hourly rainfall in india," *International Journal of Climatology*, vol. 32, no. 11, pp. 1730–1744, 2012.

[42] D. B. Rubin, "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of educational Psychology*, vol. 66, no. 5, p. 688, 1974.

[43] ——, "Matching to remove bias in observational studies," *Biometrics*, pp. 159–183, 1973.

[44] E. I. Im *et al.*, "A note on derivation of the least squares estimator," Tech. Rep., 1996.

[45] "Drive with Uber," 2019, https://www.uber.com/en-SG/drive/resources/incentives/.