Missouri University of Science and Technology
Scholars' Mine

Computer Science Faculty Research & Creative Works

Computer Science

01 Jan 2022

Rssafe: Personalized Driver Behavior Prediction for Safe Driving

Bhumika

Debasis Das

Sajal K. Das
Missouri University of Science and Technology, sdas@mst.edu

Follow this and additional works at: https://scholarsmine.mst.edu/comsci_facwork

Part of the Computer Sciences Commons

Recommended Citation

Bhumika et al., "Rssafe: Personalized Driver Behavior Prediction for Safe Driving," *Proceedings of the International Joint Conference on Neural Networks*, Institute of Electrical and Electronics Engineers, Jan 2022.
The definitive version is available at https://doi.org/10.1109/IJCNN55064.2022.9892982

This Article - Conference proceedings is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Computer Science Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

RsSafe: Personalized Driver Behavior Prediction for Safe Driving

Bhumika and Debasis Das
Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur
Jodhpur, Rajasthan, India
{bhumika.1, debasis}@iitj.ac.in

Sajal K. Das
Department of Computer Science
Missouri University of Science and Technology
Rolla, MO, USA
sdas@mst.edu

2022 International Joint Conference on Neural Networks (IJCNN) | 978-1-7281-8671-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/IJCNN55064.2022.9892982

***Abstract*—While the increased demand for taxi services like Uber, Lyft, Hailo, Ola, Grab, Cabify etc. provides livelihood to many drivers, the desire to raise income forces the drivers to work very hard without rest. However, continuous journeys not only affect their health, but also lead to abnormal driving behavior such as rash driving, swerving, side-slipping, sudden brakes, or weaving, leading to accidents in the worst cases. Motivated by the severity of rising accidents and health issues among drivers, this paper proposes a recommendation system, called *RsSafe*, for the safety of drivers. Aiming to improve the driving quality and the driver's experience, *RsSafe* suggests that the driver accepts or rejects the next trip based on the predicted driving behavior. In particular, we propose a fusion architecture that learns to predict the driver's behavior for the next trip using information from multiple streams. This architecture consists of Multi-task Learning with Attention (MTLA) that captures individual drivers' personality traits to deal with the adaptability of system. We use publicly available naturalistic driving behavior analysis dataset, namely the UAHDriveSet, results show that the MTLA predicts with F-measure score of 96% ; and outperforms the baseline as well as state-of-the-art models.**

***Index Terms*—Trip Recommendation, Personalized Driver Behavior, Multi-task Learning, and Driving Behavior Prediction.**

I. Introduction

The current transportation business scenarios lack an adequate system that ensures safe driving recommendations for taxi drivers with maximum hours of service. According to the statistics[1], there was a total of 4,49,002 road accidents claiming 1,51,113 lives in India in 2019. For such road accidents, taxi drivers constitute a significant fraction; and long working hours is one of the major causes of accidents [1]. Taxi companies offer high incentives to complete more trips that encourage taxi drivers to accept frequent trips without taking rest, but this also leads to rash driving behavior. By identifying driving behavior more accurately, such accidents can be reduced from 10% to 20% [2]. Given the rising number of road accidents every year, our motivation in this paper is to provide a safe and accurate driving system that provides an appropriate balance between income and health [3].

There exists research that provides efficient recommendation systems to reduce trip time, waiting time, overall trip cost for customers while minimizing idle time and increasing profit for taxi drivers. However, these recommendation systems pay less attention towards the road safety and drivers' health. In our proposed recommendation system, called *RsSafe*, we address the problem of safety and drivers' health by predicting driving behavior. In the literature, the driving behavior has been extensively studied, revolving around the drivers' attention, intention, behavior, drowsiness, etc. These studies have a common objective of understating the driving status using physiological and psychological data, thus providing a safe driving system. In this paper, we analyze the drivers' behavior for the task of prediction using sensor recorded data. Although the authors in [4]–[7] also predicted drivers' behavior, our proposed recommendation system aims to predict the behavior before the trip instead of during the trip. Moreover, the behavior prediction in existing works is typically done using On-Board Unit (OBU), a communication device mounted on the vehicles and in-vehicles embedded sensors that are difficult to deploy at large scale due to the underlying cost. On the other hand, our proposed recommendation system predicts the drivers' behavior using smartphone sensor recorded data. Research challenges to predicting the driver behavior include: (1) Identifying factors that contribute to the driver's behavior; (2) A generic model that captures individual personality traits; and (3) Sparsity of data. We address these issues by identifying direct and indirect factors based on existing studies [8] contributing to the driver's behavior. For sparsity and generalization, we use a multi-task learning model that learn shared representation among tasks and identify the personality traits of the individual driver collected using smartphone sensors.

The major contributions of this paper are:

- Predict driver behavior before the trip to provide a safe driving system using direct and indirect factors from multi-stream data like the trip, road, weather, dangerous maneuvers, and sensor data.
- Develop an end-to-end Multi-task Learning with Attention (MTLA) based model using smartphone sensor recorded data. This is a novel approach as MTLA with Long Short-Term Memory (LSTM) as underlying configuration has not been widely explored to predict driving behavior from multivariate time-series sensor data.

[1]https://morth.nic.in/road-accident-in-india

978-1-7281-8671-9/22/$31.00 ©2022 IEEE
Authorized licensed use limited to: Missouri University of Science and Technology. Downloaded on February 17,2023 at 20:31:07 UTC from IEEE Xplore. Restrictions apply.

- Recommend driver with the option to accept or reject the next trip based on historical driver behavior.
- Perform ablation study to show the importance and necessity of different features under consideration. From the ablation study, we observe that dangerous maneuver execution significantly affects the driving behavior. Also, validate design choices of MTLA through ablation study and find that MTL with LSTM combination performs better.
- Demonstrate superior performance in terms of F-measure, macro F-measure, micro F-measure, average receiver operating characteristic curve (ROC) as 96%, 92%, 94%, ans 97% for UAH-DriveSet [9] compared to baseline and state-of-the-art methods.

The rest of the paper is organized as follows. Section II gives an overview of related work divided into smartphone based driver behavior, and recommendation systems. Section III defines the problem formally, while Section IV presents in detail the proposed *RsSafe* methodology including data collection, feature extraction, driver behavior model, and trip recommendation. Experiment evaluation is discussed in Section V. Finally, Section VI concludes the paper with directions of future research.

II. Related Work

This section briefly reviews the relevant literature in the context of this paper view on (1) Smartphone based driver behavior analysis, and (2) Recommendation systems for drivers. Detail of the systems under these categories given below:

A. Smartphone based Driver Behavior Analysis

A survey on various techniques for driver behavior analysis using machine learning methods on smartphone sensor data, is presented in [10]. Using support vector machine (SVM) and neural network (NN), a method is proposed in [11] for detecting with an accuracy of 95% fine-grained abnormal driver behavior, such as weaving, swerving, sideslipping, fast U-turn, turning with a wide radius, and sudden braking. In [4], a method is presented for driving maneuver prediction before they occurred to alert the drivers. Specifically, Long Short-Term Memory (LSTM) model was adapted because it can automatically capture temporal relations. The work in [12] uses LSTM with attention to detect driver behavior, and compares the model with simple LSTM and Multilayer Perceptron (MLP) to show the effect of attention layer. Inspired by the result obtained after adding the attention layer, we also incorporate the attention mechanism in our proposed approach *RsSafe*. However, all the above methods do not consider personalization factor, on the other hand we use Multi-task Learning (MTL) with attention to estimate the personalized driving behavior. (Note that existing systems have not yet used this technique for behavior prediction.) In [13] the driver behavior model is personalized, based on the past maneuvers execution and adaption method is applied for prediction whereas in our case we are using MTLA based technique to personalize recommendation.

B. Recommendation Systems for Drivers

Stress affects driving behavior based recommendation system, as studied in [3], where the authors used multi-task learning neural network (MTL-NN) for stress detection. They developed an android based mobile application that recommends to accept or reject the trip request based on the current stress level of the driver. On the other hand, we recommend solely from the driver behavior without considering stress. A multi-modal and adaptive to the situational context recommendation system is proposed in [14] which also optimizes the framework for real-time and large scale route queries. In [15], authors' presented a recommendation system that satisfies the demand for taxi users and driver expectations. The authors in [16] proposed an online model that provides the balance between system efficiency and driver equality in rideshare. Existing machine learning based methods do not capture the drivers' personality traits; the recommendation systems mostly address the issue of waiting time and trip time for both customers and drivers. On the other hand, in this paper our proposed recommendation system, *RsSafe*, captures the drivers' behavior at a personalized level using smartphone and trip-related data to provide a safe journey. *RsSafe* helps the drivers with options: accept or reject a trip.

III. Problem Statement

We propose to develop an adaptable and economical system, called *RsSafe* that provides safety to the drivers, passengers, and entities outside the taxis by predicting the driver's behavior and accordingly recommending the next trip. The problem is stated as follows.

Let D_i represent a driver where $i = 1, 2, \ldots, n$ denotes the driver's identity. Let B_i denote the current behavior of a driver D_i, in our case, behavior is aggressive (A_{gg}), drowsy (D_{rw}), and normal (N_{ml}). Let TC_j represent the completed trip where $j = 1, 2, \ldots, c$ is the number of trips taken by the driver. *RsSafe* suggests to accept or reject the trip TC_{j+1} for driver D_i by predicting the behavior B_{i+1}. Then, the recommendation is represented by following expressions:

$$R_{rec} : B_{i+1} = A_{gg} | D_{rw} \rightarrow R_{rej} \quad (1)$$

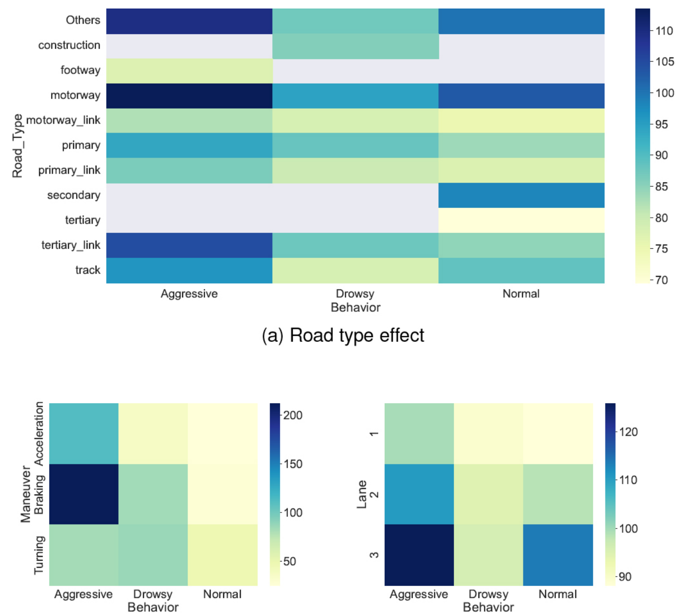
$$R_{rec} : B_{i+1} = N_{ml} \rightarrow R_{acc} \quad (2)$$

where the function R_{rec} recommends R_{rej} or R_{acc} on the basis of $B_{i+1} = A_{gg}|D_{rw}$ or $B_{i+1} = N_{ml}$ for the trip TC_{j+1} where A_{gg}, D_{rw}, and N_{ml} respectively represent the aggressive, drowsy and normal behavior of the driver for the next trip. The problem is further subdivided into four parts. The first part is the dataset collection that consists of trips, environment, dangerous maneuver, road data and sensor recorder data. The second part is to extract direct and indirect features that affect driver behavior. Since the role of driving behavior is significant throughout the trip to ensure safety, the third part of the problem is to predict the driver's behavior before the trip. The last part is to recommend the driver to accept or reject before the next trip according to the predicted behavior.

Authorized licensed use limited to: Missouri University of Science and Technology. Downloaded on February 17,2023 at 20:31:07 UTC from IEEE Xplore. Restrictions apply.

TABLE I
BASE FEATURES USED FOR MTLA

Feature Type	Features	Description
Sensor Data (S_d)	GPS data	Speed, latitude, longitude, altitude, speed (km/h)
	Accelerometer	Acceleration in X, Y, Z (Gs) axis and roll, pitch, yaw (degrees)
	Distance Traveled	First trip to current trip distance covered by the driver
Trip Data (T_d)	Time Traveled	Time for which the driver traveled from the first trip
	Completed Trips	Number of trips completed by the driver till now
	Rest Time	Rest taken by the driver after last trip
Road Data (R_d)	Allowed speed	Maximum allowed speed of current road (km/h)
	Road Type	Motorway/M_link/Primary/P_link
	Number of lanes	1/2/3 lane(s)
Environment Data (E_d)	Humidity	Humidity in percentage (%)
	Temperature	Temp in Celsius
	Pressure	Pressure in millibar
Dangerous Maneuver (D_m)	Type	Braking/ Turning/ Acceleration
	Level	Low/ Medium/ High



(a) Road type effect

(b) Maneuver execution

(c) Lane effect

Fig. 1. Relationship between features and behavior

IV. METHODOLOGY FOR RSSAFE

In our proposed approach, individual driving behavior characteristics are automatically learned by dividing the task of developing *RsSafe* into four stages as: (1) Data Collection, (2) Feature Extraction, (3) Modeling Driver Behavior, and (4) Trip Recommendation, as described below.

A. Data Collection

We are using publicly available UAH-DriveSet data [9]. This dataset is collected by a driving monitoring app, named DriveSafe[2] that has data of six drivers in the age group of 20-50 years, among which five are male and one is a female driver. Drivers cover two different routes in Madrid (Spain), the first one is 25 km round-trip in a motorway type of road with 120 km/h speed limit and has 3 lanes in each direction. The second one is 16 km round-trip in a secondary type of road with 90 km/h speed limit and has single lane in each direction. The driving data is for 500 minutes and labeled with driving behavior such as normal, drowsy, and aggressive.

B. Feature Extraction

According to the statistics[3], significant factors causing road accidents depend on the driver, vehicle, road conditions, or weather conditions which are captured by our proposed solution. We categorize the feature type as sensor data (S_d), trip data (T_d), road data (R_d), and environmental data (E_d), and danger maneuver (D_m). These features are called base features as described in Table I.

Sensor Data (S_d): The dataset contains GPS recorded data at 1Hz with tuple timestamp, speed, latitude, longitude, altitude, course, difcourse, etc. Also the calibrated inertial sensor data recorded with timestamp, acceleration X, Y and Z, roll, pitch, yaw information collected by fixing phone on the windshield.

Trip Data (T_d): The idea of taking trip data as features is that they reveal important properties of upcoming trips like distance traveled, time traveled, completed trips, and rest time, etc. which affect the driver behavior as surveyed in [3] for over 30 drivers.

Road Data (R_d): The road environment has a significant impact on the driving behavior; the Fig. 1a shows the relationship between driver behavior and road type. We divide the road environment category into the road type, maximum allowed speed, and number of lanes. The information is presented in the dataset and extracted using OpenStreetMap API[4].

Environmental Data (E_d) The weather condition affects the road surface condition and the driver's visibility, thereby increasing the chances of accidents. Adverse weather conditions such as heavy rain, thick fog and hail storms make driving riskier as the visibility reduces, and the road surface gets slippery. Weather details like temperature, wind, humidity, and pressure collected using VisualCrossing API[5] for every 6 minute of interval.

Danger Maneuver (D_m) The dataset contains the danger maneuver in processed data folder with tuple timestamp, type, level, and location information. The relationship between maneuver execution and driver behavior is shown in Fig. 1b. Since all the input features affecting the model performance are not in the same range, the normalization of the extracted features is necessary. We use the method of standardization for normalization where the feature vector is subtracted by its mean and then divides the result with standard deviation, represented as $\psi = \frac{X-\mu}{\sigma}$. Here, ψ denotes the normalized feature vector, X is the input vector, μ represents the mean, and σ denotes the standard deviation.

[2]http://www.robesafe.uah.es/personal/eduardo.romera/uah-driveset/
[3]http://jhtransport.gov.in/causes-of-road-accidents.html
[4]https://www.openstreetmap.org/
[5]https://www.visualcrossing.com/weather-data

Authorized licensed use limited to: Missouri University of Science and Technology. Downloaded on February 17,2023 at 20:31:07 UTC from IEEE Xplore. Restrictions apply.

C. Modeling Driver Behavior

After feature extraction, we model the driver behavior using *multi-task learning with attention* (MTLA). Modeling the driver behavior is challenging because of the change in driving conditions and sequential driving behavior. Modeling driving behavior in dynamic driving conditions is difficult because direct and indirect parameters affect driving over time. We use the LSTM architecture to identify the driving pattern in a dynamically changing environment to incorporate temporal dependencies, thus predicting future driver behavior more accurately. Another challenge is that every driver has a different behavior or style (e.g., aggressive vs. courteous driving) although the driving condition may be the same. In contrast, another driver is affected by changes in the driving environment that result in aggressive driving. With a goal to capture the difference, we use the MTLA based model to personalize the driving behavior because MTL is effective for modeling closely correlated tasks. It allows information sharing across tasks, especially with sparse data tackled alone (e.g. driving data missing due to sensing error). We formulate driving behavior prediction as:

1) MTLA Model: For predicting driver behavior, we need a quantitative measurement from the driving environment. As the dataset is labeled with the behavior, we use a supervised deep learning method on the extracted features and introduce MTLA architecture for predicting driver behavior during the trip. Multi-task learning can be incorporated into any neural network. In the following we demonstrate using a LSTM recurrent neural network.

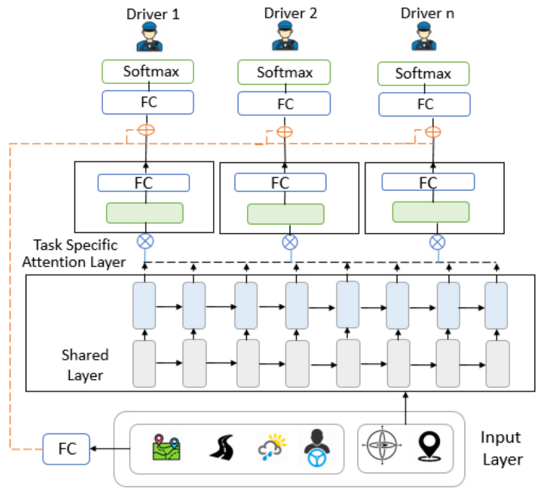
(a) Architecture Design

We use multi-task learning (MTL) model [17] that has been used in many recommendation systems to capture personalized information. MTL can learn different tasks in a single model instead of training individual neural network for each task that is more efficient not only in terms of memory and inference speed, but also in terms of data. In our case, a single driver behavior prediction is one task. i.e., D_i where $i = 1, 2, \ldots, n$ are number different of tasks. The input to the model are the features $(S_d, T_d, R_d, E_d, D_m)$, i.e., sensor data, trip data, road environment, environmental data, and danger maneuver. Formally, suppose there are n supervised learning tasks D_i where $i = 1, 2, \ldots, n$ and each task is associated with dataset $DataSet = (x_j^i, y_j^i)_{j=1}^{m_i}$ where x_j^i is the instance and y_j^i is the corresponding label. Here, m_i is pair of data instances and labels for each task D_i. We use a shared bottom approach of multi-task learning. The first layer of MTL is shared layer that captures common features for all the drivers affecting the driver behavior (i.e., taking turns along the roads, lane following, responding to traffic signals). The next layer is task-specific layer called tower that captures individual personality traits for each driver (i.e., Completed trips, rest data, distance traveled) as shown in 2a. Additionally in our case, we introduce the concept of 'attention' in the task-specific layer. The attention mechanism helps apply the mask on shared layers, to learn task-specific features. Thus, the attention mechanism acts as a feature selector, whereas the shared layer learns common features across all tasks. The proposed architecture is described in Figure 2a, where the task-specific attention layer selects features from the shared layer. Here, input to the shared layers is sensor data (S_d) whereas context data (i.e., trip data (T_d), road environment (R_d), environment data (E_d), and danger maneuver (D_m)) is passed to task specific layer through fully connected (FC) layer. Formally, the MTL shared bottom approach can be represented as:

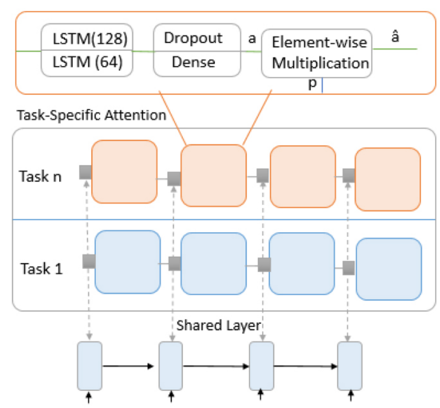
$$y_i = h^i(f(x)) \tag{3}$$

Here, y_i is the output parameter for each driver. Given i^{th} driver, the model comprised of shared bottom network as function f, and n tower networks h^i, where $i = 1, 2, \ldots, n$ for each driver respectively.

We use LSTM [18] blocks in shared layer, which allows effective representation learning from sequential input layer. In multi-tower network also, we have used LSTM blocks to capture individual personality traits. The use of LSTM as the underlying configuration in MTL because LSTM deals with vanishing gradient problem of traditional recurrent neural network (RNN) and works better with time-series data. At



(a) Driver behavior prediction using MTLA



(b) Task specific attention module

Fig. 2. Proposed architecture for driver behavior classification

Authorized licensed use limited to: Missouri University of Science and Technology. Downloaded on February 17,2023 at 20:31:07 UTC from IEEE Xplore. Restrictions apply.

each timestep t, we have input x^t (i.e., features), hidden state h^t and cell state c^t. The hidden state also acts as a memory block and operations of memory block are controlled by three gates, namely forgot, input and output gate. The operation of typical LSTM is defined using the following equations at each timestep:

$$\begin{aligned}
f^t &= \sigma(w_f h^{(t-1)} + u_f x^t + b_f) \\
i^t &= \sigma(w_i h^{(t-1)} + u_i x^t + b_i) \\
o^t &= \sigma(w_o h^{(t-1)} + u_o x^t + b_o) \\
\hat{c}^t &= tanh(w_c h^{t-1} + u_c x^t + b_c) \\
c^t &= f^t \cdot c^{t-1} + i^t \cdot \hat{c}^t \\
h^t &= o^t \cdot tanh(c^t)
\end{aligned} \tag{4}$$

where w_f, u_f, w_i, u_i, w_o, u_o, w_c, u_c are weights and b_f, b_i, b_o, b_c are biases. While σ and $tanh$ are activation functions, f^t is the forget gate, i^t is the input gate, and o^t is the output gate. Here, $\hat{c}^t$ represents new cell state and c^t represent the cell state; whereas dot $(\cdot)$ represents element-wise product operation.

(b) Task Specific Attention Module

The limitation of neural network based architectures is that they represent fixed length internal representation, which is not good for representing long dependencies. In our case, the driving behavior of a trip may exhibit complex dependencies from past events like turn (left, right, u-turn), breaking, reverse etc. In order to avoid the situation that only the last hidden vector is utilized to represent the driving behavior, the proposed *RsSafe* system utilizes attention mechanism that selects the most important signals to capture short and long distance dependencies. Moreover, the attention mechanism has shown better performance in machine translation and image analysis tasks to selectively focus on part of the important information. Motivated by this, we use attention mechanism at the task specific layer as presented in [19] for multivariate time-series classification, whereas the authors in [19] used task specific attention for image-to-image predictions and image classification. On the other hand, we have explored the use in sequential data for driver behavior prediction task. The intuition behind adding the attention layer in the task specific layer is to particularly select task related data from the shared layer. As such, the attention can be a feature selector from shared layer for task-specific layer, allows expressive representation to be learned for generalization across tasks, whilst allowing discriminative representation to be tailored for each individual task. The attention is added to learn task specific as shown in Fig. 2b features with one attention per task. Let p^j represent the shared features in the j^{th} hidden layer of i^{th} tower, and let a_i^j, t denote the learned attention mask of this layer for task i, at time $t-1$. Then task specific features $\hat{a}_i^j$ are:

$$\hat{a}_i^j = f_{att}(a_i^j, p^j) \tag{5}$$

where f_{att} represents element-wise multiplication as shown in Fig. 2b.

$$a_i^j = g_i^j([p^j; u^j(a_i^{j-1})]) \tag{6}$$

The tower network consists of a sequence of a stacked LSTM layers (u^j, g_i^j) dense layer and softmax (SM) layer. We also add the context information in task specific layer, described as follows:

$$\varrho = \varphi(T_d, R_d, E_d, D_m) \oplus \gamma(h^i(f(x))) \tag{7}$$

Here, $\oplus$ denotes the concatenation operation. The functions φ and γ denotes fully connected neural networks.

(c) The Model Objective

The model objective/loss function is the one that needs to be optimized. Compared to standard single-task learning, MTL training pose challenges to balance loss across different tasks. A Multi-task weighted loss function enable learning of all tasks, without allowing difficult task to suppress. In our case, all the tasks have objective to predict driver behavior, therefore there is no need to balance loss with weights. In multi-task learning with n tasks, the input features x and labels as y_i = 1,2,. . . ,m the loss function is defined as,

$$L_{total}\Big((x_j^i, y_j^i)_{j=1}^{m_i}\Big) = \sum_{i=1}^{n} L^i\big(x_j^i, y_j^i\big) + \Omega \tag{8}$$

Here, L^i denotes task-specific losses and Ω is the $L_2 = ||w||_2^2$ regularization term that helps to remove high bias and variance, w denotes learnable weights. Total loss $\big(L_{total}((x_j^i, y_j^i)_{j=1}^{m_i})\big)$ is the sum of losses over all the n tasks (i.e., number of drivers). In our case $L^i\big(x_j^i, y_j^i\big)$ is cross entropy loss because we are dealing with classification problem.

$$L^i\big(x_j^i, y_j^i\big) = -\sum_{k=1}^{3} y_{j,k}^i log(P_{j,k}^i) \tag{9}$$

where $y_{j,k}^i$ and $P_{j,k}^i$ are the ground truth and predicted score for each class k i.e., (A_{gg}, D_{rw}, N_{ml}).

D. Trip Recommendation

This section explains the trip recommendation use case of driver behavior prediction with the help of following scenario. Suppose the driver receive a new trip request, *RsSafe* suggest recommendation options to driver such as accept or reject based on past journeys' behavior using weighted moving average (WMA) [20] which puts more weight on recent driving behavior and less on past driving behavior.

$$WMA = \frac{(B_i * i) + (B_{i-1} * i-1) + \ldots + (B_1 * 1)}{(i + i-1 + \ldots + 1)} \tag{10}$$

The past behavior prediction model is used to predict the driving behavior for next journey. In the recommendation, two scenarios arise. The first one is if the driver's behavior is drowsy or aggressive, the system recommends to reject the trip. Second, if the driver behavior is normal, then the system recommends to accept the trip. We can write the driver behavior and recommendation relation as: $B_{i+1} \bowtie R_{rec}$

Authorized licensed use limited to: Missouri University of Science and Technology. Downloaded on February 17,2023 at 20:31:07 UTC from IEEE Xplore. Restrictions apply.

where B_{i+1} can be $A_{gg}|D_{rw}|N_{ml}$ i.e., aggressive, drowsy and normal. As in equations (1) and (2), $A_{gg}|D_{rw} \rightarrow R_{rej}$ and $N_{ml} \rightarrow R_{acc}$ where R_{rej} denotes reject recommendation while R_{acc} represents accept recommendation.

V. Experimental Study

First, we introduce experimental setup, baselines and metrics for evaluation. Then the following research questions are used to guide our experiments:

RQ1: How does the MTLA perform compared to existing baseline driver behavior classification methods?

RQ2: How does the MTLA perform compared to existing state-of-the-art driver behavior classification methods?

RQ3: How does each feature contribute to the performance?

RQ4: How effective is each design choice in MTLA?

RQ5: How does the hyper-parameter setting effect the performance of MTLA?

A. *Experimental Setup*

The experiments are performed on Google Colab with CPU Intel Xeon 2.20 GHz and GPU specification Tesla P100-PCIE-16GB. Memory space allotted by Colab environment was 12 GB RAM and Hard disk space 34 GB. We implemented the model in Python using PyTorch toolbox. In MTLA each task specific network learns from shared bottom layers; in our implementation we used LSTM with 4 hidden layers, each of size 128 units. As six drivers are involved in the UAH-DriveSet dataset, each driver has different task specific layers (i.e., 6 tower networks are used), which comprised LSTM with 2 layers, each of 128, 64 hidden units. Fully connected layer (FC), each of size 100 is used in task specific layer. We use dropout layer with 0.2 dropout ratio to avoid overfitting. Adam optimizer is used for training with learning rate of 0.0001 and run for 200 epochs.

B. *Evaluation Metrics*

Performance measures are used to evaluate the model quantitatively. We use performance measures, namely, accuracy and F-measure, as metrics for evaluation as they are popular metrics for classification. Accuracy is defined as ratio of number of instances correctly predicted as normal, drowsy and aggressive upon total instances of driving behavior given as $A_{acc} = \frac{P_r}{T}$. Here, A_{acc} represent the accuracy, P_r is the correctly predicted instance and T denotes total instances. F measure is calculated using the weighted harmonic mean of precision (P) and recall (R) given as $F = 2\frac{P*R}{P+R}$. Here, P is the ratio of correctly classified relevant driver behavior and actual behavior. R is the ratio of correctly classified relevant driver behavior and predicted behavior. We also evalutate our model using ROC curve which plots true positive rate (TPR) vs. false positive rate (FPR) on different thresholds.

C. *Evaluation*

In the proposed methodology, the dataset is divided into 80:20 training and test ratio, with 20% of training data used for the validation set. We obtain the F-measure score as 0.96, micro F-measure as 0.94, and macro F-measure as 0.92 for predicting driving behavior in the UAH-DriveSet dataset. We also tried to compute the ROC curve as shown in Fig. 3 with 0.97 macro and micro average ROC curve. The normalized confusion matrix is given in Fig. 4, where sum of each row is 1.00 represents 100% of instances of particular category. Out of total, 0.97% normal behavior instances are classified accurately, 0.94% aggressive behavior instances are classified accurately, and 0.95% drowsy behavior instances are classified accurately.

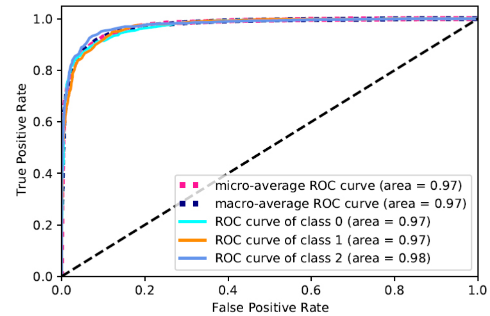


Fig. 3. ROC curve of MTLA

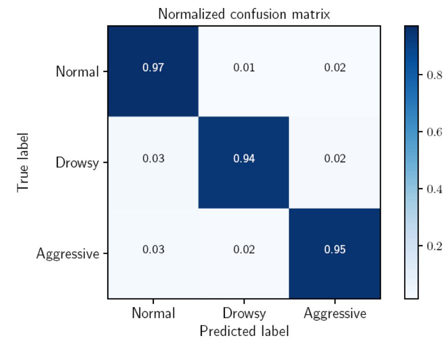
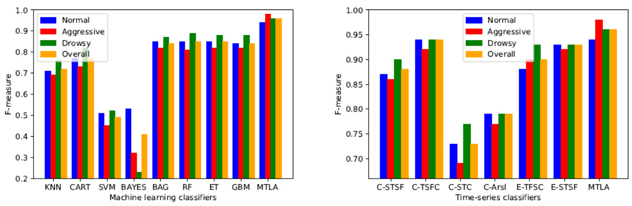


Fig. 4. Confusion matrix of driver behavior classification

*1) **RQ1: Comparison with existing baseline models**:* To show the effectiveness, we compare with machine learning, time-series, and multi-task baseline models.

- **Comparison with machine learning baseline models** First we compare the MTLA performance with existing machine learning models as shown in Fig. 5a. We include different classifiers for comparisons like K-nearest neighbors (kNN), decision tree (CART), support vector machine (SVM), naive bayes (BAYES), bagging classifier (BAG), random forest (RF), extra tree classifier (Extra-Trees), and gradient boosting classifier (GBM). From Fig. 5a, we observe that the ensemble learning based models are showing better results because they aggregate results of individual weak classifiers based on different strategies.
- **Comparison with time-series baseline models** To show the effectiveness of MTLA, we compare the model with several baselines that work with multivariate time-series

Authorized licensed use limited to: Missouri University of Science and Technology. Downloaded on February 17,2023 at 20:31:07 UTC from IEEE Xplore. Restrictions apply.



(a) Machine learning models (b) Time-series models

Fig. 5. Comparison with existing baseline models

TABLE II
COMPARISON WITH STATE-OF-THE-ART MODELS

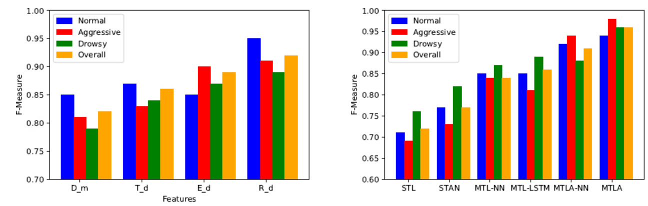
Model	Year	Technique	F Measure
Saleh et al. [21]	2017	LSTM	0.91
Xie et al. [22]	2019	Random Forest	0.70
Moukafih et al. [23]	2019	LTSM-FCN	0.95
Pjetri et al. [24]	2019	Random Forest	0.87
Yi et al. [25]	2019	Random forest	0.91
Schlegel et al. [26]	2021	HDC	0.94
RsSafe	2022	MTLA	0.96

classification using *sktime*[6] library as shown in Fig. 5b. We train the time-series column concatenate based supervised timeSeries forest classifier (C-STSF), timeSeries forest classifier (C-TSFC), shapelet transform classifier (C-STC), arsenal (C-Arsl), and column ensemble based TimeSeries Forest Classifier (E-TFSC), Supervised TimeSeries Forest classifer (E-STSF) on the same feature set as MTLA. The comparison in Fig. 5b shows that the C-STC provides worst performance, and MTLA provides best performance.

*2) **RQ2: Comparison with existing state-of-the-art models**:* We have compared our proposed model with existing driver behavior prediction in terms of F-measure and comparison details given in Table II, where our proposed model outperforms existing models. In [21] raw smartphone sensor data are used as input to LSTM that classify driving behavior with F1-Score of 0.91; and [27] achieved F1-score of 0.87 that used hand-crafted features with random forest as the classification technique. In [23] is proposed a novel LSTM Fully Convolutional Network (LSTM-FCN) to classify aggressive behavior with F1-score of 0.95; and [24] used random forest on the subset of features to classify behavior with F1-score of 0.87. In [25] random forest is applied to personalize the driver state recognition system. Deep learning based approaches work better than classical machine learning approaches as observed in Table II.

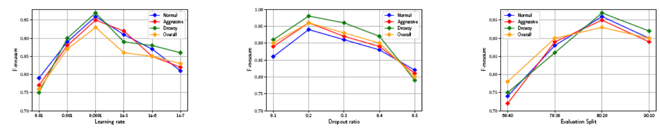
*3) **RQ3: Evaluating feature contribution** :* The ablation study is performed to understand the importance of features. In this study, features are removed, and then their effect is observed on the accuracy. The Fig. 6a shows the impact of the dangerous maneuvers type feature is most significant among the considered feature set, and road environment data have the

[6]https://sktime.org/



(a) Ablation study (b) MTL techniques

Fig. 6. Result of ablation study (left) and comparison with different MTL variants (right)



(a) Learning rate (b) Dropout ratio (c) Evaluation split

Fig. 7. Effect of hyperparameter sensitivity

least effect. The dangerous maneuvers comprised other basic features such as turning, braking acceleration that relate to driver behavior directly, whereas the road type feature consists road type, maximum speed, and number of lanes.

*4) **RQ4: Evaluating design choice** :* Another kind of comparison shows the design effectiveness of multi-task learning based approach. We compare with Single task learning (STL), Single task with attention (STAN), Multi-task with neural network (MTL-NN), Multi-task with LSTM (MTL-LSTM), Multi-task attention with neural network (MTLA-NN), and Multi-task attention with LSTM (MTLA) as shown in Fig. 6b. We adjust the layers and hidden units in each model for the best performance. The softmax (SM) function is used in the output to assign probabilities to different behavior classes. There is a difference of 24% in the STL model and MTLA, inferring that the personalized multi-task approach works better than the generic model for all drivers.

*5) **RQ5: Evaluating hyper-parameter sensitivity** :* Different hyper-parameters are used in RsSafe, such as learning rate, dropouts, and evaluation splits. The effect of these hyper-parameters are described below:

- **Effect of learning rate** The learning rate is one of the important hyperparameter to find minima of loss function, different setting and corresponding F-measure is shown in Fig. 7a. We start with the learning rate of 0.01 and decrease with $1/10^e$, where $e \in [1, 2, \ldots, n]$ denotes the power. For our model, we choose 0.0001, as the Fig. 7a shows that this learning rate is providing highest F-measure.
- **Effect of dropout ratio** Dropout is one of the important regularization techniques that prevents the model from overfitting. The effect of dropout ratio selection is shown in Fig. 7b, where dropout rate of 0.2 provides highest F-measure and as we increase the dropout rate model performance degrades abruptly.
- **Effect of evaluation split ratio** For training the machine learning models, we split the dataset into training, testing,

Authorized licensed use limited to: Missouri University of Science and Technology. Downloaded on February 17,2023 at 20:31:07 UTC from IEEE Xplore. Restrictions apply.

and validation part, therefore we select the split as 80:20 on the basis of results shown in Fig. 7c, where 80% is used for training and 20% is used for testing.

VI. Conclusion

With an increasing demand for taxi services, the services provider form policies to attract more taxi drivers. Now, to provide safe journeys, a system is required to predict drivers' behavior before the trip. In this paper, we presented an adaptable recommendation system, called *RsSafe* that provides suggestions on the trip (e.g., to accept or reject) based on the driver's behavior. We proposed the multi-task learning with attention (MTLA) based driver behavior prediction technique to make *RsSafe* more adaptable by personalizing the trip recommendations. MTLA predicts driver behavior output as aggressive, drowsy, and normal; and achieves an accuracy of 96%, 97% for UAH-DriveSet in terms of F-measure and ROC respectively. We compare our proposed method with baseline models, state-of-the-art approaches, and different MTL based techniques; the proposed model outperforms. We also perform the ablation study to show the need and importance of features. *RsSafe* has a vast area of applications in fleet management, insurance services, taxi services, and advanced driver-assistance systems.

This work is limited to taxi drivers only. In future, we plan to extend our work for privately owned vehicles that recommend drivers with alerts to reduce the risk of accidents. Although the model predicts behavior with high accuracy, indirect variables are challenging to capture, like personal or family issues of drivers and other hidden factors that also affect the driving behavior. To enhance the scalability of the system i.e., not linearly dependent on the data size (i.e. number of drivers), we can cluster the drivers according to personalities and then predict the behavior.

References

[1] C.-E. Havârneanu, C. Măirean, and S.-A. Popuşoi, "Workplace stress as predictor of risky driving behavior among taxi drivers. the role of job-related affective state and taxi driving experience," *Safety science*, vol. 111, pp. 264–270, 2019.

[2] P. I. Wouters and J. M. Bos, "Traffic accident reduction by monitoring driver behaviour with in-car data recorders," *Accident Analysis & Prevention*, vol. 32, no. 5, pp. 643–650, 2000.

[3] R. Verma, B. Mitra, and S. Chakraborty, "Avoiding stress driving: Online trip recommendation from driving behavior prediction," in *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pp. 1–10, 2019.

[4] A. Jain, A. Singh, H. S. Koppula, S. Soh, and A. Saxena, "Recurrent neural networks for driver activity anticipation via sensory-fusion architecture," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3118–3125, 2016.

[5] N. Arbabzadeh and M. Jafari, "A data-driven approach for driving safety risk prediction using driver behavior and roadway information data," *IEEE transactions on intelligent transportation systems*, vol. 19, no. 2, pp. 446–460, 2017.

[6] A. Kashevnik, I. Lashkov, and A. Gurtov, "Methodology and mobile application for driver behavior analysis and accident prevention," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 6, pp. 2427–2436, 2019.

[7] V. Petraki, A. Ziakopoulos, and G. Yannis, "Combined impact of road and traffic characteristic on driver behavior using smartphone sensor data," *Accident Analysis & Prevention*, vol. 144, p. 105657, 2020.

[8] J. Guo, U. Kurup, and M. Shah, "Is it safe to drive? an overview of factors, metrics, and datasets for driveability assessment in autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[9] E. Romera, L. M. Bergasa, and R. Arroyo, "Need data for driver behaviour analysis? presenting the public uah-driveset," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 387–392, IEEE, 2016.

[10] F. Lindow and A. Kashevnik, "Driver behavior monitoring based on smartphone sensor data and machine learning methods," in *2019 25th Conference of Open Innovations Association (FRUCT)*, pp. 196–203, IEEE, 2019.

[11] J. Yu, Z. Chen, Y. Zhu, Y. Chen, L. Kong, and M. Li, "Fine-grained abnormal driving behaviors detection and identification with smartphones," *IEEE transactions on mobile computing*, vol. 16, no. 8, pp. 2198–2212, 2016.

[12] S. M. Kouchak and A. Gaffar, "Detecting driver behavior using stacked long short term memory network with attention layer," *IEEE Transactions on Intelligent Transportation Systems*, 2020.

[13] H. Dang and J. Fürnkranz, "Using past maneuver executions for personalization of a driver model," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 742–748, IEEE, 2018.

[14] H. Liu, Y. Tong, P. Zhang, X. Lu, J. Duan, and H. Xiong, "Hydra: A personalized and context-aware multi-modal transportation recommendation system," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2314–2324, 2019.

[15] X. Wan, H. Ghazzai, and Y. Massoud, "Incremental recommendation system for large-scale taxi fleet in smart cities," in *2019 IEEE International Conference of Vehicular Electronics and Safety (ICVES)*, pp. 1–6, IEEE, 2019.

[16] Y. Xu and P. Xu, "Trade the system efficiency for the income equality of drivers in rideshare,"

[17] R. Caruana, "Multitask learning," *Machine learning*, vol. 28, no. 1, pp. 41–75, 1997.

[18] S. Hochreiter and J. Schmidhuber, "Lstm can solve hard long time lag problems," in *Advances in neural information processing systems*, pp. 473–479, 1997.

[19] S. Liu, E. Johns, and A. J. Davison, "End-to-end multi-task learning with attention," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1871–1880, 2019.

[20] J. S. Hunter, "The exponentially weighted moving average," *Journal of quality technology*, vol. 18, no. 4, pp. 203–210, 1986.

[21] K. Saleh, M. Hossny, and S. Nahavandi, "Driving behavior classification based on sensor data fusion using lstm recurrent neural networks," in *IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–6, 2017.

[22] J. Xie and M. Zhu, "Maneuver-based driving behavior classification based on random forest," *IEEE Sensors Letters*, vol. 3, no. 11, pp. 1–4, 2019.

[23] Y. Moukafih, H. Hafidi, and M. Ghogho, "Aggressive driving detection using deep learning-based time series classification," in *IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)*, pp. 1–5, 2019.

[24] A. Pjetri, M. Simoncini, F. Sambo, A. Lori, and F. Schoen, "Light footprint driving behaviour classification," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 1186–1191, IEEE, 2019.

[25] D. Yi, J. Su, C. Liu, M. Quddus, and W.-H. Chen, "A machine learning based personalized system for driving state recognition," *Transportation Research Part C: Emerging Technologies*, vol. 105, pp. 241–261, 2019.

[26] K. Schlegel, F. Mirus, P. Neubert, and P. Protzel, "Multivariate time series analysis for driving style classification using neural networks and hyperdimensional computing," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, pp. 602–609, IEEE, 2021.

[27] J. Xie, A. R. Hilal, and D. Kulić, "Driving maneuver classification: A comparison of feature extraction methods," *IEEE Sensors Journal*, vol. 18, no. 12, pp. 4777–4784, 2017.

Authorized licensed use limited to: Missouri University of Science and Technology. Downloaded on February 17,2023 at 20:31:07 UTC from IEEE Xplore. Restrictions apply.