Developing an Interpretable Driver Risk Assessment Model to Increase Driver Awareness Using In-Vehicle Records

Min Zheng[1,2] and Jingyu Zhang[1,2](✉)

[1] Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China
zhangjingyu@psych.ac.cn

[2] Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract. Providing feedback to drivers on their risky driving behaviors is an important method to improve drivers' awareness in reducing future accidents. However, it is hard to identify risk-prone behaviors and explain them to drivers. In the present study, we used driving log from 103370 electric vehicles equipped with L2-assisted driving functions. We used 28 explainable features to establish a binary classification model of accidents and eight features can be used to establish an acceptable model. Further, we developed an easy-to-understand safety score formula using these eight features. Through this accurate and transparent feedback, we may improve drivers' safety awareness without undermining their trust in the L2 and higher level automated vehicles. This will not only reduce accidents but enable them to adapt to the development of automated driving technology in a smoother manner.

Keywords: High-risk Driving Behavior Features · Automated Driving Technology · Logic Regression Model · Big Data Analysis

1 Introduction

Driving safety in the stage of man-machine co-driving is a difficult problem that needs to be solved in the development of automatic driving technology [1]. The purpose of a large number of previous studies is to design the machine to adapt to people on the basis of understanding the features of people [1], and let people match the machine to some extent to fill up the immaturity of technological development.

In-car warning functions such as DMS (Driver Monitoring System), HOD (Hands Off Detection), FCW (Forward Collision Warning), and AEB (Autonomous Emergency Braking) are sent out to remind the driver to adjust their behavior. They are important ways to provide drivers with timely feedback on their just-performed risky driving behaviors to improve their safety. However, an long-term evaluation of drivers' behaviors is also needed to increase drivers' overall awareness of their style and skills.

In order to develop an effective tool for that purpose, several principles should be met. First, such an evaluation must be based on continuous and unintrusive measures, so

C. Stephanidis et al. (Eds.): HCII 2023, CCIS 1836, pp. 131–136, 2023.
https://doi.org/10.1007/978-3-031-36004-6_18

it can be updated quickly as the drivers' behavior changes. Second, such an evaluation must reflect the actual risk of accidents, so the drivers can really trust the evaluation. Third, such an evaluation must be based on an easy-to-understand criterion, so the drivers can grasp their meaning intuitively and are glad to use it.

This study sought to make an initial step in developing such a tool. We used real vehicle records of 103370 electric vehicles equipped with L2-assisted driving functions. We first filtered the features by experts to keep the most explainable ones. Then we used these selected features to establish a model to predict accidents. Finally, we developed a safety evaluation score based on the established prediction model and developed a user-friendly interface for drivers to understand the meaning of that score.

2 Method

2.1 Data Collecting

The data used the real vehicle driving and accident data of a certain automobile enterprise within two months. More than 400 features collected from the in-car detector were used as the source of data, and all accidents were restored as far as possible by collecting data from the following three sources: (1) airbag ejection record; (2) insurance report record; (3) repair work order records.

The number of subjects in a month was 103370, and the number of accidents in the previous month was 1598. The number of accidents in the next month was 1286.

2.2 Extracting Explainable Features Through Expert Evaluation

We asked 10 experts to first choose the dimensions that were important in resulting in an accident and are explainable to ordinary drivers. They were asked to make their own judgment separately and then discuss to reach an agreement. They agreed there were five dimensions including active safety warning, attention state, acceleration and deceleration behavior, sharp turning behavior and car-following behavior. Then 28 features were selected from the five dimensions.

2.3 Using the Explainable Features to Establish an Accident Prediction Model

In order to achieve better robustness, the number of features must be reduced further. We calculated the IV (Importance Value) to choose the more relevant feature [3]. We selected a total of eight features based on the criterion of selecting the feature with the highest IV value in each category. The value of eight features were transformed into WOE (Weight of Evidence) value which were used to build the binary classification model with the same standard of measurement. From the result we got the coefficient of each feature.

The coefficient were important because they were needed for the calculation of safety score. We then validated the model through AUC value. The AUC values for the previous month and the following month were 0.75 and 0.72 respectively which were considered acceptable [4].

2.4 Establishing a User-Friendly Safety Score

"y(x)" was the probability that the sample was predicted to be 1,"1 − y(x)" was the probability that the sample was predicted to be 0. The "odds" or "y(x)/1 − y(x)" was the ratio of accident rate to non-accident rate.

$$\begin{aligned} odds &= \frac{y(x)}{1-y(x)} \\ &= \left(e^{a_1}\right)^{x_1} \times \left(e^{a_2}\right)^{x_2} \times \ldots \left(e^{a_n}\right)^{x_n} \times e^{b} \end{aligned} \tag{1}$$

x_1, x_2, …, x_n was the value of WOE which was the feature after discretization transformation and a_1~a_n is coefficients. The calculation formula of odds was obtained by substituting the coefficient:

$$\begin{aligned} odds = 1.0089 \times 1.4589^{\mathrm{dec3-2-woe}} \times 1.9338^{\mathrm{acc3-1-woe}} \times \\ 1.4412^{\mathrm{st3-2-woe}} \times 2.2853^{\mathrm{zgc3-woe}} \times 2.7437^{\mathrm{AEB-woe}} \times \\ 2.2526^{\mathrm{HOD-woe}} \times 1.8626^{\mathrm{DMS-woe}} \times 1.5123^{\mathrm{FCW-woe}} \end{aligned} \tag{2}$$

We used a linear regression formula to convert odds into easy-to-understand score:

$$\mathrm{Score} = \mathrm{A} - \mathrm{B} \times \mathrm{odds} \tag{3}$$

"A" and "B" were constants. We got the A and B values according to the needs of the score distribution. Finally, each driver could get two sets of scores, that were total score and different driving behavior scores.

From the eight features we could realize short distance from the vehicle in front, urgent acceleration and deceleration, and sharp turning had strong correlation with accidents, as well as attention status and hands-off the steering wheel. These six behaviors were the key behaviors leading to accidents. The driver's six driving behavior scores were calculated from formula (3) except car-following score. The car-following score was calculated from scores of zgc3, FCW and AEB.

We used a user-centered design approach to develop such a tool. The results were conveyed to the drivers in a way that was easy to understand.

3 Results

3.1 Explainable Features Extracted

Among the on-board features, AEB has the strongest correlation, followed by urgent acceleration, sharp turning, urgent deceleration, FCW, HOD, DMS, and the longitudinal car-following (Table 1).

Table 1. Features sorted based on IV values

Dimension	Category	Feature	IV
Active safety warning	AEB	**AEB**	**0.1904**
Acceleration and deceleration behavior	Acceleration	**acc3-1**	**0.1415**
		acc2-1	0.1361
		acc3-2	0.1353
		acc2	0.1137
		acc3	0.1089
		acc1	0.089
Sharp turning	Sharp turning	**st3-2**	**0.123**
		st2-1	0.1008
		st2	0.0752
		st3	0.0562
		st3-1	0.056
		st3	0.0267
Acceleration and deceleration behavior	Deceleration	**dec3-2**	**0.1198**
		2-Dec	0.1122
		dec2-1	0.0801
		3-Dec	0.0748
		dec3-1	0.0708
		1-Dec	0.0701
Active safety warning	FCW	**FCW**	**0.1182**
Attention status warning	HOD	**HOD**	**0.0985**
	DMS	**DMS**	**0.0696**
Car-following behavior	Car-following	**zgc3**	**0.0097**
		zgc2	0.0057
		zgc1	0.0026
		cgc3	0.0008
		cgc2	0.0007
		cgc1	0.0001

3.2 Accident Prediction Model

A binary classification model was established to quantitatively evaluate the accident risk of each driver and passed the validation (Table 2).

Table 2. Logit Regression Results

Dep. Variable:		label		No. Observations:		142480
Model:		Logit		Df Residuals:		142471
Method:		MLE		Df Model:		8
Date:		Wed, 15 Feb 2023		Pseudo R-squ.:		0.1546
Time:		18:25:29		Log-Likelihood:		−83496
converged:		True		LL-Null:		−98759
Covariance Type:		nonrobust		LLR p-value:		0.000
	coef	std err	z	P > |z|	[0.025	0.975]
const	0.0089	0.006	1.480	0.000	−0.003	0.021
dec3-2	0.3777	0.016	22.924	0.000	0.345	0.410
acc3-1	0.6595	0.013	52.338	0.000	0.635	0.684
st3–2	0.3655	0.016	22.914	0.000	0.334	0.397
zgc3	0.8265	0.042	19.640	0.000	0.744	0.909
AEB	1.0093	0.011	87.814	0.000	0.987	1.032
HOD	0.8121	0.018	44.207	0.000	0.776	0.848
DMS	0.622	0.021	29.138	0.000	0.580	0.664
FCW	0.4136	0.017	24.261	0.000	0.380	0.447

3.3 The Calculation and the Interface of the Safety Score

Scoring rules were obtained by calculating the accident rate odds. Two sets of scores were got. The car could convey it in a way that was easy for drivers to understand through the interface (Table 3).

Table 3. The results conveyed to drivers(example)

Score basis	Shown to driver	
	total score: 70	better than 50% drivers
zgc3, FCW, AEB	car-following score: 65	better than 45% drivers
acc3-1	urgent acceleration score: 77	better than 54% drivers
dec3-2	urgent deceleration score: 80	better than 60% drivers
st3-2	sharp turning score: 60	better than 47% drivers
DMS	attention status score: 68	better than 50% drivers
HOD	hands-off score: 70	better than 50% drivers

4 Discussion

Because the cause of the accident was not only human factors, but also the factors of the vehicle itself, as well as the factors of scene and road conditions, other traffic participants, traffic management, etc. And although there were literatures on human factors showed that gender, age, driving age, cognitive ability, attitude, personality traits, etc. can significantly affect driving behavior [5], it was difficult to collect data on these factors because of the principle of privacy protection and minimizing the disturbance to users. If only human behaviors were taken into account without other factors, the final effect would be limited. Therefore, the next task will be establishing a complete set of assessment system consisting of vehicle safety assessment, road safety assessment, scene safety assessment, etc., to comprehensively consider the safety score. Only after trying to break through the bottleneck in other aspects can the effect of safety score be further improved.

5 Conclusion

In this study, we made an initial step by using real vehicle recording data to establish a safety score that is continuous, explainable and predictive of accidents. We hope the drivers can use this can of feedback to improve their driving awareness.

Acknowledgment. This study was supported by the National Natural Science Foundation of China (Grant No. T2192932).

References

1. Third Global Ministerial Conference on Road Safety. Stockholm Declaration. Stockholm, Sweden (2020)
2. Huang, C., Hang, P., Hu, Z.X., Lv, C.: Collision-probability-aware human-machine cooperative planning for safe automated driving. IEEE Trans. Veh. Technol. **70**, 9752–9763 (2021)
3. Mays, E., Lynas, N.: Credit Scoring for Risk Managers: The Handbook for Lenders. CreateSpace Independent Publishing Platform (2011)
4. Fawcett, T.: An introduction to ROC analysis. Pattern Recogn. Lett. **27**, 861–874 (2006)
5. Wang, X.: Traffic Safety Analysis Shanghai. Tongji University Press (2022)