REPORT OF RESEARCH RESULTS

A. Title： Multi-dimensional Driving Risk Assessment with Limited Dynamic Vehicle Information

B. Primary Researcher： Tsai Chen Wu, Affiliation: wuwwu0814@gmail.com
Co-researchers: Wei-Hsun Lee, Affiliation: leews@mail.ncku.edu.tw

C. Summary：

Traffic accidents are predominantly caused by driver-related factors, the early detection of aggressive driving behaviors is crucial for accident prevention and enhancing road safety. This study aims to assess driving risk using GPS data, offering a practical and scalable solution compared to more expensive sensor based systems. However, using only GPS data for driving behavior analysis presents several challenges. Addressing to the problem of the limitations of GPS data, and the lack of risk groundtruth, this research proposes a comprehensive driving risk assessment framework that assesses driving risk using GPS data. The framework includes data preprocessing, risk labeling using the Naturalistic Driving Study dataset, and risk assessment utilizing only GPS data. A hierarchical risk-aware feature extraction method is developed to bridge the gap between raw GPS data and driving risk, processing GPS data through multiple layers. A multi dimensional risk assessment model, incorporating five axes of risk indicators such as speed, acceleration, deceleration, lateral movement, and safe distance, is used for labeling driving risks. The study further employs both rule-based and supervised learning models for risk assessment using only GPS data, with synthetic minority over-sampling to address data imbalance.

Experimental results show that the hierarchical feature extraction method significantly improves predictive performance. The framework achieves high accuracy in identifying aggressive driving behaviors, with F1-scores around 70 80% for various risk dimensions. Despite the limitations of GPS data, the findings demonstrate that GPS technology is a valuable tool for driving risk assessment, offering high accessibility and ease of implementation. The practical implications of this work suggest potential for widespread use in fleet management and real time driver monitoring systems to improve road safety.

D. Aim of Resaerch：

1. Proposing a driving risk assessment framework using only GPS data

Understanding the importance of safe driving and the advantages of GPS sensor, driving risk assessment framework for coaches that only requires GPS data is proposed. The vehicle dynamic information collected by sensor is incorporated into trained classification model to predict the corresponding risk level of the coach drivers.

2. Developing a multi-dimensional driving risk assessment model

This study aims to create a model that can evaluate the risks associated with driving using multiple indicators. The model is advantageous because it is easy to understand, identifies the specific type of risky driving based on each indicator's score, and can be used with various types of data from different sensors.

3. Overcoming possible problems when using GPS data to assess driving risk

The defects of GPS data include less signals, low data collection frequency, and low correlation between raw data and driving behavior; This study hopes to overcome the above obstacles by extracting more behavior-related features from GPS data to explore potential driving risk.

E. Method of Research & Progression：

The research framework consists of four primary phases: data preprocessing, feature engineering, driving risk labeling, and driving risk assessment using GPS. During the data preprocessing stage, the collected data is subjected to cleansing techniques and window sliding techniques. Feature engineering involves extracting and calculating raw features across multiple layers to identify driving risk. In driving risk labeling, multi-dimensional risk assessment model is proposed to assess driving risk with multiple sensors. Lastly, two method are proposed to assess driving risk with only GPS data.

F. Results of Research：

During the risk assessment using GPS data, the following four key findings were observed：

1. Both the rule-based and learning-based GPS risk models are effective to some extent in identifying risks. However, with the cross validation with risk label, the learning-based model outperforms the rule-based model, achieving AUC of 73.4% and accuracy of 76.4%. In comparison, the rule-based model achieves AUC of 66% and accuracy of 72%.
2. The process of feature extraction significantly enhances the predictive performance of the models.
3. For the detection of aggressive acceleration and deceleration, a window size of 3 seconds is found to be optimal. On the other hand, a window size of 5 seconds is optimal for identifying lateral movement and situations where the driver fails to maintain a safe distance. For identifying overall risk level, 3 seconds are found to be the most suitable.
4. The algorithms XGBoost, Random Forest, and voting deliver the best performance in predicting driving risk in this research.

The performance of risk assessment varies across different risk dimensions. For rapid acceleration, deceleration, lateral movement, and not keeping a safe distance,

the GPS model achieves accuracies of 86.8%, 85.7%, 78.7%, and 79.4%, respectively. For overall risk, an accuracy of 76.4% is achieved.

G. Future Areas to Take Note of, and Going Forward：

This method can be applied in many practical applications, especially those that require evaluating and monitoring driving behavior. For example, in the insurance industry, assessing the risk level of drivers and developing personalized insurance policy plans is crucial. Using GPS functionality and this method to collect driver behavior data can help insurance companies better understand the behavior and risk level of their drivers and provide more personalized insurance solutions.

Furthermore, this method can also be applied to driver training and safety management. By analyzing GPS data and using this method, drivers can receive feedback on their driving behavior and personalized training and recommendations can be provided. At the same time, this method can also be used for fleet management to monitor driver behavior, improve driving safety, and reduce accident risk.

H. Means of Official Announcement of Research Results：

Proposed as a dissertation of master's degree of Graduate program of Department of Transportation and Communication Management Science, National Cheng Kung University.