



## ABSTRACT

Intelligent Transport System is the system of detecting, analyzing, monitoring, and communications technologies to improve safety, mobility, and efficiency. At present, vehicles are increasing exponentially on the road, which leads to the accident on the roads. In this work, we focused on human behaviors of driving and death rates due to an accident on the road. For the driver, the problems caused by fatigue, sleepiness, and inattention. The Identify Driver Behavior system makes drivers or road users improve their behavior of driving are efficiently and safely, costs less in fuel, and being good with environmental. In recent decades, datasets are critical and valuable in several fields, for instance, economics, marketing or trading, and others. Researchers are interested in these exiting the knowledge from these data, especially driver behavior analysis. The objective of this work is implementing the classifiers using data mining techniques to build the classifier model. The proposed system is using two resampling techniques such as up sampling and down sampling with an imbalanced data and combine with SMOTE and ADASYN and we will compare the performance of each classifier. The classifier includes five algorithms: random forest, decision tree, linear discriminant analysis, support vector machine, and logistic regression. From evaluation, the experiment result shows that it can solve the problem of imbalanced data. The resampling dataset from SMOTE combined with the Majority class shows the highest score from the Random forest classifier algorithm with the accurate 98.3% accuracy and Decision Tree classifier algorithm 97.7% accuracy. The resampling dataset from ADASYN combined with the Majority is also a high score from RFC is 93.1% accuracy, and DT is 92.8% accuracy. By resampling the dataset and applying it to each algorithm, classification drivers' behavior using the random forest algorithm is suitable for predicting imbalanced datasets. Therefore, the experiment result of SMOTE combined with minority and ADASYN combined with minority shows a fair accuracy of approximately 70%.

*Keywords—ITS, Supervised learning, Binary classification, Resample data, ADASYN, SMOTE, Hyperparameters.*