



A data mining approach to investigate the factors influencing the crash severity of motorcycle pillion passengers



Ali Tavakoli Kashani, Rahim Rabieyan, Mohammad Mehdi Besharati*

School of Civil Engineering, Iran University of Science & Technology, Tehran, Iran

ARTICLE INFO

Article history:

Received 15 April 2014

Received in revised form 21 July 2014

Accepted 17 September 2014

Available online 7 October 2014

Keywords:

Motorcycle pillion passengers

Crash severity

Classification and regression trees

ABSTRACT

Introduction: Motorcycle passengers comprise a considerable proportion of traffic crash victims. During a 5 year period (2006–2010) in Iran, an average of 3.4 pillion passengers are killed daily due to motorcycle crashes. This study investigated the main factors influencing crash severity of this group of road users. **Method:** The Classification and Regression Trees (CART) method was employed to analyze the injury severity of pillion passengers in Iran over a 4 year period (2009–2012). **Results:** The predictive accuracy of the model built with a total of 16 variables was 74%, which showed a considerable improvement compared to previous studies. The results indicate that area type, land use, and injured part of the body (head, neck, etc.) are the most influential factors affecting the fatality of motorcycle passengers. Results also show that helmet usage could reduce the fatality risk among motorcycle passengers by 28%. **Practical Applications:** The findings of this study might help develop more targeted countermeasures to reduce the death rate of motorcycle pillion passengers.

© 2014 National Safety Council and Elsevier Ltd. All rights reserved.

1. Introduction

Traffic crashes are considered as a major public health problem worldwide, claiming 1.27 million annual deaths and between 20 and 50 million injuries (de Oña, López, & Abellán, 2013). Similarly, traffic crashes are one the main causes of death among the Iranian people. In addition, motorcycle is a popular transportation mode in Iran and is, unfortunately, involved in a significant proportion of fatal crashes. According to the Iran Forensic Medicine Organization report, motorcycle passengers are involved in a substantial proportion of fatal crashes. Based on the statistics, 30,901 motorcyclists have been killed in the traffic crashes during 2006 to 2010, which is about 25.7% of the total traffic crash fatalities occurring during this period. Furthermore, over 20% of the motorcyclists killed in these crashes are motorcycle passenger (and not riders) with an average age of only 26 years old. In other words, during this 5 year period, an average of 3.4 pillion passengers (except riders) have been killed due to motorcycle crashes every day in Iran. The high proportion of mortality and severe injury among this group of road users necessitates more investigations on their crash characteristics. Unfortunately, so far no studies have been conducted to address this issue in Iran. Therefore, the objective of this study was to identify the factors influencing the crash severity of pillion passengers by the means of CART method.

Classification and Regression Trees (CART) method applied in this study is a commonly used data mining technique. CART is a non-parametric model with no pre-defined relationships between the target variable and predictors. Moreover, decision tree models can identify and easily explain the complex patterns associated with crash risk and do not need to specify a functional form (Kashani & Mohaymany, 2011). Hence, CART models are good alternatives for analyzing the injury severity of traffic crashes (Chang & Chien, 2013).

In the next section of this paper, previous literature on motorcycle crash severity is reviewed. Section 3, represents an introduction to CART models and evaluation methods as well as a description of the study data. Next, the model results and discussions are presented in Section 4. Finally, the conclusion is described based on the study results.

2. Literature review

Although several previous studies have investigated the crash severity of vehicle passengers (Kashani & Mohaymany, 2011; Newgard, Lewis, & Jolly, 2002), the crash characteristics of the motorcyclists are substantially different from four-wheeled vehicles as the former use helmet rather than a safety belt and they are completely unprotected. So far, the effect of several variables on the injury severity of motorcyclists has been investigated in the previous studies. Such variables can be categorized into four classes: human factors, vehicle factors, crash characteristics, and roadway environmental attributes. The human factors include the rider

* Corresponding author at: School of Civil Engineering Iran University of Science and Technology Narmak, Tehran, Iran, P.O. Box 16846.

E-mail addresses: alitavakoli@iust.ac.ir (A. Tavakoli Kashani), rahimrabie@civileng.iust.ac.ir (R. Rabieyan), mehdi_besharati@civileng.iust.ac.ir (M.M. Besharati).

characteristics such as age, gender, riding license, safety helmet usage, and alcohol consumption. Several previous studies have reported that the rider age might influence the motorcycle crash severity (Cafiso, La Cava, & Pappalardo, 2012; Jou, Yeh, & Chen, 2012; Savolainen & Mannering, 2007). Elsewhere, (Evans & Frick, 1988; Mayrose, 2008; Ranney, Mello, Baird, Chai, & Clark, 2010; Savolainen & Mannering, 2007) studies have argued that helmet usage might play a significant role in reducing the injury severity of motorcycle passengers. Moreover, unlicensed male riders and consumption of alcoholic drinks can also be regarded as the factors influencing the severity of motorcycle crashes (Huang & Lai, 2011; Jou et al., 2012; Kasantikul, Ouellet, Smith, Sirathranont, & Panichabhongse, 2005; Savolainen & Mannering, 2007). However, in their studies, Quddus, Noland, and Chin (2002) and Rifaat, Tay, and de Barros (2012) have reported that the severity of crashes is higher for female riders.

On the other hand, some physical characteristics of motorcycles such as the size and production year were also found to significantly affect the severity of motorcycle crashes (Jou et al., 2012; Savolainen & Mannering, 2007). Vlahogianni, Yannis, and Golias (2012) suggested that roadway and environmental factors such as geometric design, road type, pavement condition, weather condition, area type (i.e., urban or rural), and illumination might also affect the severity of motorcycle crashes. For instance, we refer to studies conducted by Li, Doong, Huang, Lai, and Jeng (2009) and Savolainen and Mannering (2007) where they have found that the severity of crashes increased in rural areas, suburban regions, and dark conditions. Furthermore, the severity of crashes is reported to substantially reduce in bad weather, particularly in winters and wet pavements (Pai & Saleh, 2007; Rifaat et al., 2012; Savolainen & Mannering, 2007). Finally, the crash characteristics such as crash type (e.g., single- or multi-vehicle crash) and collision type (e.g., head-on or rear-end collisions) might also influence the severity of crashes. For instance, Savolainen and Mannering (2007) have indicated that if the crash is caused due to head-on collisions in multi-vehicle crashes, the probability of passenger fatality will be approximately six times greater.

In their study, Vlahogianni et al. (2012) presented a review on the literature pertaining to motorcycle crash. They showed that pillion passengers are not adequately addressed in the previous studies. Thus, influential factors affecting the fatality of pillion passengers were identified in the current study. Another innovation of this study is its spatial extent. Urban and rural crashes were investigated simultaneously in order to identify the interaction effects of the area type and other factors upon the crash severity.

3. Materials and methods

3.1. Model

To investigate the influence of factors on the severity of crashes, the dependent variable (injury type) was categorized into two levels of fatal and injury. Regarding the type of dependent variable in the present study, the Classification and Regression Trees (CART) was employed to model the effect of each variable on probability of crashes being fatal.

Fig. 1 shows the principle of the CART method in developing the classification tree. First, all of the data are concentrated at a node located at the top of the tree. Then this so-called “root node” is divided into two child nodes on the basis of a predictor variable (splitter) that maximizes the homogeneity (i.e., purity) of the two child nodes. In fact, the data in each child node are more homogenous than those in the upper parent node. This process is continued repeatedly for each child node until all the data in each node have the greatest possible homogeneity. This node is called a terminal node or “leaf” and has no branches. In other words, the principle behind tree growing is to recursively partition the target variable to minimize “impurity” in the terminal nodes. The most common measure of node impurity is the Gini criterion.

The Gini criterion is used to quantify the homogeneity based on computing the proportion of data that belong to a class. The Gini index is defined as:

$$g(t) = \sum_{i \neq j} p(j|t)p(i|t) \quad (1)$$

where i and j are categories of the target field, and:

$$p(j|t) = \frac{p(j,t)}{p(t)}, \quad p(j,t) = \frac{\pi(j)N_j(t)}{N_j}, \quad p(t) = \sum_j p(j,t) \quad (2)$$

where $\pi(j)$ is the prior probability value for category j , $N_j(t)$ is the number of records in category j of node t , and N_j is the number of records of category j in the root node. Note that when the Gini index is used to find the improvement for a split during tree growth, only those records in node t and the root node with valid values for the split-predictor are used to compute $N_j(t)$ and N_j , respectively.

In the CART method, tree growth will continue until there are only similar observations in each terminal node. To decrease its complexity, the tree is pruned using a cost-complexity measure that combines the precision criteria as opposed to complexity in the number of nodes

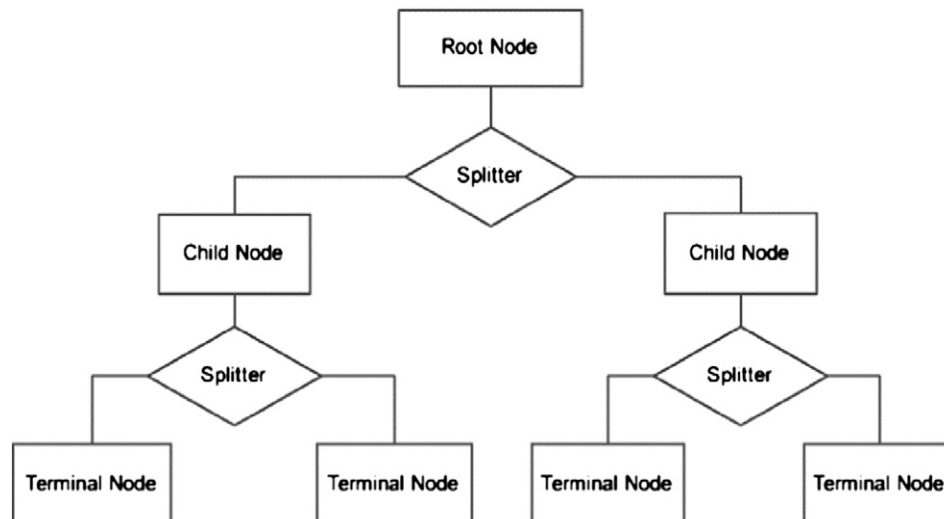


Fig. 1. General structure of a decision tree.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات