

Review of: "Crash severity analysis of vulnerable road users using machine learning"

Apostolos Ziakopoulos¹

¹ National Technical University of Athens

Potential competing interests: The author(s) declared that no potential competing interests exist.

Review for PLOS ONE

Paper ID: journal.pone.0255828

Paper Title: Crash severity analysis of vulnerable road users using machine learning

Review Type: First Review

Suggestion: Accept with major revisions. The following improvements are proposed:

The paper deals with a very interesting subject, exploring several Machine Learning algorithms for binary classification purposes of vulnerable road user injury severity. It uses appropriate methodologies, and it is very well written. The authors are commended for their effort overall. The research results are useful, though not extremely novel. There are some issues in the present form of the paper that should be addressed before it is accepted for publication.

- The abstract of the study can be made more compact, to increase its comprehensiveness. The first two sentences can be removed, for instance.
- The literature review should provide more details regarding the methodologies used in the reviewed studies (for instance, what were the algorithmic variations and the selected performance indicators?).
- The topic of the paper is well studied, and the literature review covers it adequately. As further suggestions, the authors are urged to consider some additional studies which are recent and pertinent for the literature review:

A study investigating spatio-temporal VRU crash patterns and discussing relevant risk factors:

Ouni, F., & Belloumi, M. (2018). Spatio-temporal pattern of vulnerable road user's collisions hot spots and related risk factors for injury severity in Tunisia. *Transportation research part F: traffic psychology and behaviour*, 56, 477-495.

A study modelling VRU crash severity with binary logistic regression:

Olszewski, P., Szagała, P., Rabczyński, D., & Zielińska, A. (2019). Investigating safety of vulnerable road users in selected EU countries. *Journal of safety research*, 68, 49-57.

A study regarding powered-two-wheeler crash severity with weather and traffic data:

Theofilatos, A., & Ziakopoulos, A. (2018). Examining injury severity of moped and motorcycle occupants with real-time traffic and weather data. *Journal of Transportation Engineering, Part A: Systems*, 144(11), 04018066.

An in-depth study regarding two-wheeler crashes (motorcycles and bicycles):

Brown, L., Morris, A., Thomas, P., Ekambaram, K., Margaritis, D., Davidse, R., ... & Wadji, F. (2021).

Investigation of accidents involving powered two wheelers and bicycles–A European in-depth study. *Journal of safety research*, 76, 135-145.

- A small inconsistency exists in the Methodology section: Firstly it is mentioned that “The crash data for VRU groups were collected through for the years 2013 to 2019.” Followed by “The crash data reported in QLD does not record any injury crashes since 2010.” Is the QLD database different than the TMR database? This point is not clear. And a statement follows stating that “...the data has injury severity levels information for the crashes resulting in casualty only and are reported as four scale injury severity levels...”. Probably the correct statement is: “The crash data reported in QLD does not record any non-injury crashes since 2010”. In any case, please clarify this point.
- For Table 1, the reviewer is very skeptical of vehicle/driver condition groupings. Different phenomena are presented as exclusive, which is not representative of reality. A rider can wear their helmet and still be inattentive or fatigued, for instance. Binary coding of each of these categories might lead to increased model predictions. The authors are urged to at least explore this venue. Furthermore, the authors are urged to add 100% totals in each row group for clarity on that table.
- The first paragraph of the Classification Approaches section is better suited for the literature review. It is suggested that the authors move it there and only keep a small introductory paragraph explaining their selection of ML algorithms. The terms in quotes on the last paragraph of the section (e.g. ‘rbf kernel’, ‘grid-search’ ‘C’ and ‘γ’ parameters, using the ‘grid-search’ algorithm), which the reviewer assumes are algorithmic/coding parameters, require explanations for the unfamiliar reader.
- For the KNN algorithm, the authors fine-tuned with respect to the K parameter as shown in Figure 1. However, there is no mention of how many combinations they tested for the other methodologies. Did the grid search lead to combinations in the order of 10^3 , or much more? Please provide the numbers if

you can, as it is important to obtain an idea of how deep the algorithm was explored. A mention of the runtime required would be beneficial as well.

- Does Figure 4 refer to Gini importance? If so, please state it clearly.
- The results from Figure 5 are counter-intuitive. How do the authors explain that younger ages are correlated with lower probability of severe crashes? This finding appears different from what the literature suggests. Please elaborate.
- Overall, from the reported results, model performance appears to be medium, with the main problem being low specificity. This is not reflected enough in the limitations paragraph. The authors are asked to re-write this paragraph with a focus on how these figures could be improved for the models to be more useful for predictions.
- The English language needs minor revisions throughout the paper to avoid small mistakes and to make the text flow more naturally.