

## Peer Review

# Review of: "R+R: Security Vulnerability Dataset Quality Is Critical"

Omid Bushehrian<sup>1</sup>

1. Shiraz University of Technology, Iran, Islamic Republic of

1- The “vulnerability detection systems” need to be more elaborated in the Introduction section. It is suggested that you add a figure to the Introduction section to illustrate the input/output of such systems. Adding a block diagram for such systems could be beneficial.

2- The authors discussed different dataset quality factors. Please also mention the degree of “balancedness” in each dataset. Are the datasets balanced or imbalanced regarding the classes of the samples?

3- The authors explained the models in section 3.2. The reader can better understand the models if the block diagram of models is also presented using figures.

4- To have better paper organization, it is suggested to move the content of “section 3.4” to “section 4”. “Section 4” should explain the experimental setup in addition to the results.

5- It is suggested that authors give a preview of the datasets used for the experiments by providing a few samples of them in the paper.

6- Why did the authors not use the F1-score, which is a better performance metric compared to the accuracy that is very sensitive to imbalanced data?

## Declarations

**Potential competing interests:** No potential competing interests to declare.