

A novel outlier detection approach for univariate datasets using deep neural networks

Olgun Aydin¹

Gdańsk University of Technology, Poland

Abstract

An outlier is a data point that appears distinctly different from other values in the dataset. Outliers occur due to changes in system behavior, human error, machine error, or natural deviations in populations.

Dealing with outliers is one of the key steps before moving forward with having any type of predictions. Before having predictions, building predictive models, outliers should be carefully analyzed in order not to reveal situations that could lead to potentially disastrous consequences [1]. Initially, the question of whether it is appropriate to remove outliers should be addressed [2]. Outliers in univariate datasets can be detected by using methods such as: location & scale based intervals, boxplot based methods, statistical tests and fitting mixture models to the dataset [3][4].

Main goal of this study is proposing a different perspective on outlier detection methods for univariate datasets. The method is combination of location & scale based methods and deep neural networks. The proposed method and traditional methods have been performed on simulated data and it is observed that the proposed method perform better for detecting extreme values and outliers than the traditional methods.

Keywords

Outliers, Deep neural networks, Outlier detection.

Acknowledgements

References

- [1] Hodge, V. J., Austin J., (2004), A survey of outlier detection methodologies, *Artificial Intelligence Review*. 22

- [2] Stephen W. Scheff, (2016). Chapter 9 - Outliers and Missing Data, Editor(s): Stephen W. Scheff, *Fundamental Statistical Principles for the Neurobiologist*, Academic Press, Pages 183-199, ISBN 9780128047538, <https://doi.org/10.1016/B978-0-12-804753-8.00009-9>.
- [3] D’Orazio, M.(2017), Outlier Detection in R: Some Remarks, uRos 2017, Bucharest/ROMANIA
- [4] D’Orazio, M. (2017), univOutl: Detection of Univariate Outliers. <https://cran.r-project.org/web/packages/univOutl/univOutl.pdf>