DSSV 2023 Data Science, Statistics & Visualisation
ECDA 2023 European Conference on Data Analysis

On the comparison of unsupervised anomaly detection algorithms

E. Ramentol[1], M. S. Kamalakar[1], P. M. Hegde[2], and S. Schwaar[1]

[1]Fraunhofer Institute for Industrial Mathematics ITWM, Department for Financial Mathematics, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany

[2]Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau, RPTU in Kaiserslautern, Gottlieb-Daimler-Straße, 67663 Kaiserslautern,Germany

Anomaly detection has become in the last few years a highly demanded task by companies. Thousands of fraudulent banking transactions are carried out every day, millions of users have behaviours that do not conform to the rules of specific applications, security systems are violated, patients with rare diseases appear in medical consultations, cameras capture 'weird things', and many times, anomaly detection algorithms seem blind to them.

Despite the continuous evolution of anomaly detection algorithms, there are still numerous problems to be solved. Years of work and effort by the scientific community have resulted in a considerable number of powerful algorithms for detecting anomalies [1, 2, 3, 4, 5, 6]. In our opinion, however, the problem is far from being solved. We will focus on two main problems, that are summarized in the following research questions:

1. How should we proceed when faced with a completely unsupervised anomaly detection problem?
2. How can we compare the performance of anomaly detection algorithms if the data is not labelled?

To solve these questions, we will take advantage of the rankings obtained by the algorithms. We will group the algorithms according to their similarity. To measure the algorithm similarity, a new metric will be introduced in this paper. The proposed metric is ahead of those of the state of the art in that the comparison can be as strict or soft as desired. Our new metric, which we have called "agreement degree" quantifies to what degree the scores obtained by the algorithms are similar. That is, to what degree they agree on the level of the anomaly of each example in the data set.

In our experimental study, we used 20 datasets and 11 anomaly detection algorithms [7]. For the algorithm similarity comparison we use our metric together with Kendall [8] and Spearman [9] correlations. The experimental study shows the superiority of the new metrics with respect to those of the state of the art. Our metric achieves a better separation between groups of algorithms based on their similarity. This means that it is possible to decant the algorithms with the best from those with the worst performance.

We present the new metric together with our study carried out on 20 data sets and show the advantages of our metric.

[1] H.-P. Kriegel, M. Schubert, and A. Zimek, "Angle-based outlier detection in high-dimensional data," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, (New York, NY, USA), pp. 444–452, Association for Computing Machinery, 08 2008.

[2] Z. He, X. Xu, and S. Deng, "Discovering cluster based local outliers," *Pattern Recognition Letters*, vol. 24, pp. 1641–1650, 06 2003.

[3] A. Lazarevic and V. Kumar, "Feature bagging for outlier detection," in *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, KDD '05, (New York, NY, USA), pp. 157–166, Association for Computing Machinery, 2005.

[4] M. Goldstein and A. R. Dengel, "Histogram-based outlier score (hbos): A fast unsupervised anomaly detection algorithm," in *KI-2012: Poster and Demo Track*, vol. 9, pp. 59 – 63, 2012.

[5] F. T. Liu, K. Ting, and Z.-H. Zhou, "Isolation forest," in *2008 Eighth IEEE International Conference on Data Mining*, pp. 413–422, 2008.

[6] M. Breunig, H.-P. Kriegel, R. Ng, and J. Sander, "Lof: Identifying density-based local outliers," in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, SIGMOD '00, (New York, NY, USA), pp. 93–104, Association for Computing Machinery, 2000.

[7] Y. Zhao, Z. Nasrullah, and Z. Li, "Pyod: A python toolbox for scalable outlier detection," *Journal of Machine Learning Research*, vol. 20(96), pp. 1–7, 2019.

[8] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.

[9] C. E. Spearman, "The proof and measurement of association between two things," *The American Journal of Psychology*, vol. 15, no. 1, pp. 72–101, 1904.