



# nonconform: Conformal Anomaly Detection (Python)

Oliver Hennhöfer <sup>1</sup>

<sup>1</sup> Intelligent Systems Research Group (ISRG), Karlsruhe University of Applied Sciences (HKA),  
Karlsruhe, Germany

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

## Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

## Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

## License

Authors of papers retain copyright,  
and release the work under a  
Creative Commons Attribution 4.0  
International License ([CC BY 4.0](#)).

## Summary

nonconform is a Python package that implements conformal anomaly detection methods (Laxhammar & Falkman, 2010) to provide statistically principled uncertainty quantification for unsupervised anomaly detection. The ability to quantify uncertainty is a fundamental requirement for AI systems in safety-critical domains, where reliable decision-making is essential. Based on the principles of conformal inference (Lei & Wasserman, 2013; Papadopoulos et al., 2002; Vovk et al., 2005), nonconform converts the anomaly scores from underlying anomaly detection models into statistically valid  $p$ -values. These  $p$ -values can be systematically adjusted using methods that control the False Discovery Rate (FDR) (Bates et al., 2023; Benjamini & Hochberg, 1995), allowing users to move beyond arbitrary thresholds. By calibrating detector models to align anomaly scores with their empirical false alarm rates, this approach provides statistical guarantees. The library integrates with the popular pyod library (Chen et al., 2024; Zhao et al., 2019), making it easy to apply these techniques to a wide range of anomaly detection models.

## Statement of Need

A primary challenge in anomaly detection is managing the rate of false positives. In many applications, frequent false alarms can lead to *alert fatigue*, ultimately rendering a detection system impractical. For research in areas such as fraud detection, medical diagnostics, and industrial quality control, controlling the proportion of false positives is crucial. nonconform addresses this by framing anomaly detection as a set of statistical hypothesis tests, enabling the control of the False Discovery Rate (FDR).

The package provides wrappers for various anomaly detection models (e.g., Autoencoder, IsolationForest, One-Class SVM) and includes a range of conformalization strategies. These strategies can compute classical conformal  $p$ -values or modified *weighted* conformal  $p$ -values (Jin & Candès, 2023), which are suitable even in low-data regimes (Hennhofer & Preisach, 2024). The need for weighted conformal  $p$ -values arises when the statistical assumption of exchangeability is violated due to covariate shift between calibration and test data. By providing these tools, nonconform enables researchers to create anomaly detectors with outputs that can be statistically controlled to cap the FDR at a nominal level. The methods support batch, batch-streaming, and online deployments, making them adaptable to various research settings.

The core assumption for the methods in nonconform is that the data is exchangeable, meaning the joint probability distribution is invariant to the order of observations. This assumption is suitable for many cross-sectional data analysis tasks but not for time-series data where temporal ordering is informative.

## Acknowledgements

- This work was conducted in part within the research projects *Biflex Industrie* (grant number 01MV23020A) and *AutoDiagCM* (grant number 03EE2046B) funded by the *German Federal Ministry of Economic Affairs and Climate Action (BMWK)*.
- Bates, S., Candès, E., Lei, L., Romano, Y., & Sesia, M. (2023). Testing for outliers with conformal p-values. *The Annals of Statistics*, 51(1). <https://doi.org/10.1214/22-aos2244>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.2307/2346101>
- Chen, S., Qian, Z., Siu, W., Hu, X., Li, J., Li, S., Qin, Y., Yang, T., Xiao, Z., Ye, W., Zhang, Y., Dong, Y., & Zhao, Y. (2024). PyOD 2: A python library for outlier detection with LLM-powered model selection. *arXiv Preprint arXiv:2412.12154*.
- Hennhofer, O., & Preisach, C. (2024). Leave-One-Out-, Bootstrap- and Cross-Conformal Anomaly Detectors. *2024 IEEE International Conference on Knowledge Graph (ICKG)*, 110–119. <https://doi.org/10.1109/ICKG63256.2024.00022>
- Jin, Y., & Candès, E. J. (2023). Model-free selective inference under covariate shift via weighted conformal p-values. <https://api.semanticscholar.org/CorpusID:259950903>
- Laxhammar, R., & Falkman, G. (2010). Conformal prediction for distribution-independent anomaly detection in streaming vessel data. *Proceedings of the First International Workshop on Novel Data Stream Pattern Mining Techniques*, 47–55. <https://doi.org/10.1145/1833280.1833287>
- Lei, J., & Wasserman, L. (2013). Distribution-free prediction bands for non-parametric regression. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 76(1), 71–96. <https://doi.org/10.1111/rssb.12021>
- Papadopoulos, H., Proedrou, K., Vovk, V., & Gammerman, A. (2002). Inductive confidence machines for regression. In *Machine learning: ECML 2002* (pp. 345–356). Springer Berlin Heidelberg. [https://doi.org/10.1007/3-540-36755-1\\_29](https://doi.org/10.1007/3-540-36755-1_29)
- Vovk, V., Gammerman, A., & Shafer, G. (2005). *Algorithmic learning in a random world*. Springer-Verlag. ISBN: 0387001522
- Zhao, Y., Nasrullah, Z., & Li, Z. (2019). PyOD: A python toolbox for scalable outlier detection. *Journal of Machine Learning Research*, 20(96), 1–7. <http://jmlr.org/papers/v20/19-011.html>