ABSTRACT CONTRIBUTION

Testing the performance of catch22 time-series features for classifying short, noisy, and periodic signals

Consuelo Rojas[1,*], Ulrich Parlitz[2,3] and Cristina Masoller[1]

[1]Universitat Politècnica de Catalunya, Dept. of Physics. Terrassa, Spain and [2]Biomedical Physics Group, Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany and [3]Institute for the Dynamics of Complex Systems, University of Göttingen, Göttingen, Germany

[*]consuelo.belen.paz.rojas@upc.edu

Key words: Time series analysis; machine Learning; feature extraction

In this study, we evaluate the performance, data requirements, and computational efficiency of the *Catch22*[1] feature extraction method in comparison with alternative time-series representations for a classification task. The datasets were generated using sinusoidal signals and nonlinear dynamics from the FitzHugh-Nagumo (FHN) model. We systematically varied the signals duration, resolution, noise, class separation and sample size of the data set to review their impact on classification performance.

Six data representations were analyzed using support vector machine (SVM) classifier: raw signals, Principal Component Analysis (PCA) applied to raw data, Fourier-based features (FFT), PCA applied to FFT features, *Catch22* features, and PCA applied to *Catchh22* features. The results indicate that *Catch22* provides good discriminative capability, even in low-information scenarios. For example, FHN signals with only 21 data points remain separable when parameter differences are on the order of 10%-20%. And the feature extraction and the classification are computational efficient, with a total processing times below 50ms for 200 samples of 21 points each. However, despite its robustness, *Catch22* does not outperform simpler representations, such as raw signals or FFT features. This suggest that even short and noisy signals retain sufficient structure to enable accurate classification with relatively simple decision boundaries.

Acknowledgements

This work was funded by the European Union, under HORIZON-MSCA-2022-DN, *Improving BiomEdical diagnosis through LIGHT-based technologies and machine learning* "BE-LIGHT" (GA nº 101119924 - BE-LIGHT) project. C. M. was also funded by Agencia Estatal de Investigación (PID2024-160573NB-I00) and Institució Catalana de Recerca i Estudis Avançats (ICREA Academia).

References

1. Carl H. Lubba, Sarab S. Sethi, Philip Knaute, Simon R. Schultz, Ben D. Fulcher, and Nick S. Jones. catch22: Canonical time-series characteristics. *Data Mining and Knowledge Discovery*, 33(6):1821–1852, 2019.

Compiled on: Tuesday 21st April, 2026.