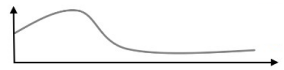


# Gradual patterns to improve explainability of black box models for time series classification

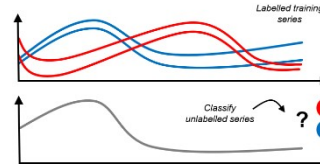
Durande Kamga, Anthony Fleury, Jerry Lonlac, Engelbert Mephu

# Time Series Classification (TSC)

Heart rate monitoring  
Always pay attention to heart health

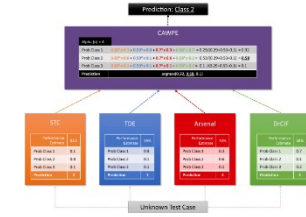


Time series

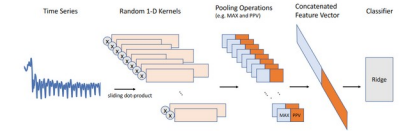


TSC

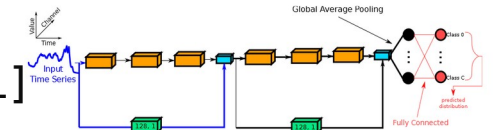
HiveCote2 [1]



ROCKET[1]:



InceptionTime [1]

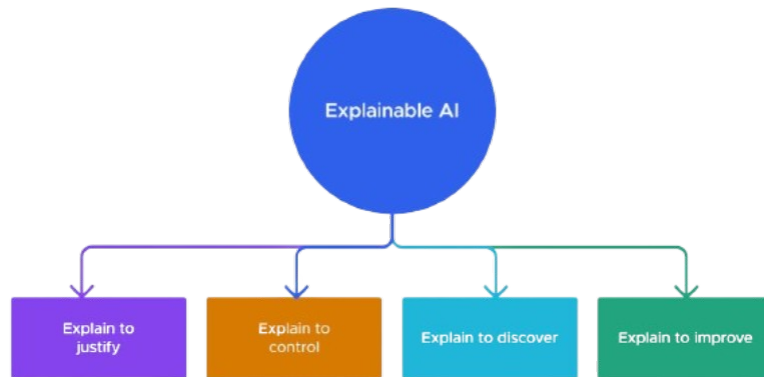


Should we trust one of these models and use it in critical applications?

# Explainable AI

**Explainable AI** is the set of methods and models that make the behaviour and predictions of machine learning systems understandable to humans [2]

Explainability is the ability of an ML model to provide the elements that influenced its decision[2]



**These methods can be**

- Interpretable by design vs Post-Hoc
- Local vs global

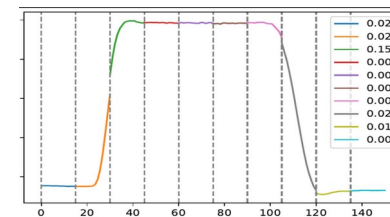
**The most used methods in general**

- Local approximation of complex model with a linear one

**SHAP and KernelSHAP**

- Marginal contribution of each feature with strong theoretical from game theory

**The most used methods for time series**



No relation between classes

# Gradual pattern

**Gradual patterns** can be used to express knowledge in the form of "more/less A, more/less B" between quantitative variables.[3]

sugar	sport(h)	stress
2.1	3	82
8.0	5	25
3.0	2	135
4	1	160

, )

**Gradual patterns to improve explainability of black-box models for time-series classification**

Durand B., Kampa N., Engelbert Mephu N., Anthony Fleury, Jerry Lonlac

### Context and Definitions

**Time series:** a sequence of ordered continuous values

**Time Series Classification (TSC)** involves training a model from a collection of time series in order to predict a target variable[1]

**State of the art models for TSC**

ROCKET[1], InceptionTime [1], HiveCost2 [1]

Should we trust one of these models and use it in critical applications? Great accuracy but how can we understand model prediction?

### Explainable AI to the rescue

**Explainable AI (XAI)** is the set of methods and models that make the behaviour and predictions of machine learning systems understandable to humans[2]

**Family:**

- Features importance
- Rules
- Counterfactuals

**XAI on TSC**

Grad-CAM, LASTS

# References

- [1] M. Middlehurst, P. Schäfer, et A. Bagnall, « Bake off redux: a review and experimental evaluation of recent time series classification algorithms ». arXiv, 25 avril 2023. doi: [10.48550/arXiv.2304.13029](https://doi.org/10.48550/arXiv.2304.13029).
- [2] S. Ali *et al.*, « Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence », *Information Fusion*, vol. 99, p. 101805, nov. 2023, doi: [10.1016/j.inffus.2023.101805](https://doi.org/10.1016/j.inffus.2023.101805).
- [3] L. Di-Jorio, A. Laurent, et M. Teisseire, « Mining Frequent Gradual Itemsets from Large Databases », in *Advances in Intelligent Data Analysis VIII*, N. M. Adams, C. Robardet, A. Siebes, et J.-F. Boulicaut, Éd., in *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2009, p. 297-308. doi: [10.1007/978-3-642-03915-7\\_26](https://doi.org/10.1007/978-3-642-03915-7_26).

# THANKS