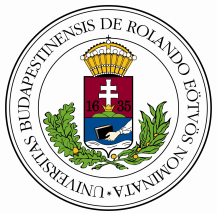
Time-series in HP Initialization of ML Techniques

Tomáš Horváth, Rafael G. Mantovani, André C. P. L. F. de Carvalho



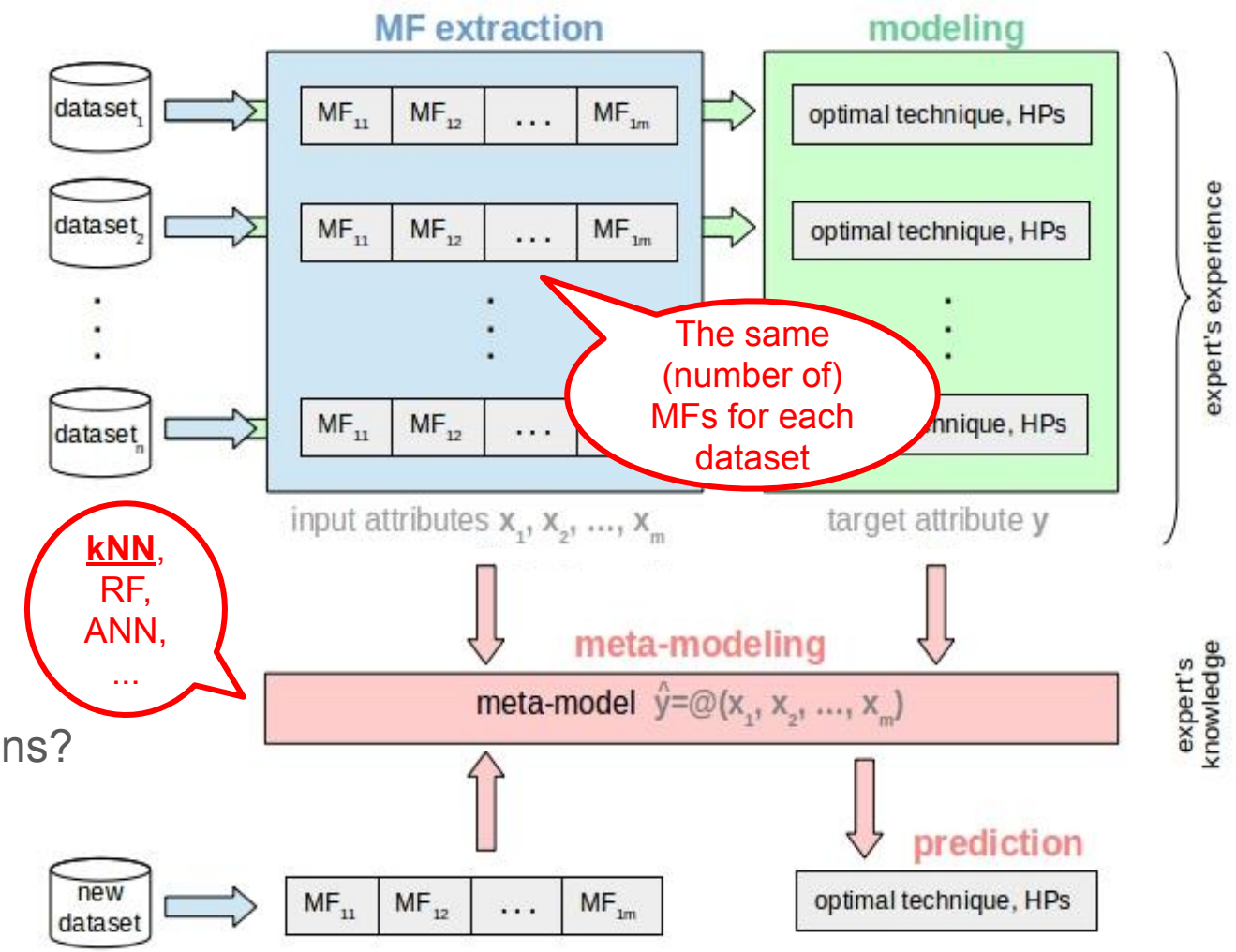
22nd International Conference on Intelligent Data Engineering and Automated Learning
The University of Manchester, UK
25-27 November 2021
(Virtual)

Metalearning

Decisions about
- used models
- HP settings,

are based on our
- **experience** and
- **knowledge**.

How to support decisions?
- Metalearning



Metafeatures

MF type	Abbr	#	Description
Simple	SL	17	Simple measures
Statistical	ST	7	Statistics measures
Inf. theoretic	IT	8	Information the sures
Landmarking	LM	9	Performanc e ML algorithms
Model-based	MB	17	Featur cted from decision trees
Time	TI	5	Ex time of some ML algorithms
Complexity	CO	14	Measures analyzing complexity
Complex Network	CN	9	Complex network property measures

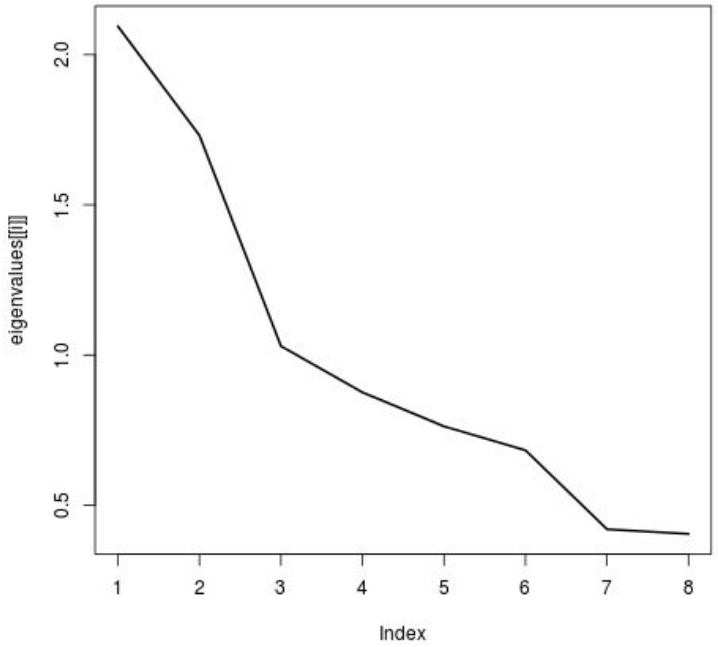
"NO FREE LUNCH"

RQ: PCA-based MFs for HP Initialization?

- **No tuning** needed
  - e.g. which MFs to use
- Relatively good **descriptive** power
  - important dimensions/variances
- Often, used in **pre-processing**

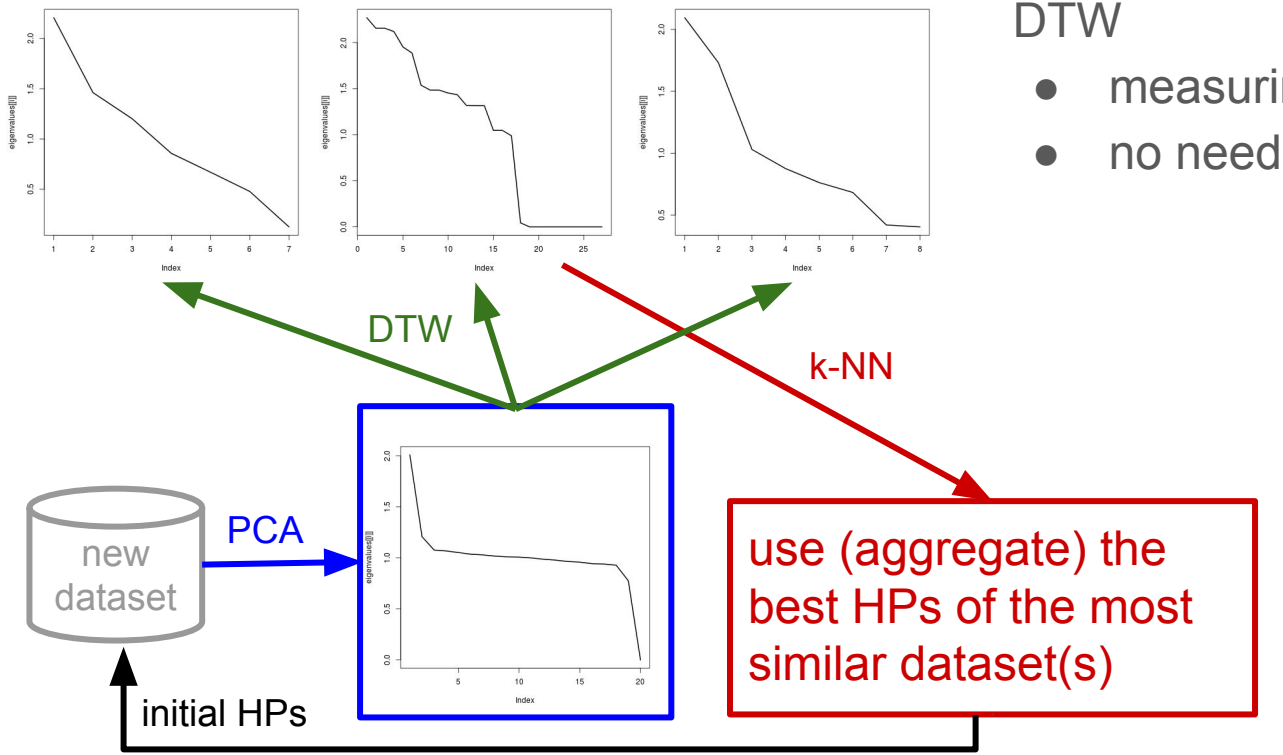
How to get equal number of MFs?

- selection
- aggregation (binning)



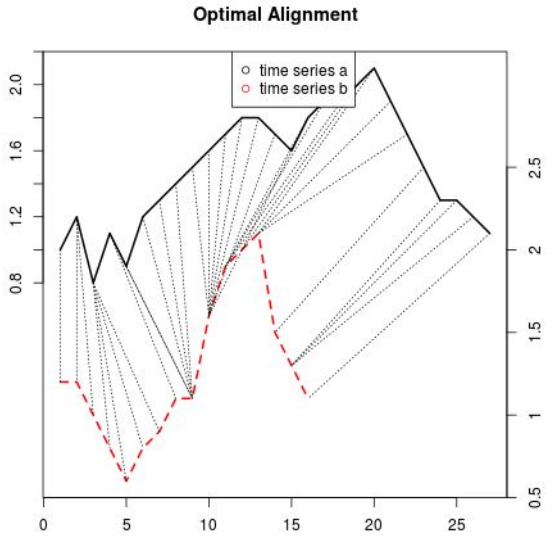
Limited use of PCA-based MFs in the literature

kNN using Dynamic Time Warping



DTW

- measuring **time-series similarity**
- no need for equal length series



Experiment Settings

Use-case

- 50 multi-class datasets
- HP initialization of SVM and DT
- balanced accuracy

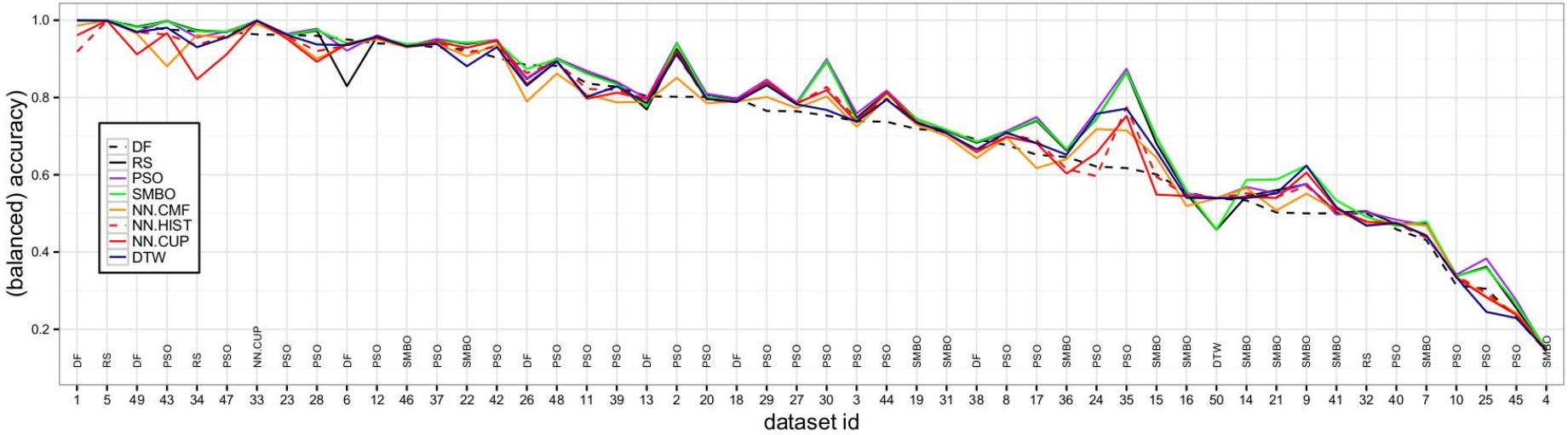
Baselines

- default HPs (**DF**)
- tuned HPs using (30 runs)
  - **RS** with 2500 (SVM) and 5000 (DT) trials
  - **PSO** (2500, 5000 trials)
  - **SMBO** (200 trials using RF)

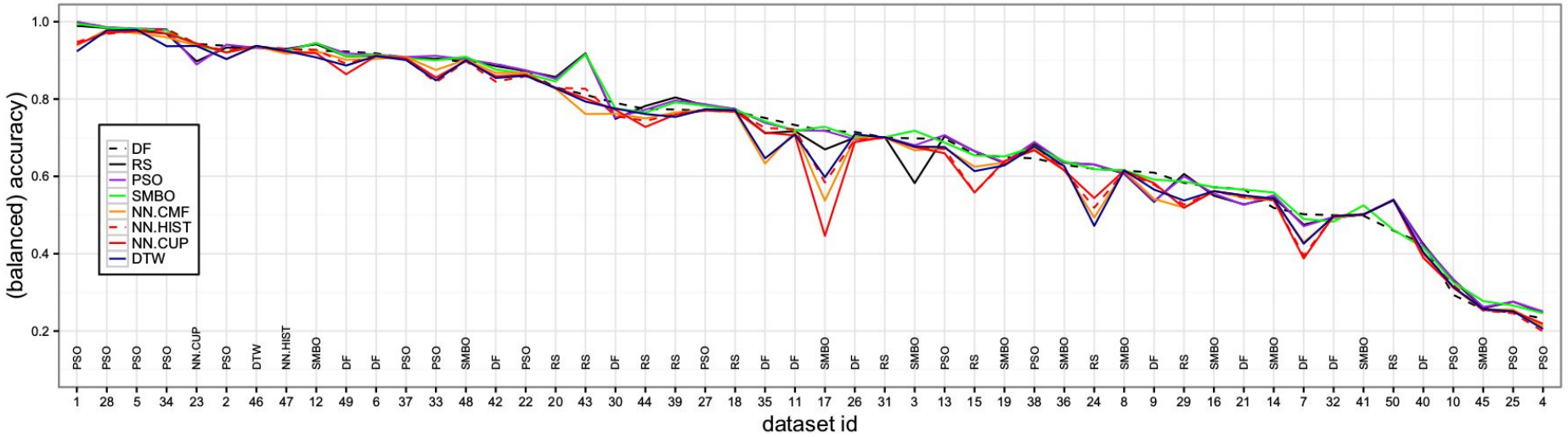
- kNN metalearner (**CMF**)
  - 255 different MF vectors
    - all the combinations of 8 MF types (5 to 86 MFs)
  - 2 types of PCA MF aggregations (**NN-HIS**, **NN-CUP**)
  - 4 vector similarity measures
    - Euclidean dist., inner product, cosine similarity, Pearson correlation

In total, around 1.6 billion computations.

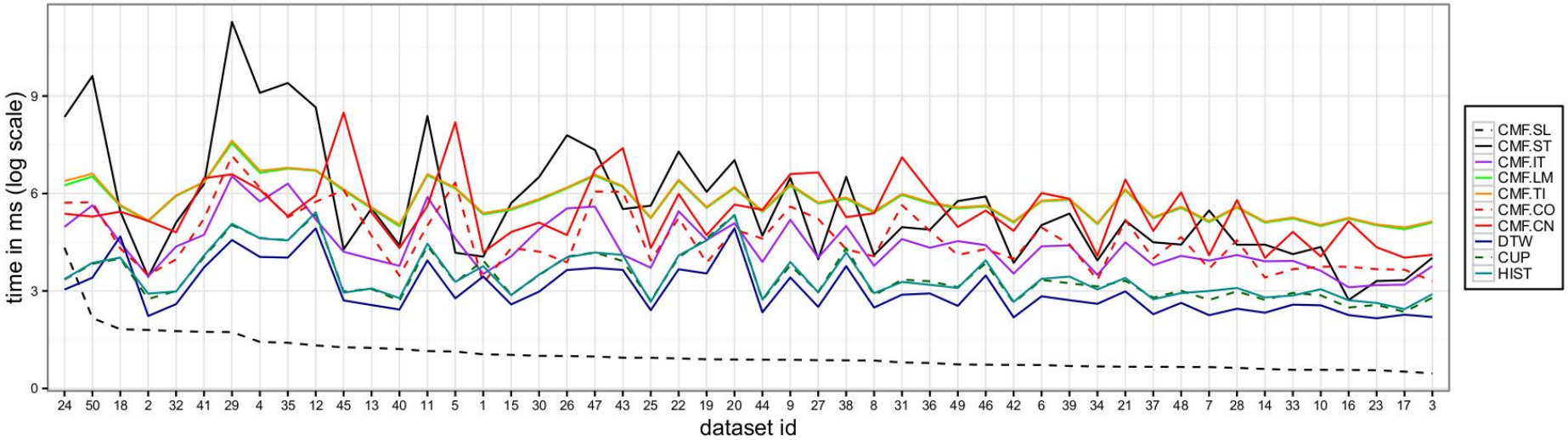
Results for SVM



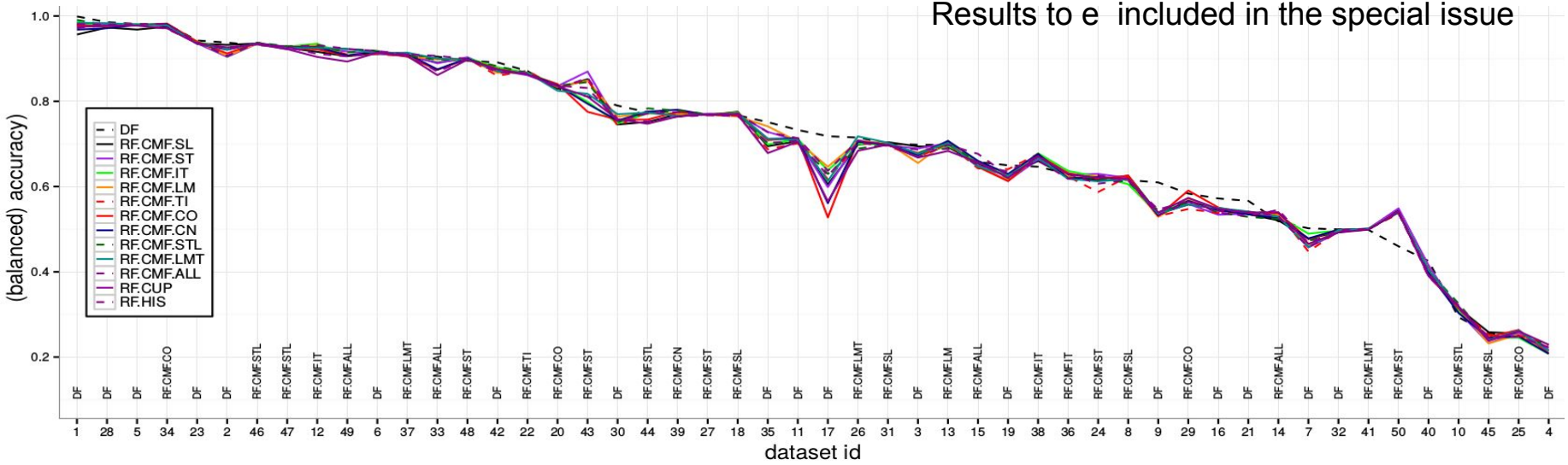
Results for DT



Results: Runtime for MF Extraction



Results: No Free Lunch



Other results in HP tuning

An empirical study on hyperparameter tuning of decision trees

- SMBO vs. black-box optimization
- IEEE, ArXiv.org

Effects of Random Sampling on SVM Hyper-parameter Tuning

- influence of sampling on HP tuning
- Springer

Thanks for your attention!

Best Paper Award

IS PRESENTED TO

Tomáš Horváth, Rafael G. Mantovani and André C. P. L. F. de Carvalho

FOR THE PAPER ENTITLED:

TIME-SERIES IN HYPER-PARAMETER INITIALIZATION OF MACHINE LEARNING TECHNIQUES

Springer
FINANCIAL SPONSOR

MANCHESTER 1824
The University of Manchester

IEEE Computational Intelligence Society
UK & IRELAND
TECHNICAL CO-SPONSOR

IDEAL 2021 GENERAL CO-CHAIRS
HUJUN YIN, DAVID CAMACHO, PETER TINO

27TH NOVEMBER 2021

tomas.horvath@inf.elte.hu