UNIVERSIDAD DE MÁLAGA | uma.es

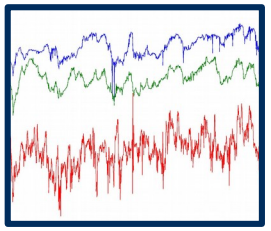
Comparing Deep Recurrent Networks Based on the MAE Random Sampling, a First Approach

Andrés Camero andrescamero@uma.es

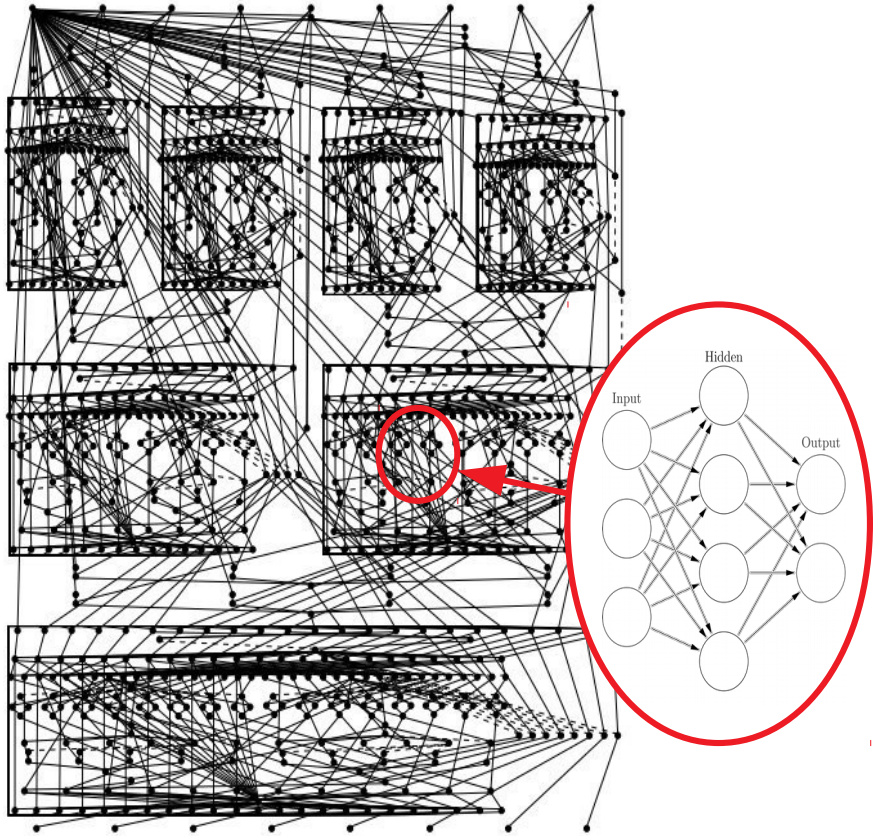
Jamal Toutouh jamal@lcc.uma.es

Enrique Alba eat@lcc.uma.es

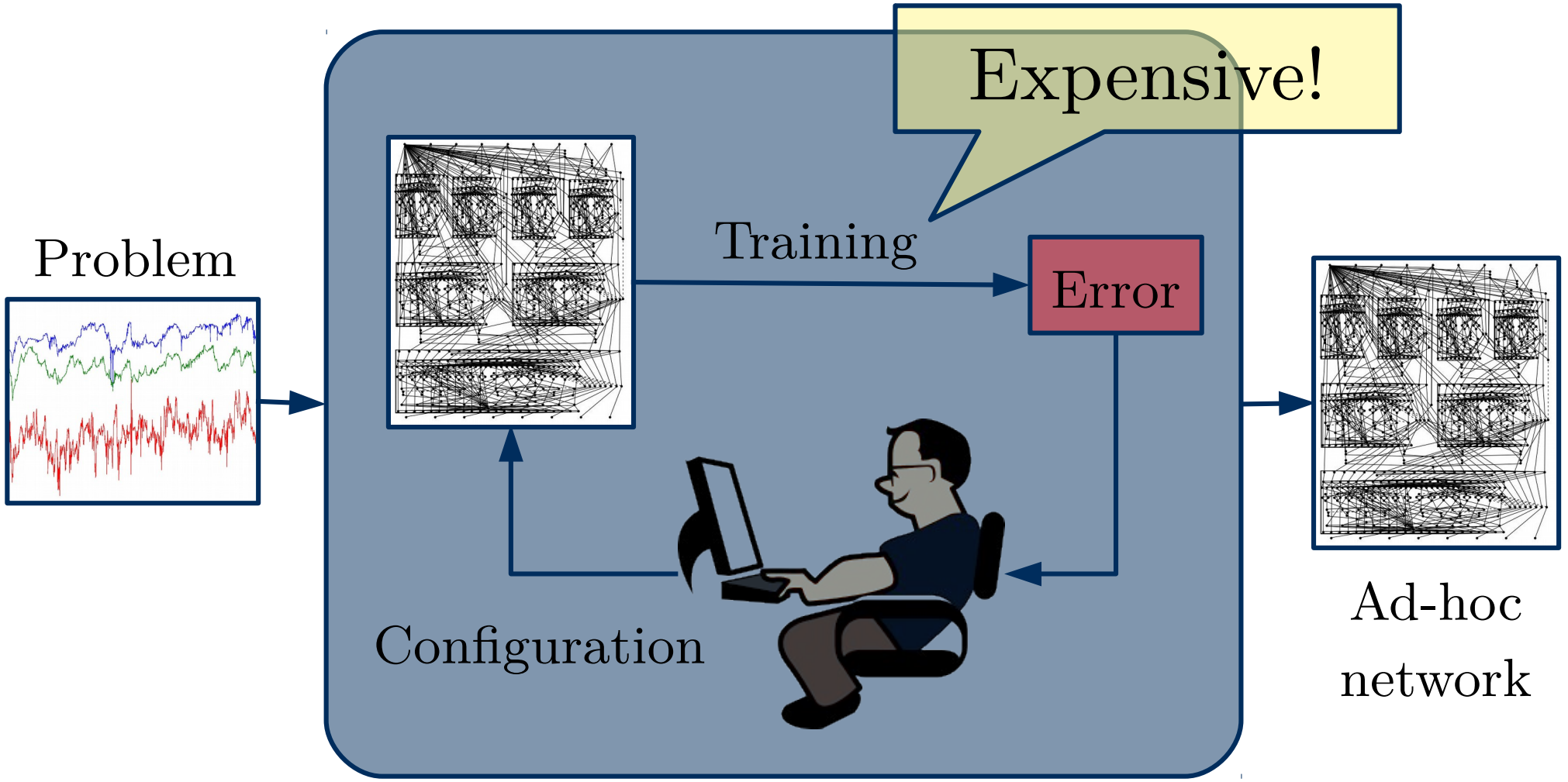
There is no silver bullet



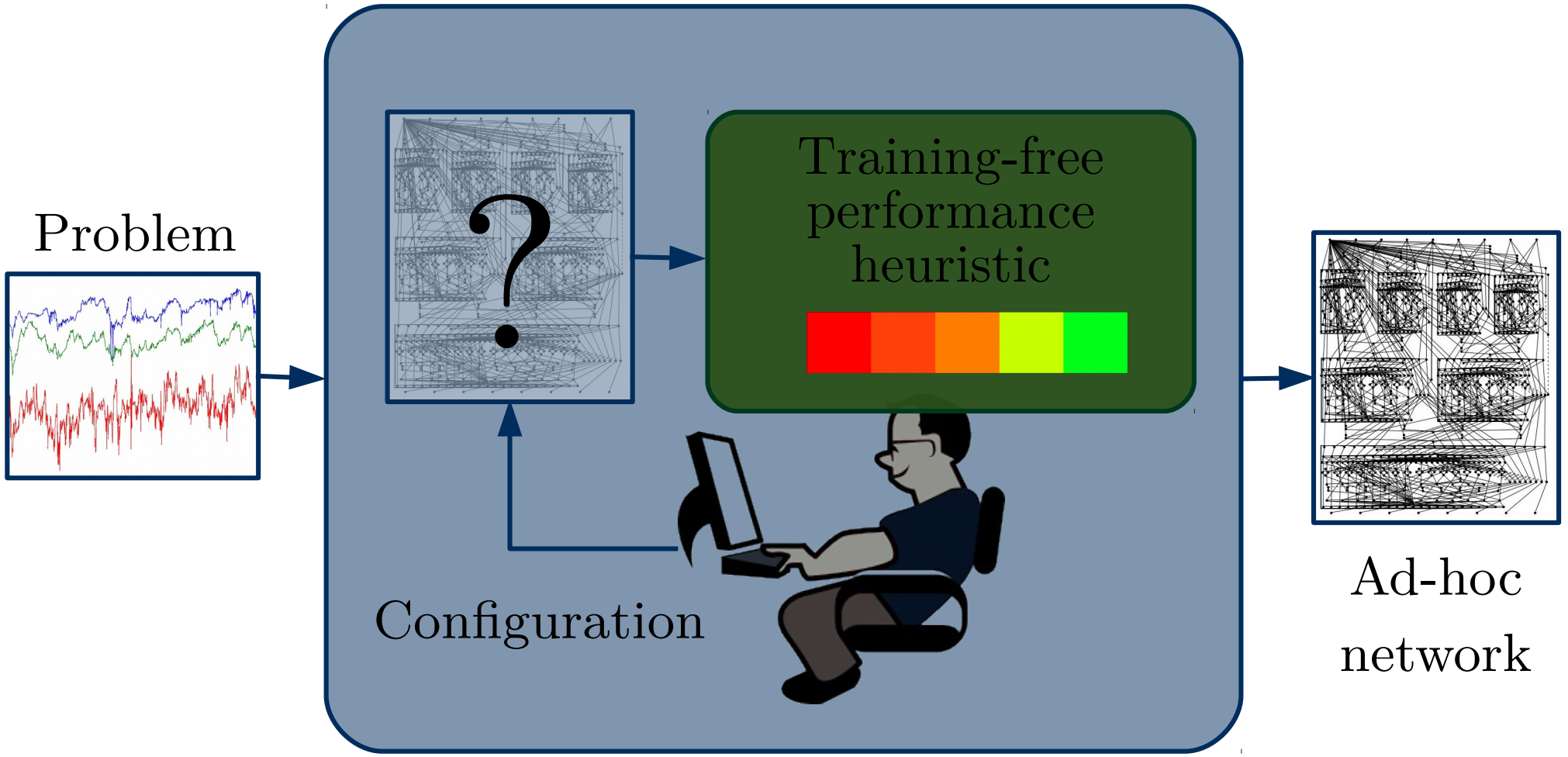
Lorem ipsum dolor sit amet, consectetuer adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoret dolore magna aliquam erat volutpat. Ut



Tailored architectures

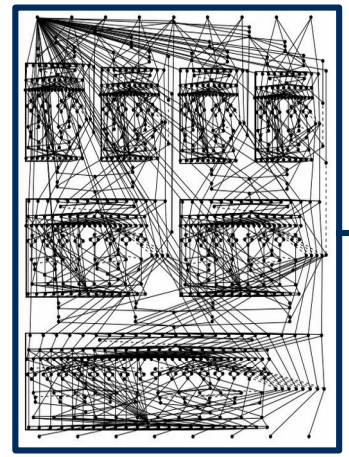


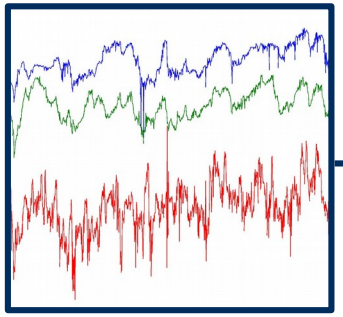
Tailored architectures

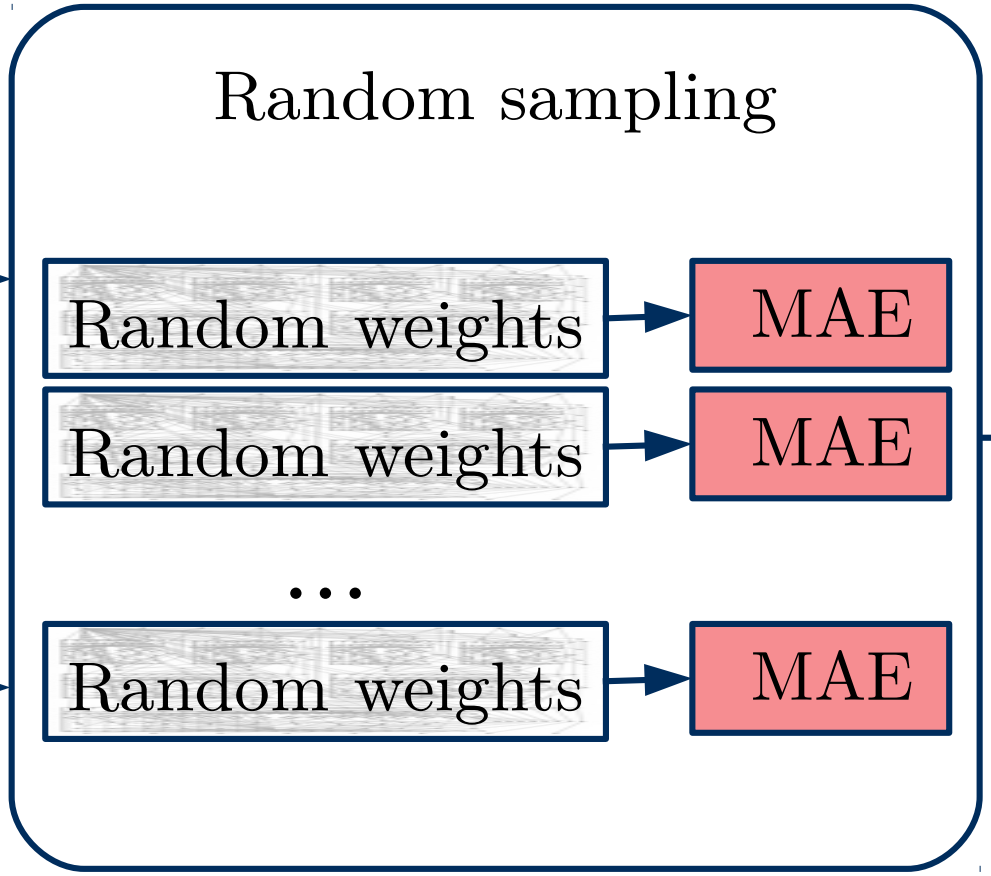


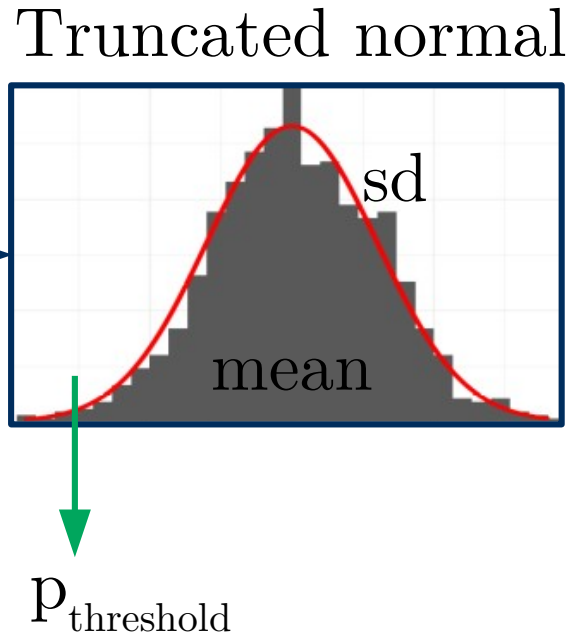
MAE random sampling*

...extended to multiple-hidden-layers





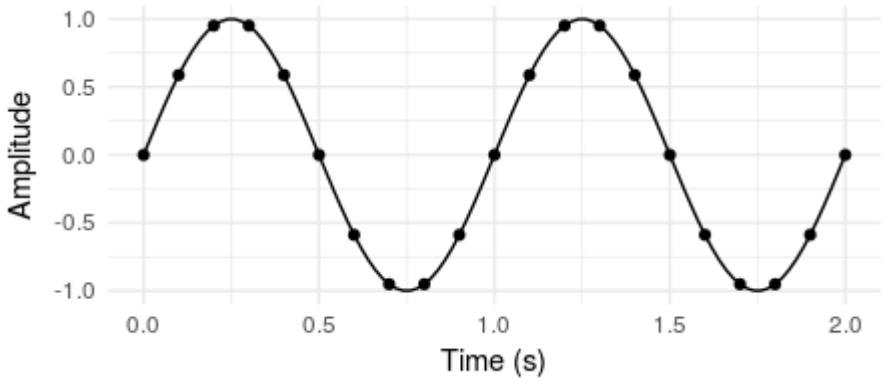




* Camero, A., Toutouh, J., Alba, E.: Low-cost recurrent neural network expected performance evaluation. arXiv preprint arXiv:1805.07159 (may 2018)

The experiment

The problem

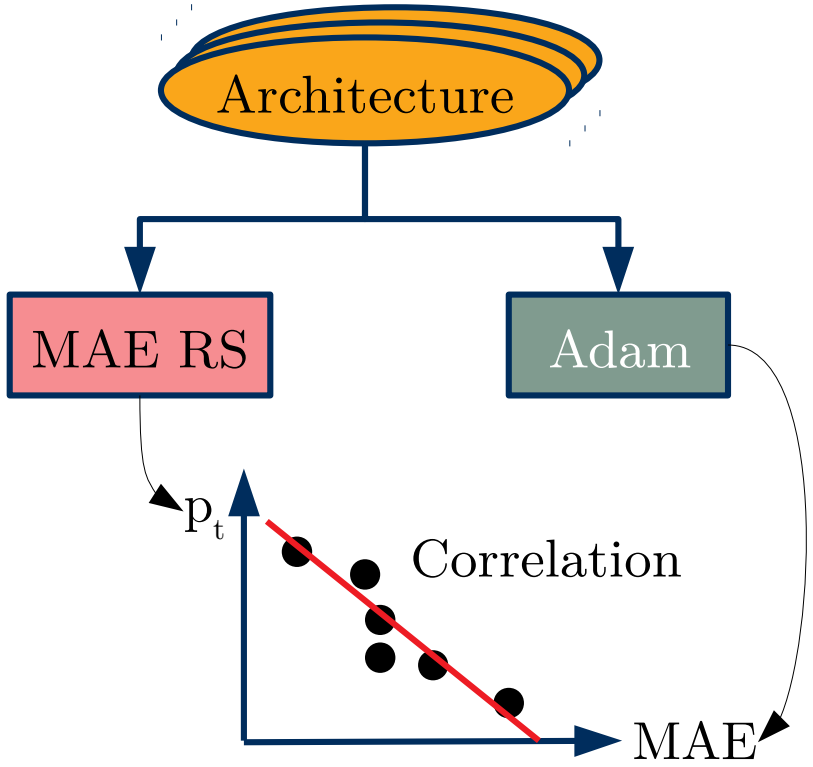


$$y(t) = A \cdot \sin(2\pi \cdot f \cdot t + \phi)$$

The architectures

1, 2 and 3 stacked LSTM layers
+ fully connected *tanh* layer

The comparison



Correlation

	Epochs	Cor Mean	Cor Sd	Cor log $p_{0.01}$
Single stacked LSTM layer	1	-0.447	-0.317	-0.211
	10	-0.726	-0.431	-0.321
	100	-0.790	-0.641	-0.650
	1000	-0.668	-0.458	-0.515
Two stacked LSTM layers	1	-0.086	-0.135	-0.171
	10	-0.450	-0.632	-0.635
	100	-0.709	-0.827	-0.905
	1000	-0.695	-0.843	-0.922
Three stacked LSTM layers	1	-0.334	-0.447	-0.475
	10	-0.546	-0.724	-0.745
	100	-0.720	-0.869	-0.906
	1000	-0.130	-0.873	-0.911

Table 1. Correlation between the MAE (Adam) and the MAE random sampling results.

Memory and time

	Mean time [s]	Sd time	Mean mem [MB]	Sd mem
Adam 1000 epochs	996	0.006	127	6.338
MAE random sampling	6	0.001	150	98.264

Table 2. Time and memory usage comparison

So far...

- ✔ MAE random sampling seems to be a good low-cost, training-free, rule of thumb method for predicting the performance, but...

- ✔ We need more evidence!

UNIVERSIDAD DE MÁLAGA | uma.es

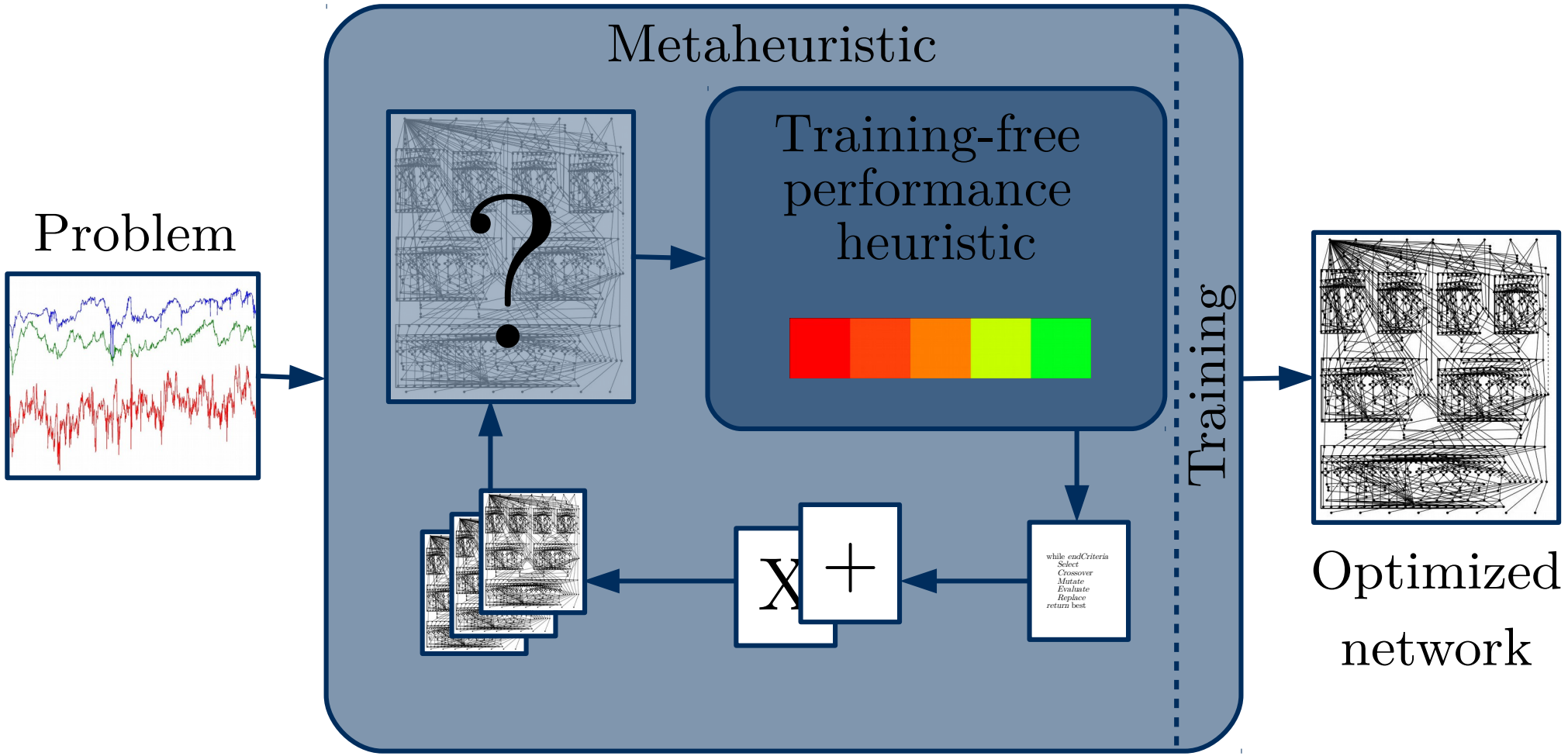
Comparing Deep Recurrent Networks Based on the MAE Random Sampling, a First Approach

Andrés Camero andrescamero@uma.es

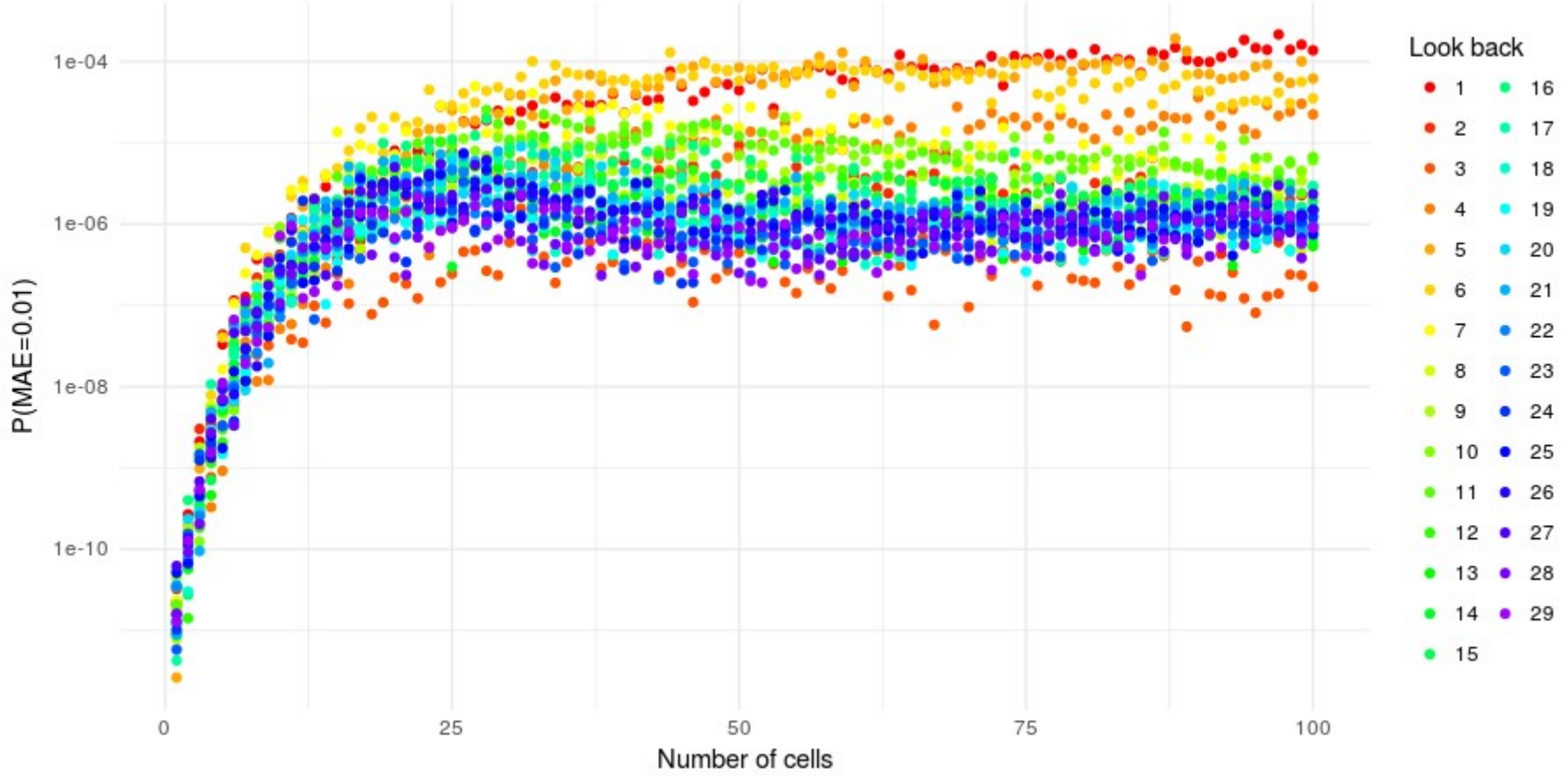
Jamal Toutouh jamal@lcc.uma.es

Enrique Alba eat@lcc.uma.es

Low cost network optimization



Single stacked LSTM layer



Two stacked LSTM layers