

## **Automatic data imputation in time series processing using neural networks for industry and medical datasets**

- Juan Ignacio Porta,<sup>1,2</sup> Martin Ariel Dominguez,<sup>2</sup> Francisco Antonio Tamarit<sup>1,2</sup>

<sup>1</sup>*Instituto de Física Enrique Gaviola de Córdoba, CONICET-UNC*

<sup>2</sup>*Facultad de Matemática Astronomía y Física - Universidad Nacional de Córdoba*

Time series classification and regression techniques help solve problems in many knowledge areas, including medicine, electronics, industry, and even music. When we apply them to real-life issues, a common obstacle is the lack of data in intervals within a time series. Usually, to solve it, the missing data is populated with information highly dependent on available datasets, which requires prior analysis. This paper addresses the problem in a novel way, automatically filling the missing data using a mixture of techniques and letting the prediction model decide which filling is better. We tested our approach for classification in industrial and medical datasets and for regression, we used a dataset containing COVID-19 information.

Our results are very competitive, and our approach improves the state-of-the-art models. We obtain better performance in all the experiments for the selected quality measures. Most importantly, the improvement is more statistically significant when the amount of missing data is higher.