

Towards the Analysis of Weak Coherent, Multidimensional Temporal Data with Deep Learning

Patrik Reiziger^{1*}, Bálint Gyires-Tóth¹

¹Budapest University of Technology and Economics, Department of Telecommunications and Media Informatics

*E-mail: rpatrik96@sch.bme.hu

Due to the advancements of information technologies, a vast amount of temporal data is available for analysis in production facilities, life sciences or in everyday life. As the information collected by sensor networks, self-driving cars or the stock market is rather diverse, extracting useful features to help decision-makers is very difficult. The problem is of multiple nature: first, the acquired information differs in the captured timescale, thus, their direct comparison is often not possible. On the other hand, the useful features are often represented by the fraction of the data sources.

As the relationships imposed by the underlying process are hard to decipher for humans, intelligent solutions are needed to automate and simplify data analysis. Due to the fact that in recent years solutions provided by the application of artificial intelligence was proven to be rather successful in several scientific domains including temporal data analysis – even for multiscale datasets, as in [1,3] –, we decided to exploit the possibilities of this toolbox during analysis. The last advancements include novel neural solutions, among others the Seq2Seq model of [4] or the Neural Transducer [2], which will also be used for our research.

In this work, we present our preliminary findings of how deep neural networks can be used to identify the core components of multidimensional temporal data, furthermore, the prediction capability of the networks will also be analyzed. Our comparison includes standard benchmarks such as the sum task, which investigates how well the network is able to estimate the sum of some specific entries within a window, given by a mask vector – with the help of which, the capability of the network to extract the important features can be assessed.

The results are compared to a collection of standard methods, such as naïve prediction, rolling-mean prediction, etc. Which is rather important, as in the case of sequence modeling, only evaluating the value of the objective function may imply a good result. But it should be noted that as in most cases (at least piecewise) continuous functions are to be estimated, even the naïve approach results in a low loss value. Thus, investigating the confusion matrix, and more importantly, the dynamics of the prediction compared to the target is crucial.

Acknowledgement

The research considered here has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

- [1] Chung, Junyoung, Sungjin Ahn, and Yoshua Bengio. "Hierarchical multiscale recurrent neural networks." *arXiv preprint arXiv:1609.01704* (2016).
- [2] Jaitly, N., Sussillo, D., Le, Q. V., Vinyals, O., Sutskever, I., Bengio, S. "A Neural Transducer." *arXiv preprint arXiv:1511.04868* (2015).
- [3] : Koutnik, J., Greff, K., Gomez, F., & Schmidhuber, J. (2014, January). A Clockwork RNN. In *International Conference on Machine Learning* (pp. 1863-1871).
- [4] Sutskever, Ilya, Oriol Vinyals, and Quoc V. Le. "Sequence to sequence learning with neural networks." *Advances in Neural Information Processing Systems*. 2014.