Analysis of multilayer perceptron based approach in studying brain neuronal connectivity networks

Nikita M. Smirnov
Neuroscience and Cognitive Technology Lab, Center for Technologies in Robotics and Mechatronics Components Innopolis University
Innopolis, Russia

Pavel S. Tishkin
Neuroscience and Cognitive Technology Lab, Center for Technologies in Robotics and Mechatronics Components Innopolis University
Innopolis, Russia

Evgeny A. Lutanin
Neuroscience and Cognitive Technology Lab, Center for Technologies in Robotics and Mechatronics Components Innopolis University
Innopolis, Russia

***Abstract*—In this paper, we present an analysis of modern tools for processing brain neuronal connectivity graphs — feed-forward artificial neural network. Different configurations of multilayer perceptron were considered and tested against an inference of functional dependence in model systems, specifically on Rossler oscillators generated data. Graphs of training, test and dynamics were considered. Furthermore, analysis and comparison of the r-squared scores and executed time were collected to measure the efficiency of the configurations.**

I. Introduction

Considering two coupled oscillatory systems we expect to see a functional dependence with increasing of coupling strength. Its explicit form varies from system to system and can be very complicated [1], [2]. Therefore classical mathematical methods can depict this dependence only partially [3]–[5]. On the other side, feed-forward multilayer perceptron (FF MLP) can infer nonlinear functions serving as a universal approximator.

Other systems under study could be neuronal ensembles interacting with each other. In Ref. [6] Frolov et al proposed a novel method the essence of which is to use FF MLP to detect functional connectivity between epileptic ECoG channels' regions. We applied this method to EEG dataset and enhanced it, reducing time complexity to make this approach more applicable for large amount of data.

II. MLP Configuration and its efficiency

Inheriting configuration from [6] to setup network we used Adam optimizer with a rate 0.005. To avoid possible issues related to overfitting and overestimation, the model was dropped if the difference between loss function on training set and on validation set was more than 0.01 for 10 consecutive epochs. The training process stopped in case the model fails to decrease the value of a loss function for more than 10^{-7} for 15 consecutive epochs.

To test our FF MLP configurations to detect functional dependence in coupled chaotic model systems we considered a pair of coupled Rössler oscillators which is a classical nonlinear model for the study of synchronous behavior.

This work is supported by the Center for Technologies in Robotics and Mechatronics Components (Innopolis University, Russian Federation).

TABLE I
Mean R^2 scores and execution times for different configurations of FF MLP

Configuration	R^2 score	Execution time (s)
10 softmax + 10 softmax	0.98	44.47
10 elu + 10 tanh	0.97	18.6
30 elu + 40 tanh	0.97	15.0

$$\ddot{x}_{1,2} = -\omega_{1,2}y_{1,2} - z_{1,2} + \epsilon_{1,2}(x_{2,1} - x_{1,2}), \quad (1)$$
$$\ddot{y}_{1,2} = \omega_{1,2}x_{1,2} + ay_{1,2}, \quad (2)$$
$$\ddot{z}_{1,2} = p + z_{1,2}(x_{1,2} - c) \quad (3)$$

The control parameters $a = 0.15$, $p = 0.2$, and $c = 10$ have been set for both systems, while $\omega_1 = 0.99$ and $\omega_2 = 0.95$ by the analogy with [6]. For testing we used unidirectional coupling, therefore $\epsilon_1 = 0$ and $\epsilon_2 = \epsilon$.

The model dataset consists of 50×25 points per variable. Data is normalized in range $[0, 1]$, shuffled and separated equally into train and validation sets.

We started off with the configuration given in [6]. MLP consists of 2 hidden layers, each containing 10 softmax units. The number of both inputs and outputs is determined by the embedding dimensions of coupled systems. Output artificial neurons have a linear activation function.

Inherited MLP configuration succeeds in recognizing functional dependence when epsilon exceeds GS (generalized synchronization) threshold of 0.11. The connection is established in case of R^2 score more than 0.5 and it is equal to 0.98. In case ϵ_1 is equal 0.03 it is not possible to find a connection. Hence, R^2 is equal to 0.3. This advocates that the model is performing adequately on the given dataset.

Despite the fact that these results are proving this configuration is applicable there is one big issue – a time of execution. It performs for 42 seconds per each equivalent model dataset

978-1-7281-7286-6/20/$31.00 ©2020 IEEE

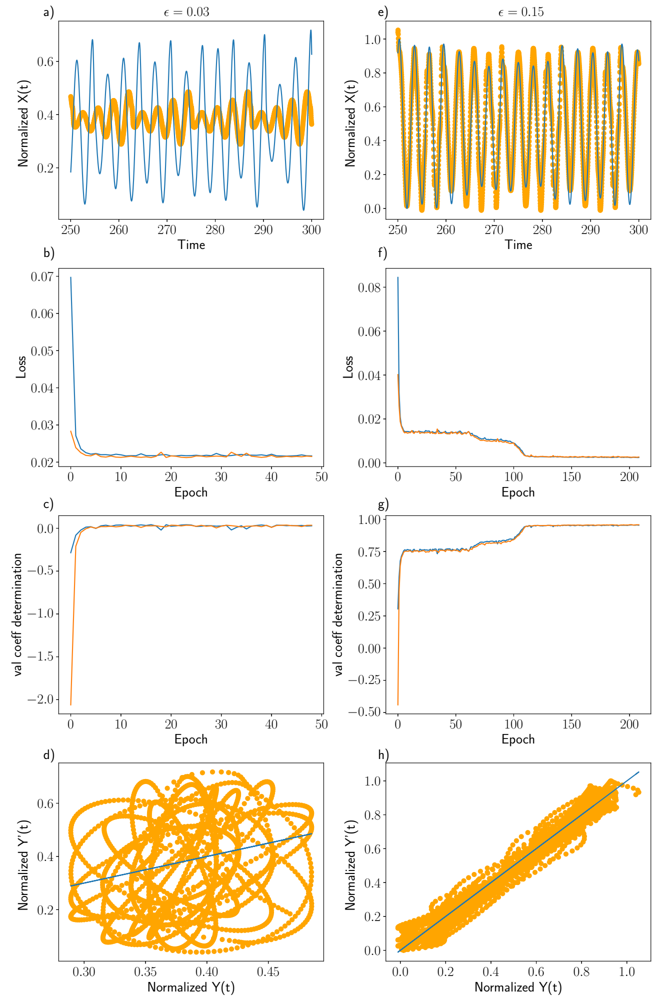

Fig. 1. Inference of functional dependence using MLP with 30 elu neurons and 40 tanh neurons (a) (e) Test sample of response Rossler oscillator time series (blue curve) and its prediction x'2 via ANN (orange points) (b) (f) History of loss function per epoch on training set (blue curve) and on validation set (orange curve) (c) (g) History of metric per epoch on training set (blue curve) and on validation set (orange curve) (d) (h) Regression analysis of x2 variable prediction by ANN model

(Fig.1). There is no problem with small real datasets (eg. ECoG data with 6 channels), but it is not applicable to EEG or MEG datasets. If the performance of the calculations is the same for EEG channels, there will be approximately 330 days for data processing (10 windows per channel, 10 channels, 15 trials, 10 test subjects, 1 hand, 1 group for 42 second per window). To decrease execution time we implemented several different configurations and had chosen the most efficient one. We will use the same argumentation for further observations of different models. Next model also has 2 hidden layers, elu serves as an activation of the first hidden layer and tanh for the second hidden layer.

This model also succeeds in recognizing functional connectivity. It has no overfitting and R-squared score is 0.971 (Fig.1). In case of $\epsilon_1 = 0.03$ R^2 score is equal 0.3 which denotes no connectivity. As such, this model can also be considered for evaluation of efficiency. To compute the task the model spent 20.5 s, which is twice less than the previous model.

A slight modification of the previous model is to make 30 elu neurons and 40 tanh neurons (Fig.2). This model has no overfitting, R^2 is equal 0.98 and 0.3 for $\epsilon_2 = 0.15$ and $\epsilon_2 = 0.03$ respectively. The task was computed in 10.2 s.

According to the time spent by models in computing 11 different intervals of time from oscillators, it can be concluded that the configurations with 10 elu neurons + 10 tanh and 30 elu + 40 tanh infer equivalent results and their average execution times are 2.3-2.5 times less than softmax configuration execution time. It is worth mentioning that 30 elu + 40 tanh has sufficient problems with its design. On some intervals it stops at the local minimum of loss function that is not optimal. In the above statistics, patience of learning was changed from 15 to 45. It is visible from the educational graph. Due to changes to the model, some intervals tend to execute for 20-26 seconds. It is not possible to identify the amount of those intervals before computation, but it performs faster than previous configuration on the average.

III. Application to EEG dataset

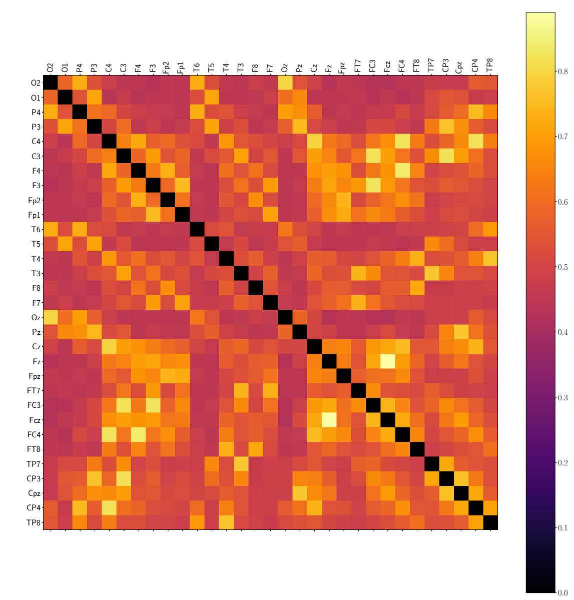

Fig. 2. Matrix with R^2 score for each pair of channels averaged over all volunteers and trials.

Tested EEG dataset was taken from “Frolov N S et al, EEG dataset for the analysis of age-related changes in motor-related cortical activity during a series of fine motor tasks performance”[1]. EEG data were recorded for 10 volunteers with 15 trials of 12 seconds each with a sampling frequency rate 250 Hz. To infer functional dependence between one EEG

[1]https://figshare.com/articles/EEG-dataset-for-the-analysis-of-age-related-changes-in-motor-related-cortical-activity-during-a-series-of-fine-motor-tasks-performance//12301181/1

channel and another we try to predict the brain state in the first channel area based on one in the second channel area using the chosen MLP configuration.

EEG dataset was filtered by the 5th-order Butterworth filter with cut-off points at 4 Hz and 8 Hz. We estimated the parameters of the embedding space of the experimental signals using mutual information approach [7] to determine the delay time τ and the false nearest neighbor method [8] to determine the embedding dimension D for each channel trial. Thus, the architecture of MLP is such that it contains $\hat{D}$ inputs and $\hat{D}$ outputs with $\hat{D}$ equal maximum embedding dimension for one trial between two channels. The same is for time delay – for each trial, we computed maximum time delay $\hat{\tau}$ between two channels. With these parameters, we reconstructed phase space of each trial between two channels, separated trials into 10 bins (discarding the baseline). Each bin was shuffled and separated into train and validation sets and put into train. Thus we computed connectivity matrix for each bin, trial, pair of channels, and volunteer.

The matrix with R^2 score for each pair of channels averaged over all volunteers and trials is shown in Fig. 3.

IV. Conclusion

The presented results contributed in such the multidisciplinary fields of science as mathematics, physics and neuroscience. Based on previous paper [6] we studied the novel method for processing brain neuronal connectivity with the help of feed-forward artificial neural network. We analyzed the different configurations of multilayer perceptron and tested against an inference of functional dependence in model coupled Rossler oscillators. Furthermore, we applied the method to real EEG data set corresponding motor execution of human subjects.

References

[1] N. F. Rulkov, M. M. Sushchik, L. S. Tsimring, and H. D. Abarbanel, "Generalized synchronization of chaos in directionally coupled chaotic systems," *Physical Review E* **51**(2), p. 980, 1995.

[2] O. I. Moskalenko, A. A. Koronovskii, A. E. Hramov, and S. Boccaletti, "Generalized synchronization in mutually coupled oscillators and complex networks," *Physical Review E* **86**(3), p. 036216, 2012.

[3] H. D. Abarbanel, N. F. Rulkov, and M. M. Sushchik, "Generalized synchronization of chaos: The auxiliary system approach," *Physical Review E* **53**(5), p. 4528, 1996.

[4] B. S. Dmitriev, A. E. Hramov, A. A. Koronovskii, A. V. Starodubov, D. I. Trubetskov, and Y. D. Zharkov, "First experimental observation of generalized synchronization phenomena in microwave oscillators," *Physical Review Letters* **102**(7), p. 074101, 2009.

[5] A. A. Koronovskii, O. I. Moskalenko, and A. E. Hramov, "Nearest neighbors, phase tubes, and generalized synchronization," *Physical Review E* **84**(3), p. 037201, 2011.

[6] N. Frolov, V. Maksimenko, A. Lüttjohann, A. Koronovskii, and A. l. Hramov, "Feed-forward artificial neural network provides data-driven inference of functional connectivity," *Chaos: An Interdisciplinary Journal of Nonlinear Science* **29**(9), p. 091101, 2019.

[7] M. S. Roulston, "Estimating the errors on measured entropy and mutual information," *Physica D* **125**, p. 285, 1999.

[8] M. B. Kennel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Phys. Rev. A* **45**, p. 3403, 1992.