Data Analytics in Neuroscience

In behavioural neuroscience, many studies have examined how animals behave inside different types of constrained environments (arenas) and how they learn to solve specific tasks [1–3] in order to study mechanisms of cognition, spatial learning and memory and how different conditions, medical disorders or treatments affects them [4–6]. The goal is to discover general principles that can be applied to various species including humans [7–9]. The analysis pipeline of these experimental procedures usually involves: (a) recording the animal path inside the arena and collecting trajectory data; (b) use these data to compute various measurements; (c) use these measurements to quantify the different animal behaviours and detect difference among animal groups for example treated and untreated animals. Nevertheless, machine learning methods have proven better on identifying behavioural differences among animal groups [10, 11], but existing applications are limited to specific experiments, they might result to loss of behavioural information and they require meta-parameter tuning thus knowledge of the machine learning field in order to be used [12]. In addition, given the plethora of the various machine learning algorithms and proposed frameworks it is difficult to select which one(s) are more appropriate to analyze data from a specific custom procedure.

Tutorial overview

This tutorial will be split into two parts. Firstly, the RODA software [13] for detailed behavioural analysis inside the Morris Water Maze experimental procedure [10, 12] will be presented and hands-on experience will be provided. Focus will be given on RODA's pipeline of analysis and the underlying machine learning techniques. Secondly, explanation will be given on how the methods of RODA can be applied to other experiments beyond the Morris Water Maze and how its machine learning methods can be adopted to analyze high-dimensional datasets in an unsupervised manner not only for obtaining observations grouping (clusters) but also for gaining information about the importance of the data attributes (feature selection and weighting).

Hands-on exercise 1

Participants will use a demo version of the RODA software to analyze a simplified toy dataset from the recent study of [14] and investigate how differential corticosterone responsiveness affect learning and memory.

Hands-on exercise 2

Participants will use a graphical user interface (GUI) and have the ability to try different K-Means inspired algorithms and algorithmic auto-tuning techniques. Then they will use the Sparse K-Means algorithm to analyze the same dataset overloaded with extra data attributes related to path analysis and examine the different behavioural motifs and their dominant data attributes. The hypothesis is that detailed analysis on the path features can reveal a number of different behaviours, regardless of the animal and the experimental procedure, which might go unnoticed using stereotypical behavioural labels as is the case in many existing analysis frameworks.

For the hands-on exercises: *Please note that RODA and the clustering GUI are written in MATLAB and require MATLAB version R2016a or higher in order to run. For users that do not have MATLAB, installers will be provided for Windows, MAC and possibly Linux. The installers require MATLAB runtime environment which is free but might need some time to be installed. In case you do not have MATLAB we highly recommend the installation of MATLAB Runtime Version: R2019a (9.6) prior to the tutorial.*

References

[1] R. Morris, "Developments of a water-maze procedure for studying spatial learning in the rat," *Journal of neuroscience methods*, vol. 11, no. 1, pp. 47–60, 1984.

[2] C. Andrade, M. Alwarshetty, S. Sudha, and J. S. Chandra, "Effect of innate direction bias on t-maze learning in rats: implications for research," *Journal of neuroscience methods*, vol. 110, no. 1-2, pp. 31–35, 2001.

[3] S. Pol-Bodetto, H. Jeltsch-David, L. Lecourtier, N. Rusnac, C. Mam-Lam-Fook, B. Cosquer, K. Geiger, and J.-C. Cassel, "The double-h maze test, a novel, simple, water-escape memory task: acquisition, recall of recent and remote memory, and effects of systemic muscarinic or nmda receptor blockade during training," *Behavioural brain research*, vol. 218, no. 1, pp. 138–151, 2011.

[4] M. D. Lindner, "Reliability, distribution, and validity of age-related cognitive deficits in the morris water maze," *Neurobiology of learning and memory*, vol. 68, no. 3, pp. 203–220, 1997.

[5] C. Sandi, H. A. Davies, M. I. Cordero, J. J. Rodriguez, V. I. Popov, and M. G. Stewart, "Rapid reversal of stress induced loss of synapses in ca3 of rat hippocampus following water maze training," *European Journal of Neuroscience*, vol. 17, no. 11, pp. 2447–2456, 2003.

[6] J. Ordy, G. Thomas, B. Volpe, W. P. Dunlap, and P. Colombo, "An animal model of human-type memory loss based on aging, lesion, forebrain ischemia, and drug studies with the rat," *Neurobiology of aging*, vol. 9, pp. 667–683, 1988.

[7] C. H. Brennan, "Zebrafish behavioural assays of translational relevance for the study of psychiatric disease," *Reviews in the Neurosciences*, vol. 22, no. 1, pp. 37–48, 2011.

[8] V. Bohbot, R. Jech, E. Ruzicka, L. Nadel, M. Kalina, K. Stepankova, and J. Bures, "Rat spatial memory tasks adapted for humans: characterization in subjects with intact brain and subjects with medial temporal lobe lesions," *Physiol Res*, vol. 51, no. Suppl 1, pp. S49–S56, 2002.

[9] A. M. Daugherty, A. R. Bender, P. Yuan, and N. Raz, “Changes in search path complexity and length during learning of a virtual water maze: Age differences and differential associations with hippocampal subfield volumes,” *Cerebral Cortex*, vol. 26, no. 6, pp. 2391–2401, 2015.

[10] T. V. Gehring, G. Luksys, C. Sandi, and E. Vasilaki, “Detailed classification of swimming paths in the morris water maze: multiple strategies within one trial,” *Scientific reports*, vol. 5, p. 14562, 2015.

[11] T. Illouz, R. Madar, Y. Louzon, K. J. Griffioen, and E. Okun, “Unraveling cognitive traits using the morris water maze unbiased strategy classification (must-c) algorithm,” *Brain, behavior, and immunity*, vol. 52, pp. 132–144, 2016.

[12] A. Vouros, T. V. Gehring, K. Szydlowska, A. Janusz, Z. Tu, M. Croucher, K. Lukasiuk, W. Konopka, C. Sandi, and E. Vasilaki, “A generalised framework for detailed classification of swimming paths inside the morris water maze,” *Scientific reports*, vol. 8, no. 1, p. 15089, 2018.

[13] A. Vouros, T. V. Gehring, M. Croucher, and E. Vasilaki, *RodentDataAnalytics/mwm-ml-gen: Version 4.0.2-beta*. Zenodo, Jul 2017, https://github.com/RodentDataAnalytics/mwm-ml-gen/releases/tag/v4.0.2.

[14] D. Huzard, A. Vouros, S. Monari, S. Astori, E. Vasilaki, and C. Sandi, “Constitutive differences in glucocorticoid responsiveness are related to divergent spatial information processing abilities,” *bioRxiv*, p. 579508, 2019.