Model selection methods for estimating learning behavior in cuttlefish and octopuses

Louis Köhler | Humboldt-Universität zu Berlin

Learning, whether in animals or humans, is the process by which behaviors become better adapted to the environment (Rescorla, 1988). This process is highly individualized and is often only observable through the actions of the learner. By leveraging individual behavioral data, we can identify models that best explain this learning process (Aubert et al., 2023). To achieve this, we propose two model selection methods—a general hold-out procedure and an AIC-type criterion (Aubert et al., submitted)—both tailored for non-stationary, dependent data (Aubert et al., 2024). Theoretical error bounds for these methods are derived, showing performance close to that in the standard i.i.d. case. To demonstrate their effectiveness, we apply these methods to contextual bandit models (Auer et al., 2002; Lattimore and Szepesvári, 2020), which approximate the learning behavior of agents interacting with their environment. These contextual bandit algorithms capture how agents balance exploration and exploitation when seeking to maximize their rewards. By observing the learning process, we can estimate how an agent integrates contextual information to make decisions.

Extending these approaches, we explore their application to experimental learning data from cephalopods, specifically octopuses and cuttlefish (Jozet-Alves et al., 2013; Poncet et al., 2022). These animals are renowned for their remarkable problem-solving abilities and contextual learning in dynamic environments. By designing tasks that simulate adversarial conditions, we collect data on how these cephalopods adapt their behavior over time to achieve specific goals, such as obtaining food rewards (Jozet-Alves et al., 2013). Using our proposed model selection framework, we aim to identify the models that best capture their learning dynamics. This application not only validates our methods on non-human experimental data but also provides new insights into the cognitive processes underlying cephalopod learning. These methods and results are part of a joint collaboration with Julien Aubert (Univ. Côte d'Azur, France), Christelle Jozet-Alves (Univ. Caen Normandie, CNRS, France), Louis Köhler (Univ. Côte d'Azur, France), Luc Lehéricy (Univ. Côte d'Azur, CNRS, France) and Patricia Reynaud-Bouret (Univ. Côte d'Azur, CNRS, France).

Julien Aubert, Luc Lehéricy, and Patricia Reynaud-Bouret. On the convergence of the MLE as an estimator of the learning rate in the exp3 algorithm. In Proceedings of the 40th International Conference on Machine Learning, volume 202 of Proceedings of Machine Learning Research, pages 1244–1275. PMLR, 7 2023.

Julien Aubert, Luc Lehéricy, and Patricia Reynaud-Bouret. General oracle inequalities for a penalized log-likelihood criterion based on non-stationary data. Working paper or preprint, 5 2024. URL https://hal.science/hal-04578260.

Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. The nonstochastic multiarmed bandit problem. SIAM journal on computing, 32(1): 48–77, 2002.

Christelle Jozet-Alves, Marion Bertin and Nicola Clayton. Evidence of episodic-like memory in cuttlefish. Current biology : CB. 23, 2013 R1033-5. 10.1016/j.cub.2013.10.021.

T. Lattimore and C. Szepesv´ari. Bandit Algorithms. Cambridge University Press, 2020.

Lisa Poncet, Coraline Desnous, Cécile Bellanger and Christell Jozet-Alves. Unruly octopuses are the rule: Octopus vulgaris use multiple and individually variable strategies in an episodic-like memory task. Journal of Experimental Biology. 225, 2022. 10.1242/jeb.244234.

Robert A Rescorla. Behavioral studies of pavlovian conditioning. Annual review of neuroscience, 1988.
