TEATIME Machine Learning meeting (5th-6th September 2023) - Report

Motivation

The home-cage monitoring field is growing rapidly. We have seen a proliferation of tools and machine learning applications which allow users to track animal movement and analyse behaviour. As the field matures, we must consolidate understanding of the remaining challenges and opportunities.

Meeting report

COST TEATIME brought together 26 home-cage and machine learning experts in Jurmala, Latvia for two days to discuss the state of the field and to set future directions.

Presentations

On the first day, we organised a series of talks on the cutting-edge of home-cage machine learning applications from academic and industry delegates. We organised three talks on novel machine learning applications. Claire Witham discussed her work to automate monitoring and behavioural segmentation for macaques. Afterwards, Hannah Lapp and Kyle Severson discussed their platforms which track and assess maternal and rodent behaviour, and perform 3D post estimation, respectively.

Our remaining speakers presented integrating machine learning and hardware systems. Konrad Danielewski discussed EcoHAB, which can track mice as they move between four cages. Vivek Kumar gave us an insight into the Jackson Laboratory's automated cage monitoring hardware. Both systems featured bespoke machine learning pipelines to capture and analyse high-quality behavioural data. We also welcomed two speakers from industry: Mara Rigamonti presented Technoplast's recent efforts to track fighting in cage to monitor welfare, and Bastijn Koopmans described work at Sylics to systematically phenotype multiple lines using their PhenoTyper platform and integrated dimensionality reduction. These examples indicate that the richness of home-cage data requires big data solutions from the outset, and at this scale data management requires a considered approach to scale effectively.

Discussions

On the second day, breakout groups focused on key challenges and opportunities in the field. Some identified challenges were technical, such as trade-offs between granularity and data volume. We discussed benchmarks are available for data quality and data models, and shared thoughts on design and accessibility of home-cage systems. We may solve some of these challenges through immediate collaborations initiated at this meeting. We also considered the long term and the direction our community should take as it matures, which could require cultural shifts and technological infrastructure, some of this is discussed below. A lively roundtable concluded the session where defined some short-term goals for TEATIME and the wider community.

Short and medium-term goals

In the **short-term** we aim to carry momentum forwards by attending meetings (Measuring Behaviour, Society for Neuroscience) and facilitating discussion on the Behaviour Forum. We plan to meet

Hamish Forrest, WG3

regularly online to discuss long term outlook and collaboration on machine learning challenges as discussed in Latvia.

In the **medium-term**, we will focus on consolidation and coordination of information. We will communicate our position in upcoming publication from TEATIME members. TEATIME leadership can coordinate further action and align with aims of the TEATIME Roadmap, whilst TEATIME can serve as a common ground for members of the community. We also hope to see the results of some collaborations in the medium term. Two examples initiated during our meeting are tracking pose from side-on cameras and building a robust data model to track dams and pups. Facilitating collaborations and presenting data (rather than hardware) may attract scientists who see an opportunity to answer their questions in novel home-cage systems.

Long-term outlook

Standard tools for machine learning, hardware and data management

Standard tools create opportunities to integrate with established systems (such as neuroimaging). They may also make home-cage systems more attractive to pharmaceutical companies hoping to add passive monitoring to their studies. In neuroscience, we have seen bespoke tools rise from the community to become standards in the field, and as we discussed, there is no shortage of machine learning applications in the home-cage field that could become accepted standards. However, often these are built for niche applications, rather than responding to the broad user community. Hardware systems that pair animal housing and robust data acquisition in one platform present an opportunity for industry partners (such as Syclics, or JAX). They could suit both experimental scientists (especially if open-source or modular) and technicians monitoring animal welfare. If collecting data at scale and high-throughput, standards of data quality and reproducibility will be key to uptake.

Coordinated development

As the complexity of home-cage data increases, machine learning, hardware and data management will need to develop in step. Preliminary work has been done by domain experts (as discussed at the meeting), and TEATIME's role should be to facilitate this work going forwards. A case where coordination in needed is when defining ontologies, as ideally this is done before large-scale data collection. Being that the field already has a huge variety of data types and systems, work to define ontologies will need to accommodate past and future data and promote decisions.

Funding and resourcing

TEATIME has brought together potential collaborators and represents a field with huge potential to behavioural research. Home-cage systems can require technical infrastructure and resources not usually considered in current funding applications where data standards and collection can be afterthoughts to the scientific question. There are some options, such as shared funding with companies, and building central repositories within institutions. TEATIME could support these efforts by highlighting the opportunities of home-cage systems to funding bodies (such as Welcome Trust), regulatory bodies (such as N3CRs and the FDA).

Hamish Forrest, WG3