The Effect of Co-adaptive Learning & Feedback in Interactive Machine Learning

Michael Zbyszyński
Goldsmiths, University of London
London, UK
m.zbyszynski@gold.ac.uk

Balandino Di Donato
Goldsmiths, University of London
London, UK
b.didonato@gold.ac.uk

Atau Tanaka
Goldsmiths, University of London
London, UK
a.tanaka@gold.ac.uk

ABSTRACT

In this paper, we consider the effect of co-adaptive learning on the training and evaluation of real-time, interactive machine learning systems, referring to specific examples in our work on action-perception loops, feedback for virtual tasks, and training of regression and temporal models. Through these studies we have encountered challenges when designing and assessing expressive, multimodal interactive systems. We discuss those challenges to machine learning and human-computer interaction, proposing future directions and research.

Glasgow '19, May 04, 2019, Glasgow, UK

KEYWORDS

human-centred machine learning, co-adaptation

INTRODUCTION

Human interaction with technology can be described by a process of co-adaptation [8], where human users adapt to technological tools while simultaneously shaping those tools to better fit their own needs. Co-adaptation is especially evident in HCI systems which use interactive machine learning (IML), where users are cyclically recording training examples, training new models, and evaluating model performance. Evaluation of model performance is enabled by feedback, either directly from the main output of a model or from secondary audio or visual feedback related to the model’s performance. Human users decide if a model is adequately trained, or if training data need to be adjusted before another model is trained. In this position paper, we present work where we have observed significant complexity in the human-machine interaction loop, involving co-adaptive learning from both ML models and humans engaged with these models.

Our work focuses on creating real-time, multimodal interactive systems controlled by biosignals – specifically electromyography (EMG) – along with other sensors (e.g. IMUs). EMG sensors measure the electrical activity of skeletal muscles, and together with IMUs can be used to analyse body movements and gestures. These real-time systems are a special case for IML. In contrast to the IML image classification problem discussed in Fails and Olsen[4], in our use cases both the training data and the inputs to trained models are generated on-the-fly by human performers. Furthermore, EMG offers particular challenges to HCI design. [11] Individuals have different anatomies and employ their muscles differently, so a “one-size-fits-all” interaction mapping of EMG data is very limited.

We have experimented with various methods of feedback – audio or visual, related or unrelated to the primary interface goals – and have observed that feedback can inform users about their performance as well as the performance of the particular IML models they are interacting with. Providing users with this information has the potential to tighten the interaction loop and improve the perceived quality of IML-centred interfaces. However, introducing feedback poses challenges for studying the system: Are we evaluating the model’s performance, or are we studying the user’s ability to work the model? In a tightly bound co-adaptation loop where human learning and machine learning are coupled, how can we design experiments that can effectively distinguish which effect is in force? These challenges inform our position and suggest design perspectives for such systems.

RELATED WORK

Fails and Olsen [4] define interactive machine learning to be machine learning with a human in the learning loop, observing the result of learning and providing input meant to improve the learning

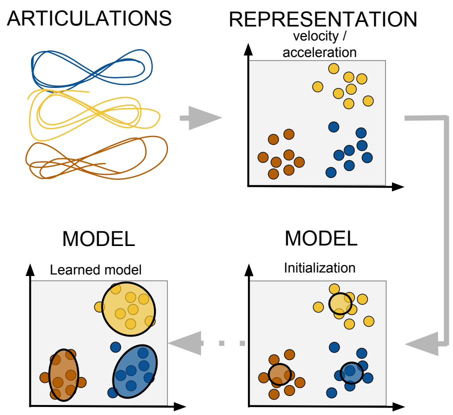

Figure 1: Learning procedure. Gestural data, velocity, and acceleration feature space is associated with an articulation label. The representation feeds a GMM initialised with the means of each class and adapted using Expectation-Maximisation.

outcome. This human engagement creates an opportunity for co-adaptation. An example of co-adaptation is Kulesza, Amershi, et al. [7] description of concept evolution: a dynamic where a user's goals change while interacting with a system. Even though a trained model might not perform as imagined, it is possible that the imperfections suggest another way to interact with a problem space. In certain real-world applications, there might not be a ground truth that would inform the accuracy of trained models, or it might change over a period of interaction. *Expressive* models are not necessarily *accurate* models.

Cartwright and Prado [3] also address concept evolution, in the context of musical tasks. They further identify the problem that editing training data sets become increasingly more tedious as the size of the data set grows, and recommend minimising the number of examples that need evaluating. Fiebrink [5] noted co-adaptation while studying the evaluation practices of end users building IML systems for real-world gesture analysis problems. She observed that users employed evaluation techniques to judge an algorithm's relative performance and improve upon trained models, as well as learning to adapt their performance to provide more effective training data. Subjects' strategies for providing training data evolved over the training sessions.

Feedback in the context of machine learning has been examined by Françoise et al. [6] who propose interactive visual feedback that exposes the behaviour and internal values of models, rather than just their results. They consider whether visualisations can improve users' understanding of machine learning and provide valuable insights into embodied interaction. Similarly, Ravet et al. [9] identify difficulties for users interacting with high-dimensional motion data and propose solutions for using ML with these data, including visual representation of the impact of learning algorithm parameter tuning on modelling performances.

EXAMPLES

Action-perception

In a study exploring feedback in an action-perception loop [10], we used Gaussian Mixture Models (GMM) trained with different orchestral conducting gestures (1-*legato*, 2-*normal*, 3-*staccato*). Participants were asked to make a simplified conducting gesture, following the beat while a melody was being played for each different articulation. The participant could rehearse until she felt confident and then record the training examples. (Figure 1). After training, the participant was presented with one of the melody versions used for training; the articulation of that version being the target articulation. The user was also provided visual feedback of the output of the model. During performance, a slider showed the fixed, target articulation value together with the current inferred one.

The study was designed with the objective of characterising the quality of trained models by evaluating accuracy during performance sessions. However, the results of that evaluation revealed

Figure 2: Virtual Task. Participant in a rest position (above) and when performing the task (below).

unexpected complexity. While algorithms were able to model participants' intended articulations, participants also adapted their performance to the system. Adaptation was evident because the accuracy of models as calculated through cross-validation against recorded examples is lower than the average accuracy measured during new performances with the models. This suggests that in a continuous action-perception loop, users responded to visual feedback by adapting their physical performance to cause a model to output the correct articulation value for given task.

Virtual tasks

We carried out a task-based study to define a simple grasping task using muscle tension. There was no machine learning in this study, but we employed auditory feedback to help subjects learn to perform muscle actions in a more consistent manner. More consistency could lead to more useful training data for IML applications. Subjects were asked to imagine holding a cup of water with just enough tension so it would not slip through their fingers. Too little tension and the cup slips, too much and it crushes or breaks. (Figure 2). The subtlety of expressive grasping could be compelling in a virtual reality scenario.

In our study, auditory feedback was provided as a secondary communication channel and compensated for the lack of haptic feedback in virtual space. The main tasks were defined using the metaphor of a glass and communicated using video of a researcher performing the same task. Auditory feedback enabled us to focus users on the specifics of their behaviours so that they would understand what they were doing and how it could lead to a result.

Workshopping regression and temporal modelling

In a more recent workshop activity, we asked users to play an imitation game to train an IML system with real-time human input. This was a *sound tracing* [2] activity; we asked participants to physically represent a sound through hand and arm gestures. These gestures were used to train different models, allowing users to compare a series of regression-based approaches with a temporal modelling algorithm. The temporal modelling implemented Hidden Markov Models to model a sequence of time-based input from beginning to end. Three different regression models looked at 1) the whole input as a single set of training examples 2) four static examples using salient anchor points from the stimulus as examples, or 3) an automated windowing system capturing short periods of dynamic input centred around the same anchor points.

The stimulus imitated in the training phase became the auditory feedback in the testing phase, with trained models controlling the parameters of the synthesizer that generated the initial stimulus. Participants in our workshop were able to try the different algorithms using a consistent IML workflow without knowing the technical details of any particular algorithm (Figure 3). They commented on the affordances of different techniques – some facilitating the reproduction of the original stimulus,

Figure 3: Workshop participants using regression and temporal modelling.

others enabling exploration – and critiqued the fluidity of response of the models. They discussed the choice of algorithms as a trade-off between faithfully reproducing the stimulus and creating a space for exploration to produce new, unexpected ways to articulate sounds.

DISCUSSION & CONCLUSION

When asked to accomplish a specific task (e.g. crumple a piece of paper, or hold a virtual glass) users are not typically aware of which muscles cause that task to be performed. There are many ways that forearm muscles can be employed to create the same apparent hand motion; users do not intentionally choose one method over another. This lack of awareness complicates the generation of training data for IML, as well as evaluation of a trained model. Users are not always aware of their exact performance, or what elements of that performance are influencing the output. Feedback can be an important tool to help users understand both their own performance and that of a model, leading to better outcomes.

By design, such feedback causes users to adapt their performance over the course of an IML session. But, co-adaptive learning complicates our ability to objectively evaluate trained models. Models respond better in interactive performance than when evaluated with recorded examples because subjects “play” them. This discrepancy suggests that in a continuous action-perception loop, users respond to feedback by adapting their physical performance to cause a model to perform properly for a given task.

Designing an adaptive system is challenging because the final stage of interaction design is placed in the hands of the users. [1] Systems should not require sophisticated understanding of machine learning from potential users. Rather, they must contextualise an evolving interaction in an exploratory space that allows a user to delve deliberately and meaningfully manipulate the affordances of the system.

Through our work, we have developed the position that interactive machine learning is an invaluable paradigm for implementing bespoke user interactions, but it needs to be contextualised in a layer of design that covers the whole UX from conceptualising and learning input actions to shaping and refining rich media outputs. Our position is relevant to IML for real-time interaction situations, such as gaming, virtual reality, or creative performance, where the user is generating new training or input data constantly and can adapt those data to the output of the system.

In response to the examples presented here, we have the following perspective on the development of IML-based interactions:

- Feedback is important for helping users evaluate and use an interactive system.
- Feedback does not need to be part of the main output of a system; it can be a secondary channel.
- Feedback leads to co-adaptation.

- Interaction design can accommodate concept evolution.
- Both humans and machines learn, together.

Real-time IML is especially appealing when it helps users to develop an expressive interaction without leaving the problem space of that interaction. The systems we design should help them consider that space, rather than distract with details of the underlying technologies. That consideration involves co-adaptive learning through evolving user goals and iteration of machine learning models. As designers, we ask: How might we present a real-time, IML workflow to users? How might we enable learning the possibilities of a given system? How might we design feedback to demonstrate the performance and potential of the system and illuminate details of the human performance?

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No 789825).

REFERENCES

[1] F. Bernardo, M. Zbyszynski, R. Fiebrink, and M. Grierson. 2017. Interactive machine learning for end-user innovation. In *2017 AAAI Spring Symposium Series*.

[2] B. Caramiaux, P. Susini, T. Bianco, et al. 2011. Gestural Embodiment of Environmental Sounds : an Experimental Study. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME'11)*. Oslo, Norway, 144–148.

[3] M. Cartwright and B. Pardo. 2016. The Moving Target in Creative Interactive Machine Learning. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. San Jose, California, USA.

[4] J. A. Fails and D. R Olsen Jr. 2003. Interactive machine learning. In *Proceedings of the 8th international conference on Intelligent user interfaces*. 39–45.

[5] R. Fiebrink, P. R. Cook, and D. Trueman. 2011. Human Model Evaluation in Interactive Supervised Learning. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '11)*. Vancouver, BC, Canada, 147–156.

[6] J. Françoise, F. Bevilacqua, and Thecla S. 2016. Supporting User Interaction with Machine Learning through Interactive Visualizations. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. San Jose, California, USA.

[7] T. Kulesza, S. Amershi, R. Caruana, D.yel Fisher, and D. Charles. 2014. Structured labeling for facilitating concept evolution in machine learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3075–3084.

[8] W. E Mackay. 1990. *Users and customizable software: A co-adaptive phenomenon*. Ph.D. Dissertation. Citeseer.

[9] T. Ravet, N. d'Alessandro, J. Tilmanne, and S. Laraba. 2016. Motion Data and Machine Learning: Prototyping and Evaluation. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. San Jose, California, USA.

[10] A. Sarasua, B. Caramiaux, and A. Tanaka. 2016. Machine learning of personal gesture variation in music conducting. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3428–3432.

[11] A. Tanaka and M. Ortiz. 2017. Gestural Musical Performance with Physiological Sensors, Focusing on the Electromyogram. In *The Routledge Companion to Embodied Music Interaction*, Micheline Lesaffre, Pieter-Jan Maes, and Marc Leman (Eds.). Routledge, 422–430.