

Human-Centered Machine Learning

Marco Gillies, Rebecca Fiebrink, Atau Tanaka, Baptiste Caramiaux, Jérémie Garcia, Frédéric Bevilacqua, Alexis Heloir, Fabrizio Nunnari, Wendy Mackay, Saleema Amershi, et al.

► **To cite this version:**

Marco Gillies, Rebecca Fiebrink, Atau Tanaka, Baptiste Caramiaux, Jérémie Garcia, et al.. Human-Centered Machine Learning. Conference on Human Factors in Computing Systems (CHI 2016), May 2016, San Jose, California, United States. pp.3558-3565, 2016, Proceedings of the 2016 CHI Conference - Extended Abstracts on Human Factors in Computing Systems. <<http://chi2016.acm.org>>. <10.1145/2851581.2856492>. <hal-01437057>

HAL Id: hal-01437057

<https://hal.inria.fr/hal-01437057>

Submitted on 17 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Human-Centered Machine Learning

**Marco Gillies,
Rebecca Fiebrink,
Atau Tanaka,
Baptiste Caramiaux,
Jérémie Garcia**
Goldsmiths, London, UK
{m.gillies, r.fiebrink,
a.tanaka, b.caramiaux,
j.garcia}@gold.ac.uk

Frédéric Bevilacqua
IRCAM, Paris, France
Bevilacqua@ircam.fr

**Alexis Heloir,
Fabrizio Nunnari**
DFKI, Saarbrücken,
Germany
{alexis.heloir,
fabrizio.nunnari}@dfki.de

Wendy Mackay
Inria; LRI, Université Paris-Sud,
Université Paris-Saclay Orsay, France
mackay@lri.fr

**Saleema Amershi,
Bongshin Lee** Microsoft
Research One Microsoft
Way, Redmond, WA 98052, USA
{samershi,bongshin}@microsoft.com

**Nicolas d'Alessandro,
Joëlle Tilmanne**
University of Mons, Belgium
{nicolas.dalessandro,
joelle.tilmanne}@umons.ac.be

Todd Kulesza
Oregon State University, USA
todd.kulesza@gmail.com

Human-Centred Machine Learning Workshop at CHI 2016
Conference on Human Factors in Computing Systems
San Jose, CA, USA
May 7-12 2016

Copyright held by the authors

Abstract

Machine learning is one of the most important and successful techniques in contemporary computer science. It involves the statistical inference of models (such as classifiers) from data. It is often conceived in a very impersonal way, with algorithms working autonomously on passively collected data. However, this viewpoint hides considerable human work of tuning the algorithms, gathering the data, and even deciding what should be modeled in the first place. Examining machine learning from a human-centered perspective includes explicitly recognising this human work, as well as reframing machine learning workflows based on situated human working practices, and exploring the co-adaptation of humans and systems. A human-centered understanding of machine learning in human context can lead not only to more usable machine learning tools, but to new ways of framing learning computationally. This workshop will bring together researchers to discuss these issues and suggest future research questions aimed at creating a human-centered approach to machine learning.

Author Keywords

Machine learning; user centered design; data.

ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous; I.2.6 [Artificial Intelligence]: Learning.

The Importance of Machine Learning

Statistical machine learning is one of the most successful areas of computer science research in recent decades. It has driven advances in domains from medical and scientific research to the arts. It provides people the ability to create new systems based on example data, for instance creating a face recognition system from a large dataset of face images, rather than by reasoning about what features make something a face and translating that reasoning into program code. This makes it possible to provide excellent performance on tasks for which it would be very difficult, if not impossible, to describe computational procedures explicitly in code.

In HCI, machine learning can facilitate the creation of user interfaces that rely on human behaviours that are difficult to reason about explicitly. For example, machine learning has been used to create virtual characters who use appropriate body language in interactions with a human [7]; the choice of body language is an example of a tacit, embodied skill which is difficult to describe in a set of explicit rules, but easy for many people to demonstrate. Likewise, machine learning can make it possible to create interfaces that recognize categories of human actions sensed with high-dimensional or noisy sensing mechanisms, or to create expressive gestural controllers with a particular “feel” [6].

Machine learning is important to HCI beyond enabling the implementation of new interfaces. For example, it can potentially democratize the design of interaction. Machine learning allows people to “program” a computer to perform a task by providing examples of how to perform the task, rather than by writing code describing the computer’s desired behaviour. This opens up the possibility for domain experts—who may not be

programmers, but who have the ability to create high quality example data for a task to create new software.

Motivating a Human-centered Approach

In practice, however, machine learning is still a difficult technology to use, requiring an understanding of complex algorithms and working processes, as well as software tools which may have steep learning curves. Patel et al. [9] studied expert programmers working with machine learning and identified a number of difficulties, including treating methods as a “black box” and difficulty interpreting results. Usability challenges inherent in both existing software tools and the learning algorithms themselves (e.g., algorithms may lack a human-understandable means for communicating how decisions are made) restrict who can use machine learning and how. A human-centered approach to machine learning that rethinks algorithms and interfaces to algorithms in terms of human goals, contexts, and ways of working can make machine learning more useful and usable.

The application of HCI methods to supporting new machine learning practices has been explored in Interactive machine learning, a term first used by Fails & Olsen [5] to describe an approach in which humans iteratively add training examples in a freeform manner to improve a model until its quality is acceptable. The term has since come to also encompass related techniques in which human users are engaged in a tight interaction loop of iteratively modifying data and/or features to improve model performance [1]. This is distinct from active learning, where an algorithm chooses from a set of unlabeled examples and queries the user to provide a label. In interactive machine learning, the user chooses what new examples to label and/or create, and so is in control of the process: the computer is

part of a human design process, rather than the human being in the loop of an algorithmic process. Enabling users to interactively edit data and edit features has been shown to improve machine learning results for tasks including image analysis [5], web-page analysis [8], and social network group creation [2].

Work by [6] demonstrates several benefits of employing interactive machine learning as a tool for creating new real-time interactive systems. System builders can demonstrate examples of human actions sensed by input devices (e.g., inertial sensors, game controllers) alongside examples of real-time control over computer processes (e.g., sound synthesis, game engines). By training supervised learning algorithms on this data using a high-level interface, these users (even those without programming or machine learning expertise) can quickly build and refine systems for real-time gestural control. Systems built in this way can lead to a more natural “feel”, as the examples encode users’ embodied practices better than coding input/output mappings [6]. Allowing users to freely edit the training data makes it possible for them to fix system mistakes by providing corrective examples, and to change the behaviour of the system in an exploratory manner (e.g., incrementally add classes, redefine classes as understanding evolves).

Past work also demonstrates ways in which a human-centered perspective leads to different approaches to evaluating, analysing, and understanding machine learning methods [1]. For instance, [6] showed that users building gestural control and analysis systems use a range of evaluation criteria when evaluating trained models, such as decision boundary shape and subjective judgements of misclassification cost. Conventional

model evaluation metrics focusing on generalisation accuracy may not capture such criteria, which means that computationally comparing alternative models (e.g., using cross-validation) may be insufficient to identify a suitable model. Users may therefore instead rely on tight action-feedback loops in which they modify model behaviour through changes to the training data, followed by real-time experimentation with models to evaluate them and inform further modifications. Users may also develop strategies for creating training sets that efficiently guide model behaviour using very few examples (e.g., placing training examples near desired decision boundaries), which results in training sets that may break common theoretical assumptions about data (e.g., that examples are independent and identically distributed) [6]. Summarising related work in a variety of application domains, [1] enumerate several properties of machine learning systems that can be beneficial to users, such as enabling users to critique learner output, to provide information beyond mere example labels, and to receive information about the learner that helped them understand it as more than a “black box.” These criteria are not typically considered when formulating or evaluating learning algorithms in machine learning research.

Further, current machine learning approaches do not fully exploit the nature of applied machine learning as a co-adaptive process, in which a human is changing computer behaviour, but the human also adapts to use a machine learning tool more effectively and adapts his or her goals in response to what is learned using the tool. A person will not necessarily start with a pre-defined concept that must be modeled as accurately as possible (as is often assumed in theory); the concept is likely to evolve during the process of selecting data and training the system. For instance, a user building a

new gesture classifier might change their goals for the number and type of gestures that should be recognised, or for the sensors used to capture the gestures. Different choices will significantly impact the accuracy of the model that can be learned, the number of training examples needed, and the ultimate usefulness of the system. Design research suggests that explicit mechanisms to support exploration, comparison of alternative prototypes, and iterative refinement are fundamentally important to enabling efficient and effective design [4, 3]. Machine learning tools should explicitly aid users in these activities. Tools must also provide effective feedback to inform subsequent user actions: to help users understand how to debug a model that has not learned a concept correctly, to understand the trade-offs between different formulations of a learning problem, and even to understand the limits of what can be learned.

In summary, there are unrealised benefits in work that attempts to further understand users' goals and ways of working, and to develop new algorithms and user-facing tools. We envision new approaches for linking human and machine, which take advantage of the richness of human expertise, and where computational processes exploit all available data to scaffold human design and understanding. At the same time, developing a human-centered approach to research and development in machine learning is challenging, due to the very different nature of theories in machine learning and HCI. The highly mathematical theory used in machine learning research has difficulty representing the messy, qualitative realities dealt with in HCI. There is also no easy way to bridge between HCI's user-centered methods for evaluating the success of new techniques, and the theoretical and quantitative approaches used in machine learning.

Goals

An immediate goal of this workshop will be to identify key research questions in the application of human-centered approaches to machine learning. For instance:

- What is the role of humans in existing machine learning systems (currently hidden)?
- What are the usability challenges of machine learning systems?
- How does a human-centered approach change the way machine learning is done?
- How does a human-centered approach relate to the current theoretical assumptions underlying machine learning?
- What new kinds of machine learning systems should we build based on human-centered research?
- How is human-centered machine learning across different domains such as the arts, science, and social data analysis?
- How can human-centered machine learning support creative work?
- Can human-centered machine learning democratize domains such as big data analytics, opening them up to deeper public engagement?

The organising committee has previously run related workshops on specific techniques (e.g., Interactive Machine Learning workshop held at IUI 2013) and application areas (e.g., the AISB 2014 workshop on Machine Learning, Expressive Movement, Interaction Design, and Creative Applications). For this workshop, we propose "human-centered machine learning" as a title that articulates a core set of values and approaches, cutting across diverse computational techniques and application areas. We therefore aim to provide a forum for discussion and

learning, to nurture and expand the community of researchers engaged in related research. This drives toward our ultimate goal of realising the full potential for helping people work more effectively and efficiently with machine learning in all application domains, for using machine learning as a tool for building new types of interactions, and for making these activities more accessible to people who are not machine learning experts.

Participants

We expect a core of participants already working on interactive machine learning, but we will encourage participation from the broader HCI and ML communities. We will advertise through a workshop website:

www.doc.gold.ac.uk/HumanCenteredMachineLearning/ and distribute the call for papers via mailing lists such as Connectionists and CHI Announcements, personal contacts with the community, and social media.

Acceptance to present at the workshop will be via short (2–4 pages) position papers. We will encourage papers that raise questions and suggest discussion rather than simply presenting work. Papers will be selected based on reviews by at least two members of the organizing committee.

Workshop structure

The workshop will have 4 panel sessions, each on a different theme, lasting roughly 1–1.5 hours each. Each panel will be comprised of 3–4 participants, chaired by one of the organisers. Panel themes could be drawn from the list of questions in the Goals section above, but the final themes will be determined based on the submissions. Each panel participant will give a short (10 minute) presentation of the work that they submitted, but the panels will also include considerable

time for discussion between the panel members and other workshop participants. The workshop will begin with a short introductory talk by the organisers (15 minutes). It will end with a general discussion period (45 minutes) (The last panel will be shorter to ensure time for discussion.)

Expected Outcomes

In the final discussion period, we will decide on mechanisms for disseminating the research questions, research methods, design principles, etc. that were highlighted by the panel discussions. This could take the form of a published position paper or “manifesto” co-authored by the participants. This could in turn be expanded into a longer review article of the area that both surveys existing research and puts forward ideas for the future. Alternatively, an outcome could be several papers by the different participants collected together in a special issue or related format such as a Frontiers research theme.

The final discussion period will also invite ideas about how to maintain a lasting community. This will minimally include a mailing list together with a possible web and/or social media presence, all of which would be open to non-attendees to join. We would like to work with this community to run a second, follow-up workshop: the CHI workshop will most likely have a bias toward HCI researchers, so we would like to have a follow-up workshop in a machine learning venue (e.g., NIPS) in order to engage effectively with and across both communities. The workshop itself will also be archived. All position papers will be archived online. The panels will be videoed and the videos made publicly available online.

Organiser Biographies

Marco Gillies is a senior lecturer at Goldsmiths, University of London. He has done research in applied and interactive machine learning in the fields of Virtual Reality, Computer Animation and Intelligent Virtual Agents. He has organised several research workshops including two workshops at the UK based Artificial Intelligence and Simulation of Behavior conference, and the BT AHRC Research Network: Digital Reconstruction in Archaeology and Contemporary Performance (of which he was director). He has been on the organising committee of Intelligent Virtual Agents (IVA) and New Interfaces for Musical Expression (NIME) and the programme committee of many conferences.

Rebecca Fiebrink is a lecturer at Goldsmiths, University of London. Her research focuses on using machine learning as a tool for designing interactive systems, especially systems for creative expression and embodied interaction. She is the author of the Wekinator software for interactive machine learning. She was the General Co-Chair of the 2014 conference on New Interfaces for Musical Expression.

Atau Tanaka is Professor of Media Computing at Goldsmiths, University of London; formerly professor at Newcastle University, and researcher at Sony Computer Science Laboratory (CSL) Paris. He creates musical instruments using sensing technology to capture movements and gestures of musicians to produce computer generated sound. He has worked at IRCAM, has been artistic ambassador for Apple Computer, and Artistic Co-Director of STEIM in Amsterdam. He is a member of the Embodied Audio Visual Interaction (EAVI) research unit at Goldsmiths.

Baptiste Caramiaux is a Marie Skłodowska Curie Fellow at McGill University (Canada) and IRCAM (France). His research focuses on understanding and modelling the cognitive processes of motor learning in musical performance, and on the design of expressive motion-based interactive systems using machine learning. He conducted academic research at Goldsmiths University of London, and was responsible for machine learning and interaction design in the London-based music tech startup Mogeex Ltd.

Jérémie Garcia is a postdoctoral researcher at Goldsmiths, University of London. His research focuses on user-centered methods to observe, design and evaluate new interactive systems able to support the most creative aspects of music composition such as free expression, interactive exploration and refinement of musical ideas.

Saleema Amershi is a researcher in the Machine Teaching group at Microsoft Research (Machine Teaching is machine learning with a focus on the human user or “teacher”). Her research lies at the intersection of human-computer interaction and machine learning. In particular, her work involves designing and developing tools to support both end-user and practitioner interaction with machine learning systems. Amershi received her Ph.D. in computer science from the University of Washington’s Computer Science & Engineering Department in 2012.

Bongshin Lee is a Senior Researcher at Microsoft Research. Her research interests include Information Visualization, Visual Analytics, Human-Computer Interaction, and User Interfaces & Interaction Techniques. Her research focuses on the design,

development, and evaluation of interactive technologies for people to create visualizations, interact with their data, and visually share data-driven stories, leveraging Natural User Interfaces (NUIs) including pen and touch. She received her Master of Science and Ph.D. in Computer Science from University of Maryland at College Park in 2002 and 2006, respectively.

Frédéric Bevilacqua is the head of the Sound Music Movement Interaction team at IRCAM in Paris. His research concerns the modeling of movement-sound interactions, and the design and development of gesture-based interactive systems. He holds a master degree in physics and a Ph.D. in Biomedical Optics from EPFL in Lausanne. From 1999 to 2003 he was a researcher at the University of California Irvine. In 2003 he joined IRCAM as a researcher on gesture analysis for music and performing arts.

Nicolas d’Alessandro is a postdoctoral researcher at UMONS and head of performative media at the Numediart Institute for Creative Technologies. He holds a PhD in Applied Sciences from UMONS Faculty of Engineering, related to gesturally-controlled synthesis of expressive speech and singing. He is co-founder of Hovertone, a startup for creative experience design.

Joelle Tilmanne is a postdoctoral researcher at UMONS and head of the motion capture and analysis research group at the Numediart Institute. She holds a PhD in Applied Sciences from UMONS Faculty of Engineering, in the field of motion capture data analysis and Hidden Markov Model based motion synthesis. She is co-founder of Hover-tone, a startup for creative experience design.

Alexis Heloir leads the Sign Language Synthesis and Interaction junior research group in Saarbrücken and is an assistant professor at the University of Valenciennes, France. His research interests are interactive control and animation of three-dimensional assets and automated generation of intelligible Sign Language utterances using avatars. He was previously a post-doc at the German Research Center for Artificial Intelligence (DFKI).

Fabrizio Nunnari is a postdoctoral researcher at the German Research Center for Artificial Intelligence (DFKI). He works in the field of digital character animation for the production of Sign Language animation. He also researches the use of Natural User Interfaces for animation authoring.

Wendy Mackay is a Research Director, at Inria, France, where she founded the InSitu research group in Human-Computer Interaction. She served as Vice President for Research at the University of Paris-Sud and as a visiting professor at Stanford University and Aarhus University. She received the ACM/SIGCHI Lifetime Achievement Service Award, is a member of the ACM SIGCHI academy, chaired CHI’13 conference and served as president of ACM SIGCHI.

Todd Kulesza recently completed a Ph.D. in computer science at Oregon State University, working under the guidance of Margaret Burnett. His research interests are in human interactions with intelligent systems, with a focus on enabling end users to personalize such systems efficiently and effectively. He was co-chair of the 2013 IUI workshop on interactive machine learning.

Call for Participation

Machine learning is one of the most important and successful techniques in contemporary computer science, with applications ranging from medical research to the arts, as well as considerable recent interest in its use for interaction design. It is often conceived in a very impersonal way, with algorithms working autonomously on passively collected data. However, machine learning is also a technology to be used by people for human goals. Human-centered machine learning explicitly recognises the role of people in machine learning, as well as reframing workflows based on situated human working practices. An understanding of machine learning in a human context can lead not only to more usable machine learning tools, but to new ways of framing learning computationally. This workshop will bring together researchers from many disciplines to discuss how a human-centered approach can be applied to machine learning.

We invite participants to submit 2–4 page position papers in the ACM Proceedings format to be submitted to humancentredML@doc.gold.ac.uk. Topics may include (but are not limited to):

- the role of humans in current machine learning
- usability challenges of machine learning
- new machine learning methodologies based on human-centered research
- new human-centered machine learning systems
- evaluation methods for human-centered machine learning
- human-centered machine learning in domains such as arts, science and social science

Papers will be reviewed by committee members and accepted authors will present at the workshop. At least one author of each accepted position paper must attend the workshop and must register for both the workshop and for at least one day of the conference. Presentations will be in a panel format to encourage discussion: 3–4 participants will present together as part of a thematic panel. Each panel participant will give a short (10 minute) presentation of their work followed by a joint discussion.

Acknowledgments

This work was partly funded by the ERC grant no. 321135 CREATIV.

References

- [1] Saleema Amershi, Maya Cakmak, W Bradley Knox, and Todd Kulesza. 2014. Power to the people: The role of humans in interactive machine learning. *AI Magazine* 35, 4 (2014), 105–120.
- [2] Saleema Amershi, James Fogarty, and Daniel S Weld. 2012. ReGroup: Interactive Machine Learning for On-Demand Group Creation in Social Networks. *Chi (2012)*, 21–30. DOI: <http://dx.doi.org/10.1145/2207676.2207680>
- [3] Bill Buxton. 2007. Sketching User Experiences: Getting the Design Right and the Right Design. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- [4] Steven P. Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L. Schwartz, and Scott R. Klemmer. 2010. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer Human Interaction* 17, 4 (Dec. 2010), 1–24. DOI: <http://dx.doi.org/10.1145/1879831.1879836>
- [5] Jerry Alan Fails and Dan R Olsen Jr. 2003. Interactive machine learning. In *Proceedings of the 8th international conference on Intelligent user interfaces (IUI '03)*. ACM,

New York, NY, USA, 39–45. DOI:

<http://dx.doi.org/10.1145/604045.604056>

[6] Rebecca Fiebrink. 2011. *Real-time Human Interaction with Supervised Learning Algorithms for Music Composition and Performance*. Ph.D. Dissertation. Princeton University, Princeton, NJ, USA.

[7] Andrea Kleinsmith and Marco Gillies. 2013. Customizing by doing for responsive video game characters. *International Journal of Human-Computer Studies* 71, 7 (2013), 775–784. DOI: <http://dx.doi.org/10.1016/j.ijhcs.2013.03.005>

[8] Todd Kulesza, Saleema Amershi, Rich Caruana, Danyel Fisher, and Denis Charles. 2014. Structured labeling for facilitating concept evolution in machine learning. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14 (2014)*, 3075–3084. DOI: <http://dx.doi.org/10.1145/2556288.2557238>

[9] Kayur Patel, James Fogarty, James A Landay, and Beverly Harrison. 2008. Investigating statistical machine learning as a tool for software development. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. ACM, New York, NY, USA, 667–676. DOI: <http://dx.doi.org/10.1145/1357054.1357160>