Sketching Machine-learning-mediated UX: Getting the Learning Right and the Right Learning

Qian Yang
Carnegie Mellon University Human Computer Interaction Institute
qyang1@cs.cmu.edu

Abstract

Machine learning (ML) applications *that directly interfaces with everyday users* is now increasingly pervasive and powerful. However, UX practitioners seem lagged behind in leveraging this not-so-new technology. ML is not yet a standard part of UX design practice, neither a part of design patterns, prototyping tools, or education. This paper is a reflection on my experience designing ML-mediated UX. I illustrate the role UX practice can play in making machine intelligence in the wild usable and valuable for everyday users: UX identifies 1) what is right machine learning to design with? 2) how to design machine learning right? The conception is a first step to untangling the tight interplay between ML and UX. I highlight the unique challenges such interplay entails and the implications for future research directions.

Introduction

Machine learning (ML) applications *that directly interface with everyday users* are now increasingly powerful, diverse and pervasive. The maturity and prevalence of this technology even catalyzed the notion of *ML is the new UX*, that is ML will be the most important way to improve user experience (UX), and ultimately would *be* the new user interface (Brownlee 2015; Kuang 2013; Yang, Zimmerman, Steinfeld, and Tomasic 2016).

However, UX practitioners seem lagged behind in leveraging this not-so-new technology. ML has not yet become a standard part of UX design practice, neither a part of design patterns, prototyping tools, or design education. UX designers do not have enough tacit understanding of ML to operationalize interaction flows that proactively adapt or evolve over time (Yang, Zimmerman, Steinfeld, and Tomasic 2016).

How should UX design practice deal with ML? How does UX incorporate and negotiate the uncertainty, temporality, and under-interpretation of ML algorithms? How to tailor machine intelligence for user needs and adequately weave them into real-world contexts?

These fundamental UX questions underlie my research and practice in HCI. Over the past 3 years as a researcher at Carnegie Mellon University, I concentrate on designing intelligent systems for real world uses. In particular, my research involves 1) designing a ML-driven software that helps clinicians with patient selection for artificial heart implant surgeries (Yang, Zimmerman, Steinfeld, Carey, et al. 2016; Yang, Zimmerman, and Steinfeld 2015); 2) adding intelligent UI adaptations to a mobile app for users with disabilities (Yang, Zimmerman, Steinfeld, and Tomasic 2016), and 3) designing autonomous driving UX. These systems interact with end users directly in high-consequence real-world scenarios, where user perception and experience of ML are vital to their success.

In this position paper, I reflect on my experience designing these systems, then map the roles UX should play in making machine intelligence usable and valuable *in the wild* for *everyday users*. I highlight the unique challenges designing ML-mediated UX entails, as well as implications for future research pathways.

Case Study: Designing ML-driven Clinical Decision Support

One of my ongoing research is designing an ML system that aids high-risk clinical decision-making – specifically, a software that helps medical professionals deciding whether and when to implant an artificial heart into an end-stage heart failure patient (Yang, Zimmerman, Steinfeld, Carey, et al. 2016; Yang et al. 2015). Previously, such systems commonly take a context-less, prototypic form: Taking in a list of patient condition measures and producing an individualized prediction of patient trajectories, such as likely survival and other post-surgical risks (Bellazzi and Zupan 2008). Interestingly, almost all these systems have failed when moving into clinical practice, despite their effectiveness in labs (Yang et al. 2015).

To better situate the system into its contexts and for its users, we chose to conduct a field study. We interviewed and observed clinicians caring for VAD patients at three different implant centers for 13 days. Our field study identified many barriers that could negatively impact the use and perceived value of machine intelligence. For example, attitudinal barriers. Physicians perceived no need for data support because they felt that they know how to effectively factor patient conditions into clinical decisions. They also lack trust in the ability of machine intelligence; instead, they rely on professional networks for actionable suggestions.

These observations in the field forced us to confront two fundamental questions (I borrow the vocabulary from Schön and DeSanctis (1986)):

1) What is the right thing to build? ("problem-setting") To overcome the attitudinal barriers, we should design machine intelligence to skill clinicians rather than deskilling them, to *make them feel better with their decision making*, rather than making the decisions for them;
2) How do we build this? ("problem-solving") By sketching service blueprints, we identified times and places ML could be helpful, and critically reshaped the form decision supports take to better fit into day-to-day clinical practice and workplace culture (Yang, Zimmerman, Steinfeld, Carey, et al. 2016).

This case study exemplifies the complexities in the nexus of designing ML-mediated UX: ML and UX exist in a symbiotic relationship to one another. UX needs the context awareness and personalizability enabled by ML; ML is also dependent upon UX design to be perceivably valuable and useful to the users.

Recently, many researchers started to investigating human components of ML. However, there still is a lack of *integrative* anatomy that is pertinent to *UX practitioners* and apposite to ML systems i*n the wild* (as oppose to the interactive machine learning works for developers(Patel 2010), artists(Fiebrink, Cook, and Trueman 2011), makers, or end users in lab settings (Amershi 2011, 2012)).

Below I propose two fundamental dimensions along which to begin a discussion on the roles UX should play when working with ML as a design material: learning the right thing and making the learning right.

UX's Role in Learning the Right Thing

> "To be design-oriented (in HCI) is to consciously seek to intervene and manipulate, aiming to convert an undesired situation into a desirable one." (Fallman 2003)

UX Goals Converses with ML Goals

The articulation of UX goals, "the desired future", is at the center of HCI design. Early investment in activities like user studies helps designers ideate on divergent framing and solutions, and establish *what is the "right thing" to design*.

In contrast, machine learning efforts emphasize *what can be (accurately)* learned given the available datasets over a designated application. The shape of the ML-mediated future seems have been mostly driven by data availability and learner performance rather than a deliberate vision. The above case study exemplifies this inherent mismatch. Consequently, many existing ML systems fail to account for the users, or to fit into its context at large (Amershi et al. 2015; Yang, Zimmerman, Steinfeld, Carey, et al. 2016).

I promote the idea that UX of ML should be planned early in the system development process, *with goals explicitly articulated and thoughtfully seeded* (Yang, Zimmerman, Steinfeld, and Tomasic 2016). I list several common lenses through which ML goals are defined. This list is by no means meant to be complete; it is, however, meant to be useful to dissect UX consequences of machine intelligence. ML systems in the wild often need to cover and balance two or more facets. (e.g. Rader and Gray 2015)

- From *users'* lens. For example, our design of the decision support tool has the explicit goal of helping clinicians feel they are doing better work, and not necessarily automating the part of work that makes them feel like an expert.
- From *service providers'* lens. For example, recommender systems often aim for longer use sessions and more user engagement.
- From a *humanistic* lens. This perspective of design goals addresses ML consequences at large. For example, how do media content rankers minimize the filter bubble effect? How to prevent "scoring" algorithms— systems that score teachers and students, sort resumes, grant or deny loans—manifest pernicious feedback loops that enforce biases? (O'Neil 2016; Rader and Gray 2015)

UX Identifies New ML Opportunities

UX design approaches have great potential in identifying new ML opportunities, because design in early phrases *is* exploration, which means UX professionals often invest time on divergent work, essentially looking around in a design space of possibilities.

Designers can differentiate ML algorithms to be interpreted in different contexts and by different audiences. UX professionals routinely consider user needs and desires, working to embody the people they made technologies for (Zimmerman, Forlizzi, and Evenson 2007). For example, in the decision support tool case, our field study revealed rich details illustrating the clinical workflow and workplace culture, which became design material for decision support UX designs to integrate and leverage.

With adequate ML literacy, designer can even potentially invent new ways of utilizing data, envisioning and operationalizing ML methods.

More work needs to be done to enable UX practitioners to leverage ML as a design material. UX education and toolkits must provide new means that help UX teams develop a tacit understanding of ML -- that is not simply teaching how ML works, but empowering them with *enough* technical literacy to be able to 1) *ideate creatively* yet *practically,* and 2) to better collaborate with machine learning experts.

UX's Role in Getting the Learning Right

UX Manifests the ML Pipeline

ML-mediated interfaces simultaneously shape user behaviors and drive its underlying machine learning pipelines, raising exciting new challenges and opportunities for designers.

Dealing with these interfaces, UX team should consider what aspects of user behavior and context to capture in order to accurately externalize and interpret the interaction traces logged; they also consider how to detect inference errors and even capture labels to enable reinforcement learning. Ideally, *UX professionals construct interfaces in such a way that the traces of user interactions can serve as a ready source for machine learning.*

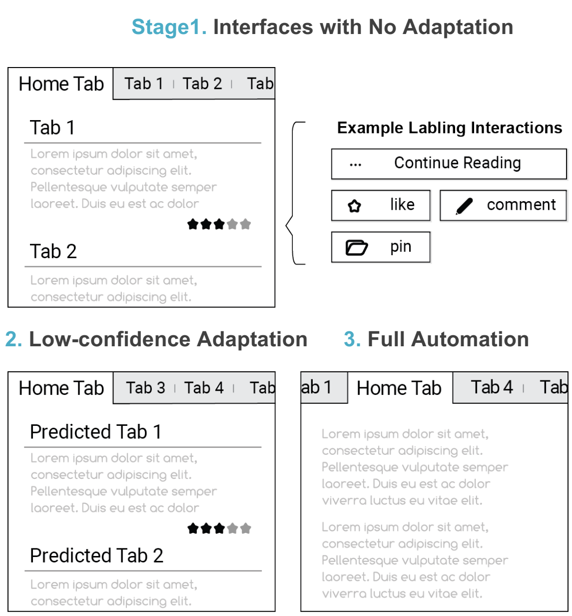

Figure 1. An Example of Adaptive UI Design Patterns - Landing Tab. On this set of interfaces, the data traces of user interactions can serve as a ready source for machine learning. (Yang, Zimmerman, Steinfeld, and Tomasic 2016)

This vision marks an exciting space for future design research in HCI. My previous work seeks solutions to this challenge in design patterns and prototyping toolkits (Figure 1, (Yang, Zimmerman, Steinfeld, and Tomasic 2016)). Others address this challenge by developing the ML system and interfaces simultaneously (e.g. (Zimmerman, Tomasic, et al. 2007), or creating an entertaining dialog that motivates users to contribute information; information that can be used as labels for personalization (e.g. (Wood 2014)).

UX Choreographs User and System Evolvement

A design that satisfies all the identified data requirements does not guarantee pleasant user experiences. In fact, creating interaction flows with two dynamic parts (users and the machine interfaces) entail many new design challenges.

UX Evolvement Over Time

ML's ability to evolve UX over time opens up exciting design opportunities for interaction and service designers. For example, data-driven personalization can serve as an important part of building and maintaining long-term customer relationships with a view toward creating customer lifetime value.

However, promises of future intelligent interactions often come at the expense of usability at the moment. In order to make confident inferences, ML systems often need to induce more explicit inputs from users. Facing a new user, how can the interaction design *induce more interactions*, in order to collect high-quality ML features and labels? Facing a user with rich use history and context data, how to find the harmony between automation and user engagement?

Appropriating UX for ML Performances

Kay, Patel, and Kientz (2015) raised the question of "how good is an 85%-accuracy classifier?", address subjectivity of ML performances. The answer to this question comes down to the tradeoff between the quality of the learning system and the risk/reward users associate with a successful or failed prediction. Designers should consider the potential benefits and costs of UX at each stage of the interface evolvement, and accordingly set an estimated accuracy threshold required for triggering the next level of personalization or automation.

Discussion

This paper elaborated on the two fundamental roles UX practice plays when designing ML-driven systems. I argue that, as most of the other emerging technologies, designing ML-mediated UX should emphasize balancing the back-end concern with ML performance and usability (getting the thing right) with an up-front investment in exploration and ideation (getting the right thing).

Like others in the HCI and UX communities, I promote the idea that ML is the new UX. I envision UX practitioners leveraging machine learning (ML) as a design material *creatively* and *thoughtfully*, guiding users and technologists toward a deliberative ML-mediated future.

More work needs to be done to advance current UX practices toward this vision. I will continue to work on understanding and supporting the conversation between ML and UX teams so that design goals can be reformulated into a problem that ML methods can directly address; I encourage the UX and HCI research community to join us in investigating new ways to evaluate the UX benefits of ML and risks of its errors; join us in developing new sketching and prototyping tools so that practitioners can better investigate and communicate their ML design ideas.

References

Amershi, Saleema. 2011. "Designing for Effective End-User Interaction with Machine Learning." P. 47 in *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology - UIST '11 Adjunct*. New York, New York, USA: ACM Press.

Amershi, Saleema. 2012. "Designing for Effective End-User Interaction with Machine Learning." *Proceedings of the 2012 ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12* 3542114:144.

Amershi, Saleema, Maya Cakmak, W.Bradley Knox, and Todd Kulesza. 2015. "Power to the People: The Role of Humans in Interactive Machine Learning.".

Bellazzi, Riccardo and Blaz Zupan. 2008. "Predictive Data Mining in Clinical Medicine: Current Issues and Guidelines." *International Journal of Medical Informatics* 77:81–97.

Brownlee, John. 2015. "Apple Finally Learns AI Is The New UI." Retrieved December 25, 2015 (http://www.fastcodesign.com/3047199/apple-finally-learns-ai-is-the-new-ui).

Buxton, Bill. 2010. *Sketching User Experiences: Getting the Design Right and the Right Design*. Morgan Kaufmann. Retrieved October 27, 2016.

Fallman, Daniel. 2003. "Design-Oriented Human-Computer Interaction." Pp. 225–32 in *Proceedings of the International Conference on Human Factors in Computing Systems (CHI'03)*.

Fiebrink, Rebecca, Perry R. Cook, and Dan Trueman. 2011. "Human Model Evaluation in Interactive Supervised Learning." P. 147 in *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. New York, New York, USA, NY, USA: ACM Press.

Kay, Matthew, Shwetak N. Patel, and Julie A. Kientz. 2015. "How Good Is 85%?: A Survey Tool to Connect Classifier Evaluation to Acceptability of Accuracy." Pp. 347–56 in Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2015).

Kuang, Cliff. 2013. "Why a New Golden Age for UI Design Is Around the Corner." Wired Magazine 21.09, August. Retrieved December 28, 2015 (http://www.wired.com/2013/08/design-and-the-digital-world/).

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown/Archetype.

Patel, K. 2010. "Lowering the Barrier to Applying Machine Learning." Pp. 355–58 in *CHI '10 Doctoral Consortium*.

Rader, Emilee and Rebecca Gray. 2015. "Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed." Pp. 173–82 in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. New York, New York, USA: ACM Press.

Schon, Donald A. and Vincent DeSanctis. 1986. "The Reflective Practitioner: How Professionals Think in Action." *The Journal of Continuing Higher Education* 34(3):29–30. Retrieved January 26, 2016.

Wood, Molly. 2014. "Beats Hopes to Serve Up Music in a Novel Way." *The New York Times*. Retrieved (http://nyti.ms/1fr8LY5).

Yang, Qian, John Zimmerman, and Aaron Steinfeld. 2015. "Review of Medical Decision Support Tools : Emerging Opportunity for Interaction Design." in *IASDR 2015 Interplay Proceedings*. Brisbane, Australia.

Yang, Qian, John Zimmerman, Aaron Steinfeld, Lisa Carey, and James F. Antaki. 2016. "Investigating the Heart Pump Implant Decision Process: Opportunities for Decision Support Tools to Help." Pp. 4477–88 in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. New York, New York, USA: ACM Press.

Yang, Qian, John Zimmerman, Aaron Steinfeld, and Anthony Tomasic. 2016. "Planning Adaptive Mobile Experiences When Wireframing." Pp. 565–76 in *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16*. New York, New York, USA: ACM Press.

Zimmerman, John, Anthony Tomasic, et al. 2007. "Vio: A Mixed-Initiative Approach to Learning and Automating Procedural Update Tasks." P. 1445 in *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*. New York, New York, USA: ACM Press.

Zimmerman, John, Jodi Forlizzi, and Shelley Evenson. 2007. "Research Through Design as a Method for Interaction Design Research in HCI."