SALEEMA AMERSHI | RESEARCH STATEMENT

The unprecedented opportunity for big data to enhance our capabilities and improve our lives is limited by our ability to *use* that data. Machine learning can give us this ability by transforming raw data into the building blocks necessary for configuring automated behaviors. However, the complexity of machine learning has largely restricted its use to experts and skilled developers. For example, trained developers can employ machine learning to automatically detect objects, organize information, and understand and predict behaviors. In contrast, ill-equipped end-users are limited to using machine learning for simple personalization based on developer-conceived notions of interest or similarity. Even within personalization systems, end-user control over the machine learning is typically and intentionally reduced to simple object labeling. **As a human-computer interaction researcher, I aim to put the full potential of machine learning in the hands of everyday people.**

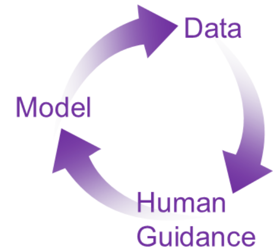

Figure 1. In the end-user interactive machine learning process, a person iteratively guides a machine towards learning a desired concept. The person then inspects the system's current understanding and decides how to proceed with further guidance. The learned concept is encoded by the machine as a classifier model that can configure automated behaviors on data.

CURRENT RESEARCH

My dissertation examines the fundamental process by which end-users interact with machine learning systems (Figure 1). In this process, a person iteratively guides a machine towards learning a desired concept and inspects feedback illustrating the machine's current understanding. The outcome of this process is a machine learned *model* which can configure automated behaviors on data.

Existing applications generally optimize for end-user flexibility during this interaction (e.g., personalization systems that allow for indiscriminate object ratings). While this simplifies the interaction, it can also create a frustrating experience if a person cannot effectively guide the system. On the other hand, a traditional active learning approach that neglects end-user needs in favor of the machine (e.g., forcing a person to label items providing the most information gain), can be equally frustrating by discounting human abilities and treating a person as a passive oracle of information. Effective solutions must therefore balance the needs of *both* the end-user and the machine.

My research target real-world problems that can benefit from end-user driven machine learning. Throughout my work, I identify challenges and opportunities for improving the interactive machine learning process and design new and balanced solutions. I also distil guiding principles applicable in a broader context, providing a foundation for future end-user interactive machine learning systems.

CueFlik: Image Classification

Keywords are a fundamentally impoverished method for characterizing images. Modern search engines therefore combine keyword queries with content-based filters (e.g., black and white images, clipart, or images with faces). But consider that no major search engine provides a filter for a "product" concept (i.e., crisp images of objects on empty backgrounds). This is not a difficult computer vision problem. Instead, the challenge is that developers cannot possibly foresee and provide all potentially desired concepts.

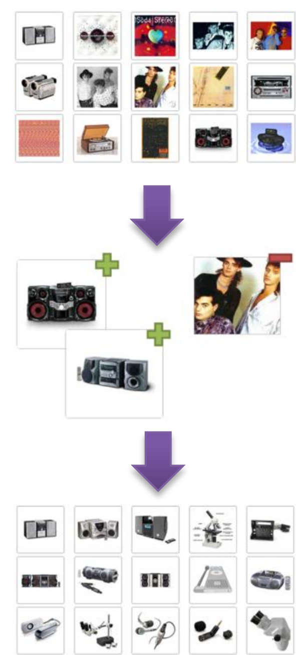

Figure 2. CueFlik is an end-user interactive machine learning system for image classification. End-users create their own content-based classifiers which they can apply to filter a set of images.

CueFlik enables end-users to define their own visual concepts for image classification. End-users train CueFlik to recognize a concept by providing examples of images with and without the desired characteristics (Figure 2). Formally, CueFlik uses these examples to learn a distance metric it then applies with a nearest-neighbor classifier.

How do we reach a shared understanding? Effective machine learning requires the end-user and machine reach a shared understanding of a desired concept. In the interactive machine learning process, this concerns both assessing the quality of the current model and further guiding the system. The standard technique is to present a person with all of the data ranked according to the current model. From this presentation, a person can evaluate the model and decide how to proceed in training. While straightforward, this technique inefficiently illustrates the current concept and provides no guidance for improving the model. An alternative is to present only the best and worst matching examples. This has the advantage of summarizing the machine's understanding and has been shown to improve resulting models. Yet, best and worst matches are extremely similar to already labeled examples (top of Figure 3). Therefore, this technique actually poorly summarizes the current model and constrains the end-user to selection of further examples that provide little additional value to the machine.

We designed two new techniques for concisely summarizing the current model while providing high-value examples to choose from during training (Figure 3). Our *global overview* technique selects representative examples that maximize the mutual information with the rest of the space. This technique has been shown to result in more informative examples than those selected at random. Our *projected overview* selects examples that illustrate variation along major dimensions of a space. We achieve this using a non-linear projection to identify principle dimensions and a novel technique for selecting examples providing coverage of a single dimension but also varying as little as possible in all other dimensions. We found that *overviews* led end-users to select better examples and improve the quality of their resulting models.

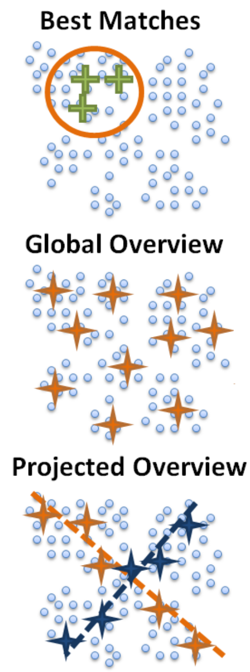

Figure 3. Our *overview* techniques of presenting high-value, representative examples of a region lead people to train better models than the best performing technique from previous work (i.e., *best matches*).

How do we enable model exploration? Prior interactive machine learning research has focused interaction on asking a person *"what class is this object?"* However, individual objects are generally ancillary to the resulting model. We therefore proposed that a person instead consider *"how will different labels for this object affect the model in relation to my goals?"* To achieve this interaction, we augmented CueFlik with support for exploring different models via revision and a history visualization for comparing alternatives. We found that end-users readily adopted revision and this led them to create better models than when revision was not available. While comparison of alternatives is a proven technique in human-computer interaction, it has not been explored in the context of people interacting with machine learning. This highlights the importance of a balanced approach to interaction design.

Our work with CueFlik has appeared in leading human-computer interaction conferences including UIST 2009 [6] and CHI 2010 [5]. In addition, given the interdisciplinary appeal of this work, CueFlik also appeared in AAAI 2011 [3].

ReGroup: Custom Access Control Groups in Online Social Networks

The prevailing approach to access control in online social networks is to pre-categorize friends into groups in advance of sharing decisions. However, usable security research has shown that pre-defined groups do not correspond to groups desired in-context (e.g., the notion of "close friends" may change depending upon the content being shared). In the absence of an appropriate group, people have little choice but to manually create a new group, share unrestrictedly, or not share at all.

Motivated by this problem, we created ReGroup, an interactive machine learning system for helping people create custom, on-demand groups in social networks (Figure 4). As a person adds members to a group, ReGroup iteratively learns a probabilistic model of group membership to assist in group creation.

Figure 4. ReGroup uses interactive machine learning to help people create on-demand groups. As a person selects group members, ReGroup suggests additional members and group characteristics for filtering.

How do we support multiple forms of interaction? Most interactive machine learning systems concentrate on examples. However, feature-based feedback has been shown to accelerate learning. ReGroup therefore uses its learned model to both suggest group members and enable interaction with features. We realized feature-based interaction via a familiar faceted browsing metaphor. As ReGroup learns a model of group membership, it suggests feature-based facets representing group characteristics. Suggested facets serve as both an explanation of group member recommendations and a tool for filtering. Because previous facet ranking solutions were unsuitable for these purposes, we developed a new decision-theoretic technique for suggesting facets. Our technique combines the likelihood of a facet-value given the current group and the expected utility of a facet-based filter represented by the information gain of a feature.

How do we generalize from incomplete information? Machine learning is a function of the information provided during training. For example, effective machine learning typically requires both positive and negative training examples. However, a person's ultimate goal with ReGroup is to create a group, meaning they are primarily focused on selecting positive examples. To mitigate this imbalance, ReGroup implicitly infers negative examples by observing when people skip over friend suggestions during the interaction process. The predictability of machine learning is also reduced by missing data – rampant in social networks. ReGroup therefore probabilistically estimates missing values conditioned on available data. ReGroup also highlights its estimates in its interface, serving as an explanation of its behavior in the face of missing data.

How do we handle unlearnable concepts? Machine learning systems suffer when the hypothesis language is too inexpressive to model the desired concept. In ReGroup, this can result in repeated suggestion of a skipped friend due to their similarity with positive examples as represented by the system (e.g., a person planning a surprise party would obviously not want to invite the guest of honor). To reduce frustration in dealing with unlearnable groups, we introduced an explicit penalty term in ReGroup's membership estimation that decreases a friend suggestion probability with every skip.

End-users found ReGroup to be a powerful complement to manual group creation. Furthermore, by facilitating on-demand group creation, ReGroup could potentially encourage better online privacy practices. ReGroup will appear at CHI 2012 [1].

CueT: Computer Network Alarm Triage

Alarm triage is the first line of defense for large computer networks. Triage refers to the grouping of a stream of low-level device health information according to shared underlying causes. Groupings are then inspected to diagnose network problems. It is critical that triage is fast and accurate so problems are identified and resolved quickly. However, current solutions require human operators to manually sift through and group thousands of alarms per day. To assist operators with alarm triage, we developed CueT, an interactive machine learning system for making triage recommendations.

How do we learn in a dynamic setting? Interactive machine learning is challenging in the alarm triage scenario because the classes of interest are not known *a priori* and evolve constantly. Therefore, we designed CueT to continually update its underlying classes based on operator actions (Figure 5). As a new alarm appears, CueT makes triage recommendations by ordering the groups currently in its system by their distance to the alarm. If an operator adds the alarm to an existing group, the corresponding class is updated. Alternatively, if the alarm is sufficiently distant from existing groups as defined by an updating threshold, CueT suggests creating a new group. Starting a new group creates a new class which can then be used for future recommendations.

How do we convey model confidence? An arbitrary ordering of comparable recommendations can present end-users with a false sense of confidence. This is particularly problematic in scenarios where accurate classification in critical. We therefore complement CueT's recommendations with a novel visualization designed to encourage inspection of comparable recommendations (Figure 6).

In an evaluation with real network operators and with real data from a large computer network, we found that operators could triage alarms faster and more accurately with CueT than with the state-of-the-art triage method. CueT was nominated for a Best Paper Award at CHI 2011 [4] and was invited to IJCAI 2011 [2].

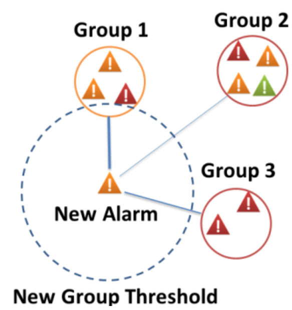

Figure 5. CueT helps operators by recommending triage actions for a stream of network alarms. Recommendations are either to triage alarms into existing groups or start a new group. In this example, there are three existing groups. The dashed circle indicates the distance thresholds at which CueT suggests an alarm might be related to a new network problem. The distance metric and the new ticket thresholds are both learned from operator actions.

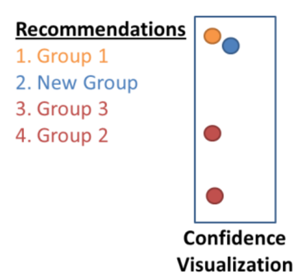

Figure 6. Triage errors can be very problematic, and a naïve ranked list can obscure ambiguity. CueT instead uses a compact confidence visualization to help draw operator attention to ambiguous alarms.

FUTURE PLANS

I envision people using machine learning to harness big data to achieve their goals and improve their lives. Realizing this future will raise new issues for human-computer interaction that I am excited to pursue.

Using Data to Solve Important Problems

Big data is a vast untapped resource that could potentially fuel the next generation of applications. I plan to aggressively pursue new opportunities for people to use this data via machine learning in wide variety of compelling application domains. For example, I believe that, if given the right tools and support, end-users could teach machines to recognize human behaviors in sensory data for health and elder-care applications, to monitor environmental impact in mobile data for green technologies, to understand human learning in Web data for online educational environments, and to sense brain and muscle activity to drive new and accessible input techniques. These are urgent problems, the solutions to which can change the way we live.

Enabling the Next Generation of Applications

Enabling end-user interactive machine learning applications in new domains will require answers to many challenging research questions. For example, how can a person guide a machine to detect complex and temporal activities? How can a machine illustrate concept understanding in sensory data? How can we incorporate domain knowledge of professionals driving some of these applications? As my dissertation work demonstrates, we must take a principled approach to these problems if we want to move beyond naïve or ad hoc solutions. I plan to advance our understanding of how to design effective end-user interaction with machine learning to provide a foundation for future researchers and application developers.

Leveraging the Masses to Scale Up the Impact and Usability of Data

The proliferation of social networks and community-driven websites is evidence of the power of the masses for creating and contributing content. So too can leveraging the masses scale up the impact and usability of data. For example, as machine learning becomes a reality in our everyday applications, people should be able to re-use already created models rather than starting from scratch. Moreover, people should be able to bootstrap, build upon and combine models to configure more sophisticated data processing and manipulation. Finally, the large-scale and high-impact problems of the future will demand the effort of crowds working in collaboration with machine learning systems (e.g., [7]). To achieve these possibilities, we must understand how people can meaningfully describe, compare, and search for existing machine learning models, how models can be generalized or transformed for new situations and purposes, and how we can create composable models to enable more powerful automation.

The past twenty years have seen the emergence of search over data. I expect the next twenty will see machine learning become an integral part of end-user interaction with data. The research needed to achieve this future will require a unique skillset: expertise in human-computer interaction, an understanding of machine learning, and an ability to collaborate with machine learning and domain experts. I plan to be a leader in this effort.

REFERENCES

1. **Amershi, S.**, Fogarty, J. and Weld, D.S. (2012) ReGroup: Interactive Machine Learning for On-Demand Group Creation in Social Networks. To Appear in *Proceedings of CHI 2012*.
2. **Amershi, S.**, Lee, B., Kapoor, A., Mahajan, R. and Christian, B. (2011) Human-Guided Machine Learning for Fast and Accurate Network Alarm Triage. *Proceedings of IJCAI 2011*, Best Papers from Sister Conferences Track, pp. 2564-2569. ***Invited.***
3. **Amershi, S.**, Fogarty, J., Kapoor, A. and Tan, D. (2011) Effective End-User Interaction with Machine Learning. *Proceedings of AAAI 2011*, Nectar Track, pp. 1529-1532.
4. **Amershi, S.**, Lee, B., Kapoor, A., Mahajan, R. and Christian, B. (2011) CueT: Human-Guided Fast and Accurate Network Alarm Triage. *Proceedings of CHI 2011*, pp. 157-166. ***Best Paper Nominee.***
5. **Amershi, S.**, Fogarty, J., Kapoor, A. and Tan, D. (2010) Examining Multiple Potential Models in End-User Interactive Concept Learning. *Proceedings of CHI 2010*, pp. 1357-1360.
6. **Amershi, S.**, Fogarty, J., Kapoor, A. and Tan, D. (2009) Overview-Based Examples Selection in Mixed-Initiative Interactive Concept Learning. *Proceeding of UIST 2009*, pp. 247-256.
7. Hoffmann, R., **Amershi, S.**, Patel, K., Wu, F., Fogarty, J., and Weld, D.S. (2009) Amplifying Community Content Creation with Mixed-Initiative Information Extraction. *Proceedings of CHI 2009*, pp. 1849-1858. ***Best Paper Nominee.***