Interactive Machine Learning for Sensing Applications

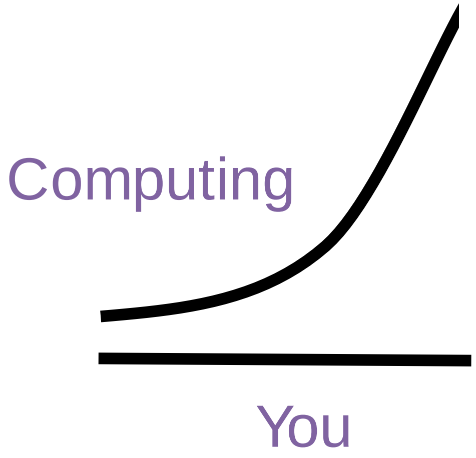
- Slides from **James Fogarty** (University of Washington) https://www.youtube.com/watch?v=qf7bPY5xDI8&feature=emb_logo
- **A Review of User Interface Design for Interactive Machine Learning**, John Dudley and Per Ola Kristensson, IUI'19
- **Power to the People: The Role of Humans in Interactive Machine Learning**, Saleema Amershi, Maya Cakmak, William Bradley Knox, Todd Kulesza, AI Magazine, 2014

Overview

- Motivation
- Interactive Machine Learning
- Case Studies
  - Ask users to check - Crayon
  - Ask users about "concept“ - CueFlik
  - Ask users to build a model - GestureScript
- Studying User Interaction with IML
  - Users are people, not oracles
  - People want to demonstrate how learners should behave
  - People naturally want to provide more than just data labels
  - People value transparency in learning systems, and transparency can help people provide better labels
- Takeaways

Opportunity of Machine Learning

Machine learning offers a unique tool for scaling human attention to new forms of data



The potential to use the power of this exponential curve to pull ourselves up

The Challenge of Machine Learning

Applying machine learning remains hard

Difficult to understand relationships between data and the behavior of machine learning algorithms

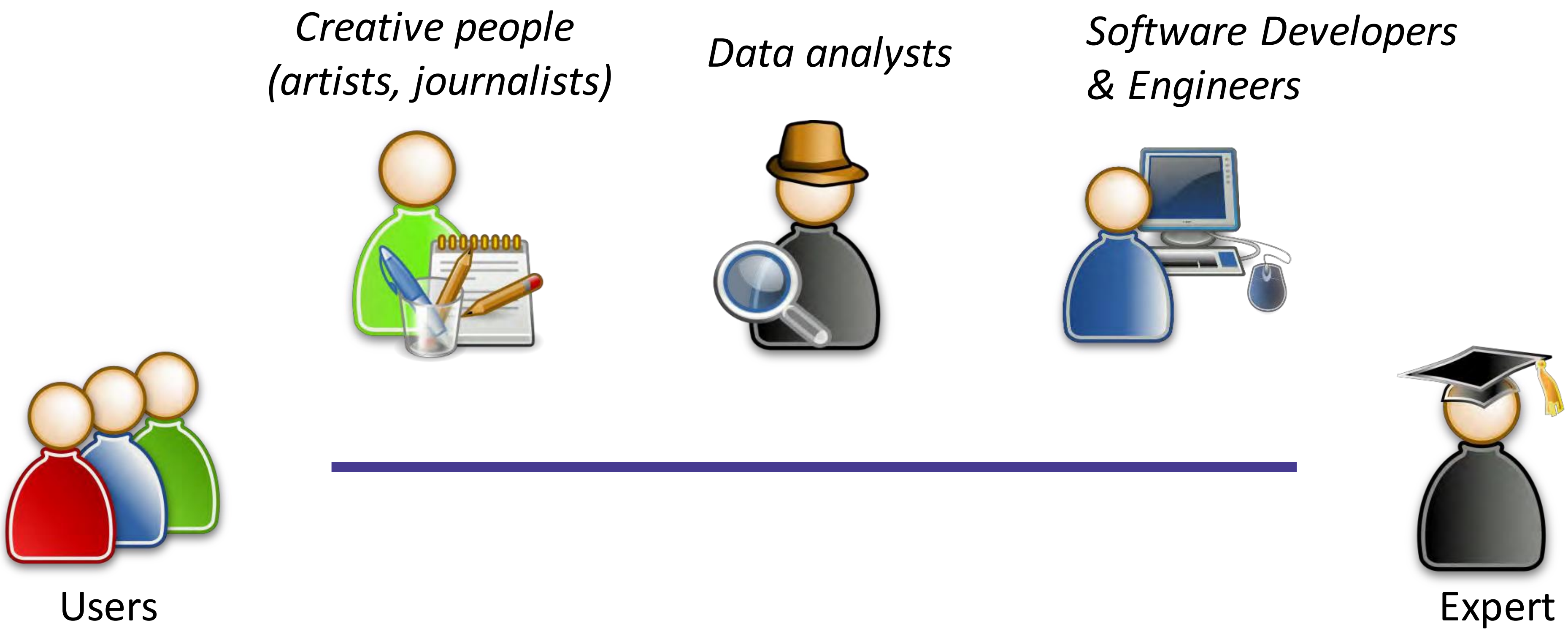
Difficult as an iterative and exploratory process

Difficult to evaluate in the context of applications

Limited to people
with highly-specialized training

Patel, Fogarty, Landay, Harrison.
Investigating Statistical Machine Learning as a Tool for Software Development. *CHI 2008*.

Everyday Interaction with Machine Learning

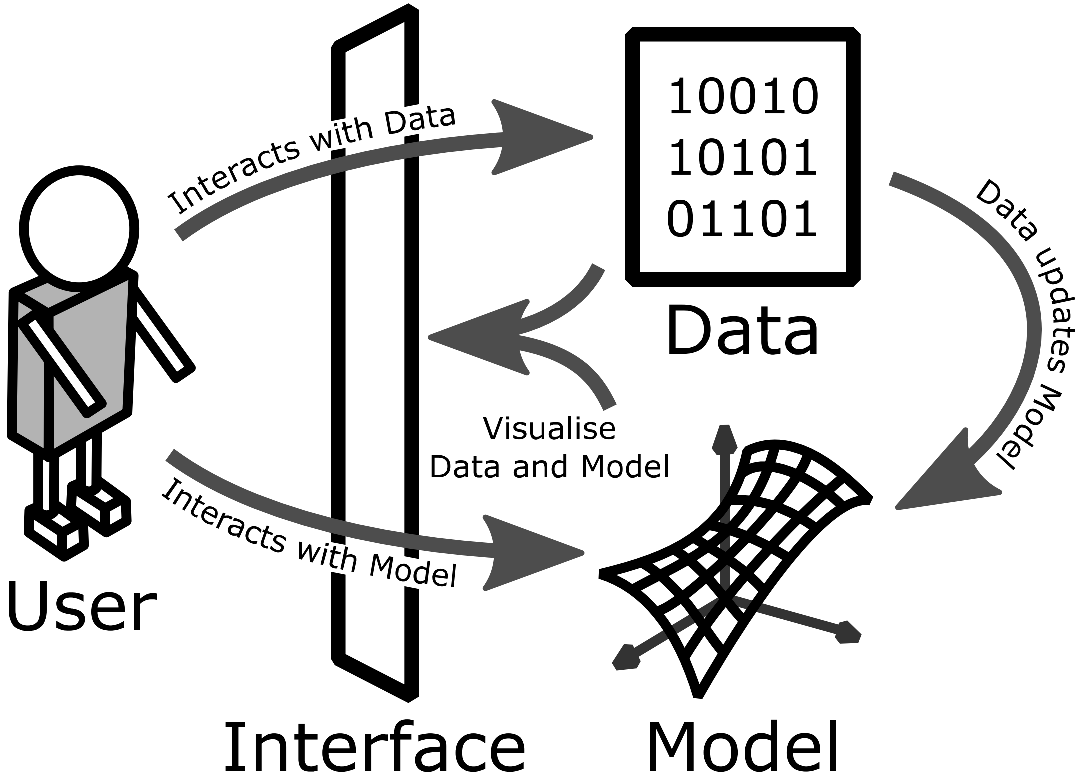


Need a range of tools for machine learning, across many domains and levels of expertise

Overview

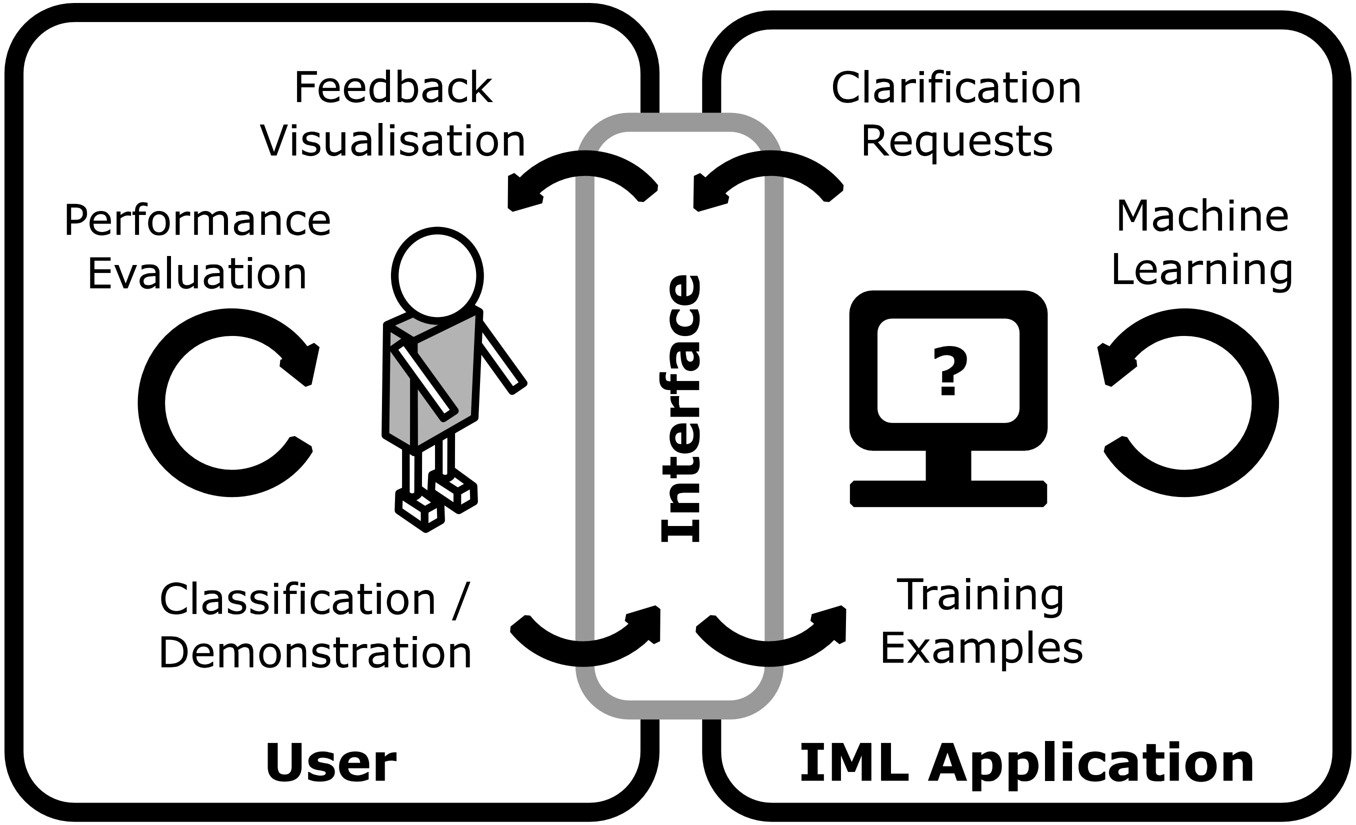
- Motivation
- Interactive Machine Learning
- Case Studies
  - Ask users to check - Crayon
  - Ask users about "concept“ - CueFlik
  - Ask users to build a model - GestureScript
- Studying User Interaction with IML
  - Users are people, not oracles
  - People want to demonstrate how learners should behave
  - People naturally want to provide more than just data labels
  - People value transparency in learning systems, and transparency can help people provide better labels
- Takeaways

Structural Breakdown of IML



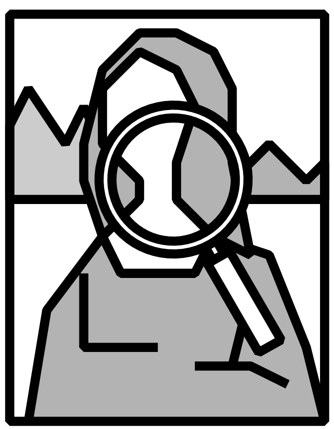
Who are the users?

Interactive Machine Learning



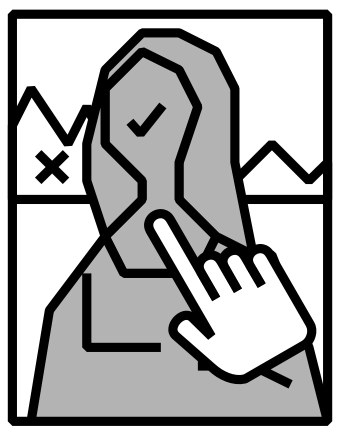
Interactive Machine Learning is an interaction paradigm in which a user or user group **refines** an ML model through iterative cycles of input and review

Composition of an IML Interface



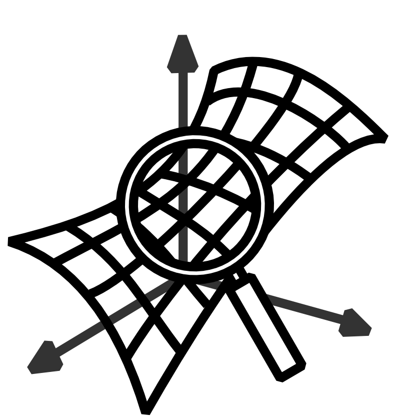
Sample Review

Visualization of output sample(s) to assess how well desired concept operates at the instance level.



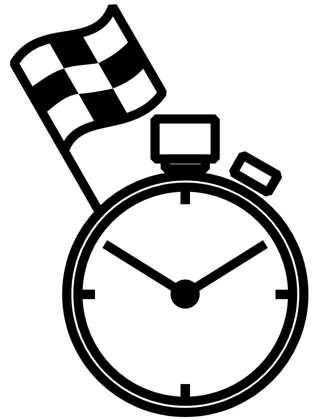
Feedback Assignment

Assignment or correction of labels and/or creation of new samples to improve match with desired concept.



Model Inspection

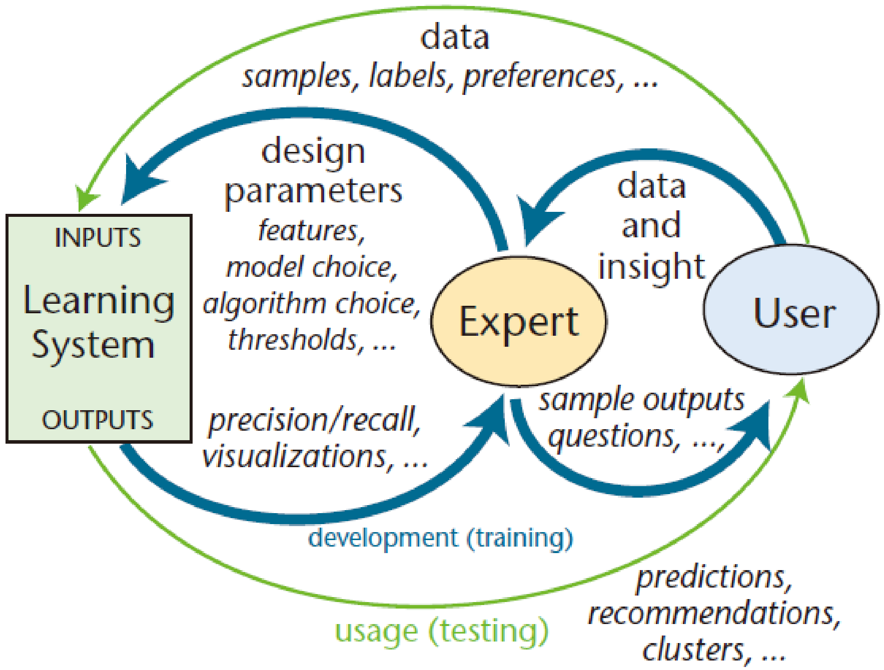
View of overall model quality and coverage to evaluate how well concept is captured.



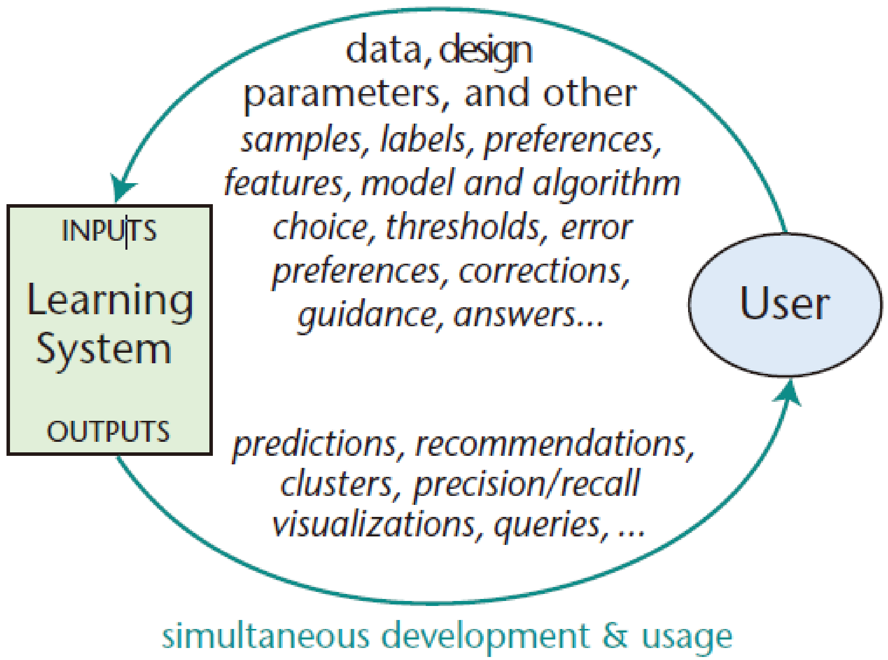
Task Overview

View of current task status contexualized by coverage of training data and improvement potential relative to cost.

Comparing processes for constructing machine learning systems



Traditional machine learning

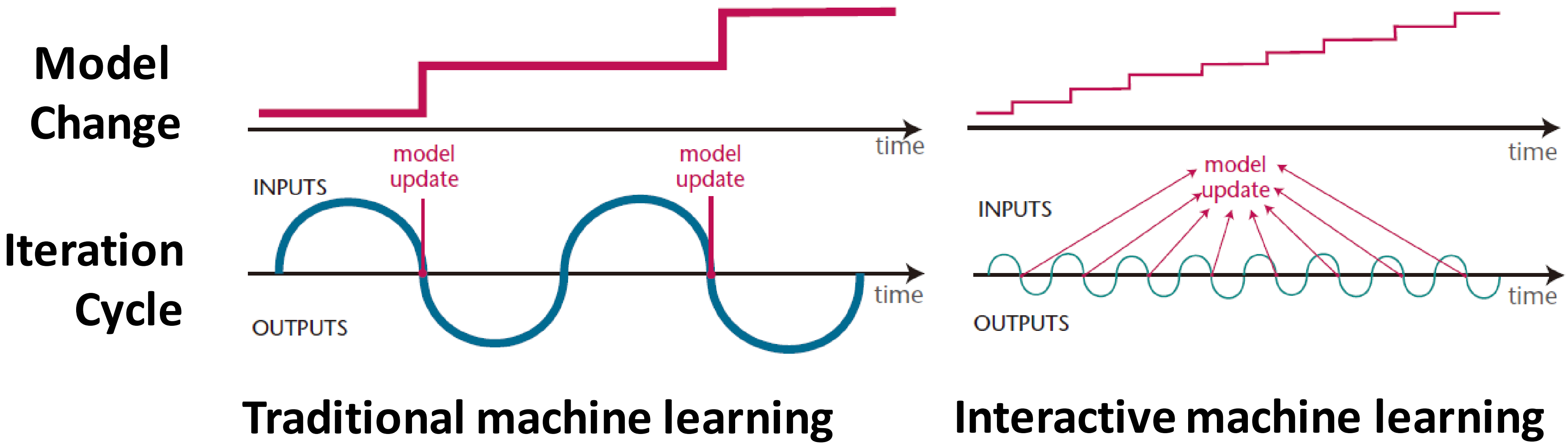


Interactive machine learning

Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the People: The Role of Humans in Interactive Machine Learning. AI Magazine, 35(4), 105-120. https://doi.org/10.1609/aimag.v35i4.2513

Interactive Machine Learning

- Rapid, focused, and incremental learning cycles result in a tight coupling between the user and the system, where the two influence one another
- As a result, it is difficult to decouple their influence on the resulting model and study such systems in isolation



Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the People: The Role of Humans in Interactive Machine Learning. AI Magazine, 35(4), 105-120. https://doi.org/10.1609/aimag.v35i4.2513

Example modes of “user interactions”

- Teaching by showing samples (labeling) (visual interactive labeling)
- Teaching by demonstration
- Teaching concepts by examples
- Teaching by reinforcement (via feedback)
- Testing learner’s state: checking/evaluating student’s learning

Towards Understanding How Humans Teach Robots, UMAP 2011

Overview

- Motivation
- Interactive Machine Learning
- Case Studies
  - Ask users to check - Crayon
  - Ask users about "concept“ - CueFlik
  - Ask users to build a model - GestureScript
- Studying User Interaction with IML
  - Users are people, not oracles
  - People want to demonstrate how learners should behave
  - People naturally want to provide more than just data labels
  - People value transparency in learning systems, and transparency can help people provide better labels
- Takeaways

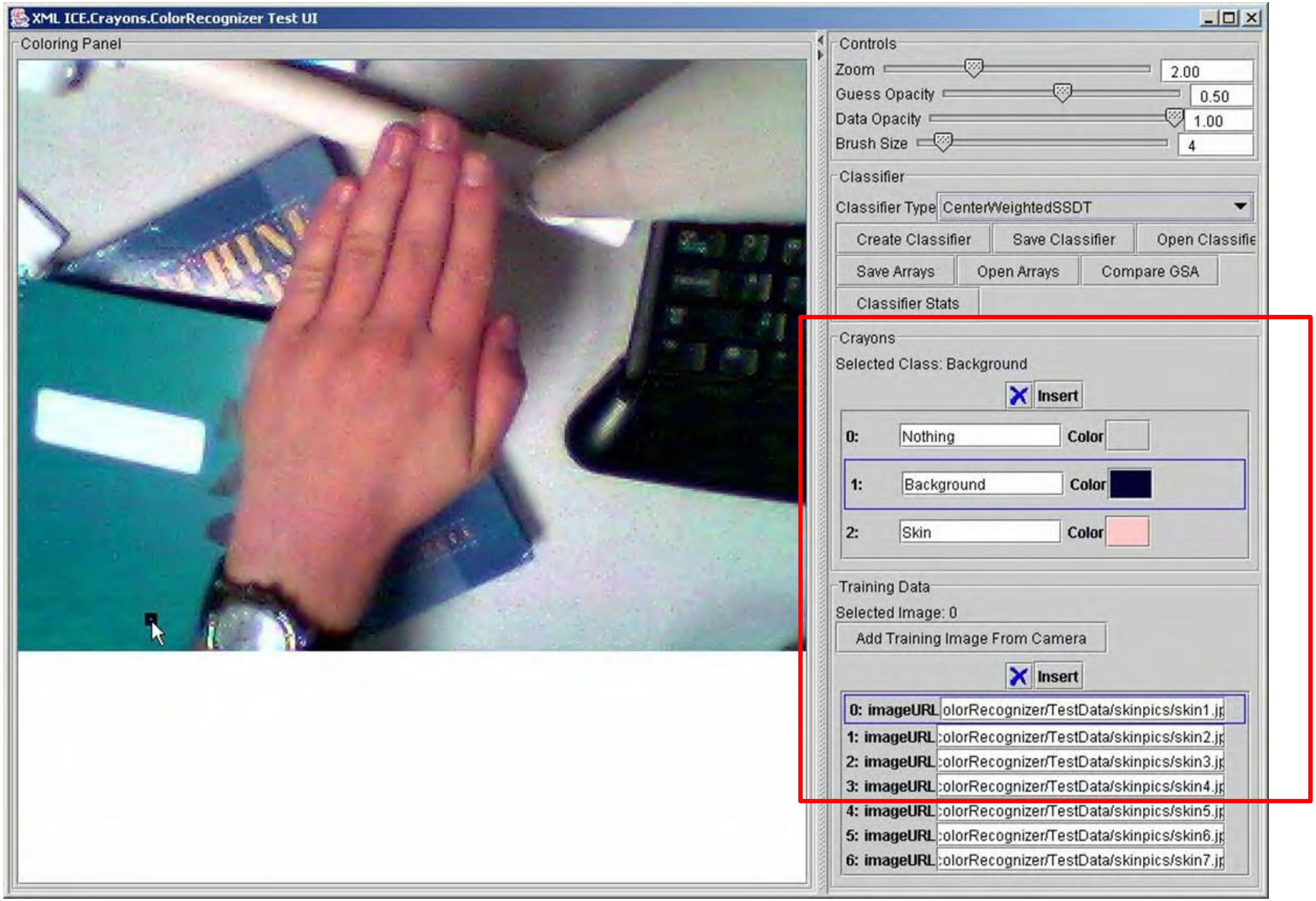
Crayons

Fails, Olsen. Interactive Machine Learning. *IUI 2003*.

Crayons

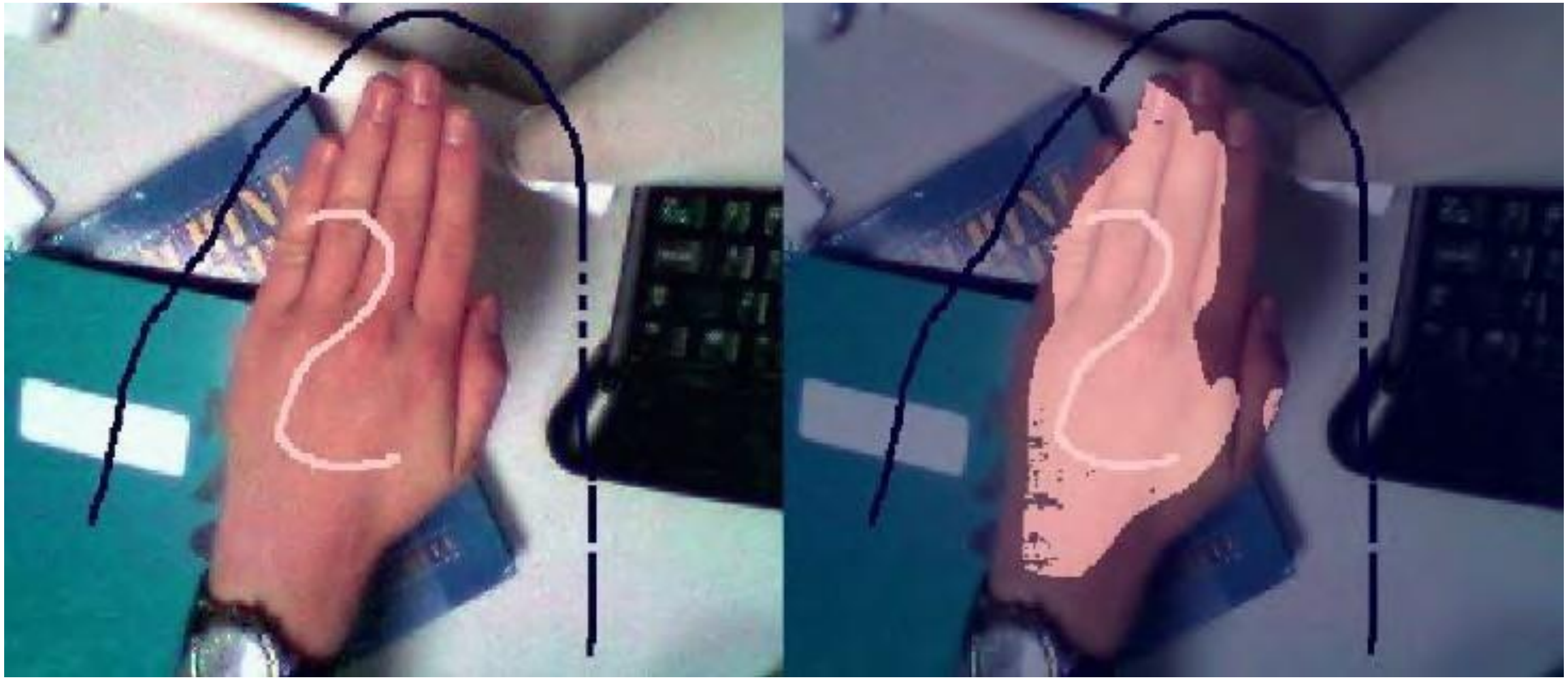
Interactive training of image segmentation

Interactive Training and Correction



Crayons

Interactive training of image segmentation

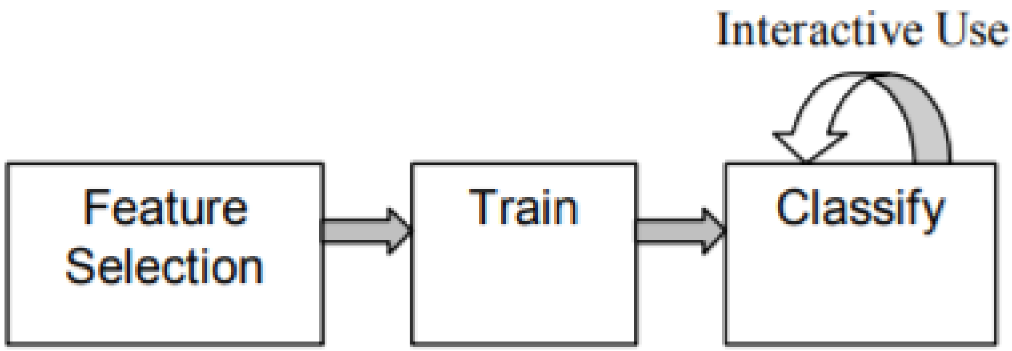


Crayons

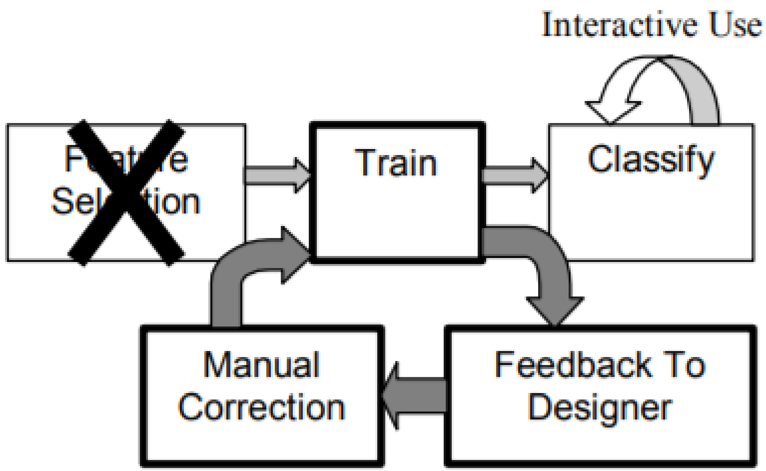
Interactive training of image segmentation

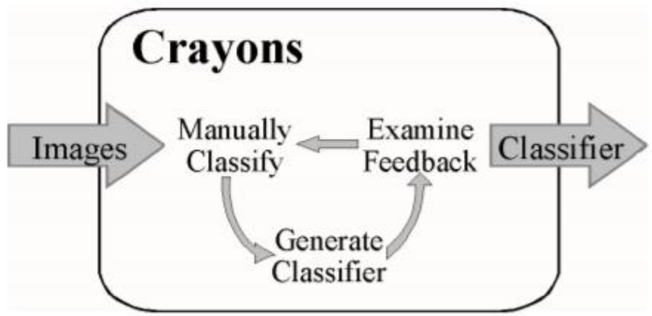
Crayons

Interactive training of image segmentation



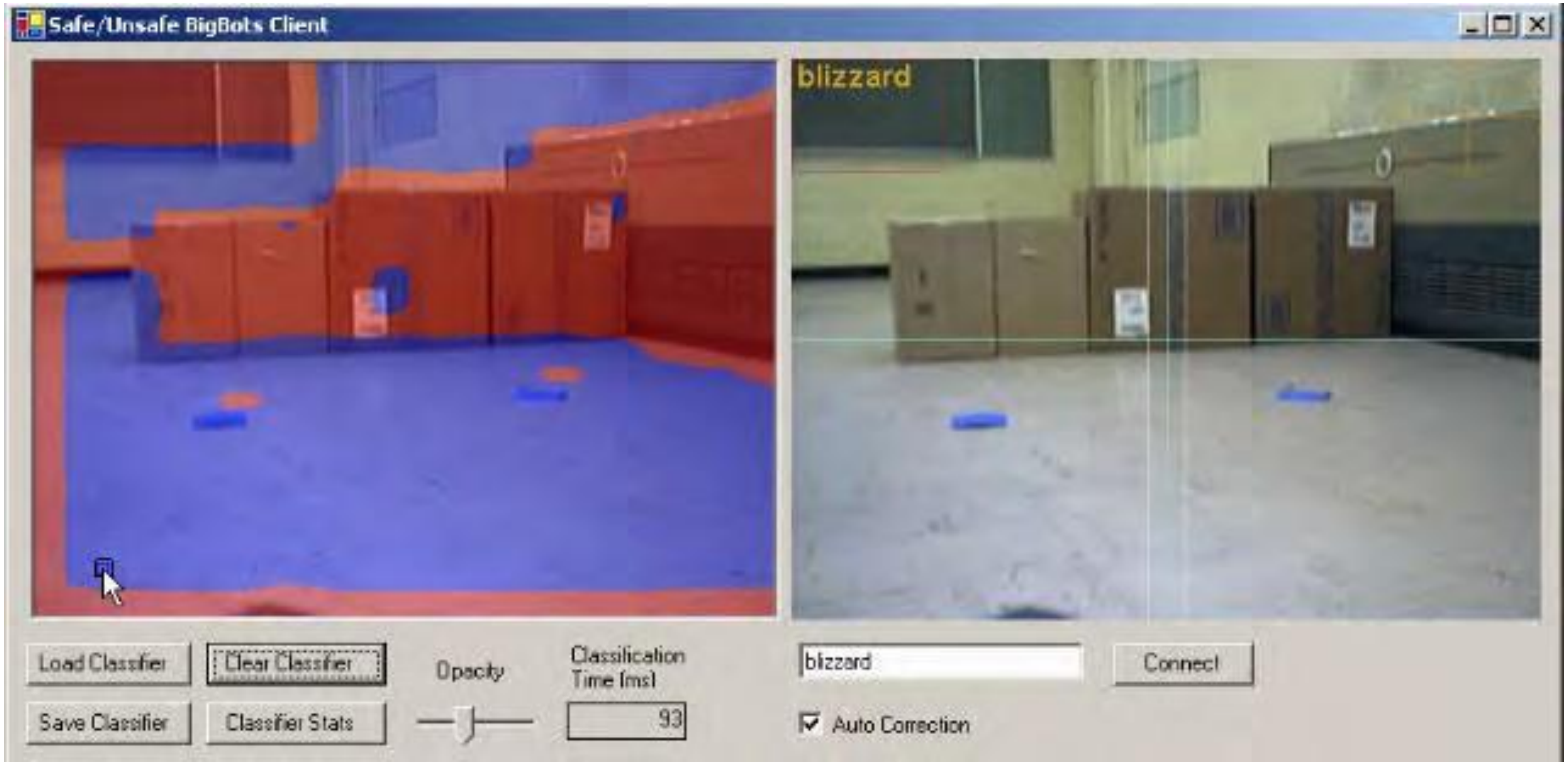
Classical machine learning model





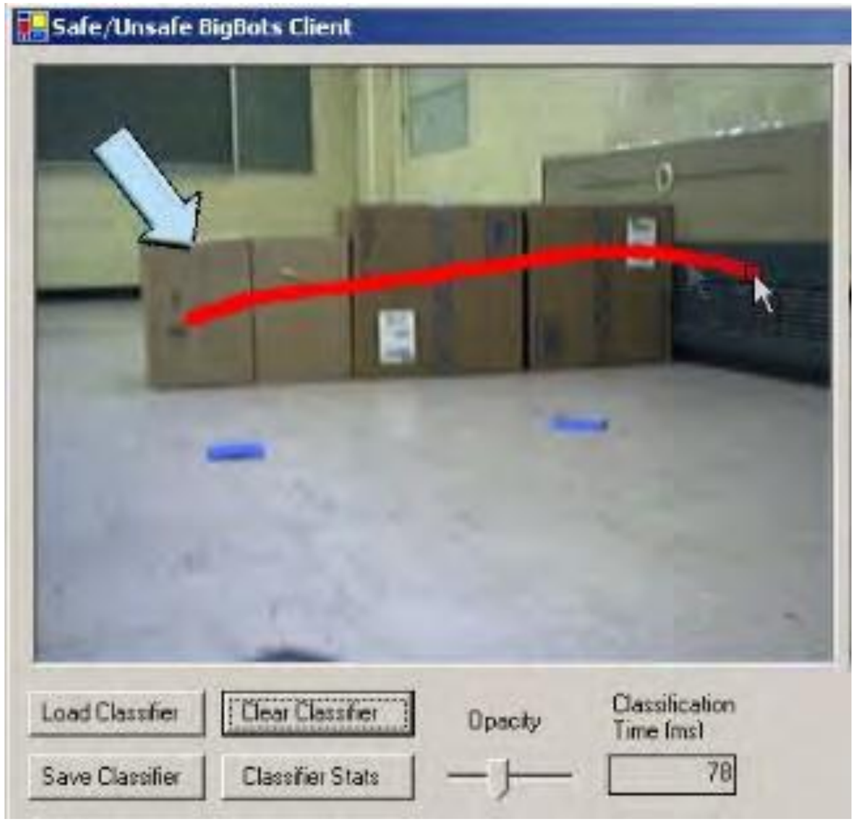
Interactive machine learning w/ Crayon

Machine Learning as Interaction

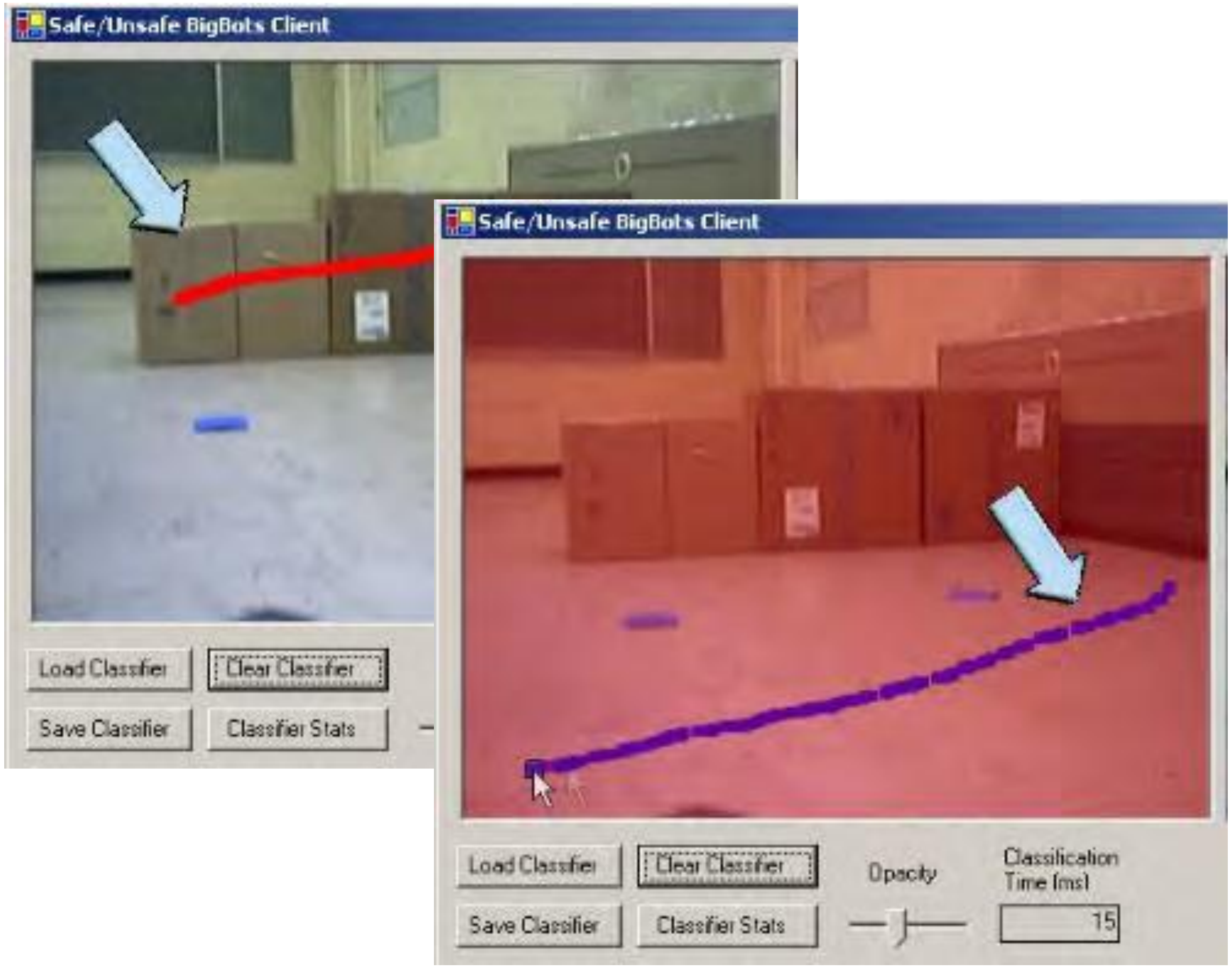


Classifier to identify safe driving paths, augmenting interactive driving controls

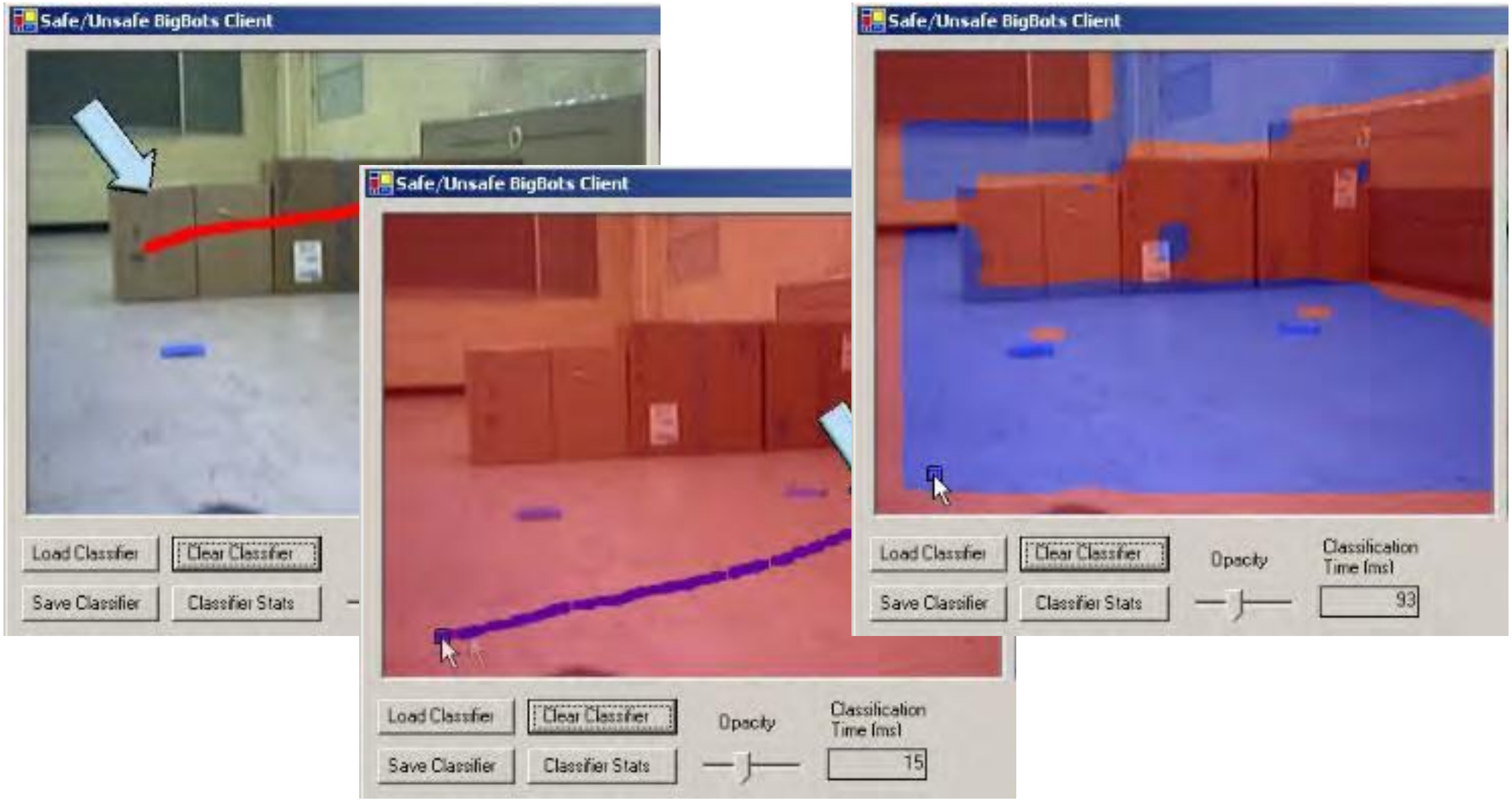
Machine Learning as Interaction



Machine Learning as Interaction



Machine Learning as Interaction



Range of Capability vs. Expertise

Domain-specific interactive tool

Easy to use

Specific capability

General purpose machine learning

High barriers to use

Flexible capabilities

CueFlik

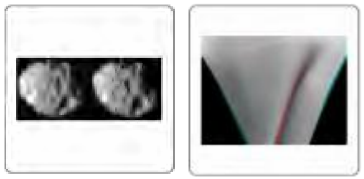
Fogarty, Tan, Kapoor, Winder. CueFlik: Interactive Concept Learning in Image Search. *CHI 2008.*

Searching for “Product” Images

Keywords are generally insufficient for visually characterizing images

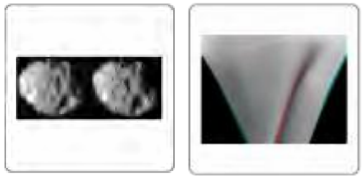
How do you query for ‘product’ images?

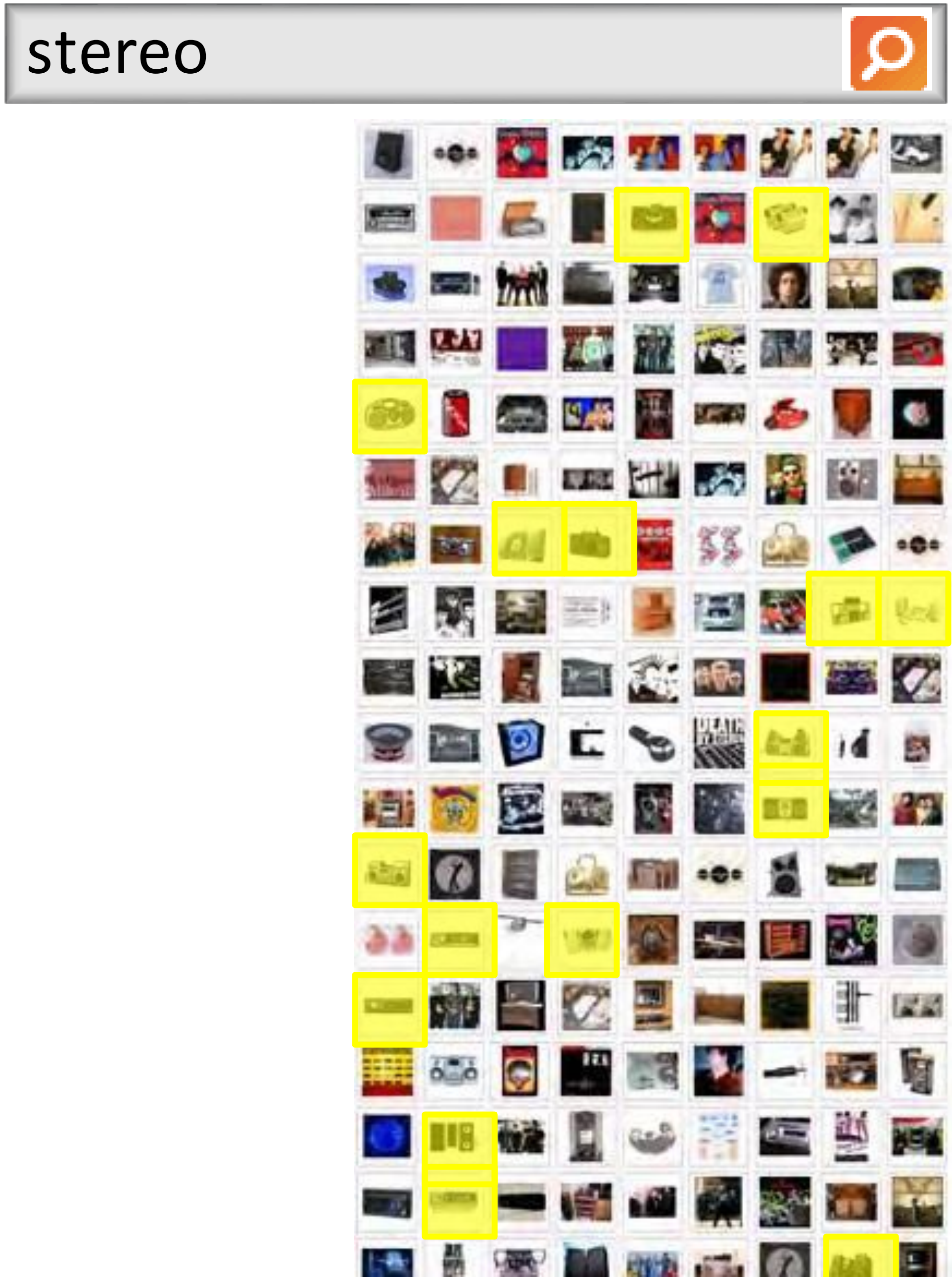
Mismatch of Keywords and Intent



Mismatch of Keywords and Intent

10% match
in top results





Mismatch of Keywords and Intent

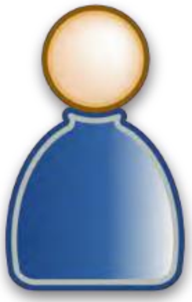
stereo on white background

Mismatch of Keywords and Intent

12% match
in top results

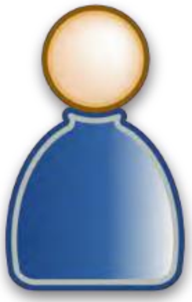
CueFlik

Interactive Concept Learning in Image Search



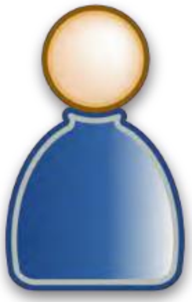
Fogarty, Tan, Kapoor, Winder. CueFlik: Interactive Concept Learning in Image Search. *CHI 2008.*

CueFlik

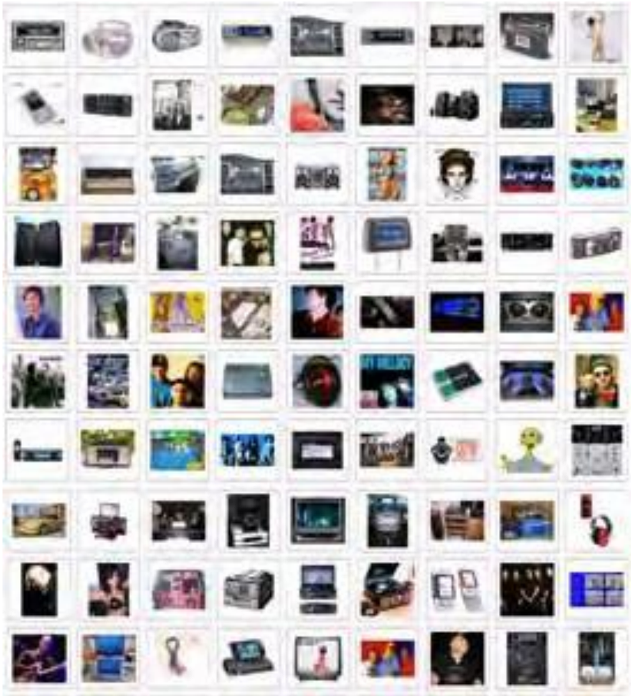


CueFlik

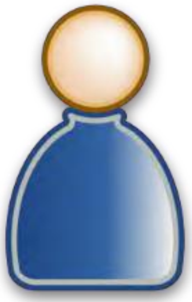
$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$



CueFlik

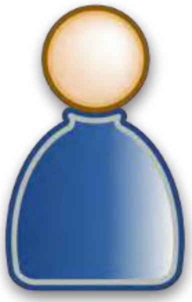


$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$



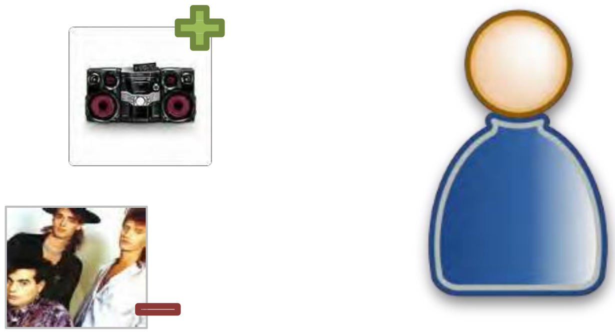
CueFlik

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$



CueFlik

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$



CueFlik

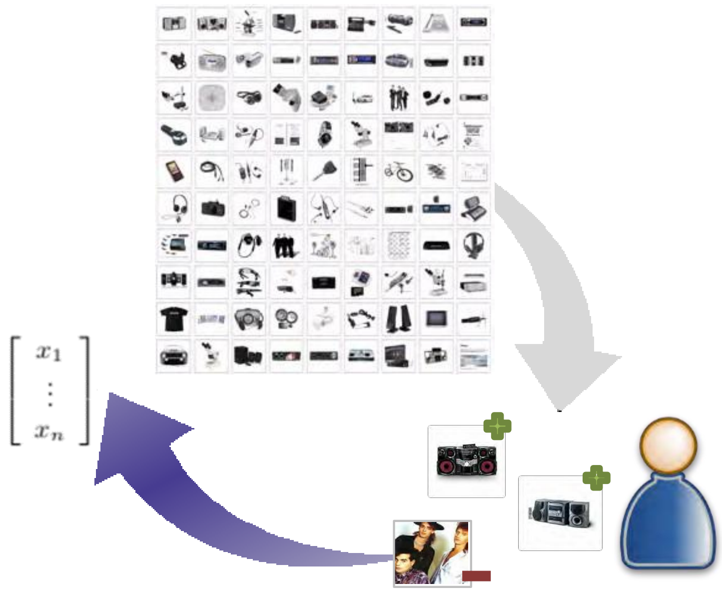
$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

CueFlik

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Defining a Language

CueFlik is interactively defining a language for interaction with data



Using examples to interactively teach **definitions of concepts**

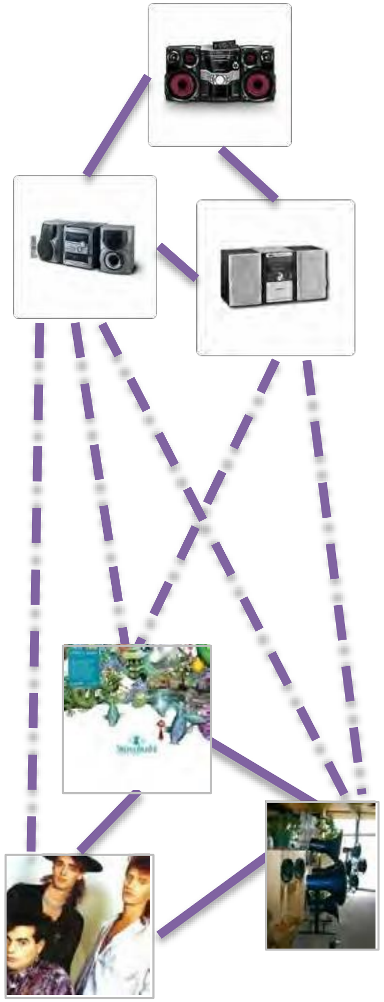
Then apply concepts, such as ‘product photo’, to interact with data in current and future tasks

Defining a Language

CueFlik learned a similarity metric, clustering examples while maximizing separation

We examined aspects of how to design example-based training to ensure agreement on concept

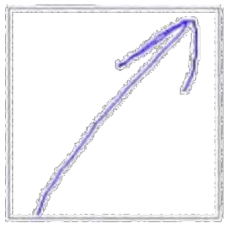
But the language is implicit, conveyed by entirely examples



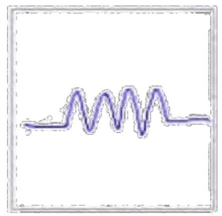
Fogarty, Tan, Kapoor, Winder. CueFlik: Interactive Concept Learning in Image Search. *CHI 2008.*
Amershi, Fogarty, Kapoor, Tan. Overview-Based Example Selection in End-User Interactive Concept Learning. *UIST 2009.*
Amershi, Fogarty, Kapoor, Tan. Examining Multiple Potential Models in End-User Interactive Concept Learning. *CHI 2010.*

GestureScript

Extends example-based demonstration through interaction with the learned representation



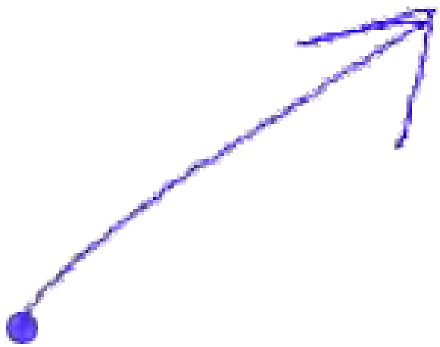
draw(Line)
draw(Head1)



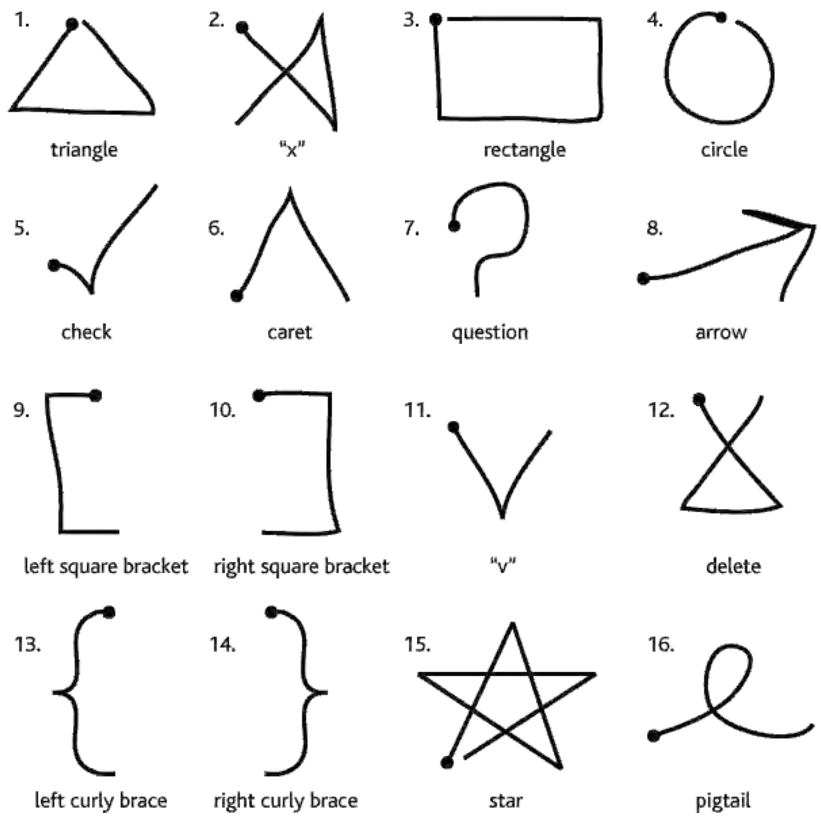
draw(Spring1Head)
repeat
  draw(Cap)
draw(Spring1Tail)

Wobbrock, Wilson, Li. Gestures without Libraries, Toolkits, or Training: A $1 Recognizer for User Interface Prototypes. *UIST 2007.*

Recognizing Symbolic Gestures



Classify

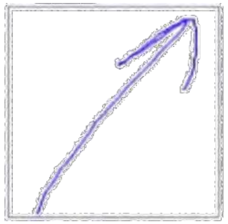


Wobbrock, Wilson, Li. Gestures without Libraries, Toolkits, or Training: A $1 Recognizer for User Interface Prototypes. *UIST 2007.*

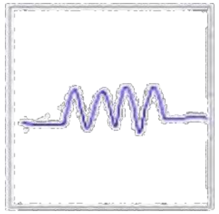
GestureScript

Extends example-based demonstration through interaction with the learned representation

Makes language explicit



draw(Line)
draw(Head1)



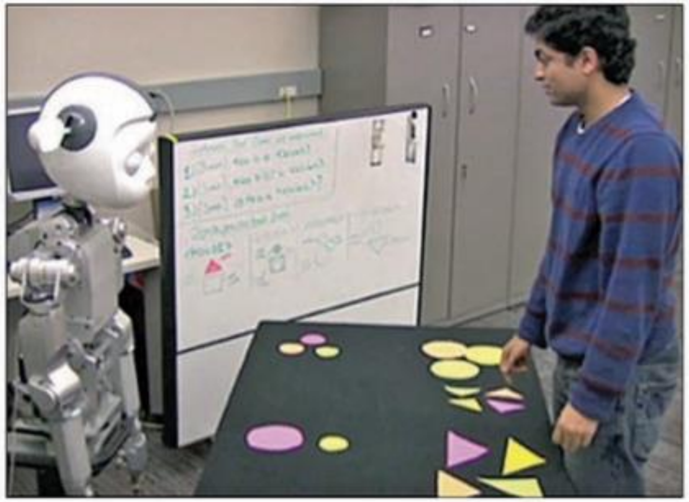
draw(Spring1Head)
repeat
  draw(Cap)
draw(Spring1Tail)

Overview

- Motivation
- Interactive Machine Learning
- Case Studies
  - Ask users to check - Crayon
  - Ask users about "concept“ - CueFlik
  - Ask users to build a model - GestureScript
- **Studying User Interaction with IML**
  - Users are people, not oracles
  - People want to demonstrate how learners should behave
  - People naturally want to provide more than just data labels
  - People value transparency in learning systems
  - Transparency can help people provide better labels
- Takeaways

Studying User Interaction with IML

1. Users are people, not oracles



Passive learning: examples are chosen and presented by the user

Active learning: particular examples are requested by the learner

Although active learning results in faster convergence, users get frustrated from having to answer the learner's long stream of questions and not having control over the interaction

Studying User Interaction with IML

2. People want to demonstrate how learners ***should*** behave
   - Teaching by showing samples (labeling) (visual interactive labeling)
   - Teaching by demonstration
   - Teaching concepts by examples
   - Teaching by reinforcement (via feedback)
   - Testing learner’s state: checking/evaluating student’s learning

Towards Understanding How Humans Teach Robots, UMAP 2011
https://link.springer.com/chapter/10.1007/978-3-642-22362-4_31

Studying User Interaction with IML

3. People naturally want to provide more than just data labels **(flexible, user-centered ways)**
   - Example: a study to understand the types of input end users might provide to if unrestricted by the interface (e.g., text classification of email messages) (Stumpf et al., 2007)
   - People naturally provided a wide variety of input types to improve the classifier's performance
     - suggesting alternative features to use
     - adjusting weights given to different features
     - modifying the information extracted from the text

Develop new ML algorithms that might better support the natural feedback that people want to provide to learners, rather than force users to interact in limited, learner-centered ways

Studying User Interaction with IML

4. People value transparency in learning systems
   - People want to learn about systems

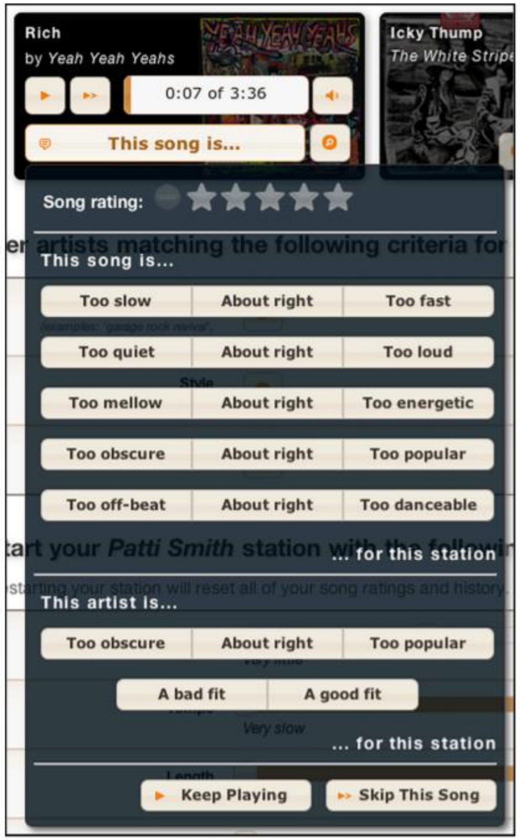


Figure 1. Users could debug by saying *why* the current song was a good or bad choice.

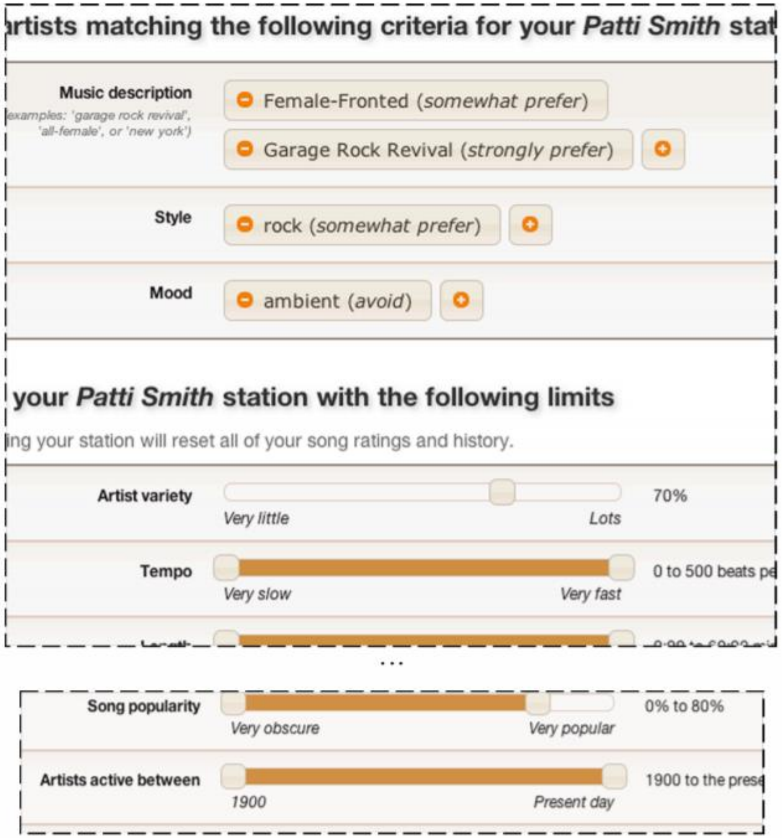


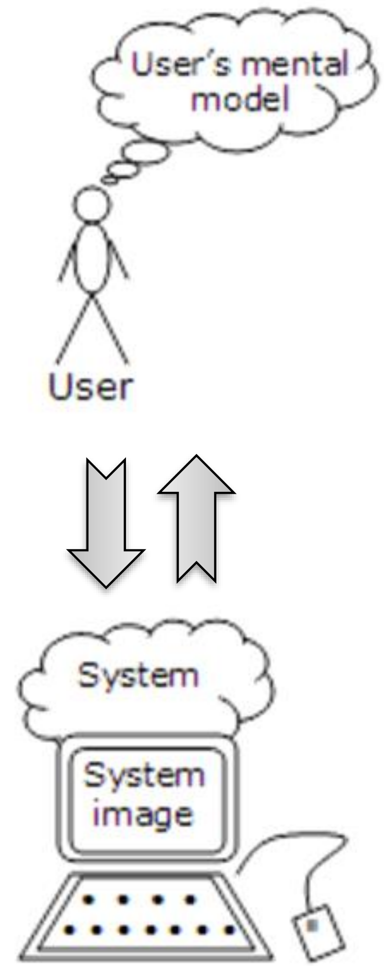
Figure 2. Participants could debug by adding guidelines on the type of music the station should or should not play, via a wide range of criteria.

Kulesza et al. Tell me more?: the effects of mental model soundness on personalizing an intelligent agent, CHI 2012

Studying User Interaction with IML

4. People value transparency in learning systems

- People want to learn about systems
  - Users **positively** evaluated a music recommender system after learning the details about the system (i.e., mental model)
  - The more participants learned about the recommender while interacting with it, the more satisfied they were with the recommender's output
  - Users are not always satisfied by "black box" learning systems
    - Sometimes they want to provide nuanced feedback to steer the system
    - And they are willing and able to learn details about the system to do so



Kulesza et al. Tell me more?: the effects of mental model soundness on personalizing an intelligent agent, CHI 2012

Studying User Interaction with IML

5. Transparency (and explanation) can help people provide better labels

- Improving labeling accuracy by providing contextual features of the sample to be labeled, along with other information (see table)
- Highest labeling accuracy occurred when the system provided sufficient contextual features and current predictions *without uncertainty information*

Dimension	Description	Activity Recognition Example
Uncertainty	Notify labeler that it is uncertain of the label	"Cannot determine your activity."
Amount of Context	Provide varying amounts of contextual information (none, sufficient, extra)	**Sufficient**: "Your feet are leaving the ground." **Extra:** "Your feet are leaving the ground together and repeatedly."
High/Low-Level Context	Give either low (sensor) level context or high (activity) level context	**Low**: "Shaking motion detected." **High**: "Your feet are leaving the ground."
Question	Ask for a label	"What activity are you doing?"
Prediction	Share the expected label for the data	"Prediction: Jumping."
User Feedback	Ask labeler to describe the important features	"How can this action be detected in the future?"

Towards Maximizing the Accuracy of Human-Labeled Sensor Data, Rosenthal and Dey, ACM IUI 2010

Studying User Interaction with IML

5. Transparency (and explanation) can help people provide better labels

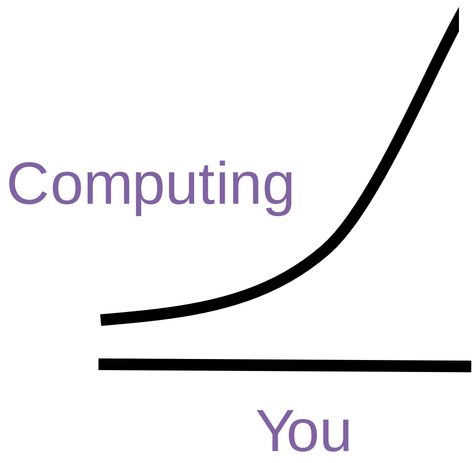
- Transparency at a more social level (why) helps
  - Displaying **the value of users' potential movie ratings** to a broader community in the MovieLens recommendation system (Rashid et al., 2006)
  - Users who were given information about **the value of their contribution to the entire MovieLens community provided more ratings** than those who were not given such information
  - Likewise, those given information about value to a group of users with similar tastes gave more ratings than those given information regarding the full MovieLens community

Overview

- Motivation
- Interactive Machine Learning
- Case Studies
  - Ask users to check - Crayon
  - Ask users about "concept“ - CueFlik
  - Ask users to build a model - GestureScript
- Studying User Interaction with IML
  - Users are people, not oracles
  - People want to demonstrate how learners should behave
  - People naturally want to provide more than just data labels
  - People value transparency in learning systems
  - Transparency can help people provide better labels
- Takeaways

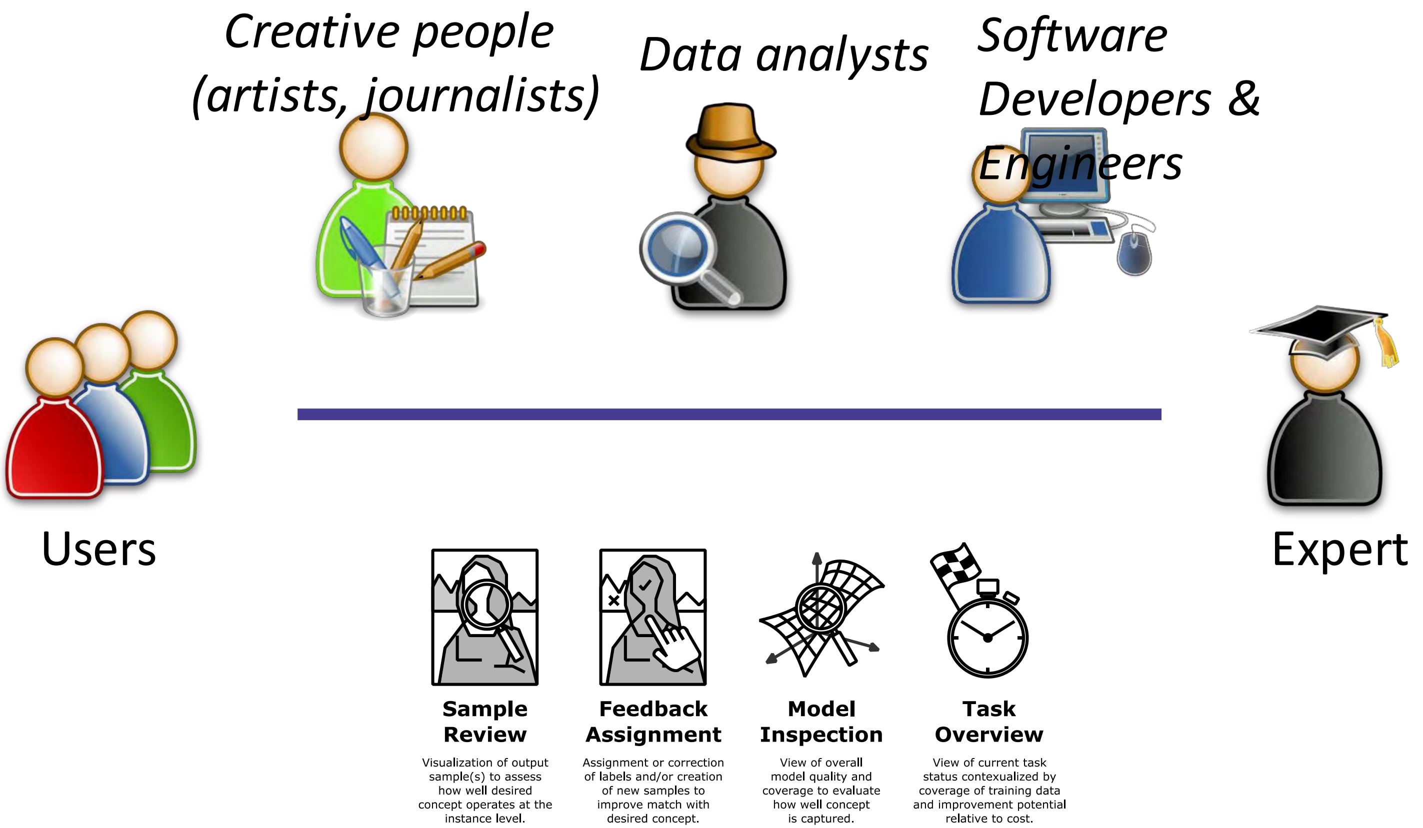
Everyday Interaction with Machine Learning

Machine learning offers a unique tool for scaling human attention to new forms of data, but its complexity and opacity are barriers to effective design for everyday interaction



The human-computer interaction and design community has proven tools, approaches, and theory for addressing this challenge

Everyday Interaction with Machine Learning



Need a range of tools for machine learning, across many domains and levels of expertise