Human Computation Architectures

Swati Mishra

Applications of Machine Learning (4AL3)

Fall 2024

McMaster University
ENGINEERING

Review

- Week 1 : Linear Regression, Gradient Descent
- Week 2: Polynomial Regression, Logistic Regression,
- Week 2-3: Machine Learning Evaluation, Data Pre-Processing
- Week 3: Support Vector Machines

What is Human Computation

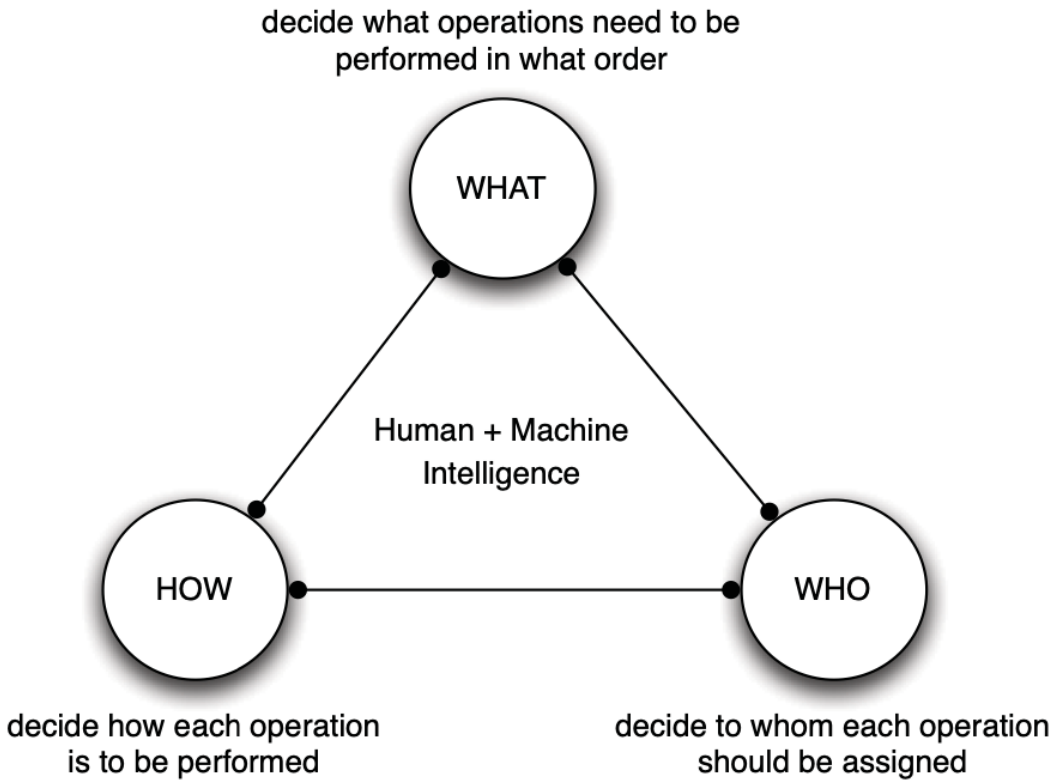
- Computation is the process of mapping some input representation using an explicit, finite set of instructions (i.e. an *algorithm*).

- **Human computation systems** can be defined as intelligent systems that organize humans to carry out the process of computation for performing:
  - the basic operations (or units of computation),
  - taking charge of the control process itself (e.g., decide what operations to execute next or when to halt the program),
  - synthesizing the program itself (e.g., by creating new operations and specifying how they are ordered).

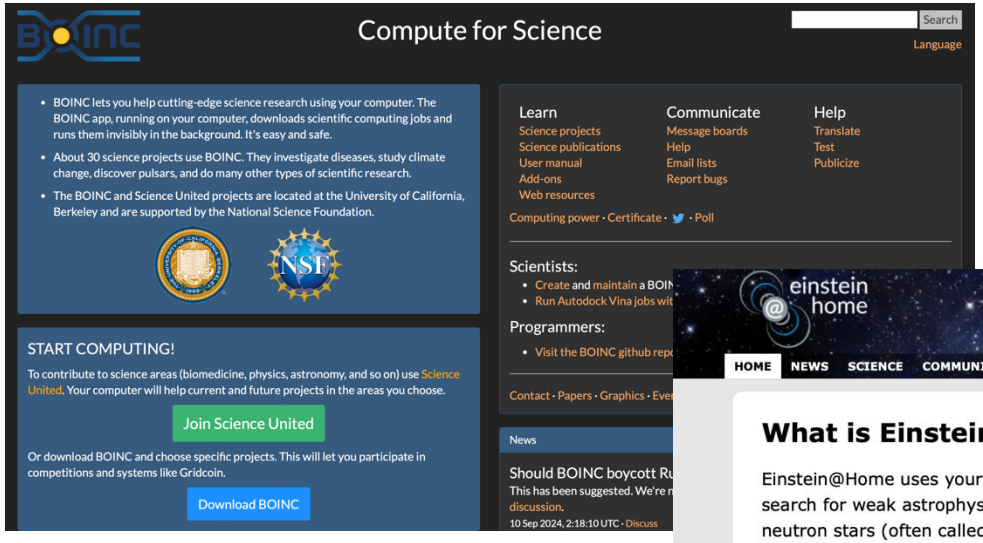
What is Human Computation

- The concept of human computation is often related to:
  - **Crowdsourcing**: The act of outsourcing tasks, traditionally performed by an employee or contractor, to an undefined, large group of people or community (a crowd) through an open call.
  - **Collective Intelligence**: A shared or group intelligence that emerges from the collaboration and com- petition of many individuals and appears in consensus decision making in bacteria, animals, humans and computer networks.
  - **Social Computing** : Technology for supporting social behavior in or through computational systems, e.g., blogs, email, instant messaging, social network services, wikis and social bookmarking. Technology for supporting computations that are carried out by groups of people, e.g., collaborative filtering, online auctions, prediction markets, reputation systems, computational social choice, tagging and verification games.

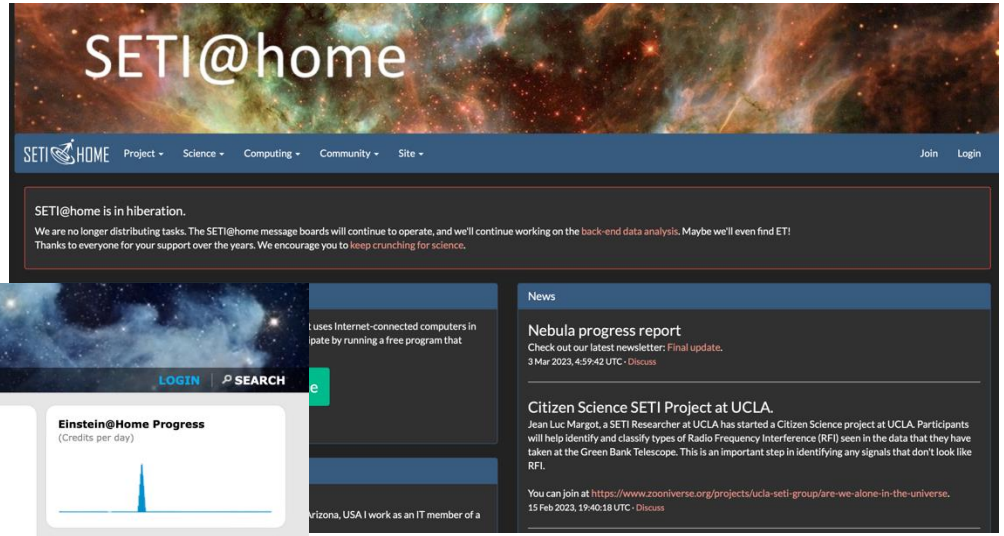
Aspects of Human Computation Systems



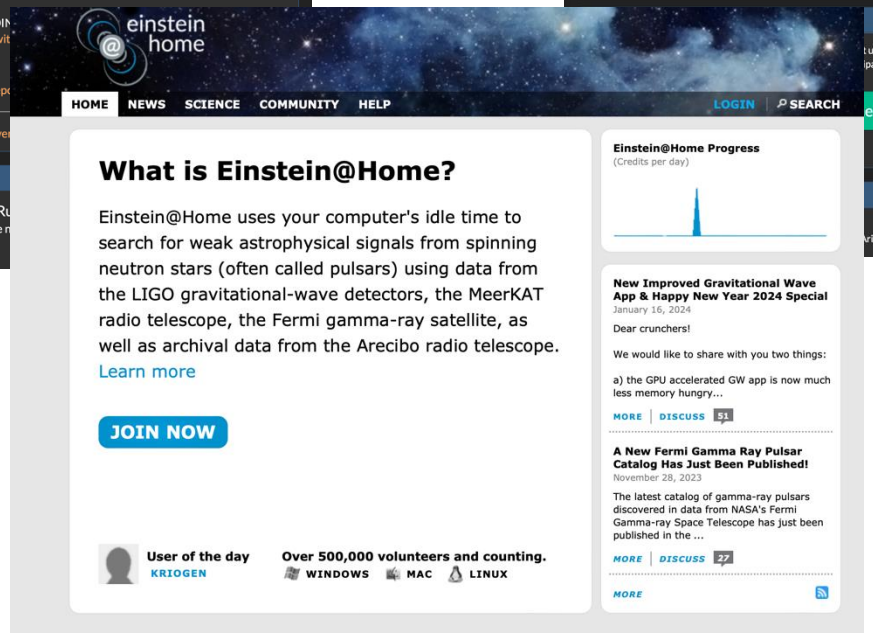
Examples of Human Computation Systems



Bionic



Seti@Home



Einstein@Home

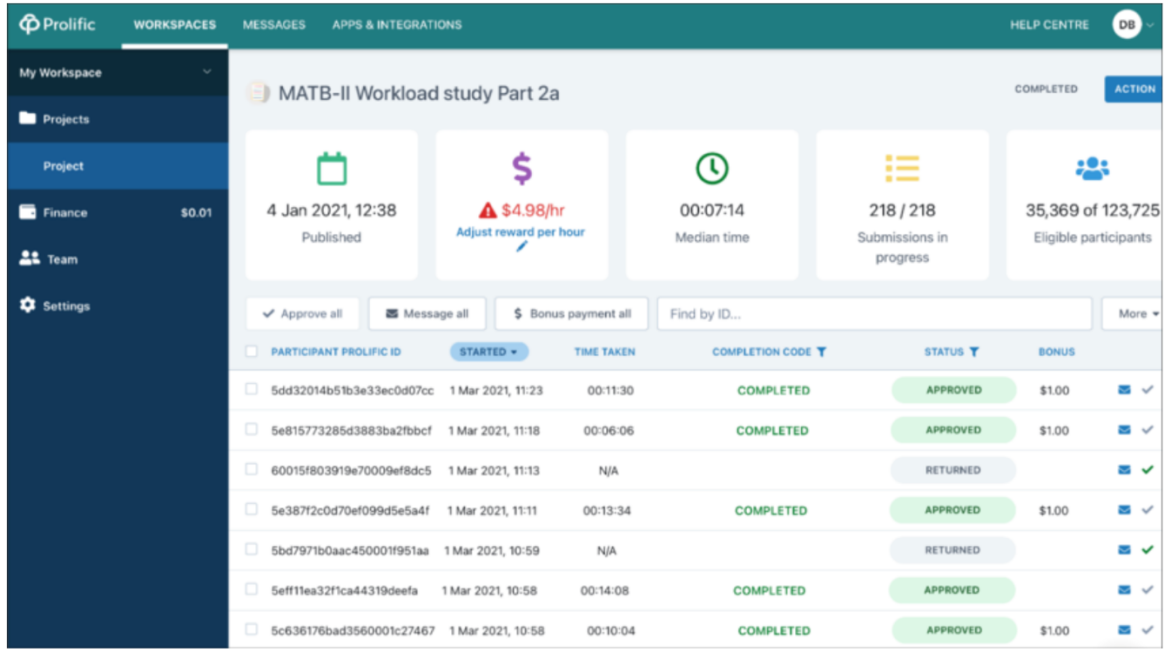
McMaster University

What is Human Computation

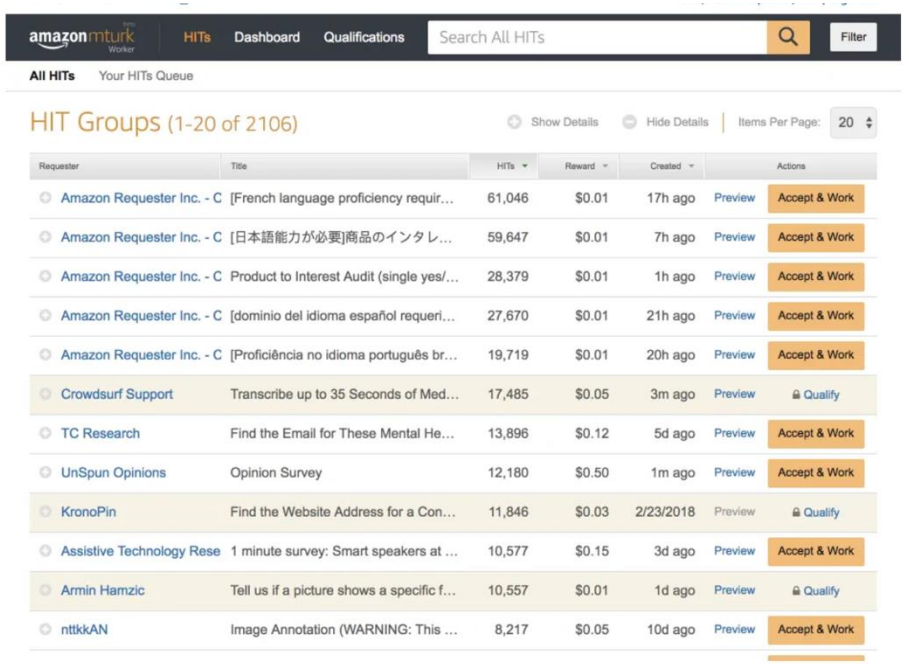
- Human computation: When humans actively decide when and where to collect the data.
  - Example: Playing game with AI

- Not Human Computation: When they have no **conscious** role in determining the outcome.
  - Example: Logging sensor data from phone.

Frameworks for Human Computation

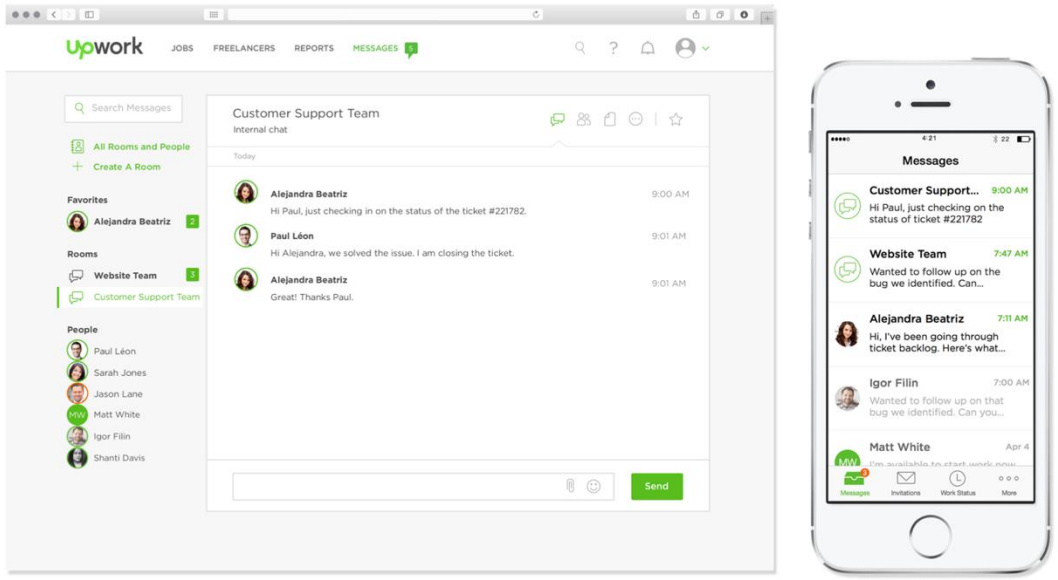


Prolific



Amazon Mechanical Turk

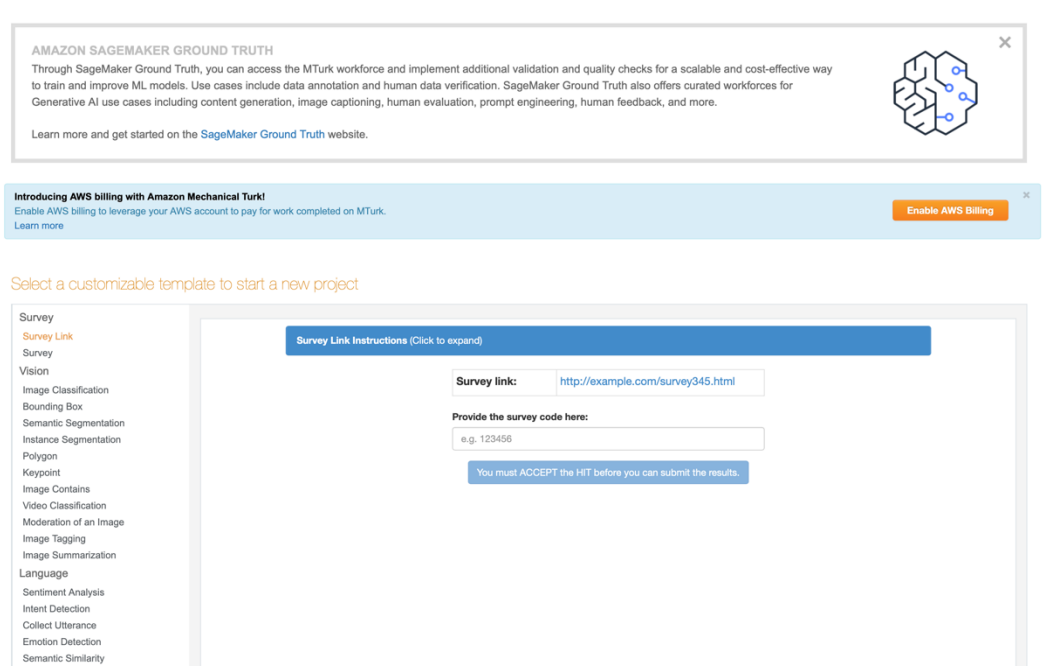
Frameworks for Human Computation

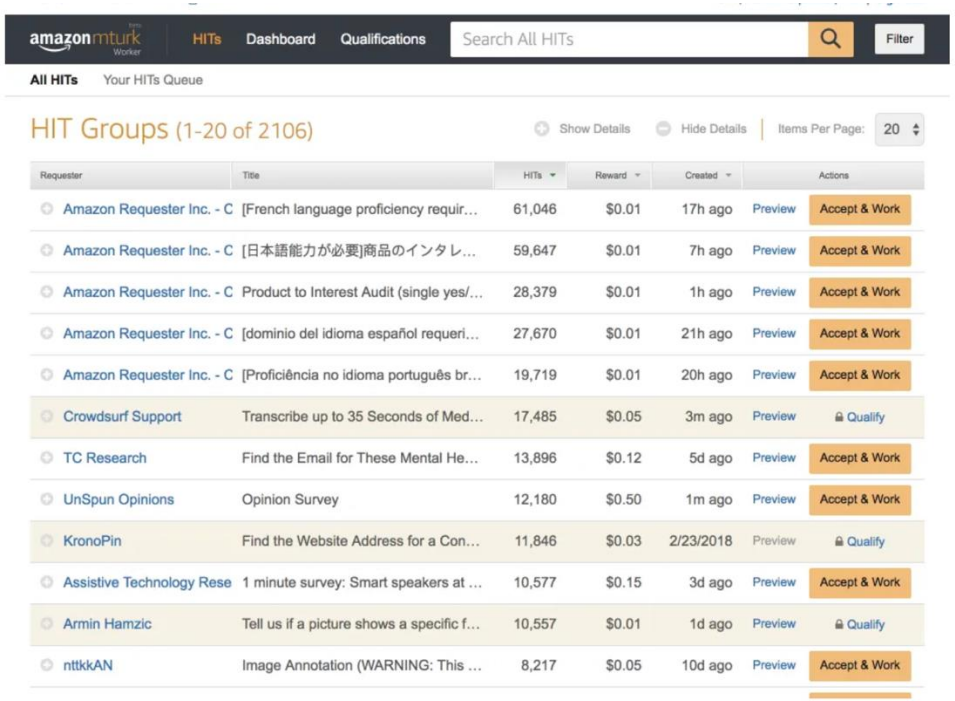


Upwork

Top Coder

Frameworks for Human Computation

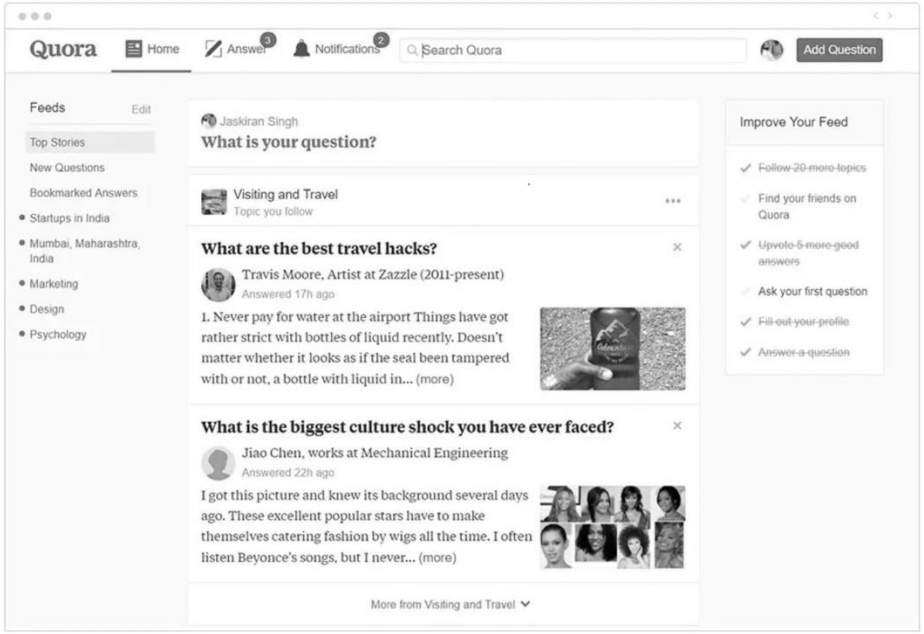




There are two sides of the crowd markets – Workers and Requesters

Wisdom of the Crowds

- It is the collective opinion of a diverse and independent group of individuals rather than that of a single expert.



Wisdom of Crowds

When is wisdom of the crowds beneficial.

- Combinatorial problems such as minimum spanning trees and the traveling salesman problem
- Ordering problems such as the order of the U.S. presidents or world cities by population.
- Multi-armed bandit problems, in which participants choose from a set of alternatives with fixed but unknown reward rates.

Human computation and Machine Learning

Large Scale Visual Recognition Challenge Case Study

IMAGENET

14,197,122 images, 21841 synsets indexed

Home Download Challenges About

Not logged in. Login I Signup

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

Competition

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) evaluates algorithms for object detection and image classification at large scale. One high level motivation is to allow researchers to compare progress in detection across a wider variety of objects -- taking advantage of the quite expensive labeling effort. Another motivation is to measure the progress of computer vision for large scale image indexing for retrieval and annotation.

For details about each challenge please refer to the corresponding page.

- ILSVRC 2017
- ILSVRC 2016
- ILSVRC 2015
- ILSVRC 2014
- ILSVRC 2013
- ILSVRC 2012
- ILSVRC 2011
- ILSVRC 2010

Workshop

Every year of the challenge there is a corresponding workshop at one of the premier computer vision conferences. The purpose of the workshop is to present the methods and results of the challenge. Challenge participants with the most successful and innovative entries are invited to present. Please visit the corresponding challenge page for workshop schedule and information.

Download

The most popular challenge is the ILSVRC 2012-2017 image classification and localization task. It is available on Kaggle. For all other data please log in or request access.

Evaluation Server

The evaluation server can be used to evaluate image classification results on the test set of ILSVRC 2012-2017. Please see here for our submission policy. Importantly, you should not make more than 2 submissions per week.

Updates

- October 10, 2019: The ILSVRC 2012 classification and localization test set has been updated. The Kaggle challenge and our download page both now contain the updated data.
- June 2, 2015: Follow-up update regarding status of the server
- May 19, 2015: Annoucement regarding the submission server

Citation

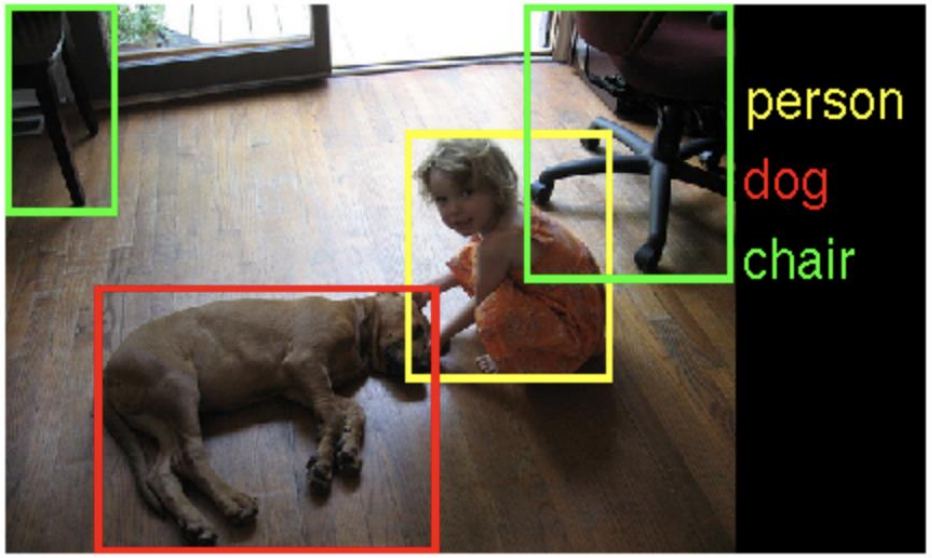
When reporting results of the challenges or using the datasets, please cite:

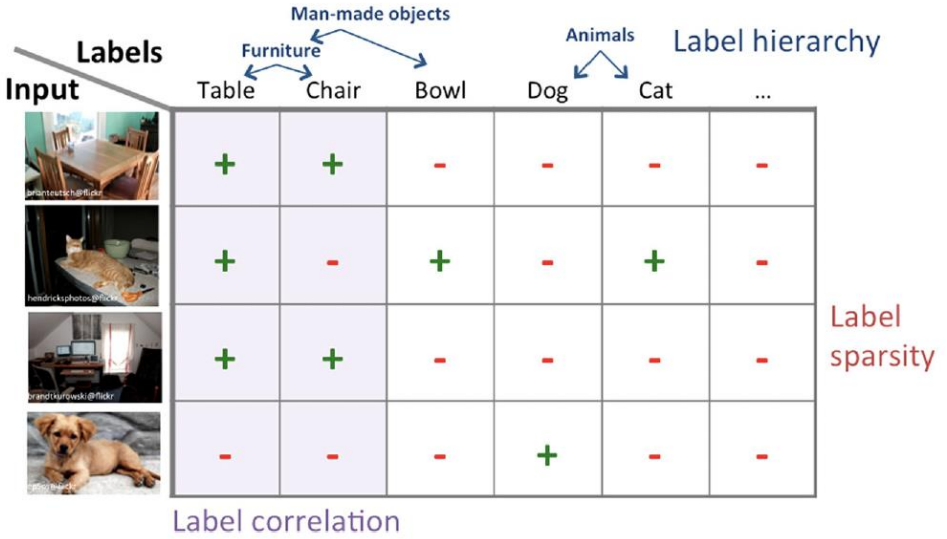
Olga Russakovsky*, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (* = equal contribution) **ImageNet Large Scale Visual Recognition Challenge**. *IJCV*, 2015. paper I bibtex I paper content on arxiv I attribute annotations

McMaster University

Human computation and Machine Learning

Large Scale Visual Recognition Challenge Case Study





Annotation task required workers to identify various objects in the image

ILSVRC2010 challenge's wining entry use stochastic SVM!

Human computation and Machine Learning

Large Scale Visual Recognition Challenge Case Study

~ 2 million images

IMAGENET

14,197,122 images, 21841 synsets indexed

Home Download Challenges About

Not logged in. Login I Signup

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

Competition

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) evaluates algorithms for object detection and image classification at large scale. One high level motivation is to allow researchers to compare progress in detection across a wider variety of objects -- taking advantage of the quite expensive labeling effort. Another motivation is to measure the progress of computer vision for large scale image indexing for retrieval and annotation.

For details about each challenge please refer to the corresponding page.

- ILSVRC 2017
- ILSVRC 2016
- ILSVRC 2015
- ILSVRC 2014
- ILSVRC 2013
- ILSVRC 2012
- ILSVRC 2011
- ILSVRC 2010

Workshop

Every year of the challenge there is a corresponding workshop at one of the premier computer vision conferences. The purpose of the workshop is to present the methods and results of the challenge. Challenge participants with the most successful and innovative entries are invited to present. Please visit the corresponding challenge page for workshop schedule and information.

Download

The most popular challenge is the ILSVRC 2012-2017 image classification and localization task. It is available on Kaggle. For all other data please log in or request access.

Evaluation Server

The evaluation server can be used to evaluate image classification results on the test set of ILSVRC 2012-2017. Please see here for our submission policy. Importantly, you should not make more than 2 submissions per week.

Updates

- October 10, 2019: The ILSVRC 2012 classification and localization test set has been updated. The Kaggle challenge and our download page both now contain the updated data.
- June 2, 2015: Follow-up update regarding status of the server
- May 19, 2015: Annoucement regarding the submission server

Citation

When reporting results of the challenges or using the datasets, please cite:

Olga Russakovsky*, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (* = equal contribution) **ImageNet Large Scale Visual Recognition Challenge**. *IJCV*, 2015. paper I bibtex I paper content on arxiv I attribute annotations

McMaster University

Human computation and Machine Learning

Human Computation tasks in Machine Learning Applications:

- Generating Binary or Categorical Labels
  - Example: a binary label indicating whether or not the website contains profanity.
  - Popular model to implement human computation architecture in this setting is Dawid-Skene model that assumes instances are homogeneous
  - A worker's probability of labeling any given instance correctly is controlled by one or more worker-specific quality parameters.

N = Number of computational tasks.
Y_n = *true* output each task is unknown.

Our goal is to estimate Y_n given an output matrix O, containing the responses from M workers.

AGGREGATING OUTPUTS

$$\begin{array}{cc} & \text{computational task} \\ & \begin{array}{cccc} 1 & 2 & \cdots & N \end{array} \\ \text{worker} \begin{array}{c} 1 \\ 2 \\ \vdots \\ M \end{array} & \begin{bmatrix} O_{11} & O_{12} & \cdots & O_{1N} \\ O_{21} & O_{22} & \cdots & O_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ O_{M1} & O_{M2} & \cdots & O_{MN} \end{bmatrix} \end{array}$$

Human computation and Machine Learning

Human Computation tasks in Machine Learning Applications:

- Generating Binary or Categorical Labels
  - Example: a binary label indicating whether or not the website contains profanity.

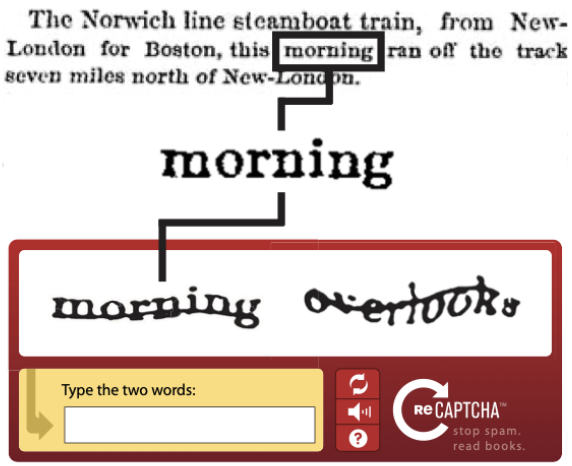
Algorithm 1 The basic EM framework of Dawid and Skene (1979).

Input: Sets of worker-generated labels for each instance
Initialize each instance's label based on a simple majority vote
repeat
  for all Workers w **do**
    Calculate w's quality parameter(s), treating each instance's current label as ground truth
  end for
  for all Instances i **do**
    Calculate the most likely label for i, treating each worker's approximated quality parameter(s) as ground truth
  end for
until Label assignments have converged
Output: The current label assignments for each instance

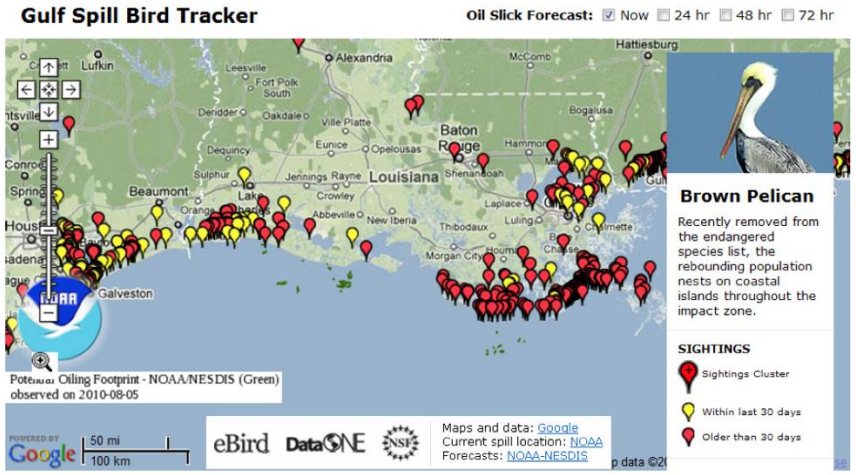
Human computation and Machine Learning

Human Computation tasks in Machine Learning Applications:

- Generating Binary or Categorical Labels
- Generating Transcriptions, Translations, and Image Annotations
  - Example: ReCAPTCHA, Citizen Science



CAPTCHA



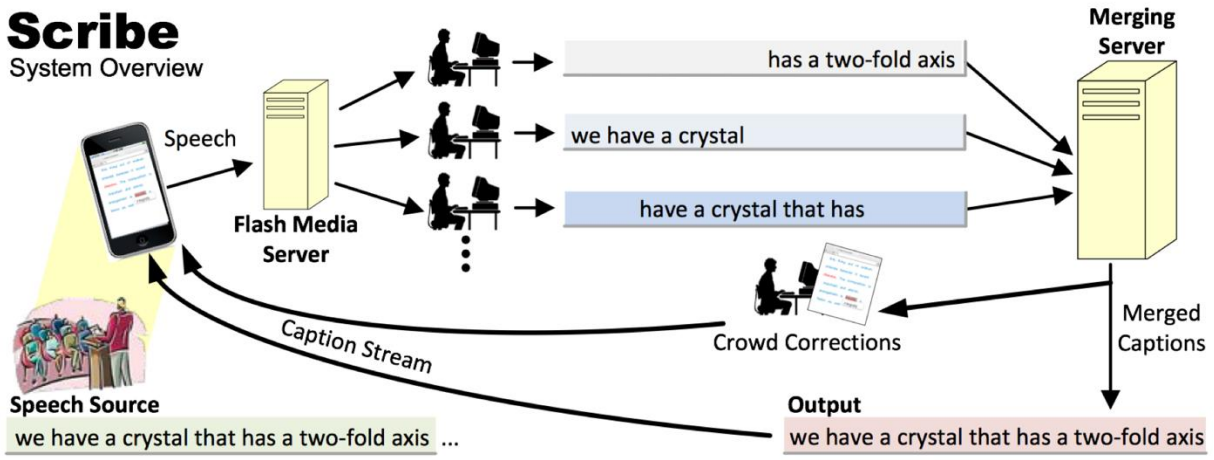
Bird sightings of birds during the Gulf Spill.

McMaster University

Human computation and Machine Learning

Human Computation tasks in Machine Learning Applications:

- Generating Binary or Categorical Labels
- Generating Transcriptions, Translations, and Image Annotations
  - Example: ReCAPTCHA, Citizen Science



The architecture of Scribe Lasecki et al. (2012).

Human computation and Machine Learning

Human Computation tasks in Machine Learning Applications:

- Generating Binary or Categorical Labels
- Generating Transcriptions, Translations, and Image Annotations
- Evaluating and Debugging Models
  - Crowdsourced evaluation is especially common for unsupervised models, which generally cannot be evaluated in terms of simple metrics like accuracy or precision because there is no objective notion of ground truth.

Human computation and Machine Learning

Human Computation tasks in Machine Learning Applications:

- Asking the right question.
  - Influence of presented information on task performance.
  - Task Granularity:
    - Well defined
    - Not cognitively overwhelming
  - Task should be done independently
  - Incentive should be adequate
  - Outcome must have quality control

BubbleView (2017)

Human computation and Machine Learning

Human Computation tasks in Machine Learning Applications:

- Asking the right question.
  - Influence of presented information on task performance.
  - Task Granularity:
    - Well defined
    - Not cognitively overwhelming
  - Task should be done independently
  - Incentive should be adequate
  - Outcome must have quality control

BubbleView (2017)

Readings

***Reference Material*:**

- Human Computation, Edith Law, Luis von Ahn
- Making Better Use of the Crowd: How Crowdsourcing Can Advance Machine Learning Research, Jennifer Wortman Vaughan, Microsoft Research