Human computation: how we can harness the power of human intelligence

Yukino Baba (Kyoto University)

The 39th Annual Meeting of the Japan Neuroscience Society, Symposium S3-B-3: Data-driven approaches to complex brain functions

2016-07-22
Human computation is an important tool for data-driven science

- Data-driven approaches require a large amount of human contributions
  - Data annotation
  - Feature engineering
  - Predictive modeling

- Good coordination of people is a key of human computation
  - Response aggregation
  - Expert discovery
Combination of humans and computers for solving large-scale complex problems

- Human computation aims to solve problems that neither can be solved by humans nor computers
- Example: large-scale book digitizing
  - **Computers** cannot read complex characters but can assign tasks to a large group of **humans**
Crowds support book digitizing

- To confirm that the users are not robots, ReCAPTCHA prompts users to read two distorted words:
  - (1) word that generated automatically
  - (2) word that computers cannot read

APPLICATION: ReCAPTCHA

The Norwich line steamboat train, from New-London to Boston, this morning ran off the track seven miles north of New-London.

For digitizing book

For filtering robots

http://irevolution.net/2013/06/17/recaptcha-for-disaster-response/
Crowds support image annotation

- ImageNet hires crowdsourcing workers to create a large-scale annotated image datasets
  - 14M images, 20K categories

---

Crowds support detailed image annotation

- Visual Genome hires crowdsourcing workers to annotate detailed information to images.

**DATA ANNOTATION: Visual Genome**

Good coordination of people is a key for harnessing the wisdom of crowds

- We cannot assume all the people are reliable or have enough skills to solve the given tasks
- Solutions using machine learning:

1. Response aggregation
2. Expert discovery
Machine learnings estimates reliability of each person without knowing the answer

- Naïve aggregation method: majority voting
- Statistical aggregation method: jointly estimates worker reliabilities and true answers

Task

“Is a bird appeared in the given image?”

<table>
<thead>
<tr>
<th>Truth</th>
<th>Predicted reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

Machine learning aggregates the ratings considering author and reviewer reliabilities

- 2-stage procedure: **authors** create outputs and **reviewers** rate the outputs
- Statistical method jointly estimates **author and reviewer reliabilities** and **true qualities of outputs**

Y. Baba et al.: Statistical Quality Estimation for General Crowdsourcing Tasks, In *KDD’13.*
Experts are identified via Google AdWords

- Advertising system automatically identifies user groups that are good matches for a given task.

The system learns that users who frequently access sites such as "HealthLine" are experts of health information.

Link to a task is shown like an ad.

Experts are identified based on their attributes

- Attribute of workers (e.g., demographics or education) can be related to worker expertise
- Machine learning method discovers related worker attributes for a given task

The system learns "Major=science" and "Gender=female" are the target group

H. Li et al., The Wisdom of minority: discovering and targeting the right group of workers for crowdsourcing, In WWW, 2014.
Crowds support feature extraction

- Feature extraction is a key process for applying machine learning methods successfully.
- Humans may capture abstract features that cannot be calculated by computers.
- Visual 20 question game: human response for predefined questions are used as features.

<table>
<thead>
<tr>
<th>Hard for humans</th>
<th>Easy for humans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finch? Bunting?...</td>
<td>Yellow Belly? Blue Belly?...</td>
</tr>
<tr>
<td><img src="image1.png" alt="Bird Image" /></td>
<td><img src="image2.png" alt="Bird Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>x1</th>
<th>x2</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>...</td>
</tr>
</tbody>
</table>

Crowds support feature engineering

- Flock: asks crowds to generate questions
  - Crowds are shown a pair of positive and negative examples and tell the difference

1. Task: “is this person lying?”

2. Prediction task: Is this person lying?
   - Crowd nominates features using examples

3. Crowd labels nominated features
   - Crowd labels nominated features
   - (Optional: add user-provided/off-the-shelf features)

4. Learn hybrid model using crowd and machine features
   - A classifier is trained


Flock: Hybrid Crowd-Machine Learning Classifiers

FEATURE EXTRACTION: Flock

- Flock: asks crowds to generate questions
  - Crowds are shown a pair of positive and negative examples and tell the difference
Crowds explore model spaces to find better predictive model

- Predictive modeling competition (e.g., kaggle)
  - Platform for leveraging **crowds of data scientists**
  - Participants build predictive models and compete for monetary rewards
Crowds of data scientists defeat experts

- Crowd of data scientists achieved far better results than “experts” in a short term

Baseline built by experts
Overtaken in 4 days

20% performance gain
Competition using a DNA sequence dataset is now open on our platform

http://universityofbigdata.net
Human computation is an important tool for data-driven science

- Data-driven approaches require a large amount of human contributions
  - Data annotation
  - Feature engineering
  - Predictive modeling

- Good coordination of people is a key of human computation
  - Response aggregation
  - Expert discovery