CrowdTruth.org

Machine-Human Computation for Harnessing Disagreement in Semantic Interpretation

Oana Inel, Khalid Khamkham, Tatiana Cristea, Anca Dumitrache, Arne Rutjes, Jelle v.d Ploeg, Lukasz Romaszko, Lora Aroyo, Robert-Jan Sips
Importance of Human Annotation

• Machines learn & are evaluated from examples

• Semantic interpretation of data is needed in all sciences

• Humans analyze examples and annotate them for the “correct” interpretation
HUMAN DISAGREEMENT IS ESSENTIAL IN HELPING MACHINES WITH SEMANTIC INTERPRETATION
disagreement can reflect
the degree of clarity in a sentence

Does each sentence express the TREAT relation?

**ANTIBIOTICS** are the first line treatment for indications of **TYPHUS**.
→ 95%

Patients with **TYPHUS** who were given **ANTIBIOTICS** exhibited side-effects.
→ 80%

*With **ANTIBIOTICS** in short supply, DDT was used during WWII to control the insect vectors of **TYPHUS**.*
→ 50%

Lora Aroyo @laroyo
GADOLINIUM agents are useful for patients with renal impairment, but in patients with severe renal failure requiring dialysis it presents a risk of nephrogenic systemic FIBROSIS.

What is the RELATION between the highlighted terms?

GADOLINIUM agents are useful for patients with renal impairment, but in patients with severe renal failure requiring dialysis it presents a risk of nephrogenic systemic FIBROSIS.

CAUSE? or SIDE EFFECT?

70%  45%
disagreement can indicate low quality workers

Does each sentence express the **TREAT** relation?

- **S1**: **ANTIBIOTICS** are the first line treatment for indications of **TYPHUS**.
- **S2**: **QUININE** is not a reliable cure for **MALARIA**.

<table>
<thead>
<tr>
<th>Worker</th>
<th>S1</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worker 1</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Worker 2</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Worker 3</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>Worker 4</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td><strong>Worker 5</strong></td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
CrowdTruth Software Components: Machines & Crowds Workflow

- **Machine Pre-processing**: optimizing crowdsourcing
- **Micro-task Templates**: reuse & optimization
- **CrowdTruth Analytics**: disagreement-based metrics

- **Novel approach** to ground truth data collection & evaluation
- **PROV** for tracking versions of data
- Reusability in **variety of annotation tasks & domains** with text, image, video (thinking about sound)
CrowdTruth Software

• Open source: https://github.com/CrowdTruth
• Web service: http://stable.crowdtruth.org

Human-Machine Computing Workflow
Training & Evaluation of Text, Images & Video Data

Lora Aroyo    @laroyo
CrowdTruth Software: Crowdsourcing Job Analytics

Overview of 6 Jobs
Overview of Jobs having Type like "RegEx", Workers > 1
Select an area to zoom. To see detailed information select individual jobs. Right click for table view. From legend select features. Adjust Y-Axis by dragging the labels (double click to return to default).

- # of inconsistent quality workers
- # of low quality workers
- # of high quality workers
- avg # of workers
- # of low quality judgements
- # of high quality judgements
- avg # of judgements
- # of unclear units
- # of clear units
- avg # of units
- job duration
- payment
- avg unit clarity
- avg # of units per worker
- avg worker agreement