Crowdsourcing

Collective Intelligence, Judgment Analysis, Citizen Science

Malay Bhattacharyya

Assistant Professor
Machine Intelligence Unit
Indian Statistical Institute, Kolkata

Guest Lecture at IIT, Kharagpur (October 04, 2018)

Collective Intelligence

Collective intelligence is a form of intelligence that emerges from the collaboration, collective efforts, and competition of many individuals [Surowiecki, 2004].

Collective Intelligence

Collective intelligence is a form of intelligence that emerges from the collaboration, collective efforts, and competition of many individuals [Surowiecki, 2004].

Types of wisdom of the crowd:

- Cognition: How faster and more reliable judgments can be made through consensus decision making from crowdsourced opinions.
- Coordination: How the crowd can be organized so as to enable them to effectively work together.
- Cooperation: How the crowd can form networks of trust.

Collective Intelligence

Actors:

- Requester: Individuals or group for whom work is done or who takes the responsibility to aggregate work.
- Worker: Someone who contributes.

Collective Intelligence

Actors:

- Requester: Individuals or group for whom work is done or who takes the responsibility to aggregate work.
- Worker: Someone who contributes.

Characteristics:

1. Directed: An individual recruits and guides a group of people to accomplish a goal.
2. Collaborative: A group gathers based on shared interest or goal and pursues a related goal together.
3. Passive: The crowd or collective may never meet or coordinate, but it is still possible to mine their collective behavior patterns for information.

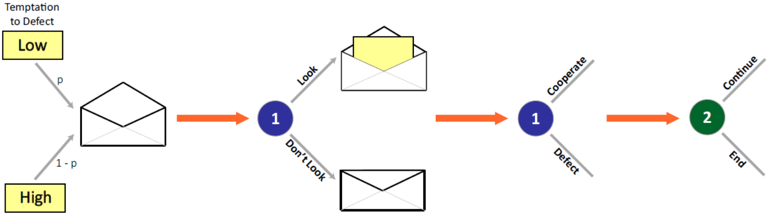
Collective Intelligence

There are several ingredients of the crowd wisdom [Woolley, 2010].

- Diversity
- Independence
- Decentralization
- Aggregation

Collective Intelligence

People who avoid looking at the costs of good acts can be trusted to cooperate in important situations, whereas those who look cannot. We find that evolutionary dynamics can lead to cooperation without looking at costs [Hoffman, 2015].



The framework of envelope game

Basics of judgment analysis

- **Opinion:** The label (annotation) marked by an annotator (basically a crowd worker) for a question.
- **Judgment:** The predicted label from a list of such opinions.
- **Gold judgment:** The actual opinion (solution) for the question.

Basics of judgment analysis

- **Opinion:** The label (annotation) marked by an annotator (basically a crowd worker) for a question.
- **Judgment:** The predicted label from a list of such opinions.
- **Gold judgment:** The actual opinion (solution) for the question.

An annotation process can be formally represented by a quadruplet $\langle Q, A, O, \tau \rangle$, which consists of the following:

- A finite set of annotators $A = \{A_1, A_2, \ldots, A_n\}$,
- A finite set of questions $Q = \{Q_1, Q_2, \ldots, Q_m\}$,
- A finite set of opinions $O = \{O_1, O_2, \ldots, O_k\}$,
- A mapping function $\tau : Q \times A \rightarrow O$.

Basics of judgment analysis

- **Opinion:** The label (annotation) marked by an annotator (basically a crowd worker) for a question.
- **Judgment:** The predicted label from a list of such opinions.
- **Gold judgment:** The actual opinion (solution) for the question.

An annotation process can be formally represented by a quadruplet $\langle Q, A, O, \tau \rangle$, which consists of the following:

- A finite set of annotators $A = \{A_1, A_2, \ldots, A_n\}$,
- A finite set of questions $Q = \{Q_1, Q_2, \ldots, Q_m\}$,
- A finite set of opinions $O = \{O_1, O_2, \ldots, O_k\}$,
- A mapping function $\tau : Q \times A \rightarrow O$.

Problem statement: Given a set of opinions obtained from the crowd for a given question, predict (after combining together) the 'gold' judgment.

Crowdsourced opinions

	Question 1	Question 2	⋯	Question m
Annotator 1	Y	—	⋯	N
Annotator 2	N	Y	⋯	Y
⋮	⋮	⋮	⋱	⋮
Annotator n	Y	U	⋯	—

An example of annotation matrix

Let us assume that ‘Y’, ‘N’ and ‘U’ denote the positive, negative and uncertain responses, respectively, whereas ‘-’ denote that no response was given.

Majority voting

Majority voting is a standard approach of prioritizing options based on the number of supports (votes).

Majority voting

Majority voting is a standard approach of prioritizing options based on the number of supports (votes).

	Question 1	**Question 2**	**Question 3**	**Question 4**
Annotator 1	-	U	Y	N
Annotator 2	U	-	U	N
Annotator 3	Y	N	-	-
Annotator 4	Y	N	U	N
Annotator 5	-	Y	N	N
Annotator 6	Y	-	N	N
Annotator 7	N	N	Y	-
Majority	Y	N	?	N

Majority voting

Majority voting is a standard approach of prioritizing options based on the number of supports (votes).

	Question 1	Question 2	Question 3	Question 4
Annotator 1	-	U	Y	N
Annotator 2	U	-	U	N
Annotator 3	Y	N	-	-
Annotator 4	Y	N	U	N
Annotator 5	-	Y	N	N
Annotator 6	Y	-	N	N
Annotator 7	N	N	Y	-
Majority	Y	N	?	N

Note: The ties are resolved arbitrarily.

Semantic majority voting

The ties may have separate interpretations due to the dependence of the opinions.

	Yes	**No**	**Unsure**
Yes	x	Unsure	Yes
No	Unsure	x	No
Unsure	Yes	No	x

Semantic majority voting

The ties may have separate interpretations due to the dependence of the opinions.

	Yes	**No**	**Unsure**
Yes	x	Unsure	Yes
No	Unsure	x	No
Unsure	Yes	No	x

Note: The aforementioned pairwise relations can be extended to k-wise relations also.

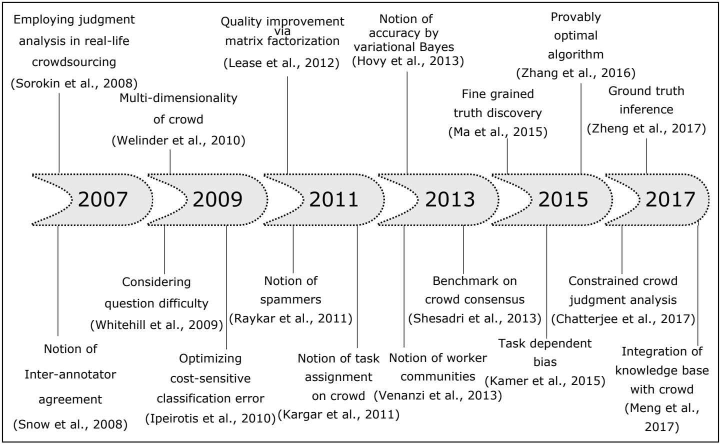
Weighted majority voting

Assigning weights to the annotators might be useful.

W1	dog	dog	dog	cat	dog	cat
W2	cat	cat	cat	cat	cat	cat
W3	dog	dog	dog	dog	hotdog	dog
W4	dog	librarian	shark	cat	dog	LION!
W5	dog	cat	shark	dog	dog	cat

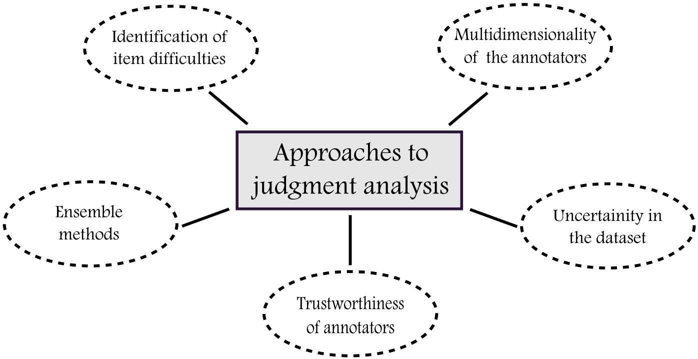
How can we identify a spammer???

Timeline


Employing judgment analysis in real-life crowdsourcing (Sorokin et al., 2008)
Multi-dimensionality of crowd (Welinder et al., 2010)
Quality improvement via matrix factorization (Lease et al., 2012)
Notion of accuracy by variational Bayes (Hovy et al., 2013)
Fine grained truth discovery (Ma et al., 2015)
Provably optimal algorithm (Zhang et al., 2016)
Ground truth inference (Zheng et al., 2017)
2007
2009
2011
2013
2015
2017
Notion of Inter-annotator agreement (Snow et al., 2008)
Considering question difficulty (Whitehill et al., 2009)
Optimizing cost-sensitive classification error (Ipeirotis et al., 2010)
Notion of spammers (Raykar et al., 2011)
Notion of task assignment on crowd (Kargar et al., 2011)
Notion of worker communities (Venanzi et al., 2013)
Benchmark on crowd consensus (Shesadri et al., 2013)
Task dependent bias (Kamer et al., 2015)
Constrained crowd judgment analysis (Chatterjee et al., 2017)
Integration of knowledge base with crowd (Meng et al., 2017)

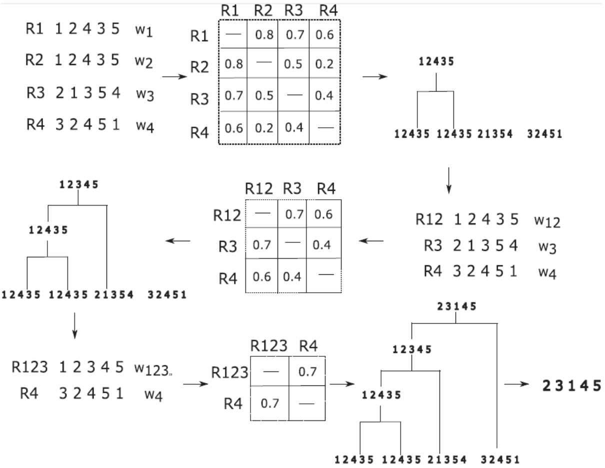
Approaches to judgment analysis

Judgment analysis is to be done on the opinions of the crowd.



Different approaches taken up by judgment analysis algorithms

Independent judgment analysis



Independent judgment analysis

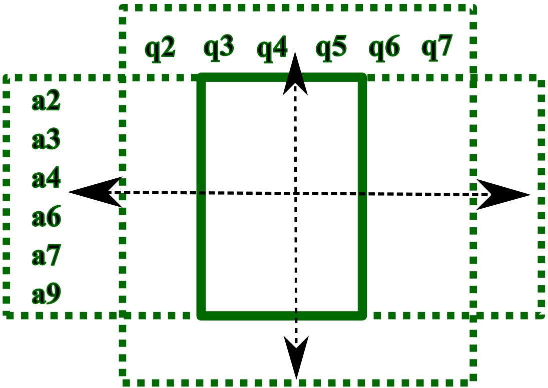
- Gain score: Due to its precedence of position with respect to another object in a particular ranking.
- Penalty score: Due to its position lagging behind in a particular ranking.

Independent judgment analysis

- Gain score: Due to its precedence of position with respect to another object in a particular ranking.
- Penalty score: Due to its position lagging behind in a particular ranking.

1. The competence of the annotators is measured based on their opinions.
2. Depending on various performance scores, the annotators are ordered.
3. Apply weighted rank aggregation approach on these multiple rankings to obtain a final aggregated ranking.
4. Use the aggregated ranking to decide the accuracy of the annotators.

Judgment analysis at scale



Judgment analysis at scale

The probability of having opinion $\hat{o}$ for the i^{th} question when 1^{st} order neighborhood is considered is

$$P^1_{(q_i=\hat{o})} = \frac{\sum\limits_{j=1,q_i\in B^j} \sum\limits_{a_x\in R(q_i=\hat{o})} \mathcal{A}(a_x) + \frac{c_{Bj}}{r_{Bj}}}{|R(q_i)|}.$$

Judgment analysis at scale

The probability of having opinion $\hat{o}$ for the i^{th} question when 1^{st} order neighborhood is considered is

$$P^1_{(q_i=\hat{o})} = \frac{\sum\limits_{j=1, q_i \in B^j} \sum\limits_{a_x \in R(q_i=\hat{o})} \mathcal{A}(a_x) + \frac{c_{Bj}}{r_{Bj}}}{|R(q_i)|}.$$

Here, $R(q_i)$ denotes the responses given by annotators on the i^{th} question. $R(q_i = \hat{o})$ describes the annotators whose opinion is $\hat{o}$ for the i^{th} question.

Judgment analysis at scale

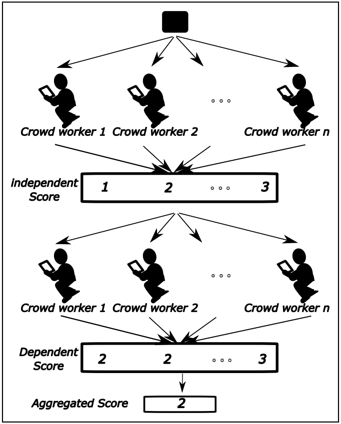
The probability of having opinion $\hat{o}$ for the i^{th} question when 1^{st} order neighborhood is considered is

$$P^{1}_{(q_i=\hat{o})} = \frac{\sum\limits_{j=1, q_i \in B^j} \sum\limits_{a_x \in R(q_i=\hat{o})} \mathcal{A}(a_x) + \frac{c_{Bj}}{r_{Bj}}}{|R(q_i)|}.$$

Here, $R(q_i)$ denotes the responses given by annotators on the i^{th} question. $R(q_i = \hat{o})$ describes the annotators whose opinion is $\hat{o}$ for the i^{th} question. So, the effect of 1^{st} neighborhood over the 2^{nd} neighborhood is

$$\frac{P^{2}_{(q_i=\hat{o})}}{P^{1}_{(q_i=\hat{o})}} = \frac{\sum\limits_{j=1, q_i \in B^{j\prime}}^{|B^{j\prime}|} \sum\limits_{a_x \in R(q_i=\hat{o})} \mathcal{A}(a_x) + \frac{c_{Bj\prime}}{r_{Bj\prime}}}{\sum\limits_{j=1, q_i \in B^{j}}^{|B^{j}|} \sum\limits_{a_x \in R(q_i=\hat{o})} \mathcal{A}(a_x) + \frac{c_{Bj}}{r_{Bj}}} = 1 + \epsilon.$$

Dependent judgment analysis



Dependent judgment analysis

- **Prior independent score:** The label (annotation) marked by an annotator (basically a crowd worker) for a question independently.
- **Posterior dependent score:** The label (annotation) marked by an annotator (basically a crowd worker) for a question after revealing the prior independent scores.

Dependent judgment analysis

An annotation process can be formally represented as a 6-tuple $\langle Q, A, I, D, \tau, \tau' \rangle$, which consists of the following:

- A finite set of questions $Q = \{Q_1, Q_2, \ldots, Q_m\}$,
- A finite set of annotators $A = \{A_1, A_2, \ldots, A_n\}$,
- A finite set of prior independent scores $I = \{i_{1j}, i_{2j}, \ldots, i_{nj}\}$,
- A finite set of posterior dependent scores $D = \{d_{1j}, d_{2j}, \ldots, d_{nj}\}$,
- A mapping function $\tau : Q \times A \rightarrow I$,
- Another mapping function $\tau' : (Q \times A) \rightarrow D$.

Dependent judgment analysis

An annotation process can be formally represented as a 6-tuple $\langle Q, A, I, D, \tau, \tau' \rangle$, which consists of the following:

- A finite set of questions $Q = \{Q_1, Q_2, \ldots, Q_m\}$,
- A finite set of annotators $A = \{A_1, A_2, \ldots, A_n\}$,
- A finite set of prior independent scores $I = \{i_{1j}, i_{2j}, \ldots, i_{nj}\}$,
- A finite set of posterior dependent scores $D = \{d_{1j}, d_{2j}, \ldots, d_{nj}\}$,
- A mapping function $\tau : Q \times A \to I$,
- Another mapping function $\tau' : (Q \times A) \to D$.

Problem statement: Given a set of prior independent scores and posterior dependent scores obtained from the crowd for a given question, predict (after combining together) the 'gold' judgment.

Dependent judgment analysis

Let $\bar{i}_j$ and $\bar{d}_j$ denote the mean of prior independent scores $\{i_{1j}, i_{2j}, \ldots, i_{nj}\}$ and posterior dependent scores $\{d_{1j}, d_{2j}, \ldots, d_{nj}\}$, respectively, given by the 1^{st}, 2^{nd}, . . . , n^{th} annotator on question j.

Dependent judgment analysis

Let $\bar{i}_j$ and $\bar{d}_j$ denote the mean of prior independent scores $\{i_{1j}, i_{2j}, \ldots, i_{nj}\}$ and posterior dependent scores $\{d_{1j}, d_{2j}, \ldots, d_{nj}\}$, respectively, given by the $1^{st}, 2^{nd}, \ldots, n^{th}$ annotator on question j.

- Confidence gap: The confidence gap of annotator k for a particular question j is defined as follows

$$|i_{kj} - d_{kj}|.$$

Dependent judgment analysis

Let $\bar{i}_j$ and $\bar{d}_j$ denote the mean of prior independent scores $\{i_{1j}, i_{2j}, \ldots, i_{nj}\}$ and posterior dependent scores $\{d_{1j}, d_{2j}, \ldots, d_{nj}\}$, respectively, given by the $1^{st}, 2^{nd}, \ldots, n^{th}$ annotator on question j.

- Confidence gap: The confidence gap of annotator k for a particular question j is defined as follows

$$|i_{kj} - d_{kj}|.$$

- Reliability: The reliability of annotator k for question j is defined as follows

$$RA_{kj} = \frac{1 + |\bar{i}_j - \bar{d}_j|}{1 + |i_{kj} - d_{kj}|}.$$

Dependent judgment analysis

- Accuracy: The accuracy of an annotator is defined as follows

$$\frac{1}{1 + |d_{kj} - \bar{d}_j|}.$$

Dependent judgment analysis

- Accuracy: The accuracy of an annotator is defined as follows

$$\frac{1}{1 + |d_{kj} - \bar{d}_j|}.$$

- Polarity: The polarity of the k^{th} annotator for the j^{th} question is defined as follows

$$PL_{kj} = \frac{(1 + |\bar{i}_j - \bar{d}_j|)}{(1 + 2 * (|i_{kj} - d_{kj}|)) * (1 + |d_{kj} - \bar{d}_j|)}.$$

Dependent judgment analysis

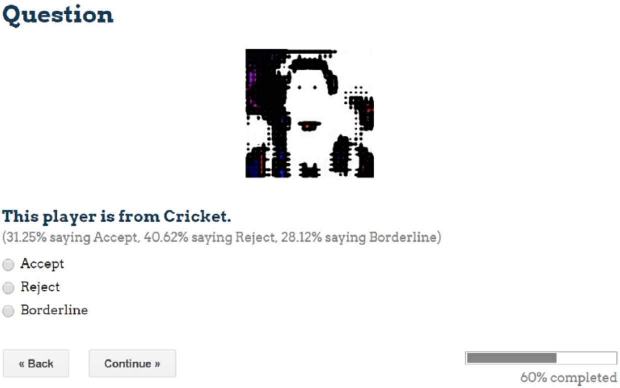
As the opinions in the first phase are independent opinions and opinions in second phase are dependent, therefore the transition of the annotators has been modeled in the transition matrix. At each time step, the final judgment of a question can be computed depending upon the stationary distribution and transition matrix.

Dependent judgment analysis

As the opinions in the first phase are independent opinions and opinions in second phase are dependent, therefore the transition of the annotators has been modeled in the transition matrix. At each time step, the final judgment of a question can be computed depending upon the stationary distribution and transition matrix.

A Markov model can be proposed for deriving the final judgment.

Dependent judgment analysis



Snapshot of a sample question in Image Recognition Task when all the independent opinions are revealed (as percentage) publicly for collecting dependent opinions.

Constrained judgment Analysis

Constrained judgment Analysis

	Question 1	Question 2
Annotator 1	{(10, 20), (22, 33), (42, 30)}	{(10, 20), (20, 30), (40, 30)}
Annotator 2	{(10, 21), (20, 30), (44, 35)}	{(40, 30), (20, 30), (10, 20)}
Annotator 3	{(10, 12), (21, 27), (27, 23)}	{(11, 20), (2, 30), (43, 33)}
Annotator 4	{(11, 22), (20, 30), (29, 50)}	{(12, 22), (20, 30), (30, 30)}
Annotator 5	{(11, 23), (20, 30), (50, 30)}	{(10, 10), (20, 30), (40, 30)}

Constrained judgment Analysis

An annotation process can be formally represented by a 4-tuple $\langle Q, A, O, \tau \rangle$, which consists of the following:

- A finite set of questions $Q = \{Q_1, Q_2, \ldots, Q_m\}$,
- A finite set of annotators $A = \{A_1, A_2, \ldots, A_n\}$,
- A finite set of opinion vectors $O = \{\{(i_{1j}^{11}, i_{1j}^{12}, \ldots, i_{1j}^{1m}), (i_{1j}^{21}, i_{1j}^{22}, \ldots, i_{1j}^{2m}), \ldots, (i_{1j}^{k1}, i_{1j}^{k2}, \ldots,$ $\mathrm{i}_{1j}^{km})\}, \{(i_{2j}^{11}, i_{2j}^{12}, \ldots, i_{2j}^{1m}), (i_{2j}^{21}, i_{2j}^{22}, \ldots, i_{2j}^{2m}), \ldots, (i_{2j}^{k1}, i_{2j}^{k2},$ $\ldots, \mathrm{i}_{2j}^{km})\}, \ldots, \{(i_{nj}^{11}, i_{nj}^{12}, \ldots, i_{nj}^{1m}), (i_{nj}^{21}, i_{nj}^{22}, \ldots, i_{nj}^{2m}),$ $\ldots, (\mathrm{i}_{nj}^{k1}, i_{nj}^{k2}, \ldots, i_{nj}^{km})\}\}$,
- A mapping function $\tau : Q \times A \rightarrow O$.

Constrained judgment Analysis

An annotation process can be formally represented by a 4-tuple $\langle Q, A, O, \tau \rangle$, which consists of the following:

- A finite set of questions $Q = \{Q_1, Q_2, \ldots, Q_m\}$,
- A finite set of annotators $A = \{A_1, A_2, \ldots, A_n\}$,
- A finite set of opinion vectors $O = \{\{(i_{1j}^{11}, i_{1j}^{12}, \ldots, i_{1j}^{1m}), (i_{1j}^{21}, i_{1j}^{22}, \ldots, i_{1j}^{2m}), \ldots, (i_{1j}^{k1}, i_{1j}^{k2}, \ldots, i_{1j}^{km})\}, \{(i_{2j}^{11}, i_{2j}^{12}, \ldots, i_{2j}^{1m}), (i_{2j}^{21}, i_{2j}^{22}, \ldots, i_{2j}^{2m}), \ldots, (i_{2j}^{k1}, i_{2j}^{k2}, \ldots, i_{2j}^{km})\}, \ldots, \{(i_{nj}^{11}, i_{nj}^{12}, \ldots, i_{nj}^{1m}), (i_{nj}^{21}, i_{nj}^{22}, \ldots, i_{nj}^{2m}), \ldots, (i_{nj}^{k1}, i_{nj}^{k2}, \ldots, i_{nj}^{km})\}\}$,
- A mapping function $\tau : Q \times A \rightarrow O$.

Problem statement: Given a set of opinion vectors obtained from the crowd for a given question, predict (after combining together) the ‘gold’ judgment.

Constrained judgment Analysis

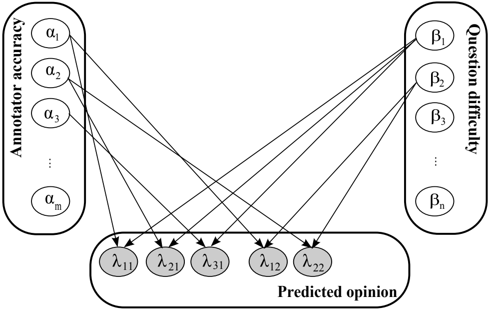
Suppose, λ_{ij} denotes the opinion given by a particular annotator i for a given component j of a question. Then the probability that the given opinion matches with the true label is given by

$$P(\lambda_{ij} = z|\alpha_i, \beta_j) = \frac{1}{1 + \exp(\frac{-\alpha_i}{\beta_j})}.$$

Constrained judgment Analysis

Suppose, λ_{ij} denotes the opinion given by a particular annotator i for a given component j of a question. Then the probability that the given opinion matches with the true label is given by

$$P(\lambda_{ij} = z|\alpha_i, \beta_j) = \frac{1}{1 + \exp(\frac{-\alpha_i}{\beta_j})}.$$



Constrained judgment Analysis

1. Remove the inconsistent solutions
2. Remove the workers who violated the constraints
3. Discretize the opinions using Bayesian binning
4. Apply a probabilistic graphical model

Constrained judgment Analysis

1. Remove the inconsistent solutions
2. Remove the workers who violated the constraints
3. Discretize the opinions using Bayesian binning
4. Apply a probabilistic graphical model

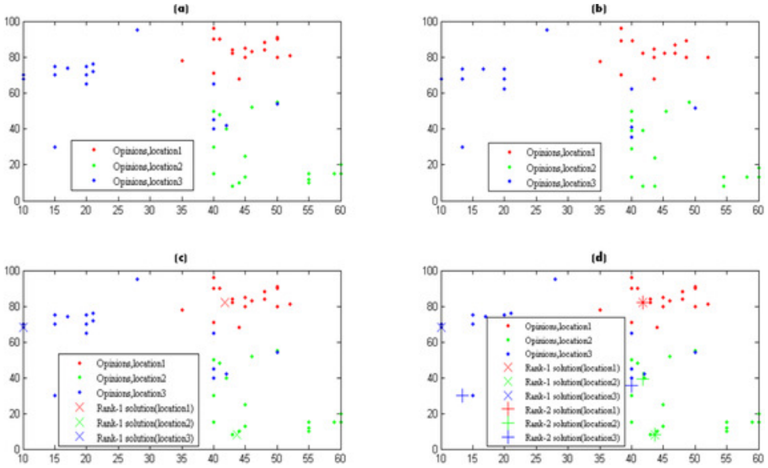
E Step:

$$\mathcal{J}_j(z) = \prod_{i \in \delta_j} P(z)^{I(R_{ij}=z)} (1 - P(z))^{I(R_{ij} \neq z)}.$$

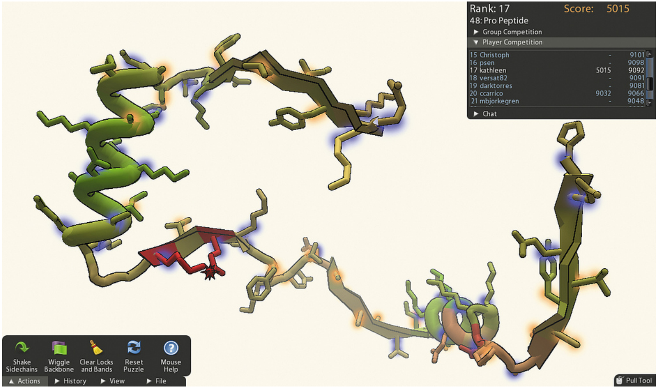
M Step:

$$\alpha_i = \frac{1}{|\sigma_i|} \sum_{i \in \sigma_i} \mathcal{J}(R_{ij}) \times \mathcal{C}.$$

Constrained judgment Analysis

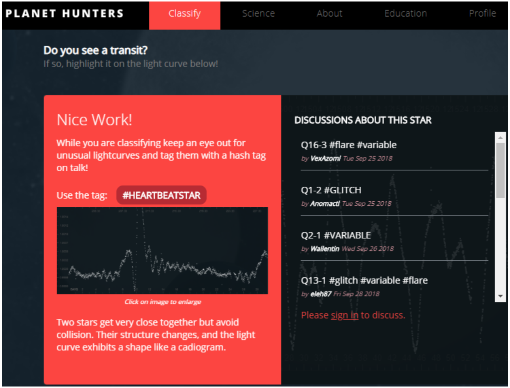


Computational biology



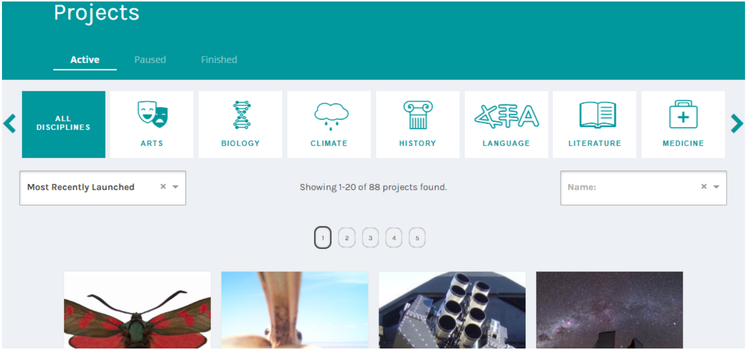
FoldIt: An online puzzle video game about protein folding

Astronomy



classifying the light curves coming from celestial objects

Innovation challenges



Zooniverse: A platform for research powered by the crowd volunteers

References

- C. Chatterjee *et al.* (2017) *Information Sciences*, 396:83-96.
- C. Chatterjee *et al.* (2017) *AAAI SSS*, AAAI Press, 16-22.
- M. Hoffman *et al.* (2015) *PNAS USA*, 112(6):1727-1732.
- J. Surowiecki (2004) *The Wisdom of Crowds*, Doubleday.
- A. W. Woolley *et al.* (2010) *Science*, 330(6004):686-688.

That's all folks!!!