

ワーカーに着目したインセンティブ設計に関する考察

中平一六^{†1} 杉村太一^{†2} 小坂隆浩^{†2}

概要: 本研究では、クラウドソーシングにおけるインセンティブ設計に着目した。インセンティブ設計とは、タスクの設計や報酬設定によりワーカーを動機付けるものである。クラウドソーシングにおける、インセンティブやメカニズムに関する研究は多く行われてきたが、依頼者が専門家でないことを考慮すると、これらの研究の中から高品質な成果物を得られるようなインセンティブを選択することは困難であると考えられる。本研究ではクラウドソーシングにおける適切なインセンティブ設計の実現を目指す。目標達成のために初期調査として、タスクの依頼者としてコスト制限がある中で、どのようなインセンティブを与えるかというタスクを実施し、各ワーカーが考案するインセンティブと個人の属性との関係の調査を行った。

キーワード: クラウドソーシング, インセンティブ, 報酬設定

1. はじめに

近年ネットワークの普及に伴い、インターネット上では様々なサービスが展開されており、Amazon Mechanical Turkに代表されるクラウドソーシングへの注目が集まっている。クラウドソーシングは、インターネットを介して不特定多数の人々の知恵や力を利用することによって問題を解決するサービスである。

クラウドソーシングでは、安価に不特定多数のワーカーを集めることが可能である一方で、得られる作業結果の品質が問題となっている。これまで、品質管理がしばしば課題として取り上げられており、その解決策が考えられてきた。

品質管理の改善手法の先行研究としては、Liu らによるワーカーに対して適切なインセンティブを与えるメカニズム(ルール)を設計することによって、品質管理を実現する方法がある[1]。また、G. Kazai らは情報検索システムにおけるラベル付けを、クラウドソーシングを用いて行い、ワーカーの属性情報とラベル付けを行う分野ごとの正答率との関係性を調査し、属性情報の有用性を示した[2]。

本研究では、クラウドソーシングにおける適切なインセンティブ設計の実現を目指す。目標達成の初期調査として、自己申告により得たタスクへの動機とワーカーの属性情報である年齢、性別、自尊心を用いて画像認識のタスクへの回答と併せて評価し、ワーカーの属性情報と動機との関係について調査し、考察する。

2. ワーカーの属性推定

ワーカーの属性推定には、機械による属性推定と人による属性推定の2種類が存在する。以下にそれらの特徴について述べる。

機械による属性推定は、例としてテストタスクの実施が

挙げられる。正解が既知なタスクにより正答率を算出し、ワーカーとタスクの相性を調査することで適正ワーカーを探す。

人による属性推定は他者評価と自己申告に分けて考えられる。他者評価は回答者以外の他のワーカーやリクエスター自身が成果物を評価する。自己申告はワーカーの属性情報や自信度を申告してもらう。ワーカーの属性情報として年齢、性別、自尊心を扱う。自尊心については次節で述べる。

Instructions:

Below is a list of statements dealing with your general feelings about yourself. There are four possible answers for each of the 10 questions, from "strongly agree" to "strongly disagree". Tap the box to indicate how strongly you agree or disagree with each statement.

| | Strongly Agree | Agree | Disagree | Strongly Disagree |
|---|----------------|-------|----------|-------------------|
| 1 On the whole, I am satisfied with myself | 3 | 2 | 1 | 0 |
| 2 At times, I think I am no good at all | 0 | 1 | 2 | 3 |
| 3 I feel that I have a number of good qualities | 3 | 2 | 1 | 0 |
| 4 I am able to do things as well as most other people | 3 | 2 | 1 | 0 |
| 5 I feel I do not have much to be proud of | 0 | 1 | 2 | 3 |
| 6 I certainly feel useless at times | 0 | 1 | 2 | 3 |
| 7 I feel that I'm a person of worth, at least on an equal plane with others | 3 | 2 | 1 | 0 |
| 8 I wish I could have more respect for myself | 0 | 1 | 2 | 3 |
| 9 All in all, I am inclined to feel that I am a failure | 0 | 1 | 2 | 3 |
| 10 I take a positive attitude toward myself | 3 | 2 | 1 | 0 |

図 1 Rosenberg Self-Esteem Scale (RSE)

2.1 Rosenberg Self-Esteem Scale (RSE)

心理学において自己に対する肯定的または否定的な態度や自己価値の感情を表す自尊心というものがある。本研究では Rosenberg が開発した自尊心尺度 RSES を用いる[4]。特徴として 10 項目という比較的簡単な尺度、半数の項目が逆転項目であり正方向と負方向の項目が混在していると

^{†1} 同志社大学
Doshisha University
^{†2} 同志社大学院
Doshisha University

いう利点がある。スコアの範囲は0～30の範囲で15～25は通常、15未満は自尊心が低い、25以上は自尊心が高いとされている。図1に質問文と採点方法を示す。

3. 調査結果

3.1 実験方法

Amazon が運営しているクラウドソーシングプラットフォームである Amazon Mechanical Turk を用いて 150 人のワーカーを対象に実験を行った。年齢、性別、タスクの動機のアンケート、Rosenberg Self-Esteem Scale を用いたテスト、画像認識の問題 4 問を 1 つのタスクとして用意し、ワーカーに依頼した。画像認識の問題には機械学習用のデータセットである CIFAR-10 を用いた。CIFAR-10 は約 6 万の画像からなり、画像を自動車、鳥、猫、鹿、犬、蛙、馬、船、トラックの 10 種類に分類するものである。

タスクの動機としては回答の多かったものを採用しており、難易度、興味、金銭、自己評価、その他の 5 つとする。

3.2 実験結果

まず各動機別の平均正解率を図 2 に示す。タスクの動機として難易度、自己評価のためにタスクを選択しているワーカーの正解率が高い結果となった。

次に各動機と個人の属性に関して自尊心についての平均正解率の分布を図 3 に示す。全体的にどの動機でも自尊心が高い、もしくは低いワーカーの正解率が高い結果となった。他の属性情報である性別に関して、女性は自己評価、男性は難易度、金銭が動機のワーカーの正解率が高い結果となった。

最後に正解率が 75%以上の自尊心が高い、もしくは低いワーカーの動機の分布図を図 4 に示す。自尊心が高いワーカーは自己評価、自尊心が低いワーカーは金銭を動機としていることが多く、興味を動機としている場合はどちらのワーカーもあまり差は見られなかった。

実験の結果、難易度、自己評価をタスクの動機としているワーカーの正解率が高く、次に行うタスクへ生かそうとしていることが考えられる。また、属性情報と動機に関して関係性が見られることがわかった。

4. まとめ

今回の調査では自己申告によるワーカーのタスクの動機と属性推定との関係について分析を行なった。

調査結果として、属性情報と動機に関して関係性が見られることがわかった。今後は調査するサンプル数を増やし複数の属性情報を組み合わせた場合や動機の有意性について調査していく予定である。

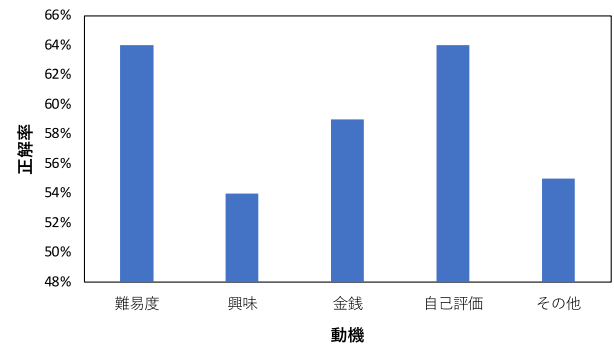


図 2 全体の平均正解率

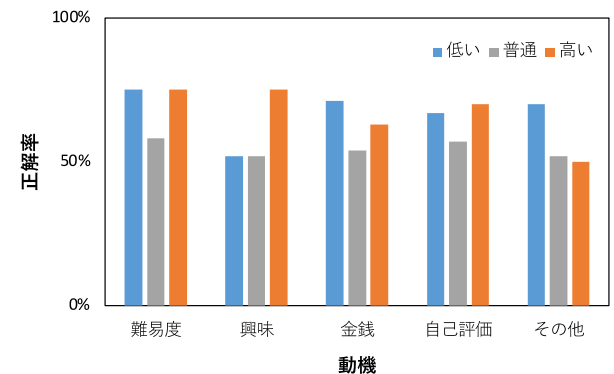


図 3 自尊心別の平均正解率

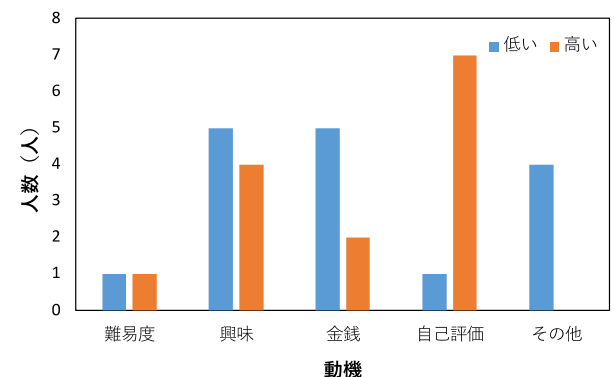


図 4 平均正解率 75%以上 (自尊心別)

参考文献

- [1] Liu, Y. and Chen, Y.: Sequential Peer Prediction: Learning to Elicit Effort using Posted Prices, in Proceedings of the 31st AAAI Conference on Artificial Intelligence, pp.495-503, 2017.
- [2] G. Kazai, The Face of Quality in Crowdsourcing Relevance Labels: Demographics, Personality and Labeling Accuracy, Proc. of the ACM Int'l conf. on Information and knowledge management, pp.2583-2686, 2012.
- [3] 櫻井, クラウドワーカーを利用したメカニズム設計のための一考察, 第 31 回人工知能学会全国大会, 2017.
- [4] Rosenberg, M., Rosenberg self-esteem scale (RSE), Acceptance and Commitment Therapy, Measures Package, 61, 1965.