

クラウドソーシングの現状と課題

清水伸幸
Nobuyuki Shimizu

ヤフー株式会社 Yahoo! JAPAN 研究所
Yahoo! JAPAN Research, Yahoo Japan Corp.

不特定多数の人々から貢献を募ることで、必要なサービス、アイデア、コンテンツを生み出す行為はクラウドソーシングと呼ばれ、近年、様々な分野で注目を集めています。本講演では、2006年にジャーナリスト Jeff Howe によって提唱されたクラウドソーシングというコンセプトの実施例とその課題、アカデミアでの様々な取り組み、また、昨年スタートしたヤフーのクラウドソーシングサービスでの取り組みについてご紹介いたします。

1. はじめに

近年、インターネットを通じて、不特定多数の人々により一定のタスクを成し遂げるクラウドソーシングと呼ばれるサービスの利用が急速に拡大しています。計算機科学分野においても、人間の知的活動を計算機で補佐するため、特に米国においては画像の注釈付けや Web コンテンツの分類などの教師データの収集や、自動要約、機械翻訳など自然言語処理システムの評価などへのクラウドソーシングの利用が始まっています。

クラウドソーシングの中でも、特にマイクロタスクと呼ばれる細分化されたタスクを行うタイプのサービスは、Yahoo!クラウドソーシングをはじめとするサービスプラットフォームによって、日本でも手軽に利用できるようになり、多くの人間の知見を集約することが可能となりました。今では、大学の一研究室が、数日で1000人以上のユーザーを動員してタグ付けなどのタスクを行うことが可能になっています。人工知能の分野でも、クラウドソーシングから得られる大量の訓練データを用いて、さらなる成果が生まれるのではないかと期待されています。

2. クラウドソーシングの課題

こうして、新しいツールとして認識されつつあるクラウドソーシングですが、従来のように専門家にタスクの依頼を行うケースと異なり、品質の高い成果物を得るためには従来とは異なるタスクの設定手法が必要であることがわかってきました。

従来と異なる点の一つは、手軽に報酬を得ることを目的として低品質の成果物を納める不誠実なユーザーや、設問の説明文を理解していないとみられるユーザーが一部存在していることです。こういったユーザーからの影響を最小限におさえ、成果物の品質を管理するため、同じタスクを複数のユーザーに依頼し多数決を取るなど、成果物を統計的に統合することにより、質の高い成果物を獲得するやり方が使われます(Dawid and Skene 1979, Whitehill et al 2009)。この他にも、テスト問題と知らせずタスク中にテストを行う手法が存在します。

従来と異なる二つ目の点は、タスクの意図を正確に伝えるため、タスクの表現に注意を払う必要がある点です。例えば、固有

名詞抽出の技術を評価するため、抽出したある単語 X について、クラウドソーシングで「X は飲食店の名前ですか?」と、ユーザーに聞いたとします。単純な質問のようですが、プラットフォーム運営の経験から、結果は少なくとも以下の3つの解釈が混じったものとなることがわかってきました。

(1) まず、店としての存在の有無にかかわらず、文字列として飲食店らしいかどうかで判断すればよい、という解釈、

(2) 次に、検索などで、店として現実に開店していることを確認して答えを出す、という解釈、

(3) そして最後に、よくある店の知名度調査だと解釈し、店の名前を以前に聞いたことがあるかないかで返答すれば良い、という解釈です。

このように、一見単純そうな質問でも、タスク依頼者の意図とタスクを行うユーザーの解釈が分かれる事態は頻繁に起こります。不誠実なユーザーの存在を疑ったところ、実は解釈のゆらぎに起因する問題だったといった不幸な誤解も起こりえます。こういったタスク設定の難しさは、企業において新人研修の効果的なシステムを作り上げるといったプロセスと類似しており、クラウドソーシングの隠れた導入コストと言えます。

クラウドソーシング利用の課題を2つほど述べましたが、解決には、人が人に物を頼むプロセスを計算機科学で支援することが必要となります。こうして得られるデータからは、人間のコミュニケーションを人工知能でモデルするにあたり新しい知見が得られるはずですが、前人未達の分野で、新たな人工知能のアプリケーションが生まれることとなるでしょう。

3. クラウドソーシング利用のベストプラクティス

これからマイクロタスク型のクラウドソーシングを利用される皆様のために、注意すべき点を挙げておきます。

1 ユーザーに実データの利用方法が想像しやすい質問の仕方をしましょう。タスク製作者が思ってもみなかったユーザーの解釈や誤解が減らせる場合が多くあります。

2 設問の説明には具体的な例を入れ、ユーザーが回答するにあたり、検索が必要ななら検索へのリンクを貼りましょう。すべき作業が具体的であればあるほど解釈の揺れが減ります。

3 自由記述のコメント欄をつけて、ユーザーが判断に迷う際にはどうして迷ったのかフィードバックを得て、次のイテレーションでのタスクの改善につなげましょう。

4 新しいタスクをクラウドソーシングに入稿する際には、小規模なデータでパイロットテストを行い、何度か試行錯誤をした上で大規模なデータを投入すると良い成果が得られます。

5 2つの質問を組み合わせて一つの答えを求める等、認知言語学的に処理が難しい聞き方を避けましょう。例:「以下のブログの内容は飲食店についてのものでしょうか?また、このお店に行ってみたいと思いますか? はい・いいえ」。

参考文献

Dawid, A. P. and Skene, A. M. (1979) "Maximum Likelihood Estimation of Observer Error-Rates Using the EM Algorithm", *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 28, No. 1, pp. 20-28

Whitehill, J., Ruvolo, P., Wu, T., Bergsma, J., and Movellan, J. (2009) "Whose Vote Should Count More: Optimal Integration of Labels from Labelers of Unknown Expertise", in *Advances in Neural Information Processing Systems* 22