

分類マイクロタスクにおけるタスク順序制御に関する研究

A Study on Task Ordering for Classification Microtasks

学籍番号：201521637

氏名：根本 千代之介

Chiyonosuke NEMOTO

計算機ネットワーク技術の発達に伴い、不特定多数の群衆に作業を委託するクラウドソーシングが多く分野で用いられている。クラウドソーシングの中でも、作業に要する時間が短いものをマイクロタスク型クラウドソーシングとよぶ。マイクロタスク型クラウドソーシングで行われる典型的な作業の1つとしてデータの分類がある。

データの分類を行うタスクにおいては、大量のデータ中からある特定のクラスのデータのみを、少数のタスクで獲得したい場合が存在する。例えば、写真に写っている家屋が被災しているか否か分類するタスクでは、迅速な支援が必要とされることから、被災家屋が写っている写真のみを少数のタスクで獲得したいといった事が考えられる。

このような場合、タスク中のデータが持っている特徴を利用し、タスクの割り当て順を変更することが有効である。例えば先ほど述べたタスクでは、写真の色などが特徴として利用できる。しかし、リクエスタのみで有効な特徴を発見することは必ずしも容易ではない。

そこで本論文では、クラウドソーシングによって大量の特徴を獲得し、その中から有効な特徴を選別しつつ選別した特徴を正例の発見に利用する手法を2つ述べる。クラウドソーシングによって特徴を入手するのは、リクエスタだけでは発見できない有効な特徴の候補を数多く入手するためである。これら2つの手法はともに、その時点までに得られているすべての回答を利用して特徴を選別し、次に割り当てるタスクを出力する。2つの手法は特徴を選別する部分のアルゴリズムが異なっており、統計情報を用いた手法では検定を利用して特徴を選別する。一方、機械学習アルゴリズムを用いた手法では、既存の機械学習アルゴリズムを用いた手法を用いて特徴を重みづけすることにより選別を行う。

シミュレーションでは、2つのクラスに分類された約6,700件のニュース記事を使用し、2つの手法それぞれについて、どれだけ少ないタスク数でどれだけ多くの正例を獲得できたかを比較した。シミュレーションの結果、タスク数が少ない段階では統計情報を用いた手法の方が機械学習アルゴリズムを用いた手法よりも有効である可能性が示唆された。

研究指導教員：森嶋 厚行

副研究指導教員：歳森 敦