



HOKKAIDO UNIVERSITY

Title	A Study on Crowdsourcing for Multi-Label Affect Annotation [an abstract of dissertation and a summary of dissertation review]
Author(s)	段, 磊
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第12181号
Issue Date	2016-03-24
Doc URL	https://hdl.handle.net/2115/61600
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	doctoral thesis
File Information	Lei_Duan_review.pdf, 審査の要旨



学位論文審査の要旨

博士の専攻分野の名称 博士 (情報科学) 氏名 段 磊

審査担当者 主 査 准教授 小山 聡
副 査 教 授 栗原 正仁
副 査 教 授 山本 雅人
副 査 教 授 小野 哲雄
副 査 教 授 川村 秀憲

学位論文題名

A Study on Crowdsourcing for Multi-Label Affect Annotation

(複数ラベル感情アノテーションのためのクラウドソーシングに関する研究)

一つの対象に複数のラベルが付与されているマルチラベルアノテーションからの学習は、機械学習の分野で近年注目を集めている。本論文はマルチラベルアノテーションの中でも、多くの応用をもつ感情アノテーションを対象として扱っている。マルチラベルアノテーションからの学習が成功するためには、品質の良いラベルが付与されたデータを、大量に準備することが必要となる。高品質なアノテーションは、専門家に依頼すれば得られるが、コストが高い上にそのような専門家がいつでも確保可能とは限らない。そこで、近年普及しつつある、インターネット上で不特定多数の作業者に仕事を依頼できるクラウドソーシングを用いることが、有望なアプローチと考えられている。

本論文はマルチラベル感情アノテーションの収集・利用における問題をクラウドソーシングを用いて解決したものであり、大きく二つの貢献が認められる。貢献の一つは、クラウドソーシングを用いた高品質かつ低コストなアノテーションの収集である。専門家とは異なり、クラウドソーシングの作業者はその能力や経験が様々であるため、作業結果の品質にもばらつきがあることが多い。高品質のアノテーションを得るために、複数の作業者に同じ作業を依頼し、その結果を多数決などで集約することがよく行われる。しかし、マルチラベルのアノテーションは可能な候補の数が多く、信頼性のある集約結果を得るためには、多数の作業者に冗長な作業を依頼する必要がある。特に主観的な感情アノテーションから、多くの人の同意を得ることができるアノテーションを得るためには、客観的なアノテーションに比べて多くの作業者に依頼を行う必要がある。そこで本論文では、比較的少数の作業者の作業結果を集約することで、多数の作業者の作業結果を集約した場合と同等のアノテーションを得るための方法を提案している。

本論文のもう一つの貢献は、クラウドソーシングを用いることで、異なる感情語彙を用いたマルチラベルアノテーション間の相互変換を可能にする方法を設計したことである。感情を表現する語彙は様々なものが提案されており、感情情報処理システムによって異なるものが用いられている。そのため、ある語彙でアノテーションされたデータを、別の語彙を用いるシステムで利用できない

という問題が生じる。本論文はクラウドソーシングを用いて、一部のデータに用いたい語彙でアノテーションを行い、語彙間の意味的なマッチングを学習することで、異なる語彙を用いるシステム間での相互運用を可能にしている。

本論文の構成は以下のとおりである。第1章では、本研究の背景が述べられ、マルチラベルアノテーションからの学習やクラウドソーシングの最近の研究について紹介されている。そして、クラウドソーシングによって得られた作業結果から、高品質なマルチラベルアノテーションを得る際の困難について述べられ、本研究の目的と貢献が示されている。

第2章では、以降の章の理解に必要な基礎知識として、感情の特徴付け、マルチラベルアノテーションからの学習、クラウドソーシングとその品質制御について説明がなされている。その後、論文で利用したデータセットと関連研究について記述されている。

第3章では、少数のクラウドソーシング作業者のアノテーションを集約し高品質なアノテーションを得るために、ラベル間の依存関係を考慮する方法が述べられている。ラベル生成過程においてラベル間の依存関係を考慮した確率モデルを用いることで、既存の方法よりも高い精度を実現している。

第4章では、一般に感情表現は一貫性と文脈に制約されることに着目し、ラベル生成過程において感情的な一貫性と文脈上の手掛かりに基づく事例間の依存関係を考慮することで、アノテーション集約の精度を向上させる方式を示している。

第5章では、ある語彙でアノテーション済みのデータの一部にクラウドソーシングで別の語彙でアノテーションを行い、そこから異なる感情語彙間での意味的なマッチングを学習することで、異なる感情語彙を用いたシステム間の相互運用を実現する方法を示している。

第6章では、本論文で提案された方法が総括され、今後の研究課題が議論されている。

これを要するに、著者は、クラウドソーシングによって得られた作業結果を、ラベル間および事例間の依存関係を考慮して集約することで、高品質な感情アノテーションを低コストで収集する方法を設計し、さらに異なる感情語彙を用いたシステム間の相互運用を実現する方法を設計したものであり、知能情報学に対して貢献するところ大なるものがある。よって著者は、博士(情報科学)の学位を授与される資格あるものと認める。