

平成 30 年 6 月 11 日現在

機関番号：17102

研究種目：基盤研究(B) (一般)

研究期間：2013～2017

課題番号：25280118

研究課題名(和文) 意見文からなる大規模テキスト集合に潜む人々の価値観を推定するための基礎的研究

研究課題名(英文) Fundamental research to estimate human values in opinion texts of large collection

研究代表者

石田 栄美 (Ishita, Emi)

九州大学・附属図書館・准教授

研究者番号：50364815

交付決定額(研究期間全体)：(直接経費) 8,600,000円

研究成果の概要(和文)：ある事柄に対する意見を述べたテキストに現れる人の価値観を自動推定する分類器 Latent Variable Model(LVM)を開発した。これを評価するため、内容分析で用いている分析方法を反映した手法を提案した。次に、語と価値観との関連を示す分類器を開発し、社会科学の観点から語の使用パターンを分析した。さらに、分類器の汎用性を検証するため、福島第一原子力発電所の災害に関する新聞記事を対象に、人手により価値観に関するカテゴリを付与した新しいテストコレクションを構築した。このコレクションに分類器を適用した結果をもとに、コレクションの構築方針を再検討し、文単位、複数のカテゴリを付与することにした。

研究成果の概要(英文)：This report describes the development and evaluation of a "Latent Variable Model" (LVM) classifier that detects evidence for specific human values at either the passage or full-document level. To determine suitability for a specific application, an extrinsic evaluation method reflecting the use of the classification results for content analysis was performed. To extend the utility of the technique, the relationship between words and human value was analyzed to discern patterns of language use. To characterize the versatility of the classifier, new test collections containing newspaper articles about the Fukushima Daiichi nuclear power disaster were developed in which human values were manually assigned as a gold standard at document level. Results obtained on these collections has informed the design of future test collections in which human annotation will be performed at sentence level and multiple categories.

研究分野：図書館情報学

キーワード：価値観推定 内容分析 価値観 自動分類 テストコレクション 国際研究者交流

1. 研究開始当初の背景

テキストの背後にある価値観の分析は情報政策や情報倫理などの分野で行われており、世論の調査や政策決定を行うときの材料とされることがある。これらに携わる研究者は、人が述べた意見文に対してその背後に潜む価値観を手により分析し、それぞれの意見がどのような価値観のもとに述べられているかを分析している。本研究グループでも、アメリカの公聴会における「ネットの中立性」に対する証言に対し、賛成・反対の立場のステークホルダーがどのような価値観を持っているかを人手による内容分析から考察した研究事例がすでにある。

内容分析は、質的分析を伴うため通常は人手により行われている。そのため、大量テキストを対象にすることは難しく、限定された範囲の分析にとどまってしまう例が多い。社会科学分野の研究者からは、大量のテキスト、例えば、ウェブ上で述べられている様々な意見を対象とした分析をしたいという要望もある。そのため、本研究グループでは、社会科学の研究者とともに内容分析の自動化の重要性と可能性について検討してきた。

本研究では、意見文にある価値観を推定する手法を開発するが、これにはテキストの自動分類(text categorization)技術を応用することができる。自動分類研究は、電子化テキストが増加したことを受けて1990年代後半から研究が始まり、とくに主題カテゴリに対する分類が研究されてきた。たとえば、書名の情報を用いて図書に日本十進分類法の分類記号を付与する例や、新聞記事に出現する語を統計的に分析し政治・経済・スポーツなどのカテゴリに分類する例などがある。これらの研究事例をもとに価値観の推定手法を開発する予定であるが、価値観の推定は、主題カテゴリの推定と異なり、価値観カテゴリと文中に出現する語の直接的な結びつきは弱いことが考えられる。つまり、価値観は文全体で暗黙的(implicit)に示されるものであり、価値観カテゴリと関連する語を特定することが難しい。また、映画や商品のレビューにおける感情分析のように肯定/否定の2値分類ではないこと、一つのテキストに対し複数のカテゴリが付与される場合が多いこともあり、これまでの主題/非主題カテゴリの分類よりも困難と考えられる。

2. 研究の目的

ある事柄に対する意見を述べたテキストの背後に潜む人の価値観を自動的に推定する手法の開発を行う。また、開発した手法が、工学的な観点からの有効性だけでなく実際に社会科学の分野における内容分析にも適用できるか、その内容分析のタスクにあわせた評価方法を検討する。同時に、分類器の学習に必要な人によるカテゴリ付与(コーディング)を自動的に推定手法によってどの程度軽減することができるかも検討する。最後に、

大規模テキスト集合を構築し、開発した価値観推定手法を適用する。

本研究での目的は上記で述べた3点であるが、それぞれの具体的な課題と方法について以下にまとめる。

(1) 意見文にある価値観推定技術の開発

主題カテゴリに対する分類の場合、サッカーという語が含まれていればスポーツのカテゴリに属する確率が高くなるというように、カテゴリと直接的に結びつく単語が存在する場合が多い。しかしながら、「正義」という価値観を重視している場合、これに直接的に結びつく語を同定するのは難しく、価値観を推定する語の識別は非常に難しいといえる。また、対象とするテキストは一文単位とするため、含まれる情報量が少ないことも、推定を難しくしている。

本研究では、これらの難しさを考慮し、より高度な統計モデルを用いた推定、情報量の少なさに対する対処など多面的な視点から様々な推定手法を検討する。

(2) 自動推定結果の新しい評価手法の提案

推定結果に対する工学的な評価は、一文に対して正しく価値観が付与されたかという観点からの評価である。一般に、精度、再現率、F値、Accuracyなどが用いられる。本研究で想定している推定手法の適用は、内容分析に対してであり、内容分析では複数の異なる立場における発言にある価値観の分布を分析している。この分布に着目し、内容分析で用いる評価方法を考慮した新しい評価の枠組みを提案する。

(3) 大規模テキスト集合への適用

開発された新しい手法が、実験で用いた小規模テキスト集合の特性に依存している可能性もある。手法の汎用性を確認するために、新しいテキスト集合を用いて検証を行う。テーマ依存性、言語依存性も検討し、またこれらの結果を用いた社会科学分野の分析も行う。その際、価値観推定手法の評価だけでなく、社会科学の分野で実際に役に立つテーマ、分析対象テキスト、コーディングフレーム等にする。

3. 研究の方法

(1) 価値観自動推定手法の開発

推定手法の開発は、まず、小規模テキスト集合を用いて実験を行い、価値観推定手法を確立する。小規模テキスト集合は、米国上院商務・科学・運輸委員会のネットの中立性に関する公聴会と連邦通信委員会によりブロードバンドネットワークに関する公聴会での証言(英語)、102文書である。価値観カテゴリはfreedom, innovation, honor, justice, social order, wealthの6カテゴリから構成されており、文書の一文ずつにこれらのカテゴリが付与されている。このコー

ディングは、社会学者によって行われたものである。これを用いて、文それぞれに対する価値観を推定する手法の開発を目指す。

一文を構成する単語の量が少ないこと、102 証言の文の数が 9,000 文程度と数が少ないということは、データがスパース(疎)ということであり、実験用集合を用いて単語の集合と価値観との関連を明確に示すことができない。これに対応するため、単語を補う方法、類似した概念を持つ単語をグループ化して扱う 2 つのアプローチを検討する。一つ目は、同じ、もしくは類似した主題について書かれたテキスト集合を用いて、単語の共出現情報を基に、元々の文章には出てこない単語を補った推定手法を検討する。一方で、他の単語を補わず、単語をいくつかのグループにわけ、そのグループとカテゴリとの関係を情報として用いた推定方法を検討する。具体的には、潜在的意味カテゴリを有する文生成モデルを用いた識別手法を適用する予定である。また、一文に付与されたカテゴリ間の独立性をカイ二乗検定を用いて調べたところ、いくつかのカテゴリ間には明らかな依存関係があることがわかっている。たとえば、innovation が付与されていると wealth も付与されていることが多い。これらの依存関係の情報を利用した手法も取り入れる予定である。これでも十分でない場合には、文脈を考慮した推定手法、複数カテゴリを推定する手法の検討などを行う。

(2) 推定結果の新しい評価方法の提案

自動推定手法の結果を評価する場合、工学的なアプローチでは、人間が行ったものと同じ結果が得られたかが基準となり、一文に対する価値観カテゴリの付与の有無を基準に評価を行う。通常は、再現率、精度、F 値などが用いられる。本研究でのベースラインの結果では、分類器 Support Vector Machine(SVM)を用いて 0.72 程度の F 値を得ている。しかし、この結果が社会科学的な分析に利用可能なものであるかの検討はなされていない。

価値観を付与したテキスト集合を用いた社会学者の分析は、すでに本研究グループで、ネットの中立性に関する賛成派と反対派の証言における各カテゴリの出現の分布を統計的検定を用いて、有意差があるかを調べている。その結果、ネットの中立性を支持するものは innovation という価値観を持っており、反対するものは wealth という価値観をもっていることが明らかになっている。自動的に行った価値観の推定の結果を、このような分析に用いるためには、一文単位の評価ではなく、分析したい属性間での分布の差をみる必要がある。本研究では、工学的な観点からの評価もするが、ある属性間での分布の違いを見るための新しい評価の枠組み、新しい評価手法を提案する。上記の分析では、分布に有意な差があるかないかの 2 値がわか

ればよいが、推定手法を改良するためには、分布に具体的にどのような差があり、人手で行ったものとどの程度異なるのかという度合いが、数値として示せることが適当である。これには、コサイン類似度係数、平均二乗誤差、JS ダイバージェンスなどの指標を応用することを想定している。

(3) 大規模テキスト集合のための価値観カテゴリの再検討

推定手法の開発は小規模テキスト集合を用いて行うが、開発手法の汎用性を検証するため、大規模テキスト集合を構築し、それらに対し価値観カテゴリの付与を行う。現在のところ、原子力発電所、大統領選挙などが考えられるが、新しく適当なテーマが出てくる可能性もあり改めて検討する。対象となるテーマが決まれば、価値観カテゴリの再検討を行う。現在の 6 カテゴリは、ネットの中立性のために構成されたカテゴリのため、異なるテーマであれば、異なるカテゴリセットが必要となる場合もあり、再検討を行う。また、人手で価値観付与を行ったサブセットを構築するため、コーディングマニュアルなども作成する。対象とするテキストは、一つのテーマについて意見が述べられているウェブ集合をクロールする予定であるが、効率的なクロール手法について検討する。検索エンジン API 等を用いて、ブログやツイッターなどを個人の意見が見られるテキストを収集する。

大規模テキスト集合を構築した後は、開発した手法のテーマ依存性、言語依存性(一部日本語テキストも対象にする)などの汎用性を検証する。依存性が認められた場合には、より汎用性の高い推定手法の検討を行う。

4. 研究成果

(1) 価値観の自動推定手法

推定手法の開発には、対象テストコレクションとして、米国でネットの中立性に関して公聴会での 102 証言(約 9,000 文)を用いた。これらの一部を用いて分類器を学習させるが、対象データでは学習に十分な量ではなく、また一文あたりの単語の数が少ないという課題がある。

そのため、まず、学習に用いるものと同内容のテキスト集合から統計的共起関係により自動抽出した語、およびシソーラス中で上位・同義関係にある語を用いて、学習用テキスト集合に対して連想語として補完することで推定性能を改善するアプローチを考案した。その結果、一般に性能が高いといわれる分類器 SVM よりも高い性能を示した。

次に、さらに推定性能の高い分類器を開発するため、文を構成する語が価値観を持ち、文の価値観は語が持つ複数の価値観の集まりで表されるというモデルを反映した分類器 Latent Variable Model(LVM)を開発した。先のネットの中立性のテストコレクション

を用いて評価したところ、SVM よりも高い性能を示した。また、対象データに対して二人目のコーダーが付与した結果よりも、主のコーダーと近い結果を得ることができた。

また、別の観点からの分類器を開発した。分類器を用いた価値観の推定は、推定対象のテキストを分類器に与えれば分類器が判断した価値観を示す。しかし、語と価値観の関係までは示さない。社会学者と議論した際に、語と価値観の関係を社会科学的観点から分析したいという要望があることがわかった。そのため、上記の分類器よりは性能は落ちるが、語と価値観との関連を示すことができるモデルを提案した。さらに、このモデルから得られた結果を用いて、社会科学的観点から分析を行った。

(2) 様々な観点からの推定結果の新しい評価手法の提案

(1)で開発した分類器は、人により価値観カテゴリが付与されたテキストが必要である。人によるコーディングは、質的分析を伴うこともあり、大量のテキストを対象にすることはできない。そのため、どの程度のノテーションされたテキストがあれば、分類器は人手によるコーディング作業と類似した結果をだすことができるかについて分析した。対象データである 102 文書のうち 1 文書から 51 文書まで 1 文書ずつ学習データを増やしていき、SVM やナイーブベイズ (NB) 分類器を学習させ、残りの 51 文書中の文の価値観を推定した。その結果、SVM に関しては 30 文書で 101 文書を使ったときの性能の 85% が得られることがわかった。NB に関しては、30 文書で 88% の性能が得られることがわかった。6 つの価値観カテゴリのうち、"honor" を除けば、追加の人によるコーディングなしにより大きなコレクションに提供できる可能性があることがわかった。

分類器の性能は、工学的な観点から精度、再現率、F 値などの尺度を用いて評価するが、これらの評価尺度は人が付与したカテゴリと同じであったかどうか、個々のインスタンス (本研究では文ごと) ごとに評価する。一方、社会学者は、個々の文にカテゴリを付与するが、それらの結果をカテゴリが付与されている割合を用いて分析する。そのため、自動推定結果の性能を F 値で示されても、それが実際に内容分析の結果として用いることができるかどうかはわからない。本研究では、この課題に対して、社会学者が実際に用いている分析手法により近いかたちの分類器の新しい評価デザインについて検討し、内容分析のタスクにあわせた評価手法 (Extrinsic evaluation) を提案した。その結果、ある程度の学習用データがあれば、LVM を用いた価値観の推定結果を内容分析に適用可能であることがわかった。

(4) 新しいテストコレクションの構築

新しいテストコレクションの構築に関しては、まず、対象とするテーマを検討した。台湾の研究協力者が原子力発電所に関する議論について、新聞記事を用いて内容分析を実施していた。原子力発電所に関する議論は、東日本大震災により福島第一原子力発電所の災害を機に日本でも議論が行われていること、また、世界にもその議論が波及している。国や言語の違いも含めて、これらの議論について人々がどのような価値観を持っているかを分析することは有益であると考え、このテーマを選択した。また、テキストの対象を新聞記事とした。日本の新聞記事の内容分析に関しては、新たに研究グループに加入してもらった。

日本語の新聞記事に対しては、2011 年から 2013 年の新聞記事を対象に 2 種のテストコレクションを構築した。一つは、“福島”と“原発 (または原子力)”を含む新聞記事の中から、ランダムに選択した 2,100 記事 (各年 700 記事) を対象に、価値観だけでなく、メディアフレーム、原発に対する立場等の他のフレームを構築し、それぞれのフレームに属するカテゴリを複数人によりコーディングした。もう一つは、同新聞記事の福島原発事故に関する社説 515 記事に対し、同様にコーディングを行った。

社説に関しては、メディアフレーム、価値観カテゴリによるコーディング結果を用いて、社会学者が言説の内容分析も行った。その結果、メディアフレームでは、全期間を通じて「政治判断」カテゴリが最も多いが、2012 年には「情報公開」カテゴリも増加し、2013 年になると「エネルギー」「環境保護」「安全保障」カテゴリなどが増え、論点が拡散することが明らかになった。一方、価値観フレームでは、多くみられたカテゴリは、2011 年は「効率性」、2012 年は「重要性」、2013 年は「社会福利」であり、変化していた。以上のように、経年により、議論の中心となるテーマやその時に重要と考えられている価値観が異なることがわかった。

(5) 分類器の汎用性の検証

開発した分類器の汎用性を検証するため、新しく構築したコレクションに対して、(1)で開発した分類器を適用した。2,100 記事に対してだけでなく、社説 515 記事に対しても行った。その結果、社説のテストコレクションを用いた場合、2,100 記事に比べて少ない記事数にも関わらずよい性能が得られた。また、SVM よりも高い性能を示した。しかしながら、ネットの中立性に関するテストコレクションを用いたときほどの十分な性能は得られなかった。

その原因を検討したところ、価値観カテゴリそのものをより洗練させることやコーディングプロセスを再検討する必要があることがわかった。コーディングプロセスに関しては、コーディング対象を記事単位から文単

位に変更し、コーディングプロセスを3段階にわけることとした。最初は分析対象となる記事かどうかの判定であり、次に記事の中でコーディング対象となる文(価値観が含まれている文)を判定し、最後にそれぞれの文に対して、コーディングするというプロセスとした。また、価値観カテゴリそのものの見直しも行った。今後は、新しいフレームを用いてコレクションの再構築を目指す。

5. 主な発表論文等 〔雑誌論文〕(計8件)

Ishita, E., Oga, T., Takayama, Y., Cheng, A.-S., Fleischmann, K. R., Oard, D. W., Tomiura, Y. "Toward Automating Detection of Human Values in the Nuclear Power Debate, Visual Presentation", Proceedings of the 80th Annual Meeting of the American Society for Information Science and Technology, pp.714-715, 2017, DOI 10.1002/pr2.2017.14505401127

大賀哲, 山腰修三, 三谷文栄, 石田栄美, 富浦洋一. 「福島原発事故をめぐるメディア言説の変容 - 毎日新聞社説(2011 - 2013年)を対象として」, メディア・コミュニケーション, vol.67, pp.119-148, 2017, <http://www.mediacom.keio.ac.jp/wp/wp-content/uploads/2017/03/9532fd5a7f63398164a59f804bdb04e0.pdf>

Takayama, Y., Tomiura, Y., Fleischmann, K. R., Cheng, A.-S., Oard, D. W., and Ishita, E. "Automatic Dictionary Extraction and Content Analysis Associated with Human Values", Information Engineering Express, Vol. 1, No. 4, pp. 107-118, 2015, <http://www.iaiai.org/journals/index.php/IEE/article/view/34/46>

Ishita, E., Oard, D. W., Fleischmann, K. R., Tomiura, Y., Takayama, Y., Cheng, A.-S., "Learning curves for automating content analysis: How much human annotation is needed?", Proceedings - 6th International Conference on E-Service and Knowledge Management (IIAI-AAI ESKM 2015), pp.171-176, 2015, DOI: 10.1109/IIAI-AAI.2015.295

Takayama, Y., Tomiura, Y., Fleischmann, K. R., Cheng, A.-S., Oard, D. W., and Ishita, E. "An Automatic Dictionary Extraction and Annotation Method Using Simulated Annealing for

Detecting Human Values", Proceedings - 6th International Conference on E-Service and Knowledge Management (IIAI-AAI ESKM 2015), 2015, DOI: 10.1109/IIAI-AAI.2015.268

Fleischmann, K. R., Takayama, Y., Cheng, A.-S., Tomiura, Y., Oard, D. W., and Ishita, E. "Thematic Analysis of Words that Invoke Values in the Net Neutrality Debate", Proceedings of iConference 2015, 6p, 2015. <http://hdl.handle.net/2142/73433>

Takayama, Y., Tomiura, Y., Ishita, E., Oard, D. W., Fleischmann, K. R. and Cheng, A.-S. "A Word-Scale Probabilistic Latent Variable Model for Detecting Human Values", Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (ACM CIKM 2014), p.1489-1498, 2014, DOI 10.1145/2661829.2661966

Takayama, Y., Tomiura, Y., Ishita, E., Wang, Z., Oard, D. W., Fleischmann, K., and Cheng, A.-S. "Improving Automatic Sentence-Level Annotation of Human Values Using Augmented Feature Vectors", Conference of the Pacific Association for Computational Linguistics (PACLING 2013), 6p. 2013.

〔学会発表〕(計2件)

高山泰博, 「テキストにおける人間の価値観の分析」, 平成29年度・シーズ発表会(～「人間中心」アプローチが生み出す地域イノベーション～), やまぐち事業化支援・連携コーディネイト会議, 山口県立大学, 日本

Takayama, Y., Tomiura, Y., Ishita, E., Wang, Z., Oard, D. W., Fleischmann, K. R., and Cheng, A.-S., "Automatic Annotation of Human Values for Sentences using Augmented Feature Vectors", Asian Summer School in Information Access (ASSIA 2013), Tsukuba, Japan.

〔その他〕

ホームページ

Research on Human Values

<http://nlp.inf.kyushu-u.ac.jp/values.html>

6. 研究組織

(1)研究代表者

石田 栄美 (ISHITA, Emi)
九州大学・附属図書館・准教授
研究者番号：50364815

(2) 研究分担者

富浦 洋一 (TOMIURA, Yoichi)
九州大学・システム情報科学研究科・教授
研究者番号：10217523

高山 泰博 (TAKAYAMA, Yasuhiro)
徳山工業高等専門学校・教授
研究者番号：30565841

(3) 連携研究者

大賀 哲 (OGA, Toru)
九州大学・法学学術院・准教授
研究者番号：90445718

(4) 研究協力者

Kenneth R. Fleischmann
米国・テキサス大学オースティン校

Douglas W. Oard
米国・メリーランド大学カレッジパーク校

An-Shou Cheng
台湾・国立中山大学