

レビュー文を対象としたあらすじ分類手法の評価

Evaluation of a Plot Classification Method for User Reviews

岩井 秀成 池田 郁 土方 嘉徳 西田 正吾
 Hidenari IWAI Kaori IKEDA Yoshinori HIJIKATA Shogo NISHIDA

大阪大学大学院 基礎工学研究科
 Graduate School of Engineering Science, Osaka University

多くのショッピングサイトにおいて、ユーザレビューが人気を集めている。しかし、小説や映画などのストーリーを持ったアイテムのレビューには、レビューの意見と同時に、ストーリーの内容（あらすじ）が書かれている場合がある。我々は、レビュー文をあらすじか否かに分類する手法を提案してきた。本研究では、レビュー文を対象としたあらすじ分類手法の評価を行う。

1. はじめに

Amazon.com や eBay のようなショッピングサイトでは、アイテムに対してレビューを簡単に作成・閲覧できる機能を提供している。レビューに書かれた他者の意見は、対象となっているアイテムの良し悪しを判断するのに役立つ。しかし、小説や映画などのストーリーを持ったアイテムに対するレビューには、レビューの意見と同時に、ストーリーの内容（本稿では「あらすじ」と呼ぶ）が書かれている場合がある。レビューによりあらすじがわかってしまうと、実際に小説や映画を見た時の楽しみや感動が減ってしまい、問題である。そこで本研究では、レビュー文を対象としたあらすじ分類手法の評価を行う。

2. 関連研究

我々の研究に最も近い研究分野として、テキストマイニングにおいて文書内の意見に着目した研究分野が挙げられる。意見の分類では、意見を肯定か否定かに分類することが注目されている。代表的な研究例として、Dave ら [1] や Wilson ら [2] の研究が挙げられる。Dave らは、Naive Bayes を用いてレビューの文を肯定または否定に分類している。Wilson らは、単語が本来持つ極性（肯定または否定）と、文脈での極性の両方を考慮し、レビューのフレーズごとに極性判定を行っている。

3. 機械学習によるあらすじ分類

単語を素性として、レビュー内の各文があらすじかどうかを分類するモデルを学習する。そして、このモデルを用いて各文に対しあらすじらしさを表すスコアを与える。機械学習アルゴリズムとして、学習されたモデルによってあらすじらしさを示すスコアを得やすいものを選択する。具体的には、Naive Bayes (NBayes), Support Vector Machine (SVM), ロジスティック回帰 (Logistic), 決定木 (D-tree), k-nearest neighbor 法 (k-NN) の 5 つを用いる。k-NN では、k を 1, 3, ..., 41 と試したが、最も結果の良かった 19 (コミック), 5 (小説), 23 (DVD) の時の結果を用いる。

各アルゴリズムにおいて最も F 値が高くなった属性数の値 (最適属性数と呼ぶことにする) で比較すると、いずれのカテゴリにおいても、NBayes, SVM, Logistic の順で良いことが

分かった。このことから、あらすじ分類に最も適したアルゴリズムは Naive Bayes であることが分かる。

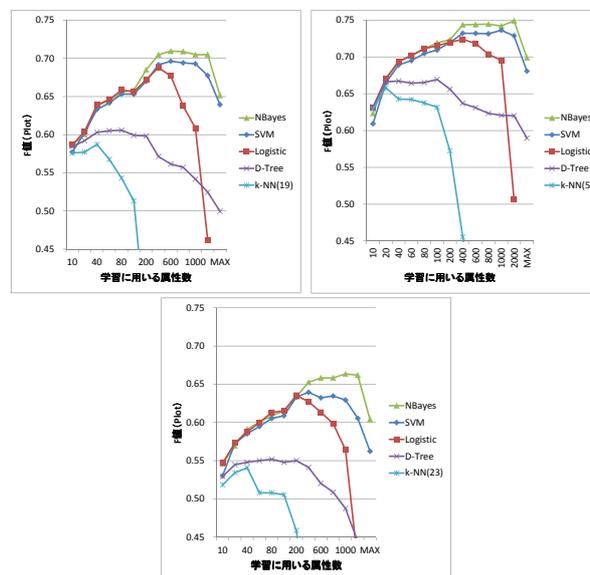


図 1: 各カテゴリ及び各アルゴリズムの F 値

4. あらすじ判定のための改善

この節では、機械学習の際にレビューの特徴を用いた工夫を行い、あらすじ分類の結果を改善することを考える。以下で、学習の際に施す改善策の詳細と、改善策を施した場合のあらすじ分類の評価結果について述べる。

4.1 人名と特有語の一般化

レビューを調査した結果、登場人物名が現れる文は比較的あらすじであることが多いということがわかっている。逆に著者名が現れる文は、あらすじになりにくいということがわかっている。しかし、登場人物名と著者名は多様に存在し、同じ語が使われることは少ない。そのため、これらの語が、分類に寄与するように学習されなかった可能性がある。改善策として人名 (登場人物名, 著者名, 俳優名とその他人名) の一般化を行う。すなわち、登場人物名 (または著者名) を指す固有名詞は、登場人物名 (または著者名) を表す抽象化したタグで表現するこ

とにする。DVD のカテゴリでは俳優名も獲得できるので、俳優名についても一般化を行う。また、レビューはアイテム紹介文に出現しない登場人物名をレビューに書くこともある。そこで、上記以外の人名を指す固有名詞についても、その他人名を表す抽象化したタグで表現する。

人名の一般化では、一般的な名前（主としてヨーロッパ系の名前）は抽出して一般化することができる。しかし、アジアやラテンアメリカなど、その他の地域の名前はカバーされていない。また、コミックにおいては、現実には使われていない名前も登場する（例えば、「ドラえもん」や「ニモ」など）。このような名前は、辞書が整備されていない上に、次々に新しい名前が誕生するため、対応することが難しい。そこで、あるアイテムへのレビュー文にしか出てこない語を、そのアイテムへの特有の単語として一般化することにした。このようなアイテム特有の単語には、上記のような特有の登場人物名が含まれるかもしれない。全 100 アイテムに対するレビューにおいて、複数のアイテムに出てきた単語を非特有語としてリスト（非特有語リストと呼ぶ）に登録した。次に、レビュー中の各文において、それぞれの単語を調べ、非特有語リストに登録されていなければ、特有語を表すタグで置き換えた。

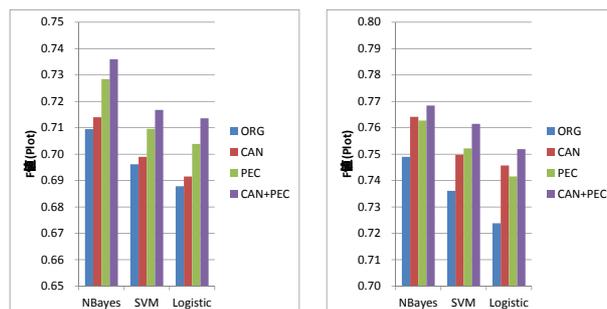
4.2 評価

前章の学習アルゴリズム間の比較実験の時と同じデータセットを用いて評価した。また、属性数を変化させて、単語の一般化なしの場合（“ORG”と表記）、人名（登場人物名、著者名、俳優名とその他人名）を一般化した場合（“CAN”と表記）、特有語を一般化した場合（“PEC”と表記）、人名と特有語を一般化した場合（“CAN+PEC”と表記）の 4 つを比較した。それぞれの最適属性数の時の結果を図 2 に示す。

まず、人名の一般化（CAN）に注目する。コミックと小説のカテゴリにおいては、ORG よりも CAN の方が F 値が高いことが分かる。また、コミックと小説を比較すると、小説の方が一般化の効果が大きいことが分かる。これは、小説の方が一般的な名前の登場人物が多いためと思われる。コミックは、辞書に載っていない名前の影響で、人名の一般化の効果があまり高くなかったと思われる。コミックと小説に対して、DVD は一般化の効果がほとんど見られない。これは、DVD の商品紹介文には非常に短いものが多かったことと、長かったとしても登場人物名が書かれていない場合がほとんどであったことが理由である。よって、人名の一般化は、コミックカテゴリと小説カテゴリで有効であることが分かった。

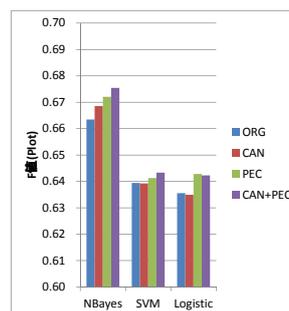
次に、特有語の一般化（PEC）について注目する。小説カテゴリでは、PEC は ORG よりも F 値が高いが、CAN とほとんど差が見られない。一方、コミックと DVD カテゴリにおいては、PEC は CAN よりも F 値が高い。このことから、小説カテゴリでは人名の一般化が十分に機能していたと思われるが、コミックと DVD では人名と認識されていなかった人名が多くあったと思われる。よって、特有語の一般化は、コミックカテゴリと小説カテゴリで有効であることが分かった。また、コミックカテゴリと DVD カテゴリにおいては、特有語の一般化は人名の一般化よりも、高い F 値が得られることも分かった。

また、特有語と人名と両方を一般化した場合（CAN+PEC）は、全てのカテゴリにおいて、最も F 値が良くなることが分かった。特に、コミックカテゴリと小説カテゴリでは、有効であることが分かった。



(a) F 値 (comic)

(b) F 値 (novel)



(c) F 値 (DVD)

図 2: 一般化（最適属性数）

5. 結論

本研究では、ストーリーを持ったアイテムに対するレビュー文を対象としたあらすじ分類手法の評価を行った。あらすじの判定には 5 つの代表的な機械学習アルゴリズム（Naive Bayes, サポートベクターマシン, ロジスティック回帰, 決定木, k-nearest neighbor 法）を適用した。結果として、Naive Bayes が最も良い結果を示した。また、あらすじかどうかの判定のための改善策として、人名（登場人物名、著者名、俳優名とその他人名）と特有語の一般化を行った。結果として、人名と特有語の一般化を行うことがあらすじかどうかの分類に有効であることが分かった。今後は、あらすじ分類に基づきあらすじを隠してレビューが見られるシステムで被験者実験を行い、ユーザ満足度の向上につながるかどうかを検証する予定である。

参考文献

- [1] Dave, K., Lawrence, S., Pennock, D.M.: Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. *Proceedings of The Twelfth International World Wide Web Conference (WWW'03)*, pp. 519-528 (2003).
- [2] Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP'05)*, pp. 347-354 (2005).