

図書を推薦するツイートの押し付けがましさに関する特徴の分析

矢田 竣太郎[†] 影浦 峯[‡]

東京大学大学院[†]教育学研究科,[‡]学際情報学府
{shuntaroy, kyo}@p.u-tokyo.ac.jp

1 関心と目的

特定の事物について消費者間でやり取りされるクチコミ (Word of Mouth; WoM) は、その事物への関心を刺激し、特に対象が商品であった場合にはその購買をも刺激する有用な情報資源としてマーケティングで活用されている [1]。精緻化見込みモデル [2] などの消費者の意思決定モデルによれば、対象の事物について消費者自身が評価や吟味できるほどの知識・技術を持っていない場合、他者による評価など、その事物に関する周辺情報に基づいて購買判断を下すとされている。このことは、WoM が当該事物の潜在的な関心層に働きかける際に特に有益であることを示唆する。検索システムや推薦システム、あるいは質問応答・対話システムにおいて、特定の事物についての詳細な情報を提示することは主な機能の 1 つであるが、レビューサイトやソーシャルメディアの普及により、クチコミはオンライン上のテキスト (electronic Word of Mouth; eWoM) として収集しやすくなっており、情報システムの提示する情報源として有用であると考えられる。例えば筆者らは、読書に好意的であるものの実際の読書頻度は少ない層に向け、図書の eWoM (例 1) を提示する図書推薦システムを開発している [3]。

例 1: 絡新婦の理はくるおいしいほど好きです シリーズダントツで好きです 好きすぎて好きしかいえない

eWoM の形態は様々であるが、中には直接的な表現で事物を推薦するものも見受けられる (例 2)。

例 2: ちなみに小説じゃないけど「ゾンビサバイバルガイド」は定期的に読み返したくなる。タイトルそのままに、ものすごく真面目な本。おすすめ一冊。 #読書

事物を推薦する度合い (推薦強度) にも幅があり、大きくなると命令や強制へと変化することもある (例 3)。

例 3: 全人類は星の王子様読め。よくわかんなくてもいいから読め。なんとなくわかったかな程度で終わってもいいからとりあえず読め。

推薦強度という観点から現象を連続的に捉えるために、こういったケースも推薦の一種と考えると、それらは受け手にとって押し付けがましいものであり、eWoM

を情報源としてシステムのユーザに提示する場合に問題となる。実際、事物を強く推薦することは一種の説得と捉えられるところ、説得理論において押し付けがましきは受け手の反発を招くと指摘されている (心理リアクタンス理論 [4])。したがって eWoM を利用するシステムはその推薦の強度を推定できることが望ましいと考えられるが、eWoM における推薦表現の強度に関わる要素についての知見は十分ではない。

そこで本研究は、主に筆者らが開発する図書推薦システムへの応用を視野に、eWoM に見られる推薦強度とその特徴を分析することを目的とする。ここで、例 3 の表現を押し付けがましく感じるのは、「読め」という命令形の表現に加え、その対象が「全人類」と広いことも関係していると思われることから、推薦する相手 (推薦対象) についても検討した。筆者らのシステムは eWoM の関心の対象として図書を扱うほか、eWoM の中でもソーシャルメディア、特に Twitter のツイートに絞っている¹。すなわち「図書に言及するツイート」(Tweets that Mention Books; TMB)、およびそのうちの「図書を推薦するツイート」(Tweets that Recommend Books; TRB) を本研究の具体的な対象とする。得られる知見は TMB に限定されるが、図書以外のメディアコンテンツについてのツイートにも一定の示唆を与えられるものと考えられる。

2 関連研究

ソーシャルメディア上でユーザが動画などのマルチメディアコンテンツについて推薦する意図を調査した研究 [5] によれば、主観的規範、プラットフォームの信頼性、使いやすさ、生産性が影響しているという。しかし推薦表現については、医療や法律の分野のガイドラインに見られるものを分析する研究があるものの [6, 7]、eWoM を含めソーシャルメディア上の推薦表現は扱われていない。

¹システムのユーザの社会資本から発信されたメッセージの方が見知らぬ他者の eWoM より良いはずとの考えからソーシャルメディア上のものを、メッセージとして短い方が潜在的関心層の情報処理負担が小さく受容されやすいはずという考えから Twitter を選択した。

3 データ

3.1 図書に言及するツイートの収集

本研究は、日本語の TMB のうち、識別可能な特定の図書に言及しているものを対象とし、図書一般や特定の図書集合についてのみ言及するものを対象外とする。なお「図書」の範囲は、すでに刊行されたかあるいは刊行される予定のある本を広く含め、電子書籍・オーディオブックなどの媒体を問わないものとするが、ISBN を付与され公共に流通しているものに限定する。加えて、Twitter 利用者による自然な推薦表現に関心を持つことから、自動化された手法によって生成されたツイートを対象から除いた。具体的には、ツイート内容をプログラムする機能が搭載されていない、Twitter 社から公式に提供されている Twitter 投稿用アプリケーションによるツイートに絞った。

TMB は Twitter REST API で読書に関連したハッシュタグを検索することで収集した。使用したハッシュタグは #読書, #読了, #書評, #本の 4 種で、図書に言及するツイートに特有の語彙分析を参考にした [8]。2018 年 6 月 16 日から 8 月 5 日にかけて収集した 1 万件のうち、目視によるアノテーションにより 8198 件の TMB を得た。このうち検索に使用したハッシュタグの出現頻度は #読書 2426 件、#読了 4626 件、#書評 411 件、#本 1517 件であった。

7964 件の TMB の中には、複数の図書に言及しているものがある。推薦表現は個々の本ごとに異なりうるため、TMB と言及された図書が 1 対 1 になるように TMB 側を複製した。ここでは便宜的に、TMB と言及図書の対をレコードと呼ぶ。8871 件のレコードに対し、次項で説明するラベリングを実施した。

3.2 図書を推薦するツイートのラベリング

TMB の中に、図書を推薦するような表現（推薦表現）が含まれる場合、それを TRB とし、以下の項目をラベリングする。

推薦表現 図書を推薦するような表現それ自体を抽出する。このとき、後述する推薦強度の解釈の一助になるため、強意表現などの修飾部分も含めた。（例：“宮本輝編の『魂がふるえるとき』のラインナップのガチさは諸手を上げておすすめしたい”）

推薦強度 推薦表現における推薦度合いの強さを次の 3 段階から選ぶ。

弱 読むことを勧める（例：おすすめ、読んでほしい、読んだ方がよい）

強 読むことを強く勧める・強制する（例：読むべき、読まなければならない、必読、絶対に～読んで欲しい）

過度 読むことを不快なまでに勧める。例えば：

- 読まないことを非難したり侮辱したりする（例：読まない人はどうかしている）
- 同調圧力を与える（例：これくらいみんな読んでよね）

推薦規模 図書をどの規模の人々に推薦しているか、次の 3 種類から選ぶ。

特定人物 名指しで特定される人物だけに推薦している（@リプライなど）

特定集団 特定の組織に属する人々や何らかの性質・属性を持つ人々を集合的に指して推薦している（例：“「ロックで独立する方法」は独立開業する人は必ず読むべき本です”）

不特定多数 特に推薦対象を示していないか、または全ての人に推薦している（例：“魚住直子先生の「超・ハーモニー」、みんな読んで。”）

推薦対象 推薦の対象となっている人々を抽出する。

特定人物 @リプライ形式であればそのユーザ名、そうでなければ該当する人物表現を抽出する

特定集団 上記の例では「独立開業する人」

不特定多数 上記の例では「みんな」

4 結果と分析

4.1 各ラベルの集計

8871 件の TMB レコードのうち、TRB レコードは 451 件 (5.1%) だった。推薦強度と推薦規模、および推薦対象のクロス集計を表 1 に示す。TRB レコードの推薦強度は、強くなるほど急激に少なかった。大部分は控えめな推薦だったということになる。推薦規模は不特定多数がおよそ 6 割を占め、残りはほぼ特定集団であった。強度と規模の関係に目立った特徴はない。

次に、推薦対象の抽出結果を確認する。TRB 全体

表 1: 推薦強度と推薦規模に付与されたラベルと推薦対象の抽出数のクロス集計（かつこ内の数値が推薦対象の抽出数）

推薦規模	推薦強度			計
	弱	強	不快	
特定人物	5 (5)	0 (0)	0 (0)	5 (5)
特定集団	160 (159)	20 (20)	6 (6)	186 (185)
不特定多数	219 (34)	35 (2)	6 (0)	260 (36)
計	384 (198)	55 (22)	12 (6)	451 (226)

のうち、推薦対象が明示的に示されていたのは約半数の 226 件である。特定人物と特定集団からは 1 件を除き²全て対象が抽出されており、不特定多数のうち推薦対象が抽出されていたのは 36 件だったことから、TRB レコードの約半数は対象を明示せず、フォロワーを漠然と想定するなどして推薦していることがわかる。特定人物を対象とする推薦が少なかったが、これは TMB をハッシュタグの検索によって収集したためと考えられる。特定人物の主なケースとして想定していたリプライ形式のツイートには、Twitter 内のトピック検索のための明示的なキーワードとして主に利用されるハッシュタグは共起しにくいといえる。

4.2 TRB の文字列長

TRB がどの程度の情報量を持つのか、本文に含まれる文字数 (URL とハッシュタグ、リプライ文字列を除く) によって近似的に計測した。推薦強度や推薦規模による TRB の長さは、図 1 に示すように、推薦強度が大きくなるほど TRB が短くなる傾向が見られた。同様に推薦対象と推薦表現の文字列長を計測すると、推薦強度が大きくなるほど推薦表現が長くなる傾向があった (図 2)。そこで、推薦対象や推薦表現を差し引いた文字列長を「推薦内容」(推薦理由などが含まれた部分) の推定値と考え、推薦強度に応じてその文字列長が変化するか確かめたところ、図 3 のようになり、推薦強度が大きいときと推薦規模が小さいときに推薦内容が少ないことが示された。

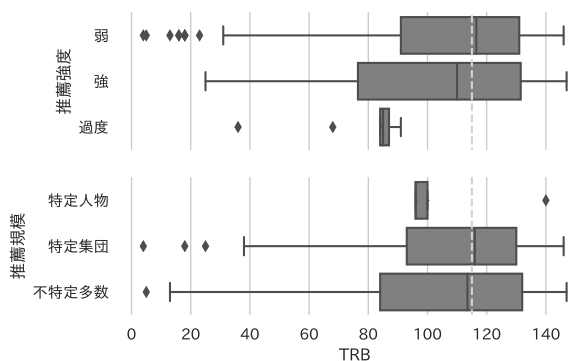


図 1: 推薦強度・推薦規模ごとの TRB 文字列長 (箱ひげ図; 点線は全 TRB における中央値)

4.3 推薦対象の類型化

推薦対象にどのような例が見られたかについて、推薦規模ごとに類型化を試みた。なお**特定人物**を対象す

²1 件の TRB が読書感想文向けの図書を推薦しており、暗黙に生徒・児童を対象としていた。

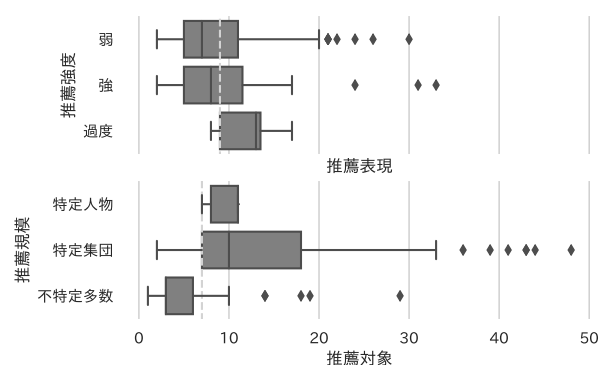


図 2: 推薦強度ごとの推薦表現の文字列長・推薦規模ごとの推薦対象の文字列長 (箱ひげ図; 点線は全 TRB におけるそれぞれの中央値)

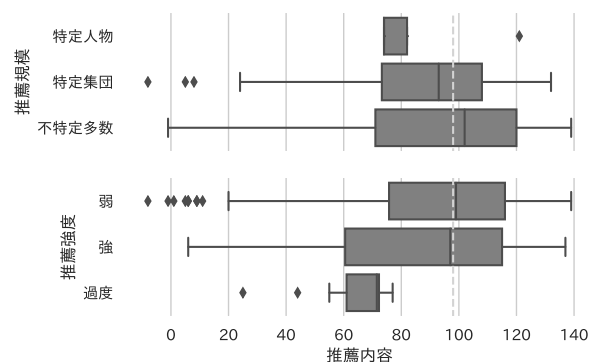


図 3: 推薦強度・推薦規模ごとの推薦内容文字列長 (箱ひげ図; 点線は全 TRB における中央値)

る場合は全て個別のユーザであるため省略する。まず推薦規模別の類型を示す。**特定集団**を対象とする場合の表現は次のように分けることができた。

未読者・不読者 “まだの方”, “未読の方”, “読んでない方”, “今、読みたい本がない、見つからないという人”, “読書慣れしてない人”, ...

特定の人口 “ちびっ子”, “10~20 代の人”, “高校生以上”, “男性”, ...

特定の関心を持つ層 “時代物が好きな人”, “ロボアニメ好き”, “他人の恋愛を覗き見してみたい人”, ...

特定の状況下にある層 “夏休みに入る前・入った人”, “「働く」ということに何らかの不安感ある人”, “生きづらい人、思考が堂々巡りして行き詰まっている人など”, ...

何らかの条件で絞っているとはいえ、集団は大きく指定されていることが見受けられる。次に、**不特定多数**において対象を宣言する場合の表現は、ほとんどが一般の全称表現 (“みんな”, “みなさん” など) であったが、それ以外について以下の 2 パターンを識別できた。

広範囲の層 “幅広い年代の方”, “性別・年齢問わず、いろんな人”, “多くの人/たくさんの人”, ...

ある集合とその補集合 “山に登る方にも登らない方にも”, “数学嫌いな人も数学好きも”, “大人も子供も”, “理系に興味のある人もない人も”, ...

4.4 推薦表現の類型化

最後に、抽出された推薦表現を分析する。推薦強度ごとに類型化を試みた。弱の強度では、大きく「提案」「希望」「推薦」「満足度の示唆」に分けられた。

提案 “ぜひご一読を!”, “読んでみてください!”, ...

希望 “読んでもらいたいです”, “読んでほしい一作です”, ...

推薦 “お薦め”, “推薦図書です”, “推しである”, ...

満足度の示唆 “読みやすいかも!”, “楽しめる小説だと思う。”, ...

強の強度は、弱で見られた表現に強意表現を付与したものと、命令表現に大別できた。

弱の表現+強意表現 “超オススメの一冊です!”, “本当におすすめです!!”, “読んでほしい!絶対!”, “とにかく読んでほしい一冊。”, “映画より断然おすすめ!”, ...

命令表現 “絶対に読んだ方がいい!”, “絶対読むべき”, “読んでください!”, “必読”, “この本を読まないとな!””, “読んで。”, ...

過度のラベルを付与された表現は、ほとんどが同調圧力をかける表現だった: “買いに行けるね!注文出来ますよ!”, “迷ってないで早く読まないんだよ”, “早く読まない!!”, ...。他に、読まない人を貶める表現 (“知らないのはヤバイよ、”) が1件と、攻撃的な表現 (“鼻先に突きつけたくなった。”) が1件あった。

5 考察

本研究では図書を推薦するツイート (TRB) を対象に、推薦の度合い (推薦強度) に応じた特徴を探索的に分析した。主な知見は (1) 母集団である図書に言及するツイートのうち約5%が TRB である、(2) TRB の約半数は対象非明示の弱い推薦である、(3) 推薦強度が強くなると、説明のための表現よりも推薦のための表現に文字数が使われる、および (4) 推薦対象と推薦表現はいくつかの類型にまとめられる、の4点である。

本研究はハッシュタグでツイートを収集したが、これによる言語表現上の偏りがあったことは否めない。本稿では省略するが、データセット中には「#読書好きと繋がりたい」などのハッシュタグが多く含まれ、

一般的なツイートよりもフォローネットワーク外への指向性もともと強い母集団であった可能性がある。

TRB から推薦強度を推定する応用を考えると、「おすすめ」や「読んでほしい」といった頻出表現と、そこに係る表現や隣接する「!」などの記号が手がかりになりそうである。しかし、押し付けがましいもの (過度の推薦強度) については本研究では少数しか集められなかった。推薦強度が弱や強のもの比べると表現の多様性が見られる可能性があり、収集方法の工夫が必要と考えられる。

謝辞

本研究は科学研究費補助金挑戦的萌芽研究「オンラインを介して『前読書家』の読書を触発する方式・環境の開発」(課題番号 16K12542) の補助を受けた。また、ツイートデータのアンテーションは株式会社バオバブ (<https://baobab-trees.com/>) に依頼した。

参考文献

- [1] 池田謙一 (2010) 『クチコミとネットワークの社会心理: 消費と普及のサービスイノベーション研究』, 東京大学出版会.
- [2] Petty, Richard E and John T Cacioppo (1988) “Communication and Persuasion: Central and Peripheral Routes to Persuasion,” *The Public Opinion Quarterly*, Vol. 52, No. 2, pp. 262–265.
- [3] Yada, Shuntaro (2014) “Development of a Book Recommendation System to Inspire ‘Infrequent Readers’,” in Tuamsuk, Kulthida, Adam Jatowt, and Edie Rasmussen eds. *The Emergence of Digital Libraries – Research and Practices*, Vol. 8839 of LNCS: Springer International Publishing, pp. 399–404.
- [4] Brehm, Sharon S. and Jack Williams Brehm (1981) *Psychological reactance: a theory of freedom and control*: Academic Press, pp.432.
- [5] Chang, Shuchih Ernest, Wei-Cheng Shen, and Chun-Hsiu Yeh (2017) “A comparative study of user intention to recommend content on mobile social networks,” *Multimedia Tools and Applications*, Vol. 76, No. 4, pp. 5399–5417.
- [6] Georg, Gersende, Hugo Hernault, Marc Cavazza, Helmut Prendinger, and Mitsuru Ishizuka (2009) “Analysing Clinical Guidelines’ Contents with Deontic and Rhetorical Structures,” in Combi, Carlo, Yuval Shahar, and Ameen Abu-Hanna eds. *Artificial Intelligence in Medicine*, pp. 86–90, Berlin, Heidelberg: Springer Berlin Heidelberg.
- [7] López, Anna Samaniego and Raquel Taranilla (2012) “Contrastive analysis of the linguistic expression of recommendation in two legal genres,” *Ibérica*, Vol. 23, No. 2012, pp. 65–88.
- [8] 矢田竣太郎・影浦峯 (2016) 「図書に言及するツイートの抽出: 素性・データ量・手法の効果に関する考察」, 『電子情報通信学会技術研究報告言語理解とコミュニケーション (NLC)』, pp. 29–34.