

アカウント推薦のための Tweet 履歴の視覚化

山田 紘司[†] 平瀬 壮大[†] 延澤 志保[†]

[†] 東京都市大学知識工学部情報科学科

1. はじめに

近年, Twitter のフォローユーザ推薦の推薦精度向上に関する検討が広く行われている[1]. 代表的な手法として, 興味の対象の類似度でアカウントの推薦を行う手法がある[2]. 本研究では, 興味の対象に対するアカウント固有の感情極性の推定と視覚化を行う手法を提案する.

2. 感情極性推定手法

2.1 提案手法の概要 対象アカウントのツイート履歴を一定期間ごとに分けた上で, 重要度に基づいてそれぞれの期間のキーワードの感情極性を推定することで, 期間ごとにキーワードに色を付けて出力する. キーワードの重要度は TF-IDF 法を基に算出し, キーワードを出力する際のフォントサイズに反映させる(図 1).

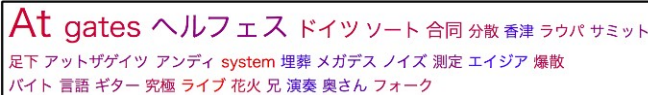


図 1 1 期間分のキーワード群出力例

2.2 ツイート履歴分割期間 キーワードの推定に最低 250 個程度のツイートが必要なこと, ツイート内容の変化の視覚化には数ヶ月単位の分割が適することから, 分割期間を新年および新年度の始まりを考慮して 3 ヶ月ごととする.

2.3 文の感情極性の推定 キーワードの感情極性はキーワードを含む文の感情極性を基に推定する. 手作業で 344 文に対して感情極性を判定したところ, 感情極性が判定できた 255 文のうち 81% が文末を用いて判定していることがわかった. 判定に要した文末語数の平均が 2.6 語であること, 文末 3 語までで判定できた文が 80% を占めたことから(図 2), 文の感情極性は文末の 3 単語で推定を行う.

感情辞書[3]には各語について「楽しい/悲しい」「嬉しい/怒り」「のどか/緊迫」の 3 本の感情軸に沿って感情値が 1 ~ 0 の範囲で記述されている. ここでは「楽しい, 嬉しい, のどか」をポジティブとみなし, キーワードごとにそのキーワードを含む文すべての感情極性を 3 軸の平均で推定する.

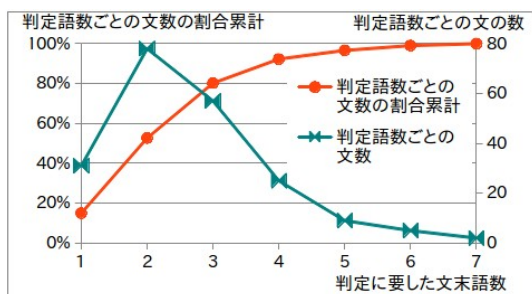


図 2 感情極性判定に要する文末語数ごとの文の数

2.4 キーワードの感情極性の推定 それぞれのキーワードを含む文のうち, 感情極性がポジティブな文の割合を各感情軸でのキーワードの感情極性値とし, 3 軸の平均をキーワードの感情極性値とする.

2.5 感情極性の視覚化 L*a*b 表色系上では, ポジティブな感情は黄から赤, ネガティブな感情は青から緑, 感情の無いものは中心に分布する[4]. これを基にポジティブを橙, ネガティブを青緑とすると中央値は感情無しとなる(図 3 左)が, 感情極性値が中央値となるのはポジティブとネガティブの数が同じ場合であり, 感情が無いわけではない. そのため軸を平行移動して赤から青とし, 中央値は紫で表す(図 3 右).

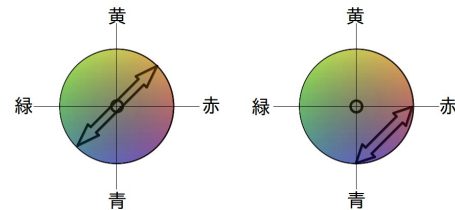


図 3 色と感情極性の軸

3. 実験結果と今後の課題

連続する 4 期間計 1,256 ツイートを用いてキーワード群の視覚化の実験を行ったところ, キーワード群の感情極性が一致または一方が中央値の平均は 78% と良好なものであった(図 4). ポジティブとネガティブの双方の文脈で出現するキーワードについて, その定義や扱い等, 今後さらに検討を行う予定である.

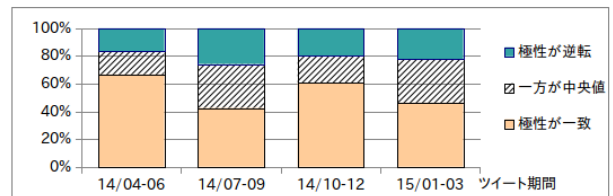


図 4 4 期間のキーワード群の感情極性の推定結果

参考文献

[1] 久米雄介, 打矢隆弘, 内匠逸, “興味領域を考慮した Twitter ユーザ推薦手法の提案と評価”, 情処研報, 2015-ICS-179(1), pp.1-8, 2015.
 [2] 土方嘉徳, “嗜好抽出と情報推薦技術”, 情報処理学会学会誌, Vol.48, No.9, pp.957-965, 2007.
 [3] 熊本忠彦, 河合由起子, 張建偉, “ユーザ印象評価データの分析に基づく印象”, 信学論, Vol.J94-D, No.3, pp.540-548, 2011.
 [4] 山下真知子, “感情からイメージされる色彩について: 一高齢者の回復期ケアを目的とした施設空間の色彩設計に関する研究— Color Images Evoked by Human Feelings”, 大手前大学論集 Vol9, pp.289-316, 2008.