

マイクロブログに投稿される内容と天気コンテキストの関係の検証

近年、人々は様々な種類の問題を解決するため、日常生活において様々なメディアから推薦情報を受信する機会が増加している。たとえば、Amazon.com社では、顧客の購買に伴うあらゆる行動に基づき顧客の興味を推測、関連する商品が推薦される。このような推薦を行うためには、ユーザの趣味嗜好を知る必要があり、もっとも単純な手段としてはPOSデータを利用することが考えられるが、全てのサービス提供者が平等に情報を得られるわけではない。

そこでTwitterなどのマイクロブログからユーザの趣味嗜好を推定する研究が行われており、同じ単語であっても異なる目的がある場合が存在し（「lose weight」という単語のトピックが、「beauty」である場合と「health」である場合がある）、それを判別するトピックモデルと呼ばれる、単語をトピッククラスターに分類するモデルの研究が存在する。さらに、ユーザのトピックは天気というコンテキスト（文脈情報）によって大きく左右すると考えられるため、目的を「トピックモデルを用いて、マイクロブログに投稿される内容と天気コンテキストの関係を検証する」こととする。気温のコンテキストとトピック、トピックと単語との関係の例をFigure 1に示す。また、結果のトピックをFigure 2に示す。

Key Words: context-aware, recommendation, topic model, weather-context, Twitter

Reference

[1] 伊藤 拓, 深澤 佑介, 朱 丹丹, 太田 順, Tweet内容に影響を与える気象条件と特徴語の抽出, 情報処理学会, 2014-MBL-73, No.1, 2014.

[2] Taku Ito, Yusuke Fukazawa, Dandan Zhu and Jun Ota, Climate Condition that Mostly Affects the Change of the Tweet Content, International Conference on Mobile Computing and Ubiquitous Networking (ICMU2015), 20 Jan 2015.

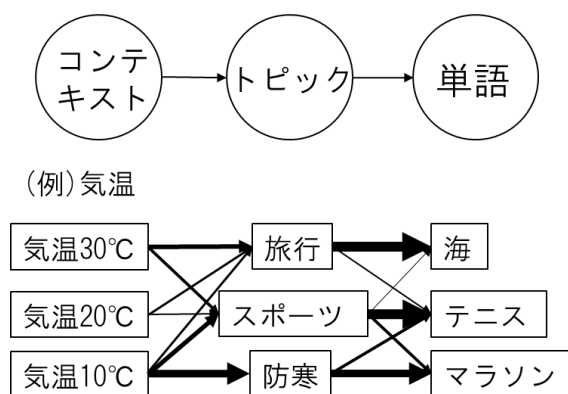


Figure 1: Relationship between temperature-context, topics, and words

22度付近 9月上旬のトピック

28°C付近 9月上旬のトピック

単語	重み係数
ヨーグルト	0.150690
ピッツァ	0.082197
食パン	0.068498
サバ	0.068498
漬物	0.054799

単語	重み係数
抹茶	0.191786
牛肉	0.123292
かき揚げ	0.095895
ハチミツ	0.082197
ソフトクリーム	0.041101

Figure 2: Topic clusters of different weather