

## うつ傾向を持つ人のツイートとブログ記事の対応分析

大澤 翔吾

近年、日本をはじめ世界中でうつ病患者の数が増加しており、2014年の時点で112万人の患者がいる。うつ病は最悪の場合、患者を自殺に追い込むこともあり早急な解決策が求められている。そこで、ソーシャルメディアのテキストからそのユーザの「うつ傾向」を推定する研究が進められている。テキストからユーザのうつ傾向を読み取ることで、うつだと疑われる患者の早期発見や早期治療に役立つと考えられる。

先行研究では一つのソーシャルメディアに焦点を当てて「うつ傾向」を持つユーザと持たないユーザを比較してうつ傾向を推定する研究は存在するが、うつ傾向をもつユーザをソーシャルメディア間で比較した研究はない。Twitter とブログで語彙の違いが出てくることで、医師などがうつ病患者を診断する際、患者の知りたい情報について Twitter とブログのどちらを参考にすればいいのか、役立てると考えた。そこで本研究では、うつ傾向をもつユーザのツイートとブログ記事に焦点を当て、出てくる語彙の特徴を比較した。その際、Twitter とブログの各ユーザと語彙の比較、テキスト量ごとにグループ分けした際の比較、Twitter とブログの使い始めた時期ごとにグループ分けした際の比較をおこなった。

「うつと診断された」というツイート、病院でうつと診断されたというブログのプロフィールをうつ傾向を持つユーザとし、Twitter ユーザ 50 件、ブログユーザ 22 件を集め、集めた。集めたテキストを形態素解析にかけ、名詞と形容詞を出し、数字や記号を抜き出し出現頻度上位 50 件の語彙を対象とした。

結果、データ量でユーザをグループ分けした場合のうつ傾向を持つ Twitter・ブログのユーザと語彙の比較した場合、ブログ記事の多いグループで「仕事」、ブログ記事が少ないグループで「病院」という語彙が見られ、Twitter のデータ量の差で表れる語彙の違いは見られなかったが、「障害」や「疾患」など自分の症状を表す語彙や「良い」「悪い」などの形容詞が多く表れた。このことから、ブログの記事では「仕事」や「病院」といった自分の日常を表す記事が書かれ、Twitter では「障害」や「発達」といった自分の疾患や「良い」「悪い」などの自分の気分や状態を表すツイートが書かれていることがわかった。

今後の課題としてユーザ数や一人当たりのデータ件数を増やすことで、研究の精度をより高めたい。

(指導教員 歳森敦)