



HOKKAIDO UNIVERSITY

Title	A Study on Deep Learning-based Humor Detecting Methods for Sentiment Analysis of Social Media [an abstract of dissertation and a summary of dissertation review]
Author(s)	栗, 達
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第14177号
Issue Date	2020-06-30
Doc URL	https://hdl.handle.net/2115/79058
Rights(URL)	https://creativecommons.org/licenses/by/4.0/
Type	doctoral thesis
File Information	Da_Li_abstract.pdf, 論文内容の要旨



学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称 博士（情報科学） 氏名 栗 達

学 位 論 文 題 名

A Study on Deep Learning-based Humor Detecting Methods for Sentiment Analysis of Social Media
(ソーシャルメディアにおける感情分析のための深層学習を用いたユーモア検出手法に関する研究)

情報通信の重要なファクタの一つであるスマートフォンは近年急速なスピードで普及しており、それに伴いソーシャルメディア (SNS) は私たちの生活に不可欠なものとなり、その中には感情情報が多くふくまれている。「Weibo」とは、中国語圏最大のソーシャルメディアであり、中国本土のみならず、日本やアメリカなど中国語話者がいる地域で幅広く利用されている。現在、自然言語処理の分野においてソーシャルメディアの感情分析は研究テーマとして注目を集めている。しかし、従来手法ではテキストのみを考慮し、絵文字などはノイズとして除外される場合が大半であった。さらに、ユーモアという感情表現が人の生活において多様な役割を果たすことが明らかとなってきた。近年、コンピュータがユーモアを理解することが可能となり、その有用性が認知されてきている。先行研究によると、SNS のユーザは絵文字を使用することで、ユーモアの感情を表す傾向が見られる。そのために、ソーシャルメディアの感情分析に着目し、絵文字を使用したユーモアの投稿を中心に、本研究を行うことは有効である。

絵文字は、感情のグラフィカル表現としてソーシャルメディアで広く使用され、重要な役割を果たしている。絵文字と同様に、インターネットスラングは、日常のオンラインコミュニケーションで使われる非公式の言語であり、感情表現として多く用いられている。スラングや絵文字がソーシャルメディアの感情分析に与える影響を十分に理解することが重要であると考えられる。したがって、本研究は絵文字の感情極性アノテーションを行い、SNS によく使用されているスラングとフェイシャル絵文字を分析し、それぞれのレキシコンを構築した。さらに、絵文字を考慮したソーシャルメディアにおける感情分析のための深層学習を用いたユーモア検出手法を提案したことが本研究の新規性である。

評価実験により、以下に述べる 3 点の絵文字極性、絵文字レキシコンとスラングレキシコンの有効性及び提案手法の有効性が確認された。1) 絵文字極性の有効性を検証するために、絵文字の極性を考慮した感情分析モデル再帰型ニューラルネットワーク Long Short-Term Memory (EPLSTM) を提案した。実験結果により、従来手法と比べ、提案手法が感情極性を予測する性能が大幅に改善することが確認された。2) 感情分析の精度を向上するために、構築した絵文字とスラングのレキシコンを形態素解析ツールに応用し、感情分析の機械学習モデルを提案した。サポートベクターマシン (SVM)、ロジスティック回帰 (LR)、単純ベイズ分類器 (NB) などの機械学習モデルを利用し、感情分析において、構築したスラングレキシコンと絵文字レキシコンが有効であることを検証した。3) 絵文字辞書とスラング辞書を用い、SNS データを深層学習モデル Attention-based Bi-directional LSTM に学習させ、細粒度ユーモア検出システム「HEMOS」(Humor-EMOji-Slang) を提案した。実験の結果、従

来手法である深層学習モデル LSTM や textCNN と比べて, 提案手法を用いることで, F 値が 6.86 ~ 20.59 ポイント上回ったことが確認された. 有意差検定により, p 値は 0.05 以下になっていることから, 有意差が認められることが確認された.

本研究により, 絵文字に関する分析の有効性及び絵文字とスラングを考慮したソーシャルメディアにおける感情分析のための深層学習を用いたユーモア検出手法の有効性が確認された.