

ソーシャルメディアを用いた深層学習による Fake News の検出

富平 準喜

近年、ソーシャルメディアの発達によってニュースは迅速に広まり、リアルタイムで情報を得られるようになってきた。以前は、テレビや新聞といったマスメディアが、第三者機関として情報の真偽を見極めた上でニュースを発信してきたのに対して、ソーシャルメディアを通して情報の真偽が明確でないニュースまでもが流通するようになった。このようなソーシャルメディアを通して発信される嘘が混在した情報は Fake News と呼ばれる。ソーシャルメディアを通して広まった Fake News はユーザを戸惑わせ、結果として社会の混乱を招く可能性があり、社会問題として考えられている。

Fake News は正しい情報と比べて伝達が早いいため、人手での確認では Fake News の拡散を食い止めることができない。ニュースの真偽を自動で判別する技術が求められ、Fake News の自動検出の研究のニーズが高まってきている。Fake News はソーシャルメディアによって拡散され深刻化していることが考えられるため、本研究ではソーシャルメディアで拡散される Fake News に着目し、ニュースを引用しているツイートを用いることで Fake News を自動判定する深層学習モデルの構築を目的とする。

Fake News と正しいニュースでは時間の経過と共に増加するその記事に賛成・反対する割合が異なる事例が確認された。そこで本研究では、ニュースに対するツイートの立場(賛成・反対など)の情報が Fake News 検出において重要であると仮定し、立場分類から Fake News 検出に適応して学習を行うことを提案する。本研究の新しい試みとしては、ユーザの反応を適切に捉えるには立場が重要であると考え、ユーザの反応を識別するモデルにおいて立場を使用した別のデータセットからの転移学習・マルチタスクラーニングを使用することである。

ある Fake News に対して、立場分類を用いて疑惑を抱いているユーザの反応を拾うことで、Fake News の自動検出の精度向上を目指す。データセットとして FakeNewsNet を使用し、加えて FakeNewsNet に手動で立場のラベリングを行ったもの、他タスクである RumorEval を使用した。転移学習においては、2 値の立場分類を行ったあとに、2 値の FakeNews の検出を行った。Multi-task learning では、loss 関数を改良することで同時に複数のデータセットを学習し、メインタスクは FakeNews の検出とした。実験のモデルとしては Transformer である BERT を使用し、データセットは FakeNewsNet とした。

提案手法の評価実験では Fake News かどうかの 2 値分類を行い、転移学習と Multi-task learning の有無の結果を比較する。実験の結果、Multi-task learning を使用したモデルでは有意差はみられなかったが、どのタスクが Fake News に影響するタスクなのかを判断することができた。また転移学習を行ったモデルでは転移学習なしの結果を上回る傾向がみられることを確認した。

(指導教員 関 洋平)