

2. 【現在までの研究状況】(図表を含めてもよいので、わかりやすく記述してください。様式の変更・追加は不可(以下同様))

- ① これまでの研究の背景、問題点、解決策、研究目的、研究方法、特色と独創的な点について当該分野の重要文献を挙げて記述してください。
- ② 申請者のこれまでの研究経過及び得られた結果について、問題点を含め①で記載したことと関連づけて説明してください。
- なお、これまでの研究結果を論文あるいは学会等で発表している場合には、申請者が担当した部分を明らかにして、それらの内容を記述してください。

●これまでの研究の背景

SNS の発展により、情報を迅速に大量取得し、拡散することで容易に共有できるようになった。その一方、悪意によって他人を騙すために作られた**フェイクニュース**が拡散されやすくなった。フェイクニュースが拡散されると、**誤った認識が広がって騙された人々が社会的損害を起こす**という問題がある。たとえば、2016 年米国大統領選挙前にフェイクニュースに騙された人々がピザ屋で銃撃事件を起こした [1]。また、今年は特に COVID-19 にまつわるフェイクニュースが広く拡散され、不安に陥った人々が買いだめを行うことが世界的に問題となった。WHO は情報の過剰な氾濫を“インフォデミック”と定義し、テドロス事務局長は**誤った情報はウイルス以上に拡散されやすい**と指摘した [2]。

●問題点

現在フェイクニュース対策として有識者が事実関係を確認する**ファクトチェック**が行われている。ただしこれは**属人的な作業**であり、拡散されてから調査されることが多く、結果を公表するまで時間がかかることからフェイクニュースと比べあまり**拡散されにくい**。また自動で検出する場合、フェイクニュースは巧妙に実際のニュースを模した形をとるため単純な**ルールベース手法では難しい**。近年ではニューステキストや添付メディア、ユーザの反応を入力にもつ**ディープニューラルネットワーク (DNN)** を利用した手法がみられる。その中で**ユーザの反応は拡散後でしか得られない**ため、早期発見を想定した場合評価対象にすることができない。

●解決策

そこで、当研究では**学習でのみユーザの反応を活用し**、テスト時は**ユーザの反応を生成し補完して、精度を落とさず早期発見を目指す**ことにした。

●研究目的・研究方法

フェイクニュース早期発見に向け、SNS 上で**ニュースに寄せられたコメントを生成することが、真偽を分類する精度の向上につながる**ことを示す。本研究はニュースと寄せられたコメントを、ニュース本文と実際に SNS 上で投稿されたコメント 3 件を 1 ユニットとして扱うことにした。

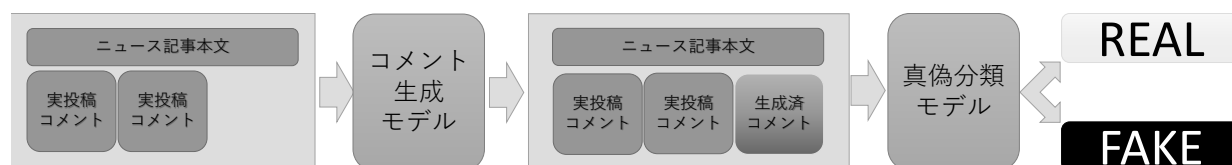


図 1: 分類タスクの流れ。コメント生成モデルで 1 件コメントを生成し真偽分類に活用する。

●特色と独創的な点

- 生成タスクを分類タスクに役立てる点
- 精度を失わずに速報性をもつことができる点
- 生成されたコメントは説明可能性にも役立てられる点

●これまでの研究経過及び得られた結果

申請者はデータセットとして FakeNewsNet[3] を使用した。これはファクトチェックによって**真偽が評価済である英文ニュース**と、それに **Twitter 上で言及された投稿 (ツイート)** 等をもつ。当研究では最低 3 件以上英文でコメントとしてツイートが寄せられたニュースを真偽で各 2000 件使用した。拡散の初期段階ではコメントの数は期待できないため、使用するコメントは各 3 件ずつ無作為に選出し、残りは対象から除外した。

生成・分類モデルはフェイクニュースを自動で作成する Grover モデル [4] を拡張する形で実装した。このモデルはフェイクニュースをドメイン・著者・投稿日・見出し・本文の 5 要素に分け、**ランダムで歯抜けにして予測させる形で生成学習を実現**したものである。今回はこれをユニットの 4 要素 (記事本文と 3 件のコメント) での実装を目指し調整を行った。訓練が完了したコメント生成モデルを使い、図 1 のように **コメントを 1 件歯抜けにさせたユニットに生成コメントを付加した上で Real か Fake か分類させた**。分類モデルは Grover が提供したものを流用した。

その結果、生成コメントを含めた場合の **Fake 記事を見抜いた割合を示す再現率 (Recall) が 0.79 と、歯抜けのまま分類させたときの 0.75 とコメントなしで分類したときの 0.62 を上回る結果を得た**。つまり、コメントを生成することで **ファクトチェックが必要な疑わしい記事をより多く検出**することができた [5]。同時に、生成されたコメントで頻出した単語の傾向において真偽で大きな違いはみられなかった。これは、投稿された **コメントのカテゴリによる出現単語傾向の差は軽微**であることを意味した。

参考文献

- [1] Guardian staff and agencies. Washington gunman motivated by fake news ‘pizzagate’ conspiracy, 12 2016.
- [2] John Zarocostas. How to fight an infodemic. *The Lancet*, Vol. 395, No. 10225, p. 676, 2020.
- [3] Kai Shu, *et al.* Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media. *ArXiv*, Vol. abs/1809.01286, 2018.
- [4] Rowan Zellers, *et al.* Defending against neural fake news. *Advances in Neural Information Processing Systems 32*, pp. 9054 – 9065. Curran Associates, Inc., 2019.
- [5] Yuta Yanagi, *et al.* Fake news detection with generated comments for news articles. *EasyChair Preprint no. 3190*, EasyChair, 2020.

3. 【これからの研究計画】

(1) 研究の背景

これからの研究計画の背景、問題点、解決すべき点、着想に至った経緯等について参考文献を挙げて記入してください。

● 2. で述べた研究状況を踏まえこれからの研究計画の背景

自然言語処理技術は、近年は前後の文脈を考慮できる BERT[1] を始め、より自然な文章が生成できるようになった。その 1 つとして、前節で述べた Grover[4] も含まれる。前節の研究では実際に Grover を拡張し **コメントを生成することでより多くのフェイクニュースを検出**することができた。これからの研究では、より拡散を抑制することに繋ぐことが可能なモデルの開発を行うことを目指す。

● 問題点・解決すべき点

生成コメントを付加して分類した場合、前節提案モデルが Fake と判断した中で実際に Fake だった割合である精度 (Precision) は 0.59 だった。これは生成コメントを付加せず分類したときの 0.68 と比べ 0.09 ポイント下回る [5]。つまり多くのフェイクニュースを検出することができても、事実に基づくニュースも同時にフェイクニュースと判定する狼少年のようなモデルとなる。同時に、生成されたコメントは文法面でのクオリティが決して良いものとは言えず、このままでは説明可能性に繋げることは難しい。これでは利用者の信用を得るのは難しく、拡散の抑制にはならない。

この **精度の向上と、不自然なコメントにより説明可能性を提供できない 2 点**を解決する必要がある。

● 着想に至った経緯

実際に早期自動検出モデルを SNS 上で運用する場合を想定した。このモデルの目的は拡散の抑制であるため、**フェイクニュース以上に利用者の信用を得なければ拡散を食い止められない**。そのためには、真偽分類の精度を上げることと、説得力向上のために説明可能性を同時に提供する必要があると考えた。

参考文献

- [1] Jacob Devlin, *et al.* BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proc. of the NAACL-HLT*, pp. 4171 – 4186, Minneapolis, Minnesota, June 2019.

(2) 研究目的・内容 (図表を含めてもよいので、わかりやすく記述してください。)

- ① 研究目的、研究方法、研究内容について記述してください。
- ② どのような計画で、何を、どこまで明らかにしようとするのか、具体的に記入してください。
- ③ 所属研究室の研究との関連において、申請者が担当する部分を明らかにしてください。
- ④ 研究計画の期間中に異なった研究機関（外国の研究機関等を含む。）において研究に従事することを予定している場合はその旨を記載してください。

象の卵の研究の目的は...

(3) 研究の特色・独創的な点

次の項目について記載してください。

- ① これまでの先行研究等があれば、それらと比較して、本研究の特色、着眼点、独創的な点
- ② 国内外の関連する研究の中での当該研究の位置づけ、意義
- ③ 本研究が完成したとき予想されるインパクト及び将来の見通し

今まで、研究者は皆「ほ乳類は卵を産まない」という生物学のいわゆる「常識」に捕われ、象が卵を産むなどということはありませんと考えていた。しかし、ほ乳類は文字通り、産まれた乳幼児に乳を与える動物の総称であり、産まれる過程を規定しているわけではない。象のように大きな動物にあっては、体内に大きな胎児をかかえて移動するよりは、卵を産んでそれを暖め、ふ化してから乳を与えて育てる方が効率的である。

したがって、今までの「常識」を打ち砕く新たな観点が、この研究の独創的な点である。

[illegible]

(研究の特色・独創的な点の続き)

むようにしよう。とにかく何か書いて文字数をかせいで、次のページまで流れ込むようにしよう。とにかく何か書いて文字数をかせいで、次のページまで流れ込むようにしよう。とにかく何か書いて文字数をかせいで、次のページまで流れ込むようにしよう。とにかく何か書いて文字数をかせいで、次のページまで流れ込むようにしよう。

(4) 研究計画

申請時点から採用までの準備状況を踏まえ、研究計画について記載してください。

この行の高さは、上の「研究の特色・独創的な点」との間隔を \vspace を用いて調整してください。



図 2: 右目用



図 3: 左目用

実はこの象の卵の像の中に本当の卵が隠されている可能性もあるため、最新の技術を用いて非破壊的に内部を調査する。

象の卵の研究の計画と方法は...

(研究計画の続き)

(5) 人権の保護及び法令等の遵守への対応

本欄には、研究計画を遂行するにあたって、相手方の同意・協力を必要とする研究、個人情報の取り扱いの配慮を必要とする研究、生命倫理・安全対策に対する取組を必要とする研究など法令等に基づく手続が必要な研究が含まれている場合に、どのような対策と措置を講じるのか記述してください。例えば、**個人情報**を伴うアンケート調査・インタビュー調査、**国内外の文化遺産**の調査等、**提供を受けた試料の使用**、**侵襲性を伴う研究**、**ヒト遺伝子解析研究**、**遺伝子組換え実験**、**動物実験**など、研究機関内外の情報委員会や倫理委員会等における承認手続が必要となる調査・研究・実験などが対象となりますので手続の状況も具体的に記述してください。

なお、該当しない場合には、その旨記述してください。

象の卵のES細胞の培養、象のクローンの生成などを行わない。象個体を現地から持ち出すことはない。ワシントン条約ならびに生物多様性条約に抵触しない。また、組換え実験は行なわないので、カルタヘナ議定書にも抵触しない。

申請者登録名 柳 裕太

4. 【研究遂行能力】 研究を遂行する能力について、これまでの研究活動をふまえて述べてください。これまでの研究活動については、網羅的に記載するのではなく、研究課題の実行可能性を説明する上で、その根拠となる文献等の主要なものを適宜引用して述べてください。本項目の作成に当たっては、当該文献等を同定するに十分な情報を記載してください。

具体的には、以下(1)～(6)に留意してください。

(1) 学術雑誌等(紀要・論文集等も含む)に発表した論文、著書(査読の有無を明らかにしてください。査読のある場合、採録決定済のものに限ります。)

著者、題名、掲載誌名、発行所、巻号、
pp 開始頁～最終頁、発行年を記載してください。

(2) 学術雑誌等又は商業誌における解説、総説

(3) 国際会議における発表(口頭・ポスターの別、査読の有無を明らかにしてください。)

著者、題名、発表した学会名、論文等の番号、場所、月・年を記載してください。(発表予定のものは除く。ただし、発表申し込みが受理されたものは記載してもよい。)

(4) 国内学会・シンポジウム等における発表

(3)と同様に記載してください。

(5) 特許等(申請中、公開中、取得を明らかにしてください。ただし、申請中のもので詳細を記述できない場合は概要のみの記載してください。)

(6) その他(受賞歴等)

応募者は過去 20 年間、7つの海を隅から隅まで航海し、浅瀬から深海まで潜り、文字通り東西南北上下の3次元でシロナガスクジラの卵の探索を行ってきた。シロナガスクジラに飲み込まれそうになったり、海賊に捕まるなどの危険な目にもあったが、それにもめげず、研究を遂行してきた強靱な能力を有する。

(1) 学術雑誌(紀要・論文集等も含む)に発表した論文、著書

(査読有り)

1. H. Yukawa¹, J. Kara², “Theory of Elephant Eggs”, Phys. Rev. Lett. **800**, 800-804 (2005).

2. F. Ehrlich, H. Yukawa¹, “You can’t Lay an Egg If You’re an Elephant”, JofUR
(www.universalrejection.org), N/A, N/A (2002).

(査読なし)

3. Kobo Abe³, H. Yukawa¹, “仔象は死んだ”, 安部公房全集, **26**, 100-200, (2004).

他 5 報

(2) 学術雑誌等又は商業誌における解説・総説

1. R. Kipling, H. Yukawa, “The Elephant’s Child (象の鼻はなぜ長い)”, Nature, **999**, 777-779, (2003).

他 2 件

(3) 国際会議における発表

1. ○ 湯川秀樹, “Theory of Elephant Eggs”, 原始殻物理国際会議、カラチ、2006 年 2 月

他 1 件

(4) 国内学会・シンポジウムにおける発表

1. ○ 湯川秀樹、朝永振一郎、「ほ乳類の真の意味」、ほ乳類学会、東京、2003 年 6 月

他 3 件

(5) 特許等

(公開中)

1. 800800 号、「クジラの卵を用いた深海潜水艇」湯川秀樹、2003 年 4 月

(6) その他

1. もうすぐもらえるで賞

申請者登録名 柳 裕太

5. 【研究者を志望する動機、目指す研究者像、アピールポイント等】

日本学術振興会特別研究員制度は、我が国の学術研究の将来を担う創造性に富んだ研究者の養成・確保に資することを目的としています。この目的に鑑み、研究者を志望する動機、目指す研究者像、アピールポイント等を記入してください。

1. 研究職を志望する動機、目指す研究者像、自己の長所等

私は、今までの世界観を全く変える発見をしたい。そのためには、あくせく金を稼ぐ普通の仕事ではなく、じっくりと研究と野球に取り組める研究職しかない。

私が理想とする研究者は、三四郎の友人でもある、野々宮宗八である。彼は俗世間の現実には煩わされることなく、地下の実験室で黙々と光の圧力の測定に取り組んでいる。彼こそが、真の研究者である。

自分の長所は、まじめで賢いことである。

2. 自己評価をする上で、特に重要と思われる事項

間もなく、ノーベル賞受賞予定。