

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、解決方策、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。

1. 研究の背景 — ブラックボックス化しやすい深層学習モデルの解釈

申請者は人工知能の核となる技術である深層学習の研究を行っており、特に人間の言語を計算機で扱う自然言語処理に対して、予測の解釈や根拠の提示について取り組んでいる。深層学習は従来の古典的な手法と比べて予測性能が非常に高い反面、具体的な予測の過程を解釈しづらく、ブラックボックス化しやすい問題がある。深層学習の予測を解釈する上で、主に「注意」の重み [1] や推論結果に対する「勾配」を元にした重要度の可視化 [2] といった予測の解釈手法が盛んに研究されている (図 1)。

申請者は、注意の重みの可視化による配信効果の高い広告文を評価する枠組みを提案しており [研究業績 (3)-2]、データマイニング分野の最難関国際会議 (採択率 約 20%) に採択された。深層学習の解釈性向上は様々な分野で重要な役割を担う可能性があり、特に申請者が取り組んできた自然言語処理と画像処理の融合 [研究実績 (3)-1, (3)-3] や医療分野における自動診断システム [研究実績 (4)-8, (4)-7] 等、結果に対する根拠の提示が重要視される課題において、我々人間が予測結果の解釈が可能になる。

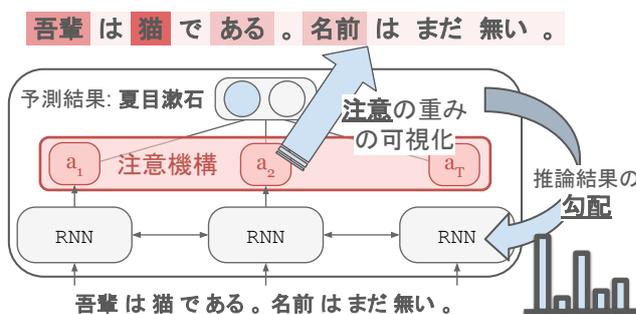


図 1: 「注意」の重みと推論結果に対する「勾配」を元にした重要度の可視化による、深層学習モデルの予測根拠の提示例。

2. 従来研究の問題点 — 深層学習モデルの解釈手法への疑問

「注意」機構を有したモデルは人間と同程度の予測性能を実現可能であり、近年の深層学習モデルには必須の技術である [3]。また「注意」の重みは推論結果の「勾配」を元にした重要度とともに深層学習モデルの予測の解釈に用いられてきたが、これらの関係性について次のような問題点が挙げられる。

1. 注意の重みはどの箇所を重点的に考慮するかを学習するが、タスクや問題設定によって勾配による重要度との関連性が強い場合と弱い場合があり、これらが必ずしも強く関連するとは言えない
2. 最先端の予測性能の実現には注意機構が必要不可欠であるが、勾配による解釈とは異なる解釈を提示する場合があります、両者の解釈を融合することについては不明な点が多い

3. 解決方策 — 新たな学習手法「関連損失」の提案と、それに基づく解釈性の高いモデルの構築

注意の重みと勾配による重要度の関連性が強くなるよう深層学習モデルを訓練すれば、複数の解釈手法が関連および融合して働き、1つの解釈によらない予測根拠の提示が期待できる。この考えに基づき、注意の重みと勾配による重要度が関連するように学習を行う新たな手法を提案し、この学習法に基づいた解釈性の高い深層学習モデルを構築した (図 2)。この手法は申請者がこれまで研究を進めてきた自然言語処理と画像処理の融合モデルや、医療分野における自動診断システムその他、深層学習モデルに対して広く適用できる汎用性の非常に高い学習手法である。

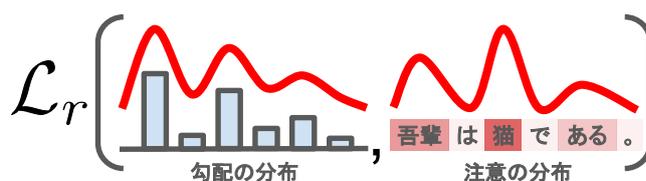


図 2: 提案手法である関連損失の模式図。単語の重要度と注意の重みが近い分布となるようにモデルを学習させることにより、解釈性の高い深層学習モデルを構築可能である。

4. 研究目的および研究方法

本研究では、自然言語処理のみならず、深層学習で重要となる本質的に解釈可能なモデルの構築が目的である。そのため、近年提案された新しい、なおかつ多様な解釈手法の取り入れを検討する。そこで、提案する新たな損失関数である関連損失の導入や、解釈手法同士の相関関係を元に解釈性を評価する。

申請者が提案する関連損失は、予測の解釈手法が相互に関連するような解釈の提示を目指している。特に先行研究で頻繁に用いられている注意による根拠と勾配による根拠を合わせることで、より解釈性の

(現在までの研究状況の続き)

高い予測根拠の提示を可能にする。深層学習の訓練時に、通常の損失に加えて提案する関連損失を導入することで、注意の分布と勾配の分布が関連するような学習を行う。

### 5. 特色と独創的な点

1. 予測の解釈に用いられる注意と勾配が関連するよう学習可能な新たな損失を設計した点
2. 予測根拠の統一的な解釈が可能となり、モデルの出力に対する解釈性が大幅に向上する点  
関連損失を中心とした成果は新規性や発展性が期待される奨励賞を受賞している [研究業績 (4)-6]。

### 6. これまでの成果

申請者のこれまでの研究成果は多岐に渡るが、最初の大きな成果は英語などの言語とは違い文字の区切りがない日本語や中国語などの言語を対象に、各文字を画像として扱う独創的かつ非常に高い性能を持つ文書解析手法の実現である [研究実績 (3)-3]。この手法を元にアラビア語の特徴的な文字形状に着目した文書分類手法 [研究実績 (3)-1] を提案し、文字形状に特徴がある言語で広く応用できることを確認した。

また、申請者はIT企業との連携により実世界のデータを利用することで、実際にどのweb広告文が購買に結びつくかの予測を高精度で行うことが出来る、極めて独自性が高く、解釈可能な深層学習モデルを実現した [研究実績 (3)-2]。上記の他にも深層学習に基づく皮膚がんの自動診断システムに関する論文 [研究実績 (4)-8] が学生奨励賞を受賞しており、幅広い分野で深層学習技術の応用に取り組んでいる。

参考文献

- [1] Bahdanau et al. “Neural machine translation by jointly learning to align and translate,” Proc. of ICLR, 2015.
- [2] Simonyan et al. “Deep inside convolutional networks: Visualising image classification model and saliency maps,” Proc. of ICLR workshop track, 2013.
- [3] Devlin et al. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” Proc. of NAACL-HLT, pp. 4171–4186, 2019.

## 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。

#### これからの研究計画の背景

予測性能の高い深層学習モデルは予測に寄与する重要な特徴を捉えていると考えられ、提示できる予測根拠にも高い解釈性が期待できる。また我々が気づかなかった新たな予測パターンの発見が可能である。高精度な予測を提供する解釈可能な深層学習モデルの構築を目指す上で、次のような問題点がある。

1. 深層学習モデルの解釈手法は一般的に摂動（ノイズ）に対して潜在的に脆弱であるため、摂動によって学習時とは異なる分布を持つ評価データに対して、その可視化結果の正当性が保証できない
2. モデルの出力に対する解釈や提示される根拠に対して、定量的な評価指標が明確ではない

#### 着想に至った経緯

深層学習は人間が知覚できない摂動に対しても反応してしまい、意図せぬ挙動や予測を示すことが知られている。これは予測の解釈として利用される注意機構も例外ではない [1]。この問題に対処すべく、学習時にあえてモデルを騙すような摂動を付加する「敵対的摂動」[2]による学習法に注目した。申請者はこの敵対的摂動が摂動に脆弱な注意機構を頑健にし、高い解釈性を有する予測根拠の提示が可能だと考えた。

深層学習における解釈性の適切な評価は、ブラックボックス化しやすいモデルを説明する上で非常に重要である。先行研究では注意の重みと推論結果の勾配による重要度に対して、順位を元にした相関が利用されている [1]。しかしながら、この指標は解釈にあまり寄与しない末端の順位付けも厳密に評価するため、しばしば誤解を招くような相関関係を示す。より本質的に解釈可能なモデルを評価するために、モデルの解釈手法が提示する根拠を適切に評価する指標の検討が必要である。

参考文献

- [1] Jain, Sarthak, and Byron C. Wallace. “Attention is not Explanation,” In Proc. of NAACL-HLT, pp. 3543–3556, 2019.
- [2] Miyato et al. “Adversarial training methods for semi-supervised text classification.” In Proc. of ICLR, 2016.

## (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。

### 研究目的

本研究では、解釈性の高い予測の根拠を提示可能な深層学習モデルの予測精度向上を目的とし、以下の2項目に焦点を当てて研究を行う。

#### 1. 摂動に頑健で解釈可能な深層学習モデルの確立

注意機構を有する深層学習モデルに対して、学習時に注意機構をあえて騙すような摂動を付加することで摂動に頑健となるよう訓練可能にする手法を確立する(図3)。提案手法の評価には、学習時に摂動を付加する部分(入力、注意)および、ノイズの種類を比較して、提案手法の有効性を議論する。提案手法は自然言語処理に関わらず、注意機構を有する深層学習モデル全般に適用可能な手法であるため、申請者がこれまで研究を進めてきた自然言語処理と画像処理の融合モデルや、医療分野における自動診断システム等の分野に対する有効性を確認する。

#### 2. 予測の解釈手法が提示する根拠に対する評価指標の確立

学習済みの注意の重みに対して、これまで解釈性の評価に使われてこなかった既存の指標や、人手のアノテーションを利用した新たな指標を元に、予測根拠に対する評価指標を確立する。人手のアノテーション解答根拠が人手で付与された既存のオープンデータセットの利用のほか、クラウドソーシングの利用を検討する。このとき、勾配とアノテーション間の一致率 [1] や認知心理学的アプローチを参考にし、既存手法や提案手法によって学習された注意の重みが人間の解釈と近いものであるかを検証する。

参考文献

- [1] Ghaeini et al. "Saliency Learning: Teaching the Model Where to Pay Attention." In Proc. of NAACL-HLT, pp. 4016-4025, 2019.

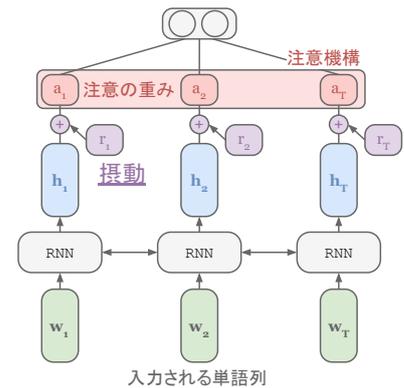


図3: 提案する摂動に頑健な解釈可能な深層学習手法の概要。注意の重みに対して摂動を付加することにより、注意機構が頑健となるように訓練する。

## (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

### 【特色・着眼点・独創点・位置づけ・意義】

#### 1. 摂動に頑健で解釈可能な深層学習モデルの確立

深層学習が摂動に脆弱な問題に対して、これまでは学習時の入力に対してランダムな摂動や敵対的摂動を付加することで、頑健性や予測精度の向上が確認されてきた。しかしながら学習時の注意の重みに対して摂動を付加したときの効果については議論されていない。本研究では、注意の重みに対して学習時に摂動を付加する新たな学習手法を提案する。摂動の種類としてランダムな摂動の他、あえてモデルを騙すような敵対的摂動を考慮することで、優れた頑健性の獲得と予測精度の向上が望める。更に、敵対的摂動を用いた学習がより明確な注意の学習を助け、注意に対して摂動が加わったとしても意味的および文法的な特徴を適切に捉えることを期待する。本提案手法は注意機構を有する深層学習モデル全般に適用可能な手法である。注意機構を有した、人間と同程度の予測性能を実現するモデルのさらなる精度向上や、より明確な注意の重みによる解釈性向上を実現できる。

#### 2. 予測の解釈手法が提示する根拠に対する評価指標の確立

深層学習モデルに対する解釈手法が提示する予測根拠を評価するために、適切な評価指標の確立は重要である。これまでの評価指標は、予測根拠として用いられてきた注意の重みと推論結果の勾配の重要度に対して、順位を元にした相関で評価している。しかしながらこの指標は厳密な順位を評価するために、解釈に寄与する箇所を適切に評価できているか疑問が残る。本研究では、既存の評価指標であるピアソ

(研究の特色・独創的な点の続き)

ンの相関による評価や、人手でアノテーションされたデータによる評価、および認知心理学的視点による評価を検討する。これらの評価指標を元に注意機構が人間の解釈と同じ特徴を捉えているか評価する。

【関連研究の中での当該研究の位置づけおよび意義】

既存研究ではモデルの予測性能と解釈性はトレードオフである。本研究では注意機構に対する敵対的摂動下での学習により、高い予測精度と解釈可能性の両立を実現する。また従来の解釈手法に対して人間と同様の解釈が学習されているかを直接的に分析するとともに新しい研究となるだろう。

【学術的および産業的の側面から期待できるインパクトや将来の見通し】

本研究の目的が達成された場合、(1) 予測性能が高くなおかつ解釈性が高い深層学習モデルの訓練および、(2) 深層学習モデルを通じて我々が気づけなかった新たな予測パターンの発見が可能である。深層学習の利用が一般的になりつつあるロボット分野や、失敗が許されない医療分野において、深層学習モデルのブラックボックス化は非常に大きな問題となる。提案する学習手法を通じて、摂動に頑健で予測精度が高く、なおかつ従来の古典的手法と同様の高い解釈性を兼ね備えた深層学習モデルによって解決可能であると考えられる。以下に学術的な側面と産業的な側面で予想されるインパクトを示す。

**学術面** 摂動に頑健で解釈可能な深層学習モデルの確立により、幅広く世界で用いられている解釈手法に対して、高い技術的貢献を実現できる。こうした貢献を通じて、注意機構を有した人間を超えるような予測性能を持つモデルを解釈することで、我々が気づけなかった新たな予測パターンの発見を期待できる。

**産業面** 本研究で確立する深層学習モデルは摂動に頑健で予測性能が高いため、実世界への迅速な応用展開が可能である。また、非常に高い解釈性を有しているため、ビジネスリスクになりうる深層学習モデルのブラックボックス化に対して、解釈性の高い予測根拠を提示することで予測の分析を容易にする。

(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。

図 4 に、本研究内容に対応する項目の実施期間を示す。また各項目の実施内容の詳細について、以下に記述する。

		1年目			2年目		
摂動に頑健で解釈可能な深層学習モデルの確立	付加する摂動の種類や位置の検討	■	■	■			
	解釈可能な既存手法の精度向上		■	■			
	異分野への応用の検討			■	■	■	■
解釈手法が提示する根拠に対する評価指標の確立	既存の解釈性評価指標による評価			■			
	新たな解釈性評価指標の考案				■	■	■
	新たな解釈性評価指標による評価				■	■	■

図 4: 本研究の年次計画概要。塗りつぶされた期間は該当項目の実施期間を表す。

1 年目

摂動に頑健で解釈可能な深層学習モデルの確立に着目して研究を行う。まず、モデルの解釈に利用される注意に対して頑健となるような摂動の付加方法について検討し、アルゴリズムの精度や自然言語処理の諸問題（特に予測の解釈が必要となるカテゴリ分類や感情分析、質問応答、自然言語推論等）への適応性の面から考察する。調査による知見から、実際に自然言語処理の諸問題に適した深層学習モデルの訓練を行うプログラムを構築し、ベースラインとの比較を通じて性能の評価を行う。この時、既存の解釈可能な深層学習を元にしたモデルの予測精度の向上に焦点を当てて研究を行う。ここから得られた知見を元に、深層学習モデルを利用可能な異分野への応用を検討する。特に申請者が以前取り組んでいた、

(研究計画の続き)

自然言語処理と画像処理の融合モデルや、失敗が許されない医療分野における深層学習を用いた自動診断システムの構築に対して、予測精度が高くなおかつ予測の解釈性が高いモデルの構築を目指す。

## 2年目

予測の解釈手法が提示する根拠に対する評価指標の確立に着目して研究を行う。まずは既存の指標を用いた解釈性の自動評価を行う。具体的には既存の評価指標であるピアソンの相関を用いて、注意の重みと推論結果に対する勾配を元にした重要度の関係性を評価する。この分析と並行して、深層学習モデルに対する新たな解釈性評価指標の考案を行う。具体的には人手でキーワードが付与されたニュース記事や、解答根拠が人手で付与されている質問応答データセットを用いて深層学習モデルを学習させ、注意の重みや勾配による重要度が人手のアノテーションとどの程度一致しているかを分析する。本研究ではモデルの解釈性を評価するために、これまで考慮されてこなかった認知心理学的視点による評価を検討する。これらの解釈性評価指標を元に、これまで検討してきた摂動に頑健な解釈可能な深層学習手法を評価する。更に、モデルが提示する予測根拠に対してクラウドソーシングを用いた人手による直接的な解釈性の評価を同時に進め、深層学習モデルに対して、我々が提案する予測の解釈手法の妥当性を評価する。

## (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。

本研究では、一般的に広く公開されているオープンデータセットを元にした性能の比較や結果の分析を行うため、個人情報の取り扱いなどの問題は発生しない。今後、企業などからデータを受け取る場合においても、個人情報を破棄した状態で受け取り、その分析結果の公開についても個人情報の保護や倫理に反することのないように十分配慮する。

申請者登録名 北田俊輔

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)

(1) 学術雑誌(紀要・論文集等も含む)に発表した論文、著書 なし

(2) 学術雑誌等又は商業誌における解説・総説 なし

(3) 国際会議における発表

1. ○ Mahmoud Daif, **Shunsuke Kitada**, Hitoshi Iyatomi, “AraDIC: Arabic Document Classification Using Image-Based Character Embeddings and Class-Balanced Loss,” In Proc. of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop (ACL 2020 SRW), Seattle, USA, July, 2020. [査読あり, ポスター発表]
2. ○ **Shunsuke Kitada**, Hitoshi Iyatomi, Seki Yoshifumi, “Conversion Prediction Using Multi-task Conditional Attention Networks to Support the Creation of Effective Ad Creative,” In Proc. of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2019), Alaska USA, August, 2019. [査読あり, ポスター発表, 採択率約 20%, データマイニング分野の最難関国際会議]
3. ○ **Shunsuke Kitada**, Ryunosuke Kotani, Hitoshi Iyatomi, “End-to-End Text Classification via Image-based Embedding using Character-level Networks,” In Proc. of IEEE Applied Imagery Pattern Recognition (AIPR 2018) workshop, Washington, D.C., USA, October, 2018. [査読あり, 口頭発表]

(4) 国内学会・シンポジウムにおける発表

1. ○ 北田 俊輔, “オーガナイズドセッション招待講演: 広告クリエイティブと機械学習技術における現状と展望”, 人工知能学会 第 34 回全国大会 (JSAI2020), 2020 年 6 月 9 日. [招待講演]
2. ○ 北田 俊輔, 彌富 仁, “解釈可能な敵対的摂動を用いた頑健な注意機構の学習,” 言語処理学会第 26 回年次大会 (NLP2020), 茨城, 3 月, 2020. [査読なし, ポスター発表]
3. ○ 青木 匠, 北田 俊輔, 彌富 仁, “文字単位の解釈可能な潜在表現の data augmentation,” 言語処理学会第 26 回年次大会 (NLP2020), 茨城, 3 月, 2020. [査読なし, ポスター発表]
4. ○ 長澤 駿太, 北田 俊輔, 彌富 仁, “Script-aware embedding を用いた文字表現の獲得,” 言語処理学会第 26 回年次大会 (NLP2020), 茨城, 3 月, 2020. [査読なし, ポスター発表]
5. ○ 北田 俊輔, 関喜史, “ウェブサービス事業者における研究開発インターン: 株式会社 Gunosy,” 第 15 回テキストアナリティクス・シンポジウム, 東京, 2019 年 9 月 27 日. [依頼講演]
6. ○ 北田 俊輔, 彌富 仁, “解釈性向上のための注意機構と損失勾配に対する関連損失の導入,” NLP 若手の会第 14 回シンポジウム (YANS2019), 北海道, 8 月, 2019. [査読なし, ポスター発表, 受賞 (下記)]
7. ○ 北田 俊輔, 彌富 仁, “頑健な皮膚腫瘍診断支援のための body hair augmentation,” 情報処理学会第 81 回全国大会 (IPJSJ2019), 福岡, 3 月, 2019. [査読なし, 口頭発表, 受賞 (下記)]
8. ○ **Shunsuke Kitada**, Hitoshi Iyatomi, “Skin lesion classification with ensemble of squeeze-and-excitation networks and semi-supervised learning,” CoRR preprint arXiv:1809.02568 (2018). [査読なし]  
(他 8 報)

(5) 特許等 なし

(6) その他

1. メディア取材, “データマイニング研究、グノシーがインターン院生主力に— 日刊工業新聞電子版”, 日刊工業新聞, 2020 年 1 月 30 日, <https://www.nikkan.co.jp/articles/view/00546182>.
2. 奨励賞 受賞, NLP 若手の会 第 14 回シンポジウム (YANS2019), 2019 年 8 月 28 日.
3. CEO 特別賞 受賞, 株式会社 Gunosy 第 7 期 下半期総会, 2019 年 7 月 26 日, <https://gunosiru.gunosy.co.jp/entry/party-7th-secondhalf>
4. 学生奨励賞 受賞, 情報処理学会 第 81 回全国大会 (IPJSJ2019), 2019 年 3 月 15 日.
5. 最優秀作品賞 受賞, Hack U 法政 2018, 2018 年 8 月 5 日, <https://hacku.yahoo.co.jp/hosei2018/>.

## 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

研究職を志望する動機は、世界の誰も知らない人工知能技術を発見および開発し、これらの技術や知見を元に日本や世界でより便利で豊かな生活への発展に役立ちたいからである。近年の人工知能技術の発展はめざましく、私たちの身の回りの生活を便利に変化させている。このような技術は最先端の研究から生まれるものであると考えているため、研究職として人工知能分野に携わっていききたい。

目指す研究者像は、最先端の人工知能技術の発見や開発をリードし、更に世界の人々が使いやすいように知見やツールを公開する実践的な研究者である。申請者は自身のサイトで最先端の自然言語処理や画像認識を中心に合計 100 編以上の詳細な論文解説を公開しており<sup>a</sup>、また深層学習に基づく皮膚がんの自動診断モデルと、そのモデルによる高速な診断システムの実装を公開している<sup>b</sup>。これらにより、最新の論文を読んで実装することで人工知能分野の動向を捉えていると自負している。また複数の招待講演や企業研究に特化したカンファレンス<sup>c</sup>の登壇により、実世界に対する知見共有を積極的に行っている。

自己の長所は、大きく次の3つである。

(1) **専門知識を元に幅広い問題を解決する力:** 申請者は深層学習を元にした自然言語処理を中心に、医用画像処理や計算機広告など多岐に渡る研究を行ってきた。これらは申請者が広く物事に興味をもって課題や問題を発見し、専門知識を元に解決できる力を有していると考えられる。

(2) **周りや密に連携するコミュニケーション力:** 研究や問題解決を行う上で、周りや密に連携することで自身が持つ力以上の実行力を発揮すると考えている。申請者は海外では一般的であるが国内ではまだ知名度が低いリサーチインターンとして株式会社 Gunosy で研究に従事している。その際に共同研究者と密に連携することで、結果的に最難関国際会議に論文が採択される研究を行うことができた。このリサーチインターンシップの成果や取り組みは日刊工業新聞に掲載された(図 5)。

(3) **アイデアの着想から実装するまでの実行力:** 申請者は日頃から最先端の深層学習手法を複数実装し公開している経験から、独創的なアイデアを着想するために日々努力している。またリサーチインターンシップという限られた時間の中での経験により、結果の考察をしながら手を動かすことで、短期間で大きな成果を上げることが出来た。このような実績から、申請者は高い生産性や実行力を有していると言える。



公開用では著作権に配慮し添付記事をぼかしました。

図 5: リサーチインターンの取り組みが掲載された記事(全文)。海外では一般的であるが国内では知名度が低く、先駆的な取り組みが注目された。

### 2. 自己評価をする上で、特に重要と思われる事項

これまでに多くの研究成果を残しており、15 件の国内会議、3 件の国際会議での発表が現時点で確定している。国際会議に関してはデータマイニング分野の最難関国際会議である KDD での論文採択があり、現在の所属研究室の学生として初めての成果である。また国内会議では、情報処理学会より奨励賞、NLP 若手の会より学生奨励賞をそれぞれ受賞している。申請者は自身の研究成果や発表資料、作成したソフトウェアなどを web サイトや外部サービスを通じて積極的に公開しており、これらより高い研究遂行能力を持つことが分かる。体外活動では外部でのイベント開催や参加を精力的に行っている。特に国内で初めてとなる競技機械学習大会である Cpaw AI Competition の作問や運営に携わった<sup>d</sup>。また申請者は複数の企業で R&D に特化したインターンシップについて積極的に取り組んでおり、実世界で応用される技術や産業界においても最先端の研究を遂行できる能力を持っていると言える。

<sup>a</sup> <https://shunk031.github.io/paper-survey/>    <sup>b</sup> <https://github.com/shunk031/libtorch-gin-api-server>

<sup>c</sup> <https://ccse.jp/2019/>    <sup>d</sup> <https://speakerdeck.com/shunk031/about-cpaw-ai-competition-in-kaggle-meetup-tokyo>