

令和 5 年 6 月 6 日現在

機関番号：12102

研究種目：若手研究

研究期間：2019～2022

課題番号：19K20333

研究課題名（和文）連語抽出手法と深層学習の融合による自然言語処理の精度向上に関する研究

研究課題名（英文）Study on Improving Performance of Natural Language Processing by Integrating Collocation Extraction and Deep Learning

研究代表者

若林 啓 (Wakabayashi, Kei)

筑波大学・図書館情報メディア系・准教授

研究者番号：40631908

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では、以下の3つの研究課題に取り組んだ。(課題A) テキストから意味のある連語をより高精度に抽出する技術として、連語の手がかりになる言語資源を活用した連語抽出モデル学習手法や、連語抽出モデルに対する効率的な教示手法を提案し、連語抽出に用いる統計モデルの基礎理論を発展させた。(課題B) 抽出された連語を活用して、文書要約、対話システムの言語理解、トピックモデリングの自然言語処理アプリケーションの精度を向上させる深層学習手法を提案した。(課題C) 後段の自然言語処理タスクの精度向上に寄与するような連語を、後段タスクの学習中に動的に抽出する手法を提案し、その効果を検証した。

研究成果の学術的意義や社会的意義

複数の単語で特定の意味を持つ連語を考慮することは、多くの自然言語処理のアプリケーションの精度を向上させるために重要な課題である。しかし、連語抽出手法の性質が、後段の自然言語処理タスクを学習する深層学習手法に与える影響については、これまで明らかにされてこなかった。本研究成果の意義は、連語の抽出と深層学習による自然言語処理タスクの精度向上を結びつける方法論を示し、その効果を明らかにしたことにある。とりわけ、連語を明示的に分析結果として提示するトピックモデリングや対話システムの言語理解タスクにおいて、直接的に応用可能な研究成果が得られたと考える。

研究成果の概要（英文）：In this study, we addressed the following three research questions. (A) To extract meaningful collocations from text with higher accuracy, we proposed a new method for training collocation extraction models by using linguistic resources and human annotator resources efficiently, and advanced a basic theory of statistical models used for collocation extraction. (B) We proposed a deep learning method that uses the extracted collocations to improve the accuracy of natural language processing applications, which are namely document summarization, language understanding in dialog systems, and topic modeling. (C) We proposed a method for dynamically extracting collocations that contribute to improving the accuracy of later-stage natural language processing tasks during the training of deep learning models for those tasks.

研究分野：機械学習，自然言語処理

キーワード：連語抽出 深層学習 能動学習 隠れマルコフモデル 文書要約 対話システム トピックモデル

1. 研究開始当初の背景

テキストデータの分析では単語を単位とした処理が行われることが多いが、単語は必ずしも現実の概念の単位と対応しておらず、このことが自然言語処理アプリケーションの精度を低下させる大きな要因の一つになっていた。このため、複数の単語で特定の意味を持つ連語を考慮することが、多くの自然言語処理のアプリケーションの精度を向上させるために重要な課題であると指摘されていた。

研究開始の当時までに、研究代表者を含む複数の研究者が、様々な自然言語処理アプリケーションのエラーを、連語の考慮によって修復できることを明らかにしていた。しかし、この結果の分析から、そもそも抽出に失敗しているために考慮されない連語が大量に残っており、精度向上の余地が未だに大きいことも示唆されていた。このため、高い網羅率で連語を発見できるような連語抽出手法の改良と、連語を考慮して後段の自然言語処理タスクを学習する深層学習手法の改良の2つの課題の解決が求められている状況であった。

2. 研究の目的

本研究の核心をなす問いは、連語抽出手法と、連語を考慮した深層学習手法のそれぞれを、どのように改良すればアプリケーションの精度を最適化できるのか、という点にある。本研究では、連語抽出手法の性質が、後段の自然言語処理タスクを学習する深層学習手法に与える影響を明らかにするとともに、それらを一つの基盤技術として融合する方法論を示すことを目的とする。この目的を達成するため、以下の3つの課題の解決に取り組んだ。

(課題 A) 自然言語処理アプリケーションの精度を高めるという観点から、連語抽出手法を改良する。

(課題 B) 抽出された連語を活用して、自然言語処理アプリケーションの精度を向上させる深層学習手法を開発する。

(課題 C) 課題 A, B で得られた知見を踏まえて、連語抽出手法と深層学習手法を一体化させた手法を構築する。

3. 研究の方法

課題 A では、以下の3つの方策に基づいて連語抽出手法を改良した手法の開発を行い、その性能や性質を検証した。

- ・ 連語の手がかりになる言語資源を活用した連語抽出モデル学習手法を開発する。
- ・ 連語抽出モデルに対する効率的な教示手法を開発する。
- ・ 連語抽出に用いる統計モデルの基礎理論を発展させる。

課題 B では、深層学習手法に連語の情報を入力する仕組みを改良することで、自然言語処理アプリケーションの精度向上を目指す。特に、文書要約タスク、対話システムの言語理解タスク、トピックモデリングの3つの自然言語処理アプリケーションにおいて、連語を考慮して精度を向上させる手法の研究を行った。

課題 C では、深層学習モデルによって学習される自然言語処理タスクを「後段タスク」と呼び、その前段として行われる連語抽出を、後段タスクの精度が向上するように最適化する手法の開発を行った。後段タスクの学習中に連語抽出を動的に行う手法を検討し、後段タスクの精度に対する影響を検証する。

4. 研究成果

(1) 連語抽出手法の改良に関する研究成果 (課題 A)

研究方法に示した3つの方策に基づいて研究を行い、それぞれ以下の成果が得られた。

連語の手がかりになる言語資源を活用した連語抽出モデル学習手法

対象とするテキストのドメインによっては、専門用語辞書や Wikipedia など、特定の意味をもつ連語のリストとして利用可能な言語資源が入手可能なことがある。このような言語資源(辞書)を利用できる連語抽出手法として最も基本的な手法は、辞書に含まれる連語と一致する文字列を抽出する文字列マッチングであるが、辞書に含まれない連語を抽出できないという問題や、誤った抽出を行う可能性がある。特に辞書が抽出した連語が誤っている場合のエラーについては、

それを修正する仕組みが既存の手法にないことが問題であった。本研究では、複数の連語抽出モデルを用いて学習と相互ラベリングを繰り返すことで、辞書マッチの誤りの可能性を考慮する手法を提案した。ベンチマーク用のデータセットを用いた実験により、提案手法が辞書マッチにおけるラベリングの誤りを減少させ、連語抽出の適合率を改善できることを示した (Kobayashi & Wakabayashi, 2020)。

連語抽出モデルに対する効率的な教示手法

連語抽出モデルの性能向上において、人からの教示情報は有用であるが、人の作業負担が大きいことが問題になる。本研究では、その負担を軽減するための研究に3つのアプローチで取り組んだ。

1つ目は、部分的なアノテーションでの教示が可能な能動学習手法である。従来は文全体の連語のアノテーションを人に依頼する必要があったが、提案手法は点推定モデルを用いることで、文の部分的なアノテーション情報での訓練を行う。どの部分のアノテーションを依頼するとモデルの訓練に効果的であるかを推定し、能動的に依頼を行う能動学習手法を提案し、その有効性を検証した (Kobayashi & Wakabayashi, 2019) (Kobayashi & Wakabayashi, 2020)。また、文のアノテーションを行う代わりに、連語の事例を言語資源として提供し、その事例を基に連語抽出モデルを訓練する能動学習手法も提案した。提案する能動学習手法は、システムが抽出した連語のうち、確信度の低いものをアノテータに提示し、その連語に対する正誤のフィードバックを得ることで、連語抽出モデルを改善する (Odakura et al., 2021)。

2つ目は、自然言語教示による連語抽出手法である。これまでに、テキスト分類のタスクにおいて、アノテータに分類の理由を併せて問い合わせることで、分類の手がかりとなる追加の情報を得る手法が知られていた。提案手法はこれを連語抽出タスクに拡張し、ある連語を連語として抽出した理由も併せてアノテータに問い合わせ、その理由を元に他の連語の抽出を試みる。実験により、自然言語教示を用いない手法に比べて、提案手法が効率的に連語抽出手法の精度を向上させることを示した (Saito et al., 2021)。

3つ目は、クラウドソーシングなどの不特定多数の作業者をを用いた連語抽出モデルの訓練である。クラウドソーシングでは、不正確な教示情報が混在した教示データが得られることが一般的である。このため、複数の作業者に同一の文のアノテーションを依頼し、その結果を統合する冗長化戦略が採用されることが多い。しかし、連語抽出モデルの教示において、この冗長化戦略が実際に有効であるかどうかは、これまで明らかにされていなかった。本研究では、連語抽出モデルの訓練において冗長化の効果を検証した。その結果、冗長化を行って高品質の教示情報を収集するよりも、不正確でも多数の文の教示データを収集する方が、最終的に得られる連語抽出モデルの性能が高まることが明らかになった (Shimizu & Wakabayashi, 2021), (Shimizu & Wakabayashi, 2022)。

連語抽出に用いる統計モデルの基礎理論

Web や書籍、新聞などからなる巨大な参照テキスト (コーパス) を、文の構文構造を考慮した統計モデルにより分析して、特別な意味がありそうな単語の並びを抽出する分割型連語抽出の統計モデルの研究を進展させた。分割型連語抽出のベースとなっている階層型隠れマルコフモデルは、系列分割を行うための階層構造が固定的であり、データに合わせて柔軟に構造を変えることができないため、外部知識や外部モデルとの連携が困難であった。本研究では、階層型隠れマルコフモデルよりも表現能力が高く、かつ柔軟な構造を持つサイレント隠れマルコフモデルを提案し、その理論と推論アルゴリズムを構築した。サイレント隠れマルコフモデルにおける系列分割は、サイレント状態 (観測値を出力しない内部状態) を潜在的に経由している箇所の推定に基づいて行われる。このモデルは数学的にもシンプルで拡張性が高いことから、連語の手がかりとなる外部知識を用いた半教師あり学習手法や、深層学習との同時学習手法など、これまで難しかった拡張を行う上で有効な基盤となることが期待される (Wakabayashi, 2019)。

(2) 連語を考慮した自然言語処理アプリケーションの精度向上に関する研究成果 (課題 B)

文書要約、対話システムの言語理解、トピックモデリングのアプリケーションを対象にして、主に以下の3つの研究成果を得た。

1つ目は、文書要約において連語を意味の単位として考慮する Semantic Content Generalization (SCG) 手法の検証である。SCG は、要約過程で変化させべきでない固有表現などに対応する連語を特殊トークンに置き換えることで、深層学習に基づく文書要約モデルが不正確な要約を出力することを防ぐ手法である。既存研究では、一般的な固有表現抽出モデルが抽出する人名、地名、組織名を対象として有効性を検証していたが、本研究ではより広い種類の連語を対象として SCG を適用して効果を検証した。また、既存研究では検証の対象となっていなかった、要約元文書の単語を直接コピーする機能を持つ Pointer Generator Network (PGN) においても SCG の効果があるかどうかを検証した。実験結果から、条件によっては、PGN においても SCG による精度向上の効果があることが明らかになったが、固有表現以外の連語を特殊トークンに置き換える手法では精度は向上しないことが示された (Wu & Wakabayashi, 2020)。

2つ目は、対話システムの言語理解において、連語の確率を明示的に考慮する手法の提案であ

る。対話システムの言語理解では、例えば「明日の午後の東京の天気を教えて」というユーザの発話文の中から、天気予報の問い合わせ条件として「明日の午後」というフレーズを抽出する必要がある。一般的な言語理解手法は、文に含まれるそれぞれの単語が抽出すべきフレーズの一部であるかどうかを教師あり学習によって分類するが、訓練データが少量であったり不正確であったりする場合に抽出精度が得られないという欠点がある。これに対して本研究では、連語の確率モデルを明示的に考慮することで、訓練データの不足に対して頑強な言語理解手法を提案した。提案手法は、ユーザの発話の中でシステムに直接与えたい情報を表す部分だけでなく、その周辺の特徴的な連語表現の確率を明示的にモデル化することで、言語理解タスクを発話全体の最も尤もらしい分割を決定する問題に帰着させる。この方式により、提案手法は特に訓練データが少量かつ不正確な時にも精度の低下を抑えることができ、頑健な性能を持つことを明らかにした (Wakabayashi et al., 2021), (Wakabayashi et al., 2022)。

3つ目は、連語を考慮したトピックモデリングの手法の提案である。連語は、トピックモデリングにおいて、ユーザーがトピックの意味を解釈する際に重要な役割を果たす。しかし、多くのトピックモデルは単語を分析の単位とすることを前提としたモデル化がされており、語彙サイズの制限とモデルの複雑さにより連語を組み込むことが困難である。本研究では、文書の埋め込みベクトルをクラスタリングすることでトピックを抽出する Top2Vec トピックモデルを基に、事前学習された文書、単語、連語の分散表現を活用したトピックモデリングの新たな手法を提案した。さらに、文書と連語の埋め込みモデルを共同で訓練する Joint Fine-tuning 法を提案し、提案手法がトピックモデリングの結果を改善することを明らかにした (Zhou & Wakabayashi, 2022)。

また、当初計画の課題遂行を通して得た知見を活かして、人からシステムへのフィードバックを連語に基づいて行うヒューマン・イン・ザ・ループ・トピックモデリング (Human-in-the-Loop Topic Modeling; HLTM) の研究を行なった。単語に比べて、連語は人にとっての解釈の単位として自然な言語表現である。既存の HLTM では、単語を単位としたフィードバックに基づいて、トピックモデルの修正を行っていた。提案手法は、文書のキーフレーズとなる連語を事前に抽出し、これを参照できるようにすることで、連語に基づいたフィードバックを可能にした。実験により、この方式が効果的に文書のトピックを目的の分類に近づけることができることを明らかにした (Khan & Wakabayashi, 2023)。

(3) 連語抽出手法と深層学習手法の融合に関する研究成果 (課題C)

後段の自然言語処理タスクの精度向上に寄与するような連語を、後段タスクの学習中に動的に抽出する手法を提案した。これまで、後段タスクのモデルの最適化と、文の単語分割を、単一のニューラルネットワークモデルの枠組みの中で同時に行うことで、後段タスクの精度を向上させるように分かち書きを最適化する手法が提案されていた。本研究ではこの枠組みを拡張し、単語分割と連語抽出のマルチレベルの分かち書きを、後段タスクの精度を向上させるように最適化する手法を提案した (図 1)。実験の結果から、後段タスクに依存した連語の抽出が可能になっていることが示されたが、後段タスクの精度自体は従来手法と大きく変化しないことが明

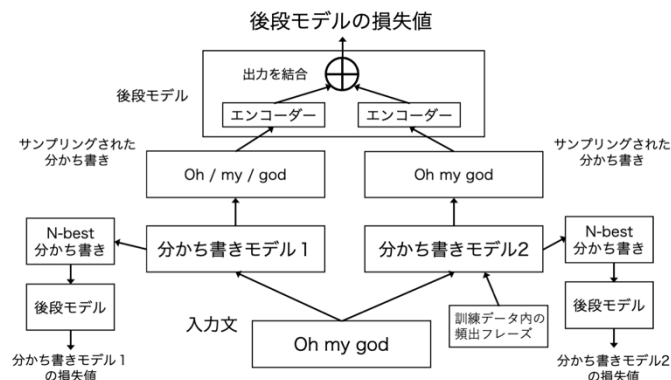


図 1 後段タスクの精度向上のためのマルチレベルな分かち書きの最適化

(小田倉 & 若林 2022)より引用)

らかになった (小田倉 & 若林, 2022)。この実験結果は、後段タスクを学習する深層学習モデルが、単語分割された入力文から連語の特徴を十分に抽出する能力があることを示唆している。

また、入力文の構文構造をグラフニューラルネットワークによってエンコードし、後段タスクの深層学習モデルの追加特徴量として用いる実験を行った。特に構文的な情報が有効に働くと期待されるタスクとして、対話システムの言語理解タスクにおける未知語の認識を対象にした実験を行ったが、後段モデル単体での精度を一貫して上回る結果は得られなかった (Yang et al., 2023)。これらの実験の結果から、近年の深層学習モデルである Transformer や再帰型ニューラルネットワークが、単語やサブワードなどのより細かいトークン列の情報から内部的に連語の特徴を十分に抽出できていることが示唆される。このことから、本研究の意義の一つは、研

究期間中に急速に進んだ深層学習モデルによって、連語の情報もある程度考慮できることを明らかにしたことにあると考えられる。

5. 主な発表論文等

〔雑誌論文〕 計18件（うち査読付論文 18件 / うち国際共著 0件 / うちオープンアクセス 16件）

1. 著者名 Rui Yang, Farnoosh Javar, Kei Wakabayashi, Johane Takeuchi	4. 巻 -
2. 論文標題 Exploring the Impact of Syntactic Structure Information on Unknown Entity Recognition in Transformer-based Natural Language Understanding	5. 発行年 2023年
3. 雑誌名 Proceedings of the 13th International Congress on Advanced Applied Informatics	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Muhammad Haseeb UR Rehman Khan, Kei Wakabayashi	4. 巻 16
2. 論文標題 Keyphrase-based Refinement Functions for Efficient Improvement on Document-Topic Association in Human-in-the-Loop Topic Models	5. 発行年 2023年
3. 雑誌名 情報処理学会論文誌データベース（TOD）	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Ryosuke Saito, Koga Kobayashi, Kei Wakabayashi	4. 巻 -
2. 論文標題 Efficient Training Method for Phrase Extraction Models using Natural Language Explanations	5. 発行年 2021年
3. 雑誌名 Proceedings of the 23rd International Conference on Information Integration and Web Intelligence	6. 最初と最後の頁 288-295
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3487664.3487703	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Fumimaro Odakura, Koga Kobayashi, Kei Wakabayashi	4. 巻 -
2. 論文標題 Active Learning for Extracting Technical Terms Covering Multiword Phrases	5. 発行年 2021年
3. 雑誌名 Proceedings of the 23rd International Conference on Information Integration and Web Intelligence	6. 最初と最後の頁 311-318
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3487664.3487706	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Ayame Shimizu, Kei Wakabayashi	4. 巻 -
2. 論文標題 Examining Effect of Label Redundancy for Machine Learning using Crowdsourcing	5. 発行年 2021年
3. 雑誌名 Proceedings of the 23rd International Conference on Information Integration and Web Intelligence	6. 最初と最後の頁 87-94
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3487664.3487677	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kei Wakabayashi, Johane Takeuchi, Mikio Nakano	4. 巻 -
2. 論文標題 Segmentation-Based Formulation of Slot Filling Task for Better Generative Modeling	5. 発行年 2021年
3. 雑誌名 Proceedings of the 12th International Workshop on Spoken Dialog System Technology	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Muhammad Haseeb UR Rehman Khan, Kei Wakabayashi	4. 巻 -
2. 論文標題 Drifting and Popularity: A Study of Time Series Analysis of Topics	5. 発行年 2021年
3. 雑誌名 Proceedings of the Seventh International Conference on Big Data, Small Data, Linked Data and Open Data	6. 最初と最後の頁 16-22
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Shimizu Ayame, Wakabayash Kei	4. 巻 3
2. 論文標題 Effect of Label Redundancy in Crowdsourcing for Training Machine Learning Models	5. 発行年 2022年
3. 雑誌名 Journal of Data Intelligence	6. 最初と最後の頁 301-315
掲載論文のDOI (デジタルオブジェクト識別子) 10.26421/JDI3.3-1	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kei Wakabayashi, Johane Takeuchi, Mikio Nakano	4. 巻 37
2. 論文標題 Robust Slot Filling Modeling for Incomplete Annotations using Segmentation-Based Formulation	5. 発行年 2022年
3. 雑誌名 人工知能学会論文誌	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Koga Kobayashi, Kei Wakabayashi	4. 巻 1
2. 論文標題 Partial Annotation Scheme for Active Learning on Named Entity Recognition Tasks	5. 発行年 2020年
3. 雑誌名 Journal of Data Intelligence	6. 最初と最後の頁 319-332
掲載論文のDOI (デジタルオブジェクト識別子) 10.26421/JDI1.3-2	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Masanori Kimura, Kei Wakabayashi, Atsuyuki Morishima	4. 巻 -
2. 論文標題 Batch Prioritization of Data Labeling Tasks for Training Classifiers	5. 発行年 2020年
3. 雑誌名 Proceedings of the 8th AAAI Conference on Human Computation and Crowdsourcing	6. 最初と最後の頁 163-167
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Yixuan Wu, Kei Wakabayashi	4. 巻 -
2. 論文標題 Effect of Semantic Content Generalization on Pointer Generator Network in Text Summarization	5. 発行年 2020年
3. 雑誌名 Proceedings of the 22nd International Conference on Information Integration and Web-based Applications & Services	6. 最初と最後の頁 72-76
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3428757.3429118	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Koga Kobayashi, Kei Wakabayashi	4. 巻 -
2. 論文標題 Mitigating Effect of Dictionary Matching Errors in Distantly Supervised Named Entity Recognition	5. 発行年 2020年
3. 雑誌名 Proceedings of the 22nd International Conference on Information Integration and Web-based Applications & Services	6. 最初と最後の頁 111-114
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3428757.3429142	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Muhammad Haseeb UR Rehman Khan, Kei Wakabayashi	4. 巻 -
2. 論文標題 Drifting and Popularity: A Study of Time Series Analysis of Topics	5. 発行年 2021年
3. 雑誌名 Proceedings of the Seventh International Conference on Big Data, Small Data, Linked Data and Open Data	6. 最初と最後の頁 16-22
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kei Wakabayashi	4. 巻 -
2. 論文標題 Silent HMMs: Generalized Representation of Hidden Semi-Markov Models and Hierarchical HMMs	5. 発行年 2019年
3. 雑誌名 Proceedings of the 14th International Conference on Finite State Methods and Natural Language Processing	6. 最初と最後の頁 98-107
掲載論文のDOI (デジタルオブジェクト識別子) 10.18653/v1/W19-3113	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Koga Kobayashi, Kei Wakabayashi	4. 巻 -
2. 論文標題 Named entity recognition using point prediction and active learning	5. 発行年 2019年
3. 雑誌名 Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services	6. 最初と最後の頁 287-295
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3366030.3366072	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Muhammad Haseeb Ur Rehman Khan, Kei Wakabayashi, Satoshi Fukuyama	4. 巻 -
2. 論文標題 Events Insights Extraction from Twitter Using LDA and Day-Hashtag Pooling	5. 発行年 2019年
3. 雑誌名 Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services	6. 最初と最後の頁 240-244
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3366030.3366090	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Xinnan Chen, Muhammad Haseeb Ur Rehman Khan, Kei Wakabayashi	4. 巻 -
2. 論文標題 Estimation Method of L2 Learners' Second Language Ability by using Features in Conversation	5. 発行年 2019年
3. 雑誌名 Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services	6. 最初と最後の頁 142-150
掲載論文のDOI (デジタルオブジェクト識別子) 10.1145/3366030.3366037	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計12件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 小田倉 史磨, 若林 啓
2. 発表標題 動作を表す言語を生成する深層学習における共同注意の有効性
3. 学会等名 第37回人工知能学会全国大会論文集
4. 発表年 2023年

1. 発表者名 Zikai Zhou and Kei Wakabayashi
2. 発表標題 Topic Modeling using Jointly Fine-tuned BERT for Phrases and Sentences
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2022年

1. 発表者名 小田倉 史磨, 若林 啓
2. 発表標題 後段タスクの精度向上のためのマルチレベルな分かち書きの最適化
3. 学会等名 第36回人工知能学会全国大会
4. 発表年 2022年

1. 発表者名 齊藤 亮将, 小林 滉河, 若林 啓
2. 発表標題 自然言語教示によるフレーズ抽出器の学習に関する研究
3. 学会等名 第13回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2021年

1. 発表者名 小田倉 史磨, 小林 滉河, 若林 啓
2. 発表標題 能動学習による複合語を考慮した専門用語抽出
3. 学会等名 第13回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2021年

1. 発表者名 清水 綾女, 若林 啓
2. 発表標題 クラウドソーシングによる訓練データセット構築における最適な冗長度の検証
3. 学会等名 第13回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2021年

1. 発表者名 大畑 直輝, 白井 匡人, 若林 啓, 劉 健全
2. 発表標題 ゼロショット文書分類向けの情報源領域から学習データの選択手法
3. 学会等名 第13回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2021年

1. 発表者名 小林 滉河, 若林 啓
2. 発表標題 遠距離教師あり固有表現抽出における辞書マッチの誤りの考慮
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 Wu Yixuan, 若林 啓
2. 発表標題 Effect of Semantic Content Generalization on Pointer Generator Network in Text Summarization
3. 学会等名 言語処理学会第26回年次大会
4. 発表年 2020年

1. 発表者名 岡 威久馬, 若林 啓
2. 発表標題 変分ベイズにおける最適解探索効率の検証
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2020年

1. 発表者名 白井 匡人, 若林 啓
2. 発表標題 文書要約における転移学習のための文書選択手法の提案
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2020年

1. 発表者名 清水 綾女, 若林 啓, 佐藤 哲司
2. 発表標題 Twitterユーザのリツイート情報を用いたトピックの可視化
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関