

令和 4 年 6 月 14 日現在

機関番号：12101

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K12093

研究課題名（和文）方向別 LSTM の共学習による半教師あり all-words WSD

研究課題名（英文）Semi-supervised all-words WSD by co-training of forward LSTM and backward LSTM

研究代表者

新納 浩幸（Shinnou, Hiroyuki）

茨城大学・理工学研究科（工学野）・教授

研究者番号：10250987

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：一般に単語は複数の語義（意味）を持つ。入力文内の各単語にその文における語義を与えるタスクが all-words WSD である。本研究ではこのタスクを解くモデルを双方向 LSTM に設定し、Co-training という半教師あり学習を用いてモデルの学習を行った。LSTM の後継モデルである BERT を利用することで本タスクが解決できるため、BERT の利用法を調査し、本タスクも含め様々なタスクに BERT を適用する方法を示した。

研究成果の学術的意義や社会的意義

自然言語処理の各種タスクは機械学習を利用することで解決できる。しかし機械学習で必要とされる訓練データの構築コストが大きいという問題がある。本研究のタスクの all-words WSD はその問題が特に顕著である。BERT は事前学習済みモデルであり、BERT を利用することで少量の訓練データから高精度のモデルを学習できる。研究課題の含め、各種タスクに BERT の利用する方法を示すことができた。

研究成果の概要（英文）：In general, a word has multiple senses (meanings). The all-words WSD is a task in which each word in an input sentence is assigned a sense in that sentence. Since this task can be solved by using BERT, the successor model of LSTM, we investigated the use of BERT and showed how to apply BERT to various tasks, including this task. We showed how to apply BERT to various tasks, including this task.

研究分野：自然言語処理

キーワード：all-words WSD 半教師あり学習 Co-training BERT

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開発当初の背景

WSD は文中の多義語の語義を識別する処理であり、意味解析を行う自然言語処理システムの根幹技術である。例えば単語「犬」には、少なくとも c_1 :動物の犬、 c_2 :スパイ、の2つの語義がある。そして入力文 $x = \text{``あいつは警察の犬だ"}$ が与えられたときに、この文中の「犬」が $C = \{c_1, c_2\}$ 内のいずれかを識別する処理が WSD である。WSD は $\text{argmax}_c p(c|x)$ を解く問題であり、分類問題そのものである。そのため WSD には、通常、教師あり学習のアプローチが取られ、高い精度で語義が識別できる。しかし教師あり学習はラベル付きデータを構築するコストが高く、現実のシステムでは対象単語を限定せざるをえない。対象単語を限定した WSD は、実際の文に適用されたとき、多くの語義の曖昧性を残してしまうため、実用的とは言えない。このため対象単語を限定せず、入力文内の全単語にその語義を付与する all-words WSD が期待されている。

2. 研究の目的

本研究の目的は高精度の all-words WSD システムを構築できる半教師あり学習手法の確立である。all-words WSD を実現するには、単純には文中の各多義語に対して WSD を行えばよいが、全多義語に対して大量のラベル付きデータを用意することは現実的には不可能である。そこで本研究では all-words WSD を系列ラベリング問題として定式化し、その問題に対する半教師あり学習を試みる。半教師あり学習では少量のラベル付きデータに加えて大量のラベルなしデータを利用して分類器を学習する学習手法である。一般に、半教師あり学習により得られる分類器の精度は、教師あり学習により得られるものと同等以上である。また all-words WSD の場合、ラベルなしデータはプレーンなテキストを意味する。このため all-words WSD に対してラベルなしデータを大量に収集することは容易であり、半教師あり学習により all-words WSD を実現する。

3. 研究の方法

all-words WSD に対応する系列ラベリング問題に対しては本研究では、双方向 LSTM のモデルを利用する。また半教師あり学習手法の一つである Co-training を all-words WSD に対して適用する。

Co-training は2つの独立した分類器 A と B を交互に学習し、互いの分類器を改善していく半教師あり学習手法である。まず少量のラベル付きデータ L_0 を用いて、分類器 A を学習し、学習できた分類器 A を用いてラベル無しデータを識別する。その識別結果のうち信頼度の高いデータだけに識別結果のラベルを付与して L_0 に追加し L_1 を作成する。次に L_1 を用いて、分類器 B を学習し、学習できた分類器 B を用いてラベル無しデータを識別する。その結果、識別結果の信頼度の高いデータだけに識別結果のラベルを付与して L_1 に追加し L_2 を作成する。このように分類器 A と B を交互に学習・適用し、ラベル付きデータを増加させていくことで分類器を改善していく。

本研究では順方向 LSTM と逆方向 LSTM を 2 つの独立した分類器として扱う。Co-training では信頼度の高いデータをラベル付きデータに追加するが、信頼度は各ラベル y_i に対して与えられるために、 x を単純にラベル付きデータに追加することはできない。本研究ではこの問題を重み付き学習により解決する。具体的には分類器 A (順方向 LSTM) において入力 $x = x_1 x_2 \dots x_n$ に対して $y = y_1 y_2 \dots y_n$ が出力されるが、各 y_i にその信頼度を意味する重み w_i を付与する。 w_i は x_i が持つ各語義に対する LSTM の出力値から求める。 w_i がある閾値以下であった場合 $w_i = 0$ とする。そして多義語に対する w_i がある閾値以上のとき、分類器 B (逆方向 LSTM) の学習に x を利用する。そのときの教師ラベルは y である。ただし学習には重みをつける。つまり x の分類器 B での出力が $z = z_1 z_2 \dots z_n$ であった場合、各 z_i とその教師ラベル y_i との誤差 $loss_i$ に重み w_i を乗じたものを誤差とする。結局 x に対して w_i を誤差として分類器 B のパラメータを更新する。この分類器 A と B を交換して、上記処理を繰り返すことで Co-training が行われる。このとき増加してゆくラベル付きデータは分類器毎に保持することに注意する。最終的に得られる分類器 A (順方向 LSTM) と B (逆方向 LSTM) を統合することで高精度の all-words WSD システムが完成する。

4. 研究の成果

本研究課題は 2018 年 4 月に採択されたが、その年の末に Google が発表した BERT により、本研究を含む自然言語処理の多くの研究が方向転換を余儀なくされた。従来の LSTM が BERT によって完全に凌駕されたからである。本研究についても、本質的には BERT により実現されている。そのため本研究課題に対しては BERT を実際に all-words WSD として利用するための手法、及びその手法を用いた他のタスクへの応用に関して研究を進めた。特に BERT が内在する MLM (Masked Language Model) は WSD を行うための核となる仕組みであり、MLM を利用したシステムを構築することで、MLM の能力やその問題点などを調査した。またタスクの領域に特化させることで BERT の小型化に関する研究も行った。以下に主な研究成果となる発表論文の要旨を述べる。

新納浩幸, 馬ブン, "BERT の Masked Language Model を用いた教師なし語義曖昧性解消", 第 27 回言語処理学会年次大会, A6-2 (2021).

本論文では BERT の MLM を利用することで、教師なしの語義曖昧性解消を試みた。MLM により推定される単語は当然、対象単語である筈だが、MLM では全ての単語に対して、その単語がマスク位置に現れる確率を付与できるため、対象単語の語義を持つ類義語が現れる確率が得られる。これによって対象単語の語義を推定できる。なお本研究では対象単語の語義を持つ類義語は分類語彙表から予め得ておく。実験では「言葉」「声」「電話」「国際」「市場」の 5 単語に対して、本手法を試み、教師なしの語義曖昧性解消手法として利用できることを確認した。

Yi Zhao, Rui Cao, Jing Bai, Wen Ma and Hiroyuki Shinnou, "Determining the Logical Relation between Two Sentences by Using the Masked Language Model of BERT", TAAI-2020 (2020).

本論文では BERT の MLM を利用することで、2 文間の関係を推定した。BERT の MLM を利用すると、文中のマスクされた単語を推定することができる。また BERT は 2 つの文 s_1 と s_2 を入力できる。そこで s_2 の直前にマスク単語を置き、"[CLS] s_1 [SEP] [MASK] s_2 [SEP]"の形の文を BERT に入力し、BERT から推定される [MASK] の単語から文 s_1 と文 s_2 の関係を推定することを試みた。具体的には関係を表す接続詞を準備し、[MASK] がそれら接続詞となる確率を求めることで 2 文間の関係を推定する。実験では [MASK] の部分が実際に接続詞であったものを対象にして、正しくマスクされた接続詞を推定できるかどうかの実験を行った。また教師あり学習で本タスクを解いた場合との比較も行った。

Youki Itoh, Hiroyuki Shinnou, "Domain-Specific Japanese ELECTRA Model Using a Small Corpus", RANLP-2021 (2021).

自然言語処理のタスクを機械学習のアプローチで解決する場合、訓練データの領域(ソース領域)とテストデータの領域(ターゲット領域)が異なるという domain shift の問題が深刻である。BERT のような事前学習済みモデルは下流のタスクによりモデルを fine-tuning するため、domain shift の問題に対処できる。さらに近年はターゲット領域のコーパスを用いて BERT の追加学習を行い、追加学習できたモデルを fine-tuning することで更なる精度向上がなされている。ただし BERT の追加学習には多大な計算機資源が必要であり、簡単に行うことはできない。また利用するターゲット領域のコーパスも大規模なものが想定されるが、そのようなコーパスが現実には入手できないことも多い。本研究では上記の問題への対処として ELECTRA の利用を試みた実験ではまず一般的な小規模の ELECTRA モデルを構築した。次にターゲット領域の小規模コーパスを利用し、その ELECTRA モデルの追加学習を行うことで領域特化型の ELECTRA モデルを構築した。構築した領域特化型の ELECTRA モデルは、BERT-base よりも小さなモデルではあるが、ターゲット領域の文書分類タスクに対して、BERT-base よりも性能がよいことを確認した。

Jing Bai, Rui Cao, Wen Ma and Hiroyuki Shinnou, "Construction of Domain-Specific DistilBERT Model by Using Fine-Tuning", TAAI-2020 (2020).

本論文も前述した論文と背景は同じである。この論文では ELECTRA ではなく蒸留のテクニックを利用した軽量版 BERT である DistilBERT を利用した。実験では Amazon レビューのデータセットを利用する。このデータセットは books, dvd, music の 3 つの領域をもつ。領域毎に既存の BERT モデルから領域に特化した DistilBERT モデルを構築した。領域毎に MASK 単語の推定問題を作成し、構築したモデルと既存 BERT モデルを比較す

ることで、領域に特化したモデルの有益性を示した。

Naoki Shibayama, Hiroyuki Shinnou, "Construction and Evaluation of Japanese Sentence-BERT Models", PACLIC-2021 (2021).

本論文では文の埋め込み表現を求めるのに適した日本語版の SentenceBERT を構築した。具体的には日本語版 SNLI (JSNLI) と公開されている 6 つの日本語 BERT をもとに 6 つの SentenceBERT を構築し、それらを比較した。その結果、東北大版 BERT と NICT 版 BERT から構築した SentenceBERT が同程度の高い性能を示した。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計53件（うち招待講演 1件 / うち国際学会 20件）

1. 発表者名 菊田尚樹, 新納浩幸
2. 発表標題 BERTを用いた文書分類タスクへのMix-Up手法の適用
3. 学会等名 第27回言語処理学会年次大会
4. 発表年 2021年

1. 発表者名 築地俊平, 新納浩幸
2. 発表標題 Tokenizerの違いによる日本語BERTモデルの性能評価
3. 学会等名 第27回言語処理学会年次大会
4. 発表年 2021年

1. 発表者名 新納浩幸, 馬ブン
2. 発表標題 BERT の Masked Language Model を用いた教師なし語義曖昧性解消
3. 学会等名 第27回言語処理学会年次大会
4. 発表年 2021年

1. 発表者名 平林照雄, 河野慎司, 古宮嘉那子, 新納浩幸
2. 発表標題 日本語の論文コーパスにおける「問題」の語義アノテーション
3. 学会等名 第27回言語処理学会年次大会
4. 発表年 2021年

1. 発表者名 伊藤陽樹, 新納浩幸
2. 発表標題 小規模コーパスを利用した領域特化型ELECTRAモデルの構築
3. 学会等名 第27回言語処理学会年次大会
4. 発表年 2021年

1. 発表者名 藤井真, 新納浩幸
2. 発表標題 BERTを利用したZero-shot学習による同音異義語の誤り検出
3. 学会等名 第27回言語処理学会年次大会
4. 発表年 2021年

1. 発表者名 平野友基, 新納浩幸
2. 発表標題 Faster-RCNNを用いたone-click supervision
3. 学会等名 情報処理学会NL研・CVIM研・PRMU研の合同研究会
4. 発表年 2021年

1. 発表者名 芝山直希, 新納浩幸
2. 発表標題 日本語 SentenceBERT の構築とその評価
3. 学会等名 情報処理学会自然言語処理研究会, NL-249-7
4. 発表年 2021年

1. 発表者名 高萩恭介, 新納浩幸
2. 発表標題 複数のBERTモデルを利用した Data Augmentation
3. 学会等名 情報処理学会自然言語処理研究会, NL-250-4
4. 発表年 2021年

1. 発表者名 河野慎司, 新納浩幸
2. 発表標題 簡易小型化BERTを利用した日本語構文解析
3. 学会等名 情報処理学会自然言語処理研究会, NL-251-20
4. 発表年 2021年

1. 発表者名 Youki Itoh, Hiroyuki Shinnou
2. 発表標題 Domain-Specific Japanese ELECTRA Model Using a Small Corpus
3. 学会等名 RANLP-2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Naoki Kikuta, Hiroyuki Shinnou
2. 発表標題 Application of Mix-Up Method in Document Classification Task using BERT
3. 学会等名 RANLP-2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Naoki Shibayama, Hiroyuki Shinnou
2. 発表標題 Construction and Evaluation of Japanese Sentence-BERT Models
3. 学会等名 PACLIC-2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Naoki Shibayama, Hiroyuki Shinnou
2. 発表標題 Construction and Evaluation of Japanese Sentence-BERT Models
3. 学会等名 PACLIC-2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Naoki Kikuta, Hiroyuki Shinnou
2. 発表標題 Application of Mix-Up Method in Document Classification Task using BERT
3. 学会等名 RANLP-2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Youki Itoh, Hiroyuki Shinnou
2. 発表標題 Domain-Specific Japanese ELECTRA Model Using a Small Corpus
3. 学会等名 RANLP-2021 (国際学会)
4. 発表年 2021年

1. 発表者名 岩本昇太, 新納浩幸
2. 発表標題 BERT の転移学習とMis-leading データの削除による識別精度の改善
3. 学会等名 第28回言語処理学会年次大会, PT4-13
4. 発表年 2022年

1. 発表者名 木村文飛, 新納浩幸
2. 発表標題 キーワード付与による画像キャプション生成
3. 学会等名 第28回言語処理学会年次大会, PT3-10
4. 発表年 2022年

1. 発表者名 田中裕隆, 新納浩幸
2. 発表標題 BERT の領域適応における複合語の語彙拡張
3. 学会等名 第28回言語処理学会年次大会, PT2-8
4. 発表年 2022年

1. 発表者名 河野慎司, 新納浩幸
2. 発表標題 簡易小型化BERTを利用した日本語構文解析
3. 学会等名 情報処理学会自然言語処理研究会, NL-251-20
4. 発表年 2021年

1. 発表者名 高萩恭介, 新納浩幸
2. 発表標題 複数のBERTモデルを利用した Data Augmentation
3. 学会等名 情報処理学会自然言語処理研究会, NL-250-4
4. 発表年 2021年

1. 発表者名 芝山直希, 新納浩幸
2. 発表標題 日本語 SentenceBERT の構築とその評価
3. 学会等名 情報処理学会自然言語処理研究会, NL-249-7
4. 発表年 2021年

1. 発表者名 平野友基, 新納浩幸
2. 発表標題 Faster-RCNNを用いた one-click supervision
3. 学会等名 情報処理学会NL研・CVIM研・PRMU研の合同研究会
4. 発表年 2021年

1. 発表者名 Jing Bai, Rui Cao, Wen Ma and Hiroyuki Shinnou
2. 発表標題 Construction of Domain-Specific DistilBERT Model by Using Fine-Tuning
3. 学会等名 TAAI-2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Hirotaka Tanaka, Rui Cao, Jing Bai, Wen Ma and Hiroyuki Shinnou
2. 発表標題 Construction of document feature vectors using BERT
3. 学会等名 TAAL-2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Yanghuizi Ou, Rui Cao, Jing Bai, Wen Ma and Hiroyuki Shinnou
2. 発表標題 Analysis of Polysemy using Variance Values for Word Embeddings by BERT
3. 学会等名 TAAL-2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Yi Zhao, Rui Cao, Jing Bai, Wen Ma and Hiroyuki Shinnou
2. 発表標題 Determining the Logical Relation between Two Sentences by Using the Masked Language Model of BERT
3. 学会等名 TAAL-2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Teruo Hirabayashi, Kanako Komiya, Masayuki Asahara and Hiroyuki Shinnou
2. 発表標題 Composing Word Vectors for Japanese Compound Words Using Bilingual Word Embedding
3. 学会等名 PAACLIC-2020 (国際学会)
4. 発表年 2020年

1 . 発表者名 Kanao Komiya, Daiki Yaginuma, Masayuki Asahara and Hiroyuki Shinnou
2 . 発表標題 Generation and Evaluation of Concept Embeddings Via Fine-Tuning Using Automatically Tagged Corpus
3 . 学会等名 PAFLIC-2020 (国際学会)
4 . 発表年 2020年

1 . 発表者名 Naoki Shibayama, Rui Cao, Jing Bai, Wen Ma and Hiroyuki Shinnou
2 . 発表標題 Evaluation of BERT Models by Using Sentence Clustering
3 . 学会等名 PAFLIC-2020 (国際学会)
4 . 発表年 2020年

1 . 発表者名 Teruo Hirabayashi, Kanao Komiya, Masayuki Asahara and Hiroyuki Shinnou
2 . 発表標題 Automatic Creation of Correspondence Table of Meaning Tags from Two Dictionaries in One Language Using Bilingual Word Embedding
3 . 学会等名 BUCC-2020 (国際学会)
4 . 発表年 2020年

1 . 発表者名 Hiroyuki Shinnou
2 . 発表標題 Use of BERT for NLP tasks by HuggingFace's transformers
3 . 学会等名 ROCLING-2020 (招待講演)
4 . 発表年 2020年

1. 発表者名 荘司響之介, 曹鋭, 白静, 馬ブン, 新納浩幸
2. 発表標題 二言語 BERT を利用したターゲット言語の教師データを必要としない感情分析
3. 学会等名 言語資源活用ワークショップ 2020
4. 発表年 2020年

1. 発表者名 趙一, 曹鋭, 白静, 馬ブン, 新納浩幸
2. 発表標題 BERT の Masked Language Model を用いた二文間の接続関係の推定
3. 学会等名 言語資源活用ワークショップ 2020
4. 発表年 2020年

1. 発表者名 欧陽恵子, 曹鋭, 白静, 馬ブン, 新納浩幸
2. 発表標題 BERT による単語埋め込み表現の分散値を用いた語義の広がり分析
3. 学会等名 言語資源活用ワークショップ 2020
4. 発表年 2020年

1. 発表者名 新納浩幸, 白静, 曹鋭, 馬ブン
2. 発表標題 Fine-Tuning による領域に特化した DistilBERT モデルの構築
3. 学会等名 第34回人工知能学会全国大会
4. 発表年 2020年

1. 発表者名 平林照雄, 古宮嘉那子, 新納浩幸
2. 発表標題 Bilingual Word Embeddingsによる短単位と長単位のアラインメント
3. 学会等名 第26回言語処理学会年次大会
4. 発表年 2020年

1. 発表者名 芝山直希, 曹銳, 白静, 馬ブン, 新納浩幸
2. 発表標題 文のクラスタリングを用いた BERT 事前学習モデルの評価
3. 学会等名 第26回言語処理学会年次大会
4. 発表年 2020年

1. 発表者名 Shin Fujii, Rui Cao, Jing Bai, Wen Ma, Hiroyuki Shinnou
2. 発表標題 Detecting Missing Translations in Neural Machine Translation Using Information Quantity in Sentences
3. 学会等名 TAAI-2019 (国際学会)
4. 発表年 2019年

1. 発表者名 Jing Bai, Rui Cao, Wen Ma, Hiroyuki Shinnou
2. 発表標題 Combination of Feature-based and Instance-based methods for Domain Adaptation in Sentiment Classification
3. 学会等名 TAAI-2019 (国際学会)
4. 発表年 2019年

1 . 発表者名 Jing Bai, Hirotaka Tanaka, Rui Cao, Wen Ma, Hiroyuki Shinnou
2 . 発表標題 Unsupervised Domain Adaptation for Sentimental Classification by Word Embeddings on the Lower Layer of BERT
3 . 学会等名 TAAL-2019 (国際学会)
4 . 発表年 2019年

1 . 発表者名 Hirotaka Tanaka, Hiroyuki Shinnou, Rui Cao, Jing Bai, Wen Ma
2 . 発表標題 Document Classification by Word Embeddings of BERT
3 . 学会等名 PAACLING-2019 (国際学会)
4 . 発表年 2019年

1 . 発表者名 Rui Cao, Jing Bai, Wen Ma, Hiroyuki Shinnou
2 . 発表標題 Semi-supervised learning for all-words WSD using self-learning and fine-tuning
3 . 学会等名 PAACLIC-2019 (国際学会)
4 . 発表年 2019年

1 . 発表者名 Kanao Komiya, Takumi Seitou, Minoru Sasaki, Hiroyuki Shinnou
2 . 発表標題 Composing Word Vectors for Japanese Compound Words Using Dependency Relations
3 . 学会等名 CICLING 2019 (国際学会)
4 . 発表年 2019年

1. 発表者名 田中裕隆, 曹鋭, 白静, 馬ブン, 新納浩幸
2. 発表標題 BERTを利用した文書の特徴ベクトルの作成
3. 学会等名 情報処理学会自然言語処理研究会 NL-243-8
4. 発表年 2019年

1. 発表者名 楊金成, 曹鋭, 白静, 馬ブン, 新納浩幸
2. 発表標題 Triple-GANによる感情分析に対する半教師あり学習
3. 学会等名 第15回テキストアナリティクス・シンポジウム
4. 発表年 2019年

1. 発表者名 芝山直希, 曹鋭, 白静, 馬ブン, 新納浩幸
2. 発表標題 日本語 Pretrained BERTモデルの比較
3. 学会等名 第15回テキストアナリティクス・シンポジウム
4. 発表年 2019年

1. 発表者名 馬ブン, 田中裕隆, 曹鋭, 白静, 新納浩幸
2. 発表標題 BERTを利用した単語用例のクラスタリング
3. 学会等名 言語資源活用ワークショップ 2019
4. 発表年 2019年

1. 発表者名 曹鋭, 田中裕隆, 白静, 馬ブン, 新納浩幸
2. 発表標題 BERTを利用した教師あり学習による語義曖昧性解消
3. 学会等名 言語資源活用ワークショップ 2019
4. 発表年 2019年

1. 発表者名 柳沼大輝, 古宮嘉那子, 新納浩幸
2. 発表標題 All-words WSDとfine-tuningを利用した分類語彙表の語義の分散表現の構築
3. 学会等名 言語資源活用ワークショップ 2019
4. 発表年 2019年

1. 発表者名 欧陽恵子, 田中裕隆, 曹鋭, 白静, 馬ブン, 新納浩幸
2. 発表標題 文書領域情報を有するBERTの階層位置に関する考察
3. 学会等名 言語資源活用ワークショップ 2019
4. 発表年 2019年

1. 発表者名 白静, 田中裕隆, 曹鋭, 馬ブン, 新納浩幸
2. 発表標題 BERT の下位階層の単語埋め込み表現列を用いた感情分析の教師なし領域適応
3. 学会等名 情報処理学会自然言語処理研究会 NL-240-17
4. 発表年 2019年

1. 発表者名 田中裕隆, 曹銳, 白静, 馬ブン, 新納浩幸
2. 発表標題 BERT による単語埋め込み表現列を用いた文書分類
3. 学会等名 情報処理学会自然言語処理研究会 NL-240-16
4. 発表年 2019年

〔図書〕 計2件

1. 著者名 新納 浩幸	4. 発行年 2021年
2. 出版社 インプレス	5. 総ページ数 240
3. 書名 PyTorch自然言語処理プログラミング word2vec/LSTM/seq2seq/BERTで日本語テキスト解析!	

1. 著者名 新納浩幸	4. 発行年 2020年
2. 出版社 オーム社	5. 総ページ数 208
3. 書名 PyTorchによる物体検出	

〔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
---------	---------