

# 深層学習による学術論文からの Structured Abstract 自動生成

橋本 快生<sup>†</sup> 井上 潮<sup>‡</sup>

<sup>†</sup> 東京電機大学 工学研究科 情報通信工学専攻 〒120-8551 東京都足立区千住旭町 5 番

<sup>‡</sup> 東京電機大学 工学部 情報通信工学科 〒120-8551 東京都足立区千住旭町 5 番

E-mail: <sup>†</sup> 19kmc17@ms.dendai.ac.jp, <sup>‡</sup> inoueu@mail.dendai.ac.jp

**あらまし** 近年、発表される学術論文の量は膨大になっており、研究者が必要とする論文を選択することが困難になっている。この問題を緩和するため、生理学・医学分野の論文を中心に **Structured Abstract** という抄録の記述方法が採用されている。これには抄録を構造化することで、閲覧者が効率的に論文の情報を得ることができる利点がある。本研究では、抽出型要約技術を利用して論文の本文から **Structured Abstract** を自動生成する手法を提案する。具体的には、自然言語処理で用いられる **BERT** モデルを科学技術分野に特化した **SciBERT** を利用して、**Structured Abstract** の各見出しに適合する文を抽出する。自動生成した要約と論文付属の要約の一致度を関連研究と比較することにより、提案手法の有効性を示す。

**キーワード** 自動要約, 抽出型要約, **Structured Abstract**, 学術論文, **SciBERT**, **BioBERT**

## 1. はじめに

近年、発表される学術論文の数は膨大になっており、高度な知識を持つ専門家であっても、論文を読み切ることは難しい。それ故に論文の選択が必要であるが、テキストが長大であり、内容が難解であることからそれも困難である。論文には通常、アブストラクトが記述されており、それには論文の内容が簡潔に示されているため、論文の選択に役立つ。しかしながら、どのような手段を用いたのか、結果がどうであったか、という点が瞬時に理解できる訳ではない。そのため、生理学・医学系論文の多くには図 1 に示すような **Structured Abstract**(構造化抄録)[1]と呼ばれるアブストラクトの記述方法が採用されている。これはアブストラクトを **Objective**(目的), **Method**(方法)といった見出しで構造化し記述されている。それぞれの項目が簡潔かつ明解な形式となっているため、理解がしやすい。

一方、自然言語処理分野では自動要約技術が開発されている。これは長大なテキストから短いテキストを生成する技術である。要約技術には抽出型要約と生成型要約の二種類の手法が存在する。抽出型要約は要約対象のテキストから重要であるとみなされた文を抽出する手法である。メリットとしては、文をそのまま抽出するという手法から文法的な間違いが発生しにくい点がある。そのため、人間が自然に読みやすい。デメリットとしては、人間がまとめるような文が生成できない点が存在する。生成型要約はモデルが単語を構成することで文を生成する手法である。メリットとしてはモデルが学習した単語を利用できることから、元のテキストの単語に依存せず、自由な表現が可能な点がある。デメリットとしては単語を構成していくという

手法から、文法的に正しくない文が生成される可能性がある。

本研究では、抽出型要約を用いて学術論文から **Structured Abstract** を自動的に生成する要約モデルの構築を検討する。論文の選択の支援を目的とするため、文法的に読みやすい抽出型要約モデルを選択した。モデルには言語モデル **BERT**[2]を利用した深層学習モデルを利用する。論文から事前学習した **BERT** である **SciBERT**[3]と **BioBERT**[4]をモデルに適用する。本論文では、**SciBERT** と **BioBERT** の有効性を関連研究と比較することで示す。

本論文は以下のように構成されている。2 章では学術論文の自動要約に関する研究を紹介する。3 章では本研究で用いるデータセットと深層学習モデルについて解説する。4 章では学習モデルへの **SciBERT** と **BioBERT** の適用方法について説明する。5 章では実験方法と評価結果、それらから得られた考察について述べる。6 章ではまとめと今後の課題について述べる。

### Abstract

**INTRODUCTION:** Although clinical use of dexmedetomidine (DEX), an alpha2-adrenergic receptor agonist admitted to intensive care units secondary to toxicological sequelae has not been well established.

**OBJECTIVES:** The primary objective of this study was to describe clinical and adverse effects observed for sedation.

**METHODS:** This was an observational case series with retrospective chart review of poisoned patients in an academic medical center. The primary endpoint was incidence of adverse effects of DEX therapy including seizures, and arrhythmias. For comparison, vital signs were collected hourly for the 5 h preceding the DEX therapy until the therapy ended. Additional endpoints included therapy duration; time within target F (RASS); and concomitant sedation, analgesia, and vasopressor requirements.

**RESULTS:** Twenty-two patients were included. Median initial and median DEX infusion rates were similar sedation. Median heart rate was lower during the therapy (82 vs. 93 beats/minute,  $p < 0.05$ ). Median systolic blood pressure was similar (111 vs. 109 mmHg,  $p = 0.745$ ). Five patients experienced an adverse effect. No additional adverse effects were noted. Median time within target RASS and duration of therapy was 6 h. Seventeen patients (77%) had concomitant use of other sedation and/or analgesia with four (23%) of the agents after DEX initiation. Seven patients (32%) had concomitant vasopressor support with four (57%) vasopressor support after DEX initiation.

**CONCLUSION:** Common adverse effects of DEX were noted in this study. The requirement for vasopressors further investigation into the safety of DEX in poisoned patients. Larger, comparative studies need to be conducted.

図 1. **Structured Abstract** の例(Mohorn ら[5]による **Abstract** を引用)

## 2. 関連研究

学術論文の自動要約を目的とする研究は数多く存在する．一般的に，論文の要約には通常のアブストラクトの生成を目的とした研究が多い．

中須賀ら[6]は学術論文の談話構造による特徴量に着目した自動要約手法を提案した．論文中の文や句の役割や話題の推移を特徴量として用いることで要約モデルの精度を向上させられることを示した．

衣川ら[7]は学術論文の章構造に着目した自動要約手法を提案した．Introduction や Related work といった章の構造を談話構造とみなし，リカレントニューラルネットワークを利用して要約生成を行った．章構造を用いて階層的に学習を行った結果，従来のモデルよりも評価を下げる結果となった．文の位置情報を伝搬させることが重要であることを示した．

安倍ら[8]は論文から 575 の形式で要約を生成する手法を提案した．575 の音韻的読みやすさを用いることで論文の把握と記憶の支援を目的としている．これは TF-IDF を用いることで単語の重要度を利用した出現頻度から要約を生成している．

Structured Abstract の自動生成に関しては麻生ら[9]による研究が挙げられる．これは Structured Abstract を採用している医学系論文をデータとして用い，ロジスティック回帰による分類器で学習させることで生成を行った．論文の文章は長大であるため，本文に TextRank や LexRank を用いて圧縮を行なったものをデータとして用いたが，モデルの精度を落とす結果となることを示した．本研究と異なる点は使用するモデルの違いと使用データの違いが挙げられる．本研究は，同じ評価指標を用いて，結果の比較を行う．

## 3. 前提

### 3.1. BERT

BERT[2]は，大規模なデータから双方向 Transformer モデルを用いて事前学習を行ったモデルである．マスクをかけたトークンを推定する Masked Language Model と，2 つの文を与え，それらが隣接しているかどうかを推定する Next Sentence Prediction という 2 つの言語タスクで事前学習をさせる．これらにより，トークンの関係と文の関係をj得る．この事前学習を行ったモデルに，様々なタスクに応じたレイヤーを追加し，Fine-tuning を行うことで高い評価を得ている．

### 3.2. SciBERT

SciBERT[3]は科学領域に特化した BERT モデルである．科学領域に特化するために，事前学習に Semantic Scholar からの論文を用いている．内部の構造は BERT と変わらないが，学習用データが論文に特化している．

データの数 は 114 万件であり，割合としてはコンピュータサイエンスが 18%，生物医学系論文が 82%含まれている．通常の BERT の語彙とこちらの語彙は重複している割合が 42%であり，一般の文章と論文の文章において使用される単語の違いが現れている．トークンの分割には SentencePiece による手法が用いられている．これらにより，生物医学系タスクで高い評価を得ている．

### 3.3. BioBERT

BioBERT[4]は生物医学領域に特化した BERT モデルである．この領域に特化するために，事前学習に生物医学系論文を用いている．SciBERT と同様，内部の構造は通常の BERT と変わらない．学習用データには英語版 Wikipedia，BookCorpus，PubMed に掲載されている論文のアブストラクト，PMC に掲載されている論文の全文を用いている．トークンの分割には WordPiece による手法が用いられている．科学領域に特化した BERT モデルであるが，Wikipedia と BookCorpus を用いている点と論文の分野が SciBERT とは異なる．

### 3.4. 学習モデル

BERTSUMEXT[10]は言語モデルの BERT を用いた抽出型モデルである．図 2 にモデルの構造を示す．抽出型要約のためにトークン単位ではなく，文単位の表現を可能にしている．複数の[CLS]トークンを各文の先頭に挿入することで，前の文の特徴を得られるよう設計されている．Token Embeddings はトークンの情報を表している．Segment Embeddings により文には奇数と偶数のラベルが与えられ，文の区別が可能になる．Position Embeddings では文章におけるトークンの位置を表している．これらを組み合わせて Transformer レイヤーで要約を行う．また，Trigram Blocking という機能を持ち，冗長性を減らしている．初めに選択された文とその他の文を比較し，連続した 3 単語の重複が無ければその他の文を選択するというフィルタリングを行う．

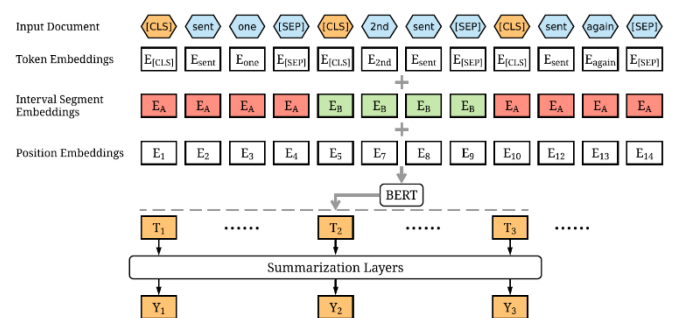


図 2. BERTSUMEXT のモデル(Liu ら[11]による図から引用)

## 4. 提案手法

第 2 章で関連研究について述べたが、**Structured Abstract** の自動生成に関する研究は少ない。そこで、本研究では要約に特化したモデルを用いて、**Structured Abstract** の自動生成モデルを作成する。麻生らの研究では、**TextRank** や **LexRank** によって前処理を行うことで、出力文の精度を上げることを期待したが、それらの前処理がむしろ精度を下げることを示した。それに倣い、本研究では圧縮等の処理を施さずにデータを学習に利用する。本研究では要約生成モデルとして **BERTSUMEXT** を用いる。通常、**BERT** の **Position Embedding** の最大長は 512 であるが、ランダムに初期化される **Position Embedding** を追加しており、論文のような長文をデータとして扱う本研究との相性は良いと考えられる。また、**BERT** を専門的な文章に対応した事前学習モデルに変更することで、論文テキストに対して評価結果が向上する可能性があると考えられる。そこで、医学系タスクで高い評価を得ている **SciBERT** と **BioBERT** を事前学習モデルとして用いる。

**BERTSUMEXT** を用いる手順を以下に示す。

- ① **Stanford CoreNLP** による学習データの前処理
- ② 前処理後の学習データの **JSON** 化
- ③ **JSON** 化されたデータを **PyTorch** に入力可能形式 (**BERT Data**) に変更
- ④ **BERT Data** を用いて学習

**SciBERT** と **BioBERT** をそれぞれ適用するためには、手順③において専用の語彙リストファイルを用いる必要がある。**SciBERT** には **Allen Institute for AI** より提供されている **scivocab**[12]を用いる。**BioBERT** には **Google** より提供されている **vocab** ファイル[13]を利用する。この **vocab** ファイルは通常の **BERT** と変わらないが、これ **BioBERT** における生物医学コーパスでの専門用語をサブワードで表すためである。また、手順④にて読み込む **BERT** のモデルを **SciBERT** と **BioBERT** にそれぞれ変更する。**SciBERT** には語彙リストファイルと同様に **Allen Institute for AI** によって事前学習されたモデルと設定ファイルを利用する。**BioBERT** には **NAVER** によって事前学習されたモデル[14]と設定ファイルを利用する。

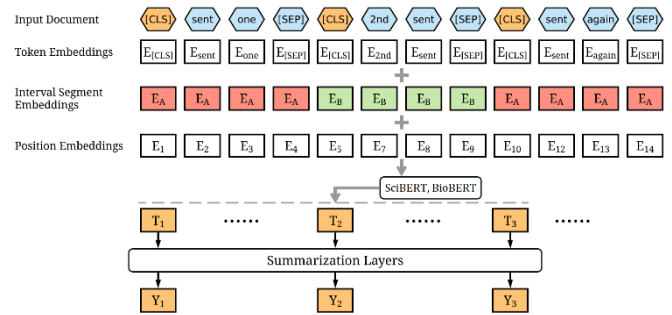


図 2. **SciBERT**, **BioBERT** をそれぞれ適用した **BERTSUMEXT** のモデル(Liu ら[11]による図を一部修正し引用)

## 5. 実験

### 5.1. 実験方法

3, 4 章で示したモデルを用いて実験を行う。前章で示した **SciBERT** を適用したモデル, **BioBERT** を適用したモデル, そして **BERT-Base Uncased** を適用したモデルでそれぞれ学習を行う。**BERT-Base Uncased** は小文字化された **Wikipedia** と **BookCorpus** から学習を行ったモデルである。これは汎用的に利用可能であり、本研究でも一定の評価が得られると考えられるため、こちらを利用した実験も行う。実装には **Pytorch** を用い、事前学習モデルの読み込みにはライブラリである **pytorch-pretrained-BERT** を利用した。

モデルのレイヤー数は Liu ら[14]の研究に倣い 2 層とし、**Trigram Blocking** を利用する。

学習用データには **CORE Dataset**[15]を用いる。このデータセットには幅広い分野の論文テキストデータが含まれている。そのため、様々な分野の論文をデータとして用いることが可能であると考えられる。これにより、モデルが対応できる論文の種類の増加が期待できる。データセットから、**Structured Abstract** を採用している英文の論文データ 5000 件を実験に用いる。テキストデータには、小文字化、文分割等の処理を行う。正解データには本文と **Structured Abstract** の見出しに最も近いと考えられる文を貪欲法による単語の一致度から算出し、本文データと対にして用いる。

生成する見出しは **Objectives**(目的), **Methods**(方法), **Results**(結果), **Conclusions**(結論)とする。麻生らの研究ではこれらの 4 つの見出しに加えて **Background**(背景)も生成しているが、**Objective** と重複する見出しであると考え、本研究では生成せず、**Objective** のデータと同様に扱う。また、学習データとテストデータの割合は 8:2 とする。

評価では 2 章で紹介した麻生らによる研究による評価結果と比較する。

## 5.2. 評価

要約モデルの評価には ROUGE[16]という手法を用いる。モデルによって生成された要約と参照された要約の単語の一致度を利用し、類似度を算出する手法である。ROUGE-N は N-gram 単位で一致度を算出する。式は以下のとおりである。

$$\text{ROUGE} - N = \frac{\sum_{S \in \{\text{Reference Summaries}\}} \text{gram}_n \in S \sum \text{Count}_{\text{match}}(\text{gram}_n)}{\sum_{S \in \{\text{Reference Summaries}\}} \text{gram}_n \in S \sum \text{Count}(\text{gram}_n)} \quad (1)$$

N は N-gram の長さを表し、 $\text{Count}_{\text{match}}$  はモデルの生成要約と参照元の要約の N-gram の最大一致度を示す。本研究では、uni-gram(N=1), bi-gram(N=2)として評価する。

また、ROUGE-L という手法も用いる。こちらは参照された要約とモデルによって生成された要約の Longest-common-subsequence を利用する。式を以下に示す。

$$R_{lcs} = \frac{LCS(X, Y)}{m} \quad (2)$$

$$P_{lcs} = \frac{LCS(X, Y)}{n} \quad (3)$$

$$\beta = \frac{P_{lcs}}{R_{lcs}} \quad (4)$$

$$\text{ROUGE} - L = \frac{(1 + \beta^2)R_{lcs}P_{lcs}}{R_{lcs} + \beta^2P_{lcs}} \quad (5)$$

$LCS(X, Y)$  は参照された要約とモデルによって生成された要約の Longest-common-subsequence である。これは二つの文章の共通部分列のうち、最長の部分列のことであり、その長さを用いる。 $m$  は参照となる要約の単語数、 $n$  は生成された要約の単語数である。それらから適合率  $P_{lcs}$ 、再現率  $R_{lcs}$  を求め、その重み付き F 値をスコアとしている。

表 1 に本研究における評価の理論値を示し、2, 3 に SciBERT と BERT-Base Uncased による評価結果をそれぞれ示す。表 1 の理論値は、Structured Abstract の見出しに対する正解データの評価値である。正解データは見出しに最も近い文であるが、著者が自由に記述した見出しと完全に一致する可能性は低い。そのため、表 1 の結果が本研究での理論上の上限値となる。表 2, 3 に示す評価結果も同じく、正解データに対する評価値ではなく、Structured Abstract の見出しに対する、モデルの生成要約の評価値である。また、SciBERT による Structured Abstract の例を掲載する。

表 1. 理論値(%)

見出し	ROUGE-1	ROUGE-2	ROUGE-L
Objectives	47.7	29.9	42.7
Methods	43.7	26.1	40.3
Results	43.3	25.1	40.2
Conclusions	57.3	40.1	52.7

表 2. SciBERT による評価結果(%)

見出し	ROUGE-1	ROUGE-2	ROUGE-L
Objectives	37.8	20.1	35.8
Methods	36.8	16.7	31.8
Results	34.4	14.0	29.5
Conclusions	43.2	25.7	40.3

表 3. BioBERT による評価結果(%)

見出し	ROUGE-1	ROUGE-2	ROUGE-L
Objectives	38.0	18.0	33.8
Methods	33.7	14.7	29.8
Results	32.7	12.8	28.8
Conclusions	44.2	26.8	40.8

表 4. BERT-Base Uncased による評価結果(%)

見出し	ROUGE-1	ROUGE-2	ROUGE-L
Objectives	35.6	15.9	31.6
Methods	33.9	14.7	30.9
Results	36.1	15.1	32.8
Conclusions	42.7	24.3	38.5

表 5. Martín-Fernández ら[17]による掲載元の Structured Abstract

<p><b>RATIONALE AND OBJECTIVES:</b></p> <p>Economic theory classifies an intervention as socially beneficial if the total Willingness to Pay (WTP) of those who gain exceeds the total Willingness to accept (WTA) of those who are harmed. This paper examines the differences in health system users' valuation of a health care service in primary care setting based on the WTP and WTA perspectives, discussing the impact of personal and service variables, including risk attitudes, on these disparities.</p> <p><b>METHOD:</b></p>
--

Six hundred and sixty two subjects who asked for care in health centres in the Region of Madrid (Spain) were interviewed, using the contingent valuation method to estimate WTP and WTA. Patient sociodemographic characteristics, health needs, satisfaction with the service and risk attitude and behaviour under risk (measured by self-reported scales and lottery games respectively) were collected. Generalised Linear Models were used to estimate the association between the explanatory variables and the WTA/WTP ratio.

#### RESULTS:

We obtained the WTA/WTP ratio for 570 subjects (mean 1.66 CI 95%: 1.53–1.79; median 1, interquartile range 1–2). People with higher education or in high social groups expressed WTA values closest to WTP. The opposite occurred in patients with the greatest health needs or who were born abroad. Self-reported expression of risk aversion appeared also related to increases in the WTA/WTP ratio. Satisfaction with the service evaluated was the most influential factor in the WTA/WTP ratio.

#### CONCLUSION:

Health need, difficulty in obtaining substitutes and satisfaction with the service could serve for profiling people averse to loss for health care services in primary care setting. Self-reported expression of risk aversion could also be related to increases in the WTA/WTP ratio. This would mean that these characteristics should be taken into account both in the design and implementation of new healthcare interventions, as in the making decision for disinvestment.

implications for the expressed values, we chose as realistic a description as possible, making reference to a good immediately obtained.

#### CONCLUSION:

Health needs and use of services were measured by the existence of chronic conditions (ones that require continuous healthcare for more than 6 months), hospital admissions in the past year (including emergency stays lasting longer than 24 hours) and the perception of health-related quality of life (HRQoL) as measured by EQ-5D. The EQ-5D results were expressed on a visual scale and turned into utilities.

### 5.3. 考察

#### 5.3.1. 科学領域 BERT モデルの効果

汎用的な BERT である BERT-Base Uncased と SciBERT, BioBERT による評価結果を比較する。SciBERT を利用したモデルにおいて、関連研究と比較すると Objectives, Methods, Conclusions の値が大きく向上している。BioBERT を利用したモデルでも同様の見出しで値が向上している。BERT-Base Uncased と比較すると Results 以外の見出しで、値が向上しており、見出しによっては特化型 BERT が有効であることを示した。Results で BERT-Base Uncased が高い値を得た原因には Results とそれ以外で使われた単語の違いによるものと推測される。Results には汎用的な単語が多く含まれていると考えられる。

#### 5.3.2. 関連研究との比較

関連研究と比較するため、表 7 に麻生らによるモデルの評価結果を示す。ROUGE-1 による評価結果のみが示されているため、同じ指標で比較を行う。採用しているデータが異なることから、厳密に比較が出来ない。特に理論値においては Results の値が関連研究による値よりも低い点に注意する必要がある。SciBERT, BioBERT を利用したモデルでは Objectives, Methods, Conclusions の値が大きく向上している。そのため、本研究で提案したモデルの性能が有効であると判断できる。関連研究における Results の値は理論値よりも 1.2 ポイント高い。麻生らによる研究では、医学系論文のみをデータとして採用している点が本研究とは異なる。その点で、本研究で扱った様々な分野の論文とは記述方法の差異が存在する。その部分が影響を及ぼした可能性がある。

表 6. モデルによって生成された Structured Abstract

#### OBJECTIVES:

The value attributed to a good or service with the CV method can be studied from the perspective of willingness to pay (WTP) or to accept (WTA).

#### METHOD:

Given the asymmetry in the distribution of the dependent variable (WTA/WTP ratio), we used Generalised Linear Models (GLM) to estimate the association between the explanatory variables and the WTA/WTP ratio.

#### RESULTS:

Given that the description of the scenario may have



表 7. 麻生らによる評価結果(%)

見出し	ROUGE-1
Objectives	28.0
Methods	31.7
Results	44.5
Conclusions	35.1

## 参 考 文 献

### 5.3.3. 参照 Structured Abstract とモデルが生成した Structured Abstract の比較

Martín-Fernández らによる研究を対象として、Structured Abstract の比較を行う。抽出型要約のメリットである、文法の正しさがどの項目でも見てとれる。Objectives では、「WTP と WTA の観点に基づく」という点が正しく抽出できている。Methods では、WTA/WTP の比に関連する箇所が正しく選択されている。Results では正しく抽出できず、全く別の文が選択されている。Conclusions では先頭 2 単語が近いものの内容としては別の文が選択されており、正しく抽出されていない。

## 6. まとめ

本稿では SciBERT と BioBERT による BERTSUMEXT を用いた Structured Abstract の生成方法を提案した。ROUGE-1 による評価では、SciBERT と BioBERT 共に 3 つの見出しで従来のモデルを上回る結果となった。

今後の課題としては使用データの調査が挙げられる。本研究では Core Dataset の Structured Abstract を含む論文をデータとして用いた。Core Dataset は広い分野の論文を含んでいるため、広い範囲に対応できるモデルの作成を検討していた。しかし、実際にデータに用いた論文の研究分野は生理学・医学系論文に偏っていると考えられる。データセットにはどの分野の論文であるかを示すラベルが存在しないため、データがどのような分野であるかが、完全に判明していない。今後、キーワード等を利用して、分析をする予定である。また、論文の章に着目し、それを利用して生成することで構造を活かすことができると考えられる。その他にも広い範囲の分野に対応した BERT を作成することで、対象を広げられることが期待できる。本研究で用いた科学分野特化型 BERT は対応分野が医学系に偏っている。そのため、BERT の対応分野を広げる必要がある。現在は、ローコストの BERT も研究されており、そういった BERT の活用が見込める。本研究では人間による論文の取捨選択を支援することを目的としている。そのため、ROUGE による評価だけでなく人間による定性的な評価が必要であり、検討をしていきたい。

- [1] 東京慈恵会医科大学 学術情報センター 編集室, <http://www.jikei.ac.jp/academic/micer/ronbun.htm>.
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova., “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 4171–4186. 2019.
- [3] Iz Beltagy, Arman Cohan, and Kyle Lo., “Scibert: Pre-trained contextualized embeddings for scientific text.” arXiv preprint arXiv:1903.10676, 2019.
- [4] Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, and Jaewoo Kang., “BioBERT: pre-trained biomedical language representation model for biomedical text mining.” arXiv preprint arXiv:1901.08746. 2019
- [5] Mohorn PL, Vakkalanka JP, Rushton W, Hardison L, Woloszyn A, Holstege C, Corbett SM. Evaluation of dexmedetomidine therapy for sedation in patients with toxicological events at an academic medical center. Clin Toxicol, 2014;52:525-530.
- [6] 中須賀謙吾, 鶴岡慶雅, “談話構造を利用した学術論文の自動要約生成”, 言語処理学会第 21 回年次大会発表論文集, pp.569-572, 2015.
- [7] 衣川和亮, 鶴岡慶雅. “学術論文の章構造に基づくニューラル自動要約モデル”, 言語処理学会第 23 回年次大会発表論文集, pp. 150-153, 2017.
- [8] 安部文紀, 寺田実 “575 の音韻的読みやすさを付与した学術論文の要約文自動生成手法”, DEIM Forum 2018 E3-1
- [9] 麻生栄樹, 松本裕治., “科学技術論文からの Structured Abstract の自動生成に関する研究”, 修士論文, 奈良先端科学技術大学院大学, 2016
- [10] Yang Liu and Mirella Lapata., “Text summarization with pretrained encoders.”, arXiv preprint arXiv:1908.08345.
- [11] Yang Liu. Fine-tune BERT for extractive summarization. arXiv preprint arXiv:1903.10318, 2019
- [12] GitHub - allenai/scibert: A BERT model for scientific text, <https://github.com/allenai/scibert>
- [13] GitHub - google-research/bert: TensorFlow code and pre-trained models for BERT, <https://github.com/google-research/bert>
- [14] GitHub - naver/biobert-pretrained: BioBERT: a pre-trained biomedical language representation model for biomedical text mining, <https://github.com/dmis-lab/biobert>
- [15] Knoth, Petr and Zdrahal, Zdenek., “CORE: three access levels to underpin open access.”, D-Lib Magazine, 18 (11/12), 2012
- [16] Chin-Yew Lin., “Rouge: A package for automatic evaluation of summaries.”, ACL'04, pp.74-81, 2004.
- [17] Martín-Fernández J, Ariza-Cardiel G, Peña-Longo bardo LM, Polentinos-Castro E, Oliva-Moreno J, Gil-Lacruz AI, Medina-Palomino H, Del Cura-González I., ““Gaining or losing”: The importance of the perspective in primary care health services valuation.”, PLOS ONE. 2017;12(12):e0188969. DOI:10.1371/journal.pone.0188969.