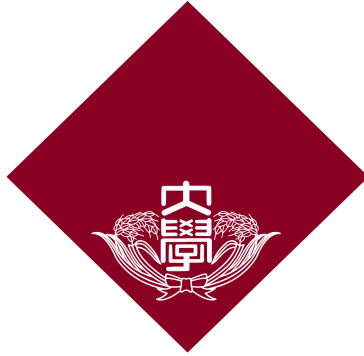


令和3年度 修士論文



# 抽出型自動要約における語彙バイアスの増幅度に関する分析

指導教員 酒井 哲也 教授  
研究指導名 情報アクセス研究

早稲田大学大学院 基幹理工学研究科 情報理工・情報通信専攻

学籍番号 5120F034

清水 嶺

提出日 2022年1月24日



## 概要

検索エンジンにおけるスニペットや、ソーシャルメディアの投稿としてニュース記事を掲載する際、ニュース記事の要約が掲載されることがある。このような要約を生成する際、自動要約が用いられることがある。一般に、ニュース記事本文には、書き手や報道機関によるニュースイベントに対する見解が含まれている。これらの見解の中には、語彙バイアスとよばれる、極性の高い単語や表現を用いることでニュースイベントに対する理解を妨げたりするものがある。自動要約をニュース記事の要約に利用する際、語彙バイアスが含まれる文を抽出しすぎてしまうことで、元の文書に含まれるバイアスを増幅してしまう可能性がある。ユーザーは、自動要約のみを読んでニュースを理解しようとする可能性があるが、要約が本文の語彙バイアスを増幅していた場合、ニュースに対する誤った理解を招いたり、ソーシャルメディアを通じた意図しない拡散を引き起こす可能性がある。そのため、自動要約は語彙バイアスを増幅していないことが望ましい。本研究では、要約アルゴリズムにおける語彙バイアスの増幅度を定量化する指標を示したのちに、既存の抽出型要約アルゴリズムを対象として語彙バイアスの増幅の程度を比較し、分析を行った。その結果、以下の結果が得られた。

- a) 増幅の程度は抽出型要約アルゴリズムによって差異がある。
- b) 抽出型要約において重要であると考えられる特徴量は、語彙バイアスを含む文の抽出しやすさに繋がる可能性がある。
- c) 提案指標は、要約アルゴリズムの評価で広く用いられる ROUGE だけでは評価できない観点での評価が可能である。



# 目次

<b>第 1 章</b>	<b>導入</b>	<b>7</b>
1.1	本研究の概要 .....	7
1.2	語彙バイアスと情報バイアス .....	8
<b>第 2 章</b>	<b>関連研究</b>	<b>11</b>
2.1	自動要約 .....	11
2.2	ニュースメディアとバイアス .....	11
2.3	機械学習におけるバイアスの問題 .....	12
<b>第 3 章</b>	<b>提案指標</b>	<b>13</b>
3.1	抽出型要約における語彙バイアスの指標 .....	13
3.2	教師あり抽出型要約における、語彙バイアスの観点からの特徴量の寄与率 .....	14
<b>第 4 章</b>	<b>実験設定</b>	<b>15</b>
4.1	実験対象となるアルゴリズム .....	15
4.2	実験に使用するデータセット .....	16
<b>第 5 章</b>	<b>抽出型要約アルゴリズムの出力における語彙バイアスの分析</b>	<b>17</b>
5.1	アルゴリズムごとの MSBS の比較 .....	17
5.2	語彙バイアスのある文を含んだ要約と語彙バイアスのある文を含まない要約の比較 .....	18
5.3	教師なし抽出型要約の結果の考察 .....	20
5.3.1	教師なし抽出型要約における、TF-IDF 特徴量が及ぼす MSBS の変化に対する分析 .....	20
5.4	教師あり抽出型要約の結果の考察 .....	22
5.4.1	抽出型要約における語彙バイアスの観点からの特徴量分析 .....	22

目次	5.4.2	BASIL データセットにおける，抽出型要約において重要な特 徴量と語彙バイアスの関係の分析 .....	23
	5.4.3	深層学習を用いた教師あり抽出型要約アルゴリズムの結果の 考察 .....	25
	5.4.4	深層学習を用いた抽出型要約における，文の位置に関する特 徴量の分析 .....	25
	5.4.5	MSBS と ROUGE の関係の分析 .....	26
第 6 章	結論		29
参考文献			33

# 第1章 導入

## 1.1 本研究の概要

ニュース記事本文には、書き手や報道機関のニュースイベントに対する見解や主張が含まれていることが多い。このような見解や主張の中には、語彙バイアスと呼ばれる、極性の高い単語や表現を用いることで、ニュースイベントに対して極端に肯定的や否定的な見解を与えるような可能性があるものが存在する。このようなニュース記事の要約を生成する際、自動要約の技術が用いられることがある。しかし、ニュース記事の自動要約は、語彙バイアスを含んだ文を多く抽出することで、本来ニュース記事に含まれていた語彙バイアスを増幅してしまう可能性がある。このような問題は、特定の状況下において望ましくない結果を引き起こす可能性がある。例えば、ソーシャルメディアにおけるニュース記事の投稿がある。Twitter<sup>\*1</sup>のようなソーシャルメディアが広く使われるようになるにしたがって、ユーザーは Web 上のニュース記事もソーシャルメディアを通して入手することが多くなった。それに伴い、ユーザーは様々な種類のニュースを偶発的に手に入れる機会が増えている。しかし、偶発的なニュースの多くはユーザーの興味を引かない [1]。また、ソーシャルメディアを通して目に入るニュースが増えたことにより、ユーザーはしばしば情報過多を経験する。このような情報過多から身を守るために、ニュースの流し読みが行われる [2]。ソーシャルメディアの投稿には、ニュース記事本文へのリンクに加えて、ニュースの要約が合わせて投稿されることがある。そのため、ニュースの流し読みを行うユーザーは、これらの要約を読んでニュースを理解する可能性がある。このとき、ニュースの要約が語彙バイアスを増幅していた場合、ユーザーはニュースイベントに対して、極端に肯定的、または否定的な見解を抱く可能性がある。さらに、このようなソーシャルメディアには再投稿機能が存在していることが多く、ニュースイベントに対する偏った理解の上で再投稿された場合、提供者が意図しない形でニュースが拡散する可能性がある。このように、要約が多く読まれる可能性のある状況においては、要約が語彙バ

---

\*1 <https://twitter.com>

## 第1章 導入

イアスを増幅していることは望ましくない。

一般的な教師あり抽出型要約アルゴリズムでは、評価関数中にこのような語彙バイアスを考慮するような制約項を設けていない。また、人によるアノテーションを行ったデータセットを用いて学習させるため、アノテーター自身の見解がラベルに反映されている可能性もある。そのため、教師あり抽出型要約アルゴリズムの要約結果は、語彙バイアスを含んでいる文を重点的に抽出している可能性がある。本研究は、抽出型要約アルゴリズムがどれほど語彙バイアスを含む文を抽出しやすいのかを調べることを目的としている。まず、抽出型要約アルゴリズムがどの程度語彙バイアスを含む文を抽出しやすいのかを定量的に評価するための指標を提案する。その後、いくつかの代表的な抽出型要約アルゴリズムの出力に対してこの指標を適用し、アルゴリズムによってどのように語彙バイアスのある文の抽出しやすさが異なっているのかを調べる。

本論文の貢献は以下の通りである。

- 提案指標を用い、実際に既存の抽出型要約アルゴリズムに対して適用することで、アルゴリズムごとに語彙バイアスを含む文の抽出しやすさが大きく異なっていることが分かった。
- 要約において重要な文を抽出するのに重要であると言われている文の位置、重要単語数、文の長さに関する特徴量は、語彙バイアスを含んだ文を抽出しやすくする可能性を示した。
- 提案指標は、要約アルゴリズムの性能を評価するための指標として広く用いられている ROUGE [3] では捉えられない観点で要約アルゴリズムを評価できる。そのため、この指標を ROUGE と合わせて使用することで、より異なった観点からの要約アルゴリズムの評価が可能であることを示した。

## 1.2 語彙バイアスと情報バイアス

メディアにおけるバイアスは、特定の文における表現や単語の極性によって導入される語彙バイアスと、極性の強い表現を用いているわけではないが、読者の事実関係の把握を妨げる可能性のある文によって導入される情報バイアスが存在する [4]。表 1.1 にそれぞれのバイアスの例を示す。

表 1.1 に示したニュースは、ポール・ライアン氏とバラク・オバマ氏によるメディアケア政策に対する議論を表している。語彙バイアスのある文の例として示した文は、“dishonest and hypocritical” という表現が語彙バイアスの根拠として示されている。この表現は強い極性を持っているため、読者にはその極性に基づいた偏見を抱かせる



A sentence with lexical bias
Obama campaign spokeswoman Lis Smith described the new Romney-Ryan ad on the subject as “dishonest and hypocritical,” considering Ryan’s own proposals for Medicare.
A sentence with informational bias
Paul Ryan went on offense Tuesday in response to criticism over his Medicare plan, using an interview with Fox News coupled with a new TV ad to claim President Obama’s health care plan treats the treasured entitlement like a “piggy bank” while the “Romney-Ryan” plan preserves it.

表 1.1 情報バイアスと語彙バイアスのある文の例

可能性がある。このように、極性の強い単語やフレーズが含まれている文は語彙バイアスのある文として扱われる。一方、情報バイアスがある文の例として示した文は、*President Obama’s health care plan treats the treasured entitlement like a “piggy bank” while the “Romney-Ryan” plan preserves it.* という一文が情報バイアスの根拠とされている。この文では、バラク・オバマ氏が制度を私物化しているかのように読者に偏見を抱かせるような内容になっているが、バイアスがあることを決定づけるような極性の強い単語やフレーズは存在しないため、語彙バイアスがある文としては扱われない。このように、文の目に見える特徴からバイアスの有無を見分けるのは難しいが、そのニュースが報じるイベントに対して異なった意見を読者に抱かせるようなものは、情報バイアスのある文として扱われる。本実験では、語彙バイアスを対象とした実験を行う。情報バイアスを機械的に捉えるのは一般的に困難とされており [5]、それは本実験においても例外ではないためである。



## 第2章 関連研究

### 2.1 自動要約

自動要約は大きく分けて抽出型要約と抽象型要約に分けることができる。抽出型要約は文書から要約に含ませるべき文をいくつか選択してつなぎ合わせることで要約文を生成する方法である。抽象型要約は、元となる文から新たに要約文を生成する方式である [6]。本実験では抽出型要約を対象とする。抽出型要約は教師ありアルゴリズムと教師なしアルゴリズムが存在する。教師なしアルゴリズムとして代表的なものは、LexRank [7] や TextRank [8] がある。これらの方法は、文グラフを構築し、固有値を抽出することで重要度のスコアリングを行う方法である。LexRank は TF-IDF をもとに文グラフを構築し、TextRank は単語の重複度をもとに文グラフを構築する。教師ありアルゴリズムは、SVM、決定木を用いた方法などが研究されていたが、近年ではデータセットの充実や、ニューラルネットの発展により、ニューラルネットを用いたものが広く研究されるようになってきている [6]。代表的なものとしては、RNN を用いた SummaRuNNer [9] や、事前学習可能な言語モデルの BERT (Bidirectional Encoder Representations from Transformers) [10] を用いた BERTSum [11] などが存在する。

### 2.2 ニュースメディアとバイアス

ニュースは一般的に提供者によるバイアスを含んでいる。ニュースにおけるバイアスは、著者や報道機関のニュースイベントに対する見解によって生み出され、ときには特定のニュースイベントに対する読者の理解を妨げる可能性がある [12]。adfontesmedia<sup>\*1</sup> では、主要なニュースメディアの思想の傾向を示しており、中立的な立場で記事を書いていないニュースメディアが多く存在することがわかる。このように、バイアスは時に望ましくない影響をもたらすため、情報科学の観点からメディアのバイアスを分析したり、バイアスの自動的な検出を行うような研究がある。このよ

---

\*1 <https://adfontesmedia.com/>

## 第2章 関連研究

うな研究の中には、政治的な観点によって生じるバイアスに取り組む研究や [5] [13] [14], ニュースイベントの事実関係を誤解させるようなバイアスに取り組む研究がある [15] [16].

### 2.3 機械学習におけるバイアスの問題

機械学習は様々なタスクにおいて用いられるが、機械学習においては、解きたいタスクに応じたデータセットを元にモデルの学習を行う。しかし、学習データ内の、何らかの保護変数に対して偏りがあった場合、モデルの出力にバイアスを含ませてしまうことがある [17]. モデルの学習に用いられるデータは Web 上から収集されたものが多いため、Web が潜在的に含んでいるバイアスをモデルの推論に反映してしまう可能性がある [18]. また、多くの深層学習を用いたタスクにおいて、事前学習を行えるモデルを用いた手法が非常に高い性能を達成している。しかし、これらのモデルの事前学習に用いたデータにバイアスが存在していた場合、出力にもバイアスが存在している可能性が指摘されている [19].

抽出型要約におけるバイアスの研究としては、ソーシャルメディアの投稿のような、さまざまな属性を持つユーザーが生成した文をもとに意見要約を行う際、特定の属性を持つユーザーの投稿のみが要約に適した文として抽出されてしまうようなバイアスに対処する研究が存在する。Dash らは、FairSumm と呼ばれる意見要約手法を提案した [20]. Shandilya らは、意見要約における公平性の問題について調査し、公平性という観点で意見要約の品質を定量化するための指標を考案した [21]. しかし、本研究で扱うような、単一文書を対象とした抽出型要約におけるバイアスの問題は研究されていない。

# 第 3 章 提案指標

## 3.1 抽出型要約における語彙バイアスの指標

抽出型要約における語彙バイアスの定量化指標として、SBS (Summary Bias Score) と呼ばれる指標を提案する。要約の対象となる文書集合  $\mathcal{D}$  があるとき、 $i$  番目の文書を  $D_i (\in \mathcal{D})$  で表す。  $B_i (\subset D_i)$  は、語彙バイアスがあるとラベリングされている文の集合を示す。また、語彙バイアスが含まれる文は  $b_{i,j} \in B_i$  とする。要約アルゴリズム  $C$  が、文書  $D_i$  における  $k$  番目の文  $s_{i,k} (\in D_i)$  に対して付与したスコアは、  $C(s_{i,k}, D_i)$  で表される。要約アルゴリズム  $C$ 、文書  $D_i$  における SBS を  $SBS(C, D_i)$  とすると、

$$SBS(C, D_i) = \frac{1}{|B_i|} \sum_{b_{i,j} \in B_i} \frac{C(b_{i,j}, D_i)}{\sum_{s_{i,k} \in D_i} C(s_{i,k}, D_i)} \quad (3.1)$$

要約アルゴリズムは、抽出型要約における文の重要度を数値としてスコアリングするものと考えることが出来る。スコアが高ければ高いほど、要約文としての重要であることを示している。SBS は、語彙バイアスが含まれている文に対して要約アルゴリズムが割り振ったスコアを正規化した値の平均である。なお、バイアスが含まれている文の数は文書によって異なっているため、語彙バイアスが含まれている文の数で正規化を行う。SBS が大きければ大きいほど、その要約アルゴリズムは語彙バイアスのある文を要約文として抽出しやすいという傾向があるということを示している。本実験でアルゴリズムごとの SBS を比較する際は、文書集合  $\mathcal{D}$  を対象とした SBS の算術平均を用いる。これは以下の式で定義される。

$$MSBS(C, \mathcal{D}) = \frac{1}{|\mathcal{D}|} \sum_{D_i \in \mathcal{D}} SBS(C, D_i) \quad (3.2)$$

要約の出力に語彙バイアスを含む文が含まれづらいことが期待される場合、MSBS が小さいアルゴリズムを選定することが望ましい。

## 3.2 教師あり抽出型要約における，語彙バイアスの観点からの特徴量の寄与率

手動で作成した特徴量に基づく教師あり抽出型要約において，特徴量と語彙バイアスの関係进行分析するために，SFPI (Summary Feature Permutation Importance) という指標を提案する．SFPI は，あるアルゴリズムに対して特定の特徴量が，どれほど MSBS を増大させるのに寄与しているのかを調べるための定量的な指標である．SFPI が正の場合，その特徴量が語彙バイアスを含む文を重要であると選択しやすくすることに寄与してしまうことを示している．SFPI が大きければ大きいほど，抽出型要約における語彙バイアスを増幅しやすい特徴量であると言える．対して，SFPI が負の場合，その特徴量が語彙バイアスを含む文を重要であると選択しにくくすることに寄与している．SFPI が小さければ小さいほど，抽出型要約における語彙バイアスを低減しやすい特徴量であると言える．この指標は，機械学習において，要約アルゴリズムの推論における重要な特徴量を定量的に扱うための Permutation Importance [22] と呼ばれる指標に着想を得ている．SFPI の計算方法を以下に示す．

アルゴリズムの評価に用いるデータセットを  $\mathcal{G}$  とする．このデータセット  $\mathcal{G}$  は表形式のデータであり，行がそれぞれの文データに対応し，列がそのデータの特徴量  $f$  を示す．これは，文書集合  $\mathcal{D}$  を表形式のデータに変換し，特徴量に関する列を付加することによって得られる．アルゴリズム  $C$  に対して，データセット  $\mathcal{G}$  を用いた際の MSBS を  $MSBS(C, \mathcal{G})$  とする． $\mathcal{G}$  において，特徴量  $f$  のデータ，すなわち列  $f$  をランダムにシャッフルしたデータセットを  $\mathcal{G}_f$  とする． $K$  は試行回数とした時， $k$  回目の試行における特徴量  $f$  の SFPI を  $MSBS(C, \mathcal{G}_f^k)$  とする．この時，アルゴリズム  $C$  において，特徴量  $f$  の SFPI となる  $SFPI(C, \mathcal{G}_f)$  は以下のように計算される．

$$SFPI(C, \mathcal{G}_f) = MSBS(C, \mathcal{G}) - \frac{1}{K} \sum_{k=0}^{K-1} MSBS(C, \mathcal{G}_f^k) \quad (3.3)$$

この式においては， $SFPI(C, \mathcal{G}_f) > 0$  の場合，特徴量  $f$  を無効にした結果 MSBS が下がっているという事実が得られるため，アルゴリズム  $C$  において特徴量  $f$  が MSBS の増大に寄与してしまっていると言える．対して， $SFPI(C, \mathcal{G}_f) < 0$  の場合，特徴量  $f$  を無効にした結果 MSBS が上がっているという事実が得られるため，特徴量  $f$  が MSBS の低減に寄与していると言える．

# 第4章 実験設定

## 4.1 実験対象となるアルゴリズム

本実験では、実験対象の抽出型要約アルゴリズムとして教師なしのものと教師ありのものを用意し、それぞれ比較を行った。教師なしのアルゴリズムとして LexRank [7] と TextRank [8] を用いる。教師ありのアルゴリズムとして SummaRuNNer [9] と BERTSum [11] を用いた。これらのニューラルネットを用いた教師ありのアルゴリズムは、双方とも二階層のモデルアーキテクチャで構成されている。一階層目は Encoder である。SummaRuNNer における Encoder は、単語を入力とし中間表現を出力する RNN で構成されている [9]。BERTSum における Encoder は、文を入力として中間表現を出力する事前学習済み BERT である [11]。二階層目は Summarization Layer である。Summarization Layer は Encoder によって得られた文の中間表現を入力とし、入力の表現をもとに要約に含ませるか否かを示す二値ラベルを出力する。SummaRuNNer は、Encoder, Summarization Layer とともに Bi-GRU を用いた。本実験では比較のために、Encoder, Summarization Layer の双方に Attention を追加したのもも実験対象として用意した。BERTSum は、Encoder として BERT と、BERT と同一の構造を持つが事前学習を行っていない Transformer Stack を用いたものを用意した。Transformer Stack を用いたものは、TransformerSum と呼ぶ。BERTSum, TransformerSum は、Summarization Layer として Bi-LSTM, Feed-forward Network で構成された Linear, そして Transformer を用いたものを用意した。SummaRuNNer は両方ともバッチサイズ 32 で 5 Epoch 学習したものを用いた。BERTSum, TransformerSum は共に、バッチサイズ 3,000 で 50,000 Epoch 学習したものを用い、最も ROUGE-1 が高いモデルを実験対象として選択した。なお、教師ありアルゴリズムはすべて CNN/DailyMail データセット [23] を用いて学習している。ただし、CNN/DailyMail データセットは、抽出型要約の学習に適したデータセットではない。そのため、抽出型要約に用いることが出来るように Cheng ら [24] によって前処理が行われたものを使用した。教師ありアルゴリズムの ROUGE スコアを表 4.1 に示す。

## 第4章 実験設定

	ROUGE-1	ROUGE-2	ROUGE-L
SummaRuNNer	22.7	9.0	14.0
SummaRuNNer (Attention)	22.7	9.0	14.0
TransformerSum (Linear)	41.3	18.6	37.6
TransformerSum (RNN)	41.1	18.3	37.4
TransformerSum (Transformer)	41.2	20.3	39.6
BERTSum (Linear)	43.3	20.3	39.6
BERTSum (RNN)	43.2	20.2	39.6
BERTSum (Transformer)	43.0	20.1	39.4

表 4.1 実験対象の教師ありアプローチの ROUGE スコア. CNN/DailyMail のうち、学習に用いていないものを 10,000 件用いて求めた. 括弧の中に記載されているものは、Summarization Layer で用いられているモデル構造である.

## 4.2 実験に使用するデータセット

本実験では、実験対象のデータセットとして BASIL (Bias Annotation Spans on the Informational Level) データセット [4] を用いる. このデータセットは、Fox News, Huffington Post, New York Times から収集したニュース記事で構成され、文ごとに語彙バイアスと情報バイアスのアノテーションが付加されている. 実際のデータ例は、1.2 節にて示した. なお、BERT を入力として用いるアルゴリズムは、文の入力長が 512 トークンに制限されている [10]. そのため、実験の公平性を満たすために、実験対象の全てのアルゴリズムにおいて、文の入力長が 512 トークン以内に収まるようにデータセットの加工を行った.



# 第5章 抽出型要約アルゴリズムの出力における語彙バイアスの分析

この章では、各抽出型要約アルゴリズムの出力における語彙バイアスを、3.1節において示した MSBS を用いて分析する。続いて、このような MSBS の差異がどのように発生しているのかを探るための分析を行う。

## 5.1 アルゴリズムごとの MSBS の比較

各アルゴリズムごとの MSBS の比較を図 5.1 に示す。

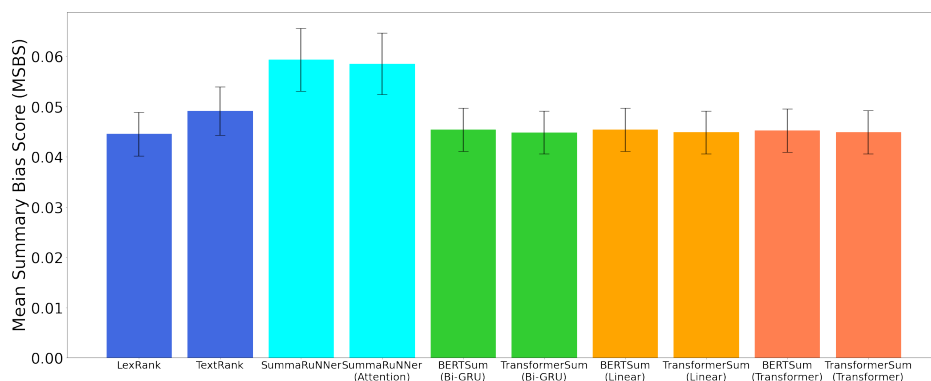


図 5.1 各アルゴリズムごとの MSBS の比較。横軸について、SummaRuNNer (Attention) は RNN に加えて Attention を用いたものである。BERTSum/TransformerSum に関して、括弧の中は Summarization Layer において用いているネットワーク構造を示す。エラーバーは、t 分布における 95% 信頼区間を示している。

図 5.1 より、教師なしアルゴリズムに関して以下のようなことが分かった。

## 第5章 抽出型要約アルゴリズムの出力における語彙バイアスの分析

- TextRank の MSBS が LexRank の MSBS と比べて高い。
- 教師ありアルゴリズムに関しては、Sentence Encoder として BERT や Transformer Stack を用いている BERTSum/TransformerSum は、RNN を用いている SummaRuNNer と比べて MSBS が低い傾向がある。
- SummaRuNNer に関しては、Attention の有無で大きな MSBS の差は発生しない。
- BERTSum/TransformerSum に関しては、Summarization Layer としてどのような構造のニューラルネットを用いても MSBS に大きな差異はない。

### 5.2 語彙バイアスのある文を含んだ要約と語彙バイアスのある文を含まない要約の比較

図 5.1 より、SummaRuNNer と、Summarization Layer として Transformer を使った BERTSum では MSBS が大きく異なっていることが分かるが、実際の出力を比較してどのような違いが見られるのかを調べる。結果を図 5.2、図 5.3 に示す。

WASHINGTON -- With presumptive Democratic presidential nominee Hillary Clinton officially reaching the number of delegates required for the nomination, President Barack Obama congratulated her on Tuesday evening, and signaled that the campaign of rival Sen. Bernie Sanders (I-Vt.) might be coming to an end.

According to the White House, Obama spoke to both Clinton and Sanders by phone following Tuesday's primaries and caucuses and will meet with Sanders on Thursday at Sanders' request.

"The President congratulated Secretary Clinton for securing the delegates necessary to clinch the Democratic Nomination for President.

Her historic campaign inspired millions and is an extension of her lifelong fight for middle-class families and children," White House Press Secretary Josh Earnest said in a statement.

"The President thanked Senator Sanders for energizing millions of Americans with his commitment to issues like fighting economic inequality and special interests' influence on our politics."

Obama will meet with Sanders in Washington on Thursday, the same day the Vermont senator plans to hold a rally there.

The meeting was arranged "at Senator Sanders' request," Earnest said, and the two will discuss "how to build on the extraordinary work he has done to engage millions of Democratic voters, and to build on that enthusiasm in the weeks and months ahead."

Obama has yet to make an endorsement in the race, but White House officials have said that it could come as early as this week.

In recent months, Obama has privately suggested that Sanders drop out of the race to help Democrats unite around Clinton and prepare for the general election.

He has also shown an eagerness to assail presumptive GOP nominee Donald Trump, which could be an important asset in helping Clinton defeat Trump in the general election.

図 5.2 SummaRuNNer によって出力された、対象に対する観点をポジティブな方向に誘導する語彙バイアスのある文を含む出力の例

赤もしくは緑でラインが引かれている文が要約アルゴリズムによって選択されたものであり、特に赤いラインが引かれている文は要約アルゴリズムによって選択された語彙バイアスが存在する文を示している。このニュースは、“ヒラリー・クリントン氏が、大統領候補として選出されるために必要な代議員の指名数を獲得したことをオバマ大統領が祝福した”という内容のニュースである。図 5.2 から、SummaRuNNer は *Her historic campaign inspired millions and is an extension of her lifelong fight for*

## 5.2 語彙バイアスのある文を含んだ要約と語彙バイアスのある文を含まない要約の比較

WASHINGTON – With presumptive Democratic presidential nominee Hillary Clinton officially reaching the number of delegates required for the nomination, President Barack Obama congratulated her on Tuesday evening, and signaled that the campaign of rival Sen. Bernie Sanders (I-Vt.) might be coming to an end.

According to the White House, Obama spoke to both Clinton and Sanders by phone following Tuesday's primaries and caucuses and will meet with Sanders on Thursday at Sanders' request.

The President congratulated Secretary Clinton for securing the delegates necessary to clinch the Democratic Nomination for President.

Her historic campaign inspired millions and is an extension of her lifelong fight for middle-class families and children," White House Press Secretary Josh Earnest said in a statement.

"The President thanked Senator Sanders for energizing millions of Americans with his commitment to issues like fighting economic inequality and special interests' influence on our politics."

Obama will meet with Sanders in Washington on Thursday, the same day the Vermont senator plans to hold a rally there.

The meeting was arranged "at Senator Sanders' request," Earnest said, and the two will discuss "how to build on the extraordinary work he has done to engage millions of Democratic voters, and to build on that enthusiasm in the weeks and months ahead."

Obama has yet to make an endorsement in the race, but White House officials have said that it could come as early as this week.

In recent months, Obama has privately suggested that Sanders drop out of the race to help Democrats unite around Clinton and prepare for the general election.

He has also shown an eagerness to assail presumptive GOP nominee Donald Trump, which could be an important asset in helping Clinton defeat Trump in the general election.

図 5.3 Summarization Layer として Transformer が使用された BERTSum によって出力された、対象に対する観点をポジティブな方向に誘導する語彙バイアスのある文を含まない出力の例

*middle-class families and children, White House Press Secretary Josh Earnest said in a statement.* という文を抽出していることが分かる。この文は、BASIL データセットにおいて“historic”という単語が語彙バイアスの原因になっているとアノテーションされている。この文が要約に含まれていることで、要約を読んだユーザーは、ヒラリー・クリントン氏に大きく肯定的な感情を抱く可能性がある。また、対象に対する観点をネガティブな方向に誘導する語彙バイアスのある文を要約文として抽出した場合の比較を図 5.4, 図 5.5 に示す。

このニュースは、“トランプ大統領が、ハンターが象牙の輸入を緩和する方針を覆した”という内容のものである。図 5.4 によると、ネガティブな方向に誘導する語彙バイアスのある文を含む出力の例は、*Laura Ingraham, a Fox News host, said Thursday on Twitter that she did not understand how the decision would not “INCREASE the gruesome poaching of elephants,” while the comedian Ellen DeGeneres started an online campaign, using the hashtag #BeKindToElephants, to raise funds for elephant conservation efforts.* という文を抽出している。BASIL データセットにおいて、この文は、“gruesome”という単語が語彙バイアスの原因になっているとアノテーションされている。この文が要約に含まれていることで、要約を読んだユーザーは、象の狩猟は絶対的に陰惨であると解釈し、象の狩猟に対して極端に否定的な印象を抱く可能性がある。

## 第 5 章 抽出型要約アルゴリズムの出力における語彙バイアスの分析

WASHINGTON — President Trump on Friday reversed the government's decision to start allowing hunters to import trophies of elephants that were killed in two African countries, pending a further review.

His evening Twitter message reversed a decision by his own administration over Zimbabwe that was announced this week and promoted as recently as Friday afternoon by the White House press secretary, Sarah Huckabee Sanders.

"Put big game trophy decision on hold until such time as I review all conservation facts," Mr. Trump tweeted.

"Under study for years. Will update soon with Secretary Zinke. Thank you!"

Ryan Zinke, the interior secretary, said in a statement later Friday night that the decision applied to two African countries, though it did not name them.

"President Trump and I have talked, and both believe that conservation and healthy herds are critical," the statement said.

"As a result, in a manner compliant with all applicable laws, rules and regulations, the issuing of permits is being put on hold as the decision is being removed."

The United States Fish and Wildlife Service had signaled its intention to end the 2014 ban, citing Zimbabwe's conservation efforts.

On Thursday, the service announced on its website that it would begin issuing permits to allow the import of elephants hunted from 2016 to 2018, with two trophies allowed per import.

Nearly two weeks ago, the department also waived the ban in a similar manner for Zambia.

Zambia and Zimbabwe have had mixed success in recent years maintaining or increasing their elephant population, according to the Great Elephant Census, a project financed by Paul G. Allen, a founder of Microsoft.

The project also found that the African elephant population shrunk nearly 30 percent from 2007 to 2014.

It is unclear why Mr. Trump reversed the decision, but the announcement faced backlash from both conservatives and liberals.

Laura Ingraham, a Fox News host, said Thursday on Twitter that she did not understand how the decision would not "INCREASE the gruesome poaching of elephants," while the comedian Ellen DeGeneres started an online campaign, using the hashtag #BeKindToElephants, to raise funds for elephant conservation efforts.

図 5.4 SummaRuNNer によって出力された、対象に対する観点をネガティブな方向に誘導する語彙バイアスのある文を含む出力の例

### 5.3 教師なし抽出型要約の結果の考察

図 5.1 から、TextRank と LexRank は MSBS の観点から結果が異なっているということが分かった。TextRank と LexRank はどちらも PageRank に着想を得たアルゴリズムであり、文同士の文の関連度をエッジの重みとした類似度グラフを構築し、隣接行列の固有ベクトルを求めることで各文の重要度スコアを求めるというものになっている。これらのアルゴリズムは、文同士の関連度を計算するために用いる特徴量が異なっている。TextRank は文同士の単語の重複度を用いており、LexRank は TF-IDF に基づいた関連度の計算を行っている。そのため、MSBS の違いは文同士の関連度を計算するための特徴量に起因していると考えられる。LexRank の MSBS が低いことから、LexRank の計算において用いられる TF-IDF が、MSBS を下げる上で有効な特徴量であるという可能性がある。

#### 5.3.1 教師なし抽出型要約における、TF-IDF 特徴量が及ぼす MSBS の変化に対する分析

本節では、教師なし抽出型要約において、TF-IDF 特徴量を考慮することが、アルゴリズム自体の MSBS を下げる上で有効であるという可能性を検証する。本実験では、

### 5.3 教師なし抽出型要約の結果の考察

WASHINGTON — President Trump on Friday reversed the government’s decision to start allowing hunters to import trophies of elephants that were killed in two African countries, pending a further review.

His evening Twitter message reversed a decision by his own administration over Zimbabwe that was announced this week and promoted as recently as Friday afternoon by the White House press secretary, Sarah Huckabee Sanders.

“Put big game trophy decision on hold until such time as I review all conservation facts,” Mr. Trump tweeted.

“Under study for years. Will update soon with Secretary Zinke. Thank you!”

Ryan Zinke, the interior secretary, said in a statement later Friday night that the decision applied to two African countries, though it did not name them.

“President Trump and I have talked, and both believe that conservation and healthy herds are critical,” the statement said.

“As a result, in a manner compliant with all applicable laws, rules and regulations, the issuing of permits is being put on hold as the decision is being removed.”

The United States Fish and Wildlife Service had signaled its intention to end the 2014 ban, citing Zimbabwe’s conservation efforts

On Thursday, the service announced on its website that it would begin issuing permits to allow the import of elephants hunted from 2016 to 2018, with two trophies allowed per import

Nearly two weeks ago, the department also waived the ban in a similar manner for Zambia

Zambia and Zimbabwe have had mixed success in recent years maintaining or increasing their elephant population, according to the Great Elephant Census, a project financed by Paul G. Allen, a founder of Microsoft.

The project also found that the African elephant population shrunk nearly 30 percent from 2007 to 2014.

It is unclear why Mr. Trump reversed the decision, but the announcement faced backlash from both conservatives and liberals.

Laura Ingraham, a Fox News host, said Thursday on Twitter that she did not understand how the decision would not “INCREASE the gruesome poaching of elephants,” while the comedian Ellen DeGeneres started an online campaign, using the hashtag #BeKindToElephants, to raise funds for elephant conservation efforts.

図 5.5 Summarization Layer として Transformer が使用された BERTSum によって出力された、対象に対する観点をネガティブな方向に誘導する語彙バイアスのある文を含まない出力の例

通常の TextRank における重要度の計算の際、TF-IDF 特徴量を考慮するようにすると MSBS がどのように変化するかを調べる。このとき、文  $s_i, s_j$  の関連度  $R(s_i, s_j)$  について、以下のような定式化を行う。

$$R(s_i, s_j) = \alpha O(s_i, s_j) + (1 - \alpha)T(s_i, s_j) \quad (5.1)$$

なお、 $\alpha = 0.0$  の時は通常の LexRank に対応し、 $\alpha = 1.0$  の時は通常の TextRank に対応する。このとき、 $O(s_i, s_j)$  は文  $s_i, s_j$  の単語の重複度で、各文を単語の集合  $w_k (\in s_i)$  とみなすとき、以下の式で表される [8]。

$$O(s_i, s_j) = \frac{|w_k | w_k \in s_i \ \& \ w_k \in s_j|}{\log(|s_i|) + \log(|s_j|)} \quad (5.2)$$

$T(s_i, s_j)$  は文  $s_i, s_j$  の TF-IDF のコサイン類似度である。実験では  $0.0 \leq \alpha \leq 1.0$  の条件で、0.1 刻みで  $\alpha$  の値を変動させ、各値ごとの比較を行う。結果は以下のようになった。

## 第5章 抽出型要約アルゴリズムの出力における語彙バイアスの分析

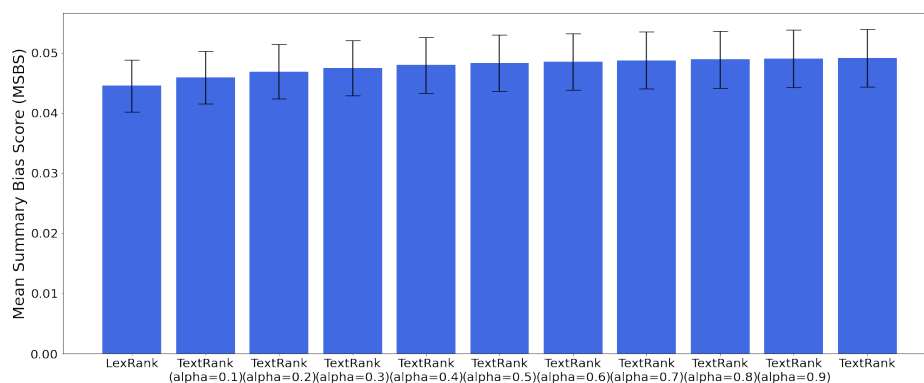


図 5.6 TF-IDF 特徴量の利用における MSBS の比較. エラーバーは, t 分布における 95% 信頼区間を示している.

図 5.6 によると,  $\alpha$  の値が小さくなるごとにわずかに MSBS が小さくなっていくことが分かった. このことから, 教師なし抽出型要約において, TF-IDF 特徴量を考慮することは, MSBS を下げる上で有効である可能性が示される.

### 5.4 教師あり抽出型要約の結果の考察

教師あり抽出型要約において重要とされているいくつかの特徴量に関して, どの特徴量が MSBS を上げることに寄与しているのかを調べる. また, 要約アルゴリズムの評価指標として広く用いられている ROUGE と MSBS の関係性について調べる.

#### 5.4.1 抽出型要約における語彙バイアスの観点からの特徴量分析

Pattanaik ら [25] によると, 教師あり抽出型要約においては, 特に TF-IDF, キーワード数, 文の位置, 文の長さといった特徴量が有効であると言われる. 本節では, これらの特徴量を考慮することが語彙バイアスの増幅に繋がっているという可能性を検証する. 具体的には, 3.2 節にて示した SFPI の計算を, それぞれの特徴量に対して行う. 4.1 節において示した教師あり要約アルゴリズムは, 全てエンドツーエンドで抽出を行う性質があるため, これらの特徴量を含ませることができない. そのため SFPI を計算することが出来ない. そこで, 本節では, これらの特徴量を含ませることが出来るアルゴリズムとして, 便宜的に Random Forest [26] を用いる. 4.1 節にて示したものと同様の方法で加工を行った CNN/DailyMail データセットを表形式に変換したものをを用いて Random Forest を学習させ, 文を要約として含ませるか否かの二値分類問題を解くように学習を行う. その後, 3.2 節で述べた SFPI を計算し, これらの

Shuffled Feature	SFPI
Sentence position	$1.69 \times 10^{-3}$
Number of keywords	$1.26 \times 10^{-3}$
Sentence length	$3.65 \times 10^{-3}$

表 5.1 Random Forest を用いた教師あり抽出型要約において、それぞれの素性ベクトルにおける特徴量の SFPI

特徴量の MSBS への寄与率を定量化する。その結果を用いて、MSBS を上げることに寄与している可能性がある特徴量を探る。SFPI の計算における試行回数  $K$  として 20 を設定した。SFPI の算出におけるデータセットとして、表敬式に変換した BASIL データセットを使用した。また、キーワードとして NLTK<sup>\*1</sup>を用いて抽出した Proper Nouns を使用した。TF-IDF は、文のベクトル化に不可欠であるため、SFPI は計算しない。結果を図 5.1 に示す。

図 5.1 より、位置、キーワード数、文の長さの SFPI を、TF-IDF をもとにして抽出した素性ベクトルに対して計算した結果、どの特徴量も正の SFPI を記録することが分かった。この結果は、抽出型要約において重要とされる特徴量を用いて要約を行うと MSBS が上がる可能性があるということを示唆している。

## 5.4.2 BASIL データセットにおける、抽出型要約において重要な特徴量と語彙バイアスの関係の分析

5.4.1 節で、抽出型要約において重要な特徴量である文の位置、重要キーワード数、文の長さの SFPI が正であり、要約の語彙バイアスを増幅する可能性について示された。本節では、BASIL データセットにおけるこれらの特徴量と語彙バイアスの関係を分析する。まず、BASIL データセットにおける位置と語彙バイアスのある文の割合を図 5.7 に示す。

図 5.7 によると、文書の 1 文目、2 文目に大きく語彙バイアスのある文が偏っていることが分かる。また、CNN/DailyMail データセットにおける参照要約に含まれる文の位置の割合を図 5.8 に示す。

図 5.8 によると、文の前半に位置する文が参照要約に含まれている傾向が見られる。このことから、CNN/DailyMail データセットを用いて学習することで、文文書の前半の文を要約文として抽出する傾向が学習される可能性があり、結果として文書の前半に含まれている傾向がある語彙バイアスのある文を抽出しやすくすることに繋がって

\*1 <https://www.nltk.org/>

## 第5章 抽出型要約アルゴリズムの出力における語彙バイアスの分析

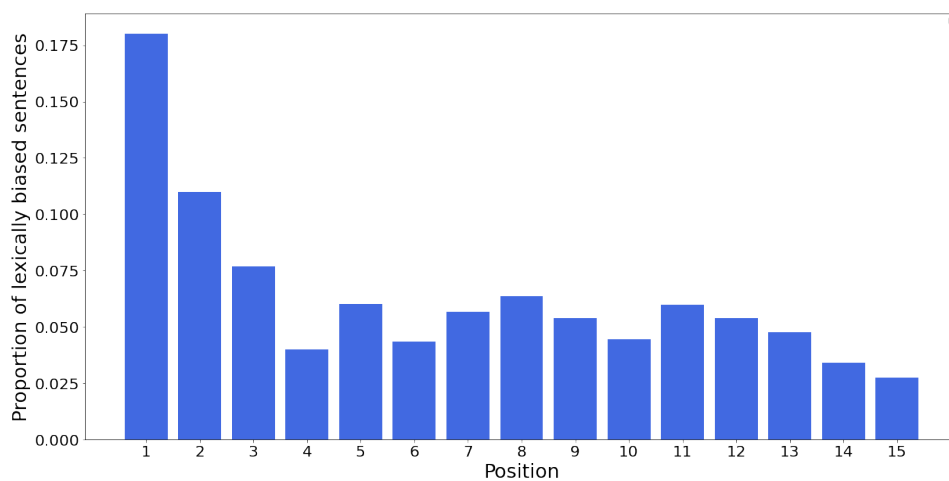


図 5.7 BASIL データセットにおける語彙バイアスのある文の位置の割合

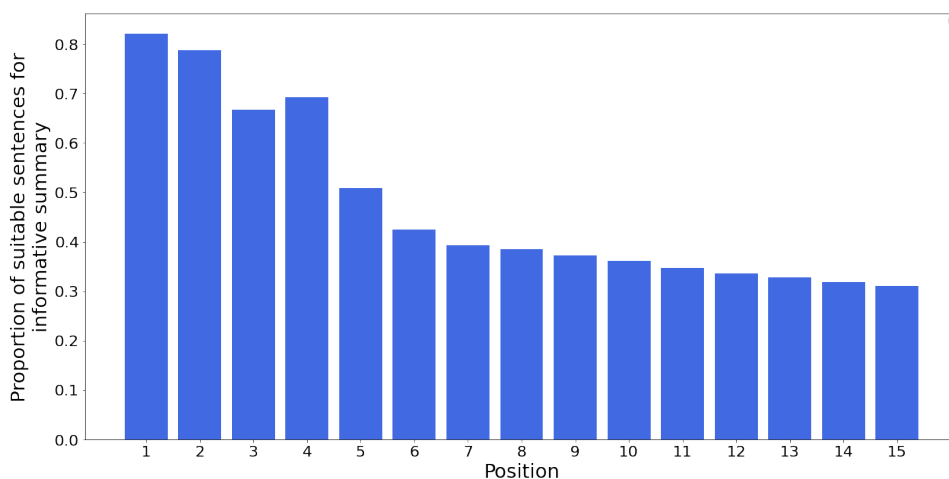


図 5.8 CNN/DailyMail データセットにおける、参照要約に含まれている文の位置の割合

いる可能性が考えられる。

また、要約として抽出された文の重要キーワード数は約 3.1 単語で、抽出されなかった文の重要キーワード数は約 1.8 単語であった。これより、重要キーワード数が多いほど要約文として抽出されやすい傾向があることが考えられる。また、語彙バイアスのない文の平均重要キーワード数は約 2.1 単語であったのに対し、語彙バイアスのある文の平均重要キーワード数は約 2.9 単語であった。そのため、重要キーワード数が多い文を抽出することが、語彙バイアスのある文を抽出しやすくすることに繋がっている可能性が考えられる。



同様に、要約として抽出された文の長さの平均は約 38.0 単語で、抽出されなかった文の長さの平均は約 24.0 単語であった。そのため、文の長さが長いほど要約文として抽出されやすい傾向があることが考えられる。また、語彙バイアスのない文の長さは平均約 27.9 単語であったのに対し、語彙バイアスのある文の長さは平均約 34.4 単語であった。このことから、文の長さが長い文を抽出することが、語彙バイアスのある文を抽出しやすくすることに繋がっている可能性が考えられる。

### 5.4.3 深層学習を用いた教師あり抽出型要約アルゴリズムの結果の考察

図 5.1 より、SummaRuNNer は教師なしアルゴリズムよりも高い傾向があり、BERT-Sum/TransformerSum は LexRank と同程度の MSBS であることが分かっている。5.4.1 節より、文の位置に関する特徴量の SFPI が正の値であるという結果が得られており、この特徴量を用いることで MSBS が上がる可能性がある。LexRank/TextRank はこの特徴量を考慮しないのに対して、SummaRuNNer は考慮している。これが SummaRuNNer の MSBS が比較的高くなっている一因として考えられる。BERT-Sum/TransformerSum も位置に関する特徴量を用いているが、その他の要因で MSBS が低くなっていると考えられる。

### 5.4.4 深層学習を用いた抽出型要約における、文の位置に関する特徴量の分析

5.4.1 節にて、要約において重要とされる特徴量の SFPI が正であるため、これらの特徴量を考慮することで MSBS が増加するという可能性を示した。深層学習を用いた抽出型要約においても、暗黙的にこれらの特徴量を考慮しており、それが語彙バイアスの増幅に繋がっていると考えられる。そこで本節では、深層学習を用いた抽出型要約において、文の位置に関する特徴量の摂動を行い、MSBS がどれほど変化するのかを検証する。本実験は文の位置に関する特徴量のみを対象としているが、これは他の特徴量に比べて摂動操作を行うのが容易なためである。SummaRuNNer/TransformerSum/BERTSum を対象とし、入力となる文の並び順をシャッフルすることで位置に関する特徴量の摂動を行う。なお、シャッフルした場合の MSBS は、5 回の試行を行った場合のそれぞれの MSBS の算術平均をとっている。結果を図 5.9 に示す。

図 5.9 より、対象となったアルゴリズムの全てのケースにおいて、入力文の並び順をシャッフルした結果、MSBS が低減していることが分かる。これにより、深層学習

## 第5章 抽出型要約アルゴリズムの出力における語彙バイアスの分析

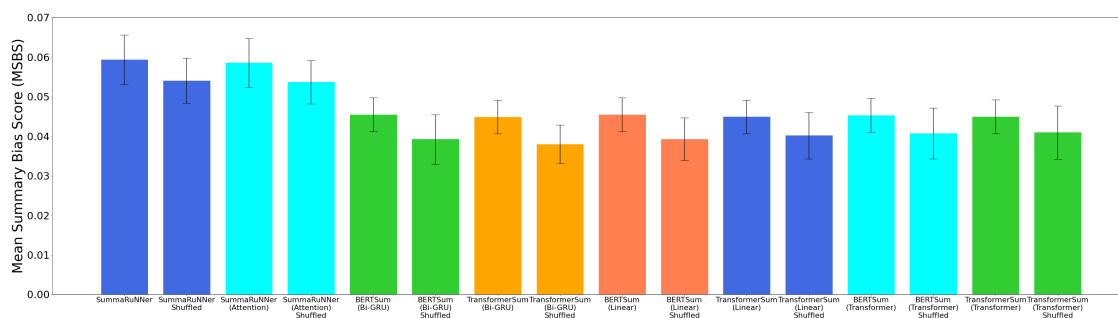


図 5.9 深層学習を用いた抽出型要約において、文を通常の並び順で与えた場合と、文の並び順をシャッフルした場合の MSBS の結果の比較. 横軸のラベルに Normal と書かれているものが通常の並び順で与えた場合を示す. Shuffled と書かれているものが文の並び順をシャッフルしたものの結果を示す. エラーバーは、t 分布における 95% 信頼区間を示している.

を用いたモデルは要約を抽出する際に、文の位置に関する特徴量を抽出しており、これらの特徴を用いないようにすることで MSBS を低減させることができることがわかる.

### 5.4.5 MSBS と ROUGE の関係の分析

MSBS が評価指標としてどのように機能するのかを ROUGE と比較することで調べる. 本実験では、BERTSum, TransformerSum を対象として、学習が進むにつれてどのように MSBS と ROUGE が変化するのかを調べた. ROUGE-1, ROUGE-2, ROUGE-L に対する分析を行ったが、すべて同じような傾向を見せたため、ROUGE-1 と MSBS の関係のみを提示する. 結果を図 5.10 に示す.

図 5.10 によると、全てのケースにおいて、学習が進むに連れて MSBS が上昇傾向を見せることが分かる. MSBS は低ければ低いほど良い指標であるため、MSBS と ROUGE は逆相関の傾向があると考えられる. この結果から、ROUGE が増加するように学習することで、アルゴリズムはより情報量の多い要約を生成することができるようになるが、同時に語彙バイアスのある文を抽出しやすくなるという可能性が示唆される. ROUGE に基づく要約アルゴリズムの評価は、CNN/DailyMail データセットに含まれる参照要約を用いて行われている. 5.4.2 節にて示したように、参照要約として抽出される文は、文書の先頭に位置する傾向が強い. そのため、文書の先頭に位置する文を抽出するという傾向を学習することで、ROUGE の改善が期待される. しかし、BASIL データセットは文書の前半に語彙バイアスを含む文を含んでいる傾向がある. これらの結果から、学習することで ROUGE が改善された結果、MSBS が上がっていると考えられる. また、MSBS を用いることで、ROUGE を補完するような

要約アルゴリズムの評価が出来る可能性がある。

5.1 節によると、TransformerSum と BERTSum に関しては、BERTSum の方が ROUGE が高いにもかかわらず、MSBS はほとんど変わらないという結果が得られている。また、表 4.1 より、BERTSum は SummaRuNNer と比べても ROUGE が大きく、MSBS が小さい。そのため、実験の対象としたアルゴリズムの中では、BERTSum が ROUGE、MSBS の両方の観点から最も優れた要約アルゴリズムであるということが分かった。Liu ら [11] によると、BERTSum は、要約文の抽出位置が前半に偏りすぎないように設計されている。これが BERTSum の MSBS が低い一因であると考えられる。

## 第5章 抽出型要約アルゴリズムの出力における語彙バイアスの分析

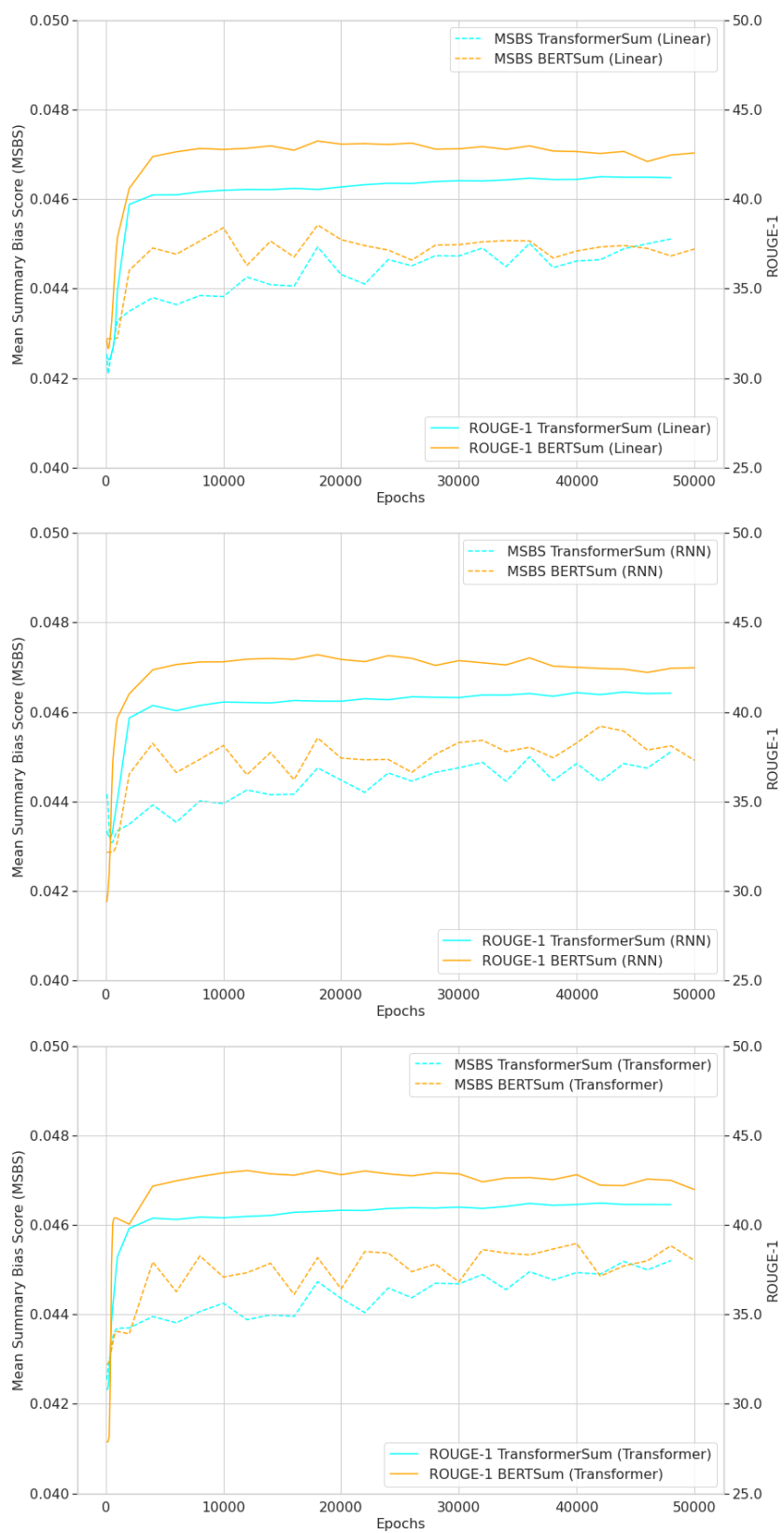


図 5.10 TransformerSum と BERTSum における MSBS と Epoch 数の関係

## 第6章 結論

本研究では、抽出型要約における語彙バイアスの増幅度合いを定量化するための評価指標を用い、既存の抽出型要約アルゴリズムに対して、生成された要約における語彙バイアスの増幅度合いの分析を行った。その結果、以下の事が分かった。

1. 抽出型要約アルゴリズムの性能は、MSBS の観点から異なっており、アルゴリズムによって語彙バイアスの増幅度合いは異なってくる事が分かった。
2. 教師ありの抽出型要約において、意味のある要約を生成する際に重要とされるいくつかの特徴量は MSBS を増大させる可能性があることを示した。また、深層学習を用いた抽出型要約において、意味のある要約を生成するために必要とされている特徴量の一つである文の位置を無視して要約を抽出することで、MSBS が低減されることが分かった。
3. ROUGE が上がるにつれて MSBS も上がるという結果が示された。つまり、MSBS と ROUGE は逆相関の傾向がある。そのため、MSBS を用いることで、ROUGE を補完するような要約アルゴリズムの評価が出来る可能性がある。また、BERTSum は ROUGE, MSBS の観点から最も優れた要約アルゴリズムであることが分かった。

今後の研究として、抽出型要約における語彙バイアスにおける対処を行いたい。本研究において、ROUGE スコアを向上させるように要約アルゴリズムの性能を改善していくことは、結果として語彙バイアスのある文を抽出しやすくする可能性について指摘した。この結果から、ROUGE の観点から要約アルゴリズムの性能改善を行いつつ語彙バイアスのある文を抽出ないようにするためには、語彙バイアスのもとなっている単語や表現を中立的なものに書き換える方法が有効であると考えられる。また、本研究は抽出型要約における語彙バイアスの分析にとどまっており、抽象型要約には言及していない。抽象型要約もまた広く使われる自動要約手法であり、分析が必要である。



# 謝辞

提案手法に関して様々な意見をくださったり、本論文の執筆にあたって熱心に指導していただいた指導教官の酒井哲也教授、および Yahoo! JAPAN 研究所の藤田澄男先生に、深い感謝を申し上げます。





## 参考文献

- [1] P.J. Boczkowski, E. Mitchelstein, and M. Matassi. “news comes across when i’ m in a moment of leisure” : Understanding the practices of incidental news consumption on social media. *New Media and Society*, Vol. 20, No. 10, pp. 3523–3539, 2018.
- [2] I. Pentina and M. Tarafdar. From "information" to "knowing": Exploring the role of social media in contemporary news consumption. *Computers in Human Behavior*, Vol. 35, pp. 211–223, 2014.
- [3] Chin-Yew Lin. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, pp. 74–81. Association for Computational Linguistics, July 2004.
- [4] L. Fan, M. White, E. Sharma, R. Su, P.K. Choubey, R. Huang, and L. Wang. In plain sight: Media bias through the lens of factual reporting. pp. 6343–6349, 2020.
- [5] W.-F. Chen, K. Al-Khatib, B. Stein, and H. Wachsmuth. Detecting media bias in news articles using gaussian bias distributions. pp. 4290–4300, 2020.
- [6] M. Gambhir and V. Gupta. Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review*, Vol. 47, No. 1, 2017.
- [7] G. Erkan and D.R. Radev. Lexrank: Graph-based lexical centrality as salience in text summarization. *Journal of Artificial Intelligence Research*, Vol. 22, pp. 457–479, 2004.
- [8] R. Mihalcea and P. Tarau. Textrank: Bringing order into texts. pp. 404–411, 2004.
- [9] R. Nallapati, F. Zhai, and B. Zhou. Summarunner: A recurrent neural network based sequence model for extractive summarization of documents. pp. 3075–3081, 2017.
- [10] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. Vol. 1, pp. 4171–4186, 2019.
- [11] Y. Liu and M. Lapata. Text summarization with pretrained encoders. pp. 3730–3740, 2020.
- [12] D.P. Baron. Persistent media bias. *Journal of Public Economics*, Vol. 90, No. 1-2,

## 参考文献

- pp. 1–36, 2006.
- [13] Wei-Fan Chen, Khalid Al-Khatib, Henning Wachsmuth, and Benno Stein. Analyzing political bias and unfairness in news articles at different levels of granularity, 2020.
- [14] M. Iyyer, P. Enns, J. Boyd-Graber, and P. Resnik. Political ideology detection using recursive neural networks. Vol. 1, pp. 1113–1122, 2014.
- [15] R. Baly, G. Karadzhov, D. Alexandrov, J. Glass, and P. Nakov. Predicting factuality of reporting and bias of news media sources. pp. 3528–3539, 2020.
- [16] H. Rashkin, E. Choi, J.Y. Jang, S. Volkova, and Y. Choi. Truth of varying shades: Analyzing language in fake news and political fact-checking. pp. 2931–2937, 2017.
- [17] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan. A survey on bias and fairness in machine learning. *ACM Computing Surveys*, Vol. 54, No. 6, 2021.
- [18] R. Baeza-Yates. Bias on the web. *Communications of the ACM*, Vol. 61, No. 6, pp. 54–61, 2018.
- [19] E.M. Bender, T. Gebru, A. McMillan-Major, and S. Shmitchell. On the dangers of stochastic parrots: Can language models be too big? pp. 610–623, 2021.
- [20] A. Dash, A. Shandilya, A. Biswas, K. Ghosh, S. Ghosh, and A. Chakraborty. Summarizing user-generated textual content: Motivation and methods for fairness in algorithmic summaries. *Proceedings of the ACM on Human-Computer Interaction*, Vol. 3, No. CSCW, 2019.
- [21] A. Shandilya, A. Dash, A. Chakraborty, K. Ghosh, and S. Ghosh. Fairness for whom? understanding the reader’s perception of fairness in text summarization. pp. 3692–3701, 2020.
- [22] A. Altmann, L. Toloşi, O. Sander, and T. Lengauer. Permutation importance: A corrected feature importance measure. *Bioinformatics*, Vol. 26, No. 10, pp. 1340–1347, 2010.
- [23] K.M. Hermann, T. Kočiský, E. Grefenstette, L. Espeholt, W. Kay, M. Suleyman, and P. Blunsom. Teaching machines to read and comprehend. Vol. 2015-January, pp. 1693–1701, 2015.
- [24] J. Cheng and M. Lapata. Neural summarization by extracting sentences and words. Vol. 1, pp. 484–494, 2016.
- [25] A. Pattanaik, S.S. Mishra, and M. Das. A comparative study of classifiers for extractive text summarization. *Advances in Intelligent Systems and Computing*, Vol. 1101, pp. 173–181, 2020.

- [26] T.K. Ho. Random decision forests. Vol. 1, pp. 278–282, 1995.