

令和三年度 卒業研究論文

題目

文書ベクトル演算に基づく考察型レポートの
自動評価とフィードバック生成に関する研究

指導教員 服部 峻

提出者 室蘭工業大学 情報電子工学系学科

氏名 細川 大和

学籍番号 18024157

提出年月日 令和4年2月10日

目次

第 1 章	まえがき	1
第 2 章	関連研究	3
2.1	「評価」	3
2.2	「フィードバック」	3
第 3 章	提案手法	5
3.1	従来手法との類似点と相違点	5
3.2	提案手法の概観	6
3.3	文書ベクトルの定義と演算	8
3.4	使用するベクトル化手法について	8
3.5	評価点の推定	10
3.6	フィードバックの生成	11
第 4 章	評価実験	12
4.1	データセット	12
4.2	評価点推定手法の精度評価	12
4.3	重回帰分析及び相関分析	12
4.4	フィードバック生成の有効性評価	13
4.5	評価点推定手法の実験結果	13
4.6	重回帰分析及び相関分析の実験結果	13
4.7	フィードバック生成の実験結果	26
第 5 章	追加実験	30
第 6 章	まとめと今後の課題	33
6.1	本稿の技術的貢献と社会的貢献	33
6.2	今後の課題	33
	参考文献	36

目次

2.1	関連研究「atama+」のシステム概観図	4
3.1	提案手法の概観	7
3.2	文書ベクトルの例	8
4.1	各手法における推定評価点の誤差（考察 2.3.1）	15
4.2	各手法における推定評価点の誤差（考察 2.3.2）	16
4.3	各手法における推定評価点の誤差（考察 2.3.3）	17
4.4	各手法における推定評価点の誤差（問題の区別なし）	18
4.5	重回帰分析における各手法の標準偏回帰係数（考察 2.3.1）	19
4.6	重回帰分析における各手法の標準偏回帰係数（考察 2.3.2）	20
4.7	重回帰分析における各手法の標準偏回帰係数（考察 2.3.3）	21
4.8	重回帰分析における各手法の標準偏回帰係数（問題の区別なし）	22
4.9	2 文書間の評価点の差と類似度の関係	25

表目次

3.1	従来手法に対する提案手法との比較	6
3.2	Word2Vec 及び Doc2Vec に使用したパラメータ設定	10
4.1	取得件数 10 件, 類似度の閾値 $t=0.5$ での実際の点数と推定評価点の不偏分散	14
4.2	取得件数 10 件, 類似度の閾値 $t=0.5$ での標準偏回帰係数	14
4.3	推定評価点 (解答例 1)	14
4.4	推定評価点 (解答例 2)	14
4.5	問題文と解答例 (考察 2.3.1)	26
4.6	解答例 1 に対するフィードバック (考察 2.3.1)	27
4.7	解答例 2 に対するフィードバック (考察 2.3.1)	27
4.8	問題文と解答例 (考察 2.3.2)	28
4.9	解答例 1 に対するフィードバック (考察 2.3.2)	28
4.10	解答例 2 に対するフィードバック (考察 2.3.2)	28
4.11	問題文と解答例 (考察 2.3.3)	29
4.12	解答例 1 に対するフィードバック (考察 2.3.3)	29
4.13	解答例 2 に対するフィードバック (考察 2.3.3)	29
5.1	gensim による解答例 1 に対するフィードバック (考察 2.3.1)	30
5.2	gensim による解答例 2 に対するフィードバック (考察 2.3.1)	31
5.3	gensim による解答例 1 に対するフィードバック (考察 2.3.2)	31
5.4	gensim による解答例 2 に対するフィードバック (考察 2.3.2)	31
5.5	gensim による解答例 1 に対するフィードバック (考察 2.3.3)	32
5.6	gensim による解答例 2 に対するフィードバック (考察 2.3.3)	32

第 1 章

まえがき

近年、AI を始めとした情報技術の発展の中で、人間が行っていた仕事を機械に置き換えることによって、時間的・経済的に効率化を図ることが様々な分野において行われている。この動きは教育の分野も例外ではなく、2019 年に文部科学省が提唱した GIGA スクール構想によって、小学校から高校までの学生に対しタブレットやノート PC などの情報端末が配布され、授業などに用いられている [1]。また、同年に告示された高等学校学習指導要領において「情報」が必修化され、小学校からプログラミング教育が行われるなど、教育と ICT の結びつきがより強くなっている [2]。さらに、大学の教育活動でも同様に ICT を用いた教育活動が行われており、例として、室蘭工業大学では moodle というオープンソースの e-ラーニングプラットフォームを用いることで、大学で開講されている講義について、出席や提出物の管理及び講義情報の提供などが行われている。

教育現場の中で、指導者である教員は、自身が開講する講義の中で問題演習や試験、レポートなど、様々な課題を受講者である学生に与えている。そして、学生は教員より与えられた課題に取り組み、教員へ提出し、教員は学生の提出した課題に対して評価の結果やフィードバックを与えている。

ここで教員が学生に与える「評価」や「フィードバック」の例として、「評価」では課題に対する評価点、「フィードバック」では、文章の添削指導、見落としているポイントを気づかせるためのヒントであるとする。このような「評価」や「フィードバック」は、学生に対して復習しやすい環境ときっかけを与え、講義で取り扱った題材について習熟度を高めることが期待出来る。しかし、1つの講義について、教員と学生の人数比は教員 1 人に対して学生が 40 人から 100 人という形式が多く、この場合、全ての学生に対して適切な評価とフィードバックを与えることは、時間と労力のコストが高いといった問題がある [3]。

本稿では、この問題を解決するため、「評価」や「フィードバック」の対象として大学の講義などで行われている、ある課題について問題点を考え解決策を考案するといった「考察型レポート」に着目する。自然言語処理技術の 1 つである文書ベクトル化を用いて「考察型レポート」をベクトル化し、差ベクトルやベクトル間の類似度の算出といったベクトル演算を用いることによって、評価とフィードバックの自動化を行い、学生に対して迅速かつ適切に評価とフィードバックを与えるシステムを提案する。また、評価とフィードバックの自動化に対して

有効な文書ベクトル演算の諸検討を行う。

第 2 章

関連研究

本章では、本稿の関連研究として、1 章で述べた「評価」と「フィードバック」について述べられている研究を挙げる。

2.1 「評価」

「評価」について、小林は国語の入試問題について、形態素解析された解答文を Doc2Vec 及び BERT を用いてベクトル化し、予め用意された模範解答のテキストとの類似度を算出し、人間が与えた評価点と、類似度との相関を求めている [4]。また、鈴木らは記述式の問題について Doc2Vec を用いてベクトル化及び模範解答との類似度を算出することで、人間が与えた評価点と類似度の相関及び有意差検定を行っている [5]。これらの研究では、Doc2Vec は多義語や言い換え表現の判定が困難であり、そのような問題に対して自動評価は難しいと述べられている。

また、安田らによって発明された特許技術では、参考文献 [6] の代表図面で示されるように、特定のベクトル化手法に依らないシステムが考案されている。ある問題に対する解答テキストについて、それぞれのベクトル化手法によって模範解答との類似度を算出し、過去データから求められた重回帰式に代入し評価を行うことで、それぞれのベクトル化手法の弱みを軽減することが期待出来る。

2.2 「フィードバック」

「フィードバック」について、atama+ と呼ばれるサービスでは、図 2.1 のように、ある単元を学習する学生に対して練習問題を与え、学生の解答が正解であった場合、反復練習を行うための類題や、次に学ぶべき項目に関わる問題を学生に対して与えている。また、学生の解答が不正解であった場合、与えた問題に関連する既習事項を解析し、学生が復習すべき既習事項を復習問題として与えている [7]。これを反復することによって、学習におけるつまづきを解消し、学習単元に至る基礎知識を十分に理解することが出来、現在学習している単元や、次に学ぶ単元に対してより効率良く学ぶことが期待出来るサービスとなっている。

このサービスは、小学校から高校までの学習範囲を取り扱う問題としており、各教科で扱われる単元項目について、関連性を内部のシステムが解析しており、学生が取り組んでいる問題とその解答によって、次に取り組むべき演習問題を推薦している。しかし、残念ながら内部処理に関わる詳しい解析技術や推薦技術は不明である。

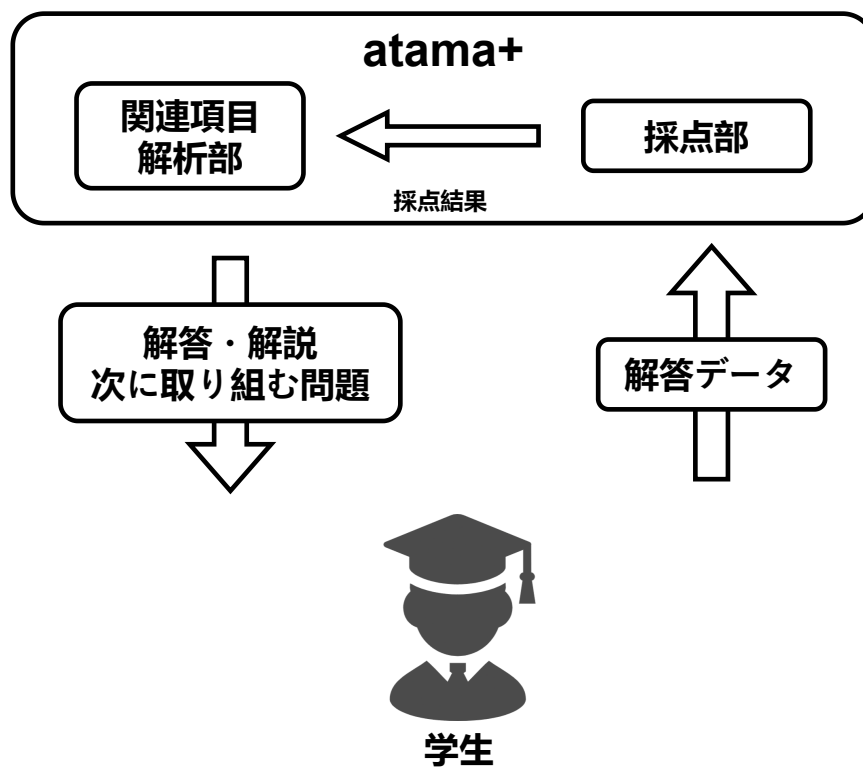


図 2.1 関連研究「atama+」のシステム概観図

第3章

提案手法

本章では、2章で述べた関連研究との類似点や相違点を挙げた上で、学生の考察型レポートに対して、自動評価とフィードバック生成を行うシステムについて提案する。

3.1 従来手法との類似点と相違点

関連研究と、本稿の提案手法の類似点及び相違点を以下で述べ、その結果を表3.1に示す。まず、「評価」について、参考文献[6]のシステムでは、提出された解答テキストに対してそれぞれのベクトル化手法での模範解答との類似度を計算し、その結果について重回帰分析を行うことで評価点の推定を行っている。後述する3.5節によって、様々なベクトル化手法を用いて、類似度を算出することや、重回帰分析を行うことで総合的な観点から判断を下すといった点で、本手法と類似している。しかし、このシステムは必ず模範解答を用意する必要があり、レポート課題のような、学生によって様々な書き方で書かれ、取り上げられるトピック数が多くなるほど、数多くの模範解答を作らなければならないと考えられる。本稿の提案手法では、模範解答または過去の採点済みレポートのどちらかが存在すれば良く、採点済みの解答データが多くなるほど多くのトピックに対して対応出来るようになるため、数多くの模範解答を作成する労力を減らすことが期待出来る。

次に、「フィードバック」について、参考文献[7]のシステムでは、提出された解答データに対してフィードバックを与えることは本手法と類似しているが、atama+では小学校から高等学校までの学習事項と対応している。一方、本稿で対象としているものは大学で課されるような「考察型レポート」であり、必ず1つの解答が存在する設問に加え、複数の答えが存在する場合や、決まり切った答えが存在しない設問についても採点対象としており、様々な問題と解答に対応が出来ることを期待出来る。

以上から、提案手法は複数のベクトル化手法を用い、模範解答だけでなく過去の採点済みデータを利用することで、答えが1つとは限らない様々な問題に対して対応し、教員の負担を減らしつつ、適切な評価とフィードバックが生成されると期待出来る。

表 3.1 従来手法に対する提案手法との比較

	参考文献 [6]	参考文献 [7]	提案手法
採点対象	様々	答えが明確な問題	考察型レポート
採点手法	複数のベクトル化手法	パターンマッチング	複数のベクトル化手法
用意するデータ	模範解答	模範解答	模範解答または過去の採点済みデータ
フィードバック	推定評価点	評価点, 解答, 解説	推定評価点, ヒント

3.2 提案手法の概観

本研究の提案システムについて、概観を図 3.1 に示す。初めに、学生が提出した未採点のレポートを入力としてシステムに与える。次に、システムは与えられたレポートの文書に対して形態素解析を行い単語に分割し、その結果を基に複数のベクトル化手法を用いてベクトル化を行う。このとき、ベクトル化手法によって事前学習が必要な場合が存在する。その場合、解答データベースや語彙データベースから分かち書きされたテキストを入力とし、学習する。さらに、ベクトル化手法によって任意の次元でベクトル化された学生レポートに対して、模範解答または過去に同様の問題に取り組み、採点済みとなったレポートに対して、提出された未採点のレポートと同様にベクトル化手法を適用する。以上の手順により、学生が提出した未採点のレポートと、保存されたデータから生成されたベクトルに基づいて、評価点の推定、及びフィードバックの生成を行う。参考文献 [6] と同様に、複数のベクトル化手法を用い、それらを総合的に扱うことで、それぞれのベクトル化手法の強みから、多角的で適切な自動評価とフィードバックの生成を行うことが期待出来る。

ここでのフィードバックは1章で述べたように、文章の添削指導や学生が見落としている点を気づかせるヒントといった情報が考えられる。システムが学生に対してフィードバックを与える場合の理想の1つとして、現在の教育現場で行われているように、関連項目を認識させつつ、学生が直すべき箇所やそれに関連するヒントなどを自然言語で出力することが考えられる。本稿では、以上のフィードバックの1つとして、学生によって提出されたレポートに対して、改善するために調べるべき重要単語を挙げ、ベクトル化手法とその演算によって求める手法について、後述する 3.6 節で述べる。

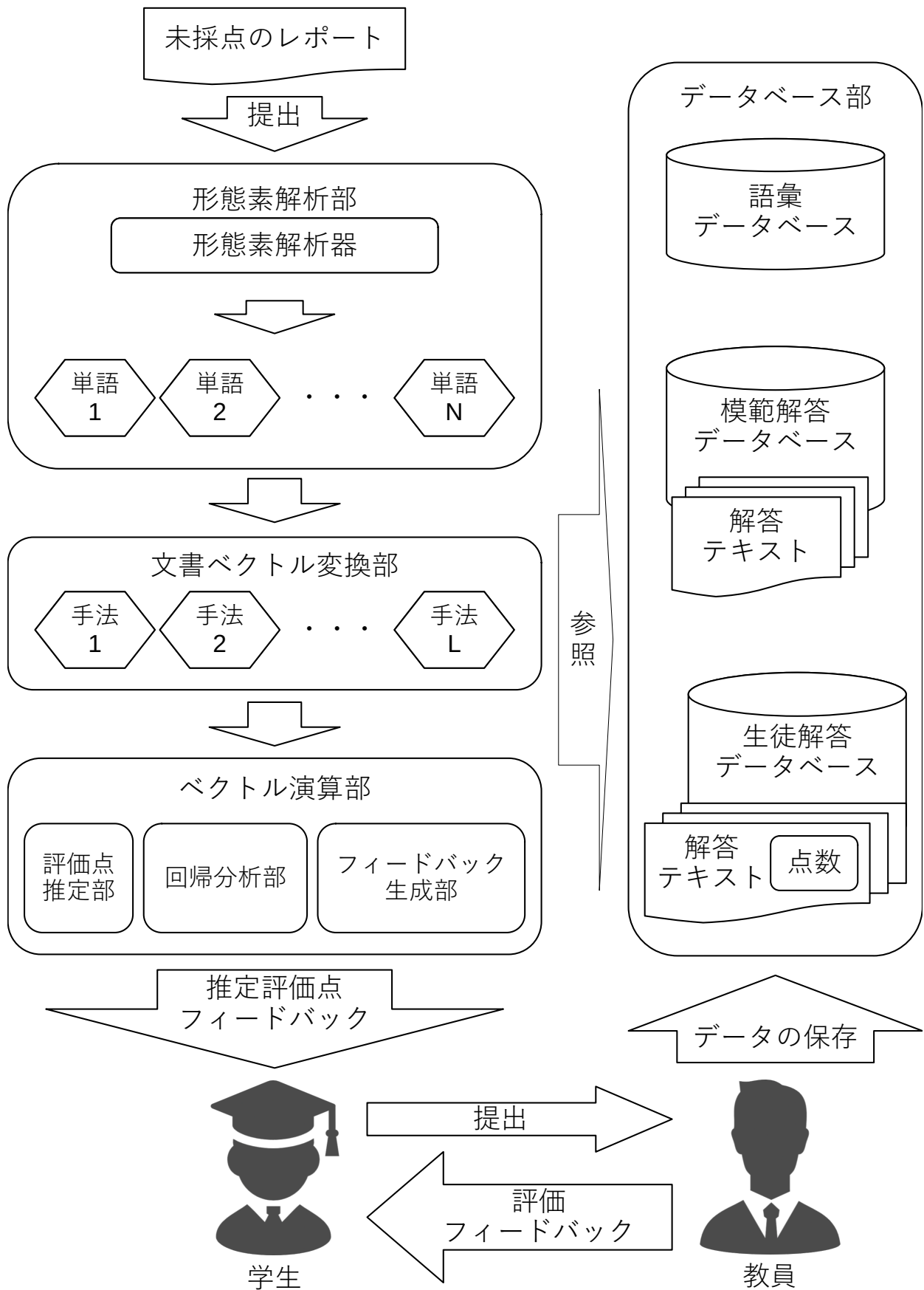


図 3.1 提案手法の概観

3.3 文書ベクトルの定義と演算

本研究で用いるベクトル化手法について、問題文、模範解答、学生のレポートの3つに対して得られた特徴ベクトルをそれぞれ \mathbf{Q} 、 \mathbf{A} 、 \mathbf{S} とし、図 3.2 に示す。

また、この3つのベクトルに対して差ベクトルを求めることで得られるベクトルは、問題文から模範解答へ至る考察、問題文から学生のレポートへ至る考察、学生の解答が模範解答へ近づく参考情報と仮定する。本稿では、この参考情報ベクトルが学生のレポートにある考察のオリジナリティを満たしつつ、学習目標を満たすことが出来ると考え、これを \mathbf{R} とおき、フィードバックの生成に用いることとする。

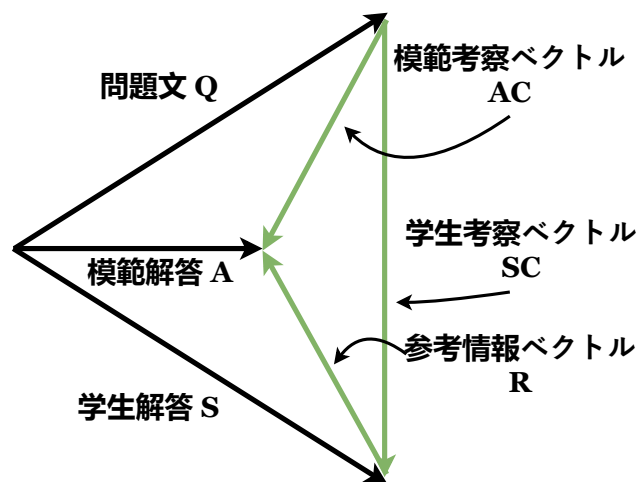


図 3.2 文書ベクトルの例

3.4 使用するベクトル化手法について

本稿では評価点の推定とフィードバックの生成について、Bag of Words (以下 BoW と表記する), Term Frequency (以下 TF と表記する), TF-IDF, Word2Vec, Doc2Vec といった5種類のベクトル化手法を使用する。なお、TF-IDF, Word2Vec, Doc2Vec の3つの手法は Python のオープンソースライブラリである gensim を用いて実装した。

BoW は、単語ごとに分割された文書に対して、使用された単語について 1 を与え、使用していない単語に対して 0 を与えるベクトル化手法である。このことから、2 文書間の類似度を算出したとき、同じ単語を多く使用している場合、類似度が高く算出されることが期待出来る。しかし、BoW の問題として、2 文書間の類似度を算出するとき、共通して使用される単語が多くなる場合、高い類似度が算出されやすいといった点が挙げられる。

それに対し、TF は、単語ごとに分割された文書に対して、使用された単語について、文書中での出現頻度を値とするベクトル化手法である。この問題に対し、TF は単語の出現頻度を値とすることで、文書内における単語の重要性を考慮したベクトル化手法と考えられる。その

ため、BoW よりも文書全体の意味を考慮した評価が期待出来る。

TF-IDF は、TF と同様に、使用されている単語について文書中での出現頻度を考慮したベクトル化手法である。これに加えて、自身以外の文書群において、自身の文書の中で使用されている単語が他の文書で何件出現しているのかといった、文書間での出現頻度である IDF (Inverse Document Frequency) との積によって表されるベクトル化手法である。TF-IDF はトピック分析や重要語検出に用いられるベクトル化手法であり、利点として、文書間での出現頻度を考慮するため、共通した話題や助詞などの文中で重要でない単語について、重みを下げることが期待出来る。本稿の実装では、参考文献 [8] で実装されている `tfidfmodel` を用いた。単語 i を含む文書 j における `tfidf` の値 $\text{weight}_{i,j}$ について、様々な定義が提案されているが、本稿では式 (3.1) で定義する。なお、 $\text{frequency}_{i,j}$ は単語 i が文書 j 中に出現した頻度を表し、 document_freq_i は単語 i が全文書 D 中に何件出現しているのかといった、文書間での単語の出現頻度を表している。本稿での全文書集合とは、図 3.1 で示される学生解答データベース、模範解答データベース、語彙データベースに含まれる文書数の合計である。

$$\text{weight}_{i,j} = \text{frequency}_{i,j} \cdot \log_2 \frac{|D|}{\text{document_freq}_i} \quad (3.1)$$

Word2Vec は、Mikolov らによって開発されたベクトル化手法で、単語に分割された文書群をテキストコーパスとして学習し、単語を任意の次元のベクトルへ埋め込むベクトル化手法である [9,10]。Word2Vec の特徴として、単語間の類似度を算出することが出来、単語ベクトルによるベクトル演算が出来ることが知られている。

一般に、Word2Vec は単語間でのベクトル演算を目的に使用されており、単語の列である文書間でのベクトル演算では用いられていない。本稿では、Word2Vec によって文書ベクトルを生成する場合、文書 j における次元 i の値 $\text{weight}_{i,j}$ について、式 (3.2) を用いて生成した。なお、 $w_{i,j,k}$ は、文書 j を単語ごとに分割したときの先頭から k 番目の単語から得られるベクトルの i 次元目の値を表している。このとき、学習手法として CBOW 法と Skip-gram 法のうち、学習精度が良くなりやすいと言われる Skip-gram 法を使用した。また、各単語を表すベクトルの次元数 M を 200 に設定した。

$$\text{weight}_{i,j} = \sum_k w_{i,j,k} \quad (3.2)$$

Doc2Vec は、Le らによって Word2Vec を基に開発されたベクトル化手法で、タグ付けされた文書を単語ごとに分割し、タグと文書を対にした文書群をテキストコーパスとして学習し、文書を任意の次元のベクトルへ埋め込むベクトル化手法である [13]。Word2Vec は単語間の類似度を算出することが出来るが、一方、Doc2Vec は文書間の類似度を算出することが出来、文書ベクトルによるベクトル演算が出来ることが知られている。本稿では学習手法として、PV-DBOW 法と PV-DM 法のうち、学習精度が良くなりやすいと言われる PV-DM 法を使用した。また、各文書を表すベクトルの次元数 M を 200 に設定した。Word2Vec と Doc2Vec について、gensim を使用して学習したときのパラメータを表 3.2 で示す。

表 3.2 Word2Vec 及び Doc2Vec に使用したパラメータ設定

	Word2Vec	Doc2Vec	備考
学習手法	Skip-gram	PV-DM	なし
vector_size	200		ベクトルの次元数 M
epoch	20		学習の反復回数
window	10		周辺単語をいくつ予測するか
sample size	10		高頻度に出現する単語について ランダムダウンサンプリングする閾値
min_count	10		学習する単語の出現頻度の下限
workers	8		学習の並列数

3.5 評価点の推定

評価点の推定を行うとき、まず、提出されたレポートに対してベクトル化手法を適用する。得られたベクトルに対して、模範解答や過去のデータを参照し類似度が高いと算出されたレポートは、実際の評価点も類似しており、差の絶対値が小さく、負の相関があると考えられる。以上の仮説から、評価点の推定を行う手法を提案する。

提出された未採点のレポートをベクトル化した文書ベクトル \mathbf{S} に対して、模範解答または過去に評価された採点済みレポートから類似度の高い文書のうち上位 n 件を抽出する。取り出された j 件目の文書 D_j について、実際に与えられた評価点と、提出された未採点のレポート \mathbf{S} とのコサイン類似度をそれぞれ $\text{point}(D_j)$, $\text{sim}(\mathbf{S}, \mathbf{D}_j)$ とする。以上の評価点と類似度を用いて、提出されたレポート \mathbf{S} の評価点を推定する関数を $\text{score}(\mathbf{S})$ とし、以下の式 (3.3) で表す。

$$\text{score}(\mathbf{S}) = \frac{\sum_{j=1}^n \text{point}(D_j) \cdot \text{sim}(\mathbf{S}, \mathbf{D}_j)}{\sum_{j=1}^n \text{sim}(\mathbf{S}, \mathbf{D}_j)} \quad (3.3)$$

また、それぞれのベクトル化手法において、類似度が高いとはどの程度を指すのか、抽出する件数が何件である方がそのベクトル化手法による点数推定として精度が良くなるのかといった最適化の問題が考えられる。このような調整をシステム上で行う場合、取得する件数 n を調整することや、類似度が設定された閾値以上であれば推定するためのデータとして取り入れるといった閾値 t などを調整することで可能であると考えられる。

さらに、それぞれのベクトル化手法について式 (3.3) を適用することで、1つのレポートに対して複数の手法による推定評価点を得ることが出来る。このことから、模範解答や過去に提出された採点済みレポートに対して評価点の推定を行い、実際に与えられた評価点を目的変数に、推定評価点を説明変数と置くことで、重回帰分析が可能となる。これによって、それぞれの手法による推定評価点に対して、標準偏回帰係数との積によって線形結合することが出来、3.2節で述べたような、それぞれのベクトル化手法の強みを利用した点数の推定が出来る時期

待出来る.

この場合の評価点は, 使用する L 個のベクトル化手法のうち ℓ 番目の手法によって推定された推定評価点 $\text{score}_\ell(\mathbf{S})$ と, 求められた偏回帰係数 $\hat{\alpha}_\ell$ の積の総和に切片 c を足した値で表すことが出来る. これを以下の式 (3.4) で表す.

$$\text{scoreEx}(\mathbf{S}) = c + \sum_{\ell=1}^L \hat{\alpha}_\ell \cdot \text{score}_\ell(\mathbf{S}) \quad (3.4)$$

3.6 フィードバックの生成

TF-IDF は重要語抽出や, トピック分析などに用いられている手法であるため, この手法は重要語抽出として適していると考えられる. また, Doc2Vec は文書や単語 (タグ) を入力に与えることで, 類似する文書を出力することが出来ることから, トピック分析などに用いることが出来ると考えられる. 少なくとも, TF-IDF と Doc2Vec の 2 つの手法について, 提出されたレポートに対して模範解答との差ベクトルを求めることで, 学生のレポートに対して足りない点として単語を出力出来るのではないかと考えた. 以上の仮説から, 学生の提出したレポートに対して足りない点を出力するような手法を提案する.

図 3.2 から, 学生のレポートベクトル \mathbf{S} と模範解答ベクトル \mathbf{A} との差ベクトル \mathbf{R} を利用して, 重要度の高い単語を上位 m 件抽出し学生に与える.

差ベクトル \mathbf{R} からフィードバックとして単語を出力する時, 任意の単語 w について与えられる重要度を $\text{importance}(w, \mathbf{R})$ とし, 式 (3.5) 及び式 (3.6) で表す. ただし, 式 (3.5) は BoW, TF, TF-IDF での重要語検出に用い, 式 (3.6) は Word2Vec, Doc2Vec で重要語検出に用いる. この時, $\mathbf{R}(w)$ はベクトル \mathbf{R} における単語 w に対応する値である, また, $\text{sim}(\mathbf{W}(w), \mathbf{R})$ は単語 w に対応する埋め込みベクトル $\mathbf{W}(w)$ と, ベクトル \mathbf{R} との類似度とする. また, 各ベクトル化手法から式 (3.5) 及び式 (3.6) によって求められた重要語について, 重要度の相加平均をレポート全体の重要語として抽出し, 学生に与える.

$$\text{importance}(w, \mathbf{R}) = \begin{cases} \text{sim}(\mathbf{S}, \mathbf{A}) \cdot \mathbf{R}(w) & (3.5) \\ \text{sim}(\mathbf{S}, \mathbf{A}) \cdot \text{sim}(\mathbf{W}(w), \mathbf{R}) & (3.6) \end{cases}$$

さらに, 3.5 節で求めた重回帰分析の結果を用いるといった手法や, 総和が 1 となる係数列による最適化, ユーザが調整した値による代入などによって, ベクトル化手法ごとに求めた単語の重要度を組み合わせることも考えられる. そのような係数を適用する場合, 式 (3.4) と同様に, 使用する L 個のベクトル化手法のうち ℓ 番目の手法によって抽出された, 単語 w の重要度を $\text{importance}_\ell(w, \mathbf{R})$ と, それぞれのベクトル化手法に設定された標準偏回帰係数 α_ℓ を掛けた値の総和によって表すことが出来る. これを式 (3.7) で表す.

$$\text{importanceEx}(w, \mathbf{R}) = \sum_{\ell=1}^L \alpha_\ell \cdot \text{importance}_\ell(w, \mathbf{R}) \quad (3.7)$$

第 4 章

評価実験

本章では、本稿の提案手法である評価点の推定及びフィードバック生成の有効性を検証するため、K-分割交差検証を行う。また、未採点のレポートと過去データによって求められた類似度に対して設定された閾値 t と、推定するために取得する件数 n を調整することによって、どのように各ベクトル化手法が評価点の推定と重要語抽出に影響を及ぼしているのか比較する。

4.1 データセット

データセットとして、室蘭工業大学で開講されている情報学応用演習 A という講義の中で、学生が提出したレポート及びその模範解答を使用する [14]。本稿では、レポート内にある設問から、3つの設問それぞれに対して、問題文 1 件、模範解答 1 件、学生の解答文 99 件、著者が作成した採点済みの解答文 2 件の合計 309 件を用いた。

また、学習などの事前処理が必要なベクトル化手法について、語彙データ及び文書データとなるテキストコーパスとして、Cirrus Search が提供している日本語版 Wikipedia のダンプデータを使用した。このとき、前処理として形態素解析に MeCab を使用し、辞書は mecab-ipadic-NEologd を使用した。

4.2 評価点推定手法の精度評価

式 (3.4) を用いて、提案手法によって推定された評価点と、教員によって実際に与えられた評価点との差の平均を求める。提出された未採点のレポートに対して評価点を推定するとき、推定対象とする文書ベクトルに対して、類似度の閾値 t と、取得する文書数の上限となる件数 n の 2 つの変数がどの様に影響するのか調べる。

4.3 重回帰分析及び相関分析

式 (3.4) 及び式 (3.7) において、実際に与えられた点数を目的変数とし、各ベクトル化手法の推定評価点を説明変数に重回帰分析を行い、算出された偏回帰係数と切片及び標準偏回帰係

数を用いることで、評価点の推定やフィードバック生成を行っている。類似度に対する閾値 t や、評価点算出に用いる文書の取得上限数などのパラメータによって、各ベクトル化手法が標準偏回帰係数にどのような影響を及ぼしているのか検討する。また、2文書間の類似度と、評価点の差を算出することで、3.5節で述べた仮定が正しいのか検討を行う。

4.4 フィードバック生成の有効性評価

式 (3.7) によって重要語を取り出し、各設問について著者が作成した解答データを基に、キーワードがどのように抽出されたのか検討を行う。

4.5 評価点推定手法の実験結果

式 (3.4) を用いて、それぞれのベクトル化手法を用いて、提案手法によって推定された評価点と、教員によって実際に与えられた評価点との差の平均について、設問ごとに分けた結果と全体での結果を図 4.1 から図 4.4 に示す。また、著者が作成した解答例についての推定結果を表 4.3 及び表 4.4 に示す。

式 (3.3) を用いることで評価点の推定を行った結果、それぞれの手法について、図 4.1 から図 4.4 について、取得件数 n に関わらず、閾値 t が高くなることで、推定評価点と実際に与えられた点数との差の平均が大きくなるのが分かる。閾値 t が高くなることで、式 (3.3) に取り入れられるレポートの数が減り、推定評価点との誤差が高まっていると考えられる。しかし、重回帰分析の結果を利用する式 (3.4) によって、取得件数 n や閾値 t に関わらず、実際に与えられる点数に近い推定評価点を求めることが可能であると分かった。

4.6 重回帰分析及び相関分析の実験結果

各ベクトル化手法が取得件数 n と類似度の閾値 t を調整することで、それぞれの手法について評価点の推定がどのようになるのか考察をするため、標準偏回帰係数がどのように変化したのかを表した結果を図 4.5 から図 4.8 に、また、評価点の推定手法の仮定が正しいのか考察するため、各設問及び問題の区別がない場合での、実際に与えられた評価点及び推定評価点の不偏分散を表 4.1 に、2文書間の評価点の差と類似度の関係を図 4.9 に示す。

各ベクトル化手法に対して、2文書間の類似度と点数の差について、図 4.9 から、それぞれの文書ベクトル化手法で点数差に関わらず類似度が正で求められやすいことが分かる。これは BoW, TF, TF-IDF の3つの手法では、定義から0以上の類似度が算出されており、Word2Vec は、式 (3.2) によって作られたベクトルは0.9以上の類似度が出やすく、出現する単語のベクトルの和では類似性と差の平均での相関が取れないことが分かる。また、Doc2Vec は同じ問題に対する解答テキストのみを比較しているため、負の値が出にくい結果となったと考えられる。

表 4.1 取得件数 10 件, 類似度の閾値 $t=0.5$ での実際の点数と推定評価点の不偏分散

問題 ID	Truth	BoW	TF	TF-IDF	Word2Vec	Doc2Vec
考察 2.3.1	204.521	486.649	208.442	595.291	14.131	733.385
考察 2.3.2	673.794	282.703	128.843	542.063	113.426	477.234
考察 2.3.3	907.566	158.190	1230.529	394.494	984.979	45.247
区別なし	910.217	812.752	800.238	938.192	660.749	951.055

表 4.2 取得件数 10 件, 類似度の閾値 $t=0.5$ での標準偏回帰係数

問題 ID	BoW	TF	TF-IDF	Word2Vec	Doc2Vec
考察 2.3.1	0.217	-0.005	0.112	0.184	0.143
考察 2.3.2	-0.009	0.088	0.070	0.341	-0.142
考察 2.3.3	-0.152	0.395	0.151	0.513	-0.036
区別なし	0.042	0.216	0.078	0.525	0.002

表 4.3 推定評価点 (解答例 1)

問題 ID	Truth	BoW	TF	TF-IDF	Word2Vec	Doc2Vec	scoreEx
考察 2.3.1	70.000	57.993	65.628	62.094	63.600	66.671	62.920
考察 2.3.2	63.000	36.500	41.350	29.510	41.700	0.000	44.547
考察 2.3.3	50.000	0.000	90.000	0.000	70.005	0.000	65.311

表 4.4 推定評価点 (解答例 2)

問題 ID	Truth	BoW	TF	TF-IDF	Word2Vec	Doc2Vec	scoreEx
考察 2.3.1	63.000	57.805	65.308	63.000	63.798	0.000	58.046
考察 2.3.2	75.000	10.000	37.964	0.000	42.200	63.000	31.682
考察 2.3.3	50.000	0.000	39.258	0.000	68.003	0.000	47.110

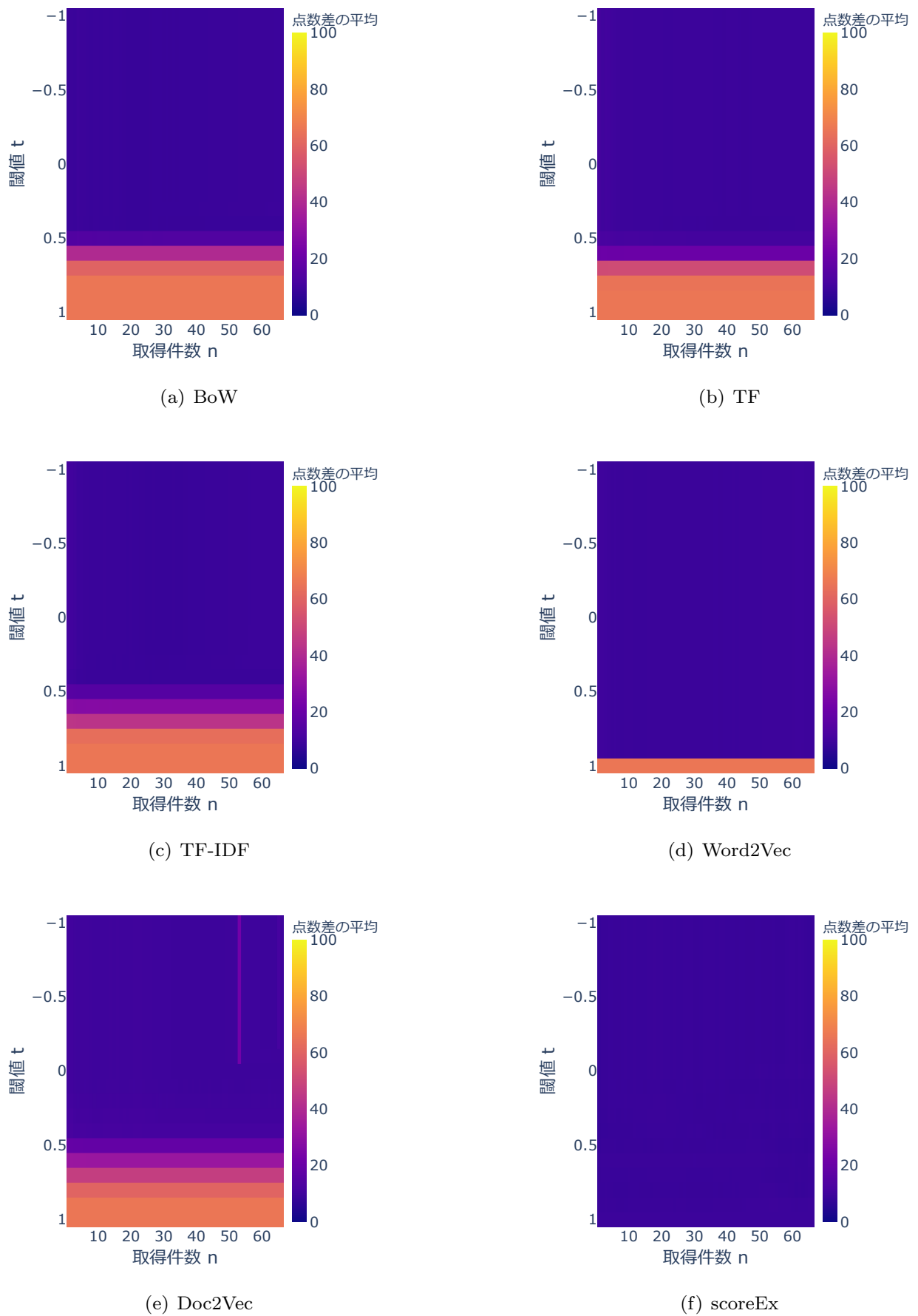
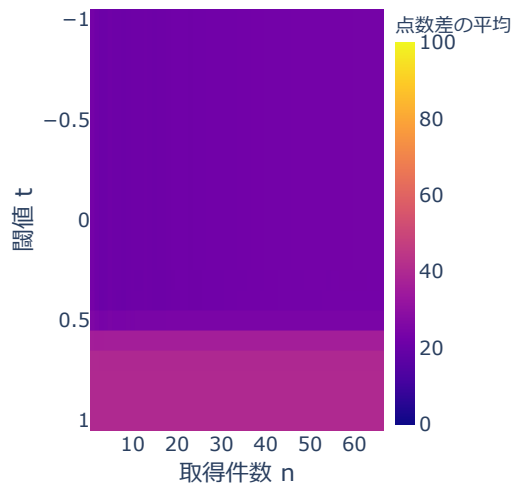
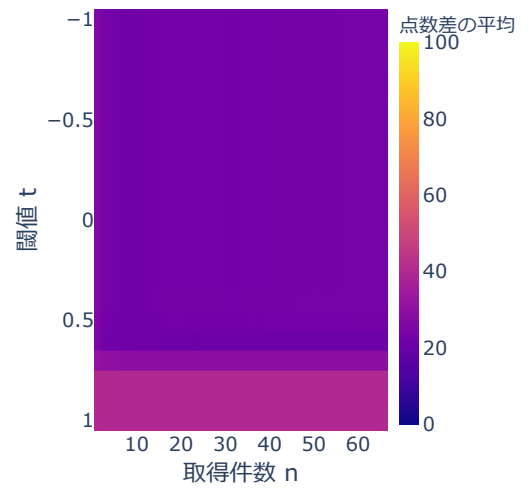


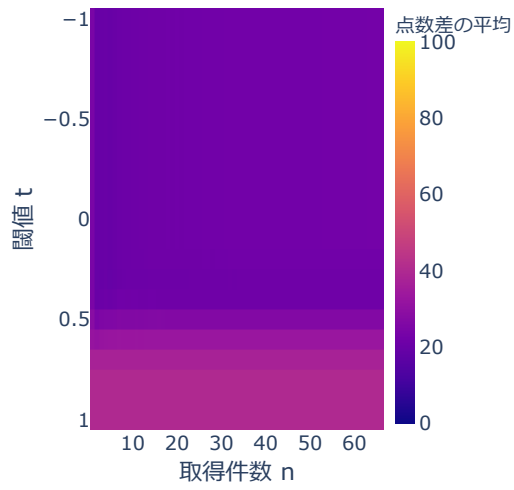
図 4.1 各手法における推定評価点の誤差 (考察 2.3.1)



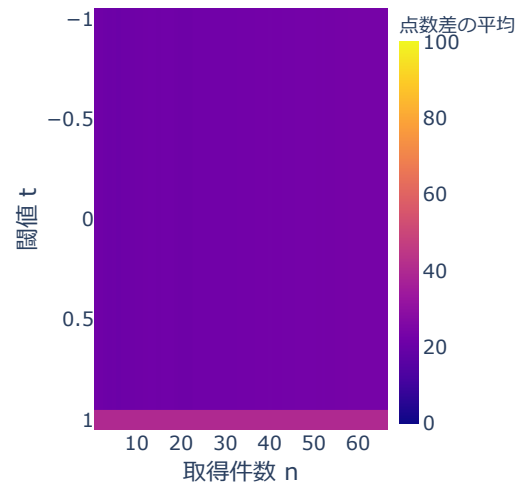
(a) BoW



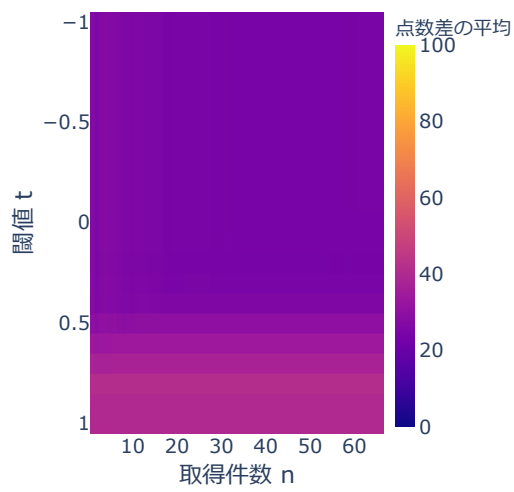
(b) TF



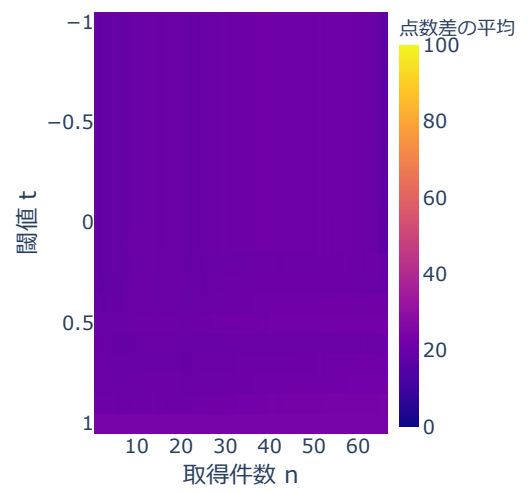
(c) TF-IDF



(d) Word2Vec

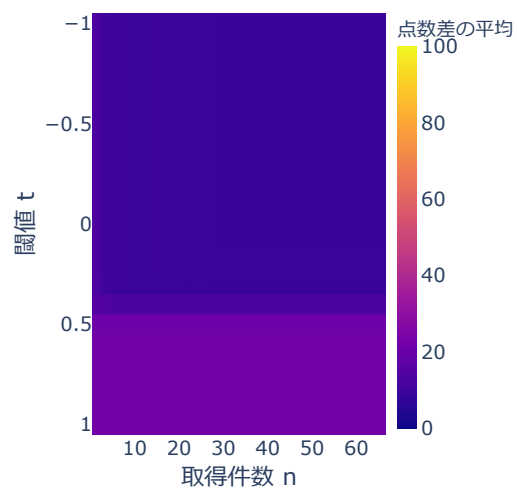


(e) Doc2Vec

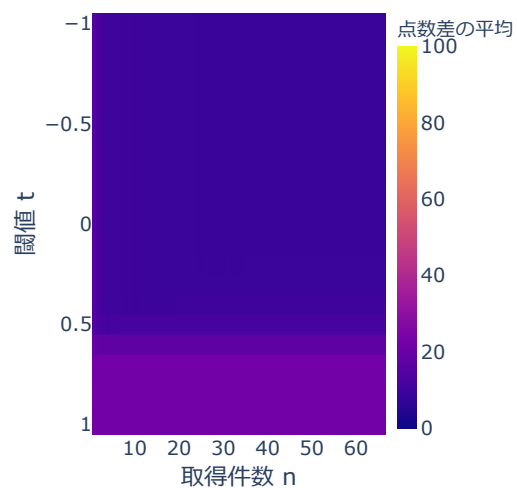


(f) scoreEx

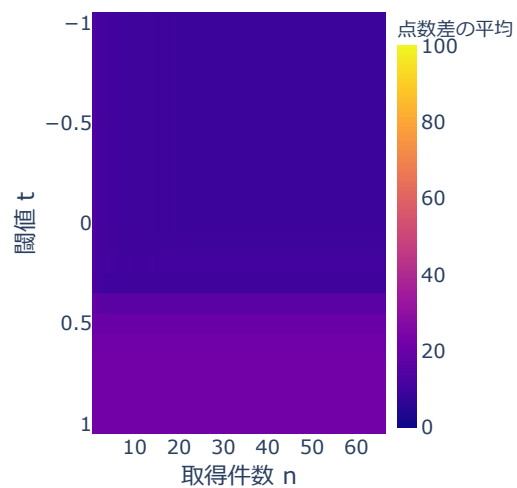
図 4.2 各手法における推定評価点の誤差 (考察 2.3.2)



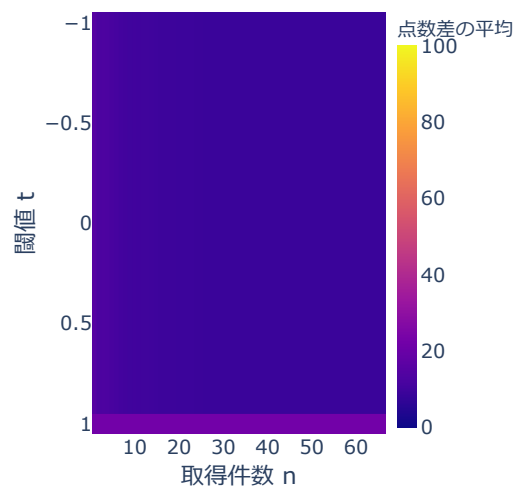
(a) BoW



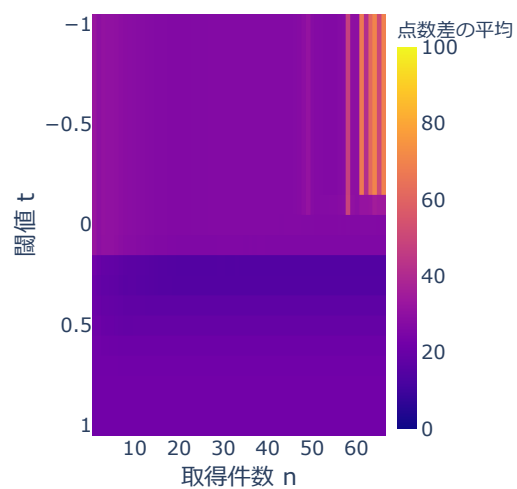
(b) TF



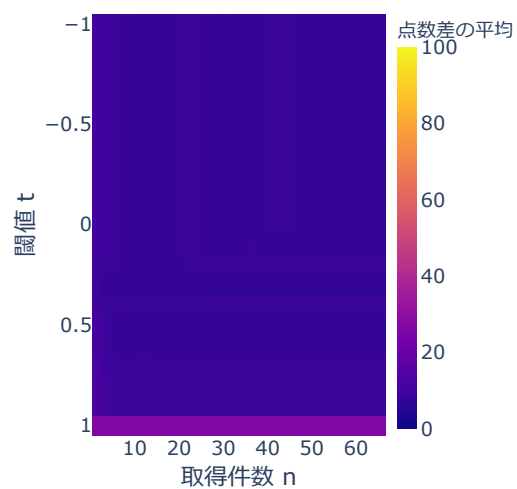
(c) TF-IDF



(d) Word2Vec



(e) Doc2Vec



(f) scoreEx

図 4.3 各手法における推定評価点の誤差 (考察 2.3.3)

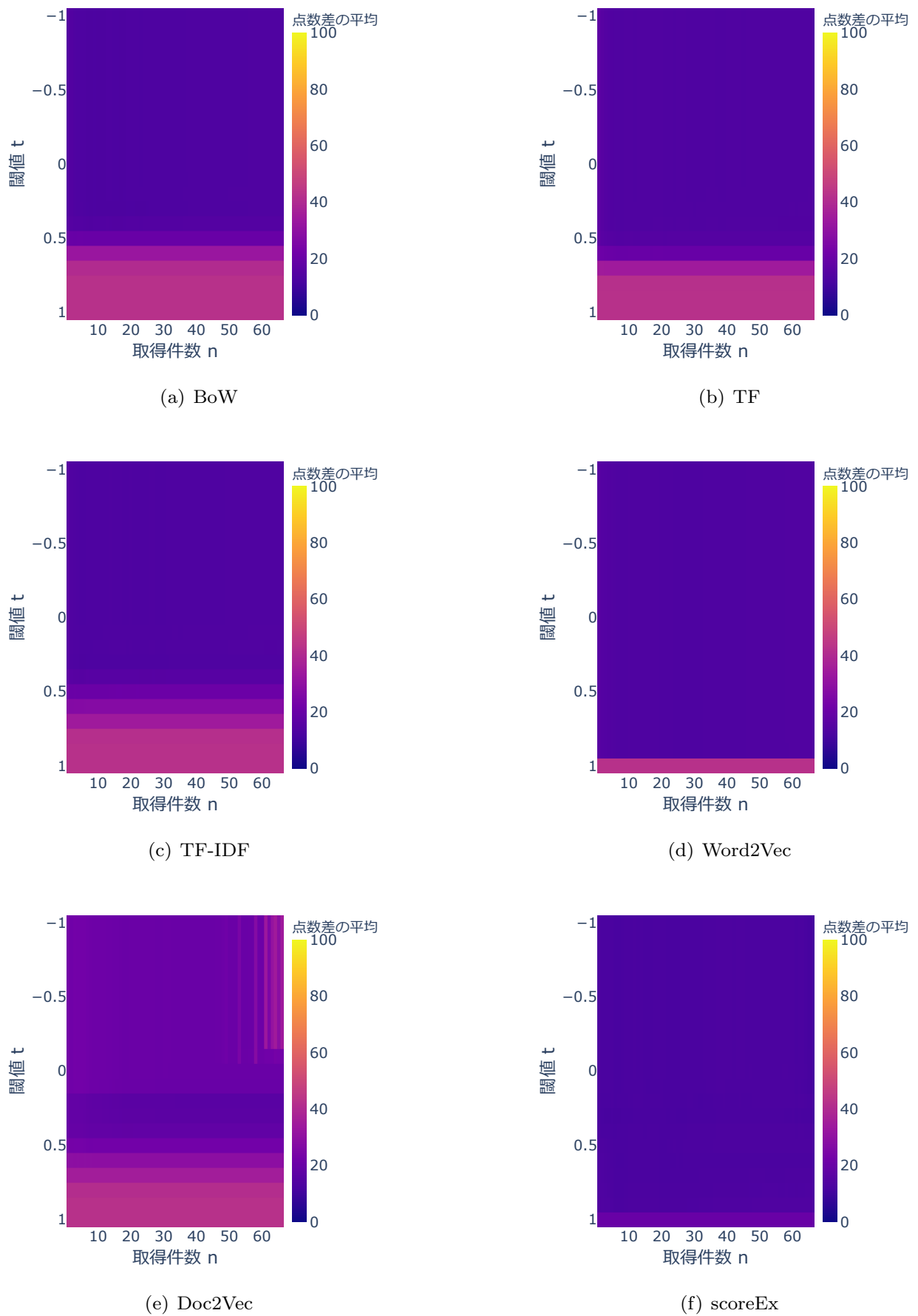


図 4.4 各手法における推定評価点の誤差（問題の区別なし）

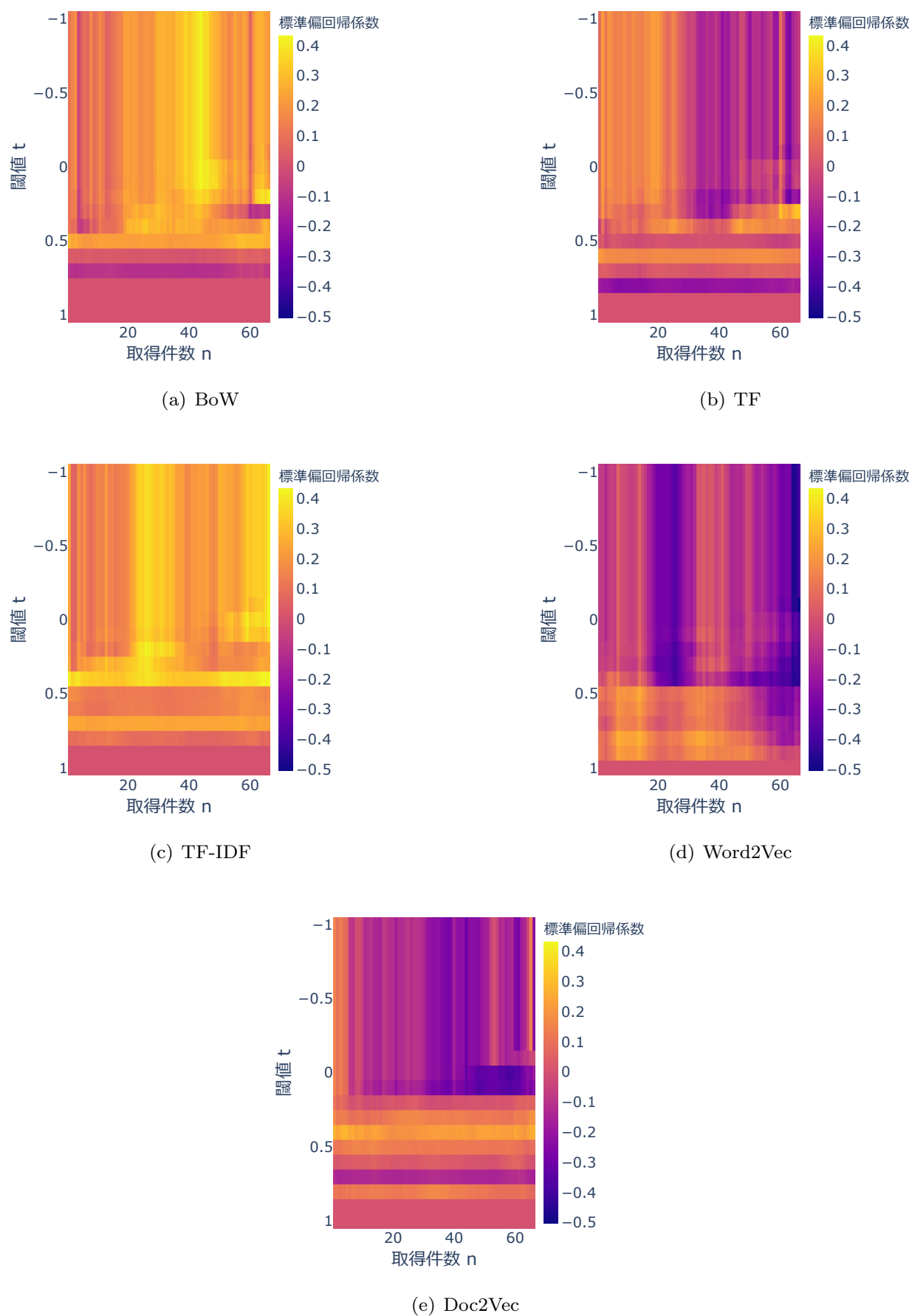


図 4.5 重回帰分析における各手法の標準偏回帰係数（考察 2.3.1）

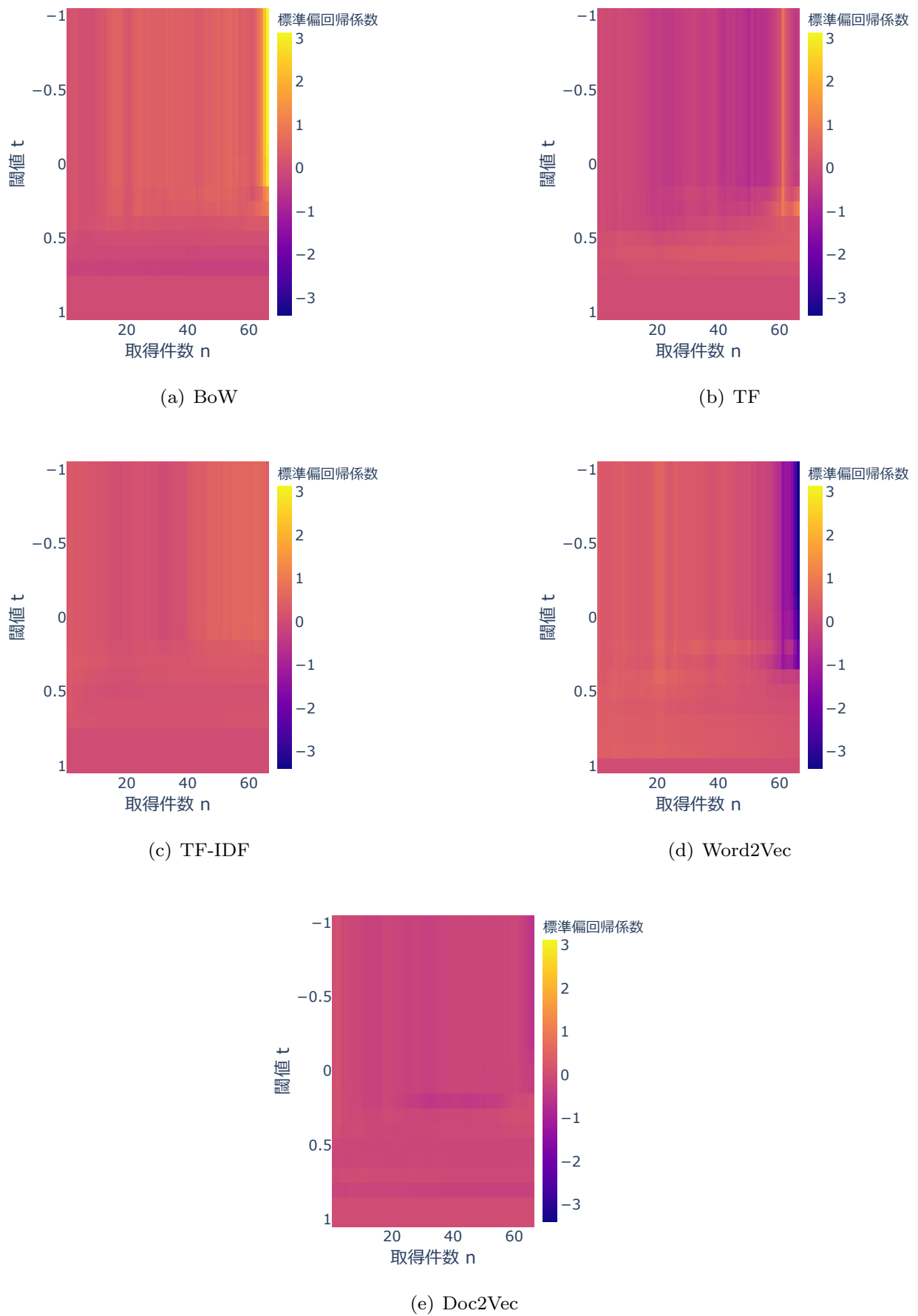


図 4.6 重回帰分析における各手法の標準偏回帰係数 (考察 2.3.2)

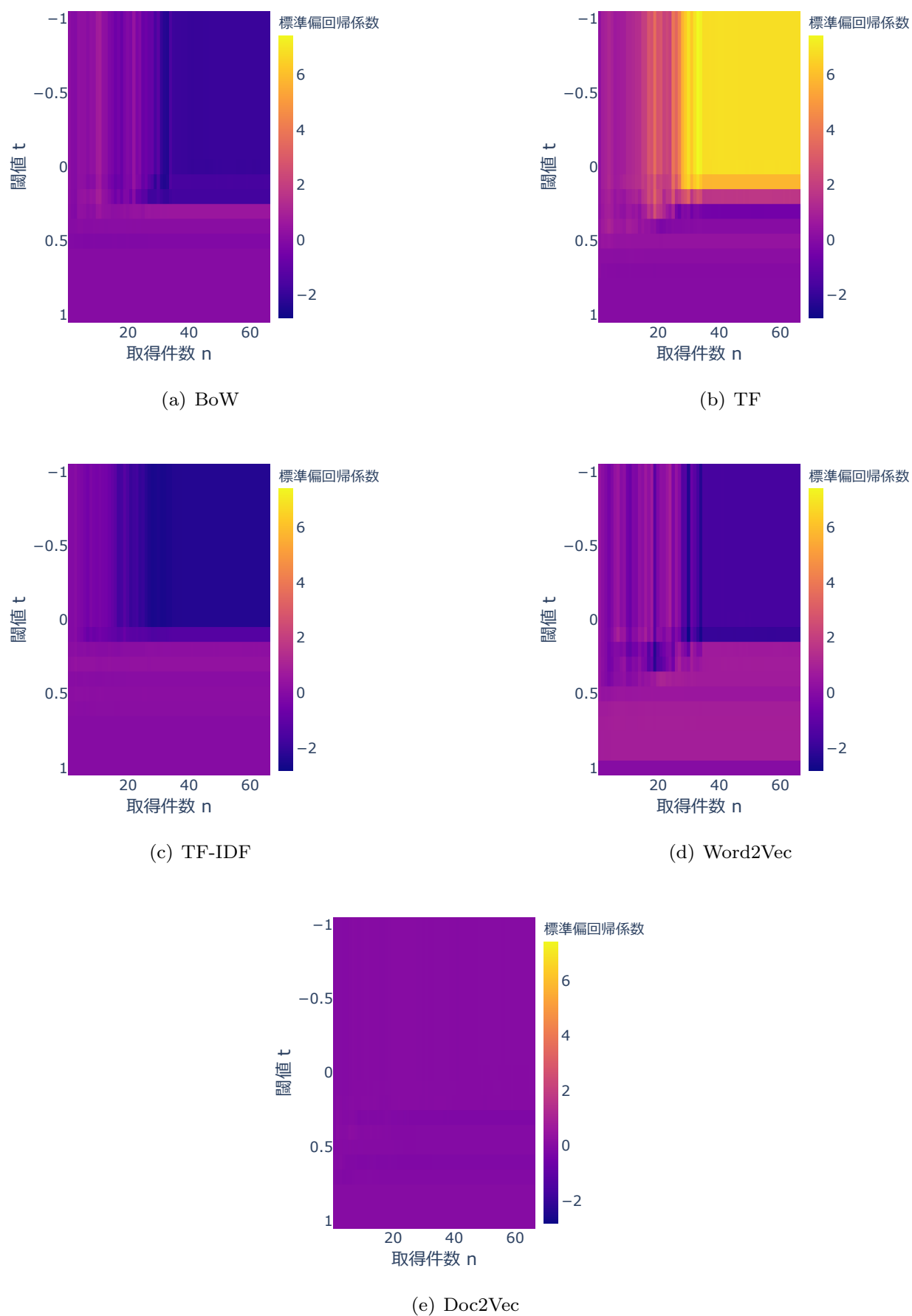


図 4.7 重回帰分析における各手法の標準偏回帰係数 (考察 2.3.3)

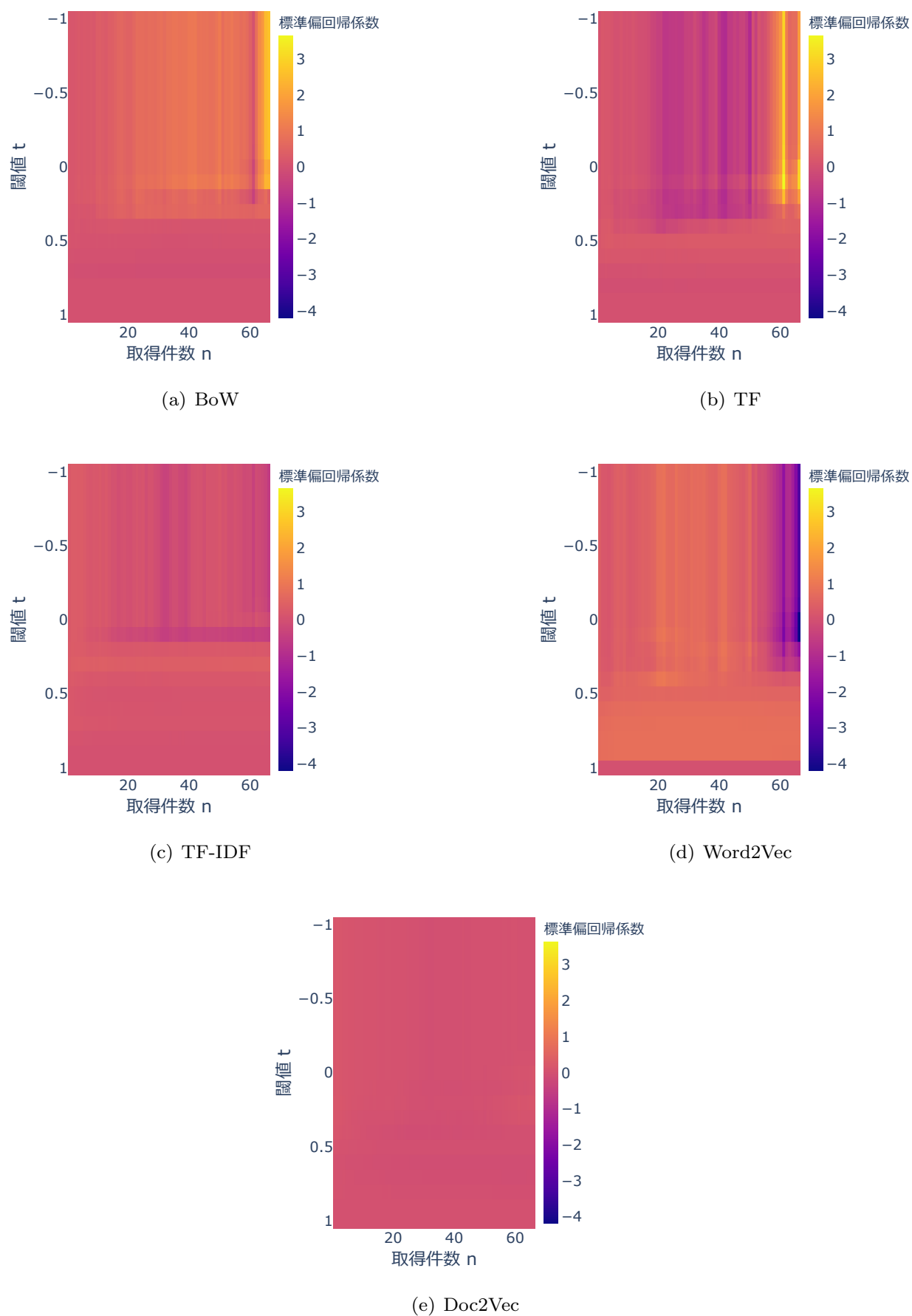
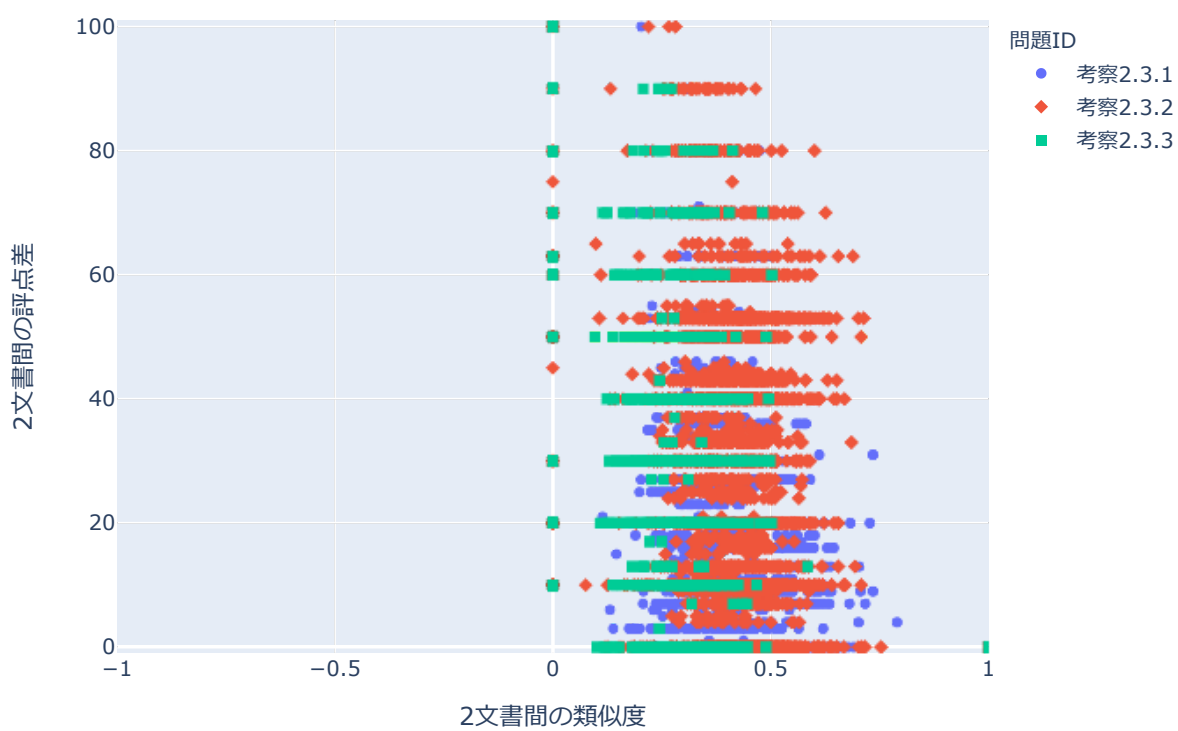
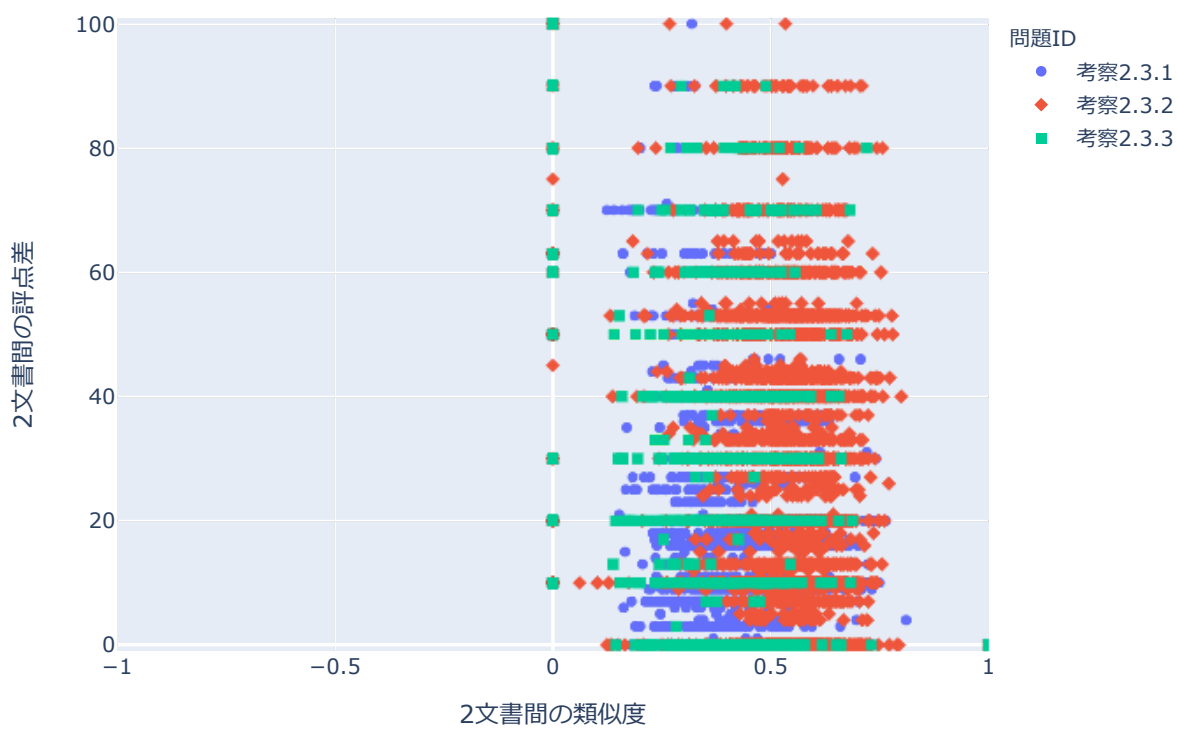


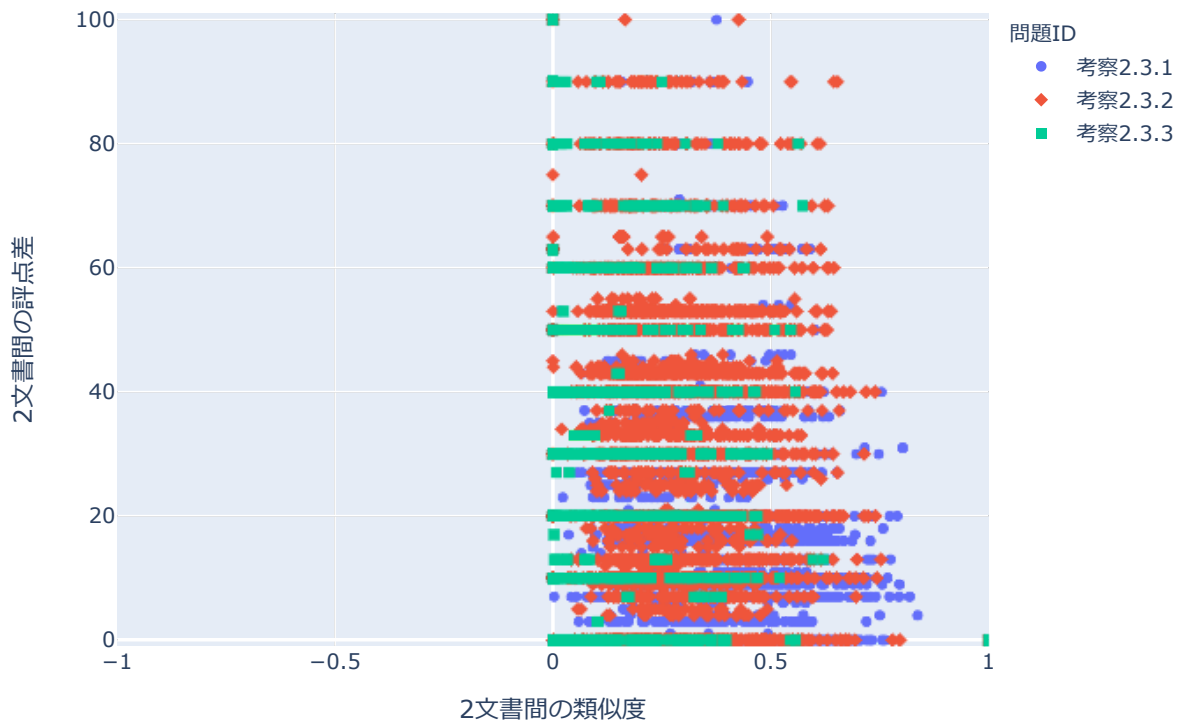
図 4.8 重回帰分析における各手法の標準偏回帰係数（問題の区別なし）



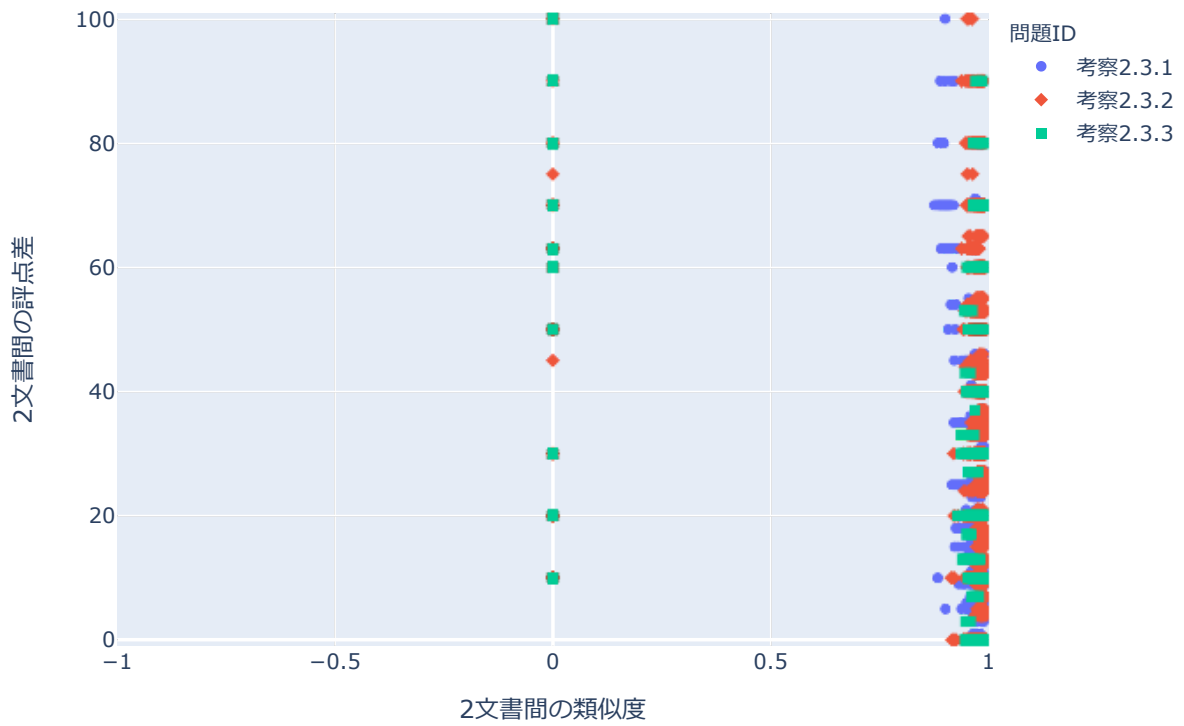
(a) BoW



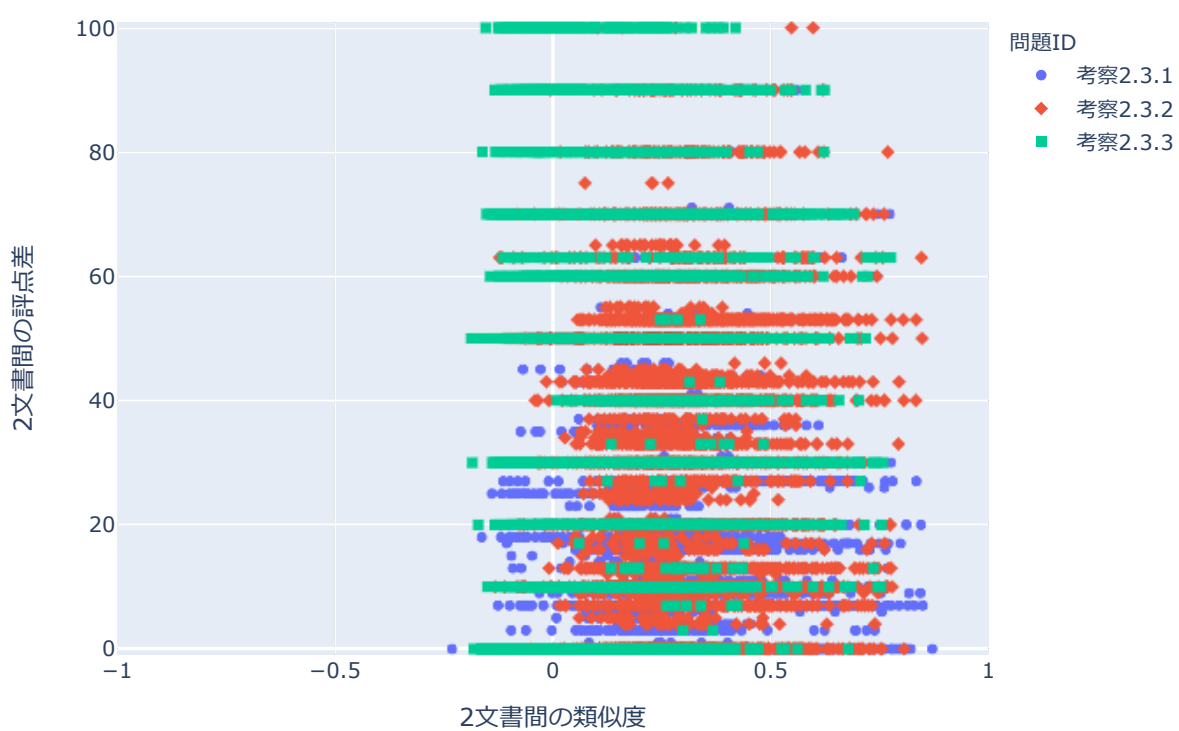
(b) TF



(c) TF-IDF



(d) Word2Vec



(e) Doc2Vec

図 4.9 2 文書間の評価点の差と類似度の関係

4.7 フィードバック生成の実験結果

式 (3.5) 及び式 (3.6) を用いて、それぞれのベクトル化手法によって求められた単語と、式 (3.7) によってそれぞれのベクトル化手法によって求められた単語に対して総合的に見て最も重要であると判断された単語群がどの様になったのか、表 4.6 から表 4.13 に示す。

フィードバックについて、BoW, TF, TF-IDF によって求められる重要語について、問題と解答と関連があるような単語を出力することが出来た。例えば、考察 2.3.1 ではランダム性がキーワードであり、ランダム性が問題なのだと気づきやすいような単語が出力されていると考えられる。また、考察 2.3.2 の結果では、表 4.10 について、TF-IDF では wing という単語が出力されており、扱っている数独という問題について X-wing 法という数独を解く考え方に繋がる可能性が考えられる。また、形態素解析による前処理の段階で、複合語の処理を高度化させることで、手法名が単一の名詞として出力させることが出来ると考えられる。考察 2.3.3 では、模範解答が一意に決まるような問題ではなかったことから、数独に関わる全般的な話題に関わるような単語が多く出力されていた。本稿では模範解答のみを使用していたが、過去データから対象の問題について、高得点者のグループを抽出して差ベクトルを取るといった手法によって、改善が出来るのではないかと考えられる。

また、フィードバックについて、本稿で提案した Word2Vec 及び Doc2Vec を用いたフィードバック生成は期待したような出力を得ることが出来なかった。これは、本稿で提案した式 (3.2) による文書ベクトル化を再検討する必要があると考えられる。

表 4.5 問題文と解答例 (考察 2.3.1)

問題文	(考察) 同じ難易度「Easy」の同じ問題に対して、同じ基本手法を適用しているにも関わらず、「solve」実行の度にアルゴリズム処理時間や空白セルアクセス回数が異なることである⇒何故ならば、(理由の分析を各自の言葉で！)
解答例 1	何故ならば、空白セルに入力する候補となる数字の順番をランダムに生成している為である。
解答例 2	何故ならば、空白セルに入力する候補を保存している ArrayList がランダムな数列で生成されている為である。

表 4.6 解答例 1 に対するフィードバック (考察 2.3.1)

BoW	(, (,), 2, Step, generateRandomNumbers, solve, sudoku, アクセス, アルゴリズム
TF	候補, 数字, 生成, 順番, ば, ランダム, 為, (, (,)
TF-IDF	アルゴリズム, sudoku, generateRandomNumbers, 解, solve, 回数, アクセス, 処理, (), Step
Word2Vec	レエブン, キメラクインケ,))》, ラズファード, overeating, VPCF, WDARL, マーガレット・コール, ザラリオ, ガンダペディア
Doc2Vec	christo, ビジネスグランドワークス, 54965532, kawasak, 赤倉_(列車), yamanashikotsu, コン・ヨンスク, ヤーギン, バケーキ
importanceEx	sudoku, generateRandomNumbers, アルゴリズム, solve, (), Step, 解, 回数, アクセス, 処理

表 4.7 解答例 2 に対するフィードバック (考察 2.3.1)

BoW	(, (,), 2, Step, generateRandomNumbers, solve, sudoku, アクセス, アルゴリズム
TF	ArrayList, 保存, 数列, 候補, 生成, ば, ランダム, 為, (, (
TF-IDF	アルゴリズム, sudoku, generateRandomNumbers, 解, solve, 回数, アクセス, 処理, (), Step
Word2Vec	レエブン, キメラクインケ,))》, ラズファード, overeating, VPCF, WDARL, マーガレット・コール, ザラリオ, ガンダペディア
Doc2Vec	バケーキ, アニソクイズ, yamanashikotsu, コン・ヨンスク, 赤倉_(列車), kawasak, 54965532, ビジネスグランドワークス, christo
importanceEx	sudoku, generateRandomNumbers, アルゴリズム, solve, (), Step, 解, 回数, アクセス, 処理

表 4.8 問題文と解答例（考察 2.3.2）

問題文	（問題点）問題の難易度が難しくなる程，巻き戻し（バックトラック）が多く発生し，空白セルに一度入力した数字のリセットも頻発してしまい，効率が悪く，アルゴリズム処理時間も膨大になっていく⇒この問題を解決する為には，（解決アプローチや方法を各自の言葉で！）
解答例 1	この問題を解決する為には，注目する空白セルに対して，解の候補を限定し，一意に定まるセルを優先的に処理する手法によって，巻き戻しやランダム性を減らす事が考えられる．
解答例 2	この問題を解決する為には，注目する空白セルに対して数字を決め打ちする事や，人間が数独を解く場合に行うような相互関係に基づく解の決め方によって，巻き戻しやランダム性を減らす事が考えられる．

表 4.9 解答例 1 に対するフィードバック（考察 2.3.2）

BoW	ランダム, 一意, 事, 処理, 性, 手法, 限定, セル, 候補, 優先
TF	入力, セル, 空白, /, 的, -, 0, 1, 3, [
TF-IDF	入力, sudoku, 左上, notty, analytics, バックトラック, 数字, 試み, wing, tech
Word2Vec	レエブン, キメラクインケ, ”)》, ラズファード, ザラリオ, overeating, WDARL, VPCF, スーパーキュウレンオー, マーガレット・コール
Doc2Vec	三浦清美, コン・ヨンスク, バケーキ, 54965532, yamanashikotsu, アニソクイズ, 赤倉_(列車), christo, ビジネスグランドワークス
importanceEx	入力, セル, 空白, /, sudoku, 的, 左上, 数字, 試み, -

表 4.10 解答例 2 に対するフィードバック（考察 2.3.2）

BoW	ランダム, 事, 場合, 性, 決め, 決め方, 相互, 関係, セル, 人間
TF	入力, セル, 空白, /, 的, -, 0, 1, 3, [
TF-IDF	入力, sudoku, 左上, notty, analytics, バックトラック, 試み, wing, セル, 候補
Word2Vec	レエブン, キメラクインケ, ”)》, ラズファード, overeating, ザラリオ, スーパーキュウレンオー, WDARL, VPCF, случаях
Doc2Vec	デス・ナイト, バケーキ, kawasak, コン・ヨンスク, 54965532, アニソクイズ, 赤倉_(列車), christo, ビジネスグランドワークス
importanceEx	入力, セル, 空白, /, 的, sudoku, 左上, 試み, 候補, -

表 4.11 問題文と解答例 (考察 2.3.3)

問題文	(以上の他にも、各自で考察、問題分析する！)
解答例 1	Step(3) の setCellValue 関数による空白セルの入力及び、Step(5) での入力済みセルへの初期化の時 GUI 処理が行われている。これによって巻き戻しの回数が多くなるにつれ、処理時間が長くなる事が考えられる。この問題を解決する為に、問題を解き終わった時のみ GUI 処理を行う事が考えられる。
解答例 2	解の候補を限定し、一意に入力できるセルを全て入力した後に、解の候補が絞り切れない場合が存在すると考えられる。この問題を解決する為には、空白セルに存在する複数存在する解の候補に対して並列処理を行うことで、逐次処理に比べ処理時間を削減することが考えられる。

表 4.12 解答例 1 に対するフィードバック (考察 2.3.3)

BoW	(,), 3, 5, GUI, Step, setCellValue, これ, 入力, 処理
TF	解, 度, 性, 的, 難易, アクセス, アルゴリズム, バックトラック, ランダム, 一般
TF-IDF	解, 難易, バックトラック, 短所, アルゴリズム, 長所, 模範, 解答, 最短, ランダム
Word2Vec	キメラクインケ, レエブン, overeating, ")》, WDARL, マーガレット・コール, 蕭珪, 》, 翹錠, ザリユース
Doc2Vec	アメリカ独立戦争の情報戦略, 坂齋小一郎, 貞山堀, JOD, コラブ, 甲斐智陽, 大友康匠, christo, Boiz, :en:The_Molly_Maguire_(film)
importanceEx	解, 難易, 度, 性, 的, バックトラック, 短所, アルゴリズム, 長所, 模範

表 4.13 解答例 2 に対するフィードバック (考察 2.3.3)

BoW	こと, 一意, 並列, 候補, 入力, 全て, 処理, 切れ, 削減, 存在
TF	事, 度, 性, 的, 難易, アクセス, アルゴリズム, バックトラック, ランダム, 一般
TF-IDF	難易, バックトラック, 短所, アルゴリズム, 長所, 模範, 解答, 最短, ランダム, 膨大
Word2Vec	キメラクインケ, レエブン,))》, overeating, 蕭珪, WDARL, VPCF, 》, マーガレット・コール, Н а ч а л ь н ы й
Doc2Vec	54965532, 連名, kufura, Boiz, ジョアン・マリオ, :en:The_Molly_Maguire_(film), christo
importanceEx	難易, 事, 度, 性, 的, バックトラック, 短所, アルゴリズム, 長所, 模範

第 5 章

追加実験

4.7 節において、Word2Vec 及び Doc2Vec による単語の推定が期待していた結果とはならなかった。本章では追加実験として、gensim の Word2Vec 及び Doc2Vec モデルに実験的に実装されている predict_output_words 関数を用いた重要語抽出を行った結果を示す。

4.7 節と比較して、Word2Vec 及び Doc2Vec から、与えられた入力に関わる重要単語について出力することが出来た。しかし、実装されている内容について未調査な部分があるため、今後の研究の中で考察を深めていく予定である。

表 5.1 gensim による解答例 1 に対するフィードバック (考察 2.3.1)

BoW	(, (,), 2, Step, generateRandomNumbers, solve, sudoku, アクセス, アルゴリズム
TF	候補, 数字, 生成, 順番, ば, ランダム, 為, (, (,),)
TF-IDF	アルゴリズム, sudoku, generateRandomNumbers, 解, solve, 回数, アクセス, 処理, (), Step
Word2Vec	アルゴリズム, solve, 空白, バックトラック, 代入, 乱数, セル, 再帰, 当り, ランダム
Doc2Vec	アルゴリズム, セル, 空白, 代入, バックトラック, 入力, 数字, 乱数, 難易, 解
importanceEx	アルゴリズム, solve, sudoku, generateRandomNumbers, (), Step, 解, 回数, アクセス, 処理

表 5.2 gensim による解答例 2 に対するフィードバック (考察 2.3.1)

BoW	(, (,), 2, Step, generateRandomNumbers, solve, sudoku, アクセス, アルゴリズム
TF	ArrayList, 保存, 数列, 候補, 生成, ば, ランダム, 為, (, (
TF-IDF	アルゴリズム, sudoku, generateRandomNumbers, 解, solve, 回数, アクセス, 処理, (), Step
Word2Vec	アルゴリズム, solve, 空白, バックトラック, 代入, 乱数, セル, 当り, 再帰, ランダム
Doc2Vec	解, 難易, 乱数, 数字, 入力, バックトラック, 代入, 空白, セル, アルゴリズム
importanceEx	アルゴリズム, solve, sudoku, generateRandomNumbers, (), Step, 解, 回数, アクセス, 処理

表 5.3 gensim による解答例 1 に対するフィードバック (考察 2.3.2)

BoW	ランダム, 一意, 事, 処理, 性, 手法, 限定, セル, 候補, 優先
TF	入力, セル, 空白, /, 的, -, 0, 1, 3, [
TF-IDF	入力, sudoku, 左上, notty, analytics, バックトラック, 数字, 試み, wing, tech
Word2Vec	代入, アルゴリズム, 空白, solve, 乱数, セル, バックトラック, 再帰, 数字, 入力
Doc2Vec	リセット, バックトラック, 配列, 乱数, 数字, 代入, アルゴリズム, 入力, 空白, セル
importanceEx	入力, 空白, セル, /, 数字, sudoku, 的, 左上, 試み, -

表 5.4 gensim による解答例 2 に対するフィードバック (考察 2.3.2)

BoW	ランダム, 事, 場合, 性, 決め, 決め方, 相互, 関係, セル, 人間
TF	入力, セル, 空白, /, 的, -, 0, 1, 3, [
TF-IDF	入力, sudoku, 左上, notty, analytics, バックトラック, 試み, wing, セル, 候補
Word2Vec	代入, アルゴリズム, 空白, solve, 乱数, セル, バックトラック, 再帰, 数字, 入力
Doc2Vec	リセット, 乱数, バックトラック, 配列, 数字, 代入, アルゴリズム, 入力, 空白, セル
importanceEx	入力, セル, 空白, /, 的, sudoku, 左上, 試み, 候補, -

表 5.5 gensim による解答例 1 に対するフィードバック (考察 2.3.3)

BoW	(,), 3, 5, GUI, Step, setCellValue, これ, 入力, 処理
TF	解, 度, 性, 的, 難易, アクセス, アルゴリズム, バックトラック, ランダム, 一般
TF-IDF	解, 難易, バックトラック, 短所, アルゴリズム, 長所, 模範, 解答, 最短, ランダム
Word2Vec	バックトラック, アルゴリズム, 代入, 乱数, solve, セル, 空白, 当り, 再帰, 数字
Doc2Vec	入力, 当り, 乱数, 代入, 効率, バックトラック, 解, セル, 難易, アルゴリズム
importanceEx	解, 難易, 度, 性, 的, バックトラック, アルゴリズム, 短所, 長所, 模範

表 5.6 gensim による解答例 2 に対するフィードバック (考察 2.3.3)

BoW	こと, 一意, 並列, 候補, 入力, 全て, 処理, 切れ, 削減, 存在
TF	事, 度, 性, 的, 難易, アクセス, アルゴリズム, バックトラック, ランダム, 一般
TF-IDF	難易, バックトラック, 短所, アルゴリズム, 長所, 模範, 解答, 最短, ランダム, 膨大
Word2Vec	バックトラック, アルゴリズム, 代入, 乱数, セル, 空白, solve, 当り, 再帰, 数字
Doc2Vec	処理, 数字, 入力, 乱数, 解, バックトラック, 効率, セル, アルゴリズム, 難易
importanceEx	難易, 事, 度, 性, 的, バックトラック, アルゴリズム, 短所, 長所, 模範

第6章

まとめと今後の課題

本稿では、様々なベクトル化手法を用いて、学生が提出する考察型レポートを文書ベクトルへ変換し、文書ベクトル演算を用いることで評価点の推定やフィードバックの生成を行う手法及びシステムを提案した。評価実験の結果から重回帰分析を行い、その結果からそれぞれのベクトル化手法の強みを活かした点数の推定を行うことが出来たが、相関分析から、文書ベクトルが類似しているならば、与えられた評価点も類似しているという仮定が間違っている可能性が高く、使用しているベクトル化手法や、推定手法を再検討する必要があることが分かった。また、フィードバックの生成について、本稿で提案した手法では期待した出力を得ることが出来ず、学生に対して全く関係のないフィードバックを与えてしまう恐れがあることが分かった。しかし、追加実験から問題に適した単語を出力することが出来ていることから、使用しているベクトル化手法による重要語分析について、検討が必要であると考えられる。

6.1 本稿の技術的貢献と社会的貢献

本稿の技術的貢献として、関連研究では考慮されていなかった、過去に提出された学生の解答を利用した評価点の推定手法及び、フィードバックの生成手法について提案したことが挙げられる。これによって、教員が必ず模範解答を用意する必要が無く、十分な過去データが存在すれば、問題に対する解答の範囲が幅広い場合についても適切に評価点を推定し、フィードバックの生成が行うことが期待出来る。

また、社会的な貢献として、教員によるフィードバックをシステムが代行し、学生が自分の取り組んだ課題の結果について評価点及び調べるべきトピックについて迅速に把握することで、教員にかかる負担を軽減し、これまで教員によるフィードバックを受け身で待つばかりであった学生が自分の力で調べて学び、成長するように促す環境を整備していることが考えられる。

6.2 今後の課題

今後追加すべき機能や改善する点として、以下の4点が考えられる。

1. 剽窃を検出する機能
2. 教員の考える評価基準によって調整を行う機能
3. 適切な日本語を用いているか判別する機能
4. フィードバックを行う対象を判別する機能

1点目は、受講している学生間で、レポートの写し合いをすることを防ぎ、学生自身の力でレポートを作成する為に必要であると考えられる。2点目は、本稿で提案したシステムでは、ベクトル化手法によってのみ点数が推定され、教員による自由な調整が出来ないようになっている。教員が考えるレポート要件として、例えば文字数などの制限を与える場合、それも考慮に入れることが出来るような機能が必要であると考えられる。3点目は、ベクトル化手法では分ち書きされた単語列を解析しており、文脈依存でないベクトル化手法が悪影響を及ぼす場合が考えられる。また、レポートとしてふさわしくない書き方などをフィードバックとして与えることも、適切なフィードバックの1つであると考えられるため、この機能は必要であると考えられる。4点目は、本稿で提案したシステムは、全員に対してフィードバックを与える様に設定されているが、実際の評価点が高い学生は、フィードバックが必要ない可能性がある。また、ただ提出することでフィードバックを基にレポートを作成する可能性も考えられる。このようなことを防ぐための機能があれば、真にフィードバックを必要としている学生に対して与えられるシステムになるのではないかと考えられる。

謝辞

本論文の作成にあたり，データセットの提供を含め，終始適切な助言を賜り，また丁寧にご指導戴いた服部峻助教に深く御礼申し上げます．また，同研究室の皆様方につきましては，研究に対する意見や論文の添削などのご指導を戴きました．心より感謝申し上げます．

参考文献

- [1] 文部科学省, GIGA スクール構想の実現パッケージ, https://www.mext.go.jp/content/20191219-mxt_syoto01_000003363_14.pdf (2021-12-16 参照) .
- [2] 文部科学省, 高等学校情報科に関する特設ページ, https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1416746.htm (2021-12-16 参照) .
- [3] Shun Hattori, Hiroyuki Kameda, Knowledge-based Compiler with e-TA for Software Engineering Education, Proceedings of the 9th Joint Conference on Knowledge-Based Software Engineering (JCKBSE'10), pp.265–278 (2010).
- [4] 小林 洋生, 自然言語処理技術を用いた答案採点支援システムの検討, https://www.ams.giti.waseda.ac.jp/data/pdf-files/2020_kobayashi_bt.pdf (2021-12-16 参照) .
- [5] 鈴木 千尋, 佐藤 直行, Doc2Vec を用いた国語記述式答案の自動評価, 情報処理学会第 81 回全国大会講演論文集, pp.135–136 (2019).
- [6] J-PlatPat 特開 2017-188039 複数の類似度算出によってテキストの点数を推定するプログラム, 装置及び方法. <https://www.j-platpat.inpit.go.jp/c1800/PU/JP-2017-188039/0F7797C29FBA9391826891398632F04813D60EB6D5855F34EAE770613829685/11/ja> (2021-12-16 参照) .
- [7] atama+. <https://corp.atama.plus/> (2021-12-16 参照) .
- [8] models.tfidfmodel - TF-IDF model - gensim - Radim Rehurek. <https://radimrehurek.com/gensim/models/tfidfmodel.html> (2021-12-16 参照) .
- [9] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean, Distributed Representations of Words and Phrases and their Compositionality, 2013. <https://arxiv.org/pdf/1310.4546.pdf> (2022-01-09 参照) .
- [10] Tomas Mikolov, Kai Chen, Greg Corrado, Jeffrey Dean, Efficient Estimation of Word Representations in Vector Space, 2013. <https://arxiv.org/pdf/1301.3781.pdf> (2022-01-09 参照) .
- [11] models.word2vec - Gensim - Radim Rehurek, <https://radimrehurek.com/gensim/models/word2vec.html> (2022-01-09 参照) .
- [12] models.doc2vec - Doc2vec paragraph embeddings — gensim, <https://radimrehurek.com/gensim/models/doc2vec.html> (2022-01-09 参照) .
- [13] Quoc Le, Tomas Mikolov, Distributed Representations of Sentences and Documents,

<https://arxiv.org/pdf/1405.4053v2.pdf> (2022-01-09 参照) .

- [14] 室蘭工業大学, 情報学応用演習 A - シラバス参照, <https://u.muroran-it.ac.jp/kyomu/syllabustop2021/2603/J4135.html> (2022-01-09 参照) .