

学習記事推薦のための推薦システムの開発と手法の評価

岡井 成遊^{1,a)} 大久保 文哉^{2,b)} 内山 英昭^{4,c)} 峰松 翼^{2,d)} 谷口 雄太^{3,e)} 島田 敬士^{2,f)}

概要: 学習要点記事を投稿したり、他者と共有したりする学習支援システムを展開することは、投稿者自身の学びの振り返りの支援のみならず、他者の学びや理解の支援にも繋がるなどの利点がある。一方で、システム利用者が他の学習者が投稿した記事を検索する際に、記事が増加するにつれて、検索者が意図する記事を発見することが困難になる。そこで本研究では、記事検索者とシステム間のインタラクティブな学習記事推薦システムを提案する。推薦記事に対する記事検索者からのフィードバックに基づいて次の記事を適応的に推薦する仕組みにより、単純な推薦手法と比較して、利用者の要望に応じた多くの記事を推薦できることを確認した。

1. はじめに

ラーニングアナリティクスの研究は多岐にわたるが、近年では、Learning Management System(LMS)の小テスト機能を利用して既習の学習内容の定着を図った研究 [1] や、学生に学習内容を振り替える日誌を書かせて内容を分析し授業改善に役立てた研究 [2] 等のような学生の学習の振り返りを支援する研究が盛んに行われている。

このような研究の一つとしてとして Learner Generated Contents(LGC)に関する研究が進められている [3]。LGCとは、「学習者が学習過程で生成した、知識や学習方法などのコンテンツ」のことである。学習者は自らの学習を振り返り、教材を理解する上で要点となるものを図、文章、画像などにまとめた1つの記事としてLGCを管理するシステム(以下、「LGCシステム」とする)に投稿する。また、投稿された記事の中から、学習者は自由に記事を閲覧することもできる。記事を閲覧することで、自分が理解できなかった点、難しいと感じた点についての理解が深まり、彼らの弱点克服や、学びの広がりへの貢献できることが期待される。LGCは、学習記事を通して知識を共有する仕組みであり、学習者自身、並びに他の学習者の理解を促進するという観点から協調学習 [4] の一つであると言える。

現状のLGCシステムには、ユーザの求める記事を探すことが容易ではないという課題がある。その理由は、システムからユーザへの検索結果の提示の方向性と、テキストデータだけに頼った検索システムとなっていることが挙げられる。LGCシステムが提供する記事の検索機能として、全文検索とハッシュタグ検索がある。これらの検索手法は、あるキーワードを入力することでそのキーワードを含む(に紐づいた)記事を検索結果としてユーザに表示するものである。しかし、これらの機能のみでは各記事のテキストデータだけに依存した検索となる。そのため、ユーザが学習する上で効果的でない記事が推薦される場合がある。また、検索結果をユーザに提示した後にユーザのフィードバックを収集し、よりユーザが満足する記事を提示するといったフォローアップは行わない。そのため、所望の記事が見つからない場合、どのようなキーワードが検索結果の改善につながるかを自ら考えた上で、再度検索を行う必要がある。しかし、調べたい科目に関する知識が浅い場合は、ユーザ自らの力で検索結果を改善するキーワードを考え出すことは困難である。

本研究では、このような問題を解決するために、インタラクティブなコミュニケーションを図る学習記事推薦システムと、ユーザの推薦結果に対するフィードバックの蓄積により得られるデータ(以下、「属性」とする)を考慮する推薦手法を提案する。開発した学習記事推薦システムでは、ユーザが推薦結果に満足できない場合、推薦結果に対するフィードバックを基に推薦結果を改善して再提示する。これにより、ユーザが詳しい知識を持たない分野について調べる場合でも、フィードバックを基に最終的に満足できる結果を提供することを目指す。このような複数回の

¹ 九州大学 大学院システム情報科学府 情報理工学専攻 修士課程

² 九州大学 大学院システム情報科学研究院

³ 九州大学 情報基盤研究開発センター

⁴ 奈良先端科学技術大学 先端科学技術研究科

a) okai@limu.ait.kyushu-u.ac.jp

b) fokubo@ait.kyushu-u.ac.jp

c) hideaki.uchiyama@is.naist.jp

d) minematsu@limu.ait.kyushu-u.ac.jp

e) taniguchi.yuta.941@ait.kyushu-u.ac.jp

f) atsushi@ait.kyushu-u.ac.jp

推薦を行うシステムにおいては、ユーザとシステムとのインタラクションが複数回発生するため、ストレスを感じさせないインターフェイスを設計することが重要である。そのため、ユーザとの自然なやり取りを可能にする対話型コミュニケーションをベースとしたチャットボットを用いることを提案する。本システムでは、日本国内での普及率の高さから LINE が提供する公式アカウントサービス [5] を用いる。提案する推薦手法では、既存のテキスト解析技術により得られる各記事の単語の重要度に基づく特徴と、より抽象的な特徴に着目した属性を組みあわせる。これにより、既存の推薦手法では対応しきれなかったユーザの意図を考慮した推薦を目指す。

2. 関連研究

推薦システムが教育分野において活発になり始めたのは、E-learning システムの導入が行われ始めた 2000 年代初期である [6]。中山ら [7] は、デジタル教材のページ毎に関連する情報 (web ページ等) を推薦するシステム K-TIPS を考案した。K-TIPS は実証実験によって、推薦による学習支援の効果と推薦結果を教員にフィードバックすることによる教育改善の可能性を示した。Wang ら [8] は、オンラインフォーラムを利用して英語を学習している大学生に対して、全ユーザのブラウザの閲覧ログを基に協調フィルタリングを行う推薦システムを開発した。実験結果の分析から、推薦機能ありのフォーラムを用いた学生の方がフォーラムに投稿されている記事を長く閲覧する傾向があることと、文章を要約して書く力が向上していることが示された。Gauth ら [9] は、コンテンツベースフィルタリングと優秀な学習者による教材評価を組み合わせた新たな推薦システムを提案し、コンテンツベースフィルタリングあるいは協調フィルタリングのみを用いたシステムと比較することで、提案システムの方が高い精度で教材を推薦することが可能であることを示した。

推薦システムにおいて、協調学習での成果物を推薦する事例はこれまでに報告されていない。また、教育分野における推薦システムはシステムからユーザへ一方通行に推薦結果を提示する場合がほとんどである。そのため、インタラクティブなコミュニケーションを実現した点において本推薦システムは独自性を持つ。また、推薦手法の観点から見ても、本研究での属性と同じ観点に着目した推薦手法は報告されていない。

3. 学習記事推薦システム

本稿で提案する学習記事推薦システムはユーザとシステムとの双方向のインタラクションを用いて、多様なユーザのニーズに適応することを目指す。ユーザとのインタラクションはチャットボット (以下、「teaching bot」と呼ぶ) を通して行う。システムはインタラクションを通して、ユー

2020-08-13 14:20:57
4つのフーリエ変換のまとめ

フーリエ級数展開: 「連続」「周期的」な信号を「離散的」「非周期的」な信号に変換する
フーリエ変換: 「連続」「非周期的」な信号を「連続」「非周期的」な信号に変換する
離散時間フーリエ変換: 「離散的」「非周期的」な信号を「連続」「周期的」な信号に変換する
離散フーリエ変換: 「離散的」「周期的」な信号を「離散的」「周期的」な信号に変換する

離散時間フーリエ変換はフーリエ級数の逆である。実際、複素フーリエ級数における複素フーリエ係数を求める式において、 t を Ω にして、 $T_0 = 2\pi(\omega_0 = 1)$ とすると、逆離散時間フーリエ変換と同じ形になる。また、複素フーリエ級数における周期信号を表す式も同様に、 t を Ω にして、 $T_0 = 2\pi(\omega_0 = 1)$ とすると、離散時間フーリエ変換と同じ形になる。

図 1 公開記事の例

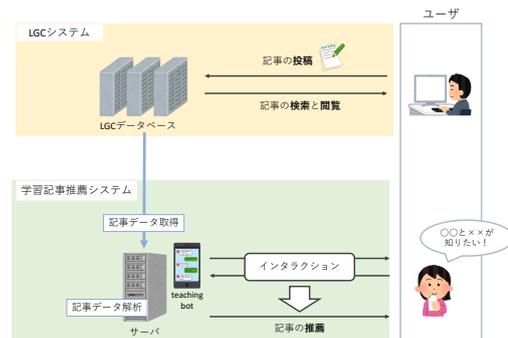


図 2 LGC システムと学習記事推薦システムの関係

ザからの推薦記事に対するフィードバックを収集する。また、学習記事推薦システムは、LGC システムから定期的な記事データの取得と、解析を行う。記事データの解析手法や、フィードバックの分析手法については、4 章で説明を行う。

3.1 LGC システム

LGC システムは Web アプリケーションであり、記事の投稿・検索機能を提供している。ユーザによって投稿された記事は科目ごとにデータベース (以下、「LGC データベース」と呼ぶ) を用いて、管理される。投稿された記事の中で、教員あるいは Teaching Assistant (TA) による確認を受けた記事のみ他の学生に公開される。公開された記事の一例を図 1 に示す。

3.2 LGC システムと学習記事推薦システムの関係

本稿で提案する学習記事推薦システムは、30 分に一度、記事データの取得と解析を行い、LGC システムとの記事データの同期を図る。学習記事推薦システムは、インタラクションを通じた記事の推薦に特化したシステムである。記事データの解析方法は、4 章にて説明する。LGC システムと学習記事推薦システムの関係を図 2 に示す。

3.3 インタラクションの流れ

学習記事推薦システムとユーザ間のインタラクションの流れを図 3 に示す。最初に、ユーザは属性を入力する。属性は、「基本」、「発展」、「まとめ」、「その他」の 4 つから構成

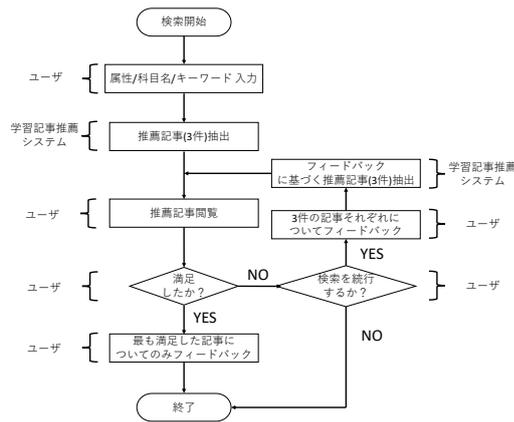


図3 インタラクションの流れ

される。ユーザは自らが所望する記事に適合する属性を一つ選択する。次に、科目名、キーワードを入力する。入力されたデータを基に学習記事推薦システムは、推薦記事を3件抽出する。その後、ユーザは推薦された記事を閲覧する。推薦結果に対して満足した場合、ユーザは3件の記事の中で最も満足した記事がどれであるかをフィードバックする。一方、推薦結果に対して満足できず追加での推薦を所望する場合、推薦された記事それぞれに対して、フィードバックを与える。具体的には、ユーザの所望する記事に対して、3件の記事それぞれが要求に「近い」か、「遠い」かをフィードバックする。フィードバックを分析し、システムは推薦記事の再抽出、再提示を行う。フィードバックの具体的な分析方法は4章にて説明する。

4. 推薦手法

記事を検索する際に全文検索のような部分文字列の一致のみを考慮した手法では、ユーザの求める記事を推薦できるとは限らない。そこで本論文では、各記事が持つ特徴をベクトルの形式で抽出するという手法を提案する。また、ユーザの推薦システムに対するクエリも記事が持つ特徴ベクトルと同次元のベクトル(以下、「ユーザベクトル」とする)として表現し、ユーザベクトルと類似度の高い記事を抽出する。その後、抽出された記事の中でユーザが指定した属性に沿った記事を推薦する。本章の構成として、4.1節にて記事データを特徴ベクトルに変換する手法について述べる。4.2節にて、提案した属性について述べる。最後に、4.3節にて特徴ベクトルと属性の組み合わせ手法について説明する。

4.1 記事データの特徴ベクトル化

記事データはHTML形式のテキストデータとして保存されている。前処理としてまず、形態素解析エンジンMecab[10]を用いて名詞のみを抽出し、ストップワードの除去[11]を行う。その後、TF-IDF[12]を用いて、各記事

contents_id	title	basic	development	summary	other
129	極、零点について	39	1	15	3
130	サンプリング	37	2	14	3

図4 属性値がデータベースに格納されている例

のTF-IDF値を計算し、それらを基に各記事をベクトル化する処理を行う。TF-IDFに関する詳しい説明は参考文献を参照されたい。

4.2 各記事が持つ属性値

各記事は各属性に関する整数値の重みを一意に持ち、それらは学習記事推薦システムサーバのローカルデータベースに格納されている。データベースに格納されているレコードの例を図4に示す。ここで、「基本」属性は“basic”，「発展」属性は“development”，「まとめ」属性は“summary”，「その他」属性は“other”というカラム名と対応している。また“contents_id”はユーザがLGCシステムに対して記事を投稿した際に一意に発行される識別子である。属性値は全て0で初期化されており、それぞれの属性値は、3.3節で述べたユーザが推薦結果に満足した場合のフィードバックを受け取った際に更新される。具体的には、もし「基本」属性を選択したユーザが最も満足した記事の“contents_id”が“129”であるならば“contents_id”が“129”の記事が持つ「基本」についての属性値に対して、1を加算する処理を行う。

4.3 推薦アルゴリズム

本システムでは、1度目の推薦と2度目以降の推薦で異なるアルゴリズムを用いて、推薦記事を抽出しているが、2つのアルゴリズムは共通したコンセプトの基に設計されている。それはテキストデータから取り出せる情報を基に6つの記事を抽出し、さらにユーザが求めている属性に関する重みが最も大きい3つの記事を抽出することである。これにより、ユーザが求めているキーワードと属性を共に満たす記事の抽出を行っている。ユーザの満足する記事を提供するためには、キーワード以外に、読みやすさや記事を理解するための前提知識といった、様々な要素がユーザのニーズに適合している必要がある。提案するアルゴリズムではキーワードに加えて、属性を考慮することで、より多様なユーザのニーズに対応することを目指す。以下1回目と2回目以降の推薦で用いるアルゴリズムについてより詳しく述べる。

4.3.1 1度目の推薦

1度目の推薦に用いられる抽出アルゴリズムについて以下に手順を示す。1から2の手順は、従来の文書検索システムでよく用いられるキーワードによる抽出であり、3から4の手順は本研究独自の属性による抽出である。

- (1) ユーザが指定した科目内で投稿されている記事から、指定されたキーワードについての TF-IDF 値の総和を求めてそれらを降順にソートする。
- (2) 上位 6 件の記事を抽出する。
- (3) 抽出した 6 件の記事について、ユーザが指定した属性値で降順にソートする。
- (4) 上位 3 件の記事を推薦する。

4.3.2 2 回目以降の推薦

1 回目の推薦アルゴリズムでは、手順 1 から 2 において TF-IDF 値の総和を基に記事を抽出していたが、2 回目以降ではユーザベクトルを基に記事の抽出を行う。このユーザベクトルは、科目内の記事が持つ特徴ベクトルと同じ次元数を持ち、ユーザが求める記事の特徴を反映している。ユーザベクトルは初期ベクトルを \vec{u}_1 、1 回目に推薦される記事のベクトル集合を D_1 として式 1 に基づき初期化される。ユーザベクトルの更新は、文書検索システムにおいて用いられることの多いロッチオの式 [13] に基づき行う。 i 回目に推薦された記事集合のうち、「近い」という評価を受けた記事 (適合記事) 集合を D_{pi} 、「遠い」という評価を受けた記事 (不適合記事) 集合を D_{ni} として、式 2 により更新される。なお、 α, β, γ はハイパーパラメータであり、一般に適合記事から得られる情報の方が不適合記事から得られる情報よりも多いことを考慮して、 $\alpha = 1.0, \beta = 0.9, \gamma = 0.1$ としている。

$$\vec{u}_1 = \frac{1}{|D_1|} \sum_{\vec{a} \in D_1} \vec{a} \quad (1)$$

$$\vec{u}_{i+1} = \alpha \vec{u}_i + \beta \frac{1}{|D_{pi}|} \sum_{\vec{a} \in D_{pi}} \vec{a} - \gamma \frac{1}{|D_{ni}|} \sum_{\vec{a} \in D_{ni}} \vec{a} \quad (i \in R) \quad (2)$$

ロッチオの式を基にユーザベクトルを更新することで、不適合記事に含まれる単語の重みを小さく、適合記事に含まれる単語の重みを大きくすることができる。次に、更新されたユーザベクトルと各記事が持つ特徴ベクトルのコサイン類似度を計算する。ロッチオの式を基に更新を繰り返すことで、記事が抽出される記事が含むキーワードが徐々にユーザの要求を満たすものになることが期待できる。

2 回目以降の推薦ではまず、更新されたユーザベクトルと各記事の特徴ベクトルのコサイン類似度を基に、6 件の記事を抽出する。その後、6 件の記事から属性値を基に 3 件の記事を抽出する。

以上の説明を踏まえて、2 回目以降の推薦アルゴリズムの手順を以下にまとめる。

- (1) 式 2 を基に、ユーザベクトルを更新する。
- (2) 1 回目の推薦以降まだ推薦していない記事について、それぞれユーザベクトルとのコサイン類似度を計算する。
- (3) コサイン類似度について降順にソートする。

表 1 各手法の違い

	ベクトルの更新	属性によるソート
手法 1	×	×
手法 2	○	×
手法 3	×	○
提案手法	○	○

- (4) 上位 6 件の記事を抽出する。
- (5) 抽出した 6 件の記事について、ユーザが指定した属性値で降順にソートする。
- (6) 上位 3 件の記事を推薦する。

5. 実験

提案手法により推薦精度が向上することを調査するための比較実験を行った。実験の概要としては、まず初めにキーワードを用意し、次にそれぞれのキーワードと関連が深い記事を、推薦されることが望ましい記事 (正解記事) として用意する。その後、キーワードが提案手法と提案手法の処理の一部を削除した比較手法に入力された場合についてシミュレーションを行う。最後に、各手法で正解記事を推薦することができたか何回目に推薦できたかを調べるというものである。実験の対象となる記事は、「デジタル信号処理」の講義について扱った記事である。本章の構成として、比較対象となる手法を 5.1 節、正解記事の抽出方法等の実験設定を 5.2 節に、最後に結果を 5.3 節で述べる。

5.1 比較手法

実験では提案手法に加えて以下の 3 種類の比較手法を用いた。各手法の違いは表 1 の通りである。この表において、○は推薦アルゴリズム内で用いられることを、×は用いられないことを意味する。

以下に各手法の推薦記事抽出手順を示す。手法 1 から 3 におけるユーザベクトルの初期化は式 1 に従って行った。

5.1.1 手法 1 : ベースライン

フィードバックを基にしたベクトルの更新、属性によるソートを行わない手法である。推薦アルゴリズムは以下の通りである。

1 回目の推薦

- (1) ユーザが指定した科目内で投稿されている記事から、ユーザが指定したキーワードについての TF-IDF 値の総和を求めてそれらを降順にソートする。
- (2) 上位 3 件の記事を推薦する。

2 回目以降の推薦

- (1) 1 回目の推薦以降まだ推薦していない記事について、それぞれユーザベクトルとのコサイン類似度を計算する。
- (2) コサイン類似度について降順にソートする。
- (3) 上位 3 件の記事を推薦する。

5.1.2 手法 2：ベクトル更新あり

フィードバックを基にしたベクトルの更新を行い、属性によるソートは行わない手法である。推薦アルゴリズムは以下の通りである。

1 回目の推薦

(1) ユーザが指定した科目内で投稿されている記事から、ユーザが指定したキーワードについての TF-IDF 値の総和を求めてそれらを降順にソートする。

(2) 上位 3 件の記事を推薦する。

2 回目以降の推薦

(1) 式 2 を基にユーザベクトルを更新する。

(2) 1 回目の推薦以降まだ推薦していない記事について、それぞれユーザベクトルとのコサイン類似度を計算する。

(3) コサイン類似度について降順にソートする。

(4) 上位 3 件の記事を推薦する。

5.1.3 手法 3：属性あり

フィードバックを基にしたベクトルの更新は行わず、属性によるソートを行う手法である。推薦アルゴリズムは以下の通りである。

1 回目の推薦

(1) ユーザが指定した科目内で投稿されている記事から、ユーザが指定したキーワードについての TF-IDF 値の総和を求めてそれらを降順にソートする

(2) 上位 6 件の記事を抽出する。

(3) 抽出した 6 件の記事について、ユーザが指定した属性値で降順にソートする。

(4) 上位 3 件の記事を推薦する。

2 回目以降の推薦

(1) 1 回目の推薦以降まだ推薦していない記事について、それぞれユーザベクトルとのコサイン類似度を計算する。

(2) コサイン類似度について降順にソートする。

(3) 上位 6 件の記事を抽出する。

(4) 抽出した 6 件の記事について、ユーザが指定した属性値で降順にソートする。

(5) 上位 3 件の記事を推薦する。

5.2 実験設定

5.2.1 学生に対する事前アンケート

正解記事の抽出に必要なデータを得るために、学生に対して実験の対象記事についてのアンケートを行った。アンケートの協力者は、2020 年度前期「デジタル信号処理」を受講した学生 40 名である。アンケート項目は以下の通りである。A1 から A3 に関しては数字が大きいほど強く当てはまるとした 7 段階評価、B1 から B4 については当てはまるものにチェックを入れる方式で行った。

A1 文章が分かりやすく書かれている。

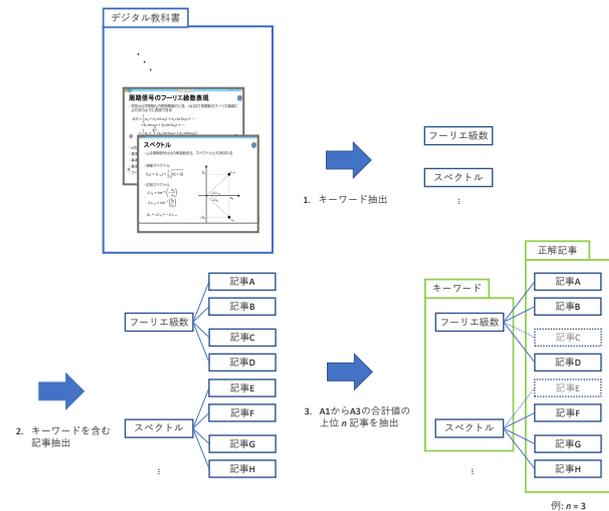


図 5 キーワードと正解記事の抽出

A2 記事の内容が分かりやすく説明されている。

A3 記事によって授業内容の理解が促されると思う。

B1 基本的な内容を扱っている。

B2 発展的な内容を扱っている。

B3 学習内容をまとめている。

B4 その他

A1 から A3 についての評価値の合計は各記事の質を評価したものであり、後述の正解記事を抽出する際に利用した。B1 から B4 についての各評価値の合計は属性を評価したものであり提案手法と手法 3 における各記事の属性値として利用した。

5.2.2 キーワードと正解記事の抽出

デジタル教科書からキーワードを抽出し、それぞれのキーワードに対して正解記事を以下で述べる手順で紐づけた。キーワードは、「デジタル信号処理」について扱ったデジタル教科書の各ページで重要度が最も高い単語を抽出^{*1}し、不適切な単語 (例: サンプル, 三角関数) を省いた 29 単語である。実験は、抽出する記事の件数 n を 1 から 8 の各整数値に設定して行った。正解記事の抽出アルゴリズムについて手順を以下に示す。また、キーワードの抽出から正解記事の抽出までの手順を図 5 に示す。

(1) キーワードを含む記事を抽出

(2) 抽出された記事をアンケート項目 A1 から A3 の合計値で降順にソートする。

(3) 上位 n 件をそのキーワードの正解記事とする。

5.2.3 実験方法

各手法でキーワードを入力したときに正解記事を推薦できるかシミュレーションを行った。提案手法と手法 3 において属性値を考慮する場合は、4 つの属性それぞれが選択された場合についてシミュレーションし、提案手法と手法 2 においてフィードバックを考慮する場合は、すべての

*1 重要度の算出方法としては TF-IDF を用いた

での正解記事集合 Get Correct Article Set(GCAS), 最小平均 Minimum Average(MinAve), 最大平均 Maximum Average(MaxAve), 総平均 All Average(AllAve) は以下の式で定義される. 以下では, 紙面の都合上, *keyword* を *k* として, *choice* を *c* として表す.

$$PKS = \{k \mid k \in KS \wedge |GVCS(k)| > 0\} \quad (7)$$

$$GVCS(k) = \{c \mid c \in GCS \wedge IP(c, k)\} \quad (8)$$

$$IP(c, k) = |GAS(c, k) \wedge GCAS(k, n)| > 0 \quad (9)$$

$$MinAve = \frac{\sum_{k \in PKS} \text{Min}(\{GC(c, k) \mid c \in GVCS(k)\})}{|PKS|} \quad (10)$$

$$MaxAve = \frac{\sum_{k \in PKS} \text{Max}(\{GC(c, k) \mid c \in GVCS(k)\})}{|PKS|} \quad (11)$$

$$AllAve = \frac{\sum_{k \in PKS} \sum_{c \in GVCS(k)} GC(c, k)}{\sum_{k \in PKS} |GVCS(k)|} \quad (12)$$

5.3 結果

n に対する推薦可能なキーワード数の関係を図 10, *n* に対する最小平均の関係を図 11, *n* に対する最大平均の関係を図 12, *n* に対する総平均の関係を図 13 に示す. 推薦可能なキーワード数については, 提案手法が各 *n* において最も高いという結果となった. このことより, 提案手法がより幅広いキーワードからの推薦に対応できることが分かる. また, 最小平均については各手法間に大きな差は見られず, 最大平均, 総平均については提案手法が比較手法に対して 0.3 から 0.5 回程度多いという結果となった. このことから, 提案手法は比較手法に対して, 推薦回数の増加を 1 回未満に抑えつつ, 推薦可能なキーワード数の増加を達成したことが分かる. また, 最小平均については, 増加が抑えられていることから, ユーザの選択次第で, 比較手法と変わらない回数で所望の記事を推薦できることが分かる.

6. まとめ

本研究では, インタラクティブな学習記事推薦システムの開発に加えて, 単語ベースの特徴と, 属性を組み合わせることによる新たな推薦手法を開発した. また, 実験結果を分析することで, 提案手法がより幅広いキーワードから記事を推薦できることを示した. 今後の課題として,

- (1) 特徴ベクトルの次元数の削減
- (2) 講義内での実証実験を用いた評価
- (3) 推薦アルゴリズムの改善

が挙げられる. 課題 1 については学習記事推薦システムが現在採用している手法では, 特徴ベクトルの次元数は記事が属する科目の単語集合の要素数と同じになる. このた

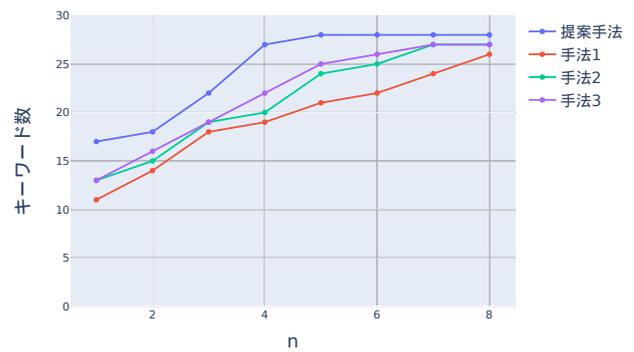


図 10 *n* に対する推薦可能なキーワード数の関係

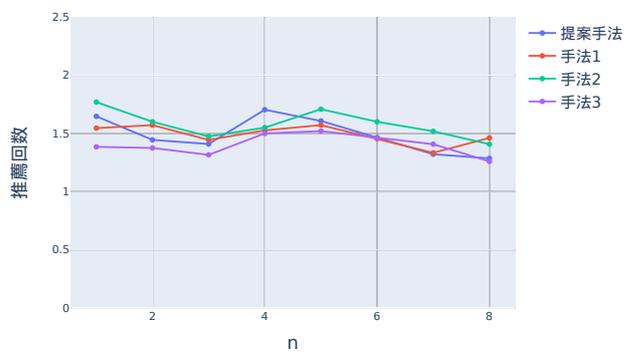


図 11 *n* に対する最小平均の関係

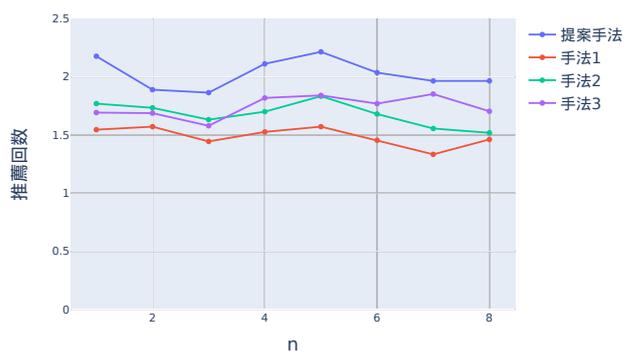


図 12 *n* に対する最大平均の関係

め, 一部科目については, 記事が持つ特徴ベクトルの次元数が非常に大きくなる可能性がある. この場合, 類似度の計算に多くの時間を消費してしまう. 現在の実装方法では, 記事が出てくるまでの待機時間は 6 秒から 8 秒ほどであった. 計算時間を削減するために, 出現数が極端に少ない単語の切り捨てや, PCA などの次元圧縮法を用いることが考えられる.

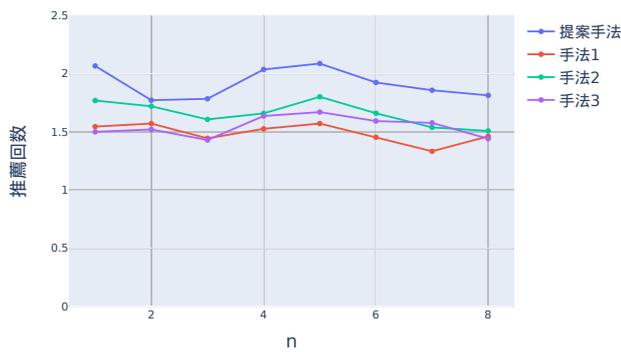


図 13 n に対する総平均の関係

課題 2 については今後、講義の受講者に学習記事推薦システムを利用してもらい、推薦情報の受率率を検証していく。また、所望の推薦情報が得られるまでの過程を分析し、精度向上を図る。さらに、多くの学生が求める記事を教員にフィードバックすることで、授業内容の改善を行う。投稿記事推薦システムが講義内で利用されることで、デジタル教科書の学習ログと学習記事推薦システムの利用ログの2つのデータを組み合わせた分析が可能となる。即ち、どのような学習をする学生がどのような疑問点を持ち学習記事推薦システムに質問したかが分かる。分析を進めることで、学生が学習記事推薦システムを利用する前に予測して記事を推薦することを目指す。その上で、ユーザにいつ、どれくらいの量の推薦を行うことで、どの程度の学習効果を得られるかについて調査を行う。

課題 3 について、現在の推薦アルゴリズムは、まず各記事が含む単語に注目した特徴ベクトルによる抽出を行い、さらに属性値による抽出を行うという2段階の抽出を行っている。今後の方針として、これらの情報を同時に抽出するアルゴリズムを開発することを考えている。各記事が含む単語と、属性を同時に反映できる新たな指標を開発することができれば、その指標に基づき、1回の抽出で記事を絞り込むことが可能となる。今後、様々な推薦システムを調査し、実現方法を検討していく予定である。

謝辞 本研究は、JST AIP 加速研究 JPMJCR19U1, JSPS 科研費 JP18H04125 の支援を受けたものである。

参考文献

- [1] 山本朋弘, 菅木禎史, 宇佐川毅, 清水康敬: 小学生対象の分数での LMS のテスト機能を活用した振り返り学習の展開と評価, 日本教育工学会論文誌, Vol. 36, No. Suppl., pp. 157-160 (オンライン), DOI: 10.15077/jjet.KJ00008609801 (2012).
- [2] Taniguchi, Y., Okubo, F., Shimada, A. and Konomi, S.: EXPLORING STUDENTS' LEARNING JOURNALS WITH WEB-BASED INTERACTIVE REPORT TOOL, 14th INTERNATIONAL CONFERENCE on COGNITION AND EXPLORATORY LEARNING IN

- [3] Uchiyama, H., Ishita, E., Watanabe, Y., Tomiura, Y., Shimada, A., and Yamada, M.: A framework for sharing learner generated contents in collaborative learning, *Proceedings of the 9th Asia-Pacific Conference on Library & Information Education and Practice (A-LIEP 2019)*, pp. 477-479 (2019).
- [4] Laal, M. and Laal, M.: Collaborative learning: what is it?, *Procedia-Social and Behavioral Sciences*, Vol. 31, pp. 491-495 (2012).
- [5] LINE: LINE for Business, <https://www.linebiz.com/jp/login/>. 最終閲覧日:2021 年 4 月 15 日.
- [6] Lu, J., Wu, D., Mao, M., Wang, W. and Zhang, G.: Recommender system application developments: A survey, *Decision Support Systems*, Vol. 74, pp. 12 - 32 (online), DOI: <https://doi.org/10.1016/j.dss.2015.03.008> (2015).
- [7] Nakayama, K., Shimada, A., Minematsu, T., Taniguchi, Y. and Taniguchi, R.: K-TIPS:Knowledge extension based on Tailor-made Information Provision System, *16th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2019)* (2019).
- [8] Wang, P.-Y. and Yang, H.-C.: Using collaborative filtering to support college students' use of online forum for English learning, *Computers & Education*, Vol. 59, No. 2, pp. 628 - 637 (オンライン), DOI: <https://doi.org/10.1016/j.compedu.2012.02.007> (2012).
- [9] Ghauth, K. I. and Abdullah, N. A.: The effect of incorporating good learners' ratings in e-Learning content-based recommender System, *Journal of Educational Technology & Society*, Vol. 14, No. 2, pp. 248-257 (2011).
- [10] Kudo, T., Yamamoto, K. and Matsumoto, Y.: Applying Conditional Random Fields to Japanese Morphological Analysis, *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, Association for Computational Linguistics, pp. 230-237 (online), available from (<https://www.aclweb.org/anthology/W04-3230>) (2004).
- [11] 大島裕明: SlothLib : Web 検索研究のためのプログラミングライブラリ, 日本データベース学会, 2008, pp. 113-116 (オンライン), 入手先 (<https://ci.nii.ac.jp/naid/10027607990/>) (2008).
- [12] Tata, S. and Patel, J. M.: Estimating the selectivity of tf-idf based cosine similarity predicates, *ACM Sigmod Record*, Vol. 36, No. 2, pp. 7-12.
- [13] 北研二, 津田和彦, 獅々堀正幹: 情報検索アルゴリズム, 共立出版 (2002).