

学生の復習支援に向けた小テストに関連する電子教材ページの自動抽出

石川 高志^{1,a)} 島田 敬士^{2,b)} 峰松 翼^{2,c)} 谷口 倫一郎^{2,d)}

概要：講義の一環として学生に小テストを課す授業が多く存在する。学生はそれらの小テストに取り組み、その後教科書の内容と照らし合わせて復習を行う。そのような復習作業の効率化を図るべく、本研究では小テストの設問文から自然言語処理の手法を用いて自動的に教科書内の関連ページを抽出するシステムを作成し、より正しく関連ページを抽出できるように精度の向上を図った。また、実際の講義において、抽出した関連ページを小テスト後の学生に提示し、どの程度理解の助けになるのかアンケート調査を行った。

1. はじめに

近年、情報処理技術の発展に伴って様々な分野で ICT (Information and Communication Technology) 化が進んでいる。教育分野においても電子媒体の利用促進により教育データが収集・蓄積され、分析することが可能になってきた [1]。多くの教育機関で学習者が個人の端末を持参して教育を受ける BYOD (Bring Your Own Device) が導入され、moodle [2] などの e-learning システムや、BookRoll などの [3] 電子教材配信システムが利用されている。これらを利用する利点として、インターネット上での出席管理や電子教材の配布が可能となることによる教育活動の効率化や、電子教材の閲覧履歴や小テスト、課題の提出などの学生の学習活動の記録を自動的に収集し分析できる点が挙げられる。そのような利点から教育の ICT 化への関心が高まり、情報処理技術を用いて学生の学習活動を支援するための研究が盛んに行われている。

例えば、CAI (Computer Assisted Instruction) を用いた学習支援や日誌の内容分析、知識マップによる知識構造の可視化など様々なアプローチによる研究が行われている。特に CAI に関しては、問題解決学習、ゲーム・シミュレーション学習、情報検索型学習、ドリル学習など教師対学生の形態では難しい効果的な学習の形が提案され、様々な教育現場で利用されている [4]。ただし既存の CAI の多くは学校単位やクラス単位での学習支援に焦点を当てており、それ

らの CAI を用いる集団に属する学生は同じような学習カリキュラムに従うこととなる。

一方、学生の学習スタイルや理解度には個人差があり、学習効果の最大化を目指すためにはそれぞれに合わせた学習内容を提供することが求められるという観点から、近年特に個人の学習状況により特化した支援を行う Adaptive Learning に注目が集まっている。例えば学生と教師の間での web 上でのコミュニケーションを利用した学習支援に関する研究 [5] や、学習者の知識や学習スタイルに合わせてそれぞれに適した教材を提案する研究 [6] など、技術の進歩とともに個人に適応する様々な研究が行われている。

そのような Adaptive Learning に着目したときに、それぞれの学生の理解度を確認する手段として講義における小テストの実施が挙げられ、実際に教育の ICT 化が進む各大学では授業時間中に e-learning システム上での小テストを実施する講義が増えている。小テストは電子教材配信システムに登録された電子教材に基づいて作成されており、学生は小テスト後のレビュー画面で自分の間違えた設問の有無やその内容を確認し、復習する場合はその設問に関連する電子教材のページを見て復習作業を行う。しかし、その関連ページというのは潜在的に教材中に存在しているが、小テスト作成者によってどの設問とどのページが関連しているかが明示されていないため、学生は教材を 1 から読み直し設問文や正答から関連ページを探すという作業を行う必要がある。そのような時間を要する復習作業は講義時間内に収まらず後回しにされる、もしくは復習自体が行われない場合も多い。

この問題の解決策として、間違えた設問内容を反映した復習用の資料を作成し、学生個人個人に提供することがで

¹ 九州大学 工学部 電気情報工学科

² 九州大学 大学院システム情報科学研究所

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

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

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

d) rin@ait.kyushu-u.ac.jp

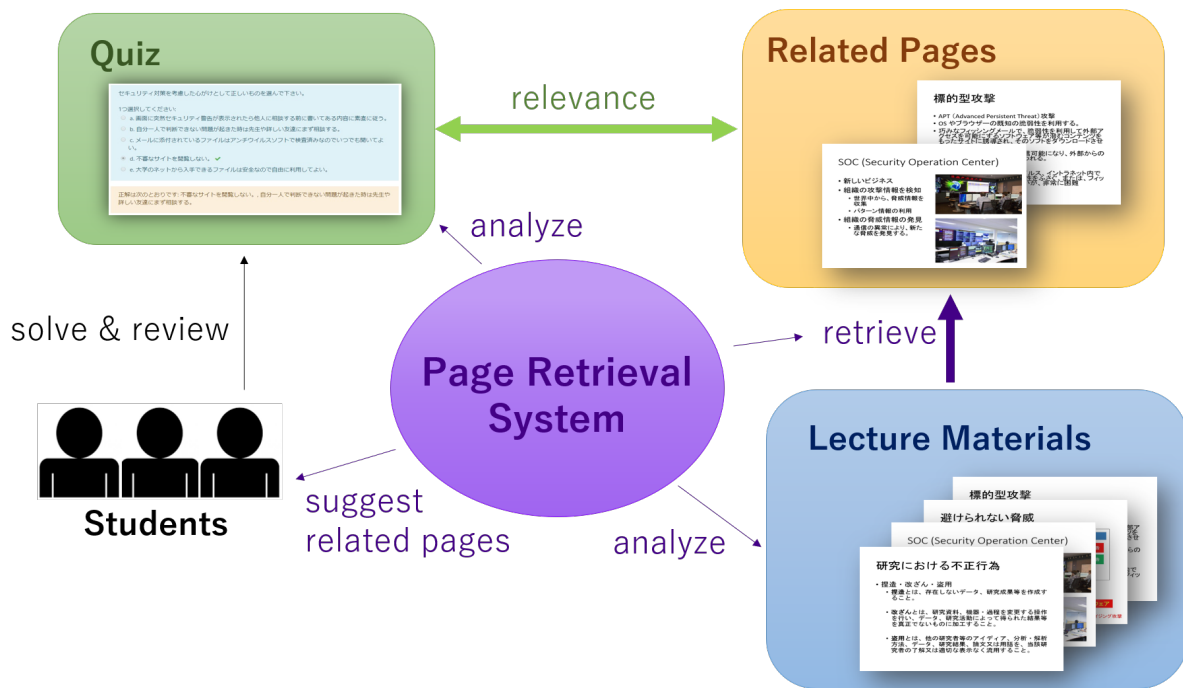


図 1: システム概略図

きれば各学生の理解が不十分な点を効率よく補えることが期待される。しかし、学生一人一人の小テスト結果を確認し指導教員がそれぞれに合わせて新しく資料を作るとするのは、膨大な労力と時間がかかり現実的でない。一方で小テストは講義で用いられる教材に基づいて作られているため、テキスト解析などの情報処理技術を適用し、元教材から自動的に各設問に関連する情報を検索することができれば、復習に必要な資料を再構成することができる可能性が考えられる。

そこで本研究では小テスト後の復習用資料作成に向けて、自然言語処理の手法である TF-IDF [7] と Word2Vec [8] を用いて、小テストの設問文から自動的に教材内の関連ページを抽出するシステムを作成する。その後複数の実験を行うことにより、より正しく関連ページを抽出できるようにシステムの精度向上を図る。また、実際の講義において、抽出した関連ページを小テスト後の学生に提示し、適切なページが抽出されているか、どの程度理解の助けになるのかといった点についてアンケート調査を行う。本研究の貢献は以下のとおりである。

- 小テストと教材のテキストデータから、小テストに関連する教材のページを抽出するシステムを提案する。
- word2vec を用いて教材の各ページと小テストの設問文のベクトル化を行うことの、関連ページ抽出に対する有効性を示す。
- 小テスト後のレビュー画面に関連ページを表示することによる、学生の復習効果の向上に対する有効性を示す。

2. 関連ページ抽出システム

2.1 システム概要

1 節に述べたように、TF-IDF と Word2Vec, MeCab を用いて電子教材と小テストのテキストを分析し、各設問との関連が高く、解答に必要な内容が記述されたページを出力するシステムを作成する。図 1 はシステムの概略図であり、小テストの設問文と BookRoll の電子教材を分析して小テストに関連するページを抽出し、学生に提示する仕組みを表している。システムの入力には電子教材のテキストデータと小テストのテキストデータを用い、出力として電子教材のページ番号を、小テストの設問ごとに各設問への関連度が高い順に並べたリストが得られるようにシステムを設計する。システムには東北大学の乾・鈴木研究室が公開している日本語 Wikipedia から作成された訓練済み Word2Vec [9] により生成された 200 次元の単語ベクトルを用いる。このベクトルの集合を W とすると、 W は日本語の Wikipedia 記事に存在する 100 万を超えるすべての単語を 200 次元のベクトルで表現している。関連ページ抽出の手順を 2.1.1~2.1.3 に示す。

2.1.1 教材のテキスト分析

BookRoll 教材の各ページの特徴ベクトル計算について説明する。BookRoll の教材は pdf データであるため、分析に用いるテキストデータはスクレイピングして余分な記号等を取り除くことができる。このテキストデータを分析し、次のようにして各ページのテキストデータのベクトル化を行う。

- i). 単語抽出:教材の各ページ(ページ j とする)内のテキストデータ p_j を MeCab を用いて形態素解析することにより, ページに含まれるすべての単語 t_i ($i = 1, 2, \dots, n_j$) を重複なく抽出し, 単語集合 $T_{p_j} = \{t_1, t_2, \dots, t_{n_j}\}$ を生成する.
- ii). TF-IDF の計算: p_j を 1 つ 1 つの文書とみなし, T_{p_j} に含まれる各単語 t_i と p_j の TF-IDF の値 $\text{TF-IDF}(t_i, p_j)$ をそれぞれ計算する.
- iii). ベクトル化: 前述の 200 次元訓練済み単語ベクトル集合 W を用いて p_j のベクトル化を行う. W から各 t_i に対応するベクトル \mathbf{v}_{t_i} を抜き出し, ページ j について, $\text{TF-IDF}(t_i, p_j)$ の値を重みとした \mathbf{v}_{t_i} の重み付き平均を計算する. この計算により得られる 200 次元のベクトルを p_j の特徴ベクトル \mathbf{v}_{p_j} とすると, \mathbf{v}_{p_j} は式 1 のように表される.

$$\mathbf{v}_{p_j} = \frac{\sum_i \text{TF-IDF}(t_i, p_j) \mathbf{v}_{t_i}}{\sum_i \text{TF-IDF}(t_i, p_j)} \quad (t_i \in T_{p_j}, \mathbf{v}_{t_i} \in W) \quad (1)$$

- i). ~ iii). を教材の全てのページについて行い, 各ページの特徴をベクトルで表現する.

2.1.2 小テストのテキスト分析

小テストの各設問のテキストデータに対しても, 教材のテキスト分析と同様に単語を抽出し, 同様の訓練済み単語ベクトル集合 W を用いて各設問の特徴ベクトルを生成するが, 教材の各ページとは異なり各設問のテキストが短いので, TF-IDF の計算は行わずに以下のように設問のテキストデータのベクトル化を行う.

- i). 単語抽出:小テストの各設問(設問 k とする)内のテキストデータ q_k を MeCab を用いて形態素解析することにより, 設問に含まれるすべての単語 t_i ($i = 1, 2, \dots, n_k$) を重複なく抽出し, 単語集合 $T_{q_k} = \{t_1, t_2, \dots, t_{n_k}\}$ を生成する.
- ii). ベクトル化: W から各 t_i に対応する 200 次元ベクトル \mathbf{v}_{t_i} を抜き出し, それら n 個の平均ベクトルを設問文の特徴ベクトル \mathbf{v}_{q_k} とする.

$$\mathbf{v}_{q_k} = \frac{\sum_{i=1}^{n_k} \mathbf{v}_{t_i}}{n_k} \quad (t_i \in T_{q_k}, \mathbf{v}_{t_i} \in W) \quad (2)$$

- i)., ii). をすべての設問について行うことにより, 教材の各ページと小テストの各設問を同じ次元のベクトルで表現することができ, ベクトルの比較によって各ページと各設問の関連度を計算することができるようになる.

2.1.3 関連ページ抽出

200 次元にそろえた BookRoll 教材の各ページと小テストの設問文のベクトル \mathbf{v}_{p_j} , \mathbf{v}_{q_k} は, コサイン類似度を比較することでベクトルの方向の類似度を知ることができる.

ベクトルの方向に近いほど, そのページと小テストの設問文が近い特徴を持つといえる. 各ページのテキストデータ p_j と小テストの各設問のテキストデータ q_k を入力としベクトル化して \mathbf{v}_{p_j} , \mathbf{v}_{q_k} を得る. その後 \mathbf{v}_{p_j} と \mathbf{v}_{q_k} のコサイン類似度を計算し k を固定して (j を動かして) 類似度が高い順にソートし, ランキングのリストを作成する (図 2).

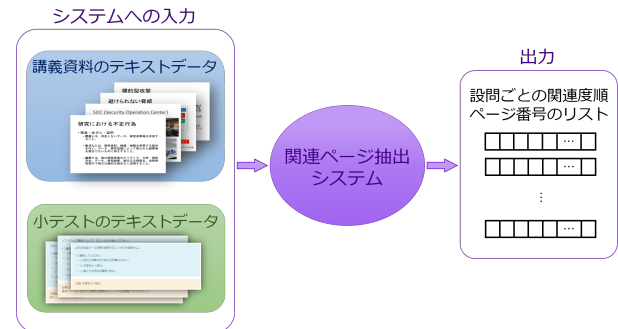


図 2: システム入出力

3. 実験

3.1 実験 1: 教材と設問文の類似度評価による正解ページ推定

3.1.1 実験内容

作成したシステムの精度を測るため, 「サイバーセキュリティ基礎論」の教材と小テストを用いて実験を行った. 講義は全 8 回で, 各回ごとに教材とそれに対応する小テストが存在する. 各設問は選択問題で, 3~7 個の選択肢が与えられている. 作成したシステムは, 小テストの各設問に対し関連度順に教材のページ番号のリストを出力するが, システムの精度を測るには, 出力した関連度の上位のページが小テストの解答に必要な内容が書かれた適切なページであるかを調べる必要がある. そこで, 設問ごとにその適切なページ(正解ページとする)をあらかじめ調べ, ページ番号のリストとして用意した. 正解ページについては, 教材の内容を理解できる大学生が該当教材の中から設問に関連するページを選択した. ただし, 小テストの問題がそのまま教材のみが更新されている場合があり, 解答に必要な内容が教材のどのページにも載っていない設問文が存在したため, そのような設問は実験から除外した.

教材のと小テストのテキスト分析は講義回ごとに行い, 各設問に対する出力のページ番号のリストはその設問に対応する講義回の教材から生成した. 正解ページには, 設問文のテキストと似た内容が書かれていることが多いが, 選択肢のテキストの内容が含まれていることもある. そこで, システムへの適切な入力を調べるため, 入力に用いるテキストデータとして 1) 設問文のみ, 2) 設問文とすべての選択肢, 3) 設問文と正しい選択肢の 3 つのパターンで分析した. 2) と 3) のように設問文と選択肢の複数のテキ

ストデータを1つにまとめる際には、テキストデータを連結し1つのテキストデータとして扱った。1), 2), 3)のそれぞれの場合について、教材のページ番号を、設問文との関連度が高い順に並べ、用意しておいた正解ページを全体の何番目に出力できたかを調べた。システムが上位 x ページ以内に正解ページを出力できたのは、全ての設問文のうちの $y\%$ というデータを x を1から2刻みでプロットし、Cumulative Match Curve(CMC) グラフを作成した。

3.1.2 実験結果

図3に結果を示す。青のグラフが1) 設問文のみ、橙が2) 設問文とすべての選択肢、緑が3) 設問文と正しい選択肢を入力として与えたグラフである。多少の逆転はあるが、おおむね3), 2), 1)の順に精度が良いことがわかる。3)の場合、実験に使った設問文のうち、全体の80%を上位3ページ以内に、全体の90%を上位5ページ以内に出力できたことがわかる。3), 2), 1)の順番となったことから、設問文に含まれる単語のみを用いて関連ページを検索するよりも、選択肢、特に正解の選択肢に含まれる単語を加味して関連ページを検索する方が高い確率で正解ページを抽出できるということがわかる。3)で上位5ページ以内に出力できなかった設問について、その原因を調べるため実際に該当する設問を確認すると、設問文や選択肢が長い文章となっておりキーワードではない単語が多く含まれている場合や、正解ページ内のテキスト量が多く、キーワードが含まれているものの他の単語の影響を大きく受けたと考えられる場合、正解ページ内に書かれた正解選択肢を選ぶために必要な情報が、テキストでなく図で表現されている場合があった。図の中に書かれたテキストはスクレイピングによって得ることができないため、本研究の手法では正解の内容が図で表現されたページを抽出することは不可能であるが、それ以外の原因については、入力データの最適化によって上位に出力できる可能性があり、正解ページを抽出する確

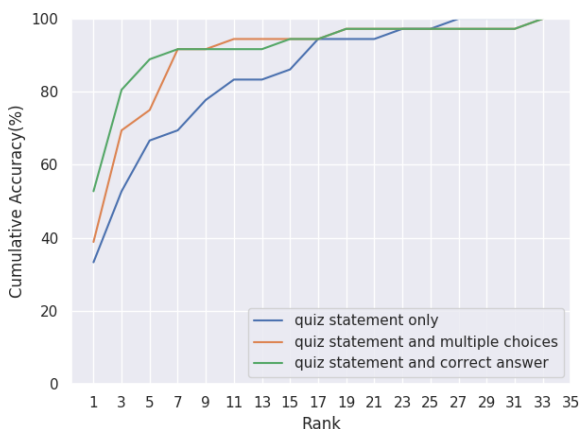


図3: 実験1の結果

率をさらに上げられるのではないかと考えられる。

3.2 実験2: 実際の講義でのアンケートに向けた事前実験

3.2.1 実験内容

実験1の結果を踏まえて、実験1の3)の手法を用いた場合のシステムの精度についての評価実験を学部1年の学生6人を対象として行った。実験2では「情報科学」の教材と小テストを用いた。あらかじめ、「情報科学」の講義資料から「情報科学」の6回分の講義資料(「D-01-アンケート」, 「D-03-ベクトル・距離・類似度」, 「D-02-検定・相関」, 「D-05-回帰・時系列」, 「D-06-画像」および「D-07-可視化」)とそれぞれの資料に基づいた小テスト各3問ずつの計18問から、各設問ごとに関連度の高い上位6ページをシステ



図4: レビュー画面

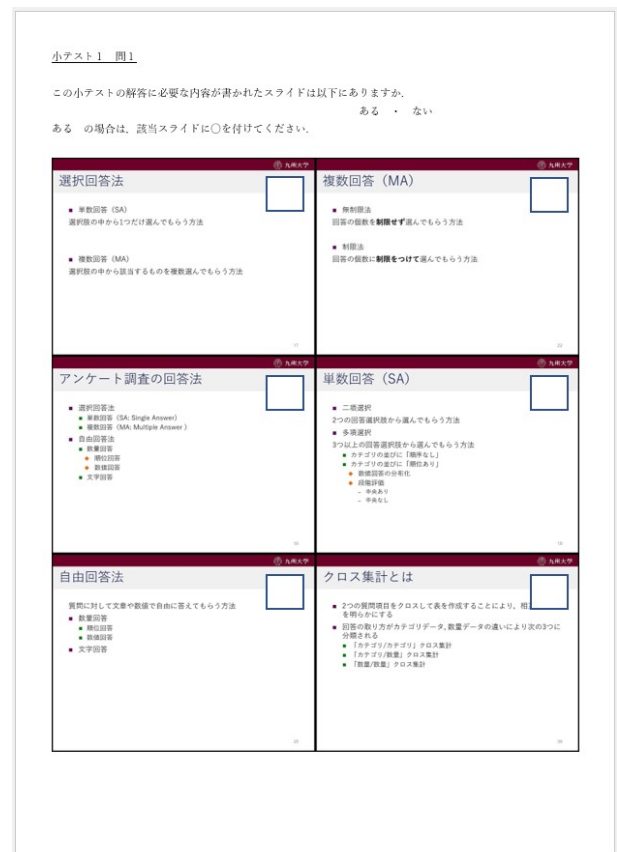


図5: アンケート用紙

ムによって抽出した。その後学生 6 人に小テストを受けてもらい、抽出されたページを見て小テストを解くために必要な知識が書かれた適切なページだと判断できるものを選択するよう指示した。小テストは moodle 上で受けてもらい、そのレビュー画面 (図 4) と図 5 のような A4 の用紙を見ながら用紙の関連ページに○をつけてもらう形で実施した。左上, 右上, 中段左, 中段右, 左下, 右下の並びで関連度の高い順にページを配置した。ただし, 先入観を排除するため, 配置の順番は学生 6 人には伝えていない。適切なページが複数ある場合はすべて選択, 適切なページがないと思われる場合は「なし」という項目を選択することとした。また, 実際の講義内での実験に向けて, 「表示されていた関連ページの枚数 (6 枚) は適切でしたか?」というアンケートにも回答してもらい, 「多い」または「少ない」を選択した場合には, 適切だと思われるページ数を記入してもらった。

3.2.2 実験結果

表 1, 表 2 に結果を示す。表 1 の各選択肢は図 5 の用紙の 6 枚の関連ページを表しており, 関連度の高い順に 1 枚目~6 枚目としている。表内の数値は 6 人のうち何人が各ページを設問に関連するページと回答したかを示している。複数選択可のため各設問ごとの人数の合計が 6 人となっていないことに注意されたい。表 1 の関連ページ選択の結果は, 半数である 3 人以上が「なし」を選択している設問が 18 問中 8 問あり, 平均して 2.39 人が「なし」を選択するという精度の悪さを示す結果となった。ただし実験 1 でも述べたように, 解答に必要な内容が教材のどのページにも載っていない設問も存在する可能性があるが, 実際の講義内での運用を見据えて関連ページをあらかじめ人手で抽出せずにすべての設問を用いて実験を行ったため, そのような関連ページがない問題によって精度が悪くなっている可能性がある。

表 2 の表示するページ数の評価結果は, 6 人中 5 人が「多い」を選択するという結果となった。適切だと思われるページ数として, 5 人のうち 2 人が「4 ページ」, 1 人が「3~4 ページ」, 2 人が「3 ページ」と記述していた。これらの意見から, 表示するページ数は 3 ページまたは 4 ページが適当であると考えられる。また, 表示するページ数が少なくなればなるほどその中に正解ページが入っている確率は低くなるため, 正解ページを表示できる可能性を高めるためには, 表示ページ数は多いほうがよい。以上を踏まえて, 表示するページ数の目標を 4 ページとした。

3.3 実験 3: 正答のテキストデータに対する重みづけ

3.3.1 実験内容

実験 1 において, システムの入力に設問文のテキスト

データのみを用いるよりも設問文と正解の選択肢のテキストデータを連結して入力として与えた方が良い結果が得られた。しかし, 実験 2 で実験 1 の 3) の手法を用いて「情報科学」の講義資料と小テストに対してシステムの精度を調べたが, 1 問あたり平均して 2.39 人が 6 枚の中に関連するページがないという項目を選択するという精度が高いとは言えない結果となった。また, 実際に小テストの設問文および選択肢と教材を確認したところ, 解答に必要な内容が書かれたページには, 正解の選択肢に含まれる単語がそのまま書かれている場合が多く見受けられた。そこで, 単に設問文と正解の選択肢のテキストデータを連結するのではなく, 設問文のテキストデータと正解の選択肢のテキストデータをそれぞれ 4.2 の手法でベクトル化し, それぞれのベクトルに重みづけしてどの程度正解の選択肢を重要視すればよいかを検証することにした。実験 1 と同様の講義資料および小テストを用い, 設問文のベクトルを v_s , 正解選択肢のベクトルを v_a , 重みを β とし, 式 (3) のように重みづけ平均を取る。計算結果のベクトル v_{in} を入力に用い, β を 0 から 1 まで 0.05 刻みで動かして出力の平均ランクを算出した。平均ランクとは, ページ番号を関連度順に並べたリストの中で, 正解ページが上から何番目にランクインしたかという順位的全設問での平均を取った値で, その値が小さいほど精度が高いことを表す。設問 k の正解ページのページ番号を a_k , 設問 k についての出力ランキングのリストを $r_k = (r_1, r_2, \dots, r_m)$ (m は教材のページ数), 設問 k の正解ページが r_k の何番目にランクインしたかというランクを b_k とし, 平均ランクを \bar{b} とすると, b_k および \bar{b} は式 (4), (5) のように表される。

$$v_{in} = \beta \cdot v_s + (1 - \beta)v_a \quad (3)$$

$$b_k = l \text{ (ただし } a_k = r_l) \quad (4)$$

$$\bar{b} = \frac{\sum_{k=1}^n b_k}{n} \quad (5)$$

3.3.2 実験結果

表 3 に結果を示す。 $\beta = 0.3$ のときに一番平均ランクが小さい, すなわち精度が高いことが表から分かる。 $\beta = 0.3$ のとき, 設問文のベクトルと正解選択肢のベクトルを 3:7 で重みづけして平均していることになり, 正解選択肢のテキストデータを重要視した方が良い結果が得られるということが分かった。もちろん, 講義資料や小テストの異なる講義ごとにこの重みの比は異なる可能性があるが, おおよそ $\beta = 0.2 \sim 0.4$ 程度が良いのではないかと考えられる。

表 1: 関連ページ選択の結果 (実験 2)

	1 枚目	2 枚目	3 枚目	4 枚目	5 枚目	6 枚目	なし
全設問の平均	2.22	1.28	0.50	0.61	0.06	0.17	2.39 (人)

表 2: 表示ページ数の適切性 (実験 2)

多い	適切	少ない	合計
5	1	0	6 (人)

表 3: 正答に重みづけしたときの平均ランク

β	0	0.05	0.1	0.15	0.2	0.25	0.3
平均ランク	4.14	3.92	3.81	3.75	3.58	3.36	3.31
β	0.35	0.4	0.45	0.5	0.55	0.6	0.65
平均ランク	3.36	3.42	3.39	3.56	3.75	3.89	3.75
β	0.7	0.75	0.8	0.85	0.9	0.95	1.0
平均ランク	3.92	4.14	4.33	4.69	5.14	5.5	6.08

3.4 実験 4: 実際の講義内でのアンケート

3.4.1 実験内容

実験 1~3 の結果を踏まえて、九州大学 1 年生向けの基幹教育科目「情報科学」2 クラスの講義内で、小テスト後のレビュー画面で各設問の正答に関連するものとしてシステムによって出力された上位 4 ページを図 6 のように表示し、表示されたページについて、実験 2 と同じように 1~4 の何番目に適切なページが出力されているか、または適切なページがその中にないかというアンケートを実施した。また、レビュー画面に実際に表示されることについて、成績の向上に役立ちそうか、表示されるページは適切か(多いか、少ないか)といった内容のアンケートも同時に行なった。アンケートは、実験 2 のときと異なり moodle 上のアンケートフォームを用いて行った(図 7,8)。各クラスとも全 15 回の講義のうちの第 13 回、第 14 回の 2 週分の講義でアンケート調査を行った。表示するページ数(4 ページ)は実験 3 の結果を踏まえて決定した。アンケートの回答者数は、1 週目は各クラス 130 人と 72 人の計 202 人、2 週目は各クラス 113 人と 74 人の計 187 人である。実験 4 の正解選択肢に対する重みづけの実験によって得られた結果から、2 番目に精度が良かった $\beta = 0.25$ の重みづけを用いて関連ページを抽出した。2 番目によりパラメータとしたのは、「成績の向上に役立ちそうか」というアンケート内容を踏まえ、最も良いパフォーマンスの際の有効性を検証するよりも多少精度が悪い場合にも有効であるかどうかを検証することの方が有意義だと考えたためである。1 週目(第 13 回)の講義では情報科学の講義資料から「D-01-アンケート」、「D-02-検定・相関」の 2 単元を、2 週目(第 14 回)の講義では「D-05-帰帰・時系列」、「D-07-可視化」の 2 単元を扱っており、各回ごとに授業終わりに 6 問の小テストが用意されていた。ただし実験 3 で用いた講義資料と講義名



図 6: レビュー画面での関連ページ表示

は同じ「情報科学」であるが、担当教員の異なる講義のため講義資料の中身は一部異なる。それらの教材と小テストの設問文をシステムに入力し、出力結果を用いてアンケートを作成した。

アンケート内容は、図 7, 図 8 にあるとおり、「以下の(システムが出力した上位)4 ページの中に、解答に必要な知識が書かれたページはあると思いますか?ある場合はそのページを(複数ある場合はすべて)選んでください。ない場合は「なし」を選んでください。」というシステムの精度についての質問を小テストの各設問ごとに 6 問、「実際に小テストのレビュー画面に関連ページが表示されることは、成績の向上に役立つと思いますか?」、「表示されていた関連ページの枚数(4 枚)は適切でしたか?」、「その他、関連ページ自動抽出について何か気づいた点があればご自由にお書きください。」という全体的な質問を 3 問の計 9 問を講義 2 回の各回ごとに用意した。

3.4.2 実験結果

表 4, 表 5, 表 6 にアンケート結果を示す。

表 4 の各設問におけるシステムの精度についてのアンケートは、学生全体の何%がアンケート選択肢の各ページを設問に関連するページと回答したかを計算しており、複数選択可のため割合の合計が 100%となっていないことに注意されたい。例えば「小テスト 1 問 1」の「1 枚目」の項目が 64%となっているのは、アンケートに回答した全学

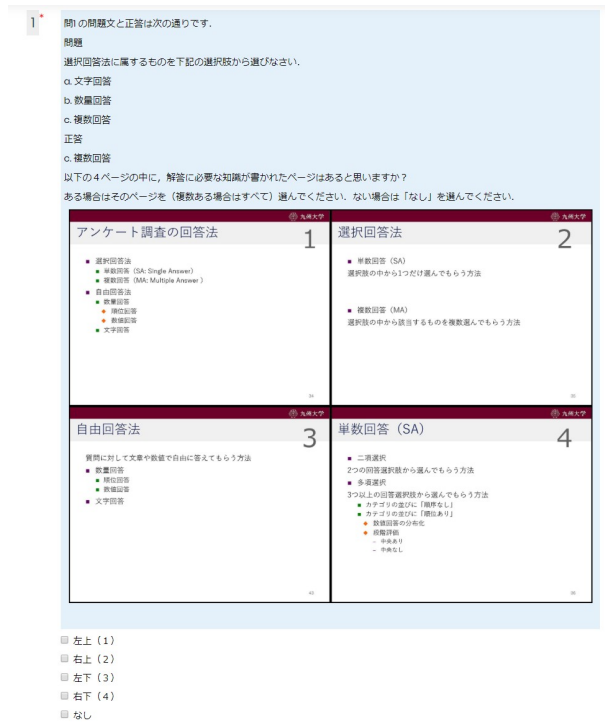


図 7: アンケート画面 1

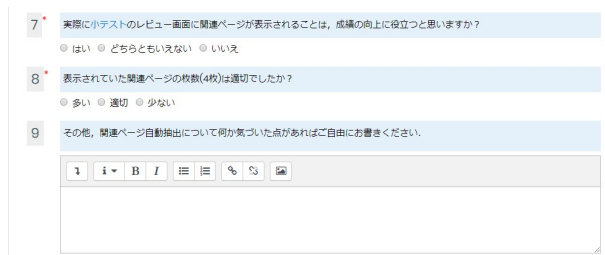


図 8: アンケート画面 2

生のうち 64%が、その設問でシステムが 1 番目に出力したページについて、その設問に関連のあるページだと判断したということを示している。テスト 1(1 週目の小テスト)では、おおむね上位に出力されたページほど多くの学生が関連があるページだと回答しており、平均して 9 割以上の学生が「なし」以外、すなわち 1~4 枚目の少なくとも 1 つに票を入れていることがわかる。一方テスト 2(2 週目の小テスト)では、問 1 と問 4 で 4 割以上の学生が「なし」を選択しているなどシステムの精度が悪かった設問があることが分かり、「なし」以外に票を入れているのは平均して 7 割弱という結果になった。全体としては平均して 8 割程度の学生が「なし」以外に票を入れており、システムが関連ページを抽出できる確率は 8 割程度だとこの実験から推定できる。

表 5 の成績向上に対する効果が期待できるかという問いに対する結果を見ると、テスト 1, 2 ともに 8 割以上の学生が「はい(役に立つ)」を選択しており、学生目線でよい効果が得られると感じられるシステムであることが分かる。

表 4: 関連ページ選択の結果(実験 4)

	1 枚目	2 枚目	3 枚目	4 枚目	なし
テスト 1 問 1	63.9%	55.1%	21.0%	11.2%	3.4%
テスト 1 問 2	35.6%	71.2%	21.5%	29.8%	4.4%
テスト 1 問 3	73.2%	21.5%	26.8%	14.1%	9.8%
テスト 1 問 4	31.7%	51.2%	34.1%	29.8%	12.7%
テスト 1 問 5	18.0%	20.5%	75.1%	17.1%	9.8%
テスト 1 問 6	67.3%	41.0%	19.5%	11.7%	10.2%
テスト 2 問 1	28.2%	27.1%	19.7%	16.0%	50.0%
テスト 2 問 2	25.5%	59.6%	30.3%	13.3%	15.4%
テスト 2 問 3	25.0%	21.8%	36.7%	19.1%	38.3%
テスト 2 問 4	36.2%	19.1%	19.1%	33.0%	34.6%
テスト 2 問 5	12.8%	75.0%	21.8%	19.7%	9.0%
テスト 2 問 6	26.1%	44.1%	26.1%	12.8%	39.4%
平均(テスト 1)	48.3%	43.4%	33.0%	18.9%	8.4%
平均(テスト 2)	25.6%	41.1%	25.6%	19.0%	31.1%
平均(全体)	37.0%	42.3%	29.3%	19.0%	19.7%

表 5: 成績向上に対する効果(実験 4)

	はい (役に立つ)	どちらとも いえない	いいえ (役に立たない)
テスト 1	83.9%	10.7%	5.4%
テスト 2	76.6%	19.7%	3.7%
平均	80.2%	15.2%	4.5%

表 6: 表示ページ数の適切性(実験 4)

	多い	適切	少ない
テスト 1	31.7%	68.8%	1.0%
テスト 2	27.4%	67.3%	5.3%
平均	29.6%	68.0%	3.1%

特に、あまり精度の良くなかったテスト 2 においても、小テスト 1 のときとあまり変わらず 8 割以上の学生が役に立つと感じており、適切な関連ページが表示されていない設問があったとしても少なからず役に立つのでは、と考える学生が存在することが読み取れる。

表 6 の表示するページ数については、67%程度の学生が「適切」、30%程度の学生が「多い」を選択していることがわかる。多いと感じる学生も少なくはないが、表示するページ数を減らすとその分適切なページを表示できる確率が下がってしまうため、システムの精度も考慮すると 4 枚という表示枚数は適切といえるのではないかと考えられる。

また、「その他、関連ページ自動抽出について何か気づいた点があればご自由にお書きください。」という問いに対する回答としては、「関連ページをつけてくれた方が復習しやすいのでぜひつけてほしい」、「他の授業でも活用してほしい」といった前向きな意見や、「(スライドの画像の中で)特に重要なところをマーキングしてあるとより分かりやすい」、「不正解の選択肢に関連するページも見せた方が分かりやすいのでは」、「ページの画像のみでなく文章による説

明も欲しい」といった改善点を挙げる意見が見られた。

「(スライドの画像の中で)特に重要なところをマーキングしてあるとより分かりやすい」という意見に関しては、word2vecを用いて設問文の内容に近い単語をページ内から探すなどの手法で実現可能と考えられるため、今後実際に重要語のマーキングを実装し評価実験を行うことによってシステムの改善を図れる可能性がある。「不正解の選択肢に関連するページも見せた方が分かりやすいのでは」という意見については、技術的には本研究で作成したシステムにより可能であるが、不正解の選択肢の数に応じて抽出するページの枚数が増えてしまうという問題点があるため、表示するページの枚数を増やしてでも不正解に関連するページを表示する方が良いかどうかは検証の余地がある。「ページの画像のみでなく文章による説明も欲しい」という意見に関しては、重要語のマーキングと同じように、重要語を含む1文1文を関連ページから抽出することは可能だと考えられるが、文章の自動生成となるとその技術自体が大きな研究テーマになってしまうため現段階では実装は難しい。しかし技術が進めばキーワードからの文章の自動生成も可能になることが期待できるため、課題として取り組む価値があると考えられる。

4. まとめ

本研究では、BookRoll上のデジタル教材およびmoodle上の小テストのテキストデータを分析し、小テストの各設問を解くために理解すべき内容が記載されているページを抽出するシステムを作成した。その後4つの実験を通して、システムが抽出するページの妥当性や、小テスト後のレビュー画面に関連ページを表示することの意義を学生からの評価によって検証した。小テストの中には、更新された最新の講義資料には関連ページが存在しないような設問もあったが、そのような設問も含め平均して8割程度の確率でシステムが適切な関連ページを上位4ページ以内に抽出できるという結果が得られた。システムの入力としては、サイバーセキュリティの講義資料および小テストの分析結果から、設問文のテキストデータと正解選択肢のテキストデータのそれぞれをベクトル化し3:7程度の割合で加重平均を取ると精度よく関連ページを抽出できることが分かり、他の講義についても同程度の割合の加重平均により最適化できるのではないかと推察された。

今後の課題としては、3.4節で述べたように、関連ページ内の重要語のマーキング、不正解選択肢に関連するページの表示、文章による解説の作成など、学生からの意見を取り入れて改善できる点が存在する。また、本研究では関連ページをレビュー画面上に画像として表示することで学生に提示したが、別の手段として、レビュー画面にBookRoll内の関連ページへのリンクを貼るという方法も考えられる。この手段を用いると、学生が復習時にそのリンクにアクセ

スしたときにBookRollの閲覧履歴をログとして残せるという点や、BookRoll上で関連ページの周辺のページも容易に閲覧できるといった点で大きなメリットがある。これらの改善案を検討しフィードバックを繰り返すことで、学生の復習活動を支援する仕組みを今後も構築、改善していく。

謝辞 本研究は、JST AIP 加速課題 JPMJCR19U1、科研費基盤研究(A) JP18H04125、基盤研究(B) 19H01716の支援を受けた。

参考文献

- [1] Rebecca Ferguson. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, Vol. 4, No. 5/6, pp. 304–317, 2012.
- [2] Martin Dougiamas. Moodle: A virtual learning environment for the rest of us. *TESL-EJ*, Vol. 8, No. 2, pp. 1–8, 2004.
- [3] Hiroaki Ogata, Chengjiu Yin, Misato Oi, Fumiya Okubo, Atsushi Shimada, Kentaro Kojima, and Masanori Yamada. E-book-based learning analytics in university education. In *International Conference on Computer in Education (ICCE 2015)*, pp. 401–406, 2015.
- [4] Jack A Chambers and Jerry W Sprecher. *Computer-assisted instruction: Its use in the classroom*. Prentice Hall Direct, 1983.
- [5] Mei-Yu Chang, Wernhuar Tarng, and Fu-Yu Shin. The effectiveness of scaffolding in a web-based, adaptive learning system. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, Vol. 4, No. 1, pp. 1–15, 2009.
- [6] Carla Limongelli, Filippo Sciarone, and Giulia Vaste. Ls-plan: An effective combination of dynamic courseware generation and learning styles in web-based education. In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pp. 133–142. Springer, 2008.
- [7] Juan Ramos, et al. Using tf-idf to determine word relevance in document queries. In *Proceedings of the first instructional conference on machine learning*, Vol. 242, pp. 133–142. Piscataway, NJ, 2003.
- [8] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pp. 3111–3119, 2013.
- [9] Masatoshi Suzuki, Koji Matsuda, Satoshi Sekine, Naoaki Okazaki, and Kentaro Inui. A joint neural model for fine-grained named entity classification of wikipedia articles. *IE-ICE Transactions on Information and Systems*, Vol. 101, No. 1, pp. 73–81, 2018.