

# シラバスデータを活用した用語からの科目推薦手法の検討

橋本菜々子<sup>\*1</sup>, 越智洋司<sup>\*2\*3</sup>

<sup>\*1</sup> 近畿大学大学院総合理工学研究科, <sup>\*2</sup> 近畿大学情報学部, <sup>\*3</sup> 近畿大学情報学研究所

## Study of a method for recommending courses from terminology using syllabus data

Nanako Hashimoto<sup>\*1</sup>, Youji Ochi<sup>\*2\*3</sup>

<sup>\*1</sup> Graduate School of Science and Engineering, Kindai University

<sup>\*2</sup> Faculty of Informatics, Kindai University

<sup>\*3</sup> Cyber Informatics Research Institute, Kindai University

大学などの教育機関では履修登録の際にシラバスが活用されていることが多い。なぜならシラバス内には、授業名・担当教員・到達目標・成績評価基準・各回の授業計画などが記載されており、履修登録の判断材料となるためである。しかし学びたい専門分野に対する科目を選定するためには、大量のシラバス情報の中から判別する必要がある。そこで本研究ではシラバスデータを機械学習により学習させ、科目推薦を提案することによる教育支援を行い、手法の妥当性の評価を行った。また今後の研究として科目の順序性についての検討も行う。

キーワード: 機械学習, 自然言語処理, シラバスデータ, 推薦科目, 情報推薦

### 1. はじめに

近年、多くの大学でシラバスを用いた履修登録が行われており、卒業単位を獲得するために重要な役割を持っている。またシラバスデータを Web 上で公開している大学がほとんどであり、学生はいつでもシラバス内の科目を容易に検索・閲覧することが出来る。しかし、シラバス内の科目は多数存在し、自分がどの科目を履修すればいいか判断することが出来ないといった問題点がある。具体的には、自分が学びたいと思っている分野はどの科目を履修すればいいのかが分からないという問題点である。そこで本研究では、学びたい分野に対する「推薦科目」を提案することにより、学生への教育支援を行う。

### 2. 導入

#### 2.1 先行研究

これまでもシラバスデータを活用した研究が行われてきた。

吉崎ら<sup>(1)</sup>は科目間の類似性の分析を目的として、Doc2Vec を用いたシラバス内における科目間の関連性の分析を行い、ユーザーが容易に目的となる講義を探し出せるシステムの提案を行った。

竹森ら<sup>(2)</sup>は、Doc2Vec によって獲得した科目情報の分散表現についてクラスタリングを行い、各クラスタに属する科目群の特徴を可視化し、ユーザーに興味のあるものを選択してもらい、必要に応じてキーワードを入力してもらうことでユーザーの学びの志向とマッチする科目を選択しやすいように支援を行うシステムを提案した。

藤本ら<sup>(3)</sup>は時間割を組む上で一番効率の良い履修コースを求めることを目的とし、各研究室に配属されている学生が履修した科目の履修データや、学生の主観的調査に基づくカリキュラム選定を行うシステムを提案した。

宮脇ら<sup>(4)</sup>は、学生が履修すべき科目を適切に選択できるように補助することを目的とし、TF-IDF と sentence2vec の 2 つの手法により C プラン（学生が



乾・岡崎研究室<sup>(7)</sup>によって作られた日本語版 Wikipedia の本文全文を学習させたものを使用した。表 1 に具体例を挙げる。例として、ある科目データに「導関数」という単語があったとする。Word2Vec モデルを用いるとその単語の類似単語 10 個を提案することが出来るため、その単語を科目の学習データとして追加する。

表 1 該当単語に対する類似単語

「導関数」に対する類似単語	
1 常微分方程式	6 内積
2 テンソル	7 多項式
3 外積代数	8 運動方程式
4 微分方程式	9 微分
5 テンソル場	10 漸化式

## 4. 評価実験

### 4.1 推薦科目の評価

本研究では、シラバス内の科目に含まれている文書ベクトルと検索する用語ベクトルの類似度計算を行い、類似度が高い科目を推薦科目として導出することができると仮定し、推薦科目の妥当性について評価を行う。ユーザーが検索する用語は近畿大学工学部電気電子工学科 2021 年度版の卒業論文予稿で用いられている用語を使用し、類似度の高い上位 3 科目を推薦科目として提案を行った。また、類似度の低い下位 3 科目に対しても同様に推薦科目として提案を行い、次節で示す手法によって妥当性の評価を行った。

### 4.2 評価手法

実験で用いた用語は、前述の通り近畿大学工学部電気電子工学科の論文中から選択しているため、その論文を作成した研究室の担当教員に「この研究を行うために必要な知識は、提示している科目に関係があるか?」というアンケートを 4 段階評価で行った。評価基準を表 2 に示す。また、1,2 の評価は「履修を求め」、3,4 の評価は「履修を求めない」と 2 値分類を行い、次節で評価分析を行う。

### 4.3 評価結果

本研究では、研究室に配属されている担当教員の先生方に協力していただいた。その結果を表 2 に示す。

また、2 値分類による推薦科目評価の再現率・適合率・F 値の計算を表 3 に示す。

表 2 推薦科目の評価基準

	評価基準
1	関係がある (履修を求める)
2	どちらかといえば関係がある (履修したほうが良い)
3	どちらかといえば関係がない (履修しなくてもよい)
4	関係がない (履修する必要がない)

### 4.4 評価分析

表 3 より、すべての組み合わせにおいて  $0.514 \pm 0.047$  程度の精度であった。宮脇らの研究では、sentence2vec を使用した推薦科目提案で適合率 35.2%であったが、本研究で行った評価では適合率は最大 56.7%であり、提案手法の有効性が示されたと考える。しかし、竹森らの研究による評価では適合率 64%であり、それより高い精度には至らなかった。精度向上のためにはデータ量の増加が必須であると考えられる。表 4 に今回シラバス内の項目ごとの単語量を示す。機械学習の中でも Doc2Vec のような教師なし学習による特徴量抽出タスクで重要になるのがデータであり、その質・量によって生成されるベクトルも変化すると考えられる。今回はモデル自体に「偏り」を表現するために 1 つの学科データのみを学習データに使用したが、今後は全学部・全学科のデータを使用した汎用性の高いモデルによる推薦科目提案を行いたい。

表 3 推薦科目の評価結果

	上位 3 科目		下位 3 科目	
	1,2	3,4	1,2	3,4
授業概要	<b>51</b>	39	56	34
授業計画	42	48	35	55
授業概要・計画	42	48	32	<b>58</b>

表 4 2 値分類による評価結果

	再現率	適合率	F 値
授業概要	0.477	<b>0.567</b>	<b>0.518</b>
授業計画	0.545	0.467	0.503
授業概要・計画	<b>0.568</b>	0.467	0.512

表 5 各項目の単語数

	最大値	最小値	中央値
授業概要	1965	212	836
授業計画	4325	1171	2673
授業概要・計画	5966	374	2404

## 5. 今後の展望

今後の展望として、科目の順序を提案するシステムの構築を行っていく。本研究では教育支援を目的として行ってきたが、教育の重要な1つのポイントとして学習順序があるためである。そこで科目の順序性を学習できるようにすれば、さらなる教育支援が可能であると思われる。順序性の学習として Yongjing ら<sup>(8)</sup>は、文の意味表現を正確に学習するためにグラフィカルネットワークを採用した、新規かつ柔軟なグラフベースの文整列モデルを提案した。また Golestani ら<sup>(9)</sup>は、文の依存関係をとらえるユニバーサルトランスフォーマーに基づく言語モデルを提案した。このような「文」の一貫性を学習する研究から派生し、「文章」を学習させる研究が可能であると考えられる。

## 6. まとめ

本稿では、教育支援を目的とした推薦科目の提案と、今後の展望として科目の順序性についての検討を述べた。推薦科目の提案では、再現率・適合率・F 値による精度が 0.5 程度となり他の推薦科目提案研究と比較してこの手法の有用性が考えられる。また、さらに精度を上げる方法としてデータの質・量の向上が挙げられる。科目の順序に関する検討では、文の一貫性をモデル化する手法から応用することが出来る可能性を述べた。科目の順序性を考慮した手法が有用であると示すことが出来れば、さらなる教育支援に繋がると考えられる。具体的には「学びたい知識を得るためにはこの科目 A を学んだ後に、科目 B を勉強すれば良い」と提示することが出来れば、ユーザーは気軽に相談や支援を受けることが可能になると考える。

## 謝辞

本研究は JSPS 科研費 22K12304 の助成を受けた。

## 参考文献

- (1) 吉崎辰悟, 越智洋司: “シラバスデータを利用した科目間の関連性の分析”, 日本教育工学会研究報告集, Vol.18, No.5, pp9-14 (2018)
- (2) 竹森汰智, 亀井清華: “履修支援のための Doc2Vec を用いた科目推薦システム”, 情報処理学会誌, Vol.12, No.4, pp1-14 (2019)
- (3) 藤森高志, 平石邦彦, 三浦元喜, 國藤進: “研究室配属のためのシラバスシステムの提案と試作”, 情報処理学会誌, Vol.34, No.1, pp67-72 (2006)
- (4) 宮脇克典, 池田雄斗, 福本加奈恵, 水野創太, 白松俊: “科目区分ダイアグラム検索システムにおけるテキスト類似度に基づく科目推薦機構の試作”, 人工知能学会全国大会論文集, pp1-3 (2017)
- (5) Tomas Mikolov, Quoc V. Le: “Distributed Representations of Sentences and Documents”, IMCL, pp1188-1196 (2014)
- (6) MeCab, <http://taku910.github.io/mecab/>, (2022年6月11日確認)
- (7) 鈴木正敏, 松田耕史, 関根聡, 岡崎直観, 乾健太郎: “Wikipedia 記事に対する拡張固有表現ラベルの多重付与”, 第 22 回言語処理学会, pp797-800 (2016)
- (8) Yongjing Yin, Linfeng Song, Jinsong Su, Jiali Zeng, Chulun Zhou, Jiebo Luo: “Graph-based Neural Sentence Ordering”, IJCAI (2019)
- (9) Melika Golestani, Seyedeh Zahra Razavi, Zeinab Borhanifard, Farnaz Tahmasebian, Hesham Faili: “Using BERT Encoding and Sentence-Level Language Model for Sentence Ordering”, The 24<sup>th</sup> International Conference of Text, Speech and Dialogue (2021)