

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

卒業生の就職先データを活用したキャリアパス指向の履修科目支援システム

Development of a Recommendation and Learning Support System Based on Teaching and Learning Data Considering Review Reliability for GPA Improvement

山本 藤也 (Touya Yamamoto)
t815043@st.pu-toyama.ac.jp

富山県立大学 工学部 情報システム工学科

Teams, 13:50-14:05 Wednesday, February 16, 2022.

1.1 研究の背景

2/20

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

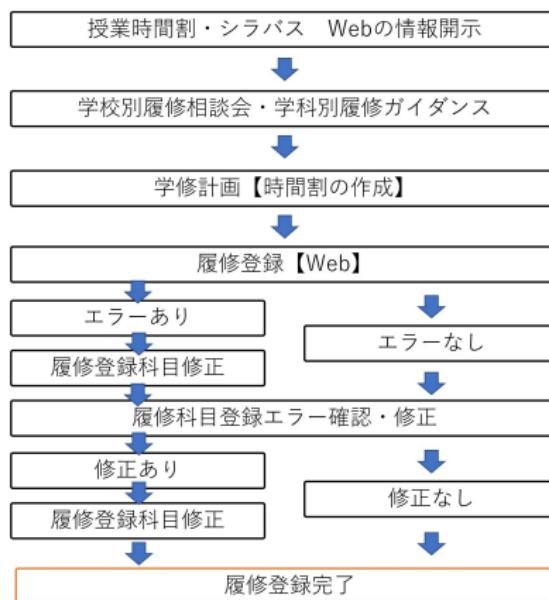


図1: 履修登録までの流れ

表2: 成績からGPへ変換

評価	数字	点数	合否
S	4	90点以上	合格
A	3	80点以上 90点未満	
B	2	70点以上 80点未満	
C	1	60点以上 70点未満	
不可	0	60点未満	不合格
空白		履修なし	

GPAの計算方法

$$GPA = \frac{(S\text{の数}) \times 4 + (A\text{の数}) \times 3 + (B\text{の数}) \times 2 + (C\text{の数}) \times 1}{\text{履修した単位数}}$$

1.2 研究の目的

3/20

過去の卒業生の教学データに対して分析を行い、成績評価の予測を行い、高いGPAを獲得できるように取得単位を推薦するとともに、全ての科目について良い成績がとれるようにWeb上から関連情報を推薦するようなシステムの開発を行う。

- はじめに
- 教学データ分析と情報推薦
- 授業科目の成績予測と教材の最適化
- 予測評価値からの適切な授業科目の推薦
- 提案手法
- 数値実験並びに考察
- おわりに

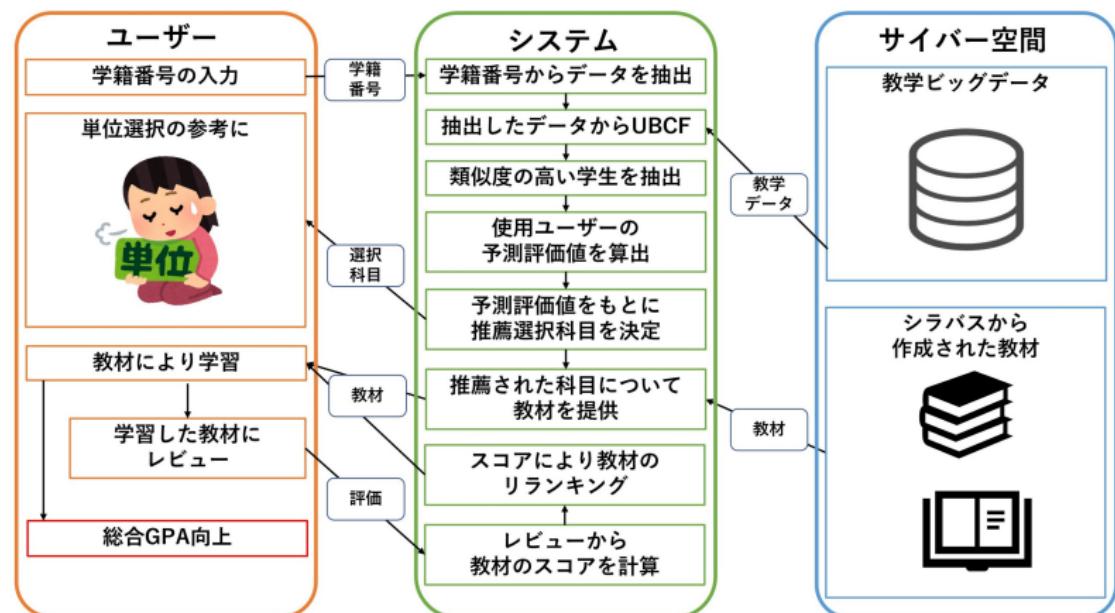


図2: 提案システムの流れ

2.1 教学におけるビッグデータ・アナリティクス

4/20

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

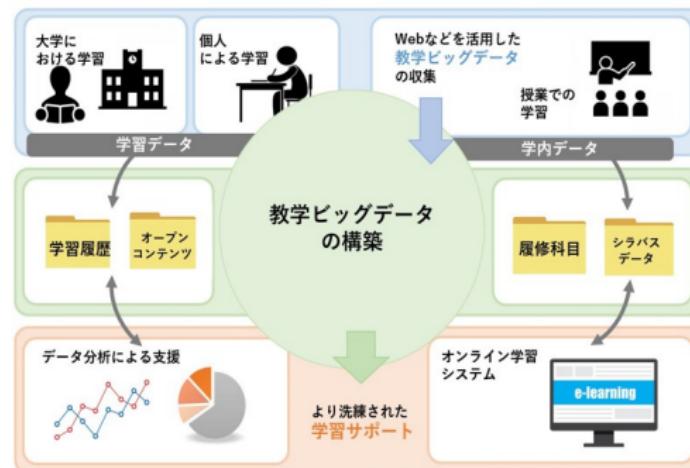


表2: 教学データの例

取得時期	教学データ	内容
入学前	出身高校	課程差別, 判定値, etc.
	入試情報	入試区分, 成績
	入学前学習	取組状況, 提出物
入学時	導入教育	オリエン欠席, テスト結果, etc.
各セメスター	履修登録	履修科目
	授業	出席状況, 課題提出, etc.
	学生生活	部活, アルバイト, etc.
	成績	科目成績, GPA, etc.
4年次	就職活動	活動履歴, 内定状況
卒業後	卒業後	満足度, アンケート

6の特徴

1. データ量はそれほど大きくない
2. 対象人数は少ないがデータの種類が急増している
3. 匿名性が低い
4. データの意味解釈が容易
5. 因果関係は複雑
6. 多様化・細分化が進行している

¹松田岳士, 渡辺雄貴, “教学 IR, ラーニング・アナリティクス, 教育工学”, 日本教育工学会論文誌, Vol. 41, No. 3, pp. 199-208, 2017.

2.2 情報推薦と協調フィルタリング

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

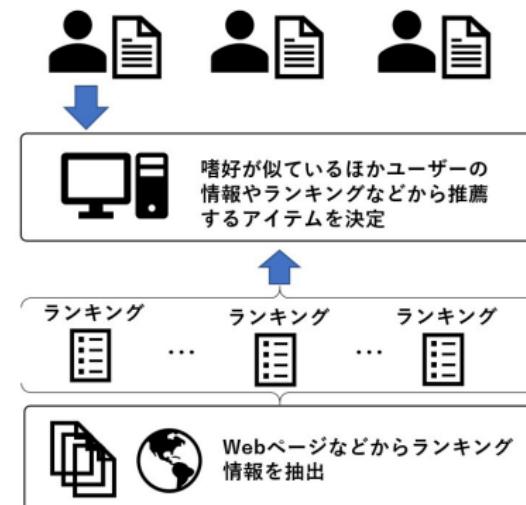


図4: 情報推薦

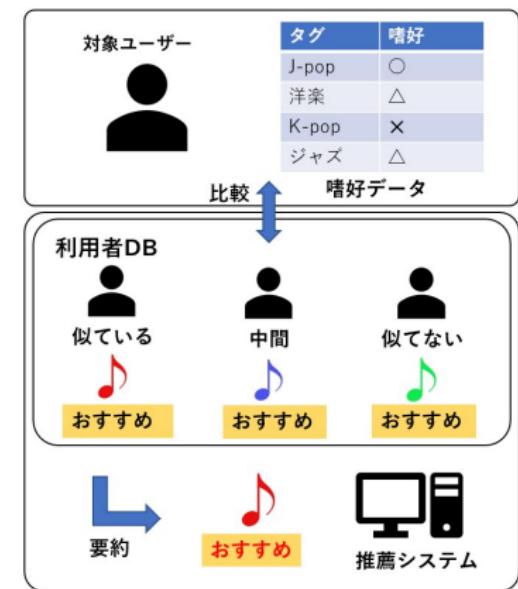


図5: 協調フィルタリング

²神島敏弘, “推薦システムのアルゴリズム”, <https://www.kamishima.net/archive/recsysdoc.pdf>, 2021.10.28

2.3 レビューの信頼性の判断支援

6/20

- はじめに
- 教学データ分析と情報推薦
- 授業科目の成績予測と教材の最適化
- 予測評価値からの適切な授業科目の推薦
- 提案手法
- 数値実験並びに考察
- おわりに

AmazonなどのECサイトにおいてユーザーから商品へのレビューは重要な役割を担っている。レビューの価値を使用し偽のスパムレビューを投稿し商品の評価を不当に操作する行為が問題となっている。その解決のためにレビューが信頼できるものなのかを判断するための研究が行われている。



図6：国内の主なECサイト



図7：ECサイトにおけるレビュー



図8：スパムレビューの投稿
サクラグループ



図9：サクラグループによるスパムレビューの投稿

3.1 協調フィルタリングからの科目的推薦

7/20

ユーザーベース協調フィルタリング (UBCF) における利用者=学生、アイテム=科目、評価値=成績と置き換えて学生に対して成績を予測し、科目的推薦を行う。

1. はじめに

2. 教学データ分析
と情報推薦

3. 授業科目の成績
予測と教材の最
適化

4. 予測評価値から
の適切な授業科目
の推薦

5. 提案手法

6. 数値実験並びに
考察

7. おわりに

嗜好データにおけるUBCF

表3:評価値行列

	1: そば	2: ラーメン	3: うどん	4: パスタ
1: 高橋		2	3	2
2: 山本	1		3	
3: 伊藤	2		2	1
4: 吉田	1	3	3	2

吉田はうどんに評価3を与えて
いるのでうどんが好きだといえる

$$\rho_{ax} = \frac{\sum_{y \in \mathcal{Y}_{ax}} (r_{ay} - \bar{r}_a')(r_{xy} - \bar{r}_x')}{\sqrt{\sum_{y \in \mathcal{Y}_{ax}} (r_{ay} - \bar{r}_a')^2} \sqrt{\sum_{y \in \mathcal{Y}_{ax}} (r_{xy} - \bar{r}_x')^2}}$$

Pearson相関でユーザー同士の
嗜好の類似度を求める

$$\hat{r}_{ay} = \bar{r}_a + \frac{\sum_{x \in \mathcal{X}_y} \rho_{ax} (r_{xy} - \bar{r}_x')}{\sum_{x \in \mathcal{X}_y} |\rho_{ax}|}$$

評価してない食べ物を予測

教学データにおけるUBCF

表4:成績値行列

	1: 国語	2: 数学	3: 化学	4: 物理
1: 高橋	1	3		3
2: 山本		1	3	
3: 伊藤	2	1	3	1
4: 吉田	1	3	2	

高橋は数学で評価3を取得して
いるので数学が得意といえる

$$\rho_{ax} = \frac{\sum_{y \in \mathcal{Y}_{ax}} (r_{ay} - \bar{r}_a')(r_{xy} - \bar{r}_x')}{\sqrt{\sum_{y \in \mathcal{Y}_{ax}} (r_{ay} - \bar{r}_a')^2} \sqrt{\sum_{y \in \mathcal{Y}_{ax}} (r_{xy} - \bar{r}_x')^2}}$$

Pearson相関で学生同士の
成績の類似度を求める

$$\hat{r}_{ay} = \bar{r}_a + \frac{\sum_{x \in \mathcal{X}_y} \rho_{ax} (r_{xy} - \bar{r}_x')}{\sum_{x \in \mathcal{X}_y} |\rho_{ax}|}$$

取得していない科目を予測

予測成績が高い科目を
優先的に推薦

3.2 シラバスからの教材作成

8/20

1. はじめに

2. 教学データ分析 と情報推薦

3. 授業科目の成績 予測と教材の最 適化

4. 予測評価値から の適切な授業科目 の推薦

5. 提案手法

6. 数値実験並びに 考察

7. おわりに

学校で配布されているシラバスを活用し、Web 上の授業内容に関連した情報をホームページと YouTube から取得し、学生に提示する。このようにして作成された情報を教材と呼ぶ。

授業科目名	IoTシステムデザイン	
授業科目名（英語）	Design of IoT System	
科目区分	専門	
記述年数	工学科 3年	
担当教員		
姓	氏名	所属
本物語	◎ 本木 健嗣	情報システム工学科
開講学期	後期	
単位数	2	
単位区分	選択	
開講する学部・教育目標	電子・情報工学科 (3)	
授業の目標／授業概要	<p>IoTは、モノや環境がインターネットに接続され、その状態が把握できるシステムであり、これからのAIやデータ社会の基礎となる技術である。IoTで得られるデータを元にIoT（技術）だけを理解しては技術者としてではなく社会課題の解決に貢献することできなくなる。そのため本授業では、社会課題の発見から解決までの探し、そしてシステムの設計の基礎となるセンサ技術や通信技術について網羅的、集めたデータをどのように分析するかについて授業を行なう。</p>	
学生の到達目標	<p>IoTの概要と、IoTに関する基礎知識を理解する。 無線LAN技術の基礎を理解する。(2週) センサネットワークなどの無線技術の応用について理解する。 IoTどのようにして地域課題解決に貢献か、デザイン思考による問題解決手法を身につける。</p>	
授業計画	<p>IoTを実現するための主な技術の概要を理解する。(2週) 無線LAN技術の基礎を理解する。(2週) センサネットワークの基礎を理解する。(2週) IoTの概要と、IoTの応用について理解する。(2週) センサ、グローバル・センシングによる地域課題の発見 (2週) IoT地域課題解決するためのIoTシステムデザイン手法を学ぶ (4週)</p>	
キーワード	無線LAN、アドホックネットワーク、センサネットワーク、IoT	
成績評価基準	小テスト (30%)、プレゼンテーション (20%)、レポート (複数回 50%) によって総合評価する。	
教科書・教材参考書等	認可資料を用いて行う。参考書：別途指定する。	
開講科目・選修条件等	本講義の受講にあたっては、インターネット工学および情報伝送方式を理解しておくことが望ましい。	
履修上の注意事項や学習上の助言	講義中の質問は理解できなかった場合には必ず質問する。 資料では、実験結果のある授業資料または主として実験的教科書から構成される授業資料です。	
学生からの質問への対応方法	質問は基本的に回答する。また、定期の質問があれば毎回、文書どちらもでも随時受け付けるが、事前に電子メールで相談内容を連絡して予めて下さい。電子メールアドレス： imabot@tpu.tuhs.tuhs.ac.jp	

図10 : Webシラバス

①シラバスから授業計画をスクレイピング

- IoTを実現するための主要技術の概要を理解する。(3週)
- 無線LAN技術の基礎を理解する。(2週)
- センサネットワーク関連技術の基礎を理解する。(2週)
- IoT構成する技術を理解する。(2週)
- 調査、グループディスカッション等による地域課題の発掘 (2週)
- 地域課題を解決するためのIoTシステムデザイン手法を学ぶ (4週)

②スクレイピングの結果をGoogleとYouTubeで検索



③検索結果の上位を取得し教材に

図11 : 教材作成の流れ

Youtubetitel

- 【中学英語】40分で学べる中1英語全解説授業 <https://www.youtube.com/watch?v=ebQ8gcXsLg>
 【英語】中1-3一般動詞(基本編) <https://www.youtube.com/watch?v=iDftdgCcBiE>
 【高校英語】文法の基礎を総まとめ <https://www.youtube.com/watch?v=Npn-eQdd3s>
 【99%の人が勘違い】英文法20選【完全イメージ化】 <https://www.youtube.com/watch?v=vnk-UB18iZ4>
 HTitle
 HUrl
 名前の前に形容詞を並べる順番のルール【ラングラント】<https://www.langland.co.jp/english/column/english-column1/>
 誰でも英語の文法をマスターできる魔法の3ステップと <https://www.manabi-vitamin.net/jhs-english/>
 英語 文法 - 東京外国语大学 <http://www.cooling.tufs.ac.jp/mt/en/gmed/>
 英語の文法を学ぶための順番と確実に覚える方法 <https://www.alohaenglish.jp/english-grammar/>

図12 : 取得された教材

3.3 レビュー信頼性の指標

9/20

レビューの信頼性の指標として類似性、協調性、集中性、情報性の4つの指標を定義している³.

1. はじめに

2. 教学データ分析
と情報推薦

3. 授業科目の成績
予測と教材の最
適化

4. 予測評価値から
の適切な授業科目
の推薦

5. 提案手法

6. 数値実験並びに
考察

7. おわりに

類似性

・複製されたレビューには多くのスパムが含まれていることがわかっている
➡ レビューの文章の類似度を図る指標として類似性スコアを定義する。

bigramで区切ったレビュー l_i と l_j の類似度

$$\text{sim}(l_i, l_j) = \frac{|X_{l_i} \cap X_{l_j}|}{|X_{l_i} \cup X_{l_j}|}$$

列について最大値を類似度とする

$$S_{\text{score}}(l_i) = \max_{l_j} (\max(\text{sim}(l_i, l_j)) | j \neq i, j = 1, 2, \dots, n)$$

$n \times n$ の類似度行列を作成

	l_1	l_2	..	l_n
l_1	0	$\text{sim}(l_1, l_2)$	$\text{sim}(l_1, l_n)$	
l_2	$\text{sim}(l_2, l_1)$	0	$\text{sim}(l_2, l_n)$	
:				
l_n	$\text{sim}(l_n, l_1)$	$\text{sim}(l_n, l_2)$..	0

類似度をスコア0から5の範囲で正規化

$$S_{\text{score}}(l_i) = 5 \cdot S_{\text{core}}(l_i)$$

各行について最大値を求める

集中性

・スパムレビューは時間的に集中して投稿される傾向にある
➡ レビューが集中的に投稿されているかを測る指標として集中性スコアを定義する。



レビューが頻繁に投稿されている日についてレビューが投稿に投稿されている時間を求める

レビュー l_i の集中性スコアを算出
 $T_{\text{score}}(l_i) = \ln(\text{size}(g_b))$

集中性スコアをスコア0から5の範囲で正規化
 $T_{\text{score}}(l_i) = \frac{5 \cdot T_{\text{score}}(l_i)}{\max(T_{\text{score}}(l_j) | j = 1, 2, \dots, N)}$

レビューが頻繁に投稿されている時間についてその時間に投稿された
レビューの数 $\text{size}(g_b)$ を求める

情報性

・レビュー本文がinformativeであるほどスパムである可能性が低い
➡ レビューがinformativeであるかを測る指標として情報性スコアを定義する。

レビュー本文から名詞を抽出

今は夏。彼女はそれを悪い出す。 今、夏、彼女

抽出した名詞がほかのレビューで使用されている
回数をカウント

今 = 4回、夏 = 3回、彼女 = 1回の時
[今, 夏, 彼女] = [4, 3, 1]

情報性スコアを算出

$$I_{\text{score}}(l_i) = \ln \left(1 + \sum_{j=1}^{|K_i|} \ln \left(\frac{o}{df(term_j)} \right) \right)$$

記号

レビュー l_i と同じジャンルに属するレビュー l_j の数: o

レビュー l_i に出現する名詞集合: K_i
[今, 夏, 彼女]

レビュー l_i と同じジャンルのレビュー集合において
 $term_j \in K_i$ を含むレビューの数: $df(term_j)$
→ [4, 3, 1]

協調性

・サクラはグループを組んでスパムレビューを投稿する傾向にある
➡ レビューがサクラによって投稿されたかを測る指標として協調性スコアを定義する。



支持度数 $support(g_c)$ とユーザーID数 $\text{size}(g_c)$ で協調度を計算
 $collaborate(g_c) = support(g_c) \cdot \text{size}(g_c)$



レビュー l_i の協調性スコアを算出
 $C_{\text{score}}(l_i) = \begin{cases} \ln(\max_{g_c \in G_{l_i}} (collaborate(g_c))) & |G_{l_i}| \neq \emptyset \\ 0 & |G_{l_i}| = \emptyset \end{cases}$



協調性スコアをスコア0から5の範囲で正規化
 $C_{\text{score}}(l_i) = \frac{5 \cdot C_{\text{score}}(l_i)}{\max(C_{\text{score}}(l_j) | j = 1, 2, \dots, N)}$



支持度数 4, IDの数 6
協調性スコアの決定方法

³ 伊木惇, 龟井清華, 藤田聰, “レビューを対象とした信頼性判断支援システムの提案”, 情報処理学会論文誌, Vol. 55, No. 11, pp. 2461-2475, Nov 2014.

4.1 予測評価値からの適切な授業科目の推薦

10/20

- はじめに
- 教学データ分析と情報推薦
- 授業科目の成績予測と教材の最適化
- 予測評価値からの適切な授業科目の推薦
- 提案手法
- 数値実験並びに考察
- おわりに

大学には卒業要件単位が存在し卒業要件単位を満たさなければ卒業できない。ここでいう適切とは、「予測成績が高く、卒業要件単位、必修科目、選択必修科目のすべてを満たす」ということである。

○論理回路	◇	半	2	4単位以上(※) 修得すること
○アルゴリズムとデータ構造	◇	半	2	
○インターネット工学	◇	半	2	
○情報理論	◇	半	2	
○電気回路1	◇	半	2	4単位以上(※) 修得すること
○電子回路1	◇	半	2	
○電磁気学1	◇	半	2	
○電子物理	◇	半	2	

図14: 選択必修単位

区分	卒業要件単位	卒業研究履修条件単位	指定科目履修条件単位
総合科目	人間 2単位以上	2単位以上	教養小計 13単位以上(注3)
	社会・環境 6単位以上	6単位以上	
	言語・文化 4単位以上	4単位以上	
	精神・身体 3単位以上	3単位以上	
総合科目計	19単位(注3)	44単位	38単位
基礎科目	13単位	13単位以上	教養小計 13単位以上(注3)
	英語 10単位	6単位以上(注3)	
	第2外国語 2単位	2単位以上	
キャリア形成科目	7単位	3単位(注2)	
専門基礎科目	卒業研究以外 71単位	専門小計 69単位(注2)	
専門共通科目	卒業研究 8単位	79単位	—
専門科目	合計	130単位	70単位
		110単位	

図15: 卒業要件単位

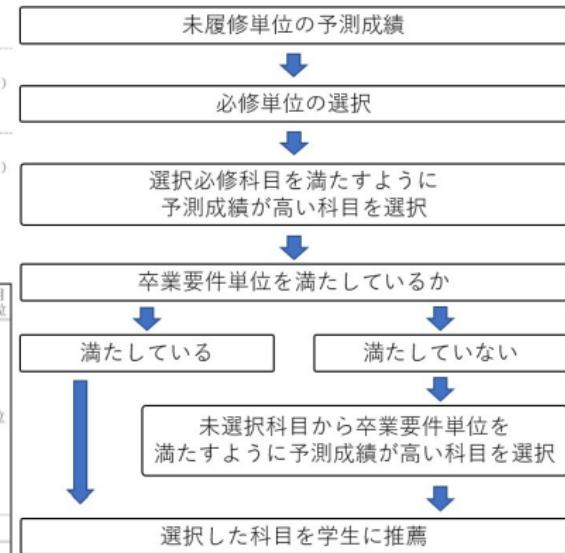


図16: 適切な授業科目の選択

4.2 関連資料のアップデート

11/20

- はじめに
- 教学データ分析と情報推薦
- 授業科目の成績予測と教材の最適化
- 予測評価値からの適切な授業科目の推薦
- 提案手法
- 数値実験並びに考察
- おわりに

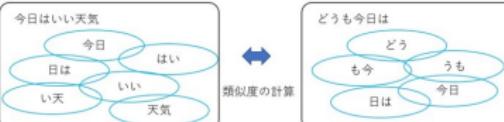
学生はこのような Web 学習システムにおいて互いに協調し合ってスパムレビューを投稿するとは考えにくいといった理由から、学生のレビューに対して類似性、集中性、情報性の 3 つの指標で各レビューのスパム性の判断を行い、信頼性スコアを算出する。

類似性

表5: n-gramによる文字の区切り

- n = 1: unigram 今日は/は/い/い/天/気/で/す/ね/。
- n = 2: bigram 今日/は/は/い/い/い/天/気/で/す/ね/。
- n = 3: trigram 今日は/は/い/は/い/い/い/天/気/天/気/で/氣/す/ね/。

・ Bigramで区切った文字を比較



集中性

Step 1: 日足によりバーストしている日を求める

Step 3: バーストしている時間帯に投稿されたレビューの投稿数を調べる

Step 2: バーストした日について15分足でバーストしている時間帯を求める

Step 4: 投稿数によって集中性を計算

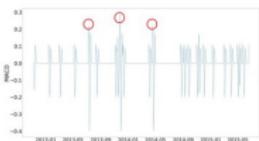


図19: 日足によるMACDの計算

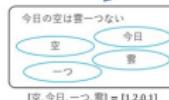
図20: 15分足によるMACDの計算

情報性



$[今日,いい,天気] = [2,0,0]$

他の文章中に名詞が何回使われているかをカウント



$[空,今日,一つ,雲] = [1,2,0,1]$

図18: 名詞の使用回数からの情報性の計算

信頼性

・ 指標からスパム性スコアを計算

類似性: S_score_{norm}

情報性: I_score_{norm}

集中性: C_score_{norm}

スパム性スコア: F_score

・ スパム性スコアとレビュー評価値から信頼性スコアを計算

教材に対するスパム性スコアの平均
レビュー評価値の平均 = 信頼性スコア: T_score

5.1 ロジスティック回帰とキャリア推薦

12/20

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

目的：どの科目がキャリアに有効かを知ること

過去の卒業生の 成績データ と 就職先データ を用い、特定の業種に進むためにどの科目が重要だったのかを分析する。

そのために、学生ごとに 適性スコア を計算し、合格の可能性を予測するモデルを構築する。

(ここにフロー図などの画像を挿入)

5.2 予測の仕組み (1)

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

ステップ1：適性スコア z の計算

学生の成績と科目の 重要度 を掛け合わせ 適性スコア z を計算する。このスコアが高いほど、その業種への適性が高いと判断される。

$$z = (w_1 \cdot x_1) + (w_2 \cdot x_2) + \cdots + (w_p \cdot x_p) + b$$

これをベクトルで書くと次のようになる。

$$z = \mathbf{w}^\top \mathbf{x} + b$$

- z ：ある学生の 適性スコア
- \mathbf{x} ：学生の成績ベクトル (x_1 : 科目 1 の成績, ...)
- \mathbf{w} ：各科目の 重要度（重み） ベクトル（モデルが学習する値）
- b ：全体のスコアを調整する 基本点（バイアス）

5.2 予測の仕組み (2)

- はじめに
- 教学データ分析と情報推薦
- 授業科目の成績予測と教材の最適化
- 予測評価値からの適切な授業科目の推薦
- 提案手法
- 数値実験並びに考察
- おわりに

ステップ 2：合格可能性 P への変換

適性スコア z を 0% から 100% の 合格可能性 P に変換するため、シグモイド関数 という便利な関数を用いる。

$$P(y = 1 \mid \mathbf{x}) = \frac{1}{1 + e^{-z}}$$

- $P(y = 1 \mid \mathbf{x})$ ：成績 \mathbf{x} の学生がターゲット業種に就職する ($y = 1$) 確率
- e ：ネイピア数 (約 2.718)
- z ：前スライドで計算した適性スコア

5.3 学習の仕組み (1)

最適な 重要度 w の発見方法

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

モデルは学習の際、以下の コスト関数 J を最小化するような最適な 重要度 w を探す。

$$J(\mathbf{w}, b) = \underbrace{-\frac{1}{m} \sum_{i=1}^m [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)]}_{\text{予測のズレ}} + \underbrace{\alpha \sum_{j=1}^p |w_j|}_{\text{複雑さへのペナルティ}}$$

- 第 1 項（予測のズレ）：モデルの予測 p_i と正解 y_i が違うほど、この値は大きくなる。この部分は 対数尤度損失 と呼ばれる。
- 第 2 項（複雑さへのペナルティ）：多くの科目の重要度 $|w_j|$ が大きいほど、この値は大きくなる。

モデルはこの両方の合計が最小になるような、最もバランスの良い 重要度 w を見つけ出す。

5.3 学習の仕組み (2)

特別なルール：L1 正則化の効果

- はじめに
- 教学データ分析と情報推薦
- 授業科目の成績予測と教材の最適化
- 予測評価値からの適切な授業科目の推薦
- 提案手法
- 数値実験並びに考察
- おわりに

コスト関数の第2項は L1 正則化項 と呼ばれ、モデルに特別な制約を与える。

$$\text{ペナルティ項} = \alpha \sum_{j=1}^p |w_j|$$

コスト関数全体を最小化する過程で、このペナルティ項を小さくしようとする力が働く。その結果、予測への貢献度が低い科目の重要度 w_j は、積極的に 0 に設定される。

- $\sum_{j=1}^p |w_j|$ ：全科目の 重要度（重み） の絶対値の合計。
- α ：ペナルティの強さを決めるパラメータ。この値が大きいほど、多くの科目の重みが 0 になる。

この仕組みのおかげで、本当に意味のある キー科目 だけが選び抜かれ、質の高い推薦リストが作成される。

数値実験と考察

1. はじめに
2. 教学データ分析
と情報推薦
3. 授業科目の成績
予測と教材の最
適化
4. 予測評価値から
の適切な授業科目
の推薦
5. 提案手法
6. 数値実験並びに
考察
7. おわりに

4.3 ロジスティック回帰と業種の予測

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

学生の成績や履修情報などの特徴量から、どの業種に就職する可能性が高いかを予測するために、ロジスティック回帰モデルを用いる。確率的に分類するため、以下のような条件付き確率を用いる：

$$P(\text{業種} = y \mid \mathbf{X} = \mathbf{x})$$

各記号の意味は以下のとおりである。

- $P(\text{業種} = y \mid \mathbf{X} = \mathbf{x})$ ：入力 \mathbf{x} に対し、業種が y (例：IT 業、製造業など) である確率
- 業種：学生の就職先の業界カテゴリ (目的変数)
- y ：業種の 1 カテゴリ (例： $y = \text{IT}$ など)
- \mathbf{X} ：学生に関する説明変数ベクトル (履修科目や GPA など)
- \mathbf{x} ：ある学生の実際の特徴量

4.3 ロジスティック回帰と業種の予測

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

実際のロジスティック回帰モデルでは、次のようなソフトマックス関数を使い、複数業種に対して分類を行う：

$$P(\text{業種} = y \mid \mathbf{x}) = \frac{\exp(\mathbf{w}_y^\top \mathbf{x})}{\sum_{k=1}^K \exp(\mathbf{w}_k^\top \mathbf{x})}$$

- \mathbf{w}_y ：業種 y に対応する重みベクトル（学習されるパラメータ）
- K ：分類対象の業種数
- \exp ：指数関数

おわりに

1. はじめに
2. 教学データ分析と情報推薦
3. 授業科目の成績予測と教材の最適化
4. 予測評価値からの適切な授業科目の推薦
5. 提案手法
6. 数値実験並びに考察
7. おわりに

学生が高いGPAを修めることができるよう、情報推薦、教材による学習を受けられるシステムの開発を行った。

- ・推薦された科目について推薦の有効性を示した。
- ・信頼性スコアを反映させることで教材のランキングが信頼できるものにした。

今後の課題

- ・単純な授業計画による検索では参考になりにくいサイトが多くヒットするときがある。
➡ 第十五回目「まとめ」など教材の作成を見直す必要がある教材が存在する。
- ・HTML内で上位3件から外れてしまうとなかなか再表示がされない。
➡ 学生に見せるHTMLの表示の仕方を改善する必要がある。
- ・有効性検証の際に同一サンプルを使用し、有効性を示した。
➡ サンプル数を増やし、改めて検証を行う必要がある。