

機械学習を用いた行列の問題を対象とした誤答原因推定機能の開発

Development of Functions for Estimating Cause of Incorrect Answer for Matrix Problems by Machine Learning

○齋藤 大¹, 星野 貴弘²

*Dai Saito¹, Takahiro Hoshino²

In recent years, education methods using ICT has progressed. It has become possible to perform individual learning that depend on the learners using digital teaching materials. However, these teaching materials are difficult to provide instruction suitable for learners. Because these teaching materials cannot determine the cause of the incorrect answer of the learner. Therefore, we develop the learning support system implemented function to estimate cause of incorrect answer by machine learning. Using the function, the system can automatically provide learning supports depending on types of incorrect answer to students. In addition, we evaluated the accuracy of determining the cause of incorrect answer for the function developed using machine learning.

1. はじめに

一般的な理工系の大学では、線形代数学が履修科目として授業に組み込まれている。しかし、2012年の学習指導要領の改訂により、線形代数学における行列の基礎的内容は高校数学では扱われなくなった^[1]。この影響を受けて、現行の学習指導要領で学習した学生は学習指導要領が改訂される前の学生に比べて線形代数学を学ぶにあたっての基礎知識が不足している。

一斉授業では、授業で扱う問題は学生の理解度に関わらず一般的に同じ問題が用いられる。しかし、大学入試の多様化の影響により、学生の数学に対する知識・理解度にばらつきがあるため、これらが不足している学生は徐々に授業についていけなくなってしまう。この問題を改善するためには、一人一人の知識・理解度に合わせて指導を行う必要があるが、大学の教育スタッフの時間、人数から困難である。

近年、ICTを活用した教育方法の導入が進んでおり、デジタル教材^[2]によって自分に合った進度で学習する個別学習が可能となっている。しかし、一般的なデジタル教材では学習者の誤答原因がわからないため、学

習指導を適切に行えていない可能性がある。誤答原因が単純な計算ミスか解くための知識が不足しているかによって必要となる学習支援は異なると考えられる。

本研究では、機械学習を用いた線形代数学における行列問題を対象とした誤答原因の推定機能の開発を行った。学生の誤答から2つの手法で教師データを作成し、誤答原因の判別性能の検証を行った。この判別結果を用いることで学習者の誤答原因に合わせた必要な知識の提示を行うことが可能になる。本稿では、開発した学習支援システムの全体像について述べた後、誤答推定機能の詳細について述べる。

2. 学習支援システム

学習者が本システムを利用する場合の操作の流れをFig.1に示す。学習者は取り組む単元を選択し、(a)の入力部に解答の入力を行う。(b)の解答ボタンを押すことで学習者は解答の正誤判定結果を確認できる。この時、誤答を入力した学習者にはその原因に合わせた必要な知識の提示を行う。

誤答推定機能では機械学習で学生の誤答について誤答原因を判別する。誤答原因がケアレスミスと判別さ

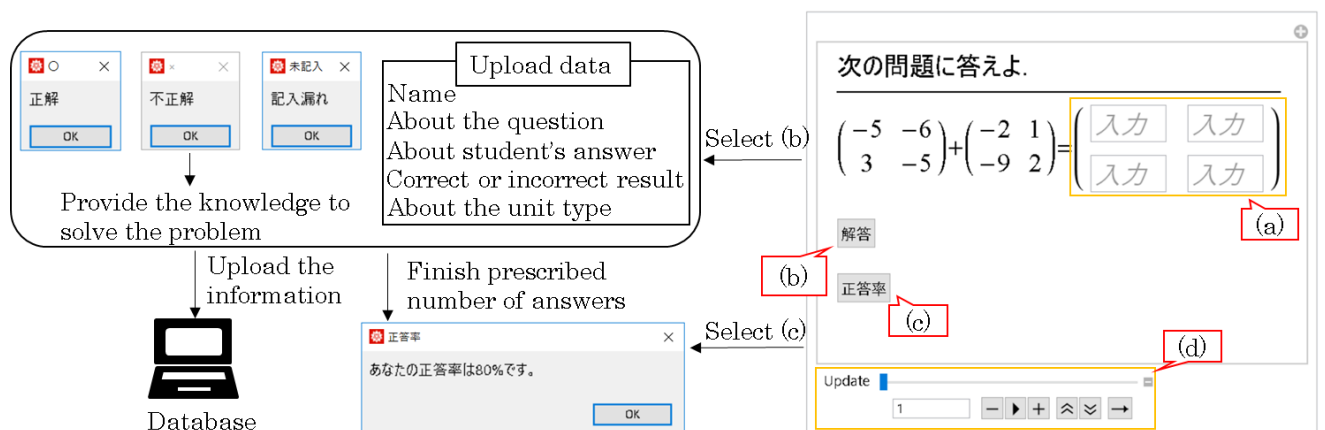


Fig.1 Flow of operation of the proposed system

1 : 日大理工・院[前]・電気

2 : 日大理工・教員・電気

れた学習者には解答の誤答部分を具体的に指摘し、解法の理解不足と判別された学習者には一般的な解法を提示することで知識の支援を行う。

解答ボタンを押すと、データベースに学習者の解答情報が記録をされる。教員が指定した問題数を解答し終わると、取り組んだ問題の単元の正答率を自動的に出力する。学習者は(c)の正答率ボタンを押すことでも、現段階での正答率を確認できる。(d)をスライドさせることによって新しい問題が生成される。

3. 誤答原因推定機能

〈3-1〉 誤答の定義

本機能では機械学習を用いて誤答原因の推定を行う。過去の授業の演習で実際に学生が行った誤答をケアレスミス、解法の理解不足の2つに定義した。ここでケアレスミスとは、解法自体は正しいものの単純な四則演算の計算間違いや数値・符号の記載ミスと定義し、解法の理解不足とは正しい解法とは異なる方法で計算している誤答と定義した。学生の実答から実際に計算過程などを見てどちらの誤答に分類されるか判断し、抽出された誤答情報を訓練データとして入力した。

〈3-2〉 線形 SVM

線形 SVM^[3]は、訓練データに対して空間内に分類境界を形成し、クラスを分類する手法である。訓練データから、分類境界と最も近いデータ点との距離が最大となるマージンの最大化を実現することで最適なパラメータを求める。本研究ではこのパラメータは、Python のデフォルトで搭載されているメソッドより自動的に決定した。

〈3-3〉 RBF カーネルを用いた SVM

線形 SVM では線形の分類境界で分類を行っていたが、カーネル関数^[3]を用いることで非線形な分類境界による分類を可能にしたモデルである。本研究ではカーネル関数としてラジアル基底関数を用いた。このラジアル基底関数のパラメータを変更することでモデルの複雑度を調整できる。この調整はグリッドサーチより行った。本研究では、ラジアル基底関数のカーネル法^{[3][4]}を用いた SVM(以下 RBF カーネルと記述する)を用いた。

4. 検証方法・検証結果

〈4-1〉 検証方法

本研究では分類した誤答情報を変数に置き換えたものを訓練データとして使用した。訓練データは、授業を受講した学生の誤答から抽出した 38 種類の誤答パターンで構成された訓練データと、上述の訓練データの担当教員とは異なる教員の授業を受講した学生の誤答から得た 12 種類の誤答パターンを加え、計 50 種類の誤答パターンで構成された訓練データを使用した。この 2 つの訓練データにつ

Table.1 Verification result

Method	RBF Kernel		Linear SVM	
	50	38	50	38
Number of test data	50	38	50	38
K-foldcross validation[%]	82.6	82.4	65.1	84.9

いて Python より RBF カーネルと線形 SVM でそれぞれ教師データを作成した。この訓練データから K-分割交差検証を用いて判別精度の検証を行った。この手法で 50 回判別結果を出力し、その平均値を判別精度とした。

〈4-2〉 検証結果

検証結果を Table.1 に示す。RBF カーネルで得られた教師データの判別結果は、データ数に関わらず、比較的高い精度となった。38 種類の訓練データの誤答パターンと、この訓練データに加えた異なる教員の授業で学んだ学生から得た 12 種類の誤答パターンは、傾向が大きく異なったため、線形 SVM による判別が困難になったと考えられる。

これらの結果から、誤答パターンが多様化した場合、RBF カーネルを用いて教師データを作成した方が精度は安定すると考えられる。本機能は、将来的には様々な教員の授業で抽出される誤答を教師データとすることが想定され、誤答パターンは多様化すると考えられるため RBF カーネルでの運用が適していると考えられる。

5. まとめと今後の課題

本研究では学習者の誤答原因を機械学習で推定することにより、理解度に合わせた学習指導を行う誤答推定機能の開発を行った。また、いくつかの機械学習モデルを利用して誤答推定機能の判別精度を検討した。その結果、RBF カーネルを用いた誤答推定機能が有用であることがわかった。

今後の課題として、学習者の理解度を問題の正答率以外の指標でソフトウェアが自動的に推定する機能を開発予定である。

参考文献

- [1] 文部科学省：「高等学校学習指導要領」，実教出版株式会社，pp.53-63(2012)
- [2] 小野雄一：「外国語教育におけるデジタル教材の自動配信システムの構築と評価」，電気学会論文誌 C 分冊，pp.94-103(2013)
- [3] 竹内一郎 鳥山昌幸：「サポートベクトルマシン」，講談社，pp.4-22，77-78(2015)
- [4] Christopher M Bishop：「パターン認識と機械学習 下ベイズ理論による統計的予測」，丸善出版，pp.1-12 (2012)