

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 4 日現在

機関番号：32670

研究種目：基盤研究(C) (一般)

研究期間：2013～2015

課題番号：25330419

研究課題名(和文) eラーニングでのビッグデータに適応可能な学習支援システムの開発

研究課題名(英文) Development of learning support system for big data in e-learning

研究代表者

小川 賀代 (Ogawa, Kayo)

日本女子大学・理学部・教授

研究者番号：20318794

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：高等教育機関を中心にLMSが普及し、学習履歴データが蓄積され、ビッグデータになりつつある。これらの情報を解析することで、学習活動の改善、将来的な能力の予測などに活用する期待が高まっている。本研究では、eラーニングにおいて蓄積された情報から、個人の学習傾向を抽出し、個人に適した学習支援システムの構築を目指した。

学習傾向の抽出は、k-means法を基にシステムティックに特徴分類する手法を開発した。学習結果の予測は、属性を考慮して可視化できるCR-SOMを用いた手法を開発した。実データに適応した所、SOMよりも同じ属性のデータが密集する傾向が得られ、識別率の高い予測に繋がる結果が得られた。

研究成果の概要(英文)：The LMS has become popular mainly among institutions of higher education and the learning history data has been accumulated, which is becoming big data. Analyzing the information is increasingly expected to be utilized for the improvement of learning activities and the prediction of future abilities. This study aimed to construct a learning support system suitable to individuals by extracting learning tendencies of individuals from the information accumulated in e-learning. To extract the learning tendencies, a method that categorizes the characteristics systematically based on the k-means method was developed. To predict the learning results, a method utilizing CR-SOM that can consider and visualize the elements was developed. As the result of adjusting to actual data, data of the same elements tended to concentrate closely more than SOM, so the result that leads to a prediction with high identification rate was obtained.

研究分野：情報システム

キーワード：学習支援システム eラーニング Learning Analytics ログ解析 SOM

1. 研究開始当初の背景

高等教育機関を中心にLMSが普及し、学習履歴データがサーバー上に蓄積され始めた。また、eポートフォリオの導入も進み、教育に関する様々なデータが蓄積されビッグデータになりつつある。しかしながら、これらの情報は十分に活用できていない状況にある。

また近年、eラーニングを取り巻く状況は、インフラが整ったことを受け、教育の質・効果についても要求され始めており、蓄積された膨大な情報の分析結果を、学習活動の改善、達成度の評価、将来的な能力の予測などに活用する期待が高まっている。

2. 研究の目的

本研究では、eラーニングにおいて蓄積された膨大な情報から、個人の学習傾向を抽出し、個人に適した学習支援システムの構築を目指す。

3. 研究の方法

個人の学習傾向を抽出し、個人に適した学習システムの構築を行うために、以下の小目標を設定する。

(1) 学習パターンの分類(クラスタリング)方法の検討

最もよく用いられるクラスタリング手法としてk-means法が挙げられるが、クラスタ数を経験的に決定しており、解析者によって異なる結果となってしまう。そこで、クラスタ数を決定する指標を用意し、自然な分布を反映したクラスタ数が得られるアルゴリズムを構築する。また、各クラスタが学習傾向を抽出しているかどうかの検証も行う。

(2) 可視化の検討

1 個人が有する情報は、LMSに残された学習ログ(アクセス回数やページ滞在時間など)や成績、それ以外での活動記録など多次元に亘るため、クラスタリング後のクラスタ間の関係、個人の集団における位置づけを理解するためには可視化が必要となる。そこで、自己組織化マップ(SOM)を用いた可視化の検討を行う。また、解析するデータの時間的粒度を変化させることで、学習支援につながる解釈が得られるか検討を行う。

(3) 学習結果や行動の予測

データマイニングによって得られた解析結果が、能力予測となっているかどうか蓄積データから検証を行う。それを行う手法については、(2)で可視化に用いていたSOMをベースに開発を行う。SOMは、属性(成績や成果)を考慮しないで可視化を行っており、新しいデータに対して成績や成果の予測が困難であると考えられる。そこで、SOMに属性を考慮したHyper SOMを用いた予測の検討を行った。更に、研究分担者が開発した属性を考慮したCR-SOM(context-relevant

self-organizing map)を用いて検討を行った。

(4) 実データへの適用

様々な分野においても提案手法が適応可能かを検証するために、実際にLMSで取得したデータを用いて実証実験を行う。

4. 研究成果

小目標ごとの研究成果を示す。

(1) 学習パターンの分類(クラスタリング)方法の検討

個人の学習傾向を抽出するための解析手法を提案した。最もよく用いられるクラスタリング手法としてk-means法が挙げられるが、クラスタ数を経験的に決定しており、解析者によって異なる結果となってしまう。そこで、クラスタ数決定において、3つの異なる指標の結果を統合して1つのクラスタ数を示す自動決定プログラムの開発を行った。

複数のデータセットを開発システムに適用し、開発システムから算出されたクラスタ数に基づいて、各クラスタのLMSの項目のデータの平均値をレーダーチャートで示したところ、特徴が明確に分類されていることが確認できた。これにより、本手法によって、学習パターンをシステムティックに分類することが可能となった。企業から提供してもらった2679人の学習履歴データを適用したときの結果を図1に示す。

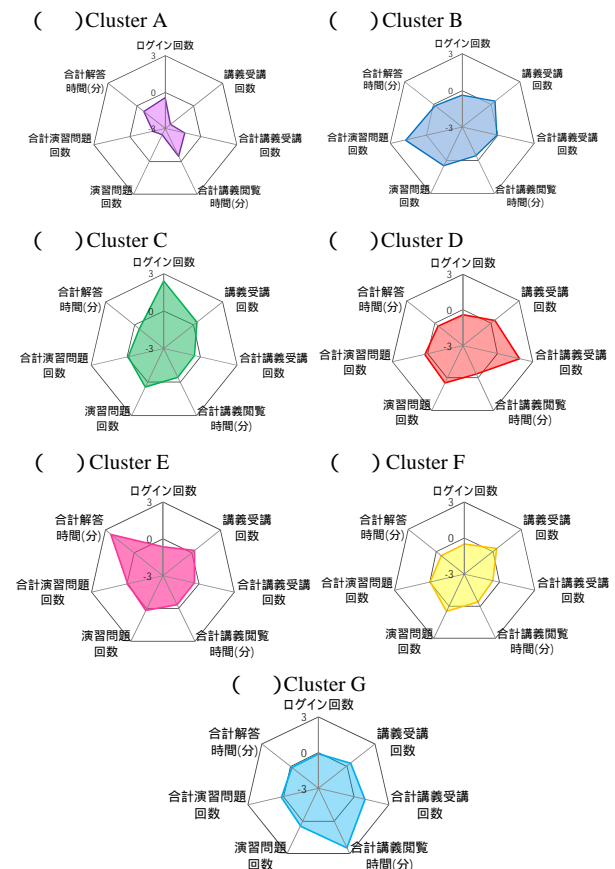


図1 クラスタ数7のレーダーチャート

(2) 可視化の検討

学習履歴データは、多次元データであり、クラスタ分類が単純にできる性質のものではない。適切に個人に適した学習システムを構築するためには、個人の集団における位置づけを理解する必要がある。そこで、多次元データを2次元に写像可能なSOMを用いて可視化を行った。図2にSOMの結果を示す。各クラスタがSOM上のどこに分布するかを確認するために、クラスタごとに異なる色で示した。結果からわかるように、クラスタ毎にまとまって分布していることが確認でき、可視化が行えたと同時に、クラスタ分析結果の妥当性も示すことができた。

また、LMSに蓄積されている学習履歴データの時間の粒度を変化させて解析を行ったところ、SOMによる可視化において、粒度を細かくすると、クラスタがより密集化する傾向が得られた。

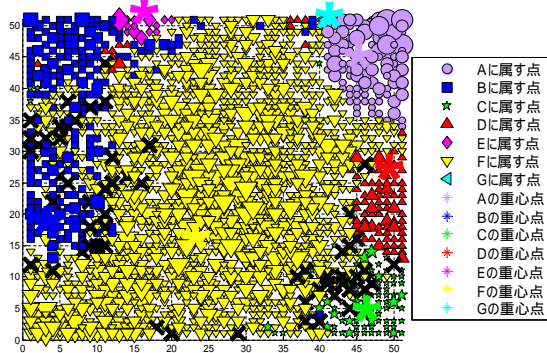


図2 SOMによる学習履歴データの可視化

(3) 学習結果や行動の予測

学習履歴データから属性の予測を行うために、SOMとハイパーSOMを用いて検証を行った。これまで、学習者の位置づけにSOMを用いており、学習傾向が近いものは、近くにマッピングされていた。この性質を利用すれば、学習傾向の予測ができると期待される。しかし、学習支援において、学習傾向から成績の予測ができることが望ましい。そもそもSOMは、属性を考慮に入れず可視化を行う手法である。そこで、属性を考慮したハイパーSOMを用いて、学習傾向から成績(属性)の予測を試みた。解析に使用したデータは、資格試験のためのeラーニングにおける学習履歴データを用い、受講者は計237名である。コンテンツは全10章からなっており、2-3, 4-6, 7-10章の学習後に章末テストとアンケート、全章学習後に修了テストが設けられている。今回使用したデータ変数は、講座学習回数、アンケート、章末テストのテスト点数、第1回模擬試験の受験回数、修了テストの受験回数であり、属性は、修了テストのテストの結果を偏差値で表し、3クラスに分けたものを用いた。本研究では、10分割交差検定によって、HyperSOMが学習者の属性の予

測が可能かどうかシミュレーション実験を行った。その結果を図3に示す。

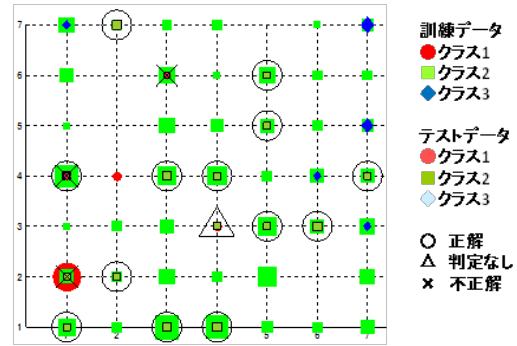


図3 HyperSOMの出力結果

各コードベクトルの属性は、マップされた教師データから決定し、属性の予測は、コードベクトルの属性とマップされたテストデータの属性が一致する場合を正解(印)、異なる場合を不正解(x印)、コードベクトルに教師データが入らず、属性の決定ができなかった所にテストデータがマップされた場合を判定なし(印)としてマップ上に表示した。その結果、予測率76.2%の結果が得られた。

更に、予測率の更なる向上を目指して、CRSOMを用いた、新たな手法についての検討を行った。従来のSOMの学習過程において、データのラベルは考慮されないが、CRSOMは、制限付きRBF(rRBF)ネットワークによって2次元に写像でき、非線形関係の可視化が可能な手法である。この提案システムを、一般的な評価用データセットに適用したところ、同じ属性をもつデータが密集して集まる傾向が得られた。さらに、実践的な複数の学習履歴データに対しても適用したところ、同じ属性をもつデータが密集して集まる傾向が得られた。

(4) 実データへの適用

本研究で開発してきたシステムが適用可能かどうかを検証するために、表1に示す4つのコンテンツのデータセットを用いて行った。

表1 実験で用いたデータセット

コンテンツ	人数	変数の数	属性の数
物理学概論	72	8	2
物理学概論	72	8	3
簿記	145	11	2
英語	211	5	2

いずれも、SOMよりもCRSOMで可視化した時に、属性ごとに密集する傾向が得られた。しかし、識別結果については、本研究で提案したrRBFが最近傍法や多層パーセプトロンと比べ、同等の結果となった。今後、ビッグデータに対して適応し、識別結果に差異が出てくるか実証実験を行っていく。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 5 件)

P. Hartono, P. Hollensen, T. Trappenberg, Learning-Regulated Context Relevant Topographical Map, IEEE Trans. on Neural Networks and Learning Systems, refereed paper, Vol. 26, No. 10, 2015, 2323-2335
DOI: 10.1109/TNNLS.2014.2379275

小川賀代、キャリア支援におけるポートフォリオ活用 持続可能なシステムに向けて (解説論文) 教育システム情報学会誌、査読有、32 巻、2015、27-36
DOI: 10.14926/jsise.32.27

P. Hartono and K. Ogawa, Visualizing learning management system data using context-relevant self-organizingmap", IEEE International Conference on Systems, Man, and Cybernetics, refereed paper, 2014, 3502-3506
DOI: 10.1109/SMC.2014.6974469

小川賀代、ピトヨ ハルトノ、LMS 蓄積データを用いた学習特徴の抽出における変数の粒度の検討、情報処理学会研究報告、査読無、vol. 2014-CLE-13、No.4、2014、1-5

石川晶子、小川賀代、ピトヨ ハルトノ、学習履歴データを活用した学習者の特性抽出手法の検討、教育システム情報学会誌、査読有、Vol.31、2014、185-196
DOI: 10.14926/jsise.31.185

[学会発表](計 10 件)

小川賀代、ピトヨ ハルトノ、学習履歴データの属性を考慮した可視化による学習予測の検討、第 41 回教育システム情報学会全国大会、2016 年 8 月 29-31 日、帝京大学宇都宮キャンパス (栃木県・宇都宮市)

小川賀代、e ポートフォリオによる学びの支援と質保証、大学・高校実践ソリューションセミナー 2015 東京、2015 年 11 月 18 日、株式会社内田洋行新川本社 (東京都・中央区)

秋本尚美、小川賀代、ピトヨ ハルトノ、HyperSOM による可視化を用いた学習支援への適応、平成 26 年度電子情報通信学会東京支部学生会研究発表会、2015 年 2 月 28 日、明治大学 (東京都・千代田区)

P. Hartono and K. Ogawa, Visualizing learning management system data using context-relevant self-organizingmap, IEEE International Conference on Systems, Man, and Cybernetics, 2014 年 10 月 5-8 日、

San Diego (USA)

小川賀代、ピトヨ ハルトノ、LMS 蓄積データを用いた学習特徴の抽出における変数の粒度の検討、第 13 回教育学習支援情報システム(CLE)研究発表会、2014 年 5 月 16-17 日、京都大学 (京都府・京都市)

小川賀代、LMS の蓄積データを用いた Learning Analytics、大学教育研究フォーラム、2014 年 3 月 19 日、京都大学 (京都府・京都市)

田中友里、小川賀代、ピトヨ ハルトノ、クラスター分析を用いた学習特徴の抽出、平成 25 年度電子情報通信学会東京支部学生会研究発表会、2014 年 3 月 1 日、東海大学高輪キャンパス (東京都・港区)

小川賀代、e ラーニングシステムを利用した学習履歴データの分析 ~ その特徴とこれからの課題、国際ワークショップ 社会のイノベーションを誘発する情報システム セッション 7「データ中心教育イノベーション」、2014 年 2 月 5 日、国立情報学研究所 (東京都・千代田区)

小川賀代、田中友里、ピトヨ ハルトノ、学習履歴データを活用した学習傾向の把握と可視化の検討、日本教育工学会 第 29 回全国大会、2013 年 9 月 23 日、秋田大学手形キャンパス (秋田県・秋田市)

小川賀代、e ポートフォリオによるキャリア支援、第 4 回 Mahara オープンフォーラム、2013 年 9 月 14 日、東京学芸大学 (東京都・小金井市)

6. 研究組織

(1) 研究代表者

小川 賀代 (Ogawa, Kayo)
日本女子大学・理学部・教授
研究者番号: 20318794

(2) 研究分担者

ハルトノ ピトヨ (Hartono, Pitoyo)
中京大学・工学部・教授
研究者番号: 90339747