

## 4L-04 機械学習を用いた Portfolio Assessment の特性分析

松居辰則 岡本敏雄

電気通信大学 大学院情報システム学研究所

### 1. はじめに

新しい学習評価観、すなわち問題解決評価観にもとづき、特に、総合的な学習における評価手法として Portfolio の概念を用いた評価手法が注目されている[1]. Portfolio とは「カリキュラムの一部分ないしは多領域における学習者全般にわたる努力、進歩、学業達成を長期的に表現する学習者の目的的な収集」と定義されている[2]. すなわち、Portfolio 評価とは、従来の測定即評価による定量的表現を重視した評価手法ではなく、学習者の学習過程・達成状態などの定性的表現をも重視した評価手法である。さらに、Portfolio 評価における学習者の収集活動は、学習者にその趣旨を伝え、彼らが自らの学習成果の選択なり質を決定するように学習者中心に収集が行われるとともに、他方では教授者自身が収集の目的を明確に定めて収集項目を決定することが重要であるとされている[3].

そこで、本研究では探求学習環境における学習者の行動制御 (Navigation) を目的として、Portfolio 評価の概念を用いた学習者モデルの構築手法を提案する。本手法は学習履歴から機械学習、特に決定木学習の手法を用いて学習者の学習過程・達成状態の構造を抽出することを目標とする。具体的には学習履歴データ群から学習項目を構成する属性間の相関ルールを抽出し、相関ルールの支持度から属性間の因果関連にもとづいた関連構造を構築する。

本稿では学習履歴データからの学習者の学習過程・達成状態を抽出し、構造を表現する手法に関して述べる。

### 2. Portfolio Assessment の定式化

本研究においては、Portfolio 評価を測定 (Assessment) の部分と評価 (Evaluation) の部分を明確に区別しており、本稿で扱う内容は測定 (Assessment) に関する部分に相当する。ここで、Portfolio Assessment (PA) の定式化を行う。本研究では教材構造がハイパー構造をもつ学習システムにおける学習者の学習履歴を Portfolio と捉える。そして Portfolio (学習履歴) から学習者固有の学習過程・達成状態 (学習構造) を定性的に表現することを Portfolio Assessment (PA) と定義する。

### 2.1 教材構造・学習項目の表現

まず、教材構造  $LS$  は学習項目  $I_i$  ( $1 \leq i \leq N$ ) と、教授者の教授戦略・教授目的にしたがって記述された項目間の構造によって記述される。項目間の構造は学習項目間の意味的な順序関連 (難易度、脈絡性など)  $r_{ij}$  ( $1 \leq i, j \leq N$ ) を指す。すなわち、教材構造  $LS$  は次のようなグラフ表現によって記述される。

$$LS = \langle I_i, r_{ij} \rangle \quad (1 \leq i, j, k \leq N : N \text{ は学習項目数})$$

次に、学習項目  $I_i$  は教授者の想定する項目の詳細や学習目的を説明属性  $a_l$  ( $1 \leq l \leq m$ ) とするベクトルデータとして表現される。

$$I_i = \langle a_{i1}, a_{i2}, \dots, a_{il}, \dots, a_{im} \rangle$$

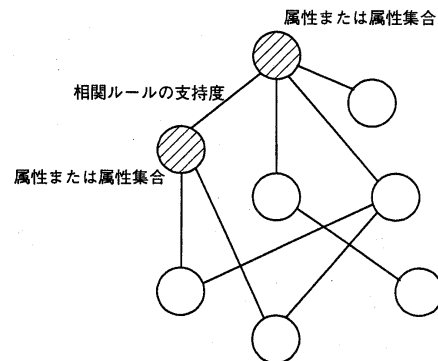


図1 学習構造  $SP^s$

### 2.2 学習履歴と学習構造の表現

学習者  $s$  の学習項目  $I_i$  に対する学習履歴  $his_i^s$  は各属性の属性値と1つの分類属性  $c_i$  の属性値 (クラス) の組として以下のように表現される。

$$his_i^s = \langle a_{is}, a_{is}, \dots, a_{il}, \dots, a_{im}, c_i \rangle^s$$

分類属性は学習者の当該学習項目に対する終了時における満足度や主観的な記述、定量的な評価結果などが用いられる。

次に、学習者  $s$  の学習構造  $PS^s$  を定義する。学習項目全体の部分  $N^s$  個 ( $N^s < N$ ) の学習履歴データ集合を  $his_{N^s, CN}^s$  とする。

Characteristic Analysis of Portfolio Assessment using Machine Learning Method

Tatsunori Matsui and Toshio Okamoto

Graduate School of Information Systems, The University of Electro-Communications

1-5-1 Chofugaoka Chofu, Tokyo 182-8585 JAPAN

$$his_{N^s \times C^N}^s = \left[ \begin{array}{cccc} & \dots & & \\ a_{i1}, a_{i2}, \dots, a_{ij}, \dots, a_{im}, c_i & & & \\ & \dots & & \end{array} \right]_{i \in N^s}$$

学習構造は学習履歴データ集合  $his_{N^s \times C^N}^s$  から抽出された説明属性間の相関ルールの支持度を属性間の関連性の程度と解釈し説明属性間の関係構造を求め、これを学習構造  $SP^s$  とする。学習構造  $SP^s$  のイメージを図1に示す。

### 3. 履歴データ集合からの構造抽出

履歴データ集合  $his_{N^s \times C^N}^s$  からの属性間の相関ルールの抽出は、履歴データ集合  $his_{N^s \times C^N}^s$  をトランザクションデータ形式に記述変換することにより実現する。

#### 3.1 相関ルールの定義

履歴データ集合  $his_{N^s \times C^N}^s$  をトランザクションデータ形式に変換したものの要素をアイテムとしアイテムの集合を  $T$  とする。また、トランザクションデータベースを  $D$  とする。アイテム集合  $X$  の支持度  $support(X)$  は  $D$  全体に対し  $X$  を含む割合を表す。

相関ルールは  $X \Rightarrow Y$  で表現される。ここで、  
 $X, Y \subset T, X \cap Y = \emptyset$

とする。相関ルールは支持度、確信度の2つのパラメータをもち、これらの値に閾値を設定し閾値との関係により相関ルールの有意性を判定する。

相関ルール  $X \Rightarrow Y$  の支持度  $support(X \Rightarrow Y)$  は  $D$  全体に対し  $X$  と  $Y$  が共起するトランザクションの割合  $support(X \cup Y)$  により定義する。この支持度が高いほど、相関ルールが表現する属性間のパターンが多く現れることになる。

確信度  $confidence(X \Rightarrow Y)$  は  $D$  の中で  $X$  を含むトランザクションのうち、 $X$  と  $Y$  が共起する割合、すなわち、

$$confidence(X \Rightarrow Y) = \frac{support(X \cup Y)}{support(X)}$$

によって定義される。確信度は相関ルールの条件部と結論部は同じ事例内に現れる共起性を表現する。すなわち、相関ルールの確信度が高いほど、このルールがより信頼性の高い結論を導くことを意味している。

#### 3.2 相関ルール抽出アルゴリズム

相関ルールの抽出は閾値によって指定された最小支持度と最小確信度を満足するすべてのルールを発見することになる。相関ルールは次の2つのステップで抽出される。

**step 1** 最小支持度を満足するアイテム集合を総べて抽出し、ラージアイテム集合を構成。

**step 2** step 1 で構成したラージアイテム集合から最小確信度を満たす相関ルールを導出。

相関ルール抽出においては、step 2 は step 1 で構成したラージアイテム集合に対してルールを導出する処理であり、比較的容易な処理である。これに対して、step 1 はトランザクションデータベースを繰り返し検索し、多くのアイテム集合の支持度を計算するため多くの処理時間を必要とする[4]。よって、本研究では効率性の観点から Apriori アルゴリズム[5]を採用している。

### 4. 探求学習支援システムへの実装

本研究では学習構造  $SP^s$  を探求学習環境における学習者モデルと位置付け、探求学習支援システムへの実装を試みている。特に、学習項目の特徴や目的を表現する説明属性は教授者があらかじめ設定するが、学習構造  $SP^s$  から関連構造の意味付けを学習者に求めることも可能である。これはインターフェースレベルでの実現となる。また、教材構造  $LS$  との構造的な比較をすることにより、学習者の学習過程・達成状態の評価も可能となる。これらの点は構想段階ではあるが、Portfolio 学習、Portfolio 評価が目指すところの、学習者中心の学習評価も可能であると考えられる。

### 5. まとめ

本稿では学習履歴を Portfolio として捉え、そこから相関ルールを抽出することにより、学習項目の説明属性の関連構造を構築する手法に関して述べた。本手法はトランザクションデータ形式に基いているため、説明属性の追加・削除および学習構造の融合・分解も可能である。この点を考慮すると、分散協調学習環境におけるグループモデルの構築への発展を期待することができる。この点に関するアルゴリズムは稿を改めて報告する

#### 参考文献

- [1] 小田勝己, "総合的な学習に適したポートフォリオ学習と評価", 学事出版(1999).
- [2] Sandra L. Shurr, Julia Thomason and Max Thompson, "Teaching At The Middle Level", D.C.Health and Company, pp.358(1995).
- [3] Gordon F. Vars, "The Effects of Interdisciplinary Curriculum and Instruction", Annual Review of Research for School Leaders(1996).
- [4] 喜連川優, "データマイニングにおける相関ルール抽出技法", 人工知能学会誌, Vol.12, No.4, pp.513-520(1997).
- [5] Argrawal, R. and Srikant, R., "Fast Algorithms for mining association rules", Proc. of the 20<sup>th</sup> VLDB Conference, pp487-499(1994).