

データマイニングを活用した学習方略フィードバックシステムの開発と評価

Developmet and Evaluation of a Feedback System of Learning Strategy to Students Utilizing Data Mining Technology.

松河秀哉*¹・北村智*²・永盛祐介*³・久松慎一*²・山内祐平*²・中野真依*⁴・金森保智*⁴・宮下直子*⁴
Hideya MATSUKAWA*¹・Satoshi KITAMURA*²・Yusuke NAGAMORI*³・Shinichi HISAMATSU*²
・Yuhei YAMAUCHI*²・Mai NAKANO*⁴・Yasutomo KANAMORI*⁴・Naoko MIYASHITA*⁴

大阪大学大学教育実践センター*¹・東京大学大学院情報学環*²・筑波大学大学院人間総合科学研究科*³・株式会社ベネッセコーポレーション*⁴

Institute for Higher Education Research and Practice*, Osaka University*¹, Interfaculty Initiative in Information Studies, The University of Tokyo*², Graduate School of Comprehensive Human Sciences, University of Tsukuba*³, Benesse Corporation*⁴

〈あらまし〉本研究では、高校生から得られたデータに基づいて、データマイニングを活用した学習方略フィードバックシステム「学習ナビ」を開発した。システムの試験運用をふまえ、(a)モデルの妥当性 (b) 学習ナビで利用したメタファの有効性 (c) ユーザからの主観的評価の観点から評価を行い、以下の結果を得た。(a)モデルが仮定する学力差が評価モニタにもみられ、モデルの妥当性が示唆された。(b)学習方略の達成度を表す信号機メタファについて、解説画面の閲覧時間の差から有効性が確認された。学習方略の順序性を表す一本道メタファは、評価モニタの約半数の理解を得た。(c)一部のユーザからアニメーションの長さを指摘された以外は、システム全体として好意的な評価を得た。

〈キーワード〉 学習方略・Web診断システム・データマイニング

1. はじめに

筆者らは、「上手な勉強の仕方が分からない」高校生が多数を占める状況に対応するため、学習方略フィードバックするシステム「学習ナビ」を開発した(松河ほか 2006)。

学習ナビは、web上で入力された学習者の現在の学習方略をもとに、その方略が適切なものであるかを判定し、結果についてわかりやすいメタファを用いてフィードバックするwebベースのシステムである。

システムは、1)学習者にいくつかの学習方略を提示し、それらの方略をどの程度活用しているかを質問するwebページと、2)そこで得られた情報をもとにサーバ側で学習者の学習方略の適切さを判定する部分(以降、ストラテジーエンジン)、3)クライアント側でその判定結果をわかりやすく視覚的にフィードバックする部分(以降、ヴィジュアルイザー)、4)各エンジンからの情報を記録し、相互に提供するためのデータベースより構成される。

「学習ナビ」の特徴は、ストラテジーエンジンを設計するにあたって、約1300人の高校

生の学力と学習方略のデータを収集し、そのデータにデータマイニングの一手法である回帰二進木分析を適用した結果を用いたこと、ヴィジュアルイザーに一本道と信号機のメタファを組み込んで、視覚的で分かりやすいフィードバックを目指したことである。

本研究では、「学習ナビ」の運用を通して得られたログデータや、利用者に対する質問紙調査を通して、ストラテジーエンジンに組み込んだモデルの妥当性と、ヴィジュアルイザーに組み込んだメタファの有効性について評価を行う。

2. 「学習ナビ」の運用

開発したシステム「学習ナビ」は、2006年11月17日午前10時から同年12月26日午前0時の期間、株式会社ベネッセコーポレーションのWebサービスである「マナビジョン」内のWebサイトで試験運用された。

そのWebサイトにて、ユーザは学力テストを受験したあとに、「学習ナビ」を利用できた。学力テストは英語・数学・国語の三教科が準備されており、問題はPDFファイルで提供さ

れた。ユーザは問題をダウンロードしてテストに取り組み、解答を Web 上で入力する。解答提出後、その教科の「学習ナビ」が利用可能になる。

試験運用にあわせて、2006年8月から10月に「学力・学習力到達度診断テスト」を受験したベネッセコーポレーションの進研ゼミ高1講座会員から、評価モニタ300名を募集した。評価モニタには英語についての「学習ナビ」を利用した上での質問紙への回答を求めた。教科を英語としたのは、文系・理系によって教科への関与度に差異が生じないようにするためである。

評価モニタの学力テストの解答入力は、事前に郵送した入力表に基づき評価モニタ自身が入力した。入力表は評価モニタ自身の「学力・学習力到達度診断テスト」での解答から作成された。評価モニタは英語の「学習ナビ」を利用後、事前に郵送した自記式の質問紙に回答した。回答された質問紙は、評価モニタからの返送を受け、2006年12月29日まで受け付けた。質問項目の詳細は「3. 『学習ナビ』の評価」において述べる。

質問紙を回収できた評価モニタは最終的に217名であった。サーバに残った「学習ナビ」の利用状況を示すログデータに欠損のあったIDは、分析対象から除いた。最終的に206名がログデータの分析対象となり、質問紙のデータにも欠損のない対象は197名であった。また、「マナビジョン」の会員である高校1年生も一般ユーザとして利用対象となった。一般ユーザは英語・数学・国語の三教科について利用可能であった。

三教科すべての「学習ナビ」を利用した一般ユーザには、Web上でサイト評価に関する質問紙への回答協力が求められた。この質問紙には、「学習ナビ」の評価に関する項目が含まれていた。質問項目の詳細は「3. 『学習ナビ』の評価」において述べる。

試験運用期間中に同Webサイトにアクセスした一般ユーザは約1000名であった。そのうち三教科利用は約600名、二教科利用は約50名、一教科利用は約160名であった。サイト評価に関するWeb質問紙に回答した一般ユーザは268名であり、全項目に欠損のない分析対象は235名となった。

3. 「学習ナビ」の評価

3.1. 評価の観点

評価モニタによる質問紙への回答データと「学習ナビ」の利用ログデータ、および一般ユーザによるWeb上での評価質問紙の回答データを分析に用いた。

「学習ナビ」は、ユーザに対して適切な学習方略に関する情報を分かりやすく伝達することを目的とする。そこで、評価の観点を(1)モデルの妥当性、(2)信号機メタファの有効性、(3)一本道メタファの有効性の3点に設定し、さらに総合的評価のために(4)ユーザからの主観的評価、を加えた。

3.2. モデルの妥当性の検討

モデルの妥当性の検討は、評価モニタの学力テスト結果と学習方略に関するWeb質問紙への回答を用いた。

妥当性の検討の対象となった英語のモデルは図1である。このモデルにおける最適ノードは図1の(1)と(2)である(最適ノードの算出方法に関しては松河ほか(2006)を参照)。

モデルの仮定では、決定木のルートノードからの第一分岐点で2つの最適ノードとは異なる側へ分岐をしたノード群(図1のノード1)がもっとも学力が低くなる。第二の分岐点で最適ノードとは異なる側へ分岐をしたノード群(図1のノード2)が次に学力が低くなる。

図1のノード3内でも学力に差異が生じることが仮定されるが、分析対象のサンプル数が206であるため、ノード3からさらに分割を行った場合、一群あたりのサンプル数が減

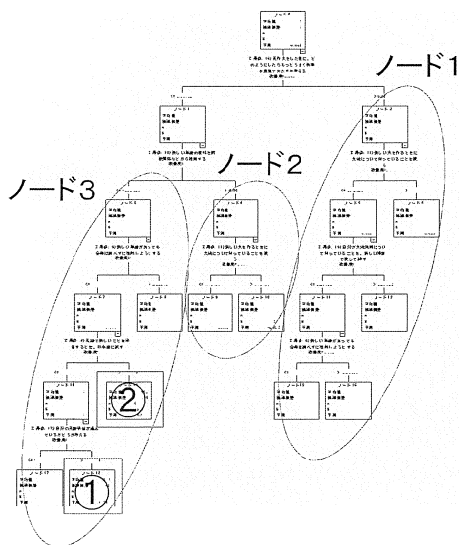


図1 モデルに順じたノードの分類

少しすぎる（ノード3のサンプル数は45）。

ここでは、図1におけるノード1, 2, 3の3つのノード間で学力に差異があるかを検討した。

3つのノード間での学力の差を検討するため、独立変数をノード1, 2, 3の3群、従属変数を学力値として、Kruskal-Wallis検定を行った。Kruskal-Wallis検定は全体の観測値を小さい順に並べた際の順位を用いて検定を行なうため、各群の順位の平均値（平均順位）に着目する。ノード1の平均順位は94.29、ノード2の平均順位は110.27、ノード3の平均順位は119.76であった。検定の結果、3つのノード間で学力の有意な差が認められた（ $\chi^2(2)=6.67, p<.05$ ）。

モデルの仮定では、ノード3の学力がもっとも高く、次いでノード2、そしてノード1の学力がもっとも低くなる。それぞれのノードの平均順位はこの仮定に合致しており、検定により5%水準での有意な差が認められた。

評価モニタの少なさのため、十分な妥当性検証が行えたとはいえないが、この結果から、「学習ナビ」のモデルは一定程度、妥当という示唆をえた。

3.3. 信号機メタファの有効性の検討

信号機メタファの有効性の検討は、評価モニタの「学習ナビ」の利用ログデータを用いて行った。

ユーザが「学習ナビ」の提示する学習方略の解説文を入念に読んだ場合、その解説画面の閲覧時間が長くなると考えられる。閲覧時間には解説文自体の長さや内容の難易度の影響も考えうる。だが、各信号で文字量に大きな差はなく、例えば英語のSTEP1の解説文は青信号149文字、黄信号141文字、赤信号143文字である。また、学習方略の基本的な説明は信号の色によらずほぼ同一であるため、難易度に大きな差はない。よって、評価指標に解説文の閲覧時間を用いた。各学習方略の解説画面（STEP）の閲覧時間を、各STEPに移るボタンをユーザが押してから、次のSTEPまたは画面に移るボタンを押すまでの時間と定義した。利用ログデータから各STEPの閲覧時間を算出した。

各STEPでの信号の色による閲覧時間の差を検討するために、各STEPでの信号の色を独立変数、各STEPの閲覧時間を従属変数と

してKruskal-Wallis検定を行った。各STEPでの各信号の人数、信号の色ごとの平均値、平均順位、および検定の結果を表1に示す。

検定の結果、STEP5は0.1%水準で、STEP2、STEP3は5%水準で、STEP1は10%水準で、信号の色による閲覧時間の有意差が認められた。STEP4では、信号の色による閲覧時間の有意差は認められなかった。

閲覧時間の分布は正規分布からの逸脱が大きく、黄信号、青信号の利用者の中に著しく閲覧時間の長いものがみられたため、平均値では赤信号の閲覧時間が短い場合も生じた（表1）。だが表1の平均順位に着目すると、どのSTEPでも青信号の平均順位がもっとも低く、黄信号、赤信号の順に高い。このことは、分布形状や外れ値の影響の小さい順位情報からみて、群の傾向として青信号のユーザの閲覧時間が短く、赤信号のユーザの閲覧時間が長かったことを意味する。黄信号・赤信号の人数が少なかったSTEP4を除いて、信号の色によって閲覧時間に統計的有意差が認められたことから、信号機メタファは有効に機能していたと考えられる。

3.4. 一本道メタファの有効性の検討

一本道メタファの有効性を検討するために、評価モニタに対して、「[学習ナビ]で示された勉強の方法のうち、まず初めにやってみなくてはいけない勉強方法はどれだと思いましたか」という項目で、自由記述回答を求めた。

一本道メタファが有効ならば、ユーザは最も黄信号・赤信号が出たSTEPの学習方略から取り組むことが効率的であると理解できると考えられる。

そこで、前述の自由記述回答と、最初に黄信号・赤信号が出たSTEPの提示内容が一致していたかを検討した。最初に黄信号・赤信号が出たSTEPの提示内容は利用ログデータから判別した。分析対象は利用ログデータに

表1 信号の色と各STEPの閲覧時間

	STEP1			STEP2			STEP3		
	人数	平均値	平均順位	人数	平均値	平均順位	人数	平均値	平均順位
青信号	92	23.01	93.48	76	14.01	88.05	128	15.05	95.33
黄信号	107	31.93	110.69	105	16.52	112.07	78	17.83	116.91
赤信号	7	23.57	125.29	25	13.72	114.48	0	?	?
χ^2	5.10			8.14			6.37		
d.f.	2			2			1		
	STEP4			STEP5					
	人数	平均値	平均順位	人数	平均値	平均順位			
青信号	189	14.86	101.35	55	10.53	64.38			
黄信号	11	14.55	124.41	116	19.68	101.57			
赤信号	6	13.33	132.75	9	18.00	107.39			
χ^2	3.05			20.05					
d.f.	2			2					

χ^2 値はKruskal-Wallis検定の結果

表2 ユーザからの主観的評価の結果

	評価モニタ (n=197)				一般ユーザ (n=235)			
	平均値	中央値	S.D.	z	平均値	中央値	S.D.	z
[学習ナビ]の操作で迷うことがあった	3.49	4.00	1.35	-4.70 p<.01	3.99	4.00	1.20	-9.77 p<.01
[学習ナビ]のアニメーションは長いと思った	3.51	4.00	1.06	-5.95 p<.01	2.51	2.00	1.21	-5.57 p<.01
[学習ナビ]で示された勉強の方法は自分のために選ばれたものだった	3.49	4.00	0.89	-6.67 p<.01	3.34	3.00	1.01	-4.76 p<.01
[学習ナビ]で示された勉強の方法は信じられるものだった	3.93	4.00	0.72	-10.82 p<.01	3.76	4.00	0.89	-9.60 p<.01
[学習ナビ]で示された勉強の方法は、どれも大切だと思った	4.07	4.00	0.80	-10.77 p<.01	3.91	4.00	0.88	-10.62 p<.01
自分ほどよい勉強の方法をたればいいかわかった	3.77	4.00	0.92	-8.78 p<.01	3.68	4.00	0.99	-8.43 p<.01
どの勉強方法からやってみればいいのかわかった	3.56	4.00	0.99	-6.84 p<.01	3.54	4.00	0.98	-7.18 p<.01
自分がよくできている勉強の方法がわかった	3.68	4.00	0.91	-8.27 p<.01	3.75	4.00	0.93	-9.36 p<.01
自分があまりできていない勉強の方法がわかった	3.90	4.00	0.90	-9.58 p<.01	3.89	4.00	0.91	-10.37 p<.01
[学習ナビ]で示された勉強の方法を実際にやってみようと思った	3.97	4.00	0.86	-10.06 p<.01	3.84	4.00	0.99	-9.58 p<.01
[学習ナビ]の利用は楽しかった	3.53	4.00	0.98	-6.62 p<.01	3.75	4.00	1.03	-8.76 p<.01
[学習ナビ]で示された勉強方法はいい勉強方法だと思った	3.92	4.00	0.80	-10.28 p<.01	3.86	4.00	0.93	-10.01 p<.01
数学についても[学習ナビ]を使ってみたい	4.07	4.00	0.95	-9.93 p<.01	?	?	?	?
国語についても[学習ナビ]を使ってみたい	3.97	4.00	0.97	-9.49 p<.01	?	?	?	?

z 値は3]を比較基準値としたWilcoxonの符号付順位検定の結果

も質問紙のデータにも欠損値のない 197 名のうち、一度は黄信号・赤信号が提示された 190 名とした。

この提示内容と自由記述回答の内容の一致について、第二著者と本研究に参加していない第三者が独立に判定した。二人の判定の一致率は 93.7 %であった。例えば、「英文は、調べる前にまず読もう」という見出しの学習方略に、「英文を読むときに英語を一回一回調べるのではなく一度全体を読んでみる」という回答だった場合などに提示内容と回答内容が一致したと判定した。

判定の結果から、49.5 %のユーザにおいて、最初に黄信号・赤信号が示された STEP の提示内容と、前述の質問に対する自由記述回答の内容が一致していた。

ここから、有効性がなかったわけではないが、一本道メタファの理解はユーザの半数程度にとどまったといえる。この理解率には改善の余地があるといえる。

3.5. ユーザからの主観的評価

「学習ナビ」の主観的評価を把握するために、ユーザに「学習ナビ」利用後に質問紙調査への回答を求めた。評価モニタからは「英語」の「学習ナビ」を利用後の主観的評価データを得た。一般ユーザからは、三教科すべての「学習ナビ」を利用後の主観的評価データを得た。各項目は、「とてもよくあてはまる」から、「まったくあてはまらない」までの5段階尺度で評定得た。回答結果は「学習ナビ」に対して肯定的な場合に値が大きくなるよう適宜反転を行った。なお、実施の制約上、評価モニタと一般ユーザの質問項目数は一致し

ていない(質問項目は表2に示す)。

評価モニタからの回答結果、一般ユーザからの回答結果に関して、平均値、標準偏差、中央値、および尺度の中央値にあたる「3」を比較基準値とした Wilcoxon の符号付順位検定の結果を表2に示す。

表2から、評価モニタ、一般ユーザともほとんどの項目で、平均値は3より大きく、中央値は4であった。「学習ナビ」は評価モニタからも、一般ユーザからも好意的に捉えられたと考えられる。

だが、一般ユーザでは「学習ナビのアニメーションは長いと思った」の項目に対し、平均値が2.51、中央値が2であった。アニメーションの長さには一般ユーザから否定的評価を受けたといえる。「学習ナビ」のアニメーションにはメタファ部分と関連のないオープニングアニメーションが含まれる。一般ユーザには、三教科すべての利用後に評価を求めたため、アニメーションが過剰に感じられた可能性がある。

参考文献

松河秀哉・北村智・永盛祐介・久松慎一・山内祐平・中野真依・金森保智・宮下直子(2006) データマイニングを活用した学習方略フィードバックシステムの開発 日本教育工学会第20回講演論文集: 63 - 66

* 本研究は、東京大学大学院情報学環ベネッセ先端教育技術学講座のプロジェクトとして、(株)ベネッセコーポレーションと共同で行われている。

