

e-Learningにおける適応的コンテンツ配信制御

吉田 賢史¹⁾・宮崎 光二²⁾・中上 香代子³⁾・中山 弘隆³⁾

自らの試行錯誤を通じ学習することは、学習者にとって大切であるが、そのような実験的学習は多くの時間を必要とするため、50分という授業時間内に、学習者一人ひとりの実験内容に適した指導をおこなうことは難しい。また、学習進度の速い・遅いという個人差に対応することも難しい。

そこで、我々は、そのような時間と場所の制約にとらわれないe-Learningの特性を生かし、学習者の試行錯誤という実験入力に対し臨機応変に対応し、その学習者に適したコンテンツを配信するシステムの構築を試みた。

しかし、学習者の実験入力情報による配信コンテンツの決定は、小テストなどとは異なり正解・不正解のように2値的に分かれるものではなく様々な可能性が存在するため困難である。

我々は、高校1年生の数学「2次関数のグラフとその平行移動」を題材に試行錯誤を繰り返して頂点の求め方を発見させる試作コンテンツを作成し、そこで得られた入力データをもとに、学習者の状況に適した配信をおこなうための単純2値による判別、k-NN (k-Nearest Neighbor) 法およびSVM (Support Vector Machine) 法による判別手法を試み、検討をおこなった。

本論文では、学習者の状況に応じて適切なコンテンツが配信されるe-Learningシステムの構造および配信制御の手法について報告する。

キーワード

実験数学、発見的学習、SVM (Support Vector Machine)、WBT (Web Based Training)、LMS (Learning Management System)

1. はじめに

1.1 e-Learningへの取り組み

e-Learningは、企業をはじめ大学でも活用されている。特に、企業においては一定の知識を社員に身につけるために、解説を聞きながら演習を繰り返すというWBT (Web Based Training) が有効とされ導入が進んでいる。最近では演習成績などにより、配信コンテンツを制御したり、既習教材をリスト消去させるLMS (Learning Management System) も存在する。

問題演習について金西ら (2003) は、同じ演習を繰り返しても、答えを暗記してしまうことから、演習内容にバラエティを持たせるため問題を自動生成させる機構について報告している (金西ほか 2003)。また、記述式問題に対する自動評価について宮地ら (2005) は、講義スライドから記述式問題を自動生成し、学習者の解答を自動評価する機構について報告している (宮地ほか 2005)。

大学におけるWBT以外の利用では、主に講義の補助や講義の再現として活用されているが、資格試験対策のための利用などに比べて学習者の目的意識が希薄なため、学習への集中力の持続が難しいと言われている。

学習への集中力を配信されるコンテンツ画面への集中力ととらえて、安藤ら (2004) はアイマークレコーダーによる画面への集中力の持続に関して計測し、18分頃に集中力が欠けると報告している。また、画面への集中力を引き戻すため、学習コンテンツに大きく変化する画面効果を付けることが有効であると報告している (安藤ほか 2004)。

初等中等教育においても、ウェブページを活用した授業実践がおこなわれ、学習者のモチベーションの向上に関する実践が報告されている。試行錯誤などの実験やシミュレーションは、学習者一人ひとりの思考にもとづいて実験ができるため、関心・意欲・態度などの情意面での効果が期待されている。

本論文では、島根式数学における情意検査 (中学) を中学生に実施し、e-Learningによって情意に変化が認められるか調べた。その結果については3.2節で報告する。

しかし、大橋ら (2005) は、「高等学校の教育のあり方」

¹⁾ 甲南大学自然科学研究科 情報・システム科学専攻

²⁾ (株)ニルバーナテクノロジー

³⁾ 甲南大学 理工学部 情報・システム工学科

の中で、数学教育の現状と問題点についてふれている。高等学校教育課程調査を引用し、コンピュータの利用は7%程度であり、生徒が主体的に活動する場面を作っていると答えた教員の割合は36%程度であったと報告している(大橋ほか 2005)。

このような現状と授業実践研究とのギャップを埋める必要があるが、学習者主体の授業を、限られた時間内に展開することは困難である。それは、学習者の個々の学習能力や塾その他の学校以外の学習による知識の差によって学習者の状況が異なるためである。

1.2 学習活動と配信制御

試行錯誤などの実験的な学習活動、発見的学習は、e-Learningにおいて展開すれば、学習者一人ひとりがそれぞれのペースで学習を進めることが可能である。

ところが、そのような学習は、教員が側にいて適切な助言をしなければ、単なる遊びになる可能性が高い。また、教員の助言無しに実験を繰り返しても、学習者は学習の方向性を見失い、楽しさのみが記憶に残る可能性が高く、学習目標到達のための学習になり難い。

そこで、我々は、学習者に試行錯誤や実験などの活動をさせ、その実験の入力内容によって適切な教材を配信することはできないかと考え、入力情報のパターン化とそれによるコンテンツ配信制御の手法を提案する。

このような学習者の状況に応じた教材配信については、國宗ら(2004)が研究報告しているが、我々の提案方法とは異なり、教員が手動により配信する教材を制御している(國宗ほか 2004)。

また、小テストなどをおこない成績によって理解不足の単元へ戻って教材を配信させるe-Learningシステムも存在するが、我々が対象とする配信制御は、点数などの評価による制御ではなく、実験という正解・不正解のように2値的に分けられないものを対象としている点で異なる。

集中力の持続に関しては、画面効果という手法ではなく、実験などの学習者の能動的活動によって持続させられないかと考えている点で前述の安藤らの手法とアプローチが異なる。

自発的に学習させるために、e-Learningを用いて講義と連携させることで学習展開する教育方法が荒川ら(2005)から提案されている。この方法は、事前学習としてアンケートや問題に答え、事後学習としてアンケートおよび応用問題に取り組む形式となっており(荒川ほか 2005)、コンテンツが自動配信制御されない点で我々の研究とは異なる。

我々は、試行錯誤などの実験入力値をもとに学習者の状況に応じたコンテンツを配信するための入力データのパターン化の手法、および、学習者がどのような実験をおこなっているかを判断する手法についていくつか試行

した。

本論文では、学習者の実験入力データをもとに、学習者の状況を判断する最も適切な手法について検討をおこなった結果を報告する。

2. システムの概要

システムは、学習者の試行錯誤による実験入力を処理し実験の結果を学習者に返す実験処理サーバ、および、その実験入力値を記録するデータベースサーバ、さらに、そのデータベースから学習者の入力情報を定期的に取得し学習者の入力情報を解析するサーバ、学習者が学習を進めるクライアントコンピュータと教材のwebページを配信するHTTPサーバで構成されている(図1)。

教材は、プレゼンテーションソフトなどで作成された解説が配信された後(図2)、学習者の試行錯誤を促すwebページが表示される(図3)。学習者は、このページでいろいろな数値を入力しグラフの変化を観察することになる。そこで入力された実験入力値は、データベースサーバと実験処理サーバに送られる。実験処理サーバは、学習者の実験入力ごとにその結果を学習者に返す。さらに、実験入力処理サーバは定期的にデータベースサーバから学習者の入力を取得し、その実験内容を解析、学習者の実験内容に応じて次の指導のためのコンテンツを決定し、HTTPサーバに対して次のコンテンツ配信の要求を送る。

学習者は自らの実験をおこなうことにより、能動的に学習を進めることができる。また、実験の内容によって配信される指導のコンテンツが変化するため、学習者は学習目標を見失うことがない。

我々は、高校1年生の数学の単元「二次関数のグラフ

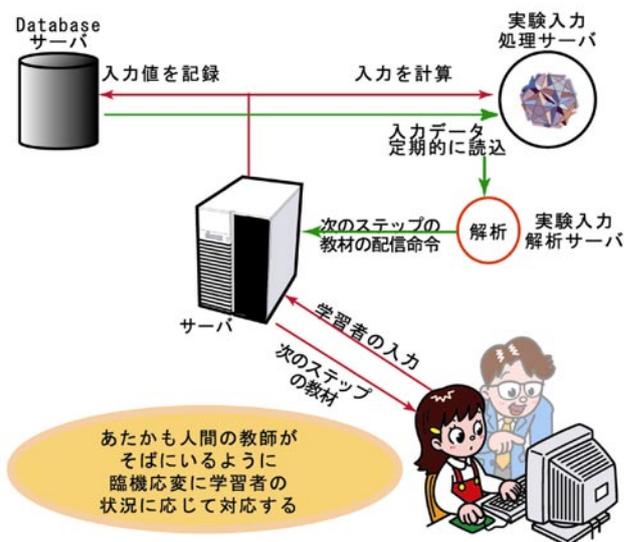


図1 システムの概要

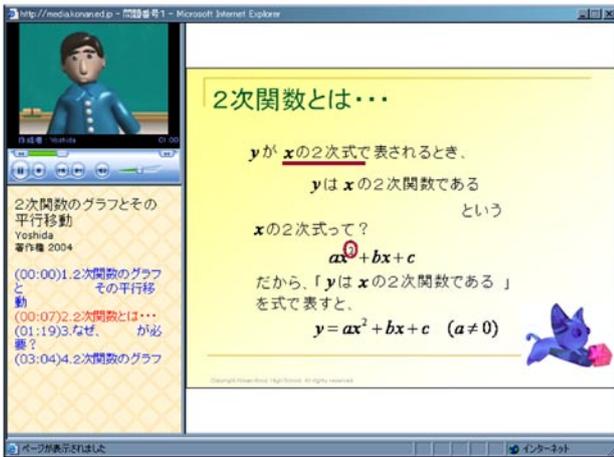


図2 コンテンツ画面

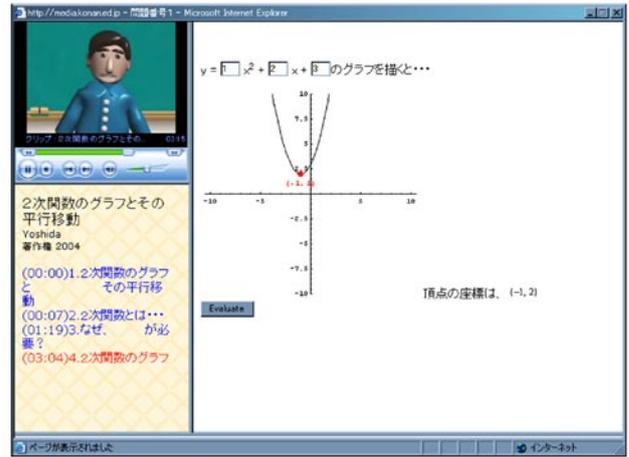


図3 実験入力画面

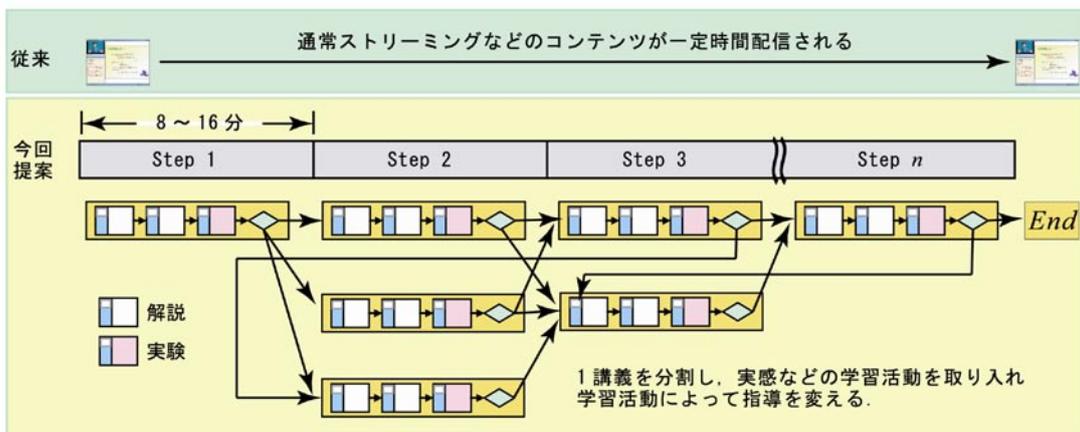


図4 コンテンツ配信の流れ

とその平行移動」を題材にコンテンツを試作した。

今回試作したコンテンツは、各パートのコンテンツの解説時間を5分以内とし、各パートの終わりに実験を促す構成とした(図4)。

3. 入力情報のパターン化と配信制御

3.1 変化差異による単純2値化

さて、学習者の実験入力値から次の指導を決定するために、学習者の入力値をいくつかのパターンに分類する必要がある。

入力させる値が m 箇所あり、 n 回の入力実験をおこなったとする。このとき、 t 回目の入力 $I_t (t = 1, 2, 3, \dots, n)$ を次のようにあらわすことにする。 a_{ti} を実験の入力値として、

$$I_t = (a_{t1}, a_{t2}, a_{t3}, \dots, a_{tm}) \quad (1)$$

とする。次に、 t 回目と $t+1$ 回目の入力値の差を

$$\Delta_t = I_{t+1} - I_t \quad (t \geq 1) \quad (2)$$

とする。

Δ_t の各成分 $\delta_{ti} (\delta_{ti} = a_{(t+1),i} - a_{ti})$ に対し、関数 f を次のように定義する。

$$f(x) = \begin{cases} 1 & (x \neq 0) \\ 0 & (x = 0) \end{cases} \quad (3)$$

また、ベクトル $y = (y_1, y_2, \dots, y_p)$ に対し、 $f(y) = (f(y_1), f(y_2), \dots, f(y_p))$ と定める。この f によって、連続した2回の入力から変化させた箇所を1とし、変化させなかった箇所を0とするベクトルが得られる。

このようにして得られた Δ_t から、次のように、学習者の n 回の入力のパターンを決定する。

$$P_n = (f(\sum_{i=1}^{n-1} \delta_{t1}), f(\sum_{i=1}^{n-1} \delta_{t2}), \dots, f(\sum_{i=1}^{n-1} \delta_{tm})) \quad (4)$$

これによって、入力パターンが生成される。

試作コンテンツでは、 $y = ax^2 + bx + c$ の a, b, c の値をいろいろ変化させてグラフを描かせるという実験をおこなわせている。例えば、3回の a, b, c の入力が $(1, 2, 3), (1, 3, 3), (1, 4, 4)$ である場合、 $(1, 2, 3), (1, 3, 3)$ から $\Delta_1 = (0, 1, 0)$ 、次に $(1, 3, 3), (1, 4, 4)$ から $\Delta_2 = (0, 1, 1)$ を得る。この Δ_1, Δ_2 から

$$P_3 = (f(0), f(2), f(1)) = (0, 1, 1)$$

というパターンを得る。この入力パターン $(0, 1, 1)$ から、 c の値を変化させて試行していると判断する。今回使用

した教員側の指導コンテンツ配信の分類ケースは、以下の5通りである。実験入力パターンが

- ・ (1, 0, 0)の場合、*a*に注目している
- ・ (0, 1, 0)の場合、*b*に注目している
- ・ (0, 0, 1)の場合、*c*に注目している
- ・ (0, 0, 0)の場合、操作の仕方がわからない
- ・ いずれにも当てはまらない場合、どの文字にも注目していない

と判断する。

このそれぞれのパターンに対して、コメントを付け加えたコンテンツが配信され、何度も実験を重ねる構成になっており、最終的に2次関数のグラフの頂点は平方完成することによって求められることを気づかせる発見的学習の要素をもたせている。

ここで、「操作の方法がわからない」と判断された場合、操作方法の解説が配信され、「どの文字にも注目していない」と判断された場合には、一つの文字に注目させグラフの変化に注目させるコンテンツが配信されるようになっている。

また、今回の1回の実験に対する実験の回数は5回とした。その理由として、我々が教室で学習者の実験の様子を観察し指導する際、入力毎に指導するのではなく、何回か試行錯誤している様子を見てから指導を行うからである。そのような指導をシステムの機能として持たせようと試みた。ただ、明らかに間違えた入力をおこなったときには、すぐに指導しなければならない。例えば、今回のコンテンツでは「 $a \neq 0$ 」の解説あるので、 $a = 0$ を入力した段階で、話を聞いていなかったことを注意するコンテンツが配信されるようになっている。

3.2 システム設計の意義と学習者の反応

システムを構築しても学習効果が現れなければ、有効なシステムとはいえない。この学習効果とは、WBTのようにトレーニングによってスキルが向上するというものではなく、学習者の興味や関心、意欲などの内面的変化が認められるということである。

我々は、コンテンツを実際の生徒に試用させ、学習効果を測定する必要があると考えた。学習効果が現れなければ、コンテンツの展開およびシステムの構成を再考しなければならない。

そこで、システムの動作確認を兼ねて、このシステムを中学3年数学の単元「放物線」を学習した中学生希望者21名に試用させた。

中学3年生を選んだ理由は、学習塾などの影響で単元に対する先入観がない状態で情意検査をおこなうためである。

学習到達度を測るために、事前テストおよび事後テストをおこない平均点の推移により学習効果があったかどうかを測定した。そのテストの内容は、 $y = (x - p)^2$ の形

式で*p*が正の場合・負の場合で頂点を求める問題、 $y = x^2 + bx + c$ から頂点の座標を算出する問題の3問とした。

興味や関心、意欲のような成績とは異なった尺度として、島根式数学に対する情意検査のアンケート項目を用い、そのポイントの平均点の差異により情意面のどの点に影響があったかを測った。情意アンケートは、

- ・ 数学における自己概念……………9項目
- ・ 数学における達成動機……………18項目
- ・ 数学への不安……………9項目
- ・ 数学に対する態度……………9項目

の計45項目についておこない、授業は、次のような手順でおこなった。

- ・ 計算テスト (事前テスト) [5分]
- ・ PC操作方法 [10分]
- ・ 情意検査 [15分]
- ・ e-Learning [90分]
- ・ 情意検査 [15分]
- ・ 計算テスト (事後テスト) [5分]

3.2.1 事前事後テストおよび情意アンケート

事前テストおよび事前テストの正答率の平均点の変化は、

事前テスト	平均30.16点
事後テスト	平均90.48点

であった。これは、最後に共通して配信される「まとめ」のコンテンツを終えたところで、正答率が90%以上になるまで、出題が繰り返される小テストを設けていたため、テストの成績が向上したと考えられる。

情意アンケートに関して、45項目全てについて有意水準0.05でt-検定をおこなった。その結果以下の7項目について変化が認められた (表1)。

変化がみられたアンケートの項目は、

- Q2 私は、自分の今の数学の成績に満足している。
- Q6 私は、数学の勉強で、他の人から尊敬されるような人間であると思います。
- Q8 数学の問題を解くとき、失敗するのではないかと恐れたりしません。
- Q11 数学の勉強では、新しい課題を自分から見つけて取り組みます。
- Q15 数学の勉強をきちんとし、数学の能力を身に着けたいと思います。
- Q26 数学のテストの勉強がある時は、他のことができなくても気になりません。
- Q42 数学の勉強は、私に考えることの喜びを与えてくれるので、好きです。

である。

3.2.2 情意アンケートの結果

Q2、Q6に関しては、問題が解けたことへの満足感が、アンケート結果に影響を及ぼしたのではないかと考えられる。Q8は、意外な結果となった。通常授業では、失敗することを「恥ずかしい」と感じ、自らあまり試行錯誤をおこなおうとせず、正解をすぐに尋ね、誰かの解答を待つという受け身の学習が多く見られる。

ところが、今回のコンテンツを受講する前と後で、失敗に対しての意識に変化が見られた。これは、グラフを描くという実験を繰り返し、思い通りにグラフが描けないことを何度も体験したからではないかと考えられる。

Q11に関してもQ8と同様、実験を繰り返しおこなうことで、グラフを描くとき、どのパラメータを変化させると、グラフはどう変化するかなどの推測と検証が学習者自身のなかで繰り返された結果ではないかと考えている。

Q15に関しては、今回のシステムによる学習がマイナスに働いた。これは、通常授業であれば、板書を写すため、授業そのものが形として残る。しかし、今回の試験運用では、ノートを取らず画面を見ながら学習を進めた。そのため、学習後に形として残るものが無く、このようなアンケート結果になったのではないかと推測される。教育的な配慮として、コンテンツを見るだけでなくノートを併用し、解説や実験内容をメモさせる必要があったのではないかと考えている。

Q26については、実験を繰り返すという学習形態がゲーム感覚に似ているため、集中して取り組めたのではないかと考えられる。Q42については、通常授業では受

け身になりがちであるが、今回のシステムの場合、実験を繰り返しおこなわせ、法則を見つけ出させるように指導コンテンツが配信されたためではないかと考えている。

学習者一人ひとりの実験によって配信されるコンテンツが異なるため、隣の学習者同士で「なぜ、そのような解説が出たの？」という議論が展開される場面も見られた。これは、Role Playing Gameのように学習者個々の思考に基づいたストーリー展開でコンテンツが配信されるため、学習者は面白さを感じていたように思われる。

今回のシステムの試用結果から学習者は、主体的に学習する機会が与えられ、能動的に学習を進めることができたと思われる。またそれにより、学習者の情意面における変化が認められた。

この結果から、e-Learningが学習者の情意に働きかけがあったと判断できる。

さて、中学生21人の実験入力では、正確な配信制御がおこなわれたがどうか検討するには不十分である。

4. 単純2値化の問題点

そこで、より適正な配信制御の方法の検討のために、新たに高校1年生114人に同様のコンテンツを試用させ、一人ひとりの実験入力データをもとに制御方法を検討した。

3.1節の5つのケースにパターン化し配信制御する場合、図5のように、5回の a, b, c の入力 $I_1 = (1, 2, 3)$ 、 $I_2 = (2, 2, 3)$ 、 $I_3 = (2, 3, 4)$ 、 $I_4 = (1, 2, 3)$ 、 $I_5 = (1, 0, 0)$ に対し、

$$I_1, I_2 \text{ から}$$

$$\Delta_i = I_2 - I_1 = (1, 0, 0)$$

$$f(\Delta_1) = (f(1), f(0), f(0)) = (1, 0, 0)$$

同様に、 I_2, I_3 から $\Delta_2 = (0, 1, 1)$ 、 I_3, I_4 から $\Delta_3 = (1, 1, 1)$ 、 I_4, I_5 から $\Delta_4 = (0, 1, 1)$ を得る。

この得られた、 $\Delta_i (i=1, 2, 3, 4)$ から

表1 情意アンケート結果

	Q 2		Q 6		Q 8	
	前	後	前	後	前	後
平均	2.19	2.81	1.81	2.48	3.00	3.76
分散	1.17	1.03	0.81	1.12	1.26	1.04
自由度	20	20	20	20	20	20
確率	0.006		0.012		0.032	

	Q 11		Q 15		Q 26	
	前	後	前	後	前	後
平均	1.86	2.95	4.38	3.76	2.24	2.95
分散	0.85	0.97	0.74	1.18	1.34	1.28
自由度	20	20	20	20	20	20
確率	0.001		0.012		0.048	

	Q 42	
	前	後
平均	2.33	2.76
分散	0.86	0.70
自由度	20	20
確率	0.035	

	実際の入力値			パターン生成			
	a	b	c	変化	a	b	c
1回目	$y = \boxed{1}x^2 + \boxed{2}x + \boxed{3}$			1回目	1	0	0
2回目	$y = \boxed{2}x^2 + \boxed{2}x + \boxed{3}$			→2回目			
3回目	$y = \boxed{2}x^2 + \boxed{3}x + \boxed{4}$			2回目	0	1	1
4回目	$y = \boxed{1}x^2 + \boxed{2}x + \boxed{3}$			→3回目			
5回目	$y = \boxed{1}x^2 + \boxed{0}x + \boxed{0}$			3回目	1	1	1
				→4回目			
				4回目	0	1	1
				→5回目			
				パターン	1	1	1

図5 パターンの決定

$$\sum_{i=1}^4 f(\Delta_i) = (2, 3, 3)$$

$$P_5 = (f(2), f(3), f(3)) = (1, 1, 1)$$

という実験入力パターンが得られ、 a, b, c の全ての値を変化させて実験していると判断する。

しかし、例えば $I_1 = (1, 2, 3), I_2 = (2, 4, 3), I_3 = (1, 4, 4), I_4 = (1, 4, 5), I_5 = (1, 4, 6)$ の場合、教員は3回目からパラメータ c に注目し始めていると判断するが、システムから得られるパターンは、 $(1, 1, 1)$ となり、注目しているパラメータはないと判断される。これでは、教員が目的とする指導になっていない。

そこで、教員が入力データを見て、どのような実験をしているかを教員自身が判別した(表2)。これを、「教員が判別した結果」と呼ぶことにする。表2において分類したカテゴリを数値で表した。教員の判別と数値の対応は以下のような対応である。

- a に注目していると判別……………1
- b に注目していると判別……………2

表2 学習者114人の入力パターンと教員の判別

1→2			2→3			3→4			4→5			教員の判別
a	b	c	a	b	c	a	b	c	a	b	c	
0	1	1	0	1	1	0	0	1	0	1	1	4
1	0	0	0	1	0	0	0	1	1	1	1	4
1	0	1	1	1	0	1	0	1	0	0	1	4
1	1	1	1	0	0	0	1	0	0	1	0	2
1	0	0	0	1	0	0	1	0	0	0	0	2
1	1	1	1	0	0	1	1	0	0	1	1	4
0	0	0	0	1	0	1	0	1	1	1	1	4
0	1	0	0	1	1	0	0	1	0	0	1	3
0	1	0	0	1	0	0	0	1	0	1	1	2
1	0	0	0	0	0	1	0	0	0	1	0	1
0	0	0	0	0	0	0	0	0	0	0	0	5
0	1	1	1	0	0	1	1	1	1	1	0	4
1	1	1	1	1	0	1	0	0	0	1	0	4
0	1	1	0	0	0	1	0	0	1	0	0	1
1	1	1	0	0	0	0	0	0	0	0	0	5
0	0	0	0	1	0	0	0	1	1	0	0	4
0	0	0	0	0	0	0	0	0	0	0	0	5
1	1	1	0	1	1	0	1	1	1	0	1	4
0	0	0	0	1	1	1	1	0	0	0	0	4
0	0	0	0	0	0	1	1	1	0	0	0	5
0	0	0	1	1	1	1	0	0	1	0	0	1

- c に注目していると判別……………3
- 注目している文字がないと判別 ……4
- 操作の方法がわからないと判別 ……5

教員の判別の基準として連続して2回以上1種類のパラメータを変化させたとき、そのパラメータに注目していると判断した。また、注目しているパラメータが5回の入力のうち前半と後半で異なる場合は、後半の入力パターンを実験のパターンとした。また、変化させるパラメータが全て0あるいは、0を多く含む場合、操作の方法がわからないという判別基準により分類がおこなった。

これに対し、3.1節の方法で判別した結果を「サーバの判別結果」と呼ぶことにする。学習者の入力情報を整理し、教員が判別した結果とサーバが判別した結果を比較すると、サーバの判別結果に対し教員側の判別と一致しない配信が67件存在し、制御対象全体のデータ数に対する誤配信率は58.77%であった(表3)。

「注目文字なし」の誤配信率が0.00%であることから、本来「注目文字なし」ではない実験入力に対して「注目文字なし」と判別しているケースが多く含まれていると考えられる。

そこで、「注目文字なし」に実験していると分類された中から、システムが誤って分類したと考えられるパターンを調べた(表4、5)。

学習者AとBにおいてシステムによる判別は、ともに「注目文字なし」の実験をしていると判断されているが、学習者Bは、3回目から1つのパラメータ c に注目して実験している。この場合、グラフが y 軸方向に平行移動していることに気づいていると判断し、グラフが y 軸方向に平行移動に関する解説を配信することが適切である。

一方、学習者Aは、 c 以外のパラメータも同時に変化

表3 単純2値による判別と誤配信率

分類カテゴリ	教員による判別	システムの誤配信	誤配信率(%)
aに注目	13	12	92.31
bに注目	9	8	88.89
cに注目	1	1	100.00
注目文字なし	45	0	0.00
操作方法不理解	46	46	100.00
全体	114	67	58.77

表4 5回の実験入力

学習者	1回目			2回目			3回目			4回目			5回目		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
A	1	2	3	2	2	3	2	3	4	1	2	3	1	0	0
B	1	2	3	2	4	3	1	4	4	1	4	5	1	4	6

表5 5回の実験入力のパターン化

学習者	1 → 2			2 → 3			3 → 4			4 → 5			Pattern			判別	
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c		
A	1	0	0	0	1	1	1	1	1	0	1	1	1	1	1	1	注目文字なし
B	1	1	0	1	0	1	0	0	1	0	0	1	1	1	1	1	注目文字なし

表6 変化頻度によるパターン化

学習者	1 → 2			2 → 3			3 → 4			4 → 5			Pattern		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
A	1	0	0	0	1	1	1	1	1	0	1	1	0.50	0.75	0.75
B	1	1	0	1	0	1	0	0	1	0	0	1	0.50	0.25	0.75

させているため、法則を発見するためには一つの文字に注目するよう解説を配信することが適切である。このように、単に変化させた箇所に注目するだけでは、正しく学習者の実験を判断することができない。

4.1 分岐方法の改善

そこで、より多く変化させたパラメータを取得するため、変化の頻度を考慮するように式(4)を次のように変形する。

$$P_n = \left(\frac{1}{n-1} \sum_{i=1}^{n-1} f(\delta_{t,i}), \frac{1}{n-1} \sum_{i=1}^{n-1} f(\delta_{t,2}), \dots, \frac{1}{n-1} \sum_{i=1}^{n-1} f(\delta_{t,m}) \right) \quad (5)$$

式(5)によって、変化の度合いによって、実験のパターンを詳細に分類することができる。表5のパターンは、表6となる。

このように変形して得られたデータをもとにk-NNを用いて判別をおこなった。判別には、a, b, cについて注目していると判断する(1, 0, 0)、(0, 1, 0)、(0, 0, 1)のパターン、すべてパラメータを変化させ注目している文字がないとするパターン、操作方法がわからないとする(0, 0, 0)の5種のパターンを用意した。

4.1.1 k-NN法について

パターン P_i に対して、全標準パターン $T_1 \cup T_2 \cup T_3 \cup \dots \cup T_n$ の中から P_i に最も距離の短い順に k 個のパターンを選び、その集合を $\{p_1, p_2, \dots, p_k\}$ とする。これを P_i の k -最近傍という。 k -最近傍の要素がそれぞれどのパターン

に属するか調べ、パターン T_n に属する要素の数を $v(n)$ とする。

$$v(n) = |\{p_1, p_2, \dots, p_k\} \cap T_n|$$

この数を用いて多数決によって P_i のパターンを決定する。すなわち、

$$\max\{v(1), v(2), \dots, v(k)\} = v(m)$$

ならば、 P_i は T_m に属すると定める。 $\{v(1), v(2), \dots, v(k)\}$ に最大値が複数個存在するときには、最も距離の近い標準パターンを選ぶことにする。

(1, 0, 0)、(0, 1, 0)、(0, 0, 1)、(0, 0, 0)、(1, 1, 1)を標準パターンとして、k-NN法により分類を試みた。

4.1.2 k-NN法による配信制御

k-NN法による判別の誤配信率は、表7のように、 $k=4$ のとき35.09%となり、精度が向上した。

さらに、配信精度を上げるため、再び学習者の入力パターンと配信結果を調査した結果、表8、9から学習者Bと学習者Cの差異を正しく配信制御することができないことが判明した。学習者Bは、3回目から1つのパラメータ“c”に注目しているといえるが、学習者Cは、3回目から4回目にかけてパラメータ“a”および“c”を変化

表7 k-NN法による誤配信率 (変化頻度)

	k=1	k=2	k=3	k=4	k=5
誤配信率(%)	66.67	66.67	36.84	35.09	47.37

表8 学習者実験入力データ

学習者	1回目			2回目			3回目			4回目			5回目		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
B	1	2	3	2	4	3	1	4	4	1	4	5	1	4	6
C	2	0	0	1	1	0	1	1	1	2	1	-1	2	1	1

表9 変化頻度によるパターン化

学習者	1→2			2→3			3→4			4→5			Pattern		
	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
B	1	1	0	1	0	1	0	0	1	0	0	1	0.50	0.25	0.75
C	1	1	0	0	0	1	1	0	1	0	0	1	0.50	0.25	0.75

表10 教員が作成したトレーニングデータ

教員が作成したパターン				教員の判別
1→2	2→3	3→4	4→5	
a b c	a b c	a b c	a b c	
1 0 0	1 0 0	1 0 0	1 0 0	1
1 1 0	1 0 0	1 0 0	1 0 0	1
:				
1 1 1	1 0 1	1 0 0	1 0 0	1
0 1 0	0 1 0	0 1 0	0 1 0	2
1 1 0	0 1 0	0 1 0	0 1 0	2
:				
1 1 1	0 1 1	0 1 0	0 1 0	2
0 0 1	0 0 1	0 0 1	0 0 1	3
1 0 1	0 0 1	0 0 1	0 0 1	3
:				
1 1 1	0 1 1	0 0 1	0 0 1	3
:				
0 0 0	0 0 0	0 0 0	0 0 0	5

表11 k-NN法による誤配信率 (入力順序を考慮)

	k=1	k=2	k=3	k=4	k=5
誤配信率 (%)	61.40	61.40	46.49	40.35	36.84

させていたため、一つのパラメータ“c”に注目して実験しているとは言い難い。

そこで、実験の順序を考慮して配信制御をおこなうため、データの変化頻度による分類ではなく4回の変化を式(6)のように表す。また、表10のようにトレーニングデータを教員自身が自らの経験をもとに634生成し、それを標準パターンとして分類を試みた。

$$v = (f(\delta_{1,1}), \dots, f(\delta_{1,m}), f(\delta_{2,1}), \dots, f(\delta_{2,m}), \dots, f(\delta_{(n-1),1}), \dots, f(\delta_{(n-1),m})) \quad (6)$$

配信結果は、表11のようにk=5のとき誤配信率が最低となったが、表7に比べ精度を下げる結果となった。

4.2 SVMによる配信制御

k-NN法では、パターンの分布によって最適なkの値を定めることが困難である。そこでSVM (Support Vector Machine) による分類を試みた。

表12 SVMトレーニングデータ

x_i				分類 カテゴリ	$D_1(z)$ y_i	...	$D_s(z)$ y_i
1→2	2→3	3→4	4→5				
a b c	a b c	a b c	a b c				
1 0 0	1 0 0	1 0 0	1 0 0	1	+1	...	-1
1 1 0	1 0 0	1 0 0	1 0 0	1	+1	...	-1
:							
1 1 1	1 0 1	1 0 0	1 0 0	1	+1	...	-1
0 1 0	0 1 0	0 1 0	0 1 0	2	-1	...	-1
1 1 0	0 1 0	0 1 0	0 1 0	2	-1	...	-1
:							
1 1 1	0 1 1	0 1 0	0 1 0	2	-1	...	-1
0 0 1	0 0 1	0 0 1	0 0 1	3	-1	...	-1
1 0 1	0 0 1	0 0 1	0 0 1	3	-1	...	-1
:							
1 1 1	0 1 1	0 0 1	0 0 1	3	-1	...	-1
:							
0 0 0	0 0 0	0 0 0	0 0 0	5	-1	...	+1

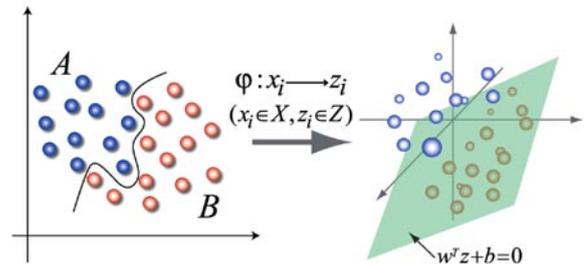


図6 原空間Xから高次元特徴空間Zへの非線形写像

4.2.1 SVMについて

n次元実空間上に2つのクラスA、Bがあり、このいずれかに属するようなデータ $x_i (i=1, 2, 3, \dots, m)$ があるとす。

クラスAの要素 x_i に対して $y_i = +1$ を与え、クラスBの要素 x_j に対して $y_j = -1$ を与えることにする。ここで、クラスAとBを超平面で分離することを考える。この2つのクラスが線形分離不可能な場合は、原空間Xからある高次元特徴空間Zへの非線形写像、

$$\phi: x \rightarrow z (x \in X, z \in Z)$$

によって図6のように線形分離可能な状態にする。

このとき、Zにおける分離超平面を

$$D(z) = w^T z + b = 0$$

とする。AとBを分離する超平面は一般に数多く存在するが、SVMではマージン (分離超平面と各クラスとの最短距離) を最大化することを目的とする。クラスA、

Bはそれぞれ

$$w^T z + b \geq 1 \quad (z_i \in \varphi(A)) \quad (7)$$

$$w^T z + b \geq 1 \quad (z_i \in \varphi(B)) \quad (8)$$

を満たすとすると、SVMは次のような数理計画問題として定式化できる。

$$\begin{aligned} \text{[SVM] Minimize} \quad & \|w\|_q \\ \text{Subject to} \quad & y_i D(z_i) \geq 1 \\ & (i = 1, 2, \dots, m) \end{aligned} \quad (9)$$

ここで、問題(9)の制約条件において、等号を成立させるような z_i をサポートベクターという。また、マージンの大きさを測る距離として、 l_2 ノルムを用いると問題(9)は2次計画問題(10)として定式化される。

$$\begin{aligned} \text{[QPSVM] Minimize} \quad & \frac{1}{2} w^T w \\ \text{Subject to} \quad & y_i D(z_i) \geq 1 \\ & (i = 1, 2, \dots, m) \end{aligned} \quad (10)$$

ここで、 l_2 ノルムによる2次計画問題(10)は、計算を容易にするため、その双対問題を取り扱う。問題(10)の双対問題は以下のように定式化することができる。

[QPSVM]

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2} \sum_{i,j=1}^m y_i y_j \alpha_i \alpha_j z_i^T z_j - \sum_{i=1}^m \alpha_i \\ \text{Subject to} \quad & \sum_{i=1}^m y_i \alpha_i \geq 1 \\ & \alpha_i \geq 0 \quad (i = 1, 2, \dots, m) \end{aligned} \quad (11)$$

ここで、問題(11)において、目的関数に $z_i^T z_j = \varphi(x_i)^T \varphi(x_j)$ が存在するが、内積演算を保存するようなカーネル関数で置き換えることにより計算の労力を削減することができる。これをカーネルトリックという。カーネル関数は

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) = z_i^T z_j \quad (12)$$

で定義される。ただし、カーネル関数はマーサーの定理を満たすものとする。このカーネル関数により、 $D(z)$ は

$$D(z) = \sum_{i=1}^m y_i \alpha_i K(x_i, x) + b = 0 \quad (13)$$

と変形することができる。

本研究では、カーネル関数 $K(x_i, x)$ を

$$K(x_i, x) = \exp\left\{-\frac{\|x_i - x\|_2^2}{r}\right\} \quad (14)$$

とした。ただし、 r を任意の定数とする。

また、表10のように、想定される学習者の入力パターンとそれに対する教員の判別による分類カテゴリをトレーニングデータとして用意する。

学習者の入力パターン p に対して、カテゴリ k とその他のカテゴリを判別する関数を

$$D_k(\varphi(p)) = f_k(p) = \sum_{i=1}^m y_i \alpha_i K(x_i, p) + b \quad (15)$$

とする。この $f_k(p)$ が最大となるカテゴリ k に p が属すると判定する。

表13 SVM判別に使用したデータ数

分類カテゴリ	教員による判別	システムの誤配信	誤配信率(%)
aに注目	13	8	61.54
bに注目	9	7	77.78
cに注目	1	1	100.00
注目文字なし	45	0	0.00
操作方法不理解	46	13	28.26
全体	114	29	25.44

4.2.2 SVMによる分類結果

SVMでは、分離超平面を決定する際、トレーニングデータを必要とする。そこでk-NN法と同じトレーニングデータ(表10)を使用し判別をおこなった。

SVMによる配信制御の結果は、表13に示すとおり、配信精度が向上したといえる。

しかし、「注目文字なし」の誤配信率が0.00%であることから、他のカテゴリに判別されるべきものを含んでいると考えられる。

つまり、「注目文字なし」と他のカテゴリを分離する平面 $f_k(p) = 0$ が適切ではないと考え、 $k = 4$ における式(15)を式(16)のように ε だけ平行移動させ、実際の学習者の実験入力データを用いて、配信精度を上げるための微調整をおこなった。(図7、図8)。

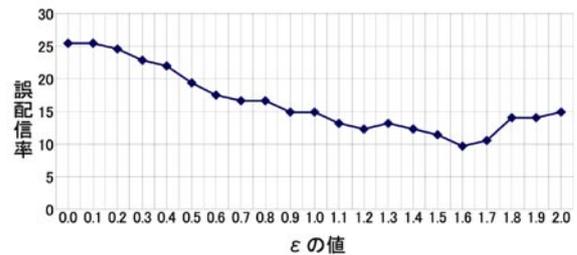


図7 分離超平面を ε だけ平行移動したときの全体の誤配信率

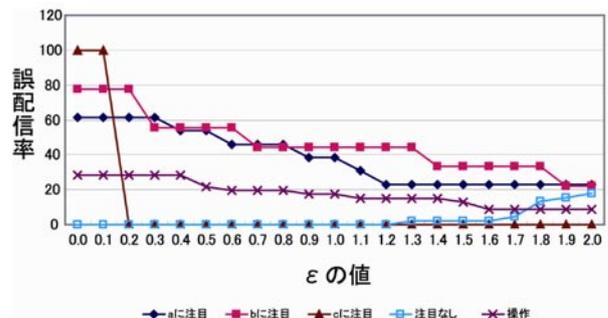


図8 分離超平面を ε だけ平行移動したときのカテゴリ別の誤配信率

表14 SVMによる誤配信率 ($\epsilon=0.0$, $\epsilon=1.6$)

分類カテゴリ	$\epsilon = 0.0$		
	教員による判別	システムの誤配信	誤配信率(%)
aに注目	13	8	61.54
bに注目	9	7	77.78
cに注目	1	1	100.00
注目文字なし	45	0	0.00
操作方法不理解	46	13	28.26
全体	114	29	25.44

分類カテゴリ	$\epsilon = 1.6$		
	教員による判別	システムの誤配信	誤配信率(%)
aに注目	13	3	22.08
bに注目	9	3	33.33
cに注目	1	0	0.00
注目文字なし	45	1	2.22
操作方法不理解	46	4	8.70
全体	114	11	9.65

$$D_4(\varphi(\hat{p})) = f_4(\hat{p}) = \sum_{i=1}^m y_i \alpha_i K(x_i, \hat{p}) + b + \epsilon \quad (16)$$

その結果、 $\epsilon=1.6$ において、全体の誤配信率をさらに下げることができた(表14)。

5. 終わりに

今回は、高校の数学教材を題材に配信実験を行ったが、試行錯誤という正解・不正解のように2値的に分類することができない実験入力に対し、SVM法を用いることによりシステム自身が自動的に学習者の実験内容を把握しその学習者に適したコンテンツを配信させる機能の組み込みが可能であることがわかった。この機能は、実験的数学だけでなく、情報検索の指導において検索語の選び方によって指導方法を変えたコンテンツを配信したり、アンケートなどにより学習者の得意・不得意を取得して、それらを考慮して適切な指導コンテンツを配信するなど、他教科への応用も考えられる。

コンテンツ配信制御の手法に関してはSVMを用いた配信制御が最も精度がよかった。その理由として、k-NN法の場合、学習者の実験入力を平均化することによって分類した結果、いつ変化させたのかという時間的変化の情報が失われたからと考えられる。今回のシステムのように、学習者の実験入力情報から指導コンテンツを決定する手法としてSVM法が有効であるといえる。

しかしながらSVMの場合、分離超平面を決定するために多くのトレーニングデータ(配信の理想値)が必要となる。つまり、トレーニングデータによって最適化さ

れるため、学習者の全パターンをトレーニングデータとして用意すれば、誤配信は0%となる。しかし、今回の実験入力におけるパターンは 2^{15} 通りあり、すべてを列挙することは不可能である。今回の実験では634のトレーニングデータを教員が用意し、決定された分離超平面を平行移動することで微調整をおこなったが、少ないトレーニングデータからいかに追加学習させ精度を上げられるかが今後の課題である。

謝辞

本研究の一部は、日本文部科学省オープン・リサーチ・センター整備事業(平成16年~平成20年)による私学助成を得ておこなわれた。

(平成17年6月15日受付)

参考文献

- [1] 荒川雅裕、植木泰博、冬木正彦、“授業支援型e-LearningシステムCEASを活用した自発学習促進パイラル教育法”、日本教育工学会 Vol. 28, No. 4, pp.311-321, 2005
- [2] 安藤雅洋、植野真臣、“アイマークレコーダを用いたeラーニングのコンテンツ評価”、教育システム情報学会研究報告 Vol. 19, pp.11-18, 2004
- [3] 伊藤俊彦、“島根式数学に対する情意的特性検査(Shimane-ACTM)について”、全国数学教育学会誌、数学教育学研究第1号、pp.93-99, 1995
- [4] 大橋志津江、村田尚志、一楽重雄ほか、“高等学校の教育のあり方”、日本数学教育学会第85巻、第5号、pp.20-29, 2003
- [5] 金西計英、林賢太郎、光原弘幸、矢野米雄、“教材知識に基づきWBT上で演習問題を生成する機能の実現”、教育システム情報学会 Vol. 20, No. 2, pp.71-82, 2003
- [6] 國宗永佳、新村正明、和崎克己、不破 泰、師玉康成、中村八東、“学習進捗状況に基づいたe-Learning教材表示制御システム”、教育システム情報学会研究報告 Vol. 19, pp.43-48, 2004
- [7] 中上香代子、中山弘隆、“ファジィサポートベクターマシンによる多値分類について”、第23回計測自動制御学会九州支部学術講演会予稿集、pp.129-130, 2004
- [8] 中上香代子、吉田賢史、中山弘隆、“Support Vector Machineによるe-Learningの適正コンテンツ決定”、第37回計測自動制御学会北海道支部学術講演会論文集、pp.111-114, 2005
- [9] 長崎栄三、瀬沼花子、“OECD生徒の学習到達度調査2003年調査の国際結果-15歳児の数学的リテラシー-”、日本数学教育学会第87巻、第1号、pp.17-26, 2005
- [10] 宮地 功、姚 華平、吉田幸二、“e-Learningにおける問題自動生成機構の構想”、教育システム情報学会研究報告 Vol. 19, No. 6, pp.99-103, 2005
- [11] 森田正康、“eラーニングの〈常識〉”、朝日新聞社、東京都、pp.78-90, 2002



よしだ けんじ
吉田 賢史

1997年甲南大学自然科学研究科情報・システム科学専攻修士課程修了
現在、甲南大学自然科学研究科情報・システム科学専攻博士課程、甲南高等学校・中学校教諭
現在の研究：機械学習の理論と応用、教育システム工学、知能情報学、数学教育、情報教育
所属学会：コンピュータ利用教育協議会、教育情報学会、情報処理学会、電子通信情報学会、計測自動制御学会、e-Learning学会、人工知能学会、教育システム情報学会、教育情報学会、日本教育工学会、日本数学教育学会



みやざき こうじ
宮崎 光二

2000年甲南大学大学院自然科学研究科情報システム科学専攻修了（理学博士）
現在、(株)ニルバーナテクノロジー勤務、甲南大学、大阪商業大学、関西学院大学・非常勤講師
現在の研究：ソフトウェアコンピューティング、シミュレーション、知能情報学、情報通信ネットワーク
所属学会：計測自動制御学会



なかがみ かよこ
中上 香代子

2003年甲南大学理学部応用数学科卒、2005年甲南大学大学院自然科学研究科情報・システム科学専攻修了
現在、甲南高等学校・中学校非常勤講師、甲南大学理工学部研究生
現在の研究：多目的最適化、機械学習の理論と応用、知能情報学、数学教育、情報教育



なかやま ひろたか
中山 弘隆

1974年京都大学大学院工学研究科博士課程数理工学専攻修了（工学博士）
現在、甲南大学 理工学部 情報システム工学科教授
現在の研究：多目的意思決定の理論と応用、機械学習の理論と応用、人工生命とその応用
所属学会：計測自動制御学会、システム制御情報学会、日本オペレーションズリサーチ学会、電子情報通信学会、International Neural Network Society, IEEE

Control of adaptive presentations of instructive contents of e-Learning systems

Kenji Yoshida¹⁾ · Koji Miyazaki²⁾ ·
Kayoko Nakagami³⁾ · Hiroataka Nakayama³⁾

It is important for students to learn through trial and error. However, in actual classes it is difficult to utilize an experimental learning process because it takes a lot of time. E-Learning systems are expected to provide good learning environments, especially for slow learners. But the existing e-Learning system cannot control such input data.

Therefore, we tried to build an e-learning system which emphasizes the trial and error process. This system delivers the appropriate contents to the student every trial. We used the contents of mathematics for senior high school students. And we tried different methods based on students' inputs and examined the results.

This paper discusses how to serve the appropriate contents to each learner depending on the situation.

Keywords

Experimental Mathematics, Heuristic Learning, Support Vector Machine (SVM), Web Based Training (WBT), Learning Management System (LMS)

¹⁾ Department of Information Science and Systems Engineering, Graduate School of Natural Science, Konan University

²⁾ Nirvana Technorogy Inc.

³⁾ Department of Information Science and Systems Engineering, Faculty of Science and Engineering, Konan University