

LMSの活動ログを利用した学生の分類の試み：情報系科目をケーススタディとして

石井 純平¹ 平川 正人^{2,a)} 山田 泰寛^{2,b)}

概要：Learning Management System(LMS)を用いて行う反転授業では、資料の閲覧や小テストの受験、課題の提出等の授業外時間の学習がLMSのログとして蓄積される。反転授業において授業外時間の学習は重要であるが、学習時間や頻度は学習者によって異なるため、授業外時間の取り組みが成績に影響を与えられ考えられる。本研究では、LMSの活動ログを用いて、学習者を学習傾向に応じて分類することを目的に、一つの情報系科目をケーススタディとして、二つの手法による学生の分類を試みた。一つ目の手法は、学習の反復と学習のペースという二つの指標を用いて、学習意欲という観点に基づく分類であり、その分類結果と成績との関連を調べた。二つ目の手法は、特定の行動の行動回数や2つの行動の順序などの指標を説明変数とし、成績を目的変数として、決定木による分類を行った。

キーワード：LMS, ラーニング・アナリティクス, 反転授業, ログ解析, アクティブラーニング

1. はじめに

近年、教育機関における教育方法の一つとして、アクティブラーニングが注目されている。文部科学省はアクティブラーニングを、「教員による一方的な講義形式の教育とは異なり、学修者の能動的な学修への参加を取り入れた教授・学習法の総称」と定義している [1]。

アクティブラーニングを実施する手段の一つとして、反転授業という授業形態が注目されており、教育機関での導入事例も増えてきている。反転授業では、学習者は基礎的な内容の学習を事前に行い、授業時間ではより発展的な内容の学習を行う。学習者は、自分のペースで事前学習を行ったり、必要な時に反復して学習を行うことができるなどの利点がある。しかし、これまで授業時間を使って学習していた内容を事前に個々で学習する必要があるため、学習者の負担がより大きなものとなったり、学習者の取り組みによっては学習者の理解に差が生じる可能性がある。このため、教師側にもよりきめ細やかな対応が求められることになる。

反転授業では、Learning Management System(LMS)を用いて、授業時間外の学習を行うことがある。LMSとは、e-ラーニングに必要な教材などを管理するシステムのことで、例として、Moodle, BookLooper, Blackboard等が挙げられる。LMSでは、資料の保存、配布や、テストの実施、課題の提出、動画の配信などを管理することができる。

LMSには、学習者の活動がログとして蓄積される。LMSのログを分析することで、授業の理解度が低い学習者や脱落する学習者の発見、成績区分ごとの特徴などが捉えられれば、教師が学習者を指導する一助になる。

筆者らは、LMSにおける学習者の活動ログを分析することで、学習者を複数のグループに分類すること、脱落する学習者の発見、全ての授業を終了した時に不合格となる学習者の発見、成績区分ごとの学習行動の特徴を抽出することを目指している。特に、授業の早い段階で不合格となる学習者が予測できれば、教師がきめ細かい対応を取ることが可能になる。本稿では、一つの情報系科目をケーススタディとして、Moodleのログを用いて二つの手法によって学習者の分類を試みた。Moodleのログには、時間、ユーザ名、学習行動などが保存されている。

一つ目の手法は、学習意欲という観点に基づいて学習者の分類を行った。初めに、Moodle上の各トピックにおける資料の閲覧やテストの受験について、学習の反復に着目して学習者を分類した。次に、資料の閲覧やテストの受験について、学習のペースに着目して学習者の分類を行い、

¹ 島根大学総合理工学部
Interdisciplinary Faculty of Science and Engineering, Shimane University, Shimane 690-8504, Japan

² 島根大学大学院総合理工学研究科
Interdisciplinary Graduate School of Science and Engineering, Shimane University, Shimane 690-8504, Japan

a) hirakawa@cis.shimane-u.ac.jp

b) yamada@cis.shimane-u.ac.jp

成績との関連を調べた。二つ目の手法は、資料の閲覧とテストの受験回数や2つの行動の順序に関する指標を説明変数とし、成績を目的変数として、決定木によって学習者の分類を行った。

本稿は以下のように構成される。2節で、関連研究について述べる。3節では、本研究で対象とする授業で用いられたLMSであるMoodleについて述べる。4節では、本研究で対象とする授業の概要について述べる。5節では、学習意欲という観点に基づいた学習者の分類について述べる。6節では、決定木を用いた学習者の分類について述べる。7節で考察について述べ、8節において、まとめと今後の課題について述べる。

2. 関連研究

中野らは、Webアプリケーションサーバソフトのログを分析する前処理として、ログから学習者のデータを抽出する方法やデータベースの活用について検討している [2]。

データマイニングやテキストマイニングを用いて、LMSにおける学習者のデータを分析する研究が発表されている。廣川らは、Booklooperの学習履歴を複数の観点から分析するためにクロス表検索エンジンを実装している [3]。また、MindMapシステムを用いて、コメント中のキーワードとその関連を可視化している。小柏らは、Moodle上で実施した小テストの得点をもとに学生を複数のクラスに分類し、各クラス内の学生の問題ごとの平均点を可視化している [4]。同じクラスに属する学生は、小テストの問題において類似した誤りをしていると考えられる。また、各クラスの特徴の可視化によって、教師が学生の理解状況を容易に把握できるようになると結論づけている。

次に、LMSのログを分析することで学習者の特徴を捉える研究がある。山川は、一つの授業を対象に、LMSのログから得られる、ビデオ授業視聴、スライド閲覧、確認テスト受験という三つの学習行動の回数から学生を複数のクラスに分類している [5]。レポートや試験のような成績を決める要素について、各クラスにおける平均点を比較することで、各クラスの特徴や学習者の行動特性の仮説を立てている。土橋らは、一つの授業を対象に、Moodle上の教材へのアクセス数と授業時間に行う小テストの相関を調べ、教材の学習効果を検証している [6]。Moodleのログを分析する研究としては、[7-9]などの研究がある。

本稿では、Moodleのログを用いて、学習行動の反復やペース、学習行動の回数や2つの行動の順序に着目して学習者の分類を行う。

3. Moodle

本稿では、LMSであるMoodleのログを利用する。Moodleにはコース（授業）が設置され、1コース内は週ごと、あるいは項目（トピック）ごとに分けられたセクションに

```
2013/10/15 18:7:49,User1,cource view,Topic 2
2013/10/15 21:13:22,User1,resource view,Topic 2
2013/10/16 22:27:11,User1,quiz attempt,Topic 2
2013/10/16 17:48:21,User2,quiz attempt,Topic 2
2013/10/16 17:59:20,User2,quiz view summary,Topic 2
```

図 1 Moodle のログ

Fig. 1 Moodle log data

よって構成される。学習者は、Moodle上に設置されたコースを受講し、トピック毎に用意された資料を閲覧したり、テストを受験することで、授業の予習や復習を行うことができる。また、学習者が課題の提出を行うこともできる。

図 1 は、Moodleにおける学習者のログの例である。カンマ区切りで左から時間、ユーザ名、操作（行動）、備考を表している。

4. 対象授業

本稿では、鳥根大学において2013年度後期に開講された情報系科目の「ヒューマン・コンピュータ・インタラクション」を対象とした。授業の受講者は88名であり、Moodleを利用して学習を行った学生は76名であった。期末試験を受けていない学生は未修として扱われる。この授業は反転授業を行っており、次の授業までに、Moodle上で提供されている資料やテストを利用して予習をする必要がある。教室での授業は座学及び演習形式であり、授業中にはMoodleを使用しない。また、15回の授業の後に期末試験が行われる。

Moodle上のコースは、トピック2から15の計14個のトピックから構成されている。それぞれのトピックが1回の授業に対応し、各トピックごとに授業資料が保存されている。また、各トピックに用意されているテストは自身の習熟度を確認するためのものであり、強制ではなく成績評価には加味されない。テストは学生が好きな時に、何度でも受験することができる。学生は学習行動として、Moodle上で資料の閲覧(pdfによる授業資料もしくはビデオの閲覧)かテストの受験を行う。

成績は、グループによる1回のレポートと実習活動内容、個人による2回のレポート及び期末試験の総合計によって算定される。

5. 学習意欲による学生の分類

本節では、Moodleのログを学習意欲という観点から分析し、学生の分類を試みる。学習意欲の低い学生を早期に発見したり、不合格となる学生の特徴を把握することができれば、教師はその学生に対応することができる。反転授業において、学生は事前学習を各人の状況に応じた時間配分、反復回数で進めることができるという利点がある。一方で、その時間選択の幅の広さや、教師という監督者が傍らにいない点から、学生の学習意欲の維持が難し

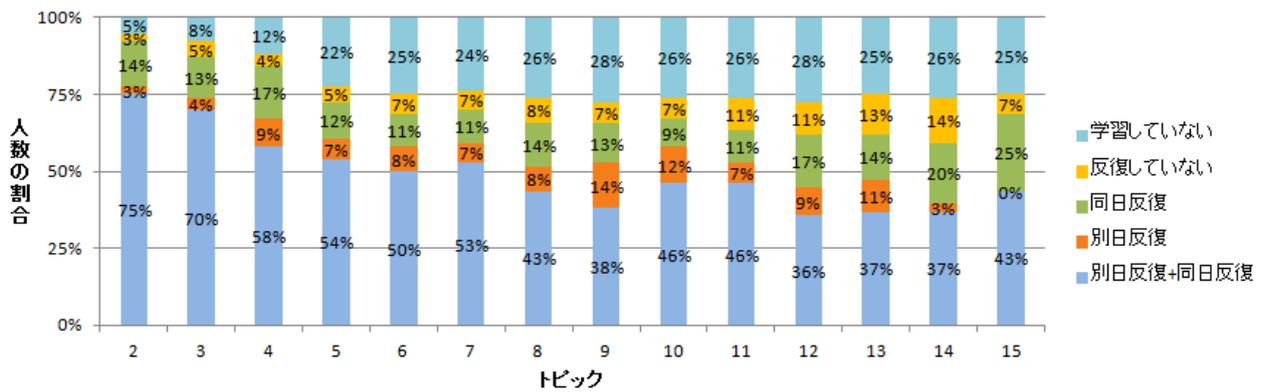


図 2 学習の反復に関する各カテゴリの人数の割合

Fig. 2 Ratio of the number of students in categories of the repetition patterns in learning

い場合があり，従来の授業形態以上に学生の理解に格差が生じることがある．このことから，学習意欲を表す指標を提案し，その指標を用いて学生の分類を試みる．

5.1 学習の反復に着目した分析

従来の授業形態では，授業時間に一度授業に参加するという形で，全員が同じ時間だけ学習する機会を得る．一方で，反転授業においては，学習者が授業までに事前に最低でも一度学習することが必要となる．また，学習の反復は，内容を理解するため，忘れないようにするため，思い出すためなど，様々な理由から重要である．授業時間外の学習が個々に委ねられている以上，学習の有無及び反復は学習意欲と結び付きが強いと考える．よって，学習意欲を測るために学習の反復に着目する．

Moodle のログに，資料の閲覧 (pdf による授業資料もしくはビデオの閲覧) として “resource view”，もしくは，テストの受験が終わったとして “quiz close attempt” が表われた時に学習が行われたと判断する．学習の反復は，同じ日に複数回学習するという場合 (同日反復と呼ぶ) と，別日に再度学習する場合 (別日反復と呼ぶ) の 2 つの場合を考える．同日反復は内容の理解度の向上，別日反復は内容の想起及び定着というように，それぞれの場合における行動目的は異なる．トピック毎に一日に複数回学習している場合を同日反復，別日に再度学習している場合を別日反復とし，トピック毎に学習者を以下の 5 つのカテゴリに分類する．

- 別日反復と同日反復を両方実行
- 別日反復
- 同日反復
- 反復していない (1 回のみ学習)
- 学習していない

ここで，一日は午前 5 時から次の日の午前 5 時までとした．

5.1.1 分析結果

ヒューマン・コンピュータ・インタラクションの授業を対象に，学習の反復による学生の分類を調べた．対象とする学生は，Moodle を利用して学習を行った 76 名とした．図 2 は，トピック毎に学習の反復について 5 つのカテゴリに学生の分類した時の人数の割合である．横軸が各トピックを表し，縦軸が各カテゴリの人数の割合である．後のトピックになるほど，半期の授業において後で学習することになる．

図 2 より，学習していない人の割合がトピック 6 まで増加し，トピック 6 以降はほとんど増減しない．トピック 6 までの早い段階で，学習を行わない学生が増え，学習意欲が低下することが分かる．ただし，ある特定のトピックについてのみ学習していない学生もいるため，トピック 6 を学習していない学生全員が，その後のトピックを学習していないとは言えない．教師は，学生に学習行動を取らせ，学習意欲を維持させるためには，早い段階において何か対応を取る必要がある．

また，後のトピックになるほど別日反復を行う時間が無くなることから，別日反復と同日反復を両方実行している学生と，別日反復を行う学生の合計が減少していることが分かる．

トピックに対応する授業が行われている期間と期末試験の直前の期間に学習行動が行われた場合は，別日反復と見なされる．分析において，授業が行われている期間にリアルタイムに反復を行わない学生を発見することを目的とする場合は，上のような行動を取る学生は，別日反復ではなく，同日反復が反復していないと見なされるため，別日反復はこの目的には適さないと考える．

別日反復は時間の経過による知識の喪失を考えると，実施の早いトピックほど重要度が高くなることを考慮にいたった方が良く考える．

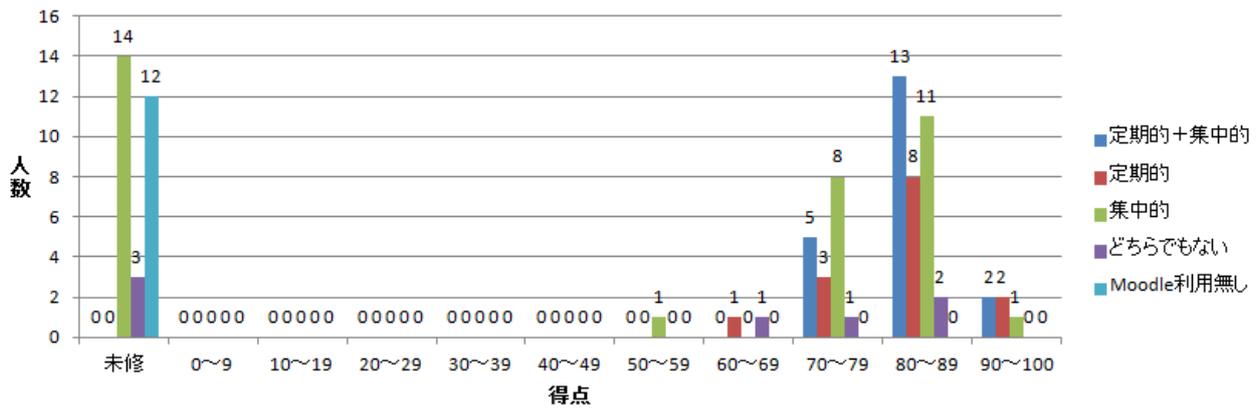


図 3 学習のペースに関するカテゴリ毎の得点別の人数分布
Fig. 3 The number of students in each category of the pace in learning

5.2 学習のペースに着目した分析

学習者が学習意欲を常に持ち続けているとは限らないため、学習の継続性や学習行動の時間による偏りを見ることが、学習意欲の検出に繋がると考える。よって、本節では学習者の学習のペースに着目する。

学習のペースは、定期的な学習と集中的な学習の2つを考える。学習者が各月において、学習行動(資料の閲覧もしくはテストの実施)を行っている場合、定期的に学習を行っていると思えず、Moodleのログに、資料の閲覧(pdfによる授業資料もしくはビデオの閲覧)として“resource view”、もしくは、テストの受験が終わったとして“quiz close attempt”が表われた時に学習行動があったとする。例えば、3ヶ月の学習行動の回数が、(10, 10, 1)の学習者と(10, 0, 10)の学習者がいた場合、前者は定期的な学習を行っており、後者は定期的な学習ではない。

また、学習者が以下の条件を満たす場合に集中的に学習を行っていると思えず。

$$\frac{\max}{\text{sum}} \geq 0.30 \quad (1)$$

ここで、 \max とは、学習者の一日当たりの学習行動の回数の最大を表わし、 sum とは学習行動の回数の合計を表わす。ここで、一日は午前5時から次の日の午前5時までとした。例えば、 $(\max, \text{sum}) = (40, 100)$ の学習者と $(\max, \text{sum}) = (10, 100)$ の学習者がいた場合、前者は集中的な学習を行っており、後者は集中的な学習ではない。

以上より、学習者を以下の4つのカテゴリに分類する。

- 定期的と集中的を両方実行
- 定期的
- 集中的
- どちらでもない

5.2.1 分析結果

ヒューマン・コンピュータ・インタラクションの授業を対象に、学習のペースによる学生の分類と成績の関連を調べた。対象とする学生は、受講登録のあった全学生88名

とした。この授業は10月から2月まで行われているため、学生がこの5ヶ月間の各月において学習行動がある場合に、定期的に学習していると見なした。

図3は学習のペースについて学生の分類を行った際の、4つのカテゴリの総合計別の人数分布である。横軸に総合計の得点区分、縦軸に各得点区分における各カテゴリの人数を表している。

「定期的と集中的」、「定期的」に含まれた学生は、全員が合格(60以上)となっている一方で、「集中的」に含まれた学生は、未修、不合格(60未満)、合格(60点以上)の3個所に分かれた。これは集中的な学習であっても、学習行動の回数の差が結果に影響したのではないかと予想する。また、各得点区分において、各カテゴリが混在する結果となった。よって、本稿の定期的と集中的の定義では、不合格と合格、もしくは合格の中でも秀、優、良、可を分類することはできなかった。

定期的な学習の定義については、3ヶ月の学習行動の回数が、(1, 1, 1)の学生と(20, 20, 20)の学生がいた場合、学習回数に大きな差があるにもかかわらず、両者とも定期的な学習を行っていると思われる。集中的な学習の定義についても、 $(\max, \text{sum}) = (1, 2)$ の学生と $(\max, \text{sum}) = (50, 100)$ の学生がいた場合、こちらも学習回数に大きな差があるにもかかわらず、両者とも集中的な学習を行っていると思われる。よって、学習回数も考慮に入れて、定期的と集中的の定義を行う必要があると考える。

また、定期的な学習の判断を1ヶ月単位で考えているため、月に1回のみ学習であっても定期的と思われる。定期的を判断する期間は短くすべきであったと考える。学習意欲の低い学生を早期に発見することを目的とする場合は、現在の定義では難しいと考える。

ヒューマン・コンピュータ・インタラクションの成績について、期末試験を受けて不合格になった学生は極めて少数だった。このため、不合格を分類するようなパターンを作

ることは難しかった。また、合格者の中で、秀，優，良，可の10点刻みを分類するようなパターンを作ることも難しかった。

定期的な学習と集中的な学習の定義を再考し、成績が秀，優，良，可，不可，未修の学生数がある程度平均的になる授業に対して実験を行うことが今後の課題である。

6. 決定木による分析

2つ目の学習者の分類の試みとして、複数の指標を説明変数、成績を目的変数として決定木による学習者の分類を行う。

本稿で対象としている授業では、Moodle 上での学生の学習行動は、基本的に資料の閲覧とテストの受験である。よって、本節では、テストを受験した回数や資料を閲覧した回数と、資料を閲覧した後にテストを受験した回数のような学習者の学習行動の推移に着目する。これまでと同じように、資料の閲覧 (pdf による授業資料もしくはビデオの閲覧) は “resource view”，テストの受験は “quiz close attempt” が、Moodle のログに表われた時に行動をカウントする。説明変数として以下の6つを用意した。

- quiz：テストを受験した回数
- resource：資料を閲覧した回数
- rtoq：資料を閲覧した後、テストを受験した回数
- qtor：テストを受験した後、資料を閲覧した回数
- qtoq：テストを受験した後、再度テストを受験した回数
- rtor：資料を閲覧した後、再度資料を閲覧した回数

rtoq, qtor, qtoq, rtor は、前後の行動が同じトピックにおいて行われた場合のみカウントする。また、行動は一日単位で区切りを設けているため、午前5時を境として、これを跨いで行動を行った場合は除くものとする。

目的変数に用いる成績は、総合計 (100 点満点) を用い、以下の4つのいずれかの値を取るものとする。

- #80-100 (80 点以上)
- #60-79 (80 点未満 60 点以上)
- #0-59 (60 点未満・不合格)
- #E (未修)

6.1 分析結果

ヒューマン・コンピュータ・インタラクションの授業を対象として、前述の6個の説明変数と、目的変数を用いて、R 言語を用いて決定木の作成を行った。6個の説明変数と決定木を用いることで、成績ごとに学生を分類できるかを調べた。対象とする学生は、受講登録のあった全学生 88 名とした。

付録 A.1 の図 A.1 に作成された決定木を載せる。葉における4つの数字は、左から、#0-59, #60-79, #80-100, #E に分類される学生の人数を表す。葉に付けられた文字

表 1 決定木における各得点区分の学生が含まれる葉の数
Table 1 The number of leaves with the range of score in the generated decision tree

得点区分	葉の数
#80-100	11
#60-79	9
#0-59	1
#E	3

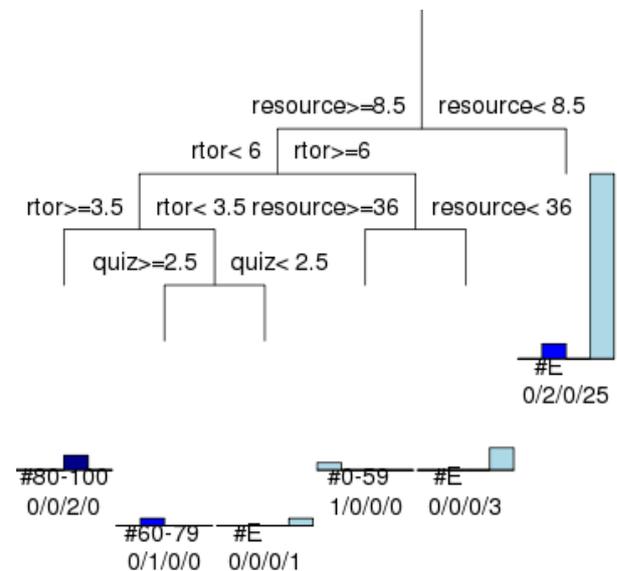


図 4 作成された決定木の一部 (quiz < 7)

Fig. 4 A part of the generated decision tree (quiz < 7)

列は、その葉において最も学生が多い得点区分である。また、分岐における文字列は、その分岐において学生を分類するためのルールを表わしている。

表 1 は、各得点区分の学生が含まれる葉の数を示している。#80-100 に含まれる学生を区別するためには、11 個のルールが必要である。一方で、#E (未修) の学生を区別するためには、3 個のルールのみでよい。

決定木より、テストを 7 回以上受験している学生は全員合格であった。一方で、不合格もしくは未修の学生はテストが 7 回未満だった。図 4 は、最初の分岐で quiz < 7 となった決定木の一部を示している。quiz < 7 の部分木の中に、不合格にあたる #0-59 と未修にあたる #E の学生が全て含まれている。このことから、対象授業において強制ではなかったテストの受験の有無と回数が、可否を分ける大きな要因となっていたことが分かった。

可否の境界である 60 点付近となる #0-59 と #60-79 を分けるルールを探したかったが、不合格の学生は極めて少数であったため、#0-59 を区別するためのルールが一般性のあるものとは言えない。

また、一つの葉を除いて、葉に対応する学生は、ある一つの得点区分に含まれる学生であった。よって、6 つの説明変数を用いて、各得点区分の学生を分類することができ

た．作成された決定木より，この授業では，6つの説明変数の内，テストを受験した回数が重要であったことが分かった．

7. 考察

分析において，対象とする授業の各得点区分における学生数の問題があった．未修を除いて不合格の学生は極めて少数であったため，不合格の学生の特徴を一般的な特徴として捉えることは適切ではなかった．

次に，Moodle のログに記録される情報に問題があった．分析において，資料を閲覧した回数を用いているが，資料の pdf ファイルを Moodle 上で閲覧するのではなく，手元に保存して閲覧した場合，資料を閲覧した回数を正しくカウントできない問題があった．学習者の学習行動を正確に記録できれば，より正確な分析を行うことができると考える．

学習のペースによる分析における定期的な学習や集中的な学習のように，分析に用いた指標は検討する余地がある．また，授業の構成によって分析手法が変わる可能性があるため，授業に応じた分析手法の考案は今後の課題である．

筆者らは，LMS のログを分析することで学生の分類を行うこと，脱落する学生を早期に発見することを目指している．この他に，LMS のログを分析することは，教師側が自分の授業について分析することに役立つと考える．例えば，本稿で対象とした授業では，テストを受講した回数によって，合格者と不合格者が分かれていた．これは，テストを受験した回数が多いほど合格しやすいようになっており，テストの問題設定などが適切に行われていたためでないかと考える．

8. おわりに

本稿では，一つの情報系科目を対象に二つの手法による学生の分類を試みた．学習の反復による分析では，トピック 6 までの早い段階で，学習を行わない学生が増え，学習意欲が低下することが分かった．決定木による分析では，6つの説明変数を用いて，各得点区分の学生を分類することができた．

本稿では一つの情報系科目を対象としたが，指標の再考や分析の対象とする授業を増やすことなどが今後の課題である．また，授業の構成によって分析内容が変わると考えられるため，授業に応じた指標や分析手法の考案も課題である．

謝辞 本研究の一部は，島根大学・地域志向教育研究経費「Ruby・OSS を活用した学習支援システムの開発と e-Learning コンテンツ化の推進」による．

参考文献

- [1] 中央教育審議会，新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学へ～（答申），http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1325047.htm，用語集，pp. 37，平成 24 年 8 月 28 日．（参照 2015-1-6）
- [2] 中野裕司，久保田真一郎，松葉龍一，戸田真志，永井孝幸，右田雅裕，武蔵泰雄，喜多敏博，杉谷賢一，LMS 等を利用したオンライン科目の学習ログの抽出及び分析環境の検討，研究報告教育学習支援情報システム（CLE），2014-CLE-14，No. 7，pp. 1-6，2014．
- [3] 廣川佐千男，殷成久，島田敬士，大久保文哉，緒方広明，教育用デジタルコンテンツの学習ログの分析，信学技報，vol. 114，no. 339，AI2014-21，pp. 31-35，2014．
- [4] 小柏香穂理，浜本義彦，王躍，刈谷文治，小河原加久治，形成的評価を支援する可視化ツールの開発-Moodle の小テストモジュールを活用して-，情報処理学会研究報告，コンピュータと教育研究会報告，2013-CE-122(22)，pp. 1-4，2013．
- [5] 山川修，LMS の学習履歴からわかる学習者の行動特性，教育システム情報学会第 38 回全国大会講演論文集，pp.21-22，2013．
- [6] 土橋喜，Moodle 上の教材と小テストによる学習データの分析と授業改善への効果，電子情報通信学会技術研究報告．ET，教育工学，111(85)，pp. 7-12，2011．
- [7] T. Hecking, S. Ziebarth, and H. U. Hoppe, Analysis of Dynamic Resource Access Patterns in a Blended Learning Course, In Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, pp. 173-182, 2014.
- [8] C. Romero, P. González, S. Ventura, M. J. del Jesus, and F. Herrera, Evolutionary algorithms for subgroup discovery in e-learning: A practical application using Moodle data, Expert Systems with Applications, Vol. 36, pp. 1632-1644, 2009.
- [9] C. Romero, S. Ventura, and E. García, Data mining in course management systems: Moodle case study and tutorial, Computers & Education, Vol. 51, pp. 368-384, 2008.

付 録

A.1 6.1 節における決定木

図 A-1 は，6.1 節の実験において作成された決定木である．葉における 4 つの数字は，左から，#0-59，#60-79，#80-100，#E に分類される学生の人数を表す．葉に付けられた文字列は，その葉において最も学生が多い得点区分である．また，分岐における文字列は，その分岐において学生を分類するためのルールを表わしている．

