

学びの姿勢が学習成果に与える影響の探究

— 自己・授業評価アンケートからの受講態度モデル化の試み —

An Investigation of the Contributions of Students' Study Attitudes to their Learning Outcomes

— Attempting to Modelling Students with their Study Attitudes from Self/Lecture Evaluation Questionnaire —

南 俊朗, 大浦 洋子

Toshiro Minami and Yoko Ohura

【要 約】

学生がより良く学ぶ大学を実現するために従来 FD 活動が推進され、ラーニングコモンズなどの学びの場が整備されてきた。そのような大学・教員側の努力にも関わらず、残念ながら学生の学力が十分向上しているようには見えない。学生側の学びへの意識が十分でないことが、その大きな要因であるかも知れない。この仮説の妥当性を検証するために、我々は授業データの解析を通じて学生の意識を理解し、その改善策を模索するアプローチを取ってきた。本稿では、自己・授業評価アンケートへの学生の回答に用いられた語句を、授業に対する学生の態度との関連性から 4 つのタイプに分け、どのタイプに属する語句をどの程度用いたかを解析することにより学生の学びへの姿勢を評価・モデル化し、その違いが学習成果である試験の成績にどのような影響を与えているかを考察する。このようなアプローチを通じて、より客観的に学生の現状を把握し、より効果的な学びの実現に繋げていくことを目指している。

キーワード： 授業データ解析, 学びへの姿勢のモデル化, テキスト解析, 自己・授業評価アンケート

[Abstract]

In order to let a university become the place where students learn more effectively, university has been carrying out various kinds of activities including preparing the good study environment such as learning commons. However, despite of such efforts, students' learning achievement does not seem to increase suitably. We presume the main cause of this problem may be on the students' side; for example, their poor attitudes to learning or their lack of curiosity, etc. In order to find a solution to this problem, we are taking an approach of modeling students' learning attitudes based on the lecture data. In this paper, we discuss the learning attitudes of students by analyzing the answer texts to self and lecture evaluation questionnaire. We assign one of the four types to each word that appear in the texts, so that the types somehow represent the student's attitude to lectures. By analyzing how their differences on word type ratios affect the learning outcome, or examination grade, of students. By taking such an approach we wish to realize an ideal learning environment for students.

Keywords: Lecture Data Analysis, Modeling Learning Attitude, Text Data Analysis, Evaluation of Self and Lecture Questionnaire

1 はじめに

大学生の学力低下が指摘されて久しい[4]。その背景には多くの要因が存在するものと考えられる。それぞれの要因の影響がどの程度であるかを明確にすることは困難であるが、いずれにしても、高等教育を受けるのに十分な学力のないまま大学生になってしまった学生が珍しくない状況である。

その対策として大学側は、リメディアル教育による学生の基礎学力の底上げを図り、初年次教育を強化し、教員の教育力を増強させるためのFD (Faculty Development) 活動を導入し、と様々な工夫を行ってきた。しかし、そのような対策にもかかわらず、状況の好転には繋がっていない。

このような現象や教員である我々の日常の授業時の印象や、これまで行ってきた研究結果などを踏まえると学生の学力低下の最大の原因は教員の教育力や学生の基礎学力などにあるというよりも、学生の持つ意識や意欲などの学びに対する姿勢や学習態度にあるように見える。

実際、学力の低い学生達を観察すると、彼らは学力以前に学ぶことへの動機づけが十分でなかったり、学ぶとはどういうことかを意識していなかったりと、学びの基礎となるべき知識に対する好奇心、ノートを作成したり復習や予習をしっかりとやるなどの心の姿勢に問題がある場合が多い。

この仮説が正しいとすれば、学生の学力低下の原因を根源的に探究し、その対策を取るためには、学生の学ぶ意欲や新しい知識に対する好奇心や研究心、探索心などの心の在りようをも含めた原因追求が必要であることになる。

様々なデータから役立つ知識を発見する手法はデータマイニング (Data Mining) という呼称で広く研究されてきた[7]。近年は Web システムの発展および普及により、アクセスログなどのデータが大量に得られるようになったことを背景に、ビッグデータ (Big Data) からのデータマイニングが注目を集めている[17]。

教育への応用を意図したデータマイニング、あるいは、データ解析の技術に関する研究は KDD (Knowledge Discovery and Data Mining) [7] や EDM (Educational Data Mining) [8][18][19] などの名称で研究が行われている。本稿の目的である授

業データの解析による有用な知見の獲得を目指す研究もこれらの分野の一部と見なすことができる。

論文[19]は e-Learning システムから得られるデータを用いて学生を分類する問題に対するデータマイニング手法を比較した。その目的は学生の成績を事前に予測することにある ([21] や合田等の研究[1][9][10] も同様)。学生の学習成果と見なせる最終成績を予測することで、問題のありそうな学生を早期に発見し、対策を講じることができる。

本稿の研究目的も概ね同様であるが、我々の主な興味は得られたデータから直接的に成績を予測すること自体ではない。成績云々以前の問題として、学生のやる気や勤勉さなどの心の姿勢や性格、あるいは心理的な傾向などをモデル化し、それに基づき、いわば間接的に成績を予測し、問題のあることが予測できる学生に対して根源的にケアしたり、望ましい学生像を明らかにすることにより、それ以外の学生に、あるべき姿を提示することなどを目指している。

我々は、このような認識を背景に、学生の受講状況を客観的に把握することを目的に、授業データ解析の研究を進めてきた[4][5][11]-[14]。たとえば、自己・授業評価アンケートへの学生の回答テキストの解析より、学修成果と言える試験成績上位の学生達は、狭義の授業内容を越えて、より広い視野から授業テーマを位置付け、また、より広く、授業テーマの周辺知識を含めた学びをしているらしいとの結果が得られている[5]。このような研究結果を受けて、本稿では、授業に対する学生の姿勢が、受講成果にどのような影響を与えているかに関して、より詳細に解析する。

このような目的のために、本稿は以下、次のように構成される。まず、第2節では、自己・授業評価アンケートへの学生からの回答テキストを対象に従来行ってきた解析結果の概要を示す。その結果、学生の成績と学生が回答に用いた語句の性格間には一定の関連性が認められた。その結果を踏まえて、第3節では、本稿のメインテーマとして、語句の性格を4種類にパターン化し、学生の使用語句のパターンと成績の関係を探究する。最後に、第4節では、本稿のまとめ及び今後の研究への展望を示す。

2 自己・授業評価アンケート解析

本節では、本稿で示す研究の背景や研究の位置付けを明確化することを目的に、従来の研究結果を概観する。まず、第 2.1 節において授業の概要および自己・授業評価アンケートの具体的内容を示す。それに引き続き、第 2.2 節では（点数などの）数値データに対する解析結果の概要を紹介する。最後の第 2.3 節においては、アンケートへの回答テキストのデータ解析により得られた内容を示し、第 3 節における更に詳細な解析へと繋げる。

2.1 授業および自己・授業アンケートの概要

本稿で対象としているのは「情報検索演習」という授業に関するデータである。本授業に関連したデータとしては、期末試験の成績点や出席・宿題による平常点、そして、期末に実施した自己・授業評価アンケートへの学生の回答であるテキストデータであり、授業を受講し、評価アンケートに回答した学生は 35 名である。本授業科目の目的は、15 回の授業を通して学生が情報検索に関する十分な知識やスキルを身につけることである。

期末試験は 3 問から構成される。第 1 問は、検索エンジンのサービスを提供している代表的な Web サイトを列挙し、その特徴などを要約して説明する問題である。本問は適切なキーワードを発想する力と、得られた情報を簡潔に要約する力の両方を評価するための問題であり、普段の宿題に真剣に取り組んできた学生には比較的容易な問題であると考えられる。

第 2 問は、電子書籍やオンラインの資料を提供しているサイトを見つける問題である。この問題も適切なキーワードを発想する力をはじめ、サイトの特徴を短時間で読み取ることを要求している。

第 3 問は、インターネットにおける情報犯罪に関して事例を探索し、また、その防止策などに関して議論する問題である。本問題の場合は、一般的な検索能力だけではなく、与えられたテーマに関して事実に基づいて自分の意見をまとめ、それを表現する能力も問われている。

本試験問題は、基本的には 90 分程度の試験時間内で、適切な情報検索を実施し、その結果をまと

める能力を定期試験で評価できるように設定されており、授業中の演習や宿題への課題に真面目に取り組んでいれば、相応の成績が取れるような問題である。

一方自己・授業評価アンケートは授業の終盤に、学生達に当期の授業全体を振り返ってもらう意図で実施されたものである。本アンケートは大学において F D 活動の一環として全学いっせいに実施される授業（評価）アンケートの多くとは異なり、記名式のアンケート調査である。

そのため、無記名のアンケートと比較すると学生達はある程度の責任感を持って、質問に回答しているものと見られる。もちろん、自分に関する良い評価にも悪い評価にもある程度の脚色を加えられることは予想される。しかし、著者等の経験によるとマイナス評価の程度を若干弱めるような脚色はあるもののマイナスの評価をプラスと評価するような虚偽評価はまず見当たらない。

自己・授業評価アンケートは全部で 12 の質問項目からなる。そのうち 5 問は授業に関する質問であり、6 問は学生自身に関する質問、そして残り 1 問は一般的なコメントを求めている。本アンケートの質問項目は次の通りである：

【授業の評価】

(Q1) あなたはこの授業でどういうことを学びましたか？それはあなたの役に立ちましたか？

(Q2) 授業全般について、良かったところはどこでしょうか？

(Q3) 逆に授業についての問題点は何でしょうか？それらはどう改善したらよいでしょうか？

(Q4) 上記の評価を総合して、あなたはこの授業に 100 点満点の何点つけますか？

(Q5) その他、授業に関して気づいた点や感想などを書いてください。

【自分自身の評価】

(Q6) この授業に対するあなたの学習態度について良かったところはどこでしょうか？どのような努力をしたでしょうか？

(Q7) この授業に対するあなたの学習態度の反省点は何でしょうか？今後の自分の学習態度はどのように改めたらよいでしょうか？

(Q8) あなたは授業の狙いの理解に努め熱心に取り

組みましたか？自分の受講態度を振り返って評価してください。

☐ とても良い, ☐ 良い, ☐ 普通, ☐ 余り良くない, ☐ 悪い

(Q9) あなたは授業中、担当講師に質問しましたか？質問に対して講師は適切に答えたでしょうか？
☐ 何度も質問した, ☐ 1度だけ質問した, ☐ 質問したかったができなかった, ☐ 質問なし。

(Q10) 授業後に疑問点を解消するための調査をしたことがありますか？

☐ よく調べた, ☐ たまには調べた, ☐ 疑問点を調べたことはない, ☐ 疑問を感じない。

(Q11) 上記の評価を総合して、あなたはこの授業に対する自分の取り組みについて100点満点の何点つけますか？

【その他の質問】

(Q12) その他、これまでの項目以外のことについての意見や気づいたことなどを書いてください。

2.2 自己評価と努力・成果との関連解析からの知見

図1に自己評価(Q11回答)と試験点(学修成果)の相関関係を示す。興味深いことに相関係数は負値(-0.1)である。すなわち、本例では成果が大きいほど自己評価が低下する傾向があることになる。この事実は何を意味しているのであろうか？

成績が平均点より高い学生のほとんどの成績は70台である。一方満足度に関しては40程度から80程度の範囲内に広く分散している。高成績グループに関して興味深いのは、成績が高い学生の満足度が低いこと、すなわち成績と満足度の間に負の相関があることである。その意味では(図右上に位置する)学生Aは例外的である。

全体的に負の相関になるというこの現象は、これらの学生の自分自身に対する要求レベルが高いことを反映しているためであると解釈できる。これは自己向上心が高いと言い換えることもできる。満足度の範囲が広いことを併せて考えると、自分の成果に満足し、自身の努力を満足感を持って振り返る気持ちと、もっと努力することでもっと高い成果を上げることができたと反省する気持ちの

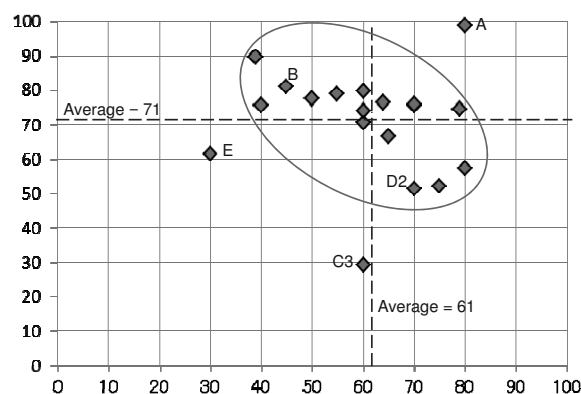


図1: 自己評価(x軸)と試験点(y軸)の相関関係

両方が絡み合った結果、満足度ないしは自己評価に関して、結果として前者を強調した判断をする学生と、逆に後者を強調して判断する学生とに分かれたものと見るができる。

一方、低成績グループに属する学生達の多くは平均以上の満足度評価になっている。グループ内では成績と満足度の関係が弱く、学生個人の判断のばらつきが大きい。学生Eはいくぶん例外的である。本学生の成績は平均値を少し下回る程度であるにもかかわらず、満足度は最低値を付けている。これはEが低い成績のグループに属しているものの、高成績グループの学生と同様の基準で自己判断を行った結果であると考えることができる。

学生C3やD2は低成績グループの典型例である。たとえばD2の場合、成績に関してはC3の次に低いにもかかわらず満足度に関しては平均よりかなり高値である。このような学生は、自分自身を客観化する能力が低いか、もしくは事実に基づいて客観的に評価しようという意識が弱いことが、その理由であると推察できる。もしかしたら、自分の状況を十分認識した上で、あえて現実以上に良い評価を与えたいという気持ちが働き、実際の自己認識以上の評価をしているのかも知れない。そうすることで、自身への自信を取り戻したいという補正の気持ちが反映されている可能性がある。

以上の考察は数値化された評価のみに基づいている。学生の評価情報に関して更に詳細な解析を行うためには、テキストによる回答部分を数量化するなど更なる追究が必要である。テキストを対象に解析した例を次節に示す。

表 1: アンケート回答中の頻出語句 (トップ 30)

語句	頻度	語句	頻度	語句	頻度
検索	88	仕方	16	考える	8
授業	37	調べる	16	行う	8
情報	37	キーワード	13	得る	8
思う	34	様々	11	さまざま	7
図書館	33	使う	10	感じる	7
学ぶ	32	役に立つ	10	機能	7
知る	30	必要	9	結果	7
自分	21	利用	9	大切	7
方法	21	インターネット	8	機会	6
今	17	パソコン	8	今回	6

2.3 総合評価テキスト解析からの知見

学生が授業で学んだ内容を適切な言葉で記述できるか否かは、授業内容の理解度を反映しており、試験結果にも影響があると考えられる。その可否を確認するために、自己・授業評価アンケートの自由記述項目と試験点の関連性を調べた。

第 2.1 節で示した本アンケートの自由記述項目中、設問「(Q1) あなたはこの授業でどういうことを学びましたか？それはあなたの役に立ちましたか？」への回答結果を解析対象とする。本設問は授業で学んだ内容を総括し学生自らの言葉で表現することを求めている設問であり、学習成果が良く表れていることが期待できる質問項目である。

アンケートの自由記述項目への回答のような、文章で表現されたテキストデータを解析対象とするためには、何らかの方法により質的データである原テキストを数量化 [2] し、解析可能なデータを生成する必要がある。高い質のデータを得るために、人手で内容进行评估し、その評価データを分析に用いることは良く行われている。

我々は客観的テキストデータ解析の第一歩として、フリーソフトの解析ツールである KH Coder[3] を用いて質問 (Q1) への回答テキストの解析を行った。表 1 に、得られた語句とその頻度の一部を示す。予想通り、「検索」、「授業」、「情報」など授業との関連性の高い語句や、「思う」、「知る」、などの一般的に良く用いられるであろう語句の使用頻度が高い。

質問 (Q1) への回答に現れた学生と使用語句の対応分析の結果を図 2 に示す。用いられた語句との相対的距離などを考慮し、目視によりグループ分けを行った結果、5つのグループに分かれた。そ

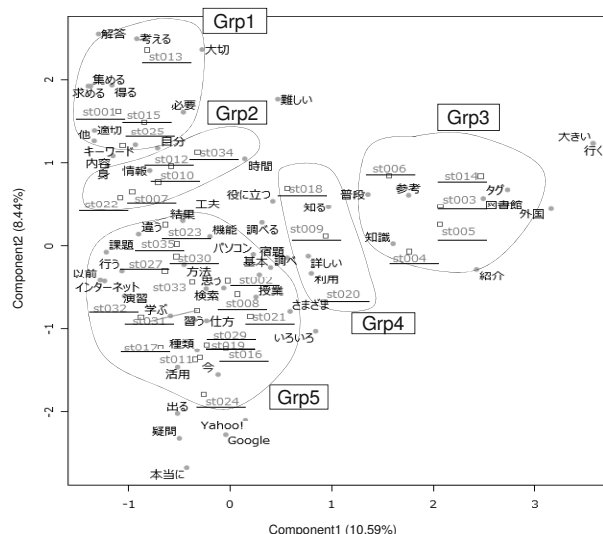


図 2: 学生と語句の対応分析図 (下線付きは学生)

表 2: グループの成績比較

グループ	人数	平均	分散
Grp1	4	65.2	27.3
Grp2	5	70.5	98.8
Grp3	5	83.5	107.2
Grp4	3	69.8	68.7
Grp5	16	59.3	335.3

れらに対して図のように Grp1 から Grp5 の名称をつけた。

得られた 5つのグループに対して、それに属する学生の人数や成績の平均と分散を表 2 に示す。グループの平均点は Grp3 の 83.5 点を最高に、Grp5 の 59.3 点まで約 24 点の差が存在する。この結果より、期末試験の成績と、学生が質問 (Q1) への回答で用いた語句の間には何らかの関連があることが認められる。

表 3 に成績最高グループ Grp3 の学生が用いた語句と最低グループ Grp5 の学生が用いた語句の例を示す。Grp3 の学生は「外国」「海外」「日本」「国内」などの授業に直結してはいない語句を多く用いている。これらの語句は、授業中に学生への参考として、国内外の図書館を紹介したことに言及した結果用いられたものである。これらの学生は、専門用語や特徴語も良く用いている。

表 3: Grp3 と Grp5 の学生が用いた特徴的語句

Grp3

st005 99	st006 76	st003 77	st004 76	st014 90
外国 ● 図書館 ▲ 最新 取り組み 世界 見る IC ■ タグ ■ さまざま 感じる	図書館 ▲ 個性 概要 就く 関係 プラス 全国 同時に レファレンス ■ 又	組む 国 ボックス HP 閉館 ■ 蔵書 ■ 載る ルート 取り組む 県	見せる 限る 興味深い 写真 導入 身近 日本 著作 ● 毎回 学べる	行く 海外 自動 特に 全く 貸出 ■ 電子 ● 大きい 普段 図書館 ▲

Grp5 (5 students out of 16)

st019 52	st030 27	st031 34	st032 29	st035 51
昔 疑問 本当に 今 思う 学ぶ 応える 答え 事前 現在	結果 行う 知る 情報 ▲ 学ぶ 検索 ▲	単語 以来 アプローチ 意欲 刺激 全然 異なる 頻? デスク たくさん	インターネット ● 活用 自分 先 使用 就職 今後 受ける 演習 以前	引き出す 自宅 作業 多々 慣れる 聞く 触れる 生かす 漠然と 奨学生

一方、Grp5 の学生達が良く用いた語句は「覚える」「学ぶ」「役に立つ」など学習に関連した表現が目につく。専門用語の使用は少ない。興味深いのは、成績最下位グループの学生達が用いている語句は授業に関連したものが多く、一見すると真面目に授業に取り組んでいる印象を受ける。

しかし、これらの学生の成績が振るわないことや成績上位者の授業への態度を勘案すると、真面目に授業を受けているように見えるこれらの学生達には、授業を通して真の学力を身につける力に問題があることが疑われる。すなわち、これらの学生は、表面的には授業に出席し、真剣に課題に取り組んでいるように見え、恐らくは学生自身も自分は真面目に学習していると信じているであろう。それにも関わらず、実は身につくような学びが実現できていないものと考えられる。

このようなことが起こるのは、これらの学生が、形としての授業や形としての学習にとらわれ過ぎていることが一因であると考えられる。すなわち、これらの学生は、授業の意義を狭い意味に限定して理解し、いかにも授業テーマに直結した内容にのみ注意を集中し、それ以外の授業内容の補足的な内容や、授業内容に関連した周辺分野に関して授業担当者が話している内容を、学ぶべき事柄で

はなく、いわゆる雑談として捉えており、一見授業とは直結しない内容からも何らかを学ぼうという意識に欠けているのであろう。

このような視野の狭さゆえに、学んだことを単に知識として記憶しようという傾向が強く、応用力が十分養成されていない恐れがある。もし、この解釈が正しいとすれば、授業担当教員は、授業の効果を高めるためにも、広い視野で学ぶことの重要性についても学生に働きかける必要がある。

3 評価テキスト出現用語のタイプと成績との関連分析

前節で述べたように、これまでの解析により、成績上位グループの学生は、広い視野を持って授業を受けており、また、授業テーマに直結した内容以外にも注意を向けつつ授業に臨んでいる状況がうかがえた。その反対に、成績下位グループの学生は、授業テーマに直結した内容に注意を集中するあまり、一見すると授業テーマに直結しているようには見えないものの、実際は授業内容を十分に身につけるために重要なことがらを軽く見ている傾向が見受けられるとの結果が得られた。

更に、授業に対するこのような態度の違いは、自己・授業評価アンケートへの回答に現れる語句のタイプの違いとして表れていることが観察された。これらの結果を受けて、本節では、学生が用いた語句のタイプの違いが実際の程度学生の試験成績に影響を与えているかを、更に検証する。

まず、第 3.1 節で、本節での解析の流れを概説する。次に第 3.2 節で、学生が用いた語句に 4 つのタイプのいずれかを割り当て、学生と語句の間の“用いる－用いられる”関係を利用して、学生の成績に対するタイプ毎の“成績貢献度”を計算する。

更に第 3.3 節では、語句の成績貢献度から“用いる－用いられる”関係を逆向きに利用して各学生の成績を見積もる。すなわち、成績貢献度の高いタイプの語句を多く用いた学生の成績は高く見積もられることになる。

残念ながら、平均値を合わせるように調整された成績推定の最初の試みの結果は精度良くタイプ

の成績貢献度を説明していないという結果が得られた。それを改善するために、成績の最大値と最小値が一致するような補正を施した結果、平均値を合わせる補正よりも良い結果が得られた。

3.1 解析の流れ

従来の観察結果を受け、学生がアンケート項目(Q1)への回答に用いた語句のタイプとして次の4種類を設定する：

- A: 授業テーマに直結した語句。
たとえば、「情報検索」「検索法」「(検索)キーワード」「パソコン」等。
- B: 授業テーマには直結していないものの、授業で取り上げた話題に関連した語句。
たとえば、「IC タグ」「(図書館) 紹介」「自動貸出機/返却機」等。
- C: 学生の回答に頻繁に現れる一般語。
たとえば、「知る」「行く」「方法」「思う」等。
- D: 学生の回答に頻繁には現れていない一般語。
たとえば、「使う」「内部」「取り組む」等。

これらの4種類のタイプに基づき、本稿における成績推定の処理は次のような流れで処理される：

- (i) 学生使用の語句に上記4種類のタイプのいずれかを割り当てる。ここに、授業担当者が授業テーマに直接係る説明などで用いたか否かによりタイプAとタイプBに分別した。これら以外の一般語は出現頻度に応じてタイプCとタイプDに分別した。
- (ii) 学生が用いた語句のタイプ別割合に応じて、その学生の成績をタイプに割り当てる。
- (iii) 各学生から割り当てられたタイプ別値の合計をそのタイプの“成績貢献度”として求める。以上が、本節(本稿)における処理の前半部分である。
- (iv) 各学生の使用語句のタイプ別割合に応じて、前項の結果得られたタイプの成績貢献度の重みで学生の成績推定値を見積もる。推定値を正規化することにより、学生が用いた語句のタイプから推定された当該学生の推定成績が得られる。

表 4: 使用語句へのタイプ割り当て

語句	タイプ	語句	タイプ	語句	タイプ
パソコン	A	利用	D	組む	D
検索術	A	仕方	C	思う	C
国や県	D	載る	D	蔵書	A
図書館	A	内部	D	カウンター	A
知る	C	取り組む	D	返却ボックス	A
検索方法	A	IC タグ	B	比べる	A
HP	A	訪れる	D	自動貸出機	B
行く	C	機会	B	検索する	A
開館日	A	対象	D	大切	C
閉館時間	A	ルート	D	検索項目	A

- (v) 推定された成績がどの程度妥当であるかを、元の成績と推定成績を比較することにより評価する。この精度が、語句のタイプがどの程度の影響を成績に与えているかを表すことになる。

以下、第3.2節で、本流れの(i)から(iii)を、実際のデータに基づき実施し、(iv)と(v)を第3.3節で実施する。

3.2 用語のタイプと成績貢献度の重みづけ

本節は次の流れで議論を進める。まず、(i) 学生が用いた語句に4つのタイプのいずれかを割り当てる。次に、(ii) 各学生において、その成績を、使用割合に応じて語句タイプに配分する。そして、(iii) 各学生から配分された値の合計を、そのタイプの“成績貢献度(貢献度合)”とする[15][16]。更に、(iv) 学生のタイプ使用割合を重みとしてタイプの成績貢献度より各学生の成績を推定する。(v) 学生の元の成績と推定された成績を比較する。

表4に語句へのタイプ割り当て例を示す。なお、出現回数3回以上の一般語を頻繁語であると定めた。この基準によると一般語の約20%が頻繁語となる。

次に、学生が用いた語句の頻度と、その語句のタイプ割り当てを基に、それぞれの学生における語句タイプの使用に対する割合を求める。

図3に学生毎および全体の使用語句タイプの割合を示す。学生はst003からst030へと成績の降順に配置してある。タイプ別に傾向を概観すると、まずAに関しては、成績上位の学生の使用割合が多く、成績中位に向かって使用割合が減少する傾向

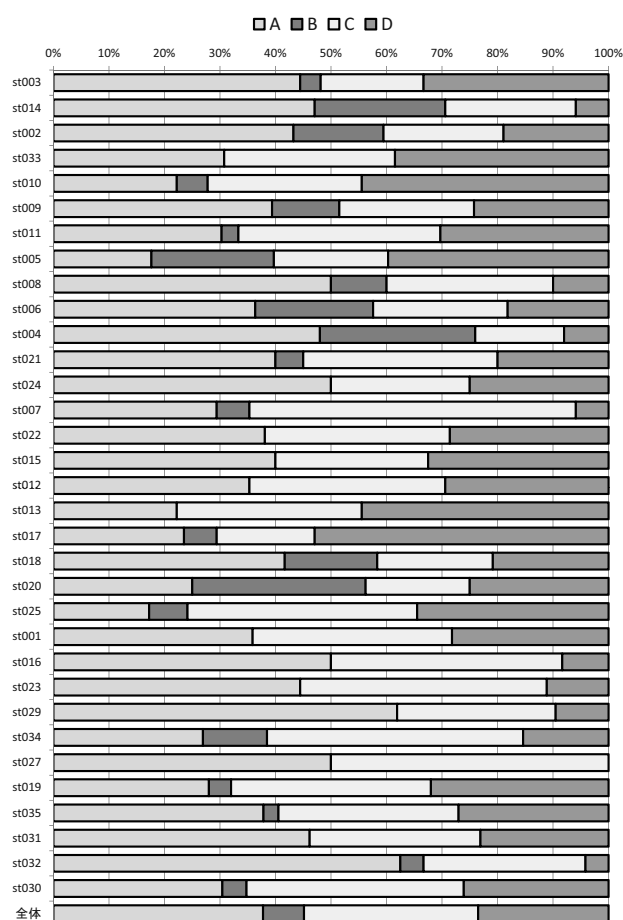


図 3: タイプ A からタイプ D に関する学生別および全体の利用割合 (100%積み上げ表示)

向が見られる。ところが、その後、成績中位から成績下位に向かって再び増加の傾向が見られる。このような観察結果からは、タイプ A (授業に直結した語句) を多く用いるほど成績が向上するとも、低下するともには判断できない。

タイプ B (授業関連語句) に関しては、全般的な傾向として、成績上位の学生が良く用いていることが図から読みとれる。しかし成績中の下グループの学生の中にタイプ B の語句を多く用いている学生が存在するため、成績との相関が非常に強いとは言えない。

タイプ C (高頻度の一般語) に関しては、成績中位グループの使用割合が高い傾向を読み取ることができる。子細に検討すると、成績下位グループと成績中位の上位グループの使用が多い傾向が読み取れる。

タイプ D (低頻度の一般語) に関しては、使用割合の高い山が、概ね 3 つほど認められる。すなわち、成績下位の学生、成績中位 (中の中)、そして成績中の上もしくは上位の下位の学生である。平均化すると中位の成績となり、成績との相関は少ないように見えるものの、子細に検討すると、成績上位から中の上の山の方が、成績下位の山よりも割合の値が大きく、また、範囲が広いので、全体的には成績上位者の使用傾向が高いと判断される。

このような観察結果を確認するために、使用語句のタイプと学生の成績の関連を、より客観的な手段で調べる。そのために、次のように“タイプの成績貢献度”なる概念を定め、また、それを用いて各学生の成績を推定する。

- 学生それぞれが使用した語句のタイプ別割合を求める。
- 学生の試験成績をタイプ別割合に応じて配分する。具体的には、ある学生の試験成績とその学生が用いた語句のタイプの割合をかけた値を、その学生に関するそのタイプの成績貢献度とする。これは、学生が用いた語句が、そのタイプの成績貢献度に応じて当該学生に与えられ、その合計が学生の成績となるというモデルに基づき、学生の成績より逆向きにタイプの成績貢献度を計算したものである。
- 各タイプに対して、各学生に関する、そのタイプへの成績貢献度の平均値を求める。そして、4 つのタイプの成績貢献度の総和に対する各タイプの割合を、そのタイプの (相対的) 成績貢献度とする。
- 前項で求めた、全体に対するタイプの成績貢献度を用いて、各学生の使用語句のタイプの割合にタイプの成績貢献度をかけた値の合計を、当該学生の推定成績元データとする。
- 推定成績元データの値の平均値が、元の成績の平均値と一致するように係数を設定し、その値をかけた結果を当該学生の推定成績とする。

図 4 に 4 つのタイプの頻度と貢献度の割合の比較を示す。これは上記、ステップ (c) の結果得られたタイプに対する成績貢献度である。比較のための頻度の割合と比較すると、タイプ A とタイプ C

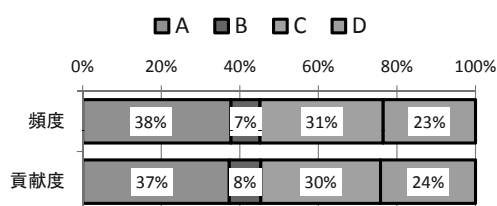


図 4: タイプ出現頻度割合と成績貢献度割合の比較

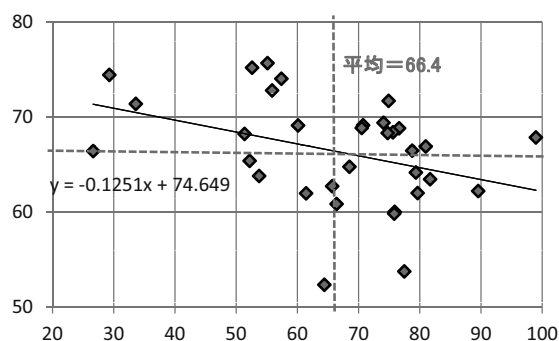


図 5: 元の成績 (x 軸) と使用語句タイプの貢献度より推定された成績 (y 軸) の相関

の割合が、それぞれ 1 ポイント減少しており、その代わりとして、タイプ B とタイプ C の割合が、それぞれ 1 ポイント増加している。これらの減少・増加は、成績上位の学生が、授業関連語句や頻度の高くない一般語を良く用いている傾向があるとの観察結果と一致している。しかし、その違いは 1 ポイント程度とさほど大きくはない。

学生の元の成績と上記ステップ (e) の結果得られた推定成績の相関を図 5 に示す。図によると、双方には負の相関があり、成績の推定値としては、的外れなものになっている。これは図 3 に対する観察結果への客観的裏付けとなっている。

すなわち、成績上位の学生と成績下位の学生の使用する語句タイプの傾向が比較的似ており、語句のタイプからの成績推定値に大きな差が生じていない。むしろ、成績上位の学生達の成績推定値は平均値である 66.4 か、それ以下として見積もられている。

それに対して、推定成績で上位を獲得している学生は、元の成績では、中の下のクラスに属している。これらの学生は、タイプ B やタイプ D の語句を良く用いており、その結果推定成績が高めに出自している。

一方、成績上位の下、もしくは、成績中位の上の学生達は、逆にタイプ A やタイプ C の語句を用いる傾向があり、その結果、実際の成績に対して、推定成績は若干低めになっている。

3.3 成績推定法の改善

前節で用いた成績推定方法では、補正のための係数をかけて推定成績値と元の成績の平均値が一致するようにしたもの、相関係数が負となってしまう、十分な精度での成績推定ができなかった。本節では、補正の仕方を改善することにより、より精度の高い成績推定方法を提案する。

本節のアイディアは、成績の最大値と最小値が一致するように補正することにより、平均一致の手法が持っていた欠点である成績のばらつきの大きさを調整し、成績全体を妥当な範囲に収めようというものである。

本補正手法は与えられた成績推定値 x より補正された成績推定値 y を次の式により求める。

$$y = OMin + (OMax - OMin) \times \frac{x - SMin}{SMax - SMin}$$

ここに、 $OMin$ および $OMax$ はそれぞれ元の成績の最小値および最大値を、また、 $SMin$ および $SMax$ はそれぞれ推定された成績の最小値および最大値を示す。

本式は、推定成績の最大値 $SMax$ を与えると、元の成績の最大値 $OMax$ を、そして、推定成績の最小値 $SMin$ を与えると、元の成績の最小値 $OMin$ を返す。

図 6 に元の成績と最大値と最小値の両者が一致するように補正した推定成績の相関を示す。これまで観察された現象が本図で明確化されている。図に示されているように、学生全体は元の成績により概ね 4 つのグループに分割できる。成績上位グループは学生 st003 と st014 からなる。これらの学生は元の成績は 90 点以上であるものの推定成績は 70 点台とかなり低下している。

成績順で第 2 のグループは元の成績で平均点である 60 点台上位から 80 点以下の学生達から構成される。例外的に st007 の学生は元の成績は 70 点台あるにも関わらず、推定成績ではかなり低い成

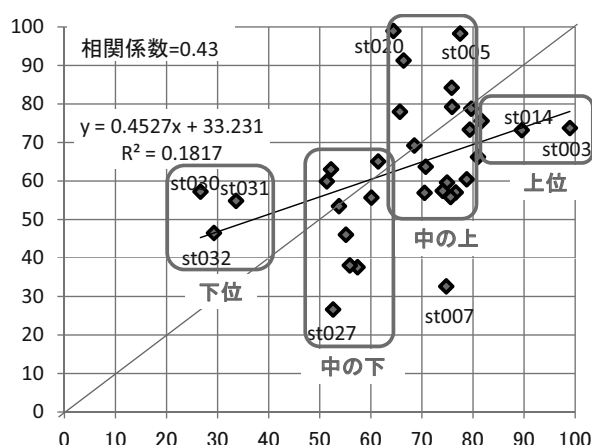


図 6: 元の成績 (x 軸) と推定成績 (y 軸) の相関 (最大・最小値が一致となる補正適用)

績となっている。しかし、他の学生達は概ね上位 (平均点の少し下から最高点近く) の推定値となっている。学生全体での推定成績が最高値に極めて近い学生 st020 や st005 は、このグループに属している。すなわち、広い観点から授業を受け、周辺知識への興味が高く、また、多くの学生が用いているような一般的な表現ではなく、独自性の高い授業評価コメントをしている学生達のグループである。

第3のグループは、元の成績で中の下を占めるグループである。このグループの学生達の推定成績は、平均点の少し上から最低点 (学生 st027) に至り、推定成績はかなりばらつき、その幅は第2グループと同様にかかなり大きい。元の成績は、さほど低くないものの、推定成績が低いことから判断して、授業に直結した語句を多く用い、また、多くの学生が用いているありふれた表現を用いてコメントした学生グループであると考えられる。このグループの学生達には、もっと自分独自の視点で物事を見、考えることを促す必要があろう。

第4のグループは st020, st031, そして st032 の3名で構成されている。これらの学生達は、成績は最下位に位置するものの、コメントへの評価としては、平均か、その少し下に位置している。更に詳しく調べる必要はあるものの、これらの学生は、潜在的实力は備えていながら、授業を受ける態度などに問題があるなどの理由により、成績が振るわなかった可能性があるものと見られる。

4 まとめと将来への課題

我々は、近年指摘される学生の学力低下の大きな要因として、学生の授業態度に注目している。学生の授業態度の状況を客観的に把握するために、本稿では、期末に実施した自己・授業評価アンケートへの学生の回答と期末試験の成績の関係を解析した。

具体的には、本稿ではアンケートの中の、質問 (Q1) 「あなたはこの授業でどうのことを学びましたか？それはあなたの役に立ちましたか？」への回答で用いられた語句を4つのタイプに分類し、タイプ毎の利用頻度に焦点を当てて解析した。

4つのタイプは、A: 授業に直結した語句、B: 授業に関連した語句、C: 学生の回答に多く現れる一般語、そして、D: タイプC以外の一般語、である。これらのタイプは、質問 (Q1) への回答に対する従来の解析の結果、成績との関連が高いものと推測された語句のタイプ分けに準じて設定されたものである。

本稿の解析の結果、全般的には、従来の予測通り、タイプBとタイプDの語句に関しては、これらを多く用いる学生ほど成績が高く、逆にタイプAとタイプCの語句に関しては、それらを多く用いる学生ほど成績が低い傾向があることが確かめられた。

しかし、これらの傾向は概ね妥当であるという程度のものであった。より詳細な解析の結果、事情はもっと複雑であることが分かった。すなわち、タイプBやタイプDを多用しているのは、成績トップの学生達ではなく、それに次ぐ成績上の中から中の下の学生達であることが見出された。成績トップの学生達は、このような傾向はさほど強くなく、語句の使用に関しては平均的であった。

また、タイプAやタイプCを多用しているのは、成績最下位の学生達ではなく、その少し上、すなわち、成績が中の下に属する学生達であることが見出された。成績最下位の学生達は、成績トップの学生達と同様に、中の下グループの学生達とは異なり、語句の使用に関してはむしろ平均的であった。

以上の結果から、語句の使用状況だけで成績を判断することは、当然ながら、かなりの誤差を伴うことが判明した一方、学生の成績層により特徴

的な語句の用い方が存在することも同時に明らかになった。

今後は、次のような課題の解決策を追究するために、本稿で紹介したようなアプローチによる研究を更に継続していく必要がある：

- (1) 本稿で用いた成績推定法は、まだまだナイーブなものである。語句タイプの使用の差をもっと強調した重みづけで成績推定することにより、精度をもっと上げることができる可能性がある。
- (2) 使用語句のタイプとして、本稿とは異なるタイプを設定することにより、本稿で示されたタイプ分けよりも、より良く成績を説明できる可能性がある。そのようなタイプ分けを今後も追究していくことが重要である。
- (3) 現在解析に用いているデータ以外のデータによる同様な解析を行い、比較することにより、授業のテーマや性格の違いによる差異や、同じ授業であってもクラスによる差異などを明らかにすることも重要な研究テーマである。
- (4) 本稿ではもっぱら授業データ解析のみをテーマとして取り扱ってきた。大学に絞ったとしても、データ解析が有効な分野は、図書館データや教務データなど数多くの対象候補が存在する。これら全体を総括した大学 IR (Institutional Research, 機関研究) の一環として授業データ解析を捉えることも重要である [6]。

本稿で示された授業データ解析手法に関する研究は始まったばかりであり、当該分野は未だ初歩的段階にある。それにも係らず、いくつかの示唆に富んだ結果が既に得られている。その潜在的な可能性の高さを考えると、今後ますます発展させるべき研究分野であると言える。

現在の多様化する時代において、留学生も含め多様化する学生に対する適切な教育環境を整備していくためには、本研究は、授業データのみならず、教務データなどを含んだ、解析対象をより拡大する方向での進展が重要である。これからは、大学 IR、特に教学 IR を意識しつつ、授業データ解析技術の研究開発が展開していくことに期待したい。

謝辞

本研究の一部は文科省の科研費、基盤研究 (C)、24500318, 2014 として実施されました。

参考文献

- [1] 合田 和正, 峯 恒憲: PCN 法による自己評価コメントの分析からの改善可能な学生の発見, 教育システム情報学会 (JSiSE) 特集論文研究会研究報告, Vol.26, No.7, pp.51-58. (2012)
- [2] 樋口耕一: テキスト型データの計量的分析: 2つのアプローチの峻別と統合, 理論と方法, 19(1), pp.101-115. (2004)
- [3] 樋口耕一: KH Coder 2.x リファレンス・マニュアル. (2012) <http://khc.sourceforge.net/>
- [4] 南 俊朗, 大浦 洋子: 学生の成長を助ける学習支援への模索―授業データ解析による支援方法発見への試み―, 九州情報大学研究論集, 第 14 巻, pp.39-50. (2012)
- [5] 南 俊朗, 大浦 洋子: 授業データ解析による授業改善策発見を目指して―努力・成果・評価の関連性からのアプローチ―, 九州情報大学研究論集, 第 15 巻, pp. 1-16. (2013)
- [6] 南 俊朗: 図書館マーケティングから大学 IR (機関研究) へ―図書館データ解析への期待―, 九州大学附属図書館研究開発室年報 2013/2014, pp. 8-17. (2014)
- [7] ACM SIGKDD: ACM Special Interest Group on Knowledge Discovery and Data Mining. <http://www.sigkdd.org/>
- [8] Educational Data Mining Society: <http://www.educationaldatamining.org/>
- [9] K. Goda and T. Mine: Analysis of Students' Learning Activities through Quantifying Time-Series Comments, Proc. 15th Annual KES Conference (KES'2011), Part II, Lecture Note in Artificial Intelligence (LNAI 6882), pp.154-164. (2011)

- [10] K. Goda, S. Hirokawa, and T. Mine, Automated Evaluation of Student Comments on Their Learning Behavior, 12th International Conference on Advances in Web-Based Learning (ICWL 2013), pp.131–140. (2013)
- [11] T. Minami and Y. Ohura: Toward Learning Support for Decision Making—Utilization of Library and Lecture Data—, The 4th KES International Conference on Intelligent Decision Technologies (KES-IDT'2012), Springer Smart Innovation, Systems and Technologies 16, pp.137–147. (2012)
- [12] T. Minami and K. Baba: Investigation of Interest Range and Earnestness of Library Patrons from Circulation Records, Proc. International Conference on e-Services and Knowledge Management in IIAI-AAI 2012, IEEE CPS, pp.25–29. (2012)
- [13] T. Minami and Y. Ohura: An Attempt on Effort-Achievement Analysis of Lecture Data for Effective Teaching, Database Theory and Application (DTA 2012), T.-h. Kim et al. (Eds.): EL/DTA/UNESST 2012, CCIS 352, Springer, pp.50–57. (2012)
- [14] T. Minami and Y. Ohura: Towards Development of Lecture Data Analysis Method and its Application to Improvement of Teaching, 2nd International Conference on Applied and Theoretical Information Systems Research (2nd ATISR 2012), 14pp. (2012)
- [15] T. Minami and Y. Ohura, A Correlation Analysis of Student's Attitude and Outcome of Lectures—Investigation of Keywords in Class-Evaluation Questionnaire—, The 7th International Conference on Database Theory and Application (DTA 2014) in Advanced Science and Technology Letters (ASTL), Vol. 73 (FGCN 2014), pp.11–16. (2014)
- [16] T. Minami and Y. Ohura: Towards Improving Students' Attitudes to Lectures and Getting Higher Grades—With Analyzing the Usage of Keywords in Class-Evaluation Questionnaires—, The Seventh International Conference on Information, Process, and Knowledge Management (eKNOW 2015). (2015) (to appear)
- [17] S.-H. Myaeng: Towards a memory: human experience mining and semantic social networks, 3rd IIAI International Conference on e-Services and Knowledge Management (IIAI ESKM 2012), Keynote Speech (2012)
- [18] C. Romero and S. Ventura: Educational data mining: A survey from 1995 to 2005, Expert Systems with Applications 33, pp.135–146. (2007)
- [19] C. Romero, S. Ventura, P.G. Espejo, and C. Hervás: Data mining algorithms to classify students, Proc. 1st International Conference on Educational Data Mining (EDM 2008), pp.8–17. (2008)
- [20] L. Talavera and E. Gaudioso: Mining student data to characterize similar behavior groups in unstructured collaboration spaces, Proc. Workshop on Artificial Intelligence in CSCW, 16th European Conference on Artificial Intelligence (ECAI 2004), pp.17–23. (2004)
- [21] N. Thai-Nghe, L. Drumond, T. Horvath, and L. Schmidt-Thieme: Multi-Relational Factorization Models for Predicting Student Performance, in KDD 2011 Workshop: Knowledge Discovery in Educational Data, as part of 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 10pp. (2011)