

テキストマイニング技術の活用へ向けた実験支援システム

丸山 広[†] 中村 正太[†] 高橋 知七[†] 中村 太一[†]
東京工科大学[†]

1. まえがき

情報処理技術の発展と普及に伴い、情報処理の技術分野が基礎から応用へと広がっていることは周知の事実である。このような動向に対応するため、情報工学分野のカリキュラムでは、コンピュータを如何に利用し、付加価値を生み出すか、といった利用技術の習得の重要性が認識されるようになってきている。マーケット分析のような決まった答えを求めない実験では、事象が確率の要素を含んでおり、得られた解釈の多様性のため結果の多様性を許容する必要がある。そこで実験講師が学生一人ひとりへフィードバックする個別指導が必須となるが、多人数の学生を擁する大学においては難しい。

この問題に対処するため、本研究では、レポートの形式チェックに加え、テキストマイニングの技術を使いレポートの内容に対して改善点を指摘することで実験講師の負担を削減、また学生の実験進捗をモニタし、学生個人の理解度に合わせた指導を実現することを目的とした実験支援システムを提案する。

2. 学生実験のカリキュラム

情報系学部2年生を対象とする実験で、膨大な文書を分析して内容を瞬時に把握したい、というニーズに応える技術、テキストマイニングの基礎手法を3人1組で学び、その分析結果に対する学生の見解をレポートとして求めるカリキュラムである。実験中にはTA(Teaching Assistant)が担当学生の進捗を把握し、かつ途中成果物を見て指導できる様に、学生は設定された課題を実験支援システムを通して順次提出する。

3. 提案システムの基本的要件

本研究は、指導する実験講師とTAの立場からは、人でなければできない個別指導の時間を確保し、指導を受ける学生には、自主的な実験参

画を促し、学習効果を高められる実験環境を構築することを目的とする。

個別指導充実の第1歩として遅れている学生へのフォローに着目し、各学生の実験の進捗を把握する支援を行う。レポートの添削においては形式の指導に多くの時間が割かれているという現状があり、内容の指導を充実させるためにはレポート品質向上支援が求められる。

提案システムの基本要件を次に示す。

(1) 実験の進捗把握

課題を提出することで、学生にとって、TA側が進捗を把握するためだけでなく、自分の進捗が分かりデータのバックアップにもなり、双方の利益になるような形で進捗を把握できる。

TAは担当グループを持ち、自分の担当学生の進捗を把握でき、また担当範囲と全体の進捗を俯瞰することで指導ペースの調節ができる。

(2) レポート品質向上支援

添削と修正のサイクルを繰り返すことがレポート作成能力の向上につながると考えられるが、人手で何度も添削指導を行うことは難しい。そこでシステムによるレポートの添削を行う。また基本的な章立てを定義に含めることで、システムは特定の章だけのテキスト情報を抽出することができ、章に合わせたテキストマイニング技術を適用することが可能となる。更に、添削を行ったレポートデータを残しておくことで、添削前と添削後のレポートが収集でき、どのように修正してくるかを分析し、個性を把握するといった有益な情報を得ることができる。

(3) レポートの電子管理

実験レポートを電子ファイルで提出するので、印刷の手間が削減でき、レポート提出の管理が容易になる。また学生は自分の実験レポートの受理状況を知ることができ、不注意による未提出を回避できる。

(4) 個別指導方針の提案

課題の提出やレポートの添削、講義資料のダウンロードといった実験に関わる学生の行動ログを分析する。更にレポートを分析することで学生一人ひとりに合わせた指導方針を提案する。

4. 機能要件

3で述べた基本的要件を満たすWebベースの

The IT assistance system that collaborate with Text mining for the program of experiment.

[†]Hiroshi MARUYAMA [†]Syouta NAKAMURA [†]Tomona TAKAHASHI [†]Taichi NAKAMURA; Tokyo University of Technology

実験支援システムの機能を図1に示す。

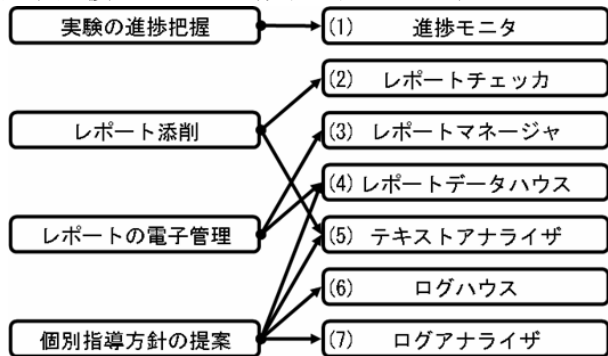


図1 基本的要件とそれを満たすための機能

具体的には、

- (1) 学生にとっては自分がどこまで進んでいるか、TA にとっては担当学生が課題をどこまで提出しているか、そして担当学生と実験全体の進捗を俯瞰できる進捗モニタ、
 - (2) レポートの形式をチェックし、内容とは関係ない部分での添削指導を行うレポートチェック、
 - (3) レポートの提出状況を閲覧でき、また未提出学生一覧の閲覧や再レポート、遅れレポートの受理と言った管理を行うレポートマネージャ、
 - (4) 添削したレポートと添削結果、提出したレポートを蓄積するレポートデータハウス、
 - (5) レポートを分析するテキストアナライザ、
 - (6) 学生の実験に関係する行動のログを収集し、蓄積するログハウス、
 - (7) 行動ログを分析するログアナライザ、
- を有し、実験遂行の労力削減に加え、個別指導の充実を支援するシステムを開発する。

5. 分析結果

レポートの提出状況と、課題提出時刻、能力に関する自己評価の関係についてログとアンケートを分析した結果を表1に示す。

表1 分析結果

	該当人数	課題提出時刻差分(秒)	アンケート回答数	レポート自己評価	総合自己評価
計	126	0.00	75	3.41	3.20
A	62	-280.53	36	3.58	3.19
B	12	161.85	7	3.14	3.14
C	36	324.66	23	3.52	3.52
D	14	239.46	9	2.67	2.44
E	2	205.04	0	0.00	0.00

レポートは3回あり、分析では提出状況別にグループ分けを行っている。レポート提出状況はA「受理を含む（BとCの学生を除く）」、B

「自主再提出を含む」、C「遅れ提出を含む」、D「3回再提出」E「3回未提出」を示している。課題提出時刻差分とは各課題の提出時刻中央値に対する差分を示し、負の値は平均よりも早いことを示している。自己評価は5段階となっており、5「向上した」、4「少し向上した」、3「どちらでもない」、2「あまり向上しなかった」、1「向上しなかった」である。

分析結果から、次の事が読み取れる。

- (1) Aは課題提出時刻が平均よりも約5分早い
- (2) Cは課題提出時刻が平均よりも約5分遅い

このことから、課題の提出時刻からレポートに対する取り組み姿勢の傾向を知ることができると思われる。また、A以外はBを除き、レポート提出状況に問題があり、その理由をレポートが完成しなかった、と仮定すると、レポートの構成や考察のヒント等を必要とする学生を実験当日中にピックアップすることができる。Bの学生群は意欲が高いと考えられ、これらのことから課題提出が平均よりも遅い学生に重点をおいて指導する効果が高いと言える。

自己評価結果から、以下が推測されるが、自己評価の理由はデータが少なく分析できていない。

- A) レポートに対する自己評価は高いが、総合に関しては高くない。総合に関しては難しく感じておらず自信には繋がらなかった。
- B) 共に評価が低い。自己評価が厳しい。
- C) 共に評価が高い。実験を難しく感じており、こなしたことで自信に繋がった。
- D) 共に評価が低い。レポートがすべて再提出であることから、自信が持てるようになるには至っていない。

実験で能力と共に自信をつけてもらうためには、Dに該当する学生を課題提出状況やレポートの内容を分析することで早期に特定し、何らかの指導方針に沿って個別指導を行うことが必要となる。

6. 今後へ向けて

形式だけでなく内容に踏み込んだレポート添削機能やリアルタイムに分析・フィードバックを行う機能を開発すると共に、心理学や教育学の知見を取り入れた個別指導の充実を図る。

参考文献

- [1] 丸山広, 中村太一: 学生実験支援システムの構築と運用—成績とシステムアクセス頻度, 習熟の過程—, 情報処理学会 研究報告, 2004-CE-76, pp. 25-32, 2004.