

平成 27 年度 修士論文概要

主査	舟橋 健司	副査	黒柳 奨	研究室	舟橋研究室
入学年度	平成 26 年度	学籍番号	26413505	氏名	伊藤 雄真

論文題目 学生の将来の成績レベル予測の向上及び特徴分析に関する研究

Study about Improvement of Forecasting Future Student's Academic Level and Analyzing Student's Feature

1 はじめに

近年 ICT の進展により生成・収集・蓄積等ができるデータ（ビッグデータ）が注目されており、教育の分野においてもデータを収集し分析するといった事例は珍しいものではなくなっている。名古屋工業大学でも IC カード出欠管理システムやコースマネジメントシステムで得られたデータを用いてデータマイニングによる分析や予測を行う研究がされており、本研究では将来の学生の成績レベルをベイジアンネットワークを用いて予測することで、予測された学生がどのような特徴を持っているか構築されたモデルから調べている。また特徴を見つけたとしても予測がしっかりできていなければ正確であるとは言えないので予測的中精度の向上を目指した。今まで予測に用いていたデータは A 年度入学者のデータだけだったが、データ数が少ないとたとえ予測的中精度が高い値を示したとしても他のデータで試した時に同じ精度が出るかわからないため、B 年度入学者のデータ、C 年度入学者のデータを加えた。これにより 2 学年度分のデータを学習データとして、余りの 1 学年度を評価データとして予測を行うといったことが確認できるため、考えられる各パターンについて検証を行った。

2 本研究で用いたデータ

本研究では、出欠データと打刻データと成績データを用いて予測を行っている。出欠データに関してはシステム上の問題からそれぞれ正確でないデータ部分が存在するので、出欠データを打刻データで補正することにより出欠補正データとして予測に用いる。出欠補正データからは月毎の打刻回数に関する変数、 D_I [1] に関する変数、入室時間と打刻時間の差の平均秒数に関する変数を拡張して用いる。ここで D_I について説明する。 D_I はある学生の 2 週間毎の打刻数を比べたときの出席率を表しており式 (1) で求められる。 D_A, D_B はある学生における直近 2 週間の打刻回数、さらにその 2 週間前の打刻回数を、 \bar{D}_A, \bar{D}_B は全学生における直近 2 週間の打刻回数の平均、さらにその 2 週間前の打刻回数の平均を表している。

$$D_I = \frac{D_A / \bar{D}_A}{D_B / \bar{D}_B} \quad (1)$$

成績データからは 1 年前期、後期における GPA 及び各学期の教科別 GPA を変数として用いる。これより出欠補正データから 33 変数、成績データから 14 変数を用いて予測する。予測開始時期は 1 年次終了時点とし 2 年次終了時点における学生の成績レベルの予測を行う。目的変数は離散化を行い、2 年次終了時点における総合 GPA (0~4) を等間隔に 5 等分したものを成績レベル (S,A,B,C,D) として設定している。S は成績が優秀であることを表しており A, B と進むにつれて成績が悪くなっていることを表している。

3 ベイジアンネットワークによる予測結果

3 学年度分のデータを用いて次のパターンにおける予測を leave one out 法による検証で行った。

- A,B 年度のデータを学習データとした予測
- B,C 年度のデータを学習データとした予測
- A,C 年度のデータを学習データとした予測

また各パターンで予測を行う際、leave one out 法 (LOO) による検証だけでなく、3 学年度のうち学習データで使用していない残りの 1 学年を評価データとし検証を行うホールドアウト法 (HO) による予測も行った。表 1 に各パターンの各検証方法で最も良かった予測的中精度を示す。

表 1: 各パターンにおける予測的中精度

LOO		HO		
学習, 評価	精度	学習	評価	精度
A+B	80.77 %	A+B	C	79.52 %
B+C	77.18 %	B+C	A	81.29 %
A+C	76.56 %	A+C	B	73.65 %

表 1 より予測的中精度が最も良いパターンは,leave one out 法では A,B 年度による学習データの予測, ホールドアウト法では B,C 年度による学習データの予測であることが分かった. またそのときの予測結果の詳細について表 2, 表 3 に示す. 縦は実際, 横は予測を表す.

表 2: A,B 年度データによる予測 (左:LOO, 右:HO)

	S	A	B	C	D
S	7	7	0	0	0
A	2	106	14	0	0
B	0	17	118	10	1
C	0	0	9	31	2
D	0	0	0	3	11

	S	A	B	C	D
S	3	4	0	0	0
A	0	52	6	0	0
B	0	10	55	6	0
C	0	0	4	19	1
D	0	0	0	3	3

A,B 年度の学習データによる予測結果は, 表 2 より leave one out 法では成績レベル S の予測的中精度は 5 割程度だが, 他の成績レベルに関してはある程度予測できているのに対し, ホールドアウト法では成績レベル D も 5 割程度しか予測できていないことが分かり, 各パターンの中で leave one out 法による予測的中精度が一番高いが成績優秀者や成績不遇者を見つけるには適していないことが分かった.

表 3: B,C 年度データによる予測 (左:LOO, 右:HO)

	S	A	B	C	D
S	13	3	0	0	0
A	5	100	14	0	0
B	0	24	104	14	1
C	0	0	7	28	5
D	0	0	0	3	12

	S	A	B	C	D
S	4	1	0	0	0
A	4	48	9	0	0
B	0	6	63	4	1
C	0	0	5	19	2
D	0	0	0	0	5

B,C 年度の学習データによる予測結果は, 表 3 より leave one out 法では成績レベル S と D の予測的中精度が高いことが確認できる. またホールドアウト法でも成績レベル S,D の予測的中精度は高く, この学習データは成績優秀者や成績不遇者を見つけるのに適していることが分かった.

4 特徴分析

本研究では成績優秀者や成績不遇者が持つ特徴を見つけないので B,C 年度の学習データによる予測の結果, 構築されたモデルよりどのような特徴を持っているのか調べた. モデルの一部分を図 1 に示す.

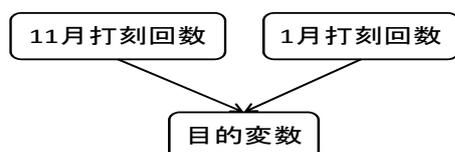


図 1: B,C 年度データによる BN モデル (抜粋部分)

図 1 より目的変数は 11 月打刻回数と 1 月打刻回数の変数によって確率的に表現できることがわかる. このときの確率結果について表 4 に示す.

表 4: B,C 年度データのモデルによる成績レベル表現

11月打刻回数	1月打刻回数	S	A	B	C	D
≤ 59.5	≤ 16	3.4%	3.4%	3.4%	24.1%	65.5%
≤ 59.5	16 - 39.5	2.6%	2.6%	28.2%	53.8%	12.8%
≤ 59.5	39.5 <	4%	28%	20%	20%	28%
59.5 - 78.5	16 - 39.5	3.4%	3.4%	44.8%	37.9%	10.3%
59.5 - 78.5	39.5 <	2.4%	32.4%	52.9%	11.9%	0.3%
78.5 <	39.5 <	9.6%	49.5%	38.1%	2.5%	0.4%

表 4 より 11 月打刻回数が 59.5 より低いと 1 月打刻回数にもよるが全体的に成績レベルが悪くなりやすいことがわかる. また 11 月打刻回数が 78.5 より高いと成績レベルは良くなりやすいことがわかる.

5 むすび

本研究では, 卒業生データを 3 学年度分用意し, ペイジアンネットワークによる予測技術を用いて検証を行った. 予測の結果, どのパターンも約 8 割ほどの予測的中精度を持っていることが分かった. またその時の成績レベル予測表をみると B,C 年度のデータによる予測に関して成績レベル S,D の予測的中精度が leave one out 法, ホールドアウト法共に高く成績優秀者や成績不振者を予測したい場合に最も適していることが分かった. またペイジアンネットワークによって作成されたモデルから, 予測したい成績レベルは 11 月打刻回数と 1 月打刻回数から確率的に表現できることが分かり, 打刻データの有用性を見つけることができた.

今後の課題として, 他学科のデータや他大学のデータを用いて本研究で行った操作を同様に検証することや, 新たなデータによる変数の追加を行い新たな変数による特徴を見つけることが挙げられる. また本研究で用いている打刻に関するデータは, 最近では出欠確認システムを導入している大学も多いためカリキュラムに縛られない変数としては有用であると言えるので, 家から大学までの通学手段や通学にかかる時間を考えた打刻データによる特徴を見つけていきたいと考えている.

参考文献

- [1] 松尾啓志, “IC カード出欠システムを用いた不登校学生早期把握と災害時人情報把握への取り組み”, サイエンティフィック・システム研究会教育環境文科会, 第 1 回会合 (2012)