

バイジアンネットワークによる要注意学生の  
半期毎の発見精度に関する検証実験

メディア系

舟橋研究室

No. 22115121

平田 大智

## 1 はじめに

近年、教育における分野でも電子化が進んでおり名古屋工業大学では、双方向型支援教育システム [1] が導入され、各学生のデータが蓄積されている。従来研究ではこれらのデータを利用し、修学指導が必要な学生いわゆる要注意学生の予測を行った。本研究では、より精度の高い予測モデルを構築するため、半期ごとのデータを利用した発見モデルの構築及び精度検証を行った。半期ごとの予測により、要注意学生に見られる特徴である急激な成績低下をより捉えることができ、多くの要注意学生を発見することが可能となる。

## 2 本研究で用いた手法

予測モデルの構築には未来予測の手法として用いられるバイジアンネットワークを利用した。バイジアンネットワークによる出力は、事後確率が閾値を越えたか否かで判定されるため、本研究では閾値を 50 %, 30 %, 事前確率の 3 つに設定した。また、用いる変数により構築されるモデルの予測精度は変わるため、属性選択を利用することでより精度の高くなる変数の取捨選択を行った。

## 3 本研究で用いたデータ

モデル構築に用いるデータは、名古屋工業大学を卒業した 2 年度分の 338 名の学生データである。元データのままでは、モデル構築に利用しがたいため、科目別 GPA データと秀や優などの各評価の獲得数をまとめた獲得成績数データ、月毎の打刻回数をまとめた打刻データの 3 種類に拡張しモデル構築に利用した。

## 4 半期ごとの要注意学生発見モデルの構築

半期毎にモデル構築をするにあたり、予測時期までの全てのデータを利用するモデルと半期ごとのデータのみを利用するモデル、半期毎のデータに前回の予測結果を加えたモデルの 3 つのパターンで発見モデルを構築し、精度検証を行った。用いるデータに関しては、科目別 GPA のみのモデルと科目別 GPA、獲得成績数、打刻データの 3 種のデータを利用したモデル構築を行った。モデルの精度検証に、Leave one out 法を利用し再現率と適合率の調和平均である F 値を比較指標とした。予測対象となる学生数は 302 名であり、その中で発見すべき要注意学生は 25 名である。

表 1, 2 は実際に構築されたモデルで各時期において最も高い精度を示した手法をまとめたものである。予測時期が遅くなるほど利用するデータ数が増えるため予測精度を示す F 値が向上していることが確認できる。データの利用範囲については予測時期により異なるため、それぞれの場合で適切な変数を選択する必

要があることが確認できる。また、用いるデータに獲得成績数と打刻データを加えた事により、要注意学生の発見数の向上が見られる。これは科目別 GPA データのみでは発見できなかった学生が獲得成績数と打刻データによって発見されていることを示しており、要注意学生の発見において獲得成績数と打刻データは有用なデータであることが確認できる。

表 1: 科目別 GPA のみを利用した各予測時期のモデル精度まとめ

予測時期	利用範囲	閾値	発見数	F 値
1 年前期	のみ	30 %	9 名	0.300
1 年後期	までの属性選択	50 %	14 名	0.491
2 年前期	のみ	50 %	14 名	0.475
2 年後期	のみ	50 %	17 名	0.607

表 2: 科目別 GPA、獲得成績数、打刻データを利用した各予測時期のモデル精度まとめ

予測時期	利用範囲	閾値	発見数	F 値
1 年前期	のみの属性選択	50 %	8 名	0.348
1 年後期	のみの属性選択	30 %	16 名	0.432
2 年前期	までの属性選択	50 %	18 名	0.554
2 年後期	までの属性選択	50 %	18 名	0.600

実際に本研究の提案モデルを 2 年後期に用いると、全体に行う場合の約 10 分の 1(35/302) の対象で、実際に修学指導を与えるべき学生の 72 % に対して修学指導を行うことができる。また、上記表からは確認できない各時期において新たに発見された学生数と発見された学生の累積人数を把握するため、半期ごとに 25 名の要注意学生が各モデルにおいてどのように発見されているのか検証したところ、半期ごとのデータによる予測のほうが累計発見数が多くなり、25 名中 22 名の要注意学生が 2 年後期までに発見されていることが確認できた。

## 5 むすび

本研究ではバイジアンネットワークを利用し、半期ごとに予測を行うことで要注意学生の発見精度の向上及び発見精度に関する検証を行った。2 年次後期までのデータを用いることで正確な予測が可能であることを示したが、成績低下が予測される学生に対し修学指導を与える場合、その時期は出来る限り早いことが望まれる。そのため、要注意学生の早い段階での特徴を捉え、予測精度のともなった早期発見モデルの構築が必要である。

### 参考文献

- [1] 伊藤宏隆, 舟橋健司, 中野智文, 内匠逸, 松尾啓志, 大貫徹, “名古屋工業大学における Moodle の構築と運用”, メディア教育研究, 4 巻, 2 号, 15-21(2008)