平成23年度 卒業論文

ベイジアンネットワークを用いた学生の修学傾向予測とその有用性の検証

指導教員
舟橋　健司　准教授
伊藤　宏隆　助教

名古屋工業大学　工学部　情報工学科
平成19年度入学　20115016番

名前　伊藤　圭佑

目 次

第1章　はじめに

2006年から始まった概算要求プロジェクト「学びの場の構築」において, 教育支援システムとしてICカード出欠管理システムと, コースマネジメントシステムが導入されている [1]. 前者では, 学生がICカードによる出席打刻をすることにより, 学生の出欠状況が学内ネットワークを通じ, サーバにデータとして蓄積される. 後者のシステムでは, 教員側が講義に関するレポートや小テストをweb上で課し, 学生がwebブラウザを通じてそれらを提出することにより, レポート, 小テストの結果や評価がサーバに蓄積される.

「学びの場の構築」プロジェクトの目的の1つが出欠データ, 学習データ及び成績データを統合したデータマイニングによる双方向学習支援システムの構築である [2]. これまで, これらのデータは教員が学生の最終成績を下す際の評価指標として用いられてきた. しかし, 情報化社会とも言われる昨今, 膨大に集積されたデータを成績指標のためだけではなく, もっと他の活用法を見出したいという向きが強まってきている. そのうちの1つの考えとして, 学生に関するデータから一人一人の修学傾向を読み取り, 何かしらの学習指導を行いたいというアイデアがある. 従来では, 学生はデータを産み出すだけの存在であり, そのフィードバックを得られていなかったが, このアイデアが現実のものとなれば, 学生側と教員側の相互関係が充実し, 学習環境の向上が期待できる. 例えば, 学生の出席状況, 課題の提出状況, 過去の講義の最終成績などのデータからある傾向を見つけ出し, その導き出された傾向によって然るべき助言を与えることで, 大学教育から脱落する危険のある学生への早期の警告や, ある特定の分野に特化しうる学生の能力を最大限に引き出すことが望めると考えられている. このような構想は様々な教育機関でも議論されており, 研究対象として熱を帯びている. 事例は枚挙に暇がないが, いくつか挙げると, 学生による授業アンケートをもとに学生個人の最終成績や, 学習状況の相互関係性を調査したもの [3] や, 講義の出席状況や課題提出状況からある学生の講義最終成績を予測したもの [4] などがある. 手法や多少の差異はあれど, これらに共通するものは, 事前に分かっているもの (データ) から, 未知の事象 (狭義的には成績) を予測するという試みである.

特に [4] の研究では, 手法としてニューラルネットワークが採用されている. ニューラルネットワークは, 人間の有する脳の構造をモデリングした関数群である. 主に予測や分類などの手法として活躍しており, 特にその予測精度は他の手法よりも優れているとされている. しかし, ニューラルネットワークの計算過程はほぼブラックボックスであり, ある予測結果に対する根拠を説明することは難しい. 例えば,「あなたの成績は段々悪くなるだろう」という予測結果を学生が受け取ったとして, その根拠が分からなければ対応に困ってしまうだろう. どうしても説得力に欠けてしまう. また, ニューラルネットワークの出力は, 意味性を持たない数値形式であるため, 表現として柔軟性が欠けているし, ニューラルネットワークの知識を持ち合わせいない人間に対し, 十分な理解が可能である形式に変換する手間が必要である. 前述の [4] の論文では, ニューラルネットワークによる出力結果をもとに,k-means法によるクラスタリング手法により, 表現力の向上と説得力の付加を試みているが,k-means法による分析は専門的な知識による解釈がその度に必要であり, システムユーザ側の作業量も増加し利便性に難があるといえる.

そこで, 成績予測の手法にベイジアンネットワークを採用することを提案する. ベイジアンネットワークは複数の変数をノードとした有向グラフと, 変数間の条件付確率で定義される確率モデル

である. 知識発見, 予測, 分類の手法として活用されている. 蓄積されたデータをもとに適切なベイジアンネットワークを構築し, 実際の成績データを入力とすることで, 以降の成績を予測することを考える.

ベイジアンネットワークを予測の手法として採用する理由として, 出力結果の直感的な分かりやすさと, 計算過程が参照できるという点が挙げられる. ベイジアンネットワークの出力形式はある事象の確率値で表現されるため, その解釈に苦労することは少ないと考えられる. また, 前述のニューラルネットワークに対し, ベイジアンネットワークは計算過程が参照できるため, 何故そのような予測結果になったのかを必要に応じて説明することが可能である. 例えば,「あなたの成績は段々悪くなるだろう」という予測に対して,「理系教科に苦手意識がある」「講義の脱落が多すぎる」などのような論拠を与えることができる. これにより予測結果の説得力が増し, 学生も今後の学習への対応を容易に決定することができる. また,WEB 上での実装も可能であり, システム実現の支障にならない. さらに, ベイジアンネットワークは知識発見 (データマイニング) の手法としても大変有望であり [5], 予測システムを活用する過程で思わぬ知識を獲得できたり, 学習状況に関する問題が発生したとき, 構築された最適モデルに観測情報を意図的に入力し, 変数の確率値の変化を観察することにより, 問題の原因追求の助けにもなる. ベイジアンネットワークの特徴は, 当研究の目的と符合する事柄が多い.

ただし, ベイジアンネットワークには「グラフ構造をどう決定するか」という潜在的問題を抱えている. 前述の通り, ベイジアンネットワークは有向グラフの 1 つである. このグラフ構造は変数間の関係性ともいえるが, これを無闇に決定しても, 期待した予測結果を得ることはできない. ベイジアンネットワークを構築するときは, 一般的に, 専門的な知識をもとに各変数の因果関係を矢印に見立て有向グラフを構築する. つまり, 各変数の因果関係を把握していれば, グラフ構造は決定できる. 人間は数ある生体のなかでも, 事象の関係性を見抜くのに優れている存在であるが, その対象が大量の多変数データとなれば話は別である. そこで, グラフ構造に用いる変数とそれらの因果関係を, データマイニングの手法によって獲得する. データマイニングとは計算機による計算による解析技術であり, 多数のレコードで形成されたデータから特殊な知識を抽出するのに有望な技術および理論群である. データマイニングの概要と採用した具体的手法は後の第 2 章で述べるものとする. データマイニングにより, 各事象間の因果関係を調べたり, 下準備として連続値であるデータを有意な区分に離散化したりする.

本研究では, 学習データとして, ある年度入学の学生の 2 年分を用いている. これを元に学習指導に適切なベイジアンネットワークを構築し,1 年終了時点の成績データを入力として,2 年次の成績予測を行う. 多様な学習指導を目指すため, 複数のモデルを構築した. それぞれのモデルがどれほどの予測精度を持つのかは,leave one out 法により予測的中率を評価することで検証した. また, 評価されたモデルが学習指導に応用することができるかどうかを様々な観点から議論した.

本論文の構成を説明する. 次の第 2 章ではベイジアンネットワークと, 本研究で用いたデータマイニングの手法を概説する. 第 3 章ではベイジアンネットワークに必要な変数の定義について述べる. 第 4 章では学習指導の構想に沿った最適モデルを構築し, 構築されたモデルを評価及び検証する. 第 5 章では本研究の課題点と展望を述べて, 本論文の締めとする.

ちなみに, 本研究では, 個人を特定できる情報 (氏名や学籍番号など) を一切排除した上で研究をおこなっており, この研究報告によって個人情報が侵害されることはないことをここに付記する.

第2章　本研究に用いる手法の概論

本研究では, 予測手法としてベイジアンネットワークを採用している. また, そのベイジアンネットワーク [6][7][8] を構築するにあたり, データマイニング [9][10] の手法を取り入れている. 本章では, それらの概要を記している.

2.1　ベイジアンネットワーク

2.1.1　概要

ベイジアンネットワークは, 複数個の確率変数の定性的な依存関係をグラフ構造で表現する確率モデルである [6]. モデル上での確率計算により, 推論や不確実性を含む事象の予測などが可能となる. この手法は, 不確実性を扱う問題, 例えば障害診断やヒューマンモデリングなどに適用されている.

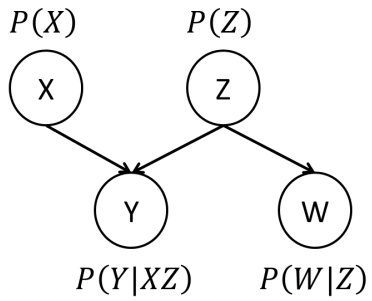

図 2.1: ベイジアンネットワークの例

ベイジアンネットワークは (1) 確率変数 (2) 因果関係を矢印と見立てた有向グラフ (3) 条件付確率及び事前確率 で定義される. 図 2.1 のようなベイジアンネットワークを仮定した場合, 構成要素としては, 確率変数 {X,Y,Z,W}, 条件付確率及び事前確率 $\{P(X), P(Y|X,Z), P(Z), P(W|Z)\}$, そして図にあるような有向グラフで表現される. 事前確率及び条件付確率の計算は, その発生確率が明らかであるときはユーザが設定し, そうでない場合はデータから推定する.

この 3 つの要素の決定は, ベイジアンネットワークの決定を意味すると言える. ただし ,3 要素の決定はユーザの判断に委ねられているため, 最適なベイジアンネットワークを構築する際は, ユーザによる適切な判断を要する. どのような判断をすればよいのか, という問題は方々で議論されている. 最近ではデータそのものから最適なベイジアンネットワークを構築するアルゴリズムなどが多数研究されているが, そのデータそのものをどのような形式にすればよいのか, という疑問も生まれており, ベイジアンネットワーク構築の完全なる自動化は難しいものとされている. そこで, データそのものの形や,3 要素の決定の段階に, データマイニングの考え方を取り入れるという向き

が強まってきている. データマイニングによって, もとあるデータをより有用な形式に書き換えることで, ベイジアンネットワーク構築を支援することができる.

2.1.2　ベイジアンネットワークによる予測

ベイジアンネットワークの応用例に, 未知事象の推論, 予測が挙げられる. それらは, ベイジアンネットワークによる確率計算によって実現されている. 一般的には, 観測された情報 e から, 求めたい確率変数 X の確率値, すなわち事後確率 $P(X|e)$ を求めることで, ある状況下の確率変数 X の期待値や確信度を評価する. この一連の確率計算を, ある未知の事象の決定と解釈している.

例えば, 図 2.1 のモデルにおける全確率変数の結合確率 $P(X, Y, Z, W)$ は, モデル構造を利用して,

$$P(X, Y, Z, W) = P(X)P(Y|X, Z)P(Z)P(W|Z) \tag{2.1}$$

と表現できる. ベイジアンネットワークを利用した予測は, この結合確率の周辺化によって周辺事後確率を計算することで行われる. 例えば $Y = 1, W = 1$ が事前の情報として与えられたとき, $X = 1$ となる事後確率を求めたい場合は, 以下のような計算を行う.

$$\begin{aligned} P(X = 1|Y = 1, W = 1) &= \frac{P(X = 1, Y = 1, W = 1)}{P(Y = 1, W = 1)} \\ &= \frac{\sum_Z P(X = 1)P(Y = 1|X = 1, Z)P(Z)P(W = 1|Z)}{\sum_X \sum_Z P(X)P(Y = 1|X, Z)P(Z)P(W = 1|Z)} \\ &= \frac{P(X = 1)\sum_Z P(Y = 1|X = 1, Z)P(Z)P(W = 1|Z)}{\sum_X P(X)\sum_Z P(Y = 1|X, Z)P(Z)P(W = 1|Z)} \end{aligned} \tag{2.2}$$

この計算により, 観測情報 $e = \{Y = 1, W = 1\}$ から, 事後確率 $P(X|e)$ を得ることができる.

しかし, このような周辺化による確率計算では確率変数の数に応じて, 計算コストが指数オーダーで増加する. そこで計算コストを削減するため, 網羅的な計算方法である周辺化ではなく, あるノードとその親ノードと子ノードに注目した局所的確率計算により事後確率を得る方法を採用する. この方法は確率伝播法と呼ばれている. 確率伝播法の解説のため, ある複雑な構造を持つモデルの一部分を抜粋し, その部分における確率計算を考える.

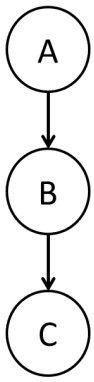

図 2.2: 単純なモデルの一部分

ここで, 確率変数 B の事後確率を求めたいとする. このとき, ノード A はノード B の親ノードであり, ノード C はノード B の子ノードである.B の親ノード群 (B よりも上流に存在するノード群)

に入力される観測情報と,B の子ノード群 (B よりも下流に存在するノード群) に入力される観測情報としてそれぞれ e^+, e^- を与える. このとき, ベイズの定理より, 事後確率 $P(B|e^+, e^-)$ は以下のように計算できる.

$$P(B|e^+, e^-) = \frac{P(e^-|e^+, B)P(B|e^+)}{P(e^+, e^-)}$$

このとき, $\frac{1}{P(e^-|e^+)}$ は確率変数 B に依らないものだから, 定数 $\frac{1}{P(e^-|e^+)} = \alpha$ として扱える. よって求めたい事後確率は,

$$P(B|e^+, e^-) = \alpha P(B|e^+)P(e^-|B) \tag{2.3}$$

と記述できる. ここで親ノードからの寄与確率値を $P(B|e^+) = \pi(B)$ とおき, 子ノードからの寄与確率値を $P(e^-|B) = \lambda(B)$ とおく. まず,$\pi(B)$ は定義済みである $P(B|A)$ と $P(A|e^+)$ の周辺化により計算が可能である.

$$\pi(B) = \sum_A P(B|A)P(A|e^+) \tag{2.4}$$

このとき, ノード A が親ノードを持たない最上流の親ノードであれば, 予め用意された事前確率を与える. 観測値が与えられているのならば, その値は決定できる. それ以外の場合, つまりノード A には入力が無く, かつノード A の上流に親のノードが存在するとき, 式 (2.4) を再帰的に適用することでその値を求めることが出来る.

同様に $\lambda(B)$ も, 定義済みの $P(C|B)$ とノード C の状態の周辺化によって計算が可能である.

$$\lambda(B) = \sum_C P(C|B)P(e^-|B, C) \tag{2.5}$$

ここで, 下流に与えられた情報 e^- と変数 B は独立関係にあることを考慮すると, 以下のように式を書き換えることが出来る.

$$\lambda(B) = \sum_C P(C|B)P(e^-|C) \tag{2.6}$$

この式の計算法法は $\pi(B)$ の場合とほぼ同様であり, 式が定まるまで再帰的に適用される.

このように, あるノードの上流と下流からの確率伝播を局所的に考慮することで, 計算量が大幅に削減される.

ただし, 確率伝播法はどのようなグラフ構造でも厳密な値を算出するとは限らない. ベイジアンネットワークを無向グラフとみなしたとき, 閉路が存在しないものを singly connected と呼び, そうでない, すなわち閉路が存在するものを multiply connected と呼ぶ. グラフ構造が singly connected である場合, 確率伝播法で厳密解を算出することができるが,multiply connected の場合は計算が収束する保証がない.

しかし,multiply connected であるグラフに対して確率伝播法を適用し, ループを避けるため計算回数を任意に定めてみると, 厳密解は得られなくとも大抵の場合は近似解を得ることが可能であることが指摘されている. さらに,multiply connected なグラフを等価な singly connected に変換した上で確率伝播法を適用するという,Janction Tree アルゴリズムという確率計算法など, 多数提案されている. 商用ソフトウェアである Hugin[11] はこのアルゴリズムを応用し, ベイジアンネットワークの確率計算システムを構築した. これら研究報告により, ベイジアンネットワーク関連の技術発展及びシステム開発が多方で進められている.

2.1.3　有向グラフ構造とその学習

ベイジアンネットワークは有向グラフの 1 つである. ベイジアンネットワークを予測の手法として用いたいとき, 有向グラフ構造はその結果を大きく左右する. 一般的に, 経験則やユーザが所持している専門的知識により手作業でモデルを構築していくが, 特別な知識を持ち合わせていなくても, データそのものからグラフ構造を決定したいという要求もある. ベイジアンネットワークにおける有向グラフ構造にはいくつかの種類がある. それぞれの構造の簡単な説明と学習方法についてここでは述べる.

Naive Bayes

Naive Bayes は目的変数を木構造における根の部分に置き, 他の変数を根ノードの葉としたものである. ベイジアンネットワークにおいて最もシンプルなグラフ構造を有していると言える.

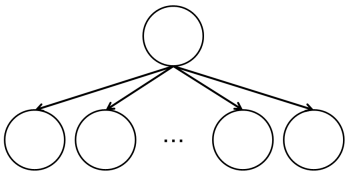

図 2.3: Naive Bayes の例

Naive Bayes の簡易な構造から, 構造学習の作業が不要であり, 条件付確率の推定だけを行えば良い. ただし, 葉となるノードを無闇に増やしたり, 根拠もなく選択をしたとしても予測精度が上がる保証は無く, むしろ下がる恐れもある. 故に目的変数を説明するに適切な変数を選択しなければならない.

TAN

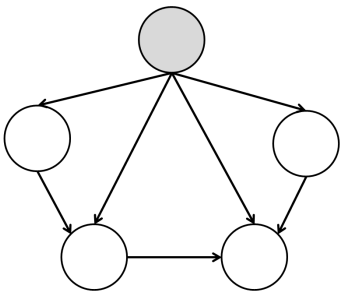

図 2.4: TAN の例

TAN は Tree Augmented Network の略記である.Naive Bayes と似た基本的構造を有しているが, 目的変数以外の変数が目的変数以外にもう 1 つの変数を親ノードとして選択しなければならないという制約を課している. その選択基準として条件付相互情報量が考慮されている. ある確率変

数 X_i, X_j の相互情報量 $I(X_i, X_j)$ は, 確率値 Pr を用いて

$$I(X_i, X_j) = -\sum_{x_i \in X_i} \sum_{x_j \in X_j} Pr(x_i, x_j) log \frac{Pr(x_i, x_j)}{Pr(x_i) Pr(x_j)} \tag{2.7}$$

と表すことができる. ここで, 目的変数を C, それ以外の 2 つの変数を X_i, X_j としたとき, 目的変数 C が与えられたという条件付相互情報量は

$$I(X_i, X_j|C) = -\sum_{x_i \in X_i} \sum_{x_j \in X_j} \sum_{c \in C} Pr(x_i, x_j, c) log \frac{Pr(x_i, x_j|c)}{Pr(x_i|c) Pr(x_j|c)} \tag{2.8}$$

と記述できる. この値が最大となる X_i, X_j を求めることにより,TAN のグラフ構造を決定することができる.

Free Network

親ノード, 子ノード数の制限がないグラフ構造を有したモデルを,Free Network と総称する. 例えば, 図 2.1 も Free Network と呼ばれるものの 1 つである. 制限がないと雖も, 無計画に複雑なグラフを構築したところで期待する予測精度が得られるとは限らない. また, あるノードに対する親ノードが増えるにつれ, 必要となる条件付確率が爆発的に増え, 条件付確率値に欠損が生まれる可能性もある. そのため, 親ノードの個数などを制限した上で構造学習をする場合が多い.

Free Network の構造学習法は各方面で研究されており, 学習アルゴリズムも多数発表されている. 上記の TAN と同様に, 各変数間の相互情報量を考慮することで接続する変数を決定したものもある. 全般的に, 局所的に最適な木構造を生成することにより, 結果的にある程度適したグラフ構造を取得する手法がほとんどである. より優れたグラフ構造を得るには, やはり元となるデータを整備することが重要である.

本研究では, 用いるデータの総量が一般的に少量であるため, ベイジアンネットワークのグラフ構造として主に Naive Beyes と TAN を採用している.

2.2　データマイニング

データマイニングとは, 多量のデータから有用な知識を発掘する技術の総称である. 近年の計算機の安価傾向と性能向上により, 単体の計算機でも多量のデータを処理することが可能となったため, 急速に発展を遂げている分野である. データマイニングは以下の過程により実行される.

1. データの収集及び処理
2. パターン発見
3. 発見されたパターンの解釈

本研究において,1. のデータは学生の成績データに相当し,2. で得られるパターンはベイジアンネットワークを構築するに適切なデータ形式や正規化方法などに相当し,3. の解釈は結果として構築されたモデルの評価として読み取れる. ベイジアンネットワークの構築の過程と符号する点が多い. 最適なベイジアンネットワークの構築の支援として, データマイニングの理論や手法が有効であるため, 本研究ではベイジアンネットワークの構築の際にいくつかのデータマイニングに関連する手法や理論を採用している.

以下より, 主に採用したデータマイニングの具体的手法を概説する.

2.2.1　情報利得

本研究では, 元となるデータから多くの変数を定義している. しかし, ベイジアンネットワークの構築において, 変数が多ければ多いほど良いというわけではなく, 不必要な変数は予測の際のノイズになることがある. これを避けるため, 数ある変数の中から, 必要な変数を取捨選択する必要があり, この処理は「属性選択」というデータマイニングに内包される処理の 1 つである.

属性選択をする際には, 何かしらの指標を設けなければいけない. その指標の 1 つとして情報利得が挙げられる. 情報利得は,「2 つの確率分布の距離」とよく記述される (ただし距離の公理を満たしていないため, あくまで表現としての「距離」である). 情報利得の定義は 2 つの確率分布 P, Q と, 以下の式で与えられる.

$$D(P||Q) = \sum_{x} P(x) log \frac{P(x)}{Q(x)} \tag{2.9}$$

$D(P||Q)$ が 2 つの確率分布 P, Q の情報利得である, と表すことができる.

情報利得を用いた変数選択手法として,CFS(correlation besed feature selection)[12] が挙げられる. ある変数と関連性の高い変数を抜粋する際の指標として有効である.CFS は以下の式で与えられる. k は変数の個数,Z は目的変数を指す. この CFS を最大化するような変数 Y_i を抜粋する.

$$CFS = \frac{\sum_{i=1}^{k} SU(Y_i, Z)}{\sqrt{k + \sum_{i=1}^{k} \sum_{j \neq i, j=1}^{k} SU(Y_i, Y_j)}} \tag{2.10}$$

ちなみに SU は情報量 H と情報利得 D で求めることができる.

$$SU(Y, Z) = 2 * \frac{D(Y||Z)}{H(Y) + H(Z)}$$

2.2.2　決定木

決定木は主に予測や分類に用いられる木構造のグラフである. 利点として, 視覚的な直感により, 変数間の関係性などが理解できることや, 原理が単純であるがゆえに応用が比較的容易であることが挙げられる. 目的変数が離散型, カテゴリ型である場合, 決定木を分類木と呼び, 数値型である場合は回帰木と呼ばれる. 分類木は, 木の根から葉までに記述されたルールに沿って葉の方に下っていくことにより, 事例がどのようなグループに属するか分類するものである. 分類木の場合は, 目的変数を必ずカテゴリ型にしなければならない. つまり予測したい変数が数値型であった場合は離散化の必要性が生じる. 目的変数以外のものは, その変数の型を問わない. 図 2.4 は有名な決定木の出力例である. 天気, 湿度, 風の強さによって, 顧客がゴルフをするかしないかを予測する分類木である. 湿度のような数値型も扱えることに注目したい.

決定木はしばしば, データから学習することで取得される. この行為を決定木学習という. 決定木を構築する際, 全パス長を最短にしつつ予測精度を高く保つことが理想とされるが, この問題は NP 困難であるとされている. ただし各属性間の情報利得を最大にするような決定木を構築することによって, ある程度実用的な大きさの決定木を得ることができる. 以下に, 決定木構築の大まかな仮定を記述する.

1. 全体の情報利得が最大となるような変数を根ノードにセットし, 変数の属性値に応じて分岐を作る.
2. データ集合を各分岐に応じて部分集合に分割して子ノードを作成し, その子ノードを根ノードにする.
3. 1.2. のプロセスを再帰的に繰り返し, 決定木を成長させる.
4. 子ノードすべての事象が同一の属性値に属していれば, 決定木の成長を止める.

ただし, 可能な限り決定木を成長させると, 決定木が必要以上に大きくなり, 視覚的直感による理解が損なわれることや, オーバーフィッティングの問題が発生することがある. そのため, ある程度運用可能な分類精度が期待できるならば, 決定木構築後に必要性に乏しい葉を切り落とすことがある. この行為は枝刈りと呼ばれ, 前述の問題を解消する手立ての一つである.

決定木は数値型変数を, 目的変数分類における有意な区分に離散化する役目も果たす. この決定木の特性が本研究に用いる数値型データの離散化に一役買っている.

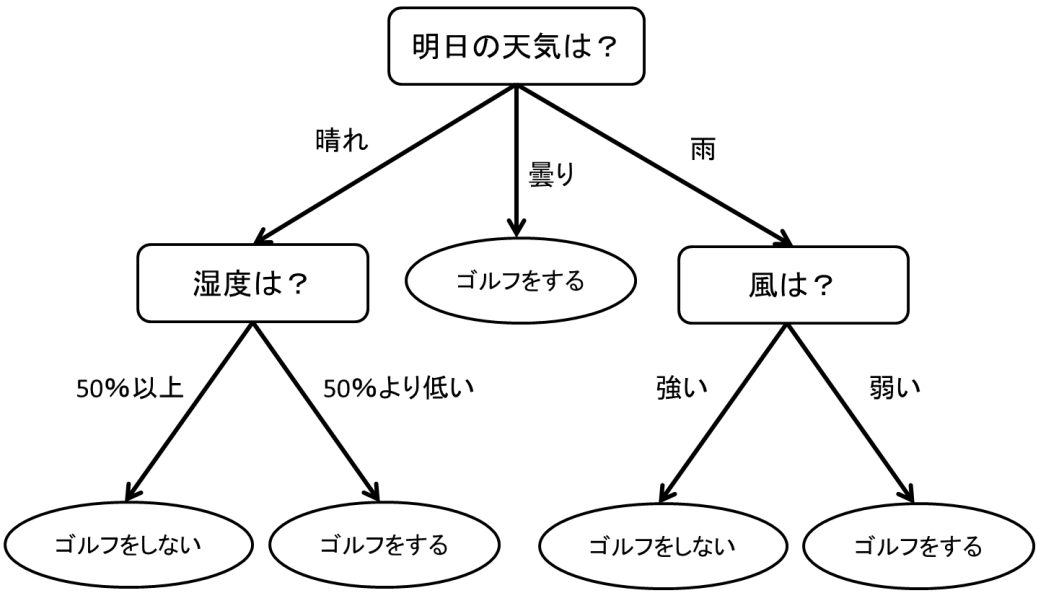
明日の天気は？
晴れ
曇り
雨
湿度は？
ゴルフをする
風は？
50%以上
50%より低い
強い
弱い
ゴルフをしない
ゴルフをする
ゴルフをしない
ゴルフをする

図 2.5: 決定木の出力例

第3章　モデル構築に用いるデータについて

本研究では学習指導の手法としてベイジアンネットワークを採用している. ベイジアンネットワークによる予測を行う前に, 未知事象を予測をするに最適なベイジアンネットワークの構築しなければならないが, この試みが期待する成果を挙げられるかどうかは, 用いるデータの質が大きく左右すると言っても過言ではない. 本章では, 元となるデータの概要と, その正規化について解説をする.

3.1　元となるデータの概要

ベイジアンネットワークを構築するには, ある程度の量を有するデータを用いて, 確率変数, 条件付確率, 変数間の依存関係とも言える有向グラフ構造を学習せねばならない. 用いるデータでは番号, 成績, 授業名, 開講学期が 1 レコードとなっている.

学籍番号は, 個人が特定できないように予め暗号化されている. 専門教科と演習授業に関しては,「専門 1」「演習 1」のように講義が特定できないよう名前が伏せられている. このような形式のレコードが合わせて 8000 レコード以上ある. しかし, このままの状態では情報量に乏しく, 満足できるベイジアンネットワークのモデルを構築できそうにないため, 講義シラバス等をもとにデータを拡張及び正規化する.

3.2　データの拡張

ベイジアンネットワークの構築に用いる学習データは, 情報量に乏しく整然としていないものよりも, 情報量が豊富で正規化されたものの方が適している. 図 3.1 のような単純な形式を持つデータ群から, より学習用に優れたデータに拡張及び正規化する.

ここで,GPA と呼ばれる成績指標について説明を加える.GPA とは数多くの大学で採用されている成績評価方法の 1 つである. 欧米では積極的に重要されており, 近年では GPA を学力の指標として導入する日本の大学も増えてきており, データ提供元である名古屋工業大学もその例外ではない.

GPA は成績評価に割り振られた得点と, 講義によって定められている単位数によって計算する事ができる. 前述した成績段階の表記である S,A,B,C,D,X にそれぞれ,4 点,3 点,2 点,1 点,0 点,0 点を成績得点として割り振り, 以下の計算式により,GPA の値を算出する.

$$GPA = \frac{\sum_{\text{受講した講義全て}} (\text{講義の成績得点}) * (\text{講義の単位数})}{\sum_{\text{受講した講義全て}} (\text{講義の単位数})} \tag{3.1}$$

例えば, 受講した全ての講義で成績 S を獲得したならば, 得られる GPA 値は 4 となる. 逆に, 受講した全ての講義で不合格, もしくは失格となった場合, 得られる GPA 値は 0 となる.GPA の特徴として, ただ闇雲に多くの講義を受講しても, 全ての講義で単位を取らなければ GPA は一向に良くな

らないということが挙げられる. だから, 受講数が多ければ良いというわけではなく,1 つ 1 つの講義に真面目に取り組まなければならない.GPA 値は学生の受講態度の評価基準になると言える.

元のデータ形式では, 成績評価と講義の種類を調べることができるため, 式 (3.1) より, 学生個人の GPA 値を計算することができる. ちなみに, 専門教科は講義に関わらず単位数が 2 であるので,GPA 値の計算が講義の匿名性に影響されることはない.1 年終了時の GPA 値,2 年終了時の GPA 値の計算はもちろんのこと, 各授業の教科と開催時期が元データから参照できるため, ある教科の講義成績のみで算出された GPA 値, 前期分の GPA 値, 後期分の GPA 値というように, 様々な授業区分やある期間における GPA 評価を得ることができる. また, ある期間内に取得した S 成績の個数といった値も計算によって獲得できる. 以上のような作業によって得られた値の群を 1 つの変数とすることで, データ全体の情報量を増強していく. 以下の表 3.1 に, 考えられる GPA 値のパターンとその他計算可能であった数値を列挙した. ここで,「入力」の欄に○がついているものは, 成績予測時に入力データとして用いることができる変数であり, 対して○がついていないものは, 目的変数になるか, モデルの学習時のみに用いられる変数である.

表 3.1: 変数の定義一覧その 1

番号	表記	意味	入力
1	1 年 GPA	1 年次に受講した講義の GPA 値	○
2	1 年前期 GPA	1 年次の前期に受講した講義の GPA 値	○
3	1 年後期 GPA	1 年次の後期に受講した講義の GPA 値	○
4	1 年体育 GPA	1 年次に受講した体育教科 GPA 値	○
5	1 年体育前期 GPA	1 年次の前期に受講した体育教科 GPA 値	○
6	1 年体育後期 GPA	1 年次の後期に受講した体育教科 GPA 値	○
7	1 年人文 GPA	1 年次に受講した「人間文化」に分類される講義の GPA 値	○
8	1 年人文前期 GPA	1 年次の前期に受講した「人間文化」に分類される講義の GPA 値	○
9	1 年人文後期 GPA	1 年次の前期に受講した「人間文化」に分類される講義の GPA 値	○
10	1 年外国語 GPA	1 年次に受講した外国語に関する講義の GPA 値	○
11	1 年外国語前期 GPA	1 年次の前期に受講した外国語に関する講義の GPA 値	○
12	1 年外国語後期 GPA	1 年次の後期に受講した外国語に関する講義の GPA 値	○
13	1 年理系 GPA	1 年次に受講した理系基礎 (数学, 理科系) の講義の GPA 値	○
14	1 年前期理系 GPA	1 年次の前期に受講した理系基礎 (数学, 理科系) の講義の GPA 値	○
15	1 年後期理系 GPA	1 年次の後期に受講した理系基礎 (数学, 理科系) の講義の GPA 値	○
16	1 年数学 GPA	1 年次に受講した数学系の講義の GPA 値	○
17	1 年数学前期 GPA	1 年次の前期に受講した数学系の講義の GPA 値	○
18	1 年数学後期 GPA	1 年次の後期に受講した数学系の講義の GPA 値	○
19	1 年理科 GPA	1 年次に受講した理科系の講義の GPA 値	○
20	1 年理科前期 GPA	1 年次の前期に受講した理科系の講義の GPA 値	○

21	1 年理科後期 GPA	1 年次の後期に受講した理科系の講義の GPA 値	○
22	1 年専門 GPA	1 年次に受講した専門科目の講義の GPA 値	○
23	1 年専門前期 GPA	1 年次の前期に受講した専門科目の講義の GPA 値	○
24	1 年専門後期 GPA	1 年次の後期に受講した専門科目の講義の GPA 値	○
25	1 年 S	1 年次に獲得した成績評価 S の個数	○
26	1 年前期 S	1 年次の前期に獲得した成績評価 S の個数	○
27	1 年後期 S	1 年次の後期に獲得した成績評価 S の個数	○
28	1 年 A	1 年次に獲得した成績評価 A の個数	○
29	1 年前期 A	1 年次の前期に獲得した成績評価 A の個数	○
30	1 年後期 A	1 年次の後期に獲得した成績評価 A の個数	○
31	1 年 B	1 年次に獲得した成績評価 B の個数	○
32	1 年前期 B	1 年次の前期に獲得した成績評価 B の個数	○
33	1 年後期 B	1 年次の後期に獲得した成績評価 B の個数	○
34	1 年 C	1 年次に獲得した成績評価 C の個数	○
35	1 年前期 C	1 年次の前期に獲得した成績評価 C の個数	○
36	1 年後期 C	1 年次の後期に獲得した成績評価 C の個数	○
37	1 年 D	1 年次に獲得した成績評価 D の個数	○
38	1 年前期 D	1 年次の前期に獲得した成績評価 D の個数	○
39	1 年後期 D	1 年次の後期に獲得した成績評価 D の個数	○
40	1 年 X	1 年次に獲得した成績評価 X の個数	○
41	1 年前期 X	1 年次の前期に獲得した成績評価 X の個数	○
42	1 年後期 X	1 年次の後期に獲得した成績評価 X の個数	○
43	2 年 GPA	2 年次に受講した講義の GPA 値	
44	2 年前期 GPA	2 年次の前期に受講した講義の GPA 値	
45	2 年後期 GPA	2 年次の後期に受講した講義の GPA 値	
46	2 年人文 GPA	2 年次に受講した「人間文化」に分類される講義の GPA 値	
47	2 年人文前期 GPA	2 年次の前期に受講した「人間文化」に分類される講義の GPA 値	
48	2 年人文後期 GPA	2 年次の前期に受講した「人間文化」に分類される講義の GPA 値	
49	2 年外国語 GPA	2 年次に受講した外国語に関する講義の GPA 値	
50	2 年外国語前期 GPA	2 年次の前期に受講した外国語に関する講義の GPA 値	
51	2 年外国語後期 GPA	2 年次の後期に受講した外国語に関する講義の GPA 値	
52	2 年理科 GPA	2 年次に受講した理科系の講義の GPA 値	
53	2 年理科前期 GPA	2 年次の前期に受講した理科系の講義の GPA 値	
54	2 年理科後期 GPA	2 年次の後期に受講した理科系の講義の GPA 値	
55	2 年専門 GPA	2 年次に受講した専門科目の講義の GPA 値	
56	2 年専門前期 GPA	2 年次の前期に受講した専門科目の講義の GPA 値	
57	2 年専門後期 GPA	2 年次の後期に受講した専門科目の講義の GPA 値	

ここで, さらに変数の種類を増やしていくことを考える. 変数の過多がモデル学習の妨げになることもあるが, 学習前に適切な変数を取捨選択することで対応していく.

例えば, 表 3.2 の変数番号 2,3 はそれぞれ 1 年前期の GPA 値と後期の GPA 値を表している. では,(変数番号 3)−(変数番号 2) を計算して得られた値は, 何を意味しているのだろうか. 厳密には断言できないが, この値が「1 年次の成績成長度合い」を表すのではないか, と考えられる. もし (変数番号 3)−(変数番号 2) の値が 0 より大きければ,(前期の GPA) < (後期の GPA) であるのは明らかであり, 成績が向上していることを示しうる. 故に, この値が大きければ直近の未来である 2 年前期の成績もさらに向上していることが期待できると考えられる. また,(変数番号 10)−(変数番号 1) の値は,(1 年次の専門 GPA)−(1 年次 GPA) に相当するが, この計算は何を意味するかを考える.1 年次の GPA とは, 言い換えれば, 各教科の GPA の平均であるから,(変数番号 10)−(変数番号 1) の値が大きければ,「その学生は専門教科が比較的得意である」と記述できる可能性が高くなる.(変数番号 3)−(変数番号 2) や (変数番号 10)−(変数番号 1) という式によって得られる値は, ベイジアンネットワークの構築において有用なものになる見込みが十分にあると言える.

このように, 複数の変数からさらに意味を持ちそうな変数を定義することが可能であり, 元のデータがより学習に有効でありそうな形へと変質していくのが確認できるだろう. 表 3.2 にある変数を因子として用いて新しく定義した変数を, 表 3.2 に示す.

表 3.2: 変数の定義一覧その 2

番号	表記	意味	入力
58	1 年後期-前期 GPA	1 年次後期の GPA 値から 1 年次前期の GPA 値を減算したもの	○
59	1 年体育後期-前期 GPA	1 年次後期に受講した体育の GPA 値から 1 年次前期に受講した体育の GPA 値を減算したもの	○
60	1 年人文後期-前期 GPA	1 年次後期に受講した「人間文化」に分類される講義の GPA 値から 1 年次前期に受講した「人間文化」に分類される講義の GPA 値を減算したもの	○
61	1 年外国語後期-前期 GPA	1 年次後期に受講した外国語に関係する講義の GPA 値から 1 年次前期に受講した外国語に関係する講義の GPA 値を減算したもの	○
62	1 年理系後期-前期 GPA	1 年次後期に受講した理系基礎の講義の GPA 値から 1 年次前期に受講した理系基礎の講義の GPA 値を減算したもの	○
63	1 年数学後期-前期 GPA	1 年次後期に受講した数学系の講義の GPA 値から 1 年次前期に受講した数学系の講義の GPA 値を減算したもの	○
64	1 年理科後期-前期 GPA	1 年次後期に受講した理科系の講義の GPA 値から 1 年次前期に受講した理科系の講義の GPA 値を減算したもの	○
65	1 年専門後期-前期 GPA	1 年次後期に受講した専門教科の講義の GPA 値から 1 年次前期に受講した専門教科の講義の GPA 値を減算したもの	○

66	1 年体育 GPA-1 年 GPA	1 年次後期に受講した体育の GPA 値から 1 年次の GPA を減算したもの	○
67	1 年人文 GPA-1 年 GPA	1 年次後期に受講した「人間文化」に分類される講義の GPA 値から 1 年次の GPA を減算したもの	○
68	1 年外国語 GPA-1 年 GPA	1 年次後期に受講した外国語に関係する講義の GPA 値から 1 年次の GPA を減算したもの	○
69	1 年理系 GPA-1 年 GPA	1 年次後期に受講した理系基礎の講義の GPA 値から 1 年次の GPA を減算したもの	○
70	1 年数学 GPA-1 年 GPA	1 年次後期に受講した数学系の講義の GPA 値から 1 年次の GPA を減算したもの	○
71	1 年理科 GPA-1 年 GPA	1 年次後期に受講した理科系の講義の GPA 値から 1 年次の GPA を減算したもの	○
72	1 年専門 GPA-1 年 GPA	1 年次後期に受講した専門教科の講義の GPA 値から 1 年次の GPA を減算したもの	○
73	1 年後期 S-1 年前期 S	1 年次後期に取得した成績評価 S の個数から 1 年次前期に所得した成績評価 S の個数を減算したもの	○
74	1 年後期 A-1 年前期 A	1 年次後期に取得した成績評価 A の個数から 1 年次前期に所得した成績評価 A の個数を減算したもの	○
75	1 年後期 B-1 年前期 B	1 年次後期に取得した成績評価 B の個数から 1 年次前期に所得した成績評価 B の個数を減算したもの	○
76	1 年後期 C-1 年前期 C	1 年次後期に取得した成績評価 C の個数から 1 年次前期に所得した成績評価 C の個数を減算したもの	○
77	1 年後期 D-1 年前期 D	1 年次後期に取得した成績評価 D の個数から 1 年次前期に所得した成績評価 D の個数を減算したもの	○
78	1 年後期 X-1 年前期 X	1 年次後期に取得した成績評価 X の個数から 1 年次前期に所得した成績評価 X の個数を減算したもの	○
79	2 年後期-前期 GPA	2 年次後期の GPA 値から 2 年次前期の GPA 値を減算したもの	
80	2 年人文後期-前期 GPA	2 年次後期に受講した「人間文化」に分類される講義の GPA 値から 2 年次前期に受講した「人間文化」に分類される講義の GPA 値を減算したもの	
81	2 年外国語後期-前期 GPA	2 年次後期に受講した外国語に関係する講義の GPA 値から 2 年次前期に受講した外国語に関係する講義の GPA 値を減算したもの	

82	2 年理科後期-前期 GPA	2 年次後期に受講した理科系の講義の GPA 値から 2 年次前期に受講した理科系の講義の GPA 値を減算したもの	
83	2 年専門後期-前期 GPA	2 年次後期に受講した専門教科の講義の GPA 値から 2 年次前期に受講した専門教科の講義の GPA 値を減算したもの	
84	2 年人文 GPA-2 年 GPA	2 年次後期に受講した「人間文化」に分類される講義の GPA 値から 2 年次の GPA を減算したもの	
85	2 年外国語 GPA-2 年 GPA	2 年次後期に受講した外国語に関係する講義の GPA 値から 2 年次の GPA を減算したもの	
86	2 年理科 GPA-2 年 GPA	2 年次後期に受講した理科系の講義の GPA 値から 2 年次の GPA を減算したもの	
87	2 年専門 GPA-2 年 GPA	2 年次後期に受講した専門教科の講義の GPA 値から 2 年次の GPA を減算したもの	
88	1 年外国語 GPA-1 年理系 GPA	1 年次に受講した外国語に関する講義の GPA から 1 年次に受講した理系基礎の講義の GPA を減算したもの	○
89	1 年外国語 GPA-1 年専門 GPA	1 年次に受講した外国語に関する講義の GPA から 1 年次に受講した専門教科の講義の GPA を減算したもの	○
90	1 年理系 GPA-1 年専門 GPA	1 年次に受講した理系基礎の講義の GPA から 1 年次に受講した専門教科の講義の GPA を減算したもの	○
91	教科別成績分散	教科別の GPA を標本として分散を計算したもの	○
92	成績評価分散	成績評価のバラつき具合を分散によって計算したもの	○
93	体育未単位あり	1 年の体育講義に成績評価 D か X が存在するかどうか.{YES, NO}	○
94	人文不合格あり	1 年の「人間文化」に関する講義に成績評価 D が存在するかどうか.{YES, NO}	○
95	人文失格あり	1 年の「人間文化」に関する講義に成績評価 X が存在するかどうか.{YES, NO}	○
96	外国語不合格あり	1 年の外国語に関する講義に成績評価 D が存在するかどうか.{YES, NO}	○
97	外国語失格あり	1 年の外国語に講義に成績評価 X が存在するかどうか.{YES, NO}	○
98	数学不合格あり	1 年の数学に関する講義に成績評価 D が存在するかどうか.{YES, NO}	○
99	数学失格あり	1 年の数学に講義に成績評価 X が存在するかどうか.{YES, NO}	○

100	理科不合格あり	1 年の理科に関する講義に成績評価 D が存在するかどうか.{YES, NO}	○
101	理科失格あり	1 年の理科に講義に成績評価 X が存在するかどうか.{YES, NO}	○
102	専門不合格あり	1 年の専門教科の講義に成績評価 D が存在するかどうか.{YES, NO}	○
103	専門失格あり	1 年の専門教科の成績評価 X が存在するかどうか.{YES, NO}	○

さらに, データの表現方法についても考える. 現時点での定義では, 各変数は絶対値表現がなされているが, この合理性を検証しなければならない. 例えば, ある GPA 値が 2.5 を示していたとして, この値に相対的な評価を下すのは難しいだろう.GPA 平均値が 2.0 であれば,「GPA 値が 2.5」という情報は「成績が比較的優秀である」と解釈できるが,GPA 平均値が 3.0 であれば,「GPA 値が 2.5」という情報は「成績が比較的劣等である」と解釈できる. 周辺の学習状況に応じてその意味が変容するのである.

そこで, 全ての属性値を偏差値に書き換えることにより, 相対的な情報も付加することを考えた. 偏差値とは, ある数値が母集団のどの位置にいるかを示す数値のことであり, 平均値が 50, 標準偏差が 10 となるように標本変数を規格化している. 数値的表現として既に整備されているものである. ある属性値 x_i に対応する偏差値 T_i は標本変数の平均値 μ と標準偏差 σ で計算することができ, その式は以下のように与えられる.

$$T_i = \frac{10(x_i - \mu)}{\sigma} + 50 \tag{3.2}$$

変数番号 1 から 103 の変数を, 式 (3.2) により全て偏差値に表現し直し, それを定義として再現し直した. 以下の表 3.3 にそれを示す. ちなみに, 絶対評価によって定義された変数と区別するため, 偏差値によって定義された変数の番号の頭に T を付け加えている.

表 3.3: 変数の定義一覧その 3

番号	意味	入力
T(1-92)	変数番号 1-92 を全て偏差値に表し直したもの	

本研究では, 絶対値による表現でなされた変数群と偏差値により表現された変数群とで, どちらが成績予測モデルの構築に適しているかを検証している.

3.3　データの正規化, 具体的な離散化方法

ベイジアンネットワークの構築の際, 連続的数値表現である目的変数を離散的表現に書き換えなければならない. そこで目的変数の離散化方法として, 以下の 3 つを考えた.

1. N 個の区分に等量分割する

   目的変数の事例数が x 個ある場合,1 つの区分の事例数が $\frac{x}{N}$ になるように (もしくは極力そのような事例数の配分になるように) 分割する. 利点として, 手間がかからない事, 目的変数の事前確率がほぼ一定の分布になることが挙げられるが, 分割された区分が有意であるかどうかはモデル構築後の評価において初めて判断できる. 本論文ではこの離散化方法を一貫して「等量分割」と呼称する.

2. 各区分の分散を最小にするように分割する

   N 個の区分に分けたとき, それぞれの区分の分散が最小となるように分割する. 利点として, 分散という判断基準を用いることで等量分割よりも有意であると考えられる離散化が可能だということが挙げられるが, データに存在する外れ値に過敏反応し極端に事例数の少ない区分を作り出してしまう恐れがある. 本論文ではこの離散化方法を一貫して「分散最小分割」と呼称する.

また, 目的変数以外の離散化は, 上記の方法に加え, 各変数間における情報利得を最大にするような離散幅を獲得する決定木の手法を用いる.

具体例を以下に示す. 例えば, 数値型である変数 A と, 属性値として $\{ YES, NO \}$ の 2 値をとる変数 B を仮定する. そして,2 つの変数に関するデータが下の表 3.4 のようになっているとする. このとき, 変数 B に関する決定木 (分類木) を生成すると, 図 3.1 のようになる. この決定木から, 数値型変数 A の有意な離散幅 $\{(-\infty : 6], (6 : 15], (15 : 28], (28 : \infty)\}$ を取得する. この一連の処理により数値型変数の離散化を完了している.

表 3.4: 変数 A,B に関するデータ

A	3	2	15	6	7	12	3	23	21	45	28	9
B	YES	YES	NO	YES	NO	NO	YES	YES	YES	NO	YES	NO

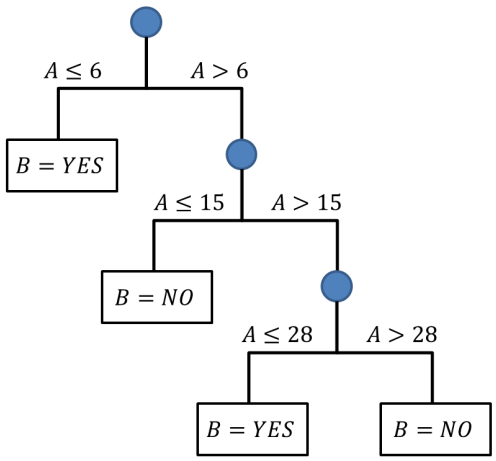

図 3.1: 変数 B の分類木

第4章　成績予測モデルの構築

成績予測による学習指導を実現するため, 拡張及び正規化を施したデータをもとに, 成績予測に適するベイジアンネットワークを構築しなければならない. まずは理想的なモデルを概論し, 本研究の目標を自己設定する. そして, どのような形での学習指導が行えるかを議論した後, それぞれの案に沿ったベイジアンネットワークモデルを構築する. 本研究では,3 つの学習指導モデルを構想し, 適したモデルを構築した後にモデル評価をすることで, 考えられた学習指導の構想が有用であるかどうかを検証した. また, モデル構築の際の学習用データとして,3 章にて説明がなされている生徒個人の成績データ群を変数としたものを利用している.

また, 構築されたベイジアンネットワークの精度評価の際に, ソフトウェア weka[13] を使用している.

4.1　予測技術を用いた適切な学習指導

本研究は, ベイジアンネットワークを用いた予測技術を学生に対する学習指導に活かしたいと目論んでいる. ここで, 適切な学習指導とは何なのかを議論しなければ, 有用性を提示することは難しい. そこで, 理想的な学習指導を仮定し, これを 1 つの目標にすることで, 有用な学習指導モデルを目指す. 主に 2 つの観点から予測技術を用いた理想的モデルを設定した.

- 高い予測精度を示すモデル

予測技術を用いた学習指導をする際, 学生側に対し, その時点において未知である情報を提供することになる. しかし, 全くの出鱈目な情報を提供しては学生側も学習指導に意義を見出せない. やはり, 提供された情報の確信度が高ければ高いほど, 学生の満足度は高くなり, 確信度の高い情報を提供することができれば, 学習環境は向上するだろう.

- 学生にとって理解しやすい情報を提供するモデル

前述した確信度の高い情報であっても, 受信する側の学生にとって意味が理解できない情報では, 学習環境の向上には繋がらない. どのような学生でも直感的に理解できるような表現がなされた学習指導ができれば, 今後の勉学に向けての対応が容易に決定できる.

本研究では上の観点を重視したモデル評価を行い, より理想的なモデルかどうかを検証している. 予測精度の評価には, 学習データをそのまま入力データとして扱う代替推定法ではなく, 学習データから 1 事例だけ抜粋し, 残りのデータでモデルを構築した後に, 抜粋したデータを入力して正しい予測ができるかどうかを, 全ての事例に適用し精度を評価する leave one out 法を採用している.leave one out 法は, データ数の少ない場合の評価方法として採用されることが多く, 本研究において適した評価方法であると言える.

また, 分かりやすい情報を提供する学習指導モデルを構築するため, 本論文では, モデル構築の前に, どのような学習指導を行い, それはどのような情報を提供しうるかを議論した. 対応するモデルを構築した後, 様々な視点から, 与える情報の有用性を検討した.

4.2　モデル 1

4.2.1　学習指導構想

学習指導を与える際,どのような学生に指導を与えるべきかを考えた. 本研究において学習指導の対象とする学生を決めなければならない. 著者はまず,1 年次から 2 年次にかけて, 成績が悪化（もしくは好転）している学生に対し, 何かしらの指導を与えることを考えた. ここで言う成績を GPA 値に置き換えれば,1 年次の GPA 値と 2 年次の GPA 値の増減を確認することで, 成績が悪化しているか, 好転しているかどうかを判断することができる. そこで 3 章での変数とは別に,(2 年次 GPA 値) – (1 年次 GPA 値) という変数を定義し, これがどのような属性値を示すかどうかを予測することにより, 学習指導内容を決定する.2 年終了時にどれだけ GPA 値が増減しているかを予測できれば, 学生は予測結果を受けて, 今後の学習態度を改めることが可能だと言える.

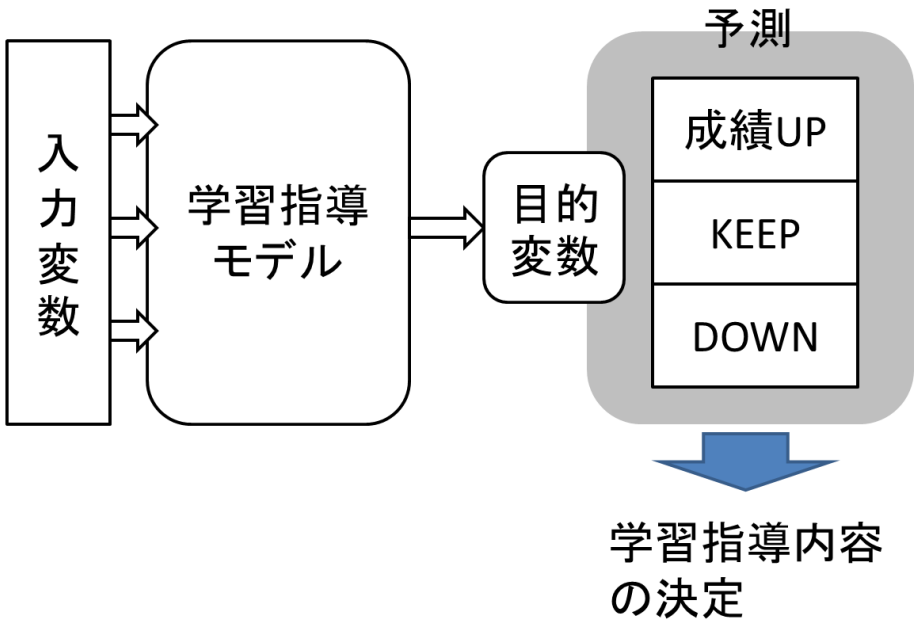

図 4.1: モデルその 1 による学習指導の構想

4.2.2　モデル構築とその精度評価

モデルの構想を練り終えたところで, 適する成績予測モデルの構築に取り掛かった. 各学生の (2 年次 GPA 値) – (1 年次 GPA 値) を計算し, この値群を番号 104 の変数として定義した. さらに, 目的変数となる変数番号 103 の変数に, 等量分割と分散最小分割を施した. 分割数は 2,3 とした. 以下の表 4.2 に離散化の結果を示す.

表 4.1: モデル 1 の構想によって新しく定義された目的変数

番号	表記	意味	入力
104	(2 年次 GPA 値) – (1 年次 GPA 値)	2 年次の GPA 値から 1 年次の GPA 値を減算したもの	

表 4.2: 目的変数の離散化結果

2 分割の場合				
離散化方法	等量分割		分散最小分割	
属性値	1	2	1	2
集計値	86	85	112	59
平均値	−0.498	0.207	−0.4	0.3
意味	成績悪化	成績好転	成績悪化	成績好転

3 分割の場合						
離散化方法	等量分割			分散最小分割		
属性値	1	2	3	1	2	3
集計値	57	57	57	31	129	11
平均値	−0.64	−0.14	0.33	−0.84	−0.06	0.69
集計値	成績悪化	維持	好転	成績悪化	維持	好転

属性値が 1 から 3 になるにつれて, 良成績であることを示している. これらを目的変数とし, 適するベイジアンネットワークのグラフ構造を学習及び決定していく.

ここでひとまず, 目的変数以外の変数集合として, 絶対値表現がなされた変数番号 1-103 を採用した. 変数の取捨選択の判断基準として, 節 2.2 にて述べた CFS を用いている. 表 4.2 の目的変数に対応するグラフ構造をそれぞれ構築し,leave one out 法による予測精度の検証をしたあと, 最も結果の良いものを各目的変数に応じたグラフ構造として選定した.

以下に, 各目的変数における最適な精度評価を得たグラフの大まかな構築工程と, その予測精度を示す. また, 構築されたグラフ構造は表中に示す図番号の各図に示している.

まず, 目的変数を 2 分割したときの最適なグラフの構築工程を示す. まず等量分割を施した場合の工程は以下の通りである.

1. 目的変数以外の変数を 3 つの区分に等量分割
2. CFS を指標として, 変数を抜粋
3. 抜粋した変数を用いて TAN を構築

また, 分散最小分割の場合は以下の通りとなる.

1. 目的変数以外の変数を 3 つの区分に分散最小分割
2. CFS を指標として, 変数を抜粋
3. 抜粋した変数を用いて Naive bayes を構築

次に, 目的変数を 3 分割したときのグラフ構築工程を示す. まず等量分割を施した場合の工程は以下の通りである.

1. 目的変数以外の変数を 2 つの区分に等量分割
2. CFS を指標として, 変数を抜粋
3. 抜粋した変数を用いて TAN を構築

また, 分散最小分割の場合は以下の通りとなる.

1. 目的変数以外の変数を 5 つの区分に分散最小分割
2. CFS を指標として, 変数を抜粋
3. 抜粋した変数を用いて Naive bayes を構築

そして, それぞれのモデルの評価を,leave one out 法による的中率の評価によって行った. 次ページの表 4.3, 表 4.4 に結果を示す.

表 4.3: 精度評価 1

等量 2 分割			
		予測出力	
		1	2
実際	1	46	40
	2	39	46
的中率		53.8%	
グラフ構造：図 4.2			

分散最小 2 分割			
		予測出力	
		1	2
実際	1	95	17
	2	32	27
的中率		71.34%	
グラフ構造：図 4.3			

表 4.4: 精度評価 2

等量 3 分割				
		予測出力		
		1	2	3
実際	1	28	15	14
	2	10	35	12
	3	15	20	22
的中率		49.7%		
グラフ構造：図 4.4				

分散最小 3 分割				
		予測出力		
		1	2	3
実際	1	4	27	0
	2	3	118	2
	3	2	4	5
的中率		74.2%		
グラフ構造：図 4.5				

次に, 偏差値による表現がなされた変数群を用いてグラフ構造を学習する. まずは目的変数である変数番号 104 の変数も偏差値に書き換え, 離散化を施した. 以下の表 4.5 に離散化結果と精度評価結果を示す. また, グラフ構造工程は, 絶対値表現の変数と偏差値表現の変数の有意性比較のため, 絶対値表現がなされた変数を用いたときと同様とする.

表 4.5: 偏差値表現の変数を用いたときの離散化結果と精度評価

2 分割の場合				
離散化方法	等量分割		分散最小分割	
属性値	1	2	1	2
集計値	86	85	112	59
平均値	42.37	57.53	44.6	60.23
意味	成績悪化	成績好転	成績悪化	成績好転

3 分割の場合						
離散化方法	等量分割			分散最小分割		
属性値	1	2	3	1	2	3
集計値	57	57	57	31	129	11
平均値	39.42	50.11	60.46	35.24	51.98	68.31
集計値	成績悪化	維持	好転	成績悪化	維持	好転

等量 2 分割			
		予測出力	
		1	2
実際	1	60	25
	2	24	62
的中率		71.34%	
グラフ構造：図 4.6			

分散最小 2 分割			
		予測出力	
		1	2
実際	1	99	13
	2	35	24
的中率		71.92%	
グラフ構造：図 4.8			

等量 3 分割				
		予測出力		
		1	2	3
実	1	28	15	14
	2	11	34	12
際	3	12	18	27
的中率		52.04%		
グラフ構造：図 4.9				

分散最小 3 分割				
		予測出力		
		1	2	3
実	1	2	29	0
	2	10	118	1
際	3	2	5	4
的中率		72.5%		
グラフ構造：図 4.10				

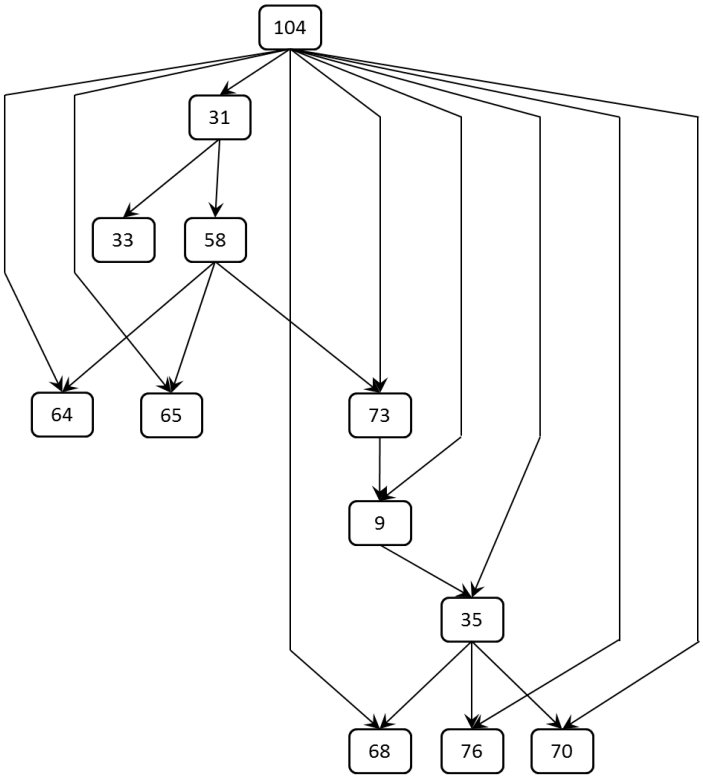

図 4.2: 目的変数 104, 等量 2 分割

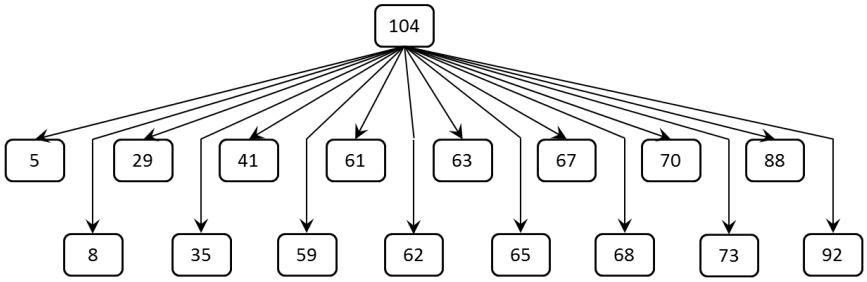

図 4.3: 目的変数 104, 分散最小 2 分割

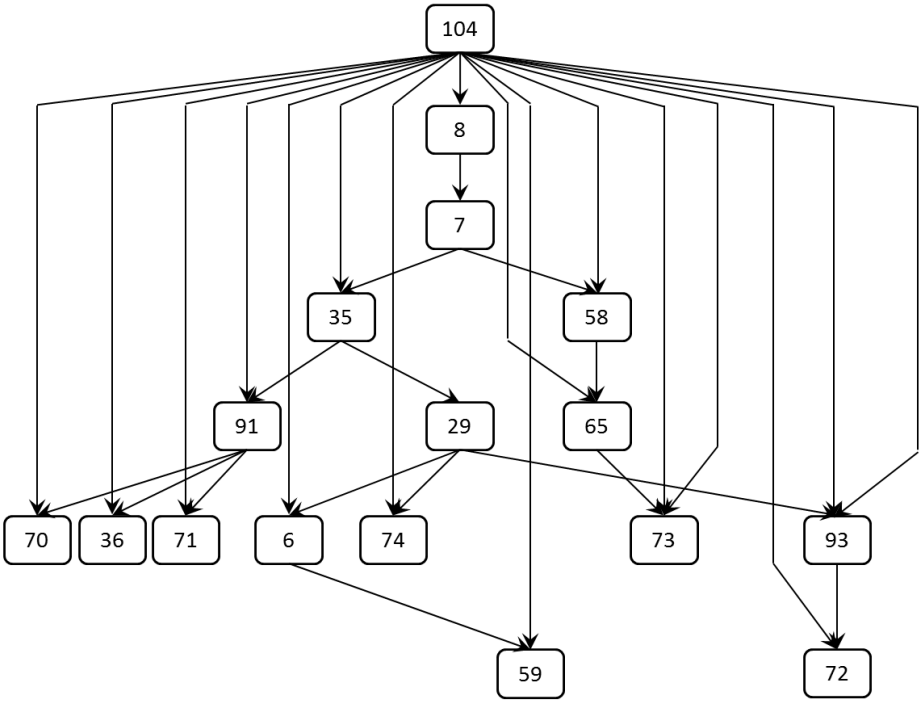

図 4.4: 目的変数 104, 等量 3 分割

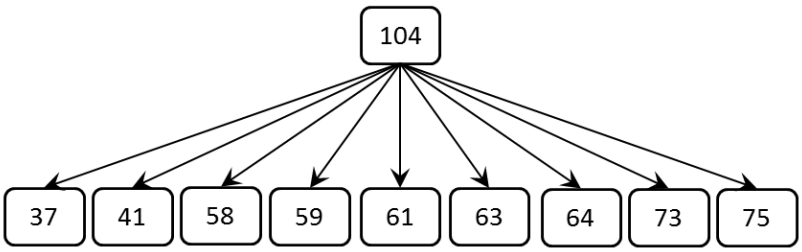

図 4.5: 目的変数 104, 分散最小 3 分割

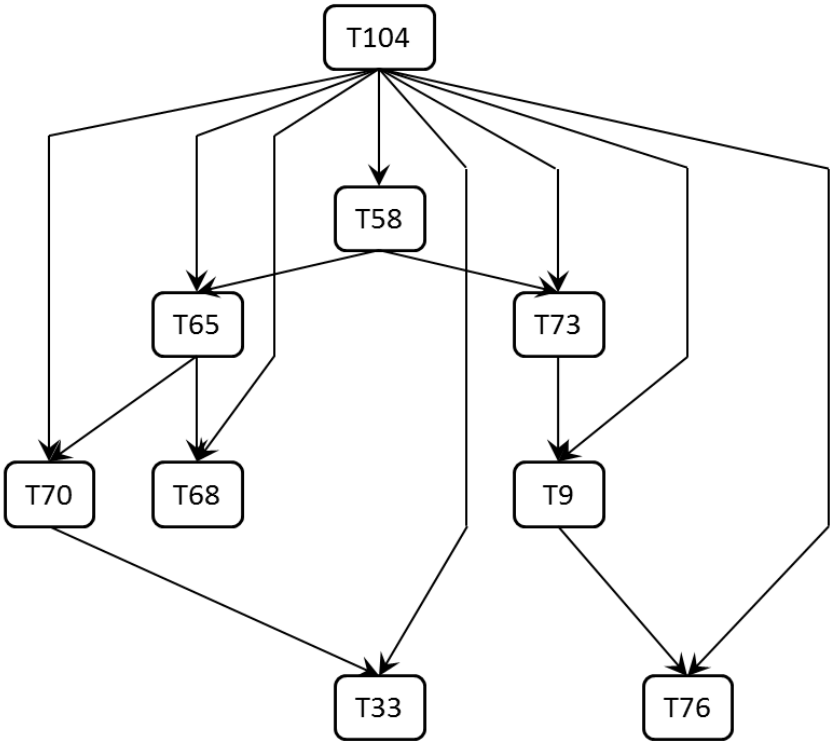

図 4.6: 目的変数 T104, 等量 2 分割

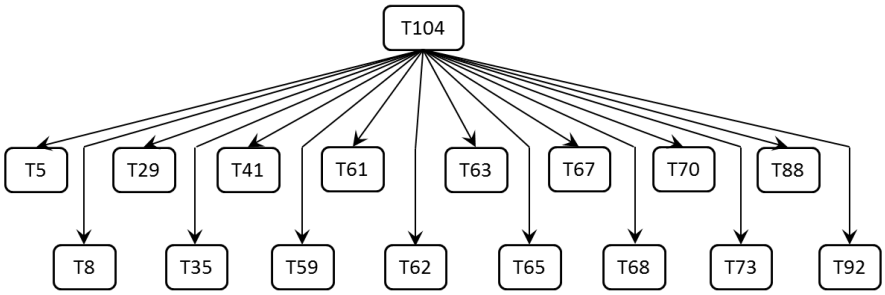

図 4.7: 目的変数 T104, 分散最小 2 分割

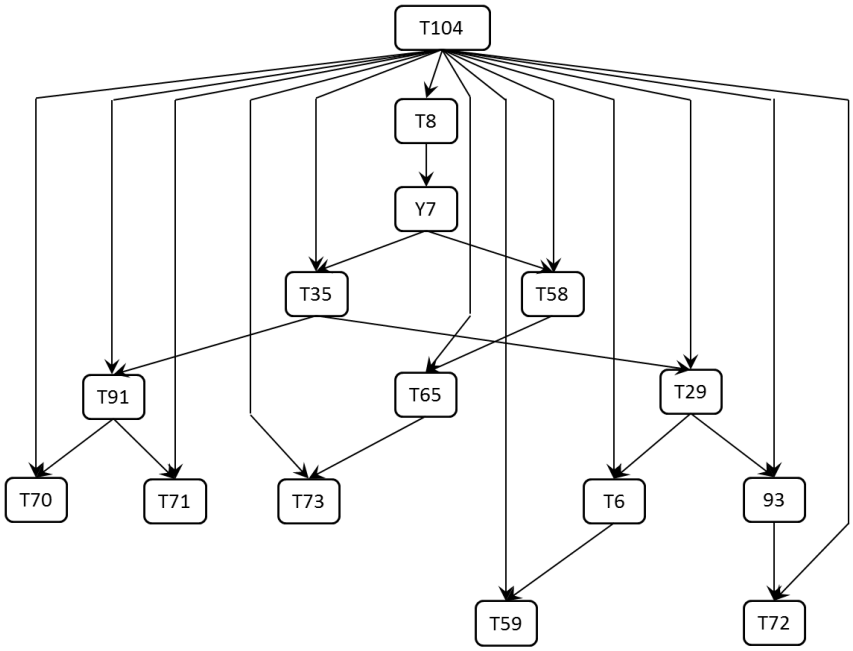

図 4.8: 目的変数 T104, 等量 3 分割

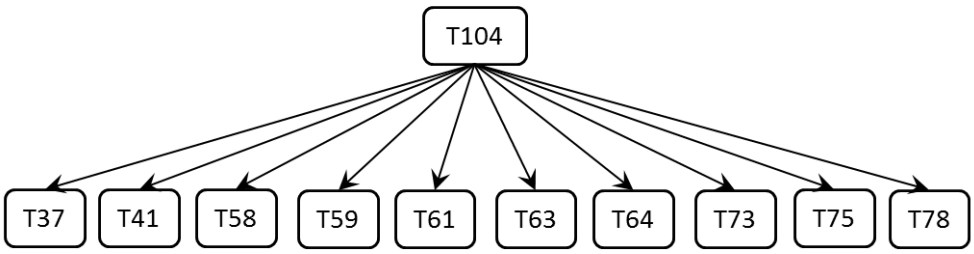

図 4.9: 目的変数 T104, 分散最小 3 分割

4.2.3　有用性の検証と考察

変数番号 104 の変数を目的変数とし, 用いるデータの表現形式, 目的変数の離散化方法に応じた最適なベイジアンネットワークのグラフ構造を得た. 予測精度だけをまとめたものを以下の表 4.6 に示す.

表 4.6: 各グラフ構造における予測精度の比較

離散化方法	等量分割				分散最小分割			
分割数	2		3		2		3	
表現形式	絶対値	偏差値	絶対値	偏差値	絶対値	偏差値	絶対値	偏差値
予測精度	53.8%	71.3%	71.3%	71.9%	49.7%	52.0%	74.2%	72.5%

目的変数を等量 2 分割する場合において, 偏差値表現の方が有意であることが分かる. また他の場合においても, 絶対値表現と偏差値表現とで精度評価が変わらないため, 表現形式を偏差値に書き換えても, 予測精度が落ちることがなく, かつ, ある条件下における最適なグラフ構造の学習の際, 偏差値表現がなされたデータを用いる方が得られたグラフ構造の予測精度が上昇する可能性があるということが, この実験から示された.

偏差値表現がなされたデータを用い, 目的変数を等量 2 分割した場合に得られたグラフ構造では, 予測精度が 70%を越えた. 出力形式は「成績が上がる」か「成績が下がる」の 2 パターンしかなく, 極めて単純な成績予測モデルである. しかし, ベイジアンネットワークの出力の表現力を応用することで, この単純性をある程度解消できる.

ベイジアンネットワークの出力形式は, ある変数の事後確率で表現される. 例えば, ベイジアンネットワークによる予測分類を利用する学生に対して, 変数番号 104 の変数の事後確率を与える. 同じ「成績が上がる」という分類判定であっても,「成績が上がる確率が 98%, 成績が下がる確率が 2%」という情報を受け取るのと,「成績が上がる確率が 56%, 成績が下がる確率が 44%」という情報を受けとるのでは, その意味合いが変わってくるだろう. このように, 予測結果に対し,「表現の含み」を持たせることができるので, より柔軟な学習指導を与えられることが期待できる. 出力例として, 付録 B の表 B.1 に, 偏差値表現のデータを用いてかつ目的変数を等量 2 分割したときの最適モデルの予測出力を記載している.

しかし, 前述の等量 2 分割の場合を別にするならば, 期待した予測精度を得られたとは言い難い. 特に, 分散最小分割を程こした場合, 予測精度は等量分割よりも高い精度を示しているが, しかし, 表 4.3, 表 4.4, 表 4.5 を確認すると, 成績を維持している学生の予測は十分行えているものの, 成績が悪化する学生と好転する学生の予測精度はあまり芳しくない. これでは, 学習指導を与えるべき対象である学生の選別ができず, 全体として有効な学習指導が行えない.

特に, この実験では「特殊な学生の検知」を実現することはできなかった. 例えば,「大学教育から脱落してしまう（またはしてしまいそうな）学生」がその 1 つであると言える. この実験で構築された成績予測モデルでは, 主に成績を維持している学生の分類予測の精度は高く, いわゆる「中間層」に分類されるかどうかの予測に優れていると言えるが, そのような学生は学習指導の対象にはなりにくい. やはり, 成績が特に悪い学生を救済したいという要望がある.

そこで, 次の学習指導のモデルでは, 学習指導の対象を予め限定しておき,「成績が芳しくない学生」の予測を試みる.

4.3　モデル 2

4.3.1　学習指導構想

モデルその 1 では成績の増減を予測することで学習指導の内容を決定していた. しかし期待以上の予測精度は得られず, 予測モデルによる学習指導の説得力に不足があるように思えた. そこで, 予測内容を再度検討し, 違ったアプローチで学習指導の内容決定を試みる.

モデル 1 の予測精度が伸び悩んだ要因として, 成績維持層以外の予測に苦しんだことが挙げられる. そこで, 学習指導を与える学生の対象をより限定し, その限定された学生を救援するような学習指導を考える. ここで, 対象とする学生を「成績が芳しくない学生」と仮定し, さらに,「成績が芳しくない学生」を「2 年後期の GPA の偏差値が低い」と意味付けする.1 年終了時に予測された 2 年後期 GPA の偏差値がある基準を下回った場合に「成績が芳しくない」と認定し, 然るべき学習指導を与える.

この章では,「成績が芳しくない学生」かどうかの判断基準となる 2 年後期 GPA の偏差値のしきい値を複数設定した後, それぞれに対応した予測モデルを構築し, 予測精度の評価をする. その結果によりモデルその 2 による学習指導構想が如何ほど有用であるかを検証する.

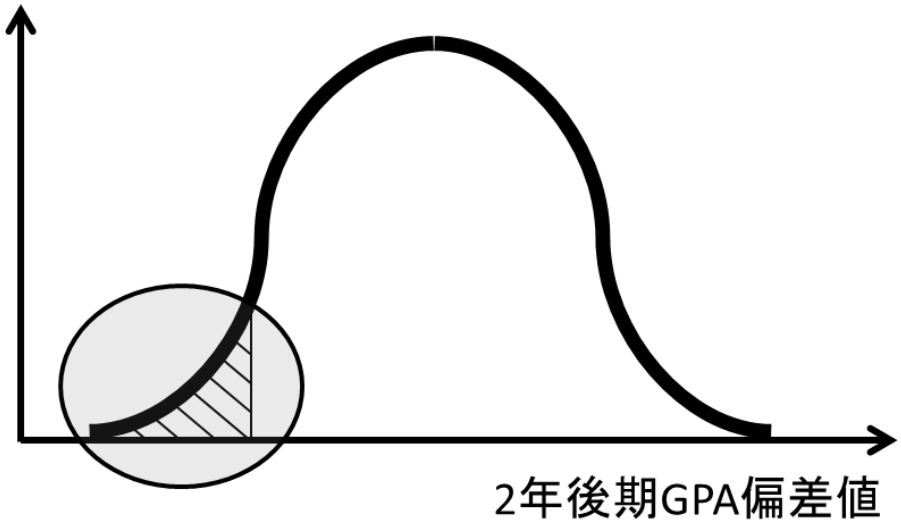

図 4.10:「成績が著しくない学生」の例

4.3.2　モデル構築とその精度評価

節 4.2 の実験結果から, モデル構築に用いるデータとして偏差値表現がなされているものを採択する. また, 変数番号 45 の「2 年後期 GPA」を偏差値変換し,「2 年後期 GPA」の偏差値が設定値を下回る学生を「成績が芳しくない学生」と認定する. ここで設定偏差値を 40 とし, この設定値に応じた予測モデルを構築する

まず, 表 4.7 にあるように目的変数として新しく変数を定義する. 属性値は YES, NO の 2 つである.

表 4.7: モデル 2 の構想によって新しく定義された目的変数

番号	意味	入力
105	2 年後期 GPA の偏差値が 40 を下回っている. $\{YES, NO\}$	

また, モデルの構築工程は以下のようになった.

1. CFS を指標として, 変数を抜粋
2. 抜粋した変数をもとに決定木 (分類木) を生成
3. 出力された決定木を参考に, 数値型変数を離散化
4. TAN を構築

得られた決定木は図 4.12 に示している. この決定木は leave one out 法による精度評価で 80.7%, 学習データをそのままテストデータとして用いる精度評価で 97%の結果を得た. 例えば, 変数番号 $T12$ は下の決定木により「40.8 より大きい」「40.8 以下」という離散値に分割がなされた. この工程を経て得られたグラフの精度評価を行った. 以下の表 4.8 にその結果を示す.

表 4.8: 偏差値表現の変数を用いたときの離散化結果と精度評価

目的変数 105			
		予測出力	
		YES	NO
実際	YES	18	7
	NO	4	142
的中率		93.56%	
グラフ構造：図 4.11			

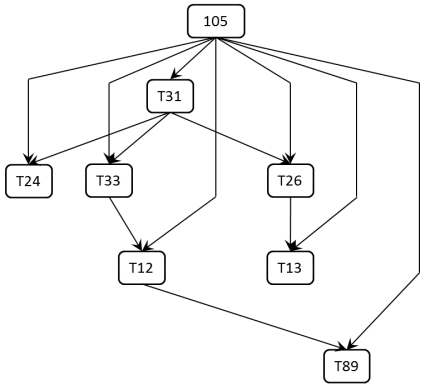

図 4.11: 目的変数 105,「成績が芳しくない生徒」の予測モデル

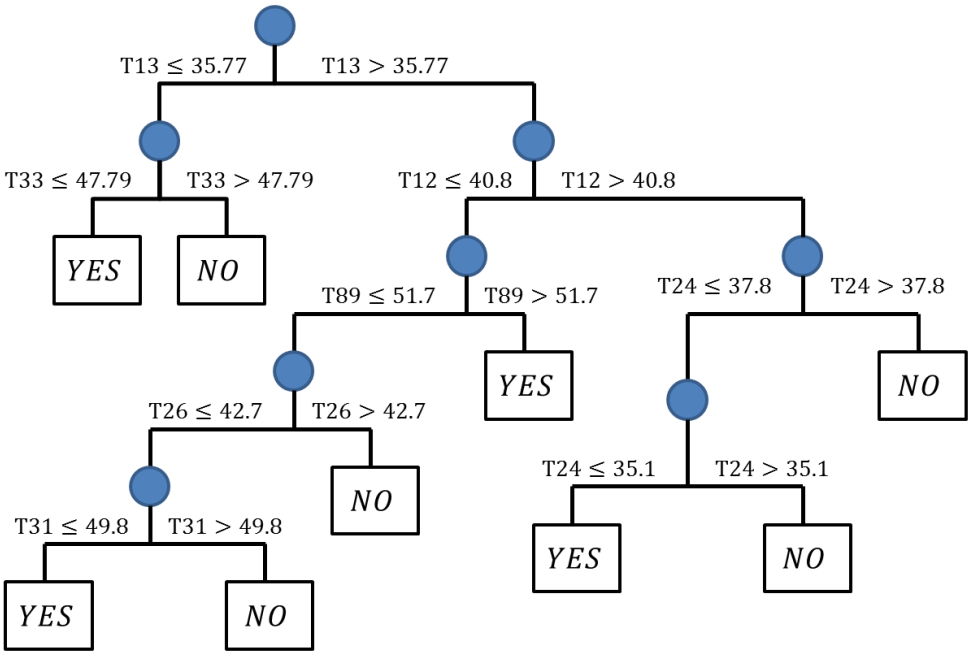

図 4.12: 生成された決定木

4.3.3　有用性の検証と考察

2 年後期の偏差値が 40 を下回るかどうかを予測した. モデル全体の予測精度としては 93.56%と非常に高く, また, 学習指導対象となりうる学生の予測 (「YES」に該当する学生) の予測については, 実際に「YES」に該当する 25 人中,18 人の学生に対し「YES」の予測をすることができた. つまり,「YES」の的中率は $\frac{18}{25} = 72\%$ となった. さらに, ベイジアンネットワークによる事後確率の出力結果を確認すると, 実際「YES」であるのに「NO」と予測してしまった場合でも, それに該当する 8 人の内の 4 人に 30%以上の「YES」事後確率を与えられている. 予測自体は間違っていても, 事後確率による表現により, 低成績の傾向を示唆するような指導が可能になる. 事後確率の出力結果は, 付録 B の表 B.2 に掲載しているので, よろしければ参照していただきたい.

これらのことから,「YES」と判定された 18 人と, 事後確率による表現でフォローされた 4 人を含めると,$\frac{22}{25} = 88\%$ の「成績が芳しくない学生」に対し, 然るべき学習指導が与えられると言える.

次に, 予測した内容によって, どのような学習指導を行えばいいかを考える. 今回, 予測した事象は「2 年後期の GPA が偏差値 40 を下回る」かどうかである. しかし, ある学生がどのような修学傾向により,2 年次の成績が低迷してしまったかは, 予測の結果だけでは分からない.

そこで,「成績が芳しくない学生」の傾向を予め想定しておき, 予測によって得られた結果とその時点で既知である情報をもとに, その「成績が芳しくない学生」がどのような傾向で成績が低迷しているかを推定し, それに対応するように学習指導内容を決定することを考える. まず,「成績が芳しくない学生」の傾向について以下の 2 つを挙げた.

(1) 1 年から 2 年にかけて, 成績が急激に悪化した

(2) 1 年次から既に成績が低迷していた

このような想定から, 傾向を推定するにおいて必要な情報は,「1 年次の GPA の偏差値」であると結論付けた. つまり,「成績が芳しくない学生」の内,「1 年次の GPA 偏差値」が基準値である 40 を上回っている学生は,(1) の傾向が当てはめられるので,「成績がさらに悪化する恐れ」を示唆するような学習指導を行えばいい. 対して,「1 年次の GPA 偏差値」が基準値である 40 を下回っている学生には,(2) の傾向が当てはめられるので,「成績が常に全体の下位に位置している」ことを警告し, 学習態度の改善を訴えかけるようなメッセージを送ればいい.

このように, 予測結果と既知の情報を複合的に用いることで多様な学習指導が可能であるといえる.

表 4.9: ベイジアンネットワークの出力例

YES	NO
34.6%	65.4%

4.4　モデル 3

4.4.1　学習指導構想

前述のモデルでは,1 年の成績から 2 年の成績を予測をするというのが大まかな流れであったが, ここで 1 つ加味しなければならないのが,1 年次から 2 年次にかけて講義の様相が変わるということである. 例えば 1 年次には線形代数, 微分積分, 力学などのような理系にとって基礎的な講義や, 専門教科に分類される講義でもその内実はオリエンテーションのような単位取得には易い講義を受講することが多い. 対して 2 年次には, 専門的に深く掘り下げた講義が増え, 高校学習の延長線上とは言い表せないような学習に迫られる.

そこで, 今度はある種類の教科を予測の対象として, 大学教育の専門性の強い講義から脱落しそうな学生を救済することを考える. 特に脱落者を多数生むと考えられる 2 年次の専門教科に関する予測を試行する. ここで,「2 年次の専門教科に苦しむ学生」を仮定すると, この集団が示す特徴として以下のものが挙げられる.

(1) 他の教科の GPA よりも専門教科の GPA の方が低い

(2) 専門教科の GPA がある基準を下回っている

言い換えれば, この傾向が見て取れる学生を予測できれば, 学習指導を与えるべき学生を選定できる. そして予測結果に応じて学習指導内容を変えれば, その成績予測モデルは学生にとって有用なものになりうる.

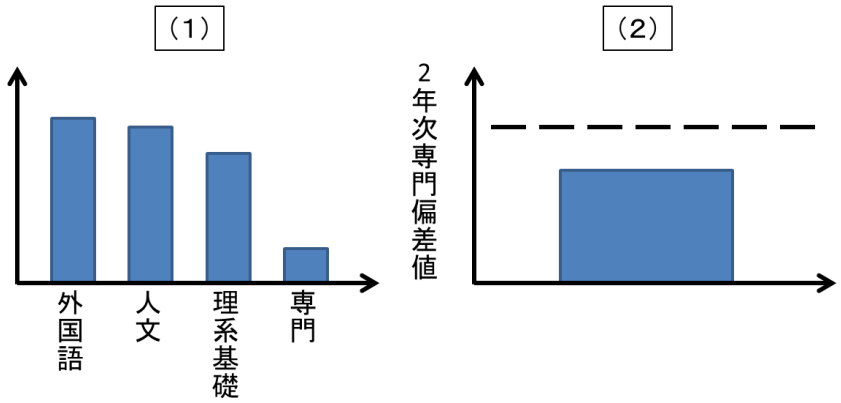

図 4.13:「2 年次の専門教科に苦しむ学生」の例

4.4.2　モデル構築とその精度評価

前述した「2 年次の専門教科に苦しむ学生」の特徴を, より具体的に表現すると, 以下のように考えることができる.

1. 変数番号 T-55(「2 年専門 GPA」の偏差値) が基準値を下回る

2. 変数番号 T-87(「2 年専門 GPA−2 年 GPA」の偏差値) が基準値を下回る

そこで, 基準値を $\{50, 45, 40\}$ とし, 上記の条件を参考にして, 変数を新たに定義した. 以下の表 4.10 にそれらを示す. そして, これらを目的変数とした予測モデルを構築することで,「2 年次の専門教科に苦しむ学生」を推定し然るべき学習指導を与えることを考える. モデル構築の後,leave one out 法による予測精度の評価を行った. 表 4.11 にその結果を示す. また, 基準値によって予測精度や学習指導の有意性が変化するかを調査し, 考察を与えた.

それぞれのモデル構築工程は, 全てにおいて以下の通りであった.

1. CFS を指標として, 変数を抜粋
2. 抜粋した変数をもとに決定木 (分類木) を生成
3. 生成された決定木を参考に, 数値型変数を離散化
4. TAN を構築

表 4.10: モデル 3 の構想によって新しく定義された目的変数

番号	意味	入力
106	変数番号 T-55 と変数番号 T-87 が共に 50 を下回っている. $\{YES, NO\}$	
107	変数番号 T-55 と変数番号 T-87 が共に 45 を下回っている. $\{YES, NO\}$	
108	変数番号 T-55 と変数番号 T-87 が共に 40 を下回っている. $\{YES, NO\}$	

表 4.11: 基準値が 50 である場合の精度評価

目的変数 106			
		予測出力	
		YES	NO
実際	YES	34	26
	NO	15	96
的中率		76.02%	
グラフ構造：図 4.14			

目的変数 107			
		予測出力	
		YES	NO
実際	YES	18	12
	NO	9	132
的中率		87.71%	
グラフ構造：図 4.15			

目的変数 108			
		予測出力	
		YES	NO
実際	YES	0	5
	NO	2	164
的中率		95.9%	
グラフ構造：図 4.16			

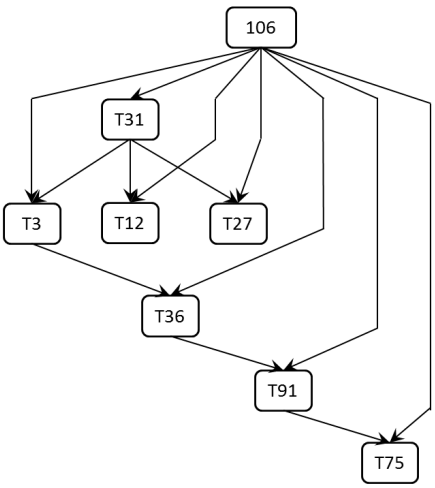

図 4.14: 目的変数 106 の予測モデル

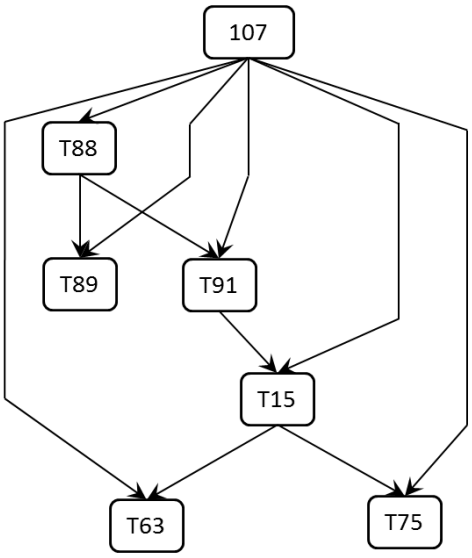

図 4.15: 目的変数 107 の予測モデル

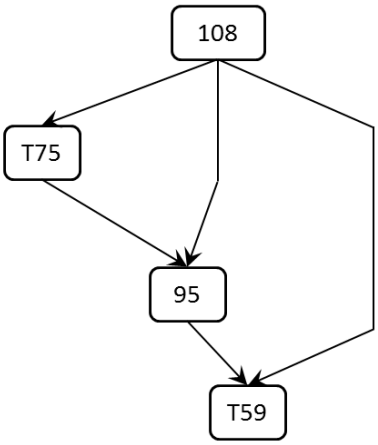

図 4.16: 目的変数 108 の予測モデル

4.4.3　有用性の検証と考察

2 年専門 GPA の偏差値と (2 年専門 GPA) − (2 年 GPA) の偏差値がともに基準値を下回っているかどうかを予測した. 今回の実験では $\{50, 45, 40\}$ の 3 つの数値を基準値として設けた. 全体の精度としては基準値が 40 の予測モデルが一番高いが,「YES」における的中率は 0%であり, 予測モデルとしては機能しているとは言えない. 原因として,「YES」に該当する事例数が 171 事例中 5 つと極端に少なく, 事象を確率的に扱うベイジアンネットワークでは拾い上げることのできない集合であったことが挙げられる.

対して, 基準値が 50,45 である予測モデルではともに 75%を越える予測精度を誇っている. また,「YES」の的中率で見ると, 基準値が 50 の場合は $\frac{34}{60} = 57\%$ であり, 基準値が 45 の場合は $\frac{18}{30} = 60\%$ であった. さらに, 節 4.3 でも述べたようにベイジアンネットワークによる確率値での表現により, 誤測に対してある程度のフォローが可能である. 例えば, 基準値が 50 である予測モデルで, 実際は「YES」であるのに「NO」と予測された学生が 26 人いるが, そのうちの 14 人に,40%を越える「YES」の確率値を与えている. これを加味すると,「YES」とちゃんと予測された学生を含めると,$\frac{34+14}{60} = 80\%$ の学生を学習指導の対象として扱える. 基準値が 50 である予測モデルの出力結果を, 付録 B の表 B.3 に掲載した.

この予測モデルにおける学習指導の内容として, この予測で「YES」と予測された学生に対し,1 年次と 2 年次とでは講義の難易度やその種類が全く異なるということをしっかりと警告し,1 年次に受講した基礎科目の復習を促すようなメッセージが相応しいと考えられる.

4.5　総括

本研究において, 複数の学習指導を構想し, それに対応したベイジアンネットワークを計 12 構築した. この実験の考察を述べ, 次への課題点を指摘する.

(1) 学習データの表現形式

3 章にて複数の変数を絶対値表現と偏差値表現とで定義した. 節 4.2 にて表現形式別のデータを用いてベイジアンネットワークを構築し, 予測精度の比較を行った. データの表現形式により, 予測精度が変わることが確認でき, 実験の結果として, 偏差値表現の方が学習データとして有意であるという結論を得ることができた. ただし, 統計的表現は偏差値の他にも数多く存在するため, 更なる比較検証が必要である.

(2) 数値型の変数の離散化

用いたデータの変数はほとんどが数値型であった. ベイジアンネットワークは数値型の変数は取り扱えないため, 離散化する必要がある. しかし, 数値型のデータを離散化すると, そのものが持っている情報量は著しく減少し, 有意性が損なわれてしまう. 本研究では変数の離散化方法として,3 章で述べた「等量分割」「分散最小分割」「生成された決定木を参考にした分割」を採用していた. この 3 つの中で特に有効であると考えられたのは生成された決定木を参考にした離散化であった. しかし, 目的変数の離散化では決定木の方法を適用できないため, その離散化の方法はユーザの判断が必要である. より有効な離散化方法を今後も模索する必要があると言える.

(3) 目的変数の選択

目的変数をもとに, ベイジアンネットワークは構築される. 構築されたベイジアンネットワークの精度は, 当然ではあるが目的変数に左右される. 本研究において, 例えば, 節 4.4 の基準値を 40 にした場合での予測モデルでは,「YES」の的中率が 0%となってしまった. これは「YES」の事例数が全事例数 171 に対して,5 つしかなかったことが原因ではないかと考えられた. ベイジアンネットワークは 2 章で述べたように, 確率変数, 有向モデル構造, そして条件付き確率で定義される. 条件付き確率は, データから推定されるため, 事例数が極端に少ないと思うような確率値が定義できず, 結果的に期待はずれの予測精度を得ることになってしまう. 該当する事例数の少ない場合の予測には, 他の予測手法を適用するのがいいと言える. 換言すれば, 目的変数の事前確率がほぼ一様である場合の予測では, ベイジアンネットワークは有効であると考えられる.

(4) 予測精度

本研究では, 構築されたモデルの予測精度, すなわち的中率をより重視した. 著者の主観評価では, 全般的に満足のいく精度評価を得ることができたと言えるが, やはり今回構築されたモデルよりも優れた予測精度を誇るモデルの構築を目指したいものである. そのためには,1. データの精錬 2. グラフ構造の学習 3. 条件付き確率の推定, この 3 つの要素をさらに追求する必要がある. 特に, 今回の研究に用いたデータは一般的には少量なものであったため, 条件付き確率の推定に難があったと考えられる. 条件付き確率の推定について, 様々な研究報告がなされているため, それらを参照し次のベイジアンネットワーク構築の参考にしたい.

(5) 学習指導モデルの実装案

将来的には, ベイジアンネットワークによる予測を用いた学習指導を現実のものとすることを考えたい. ベイジアンネットワークの運用コストは高くないため, 複数モデルの併用は難しくない. 本研究で構築したような予測モデルを複数適用することで, 多様な学習指導が可能となる. 以下の図 4.17 は学習指導モデルの構想を図示化したものである. 例えば章 4.2, 章 4.3, 章 4.4 それぞれの予測モデルを併用することで,「成績 (GPA 値) の傾向」「学習脱落可能性」「専門教科への適正」が予測されるため, 学習指導は多様化されると考えられる.

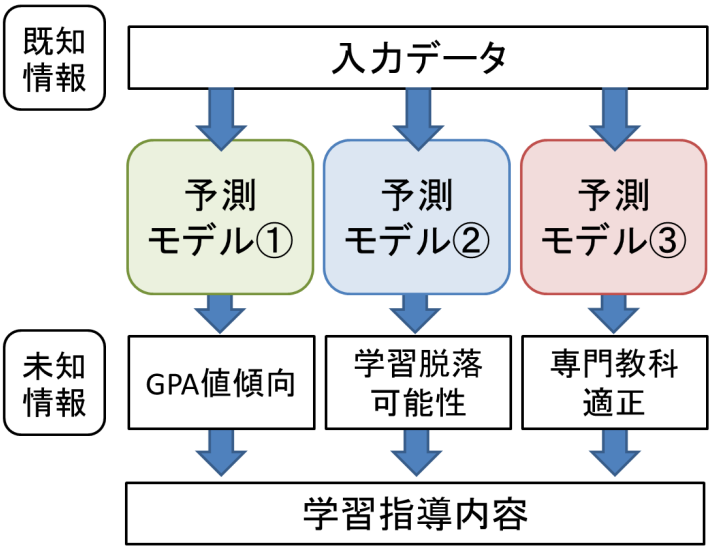

図 4.17: 学習指導のための実装モデル

第5章　むすび

本研究では, ベイジアンネットワークによる予測技術を用いて, 学生に対し然るべき学習指導を与えることを考えた.2 章では本研究で用いた理論を概説した.3 章では元となるデータの拡張と正規化について述べ,2 章にて取り上げた手法を用いた数値型変数の離散化方法について解説した.4 章では複数の学習指導案から, それに対応するベイジアンネットワークを構築した.

構築されたベイジアンネットワークを leave one out 法により評価した結果, ベイジアンネットワークの出力の確率値表現により, 節 4.3 における学習指導モデルでは学習から脱落しそうな学生の 88%, 節 4.4 における学習指導モデルでは 2 年次から開始される専門教科に躓きそうな学生の 80%に対し, 然るべき学習指導が与えられることが分かった. これにより, ベイジアンネットワークによる予測は, 学習指導の対象となる学生の選定において有用であることが示された. また, ベイジアンネットワークの構築を通して, データ形式や変数の取捨選択が予測精度に影響があることを確認できた.

ただし, ある条件下における予測モデルは, 期待していた程の予測精度を示さなかった. これは目的変数の選択の問題であったり, データそのものの問題であるかもしれない. 現時点では問題点の究明は難しいため, ベイジアンネットワークやデータ処理に関する更なる理解が必要であると言える.

また, それぞれの予測モデルは, 別年度のデータを考慮していないため, 今後は本研究の検証を含め考慮していきたい.

謝辞

本研究を進めるにあたり, 日頃から多大な御尽力を頂き, 御指導を賜わりました名古屋工業大学舟橋健司　准教授, 伊藤宏隆　助教, 山本大介　助教に深く感謝申し上げます.

また, 本研究の実験のためのデータの提供元である, 出欠システム及びコースマネージメントシステムの開発に尽力されました, 名古屋工業大学情報基盤センター長　松尾啓志　教授, 内匠逸教授, 情報基盤センター教職員の皆様に心から感謝いたします.

そして, 本研究に対し御討論, 御強力いただきました中村研究室の皆様ならびに中部大学岩堀研究室の皆様に深く感謝いたします.

最後に, 舟橋研究室のゼミにおいて舟橋研究室諸氏に多大な助言をいただきました. この場でお礼申し上げます.

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）

[2] 伊藤宏隆, 舟橋健司, 内匠逸, 松尾啓志,”IC カード出欠データと CMS 学習データを用いたデータマイニング”, 日本 e-Learning 学会誌,9 巻,pp95-108（2009）

[3] 原圭司, 高橋健一, 上田祐彰,”ベイジアンネットワークを用いた授業アンケートからの学生行動モデルの構築と考察”, 情報処理学会論文誌, 情報処理学会論文誌 51 巻, 4 号, 1215-1226 (2010)

[4] 伊藤暁人, 舟橋健司, 伊藤宏隆, ”ニューラルネットワークによる学生の成績予測とその学習指導への適用可能性の検討”, 平成 22 年度名古屋工業大学卒業研究論文, (2010)

[5] 鈴木恒一, ”大量データから知識を抽出するベイジアンネットワークの学習技術とその応用”, Softechs, L3765A, 32 巻, 1 号 ,14-17 (2011)

[6] 本村陽一, ”ベイジアンネットワーク:入門からヒューマ ンモデリングへの応用まで”, 行動計量学会セミナー資料 (2004)

[7] 鈴木護, ”ベイジアンネットワーク入門”, 培風館 (2009)

[8] 田中和之, ”ベイジアンネットワークの統計的推論の数理”, コロナ社 (2009)

[9] 石井一夫, ”図解よくわかるデータマイニング”, 日刊工業新聞社 (2004)

[10] 元田浩, 山口高平, 津本周作, 沼尾正行, ”データマイニングの基礎”, オーム社 (2006)

[11] http://www.hugin.com/

[12] Duan, S. and Babu, S, ”Processing Forecasting Queries”, *Proc. 2007 Intl, Conf. on Very Large Data Bases*, pp711-722 (2007)

[13] http://www.cs.waikato.ac.nz/ml/weka/

付 録A　ベイジアンネットワークの出力結果

この付録では, ベイジアンネットワークの出力結果を掲載している. 表の詳細を述べると, 一番左の欄から,「No」は番号,「実属性」は実際の属性値,「予測」はベイジアンネットワークが出力した予測結果,「的中」は実際の属性値と予測結果が一致しているかどうか (しているならば「○」が打たれている),「フォロー」については後述,「属性値 [・]」は出力された事後確率である. また, 出力確率値には, 予測出力として採択された属性値の値に対し,「*」のマークを追記している.

「フォロー」について,「○」があるものは, 誤予測であるが実属性の事後確率が 40%を越えているもの,「△」があるものは, 誤予測であるが実属性の事後確率が 30%を越えているものを意味している.

表 A.1: 目的変数 T104 を等量 2 分割した場合の予測出力結果

No	実属性	予測	的中	フォロー	属性値 [1]	属性値 [2]
1	1	1	○		*0.944	0.056
2	1	1	○		*0.676	0.324
3	1	1	○		*0.76	0.24
4	1	1	○		*0.844	0.156
5	1	1	○		*0.734	0.266
6	1	1	○		*0.814	0.186
7	1	1	○		*0.882	0.118
8	1	1	○		*0.773	0.227
9	1	1	○		*0.955	0.045
10	1	1	○		*0.936	0.064
11	1	1	○		*0.953	0.047
12	1	1	○		*0.942	0.058
13	1	1	○		*0.881	0.119
14	1	1	○		*0.523	0.477
15	1	1	○		*0.613	0.387
16	1	1	○		*0.908	0.092
17	1	1	○		*0.861	0.139
18	1	1	○		*0.986	0.014
19	1	1	○		*0.702	0.298
20	1	1	○		*0.743	0.257
21	1	1	○		*0.939	0.061
22	1	1	○		*0.804	0.196
23	1	1	○		*0.664	0.336
24	1	1	○		*0.892	0.108
25	1	1	○		*0.833	0.167

26	1	1	○		*0.709	0.291
27	1	1	○		*0.75	0.25
28	1	1	○		*0.85	0.15
29	1	1	○		*0.947	0.053
30	1	1	○		*0.86	0.14
31	1	1	○		*0.943	0.057
32	1	1	○		*0.946	0.054
33	1	1	○		*0.874	0.126
34	1	1	○		*0.824	0.176
35	1	1	○		*0.98	0.02
36	1	1	○		*0.822	0.178
37	1	1	○		*0.689	0.311
38	1	1	○		*0.997	0.003
39	1	1	○		*0.934	0.066
40	1	1	○		*0.84	0.16
41	1	1	○		*0.751	0.249
42	1	1	○		*0.627	0.373
43	1	1	○		*0.966	0.034
44	1	1	○		*0.807	0.193
45	1	1	○		*0.558	0.442
46	1	1	○		*0.622	0.378
47	1	1	○		*0.961	0.039
48	1	1	○		*0.806	0.194
49	1	1	○		*0.952	0.048
50	1	1	○		*0.575	0.425
51	1	1	○		*0.956	0.044
52	1	1	○		*0.987	0.013
53	1	1	○		*0.763	0.237
54	1	1	○		*0.848	0.152
55	1	1	○		*0.943	0.057
56	1	1	○		*0.698	0.302
57	1	1	○		*0.796	0.204
58	1	1	○		*0.941	0.059
59	1	1	○		*0.68	0.32
60	1	1	○		*0.592	0.408
61	1	2		○	0.45	*0.55
62	1	2		○	0.404	*0.596
63	1	2		○	0.495	*0.505
64	1	2		○	0.407	*0.593
65	1	2		○	0.48	*0.52
66	1	2		○	0.409	*0.591
67	1	2		△	0.346	*0.654

68	1	2		△	0.396	*0.604
69	1	2		△	0.338	*0.662
70	1	2		△	0.31	*0.69
71	1	2		△	0.377	*0.623
72	1	2		△	0.377	*0.623
73	1	2		△	0.319	*0.681
74	1	2		△	0.364	*0.636
75	1	2		△	0.377	*0.623
76	1	2			0.149	*0.851
77	1	2			0.148	*0.852
78	1	2			0.137	*0.863
79	1	2			0.26	*0.74
80	1	2			0.162	*0.838
81	1	2			0.143	*0.857
82	1	2			0.269	*0.731
83	1	2			0.273	*0.727
84	1	2			0.242	*0.758
85	1	2			0.291	*0.709
86	2	2	○		0.018	*0.982
87	2	2	○		0.048	*0.952
88	2	2	○		0.044	*0.956
89	2	2	○		0.198	*0.802
90	2	2	○		0.024	*0.976
91	2	2	○		0.362	*0.638
92	2	2	○		0.138	*0.862
93	2	2	○		0.096	*0.904
94	2	2	○		0.232	*0.768
95	2	2	○		0.127	*0.873
96	2	2	○		0.432	*0.568
97	2	2	○		0.149	*0.851
98	2	2	○		0.017	*0.983
99	2	2	○		0.058	*0.942
100	2	2	○		0.106	*0.894
101	2	2	○		0.062	*0.938
102	2	2	○		0.066	*0.934
103	2	2	○		0.402	*0.598
104	2	2	○		0.441	*0.559
105	2	2	○		0.087	*0.913
106	2	2	○		0.327	*0.673
107	2	2	○		0.377	*0.623
108	2	2	○		0.377	*0.623
109	2	2	○		0.165	*0.835

110	2	2	○		0.196	*0.804
111	2	2	○		0.102	*0.898
112	2	2	○		0.028	*0.972
113	2	2	○		0.057	*0.943
114	2	2	○		0.236	*0.764
115	2	2	○		0.255	*0.745
116	2	2	○		0.143	*0.857
117	2	2	○		0.242	*0.758
118	2	2	○		0.109	*0.891
119	2	2	○		0.369	*0.631
120	2	2	○		0.196	*0.804
121	2	2	○		0.151	*0.849
122	2	2	○		0.334	*0.666
123	2	2	○		0.14	*0.86
124	2	2	○		0.334	*0.666
125	2	2	○		0.096	*0.904
126	2	2	○		0.024	*0.976
127	2	2	○		0.05	*0.95
128	2	2	○		0.012	*0.988
129	2	2	○		0.15	*0.85
130	2	2	○		0.174	*0.826
131	2	2	○		0.022	*0.978
132	2	2	○		0.24	*0.76
133	2	2	○		0.143	*0.857
134	2	2	○		0.383	*0.617
135	2	2	○		0.053	*0.947
136	2	2	○		0.234	*0.766
137	2	2	○		0.235	*0.765
138	2	2	○		0.235	*0.765
139	2	2	○		0.152	*0.848
140	2	2	○		0.127	*0.873
141	2	2	○		0.053	*0.947
142	2	2	○		0.182	*0.818
143	2	2	○		0.113	*0.887
144	2	2	○		0.14	*0.86
145	2	2	○		0.462	*0.538
146	2	2	○		0.251	*0.749
147	2	2	○		0.06	*0.94
148	2	1		○	*0.576	0.424
149	2	1		○	*0.554	0.446
150	2	1		○	*0.523	0.477
151	2	1		○	*0.546	0.454

152	2	1		○	*0.598	0.402
153	2	1		○	*0.579	0.421
154	2	1		○	*0.544	0.456
155	2	1		△	*0.623	0.377
156	2	1		△	*0.684	0.316
157	2	1		△	*0.684	0.316
158	2	1		△	*0.619	0.381
159	2	1		△	*0.656	0.344
160	2	1		△	*0.66	0.34
161	2	1			*0.744	0.256
162	2	1			*0.893	0.107
163	2	1			*0.727	0.273
164	2	1			*0.785	0.215
165	2	1			*0.76	0.24
166	2	1			*0.711	0.289
167	2	1			*0.773	0.227
168	2	1			*0.792	0.208
169	2	1			*0.903	0.097
170	2	1			*0.886	0.114
171	2	1			*0.821	0.179

表 A.2: 目的変数 105 の予測モデルの出力結果

No	実属性	予測	的中	フォロー	属性値 [YES]	属性値 [NO]
1	YES	YES	○		*0.551	0.449
2	YES	YES	○		*0.578	0.422
3	YES	YES	○		*0.827	0.173
4	YES	YES	○		*0.895	0.105
5	YES	YES	○		*0.968	0.032
6	YES	YES	○		*0.98	0.02
7	YES	YES	○		*0.981	0.019
8	YES	YES	○		*0.994	0.006
9	YES	YES	○		*0.999	0.001
10	YES	YES	○		*0.999	0.001
11	YES	YES	○		*0.999	0.001
12	YES	YES	○		*0.999	0.001
13	YES	YES	○		*0.999	0.001
14	YES	YES	○		*0.999	0.001
15	YES	YES	○		*0.999	0.001
16	YES	YES	○		*0.999	0.001
17	YES	YES	○		*1	0
18	YES	YES	○		*1	0

19	YES	NO		○	0.437	*0.563
20	YES	NO		△	0.332	*0.668
21	YES	NO		△	0.332	*0.668
22	YES	NO		△	0.332	*0.668
23	YES	NO			0.005	*0.995
24	YES	NO			0.242	*0.758
25	YES	NO			0	*1
30	NO	NO	○		0.001	*0.999
31	NO	NO	○		0.001	*0.999
32	NO	NO	○		0.001	*0.999
33	NO	NO	○		0.001	*0.999
34	NO	NO	○		0.001	*0.999
35	NO	NO	○		0.001	*0.999
36	NO	NO	○		0.001	*0.999
37	NO	NO	○		0.001	*0.999
38	NO	NO	○		0.001	*0.999
39	NO	NO	○		0.001	*0.999
40	NO	NO	○		0.001	*0.999
41	NO	NO	○		0.001	*0.999
42	NO	NO	○		0.001	*0.999
43	NO	NO	○		0.001	*0.999
44	NO	NO	○		0.001	*0.999
45	NO	NO	○		0.001	*0.999
46	NO	NO	○		0.001	*0.999
47	NO	NO	○		0.001	*0.999
48	NO	NO	○		0.001	*0.999
49	NO	NO	○		0.001	*0.999
50	NO	NO	○		0.001	*0.999
51	NO	NO	○		0.001	*0.999
52	NO	NO	○		0.001	*0.999
53	NO	NO	○		0.001	*0.999
54	NO	NO	○		0.001	*0.999
55	NO	NO	○		0.001	*0.999
56	NO	NO	○		0.001	*0.999
57	NO	NO	○		0.001	*0.999
58	NO	NO	○		0.001	*0.999
59	NO	NO	○		0.001	*0.999
60	NO	NO	○		0.004	*0.996
61	NO	NO	○		0.004	*0.996
62	NO	NO	○		0.004	*0.996
63	NO	NO	○		0.004	*0.996
64	NO	NO	○		0.004	*0.996

65	NO	NO	○		0.004	*0.996
66	NO	NO	○		0.004	*0.996
67	NO	NO	○		0.004	*0.996
68	NO	NO	○		0.004	*0.996
69	NO	NO	○		0.004	*0.996
70	NO	NO	○		0.004	*0.996
71	NO	NO	○		0.004	*0.996
72	NO	NO	○		0.004	*0.996
73	NO	NO	○		0.004	*0.996
74	NO	NO	○		0.004	*0.996
75	NO	NO	○		0.004	*0.996
76	NO	NO	○		0.004	*0.996
77	NO	NO	○		0.004	*0.996
78	NO	NO	○		0.004	*0.996
79	NO	NO	○		0.004	*0.996
80	NO	NO	○		0.005	*0.995
81	NO	NO	○		0.005	*0.995
82	NO	NO	○		0.005	*0.995
83	NO	NO	○		0.005	*0.995
84	NO	NO	○		0.005	*0.995
85	NO	NO	○		0.005	*0.995
86	NO	NO	○		0.005	*0.995
87	NO	NO	○		0.005	*0.995
88	NO	NO	○		0.005	*0.995
89	NO	NO	○		0.005	*0.995
90	NO	NO	○		0.005	*0.995
91	NO	NO	○		0.005	*0.995
92	NO	NO	○		0.005	*0.995
93	NO	NO	○		0.005	*0.995
94	NO	NO	○		0.005	*0.995
95	NO	NO	○		0.005	*0.995
96	NO	NO	○		0.005	*0.995
97	NO	NO	○		0.005	*0.995
98	NO	NO	○		0.005	*0.995
99	NO	NO	○		0.005	*0.995
100	NO	NO	○		0.005	*0.995
101	NO	NO	○		0.005	*0.995
102	NO	NO	○		0.005	*0.995
103	NO	NO	○		0.005	*0.995
104	NO	NO	○		0.005	*0.995
105	NO	NO	○		0.005	*0.995
106	NO	NO	○		0.005	*0.995

107	NO	NO	○		0.006	*0.994
108	NO	NO	○		0.006	*0.994
109	NO	NO	○		0.006	*0.994
110	NO	NO	○		0.006	*0.994
111	NO	NO	○		0.006	*0.994
112	NO	NO	○		0.006	*0.994
113	NO	NO	○		0.006	*0.994
114	NO	NO	○		0.006	*0.994
115	NO	NO	○		0.006	*0.994
116	NO	NO	○		0.006	*0.994
117	NO	NO	○		0.008	*0.992
118	NO	NO	○		0.008	*0.992
119	NO	NO	○		0.008	*0.992
120	NO	NO	○		0.012	*0.988
121	NO	NO	○		0.019	*0.981
122	NO	NO	○		0.019	*0.981
123	NO	NO	○		0.02	*0.98
124	NO	NO	○		0.02	*0.98
125	NO	NO	○		0.02	*0.98
126	NO	NO	○		0.02	*0.98
127	NO	NO	○		0.02	*0.98
128	NO	NO	○		0.02	*0.98
129	NO	NO	○		0.02	*0.98
130	NO	NO	○		0.022	*0.978
131	NO	NO	○		0.022	*0.978
132	NO	NO	○		0.022	*0.978
133	NO	NO	○		0.022	*0.978
134	NO	NO	○		0.022	*0.978
135	NO	NO	○		0.023	*0.977
136	NO	NO	○		0.023	*0.977
137	NO	NO	○		0.023	*0.977
138	NO	NO	○		0.023	*0.977
139	NO	NO	○		0.023	*0.977
140	NO	NO	○		0.023	*0.977
141	NO	NO	○		0.023	*0.977
142	NO	NO	○		0.023	*0.977
143	NO	NO	○		0.023	*0.977
144	NO	NO	○		0.023	*0.977
145	NO	NO	○		0.023	*0.977
146	NO	NO	○		0.025	*0.975
147	NO	NO	○		0.025	*0.975
148	NO	NO	○		0.025	*0.975

149	NO	NO	○		0.027	*0.973
150	NO	NO	○		0.03	*0.97
151	NO	NO	○		0.03	*0.97
152	NO	NO	○		0.03	*0.97
153	NO	NO	○		0.03	*0.97
154	NO	NO	○		0.033	*0.967
155	NO	NO	○		0.044	*0.956
156	NO	NO	○		0.044	*0.956
157	NO	NO	○		0.049	*0.951
158	NO	NO	○		0.049	*0.951
159	NO	NO	○		0.049	*0.951
160	NO	NO	○		0.049	*0.951
161	NO	NO	○		0.049	*0.951
162	NO	NO	○		0.049	*0.951
163	NO	NO	○		0.049	*0.951
164	NO	NO	○		0.059	*0.941
165	NO	NO	○		0.059	*0.941
166	NO	NO	○		0.059	*0.941
167	NO	NO	○		0.193	*0.807
168	NO	NO	○		0.297	*0.703
169	NO	NO	○		0.353	*0.647
170	NO	NO	○		0.382	*0.618
171	NO	NO	○		0.382	*0.618
26	NO	YES		○	*0.569	0.431
27	NO	YES			*0.818	0.182
28	NO	YES			*0.714	0.286
29	NO	YES			*0.742	0.258

表 A.3: 目的変数 106(基準値 50) の予測モデルの出力結果

No	実属性	予測	的中	フォロー	属性値 [YES]	属性値 [NO]
1	YES	YES	○		*0.978	0.022
2	YES	YES	○		*0.978	0.022
3	YES	YES	○		*0.947	0.053
4	YES	YES	○		*0.935	0.065
5	YES	YES	○		*0.935	0.065
6	YES	YES	○		*0.935	0.065
7	YES	YES	○		*0.935	0.065
8	YES	YES	○		*0.935	0.065
9	YES	YES	○		*0.935	0.065
10	YES	YES	○		*0.935	0.065
11	YES	YES	○		*0.931	0.069

12	YES	YES	○		*0.923	0.077
13	YES	YES	○		*0.923	0.077
14	YES	YES	○		*0.908	0.092
15	YES	YES	○		*0.9	0.1
16	YES	YES	○		*0.801	0.199
17	YES	YES	○		*0.741	0.259
18	YES	YES	○		*0.741	0.259
19	YES	YES	○		*0.741	0.259
20	YES	YES	○		*0.741	0.259
21	YES	YES	○		*0.741	0.259
22	YES	YES	○		*0.714	0.286
23	YES	YES	○		*0.714	0.286
24	YES	YES	○		*0.714	0.286
25	YES	YES	○		*0.683	0.317
26	YES	YES	○		*0.683	0.317
27	YES	YES	○		*0.683	0.317
28	YES	YES	○		*0.683	0.317
29	YES	YES	○		*0.683	0.317
30	YES	YES	○		*0.683	0.317
31	YES	YES	○		*0.681	0.319
32	YES	YES	○		*0.65	0.35
33	YES	YES	○		*0.65	0.35
34	YES	YES	○		*0.65	0.35
35	YES	NO		○	0.471	*0.529
36	YES	NO		○	0.462	*0.538
37	YES	NO		○	0.443	*0.557
38	YES	NO		○	0.443	*0.557
39	YES	NO		○	0.443	*0.557
40	YES	NO		○	0.443	*0.557
41	YES	NO		○	0.443	*0.557
42	YES	NO		○	0.443	*0.557
43	YES	NO		○	0.443	*0.557
44	YES	NO		○	0.443	*0.557
45	YES	NO		○	0.443	*0.557
46	YES	NO		○	0.435	*0.565
47	YES	NO		○	0.435	*0.565
48	YES	NO		○	0.435	*0.565
49	YES	NO			0.271	*0.729
50	YES	NO			0.271	*0.729
51	YES	NO			0.221	*0.779
52	YES	NO			0.221	*0.779
53	YES	NO			0.221	*0.779

54	YES	NO			0.221	*0.779
55	YES	NO			0.221	*0.779
56	YES	NO			0.221	*0.779
57	YES	NO			0.221	*0.779
58	YES	NO			0.221	*0.779
59	YES	NO			0.221	*0.779
60	YES	NO			0.221	*0.779
61	NO	YES		○	*0.572	0.428
62	NO	YES		○	*0.572	0.428
63	NO	YES		○	*0.54	0.46
64	NO	YES		○	*0.512	0.488
65	NO	YES		○	*0.506	0.494
66	NO	YES		○	*0.506	0.494
67	NO	YES		○	*0.506	0.494
68	NO	YES		○	*0.506	0.494
69	NO	YES		○	*0.506	0.494
70	NO	YES		○	*0.506	0.494
71	NO	YES			*0.972	0.028
72	NO	YES			*0.96	0.04
73	NO	YES			*0.79	0.21
74	NO	YES			*0.79	0.21
75	NO	YES			*0.743	0.257
76	NO	NO	○		0.475	*0.525
77	NO	NO	○		0.475	*0.525
78	NO	NO	○		0.475	*0.525
79	NO	NO	○		0.475	*0.525
80	NO	NO	○		0.475	*0.525
81	NO	NO	○		0.475	*0.525
82	NO	NO	○		0.475	*0.525
83	NO	NO	○		0.475	*0.525
84	NO	NO	○		0.475	*0.525
85	NO	NO	○		0.381	*0.619
86	NO	NO	○		0.381	*0.619
87	NO	NO	○		0.381	*0.619
88	NO	NO	○		0.381	*0.619
89	NO	NO	○		0.381	*0.619
90	NO	NO	○		0.281	*0.719
91	NO	NO	○		0.281	*0.719
92	NO	NO	○		0.271	*0.729
93	NO	NO	○		0.252	*0.748
94	NO	NO	○		0.252	*0.748
95	NO	NO	○		0.243	*0.757

96	NO	NO	○		0.243	*0.757
97	NO	NO	○		0.243	*0.757
98	NO	NO	○		0.243	*0.757
99	NO	NO	○		0.243	*0.757
100	NO	NO	○		0.243	*0.757
101	NO	NO	○		0.243	*0.757
102	NO	NO	○		0.243	*0.757
103	NO	NO	○		0.243	*0.757
104	NO	NO	○		0.243	*0.757
105	NO	NO	○		0.243	*0.757
106	NO	NO	○		0.243	*0.757
107	NO	NO	○		0.243	*0.757
108	NO	NO	○		0.243	*0.757
109	NO	NO	○		0.243	*0.757
110	NO	NO	○		0.243	*0.757
111	NO	NO	○		0.243	*0.757
112	NO	NO	○		0.243	*0.757
113	NO	NO	○		0.243	*0.757
114	NO	NO	○		0.243	*0.757
115	NO	NO	○		0.243	*0.757
116	NO	NO	○		0.243	*0.757
117	NO	NO	○		0.087	*0.913
118	NO	NO	○		0.081	*0.919
119	NO	NO	○		0.081	*0.919
120	NO	NO	○		0.081	*0.919
121	NO	NO	○		0.081	*0.919
122	NO	NO	○		0.053	*0.947
123	NO	NO	○		0.053	*0.947
124	NO	NO	○		0.053	*0.947
125	NO	NO	○		0.053	*0.947
126	NO	NO	○		0.019	*0.981
127	NO	NO	○		0.019	*0.981
128	NO	NO	○		0.019	*0.981
129	NO	NO	○		0.006	*0.994
130	NO	NO	○		0.006	*0.994
131	NO	NO	○		0.006	*0.994
132	NO	NO	○		0.006	*0.994
133	NO	NO	○		0.006	*0.994
134	NO	NO	○		0.006	*0.994
135	NO	NO	○		0.006	*0.994
136	NO	NO	○		0.006	*0.994
137	NO	NO	○		0.006	*0.994

138	NO	NO	○		0.006	*0.994
139	NO	NO	○		0.006	*0.994
140	NO	NO	○		0.006	*0.994
141	NO	NO	○		0.006	*0.994
142	NO	NO	○		0.006	*0.994
143	NO	NO	○		0.006	*0.994
144	NO	NO	○		0.006	*0.994
145	NO	NO	○		0.006	*0.994
146	NO	NO	○		0.006	*0.994
147	NO	NO	○		0.006	*0.994
148	NO	NO	○		0.006	*0.994
149	NO	NO	○		0.006	*0.994
150	NO	NO	○		0.006	*0.994
151	NO	NO	○		0.006	*0.994
152	NO	NO	○		0.006	*0.994
153	NO	NO	○		0.006	*0.994
154	NO	NO	○		0.006	*0.994
155	NO	NO	○		0.006	*0.994
156	NO	NO	○		0.006	*0.994
157	NO	NO	○		0.006	*0.994
158	NO	NO	○		0.006	*0.994
159	NO	NO	○		0.006	*0.994
160	NO	NO	○		0.006	*0.994
161	NO	NO	○		0.006	*0.994
162	NO	NO	○		0.006	*0.994
163	NO	NO	○		0.006	*0.994
164	NO	NO	○		0.006	*0.994
165	NO	NO	○		0.006	*0.994
166	NO	NO	○		0.006	*0.994
167	NO	NO	○		0.006	*0.994
168	NO	NO	○		0.006	*0.994
169	NO	NO	○		0.006	*0.994
170	NO	NO	○		0.006	*0.994
171	NO	NO	○		0.006	*0.994