

## Augmented Naive Bayes Classifier の厳密学習

## Exact Learning Augmented Naive Bayes Classifier

菅原聖太 植野真臣  
Shouta Sugahara Maomi Ueno

電気通信大学大学院情報理工学研究科

Graduate school of Informatics and Engineering, The University of Electro-Communications

The purpose of this study is improving classification accuracies of Bayesian network classifiers. This study proposes exact learning augmented naive Bayes which is expected to provide stable and high accuracy. Our experiments demonstrate that the proposed method outperforms the other methods to learn discriminative models.

## 1. はじめに

ベイジアンネットワークは、離散確率変数をノードとし、ノード間の条件付き従属関係を非循環有向グラフ (Directed Acyclic Graph: DAG) で表し、同時確率分布を各ノードの親ノード集合を所与とした条件付き確率パラメータの積に分解する、確率的グラフィカルモデルである。ベイジアンネットワークにおける一つのノードを目的変数とし、その他のノードを説明変数としたベイジアンネットワーク分類器 (Bayesian Network Classifier: BNC) は、離散変数を扱う分類器として知られている [Friedman 97]。

一般にベイジアンネットワークの DAG 構造はデータから推定する必要があり、この問題をベイジアンネットワークの構造学習と呼ぶ。構造学習では、候補構造から最適な学習スコアを持つ構造を探索するスコアベースアプローチが従来から行われてきた。一般にこのアプローチでは、漸近一致性を有する、構造の周辺尤度 (Marginal Likelihood: ML) を学習スコアとして用いる。

ML を用いると、全変数の同時確率分布をモデル化する生成モデルとして BNC を学習できる。しかし、Friedman ら [Friedman 97] は、BNC の構造学習スコアとして、生成モデルではなく、説明変数を所与とした目的変数の条件付き確率分布をモデル化する識別モデルのためのスコアを用いるべきだと主張した。そのような学習スコアとして、説明変数を所与とした目的変数の条件付き対数尤度 (Conditional Log Likelihood: CLL) が提案された。しかし、CLL はスコアの性質上、学習時間が膨大になってしまう。これを解決するため、CLL を用いた BNC の近似学習法が提案されてきた [Carvalho 11, Carvalho 13, Grossman 04]。これらの近似手法で学習した BNC の方が、ML で学習した BNC よりも分類精度が高いことが報告されている。

しかし、ML 最大化より CLL 最大化の方がなぜ良いかという理由については未だ明らかにされていない。ML は推定構造に対して漸近一致性が保証されており、サンプルサイズが大きい時に一致性のない CLL に分類精度が劣るのは奇異である。また、BNC の ML は閉形式で表せるため CLL より計算効率がよく、ML を大域的に最大化する構造を探索する厳密学習を効率的に行える。先行研究の比較実験では、ML を局所的に最

大化する構造を探索する近似学習を行なっているため、探索精度の悪さが影響したのかもしれない。

そこで、本研究ではまず ML による厳密学習と CLL による近似学習によって得られた BNC の分類精度を比較する。結果として、ML 最大化による BNC は厳密学習することで、大きく精度が向上することがわかった。特にサンプルサイズが大きいときに、最も分類精度が高いことが示された。しかし、厳密学習ではサンプルサイズが小さくなると ML を最大化する BNC の分類精度が低くなり、最も単純な構造をもつ Naive Bayes よりも低い場合もあった。特に、目的変数の親変数が多く子変数が少ないような構造を学習する場合に分類精度が低くなっていることがわかった。その理由は、目的変数の親変数が多いと、パラメータ数が指数的に増えるため、一つのパラメータ学習のためのサンプルサイズが小さくなり、推定精度が悪くなってしまふからである。一方、目的変数の子変数の増加に対してパラメータ数は線形増加しかしないため、上の問題は緩和される。

この性質を利用して、本論文では、目的変数が親変数を持たず、全説明変数が必ず目的変数の子となる Augmented Naive Bayes (ANB) 構造を制約とした BNC の厳密学習を提案する。さらに変数数が大きい場合に分類に不必要な変数が影響することを避けるために、厳密学習により目的変数のマルコフブランケットを抽出し、その後、ANB 構造を厳密学習する手法を提案した。ANB はこれまで識別モデルとして扱われてきたため、ML を最大化して学習することはなかった。本論の提案は、識別モデルの学習ではなく、生成モデルとしての GBN の学習に、目的変数の親変数が増えないように ANB 構造を制約することである。ベンチマークデータによる比較実験で、提案手法が従来の分類法よりも有意に精度が向上し、小サンプルでも高い分類精度を示せた。

## 2. ベイジアンネットワーク

## 2.1 ベイジアンネットワーク

ベイジアンネットワークは、確率変数をノードとし、ノード間の条件付き従属関係を非循環有向グラフで表し、各ノードの親ノード集合を所与とした条件付き確率で表現される確率的グラフィカルモデルである。今、離散確率変数集合  $\{X_0, X_1, \dots, X_i, \dots, X_n\}$  において、各変数  $X_i$  は  $r_i$  個の状態集合  $\{1, \dots, r_i\}$  から一つの値をとるとし、各変数  $X_i$  が値  $k$  をとるとき、 $X_i = k$  と書く。また、ベイジアンネットワークの構造を  $G$  とし、 $G$  における変数  $X_i$  の親変数集合を

連絡先: 電気通信大学大学院情報理工学研究科, 調布市調布ヶ丘 1-5-1, 042-484-8585,  
{sugahara,ueno}@ai.lab.uec.ac.jp

$\Pi_i$  とする。さらに、 $\Theta$  をベジアンネットワークの条件付き確率パラメータ集合とする。ベジアンネットワークでは、次式のように同時確率分布  $P(X_0, X_1, \dots, X_n | G, \Theta)$  を各変数の条件付き確率パラメータの積に分解して表せる。

$$P(X_0, X_1, \dots, X_n | G, \Theta) = \prod_{i=0}^n P(X_i | \Pi_i, G, \Theta)$$

### 2.1.1 ベジアンネットワークの構造学習

ベジアンネットワークの構造はデータから推定する必要があり、この問題を構造学習と呼ぶ。構造学習では、候補構造から最適な学習スコアを持つ構造を探索するスコアベースアプローチが従来から行われてきた。一般に学習スコアとして周辺尤度 (Marginal Likelihood: ML) を用いる。今、全変数に値が割り当てられたサンプルが  $N$  個あり、 $t$  番目のサンプルを  $\mathbf{d}^t = \langle x_0^t, x_1^t, \dots, x_n^t \rangle$  と表し、学習データを  $D = \langle \mathbf{d}^1, \dots, \mathbf{d}^t, \dots, \mathbf{d}^N \rangle$  と表す。パラメータの事前分布がディリクレ分布と仮定すると、ML は次のように閉形式で表される。

$$P(D | G) = \prod_{i=0}^n \prod_{j=1}^{q_i} \frac{\Gamma(N'_{ij})}{\Gamma(N'_{ij} + N_{ij}^D)} \prod_{k=1}^{r_i} \frac{\Gamma(N'_{ijk} + N_{ijk}^D)}{\Gamma(N'_{ijk})} \quad (1)$$

ここで、 $N_{ijk}$  は  $D$  において  $\Pi_i$  が  $j$  番目のパターンをとったとき ( $\Pi_i = j$  と書く) に  $X_i = k$  となる頻度を表し、 $N'_{ijk}$  はディリクレ事前分布のハイパーパラメータを表す。また、 $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$ 、 $N'_{ij} = \sum_{k=1}^{r_i} N'_{ijk}$  である。Buntine[Buntine 91] は式 (1) のハイパーパラメータを  $N'_{ijk} = N'/(r_i \cdot q_i)$  とした尤度等価を満たすスコアを提案しており、このスコアは Bayesian Dirichlet equivalent uniform (BDeu) と呼ばれ、最もよく用いられるスコアである。

一方、次式に示される、ML の近似である最小記述長 (Minimum Description Length: MDL) [Rissanen 89] は、ベジアンネットワークと学習データ  $D$  の同時記述長を表す。

$$MDL(D | G, \Theta) = \frac{\log N}{2} |\Theta| - \log P(D | G, \Theta) \quad (2)$$

MDL を用いた学習では、MDL を最小にする構造を最適解とする。式 (2) の第一項は構造の複雑さに対するペナルティ項である。式 (2) の第二項は構造のデータへの当てはまりを反映するフィッティング項を表す対数尤度である。

BDeu と MDL は各変数  $X_i$  とその親変数集合  $\Pi_i$  からなる局所構造のスコアについての総積 (MDL の場合は総和) が構造全体のスコアに一致する分解可能なスコアであり、効率的な学習ができる。

## 2.2 ベジアンネットワーク分類器

ベジアンネットワークにおける一つのノードを目的変数とし、それ以外のノードを説明変数とした分類器モデルはベジアンネットワーク分類器 (BNC) と呼ばれ、高い分類精度を持つことが知られている [Friedman 97]。

### 2.2.1 ベジアンネットワーク分類器のモデル

一般に、BNC の構造学習で探索する候補構造はとりうる全ての構造であり、そのような候補構造に対して BD や MDL など最適化して学習される BNC は General Bayesian Network (GBN) と呼ばれる。大きいネットワークでは GBN の学習に膨大な時間がかかってしまうため、候補構造に制約を入れて学習することが多い。例えば、GBN の下位構造として、各説明変数が目的変数のみを親に持つと仮定する Naive Bayes や、各説

明変数が目的変数を親に持ち、説明変数間で木構造をとると仮定した Tree-Augmented Naive Bayes (TAN) [Friedman 97] などが知られている。また、Naive Bayes や TAN を一般化した、より表現力の高いモデルとして、各説明変数が目的変数を親に持つことのみを仮定する Augmented Naive Bayes (ANB) [Friedman 97] が知られている。

### 2.2.2 ベジアンネットワーク分類器の学習

BDeu や MDL で学習した BNC は、全変数の同時確率分布をモデル化する生成モデルである。しかし、Friedman ら [Friedman 97] は、BNC の構造学習には、説明変数を所与とした目的変数の条件付き確率分布をモデル化する識別モデルのためのスコアを用いるべきだと主張した。そのようなスコアとして、以下で表される、説明変数を所与とした目的変数の条件付き対数尤度 (Conditional Log Likelihood: CLL) が提案された。

$$\sum_{t=1}^N \log P(x_0^t | x_1^t, \dots, x_n^t, G, \Theta) \quad (3)$$

しかし、CLL は分解可能ではないため、効率的な構造探索アルゴリズムを用いることができない。そこで、Grossman ら [Grossman 04] は近似的な構造探索法として、構造に対しエッジを一つ追加、消去、反転のいずれかの操作を行った時に最もスコアが良くなるようなエッジを選びその操作を行うというプロセスを繰り返して構造を更新する Hill-Climbing アルゴリズム [Heckerman 95] を用いた。Hill-Climbing アルゴリズムでは、任意のエッジの追加、消去、反転のどの操作を行ってもスコアが改善されない時に更新を終了し、その時の構造を解とする。一方、Carvalho ら [Carvalho 13] は、候補構造集合に ANB を仮定し、分解可能となるよう CLL を近似した aCLL (approximate Conditional Log Likelihood) を提案した。

これらの近似手法で学習した BNCの方が、ML (BDeu や MDL など) によって近似的に学習した BNC よりも分類精度が高いことが報告されてきた。

しかし、ML 最大化より CLL 最大化の方がなぜ良いかという理由についてはこれまで明らかにされていない。ML は推定構造に対して漸近一致性が保証されており、サンプルサイズが大きい時に一致性のない CLL に分類精度が劣るのは奇異である。また、BNC の ML は分解可能であるため、ML による厳密学習は CLL による厳密学習とは異なり、現実的な時間で学習できる。先行研究の比較実験では、ML による近似学習を行っているため、探索精度の悪さが影響したのかもしれない。

そこで、次章では ML によって厳密に学習した BNC と CLL によって近似的に学習した BNC の分類精度を比較する。

## 3. GBN の厳密学習と識別モデルの比較実験

本章では、BDeu によって厳密に学習した BNC と CLL によって近似的に学習した BNC の分類精度を比較する。この実験で比較する手法は、BDeu を用いて厳密学習した GBN (GBN-BDeu)、Naive Bayes、MDL のフィッティング項を CLL に置き換えた Conditional MDL を用いて近似学習した GBN (GBN-CMDL)、各変数が最大 2 つまでしか親を持たない構造を候補として、CLL を用いて近似学習した BNC (BNC-2P)、aCLL を用いて厳密学習した TAN (TAN-aCLL)、BDeu を用いて近似学習した GBN (gGBN-BDeu) である。近似学習の構造探索法としては Hill-Climbing を用い、厳密学習の構造探索手法としては動的計画法 [Silander 06] を用いた。UCI レポジトリ [Lichman 13] の 43 個のデータセットを用いた。

表 1: GBN-BDeu と従来手法の分類精度（太字は最大の分類精度）

| No.     | Dataset                      | Variables | Sample size | Classes | Naive-Bayes   | GBN-CMDL      | BNC2P         | TAN-aCLL      | gGBN-BDeu     | GBN-BDeu      |
|---------|------------------------------|-----------|-------------|---------|---------------|---------------|---------------|---------------|---------------|---------------|
| 1       | Balance Scale                | 5         | 625         | 3       | <b>0.9152</b> | 0.3333        | 0.8560        | 0.8656        | <b>0.9152</b> | <b>0.9152</b> |
| 2       | banknote authentication      | 5         | 1372        | 2       | 0.8433        | <b>0.8819</b> | 0.8797        | 0.8761        | <b>0.8819</b> | 0.8812        |
| 3       | Hayes-Roth                   | 5         | 132         | 3       | <b>0.8182</b> | 0.6136        | 0.6894        | 0.6742        | 0.7525        | 0.6136        |
| 4       | iris                         | 5         | 150         | 3       | 0.7133        | 0.7800        | 0.8200        | 0.8200        | 0.8133        | <b>0.8267</b> |
| 5       | lenses                       | 5         | 24          | 3       | 0.7500        | <b>0.8333</b> | 0.6667        | 0.7083        | <b>0.8333</b> | <b>0.8333</b> |
| 6       | Car Evaluation               | 7         | 1728        | 4       | 0.8571        | <b>0.9497</b> | 0.9416        | 0.9433        | 0.9416        | 0.9416        |
| 7       | liver                        | 7         | 345         | 2       | 0.6319        | 0.6145        | 0.6290        | <b>0.6609</b> | 0.6029        | 0.6087        |
| 8       | MONK's Problems              | 7         | 432         | 2       | 0.7500        | <b>1.0000</b> | <b>1.0000</b> | <b>1.0000</b> | 0.8449        | <b>1.0000</b> |
| 9       | mux6                         | 7         | 64          | 2       | 0.5469        | 0.3750        | <b>0.5625</b> | 0.4688        | 0.4063        | 0.4531        |
| 10      | led7                         | 8         | 3200        | 10      | 0.7294        | 0.7366        | <b>0.7375</b> | 0.7350        | 0.7297        | 0.7294        |
| 11      | HTRU2                        | 9         | 17898       | 2       | 0.7031        | 0.7096        | 0.7070        | 0.7018        | 0.7188        | <b>0.7305</b> |
| 12      | Nursery                      | 9         | 12960       | 3       | 0.6782        | <b>0.7126</b> | 0.6092        | 0.5862        | <b>0.7126</b> | <b>0.7126</b> |
| 13      | pima                         | 9         | 768         | 9       | 0.8966        | 0.9086        | 0.9118        | <b>0.9130</b> | 0.9092        | 0.9112        |
| 14      | post                         | 9         | 87          | 5       | 0.9033        | 0.5823        | <b>0.9442</b> | 0.9177        | 0.9291        | 0.9340        |
| 15      | Breast Cancer                | 10        | 277         | 2       | <b>0.9751</b> | 0.8917        | 0.9473        | 0.9488        | 0.7058        | <b>0.9751</b> |
| 16      | Breast Cancer Wisconsin      | 10        | 683         | 2       | <b>0.7401</b> | 0.6209        | 0.6823        | 0.7184        | 0.7094        | 0.7184        |
| 17      | Contraceptive Method Choice  | 10        | 1473        | 3       | 0.4671        | 0.4501        | <b>0.4745</b> | 0.4705        | 0.4440        | 0.4542        |
| 18      | glass                        | 10        | 214         | 6       | 0.5561        | 0.5654        | 0.5794        | <b>0.6308</b> | 0.4626        | 0.5701        |
| 19      | shuttle-small                | 10        | 5800        | 6       | 0.9384        | 0.9660        | <b>0.9703</b> | 0.9583        | 0.9683        | 0.9693        |
| 20      | threeOf9                     | 10        | 512         | 2       | 0.8164        | <b>0.9434</b> | 0.8691        | 0.8828        | 0.8652        | 0.8887        |
| 21      | Tic-Tac-Toe                  | 10        | 958         | 2       | 0.6921        | <b>0.8841</b> | 0.7338        | 0.7203        | 0.6754        | 0.8340        |
| 22      | MAGIC Gamma Telescope        | 11        | 19020       | 2       | 0.7482        | 0.7849        | 0.7806        | 0.7631        | 0.7844        | <b>0.7873</b> |
| 23      | Solar Flare                  | 11        | 1389        | 9       | 0.7811        | 0.8265        | 0.8315        | 0.8229        | <b>0.8431</b> | <b>0.8431</b> |
| 24      | heart                        | 14        | 270         | 2       | <b>0.8259</b> | 0.8185        | 0.8037        | 0.8148        | 0.8222        | <b>0.8259</b> |
| 25      | wine                         | 14        | 178         | 3       | 0.9270        | <b>0.9438</b> | 0.9157        | 0.9326        | 0.9045        | 0.9270        |
| 26      | cleve                        | 14        | 296         | 2       | <b>0.8412</b> | 0.8209        | 0.8007        | 0.8378        | 0.7973        | 0.7973        |
| 27      | australian                   | 15        | 690         | 2       | 0.8290        | 0.8312        | 0.8348        | 0.8464        | 0.8420        | <b>0.8536</b> |
| 28      | crx                          | 15        | 653         | 2       | 0.8377        | 0.8346        | 0.8208        | 0.8560        | <b>0.8622</b> | 0.8591        |
| 29      | EEG                          | 15        | 14980       | 2       | 0.5778        | 0.6787        | 0.6374        | 0.6125        | 0.6732        | <b>0.6814</b> |
| 30      | Congressional Voting Records | 17        | 232         | 2       | 0.9095        | 0.9698        | 0.9612        | 0.9181        | <b>0.9741</b> | 0.9655        |
| 31      | zoo                          | 17        | 101         | 5       | 0.9802        | 0.9109        | 0.9505        | <b>1.0000</b> | 0.9505        | 0.9307        |
| 32      | pendigits                    | 17        | 10992       | 10      | 0.8032        | 0.9062        | 0.8719        | 0.8700        | 0.9253        | <b>0.9290</b> |
| 33      | letter                       | 17        | 20000       | 26      | 0.4466        | <b>0.5796</b> | 0.5132        | 0.5093        | 0.5761        | 0.5761        |
| 34      | ClimateModel                 | 19        | 540         | 2       | 0.9222        | <b>0.9407</b> | 0.9241        | 0.9333        | 0.9370        | 0.9000        |
| 35      | Image Segmentation           | 19        | 2310        | 7       | 0.7290        | 0.7918        | 0.7991        | 0.7407        | 0.8026        | <b>0.8156</b> |
| 36      | lymphography                 | 19        | 148         | 4       | <b>0.8446</b> | 0.7939        | 0.7973        | 0.8311        | 0.7905        | 0.7500        |
| 37      | vehicle                      | 19        | 846         | 4       | 0.4350        | <b>0.5910</b> | <b>0.5910</b> | 0.5816        | 0.5461        | 0.5768        |
| 38      | hepatitis                    | 20        | 80          | 2       | 0.8500        | 0.7375        | <b>0.8875</b> | 0.8750        | 0.8500        | 0.5875        |
| 39      | german                       | 21        | 1000        | 2       | 0.7430        | 0.6110        | 0.7340        | <b>0.7470</b> | 0.7140        | 0.7210        |
| 40      | bank                         | 21        | 30488       | 2       | 0.8544        | 0.8618        | 0.8928        | 0.8618        | 0.8952        | <b>0.8956</b> |
| 41      | waveform-21                  | 22        | 5000        | 3       | 0.7886        | 0.7862        | 0.7754        | <b>0.7896</b> | 0.7698        | 0.7846        |
| 42      | Mushroom                     | 22        | 5644        | 2       | 0.9957        | <b>1.0000</b> | <b>1.0000</b> | 0.9995        | <b>1.0000</b> | 0.9949        |
| 43      | spect                        | 23        | 263         | 2       | 0.7940        | 0.7940        | 0.7903        | <b>0.8090</b> | 0.7603        | 0.7378        |
| average |                              |           |             |         | 0.7764        | 0.7721        | 0.7936        | 0.7943        | 0.7867        | <b>0.7963</b> |

各手法, 各データセットに対して, 10 分割交差検証によるテストデータの平均一致率を求め, 分類精度として表 1 に示した. 各データセットに対し, 各手法の中で最も高い分類精度を太字で示している.

表 1 を見ると, GBN-BDeu はサンプルサイズの大きいデータセット 11 番, 12 番, 22 番, 32 番, 40 番について, 比較手法 6 つの中で分類精度が最も高い. この結果から, ML によって学習した BNC が必ずしも CLL によって学習した BNC より分類精度が低いとは限らないことがわかる. しかし, サンプルサイズの小さいデータセットである 3 番, 9 番, 31 番では GBN-BDeu の分類精度が著しく低い. これらのデータセットでは, GBN-BDeu の目的変数の子変数が少なく, 親変数が多かった. このように目的変数の親変数が多いと, パラメータ数が指数的に増えるため, 一つのパラメータ学習のためのサンプルサイズが小さくなり, 推定精度が悪くなってしまう. 一方で, 目的変数の子変数の増加に対してパラメータは線形増加しなないため, 上の問題は緩和される.

この性質を考慮すると, 目的変数の親変数が存在しないような制約の下で BDeu を最大化する構造を探索すれば, 分類精度が高く安定した BNC を学習すると期待できる. これを実現できる単純な手法は, 目的変数が全ての説明変数に子を持つ ANB を候補構造とすることである. しかし, ANB では, 目的変数と無関係な変数が分類に影響を与えてしまう危険があり, その場合, 一般に分類精度は下がってしまう. そこで本論では, GBN-BDeu の目的変数に影響を及ぼす変数集合である

マルコフブランケットを取り出し, それらを説明変数とした BDeu を最大化する ANB の厳密学習法を提案する.

## 4. 評価実験

本章では, 提案手法と CLL をベースとした先行研究の手法の分類精度を比較するため, リポソトリデータを用いた評価実験を行う. 提案手法とその他の変数選択を適用した各手法との有意性を示すため, 分類精度の多重検定手法として標準的に用いられる Hommel の多重検定を行った. 検定の  $p$  値を表 2 の最下部に示した. ここで, 表 2 において右端の MANB-BDeu が提案手法であり, 各手法名の頭文字 'M' は変数選択を適用したことを示している.

結果として提案手法は, GBN-BDeu の学習構造の目的変数の子変数が少なく, 親変数が多かったデータセット 3 番, 9 番, 31 番において, GBN-BDeu の分類精度を改善している. この理由は, 3 章で述べたパラメータ数増加の問題を提案手法で緩和できるからである.

さらに, 表 2 より MANB-BDeu は全比較手法に対して有意水準 5% のもとで有意に分類精度が高いことがわかった.

## 5. むすび

本論では, 最初に ML によって厳密に学習した生成モデルの BNC と CLL によって近似的に学習した識別モデルの BNC



表 2: 変数選択を用いた場合の各手法の分類精度

| No. | Dataset                      | Variables | Sample size | Classes | MNaive-Bayes  | MGBN-CMDL     | MBNC-2P       | MTAN-aCLL     | MgGBN-BDeu    | GBN-BDeu      | MANB-BDeu     |
|-----|------------------------------|-----------|-------------|---------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| 1   | Balance Scale                | 5         | 625         | 3       | <b>0.9152</b> | 0.3333        | 0.8560        | 0.8656        | <b>0.9152</b> | <b>0.9152</b> | <b>0.9152</b> |
| 2   | banknote authentication      | 5         | 1372        | 2       | 0.8433        | <b>0.8819</b> | 0.8783        | 0.8761        | 0.8812        | 0.8812        | 0.8812        |
| 3   | Hayes-Roth                   | 5         | 132         | 3       | <b>0.8333</b> | 0.6136        | 0.7197        | 0.7879        | 0.7980        | 0.6136        | <b>0.8333</b> |
| 4   | iris                         | 5         | 150         | 3       | <b>0.8267</b> | 0.7800        | 0.8200        | 0.8200        | 0.8200        | <b>0.8267</b> | <b>0.8267</b> |
| 5   | lenses                       | 5         | 24          | 3       | 0.8333        | 0.8333        | 0.8333        | 0.8333        | <b>0.8750</b> | 0.8333        | 0.8333        |
| 6   | Car Evaluation               | 7         | 1728        | 4       | 0.8559        | 0.9242        | 0.9375        | 0.9363        | <b>0.9416</b> | <b>0.9416</b> | <b>0.9416</b> |
| 7   | liver                        | 7         | 345         | 2       | <b>0.6348</b> | <b>0.6348</b> | 0.6000        | 0.5942        | 0.6000        | 0.6087        | 0.5855        |
| 8   | MONK's Problems              | 7         | 432         | 2       | 0.7500        | <b>1.0000</b> | <b>1.0000</b> | <b>1.0000</b> | 0.8194        | <b>1.0000</b> | <b>1.0000</b> |
| 9   | mux6                         | 7         | 64          | 2       | 0.5469        | 0.3750        | <b>0.6250</b> | 0.4688        | 0.3906        | 0.4531        | 0.5469        |
| 10  | led7                         | 8         | 3200        | 10      | 0.7294        | 0.7363        | <b>0.7375</b> | 0.7350        | 0.7303        | 0.7294        | 0.7294        |
| 11  | HTRU2                        | 9         | 17898       | 2       | 0.7083        | 0.7057        | 0.7044        | 0.7070        | <b>0.7305</b> | <b>0.7305</b> | 0.7227        |
| 12  | Nursery                      | 9         | 12960       | 3       | <b>0.7126</b> | <b>0.7126</b> | <b>0.7126</b> | <b>0.7126</b> | <b>0.7126</b> | <b>0.7126</b> | <b>0.7126</b> |
| 13  | pima                         | 9         | 768         | 9       | 0.9102        | 0.9046        | 0.9076        | <b>0.9141</b> | 0.9083        | 0.9112        | <b>0.9141</b> |
| 14  | post                         | 9         | 87          | 5       | 0.8996        | 0.8775        | 0.9322        | 0.9103        | 0.9258        | <b>0.9340</b> | 0.9174        |
| 15  | Breast Cancer                | 10        | 277         | 2       | <b>0.9751</b> | 0.8909        | 0.9663        | 0.9458        | 0.9429        | <b>0.9751</b> | <b>0.9751</b> |
| 16  | Breast Cancer Wisconsin      | 10        | 683         | 2       | <b>0.7184</b> | <b>0.7184</b> | <b>0.7184</b> | <b>0.7184</b> | <b>0.7184</b> | <b>0.7184</b> | 0.7166        |
| 17  | Contraceptive Method Choice  | 10        | 1473        | 3       | 0.4549        | 0.4542        | <b>0.4555</b> | 0.4535        | 0.4501        | 0.4542        | 0.4549        |
| 18  | glass                        | 10        | 214         | 6       | <b>0.5841</b> | 0.5514        | 0.5467        | <b>0.5841</b> | 0.5047        | 0.5701        | 0.5654        |
| 19  | shuttle-small                | 10        | 5800        | 6       | 0.9360        | 0.9645        | 0.9666        | 0.9605        | 0.9690        | <b>0.9693</b> | <b>0.9693</b> |
| 20  | threeOf9                     | 10        | 512         | 2       | 0.8145        | 0.8750        | 0.8750        | 0.8809        | 0.8652        | <b>0.8887</b> | 0.8711        |
| 21  | Tic-Tac-Toe                  | 10        | 958         | 2       | 0.7182        | <b>0.8476</b> | 0.7244        | 0.7213        | 0.7359        | 0.8340        | <b>0.8476</b> |
| 22  | MAGIC Gamma Telescope        | 11        | 19020       | 2       | 0.7520        | 0.7841        | 0.7807        | 0.7699        | 0.7875        | 0.7873        | <b>0.7880</b> |
| 23  | Solar Flare                  | 11        | 1389        | 9       | <b>0.8431</b> | <b>0.8431</b> | <b>0.8431</b> | <b>0.8431</b> | <b>0.8431</b> | <b>0.8431</b> | <b>0.8431</b> |
| 24  | heart                        | 14        | 270         | 2       | 0.8222        | 0.8185        | 0.8148        | 0.8259        | 0.7889        | 0.8259        | <b>0.8296</b> |
| 25  | wine                         | 14        | 178         | 3       | <b>0.9607</b> | 0.9494        | 0.9438        | 0.9494        | 0.9326        | 0.9270        | 0.9326        |
| 26  | cleve                        | 14        | 296         | 2       | <b>0.8176</b> | <b>0.8176</b> | 0.7804        | 0.8108        | 0.7905        | 0.7973        | 0.8108        |
| 27  | australian                   | 15        | 690         | 2       | 0.8536        | <b>0.8580</b> | 0.8493        | 0.8522        | 0.8507        | 0.8536        | 0.8507        |
| 28  | crx                          | 15        | 653         | 2       | <b>0.8622</b> | 0.8545        | 0.8545        | <b>0.8622</b> | 0.8576        | 0.8591        | <b>0.8622</b> |
| 29  | EEG                          | 15        | 14980       | 2       | 0.5774        | 0.6790        | 0.6389        | 0.6111        | 0.6670        | 0.6814        | <b>0.6935</b> |
| 30  | Congressional Voting Records | 17        | 232         | 2       | 0.9353        | <b>0.9698</b> | 0.9655        | 0.9397        | 0.9655        | 0.9655        | 0.9569        |
| 31  | zoo                          | 17        | 101         | 5       | 0.9406        | 0.9406        | 0.9307        | 0.9307        | <b>0.9505</b> | 0.9307        | <b>0.9505</b> |
| 32  | pendigits                    | 17        | 10992       | 10      | 0.8032        | 0.9062        | 0.8719        | 0.8700        | 0.9253        | 0.9290        | <b>0.9297</b> |
| 33  | letter                       | 17        | 20000       | 26      | 0.4536        | <b>0.5796</b> | 0.5068        | 0.5036        | 0.5636        | 0.5761        | 0.5779        |
| 34  | ClimateModel                 | 19        | 540         | 2       | 0.9259        | <b>0.9407</b> | 0.9222        | 0.9352        | 0.9370        | 0.9000        | 0.8667        |
| 35  | Image Segmentation           | 19        | 2310        | 7       | 0.7662        | 0.7848        | 0.7918        | 0.7922        | 0.8022        | 0.8156        | <b>0.8203</b> |
| 36  | lymphography                 | 19        | 148         | 4       | <b>0.8176</b> | 0.7027        | 0.7770        | 0.8041        | 0.7770        | 0.7500        | 0.8108        |
| 37  | vehicle                      | 19        | 846         | 4       | 0.4634        | 0.5816        | 0.5721        | 0.5922        | 0.5437        | 0.5768        | <b>0.6028</b> |
| 38  | hepatitis                    | 20        | 80          | 2       | <b>0.8750</b> | 0.8500        | 0.8625        | 0.8500        | 0.8625        | 0.5875        | 0.6625        |
| 39  | german                       | 21        | 1000        | 2       | 0.7210        | 0.7250        | <b>0.7350</b> | 0.7230        | 0.7230        | 0.7210        | 0.7240        |
| 40  | bank                         | 21        | 30488       | 2       | 0.8680        | 0.8955        | 0.8924        | 0.8777        | 0.8954        | 0.8956        | <b>0.8966</b> |
| 41  | waveform-21                  | 22        | 5000        | 3       | 0.7852        | 0.7912        | 0.7806        | 0.7814        | 0.7626        | 0.7846        | <b>0.7920</b> |
| 42  | Mushroom                     | 22        | 5644        | 2       | 0.9970        | 0.9991        | 0.9991        | 0.9972        | <b>1.0000</b> | 0.9949        | <b>1.0000</b> |
| 43  | spect                        | 23        | 263         | 2       | <b>0.7865</b> | 0.7303        | 0.7416        | 0.7715        | 0.7715        | 0.7378        | 0.7603        |
|     | average                      |           |             |         | 0.7867        | 0.7801        | 0.7993        | 0.7981        | 0.7961        | 0.7963        | <b>0.8074</b> |
|     | p-value                      |           |             |         | 0.0089        | 0.0054        | 0.0104        | 0.0057        | 0.0188        | 0.0301        | -             |

の分類精度を比較した。その結果、ML を最大化する生成モデルが必ずしも CLL を最大化する識別モデルより分類精度が低いとは限らないことがわかった。しかし、ML は目的変数の親変数が多く子変数が少ないような構造を学習することがあり、その場合は CLL を最大化する BNC より著しく分類精度が低くなっていることがわかった。その理由は、目的変数の親変数が多いと、パラメータ数が指数的に増えるため、一つのパラメータ学習のためのサンプルサイズが小さくなり、推定精度が悪くなってしまふからである。一方で、目的変数の子変数の増加に対してパラメータ数は線形増加しきしないため、上の問題は緩和される。次に本論では、この性質を利用して、ML を最大化する GBN から目的変数のマルコフブランケットを取り出し、それを説明変数として ML を最大化する生成モデルとしての ANB を提案した。リポトリデータセットを用いた実験を行ったところ、提案手法は ML を最大化する生成モデルの GBN の問題点を改善し、分類精度の著しい低下を防ぐことができた。さらに、提案手法は既存の CLL を最大化する識別モデルの BNC よりも分類精度が有意に高いことを示した。

## 参考文献

[Buntine 91] Buntine, W.: Theory Refinement on Bayesian Networks, in *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*, pp. 52–60, San Francisco, CA, USA (1991), Morgan Kaufmann Publishers Inc.

- [Carvalho 11] Carvalho, A. M., Roos, T., Oliveira, A. L., and Myllymäki, P.: Discriminative Learning of Bayesian Networks via Factorized Conditional Log-Likelihood, *Journal of Machine Learning Research*, Vol. 12, pp. 2181–2210 (2011)
- [Carvalho 13] Carvalho, A. M., Adão, P., and Mateus, P.: Efficient Approximation of the Conditional Relative Entropy with Applications to Discriminative Learning of Bayesian Network Classifiers, *Entropy*, Vol. 15, No. 7, pp. 2716–2735 (2013)
- [Friedman 97] Friedman, N., Geiger, D., and Goldszmidt, M.: Bayesian Network Classifiers, *Machine Learning*, Vol. 29, No. 2, pp. 131–163 (1997)
- [Grossman 04] Grossman, D. and Domingos, P.: Learning Bayesian Network classifiers by maximizing conditional likelihood, in *Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004*, pp. 361–368 (2004)
- [Heckerman 95] Heckerman, D., Geiger, D., and Chickering, D. M.: Learning Bayesian Networks: The Combination of Knowledge and Statistical Data, *Machine Learning*, Vol. 20, No. 3, pp. 197–243 (1995)
- [Lichman 13] Lichman, M.: UCI Machine Learning Repository (2013)
- [Rissanen 89] Rissanen, J.: *Stochastic Complexity in Statistical Inquiry Theory*, World Scientific Publishing Co., Inc. (1989)
- [Silander 06] Silander, T. and Myllymäki, P.: A Simple Approach for Finding the Globally Optimal Bayesian Network Structure, in *Proceedings of Uncertainty in Artificial Intelligence*, pp. 445–452 (2006)