3L1-OS-06a-2

分類問題における視点中立化

Classification with Viewpoint Neutralization

福地 一斗*1 佐久間 淳*1 神嶌 敏弘*2
Kazuto Fukuchi Jun Sakuma Toshihiro Kamishima

*1筑波大学 大学院システム情報工学研究科 コンピュータサイエンス専攻
Dept. of Computer Science, Graduate school of SIE, University of Tsukuba

*2産業技術総合研究所
National Institute of Advanced Industrial Science and Technology (AIST)

Predictions made by machine learning must be neutral with respect to some personal attributes, such as gender, race, and religious. In this paper, we introduce a probabilistic model, refered to as a viewpoint, for neutralization in order to treat cases that such attributes are unobservable. We show our proposal method enforces neutrality to a prediction model with low decrease of the prediction accuracy in evaluation section.

1. はじめに

近年の機械学習の技術の発展と普及によって，学習による予測結果が個人の生活や行動に対して重大な影響をあたえるようになってきている．しかし一方で，この学習に因る予測結果が個人に対して不公正または差別的な事がある．以下の例で，機械学習によって引き起こされる不公正な扱いを説明する．

例 1．ある企業は，従業員と求職者の個人情報を収集し，機械学習によって従業員の情報を用いて求職者の仕事の能力を推定していたとする．個人情報には，年齢，性別，種族や民族性，居住地，職歴などの情報が含まれ，雇用の採決はこれらの情報に依存することになる．

例 2．web サービスを提供しているある企業は，検索履歴やページビューなどのユーザの振舞に関する情報を収集し，ユーザの属性や好みを機械学習で持って予測していたとする．web サービスは，予測した属性や好みからユーザ個人に適した広告の表示や推薦を行う．

例 1 では，能力の推定に性別，種族，民族性などの情報も用いるため，能力の推定結果がこれらの情報に依存してしまうと，雇用の採決もこれらの情報に依存し，差別的な採用を行なってしまう．また，フィルターバブル問題という，推薦に個人のバイアスが発生してしまう問題がある [Pariser 11]．例 2 において，政治や宗教に関係した記事を正確に推薦してしまうと，バイアスが発生してしまう．例えば，政治関係の記事において支持政党に関するものばかり見ていると，記事の推薦にそれ以外の政治関係の記事が排除されてしまう危険性がある．

例 2 においてユーザに適した広告を表示するためには，性別や年齢などといったユーザの属性を読み取る必要がある．しかし一方で，性別や人種などといった属性に依存して保険の広告の表示をすると，特定の人種に保険に加入する機会を損失させる可能性がある．従って，明確に推薦による個人化と差別を区別することが難しい．

本稿では，予測を行うために観測される個人情報などを**入力**，仕事の能力などの予測の結果を**出力**と表現する．また，差別の要因になるような人種や支持政党などを**視点**と表現する．

連絡先: 福地 一斗，筑波大学 大学院システム情報工学研究科 コンピュータサイエンス専攻，茨城県つくば市天王台 1-1-1，029-853-3826，kazuto@mdl.cs.tsukuba.ac.jp

中立性に配慮した学習を行うために，既存手法では中立性を評価するために指標を設定し，それを学習に組み込むことで実現を目指している [Kamishima 12b]．既存手法では，視点の値は訓練データの中に含まれているものとして扱っているが，例 2 のようにユーザの属性は与えられなくとも予測することで差別が発生する可能性がある．

本稿では，視点の予測モデルに対して中立性を保証できるような学習法を提案する．視点の予測モデルに対して中立性を保証できることにより，先の視点の値が与えられないような状況においても用いることが可能になる．また，モデルを用いることで中立性の汎化的な性能が期待できる．

2. 視点中立

この節では，提案する η-中立性を定義し，最尤推定の枠組みにおける分類問題で中立性を保証する手法について述べる．

$\mathcal{D} = \{(x_i, y_i) \in \mathcal{X} \times \mathcal{Y}\}_{i=1}^{N}$ を確率分布 $\Pr(X, Y)$ から i.i.d. に得られた訓練データとする．ここで，X, Y は，それぞれ入力と出力を表す確率変数である．パラメータ θ でモデル化された出力の予測関数 $f(Y|X;\theta) = \Pr(Y|X)$ を考え，与えられた訓練データを用いて最尤推定を行うことで予測モデルを学習する．最尤推定では，負の対数尤度をパラメータ θ について最小化することで，推定パラメータ $\theta^* = \mathrm{argmin}_{\theta\in\Theta} L(\theta)$ を得る．ここで，負の対数尤度 $L(\theta)$ は以下の式で与えられる．

$$L(\theta) = -\sum_{(x_i,y_i)\in\mathcal{D}} \ln f(y_i|x_i;\theta). \tag{1}$$

2.1 η-中立性

中立性について考えるために，入力 X と出力 Y に加え視点 V を導入する．以下の議論では，入力 X は連続変数を仮定して話を進めるが，単純に積分を総和に変えることによって離散でも議論が成り立つ．出力 Y と視点 V については，連続と離散のどちらでも成り立つように議論を進める．視点は，出力と同様にして，入力 X から条件付確率 $\Pr(V|X)$ によって予測を行う．また，出力と視点は，入力からそれぞれ独立に予測する．そのため，X, Y, V の同時確率は以下の式で表わせる．

$$\Pr(X, Y, V) = \Pr(X)\Pr(Y|X)\Pr(V|X). \tag{2}$$

この仮定の下，出力 Y と視点変数 V の依存性を考えることにより中立性を定義する．

V と Y が統計的独立関係にある場合，任意の $y \in \mathcal{Y}, v \in \mathcal{V}$ について $\Pr(v,y)/\Pr(v)\Pr(y) = 1$ が成り立つ．もし v, y が従属しているならば，$\Pr(v,y)/\Pr(v)\Pr(y) > 1$ となる．そこで，この周辺確率の比を用いて，中立性を以下のように定義する．

定義 1 (η-中立性)．X，Y をそれぞれ入力，出力を表す確率変数，V を視点を表す確率変数とする．ある与えられた $\eta \geq 0$ について，以下の式を満たすとき確率分布 $\Pr(X,Y,V)$ は η-中立であるという．

$$\forall v \in \mathcal{V}, y \in \mathcal{Y}, \quad \frac{\Pr(v,y)}{\Pr(v)\Pr(y)} \leq 1 + \eta. \tag{3}$$

次に，$\Pr(Y|X), \Pr(V|X)$ に対し確率的予測モデルを与え，与えられたモデルの同時確率が η-中立性を満たす条件を導く．確率分布 $\Pr(Y|X), \Pr(V|X)$ は，それぞれ出力と視点の予測モデル $f(Y|X;\theta) = \Pr(Y|X), g(V|X;\phi) = \Pr(V|X)$ で記述される．従って，$f(Y|X;\theta)$ と $g(V|X;\phi)$ が与えられたとき，確率モデル $\Pr(X,Y,V)$ は以下の式で与えられる．

$$M(X,Y,V;\theta,\phi) = \Pr(X) f(Y|X;\theta) g(V|X;\phi). \tag{4}$$

視点の予測モデルは事前に与えられ，変わらないものであるとする．そのため，以下では視点の予測モデルのモデルパラメータ ϕ は省略し，$g(V|X)$ と表記する．以下の定理で，式 (4) によってモデルが与えられているときに，モデルが η-中立性を満たすために必要な条件を示す．

定理 1. 入力 X，出力 Y，視点 V の同時確率が，モデル $M(X,Y,V;\theta) = \Pr(X) f(Y|X;\theta) g(V|X)$ で与えられているとする．$\forall v \in \mathcal{V}, y \in \mathcal{Y}$ について以下の式を満たすとき，モデル M は η-中立である．

$$\int_x \Pr(x) f(y|x;\theta) \left[g(v|x) - (1+\eta)\bar{g}(v) \right] dx \leq 0, \tag{5}$$

ここで，$\bar{g}(v) = \int_x \Pr(x) g(v|x) dx$ を表す．

2.2 η-中立性の経験分布による近似

一般に入力についての確率分布 $\Pr(x)$ は得ることができないため，訓練データ $\mathcal{D}$ における頻度分布によって η-中立性の経験的な評価を行う．訓練データ $\mathcal{D}$ における頻度分布 $\tilde{\Pr}(X)$ は，以下の式で与えられる．

$$\tilde{\Pr}(X = x) = \frac{1}{N} \sum_{i=1}^{N} I(x_i = x) \tag{6}$$

ここで，$I(\cdot)$ は指示関数 (indicator function) を表す．同時確率 $\Pr(X,Y,V)$ は，頻度分布を用いて $\tilde{\Pr}(X,Y,V) = \tilde{\Pr}(X)\Pr(Y|X)\Pr(V|X)$ で近似する．$\tilde{\Pr}(X,Y,V)$ を用いて，η-中立性を経験的な近似を行った，経験的 η-中立性を定義する．

定義 2 (経験的 η-中立性)．X，Y をそれぞれ入力，出力を表す確率変数，V を視点を表す確率変数，$\tilde{\Pr}(X)$ を訓練データ $\mathcal{D}$ による X の頻度分布とする．ある与えられた $\eta \geq 0$ について，$\tilde{\Pr}(X,Y,V)$ が η-中立であるとき，確率分布 $\Pr(X,Y,V)$ は訓練データ $\mathcal{D}$ について経験的 η-中立であるという．

以下の定理によって，式 (4) のモデルが与えられた訓練データについて経験的 η-中立である条件を示す．

定理 2. 入力 X，出力 Y，視点 V の同時確率が，モデル $M(X,Y,V;\theta) = \Pr(X) f(Y|X;\theta) g(V|X)$ で与えられているとする．$\forall v \in \mathcal{V}, y \in \mathcal{Y}$ について以下の式を満たすとき，モデル M は経験的 η-中立である．

$$\sum_{i=1}^{N} f(y|x_i;\theta) \left[g(v|x_i) - (1+\eta)\tilde{g}(v) \right] \leq 0, \tag{7}$$

ここで，$\tilde{g}(v) = \frac{1}{N} \sum_{i=1}^{N} g(v|x_i)$ である．

利便性のため，中立性の条件を以下の式で表記する．

$$N(y,v) = \sum_{i=1}^{N} f(y|x_i) \left[g(v|x_i) - (1+\eta)\tilde{g}(v) \right] \leq 0. \tag{8}$$

2.3 η-中立性による最尤推定

視点中立な最尤推定では，訓練データと視点の予測モデルが与えられた時に，η-中立性を保証しつつ最尤推定を行う．出力の予測を行うために，与えられた訓練データに対して対数尤度が最大である出力の予測モデルを得たい．しかし同時に，出力の予測モデルは，与えられた訓練データと視点の予測モデルに対して経験的 η-中立でありたい．この問題は，以下に示す制約付き最適化問題として記述できる．

$$\text{min. } L(\theta) \quad \text{s.t. } N(y,v;\theta) \leq 0 \quad \forall v, y.$$

本稿では，分類問題における最尤推定の一種であるロジスティック回帰において経験的 η-中立性を保証する，η-中立ロジスティック回帰を示す．

2.4 既存の中立性指標との比較

既存研究における視点は，公正配慮型データマイニングの文脈におけるセンシティブ属性に対応し，ここではその表現を用いる．CV スコア (Caldars-Verwer score)[Calders 10] は，出力，センシティブ属性が共に 2 値の時の公正性を測る量である．センシティブ属性の値を v_+，v_- としたとき，条件付確率の差 $p(y|v_+) - p(y|v_-)$ で定義される．CV スコアは，条件付確率は与えられた訓練データから経験的に計算される．

先入観尺度 (prejudice index: PI)[Kamishima 12a] は，出力 Y とセンシティブ属性 V の間の相互情報量によって定義される．先入観尺度は，訓練データ $\mathcal{D}$ が与えられた時に以下の式で計算される．

$$PI = I(Y;V) = \sum_{i=1}^{N} \tilde{\Pr}(y_i, v_i) \ln \frac{\tilde{\Pr}(y_i, v_i)}{\tilde{\Pr}(y_i)\tilde{\Pr}(v_i)} \tag{9}$$

これらの指標を中立性を測るために用いるとき，この指標が小さな値を取るように予測モデルを学習する．一方，提案した中立な最尤推定の定義では制約項として中立性を定式化し，η-中立性を満たすように予測モデルの学習を行う．

この違いは，中立性の解釈の仕方によって生じると考えられる．CV スコアや先入観尺度は，どちらも与えられた訓練データに対して統計的な観点から測られる不公正性を減らすことを目的としている．それに対し，η-中立性は，すべての出力と視点の予測について最も不公正となるところの上限を設定することを目的としている．

3. η-中立ロジスティック回帰

この節では，提案した中立性の定義を 2 値の分類問題におけるロジスティック回帰に適用する．

3.1 η-中立ロジスティック回帰の定式化

2 値の分類問題におけるロジスティック回帰は，入力の値域は $\mathcal{X} = \mathbb{R}^d$ であり，出力の値域は $\mathcal{Y} = \{0, 1\}$ の 2 値である．また，$\boldsymbol{\theta} \in \mathbb{R}^d$ をモデルパラメータとしたとき，出力の予測モデルは以下の式で与えられる．

$$f(y|\boldsymbol{x};\boldsymbol{\theta}) = \sigma(\boldsymbol{\theta}^T\boldsymbol{x})^y(1-\sigma(\boldsymbol{\theta}^T\boldsymbol{x}))^{1-y}, \qquad (10)$$

ここで，$\sigma(a) = 1\,/\,(1+e^{-a})$ はロジスティックシグモイド関数を示す．式 (10) を出力の予測モデルとした時，対数尤度は式 (1) で与えられる．そして，η-中立ロジスティック回帰の分類問題は以下のように定式化される．

$$\text{min. } L(\boldsymbol{\theta}) \quad \text{s.t. } N(y, v; \boldsymbol{\theta}) \leq 0 \quad \forall v, y.$$

視点の予測モデル $g(v|\boldsymbol{x})$ については，任意の確率モデルを用いることができる．

3.2 最適化

ここでは，上で定式化した η-中立ロジスティック回帰の最適化手法について述べる．η-中立ロジスティック回帰の分類問題の目的関数である負の対数尤度は，凸になることが知られている．そこで，制約項の凸性を考察することで，η-中立ロジスティック回帰の最適化問題の凸性を考察する．

制約項 $N(y, v; \boldsymbol{\theta})$ は関数 f の線形結合によって表現されているため，f の凸性について調査する．f のパラメータ $\boldsymbol{\theta}$ に関する勾配とヘッセ行列は，以下の式で求められる．

$$\nabla f(y, \boldsymbol{x}; \boldsymbol{\theta}) = \Big(y - \sigma(\boldsymbol{\theta}^T\boldsymbol{x})\Big) f(y|\boldsymbol{x};\boldsymbol{\theta})\boldsymbol{x}.$$
$$\nabla^2 f(y|\boldsymbol{x};\boldsymbol{\theta}) = \alpha(\boldsymbol{x}, y, \boldsymbol{\theta}) f(y|\boldsymbol{x};\boldsymbol{\theta})\boldsymbol{x}\boldsymbol{x}^T,$$

ここで，$\alpha(\boldsymbol{x}, y, \boldsymbol{\theta}) = 2\sigma(\boldsymbol{\theta}^T\boldsymbol{x})^2 + y^2 - (2y+1)\sigma(\boldsymbol{\theta}^T\boldsymbol{x})$ である．$\alpha(\boldsymbol{x}, y, \boldsymbol{\theta}) \in \mathbb{R}$ がパラメータ $\boldsymbol{\theta}$ によって負の値を取るため，ヘッセ行列は正定値ではなく，関数 f は凸ではない．従って，$g(v|\boldsymbol{x})$ の取り方に関係なくロジスティック回帰における中立性の条件が非凸になる．

実験では η-中立ロジスティック回帰の最適化問題を解くために，応的空間拡張 (adaptive space dilation) を基にした非線形の凸最適化手法である Shor の r-algorithm を用いた [Shor 85]．制約は非凸になるが，実験によって予測精度をあまり下げることなく η-中立性を実現できることを 4. 節で示す．中立性を保証した凸性を持った学習方法の考案は，今後の課題である．

4. 実験

分類問題における提案手法の性能を評価するために Adult[*1] と，Dutch Census[*2] の 2 つのデータセットに対して実験を行った．表 1 に，各データセットについての仕様をまとめたものを示す．表中の，#Instances はデータ数，#Attributes は属性数を表しており，$\#y_+$， $\#v_+$ は，それぞれ，出力と視点変数の正例の数を表す．また，ベースラインとして，ロジスティック回帰で学習した際の出力 (Acc(y)) と視点 (Acc(v)) の正答率を示す．収入の High/Low を予測する問題において，性別を視点とした．視点の予測モデル $g(v|x)$ は，ロジスティクス回帰による予測器を用いた．

実験では，ロジスティック回帰 (LR)，視点を用いないロジスティック回帰 (LRns)，単純ベイズ法 (NB)，視点を用いない単純ベイズ法 (NBns)，Calders-Verwers の 2 単純ベイズ法 (CV2NB)[Calders 10]，正則化によるロジスティック回帰の先入観削除手法 (PR)[Kamishima 12a]，η-中立ロジスティック回帰 (VN) を比較した．PR のパラメータ λ は $\lambda \in \{0, 5, 10, 15, 20, 30\}$，VN のパラメータ η は $\eta \in \{0.00, 0.01, ..., 0.40\}$ に設定した．

表 1: 使用したデータセットの仕様.

dataset	Adult	Dutch Census
#Instances	16281	60420
#Attributes	13	10
Acc(y)	0.851	0.835
Acc(v)	0.842	0.665

表 2: 2 つの設定と各アルゴリズムにおける視点の扱い

case	method	learning of $f(x, y)$	neutrality guaratee	neutrality measure
Case 1	others	$\boldsymbol{x}, v$	v	$\hat{y}, v$
	ours	$\boldsymbol{x}, v$	$g(v\|\boldsymbol{x})$	$\hat{y}, v$
Case 2	others	$x, \hat{v}$	$\hat{v}$	$\hat{y}, v$
	ours	$\boldsymbol{x}$	$g(v\|\boldsymbol{x})$	$\hat{y}, v$

実験の評価は正答率，正規化先入観尺度 (normalized prejudice index: NPI) に加え η-中立性に基づいた中立性の指標 $\hat{\eta}$ を用いて評価を行う．指標は，5 分割交差検定を用いて算出する．NPI は，出力 Y と視点 V の間の $[0, 1]$ に正規化された相互情報量 $NPI = PI/H(Y)H(V)$ で定義される [Kamishima 12a]．ここで，$H(\cdot)$ はエントロピーである．

η-中立性に基づいた中立性の指 $\hat{\eta}$ は，以下の式で定義される．

$$\hat{\eta} = \max_{y\in\mathcal{Y}, v\in\mathcal{V}} \frac{\tilde{\Pr}(y, v)}{\tilde{\Pr}(y)\tilde{\Pr}(v)} - 1 \qquad (11)$$

$\hat{\eta}$ は，y, v の中で最も依存している組みの依存性を評価している．Y と V が独立であれば，$\hat{\eta} = 0$ となる．$\hat{\eta}$ は，予測モデルが訓練データに対して $\hat{\eta}$-中立であることを意味している．

4.1 設定

表 2 に，予測モデルの学習，中立性の保証，中立指標の評価の際に各アルゴリズムが用いる値を示す．実験では，予測モデルのみを用いた時の性能評価のために，2 つの状況を仮定する．

Case 1 は，視点の実測値が訓練データとして与えられる状況を表す．提案法は訓練データから視点の予測モデルを構築して中立化し，既存手法では視点の実測値を用いて中立化する．

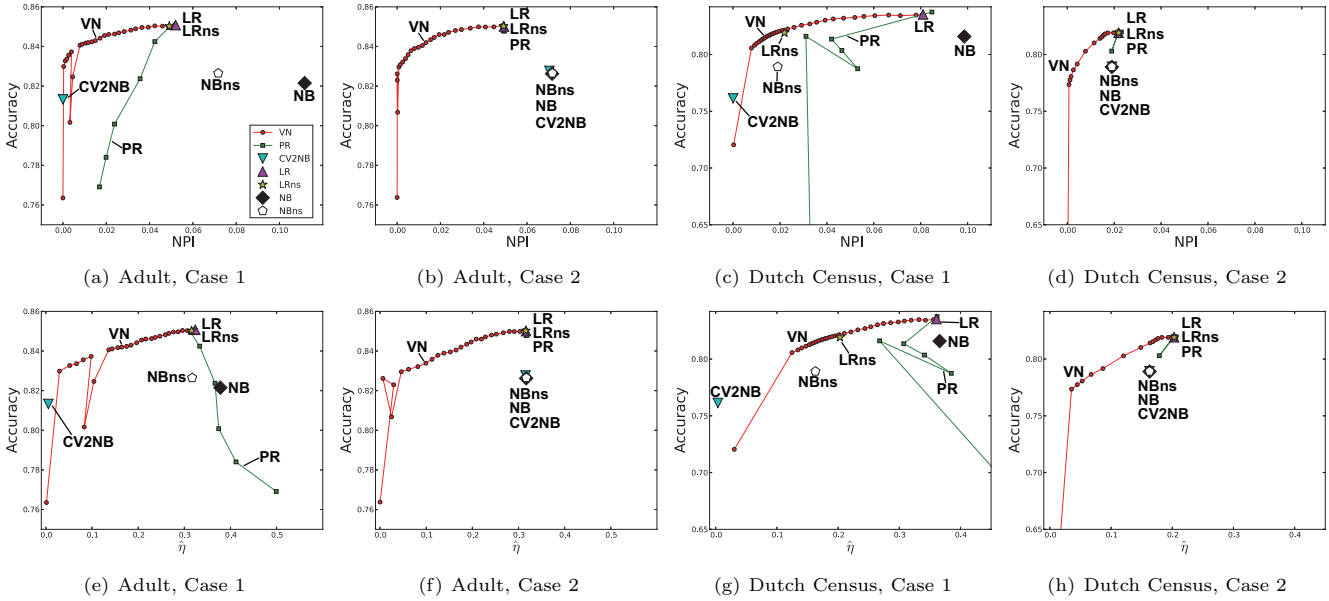
Case 2 は，視点の実測値が与えられず，視点の予測モデルのみしか与えられない状況を表す．提案法は，与えられた予測モデル $g(v|\boldsymbol{x})$ を用い，既存手法では視点の予測値 $\hat{v} = \text{argmax}_v\, g(v|\boldsymbol{x})$ を用いて中立化する．中立性指標の計算に用いる視点の値は，予測値でなく真の値 v を用いる．

4.2 結果

図 1 に実験結果を示す．上の段は中立性の評価として NPI，下の段は $\hat{\eta}$ を用いたときのグラフである．各グラフは，横軸が中立性の評価指標を表しており，縦軸が正答率を表している．中立性指標はどちらとも小さいほどよく，正答率は大きいほどよいため，グラフは左上に行くほど性能が良いといえる．

単純ベイズ，ロジスティック回帰において，視点を使わずに学習を行った NBns と LRns の結果を見ていく．既存手法における Case 1 と Case 2 の違いは，視点に実測値を使うか，予測値を使うかであり，視点を用いない NBns と LRns は 2

*1 http://archive.ics.uci.edu/ml/datasets/Adult
*2 https://sites.google.com/site/conditionaldiscrimination/



(a) Adult, Case 1 (b) Adult, Case 2 (c) Dutch Census, Case 1 (d) Dutch Census, Case 2

(e) Adult, Case 1 (f) Adult, Case 2 (g) Dutch Census, Case 1 (h) Dutch Census, Case 2

図 1: 正答率と中立性の対比．中立性の指標は, それぞれ， NPI(上段)，$\hat{\eta}$(下段)．

つ Case における違いがない．従って，Case 1 を見ると，NB，LR に比べると，NBns，LRns はどちらの指標も低くなっているが，完全に中立な 0 のラインと比べるとまだ大きい．そのため，視点を用いない学習では中立化が不十分である．

CV2NB は，Case 1 において，NPI，$\hat{\eta}$ ともに 0 付近まで下げることができている．さらに中立化に伴う正答率の低下をみると，Adult では 1%未満，Dutch Census では 5%未満程度に抑えられている．そのため，CV2NB は Case 1 において，中立化が行えているといえる．しかし，Case 2 では，CV2NB が NBns と変わらない点にあることがわかる．従って，CV2NB は Case 2 で中立化が達成できていない．

PR は，Case 1 の Adult において，パラメータを変化させると正答率が低下するが，それに伴い NPI も低下させられている．しかし，$\hat{\eta}$ はパラメータの変化に伴い大きくなる．これは，PR は NPI を直接最適化し，平均的な従属性のみを考えているため，一番従属しているところを評価する $\hat{\eta}$ は逆に大きくなったと考えられる．Case 1 の Dutch Census は，安定していない．これは，NPI がパラメータに対して非凸であり，局所最適解に陥いることが多いためであると考えられる．Case 2 では，どちらのデータセットでも，ほとんどのパラメータで PR が LR と変わらない点にある．従って，PR は Case 2 では正答率が落ちてはいないが，中立化ができていない．

提案法である VN は，Case 1 でも Case 2 でもパラメータの変化に伴って，NPI，$\hat{\eta}$ 両方共低下させることができている．Adult では制約の非凸性によって正答率が低下している点もあるが，それ以外の点では，中立化に伴う正答率の低下も 5%未満に抑えられている．CV2NB と比べると，Case 1 では，VN は CV2NB より NPI，$\hat{\eta}$ が高いため CV2NB のほうが高い中立性を達成している．従って，Case 1 での性能は CV2NB より劣るが，VN は Case 1，Case 2 両方で中立化を実現可能であり，高いトレードオフ比を実現している．

5. まとめ

学習モデルが予測分布として与えられる最尤推定の枠組みにおいて，中立性を配慮する手法の提案を行った．実験では，既存の手法においてセンシティブ属性のモデルのみ与えられている状況で，中立性が達成できないことを示した．また，提案法では予測精度と中立性のトレードオフを実現できることを示した．提案法は，既存手法では性能が著しく低下した，モデルのみが与えられている状況でも高い中立性に関する性能を達成していることを示した．

中立化には他にも応用があり，今後はその他の応用について性能評価を行いたいと考えている．また，中立化の際に非凸の制約が出てきたため，この凸化が今後の課題である．

謝辞

本研究は，最先端研究開発プログラム「超巨大データベース時代に向けた最高速データベースエンジンの開発と当該エンジンを核とする戦略的社会サービスの実証・評価」および部分的に科学研究費 24500194 の助成を受けました．

参考文献

[Calders 10] Calders, T. and Verwer, S.: Three Naive Bayes Approaches for Discrimination-Free Classification, *Data Mining and Knowledge Discovery*, Vol. 21, No. 2, pp. 277–292 (2010)

[Kamishima 12a] Kamishima, T., Akaho, S., Asoh, H., and Sakuma, J.: Fairness-aware Classifier with Prejudice Remover Regularizer, in *Proceedings of the ECML/PKDD2012, Part II*, Vol. LNCS 7524, pp. 35–50, Springer (2012)

[Kamishima 12b] Kamishima, T., Akaho, S., and Sakuma, J.: Considerations on Fairness-aware Data Mining, in *Proceedings of the IEEE International Workshop on Discrimination and Privacy-Aware Data Mining (DPADM@ICDM)*, pp. 378–385 (2012)

[Pariser 11] Pariser, E.: *The Filter Bubble: What The Internet Is Hiding From You*, Viking, London (2011)

[Shor 85] Shor, N. Z., Kiwiel, K. C., and Ruszcayǹski, A.: *Minimization methods for non-differentiable functions*, Springer-Verlag New York, Inc., New York, NY, USA (1985)