Neural Computing and Applications (2021) 33:1171–1191
https://doi.org/10.1007/s00521-020-05038-8

ORIGINAL ARTICLE

GMM discriminant analysis with noisy label for each class

Jian-wei Liu[1] · Zheng-ping Ren[1] · Run-kun Lu[1] · Xiong-lin Luo[1]

Received: 8 June 2019 / Accepted: 13 May 2020 / Published online: 1 June 2020

Abstract

Keywords

1 Introduction

[5], [2], [2]

1.

2.

[2]

[5]

1.1 Discriminant analysis based on Gaussian mixture models

1.1.1 Description of the problem with the noise labels

$$\mathbf{x} = (x_1, \ldots, x_d) \in \mathcal{X}, \qquad \mathcal{X} = \mathcal{R}^d$$

$\Omega = \{\omega_1, \omega_2, \ldots, \omega_K\}$. $\mathbf{x} \in \mathcal{X}$

$p(\mathbf{x}|\omega)$

$p(\omega)$, $\omega \in \Omega$.

$\mathcal{S}_\omega$ $\omega \in \Omega$

$$\mathcal{S}_\omega = \{\mathbf{x} \in \mathcal{X}\},\ \omega \in \Omega\ ;\ \mathcal{S} = \bigcup_{\omega \in \Omega} \mathcal{S}_\omega,\ |\mathcal{S}| = \sum_{\omega \in \Omega} |\mathcal{S}_\omega|,$$

$|\mathcal{S}|$ $|\mathcal{S}_\omega|$

$\mathcal{S}_\omega$.

$p(\mathbf{x}|\omega)$.

L_ω

$$L_\omega = \frac{1}{|\mathcal{S}_\omega|} \sum_{\mathbf{x} \in \mathcal{S}_\omega} p(\mathbf{x}|\omega)p(\omega), \quad (\omega \in \Omega),$$

$1/|\mathcal{S}_\omega|$

$\omega \in \Omega$ $\mathbf{x} \in \mathcal{S}$

$\mathbf{x} \in \mathcal{S}$,

$\tilde{\omega} \in \Omega$, ω.

$p(\tilde{\omega})$ $\tilde{\omega}$

$p(\omega)$ ω.

$p(\omega|\tilde{\omega})$ ω

$\tilde{\omega}$.

$\tilde{p}(\mathbf{x}|\tilde{\omega})$ $\mathbf{x} \in \mathcal{X}$ $\tilde{\omega}$

$$\tilde{p}(\mathbf{x}|\tilde{\omega}) = \sum_{\omega \in \Omega} p(\mathbf{x}|\omega)p(\omega|\tilde{\omega}), \mathbf{x} \in \mathcal{X},\ \tilde{\omega} \in \Omega,$$

$p(\omega|\tilde{\omega})$,

$\omega \in \Omega$

$\mathbf{x} \in \mathcal{X}$. $p(\omega|\tilde{\omega})$, $\omega \in \Omega$, $p(\tilde{\omega})$,

$\tilde{\omega} \in \Omega$ r , r b
r .

1.1.2 Gaussian mixture model

r l r r b b
r b , r , l , 22
, ,
r . r r
b r r
r b b r b r .

'04F19461292 113.904F1f 3.7Tf 1.07536580Tj /F103759 .9389 0 T25. -.0285 TD045Q

$$L = \prod_{\mathbf{x} \in \mathcal{S}} p(\mathbf{x}|\tilde{\omega}, \theta_{\tilde{\omega}}) p(\omega)$$ Qx

$$\sum_{\omega\in\Omega} q(\omega) = e^{-(1+\lambda_{G_1})} \sum_{\omega\in\Omega} p(\mathbf{x}|\omega,\theta_\omega)\gamma_{\tilde{\omega},\omega}\pi_\omega = 1$$
$$\Rightarrow \lambda_{G_1} = \log \sum_{\omega\in\Omega} p(\mathbf{x}|\omega,\theta_\omega)\gamma_{\tilde{\omega},\omega}\pi_\omega - 1.$$

λ_{G_1} ... $q(\omega)$ ...

$$q(\omega) = \frac{p(\mathbf{x}|\omega,\theta_\omega)\gamma_{\tilde{\omega},\omega}\pi_\omega}{\sum_{\omega\in\Omega} p(\mathbf{x}|\omega,\theta_\omega)\gamma_{\tilde{\omega},\omega}\pi_\omega} = p(\omega|x,\tilde{\omega})$$

... $p(\mathbf{x}|\omega,\theta_\omega) = \log \sum_{m\in\mathcal{M}} w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})$ ...

$$\begin{aligned}
\log p(\mathbf{x}|\omega,\theta_\omega) &= \log \sum_{m\in\mathcal{M}} w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})\\
&= \log \sum_{m\in\mathcal{M}} h_{\omega,m} \frac{w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})}{h_{\omega,m}}\\
&\geq \sum_{m\in\mathcal{M}} h_{\omega,m} \log \frac{w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})}{h_{\omega,m}}\\
&= \sum_{m\in\mathcal{M}} h_{\omega,m} \log w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m}) - \log h_{\omega,m}
\end{aligned}$$

... $\sum_{m\in\mathcal{M}} h_{\omega,m} = 1$ ... λ_{G_2} ... $h_{\omega,m}$ ...

$$G_2 = \sum_{m\in\mathcal{M}} h_{\omega,m} \log w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m}) - \log h_{\omega,m} + \lambda_{G_2}\left(1 - \sum_{m\in\mathcal{M}} h_{\omega,m}\right)$$

... $h_{\omega,m}$ ...

$$\log w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m}) - \log h_{\omega,m} - 1 - \lambda_{\lambda_{G_2}} = 0$$
$$\Rightarrow \log h_{\omega,m} = \log w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m}) - 1 - \lambda_{\lambda_{G_2}}$$
$$\Rightarrow h_{\omega,m} = w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m}) \cdot e^{-(1+\lambda_{\lambda_{G_2}})}$$

... $h_{\omega,m}$ ... λ_{G_2} ...

$$\sum_{m\in\mathcal{M}} h_{\omega,m} = \sum_{m\in\mathcal{M}} w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m}) \cdot e^{-(1+\lambda_{G_2})}$$
$$\Rightarrow \lambda_{G_2} = \log \sum_{m\in\mathcal{M}} w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m}) - 1.$$

... λ_{G_2} ... $h_{\omega,m}$ ...

$$h_{\omega,m} = \frac{w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})}{\sum_{m\in\mathcal{M}} w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})} = \mathbf{r}(m|\mathbf{x},\omega)$$

...

$$Q = \sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} \sum_{\omega\in\Omega} \left(p(\omega|\mathbf{x},\tilde{\omega}) \times \sum_{m\in\mathcal{M}} \mathbf{r}(m|\mathbf{x},\omega) \log \frac{w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})}{\mathbf{r}(m|\mathbf{x},\omega)} \right) + \sum_{\mathbf{x}\in\mathcal{S}} \sum_{\omega\in\Omega} p(\omega|\mathbf{x},\tilde{\omega}) \log \gamma_{\tilde{\omega},\omega} + \sum_{\mathbf{x}\in\mathcal{S}} \sum_{\omega\in\Omega} p(\omega|\mathbf{x},\tilde{\omega}) \log \pi_\omega$$

... $\Theta = \{\theta_\omega\}_{\omega\in\Omega}$, $\Gamma = \{\gamma_{\tilde{\omega},\omega}\}_{\tilde{\omega}\in\Omega}$, ... $\Pi = \{\pi_\omega\}_{\omega\in\Omega}$, ... $\theta_\omega = \{w_{\omega,m},\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m}\}_{m\in\mathcal{M}}$. ...

E step $p(\omega|\mathbf{x},\tilde{\omega})$ ... $\mathbf{r}(m|\mathbf{x},\omega)$ ...

$$p(\omega|\mathbf{x},\tilde{\omega}) = \frac{p(\mathbf{x}|\omega,\theta_\omega)\gamma_{\tilde{\omega},\omega}\pi_\omega}{\sum_{\omega\in\Omega} p(\mathbf{x}|\omega,\theta_\omega)\gamma_{\tilde{\omega},\omega}\pi_\omega}$$

$$\mathbf{r}(m|\mathbf{x}_n,\omega_n=k) = \frac{w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})}{\sum_{m\in\mathcal{M}} w_{\omega,m} g(\mathbf{x}|\omega,\boldsymbol{\mu}_{\omega,m},\boldsymbol{\Sigma}_{\omega,m})}$$

M step ... $\Theta = \{\theta_\omega\}_{\omega\in\Omega}$, $\Gamma = \{\gamma_{\tilde{\omega},\omega}\}_{\tilde{\omega}\in\Omega}$, ... $\Pi = \{\pi_\omega\}_{\omega\in\Omega}$.

1.1.3 Updating

... $\boldsymbol{\mu}_{\omega,m}$ ...

m

$$\frac{Q}{\boldsymbol{\Sigma}_{\omega,m}} = -\sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} \Big\{ p(\omega|\mathbf{x},\tilde{\omega})\, \mathbf{r}(m|\mathbf{x},\omega) \cdot \Big[\boldsymbol{\Sigma}_{\omega,m}^{-1} - \boldsymbol{\Sigma}_{\omega,m}^{-1}\big(\mathbf{x}-\boldsymbol{\mu}_{\omega,m}\big)\big(\mathbf{x}-\boldsymbol{\mu}_{\omega,m}\big)^{T}\boldsymbol{\Sigma}_{\omega,m}^{-1}\Big]\Big\} = 0 \Rightarrow \boldsymbol{\Sigma}_{\omega,m} = \frac{\sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega})\, \mathbf{r}(m|\mathbf{x},\omega)\big(\mathbf{x}-\boldsymbol{\mu}_{\omega,m}\big)\big(\mathbf{x}-\boldsymbol{\mu}_{\omega,m}\big)^{T}}{\sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega})\, \mathbf{r}(m|\mathbf{x},\omega)} \tag{11}$$

… $w_{\omega,m}$ … $\lambda_{w_{\omega,m}}$ … $\sum_{m\in\mathcal{M}} w_{\omega,m} = 1$, …

$$Q_{\lambda_{w_{\omega,m}}} = Q + \lambda_{w_{\omega,m}}\left(1 - \sum_{m\in\mathcal{M}} w_{\omega,m}\right),$$

… $Q_{\lambda_{w_{\omega,m}}}$ … $w_{\omega,m}$ …

$$\frac{Q_{\lambda_{w_{\omega,m}}}}{w_{\omega,m}} = 0 = \sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega})\, \mathbf{r}(m|\mathbf{x},\omega)\frac{1}{w_{\omega,m}} - \lambda_{w_{\omega,m}}$$
$$\Rightarrow \lambda_{w_{\omega,m}} w_{\omega,m} = \sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega})\, \mathbf{r}(m|\mathbf{x},\omega)$$
$$\overset{w_{\omega,m}}{\Rightarrow} \lambda_{w_{\omega,m}} \sum_{m\in\mathcal{M}} w_{\omega,m} = \lambda_{w_{\omega,m}} = \sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega}) \sum_{m\in\mathcal{M}} \mathbf{r}(m|\mathbf{x},\omega),$$

… $\lambda_{w_{\omega,m}}$ … $w_{\omega,m}$ …

$$w_{\omega,m} = \frac{\sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega})\, \mathbf{r}(m|\mathbf{x},\omega)}{\sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega}) \sum_{m\in\mathcal{M}} \mathbf{r}(m|\mathbf{x},\omega)} \tag{12}$$

… $\gamma_{\tilde{\omega},\omega}$ … $\lambda_{\gamma_{\tilde{\omega},\omega}}$ … $\sum_{\tilde{\omega}\in\Omega} \gamma_{\tilde{\omega},\omega} = 1$, …

$$Q_{\gamma_{\tilde{\omega},\omega}} = Q + \lambda_{\gamma_{\tilde{\omega},\omega}}\left(1 - \sum_{\tilde{\omega}\in\Omega} \gamma_{\tilde{\omega},\omega}\right),$$

… $Q_{\gamma_{\tilde{\omega},\omega}}$ … $\gamma_{\tilde{\omega},\omega}$ …

$$\frac{Q_{\gamma_{\tilde{\omega},\omega}}}{\gamma_{\tilde{\omega},\omega}} = \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega})\frac{1}{\gamma_{\tilde{\omega},\omega}} - \lambda_{\gamma_{\tilde{\omega},\omega}} = 0$$
$$\Rightarrow \lambda_{\gamma_{\tilde{\omega},\omega}} \gamma_{\tilde{\omega},\omega} = \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega}),$$
$$\overset{\gamma_{\tilde{\omega},\omega}}{\Rightarrow} \lambda_{\gamma_{\tilde{\omega},\omega}} \sum_{\tilde{\omega}\in\Omega} \gamma_{\tilde{\omega},\omega} = \lambda_{\gamma_{\tilde{\omega},\omega}} = \sum_{\mathbf{x}\in\mathcal{S}} \sum_{\tilde{\omega}\in\Omega} p(\omega|\mathbf{x},\tilde{\omega})$$

… $\lambda_{\gamma_{\tilde{\omega},\omega}}$ … $\gamma_{\tilde{\omega},\omega}$ …

$$\gamma_{\tilde{\omega},\omega} = \frac{\sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega})}{\sum_{\mathbf{x}\in\mathcal{S}} \sum_{\tilde{\omega}\in\Omega} p(\omega|\mathbf{x},\tilde{\omega})}. \tag{13}$$

… λ_{π_ω} … $\sum_{\omega\in\Omega} \pi_\omega = 1$, …

$Q\omega$

$$L' - L = \sum_{\omega \in \Omega} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) \frac{p'(\mathbf{x}|\omega) p'(\omega|\psi(\mathbf{x}))}{p(\mathbf{x}|\omega) p(\omega|\psi(\mathbf{x}))} + \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} \sum_{\omega \in \Omega} q(\omega|\mathbf{x}, \psi(\mathbf{x})) \frac{q(\omega|\mathbf{x}, \psi(\mathbf{x}))}{q'(\omega|\mathbf{x}, \psi(\mathbf{x}))}. \quad (2\ )$$

. (2)

$$I(q, q') = \sum_{\omega \in \Omega} q(\omega|\mathbf{x}, \psi(\mathbf{x})) \frac{q(\omega|\mathbf{x}, \psi(\mathbf{x}))}{q'(\omega|\mathbf{x}, \psi(\mathbf{x}))} \geq 0, \quad (2\ )$$

$$L' - L \geq \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} \sum_{\omega \in \Omega} q(\omega|\mathbf{x}, \psi(\mathbf{x})) \frac{p'(\mathbf{x}|\omega) p'(\omega|\psi(\mathbf{x}))}{p(\mathbf{x}|\omega) p(\omega|\psi(\mathbf{x}))}. \quad (2\ )$$

$$L' - L \geq \sum_{\omega \in \Omega} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) \frac{p'(\mathbf{x}|\omega)}{p(\mathbf{x}|\omega)} + \sum_{\omega \in \Omega} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) \frac{p'(\omega|\psi(\mathbf{x}))}{p(\omega|\psi(\mathbf{x}))} \geq 0. \quad (2\ )$$

$x \in \mathcal{S}$ … $\mathbf{x} \in \mathcal{S}_{\tilde{\omega}}$

$$L' - L = \sum_{\omega \in \Omega} \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) \frac{p'(\mathbf{x}|\omega)}{p(\mathbf{x}|\omega)} + \sum_{\tilde{\omega} \in \Omega} \frac{|\mathcal{S}_{\tilde{\omega}}|}{|\mathcal{S}|} \sum_{\omega \in \Omega} \left[\frac{1}{|\mathcal{S}_{\tilde{\omega}}|} \sum_{\mathbf{x} \in \mathcal{S}_{\tilde{\omega}}} q(\omega|\mathbf{x}, \tilde{\omega}) \right] \frac{p'(\omega|\tilde{\omega})}{p(\omega|\tilde{\omega})}. \quad (2\ )$$

$$p'(\omega|\tilde{\omega}) = \frac{1}{|\mathcal{S}_{\tilde{\omega}}|} \sum_{\mathbf{x} \in \mathcal{S}_{\tilde{\omega}}} q(\omega|\mathbf{x}, \tilde{\omega}), \ \omega \in \Omega, \tilde{\omega} \in \Omega \quad (2\ )$$

$$p'(\cdot|\omega) = \underset{p(\cdot|\omega)}{} \left\{ \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) p(\mathbf{x}|\omega) \right\} \quad \omega \in \Omega, \quad (2\ )$$

(2), (2)

$$\sum_{\omega \in \Omega} p'(\omega|\tilde{\omega}) \frac{p'(\omega|\tilde{\omega})}{p(\omega|\tilde{\omega})} \geq 0, \tilde{\omega} \in \Omega, \quad (\ 0)$$

. (2)

$$\frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) p'(\mathbf{x}|\omega) \geq \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) p(\mathbf{x}|\omega), \omega \in \Omega,$$

$$\frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) \frac{p'(\mathbf{x}|\omega)}{p(\mathbf{x}|\omega)} \geq 0, \omega \in \Omega. \quad (\ 1)$$

(2) … (0), (1), … (2), (2), (2) … [2 , 2].

2.1.1 Gaussian Classes with Noisy Labels

$$p(\mathbf{x}|\omega) = f(\mathbf{x}|\boldsymbol{\mu}_\omega, \boldsymbol{\Sigma}_\omega), \ \omega \in \Omega, \quad (\ 2)$$

. (2)

$$\left\{ \boldsymbol{\mu}'_\omega, \boldsymbol{\Sigma}'_\omega \right\} = \underset{\{\boldsymbol{\mu}_\omega, \boldsymbol{\Sigma}_\omega\}}{} \left[\frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} q(\omega|\mathbf{x}, \psi(\mathbf{x})) f(\mathbf{x}|\boldsymbol{\mu}_\omega, \boldsymbol{\Sigma}_\omega) \right] \quad \omega \in \Omega. \quad (\ )$$

. ( ) … [2]

$F(\mathbf{x}|\boldsymbol{\mu})$ … $\boldsymbol{\mu}$ … $\hat{\boldsymbol{\mu}}$

$$L_{\boldsymbol{\mu}} = \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} f(\mathbf{x}|\boldsymbol{\mu}) \rightarrow \quad \Rightarrow \hat{\mu} = \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x} \in \mathcal{S}} \mathbf{x}. \quad (\ )$$

$N(\mathbf{x})$ … $\mathbf{x} \in \mathcal{X}$ … $\mathcal{S}$ … $q(\mathbf{x})$ … $\mathbf{x} \in \mathcal{S}$

$$q(\mathbf{x}) = \frac{N(\mathbf{x})}{|\mathcal{S}|}, \sum_{\mathbf{x} \in \mathcal{X}} q(\mathbf{x}) = 1, (\mathbf{x} \notin \mathcal{S} \Rightarrow q(\mathbf{x}) = 0),$$

. ( )

$$L_{\boldsymbol{\mu}} = \sum_{\mathbf{x} \in \mathcal{X}} q(\mathbf{x}) F(\mathbf{x}|\boldsymbol{\mu}) \rightarrow \quad \Rightarrow \hat{\mu} = \sum_{\mathbf{x} \in \mathcal{X}} q(\mathbf{x}) \mathbf{x}. \quad (\ 5)$$

. ( ) (5), … (5)

([2]). ( ),

$$\boldsymbol{\mu}'_{\omega} = \frac{1}{\sum_{\mathbf{x}\in\mathcal{S}} q(\omega|\mathbf{x},\psi(\mathbf{x}))} \sum_{\mathbf{x}\in\mathcal{S}} q(\omega|\mathbf{x},\psi(\mathbf{x}))\,\mathbf{x}, \quad \omega \in \Omega \qquad (\;)$$

$$\boldsymbol{\Sigma}'_{\omega} = \frac{1}{\sum_{\mathbf{x}\in\mathcal{S}} q(\omega|\mathbf{x},\psi(\mathbf{x}))} \sum_{\mathbf{x}\in\mathcal{S}} q(\omega|\mathbf{x},\psi(\boldsymbol{x}))\mathbf{x}\mathbf{x}^T - \boldsymbol{\mu}'_{\omega}\boldsymbol{\mu}'^{T}_{\omega}, \quad \omega \in \Omega. \qquad (\;)$$

([1]), ([2]), ( ), ( ).

2.1.2 Class-conditional Gaussian Mixtures with Noisy Labels

([2]) $p(\mathbf{x}|\omega)$

$$p(\mathbf{x}|\omega) = \sum_{m\in\mathcal{M}_{\omega}} w_{m\omega} f(\mathbf{x}|\boldsymbol{\mu}_{m\omega},\boldsymbol{\Sigma}_{m\omega}), \qquad \sum_{m\in\mathcal{M}_{\omega}} w_{m\omega} = 1, \; \omega \in \Omega \qquad (\;)$$

$m \in \mathcal{M}_{\omega}$ $P(\mathbf{x}|\omega)$. ( ) ([1]),

$$L = \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathcal{S}} \left[\sum_{\omega\in\Omega} \sum_{m\in\mathcal{M}_{\omega}} p(\omega|\psi(\mathbf{x}))\, w_{m\omega} f(\mathbf{x}|\boldsymbol{\mu}_{m\omega},\boldsymbol{\Sigma}_{m\omega}) \right].$$

$$h(m,\omega|\mathbf{x},\psi(\mathbf{x})) = \frac{p(\omega|\psi(\mathbf{x})) w_{m\omega} f(\mathbf{x}|\boldsymbol{\mu}_{m\omega},\boldsymbol{\Sigma}_{m\omega})}{\sum_{\omega\in\Omega}\sum_{m\in\mathcal{M}_{\omega}} p(\omega|\psi(\mathbf{x}))\, w_{m\omega} f(\mathbf{x}|\boldsymbol{\mu}_{m\omega}}$$

$p'(\omega|\tilde{\omega})$, $w'_{m\omega}$

$$p'(\omega|\tilde{\omega}) = \frac{1}{|\mathcal{S}_{\tilde{\omega}}|} \sum_{\mathbf{x}\in\mathcal{S}_{\tilde{\omega}}} \sum_{m\in\mathcal{M}_{\omega}} h(m,\omega|\mathbf{x},\tilde{\omega}), \omega \in \Omega,\ \tilde{\omega} \in \Omega.$$

$w'_{m\omega}$

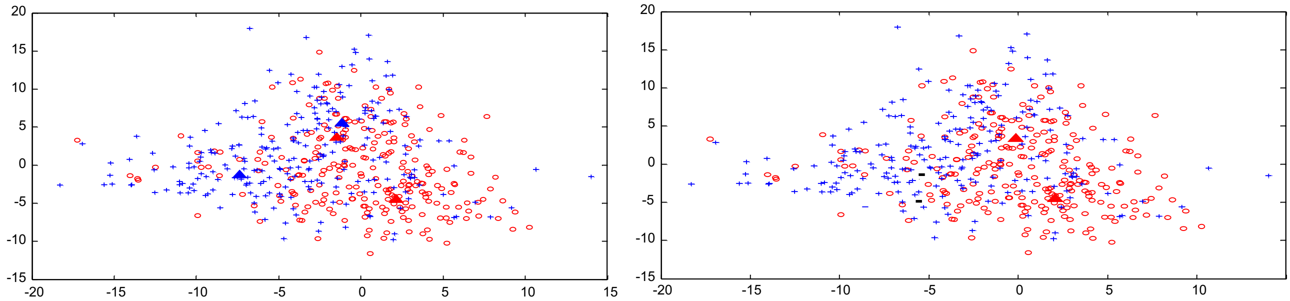
$$w'_{m\omega} = \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathcal{S}} h(m|\omega,\mathbf{x},\psi(\mathbf{x})) = \frac{1}{|\mathcal{S}|} \sum_{\mathbf{x}\in\mathcal{S}} \frac{h(m,\omega|\mathbf{x},\psi(\mathbf{x}))}{\sum_{m\in\mathcal{M}_{\omega}} h(m,\omega|\mathbf{x},\psi(\mathbf{x}))} \quad m \in \mathcal{M}_{\omega}, \omega \in \Omega,$$

3 Related work

Table 1

1	2000	0	5
2	1000	2	2
	10		
	50		
	1	1	
	2	22	2
	50	1	2
	5 000	21	

[illegible]



4 Experiments and discussion

4.1 Datasets and preprocessing

4.2 Results and discussion

1

20%

12. 0%

12. 0%

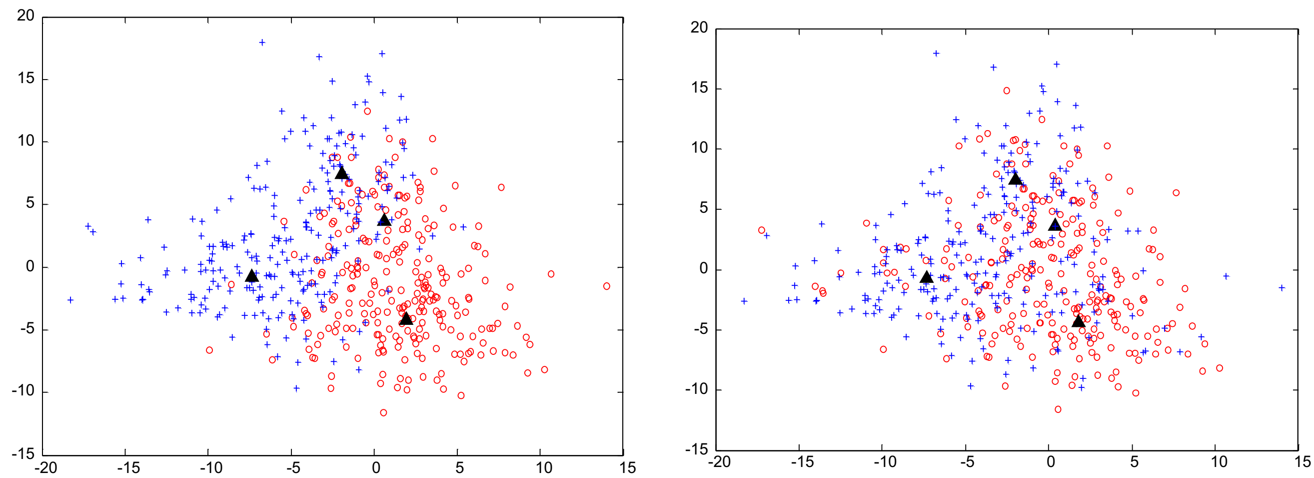


Fig. 5 2 r r b b r

Table 2 2 r

	r b b	r b b	rr r (%)
r 2	1 2, $e-0^5$ .0 − 0 1	[0, 0^5 00]	1 .00
2 20% b rr r	0. 2^5 0.21 ^{5} 0.1 0. 02	[0^5 101 0,]	12. 0

r r r r r . r r r A A , r r A r r r 20 0% r r r r r . r r bb r rr r r A r b b r r r b 5 0% r r 5 0% 5 r r 2 r r b b r r 20% b rr r 2 r b b r . b 2 r r rr r r b b r r 2 2 20% b rr r . $\gamma_{ij}(i=j)$ r b b b r . b 2 r b b r 2 r r 1, r 20% b rr r r 0. , r

r b b 0^5 . r b b r 0^5 . r r r . rr r r r 1 .00% 12. 0% r , r r . r b r r b r r r r . r r r r - r r r A r r r b 5 . r r b . r r rr r r b . A b - r b , b b rr r rr r r . r r r r r r . r b r r r rr r r r r b r r . A , r , r r b r b r . r , . r r r r r - 1 r r . r

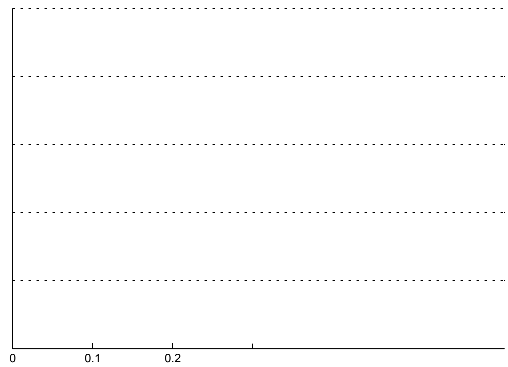

0
0.1
0.2

Table 3 r mr r r -
r r r b r

		r b r					
		0	0.1	0.2	0.	0,	0.5
1	A	0.0	0.05	0.0 1	0.0	0.11	0.22
	r A	0.0 5	0.0	0.0	0.0 5	0.0 5	0.0 5
	A	**0.005**	**0.003**	**0.003**	**0.003**	**0.030**	**0.003**
2	r	0.1 2	0.1	0.1	0.1	0.5 0	0.5 0
	r	0.1	0.1	0.1	0.1	**0.186**	0.5 0
	A	0.1 2	0.1	0.1 2	0.1	0.22	0. 1
	r A	0.1	0.1	0.1	0.2	0.5 2	0.5 2
	A	**0.128**	**0.120**	**0.126**	**0.166**	0.1	**0.182**
	A	0.	0.	0. 1	0.5 2	0.5 1	0. 1
r	r A	**0.305**	0, 0	**0.357**	0.5	0.5 0	0.5 2
	A	0.	**0.238**	**0.357**	**0.357**	**0.500**	**0.504**
r	A	0.02	0.05	0.1	0.200	0.	0.2
	r A	0.02	0.0	**0.033**	0.05	0.100	0.08
	A	**0.013**	**0.016**	**0.033**	**0.05**	**0.083**	**0.08**
	A	0.0	0.12	0.112	0.15	0.2	0. 0
	r A	0.0	**0.011**	**0.044**	0.0	**0.044**	0.0
	A	**0.033**	0.022	**0.044**	**0.033**	0.0 5	**0.076**
r	r	0.5	0. 5	0. 1	0.	0.	0.5
	r	0.5	0. 5	0. 2	0.	0. 5	0.
	A	0.2 0	**0.261**	0.2	0.5	0. 2	0.5
	r A	0. 0	0. 2	0. 0	0.	0.	0.
	A	**0.248**	**0.261**	**0.261**	**0.289**	**0.287**	**0.271**
	r	**0.130**	0.1	0.12	0.1	0.	0.5 1
	r	**0.130**	**0.183**	**0.113**	0.1	0. 1	0.
	A	0.1	0.15	0.1	**0.183**	0.2 2	0.5
	r A	0.2	0.1	0.1	0.1	0, 5	0. 1
	A	0.205	**0.183**	0.1	**0.183**	**0.188**	**0.185**
	A	**0.188**	**0.201**	**0.223**	**0.223**	0.2 0	0. 12
r	r A	0.2	0.2	0.2	0.25 1	0.25 2	0.25
	A	0.222	0.22	0.2 1	0.2 2	**0.244**	**0.244**

b r r r b

Table 4 r mr r r -
r r r b r

		A r b r					
		0	0.1	0.2	0.	0,	0.5
1	A	0.0	0.0	0.0 5	0.1	0.211	0.221
	r A	0.0 5	0.0 5	0.0	0.0	0.0 5	0.
	A	**0.005**	**0.005**	**0.005**	**0.005**	**0.026**	**0.191**
2	r	0.1 2	0.1	0.1	0.1	0.1 2	–
	r	0.1	0.1	0.1 2	0.1	0.2	–
	A	0.1 2	0.1 2	0.2	0,	0,	–
	r A	0.1	**0.128**	0.1	0.1	0.1	**0.504**
	A	**0.128**	**0.128**	**0.132**	**0.13**	**0.144**	0.5 2
r	A	0.	0.	0. 5	0.	0, 5 2	0.5 0
	r A	**0.305**	**0.238**	0.2 5	0,	0.5 2	0.5
	A	0.	0.2 1	**0.214**	**0.25**	**0.404**	**0.415**
r	A	0.02	0.01	0.1	0.1	0,	–
	r A	0.02	**0.016**	**0.016**	**0.022**	**0.033**	–
	A	**0.013**	**0.016**	**0.016**	**0.022**	**0.033**	–
	A	0.0	0.1 0	0.112	0.1	0, 0	–
	r A	0.0	0.05	0.05	**0.042**	**0.056**	–
	A	**0.033**	**0.042**	**0.042**	**0.042**	**0.056**	–
r	r	0.5	0,	0.	0. 55	0. 00	–
	r	0.5	0, 20	0, 11	0.	0. 2	–
	A	0.2 0	0.2	0. 0	0. 2	0. 1	–
	r A	0. 0	0.	0.2 1	0.	0. 2	–
	A	**0.248**	**0.261**	**0.261**	**0.261**	**0.327**	–
	r	**0.130**	**0.163**	**0.173**	**0.178**	0.1	0.
	r	**0.130**	**0.163**	0.1	0.1	0.1	0.
	A	0.1	0.1	0.2 2	0, 5	0, 5	0, 2
	r A	0.2	0.1	0.1	0.1	0, 5	0.2
	A	0.205	0.1	0.1	0.1	**0.163**	**0.221**
	A	**0.188**	**0.196**	0.2	0.	0. 0	–
r	r A	0.2	0.2	0.2 1	0,	0.5 00	–
	A	0.222	0.22	**0.226**	**0.280**	**0.296**	–

b r r r b

r r r - r 1000 , r , r 5 0. , 20% b rr r r r .

r r r r r r . r r r , r rr r b . r r r b r r .

r br

br . 1000 r r , br r 1 5 r . r r , r r rr r , r b r r r .

A r . b 5, r b r r r - - . r r r br . r r r 1000

Table 5

1	A	0/0/	0/0/
	r A	0/0/	0/0/
	A	/0/0	/0/0
2	r	0/0/	0/0/
	r	1/0/5	0/0/
	A	0/0/	0/0/
	r A	0/0/	1/1/
	A	5/0/1	/1/1
r	A	0/0/	0/0/
	r A	1/1/	2/0/
	A	/1/1	/0/2
r	A	0/0/	0/0/5
	r A	0/ /	0/ /1
	A	/ /0	1/ /0
	A	0/0/	0/0/5
	r A	1/2/	0/2/
	A	/2/1	/2/0
r	r	0/0/	0/0/5
	r	0/0/	0/0/5
	A	0/1/5	0/0/5
	r A	0/0/	0/0/5
	A	5/1/0	5/0/0
	r	1/1/	1/2/
	r	1/1/	1/2/
	A	0/1/5	0/0/
	r A	0/0/	0/0/
	A	2/1/	2/0/
r	A	/0/2	2/0/
	r A	0/0/	0/0/5
	A	2/0/	/0/2

... 10, ... 0.5 0 0 0, 0, ... 1.

5 Experimental results on large-scale datasets

... 5,000 ... 200 ... 1 5 ... 0 ... 2 5 0, ... 0%, 10%, 20%, 0%, 0%, 5 0% ... 5, 10, ... (1), (1), (1) ... (2), (2), (2) ... 11 ... 11, (1), (1), (1) ... (2), (2), (2) ... 11, ...

( ) ...

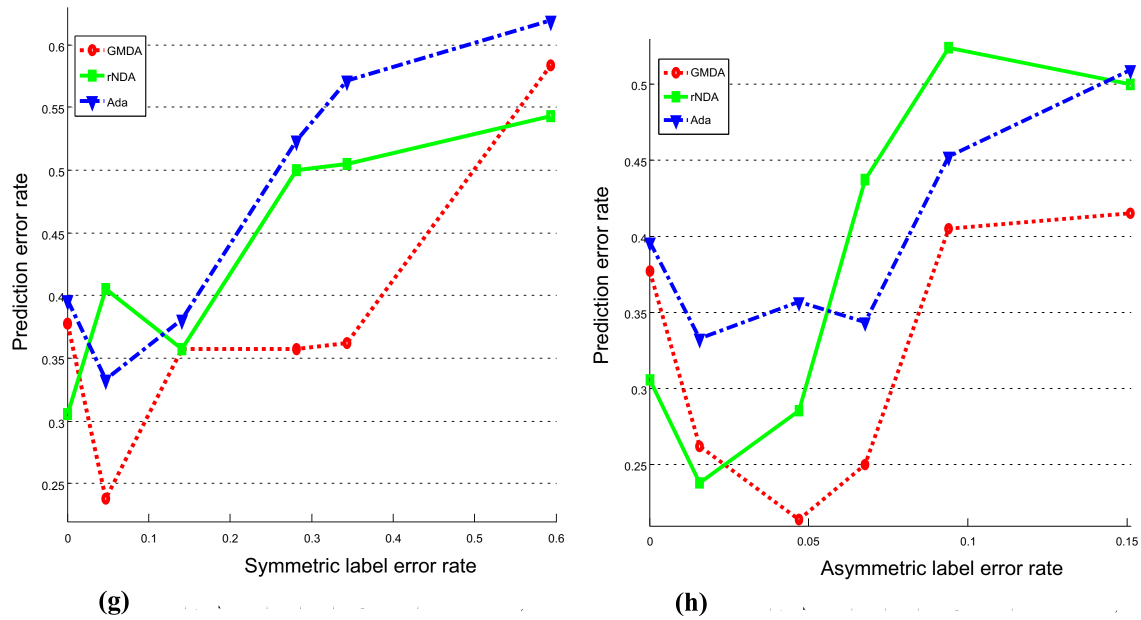


Fig. 6

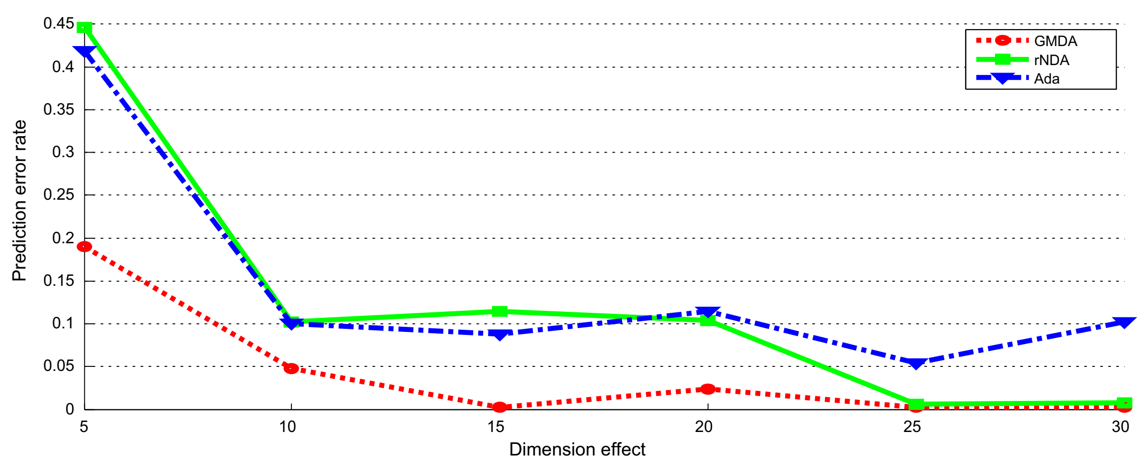


Fig. 7

Fig. 8

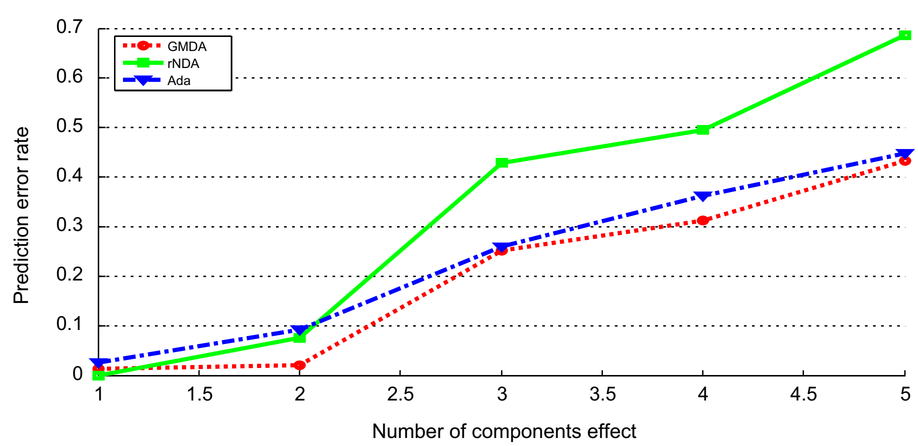


Fig. 9

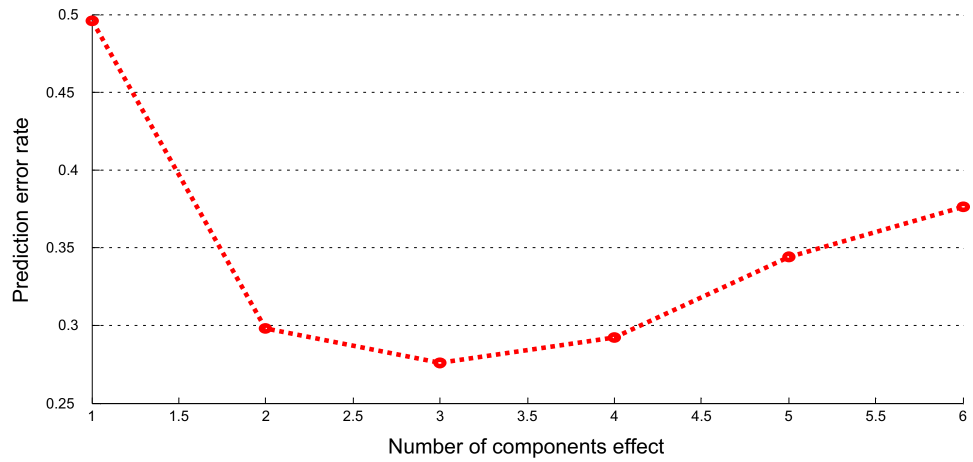


Fig. 10

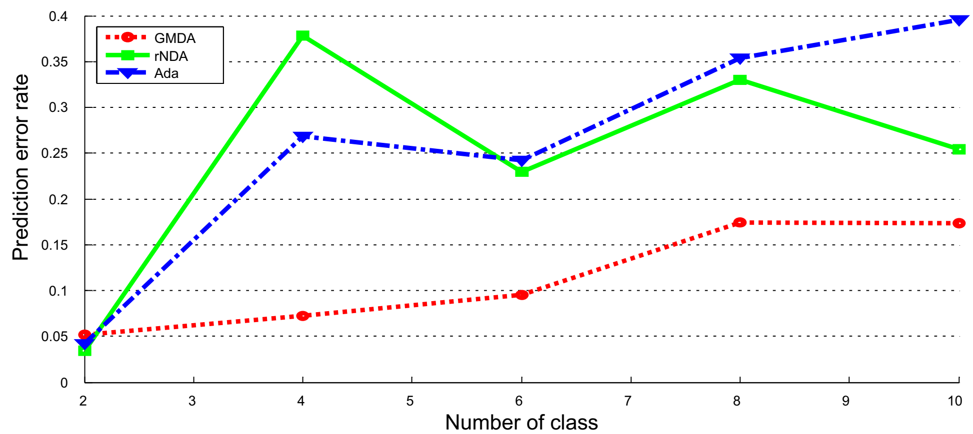


Table 6

	N	D			1	2
1	5,000	200	250	5	5 2	0
2	5,000	200	0	5	1	1
	5,000	200	250	10		5
	5,000	200	0	10	1	1
5	5,000	200	250		0	2
	5,000	200	0			20

(b) ... 250, ... 50%. ... 1%. ... () ... 0,

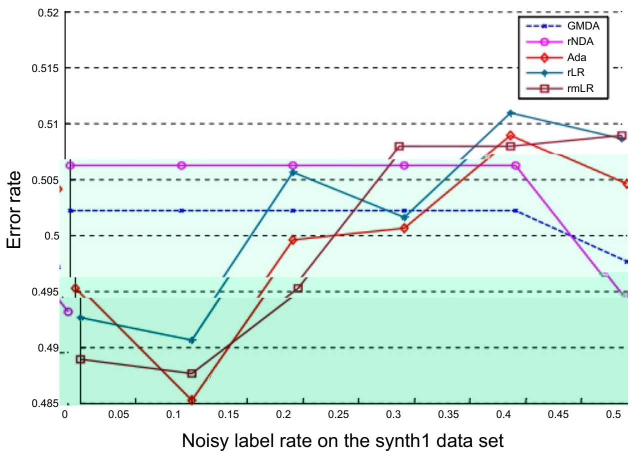
Table 7 [illegible]

[illegible]	[illegible]					
	0.0	0.1	0.2	0.[illegible]	0,[illegible]	0[illegible]
(a1) error rate with 5-cross-validation on correlated synth1 dataset						
[illegible]	[illegible]	[illegible]	[illegible]	**0.5023**	**0.5023**	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**0.4937**
[illegible]	[illegible]	**0.4853**	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	**0.4890**	[illegible]	**0.4953**	[illegible]	[illegible]	[illegible]
(b1) error rate with 10-cross-validation on correlated synth3 dataset						
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**0.4927**
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	**0.4520**	**0.4520**	**0.4520**	**0.4520**	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**0.4827**	[illegible]
(c1) error rate with 3-cross-validation on correlated synth5 dataset						
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**0.4938**
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	**0.4812**	**0.4812**	**0.4812**	**0.4812**	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**0.4904**	[illegible]
(a2) error rate with 5-cross-validation on uncorrelated synth2 dataset						
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	**0.4427**	**0.4427**	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	**0.1653**	**0.1103**	**0.3853**	[illegible]	[illegible]	**0.1653**
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
(b2) error rate with 10-cross-validation on uncorrelated synth4 dataset						
[illegible]	**0.0420**	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**0.4560**
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	**0.2313**	**0.2320**	**0.2133**	**0.4267**	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
(c2) error rate with 3-cross-validation on uncorrelated synth6 dataset						
[illegible]	**0.0424**	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	**0.4422**
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	**0.1754**	**0.1562**	**0.1902**	**0.4370**	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

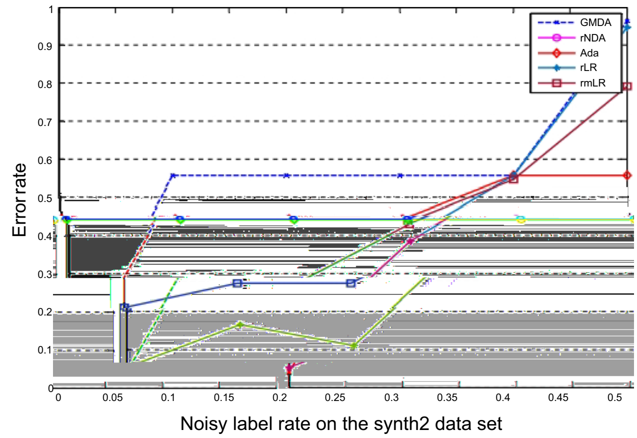
[illegible]

[illegible] [0.2, 0, [illegible]

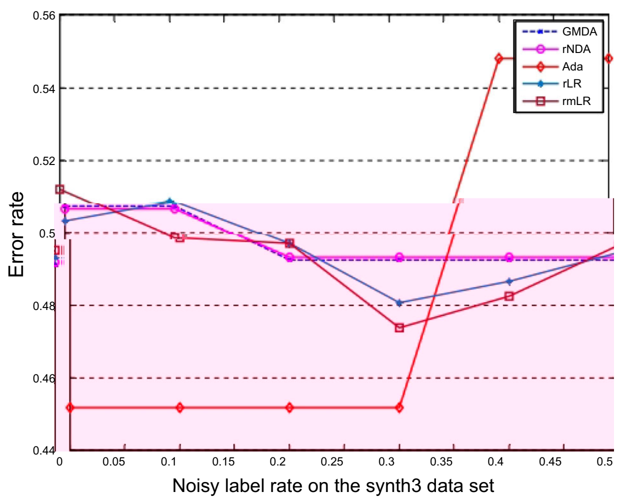
[illegible]


GMDA
rNDA
Ada
rLR
rmLR
Error rate
Noisy label rate on the synth1 data set

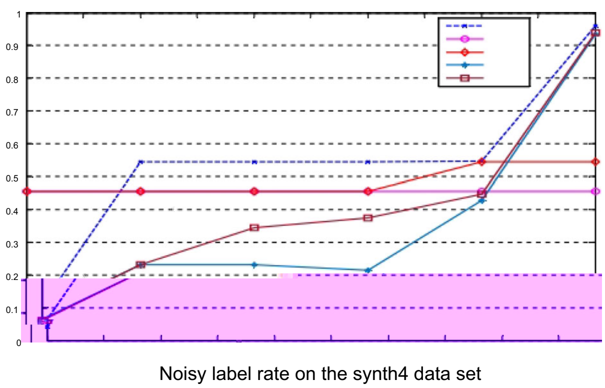
(a1) 1

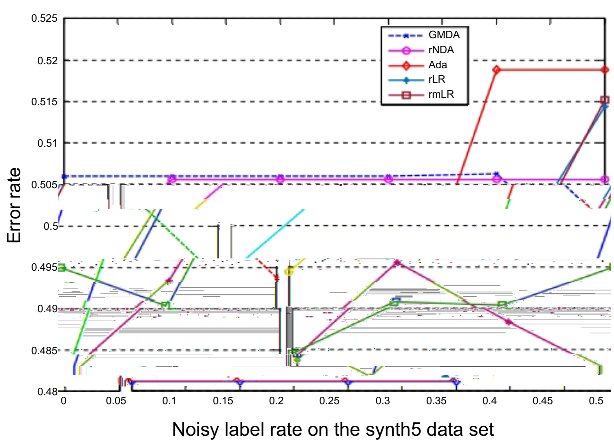

GMDA
rNDA
Ada
rLR
rmLR
Error rate
Noisy label rate on the synth2 data set

(a2)

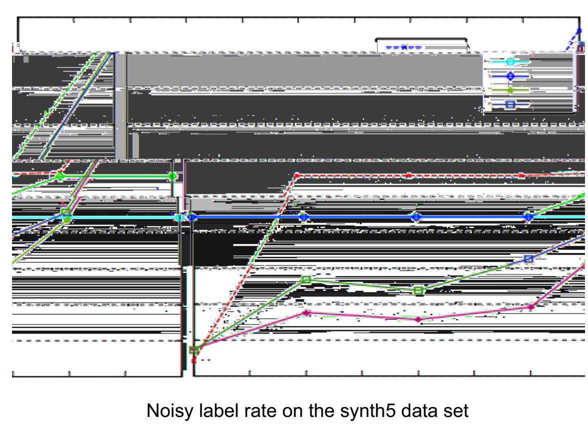

GMDA
rNDA
Ada
rLR
rmLR
Error rate
Noisy label rate on the synth3 data set

(b1)


Noisy label rate on the synth4 data set


GMDA
rNDA
Ada
rLR
rmLR
Error rate
Noisy label rate on the synth5 data set

(c1)


Noisy label rate on the synth5 data set

6 Conclusion

Acknowledgements

Compliance with ethical standards

Conflict of interest

References

[illegible] (201[illegible]) [illegible]

[illegible] (2020) [illegible] 2020

[illegible] (201[illegible]) [illegible], 201[illegible]

[illegible] (2012) [illegible] 2012, [illegible] 2012

[illegible] (2012) [illegible]

[illegible] (1[illegible]) [illegible]

[illegible] (200[illegible]) [illegible]

[illegible] (2012) [illegible] 2012, [illegible]

[illegible] (1[illegible]) [illegible]

[illegible] (201[illegible]) [illegible]

[illegible] (200[illegible]) [illegible]

[illegible] (2010) [illegible]

[illegible] (200[illegible]) [illegible] 200[illegible]), [illegible]

[illegible] (201[illegible]) [illegible]

[illegible] (20[illegible]) [illegible]

[illegible] (20[illegible]) [illegible]

[illegible] (20[illegible]) [illegible]

[illegible] (201[illegible]) [illegible]

[illegible] (200[illegible]) [illegible]

[illegible] (201[illegible]) [illegible]

Publisher's Note [illegible]