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

1.

2.

1.1.1 Description of the problem with the noise labels

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

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

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

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

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

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

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

$\mathcal{S}_\omega.$

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

L_ω

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

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

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

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

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

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

$p(\omega)$ $\omega.$

$p(\omega|\tilde{\omega})$ ω

$\tilde{\omega}.$

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

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

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

$\omega \in \Omega$

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

$\tilde{\omega} \in \Omega$

$\mathbf{f}$.

1.1.2 Gaussian mixture model

$\mathbf{f}$

, , 22 $\mathbf{x}$

,

$\mathbf{f}$.

$\mathbf{x}$

$\mathbf{f}$.

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

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

λ_{G_1} ( ), $q(\omega)$

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

,

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

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

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

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

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

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

, $h_{\omega,m}$

. . λ_{G_2}

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

. λ_{G_2} $h_{\omega,m}$

$$\begin{aligned} h_{\omega,m} &= \frac{w_{\omega,m} g\left(\mathbf{x}\ \omega, \boldsymbol{\mu}_{\omega,m}, \boldsymbol{\Sigma}_{\omega,m}\right)}{\sum_{m\in\mathcal{M}} w_{\omega,m} g\left(\mathbf{x}\ \omega, \boldsymbol{\mu}_{\omega,m}, \boldsymbol{\Sigma}_{\omega,m}\right)} \\ &= \quad (m|\mathbf{x},\omega) \end{aligned} \quad ( )$$

, . ( ) ( ) ( )

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

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

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

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

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

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

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

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

f $w_{\omega,m}$. $\lambda_{w_{\omega,m}}$

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

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

f $Q_{\lambda_{w_{\omega,m}}}$ . . $w_{\omega,m}$

, f f

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

$$\Rightarrow \lambda_{w_{\omega,m}} w_{\omega,m} = \sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega}) \quad (m|\mathbf{x},\omega)$$

$$\Rightarrow^{w_{\omega,m}} \lambda_{w_{\omega,m}} \sum_{m\in\mathcal{M}} w_{\omega,m} = \lambda_{w_{\omega,m}} = \sum_{\tilde{\omega}\in\Omega} I(\tilde{\omega}=\omega) \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega}) \sum_{m\in\mathcal{M}} \quad (m|\mathbf{x},\omega),$$

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

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

f $\gamma_{\tilde{\omega},\omega}$. $\lambda_{\gamma_{\tilde{\omega},\omega}}$

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

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

f $Q_{\gamma_{\tilde{\omega},\omega}}$ . . $\gamma_{\tilde{\omega},\omega}$

, f f

$$\frac{Q_{\gamma_{\tilde{\omega},\omega}}}{\gamma_{\tilde{\omega},\omega}} = \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega}) \frac{1}{\gamma_{\tilde{\omega},\omega}} - \lambda_{\gamma_{\tilde{\omega},\omega}} =$$ [redacted]

$$\Rightarrow \lambda_{\gamma_{\tilde{\omega},\omega}} \gamma_{\tilde{\omega},\omega} = \sum_{\mathbf{x}\in\mathcal{S}} p(\omega|\mathbf{x},\tilde{\omega})$$

$$\Rightarrow^{\gamma_{\tilde{\omega},\omega}} \lambda_{\gamma_{\tilde{\omega},\omega}} \sum_{\tilde{\omega}\in\Omega} \gamma_{\tilde{\omega},\omega} = \lambda_{\gamma}$$ [redacted]

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

. $\lambda_{\gamma_{\tilde{\omega},\omega}}$ [redacted] f

$$\gamma_{\tilde{\omega},\omega} = \frac{\sum_{\mathbf{x}\in\mathcal{S}} p(\omega}{\sum_{\mathbf{x}\in\mathcal{S}} \sum}$$ [redacted] (1)

. λ_{π_ω}

$\sum_{\omega\in\Omega} \pi_\omega$ [redacted] f

f

Q [redacted]

$= Q + \lambda$

Input: $\Theta=\{\theta_\omega\}_{\omega\in\Omega}$,

Initial $\boldsymbol{\mu}_{\omega,m}, \boldsymbol{\Sigma}$

(1)

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

$\pi(\tilde{\omega}) = |\mathcal{S}_{\tilde{\omega}}|/|\mathcal{S}|, \ \tilde{\omega} \in \Omega$

$\pi(\psi)].$

$p(\mathbf{x}|\omega)p(\omega|\psi$ (20)

$+ \ \pi(\tilde{\omega})$,

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

. (2) - -

K . f

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

f ,

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

. , f f f

,

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

, . $\boldsymbol{x} \in \mathcal{S}$

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

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

, f . f

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

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

f

. , . . (2), . (2)

f K .

f .

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

. (2) .

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

f .

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

, (2) f (0), (1), . f (2), (2), (2) . f 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) f

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

x x . ( ) f . f , f 2

x x f

. x . . ,

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

x . $\hat{\mu}$

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

f $N(\mathbf{x})$. f . f $\mathbf{x} \in \mathcal{X}$

$\mathcal{S}$ $q(\mathbf{x})$ f . f $\mathbf{x} \in \mathcal{S}$

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

. ( ) . f

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

f . ( ) ( ), f

( ) x

x (f

f 2). , f ( ),

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

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

f

f

. (1), (2), ( ), ( ).

2.1.2 Class-conditional Gaussian Mixtures with Noisy Labels

(2)

x x

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

x . ,

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

$m \in \mathcal{M}_{\omega}$ x f

- x $P(\mathbf{x}|\omega)$. ( )

(1), - f f -

f

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

f

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

(1), f x f . , $p'(\omega|\tilde{\omega})$, $w'_{m\omega}$ ( ) f

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

( ) f $w'_{m\omega}$

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

f f x f . ( ), ( ), ( ), ( ), ( ).

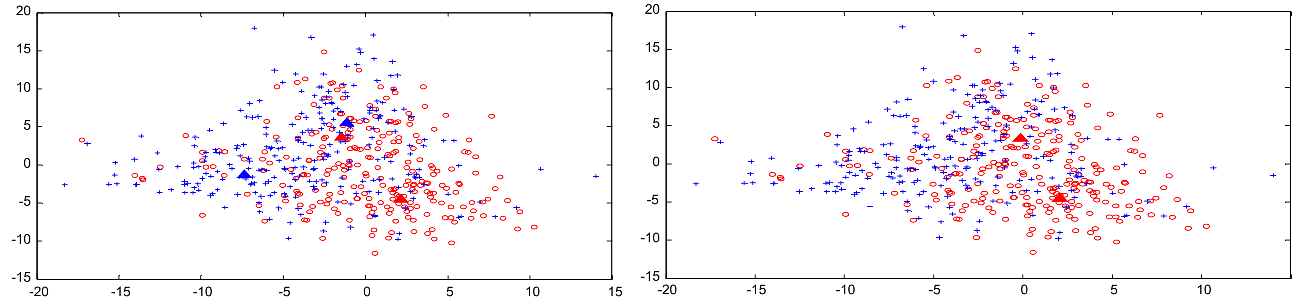
3 Related work

f f , , , , , . , . f ff f x f 11, 0 . , f x x . 11 f f x f , f x f . , f x . f x f x f x 0 . 1 , f f f ,

f f . 1 K x f f x f , , f . 12 x . f f x f f 2 , , f f , 1 . f .

Table 1 f

1	2000	0	
2	1000	2	2
	10		
	1 0		
	1	1	
	2	22	2
	0	1	2
f	000	21	



4 Experiments and discussion

C C

x . 1 f .

f

.

, , , f , ,

. .

. .

.

,

.

, f f

f . 1

20% . f

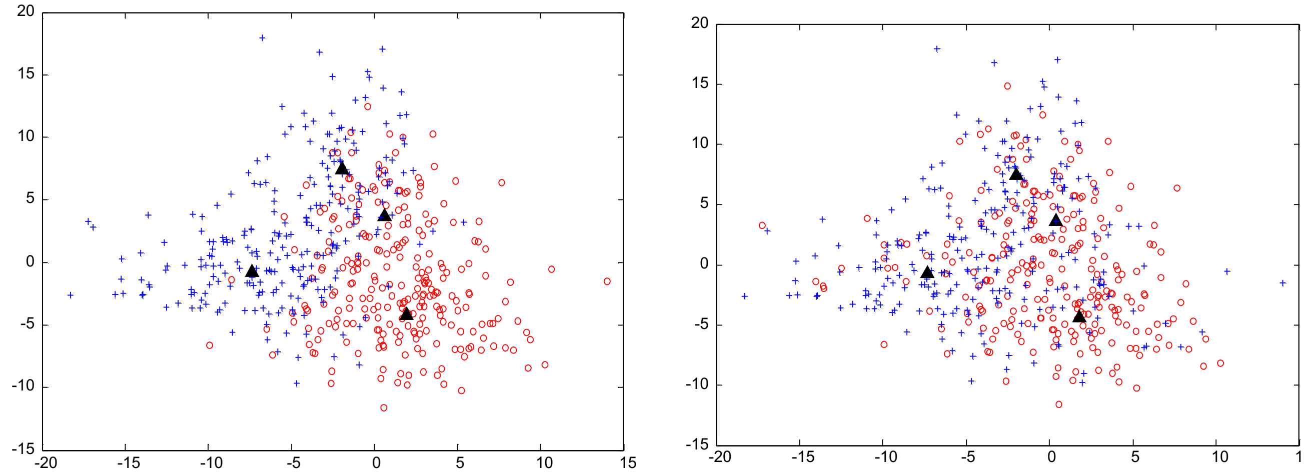
f . 1.

. 2, f f

12. 0% 12. 0% f .

f ,

f .



(a) 2 (b) 2 20%

Fig. 5 2 x

Table 2 2

			(%)
2	1 2. $e-0$ .0 -0 1	[0. 0. 00]	1 .00
2 20%	0. 2 0.21 0.1 0. 02	[0. 101 0.]	12. 0

ff . f x f 0. . ff . . 0. . f , , , . x 1 .00% 12. 0%, , ff , . , f 20 0% ff f f f f . . ff f f f x f ff f . f , , , . 0% . , f x , . 0% . , , ff 2 x . 20% 2 f . . 2 f f f 2 2 . 20% . $\gamma_{ij}(i = j)$ f f . 2 f , . 2 x 1, f f 20% 0. , 1 .

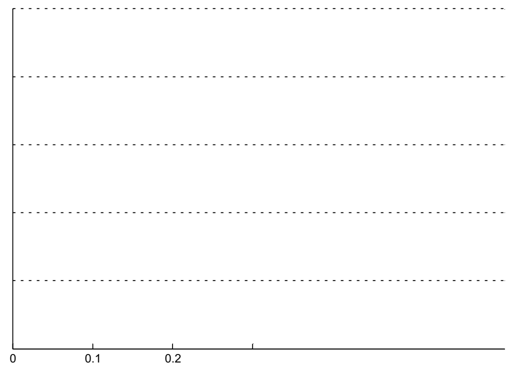

0
0.1
0.2

Table 3 x -
ff

	0	0.1	0.2	0.	0.	0.
1	0.0	0.0	0.0 1	0.0	0.11	0.22
	0.0	0.0	0.0	0.0	0.0	0.0
	0.005	**0.003**	**0.003**	**0.003**	**0.030**	**0.003**
2	0.1 2	0.1	0.1	0.1	0. 0	0. 0
	0.1	0.1	0.1	0.1	**0.186**	0. 0
	0.1 2	0.1	0.1 2	0.1	0.22	0. 1
	0.1	0.1	0.1	0.2	0. 2	0. 2
	0.128	**0.120**	**0.126**	**0.166**	0.1	**0.182**
	0.	0.	0. 1	0. 2	0. 1	0. 1
	0.305	0. 0	**0.357**	0.	0. 0	0. 2
	0.	**0.238**	**0.357**	**0.357**	**0.500**	**0.504**
	0.02	0.0	0.1	0.200	0.	0.2
	0.02	0.0	**0.033**	0.0	0.100	**0.08**
	0.013	**0.016**	**0.033**	**0.05**	**0.083**	**0.08**
	0.0	0.12	0.112	0.1	0.2	0. 0
	0.0	**0.011**	**0.044**	0.0	**0.044**	0.0
	0.033	0.022	**0.044**	**0.033**	0.0	**0.076**
	0.	0.	0. 1	0.	0.	0.
	0.	0.	0. 2	0.	0.	0.
	0.2 0	**0.261**	0.2	0.	0. 2	0.
	0. 0	0. 2	0. 0	0.	0.	0.
	0.248	**0.261**	**0.261**	**0.289**	**0.287**	**0.271**
	0.130	0.1	0.12	0.1	0.	0. 1
	0.130	**0.183**	**0.113**	0.1	0. 1	0.
	0.1	0.1	0.1	**0.183**	0.2 2	0.
	0.2	0.1	0.1	0.1	0.	0. 1
	0.20	**0.183**	0.1	**0.183**	**0.188**	**0.185**
	0.188	**0.201**	**0.223**	**0.223**	0.2 0	0. 12
f	0.2	0.2	0.2	0.2 1	0.2 2	0.2
	0.222	0.22	0.2 1	0.2 2	**0.244**	**0.244**

f

Table 4 x -
ff

	0	0.1	0.2	0.	0.	0.
1	0.0	0.0	0.0	0.1	0.211	0.221
	0.0	0.0	0.0	0.0	0.0	0.
	0.005	**0.005**	**0.005**	**0.005**	**0.026**	**0.191**
2	0.1 2	0.1	0.1	0.1	0.1 2	–
	0.1	0.1	0.1 2	0.1	0.2	–
	0.1 2	0.1 2	0.2	0.	0.	–
	0.1	**0.128**	0.1	0.1	0.1	**0.504**
	0.128	**0.128**	**0.132**	**0.13**	**0.144**	0. 2
	0.	0.	0.	0.	0. 2	0. 0
	0.305	**0.238**	0.2	0.	0. 2	0.
	0.	0.2 1	**0.214**	**0.25**	**0.404**	**0.415**
	0.02	0.01	0.1	0.1	0.	–
	0.02	**0.016**	**0.016**	**0.022**	**0.033**	–
	0.013	**0.016**	**0.016**	**0.022**	**0.033**	–
	0.0	0.1 0	0.112	0.1	0. 0	–
	0.0	0.0	0.0	**0.042**	**0.056**	–
	0.033	**0.042**	**0.042**	**0.042**	**0.056**	–
	0.	0.	0.	0.	0. 00	–
	0.	0. 20	0. 11	0.	0. 2	–
	0.2 0	0.2	0. 0	0. 2	0. 1	–
	0. 0	0.	0.2 1	0.	0. 2	–
	0.248	**0.261**	**0.261**	**0.261**	**0.327**	–
	0.130	**0.163**	**0.173**	**0.178**	0.1	0.
	0.130	**0.163**	0.1	0.1	0.1	0.
	0.1	0.1	0.2 2	0. 1	0.	0. 2
	0.2	0.1	0.1	0.1	0.	0.2
	0.20	0.1	0.1	0.1	**0.163**	**0.221**
	0.188	**0.196**	0.2	0.	0. 0	–
f	0.2	0.2	0.2 1	0.	0. 00	–
	0.222	0.22	**0.226**	**0.280**	**0.296**	–

f

x ff f -
f . f
f 1000 , , f 0.
, 20%
.
ff
ff ff .
f , .
f
.
f ff
f x f f

f .
1000
, . f 1 f .
, . ,
f
.
. , , , .
f -
x
- ff f . f .
/ 1000

Table 5 / /

	/ /	
1	0/0/	0/0/
	0/0/	0/0/
	/0/0	/0/0
2	0/0/	0/0/
	1/0/	0/0/
	0/0/	0/0/
	0/0/	1/1/
	/0/1	/1/1
	0/0/	0/0/
	1/1/	2/0/
	/1/1	/0/2
	0/0/	0/0/
	0/ /	0/ /1
	/ /0	1/ /0
	0/0/	0/0/
	1/2/	0/2/
	/2/1	/2/0
	0/0/	0/0/
	0/0/	0/0/
	0/1/	0/0/
	0/0/	0/0/
	/1/0	/0/0
	1/1/	1/2/
	1/1/	1/2/
	0/1/	0/0/
	0/0/	0/0/
	2/1/	2/0/
f	/0/2	2/0/
	0/0/	0/0/
	2/0/	/0/2

ff x f x x f f 1 10, x

f 0. 0 0 0. 0, f x f x f 1.

/ ff f ff / f f x / f x f f x

5 Experimental results on large-scale datasets

f f f x ff f x f f f 1 ,000 200 f 1 f x f f x f x f f f 0 2 0, f f x f f f f f 0%, 10%, 20%, 0%, 0%, 0% f f , 10, f , (1), (1), (1) (2), (2), (2) 11 f f / . 11, (1), (1), (1) f (2), (2), (2) f . 11, ( ) f f

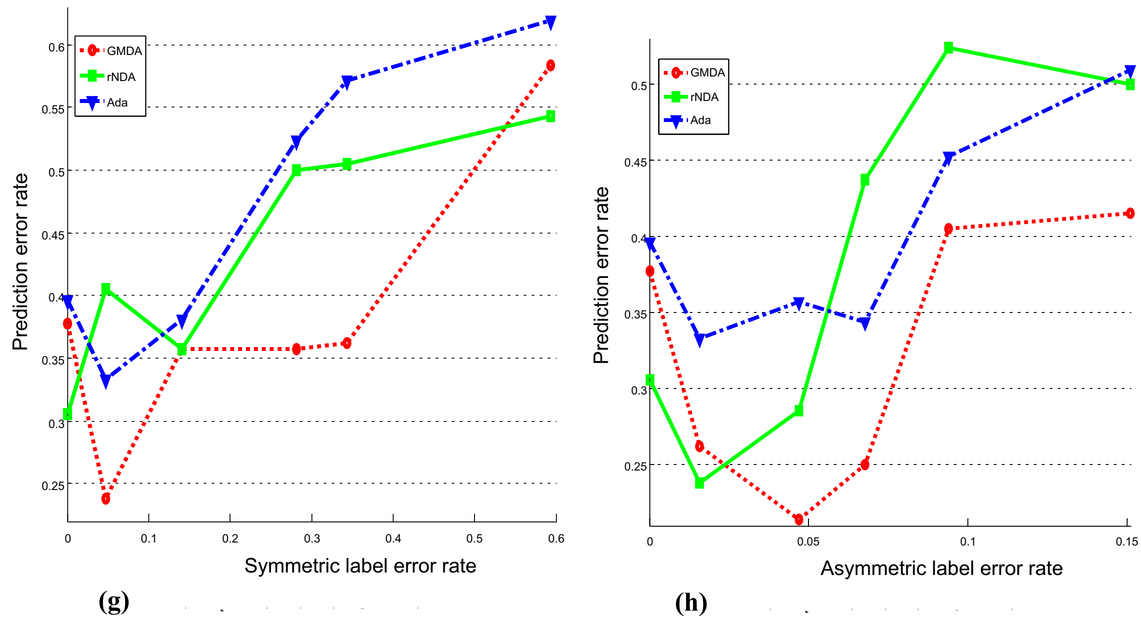


Fig. 6

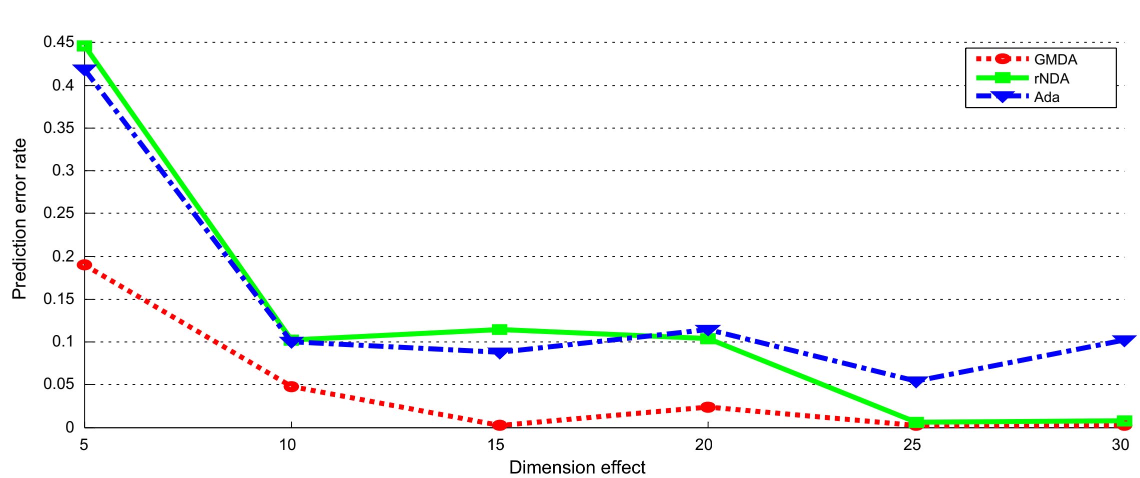


Fig. 7 ff f f

Fig. 8 ff f f f f x ( )

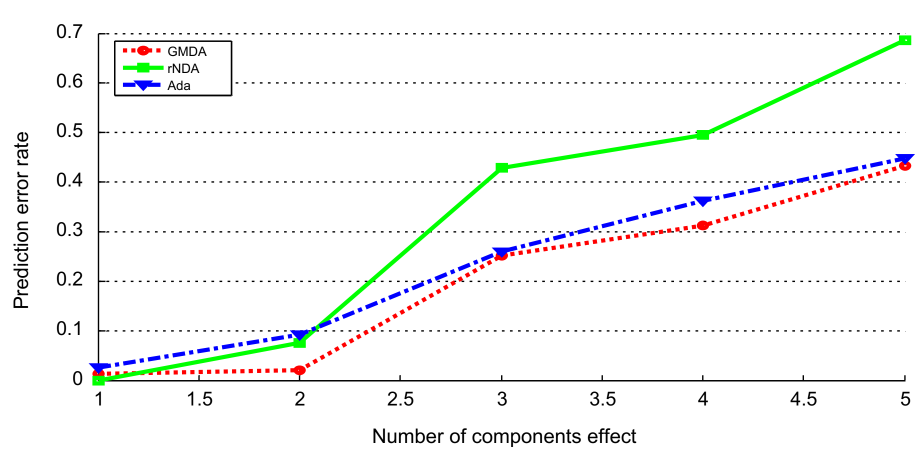


Fig. 9 ff f . f

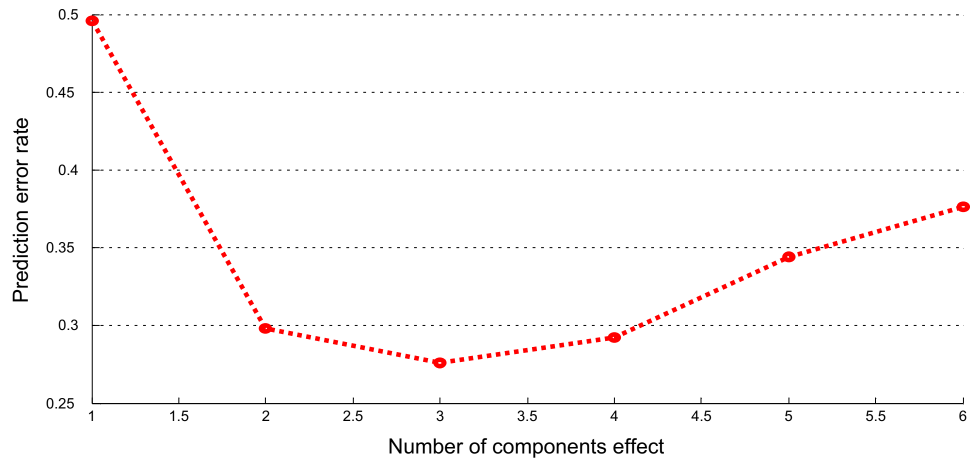


Fig. 10 ff f . f

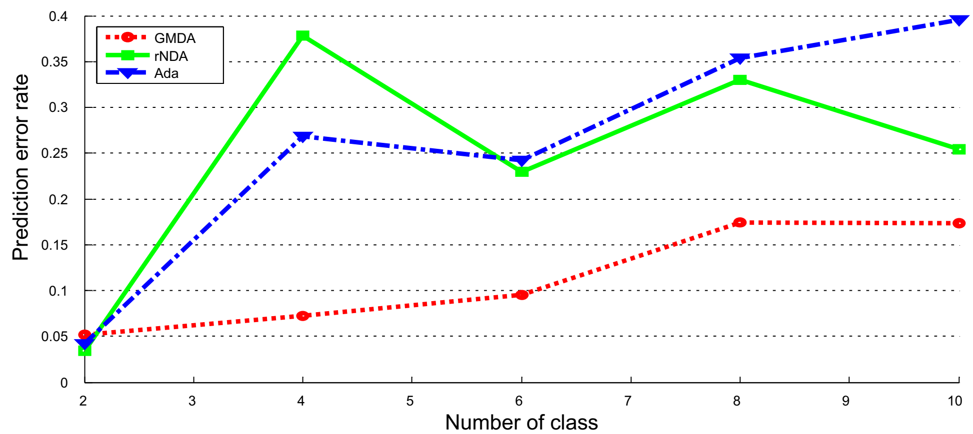


Table 6 f -

	N	*D*	-	. f	1 . f	2
1	1 ,000	200	2 0		2	0
2	1 ,000	200	0		1	1
	1 ,000	200	2 0	10		
	1 ,000	200	0	10	1	1
	1 ,000	200	2 0		0	2
	1 ,000	200	0			20

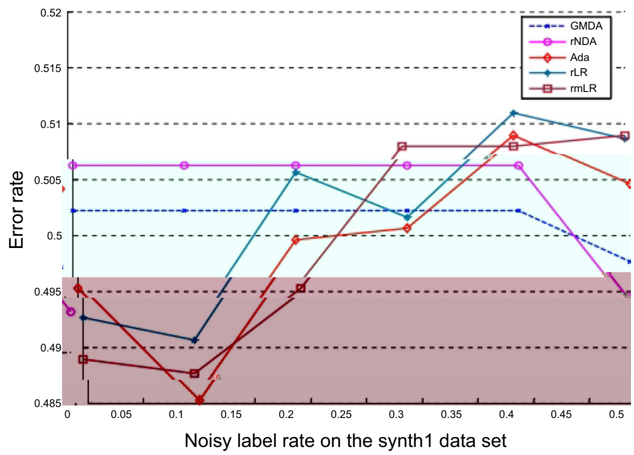
, , - ,
( ) . ff . . . - . -
. , 2 0, ff f . . ff . , f
. . f f . .
, f ( ) ff
, - , ,
0%. , , . 0, f
f . 1%. ff f .

Table 7 x x ff ff

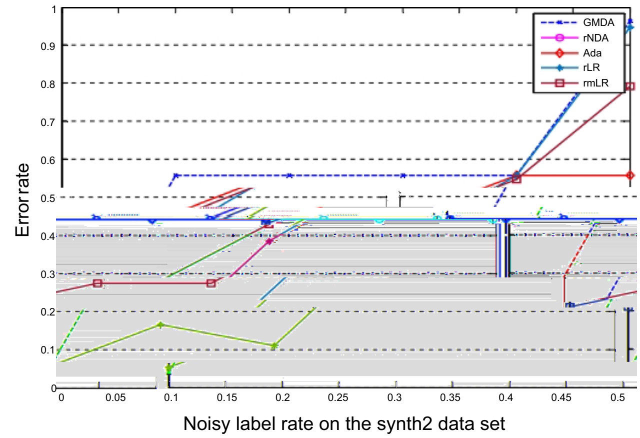
	0.0	0.1	0.2	0.	0.	0.
(a1) error rate with 5-cross-validation on correlated synth1 dataset						
	0. 02	0. 02	0. 02	**0.5023**	**0.5023**	0.
	0. 0	0. 0	0. 0	0. 0	0. 0	**0.4937**
	0.	**0.4853**	0.	0. 00	0. 0 0	0. 0
	0. 2	0. 0	0. 0	0. 01	0. 110	0. 0
	0.4890	0.	**0.4953**	0. 0 0	0. 0 0	0. 0 0
(b1) error rate with 10-cross-validation on correlated synth3 dataset						
	0. 0	0. 0	0. 2	0. 2	0. 2	**0.4927**
	0. 0	0. 0	0.	0.	0.	0.
	0.4520	**0.4520**	**0.4520**	**0.4520**	0. 0	0. 0
	0. 0	0. 0	0.	0. 0	0.	0.
	0. 120	0.	0.	0. 0	**0.4827**	0.
(c1) error rate with 3-cross-validation on correlated synth5 dataset						
	0. 0 0	0. 0 0	0. 0 0	0. 0 0	0. 0 0	**0.4938**
	0.	0. 0	0. 0	0. 0	0. 0	0. 0
	0.4812	**0.4812**	**0.4812**	**0.4812**	0. 1	0. 1
	0.	0.	0.	0. 1	0.	0. 1
	0.	0. 0	0. 0	0. 0	**0.4904**	0. 1 2
(a2) error rate with 5-cross-validation on uncorrelated synth2 dataset						
	0. 0	0. 0	0. 0	0.	0. 0	0. 0
	0. 2	0. 2	0. 2	**0.4427**	**0.4427**	0. 2
	0. 2	0. 2	0. 2	0.	0.	0. 2
	0.1653	**0.1103**	**0.3853**	0.	0.	**0.1653**
	0.2 0	0.2 0	0. 0	0.	0. 20	0.2 0
(b2) error rate with 10-cross-validation on uncorrelated synth4 dataset						
	0.0420	0. 0	0.	0. 0	0.	0. 1
	0. 0	0. 0	0. 0	0. 0	0. 0	**0.4560**
	0. 0	0. 0	0. 0	0. 0	0. 0	0. 0
	0.0	**0.2313**	**0.2320**	**0.2133**	**0.4267**	0.
	0.0 00	0.2 1	0.	0.	0.	0. 00
(c2) error rate with 3-cross-validation on uncorrelated synth6 dataset						
	0.0424	0. 2	0. 2	0. 0	0.	0. 2
	0. 22	0. 22	0. 22	0. 22	0. 22	**0.4422**
	0. 22	0. 22	0. 22	0. 22	0.	0.
	0.0	**0.1754**	**0.1562**	**0.1902**	**0.4370**	0. 2
	0.0 0	0.2	0.2	0. 2	0.	0. 112

f

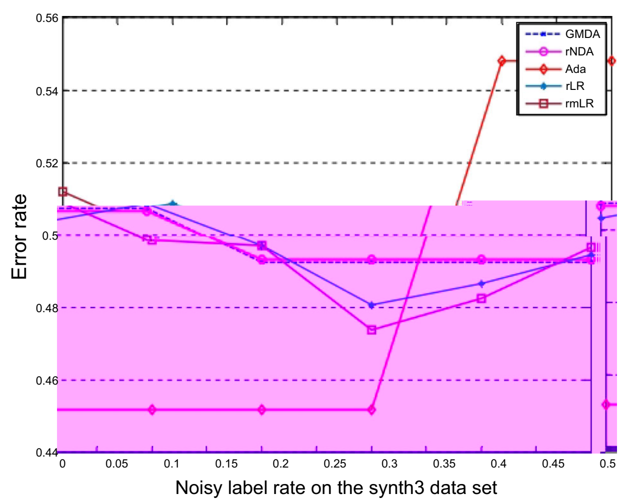
, , . . 0.2, 0. , , ff f , x , f f ff . f x . f f . x ff f , , ff f , . ff . , f , .


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

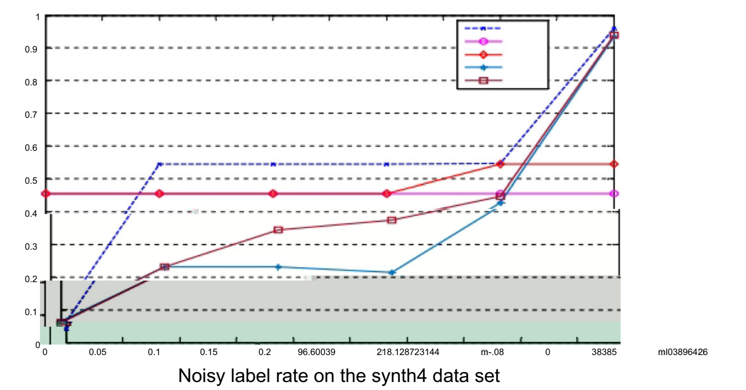
(a1) 1

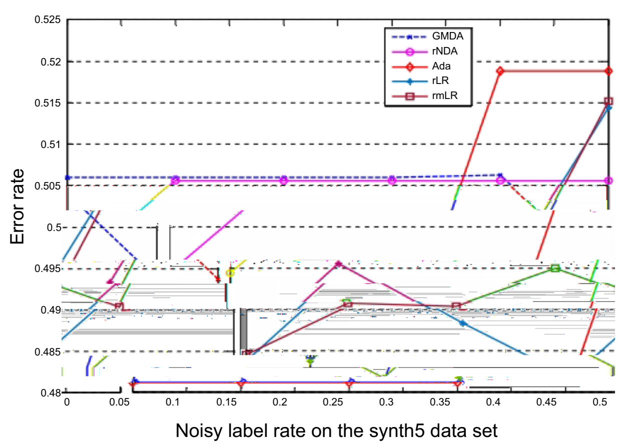

GMDA
rNDA
Ada
rLR
rmLR
Error rate
Noisy label rate on the synth2 data set

(a2) 2

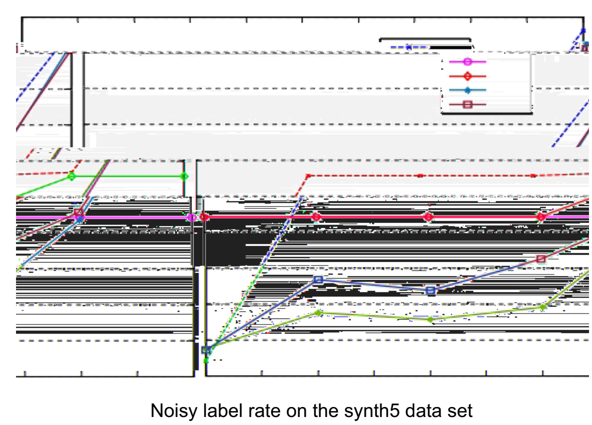

GMDA
rNDA
Ada
rLR
rmLR
Error rate
Noisy label rate on the synth3 data set

(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

x f f x ff f ff f f f ff f f f f f f x f f f x f f x x f f

Acknowledgements f

Compliance with ethical standards

Conflict of interest f f f f f f f

References

1. ff f ff f (2010) f f f , ( ) 2 - 0
2. , (201) f 2 ( ) (1) f f 11(1) 1 1 1 (200) x 1 (1) 11 (2010) f f f

f , K 2010, , , , 2 2 f (201) x f 1 11.01 1,201
, , (2001) f
, K (2012)
f , K 2012, , 1 -1 , 2012 .K (1) f f ( ) ( ) 100
10. , , (200) f 0(12)
11. , (2010) f x ( ) 2 0
12. , K , (201) f 1 0 .0
1 . , (201) f f , , x , , 2 2
1 . , , (201) x f 2 f , , 11
1 . , , , (201) x 1 0 .120
1 . , (201) ( ) 11 11 1
1 . , (1) 1 (1) 0
1 . , , K (200) f f x f 22( ) 1 2
1 . ,K f , (2012) f f f f f ,
20. , , K (2001) f f 2( ) 2 20
21. , (2000) f ( ) 2 2
22. , (1) f f f f f , , , . 0
2 . (201) f , x
2 . K (1) x x 1 ( ) 0
2 . , , (1) x f (1) 1
2 . (1 2) f x f x f .K ( ) 1 ( ) 1 1 0

[illegible]

Publisher's Note

[illegible]