TIANJIN UNIVERSITY

Hybrid Noise Oriented Multi-Label learning

混合噪声环境下的多标记学习



报告人：余子伟

天津大学计算机学院

机器学习与数据挖掘实验室

2016.11.12

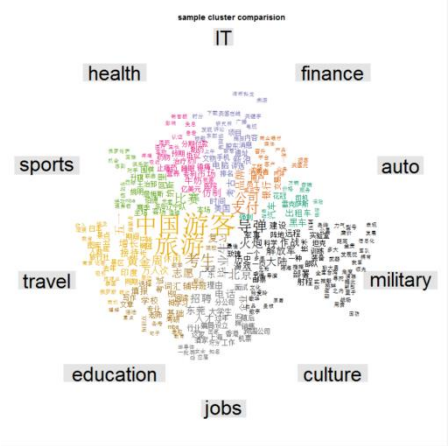
Contents



1. Background
2. Related Work
3. Our Model
4. Experiment
5. Conclusion

1. Background – Multi-label Learning

- Multi-label Learning (ML):
  - Definition: an instance can be assigned with multiple labels simultaneously
  - Applications: text categorization, functional genomics and scene classification



text categorization



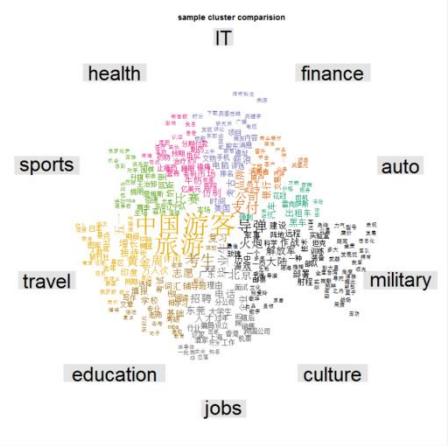
functional genomics



scene classification

1. Background – Multi-label Learning

- Multi-label Learning (ML):
  - Definition: an instance can be assigned with multiple labels simultaneously
  - Applications: text categorization, functional genomics and scene classification



text categorization



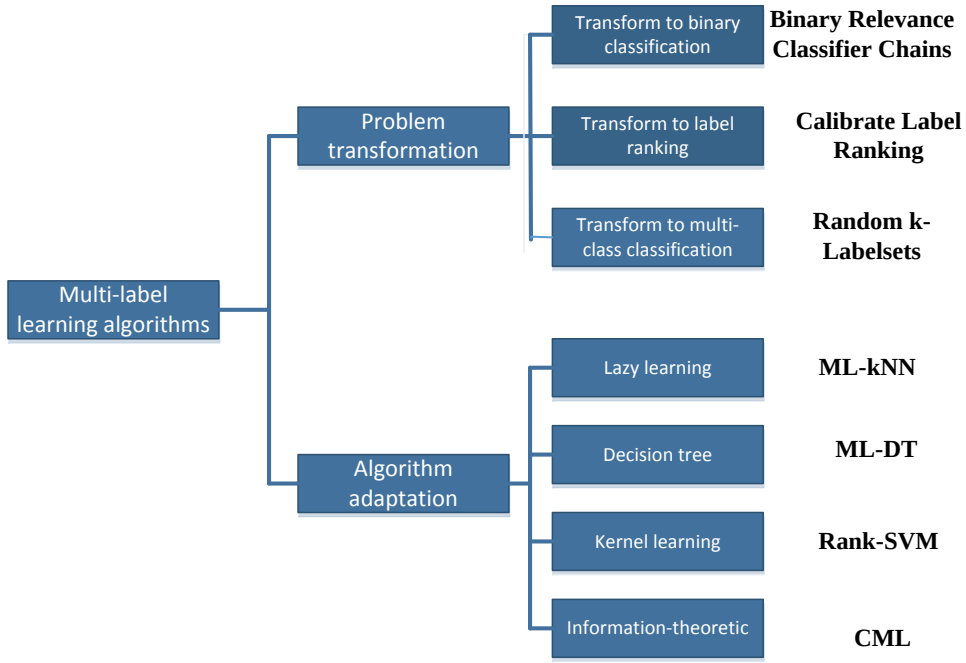
functional genomics



scene classification

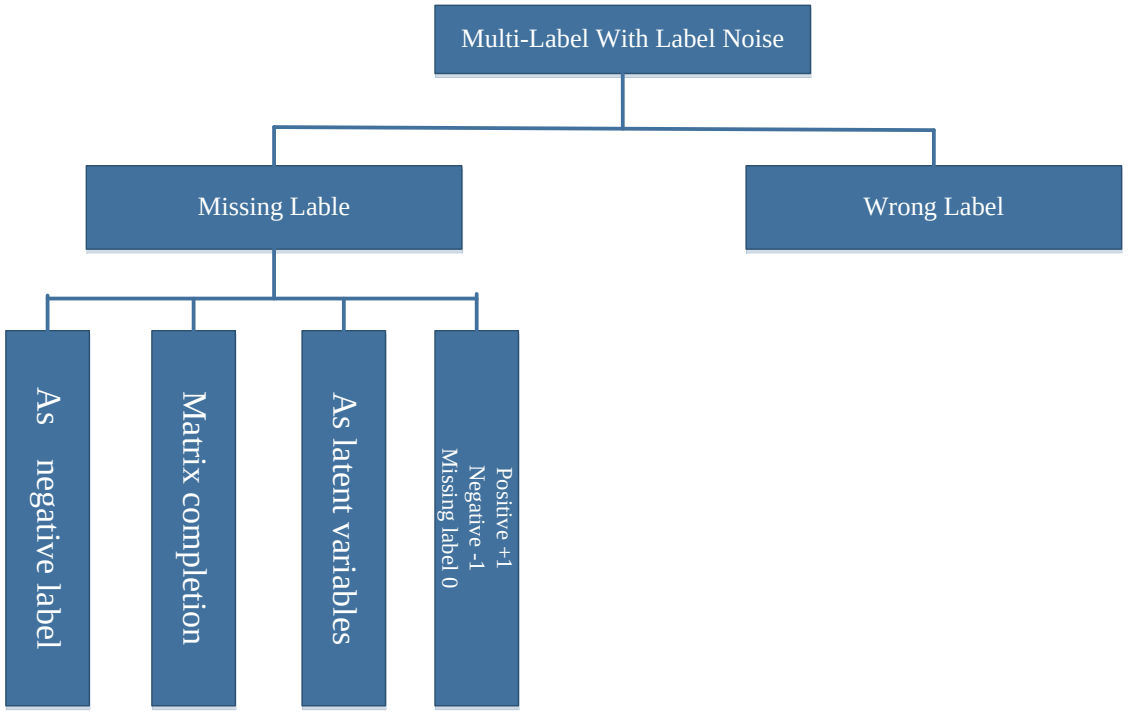
1. Background – Multi-label Learning

Traditional Multi-label Algorithm：



1. Background – Multi-label Learning

- Multi-label learning with label noise：





baby, *boy*, **child**, *eye*, *face*, *girl*, *hair*, *house*, **kid**, *mouth*, *nose*, *pink*, *smile*



anime, *ball*, *boy*, **cartoon**, **drawing**, *girl*, *group*, *hair*, *kid*, *man*, *people*, *play*, *red*, *sport*

1. Related Word– Multi-label Learning

Multi-label learning with feature noise：

- Multi-Label Dimensionality Reduction via Dependence Maximization(MDDM)：

Assumption: exist some relation between the feature description and labels associated with the same object

$\mathcal{F} \mid \phi(x) = P^T x$ ← dependence → $y \in Y$

Feature description **max** Class labels

Method:

$$\mathrm{HSIC}(\mathcal{F}, \mathcal{Y}, \mathbf{P}_{xy}) = (N-1)^{-2}\mathrm{tr}\,(\mathbf{HKHL})\,,$$

Model:

$$\max_{P} \mathrm{tr}\,(\mathbf{H}X^T PP^T X\mathbf{HL}) \quad \text{s.t.} \quad P_i^T P_j = \delta_{ij}.$$

X：D×N feature matrix,
Y：Q×N label matrix ,
P： D×d d≪D，δ_{ij}=1 when i=j, 0 otherwise.

3. Our Model – HNOML

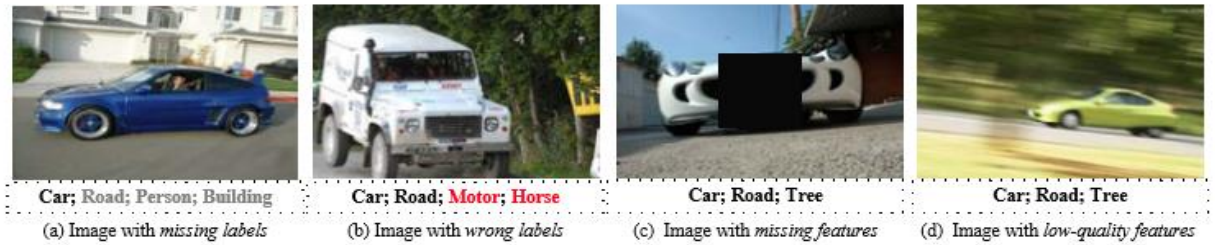
Multi-label learning with hybrid noise：

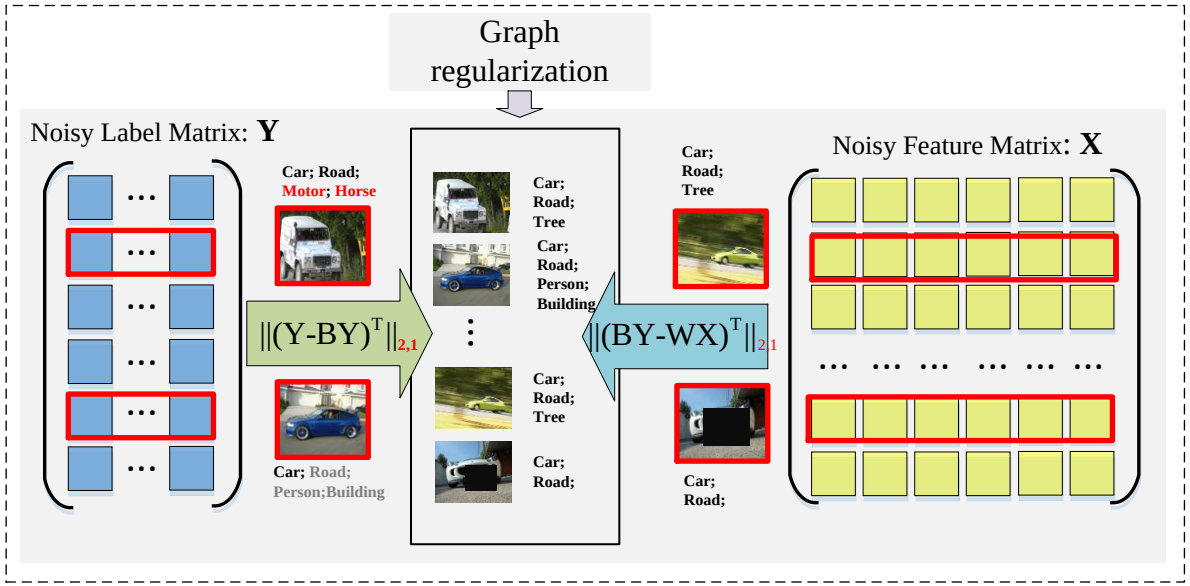
Hybrid Noise Oriented Multi-Label Learning：

- 目标函数：

$$\min_{W,B} \|(BY-WX)^T\|_{2,1}+\alpha Tr(BYLY^TB^T)+\beta\|(Y-BY)^T\|_{2,1}+\gamma\|W\|_F^2$$

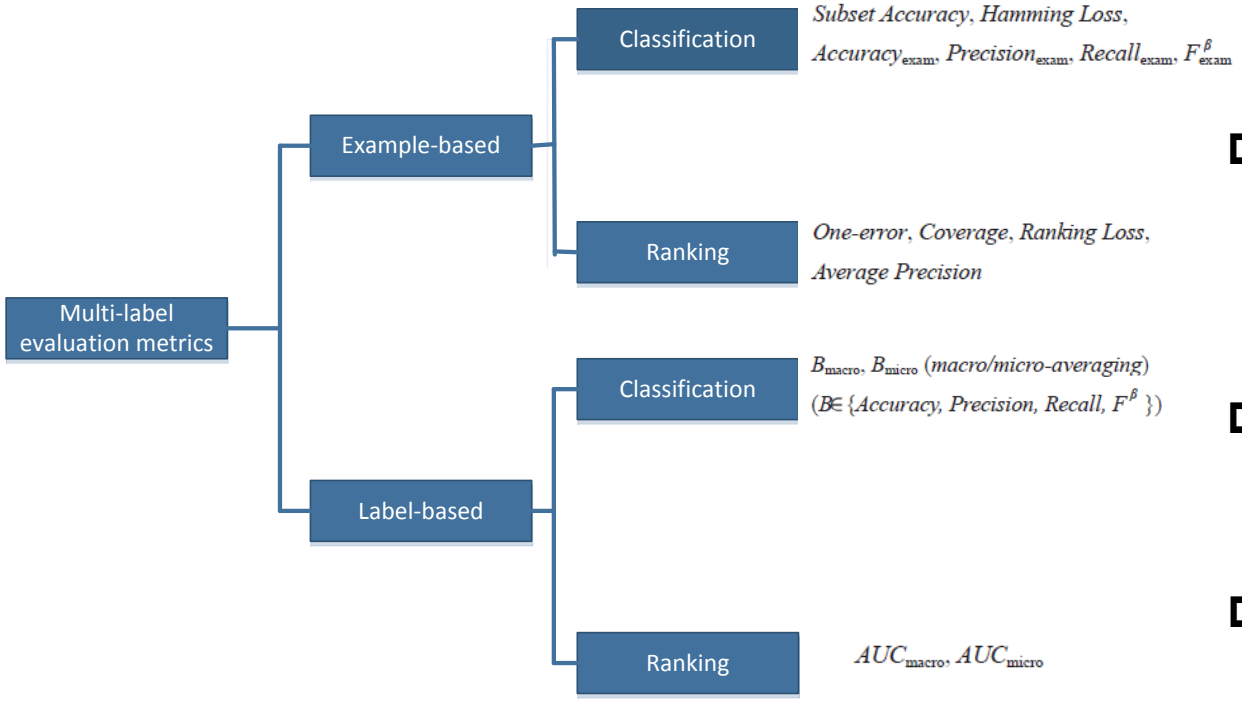
$\|(BY-WX)^T\|_{2,1}$	Structured Loss
$\alpha Tr(BYLY^TB^T)$	Label Embedding
$\beta\|(Y-BY)^T\|_{2,1}$	Label Enriching
$\gamma\|W\|_F^2$	Model Regularization





1. Background – Multi-label Learning

Multi-label evaluation metrics:



- Average Precision:

$$avgprec(f) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{|Y_i|}\sum_{y\in Y_i}\frac{|\{y' \mid rank_f(\boldsymbol{x}_i,y') \le rank_f(\boldsymbol{x}_i,y),\ y' \in Y_i\}|}{rank_f(\boldsymbol{x}_i,y)}$$

- Hamming Loss:

$$hloss(h) = \frac{1}{p}\sum_{i=1}^{p}|h(\boldsymbol{x}_i)\Delta Y_i|$$

- Ranking Loss:

$$rloss(f) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{|Y_i||\bar{Y}_i|}|\{(y',y'') \mid f(\boldsymbol{x}_i,y') \le f(\boldsymbol{x}_i,y''),\ (y',y'') \in Y_i \times \bar{Y}_i)\}|$$

- One-error:

$$one\text{-}error(f) = \frac{1}{p}\sum_{i=1}^{p}[[\arg\max_{y\in\mathcal{Y}} f(\boldsymbol{x}_i,y)] \notin Y_i]$$

- Coverage:

$$coverage(f) = \frac{1}{p}\sum_{i=1}^{p}\max_{y\in Y_i} rank_f(\boldsymbol{x}_i,y) - 1$$

4. Experiment

- Multi-label learning with hybrid noise：