Partial Multi-Label Learning with Noisy Label Identification

Ming-Kun Xie and Sheng-Jun Huang

College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics

MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, Nanjing

Partial Multi-Label Learning

☐ Partial multi-label learning (PML) deals with problems where each instance is assigned with a candidate label set, which contains multiple relevant labels and some noisy labels.

✓ More specifically, the candidate label set of each training instance indicates the following supervised information:

a) the candidate set may consist of both relevant and irrelevant labels;

b) the number of relevant labels in the candidate set is at least one; but unknown

c) label not in the candidate set are irrelevant to the instance.

Motivation

The candidate set
(Accurate ones in black)

house	tree
car	light
cloud	
flower	cat
people	

noisy labels are usually caused by some ambiguous contents of the example ➡ Identify the noisy labels by modeling them based on feature representations

The PML-NI Framework

Based on the observations mentioned above, we propose a new approach for Partial Multi-label Learning with Noisy label Identification (PML-NI), which recovers the ground-truth labeling information and identifies the noisy labels simultaneously.

Firstly, we model the noisy labels as the outputs of a linear mapping from the feature representations as follows:

$$\text{observed label}\longleftarrow\boxed{\mathbf{y}_i-\widetilde{\mathbf{y}}_i}=\widehat{\boldsymbol{V}}\mathbf{x}_i+\mathbf{s}=\boxed{\boldsymbol{V}}\phi_i$$

noisy label ground-truth noisy label identifier

However, the ground-truth label $\tilde{y}_i$ here is unknown and the equation above is thus intractable. To solve the problem, we propose a joint learning framework that can identify the noisy labels while training the multi-label classifier simultaneously

$$\min_{\boldsymbol{W},\boldsymbol{U},\boldsymbol{V}}\frac{1}{2}\|\mathbf{Y}-\boldsymbol{W}\boldsymbol{\Phi}\|_{\mathrm{F}}^2+\frac{\lambda}{2}\|\boldsymbol{W}\|_{\mathrm{F}}^2$$

$$s.t.\quad \boldsymbol{W}=\boxed{\boldsymbol{U}}+\boldsymbol{V}$$

multi-label classifier

Furthermore, to encourage the classifier and identifier to perform their individual abilities, i.e., the ground-truth label prediction and noisy label identification, we try to capture their intrinsic property and potential structure information by employing different regularizers for each of U and V .

$$\min_{\boldsymbol{W},\boldsymbol{U},\boldsymbol{V}}\frac{1}{2}\|\mathbf{Y}-\boldsymbol{W}\boldsymbol{\Phi}\|_{\mathrm{F}}^2+\frac{\lambda}{2}\|\boldsymbol{W}\|_{\mathrm{F}}^2+\boxed{\beta\Omega(\boldsymbol{U})}$$

$$+\gamma\boxed{\Psi(\boldsymbol{V})}$$

$$s.t.\quad \boldsymbol{W}=\boldsymbol{U}+\boldsymbol{V}$$

Utilize the sparsity assumption to capture a few of ambiguous feature representations which lead to the noisy labels

⬇

$$\|\boldsymbol{V}\|_0$$

Utilize the low-rank assumption to capture label correlation

⬇

$$\mathbf{rank}(\boldsymbol{U})$$

The final objective function can be re-written as follows

$$\min_{\boldsymbol{W},\boldsymbol{U},\boldsymbol{V}}\frac{1}{2}\|\mathbf{Y}-\boldsymbol{W}\boldsymbol{\Phi}\|_{\mathrm{F}}^2+\frac{\lambda}{2}\|\boldsymbol{W}\|_{\mathrm{F}}^2+\beta\mathbf{rank}(\boldsymbol{U})$$

$$+\gamma\|\boldsymbol{V}\|_0$$

$$s.t.\quad \boldsymbol{W}=\boldsymbol{U}+\boldsymbol{V}$$

Experiment

Datasets

Table 1: Characteristics of the experimental data sets.

Data set	# Instances	# Features	# Class Labels	Cardinality	Domain
music_emotion	6833	98	11	2.42	music
music_style	6839	98	10	1.44	music
birds	654	260	19	2.402	music
genbase	662	1186	27	1.252	biology
medical	978	1449	45	1.245	text
enron	1702	1001	53	3.378	text
image	2000	294	5	1.23	image
bibtex	7395	1836	159	2.402	text
corel16k	13811	500	161	2.867	image
tmc2007	21519	500	22	2.158	text

Comparison Results

When compare PML-NI approach with other methods, our algorithm shows significant superiority. It achieves the best performance in most cases.

Table 2: Experimental results of each con点击固定g approach in terms of *ranking loss*, where ●/○ indicates whether PML-NI is superior/inferior to the other method.

Data	$\alpha\%$	PML-NI	PAR-VAL	PAR-MAP	PML-LRS	ML-kNN	CPLST
music_emotion		.251 ± .009	.265 ± .008●	.253 ± .008●	.256 ± .002●	.257 ± .006●	.364 ± .009●
music_style		.141 ± .003	.157 ± .002●	.164 ± .004●	.148 ± .006●	.157 ± .005●	.232 ± .006●
birds	50%	.190 ± .014	.438 ± .058●	.285 ± .021●	.302 ± .018●	.324 ± .040●	.252 ± .012●
	100%	.207 ± .019	.400 ± .046●	.298 ± .017●	.323 ± .028●	.322 ± .019●	.283 ± .031●
	150%	.236 ± .028	.466 ± .066●	.307 ± .026●	.330 ± .014●	.331 ± .030●	.293 ± .013●
genbase	50%	.003 ± .001	.025 ± .013●	.012 ± .006●	.017 ± .004●	.008 ± .004●	.050 ± .010●
	100%	.004 ± .002	.059 ± .030●	.010 ± .004●	.017 ± .003●	.011 ± .004●	.063 ± .018●
	150%	.010 ± .003	.017 ± .008●	.011 ± .004●	.031 ± .008●	.027 ± .007●	.075 ± .016●
medical	50%	.023 ± .005	.157 ± .034●	.071 ± .015●	.048 ± .013●	.047 ± .008●	.089 ± .008●
	100%	.023 ± .007	.155 ± .035●	.074 ± .017●	.049 ± .008●	.047 ± .008●	.097 ± .010●
	150%	.025 ± .005	.147 ± .029●	.073 ± .013●	.053 ± .005●	.049 ± .005●	.102 ± .015●
enron	50%	.175 ± .013	.318 ± .070●	.188 ± .047●	.163 ± .021○	.180 ± .007●	.301 ± .019●
	100%	.176 ± .012	.376 ± .088●	.216 ± .048●	.168 ± .012○	.190 ± .011●	.294 ± .011●
	150%	.178 ± .013	.366 ± .077●	.209 ± .047●	.171 ± .021○	.196 ± .011●	.297 ± .017●
image	50%	.175 ± .005	.195 ± .045●	.267 ± .102●	.187 ± .010●	.186 ± .016●	.189 ± .019●
	100%	.178 ± .009	.198 ± .042●	.267 ± .099●	.182 ± .014●	.190 ± .012●	.189 ± .010●
	150%	.183 ± .006	.205 ± .059●	.265 ± .139●	.185 ± .015●	.212 ± .013●	.196 ± .013●
bibtex	50%	.038 ± .003	.080 ± .002●	.057 ± .001●	.042 ± .002●	.115 ± .008●	.115 ± .010●
	100%	.032 ± .002	.095 ± .006●	.062 ± .004●	.035 ± .004●	.136 ± .019●	.138 ± .002●
	150%	.033 ± .003	.098 ± .007●	.064 ± .004●	.035 ± .003●	.143 ± .011●	.151 ± .006●
corel16k	50%	.211 ± .002	.288 ± .002●	.236 ± .003●	.214 ± .003●	.264 ± .007●	.229 ± .004●
	100%	.224 ± .004	.334 ± .008●	.262 ± .005●	.226 ± .004●	.273 ± .002●	.239 ± .005●
	150%	.224 ± .006	.326 ± .007●	.258 ± .003●	.228 ± .001●	.275 ± .007●	.237 ± .005●
tmc2007	50%	.046 ± .001	.087 ± .014●	.057 ± .008●	.046 ± .001●	.075 ± .004●	.080 ± .002●
	100%	.047 ± .001	.082 ± .014●	.057 ± .009●	.047 ± .002●	.079 ± .002●	.081 ± .001●
	150%	.050 ± .002	.107 ± .023●	.060 ± .010●	.050 ± .002●	.082 ± .001●	.086 ± .001●

Table 3: Experimental results of each comparing approach in terms of *average precision*, where ●/○ indicates whether PML-NI is superior/inferior to the other method. 点击固定

Data	$\alpha\%$	PML-NI	PAR-VAL	PAR-MAP	PML-LRS	ML-kNN	CPLST
music_emotion		.598 ± .010	.607 ± .010○	.611 ± .011○	.589 ± .006●	.595 ± .007●	.506 ± .009●
music_style		.731 ± .003	.713 ± .004●	.710 ± .007●	.714 ± .008●	.717 ± .011●	.658 ± .009●
birds	50%	.507 ± .019	.413 ± .034●	.395 ± .024●	.371 ± .030●	.370 ± .037●	.451 ± .015●
	100%	.466 ± .013	.416 ± .042●	.386 ± .024●	.352 ± .033●	.366 ± .037●	.410 ± .033●
	150%	.419 ± .026	.392 ± .033●	.369 ± .023●	.344 ± .031●	.352 ± .017●	.387 ± .040●
genbase	50%	.980 ± .005	.895 ± .022●	.968 ± .020●	.860 ± .022●	.948 ± .011●	.738 ± .028●
	100%	.971 ± .010	.819 ± .039●	.965 ± .019●	.851 ± .025●	.920 ± .055●	.723 ± .030●
	150%	.922 ± .022	.897 ± .042●	.960 ± .010○	.785 ± .049●	.773 ± .069●	.612 ± .020●
medical	50%	.819 ± .010	.703 ± .021●	.737 ± .029●	.738 ± .034●	.737 ± .014●	.592 ± .027●
	100%	.809 ± .017	.680 ± .020●	.714 ± .031●	.724 ± .020●	.734 ± .014●	.568 ± .027●
	150%	.758 ± .016	.673 ± .013●	.675 ± .018●	.665 ± .014●	.664 ± .032●	.498 ± .031●
enron	50%	.563 ± .013	.297 ± .132●	.432 ± .068●	.528 ± .022●	.450 ± .017●	.350 ± .004●
	100%	.494 ± .017	.271 ± .129●	.398 ± .081●	.474 ± .019●	.412 ± .016●	.346 ± .013●
	150%	.474 ± .014	.264 ± .120●	.397 ± .058●	.453 ± .021●	.395 ± .017●	.326 ± .022●
image	50%	.780 ± .007	.770 ± .055●	.734 ± .076●	.765 ± .013●	.767 ± .015●	.766 ± .019●
	100%	.782 ± .006	.767 ± .051●	.735 ± .077●	.772 ± .016●	.763 ± .016●	.769 ± .007●
	150%	.772 ± .007	.760 ± .068●	.709 ± .150●	.770 ± .016●	.732 ± .009●	.757 ± .015●
bibtex	50%	.890 ± .008	.810 ± .009●	.831 ± .006●	.888 ± .007●	.748 ± .009●	.733 ± .017●
	100%	.889 ± .007	.763 ± .010●	.816 ± .011●	.874 ± .013●	.708 ± .028●	.621 ± .008●
	150%	.888 ± .006	.761 ± .010●	.816 ± .009●	.873 ± .006●	.697 ± .019●	.598 ± .015●
corel16k	50%	.511 ± .006	.473 ± .003●	.484 ± .003●	.511 ± .004○	.456 ± .010●	.500 ± .003●
	100%	.483 ± .007	.453 ± .006●	.454 ± .007●	.481 ± .007●	.436 ± .004●	.476 ± .005●
	150%	.487 ± .006	.458 ± .004●	.455 ± .009●	.479 ± .005●	.433 ± .009●	.475 ± .007●
tmc2007	50%	.804 ± .002	.731 ± .033●	.783 ± .022●	.803 ± .006●	.746 ± .008●	.747 ± .002●
	100%	.803 ± .003	.737 ± .035●	.785 ± .021●	.802 ± .005●	.729 ± .004●	.738 ± .005●
	150%	.793 ± .003	.676 ± .033●	.760 ± .036●	.792 ± .005●	.710 ± .005●	.721 ± .002●