Complementary Classifier Induced Partial Label Learning

Yuheng Jia[1*], Chongjie Si[2*], Min-Ling Zhang[1]
[1]School of Computer Science and Engineering, Southeast University
[2]MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University
*Equal Contribution
{yhjia, zhangml}@seu.edu.cn, chongjiesi@sjtu.edu.cn

KDD2023
AUGUST 6-10
29TH ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY & DATA MINING

INTRODUCTION

- **Insight**:
  - There are three kinds of priors that are useful for partial label learning, the correlations among instances, the mapping from instances to the candidate labels and non-candidate labels, I.e., complementary labels.
  - The majority of the existing works overlook the valuable information in complementary labels. To fully take advantage of above mentioned priors, a novel PLL method named PL-CL is proposed.
- **PL-CL**:
  - Construct an ordinary classifier and a complementary classifier.
  - Design an adversarial term to link the outputs of the two classifiers.
  - Construct an adaptive local topological graph shared by both the feature space and the label space
- **Contributions**:
  - We first propose a complementary classifier, which has never been studied in the partial label literature.
  - We propose an adversarial term to link the ordinary classifier and the complementary classifier.
  - We conducted extensive experiments to show the effectiveness of PL-CL.

METHOD

- **PL-CL**: Fig.1 shows the overall framework of PL-CL.

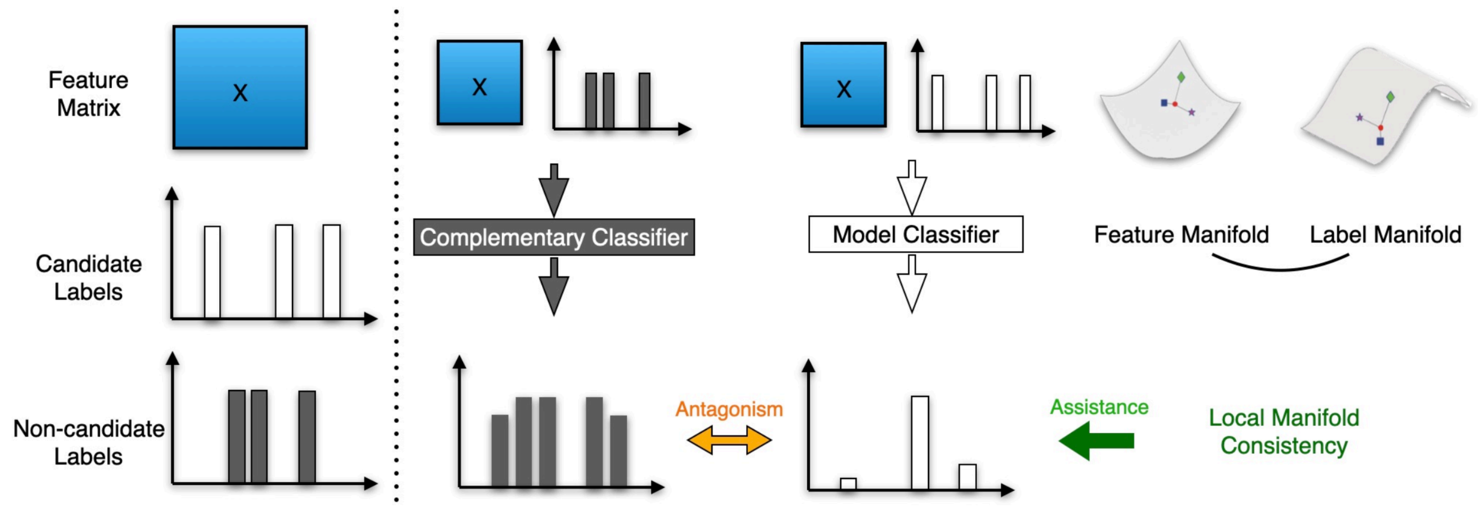

Fig.1 Overall Framework

- **PL-CL**: The loss function of PL-CL is shown as follows.

$$\min_{\substack{\mathbf{W},\mathbf{b},\hat{\mathbf{W}},\hat{\mathbf{b}},\\ \mathbf{P},\mathbf{Q},\mathbf{G}}} \left\|\mathbf{X}\mathbf{W}+\mathbf{1}_n\mathbf{b}^\mathrm{T}-\mathbf{P}\right\|_F^2+\beta\|\mathbf{1}_{n\times l}-\mathbf{P}-\mathbf{Q}\|_F^2$$

$$+\alpha\left\|\mathbf{X}\hat{\mathbf{W}}+\mathbf{1}_n\hat{\mathbf{b}}^\mathrm{T}-\mathbf{Q}\right\|_F^2+\mu\left\|\mathbf{P}^\mathrm{T}-\mathbf{P}^\mathrm{T}\mathbf{G}\right\|_F^2$$

$$+\gamma\left\|\mathbf{X}^\mathrm{T}-\mathbf{X}^\mathrm{T}\mathbf{G}\right\|_F^2+\lambda(\|\mathbf{W}\|_F^2+\|\hat{\mathbf{W}}\|_F^2)$$

$$\text{s.t.}\quad \mathbf{P}\mathbf{1}_q=\mathbf{1}_n, \mathbf{0}_{n\times l}\le\mathbf{P}\le\mathbf{Y}, \hat{\mathbf{Y}}\le\mathbf{Q}\le\mathbf{1}_{n\times l}$$

$$\mathbf{G}^\mathrm{T}\mathbf{1}_n=\mathbf{1}_n, \mathbf{0}_{n\times n}\le\mathbf{G}\le\mathbf{U}$$

EXPERIMENT

- Comparison on synthetic datasets (win/tie/loss).

	I	II	III	IV	Total
SURE	22/6/0	26/2/0	26/2/0	24/4/0	88/14/0
PL-AGGD	26/2/0	25/3/0	23/5/0	22/6/0	96/16/0
LALO	27/1/0	24/4/0	26/2/0	22/6/0	99/13/0
IPAL	28/0/0	23/5/0	26/2/0	25/3/0	102/10/0
PLDA	28/0/0	28/0/0	28/0/0	28/0/0	112/0/0
PL-KNN	28/0/0	28/0/0	28/0/0	28/0/0	112/0/0

- Comparison on real-world datasets.

Data set	FG-NET	Lost	MSRCv2	Mirflickr	Soccer Player	Yahoo!News	FG-NET(MAE3)	FG-NET(MAE5)
PL-CL	0.072 ± 0.009	0.709 ± 0.022	0.469 ± 0.016	0.642 ± 0.012	0.534 ± 0.004	0.618 ± 0.003	0.433 ± 0.022	0.575 ± 0.015
SURE	0.052 ± 0.006 •	0.693 ± 0.020 •	0.445 ± 0.021 •	0.631 ± 0.021 •	0.519 ± 0.004 •	0.598 ± 0.002 •	0.356 ± 0.019 •	0.494 ± 0.020 •
PL-AGGD	0.063 ± 0.009 •	0.683 ± 0.014 •	0.451 ± 0.012 •	0.610 ± 0.011 •	0.524 ± 0.004 •	0.607 ± 0.004 •	0.387 ± 0.015 •	0.530 ± 0.015 •
LALO	0.065 ± 0.009 •	0.680 ± 0.014 •	0.448 ± 0.015 •	0.626 ± 0.013 •	0.523 ± 0.003 •	0.600 ± 0.003 •	0.423 ± 0.020 •	0.566 ± 0.014 •
IPAL	0.051 ± 0.011 •	0.646 ± 0.023 •	0.488 ± 0.031 ○	0.527 ± 0.009 •	0.528 ± 0.003 •	0.625 ± 0.004 ○	0.349 ± 0.019 •	0.500 ± 0.019 •
PLDA	0.042 ± 0.005 •	0.289 ± 0.045 •	0.422 ± 0.013 •	0.480 ± 0.015 •	0.493 ± 0.003 •	0.380 ± 0.003 •	0.150 ± 0.012 •	0.232 ± 0.012 •
PL-KNN	0.036 ± 0.006 •	0.296 ± 0.021 •	0.393 ± 0.014 •	0.454 ± 0.015 •	0.483 ± 0.005 •	0.368 ± 0.004 •	0.288 ± 0.013 •	0.440 ± 0.016 •

- Comparison on transductive accuracy.

Data set	FG-NET	Lost	MSRCv2	Mirflickr	Soccer Player	Yahoo!News	FG-NET(MAE3)	FG-NET(MAE5)
PL-CL	0.159 ± 0.016	0.832 ± 0.019	0.585 ± 0.012	0.697 ± 0.016	0.715 ± 0.002	0.840 ± 0.004	0.600 ± 0.023	0.737 ± 0.018
SURE	0.158 ± 0.012	0.798 ± 0.019 •	0.603 ± 0.015 ○	0.652 ± 0.022 •	0.700 ± 0.003 •	0.798 ± 0.005 •	0.590 ± 0.018 •	0.727 ± 0.019 •
PL-AGGD	0.130 ± 0.015 •	0.804 ± 0.016 •	0.551 ± 0.015 •	0.653 ± 0.015 •	0.698 ± 0.003 •	0.817 ± 0.005 •	0.530 ± 0.021 •	0.679 ± 0.024 •
LALO	0.153 ± 0.016 •	0.817 ± 0.012 •	0.548 ± 0.009 •	0.675 ± 0.017 •	0.698 ± 0.003 •	0.821 ± 0.004 •	0.592 ± 0.024 •	0.730 ± 0.015 •
IPAL	0.148 ± 0.021	0.793 ± 0.017 •	0.680 ± 0.013 ○	0.586 ± 0.007 •	0.681 ± 0.003 •	0.839 ± 0.003	0.563 ± 0.021 •	0.698 ± 0.022 •
PLDA	0.042 ± 0.005 •	0.351 ± 0.060 •	0.479 ± 0.015 •	0.564 ± 0.015 •	0.493 ± 0.004 •	0.460 ± 0.009 •	0.150 ± 0.012 •	0.232 ± 0.012 •
PL-KNN	0.041 ± 0.007 •	0.338 ± 0.016 •	0.415 ± 0.013 •	0.466 ± 0.013 •	0.504 ± 0.005 •	0.403 ± 0.009 •	0.285 ± 0.016 •	0.438 ± 0.014 •

- Ablation study.

Classification Accuracy											
Kernel	Complementary Classifier	Graph	FG-NET	Lost	MSRCv2	Mirflickr	Soccer Player	Yahoo!News	FG-NET(MAE3)	FG-NET(MAE5)	Average
✗	✗	✗	0.061 ± 0.006 •	0.622 ± 0.019 •	0.381 ± 0.015 •	0.249 ± 0.010 •	0.492 ± 0.003 •	0.430 ± 0.051 •	0.402 ± 0.031 •	0.551 ± 0.024 •	0.398
✓	✗	✗	0.060 ± 0.008 •	0.654 ± 0.019 •	0.426 ± 0.017 •	0.533 ± 0.012 •	0.515 ± 0.010 •	0.526 ± 0.006 •	0.413 ± 0.026 •	0.564 ± 0.018 •	0.461
✓	✗	✓	0.057 ± 0.009 •	0.705 ± 0.023 •	0.462 ± 0.015 •	0.637 ± 0.011 •	0.530 ± 0.004 •	0.517 ± 0.046 •	0.416 ± 0.017 •	0.560 ± 0.019 •	0.485
✓	✓	✗	0.065 ± 0.008 •	0.684 ± 0.022 •	0.456 ± 0.014 •	0.635 ± 0.012 •	0.529 ± 0.003 •	0.607 ± 0.003 •	0.426 ± 0.024 •	0.566 ± 0.017 •	0.496
✓	✓	✓	0.072 ± 0.009	0.709 ± 0.022	0.469 ± 0.016	0.642 ± 0.012	0.534 ± 0.004	0.618 ± 0.003	0.433 ± 0.022	0.575 ± 0.015	0.507
Transductive Accuracy											
Kernel	Complementary Classifier	Graph	FG-NET	Lost	MSRCv2	Mirflickr	Soccer Player	Yahoo!News	FG-NET(MAE3)	FG-NET(MAE5)	Average
✗	✗	✗	0.161 ± 0.017 •	0.729 ± 0.017 •	0.414 ± 0.013 •	0.464 ± 0.008 •	0.492 ± 0.004 •	0.439 ± 0.003 •	0.580 ± 0.024 •	0.717 ± 0.022 •	0.499
✓	✗	✗	0.159 ± 0.012 •	0.748 ± 0.015 •	0.515 ± 0.018 •	0.594 ± 0.013 •	0.700 ± 0.004 •	0.774 ± 0.032 •	0.588 ± 0.021 •	0.726 ± 0.010 •	0.600
✓	✗	✓	0.141 ± 0.011 •	0.819 ± 0.017 •	0.572 ± 0.013 •	0.685 ± 0.014 •	0.714 ± 0.002 •	0.839 ± 0.004 •	0.568 ± 0.025 •	0.706 ± 0.022 •	0.630
✓	✓	✗	0.150 ± 0.015 •	0.824 ± 0.021 •	0.567 ± 0.008 •	0.687 ± 0.012 •	0.701 ± 0.003 •	0.829 ± 0.004 •	0.570 ± 0.022 •	0.710 ± 0.021 •	0.629
✓	✓	✓	0.159 ± 0.016	0.832 ± 0.019	0.585 ± 0.012	0.697 ± 0.016	0.715 ± 0.002	0.840 ± 0.004	0.600 ± 0.023	0.737 ± 0.018	0.646

Code Link: https://github.com/Chongjie-Si/PL-CL Yuheng Jia: https://cse.seu.edu.cn/2020/1114/c23024a353193/page.htm Chongjie Si: https://chongjiesi.site