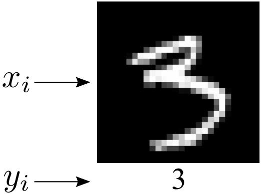
Conformal Prediction with Partially Labeled Data

Alireza Javanmardi **Yusuf Sale** **Paul Hofman** **Eyke Hüllermeier**

Institute of Informatics, LMU Munich, Germany

Munich Center for Machine Learning (MCML), Germany

Standard Setting of Multi-class Classification



- Instance space: $\mathcal{X}$
- Label space: $\mathcal{Y} = \{1, \ldots, K\}$
- Training data: $\mathcal{D} = \{\, (x_i, y_i) \,\}_{i=1}^{n} \in (\mathcal{X} \times \mathcal{Y})^n$
- A model is sought that minimizes the empirical risk, i.e.,

$$\hat{h} = \underset{h \in \mathcal{H}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} L(y_i, h(x_i))$$

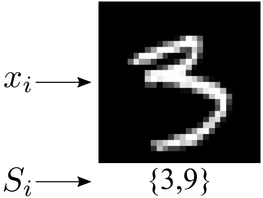
where $\mathcal{H}$ is a hypothesis space and $L : \mathcal{Y} \times \mathcal{Y} \longrightarrow \mathbb{R}$ is a loss function.

Partial Label Learning (PLL)

- Instance space: $\mathcal{X}$
- Label space: $\mathcal{Y} = \{1, \ldots, K\}$
- Training data: $\mathcal{O} = \{\, (x_i, S_i) \,\}_{i=1}^{n} \in (\mathcal{X} \times 2^{\mathcal{Y}})^n$
- **Assumption:** $y_i \in S_i, \forall x_i$
- A model is sought that minimizes the empirical risk, i.e.,

$$\hat{h} = \underset{h \in \mathcal{H}}{\operatorname{argmin}} \frac{1}{n} \sum_{i=1}^{n} L_O(S_i, h(x_i))$$

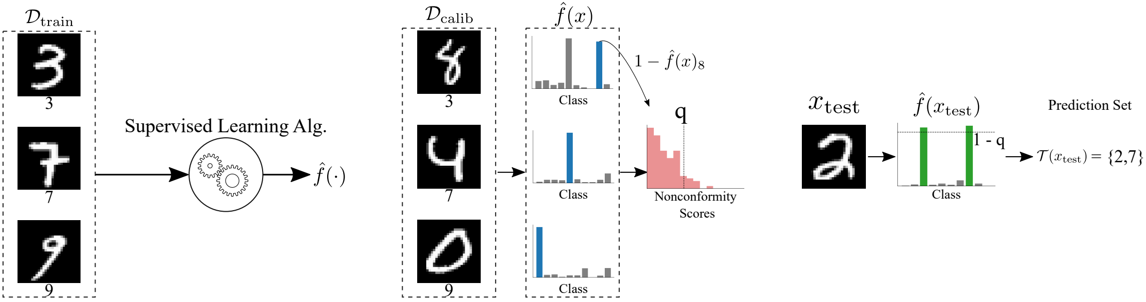
where $\mathcal{H}$ is a hypothesis space and $L_O(S, \hat{y}) = \min_{y \in S} L(y, \hat{y})$ is an extension of the loss L.



Our Goal

- In PLL:
  - Training data: ambiguous (imprecise)
  - Induced model: precise
- **Goal:** to reflect this ambiguity in the predictions
- **How:** by extending the (split) conformal prediction framework for partially labeled data

Conformalized Multi-class Classification



(a) Training (b) Calibration (c) Test

Validity

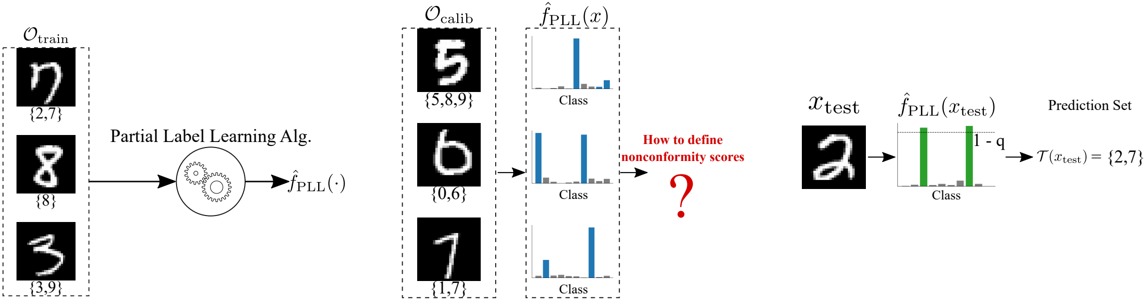
Theorem 1 (Validity[1])

If the data points in $\mathcal{D}_{calib} \cup (x_{test}, y_{test})$ *are exchangeable, then*

$$\mathbb{P}\Big(y_{test} \in \mathcal{T}(x_{test})\Big) \geq 1 - \epsilon.$$

[1] Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic learning in a random world*. Springer, 2005.

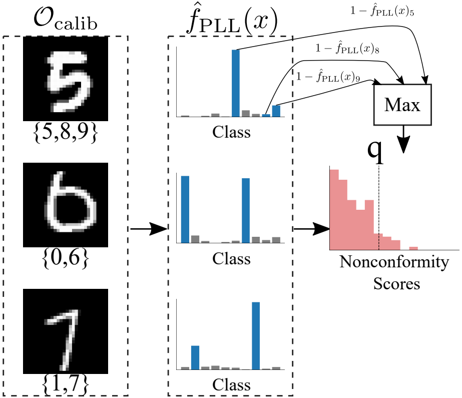
Conformal Prediction for Partial Label Learning



(a) Training (b) Calibration (c) Test

Conformal Prediction for Partial Label Learning

First Proposal: "Max" approach



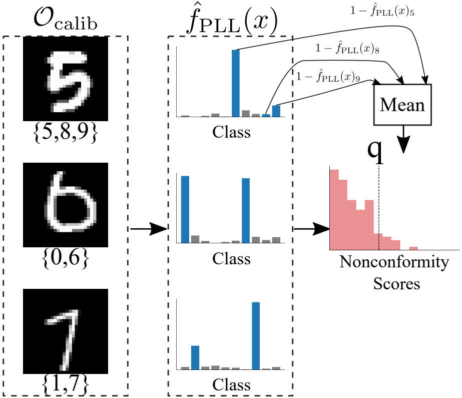
Conformal Prediction for Partial Label Learning

- $\mathcal{O}'_{\text{calib}}$: the precise counterpart of $\mathcal{O}_{\text{calib}}$

Theorem 2 (Validity of "Max" approach)

If the data points in $\mathcal{O}'_{calib} \cup (x_{test}, y_{test})$ are exchangeable, then the prediction set constructed with the "Max" approach is valid.

Second Proposal: : "Mean" approach



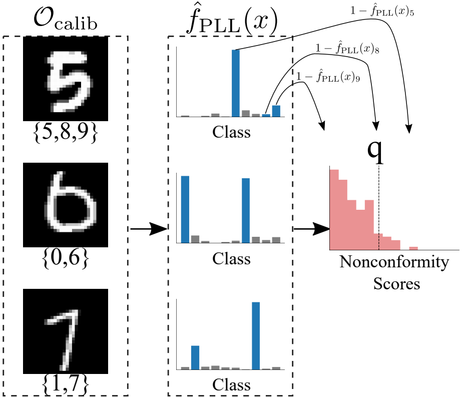
Conformal Prediction for Partial Label Learning

Theorem 3 (Validity of "Mean" approach)

If the data points in $\mathcal{O}'_{calib} \cup (x_{test}, y_{test})$ are exchangeable and $\hat{f}_{PLL}(x_j)_{y_j} \geq \frac{1}{|S_j|}$, $\forall j \in \mathcal{O}_{calib}$, then the prediction set constructed with the "Mean" approach is valid.

Conformal Prediction for Partial Label Learning

Third Proposal: "All" approach



Conformal Prediction for Partial Label Learning

- $\mathcal{E}' := \left\{1 - \hat{f}_{\text{PLL}}(x_j)_{y_j} : j \in \mathcal{O}'_{\text{calib}}\right\}$
- q': the $\lceil (1 + |\mathcal{E}'|)(1-\epsilon) \rceil$ smallest value of $\mathcal{E}'$

Theorem 4 (Validity of "All" approach)

For any $\epsilon \leq \min\left(\frac{1}{4}, \frac{|\mathcal{O}_{calib}| + |\mathcal{Y}|}{|\mathcal{Y}| \cdot (1 + |\mathcal{O}_{calib}|)}\right)$, if the points in $\mathcal{O}'_{calib} \cup \{(x_{test}, y_{test})\}$ are exchangeable and $q' \leq 0.5$, then the prediction set constructed with the "All" approach is valid.

Numerical Experiments

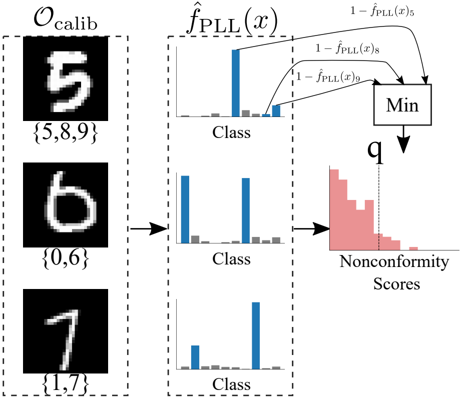
Description of the benchmark datasets

		FashionMNIST	KMNIST	MNIST
	Num. of classes	10	10	10
Avg. CSS[2]	Instance-dependent contamination	2.32	2.49	2.25
	Random contamination (p=0.7)	7.30	7.30	7.30

[2]candidate set sizes

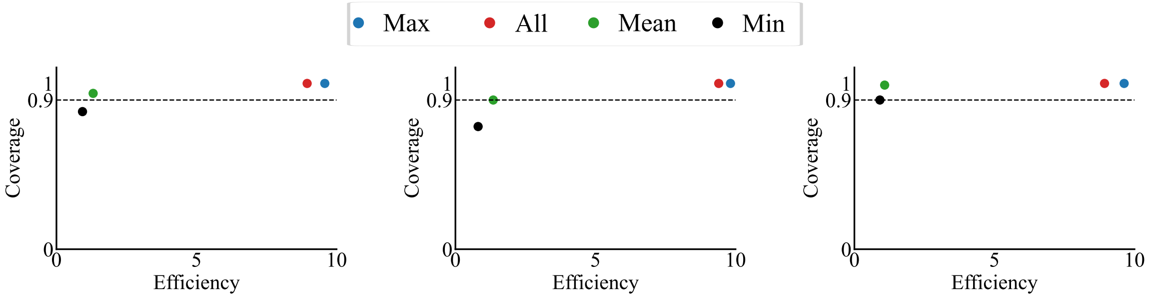
Baseline Approach

Taking minimum nonconformity score per calibration instance:



Numerical Experiments

Performance comparison of different calibration approaches on benchmark datasets with **random contamination ($p = 0.7$)**:

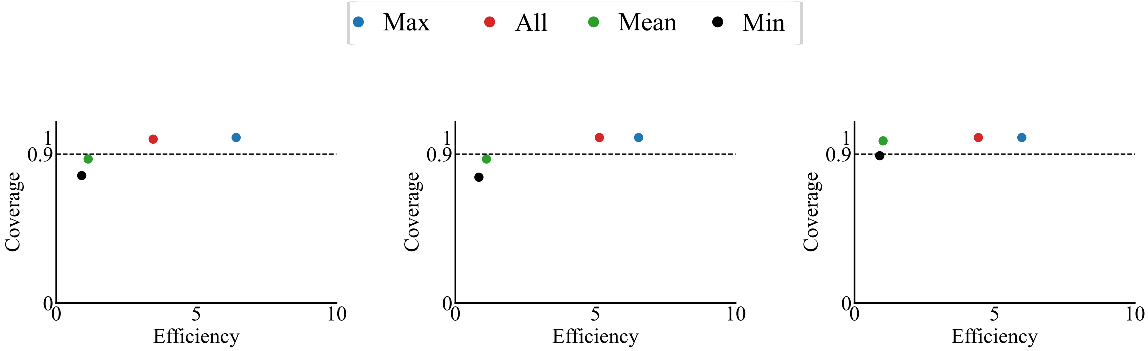


(a) FashionMNIST (b) KMNIST (c) MNIST

- Efficiency: average cardinality of prediction sets

Numerical Experiments

Performance comparison of different calibration approaches on benchmark datasets with **instance-dependent contamination**:



(a) FashionMNIST (b) KMNIST (c) MNIST

Conclusion and Future Work

- We connect two popular machine learning frameworks: conformal prediction and partial label learning.
- We theoretically show that the prediction sets constructed by the proposed approaches are valid.

Possible future work:

- There is room for developing novel approaches that could yield **more efficient results** while preserving the validity.
- It is worth exploring **nonconformity scores** other than the one used in this work.

Let's follow each other

@AlirezaJVNMRDI