Adaptive Prediction Sets with Class Conditional Coverage

Alex Derhacobian, John Guibas, Linden Li, Bharath Namboothiry

Summary

We introduce **CCAPS:** an algorithm which can **convert any black-box classifier** to output a predictive set of labels guaranteed to satisfy **class conditional coverage.**

CCAPS works regardless of how big the calibration set **(finite sample)**, works on any dataset **(distribution-free)** can wrap around any model **(model-agnostic)** and is flexible enough to allow users to choose a subset of classes on which they want the coverage guarantees to hold.

Goal

Suppose we have a labeled dataset $\{(X_i, Y_i)\}_{i=1}^n$. Then, given an unseen example (X_{n+1}, Y_{n+1}), we want to construct a set valued function $\hat{C}_{n,\alpha}: \mathcal{X} \to 2^{\mathcal{Y}}$ such that

$$\mathbb{P}\left[Y_{n+1} \in \hat{C}(X_{n+1}) \middle| Y_{n+1} = y\right] > 1 - \alpha$$

holds. This is the **class-conditional coverage** guarantee.

Method

Algorithm 1 CCAPS (Class Conditional Adaptive Prediction Sets)

Input: dataset $\{(X_i, Y_i)\}_{i=1}^n$, new data point X_{n+1}, confidence level α, number of classes K, exclusion threshold $\gamma \in [0, K-1]$
Split the training data into 2 subsets, $\mathcal{I}_{\text{train}}, \mathcal{I}_{\text{cal}}$
Split $\mathcal{I}_{\text{cal}}$ into K disjoint subsets $\mathcal{I}_{\text{cal}}^{(1)}, \ldots, \mathcal{I}_{\text{cal}}^{(K)}$, where $\mathcal{I}_{\text{cal}}^{(k)} = \{(X_i, Y_i) \in \mathcal{I}_{\text{cal}} \mid Y_i = k\}$
$\hat{\pi} \leftarrow$ black-box learning model trained on $\mathcal{I}_{\text{train}}$
for $k \in 1, \ldots, K$ **do**
    $E_i \leftarrow E(X_i, Y_i; \hat{\pi}),\ \forall\ (X_i, Y_i) \in \mathcal{I}^{(k)}$ where E is defined in (6)
    $\hat{\tau}_k \leftarrow$ the $\lceil (1-\alpha)(1 + |\mathcal{I}_{\text{cal}}^{(k)}|) \rceil$-th largest value in $\{E_i\}_{i \in \mathcal{I}_{\text{cal}}^{(k)}}$
end for
$\{\hat{\tau}_k\} \leftarrow \{\hat{\tau}_k\}$ excluding top γ largest $\hat{\tau}_k$
$\hat{\tau}_{\text{max}} \leftarrow \max_k \{\hat{\tau}_k\}$
$\hat{C}_{n,\alpha}^{SC}(X_{n+1}) \leftarrow S(X_{n+1}; \hat{\pi}, \hat{\tau}_{\text{max}})$
Output: The $1-\alpha$ class-conditional prediction set $\hat{C}_{n,\alpha}^{SC}(X_{n+1})$ for unobserved label Y_{n+1}

Formal Guarantees

Theorem (CCAPS Coverage Guarantee). Let the data be denoted by $\{(X_i, Y_i)\}_{i=1}^n$. If the data is exchangeable and the black box learning algorithm utilized to train the model in Algorithm 1 is invariant to permutations of its input samples, the output of Algorithm 1 satisfies the following for each class $k \in \{1, \ldots, K\}$:

$$\mathbb{P}\left[Y_{n+1} \in \hat{C}(X_{n+1}) \middle| Y_{n+1} = k\right] > 1 - \alpha.$$

Experiments

Hyperparameters			Coverage		Min. CC		Set Size	
Dataset	# Calib	Method	APS	CCAPS	APS	CCAPS	APS	CCAPS
MNIST (C=10)	10K	RFC	0.90	0.96	0.79	**0.90**	1.49	2.12
MNIST (C=10)	10K	MLP	0.90	0.92	0.87	**0.89**	0.89	0.98
CIFAR10 (C=10)	10K	RFC	0.90	0.93	0.85	**0.90**	5.43	6.27
CIFAR10 (C=10)	10K	MLP	0.90	0.94	0.85	**0.90**	4.94	6.68
ImageNet (C=1K)	25K	RN-152	0.90	0.99	0.52	**0.91**	6.34	296.
CIFAR-10 (C=10)	5K	MLP	0.91	0.93	0.86	**0.90**	5.28	5.5
CIFAR-10 (C=10)	2.5K	MLP	0.90	0.94	0.85	**0.90**	4.26	5.00
CIFAR-10 (C=10)	1K	MLP	0.90	0.92	0.80	**0.89**	4.33	4.99
CIFAR10-LT (C=10)	10K	RFC	0.90	0.99	0.29	**0.90**	5.31	9.65

- Compared to APS by Romano et al., CCAPS achieves coverage **for every class** for experiments when $\alpha = 0.1$.
- CCAPS provides a **finite-sample guarantee,** working for an arbitrary calibration set size.

Applications and Future Steps

- Class-conditional coverage is an important metric in both healthcare, security, and other related risky fields where reliable uncertainty quantification is key.
- Our set sizes are currently large, so *approximate conditional coverage* as a weaker guarantee could be explored.