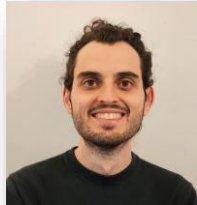


UQinMIA – Sept 2025

Uncertainty Quantification for Medical Foundation Models



Julio
Silva-Rodríguez



Ismail
Ben Ayed



Jose
Dolz

Outline

A. Vision-Language Models (VLMs)

- Contrastive Language-Image Pre-training (CLIP).
- Zero-shot and few-shot inference.
- Calibration in contrastive VLMs.
- Vision-language models for medical imaging.

B. Conformal Prediction in VLMs

- Split conformal prediction (SCP).
- Theoretical guarantees in CP.
- Benefits of foundation models for CP.
- Full conformal predictors.
- Full conformal adaptation (FCA).
- Interpretability of conformal sets.

Outline

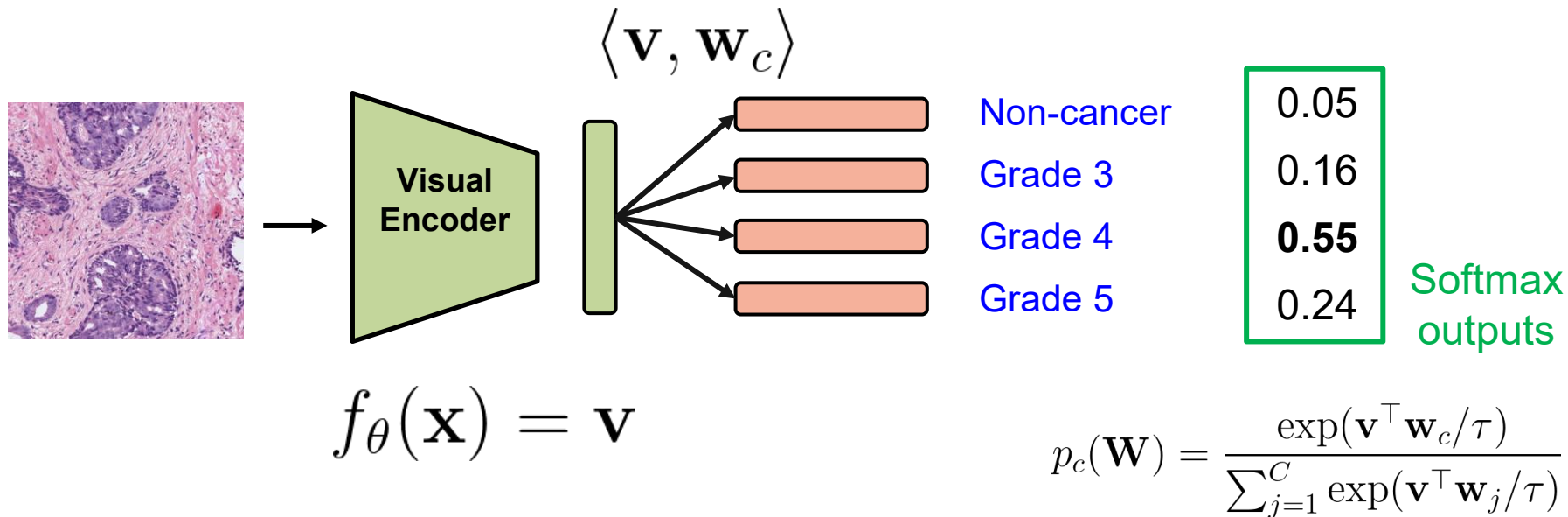
A. Vision-Language Models (VLMs)

- Contrastive Language-Image Pre-training (CLIP).
- Zero-shot and few-shot inference.
- Calibration in contrastive VLMs.
- Vision-language models for medical imaging.

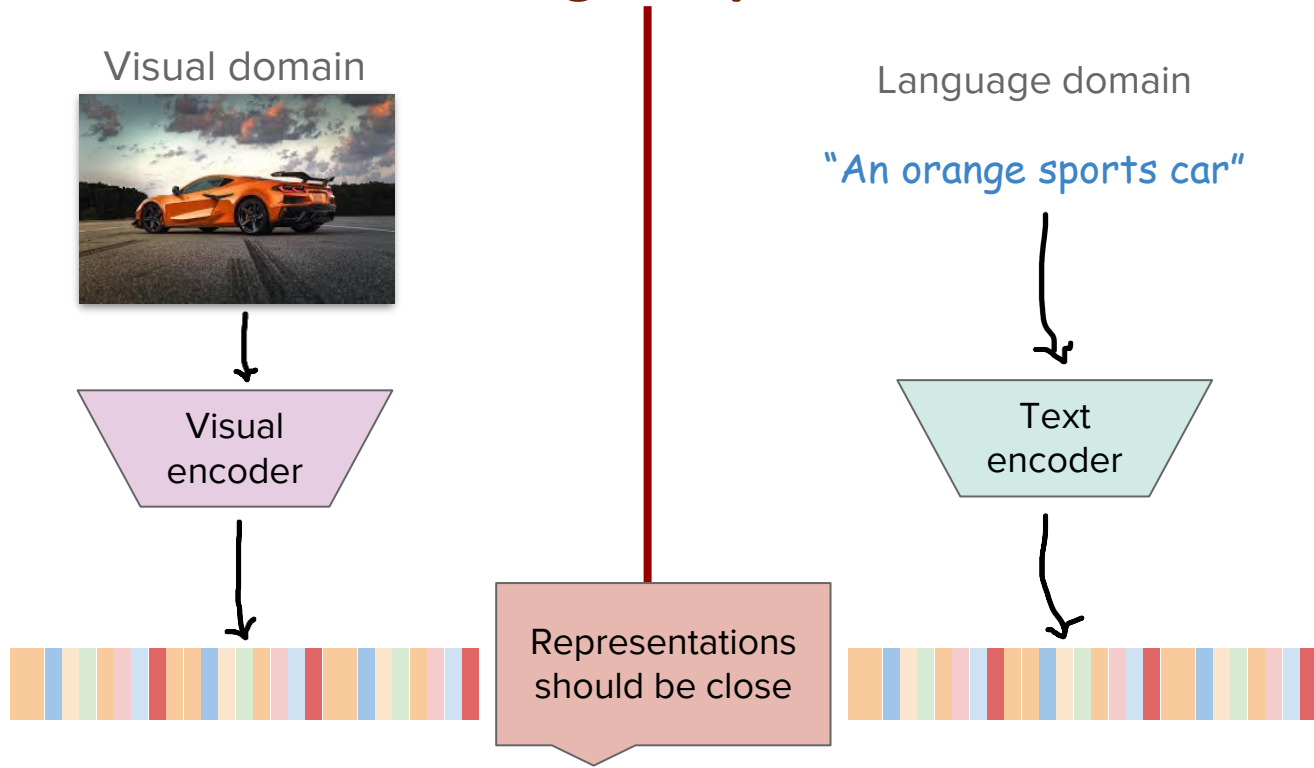
B. Conformal Prediction in VLMs

- Split conformal prediction (SCP).
- Theoretical guarantees in CP.
- Benefits of foundation models for CP.
- Full conformal predictors.
- Full conformal adaptation (FCA).
- Interpretability of conformal sets.







Dataset-focused image classifiers



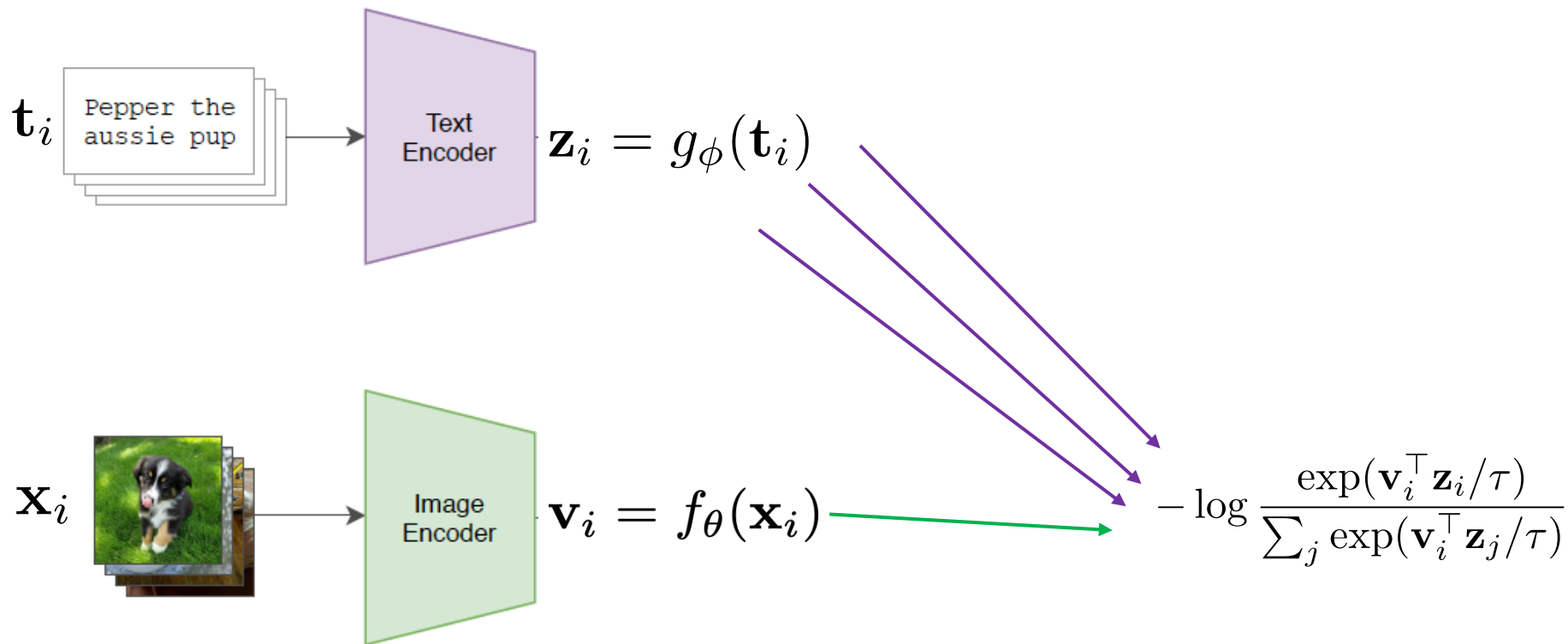
Foundational Vision-Language Models (VLMs) are transforming computer vision



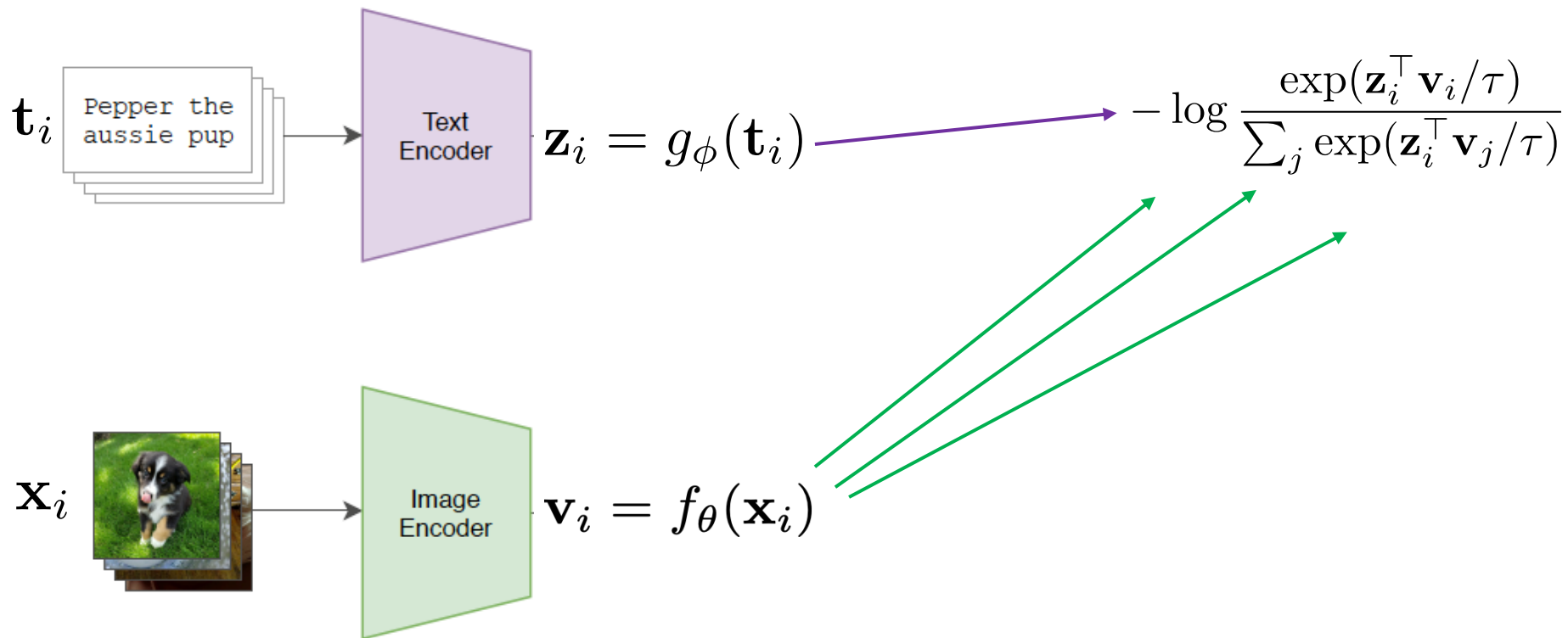
Promising zero-shot generalization (in comparison to standard task-specific learning)

	Dataset Examples	ImageNet ResNet101	Zero-Shot CLIP	Δ Score
ImageNet		76.2	76.2	0%
ImageNetV2		64.3	70.1	+5.8%
ImageNet-R		37.7	88.9	+51.2%
ObjectNet		32.6	72.3	+39.7%
ImageNet Sketch		25.2	60.2	+35.0%
ImageNet-A		2.7	77.1	+74.4%

Contrastive Language-Image Pre-training (CLIP)

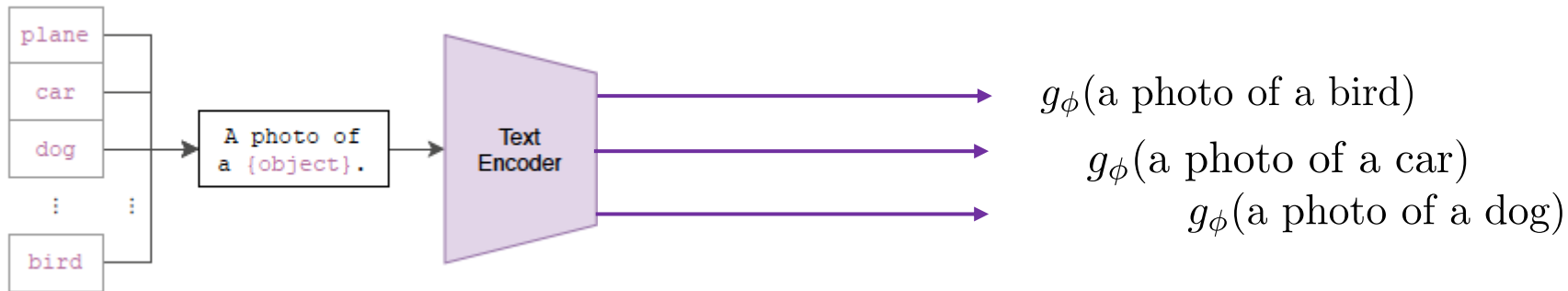


Contrastive Language-Image Pre-training (CLIP)



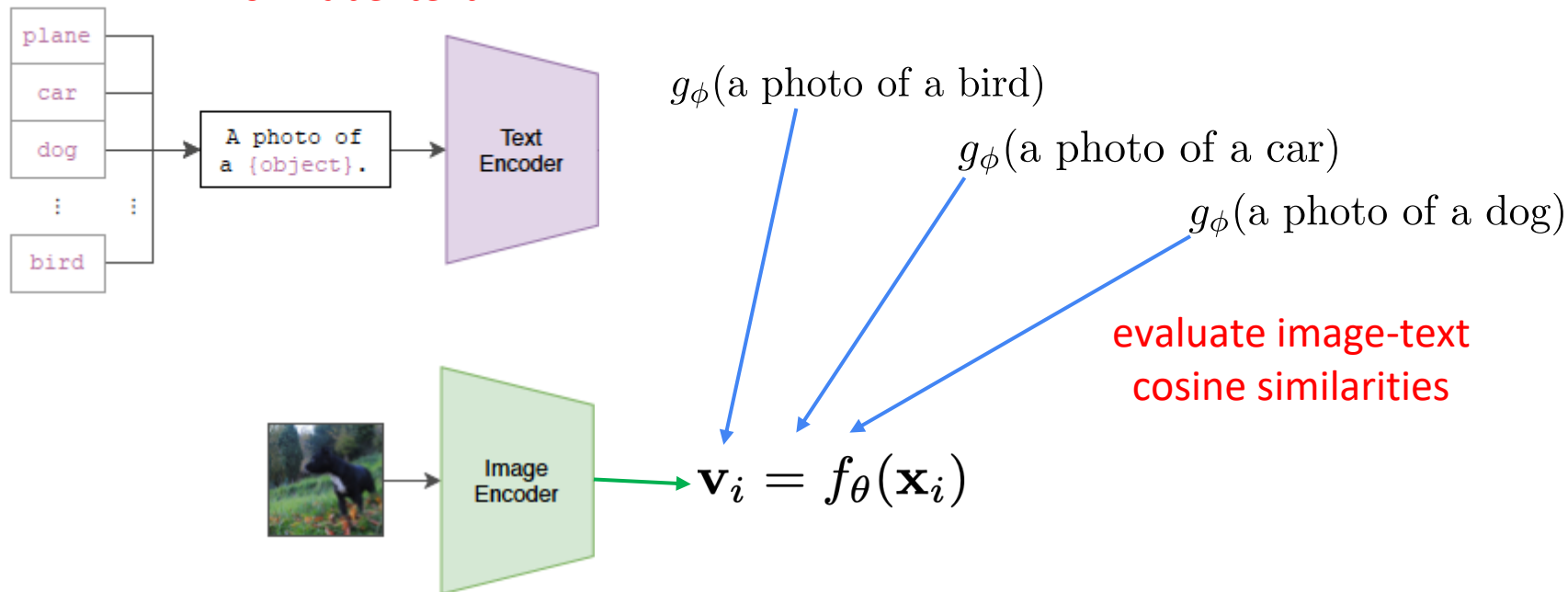
“Zero-shot” prediction for new tasks

create “class prototypes”
From label text



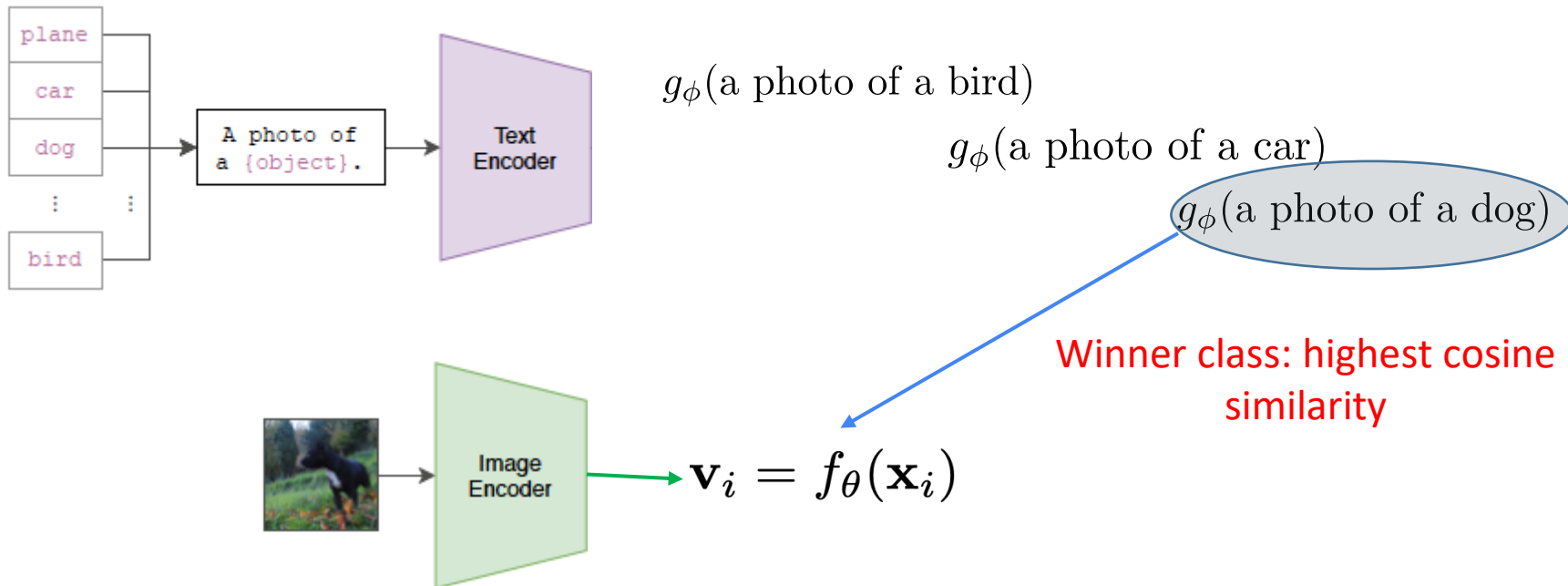
“Zero-shot” prediction for new tasks

create “class prototypes”
From label text

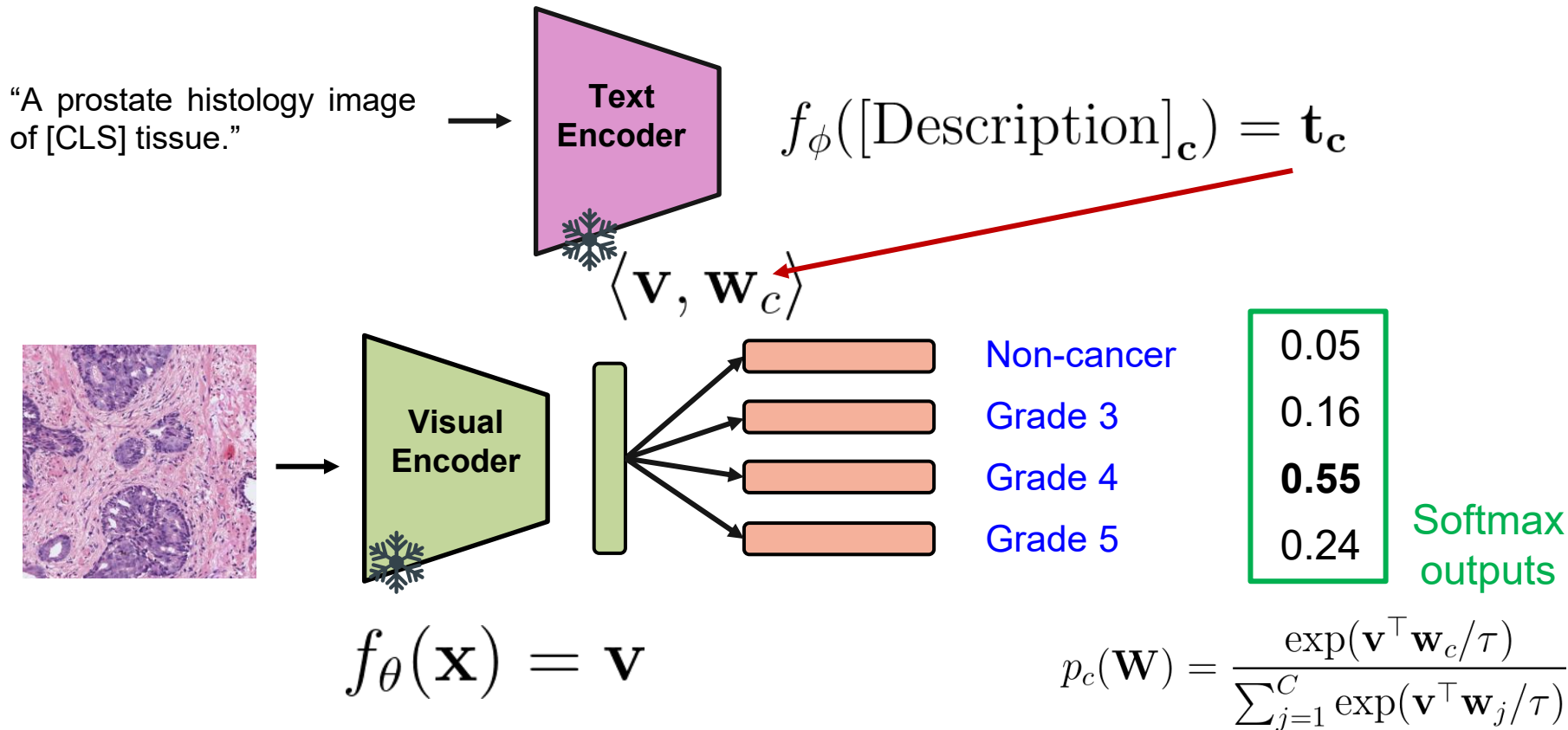


“Zero-shot” prediction for new tasks

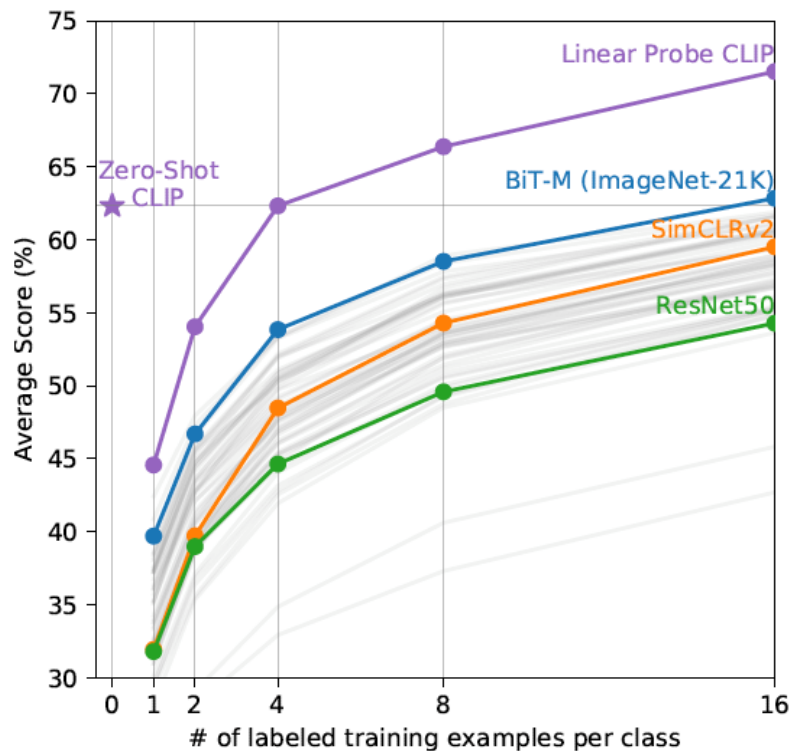
create “class prototypes”
From label text



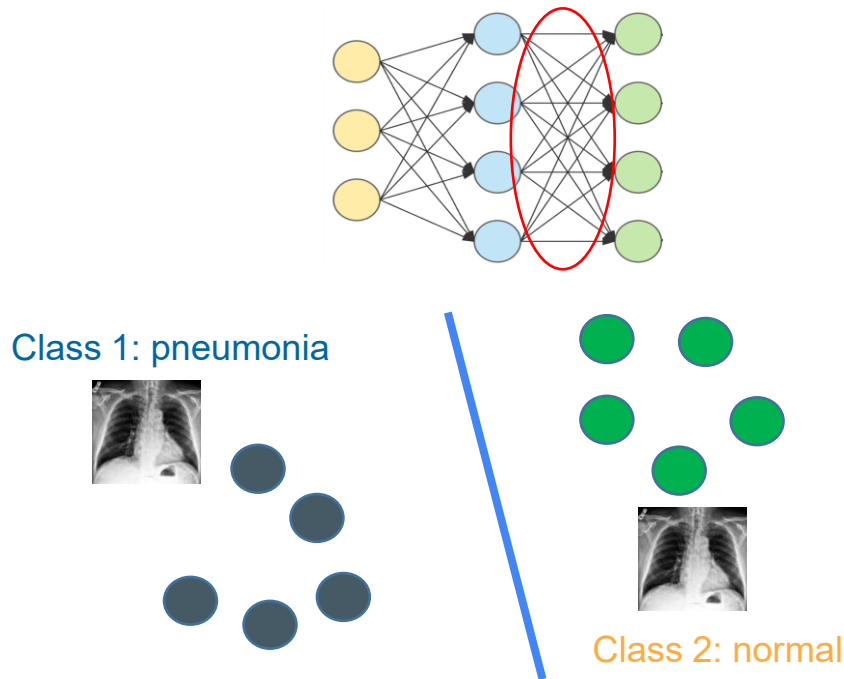
“Zero-shot” prediction for new tasks



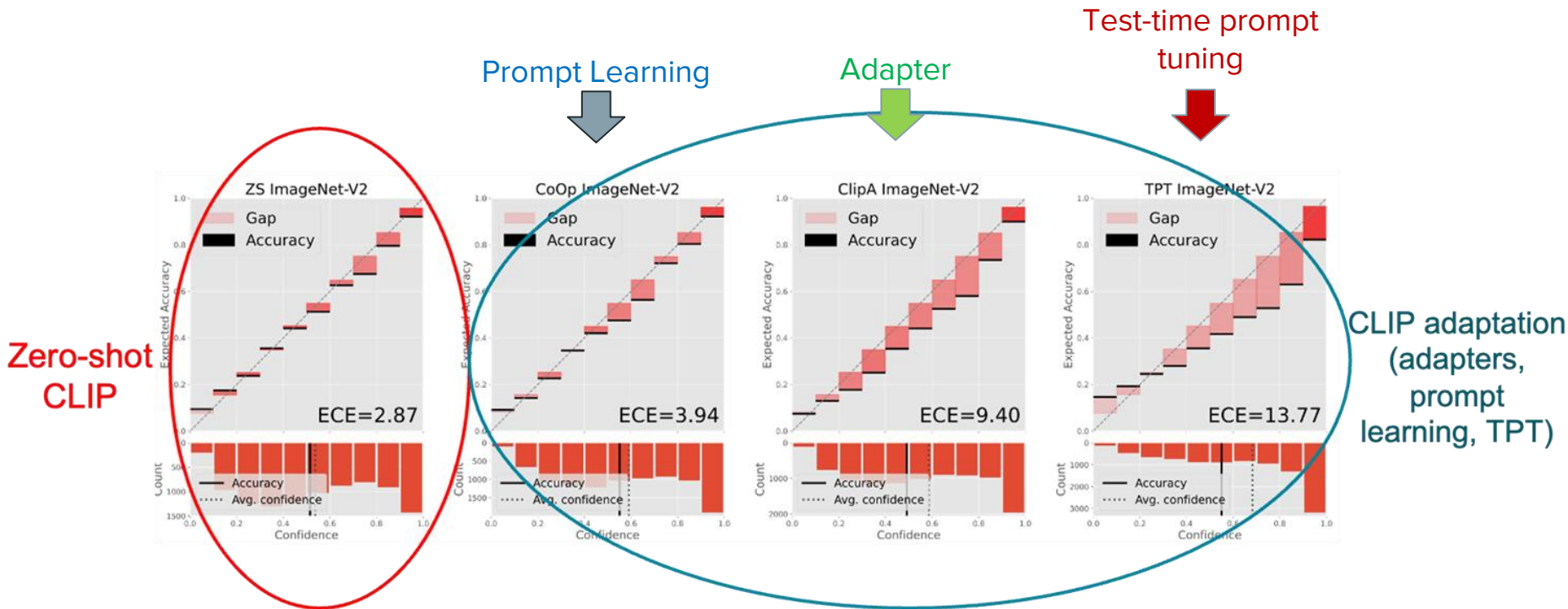
Beyond zero-shot: Few-shot generalization



Few-shot fine-tuning (linear probe)



Calibration in contrastive VLMs

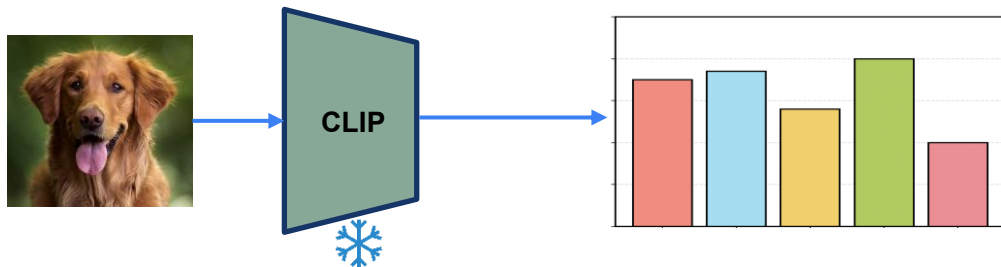


Calibration in contrastive VLMs

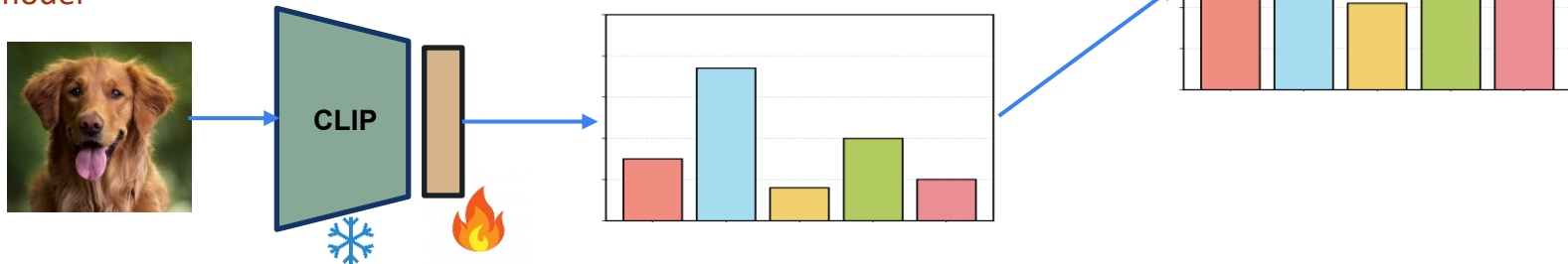
minimize $\mathcal{H}(Y, P)$

subject to $l_i^{\text{ZS-min}} \mathbf{1} \leq l_i \leq l_i^{\text{ZS-max}} \mathbf{1} \quad \forall i \in \mathcal{D},$

1: **Zero-shot** prediction

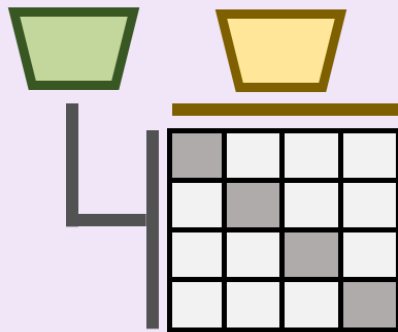


2: **Adapt** the model

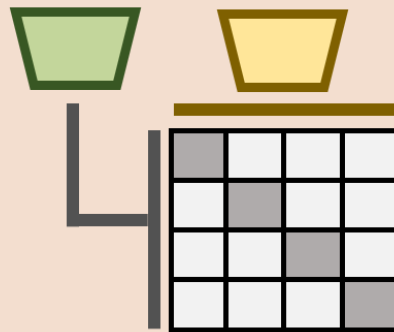


Medical Vision-Language Models

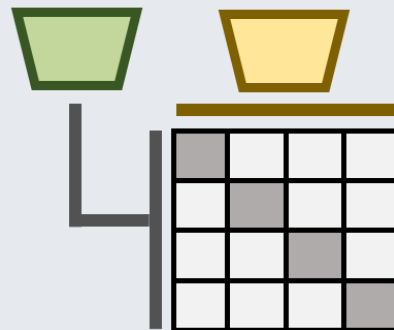
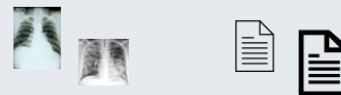
CONCH



FLAIR

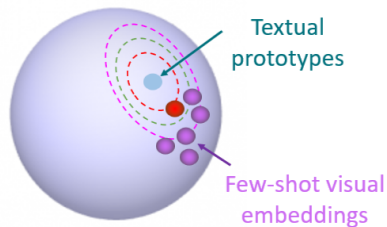
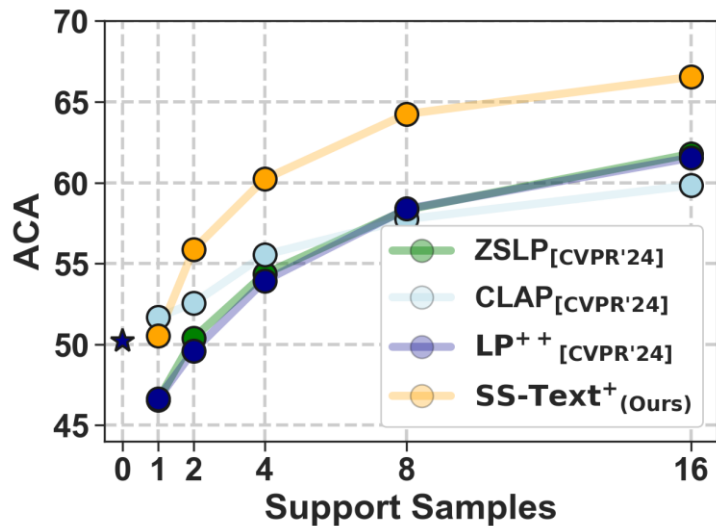


CONVIRT



Medical Vision-Language Models

$$\min_{\mathbf{W}} \mathcal{L}(\mathbf{W}) = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \ln(p_{ic}(\mathbf{W})) + \frac{\lambda}{2} \sum_{c=1}^C \|\mathbf{w}_c - \mathbf{t}_c\|_2^2.$$



SS-Text solver:

$$w_c^* = \frac{1}{\lambda N \tau} \sum_{i=1}^N y_{ic} \mathbf{v}_i + \mathbf{t}_c$$

$\lambda = 1/(K_c \tau)$

Outline

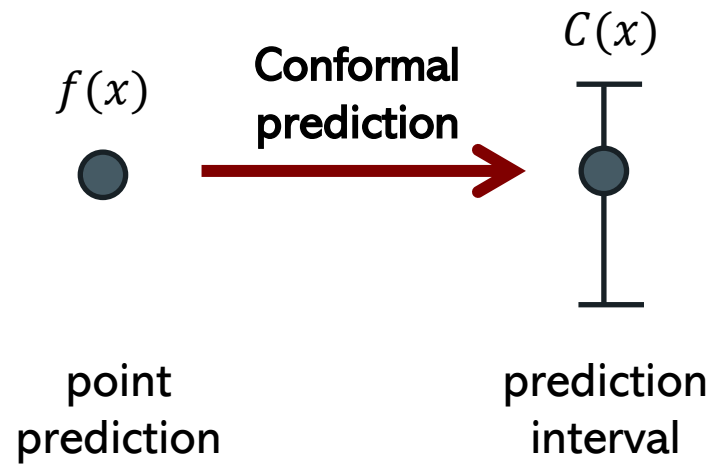
A. Vision-Language Models (VLMs)

- Contrastive Language-Image Pre-training (CLIP).
- Zero-shot and few-shot inference.
- Calibration in contrastive VLMs.
- Vision-language models for medical imaging.

B. Conformal Prediction in VLMs

- Split conformal prediction (SCP).
- Theoretical guarantees in CP.
- Benefits of foundation models for CP.
- Full conformal predictors.
- Full conformal adaptation (FCA).
- Interpretability of conformal sets.

Conformal Prediction (CP)

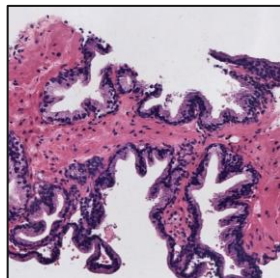


Conformal Prediction (CP)

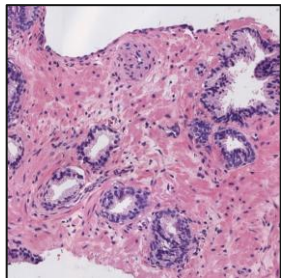
$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

Sets

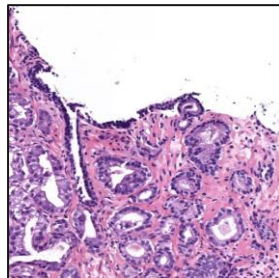
Error rate,
e.g., 10%



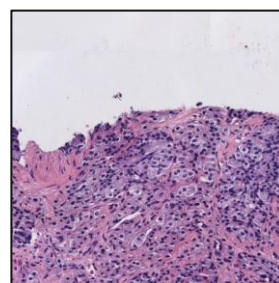
GT: NC
Set: [NC]



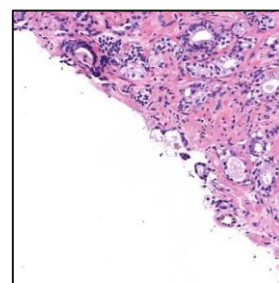
GT: G3
Set: [G3]



GT: G3
Set: [G3,G4]



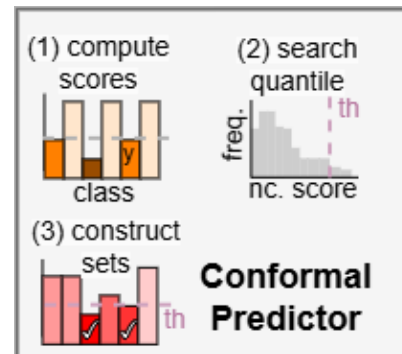
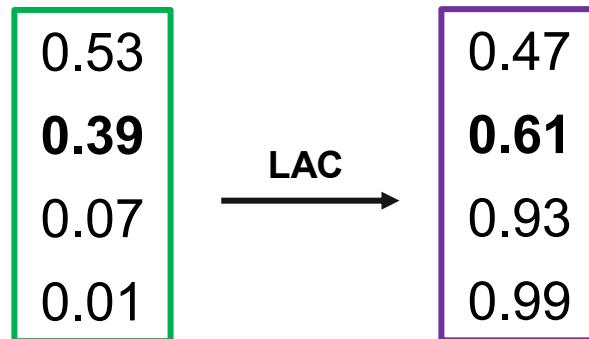
GT: G5
Set: [G5]



GT: G5
Set: [G3,G4,G5]

Split Conformal Prediction (SCP)

1. Non-conformity score from black-box classifier. $\mathcal{S}(\mathbf{x}, y) = 1 - \hat{p}_{k=y}$

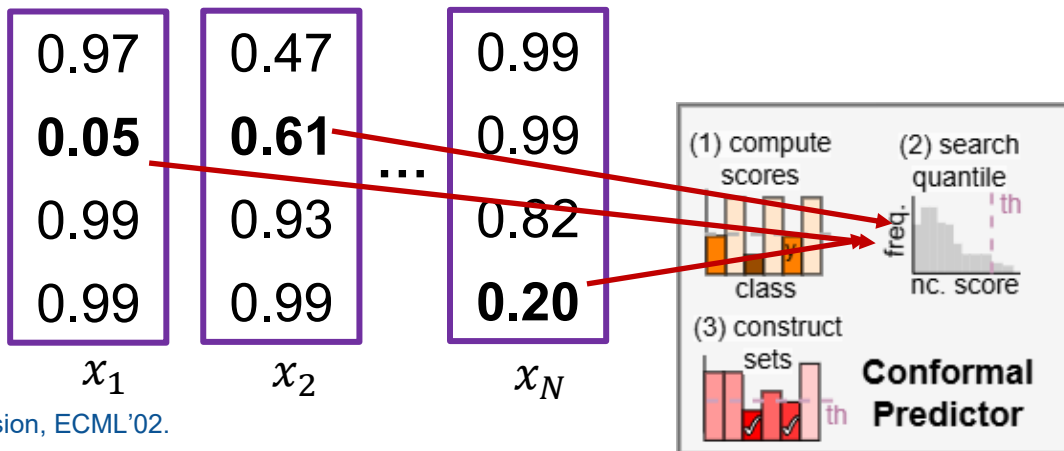


Split Conformal Prediction (SCP)

1. Non-conformity score from black-box classifier. $\mathcal{S}(\mathbf{x}, y) = 1 - \hat{p}_{k=y}$
2. Search threshold in the **true-label** $\mathcal{S}(\mathbf{x}, y)$ distribution that ensures a given coverage.



Calibration set (labeled)



Split Conformal Prediction (SCP)

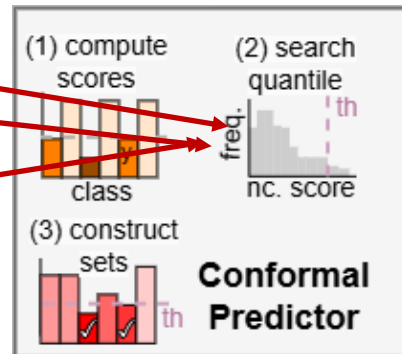
1. Non-conformity score from black-box classifier. $\mathcal{S}(\mathbf{x}, y) = 1 - \hat{p}_{k=y}$
2. Search threshold in the **true-label** $\mathcal{S}(\mathbf{x}, y)$ distribution that ensures a given coverage.



$$\hat{s} = \inf \left[s : \frac{|\{i \in \{1, \dots, N\} : s_i \leq s\}|}{N} \geq \frac{\lceil (N+1)(1-\alpha) \rceil}{N} \right]$$

Calibration set (labeled)

0.97	0.47	0.99
0.05	0.61	0.99
0.99	0.93	0.82
0.99	0.99	0.20
x_1	x_2	x_N



Split Conformal Prediction (SCP)

1. Non-conformity score from black-box classifier. $\mathcal{S}(\mathbf{x}, y) = 1 - \hat{p}_{k=y}$
2. Search threshold in the **true-label** $\mathcal{S}(\mathbf{x}, y)$ distribution that ensures a given coverage.



$$\hat{s} = \inf \left[s : \frac{|\{i \in \{1, \dots, N\} : s_i \leq s\}|}{N} \geq \frac{\lceil (N+1)(1-\alpha) \rceil}{N} \right]$$

Calibration set (labeled)

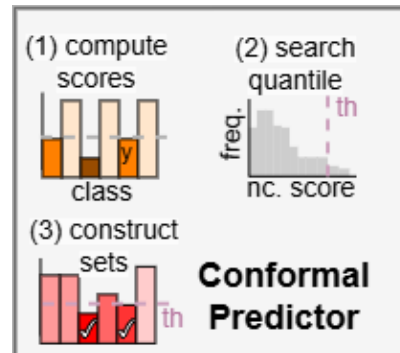
3. Create sets using the threshold as rejection criteria.



$$\mathcal{C}(\mathbf{x}) = \{y \in \mathcal{Y} : \mathcal{S}(\mathbf{x}, y) \leq \hat{s}\}$$

e.g., 0.43

Test set (unlabeled)



Theoretical guarantees in conformal prediction

$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

Coverage guarantee

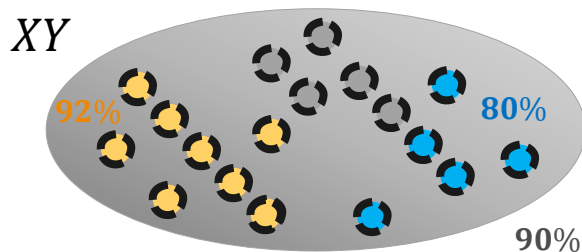
1. **Distribution-free**, e.g., no gaussian distribution required.
2. **Marginal over XY** , i.e., does not inform about specific examples/subgroups.
3. Assumes at least **exchangeability** of D_{cal} and D_{test} .
4. **Finite-sample guarantee** – holds on average across random experiments.

Theoretical guarantees in conformal prediction

$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

Coverage guarantee

1. **Distribution-free**, e.g., no gaussian distribution required.
2. **Marginal over XY** , i.e., does not inform about specific examples/subgroups.
3. Assumes at least **exchangeability** of D_{cal} and D_{test} .
4. **Finite-sample guarantee** – holds on average across random experiments.

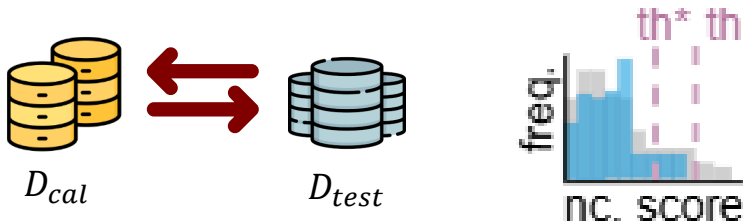


Theoretical guarantees in conformal prediction

$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

Coverage guarantee

1. **Distribution-free**, e.g., no gaussian distribution required.
2. **Marginal over XY** , i.e., does not inform about specific examples/subgroups.
3. Assumes at least **exchangeability** of D_{cal} and D_{test} .
4. **Finite-sample guarantee** – holds on average across random experiments.

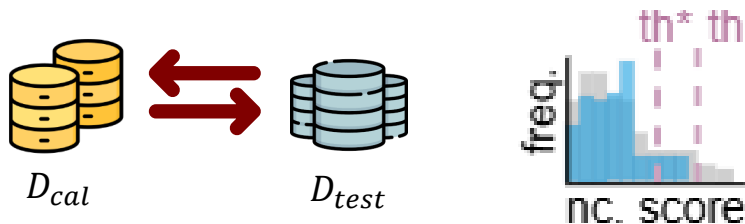


Theoretical guarantees in conformal prediction

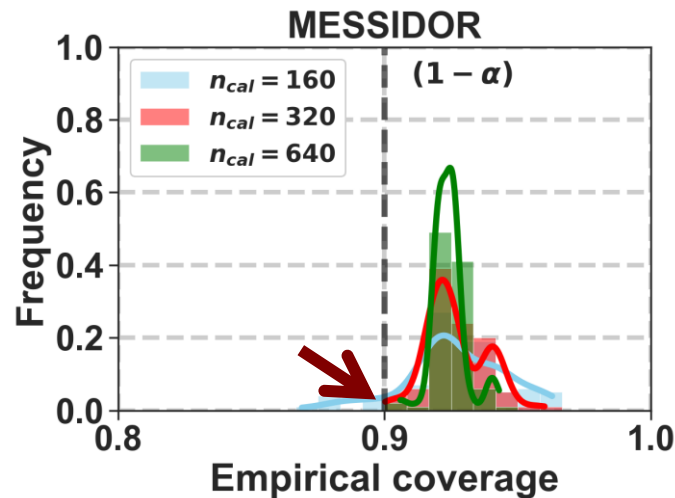
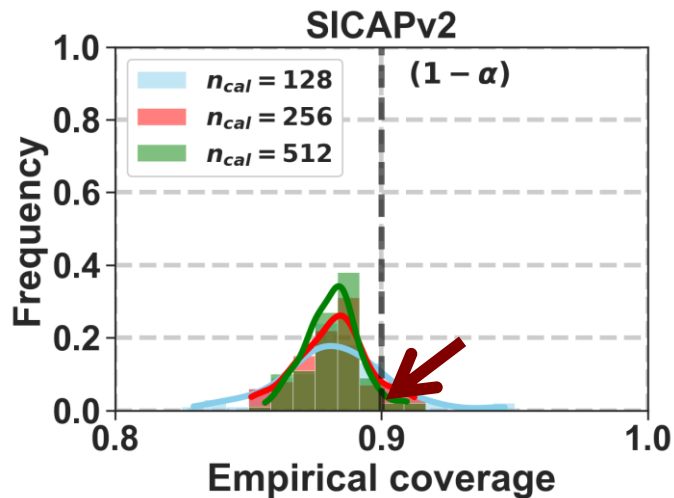
$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

Coverage guarantee

1. **Distribution-free**, e.g., no gaussian distribution required.
2. **Marginal over XY** , i.e., does not inform about specific examples/subgroups.
3. Assumes at least **exchangeability** of D_{cal} and D_{test} .
4. **Finite-sample guarantee** – holds on average across random experiments.



Theoretical guarantees in conformal prediction



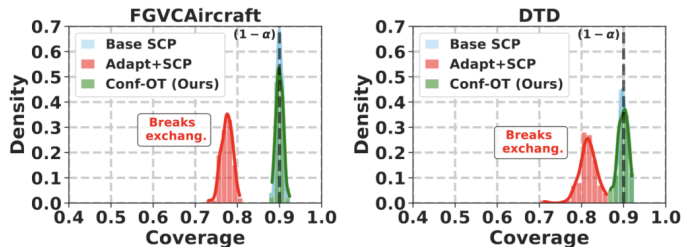
Are data samples coming from different patients necessary exchangeable?

Theoretical guarantees in conformal prediction

$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

Coverage guarantee

1. **Distribution-free**, e.g., no gaussian distribution required.
2. **Marginal over XY** , i.e., does not inform about specific examples/subgroups.
3. Assumes at least **exchangeability** of D_{cal} and D_{test} .
4. **Finite-sample guarantee** – holds on average across random experiments.



Theoretical guarantees in conformal prediction

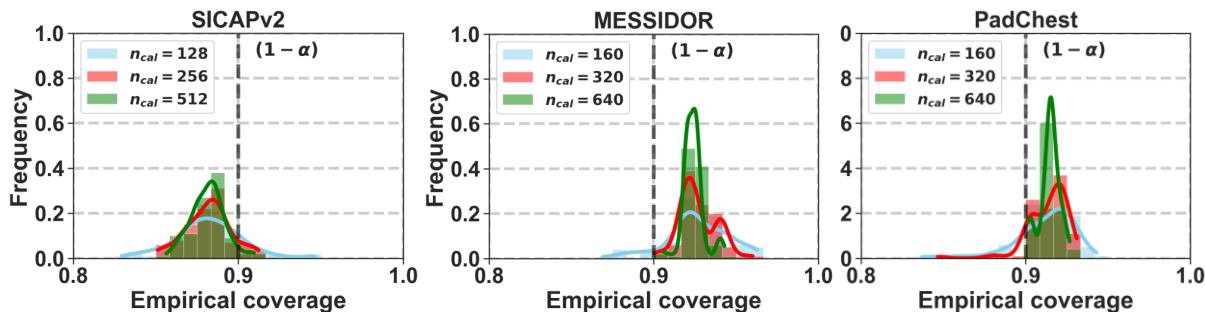
$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

Marques F. Universal distribution of the empirical coverage in split conformal prediction, [ArXiv'24](#).

Theorem 1. Under the data exchangeability assumption, for a regular conformity function, the sequence of coverage indicators $\{Z_i\}_{i \geq 1}$ is exchangeable and $m \times C_m^{(n, \alpha)}$ is distributed as a Beta-Binomial($\lceil (1 - \alpha)(n + 1) \rceil, \lfloor \alpha(n + 1) \rfloor$) random variable, to the effect that the distribution of the empirical coverage is given by

$$P\left(C_m^{(n, \alpha)} = \frac{k}{m}\right) = \binom{m}{k} \frac{n! (k + \lceil (1 - \alpha)(n + 1) \rceil - 1)! (m - k + \lfloor \alpha(n + 1) \rfloor - 1)!}{(\lceil (1 - \alpha)(n + 1) \rceil - 1)! (\lfloor \alpha(n + 1) \rfloor - 1)! (m + n)!},$$

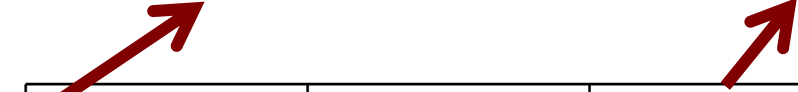
Theorem 2. Under the data exchangeability assumption, for a regular conformity function, the empirical coverage $C_m^{(n, \alpha)}$ converges almost surely, when the future batch size tends to infinity, to a random variable $C_\infty^{(n, \alpha)}$ with distribution $\text{Beta}(\lceil (1 - \alpha)(n + 1) \rceil, \lfloor \alpha(n + 1) \rfloor)$.



Theoretical guarantees in conformal prediction

$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

*“I guarantee you predictive sets with **coverage of 90%**,
with a **probability of at least 95%**, and a **1% tolerance error**”*



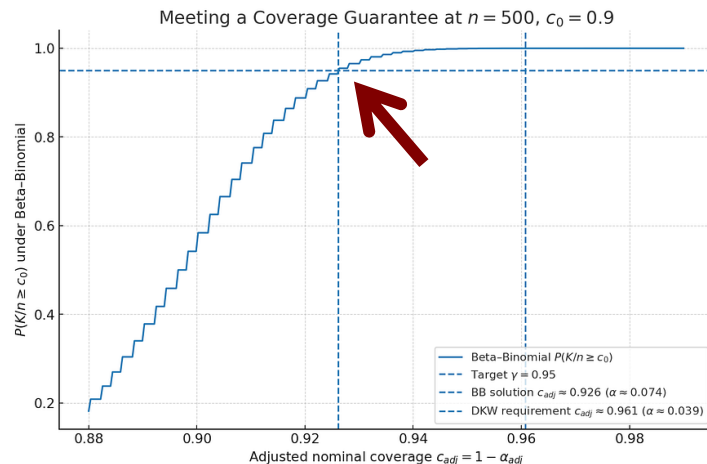
$1 - \alpha \backslash \tau$	$\epsilon = 0.1$			$\epsilon = 0.05$			$\epsilon = 0.01$			$\epsilon = 0.005$		
	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
80%	40	57	98	170	241	418	4,326	6,142	10,611	17,314	24,581	42,457
85%	30	42	77	134	189	330	3,446	4,893	8,451	13,794	19,587	33,830
90%	11	14	47	90	128	227	2,429	3,448	5,958	9,733	13,821	23,875
95%	19	19	29	22	29	97	1,270	1,806	3,132	5,125	7,278	12,578

Marques F. Universal distribution of the empirical coverage in split conformal prediction, [ArXiv'24](#).

Theoretical guarantees in conformal prediction

$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

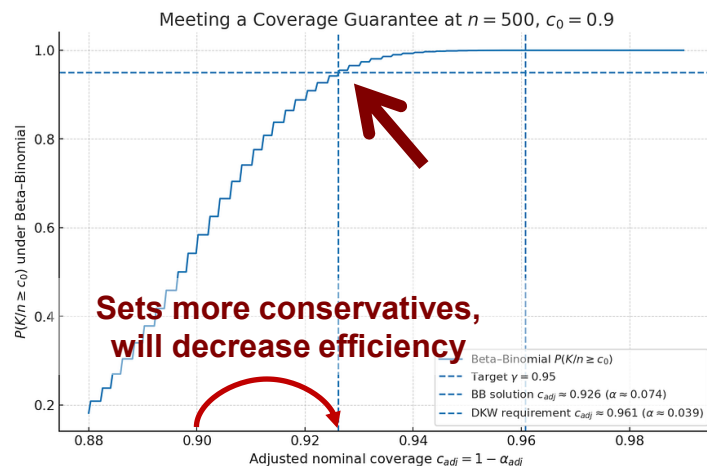
“I want predictive sets with coverage above 90%, with a probability of at least 95%, and a 1% tolerance error. I have N=500 calibration samples: which nominal coverage should I use?”



Theoretical guarantees in conformal prediction

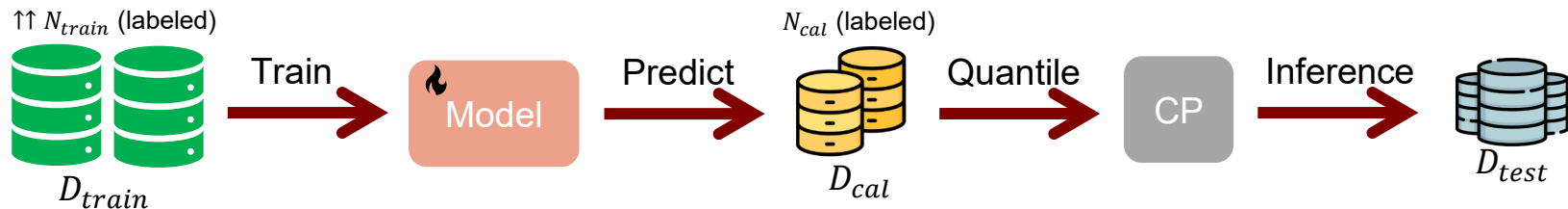
$$\mathcal{P}(Y \in C(\mathbf{x})) \geq 1 - \alpha$$

“I want predictive sets with coverage above 90%, with a probability of at least 95%, and a 1% tolerance error. I have N=500 calibration samples: which nominal coverage should I use?”



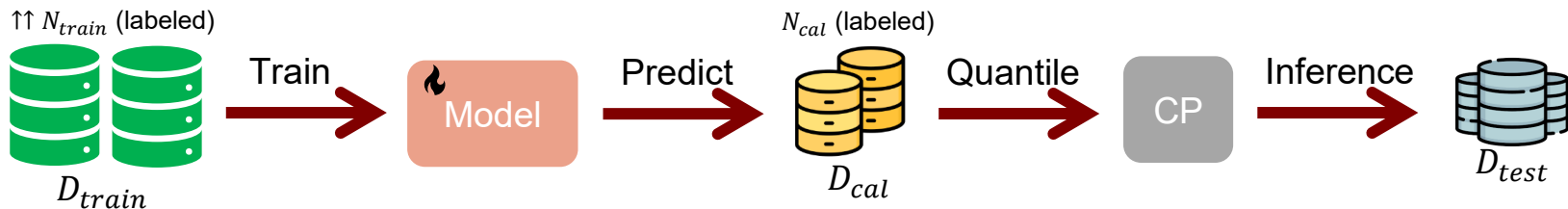
Foundation models and conformal prediction

1. Standard task-specific training scenario.

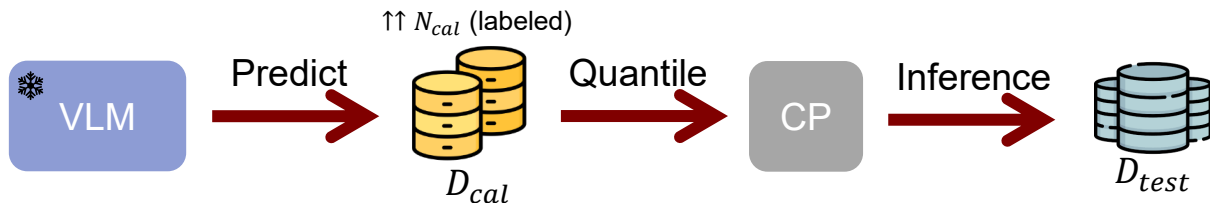


Foundation models and conformal prediction

1. Standard task-specific training scenario.

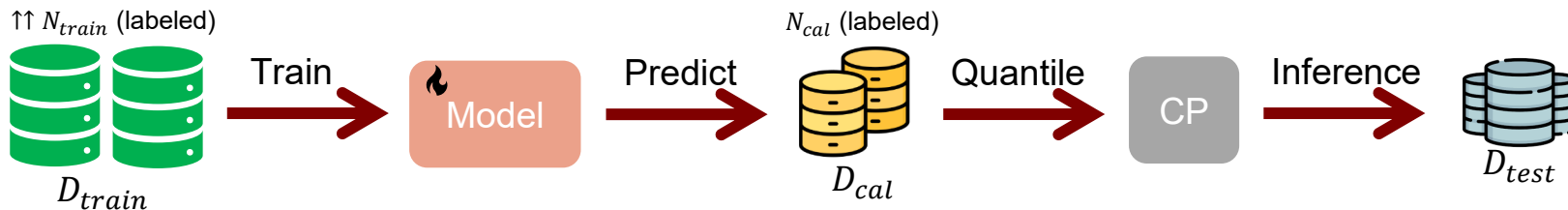


2. Modern scenario with zero-shot VLMs.

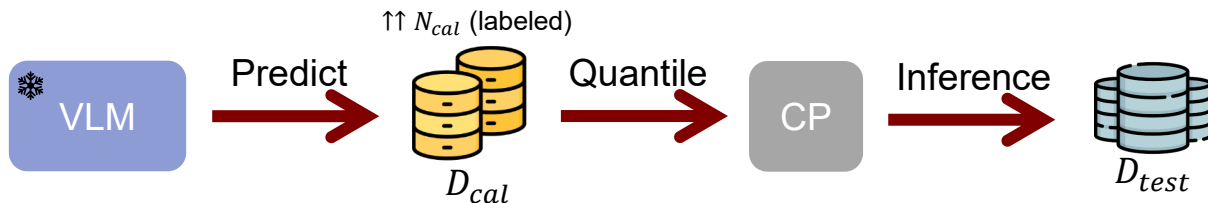


Foundation models and conformal prediction

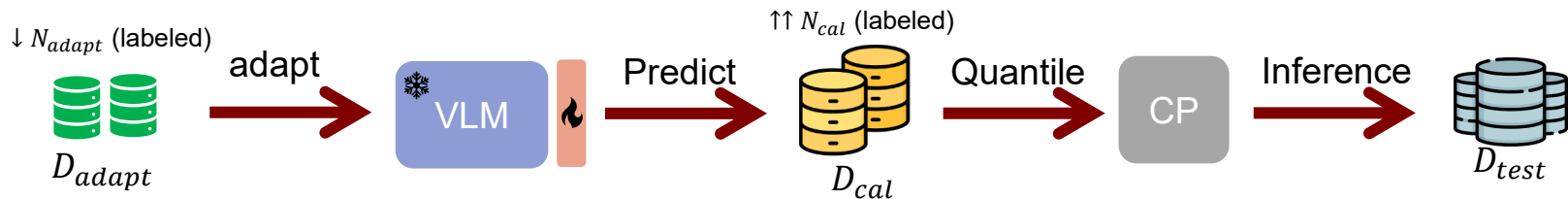
1. Standard task-specific training scenario.



2. Modern scenario with zero-shot VLMs.

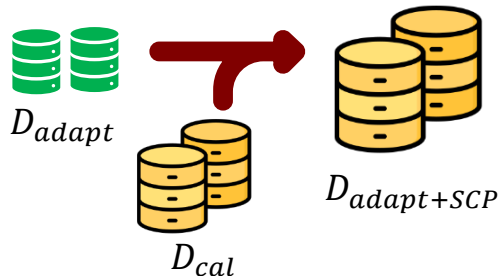


3. Modern scenario adapting foundation models.



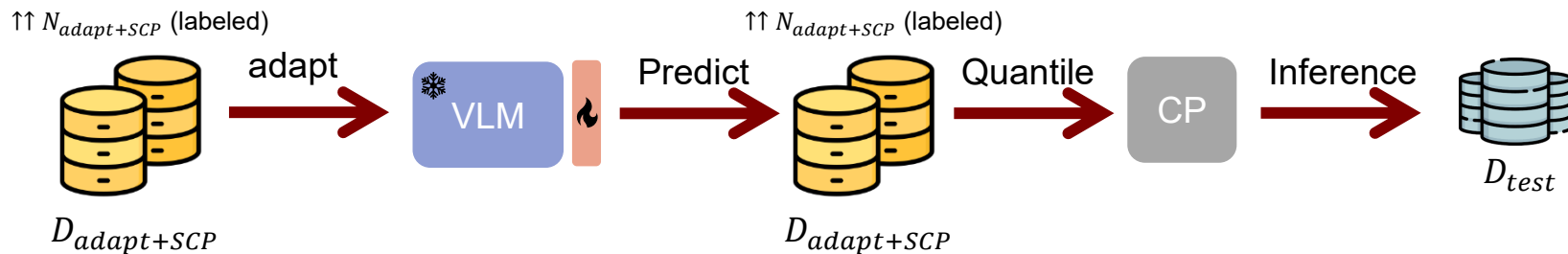
Foundation models and conformal prediction

“For enhanced **data-efficiency**, could you **adapt** and then search the CP quantile using the same **joint adapt and calibration data**?”



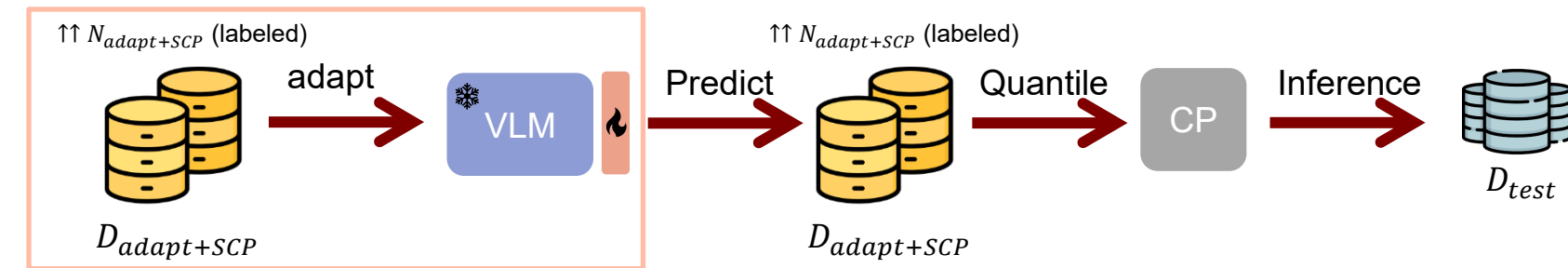
Foundation models and conformal prediction

“For enhanced **data-efficiency**, could you **adapt** and then search the **CP** quantile using the same **joint adapt and calibration data**?”

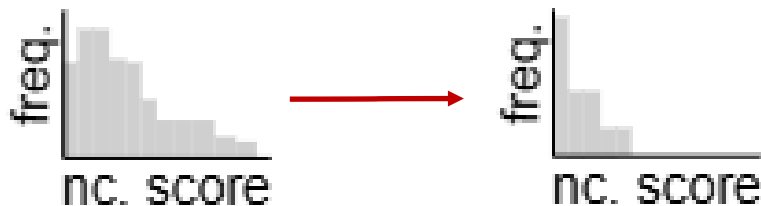


Foundation models and conformal prediction

“For enhanced **data-efficiency**, could you **adapt** and then search the CP quantile using the same **joint adapt and calibration data**?”

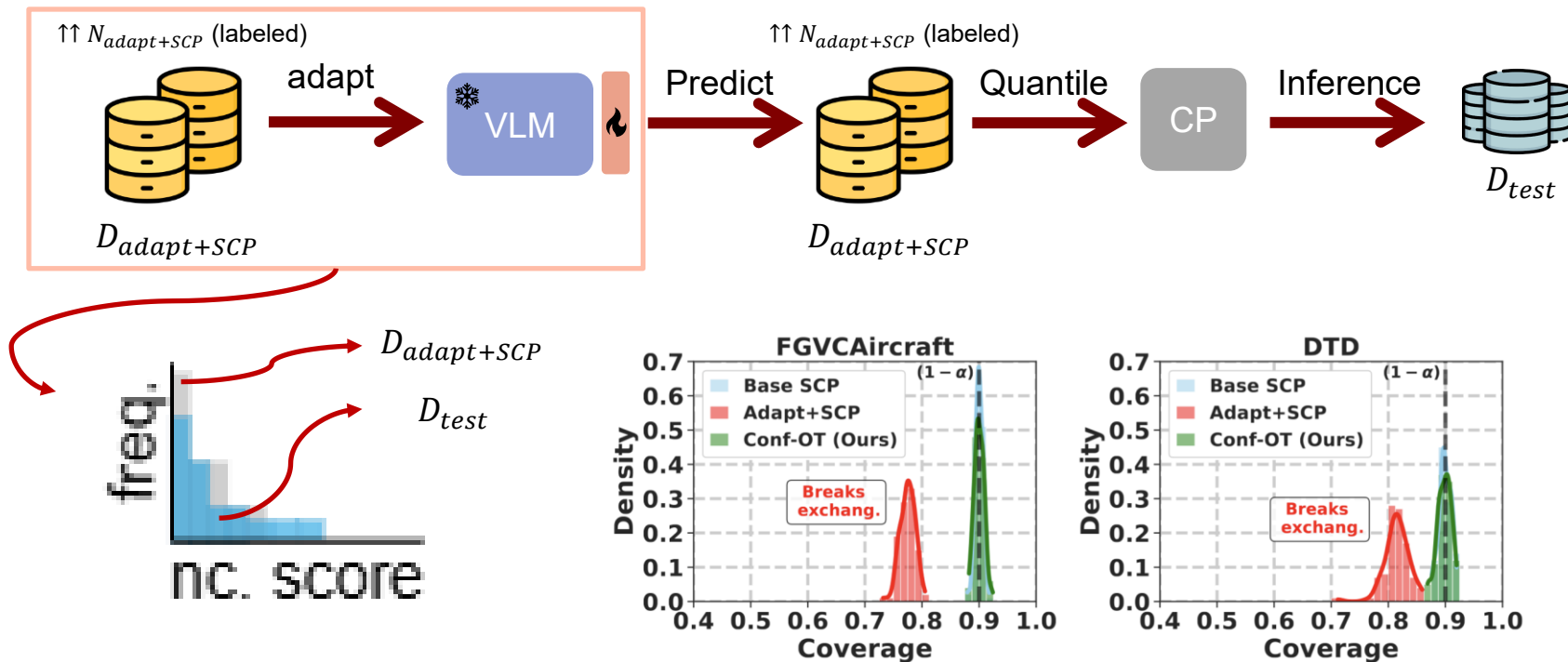


$$\min_{\mathbf{W}} \mathcal{L}(\mathbf{W}) = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \ln(p_{ic}(\mathbf{W})) + \frac{\lambda}{2} \sum_{c=1}^C \|\mathbf{w}_c - \mathbf{t}_c\|_2^2.$$



Foundation models and conformal prediction

“For enhanced **data-efficiency**, could you **adapt** and then search the CP quantile using the same **joint adapt and calibration data**?”



Full Conformal Prediction

There is life beyond vanilla split conformal predictors!!!

Transduction with Confidence and Credibility

C. Saunders, A. Gamerman, V. Vovk
Royal Holloway, University of London
Egham, Surrey, England.
{craig,alex,vovk}@dcs.rhbnc.ac.uk

Saunders et al. Transduction with Confidence and Credibility, IJCAI'99.

*Algorithmic Learning
in a Random World*

Vladimir Vovk
University of London
Egham, United Kingdom

Vovk et al. Algorithmic learning in a random world, Springer'05.

Full Conformal Prediction

$$\underbrace{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N)}_{\mathcal{D}_{train}}, (\mathbf{x}_{N+1}, ?)$$

Full Conformal Prediction

$$\underbrace{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N)}_{\mathcal{D}_{train}}, (\mathbf{x}_{N+1}, ?)$$

- 1) *We know that, for a **test sample**, the **true label** of a test point lies somewhere on the label space.*
- 2) *Let's **fit the model** with each label assignment and check if the **errors** on the test point **conform** to the training observations.*

Full Conformal Prediction

A: For each test data point...

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N), (\mathbf{x}_{N+1}, ?)$$

B: For each label...

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N), \boxed{(\mathbf{x}_{N+1}, y)}$$

Full Conformal Prediction

A: For each test data point...

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N), (\mathbf{x}_{N+1}, ?)$$

B: For each label...

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N), (\mathbf{x}_{N+1}, y)$$

1. Train model on joint dataset

$$\boxed{\pi(\cdot)^y}: y_{N+1} = y \in \mathcal{Y}$$

Full Conformal Prediction

A: For each test data point...

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N), (\mathbf{x}_{N+1}, ?)$$

B: For each label...

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N), (\mathbf{x}_{N+1}, y)$$

1. Train model on joint dataset

$$\pi(\cdot)^y : y_{N+1} = y \in \mathcal{Y}$$

2. Search quantile in training data

$$s_i^y = \mathcal{S}[\pi_i^y(\mathbf{x}), y_i]$$

3. Accept/Reject label

$$\mathcal{C}(\mathbf{x}) = \{y \in \mathcal{Y} : s^y \leq \hat{s}^y\}$$

Full Conformal Prediction

A: For each test data point...

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N), (\mathbf{x}_{N+1}, ?)$$

B: For each label...

$$(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_i, y_i), \dots, (\mathbf{x}_N, y_N), (\mathbf{x}_{N+1}, y)$$

1. Train model on joint dataset

$$\boxed{\pi(\cdot)^y} : y_{N+1} = y \in \mathcal{Y}$$

Training a model for each test sample and label combination. Computationally unfeasible

Full Conformal Adaptation (FCA)

A: For each test data point...

$(\mathbf{v}_1, y_1), \dots, (\mathbf{v}_i, y_i), \dots, (\mathbf{v}_N, y_N), (\mathbf{v}_{N+1}, ?)$

Pre-computed embeddings from a foundation model.

B: For each label...

$(\mathbf{v}_1, y_1), \dots, (\mathbf{v}_i, y_i), \dots, (\mathbf{v}_N, y_N), (\mathbf{v}_{N+1}, y)$

1. Adapt the model on joint dataset

$p(\mathbf{W}^*, \cdot)^y \vdash y_{N+1} = y \in \mathcal{Y}$

Leveraging efficient linear probing solvers, the adaptation phase takes few milliseconds

Full Conformal Adaptation (FCA)

A: For each test data point...

$$(\mathbf{v}_1, y_1), \dots, (\mathbf{v}_i, y_i), \dots, (\mathbf{v}_N, y_N), (\mathbf{v}_{N+1}, ?)$$

B: For each label...

$$(\mathbf{v}_1, y_1), \dots, (\mathbf{v}_i, y_i), \dots, (\mathbf{v}_N, y_N), (\mathbf{v}_{N+1}, y)$$

1. Adapt the model on joint dataset

$$p(\mathbf{W}^*, \cdot)^y : y_{N+1} = y \in \mathcal{Y}$$

2. Search quantile in training data

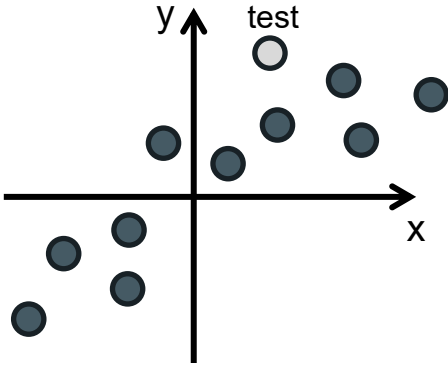
$$s_i^y = \mathcal{S} \left[p(\mathbf{W}^*, \mathbf{v}_i)^y \mid y_i \right]$$

3. Accept/Reject label

$$\mathcal{C}(\mathbf{x}) = \{y \in \mathcal{Y} : s^y \leq \hat{s}^y\}$$

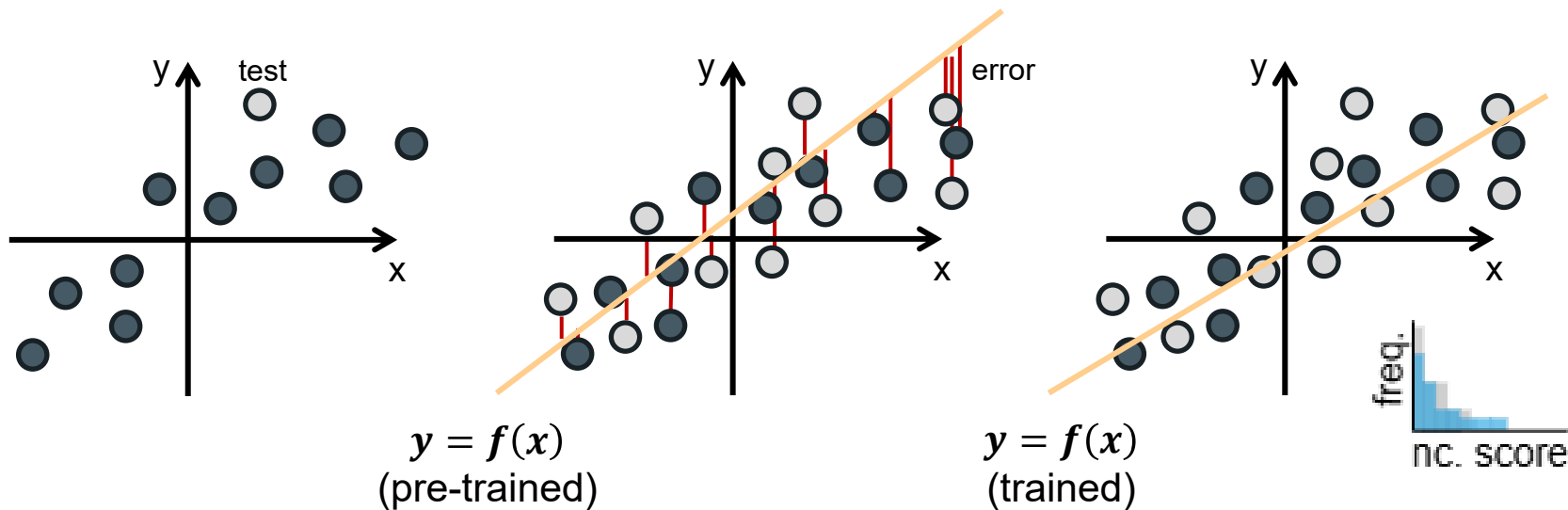
Full Conformal Prediction

Interpretation: Why does it work?



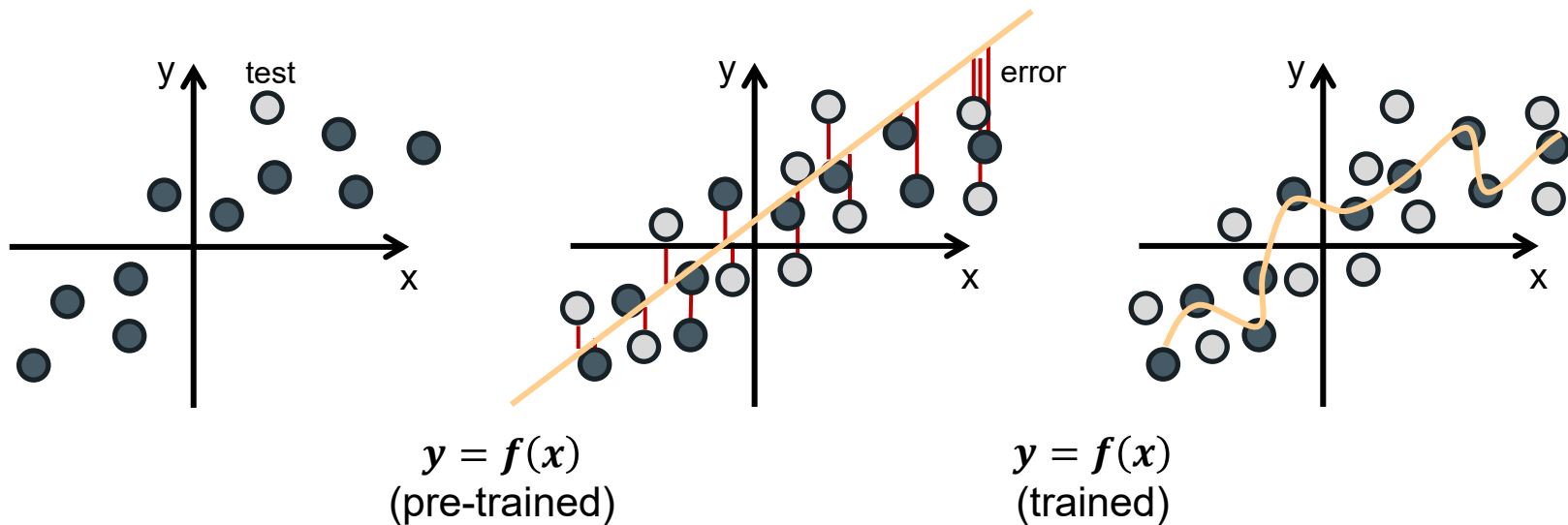
Full Conformal Prediction

Interpretation: Why does it work?



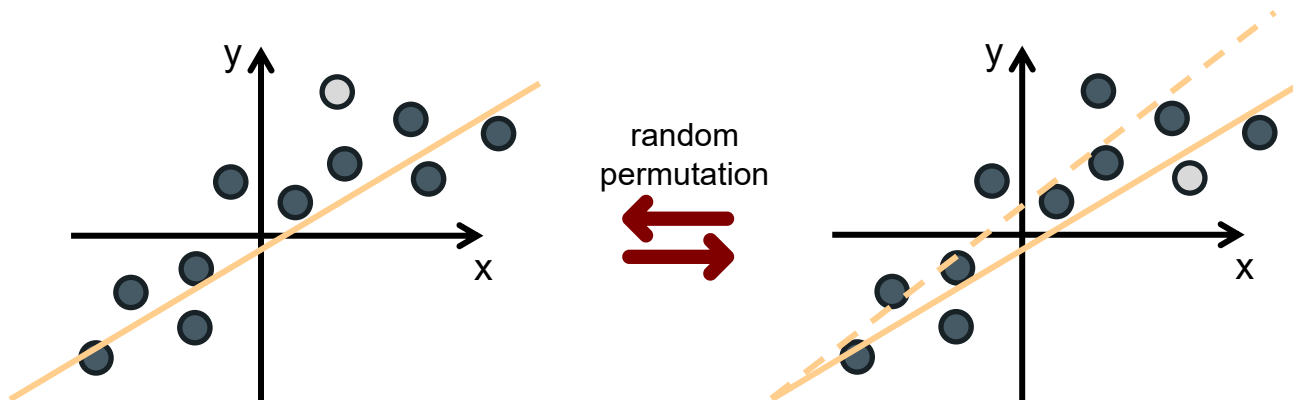
Full Conformal Prediction

Interpretation: Why does it work?



Full Conformal Prediction

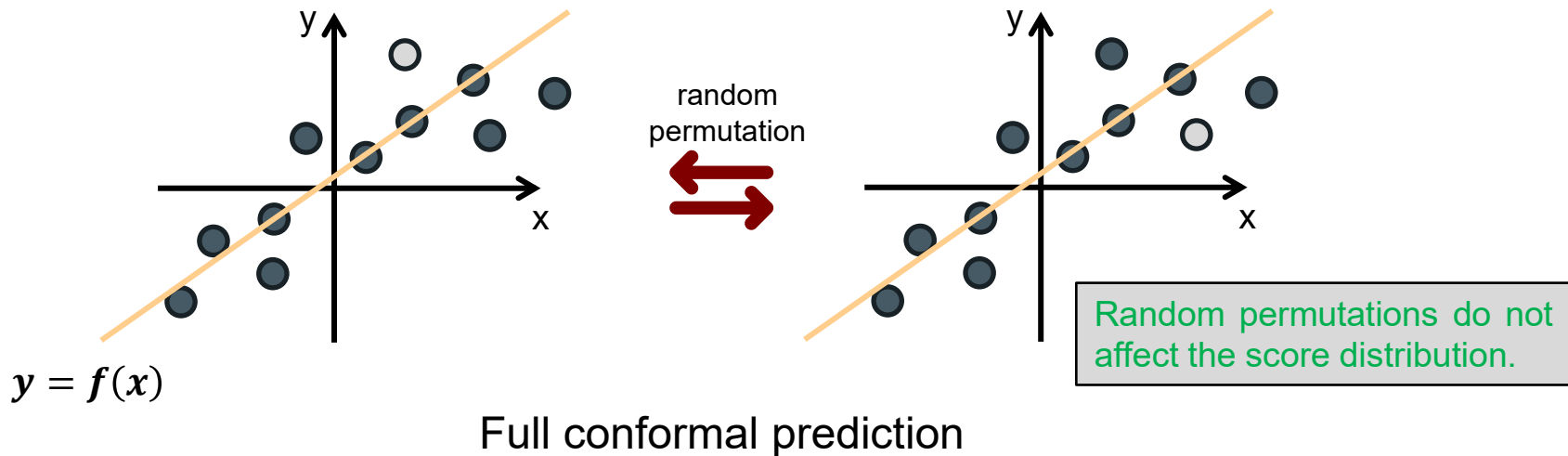
A fast trick to check exchangeability of the pipeline



Adapt + Split Conformal Prediction

Full Conformal Prediction

A fast trick to check exchangeability of the pipeline

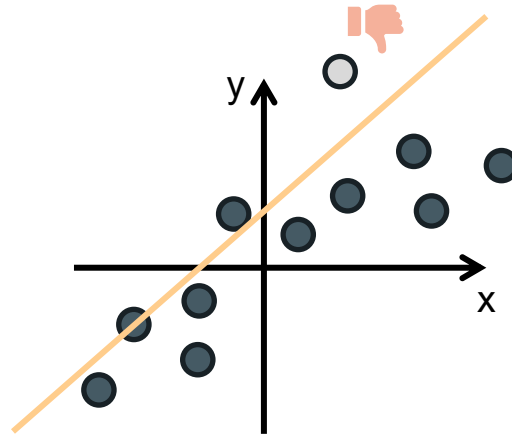


The model is trained using the test point:

$$(\mathbf{v}_1, y_1), \dots, (\mathbf{v}_i, y_i), \dots, (\mathbf{v}_N, y_N), (\mathbf{v}_{N+1}, y)$$

Full Conformal Prediction

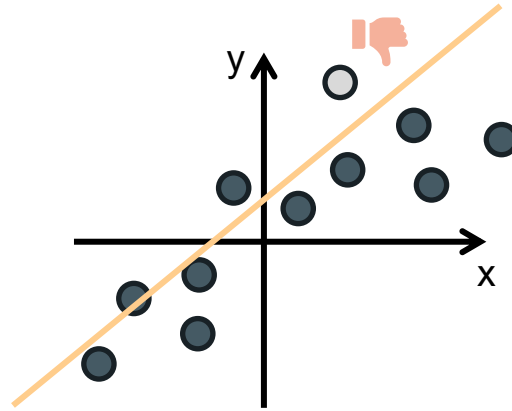
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

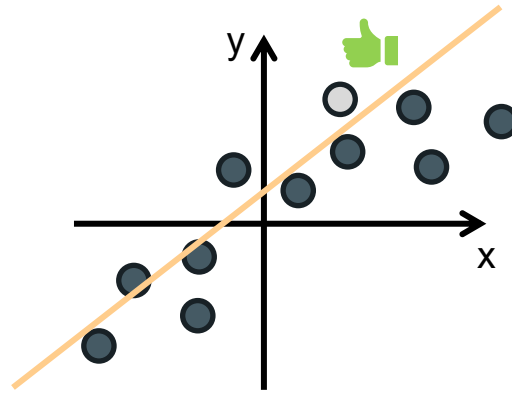
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

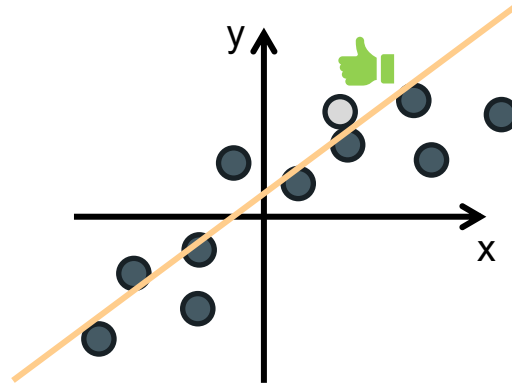
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

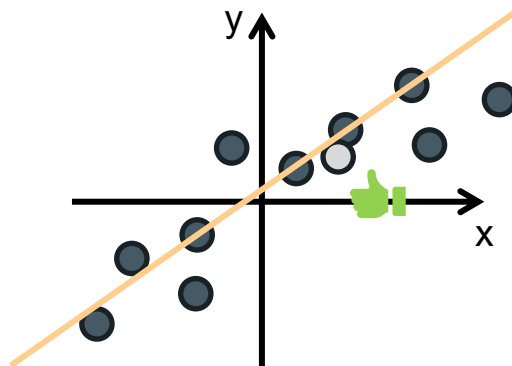
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

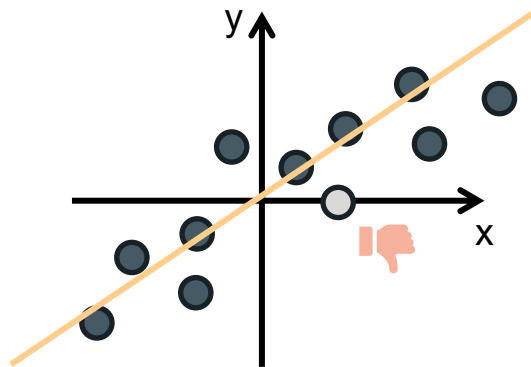
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

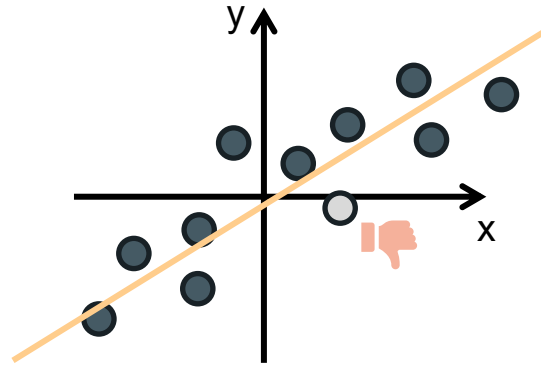
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

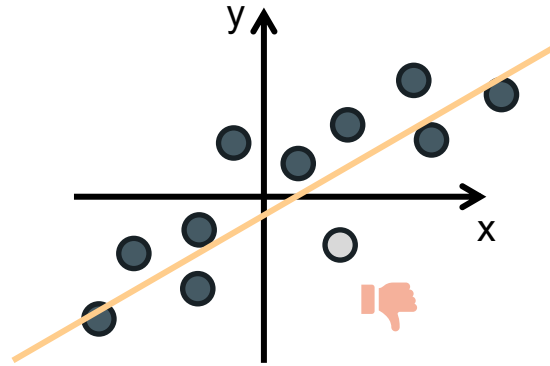
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

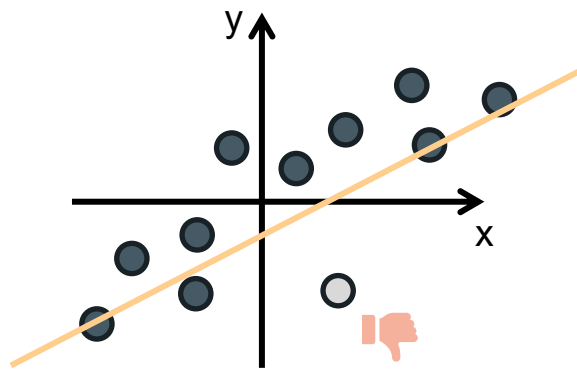
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

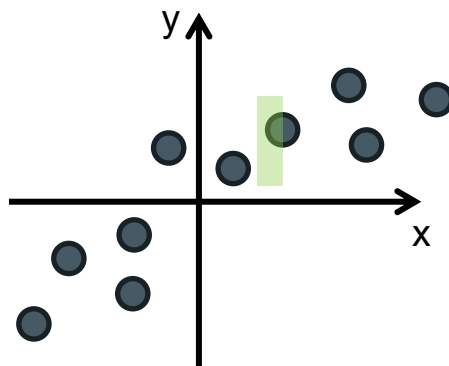
Full conformal prediction loop



Full conformal prediction

Full Conformal Prediction

Full conformal prediction loop



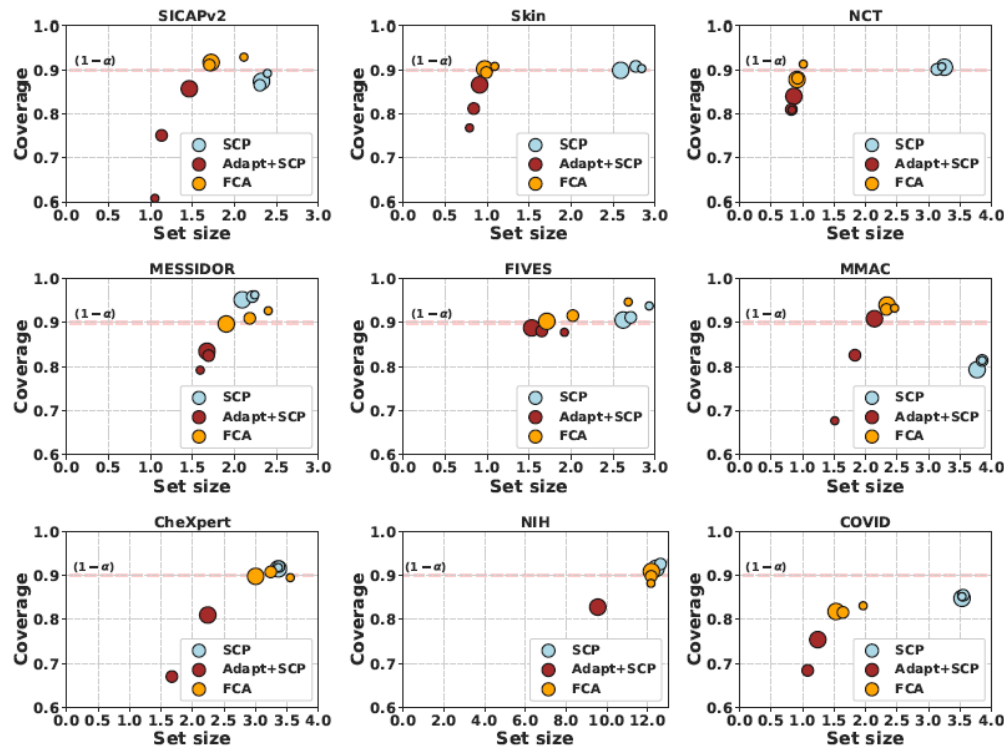
Full conformal prediction

Full Conformal Adaptation (FCA)

Method		$\alpha = 0.10$			
		ACA \uparrow	Cov.	Size \downarrow	CCV \downarrow
LAC	SCP	50.2	0.890	3.99	9.96
	Adapt+SCP	67.1 _{+16.9}	0.842	2.40 _{-1.59}	11.17 _{+1.21}
	FCA (<i>Ours</i>)	67.1 _{+16.9}	0.896	2.91 _{-1.08}	8.38 _{-1.58}

Method		$\alpha = 0.05$			
		ACA \uparrow	Cov.	Size \downarrow	CCV \downarrow
LAC	SCP	50.2	0.951	4.88	5.68
	Adapt+SCP	67.1 _{+16.9}	0.921	3.07 _{-1.81}	6.87 _{+1.19}
	FCA (<i>Ours</i>)	67.1 _{+16.9}	0.952	3.56 _{-1.32}	5.02 _{-0.66}

Average performance across tasks
(from 4 until 20 categories, 8 in average)



What about non-conformity scores?

$$\mathcal{S}_{\text{LAC}}(\mathbf{x}, y) = 1 - p_{k=y}$$

Sadinle et al., Least ambiguous set-valued classifiers with bounded error levels, Jour. American Statistical Association 2019

$$\mathcal{S}_{\text{APS}}(\mathbf{x}, y) = \rho_x(y) + p_{k=y} \cdot u$$

Romano et al., Classification with valid and adaptive coverage., NeurIPS 2020

$$\mathcal{S}_{\text{RAPS}}(\mathbf{x}, y) = \mathcal{S}_{\text{APS}}(\mathbf{x}, y) + \lambda \cdot (o(\mathbf{x}, y) - k_{\text{reg}})^+$$

Angelopoulos et al., Uncertainty Sets for Image Classifiers Using Conformal Prediction, ICLR 2021

What about non-conformity scores?

$$\mathcal{S}_{\text{LAC}}(\mathbf{x}, y) = 1 - p_{k=y}$$

Sadinle et al., Least ambiguous set-valued classifiers with bounded error levels, Jour. American Statistical Association 2019

$$\mathcal{S}_{\text{APS}}(\mathbf{x}, y) = \rho_x(y) + p_{k=y} \cdot u$$

Romano et al., Classification with valid and adaptive coverage., NeurIPS 2020

$$\mathcal{S}_{\text{RAPS}}(\mathbf{x}, y) = \mathcal{S}_{\text{APS}}(\mathbf{x}, y) + \lambda \cdot (o(\mathbf{x}, y) - k_{\text{reg}})^+$$

Angelopoulos et al., Uncertainty Sets for Image Classifiers Using Conformal Prediction, ICLR 2021

		$\alpha = 0.10$			
Method		ACA \uparrow	Cov.	Size \downarrow	CCV \downarrow
LAC	SCP	50.2	0.890	3.99	9.96
	Adapt+SCP	67.1$_{+16.9}$	0.842	2.40$_{-1.59}$	11.17$_{+1.21}$
	FCA (<i>Ours</i>)	67.1$_{+16.9}$	0.896	2.91 $_{-1.08}$	8.38$_{-1.58}$
APS	SCP	50.2	0.900	4.05	9.59
	Adapt+SCP	67.1$_{+16.9}$	0.858	2.56$_{-1.49}$	8.57$_{-1.02}$
	FCA (<i>Ours</i>)	67.1$_{+16.9}$	0.898	3.06 $_{-0.99}$	6.12$_{-3.47}$
RAPS	SCP	50.2	0.901	4.16	9.55
	Adapt+SCP	67.1$_{+16.9}$	0.856	2.55$_{-1.61}$	8.64$_{-0.91}$
	FCA (<i>Ours</i>)	67.1$_{+16.9}$	0.898	3.05 $_{-1.11}$	6.21$_{-3.34}$

What about non-conformity scores?

$$\mathcal{S}_{\text{LAC}}(\mathbf{x}, y) = 1 - p_{k=y}$$

Sadinle et al., Least ambiguous set-valued classifiers with bounded error levels, Jour. American Statistical Association 2019

$$\mathcal{S}_{\text{APS}}(\mathbf{x}, y) = \rho_x(y) + p_{k=y} \cdot u$$

Romano et al., Classification with valid and adaptive coverage., NeurIPS 2020

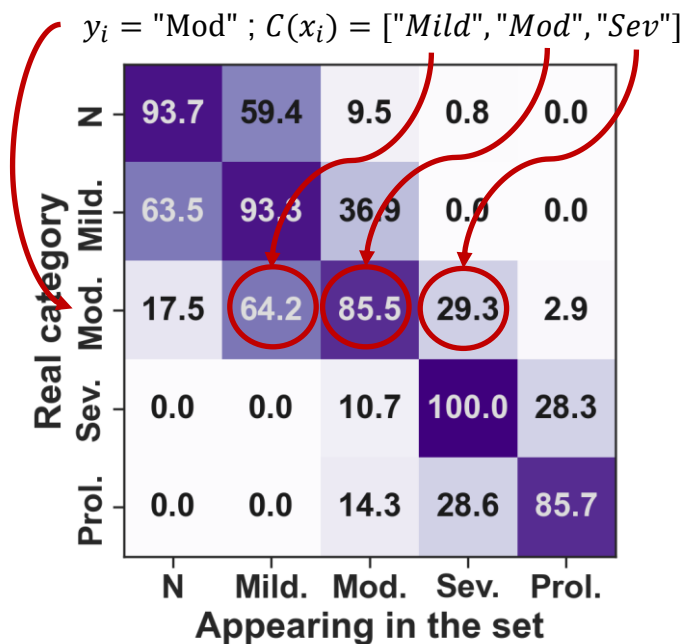
$$\mathcal{S}_{\text{RAPS}}(\mathbf{x}, y) = \mathcal{S}_{\text{APS}}(\mathbf{x}, y) + \lambda \cdot [o(\mathbf{x}, y) - k_{\text{reg}}]^+$$

Angelopoulos et al., Uncertainty Sets for Image Classifiers Using Conformal Prediction, ICLR 2021

Method		$\alpha = 0.10$			
		ACA \uparrow	Cov.	Size \downarrow	CCV \downarrow
LAC	SCP	50.2	0.890	3.99	9.96
	Adapt+SCP	67.1 _{+16.9}	0.842	2.40 _{-1.59}	11.17 _{+1.21}
	FCA (<i>Ours</i>)	67.1 _{+16.9}	0.896	2.91 _{-1.08}	8.38 _{-1.58}
APS	SCP	50.2	0.900	4.05	9.59
	Adapt+SCP	67.1 _{+16.9}	0.858	2.56 _{-1.49}	8.57 _{-1.02}
	FCA (<i>Ours</i>)	67.1 _{+16.9}	0.898	3.06 _{-0.99}	6.12 _{-3.47}
RAPS	SCP	50.2	0.901	4.16	9.55
	Adapt+SCP	67.1 _{+16.9}	0.856	2.55 _{-1.61}	8.64 _{-0.91}
	FCA (<i>Ours</i>)	67.1 _{+16.9}	0.898	3.05 _{-1.11}	6.21 _{-3.34}

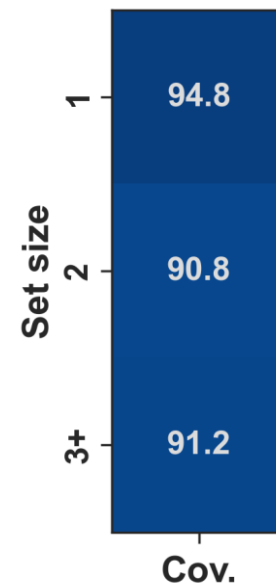
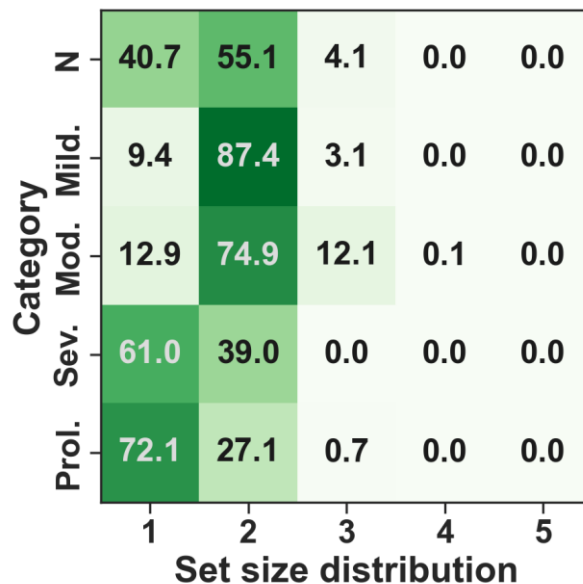
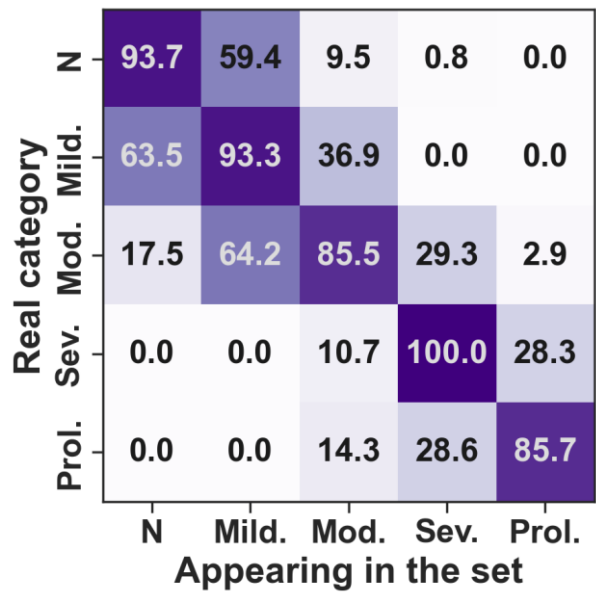
Interpretability of conformal sets

- Use-case: **diabetic retinopathy grading**. Top-1 accuracy: **71%** ; Coverage: **90%**.



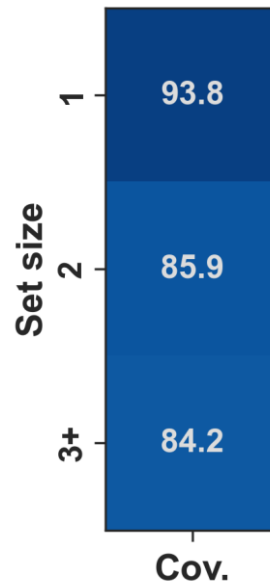
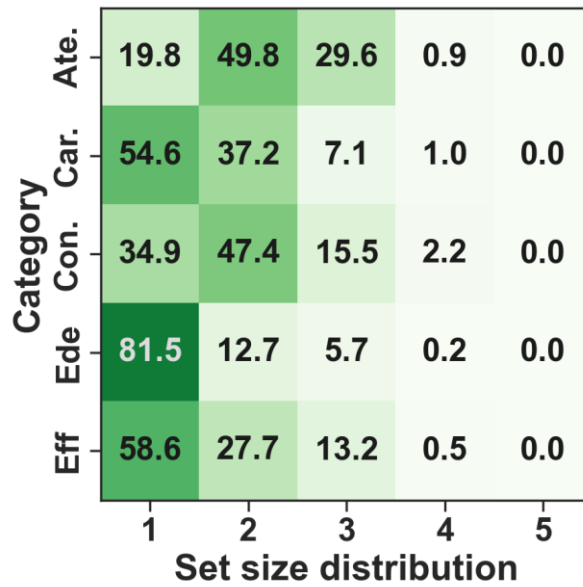
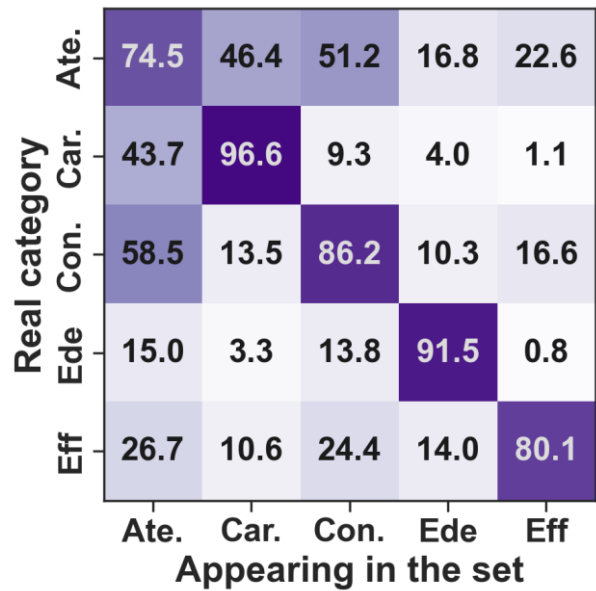
Interpretability of conformal sets

- Use-case: **diabetic retinopathy grading**. Top-1 accuracy: **71%** ; Coverage: **90%**.



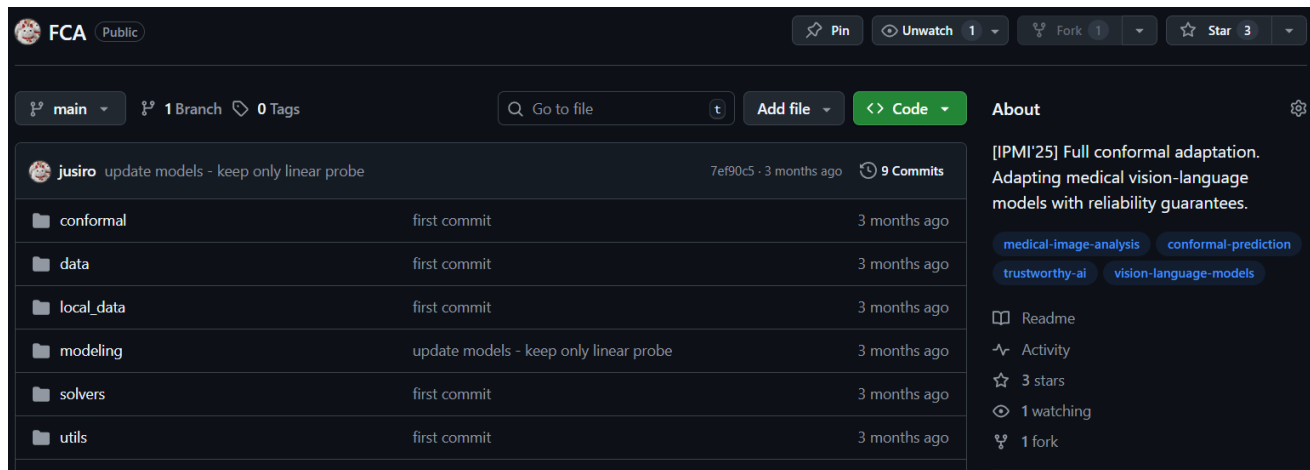
Interpretability of conformal sets

- Use-case: **chest X-ray findings classification**. Top-1 accuracy: **81%** ; Coverage: **90%**.



Implementation & benchmark publicly available

<https://github.com/jusiro/FCA>



The screenshot shows the GitHub repository page for 'FCA' by user 'jusiro'. The repository is public and has 3 stars, 1 fork, and 1 watch. The main branch is 'main' with 1 branch and 0 tags. The repository description is '[IPMI'25] Full conformal adaptation. Adapting medical vision-language models with reliability guarantees.' The repository contains several files and folders: 'conformal', 'data', 'local_data', 'modeling', 'solvers', and 'utils'. The 'modeling' folder is the most recent, updated 3 months ago. The 'About' section on the right provides more details about the project, including tags like 'medical-image-analysis', 'conformal-prediction', 'trustworthy-ai', and 'vision-language-models'. It also lists the repository's README, activity, stars, and forks.



File/Folder	Commit Message	Commit Hash	Time Ago
conformal	first commit	7ef90c5	3 months ago
data	first commit		3 months ago
local_data	first commit		3 months ago
modeling	update models - keep only linear probe		3 months ago
solvers	first commit		3 months ago
utils	first commit		3 months ago



Implementation & benchmark publicly available

<https://github.com/jusiro/FCA/blob/main/docs/awesome-miccai-conformal.md>





 **awesome-miccai-conformal** 


license **MIT** - Under construction

Curated list of awesome papers on conformal prediction for medical image analysis, published in the MICCAI community.
For a more general list of resources in conformal prediction, we highly recommend the following repository:
[valeman/awesome-conformal-prediction](#)

MICCAI 2025

Conformal forecasting for surgical instrument trajectory.
Sara Sangalli, Gary Sarwin, Ertunc Erdil, Alessandro Carretta, Victor Staartjes, Carlo Serra, Ender Konukoglu
[\[Paper\]](#) [\[Code\]](#)  Stars **0**

Trustworthy Few-Shot Transfer of Medical VLMs through Split Conformal Prediction.
Julio Silva-Rodríguez, Ismail Ben Ayed, Jose Dolz
[\[Paper\]](#) [\[Code\]](#)  Stars **2**

Conformal Prediction for Image Segmentation Using Morphological Prediction Sets.
Luca Mossina, Corentin Friedrich
[\[Paper\]](#) [\[Code\]](#)  Stars **5**