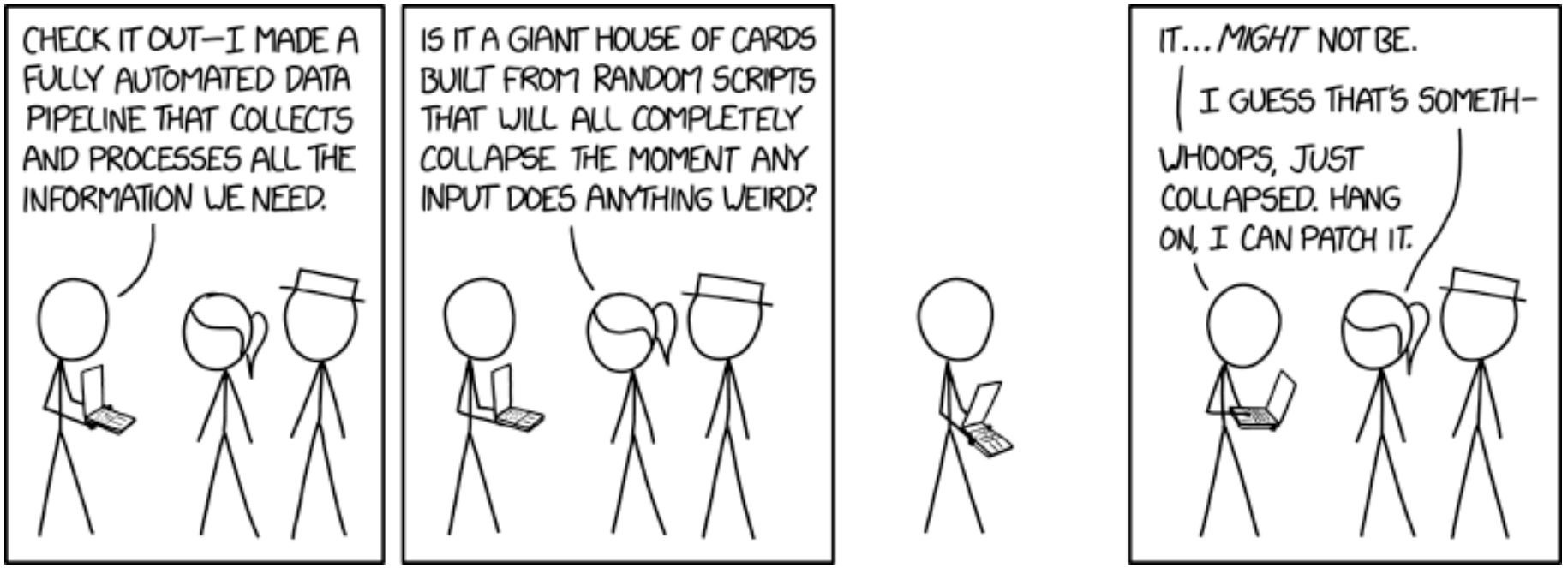


PC: https://xkcd.com/2054

November 8th

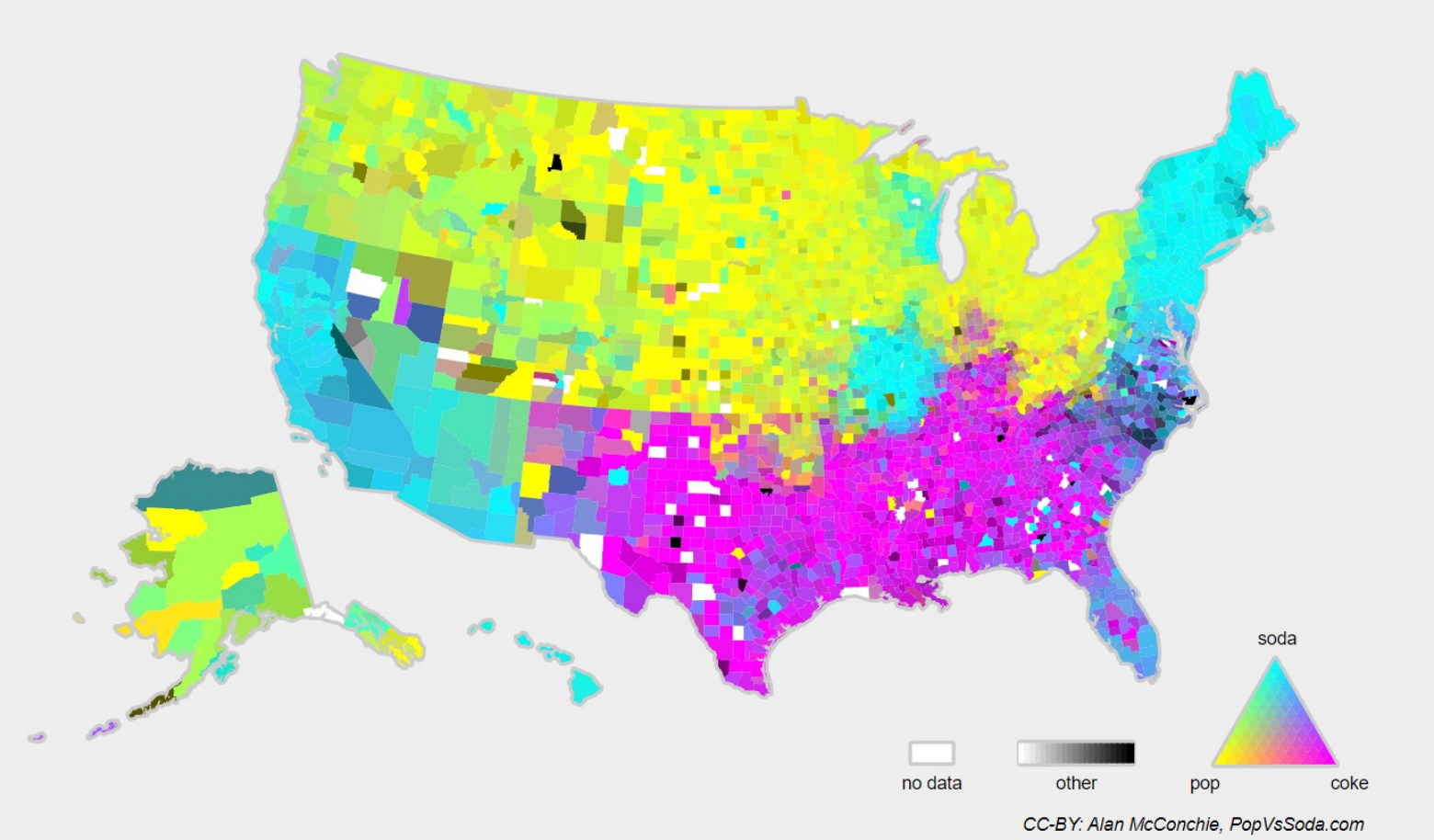
- Election day
- Option A: Cancel class
- Option B: Hold extended reading group for those that can come

Distribution shift

Eric Wong

9/8/2022

Pop vs Soda



PC: https://popvssoda.com/

Content moderation

92.2% average test accuracy, but...

Demographic	Test accuracy on non-toxic comments
Male	87.3 (0.7)
Female	89.0 (0.6)
LGBTQ	74.6 (0.5)
Christian	92.1 (0.2)
Muslim	80.9 (1.0)
Other religions	86.1 (0.1)
Black	**69.2** (1.3)
White	71.2 (1.4)

Borkan et al. 2019 “Nuanced Metrics for Measuring Unintended Bias with Real Data for Text Classification”

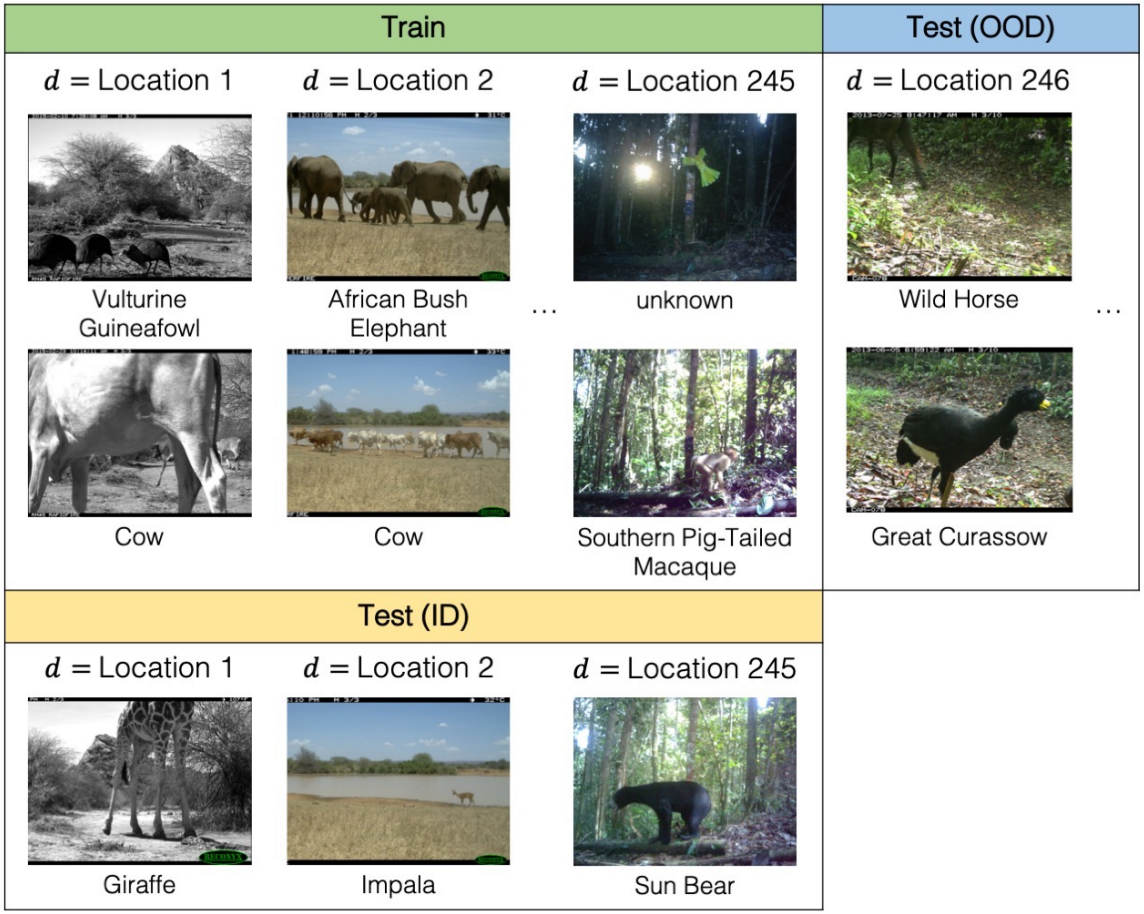
WILDS

- 10 datasets with distribution shift, ranging from ecological conservation to medical imaging.

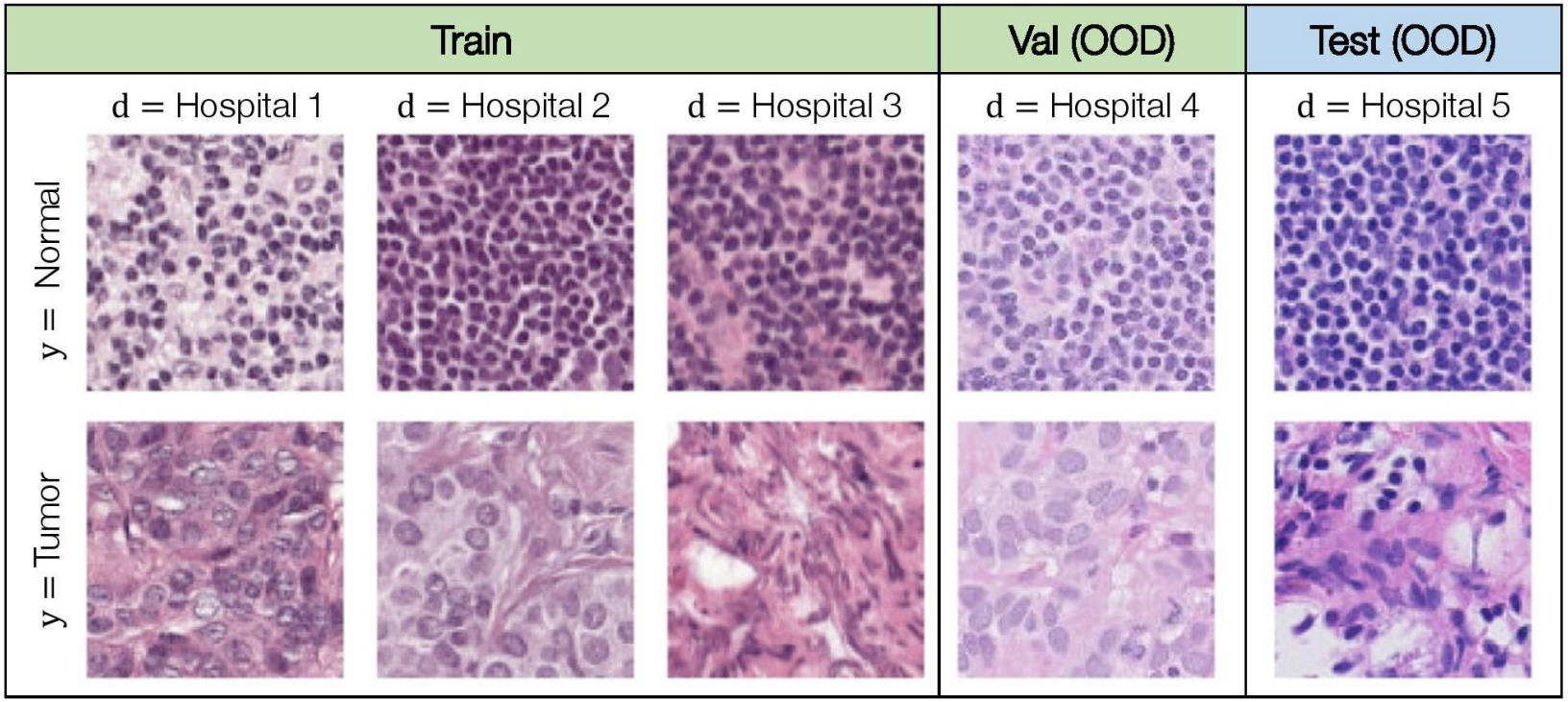
- WILDS 2.0 adds unlabeled data for 8 datasets.

Koh*, Sagawa*, Marklund, Xie, Zhang, Balsubramani, Hu, Yasunaga, Phillips, Gao, Lee, David, Stavness, Guo, Earnshaw, Haque, Beery, Leskovec, Kundaje, Pierson, Levine, Finn, Liang. WILDS: A Benchmark of in-the-Wild Distribution Shifts. ICML 2021.

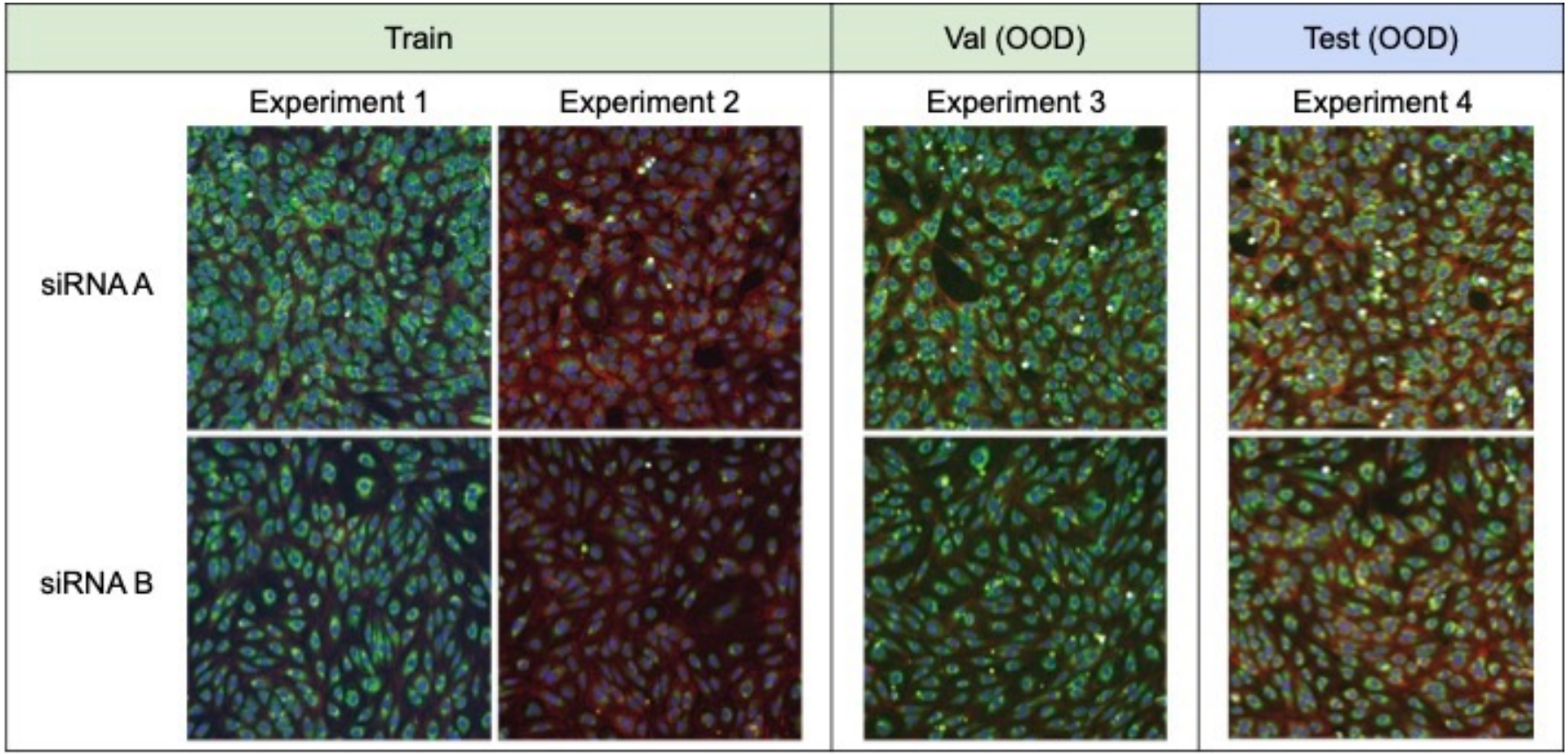
iWildCam


Train
Test (OOD)
d = Location 1
d = Location 2
d = Location 245
d = Location 246
Vulturine Guineafowl
African Bush Elephant
unknown
Wild Horse
Cow
Cow
Southern Pig-Tailed Macaque
Great Curassow
Test (ID)
d = Location 1
d = Location 2
d = Location 245
Giraffe
Impala
Sun Bear

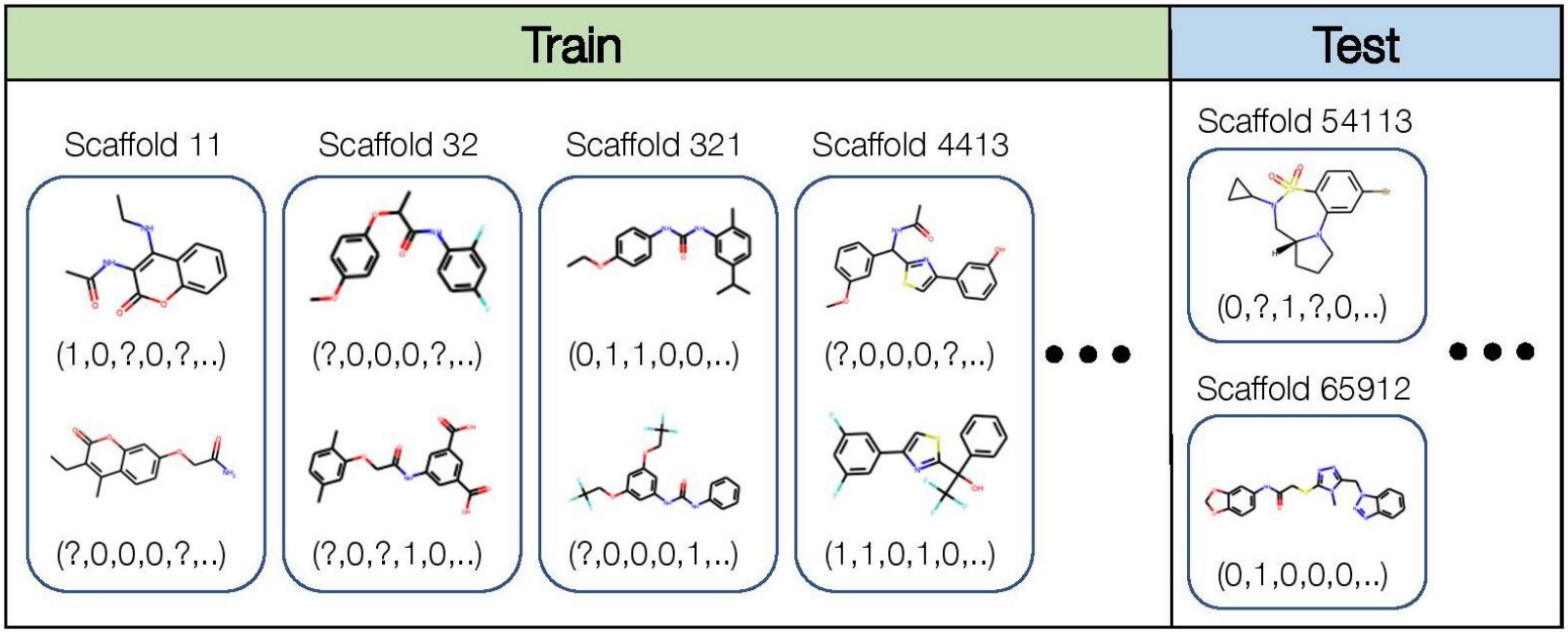
Camelyon17


Train
Val (OOD)
Test (OOD)
d = Hospital 1
d = Hospital 2
d = Hospital 3
d = Hospital 4
d = Hospital 5
y = Normal
y = Tumor

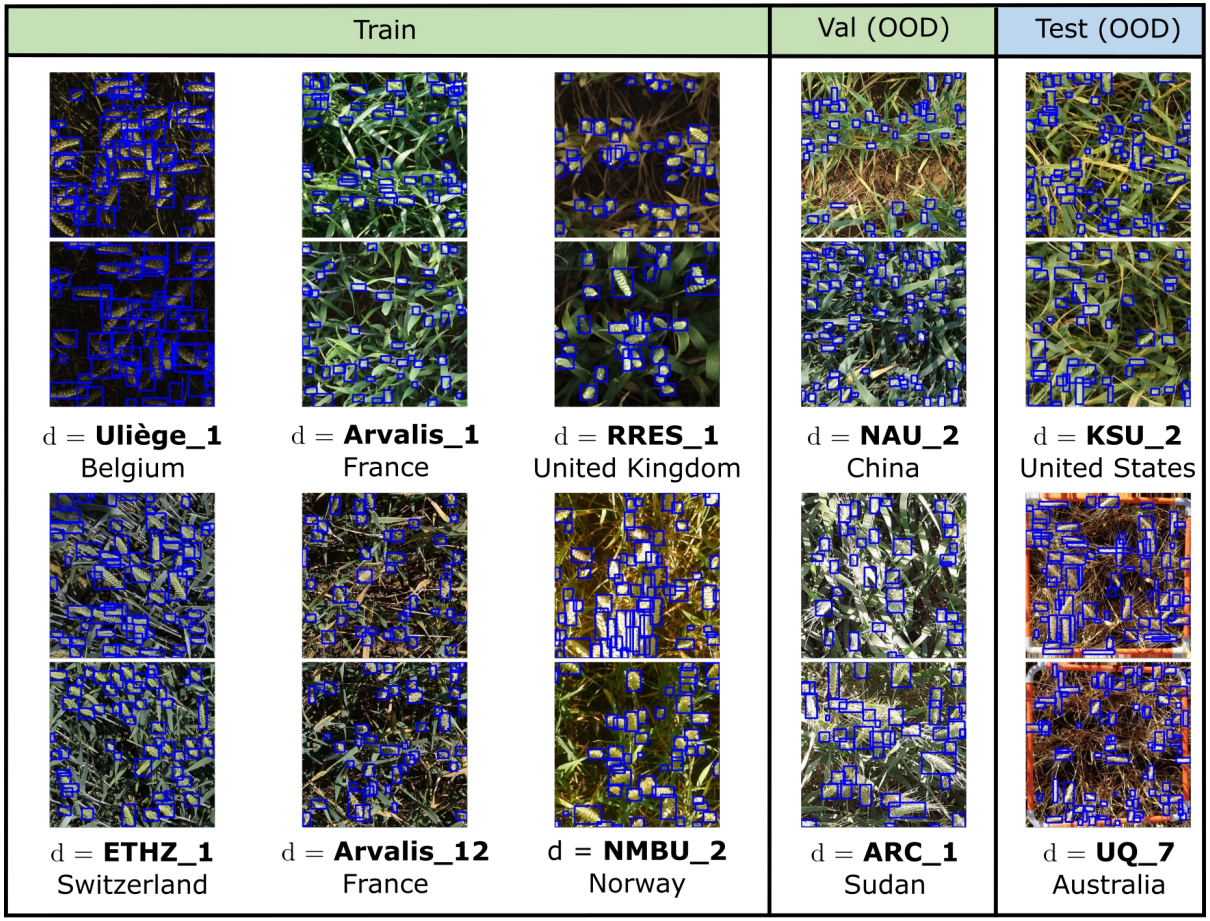
RxRx1


Train
Val (OOD)
Test (OOD)
Experiment 1
Experiment 2
Experiment 3
Experiment 4
siRNA A
siRNA B

OGB-MolPCBA


Train
Test
Scaffold 11
(1,0,?,0,?,..)
(?,0,0,0,?,..)
Scaffold 32
(?,0,0,0,?,..)
(?,0,?,1,0,..)
Scaffold 321
(0,1,1,0,0,..)
(?,0,0,0,1,..)
Scaffold 4413
(?,0,0,0,?,..)
(1,1,0,1,0,..)
Scaffold 54113
(0,?,1,?,0,..)
Scaffold 65912
(0,1,0,0,0,..)

GlobalWheat



CivilComments

Toxic	Comment Text	Male	Female	LGBTQ	White	Black	...	Christian
0	I applaud your father. He was a good man! We need more like him.	1	0	0	0	0	...	0
0	As a Christian, I will not be patronizing any of those businesses.	0	0	0	0	0	...	1
0	What do Black and LGBT people have to do with bicycle licensing?	0	0	1	0	1	...	0
0	Government agencies track down foreign baddies and protect law-abiding white citizens. How many shows does that describe?	0	0	0	1	0	...	0
1	Maybe you should learn to write a coherent sentence so we can understand WTF your point is.	0	0	0	0	0	...	0

FMoW

	Train			Test	
Satellite Image (x)					
Year / Region (d)	2002 / Americas	2009 / Africa	2012 / Europe	2016 / Americas	2017 / Africa
Building / Land Type (y)	shopping mall	multi-unit residential	road bridge	recreational facility	educational institution

PovertyMap

	Train			Test	
Satellite image (x)					
Country / Urban-rural (d)	Angola / urban	Angola / rural	Angola / urban	Kenya / urban	Kenya / rural
Asset index (y)	0.259	-1.106	2.347	0.827	0.130

Amazon

	Reviewer ID (d)	Review Text (x)	Stars (y)
Train	Reviewer 1	They are decent shoes. Material quality is good but the color fades very quickly. Not as black in person as shown.	5
		Super easy to put together. Very well built.	5
	Reviewer 2	This works well and was easy to install. The only thing I don't like is that it tilts forward a little bit and I can't figure out how to stop it.	4
		Perfect for the trail camera	5
		…	
	Reviewer 10,000	I am disappointed in the quality of these. They have significantly deteriorated in just a few uses. I am going to stick with using foil.	1
		Very sturdy especially at this price point. I have a memory foam mattress on it with nothing underneath and the slats perform well.	5
Test	Reviewer 10,001	Solidly built plug in. I have had 4 devices plugged in and all charge just fine.	5
		Works perfectly on the wall to hang our wreath without having to do any permanent damage.	5
		…	

Py150

	Repository ID (d)	Source code context (x)	Next tokens (y)
Train	Repository 1	`... from easyrec.gateway import EasyRec <EOL> gateway = EasyRec('tenant','key') <EOL> item_type = gateway. ____`	`get_item_type`
		`... response = gateway.get_other_users() <EOL> get_params = HTTPretty. ____`	`last_request`
	Repository 2	`import numpy as np ... <EOL> if np.linalg.norm(target - prev_target) > far_threshold: <EOL> norm = np. ____`	`linalg`
		`... new_trans = np.zeros((n_beats + max_beats, n_beats) <EOL> new_trans[:n_beats,:n_beats] = np. ____`	`max`
	⋮		
Test	Repository 6,001	`... if e.errno == errno.ENOENT: <EOL> continue <EOL> p = subprocess.Popen () <EOL> stdout = p. ____`	`communicate`
		`... command = shlex.split(command) <EOL> command = map(str, command) <EOL> env = os. ____`	`environ`
	⋮		