Single-sample versus case-control sampling schemes for Positive Unlabeled data: the story of two scenarios – supplementary material

Jan Mielniczuk [C]
Faculty of Mathematics and Information Science
Warsaw University of Technology
Koszykowa 75, 00-662 Warsaw, Poland
and
Institute of Computer Science
Polish Academy of Sciences
Jana Kazimierza 5, 01-248 Warsaw, Poland
jan.mielniczuk@ipipan.waw.pl

Adam Wawrzeńczyk
Institute of Computer Science
Polish Academy of Sciences
Jana Kazimierza 5, 01-248 Warsaw, Poland
adam.wawrzenczyk@ipipan.waw.pl

1. Extended result tables

In this section we present additional result tables not presented in the paper (F1 score, precision and recall for both single-sample and case-control scenario) in tables 1 through 6.

[C]Corresponding author

Table 1. Test F1 score, single-sample datasets. „Δ" indicates F1 score difference between scenario-appropriate nnPU$_{ss}$ method and ill-specified nnPU$_{cc}$ method.

c	Model	Beans	CIFAR	Chest X-ray	DogFood	EuroSAT	FashionMNIST	MNIST	Oxford Pets	Snacks
0.1	nnPUcc	66.24	94.19	92.30	79.91	81.71	97.14	95.28	74.36	68.70
	nnPUss	61.48	93.20	92.80	76.76	76.04	94.99	93.71	65.67	67.93
	Δ	-4.76	-0.99	0.50	-3.15	-5.67	-2.15	-1.56	-8.70	-0.77
0.3	nnPUcc	84.22	93.75	94.13	96.37	90.34	96.79	95.51	93.21	77.40
	nnPUss	82.71	94.97	94.95	92.60	85.15	96.96	96.80	81.24	77.20
	Δ	-1.51	1.22	0.82	-3.77	-5.19	0.17	1.29	-11.97	-0.20
0.5	nnPUcc	87.24	90.58	93.96	98.21	90.00	95.12	93.05	97.09	79.54
	nnPUss	87.65	95.96	95.41	97.18	90.01	97.98	98.26	88.85	81.74
	Δ	0.41	5.38	1.44	-1.02	0.01	2.86	5.21	-8.24	2.20
0.7	nnPUcc	88.47	85.83	92.17	98.32	86.25	85.51	87.48	96.25	79.84
	nnPUss	91.31	97.10	95.84	99.01	92.65	99.11	99.00	93.77	84.47
	Δ	2.84	11.27	3.67	0.69	6.39	13.60	11.52	-2.48	4.63
0.9	nnPUcc	88.01	82.93	89.24	98.00	82.69	76.08	75.94	94.73	79.92
	nnPUss	93.49	97.99	96.57	99.71	94.83	99.44	99.21	98.35	86.89
	Δ	5.48	15.06	7.33	1.70	12.15	23.36	23.27	3.62	6.96

c	Model	California	Credit	Electricity	Wine	20News	HateSpeech	IMDB	PoemSentiment	SMSSpam
0.1	nnPUcc	82.43	67.12	74.80	69.65	82.96	0.00	73.01	1.18	20.45
	nnPUss	82.07	66.57	73.97	68.17	81.44	0.00	71.70	2.35	28.29
	Δ	-0.36	-0.54	-0.83	-1.48	-1.52	0.00	-1.31	1.18	7.84
0.3	nnPUcc	84.39	70.78	78.62	75.52	85.70	29.84	77.50	11.54	85.11
	nnPUss	84.08	70.08	76.88	73.94	83.88	30.87	74.22	27.46	80.71
	Δ	-0.31	-0.70	-1.74	-1.58	-1.82	1.03	-3.28	15.92	-4.40
0.5	nnPUcc	84.37	70.99	80.48	78.44	85.64	42.47	79.36	29.54	91.42
	nnPUss	85.33	70.94	79.15	76.08	85.48	41.00	75.18	48.05	88.58
	Δ	0.95	-0.05	-1.34	-2.36	-0.16	-1.47	-4.18	18.51	-2.84
0.7	nnPUcc	82.93	71.02	80.58	79.47	83.94	47.28	79.57	52.70	93.89
	nnPUss	86.30	70.64	80.33	78.65	86.46	45.30	77.07	60.37	92.74
	Δ	3.38	-0.38	-0.25	-0.81	2.51	-1.98	-2.50	7.67	-1.15
0.9	nnPUcc	80.82	70.81	79.23	79.11	81.19	49.51	78.55	59.71	94.85
	nnPUss	86.37	70.05	81.14	80.29	87.62	48.39	78.77	65.19	93.67
	Δ	5.55	-0.76	1.91	1.18	6.43	-1.11	0.22	5.48	-1.18

Table 2. Test precision, single-sample datasets. „Δ" indicates precision difference between scenario-appropriate nnPU$_{ss}$ method and ill-specified nnPU$_{cc}$ method.

c	Model	Beans	CIFAR	Chest X-ray	DogFood	EuroSAT	FashionMNIST	MNIST	Oxford Pets	Snacks
0.1	nnPUcc	82.98	90.11	90.88	89.28	92.74	97.62	94.47	99.59	70.35
	nnPUss	77.50	89.90	91.21	88.30	91.31	97.83	95.15	99.72	71.62
	Δ	-5.48	-0.20	0.33	-0.98	-1.44	0.20	0.68	0.13	1.27
0.3	nnPUcc	86.94	88.41	91.84	97.43	89.56	93.79	91.53	99.53	74.45
	nnPUss	86.16	92.51	93.00	97.47	91.33	97.24	97.27	99.78	77.30
	Δ	-0.77	4.09	1.17	0.04	1.78	3.45	5.74	0.26	2.85
0.5	nnPUcc	83.14	82.81	91.72	97.77	83.92	90.71	87.04	97.68	73.93
	nnPUss	87.08	93.76	93.86	99.36	92.65	98.02	98.26	99.98	81.20
	Δ	3.93	10.95	2.14	1.59	8.73	7.31	11.22	2.30	7.27
0.7	nnPUcc	81.97	75.18	91.03	97.42	76.65	74.76	77.80	93.87	72.64
	nnPUss	90.26	95.41	94.80	99.80	93.90	98.91	98.79	99.99	84.15
	Δ	8.29	20.23	3.76	2.38	17.25	24.14	20.99	6.13	11.51
0.9	nnPUcc	81.71	70.85	90.14	96.64	70.86	61.39	61.23	90.34	71.83
	nnPUss	93.23	97.00	96.41	99.96	95.34	99.07	98.84	100.00	87.84
	Δ	11.52	26.14	6.27	3.32	24.48	37.68	37.62	9.66	16.01

c	Model	California	Credit	Electricity	Wine	20News	HateSpeech	IMDB	PoemSentiment	SMSSpam
0.1	nnPUcc	79.48	62.82	75.93	69.46	79.78	0.00	73.84	10.00	80.00
	nnPUss	80.85	63.22	76.93	69.72	79.86	0.00	74.69	20.00	99.77
	Δ	1.37	0.41	1.01	0.26	0.08	0.00	0.86	10.00	19.77
0.3	nnPUcc	78.83	63.23	77.92	71.86	79.58	57.14	75.85	48.00	95.50
	nnPUss	84.06	65.97	81.01	75.10	82.08	55.47	78.96	55.04	92.94
	Δ	5.22	2.74	3.09	3.24	2.49	-1.67	3.12	7.04	-2.56
0.5	nnPUcc	76.24	59.74	76.95	71.77	77.53	52.24	74.72	67.50	94.70
	nnPUss	85.29	62.80	81.15	76.70	83.88	52.74	80.96	70.25	92.48
	Δ	9.05	3.06	4.20	4.93	6.35	0.50	6.25	2.75	-2.22
0.7	nnPUcc	72.68	57.87	73.63	70.48	73.84	50.32	71.89	80.16	96.70
	nnPUss	86.10	59.65	81.76	77.79	84.96	53.35	81.58	76.51	93.85
	Δ	13.42	1.78	8.13	7.31	11.12	3.03	9.69	-3.66	-2.85
0.9	nnPUcc	68.84	57.12	69.57	67.99	69.14	49.43	68.30	81.64	98.23
	nnPUss	86.20	56.81	81.42	78.09	86.02	53.92	81.01	80.15	92.39
	Δ	17.36	-0.31	11.84	10.10	16.88	4.49	12.71	-1.49	-5.84

Table 3. Test recall, single-sample datasets. „Δ" indicates precision recall between scenario-appropriate nnPU$_{ss}$ method and ill-specified nnPU$_{cc}$ method.

c	Model	Beans	CIFAR	Chest X-ray	DogFood	EuroSAT	FashionMNIST	MNIST	Oxford Pets	Snacks
0.1	nnPUcc	55.95	98.66	94.04	72.75	73.05	96.68	96.10	59.53	67.40
	nnPUss	51.67	96.76	94.62	68.51	65.20	92.33	92.33	49.28	64.81
	Δ	-4.29	-1.90	0.58	-4.24	-7.85	-4.34	-3.78	-10.25	-2.60
0.3	nnPUcc	81.90	99.76	96.59	95.37	91.15	99.98	99.85	87.65	80.95
	nnPUss	79.76	97.57	97.01	88.24	79.77	96.68	96.33	68.59	77.22
	Δ	-2.14	-2.20	0.41	-7.14	-11.38	-3.30	-3.52	-19.06	-3.73
0.5	nnPUcc	92.14	99.98	96.37	98.67	97.05	100.00	99.96	96.50	86.32
	nnPUss	88.57	98.28	97.01	95.14	87.52	97.95	98.26	79.97	82.39
	Δ	-3.57	-1.70	0.63	-3.53	-9.53	-2.05	-1.70	-16.53	-3.93
0.7	nnPUcc	96.43	99.99	93.55	99.25	98.64	100.00	99.99	98.77	88.87
	nnPUss	92.62	98.85	96.91	98.24	91.44	99.31	99.21	88.28	84.83
	Δ	-3.81	-1.15	3.36	-1.02	-7.20	-0.69	-0.79	-10.49	-4.04
0.9	nnPUcc	95.71	100.00	89.12	99.45	99.27	100.00	99.99	99.59	90.39
	nnPUss	93.81	99.01	96.74	99.45	94.33	99.81	99.57	96.76	85.98
	Δ	-1.90	-0.99	7.62	0.00	-4.94	-0.19	-0.41	-2.83	-4.42

c	Model	California	Credit	Electricity	Wine	20News	HateSpeech	IMDB	PoemSentiment	SMSSpam
0.1	nnPUcc	85.77	72.17	73.73	69.90	86.41	0.00	72.20	0.62	12.46
	nnPUss	83.49	70.44	71.25	66.76	83.09	0.00	68.94	1.25	17.85
	Δ	-2.29	-1.74	-2.48	-3.14	-3.32	0.00	-3.26	0.62	5.40
0.3	nnPUcc	90.86	80.52	79.37	79.66	92.84	20.29	79.23	6.88	76.89
	nnPUss	84.18	74.99	73.17	72.94	85.77	21.43	70.02	18.75	71.38
	Δ	-6.68	-5.54	-6.20	-6.72	-7.08	1.15	-9.22	11.88	-5.51
0.5	nnPUcc	94.57	87.65	84.36	86.53	95.64	35.82	84.63	19.38	88.45
	nnPUss	85.47	81.93	77.25	75.67	87.15	33.66	70.18	37.50	85.20
	Δ	-9.10	-5.72	-7.11	-10.86	-8.50	-2.16	-14.46	18.13	-3.25
0.7	nnPUcc	96.62	92.10	89.00	91.11	97.25	44.67	89.08	40.00	91.31
	nnPUss	86.62	87.02	78.98	79.65	88.01	39.50	73.03	50.62	91.72
	Δ	-10.00	-5.08	-10.03	-11.46	-9.24	-5.17	-16.05	10.62	0.41
0.9	nnPUcc	97.89	93.27	92.00	94.60	98.34	49.67	92.42	48.12	91.71
	nnPUss	86.60	91.64	80.92	82.70	89.30	44.06	76.66	55.62	95.07
	Δ	-11.29	-1.62	-11.07	-11.90	-9.04	-5.61	-15.76	7.50	3.37

Table 4. Test F1 score, case-control datasets. „Δ" indicates F1 score difference between scenario-appropriate nnPU_{cc} method and ill-specified nnPU_{ss} method.

c	Model	Beans	CIFAR	Chest X-ray	DogFood	EuroSAT	Fashion MNIST	MNIST	Oxford Pets	Snacks
0.1	nnPUss	71.14	94.59	92.89	79.05	75.37	94.77	93.64	68.69	69.33
	nnPUcc	71.76	94.99	92.33	81.47	81.77	97.06	95.76	76.14	69.63
	Δ	0.62	0.41	-0.57	2.41	6.40	2.29	2.12	7.45	0.30
0.3	nnPUss	80.53	95.86	95.44	90.48	82.73	94.67	94.25	76.35	78.76
	nnPUcc	80.47	96.76	94.89	95.79	90.94	98.71	97.95	90.09	80.64
	Δ	-0.06	0.90	-0.55	5.31	8.21	4.03	3.70	13.73	1.87
0.5	nnPUss	84.72	92.34	95.35	93.16	79.46	88.21	86.20	82.06	73.04
	nnPUcc	84.72	97.87	95.81	97.51	94.07	99.34	98.95	94.87	84.74
	Δ	0.00	5.53	0.46	4.35	14.61	11.13	12.75	12.81	11.70
0.7	nnPUss	82.66	87.82	88.92	89.77	75.12	87.76	89.83	80.44	72.22
	nnPUcc	86.09	98.66	94.28	96.85	96.01	99.59	99.38	96.99	84.29
	Δ	3.42	10.84	5.36	7.08	20.89	11.83	9.55	16.56	12.07
0.9	nnPUss	47.27	94.06	91.77	27.68	12.50	92.43	92.08	18.05	55.93
	nnPUcc	69.56	99.32	87.59	81.60	97.23	99.77	99.59	94.50	77.50
	Δ	22.29	5.26	-4.18	53.91	84.74	7.33	7.51	76.45	21.57

c	Model	California	Credit	Electricity	Wine	20News	HateSpeech	IMDB	PoemSentiment	SMSSpam
0.1	nnPUss	82.45	66.44	74.32	69.11	82.00	0.00	72.41	7.38	36.99
	nnPUcc	82.81	67.37	75.04	71.25	83.44	0.00	73.75	5.31	30.33
	Δ	0.36	0.93	0.73	2.14	1.44	0.00	1.34	-2.07	-6.66
0.3	nnPUss	85.54	71.56	78.13	71.83	83.71	34.42	71.84	44.10	80.68
	nnPUcc	87.03	73.10	78.95	77.06	87.64	38.94	78.21	41.61	86.37
	Δ	1.48	1.54	0.82	5.23	3.93	4.52	6.38	-2.49	5.69
0.5	nnPUss	87.39	76.16	81.21	76.63	78.87	32.08	74.80	55.17	79.65
	nnPUcc	89.26	77.83	81.75	81.65	90.34	49.95	80.95	59.11	92.79
	Δ	1.87	1.67	0.55	5.02	11.46	17.87	6.15	3.94	13.14
0.7	nnPUss	90.07	80.34	83.89	79.64	80.82	6.47	78.05	45.80	47.23
	nnPUcc	91.68	82.49	84.54	85.44	93.01	56.17	84.07	66.88	94.47
	Δ	1.62	2.16	0.65	5.80	12.19	49.70	6.02	21.08	47.24
0.9	nnPUss	89.91	82.10	85.19	81.38	89.11	0.00	81.48	0.00	0.78
	nnPUcc	94.24	85.59	87.69	90.05	96.11	53.40	87.81	54.04	90.68
	Δ	4.32	3.50	2.50	8.67	6.99	53.40	6.33	54.04	89.90

Table 5. Test precision, case-control datasets. „Δ" indicates precision difference between scenario-appropriate nnPU$_{cc}$ method and ill-specified nnPU$_{ss}$ method.

c	Model	Beans	CIFAR	Chest X-ray	DogFood	EuroSAT	Fashion MNIST	MNIST	Oxford Pets	Snacks
0.1	nnPUss	69.88	92.13	92.24	86.86	89.84	98.17	95.55	96.53	69.59
	nnPUcc	70.66	91.73	91.88	85.98	91.76	98.20	95.88	97.48	68.11
	Δ	0.78	-0.40	-0.36	-0.88	1.92	0.03	0.33	0.95	-1.49
0.3	nnPUss	82.50	96.53	95.64	98.36	93.63	99.48	99.33	99.20	86.25
	nnPUcc	81.66	94.40	95.07	97.67	92.93	98.68	97.72	99.16	81.52
	Δ	-0.85	-2.13	-0.57	-0.69	-0.70	-0.80	-1.61	-0.04	-4.73
0.5	nnPUss	90.88	98.83	98.05	99.75	98.06	99.67	99.76	99.80	94.17
	nnPUcc	90.08	96.29	96.78	99.66	94.44	99.09	98.75	99.37	89.23
	Δ	-0.80	-2.55	-1.27	-0.09	-3.62	-0.58	-1.01	-0.43	-4.95
0.7	nnPUss	97.49	99.49	99.32	99.88	98.77	99.83	99.79	99.88	96.74
	nnPUcc	97.43	97.72	98.24	99.90	96.37	99.35	99.28	99.57	94.52
	Δ	-0.07	-1.77	-1.08	0.01	-2.40	-0.48	-0.51	-0.31	-2.22
0.9	nnPUss	99.32	99.62	99.73	100.00	99.75	99.90	99.78	99.90	99.27
	nnPUcc	99.53	98.89	98.73	99.94	98.11	99.64	99.58	99.80	96.76
	Δ	0.21	-0.73	-1.00	-0.06	-1.64	-0.26	-0.20	-0.10	-2.51

c	Model	California	Credit	Electricity	Wine	20News	HateSpeech	IMDB	PoemSentiment	SMSSpam
0.1	nnPUss	81.21	65.61	79.28	71.19	81.08	0.00	76.76	23.75	99.00
	nnPUcc	80.02	65.02	78.31	70.97	81.05	0.00	75.83	20.00	89.57
	Δ	-1.19	-0.59	-0.97	-0.22	-0.03	0.00	-0.94	-3.75	-9.44
0.3	nnPUss	89.47	74.41	85.39	80.97	87.17	61.88	85.07	79.24	96.27
	nnPUcc	86.19	72.45	85.09	79.43	85.56	59.57	82.71	81.66	94.78
	Δ	-3.28	-1.97	-0.30	-1.54	-1.61	-2.31	-2.36	2.42	-1.49
0.5	nnPUss	92.21	80.09	89.20	87.65	91.12	77.01	88.95	85.26	99.33
	nnPUcc	89.87	78.57	88.69	85.61	89.27	67.95	87.32	79.83	96.68
	Δ	-2.34	-1.52	-0.51	-2.05	-1.85	-9.06	-1.63	-5.43	-2.65
0.7	nnPUss	94.79	86.07	93.06	91.36	94.63	92.39	92.60	95.10	99.92
	nnPUcc	93.35	85.22	92.84	90.30	93.06	77.90	91.67	90.85	98.21
	Δ	-1.44	-0.85	-0.23	-1.06	-1.57	-14.49	-0.93	-4.25	-1.71
0.9	nnPUss	98.03	94.78	97.28	95.87	97.32	0.00	96.92	0.00	20.00
	nnPUcc	97.14	94.65	97.05	95.61	96.68	91.34	96.68	96.93	99.30
	Δ	-0.89	-0.14	-0.23	-0.27	-0.65	91.34	-0.25	96.93	79.30

Table 6. Test recall, case-control datasets. „Δ" indicates recall difference between scenario-appropriate nnPU$_{cc}$ method and ill-specified nnPU$_{ss}$ method.

c	Model	Beans	CIFAR	Chest X-ray	DogFood	EuroSAT	Fashion MNIST	MNIST	Oxford Pets	Snacks
0.1	nnPUss	76.01	97.19	93.72	72.93	64.97	91.62	91.82	53.38	69.37
	nnPUcc	76.52	98.50	93.07	77.89	73.79	95.95	95.66	62.52	71.49
	Δ	0.51	1.31	-0.66	4.96	8.82	4.34	3.84	9.14	2.12
0.3	nnPUss	80.79	95.20	95.26	83.90	74.14	90.31	89.67	62.08	72.64
	nnPUcc	81.19	99.25	94.77	94.03	89.05	98.73	98.19	82.55	79.96
	Δ	0.39	4.05	-0.49	10.13	14.91	8.42	8.52	20.47	7.32
0.5	nnPUss	80.03	86.67	92.80	87.47	66.87	79.15	75.93	69.69	59.79
	nnPUcc	80.64	99.50	94.89	95.46	93.71	99.59	99.14	90.77	80.89
	Δ	0.61	12.83	2.09	7.99	26.84	20.44	23.21	21.07	21.09
0.7	nnPUss	72.37	78.63	80.53	81.65	60.68	78.37	81.69	67.35	57.79
	nnPUcc	77.64	99.62	90.75	94.04	95.66	99.83	99.49	94.55	76.29
	Δ	5.27	20.99	10.22	12.39	34.98	21.47	17.79	27.20	18.50
0.9	nnPUss	32.86	89.08	85.04	16.67	6.82	86.01	85.50	10.13	39.37
	nnPUcc	54.70	99.74	80.01	69.99	96.38	99.89	99.60	89.75	65.74
	Δ	21.84	10.66	-5.03	53.32	89.56	13.88	14.10	79.62	26.36

c	Model	California	Credit	Electricity	Wine	20News	HateSpeech	IMDB	PoemSentiment	SMSSpam
0.1	nnPUss	83.83	67.44	69.97	67.29	82.96	0.00	68.53	4.50	23.99
	nnPUcc	85.88	70.00	72.06	71.65	86.00	0.00	71.79	3.07	19.58
	Δ	2.04	2.56	2.09	4.36	3.04	0.00	3.26	-1.43	-4.41
0.3	nnPUss	82.02	69.58	72.06	64.66	80.53	23.98	62.18	31.69	69.52
	nnPUcc	87.90	74.33	73.68	74.93	89.84	28.99	74.19	28.83	79.56
	Δ	5.88	4.75	1.63	10.27	9.31	5.01	12.01	-2.86	10.04
0.5	nnPUss	83.11	73.08	74.57	68.23	69.55	20.36	64.54	41.72	66.54
	nnPUcc	88.74	77.74	75.86	78.10	91.44	39.72	75.45	48.24	89.34
	Δ	5.62	4.66	1.29	9.88	21.90	19.36	10.90	6.52	22.80
0.7	nnPUss	85.81	75.85	76.38	70.70	70.55	3.37	67.46	30.95	31.44
	nnPUcc	90.09	80.35	77.62	81.14	92.97	44.07	77.64	53.67	91.07
	Δ	4.28	4.49	1.24	10.44	22.42	40.70	10.18	22.72	59.64
0.9	nnPUss	83.08	72.50	75.81	70.92	82.19	0.00	70.29	0.00	0.40
	nnPUcc	91.51	78.18	79.99	85.18	95.54	37.88	80.44	38.80	83.50
	Δ	8.43	5.68	4.18	14.26	13.35	37.88	10.15	38.80	83.09