

**Supplementary Table S1.** Descriptive statistics (mean  $\pm$  sd) of the models trained using different strengths and types of biases. Bias\_fn: false negative bias introduced to the training set as described in Table 1. Bias\_fp: false positive bias introduced to the training set as described in Table 1. Random: random error introduced to the training set as described in Table 1. Baseline: baseline model. Regularized: improved model with regularization applied while training.

% Level of Introduced Error	Error type	Model	Accuracy	AUC	Number of true positives	Number of true negatives	Number of false negatives	Number of false positives
0	/	Baseline	0.8376 $\pm$ 0.0142	0.794 $\pm$ 0.0199	377.6667 $\pm$ 1.8559	145 $\pm$ 10.0167	12.3333 $\pm$ 1.8559	89 $\pm$ 10.0167
		Regularized	0.8584 $\pm$ 0.0152	0.8295 $\pm$ 0.0235	368.6667 $\pm$ 5.3333	167 $\pm$ 13.5769	21.3333 $\pm$ 5.3333	67 $\pm$ 13.5769
5	Bias_fn	Baseline	0.8128 $\pm$ 0.0126	0.7974 $\pm$ 0.0191	335 $\pm$ 6.6408	172.2 $\pm$ 11.5213	55 $\pm$ 6.6408	61.8 $\pm$ 11.5213
		Regularized	0.8446 $\pm$ 0.0185	0.8138 $\pm$ 0.0287	365.4 $\pm$ 5.938	161.6 $\pm$ 16.5487	24.6 $\pm$ 5.938	72.4 $\pm$ 16.5487
	Bias_fp	Baseline	0.8109 $\pm$ 0.0109	0.7605 $\pm$ 0.0161	375.2 $\pm$ 5.1711	130.8 $\pm$ 9.4414	14.8 $\pm$ 5.1711	103.2 $\pm$ 9.4414
		Regularized	0.8625 $\pm$ 0.007	0.8464 $\pm$ 0.0022	355.2 $\pm$ 11.8042	183 $\pm$ 7.5697	34.8 $\pm$ 11.8042	51 $\pm$ 7.5697
	Random	Baseline	0.8128 $\pm$ 0.0068	0.7756 $\pm$ 0.0082	360.6 $\pm$ 4.9051	146.6 $\pm$ 4.9457	29.4 $\pm$ 4.9051	87.4 $\pm$ 4.9457
		Regularized	0.8683 $\pm$ 0.0059	0.8447 $\pm$ 0.0084	366.2 $\pm$ 4.4542	175.6 $\pm$ 5.5191	23.8 $\pm$ 4.4542	58.4 $\pm$ 5.5191
10	Bias_fn	Baseline	0.7721 $\pm$ 0.016	0.7739 $\pm$ 0.0165	299 $\pm$ 6.7231	182.8 $\pm$ 5.1711	91 $\pm$ 6.7231	51.2 $\pm$ 5.1711
		Regularized	0.8423 $\pm$ 0.007	0.819 $\pm$ 0.0171	355.8 $\pm$ 10.5612	169.8 $\pm$ 14.0656	34.2 $\pm$ 10.5612	64.2 $\pm$ 14.0656
	Bias_fp	Baseline	0.8138 $\pm$ 0.0161	0.7722 $\pm$ 0.0221	366 $\pm$ 2.5495	141.8 $\pm$ 11.0336	24 $\pm$ 2.5495	92.2 $\pm$ 11.0336
		Regularized	0.8288 $\pm$ 0.0171	0.7874 $\pm$ 0.0272	371.8 $\pm$ 5.8344	145.4 $\pm$ 16.0019	18.2 $\pm$ 5.8344	88.6 $\pm$ 16.0019
	Random	Baseline	0.7837 $\pm$ 0.0087	0.7534 $\pm$ 0.0096	341 $\pm$ 5.099	148 $\pm$ 4.7223	49 $\pm$ 5.099	86 $\pm$ 4.7223
		Regularized	0.8135 $\pm$ 0.0247	0.768 $\pm$ 0.0378	370.4 $\pm$ 7.3389	137.2 $\pm$ 21.3832	19.6 $\pm$ 7.3389	96.8 $\pm$ 21.3832

15	Bias_fn	Baseline	$0.7282 \pm 0.0184$	$0.7467 \pm 0.0168$	$262.4 \pm 11.0481$	$192 \pm 5.3759$	$127.6 \pm 11.0481$	$42 \pm 5.3759$
		Regularized	$0.8494 \pm 0.011$	$0.8397 \pm 0.0206$	$342.6 \pm 7.1666$	$187.4 \pm 13.8838$	$47.4 \pm 7.1666$	$46.6 \pm 13.8838$
	Bias_fp	Baseline	$0.7599 \pm 0.0069$	$0.6946 \pm 0.0095$	$372.8 \pm 2.1541$	$101.4 \pm 4.8642$	$17.2 \pm 2.1541$	$132.6 \pm 4.8642$
		Regularized	$0.8417 \pm 0.0077$	$0.805 \pm 0.0152$	$371.2 \pm 6.8659$	$154 \pm 10.9453$	$18.8 \pm 6.8659$	$80 \pm 10.9453$
	Random	Baseline	$0.7529 \pm 0.0056$	$0.7271 \pm 0.0028$	$323.8 \pm 8.0833$	$146 \pm 5$	$66.2 \pm 8.0833$	$88 \pm 5$
		Regularized	$0.8484 \pm 0.0108$	$0.8235 \pm 0.0167$	$360 \pm 6.9065$	$169.4 \pm 10.6986$	$30 \pm 6.9065$	$64.6 \pm 10.6986$
20	Bias_fn	Baseline	$0.6548 \pm 0.0117$	$0.6934 \pm 0.0079$	$210.2 \pm 10.9973$	$198.4 \pm 5.1827$	$179.8 \pm 10.9973$	$35.6 \pm 5.1827$
		Regularized	$0.8474 \pm 0.0039$	$0.8333 \pm 0.01$	$347 \pm 6.5955$	$181.8 \pm 8.4817$	$43 \pm 6.5955$	$52.2 \pm 8.4817$
	Bias_fp	Baseline	$0.7522 \pm 0.012$	$0.6859 \pm 0.0158$	$371 \pm 3.1623$	$98.4 \pm 7.6328$	$19 \pm 3.1623$	$135.6 \pm 7.6328$
		Regularized	$0.8054 \pm 0.0269$	$0.756 \pm 0.041$	$372 \pm 7.1554$	$130.6 \pm 22.936$	$18 \pm 7.1554$	$103.4 \pm 22.936$
	Random Random	Baseline	$0.7032 \pm 0.0051$	$0.6771 \pm 0.0081$	$304.8 \pm 7.9524$	$134 \pm 7.6616$	$85.2 \pm 7.9524$	$100 \pm 7.6616$
		Regularized	$0.8381 \pm 0.0201$	$0.8122 \pm 0.0325$	$357.2 \pm 8.2668$	$165.8 \pm 19.5765$	$32.8 \pm 8.2668$	$68.2 \pm 19.5765$
25	Bias_fn	Baseline	$0.6144 \pm 0.0158$	$0.6599 \pm 0.0109$	$186.4 \pm 13.178$	$197 \pm 4.7749$	$203.6 \pm 13.178$	$37 \pm 4.7749$
		Regularized	$0.7333 \pm 0.0376$	$0.7725 \pm 0.0255$	$240.2 \pm 29.5456$	$217.4 \pm 7.1944$	$149.8 \pm 29.5456$	$16.6 \pm 7.1944$
	Bias_fp	Baseline	$0.749 \pm 0.0098$	$0.6768 \pm 0.0133$	$376.6 \pm 2.7857$	$90.8 \pm 6.6888$	$13.4 \pm 2.7857$	$143.2 \pm 6.6888$
		Regularized	$0.8054 \pm 0.0259$	$0.7522 \pm 0.0379$	$376.4 \pm 6.063$	$126.2 \pm 20.4436$	$13.6 \pm 6.063$	$107.8 \pm 20.4436$
	Random	Baseline	$0.6724 \pm 0.0123$	$0.6528 \pm 0.0157$	$285.2 \pm 4.2237$	$134.4 \pm 7.5206$	$104.8 \pm 4.2237$	$99.6 \pm 7.5206$
		Regularized	$0.8304 \pm 0.0106$	$0.8158 \pm 0.008$	$341 \pm 16.9145$	$177.2 \pm 11.8634$	$49 \pm 16.9145$	$56.8 \pm 11.8634$