

	Attack	AUC-ROC (%)					TPR@5FPR (%)				
		APGD- L_∞		APGD- L_2		CW- L_2	APGD- L_∞		APGD- L_2		CW- L_2
Image Data	Att. ϵ										
	Detector	Low	High	Low	High		Low	High	Low	High	
MNIST (CNN)	RC	73.3	51.6	50.8	51.2	99.9	49.6	0.0	0.0	0.0	100.0
	FS	99.8	74.8	75.1	66.9	99.9	99.6	2.1	30.6	3.6	100.0
	LID	61.3	98.7	43.4	82.0	63.0	19.3	93.9	7.1	45.9	16.1
	Odds	98.7	99.7	96.2	96.3	95.5	97.3	99.9	78.4	80.5	75.2
	ML-LOO	99.7	100.0	94.1	100.0	59.0	99.0	100.0	72.5	100.0	11.5
	PN	91.1	63.6	56.2	57.2	97.2	69.4	6.8	10.9	6.6	90.8
	BAARD	97.0	98.9	93.0	97.4	96.0	84.0	97.9	56.9	85.1	75.2
CIFAR10 (ResNet18)	RC	50.2	54.8	56.8	55.0	99.4	3.7	0.0	13.4	0.0	98.5
	FS	95.4	70.1	95.1	82.0	90.4	75.3	24.3	78.7	43.0	6.4
	LID	83.4	99.3	73.2	99.0	35.5	45.6	96.9	31.5	95.4	10.9
	Odds	98.1	65.7	97.5	80.3	96.3	96.0	0.1	95.6	2.9	83.2
	ML-LOO	65.2	99.5	56.5	98.8	66.9	25.6	98.5	15.7	96.4	12.3
	PN	76.3	55.3	76.0	57.7	67.1	20.8	8.5	19.5	11.5	11.9
	BAARD	83.9	100.0	72.6	100.0	98.1	50.3	100.0	23.6	100.0	93.5
Tabular Data	Attack (Model)	PGD- L_∞ (SVM)		DTA (DT)			PGD- L_∞ (SVM)		DTA (DT)		
Banknote	RC FS, LID, etc. BAARD	-	-	-			-	-	-		
Breast Cancer	RC FS, LID, etc. BAARD	-	-	-			-	-	-		