

# Membership Inference: Results

Defense	Dataset	Acc.	Global Attacker		Local Attacker	
			Pass.	Act.	Pass.	Act.
No Defense	CIFAR100	82%	84%	91%	73%	75%
	Purchase100	84%	71%	82%	65%	68%
	Texas100	56%	65%	71%	62%	66%
Norm Bound. ( $S = 15$ )	CIFAR100	81%	-	-	72%	74%
	Purchase100	82%	-	-	64%	67%
	Texas100	55%	-	-	62%	65%
Weak DP ( $S = 15$ , $\sigma = 0.006$ )	CIFAR100	76%	-	-	70%	71%
	Purchase10	74%	-	-	62%	65%
	Texas100	50%	-	-	60%	61%
LDP ( $\epsilon = 8.6$ )	CIFAR100	68%	58%	53%	52%	55%
	Purchase100	65%	51%	62%	58%	54%
	Texas100	48%	55%	59%	56%	58%
CDP ( $\epsilon = 5.8$ )	CIFAR100	69%	-	-	58%	52%
	Purchase100	70%	-	-	53%	55%
	Texas100	45%	-	-	54%	52%

We measure attack accuracy as the fraction of correct membership predictions for unknown data points.  
(Baseline is 50%)

# Property Inference: Results

Labeled Faces In the Wild (LFW) dataset