Conversion of classifier output to probabilities:
We can convert output of classifiers d=wTx+w0 into a probability with the sigmoid function

$$\text{Sigmoid}(d) = \frac{1}{(1 + e^{\wedge}(-d))}$$

Evaluation of classifiers
1. We have been using error throughout
2. We can also evaluate probabilities with the Receiver Operating Characteristic (ROC) curve and its area under curve (AUC).

To understand ROC curve consider the data below

Y	probability	for threshold				
		.3	.5	1	0	.75
		Predicted labels				
0	.2	0	0	0	1	0
0	.6	1	1	0	1	0
0	.4	1	0	0	1	0
1	.8	1	1	0	1	1
1	.1	0	0	0	1	0
1	.7	1	1	0	1	0

fp=2,tp=2	fp=1,tp=2	fp=0,tp=0	tp=3,fp=3	fp=0 tp=1
fpr=⅔,tpr=2/3	fpr=⅓,tpr=⅔	fpr=0,tpr=0	tpr=1,fpr=1	fpr=0 tpr=1/3

The false positive (fp) is the number of wrong positive predictions. The true positive is the number of correct positive predictions. The false positive rate (fpr) is false positive divided by number of negatives (or label 0). The true positive rate is true positive divided by number of positives (or label 1).

We obtain the ROC curve by plotting the fpr on the x-axis and tpr on the y-axis.