

Introspective Classification for Robot Perception

Hugo Grimmert

Why classify?

Why classify?

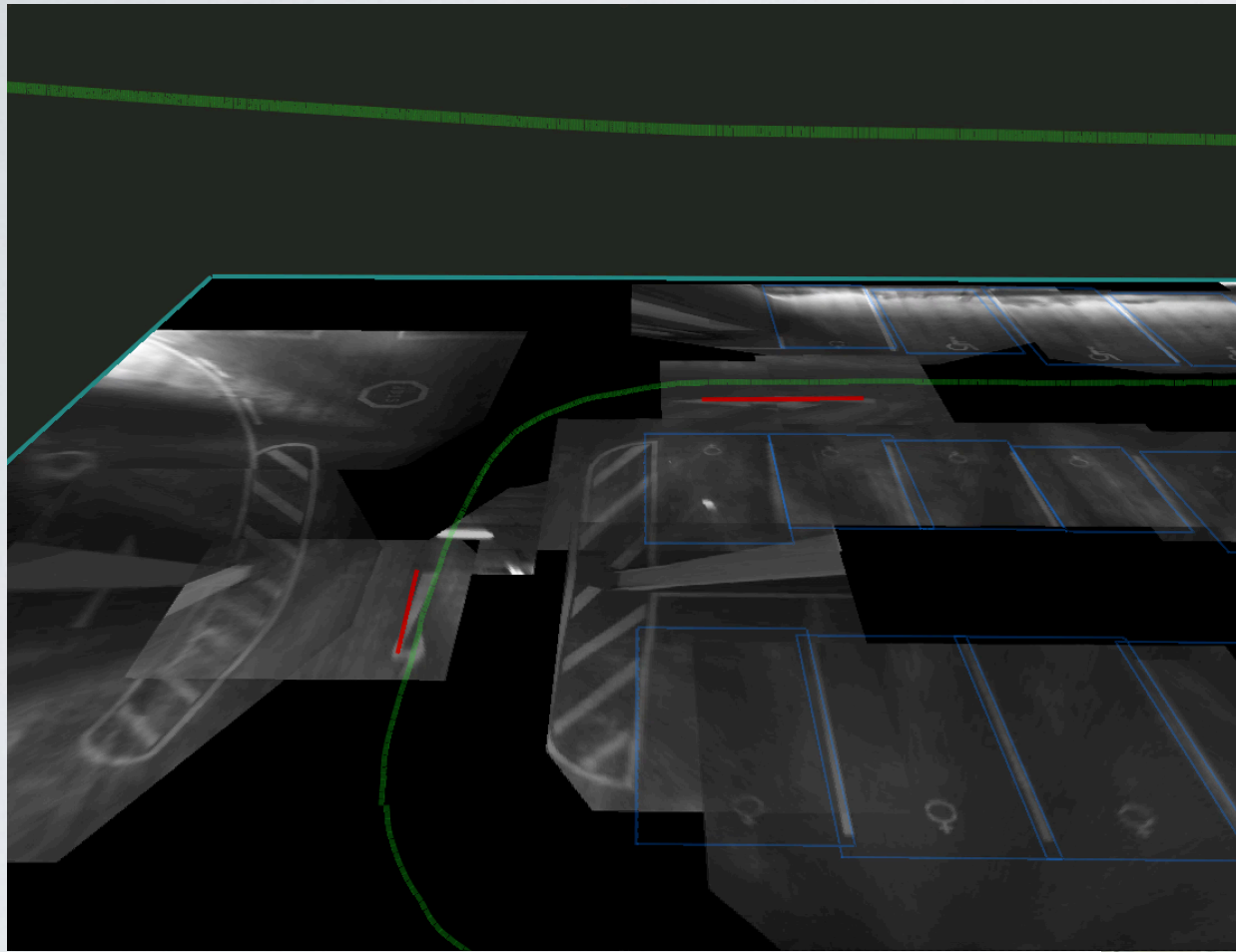


Pedestrians

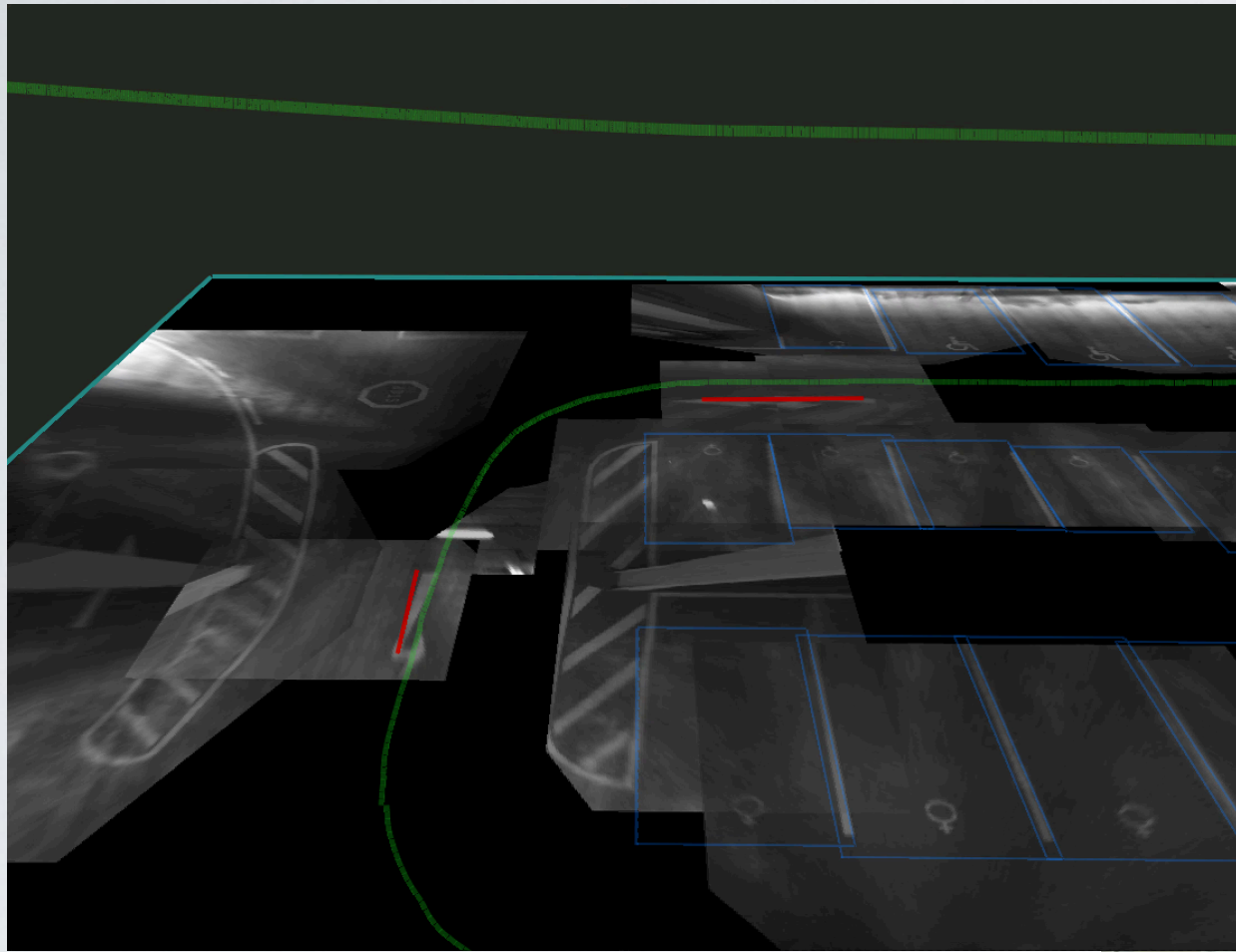
Cars



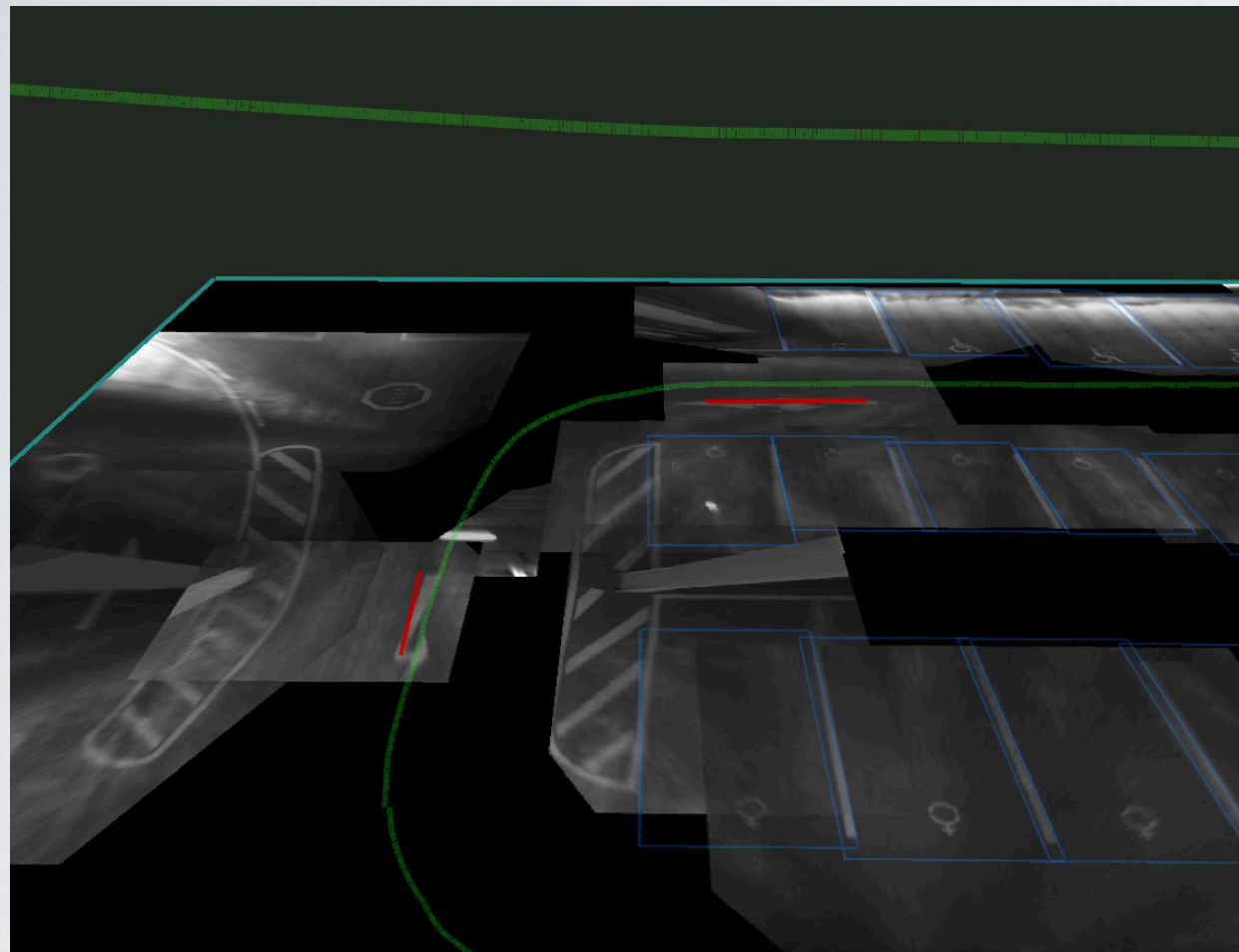
Classification is about making decisions.



Classification is about making decisions.



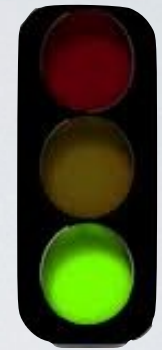
Classification is about making decisions.



Do we



or



?



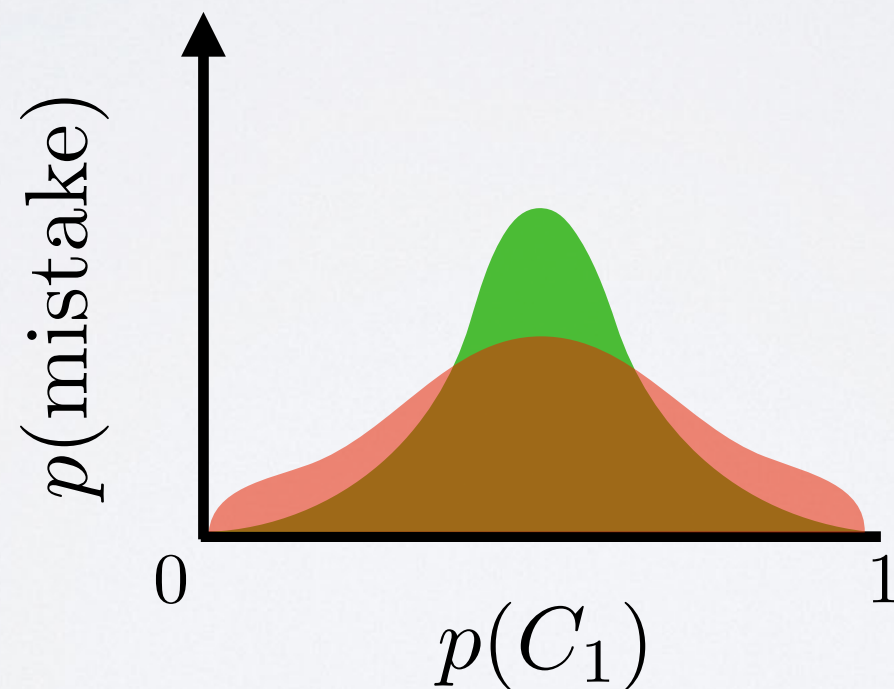
Is my robot going to make good decisions?

Is my robot going to make good decisions?

$$\frac{tp}{tp + fp} \quad \& \quad \frac{tp}{tp + fn} \quad ?$$

Is my robot going to make good decisions?

$$\frac{tp}{tp + fp} \quad \& \quad \frac{tp}{tp + fn} \quad ?$$



Overview of decision making process

Overview of decision making process

Scene



Overview of decision making process

Scene



Image



Overview of decision making process

Scene



Image



Feature

$$\mathbf{x}_k = \begin{bmatrix} 0.30 \\ 0.42 \\ -0.29 \\ \vdots \end{bmatrix}^T_{\mathbf{d}}$$

Overview of decision making process

Scene



Image



Feature

$$\mathbf{x}_k = \begin{bmatrix} 0.30 \\ 0.42 \\ -0.29 \\ \vdots \end{bmatrix}^T_d$$

Classifier

$$p(\text{cake}) = 0.8$$

Overview of decision making process

Scene



Image



Feature

$$\mathbf{x}_k = \begin{bmatrix} 0.30 \\ 0.42 \\ -0.29 \\ \vdots \end{bmatrix}^T_d$$

Classifier

$$p(\text{cake}) = 0.8$$

Costs

Cost	P	N
T	0	0
F	1	10

Overview of decision making process

Scene



Image



Feature

$$\mathbf{x}_k = \begin{bmatrix} 0.30 \\ 0.42 \\ -0.29 \\ \vdots \end{bmatrix}^T_d$$

Classifier

$$p(\text{cake}) = 0.8$$

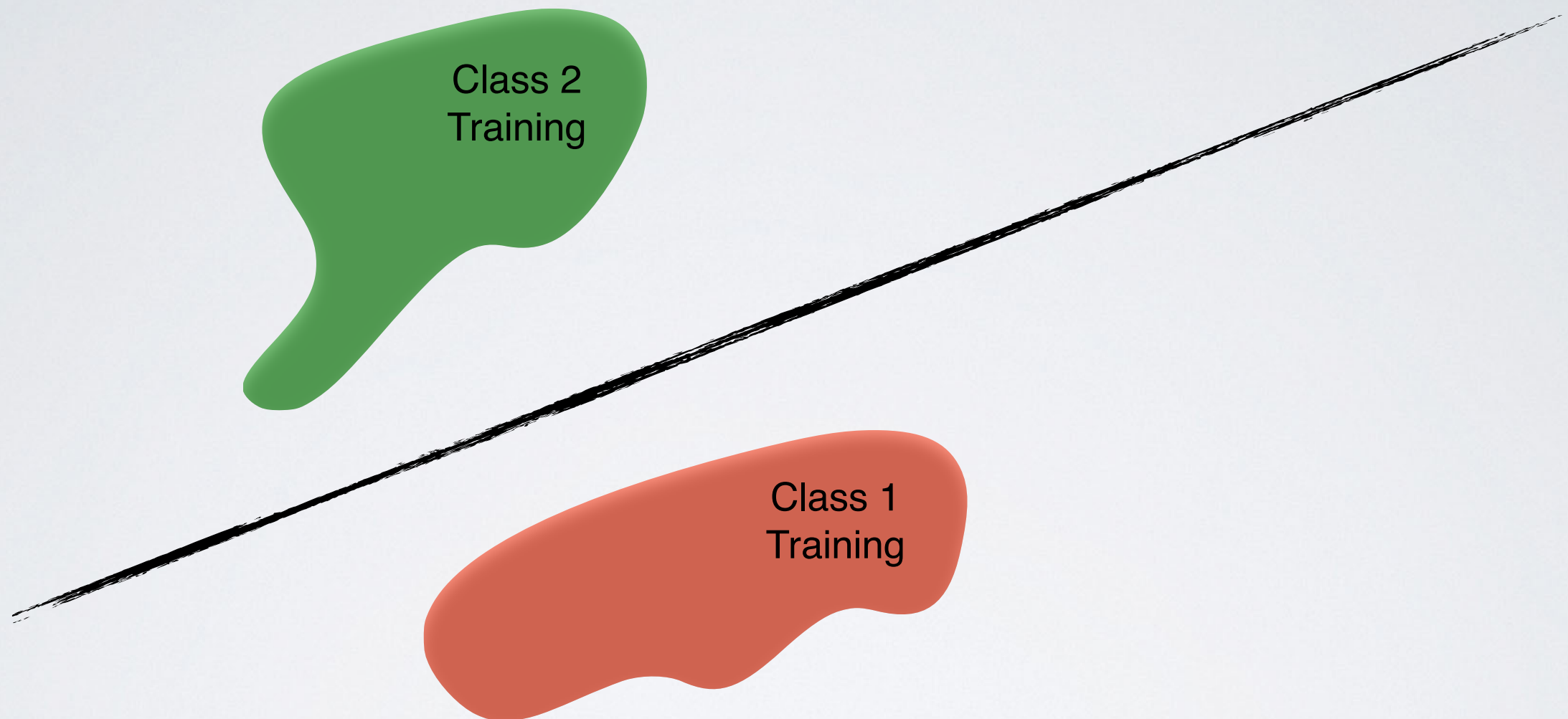
Costs

Cost	P	N
T	0	0
F	1	10

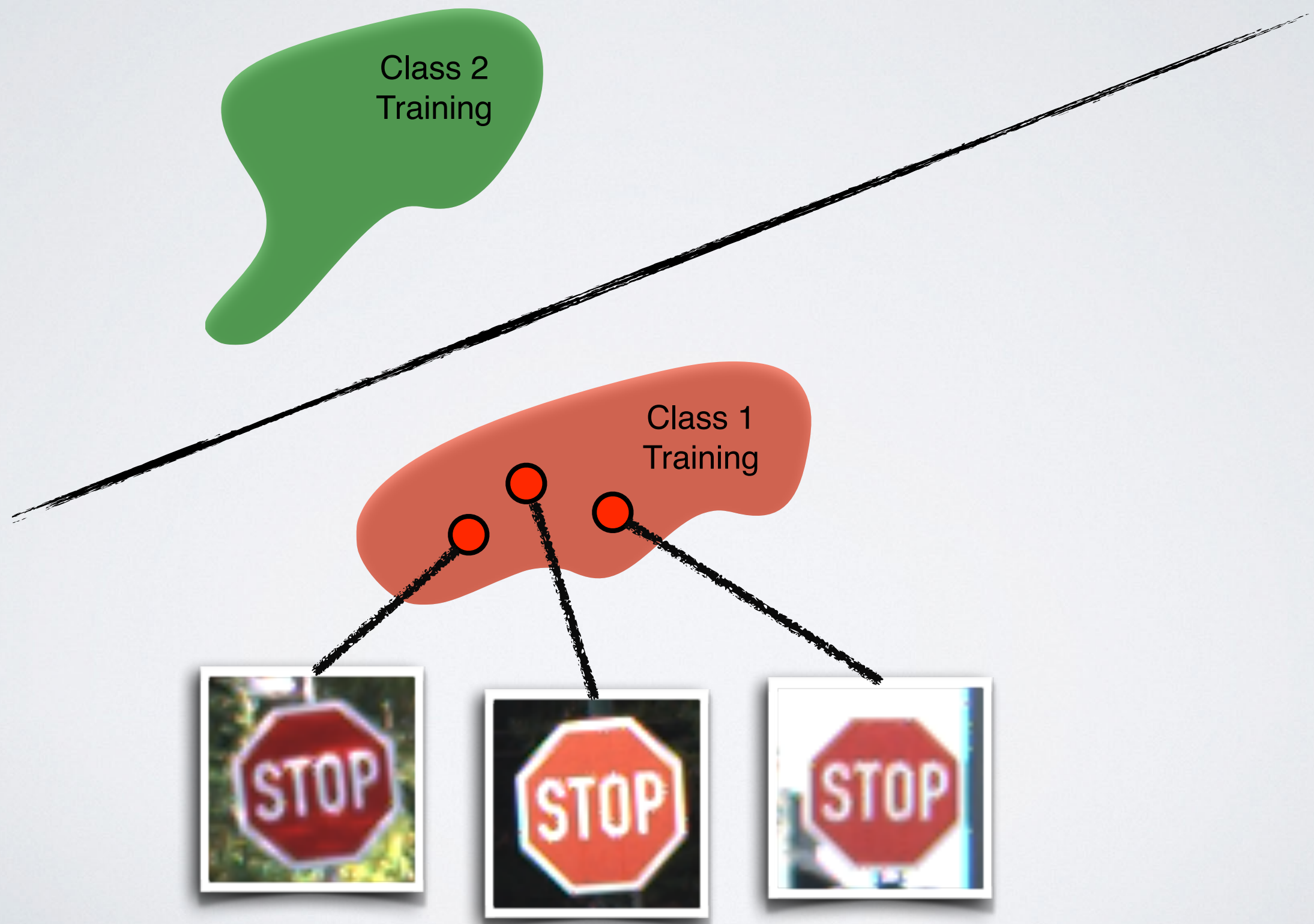
Action



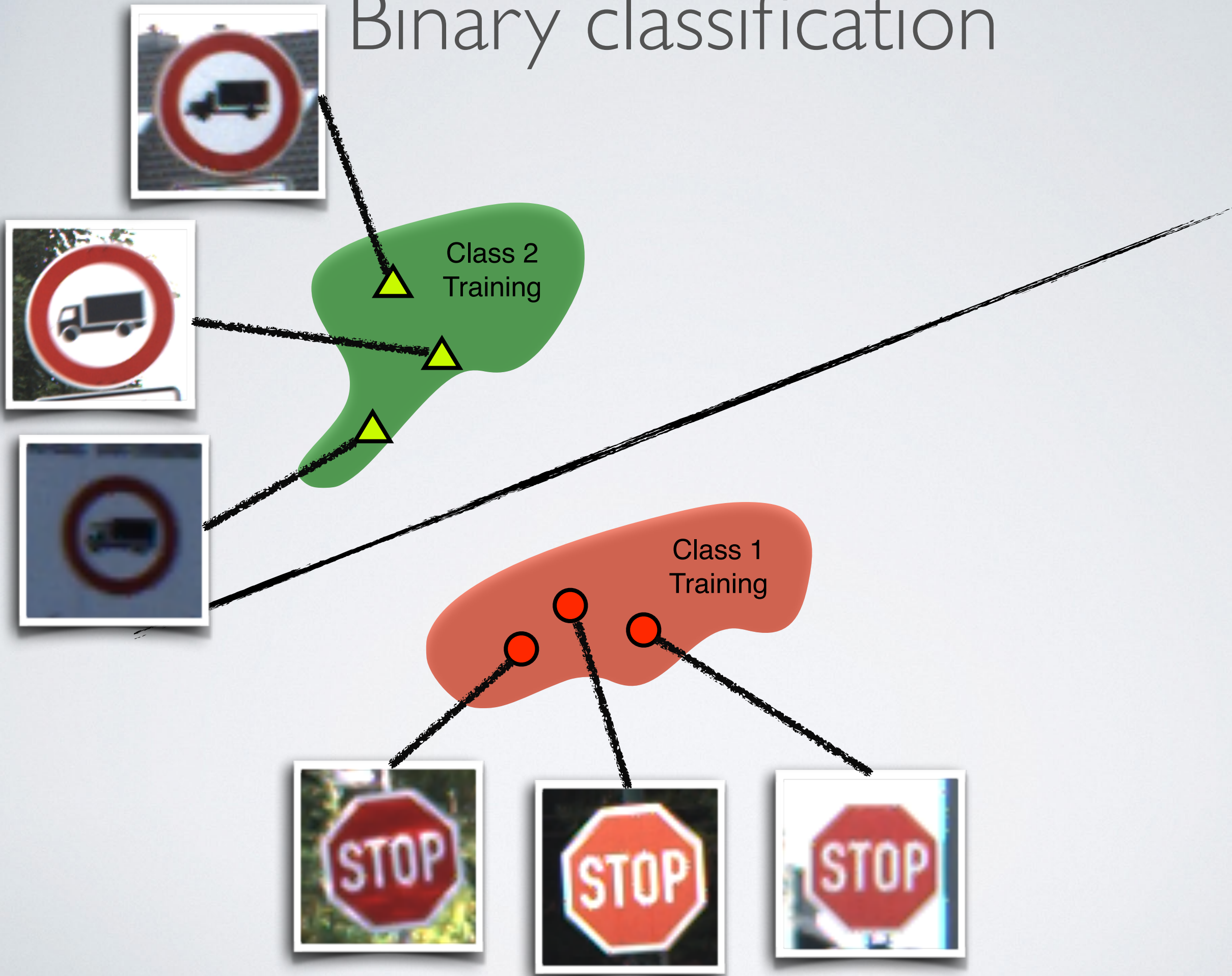
Binary classification



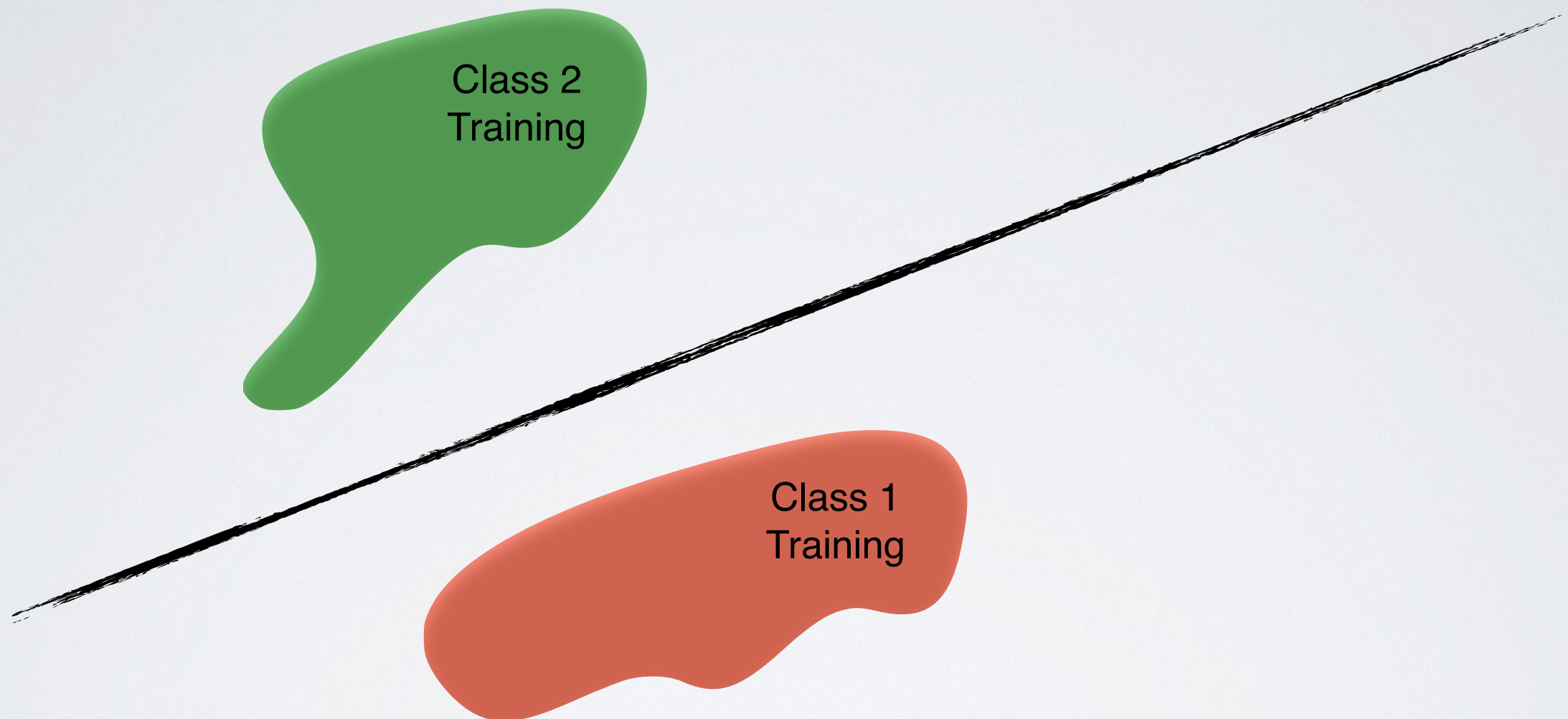
Binary classification



Binary classification



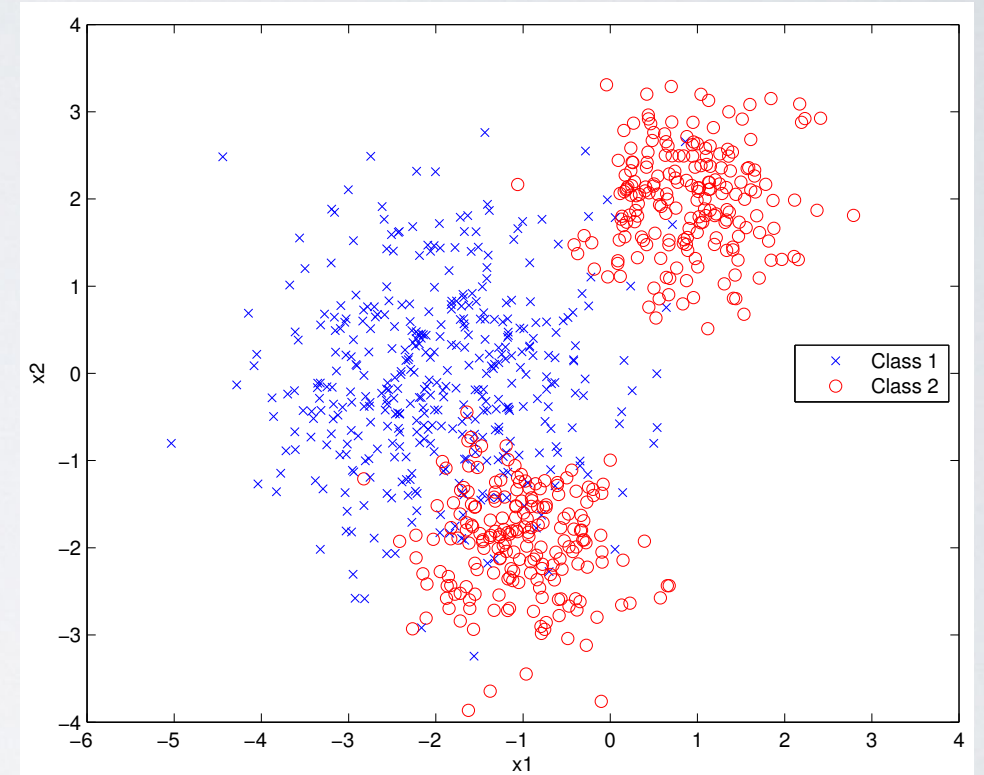
Binary classification



Non-linear classifiers

Many of the classifiers we use are inherently linear, but not all data are linearly separable

So we can augment our linear classifiers!



Cover's theorem paraphrased:

all data are linearly separable into two classes if you lift them to a sufficiently high dimension.

The common metrics for learning

Precision

$$p = \frac{\text{tp}}{\text{tp} + \text{fp}}$$

Recall

$$r = \frac{\text{tp}}{\text{tp} + \text{fn}}$$

F-measure

$$F_1 = 2 \cdot \frac{p \times r}{p + r}$$

The common metrics for learning

Precision

$$p = \frac{tp}{tp + fp}$$

Recall

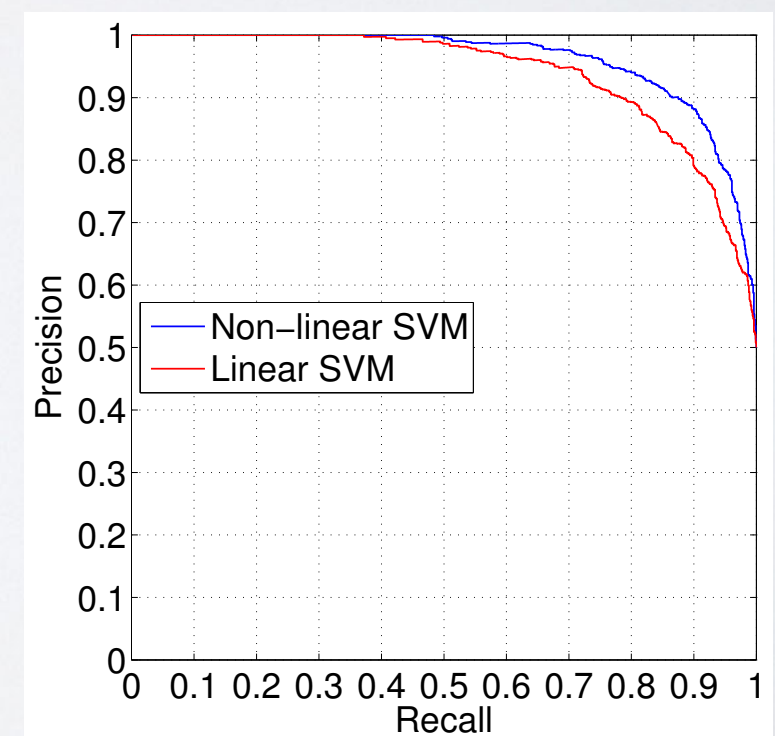
$$r = \frac{tp}{tp + fn}$$

F-measure

$$F_1 = 2 \cdot \frac{p \times r}{p + r}$$

PR curves change the *operating point*,
which is the boundary between positive and negative

scores = [0.1 0.4 0.3 0.9 0.8 0.6]



The common metrics for learning

Precision

$$p = \frac{tp}{tp + fp}$$

Recall

$$r = \frac{tp}{tp + fn}$$

F-measure

$$F_1 = 2 \cdot \frac{p \times r}{p + r}$$

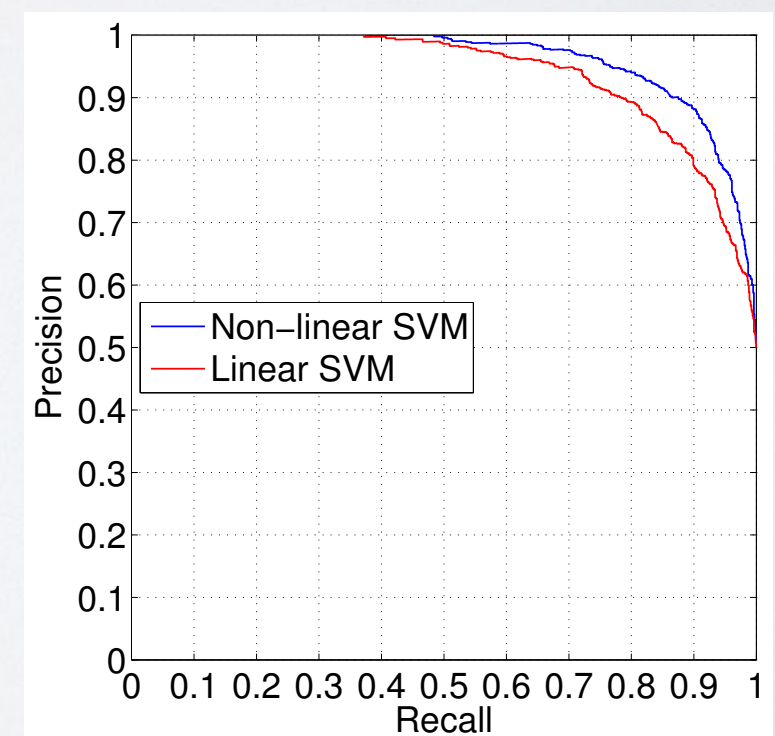
PR curves change the *operating point*,
which is the boundary between positive and negative

scores = [0.1 0.4 0.3 0.9 0.8 0.6]

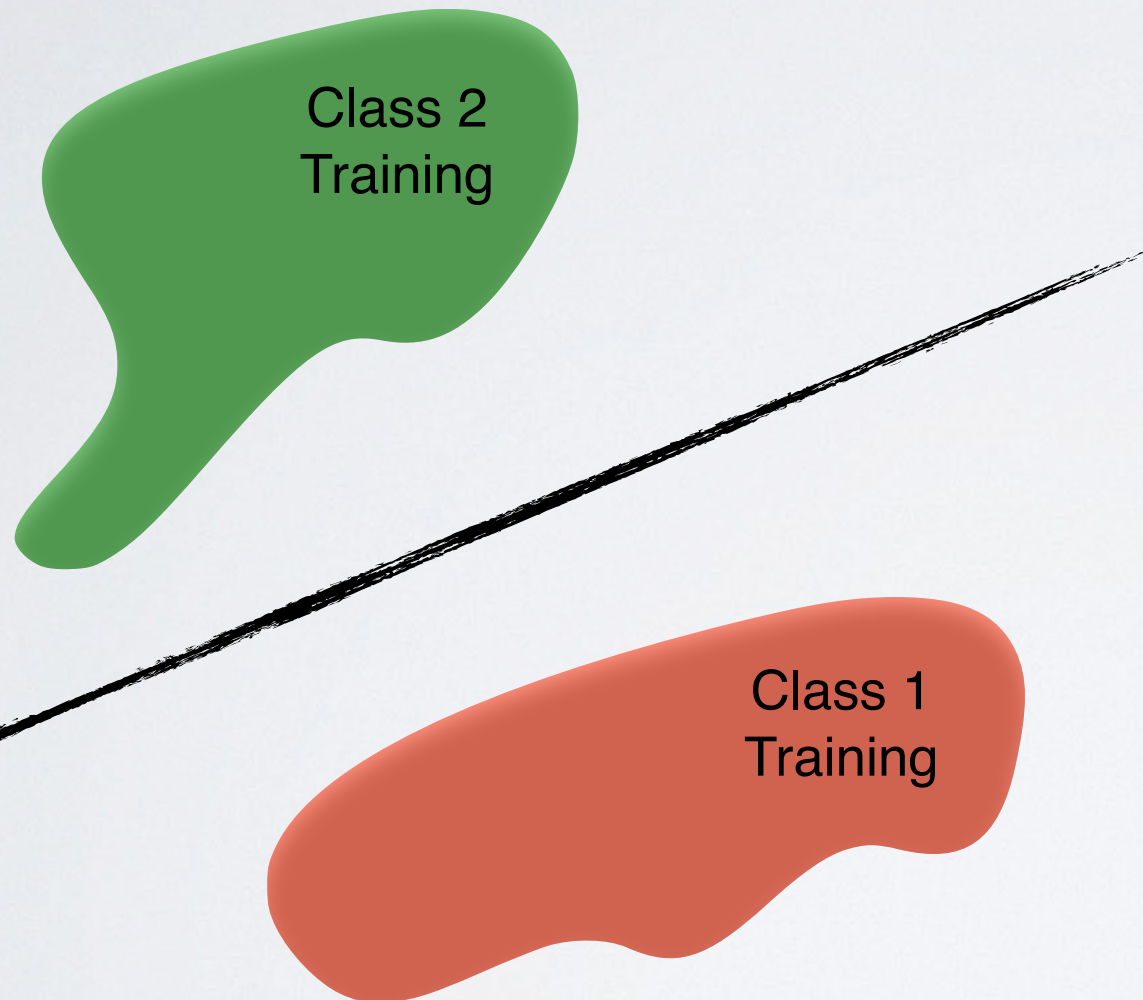
sort:

scores = [0.1 0.3 0.4 0.6 0.8 0.9]

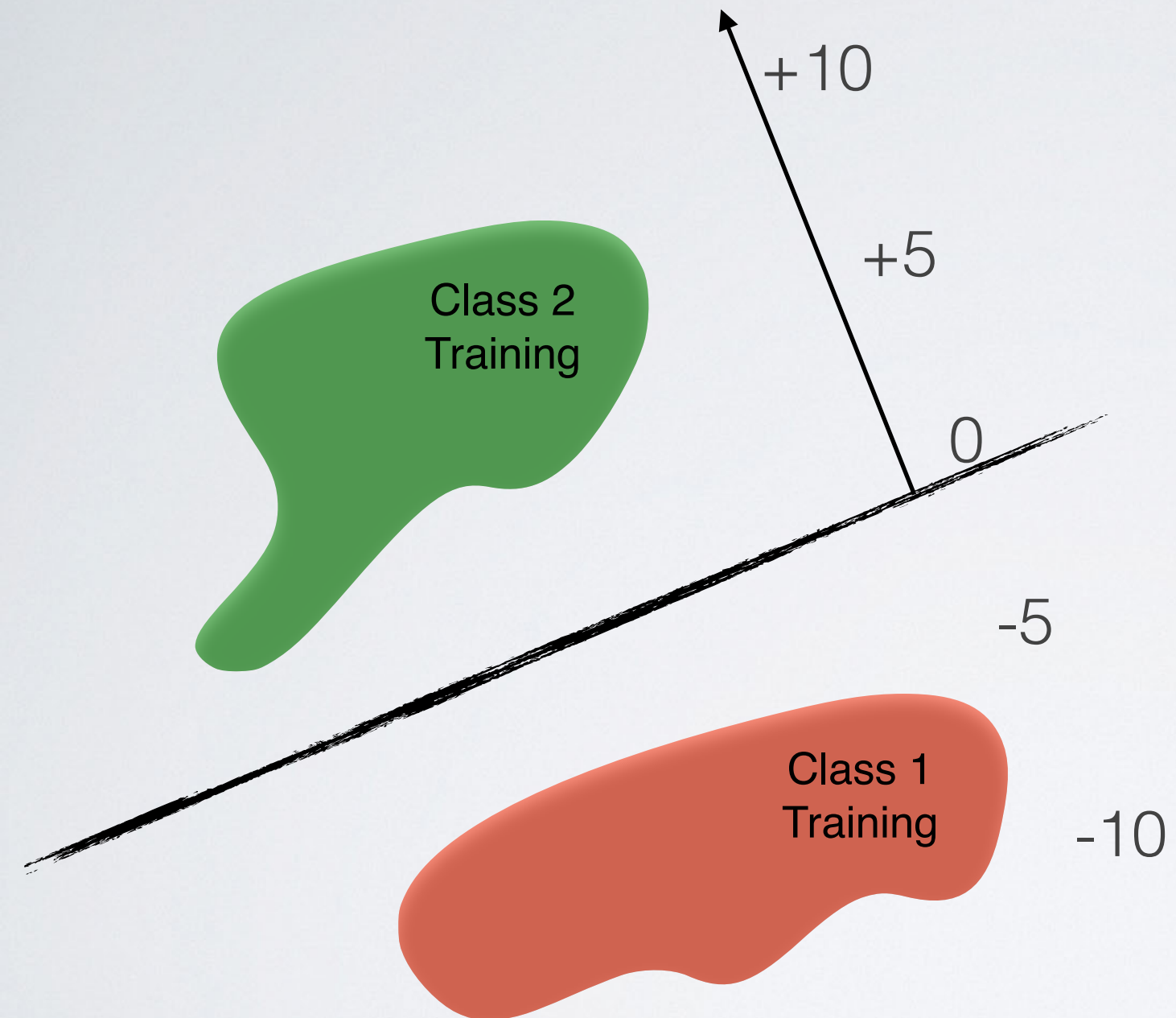
GT = [0 0 1 1 1 1]



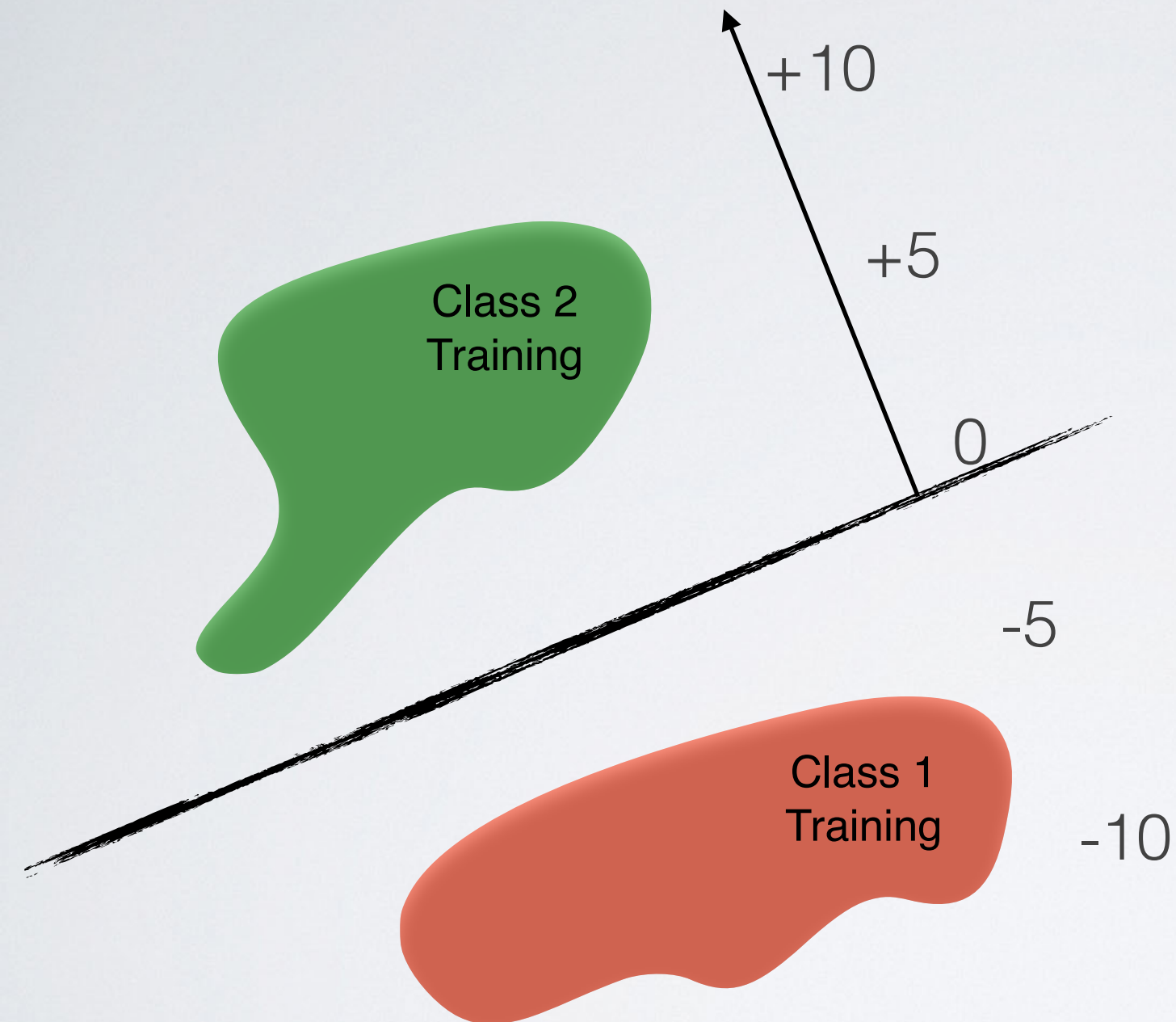
Probabilities



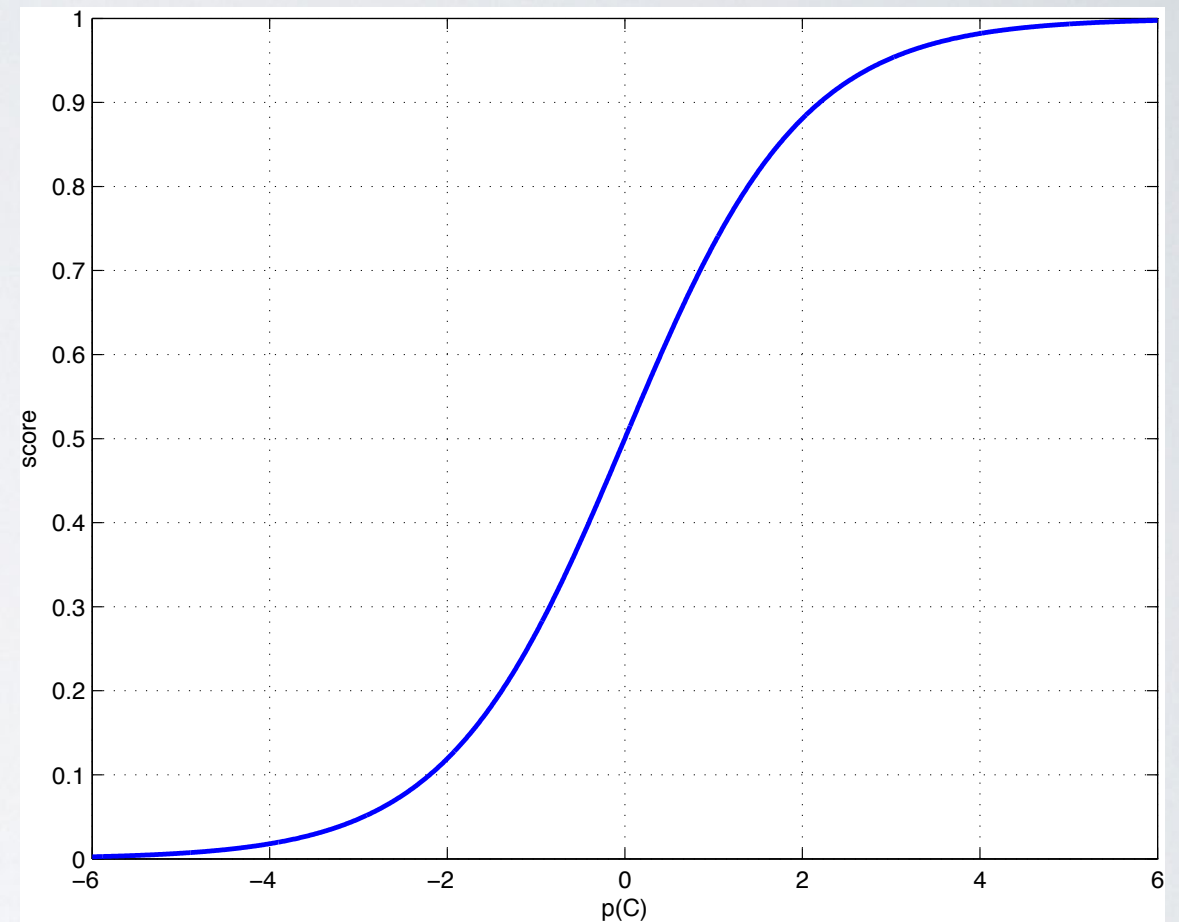
Probabilities



Probabilities



Platt's method

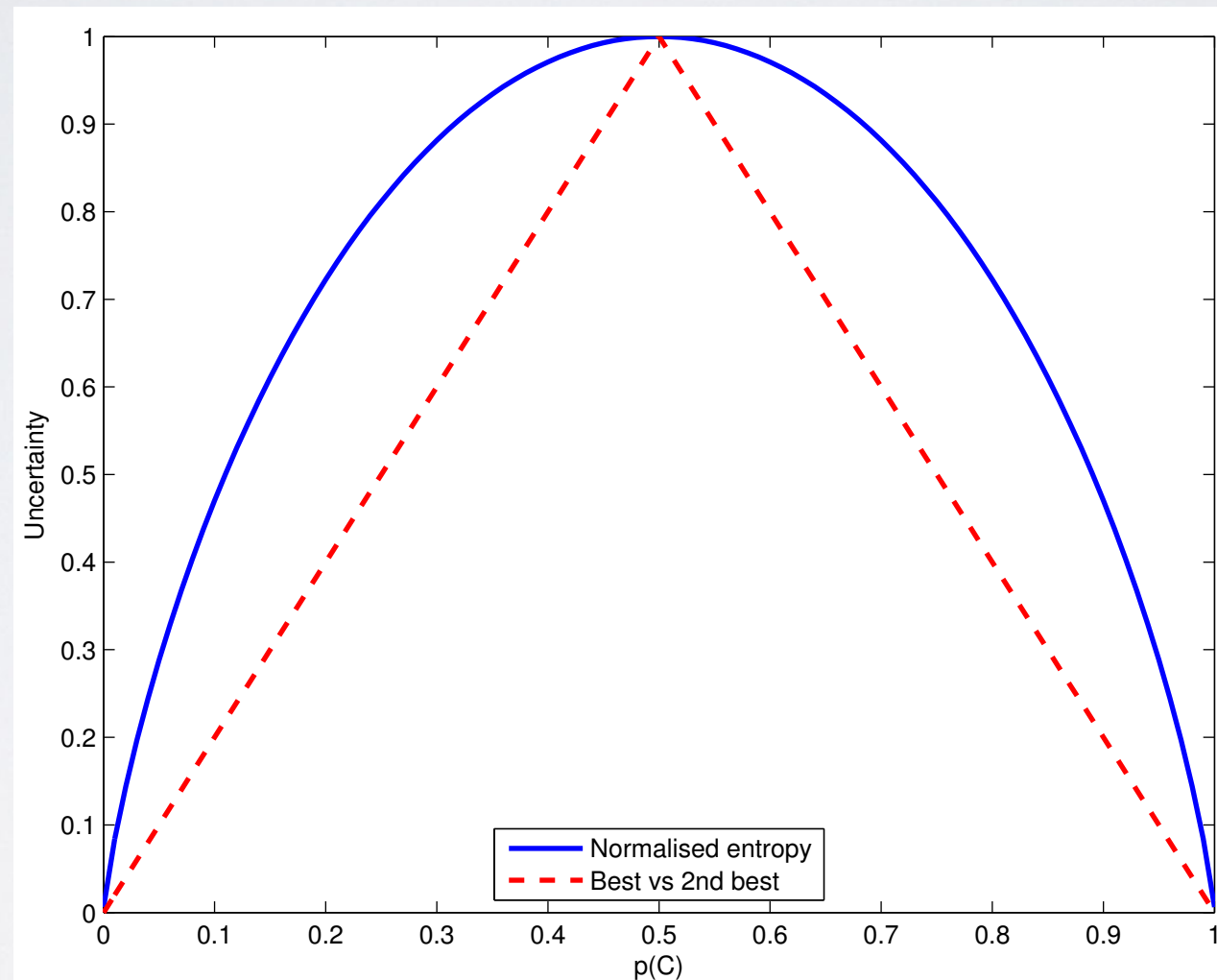


Sigmoid function

$$y = \frac{1}{1 + e^{-ax}}$$

Uncertainties

A mapping from $p(C)$ to uncertainty



Decision theory

Desired behaviour

Decision theory

Desired behaviour



Negative

$$p(C_1) \rightarrow 0$$

Decision theory

Desired behaviour



Negative

$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

Decision theory

Desired behaviour



Negative

$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

Cost functions



Cost	P	N
T	0	0
F	1	1

Decision theory

Desired behaviour



Negative

$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

Cost functions



Cost	P	N
T	0	0
F	1	1

$$\bar{L}(a) = \sum_{i=1}^{|\mathcal{C}|} L(a, C_i) p(C_i)$$

Decision theory

Desired behaviour



Negative

$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

Cost functions



Cost	P	N
T	0	0
F	1	1

$$\bar{L}(a) = \sum_{i=1}^{|\mathcal{C}|} L(a, C_i) p(C_i)$$

Decision theory

Desired behaviour



Negative

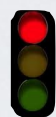
$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

Cost functions

		
Cost	P	N
T	0	0
F	1	1

$$\bar{L}(a) = \sum_{i=1}^{|\mathcal{C}|} L(a, C_i) p(C_i)$$

An arrow points from the highlighted term $L(a, C_i)$ in the equation to the 'P' column of the cost function table above.

Decision theory

Desired behaviour



Negative

$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

Cost functions



Cost	P	N
T	0	0
F	1	1

$$\bar{L}(a) = \sum_{i=1}^{|\mathcal{C}|} L(a, C_i) p(C_i)$$

Decision theory

Desired behaviour



Negative

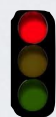
$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

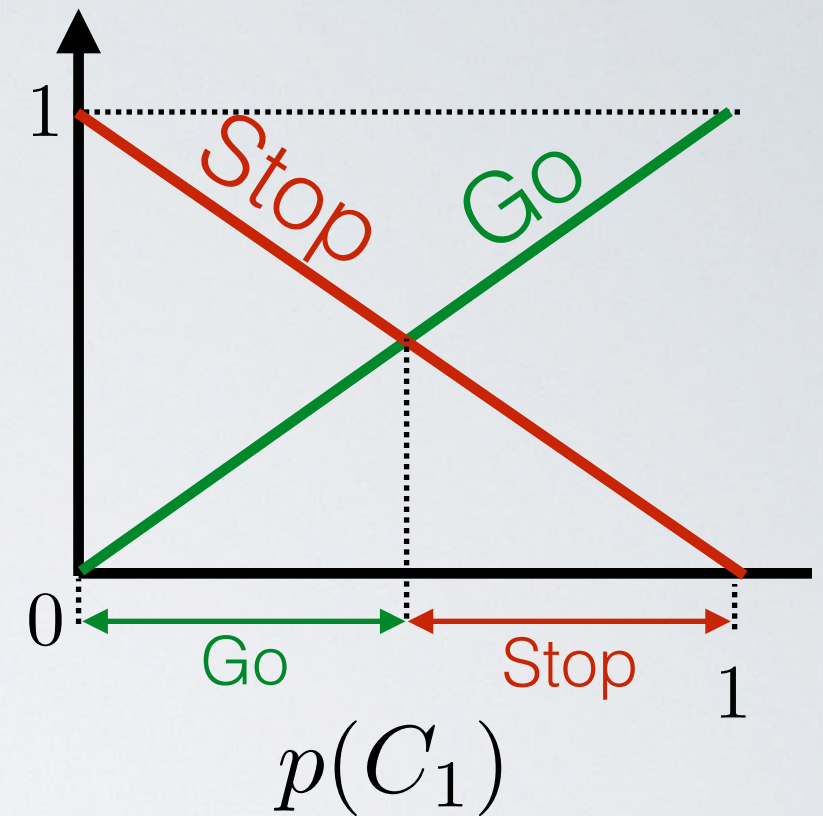
Cost functions

Cost		
	P	N
T	0	0
F	1	1

$\bar{L}(\text{action})$

$$\bar{L}(a) = \sum_{i=1}^{|\mathcal{C}|} L(a, C_i) p(C_i)$$

Expected loss



Decision theory

Desired behaviour



Negative

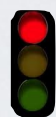
$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

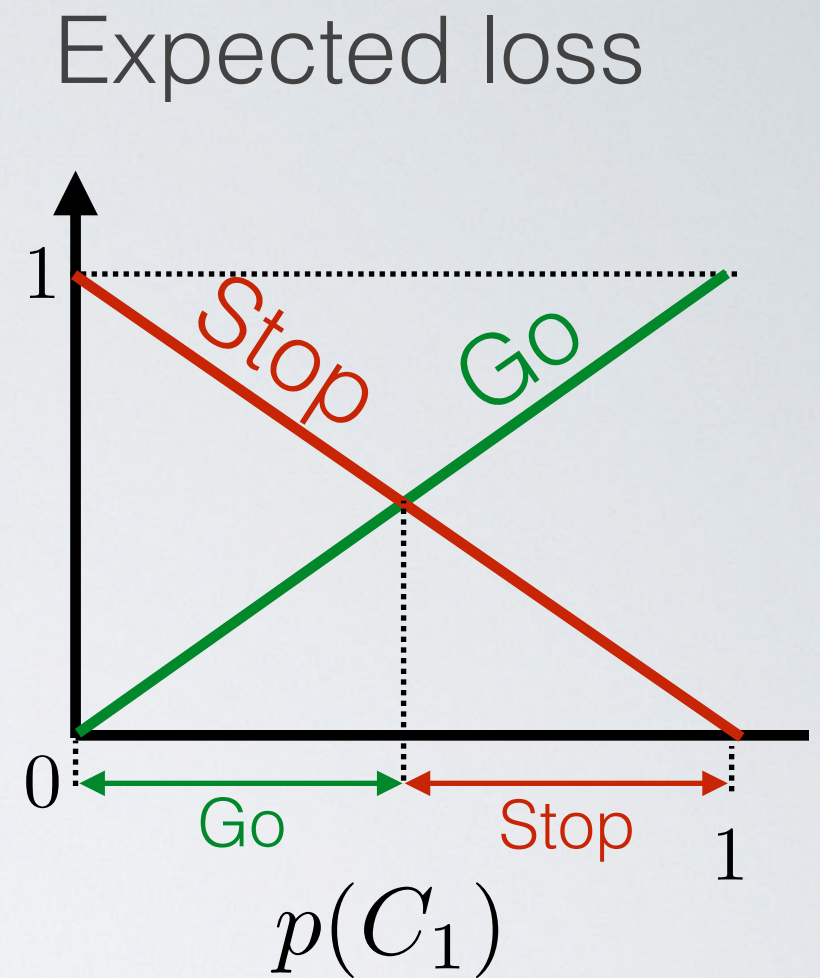
Cost functions

Cost		
	P	N
T	0	0
F	1	1

$$\bar{L}(a) = \sum_{i=1}^{|\mathcal{C}|} L(a, C_i) p(C_i)$$

Cost	P	N
T	0	0
F	1	3

$\bar{L}(\text{action})$



Decision theory

Desired behaviour



Negative

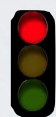
$$p(C_1) \rightarrow 0$$



Positive

$$p(C_1) \rightarrow 1$$

Cost functions

Cost		
	P	N
T	0	0
F	1	1

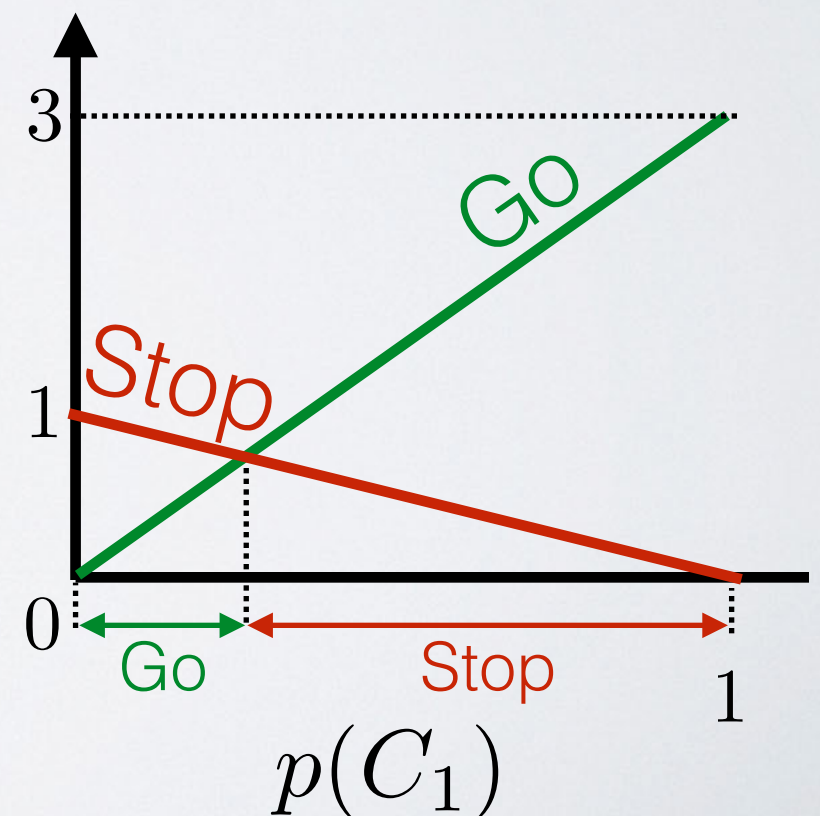
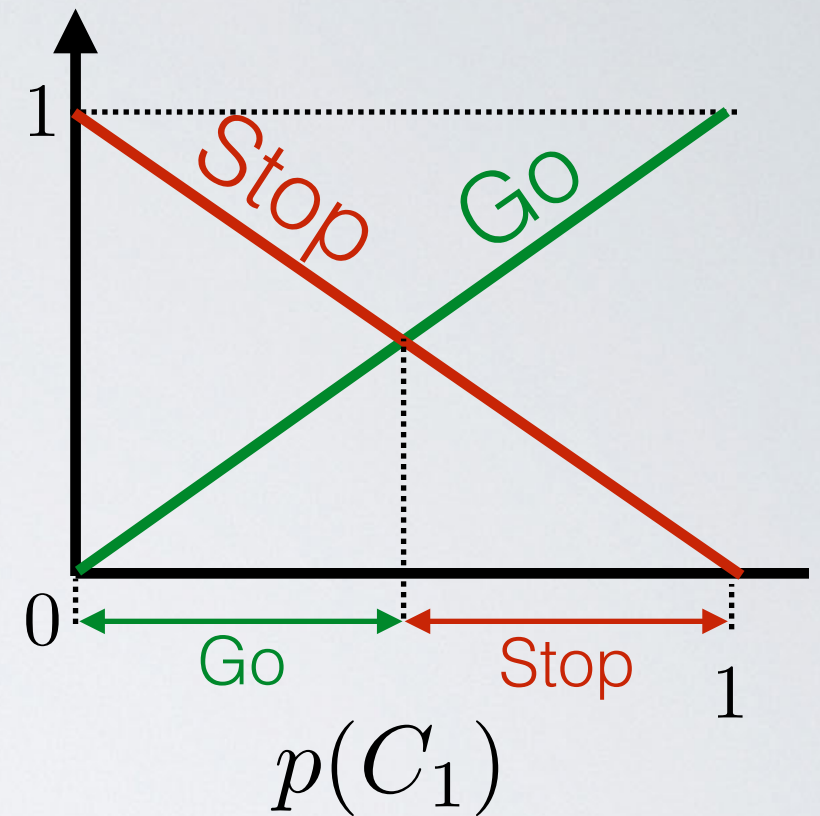
$$\bar{L}(a) = \sum_{i=1}^{|\mathcal{C}|} L(a, C_i) p(C_i)$$

Cost	P	N
T	0	0
F	1	3

$\bar{L}(\text{action})$

$\bar{L}(\text{action})$

Expected loss



Overview of decision making process

Overview of decision making process

Scene



Overview of decision making process

Scene



Image



Overview of decision making process

Scene



Image



Feature

$$\mathbf{x}_k = \begin{bmatrix} 0.30 \\ 0.42 \\ -0.29 \\ \vdots \end{bmatrix}^T_{\mathbf{d}}$$

Overview of decision making process

Scene



Image



Feature

$$\mathbf{x}_k = \begin{bmatrix} 0.30 \\ 0.42 \\ -0.29 \\ \vdots \end{bmatrix}^T_d$$

Classifier

$$p(\text{cake}) = 0.8$$

Overview of decision making process

Scene



Image



Feature

$$\mathbf{x}_k = \begin{bmatrix} 0.30 \\ 0.42 \\ -0.29 \\ \vdots \end{bmatrix}^T_d$$

Classifier

$$p(\text{cake}) = 0.8$$

Costs

Cost	P	N
T	0	0
F	1	10

Overview of decision making process



$$\mathbf{x_k} = \begin{bmatrix} 0.30 \\ 0.42 \\ -0.29 \\ \vdots \end{bmatrix}^T_d$$

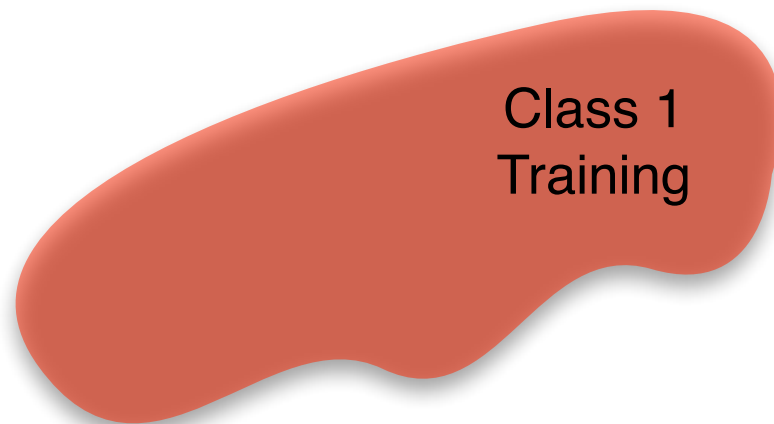
$$p(\text{cake}) = 0.8$$

Cost	P	N
T	0	0
F	1	10





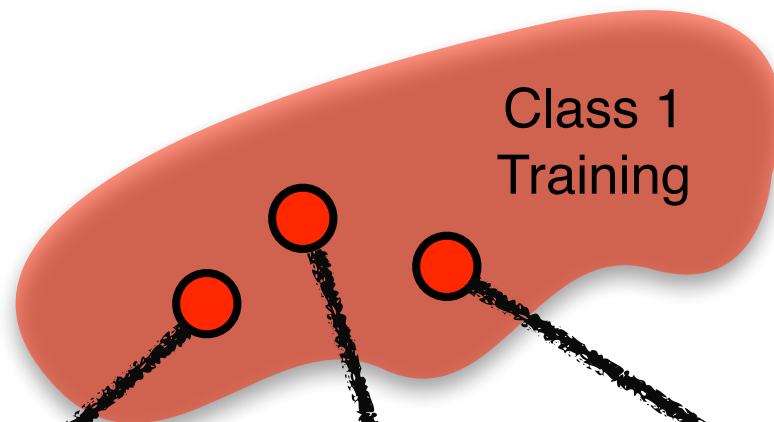
Class 2
Training



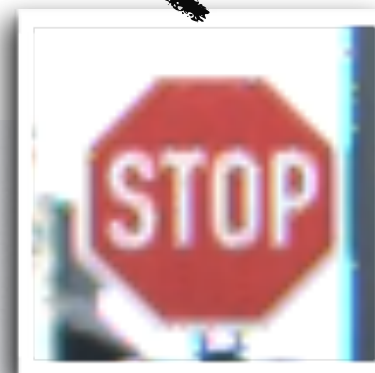
Class 1
Training

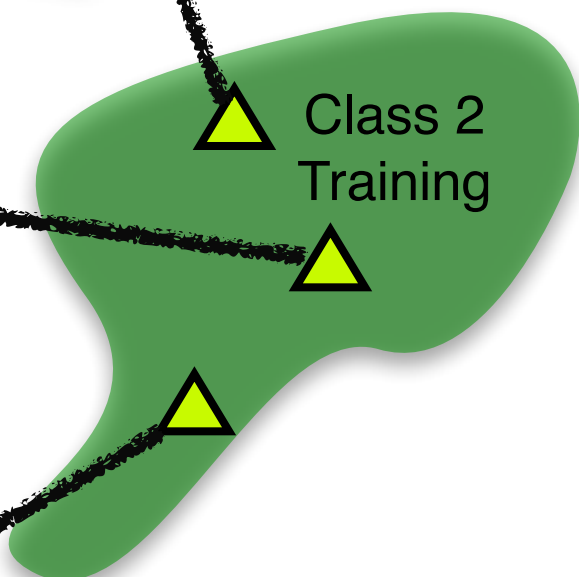


Class 2
Training

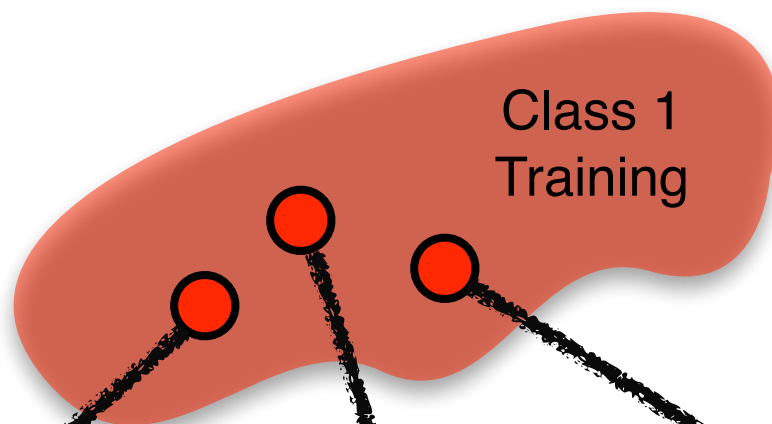


Class 1
Training

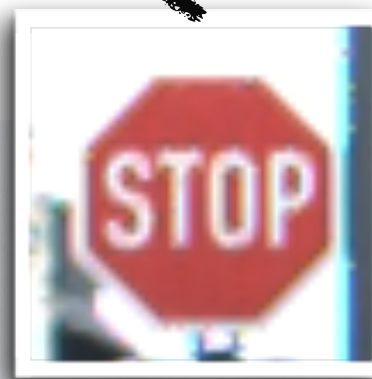


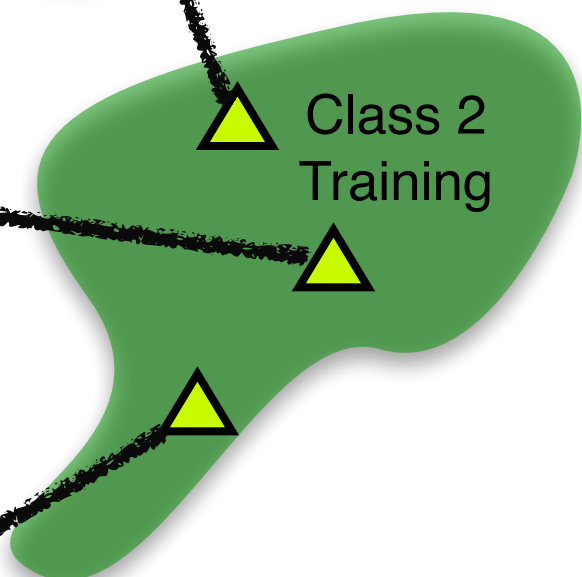


Class 2
Training

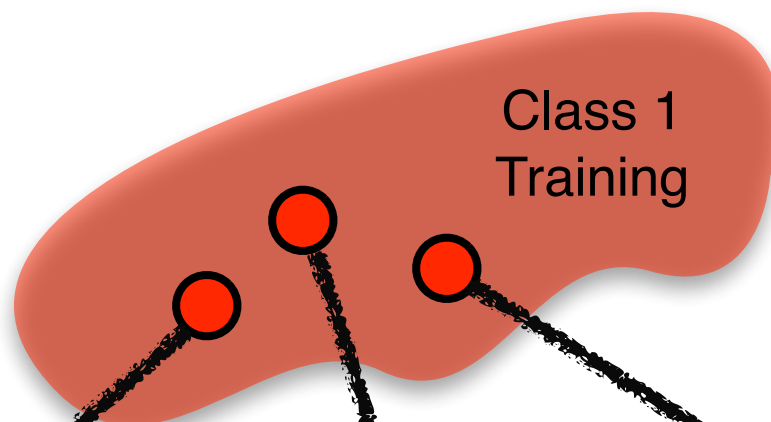


Class 1
Training

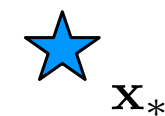
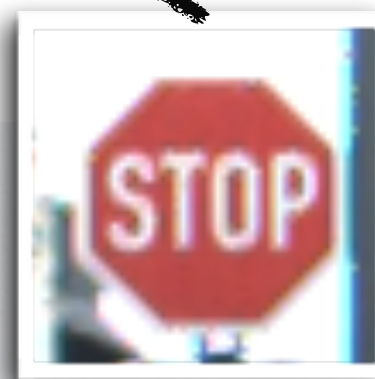


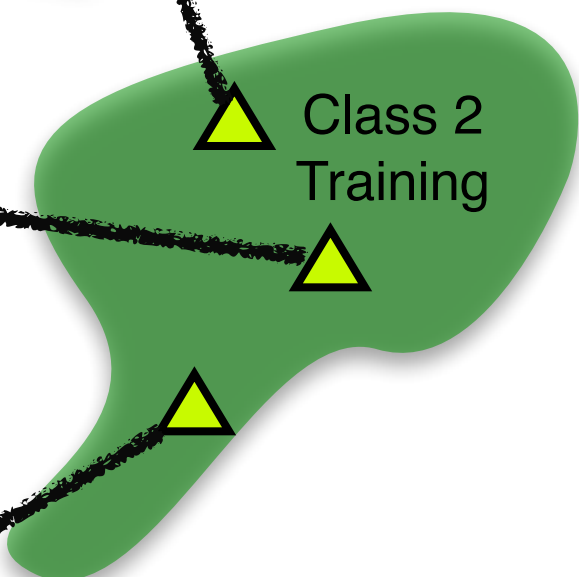


Class 2
Training

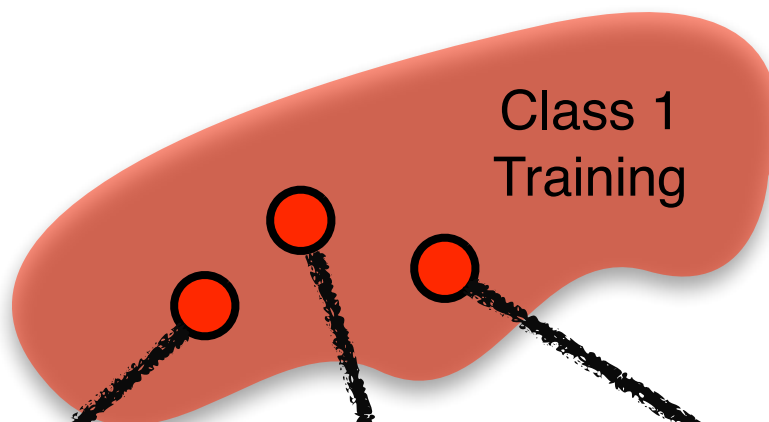


Class 1
Training

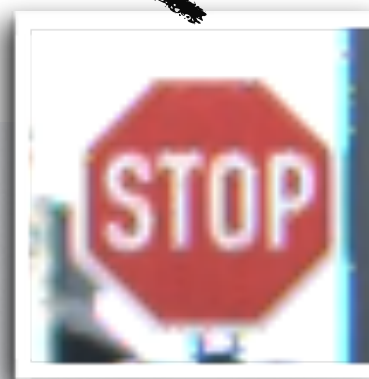




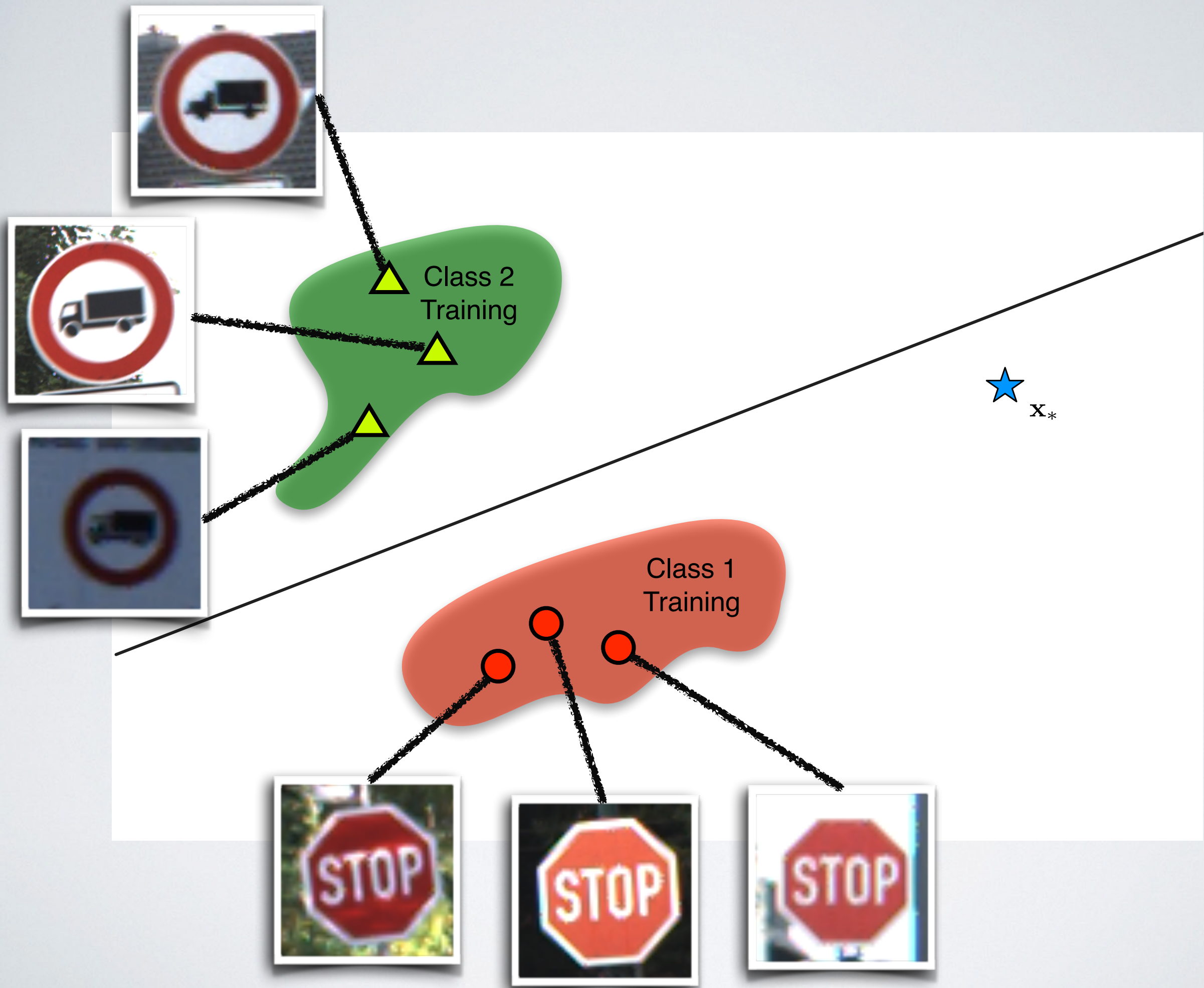
Class 2
Training

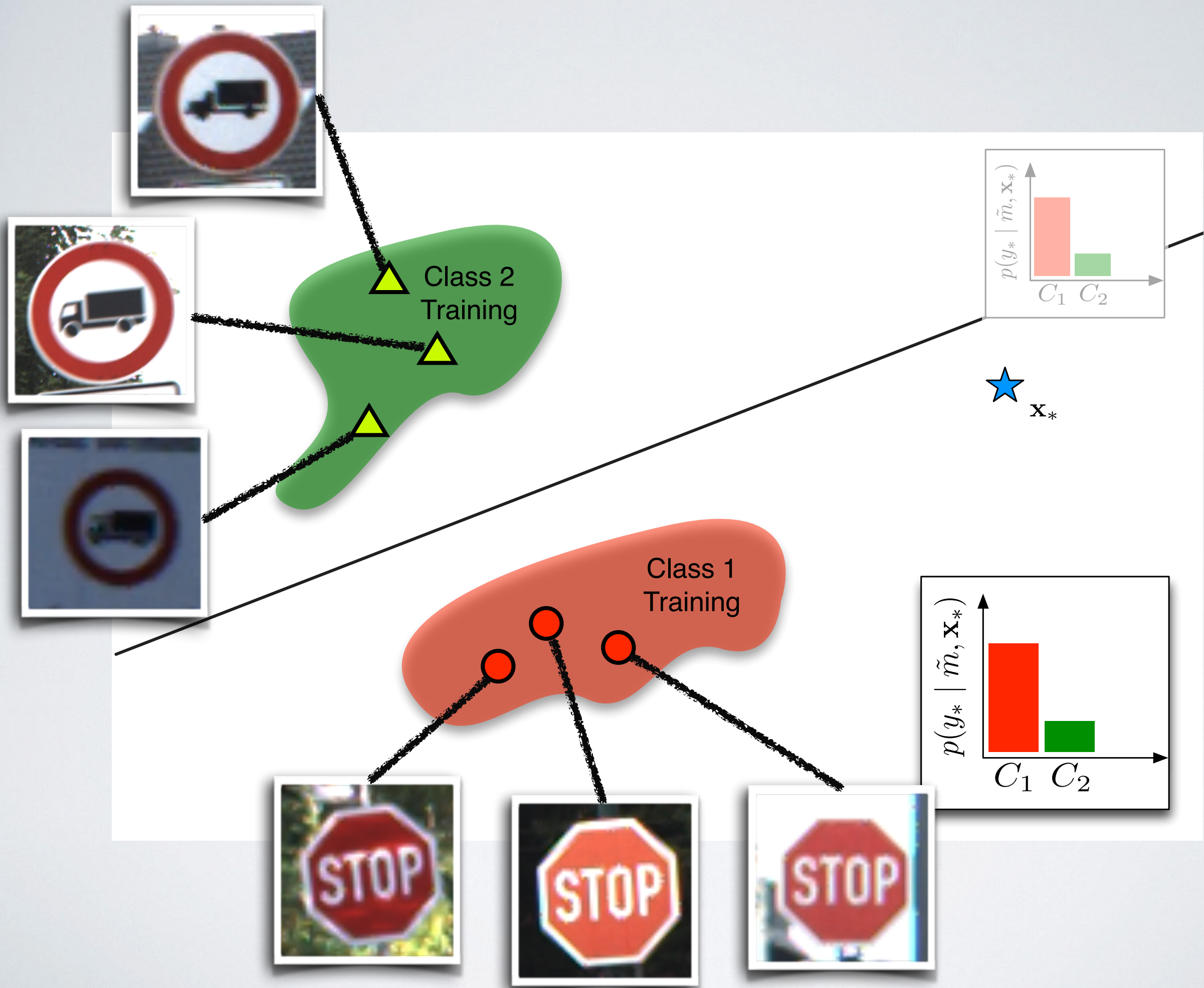


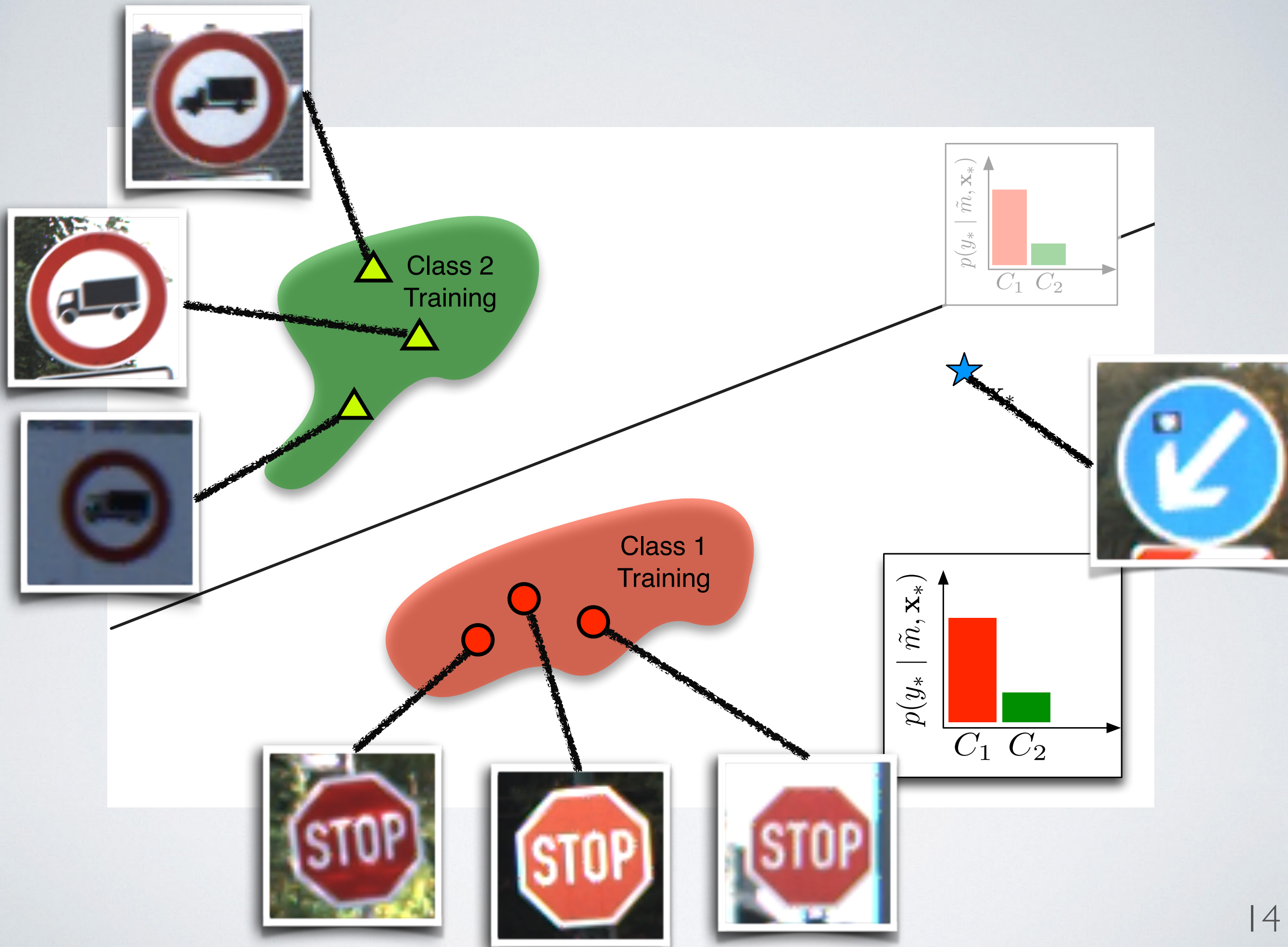
Class 1
Training

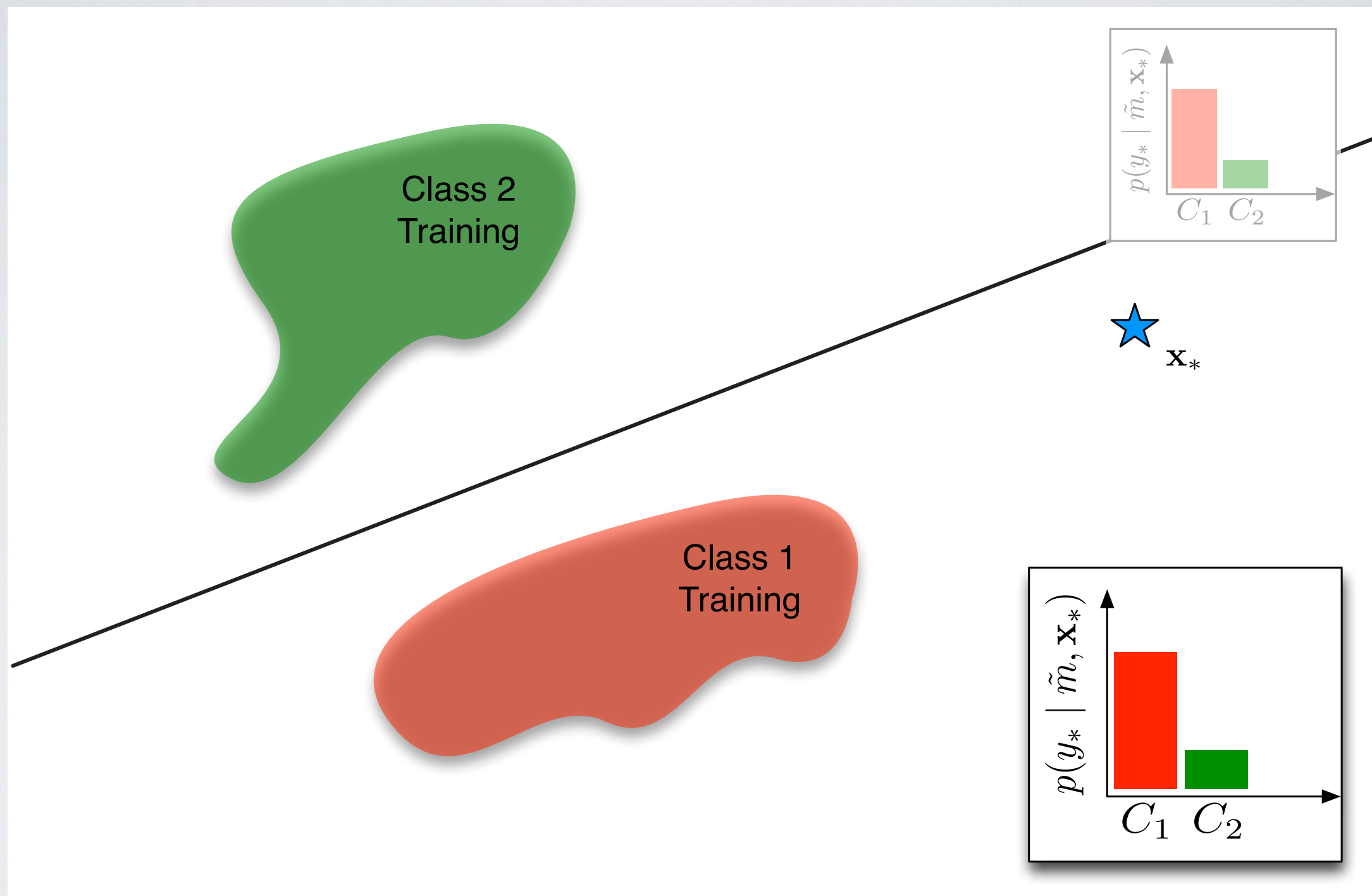


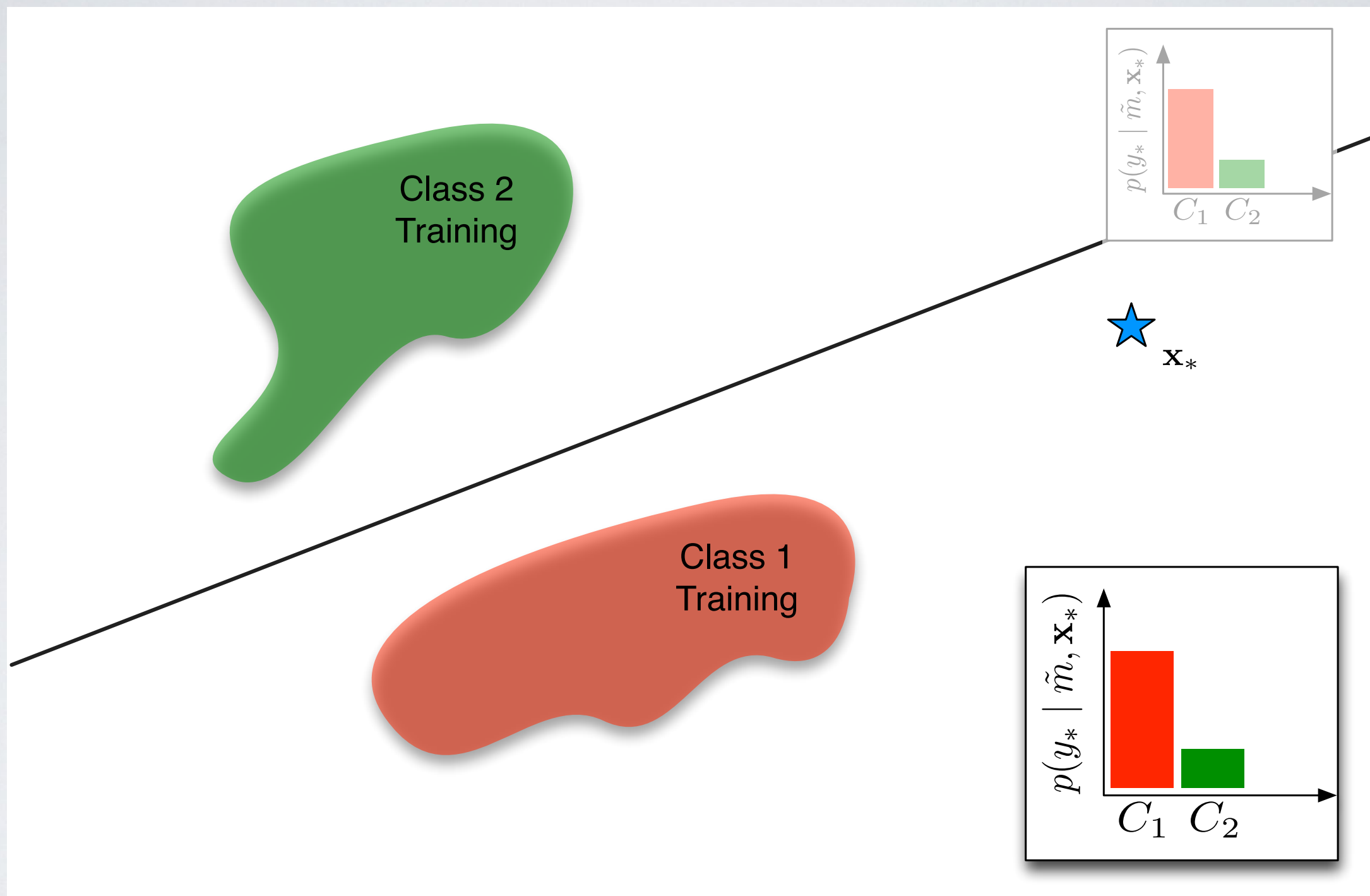
★
 x_* ?

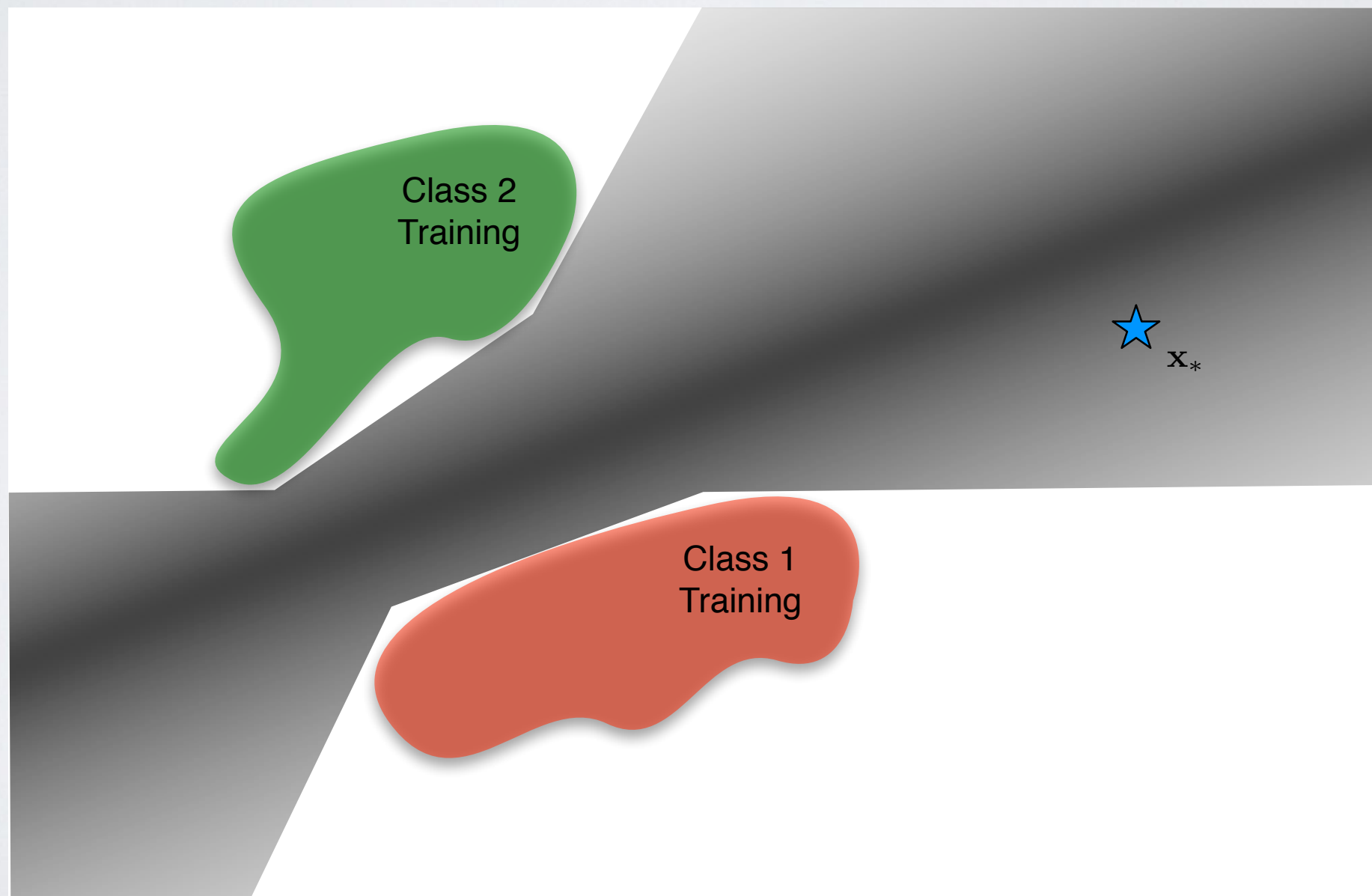
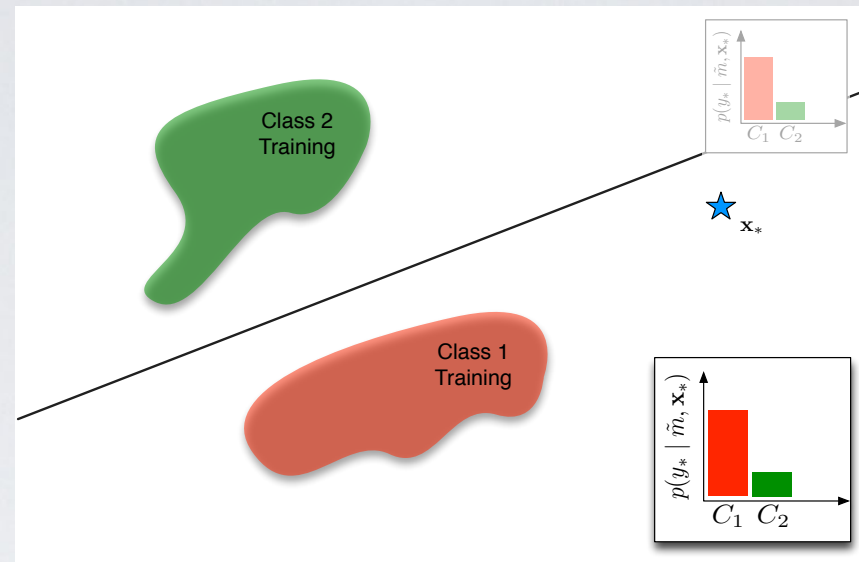


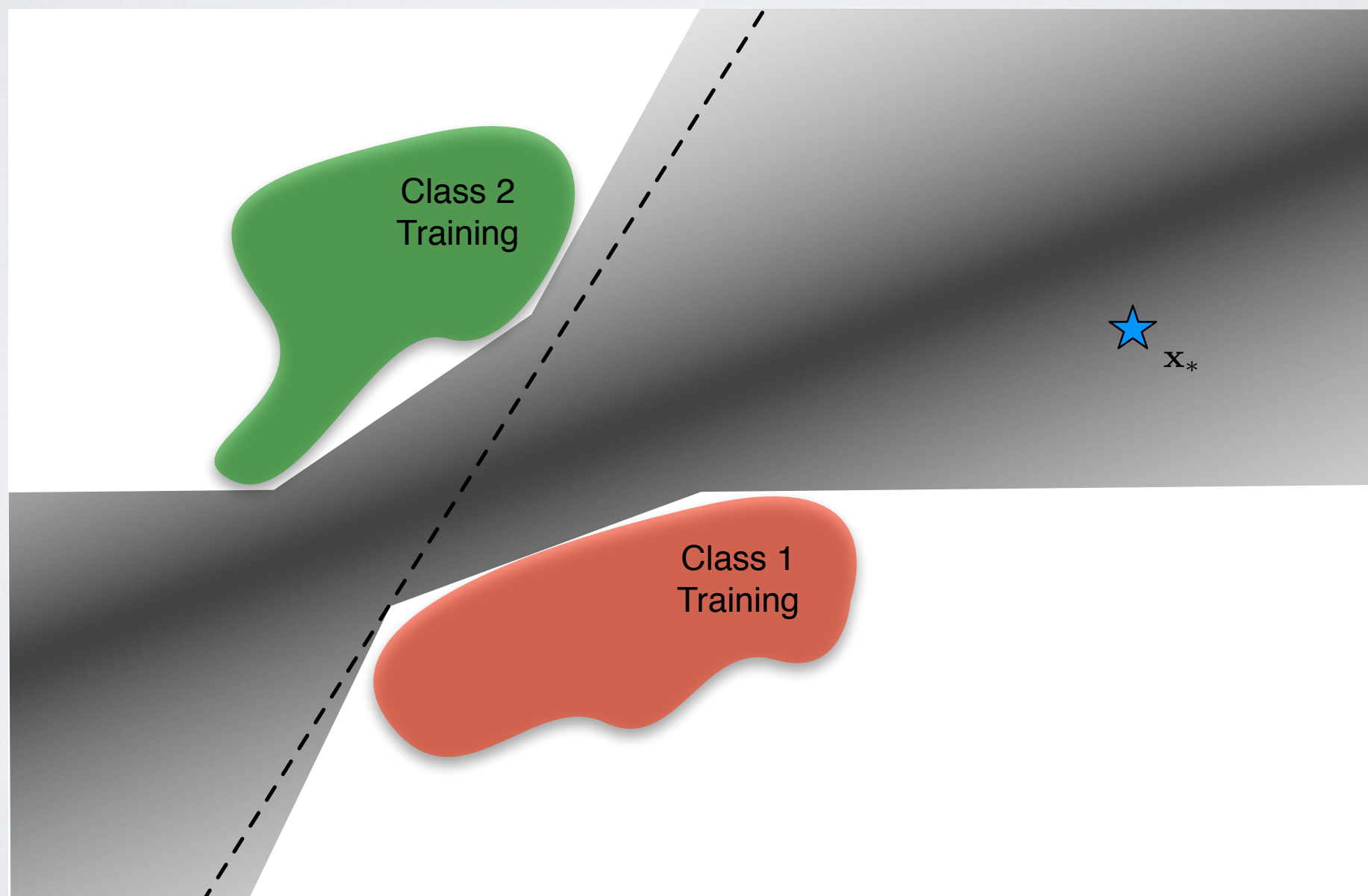
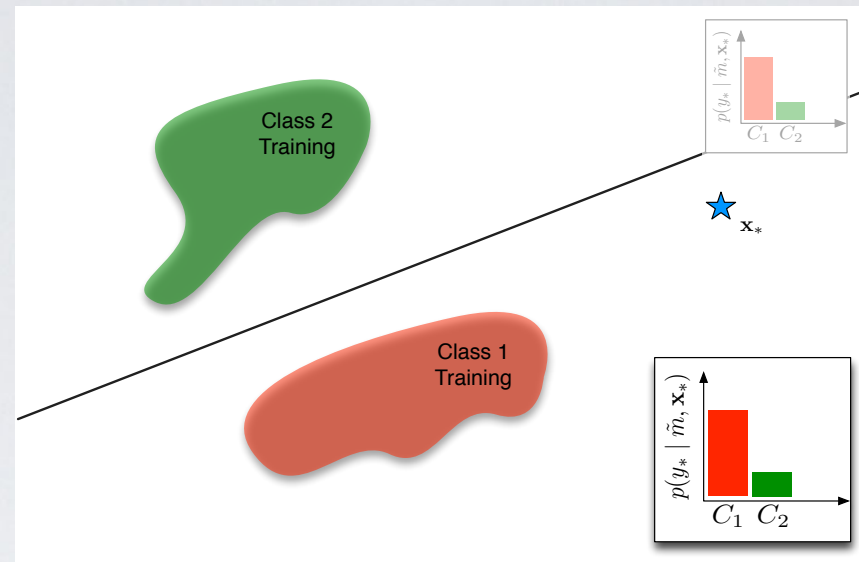


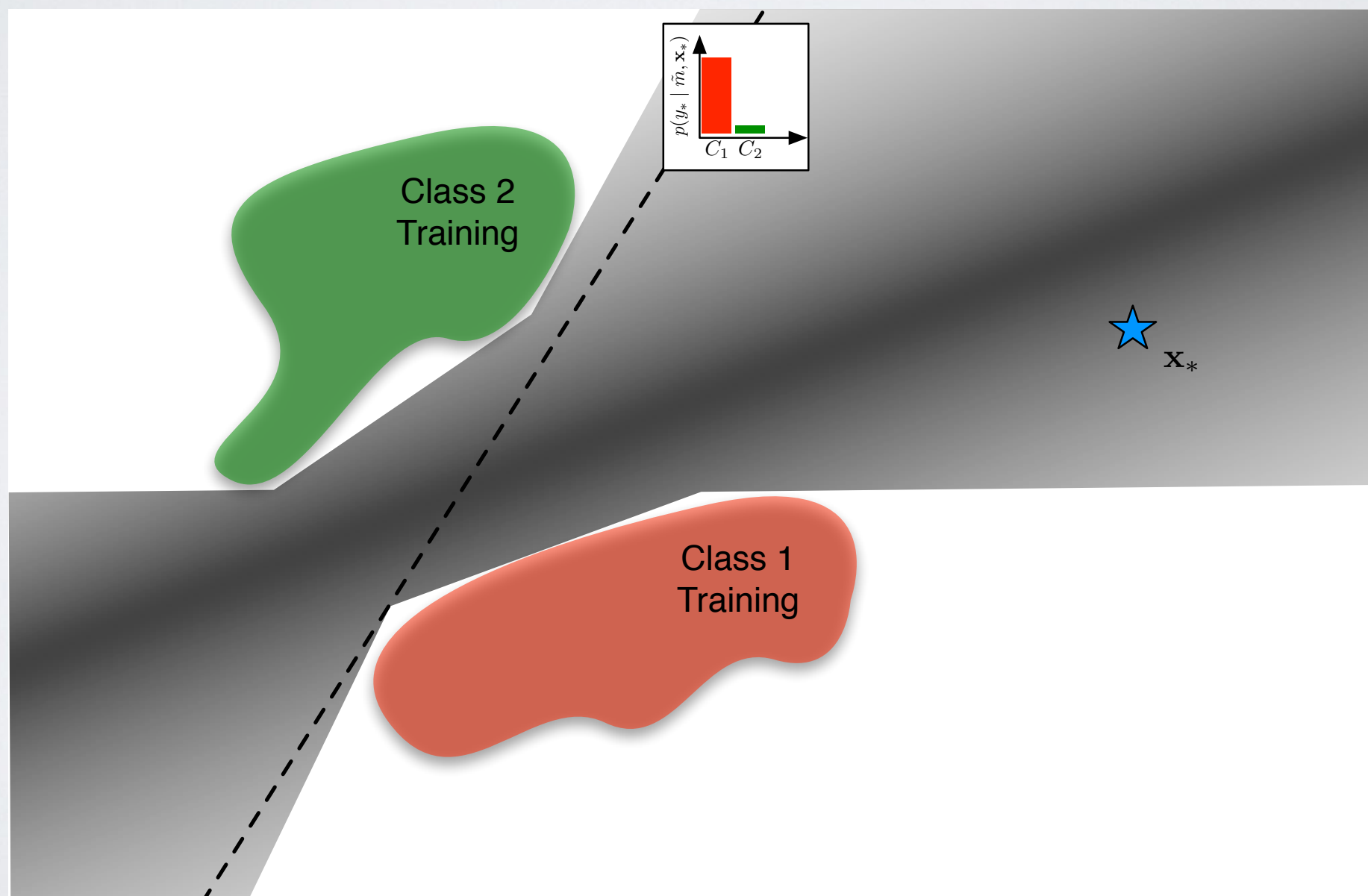
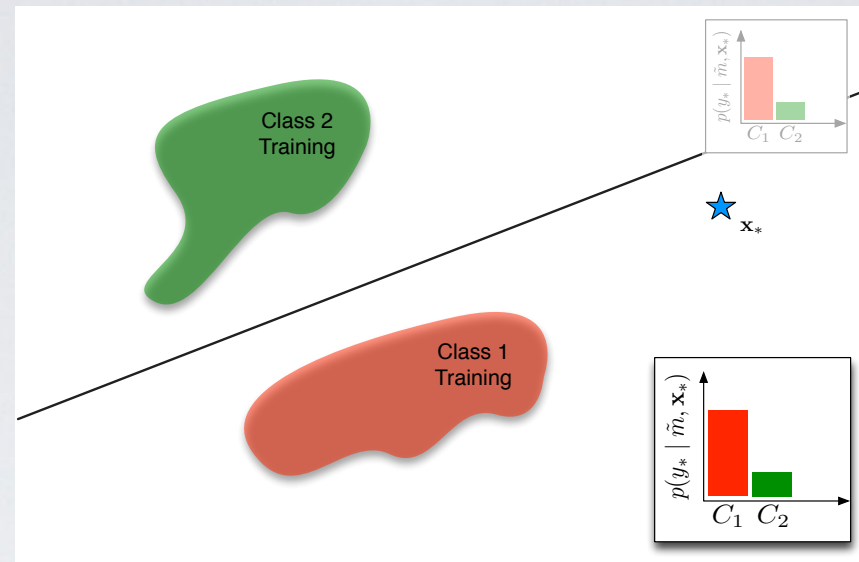


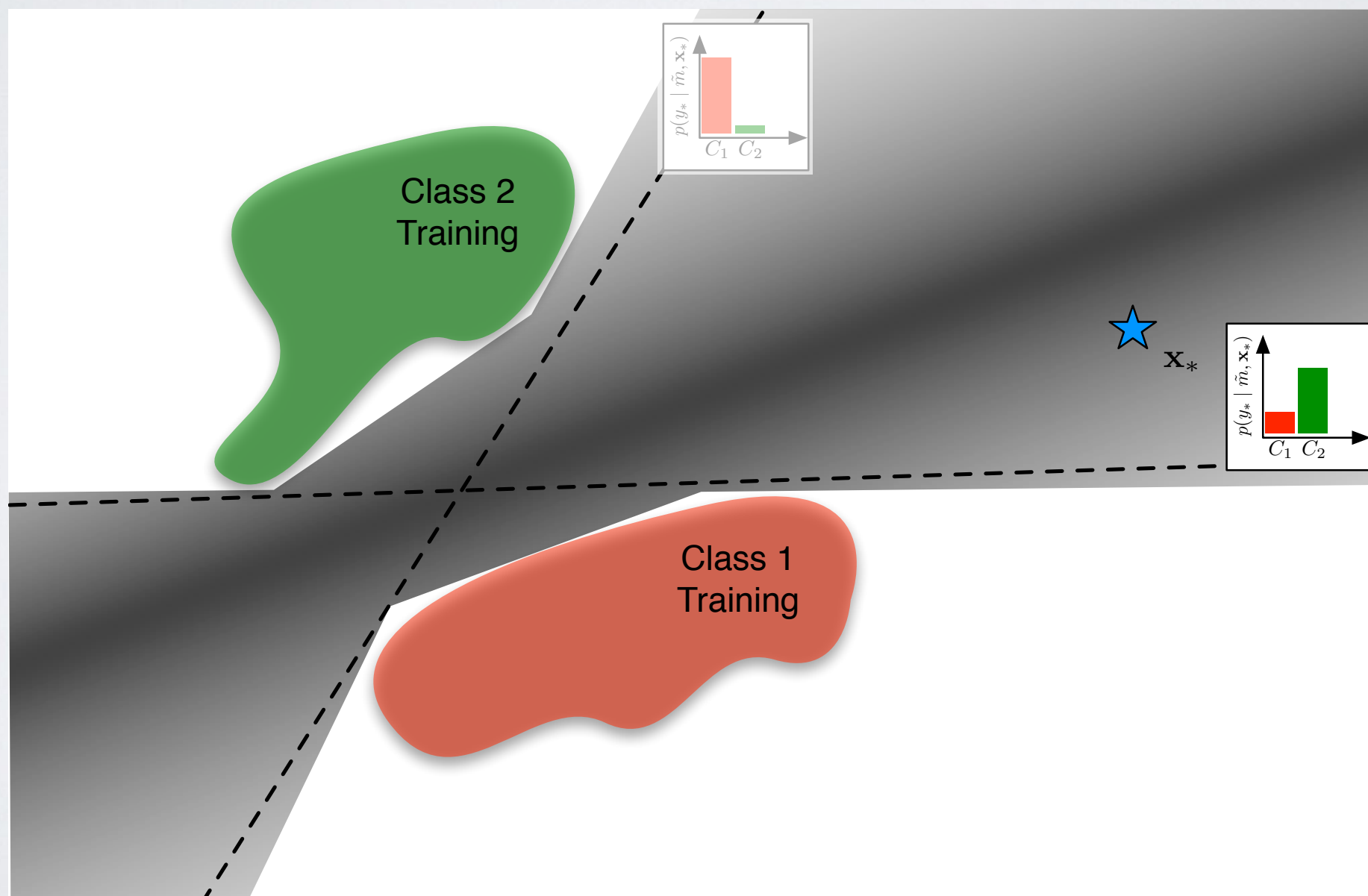
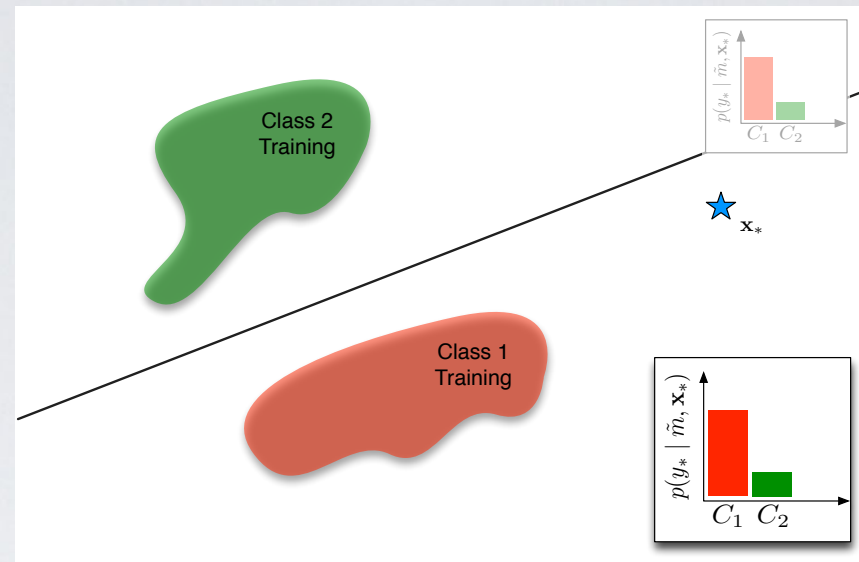


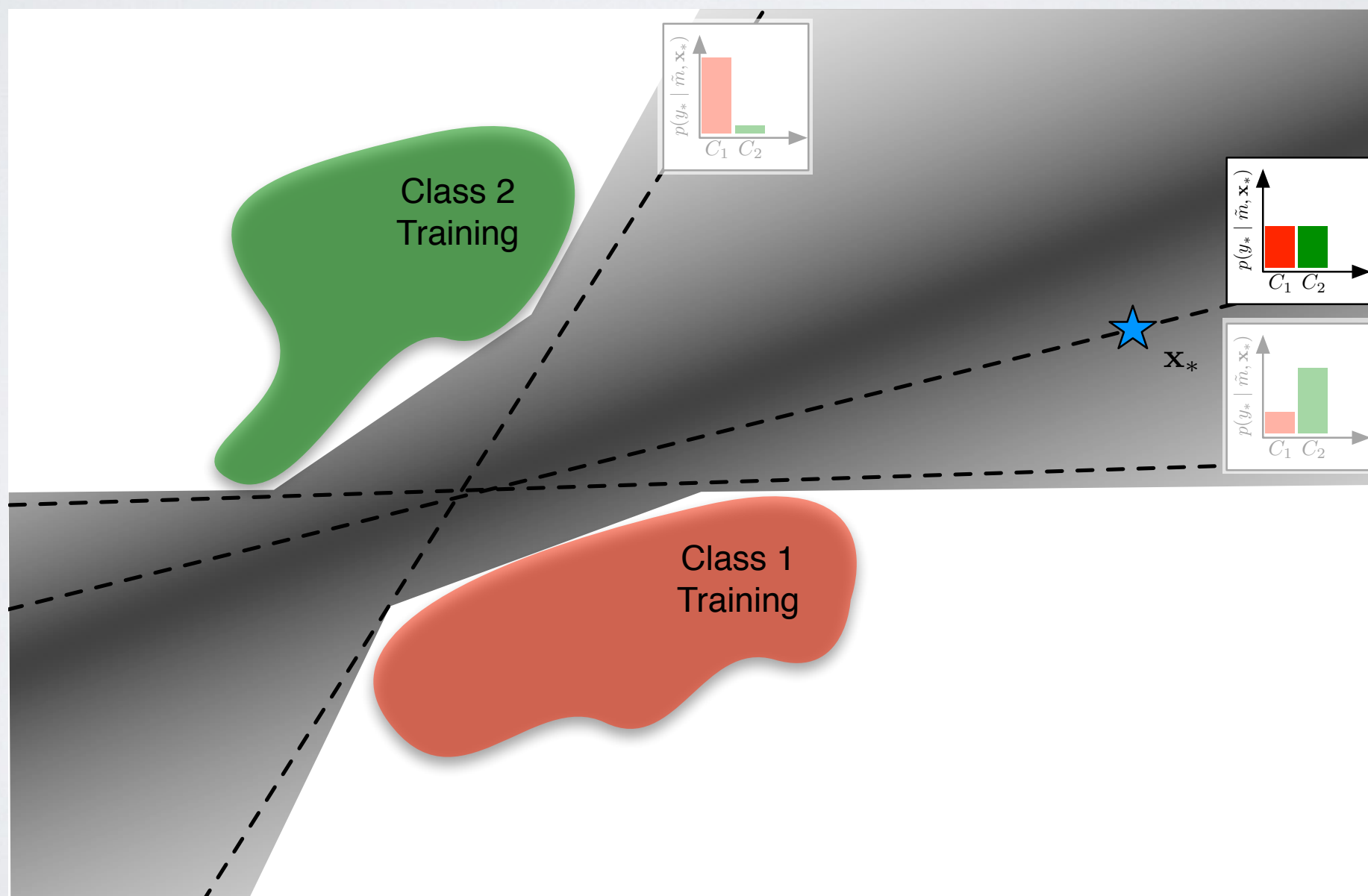
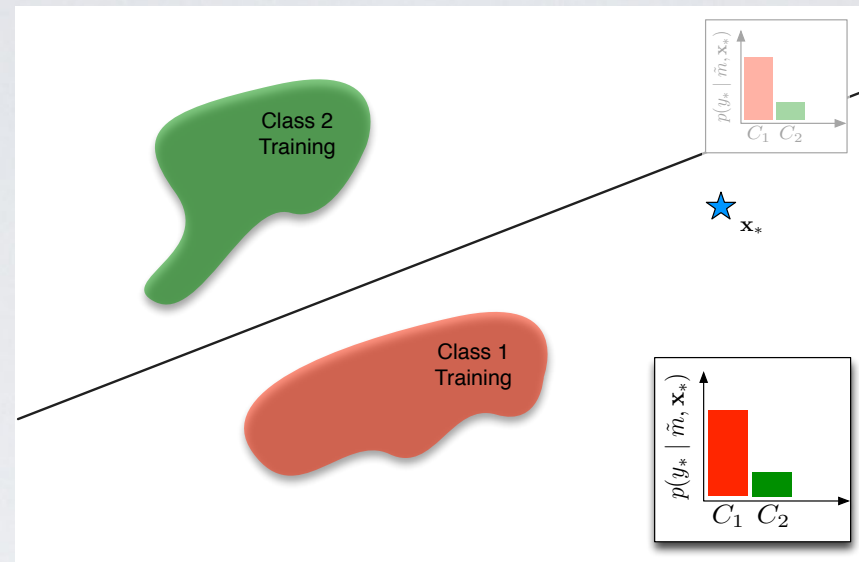


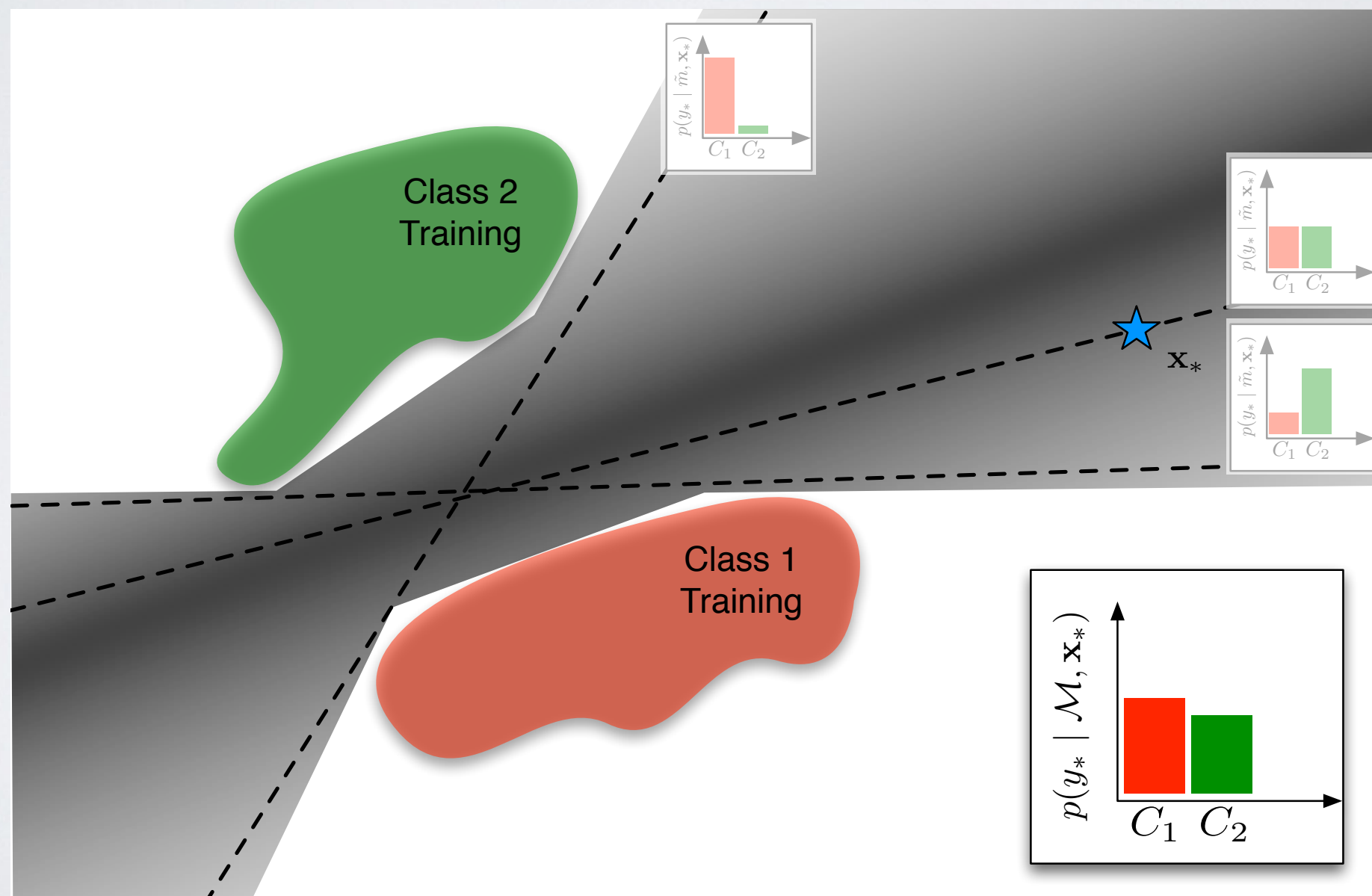
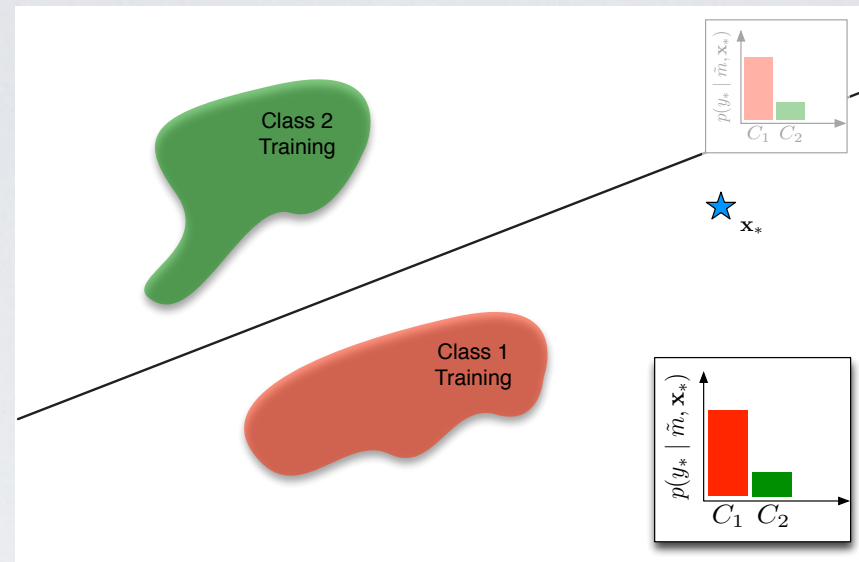


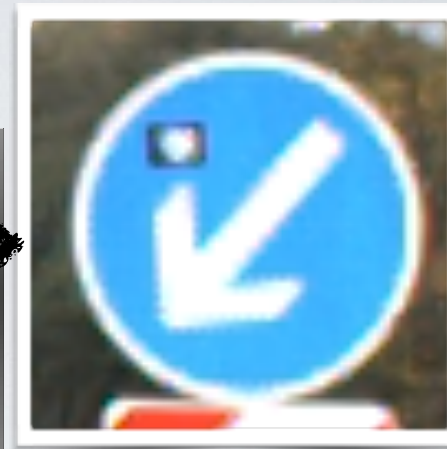
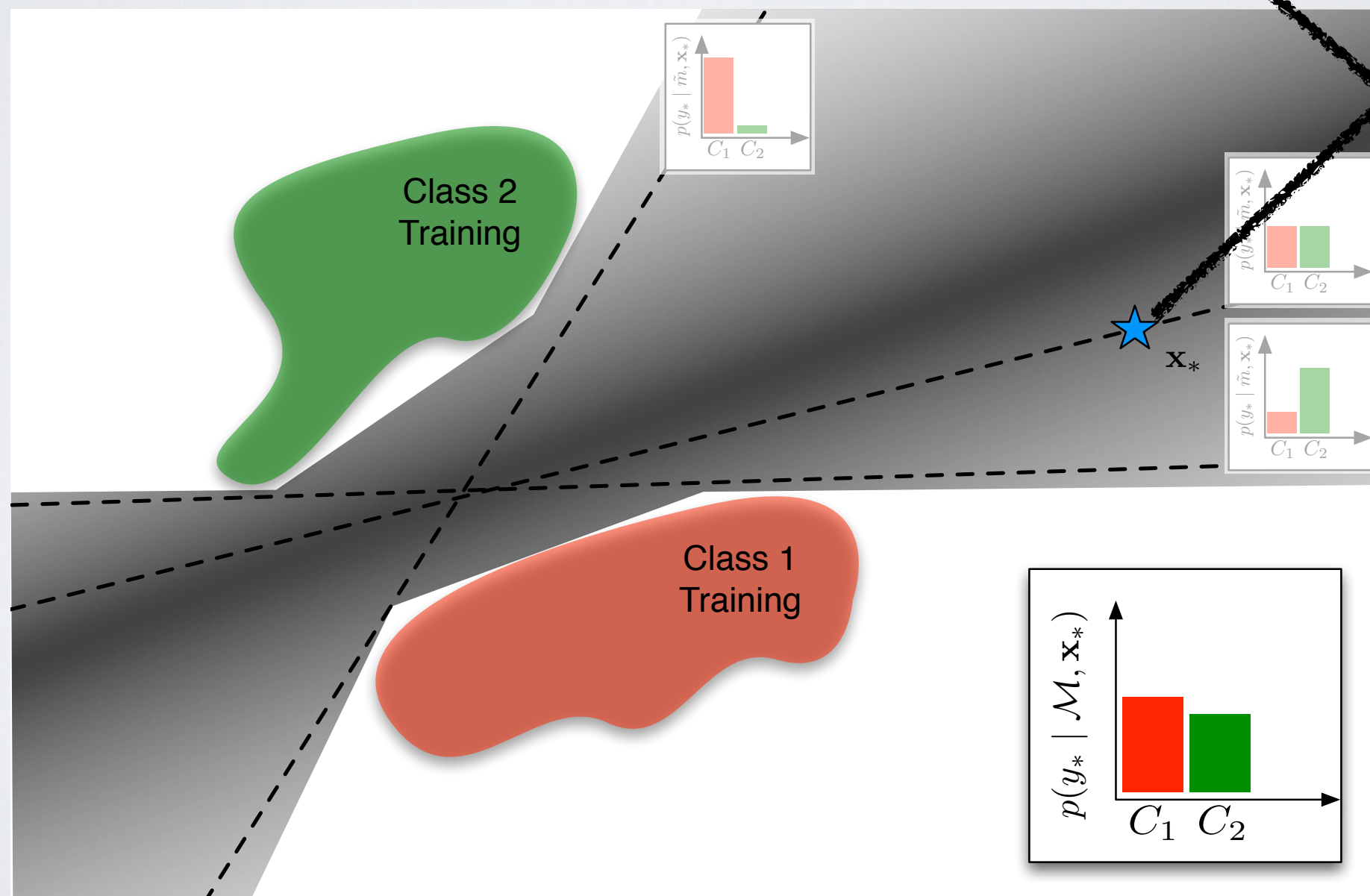
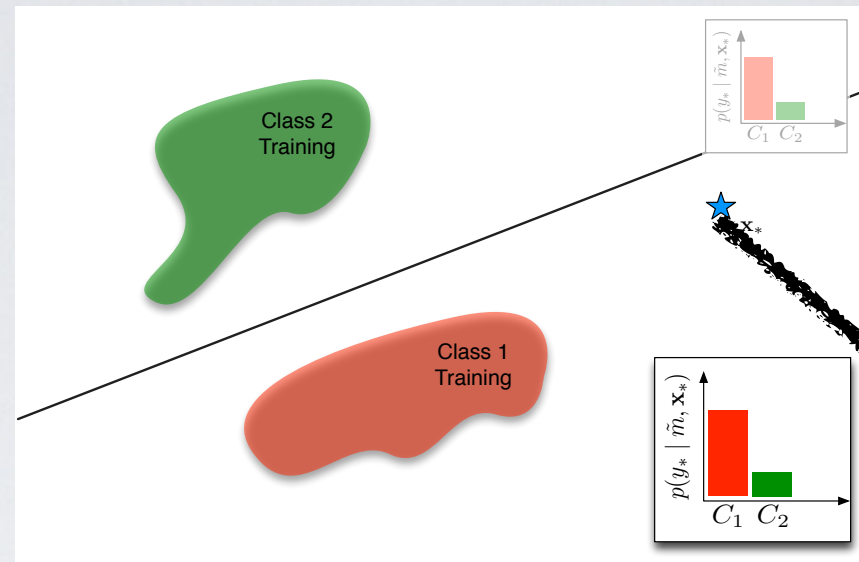








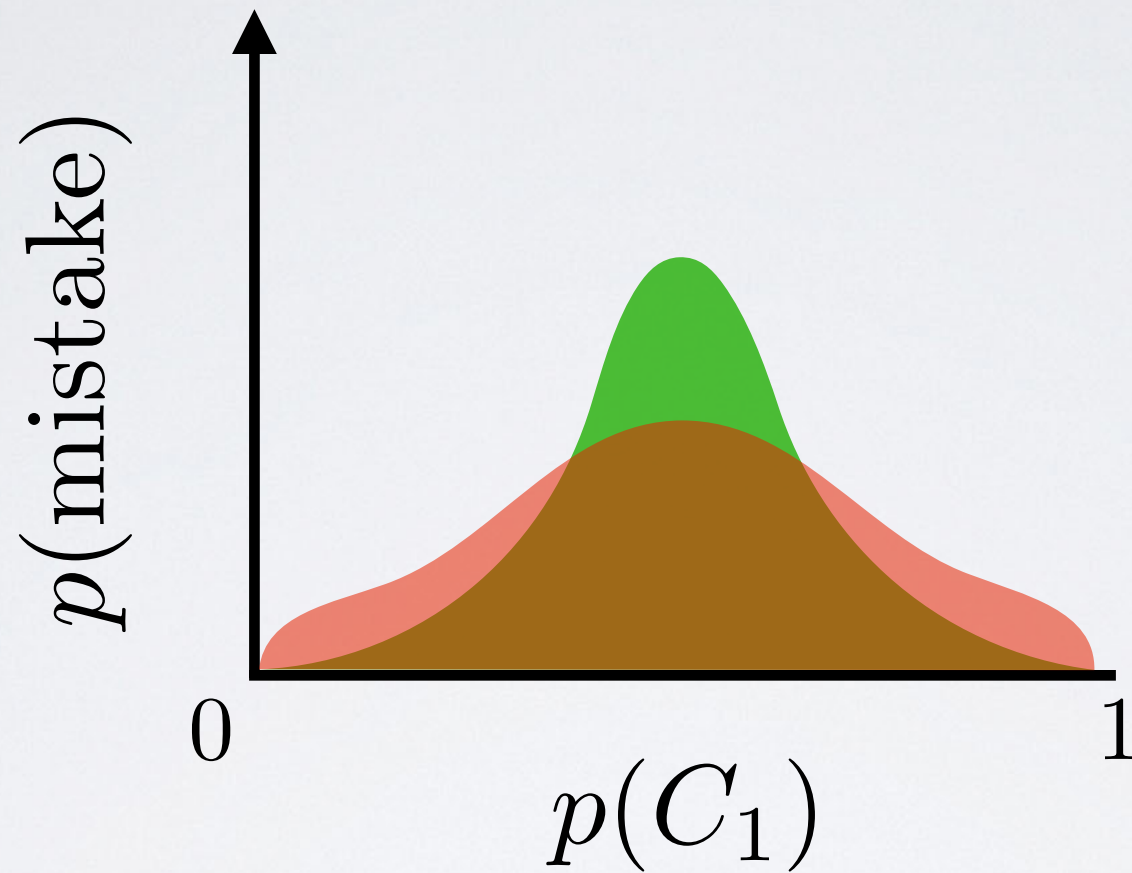




Introspection

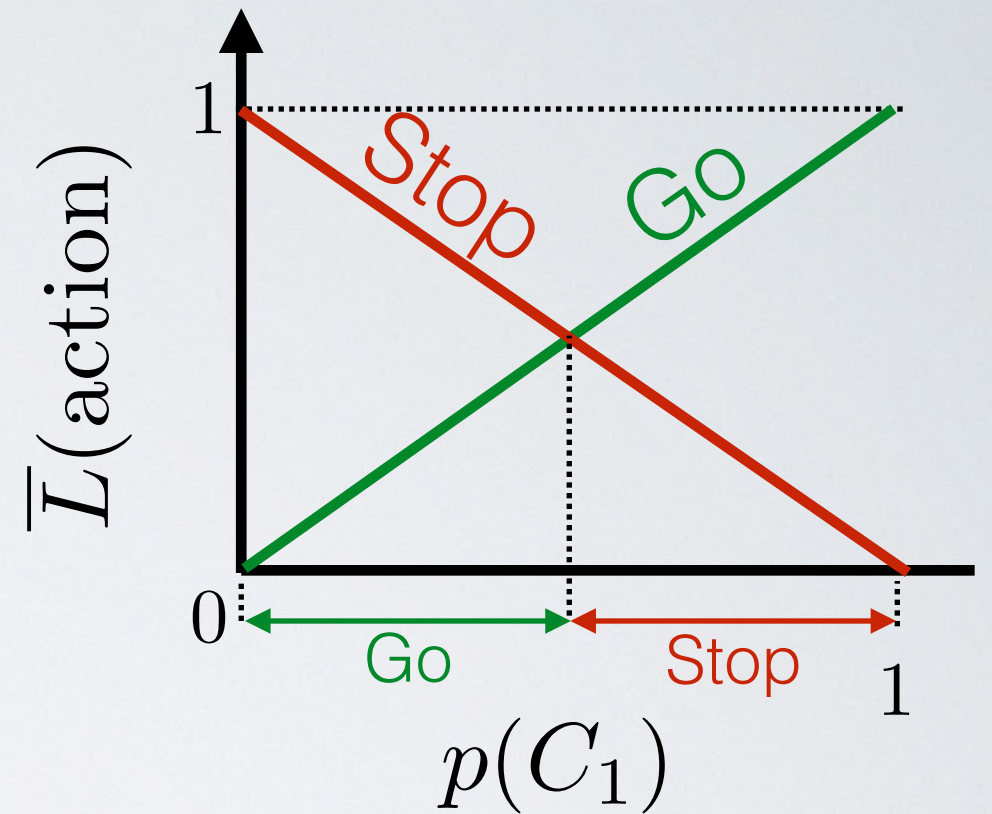
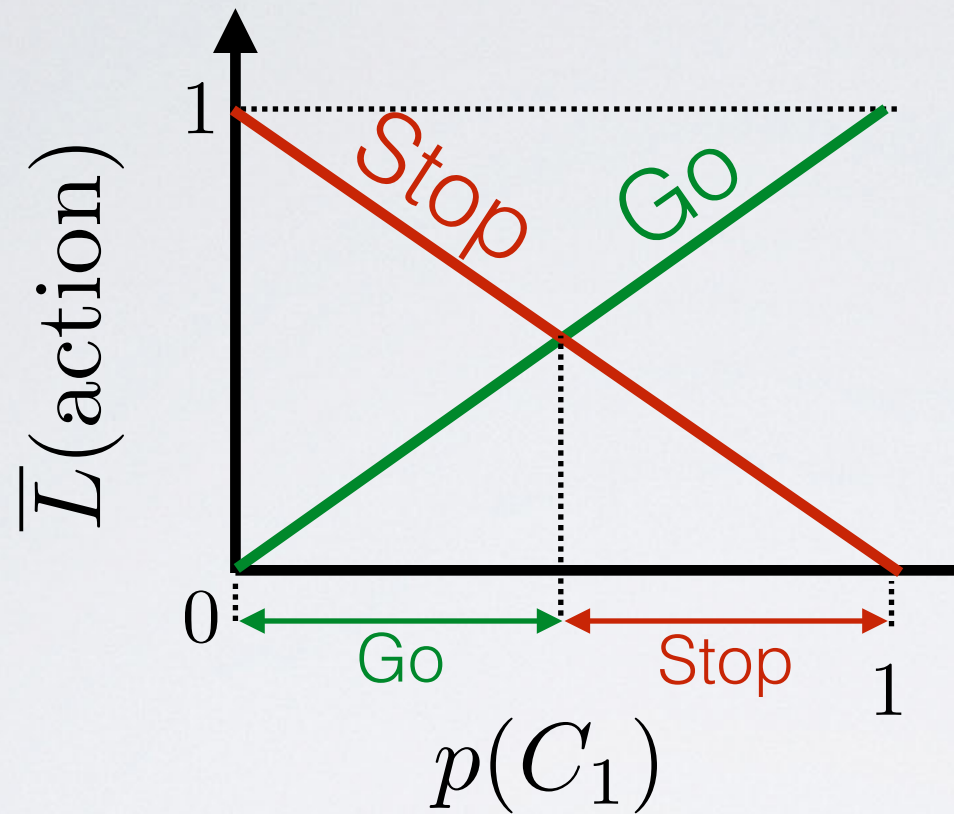
An introspective classifier is one which gives you
a **realistic** estimate of uncertainty
associated with a classification output

Introspective qualities



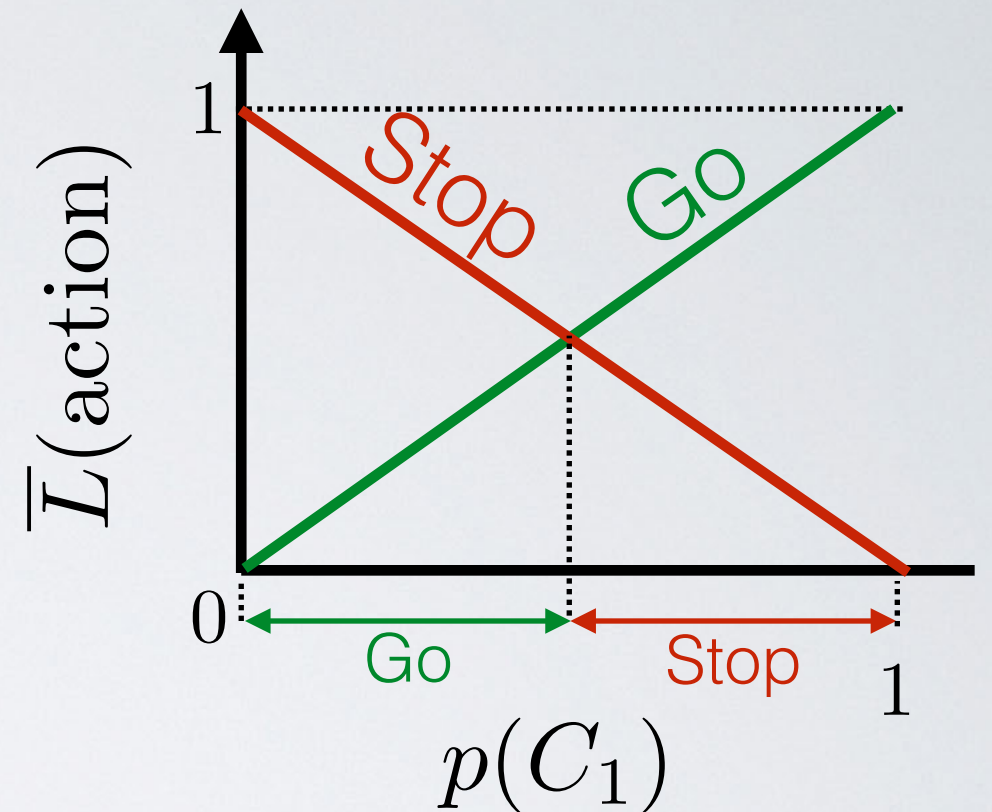
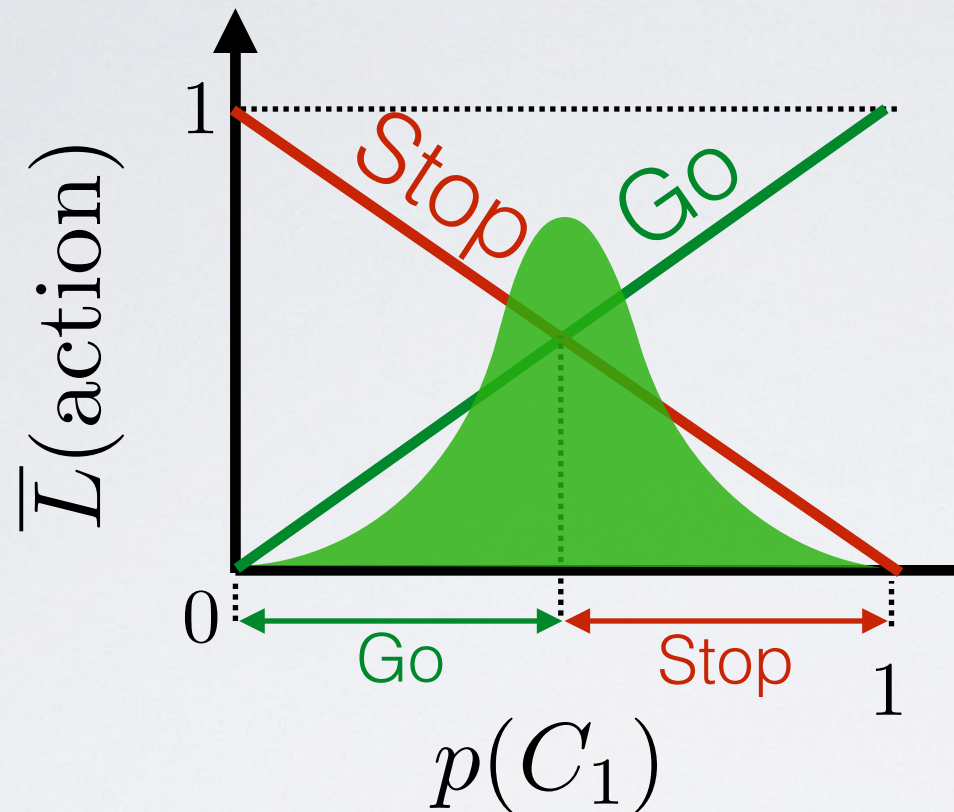
Introspective qualities

Cost	P	N
T	0	0
F	1	1



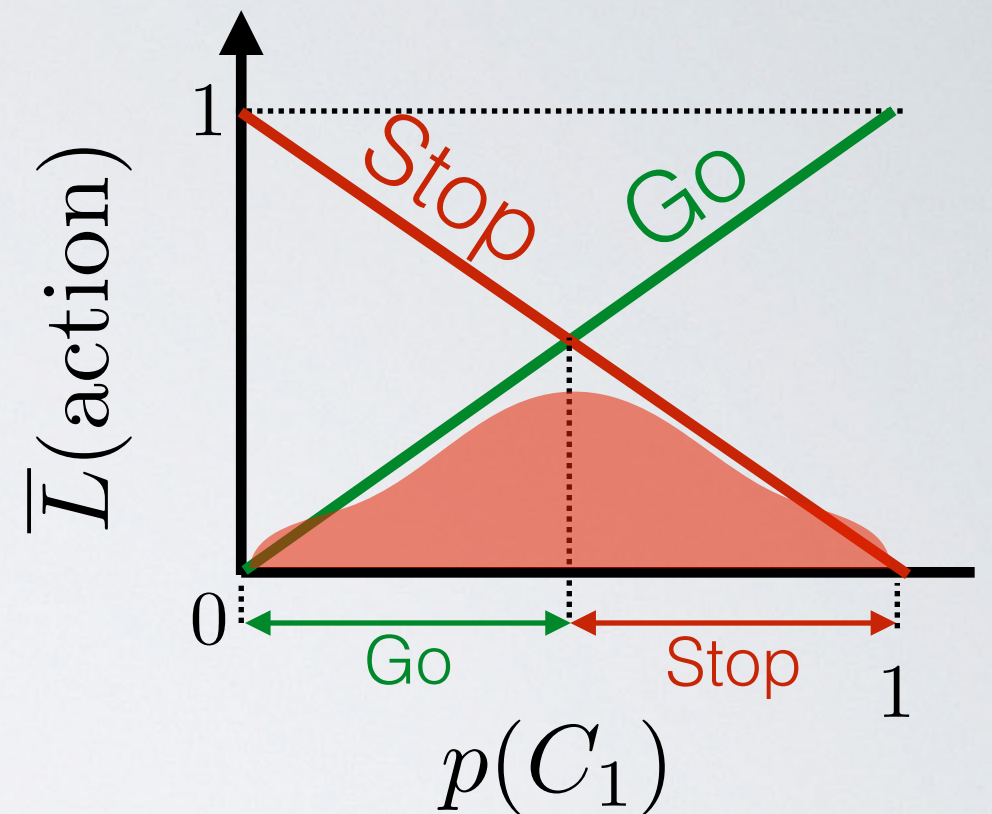
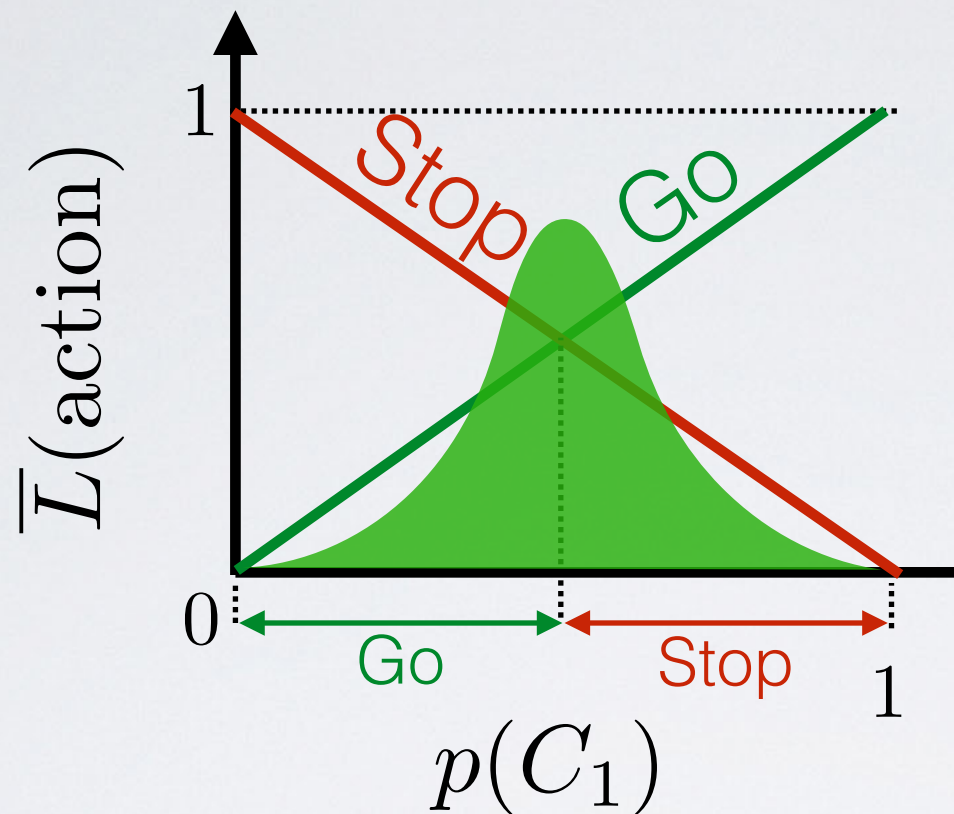
Introspective qualities

Cost	P	N
T	0	0
F	1	1



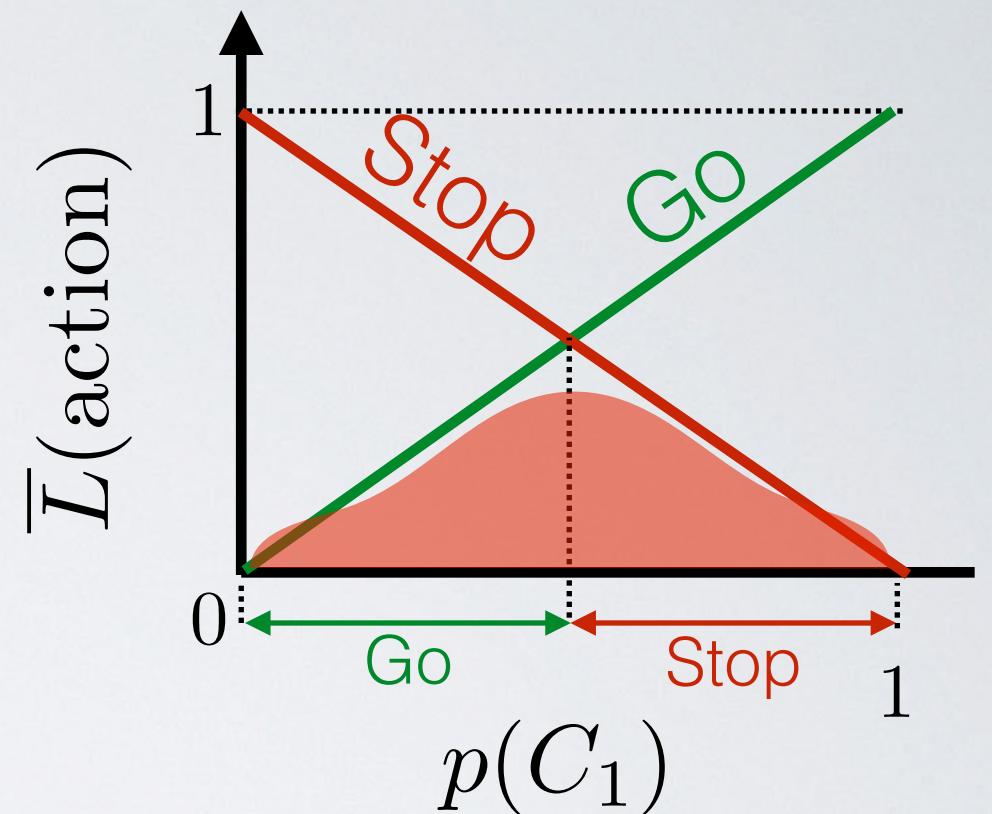
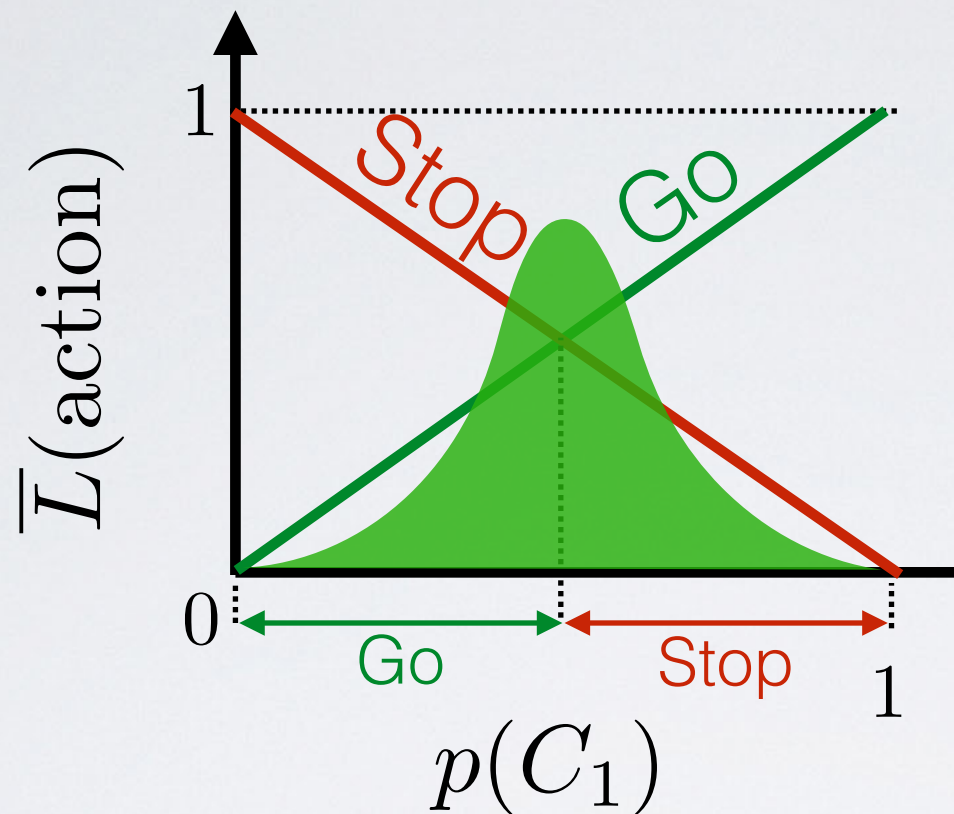
Introspective qualities

Cost	P	N
T	0	0
F	1	1

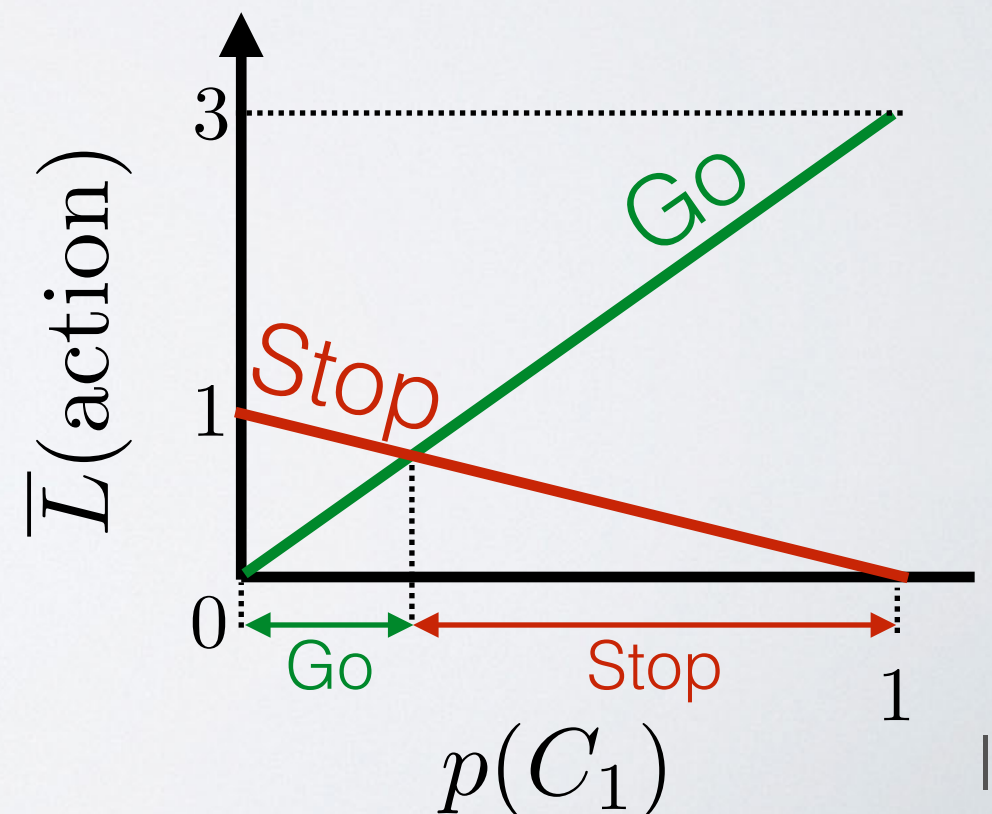
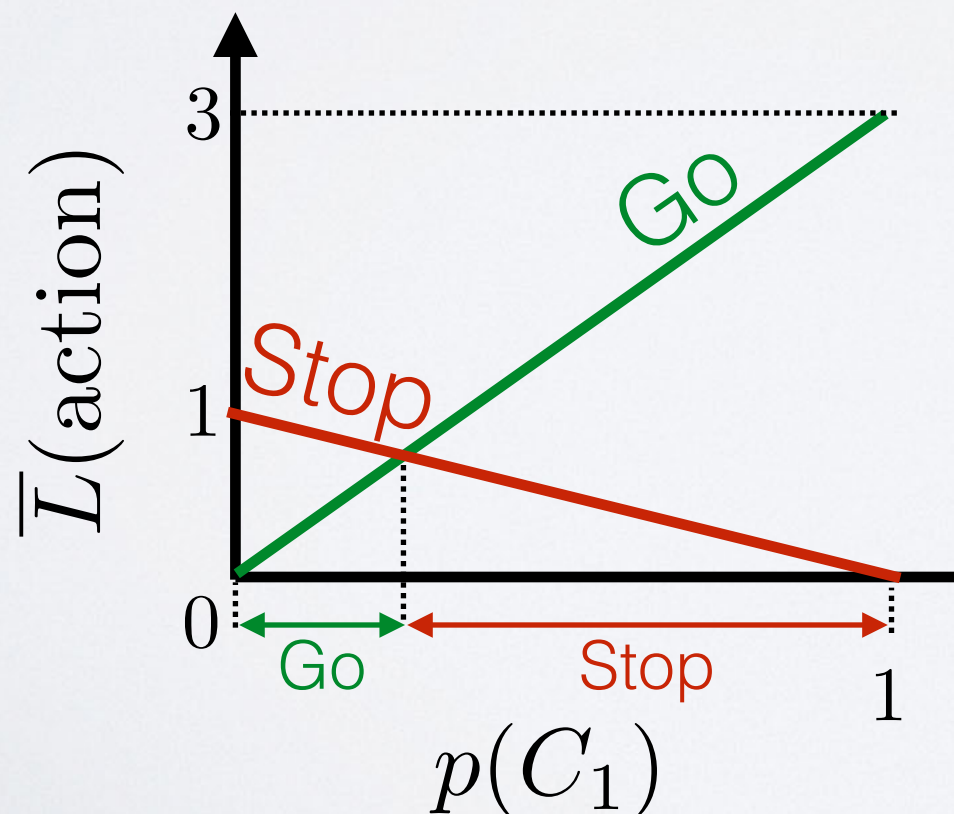


Introspective qualities

Cost	P	N
T	0	0
F	1	1

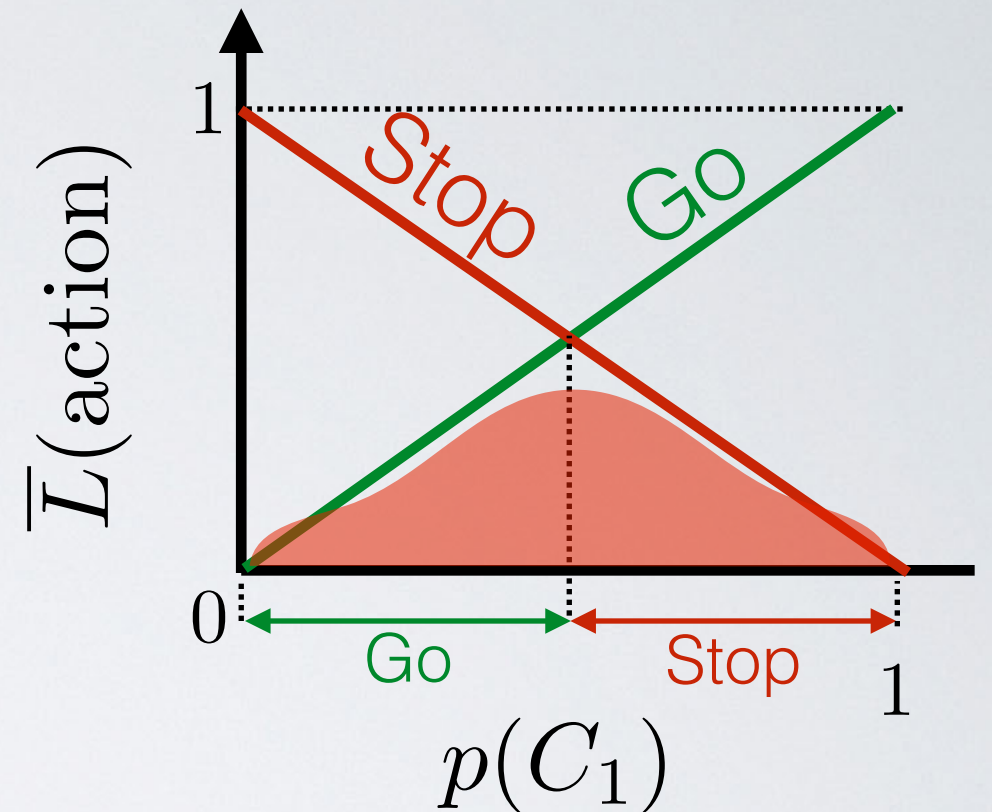
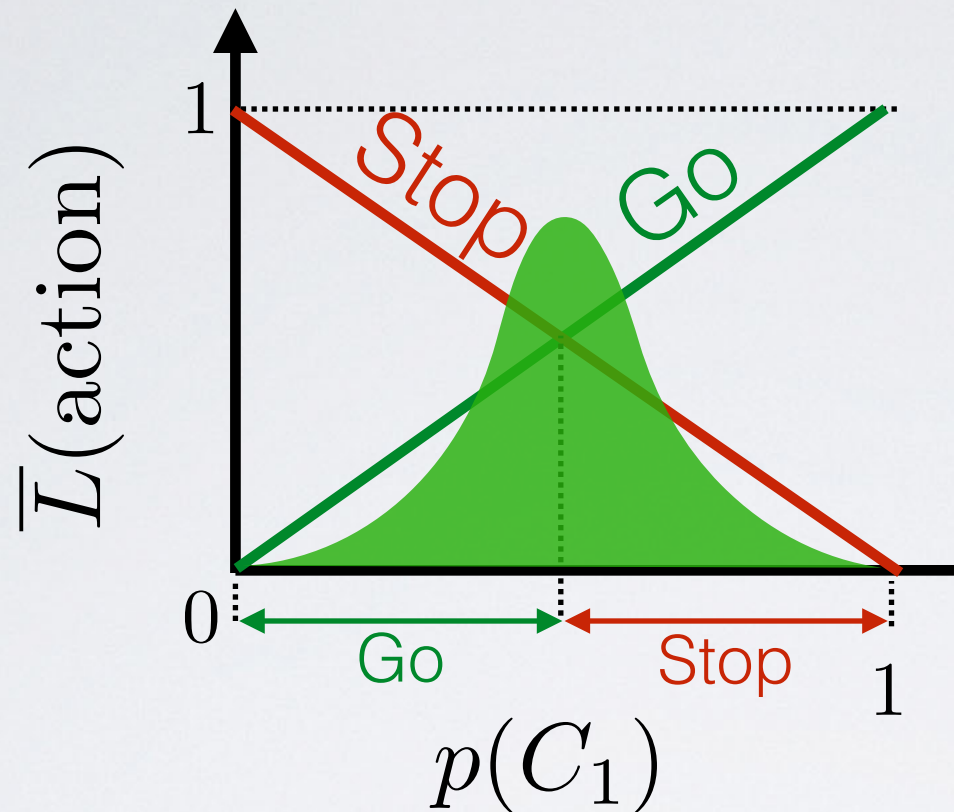


Cost	P	N
T	0	0
F	1	3

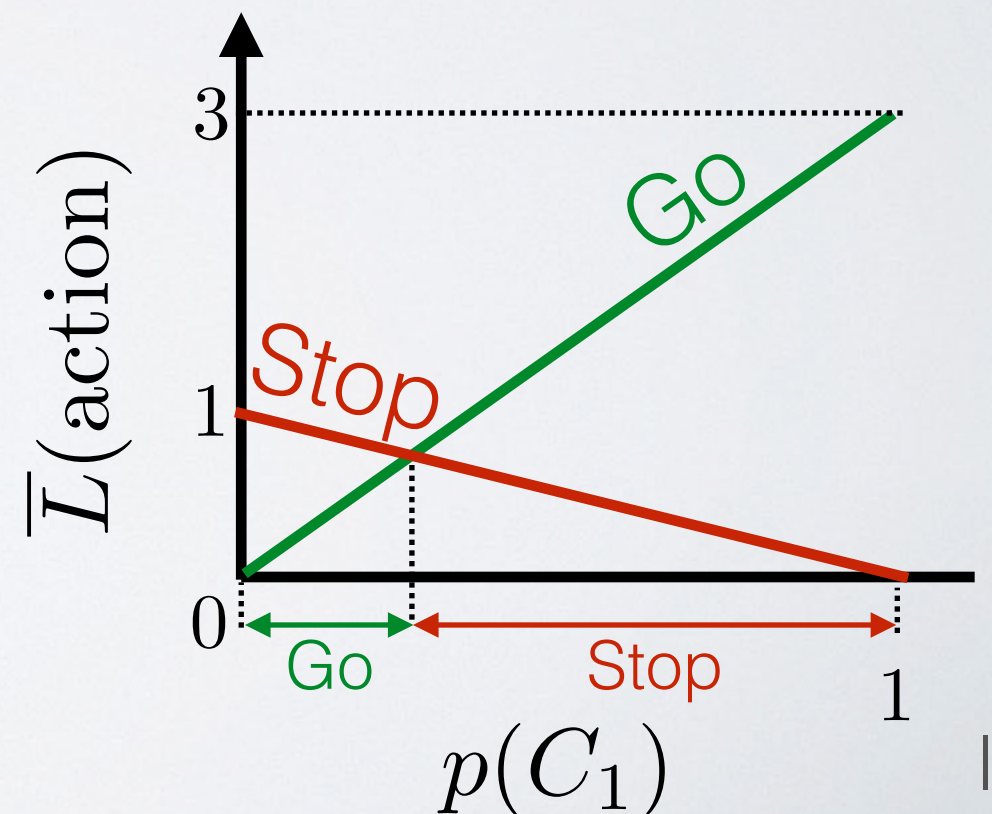
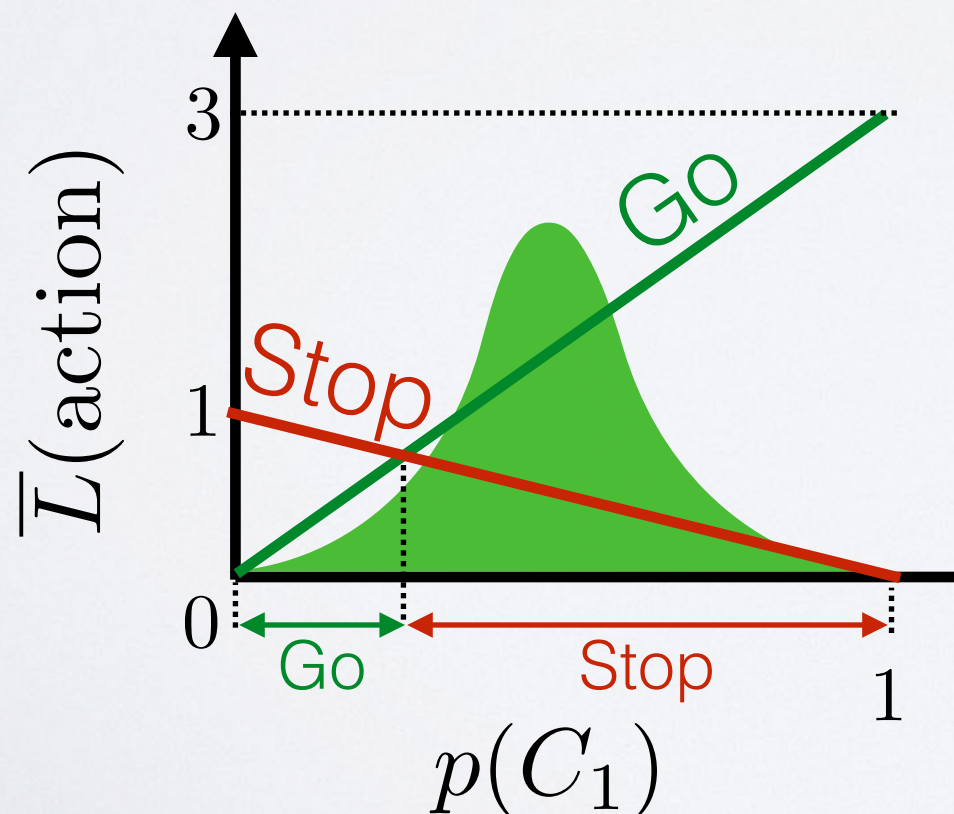


Introspective qualities

Cost	P	N
T	0	0
F	1	1

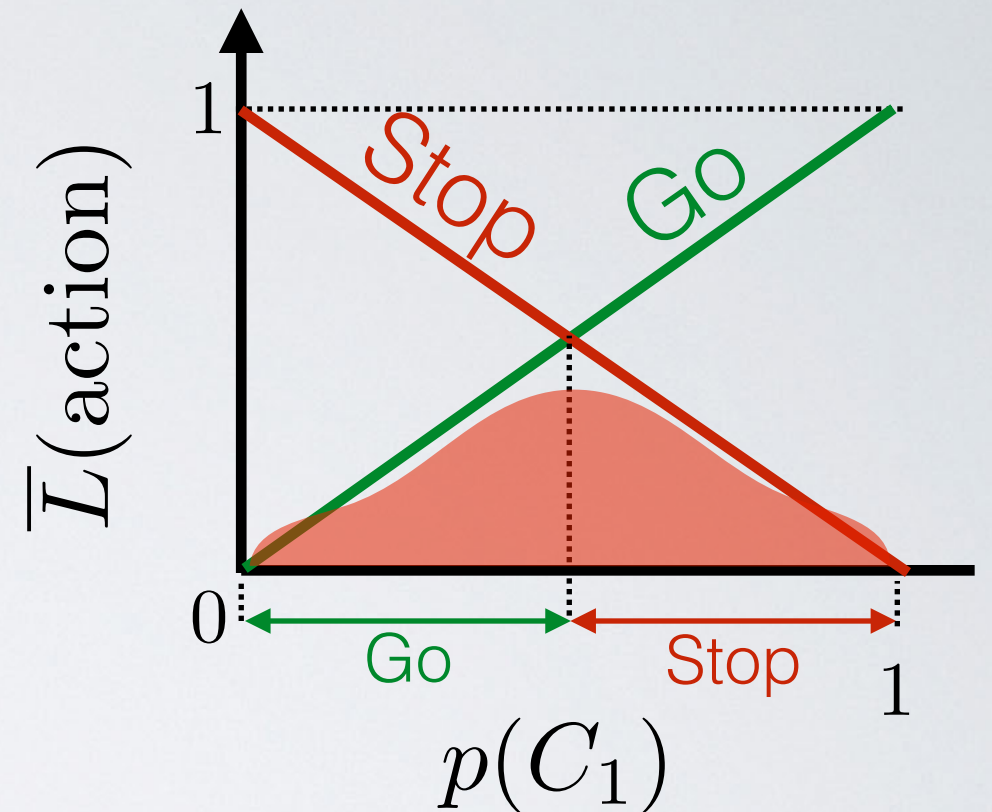
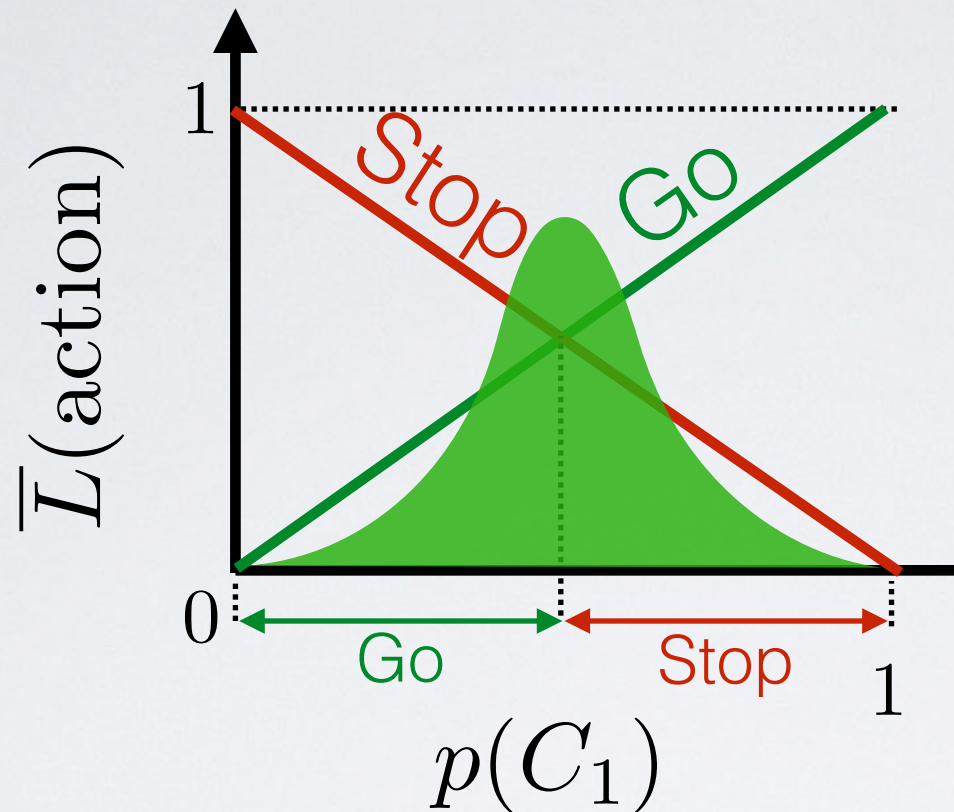


Cost	P	N
T	0	0
F	1	3

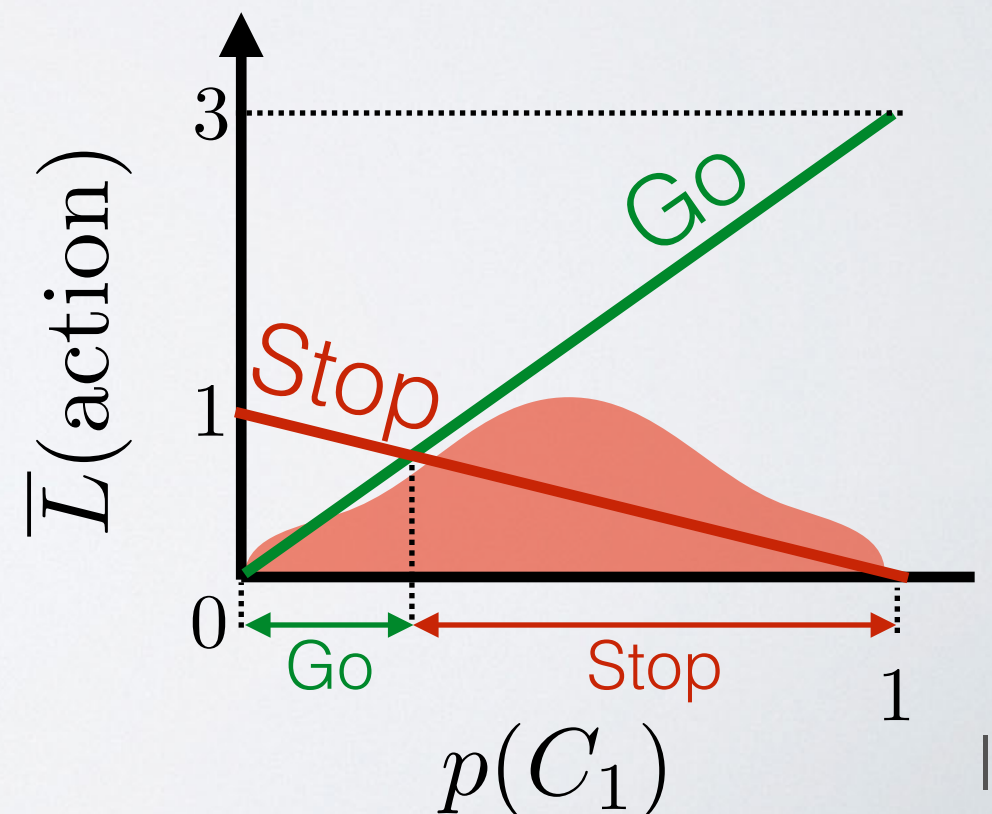
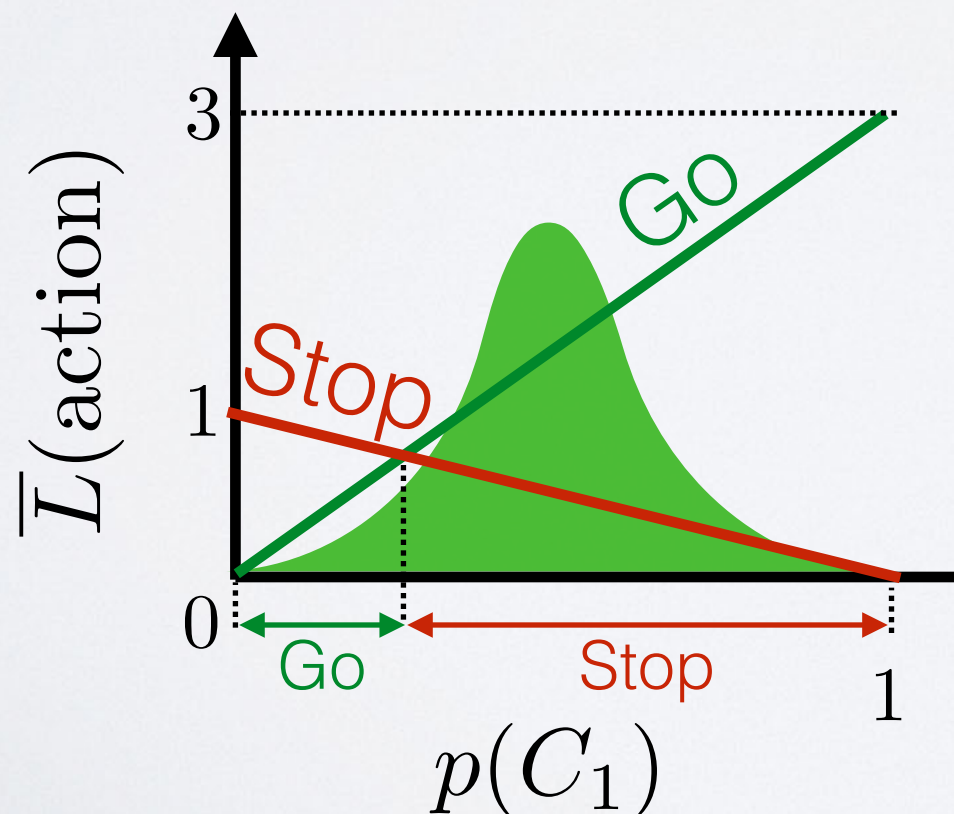


Introspective qualities

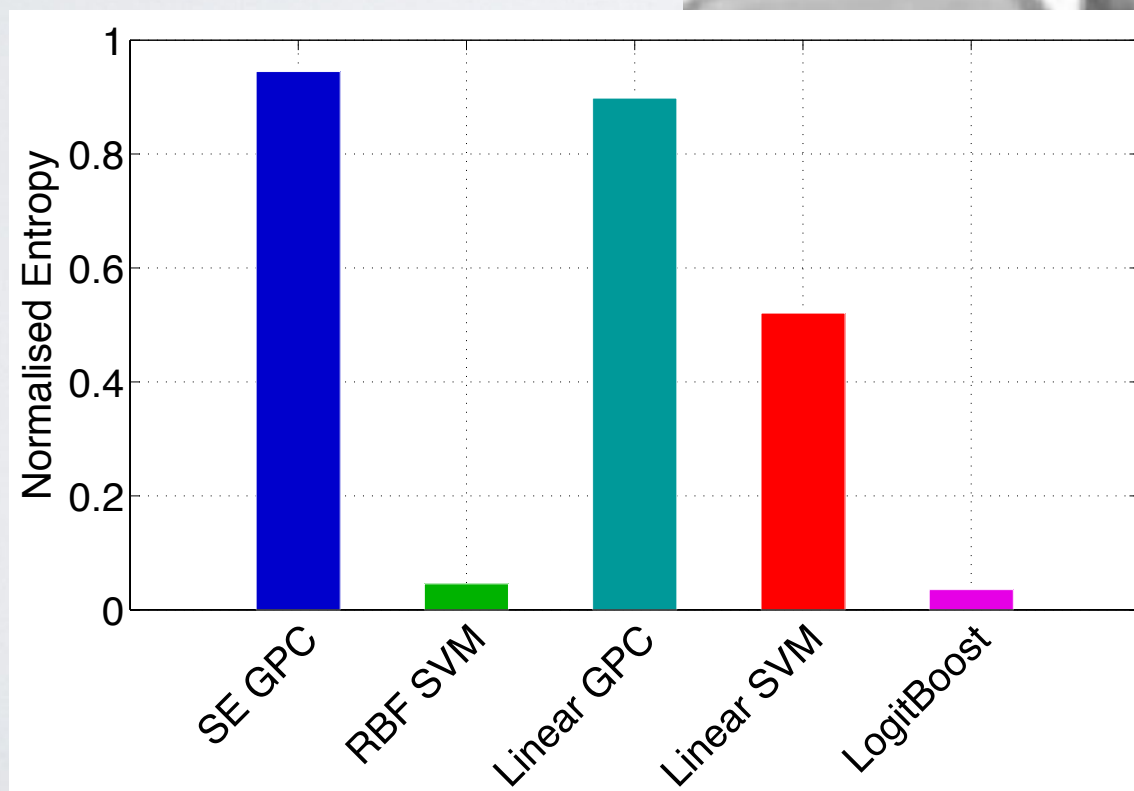
Cost	P	N
T	0	0
F	1	1



Cost	P	N
T	0	0
F	1	3







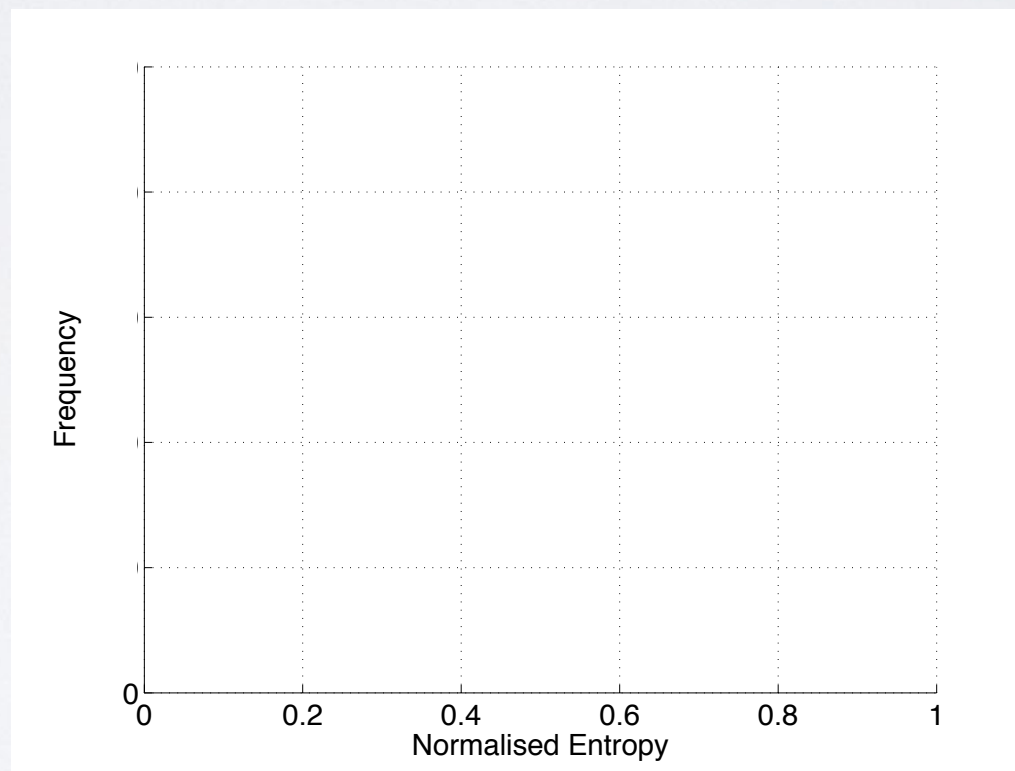
Normalised Entropy

$$H_N = - \sum_{c \in \mathcal{C}} p_c \log_{|\mathcal{C}|}(p_c)$$

a measure of *uncertainty*
normalised s.t. $H_N \in [0, 1]$

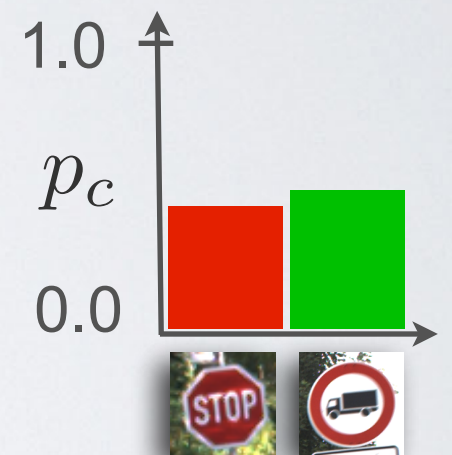
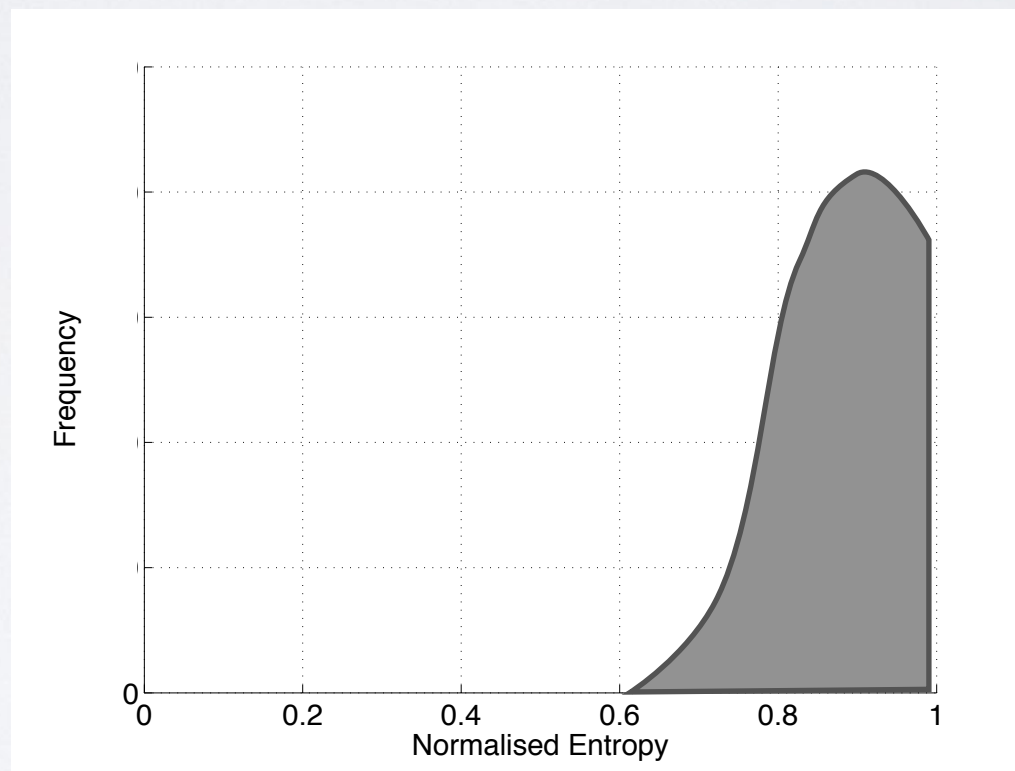
Normalised Entropy

$$H_N = - \sum_{c \in \mathcal{C}} p_c \log_{|\mathcal{C}|} (p_c) \quad \begin{array}{l} \text{a measure of } \textit{uncertainty} \\ \text{normalised s.t. } H_N \in [0, 1] \end{array}$$



Normalised Entropy

$$H_N = - \sum_{c \in \mathcal{C}} p_c \log_{|\mathcal{C}|}(p_c) \quad \text{a measure of } \textit{uncertainty} \\ \text{normalised s.t. } H_N \in [0, 1]$$

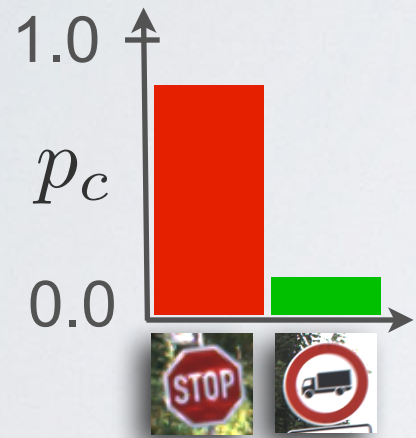


Uncertain

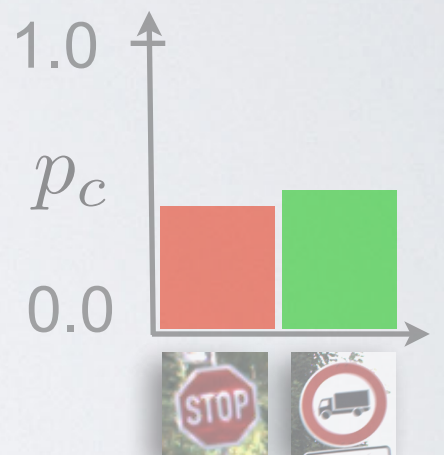
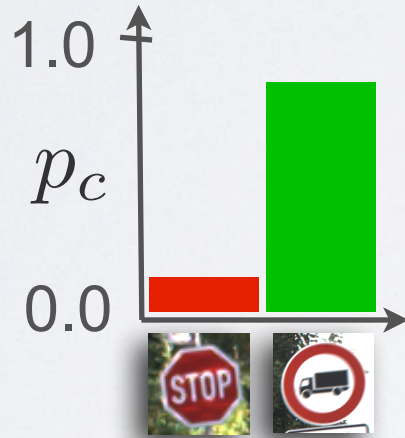
Normalised Entropy

$$H_N = - \sum_{c \in \mathcal{C}} p_c \log_{|\mathcal{C}|}(p_c)$$

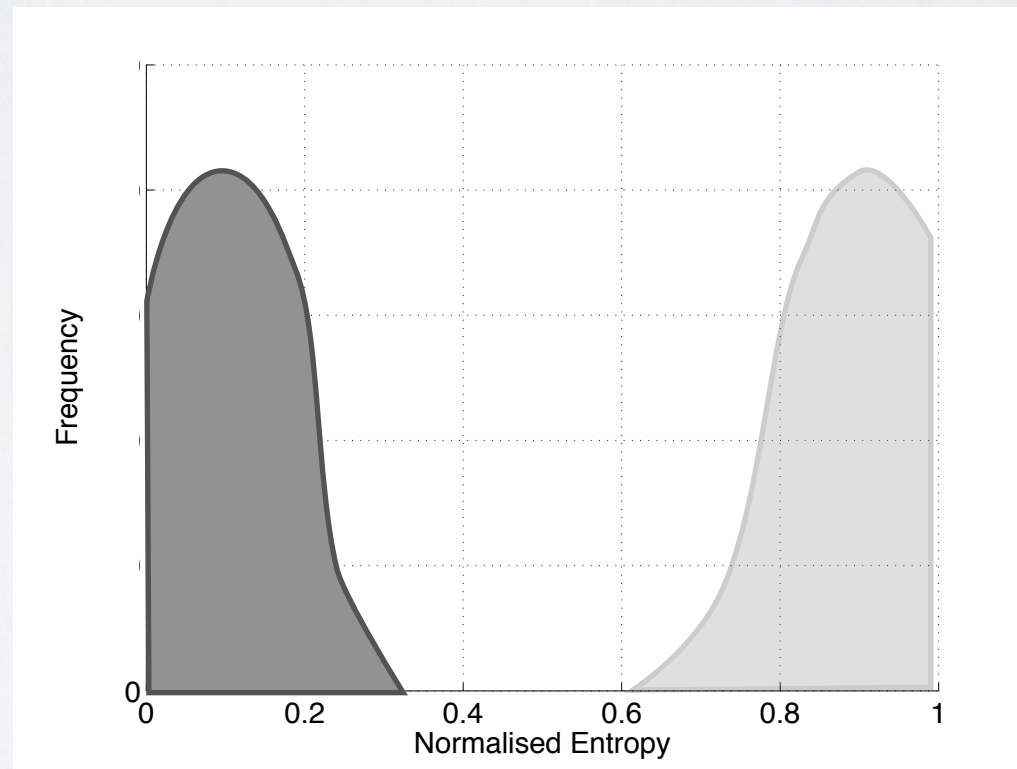
a measure of *uncertainty*
normalised s.t. $H_N \in [0, 1]$



Certain



Uncertain



Classification Of An Unseen Object Class

Classification Of An Unseen Object Class

train



against



, test on



Classification Of An Unseen Object Class

train



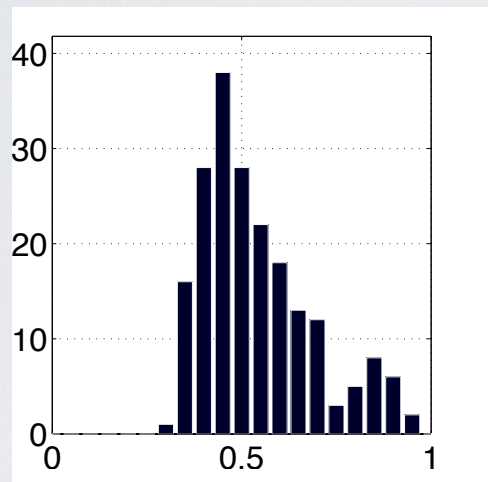
against



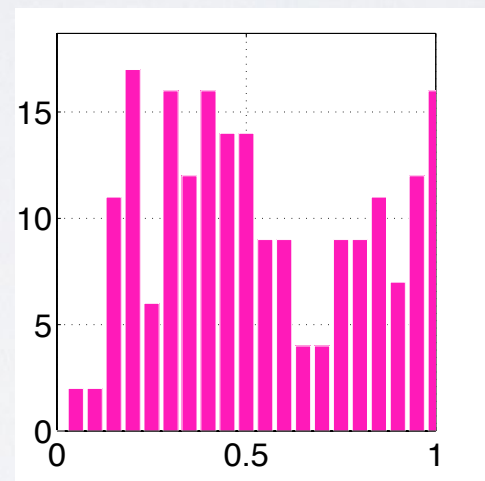
, test on



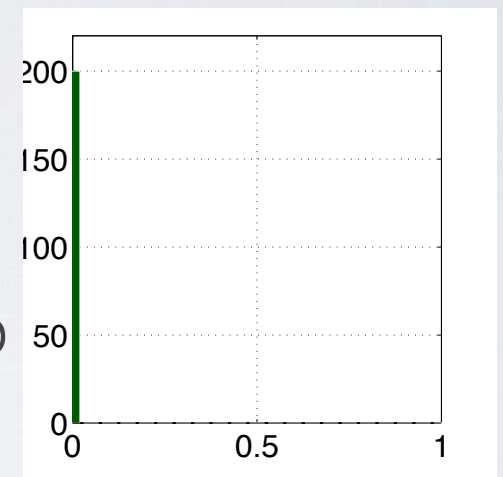
Non-linear SVM



Linear SVM



LogitBoost



Classification Of An Unseen Object Class

train



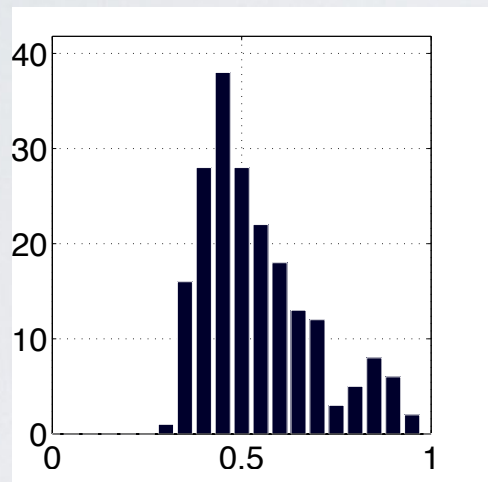
against



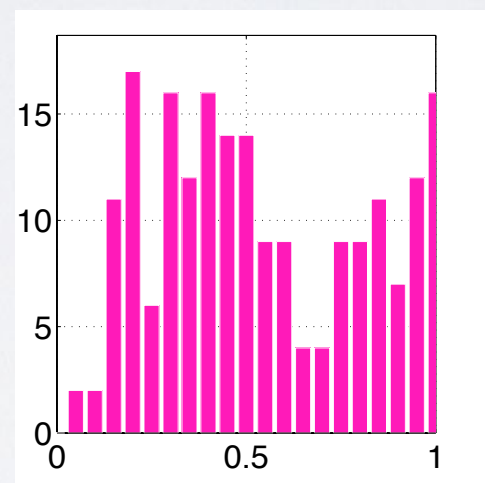
, test on



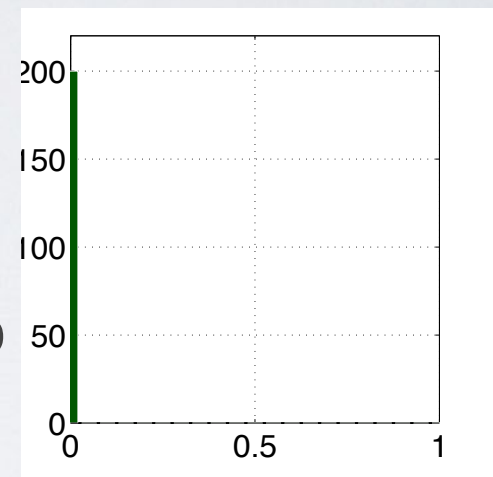
Non-linear SVM



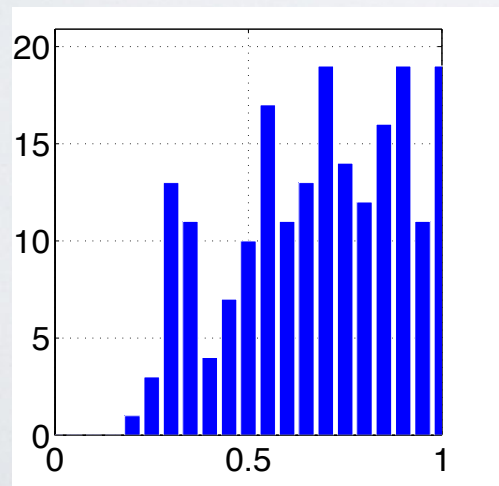
Linear SVM



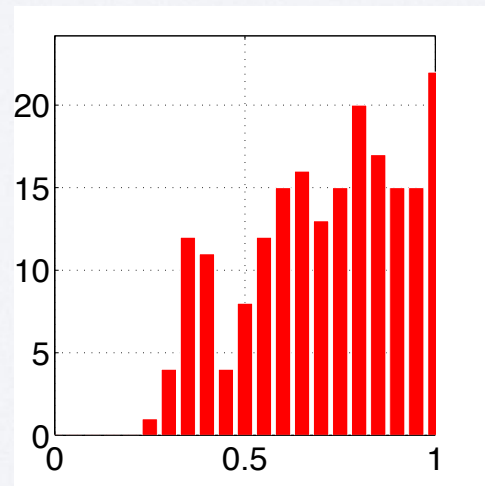
LogitBoost



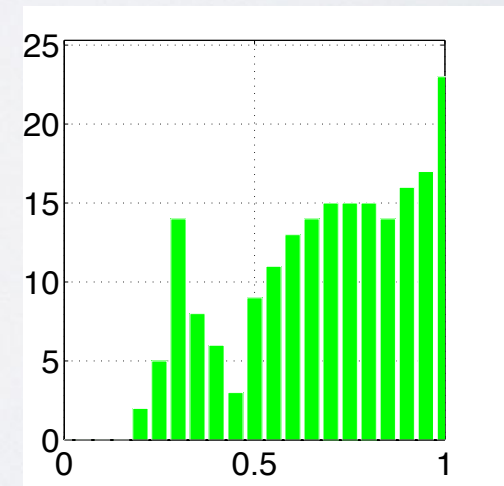
IVM



Non-linear GPC



Linear GPC



train



VS



Test Class	Classifier	Normalised Entropy	
		$\mu \pm \text{std. err.}$	$\sigma \pm \text{std. err.}$
	IVM	0.776 $\pm$ 0.081	0.145 $\pm$ 0.030
	Non-linear GPC	0.751 $\pm$ 0.087	0.152 $\pm$ 0.029
	Linear GPC	0.776 $\pm$ 0.108	0.150 $\pm$ 0.041
	Non-linear SVM	0.476 $\pm$ 0.101	0.089 $\pm$ 0.056
	Linear SVM	0.664 $\pm$ 0.122	0.250 $\pm$ 0.041
	LogitBoost	0.019 $\pm$ 0.025	0.041 $\pm$ 0.073
	IVM	0.794 $\pm$ 0.117	0.106 $\pm$ 0.026
	Non-linear GPC	0.779 $\pm$ 0.124	0.107 $\pm$ 0.024
	Linear GPC	0.777 $\pm$ 0.202	0.124 $\pm$ 0.058
	Non-linear SVM	0.537 $\pm$ 0.126	0.028 $\pm$ 0.036
	Linear SVM	0.494 $\pm$ 0.239	0.222 $\pm$ 0.049
	LogitBoost	0.016 $\pm$ 0.022	0.031 $\pm$ 0.059
	IVM	0.539 $\pm$ 0.140	0.173 $\pm$ 0.023
	Non-linear GPC	0.546 $\pm$ 0.144	0.168 $\pm$ 0.023
	Linear GPC	0.569 $\pm$ 0.166	0.177 $\pm$ 0.026
	Non-linear SVM	0.407 $\pm$ 0.077	0.076 $\pm$ 0.053
	Linear SVM	0.315 $\pm$ 0.195	0.197 $\pm$ 0.058
	LogitBoost	0.008 $\pm$ 0.004	0.012 $\pm$ 0.026
	IVM	0.579 $\pm$ 0.133	0.137 $\pm$ 0.020
	Non-linear GPC	0.577 $\pm$ 0.130	0.136 $\pm$ 0.019
	Linear GPC	0.585 $\pm$ 0.188	0.151 $\pm$ 0.029
	Non-linear SVM	0.488 $\pm$ 0.111	0.039 $\pm$ 0.034
	Linear SVM	0.177 $\pm$ 0.127	0.155 $\pm$ 0.056
	LogitBoost	0.014 $\pm$ 0.019	0.030 $\pm$ 0.056
	IVM	0.931 $\pm$ 0.025	0.080 $\pm$ 0.026
	Non-linear GPC	0.934 $\pm$ 0.021	0.079 $\pm$ 0.023
	Linear GPC	0.925 $\pm$ 0.031	0.085 $\pm$ 0.027
	Non-linear SVM	0.641 $\pm$ 0.168	0.100 $\pm$ 0.047
	Linear SVM	0.705 $\pm$ 0.142	0.212 $\pm$ 0.049
	LogitBoost	0.059 $\pm$ 0.103	0.077 $\pm$ 0.127

train



VS



Test Class	Classifier	Normalised Entropy	
		$\mu \pm \text{std. err.}$	$\sigma \pm \text{std. err.}$
	IVM	0.776 $\pm$ 0.081	0.145 $\pm$ 0.030
	Non-linear GPC	0.751 $\pm$ 0.087	0.152 $\pm$ 0.029
	Linear GPC	0.776 $\pm$ 0.108	0.150 $\pm$ 0.041
	Non-linear SVM	0.476 $\pm$ 0.101	0.089 $\pm$ 0.056
	Linear SVM	0.664 $\pm$ 0.122	0.250 $\pm$ 0.041
	LogitBoost	0.019 $\pm$ 0.025	0.041 $\pm$ 0.073
	IVM	0.794 $\pm$ 0.117	0.106 $\pm$ 0.026
	Non-linear GPC	0.779 $\pm$ 0.124	0.107 $\pm$ 0.024
	Linear GPC	0.777 $\pm$ 0.202	0.124 $\pm$ 0.058
	Non-linear SVM	0.537 $\pm$ 0.126	0.028 $\pm$ 0.036
	Linear SVM	0.494 $\pm$ 0.239	0.222 $\pm$ 0.049
	LogitBoost	0.016 $\pm$ 0.022	0.031 $\pm$ 0.059
	IVM	0.539 $\pm$ 0.140	0.173 $\pm$ 0.023
	Non-linear GPC	0.546 $\pm$ 0.144	0.168 $\pm$ 0.023
	Linear GPC	0.569 $\pm$ 0.166	0.177 $\pm$ 0.026
	Non-linear SVM	0.407 $\pm$ 0.077	0.076 $\pm$ 0.053
	Linear SVM	0.315 $\pm$ 0.195	0.197 $\pm$ 0.058
	LogitBoost	0.008 $\pm$ 0.004	0.012 $\pm$ 0.026
	IVM	0.579 $\pm$ 0.133	0.137 $\pm$ 0.020
	Non-linear GPC	0.577 $\pm$ 0.130	0.136 $\pm$ 0.019
	Linear GPC	0.585 $\pm$ 0.188	0.151 $\pm$ 0.029
	Non-linear SVM	0.488 $\pm$ 0.111	0.039 $\pm$ 0.034
	Linear SVM	0.177 $\pm$ 0.127	0.155 $\pm$ 0.056
	LogitBoost	0.014 $\pm$ 0.019	0.030 $\pm$ 0.056
	IVM	0.931 $\pm$ 0.025	0.080 $\pm$ 0.026
	Non-linear GPC	0.934 $\pm$ 0.021	0.079 $\pm$ 0.023
	Linear GPC	0.925 $\pm$ 0.031	0.085 $\pm$ 0.027
	Non-linear SVM	0.641 $\pm$ 0.168	0.100 $\pm$ 0.047
	Linear SVM	0.705 $\pm$ 0.142	0.212 $\pm$ 0.049
	LogitBoost	0.059 $\pm$ 0.103	0.077 $\pm$ 0.127

train



VS



Test Class	Classifier	Normalised Entropy	
		$\mu \pm \text{std. err.}$	$\sigma \pm \text{std. err.}$
	IVM	0.776 $\pm$ 0.081	0.145 $\pm$ 0.030
	Non-linear GPC	0.751 $\pm$ 0.087	0.152 $\pm$ 0.029
	Linear GPC	0.776 $\pm$ 0.108	0.150 $\pm$ 0.041
	Non-linear SVM	0.476 $\pm$ 0.101	0.089 $\pm$ 0.056
	Linear SVM	0.664 $\pm$ 0.122	0.250 $\pm$ 0.041
	LogitBoost	0.019 $\pm$ 0.025	0.041 $\pm$ 0.073
	IVM	0.794 $\pm$ 0.117	0.106 $\pm$ 0.026
	Non-linear GPC	0.779 $\pm$ 0.124	0.107 $\pm$ 0.024
	Linear GPC	0.777 $\pm$ 0.202	0.124 $\pm$ 0.058
	Non-linear SVM	0.537 $\pm$ 0.126	0.028 $\pm$ 0.036
	Linear SVM	0.494 $\pm$ 0.239	0.222 $\pm$ 0.049
	LogitBoost	0.016 $\pm$ 0.022	0.031 $\pm$ 0.059
	IVM	0.539 $\pm$ 0.140	0.173 $\pm$ 0.023
	Non-linear GPC	0.546 $\pm$ 0.144	0.168 $\pm$ 0.023
	Linear GPC	0.569 $\pm$ 0.166	0.177 $\pm$ 0.026
	Non-linear SVM	0.407 $\pm$ 0.077	0.076 $\pm$ 0.053
	Linear SVM	0.315 $\pm$ 0.195	0.197 $\pm$ 0.058
	LogitBoost	0.008 $\pm$ 0.004	0.012 $\pm$ 0.026
	IVM	0.579 $\pm$ 0.133	0.137 $\pm$ 0.020
	Non-linear GPC	0.577 $\pm$ 0.130	0.136 $\pm$ 0.019
	Linear GPC	0.585 $\pm$ 0.188	0.151 $\pm$ 0.029
	Non-linear SVM	0.488 $\pm$ 0.111	0.039 $\pm$ 0.034
	Linear SVM	0.177 $\pm$ 0.127	0.155 $\pm$ 0.056
	LogitBoost	0.014 $\pm$ 0.019	0.030 $\pm$ 0.056
	IVM	0.931 $\pm$ 0.025	0.080 $\pm$ 0.026
	Non-linear GPC	0.934 $\pm$ 0.021	0.079 $\pm$ 0.023
	Linear GPC	0.925 $\pm$ 0.031	0.085 $\pm$ 0.027
	Non-linear SVM	0.641 $\pm$ 0.168	0.100 $\pm$ 0.047
	Linear SVM	0.705 $\pm$ 0.142	0.212 $\pm$ 0.049
	LogitBoost	0.059 $\pm$ 0.103	0.077 $\pm$ 0.127

Any questions?

Scene

Image

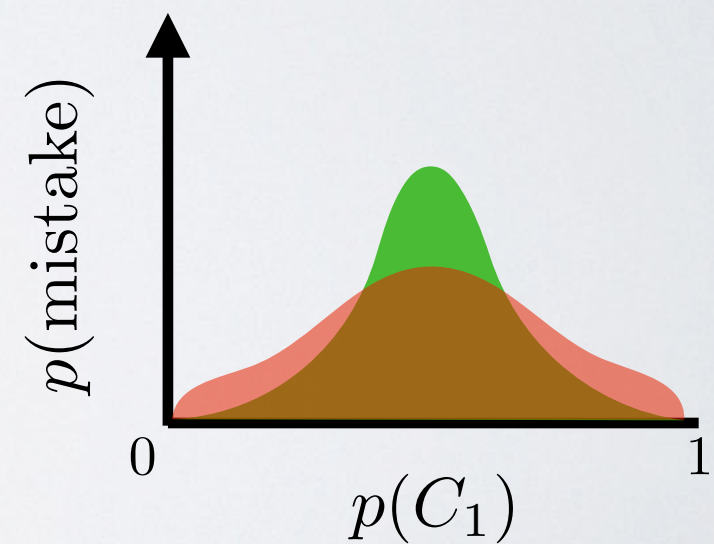
Feature

Classifier

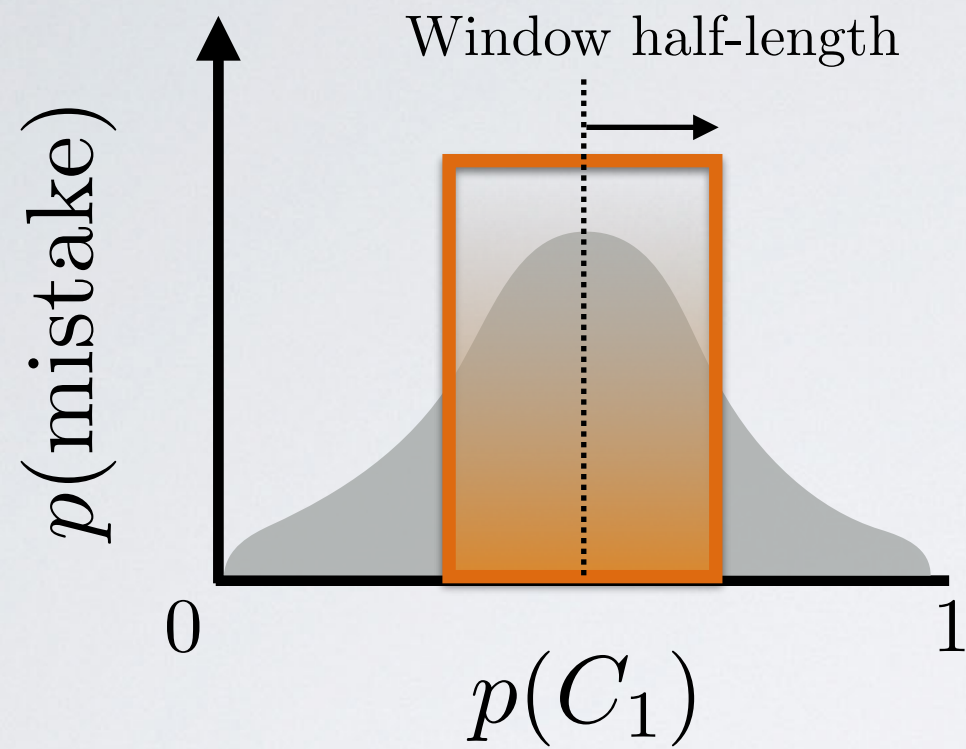
Costs

Action

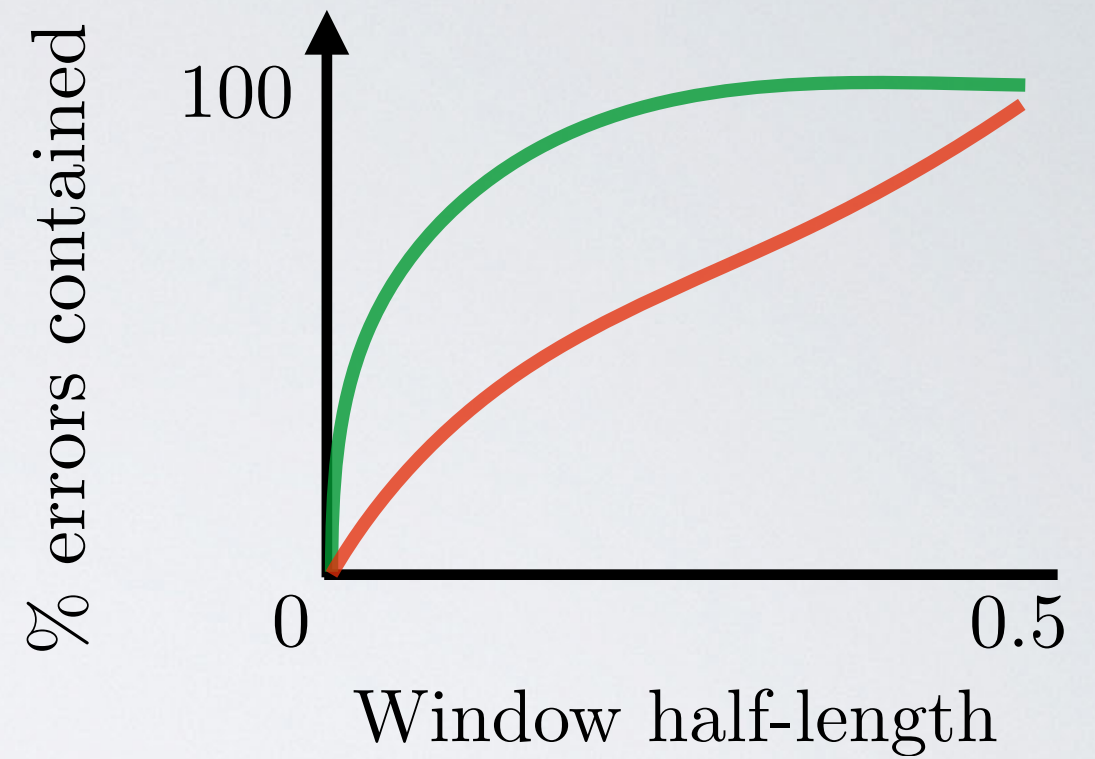
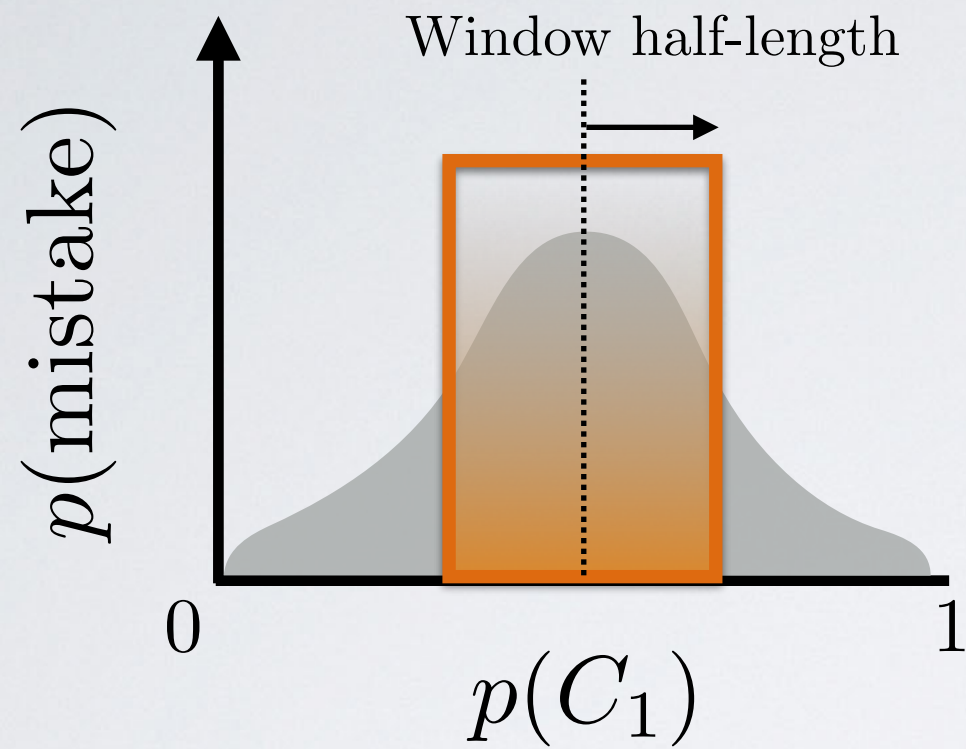
Introspection



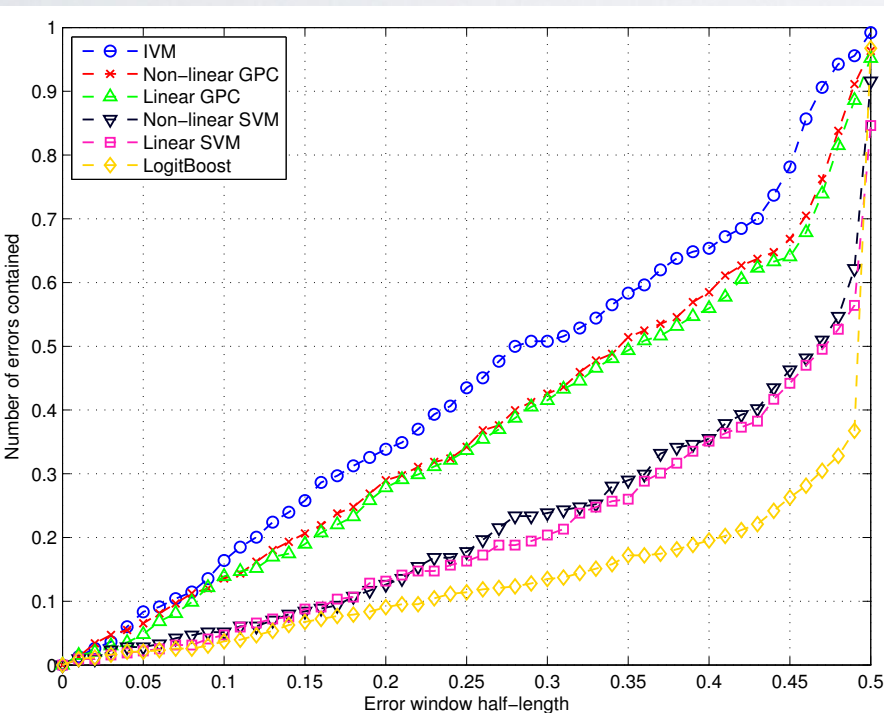
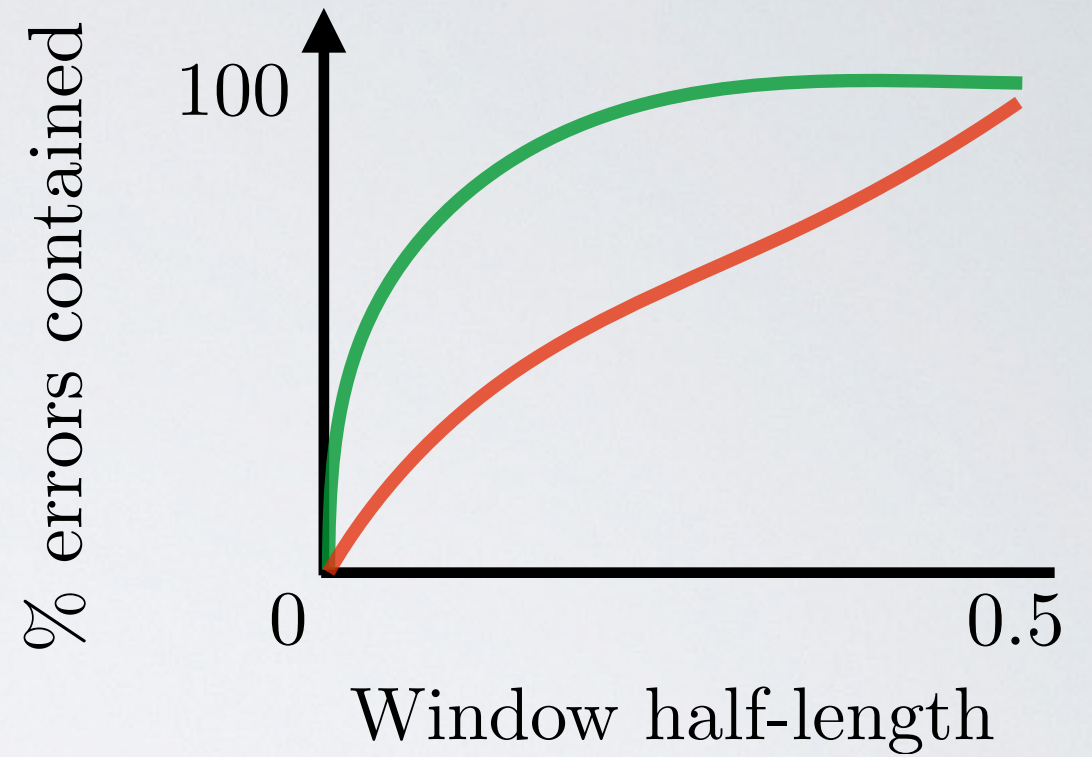
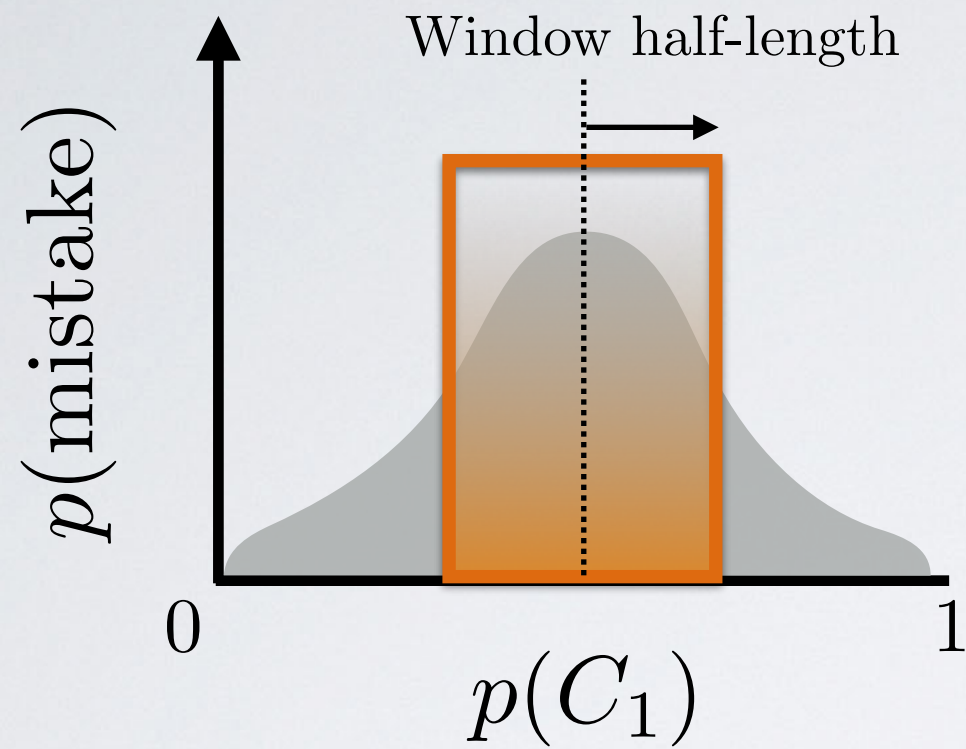
How confident are the errors?



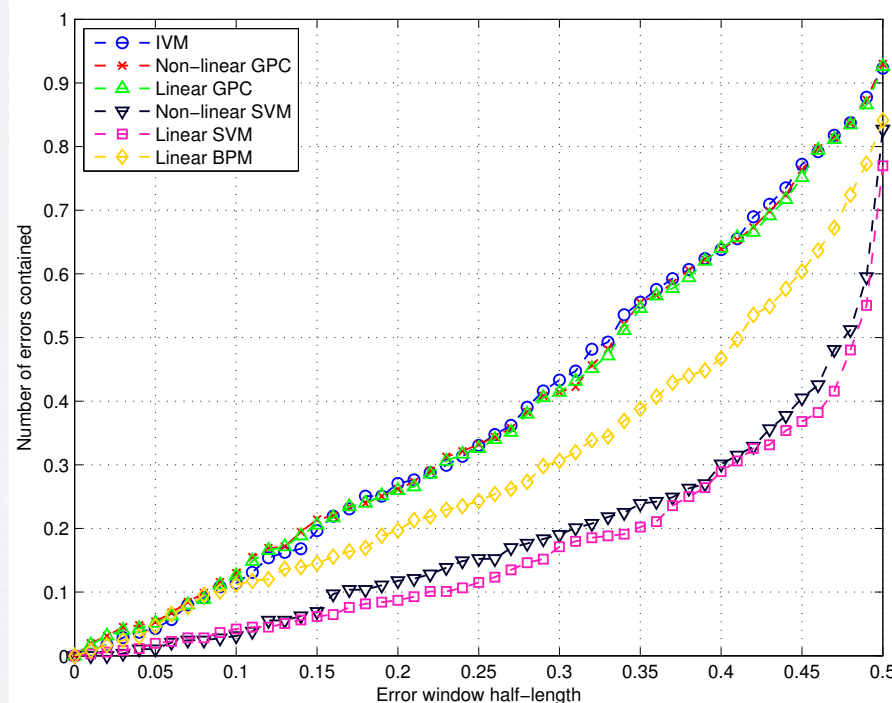
How confident are the errors?



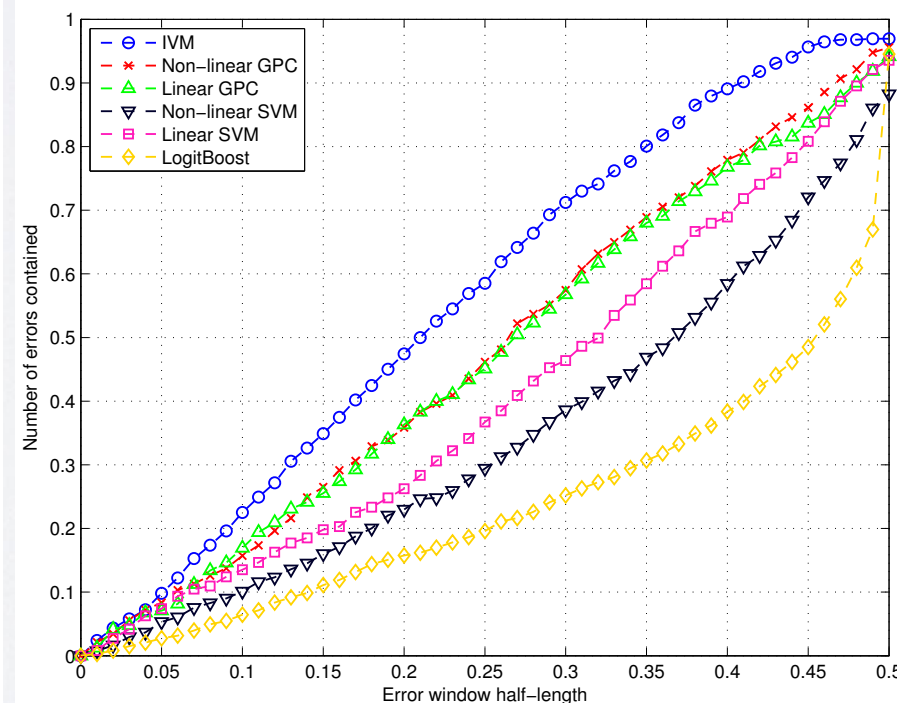
How confident are the errors?



(a) TLR



(b) DP



(c) KITTI

Fig. 13: Errors contained in central regions of varying sizes

What about the precision and recall?

Classifier	TLR			Daimler Pedestrian			KITTI		
	Precision	Recall	F_1	Precision	Recall	F_1	Precision	Recall	F_1
IVM	0.995	0.619	0.763	0.956	0.680	0.795	0.950	0.399	0.562
Non-linear GPC	0.977	0.632	0.767	0.959	0.674	0.792	0.933	0.419	0.578
Linear GPC	0.969	0.625	0.760	0.958	0.680	0.795	0.903	0.392	0.547
Non-linear SVM	0.976	0.806	0.883	0.931	0.768	0.842	0.880	0.528	0.660
Linear SVM	0.928	0.738	0.822	0.887	0.738	0.806	0.902	0.425	0.578
LogitBoost	0.972	0.587	0.732	0.965	0.446	0.610	0.871	0.317	0.465

Fig. 5: Performance the detection tasks across data sets for the classifiers.

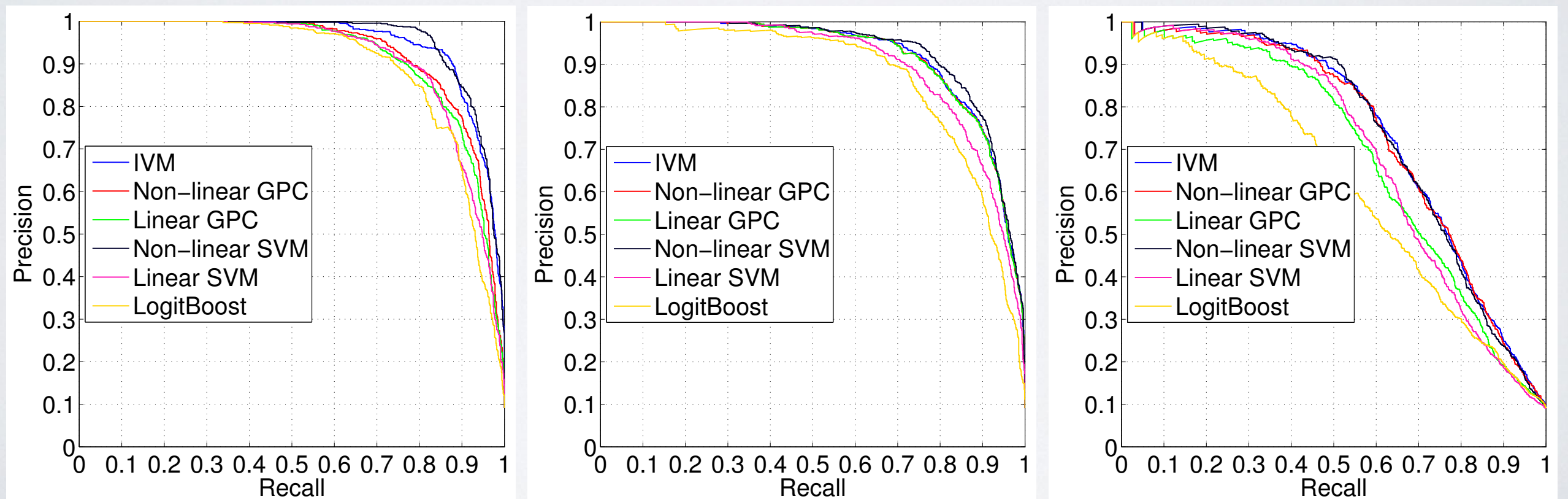
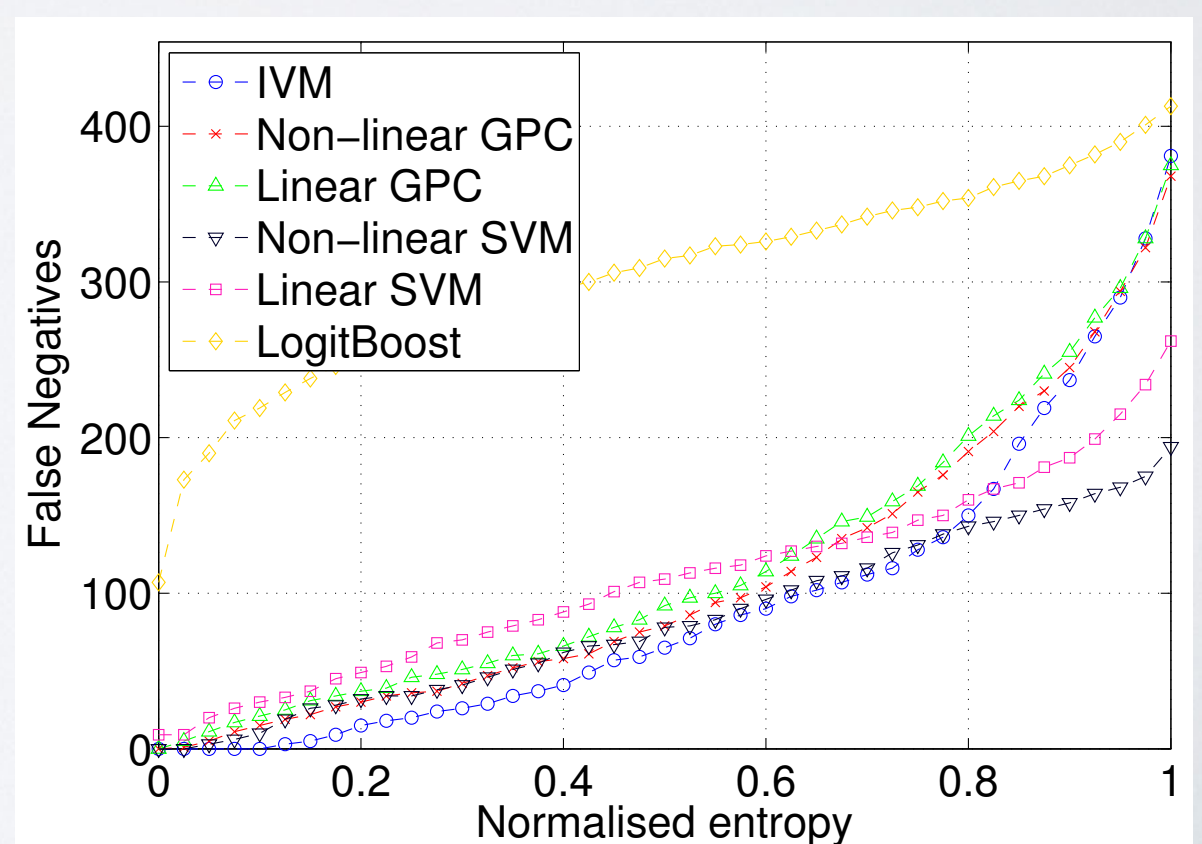
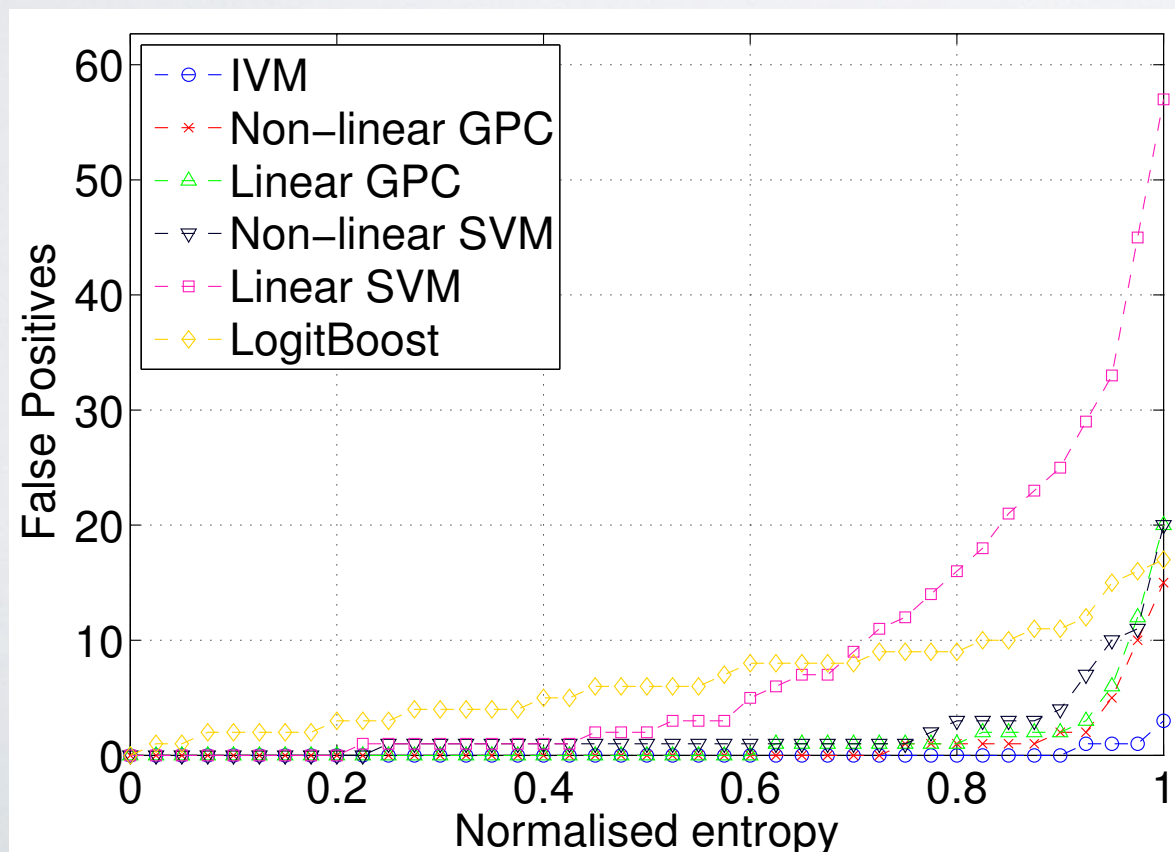
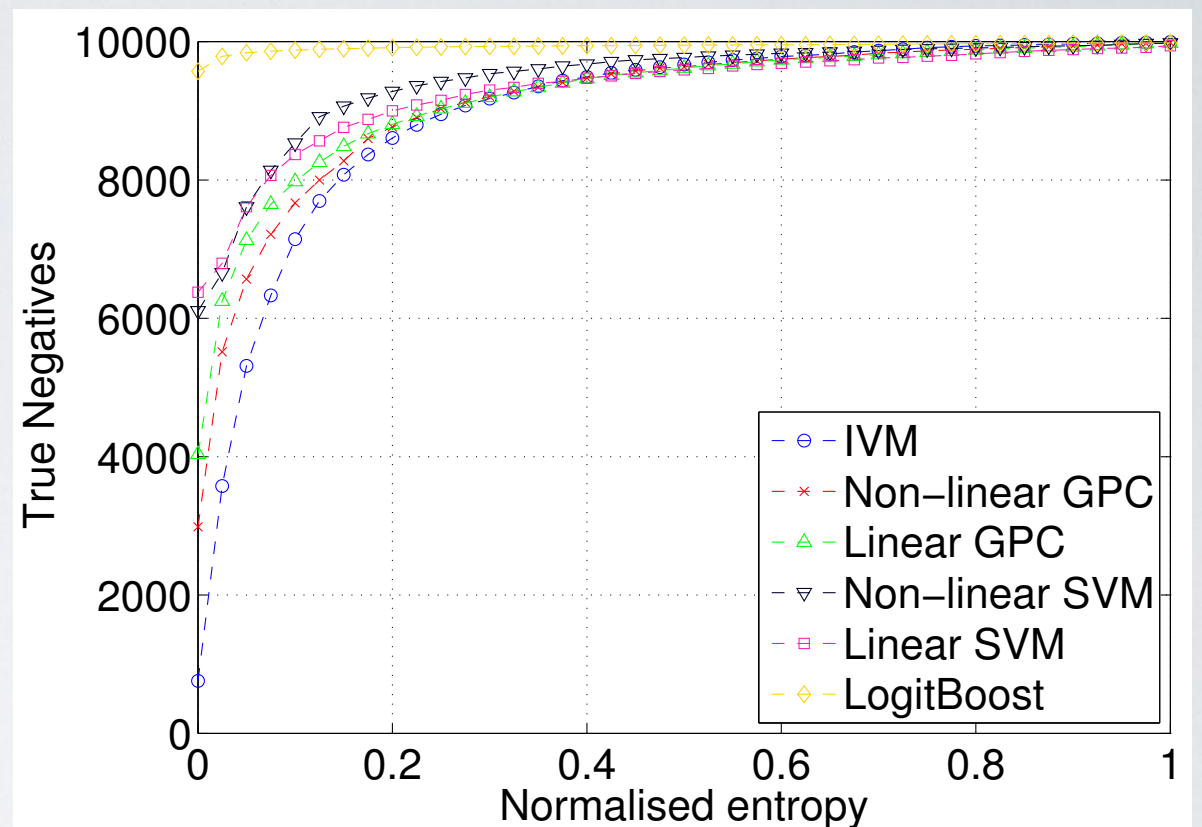
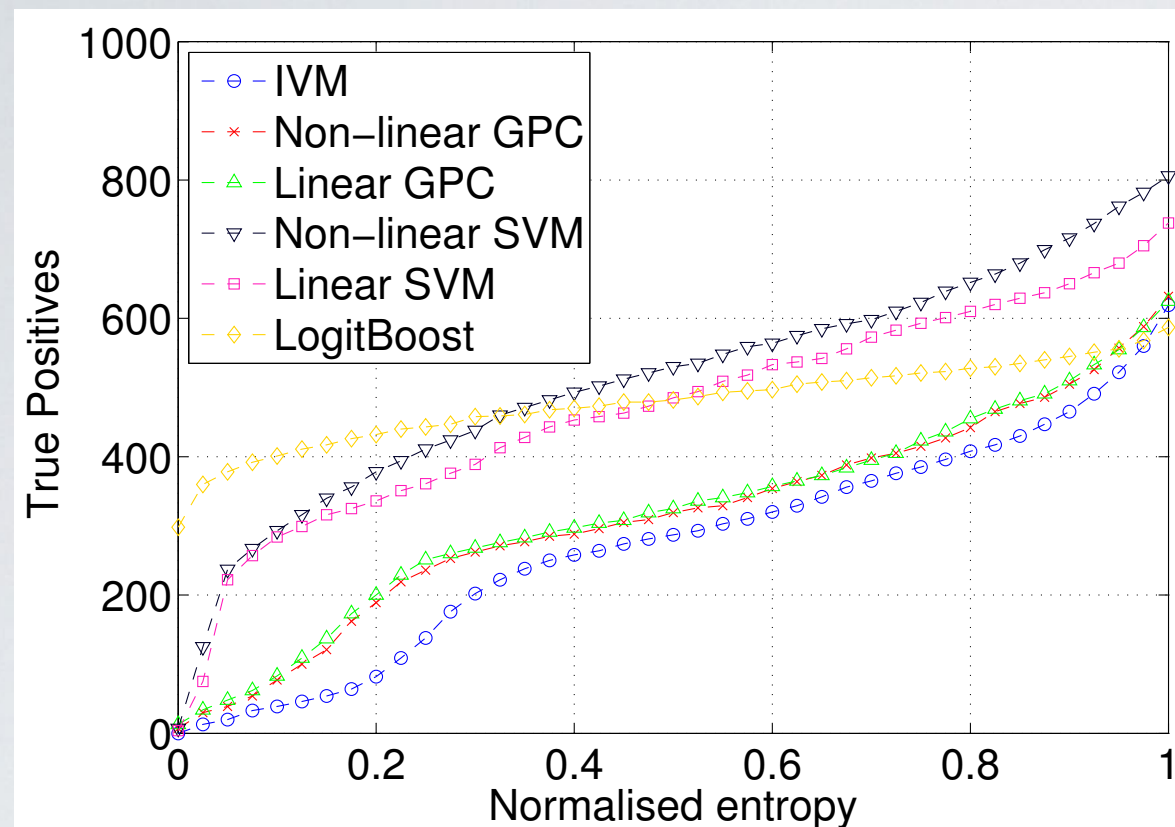
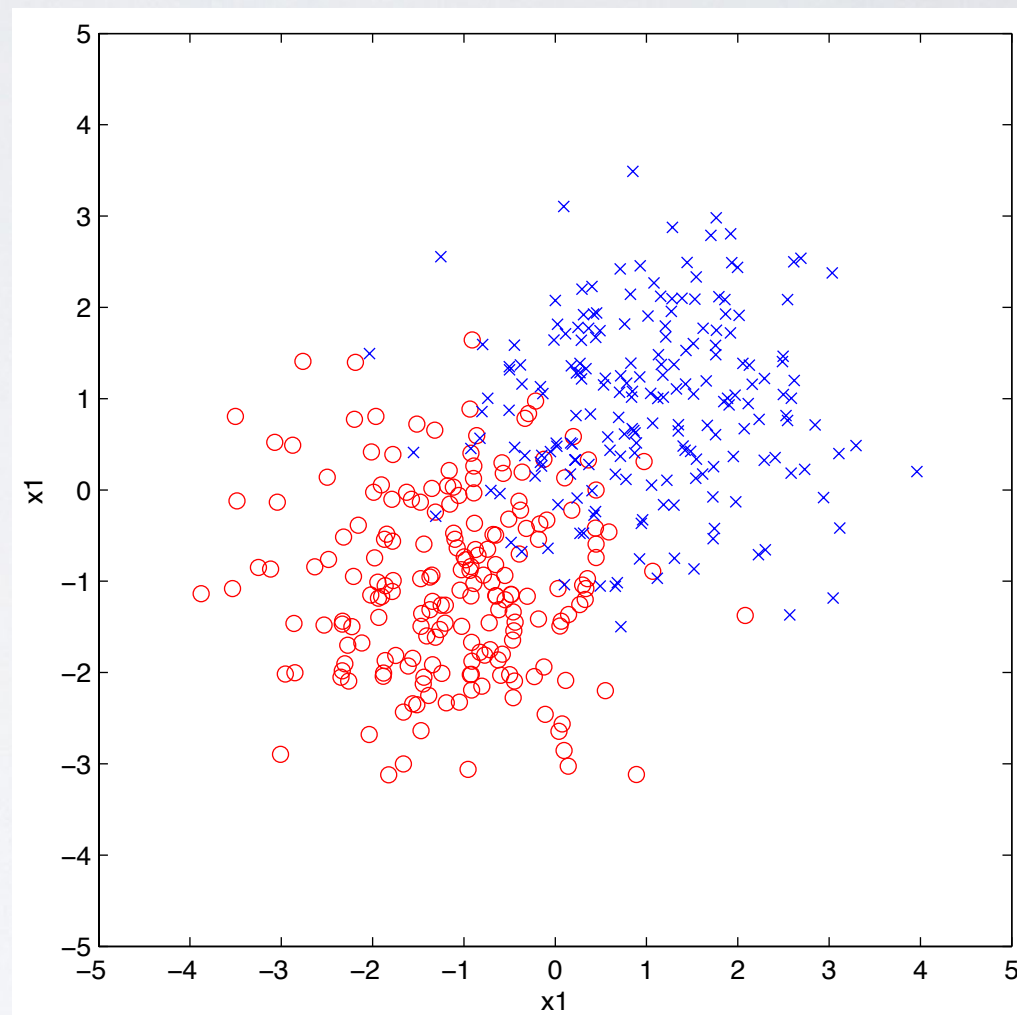


Fig. 6: Precision-recall curves for the TLR, Daimler Pedestrian and KITTI data sets (Best viewed in colour.)

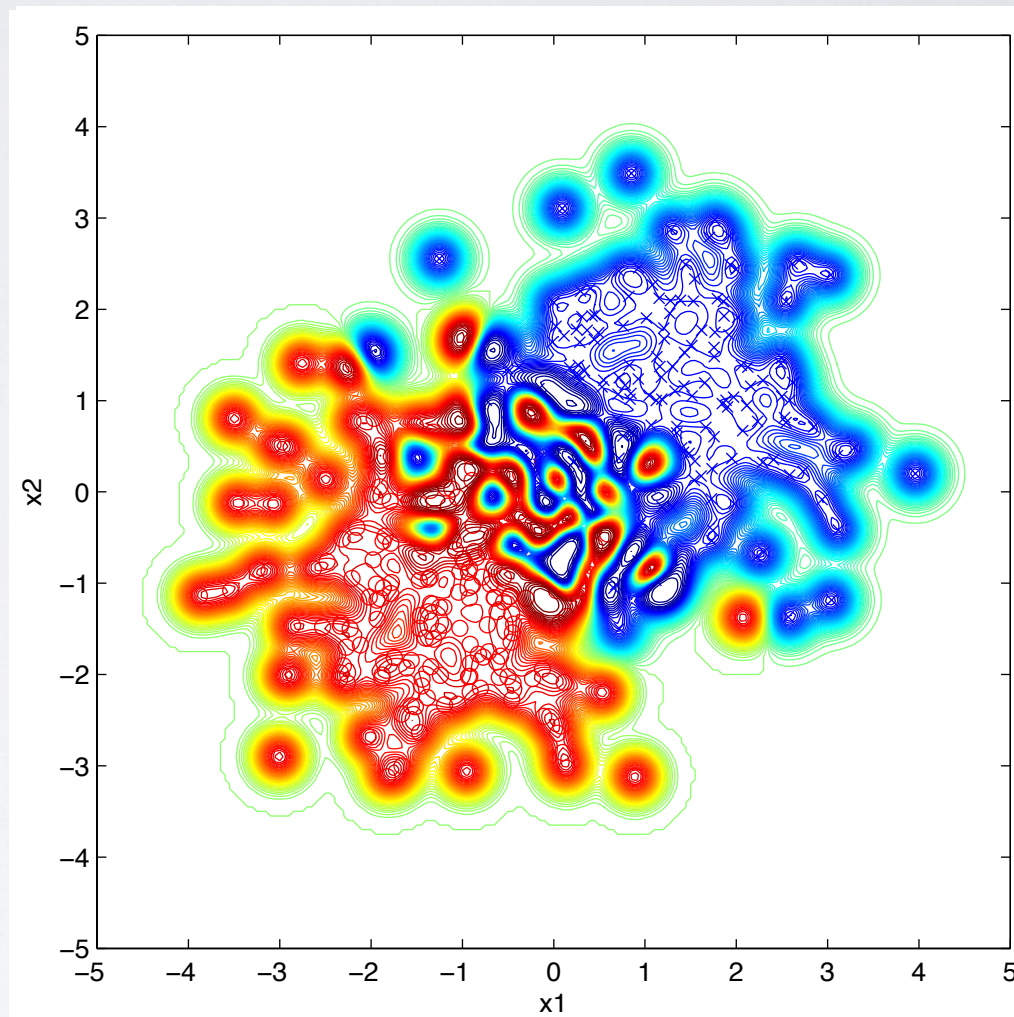
Cumulative decisions across uncertainty



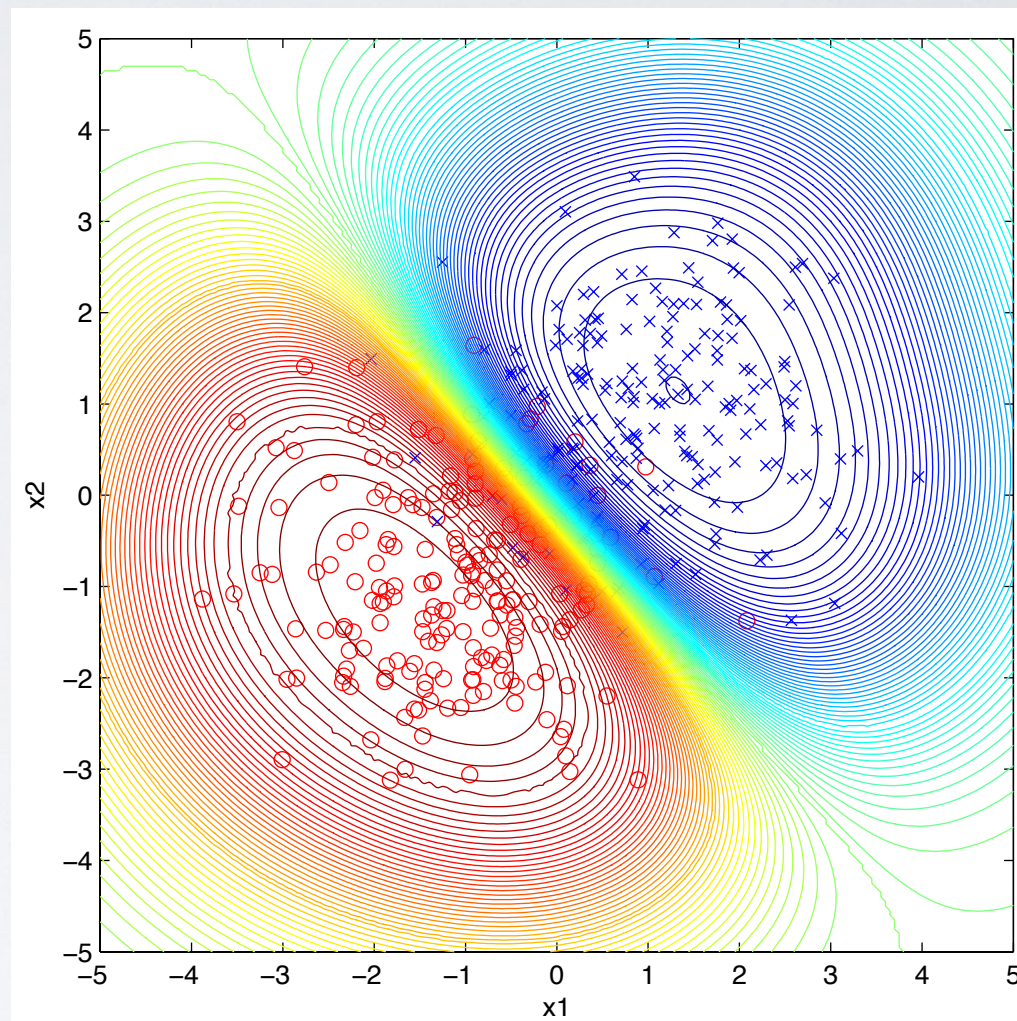
Overfitting



Overfitting

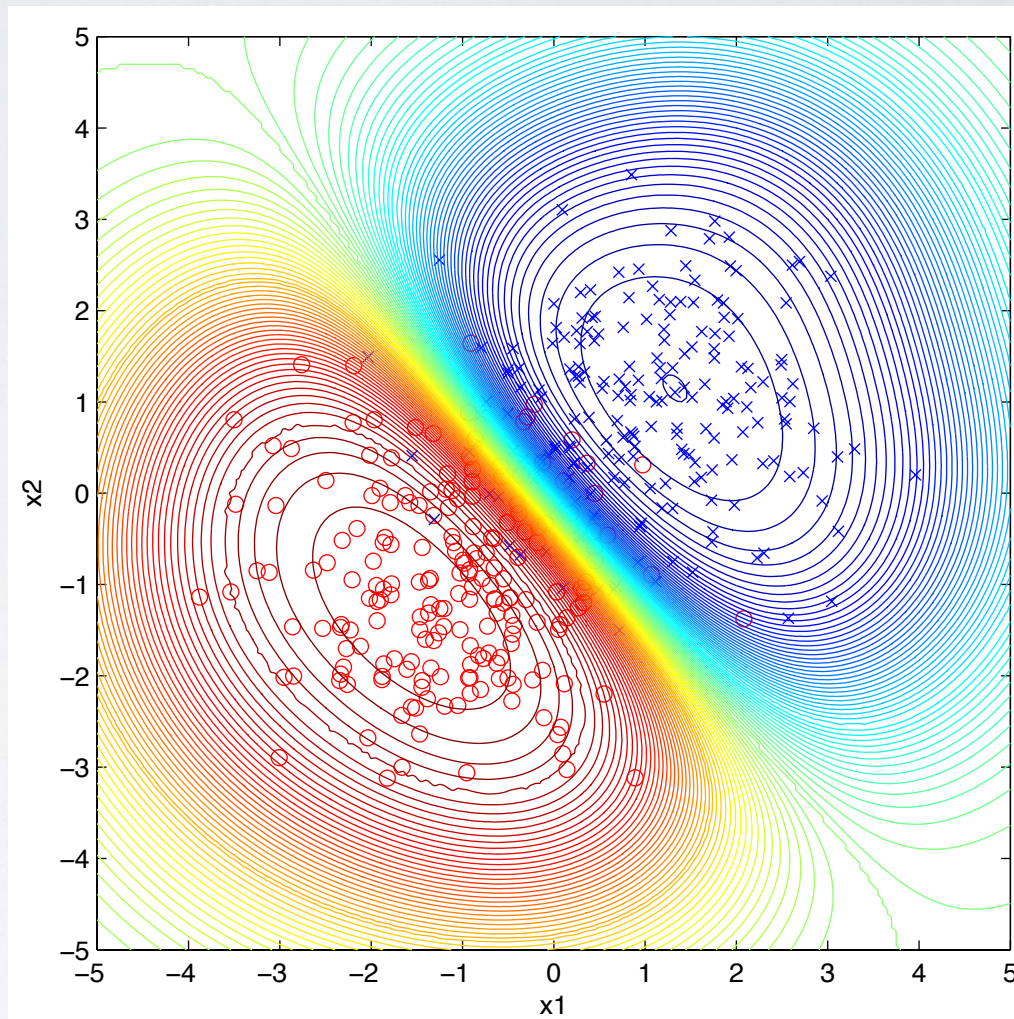


Overfitting



Overfitting

- Cross-validation
- Regularisation
- Pruning
- Bayesian priors
- Ensemble methods



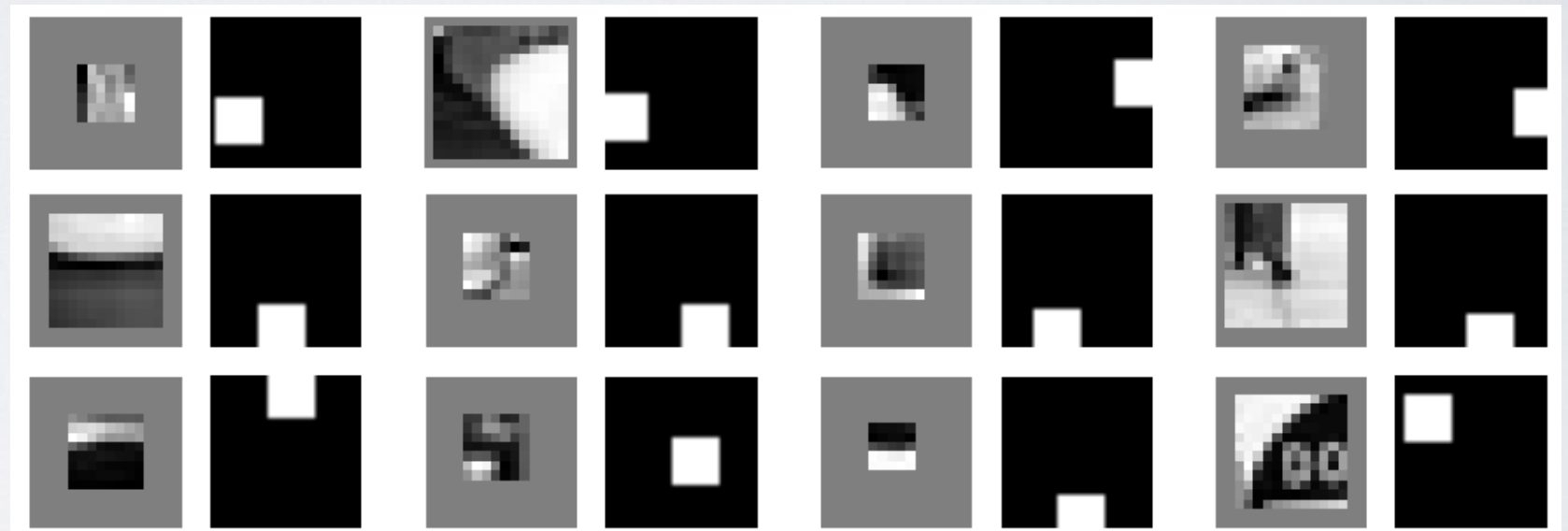
Features



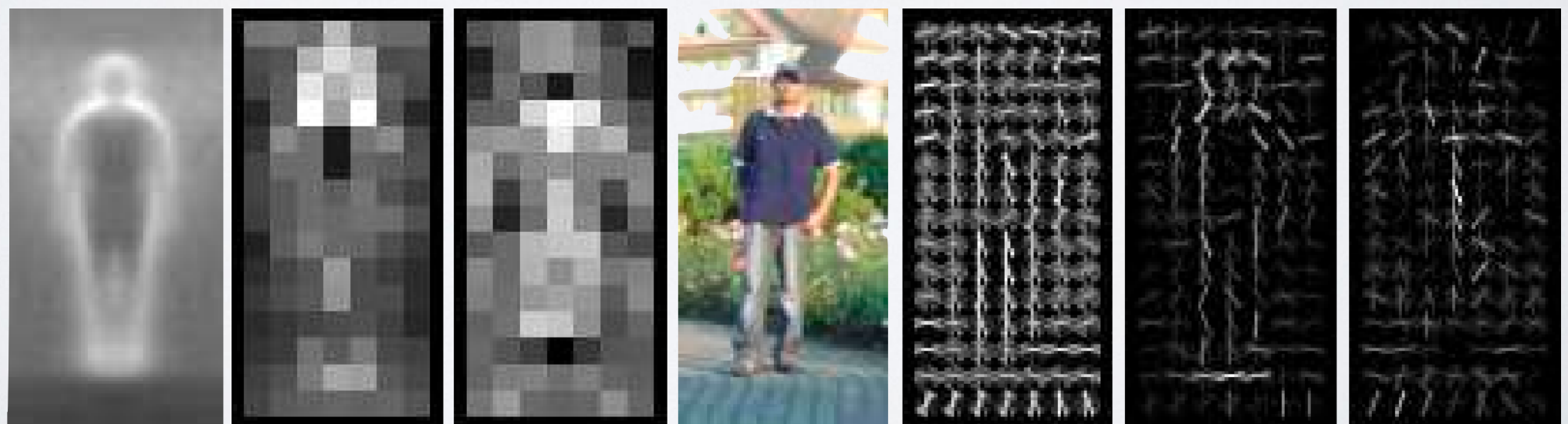
86	133	226	212	218	218	222	212	218	222	226	218
75	156	177	218	212	218	218	218	218	222	218	218
96	84	133	203	218	218	218	222	212	218	222	218
123	75	111	156	212	218	212	212	218	218	218	226
93	75	71	133	184	231	226	226	222	212	218	218
51	75	75	75	156	206	218	218	218	222	212	222
44	110	75	64	143	194	231	218	218	218	218	218
52	123	69	44	60	156	198	231	231	222	226	226
52	75	84	8	65	69	150	23	231	226	23	231
36	36	84	93	84	71	156	160	240	240	23	231
36	40	113	75	69	75	71	133	194	240	240	240
52	52	105	85	69	75	75	123	111	222	231	231
69	44	69	93	81	75	75	69	150	177	247	240
73	44	40	96	101	75	75	75	84	133	23	240

Pixel intensity

Templates



HOG



(a)

(b)

(c)

(d)

(e)

(f)

(g)

Parametric vs non-parametric classifiers

Confusing words to describe two types of classification frameworks

Parametric

- Choose a model for the data
- That model changes depending on parameters used
- That model is used *in lieu* of the data

Non-parametric

- Let the data do the talking
- Can have parameters, but typically they are less crucial