Adapting your classifiers for real-world applications

Estimating uncertainty with Background Check

Miquel Perelló Nieto - PhD Student and Research Associate

18th of February, 2022

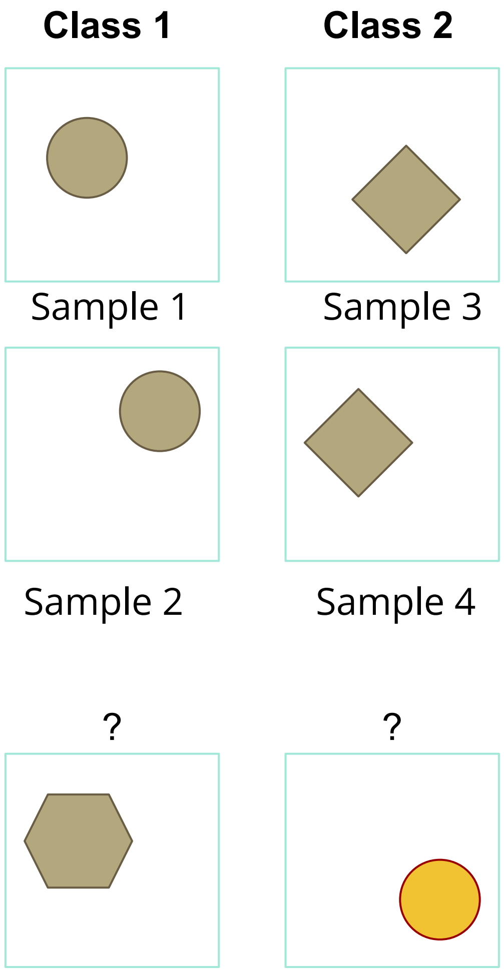
University of BRISTOL

Training data vs deployment data

- The real-world is not static
- Model assumptions
- Uncertainty quantification
- Consider the option to abstain

Related tasks

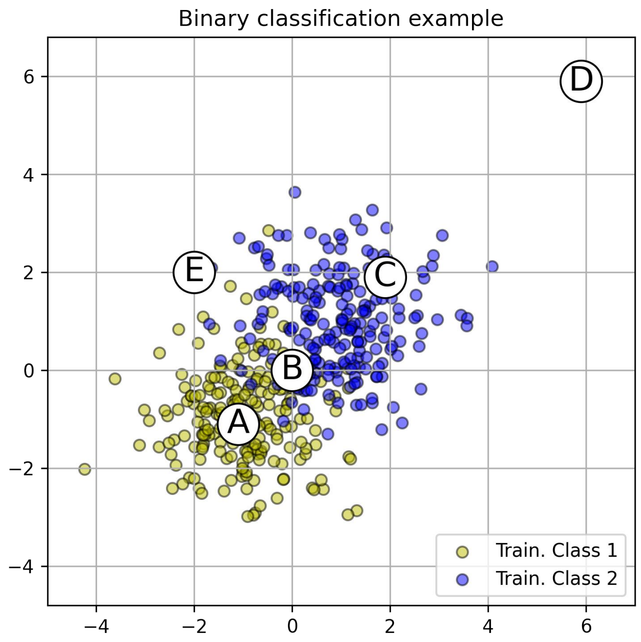
cautious classification, active learning, semi-supervised learning, novelty detection, outlier detection, anomaly detection, online learning.



Binary classification example

Binary classification problem with two features, but generalises to arbitrary number of classes and dimensions.

- A, B, and C are in **dense** regions
- E and D are in **low density** regions
- B and E are in the **decision boundary**

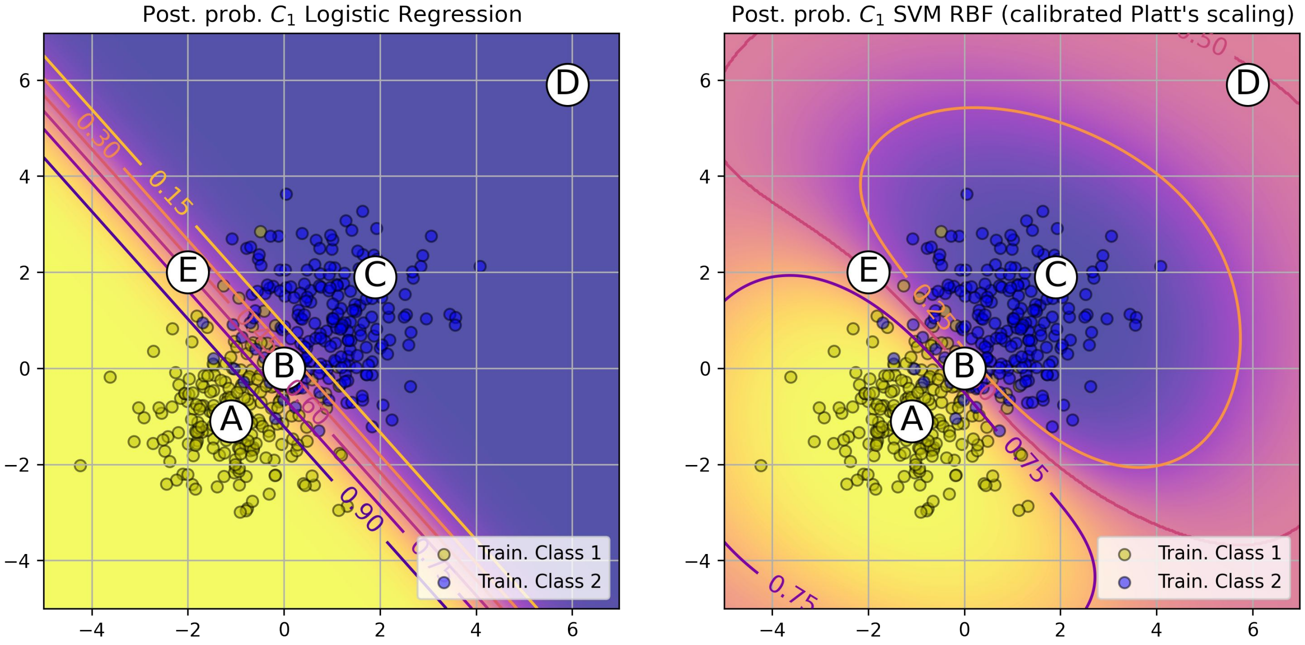


Two common classifiers

Minimization of empirical risk.

Focus on the performance in regions of high density.

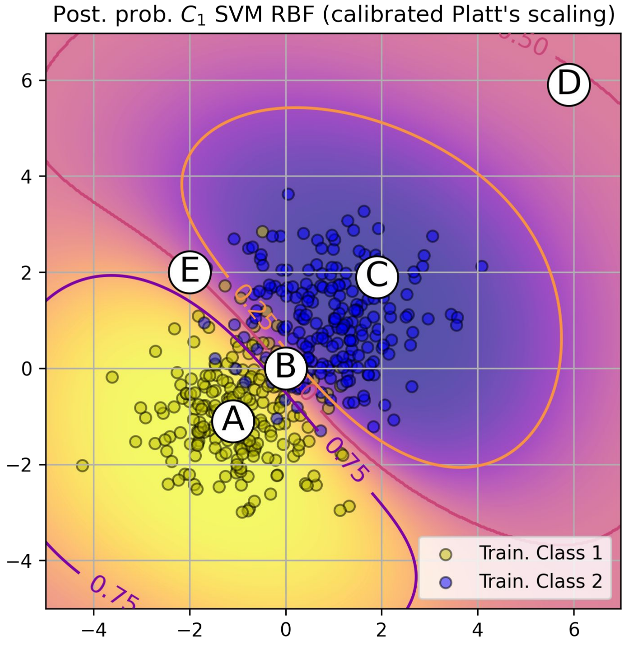
Expect same data distribution during deployment.



Interpretation of the posterior probabilities

- A is clearly from Class 1
- C is clearly from Class 2
- B, E and D are in the same issoline 0.5
- **Several examples in B**
- **No examples in D**

	$p(C_1\|x)$	$p(C_2\|x)$
A	1	.0
B	*.5*	*.5*
C	.0	1
D	*.5*	*.5*
E	*.5*	*.5*

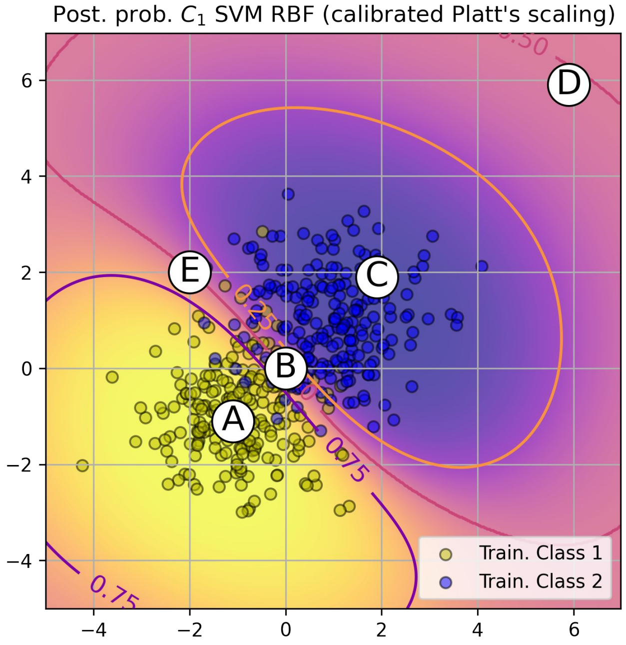


Adding an additional posterior probability (background)

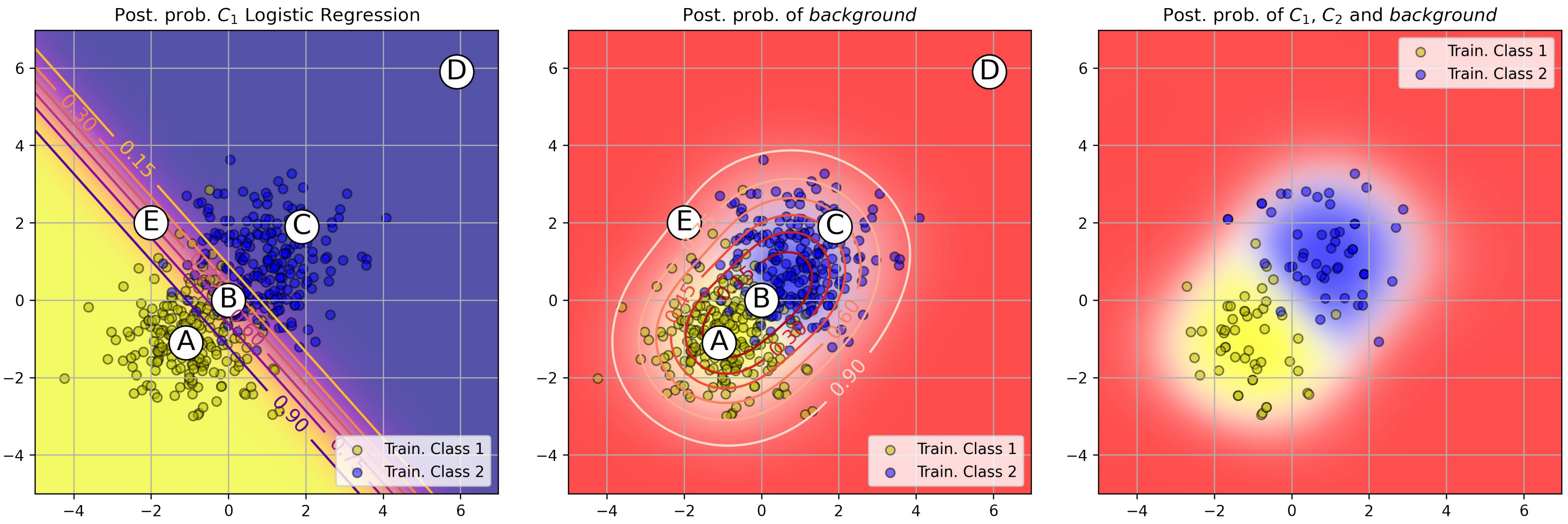
We refer to the foreground class as the known training data, and background class the rest.

- We are certain about B being ambiguous
- We are uncertain about D

	$p(C_1\|x)$	$p(C_2\|x)$	$p(b\|x)$
A	$1 \rightarrow .9$	$.0 \rightarrow .0$	$.1$
B	$.5 \rightarrow .5$	$.5 \rightarrow .5$	$.0$
C	$.0 \rightarrow .0$	$1 \rightarrow .5$	$.5$
D	$.5 \rightarrow .0$	$.5 \rightarrow .0$	1
E	$.5 \rightarrow .1$	$.5 \rightarrow .1$	$.8$



Objective: Adapt an arbitrary classifier to provide familiarity



How to adapt the probabilities in theory

Base classifier: known posterior class probabilities

$$p(f_c|f,x) = \frac{p(x|f,f_c)p(f_c|f)}{p(x|f)} \quad \text{for} \quad c = 1,\dots,C$$

We want: foreground vs background posterior probabilities

$$p(f|x) = \frac{p(x|f)p(f)}{p(x)} \qquad p(b|x) = \frac{p(x|b)p(b)}{p(x)}.$$

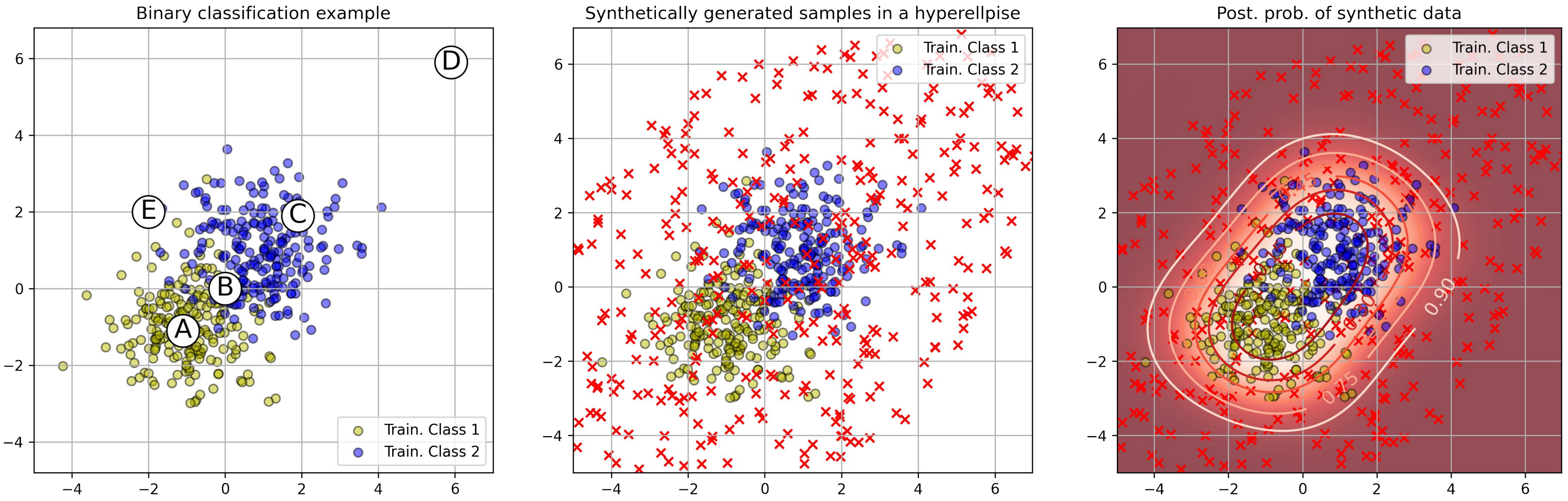
We only need the ratio between the previous probabilities

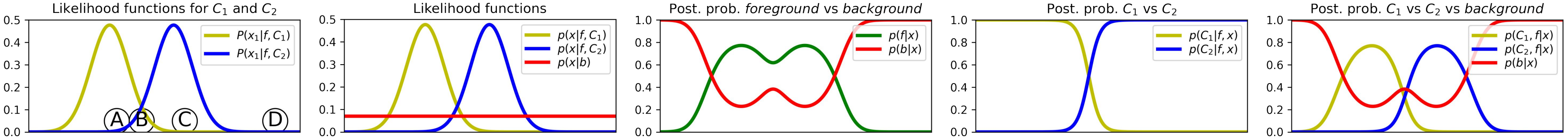
$$r(x) = p(f|x)/p(b|x)$$

We obtain posteriors for all foreground classes and background class

$$p(f_c|x) = \frac{p(f_c|f,x)r(x)}{1+r(x)} \qquad p(b|x) = \frac{1}{1+r(x)}$$

A discriminative approach and synthetic data

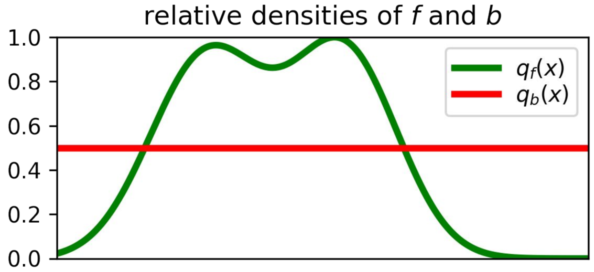




A familiarity approach and density estimation

- Estimate density of foreground (training data)
- Obtain **relative density** with respect to the maximum of foreground

$$q_f(x) = \frac{p(f,x)}{\max_x p(x,f)},$$
$$q_b(x) = \frac{p(b,x)}{\max_x p(x,f)}$$



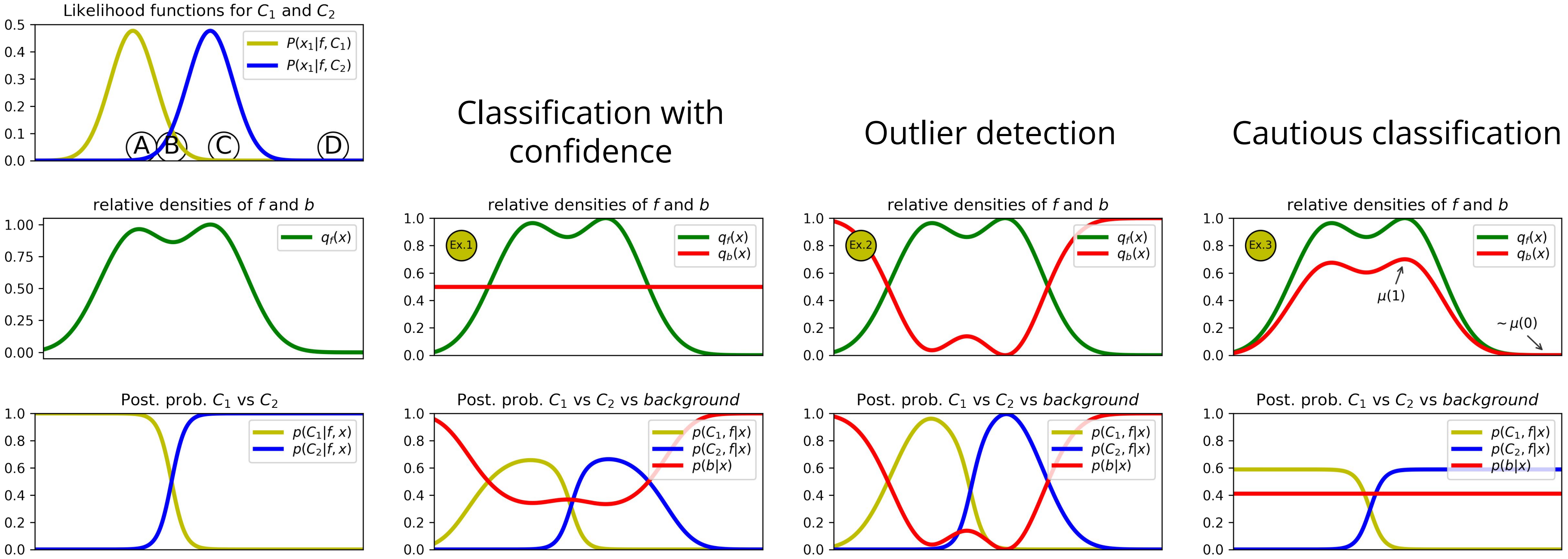
With those, we can still obtain the familiarity ratio

$$\cancel{r(x) = p(f|x)/p(b|x)} \qquad r(x) = q_f(x)/q_b(x)$$

Obtain the new posterior probabilities

$$p(f_c|x) = \frac{p(f_c|f,x)r(x)}{1+r(x)} \qquad p(b|x) = \frac{1}{1+r(x)}$$

A familiarity approach



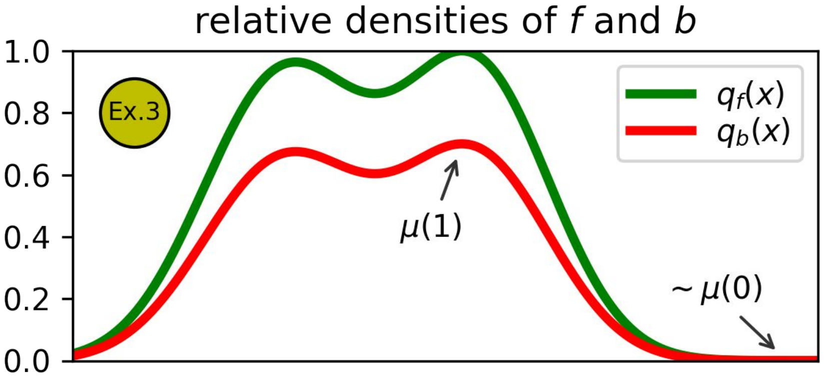
A familiarity approach and affine transformation

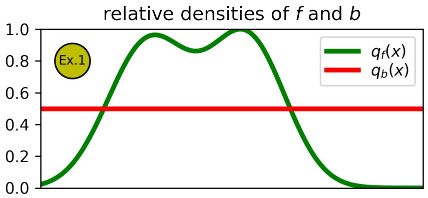
Relative density of the background as a function of the foreground

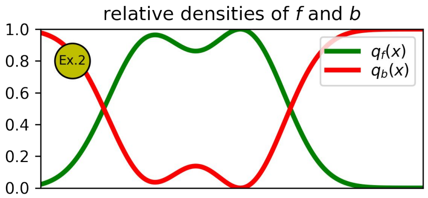
$$q_b(x) = \mu(q_f(x))$$

Parametric form with minimum and maximum values.

$$\mu(z) = (1 - z)\mu(0) + z\mu(1)$$

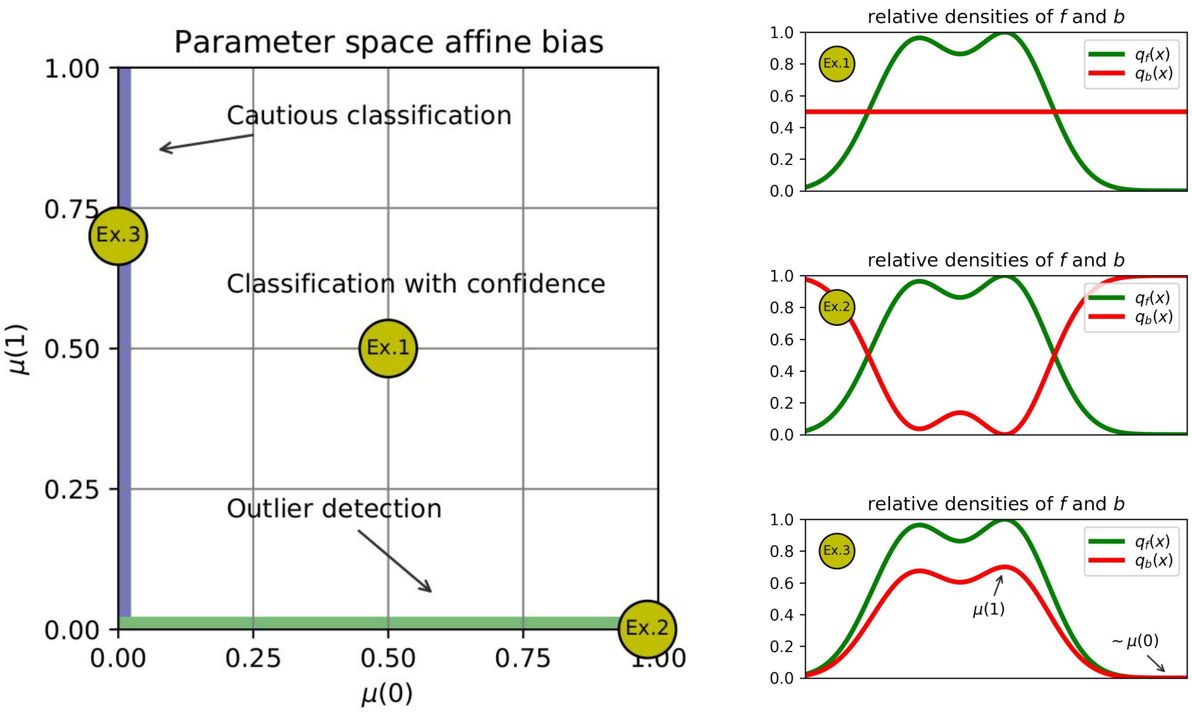






A familiarity approach and affine transformation

Other possible values



$$q_b(x) = \mu(q_f(x))$$

$$\mu(z) = (1 - z)\mu(0) + z\mu(1)$$

Results

Our tests with 41 multiclass datasets showed:

1. Significantly better performance in **classification with confidence** against a SOTA method
2. Competitive results for **outlier detection** against two specialised methods

And it is equivalent to Chow’s rule to perform **cautious classification**

Conclusion

- Consider the **model assumptions** in real-world problems (some ML algorithms make strong assumptions)
- The available data for **training may be biased**.
- We designed Background Check as a theoretical framework which can be applied to multiple problems.

More details in:

M. Perello-Nieto, T. M. S. Filho, M. Kull and P. Flach, "**Background Check: A General Technique to Build More Reliable and Versatile Classifiers**," 2016 IEEE 16th International Conference on Data Mining (ICDM), 2016, pp. 1143-1148, doi: 10.1109/ICDM.2016.0150.

reframe.github.io/background_check