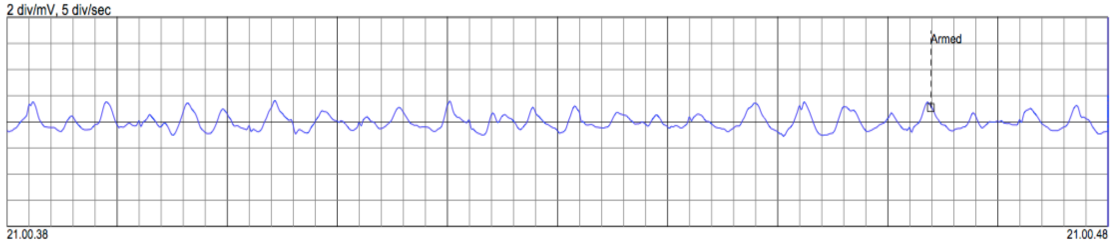
Classification with guaranteed specificity and sensitivity for medical applications

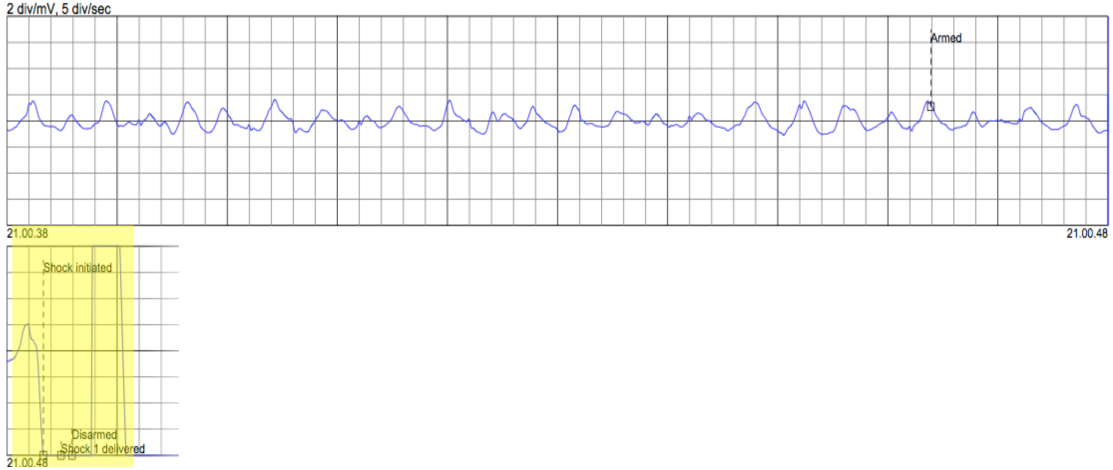
Algo Carè, Marco C. Campi, Federico A. Ramponi

Università degli Studi di Brescia

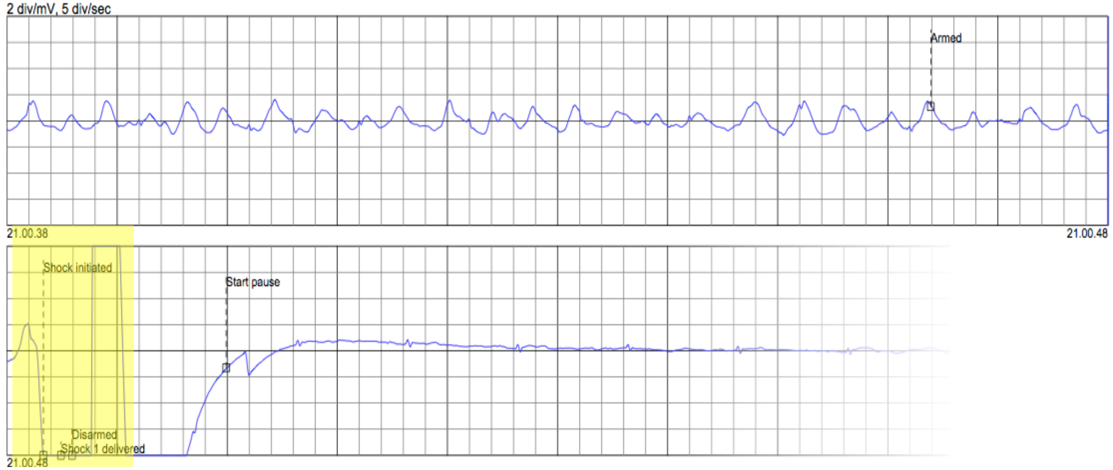
A difficult decision



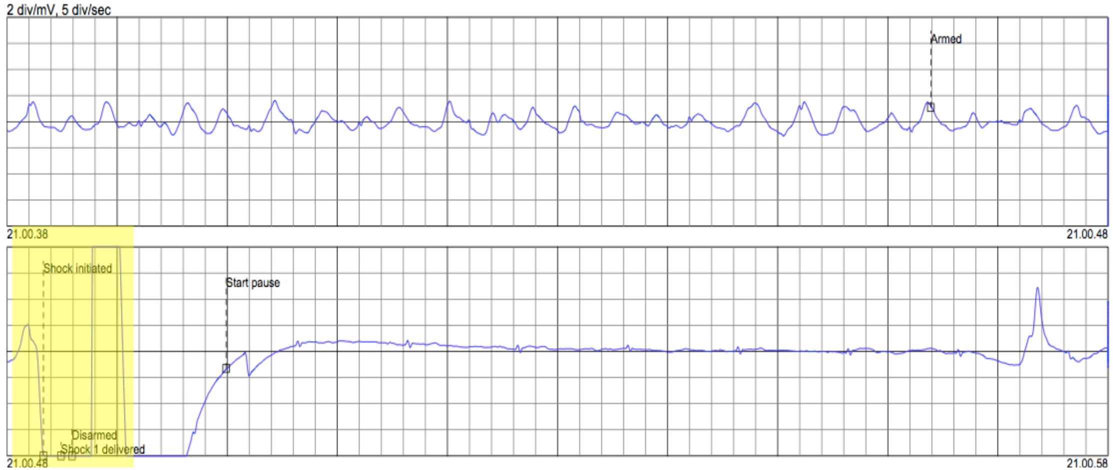
A difficult decision



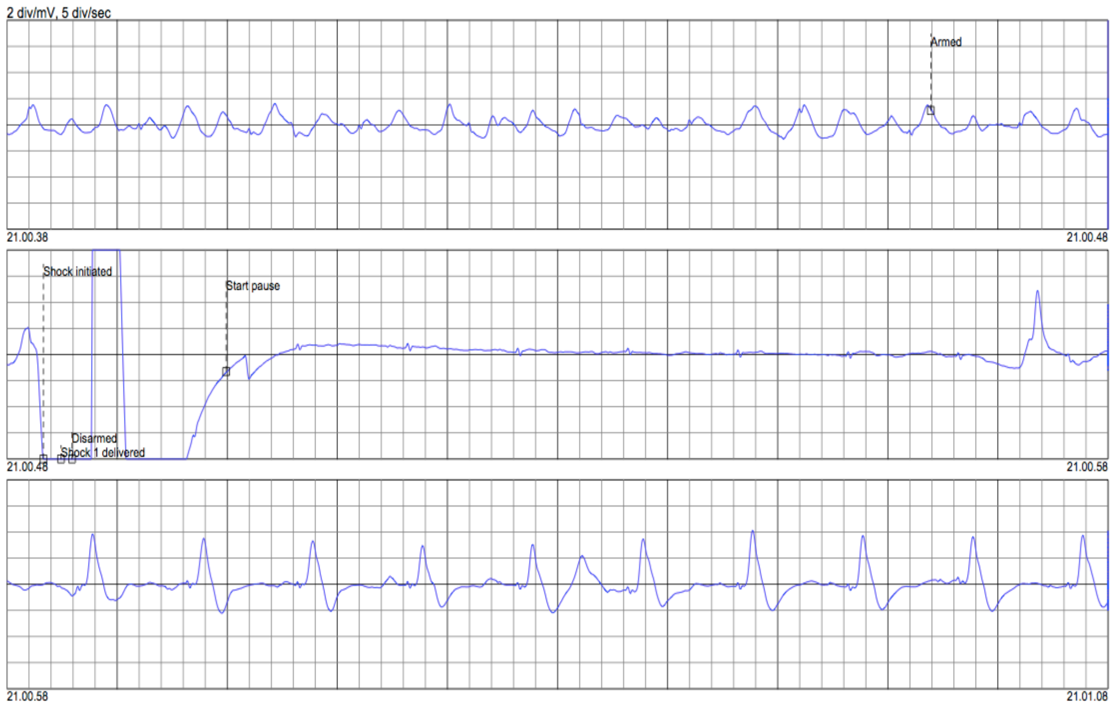
A difficult decision



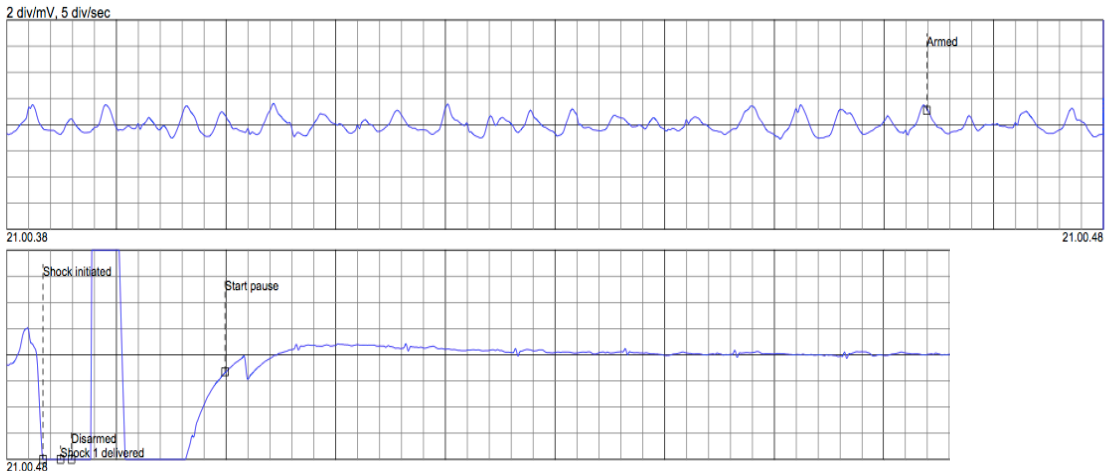
A difficult decision



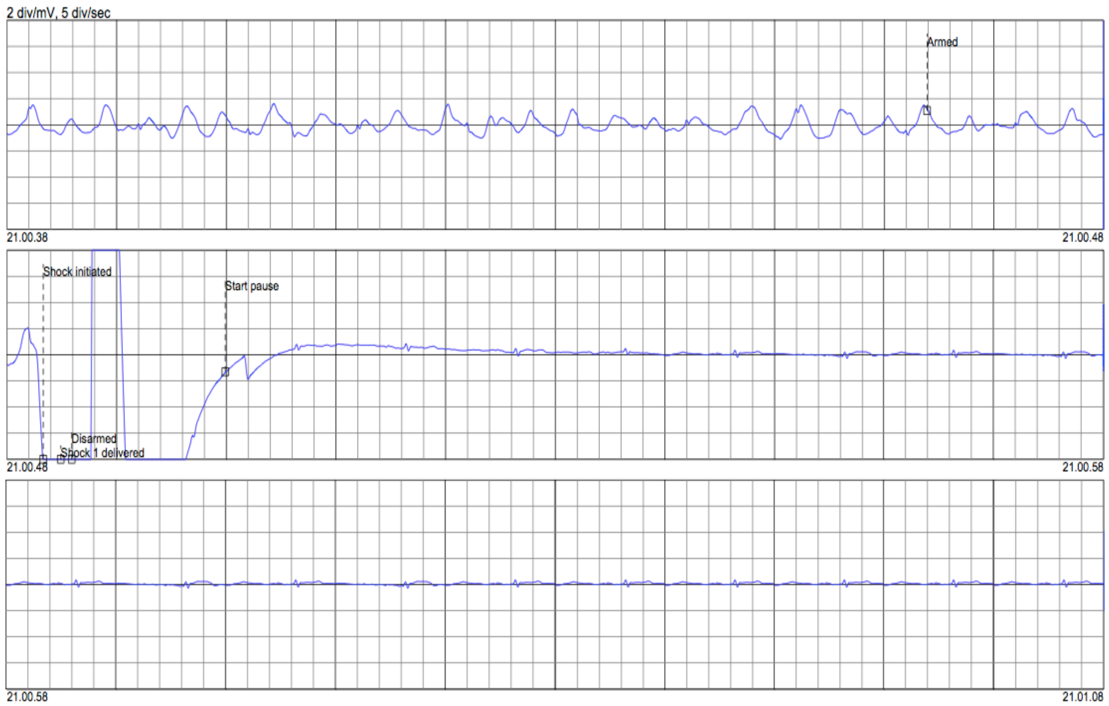
A difficult decision



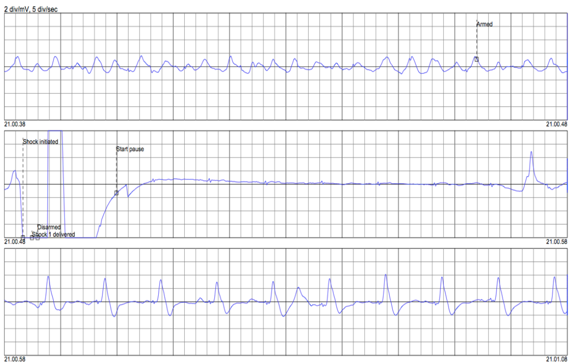
A difficult decision



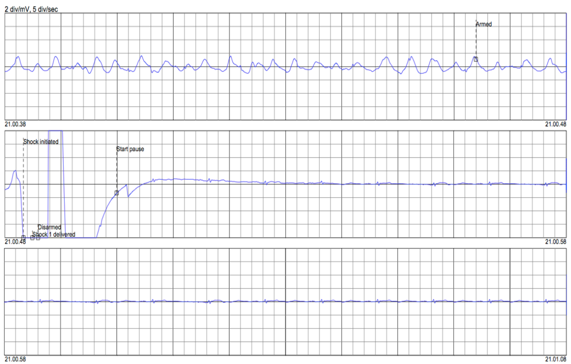
A difficult decision



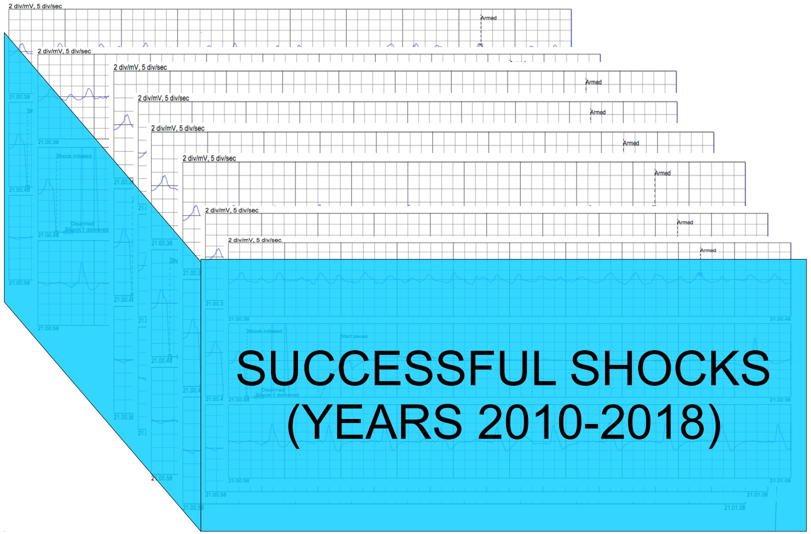
A difficult decision

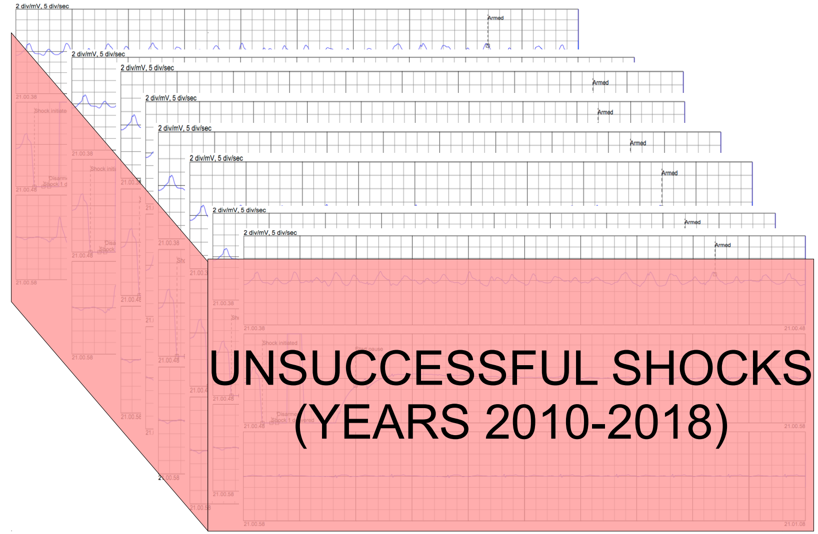


?

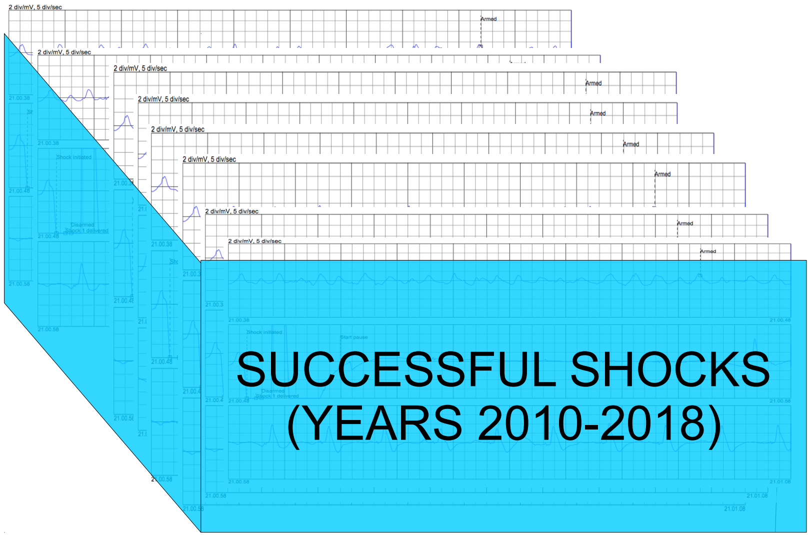


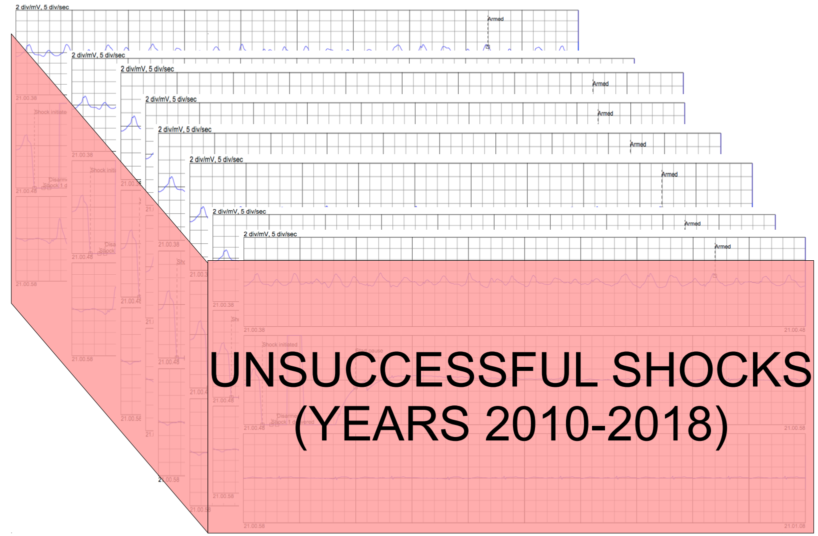
Learning from experience





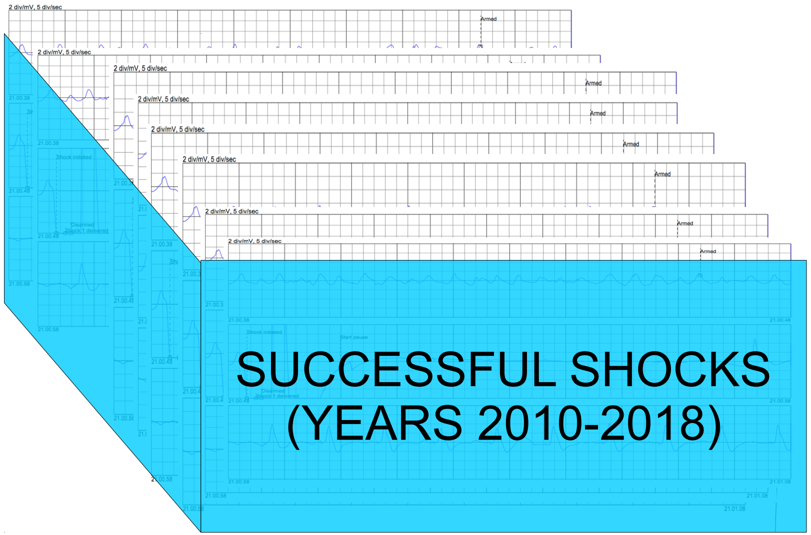
Learning from experience

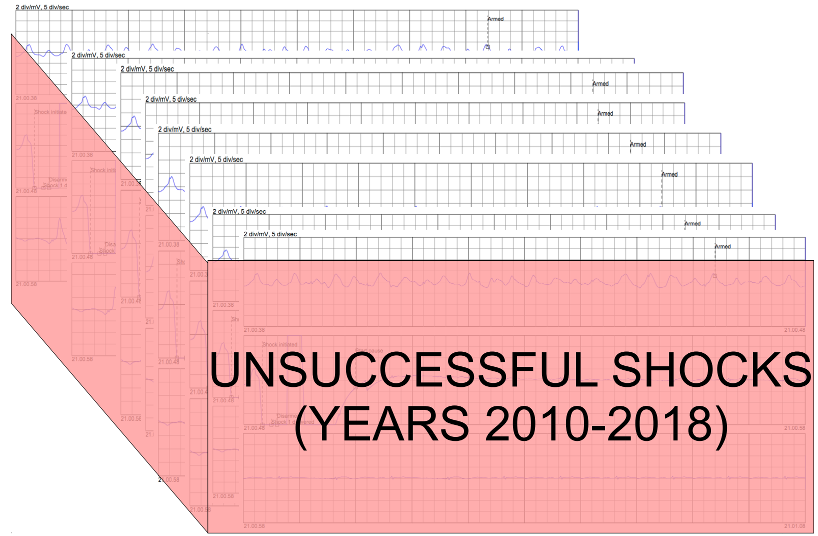




PAST

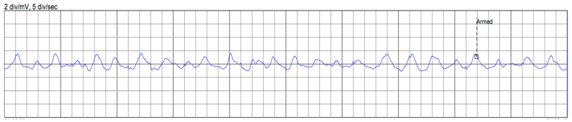
Learning from experience



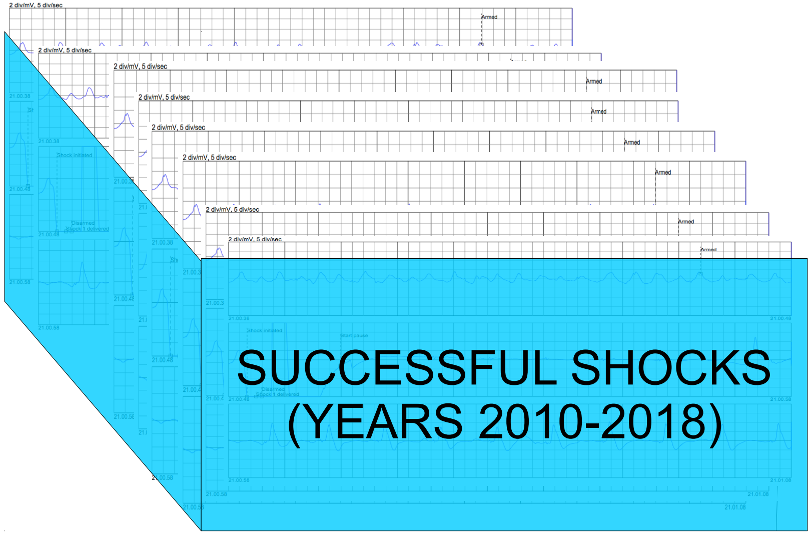


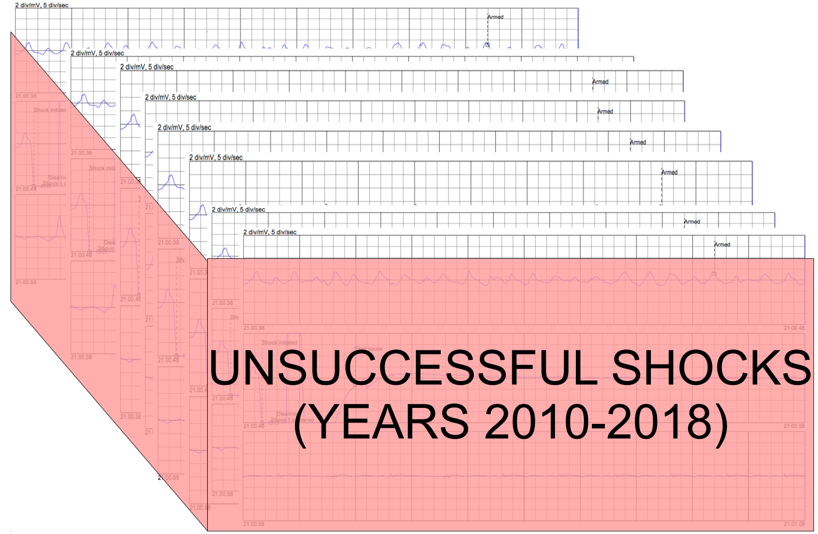
PAST

PRESENT



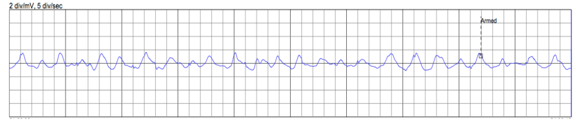
Learning from experience





PAST

PRESENT

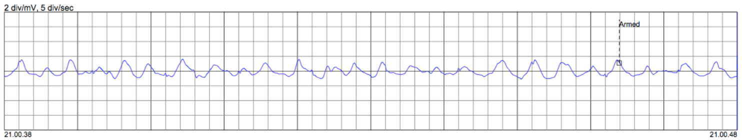


?

SUCCESS

FAILURE

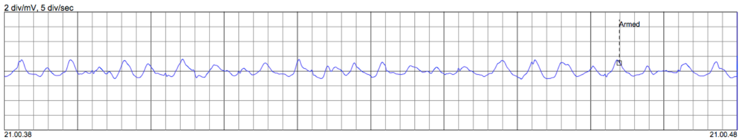
Feature extraction



- •PTT (peak to peak amplitude)
- •Amax (maximum amplitude)
- •Amin (minimum amplitude)
- •RMS (root mean square)

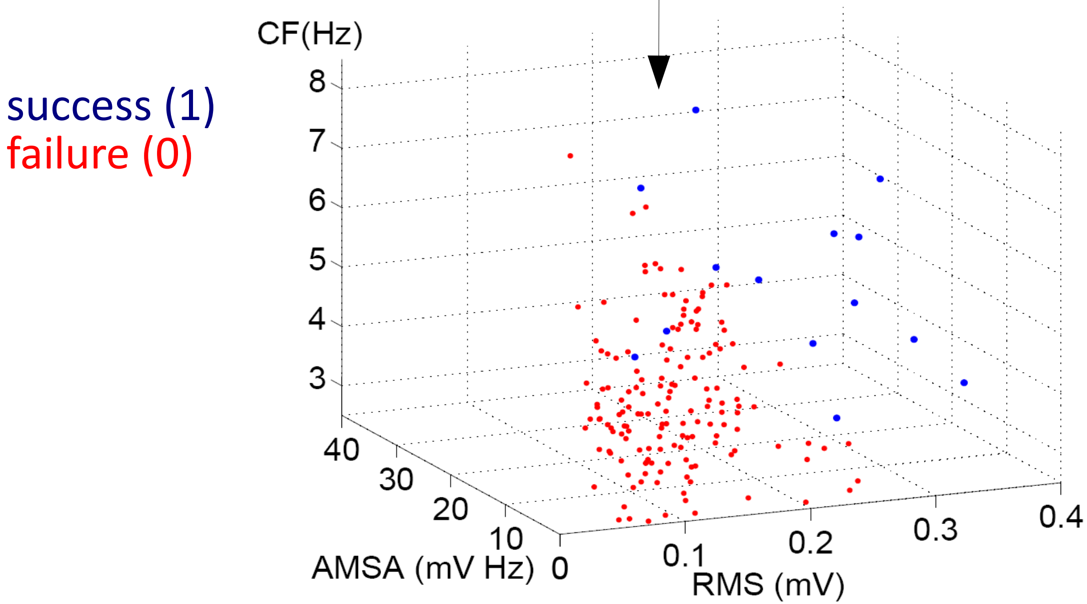
- •DF (dominant frequency)
- •CF (spectral centroid frequency)
- •EF (edge frequency)
- •AMSA (amplitude spectral area)

Feature extraction

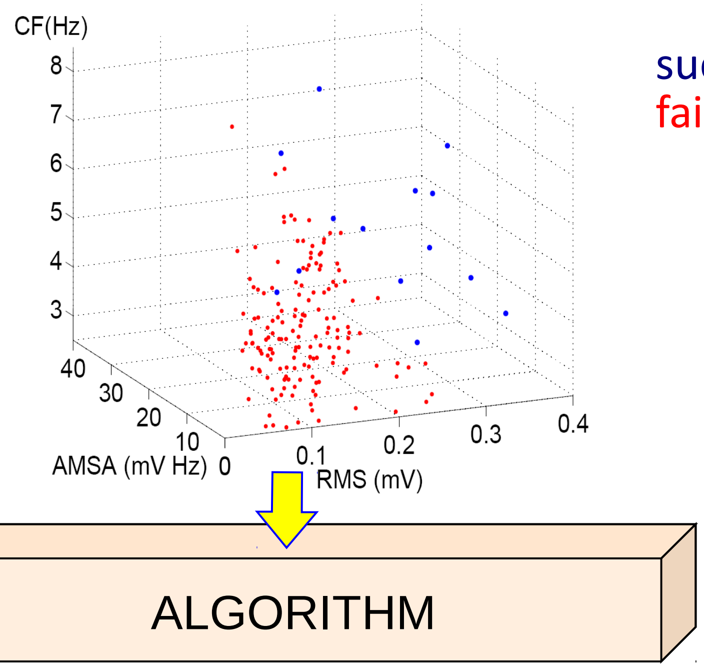


- PTT (peak to peak amplitude)
- Amax (maximum amplitude)
- Amin (minimum amplitude)
- **RMS** (root mean square)

- DF (dominant frequency)
- **CF** (spectral centroid frequency)
- EF (edge frequency)
- **AMSA** (amplitude spectral area)

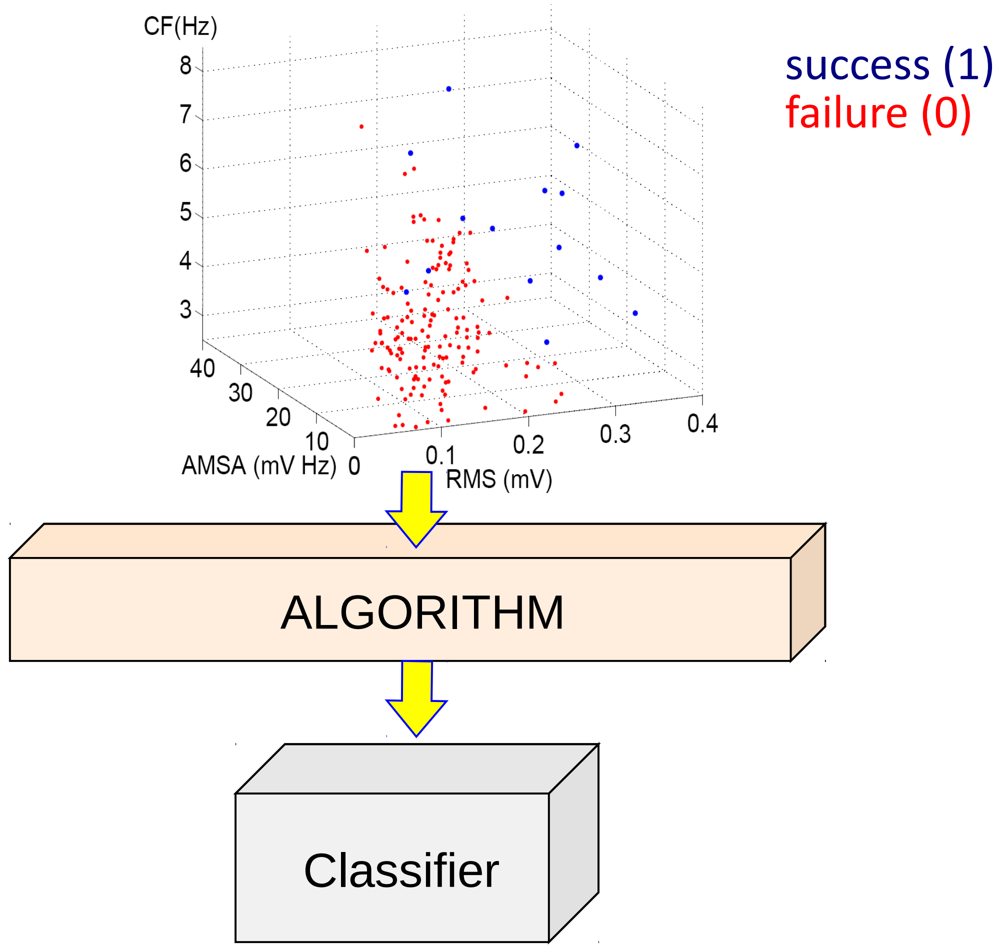


Automatic learning

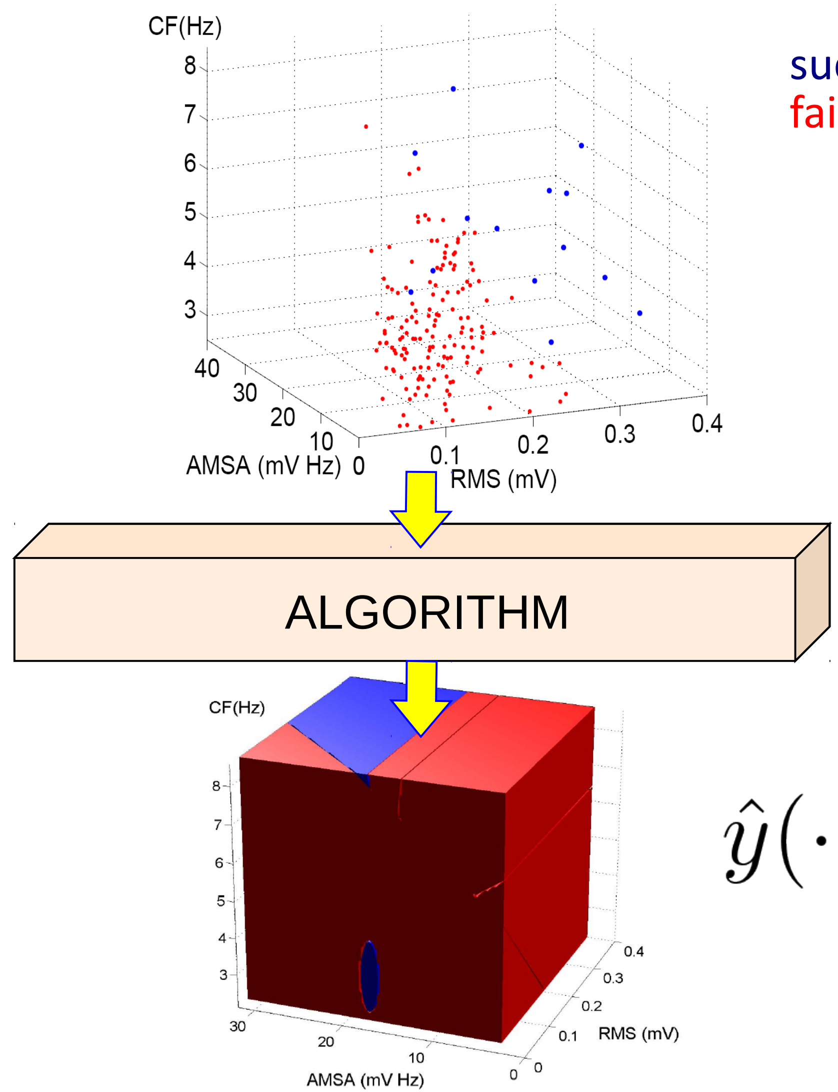


success (1)
failure (0)

Automatic learning



Automatic learning



success (1)
failure (0)

$$\hat{y}(\cdot) : \mathbb{R}^3 \to \{1, 0\}$$

Mathematical set-up

$\mathrm{x} \in \mathbb{R}^n$: patient ($n$ features)

$y \in \{1, 0\}$: true outcome

$\hat{y}(\cdot) : \mathbb{R}^n \to \{1, 0\}$ classifier

$(x, y) \sim \mathbb{P}$

$$\hat{y}(x) \neq y$$

Mathematical set-up

$\mathrm{x} \in \mathbb{R}^n$: patient ($n$ features)

$y \in \{1, 0\}$: true outcome

$\hat{y}(\cdot) : \mathbb{R}^n \to \{1, 0\}$ classifier

$$(x, y) \sim \mathbb{P}$$

$$PE(\hat{y}) := \mathbb{P}\{\hat{y}(x) \neq y\}$$

A fundamental distinction

Probability of **error**

A fundamental distinction

Probability of **error**

Pr("success" | failure)

Pr("failure" | success)

A fundamental distinction

Probability of **correct classification**

$1 - \Pr(\text{"success"} \mid \text{failure})$

$1 - \Pr(\text{"failure"} \mid \text{success})$

A fundamental distinction

Probability of **correct classification**

1 – Pr(”success” | failure)
=Pr(”failure” | failure)

1 – Pr(”failure” | success)
=Pr(”success” | success)

A fundamental distinction

Probability of **correct classification**

$1 - \Pr(\text{"success"} \mid \text{failure})$
$= \Pr(\text{"failure"} \mid \text{failure})$

specificity

$1 - \Pr(\text{"failure"} \mid \text{success})$
$= \Pr(\text{"success"} \mid \text{success})$

sensitivity

A fundamental distinction

Probability of **correct classification**

1 – Pr("success" | failure)
=Pr("failure" | failure)

specificity

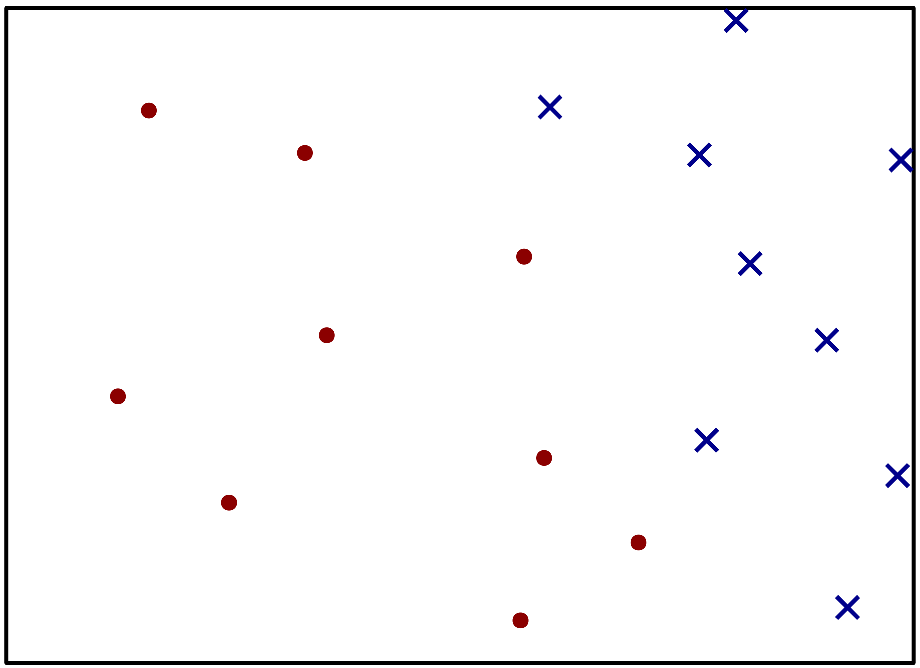
Target 50%

1 – Pr("failure" | success)
=Pr("success" | success)

sensitivity

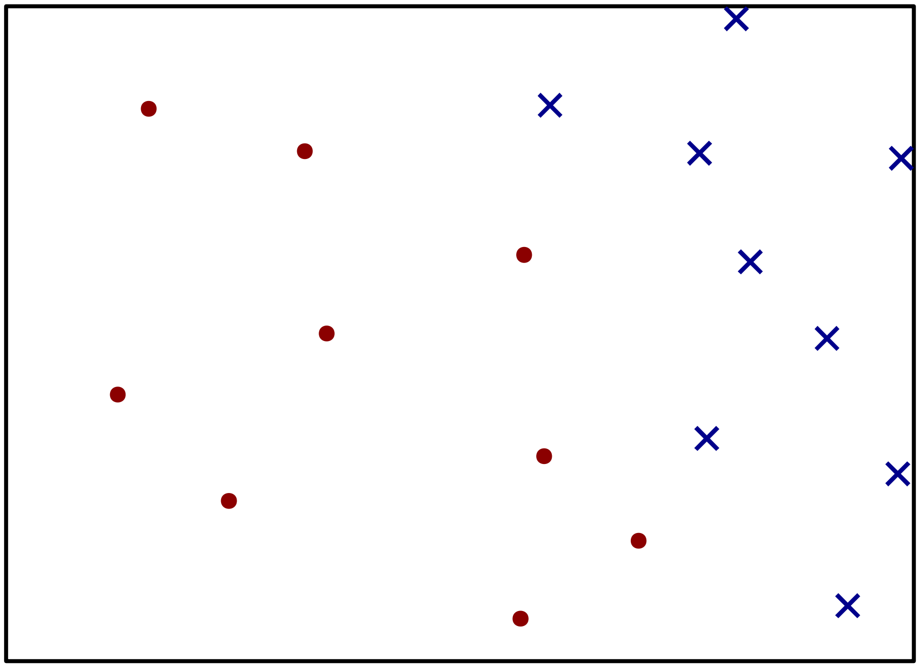
Target 95%

GEM-BALLS



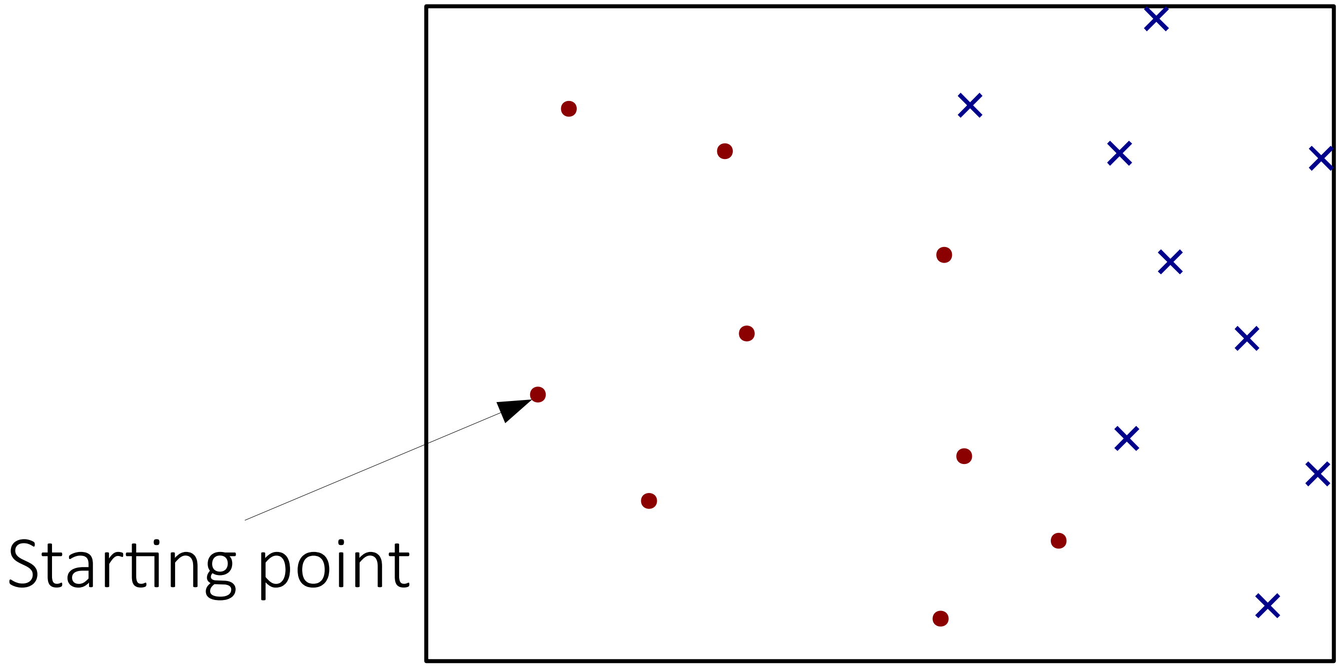
GEM-BALLS

- ● "0", "negative", "failure"
- × "1", "positive", "success"



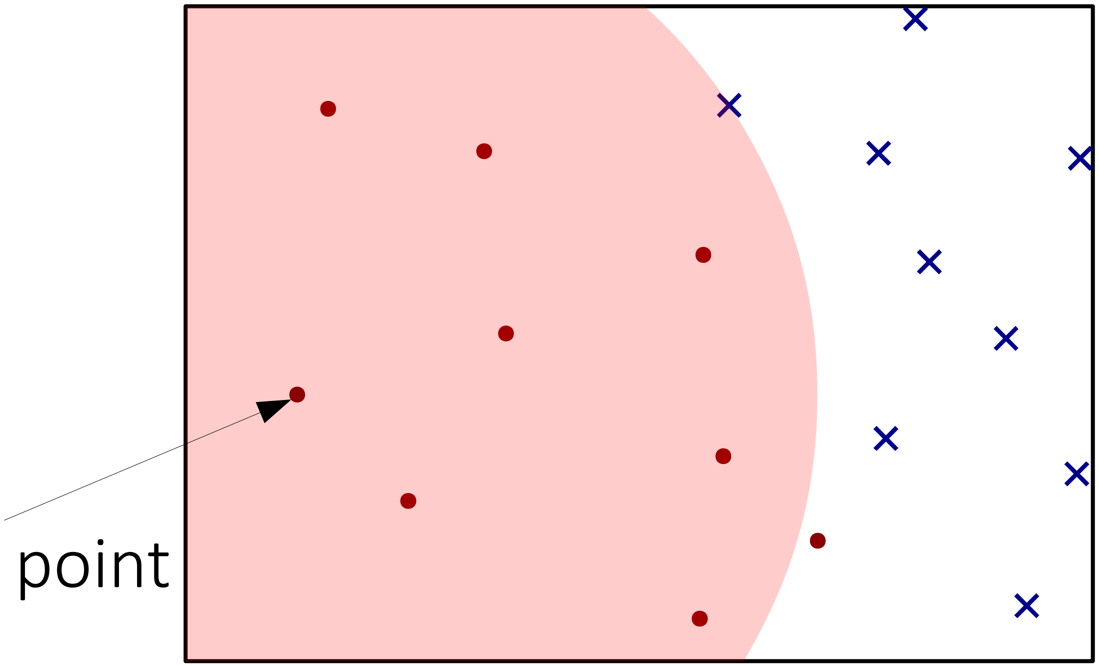
GEM-BALLS

- "0", "negative", "failure"
- × "1", "positive", "success"



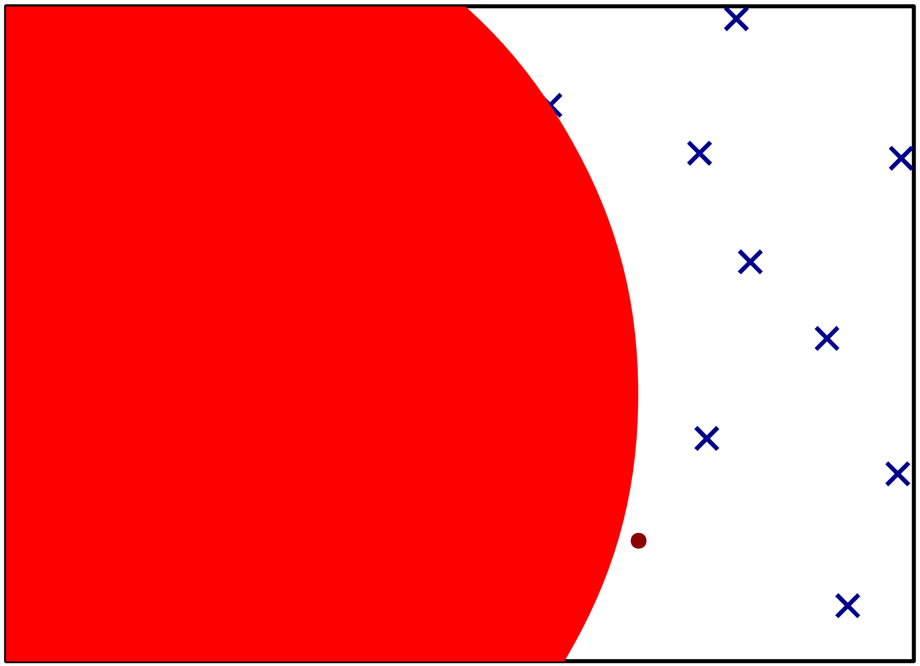
GEM-BALLS

Largest ball that does not include blue points



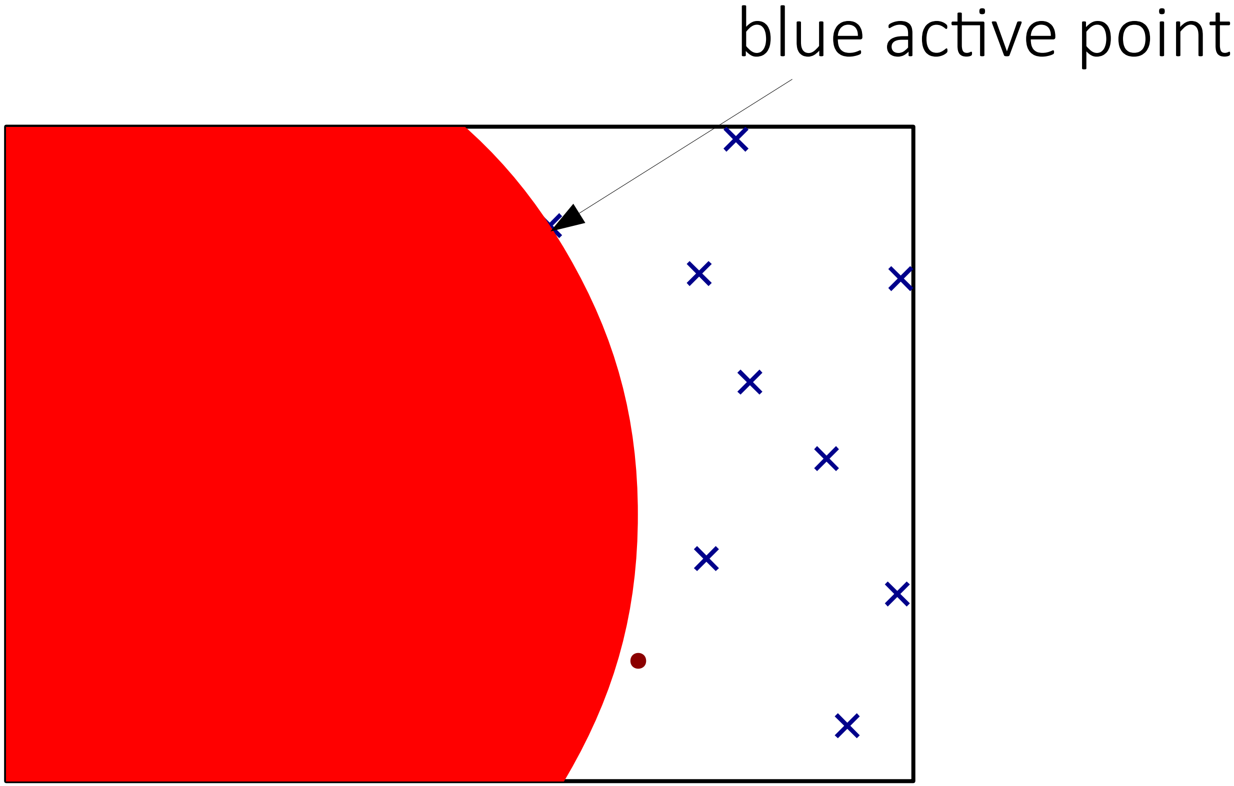
Starting point

GEM-BALLS



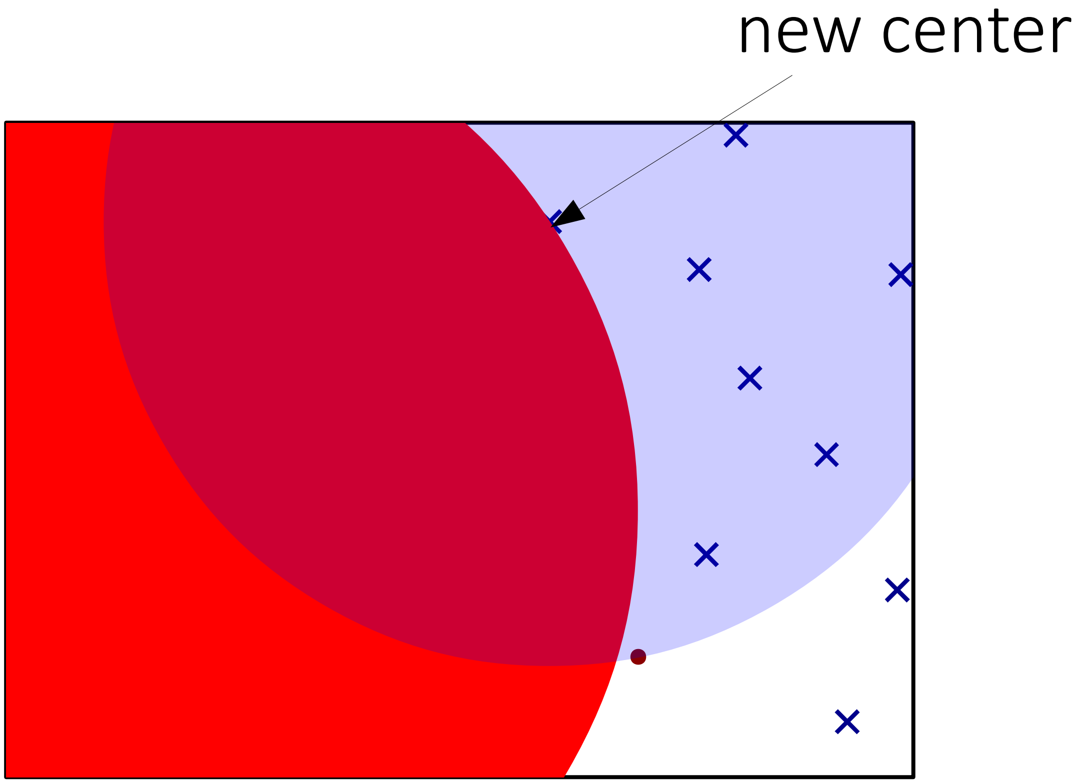
GEM-BALLS

#blue active points so far =1



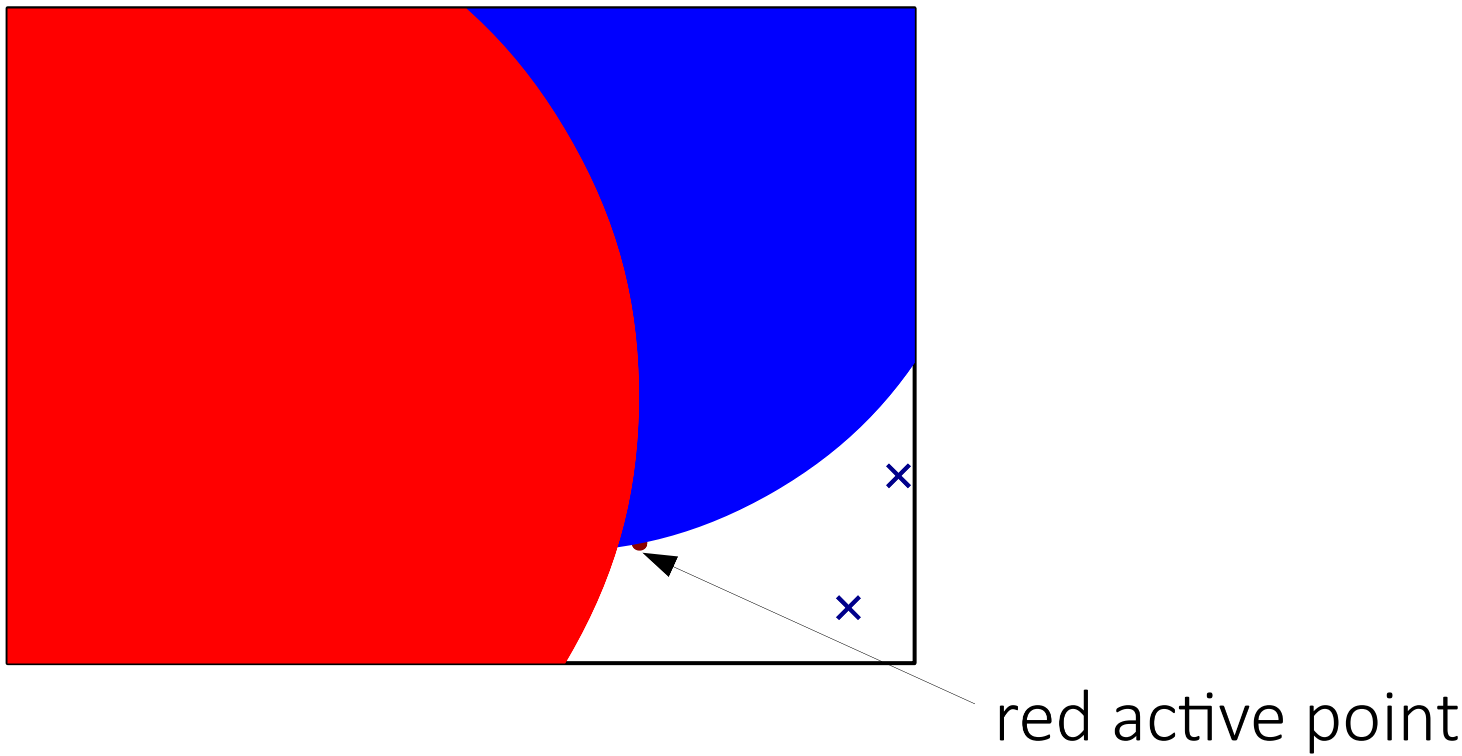
GEM-BALLS

#blue active points so far =1



GEM-BALLS

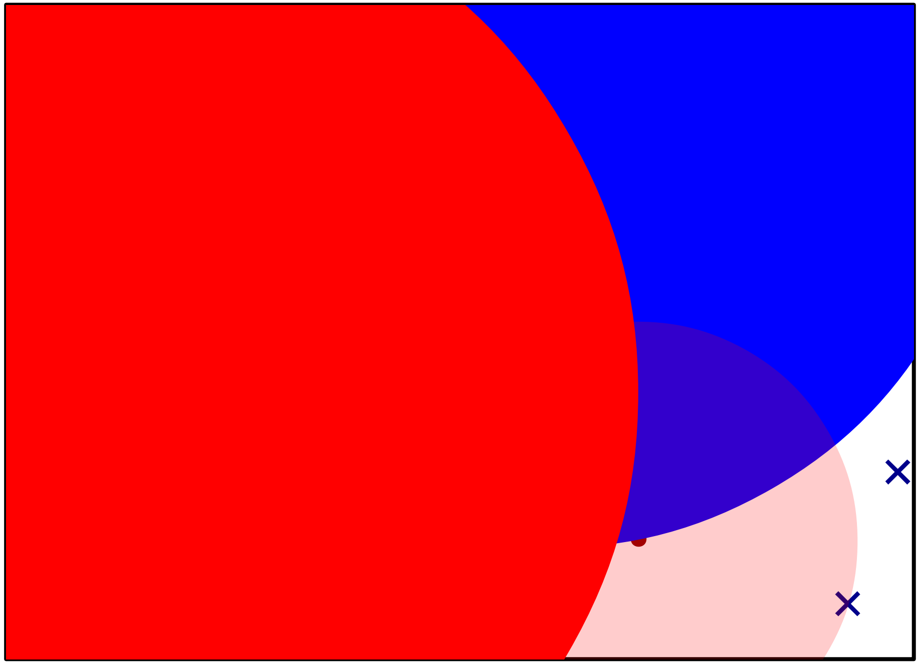
#blue active points so far =1
#red active points so far =1



GEM-BALLS

#blue active points so far =1

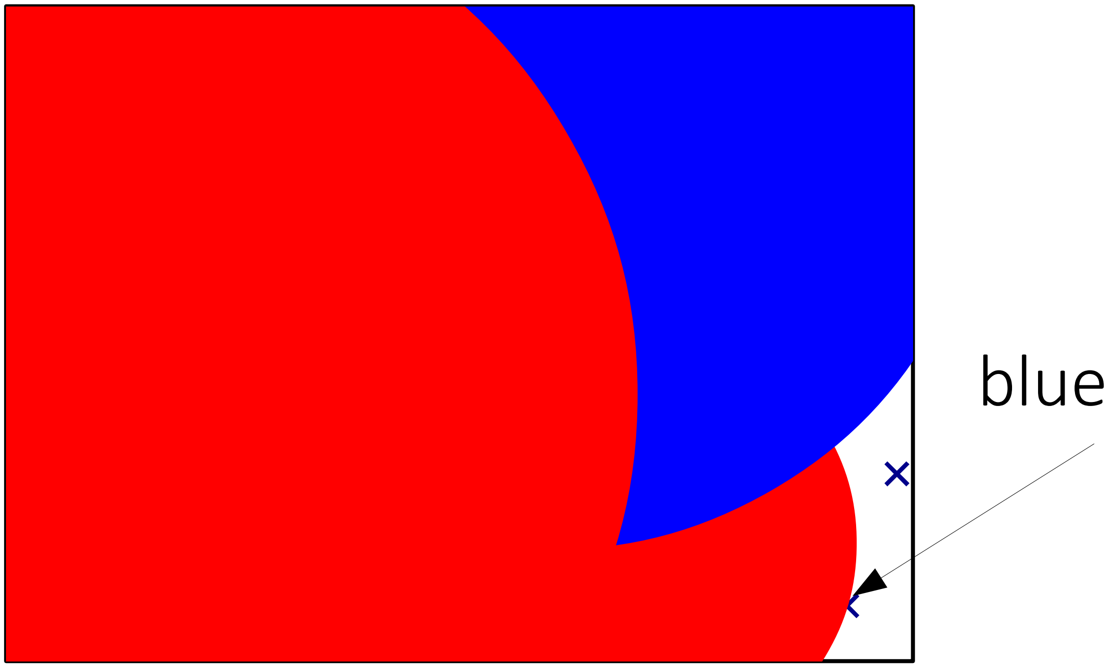
#red active points so far =1



GEM-BALLS

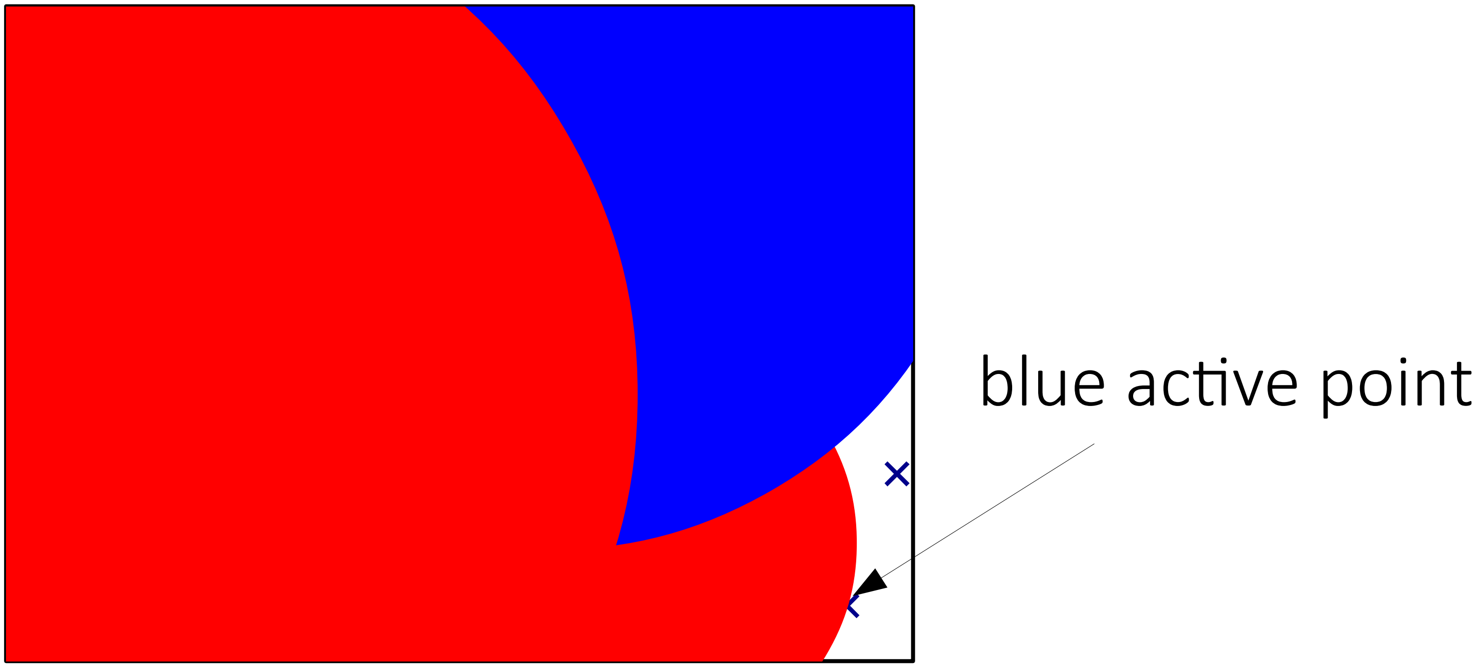
#blue active points so far =1+1

#red active points so far =1



GEM-BALLS

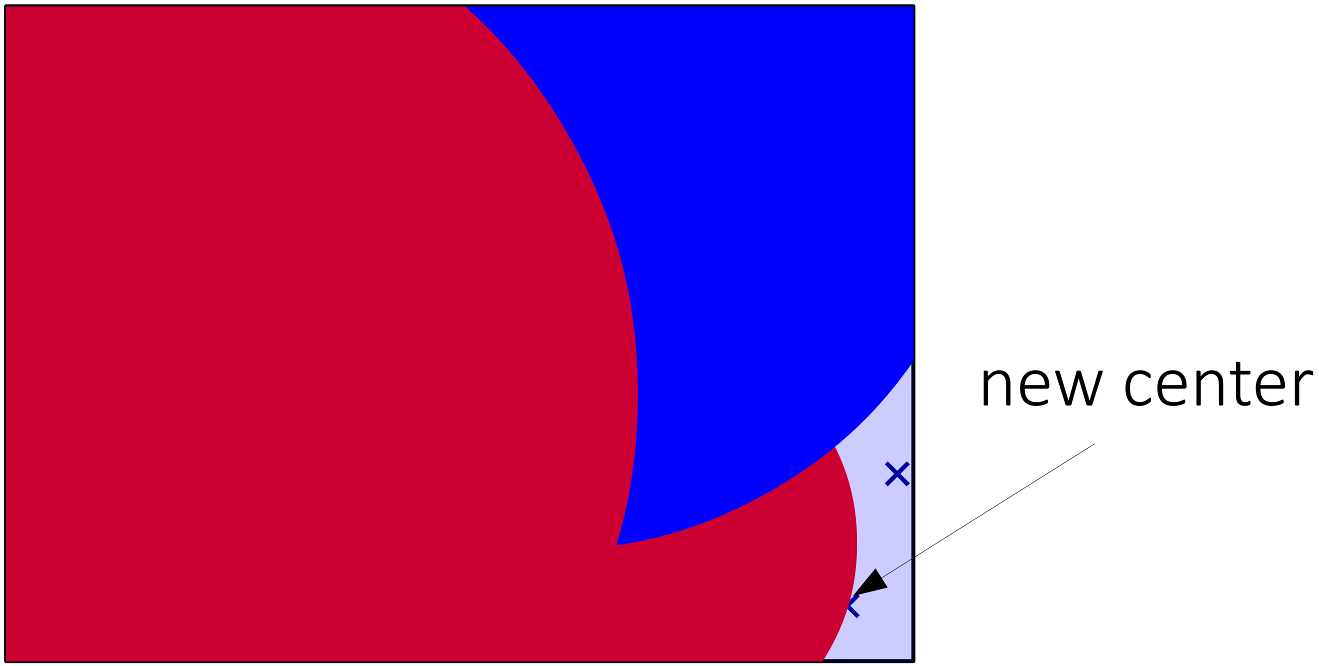
#blue active points so far =2
#red active points so far =1



GEM-BALLS

#blue active points so far =2

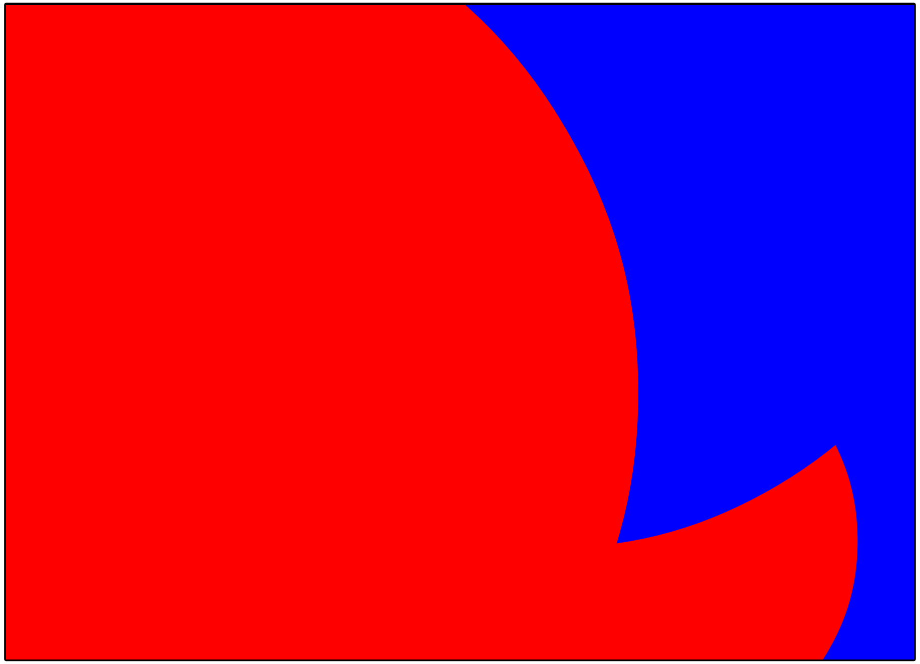
#red active points so far =1



GEM-BALLS

#blue active points so far =2

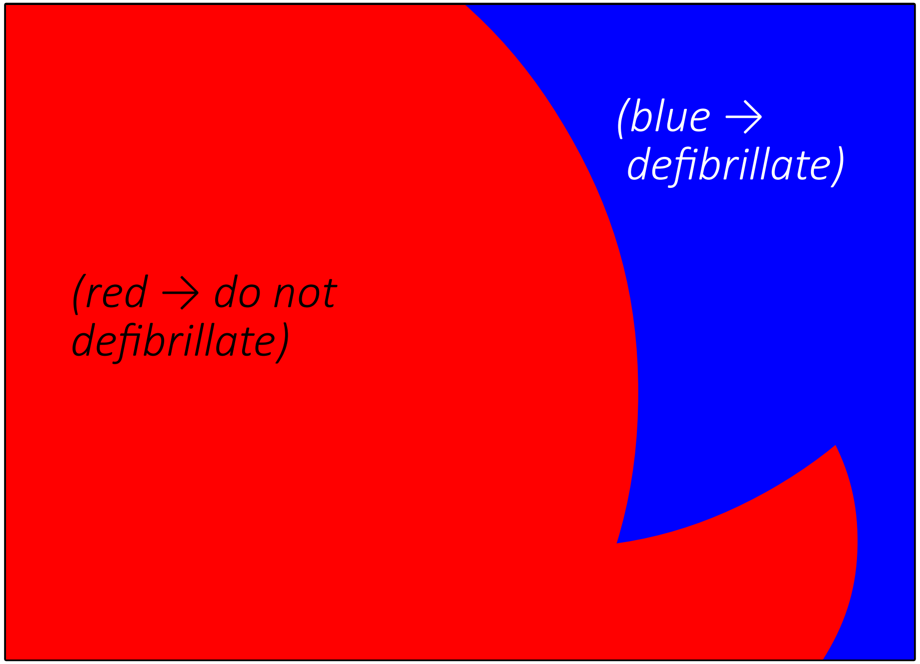
#red active points so far =1



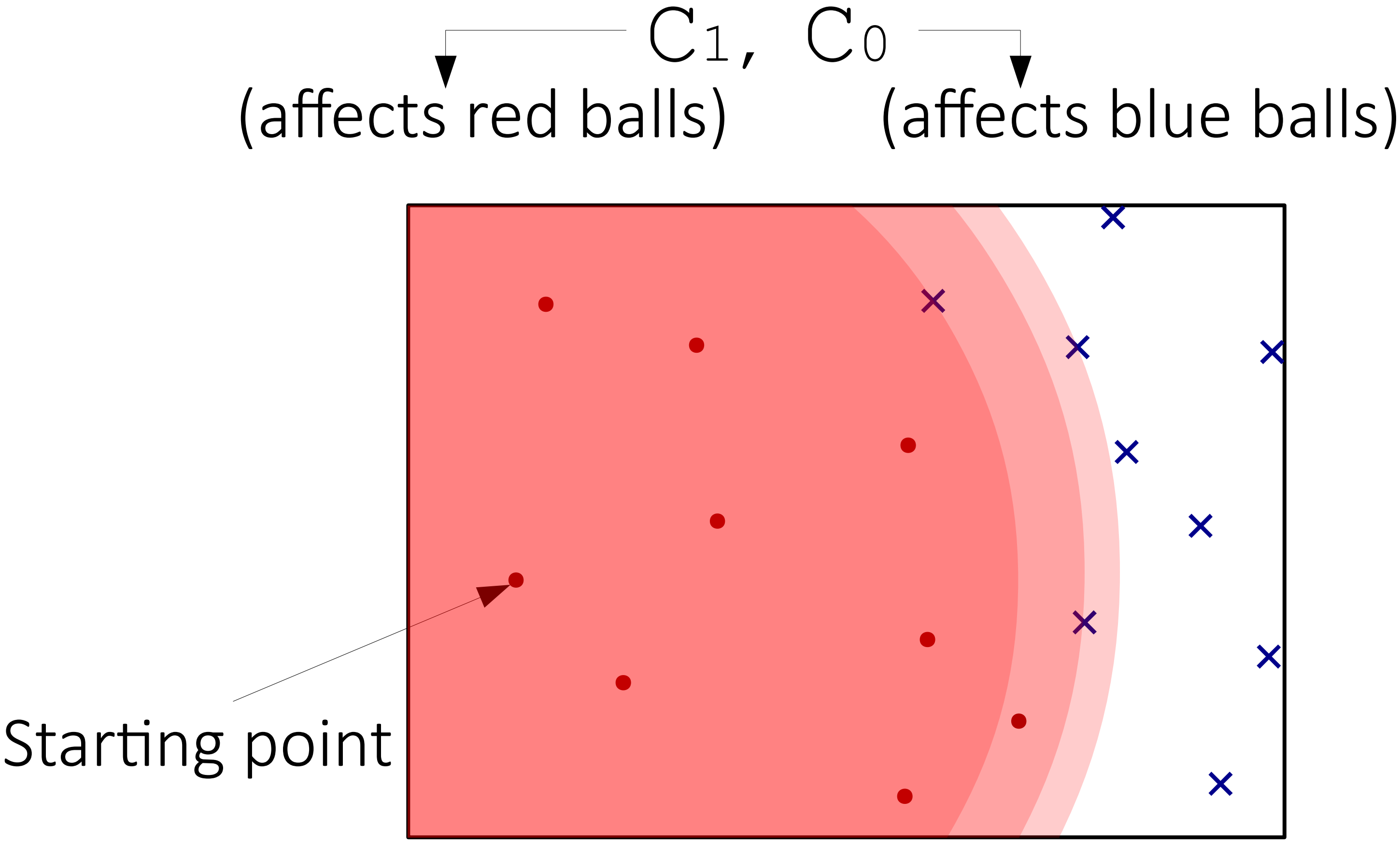
GEM-BALLS

#blue active points so far =2 → k_1

#red active points so far =1 → k_0



The tuning knob



Theorem - preliminaries

β: small (e.g. 10^{-6}) confidence parameter

Theorem - preliminaries

β: small (e.g. 10^{-6}) confidence parameter

N_1: # of positives

Theorem - preliminaries

β: small (e.g. 10^{-6}) confidence parameter

N_1: # of positives

k_1 : # active positives

Theorem - preliminaries

β: small (e.g. 10^{-6}) confidence parameter

N_1: # of positives

k_1 : # active positives

$\{\epsilon_1(d) : d = 1, \ldots, N_1\}$: are pre-computed thresholds.

Theorem - preliminaries

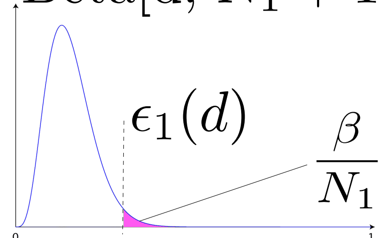
β: small (e.g. 10^{-6}) confidence parameter

N_1: # of positives

k_1 : # active positives

$\{\epsilon_1(d) : d = 1, \ldots, N_1\}$: are pre-computed thresholds.

$\epsilon_1(d)$, for $d = 1, \ldots, N_1$ is the quantile at level $1 - \frac{\beta}{N_1}$ of $\mathrm{Beta}[\mathrm{d}, \mathrm{N}_1 + 1 - d]$



Theorem - preliminaries

$\hat{y}$ (k_1)

Theorem - preliminaries

$$\mathbb{P}\{\underbrace{\hat{y}(x) = 0}_{\text{"failure"}} \mid \underbrace{y = 1}_{\text{success}}\} \qquad (k_1)$$

Theorem - preliminaries

$$\mathbb{P}\{\overbrace{\hat{y}(x) = 0}^{\text{“failure”}} \mid \overbrace{y = 1}^{\text{success}}\} \qquad \underbrace{\epsilon_1(k_1)}_{\text{precomputed threshold}}$$

Theorem - preliminaries

"failure" success

$$\underbrace{\mathbb{P}\{\hat{y}(x) = 0}_{\text{“failure”}} \mid \underbrace{y = 1}_{\text{success}}\} \leq \underbrace{\epsilon_1(k_1)}_{\text{precomputed threshold}}$$

precomputed threshold

Theorem - Statement

“failure” success

$$\mathbb{P}\{\underbrace{\hat{y}(x) = 0}_{\text{“failure”}} \mid \underbrace{y = 1}_{\text{success}}\} \leq \underbrace{\epsilon_1(k_1)}_{\text{precomputed threshold}}$$

precomputed threshold

$$\text{with confidence } 1 - \beta$$

Theorem - Discussion

"failure" success

$$\mathbb{P}\{\hat{y}(x) = 0 \mid y = 1\} \leq \epsilon_1(k_1)$$

precomputed threshold

↕

$$\text{sensitivity} \geq 1 - \epsilon_1(k_1)$$

$$\text{with confidence } 1 - \beta$$

Theorem - Discussion

$$\overbrace{\mathbb{P}\{\hat{y}(x) = 0}^{\text{“failure”}} \mid \overbrace{y = 1}^{\text{success}}\} \leq \underbrace{\epsilon_1(k_1)}_{\text{precomputed threshold}}$$

⇕

$$\text{sensitivity} \geq 1 - \epsilon_1(k_1)$$

$$\text{with confidence } 1 - \beta$$

Theorem - Statement

$$\text{sensitivity} \geq 1 - \epsilon_1(k_1)$$

$$\text{with confidence } 1 - \beta$$

Theorem - Discussion

$$\text{sensitivity} \geq 1 - \epsilon_1(k_1)$$

$$\text{with confidence } 1 - \beta$$

Theorem - Discussion

$$(x_i, y_i) \sim \mathbb{P}$$

$$\text{sensitivity} \geq 1 - \epsilon_1(k_1)$$

Theorem - Statement

$$(x_i, y_i) \sim \mathbb{P}$$

$$\mathbb{P}^N\{\text{sensitivity} \geq 1 - \epsilon_1(k_1)\} \geq 1 - \beta$$

Theorem - Statement

$$(x_i, y_i) \sim \mathbb{P}$$

$$\mathbb{P}^N\{\text{sensitivity} \geq 1 - \epsilon_1(k_1)\} \geq 1 - \beta$$

$$\mathbb{P}^N\{\text{specificity} \geq 1 - \epsilon_0(k_0)\} \geq 1 - \beta$$

Theorem - Statement

$$\text{sensitivity} \geq 1 - \epsilon_1(k_1) \;\&\; \text{specificity} \geq 1 - \epsilon_0(k_0)$$

$$\text{with overall confidence } 1 - 2\beta$$

Theorem - Statement

$$\text{sensitivity} \geq 1 - \epsilon_1(k_1) \;\&\; \text{specificity} \geq 1 - \epsilon_0(k_0)$$

$$\text{with overall confidence } 1 - 2\beta$$

Take-home message:

“GEM-BALLS is a self-testing algorithm!”

Benchmark datasets

BreastW (239 positive instances, 444 negative instances), $\beta = 5 \cdot 10^{-3}$							
$c_1 : c_0$	$1:1$	$1:2$	$1:3$	$1:5$	$1:10$	$10:100$	$1:50$
$\mathbf{k}_1 : \mathbf{k}_0$	$17:17$	$11:21$	$10:28$	$6:28$	$4:33$	$10:17$	$2:66$
$Sens:Spec$	83% : 91%	87% : 89%	88% : 87%	90% : 87%	92% : 86%	88% : 91%	94% : 76%

Pima (268 positive instances, 500 negative instances), $\beta = 10^{-3}$							
$c_1 : c_0$	$1:1$	$1:2$	$2:4$	$1:4$	$1:8$	$1:10$	$1:50$
$\mathbf{k}_1 : \mathbf{k}_0$	$125:125$	$88:175$	$96:189$	$61:245$	$38:300$	$33:324$	$9:424$
$Sens:Spec$	40% : 65%	54% : 55%	51% : 52%	65% : 41%	75% : 30%	70% : 20%	90% : 9%

Haberman (75 positive instances, 219 negative instances), $\beta = 10^{-3}$					
$c_1 : c_0$	$1:1$	$1:3$	$1:5$	$1:10$	$1:20$
$\mathbf{k}_1 : \mathbf{k}_0$	$46:46$	$28:82$	$23:114$	$14:139$	$9:176$
$Sens:Spec$	20% : 67%	41% : 49%	48% : 35%	62% : 24%	71% : 10%

Benchmark datasets

BreastW (239 positive instances, 444 negative instances), $\beta = 5 \cdot 10^{-3}$							
$c_1 : c_0$	1 : 1	1 : 2	1 : 3	1 : 5	1 : 10	10 : 100	1 : 50
$\mathbf{k}_1 : \mathbf{k}_0$	17 : 17	11 : 21	10 : 28	6 : 28	4 : 33	10 : 17	2 : 66
Sens:*Spec*	83% : 91%	87% : 89%	88% : 87%	90% : 87%	92% : 86%	88% : 91%	94% : 76%

Pima (268 positive instances, 500 negative instances), $\beta = 10^{-3}$							
$c_1 : c_0$	1 : 1	1 : 2	2 : 4	1 : 4	1 : 8	1 : 10	1 : 50
$\mathbf{k}_1 : \mathbf{k}_0$	125 : 125	88 : 175	96 : 189	61 : 245	38 : 300	33 : 324	9 : 424
Sens:*Spec*	40% : 65%	54% : 55%	51% : 52%	65% : 41%	75% : 30%	70% : 20%	90% : 9%

Haberman (75 positive instances, 219 negative instances), $\beta = 10^{-3}$					
$c_1 : c_0$	1 : 1	1 : 3	1 : 5	1 : 10	1 : 20
$\mathbf{k}_1 : \mathbf{k}_0$	46 : 46	28 : 82	23 : 114	14 : 139	9 : 176
Sens:*Spec*	20% : 67%	41% : 49%	48% : 35%	62% : 24%	71% : 10%

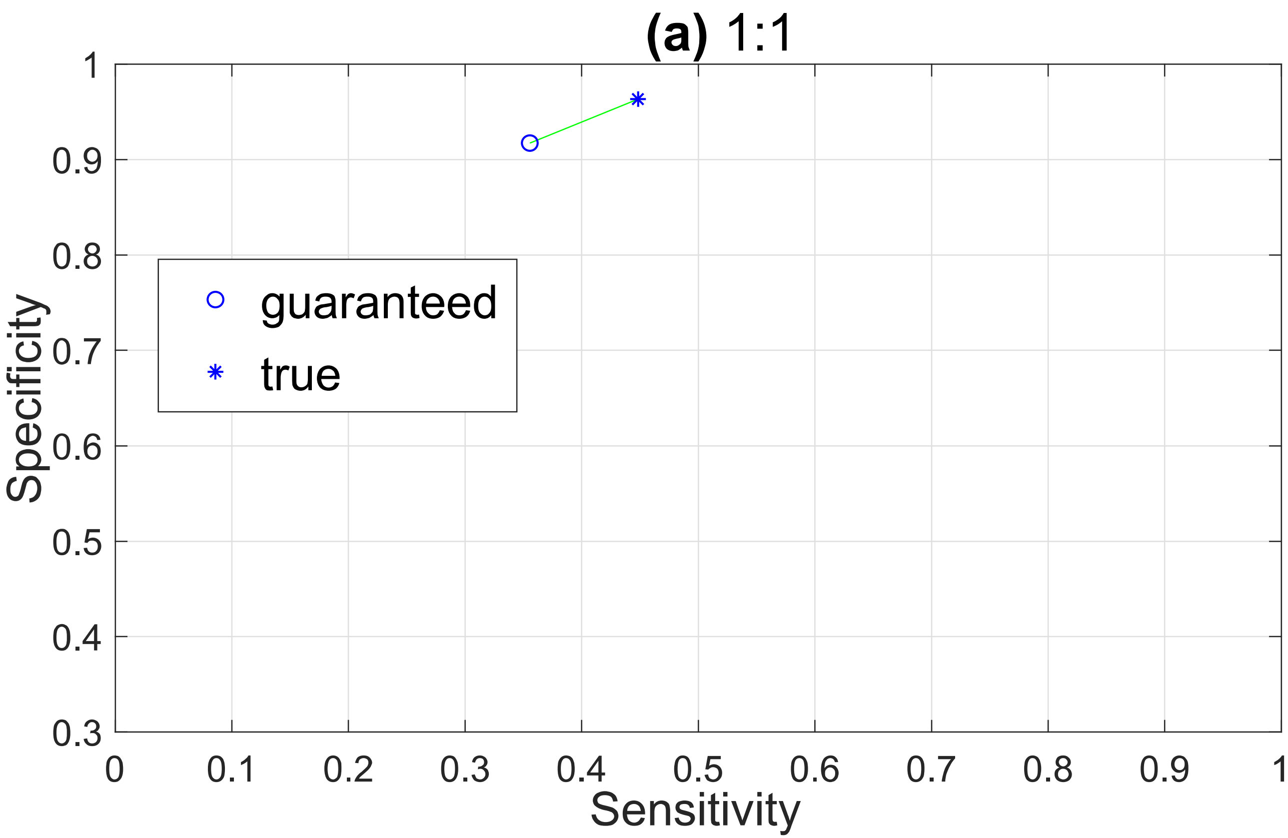
Benchmark datasets

BreastW (239 positive instances, 444 negative instances), $\beta = 5 \cdot 10^{-3}$							
$c_1 : c_0$	$1:1$	$1:2$	$1:3$	$1:5$	$1:10$	$10:100$	$1:50$
$\mathbf{k}_1 : \mathbf{k}_0$	$17:17$	$11:21$	$10:28$	$6:28$	$4:33$	$10:17$	$2:66$
$Sens$:$Spec$	83% : 91%	87% : 89%	88% : 87%	90% : 87%	92% : 86%	88% : 91%	94% : 76%

Pima (268 positive instances, 500 negative instances), $\beta = 10^{-3}$							
$c_1 : c_0$	$1:1$	$1:2$	$2:4$	$1:4$	$1:8$	$1:10$	$1:50$
$\mathbf{k}_1 : \mathbf{k}_0$	$125:125$	$88:175$	$96:189$	$61:245$	$38:300$	$33:324$	$9:424$
$Sens$:$Spec$	40% : 65%	54% : 55%	51% : 52%	65% : 41%	75% : 30%	70% : 20%	90% : 9%

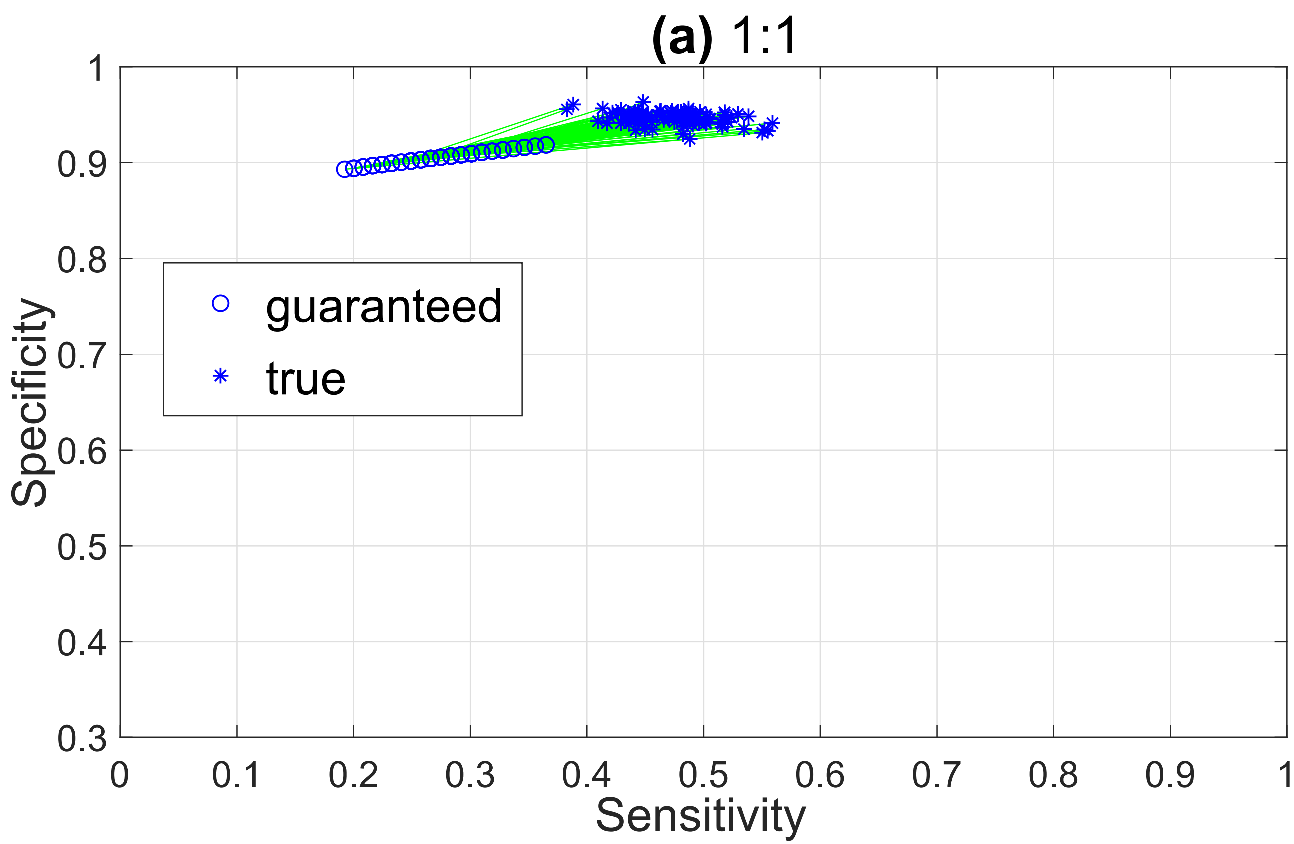
Haberman (75 positive instances, 219 negative instances), $\beta = 10^{-3}$					
$c_1 : c_0$	$1:1$	$1:3$	$1:5$	$1:10$	$1:20$
$\mathbf{k}_1 : \mathbf{k}_0$	$46:46$	$28:82$	$23:114$	$14:139$	$9:176$
$Sens$:$Spec$	20% : 67%	41% : 49%	48% : 35%	62% : 24%	71% : 10%

Tests on reproducible simulated data



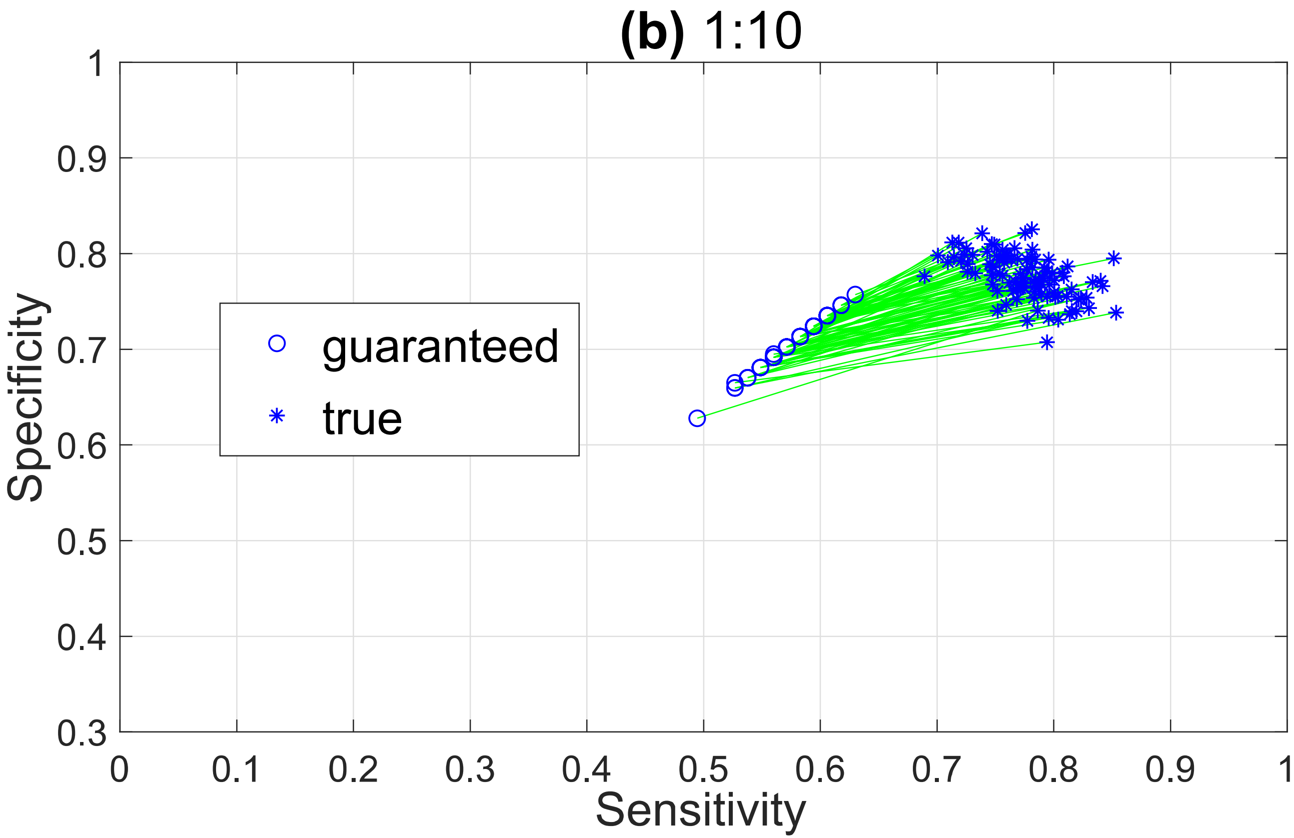
Labels are generated by using the MATLAB kstest function.
$N_0 = 1000$, $N_1 = 100$. $\beta = 10^{-3}$

Tests on reproducible simulated data



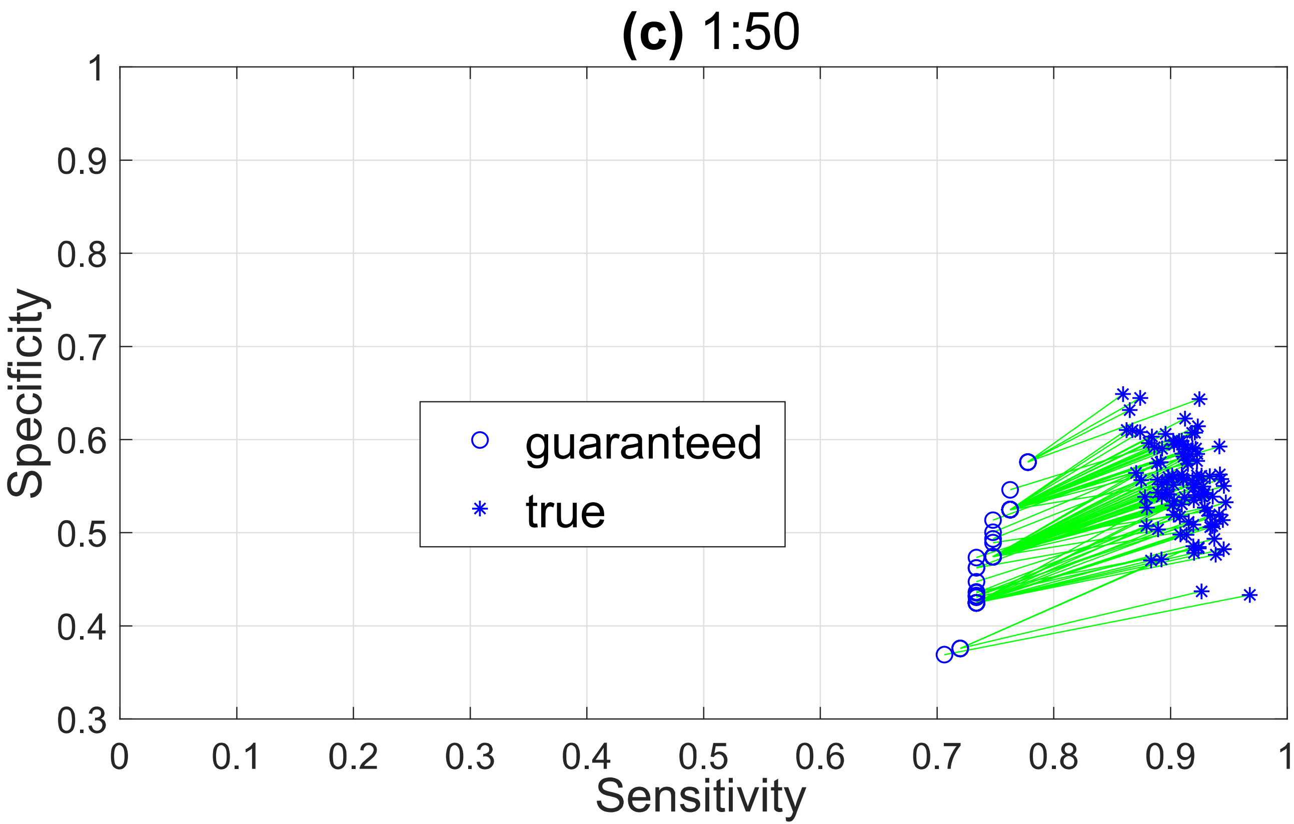
Labels are generated by using the MATLAB kstest function.
$N_0 = 1000,\ N_1 = 100.\ \beta = 10^{-3}$

Tests on reproducible simulated data



Labels are generated by using the MATLAB kstest function.
$N_0 = 1000,\ N_1 = 100.\ \beta = 10^{-3}$

Tests on reproducible simulated data



Labels are generated by using the MATLAB kstest function.
$N_0 = 1000,\ N_1 = 100.\ \beta = 10^{-3}$

Ventricular Fibrillation (VF) dataset

VF dataset (15 pos, 155 neg), $\beta = 10^{-2}$			
$c_1 : c_0$	$1:1$	$1:10$	$1:80$
$\mathbf{k}_1 : \mathbf{k}_0$	$9:9$	$5:41$	$2:90$
$Sens : Spec$	$11\% : 85\%$	$30\% : 59\%$	$51\% : 28\%$

Ventricular Fibrillation (VF) dataset

VF dataset (15 pos, 155 neg), $\beta = 10^{-2}$			
$c_1 : c_0$	$1:1$	$1:10$	$1:80$
$\mathbf{k}_1 : \mathbf{k}_0$	$9:9$	$5:41$	$2:90$
$Sens:Spec$	$11\% : 85\%$	$30\% : 59\%$	$51\% : 28\%$

Expanded VF dataset (240 pos, 2477 neg), $\beta = 10^{-3}$				
$c_1 : c_0$	$1:1$	$1:10$	$1:80$	$1:240$
$\mathbf{k}_1 : \mathbf{k}_0$	$16:16$	$1:121$	$8:568$	$4:1055$
$Sens:Spec$	$84\% : 98\%$	$86\% : 93\%$	$89\% : 73\%$	$92\% : 53\%$

SMOTE (Synthetic Minority Over-sampling Technique)

Thank you!

Are you interested in knowing more?
Write to us!

algo.care@unibs.it

Try GEM-BALLS now!
www.algocare.it/GEM-BALLS/

Bibliography

GEM-BALLS:

- “A new classification algorithm with guaranteed sensitivity and specificity for medical applications,”
  A. Carè, F.A. Ramponi, M.C. Campi
  IEEE Control Systems Letters. 2(3):393-398 [doi: 10.1109/LCSYS.2018.2840427], 2018
  www.algocare.it/GEM-BALLS/

GEM:

- “Classification with Guaranteed Probability of Error”
  M.C. Campi
  Machine Learning, 80:63-84, 2010.

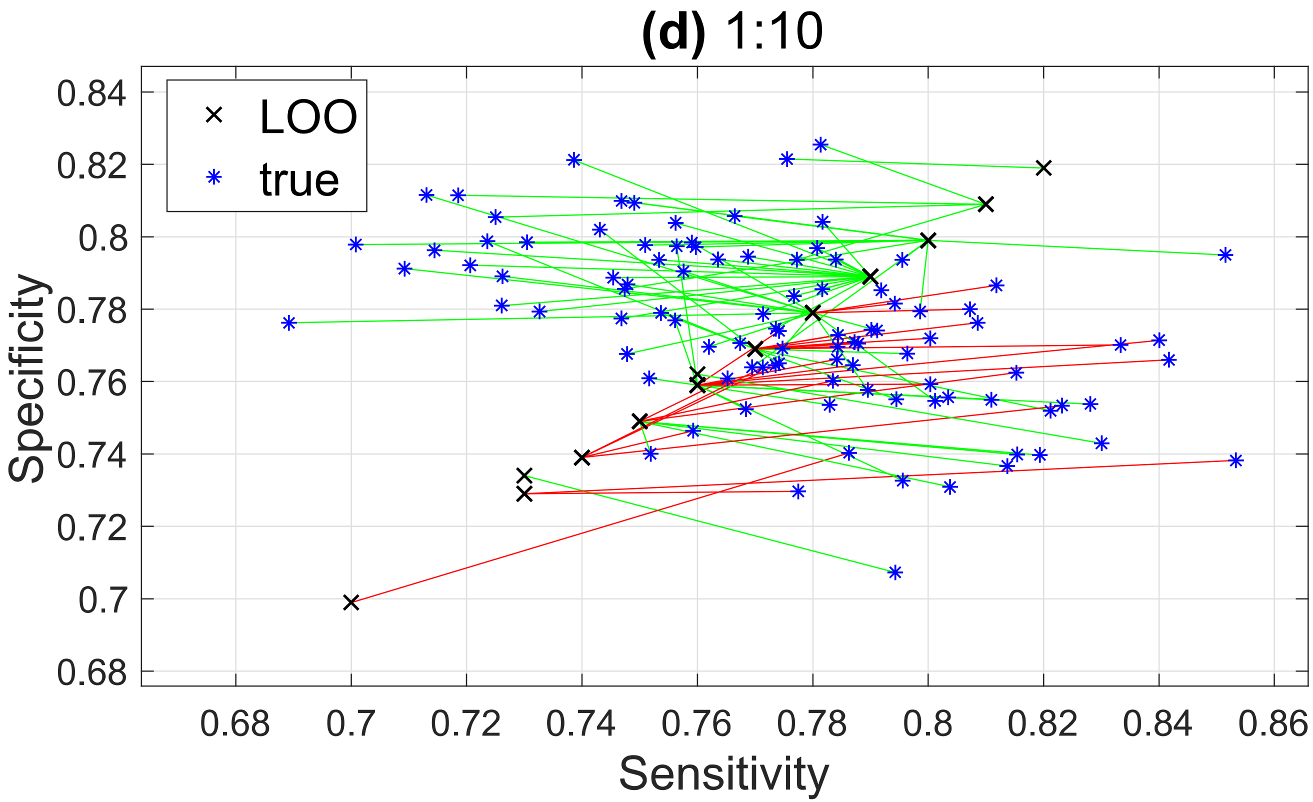
- ”Consensus and Reliability: The Case of Two Binary Classifiers”
  A.T.J.R. Cobbenhagen. A. Carè, M.C. Campi, F.A. Ramponi, W.P.M.H. Heemels
  8th IFAC Workshop on Distributed Estimation and Control in Networked Systems
  Sept. 16-17, 2019, Chicago, IL, USA

- “Ventricular defibrillation: Classification with GEM and a roadmap for future investigations”
  F. Baronio, M. Baronio, M.C. Campi, A. Carè, S. Garatti, G. Perone
  2017 IEEE 56th Annual Conference on Decision and Control (CDC), 2718-2723

Tests on reproducible simulated data

The problem is that of predicting the output y of the binary function
kstest($[x^{(1)}, \ldots, x^{(7)}]$,'Alpha',0.005)
in the MATLAB Statistics and Machine Learning Toolbox,
when the feature vector $x = [x^{(1)}, \ldots, x^{(7)}]$ is
uniformly and independently sampled over $[0, 1]^7$.

Tests on reproducible simulated data



Labels are generated by using the MATLAB kstest function.
$N_0 = 1000,\ N_1 = 100.$

Leave-one-out estimates

$$1 - \frac{k_1}{N_1}$$

is a sample-based estimate of sensitivity (leave-one-out estimate);

$$1 - \frac{k_0}{N_0}$$

is a sample-based estimate of specificity (leave-one-out estimate).

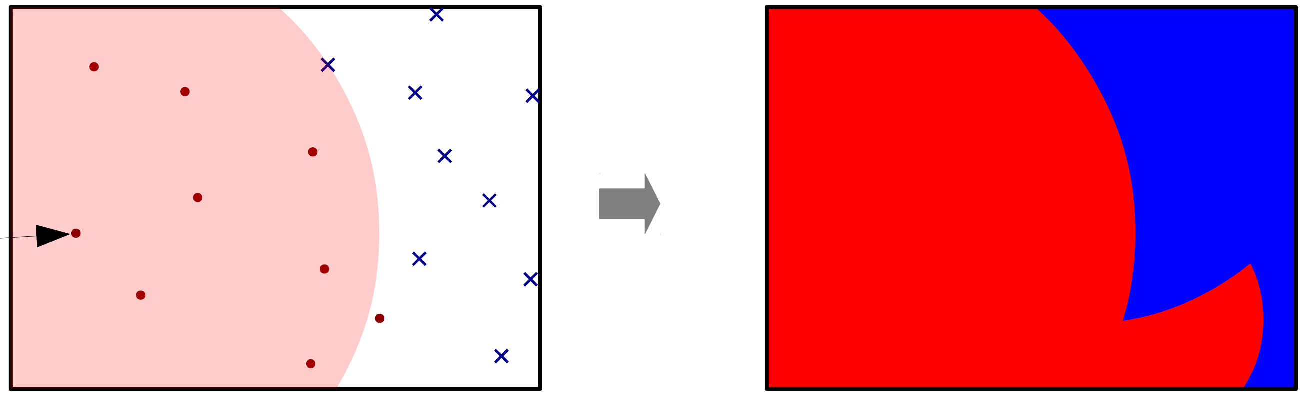
New lines of research

Pool of guaranteed classifiers:

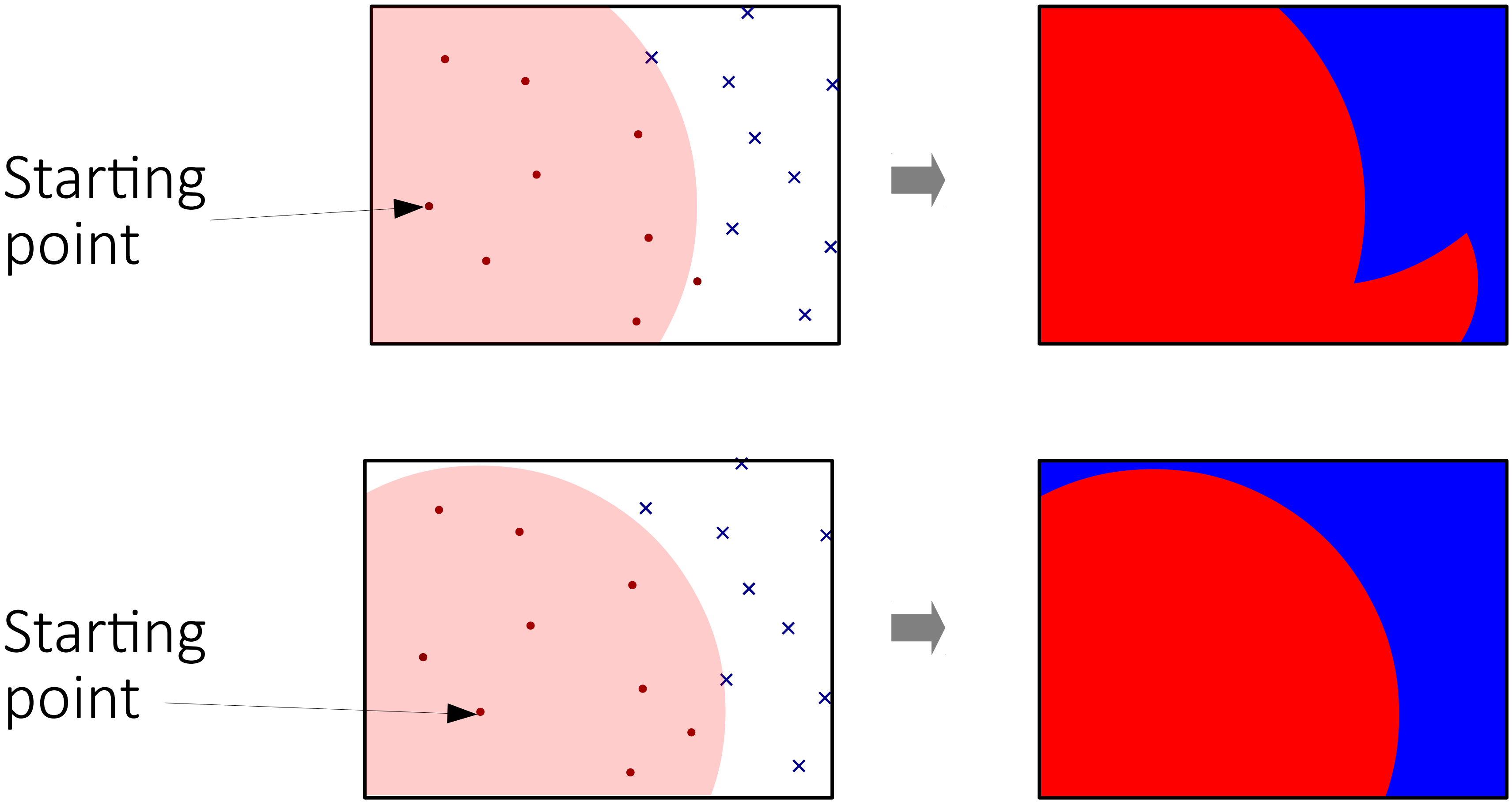
Train many guaranteed classifiers simultaneusly.

New lines of research

Starting point



New lines of research



New lines of research

Pool of guaranteed classifiers:

Train many guaranteed classifiers simultaneusly.

*Can we be **more** confident in the case of agreement?*

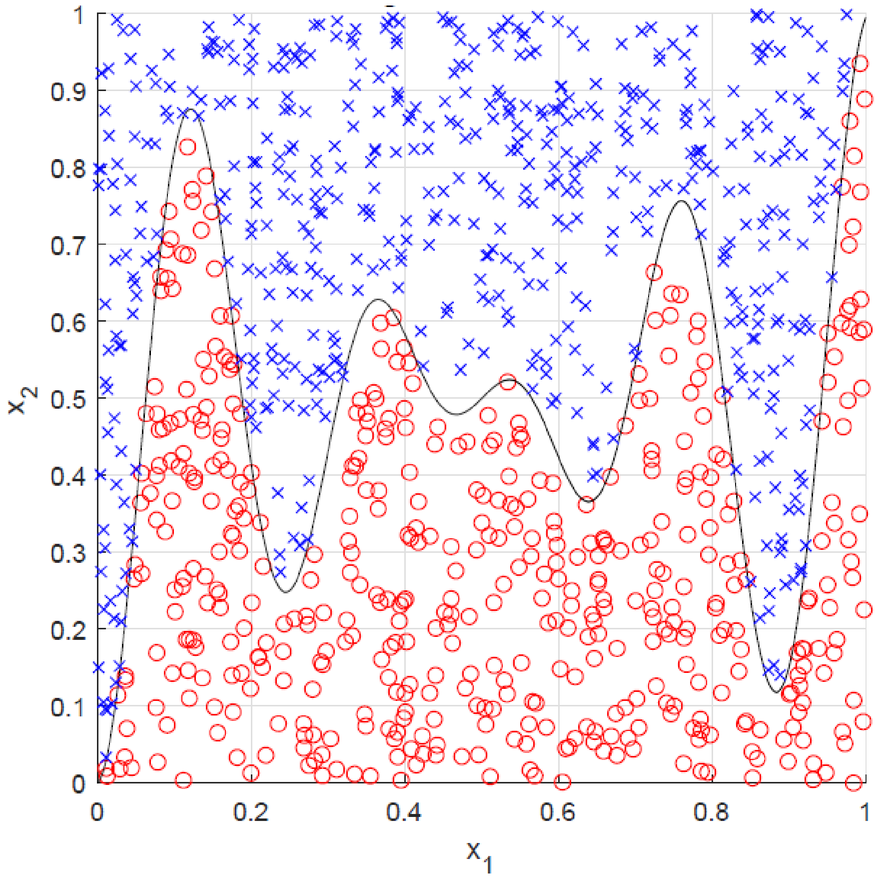
New lines of research

Theorem 2. If $PE_A + PE_B < 1$, then

$$PE_{ag} \leq \frac{PE_{best}}{1 + PE_{best} - PE_{worst}}.$$

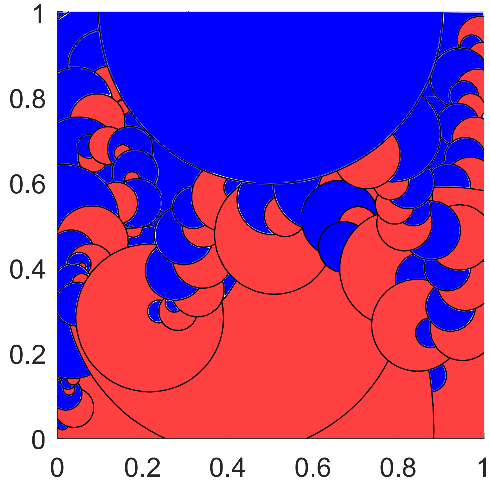
”Consensus and Reliability: The Case of Two Binary Classifiers”
A.T.J.R. Cobbenhagen. A. Carè, M.C. Campi, F.A. Ramponi, W.P.M.H. Heemels
8th IFAC Workshop on Distributed Estimation and Control in Networked Systems Sept. 16-17, 2019, Chicago, IL, USA

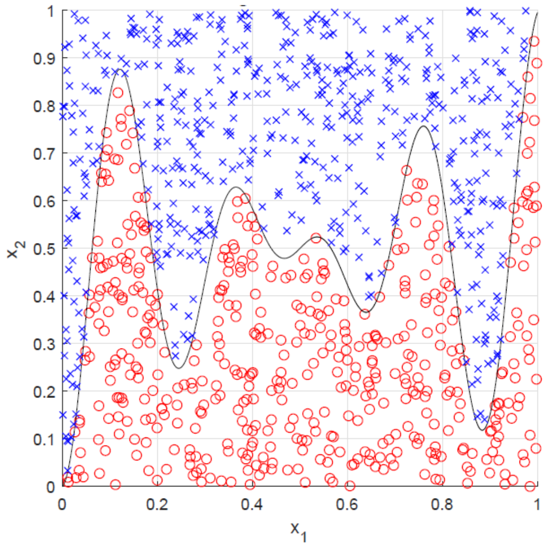
New lines of research



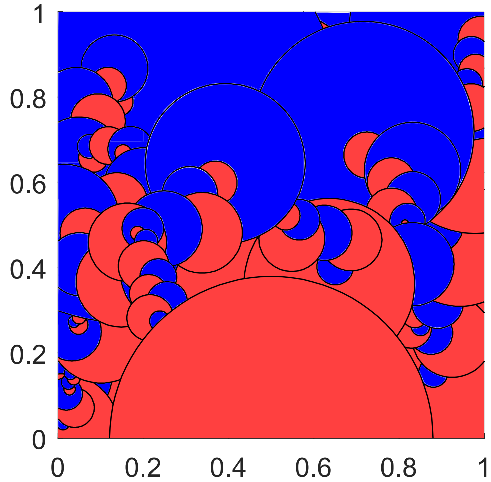
New lines of research

Classifier A

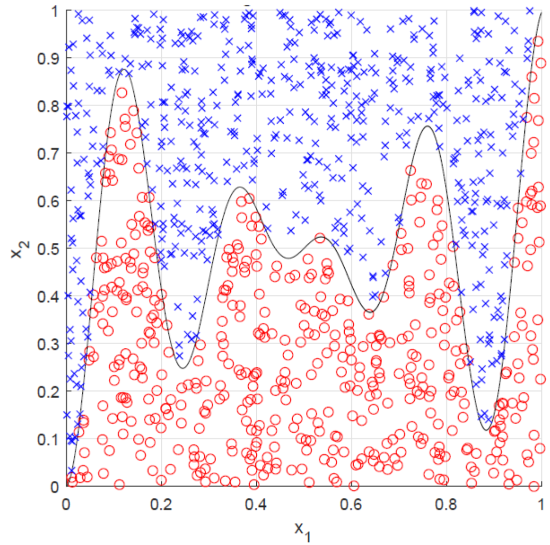




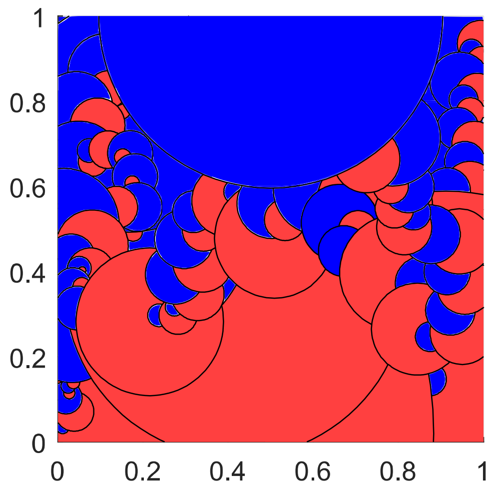
Classifier B



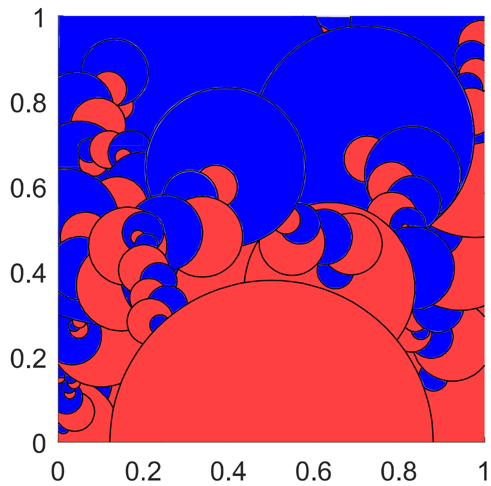
New lines of research



Classifier A



Classifier B



$$PE_A = 0.098,\ PE_B = 0.10,\ PE_{A \cap B} = 0.052$$
$$\alpha = 0.90, \text{ which yields } PE_{ag} = 0.057$$

Thank you!

Are you interested in knowing more?
Write to us!

algo.care@unibs.it

Try GEM-BALLS now!
www.algocare.it/GEM-BALLS/

Bibliography

GEM-BALLS:

- "A new classification algorithm with guaranteed sensitivity and specificity for medical applications,"
  A. Carè, F.A. Ramponi, M.C. Campi
  IEEE Control Systems Letters. 2(3):393-398 [doi: 10.1109/LCSYS.2018.2840427], 2018
  www.algocare.it/GEM-BALLS/

GEM:

- "Classification with Guaranteed Probability of Error"
  M.C. Campi
  Machine Learning, 80:63-84, 2010.

- "Consensus and Reliability: The Case of Two Binary Classifiers"
  A.T.J.R. Cobbenhagen. A. Carè, M.C. Campi, F.A. Ramponi, W.P.M.H. Heemels
  8th IFAC Workshop on Distributed Estimation and Control in Networked Systems
  Sept. 16-17, 2019, Chicago, IL, USA

- "Ventricular defibrillation: Classification with GEM and a roadmap for future investigations"
  F. Baronio, M. Baronio, M.C. Campi, A. Carè, S. Garatti, G. Perone
  2017 IEEE 56th Annual Conference on Decision and Control (CDC), 2718-2723

Leave-one-out estimates

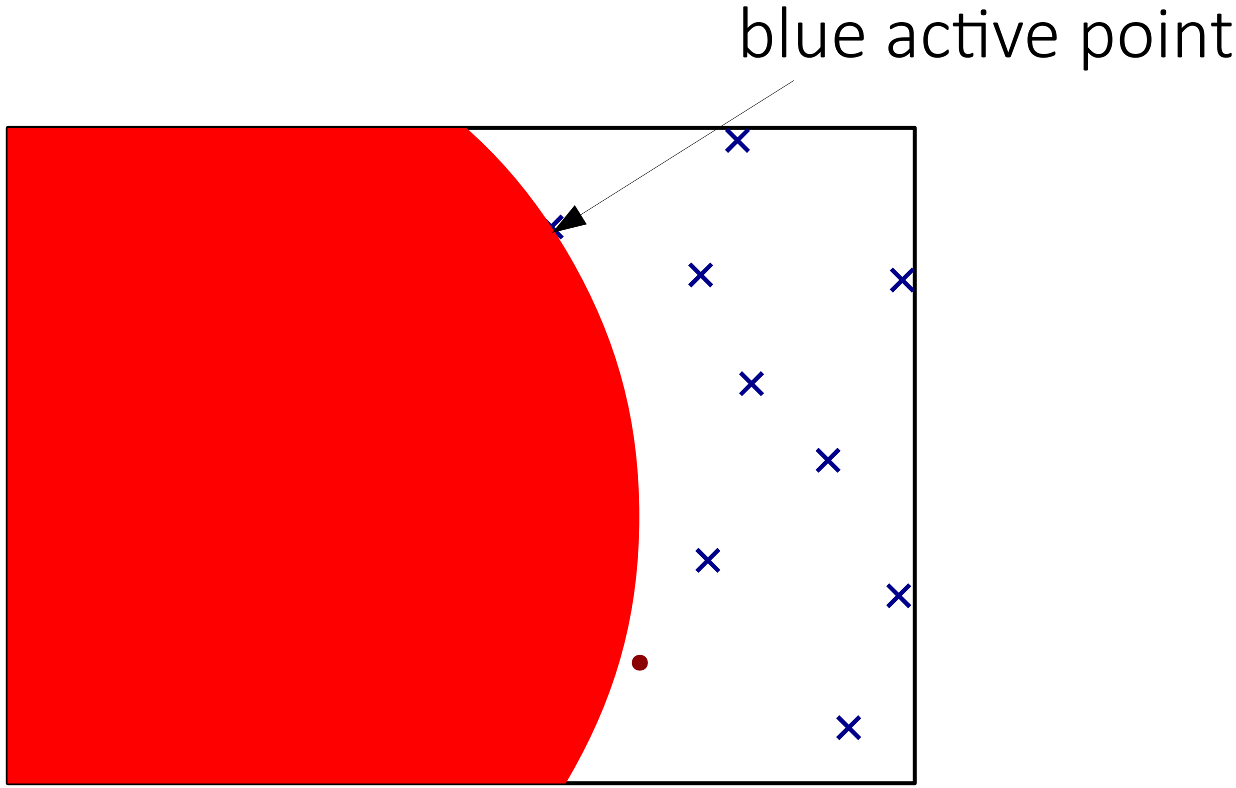
$$1 - \frac{k_1}{N_1}$$

is a sample-based estimate of sensitivity (leave-one-out estimate);

$$1 - \frac{k_0}{N_0}$$

is a sample-based estimate of specificity (leave-one-out estimate).

GEM-BALLS



GEM-BALLS