

Basics of Machine Learning

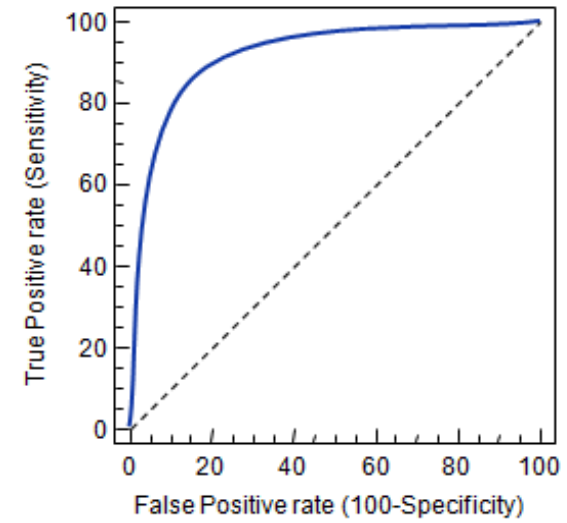
Dmitry Ryabokon, github.com/dryabokon





Lesson 17

Benchmarking



Benchmarking

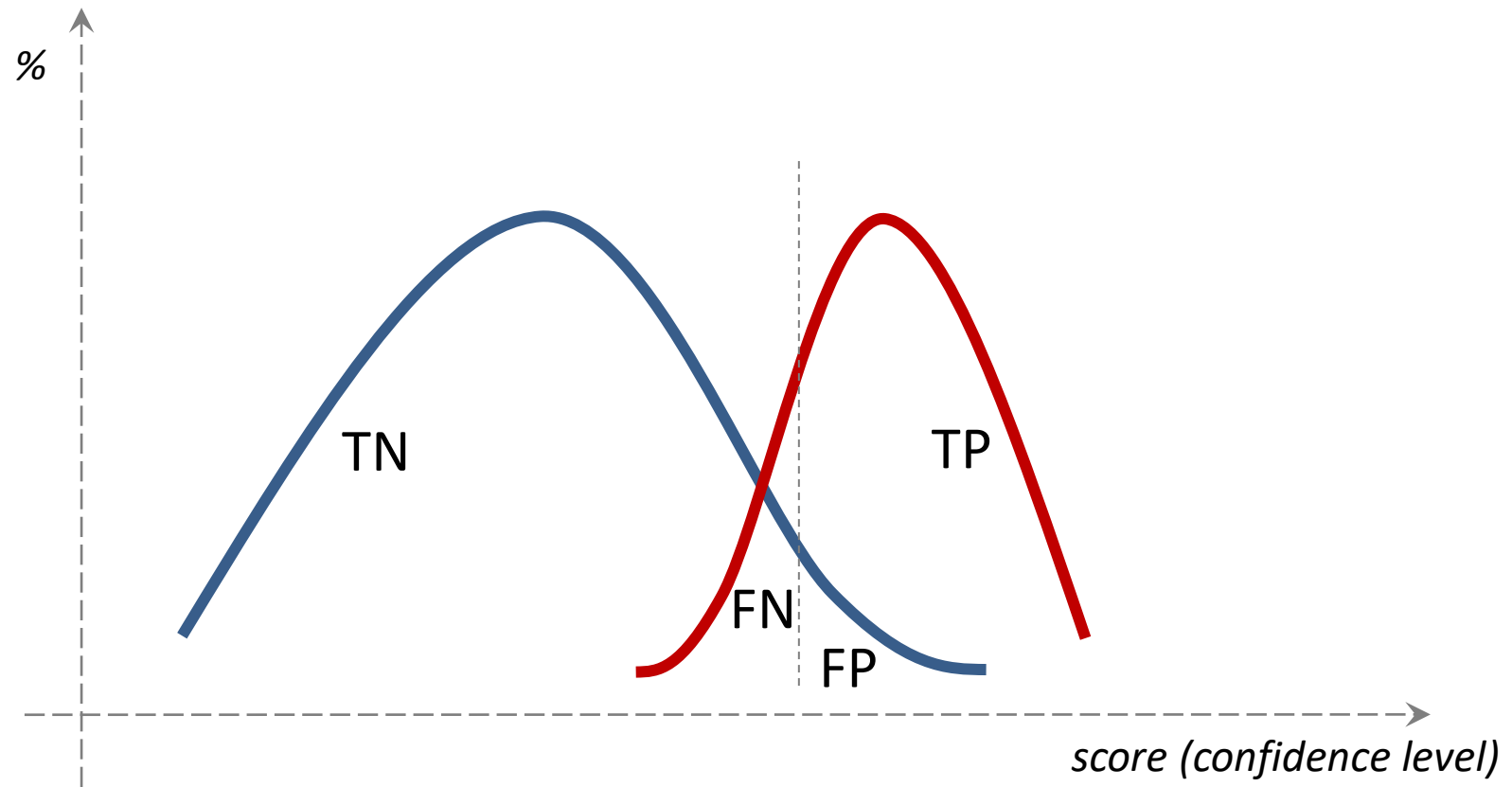
Summary

- ROC Curve
- PR-Curve
- Benchmarking the classifiers

ROC curve

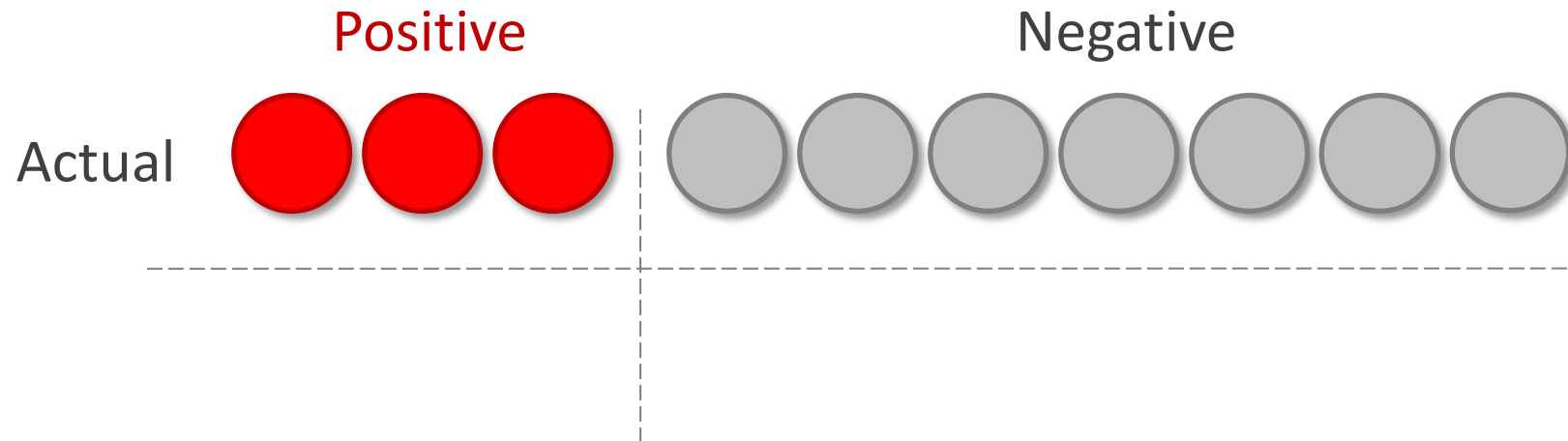


ROC curve



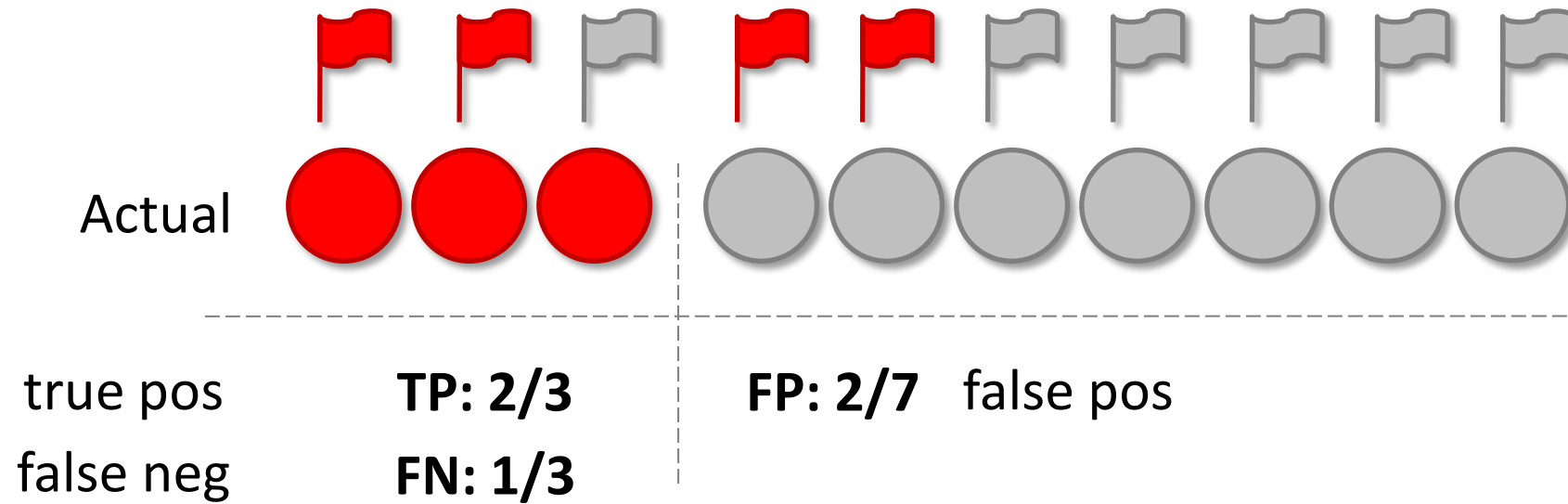
ROC curve

Error types

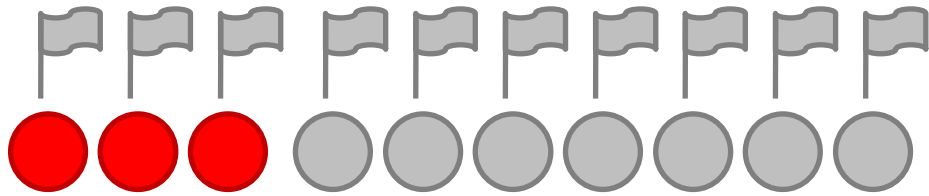


ROC curve

Error types



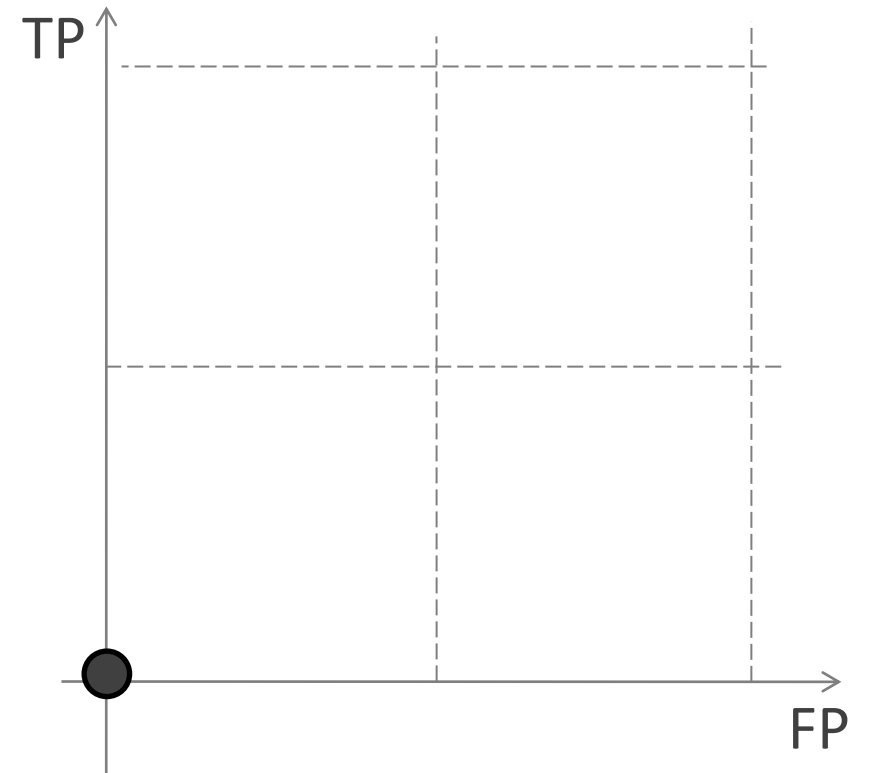
ROC curve



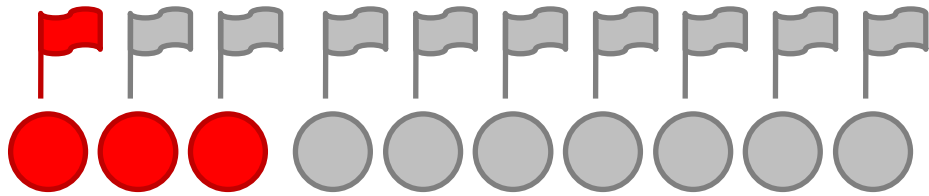
TP: 0/3

FP: 0/7

FN: 3/3



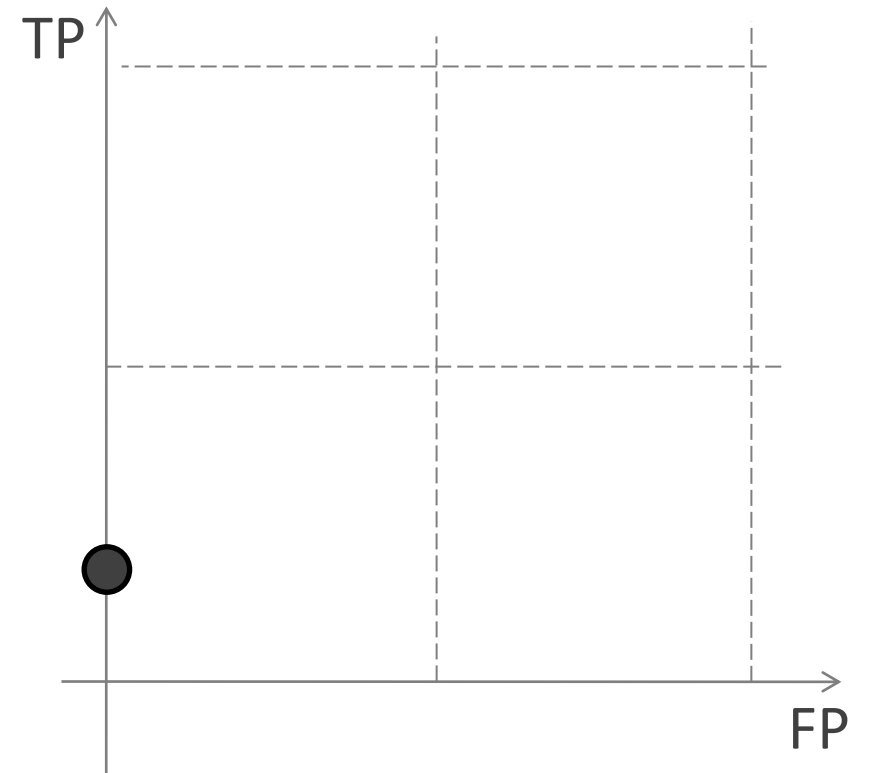
ROC curve



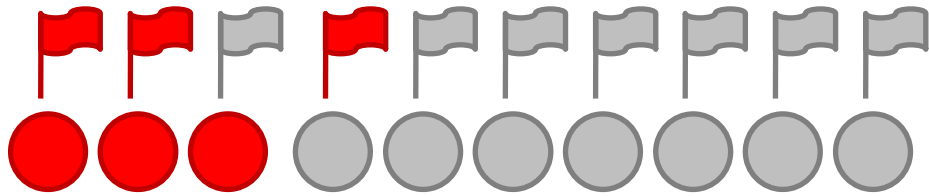
TP: 1/3

FP: 0/7

FN: 2/3



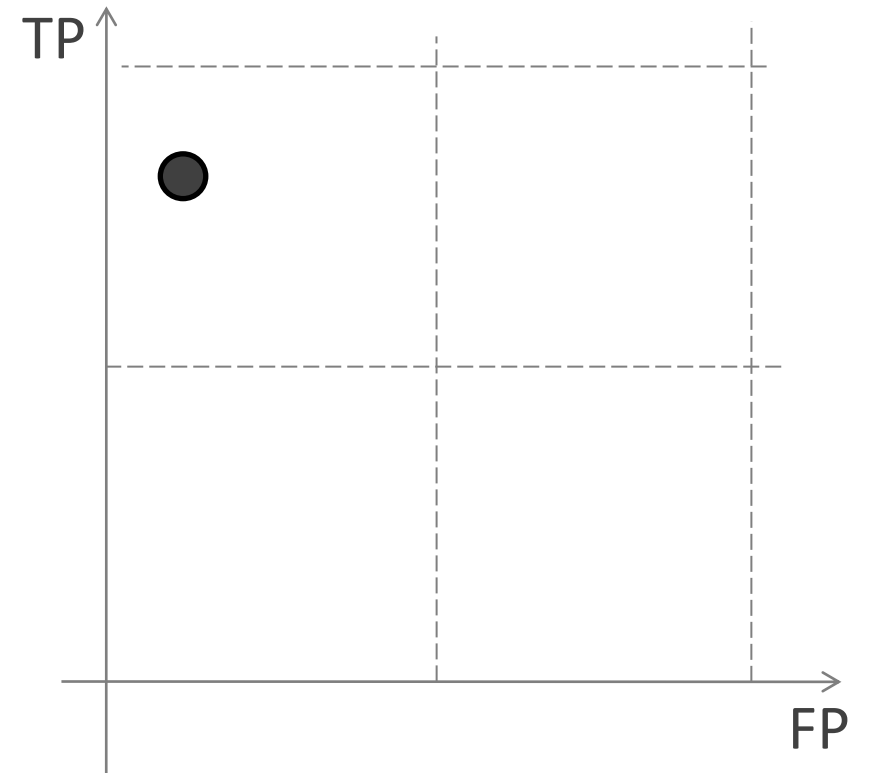
ROC curve



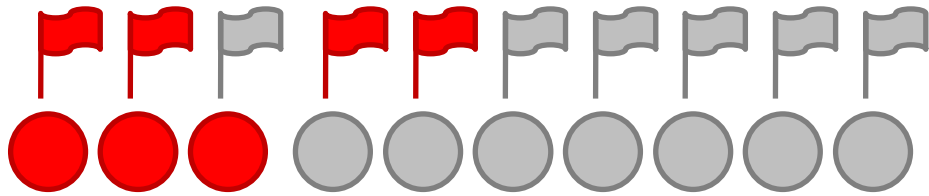
TP: 2/3

FP: 1/7

FN: 1/3



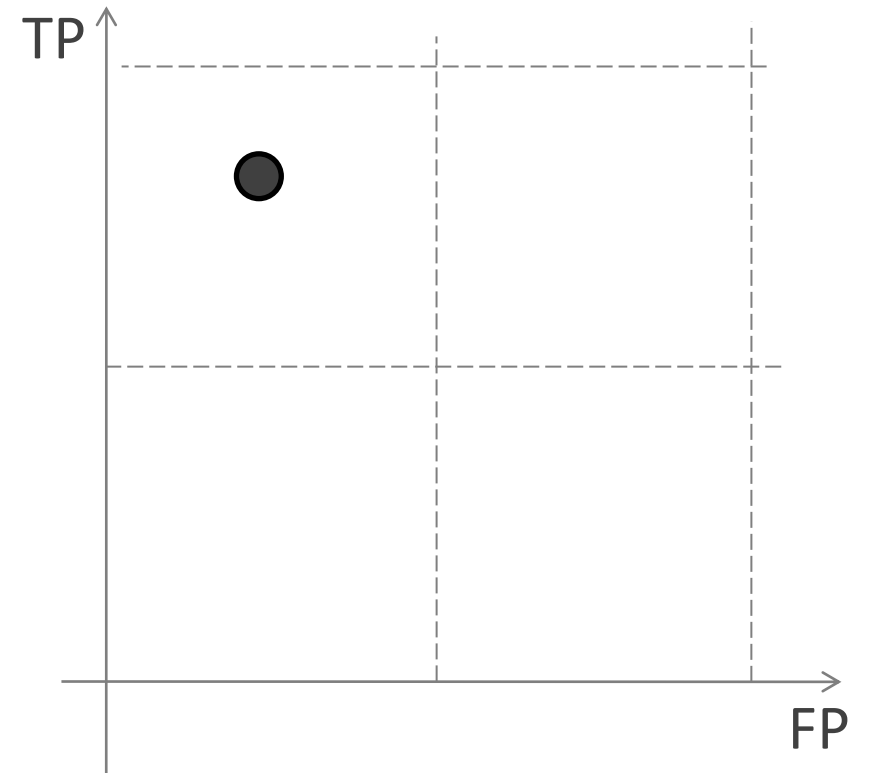
ROC curve



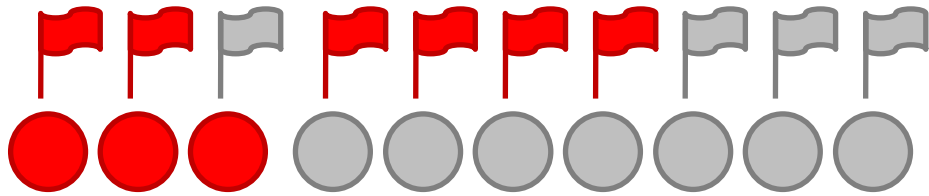
TP: 2/3

FP: 2/7

FN: 1/3



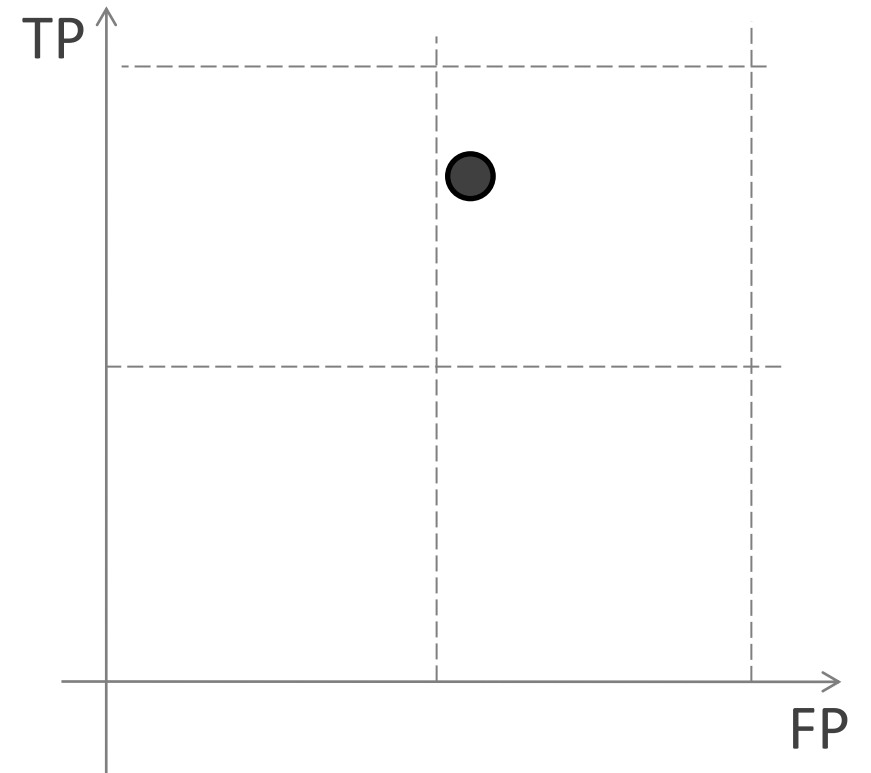
ROC curve



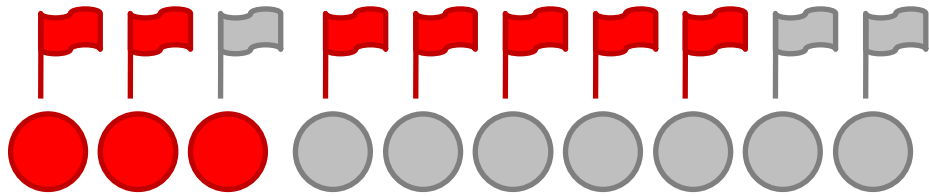
TP: 2/3

FP: 4/7

FN: 1/3



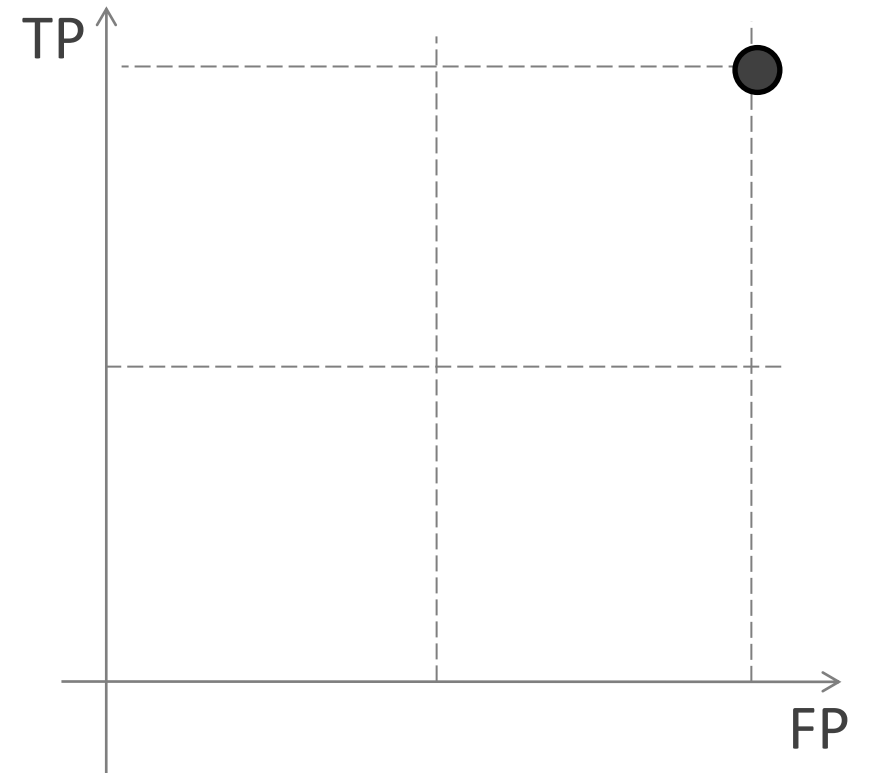
ROC curve



TP: 3/3

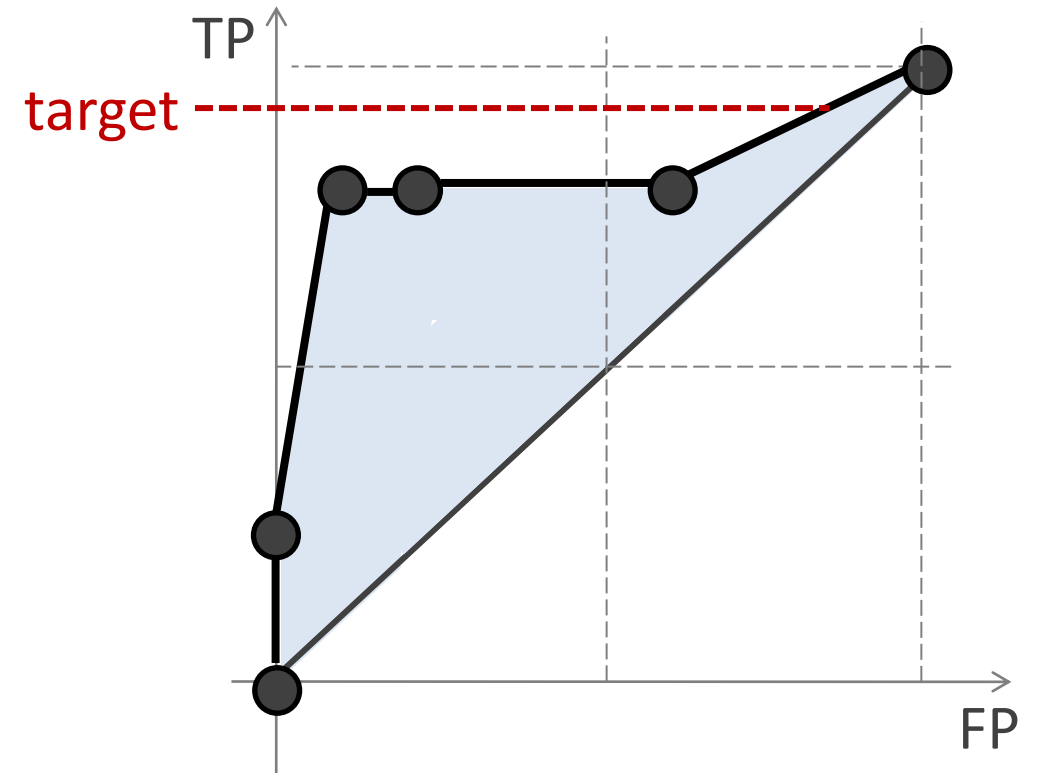
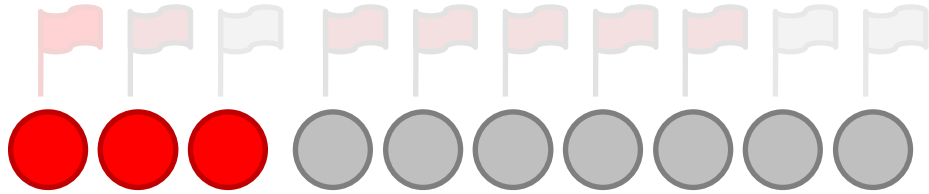
FP: 7/7

FN: 0/3



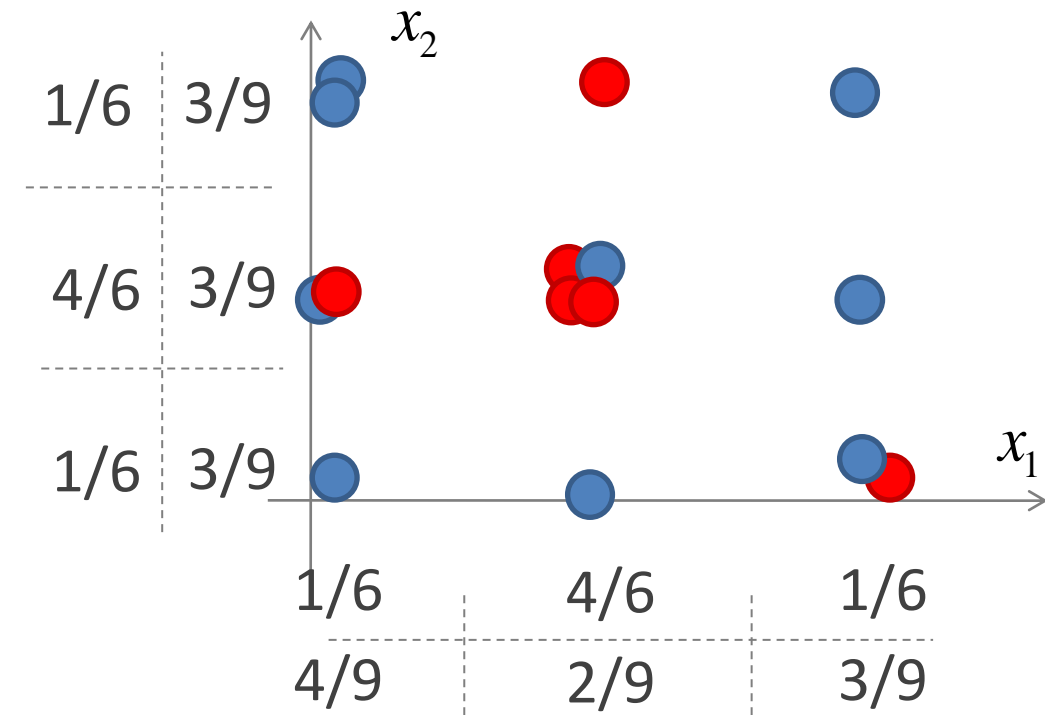
ROC curve

AUC: area under ROC curve



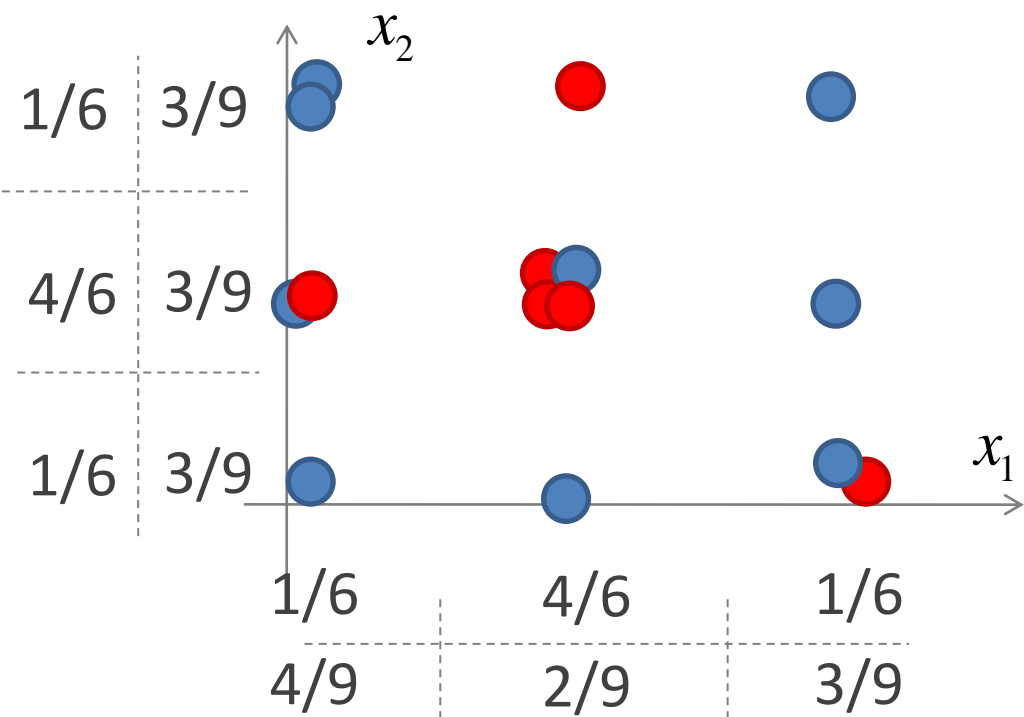
ROC curve

Example: Naive Bayesian classifier



ROC curve

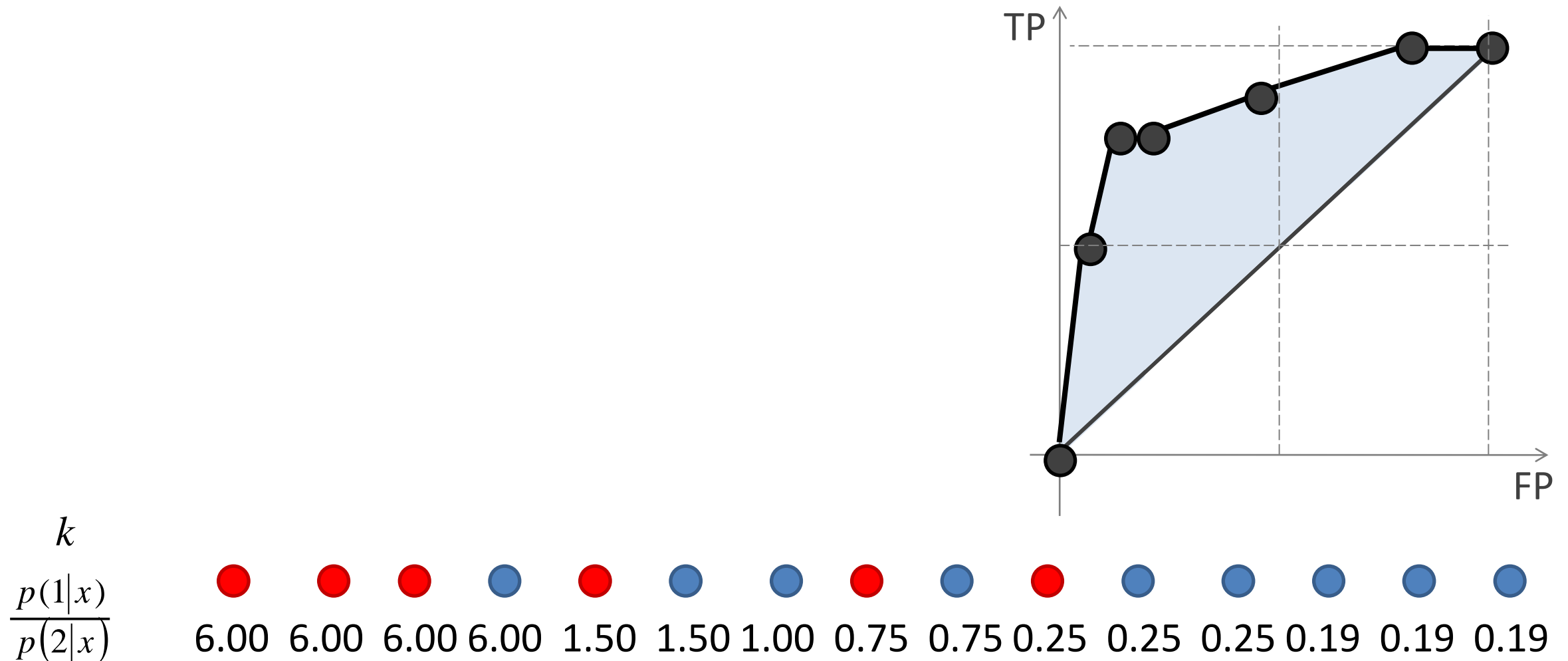
Example: Naive Bayesian classifier



x	(0,0)	(1,0)	(2,0)	(2,0)	(0,1)	(0,1)	(1,1)	(1,1)	(1,1)	(1,1)	(2,1), (0,2)	(0,2)	(1,2)	(2,2)	
k															
$\frac{p(1 x)}{p(2 x)}$	0.19	1.50	0.25	0.25	0.75	0.75	6.00	6.00	6.00	6.00	1.00	0.19	0.19	1.50	0.25

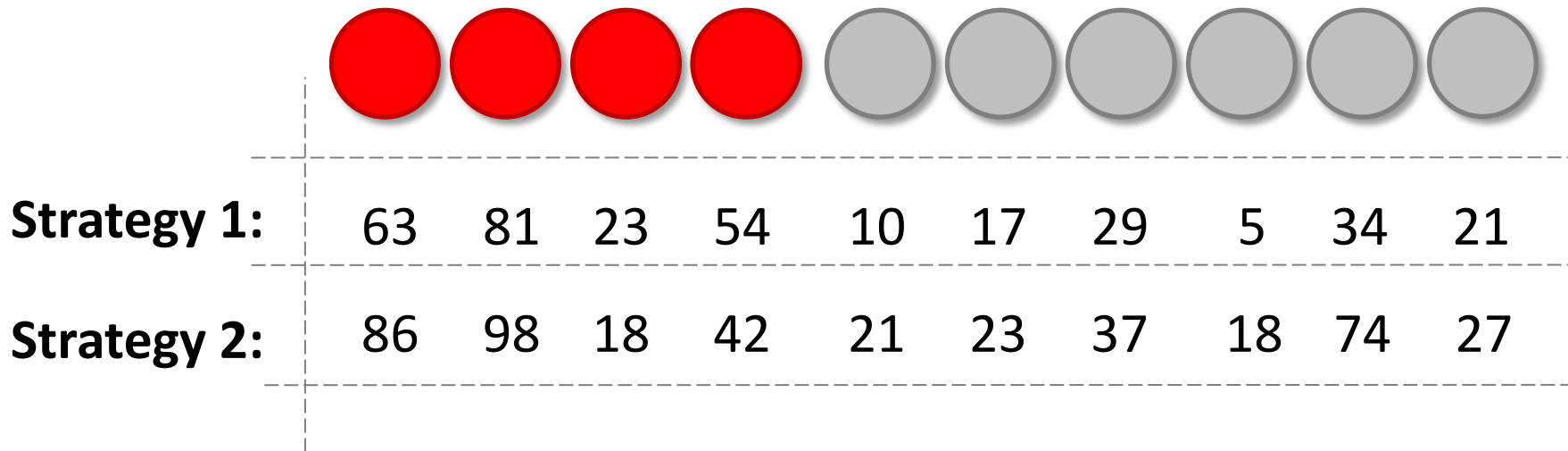
ROC curve

Example: Naive Bayesian classifier



ROC curve

The task: compare accuracy of two strategies



Precision-Recall

Precision - Recall

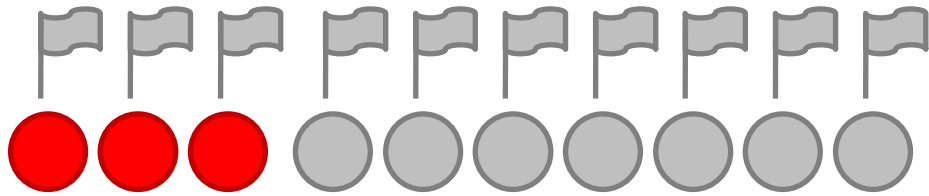
Precision: $TP / (TP+FP)$

Recall: $TP / (TP+FN)$

Accuracy: $(TP+TN) / \text{Total}$

F1-score: $2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$

Precision - Recall



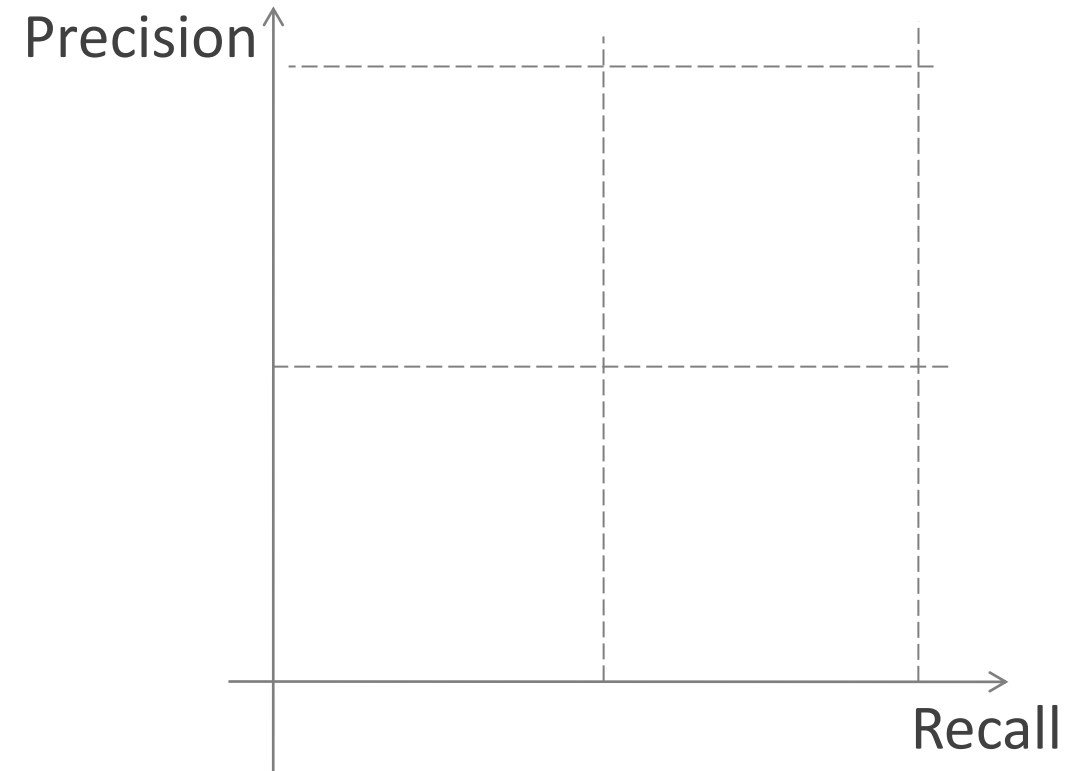
TP: 0

FP: 0

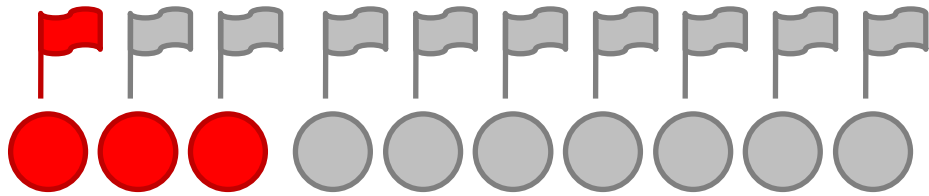
Precision: 0/0

FN: 3

Recall: 0/3



Precision - Recall



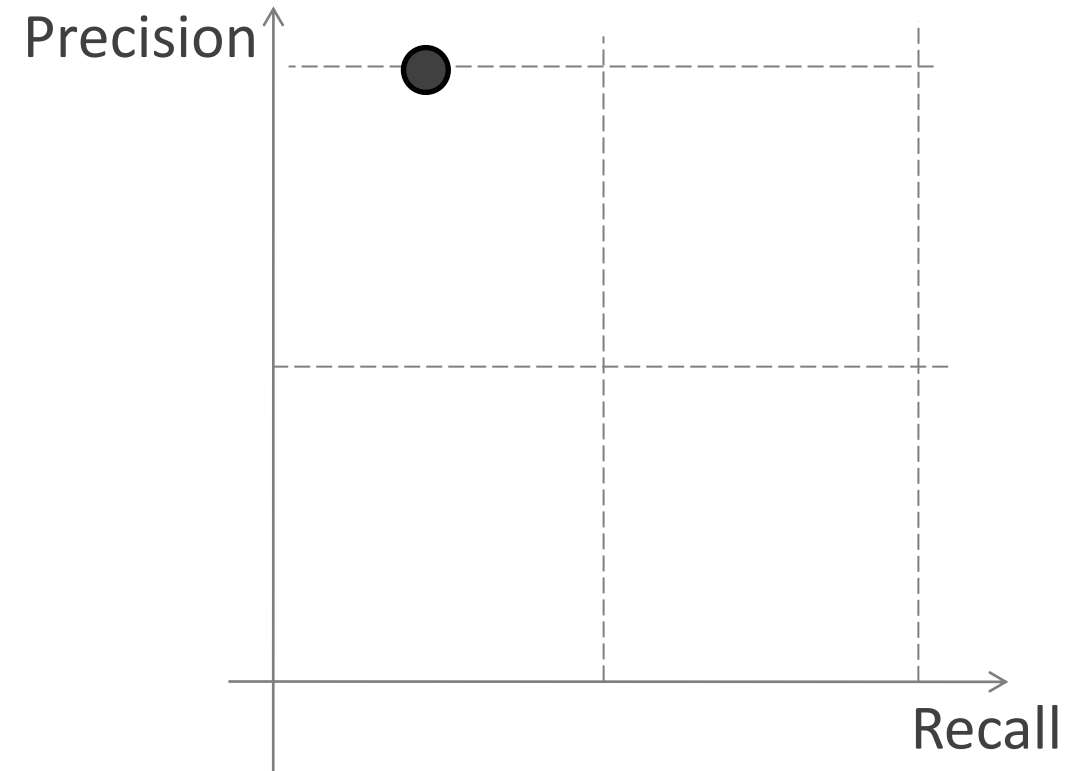
TP: 1

FP: 0

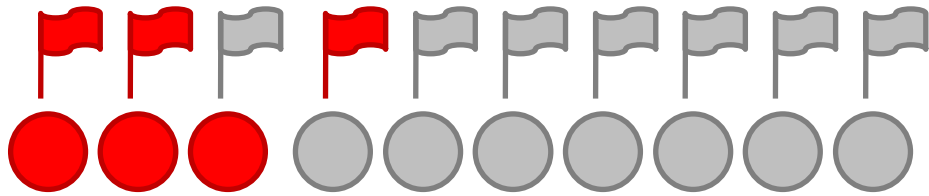
Precision: 1

FN: 2

Recall: 1/3



Precision - Recall



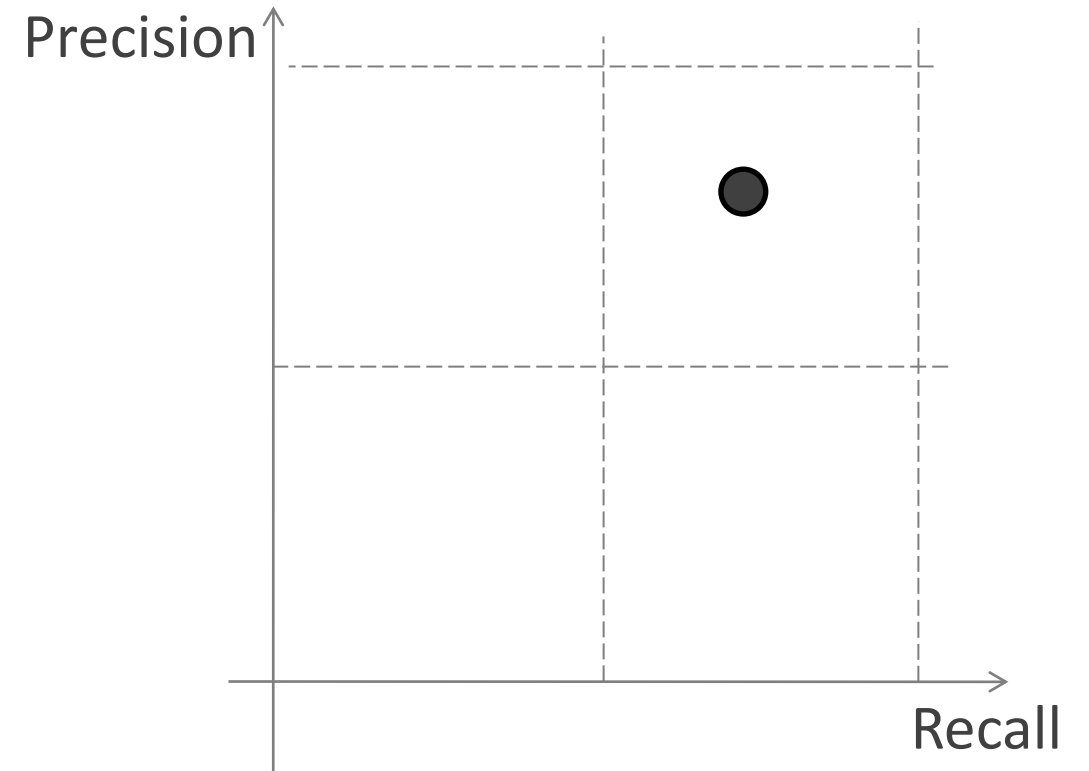
TP: 2

FP: 1

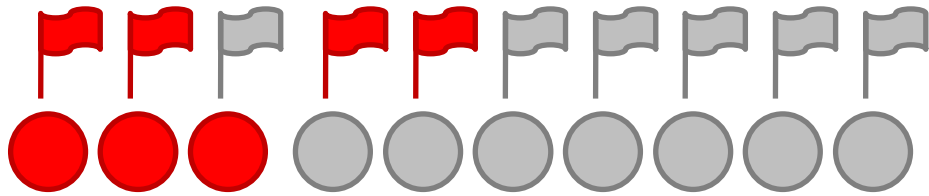
Precision: $2/3$

FN: 1

Recall: $2/3$



Precision - Recall



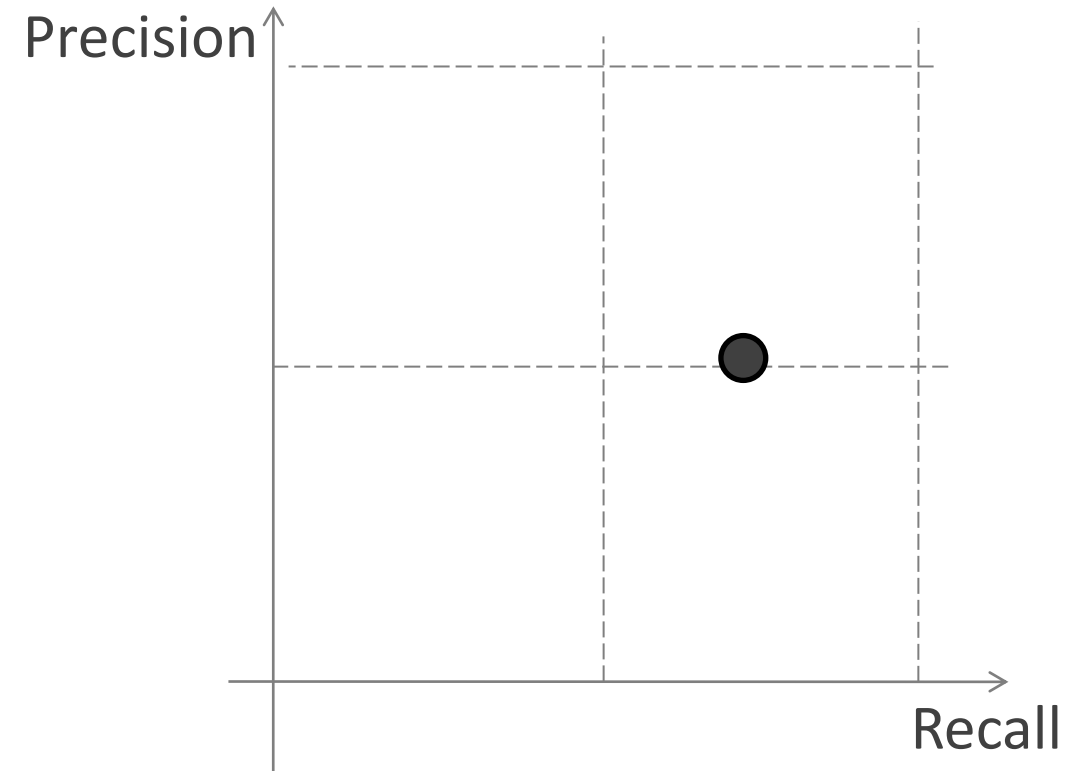
TP: 2

FP: 2

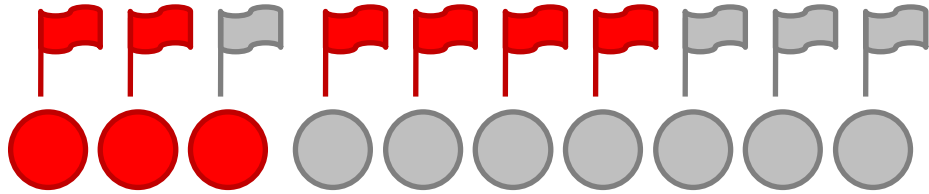
Precision: 1/2

FN: 1

Recall: 2/3



Precision - Recall



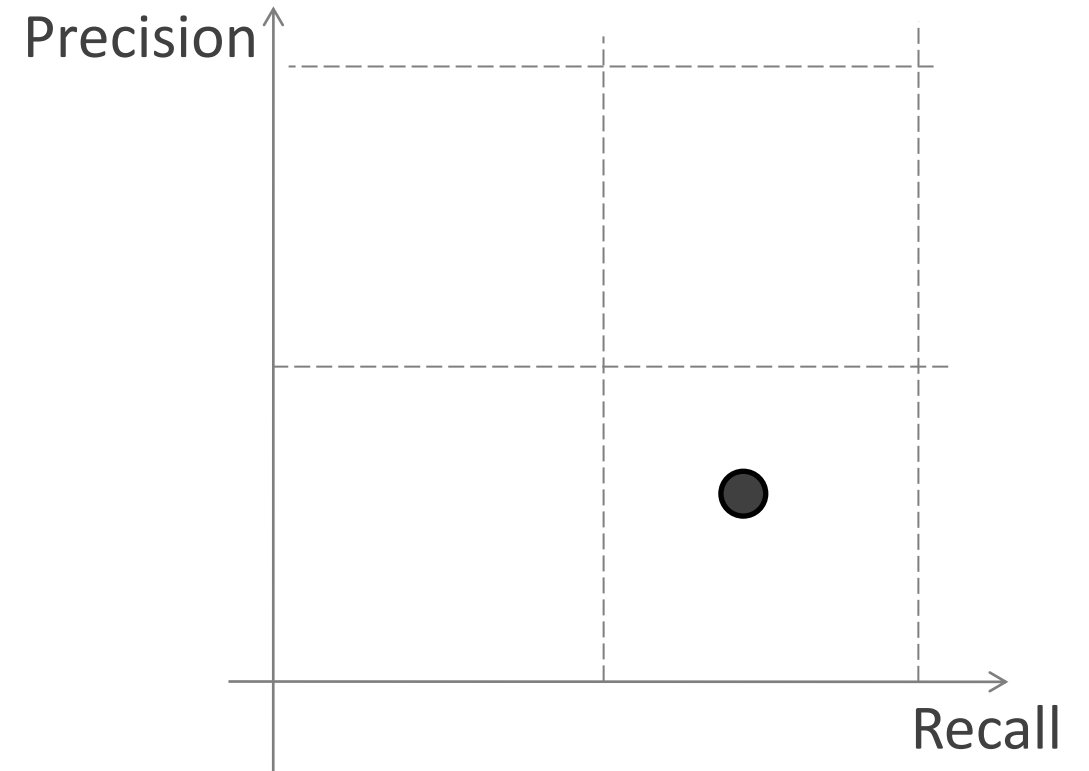
TP: 2

FP: 4

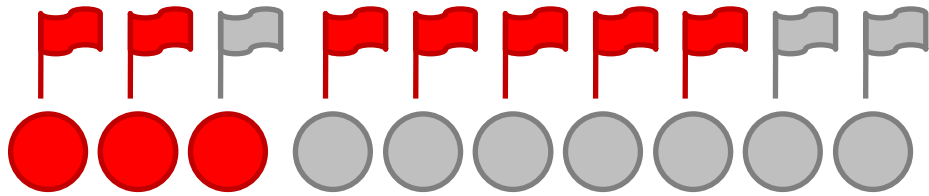
Precision: $\frac{2}{6}$

FN: 1

Recall: $\frac{2}{3}$



Precision - Recall



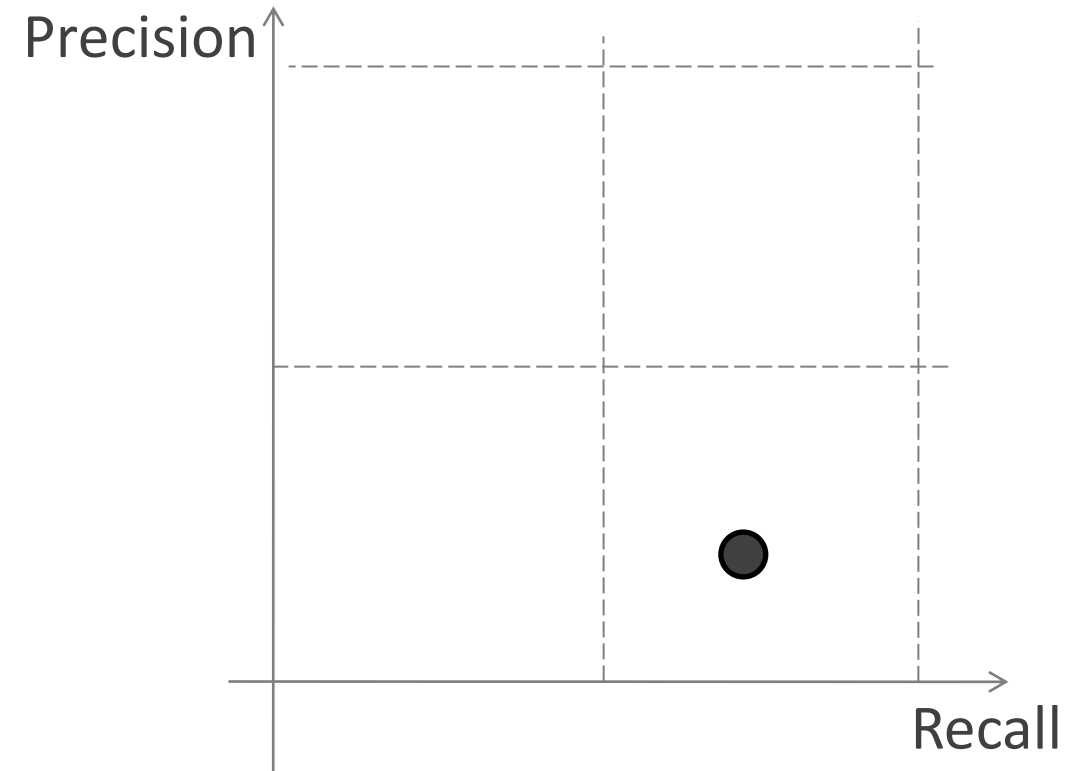
TP: 2

FP: 5

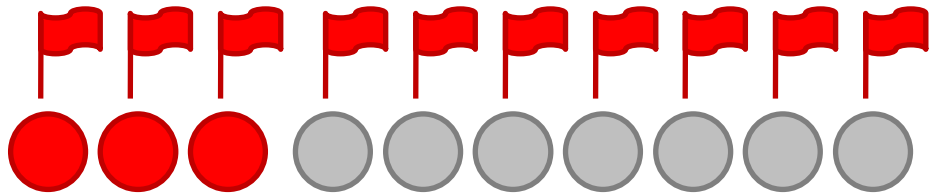
Precision: 2/7

FN: 1

Recall: 2/3



Precision - Recall



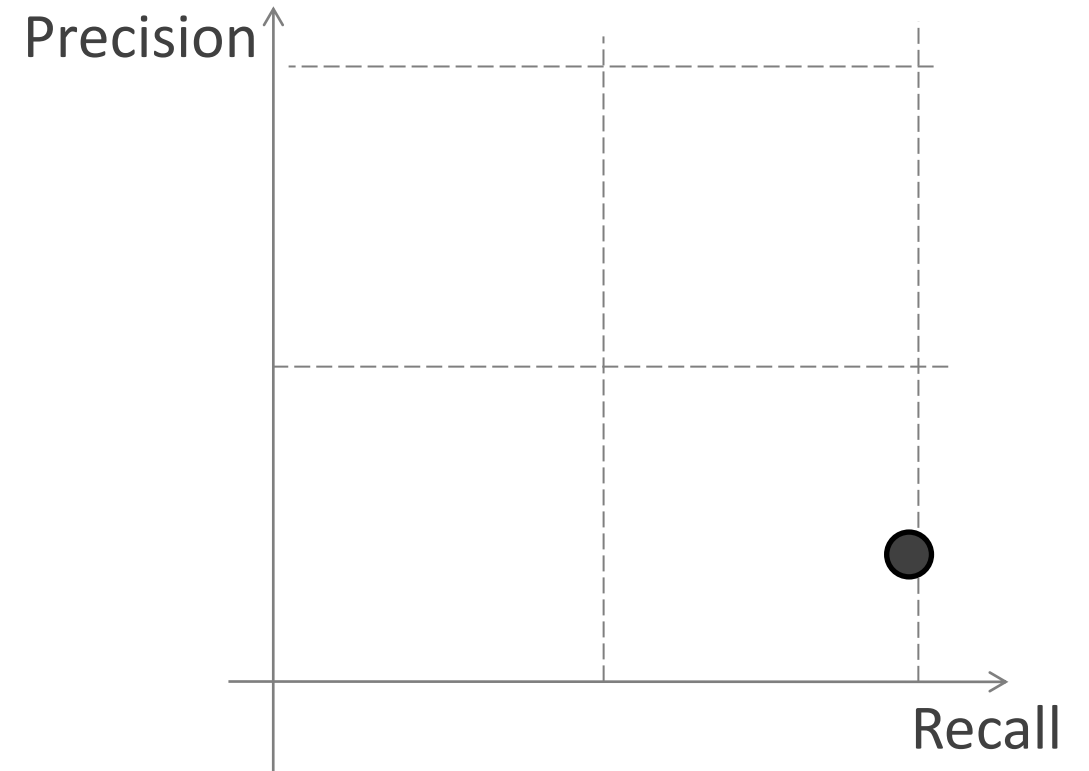
TP: 3

FP: 7

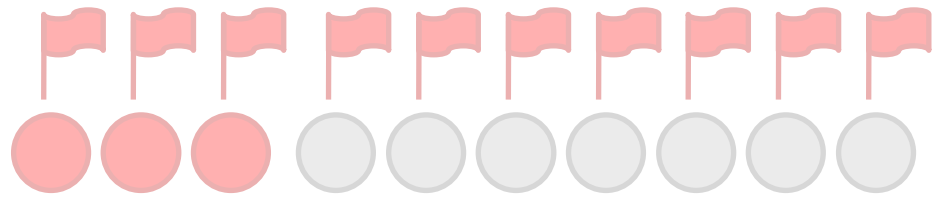
Precision: 3/7

FN: 0

Recall: 1



Precision - Recall

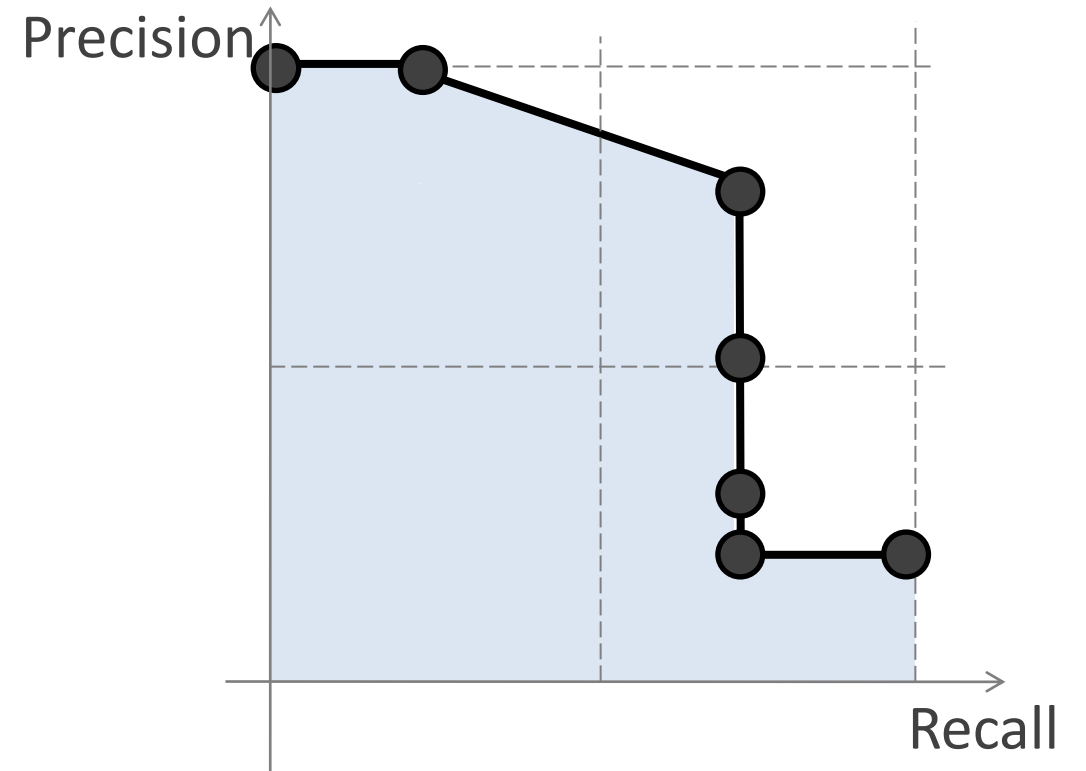


TP: 3

FP: 7

FN: 0

Recall: 1



Precision - Recall



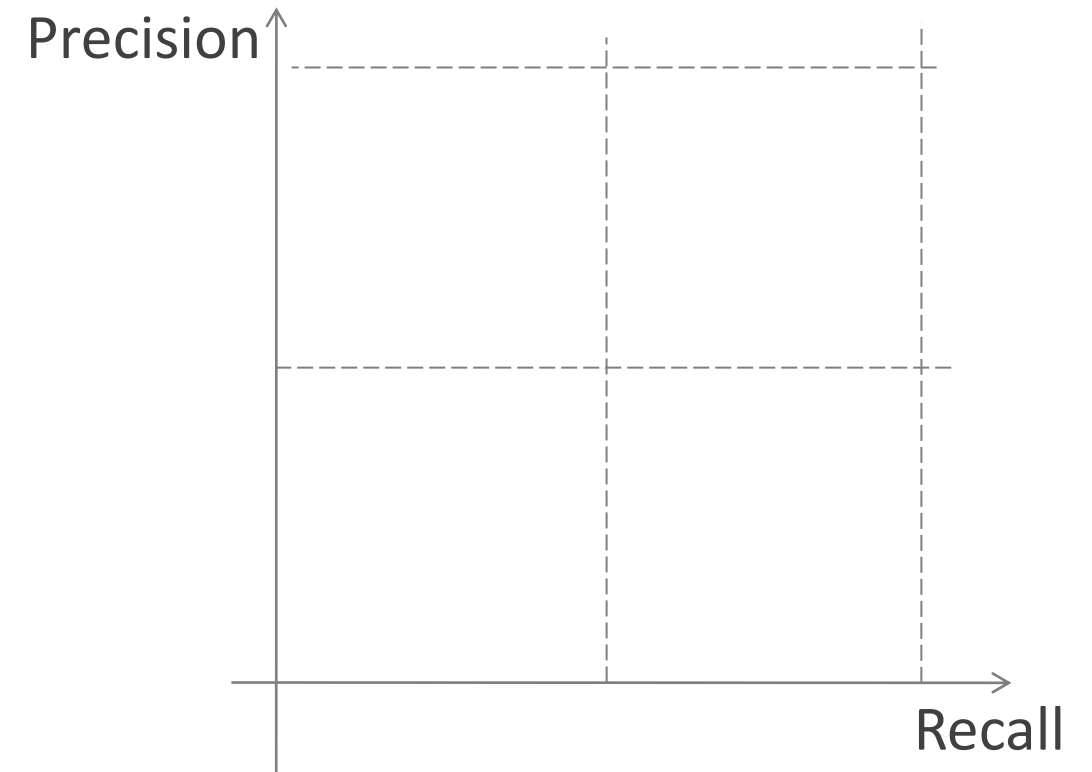
TP: 0

FP: 0

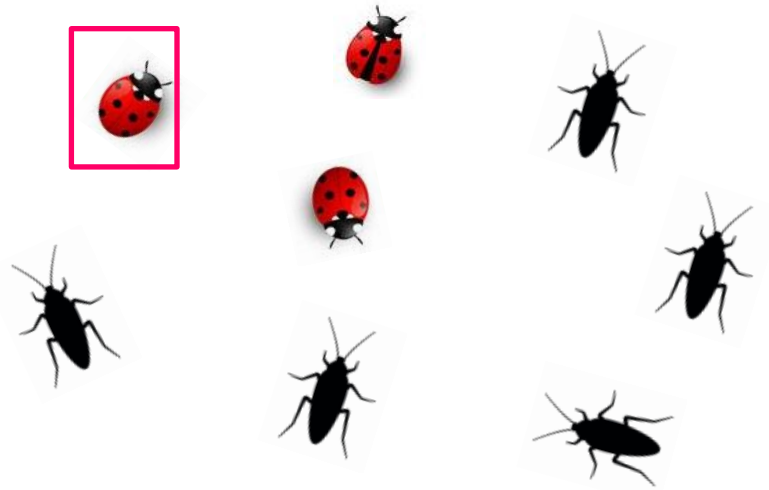
FN: 3

Precision: 0

Recall: 0



Precision - Recall



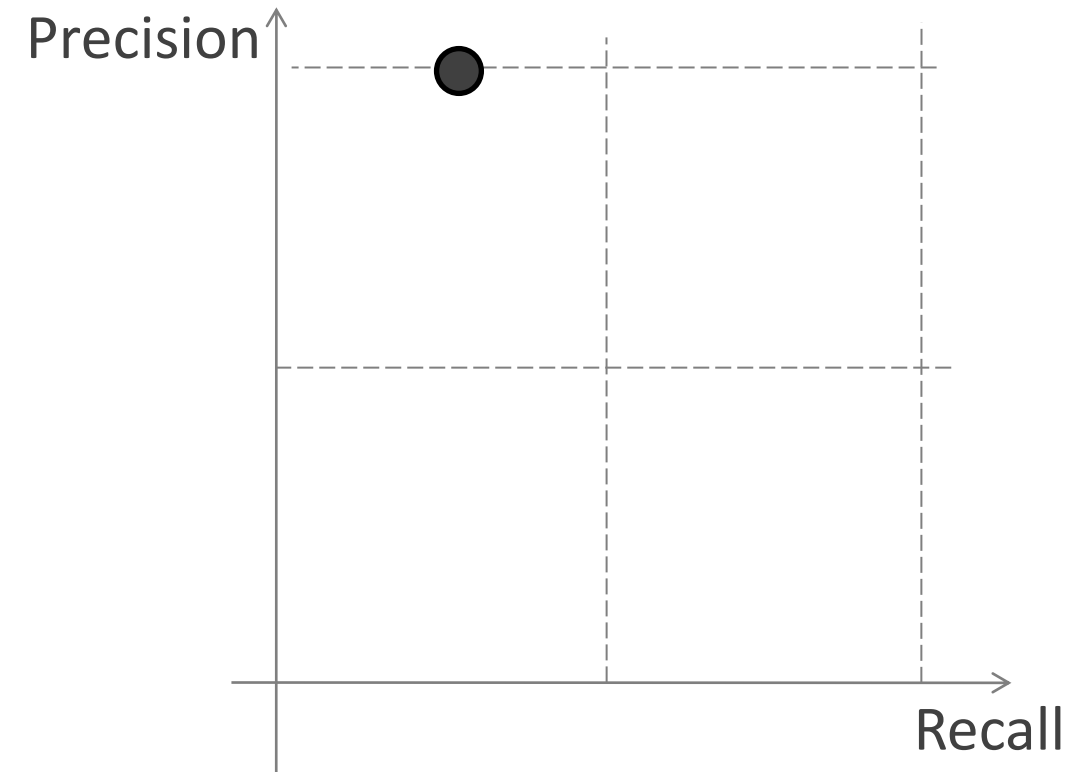
TP: 1

FP: 0

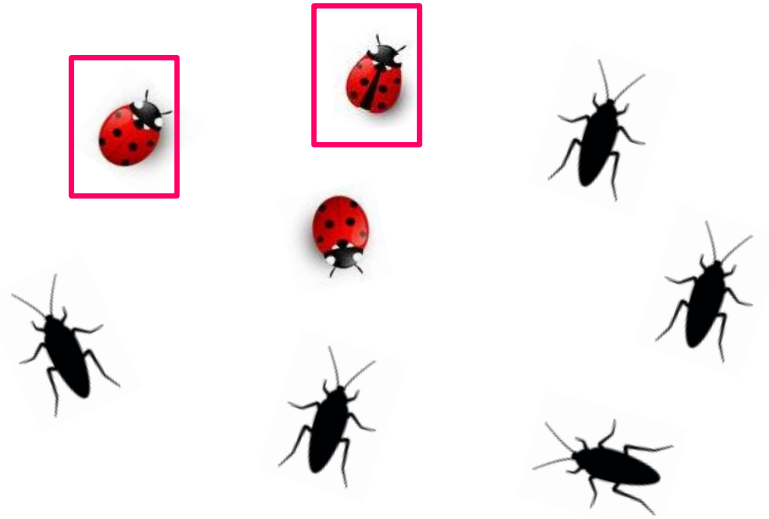
FN: 2

Precision: 1

Recall: 1/3



Precision - Recall



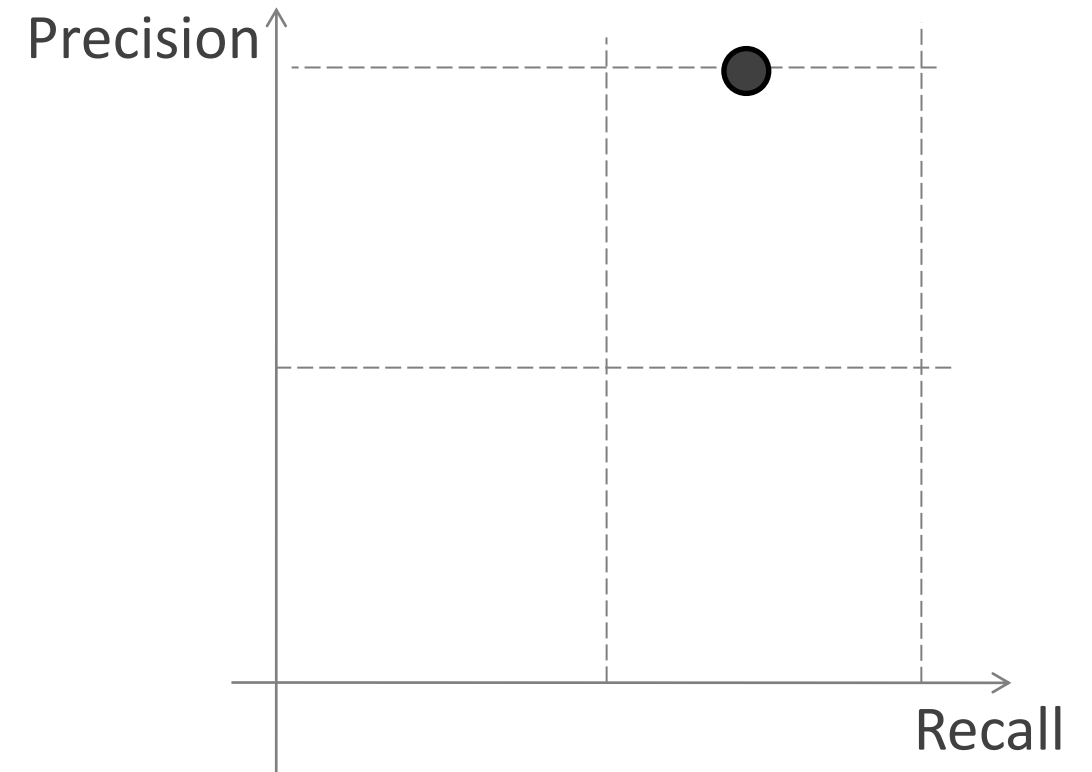
TP: 2

FP: 0

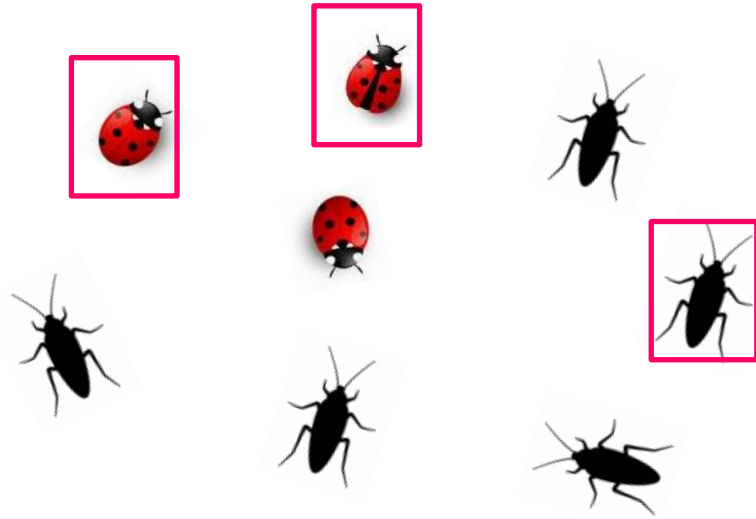
FN: 1

Precision: 1

Recall: 2/3



Precision - Recall



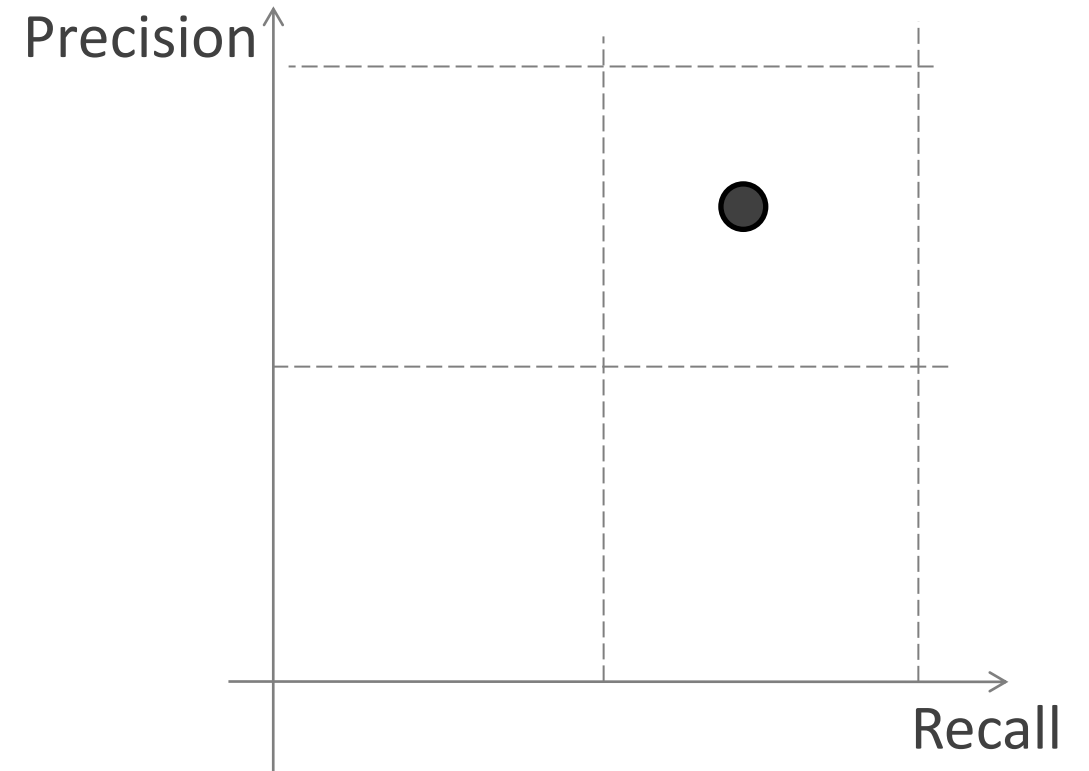
TP: 2

FP: 1

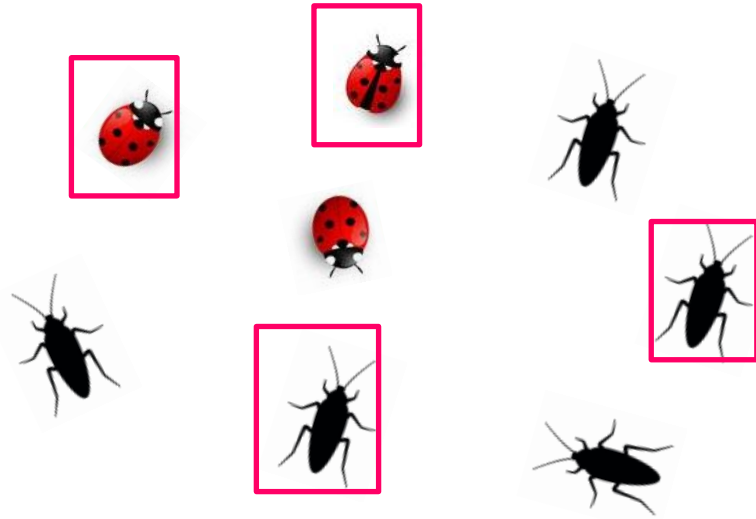
Precision: $2/3$

FN: 1

Recall: $2/3$



Precision - Recall



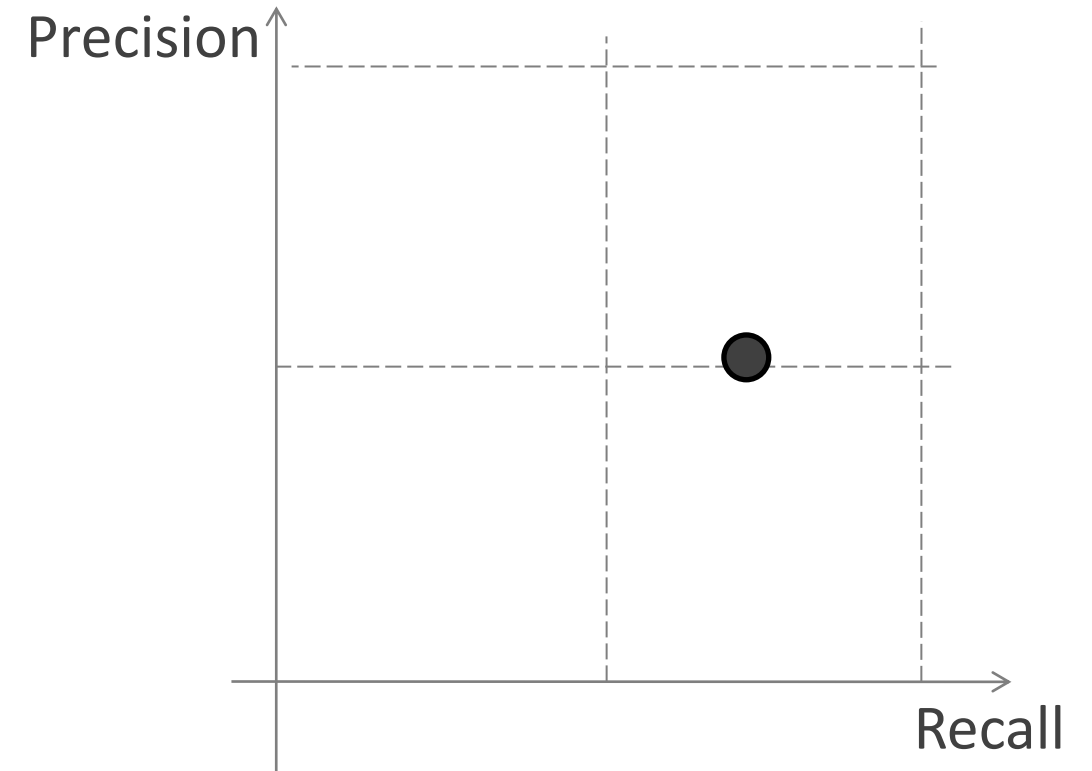
TP: 2

FP: 2

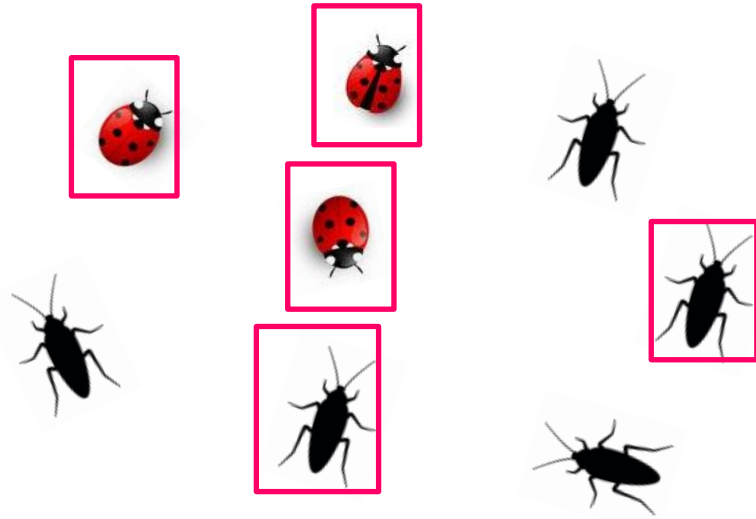
Precision: 2/4

FN: 1

Recall: 2/3



Precision - Recall



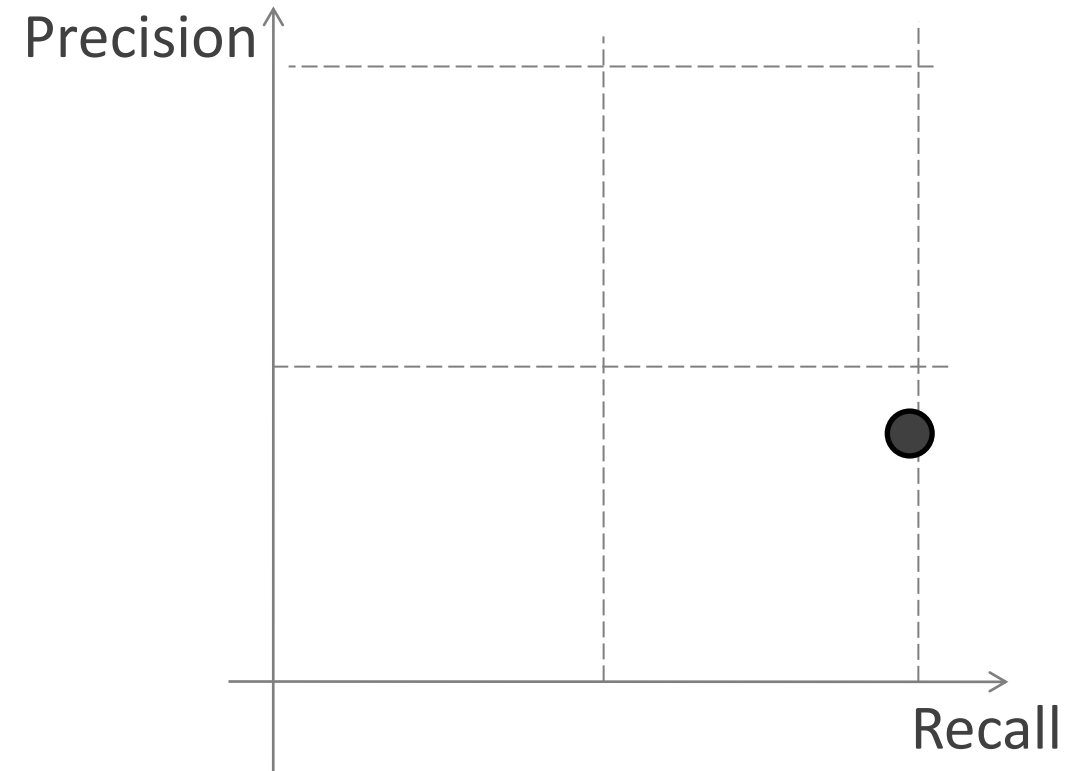
TP: 3

FP: 2

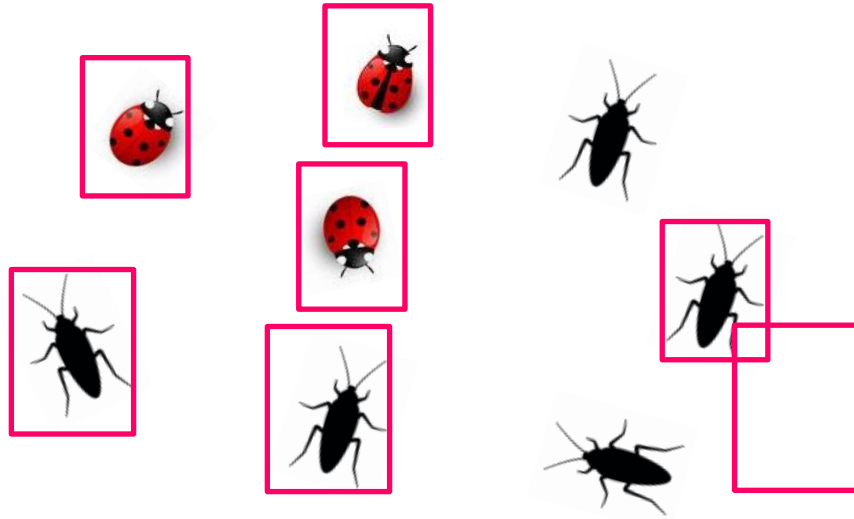
Precision: 2/5

FN: 0

Recall: 1



Precision - Recall



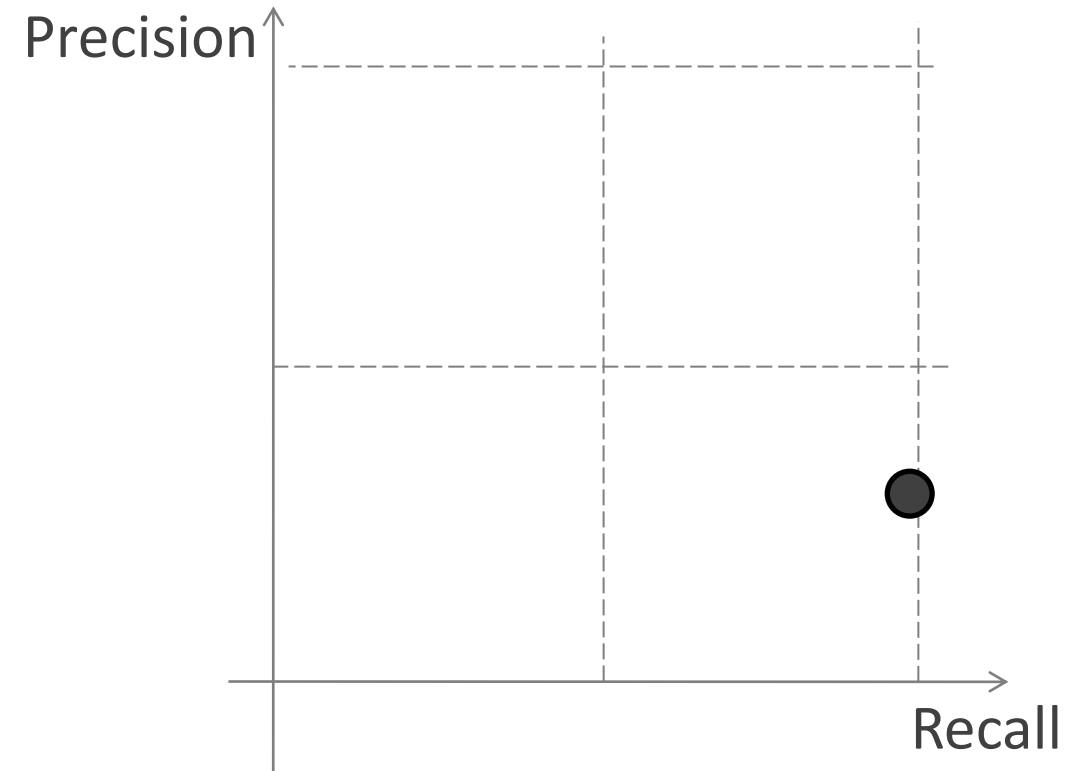
TP: 3

FP: 4

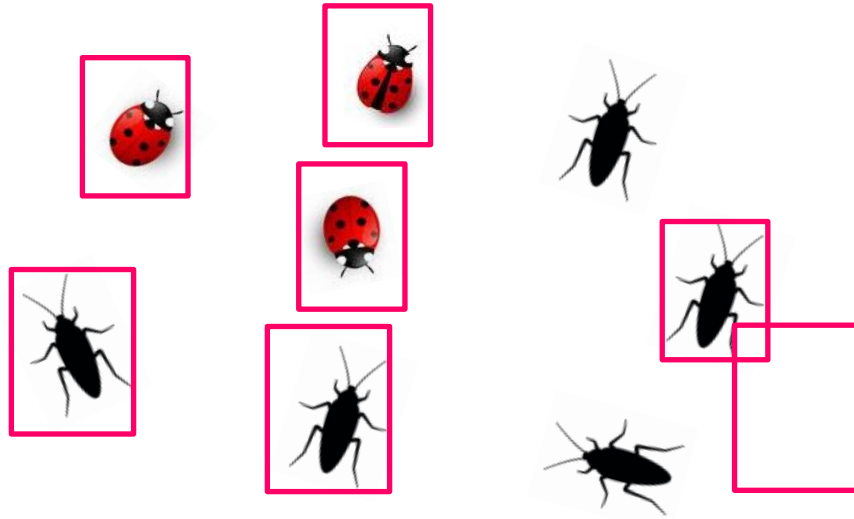
Precision: 3/7

FN: 0

Recall: 1



Precision - Recall



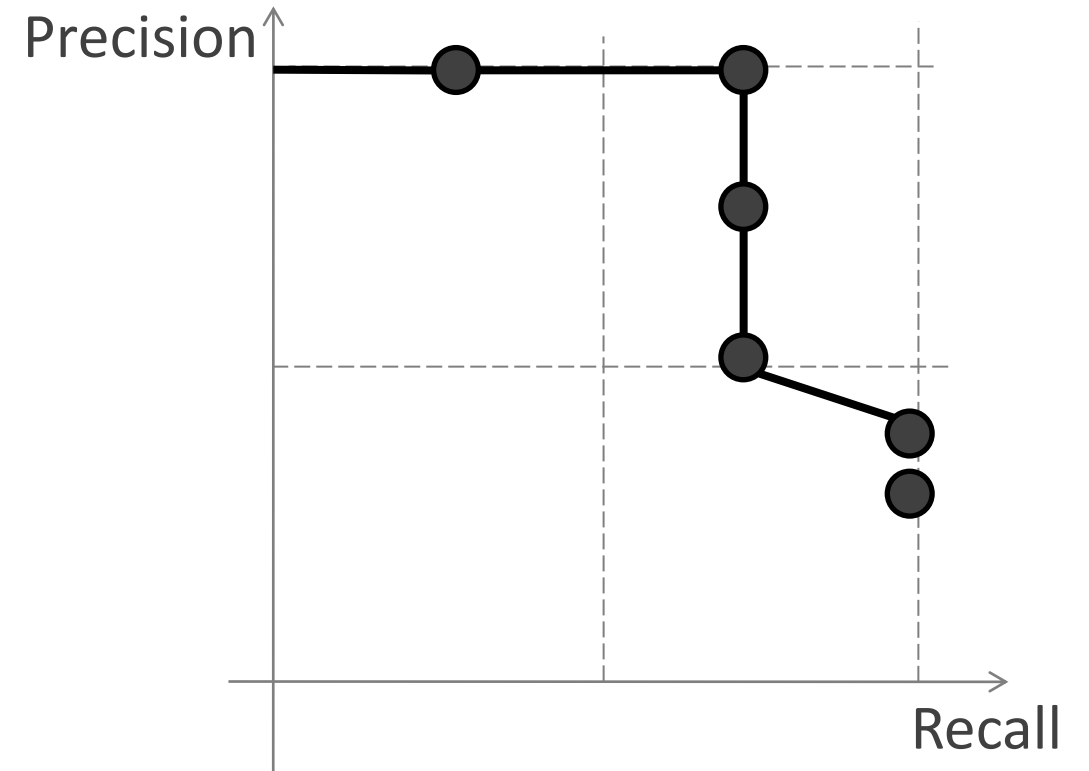
TP: 3

FP: 4

Precision: 3/7

FN: 0

Recall: 1



Benchmarking the classifiers

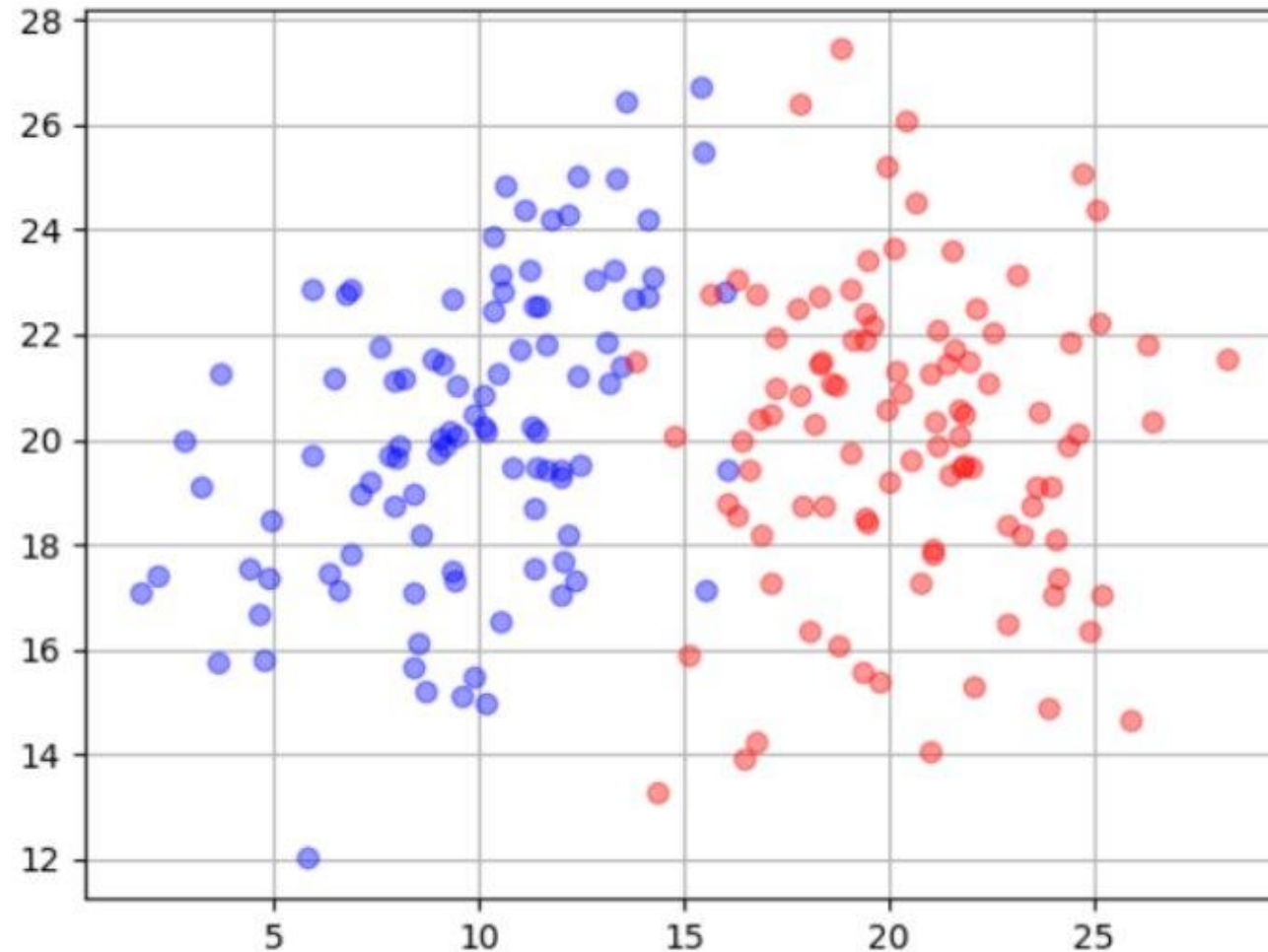
github.com/dryabokon/ML/blob/master/ex_03_03_benchmark_classifiers_k2.py
github.com/dryabokon/ML/blob/master/ex_03_03_benchmark_classifiers_multi_class.py

Benchmarking the Classifiers

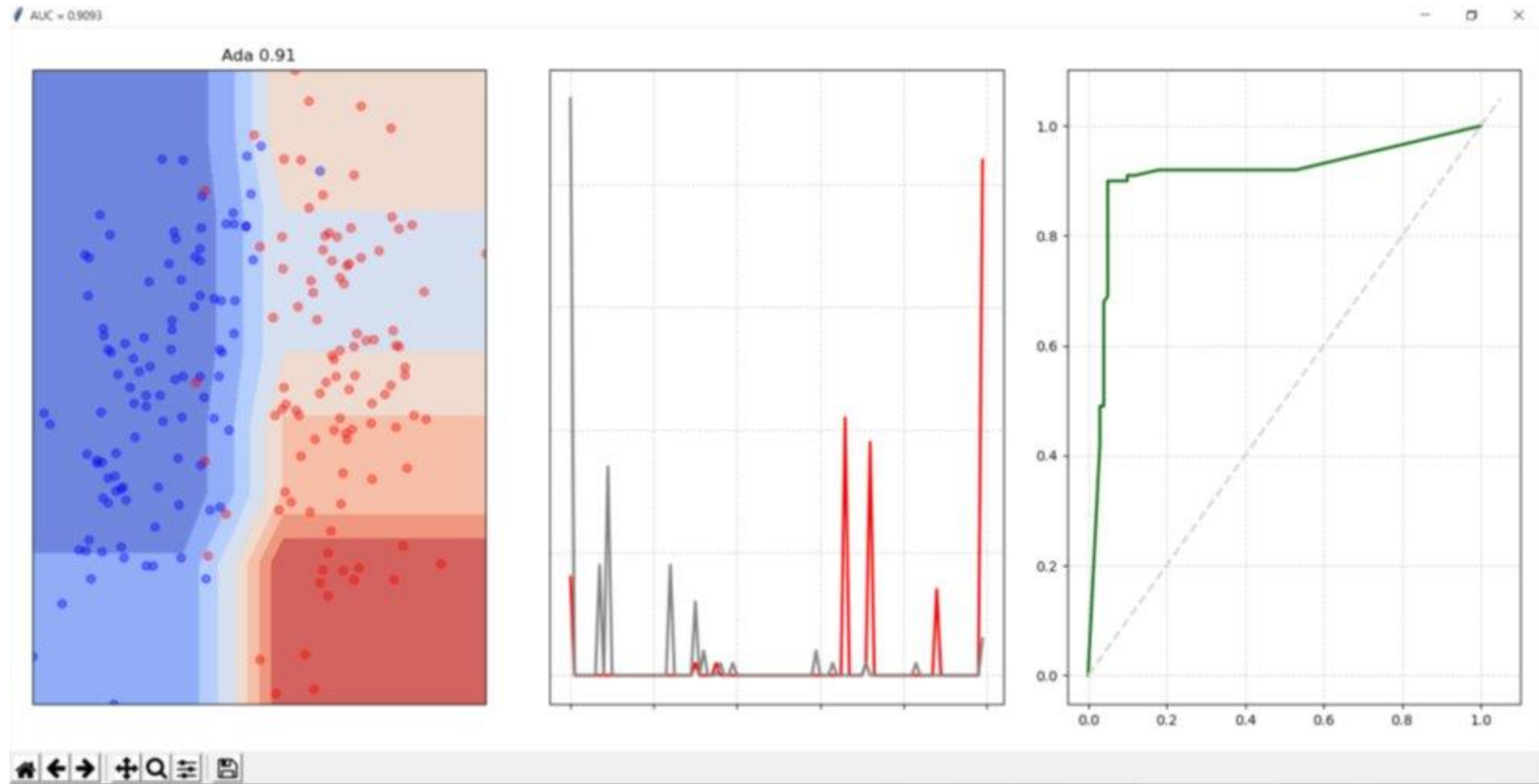
- Regression
- Naive Bayes
- Gaussian
- SVM
- Nearest Neighbors
- Decision Tree
- Random Forest
- AdaBoost
- xgboost
- Neural Net

Benchmarking the Classifiers

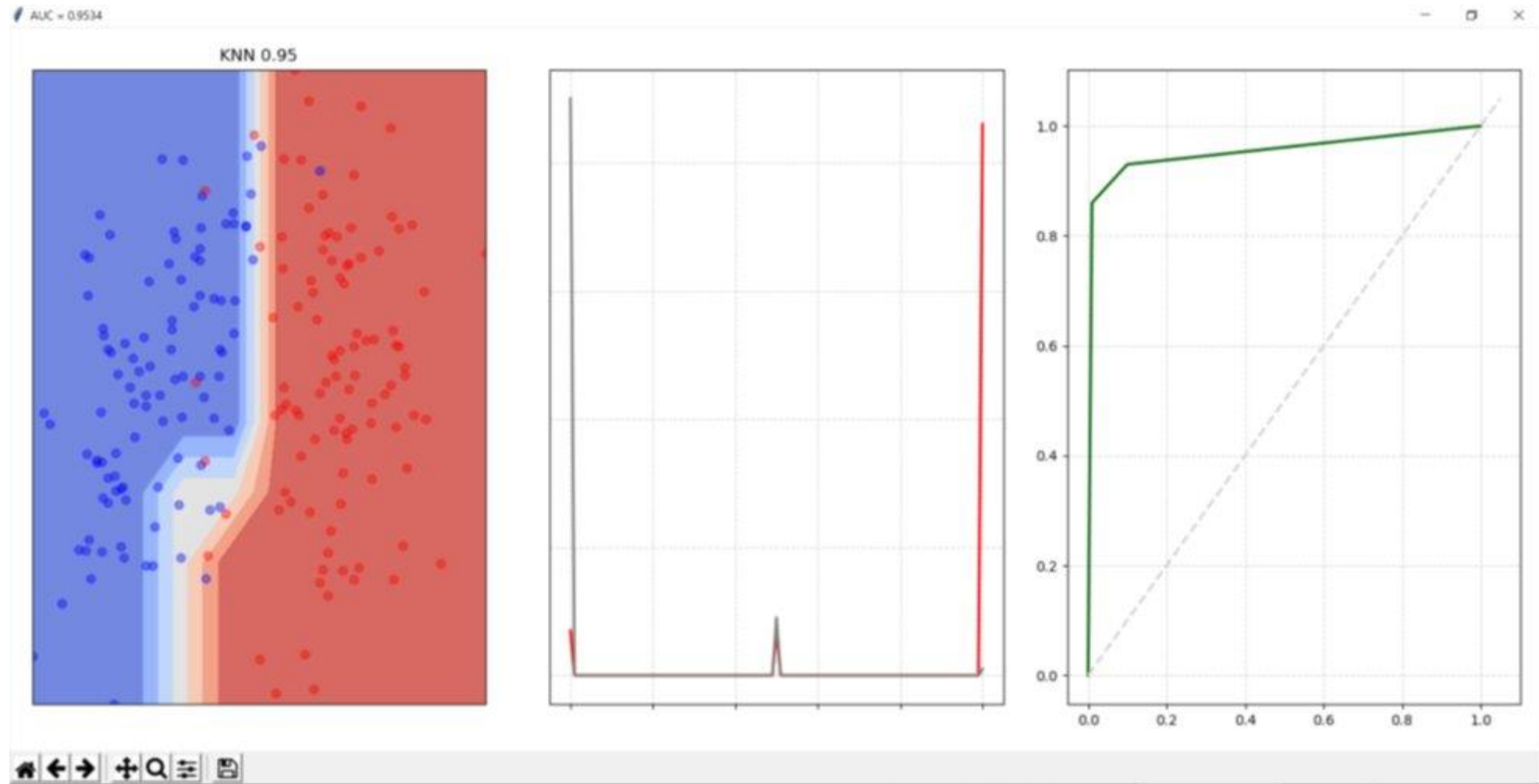
Gaussian distribution



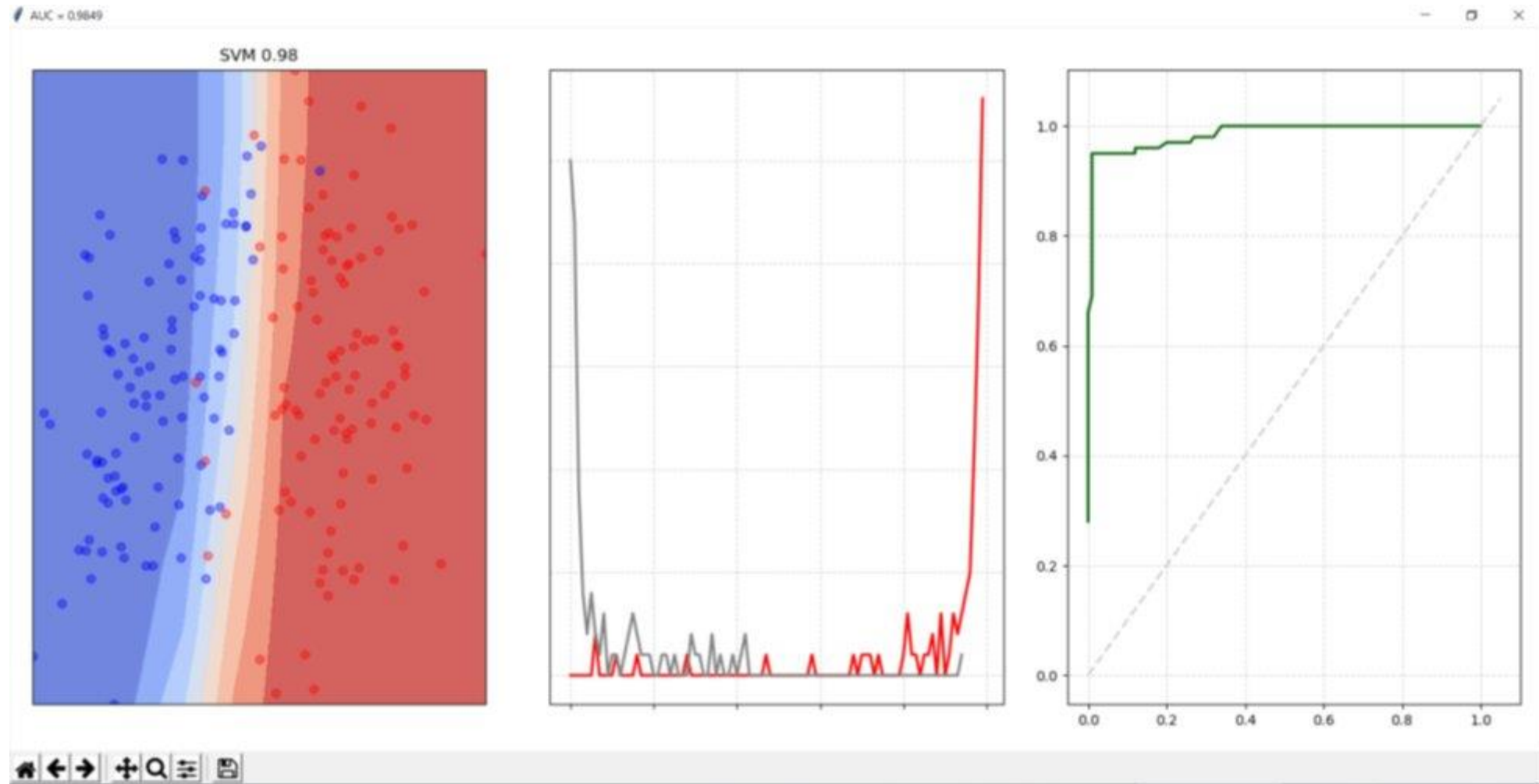
Benchmarking the Classifiers



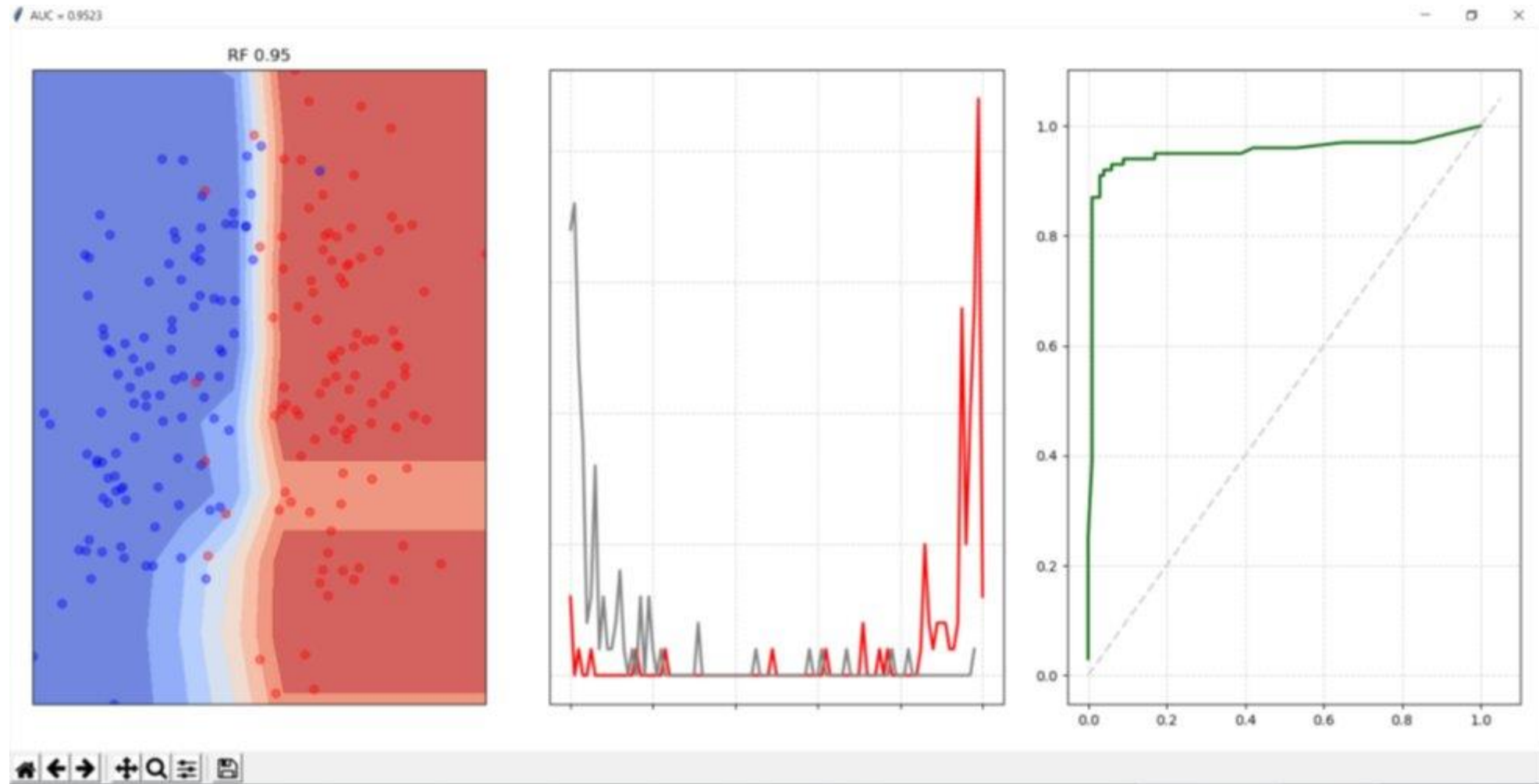
Benchmarking the Classifiers



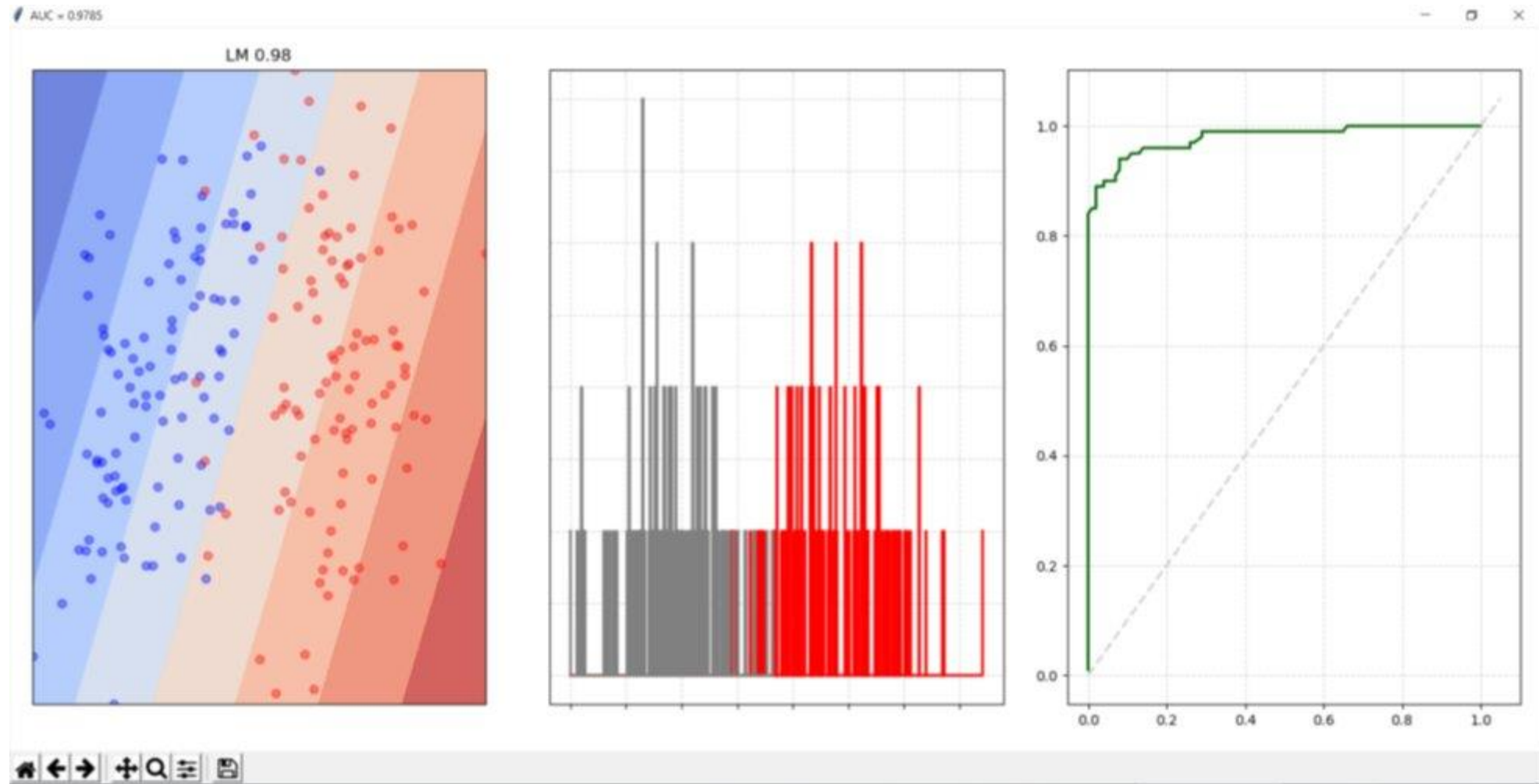
Benchmarking the Classifiers



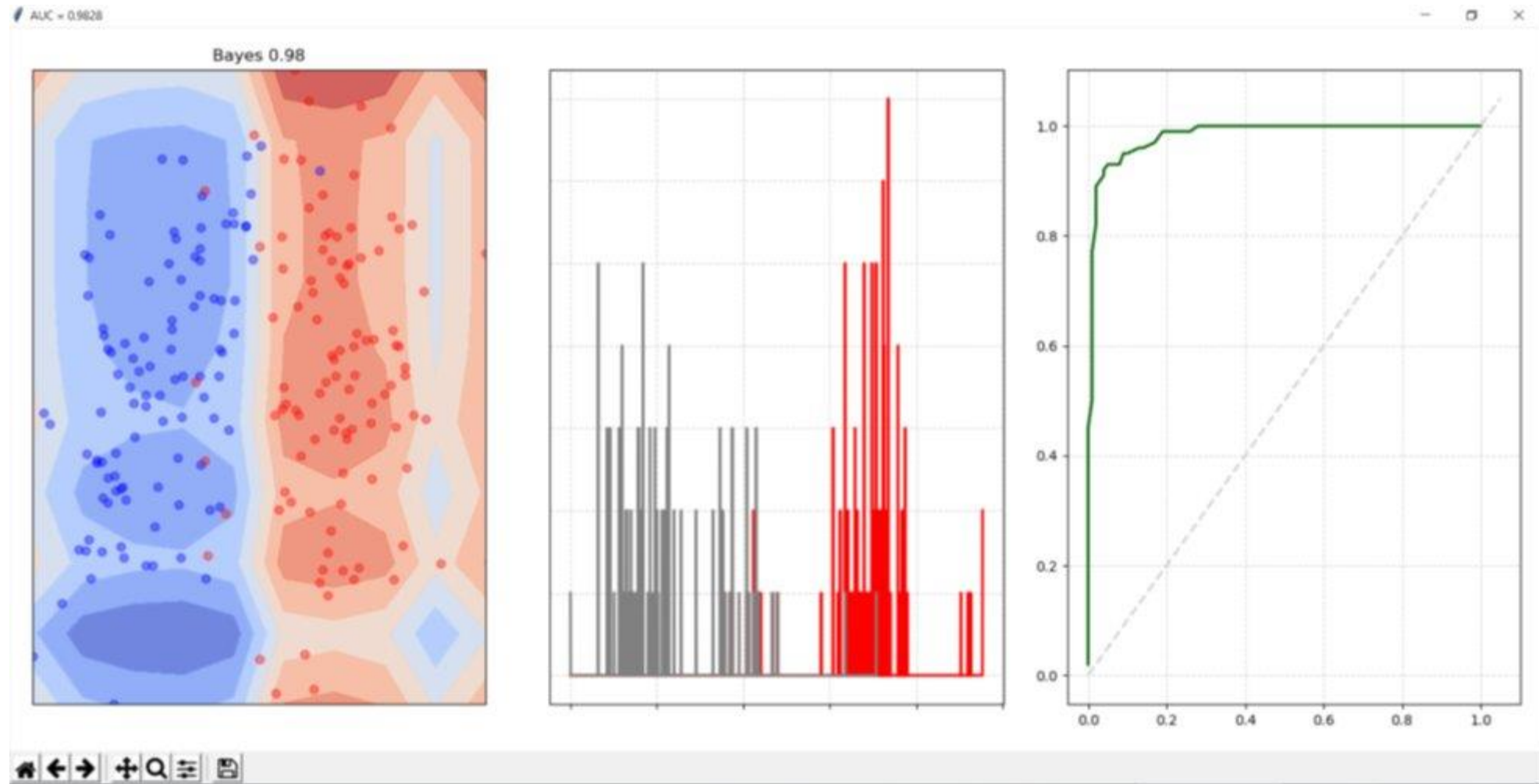
Benchmarking the Classifiers



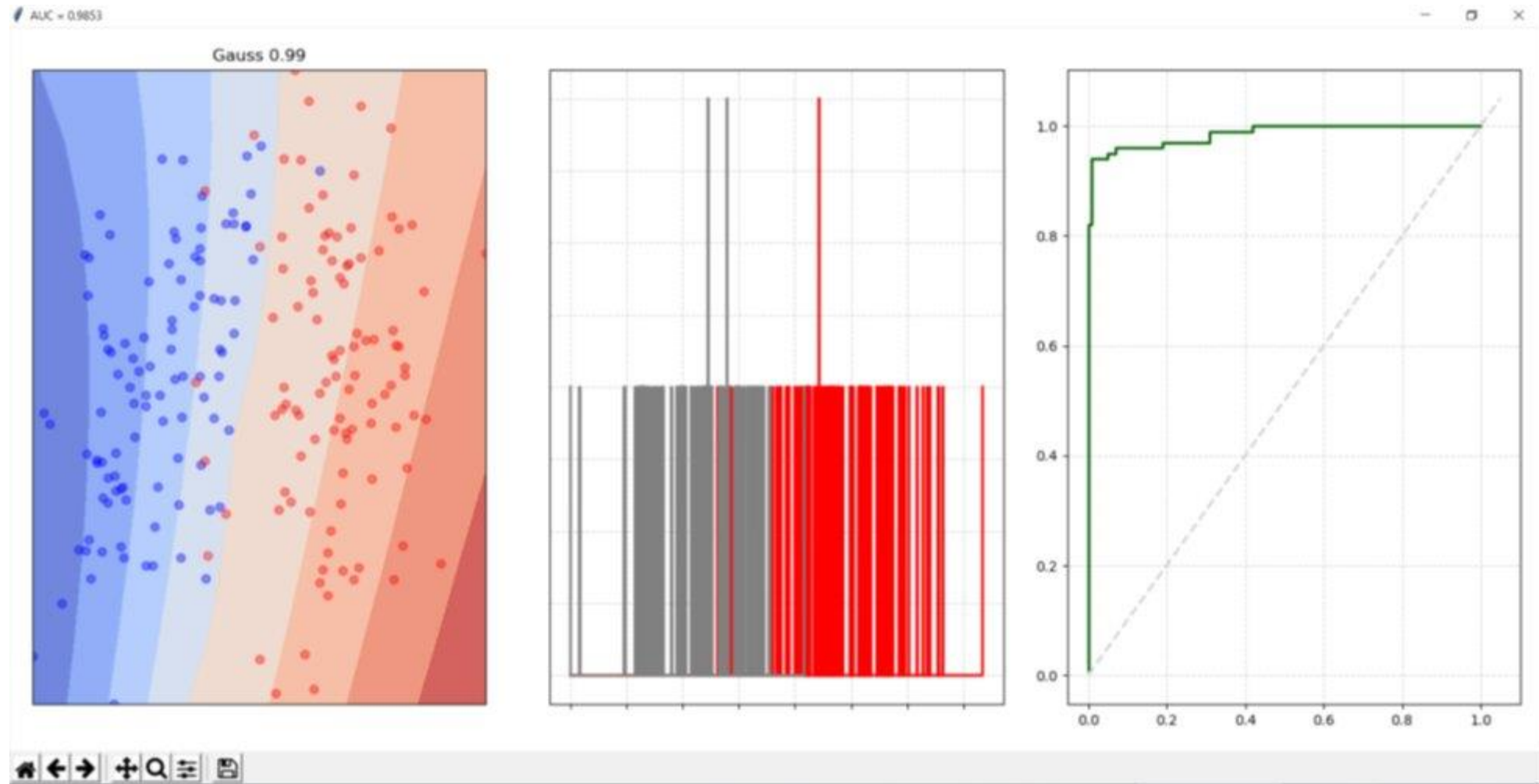
Benchmarking the Classifiers



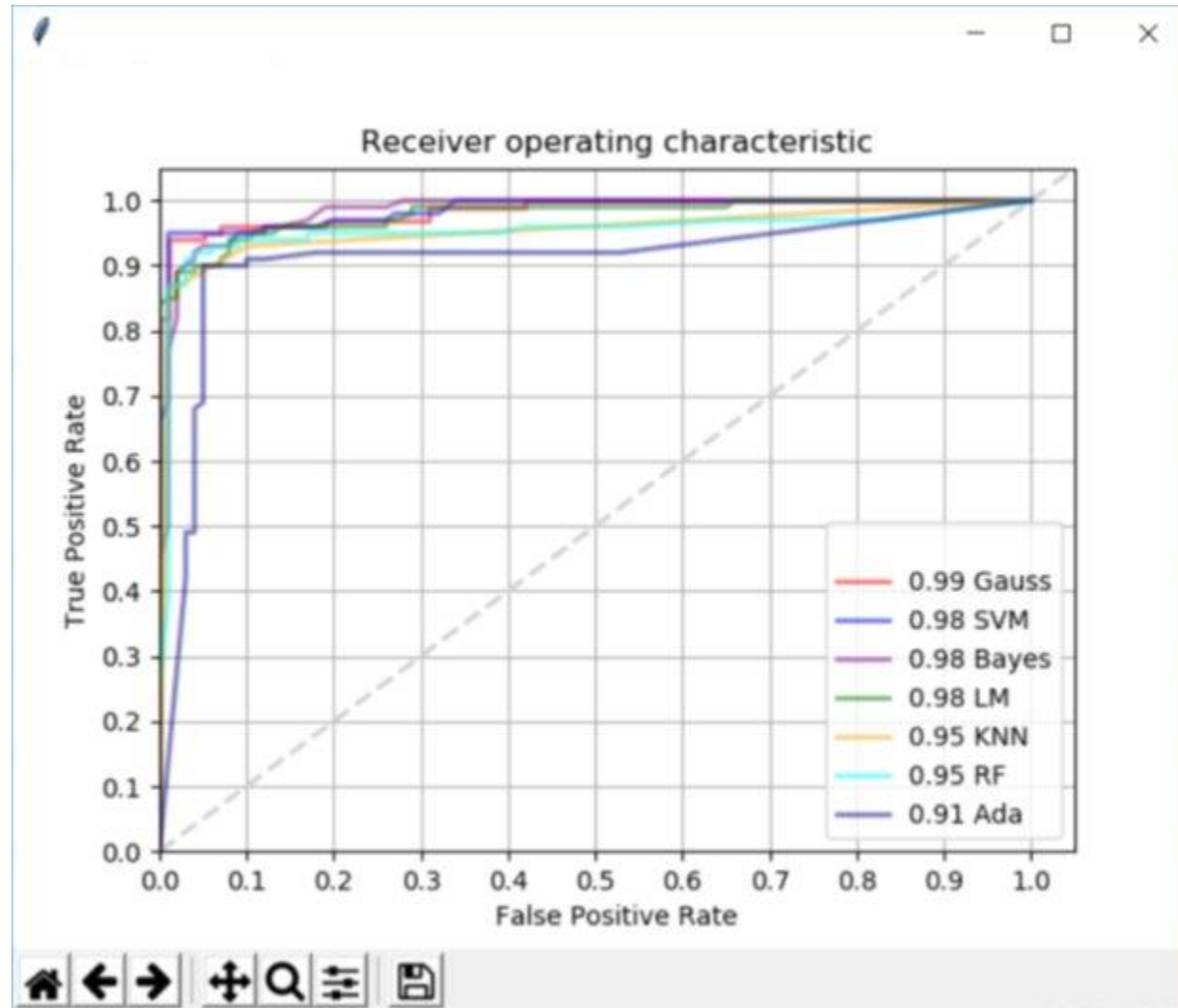
Benchmarking the Classifiers



Benchmarking the Classifiers

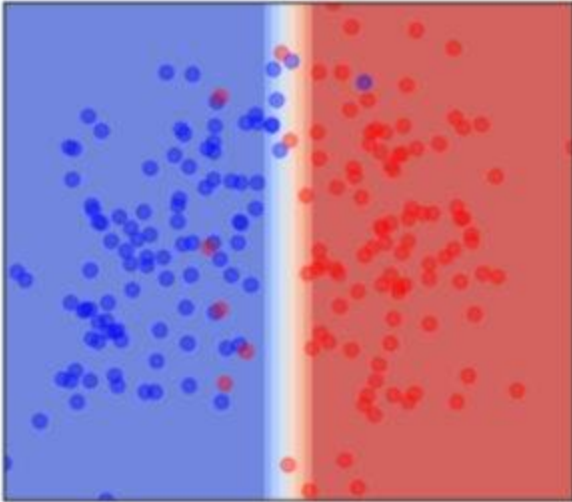


Benchmarking the Classifiers

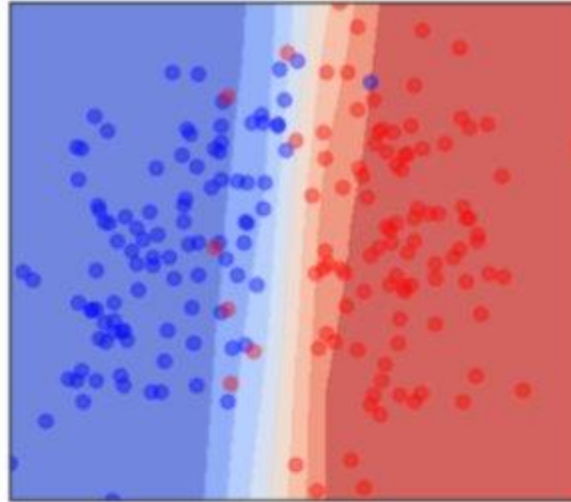


Benchmarking the Classifiers

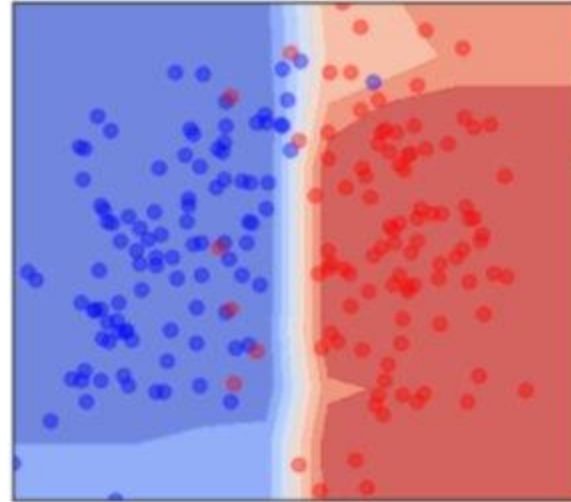
Ada 0.91



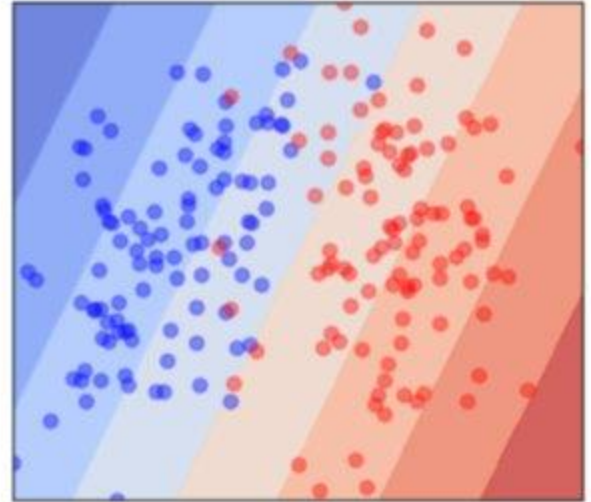
SVM 0.99



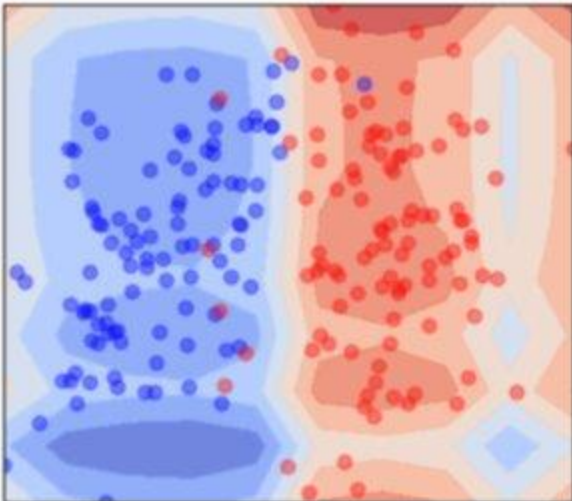
RF 0.95



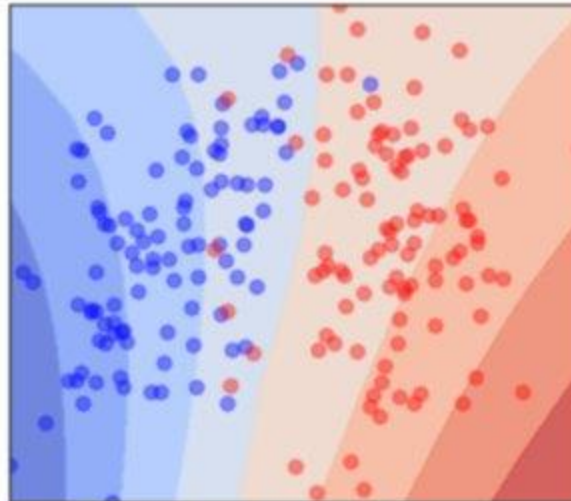
LM 0.98



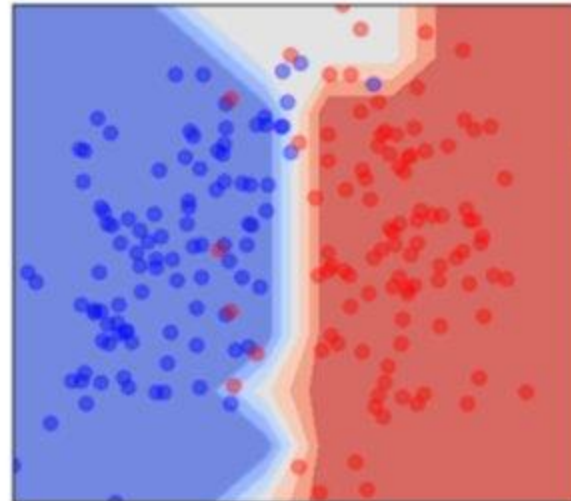
Bayes 0.98



Gauss 0.99

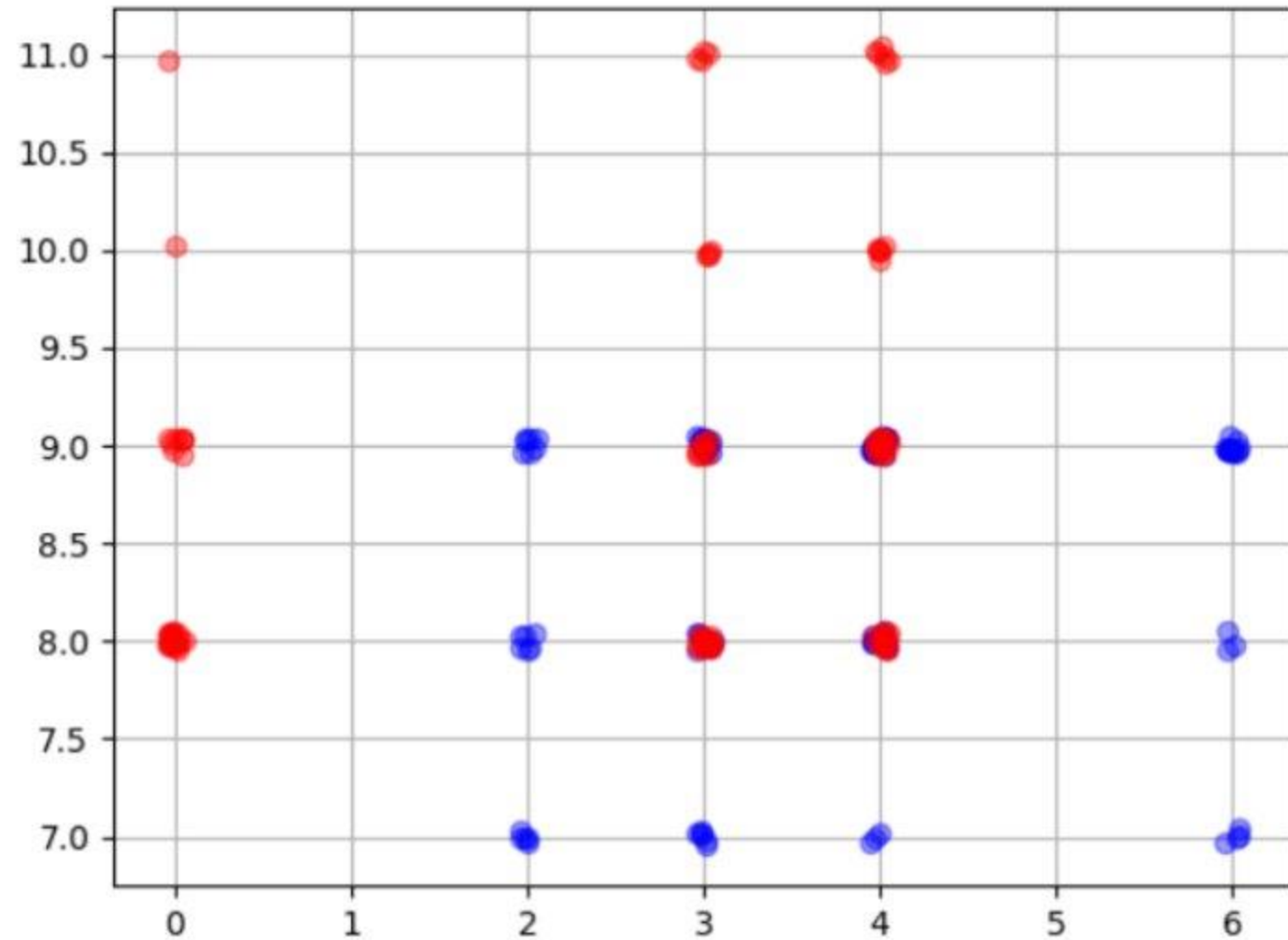


KNN 0.95



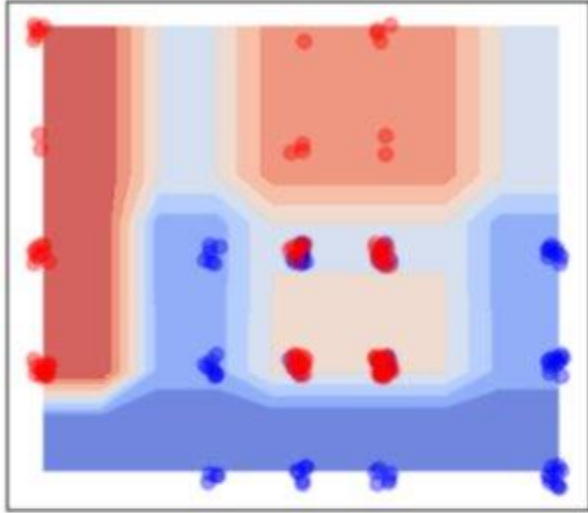
Benchmarking the Classifiers

Non-gaussian distribution

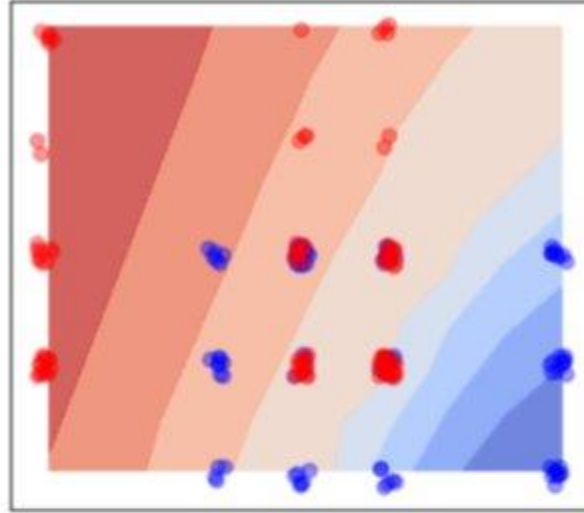


Benchmarking the Classifiers

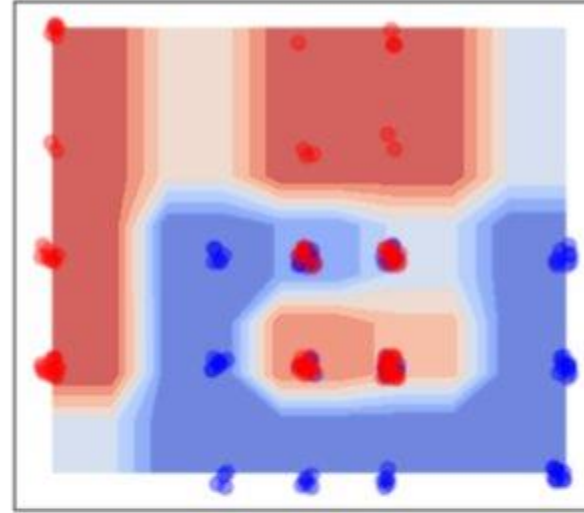
Ada 0.88



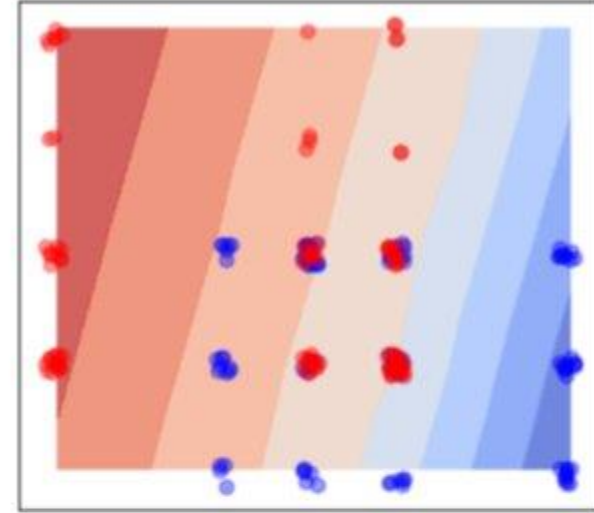
SVM 0.74



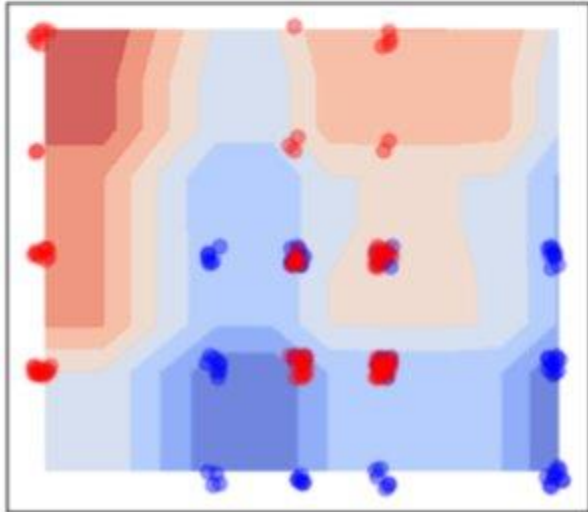
RF 0.88



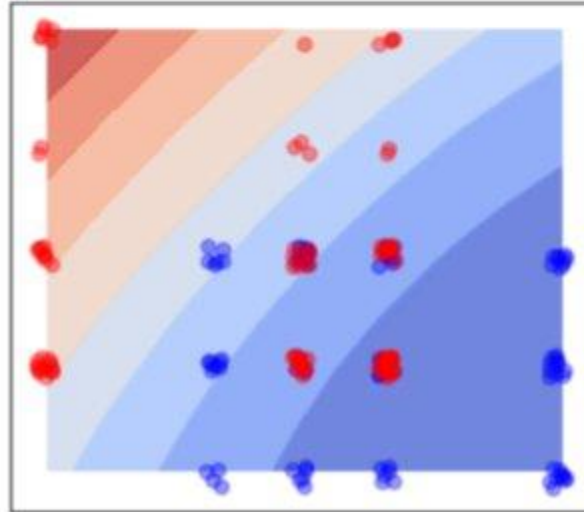
LM 0.71



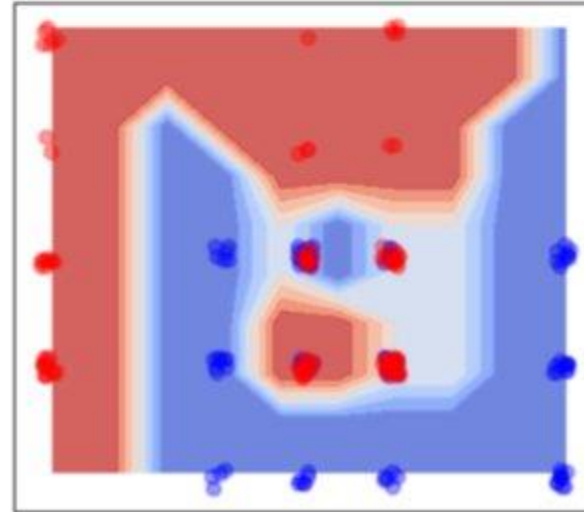
Bayes 0.90



Gauss 0.73



KNN 0.83



Benchmarking the Classifiers

