Basics of Machine Learning
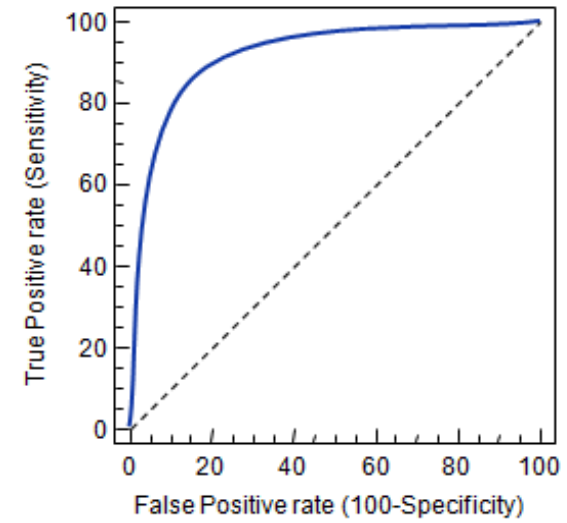
Dmitry Ryabokon, github.com/dryabokon

Lesson 17
Benchmarking

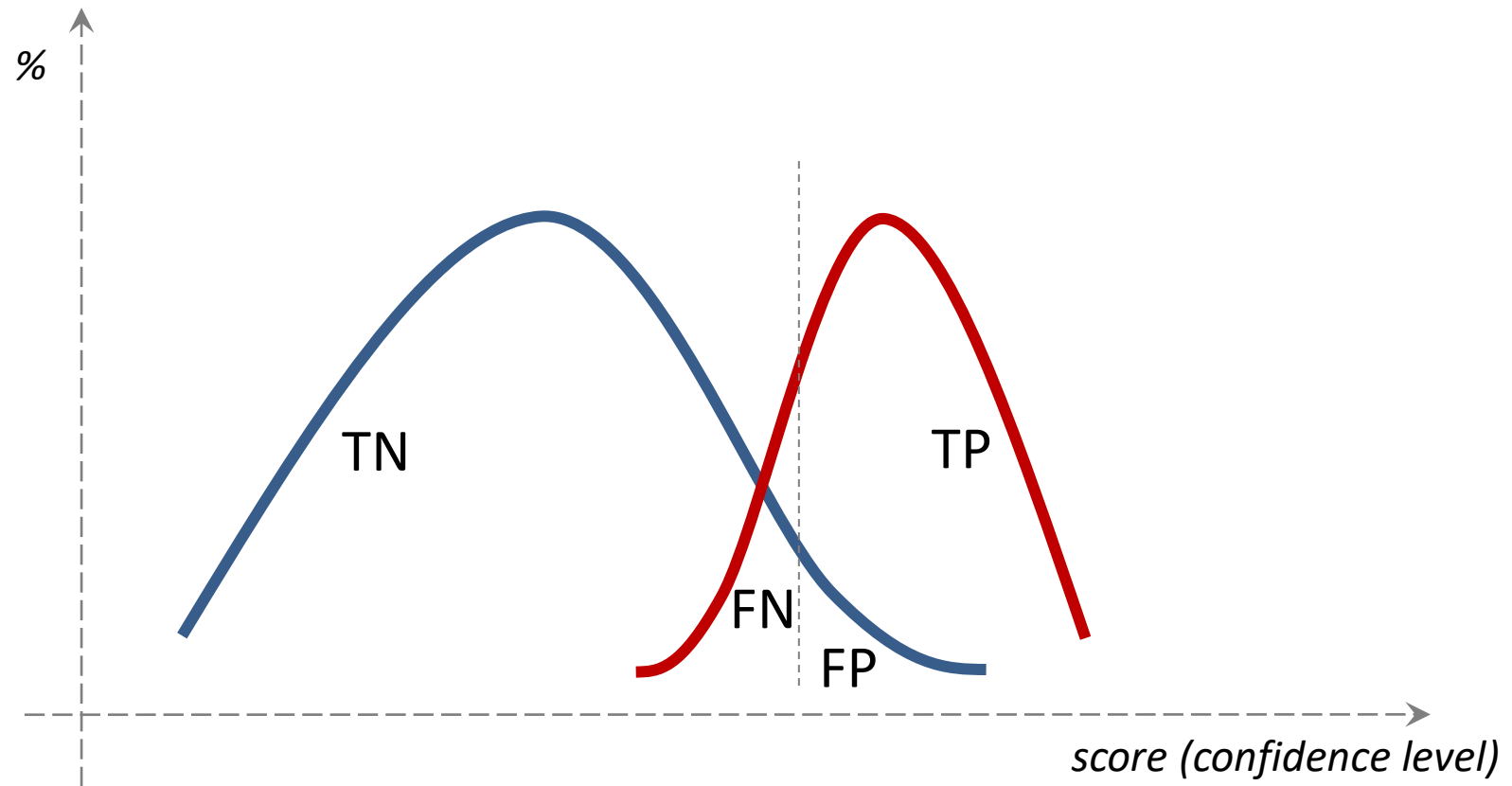# Benchmarking

Summary

- ROC Curve
- PR-Curve
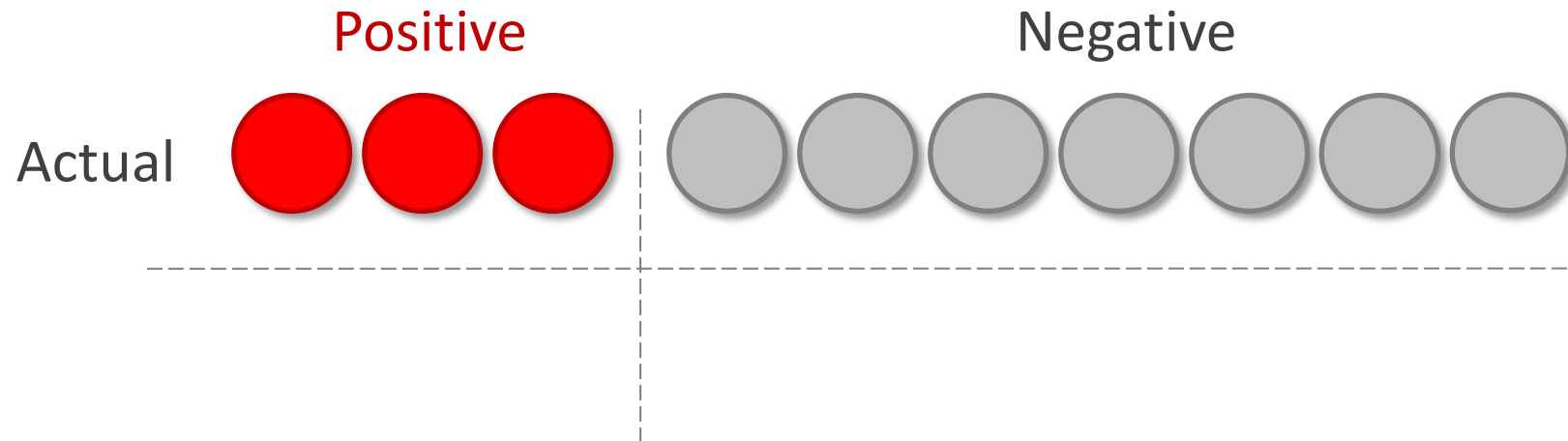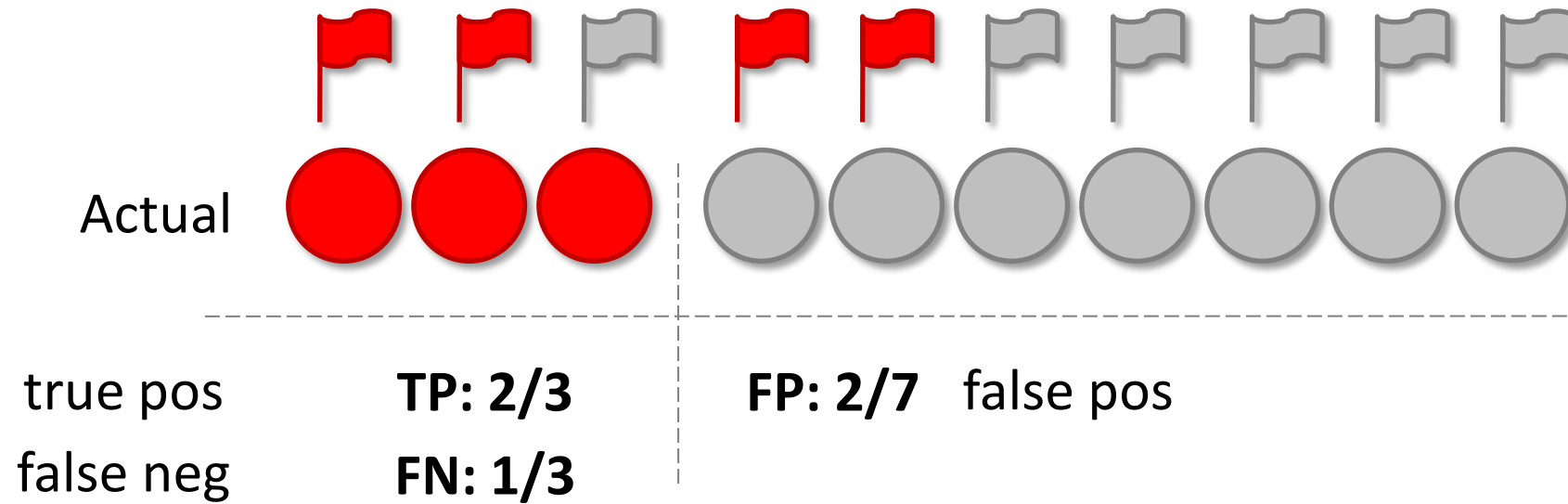- Benchmarking the classifiers

ROC curve

ROC curve

ROC curve

Error types

ROC curve

Error types

Actual

true pos   **TP: 2/3**   **FP: 2/7**   false pos

false neg   **FN: 1/3**
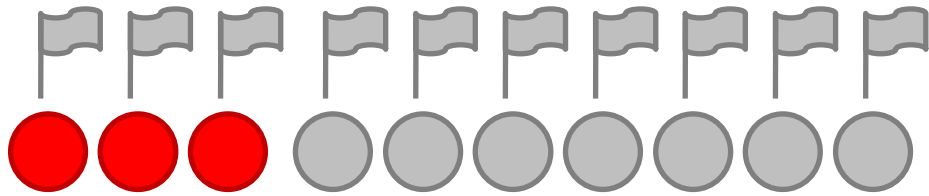
ROC curve

TP: 0/3   **FP: 0/7**

FN: 3/3
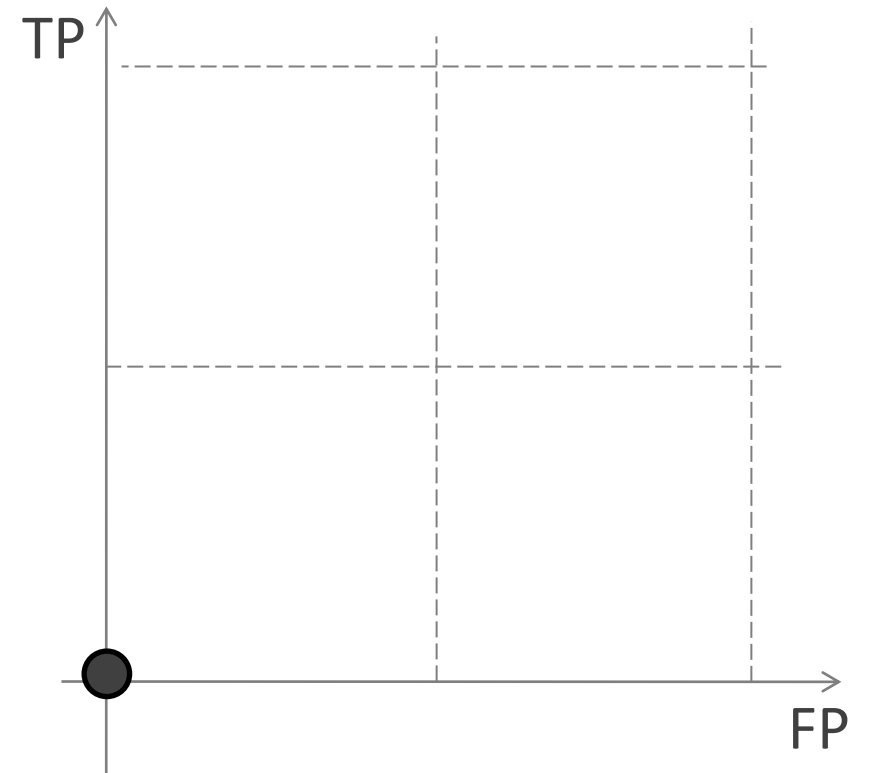
ROC curve

TP: 1/3   **FP: 0/7**

FN: 2/3

TP

FP
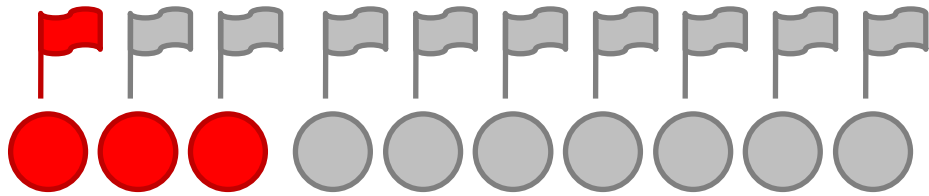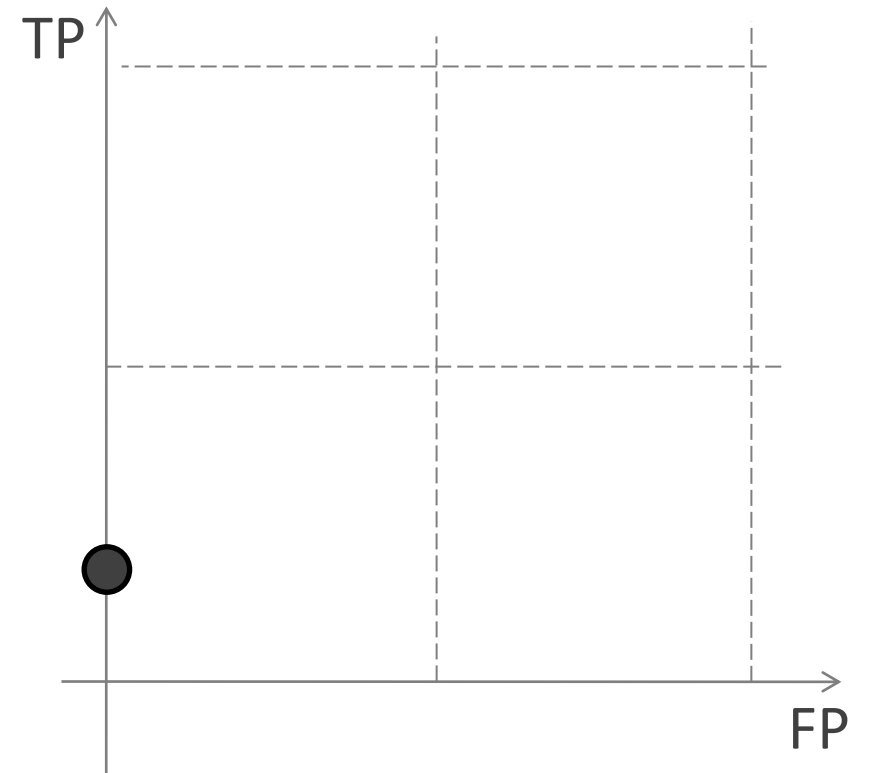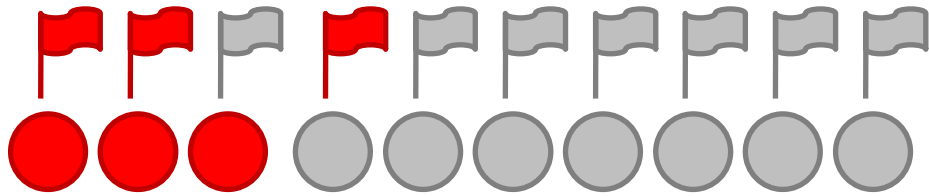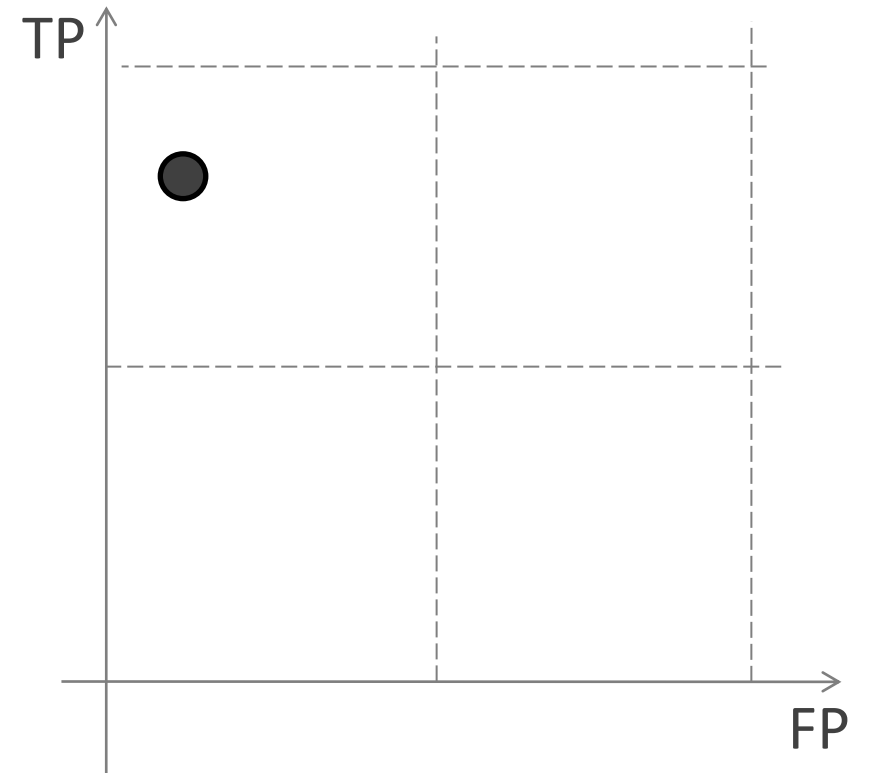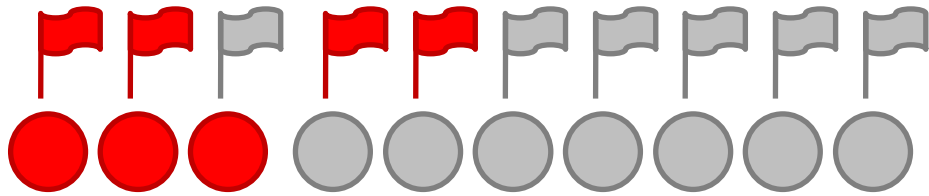
ROC curve

ROC curve

TP: 2/3   **FP: 2/7**

FN: 1/3
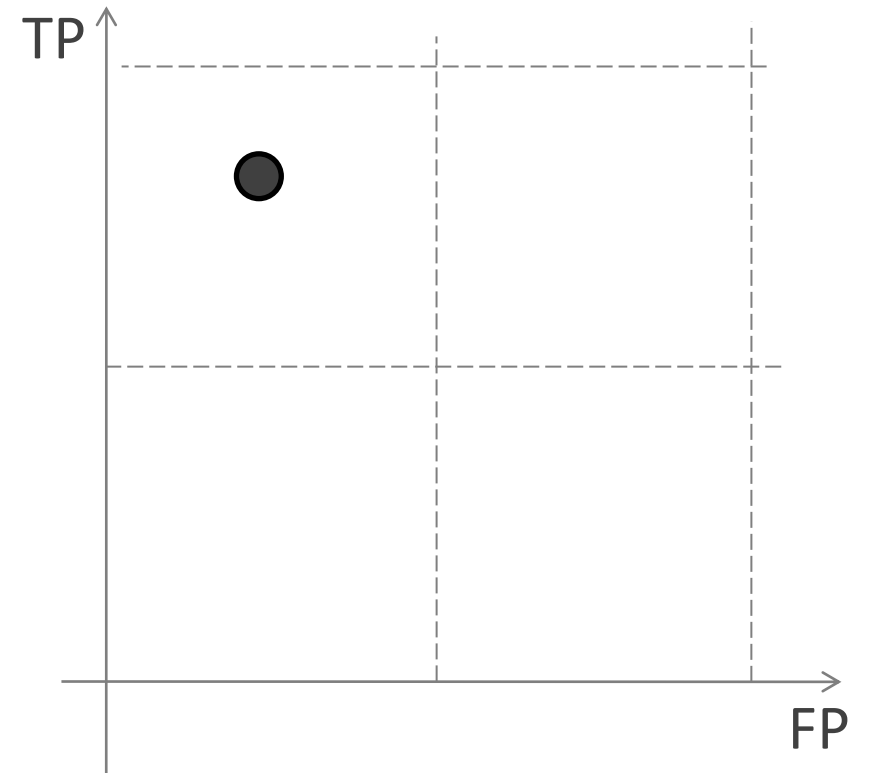
ROC curve
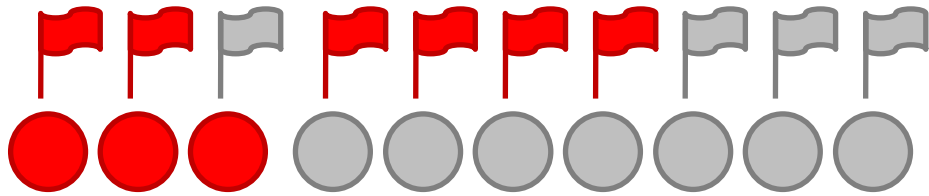
TP: 2/3 **FP: 4/7**

FN: 1/3
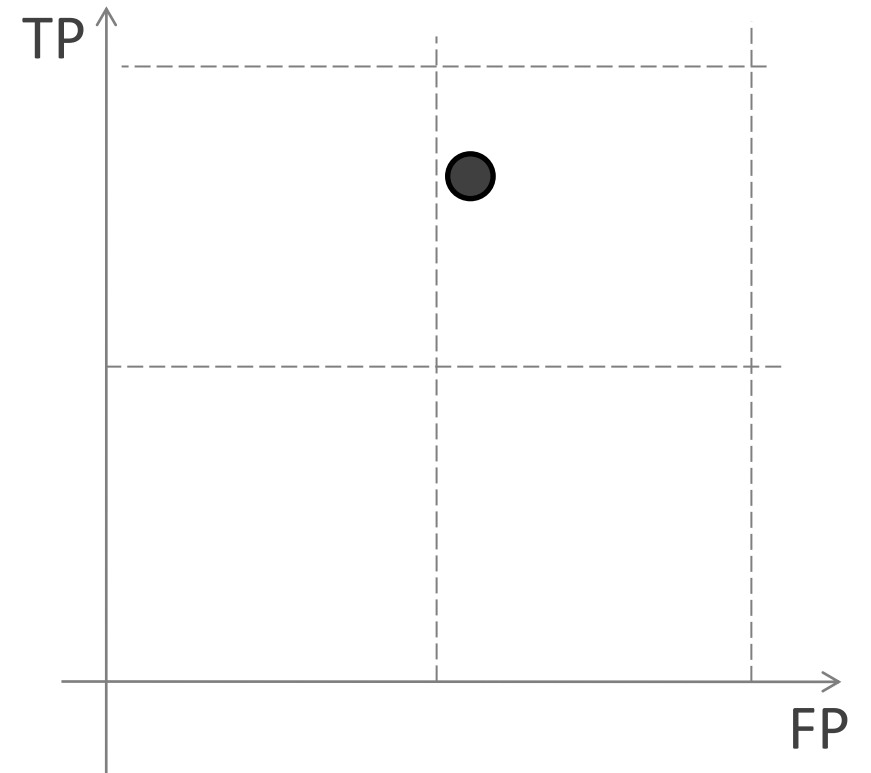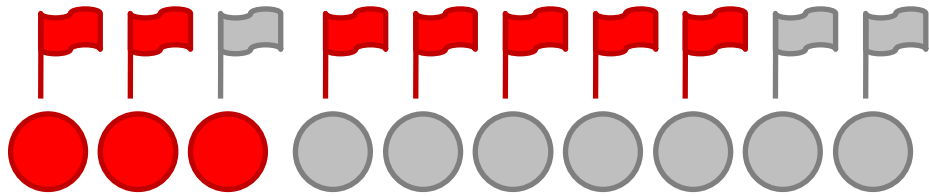
TP: 3/3　　**FP: 7/7**

FN: 0/3
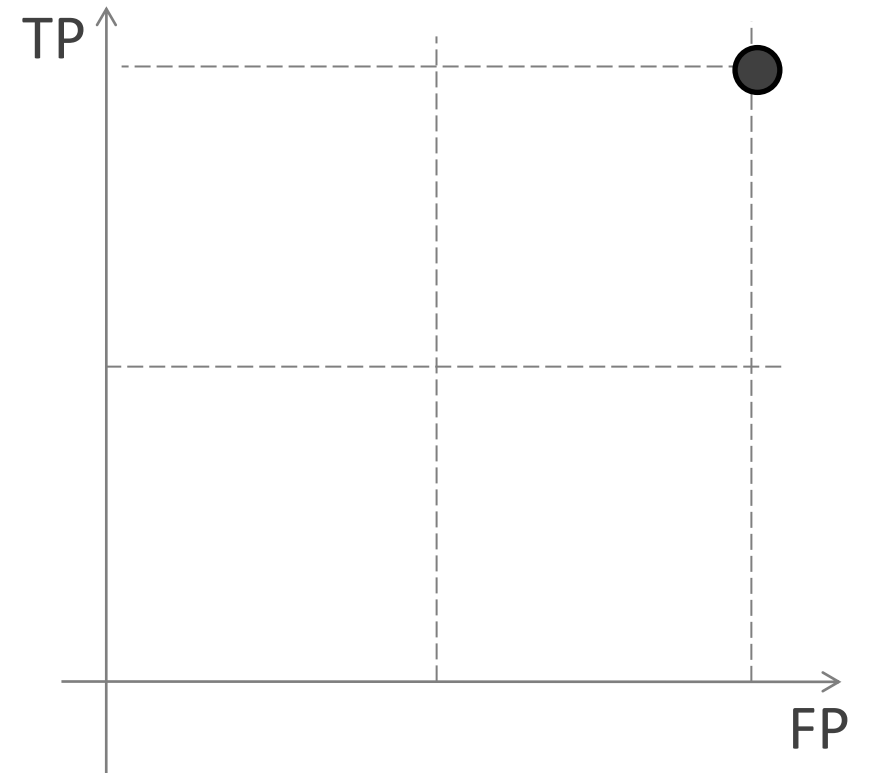
AUC: area under ROC curve
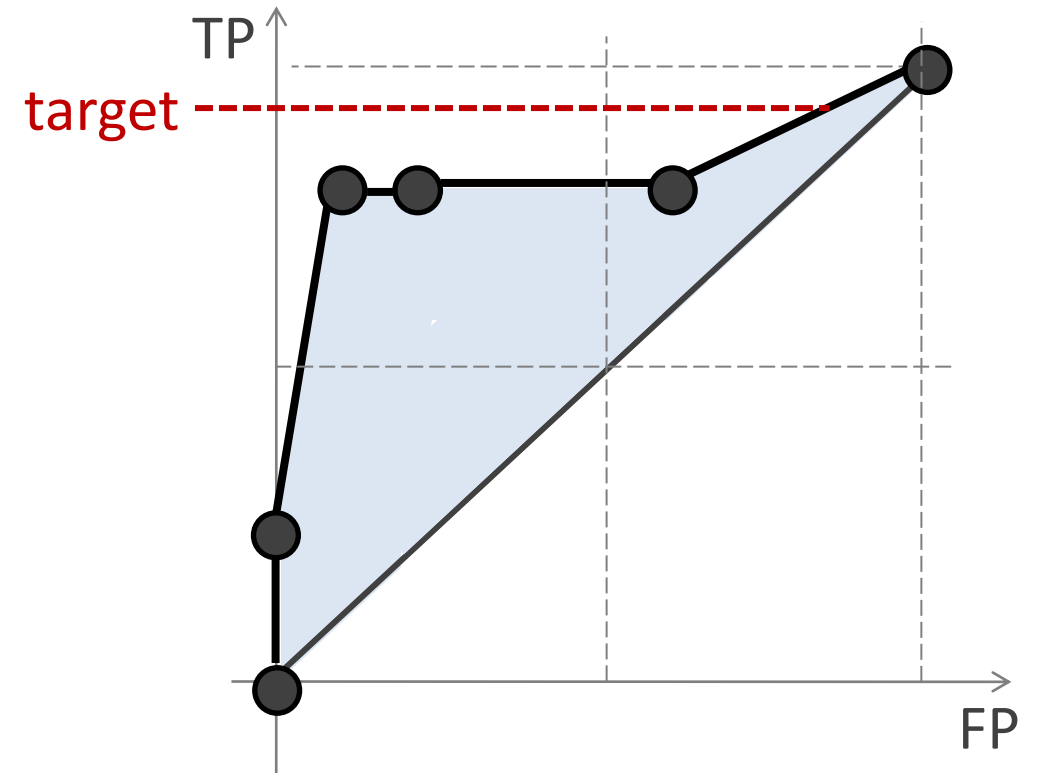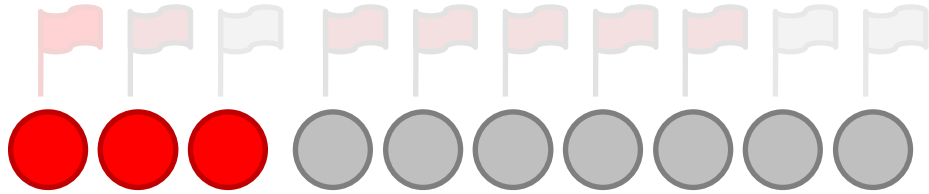
Example: Naive Bayesian classifier

Example: Naive Bayesian classifier

Example: Naive Bayesian classifier

$$k$$

$$\frac{p(1|x)}{p(2|x)}$$

6.00 6.00 6.00 6.00 1.50 1.50 1.00 0.75 0.75 0.25 0.25 0.25 0.19 0.19 0.19
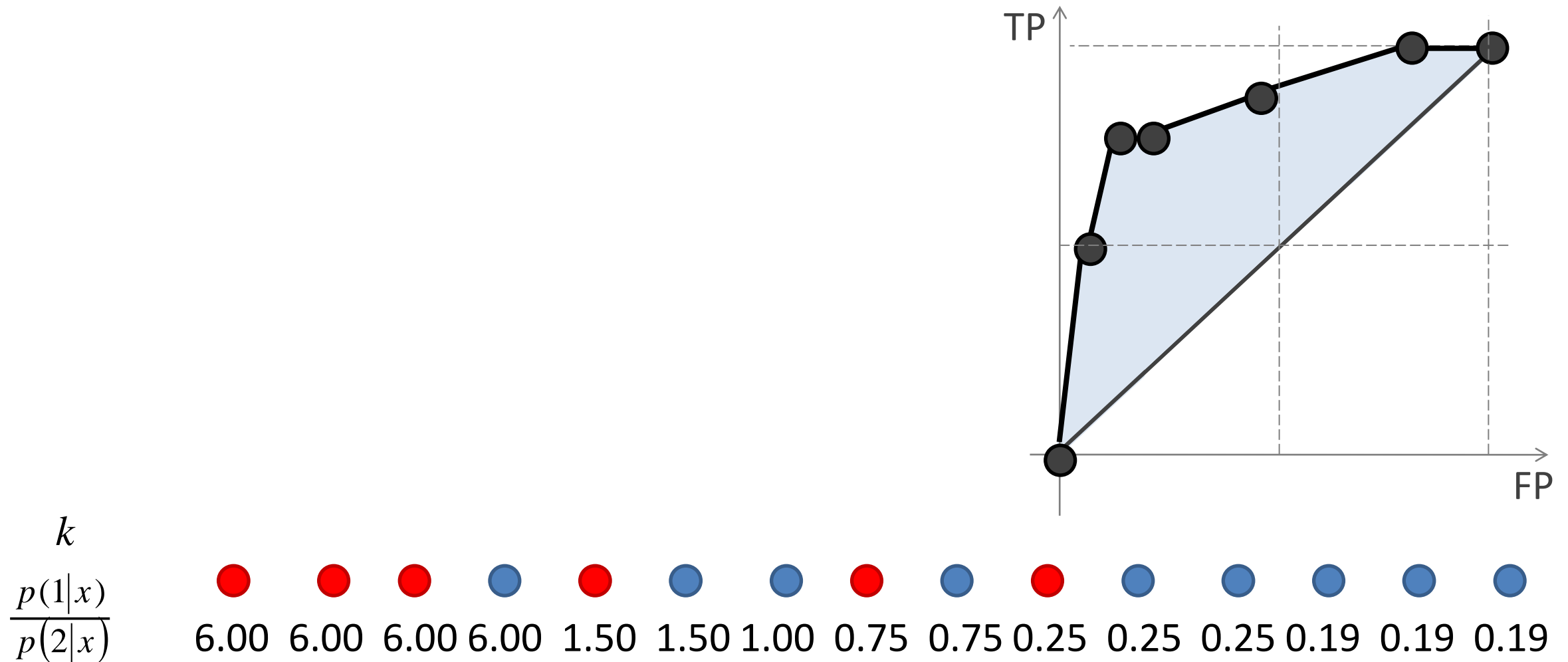
ROC curve

The task: compare accuracy of two strategies

Strategy 1:	63	81	23	54	10	17	29	5	34	21
Strategy 2:	86	98	18	42	21	23	37	18	74	27
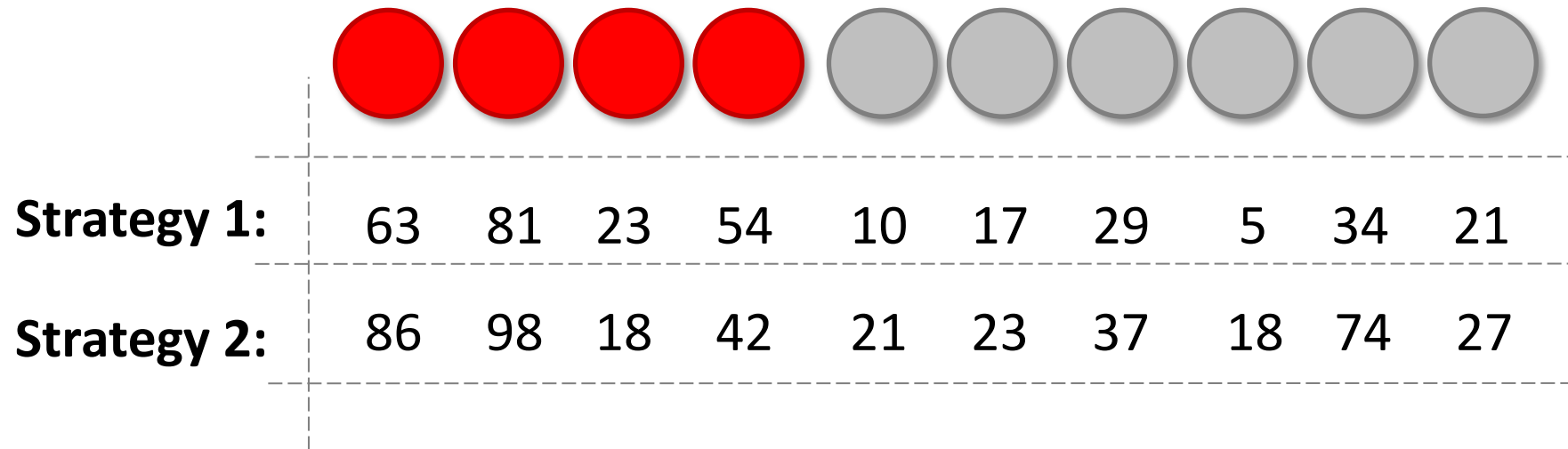
Precision-Recall

Precision - Recall

Precision: TP / (TP+FP)
Recall: TP / (TP+FN)

Accuracy: (TP+TN) /Total
F1-score: 2 x Precision x Recall / (Precision+ Recall)
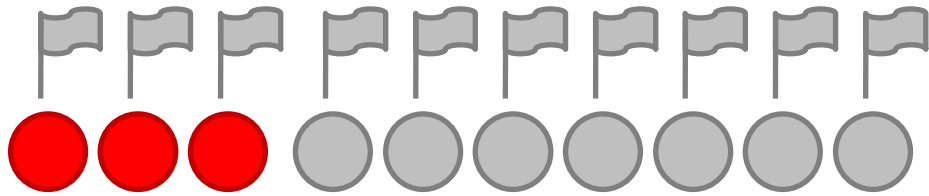
Precision - Recall

TP: 0   FP: 0   **Precision: 0/0**

FN: 3

Recall: 0/3
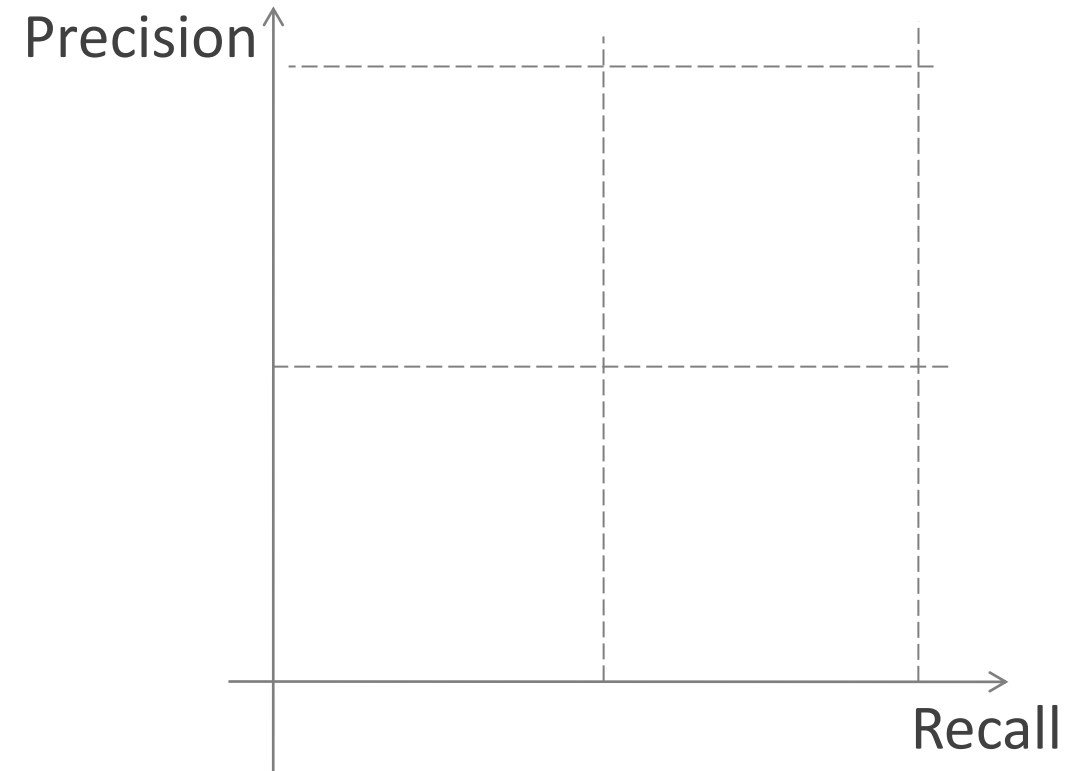
Precision - Recall
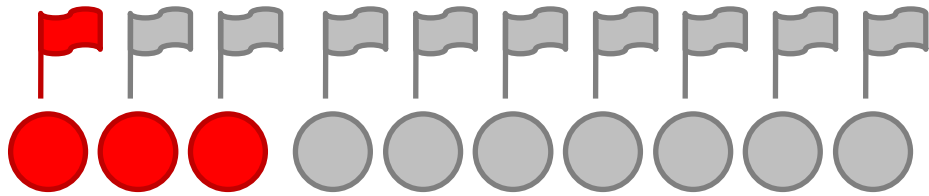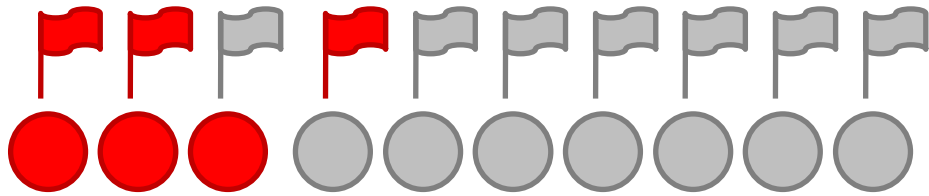

TP: 1   **FP: 0**   **Precision: 1**

FN: 2

Recall: 1/3

Precision - Recall

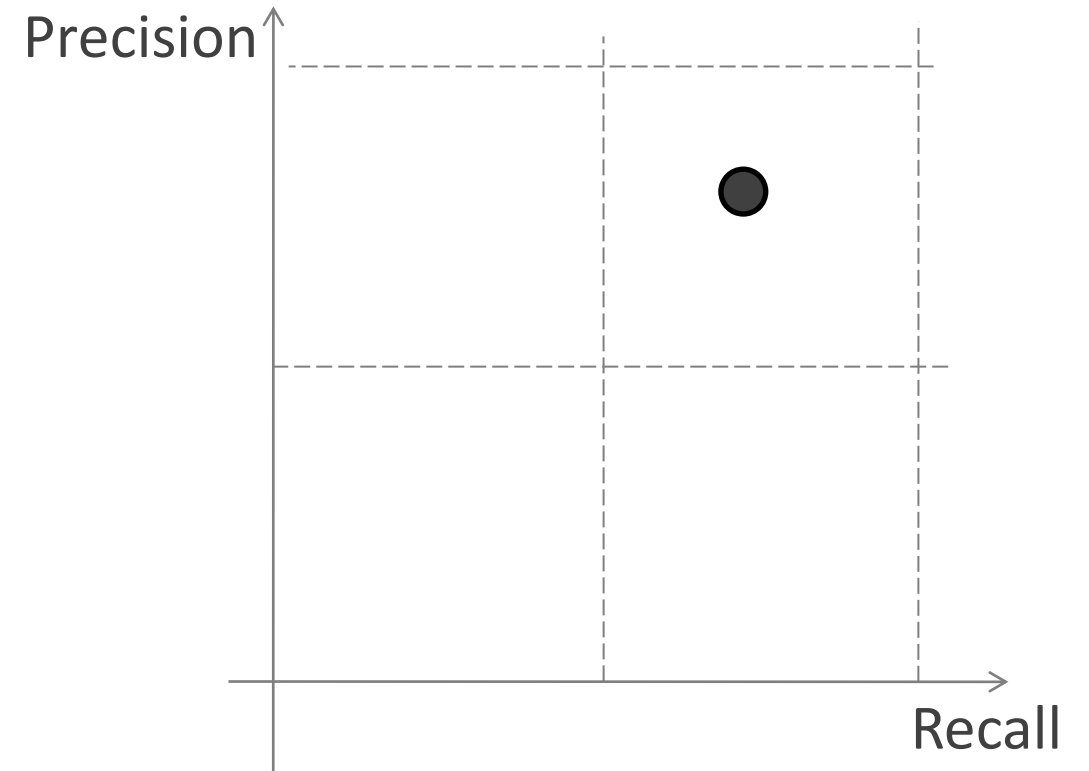TP: 2　　FP: 1　　Precision: 2/3

FN: 1
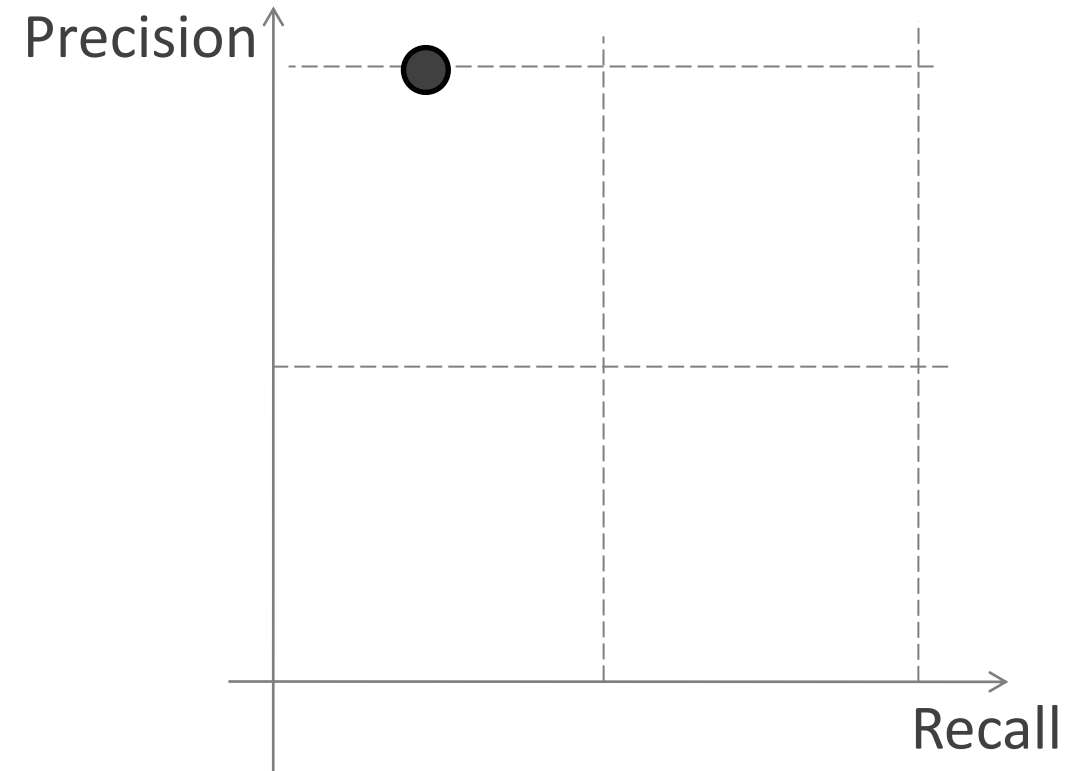
Recall: 2/3
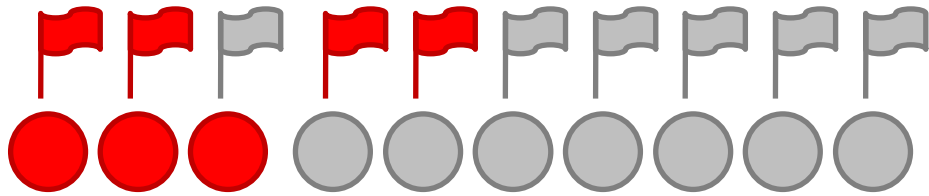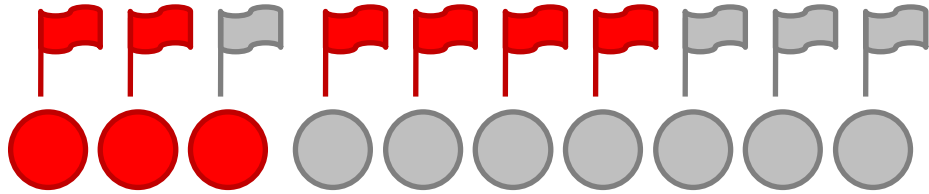
Precision - Recall

TP: 2   **FP: 2**   **Precision: 1/2**

FN: 1

Recall: 2/3
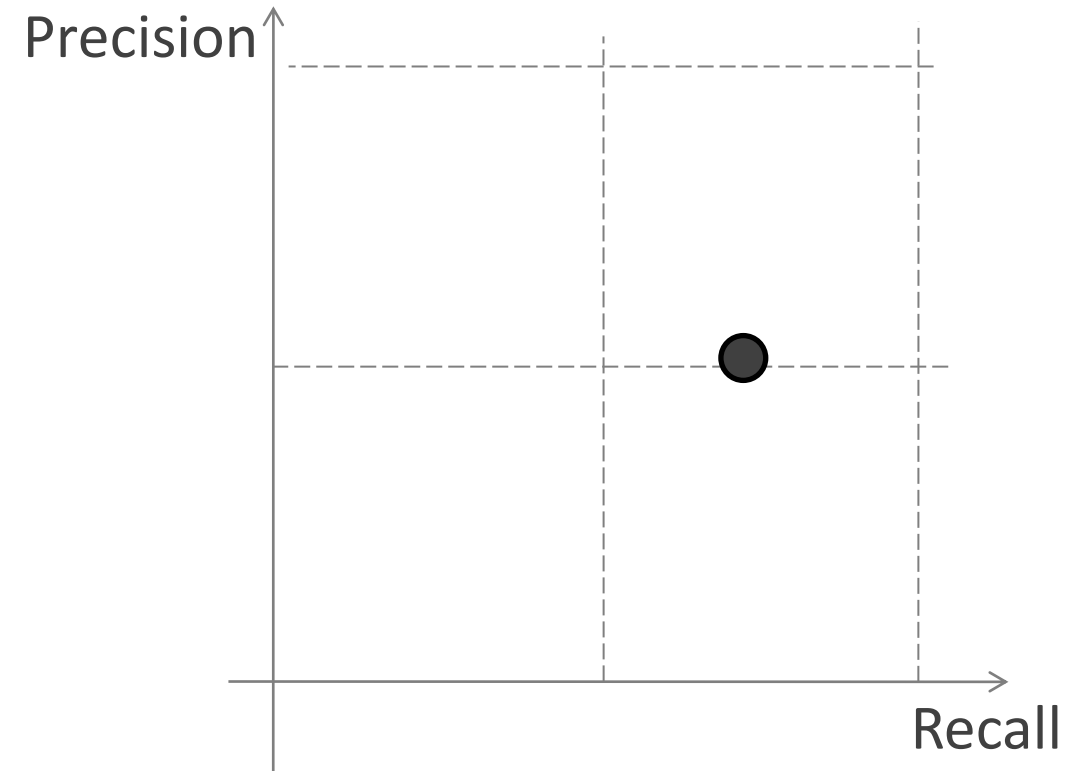
Precision - Recall

TP: 2   **FP: 4**   **Precision: 2/6**

FN: 1

Recall: 2/3
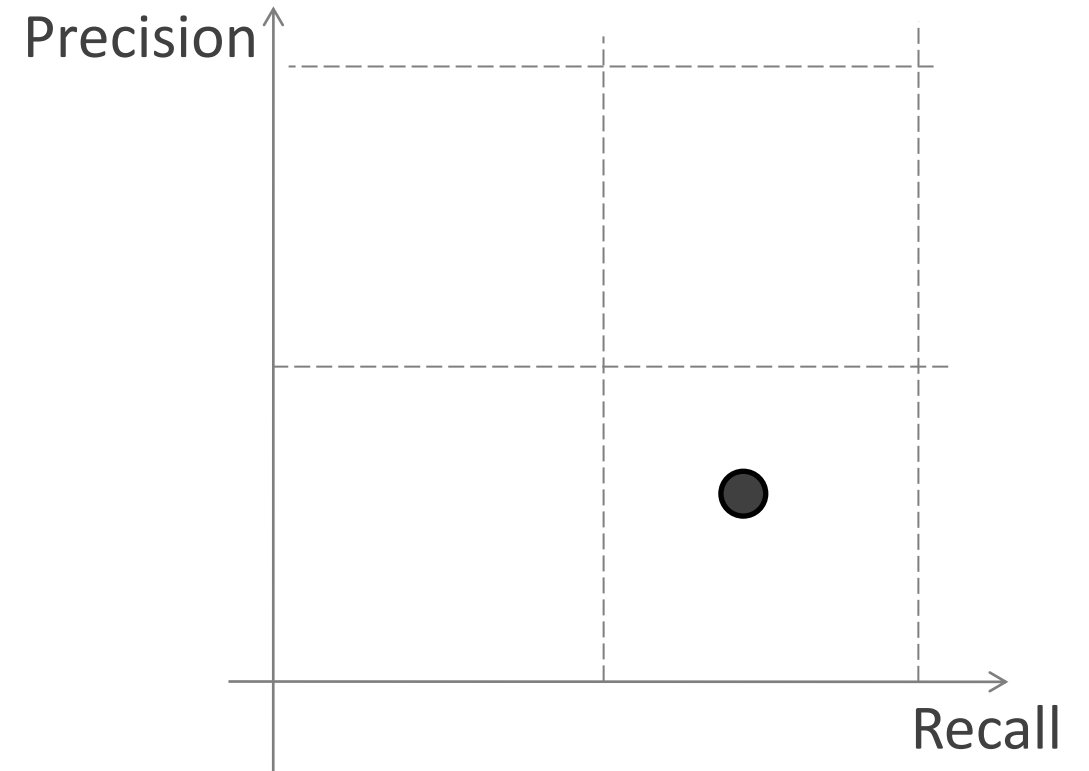
Precision - Recall
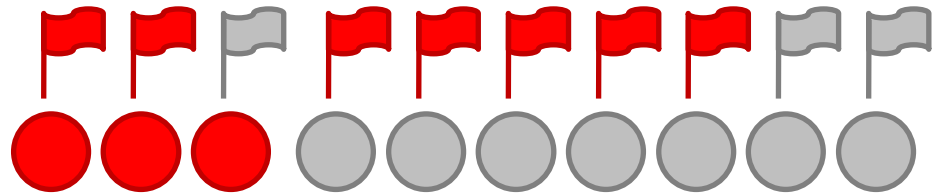
TP: 2 | FP: 5 | Precision: 2/7
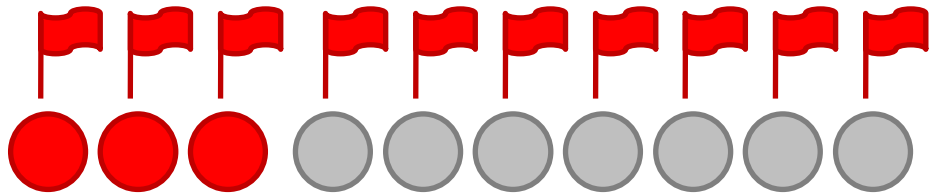
FN: 1

Recall: 2/3
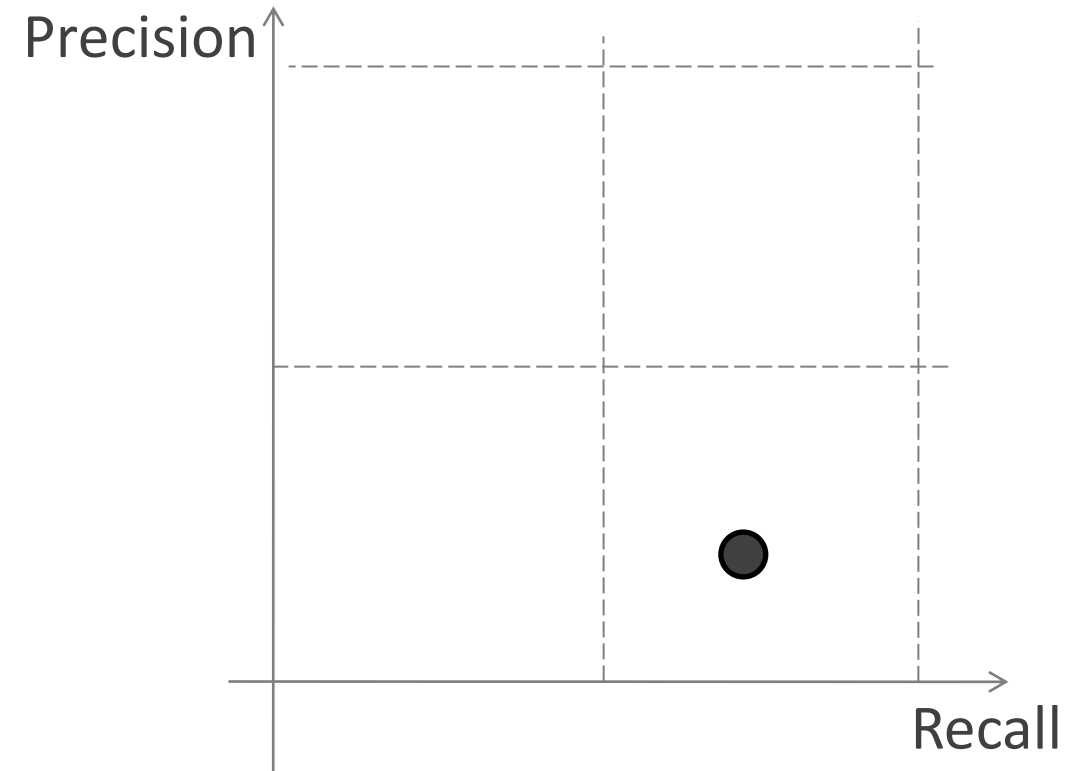
Precision - Recall

TP: 3   FP: 7   Precision: 3/7

FN: 0

Recall: 1
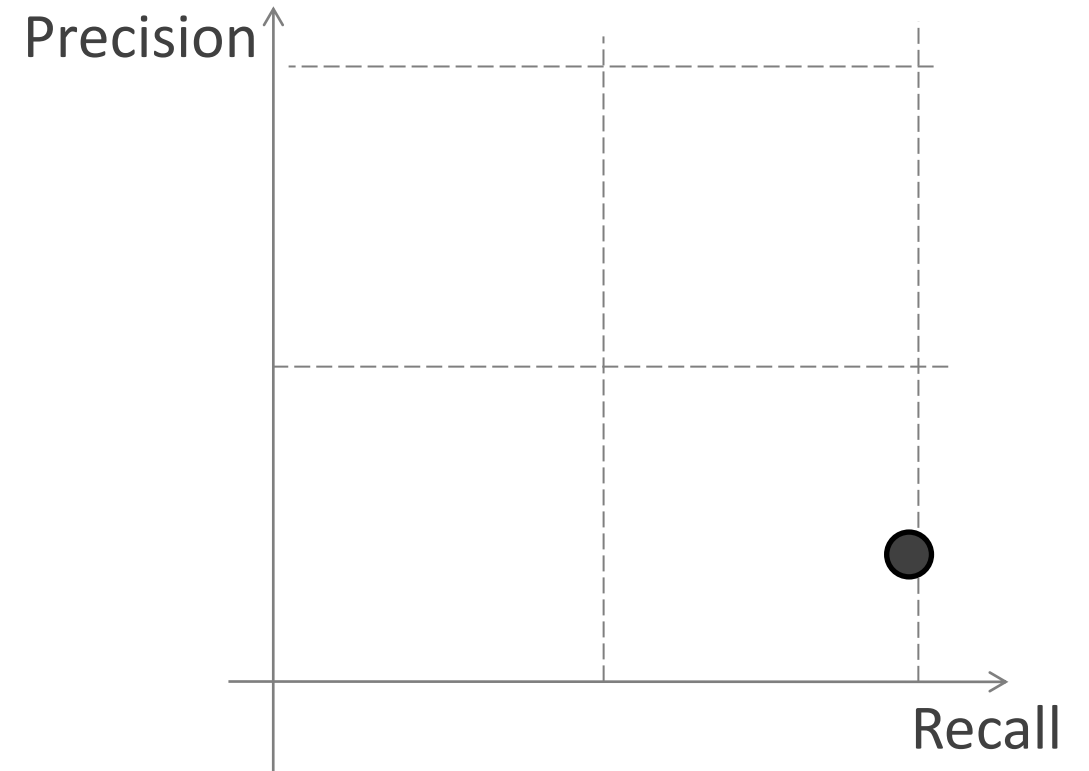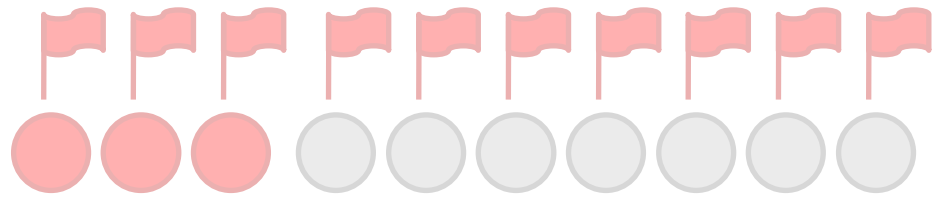
Precision

Recall

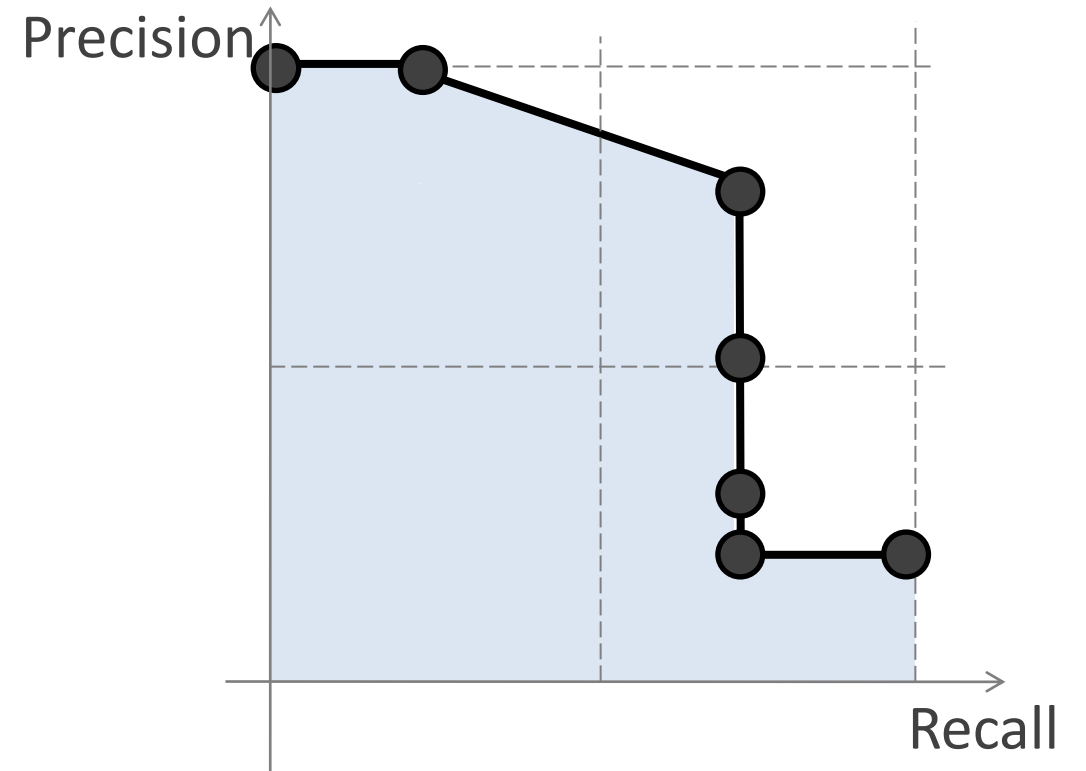# Precision - Recall

TP: 3   FP: 7

FN: 0

Recall: 1

Precision - Recall
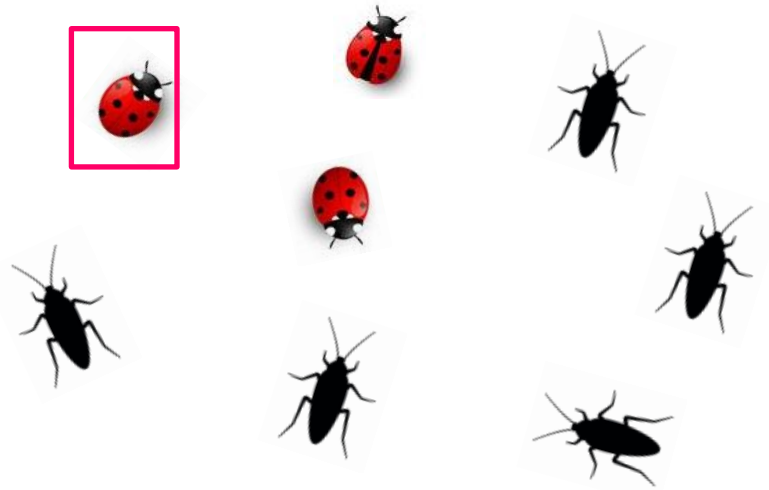
TP: 0   FP: 0   Precision: 0

FN: 3

Recall: 0

Precision

Recall

Precision - Recall
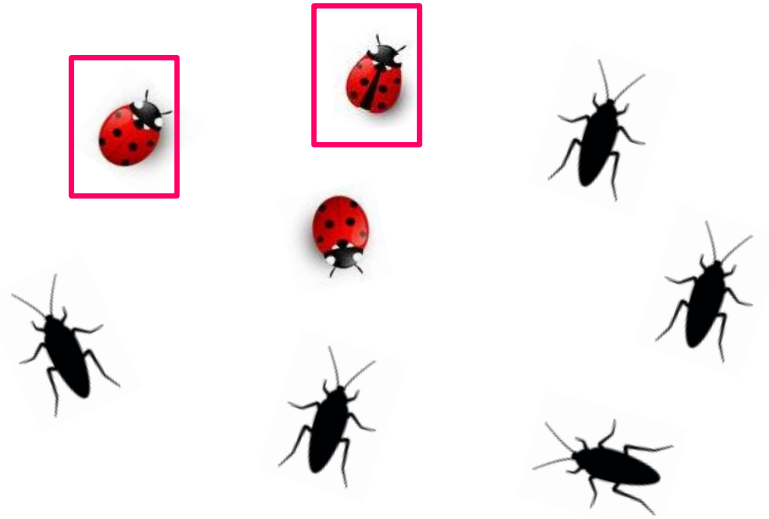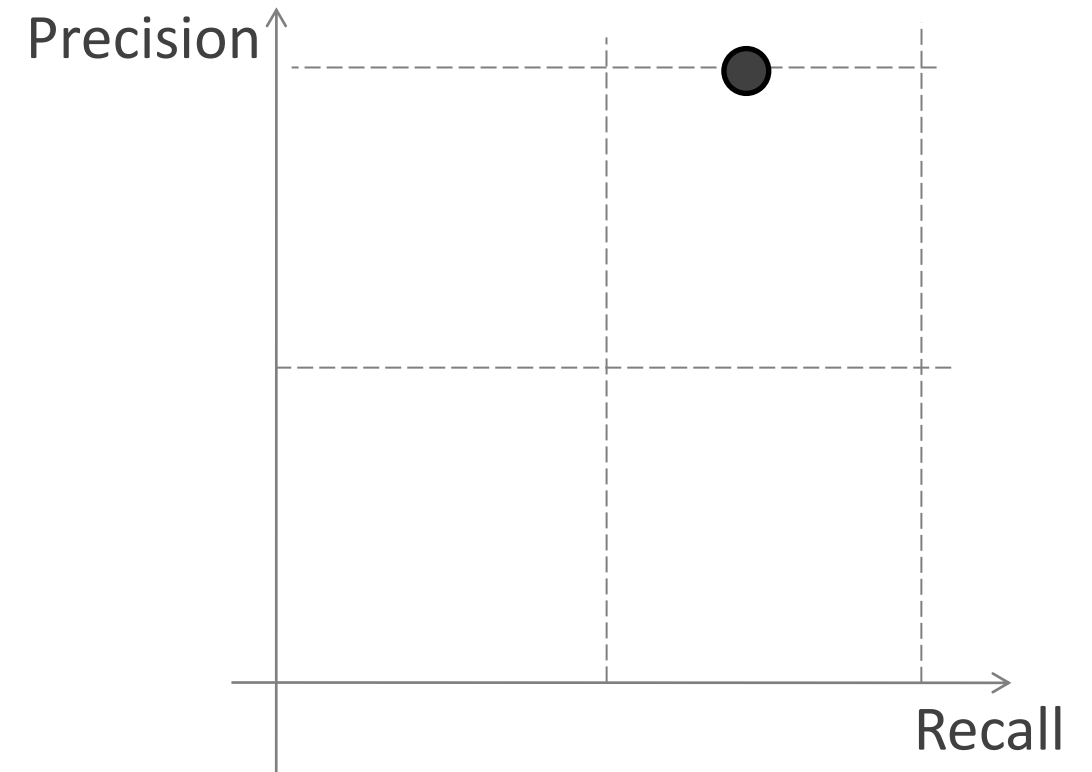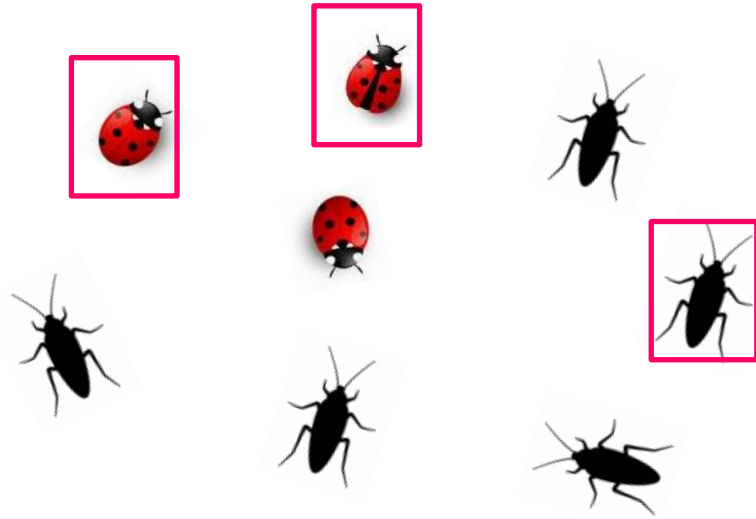
TP: 1   FP: 0   Precision: 1

FN: 2
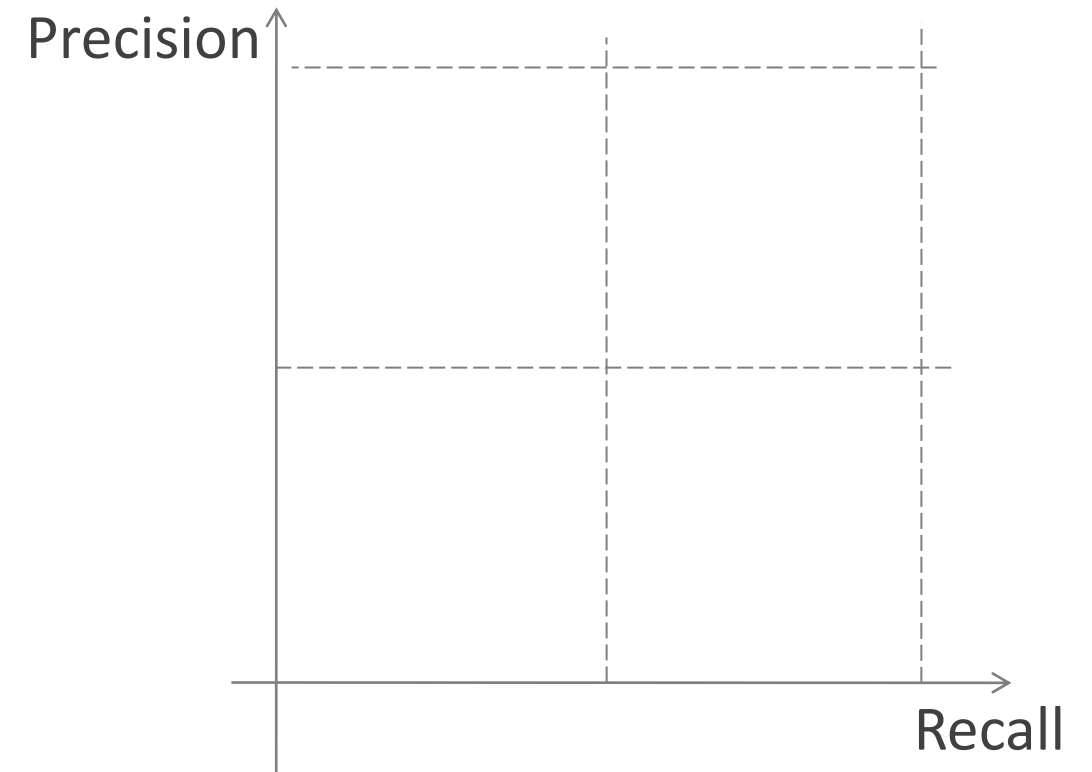
Recall: 1/3
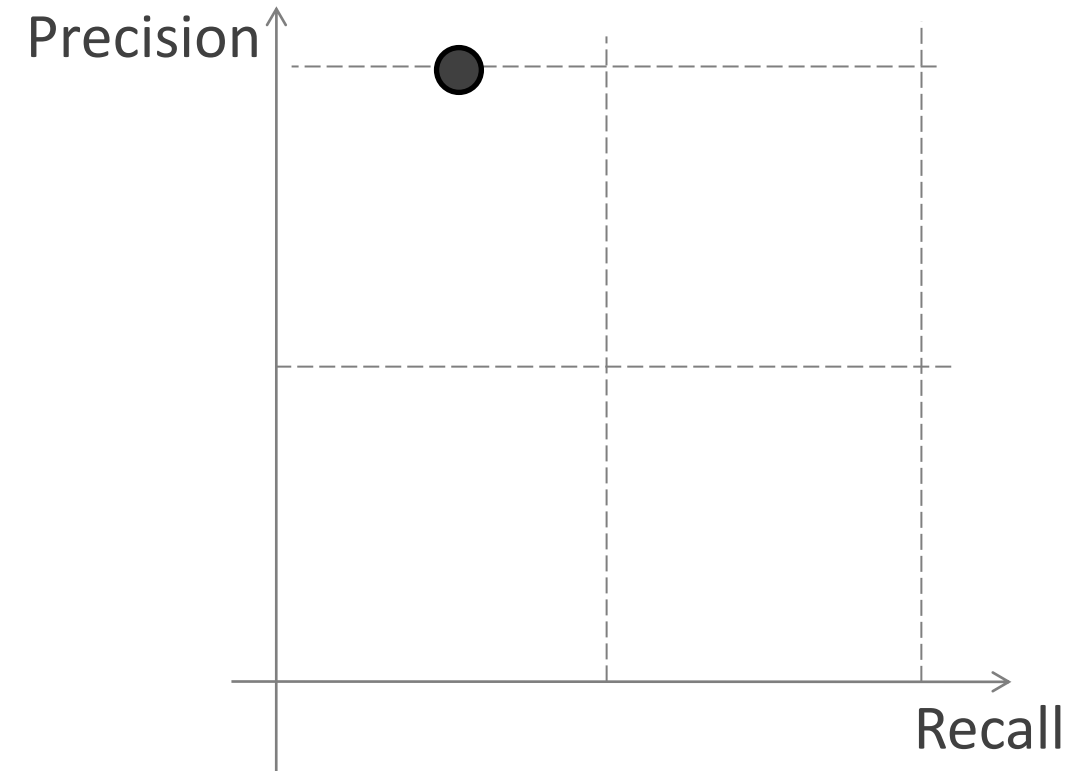
Precision

Recall

Precision - Recall

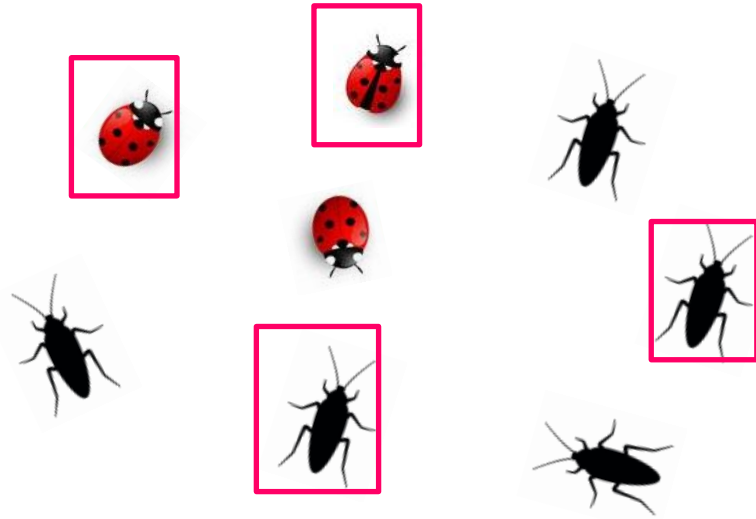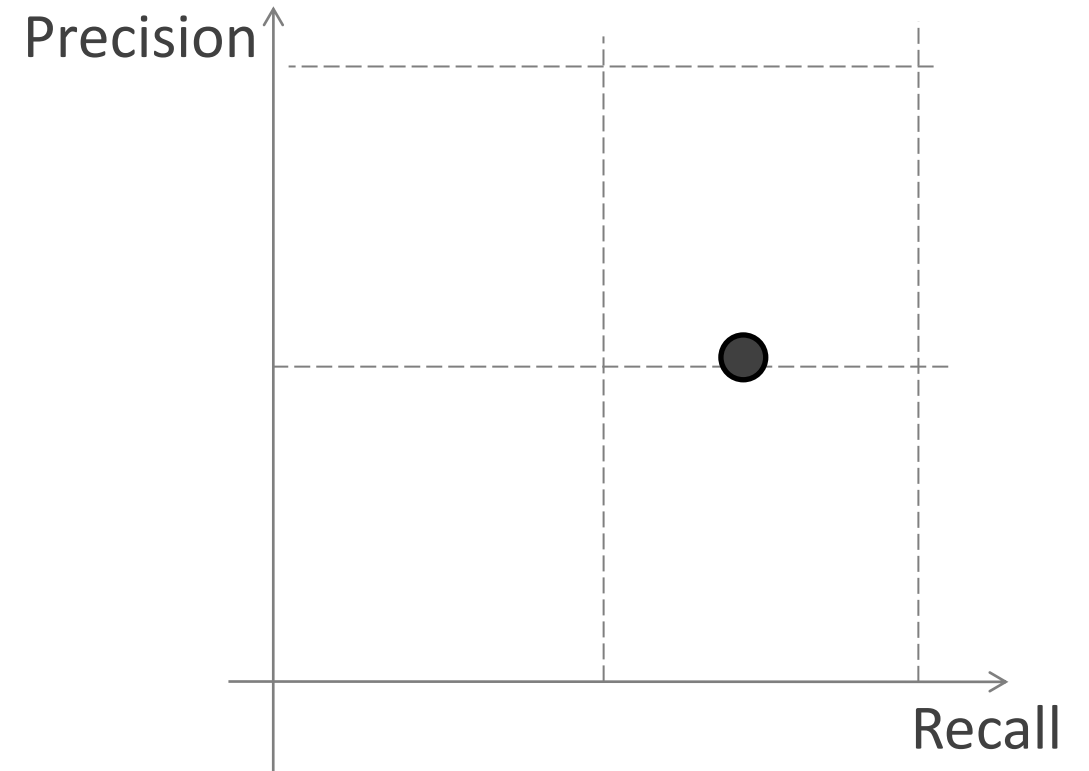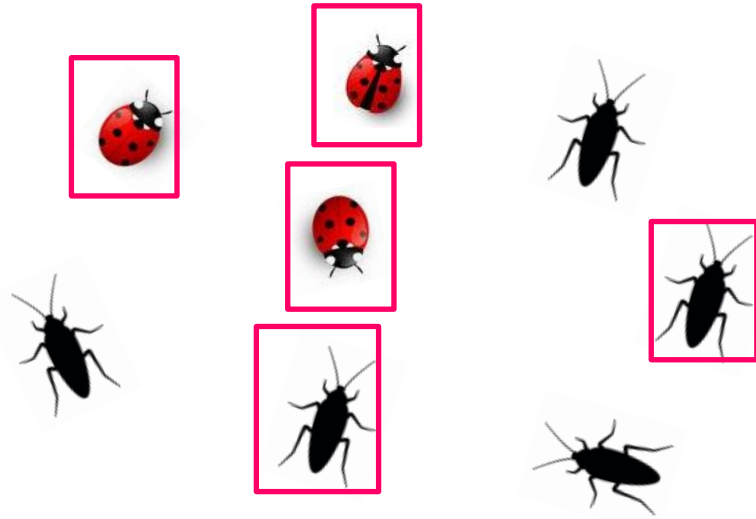TP: 2   FP: 0   Precision: 1

FN: 1

Recall: 2/3

Precision

Recall

Precision - Recall

TP: 2 **FP: 1** **Precision: 2/3**

FN: 1

Recall: 2/3
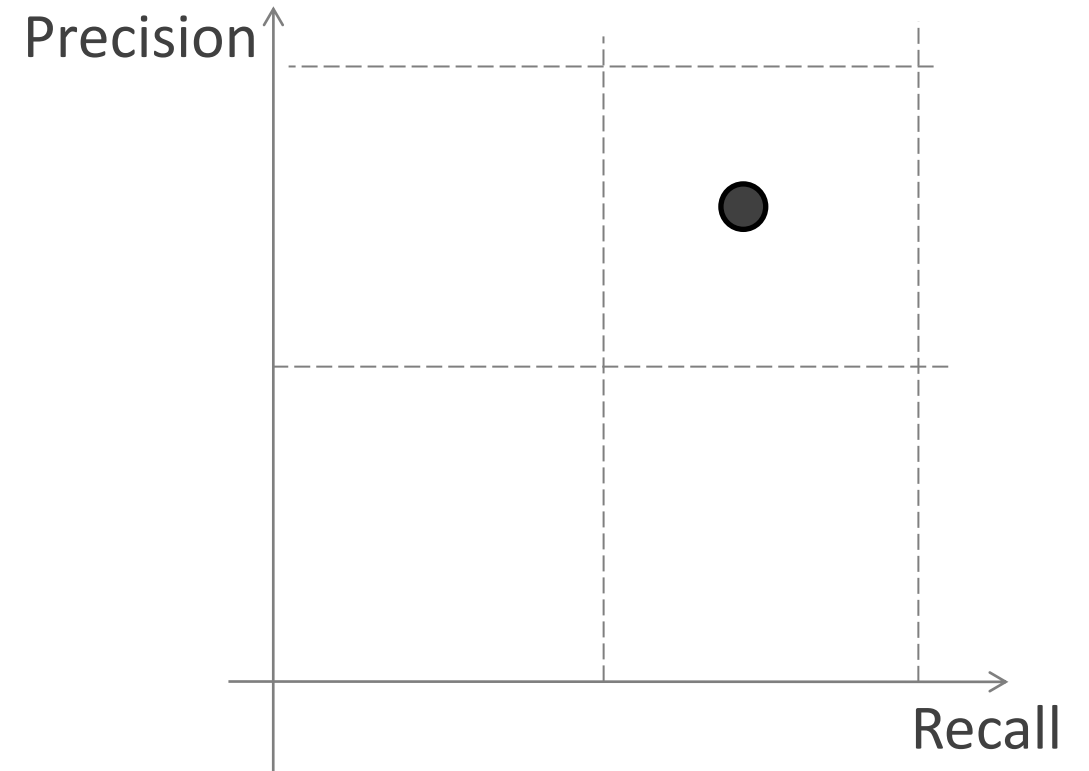
Precision

Recall

Precision - Recall

TP: 2 | FP: 2 | Precision: 2/4

FN: 1

Recall: 2/3

Precision
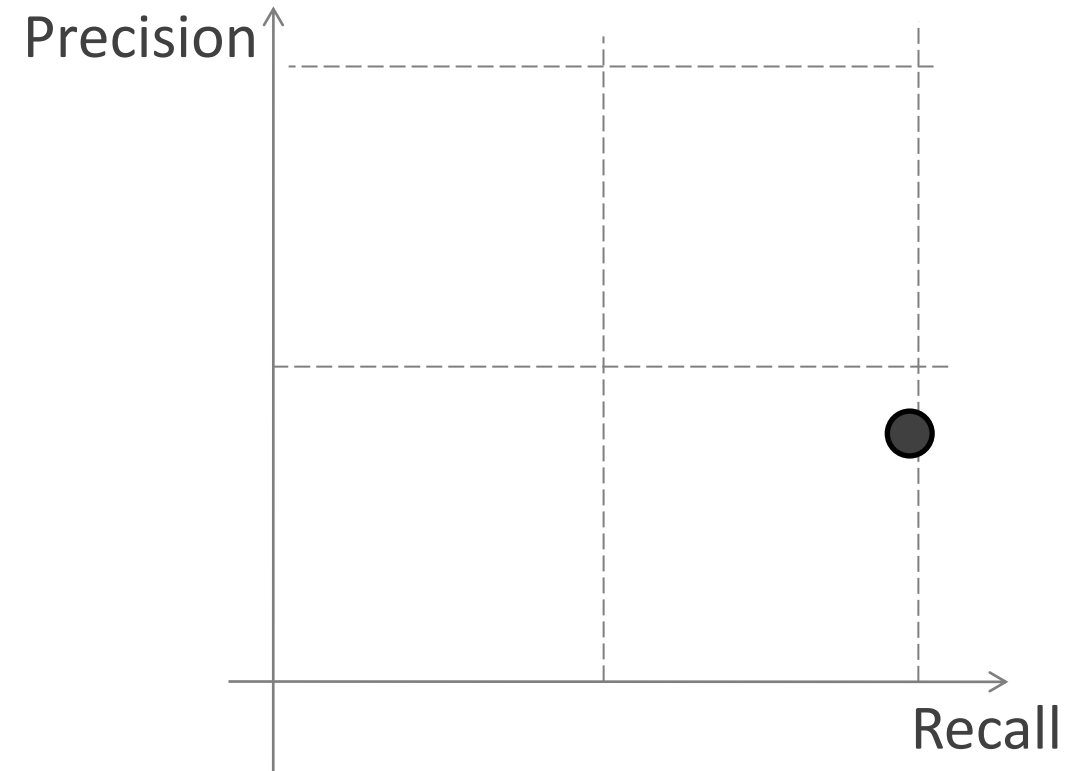
Recall
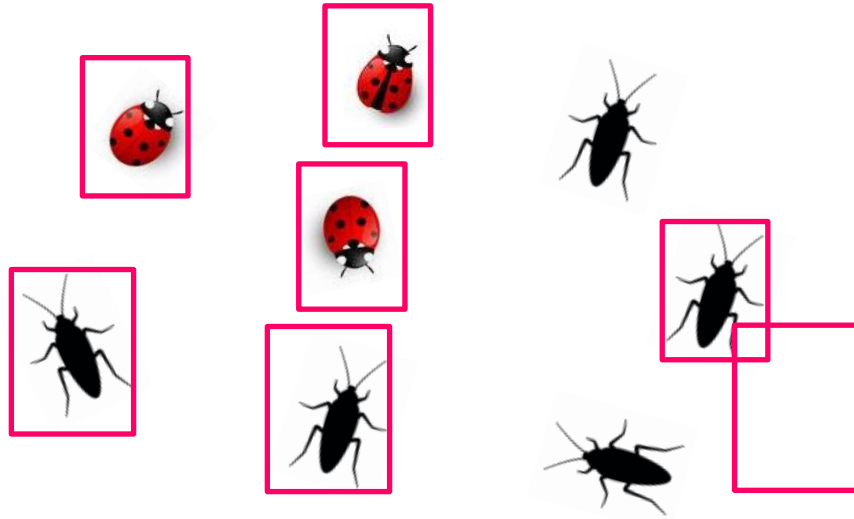
Precision - Recall
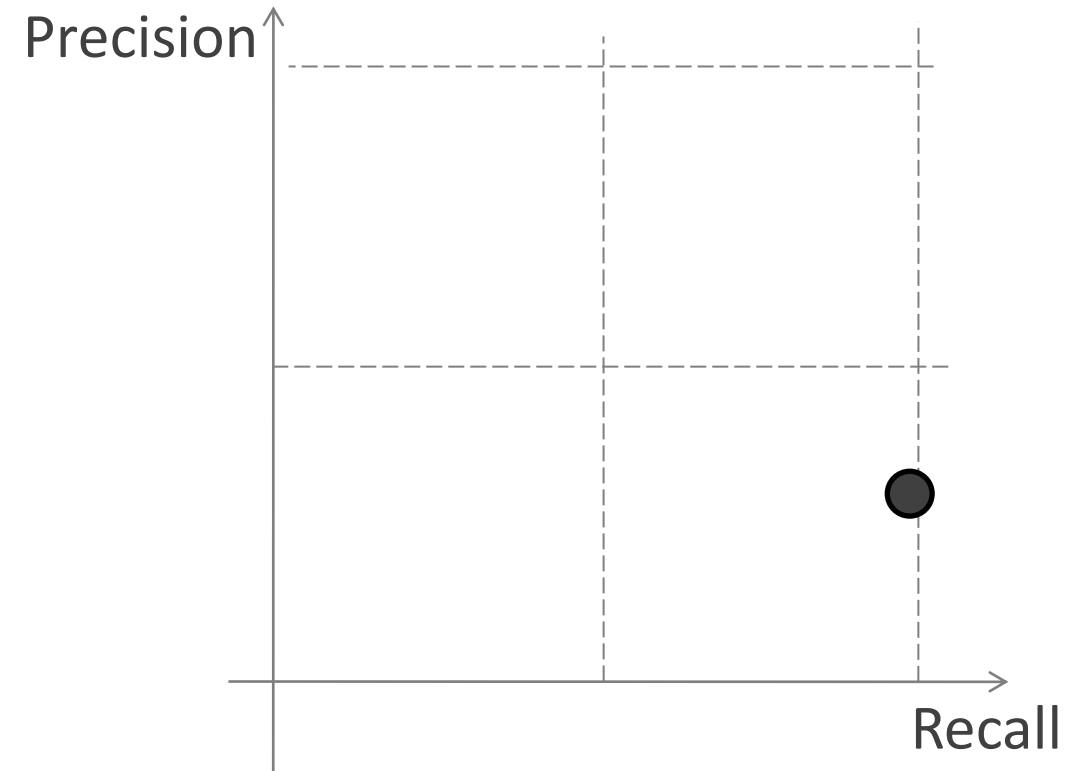

TP: 3   FP: 2   Precision: 2/5

FN: 0
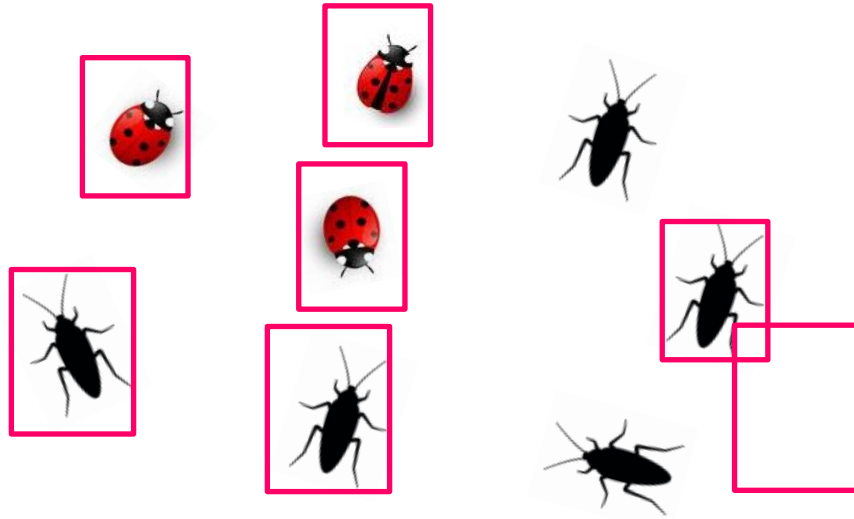
Recall: 1

Precision - Recall

TP: 3   FP: 4   Precision: 3/7

FN: 0

Recall: 1

Precision

Recall

Precision - Recall

TP: 3   **FP: 4**   **Precision: 3/7**

FN: 0

Recall: 1
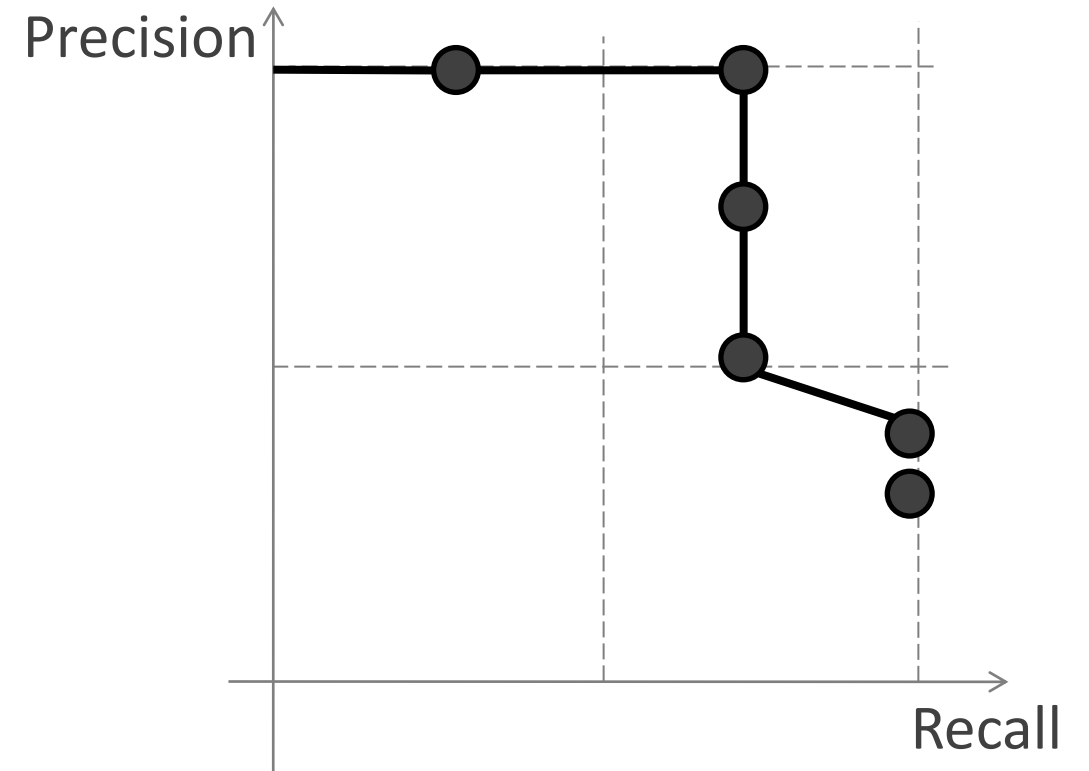
Benchmarking the classifiers

Benchmarking the Classifiers

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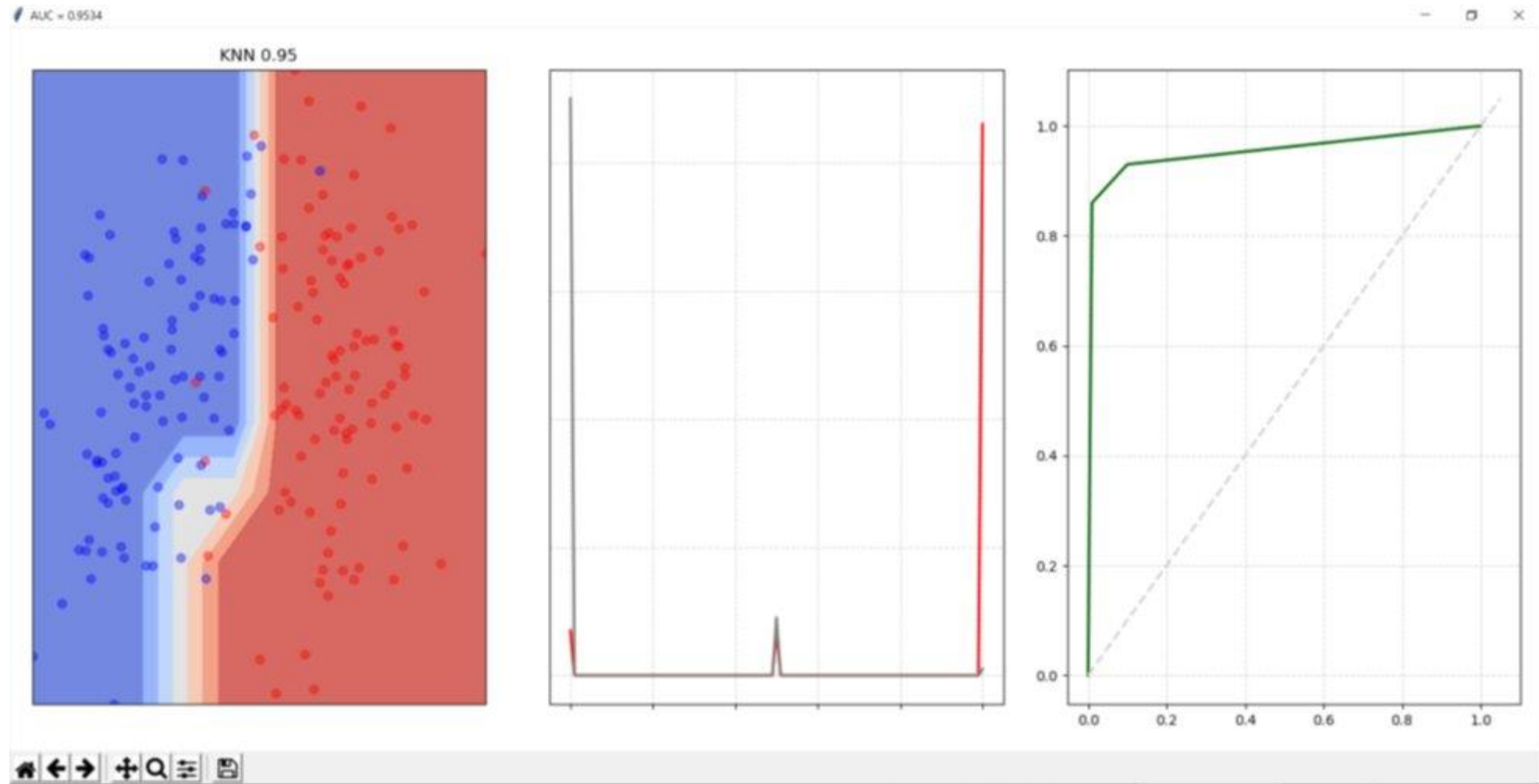- o Neural Net
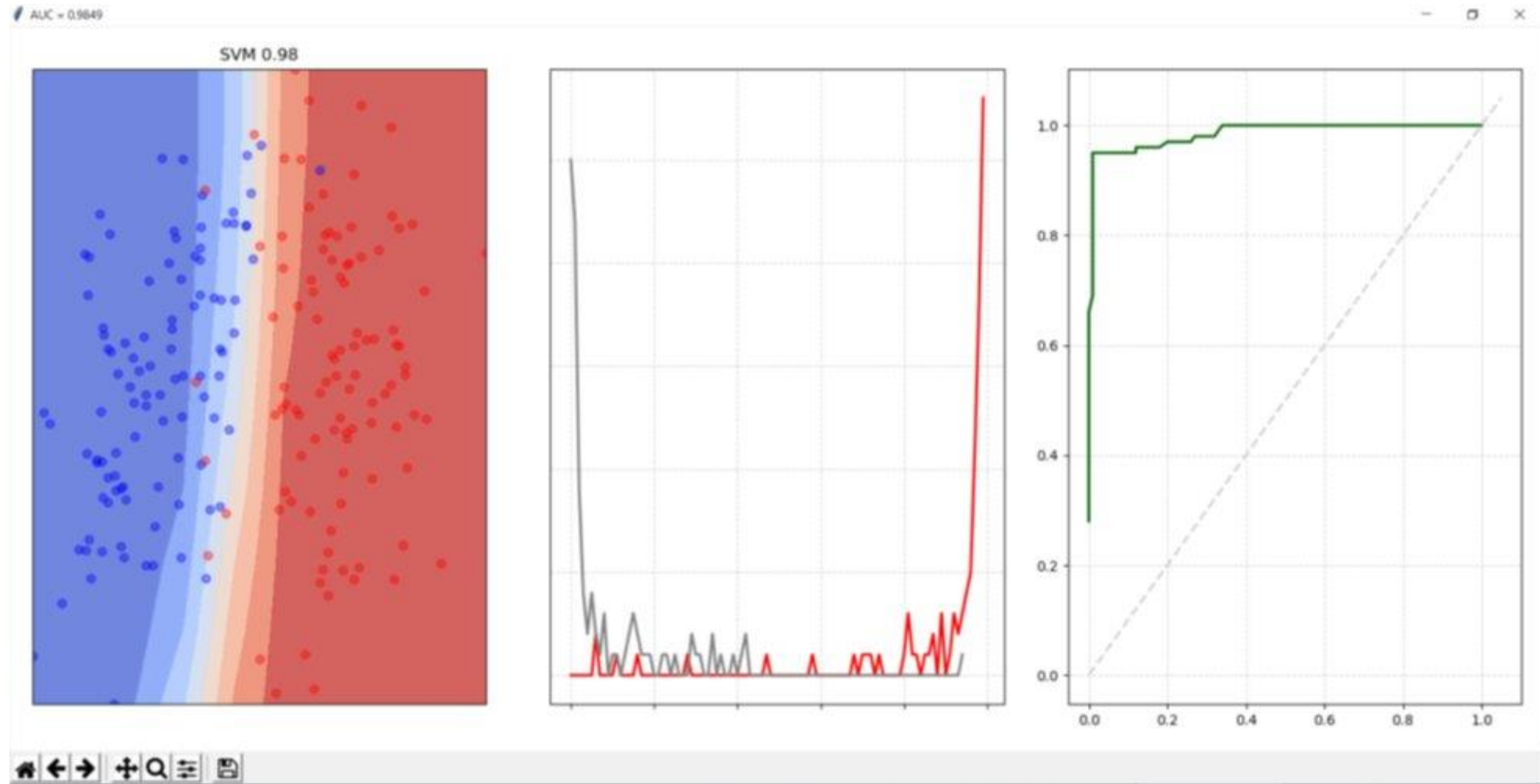
Benchmarking the Classifiers

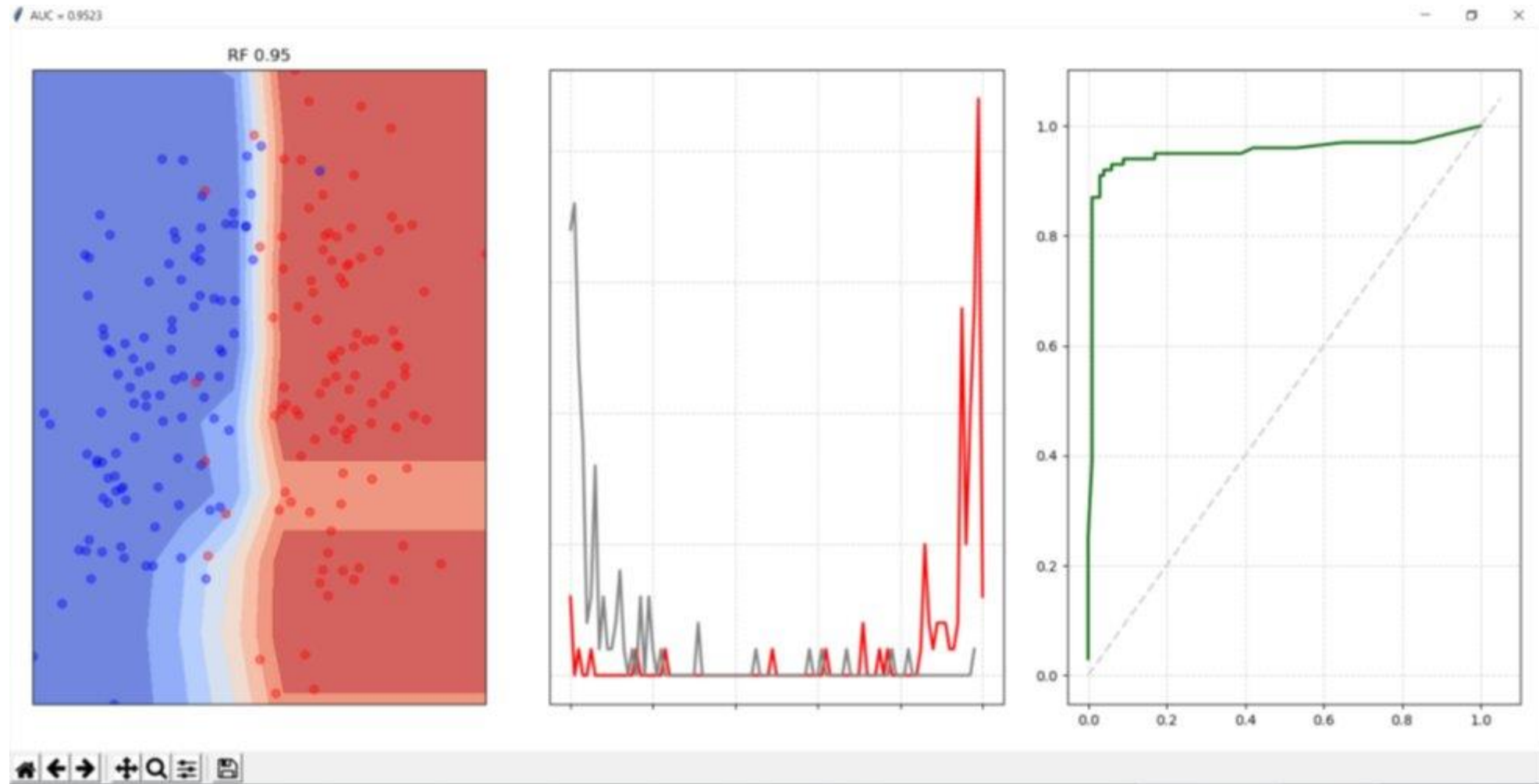Gaussian distribution
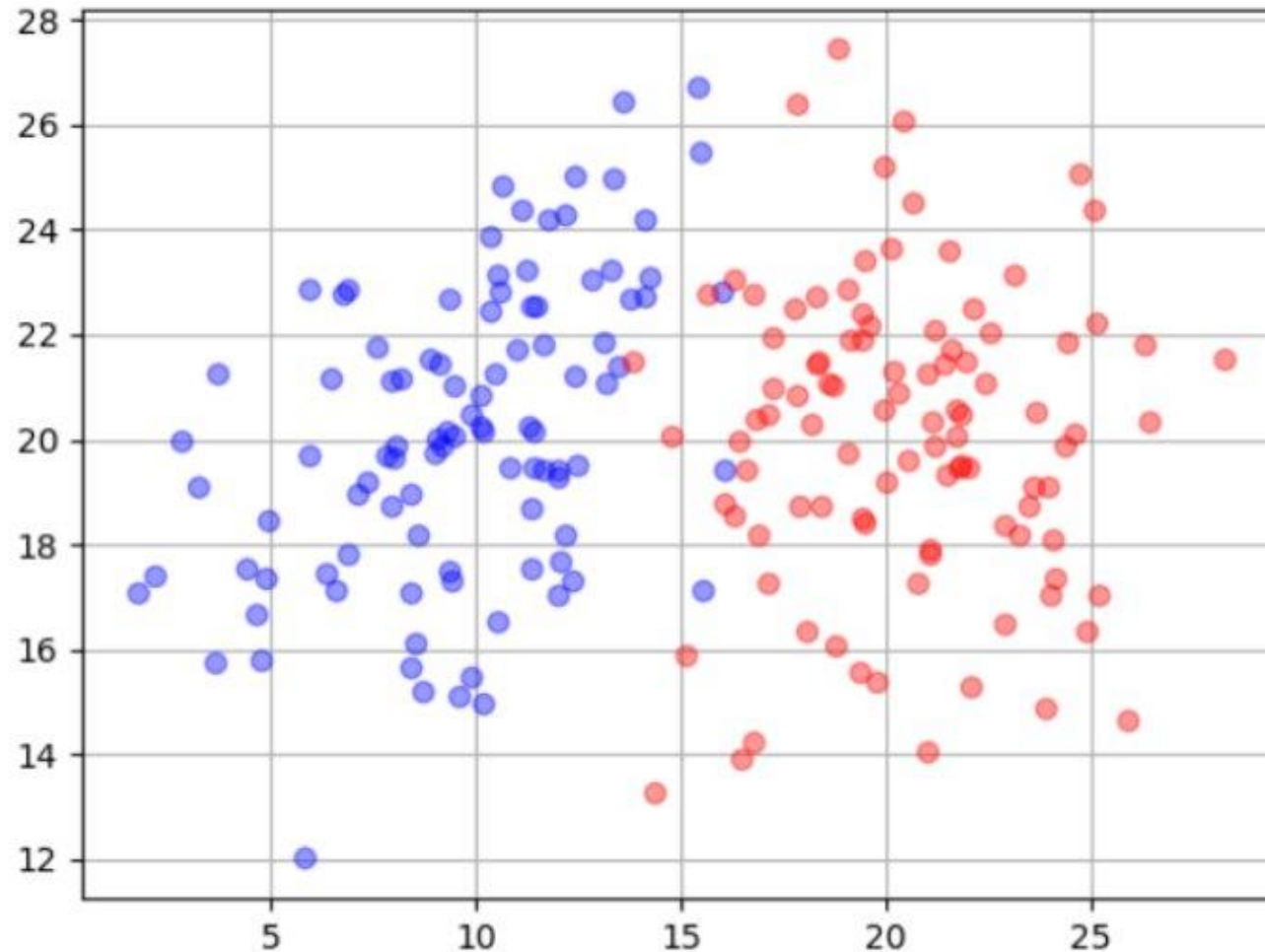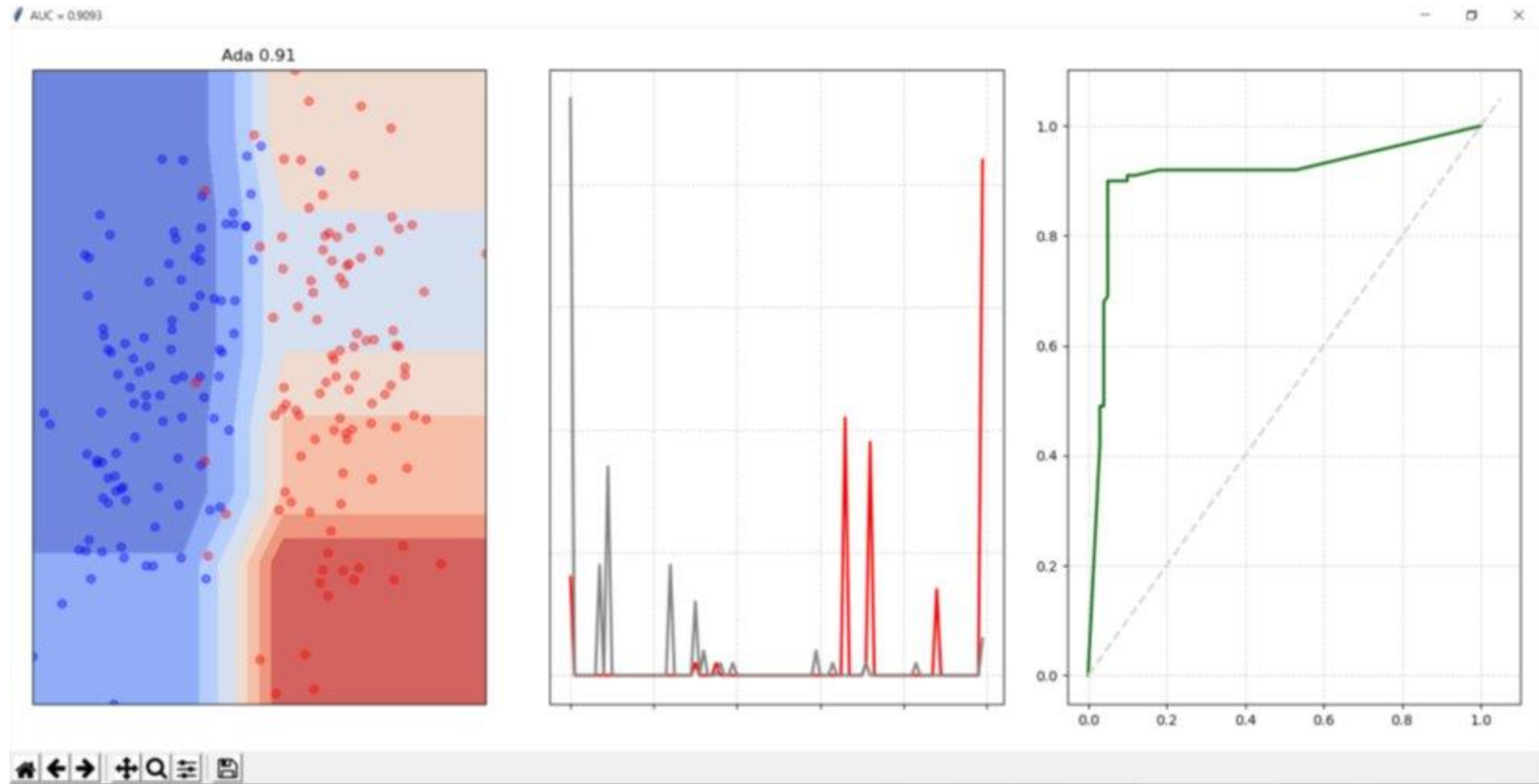
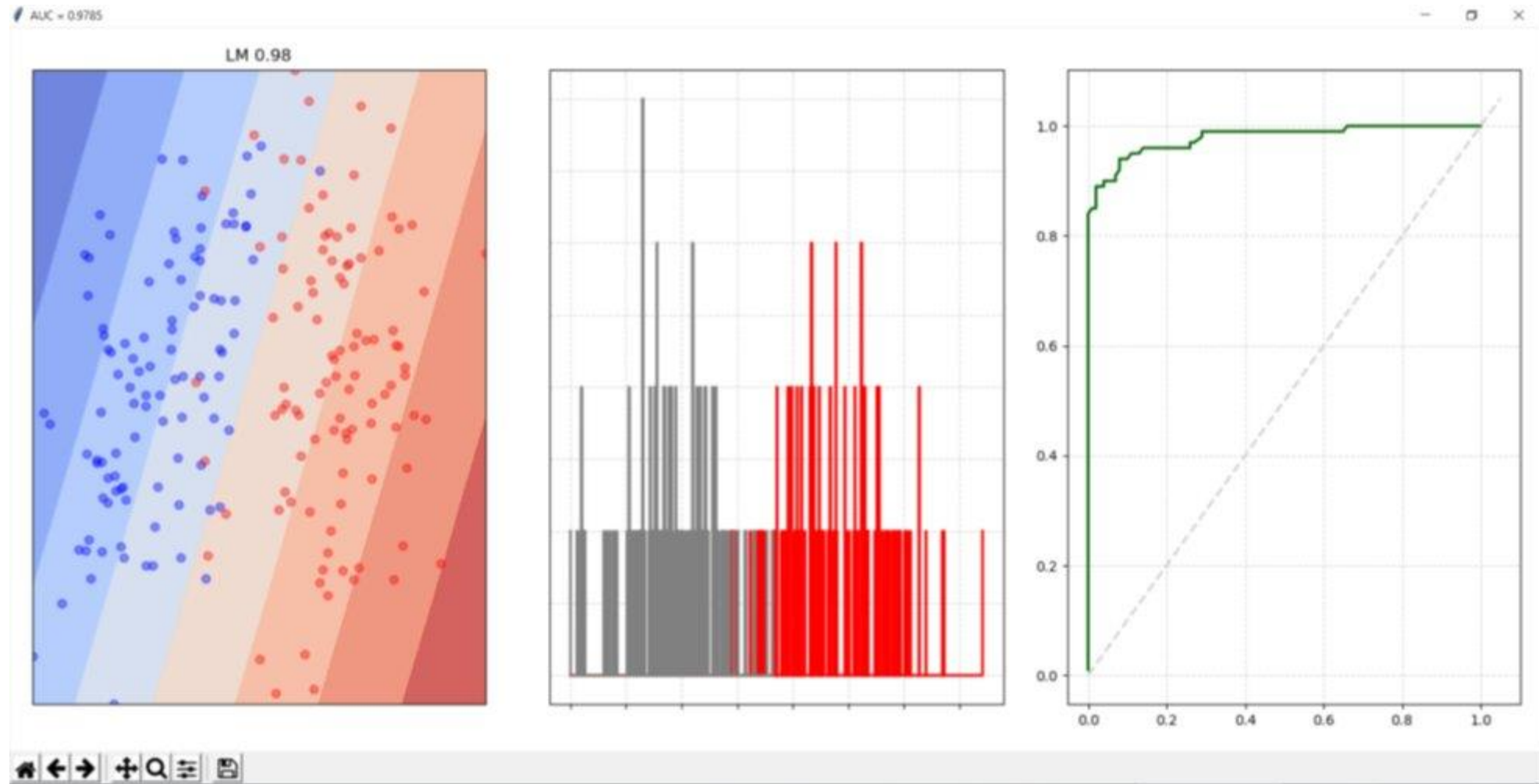# Benchmarking the Classifiers

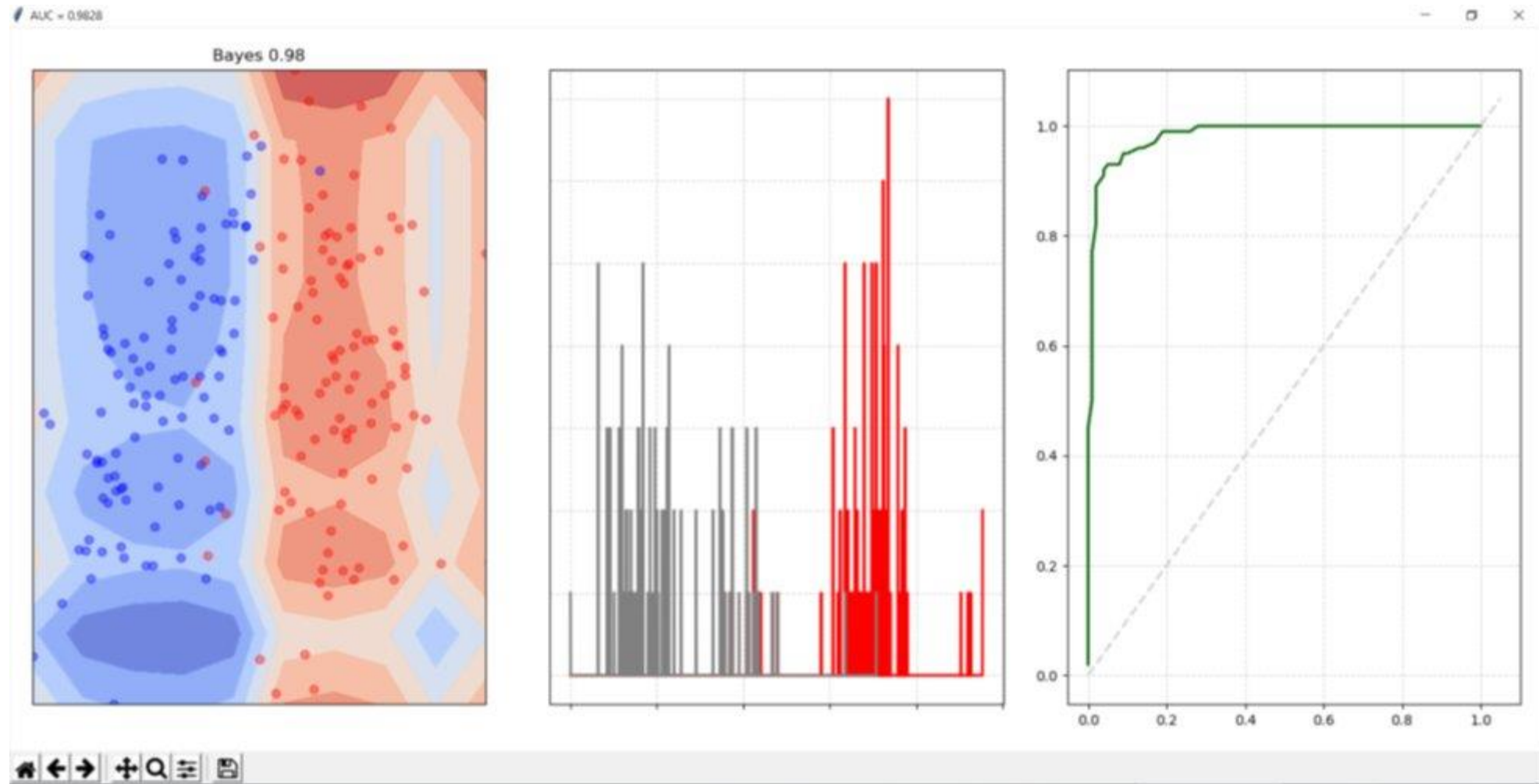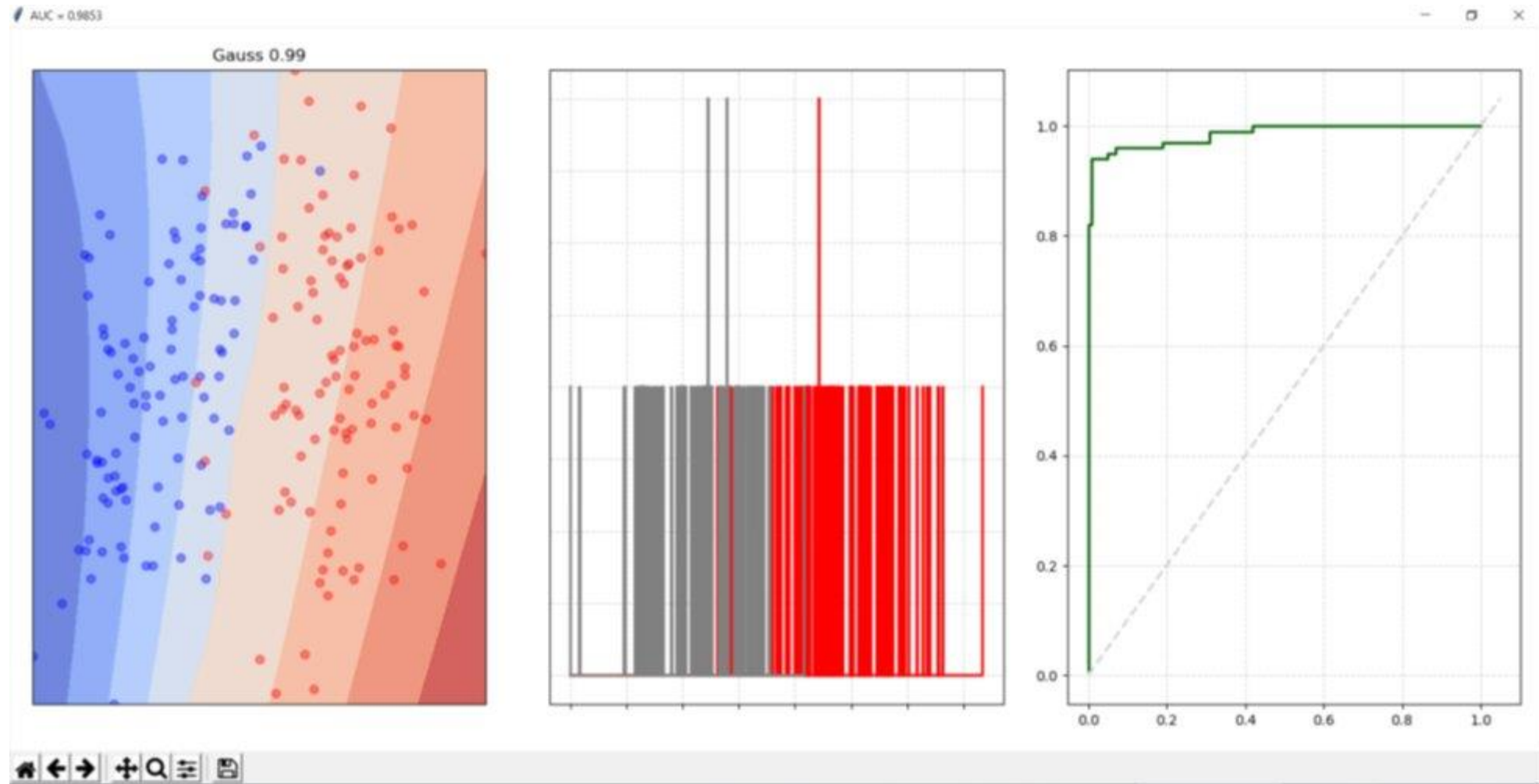# Benchmarking the Classifiers
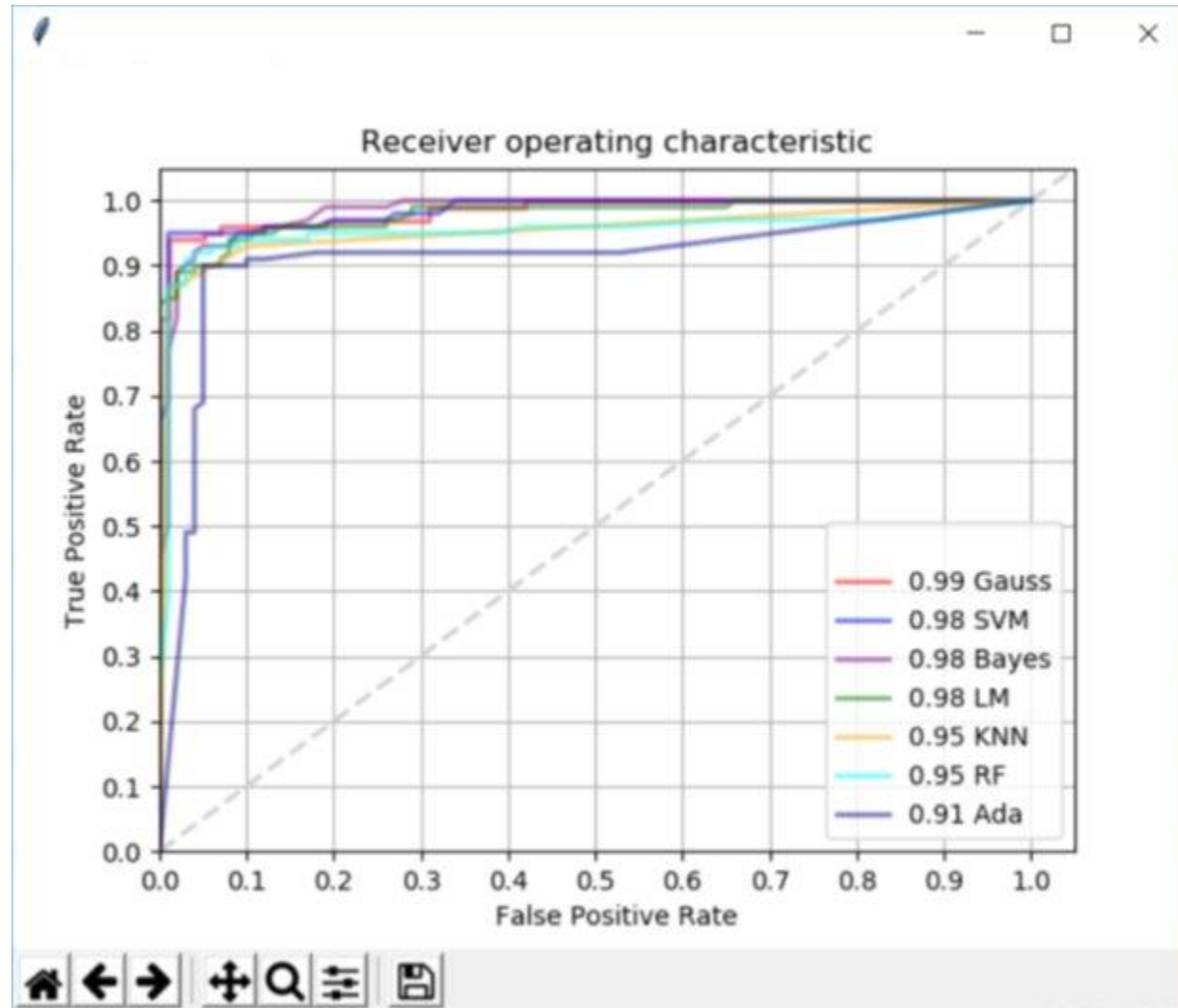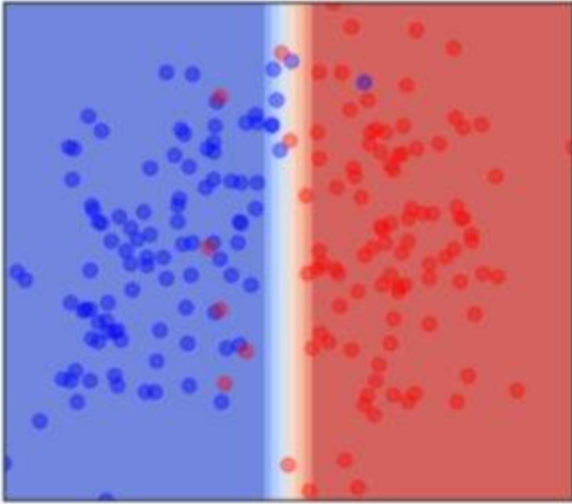
Benchmarking the Classifiers

Benchmarking the Classifiers

Benchmarking the Classifiers

Benchmarking the Classifiers

Benchmarking the Classifiers

Benchmarking the Classifiers

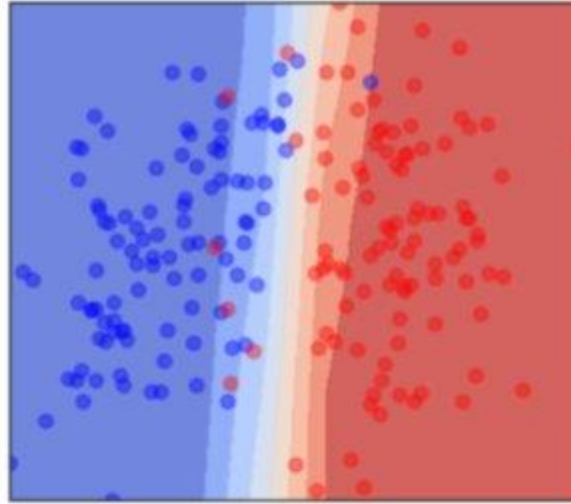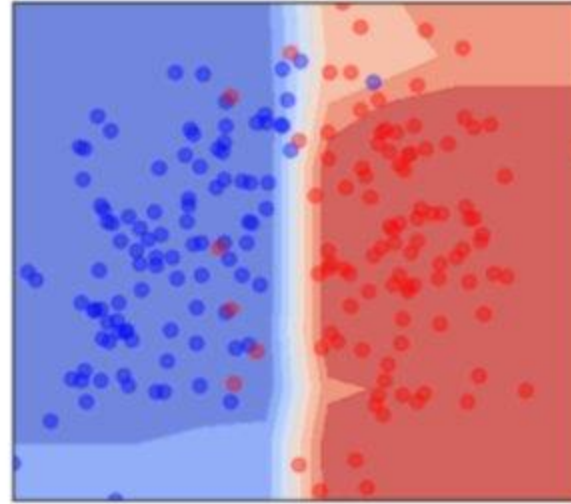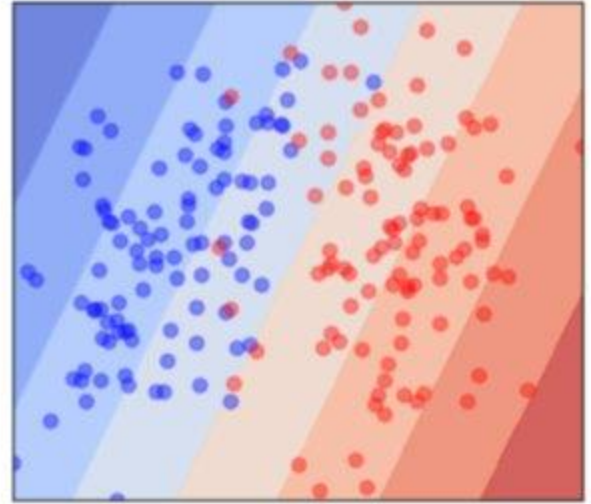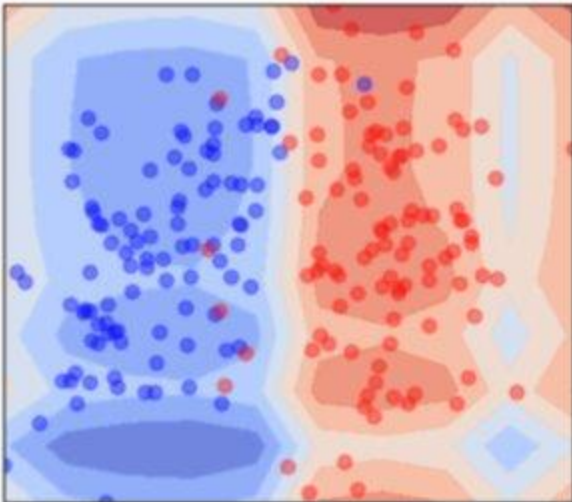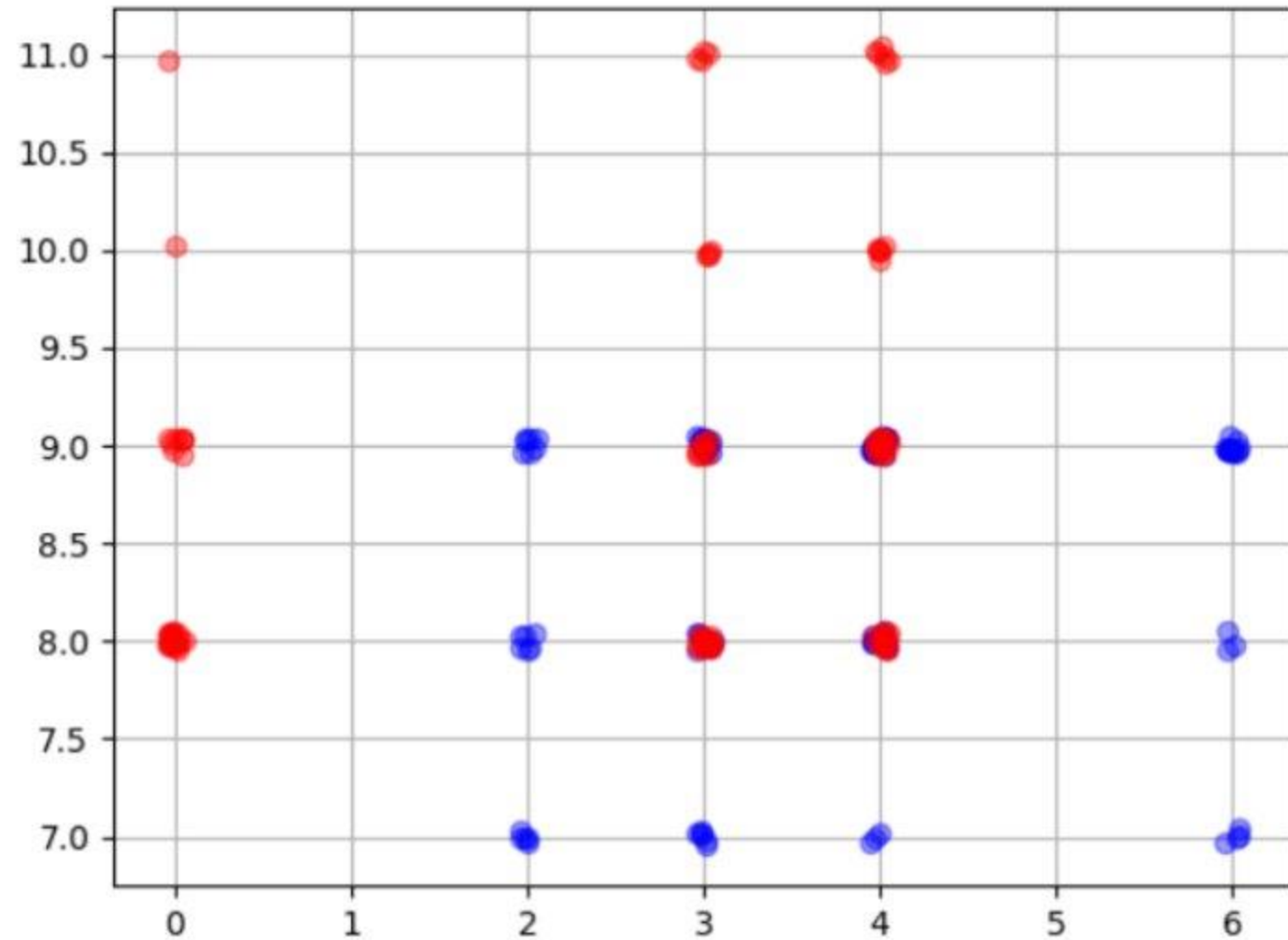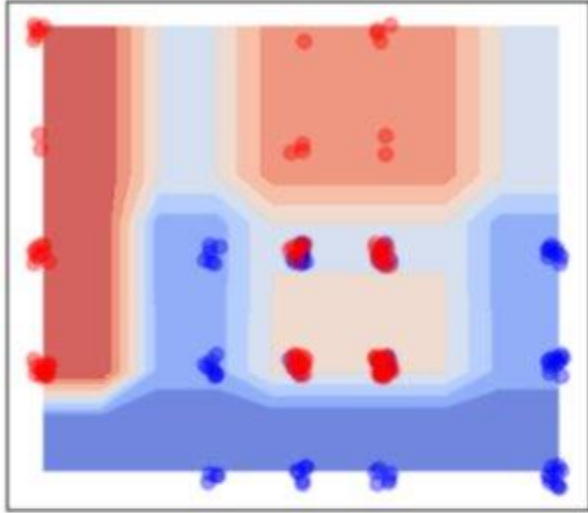# Benchmarking the Classifiers
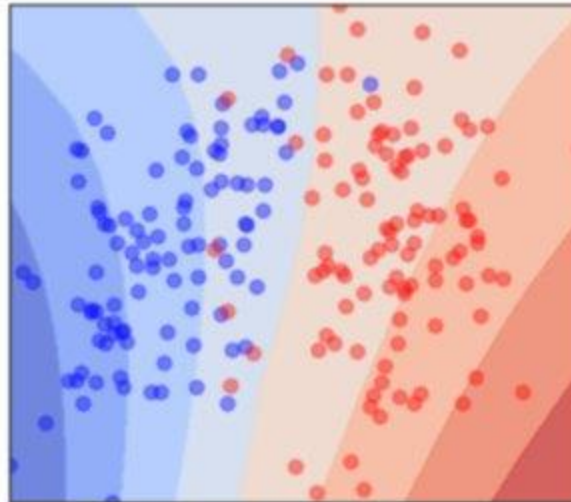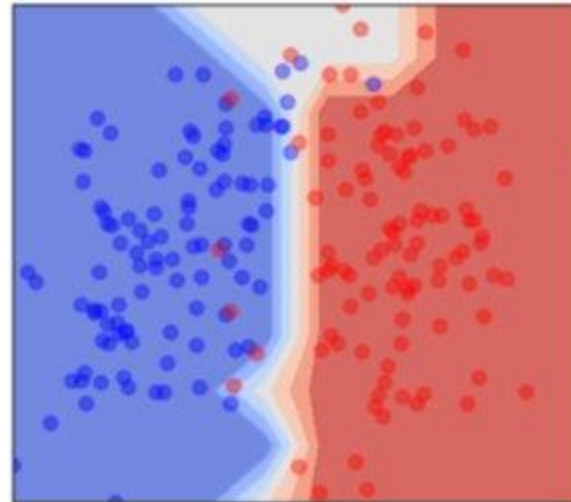
Benchmarking the Classifiers

Non-gaussian distribution

Benchmarking the Classifiers
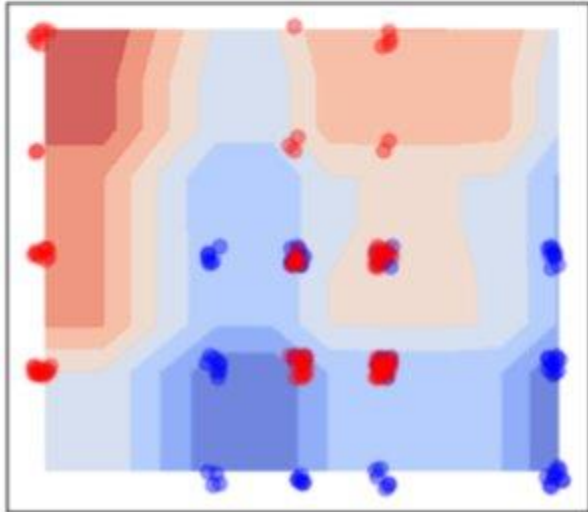
Benchmarking the Classifiers