Classifiers

Z. Straka, P. Švarný, J. Kostlivá

Today two examples:

1. Recall and Precision
2. Linear classification

Recall and Precision

Confusion matrix

		Actual class	
		P	N
Predicted class	P	TP	FP
	N	FN	TN

The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

Recall:

A: $\frac{TP}{TP+FN}$

B: $\frac{TP}{TP+TN}$

C: $\frac{TP}{FP+FN}$

D: $\frac{FP}{TP+FN}$

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

Recall:

A: $\frac{TP}{TP+FN}$

B: $\frac{TP}{TP+TN}$

C: $\frac{TP}{FP+FN}$

D: $\frac{FP}{TP+FN}$

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$

Precision:

A: $\frac{TP}{TP+FN}$

B: $\frac{FP}{FP+TN}$

C: $\frac{TP}{TP+FP}$

D: $\frac{FP}{TP+FN}$

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$

Precision:

A: $\frac{TP}{TP+FN}$

B: $\frac{FP}{FP+TN}$

C: $\frac{TP}{TP+FP}$

D: $\frac{FP}{TP+FN}$

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

For the classifier from table A count the number of members for the two classes (from the point of view of the data/reality):

A: class1: 23, class2: 32

B: class1: 38, class2: 17

C: class1: 21, class2: 34

D: class1: 23, class2: 36

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

For the classifier from table A count the number of members for the two classes (from the point of view of the data/reality):

A: class1: 23, class2: 32

B: class1: 38, class2: 17

C: class1: 21, class2: 34

D: class1: 23, class2: 36

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.

Task: Which classifier (A/B/C/D) will unnecessarily stop the car the least amount of times?

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.
Task: Which classifier (A/B/C/D) will unnecessarily stop the car the least amount of times?

A because it has the lowest $\frac{FP}{TP+FP}$

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.
Task: Which classifier (A/B/C/D) will unnecessarily stop the car the least amount of times?

A because it has the lowest $\frac{FP}{TP+FP}$

Which measure is more important for the decision, recall or precision?:

A: Recall

B: Precision

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.
Task: Which classifier (A/B/C/D) will unnecessarily stop the car the least amount of times?

A because it has the lowest $\frac{FP}{TP+FP}$

Which measure is more important for the decision, recall or precision?:

B: Precision, because $\frac{FP}{TP+FP} = 1 - \textit{Precision}$. Therefore, we are looking for the maximum Precision.

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.

Task: Which classifier (A/B/C/D) is the safest?

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.
Task: Which classifier (A/B/C/D) is the safest?

D because it has the lowest $\frac{FN}{TP+FN}$

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.
Task: Which classifier (A/B/C/D) is the safest?

D because it has the lowest $\frac{FN}{TP+FN}$

Which metric is important for this decision, recall or precision?:

A: Recall

B: Precision

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.

Task: Which classifier (A/B/C/D) is the safest?

D because it has the lowest $\frac{FN}{TP+FN}$

Which metric is important for this decision, recall or precision?:

A: Recall, because $\frac{FN}{TP+FN} = 1 - Recall$. Therefore, we are looking for the maximum Recall.

Recall and Precision

You get a few confusion matrices of 4 different classifiers. The abbreviations have the following meaning true positive (TP), false positive (FP), false negative (FN), true negative (TN).

classifier	$TP = 20$	$FP = 3$
A	$FN = 18$	$TN = 14$

classifier	$TP = 13$	$FP = 14$
C	$FN = 18$	$TN = 1$

classifier	$TP = 60$	$FP = 80$
B	$FN = 43$	$TN = 21$

classifier	$TP = 14$	$FP = 16$
D	$FN = 4$	$TN = 80$

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Let us assume that this is a table of classifiers that recognize the presence of a person in front of an autonomous car. The car stops or continues its journey.

Discussion: Suggest the simplest and safest classifier. Use chat or microphone.

Could it be used in practice?

Summary Recall and Precision

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Think about object (pedestrian) detection in images:

Recall: What is related to recall?

A: How many (what percentage of) objects/pedestrians missed?

B: How often are the (pedestrian) detections truly negative compared to truly positive detections?

C: How often are the (pedestrian) detections truly positive compared to truly negative detections?

D: How much are the (pedestrian) detections contaminated by false detections?

Summary Recall and Precision

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Think about object (pedestrian) detection in images:
Recall: What is related to recall?

- How many (what percentage of) objects/pedestrians missed?

Precision: What is related to precision?

A: How many (what percentage of) objects/pedestrians are not missed?

B: How often are the (pedestrian) detections truly negative compared to truly positive detections?

C: How often are the (pedestrian) detections truly positive compared to truly negative detections?

D: How much are the (pedestrian) detections contaminated by false detections?

Summary Recall and Precision

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Think about object (pedestrian) detection in images:
Recall: What is related to recall?

- How many (what percentage of) objects/pedestrians missed?

Precision: What is related to precision?

A: How many (what percentage of) objects/pedestrians are not missed?

B: How often are the (pedestrian) detections truly negative compared to truly positive detections?

C: How often are the (pedestrian) detections truly positive compared to truly negative detections?

D: How much are the (pedestrian) detections contaminated by false detections?

Summary Recall and Precision

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Think about object (pedestrian) detection in images:
Recall: What is related to recall?

- How many (what percentage of) objects/pedestrians missed?

Precision: What is related to precision?

D: How much are the (pedestrian) detections contaminated by false detections?

Summary Recall and Precision

- Recall: $\frac{TP}{TP+FN}$
- Precision: $\frac{TP}{TP+FP}$

Think about object (pedestrian) detection in images:

Recall: What is related to recall?

- How many (what percentage of) objects/pedestrians missed?

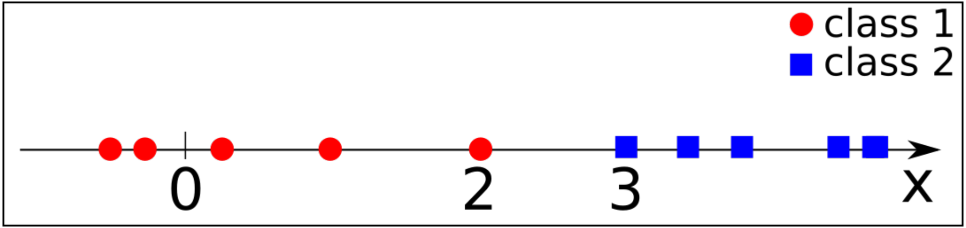
Precision: What is related to precision?

- How much are the (pedestrian) detections contaminated by false detections?

It is difficult to satisfy both.

Linear classification

1D classification



Classification is made according to:

$$s^* = \arg\max_s f_s(x)$$

where $f_s(x) = w_s x + b_s$, $w_s, b_s \in R$, $s \in \{1, 2\}$
Select a classifier with zero classification error on the given dataset:

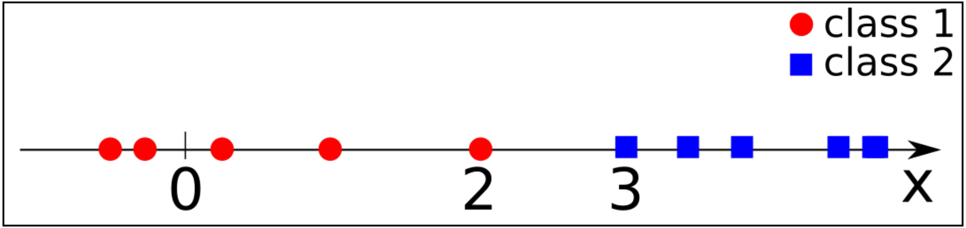
A: $w_1 = 1, b_1 = 0; w_2 = -1, b_2 = 0$

B: $w_1 = 1, b_1 = -2.4; w_2 = -1, b_2 = 2.5$

C: $w_1 = -1, b_1 = 2.5; w_2 = 1, b_2 = -2.4$

D: $w_1 = -1, b_1 = 0; w_2 = 1, b_2 = 0$

1D classification



Classification is made according to:

$$s^* = \arg\max_s f_s(x)$$

where $f_s(x) = w_s x + b_s, w_s, b_s \in R, s \in \{1, 2\}$
Select a classifier with zero classification error on the given dataset:

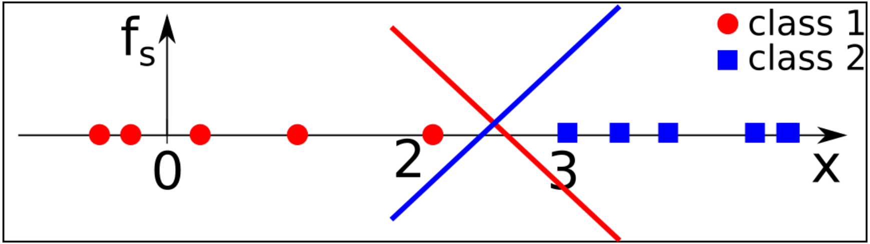
A: $w_1 = 1, b_1 = 0; w_2 = -1, b_2 = 0$

B: $w_1 = 1, b_1 = -2.4; w_2 = -1, b_2 = 2.5$

C: $w_1 = -1, b_1 = 2.5; w_2 = 1, b_2 = -2.4$

D: $w_1 = -1, b_1 = 0; w_2 = 1, b_2 = 0$

1D classification



Classification is made according to:

$$s^* = \arg\max_s f_s(x)$$

where $f_s(x) = w_s x + b_s, w_s, b_s \in R, s \in \{1, 2\}$
Select a classifier with zero classification error on the given dataset:

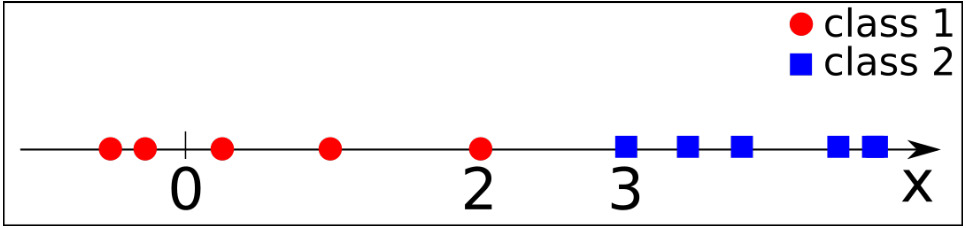
A: $w_1 = 1, b_1 = 0; w_2 = -1, b_2 = 0$

B: $w_1 = 1, b_1 = -2.4; w_2 = -1, b_2 = 2.5$

C: $w_1 = -1, b_1 = 2.5; w_2 = 1, b_2 = -2.4$

D: $w_1 = -1, b_1 = 0; w_2 = 1, b_2 = 0$

1D classification



Alternatively (only for binary classification):

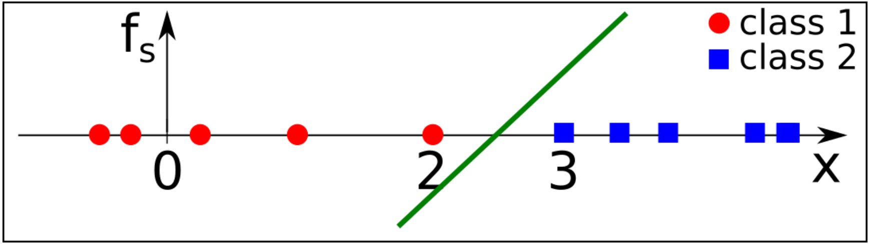
A: $s^* = (\mathtt{sgn}(x - 2.5) + 1)/2 + 1$

B: $s^* = \mathtt{sgn}(x - 2.5)$

C: $s^* = \mathtt{max}(x_i - 2.5)$

D: $s^* = \mathtt{sgn}(x - 2.5) + 1$

1D classification



Alternatively (only for binary classification):

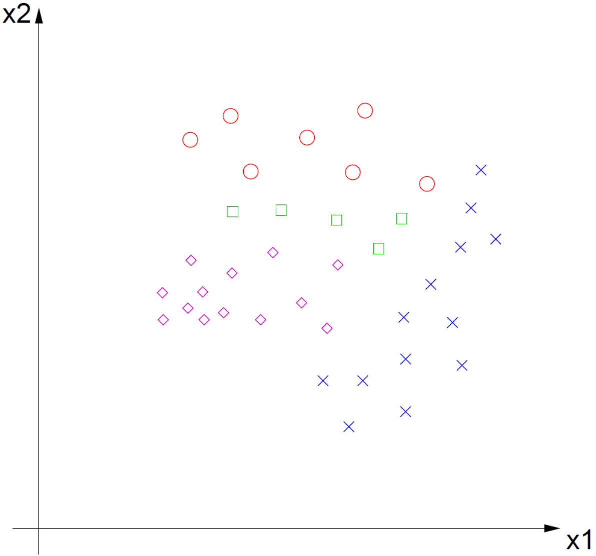
A: $s^* = (\mathrm{sgn}(x - 2.5) + 1)/2 + 1$

B: $s^* = \mathrm{sgn}(x - 2.5)$

C: $s^* = \max(x_i - 2.5)$

D: $s^* = \mathrm{sgn}(x - 2.5) + 1$

2D classification

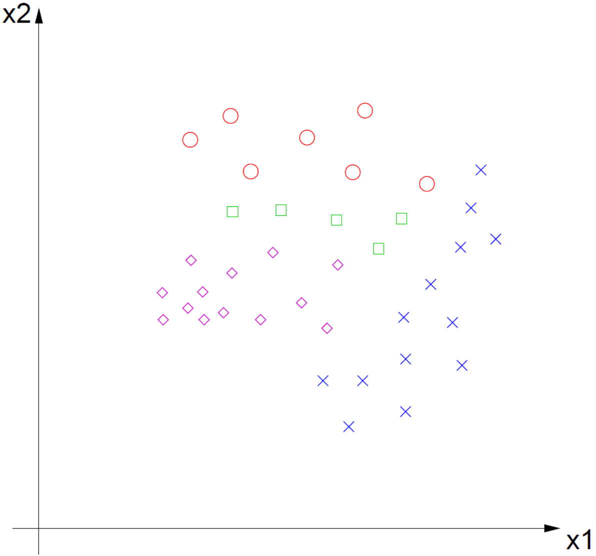


Is it possible to get zero classification errors on the given training multiset using only a linear classifier?

Yes

No

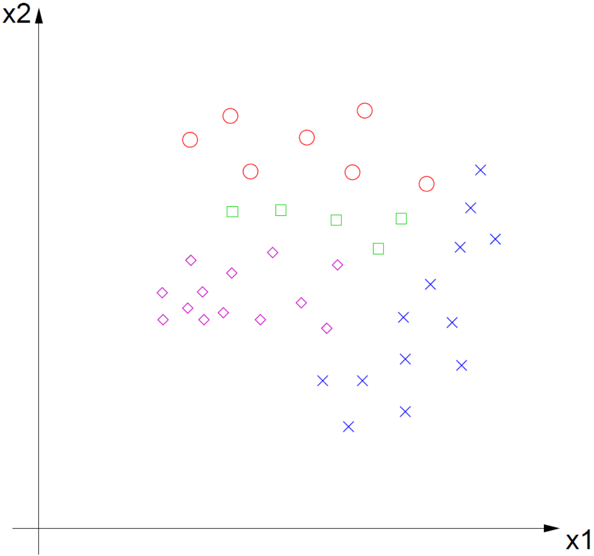
2D classification



Is it possible to get zero classification errors on the given training multiset using only a linear classifier?

- **Yes.**

2D classification



- Draw the solution.