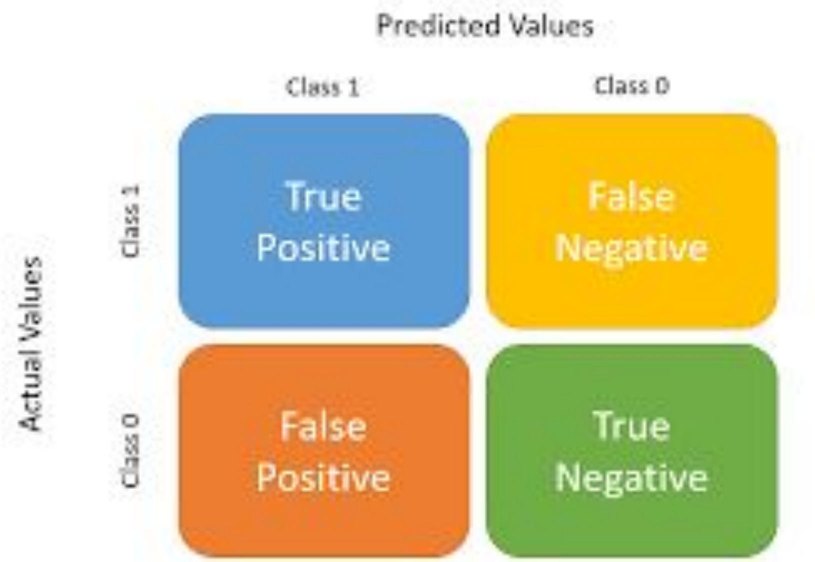
Performance indices

Classification

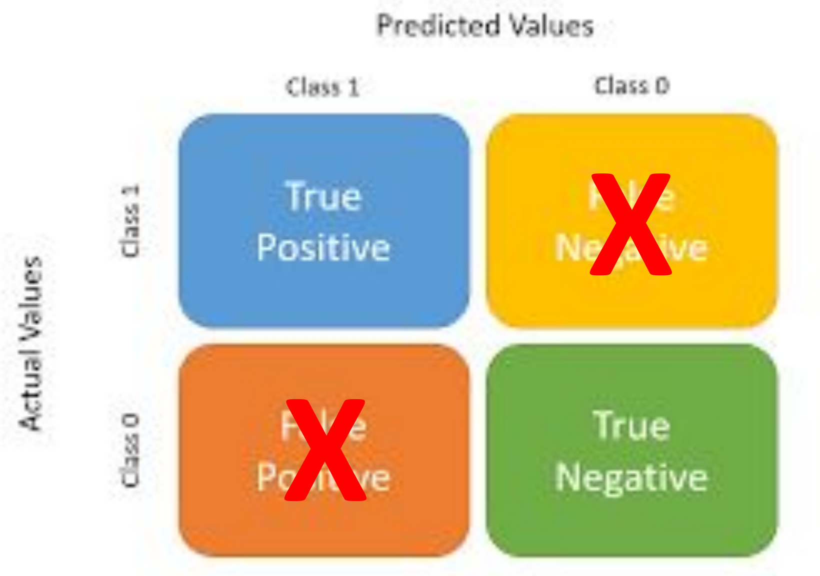
Meunier Jean-Christophe

Last updated: March 12, 2022

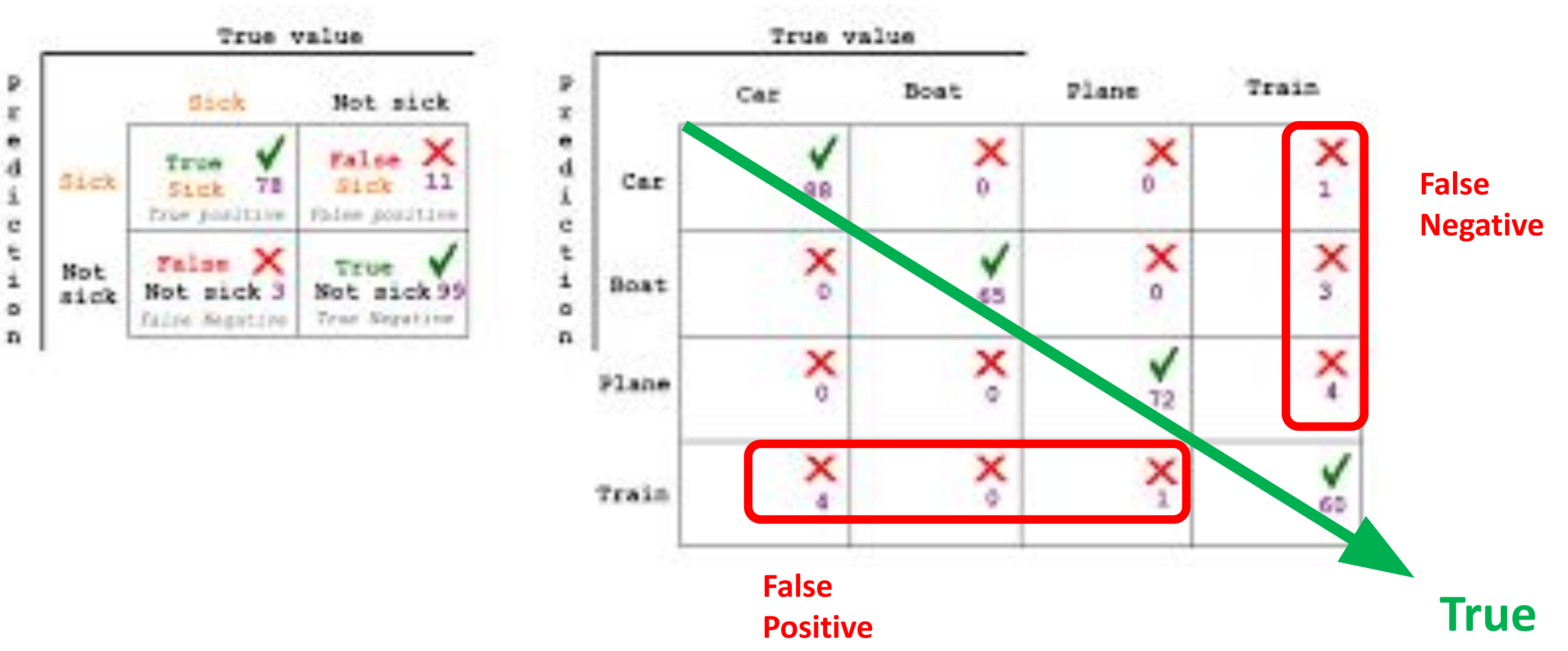
Confusion matrix



Confusion matrix



Objective of any classification task:
reaching highest rates of true values


True value
Prediction
Sick
Not sick
True Sick 78
False Sick 11
False Not sick 3
True Not sick 99
True value
Prediction
Car
Boat
Plane
Train
98
0
0
1
0
65
0
3
0
0
72
4
4
0
1
60
False Negative
False Positive
True

- Confusion !!!

		Predicted value	
		A	B
Actual value	A	TP	FN
	B	FP	TN

		Actual values	
		A	B
Predicted value	A	TP	FP
	B	FN	TN

By default presentation in scikit-learn (e.g. for confusion matrix, classification report)

Accuracy

- Percentage of good prediction

		Predicted value	
		Immune	No Imm.
Actual value	Immune	45	15
	No Imm.	13	27

Covid

In a sample of 100 observations

- 60, naturally immune
- 40, no immunization

Good predictions

$$Accuracy = \frac{TP + TN}{Total} = \frac{45 + 27}{100} = 0{,}72$$

Accuracy !!! not good indicator if strongly unbalanced classes

- No skill model predict that everyone is immune

		Predicted value	
		Immune	No Imm.
Actual value	Immune	95	0
	No Imm.	5	0

Covid

In a sample of 100 observations

- 95, naturally immune
- 5, no immunization

Good predictions

$$Accuracy = \frac{TP + TN}{Total} = \frac{95 + 0}{100} = 0{,}95$$

Accuracy **!!! not good indicator if strongly unbalanced classes**

- No skill model predict that no one is immune

		Predicted value	
		Immune	No Imm.
Actual value	Immune	0	95
	No Imm.	0	5

Covid

In a sample of 100 observations

- 95, naturally immune
- 5, no immunization

Good predictions

$$Accuracy = \frac{TP + TN}{Total} = \frac{0 + 5}{100} = 0{,}05$$

$$Precision = \frac{TP}{TP + FP}$$ → Penalize optimistic models (i.e. those with high FP rate)

Precision

- Ratio true positives (TP) on all predicted positives (TP + FP)

		Predicted value	
		Immune	No Imm.
Actual value	Immune	45	15
	No Imm.	13	27

Covid

In a sample of 100 observations

- 60, naturally immune
- 40, no immunization

Good predictions

$$Precision = \frac{TP}{TP + FP} = \frac{45}{45 + 13} = 0,78$$

Recall

$$Recall = \frac{TP}{TP + FN}$$

→ Penalize pessimistic models (i.e. those with high FN rate)

- Ratio true positives (TP) on all real/actual positives (TP + FN)

		Predicted value	
		Immune	No Imm.
Actual value	Immune	45	15
	No Imm.	13	27

Covid

In a sample of 100 observations

- 60, naturally immune
- 40, no immunization

Good predictions

$$Recall = \frac{TP}{TP + FP} = \frac{45}{45 + 15} = 0{,}75$$

Accuracy, precision, recall

Pessimistic model			
		Predicted values	
		Imm.	No i.
Actual values	Imm	35	25
	No i.	2	38

Accuracy 0.73

Precision 0.95

Most predicted positives are TP (few FP)

Recall 0.58

Many actual positives not correctly predicted (many FN)

Perfect model			
		Predicted values	
		Imm	No i.
Actual values	Imm	60	0
	No i.	0	40

Accuracy 1.00

Precision 1.00

Recall 1.00

Optimistic model			
		Predicted values	
		Imm	No i.
Actual values	Imm	58	2
	No i.	25	15

Actual positives

Predicted positives

Accuracy 0.73

Precision 0.69

Many predicted positives are not TP (many FP)

Recall 0.97

Most actual positives correctly predicted (few FN)

F1-score

$$\frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

- Harmonic mean of precision and recall
  - Harmonic mean
    - Reciprocal of the arithmetic mean of the reciprocals of the given set of observations
    - Used for averaging rates/ratios (respect proportionality links)
      - Ex : calculate the average speed of a round trip when the speed of the outbound trip is not the same as the return trip

F1-score

- Harmonic mean of precision and recall

$$F1score = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = \frac{2}{\frac{precision + recall}{precision * recall}} = 2 * \frac{precision * recall}{precision + recall}$$

$$F1score = \frac{2}{\frac{1}{\frac{TP}{TP+FP}} + \frac{1}{\frac{TP}{TP+FN}}} = \frac{2}{\frac{TP+FP+TP+FN}{TP}} = \frac{2 * TP}{2 * TP + FP + FN} \rightarrow$$

ratio/# of FP and FN respected in the harmonic mean calculation

Penalize optimistic and pessimistics models in the exact same way

F1-score

		Predicted value	
		Immune	No imm.
Actual value	Immune	45	15
	No imm.	13	27

$$Precision = \frac{TP}{TP + FP} = \frac{45}{45 + 13} = 0{,}78$$

$$Recall = \frac{TP}{TP + FP} = \frac{45}{45 + 15} = 0{,}75$$

$$F1score = 2 * \frac{precision + recall}{precision * recall} = 2 * \frac{0.78 * 0.75}{0.78 + 0.75} = 2 * \frac{0.585}{1{,}53} = 0{,}76$$

$$F1score = \frac{2 * TP}{2 * TP + FP + FN} = \frac{2 * 45}{2 * 45 + 13 + 15} = \frac{90}{118} = 0.76$$

Accuracy, Precision, Recall & F1-Score

Pessimistic model			
		Predicted values	
		Imm.	No I.
Actual values	Imm.	35	25
	No I.	2	38

Accuracy	0.73
Precision	0.95
Recall	0.58
F1-Score	0,72

Perfect model			
		Predicted values	
		Imm.	No I.
Actual values	Imm.	60	0
	No I.	0	40

Accuracy	1.00
Precision	1.00
Recall	1.00
F1-Score	1.00

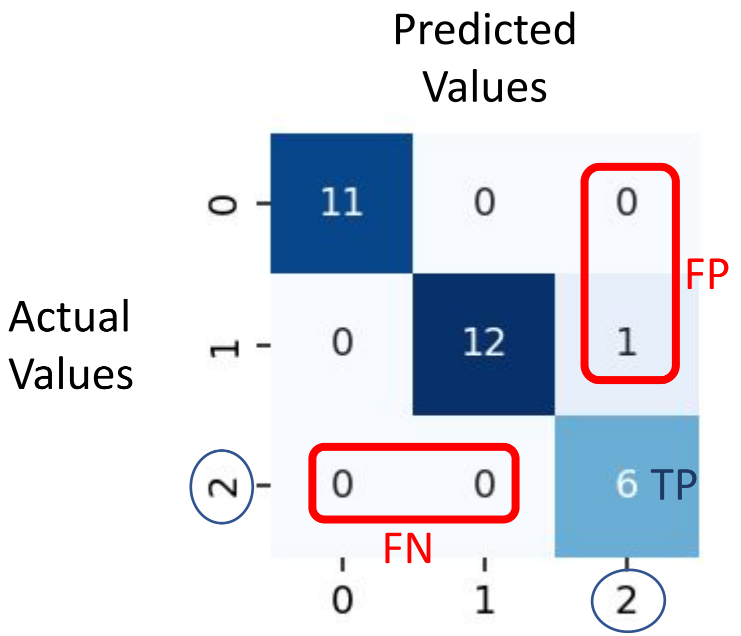
Optimistic model			
		Predicted values	
		Imm.	No I.
Actual values	Imm.	58	2
	No I.	25	15

Accuracy	0.73
Precision	0.69
Recall	0.97
F1-Score	0,81

Multiclass classification

```python
from sklearn.metrics import confusion_matrix

conf_matrix = confusion_matrix(y_test, y_pred)
fig = plt.figure (figsize=(2,2))
sns.heatmap(conf_matrix, annot=True, cmap="Blues", cbar=False)
```




```python
print(classification_report(y_test, y_pred, target_names=iris_data.target_names))
```

	precision	recall	f1-score	support
setosa	1.00	1.00	1.00	11
versicolor	1.00	0.92	0.96	13
virginica	0.86	1.00	0.92	6
accuracy			0.97	30
macro avg	0.95	0.97	0.96	30
weighted avg	0.97	0.97	0.97	30

cells size

Simple mean

Mean weighted by cells size