Evaluating Classifiers

Class	02_titanic
Created	@Nov 24, 2020 1:26 PM
Materials	
Reviewed	☐
Type	

Evaluating Classifiers

0) What we have done so far

- Introduced the concept of machine learning
- Introduced a typical machine learning workflow
- Introduced first machine learning model → Logistic Regression

1) Evaluating Classifiers

Q: What was the first step of the machine learning workflow?

Define a business goal. One component of this business goal will often be to design a *good* model.

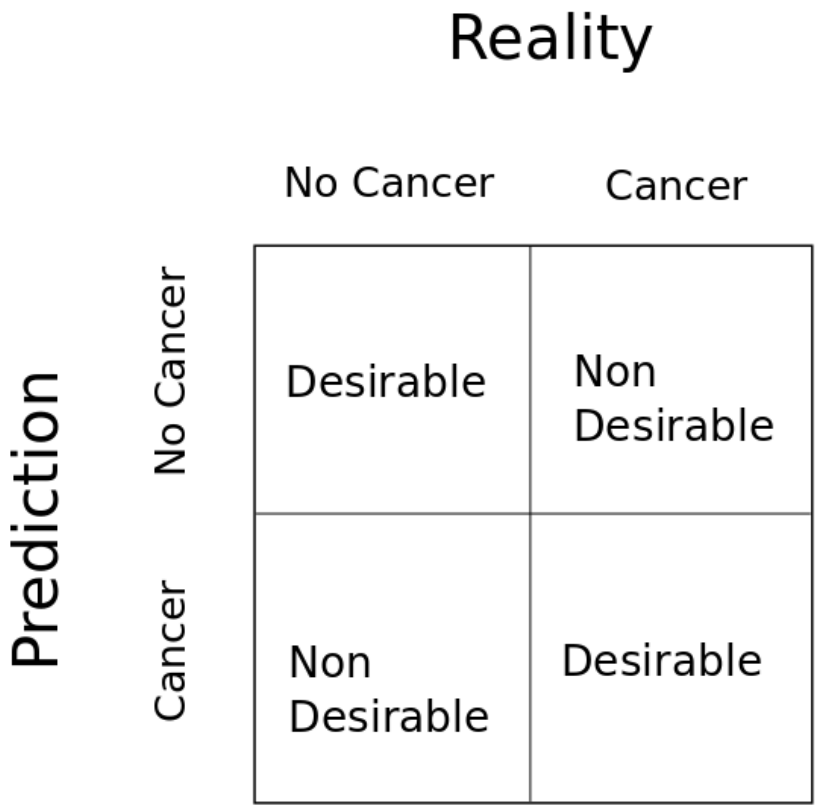
We want a good model.

How can we measure the quality of the model?

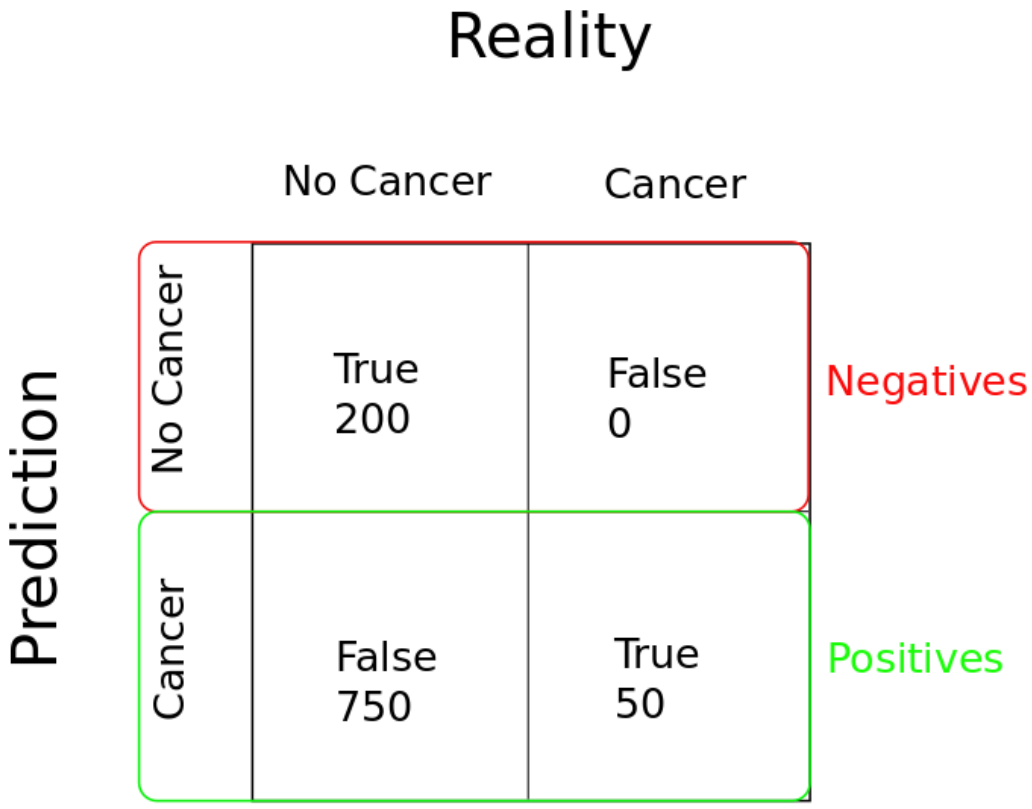
- Accuracy: Share of correctly classified observations. In the first model of the warmup the accuracy would be (50+200)/1000 = 0.25; For the second model the accuracy would be (40+890)/1000=0.93
- One important point to note here is that accuracy might not be the only metric of interest. Although model 2 is much more accurate, we might still prefer model 1 in this case because it does not miss any patients with high risk of skin cancer

- What other metrics do we have at our disposal

1.1) The Confusion Matrix



Positives and Negatives



TP: True Positive

TN: True Negative

FP: False Positive

FN: False Negative

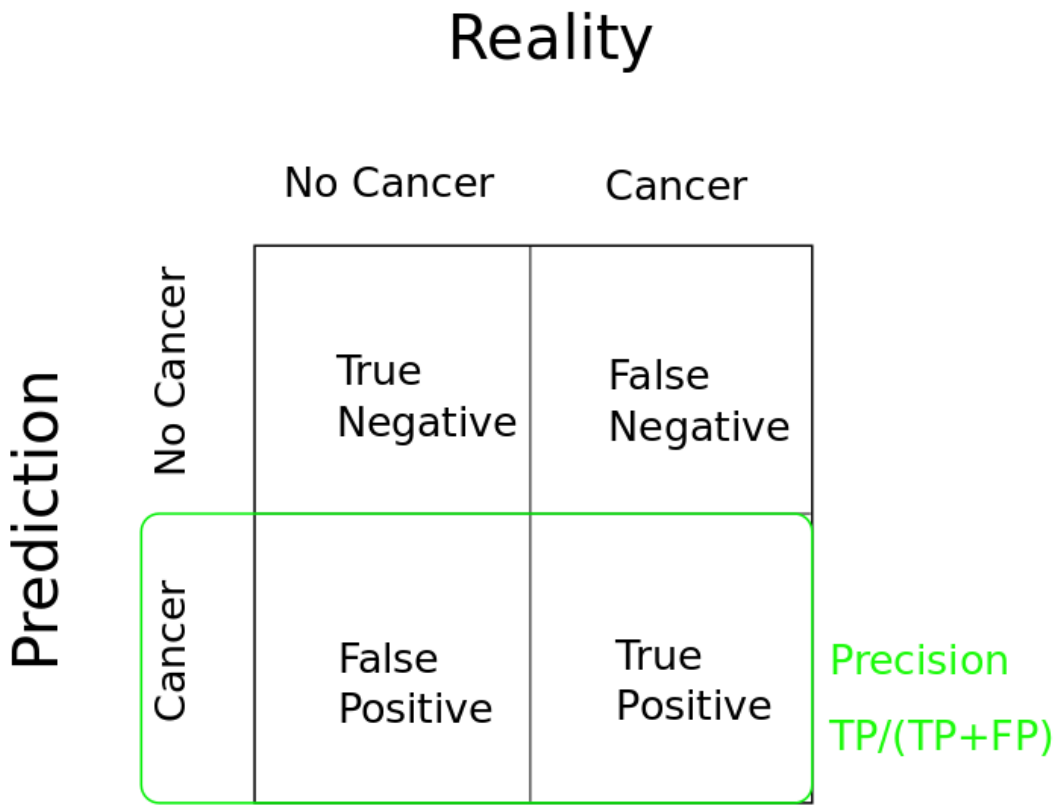
1.2) Accuracy

$$\frac{TP+TN}{TP+FP+TN+FN}$$

1.3) Precision

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

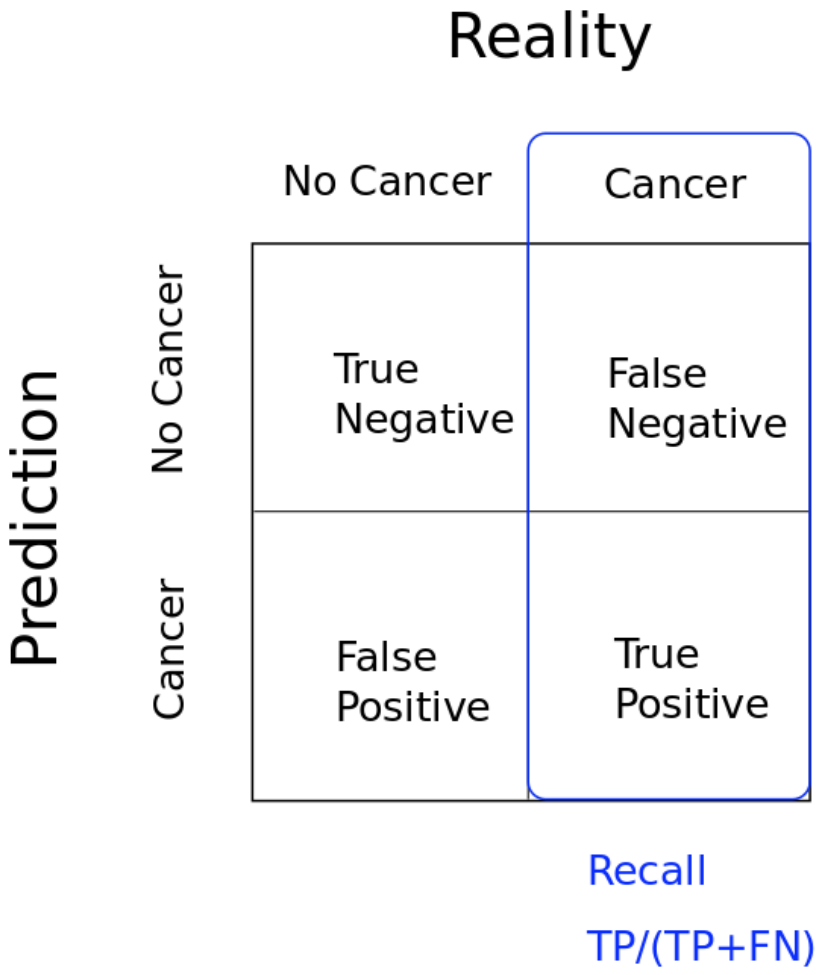
Precision answers the question: given the model classified an observation as positive, how sure can we be that this observation is actually positive?


Reality
No Cancer
Cancer
Prediction
No Cancer
Cancer
True Negative
False Negative
False Positive
True Positive
Precision
TP/(TP+FP)

1.4) Recall (Sensitivity)

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

Recall is a measure of the share of truly positive cases that our model is able to capture.


Reality
No Cancer
Cancer
Prediction
No Cancer
Cancer
True Negative
False Negative
False Positive
True Positive
Recall
TP/(TP+FN)

1.5) F1-Score

$2 * \frac{precision * recall}{precision + recall}$