Cost-based evaluation of classifiers

Lecture 21

by Marina Barsky

The Inadequacy of success rates

- As the class distribution becomes more skewed, evaluation based on success rate breaks down.
  - Consider a dataset where the classes appear in a 999:1 ratio.
  - A simple rule, which classifies every instance as the majority class, gives a 99.9% accuracy – no further improvement is needed!

- Evaluation by classification success rate assumes equal error costs---that a false positive error is equivalent to a false negative error.
  - In the real world this is rarely the case, because classifications lead to actions which have consequences, sometimes grave.

Cost-based evaluation

- In practice, different types of classification errors often incur different costs

- The rare class is often denoted as positive (HIV from test results)
- The *confusion matrix*:

		Predicted class	
		Yes	No
Actual class	Yes	True positive	False negative
	No	False positive	True negative

Terminology

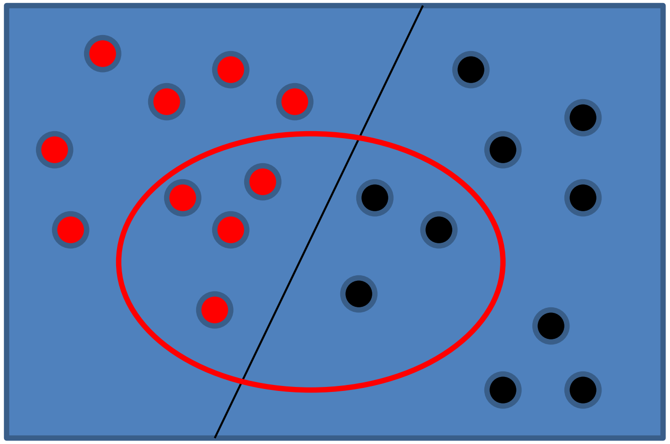
- The *confusion matrix*:

		Predicted class	
		Yes	No
Actual class	Yes	True positive	False negative
	No	False positive	True negative

True positives (TP) – the number of positive examples correctly predicted as positives
False negatives (FN) – the number of positive examples wrongly predicted as negatives
False positives (FP) – the number of negative examples wrongly predicted as positives
True negatives (TN) – the number of negative examples correctly predicted as negatives

Terminology. Fractions

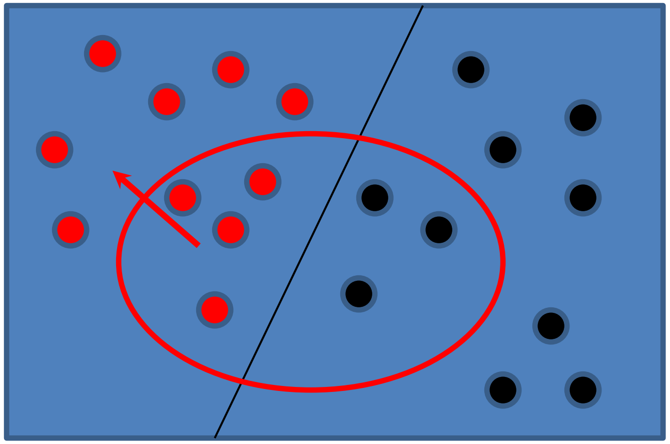
- Suppose you know what are all positive instances in your dataset (red dots)
- The classifier outputs as positives the instances inside the red oval



True Positive Fraction of All Positives

- True positive rate (fraction):
  TPF=TP/all positives
- In the example: 4 red dots predicted as positives out of 10 red dots → TPF=0.4
- Also called: ***sensitivity*** or ***recall***

High sensitivity or high recall mean that classifier found most of the relevant positive instances



Examples:

High-sensitive HIV test- if the person is sick, it will be diagnosed with high-probability:

sick $\xrightarrow[\text{high probability}]{}$ positive test

High-recall document query: the query brought most of the relevant documents

True Positive Fraction of All ***Classified as Positives***

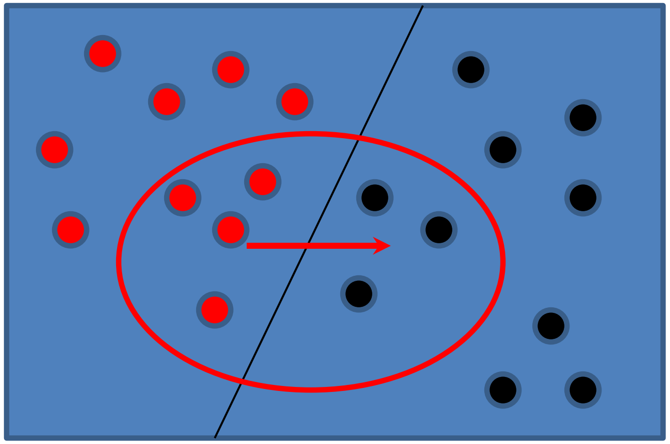
- *Precision* (fraction):

  precision=TP/(TP+FP)

- In the example: 4 red dots out of 7 total dots which are identified as positive

Precision=4/7

- High precision means that classifier returned more relevant results than irrelevant



Example:

Highly precise HIV test – whoever is classified as HIV-positive is most probably sick:

positive test → HIV

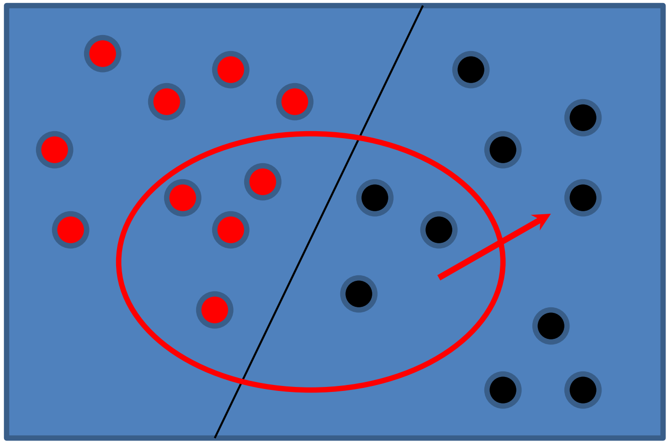
high probability

Terminology. False Positive Fractions

- False Positive Rate(fraction):

  FPF=FP/(all negatives)
- In the example: 3 black dots predicted as positives out of 10 total negative dots

  FPF=3/10
- High FPF means that classifier often classifies negative as positive



Example: mammography

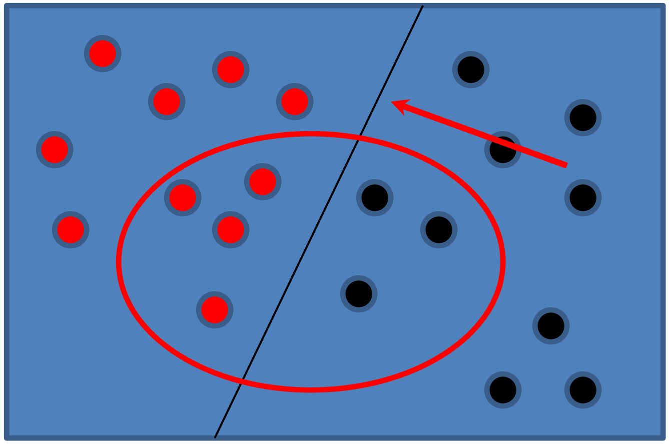
If the person is diagnosed, it is not very likely that the person is really sick:
positive test $\xrightarrow[\text{low probability}]{}$ sick

True Negative Fraction of All Negatives

- *Specificity* (fraction):

  specificity=TN/(all negatives)

- In the example: 7 black dots which are left outside of the positive prediction out of total 10 negative instances



Specificity=7/10

- High specificity means that if classifier identifies something as negative, it is a high probability that it is indeed negative

Highly-specific test means that it has a very low probability to be classified as positive, if the person is indeed negative:

$\text{negative} \xrightarrow[\text{low probability}]{} \text{positive test}$

Specificity + FPF=1.00

Incorporating the cost: Example

Cost matrix

		Predicted class	
		Class +	Class -
Actual class	Class +	-1	100
	Class -	1	0

For example, HIV diagnostic test

- A cost matrix encodes the **penalty** of classifying records of one class as another
- A negative value represents an award for making a correct classification

Counting the cost. Example

		Predicted class	
		Class +	Class -
Actual class	Class +	-1	100
	Class -	1	0

Cost matrix

		Predicted class	
		Class +	Class -
Actual class	Class +	150	40
	Class -	60	250

Confusion matrix for Classifier A

		Predicted class	
		Class +	Class -
Actual class	Class +	250	45
	Class -	5	200

Confusion matrix for Classifier B

The total cost of model A=150*(-1)+60*1+40*100=3910

The total cost of model B=250*(-1)+5*1+45*100=4255

If not take cost into account: B is better than A

		Predicted class	
		Class +	Class -
Actual class	Class +	-1	100
	Class -	1	0

Cost matrix

		Predicted class	
		Class +	Class -
Actual class	Class +	150	40
	Class -	60	250

Classifier A

		Predicted class	
		Class +	Class -
Actual class	Class +	250	45
	Class -	5	200

Classifier B

The total cost of model A=150*(-1)+60*1+40*100=3910

The total cost of model B=250*(-1)+5*1+45*100=4255

Cost matrix example 1

- HIV diagnostic test

		Predicted class	
		Class +	Class -
Actual class	Class +	-100	10000
	Class -	10	0

Person dies untreated and infects others

Cost of additional testing plus some discomfort

Cost matrix example 2

- Promotional mailing

		Predicted class	
		Class +	Class -
Actual class	Class +	-1000	1000
	Class -	1	0

Loses potential revenue

Cost of mailing

Cost matrix example 3

- Loan decisions

		Predicted class	
		Class +	Class -
Actual class	Class +	-100	10
	Class -	50	0

Loses potential revenue

bankruptcy

Cost-based classification

- Let {**p**,**n**} be the positive and negative instance classes.
- Let {**Y**,**N**} be the classifications produced by a classifier.
- Let c(**Y**,**n**) be the cost of a false positive error.
- Let c(**N**,**p**) be the cost of a false negative error.

- For an instance E,
  - the classifier computes p(**p**) and p(**n**)=1- p(**p**)
  - the decision to emit a positive classification should be:

$$[1-p(\mathbf{p})]*c(\mathbf{Y},\mathbf{n}) < p(\mathbf{p}) * c(\mathbf{N},\mathbf{p})$$

Cost-based classification: example

		Predicted class	
		Class +	Class -
Actual class	Class +	-100	10000
	Class -	10	0

Positive if:

$[1-p(\mathbf{p})]*c(\mathbf{Y},\mathbf{n}) < p(\mathbf{p}) * c(\mathbf{N},\mathbf{p})$

The probabilistic classifier returned the probability that you have HIV:

p(HIV) = 0.90

p(not HIV) = 0.10

[1-p(p)]*c(Y,n) = 0.10 * 10 = 1

p(p) * c(N,p) = 0.90 * 10000 = 9000

Classify as positive: 1<<<9000

Cost-based classification: example

		Predicted class	
		Class +	Class -
Actual class	Class +	-100	10000
	Class -	10	0

Positive if:

$$[1-p(\mathbf{p})]*c(\mathbf{Y},\mathbf{n}) < p(\mathbf{p}) * c(\mathbf{N},\mathbf{p})$$

p(HIV) = 0.3

p(not HIV) = 0.7

[1-p(p)]*c(Y,n) = 0.7 * 10 = 7

p(p) * c(N,p) = 0.3 * 10000 = 3000

Classify as positive: 7<<<3000 (still classify as positive, because it is better to be false positive than false negative in this case)