ML INFOSEC
6: Performance Measures

May 10, 2019

The setting

- A set of samples X
- A decomposition

$$X = X_+ \dot{\cup} X_-, \quad P = |X_+|, N = |X_-|$$

into sets of samples marked "positive" and "negative", resp.

- A classifier $C : X \to \{0, 1\}$
- A decomposion of X induced by C:

$$X = \hat{X}_+ \dot{\cup} \hat{X}_-$$

with

$$\hat{X}_+ = \{x \in X \mid C(x) = 1\}$$

and

$$\hat{X}_- = \{x \in X \mid C(x) = 0\}$$

True/False Positive/Negatives

Set of **True Positives**: $X_+ \cap \hat{X}_+$

$$TP = |X_+ \cap \hat{X}_+|$$

Set of **True Negatives**: $X_- \cap \hat{X}_-$

$$TN = |X_- \cap \hat{X}_-|$$

Set of **False Positives**: $X_- \cap \hat{X}_+$

$$FP = |X_- \cap \hat{X}_+|$$

Set of **False Negatives**: $X_+ \cap \hat{X}_-$

$$FN = |X_+ \cap \hat{X}_-|$$

The confusion matrix

	Positive samples (MW)	Negative samples (BW)
Classified as pos.	TP	FP
Classified as neg.	FN	TN

Parameters I

Base rate

$$BR = \frac{|X_+|}{|X|} = \frac{P}{P+N}$$

Accuracy

$$\text{accuracy} = \frac{|X_+ \cap \hat{X}_+| + |X_- \cap \hat{X}_-|}{|X|} = \frac{TP + TN}{P + N}$$

Precision

$$\text{precision} = \frac{|X_+ \cap \hat{X}_+|}{|\hat{X}_+|} = \frac{TP}{TP + FP}$$

Parameters II

Sensitivity $=$ Recall $=$ True Positive Rate

$$TPR = \frac{|X_+ \cap \hat{X}_+|}{|X_+|} = \frac{TP}{P} = \frac{TP}{TP + FN}$$

Specificity $=$ True Negative Rate

$$TNR = \frac{|X_- \cap \hat{X}_-|}{|X_-|} = \frac{TN}{N} = \frac{TN}{TN + FP}$$

False Positive Rate

$$FPR = \frac{|X_- \cap \hat{X}_+|}{|X_-|} = \frac{FP}{N} = 1 - TNR$$

Parameters III

Remark 1

We may assume that $TPR \geq FPR$ because otherwise we can replace C by $1 - C$.

Remark 2

$$\text{precision} = \frac{TPR \times BR}{TPR \times BR + FPR \times (1 - BR)}$$

Example: Let $TPR = 0.8$ and $FPR = 0.1$:

BR	Precision
0.5	0.8889
0.1	0.4706
0.01	0.0748

Scoring I

In general, classifiers are derived from score functions

$$S : X \to [0,1].$$

For each **threshold** $\tau \in [0,1]$ the score function S induces a classifier

$$C_\tau(x) = \begin{cases} 1, & S(x) \geq \tau \\ 0, & S(x) < \tau, \end{cases}$$

on X. $\hat{X}_+$ and $\hat{X}_-$ depend on τ:

$$\hat{X}_+(\tau) = \{x \in X \mid S(x) > \tau\},$$

$$\hat{X}_-(\tau) = \{x \in X \mid S(x) \leq \tau\},$$

so do all parameters introduced above.

Scoring II

The functions

$$[0,1] \ni \tau \mapsto TPR(\tau) \in [0,1]$$

and

$$[0,1] \ni \tau \mapsto FPR(\tau) \in [0,1]$$

are decreasing with

$$TPR(1) = 0, \quad FPR(1) = 0$$

The curve

$$[0,1] \ni \tau \mapsto (FPR(\tau), TPR(\tau)) \in [0,1] \times [0,1]$$

is called **Receiver Operating Characteristic** (ROC) curve. The area under [this] curve (AUC) is called ROC-AUC and is another performace measure (the closer to 1 the better).

Receiver Operating Characteristic (ROC) - Area under curve (AUC)

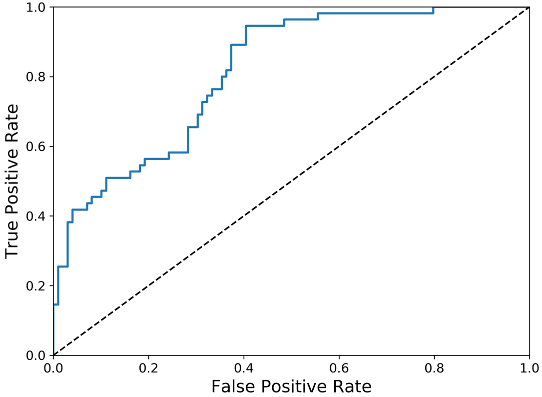

Figure: ROC curve: Gaussian Naive Bayes, Pima diabetes dataset