CISC 271   Class 22

Classification – Assessment With Confusion Matrix

Text Correspondence: Fawcett, 2006 [4]

Main Concepts:

- *Comparing labels to predictions*
- *Confusion matrix*
- *Sensitivity, specificity, Type I error, Type II error*

Sample Problem, Machine Inference: How successful is a classification algorithm?

We now understand that an optimal answer to a problem may require additional information. For example, is we use a hyperplane in binary classification, changing a value such as the bias scalar can be expected to affect the performance of an optimal classifier.

We can explore this classification performance by going beyond the number of "right" and "wrong" answers provided by the classifier. let us begin by recalling that a *label* in a binary classification problem is a value, such as +1 or -1, that represents the subset that the data vector presumably is in. The *prediction* is also a value, such as +1 or -1; the predicted class may differ from the label. The four combinations, of two labels and two predicted classes, are typically represented as a 2×2 *confusion matrix*. We will use these common abbreviations for the relevant terms:

$$
\begin{array}{lll}
\text{P} & : & \text{Positive instances, by label} \\
\text{N} & : & \text{Negative instances, by label} \\
\text{TP} & : & \text{True Positives, label is } +1 \text{ and prediction is } +1 \\
\text{TN} & : & \text{True Negatives, label is } -1 \text{ and prediction is } -1 \\
\text{FP} & : & \text{False Positives, label is } -1 \text{ and prediction is } +1 \\
\text{FN} & : & \text{False Negatives, label is } +1 \text{ and prediction is } -1
\end{array}
\tag{22.1}
$$

Terminology for the rows, columns, and entries differ considerably; this course uses the `MATLAB` convention, which is summarized in Table 22.1.

Table 22.1: A *confusion matrix*, which is an instance of a contingency table. The columns are the "true" labels of the class of a data item. The rows are the predicted classes of a data item.

		Predicted Class		**Row**
		+1	**−1**	**Totals**
Data Label	**+1**	True Positives	False Negatives	**P**
	−1	False Positives	True Negatives	**N**

Many terms that are commonly used in empirical data analysis can be derived from Definition 22.1. Some commonly encountered terms are summarized in Table 22.2.

Table 22.2: Commonly used terms that are derived from a confusion matrix.

Name	Formula	Alternative Names
Sensitivity	$\text{TPR} = \dfrac{\text{TP}}{\text{P}}$	True positive rate, hit rate, recall
Specificity	$\text{TNR} = \dfrac{\text{TN}}{\text{N}}$	True negative rate, selectivity
Accuracy	$\text{ACC} = \dfrac{\text{TP+TN}}{\text{P+N}}$	
Type I error	FP	False alarm
Type I error rate	$\text{FPR} = \dfrac{\text{FP}}{\text{N}}$	False-alarm rate
Type II error	FN	Miss
Type II error rate	$\text{FNR} = \dfrac{\text{FN}}{\text{P}}$	Miss rate
Precision	$\text{PPV} = \dfrac{\text{TP}}{\text{TP+FP}}$	Positive predictive value

The *accuracy* of a classifier is the number of predictions that match the labels. For example, in a binary classification problem with a linear SVM, the accuracy would be the proportion of the sum of the number of vectors on the positive side of the hyperplane that were given as Label +1, plus the number of vectors on the negative side of the hyperplane that were given as Label -1.

Accuracy is often useful but does not fully capture the performance of a classifier. In some

applications, such as medical diagnosis, the error rates may be of substantial importance. The *Type I error*, or as a relative value the *false positive rate*, measures the number or rate of "false alarms"; in a medical context, this may result in needless medical care such as further diagnostic tests or invasive surgical procedures. The *Type II error*, or as a relative value the *false negative rate*, measures the number or rate of "misses"; in a medical context, this may result in under-treatment and an unreasonable sense of security in both the patient and the care-giver.

Sensitivity and specificity are measures of accuracy that are restricted to only the data with, respectively, labels of +1 and -1. If a classifier has a hyper-parameter that changes its performance, the sensitivity and specificity will co-vary and can be assessed together.

In some applications, the *F-score* or *F-measure* is used. The F-score is defined as the harmonic mean of precision and recall. Because the F-score does not account for the true negatives (TN), care must be taken when using the F-score to evaluate a classifier.

22.1 Relative Confusion Matrix

In many applications, the absolute numbers of true positive instances, etc., are not used. Instead, the *rates* or relative proportions of the values from Definition 22.1 are preferred. We can use these rates, which are defined in Table 22.2, to define a "relative" confusion matrix such as that in Table 22.3.

Table 22.3: A *relative confusion matrix* is a contingency table that uses rates instead of the numbers of instances. The rows each sum to 1.

		Class	
		+1	**−1**
Label	**+1**	TPR	FNR
	−1	FPR	TNR

We can see that the relative confusion matrix of Table 22.3 has rows that sum to 1, because

$$\begin{aligned}
\text{TPR} + \text{FNR} &= \frac{\text{TP}}{\text{P}} + \frac{\text{FN}}{\text{P}} = 1 \\
\text{FPR} + \text{TNR} &= \frac{\text{FP}}{\text{N}} + \frac{\text{TN}}{\text{N}} = 1
\end{aligned} \tag{22.2}$$

Equation 22.2 has an important implication for us:

A relative confusion matrix has two degrees of freedom

A general confusion matrix, such as that of Table 22.1, has four degrees of freedom. This is because each of the entries is an independent count that depends on the labels and predicted classes. The row sums, which are **P** and **N** that are respectively the number of data with positive labels and negative labels, are dependent on the independent counts.

A relative confusion matrix has removed two of these degrees of freedom. By requiring that the TPR and FNR add up to 1, if we are given one of these rates then we can deduce the other. For example, given the TPR which is the true-positive rate, we can find the FNR as

$$\text{FNR} = 1 - \text{TPR} \tag{22.3}$$

Likewise, if we are given the FPR which is the false-positive rate, we can find the TNR as

$$\text{TNR} = 1 - \text{FPR} \tag{22.4}$$

This has a second useful implication for us:

A relative confusion matrix is a point in 2D

We will use this second implication to help us understand a set of confusion matrices as a curve in 2D.