Model Evaluation
Confusion Matrix and Prediction Accuracy Measures

“Take a step back, evaluate what is important, and enjoy life” Teri Garr

Anasse Bari, Ph.D.
Defining Confusion Matrix

- Confusion matrix is a two dimensional table that is used as a description of predictive model’s performance.

- The performance is being calculated based on a test data where labels' of data records are known.
Example

Confusion matrix for a binary classifier

<table>
<thead>
<tr>
<th></th>
<th>Predicted: No</th>
<th>Predicted: Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual: No</strong></td>
<td>50 (True Positives)</td>
<td>10 (False Negatives)</td>
</tr>
<tr>
<td><strong>Actual: Yes</strong></td>
<td>5 (False Positives)</td>
<td>100 (True Negatives)</td>
</tr>
</tbody>
</table>

There are two possible predicted classes: "yes" and "no".

If we were predicting the presence of a disease, for example, "yes" would mean they have the disease, and "no" would mean they don't have the disease.

Reference: http://www.dataschool.io/
Example

Confusion matrix for a binary classifier

In reality, 105 patients in the sample have the disease, and 60 patients do not.

Test data consists of 165 test cases.
The classifier made a total of 165 predictions (e.g., 165 patients were being tested for the presence of that disease).

<table>
<thead>
<tr>
<th>Actual: No</th>
<th>Predicted: No</th>
<th>Predicted: Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual: No</td>
<td>50 (True Positives)</td>
<td>10 (False Negatives)</td>
</tr>
<tr>
<td>Actual: Yes</td>
<td>5 (False Positives)</td>
<td>100 (True Negatives)</td>
</tr>
</tbody>
</table>

Out of those 165 cases, the classifier predicted "yes" 110 times, and "no" 55 times.
Example

Confusion matrix for a binary classifier

<table>
<thead>
<tr>
<th></th>
<th>Predicted: No</th>
<th>Predicted: Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual: No</td>
<td>50 (True Positives)</td>
<td>10 (False Negatives)</td>
</tr>
<tr>
<td>Actual: Yes</td>
<td>5 (False Positives)</td>
<td>100 (True Negatives)</td>
</tr>
</tbody>
</table>

Out of those 165 cases, the classifier predicted "yes" 110 times, and "no" 55 times. In reality, 105 patients in the sample have the disease, and 60 patients do not.

Reference: http://www.dataschool.io/
Explaining Related Terms

- **True positives (TP):** These are cases in which we predicted yes (they have the disease according to your model), and they do have the disease in reality. Instances Predicted *Correctly*

- **True negatives (TN):** We predicted no, and they don't have the disease. Instances Predicted *correctly*

- **False positives (FP):** We predicted yes, but they don't actually have the disease. (Also known as a "Type I error.") Instances Predicted *Incorrectly*

- **False negatives (FN):** We predicted no, but they actually do have the disease. (Also known as a "Type II error.") Instances Predicted *Incorrectly*
Calculating Accuracy

Prediction Accuracy: Overall, how often is the classifier correct?

\[
\frac{TP + TN}{total} = \frac{100 + 50}{165} = 0.91 = 91\%
\]
Calculating Accuracy

**Misclassification Rate:** Overall, how often is it wrong?

\[
\frac{(FP+FN)}{\text{total}} = \frac{10+5}{165} = 0.09 = 9\%
\]

Equivalent to 1 minus Accuracy also known as "**Error Rate**"
Calculating Accuracy

True Positive Rate
When it's actually yes, how often does it predict yes?

\[
\frac{TP}{actual\ yes} = \frac{100}{105} = 0.95 = 95\%
\]

also known as "Sensitivity" or "Recall"
Calculating Accuracy

Specificity
When it's actually no, how often does it predict no?

$$\text{TN/actual no} = \frac{50}{60} = 0.83 = 83\%$$

equivalent to 1 minus False Positive Rate
Calculating Accuracy

Precision
When it predicts yes, how often is it correct?
- \( \frac{TP}{predicted\ yes} = \frac{100}{110} = 0.91 = 91\% \)

Prevalence
How often does the yes condition actually occur in our sample?
- \( \frac{actual\ yes}{total} = \frac{105}{165} = 0.64 = 64\% \)
Positive Predictive Value and Negative Predictive Value

A clinician and a patient have a different question: what is the chance that a person with a positive test truly (predicted test result) has the disease?  
If the subject is in the first row in the table above, what is the probability of being in cell (TP) as compared to cell (FP)? A clinician calculates across the row as follows:

Positive Predictive Value: \( \frac{TP}{(TP+FP)} \times 100 \)

Negative Predictive Value: \( \frac{TN}{(TN+FN)} \times 100 \)

*Positive and negative predictive values are influenced by the prevalence of disease in the population that is being tested. If we test in a high prevalence setting, it is more likely that persons who test positive truly have disease than if the test is performed in a population with low prevalence.*
More to know (cont’d)

F Score (F Measure)

This is a weighted average of the true positive rate (recall) and precision.

\[ F1 = 2 \frac{pr}{p + r} \]

where \( p = \frac{tp}{tp + fp} \), \( r = \frac{tp}{tp + fn} \)

ROC Curve (Receiver Operating Characteristic Curve)

This is a commonly used graph that summarizes the performance of a classifier over all possible thresholds.

It can be generated by plotting the True Positive Rate (y-axis) against the False Positive Rate (x-axis) as you vary the threshold and heuristics for assigning observations to a given class.

(See last slide – references for more details about ROC Curves.)
Reference

• [1] Data Schooll, Simple Guide to Confusion Matrix:  
  http://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/  
  http://www.dataschool.io/roc-curves-and-auc-explained/

• [2] Jiawei Han et al.: *Data Mining: Concepts and Techniques, 3rd ed*

More Resources:

https://onlinecourses.science.psu.edu/stat507/node/71