<u>Pattern Recognition Numerical</u>

n=165	Predicted: NO	Predicted: YES
Actual: NO	50	10
Actual: YES	5	100

What can we learn from this matrix?

- 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.
- The classifier made a total of 165 predictions (e.g., 165 patients were being tested for the presence of that disease).
- 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.

Let's now define the most basic terms, which are whole numbers (not rates):

- **true positives (TP):** These are cases in which we predicted yes (they have the disease), and they do have the disease.
- **true negatives (TN):** We predicted no, and they don't have the disease.
- **false positives (FP):** We predicted yes, but they don't actually have the disease. (Also known as a "Type I error.")
- **false negatives (FN):** We predicted no, but they actually do have the disease. (Also known as a "Type II error.")

Now, added these terms to the confusion matrix, and also added the row and column totals:

n=165	**Predicted: NO**	**Predicted: YES**	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

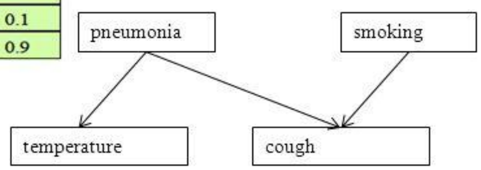
This is a list of rates that are often computed from a confusion matrix for a binary classifier:

- **Accuracy:** Overall, how often is the classifier correct?
  - (TP+TN)/total = (100+50)/165 = 0.91
- **Misclassification Rate:** Overall, how often is it wrong?
  - (FP+FN)/total = (10+5)/165 = 0.09
  - equivalent to 1 minus Accuracy
  - also known as "Error Rate"
- **True Positive Rate:** When it's actually yes, how often does it predict yes?
  - TP/actual yes = 100/105 = 0.95
  - also known as "Sensitivity" or "Recall"
- **False Positive Rate:** When it's actually no, how often does it predict yes?
  - FP/actual no = 10/60 = 0.17
- **True Negative Rate:** When it's actually no, how often does it predict no?
  - TN/actual no = 50/60 = 0.83
  - equivalent to 1 minus False Positive Rate
  - also known as "Specificity"
- **Precision:** When it predicts yes, how often is it correct?
  - TP/predicted yes = 100/110 = 0.91
- **Prevalence:** How often does the yes condition actually occur in our sample?
  - actual yes/total = 105/165 = 0.64

- **Diagnostic: Evidence:** *cough=true.* What is *P*(*pneumonia* | *cough*)?

pneumonia	
true	0.1
false	0.9

smoking	
yes	0.2
no	0.8



	temperature	
pneumonia	yes	no
yes	0.9	0.1
no	0.2	0.8

		cough	
pneumonia	smoking	true	false
true	yes	0.95	0.05
true	no	0.8	0.2
false	yes	0.6	0.4
false	no	0.05	0.95

$$P(pneumonia \mid cough) = \frac{P(cough \mid pneumonia)P(pneumonia)}{P(cough)}$$

$$= \frac{[P(cough \mid pneumonia, smoking)P(smoking) + P(cough \mid pneumonia, \neg smoking)P(\neg smoking)]P(pneumonia)}{P(cough)}$$

$$= \frac{[(.95)(.2)+(.8)(.8)](.1)}{P(cough)} = \frac{.083}{P(cough)}$$

$$= \frac{.083}{P(cough)} = \frac{.083}{.227} = .366$$

Q. While Watching a game of Champions League football in a Cafe, you observe someone who is clearly supporting Manchester United in the game. Using Bayes Rule Calculate the probability that they ~~are~~ were actually born within 30 miles of Manchester.? Assume that:

- The probability that a randomly selected person in a typical local bar environment is born within 30 miles of Manchester is $\frac{1}{20}$
- The chance that a person born within 30 miles of Manchester actually supports Manchester United is 7/10
- The probability that a person not born within 30 miles of Manchester supports Manchester United with probability 1/10

Solution:

Let M: Set of born within 30 miles of Manchester

N: " " Not " " " " " "

S: Set of Supporter of Manchester

Here, given,

$$P(M) = \frac{1}{20}$$

$$P(N) = 1 - \frac{1}{20} = \frac{19}{20}$$

$$P(M|S) = \frac{7}{10}$$

$$P(N|S) = 1/10$$

$$P(S|M) = ?$$

$$P(S|M) = \frac{P(M)\cdot P(M/S)}{P(M)\cdot P(M|S) + P(N)\cdot P(N|S)}$$

$$= \frac{\frac{1}{20} * \frac{7}{10}}{\frac{1}{20} * \frac{7}{10} + \frac{19}{20} * \frac{1}{20}}$$

$$= \frac{7}{26}$$