Naive Bayes and Evaluation Metrics Review *(find and work with a partner)*

The first two columns below represent the *predictions* of a Naive Bayes model on $n = 10$ examples from a test dataset. The third column shows the true label (this is a binary classification task).

$p(y = 0 \mid \vec{x})$	$p(y = 1 \mid \vec{x})$	pred: $t = 0.2$	pred: $t = 0.5$	pred: $t = 0.8$	y (true)
0.95	0.05				0
0.1	0.9				1
0.97	0.03				0
0.93	0.07				1
0.19	0.81				1
0.79	0.21				0
0.35	0.65				1
0.15	0.85				1
0.8	0.2				0
0.77	0.23				0

1. For each threshold t above, fill in the prediction column with 0's and 1's. A reminder that if $p(y = 1|\vec{x}) \geq t$, then we predict the label 1 (otherwise 0).

2. For each threshold, calculate the *accuracy* of the results.

3. (extra practice outside class) Create a confusion matrix for each threshold t, then compute the FPR and TPR. Finally, use these three points to create a ROC curve. Also include the two points that are always on a ROC curve: (0,0) and (1,1).

4. In the original training data for this Naive Bayes model, say there were 35 examples with label 0 and 15 examples with label 1. How would you estimate the *prior* for our model? Include LaPlace counts as well.

5. Say the first feature x_1 can take on values `red`, `green`, or `blue`, and in the data there are 19 examples where $y = 0$ and x_1 is `red`. Calculate the estimate for:

$$p(x_1 = \text{red}|y = 0) \approx \theta_{0,1,\text{red}} =$$

A reminder that for class label k, with feature j equal to value v, our likelihood estimates are:

$$\theta_{k,j,v} = \frac{N_{k,j,v} + 1}{N_k + |f_j|}$$

where $|f_j|$ is the number of possible values feature j can take on.