

Naive Bayes*(find and work with a partner)*

Say we have two tests for a specific disease. Each test (f_1, f_2) can come back either positive “pos” or negative “neg”, and the true underlying condition of the patient is represented by y ($y = 1$ is “healthy” and $y = 2$ is “disease”). We observe this training data where $n = 7$ and $p = 2$:

\mathbf{x}	f_1	f_2	y
\mathbf{x}_1	pos	neg	1
\mathbf{x}_2	pos	pos	2
\mathbf{x}_3	pos	neg	2
\mathbf{x}_4	neg	neg	1
\mathbf{x}_5	pos	neg	2
\mathbf{x}_6	neg	neg	1
\mathbf{x}_7	neg	pos	2

1. To estimate the probability $p(y = k)$, for $k = 1, 2, \dots, K$, we will use the formula:

$$\theta_k = \frac{N_k + 1}{n + K}$$

where N_k is the count (“Number”) of data points where $y = k$. Compute θ_1 and θ_2 . What would θ_1 and θ_2 be if we in fact had *no* training data?

2. To estimate the probabilities $p(x_j = v | y = k)$ for all features j , values v , and class label k , we will use the formula:

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

where $N_{j,v,k}$ is the count of data points where $x_j = v$ and $y = k$, and $|f_j|$ is the number of possible values that f_j (feature j) can take on. Fill in the following tables with these θ values.

	$f_1=\text{pos}$	$f_1=\text{neg}$
$y = 1$		
$y = 2$		

	$f_2=\text{pos}$	$f_2=\text{neg}$
$y = 1$		
$y = 2$		

3. Say we have a new data point $\mathbf{x}_{\text{test}} = [\text{neg}, \text{pos}]$. Our goal is to predict based on the Naive Bayes posterior probability:

$$p(y = k|\mathbf{x}) \propto p(y = k) \prod_{j=1}^p p(x_j|y = k).$$

In practice, we will compute this probability for each class k , based on our estimates (θ_k and $\theta_{j,v,k}$ terms). Then we will assign this data point the class label with maximum probability:

$$\hat{y} = \arg \max_{k \in \{1, 2, \dots, K\}} p(y = k) \prod_{j=1}^p p(x_j|y = k).$$

For this \mathbf{x}_{test} , compute $p(y = 1|\mathbf{x})$ and $p(y = 2|\mathbf{x})$ and then assign a prediction label \hat{y} . Why do these two probabilities not sum to 1?

4. *Confusion Matrix.* Say for a test dataset with $m = 5$, we have these true labels and predictions:

\mathbf{x}_{test}	y	\hat{y}
\mathbf{x}_1	2	1
\mathbf{x}_2	1	2
\mathbf{x}_3	1	2
\mathbf{x}_4	2	2
\mathbf{x}_5	1	1

Draw a confusion matrix for this dataset with true labels on the rows and predicted labels on the columns. Normalize so that each row sums to 1. What would an *ideal* confusion matrix look like?