Bayesian Learning: In-class exercises, part 3

1. Sixty percent of women in the US are currently married. Ninety percent of married women wear wedding rings. One percent of unmarried women wear wedding rings. Your new neighbor Julia does not wear a wedding ring. What is the probability that she is married?

Assuming Julia is in the US and is a woman:

\( h_1 = \text{Julia is married} \quad P(h_1) = .6 \)

\( h_2 = \text{Julia is not married} \quad P(h_2) = .4 \)

\( D = \text{Julia does not wear a wedding ring} \)

\( P(D \mid h_1) = .1 \)

\( P(D \mid h_2) = .99 \)

\( P(D) = (.1)(.6) + (.99)(.4) = .456 \)

\( P(h_1 \mid D) = \frac{P(D \mid h_1)P(h_1)}{P(D)} = \frac{.1(.6)}{.456} = .13 \)
The following data was collected on several different days, along with whether people rated the day as “nice” or “not nice”. **Temperature** can be either “warm” or “cold”, **Precipitation** can be either “low” or “high”, and **Foggy?** can either be “yes” or “no”.

<table>
<thead>
<tr>
<th>Day</th>
<th>Temperature</th>
<th>Precipitation</th>
<th>Foggy?</th>
<th>Rating (Class)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>warm</td>
<td>low</td>
<td>no</td>
<td>nice</td>
</tr>
<tr>
<td>D2</td>
<td>cold</td>
<td>low</td>
<td>no</td>
<td>nice</td>
</tr>
<tr>
<td>D3</td>
<td>warm</td>
<td>high</td>
<td>no</td>
<td>nice</td>
</tr>
<tr>
<td>D4</td>
<td>cold</td>
<td>high</td>
<td>no</td>
<td>nice</td>
</tr>
<tr>
<td>D5</td>
<td>cold</td>
<td>high</td>
<td>yes</td>
<td>not nice</td>
</tr>
<tr>
<td>D6</td>
<td>cold</td>
<td>low</td>
<td>yes</td>
<td>not nice</td>
</tr>
</tbody>
</table>

Suppose now you are given a new description of a day:

D7 warm high yes ?

How would a Naive Bayes classifier, trained on D1-D6, classify D7? Use probabilities with Laplace smoothing. Note: You only need to compute the probabilities relevant to the test example.

\[
P(\text{nice}) = \frac{4}{6} \quad \quad P(\text{not-nice}) = \frac{2}{6}
\]

\[
P(\text{Temp=\text{warm} | \text{nice}}) = \frac{2}{4} \rightarrow \frac{3}{6} \quad \quad P(\text{Temp=\text{warm} | \text{not-nice}}) = \frac{0}{2} \rightarrow \frac{1}{4}
\]

\[
P(\text{Precipitation=\text{high | nice}}) = \frac{2}{4} \rightarrow \frac{3}{6} \quad \quad P(\text{Precipitation=\text{high | not-nice}}) = \frac{1}{2} \rightarrow \frac{2}{4}
\]

\[
P(\text{Foggy=\text{yes | nice}}) = \frac{0}{4} \rightarrow \frac{1}{6} \quad \quad P(\text{Foggy=\text{yes | not-nice}}) = \frac{2}{2} \rightarrow \frac{3}{4}
\]

D7 Positive: \((\frac{4}{6})(\frac{3}{6})(\frac{3}{6})(\frac{1}{6}) = .02778\)

D7 Negative: \((\frac{2}{6})(\frac{1}{4})(\frac{2}{4})(\frac{3}{4}) = .03125\)

Class is Negative
3. Create a probabilistic model for Gaussian Naïve Bayes using the following training set.

<table>
<thead>
<tr>
<th>Instance</th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>.1</td>
<td>.4</td>
<td>+</td>
</tr>
<tr>
<td>$x_2$</td>
<td>.3</td>
<td>.6</td>
<td>+</td>
</tr>
<tr>
<td>$x_3$</td>
<td>.2</td>
<td>.2</td>
<td>+</td>
</tr>
<tr>
<td>$x_4$</td>
<td>.7</td>
<td>.1</td>
<td>−</td>
</tr>
<tr>
<td>$x_5$</td>
<td>.8</td>
<td>.3</td>
<td>−</td>
</tr>
</tbody>
</table>

\[
\mu_{F_i, POS} = \frac{.1 + .3 + .2}{3} = .2 \quad \sigma_{F_i, POS} = \sqrt{\frac{(2-.1)^2 + (2-.3)^2 + (2-.2)^2}{3}} = .082
\]

\[
\mu_{F_2, POS} = \frac{.4 + .6 + .2}{3} = .4 \quad \sigma_{F_2, POS} = \sqrt{\frac{(4-.4)^2 + (4-.6)^2 + (4-.2)^2}{3}} = .163
\]

\[
\mu_{F_1, NEG} = \frac{.7 + .8}{2} = .75 \quad \sigma_{F_1, NEG} = \sqrt{\frac{(75-.7)^2 + (75-.8)^2}{2}} = .05
\]

\[
\mu_{F_2, NEG} = \frac{.1 + .3}{2} = .2 \quad \sigma_{F_2, NEG} = \sqrt{\frac{(2-.1)^2 + (2-.3)^2}{2}} = .1
\]

(b) Use the model you created in part (a) to determine $class_{NB}$ for the following new instances:
<table>
<thead>
<tr>
<th>Instance</th>
<th>$F_1$</th>
<th>$F_2$</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_6$</td>
<td>.9</td>
<td>.4</td>
<td>Class is Negative</td>
</tr>
</tbody>
</table>

$x_6$ : Positive

$$P(POS) = .6$$

$$PDF(x_1 = .9 \mid POS) = \frac{1}{\sqrt{2\pi (.082)}} e^{-\frac{(9 -.2)^2}{2(.082)^2}} = 7.3 \times 10^{-16}$$

$$PDF(x_2 = .4 \mid POS) = \frac{1}{\sqrt{2\pi (.163)}} e^{-\frac{(4 -.4)^2}{2(.163)^2}} = 2.45$$

$$\ln(.6) + \ln(7.3 \times 10^{-16}) + \ln(2.45) = -34.5$$

$x_6$ : Negative

$$P(NEG) = .4$$

$$PDF(x_1 = .9 \mid NEG) = \frac{1}{\sqrt{2\pi (.05)}} e^{-\frac{(9 -.5)^2}{2(.05)^2}} = .089$$

$$PDF(x_2 = .4 \mid NEG) = \frac{1}{\sqrt{2\pi (.1)}} e^{-\frac{(4 -.2)^2}{2(.1)^2}} = .539$$

$$\ln(4) + \ln(.089) + \ln(.539) = -3.95$$

Class is Negative.