1. Suppose you receive an e-mail message whose subject line is in all capital letters. You have been keeping statistics on your e-mail, and have found that while only 5% of the total e-mail messages you receive are spam, 15% of the spam messages have an all-caps subject line, and 2% of the non-spam messages have an all-caps subject line.

According to Bayes’ theorem, what is the probability the message you received is spam?

Solution:

Data $D$ = email has all-caps subject line

$P(\text{Spam}) = .05$

$P(\text{NotSpam}) = .95$

$P(D \mid \text{Spam}) = .15$

$P(D \mid \text{NotSpam}) = .02$

$P(D) = P(D \mid \text{Spam})P(\text{Spam}) + P(D \mid \text{NotSpam})P(\text{NotSpam})$

$= (.15)(.05) + (.95)(.02) = .0265$

$P(\text{Spam} \mid D) = \frac{P(D \mid \text{Spam})P(\text{Spam})}{P(D)} = \frac{(.15)(.05)}{.0265} = .28$
2. You’ve been hired by Amazon to work on their recommendation system for books.

Your first task is to see if you can predict the recommendations of a single user, using the following two attributes: **Genre (Romance, Self-Help, or Thriller)** and **Price (High, Medium, or Low)**. The classification is whether the user recommended the book or not. You collect the following training set.

<table>
<thead>
<tr>
<th>Book</th>
<th>Genre</th>
<th>Price</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>Romance</td>
<td>Low</td>
<td>Recommended</td>
</tr>
<tr>
<td>B2</td>
<td>Romance</td>
<td>Medium</td>
<td>Recommended</td>
</tr>
<tr>
<td>B3</td>
<td>Thriller</td>
<td>Low</td>
<td>Recommended</td>
</tr>
<tr>
<td>B4</td>
<td>Thriller</td>
<td>Low</td>
<td>Recommended</td>
</tr>
<tr>
<td>B5</td>
<td>Self-Help</td>
<td>High</td>
<td>Not Recommended</td>
</tr>
<tr>
<td>B6</td>
<td>Romance</td>
<td>High</td>
<td>Not Recommended</td>
</tr>
<tr>
<td>B7</td>
<td>Self-Help</td>
<td>High</td>
<td>Not Recommended</td>
</tr>
</tbody>
</table>

Show how a Naïve Bayes classifier, with Laplace smoothing of probabilities, would classify the following new example as “Recommended” or “Not Recommended”

<table>
<thead>
<tr>
<th>Book</th>
<th>Genre</th>
<th>Price</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>B8</td>
<td>Self-Help</td>
<td>Medium</td>
<td></td>
</tr>
</tbody>
</table>

Show your work.

**Solution:**

\[
P(\text{Rec}) = \frac{4}{7} \quad P(\text{NotRec}) = \frac{3}{7}
\]

\[
P(\text{Self Help}|\text{Rec}) = \frac{0}{4} \rightarrow \frac{1}{7} \quad P(\text{Medium}|\text{Rec}) = \frac{1}{4} \rightarrow \frac{2}{7}
\]

\[
P(\text{Self Help}|\text{NotRec}) = \frac{2}{3} \rightarrow \frac{3}{6} \quad P(\text{Medium}|\text{NotRec}) = \frac{0}{3} \rightarrow \frac{1}{6}
\]

\[
\text{Rec: } \left(\frac{4}{7}\right) \left(\frac{1}{7}\right) \left(\frac{3}{7}\right) = .023 \quad \text{NotRec: } \left(\frac{2}{7}\right) \left(\frac{3}{6}\right) \left(\frac{1}{6}\right) = .036
\]

Class is “NotRec”
3. (a) Consider the following training data, where each instance \( x \) is described by two features, \( F_1 \) and \( F_2 \). The possible classes are POS and NEG.

<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>7</td>
<td>1</td>
<td>POS</td>
</tr>
<tr>
<td>( x_2 )</td>
<td>6</td>
<td>2</td>
<td>POS</td>
</tr>
<tr>
<td>( x_3 )</td>
<td>8</td>
<td>3</td>
<td>POS</td>
</tr>
<tr>
<td>( x_4 )</td>
<td>4</td>
<td>6</td>
<td>NEG</td>
</tr>
<tr>
<td>( x_5 )</td>
<td>5</td>
<td>5</td>
<td>NEG</td>
</tr>
<tr>
<td>( x_6 )</td>
<td>6</td>
<td>4</td>
<td>NEG</td>
</tr>
</tbody>
</table>

Give the probabilistic model that Gaussian Naive Bayes would compute—that is, the prior class probabilities and parameters of the relevant Gaussian distributions.

**Hint:** Recall that \( \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i-x)^2}{n}} \)

**Solution:**

\[
P(\text{POS}) = .5 \quad P(\text{NEG}) = .5
\]

\[
\mu_{F_1,\text{POS}} = 7 \quad \sigma_{F_1,\text{POS}} = \sqrt{\frac{(7-7)^2 + (6-7)^2 + (8-7)^2}{3}} = .816
\]

\[
\mu_{F_2,\text{POS}} = 2 \quad \sigma_{F_2,\text{POS}} = .816
\]

\[
\mu_{F_1,\text{NEG}} = 5 \quad \sigma_{F_1,\text{NEG}} = .816
\]

\[
\mu_{F_2,\text{NEG}} = 5 \quad \sigma_{F_2,\text{NEG}} = .816
\]
(b) Using the formula for a Gaussian distribution,

\[ N(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]

show how the Gaussian Naive Bayes classifier from part (a) would classify the instance \( x_7 = (8, 4) \).

**Solution:**

**Positive case:**

\[
P(F_1 = 8|POS) = \frac{1}{(0.816)\sqrt{2\pi}} e^{-\frac{(8-7)^2}{2(0.816)^2}} = \frac{1}{2.045} e^{-\frac{1}{1.33}} = .23
\]

\[
P(F_2 = 4|POS) = \frac{1}{2.045} e^{-\frac{(4-2)^2}{1.33}} = .024
\]

\[
\frac{1}{2} (.23)(.024) = .00276
\]

**Negative case:**

\[
P(F_1 = 8|NEG) = \frac{1}{(0.816)\sqrt{2\pi}} e^{-\frac{(8-5)^2}{2(0.816)^2}} = \frac{1}{2.045} e^{-\frac{9}{1.33}} = .00056
\]

\[
P(F_2 = 4|NEG) = \frac{1}{2.045} e^{-\frac{(4-5)^2}{1.33}} = .23
\]

\[
\frac{1}{2} (.00056)(.23) = .0000644
\]

**Class is Positive.**