CIS4526 Written Assignment II Answers

1. (a) For the first step, the Weekend attribute achieves the max information gain:
   Weekend: \( \text{mrH(pr) + mlH(pl)} = 8H(1/2) + 3H(0) = 8 \)
   Weather: \( \text{mrH(pr) + mlH(pl)} = 5H(1/5) + 6H(1/2) \approx 9.6 \)
   Company: \( \text{mrH(pr) + mlH(pl)} = 4H(1/4) + 7H(3/7) \approx 10.1 \)
   Therefore we first split on weekend attribute.

   If weekend = NO: then Go Hiking = NO.

   If weekend = YES, we need to choose second attribute to split on:
   Weather: \( \text{mrH(pr) + mlH(pl)} = 4H(1/4) + 4H(1/4) \approx 6.4 \)
   Company: \( \text{mrH(pr) + mlH(pl)} = 5H(2/5) + 3H(1/3) \approx 7.6 \)
   Therefore the second attribute will be Weather attribute, and third one will be Company attribute. The decision tree will be as follows:

   (b) probability is 0

   (c) probability is 1/3
2. (a) \( p(B|M) = \frac{p(M|B) \cdot p(B)}{p(M)} \); use Bayes rule
   
   \[ = 0.90 \times 0.01 / 0.10899 \]
   
   \[ = 0.0826 \]

   Based on MAP, \( p(B|M) < p(\sim B|M) \), so the decision is no cancer.
   
   Based on MLE, since the prior probability will not be considered, namely \( p(B) = p(\sim B) \), we will make the opposite decision that there is cancer. The difference lies in whether one takes into account the prior (hence the false positives).

(b) \( p(B|\sim M) = \frac{p(\sim M|B) \cdot p(B)}{p(\sim M)} \); use Bayes rule

\[ = 0.10 \times 0.01 / 0.89 \]

\[ = 0.0011 \]

(c) given: \( p(B) = 0.01; p(M|B) = 0.90; p(\sim M|\sim B) = 0.899 \)

so: \( p(\sim B) = 0.99; p(\sim M|B) = 0.10, p(M|\sim B) = 0.101 \)

\[ p(M) = p(M^B) + p(M^{\sim B}) \]

\[ = p(M|B) \cdot p(B) + p(M|\sim B) \cdot p(\sim B) \]

\[ = 0.9 \times 0.01 + 0.101 \times 0.99 = 0.10899 \]

Finally the probability of a negative mammography test \( P(\sim M) = 1 - p(M) = 0.89 \)

One can use the following illustration to better your understanding in computing the prior, joint, and conditional probabilities.
(a) \[ \alpha P(S\text{ound, Fur, Color } | \text{Class}) P(\text{Class}) \]
\[ = \alpha P(S\text{ound } | \text{Class}) P(\text{Fur } | \text{Class}) P(\text{Color } | \text{Class}) P(\text{Class}) \]

(b) answers below

\[
\begin{align*}
P(\text{Dog}) &= \frac{1}{2} \quad P(\text{Cat}) &= \frac{1}{2} \\
P(\text{Sound}=\text{Meow } | \text{Class}=\text{Dog}) &= \frac{1}{4} \quad P(\text{Sound}=\text{Meow } | \text{Class}=\text{Cat}) &= \frac{3}{4} \\
P(\text{Sound}=\text{Bark } | \text{Class}=\text{Dog}) &= \frac{3}{4} \quad P(\text{Sound}=\text{Bark } | \text{Class}=\text{Cat}) &= \frac{1}{4} \\
P(\text{Fur}=\text{Coarse } | \text{Class}=\text{Dog}) &= \frac{3}{4} \quad P(\text{Fur}=\text{Coarse } | \text{Class}=\text{Cat}) &= \frac{1}{4} \\
P(\text{Fur}=\text{Fine } | \text{Class}=\text{Dog}) &= \frac{1}{4} \quad P(\text{Fur}=\text{Fine } | \text{Class}=\text{Cat}) &= \frac{3}{4} \\
P(\text{Color}=\text{Brown } | \text{Class}=\text{Dog}) &= \frac{1}{2} \quad P(\text{Color}=\text{Brown } | \text{Class}=\text{Cat}) &= \frac{1}{2} \\
P(\text{Color}=\text{Black } | \text{Class}=\text{Dog}) &= \frac{1}{2} \quad P(\text{Color}=\text{Black } | \text{Class}=\text{Cat}) &= \frac{1}{2} \\
\end{align*}
\]

(3) \[
P(\text{Class}=\text{Dog } | \text{Sound}=\text{Bark } \land \text{Fur}=\text{Coarse } \land \text{Color}=\text{Brown})
\[ = \alpha(3/4)(3/4)(1/2)(1/2) = 9/10 \]
\[
P(\text{Class}=\text{Cat } | \text{Sound}=\text{Bark } \land \text{Fur}=\text{Coarse } \land \text{Color}=\text{Brown})
\[ = \alpha(1/4)(1/4)(1/2)(1/2) = 1/10 \]