CSE 151A - Discussion 03

Quick Review

Conditional Independence
Events A and B are conditionally independent given C if:

$$P(A, B|C) = P(A|C) \times P(B|C)$$
$$P(A|B, C) = P(A|C)$$
$$P(B|A, C) = P(B|C)$$

Mutual Conditional Independence
Events $X_1 \dots X_d$ are mutually conditionally independent given Y if:

$$P(X_1 \dots X_d|Y) = P(X_1|Y) \times \cdots \times P(X_d|Y)$$

Naïve Bayes Classifier : Predict class y_i that maximizes $P(Y = y_i|\vec{X} = \vec{x})$ for some input $\vec{x} = x_1 \dots x_d$.
Like before, $\text{argmax}_{y_i} P(Y = y_i|\vec{X} = \vec{x})$ is equivalent to $\text{argmax}_{y_i} P(\vec{X} = \vec{x}|Y = y_i)P(Y = y_i)$
Assuming mutual conditional independence of $\vec{X} = X_1 \dots X_d$ given Y:

$$\text{argmax}_{y_i} P(Y = y_i|\vec{X} = \vec{x}) = \text{argmax}_{y_i} P(X_1 = x_1|Y = y_i) \times \cdots \times P(X_d = x_d|Y = y_i) \times P(Y = y_i)$$

Logarithm Properties
We can use log probabilities to prevent underflow issues by computing sums instead of products:

$$\log \left[P(X_1 = x_1|Y = y_i) \times \cdots \times P(X_d = x_d|Y = y_i) \times P(Y = y_i)\right]$$
$$= \log \left[\prod_{j=1}^{d} P(X_j = x_j|Y = y_i) \times P(Y_i = y_i)\right]$$
$$= \left[\sum_{j=1}^{d} \log P(X_j = x_j|Y = y_i)\right] + \log P(Y_i = y_i)$$

Linear Regression
With $H(x) = w_1x + w_0$,

$$MSE = R_{sq}(w_1, w_0) = \frac{1}{n}\sum_{i=1}^{n}((w_1x + w_0) - y_i)^2$$

Least squares solution,

$$w_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad w_0 = \bar{y} - w_1\bar{x}$$

Linear Algebra

Problem 1.

Given the following events, use intuition to determine if A and B are conditionally independent given C:

a) $A:$ Are people using umbrellas?
$B:$ Is the sidewalk wet?
$C:$ Is it raining outside?

b) $A:$ Was there an earthquake?
$B:$ Was there a burglary?
$C:$ Did the house alarm go off?

Solution:
a) **Yes**. A and B are conditionally independent given C.
If we know the outcome of C, then calculating $P(A|C) = P(A|B, C)$. In simpler terms, the probability that people are using umbrellas given that it is raining outside is not changed when we are also told that the sidewalk is wet. Knowing that the sidewalk is wet did not give us any additional information.
$P(\text{umbrellas}|\text{raining}) = P(\text{umbrellas}|\text{wet sidewalk}, \text{raining})$

b) **No**. A and B are not conditionally independent given C.
In this case, $P(A|C) \neq P(A|B, C)$. If we are trying to predict A given the outcome of C, that probability will not be the same if we are also told the outcome of B. In simpler terms, the probability that there was an earthquake given that the house alarm is going off, is not the same as it would be when we are also told that there was a burglary. Knowing that a burglary occurred gives us lots of relevant information and explains away the likelihood that there was an earthquake.
$P(\text{earthquake}|\text{alarm}) \neq P(\text{earthquake}|\text{burglary}, \text{alarm})$

Problem 2.

A local hospital has collected the symptoms and diagnoses of 20 sick patients. Your friend is sick with a fever, dry cough, and a sore throat, but they are not sneezing. Use the data below and a Naïve Bayes classifier to predict whether your friend has COVID-19, the Cold, or the Flu. Be sure to clearly state your prediction as one of the three illnesses. Show your work.

Patient #	Illness	Fever	Dry Cough	Sore Throat	Sneezing
1	COVID-19	**Yes**	**Yes**	No	No
2	COVID-19	**Yes**	**Yes**	**Yes**	No
3	COVID-19	**Yes**	**Yes**	No	No
4	Cold	No	**Yes**	**Yes**	**Yes**
5	Cold	No	**Yes**	**Yes**	No
6	Cold	No	No	**Yes**	**Yes**
7	Cold	No	No	No	**Yes**
8	Cold	**Yes**	**Yes**	**Yes**	**Yes**
9	Cold	No	**Yes**	**Yes**	**Yes**
10	Cold	No	No	**Yes**	**Yes**
11	Cold	No	No	**Yes**	**Yes**
12	Cold	No	**Yes**	**Yes**	**Yes**
13	Cold	No	**Yes**	No	**Yes**
14	Flu	**Yes**	No	No	No
15	Flu	**Yes**	**Yes**	**Yes**	No
16	Flu	**Yes**	**Yes**	**Yes**	No
17	Flu	**Yes**	**Yes**	**Yes**	No
18	Flu	No	No	**Yes**	No
19	Flu	**Yes**	**Yes**	**Yes**	No
20	Flu	**Yes**	**Yes**	No	No

Solution: We estimate the following probabilities from the table:

$$\begin{aligned}
P(\text{Fever}|\text{COVID-19}) &= 1 \\
P(\text{Cough}|\text{COVID-19}) &= 1 \\
P(\text{Sore Throat}|\text{COVID-19}) &= 1/3 \\
P(\text{No Sneezing}|\text{COVID-19}) &= 1 \\
P(\text{COVID-19}) &= 3/20
\end{aligned}$$

$$\begin{aligned}
P(\text{Fever}|\text{Cold}) &= 1/10 \\
P(\text{Cough}|\text{Cold}) &= 6/10 \\
P(\text{Sore Throat}|\text{Cold}) &= 8/10 \\
P(\text{No Sneezing}|\text{Cold}) &= 1/10 \\
P(\text{Cold}) &= 10/20
\end{aligned}$$

$$\begin{aligned}
P(\text{Fever}|\text{Flu}) &= 6/7 \\
P(\text{Cough}|\text{Flu}) &= 5/7 \\
P(\text{Sore Throat}|\text{Flu}) &= 5/7 \\
P(\text{No Sneezing}|\text{Flu}) &= 1 \\
P(\text{Cold}) &= 7/20
\end{aligned}$$

Therefore:

$$
\begin{aligned}
&P(\text{COVID-19}|\text{Fever, Cough, Sore Throat, No Sneezing}) \\
&\quad \propto P(\text{Fever}|\text{COVID-19}) \\
&\qquad \times P(\text{Cough}|\text{COVID-19}) \\
&\qquad \times P(\text{Sore Throat}|\text{COVID-19}) \\
&\qquad \times P(\text{No Sneezing}|\text{COVID-19}) \\
&\qquad \times P(\text{COVID-19}) \\
&= 1 \cdot 1 \cdot \frac{1}{3} \cdot 1 \cdot \frac{3}{20} = \frac{3}{60}
\end{aligned}
$$

$$
\begin{aligned}
&P(\text{Cold}|\text{Fever, Cough, Sore Throat, No Sneezing}) \\
&\quad \propto P(\text{Fever}|\text{Cold}) \\
&\qquad \times P(\text{Cough}|\text{Cold}) \\
&\qquad \times P(\text{Sore Throat}|\text{Cold}) \\
&\qquad \times P(\text{No Sneezing}|\text{Cold}) \\
&\qquad \times P(\text{Cold}) \\
&= \frac{1}{10} \cdot \frac{6}{10} \cdot \frac{8}{10} \cdot \frac{1}{10} \cdot \frac{10}{20} = \frac{480}{200{,}000}
\end{aligned}
$$

$$
\begin{aligned}
&P(\text{Flu}|\text{Fever, Cough, Sore Throat, No Sneezing}) \\
&\quad \propto P(\text{Fever}|\text{Flu}) \\
&\qquad \times P(\text{Cough}|\text{Flu}) \\
&\qquad \times P(\text{Sore Throat}|\text{Flu}) \\
&\qquad \times P(\text{No Sneezing}|\text{Flu}) \\
&\qquad \times P(\text{Flu}) \\
&= \frac{6}{7} \cdot \frac{5}{7} \cdot \frac{5}{7} \cdot 1 \cdot \frac{7}{20} = \frac{1{,}050}{6{,}860}
\end{aligned}
$$

Since the last probability is the largest, our prediction is that our friend has the flu.

Problem 3.

Suppose that 1% of all computer science students cheat and that a particular cheat-detecting software has a 5% false positive rate and a 10% false negative rate.
Let $A \in \{0, 1\}$ indicate whether a student is cheating, and let $B \in \{0, 1\}$ indicate the result of the cheat-detecting software.

a) Using this information, determine the probability tables for $P(A)$ and $P(B|A)$.
$P(A = 1) =$ $\qquad$ $P(A = 0) =$

A	$P(B = 1\|A)$	$P(B = 0\|A)$
0		
1		

b) The cheat-detecting software just reported that a student named Luke tested positive for cheating. What is the probability that Luke actually cheated?

Solution:

a)

$P(A = 1) = 0.01$ $\qquad$ $P(A = 0) = 0.99$

A	$P(B = 1\|A)$	$P(B = 0\|A)$
0	0.05	0.95
1	0.90	0.10

$P(A = 1) = 0.01$ is given to us from the problem statement.

$P(A = 0) = 1 - P(A = 1) = 1 - 0.01 = 0.99$

The false positive rate, $P(B = 1|A = 0)$, is given as 0.05.

Thus, $P(B = 0|A = 0) = 1 - P(B = 1|A = 0) = 0.95$.

The false negative rate, $P(B = 0|A = 1)$, is given as 0.10.

Thus, $P(B = 1|A = 1) = 1 - P(B = 0|A = 1) = 0.90$.

b) **15.4%**

Here we are trying to compute $P(A = 1|B = 1)$.

$P(A = 1|B = 1) = \frac{P(B=1|A=1)P(A=1)}{P(B=1)}$ by Bayes Theorem

$= \frac{P(B=1|A=1)P(A=1)}{\sum_j P(A=a_j,B=1)}$ by marginalization

$= \frac{P(B=1|A=1)P(A=1)}{\sum_j P(A=a_j)P(B=1|A=a_j)}$ by definition

$= \frac{P(B=1|A=1)P(A=1)}{P(A=0)P(B=1|A=0)+P(A=1)P(B=1|A=1)}$

$= \frac{(0.90)(0.01)}{(0.99)(0.05)+(0.90)(0.01)}$ by substitution

$= 0.154$.

Problem 4.

Here is a dataset of (x, y) pairs:

x	y
4	1
5	6
9	8

a) Given the above data, fit a linear prediction rule by computing w_0 and w_1.

b) Use the prediction rule from part a) to estimate the y value that corresponds to an input of $x = 2$.

c) Calculate the MSE of the prediction rule from part a) with respect to the given (x, y) pairs.

Solution:

a)

$\bar{x} = \frac{4+5+9}{3} = 6$

$\bar{y} = \frac{1+6+8}{3} = 5$

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$
4	1	-2	-4	8	4
5	6	-1	1	-1	1
9	8	3	3	9	9

$w_1 = \frac{8-1+9}{4+1+9} = \frac{16}{14} = \frac{8}{7}$

$w_0 = 5 - \frac{8}{7}(6) = \frac{-13}{7}$

$y = \frac{8}{7}x - \frac{13}{7}$

b) **0.42**
We have $y = \frac{8}{7}x - \frac{13}{7}$ with $x = 2$.

$y = \frac{8}{7}(2) - \frac{13}{7}$

$= \frac{3}{7} = 0.42$

c) **2.57**
$MSE = \frac{1}{n}\sum_{i=1}^{n}((w_1 x + w_0) - y_i)^2$

$= \frac{1}{3}\left[((\frac{8}{7}(4) - \frac{13}{7}) - 1)^2 + ((\frac{8}{7}(5) - \frac{13}{7}) - 6)^2 + ((\frac{8}{7}(9) - \frac{13}{7}) - 8)^2\right]$

$= \frac{1}{3}\left[(\frac{19}{7} - 1)^2 + (\frac{27}{7} - 6)^2 + (\frac{59}{7} - 8)^2\right]$

$= \frac{1}{3}\left[(\frac{12}{7})^2 + (-\frac{15}{7})^2 + (\frac{3}{7})^2\right]$

$= \frac{18}{7} = 2.57$