CS412 Machine Learning
HW 2 – Probabilities – Bayesian Learning
100pts

- **Please TYPE your answer or write legibly by hand (pts off if it is hard to read).**
- **Use this document to type in your answers** (rather than writing on a separate sheet of paper), so as to keep questions, answers and and grades together to facilitate grading.
- **SHOW all your work for partial/full credit.**
- **Allocated spaces should be enough for your answers** (unnecessarily long and irrelevant answers may loose points)

1) 20 pt - Suppose that we have 3 colored boxes r (red), b (blue) and g (green).
Box r contains 9 apples, 5 oranges and 3 limes;
Box b contains 8 apples, 4 oranges and 1 limes;
Box g contains 5 apples, 2 oranges and 6 limes.

Assume a process **where we pick a box first and then pick a fruit from the selected box**. A box is chosen at random according to the following probability of being selected: p(r) = p(b) = 0.3 and p(g)=0.4 and a piece of fruit is selected from the **chosen** box randomly.

a) 10 pt – What is the **probability of selecting an orange**?

Probability of selecting orange = P(orange and red box) or P(orange and blue box) or P(orange and green box)
= (0.3 * 5/17) + (0.3 * 4/13) + (0.4 * 2/13)
= 0.088 + 0.092 + 0.062
= 0.242

b) 10pt - If we **observe that the selected fruit is an orange**, what is the probability that it came from the red box?

P(RED_BOX|ORANGE) = (P(ORANGE | RED_BOX) * P(RED)) / P(ORANGE)
= (5/17 * 0.3) / 0.242
= 0.88 / 0.242
= 0.365

2) 40 pt - For a 2-dimensional input space, we are given the following class conditional probability densities. Assume that x_1 and x_2 are conditionally independent given class names.

$p(x_1|C_1) = 1/3$ for $0 <= x_1 <= 4$
0 elsewhere

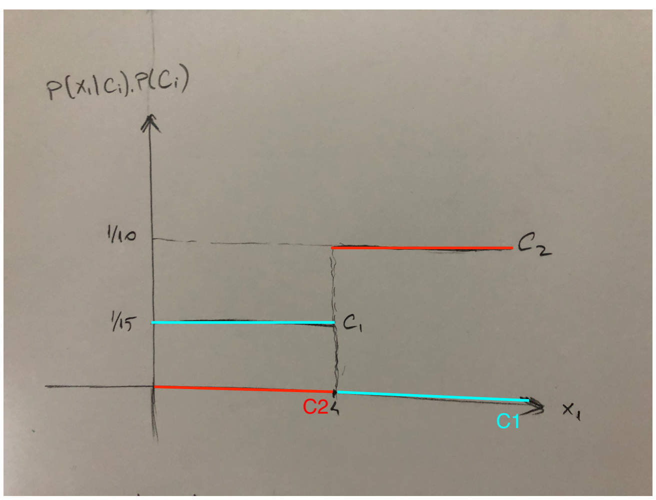
$p(x_2|C_1) = 1/4$ for $0 <= x_2 <= 4$
0 elsewhere

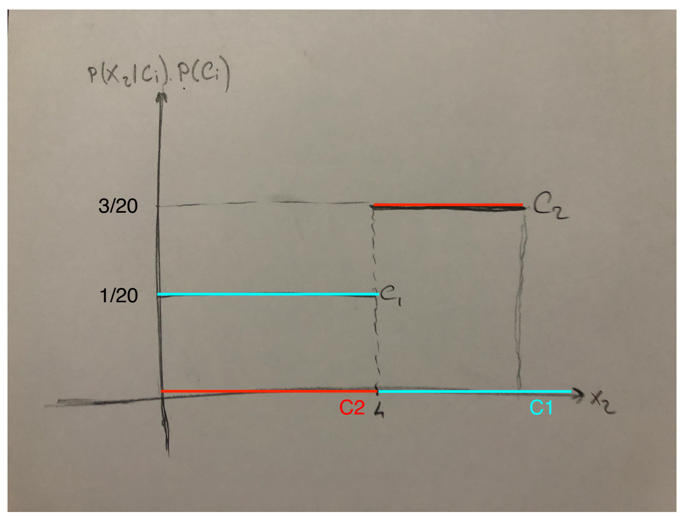
$p(x_1|C_2) = 1/2$ for $4 < x_1$
0 elsewhere

$p(x_2|C_2) = 3/4$ for $4 < x_2$
0 elsewhere

Assume $P(C_1)=P(C_2)=0.2$ and $P(C_3)=0.6$.

a) **12pt – Draw the corresponding pdfs for x_1 and x_2, being as precise as possible** (e.g. **label axes and important points on the axes**). You can draw by hand, take a picture and include here as image.





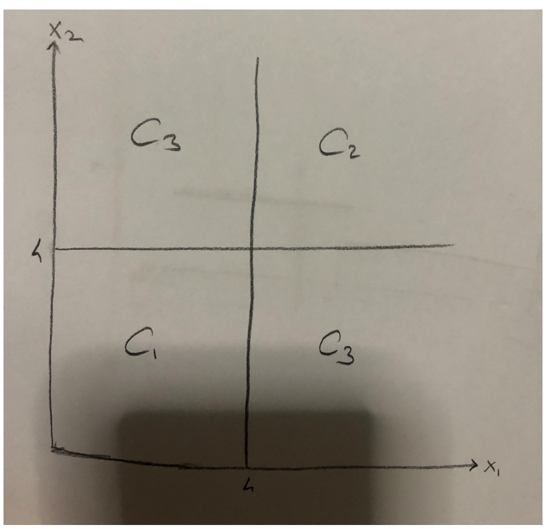
b) **10pts - Develop a classification strategy for given feature values** (just looking at the graph – no formula), just complete the sentence(s):

if x_1 is in the region 0 <= x_1 <= 4, and x_2 is in the region 0 <= x_2 <= 4, I will classify it as C_1;

if x_1 is in the region 4 < x_1, and x_2 is in the region 4 < x_2 , I will classify it as C_2;

Otherwise, I will classify it as C_3;

c) 8pts - Draw the decision regions.



d) 10pts – Give a one line qualitative answer (no precise numbers/thresholds...) **& reasoning for each case below (**how your decision changes or whether it doesn't).

- Would your decision strategy change if $P(C_1)=0.8$ and $P(C_2)=P(C_3)=0.1$?

No, there is no overlap between boundaries of different classes in pdfs. So, probabilities are not important in the decision of classes.

- How about if it was the reverse $P(C_1)=P(C_3)=0.1$ and $P(C_2)=0.8$?

No, there is no overlap between boundaries of different classes in pdfs. So, probabilities are not important in the decision of classes.

3) 40pts – NAIVE BAYES

a) 10pts – Given that two random variables X and Y are conditionally independent given C, circle True or False (2pts for each correct answer; -1pts each wrong answer):

- $P(X \mid Y) = P(X)$ — True / False
- $P(X \mid Y, C) = P(X \mid Y)$ — True / False
- $P(X, C \mid Y) = P(X \mid Y)$ — True / False
- $P(X, Y \mid C) = P(X \mid C)\, P(Y \mid C)$ — True / False
- $P(X, Y, C) = P(X \mid C)\, P(Y \mid C)\, P(C)$ — True / False

b) 24pts - Using the Mammal dataset given below, **how would you classify the animal that give birth, cannot fly, sometimes live in water, and has not legs, using Naive Bayes classifier *without any smoothing*. Show your work** (e.g. indicate class conditional attribute probabilities under the given table in the next page and just transfer them here).

P(mammal| x) = P(Give Birth=Yes | Mammal) * P(Can Fly=No | Mammal) * P(Live in Water=Sometimes | Mammal) * P(Have Legs=No | Mammal) * P(Mammal)
= 6/7 * 6/7 * 0 * 2/7 * 7/20
= 0

P(non-mammals| x) = P(Give Birth=Yes | non-Mammal) * P(Can Fly=No | non-Mammal) * P(Live in Water(sometimes) | non-Mammal) * P(Have Legs=No | non-Mammal) * P(non-Mammal)
= 1/13 * 10/13 * 4/13 * 4/13 * 13/20
= 0.0036

Decision: non-mammal

Name	Give Birth	Can Fly	Live in Water	Have Legs	Class
human	yes	no	no	yes	mammals
python	no	no	no	no	non-mammals
salmon	no	no	yes	no	non-mammals
whale	yes	no	yes	no	mammals
frog	no	no	sometimes	yes	non-mammals
komodo	no	no	no	yes	non-mammals
bat	yes	yes	no	yes	mammals
pigeon	no	yes	no	yes	non-mammals
cat	yes	no	no	yes	mammals
leopard shark	yes	no	yes	no	non-mammals
turtle	no	no	sometimes	yes	non-mammals
penguin	no	no	sometimes	yes	non-mammals
porcupine	yes	no	no	yes	mammals
eel	no	no	yes	no	non-mammals
salamander	no	no	sometimes	yes	non-mammals
gila monster	no	no	no	yes	non-mammals
platypus	no	no	no	yes	mammals
owl	no	yes	no	yes	non-mammals
dolphin	yes	no	yes	no	mammals
eagle	no	yes	no	yes	non-mammals

Write here the estimated probabilities (you should only write those related to the question for simplicity):

P(Give Birth=Yes | Mammal) = 6/7
P(Can Fly=No | Mammal) = 6/7
P(Live in Water=Sometimes | Mammal) = 0
P(Have Legs=No | Mammal) = 2/7
P(Mammal) = 7/20

P(Give Birth=Yes | non-Mammal) = 1/13
P(Can Fly=No | non-Mammal) = 10/13
P(Live in Water(sometimes) | non-Mammal) = 4/13
P(Have Legs=No | non-Mammal) = 4/13
P(non-Mammal) = 13/20

c) 6pts - Without re-doing the whole process, calculate the probabilities for P(Live in Water|mammals), P(Have Legs|mammals) and P(Give Birth|non-mammals) using Laplace smoothing:

P(Live in Water=Sometimes|mammals) = 1/10

P(Have Legs=No|mammals) = 3/9

P(Give Birth=Yes|non-mammals) = 2/15