CS412 Machine Learning
HW 3 – Probabilities – Bayesian Learning
100pts

- **Use this document to type in your answers after the quetions** (rather than writing on a separate sheet of paper), so as to keep questions, answers and and grades together to facilitate grading.
- **TYPE your answer or write <u>legibly</u> by hand** (pts will be taken off, if your writing is not clear).
- **SHOW all your work for partial/full credit.**
- **Allocated spaces should be enough for your answers** (unnecessarily long and irrelevant answers may loose points)

Student Name: Aleyna Kütük
ID: 23991

1) 20 pt - Suppose that we have 3 colored boxes r (red), b (blue) and g (green).
Box r contains 8 apples, 1 oranges and 1 limes;
Box b contains 5 apples, 5 oranges and 0 limes;
Box g contains 3 apples, 3 oranges and 4 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) = 0.2 and p(g) = 0.6 and a piece of fruit is selected from the **chosen** box randomly.

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

probability of selecting an lime after selecting box r: p(x=red,y=lime) = 0.2 * 1/10= 0.02

probability of selecting an lime after selecting box b: p(x=bluebox,y=lime) = 0.2 * 0/10= 0

probability of selecting an lime after selecting box g: p(x=greenbox,y=lime) = 0.6 * 4/10= 0.24

For all boxes (for all x values), I calculated selecting a lime and then sum over x:

$P(y = lime) = \sum_x p(x, y)$ =0.02 + 0 + 0.24 = 0.26

b) 10pt - If we **observe that the selected fruit is a lime**, what is the probability that it came from the green box?

From a), the prob of selecting an lime from any box: 0.02 + 0 + 0.24 = 0.26 and the prob of selecting

lime from greenbox is 0.24.

So, the conditional probability of selecting from greenbox when we know that we have a lime:

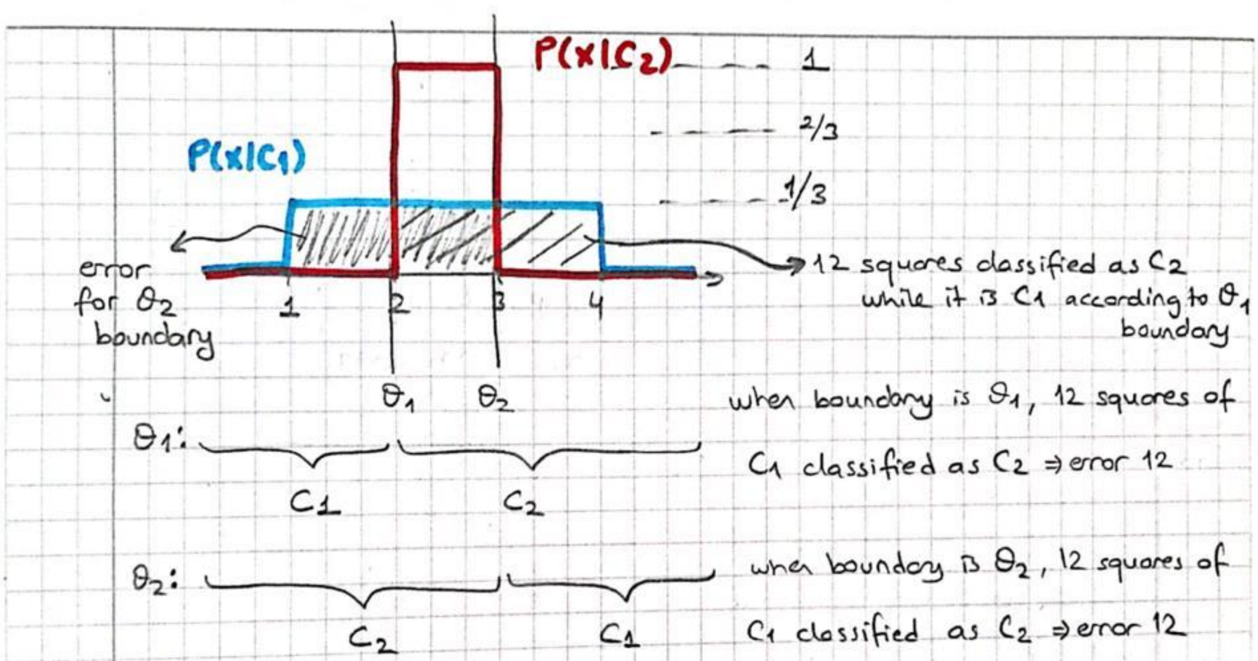
$P(\text{x=greenbox} \mid \text{y=lime}) = \frac{P(\text{x=greenbox,y=lime})}{P(y=lime)} = \frac{0.24}{0.26}$

2) 40 pt - For a **1-dimensional** input x, assume we are given the following class conditional probability densities as follows:

$p(x\|C_1) = 1/3$	for $1 <= x <= 4$	$p(x\|C_2) = 1$	for $2 <= x <= 3$
0	elsewhere	0	elsewhere

Assume $P(C_1)=P(C_2)=0.5$.

a) 15pt – Draw the corresponding $p(x|C_i)$, **being <u>as precise/specific as possible</u>** (e.g. label axes and important points on the axes). You can draw by hand, take a picture and include here as image.



θ1 and θ2 are two possible boundaries to split into two regions so that the total error can be minimum (12 squares are predicted incorrectly)

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

`if x is in the region x <= 2, I will classify it as C1;

if x is in the region 2 <= x, I will classify it as C2;

(boundary 1 which I draw on the picture at part a), and notice that boundary 2 is similar beacuse of the symmetry of graphs. Similarly for other possible boundary 2, if x is in the region x <= 3, I will classify it as C2; if x is in the region 3 <= x, I will classify it as C1. Both boundary 1 and 2 gives the same error.)

c) 5pts - Draw the decision regions on the above figure.

I draw them on the picture above see the part a). Both boundary 1 and 2 give the same result because the graphs are symmetric according to imaginary x=2.5 axis in the middle.

So, boundary 1 → if x <=2 classify as C1 and if x=>2 classify as C2
or boundary 2 → if x <=3 classify as C2 and if x=>3 classify as C1

d) 10pts – Give a **one line qualitative answer** (no precise numbers/threshold**s...)** about **how your decision changes or whether it doesn't.**

- Would your decision strategy change if $P(C_1)=0.9$ and $P(C_2)=0.1$?

Yes, decision strategy will change because we are always interested in posterior probabilities and we ignored the step of multiplying the P(x|Ci) with their related prior probabilities in the previous case when P(C1) = P(C2) = 0.5.
However, in this case when we multiply P(x|C1)* P(C1) and P(x|C2)* P(C2) to find Likelihood x Prior, we increase posterior of class C1 and class C1 becomes higher than class C2 on the graph due to P(C1) > P(C2).

P(x|C1)* P(C1) = 1/3 * 9/10 = 0.3 for 1 <= x <= 4
P(x|C2)* P(C2) = 1 * 1/10 = 0.1 for 2 <= x <= 3

Decision boundary is now at x=1 or x=4 axis, both are okay due to symmetry. For boundary at x=1, I would call if x <=1 classify as C2 and if x=>1 classify as C1.

- How about if it was the reverse $P(C_1)=0.1$ and $P(C_2)=0.9$?

No, decision strategy will not change because when we multiply Likelihoods with their related priors and it results that Likelihood x Prior of class C2 becomes much higher compared to class C1 on the graph. Then, decision boundary is still the same with P(C1)=P(C2)=0.5 case.

P(x|C1)* P(C1) = 1/3 * 1/10 = 0.033   for 1 <= x <= 4
P(x|C2)* P(C2) = 1 * 9/10 = 0.9   for 2 <= x <= 3

Decision boundary is still at x=2 or x=3 axis. For boundary at x=2, I would call if x <=2 classify as C1 and if x=>2 classify as C2.

3) 40pts – NAIVE BAYES

a) 15pts – 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| Y) = P(X)   True / False

(because Y is dependent with X when C is not given like in the above case)

- P(X | Y, C) = P(X | Y)   True / False

(because P(X | Y, C) = P(X|C) , not P(X|Y))

- P(X , C | Y) = P(X | Y)   True / False

(because P(X,C | Y) = P(C| X,Y) P(X|Y))

- P(X ,Y | C) = P(X | C) P(Y | C)   True / False

(because X and Y are independent so by using Naive Bayesian approach, we can consider them separately)

- P(X ,Y , C) = P(X | C) P(Y | C) P(C)   True / False

(because $\frac{P(X,Y,C)}{P(C)} = P(X,Y\,|C) = P(X|C)P(Y|C)$ so it is true)

b) 20pts - Using the PlayTennis data given below (and in the lecture slides), **how would you classify x=(Overcast, Mild, Normal, Strong), 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(Yes| x) = ?

P(X|Play=Yes)= P(O=Overcast|Play=Yes) * P(T= Mild|Play=Yes)* P(H=Normal|Play=Yes)* P(W=Strong|Play=Yes) = $\frac{4}{9} . \frac{4}{9} . \frac{6}{9} . \frac{3}{9}$ = 0.04389

P(Play=Yes) = $\frac{9}{14}$

P(Yes| x) = P(X|Play=Yes)* P(Play=Yes) = 0.04389* $\frac{9}{14}$ = 0.028218

P(No| x) = ?

P(X|Play=No)= P(O=Overcast|Play=No) * P(T= Mild|Play=No)* P(H=Normal|Play=No)* P(W=Strong|Play=No) = $\frac{0}{5} . \frac{2}{5} . \frac{1}{5} . \frac{3}{5}$ = 0

P(Play=No) = $\frac{5}{14}$

P(No| x) = P(X|Play=No)* P(Play=No) = 0* $\frac{5}{14}$ = 0

In conclusion, when we compare posterior probabilities, we select the higher posterior and class label predicted for x is Play=Yes

Decision: Play = Yes

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Write here the estimated probabilities (write those related to the question):

P(Outlook=Overcast | Yes) = $\frac{4}{9}$

P(Temperature= Mild| Yes) = $\frac{4}{9}$

P(Humidity= Normal| Yes) = $\frac{6}{9}$

P(Wind=Strong | Yes) = $\frac{3}{9}$

P(Play=Yes) = $\frac{9}{14}$

P(Outlook=Overcast | No) = $\frac{0}{5}$

P(Temperature= Mild| No) = $\frac{2}{5}$

P(Humidity= Normal| No) = $\frac{1}{5}$

P(Wind=Strong | No) = $\frac{3}{5}$

P(Play=No) = $\frac{5}{14}$

c) 5pts - Indicate the values **for only the following probabilities** estimated during Naive Bayes training, **using Laplace smoothing:**

P(Outlook=Overcast|No) = $\frac{0+1}{5+3} = \frac{1}{8}$ (3 possible Outlook values:Sunny, Rain and Overcast)

P(Humidity=Normal|Yes) = $\frac{6+1}{9+2} = \frac{7}{11}$ (2 possible Humidity values: Normal and High)