Assignment 4

a) The assignments have to be done individually.

b) The assignments have to be answered in English.

c) The answers have to be uploaded to Toledo trough "Assignments", not using any other Toledo tool and not by email.

d) Clearly indicate your name and student number in the uploaded answer.

e) Indicate the time that you have (approximately) spend on the assignment. This will not be taken into account in any way for the quotation but will give us an idea of the load of the assignments. We estimate a load of maximally two hours and half per assignment.

f) You will be rewarded for correct answers, not for the format of your answer.

g) The name of the uploaded documents should be:

- **rXXXXXXX.pdf** for the solution sheet
- **rXXXXXXX-calc.pdf** for the calculations

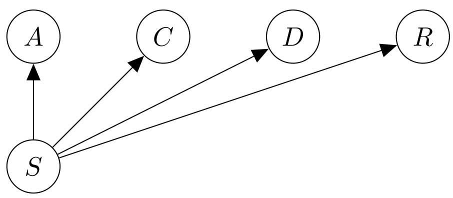
where you replace XXXXXXX with your student number.

h) **Use the template of the solution sheet to submit your solution.** In case no printer is available, feel free to mimic this template by hand.

1 Naive bayes and parameter estimation

We are interested in estimating the parameters of a simple Naive Bayes classifier for movies. The goal is predicting whether movies will be commercially successful or not (S), based on whether they contain the following elements (encoded as binary variables):

- Action A
- Comedy C
- Drama D
- Romance R



1. Given the following fully-observed dataset $\mathcal{D}_{full}$, compute the probability of success of movies with the following elements:

   (1,0,0,1)

   M_1 Contains both action and romance, but no comedy nor drama elements

   M_2 Has a bit of everything: action, comedy, drama, romance (1,1,1,1)

   after estimating the parameters using 2 different methods:

   (a) Maximum likelihood estimation

   (b) Maximum a-posteriori estimation with $Beta(\theta|\alpha = 1, \beta = 2)$ (α, β being the virtual counts for value 1 and 0 respectively) on all the parameters.

$\mathcal{D}_{full}$				
A	C	D	R	**S**
0	0	0	1	0
0	0	0	1	1
0	1	1	0	0
0	1	0	0	0
0	1	0	1	1
1	0	1	1	1
1	1	0	0	1
1	1	0	1	0

2. Given the following partially-observed dataset, use Expectation-Maximization over $\mathcal{D} = \mathcal{D}_{full} \cup \mathcal{D}_{partial}$ for estimating the parameters, writing the parameter estimates for the first 2 iterations. Assume an initial flat (uniform) distribution on each variable when computing the first E-step.

$\mathcal{D}_{partial}$				
A	C	D	R	**S**
0	0	0	?	0
0	1	?	0	0
1	?	1	1	1

Name: Mohammad Sadil Khan

Student ID: r0873969

Solution

Fill in each answer in the respective box. Intermediate computations must be submitted on a separate page.

1.a

$p(S = 1|M_1) \approx 0.86$

$p(S = 1|M_2) \approx 0.67$

1.b

$p(S = 1|M_1) \approx 0.789$

$p(S = 1|M_2) \approx 0.625$

2

E-step

A	C	D	R	S	weight at iteration 1	weight at iteration 2
0	0	0	**0**	0	1/2	7/12
0	0	0	**1**	0	1/2	5/12
0	1	**0**	0	0	1/2	3/4
0	1	**1**	0	0	1/2	1/4
1	**0**	1	1	1	1/2	1/2
1	**1**	1	1	1	1/2	1/2
0	0	0	1	0	1	1
0	0	0	1	1	1	1
0	1	1	0	0	1	1
0	1	0	0	0	1	1
0	1	0	1	1	1	1
1	0	1	1	1	1	1
1	1	0	0	1	1	1
1	1	0	1	0	1	1

M-step

parameters	parameters at iteration 1	parameters at iteration 2
$\theta_{A=1\|S=0}$	0.167	0.167
$\theta_{A=1\|S=1}$	0.6	0.6
$\theta_{C=1\|S=0}$	0.67	0.67
$\theta_{C=1\|S=1}$	0.5	0.5
$\theta_{D=1\|S=0}$	0.25	0.21
$\theta_{D=1\|S=1}$	0.4	0.4
$\theta_{R=1\|S=0}$	0.417	0.403
$\theta_{R=1\|S=1}$	0.8	0.8
$\theta_{S=1}$	5/11	5/11

Name: Mohammad Sadil Khan
Student ID: r0873969

1.a)
$M_1 = (1,0,0,1)$

$M_2 = (1,1,1,1)$

Given S, A, C, D, R are conditionally independent of each other.

Using ML, we have

$P(A=1|S=1) = \frac{2}{4}$

$P(C=1|S=1) = \frac{2}{4}$

$P(D=1|S=1) = 1/4$

$P(R=1|S=1) = 3/4$

$P(S=1) = \frac{1}{2}$

$P(A=1|S=0) = \frac{1}{4}$

$P(C=1|S=0) = \frac{3}{4}$

$P(D=1|S=0) = \frac{1}{4}$

$P(R=1|S=0) = \frac{1}{2}$

$P(S=0) = \frac{1}{2}$

For M_1

$P(M_1|S=1,D_{full}) = P(A=1, C=0, D=0, R=1|S=1, D_f)$

$= P(A=1|S=1,D_f)P(C=0|S=1,D_f)P(D=0|S=1,D_f)P(R=1|S=1,D_f)$

$= \frac{1}{2} \times \frac{1}{2} \times \frac{3}{4} \times \frac{3}{4}$

$= \frac{9}{64}$

$P(S=1|M_1,D_{full}) \quad \frac{P(M_1|S=1,D_{full})P(S=1|D_{full})}{P(M_1|D_{full})}$

$$= \frac{\frac{9}{64} \times \frac{1}{2}}{\frac{9}{64} \times \frac{1}{2} + \frac{1}{4} \times \frac{1}{4} \times \frac{3}{4} \times \frac{1}{2} \times \frac{1}{2}}$$

$$= \frac{\frac{9}{64}}{\frac{9}{64} + \frac{3}{128}} = \frac{18}{21} = \frac{6}{7}$$

$$\approx 0.86$$

For M_2

$M_2 = (1,1,1,1)$

$$P(s=1 \mid M_2, D_{full}) = \frac{P(M_2 \mid s=1, D_{full})\, P(s=1 \mid D_{full})}{P(M_2 \mid D_{full})}$$

$$= \frac{\frac{1}{2} \times \frac{1}{2} \times \frac{1}{4} \times \frac{3}{4} \times \frac{1}{2}}{\frac{1}{2} \times \frac{1}{2} \times \frac{1}{4} \times \frac{3}{4} \times \frac{1}{2} + \frac{1}{4} \times \frac{3}{4} \times \frac{1}{4} \times \frac{1}{2} \times \frac{1}{2}}$$

$$= \frac{\frac{3}{64}}{\frac{3}{64} + \frac{3}{128}} = \frac{6}{9} = \frac{2}{3}$$

$$\approx 0.67$$

b) Using MAP estimation

Let $\theta_{AS_i} = P(A=1 \mid S=i, \theta_A)$ $\therefore$ $\theta_{AS_i} \sim \beta(\theta_{AS_i} \mid 1,2)$

$\theta_{CS_i} = P(C=1 \mid S=i, \theta_B)$ $\therefore$ $\theta_{CS_i} \sim \beta(\theta_{CS_i} \mid 1,2)$

$\theta_{DS_i} = P(D=1 \mid S=i, \theta_C)$ $\therefore$ $\theta_{DS_i} \sim \beta(\theta_{DS_i} \mid 1,2)$

$\theta_{RS_i} = P(R=1 \mid S=i, \theta_D)$ $\therefore$ $\theta_{RS_i} \sim \beta(\theta_{RS_i} \mid 1,2)$

$\theta_{S_i} = P(S=i \mid \theta_S)$ $\therefore$ $\theta_{S_i} \sim \beta(\theta_{S_i} \mid 1,2)$

$\forall i = \{0,1\}$

If $\theta \sim \beta(\theta \mid \alpha, \beta)$, MAP estimate of θ is $\frac{\alpha-1}{\alpha+\beta-2}$

$$P(\theta_{S_1} \mid D_f) = \beta(\theta_S \mid 5, 6)$$

$$P(S=1 \mid D_f) = \frac{5}{11}$$

$$P(S=0 \mid D_f) = \frac{6}{11}$$

$$P(\theta_{AS_1} \mid D_f) = \beta(\theta_{AS_1} \mid 1 + \#(A=1, S=1),\ 2 + \#(A=0, S=1))$$

$$= \beta(\theta_{AS_1} \mid 3, 4)$$

$$P(A=1 \mid S=1, D_f) = \frac{2}{5}$$

$$P(\theta_{AS_0} \mid D_f) = \beta(\theta_{AS_0} \mid 2, 5)$$

$$P(A=1 \mid S=0, D_f) = \frac{1}{5}$$

$$P(\theta_{CS_1} \mid D_f) = \beta(\theta_{CS_1} \mid 3, 4)$$

$$P(C=1 \mid S=1, D_f) = \frac{2}{5}$$

$$P(\theta_{CS_0} \mid D_f) = \beta(CS_0 \mid 4, 3)$$

$$P(C=1 \mid S=0, D_f) = \frac{3}{5}$$

$$P(\theta_{DS_1} \mid D_f) = \beta(\theta_{DS_1} \mid 2, 5)$$

$$P(D=1 \mid S=1, D_f) = \frac{1}{5}$$

$$P(\theta_{DS_0} \mid D_f) = \beta(\theta_{DS_0} \mid 2, 5)$$

$$P(D=1 \mid S=0, D_f) = \frac{1}{5}$$

$$P(RS_1 \mid D_f) = \beta(RS_1 \mid 4, 3)$$

$$P(R=1 \mid S=1, D_f) = \frac{3}{5}$$

$$P(RS_0 \mid D_f) = \beta(RS_0 \mid 3, 4)$$

$$P(R=1 \mid S=0, D_f) = \frac{2}{5}$$

For M_1

$$P(S=1 \mid M_1, D_f)$$

$$= \frac{P(M_1 \mid S=1, D_f)\, P(S=1 \mid D_f)}{P(M_1)}$$

$$= \frac{P(1,0,0,1 \mid S=1, D_f)\, P(S=1 \mid D_f)}{P(1,0,0,1 \mid S=1, D_f)\, P(S=1 \mid D_f) + P(1,0,0,1 \mid S=0, D_f)\, P(S=0 \mid D_f)}$$

$$= \frac{\frac{2}{5} \times \frac{3}{5} \times \frac{4}{5} \times \frac{3}{5} \times \frac{5}{11}}{\frac{2}{5} \times \frac{3}{5} \times \frac{4}{5} \times \frac{3}{5} \times \frac{5}{11} + \frac{1}{5} \times \frac{2}{5} \times \frac{4}{5} \times \frac{2}{5} \times \frac{6}{11}}$$

$$= \frac{2 \times 3^2 \times 4 \times 5}{2 \times 3^2 \times 4 \times 5 + 2^2 \times 4 \times 6}$$

$$= \frac{2^3 \times 3^2 \times 5}{2^3 \times 3^2 \times 5 + 2^3 \times 2 \times 3 \times 2}$$

$$= \frac{15}{15 + 4}$$

$$= \frac{15}{19} = 0.789$$

$$P(S=1 \mid M_2, D_f)$$

$$= \frac{P(M_2 \mid S=1, D_f)\; P(S=1 \mid D_f)}{P(M_2)}$$

$$= \frac{\frac{5}{11} \times \frac{2\times 2\times 1\times 3}{5^4}}{\frac{5}{11} \times \frac{2\times 2\times 1\times 3}{5^4} + \frac{6}{11} \times \frac{1\times 3\times 1\times 2}{5^4}}$$

$$= \frac{5\times 2^2\times 3}{5\times 2^2\times 3 + 6\times 3\times 2}$$

$$= \frac{5\times 2}{5\times 2 + 6} = \frac{10}{16} = \frac{5}{8} \simeq 0.625$$

27 : If, $\theta \sim U(0,1)$

$$\theta_{mean} = \frac{1}{2}$$

Let the parameters be

$$P(A=1|S=0) = \theta_{A=1|S=0} = 1/2$$

$$P(A=1|S=0) = \theta_{A=1|S=1} = 1/2$$

$$P(C=1|S=0) = \theta_{C=1|S=0} = 1/2$$

$$P(C=1|S=1) = \theta_{C=1|S=1} = 1/2$$

$$P(D=1|S=0) = \theta_{D=1|S=0} = 1/2$$

$$P(D=1|S=1) = \theta_{D=1|S=1} = 1/2$$

$$P(R=1|S=0) = \theta_{R=1|S=0} = 1/2$$

$$P(R=1|S=1) = \theta_{R=1|S=1} = 1/2$$

$$P(S=1) = \theta_{S=1} = 1/2$$

First Iteration

E-step

$$P(R=0|S=0) = 1 - \theta_{R=1|S=0} = \frac{1}{2}$$

$$P(R=1|S=0) = \theta_{R=1|S=0} = \frac{1}{2}$$

$$P(D=0|S=0) = 1 - \theta_{D=1|S=0} = \frac{1}{2}$$

$$P(D=1|S=0) = \theta_{D=1|S=0} = \frac{1}{2}$$

$$P(C=0|S=1) = 1 - \theta_{C=1|S=1} = \frac{1}{2}$$

$$P(C=1|S=1) = \theta_{C=1|S=1}$$

M-step

$$\theta_{A=1|S=0} = \frac{1}{6}$$

$$\theta_{A=1|S=1} = \frac{1+2}{1+4} = \frac{3}{5} = 0.6$$

$$\theta_{C=1|S=0} = \frac{1+3}{6} = \frac{4}{6} = \frac{2}{3}$$

$$\theta_{C=1|S=1} = \frac{2+\frac{1}{2}}{5} = \frac{1}{2} = 0.5$$

$$\theta_{D=1|S=0} = \frac{1.5}{6} = \frac{1}{4} = 0.25$$

$$\theta_{D=1|S=1} = \frac{2}{5} = 0.4$$

$$\theta_{R=1|S=0} = \frac{2.5}{6} = \frac{25}{60} = \frac{5}{12}$$

$$\theta_{R=1|S=1} = \frac{4}{5} = 0.8$$

$$\theta_{S=1} = \frac{5}{11}$$

Second Iteration

E-step

$$P(R=0|S=0) = 1 - \theta_{R=1|S=0} = \frac{7}{12}$$

$$P(R=1|S=0) = \theta_{R=1|S=0} = \frac{5}{12}$$

$$P(D=0|S=0) = 1 - \theta_{D=1|S=0} = 0.75$$

$$P(D=1|S=0) = \theta_{D=1|S=0} = 0.25$$

$$P(C=0|S=1) = 1 - \theta_{C=1|S=1} = \frac{1}{2}$$

$$P(C=1|S=1) = \theta_{C=1|S=1} = \frac{1}{2}$$

M-Step

$$\theta_{A=1|S=0} = \frac{1}{6}$$

$$\theta_{A=1|S=1} = \frac{3}{5}$$

$$\theta_{C=1|S=0} = \frac{4}{6} = \frac{2}{3} = 0.67$$

$$\theta_{C=1|S=1} = \frac{2.5}{5} = 0.5$$

$$\theta_{D=1|S=0} = \frac{1.25}{6} = 0.21$$

$$\theta_{D=1|S=1} = \frac{2}{5} = 0.4$$

$$\theta_{R=1|S=0} = \frac{2+\frac{5}{12}}{6} = 0.403$$

$$\theta_{R=1|S=1} = \frac{4}{5} = 0.8$$

$$\theta_{S=1} = \frac{5}{11}$$