Instructions

- You have 90 minutes to complete the test.
- Make sure that your test has a total of 7 pages and is not missing any sheets, then write your full name and student n. on this page (and all others if you want to be safe).
- The test has a total of 5 questions, with a maximum score of 20 points. The questions have different levels of difficulty. The point value of each question is provided next to the question number.
- *If you get stuck in a question, move on.* You should start with the easier questions to secure those points, before moving on to the harder questions.
- *No interaction with the faculty is allowed during the test.* If you are unclear about a question, clearly indicate it and answer to the best of your ability.
- Please provide your answer in the space below each question. If you make a mess, clearly indicate your answer.
- The test is open book and open notes. You may use a calculator, but any other type of electronic or communication equipment is not allowed.
- Good luck.

Question 1. (**7 pts.**)

Consider the dataset $\mathcal{D}$ represented in Table 1, which is a subset of the IRIS dataset that you used in the lab. Each datapoint $p \in \mathcal{D}$ corresponds to a type of iris flower, as indicated in Table 1.

Table 1: Dataset and corresponding class information.

Petal length (cm)	4.7	4.9	5.0	5.3	5.5	5.9
Petal width (cm)	1.6	1.5	1.7	1.9	1.8	2.1
Type	Versicolor	Versicolor	Versicolor	Setosa	Setosa	Setosa

a) **(3 pts.)** Suppose that each data point $p \in \mathcal{D}$ is described by the following binary attributes:

- For each datapoint p, $\phi_1(p)$ is 1, if the petal length of p is larger than or equal to 5cm, and 0 otherwise.
- For each datapoint p, $\phi_2(p)$ is 1, if the petal width of p is larger than 1.8cm, and 0 otherwise.

Compute all parameters of the Naïve Bayes classifier for the dataset above with attributes ϕ_1 and ϕ_2.

b) **(3 pts.)** Suppose now that each data point $p \in \mathcal{D}$ is described by the two attributes in Table 1. Compute all parameters of the Naïve Bayes classifier.

c) **(1 pts.)** Determine the accuracy of the classifier from a) using the following test set: **Note:**

Table 2: Test set.

Petal length (cm)	4.5	5.0	5.1
Petal width (cm)	1.7	1.5	1.6
Type	Setosa	Setosa	Versicolor

If a datapoint falls in the decision boundary, consider it as being classified "half-correctly".

Solution 1.

a) We first rewrite the datapoints in Table 1 using the provided features, yielding the following table:

ϕ_1	0	0	1	1	1	1
ϕ_2	0	0	0	1	0	1
Label	V	V	V	S	S	S

We can now compute the Naive Bayes parameters. Starting with the prior probabilities,

$$\mathbb{P}\left[\pi(x) = V\right] = \frac{N_V}{N} = \frac{3}{6} = 0.5 \qquad \mathbb{P}\left[\pi(x) = S\right] = \frac{N_S}{N} = 0.5.$$

As for the feature likelihoods we get, for feature ϕ_1,

$$\mathbb{P}\left[\phi_1(x) = 0 \mid \pi(x) = V\right] = \frac{N_{\neg\phi_1}}{N_V} = \frac{2}{3} \qquad \mathbb{P}\left[\phi_1(x) = 0 \mid \pi(x) = S\right] = \frac{N_{\neg\phi_1}}{N_S} = \frac{0}{3} = 0$$
$$\mathbb{P}\left[\phi_1(x) = 1 \mid \pi(x) = V\right] = \frac{N_{\phi_1}}{N_V} = \frac{1}{3} \qquad \mathbb{P}\left[\phi_1(x) = 1 \mid \pi(x) = S\right] = \frac{N_{\phi_1}}{N_S} = \frac{3}{3} = 1$$

and for feature ϕ_2,

$$\mathbb{P}\left[\phi_2(x) = 0 \mid \pi(x) = V\right] = \frac{N_{\neg\phi_2}}{N_V} = \frac{3}{3} = 1 \qquad \mathbb{P}\left[\phi_2(x) = 0 \mid \pi(x) = S\right] = \frac{N_{\neg\phi_2}}{N_S} = \frac{1}{3}$$
$$\mathbb{P}\left[\phi_2(x) = 1 \mid \pi(x) = V\right] = \frac{N_{\phi_2}}{N_V} = \frac{0}{3} = 0 \qquad \mathbb{P}\left[\phi_2(x) = 1 \mid \pi(x) = S\right] = \frac{N_{\phi_2}}{N_S} = \frac{2}{3}.$$

b) The priors are, once again,

$$\mathbb{P}\left[\pi(x) = V\right] = \frac{N_V}{N} = \frac{3}{6} = 0.5 \qquad \mathbb{P}\left[\pi(x) = S\right] = \frac{N_S}{N} = 0.5.$$

Since the points in the training set are originally represented as continuous-valued attributes, we now represent the feature likelihoods as (univariate) normal distributions. Since a normal distribution is fully specified by its mean and variance, we get, for the length attribute,

$$L \mid V \sim \mathsf{Normal}(\mu_{LV}, \sigma_{LV})$$
$$L \mid S \sim \mathsf{Normal}(\mu_{LS}, \sigma_{LS}),$$

where

$$\mu_{LV} = \frac{4.7 + 4.9 + 5.0}{3} = 4.87$$
$$\mu_{LS} = \frac{5.3 + 5.5 + 5.9}{3} = 5.57$$
$$\sigma^2_{LV} = \frac{(4.7 - 4.87)^2 + (4.9 - 4.87)^2 + (5.0 - 4.87)^2}{2} = 0.02$$
$$\sigma^2_{LS} = \frac{(5.3 - 5.57)^2 + (5.5 - 5.57)^2 + (5.9 - 5.57)^2}{2} = 0.09$$

As for the length attribute, we have

$$W \mid V \sim \mathsf{Normal}(\mu_{WV}, \sigma_{WV})$$
$$W \mid S \sim \mathsf{Normal}(\mu_{WS}, \sigma_{WS}),$$

where

$$\mu_{WV} = \frac{1.6 + 1.5 + 1.7}{3} = 1.6$$
$$\mu_{WS} = \frac{1.9 + 1.8 + 2.1}{3} = 1.93$$
$$\sigma^2_{WV} = \frac{(1.6 - 1.6)^2 + (1.5 - 1.6)^2 + (1.7 - 1.6)^2}{2} = 0.01$$
$$\sigma^2_{WS} = \frac{(1.9 - 1.93)^2 + (1.8 - 1.93)^2 + (2.1 - 1.93)^2}{2} = 0.02.$$

c) To determine the accuracy of the classifier, we classify each of the three points provided and compare the obtained classifications with those provided. We start by describing the three datapoints, henceforth denoted as x_1, x_2 and x_3, in terms of features ϕ_1 and ϕ_2, to yield

	x_1	x_2	x_3
ϕ_1	0	1	1
ϕ_2	0	0	0

From the classifier in a) we have, for x_1,

$$\begin{aligned}
\mathbb{P}\left[\pi(x_1) = V \mid \phi_1(x) = 0, \phi_2(x) = 0\right] &\propto \mathbb{P}\left[\pi(x_1) = V\right] \mathbb{P}\left[\phi_1(x) = 0 \mid \pi(x) = V\right] \mathbb{P}\left[\phi_2(x) = 0 \mid \pi(x) = V\right] \\
&= \frac{1}{2} \times \frac{2}{3} \times 1 = \frac{1}{3} \\
\mathbb{P}\left[\pi(x_1) = S \mid \phi_1(x) = 0, \phi_2(x) = 0\right] &\propto \mathbb{P}\left[\pi(x_1) = S\right] \mathbb{P}\left[\phi_1(x) = 0 \mid \pi(x) = S\right] \mathbb{P}\left[\phi_2(x) = 0 \mid \pi(x) = S\right] \\
&= \frac{1}{2} \times 1 \times \frac{1}{3} = \frac{1}{6},
\end{aligned}$$

and x_1 is classified as Versicolor. Repeating the process to the other two points, we get

$$\begin{aligned}
\mathbb{P}\left[\pi(x_2) = V \mid \phi_1(x) = 0, \phi_2(x) = 0\right] &\propto \frac{1}{2} \times \frac{1}{3} \times 1 = \frac{1}{6} \\
\mathbb{P}\left[\pi(x_2) = S \mid \phi_1(x) = 0, \phi_2(x) = 0\right] &\propto \frac{1}{2} \times 1 \times \frac{1}{3} = \frac{1}{6} \\
\mathbb{P}\left[\pi(x_3) = V \mid \phi_1(x) = 0, \phi_2(x) = 0\right] &\propto \frac{1}{2} \times \frac{1}{3} \times 1 = \frac{1}{6} \\
\mathbb{P}\left[\pi(x_3) = S \mid \phi_1(x) = 0, \phi_2(x) = 0\right] &\propto \frac{1}{2} \times 1 \times \frac{1}{3} = \frac{1}{6}
\end{aligned}$$

and the two points fall in the decision boundary and can be classified into any of the two classes. Considering that these points can fall into either class with probability 0.5, the accuracy is given by

$$\text{Acc} = \frac{\text{\# Correct labels}}{\text{\# Points}} = \frac{0.5 + 0.5}{3} = 0.33\%.$$

Question 2. **(2 pts.)**

Both Naïve Bayes and logistic regression are probabilistic classifiers. However, the two approaches are quite different in the type of classifier that they determine. Discuss the main differences between the two approaches with respect to the type of classifier determined by each approach.

Solution 2.

The key difference between the two classifiers is that LR builds a discriminative model (i.e., determines the probabilities $\pi(a \mid x)$) while NB builds a generative model (i.e., determines the joint probabilities $\mathbb{P}\left[\mathrm{x} = x, \mathrm{a} = a\right]$).

Question 3. **(8 pts.)**

Consider a two-state, two-action MDP $(\mathcal{X}, \mathcal{A}, \{\mathbf{P}_a\}, c, \gamma)$, where $\mathcal{X} = \{A, B\}$, $\mathcal{A} = \{a, b\}$ and $\gamma = 0.95$. Suppose that an agent observed the two following trajectories:

- $\{(A, a, 1, A), (A, b, 1, B), (B, a, 0, A)\}$
- $\{(B, b, 0, B)\}$

a) **(3 pts.)** Suppose that the agent uses model-based RL to compute Q^*. What is the resulting Q-function? Indicate all intermediate computations.

Note: Initialize the transition probability matrices to the identity matrix, and the cost matrix to $\mathbf{0}$. Also, initialize all visit counters to 1.

b) **(3 pts.)** Repeat the previous question but now using Q-learning. Namely, compute the resulting Q-function and indicate all intermediate computations.

Note: Initialize the Q-function to $\mathbf{0}$ and use $\alpha = 0.2$.

c) **(1 pts.)** For each of the Q-functions computed in a) and b) indicate a corresponding greedy policy.

d) **(1 pts.)** Knowing that the optimal Q-function for the MDP is

$$\mathbf{Q}^* = \begin{bmatrix} 7.1 & 6.7 \\ 6.1 & 5.7 \end{bmatrix},$$

discuss the error in Q in each of the two approaches after 4 iterations, both in terms of the resulting greedy policy and of the Q-values.

Solution 3.

a) We depart from the matrices:

$$\mathbf{P}_a^{(0)} = \mathbf{P}_b^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \boldsymbol{Q}^{(0)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \boldsymbol{C}^{(0)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \boldsymbol{N}^{(0)} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Using the transitions provided, we get:

Step 1:

$$\mathbf{P}_a^{(1)}(A \mid A) = 1 + \frac{1}{2}(1-1) = 1 \qquad \mathbf{P}_a^{(1)}(B \mid A) = 0 + \frac{1}{2}(0-0) = 0$$

$$\boldsymbol{C}^{(1)}(A, a) = 0 + \frac{1}{2}(1-0) = \frac{1}{2} \qquad \boldsymbol{Q}^{(1)}(A, a) = \frac{1}{2} + 0.95 \times \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \frac{1}{2}.$$

Step 2:

$$\mathbf{P}_b^{(2)}(A \mid A) = 1 + \frac{1}{2}(0-1) = \frac{1}{2} \qquad \mathbf{P}_b^{(2)}(B \mid A) = 0 + \frac{1}{2}(1-0) = \frac{1}{2}$$

$$\boldsymbol{C}^{(2)}(A, b) = 0 + \frac{1}{2}(1-0) = \frac{1}{2} \qquad \boldsymbol{Q}^{(2)}(A, b) = \frac{1}{2} + 0.95 \times \begin{bmatrix} 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \frac{1}{2}.$$

Step 3:

$$\mathbf{P}_a^{(3)}(A \mid B) = 0 + \frac{1}{2}(1-0) = \frac{1}{2} \qquad \mathbf{P}_a^{(3)}(B \mid B) = 1 + \frac{1}{2}(0-1) = \frac{1}{2}$$

$$\boldsymbol{C}^{(3)}(B, a) = 0 + \frac{1}{2}(0-0) = 0 \qquad \boldsymbol{Q}^{(3)}(B, a) = 0 + 0.95 \times \begin{bmatrix} 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} = 0.24.$$

Step 4:

$$\mathbf{P}_b^{(4)}(A \mid B) = 0 + \frac{1}{2}(0 - 0) = 0 \qquad \mathbf{P}_b^{(4)}(B \mid B) = 1 + \frac{1}{2}(1 - 1) = 1$$

$$\boldsymbol{C}^{(4)}(B, b) = 0 + \frac{1}{2}(0 - 0) = 0 \qquad \boldsymbol{Q}^{(4)}(B, b) = 0 + 0.95 \times \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0 \end{bmatrix} = 0.$$

The resulting Q-function is, thus,

$$\mathbf{Q}^{(4)} = \begin{bmatrix} 0.5 & 0.5 \\ 0.24 & 0 \end{bmatrix}.$$

b) With Q-learning, we get:

Step 1:

$$Q^{(1)}(A, a) = 0 + 0.2 \times (1 + 0.95 \times 0 - 0) = 0.2$$

Step 2:

$$Q^{(2)}(A, b) = 0 + 0.2 \times (1 + 0.95 \times 0 - 0) = 0.2$$

Step 3:

$$Q^{(3)}(B, a) = 0 + 0.2 \times (0 + 0.95 \times 0.2 - 0) = 0.04$$

Step 4:

$$Q^{(4)}(B, b) = 0 + 0.2 \times (0 + 0.95 \times 0 - 0) = 0.$$

The resulting Q-function is, thus,

$$\mathbf{Q}^{*} = \begin{bmatrix} 0.2 & 0.2 \\ 0.04 & 0 \end{bmatrix}.$$

c) The set of greedy policies is the same in both cases, the same. One possible greedy policy is, for example,

$$\pi_g = \begin{bmatrix} 0.5 & 0.5 \\ 0.0 & 1.0 \end{bmatrix}.$$

d) In terms of greedy policy, both methods yield the same result after 4 iterations. The greedy policy selects the correct action in state B but not in state A. As for the error in terms of absolute value, the Q-values for the model-based approach are closer to the optimal ones, thus yielding a smaller error in terms of absolute value.

Question 4. (2 pts.)

Consider an agent that must, at each time-step t, select one of three possible actions: a, b and c. After selecting the action, the agent pays a cost c_t that depends on the action selected at time-step t. Suppose that the agent selected the actions:

$$a \qquad b \qquad c \qquad c \qquad a \qquad b \qquad c \qquad b$$

and incurred the costs

$$0.5 \qquad 0.5 \qquad 1.0 \qquad 0.2 \qquad 0.3 \qquad 0.5 \qquad 0.2 \qquad 0.5.$$

Indicate the action selected by the agent at the next time-step if the agent follows the UCB algorithm. Include all relevant computations.

Note: Keep in mind that the log used in UCB is the *natural logarithm.*

Solution 4.

From the data provided, we have that:

$$\hat{c}(a) = 0.4 \qquad N_a = 2$$
$$\hat{c}(b) = 0.5 \qquad N_b = 3$$
$$\hat{c}(c) = 0.47 \qquad N_c = 3.$$

Using the UCB algorithm, we have that:

$$Q(a) = \hat{c}(a) - \sqrt{\frac{2\log t}{N_a}} = 0.4 - \sqrt{\log 8} = 0.4 - 1.442 = -1.042$$
$$Q(b) = \hat{c}(b) - \sqrt{\frac{2\log t}{N_b}} = 0.5 - \sqrt{\frac{2}{3}\log 8} = 0.5 - 1.177 = -0.677$$
$$Q(c) = \hat{c}(c) - \sqrt{\frac{2\log t}{N_c}} = 0.47 - \sqrt{\frac{2}{3}\log 8} = 0.47 - 1.177 = -0.711,$$

and UCB will select action a.

Question 5. (1 pts.)

Briefly discuss the main distinction between the EWA algorithm and the EXP3 algorithm, both in terms of the algorithm, performance guarantees and class of problems they are designed to address.

Solution 5.

EXP3 is an adaptation of EWA to multiarmed bandit problems. In other words, while EWA observes, at each step, the cost incurred by all actions, EXP3 observes only the cost associated with the action it selected. Therefore, at each step, EXP3 only updates the weight associated with the action selected, while EWA updates all weights. The regret of the two algorithms, however, is not too different, differing only by a factor of $\sqrt{|\mathcal{A}|}$.