DIPARTIMENTO DI ELETTRONICA, INFORMAZIONE E BIOINGEGNERIA

Politecnico di Milano

Machine Learning (Code: 097683)
January 22, 2018

Name:	
Surname:	
Student ID:	
Row:	Column:

Time: 2 hours 30 minutes | Prof. Marcello Restelli | Maximum Marks: 34

- The following exam is composed of **10 exercises** (one per page). The first page needs to be filled with your **name, surname and student ID**. The following pages should be used **only in the large squares** present on each page. Any solution provided outside these spaces will not be considered for the final mark.
- During this exam you are **not allowed to use electronic devices** like laptops, smartphones, tablets and/or similar. As well, you are not allowed to bring with you any kind of note, book, written scheme and/or similar. You are also not allowed to communicate with other students during the exam.
- The first reported violation of the above mentioned rules will be annotated on the exam and will be considered for the final mark decision. The second reported violation of the above mentioned rules will imply the immediate expulsion of the student from the exam room and the **annulment of the exam**.
- You are allowed to write the exam either with a pen (black or blue) or a pencil. It is your responsibility to provide a readable solution. We will not be held accountable for accidental partial or total cancellation of the exam.
- The exam can be written either in **English** or **Italian**.
- You are allowed to withdraw from the exam at any time without any penalty. You are allowed to leave the room not early than half the time of the duration of the exam. You are not allowed to keep the text of the exam with you while leaving the room.
- **Three of the points will be given on the basis on how quick you are in solving the exam. If you finish earlier than** 45 **min before the end of the exam you will get** 3 **points, if you finish earlier than** 30 **min you will get** 2 **points and if you finish earlier than** 15 **min you will get** 1 **point (the points cannot be accumulated).**

Ex. 1	Ex. 2	Ex. 3	Ex. 4	Ex. 5	Ex. 6	Ex. 7	Ex. 8	Ex. 9	Ex. 10	Time	Tot.
/ 5	/ 5	/ 5	/ 2	/ 2	/ 2	/ 2	/ 2	/ 3	/ 3	/ 3	/ 34

Exercise 1 (5 marks)

Describe the supervised learning technique denominated **Support Vector Machines** for classification problems.

Exercise 2 (5 marks)

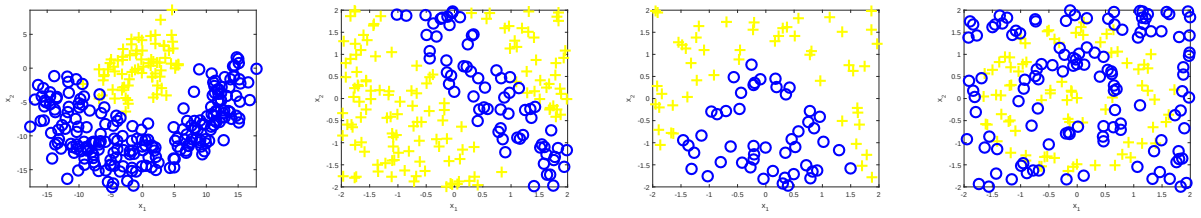
Describe the supervised learning technique denominated **logistic regression** for classification problems.

Exercise 3 (5 marks)

Describe the **policy iteration** technique for control problems on Markov Decision Processes.

Exercise 4 (2 marks)

For which ones of the following datasets you would use the kernel trick to represent your data? Would you use some other methodology? Provide motivation for your choice.



(a) (b) (c) (d)

Figure 1: Different datasets.

a. One might use polar coordinates to linearly separate the two classes.

b. There is no clear single separating surface between the two classes (not even a non-linear one), therefore the use of kernels might be a good idea.

c. A quadratic surface might be sufficient to separate the two classes.

d. The data presents a regularity, but even with a coordinate change we would not be able to separate the two classes. A kernel could do the trick.

Exercise 5 (2 marks)

Tell whether the following statements about the perceptron algorithm for classification are true or false.

1. Shuffling the initial data is not fundamental for the perceptron optimization procedure;
2. There exists a unique solution to the minimization of the perceptron loss when the data are linearly separable;
3. The choice of the learning rate α does not influence the learning procedure (asymptotically);
4. We are guaranteed that the overall loss of the method is strictly decreasing over time.

Motivate your answer.

1. FALSE: the online procedure to train the perceptron is dependent on the order of the data and on the initialization we use.
2. FALSE: if the data are linearly separable we are assured that the method converges to a solution, but any hyperplane separating the two classes would be a feasible solution.
3. TRUE: even if the choice of α influences the initial part of the learning procedure, its influence decreases as the number of iteration increases.
4. FALSE: we are only guaranteed that the loss over the processes datum does not increase.

Exercise 6 (2 marks)

Tell whether the following statements are true or false and motivate your answers.

1. Applying MC estimation on a single episode, you extract a number of samples for the value function equal to the length of the episode;
2. Applying MC estimation on a single episode, you extract a number of samples for the value function less or equal to the number of states of the MDP;
3. TD cannot be used in the case we are analysing an MDP with no terminal state;
4. MC every visit is a consistent, but biased, estimator for the state value function of the MDP.

1. FALSE: if we are using MC first visit since we are not counting twice the same state, TRUE: if we are using MC every visit, which computes a sample for the value function for each state visited during the episode.
2. TRUE: for MC first visit, FALSE: for MC every visit, for the same reasons of question 1.
3. FALSE: the TD update requires only to have a new visited state and, differently from MC does not require to have the complete episode to be applied.
4. TRUE: it is biased since it is considering the same visited state multiple times, but if the number of samples is large enough it converges to the unbiased estimates (consistent estimator), e.g., provided by MC first visit.

Exercise 7 (2 marks)

Tell whether the following problems can be modeled as a MAB or RL problem and explain why. Specify the MDP you would use in the different settings.

1. Play optimally a poker game;
2. Escape from a maze;
3. Place ads in a search engine page;
4. Select the most effective drug during a clinical test.

Since MAB is a specific case of RL, every time we can use a MAB approach we can also use a RL approach too.

1. RL: the MDP is partially observable, the actions are the poker actions (e,g,. Fold, Check, Call, Raise), the states are represented by each card in the players hands and on the table, the reward is the amount of money you make during a hand.
2. RL: the MDP is partially observable, the actions are the directions one might take in each state, the state is your position in the maze, the transition are deterministic, the reward is one only if you are able to exit the maze.
3. MAB: the actions are the different ads to be displayed and the reward is stochastic (one if clicked, zero if not).
4. MAB: the actions are the different drugs and the reward is the patient status (one if the patient is feeling better, zero if not).

Exercise 8 (2 marks)

Consider the modelization of a classification problem as a sequential decision making problem. In the specific, assume that the observation of the state is the input vector, the action is the prediction provided by an SVM and the reward is zero if we mis-classify the input and 1 if we correctly classify it. Formally:

$$o_i \leftarrow x_i$$
$$a_i \leftarrow \hat{y}_i$$
$$r_i \leftarrow 1 - |t_i - \hat{y}_i|$$

Does the correspondence between the aforementioned classification problem and the sequential decision problem make sense? Comment adequately your answer.

The modelization makes sense, since an MDP is a more general model than the one we use for classification. None the less, since a classification problem does not show any temporal dependency among the samples (since they are i.i.d.) the use of an MDP to model this problem is an overcomplicated way to solve it.

Exercise 9 (3 marks)

Consider the following results for a linear regression model:

```
Linear regression model:
    y ~ 1 + x1 + x2
Estimated Coefficients:
                    Estimate        SE          tStat        pValue
                    --------    ---------    -------    -----------
    (Intercept)    0.040939      0.14883    0.27507         0.78385
    x1               1.9971    0.0052522     380.24    9.6971e-156
    x2               983.95       50.635     19.432     3.2493e-35
Number of observations: 100, Error degrees of freedom: 97
Root Mean Squared Error: 1.48
R-squared: 0.999,  Adjusted R-Squared 0.999
F-statistic vs. constant model: 7.32e+04, p-value = 6.64e-155
```

1. Do you think that the model has some problem? Why? Propose a solution to solve the model problem.
2. Is the number of samples enough w.r.t. the complexity of the model? Why?
3. Is the intercept meaningful for the model? Why?

1. Both the features seems to be meaningful for the problem (the p-values are almost zero), but their magnitude is significantly different. Maybe the data have not been normalized before running the linear regression. This could be a good way of solving the problem.
2. We have 97 degree of freedom for the model, which is meaningful since we have 100 observation and 2 parameters (not counting the constant, see answer 3). the rule of thumb is that we need 10^c samples, where c is the number of parameters, which is the current case.
3. Since its p-value is high $p = 0.78$, it is likely that the model does not require to consider a constant term.

Exercise 10 (3 marks)

The decision stumps in 2D are the classifiers that use either a single horizontal or a single vertical line in 2D to separate points in a plane. Show that the VC dimension of decision stumps in 2D is 3.

Since the decision stumps in 2D are a model which is less flexible than linear boundaries, which have $VC = 3$, they should have $VC(\mathcal{H}) \leq 3$. The proof that $VC(\mathcal{H}) = 3$ is by enumeration: