DIPARTIMENTO DI ELETTRONICA, INFORMAZIONE E BIOINGEGNERIA

Politecnico di Milano

Machine Learning (Code: 097683)

July 2, 2019

Surname:	
Name:	
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 either outside these spaces or **without a motivation** will not be considered for the final mark.
- During this exam you are **not allowed to use electronic devices**, such as 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 **Bayesian linear regression** technique and compare it with the **Ridge regression** technique.

Exercise 2 (5 marks)

Describe what is a **valid kernel** function.

Exercise 3 (5 marks)

Describe and compare the **Value Iteration** algorithm and the **Policy Iteration** algorithm.

Exercise 4 (2 marks)

```
load iris_dataset.mat;
x = zscore(irisInputs');
t = irisTargets(1,:)';
method_1 = mnrfit(x, t+1); % The plus one is for the standard of
   mnrfit method
method_2 = perceptron;
method_2 = train(method_2, x', t');
method_3 = fitcnb(x, t);
```

1. Describe the instructions of the `Matlab` code above. What problem are we solving?
2. Provide motivations for the use of each one of the above methods over the others when solving the considered problem.
3. Do you think that the code is complete or you would add some other operations to properly solve this problem?

1. The following instructions load the iris dataset, normalize the input data, and uses them to train three different kind of classifiers. At first, it uses a logistic regression, then a perceptron and, at last, it considers a Naive Bayes classifier.
2. The logistic regression has a convex loss function, therefore there exists a single parameter vector realizing the minimum of the loss. The perceptron learning algorithm provides a simple update rule, whose process converges if the dataset is linearly separable. The Naive Bayes has the pros to be a generative model, and, therefore, it allows to generate new data from the model one trains.
3. Since the classes in the iris dataset are ordered, it might be a good idea to shuffle the data for the perceptron learning process to be effective. Moreover, one might argue that the three models has only been trained and no analysis have been conducted on their performance. Therefore, a model selection procedure might complete the process.

Exercise 5 (2 marks)

Tell if the following statements about Reinforcement Learning (RL) are true or false. Motivate your answers.

1. The value function estimation provided by $TD(0)$ is equivalent to the Monte Carlo one.
2. The use of an ε-greedy policy in control RL problems is required to incentivize exploitation.
3. Eligibility traces are used to distribute the instantaneous reward over multiple time steps.
4. Importance sampling allows to re–use experience generated by old policies in RL algorithms.

1. FALSE: $TD(0)$ is equivalent to Temporal difference.
2. FALSE: the fact that it chooses a random policy with probability ε incentivizes the exploration of possibly sub-optimal action by the policy.
3. TRUE: it allows to update more than one value functions after each step.
4. TRUE: it reweights the samples to w.r.t. a sampling policy and a target one, to give the correct importance to each step.

Exercise 6 (2 marks)

Categorize the following ML problems:

1. Intrusion detection (determine if a system under attack by malicious users).
2. Group malware w.r.t. homogeneous characteristics.

For each one of them suggest a set of features which might be useful to solve the problem and a method to solve it.

1. Intrusion detection: we would like to determine if a system is under attack or not, therefore we are solving a binary classification problem. The features might be the amount of requests we get from external users and/or the amount of denied requests the firewall provides you. A method to solve this problem is SVM.

2. Group malware w.r.t. homogeneous characteristics: since we do not have any label corresponding to each malware we are in an unsupervised learning problem, specifically, in a clustering one. The features one might use are the dimension of the code, the language of the comments in the source files, the type of data the malware is accessing. We can use any clustering method like K-means (which has not been covered by the course and is not required as answer). In the case we assume to have labels for the malwares we will resort to multi-class classification algorithms, e.g., Naive Bayes.

Exercise 7 (2 marks)

Comment on the following statements about Gaussian Processes (GP). Assume to have a dataset generated from a GP $D = (x_i, y_i)_{i=1}^N$. Motivate your answers.

1. GPs are parametric methods.
2. The computation of the estimates of the variance of the GP $\hat{\sigma}^2(x)$ corresponding to the input x provided by D does not require the knowledge of the samples output $(y_1, \ldots, y_N)$.
3. In the neighbourhood of the input points $(x_1, \ldots, x_N)$, we observed the variance of the GP gets smaller and smaller as we collect more samples.
4. The complexity of the computation of the estimates of the mean $\hat{\mu}(x)$ and variance $\hat{\sigma}^2(x)$ scales as N^3, i.e., cubically with the number of samples N.

1. FALSE: they require to store the gram matrix whose dimension depends on the number of samples.
2. TRUE: it requires only the gram matrix and the computation of the kernel on the new point.
3. TRUE: the uncertainty we have around the sampled points decreases as we get more and more samples.
4. TRUE: indeed, it requires the inversion of the gram matrix which has N^3 computational cost.

Exercise 8 (2 marks)

Tell if the following statements about Multi-Armed Bandit are true or false and provide motivations for your choice.

1. The MDP corresponding to a MAB setting has no actions.
2. One of the Bayesian approaches to solve the MAB setting resorts to a sampling procedure from posterior distributions.
3. The use of an ε-greedy algorithm to solve the MAB setting is a viable solution and provides an expected pseudo-regret of order $O(\log(\log(T)))$, where T is the number of rounds.
4. The lower bound on the weak regret of the adversarial MAB setting is of order of $\Omega(\sqrt{T})$.

1. FALSE: determining the correct action among a finite set is the goal of MAB algorithms.
2. TRUE: it is Thompson Sampling, which selects the arm to pull depending on the largest sample generated from posterior distributions.
3. FALSE: The stochastic MAB problem has a lower bound of $O(\log(T))$, therefore, no algorithm can achieve a better regret bound.
4. TRUE: there exist a theoretical result which states exactly that the weak regret of the adversarial MAB setting is of order of $\Omega(\sqrt{T})$.

Exercise 9 (3 marks)

Given an MDP with four states $\mathcal{S} = \{A, B, C, D\}$ (D is terminal), three actions $\mathcal{A} = \{u, d, s\}$, given a policy π, and given the following episodes:

$$(A, u, 3) \to (C, s, 2) \to (B, d, 1) \to (D)$$
$$(B, d, 2) \to (C, u, 1) \to (D)$$
$$(A, u, 1) \to (B, u, -1) \to (A, s, 1) \to (D)$$

1. Compute the state values functions resorting to MC with every-visit and first-visit approach.
2. Assuming a discount factor $\gamma = 0.5$ and $\alpha = 0.1$, compute the same values by resorting to TD? Assume to start from a zero value for each state.
3. Compute the action-value function resorting to MC with every-visit.

1. MC every-visit:

$$V(A) = \frac{6+1+1}{3} = \frac{8}{3} \qquad V(B) = \frac{1+3+0}{3} = \frac{4}{3} \qquad V(C) = \frac{3+1}{2} = 2$$

MC first-visit:

$$V(A) = \frac{6+1}{3} = \frac{7}{3} \qquad V(B) = \frac{1+3+0}{3} = \frac{4}{3} \qquad V(C) = \frac{3+1}{2} = 2$$

2.

$$V(s_t) \leftarrow V(s_t) + \alpha(R_t + \gamma V(s_{t+1}) - V(s_t))$$
$$V(A) \leftarrow V(A) + 0.1(R_t + 0.5V(C) - V(A)) = 0 + 0.1(3 + 0.5 \cdot 0 - 0) = 3/10$$
$$V(C) \leftarrow V(C) + 0.1(R_t + 0.5V(B) - V(C)) = 0 + 0.1(2 + 0.5 \cdot 0 - 0) = 2/10$$
$$V(B) \leftarrow V(B) + 0.1(R_t + 0.5V(D) - V(B)) = 0 + 0.1(1 + 0.5 \cdot 0 - 0) = 1/10$$
$$V(B) \leftarrow V(B) + 0.1(R_t + 0.5V(C) - V(B)) = 1/10 + 0.1(2 + 0.5 \cdot 2/10 - 1/10) = 3/10$$
$$V(C) \leftarrow V(C) + 0.1(R_t + 0.5V(D) - V(C)) = 2/10 + 0.1(1 + 0.5 \cdot 0 - 2/10) = 28/100$$
$$V(A) \leftarrow V(A) + 0.1(R_t + 0.5V(B) - V(A)) = 3/10 + 0.1(1 + 0.5 \cdot 3/10 - 3/10) = 385/1000$$
$$V(B) \leftarrow V(B) + 0.1(R_t + 0.5V(A) - V(B)) = 3/10 + 0.1(-1 + 0.5 \cdot 385/1000 - 3/10) = 0.1892$$
$$V(A) \leftarrow V(A) + 0.1(R_t + 0.5V(A) - V(D)) = 385/1000 + 0.1(1 + 0.5 \cdot 0 - 385/1000) = 0.4465$$

3.

$$\begin{array}{lll} Q(A,u) = \frac{6+1}{2} = \frac{7}{2} & Q(A,d) = N.A. & Q(A,s) = 1 \\ Q(B,u) = 0 & Q(B,d) = \frac{1+3}{2} = 2 & Q(B,s) = N.A. \\ Q(C,u) = 1 & Q(C,d) = N.A. & Q(C,s) = 3 \end{array}$$

Exercise 10 (3 marks)

Given the following results given by the use of Linear Regression on a dataset:

```
Estimated Coefficients:
                   Estimate        SE         tStat        pValue
                   --------    --------     -------     ----------
(Intercept)        -0.88892     0.42963      -2.069       0.084001
x1                   2.0714    0.057002       36.34     2.8962e-08
x2                   3.5689     0.47751      7.4739     0.00029617
x3                  -0.34271     0.13587     -2.5224       0.045143
Number of observations: 10, Error degrees of freedom: 6
Root Mean Squared Error: 0.196
R-squared: 0.996,  Adjusted R-Squared 0.994
F-statistic vs. constant model: 499, p-value = 1.39e-07
```

answer the following questions and motivate your answers.

1. Do you think that all the features are significant?
2. Do you think that at least one of the features is significant?
3. How much is the RSS for this model?
4. How much variance is explained by the model?
5. How much error is this model making on average on a new data point?
6. Do you think that the model is sound?

1. NO, it seems that x_1, x_2 and x_3 are significant, according to the p-value (which is smaller than 0.05). One might argue that at a confidence level of 0.1 also the intercept is significant.
2. YES: since the F-statistic has a low p-value.
3. $RSS = RMSE^2 \cdot n = 0.196^2 \cdot 10 = 0.3842$.
4. According to the R^2 index the 99.6% of the variance has been explained by the model.
5. This is the $RMSE$, therefore, 0.196.
6. It seems that a model with 4 parameters for a dataset with 10 samples might be a little bit too complex. One should check if this model is overfitting before being sure to use it.