26/06/2020

Q1

Categorize the following ML problems. For each one of them suggest a proper method to solve it and the set of features/actions/states which might be useful to solve the problem. Motivate your answers.

- Play optimally a bridge game.
Reinforcement Learning, State representation would include (1) the result of the bidding stage, (2) my cards, (3) previous played cards. Action consists of choosing the card to play. Sarsa or Q-Learning could be used.

- Escape from a labyrinth.
RL, State could be the agent's perception (as well an estimate of the position). Actions are the ones that allow us to navigate the labyrinth. Sarsa or Q-Learning could be used.

- Display an advertisement on a search engine page.
It can be modeled as a MAB, where the choice of the advertisement is the action. UCB1 could be used. Contextual MAB can be also used to model the search query as context.

- Recognise handwritten digits and characters.
Classification problem. Features would be the image of the digit/character to recognize. Convolutional Neural Networks are a common choice for this problem.

- Identify the most informative features from an image.
This could be modeled as a feature extraction problem (e.g., using PCA or applying filters to images). Alternatively we could see it also as a feature selection problem, in the case we have a supervised problem for which ranking the importance of the features.

- Predicting the value of motorcycles for second-hand retailers.
Regression Problem. Features could be some characteristics of the bike, e.g., mileage, age. We can use Linear Regression.

- Design an automatic medical system providing a diagnosis based on symptoms.
Classification problem. Features are symptoms. Any classification algorithm like the perceptron or SVM can be used.

- Determine which are the most important factors (age, gender, geographical position) when trying to predict the most suitable person to hire.
This is a feature selection problem: assuming we have labeled samples (hired/not hired) or a target (e.g., productivity of employees) we can apply a wrapper feature selection algorithm, such as Forward Stepwise Selection.

- Estimate time of arrival of an airplane.
This can be solved also using physics, but we could model this also as a regression problem if we want to learn from historical data. Features would be the route, time of the day, weather data, takeoff delay. Linear Regression can be used.

- Identify the country given its flag.
This can be modeled as a Classification problem, where we want to identify a country from a picture of its flag. Convolutional Neural Networks are a common choice for this kind of problem. One might also argue that the problem cannot be solved with ML since we have a single sample (flag) for each country.

- Tell if a bank is going to approve a credit line for a client.
This is a binary classification problem. Features would be data about the client, e.g., income, age, other credit lines, guarantees. We can use Logistic Regression.

Q2

Answer to the following questions about regression and classifications, providing motivations.

- You are asked to implement an algorithm to solve a regression problem. The training will be performed on a device with constrained resources: which solution would you propose?
I could train a linear regression model using gradient descent (LMS algorithm).
Also non parametric methods can be used as they do not require expensive training.

- You are asked to implement an algorithm to solve a regression problem. The training will be performed on a server with a great computational power, but the resulting model should be run on a device with constrained resources: which solution would you propose?
I could train a linear regression model using least squares or lasso (to reduce the number of non-zero weights). The resulting model can be run without a large computational power.
Alternatively, I could train a Support Vector Machine for regression: expensive to train but not very expensive during test.

- You are asked to implement an algorithm to solve a classification problem. The training will be performed on a server with a great computational power, but the resulting model should be run on a device with constrained resources: which solution would you propose?
I could train a Support Vector Machine: expensive to train but not very expensive to compute the prediction. In general any parametric method would be a good choice since the prediction is usually light to be evaluated.

- You decide to use a perceptron to classify your system, but you find out that the two classes are not linearly separable in the current feature space. What would you do?
I could either use a different method (e.g., SVM) or apply a feature mapping to input space to make the problem linearly separable into the feature space.

- The company you work for wants to extend your classification system to the online setting. If you are using a perceptron, which kind of changes should you make? What if you are using a logistic regression instead?
In the perceptron, if the new sample is classified correctly, I would not do anything. Otherwise, I will perform the update rule on the new sample and then run the algorithm again on the whole dataset until convergence. In the case of logistic regression, I can simply perform the update rule on the new sample. Eventually, also in this case I can also go through the entire dataset again. In both cases we use the online update offered by the training of the two methods.

- You are asked to extend your regression system to the online setting. Which are the available options? Can you exploit the results obtained before from the offline setting?
If I am using Bayesian Linear Regression, I can use the old data as prior and update the model with new data. Alternatively I can use a gradient descent training algorithm (e.g., LMS) to train incrementally my model.

- You have trained two different classification models and you have computed their confusion matrices. Which figure of merits should you look at to choose the best model for an anomaly detection task?
In an anomaly detection task, I am usually interested in minimizing the False Negative. On the other hand this might easily lead to many False Positives. Hence it is necessary to find a good tradeoff between these two errors. One way of doing that is by looking at other index like the F1 score.

- You have performed a regression on stock prices time series. Your model has adjusted R^2 equal to 0.01. Do you think that the predictions of this model are informative for the specific application?
An adjusted R^2 equal to 0.01 means that there is basically no relationship between my input variable and the target (the stock price). So I cannot use the output of this prediction model for any trading application.

- While solving a regression task, you find that X^TX matrix is singular. What does it imply? Can you still solve the problem?
In this case I cannot compute OLS, because I am not able to invert the X^TX matrix. This happens when some variables are linearly dependent. To solve this problem I can either find and get rid of dependent variables or I can use ridge regression.

- Would it make sense to use classification algorithms to solve a Multi-Armed bandit problem?

In principle, we can use a classification algorithm and solve it as on online classification problem. However we would face a major problem: in MAB we don't know in advance the best arm and we don't get this information as feedback after choosing an arm. In addition classification algorithms are designed to exploit and not to explore.

- Is it correct to say that the RL policy evaluation problem is a regression task?
No, it is not correct, as the targets (the state values I am trying to learn) are not provided in the training data but they are learned through interaction and there are dependences between their values.

Q3

Tell if the following statements about the bias-variance dilemma (and related topics) are true or false. Motivate your answers.

- If a perceptron classifier does not achieve the desired performance (on a test set) one might consider an SVM with a linear kernel as a method to improve the performances.
TRUE: Even if SVM with linear kernel still find a linear boundary, it might find a better solution in terms of test error, because will find the maximum margin separating hyperplane and soft margin can be used to improve generalization.

- If a logistic regression does not achieve the desired performance (on a test set) one might consider an SVM with a Gaussian kernel as a method to improve the performances.
FALSE: if performance is good on the training set, it means we are overfitting data. SVM is a more complex data and it will probably result in even worse overfitting.

- If the performance on the training and the test sets are getting almost the same as we are using more and more data for training, but both are worse than desired, the model might be too simple for the task.
TRUE: increasing training data will reduce the variance of the model, hence the error on both training set and test set are likely to depend on bias (in fact, being the performance bad also on training set we are not overfitting the noise in data). Using a more complex model seems the proper choice.

- Given a fixed training set, increasing the complexity of the model always improves its generalization capabilities.
FALSE: the more complex is the model, the more data I usually need to train it in order to avoid overfitting.

- It is a good idea to increase the number of samples used for training if we decided to increase the model complexity.

TRUE: the more complex is the model, the more data I usually need to train it in order to avoid overfitting.

- Adding new features to the model might help if the model has a bias with respect to the real process generating data.
TRUE: the bias might be due to some features the model is not using as input.

- The use of cross-validation decreases the variance of the model.
TRUE/FALSE: cross-validation is used only to assess the performance of the model and does not reduce their variance, however through a proper assessment I can hopefully select a model with lower variance.

- Regularization techniques are able to reduce the variance of a model.
TRUE: regularization acts as constraints or penalty on parameters variability and hence reduce the variance of the model

- The use of the soft-margin SVM decreases the variance of the method (compared to hard-margin SVM).
TRUE: Hard margin allows no error and can easily overfit training data, while soft margin allowing some errors in the training set might lead to a model with better generalization.

- The use of the error on a validation set is suggested when we have a large dataset and we want to discriminate the performance of a set of (computationally) simple models.
FALSE: with computationally simple models, cross-validation is suggested to get a better estimate of the test error.

- If a set of hypothesis spaces have finite VC dimension, we can use the minimization of the training error to select the best model (among the given set).
TRUE/FALSE: a finite VC dimension means that we cannot overfit completely the training set, however it is still possible that training error is not an unbiased estimate of test error (especially if training set is small with respect to the VC dimension)

- If we focus on a specific task, it does not exist an ML algorithm that performs better than the others.
FALSE: If we focus on a specific task, a single ML algorithm can outperform others. This is not true if we, instead, consider all the possible problems (no free lunch theorem).

Q4

Tell if the following statements about Reinforcement Learning are true or false, and motivate your answers.

- For policy evaluation, if I have a simulator of the RL task, I can use DP, but not Monte Carlo or TD.
FALSE: DP requires the knowledge MDP model, instead a simulator can be used to simulate experience and thus can be used with MC and TD.

- Off-policy learning, Exploring Starts, and Soft-policies are three ways to deal with the Exploration-Exploitation dilemma.
TRUE: These three approaches allows to keep exploring while trying to find optimal policy

- The sparsity of the reward can be a problem for on-policy algorithms.
TRUE: Despite reward sparsity is a general issue for RL algorithms, an on-policy algorithm might not be able to explore enough to reach states with higher reward, while an off-policy algorithm might be more aggressive in the exploration.

- TD based control algorithms can be applied also to the MAB setting, while MC ones cannot.
TRUE/FALSE: Both MC and TD are not specifically designed to solve MAB setting. On the other hand we can argue that MAB setting is not an episodic task and, thus, MC is not suited for it, while TD could be applied in principle.

- If the actions do not affect the dynamics of the system, an RL problem is equivalent to a contextual bandit.
TRUE: In this case we have no control on the state and actions only affect the reward achieved.

- I can use Q-Learning with a greedy policy in an Adversarial Bandit context.
FALSE: because using a deterministic policy in an adversarial problem is not recommended.

- Sarsa, as Value Iteration in DP, is based on the Bellman Expectation Equation.
FALSE: Sarsa is based on Bellman Expectation Equation, while Value Iteration exploits Bellman Optimality Equation.

- Applying a TD approach to a problem, I can better exploit the markovianity of the state.
TRUE: TD does bootstrapping and, thus, it exploits the markov property of the state.

- A Markov Decision Process can always be solved analytically.
FALSE: Even if we have a full knowledge of the MDP model, an analytical solution can be computationally infeasible.

- Differently from Q-learning, SARSA cannot handle the exploration-exploitation trade-off.
FALSE: In SARSA the exploration-exploitation trade-off is handled using epsilon-soft policies

- In an MDP a stochastic policy cannot be optimal.

FALSE: In a finite MDP, for sure there is at least an optimal deterministic policy. This, however, does not prevent the existence of a stochastic optimal policy.

Domanda **22**

Risposta non data

Punteggio max.: 7,00

Consider the set of trajectories below, which are obtained by running a given policy in an MDP with three states S={A, B, C} (C is terminal) and two actions A={up, down}.
A trajectory is a sequence of (S_t,A_t,R_t):

(A,up,2) -> (A,down,4) -> (B,down,4) -> (C)
(B,up,2) -> (A,down,4) -> (C)

1. Compute the value functions V(A) and Q(A, down) by resorting to first-visit Monte-Carlo evaluation.
2. Compute the value function V(A) by resorting to Temporal-Difference evaluation. Assume to start from zero value for each state, α=0.5, γ=1.
3. Consider the values V(A) obtained with Monte Carlo and Temporal Difference. Can you tell if, given an infinite number of trajectories, the computed values would be different?

1)
V(A) = (10 + 4) / 2 = 7
Q(A,down) = (8 + 4) / 2 = 6

2)
V(A) = 0 + 0.5 * (2 + 0 - 0) = 1
V(A) = 1 + 0.5 * (4 + 0 - 1) = 2.5
V(A) = 2.5 + 0.5 * (4 + 0 - 2.5) = 3.25

3)
Since TD and MC are consistent estimators, they will converge to the same value V^\pi (A)

Domanda **23**

Risposta non data

Punteggio max.: 7,00

Consider the set of trajectories below, which are obtained by running a given policy in an MDP with three states S={A, B, C} (C is terminal) and two actions A={up, down}.
A trajectory is a sequence of (S_t,A_t,R_t):

(A,down,4) -> (B,up,2) -> (C)
(A,up,2) -> (B,down,4) -> (A,down,4) -> (C)

1. Compute the value functions V(A) and Q(A, down) by resorting to first-visit Monte-Carlo evaluation.
2. Compute the value function V(A) by resorting to Temporal-Difference evaluation. Assume to start from zero value for each state, $\alpha=0.5$, $\gamma=1$.
3. Consider the values V(A) obtained with Monte Carlo and Temporal Difference. Can you tell if, given an infinite number of trajectories, the computed values would be different?

1)
V(A) = (6 + 10) / 2 = 8
Q(A,down) = (6 + 4) / 2 = 5

2)
V(A) = 0 + 0.5 * (4 + 0 - 0) = 2
V(B) = 0 + 0.5 * (2 + 0 - 0) = 1
V(A) = 2 + 0.5 * (2 + 1 - 2) = 2.5
V(A) = 2.5 + 0.5 * (4 + 0 - 2.5) = 3.25

3)
Since TD and MC are consistent estimators, they will converge to the same value V^\pi (A)

Domanda **24**

Risposta non data

Punteggio max.: 7,00

Consider the set of trajectories below, which are obtained by running a given policy in an MDP with three states S={A, B, C} (C is terminal) and two actions A={up, down}.
A trajectory is a sequence of (S_t,A_t,R_t):

(A,up,2) -> (A,down,8) -> (B,down,4) -> (C)
(B,up,2) -> (A,down,4) -> (C)

1. Compute the value functions V(A) and Q(A, down) by resorting to first-visit Monte-Carlo evaluation.
2. Compute the value function V(A) by resorting to Temporal-Difference evaluation. Assume to start from zero value for each state, α=0.5, γ=1.
3. Consider the values V(A) obtained with Monte Carlo and Temporal Difference. Can you tell if, given an infinite number of trajectories, the computed values would be different?

1)
V(A) = (14 + 4) / 2 = 9
Q(A,down) = (12 + 4) / 2 = 8

2)
V(A) = 0 + 0.5 * (2 + 0 - 0) = 1
V(A) = 1 + 0.5 * (8 + 0 - 1) = 4.5
V(A) = 4.5 + 0.5 * (4 + 0 - 4.5) = 4.25

3)
Since TD and MC are consistent estimators, they will converge to the same value V^\pi (A)

Domanda **25**

Risposta non data

Punteggio max.: 7,00

Consider the set of trajectories below, which are obtained by running a given policy in an MDP with three states S={A, B, C} (C is terminal) and two actions A={up, down}.
A trajectory is a sequence of (S_t,A_t,R_t):

(A,down,4) -> (B,up,2) -> (C)
(A,up,2) -> (B,down,4) -> (A,down,8) -> (C)

1. Compute the value functions V(A) and Q(A, down) by resorting to first-visit Monte-Carlo evaluation.
2. Compute the value function V(A) by resorting to Temporal-Difference evaluation. Assume to start from zero value for each state, α=0.5, γ=1.
3. Consider the values V(A) obtained with Monte Carlo and Temporal Difference. Can you tell if, given an infinite number of trajectories, the computed values would be different?

1)
V(A) = (6 + 14) / 2 = 10
Q(A,down) = (6 + 8) / 2 = 7

2)
V(A) = 0 + 0.5 * (4 + 0 - 0) = 2
V(B) = 0 + 0.5 * (2 + 0 - 0) = 1
V(A) = 2 + 0.5 * (2 + 1 - 2) = 2.5
V(A) = 2.5 + 0.5 * (8 + 0 - 2.5) = 5.25

3)
Since TD and MC are consistent estimators, they will converge to the same value V^\pi (A)

Domanda **26**

Risposta non data

Punteggio max.: 7,00

Consider a linear, hard-margin, two-class SVM classifier defined by parameters w = [1, 2], b = -2.
Answer the following questions providing adequate motivations.

1) Provide the analytical formula of the boundary and the margins.
2) How the point x1 = [1; 3/4] is classified according to the trained SVM?
3) Assume to collect a new sample x2 = [1; 3] in the negative class. Do you need to retrain the SVM? In which case would you resort to a soft-margin version?

1)
x + 2*x2 - 2 = 0 (boundary)
x + 2*x2 - 3 = 0 (positive margin)
x + 2*x2 - 1 = 0 (negative margin)

2)
1 + 3/2 - 2 = - 0.5 -> negative class

3)
1 + 6 - 2 = + 5 -> it would be misclassified, thus you need to retrain and x2 is a SV.
You would use a soft-margin if the problem is no more lienearly separable.

Domanda 27

Risposta non data

Punteggio max.: 7,00

Consider a linear, hard-margin, two-class SVM classifier defined by parameters w = [-1, 2], b = 2.
Answer the following questions providing adequate motivations.

1) Provide the analytical formula of the boundary and the margins.
2) How the point x1 = [1; -3/4] is classified according to the trained SVM?
3) Assume to collect a new sample x2 = [1; -3] in the positive class. Do you need to retrain the SVM? In which case would you resort to a soft-margin version?

1)
- x + 2*x2 + 2 = 0 (boundary)
- x + 2*x2 + 1 = 0 (positive margin)
- x + 2*x2 + 3 = 0 (negative margin)

2)
- 1 - 3/2 + 2 = - 0.5 -> positive class

3)
- 1 - 6 + 2 = - 5 -> it would be misclassified, thus you need to retrain and x2 is a SV.
You would use a soft-margin if the problem is no more lienearly separable.

Domanda **28**

Risposta non data

Punteggio max.: 7,00

Consider a linear, hard-margin, two-class SVM classifier defined by parameters w = [-1, 2], b = -2.
Answer the following questions providing adequate motivations.

1) Provide the analytical formula of the boundary and the margins.
2) How the point x1 = [-1; 3/4] is classified according to the trained SVM?
3) Assume to collect a new sample x2 = [-1; 3] in the positive class. Do you need to retrain the SVM? In which case would you resort to a soft-margin version?

1)
$- x + 2*x2 - 2 = 0$ (boundary)
$- x + 2*x2 - 3 = 0$ (positive margin)
$- x + 2*x2 - 1 = 0$ (negative margin)

2)
$+ 1 + 3/2 - 2 = 0.5$ -> positive class

3)
$1 + 6 - 2 = 5$ -> it would be correctly classified, thus you do not need to retrain.
You do not need a soft-margin since the problem is linearly separable.

Domanda **29**

Risposta non data

Punteggio max.: 7,00

Consider a linear, hard-margin, two-class SVM classifier defined by parameters w = [1, 2], b = 2.
Answer the following questions providing adequate motivations.

1) Provide the analytical formula of the boundary and the margins.
2) How the point x1 = [-1; -3/4] is classified according to the trained SVM?
3) Assume to collect a new sample x2 = [-1; -3] in the positive class. Do you need to retrain the SVM? In which case would you resort to a soft-margin version?

1)
x + 2*x2 + 2 = 0 (boundary)
x + 2*x2 + 3 = 0 (positive margin)
x + 2*x2 + 1 = 0 (negative margin)

2)
- 1 - 3/2 + 2 = 0.5 -> positive class

3)
- 1 - 6 + 2 = - 5 -> it would be misclassified, thus you need to retrain and x2 is a SV.
You would use a soft-margin if the problem is no more lienearly separable.

Domanda **30**

Risposta non data

Non valutata

Submit your test.

Scegli un'alternativa:

- a. I want my test to be graded
- b. I want to whidraw from the test