ML Exam Questions 9/2/2021

Question 1 (Categorize ML problems)

Categorize the following machine learning problems. For each one of them, suggest a technique to solve it and some features or actions/states pairs that might provide useful information for the specified technique.

Predict the outcome of a tennis match.
The outcome could be either binary (Win/Loss) or the actual score (for each player). The former is a classification problem and we can resort to SVM to solve it. If the outcome is the actual score (for each player) it is a regression problem and we might use Kernel Regression to solve it. In both cases, the features might be: previous results between the players, current ATP ranking, and the statistics of the players.

Select the players which are worthy of taking a specific role in a team.
We want to determine which player will get or not a spot in the team, therefore it is a classification problem, a multiclass one if we want to also select the role for each player. We can solve it with logistic regression using the players' statistics as input.

Evaluate the performance of a robot squad in a specific environment.
We are in an RL evaluation problem. We want to understand how good is a specific policy, which can be performed using either MC or TD estimation techniques. Here the actions are the movements of the robots, the actions allowed in the environment, and the states are the energetic state of each robot and their position in the environment.

Identify groups of similar songs (in terms of genre) from a large database.
We want to generate some grouping of samples starting from an unlabeled dataset. This is a clustering problem. The technique one might use is K-means and the features are the songs' length, author, and information about the release.

Determine the critical factors for the development of a disease.
We are facing a problem to understand the features that are the most relevant for a phenomenon, which is a feature selection problem. We can use any wrapper or filtering techniques. Here the features we want to test are the ones related to each potential patient and its geographical position.

Learn how to provide an automated customer service.
Assuming to have different pre-recorded answers, we might resort to a control RL technique, to learn the proper action (where each action corresponds to a specific answer) according to the situation (specific user and related context).

Deciding the most promising investment strategy, among a given set.
If we assume the environment is stationary, It is a MAB problem, solvable with e.g., UCB1 using the different treatment as arms. Instead, if we assume some opponents are playing against us, we should rely on adversarial MAB techniques, e.g., EXP3.

Prediction of the sales in the next month for a bakery shop.
If we assume the environment is stationary, we can use regression techniques, like linear regression. As features, one might choose the sales on the previous month, weather conditions, etc. Otherwise, one should rely on time series forecasting techniques.

Question 2 (RL)

Tell whether the following statements about MDP and RL are true or false. Motivate your answers.

1) Policy Evaluation always outputs the optimal value function.
FALSE: Policy Evaluation goal is to find the correct value function corresponding to the policy we are evaluating. The optimal value function is the value function corresponding to the optimal policy instead: in order to obtain it, we have to solve the MDP (e.g., with Value-Iteration).

2) The value function may decrease on some steps of Policy Iteration, but in the end, the algorithm outputs the optimal one.
FALSE, the policy improvement theorem guarantees that, at every step, the new policy is better than the previous one.

3) Employing a discount factor in the computation of the cumulative return in MDPs is only a mathematical trick to ensure the convergence of the return.
FALSE, the discount factor can also be interpreted as the probability for an episode to terminate at each step, or as how much we value immediate rewards w.r.t. future ones.

4) Dealing with the exploration-exploitation tradeoff is more crucial in SARSA than Q-learning.
TRUE, being SARSA an on-policy algorithm, the learned policy is used to collect samples too, hence it has to balance between obtaining high returns and exploring the state and action space. Q-learning is an off-policy algorithm, hence an explorative policy can be employed with a greedy learned policy.

5) The optimal policy of a Multi-Armed Bandit problem can be found with Value-Iteration.
TRUE, provided that I know the reward distributions of each arm, however, in that case, it would be probably simpler to find the mean of each one and then to choose the best.
(FALSE, in order to use value-iteration I have to know the model, hence the distributions, which are usually unknown to the learner in the MAB setting)

6) Given an MDP with a certain reward function, there is only a single policy that is optimal for it, and, for each optimal policy, there is only a single reward function for which it is optimal.
FALSE, for each MDP we are guaranteed that there is always at least one optimal policy, but this does not prevent having more optimal policies. Moreover, the same policy can still be optimal if we modify the reward function, sometimes this procedure can also speed up learning (reward reshaping).

7) In an MDP, for each state, we can always choose an action that is optimal, independently from the time, and from the past history.
TRUE, this is guaranteed from the existence of a stationary, Markovian, and deterministic optimal policy for each MDP.

8) It is always better to perform Policy Evaluation by solving a linear system instead of using Dynamic Programming.
FALSE, when the state space is small we can use the closed-form solution. However, even if the linear system approach offers the exact solution, for very large problems it can be impractical to solve it, hence, we can resort to using the approximated solution offered by DP

9) All sequential decision problems can be modeled as MDPs.
FALSE, to model a problem as an MDP, we have to assume that the environment state is fully observable and Markovian, hence the current state of the environment should be completely determined by the current observation made by the agent. In some cases, we can "transform" the state of the environment to make it Markovian.

10) As many policies can be optimal, there can be multiple optimal value functions in an MDP.
FALSE, the only optimal value function is the unique fixed point of the Bellman Optimal Operator, and all optimal policies share the same value function.

Question 3 (Model Selection)

Answer to the following questions about the bias-variance decomposition, model selection, and related topics. Motivate your answers.

You trained two models on a problem with 5 features: Model A using all the 5 features and Model B using only 3 features. Assuming they have similar performances on the training set, do you expect Model A to perform better on the test set?
NO, Model A is more complex and hence has a larger variance and probability of overfitting training data. So it would probably result in a worse test error.

You trained two models on a problem with 9 features: Model A using only 5 features and Model B using all the 9 features. Do you expect Model A to have a smaller training error than Model B?
NO, Model A is simpler and will probably have a larger bias resulting in a larger training error.

You trained a K-NN classifier and the performance on the validation set is much worse than the one on the training set. Would you increase the value of K?
YES, the model is overfitting training data. Increasing K will decrease variance and could reduce overfitting.

You trained a model with ridge regression and the performance on the validation set is much worse than the one on the training set. Would you decrease the regularization coefficient?
NO, the model is overfitting training data. Decreasing the regularization coefficient will increase variance and possibly also overfitting.

You used 10-fold cross-validation to tune the hyper-parameter K of a classifier. The model trained on the third fold with K=3 achieved the best performance overall. Based on this, would you set K=3?
NO, you should choose the value of K based on the average performance compute on all the 10 folds.

You used 10-fold cross-validation to select a classification model among several ones. Once you selected the model, would it be a good idea to re-train it on the whole dataset (i.e., all the 10 folds together)?
YES, cross-validation provides an unbiased estimate of the test error of each model. Once selected the model, it still is a good idea to use the whole data to train a possibly better model.

You trained 10 regression models applying different basis functions to the problem inputs. Assuming you don't have enough time to perform additional training, would you select the model with the lowest training error?

NO, the training error does not provide a useful estimate of the true error (test error) of the model. You could, instead, use some adjusted error measure to select the model.

You need to assess the performance of a model on a very large dataset. You discover that training your model on the whole dataset is not very (computationally) expensive. Would you use Leave-one-out (LOO) cross-validation?
NO, because even if a single training process is not expensive, if the dataset is very large, LOO will not be feasible. Instead, we should use K-fold cross-validation.

Code 1 (MAB)

Which algorithm is the following code implementing for the MAB setting? Describe in detail (line by line) what are the operations it performs and answer the following questions. Motivate your answers.

```
1  for tt = 1:T
2      for ii = 1:n_arms
3          hat_r(ii) = beta_dist(ii).random();
4      end
5      [~, ind(tt)] = max(hat_r);
6
7      outcome = mathcal_R(ind(tt)).random();
8
9      beta_dist(ind(tt)).a = beta_dist(ind(tt)).a + outcome;
10     beta_dist(ind(tt)).b = beta_dist(ind(tt)).b + 1 - outcome;
11 end
```

The algorithm implemented is Thompson Sampling, a Bayesian algorithm designed to tackle the stochastic MAB setting. It provides theoretical results about the regret in the order of log(T). In the code, at each round tt, we draw a sample hat_r(ii) from the posterior distribution beta_dist(ii) inferred for each arm (Lines 2-4), it selects the largest sample (Line 5), observes an outcome (Line 7), and updates the posterior with this outcome (Lines 9-10).

The algorithm updates the distributions of all the arms at each time step.
FALSE: only the arm corresponding to ind(tt) is updated at round tt.

This algorithm provides a regret bound of order log(T) for stochastic MAB and a regret bound of order sqrt(T) for adversarial ones.
FALSE: the TS algorithm has been designed for the stochastic MAB problem and provides a regret bound only for this setting. No known results are available for the adversarial setting.

As we proceed with the learning procedure, the distributions beta_dist() are getting more and more concentrated to the real mean of the process.
TRUE: as we increase the number of samples coming from an arm, the posterior of the beta shrinks to a single value, which is an estimate of the true mean of the arm reward.

As we proceed with the learning procedure the distributions mathcal_R() are getting more and more concentrated to the real mean of the process.
FALSE: the reward of an arm is stationary over time. It is not influenced by the algorithm we run on the environment.

The procedure is able to model processes in which the reward is Gaussian.

FALSE: this approach has been designed for Bernoulli rewards. If one wants to use Gaussian rewards, she/he should use a pair prior-posterior conjugate allowing to use a Gaussian as posterior.

The algorithm is using a specific pair of distributions called prior-posterior conjugates.
TRUE: we use the information contained in the prior and improve them with the ones provided by the samples. The fact that the two distributions are conjugates allows us to provide a procedure that updates efficiently the prior.

This procedure can be applied to generic RL control problems.
FALSE: the MAB setting takes into account a specific RL model with a single state. Using MAB for a generic RL problem requires extending the techniques used in this setting.

There might exist a more effective procedure that provides a regret of order less than log(T).
FALSE: there exists a lower bound for the MAB stochastic setting, telling that it is not possible to design procedure getting less than C * log(T) regret, being C>0 a constant.

Exercise 1 (SVM)

A)

Consider a linear, hard-margin, two-class SVM classifier defined by parameters w = [-3, -5], b = -5.

Answer the following questions, providing adequate motivations (including the math).

1) Provide the analytical formula of the boundary and the margins.

2) How is the point p_1 = [0.8; -1.0] classified according to the trained SVM?

3) Assume to collect a new sample p_2 = [1.0; -1.3] in the negative class. Do you need to retrain the SVM?

1)

-3*x_1 + -5*x_2 + -5 = 0 (boundary)

-3*x_1 + -5*x_2 + -6 = 0 (positive margin)

-3*x_1 + -5*x_2 + -4 = 0 (negative margin)

2) -3*0.8 + -5*-1.0 + -5 = -2.4 -> negative

3) -3*1.0 + -5*-1.3 + -5 = -1.5 -> negative

No need to retrain, the point is correctly classified.

You would use a soft-margin if the problem is no more linearly separable.

B)

Consider a linear, hard-margin, two-class SVM classifier defined by parameters w = [4, -1], b = -5.

Answer the following questions providing adequate motivations (including the math).

1) Provide the analytical formula of the boundary and the margins.

2) How is the point p_1 = [0.4; -2.7] classified according to the trained SVM?

3) Assume to collect a new sample p_2 = [0.8; -4.9] in the negative class. Do you need to retrain the SVM?

1)

4*x_1 + -1*x_2 + -5 = 0 (boundary)

4*x_1 + -1*x_2 + -6 = 0 (positive margin)

4*x_1 + -1*x_2 + -4 = 0 (negative margin)

2) 4*0.4 + -1*-2.7 + -5 = -0.7 -> negative

3) 4*0.8 + -1*-4.9 + -5 = 3.1 -> negative

it would be misclassified, thus you need to retrain and p_2 is a SV.

You would use a soft-margin if the problem is no more linearly separable.

C)

Consider a linear, hard-margin, two-class SVM classifier defined by parameters w = [4, -5], b = -4.

Answer the following questions providing adequate motivations (including the math).

1) Provide the analytical formula of the boundary and the margins.

2) How is the point p_1 = [0.2; -0.7] classified according to the trained SVM?

3) Assume to collect a new sample p_2 = [0.9; -0.6] in the negative class. Do you need to retrain the SVM?

1)
4*x_1 + -5*x_2 + -4 = 0 (boundary)
4*x_1 + -5*x_2 + -5 = 0 (positive margin)
4*x_1 + -5*x_2 + -3 = 0 (negative margin)

2) 4*0.2 + -5*-0.7 + -4 = 0.3 -> positive
3) 4*0.9 + -5*-0.6 + -4 = 2.6 -> positive
it would be misclassified, thus you need to retrain and p_2 is a SV.
You would use a soft-margin if the problem is no more linearly separable.

D)

Consider a linear, hard-margin, two-class SVM classifier defined by parameters w = [4, 3], b = -1.

Answer the following questions providing adequate motivations (including the math).

1) Provide the analytical formula of the boundary and the margins.

2) How is the point p_1 = [0.7; -0.7] classified according to the trained SVM?

3) Assume to collect a new sample p_2 = [0.0; -0.9] in the negative class. Do you need to retrain the SVM?

1)
4*x_1 + 3*x_2 + -1 = 0 (boundary)
4*x_1 + 3*x_2 + -2 = 0 (positive margin)
4*x_1 + 3*x_2 + 0 = 0 (negative margin)

2) 4*0.7 + 3*-0.7 + -1 = -0.3 -> negative
3) 4*0.0 + 3*-0.9 + -1 = -3.7 -> negative
No need to retrain, the point is correctly classified.

You would use a soft-margin if the problem is no more linearly separable.

Exercise 2 (Linear Regression)

A)

Consider the following coefficients estimated for a linear regression:

	Estimate	SE	tStat	pValue
(Intercept)	8.8471e-16	0.02053	4.3093e-14	1
x1	-0.22815	0.051551	-4.4258	1.869e-05
x2	0.12998	0.027838	4.6692	6.793e-06
x3	1.2163	0.056478	21.536	3.6039e-47

Number of observations: 150, Error degrees of freedom: 146
Root Mean Squared Error: 0.251
R-squared: 0.938, Adjusted R-Squared: 0.937
F-statistic vs. constant model: 737, p-value = 6.2e-88

Answer the following questions:

1. Do you think that at least one of the features is significant?
2. Do you think that all the features are 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?

Motivate your answers.

1. Yes: since the F-statistic has a low p-value.
2. No, it seems that x1, x2, and x3 are significant, according to the p-value (which is smaller than 0.05).
3. RSS = RMSE^2 n = 0.251^2 * 150 = 9.45
4. According to the R2 index the 93.8% of the variance has been explained by the model.
5. This is the RMSE, therefore: 0.251.
6. Yes, we have 150 samples and 4 variables, however, one may want to verify whether it is possible to obtain better results with a simpler model, e.g., only using the third feature.

B)

Consider the following coefficients estimated for a linear regression:

	Estimate	SE	tStat	pValue
(Intercept)	8.6188e-16	0.030949	2.7848e-14	1
x1	0.3429	0.03491	9.8225	8.529e-18
x2	1.5151	0.12063	12.56	5.4099e-25
x3	-0.51847	0.11715	-4.4258	1.869e-05

Number of observations: 150, Error degrees of freedom: 146
Root Mean Squared Error: 0.379
R-squared: 0.859, Adjusted R-Squared: 0.856
F-statistic vs. constant model: 297, p-value = 6.28e-62

Answer the following questions:

1. Do you think that at least one of the features is significant?
2. Do you think that all the features are 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?

Motivate your answers.

1. Yes: since the F-statistic has a low p-value.
2. No, it seems that x1, x2, and x3 are significant, according to the p-value (which is smaller than 0.05).
3. RSS = RMSE^2 n = 0.379^2 * 150 = 21.5461
4. According to the R2 index the 85.9% of the variance has been explained by the model.
5. This is the RMSE, therefore: 0.379.
6. Yes, we have 150 samples and 4 variables, however, one may want to verify whether it is possible to obtain better results with a simpler model, e.g., only using the first two features.

C)

Consider the following coefficients estimated for a linear regression:

	Estimate	SE	tStat	pValue
(Intercept)	-0.0050514	0.041701	-0.12113	0.90384
x1	0.46141	0.13661	3.3776	0.0010614
x2	0.10669	0.1695	0.62943	0.53058
x3	0.38179	0.10309	3.7033	0.0003571

Number of observations: 100, Error degrees of freedom: 96
Root Mean Squared Error: 0.415
R-squared: 0.835, Adjusted R-Squared: 0.83
F-statistic vs. constant model: 160, p-value = 5.07e-37

Answer the following questions:
1. Do you think that at least one of the features is significant?
2. Do you think that all the features are 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?

Motivate your answers.

1. YES: since the F-statistic has a low p-value.
2. NO, it seems that x1, x3, and the are significant, according to the p-value (which is smaller than 0.05), while feature x2 and the intercept are not.
3. RSS = RMSE^2 n = 0.415^2 * 100 = 17.2225
4. According to the R2 index the 83.5% of the variance has been explained by the model.
5. This is the RMSE, therefore: 0.415.
6. Yes, we have 100 samples and 4 variables, however, it would be better to fit again the model removing the features that are not significant.

D)

Consider the following coefficients estimated for a linear regression:

	Estimate	SE	tStat	pValue
(Intercept)	-0.0059449	0.040643	-0.14627	0.88402
x1	-0.11639	0.057655	-2.0186	0.046343
x2	-0.05157	0.13343	-0.3865	0.69999
x3	0.87878	0.14193	6.1915	1.5109e-08

Number of observations: 100, Error degrees of freedom: 96
Root Mean Squared Error: 0.404
R-squared: 0.842, Adjusted R-Squared: 0.837
F-statistic vs. constant model: 168, p-value = 7.09e-38

Answer the following questions:

1. Do you think that at least one of the features is significant?
2. Do you think that all the features are 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?

Motivate your answers.

1. YES: since the F-statistic has a low p-value.
2. NO, it seems that x1 and x3 are significant, according to the p-value (which is smaller than 0.05), while x2 and the intercept are not.
3. RSS = RMSE^2 n = 0.404^2 * 100 = 16.3216
4. According to the R2 index the 84.2% of the variance has been explained by the model.
5. This is the RMSE, therefore: 0.404.
6. Yes, we have 100 samples and 4 variables, however, one may want to verify whether it is possible to obtain better results with a simpler model, e.g., only using the significant features.