CS771A: Introduction to Machine Learning
Users Online : 232

☰ Quiz 2

Submission Deadline
23/11/2020 19:05

Instructions

Timed: No

● **Please do not open the quiz in multiple tabs/browsers. All your responses may not be recorded.**

1) There are a total of **24 questions**, and the question paper is worth a total of **35 marks**.

2) Total duration for the quiz is **35 minutes** (with standard compensatory time for DAP students)

3) Multiple choice questions are each of 2 marks and can have more than one correct answer. Please mark all options that you think are correct. There is no negative marking for such questions.

4) True/False questions are each of 1 mark and have negative marking (1 negative mark for an incorrect answer)

5) For fill-in-the-blank questions, please write your final answer in the text box without any whitespaces before or after your answer. There is no negative marking for such questions.

6) For all questions, assume only the standard scenarios and not very rare/pathological cases. Specifically, consider all the problems only in the context of "normal" scenarios/situations that we have discussed in the course. In particular, for the purpose of this quiz at least, referring to any obscure research papers during or before/after the quiz is discouraged/unnecessary. :)

Wishing you the very best for the quiz.

Quiz Opens: 23/11/2020 18:30

This is only a quiz preview. As an instructor, only you can preview the quiz, not students.

RESUME QUIZ

CS771A: Introduction to Machine Learning
Users Online : 232   00:36:12

Submission Deadline : 23/11/2020 19:05

Quiz 2

Q.1 Standard K-means gives us a K-dimensional real-valued feature representation for each input.

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.2 Which of these is true for the K-means clustering algorithm? Select all correct options.

Max. score: 2; Neg. score: 0

When run twice with the exact same initialization, it will output the same clustering both times.

The running time is dependent on the number of clusters we want.

Kernelized K-means is in general slower than standard K-means

When initialized with K-means++ initialization, it will converge to the globally optimal solution.

Q.3 Unlike standard K-means, the kernelized K-means does not optimize any loss function to find the clustering.

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.4 Which of these classification approaches can potentially learn nonlinear boundaries even without kernelization? Select all correct options.

Max. score: 2; Neg. score: 0

Nearest Neighbors

Support Vector Machine

Q.5 Which of the following is true for the Gaussian (RBF) kernel? Select all correct options.

Max. score: 2; Neg. score: 0

   We can store a finite set of points in the kernel-induced feature space using a finite amount of storage

   We can compute the inner product between any two points in the kernel-induced feature space in finite amount of time and with finite amount of storage.

   We can compute and store the mean of a finite set of points in the kernel-induced feature space using a finite amount of storage.

   We can compute the Euclidean distance between any two points in the kernel-induced feature space in finite amount of time and with finite amount of storage.

Q.6 A generative model can be used for which of the following problems? Select all correct options.

Max. score: 2; Neg. score: 0

   Binary Classification

   Singe-output (scaled-valued) Regression

   Multi-output Regression

   Multiclass Classification

Q.7 If the number of training examples and features is the same and if we are using the same kernel function, then at test time, a kernelized SVM will be faster than kernelized ridge regression

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

   false

   true

Q.8 Assuming $\mathbf{S}$ to be the DxD covariance matrix of the inputs, the unit-norm vector $\boldsymbol{u}$ that minimizes the quantity $\boldsymbol{u}^\top \boldsymbol{S} \boldsymbol{u}$ will be the first principal component given by PCA

Max. score: 1; Neg. score: 1

false

true

Q.9 K-means clustering can be viewed as matrix factorization of the form $X = Z\mu$ where X is NxD, Z is NxK and μ is KxD. The number of nonzeros in the Z matrix would be

Max. score: 2; Neg. score: 0

> Enter your answer here

Q.10 Given data from two classes, suppose we train a linear binary SVM and obtain a weight vector $\mathbf{w}_1$. Now suppose we throw away the training examples from one of the classes and train a linear one-class SVM and obtain a weight vector $\mathbf{w}_2$ for this one-class SVM problem. The two weight vectors will be different in general.

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.11 The soft K-means clustering algorithm can handle non-spherical clusters

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.12 Standard PCA can not be used to increase the data dimensionality, i.e., number of projection directions (K) can't be larger than the data dimensionality (D).

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

CS771A: Introduction to Machine Learning
Users Online : 232   00:36:12

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.14 Generative classification with Gaussian class-conditionals can only learn linear decision boundaries.

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.15 Assuming N inputs in D dimensions, the K-means objective function's value will be the smallest possible if K is equal to:

Max. score: 2; Neg. score: 0

1

N

$\log_2(N)$

D

Q.16 Which of the following is true about the standard K-means clustering? Select all correct options.

Max. score: 2; Neg. score: 0

It assumes clusters to be of roughly equal sizes

When it converges, the cluster means coincide with K of the data points.

It assumes that the decision boundary between any pair of clusters is linear

It can handle arbitrary-shaped clusters

Trying to make training error smaller would typically lead to reduced margin

For the same training data, a soft-margin SVM would typically give a smaller margin than a hard-margin SVM

For the same training data, a soft-margin SVM would typically give a larger margin than a hard-margin SVM

Trying to make training error smaller would typically lead to increased margin

---

Q.18 Even if we keep all D principal components in PCA (assuming original data is D dimensional), the reconstruction error is not guaranteed to be zero

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.19 Which of these clustering algorithms require the number of clusters to be specified? Select all correct options.

Max. score: 2; Neg. score: 0

K-means

DBSCAN

Spectral clustering

Agglomerative clustering

Q.20 Using the landmarks-based approach for extracting kernel based features, the largest size of the feature vector that we can construct for each input if the training set size is N and the dimensionality is D will be:

Max. score: 2; Neg. score: 0

Enter your answer here

Q.21 Which of the following kernels will have finite dimensional feature maps? Select all correct options.

Linear kernel

Sum of a quadratic kernel and Gaussian kernel

Sum of a linear kernel and Gaussian kernel

Quadratic

Q.22 The weight vector of a kernelized SVM, since it is of the form $w = \sum_{n=1}^{N} \alpha_n y_n \phi(x_n)$, can never be stored as a finite dimensional vector

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.23 Logistic and softmax classification are examples of generative classification.

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

Q.24 The solution found by SVM for a linearly separable case is unique.

Max. score: 1; Neg. score: 1

[Imp. Note: If you wish to skip the question, you should do so immediately. Once an option is chosen (either true or false), you can not skip the question at a later stage.]

false

true

SAVE   SUBMIT