

Data Mining Quiz

Quiz for Data Mining course

- Q.1) The problem of finding hidden structure in unlabeled data is called...**
- A. ☐ Supervised learning
B. ☐ Unsupervised learning
C. ☐ Reinforcement learning
- Q.2) Task of inferring a model from labeled training data is called**
- A. ☐ Unsupervised learning
B. ☐ Supervised learning
C. ☐ Reinforcement learning
- Q.3) Some telecommunication company wants to segment their customers into distinct groups in order to send appropriate subscription offers, this is an example of ...**
- A. ☐ Supervised learning
B. ☐ Data extraction
C. ☐ Seriation
D. ☐ Unsupervised learning
- Q.4) Self-organizing maps is an example of ...**
- A. ☐ Unsupervised learning
B. ☐ Supervised learning
C. ☐ Reinforcement learning
D. ☐ Missing data imputation
- Q.5) You are given data about seismic activity in Japan, and you want to predict a magnitude of the next earthquake, this is in an example of ...**
- A. ☐ Supervised learning
B. ☐ Unsupervised learning
C. ☐ Seriation
D. ☐ Dimensionality reduction
- Q.6) Assume you want to perform supervised learning and to predict number of newborns according to size of storks' population (<http://www.brixtonhealth.com/storksBabies.pdf>), it is an example of ...**
- A. ☐ Classification
B. ☐ Regression
C. ☐ Clustering
D. ☐ Structural equation modeling
- Q.7) Discriminating between spam and ham e-mails is a classification task, true or false?**
- A. ☐ True
B. ☐ False
- Q.8) In the example of predicting number of babies based on storks' population size, number of babies is ...**
- A. ☐ outcome
B. ☐ feature
C. ☐ attribute
D. ☐ observation
- Q.9) It may be better to avoid the metric of ROC curve as it can suffer from accuracy paradox.**
- A. ☐ True
B. ☐ False
- Q.10) A hundred people were tested for HIV. 40 of them recieved positive answers, however only 25 had the disease. Fill in the confusion matrix below:**

	Choice		Match
A	True positives	1	0
B	True negatives	2	15
C	False positives	3	60
D	False negatives	4	25