

Department of Computer Science and Engineering

Indian Institute of Technology Jodhpur

CSL 7620: Machine Learning

Quiz 3

Total marks: 10

Set A- Answers

Time: 10 minutes

Instructions:

- During the exam, invigilators will not address any queries. If you encounter anything unclear or incorrect in a question, make a reasonable assumption and proceed
 - Each question may have multiple correct options. You have to mark all the correct options. Otherwise, no marks will be awarded.
 - Mark the correct choices with a pen on the OMR sheet. No correction is allowed in the marking.
-

Answers are marked in bold. [Each question carries 1 mark]

01. In Independent Component Analysis (ICA), which statement is true?

- a. ICA uses orthogonal transformations only
- b. ICA aims to maximize statistical independence**
- c. ICA requires labeled data
- d. None of the others

02. Which of the following assumptions is essential for ICA to successfully separate mixed signals?

- a. The source signals must be Gaussian distributed.
- b. The number of sources must be greater than the number of observed mixtures.
- c. The source signals must be non-Gaussian.**
- d. None of the others.

03. Which of the following statements about **Decision Trees** are true?

- a. They can handle both categorical and numerical data**
- b. Nine of the others
- c. They can be prone to overfitting on training data**
- d. They can model non-linear relationships**

04. Which criterion is typically not **used by decision trees** for feature selection at a node?

- a. Gini Index
- b. Entropy / Information Gain
- c. Chi-square test**
- d. Variance Reduction

05. Which of the following techniques helps in **reducing overfitting** in decision trees?

- a. Limiting the maximum depth of the tree**
- b. None of the others
- c. Increasing the number of features used per split
- d. Setting a minimum number of samples per leaf**

-
06. A node contains 8 samples: 5 positive and 3 negative. Compute the **entropy** (take base 2 for logarithm).
- a. None of the others
 - b. 0.81
 - c. **0.95**
 - d. 1.00
07. The 'randomness' in a standard Random Forest algorithm comes from which of the following two sources?
- a. Adding random noise to the training labels.
 - b. **Bootstrap aggregation (bagging) of the training samples.**
 - c. **Random feature selection at each split (feature bagging).**
 - d. Randomly pruning the trees after they are fully grown.
08. What is the primary theoretical difference in the *objective* of Bagging (e.g., Random Forest) versus Boosting (e.g., AdaBoost, GBM)?
- a. **Bagging trains independent learners in parallel, while Boosting trains dependent learners sequentially.**
 - b. Both methods are sequential, but Bagging re-weights samples while Boosting re-weights features.
 - c. Bagging is a sequential method to reduce bias, while Boosting is a parallel method to reduce variance.
 - d. Both methods are parallel, but Bagging uses deep trees (low bias) while Boosting uses shallow trees (high bias).
09. A Random Forest is built with N trees, where each tree is grown on a bootstrap sample (size N, with replacement). What is the approximate probability that a specific training sample x_i is *not* used in the training of a specific tree T_b (i.e., is 'out-of-bag')?
- a. It depends entirely on the max_features parameter.
 - b. $\approx 50\%$
 - c. **None of the others**
 - d. $\approx 63.2\%$
10. Which of the following are key mechanisms of the AdaBoost algorithm?
- a. Each learner h_m is trained on a different bootstrap sample of the original data.
 - b. Each new learner h_m is trained to predict the pseudo-residuals of the previous ensemble f_{m-1} .
 - c. **After each step m, the sample weights are updated, increasing the weights of misclassified samples.**
 - d. **The final prediction is a weighted vote of all learners, where α_m (the weight) is higher for learners with lower weighted error ϵ_m .**