



**Course Name: Introduction to Machine Learning**  
**Assignment – Week 7 (Computational Learning theory, PAC Learning, Sample Complexity, VC Dimension, Ensemble Learning)**

**TYPE OF QUESTION: MCQ/MSQ**

**Number of Question: 8**

**Total Marks: 8X2 = 16**

1. Which of the following option is / are correct regarding the benefits of ensemble model?

- 1. Better performance
- 2. More generalized model
- 3. Better interpretability

- A) 1 and 3
- B) 2 and 3
- C) 1 and 2**
- D) 1, 2 and 3

**Answer: C** (1 and 2 are the benefits of ensemble modelling. Option 3 is incorrect because when we ensemble multiple models, we lose interpretability of the models).

2. In AdaBoost, we give more weights to points having been misclassified in previous iterations. Now, if we introduced a limit or cap on the weight that any point can take (for example, say we introduce a restriction that prevents any point's weight from exceeding a value of 10). Which among the following would be an effect of such a modification?

- A) We may observe the performance of the classifier reduce as the number of stages increase.
- B) It makes the final classifier robust to outliers.**
- C) It may result in lower overall performance.**
- D) None of these.

**Answer: B, C** (Outliers tend to get misclassified. As the number of iterations increase, the weight corresponding to outlier points can become very large resulting in subsequent classifier models trying to classify the outlier points correctly. This generally has an adverse effect on the overall classifier. Restricting the weights is one way of mitigating this problem. However, this can also lower the performance of the classifier).

3. Which among the following are some of the differences between bagging and boosting?

- A) In bagging we use the same classification algorithm for training on each sample of the data, whereas in boosting, we use different classification algorithms on the different training data samples.



- B) **Bagging is easy to parallelize whereas boosting is inherently a sequential process.**
- C) **In bagging we typically use sampling with replacement whereas in boosting, we typically use weighted sampling techniques.**
- D) **In comparison with the performance of a base classifier on a particular data set, bagging will generally not increase the error whereas as boosting may lead to an increase in the error.**

**Answer:** Options (B), (C) and (D) are correct.

4. What is the VC-dimension of the class of circle in a 4-dimensional plane?

- A) 3
- B) 4
- C) **5**
- D) 6

**Answer:** C is the correct option.

5. Considering the AdaBoost algorithm, which among the following statements is true?

- A) In each stage, we try to train a classifier which makes accurate predictions on any subset of the data points where the subset size is at least half the size of the data set.
- B) **In each stage, we try to train a classifier which makes accurate predictions on a subset of the data points where the subset contains more of the data points which were misclassified in earlier stages.**
- C) The weight assigned to an individual classifier depends upon the number of data points correctly classified by the classifier.
- D) **The weight assigned to an individual classifier depends upon the weighted sum error of misclassified points for that classifier.**

**Answer:** B, D (The classifier chosen at each stage is the one that minimizes the weighted error at that stage. The weight of a point is high if it has been misclassified more number of times in the previous iterations. Thus, maximum error minimization is performed by trying to correctly predict the points which were misclassified in earlier iterations. Also, weights are assigned to the classifiers depending upon their accuracy which again depends upon the weighted error (for that classifier).



6. Suppose the VC dimension of a hypothesis space is 6. Which of the following are true?

- A) At least one set of 6 points can be shattered by the hypothesis space.
- B) No sets of 6 points can be shattered by the hypothesis space.
- C) All sets of 6 points can be shattered by the hypothesis space.
- D) No set of 6 points can be shattered by the hypothesis space.

**Answer:** A, D (From the definition of VC dimension)

- If there exists at least one subset of  $X$  of size  $d$  that can be shattered then  $VC(H) \geq d$ .
- If no subset of size  $d$  can be shattered, then  $VC(H) < d$ .
- From the above facts, options A and D are correct.

7. Ensembles will yield bad results when there is a significant diversity among the models.  
Write True or False.

- A) True
- B) False

**Answer:** B

Ensemble is a collection of diverse set of learners to improve the stability and the performance of the algorithm. So, more diverse the models are, the better will be the performance of ensemble.

8. Which of the following algorithms are not an ensemble learning algorithm?

- A) Random Forest
- B) Adaboost
- C) Gradient Boosting
- D) Decision Tress

**Answer:** D.

Decision trees do not aggregate the results of multiple trees, so it is not an ensemble algorithm.