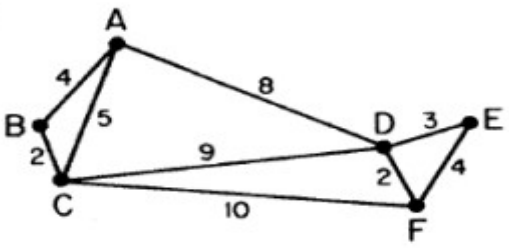
Agnel Charities'

Fr. C. Rodrigues Institute of Technology, Vashi,
Navi-Mumbai

Department of Computer Engineering
Question Bank

Subject : Machine learning

Sem- VII- A&B

Q. 1	**Weightage- 2M**
1.	Exemplify cross validation.
2.	Draw k-fold method process diagram to illustrate it's working.
3.	Give two real-world examples where it can be applied effectively. Discuss the advantages of ensemble learning in these scenarios.
4.	Exemplify XGBoost.
5.	Define the curse of dimensionality and its impact on machine learning algorithms.
6.	How does boosting improve model performance?
7.	Justify how dimensionality reduction techniques can be advantageous in real-world applications. Provide two specific examples to support your explanation.
8.	Mention the key point why PCA is mostly preferred technique.
9.	Elaborate how ensemble learning is different from traditional method of ML model building.
10.	Discuss the concept of decision stumps in stump-based ensemble learning.
11.	Exemplify the role of eigenvalues and eigenvectors in PCA.
12.	Exemplify Linear discriment analysis.
13.	List two main categories of dimensionality reduction techniques and briefly describe the purpose of each category.
Q. 2	**Weightage- 5 M**
14.	Compare and contrast bagging and boosting techniques. Include their basic principles, how they address model variance, and their differences in terms of ensemble member creation.
15.	Briefly describe K-fold cross-validation and its purpose in model evaluation.
16.	Compare PCA with LDA and SVA.
17.	Apply graph based minimum spanning tree method with delete branch with maximum weight strategy to find the clusters. 

18.	Apply the mentioned distance measure to find the distance. i) Jaccard distance A={1,2,3,4,5,8,9} B={0,1,3,4,5,6,7} ii) Euclidean distance object1= {2,3,0,5,6} object2={0,3,3,8,9} iii) Hamming distance A= 111000111 B=11110000 iv) No of Edit operations to convert " Hello" and "Hallow" v) Manhattan distance object1= {2,3,0,5,6} object2={0,3,3,8,9}
19.	Compare the boosting approach and the bagging approach in ensemble learning. Highlight their differences in terms of how they assign weights to misclassified samples and how they update models.
20.	Outline the process of bagging in Random Forest and discuss how it prevents overfitting.