COMP

Dec 24

Duration: 3 Hours **[Max. Marks: 80]**

N.B.: (1) Question No 1 is Compulsory.
(2) Attempt any **THREE** questions out of the remaining **FIVE**.
(3) All questions carry equal marks.
(4) Assume suitable data, if required and state it clearly.

Q1 Attempt any **FOUR** **[20]**

A Explain Training error and Generalization error.
B Differentiate between Supervised and unsupervised Learning
C Differentiate between Linear regression and Logistic regression.
D Explain issues in Machine learning.
E Explain performance evaluation metrics for the classification.

Q2 A Demonstrate MST algorithm along with example. **[10]**
B Explain Logistics regression and performance evaluation metrics. **[10]**

Q3 A Demonstrate steps to design a Machine Learning application. **[10]**
B What is over fitting, under fitting and Bias variance trade-off with reference to Machine learning? **[10]**

Q4 A Demonstrate Ensemble learning based Random Forest algorithm in detail. **[10]**
B Suppose we want Gini index to decide whether the car will be stolen or not. The target classification is "car is stolen?" which can be Yes or No, create a decision tree for the given data below. **[10]**

Car no	Colour	Type	Origin	Stolen ?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes

Car no	Colour	Type	Origin	Stolen ?
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Q5 A Give steps to design PCA dimensional reduction technique along with an example. **[10]**
B Demonstrate DBSCAN algorithm along with example. **[10]**

Q6 Write detailed note on following. (**Any TWO**) **[20]**

A Write a short note on XGBoost ensemble method.
B Explain support vector machine as constraint optimization problem.
C SVM Kernel trick

****