National Institute of Technology Kurukshetra
Machine Learning and Data Analytics-AIIC 221 (Civil Engineering)
Mid-Term Examination-II, Odd Semester, 2025-26

Time: 50 Minutes **Total Marks: 15**

Note: Attempt all the questions. If require any missing data; then choose suitably.

Q1. A) What is the curse of dimensionality? Why does it occur and how it affects machine learning? **[2 marks]**

B) Explain forward selection for feature selection in dimensionality reduction. **[2 marks]**

Q2. Consider the car theft dataset given in the table below. The attributes of the dataset are Colour, Type, Origin, and the target, Stolen can be either Yes or No. Use Naïve Bayers classification to predict whether a Red colour SUV of Domestic origin will be stolen? **[4 marks]**

Example No.	Color	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
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

Q3. Consider the following training examples.

Instances	Attribute1	Atrribute2	Class
1	True	T	+
2	True	T	+
3	True	F	-
4	False	F	+
5	False	T	-
6	False	T	-

Use information gain measure to create a decision tree with the above training examples. **[5 marks]**

Q4. A spam detection system detects 60 spam emails out of 150 emails, out of which 40 emails were actually spam emails. There were 70 spam emails and 80 non-spam emails originally. Find precision, recall, accuracy and F1 score of the spam detection system. **[2 marks]**