Roll No: | | | | | | | | | | | |

PRESIDENCY UNIVERSITY

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

Presidency University Act, 2013 of the Karnataka Act No. 41 of 2013 | Established under Section 2(f) of UGC Act, 1956
Approved by AICTE, New Delhi | Approved By BCI
Bengaluru

Mid Term Examinations - October 2026

Date: 05/10/2026 **Time:** 11:45 AM - 01:15 PM

Course Code: CBD3407 **Course Name:** Feature Engineering And Model Selection

Semester: Seventh Semester **Max. Marks:** 50 **Weightage:** 50%

CO - Levels	CO1	CO2
Marks	46	44

PART A: Answer Following Questions. 10 M

Q.N.	Questions	M	CO	BT
1	Define feature engineering. State any one role of features in machine learning.	2	CO1	BT2
2	What is data profiling? Mention any two data quality issues.	2	CO1	BT2
3	Differentiate between standardization and Min-Max normalization.	2	CO1	BT2
4	What is feature extraction? Why is it important in machine learning?	2	CO2	BT1
5	What is TF-IDF? Mention the two main components used in its calculation.	2	CO2	BT2

PART B : Answer Any 1 Following Questions. 10 M

Q.N.	Questions	M	CO	BT
6	Explain the complete data cleaning and preprocessing workflow. Discuss different types of data quality issues and appropriate strategies to address them.	10	CO1	BT3
7	A dataset contains missing values in numerical and categorical attributes. Explain different missing value imputation techniques and demonstrate how you would select an appropriate technique for each type of attribute.	10	CO1	BT3

PART C: Answer Any 1 Following Questions. 10 M

Q.N.	Questions	M	CO	BT
8	Explain the statistical methods for outlier detection. Demonstrate the IQR-based method with a suitable numerical example and explain how detected outliers can be handled.	10	CO1	BT3
9	Explain log transformation and power transformation. Using a suitable example, show how data transformation can help reduce skewness.	10	CO1	BT3

PART D: Answer Any 1 Following Questions. 10 M

Q.N.	Questions	M	CO	BT
10	Explain numerical feature construction techniques such as aggregation, binning, and mathematical transformations. Illustrate each technique with a suitable example.	10	CO2	BT3
11	Explain frequency encoding and count encoding for categorical variables. Using a suitable dataset example, demonstrate how categorical values can be converted into numerical features.	10	CO2	BT3

PART E: Answer Any 1 Following Questions. 10 M				
Q.N.	**Questions**	**M**	**CO**	**BT**
12	Explain the Bag of Words (BoW) approach for text feature extraction. Construct a vocabulary and demonstrate how a set of documents can be represented numerically using BoW.	10	CO2	BT3
13	Explain the fundamentals of image feature extraction. Discuss color, texture, and shape features and explain how numerical descriptors can be constructed from an image.	10	CO2	BT2