PRESIDENCY UNIVERSITY
BENGALURU

Department of Research & Development
Mid - Term Examinations - SEPTEMBER 2024

Odd Semester: Ph.D. Course Work	**Date**: 28 /09/2024
Course Code: CSE844	**Time**: 2:00pm – 3:30pm
Course Name: Data Mining	**Max Marks**: 50
Department: SOCSE	**Weightage**: 25%

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 4 marks.		4Qx5M=20M
1	List and briefly describe the main stages of the Data Mining process	**5 Marks**
2	Explain the importance of data quality in Data Mining	**5 Marks**
3	What is the difference between candidate generation and pruning in the Apriori Algorithm?	**5 Marks**
4	Briefly describe the significance of association rules in Data Mining.	**5 Marks**

Part B

Answer ALL Questions. Each question carries 15 marks.		2QX15M=30M
5	What are similarity and dissimilarity measures, and why are they important in Data Mining?	**15 Marks**
6	Discuss the challenges in mining frequent patterns and how the FP-Growth algorithm addresses these challenges	**15 Marks**