

Instructions:

- (1) No queries will be answered during the quiz.**
- (2) There will not be any compensatory quiz.**
- (3) The duration of this quiz is 45 minutes (+5 minutes to upload for the online students).**
- (4) Cases of plagiarism and late submissions will be penalized.**

Q1. [10 marks] Briefly discuss how the concept of frequent itemset mining can be used in a machine learning task.

Q2. [10 marks] In the page rank algorithm, we assume that each page (data point) is connected to a subset of pages from the given set in the form of a directed graph. Suppose we are also given a category label associated with each page, from a fixed vocabulary $Y = \{y_1, \dots, y_p\}$. Briefly discuss how the conventional power iteration algorithm can be adapted to include this additional category information while computing page ranks.

Q3. [10 marks] Share three conceptual learnings from assignment-1 (on implementation and analysis of the BFR and CURE algorithms). Also discuss one technical reason (as per your understanding) for why these algorithms are generally not adopted for clustering while dealing with particularly small-scale datasets over other popular clustering algorithms such as k-means and agglomerative clustering. (Note: The reason should be other than that of the overall simplicity of the latter algorithms.)