Roll No. | | | | | | | | | | | |

PRESIDENCY UNIVERSITY
BENGALURU

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

Odd Semester: Ph.D. Course Work	**Date**: 27 /09/2024
Course Code: **CSE881**	**Time**: 10:00am – 11:30am
Course Name: Mining the Social Web Data- Facebook, Twitter, LinkedIn, Instagram	**Max Marks**: 50
Department: SOCSE&IS	**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 5 marks.		4Qx5M=20M
1	Explain the difference between structured and unstructured data. Provide examples of both types and discuss how each is stored, managed, and processed in data systems.	**5 Marks**
2	Describe various methods and techniques used to analyze structured data. Discuss how tools like SQL, data warehousing, and machine learning models help in deriving insights from structured data. Include examples of specific techniques such as regression analysis, classification, and clustering.	**5 Marks**
3	Explain the significance of Python in data collection and analysis. Discuss how Python's syntax and libraries make it an effective language for automating tasks like data extraction and analysis from various sources such as Twitter and YouTube.	**5 Marks**
4	Discuss the methods used to collect data from social media platforms like Twitter and YouTube using Python. Explain the role of APIs, and describe how Python can interact with the Twitter API and YouTube API to retrieve specific data (e.g., tweets, video statistics).	**5 Marks**

Part B

Answer ALL Questions. Each question carries 15 marks.		2QX15M=30M
5	Identify and explain different visualization techniques used for representing data in an exploratory analysis and final presentation. Discuss how visualizations like	**15 Marks**

	histograms, scatter plots, heatmaps, and dashboards can help reveal patterns and trends in the data.	
6	Explain the role of Python packages in data collection and analysis. Discuss the steps involved in preparing tools for data extraction and analysis, including installing necessary packages such as tweepy for Twitter or google-api-python-client for YouTube, and explain their functions in the process.	**15 Marks**