

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

School of Computing Science and Engineering

BCA with Industry Oriented Specialization in Data Analytics

Mid Term Examination - Nov 2023

Duration : 90 Minutes

Max Marks : 50

Sem III - E1UE301B - Business Data AnalyticsGeneral Instructions*Answer to the specific question asked**Draw neat, labelled diagrams wherever necessary**Approved data hand books are allowed subject to verification by the Invigilator*

- 1) What are the key steps involved in data preprocessing? K2 (2)
- 2) Explain the difference between structured and unstructured data, and provide examples of each? K1 (3)
- 3) Evaluate the importance of data visualization in data mining. How does it contribute to the decision-making process? K2 (4)
- 4) Explain the concept of a decision tree in classification? K2 (6)
- 5) Describe the K-means clustering algorithm and its basic steps. How does it determine cluster centroids? K3 (6)
- 6) What is Principal Component Analysis (PCA), and how does it contribute to dimensionality reduction in data analysis? What are the steps involved in performing PCA on a dataset K3 (9)
- 7) What is a big data framework, and what is its primary purpose in the context of big data analytics? Provide examples of popular big data frameworks used in the industry. K4 (8)
- 8) Evaluate the impact of big data frameworks on reducing the time and cost of data processing in various industries. K4 (12)

OR

Define data ethics and explain its importance in the context of data collection, storage, and analysis. K4 (12)