K. K. Wagh Institute of Engineering Education & Research, Nashik

(An Autonomous Institute From A.Y. 2022-23)

WINTER-2025	
Exam Seat No.:	
Academic Year:2025-2026	Semester:V
Class:TY	Program:B.Tech
Branch Code:ETC	Pattern:2023
Name of Course:Data Science using Python	Course Code:2302309B
Max. Marks:60	Duration:2.30 Hrs.

Instructions: Candidates should read carefully the instructions printed on the Question Paper and on the cover page of the Answer Book, which is provided for their use.

1. This question paper contains 02 pages.
2. Answer to each new question is to be started on a new page.
3. Assume suitable data wherever required, but justify it.
4. Draw the neat labelled diagrams, wherever necessary.
5. The last columns indicates the Course Outcome and level of Blooms Taxonomy of the Question/sub-question.

		Marks	CO
Question No. 1			
1a)	Write a Python program with Pandas to import a CSV, display summary statistics, and detect missing values in the data.	(6)	CO1
Question No. 2			
2a)	For a distribution Karl Pearson's coefficient of skewness is 0.64, standard deviation is 13 and mean is 59.2 Find mode and median. Illustrate the same using Python.	(6)	CO2
Question No. 3			
3a)	Explain the different stages of data preprocessing. Why is it considered critical?	(8)	CO3
	OR		
3b)	Discuss challenges in real-world data and methods to handle noisy, inconsistent, and duplicate data.	(8)	CO3
3c)	Describe the role of EDA in preprocessing. How does visualization help?	(8)	CO3
	OR		
3d)	Compare Mean/Median, KNN, and Regression-based Imputation.	(8)	CO3
Question No. 4			
4a)	Explain the architecture and key components of TensorFlow. How does it manage computational graphs and automatic differentiation?	(8)	CO4
	OR		
4b)	Write a Python program using **Keras Sequential API** to create a simple feed forward neural network for binary classification. Explain each layer used in the model.	(8)	CO4
4c)	Compare TensorFlow, Keras, and PyTorch	(8)	CO4
	OR		
4d)	Demonstrate with code how to define, compile, and train a neural network in **PyTorch** for classifying handwritten digits (MNIST dataset).	(8)	CO4

Question No. 5

5a)	Explain the process of ECG signal classification using Machine Learning in MATLAB. Outline steps for data acquisition, feature extraction (e.g., R-R interval), training classifiers, and accuracy evaluation.	(8)	CO5

OR

5b)	Write MATLAB code for Image Segmentation using K-Means Clustering. Explain how pixel intensity and color features are used for segmentation and visualize the output before and after clustering.	(8)	CO5
5c)	Explain the step-by-step process of building a **Sentiment Analysis model using LSTM** in Python. Describe the stages of text preprocessing, model creation, and evaluation.	(8)	CO5

OR

5d)	Write and explain Python code using **Keras** to preprocess textual data for sentiment analysis — include to kenization, padding, and vocabulary creation.	(8)	CO5

····· End of question paper·····