EPCET LIBRARY BANGALORE-49

CBCS SCHEME

USN | | | | | | | | | | |

21AD62

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2026
Data Science and its Applications

Time: 3 hrs. Max. Marks: 100

Note: ***Answer any FIVE full questions, choosing ONE full question from each module.***

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg, 42+8 = 50, will be treated as malpractice.

Module-1

1 a. Create a Python code that calculate the Euclidean distance between two vectors V = [2, 3, 4] and W = [1, 1, 1] and describe the procedures in brief. **(10 Marks)**

b. Compare and contrast different measures of central tendency in statistics in terms of computational methods, sensitivity to data changes and suitability for handling outlines with example. **(10 Marks)**

OR

2 a. Show how to plot and compare multiple probability density functions with a normal distribution using matplotlib code : Interpret the plotted graphs interms of changing μ and σ. **(10 Marks)**

b. Discuss how the Central limit theorem applies to binomial random variable. Provide an explanation of how a Prinomial (n, p) random variable approximates to a normal distribution as n becomes large. **(10 Marks)**

Module-2

3 a. Describe the methodology to determine a symmetric interval around the mean of a distribution that captures specified probability such as 60% in the central region by identifying the cutoff points and the significance level at 5% or 1%. **(10 Marks)**

b. Describe the concept of gradient descent in the context of optimizing model parameters. How does gradient descent iteratively adjust the parameters to minimize the loss function? **(10 Marks)**

OR

4 a. Illustrate tqdm library by considering an example to find all prime numbers up to a given integer. **(10 Marks)**

b. Explain cleaning and managing of data with an example. **(10 Marks)**

Module-3

5 a. Discuss bias – variance tradeoff in detail. **(10 Marks)**

b. Describe the implementation procedure of Naïve Bayes theorem to model a sophisticated spam filter. **(10 Marks)**

OR

6 a. Write Python code to apply the logistic regression model and check of goodness of fit. **(10 Marks)**

b. Explain the concept of support vector machines in detail. **(10 Marks)**

Module-4

7 a. Explain feed – forward neural networks with a neat sketch. **(10 Marks)**

b. Write a python code to compute softmaxes and cross-entropy in deep learning. **(10 Marks)**

OR

8 a. Construct linear layers with implementation in python. **(10 Marks)**

b. Discuss the underlying principles of bottom – up hierarchical clustering. **(10 Marks)**

Module-5

9 a. What are n – gram models? Demonstrate the implementation of bigram and trigram models in python. **(10 Marks)**

b. Explain Gibbs sampling and its application in generating samples form multidimensional distribution. **(10 Marks)**

OR

10 a. Differentiate between Centrality and Closeness centrality in the context of social network analysis. **(10 Marks)**

b. Write short notes on :

i) User based collaborative filtering.

ii) Item – based collaborative filtering. **(10 Marks)**

* * * * *

VTU-01-06-2026 08:46:30am - EP - EP - EP