

December 2023

B.Tech (CSE)

B.Tech VII SEMESTER

## Data Analytics (PEC-CS-D-702(I))

Time: 3 Hours

Max. Marks:75

- Instructions:**
1. It is compulsory to answer all the questions (1.5 marks each) of Part -A in short.
  2. Answer any four questions from Part -B in detail.
  3. Different sub-parts of a question are to be attempted adjacent to each other.

PART -A

- Q1 (a) Differentiate between Descriptive and Inferential Statistics. (1.5)
- (b) What is Data Validation? (1.5)
- (c) Explain dimensionality reduction. (1.5)
- (d) What is overfitting? (1.5)
- (e) What is CART used for? (1.5)
- (f) Differentiate between Regression and Classification. (1.5)
- (g) Advantages of KNN algorithm. (1.5)
- (h) What is Active Learning? (1.5)
- (i) What is hypothesis testing? (1.5)
- (j) Applications of Data Analytics. (1.5)

PART -B

- Q2 (a) Differentiate between Univariate , Bivariate and Multivariate Analysis (5)
- (b) What is Data Analytics? Explain the challenges associated with Big Data Analytics. (10)

- Q3 (a) How does regression help in data analytics? Formulate the linear regression equation using the given data : (15)

X	2	4	6	8
Y	3	7	5	10

- Q4 (a) How does OLS regression helps to minimize error in Regression? (5)
- (b) What is dispersion in Data Analytics? Explain various measures of dispersion. (10)

- Q5 (a) What is Regularization? Explain regularization techniques. (5)
- (b) Explain ways to validate a model. (10)

- Q6 (a) What are ensemble methods? Explain their need and types. (10)
- (b) Explain Linear Discriminant Analysis. (5)

- Q7 (a) Elaborate the bias variance dichotomy. (10)
- (b) How does Reinforcement Learning contribute to data analytics? (5)