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

18AD62

Sixth Semester B.E. Degree Examination, June/July 2023
Data science and its Applications

Time: 3 hrs.

Max. Marks: 100

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

Module-1

1 a. What is data science? How can data science contribute to solving real-world problems? **(06 Marks)**

b. What are the different types of data visualizations that can be created using mat plot lib? **(06 Marks)**

c. Explain Bayes's theorem and its applications in data science. **(08 Marks)**

OR

2 a. Explain the concept of linear algebra and its significance in data science. **(06 Marks)**

b. Explain Simpson's paradox and provide an example where it might occur. **(06 Marks)**

c. What is the central limit theorem and why is it important in statistical inference? **(08 Marks)**

Module-2

3 a. What is the difference between hypothesis and inference in statistics? **(06 Marks)**

b. Discuss the concept of p-hacking and its implications in statistical analysis. **(06 Marks)**

c. Describe techniques for cleaning, muning and manipulating data. **(08 Marks)**

OR

4 a. Explain the concept of statistical hypothesis testing and its steps. **(06 Marks)**

b. How can APIs be used to access and retrieve data for analysis? **(06 Marks)**

c. Describe the idea behind estimating the gradient in gradient descent. **(08 Marks)**

Module-3

5 a. What is machine learning and how does it differ from traditional programming? **(06 Marks)**

b. How does the K-nearest neighbors algorithm work in classification tasks? **(06 Marks)**

c. Explain the concept of simple linear regression and its components. **(08 Marks)**

OR

6 a. Explain the concepts of overfitting and underfiiting in machine learning. **(06 Marks)**

b. What are support vector machines and how do they work in classification tasks? **(06 Marks)**

c. Describe multiple regression and how it extends simple linear regression. **(08 Marks)**

Module-4

7 a. What is a decision tree and how does it work in machine learning? **(06 Marks)**

b. Explain the architecture and working principles of feed – forward neural networks. **(06 Marks)**

c. What is deep learning and how does it differ from traditional neural networks. **(08 Marks)**

OR

8 a. How do random forests work and what are their advantages in machine learning. **(06 Marks)**

b. Explain K-Means clustering. **(06 Marks)**

c. Explain the back propagation algorithm and its role in training neural networks. **(08 Marks)**

<u>Module-5</u>

9 a. What is natural language processing (NLP) and how is it used in machine leaning. **(06 Marks)**
   b. What is Gibbs sampling and how is it used in probabilistic models and text analysis.
   (06 Marks)
   c. Explain Recurrent Neural Networks (RNNs) and their applications in sequential data processing. **(08 Marks)**

OR

10 a. Explain directed graphs and page rank algorithms. **(06 Marks)**
   b. Explain the concept of Manual Curation. **(06 Marks)**
   c. Explain item – Based collaborative filtering and its approach. **(08 Marks)**

* * * * *