M.Sc. DEGREE EXAMINATION,NOVEMBER 2019
I Year II Semester
Data Science using Python

Time : 3 Hours **Max.marks :75**

Section A $(10 \times 2 = 20)$ Marks

Answer any **TEN** questions

1. What is the main goal of python programming?
2. State any two features of python.
3. How to sample the data in python?
4. Write the steps to insert data frames values into the table.
5. What is a subplot?
6. When to use scatter plots?
7. What are the two components of dimension reduction?
8. Define variance and covariance.
9. How to cross validate a model?
10. What is variable transformation?
11. What is feature elimination?
12. When to use data aggregation?

Section B $(5 \times 5 = 25)$ Marks

Answer any **FIVE** questions

13. Why is python programming language important in data science?
14. List and explain the commonly used data structure in pandas?
15. State the standard way of displaying the distribution of data based on the five number summary.
16. Discuss about the types of outliers.
17. How to implement stochastic gradient descent?
18. How to use cluster analysis for spotting outliers?
19. Estimate the probability density function using Gaussian distribution.

Section C $(3 \times 10 = 30)$ Marks

Answer any **THREE** questions

20. What is Anaconda?Write the various steps for installing and managing the packages with anaconda?
21. Discuss briefly about slicing and dicing in manipulating binary data.
22. What are the various commonly used measures of descriptive statistics?
23. What is classification and explain its types?
24. Explain the Support Vector Machine module of learning algorithms?

M.Sc. DEGREE EXAMINATION,NOVEMBER 2019
I Year II Semester
Data Science using Python

Time : 3 Hours **Max.marks :75**

Section A $(10 \times 2 = 20)$ Marks

Answer any **TEN** questions

1. What is the main goal of python programming?
2. State any two features of python.
3. How to sample the data in python?
4. Write the steps to insert data frames values into the table.
5. What is a subplot?
6. When to use scatter plots?
7. What are the two components of dimension reduction?
8. Define variance and covariance.
9. How to cross validate a model?
10. What is variable transformation?
11. What is feature elimination?
12. When to use data aggregation?

Section B $(5 \times 5 = 25)$ Marks

Answer any **FIVE** questions

13. Why is python programming language important in data science?
14. List and explain the commonly used data structure in pandas?
15. State the standard way of displaying the distribution of data based on the five number summary.
16. Discuss about the types of outliers.
17. How to implement stochastic gradient descent?
18. How to use cluster analysis for spotting outliers?
19. Estimate the probability density function using Gaussian distribution.

Section C $(3 \times 10 = 30)$ Marks

Answer any **THREE** questions

20. What is Anaconda?Write the various steps for installing and managing the packages with anaconda?
21. Discuss briefly about slicing and dicing in manipulating binary data.
22. What are the various commonly used measures of descriptive statistics?
23. What is classification and explain its types?
24. Explain the Support Vector Machine module of learning algorithms?