M.Sc. DEGREE EXAMINATION,APRIL 2020
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. Why to learn python for data science?
2. Name any two used libraries for data science.
3. What is the use pymongo package in python?
4. How do extract the range based subsets in classification?
5. What does figure method describes in matplotlib?
6. Write the use of numpy library in python.
7. What is the result of pip command in python?
8. Analyse the disadvantage of k means clustering algorithm.
9. Define the term loocv.
10. Write about Ridge regression.
11. Differentiate between loc and iloc in indexing structure of python.
12. What is the use of seaborn in python?

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

Answer any **FIVE** questions

13. State and explain python IDEs which are used for data science.
14. Aggregate sum and min function across all the columns in data frame using a sample dataset.
15. Discuss about plot manipulation with detailed examples.
16. Differentiate flat and hierarchical clustering with examples.
17. Discuss about Variable transformation with example.
18. Explain the concept of detecting outliers in data.
19. When can we use the principle component analysis and how to use?

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

Answer any **THREE** questions

20. Python ecosystem is growing and may become dominant platform for machine learning. Justify.
21. How to handle categorical data in python and use the pre-processing technique to build machine learning models?
22. How do you use data exploratory for data analysis?
23. Differentiate logistic and linear regression with examples.
24. Briefly explain the concept and implementation of Support vector machine in python.

M.Sc. DEGREE EXAMINATION,APRIL 2020
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. Why to learn python for data science?
2. Name any two used libraries for data science.
3. What is the use pymongo package in python?
4. How do extract the range based subsets in classification?
5. What does figure method describes in matplotlib?
6. Write the use of numpy library in python.
7. What is the result of pip command in python?
8. Analyse the disadvantage of k means clustering algorithm.
9. Define the term loocv.
10. Write about Ridge regression.
11. Differentiate between loc and iloc in indexing structure of python.
12. What is the use of seaborn in python?

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

Answer any **FIVE** questions

13. State and explain python IDEs which are used for data science.
14. Aggregate sum and min function across all the columns in data frame using a sample dataset.
15. Discuss about plot manipulation with detailed examples.
16. Differentiate flat and hierarchical clustering with examples.
17. Discuss about Variable transformation with example.
18. Explain the concept of detecting outliers in data.
19. When can we use the principle component analysis and how to use?

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

Answer any **THREE** questions

20. Python ecosystem is growing and may become dominant platform for machine learning. Justify.
21. How to handle categorical data in python and use the pre-processing technique to build machine learning models?
22. How do you use data exploratory for data analysis?
23. Differentiate logistic and linear regression with examples.
24. Briefly explain the concept and implementation of Support vector machine in python.