(assume the documents as per your own). (use single command)

iv) Retrieve all documents having author "Suryakumar Dutta" and title "Introduction to Maching learning models".

v) Set the new title "Introduction to JSP and Servlet" of the documents whose earlier title was "Java : overview".

5. Write the following statements in Python :   5x2=10

    i)   Declare a varible "my_list" with number 1 to 7.

    ii) Append a new number 8 in the "my_list".

    iii) Create an array with the numbers : 10,20,30,40 in the variable "my_arr".

    iv) Display the "my_list" using series.

    v)   Import Sklearn & import Train-test dataset.

6. What is Apache Hadoop ? What are main modules of Hadoop ? Discuss the master/slave architecture in HDFS.

                                                    2+3+5=10

7. Write short notes on any *two* of the following :   2x5=10

    (a) Supervised learning

    (b) Clustering and Association Rule learning

    (c) Hive and Pig

    (d) Seaborn and Matplotlib in Python.

MASTER OF LIBRARY AND INFORMATION SCIENCE

(DIGITAL LIBRARY) EXAMINATION, 2023

(2nd Year, 1st Semester)

BIG DATA ANALYTICS

Course : MLDL-11E

Time : Two Hours   Full Marks : 50

The Figures in the margin indicate full marks.

Answer any five questions from the following

1. Define the terms "Data Science" and "Big data". Make a comparative study between "data analysis" and "data analytics".   3+7=10

2. What is "Machine learning" ? State the difference between machine learning and deep learning. How does machine learning work ?   2+3+5=10

3. What is "NoSQL" database ? How does it differ from RDBMS ? Discuss the different types of NoSQL ?

                                                    2+3+5=10

4. Write the following commands of MongoDB :   5x2=10

    i)   Create a database "mldl".

    ii) Create a collection "articles".

    iii) Insert three documents in the collection "article" with the parameters : title, author, abstract, url, keywords and likes.