Register Number:……..

KUMARAGURU
college of technology
character is life

B.E / B.TECH DEGREE EXAMINATIONS: NOV/DEC 2023

(Regulation 2018)

Third Semester

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

U18AII3205**:** Data Collection and Data Management

COURSE OUTCOMES

CO1: Explain basic database concepts, applications, data models, schemas, and instances.

CO2: Understand the concept of handling unstructured data.

CO3: Explain the various data collection methodologies such as map, filter, and List comprehension

CO4: Apply mapreduce in real-world applications.

Time: Three Hours **Maximum Marks: 100**

Answer all the Questions:-

PART A (10 x 2 = 20 Marks)

(Answer not more than 40 words)

1.	Differentiate data models, schemas, and instances?	CO1	[K_1]
2.	Construct a simple ER diagram for a hospital management system, including entities like "Patient," "Doctor," "Nurse," "Department," and "Appointment," also establish relationships between them.	CO1	[K_3]
3.	Write the steps involved in converting the ER model to a relational schema.	CO1	[K_2]
4.	Identify the challenges of handling unstructured data in a database system.	CO2	[K_2]
5.	Convert the given XML document to its corresponding JSON format.	CO2	[K_2]

```
<employees>
  <employee>
    <id>101</id>
    <name>John Doe</name>
    <position>Software Engineer</position>
  </employee>
  <employee>
    <id>102</id>
    <name>Jane Smith</name>
    <position>Project Manager</position>
  </employee></employees>
```

6.	Using map, create a list assigned to the variable squared_numbers that squares each element in the given list. numbers= [1,2,3,4,5]		CO3	[K_2]
7.	Fill the blank in the given Python program using the filter to create a list assigned to the variable positive_numbers that filters out negative numbers from the list numbers. def keep_positives(numbers): _____#Fill the code here print(keep_positives([-2, 5, -8, 10, -1, 3]))		CO3	[K_2]
8.	Write a Python list comprehension to create a list assigned to the variable squared_evens that squares each even number from the list numbers. numbers = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] [Hint: output [1, 4, 9, 16, 25]]		CO3	[K_2]
9.	What is the role of data centers in cloud computing?		CO4	[K_1]
10.	How does MapReduce handle algorithmic tasks in distributed computing?		CO4	[K_1]

Answer any FIVE Questions: -
PART B (5 x 16 = 80 Marks)
(Answer not more than 400 words)

11.	a)	Explain the purpose of database systems and discuss the key principles of Relational Database Management Systems (RDBMS). Illustrate with examples.	8	CO1	[K_1]
	b)	Discuss the handling of unstructured data in database systems, give examples of scenarios where unstructured data is encountered, and explain strategies to manage such data effectively.	8	CO2	[K_1]
12.	a)	Explain the Database Development Life Cycle (DBLC) and its stages. Elaborate on the importance of each stage in the successful creation and maintenance of a database system.	8	CO1	[K_1]
	b)	Illustrate the concept of scoring, weighting, and vector space in information retrieval and explain their roles in improving the accuracy of search results with examples.	8	CO2	[K_2]
13.	a)	Compare and contrast MongoDB with traditional relational databases with examples.	8	CO2	[K_2]

	b)	Consider the given nested data structure given below. nested = {'data': ['finding', 23, ['exercises', 'hangout', 34]], 'window': ['part', 'whole', [], 'sum', ['math', 'calculus', 'algebra', 'geometry', 'statistics',['physics', 'chemistry', 'biology']]]} Write the python code to check the following conditions. 1. Check to see if the string 'data' is a key in nested, if it is, assign True to the variable data, otherwise assign False. 2. Check to see if the integer 24 is in the value of the key data, if it is then assign to the variable twentyfour the value of True, otherwise False. 3. Check to see that the string 'whole' is not in the value of the key window. If it's not, then assign to the variable whole the value of True, otherwise False. 4. Check to see if the string 'physics' is a key in the dictionary nested. If it is, assign to the variable physics, the value of True, otherwise False. Print the final output.	8	CO3	[K2]
14.	a)	Analyze the role of web crawling and web APIs in gathering unstructured data with the workings of how web crawling is implemented.	8	CO2	[K3]
	b)	Compare and contrast the functionalities of the Map and Filter functions in Python with examples illustrating their applications in data processing.	8	CO3	[K2]
15.	a)	Describe the role of the Filter function in Python and write a code example using the filter function to extract even numbers from a list.	8	CO3	[K2]
	b)	Explain the concept of Cloud Computing and its relationship to MapReduce and how MapReduce utilizes cloud resources.	8	CO4	[K2]
16.	a)	Explain the concept of Shallow Copies and Deep Copies in Python lists and Provide examples illustrating the differences between them.	8	CO3	[K2]
	b)	Implement a MapReduce algorithm in Python to process a dataset. Discuss the steps involved and demonstrate the application using a practical example.	8	CO4	[K3]

****