Register Number:……..

KUMARAGURU
college of technology
character is life

ENTER **B.E/B.TECH DEGREE EXAMINATIONS: NOV/DEC 2022**

(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 concepts 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.	Infer ETL.	CO1	[K_2]
2.	If R={W, X,Y,Z } and Functional dependencies are F={ XYZ → W, XY → ZW and X → YZW} List all the candidate keys.	CO1	[K_4]
3.	List some of the Web Crawling tools used by you.	CO1	[K_1]
4.	<?xml version="1.0" encoding="ISO-8859-1"?> <note> <to>Elon Musk</to> <from>John Walker</from> <heading>Reminder</heading> <body>Gentle reminder for Interview slot</body> </note> Represent the above XML code into JSON format	CO2	[K_4]
5.	Summarize Information retrieval	CO2	[K_2]
6.	Inspect usage of map after filter.	CO3	[K_4]
7.	Distinguish Deep copy from Shallow copy.	CO3	[K_4]
8.	Interpret different types of Data Center	CO3	[K_1]
9.	Rephrase Communication Cost Model	CO4	[K_2]
10.	Criticize the challenges faced by data in the cloud	CO4	[K_5]

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

11.	a)	Outline the Architecture of DBMS	8	CO1	[K_2]
	b)	Demonstrate your understanding of normalization. With suitable example justify different forms of normalization.	8	CO1	[K_2]
12.	a)	Drawing of ER model of university database application considering the constraints • A university has many departments. • Each department has multiple instructors (one person is HOD). Here the HOD refers to the head of department. • An instructor belongs to only one department. • Each department offers multiple courses, each subject is taught by a single instructor. • A student may enroll for many courses offered by different departments.	16	CO2	[K_3]
13.	a)	Define the following methods in MangoDB 1. update() 2. Save() 3. insert() 4. find()	10	CO2	[K_4]
	b)	Infer some of the web APIs and its usage	6	CO2	[K_2]
14.	a)	Analyze Nested Data and Nested Iterations	8	CO3	[K_4]
	b)	Demonstrate your understanding of List Comprehension	8	CO3	[K_2]
15.	a)	Examine HADOOP Ecosystem.	8	CO4	[K_5]
	b)	Illustrate Regular expression	8	CO4	[K_2]
16.	a)	All of us are familiar with the disaster that happened on April 14, 1912. The big giant ship of 46000-ton in weight got sink-down to the depth of 13,000 feet in the North Atlantic Ocean. Analyze the Titanic Disaster dataset and find the average	16	CO4	[K_4]

Please indicate knowledge level (K_1toK_6) and Course Outcome level (CO1 to CO5) against each question for each subdivision.

age of male and female persons died in this disaster with MapReduce Hadoop. Need proper Analysis with sample code snippet.

[Titanic DataSet for your reference

Passanger id, Survived(0/1), Pclass, Name, Sex, Age,SibSp, Parch, Ticket, Fare, Cabin, Embarked

1,0,3,"Braund Mr. Owen Harris",male,22,1,0,A/5 21171,7.25,,S,

2,1,1,"Cumings Mrs. John Bradley (Florence Briggs Thayer)",female,38,1,0,PC 17599,71.2833,C85,C,

3,1,3,"Heikkinen Miss. Laina",female,26,0,0,STON/O2. 3101282,7.925,,S,

4,1,1,"Futrelle Mrs. Jacques Heath (Lily May Peel)",female,35,1,0,113803,53.1,C123,S,

5,0,3,"Allen Mr. William Henry",male,35,0,0,373450,8.05,,S,

6,0,3,"Moran Mr. James",male,,0,0,330877,8.4583,,Q,

7,0,1,"McCarthy Mr. Timothy J",male,54,0,0,17463,51.8625,E46,S,

8,0,3,"Palsson Master. Gosta Leonard",male,2,3,1,349909,21.075,,S,

9,1,3,"Johnson Mrs. Oscar W (Elisabeth Vilhelmina Berg)",female,27,0,2,347742,11.1333,,S,

10,1,2,"Nasser Mrs. Nicholas (Adele Achem)",female,14,1,0,237736,30.0708,,C,

11,1,3,"Sandstrom Miss. Marguerite Rut",female,4,1,1,PP 9549,16.7,G6,S,

12,1,1,"Bonnell Miss. Elizabeth",female,58,0,0,113783,26.55,C103,S,

]

****

Please indicate knowledge level (K_1toK_6) and Course Outcome level (CO1 to CO5) against each question for each subdivision.