Subject Code: R23xxxx

R23

KALLAM HARANADHAREDDY INSTITUTE OF TECHNOLOGY (AUTONOMOUS)

II B. Tech I Semester (R23)

COURSE NAME

(Common to)

QUESTION BANK

Course Advisor/Moderator: Dr. B Tarakeswara Rao

Q. No.			CO	BTL	Total Marks
UNIT-I Short Answer Questions					
1		Define Data Science. Write components of Data Science.	CO1	1	2M
2		What are the advantages and disadvantages of Data Science.	CO1	1	2M
3		What are the applications of Data Science.	CO1	1	2M
4		Explain the Structured data in facets.	CO1	2	2M
5		Explain Natural Language in facets.	CO1	2	2M
6		Build the neat diagram of Data Science Process.	CO1	3	2M
7		Explain NoSQL Databases.	CO1	2	2M
8		Explain External Data	CO1	2	2M
9		Explain data cleansing.	CO1	2	2M
10		Explain Transforming data.	CO1	2	2M
UNIT-I Long Answer Questions					
1	i	Define Data Science. What are the applications of Data Science?	CO1	1	5M
	ii	Explain various components in Data Science.	CO1	2	5M
2.	i	List out the various jobs/data scientist roles in Data Science	CO1	1	5M
	ii	What are the benefits and uses of Data Science	CO1	1	5M
3	i	Explain various facets of data	CO1	2	5M
	ii	Explain Data Science process with neat diagram.	CO1	2	5M
4	i	Explain Big Data ecosystem and Data Science.	CO1	2	5M
	ii	Explain the overview of Data Science process with neat diagram.	CO1	2	5M
5	i	Explain data cleansing, data integration and data transformation.	CO1	2	5M
	ii	Explain How to presenting findings and building applications on top of them in data science.	CO1	2	5M

Q. No.			CO	BTL	Total Marks
UNIT-II Short Answer Questions					
1		Define Machine Learning. Write types of Machine Learnings.	CO2	1	2M
2		What are the differences between Data Science and Machine Learning.	CO2	1	2M
3		List Applications for Machine Learning in Data Science.	CO2	1	2M
4		Explain Exploratory Data Analysis.	CO2	2	2M
5		Explain Model Validation.	CO2	2	2M
6		Define Classification. Write four classification algorithms.	CO2	1	2M
7		What are the advantages and disadvantages of Unsupervised Machine Learning	CO2	1	2M
8		Explain Mapreduce.	CO2	2	2M
9		Explain Sparse Data.	CO2	2	2M
10		Explain How to reduce your computing needs.	CO2	2	2M
UNIT-II Long Answer Questions					
1	i	Explain various Applications for machine learning in data science.	CO2	2	5M
	ii	Explain the role of machine learning in Data Science	CO2	2	5M
2.	i	Where machine learning is used in the data science process and explain.	CO2	1	5M
	ii	Explain various Python tools used in machine learning.	CO2	2	5M
3	i	Explain various modeling phases in Modeling Process.	CO2	2	5M
	ii	Explain Semi-Supervised Machine Learning.	CO2	2	5M
4	i	What are the problems you face when handling large data	CO2	1	5M
	ii	Explain various General techniques for handling large volumes of data	CO2	2	5M
5	i	Explain General programming tips for dealing with large data sets	CO2	2	5M
	ii	Explain Case studies on ds projects for predicting malicious urls for building recommender systems	CO2	2	5M

Q. No.			CO	BTL	Total Marks
UNIT-III Short Answer Questions					
1		List out the different components of Hadoop	CO3	1	2M
2		Explain how Hadoop achieves parallelism	CO3	2	2M
3		How does Spark solve the problems of MapReduce?	CO3	1	2M

4		Explain the different components of the Spark ecosystem	CO3	2	2M
5		List out the common Hadoop file system commands	CO3	1	2M
6		Build the neat diagram of NoSQL and NewSQL databases	CO3	3	2M
7		Explain the differences between ACID and BASE	CO3	2	2M
8		Explain graph databases	CO3	2	2M
9		List out the biggest types of NoSQL databases	CO3	1	2M
10		Explain Wide-Column databases	CO3	2	2M
UNIT-III Long Answer Questions					
1	i	Explain a frame work for storing and processing large data sets	CO3	2	5M
	ii	Explain how to replacing MapReduce for better reference in Spark	CO3	2	5M
2.	i	Summarize a case study on Assessing risk when loaning money	CO3	2	10M
3	i	Explain the core principle of relational databases in ACID	CO3	2	5M
	ii	Explain the problem with DBs on many nodes in CAP theorem	CO3	2	5M
4	i	Explain the BASE principles of NoSQL databases	CO3	2	5M
	ii	Explain types of NoSQL databases	CO3	2	5M
5	i	Explain the case study on what disease is that?	CO3	2	10M

Q. No.			CO	BTL	Total Marks
UNIT-IV Short Answer Questions					
1		Define entities with a neat diagram	CO4	1	2M
2		Define relationships	CO4	1	2M
3		Explain Properties	CO4	2	2M
4		Define edges	CO4	1	2M
5		List out Text Mining applications	CO4	1	2M
6		Define Stemming	CO4	1	2M
7		Explain Lemmatization	CO4	2	2M
8		Explain Stop Word filtering	CO4	2	2M
9		Explain Tokenization	CO4	2	2M
10		Explain POS Tagging	CO4	2	2M
UNIT-IV Long Answer Questions					
1	i	Explain connected and graph databases with a neat diagram	CO4	2	5M
	ii	Illustrate why and when should we use a graph database	CO4	2	5M
2.	i	Explain a graph database in Neo4j	CO4	2	5M
	ii	Explain a graph query language in Cypher	CO4	2	5M

3	i	Explain connected data example on recipe recommendation engine	CO4	2	5M
	ii	Explain text mining in the real world	CO4	2	5M
4	i	Explain text mining techniques	CO4	2	5M
	ii	Explain the research goal in the data science process overview for Reddit classification	CO4	2	5M
5	i	Explain the Data retrieval in the data science process overview for Reddit classification	CO4	2	5M
	ii	Explain Presentation and automation data science process overview for Reddit classification	CO4	2	5M

Q. No.			CO	BTL	Total Marks
UNIT-V Short Answer Questions					
1		Explain a new viewpoint on your data	CO5	2	2M
2		Explain a real time dashboard	CO5	2	2M
3		Explain JavaScript-based dashboards	CO5	2	2M
4		Define Crossfilters	CO5	1	2M
5		Explain about dc.js	CO5	2	2M
6		List out few reasons against developing your own application	CO5	2	2M
7		Explain about reduceCount() and reduceSum() in MapReduce	CO5	2	2M
8		Discuss about HighCharts	CO5	6	2M
9		Explain about reduceAdd() and reduceRemove()	CO5	2	2M
10		Explain about Prototyping	CO5	2	2M
UNIT-V Long Answer Questions					
1	i	What are the data visualization options and explain in detail	CO5	1	10M
2.	i	How to setting up everything in the JavaScript MapReduce library	CO5	1	10M
3	i	Explain unleashing crossfilters to the medicine data set	CO5	2	10M
4	i	How to create an interactive dashboard with dc,js	CO5	1	10M
5	i	Explain in detail the dashboard development tools	CO5	2	10M

Course Advisor BoS Chairman

****