

Printed Pages:02

University Roll No:

End Term Examination, Odd Semester 2021-22
B.Tech (CSE/ CCV/ DA/ CSF/ IoT), IIIrd Year, Vth Semester

BCSC 0007: Introduction to Microprocessors

Time: 3 Hour

Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

$5 \times 4 = 20$ marks

- (I) What is the difference between CALL and JUMP instruction? Explain all the types of conditional JUMP instructions.
- (II) What are the drawbacks of 8085 Microprocessor as compare to 8086 Microprocessor? Write any four drawbacks.
- (III) Draw the architecture of 8086 microprocessor.
- (IV) Explain all the rotate instructions of 8085 microprocessor instruction set. With suitable diagram.
- (V) Draw the flag register with all flags of 8086 microprocessor. Explain each flag.
- (VI) Calculate the sum of series of numbers. The length of the series is in memory location 2000H and the series begins from memory location 2001H. Consider the sum to be 8 bit numbers. Store the sum at memory location 3000H.

Section- B

Note: Attempt ALL Questions.

$5 \times 3 = 15$ marks

- (I) Explain the function of execution unit (EU) and bus interface unit (BIU) in 8086 microprocessor.
- (II) Write an assembly language program for multiply two 8 bit numbers, Data are stored in memory location 3000H and 3001H. Result store in memory location 4000H.
- (III) Write a program for 10 bytes of data are stored in memory location at 2001H to 200AH. Transfer this entire block of data to new memory location starting at 3001H.

- (IV) Define the concept of pipelining and non-pipelining. What is the difference between stack and queue?
- (V) Explain the functionality of following instructions:
(a) CMP B (b) EI (c) RET

Section- C

Note: Attempt ANY THREE Questions.

$3 \times 5 = 15$ marks

- (I) Explain the concept of memory segmentation in 8086 microprocessor. How the physical address can be calculated?
- (II) Explain the types of addressing modes of 8086 microprocessor. Also give suitable example of each addressing mode.
- (III) Write a program to find the largest number in a given series. Length of the series given in memory location 2000H, Series starts from memory location 2001H to 2010H. Result store in memory location 3000H.
- (IV) Explain the functionality of following instructions:
(a) PUSH PSW (b) POP PSW (c) IN Port address
(d) OUT port address (e) NOP

Ceman
ETrdVH CS, EC

Printed Pages: 22 University Roll No.....

End-Term Examination, Odd Semester 2021-22

B.Tech. CSE and ECE, III / IV: Year, V / VII: Semester

BCSE.0904: Machine Learning for IOT

Time: 1.5-hours Maximum Marks: 50

Section A

Note: Attempt all questions. $30 \times 1 = 30$ Marks

1. A _____ is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility.
 - A. Decision tree
 - B. Graphs
 - C. Trees
 - D. Neural Networks
2. Decision Tree is a display of an algorithm.
 - A. TRUE
 - B. FALSE
3. Which of the following is/are true about bagging trees?
 1. In bagging trees, individual trees are independent of each other.

2. Bagging is the method for improving the performance by aggregating the results of weak learners.
- A. 1
 - B. 2
 - C. 1 and 2
 - D. None of these
4. Data Analysis is a process of:
- A. Inspecting data
 - B. Cleaning data
 - C. Transforming data
 - D. All of Above
5. Which of the following is/are true about boosting trees?
- 1. In boosting trees, individual weak learners are independent of each other
 - 2. It is the method for improving the performance by aggregating the results of weak learners
- A. 1
 - B. 2
 - C. 1 and 2
 - D. None of these
6. Does Facebook uses “Big Data” to perform the concept of Flashback?
- A. TRUE

B. FALSE

7. Which of the following is not a major data analysis approach?

- A. Data Mining
- B. Predictive Intelligence
- C. Business Intelligence
- D. Text Analytics

8. What is Decision Tree?

- A. Flow-Chart
- B. Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label
- C. Flow-Chart & Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label
- D. None of the mentioned

9. Which of the following is/are true about Random Forest and Gradient Boosting ensemble methods?

- 1. Both methods can be used for classification task.
- 2. Random Forest is use for classification whereas Gradient Boosting is use for regression task.
- 3. Random Forest is use for regression whereas Gradient Boosting is use for Classification task.

4. Both methods can be used for regression task.

- A. 1
- B. 2
- C. 3
- D. 1 and 4

10. In Random forest you can generate hundreds of trees (say T1, T2Tn) and then aggregate the results of these tree. Which of the following is true about individual(Tk) tree in Random Forest?

- 1. Individual tree is built on a subset of the features
 - 2. Individual tree is built on all the features
 - 3. Individual tree is built on a subset of observations
 - 4. Individual tree is built on full set of observations
- A. 1 and 3
 - B. 1 and 4
 - C. 2 and 3
 - D. 2 and 4

11. The Process of describing the data that is huge and complex to store and process is known as:

- A. Analytics
- B. Data mining
- C. Big data
- D. Data warehouse

12. Decision Trees can be used for Classification Tasks.
- A. True
 - B. False
13. Which of the following is true about “max_depth” hyperparameter in Gradient Boosting?
- 1. Lower is better parameter in case of same validation accuracy.
 - 2. Higher is better parameter in case of same validation accuracy.
 - 3. Increase the value of max_depth may overfit the data.
 - 4. Increase the value of max_depth may underfit the data.
- A. 1 and 3
 - B. 1 and 4
 - C. 2 and 3
 - D. 2 and 4
14. How many main statistical methodologies are used in data analysis?
- A. 2
 - B. 3
 - C. 4
 - D. 5
15. In descriptive statistics, data from the entire population or a sample is summarized with?

- A. Integer descriptor
- B. floating descriptor
- C. numerical descriptor
- D. decimal descriptor

16. Choose from the following that are Decision Tree nodes?

- A. Decision Nodes
- B. End Nodes
- C. Chance Nodes
- D. All of the mentioned

17. ____ have a structure but cannot be stored in a database.

- A. Structured
- B. Semi Structured
- C. Unstructured
- D. None of these

18. Data generated from online transactions is one of the example for volume of big data.

- A. TRUE
- B. FALSE

19. Files are divided into ____ sized Chunks.

- A. Static
- B. Dynamic

- C. Fixed
- D. Variable

20. Which of the following are the advantage/s of Decision Trees?
- A. Possible Scenarios can be added
 - B. Use a white box model, if given result is provided by a model
 - C. Worst, best and expected values can be determined for different scenarios
 - D. All of the mentioned
21. Which of the following algorithm doesn't uses learning Rate as one of its hyperparameter?
- 1. Gradient Boosting
 - 2. Extra Trees
 - 3. AdaBoost
 - 4. Random Forest
- A. 1 and 3
 - B. 1 and 4
 - C. 2 and 3
 - D. 2 and 4
22. For improving supply chain management to optimize stock management, replenishment, and forecasting:
- A. Descriptive
 - B. Diagnostic

- C. Predictive
- D. Prescriptive

23. _____ as a result of data accessibility, data latency, data availability, or limits on bandwidth in relation to the size of inputs:

- A. Computation-restricted throttling
- B. Large data volumes
- C. Data throttling
- D. Data Parallelization

24. As an example, an expectation of using a recommendation engine would be to increase same-customer sales by adding more items into the market basket.

- A. Lowering costs
- B. Increasing revenues
- C. Increasing productivity
- D. Reducing risk

25. Decision trees can handle _____

- A. High dimensional data
- B. low dimensional data
- C. medium dimensional data
- D. none of these

26. In random forest or gradient boosting algorithms, features can be of any type. For example, it can be a continuous feature or a categorical feature. Which of the following option is true when you consider these types of features?

- A. Only Random forest algorithm handles real valued attributes by discretizing them.
- B. Only Gradient boosting algorithm handles real valued attributes by discretizing them.
- C. Both algorithms can handle real valued attributes by discretizing them.
- D. None of these

27. Which storage subsystem can support massive data volumes of increasing size.

- A. Extensibility
- B. Fault tolerance
- C. Scalability
- D. High-speed I/O capacity

28. Which of the following is not strength of Decision Tree?

- A. able to generate understandable rules
- B. able to handle both continuous and categorical variables
- C. less appropriate for estimation tasks
- D. perform classification without requiring much computation

29. Movie Recommendation systems are an example of

1. Classification
 2. Clustering
 3. Reinforcement Learning
 4. Regression
- A. 2 only
B. 1 and 3
C. 1 and 2
D. 2 and 3

30 Decision tree is an _____ algorithm.

- A. Supervised learning
- B. Unsupervised learning
- C. Both
- D. None of these

Section B

Note: Attempt all questions.

40 x 0.5 = 20 Marks

31. Sentiment Analysis is an example of

1. Regression
2. Classification
3. clustering
4. Reinforcement Learning

- A. 1, 2 and 4
- B. 1, 2 and 3
- C. 1 and 3
- D. 1 and 2
32. If $Y = 5 - 0.7X$, then the value of Y intercepts is equal to:
- A. -0.7
- B. 5
- C. $0.2X$
- D. All of the above
33. The branch of statistics that deals with the development of particular statistical methods are classified as:
- A. industry statistics
- B. economic statistics
- C. applied statistics
- D. applied statistics
34. In Decision-tree algorithm at the beginning, we consider the whole training set as _____.
- A. leaf
- B. root
- C. steam
- D. none of these

35. Which of the following is true about regression analysis?

- A. answering yes/no questions about the data
- B. estimating numerical characteristics of the data
- C. modeling relationships within the data
- D. describing associations within the data

36. What is the maximum depth in a decision tree?

- A. the length of the longest path from a root to a leaf
- B. the length of the shortest path from a root to a leaf
- C. the length of the longest path from a root to a sub-node
- D. None of these

37. Text Analytics; also referred to as Text Mining?

- A. True
- B. False
- C. Can be true or False
- D. Cannot say

38. _____ is the measure of uncertainty of a random variable, it characterizes the impurity of an arbitrary collection of examples.

- A. Information Gain
- B. Gini Index
- C. Entropy
- D. none of these

39. What is a hypothesis?

- A. A statement that the researcher wants to test through the data collected in a study.
- B. A research question the results will answer.
- C. A theory that underpins the study.
- D. A statistical method for calculating the extent to which the results could have happened by chance.

40. NumPy is often used along with packages like?

- A. Node.js
- B. Matplotlib
- C. SciPy
- D. Both B and C

41. The most important object defined in NumPy is an N-dimensional array type called?

- A. ndarray
- B. narray
- C. nd_array
- D. darray

42. _____ is a metric to measure how often a randomly chosen element would be incorrectly identified.

- A. Information Gain

- B. Gini Index
- C. Entropy
- D. none of these
43. Which of the following is not weakness of Decision Tree ?
- A. able to generate understandable rules
- B. can be computationally expensive to train
- C. less appropriate for estimation tasks
- D. prone to errors in classification problems with many class
44. What will be output for the following code?
- ```
import numpy as np
a = np.array([1,2,3])
```
- A. [[1, 2, 3]]
- B. [1]
- C. [1, 2, 3]
- D. Error
45. If a dimension is given as \_\_\_\_\_ in a reshaping operation, the other dimensions are automatically calculated.
- A. Zero
- B. One
- C. Negative one
- D. Infinite

46. What will be output for the following code?

```
import numpy as np
dt = dt = np.dtype('i4')
print dt
```

- A. int32
- B. int64
- C. int128
- D. int16

47. Which of the following Numpy operation are correct?

- A. Mathematical and logical operations on arrays.
- B. Fourier transforms and routines for shape manipulation.
- C. Operations related to linear algebra.
- D. All of the above

48. What is the syntax for dtype object?

- A. numpy.dtype(object, align, copy, subok)
- B. numpy.dtype(object, align, copy)
- C. numpy.dtype(object, align, copy, ndmin)
- D. numpy\_dtype(object, align, copy)

49. What is the range of uint32 data type?

- A. (-2147483648 to 2147483647)

- B. (-32768 to 32767)
- C. (0 to 65535)
- D. (0 to 4294967295)

50. What will be output for the following code?

```
import numpy as np
a = np.array([1,2,3,5,8])
print (a.ndim)
```

- A. 0
- B. 1
- C. 2
- D. 3

51. What is the correct code to install numpy in the linux system containing python3?

- A. pip numpy install python3
- B. pip3 install numpy
- C. pip install numpy
- D. python3 pip3 numpy install

52. fetch numpy as np

```
np.array(list)
```

Is it true to import numpy module like this?

A. Yes, true

B. Syntax Error

53. Is the following statement true?

numpy array can be converted to the list in python3?

A. False

B. True

C. Can't say

D. None of the above

54. Correct syntax of the reshape () function in Numpy array python is

A. array.reshape(shape)

B. reshape(shape,array)

C. reshape(array,shape)

D. reshape(shape)

55. Suppose, your target variable is the price of a house using Decision Tree. What type of tree you need to predict the target?

A. Classification tree

B. Regression tree

C. Clustering tree

D. Dimensionality reduction tree

56. A \_\_\_\_\_ is a collection of data values and operations that can be applied to that data.

- A. Data Structure
- B. Data Frame
- C. Table
- D. None of the above

57. Which of the following statement is wrong?

- A. We can create Series from Dictionary in Python.
- B. Keys of dictionary become index of the series.
- C. Order of indexes created from Keys may not be in the same order as typed in dictionary.
- D. All are correct

58. Decision tree learners may create biased trees if some classes dominate. What's the solution of it?

- A. balance the dataset prior to fitting
- B. imbalance the dataset prior to fitting
- C. balance the dataset after fitting
- D. No solution possible

59. The data label associated with a particular value of Series is called its \_\_\_\_\_

- A. Data value
- B. Index

- C. Value
- D. None of the above

60. Decision trees are also known as CART. What is CART?

- A. Classification and Regression Trees
- B. Customer Analysis and Research Tool
- C. Communication Access Real-time Translation
- D. Computerized Automatic Rating Technique

61. Bootstrap and Aggregation, commonly known as \_\_\_\_\_

- A. Information Gain
- B. bagging
- C. Entropy
- D. none of these

62. Random Forest has \_\_\_\_\_ as base learning models

- A. multiple decision trees
- B. bagging
- C. Entropy
- D. none of these

63. \_\_\_\_\_ helps improve machine learning results by combining several models.

- A. Machine Learning

- B. bagging
  - C. Entropy
  - D. Ensemble learning
64. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?
- A. Decision Tree
  - B. Regression
  - C. Classification
  - D. Random Forest
65. To find the minimum or the maximum of a function, we set the gradient to zero because:
- A. The value of the gradient at extrema of a function is always zero
  - B. Depends on the type of problem
  - C. Both A and B
  - D. None of the above
66. The most widely used metrics and tools to assess a classification model are:
- A. Confusion matrix
  - B. Cost-sensitive accuracy

- C. Area under the ROC curve
  - D. All of the above
67. Which of the following is a good test dataset characteristic?
- A. Large enough to yield meaningful results
  - B. Is representative of the dataset as a whole
  - C. Both A and B
  - D. None of the above
68. Which of the following is a disadvantage of decision trees?
- A. Factor analysis
  - B. Decision trees are robust to outliers
  - C. Decision trees are prone to be overfit
  - D. None of the above
69. How do you handle missing or corrupted data in a dataset?
- A. Drop missing rows or columns
  - B. Replace missing values with mean/median/mode
  - C. Assign a unique category to missing values
  - D. All of the above
70. What is the purpose of performing cross-validation?
- A. To assess the predictive performance of the models

- B. To judge how the trained model performs outside the sample on test data
- C. Both A and B – answer
- D. None of the above

Printed Pages:4

University Roll No: .....

**End Term Examination, Odd Semester 2021-22**

**B. Tech, CSE, Year-III/IV, Semester-V/VII**

**BCSE 0105: Machine Learning**

**Time: 3 Hour**

**Maximum Marks: 50**

**Section- A**

**Note: Attempt ANY FIVE Questions.**

**$5 \times 4 = 20$  marks**

- (I) What is categorical data? Discuss one-hot encoding to convert categorical data to numerical data using an example.
- (II) What is K-means clustering? How to decide the optimal number of K in the K-means Algorithm?
- (III) What do you mean by the Logistic Regression? Explain the hypothesis representation and cost function for Logistic Regression.
- (IV) What is the general principle of an ensemble method and discuss bagging and boosting concepts in an ensemble method?
- (V) The values of independent variable X and dependent variable Y are given below:

| X | Y |
|---|---|
| 0 | 2 |
| 1 | 3 |
| 2 | 5 |
| 3 | 4 |
| 4 | 6 |

Find the least square regression line  $y = m*x + c$ . Estimate the value of Y when X is 10.

- (VI) Differentiate between ‘Over fitting’ and ‘Under fitting’ in Machine learning? Explain how can you avoid them?

### Section- B

Note: Attempt ALL Questions.

5 x 3 = 15 marks

- (I) Briefly describe and give an example of the approach to hierarchical clustering.
- (II) Explain the concept of bias and variance tradeoff.
- (III) Why is naive Bayesian classification called “naive”? Briefly outline the major ideas of naive Bayesian classification.
- (IV) Differentiate between feature selection and feature extraction. Enlist one method for each.
- (V) Explain SVM classifier. What are the support vectors, hyper plane and kernel in SVM?

### Section- C

Note: Attempt ANY THREE Questions.

3 x 5 = 15 marks

- (I) Suppose there are several objects (4 types of medicines) and each object has two attributes or features as shown in the table below. Apply the K-means clustering algorithm to group these objects into the K=2 group of medicine based on the two features (pH and weight index).

| Object     | X: Weight index | Y:pH |
|------------|-----------------|------|
| Medicine A | 1               | 1    |
| Medicine B | 2               | 1    |
| Medicine C | 4               | 3    |
| Medicine D | 5               | 4    |

- (II) Suppose you have given a dataset of weather conditions and corresponding target variable "Play". So, using this dataset you need to decide that whether we should play or not on a particular day according to the outlook using Naïve Bayes Classifier.

Problem statement: If the outlook is sunny, then the Player should play or not?

| Outlook  | Play |
|----------|------|
| Rainy    | Yes  |
| Sunny    | Yes  |
| Overcast | Yes  |
| Overcast | Yes  |
| Sunny    | No   |
| Rainy    | Yes  |
| Sunny    | Yes  |
| Overcast | Yes  |
| Rainy    | No   |
| Sunny    | No   |
| Sunny    | Yes  |
| Rainy    | No   |
| Overcast | Yes  |
| Overcast | Yes  |

(III) Explain the AdaBoost Algorithm with an example.

(IV) What is the use of Regularization? Explain L1 and L2 Regularizations.

but algorithms reduce to simple a noisy objective function. (II)  
research with some of "old" old ideas now gain importance  
it is too to valq. Many people say that it has been long  
since we have had a lot of progress in machine  
learning. But there are still many applications of the approach  
of learning from data.

Today we will learn about AdaBoost which is a very good  
example of how old ideas can be used in machine learning.

| Ques | Ans                                    |
|------|----------------------------------------|
| Q1   | What is AdaBoost?                      |
| Q2   | What is the main idea behind AdaBoost? |
| Q3   | What is the output of AdaBoost?        |
| Q4   | What is the weight vector?             |
| Q5   | What is the weight of a classifier?    |
| Q6   | What is the weight of a feature?       |
| Q7   | What is the weight of a bias?          |
| Q8   | What is the weight of a classifier?    |
| Q9   | What is the weight of a feature?       |
| Q10  | What is the weight of a bias?          |
| Q11  | What is the weight of a classifier?    |
| Q12  | What is the weight of a feature?       |
| Q13  | What is the weight of a bias?          |
| Q14  | What is the weight of a classifier?    |
| Q15  | What is the weight of a feature?       |
| Q16  | What is the weight of a bias?          |
| Q17  | What is the weight of a classifier?    |
| Q18  | What is the weight of a feature?       |
| Q19  | What is the weight of a bias?          |
| Q20  | What is the weight of a classifier?    |
| Q21  | What is the weight of a feature?       |
| Q22  | What is the weight of a bias?          |
| Q23  | What is the weight of a classifier?    |
| Q24  | What is the weight of a feature?       |
| Q25  | What is the weight of a bias?          |
| Q26  | What is the weight of a classifier?    |
| Q27  | What is the weight of a feature?       |
| Q28  | What is the weight of a bias?          |
| Q29  | What is the weight of a classifier?    |
| Q30  | What is the weight of a feature?       |
| Q31  | What is the weight of a bias?          |
| Q32  | What is the weight of a classifier?    |
| Q33  | What is the weight of a feature?       |
| Q34  | What is the weight of a bias?          |
| Q35  | What is the weight of a classifier?    |
| Q36  | What is the weight of a feature?       |
| Q37  | What is the weight of a bias?          |
| Q38  | What is the weight of a classifier?    |
| Q39  | What is the weight of a feature?       |
| Q40  | What is the weight of a bias?          |
| Q41  | What is the weight of a classifier?    |
| Q42  | What is the weight of a feature?       |
| Q43  | What is the weight of a bias?          |
| Q44  | What is the weight of a classifier?    |
| Q45  | What is the weight of a feature?       |
| Q46  | What is the weight of a bias?          |
| Q47  | What is the weight of a classifier?    |
| Q48  | What is the weight of a feature?       |
| Q49  | What is the weight of a bias?          |
| Q50  | What is the weight of a classifier?    |

**Printed Pages: 03**      University Roll No: .....  
**End Term Examination, Odd Semester 2021-22**  
**B.Tech., CSE, III<sup>rd</sup> year, V<sup>th</sup> semester**  
**BCSE0157: Introduction to Big Data Analytics**

**Time: 3 Hour**      **Maximum Marks: 50**

**Section- A**

**Note: Attempt ANY FIVE Questions.**       **$5 \times 4 = 20$  marks**

- (I) What do you mean by Big Data Analytics? Write two real life applications with brief scenario of Big Data Analytics. How we can say whether particular data is big data or not?
- (II) What are three types of Digital Data? How they are different from each other? Classify the following in three types of Digital Data with appropriate reasons:
  - a) Images
  - b) CCTV Footage
  - c) Facebook post with hashtag
  - d) Dataset in JSON format of Bing Search Result Page
  - e) Dataset of GLA University student's personal information
  - f) Data extracted from Big Bazaar billing system
- (III) What is Hadoop? Explains Daemons of Hadoop? Explain the working of File read in Hadoop?
- (IV) Discuss various types of Analytics with help of example in Tabular format.
- (V) Explain types of NoSQL with example and how each type is different from other?
- (VI) How mongoDB differentiate among different documents? What makes mongoDB more popular among IT professional over conventional DBMS systems?

## Section- B

Note: Attempt ALL Questions.

$5 \times 3 = 15$  marks

- (I) Explain SSTable, MemTable, Bloom-Filter with their importance in Cassandra?
- (II) Explain MongoDB's CRUD operation by assuming suitable dataset.
- (III) Explain collection datatype in Apache Cassandra?
- (IV) Briefly explain Hadoop Ecosystem and draw it diagram?
- (V) Discuss MapReduce and WordCount problem with help of diagram.

## Section- C

Note: Attempt ANY THREE Questions.

$3 \times 5 = 15$  marks

- (I) Consider you are CTO of ABC company, this company had design an app for hosting microvideos on internet. Company had collected lot of usage data of users such as time they spent, what videos they like and many more things such as gender, morning, evening, afternoon when they spent time on App. You keep on storing this data but now you want to harness this data for decision making. Answer the following:
  - a) What tool will you use? Explain briefly the features of that tools?
  - b) So far, your employees only worked on SQL. Do they need any training for this tool of your choice? Discuss the problems they can solve after getting training?
  - c) Is there any tool that not require any training? If yes, Explain that tool briefly.

- d) How will the tool of your choice that you have mentioned in part (a) is taking care of replication? Discuss it vividly, if possible, with diagram.
- e) How can the tool of your choice help in making analytics fast explain your observation with an example?
- (II) Answer the following (assume suitable data):
- Write a query in mongoDB to fetch unique document
  - Write a query to display no. of documents in MongoDB
  - Write a query to display last two document in a collection in MongoDB
  - Write query in Cassandra to create keyspace
  - Write a query to create any two tables in Cassandra
- (III) Explain the importance of YARN component and its architecture with help of diagram.
- (IV) Write commands for following:
- Justify the importance of Pig on Hadoop Mapreduce
  - Write any program in Pig Latin using various commands (filter, load, Dump, store, foreach,..etc.)

Printed Pages: University Roll No: .....

End Term Examination, Odd Semester 2021-22

B.Tech (CSE), III Year V Semester

BCSE 0207: Cloud Computing

Time: 3 Hour

Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

$5 \times 4 = 20$  marks

- (I) How utility computing benefits user?
- (II) Imagine that you are running a website that is subject to wide fluctuations in demand, such as a ticket-selling system for events. How would you plan for migrating the system from traditional to cloud?
- (III) You are a member of a 4 - member project team. During the covid time, all the members are stuck in their native places. With no means of physical meeting, your project has come to a halt. Give the solution how cloud services can help you and your team in this pandemic time.
- (IV) Compare the different types of storage available in OpenStack? How a user can choose the storage backend for OpenStack?
- (V) What's the difference between hybrid cloud and multi cloud? How can organizations make the multi cloud solutions work together? What are the challenges that the organization will face?
- (VI) Discuss cloud security with respect to infrastructure level and network level?

### Section- B

Note: Attempt ALL Questions.

5 x 3 = 15 marks

- (I) Why traditional security measures do not work in cloud computing? Explain any one method that can be used for cloud security?
- (II) What are the areas for security concerns in cloud computing? Explain each in brief.
- (III) Why a user faces more security challenges after installing the VMM on a OS compare to one who have installed directly on hardware?
- (IV) Why virtual security is more hard to attain? How does it differ from physical security?
- (V) Explain what is SOA governance and what are its functions. Are web services and SOA the same? Explain.

### Section- C

Note: Attempt ANY THREE Questions.

3 x 5 = 15 marks

- (I) Not having experts in the domain of Multi cloud is restricting its growth. Explain the statement. How does the security concept changes to when data is at rest and when data is at transit? Discuss the security approach to both scenarios.
- (II) Having a large number of virtual machine may increase the security risks associated with system. Discuss the problem by citing the virtual threats that may come due to this problem. A malicious VM can bring the entire system down. Justify.
- (III) What is the problem that is associated with when some resources and data previously confined to a private network are now exposed to the Internet, and to a shared public

network belonging to a third-party cloud provider?

- (IV) How DoS and DDoS attack can cause harm to Cloud Servers? In cloud who is responsible for security? Justify.

Printed Pages:3 University Roll No: .....  
End Term Examination, Odd Semester 2021-22  
Program (AIML), Year: III, Semester: 5<sup>th</sup>  
Subject Code (BCSE0703): Fundamental of Data Mining and  
Predictive Modeling

Time: 3 Hour Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions. 5 x 4 = 20 marks

- (I) What is CNN? Draw and explain detail on the architecture of CNN?
- (II) What is association rule mining, explain detail with an example. Define and discuss the support and confidence with its mathematical expression.
- (III) Using Apriori algorithm, generate rules for the following transaction of the given data. Consider the value of Support= 50%, and Confidence= 75%.

| Transaction ID | Item purchased |
|----------------|----------------|
| 1              | A, B, C, D     |
| 2              | A, B, D        |
| 3              | A, E, F        |
| 4              | A, D, E        |
| 5              | B, D, E        |

- (IV) What is ROC? Consider few data (five from each) of two different classes and using logistic regression function with different threshold value trying to classify them. Draw the ROC plot between false positive rate with true positive rate.
- (V) Define what is pooling, padding and their significance in convolution neural network. With an appropriate example, discuss on Max-pooling and Average-pooling.

- (VI) What LSTM, explain with an example. Why we used LSTM in recurrent neural network.

### Section- B

Note: Attempt ALL Questions.

$5 \times 3 = 15$  marks

- (I) What is skewness? How different value of skewness help for extraction of feature in a given dataset, explain?
- (II) What do you mean by Bayesian information criterion, and how it differs from Akaike information criterion?
- (III) In the below figure the input data is given as  $4*4$  matrix, with the  $3*3$  filter. Perform the convolution operation with an assumption that the filter follows stride-1 path.

|   |   |   |   |
|---|---|---|---|
| 2 | 0 | 1 | 1 |
| 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 |
| 0 | 3 | 0 | 0 |

|   |   |   |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 0 | 0 |
| 0 | 1 | 0 |

- (IV) What do you mean by AUC? Discuss its significance for different types of ROC plot.
- (V) If a given data set having two different variable and for this data the covariance matrix is given below having dimension  $2*2$ . Using this data find the eigen value for the matrix.

|                     |       |       |
|---------------------|-------|-------|
| Covariance (x, y) = | 0.616 | 0.615 |
|                     | 0.615 | 0.716 |

### Section- C

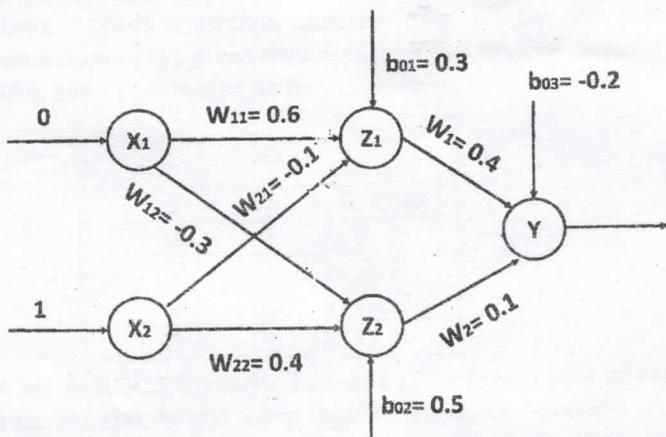
Note: Attempt ANY THREE Questions.

$3 \times 5 = 15$  marks

- (I) What is empirical mode decomposition and IMFs in a time

series data? Explain all five steps in EMD algorithm to find the residual values with an example.

- (II) What are different types of test that one can used for evaluation of predictive model. Step by step process explain on KS-test by considering below data, to check the null hypothesis is true or not. Assuming significant value  $\alpha = 0.05$ ,  $D_\alpha = 0.513$ , is the critical value and data set,  $X = 0.16, 0.95, 0.06, 0.52$ , and  $0.29$ .
- (III) Consider the below back-perceptron model and find the error in output "y" and weight " $W_1$ ". Given that the learning rate = 0.25, target value = 1, activation function= binary sigmoid.
- (IV) Draw and describe detail architecture of RNN. Discuss on exploding gradient problem and how one can eliminate this issue using different techniques, explain in brief?



**Printed Pages: 02**

**University Roll No: .....**

**End Term Examination, Odd Semester 2021-22**

**B. Tech.(CCV) Year: III, Semester: V**

**Cloud Computing Architecture & Deployment Models : BCSE 0503**

**Time: 3 Hour**

**Maximum Marks: 50**

**Section- A**

**Note: Attempt ANY FIVE Questions.**

**$5 \times 4 = 20$  marks**

- (I) An organization wants PowerVM, KVM and ESXi hypervisors to provision VMs. It is rather difficult to manage such heterogeneous system with a single UI. Suggest a solution to manage such heterogeneous system?
- (II) Effective governance is essential to guiding management processes and decision making to deliver IT services in accordance with the needs of the organization. Identify some standards that support the governance of cloud computing.
- (III) The mayor of a town wants that everything should be online with its user always accessing the mail on go. The city has a population over 1 million. City employees fulfill a range of important functions- from policing streets to supplying water and power to city residents and businesses, and from operating libraries to designing and building wastewater treatment plants and other public facilities. The mayor wants to provide all these employees with modern tools that help them do their jobs. The mayor want to reduce the response time of the primary services, users having option of complaining either in the mode of text or by uploading picture. Examine the requirements of the mayor and identify the deployment model and cloud services. Also, justify your answer.
- (IV) An organization want a storage solution with no single point of failure. Identify a storage solution that match the requirement of the organization, also mention other features of the storage solution in support of your answer.
- (V) What are the different types of Network traffic exist in OpenStack environment? What is ML2 plugin? Also explain the types of drivers used in ML2 plugin.

- (VI) Differentiate between different types of storage used in OpenStack architecture.

### Section- B

*Note: Attempt ALL Questions.*

**5 x 3 = 15 marks**

- (I) Networking is a very important part of the overall security of the UEC system. Depending on the level of security that a setup may require, explain the networking modes that are available at set-up time.
- (II) What is GFO? Also, compare different types of Gluster volumes based on their advantages and disadvantages.
- (III) Model an OpenNebula based highly available virtual data center? Justify it with the help of example.
- (IV) Cloud Orchestration Layer automates provisioning of cloud services using policy-based tools. Explain HOT and its various components using a template.
- (V) Summarize the VM creation process in OpenStack using different components of nova service.

### Section- C

*Note: Attempt ANY THREE Questions.*

**3 x 5 = 15 marks**

- (I) Design a Tenant Model Architecture of OpenStack with three nodes configuration using appropriate openstack components/ services.
- (II) Differentiate between Eucalyptus and OpenNebula based on customizability, DHCP, user security, internal security, networking issues and unique features of cloud environment.
- (III) Assume that you have used the Eucalyptus architecture to design a private cloud with five physical machines. The components of the Eucalyptus architecture are: Node Controller (NC), Cluster Controller (CC), Walrus Storage Controller (WS3), Storage Controller (SC) and Cloud Controller (CLC). How you will use the five physical machines to design a private cloud with two clusters? (Note: Use diagram to explain your ideas)
- (IV) How object storage is different from block storage? Explain the significance of ring, proxy server, zones and container defined in swift service of openstack.

**Printed Pages: 02**

**University Roll No: .....**

**End Term Examination, Odd Semester 2021-22**

**B. Tech. (CSF), Year: III, Semester: V**

**Cloud Architecture & Deployment Models : BCSE 0604**

**Time: 3 Hour**

**Maximum Marks: 50**

**Section- A**

**Note: Attempt ANY FIVE Questions.**

**$5 \times 4 = 20$  marks**

- (I) Effective and efficient management of hybrid cloud environments require rich Cloud Management Platform. Suggest a Reference Architecture Model for Hybrid Cloud that enable visibility, simplify management, and optimize the utilization of resources?
- (II) Identify the types of virtualization an organization must use to set up its offices in different part of the world. Justify your answer with proper reasoning. Also, discuss the cost impacts of proposed virtualization.
- (III) A school having nice computer lab facility wants to go on cloud so as to decrease the cost of licensing software. The infrastructure that the school is having is updated and is in good condition. However, they are reluctant to move on cloud, thinking it to be expensive. Neither they want their database to move out of premises. Examine the requirements of the school and identify the deployment model and cloud services. Also, justify your answer.
- (IV) An organization want a storage solution with no single point of failure. Identify a storage solution that match the requirement of the organization, also mention other features of the storage solution in support of your answer.
- (V) What are the different types of Network traffic exist in OpenStack environment? What is ML2 plugin? Also explain the types of drivers used in ML2 plugin.
- (VI) Differentiate between different types of storage used in OpenStack architecture.

### Section- B

Note: Attempt ALL Questions.

**5 x 3 = 15 marks**

- (I) Networking is a very important part of the overall security of the UEC system. Depending on the level of security that a setup may require, explain the networking modes that are available at set-up time.
- (II) What is GFO? Also, compare different types of Gluster volumes based on their advantages and disadvantages.
- (III) Model an OpenNebula based highly available virtual data center? Justify it with the help of example.
- (IV) Cloud Orchestration Layer automates provisioning of cloud services using policy-based tools. Explain HOT and its various components using a template.
- (V) Summarize the VM creation process in OpenStack using different components of nova service.

### Section- C

Note: Attempt ANY THREE Questions.

**3 x 5 = 15 marks**

- (I) Design a Tenant Model Architecture of OpenStack with three nodes configuration using appropriate openstack components/ services.
- (II) Differentiate between Eucalyptus and OpenNebula based on customizability, DHCP, user security, internal security, networking issues and unique features of cloud environment.
- (III) Assume that you have used the Eucalyptus architecture to design a private cloud with five physical machines. The components of the Eucalyptus architecture are: Node Controller (NC), Cluster Controller (CC), Walrus Storage Controller (WS3), Storage Controller (SC) and Cloud Controller (CLC). How you will use the five physical machines to design a private cloud with two clusters? (Note: Use diagram to explain your ideas)
- (IV) How object storage is different from block storage? Explain the significance of ring, proxy server, zones and container defined in swift service of openstack.

Printed Pages:3

University Roll No: .....

End Term Examination, Odd Semester 2021-22

Program (DA), Year: III, Semester: 5<sup>th</sup>

Subject Code (BCSE0553): Data Mining and  
Predictive Modeling

Time: 3 Hour

Maximum Marks: 50

### Section- A

Note: Attempt ANY FIVE Questions.

5 x 4 = 20 marks

- (I) Define what is pooling, padding and their significance in convolution neural network. With an appropriate example, discuss on Max-pooling and Average-pooling.
- (II) Using Apriori algorithm, generate rules for the following transaction of the given data. Consider the value of Support= 50%, and Confidence= 75%.

| Transaction ID | Item purchased |
|----------------|----------------|
| 1              | A, B, C, D     |
| 2              | A, B, D        |
| 3              | A, E, F        |
| 4              | A, D, E        |
| 5              | B, D, E        |

- (III) What is ROC? Consider few data (five from each) of two different classes and using logistic regression function with different threshold value trying to classify them. Draw the ROC plot between false positive rate with true positive rate.
- (IV) What is association rule mining, explain detail with an example. Define and discuss the support and confidence with its mathematical expression.
- (V) What is CNN? Draw and explain detail on the architecture of CNN?
- (VI) What LSTM, explain with an example. Why we used LSTM in recurrent neural network.

### Section- B

Note: Attempt ALL Questions.

5 x 3 = 15 marks

- (I) What do you mean by Bayesian information criterion, and how it differs from Akaike information criterion?
- (II) What is skewness? How different value of skewness help for extraction of feature in a given dataset, explain?

- (III) In the below figure the input data is given as 4\*4 matrix, with the 3\*3 filter. Perform the convolution operation with an assumption that the filter follows stride-1 path.

|   |   |   |   |
|---|---|---|---|
| 2 | 0 | 1 | 1 |
| 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 |
| 0 | 3 | 0 | 0 |

|   |   |   |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 0 | 0 |
| 0 | 1 | 0 |

- (IV) What do you mean by AUC? Discuss its significance for different types of ROC plot.  
(V) If a given data set having two different variable and for this data the covariance matrix is given below having dimension 2\*2. Using this data find the eigen value for the matrix.

$$\text{Covariance } (x, y) =$$

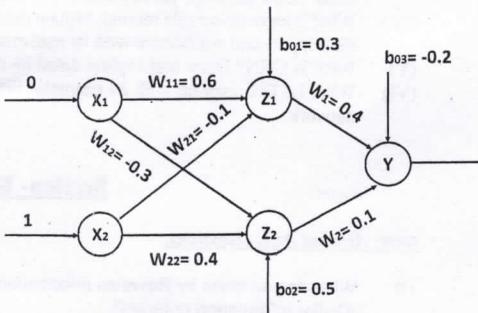
|       |       |
|-------|-------|
| 0.616 | 0.615 |
| 0.615 | 0.716 |

### Section- C

Note: Attempt ANY THREE Questions.

3 x 5 = 15 marks

- (I) What is empirical mode decomposition and IMFs in a time series data? Explain all five steps in EMD algorithm to find the residual values with an example.  
(II) Consider the below back-perceptron model and find the error in output "y" and weight "W1". Given that the learning rate = 0.25, target value = 1, activation function = binary sigmoid.



- (III) Draw and describe detail architecture of RNN. Discuss on exploding gradient problem and how one can eliminate this issue using different techniques, explain in brief?

- (IV) What are different types of test that one can used for evaluation of predictive model. Step by step process explain on KS-test by considering below data, to check the null hypothesis is true or not. Assuming significant value  $\alpha=0.05$ ,  $D_\alpha=0.513$ , is the critical value and data set,  $X=0.16, 0.95, 0.06, 0.52$ , and  $0.29$ .

**Printed Pages:** \_\_\_\_\_ **University Roll No:** .....

**End Term Examination, Odd Semester 2021-22**

**B.Tech (IoT), III, V**

**BCSE 0653: Descriptive Analytics for IoT**

**Time: 3 Hour**

**Maximum Marks: 50**

### Section- A

**Note: Attempt ANY FIVE Questions.**

**$5 \times 4 = 20$  marks**

- (I) Enlist BI tools and discuss any one in detail.
- (II) Explain how to maximize value for BI in details.
- (III) How different analytics play a vital role in BI? Explain with example.
- (IV) What is virtual warehouse? Can data warehouse alternative for BI? Justify your answer with example?
- (V) Explain phases of ETL?
- (VI) How can you say a data as big data? Explain 3V's of Gardner's theory of Big Data?

### Section- B

**Note: Attempt ALL Questions.**

**$5 \times 3 = 15$  marks**

- (I) Differentiate between data mart and data warehouse?
- (II) Define Data cube? What is list report and map report?
- (III) Mention the key trends in IoT analytics market according to Gartner?
- (IV) Differentiate between centralized and decentralized architecture?
- (V) Write short note on IBM Infosphere?

### Section- C

Note: Attempt ANY THREE Questions.

$3 \times 5 = 15$  marks

- I. Explain difference between OLAP and OLTP with example?
- II. Explain the dashboard design principle in details?
- III. Draw and Explain the concept of IoT data life cycle management?
- IV. What is Meta Data Model and write steps involved in metadata model?

Printed Pages: 02

Univ. Roll No:.....

End Term Examination, Odd Semester 2021-22  
B. Tech. (CSE-All), III-Year, V-Semester  
BCSC 0011 : Theory of Automata & Formal Languages

Time : 3 Hour

Max. Marks : 50

SECTION A

Note: Attempt ANY FIVE Questions. (5x4 =20 marks)

- I. Find the Reduced Grammar equivalent to the grammar G whose productions are:  
 $S \rightarrow AB|CA, \quad B \rightarrow BC|AB, \quad A \rightarrow a, \quad C \rightarrow aB|b.$
- II. Design a Push Down Automaton that accepts the language  $L = \{a^n c a^n | n \geq 1\}$  by final state.
- III. State Pumping lemma for regular languages. Use it to prove that the language  $L = \{a^p | p \text{ is prime}\}$  is not regular.
- IV. Design a Turing Machine that accepts the language  $L = \{a^n b^n | n \geq 1\}$
- V. Show that complement of a recursive language is also recursive.
- VI. Design a Turing machine to compute the sum of two positive integers  $n$  and  $m$ , where  $n, m \geq 1$ .

SECTION-B

Note: Attempt ALL Questions. (5 x3 =15 marks)

- I. Convert the following Context Free grammar to equivalent CNF.

$$S \rightarrow aAbB \quad A \rightarrow Ab|b \quad B \rightarrow Ba|a$$

- II. Convert the following Context Free Grammar to equivalent GNF.  $S \rightarrow YY|0$     $Y \rightarrow SS|1$

III. Write a short note on Universal Turing Machine (UTM).

IV. Describe halting problem of Turing machine.

V. Discuss Church Turing Thesis in detail. Why is it not called a theorem? Explain.

## SECTION C

*Note: Attempt ANY THREE Questions.*

(3x5=15 marks)

- I. Q.1 If a language L and its complement  $\bar{L}$  are both recursively enumerable then prove that L is recursive.
  - II. Given a Context Free Grammar  $G = (\{S, A, B\}, \{0, 1\}, P, S)$  with the production set as follows:

$$S \rightarrow 0B|1A \quad A \rightarrow 0S|1AA|0 \quad B \rightarrow 1S|0BB|1$$

Design a PDA M corresponding to this CFG.

- III. Design a Turing Machine to accept the language  $L = \{WW^R \mid W \in \{a, b\}^*\}$

IV. Discuss any five variants of the Turing machine with proper model diagram and transition function.

Printed Pages: 1

University Roll No.....

End Term Examination, Odd Semester 2021-22  
B.Tech CSE, Year: III, Sem: V  
BCSE0251: Full Stack Using Scripting Technologies

Time: 03 Hours.

Max. Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

$5 \times 4 = 20$  marks

- (I) What is DOM? Explain its Nodes Types?
- (II) Explain various Dialog Boxes that can be created using Client-Side Scripting Language.
- (III) Explain Text Shadow and Box Shadow with Examples.
- (IV) Explain any Two JQuery Getter and Setter Methods with Examples.
- (V) What is document.ready( ) method? Explain its importance in JQuery.
- (VI) What is the use of test( ) method in JavaScript? Write an Example.

Section- B

Note: Attempt ALL Questions.

$5 \times 3 = 15$  marks

- (I) Write a program in JavaScript to generate a unique random number and display in the heading-2.
- (II) What is Variable or Function Hoisting? Explain with the help of Examples.
- (III) What is the purpose of Container, Row and Column in Bootstrap Framework?
- (IV) Write a JavaScript code to create a new paragraph element under the div dynamically.      <div id="demo"></div>
- (V) Write a JavaScript code to add today's date in a web page dynamically.

Section- C

Note: Attempt ANY THREE Questions.

$3 \times 5 = 15$  marks

- (I) Difference between Global Scope, Function Scope and Block Scope? Explain Var, Let and Const keywords with respect to above stated scopes.
- (II) Write an example of Function Expression, Arrow Function and Self Invoking Function.
- (III) What is AJAX? Write code for add JQuery AJAX content in the page.
- (IV) Write JavaScript code to check a password between 6 to 20 characters which contain at least one numeric digit, one uppercase and one lowercase letter.

## End-Term Examination, Even Semester 2021-22

B. Tech. (CS), 3<sup>rd</sup> Year, 5<sup>th</sup> Semester

BCSE0254: PHP Scripting Language

Time: 3 Hours

Maximum Marks: 50

Section- A**Note: Attempt Any Five Questions.*****5 x 4 = 20 Marks***

- Q1. Illustrate the working of Heredoc and Nowdoc strings with an example.
- Q2. Write a complete HTML table tag and CSS code to draw the following output on the browser screen. IBM and Total Amount must be centrally aligned inside the cell. Also follow the alternate row color pattern with using CSS (dark gray and white).

| Name         | Designation       | Salary   | Company |
|--------------|-------------------|----------|---------|
| Steve        | Manager           | 1,00,000 |         |
| SURAJ        | Assistant Manager | 50,000   |         |
| Khushboo     | Analyseist        | 65,000   |         |
| Kartik       | Worker            | 20,000   |         |
| Saksham      | Worker            | 20,000   |         |
| Total Amount |                   | 2,55,000 | IBM     |

- Q3. Design a webpage using PHP that performs the following operations on the given string  
**"Was it a BMW or an audi car I saw".**
- Convert all characters to lower case letters.
  - Reverse the string produced by the above task.
  - Remove all spaces from the string produced.
  - Display the output string on the browser screen.
- Q4. What will be the output of the following code? Explain with justification.
- a)
- ```
<? php
    define("value","5");
    $value=value+10;
    echo value;
?>
```

b)

```

<?php
class Ibm{
    public $rooms = 5;
    public function __destruct()
    {
        echo $this->rooms=4;
    }
}
$obj = new Ibm();
echo $obj->rooms;
?>
```

- Q5. Design a webpage having a button to find the area of a triangle by heron's formula using a function with argument with return type and the result must be shown in the input type field which is not editable. Parameter list must be of float type.
 $(\sqrt{s(s-a)(s-b)(s-c)})$.
- Q6. How many types of modes are being there in the fopen() function. Elaborate all of them.

Section- B

Note: Attempt All Questions.

5 x 3 = 15 Marks

- Q1. Explain private, protected and public access modifiers in brief.
- Q2. Name any two magic functions used in PHP and write a short note on any one of them. Explain with an example.
- Q3. a) Explain the role of **\$this**.
 b) Why we required connection string in PHP.
- Q4. A company wants the users to register for an event with the following data on its website: **Name, Email, Mobile Number, Gender and password**. Design a webpage that displays the registration page with appropriate fields. However, the webpage shall not send the data to server unless the following validations are checked:
 a) No field shall be left blank
 b) The email shall be in the correct format
 c) The password must contain more than 8 characters.
 d) For gender you have to use radio buttons.
- Q5. Explain the working of following:
 a) **\$_REQUEST**
 b) **\$_SERVER**
 c) **Error_reporting()**

Note: Attempt Any Three Questions.**3 x 5 = 15 Marks****Q1.**

- a) Suppose you are the owner of a website and you want to get the client machine's name as well the IP address of the client machine write a PHP script to get these values and stored them in cookies.
- b) Explain the role of any six parameters required for Cookies creation.

Q2. In a website there are two pages one is "index.html" which consist a form having various input fields such as "name", "email", "password" etc. along with three buttons "Save data", "read data" and "reset". Write a PHP script "new.php" for the mentioned html page to store data inside the database while clicked save data button, show all the stored data on the browser screen in the tabular format when clicked on read data button also empty the content of input fields when user clicked on reset button.

Q3. Here is an attached database data table which consist of ID and username value of registered users.

Id	Username	Update	Delete
1	kartik	Update	Delete
2	rohit	Update	Delete
3	manish	Update	Delete
4	prakhar	Update	Delete
5	vishal	Update	Delete
6	garry	Update	Delete
7	harry	Update	Delete
8	prashant	Update	Delete

You have to write the PHP script for update button and delete button.

- a) consider the name of the buttons as `update_btn` and `delete_btn`
- b) Both the buttons are of submit type.
- c) Whenever user clicked on update button it will update that username value for the particular ID.
- d) When user clicked on delete button it will delete the complete row

Q4 Write a PHP script to create various web page for a particular website (registration, login, home, logout) and show the concept of Session variable to display the user name on each and every page until the logout button clicked.

Printed Pages: 02

University Roll No:

End Term Examination, Odd Semester 2021-22

B.Tech -AIML, III Year, V Semester

BCSE0704: Computational Linguistics and Natural Language Processing
Time: 3 Hour

Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

5 x 4 = 20 marks

- (I) What are word vectors? Where Do Word Vectors Come from?
- (II) Consider the following texts
T1: Julie loves me more than Linda loves me
T2: Jane likes me more than Julie loves me
Compute the similarity between T1 and T2 in terms of word counts and ignoring word order.
- (III) Explain any three text similarity measure techniques? Consider two sets $A = \{0, 1, 2, 5, 6\}$ and $B = \{0, 2, 3, 5, 7, 9\}$. Compute how similar are the given sets based on any similarity technique?
- (IV) A database contains 80 records on a particular topic. A search was conducted on that topic and 60 records were retrieved. Of the 60 records retrieved, 45 were relevant. Calculate the precision and recall scores for the search.
- (V) Explain the concept of sentence segmentation with an example?
- (VI) Differentiate between stemming and lemmatization with an example?

Section- B

Note: Attempt ALL Questions.

5 x 3 = 15 marks

- (I) Let's say I searched on Google for "what is precision and recall?" and in less than a minute I have about 15,600,000 results. Let's say out of these 15.6 million results, the relevant links to my question were about 2 million. Assuming there were also about 6 million more results that were relevant but weren't returned by Google, Compute Precision and Recall for such systems?

- (II) Explain the different logical steps used in natural language processing
- (III) Discuss various types of sentiment analysis?
- (IV) Describe Question Answering System? Discuss the different types of Question Answering Systems?
- (V) How NLP Makes Communication Easy Between Humans and Computers?

Section- C

Note: Attempt ANY THREE Questions.

$3 \times 5 = 15$ marks

- (I) We are looking to develop a machine learning algorithm to predict whether someone will pay a loan back or not.
 - A) What is the positive class?
 - B) What would a recall of 75% mean?
 - C) What would a precision of 85% mean?
- (II) What is the role of regular expression in NLP? Explain the use of Special Sequences in Regular Expressions. Explain the meaning of following sequences that come up with Regular expressions.
 - A) \b
 - B) \d
 - C) \D
 - D) eas+y
 - E) eas?
- (III) Consider a case insensitive query and document collection with a query Q and a document collection consisting of the following three documents:

Q: "gold silver truck"

D1: "Shipment of gold damaged in a fire"

D2: "Delivery of silver arrived in a silver truck"

D3: "Shipment of gold arrived in a truck"

Compute the similarity coefficient between the query Q to Document D1, D2 & D3
- (IV) Explain the different levels of sentiment Analysis?

Printed Pages: 02

University Roll No:

End Term Examination, Odd Semester, 2021-22

Program: B.Tech. Branch: CSE-CCV Year: III Semester: V

Subject Code: BCSE0510 Subject Name: Container Orchestration & Infrastructure Automation

Time: 3 Hours

Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

5 x 4 = 20 marks

- (I) What are docker images? Give commands to pull docker images from docker hub for ubuntu, ubuntu:16.04 and redis image.
- (II) Describe the lifecycle of Docker container with proper diagrammatic presentation and appropriate commands.
- (III) Show how you would create a container from an image with the help of proper description. And also explain Dockerfile with the help of commands.
- (IV) What is Docker-compose? Explain with the help of commands.
- (V) What is the role of Kube-apiserver? What are the main differences between the Docker Swarm and Kubernetes?
- (VI) What is "Continuous Integration" with reference to Jenkins? And also explain "Continuous Delivery" and "Continuous deployment".

Section- B

Note: Attempt ALL Questions.

5 x 3 = 15 marks

- (I) What are the main features of Docker compose?
- (II) What is the role of Load Balancer in Kubernetes?
- (III) How does Docker swarm work? Explain with the help of an example.
- (IV) What will happen if docker swarm manager goes down? Suggest an appropriate solution for this incident happened in any company.

- (V) To get started with Kubernetes development, you can use Minikube. What does Minikube stand for in the above statement? Give commands for Minikube start, stop and delete with proper description.

Section- C

Note: Attempt ANY THREE Questions.

3 x 5 = 15 marks

- (I) What are the various K8's services running on nodes and describe the role of each service? How details about the Pod's container: IP address, the ports used and a list of events related to the lifecycle of the Pod could be displayed with the help of commands?
- (II) What is kube-proxy? Explain the working of the master node in Kubernetes? Which command is used to delete the deployment?
- (III) What is a pod in Kubernetes? What are the types of controller managers? How will you display cluster information with the help of commands?
- (IV) What are etcd, Kubelet and Kubectl? Create the deployment with the help of commands for two images ubuntu and nginx. And also create containers inside the pods.

Printed Pages: University Roll No:

End Term Examination, Odd Semester 2021-22

B.Tech. CSE (Specialization in CSF), III Year, V Semester

BCSE0603: Information Security Audit & Monitoring

Time: 3 Hour

Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

$5 \times 4 = 20$ marks

- (I) Cyber Security and Information Security, are they related to each Other? Discuss.
- (II) What are the features of a good GRC Application Software?
- (III) Discuss the methodology adopted for Business – Information Security Alignment.
- (IV) Why encryption is required in transmission of cardholder data across open, public networks? and also describe how it can be achieved?
- (V) Briefly describe the steps involved in Security Assurance Audit Planning. And also discuss the advantages and disadvantages of Audit Planning.
- (VI) Why Auditing is required in Security concepts? Discuss various auditing tools relied upon by auditors.

Section- B

Note: Attempt ALL Questions.

$5 \times 3 = 15$ marks

- (I) How the documentation of the key-management processes and procedures for cryptographic keys used for encryption of cardholder data is useful.
- (II) Why security is important for Indian government? Enlist the

baseline control requirements as mandated by RBI for security in BFSI.

- (III) What is the Business Value of auditing? Also enlist types of auditing activities that may be utilized for critical controls in an organization.
- (IV) Discuss the requirements and methods to ensure proper user-authentication management for non-consumer users and administrators on all system components.
- (V) Describe the IT Act of India and also discuss the advantages of Cyber Laws.

Section- C

Note: Attempt ANY THREE Questions.

3 x 5 = 15 marks

- (I) What is the main objective behind PCI-DSS? Enlist and discuss the Control Objectives specified in PCI-DSS in form of logically related groups of control requirements.
- (II) How can an organization restrict physical access to cardholder data? Discuss the scenario in which the organization should protect devices that capture payment card data via direct physical interaction with the card from tampering and substitution.
- (III) How are network vulnerability scans and penetration testing helpful to meet the control objective as regularly monitor and test networks of PCI-DSS? Also discuss the frequency of network vulnerability scans and penetration testing to achieve the requirement of PCI-DSS.
- (IV) Briefly discuss the roles and responsibility of Core Auditor. What are the differences between the roles of an Auditor and Administrator?

Printed Pages: 2

University Roll No:

End Term Examination, Odd Semester 2021-22

B.Tech, CSE(DA), IIIrd Year, Vth Semester

BCSE0558: Enterprise Business Intelligence and Data Warehousing

Time: 3 Hour

Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

5 x 4 = 20 marks

- (I) Differentiate Database with Data Warehouse and also Explain Characteristics of Data Warehouse.
 - (II) Explain Star Schema with its any two advantages and also explain any scenario where it should be used instead of any other available schema?
 - (III) Explain factors to construct Data warehouse instead of RDBMS for their analysis in Business?
 - (IV) What would be the recommended approaches that you would suggest as BI architect to design any BI system .
 - (V) Differentiate between Slice and Dice operations with example?
 - (VI) Compare Ralf Kimball and Inmon approach and also justify which approach is better for given IT oriented conditions.

Section- B

Note: Attempt ALL Questions.

5 x 3 = 15 marks

- (I) Discuss the need of Data Profiling and also explain its various types also with any suitable real time example.
 - (II) Explain various dimensions of Data Quality with any real time example

- (III) "Meta data is necessary for Data warehouse" Justify this statement with few important reasons
- (IV) What do you understand by cloud computing? Comment on Business Intelligence association with Cloud Computing in the context of Real Time Industry
- (V) Draw and explain Metadata Framework of Metadata Management System and also explain various types of Meta Data.

Section- C

Note: Attempt ANY THREE Questions.

3 x 5 = 15 marks

- (I) Krishna is working on project of developing a data warehouse for E- Commerce Website which should consist at least four dimensions: Product Category, Product Name ,Shipping Details ,Product Company and two measures count and avg_sale .Draw Snowflake schema for the same.
- (II) Radhika is working on a Data Ware house for a University, which should consist of at least four dimensions: student, semester, course and instructor, and two measures count and avg_grade. Draw star and fact constellation schema diagram for the above Data Warehouse.
- (III) Draw BI Sample architecture and also explain various types of Digital Data with any case study.
- (IV) Discuss Need of OLTP and Compare OLAP and OLTP and also explain various operations of OLAP Data warehouse Cube.

Printed Pages: 1

University Roll No:

End Term Examination, Odd Semester 2021-22

B.Tech, CS-IoT, III, V

BCSE0654: Ipv6 Analysis and Applications

Time: 3 Hour

Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

$5 \times 4 = 20$ marks

- (I) What are the main purpose of IPv6 addressing and what are the role IPv6 in IoT?
- (II) Write any four differences between IPv4 and IPv6?
- (III) What are the main role of extension header?
- (IV) Write about all three categories of IPv6 addressing.
- (V) What are the possible ways to represent the IPv6 address?
- (VI) What are the categorization of Unicast addresses?

Section- B

Note: Attempt ALL Questions.

$5 \times 3 = 15$ marks

- (I) What are the possible representation of given IPv6 address
2001:DB8:0000:0056:0000:ABCD:EF12:1234.
- (II) It is possible to balance the load of traffic in Anycast Addressing, give your explanation.
- (III) What are the purpose of Multicast Addressing in IPv6?
- (IV) What are the main challenges to design a routing protocol for IPv6 addressing?
- (V) What are the main purpose of Routing Table in routing protocol?

Section- C

Note: Attempt ANY THREE Questions.

$3 \times 5 = 15$ marks

- (I) How RIPng routing protocol works, explain all steps with flow chart?
- (II) Draw and explain the message format of RIPng and also write the main limitation of RIPng protocol?
- (III) What are the differences between OSPF for IPv4 and OSPF for IPv6? What is OSPFv3?
- (IV) Explain 6LoWPAN protocol in detail. Write about IP Header compression in 6LoWPAN.

Printed Pages: 03

University Roll No:

End Term Examination, Odd Semester 2021-22

Program: B.Tech. (CSE) Year: III Semester: V
Subject Code:BCSC0012, Subject: Design & Analysis of Algorithms
Time: 3 Hours Maximum Marks: 50

Section- A

Note: Attempt ANY FIVE Questions.

5 x 4 = 20 marks

- (I) Describe the algorithm analysis for merge sort in best case and worst case.
 - (II) Write properties of B-tree. Explain 2-3 trees.
 - (III) For $T(n) = 7T(n/2) + 9n^2$ Solve the recurrence relation and find the time complexity.
 - (IV) Explain matrix chain multiplication with suitable example.
 - (V) Find all possible set of feasible solutions for 4-queens problem using backtracking strategy.
 - (VI) Explain graph coloring problem with suitable example.

Section- B

Note: Attempt ALL Questions.

$5 \times 3 = 15$ marks

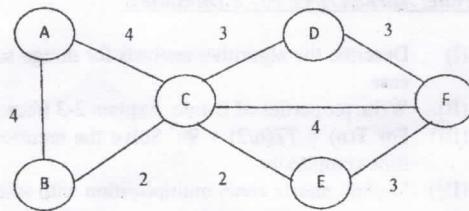
- (I) Discuss the concept of optimal substructure and overlapping sub problems in dynamic programming.
 - (II) Consider following instance of knapsack problem:
 $n=3, m=12, (p_1, p_2, p_3) = (2, 4, 10)$ and $(w_1, w_2, w_3) = (4, 6, 8)$
Find the optimal solution for given problem and list all feasible solutions.
 - (III) Explain sum of subset problem.
 - (IV) Compare Dijkstra, Bellman Ford and Floyd-Warshall algorithm.
 - (V) Find a solution for the Activity Selection problem using a suitable example.

Section-C

Note: Attempt ANY THREE Questions.

$3 \times 5 = 15$ marks

- (I) Construct minimum cost spanning tree for the given graph and compute cost using Prim's and Kruskal algorithm.



- (II) Consider a scenario in which 4 items are given along with weights and values/profits as follows. Solve the following instance of 0/1 knapsack using dynamic programming. The capacity of knapsack W is 10.

Item	1	2	3	4
Weight	4	7	5	3
Value/Profit	40	42	25	12

- (III) Apply Floyd-Warshall algorithm to find shortest path between every pair of vertices in the given graph.

0	∞	3	∞
2	0	∞	∞
∞	7	0	1
6	∞	∞	0

- (IV) How branch and bound technique differ from dynamic Programming?
Solve travelling salesman problem using branch & bound technique
for following graph where A, B, C and D represent different cities
which are visited by the salesman.

