Year - April 2024

TYBSC-CS

SEM - VI

No. of pages - 15

CS

TYBSC - CS - Sem VI

(2 ½ Hours) [Total Marks: 75]

N.B. 1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

Q. 1 Attempt All

(a) Select the correct alternative from the options given: **(10M)**

(i) One of the main issues of IR is _____
(a) Scale (b) Size
(c) Quality (d) Relevance

(ii) Dictionary data structure can be _____.
(a) Hash table and Tree (b) Linked list and Tree
(c) Arrays (d) Stack

(iii) _____ was named after the Larry page , one of the founders of Google.
(a) Page Rank (b) Ranking
(c) Permuterm (d) Boolean Retrieval

(iv) _____ function is responsible for consolidating the results produced by each of the map() function
(a) Mapper() (b) Reducer()
(c) Reduce() (d) Map()

(v) PPC stands for _____
(a) Pay per context (b) Pay Per customer
(c) Pay per click (d) Pay per content

(vi) SERP stands for _____
(a) Search Engine Records Page (b) Search Engine Result Page
(c) Search Engine Ranked Page (d) Search Engine Repeated Page

(vii) Web search engines stores the information about many web pages by a _____.
(a) Web Indexer (b) Web Organizer
(c) Web Router (d) Web Crawler

8800096A7F25342DD6B57498F1F38855

(viii) The _____ is categorized as an unknown segment of the Deep Web which has been purposely kept hidden & is inaccessible using standard web browsers.

(a) Haunted web
(b) Dark web
(c) World Wide Web
(d) Surface web

(ix) _____ filtering recommends the products which are similar to the ones the user has liked in the past.

(a) Context Based filtering
(b) Collaborative filtering
(c) Content based filtering
(d) None of these

(x) Recommendation system uses _____.

(a) User's History of search
(b) User's Profile
(c) Both users history and users profile
(d) Neither users history nor users profile

(b) **Fill in the blanks by selecting from the pool of options:** **(5M)**
(position, surface web, Isolated term distance, 1, Corpus, degree, dark web, Levesthein distance, Lexicon, 0)

(i) The normal google search is a part of _____

(ii) Given the two strings s1 and s2 the edit distance between them is known as _____

(iii) A similarity measure is a function that computes the _____ of similarity between two vectors.

(iv) A dictionary of terms is sometime also referred as _____.

(v) Initially Page Rank of all web pages is _____.

Q. 2 Attempt the following (Any THREE) **(15M)**

(a) Define Information Retrieval. Discuss and differentiate between the terms precision and recall with respect to information retrieval.

(b) Discuss the issues and challenges in information retrieval.

(c) Draw the inverted index for following document.
Doc1-new home sales top forecast
Doc2-home sales rise in july
Doc3-increase in home sales in july
Doc4-july new home sales rise

(d) Explain how k-gram index can be used for spelling correction.

(e) With the help of a diagram, explain the components of an information retrieval system.

(f) Write a short note on phonetic correction.

8800096A7F25342DD6B57498F1F38855

Q. 3	**Attempt the following (Any THREE)**	**(15M)**
(a)	Write a short note on Cross - Lingual Information Retrieval.	
(b)	Explain the concept of collaborative filtering, with the help of an example.	
(c)	Discuss the HITS Algorithm in detail.	
(d)	What is invisible web? Explain in detail.	
(e)	What is text summarization? Discuss the stages of summarization.	
(f)	Write a short note on the Page Rank Algorithm.	

Q. 4	**Attempt the following (Any THREE)**	**(15)**
(a)	Explain the web search architecture.	
(b)	State and explain the challenges in XML retrieval.	
(c)	Discuss the characteristics of web search engine?	
(d)	Write a short note on Search Engine Optimization.	
(e)	Explain the vector space model for XML retrieval.	
(f)	List and explain various SPAM Techniques.	

Q. 5	**Attempt the following (Any FIVE)**	**(15)**
(a)	Compare Information retrieval and Data retrieval.	
(b)	Define following terms i) Posting list ii) Corpus iii) Lexicon	
(c)	Differentiate between Content based and Collaborative Filtering	
(d)	Explain the following in brief: i) hubs ii) authorities	
(e)	Discus the challenges in Question and Answering.	
(f)	Define the following i) cloaking ii) stopping iii) stemming	
(g)	Discuss the role of XML in Information Retrieval.	
(h)	Compare Text Centric vs Data Centric XML.	

****

8800096A7F25342DD6B57498F1F38855

(2 ½ Hours) [Total Marks: 75]

N.B. 1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

Q. 1 Attempt All

(a) Select the correct alternative from the options given: (10M)

(i) In _____ protocols, transmission schedules can be comp such that there will be no collision at receiver

(a) Contention based (b) scheduled based
(c) ARQ (d) none of these

(ii) Multihop network operates on _____ principle.

(a) store packet (b) store and forward packet
(c) forward packet (d) receive and update packet

(iii) Frequency of ground wave is _____.

(a) >30MHz (b) 2-130 MHZ
(c) < 2 MHz (d) >200MHz

(iv) A transceiver that is ready to receive but not currently receiving anything is called _____ state.

(a) sleep (b) idle
(c) receive (d) transmit

(v) A delay of detection of error and reporting it to all / any interested nodes is called _____.

(a) event detection delay (b) processing delay
(c) turn delay (d) event correction delay

(vi) Dwell time is the time for_____.

(a) Hand off (b) A call within the cell
(c) Waiting for channel allocation (d) Survival

(vii) Antenna having equal radiations in all directions is called _____

(a) Isotropic (b) Directional
(c) Dipole (d) Monopole

648CAF94A418F5E0DDB9E9E64563DE3E

(viii) _____ supports modularity and event-based programming by the concept of components.

(a) nesC (b) TinyOS
(c) TOSSIM (d) NETSIM

(ix) A satellite signal transmitted from a satellite transponder to earth's station is _____.

(a) uplink (b) downlink
(c) terrestrial (d) earthbound

(x) IMEI stands for ____

(a) Internal Mobile Equipment Identity (b) International Mobile Equipment Identity
(c) Intra Mobile Enable Identity (d) International Mobile Enable Identity

(b) **Fill in the blanks by selecting from the pool of options:** (5M)
(Power Amplifier, sink, Noise figure, Gain, Controller unit, Guard bits)

(i) The _____ is the ratio of the output signal power to the input signal power.

(ii) An entity where information is needed is called _____.

(iii) The _____ accepts upconverted signals from the IF or baseband part and amplifies them for transmission over the antenna.

(iv) The core of WSN is _____.

(v) _____ are utilized to allow synchronization of the receivers between different slots and frames.

Q. 2 Attempt the following (Any THREE) (15M)

(a) State and explain any 5 characteristics of Transceiver.
(b) Explain the need of gateways in WSN.
(c) What is a MANET? Explain important characteristics of MANETs.
(d) Explain different applications of WSN.
(e) Define Wireless Sensor Network. Write a short note on advantages of WSN.
(f) What are the characteristic requirements of WSN?

648CAF94A418F5E0DDB9E9E64563DE3E

Q. 3	**Attempt the following (Any THREE)**	(15M)
(a)	Describe in brief S-MAC	
(b)	What is Low Energy Adaptive Clustering Hierarchy(LEACH)?State its advantages and disadvantages	
(c)	How Power efficient gathering in sensor information systems (PEGASIS) routing protocol works?	
(d)	What are different routing strategies used in Wireless Sensor Networks?	
(e)	Explain in brief Fixed Assignment MAC protocols used in WSN	
(f)	Explain Sensor Protocols for Information via Negotiation (SPIN).	

Q. 4	**Attempt the following (Any THREE)**	(15)
(a)	What is Frequency hopping and Direct Sequence Spread Spectrum techniques?	
(b)	Explain any two types of Multiplexing.	
(c)	Explain DECT system architecture reference model.	
(d)	What is handover? Explain types of handovers in satellite system?	
(e)	Explain the types of Satellites?	
(f)	Write short note on GPRS architecture.	

Q. 5	**Attempt the following (Any FIVE)**	(15)
(a)	Differentiate between Ad-hoc networks and Wireless Sensor Networks.	
(b)	With neat diagram explain Sensor node architecture.	
(c)	Explain Piconet and Scatternet.	
(d)	Explain feasibility of using TCP/UDP for WSN.	
(e)	What are different types of antennas? Explain any two.	
(f)	Explain GSM architecture with diagram.	
(g)	List and explain the types of Mobility.	
(h)	What is modulation? State and explain its types.	

****

648CAF94A418F5E0DDB9E9E64563DE3E

T4BSC CS Sem VI

(7)

(2½ Hours)

[Total Marks: 75]

N B 1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

1. Attempt *any four* of the following: **20**

a. What is information retrieval example? What are the characteristics of information retrieval.

b. What are the components and What are the major challenges faced in Information Retrieval.

c. What is edit distance, and how is it used in measuring string similarity with suitable example.

d. Explain the process of constructing an inverted index. How does it facilitate efficient information retrieval?

e. What is relevance feedback in the context of retrieval models.

f. Explain Vector space model. Discuss TF-IDF, cosine similarity.

2. Attempt *any four* of the following : **20**

a. Define text categorization and explain its importance in information retrieval systems.

b. How can clustering be utilized for query expansion and result grouping in information retrieval systems.

c. Explain the effectiveness of K-means and hierarchical clustering in text data analysis.

d. Explain the architecture of a web search engine. What are the components involved in crawling and indexing web pages.

e. What is the role of supervised learning techniques in learning to rank and their impact on search engine result quality.

f. Discuss the difference between the PageRank and HITS algorithms.

3. Attempt *any four* of the following : **20**

a. Explain breadth-first and depth-first Web page crawling Techniques?

b. Define near-duplicate page detection and its significance in web search. Explain the challenges associated with identifying near-duplicate pages.

c. Describe common techniques used in extractive text summarization.

d. What are Challenges associated with question answering.

e. Define collaborative filtering and content-based filtering in recommender systems.

f. Explain different approaches to machine translation, including rule-based, statistical, and neural machine translation models.

X51Y4D804BX51Y4D804BX51Y4D804BX51Y4D804B

4. **Attempt *any five* of the following :**

a Discuss the steps involved in the Soundex Algorithm for phonetic matching.

b Construct 2-gram, 3-gram and 4-gram index for the following terms:
a. banana
b. pineapple
c. computer

c Discuss the Naive Bayes algorithm for text classification. How does it work, and what are its assumptions.

d Discuss how link analysis can be used in social network analysis and recommendation systems.

e Discuss challenges in abstractive text summarization.

f Describe the role of test collections and benchmarking datasets in evaluating IR systems.

X51Y4D804BX51Y4D804BX51Y4D804BX51Y4D804B

(9)

TYBSc - Sem VI - CS

(2 ½ Hours) **[Total Marks: 75]**

N.B. 1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

Q. 1 **Attempt ANY FOUR from the following:** **(20M)**

(a) Explain the structure of SOAP message.
(b) Write a web service method that accepts two numbers as parameters and returns the largest of those numbers using JAX-WS. (Write only the Webservice method block)
(c) Differentiate between Parallel Computing and Distributed Computing.
(d) List and explain the various Http methods required for creating RESTful Web Services.
(e) Explain the different characteristics of Virtualized environment.
(f) Explain virtualization using KVM.

Q. 2 **Attempt ANY FOUR from the following:** **(20M)**

(a) Explain the advantages of Cloud Computing.
(b) Describe the Cloud Computing Reference Model in detail.
(c) Explain open challenges of Cloud Computing.
(d) Write about various Security Services of Cloud.
(e) Explain the cloud security design principles.
(f) Explain Cloud Security Policy Implementation.

Q. 3 **Attempt ANY FOUR from the following:** **(20M)**

(a) Explain CloudSim architecture in detail. Also draw the necessary diagram for the same.
(b) Explain different features of CloudSim.
(c) Write about the working platform for CloudSim.
(d) Explain the key components of OpenStack?
(e) What is DevStack? Explain the installation steps.
(f) Explain Components and services of AWS.

Q. 4 **Attempt ANY FIVE from the following:** **(15M)**

(a) Write in short about JAX-WS.
(b) What are the type of Cloud? Write in short about any one
(c) Write a short note on GridSim.
(d) Describe oVirt in short.
(e) What are the security best practices for AWS.
(f) What is an OpenStack.

****

X51Y063165X51Y063165X51Y063165X51Y063165

(10)

TYBSC - Sem VI - CS

(2 ½ Hours) **[Total Marks: 75]**

N.B. 1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

Q. 1 Attempt ANY FOUR from the following: **(20M)**

(a) Explain the structure of SOAP message.
(b) Write a web service method that accepts two numbers as parameters and returns the largest of those numbers using JAX-WS. (Write only the Webservice method block)
(c) Differentiate between Parallel Computing and Distributed Computing.
(d) List and explain the various Http methods required for creating RESTful Web Services.
(e) Explain the different characteristics of Virtualized environment.
(f) Explain virtualization using KVM.

Q. 2 Attempt ANY FOUR from the following: **(20M)**

(a) Explain the advantages of Cloud Computing.
(b) Describe the Cloud Computing Reference Model in detail.
(c) Explain open challenges of Cloud Computing.
(d) Write about various Security Services of Cloud.
(e) Explain the cloud security design principles.
(f) Explain Cloud Security Policy Implementation.

Q. 3 Attempt ANY FOUR from the following: **(20M)**

(a) Explain CloudSim architecture in detail. Also draw the necessary diagram for the same.
(b) Explain different features of CloudSim.
(c) Write about the working platform for CloudSim.
(d) Explain the key components of OpenStack?
(e) What is DevStack? Explain the installation steps.
(f) Explain Components and services of AWS.

Q. 4 Attempt ANY FIVE from the following: **(15M)**

(a) Write in short about JAX-WS.
(b) What are the type of Cloud? Write in short about any one
(c) Write a short note on GridSim.
(d) Describe oVirt in short.
(e) What are the security best practices for AWS.
(f) What is an OpenStack.

****

X51Y063165X51Y063165X51Y063165X51Y063165

TY BSc - CS - Sem - VI

(2 ½ Hours) [Total Marks: 75]

N.B. 1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

Q. 1 Attempt ANY FOUR from the following: **(20M)**

(a) Explain the Types of Ethical Hacking
(b) Explain Hacking Technology and its types in detail
(c) Define Hacktivism and explain ways to manifest it
(d) Explain any 5 Hacker Classes
(e) What are the ways to conduct Ethical Hacking
(f) Write a short note on foot printing

Q. 2 Attempt ANY FOUR from the following: **(20M)**

(a) Ex[plain Active and Passive Sniffing
(b) Explain ARP Poisoning in Detail
(c) Explain DNS Spoofing Techniques in detail
(d) Explain the working of DOS attack in detail
(e) Explain Smurf attack in detail
(f) Explain types of Session Hijacking

Q. 3 Attempt ANY FOUR from the following: **(20M)**

(a) Explain methods involved in Google Hacking
(b) Define the term Authentication and its types
(c) Define SQL injection and names its types
(d) Explain Buffer –Overflow and its types.
(e) Explain WEP in detail
(f) Explain the working of wireless sniffing

Q. 4 Attempt ANY FIVE from the following: **(15M)**

(a) Explain the methods to perform information Gathering
(b) Define Password and its types
(c) Explain Mutation Techniques
(d) Define the terms (i)Scanning (ii)Enumeration
(e) What is Web Server Hardening?
(f) Write a short note on Rogue Access Point

****

X51Y30A04EX51Y30A04EX51Y30A04EX51Y30A04E

(12)

(2 ½ Hours) [Total Marks: 75]

N.B. 1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

Q. 1 Attempt ANY FOUR from the following: **(20M)**

(a) Define Data Science. Discuss in detail applications of Data Science.
(b) Explain Data Warehousing (DW) and Data Mining (DM) in detail.
(c) Explain in detail different Data Sources.
(d) Describe in detail Data Transformation.
(e) Define Data Wrangling. Discuss Data Wrangling Techniques.
(f) Discuss about Feature Engineering and Time Series Data.

Q. 2 Attempt ANY FOUR from the following: **(20M)**

(a) Explain Data Visualization techniques in detail.
(b) Define Descriptive Statistics. Explain Mean, Median, Mode and Standard Deviation in detail.
(c) Explain in detail Classification and Regression analysis.
(d) Define bias, variance and discuss about bias-variance tradeoff.
(e) Discuss different techniques for evaluating model performance.
(f) Explain Ensemble Learning with Bagging and Boosting.

Q. 3 Attempt ANY FOUR from the following: **(20M)**

(a) Explain storytelling in analysis in detail.
(b) Discuss visualization tools in detail
(c) List and Discuss Data Management Activities.
(d) Elaborate the concept of Data Governance.
(e) Illustrate Extraction, Transformation and Load (ETL) in detail.
(f) Give the importance of Data Quality.

Q. 4 Attempt ANY FIVE from the following: **(15M)**

(a) Differentiate between structured and unstructured data.
(b) Explain Hyperparameter Tuning.
(c) Define Accuracy and Explain in brief.
(d) Discuss any three libraries of Data Science.
(e) Differentiate between underfitting and overfitting.
(f) Define Precision, Recall and F1-Score.

****

X51Y0434F3X51Y0434F3X51Y0434F3X51Y0434F3

(19)

Q. 4 **Attempt the following (Any THREE)** **(15)**

(a) What are the forensic duplication tool requirements?

(b) Explain different types of reports.

(c) What are the offences and the punishment in IT ACT 2000?

(d) List the guidelines to document and prepare evidence.

(e) What is deposition? Explain its types and any two guidelines for testifying at a deposition.

(f) Define evidence. What are the types of evidence?

Q. 5 **Attempt the following (Any FIVE)** **(15)**

(a) Explain the Investigation Triad.

(b) Explain the process of computer forensics.

(c) Write a short note on Email Servers.

(d) Discuss Email spoofing.

(e) Define the following terms as per IT Act:

i) Access

ii) Addressee

iii) Certifying Authority

(f) Describe the process of restoring the forensic duplicate

(g) Explain how to give evidence in court.

(h) What are the requirements to set up a workstation for computer forensics?

****

2A989FF6B32CF6AC02004EF9B07D9FB7

TYBSC - CS - Sem VI

(2 ½ Hours) [Total Marks: 75]

1) All questions are compulsory.
2) Figures to the right indicate marks.
3) Illustrations, in-depth answers and diagrams will be appreciated.
4) Mixing of sub-questions is not allowed.

Q. 1 Attempt ANY FOUR from the following: **(20M)**

(a) Describe the framework of Data Warehouse with suitable diagram.

(b) How would you describe Dimensional Model. Also explain steps to design Dimensional Model.

(c) Compare OLTP V/S OLAP.

(d) Elaborate parts of Enterprise Data Warehouse (EDW).

(e) Design a Star Schema for Data Warehouse of Bank Management System.

(f) When do you prefer Hybrid Data Warehouse model? Justify your answer.

Q. 2 Attempt ANY FOUR from the following: **(20M)**

(a) What kind of data can be mined form Data Mining?

(b) Describe various Data Mining metrics.

(c) What do you understand from Statistical Description of data.

(d) Why is data cleaning needed?

(e) For the given following data, apply Apriori algorithm. (Min. Support= 50%, Confidence threshold = 70%)

TID	Items
100	1 3 4
200	2 3 5
300	1 2 3 5
400	2 5

Dataset D

(f) How will you construct FP growth tree for the following transactions containing set of item = {a,b,c,d……} (min support=3)

T_id	Items bought
1	f,a,c,d,g,i,m,p
2	a,b,c,f,l,m,o
3	b,f,j,h,o
4	b,c,k,s,p
5	a,f,c,e,l,p,m,n

X51Y578609X51Y578609X51Y578609X51Y578609

15

Q. 3 Attempt ANY FOUR from the following: **(20M)**

(a) Differentiate between classification and prediction

(b) Explain linear regression with example.

(c) How can measure performance of classifier? Explain in brief.

(d) List the steps involved in k-means clustering algorithm.

(e) Apply ID3 on the following Training Dataset from the given database and extract the classification from the tree.

Age	Income	Student	Credit_Rating	Class:buys_computer
<=30	High	No	Fair	No
<=30	High	No	Excellent	No
31...40	High	No	Fair	Yes
>40	Medium	No	Fair	Yes
>40	Low	Yes	Fair	Yes
>40	Low	Yes	Excellent	No
31....40	Low	Yes	Excellent	Yes
<=30	Medium	No	Fair	No
<=30	Low	Yes	Fair	Yes
>40	Medium	Yes	Fair	Yes
<=30	Medium	Yes	Excellent	Yes
31...40	Medium	No	Excellent	Yes
31...40	High	Yes	Fair	Yes
>40	Medium	No	Excellent	No

(f) What are some real-world applications where clustering algorithms play a significant role in different industries and domains?

Q. 4 Attempt ANY FIVE from the following: **(15M)**

(a) Elaborate features of Data warehouse.

(b) Define attribute and Enlist its types

(c) What are the requirements a clustering algorithm should satisfy?

(d) What can you say about granularity of facts?

(e) Explain the terms Support ,Confidence and Risk in Association Rule metrics

(f) State the applications of Classification.

****

X51Y578609X51Y578609X51Y578609X51Y578609