



55

MCKV Institute of Engineering

Paper Code: PE-IT602A

Paper Name: Artificial Intelligence

Time: 3 Hours

Full Marks: 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any **ten** of the following: $10 \times 1 = 10$
 - i. The time complexity and space complexity for bidirectional breadth first search technique, respectively are (with branching factor b and depth d)
 - a) $(b^d), O(b^d)$
 - b) $(b^{d/2}), O(b^{d/2})$
 - c) $(b^{d/2}), O(b^d)$
 - d) $(b^d), O(b^{d/2})$
 - ii. A process of eliminating existential quantifier is known as
 - a) Resolution
 - b) Skolemisation
 - c) Unification
 - d) None of these
 - iii. What are Agent?
 - a) An agent is anything that can perceive its environment through sensors
 - b) An agent is anything that can change its environment through sensors
 - c) An agent is anything that can control its environment through sensors
 - d) None of the Above
 - iv. Searching techniques are used for
 - a) Goal node searching
 - b) Optimization of search space
 - c) Finding goal distance of the goal node from start node
 - d) All of these.
 - v. NLP (with respect of AI) stands for
 - a) Natural Linear Processing
 - b) Natural Language Processing
 - c) Natural Linear Programming
 - d) Natural Language Programming.
 - vi. What is the full form of PEAS?
 - a) Perceptual Measure, Environment, Actuators, and Sensors
 - b) Performance Measure, Environment, Actuators, and Sensors
 - c) Performance Measure, Entity, Actuators, and Sensors
 - d) Performance Measure, Environment, Agent Function, and Sensors

- vii. Uninformed search is also known as
 a) Brute force search b) Hill Climbing search
 c) Worst case search d) Blind search
- viii. A Bayesian network is a
 a) Tree b) Directed graph c) Undirected graph d) None of these
- ix. Lata is slightly ill". The statement can be complexly expressed in
 a) FOPL b) Propositional logic c) Fuzzy logic d) None of these
- x. Rationality of an agent does not depend on?
 a) performance measures b) Percept Sequence c) reaction d) actions
- xi. The term 'Optimality', so far one of the performance measuring indices of any search technique is concerned, refers to
 a) Time complexity b) Space complexity c) Both (a) and (b) d) None of these.
- xii. How the new state is generated in genetic algorithm?
 a) Compression b) mutation c) Cross over d) Both a) and b)

Group - B

(Short Answer Type Questions)

Answer any *three* of the following

3×5=15

1. Explain when does simulated annealing algorithm behave like hill climbing?
 [Module3/C02/Analyze/HOCQ] [5]
2. Compare blind search and heuristic search.
 [Module3/C02/Understand-LOCQ] [5]
3. Genetic algorithm, how do you obtain the new chromosome (solution) from the old one?
 [Module3/C02/Analyze/HOCQ] [5]
4. Write down the differences between conventional set and fuzzy set.
 [Module3 /C02/Understand/IOCQ] [5]
5. Discuss how you evaluate any search technique.
 [Module3/C02/Understand/IOCQ] [5]



Group - C

(Long Answer Type Questions)

Answer any *three* of the following

3×15=45

6. a) Describe DFS with suitable example.
 [Module3/C02/Remember /LOCQ] [5]
- b) When would best-first search be worse than simple breadth-first search?
 [Module3/C02/Understand/IOCQ] [5]

c) Discuss on Bidirectional search technique.

[Module3/C02/Understand/IOCQ]

[5]

8. Consider the 3-puzzle problem shown in figure 1.

2	3
1	

Initial

1	2
3	

Final

Figure 1

Possible operators (in order) are: up, down, left, right; Assume that repeated states are not detected.

a) Draw the search tree using breadth first search.

b) Would depth first search find the goal? Explain.

c) How many nodes would be generated if Iterative Deepening is used starting with depth increment one?

[Module3/C02/Apply/HOCQ]

[7+4+4]

9. a) Consider a water-jug problem. You are given two jugs: a 4-gallon one and a 3-gallon one. Neither has any measuring mark on it. There is a pump that can be used to fill the jugs with water. How can you get exactly 2 gallons of water into the 4-gallon jug? Discuss the Water Jug problem as state space search.

[Module1/C01/Apply/IOCQ]

[7]

b) Discuss how different types of agents are classified?

[Module2/C01/Remember/LOCQ]

[5]

c) Discuss about the environment of agents?

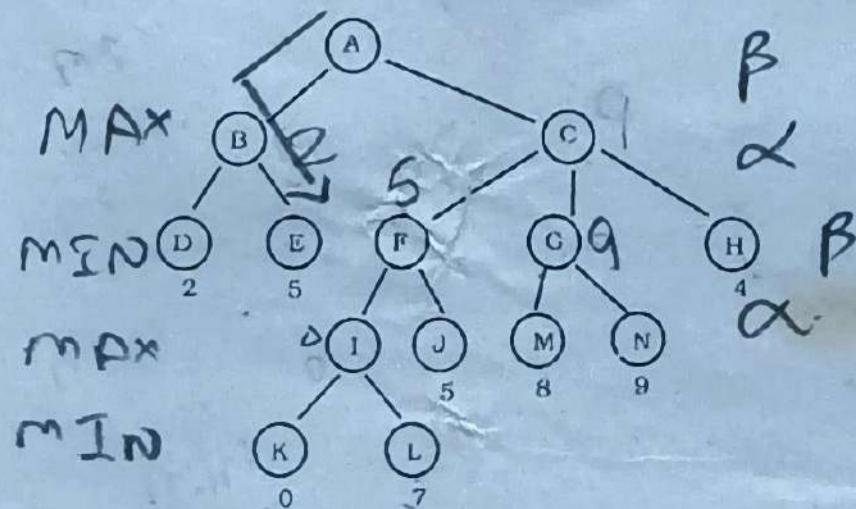
[Module2/C01/Remember/LOCQ]

[3]

10. Consider the following game tree in which static scores are all from first player's point of view.

a) Which would be his best first move if MINIMAX algorithm is used?

b) Which branches will be pruned if α - β pruning algorithm is used?



[Module3/C02/Apply/HOCQ]

11. Write Short notes on **any three** of the following :-

[3×5]

- ✓ a) Learning
- b) Dempster-Shafer theory.
- ✗ c) A* algorithm.
- ✓ d) Alpha-beta Cutoffs
- ✓ e) Natural Language Processor.

[Module6/C04/Remember/LOCQ]
[Module5/C03/Remember/LOCQ]
[Module3/C02/Remember/LOCQ]
[Module3/C02/Remember/LOCQ]
[Module3/C02/Remember/LOCQ]



Time Allotted: 3 Hours

Full Marks: 70

*The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable.*

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any **ten** of the following:

(i) What is the full form of PEAS?

 - a) Perceptual Measure, Environment, Actuators, and Sensors
 - b) Performance Measure, Environment, Actuators, and Sensors
 - c) Performance Measure, Entity, Actuators, and Sensors
 - d) Performance Measure, Environment, Agent Function, and Sensors

(ii) Which of the following is not Properties of Environment?

 - a) Discrete / Continuous
 - b) Static / Dynamic
 - c) Deterministic / Non-deterministic
 - d) No agent / Multiple agents

(iii) Which instruments are used for perceiving and acting upon the environment?

 - a) Sensors and Actuators
 - b) Sensors
 - c) Perceiver
 - d) Perceiver and Sensor

(iv) An AI system is composed of?

 - a) agent
 - b) environment
 - c) Both a) and b)
 - d. None of these

(v) What are Agent?

 - a) An agent is anything that can perceive its environment through sensors
 - b) An agent is anything that can change its environment through sensors
 - c) An agent is anything that can control its environment through sensors
 - d) None of the Above

(vi) Rationality of an agent does not depends on?

 - a) performance measures
 - b) Percept Sequence
 - c) reaction
 - d) actions

(vii) Searching techniques are used for

 - a) Goal node searching
 - b) Optimization of search space
 - c) Finding goal distance of the goal node from start node
 - d) all of these

(viii) NLP (with respect of AI) stands for

- a) Natural Linear Processing b) Natural Language Processing
c) Natural Linear Programming d) Natural Language Programming

(ix) Algorithm that gives optimal solution

- a) Hill climbing b) BFS c) Blind search d) A*

Fuzzy logic is a form of

- a) Two valued logic b) Crisp set logic
b) Any valued logic d) Binary set logic

Which is NOT a heuristic search

- a) Steepest Ascent Hill Climbing b) A* search
b) Depth first search d) Simulated annealing

Informed Search is also known as

- a) Greedy search b) Hill Climbing search c) Blind search d) None of these

Group - B

(Short Answer Type Questions)

Answer any three of the following

10. a) Components of AI production system.

[Module 1/001/Understand-100Q]

$$3 \times 5 = 15$$

b) Difference between Goal-Based Agent & Utility-Based Agent.

[Module 1/002/Understand-100Q]

$$5$$

c) Contrast Best First and Hill climbing search.

[Module 1/003/Understand-100Q]

$$5$$

d) Overview of knowledge representation methods.

[Module 1/004/Understand-100Q]

$$5$$

e) Steps involved in forward chaining.

[Module 1/005/Understand-100Q]

$$5$$

Group - C

(Long Answer Type Questions)

Answer any three of the following

$$3 \times 15 = 45$$

11. a) What is Artificial Intelligence? Explain how an AI system is different from a conventional computer system.

[Module 2/001/Understand-100Q]

b) Challenges and limitations of artificial intelligence?

[Module 2/002/Understand-100Q]

c) Advantages and disadvantages of hill climbing search procedure.

[Module 2/003/Understand-100Q]

$$(2+4)+5+4$$

b) Consider a maze type problem. You are given two paths. Explain which path is better for any searching mechanism. There is a junction after 2nd step. In which case can you get stuck? Explain the steps for solving the problem as best you can.

[Module 2/004/Apply-100Q]

c) Discuss no agents as search procedure.

[Module 2/005/Understand-100Q]

D-2-5-5
d) i) Describe agent. Write down the basic types of the agent.

[Module 2/001/Remember-100Q]

ii) Discuss about the environment of agents.

[Module 2/002/Remember-100Q]

iii) Compare blind search and heuristic search.

[Module 3/002/Understand-100Q]

10. a) Represent the following using FOL [First Order Logic]

i. All birds fly.

ii. Every man respects his parent.

iii. Some boys play cricket.

iv. Not all students like both Mathematics and Science.

v. Only one student failed in Mathematics.

[Module 4/003/Apply-100Q]

b) What are the issues involved in relating an appropriate knowledge representation scheme to a particular application?

[Module 4/003/Understand-100Q]

D-2-5-5

11. a) Describe Dempster-Shafer theory

[Module 5/003/Remember-100Q]

b) Describe production system with an example.

[Module 5/003/Understand-100Q]

c) Discuss the benefits of production system.

[Module 5/003/Remember-100Q]

5

-

10

-

15

-

MCKV Institute of Engineering
Paper Code: PE-IT602A
Paper Name: Artificial Intelligence

Time Allotted: 1 Hour**Full Marks: 30**

*The figures in the margin indicate full marks.
 Candidates are required to give their answers in their own words as far as practicable.*

Group - A**(Multiple Choice Type Questions)**

1. Choose the correct alternatives for any **five** of the following: **5×1=5**
- Uninformed search is also known as
 - Brute force search b. Hill Climbing search c. Worst case search d. ✓ Blind search
 - An Algorithm that gives optimal solution is
 - Hill Climbing
 - BFS
 - Blind search
 - ✓ A*
 - What is meant by simulated annealing in Artificial Intelligence?
 - Return optimal solution when there is a proper cooling schedule
 - Return optimal solution when there is no proper cooling schedule
 - It will not return optimal solution when there is a proper cooling schedule
 - None of these
 - Depth first search procedure uses
 - AND graph
 - ✓ OR graph
 - AND and OR graph
 - None of these
 - Which search method takes less memory?
 - Breadth first search
 - Depth first search
 - Both a. and b.
 - Linear search
 - The time complexity and space complexity for bidirectional breadth first search technique, respectively are (with branching factor b and depth d)

a. $O(b^d)$	b. $O(b^{d/2})$
c. $O(b^{d/2})$	d. $O(b^d)$

Group - B**(Short Answer Type Questions)**Answer any **two** of the following**2×5=10**

2. Compare blind search and heuristic search. [Module3/CO2/Understand-LOCQ] 5
3. Discuss on Bidirectional search technique. [Module3/CO2/Understand-LOCQ] 5
4. When would best-first search be worse than simple breadth-first search? [Module3/CO2/Understand/IOCQ] 5

Group - C

(Long Answer Type Questions)

$1 \times 15 = 15$

Answer any **one** of the following

5. a. Describe hill climbing algorithm. [Module3/C02/Understand/LOCQ]
- b. Write down the disadvantages of hill climbing search procedure
[Module3/C02/Understand/LOCQ]. [7 + 8]
6. a. Explain when does simulated annealing algorithm behave like hill climbing?
[Module3/C02/Analyze/HOCQ]
- b. Discuss why uninformed search is called blind search?
[Module3/C02/Understand/LOCQ] [7+8]



MCKV Institute of Engineering

Paper Code: PE-IT602A

ARTIFICIAL INTELLIGENCE

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any **five** of the following:

5×1

(i) What is the full form of PEAS?

- a. Perceptual Measure, Environment, Actuators, and Sensors
- b. Performance Measure, Environment, Actuators, and Sensors
- c. Performance Measure, Entity, Actuators, and Sensors
- d. Performance Measure, Environment, Agent Function, and Sensors

(ii) Which of the following is not Properties of Environment?

- a. Discrete / Continuous b. Static / Dynamic
- c. Deterministic / Non-deterministic d. No agent / Multiple agents

(iii) Which Instruments are used for perceiving and acting upon the environment?

- a. Sensors and Actuators b. Sensors c. Perceiver d. Perceiver and Sensor

(iv) An AI system is composed of?

- a. agent b. environment c. Both A and B d. None of the Above

(v) What are Agent?

- a. An agent is anything that can perceive its environment through sensors
- b. An agent is anything that can change its environment through sensors
- c. An agent is anything that can control its environment through sensors
- d. None of the Above

(vi) Rationality of an agent does not depends on?

a. performance measures

b. Percept Sequence

c. reaction

d. actions

Group - B

(Short Answer Type Questions)

Answer any **two** of the following

2×5

2. Explain the main challenges and limitations of artificial intelligence? 5

[Module1/CO1/Understand/LOCQ]

3. Describe Artificial Intelligence? Explain how an AI system is different from a conventional computing system. [Module1/CO1/Understand/LOCQ] [2+3]

4. Describe agent. Write down the disadvantages of table driven agent? [2+3]

[Module2/CO1/Remember/LOCQ]

Group - C

(Long Answer Type Questions)

Answer any **one** of the following

1×15

5. a) Discuss about the environment of agents? [Module2/CO1/Remember/LOCQ] 7

b) Explain the role of performance measure in the design of an intelligent agent, and how is it used to evaluate the success of the agent? [Module2/CO1/Understand/LOCQ] [3+5]

6. a) Consider a water-jug problem. You are given two jugs: a 4-gallon one and a 3-gallon one. Neither has any measuring mark on it. There is a pump that can be used to fill the jugs with water. How can you get exactly 2 gallons of water into the 4-gallon jug? Discuss the Water Jug problem as state space search. [Module1/CO1/Apply/LOCQ] 8

b) Exemplify the necessary component to define an AI problem with example. [Module1/CO1/Understand/LOCQ] 7

Observation
Acquisition
Utility



MCKV Institute of Engineering

Paper Code: PE-IT602A
Artificial Intelligence

Full Marks: 30

Time Allotted: 1 Hour

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

5×1

1. Choose the correct alternatives for any **five** of the following:

- i. Uninformed search is also known as
 - a. Brute force search
 - b. Hill Climbing search
 - c. Worst case search
 - d. Blind search
- ii. The time complexity and space complexity for bidirectional breadth first search technique, respectively are (with branching factor b and depth d)
 - a. $(b^d), O(b^d)$
 - b. $(b^{d/2}), O(b^{d/2})$
 - c. $(b^{d/2}), O(b^d)$
 - d. $(b^d), O(b^{d/2})$
- iii. The term 'Optimality', so far one of the performance measuring indices of any search technique is concerned, refers to
 - a. time complexity
 - b. space complexity
 - c. both (a) and (b)
 - d. none of these
- iv. Agents are
 - a. Autonomous
 - b. Adaptive
 - c. Both a. and b.
 - d. none of these
- v. Searching techniques are used for
 - a. Goal node searching
 - b. Optimization of search space
 - c. Finding goal distance of the goal node from start node
 - d. all of these.
- vi. Which search method takes less memory?
 - a. Breadth first search
 - b. Depth first search
 - c. Both a. and b
 - d. Linear search

Group - B

(Short Answer Type Questions)

Answer any *two* of the following

2×5

2. Discuss Bidirectional search technique. [Module3/CO2/Understand-LOCQ] 5
3. Compare blind search and heuristic search. [Module3/CO2/Understand-LOCQ] 5
4. Write down the disadvantages of hill climbing search procedure [Module3/CO2/Understand/LOCQ]. 5

Group - C

(Long Answer Type Questions)

1×15

Answer any *one* of the following

5. a) Explain DFS Algorithms with an example. [Module3/CO2/Understand/LOCQ] [10+5]
b) Discuss why uninformed search is called blind search?
6.a) Distinguish between Goal-Based Agent & Utility-Based Agent.
[Module2/CO1/Understand/LOCQ]
b) Explain State Space Search. [Module3/CO2/Understand/LOCQ] [10+5]
- state what is it's condition → next step goal agent*



MCKV Institute of Engineering

Paper Code: PE-IT602A

Paper Name: Artificial Intelligence

Full Marks: 30

Time Allotted: 1 Hour

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

5x1=5

1. Choose the correct alternatives for any **five** of the following:

- i. What is the full form of PEAS?
 - a. Perceptual Measure, Environment, Actuators, and Sensors
 - b. Performance Measure, Environment, Actuators, and Sensors
 - c. Performance Measure, Entity, Actuators, and Sensors
 - d. Performance Measure, Environment, Agent Function, and Sensors
- ii. Which instruments are used for perceiving and acting upon the environment?
 - a. Sensors and Actuators b. Sensors c. Perceiver d. Perceiver and Sensor
- iii. An AI system is composed of?
 - a. Agent b. Environment c. Both A and B d. None of the Above
- iv. NLP (with respect of AI) stands for
 - a. Natural Linear Processing b. Natural Language Processing
 - c. Natural Linear Programming d. Natural Language Programming
- v. Which is used to improve the agent's performance?
 - a. Perceiving b. Observing c. Learning d. Sequence
- vi. Rationality of an agent does not depend on?
 - a. Performance measures b. Percept Sequence c. Reaction d. Actions

Group - B

(Short Answer Type Questions)

2x5=10

Answer any **two** of the following

2. What is Artificial Intelligence? [CO1/ Remember /LOCQ] Explain how an AI system is different from a conventional computing system. [CO1/Understand /LOCQ]
- 3 State the Characteristics of an AI problem. [CO1/Analysis /HOCQ]
4. What are the different types of agent? Discuss about the environment of agents? [CO1/Understand /LOCQ]

2+3=5

5

2+3=5

Group - C

(Long Answer Type Questions)

1x15=15

Answer any **one** of the following

- 5.a) What is an agent? [CO1/Remember /LOCQ] What are the disadvantages of table driven agent? [CO1/Understand/LOCQ]
- b) How do you evaluate any search technique? [CO2/ Analysis/HOCQ]
- c) Consider a water-jug problem. You are given two jugs: a 4-gallon one and a 3-gallon one. Neither has any measuring mark on it. There is a pump that can be used to fill the jugs with water. How can you get exactly 2 gallons of water into the 4-gallon jug? Discuss the Water Jug problem as state space search. [CO2/ Create /HOCQ]
6. a) What is the role of performance measure in the design of an intelligent agent [CO1/Understand /LOCQ] and how is it used to evaluate the success of the agent? [CO1/Evaluate /HOCQ]
- b) Demonstrate on 'agents as search procedure'? [CO1/ Apply / IOCQ]

5+5+5=15

3+5+7=15



MCKV Institute of Engineering

Paper Code: OE-CS601H

Paper Name: E-Commerce and ERP

Full Marks: 70

Time Allotted: 3 Hours

The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

$10 \times 1 = 10$

1. Choose the correct alternatives for any **ten** of the following:
- (i) What is the abbreviation of CGI?
a) Computer Gateway Interface b) Common Gateway Interface
c) Computer Gateway Interconnection d) Common Gateway Interconnection
- (ii) Which of the following is not a key element of the business model?
a) Competitive advantage b) Market strategy c) Universal standards d) Value Proposition
- (iii) _____ is a function of E commerce.
a) Marketing b) Supply Chain c) Finance d) All of these
- (iv) What is 'EOQ' in Inventory control?
a) Economic Occurring Quantity b) Economic Order Quantity
c) Economic Over Quantity d) Economic of Quantity
- (v) The initial stage of the supply chain process is the _____.
a) Sourcing Stage b) Organizing Stage c) Planning stage d) Directing Stage
- (vi) What is the full form of SEO?
a) Search Engine Optimal b) Social Engine Optimization
c) Search Engine Optimization d) Social Engine Optimal
- (vii) Which of the following is the least strong security encryption standard?
a) WPA3 b) WPA2 c) WPA d) WEP
- (viii) Unique value auction is mainly applies to?
a) New products b) Second hand products c) Engineering products d) None of these
- (ix) Which type of products is lesser purchased using ecommerce?
a) Automobiles b) Books c) Software d) None
- (x) Who protects system from external threats?
a) firewall b) EDI c) ERP d) Script kiddies

a a a b d

(xi) In Electronic cash payment

- a) A debit card payment system is used
- b) A customer buys several electronic coins which are digitally signed by coin issuing bank
- c) A credit card payment system is used
- d) RSA cryptography is used in the transactions

(xii) Master scheduling is the process that

- a) Determines the direction of business.
- b) Creates the capacity.
- c) Involves the functional planning.
- d) Provides specification for manufacturing.

Group - B

(Short Answer Type Questions)

Answer any **three** of the following

$3 \times 5 = 15$

2 Explain the advantages and disadvantages of ecommerce.

[Module1/CO1/Understand-LOCQ]

3. Define EDI? Discuss its layered structure.

[Module2/CO2/Remember-LOCQ]

$2+3$

4. Illustrate what is the impact of website traffic?

[Module3/CO2/Apply-LOCQ]

5. Describe the advantages and disadvantages of a Smart Card?

[Module4/CO2/Understand-LOCQ]

6. Define ERP? Discuss Different Types Of ERP?

[Module6/CO2/Remember-LOCQ]

$2+3$

Group - C

(Long Answer Type Questions)

Answer any **three** of the following

$3 \times 15 = 45$

Explain the different phases in an e commerce market element. What are the four types of

consumer oriented applications of E-commerce? Discuss them briefly.

[Module1/CO2/Understand-LOCQ]

$5+10$

Explain the evolution of different types of web servers on the internet.

[Module3/CO3/Understand-LOCQ]

15

Discuss the different online marketing strategies in EC.

[Module4/CO4/Understand-LOCQ]

7

Discuss about how search engines can be used to reach the customers.

[Module2/CO1/Apply-LOCQ]

8

10. Describe the reasons for the explosive growth of the ERP market? List the direct benefits of ERP systems?

[Module6/CO3/Remember-LOCQ]

11. Write notes on the following (**Any Three**):

- a) Digital Signature
- b) Risk Implementations of ERP
- c) Digital Certificates
- d) Public Key in Cryptography
- e) Related technologies of ERP

[Module5/CO3/Remember-LOCQ]



MCKV Institute of Engineering

Paper Code: OE-CS601H

Paper Name: E-Commerce and ERP

Time Allotted: 3 Hours

Full Marks: 70

*The figures in the margin indicate full marks.
Candidates are required to give their answers in their own words as far as practicable.*

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any **ten** of the following: $10 \times 1 = 10$
- (i) What is the full form of EDI?
 - (a) Electronic Data Interchange
 - (b) Electronic Direct Interchange
 - (c) Electronic Data Information
 - (d) Electronic Data Interconnection
- (ii) The E-payment mechanism widely used in B2B is
 - (a) Credit card
 - (b) Master card
 - (c) E-cheque
 - (d) Debit card
- (iii) Smart card is better protected than other cards using
 - (a) Hub
 - (b) Firewall
 - (c) Encryption
 - (d) All of these
- (iv) _____ is not a major type of electronic commerce.
 - (a) Consumer to Business
 - (b) Business to Consumer
 - (c) Business to Business
 - (d) Consumer to Consumer
- (v) Which of the following is not a party of SCM?
 - (a) Suppliers
 - (b) Manufacturers
 - (c) Distributor
 - (d) Customers
- (vi) Which of the following modules are included in ERP system?
 - (a) Materials management
 - (b) Financial accounting
 - (c) Human resource management
 - (d) All of the above
- (vii) Using _____ you improve your site's structure and content and conduct promotional activities to enhance your ranking on search engines.
 - (a) Search Engine Optimization
 - (b) Social Media Marketing
 - (c) Social Media Optimization
 - (d) Search Engine Marketing
- (viii) Where did the term "hacker" originate?
 - (a) MIT
 - (b) New York University
 - (c) Harvard University
 - (d) Bell's Lab

- (ix) Master scheduling is the process that
 (a) Determines the direction of business.
 (c) Involves the functional planning.
- manufacturing.
- (x) Real time communication of information regarding requirements and availability of logistics service is the core of
 (a) Right response (b) Right quality. (c) Right quantity (d) Right value.
- (xi) Secure Sockets Layers does which of the following.
 (a) Creates a secure, private connection to a web server (b) Encrypts information
 (c) Sends information over the internet (d) All of them
- (xii) Which type of products is lesser purchased using ecommerce?
 (a) Automobiles (b) Books (c) Softwares (d) None of these

Group - B
(Short Answer Type Questions)

Answer any *three* of the following $3 \times 5 = 15$

1. Describe the challenges faced by the traditional marketing of a product as compared to e-commerce marketing.
(Module1/C01/Understand/IOCQ) [5]
2. Demonstrate what are the components of security audit report.
(Module2 /C03/Apply/IOCQ) [5]
3. Describe what is a Manufacturer Model? Give an example for this model.
(Module3/C02/Understand/IOCQ) [5]
4. Explain the working of E-Cash payment system with a neat diagram.
(Module4/C02/Understand/IOCQ) [5]
5. Justify the need of collaborative computing in e-commerce.
(Module3/ C04/Apply/HOCQ) [5]

Group - C
(Long Answer Type Questions)

Answer any *three* of the following $3 \times 15 = 45$

6. Justify the different protocols in the information distribution layer of e commerce.
 Describe the tools available to publish the multimedia content in e commerce. What is HTML? Why is it used in e commerce?
(Module3/C03/Remember/HOCQ) [5+5+2+3]

8. Demonstrate security and the Web in e-commerce. Discuss about how search engines can be used to reach the customers.
(Module2/C01/Apply/IOCQ) [7+8]

9. Describe the necessary conditions a hash function used in digital signature should satisfy? Why is a certifying authority required in Ecommerce? How a certifying authority does perform its tasks?

(Module5/C03/Understand/IOCQ) [5+5+5]

- 10.a) Describe collaborative E-Commerce. Discuss its advantages.
 b) Discuss the procedures for Collaborative E-Commerce.
Module3/C01/Remember/LOCQ

11. Define what is Supply Chain Management. Why is it important in e-business environment? Illustrate the role does SCM play in Business to Business model i.e B2B e-Commerce.

(Module4/C02/Understand/IOCQ) [5+5+5]

Faster delivery control
 Inventory management
 Reduce cost
 Customer Satisfaction
 Supply management



MCKV Institute of Engineering

Paper Code: OE-CS601H

Paper Name: E Commerce & ERP

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any **five** of the following: 5×1=5

(i) Which type of E - Commerce is Amazon?

- 3 A) B2B B) B2C C) G2B D) G2C

(ii) The E-payment mechanism widely used in B2B is

- 5 A) Credit card B) Master card C) E-Cheque D) Debit card

(iii) _____ is a function of E commerce.

- 15 A) Marketing B) Supply Chain C) Finance D) All of the above

(iv) Which of the following is not a user of B2C E-commerce?

- 6 A) Distributors B) Manufacturers C) Publishers D) None of the above

(v) Which of the following is not a party of SCM?

- 8 A) Suppliers B) Manufacturers C) Distributors D) Customers

(vi) Systems that involve interaction with a shared database can be considered as-

- 4 A) software-based B) transaction-based C) server-based D) client-based

Group - B**(Short Answer Type Questions)**Answer any *two* of the following

2×5=10

- (2.) Describe the difference between Traditional Commerce with E Commerce. [(CO1)(Understand/LOCQ)]
- (3.) Illustrate EDI System with its Challenges. Apply [(CO2)(Apply/LOCQ)]
- (4.) Define different types of network security issues. [(CO1)(Remember/LOCQ)]

Group - C**(Long Answer Type Questions)**

1×15=15

- Answer any *one* of the following
- (5.) Explain Wireless Application Protocol and its related Architecture in context of Hand held devices. [(CO3)(Understand/LOCQ)]
- (6.) Describe the Key Components of ERP? Explain How Companies benefit from ERP? Summarise the most difficult challenges faced while working with ERP? [(CO3)(Understand/LOCQ)]



MCKV Institute of Engineering

Paper Code: OE-CS601H

Paper Name: E-Commerce and ERP

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

5×1=5

1. Choose the correct alternatives for any **five** of the following:

- (i) Secure Sockets Layers does which of the following?
 (A). Creates a secure, private connection to a web server (B). Encrypts information
 (C). Sends information over the internet (D). All of the above

- (ii) Who protects system from external threats?
 (A). firewall (B). EDI
 (C). ERP (D). Script kiddies

- (iii) _____ is a Japanese term meaning continuous improvement.
 (A) JIT (B) Kaizen
 (C) TQM (D) Taguchi Map

- (iv) In Electronic cash payment
 (A). A debit card payment system is used (B). A customer buys several electronic coins which are digitally signed by coin issuing bank
 (C). A credit card payment system is used (D). RSA cryptography is used in the transactions

- (v) The initial stage of the supply chain process is the _____.
 (A). Sourcing Stage (B). Organizing Stage
 (C). Planning Stage (D). Directing Stage

- (vi) Which of the following is not a key element of the business model?
 (A). Competitive advantage (B). Market strategy
 (C). Universal standards (D). Value Proposition

Group - B

(Short Answer Type Questions)

2×5=10

Answer any **two** of the following

- (2.) Explain the working of E-Cash payment system with a neat diagram.

(3.) Define ERP? Discuss Different Types Of ERP?

MCKVIE/B.TECH/IT/CA-4/OE-CS601H/2022-23

(4.) Describe the few major applications of cryptography in the modern world?

Module6/C02/Remember/LOCQ

Module5/C03/Understand/LOCQ

Group - C

(Long Answer Type Questions)

Answer any **one** of the following

$1 \times 15 = 15$

(5.) Discuss the main differences between electronic cheque payment and credit card payment in E-Commerce?

Explain cheque transaction protocol used in Ecommerce. (8+7)

Module4/C04/Understand/LOCQ

(6.) Define what is Supply Chain Management? Why is it important in e-business environment?

Illustrate the role does SCM play in Business to Business model i.e B2B e-Commerce. (4+3+8)

Module4/C02/Understand/LOCQ



MCKV Institute of Engineering
Paper Code: OE-CS601H
Paper Name: E Commerce & ERP

Full Marks: 30**Time Allotted: 1 Hour***The figures in the margin indicate full marks.**Candidates are required to give their answers in their own words as far as practicable.***Group - A****(Multiple Choice Type Questions)** **$5 \times 1 = 5$** 1. Choose the correct alternatives for any **five** of the following:

- (i) What is the abbreviation of CGI?
 (A). Computer Gateway Interface (B). Common Gateway Interface
 (C). Computer Gateway Interconnection (D). Common Gateway Interconnection
- (ii) The E-payment mechanism widely used in B2B is
 (A) Credit card (B) Master card (C) E-Cheque (D) Debit card
- (iii) What is 'EOQ' in Inventory control?
 (A). Economic Occurring Quantity (B). Economic Order Quantity
 (C). Economic Over Quantity (D). Economic of Quantity
- (iv) Which of the following is not a user of B2C E-commerce?
 (A) Distributors B) Manufacturers C) Publishers D) None of the above
- (v) Which of the following is not a party of SCM?
 (A) Suppliers B) Manufacturers C) Distributors D) Customers
- (vi) Organizations are using the ERP system in order to improve their
 (A) Business transactions (B) Worth (C) Knowledge (D) All of the above

WAP

gateway
proxy
router

[P.T.O]

Group - B

(Short Answer Type Questions)

$2 \times 5 = 10$

Answer any **two** of the following

(2.) Describe the difference between Traditional Commerce with E Commerce.

[Module1/C01/Remember/LOCQ]

[Module3/ C03/Understand/IOCQ]

(3.) Identify web traffic analysis in e-commerce?

[Module2/C01/Understand/LOCQ]

(4.) Discuss how E-Commerce is helpful to business success.

Group - C

(Long Answer Type Questions)

$1 \times 15 = 15$

Answer any **one** of the following

(5.) Explain Wireless Application Protocol and its related Architecture in context of Hand held devices.

[Module 2/C03/Understand/LOCQ]

(6.) Demonstrate security and the Web in e-commerce? Discuss about how search engines can be used to reach the customers. 7+8

[Module2/C01/Apply/IOCQ]



MCKV Institute of Engineering

Paper Code: OE-CS 601H

E Commerce & ERP

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any **five** of the following:

5×1

(i) Now day's organizations are paying attention supply chain management for...
(A)Proper supply (B)Value addition (C)Transportation (D)All the above

(ii) Many times..... is not included in SCM
(A)Material handling (B)Retail management (C)Sales management (D)None of these

(iii) Who protects system from external threats?
(A)Firewall (B)EDI (C) ERP (D)Script kiddies

(iv) In Electronic cash payment
(A)A debit card payment system is used (B)A customer buys several electronic coins which are digitally signed by coin issuing bank
(C) A credit card payment system is used (D)RSA cryptography is used in the transactions

(v) Secure Sockets Layers does which of the following?
(A)Creates a secure, private connection to a web server (B)Encrypts information
(C)Sends information over the internet (D) All of the above

(vi) The upstream supply chain is:
(A)Exclusively inside an organization. (B)Involved with procurement of material from suppliers.
(C)The distribution of products or delivery of services to customer
(D)Both the first and third answer above.

Group - B

(Short Answer Type Questions)

Answer any **two** of the following

2×5

(2.) Explain the characteristics of internet-based EDI?

Module4/CO2/Understand/IOCQ

(3.) Describe the few major applications of cryptography in the modern world?

Module5/CO3/Understand/LOCQ

(4.) Describe the limitations of ERP in E-Commerce?

Module6/CO3/Understand/IOCQ

Group - C

(Long Answer Type Questions)

Answer any **one** of the following

1×15

(5.) Define what is Supply Chain Management? Why is it important in e-business environment?

Illustrate the role does SCM play in Business to Business model i.e B2B e-Commerce. 4+3+8

Module4/CO2/Understand/IOCQ

(6.) Write notes on 3x5

a) Digital Signature b) Smart Cards c) Related technologies of ERP

Module5/CO3/Remember/LOCQ

Max Eff of Pure Aloha ~

$$N_{PA} = 6e^{-2b_1} = f(n)$$

$$\Rightarrow \frac{dn_{PA}}{db_1} = 0, \Rightarrow \frac{d(6e^{-2b_1})}{db_1} = 0$$

$$\Rightarrow 6 \left(\frac{de^{-2b_1}}{db_1} \right) + e^{-2b_1} \cdot \frac{db_1}{db_1^2} = 0$$

$$= 6 / (e^{-2b_1} \cdot 2) + e^{-2b_1} = 0$$

$$\Rightarrow -2b_1 + 1 = 0$$

$$\Rightarrow b_1 = \frac{1}{2}$$

$$N_{PA} | b_1 = \frac{1}{2} = \frac{1}{2} e^{-1} = 0.184 = 18.4\%$$

$$b_1 = \frac{1}{2} \rightarrow T_f \quad | \quad t = 2T_f - 1 \text{ short in transmitting}$$



MCKV Institute of Engineering

Paper Code: PE-IT601B

Paper Name: Data Analytics and Big Data

Time Allotted: 3 Hours

Full Marks: 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any **ten** of the following: $10 \times 1 = 10$
 - (i) Which command lists the blocks that make up each file in the filesystem.
 - a) hdfs fsck / -files -blocks
 - b) hdfs fsck / -blocks -files
 - c) hdfs fchk / -blocks -files
 - d) hdfs fchk / -files -blocks
 - (ii) For the frequently accessed HDFS files the blocks are cached in
 - a) the memory of the datanode
 - b) in the memory of the namenode
 - c) Both a) and b)
 - d) In the memory of the client application which requested the access to these files.
 - (iii) Under Hadoop High Availability, Fencing means
 - a) Preventing a previously active namenode from start running again.
 - b) Preventing the start of a failover in the event of network failure with the active namenode.
 - c) Preventing the power down to the previously active namenode.
 - d) Preventing a previously active namenode from writing to the edit log.
 - (iv) In a Hadoop cluster, what is true for a HDFS block that is no longer available due to disk corruption or machine failure?
 - a) It is lost for ever
 - b) It can be replicated from its alternative locations to other live machines.
 - c) The namenode allows new client request to keep trying to read it.
 - d) The Mapreduce job process runs ignoring the block and the data stored in it.
 - (v) Which utility is used for checking the health of a HDFS file system?
 - a) fchk
 - b) fsck
 - c) fsch
 - d) fcks
 - (vi) Which of these provides a Stream processing system used in Hadoop ecosystem?
 - a) Solr
 - b) Tez
 - c) Spark
 - d) Hive
 - (vii) A file in HDFS that is smaller than a single block size
 - a) Cannot be stored in HDFS.
 - b) Occupies the full block.
 - c) Occupies only the size it needs and not the full block.
 - d) Can span multiple blocks.

- (viii) As compared to RDBMS, Hadoop
 a) Has higher data Integrity.
 b) Does ACID transactions
 c) IS suitable for read and write many times
 d) Works better on unstructured and semi-structured data.

- (ix) The type of data Hadoop can deal with is
 a) Structured b) Semi-structured c) Unstructured d) All of these

- (x) The replication factor of a file in HDFS can be changed using
 a) chancerep b) rerep c) setrep d) xrep

- (xi) For a HDFS directory the replication factor(RF) is
 a) same as the RF of the files in that directory b) Zero c) 3 d) Does not apply.

- (xii) Data locality feature in Hadoop means
 a) store the same data across multiple nodes.
 b) relocate the data from one node to another.
 c) co-locate the data with the computing nodes.
 d) Distribute the data across multiple nodes.

Group - B

(Short Answer Type Questions)

Answer any **three** of the following

3×5=15

Distinguish between Supervised Learning & Unsupervised Learning with examples.

[Module1/C01/Understand-LOCQ]

5

Describe on what concept the Hadoop framework works?

[Module2/C03/Understand-HOCQ]

5

Explain the benefits of Apache Pig over MapReduce?

[Module5/C03/Understand-LOCQ]

5

Define "block" in HDFS? What is the default block size in Hadoop 1 and in Hadoop 2?

In it be changed ?

[Module3/C04/Remember-LOCQ]

5

Is it okay to optimize algorithms or codes to make them run faster? If yes, explain why?

[Module4/C05/Analyse-HOCQ]

5

Group - C

(Long Answer Type Questions)

Answer any **three** of the following

3×15=45

③ Describe 4 major classifications of Data Analytics.

[Module1/C01/Understand-LOCQ]

5

Differentiate between data analytics & data analysis.

[Module1/C01/Understand-LOCQ]

5

- ④ Describe data product? Explain why data is important?
 [Module1/C01/Remember-LOCQ]

7+4+2+2

- 8.a) Describe what does a "MapReduce Partitioner" do?
- b) Define the "Combiner". Is it mandatory?
- c) Explain the phases of MapReduce Program.
 [Module4/C04/Understand-LOCQ]

5+5+5

- 9.a) Explain how the DataNode knows the location of NameNode?
 [Module3/C04/Understand-HOCQ]

5

- b) Define the active and passive "NameNodes"?
 [Module3/C04/Remember-LOCQ]

5

- c) State the commands :- i) to copy data from the local system onto HDFS? ii) to copy data from the HDFS to local system.
 [Module3/C04/Understand-LOCQ]

2.5×2

- 10.a) Write a Pig Script to count each word from a large csv (comma separated) file.

10

[Module5/C05/Apply-HOCQ]

- b) How does NameNode tackle DataNode failures?
 [Module3/C04/Understand-HOCQ]

5

- 11.a) Calculate the following considering the Market Basket transactions given below :-

TID	ITEMS
1	Bread, Milk
2	Bread, Diaper, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

- i. What is the k-itemset of {Milk, Bread, Diaper}?

- ii. Calculate support count of {Milk, Bread, Diaper}?

- iii. What is support of {Milk, Bread, Diaper}?

- iv. What is confidence of {Milk, Bread, Diaper}?

- v. Check whether association rule {Milk, Bread} → Diaper is valid association rule or not. The threshold values of support & confidence are 0.3 & 0.55 respectively.
 [Module1/C01/Apply-HOCQ]

2

- b) Explain "Big Data" & 3V's In Big data?
 [Module2/C02/Remember-LOCQ]



MCKV Institute of Engineering

Paper Code: PE-IT601B

(60)

Paper Name: Data Analytics and Big Data

Time Allotted: 3 Hours

Full Marks: 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any **ten** of the following: **10×1=10**
- i. Which property can be used to change the replication factor?
 - a) dfs.update.replica
 - b) dfs.replication **12**
 - c) changeReplicationFactor
- ii. Users can specify the number of reducers **18**
 - a) True
 - b) False
- iii. The datanode and namenode are respectively
 - a) Master & Worker nodes **24**
 - b) Worker & Master Nodes
 - c) Both are worker nodes
 - d) None of these
- iv. What is the input to the Mapper?
 - a) Records from InputSplit **16**
 - b) Key Value pair from input
 - c) InputFormat will group the keys and send to Mapper
- v. The property used to set the default filesystem for Hadoop in core-site.xml is
 - a) filesystem.default
 - b) fs.default
 - c) fs.defaultFS
 - d) hdfs.default
- vi. The following is not permitted on HDFS files
 - a) Deleting
 - b) Renaming
 - c) Moving
 - d) Executing.
- vii. Data locality feature in Hadoop means
 - a) store the same data across multiple nodes
 - b) relocate the data from one node to another
 - c) co-locate the data with the computing nodes
 - d) Distribute the data across multiple nodes
- viii. HDFS block size is larger as compared to the size of the disk blocks so that
 - a) Only HDFS files can be stored on the disk used
 - b) The seek time is maximum
 - c) Transfer of a large file made over multiple disk blocks is not possible
 - d) A single file larger than the disk size can be stored across many disks in the cluster

ix. What is the main problem faced while reading and writing data in parallel from multiple disks?

- a) Processing high volume of data faster ✓ b) Combining data from multiple disks
 - c) The software required to do this task is extremely costly
 - d) The hardware required to do this task is extremely costly
- x. The number of tasks a task tracker can accept depends on
- ✓ a) Maximum memory available in the node b) Not limited
 - c) Number of slots configured in it d) As decided by the job Tracker
- xi. Which of the following is not a fencing mechanism for a previously active namenode?
- a) Disabling its network port via a remote management command.
 - ✓ b) Revoking its access to shared storage directory
 - c) Formatting its disk drive ✓ d) STONITH
- xii. Which command lists the blocks that make up each file in the filesystem.
- a) hdfs fsck / -files -blocks ✓ b) hdfs fsck / -blocks -files
 - c) hdfs fchk / -blocks -files d) hdfs fchck / -files -blocks

Group - B

(Short Answer Type Questions)

Answer any **three** of the following

3×5=15

1. What happens when 2 clients try to access the same file in HDFS?

[Module3/C04/Understand-HOCQ] [5]

2. Explain why NameNode does not persist information of block location?

[Module3/C04/Analyze-LOCQ] [5]

3. Define "SerDe" in "Hive"?

[Module5/C03/Remember-LOCQ] [5]

4. Distinguish between "HBase" and "Relational Databases".

[Module5/C03/Understand-LOCQ] [5]

5. Explain the benefits of Apache Pig over MapReduce?

[Module5/C03/Understand-LOCQ] [5]

Group - C

(Long Answer Type Questions)

Answer any **three** of the following

3×15=45

6. Describe Architecture of HDFS with its components.

[Module3/C04/Remember-LOCQ] [10]

b) Distinguish node, rack, cluster, data center.

[Module2/C02/Understand-LOCQ] [5]

8. a) Describe the configuration parameters of "MapReduce" program.

[Module4/C04/Understand-LOCQ] [5]

b) Define the components of Apache HBase?

[Module5/C03/Remember-LOCQ] [5]

c) Explain the benefits of HDFS over local file system?

[Module3/C04/Understand-HOCQ] [5]

9. a) Calculate the following considering the Market Basket transactions given below:-

TID	ITEMS
1	Bread, Milk
2	Bread, Diaper, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

i. What is the k-itemset of {Milk, Bread, Diaper}?

ii. Calculate support count of {Milk, Bread, Diaper}?

iii. What is support of {Milk, Bread, Diaper}?

iv. What is confidence of {Milk, Bread, Diaper}?

v. Check whether association rule {Milk, Bread} → Diaper is valid association rule or not.
The threshold values of support & confidence are 0.3 & 0.55 respectively.

[Module1/C01/Apply-HOCQ] [2×5]

b) Is Data Analysis & Data Analytics synonymous? If yes, describe their similarity or if no, describe why not synonymous.

[Module1/C02/Understand-LOCQ] [1+4]

10. a) Calculate the total number of movie rating of all available rating (1 to 5) using MAP-REDUCE program in python from the Movie -Lens (u.data) Database.

[Module4/C05/Apply-HOCQ] [10]

b) Write a pig command

- ✓ i) to load the file on HDFS to pigstorage for analyzing.
- ii) to store the above loaded data in a file then see the content of the file.

[Module5/C03/Understand-LOCQ] [2.5×2]

11. a) Name the three modes in which Hadoop can run.

[Module2/C02/Remember-LOCQ] [5]

b) Describe a priori principle. Generate frequent item set lattice with 4 items A, B, C, D.

[Module1/C01/Apply-LOCQ] [2+3]

c) Explain "Big Data" & 3V's in Big data?

[Module2/C02/Remember-LOCQ] [5]



MCKV Institute of Engineering

Paper Code: PE-IT601B

Data Analytics & Big Data

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

- I. Choose the correct alternatives for any five of the following: 5×1=5
- ii) An item set satisfying the support criterion is known as:
a) Frequent Item set b) Confident Item set c) Accurate Item set d) Reliable Item set
- ii) What happens to local file system when you install HDFS?
a) Local file system and HDFS co-exist b) Local file system is replaced by HDFS
- iii) You have an ETL job in Oracle that takes 2 hours to complete. Business users are requesting you to bring down the execution time of the job to 1 hour. Best Approach: It is recommended to add additional servers to solve this problem?
a) True b) False
- iv) A good clustering is one having:
a) Low inter-cluster distance & low intra-cluster distance
b) Low inter-cluster distance & high intra-cluster distance
c) High inter-cluster distance & low intra-cluster distance
d) High inter-cluster distance & high intra-cluster distance
- v) If a record data matrix has reduced number of columns after transformation, the transformation has performed:
a) Data Sampling b) Dimensionality Reduction c) Noise Cleaning d) Discretization
- vi) What is the default replication factor of a block on HDFS?
a) 1 b) 2 c) 3

Group - B
(Short Answer Type Questions)

Answer any two of the following

$2 \times 5 = 10$

2. Explain "Big Data" & 3V's in Big data? [Module2/CO2/Remember-LOCQ]
3. Distinguish between HADOOP & RDBMS. [Module2/CO2/Understand-LOQC]
4. Explain all the levels of Data. [Module1/CO1/Remember-LOQC]

Group - C

(Long Answer Type Questions)

Answer any **one** of the following

$1 \times 15 = 15$

5. a) Discuss the major types of data analytics.
b) Is Data Analysis & Data Analytics synonymous? If yes describe their similarity or if no describe why not synonymous. [Module1/CO1/Remember-LOQC] 6+4
c) Define "block" in HDFS? What is the default block size in Hadoop2? Calculate the no. of blocks required for 680 MB file in Hadoop distributed file system for Hadoop 2. [Module2/CO2/Evaluate-IOQC] 2 +1+2

6. a) Consider the following Market Basket transactions and answer the questions:-

7. TID	8. ITEMS
9. 1	10. Bread, Milk
11. 2	12. Bread, Diaper, Beer, Eggs
13. 3	14. Milk Diaper, Beer, Coke
15. 4	16. Bread, Milk, Diaper, Beer
17. 5	18. Bread, Milk, Diaper, Coke

- i. What is the k-itemset of {Milk, Bread, Diaper}?
ii. Calculate support count of {Milk, Bread, Diaper}?
iii. What is support of {Milk, Bread, Diaper}?
iv. What is confidence of {Milk, Bread, Diaper}?
v. Check whether association rule {Milk, Bread} \rightarrow Diaper is valid association rule or not. The threshold values of support & confidence are 0.3 & 0.55 respectively.
b) Describe Apriori principle. Generate frequent itemset lattice with 4 items A, B, C, D.

[Module1/CO1/Evaluate-IOQC] (2*5+2+3)



MCKV Institute of Engineering

Paper Code : PE-IT601B

Paper Name : Data Analytics and Big Data

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any five of the following: $5 \times 1 = 5$
 - i) An item set satisfying the support criterion is known as: ✓
 - a) Frequent Item set
 - b) Confident Item set
 - c) Accurate Item set
 - d) Reliable Item set
 - ii) If A, B are 2 sets of items and A is a subset of B. Which of the following statement is always true: ✓
 - a) $\text{support}(A) \leq \text{support}(B)$
 - b) $\text{support}(A) \geq \text{support}(B)$
 - c) $\text{support}(A) = \text{support}(B)$
 - d) $\text{support}(A) \neq \text{support}(B)$
 - iii) You have an ETL job in Oracle that takes 2 hours to complete. Business users are requesting you to bring down the execution time of the job to 1 hour. Best Approach: It is recommended to add additional servers to solve this problem? ✓
 - a) True
 - b) False
 - iv) A good clustering is one having: ✓
 - a) Low inter-cluster distance & low intra-cluster distance
 - b) Low inter-cluster distance & high intra-cluster distance
 - c) High inter-cluster distance & low intra-cluster distance
 - d) High inter-cluster distance & high intra-cluster distance
 - v) If a record data matrix has reduced number of columns after transformation, the transformation has performed: ✓
 - a) Data Sampling
 - b) Dimensionality Reduction
 - c) Noise Cleaning
 - d) Discretization

Q1) If a record data matrix has reduced number of rows after transformation, the transformation has performed:

- A. Data Sampling
- B. Dimensionality Reduction
- C. Noise Cleaning
- D. Discretization

Group - B

(Short Answer Type Questions)

Answer any **two** of the following

2×5=10

Q2. Assuming that we consider one data problem as big data problem by considering only the volume, but google or other company may not consider this. Why? [(CO2)]

Q3. Distinguish between HADOOP & RDBMS.

[(CO2) (Understand/LOQC)]

Q4. Explain all the levels of Data.

[(CO1) (Remember/LOQC)]

Group - C

(Long Answer Type Questions)

Answer any **one** of the following

1×15=15

Q5. a) Discuss the major types of data analytics.

b) Is Data Analysis & Data Analytics synonymous? If yes describe their similarity or if no describe why not synonymous.

[(CO1) (Remember/LOQC)]

(8+7)

Q6. a) Consider the following Market Basket transactions and answer the questions:-

7. TID	8. ITEMS
9. 1	10. Bread, Milk
11. 2	12. Bread, Diaper, Beer, Eggs
13. 3	14. Milk Diaper, Beer, Coke
15. 4	16. Bread, Milk, Diaper, Beer
17. 5	18. Bread, Milk, Diaper, Coke

i. What is the k-itemset of {Milk, Bread, Diaper}?

ii. Calculate support count of {Milk, Bread, Diaper}?

iii. What is support of {Milk, Bread, Diaper}?

iv. What is confidence of {Milk, Bread, Diaper}?

v. Check whether association rule {Milk, Bread} → Diaper is valid association rule or not. The threshold values of support & confidence are 0.3 & 0.55 respectively.

vi. Describe Apriori principle. Generate frequent itemset lattice with 4 items A, B, C, D.

[(CO1) (Evaluate/IOQC)] (2×5+2+3)

MCKV Institute of Engineering

Paper Code: PE-IT601B

Data Analytics & Big Data

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any five of the following: 5x1=5

- i) The type of data Hadoop can deal with
a) Structured b) Semi-structured c) Unstructured d) All of the above

- ii) Pig Latin instructions are converted in to

- a) MapReduce jobs b) Pig jobs

- iii) Assume the dataset you would like to use for your Hive table is already used by a MapReduce job and Pig. What is the safest table type to use when creating a table in Hive?

- a) External b) Internal

- iv) Users can specify the number of reducers

- a) True b) False

- v) How do you plan to recover from data loss?

- a) Replicate each blocks to more than one node b) It is impossible to recover from data loss

- vi) What is the input to the Reducer?

- a) One key value pair for each record in the input split
b) Key and list of values grouped by key from all the Mappers
c) All records from the InputSplit

$$VFP = 3 \times 10 + 8 \times 10 + 5 + 15 + 3 \times 13 + 15 \times 2 + 7 \times 2$$

Support high degree of parallelism
can be varying 0 or 5.
Total degree influence = 545 to 10
Value adjustment factor = $F = 0.01 + 0.05$

$$\text{adjust F} = 0.01 + 0.05 = 0.06$$

$$\text{points} = VFP \times VF = 287 \times 0.06$$

$$EF = 2.4(32) + 0.05 = 222.25$$

$$\text{development time} = 2.5(91)^{0.3} \text{ months}$$

$$= 1.4 \times 1500$$

$$= Rs 21000$$

Group - B

(Short Answer Type Questions)

Answer any two of the following

$2 \times 5 = 10$

2. Explain the benefits of Apache Pig over MapReduce?

[Module5/CO3/Understand-IOCQ]

5

3. Distinguish between Supervised & UnSupervised Learning with examples

[Module2/CO2/Understand-LOQC]

5

4. Explain all the levels of Data.

[Module1/CO1/Remember-LOQC]

5

Group - C

(Long Answer Type Questions)

Answer any **one** of the following

$1 \times 15 = 15$

5. a) Explain how the DataNode knows the location of NameNode?

[Module3/CO4/Understand-HOCQ]

5

b) Define the active and passive "NameNodes"?

[Module3/CO4/Remember-LOCQ]

5

c) *State the commands :-i) to copy data from the local system onto HDFS? ii) to copy data from the HDFS to local system.*

[Module3/CO4/Understand-LOCQ] $2.5 \times 2 = 5$

6. a) Explain the phases of MapReduce Program.

b) Define the "Combiner". Is it mandatory?

[Module4/CO4/Understand-IOCQ]

5+5

c) Explain collaborative filtering?

[Module2/CO2/Understand-LOQC]

5

MCKV Institute of Engineering

Paper Code : PE-IT601B

Paper Name : Data Analytics and Big Data

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any five of the following: $5 \times 1 = 5$

i) An item set satisfying the support criterion is known as:

- a) Frequent Item set b) Confident Item set c) Accurate Item set d) Reliable Item set

ii) How do you plan to recover from data loss?

- a) Replicate each blocks to more than 1 node b) It is impossible to recover from data loss.

iii) Why not use Solid State Drive (SSD) instead of Hard Disk Drive (HDD) to improve data access rate?

- a) SSDs are not reliable b) SSDs are expensive c) SSDs are low quality

iv) A good clustering is one having:

- a) Low inter-cluster distance & low intra-cluster distance

- b) Low inter-cluster distance & high intra-cluster distance

- c) High inter-cluster distance & low intra-cluster distance

- d) High inter-cluster distance & high intra-cluster distance

v) If a record data matrix has reduced number of columns after transformation, the transformation performed:

- a) Data Sampling b) Dimensionality Reduction c) Noise Cleaning

vi) If a record data matrix has reduced number of rows after transformation, the transformation performed:

- a) Data Sampling b) Dimensionality Reduction

- c) Normalization d) Discretization

Group - B**(Short Answer Type Questions)**Answer any ***two*** of the following

2×5=10

2. What is data product? Why data is important?
3. Distinguish between HADOOP & RDBMS.
4. Explain all the levels of Data.

[(CO2) (Understand/LOQC)]

[(CO2) (Understand/LOQC)]

[(CO1) (Remember/LOQC)]

Group - C**(Long Answer Type Questions)**Answer any ***one*** of the following

1×15=15

5. a) Discuss the major types of data analytics.
 b) Explain 3V. Distinguish between data analytics & data analysis.

[(CO1) (Remember/LOQC)]

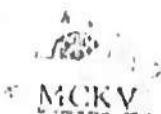
(8+3+4)

6. a) Consider the following Market Basket transactions and answer the questions:-

7. TID	8. ITEMS
9. 1	10. Bread, Milk
11. 2	12. Bread, Diaper, Beer, Eggs
13. 3	14. Milk, Diaper, Beer, Coke
15. 4	16. Bread, Milk, Diaper, Beer
17. 5	18. Bread, Milk, Diaper, Coke

- i. What is the k-itemset of {Milk, Bread, Diaper}?
- ii. Calculate support count of {Milk, Bread, Diaper}?
- iii. What is support of {Milk, Bread, Diaper}?
- iv. What is confidence of {Milk, Bread, Diaper}?
- v. Check whether association rule {Milk, Bread} → Diaper is valid association rule or not. The threshold values of support & confidence are 0.3 & 0.55 respectively.
- What is support, support count & frequent item set?

[(CO1) (Evaluate/LOQC)] (2×5+2+2+1)



MCKV Institute of Engineering
Paper Code :PE-IT601B

Paper Name :Data Analytics and Big Data

Time Allotted: 1 Hour

Full Marks: 30

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for any five of the following: $5 \times 1 = 5$
 - i. What happens to local file system when you install HDFS?
 - a) Local file system and HDFS co-exist ✓
 - b) Local file system is replaced by HDFS
 - ii. Which property can be used to change the replication factor?
 - a) dfs.update.replica
 - b) dfs.replication
 - c) changeReplicationFactor 12
 - iii. Pig Latin instruction are converted into
 - a) MapReduceJobs ✓ b) Pig Jobs
 - iv. The datanode and namenode are respectively
 - a) Master & Worker nodes ✓
 - b) Worker & Master Nodes
 - c) Both are worker nodes
 - d) None of these 24
 - v. The property used to set the default filesystem for Hadoop in core-site.xml is
 - a) filesystem.default ✓
 - b) fs.default 98
 - c) fs.defaultFS
 - d) hdfs.default
 - vi. What is the benefit of Combiner?
 - a) Reduce the amount of input to reducers ✓
 - b) Reduce the amount of input to mappers
 - c) None of the above

Group - B**(Short Answer Type Questions)**Answer any **two** of the following**2×5=10**

2. Explain the benefits of Apache Pig over MapReduce? [Module5/CO3/Understand-LOCQ] 5
3. Distinguish node, rack, cluster, data center. [Module2/CO2/Understand-LOCQ] 5
4. "Storing a file in HDFS means a lot of more involvements internally"---Explain this. [Module3/CO4/Analyze-HOCQ] 5

Group - C**(Long Answer Type Questions)**Answer any **one** of the following**1×15=15**

Ratio of
ordinal
nominal

Validity
Variable

5. a) Discuss the major types of data analytics. [Module1 /CO1/Remember-LOCQ]
- b) Explain 3V. Distinguish between data analytics & data analysis. [Module2/CO2/Remember-LOCQ] (8+3+4)
6. a) Calculate the total number of movie rating of all available rating (1 to 5) using MAP-REDUCE program in python from the Movie -Lens (u.data) Database. [Module4/CO5/Apply-HOCQ] 10
- b) What is support, support count & frequent itemset? [Module1/CO1/Remember-LOQC] (2+2+1)

2
128
15
4