6E7101

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7101

B.Tech. VI-Sem. (Main/Back) Exam. May, 2025

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

6AID3-01 Digital Image Processing

CS, IT, AID, CAI, CDS, CCS, CIT, CSD

Time : 3 Hours **Maximum Marks : 70**

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting material is permitted during examination.

(Mentioned in Form No. 205)

1. ..... 2. .....

PART-A [10x2=20]

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. Discuss about contrast stretching and intensity slicing.

Q.2. What are the types of image enhancement available?

Q.3. What are the types of edge detection operators available?

Q.4. What is the role of Hough Transform in global processing?

Q.5. Why noise is always considered to be additive, in images?

Q.6. What is the role of Fourier transform in the context of image processing?

Q.7. How do frequency domain filters contribute to image transformation?

Q.8. What is the significance of Wavelet Transforms?

Q.9. What is the concept of degradation function?

Q.10. Identify the types of redundancies in image compression.

PART-B [5x4=20]

(Analytical/Problem-solving Questions)

Attempt any five questions

Q.1. Show the bit plane slicing of the following image :

7	6	5
4	3	2
1	1	0

Q.2 Prove whether the Fourier Transform of a discrete time signal is continuous or discrete. Explain with an example.

Q.3. Explain HIS color model with an appropriate figure. Explain the conversion procedure from RGB color space to HSI color space.

Q.4. Describe how homomorphic filtering separates illumination and reflectance components.

Q 5. Write notes on Run length encoding and Shift codes.

Q.6. Explain any two boundary representation schemes and illustrate with an example.

Q.7. Discuss image segmentation based on various thresholding techniques.

PART-C [3x10=30]

(Descriptive/Analytical/Problem-solving/Design Question)

Attempt any three questions

Q1. Describe histogram equalization. Obtain histogram equalization for the following image segment of size. Write the inference on the image segment before and after equalization :

20	20	20	18	16
15	15	16	18	15
15	15	19	15	17
16	17	19	18	16
20	18	17	20	15

Q.2 Decode the arithmetic coded message, 0.23355 for the coding model :

Symbol	Probability
A	0.2
E	0.3
I	0.1
O	0.2
u	0.1
!	0.1

And also explain LZW coding with an example and explain Redundancies and their removal methods.

Q.3 Consider the image 'I' below and the filters 'F' and 'L'

'I'			'F'			'L'		
1	1	1		1/8			1	
1	8	1	1/8	1/2	1/8	1	–4	1
1	1	1		1/8			1	

(a) Correlate the image 'I' with the filter 'F' and compute the output image.

(b) Apply filter 'L' to the same image 'I' to produce a 3 by 3 output image.

Q.4 Explain following terms with example :

(a) Lookup Table (b) Additive vs. Subtractive color system

(c) Hue and Saturation (d) Hue-Max-Min-Diff (HMMD) Color Space

Q. 5. Consider an image- Show the output of any edge detection algorithm :

1	2	5
5	5	5
5	3	2

----- × -----

6E7102

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7102

B.Tech. VI-Sem. (Main/Back) Exam., May-2025

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

6AID4-02 Machine Learning

CS, IT, AID, CAI, CDS, CCS, CIT, CSD

Time : 3 Hours **Maximum Marks : 70**

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting material is permitted during examination. (Mentioned in Form No. 205)

1. ..... 2. .....

PART-A [10x2=20]

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. What is the significance of standard deviation in Machine Learning?

Q.2. Differentiate between overfitting and underfitting.

Q.3. What is Confusion matrix?

Q.4. What are the applications of Support Vector Machine (SVM)?

Q.5. What is the working principle of Naive Bayes Classifier?

Q.6. What are the advantages and disadvantages of hierarchical clustering?

Q.7. What is a dendrogram in hierarchical clustering?

Q.8. Calculate the variance and standard deviation of the following data :

4, 2, 5, 8, 6

Q.9. What does Mean Squared Error (MSE) tell you?

Q.10. Explain the straight line equation $y = mx + c$ with respect to Machine learning.

PART-B [5x4=20]

(Analytical/Problem-solving Questions)

Attempt any five questions

Q.1. Explain the concept of deep learning in detail.

Q.2 What do you mean by reinforcement learning? Also explain the model based reinforcement learning.

Q.3. Explain Markov Decision Process (MDP) also explain policy iteration and value iteration.

Q.4. Explain the concept of feature selection used in Machine Learning.

Q.5. How the decision tree algorithm works? Explain it with the help of an example. What is the role of entropy and information gain in Decision tree?

Q.6. Explain the working principle of Probabilistic clustering in Machine Learning with example.

Q.7. What is the use of PCA in Machine Learning? Give the steps of PCA algorithm.

PART-C [3x10=30]

(Descriptive/Analytical/Problem-Solving/Design Question)

Attempt any three questions

Q1. Write short notes on the following :

(a) Perceptions

(b) Multilayer Network

(c) Back Propagation

(d) Artificial Neutral Network

Q.2 What is the role of State-Action-Reward-State-Action (SARSA) in machine learning? Differentiate between SARSA and Q-learning.

Q.3 Explain Principal Component Analysis (PCA) in machine learning? How to evaluate machine learning algorithm?

Q.4 Consider an example where five weeks sales data (in thousand) is given as shown in table :

x_i (Weeks)	y_i(Sales in thousand)
1	1.2
2	1.8
3	2.6
4	3.2
5	3.8

Apply linear regression algorithm to predit 7th and 12th week sales.

Q. 5. Consider the following dataset which shows BMI, Age, and Sugar level (0 represent "No Sugar", 1 represent "Sugar"). Assume k = 3, apply k-nearest neighbor classifier to predict the diabetic patient with the given feature BMI, Age.

Test example of BMI = 43.6, Age = 40, predict the patient has sugar or no sugar:

BMI	AGE	SUGAR
33.6	50	1
26.6	30	0
23.4	40	0
43.1	67	0
35.3	23	1
35.9	67	1
36.7	45	1
25.7	46	0
23.3	29	0
31	56	1

----- × -----

6E7103

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7103

B.Tech. VI-Sem. (Main/Back) Exam., May-2025

COMPUTER SCIENCE AND ENGINEERING (ARTIFICIAL INTELLIGENCE & DATA SCIENCE)

6CAI4-03/Information Security Systems

CS, IT, AID, CAI, CIT, CCS, CDS

Time : 3 Hours Maximum Marks : 70

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting material is permitted during examination. (Mentioned in Form No. 205)

1. ..... 2. .....

PART-A

[10x2=20]

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. What is difference between Cryptography and Cryptanalysis?

Q.2. What is difference between stream and block ciphers?

Q.3. What are Confusion and Diffusion in Cryptography?

Q.4. What are strength of DES algorithm?

Q.5. What is difference between public key and private key cryptosystems?

Q.6. Write any two applications of public key cryptography.

Q.7. What is cryptographic hash function? Write its any two properties.

Q.8. What are the requirements of Message Authentication Codes?

Q.9. Write any four general means of authenticating a user's identity.

Q.10. What is HTTPS?

PART-B [5x4=20]

(Analytical/Problem solving questions)

Attempt any five questions

Q.1. Encrypt the message "Code" using the Hill cipher with the key $\begin{bmatrix} 3 & 2 \\ 8 & 5 \end{bmatrix}$ and also decrypt the ciphertext to original plaintext.

Q.2 Explain AES key expansion algorithm with suitable diagram.

Q.3. Perform encryption and decryption using the elgamal algorithm for the following:

$q = 19; \alpha = 13, X_A = 5; M = 15; k = 6$

Q.4. Explain Cipher-Based Message Authentication Code (CMAC) with suitable diagram.

Q 5. Explain SSL Record Protocol with suitable diagram.

Q.6. Explain the Electronic Code Book (ECB) block cipher mode of operation.

Q.7. Explain the application of cryptographic hash function for digital signature.

PART-C [3x10=30]

(Descriptive/Analytical/Problem Solving/Design questions)

Attempt any three questions

Q.1. What are security attacks? Explain different types of security attacks.

Q.2. Explain the internal structure of single round of DES algorithm.

Q.3. Explain the RSA algorithm and using this algorithm perform the encryption and decryption for the following :

$p = 5; q = 31; e = 13; M = 5$

Q.4. Explain the digital signature algorithm with the diagrams showing functions of signing and verifying.

Q. 5. Explain the Public-key certificates technique for distribution of public keys.

----- × -----

6E7104

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7104

B.Tech. VI-Sem. (Main/Back) Exam. May - 2025

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

6AID4-04 Computer Architecture and Organization

CS, IT, AID, CAI, CDS, CCS, CIT, CSD

Time : 3 Hours **Maximum Marks : 70**

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting material is permitted during examination. (Mentioned in Form No. 205)

1. ..... 2. .....

PART-A [10x2=20]

(Answer should be given up to 25 words only)
All questions are compulsory

Q.1. What is a priority interrupt?

Q.2. What are the most common fields of an Instruction format?

Q.3. What is immediate addressing mode?

Q.4. What are the major characteristics of a pipeline?

Q.5. Define latency time.

Q.6. Explain fixed point representation.

Q.7. Define multi-processing.

Q.8. Represent $(70)_{10}$ in a signed magnitude format and one's complement form.

Q.9. Write a quick note on subroutine.

Q.10. Define superscalar processor.

PART-B [5x4=20]

(Analytical/Problem solving questions)

Attempt any five questions

Q.1. Explain the functional units of an architecture of computer with diagram.

Q.2 Why interprocess synchronization needed? Explain.

Q.3. What are logical microoperatios? Explain about applications of logical microoperation.

Q.4. Differentiate between hardwired and microprogrammed control unit.

Q 5. Explain about serial communication.

Q.6. Explain the need of cache memory. Also explain Hit Ratio.

Q.7. Describe the algorithm for integer division with suitable example.

PART-C [3x10=30]

(Descriptive/Analytical/Problem Solving/Design questions)

Attempt any three questions

Q1. Define parallel processing and explain the Flynn's classification of computer with suitable diagram.

Q.2 Explain different types of computer registers with common bus system with a neat diagram.

Q.3 Explain the Booth Algorithm with its flowchart. Perform multiplication of (-13) and (+9) using Booth Algorithm.

Q.4 What is DMA? Draw the Block diagram of DMA controller and explain it.

Q. 5. (i) How many 128×8 RAM chips are needed to provide a memory capacity of 2048 bytes?

(ii) How many lines of address bus must be used to access 2048 bytes of memory? How many of these lines will be common to all chips?

(iii) How many lines must be decoded for chip select?

(iv) Specify the size of the decoders?

----- × -----

6E7105

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7105

B.Tech. VI-Sem. (Main/Back) Exam. May–2025

Artificial Intelligence and Data Science

6AID4-05 / Principles of Artificial Intelligence

CS, IT, AID, CAI

Time : 3 Hours **Maximum Marks : 70**

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting material is permitted during examination.

(Mentioned in Form No. 205)

1. ..... 2. .....

PART-A **[10×2=20]**

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. Define Rational Agent and Intelligent Agent.

Q.2. What is the role of Heuristic function in informed search?

Q.3. List the three main components of a Constraint satisfaction problem.

Q.4. What is rule based knowledge base?

Q.5. Construct a truth table for the compound proposition $(p \rightarrow q) \rightarrow (q \rightarrow p)$.

Q.6. What are support vectors in SVM?

Q.7. What are Neural Networks?

Q.8. Define Universal and Existential quantifiers in FOL with examples.

Q.9. What is initial and goal state in water jug problem?

Q.10. Define uncertainty and list the causes of uncertainty.

PART-B [5x4=20]

(Analytical/Problem-solving Questions)

Attempt any five questions

Q.1. Explain the differences between Breadth First Search and Depth First Search algorithms with the help of an example.

Q.2. Compare forward chaining and backward chaining used in expert systems.

Q.3. State Bayes' theorem and explain its significance in probability estimation.

Q.4. What is Robotics? Mention two real-world applications of robots.

Q.5. What is Market Basket analysis and how it is used in recommendation systems?

Q.6. Explain the difference between constants, variables, predicates, and functions in FOL. Convert the following sentence into FOL : "all humans are mortal".

Q.7. Solve the following 3×3 sliding puzzle given with the initial and goal state.

Initial State Goal State

1	2	3		1	2	3
8		4	⇒	4	5	6
7	6	5		7	8	

PART-C [3x10=30]

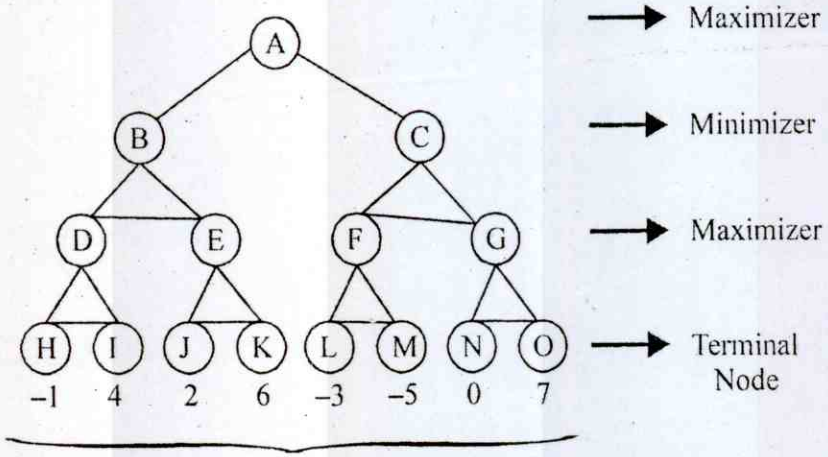
(Descriptive/Analytical/Problem-Solving/Design Questions)

Attempt any three questions

Q.1. Write the algorithm for the A* search technique. Using a suitable example, explain how A* search is more effective than the Greedy Best-First Search technique.

Q.2. Explain the concepts of Supervised Learning and Unsupervised Learning in Machine Learning. Discuss their key differences with suitable examples. Additionally, compare the advantages and limitations of both approaches.

Q.3. Given the following game tree, apply the Minimax algorithm to determine the optimal move for the maximizing player. Then, apply Alpha-Beta Pruning to the same tree and show which nodes are pruned. Compare the number of nodes evaluated in both cases. Compare time efficiency in terms of the number of nodes processed.



Terminal Values

Q.4. Explain the steps involved in Natural Language Processing (NLP). Provide examples of how techniques like tokenization, part-of-speech tagging, named entity recognition (NER), and sentiment analysis are used in NLP. Discuss any one example highlighting the importance on NLP in real-world applications.

Q.5. Consider the following knowledge base in First-Order Logic (FOL) :

$\forall x \left(\text{Cat}(x) \rightarrow \text{Animal}(x)\right)$

$\forall x \left(\text{Animal}(x) \rightarrow \text{Living}(x)\right)$

$\forall x \left(\text{Living}(x) \wedge \text{Friendly}(x) \rightarrow \text{Pet}(x)\right)$

Cat(Tom)

Friendly(Tom)

Task : Using Theorem Proving in First-Order Logic. Prove that Pet(Tom).

----- × -----

6E7106

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7106

B.Tech. VI-Sem. (Main/Back) Exam., May–2025

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

6AID4-06 Cloud Computing

CS, IT, AID, CAI, CCS, CIT, CSD

Time : 3 Hours **Maximum Marks : 70**

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting materials is permitted during examination. (Mentioned in Form No. 205)

1. ..... 2. .....

PART-A **[10×2=20]**

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. Briefly explain the historical journey of the development of Cloud Computing. [2]

Q.2. What are the ethical issues in cloud computing`? [2]

Q.3. Briefly explain the concept of : [2]

(a) Ubiquitous Cloud

(b) Internet of Things

Q.4. Define virtualization techniques used in Cloud Computing. What are the advantages of virtualization in Cloud Computing? [2]

Q.5. What are the fundamentals of cloud information security? [2]

Q.6. What are the cloud security services in cloud computing? [2]

Q.7. What are the threats in cloud and how it can be identified? [2]

Q.8. What is the vision and components of cloud computing? [2]

Q.9. What are the enabling technologies used in cloud computing? [2]

Q.10. Explain the layers and service models used in cloud computing. [2]

PART-B [5×4=20]

(Analytical/Problem solving questions)

Attempt any five questions

Q.1. What is the role of Data Center in the Cloud Computing? What are the components required in data center architecture? [4]

Q.2. Explain the concept of Kernel based virtual machine used in Cloud Computing. [4]

Q.3. What is Network support in cloud computing? [4]

Q.4. What are the design principles and policy implementation used for securing the cloud? [4]

Q.5. Explain Google App Engine. [4]

Q.6. What are the challenges faced in cloud computing? [4]

Q.7. Write short notes on the following : [4]

(a) Protein Structure Prediction

(b) Data Analysis

PART-C [3×10=30]

(Descriptive/Analytical/Problem Solving/Design questions)

Attempt any three questions

Q.1. (a) What is the role of virtualization in Cloud Computing? What is hyper visor virtual machine? What are the characteristics of virtualization? [5]

(b) Write a short note on virtualization of a : [5]

(i) Server

(ii) Network

Q.2. (a) Write a short note on : [6]

(i) PasS

(ii) IaaS

(iii) Hadoop

(b) Explain the architectural design of compute and storage cloud. What is the role of Map Reduce in Hadoop? [4]

Q.3. (a) Explain Microsoft Azure Design in detail. [5]

(b) What is the role of Aneka in the integration of Public and Private Clouds? [5]

Q.4. (a) What are the challenges, Risk and approaches during migration of data into cloud? [5]

(b) Evaluate the cloud's business impart and economics. [5]

Q.5. (a) What challenges we are facing in the security of cloud computing? Explain the architecture of cloud computing security. [5]

(b) Explain Service Level Agreement (SLA)with the help of diagram and also explain how the trust management can be done in cloud computing security. [5]

----- × -----

6E7136

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7136

B.Tech. VI-Sem. (Main/Back) Exam., May–2025

Computer Science and Engineering

6CS5-11/-Distributed System (El-I)

CS,IT

Time : 3 Hours **Maximum Marks : 70**

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting material is permitted during examination. (Mentioned in Form No. 205)

1. ..... 2. .....

PART-A **[10×2=20]**

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. What is distributed system?

Q.2. What are the key features of distributed system?

Q.3. Differentiate between the RPC and RMI.

Q.4. Define Load Balancing in distributed system.

Q.5. Differentiate between the process and thread.

Q.6. List two advantages of dynamic load sharing.

Q.7. What is distributed shared memory?

Q.8. How does distributed shared memory improve system performance?

Q.9. What is concept of Faults?

Q.10. What are the key issues in managing replicated data?

PART-B **[5×4=20]**

(Analytical/Problem Solving Questions)

Attempt any five questions

Q.1. Discuss different types of operating systems used in distributed system.

Q.2. Explain the role of distributed computing environment.

Q.3. Explain Inter-Process Communication Mechanism.

Q.4. What is distributed deadlock handling? Explain it.

Q.5. Explain the distributed process implementation.

Q.6. Explain the concept of memory consistency models. Why are they important?

Q.7. Discuss the randomized distributed agreement approach and its advantages.

PART-C [3×10=30]

(Descriptive/Analytical/Problem-Solving/Design Questions)

Attempt any three questions

Q.1. Explain the Architecture Models. How design issue affect the system performance?

Q.2. Explain the language mechanisms for synchronization.

Q.3. Discuss the design and implementation of a distributed file system.

Q.4. Explain the implementation of distributed shared memory systems.

Q.5. What is the Byzantine Agreement? Explain different approaches to solve it.

----- × -----

6E7107

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7107

B.Tech. VI-Sem. (Main/Back) Exam. - 2025

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

6AID5-11 Artificial Neural Network

AID, CAI, CDS

Time : 3 Hours **Maximum Marks : 70**

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting materials is permitted during examination. (Mentioned in Form No. 205)

1. ..... 2. .....

PART-A **[10x2=20]**

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. What is the model of a neuron in a neural network?

Q.2. What is Boltzmann Learning?

Q.3. What are linear least square filters in neural networks?

Q.4. What is the relation between the Perceptron and Bayes classifier for a Gaussian Environment?

Q.5. What is Cross-validation in neural network training?

Q.6. What are the limitations of Back Propagation learning?

Q.7. What is the SOM algorithm?

Q.8. Explain learning vector quantization.

Q.9. What are Neurodynamical models?

Q.10. Explain Hopfied model.

PART-B **[5x4=20]**

(Analytical/Problem solving questions)

Attempt any five questions

Q.1. What is a Neural Network?

Q.2. Explain Hebbian learning in Neural Networks.

Q.3. Explain Perceptron convergence theorem.

Q.4. How does the XOR problem demonstrate the need for multilayer perceptrons?

Q.5. Explain Hessian matrix in neural networks.

Q.6. Define supervised learning in neural networks.

Q.7. What are the two basic feature mapping models in Neural Networks?

PART-C [3x10=30]

(Descriptive/Analytical/Problem Solving/Design questions)

Attempt any three questions

Q.1. What is memory based learning, and how it differs from other learning techniques?

Q.2. What are Learning Rate Annealing Techniques? How do they improve perceptron learning?

Q.3. How does pruning help optimize deep neural networks for deployment?

Q.4. Compare Self-Organizing Maps (SOM) with other clustering techniques like K-means.

Q.5. Explain the concept of dynamical systems in the context of neural networks and describe their significance in modeling neural activity.

----- × -----

6E7108

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E7108

B.Tech. VI-Sem. (Main/Back) Exam. - 2025

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

6AID5-12 / Natural Language Processing (NLP)

AID, CAI

Time : 3 Hours **Maximum Marks : 70**

Instructions to Candidates :

Attempt all ten questions from Part-A, five questions out of seven questions from Part-B and three questions out of five questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting material is permitted during examination.

(Mentioned in Form No. 205)

1. .....NIL..... 2. .....

PART-A **[10x2=20]**

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. Define Natural Language Processing (NLP).

Q.2. List three common applications of NLP.

Q.3. What is the ArgMax function used for in NLP?

Q.4. Define syntactic collocations.

Q.5. What is term weighting, and why is it important in text processing?

Q.6. What is WordNet?

Q.7. Define Keyphrase Extraction.

Q.8. What is the purpose of PCFGs in NLP?

Q.9. What is Text Clustering?

Q.10. Explain the concept of Sequence Labeling.

PART-B **[5x4=20]**

(Analytical/Problem solving questions)

Attempt any five questions

Q.1. Define the term Machine Learning. Also explain the relationship between Machine Learning and NLP.

Q.2. What is WordNet Similarity? Explain Resnick's work on WordNet similarity measures.

Q.3. What is an IPython Notebook? Discuss the advantages of using IPython Notebooks for NLP projects.

Q.4. What is top-down parsing? Explain the role of parsing algorithms in syntactic analysis.

Q.5. Develop a method of automatically expand user queries semantically related terms from WordNet.

Q.6. What is a Probabilistic Context-Free Grammar (PCFG)? Discuss the training issues associated with PCFGs.

Q.7. Define Graphical Models. Also explain the role of graphical models in sequence labeling.

PART-C [3x10=30]

(Descriptive/Analytical/Problem Solving/Design questions)

Attempt any three questions

Q.1. Describe the role of Part-of-Speech (POS) tagging in syntactic analysis. Analyze the impact of POS tagging accuracy on the performance of downstream NLP tasks.

Q.2. Describe the importance of tokenization in text preprocessing. Implement a tokenizer that can handle punctuation, contractions, and special characters in a given text.

Q.3. Develop a script to generate basic statistics (e.g., word count, sentence count, vocabulary size) for a give text collection.

Q.4. Develop a script to automatically extract keyphrases from a given mystery text and rank them by relevance.

Q.5. Design a system that integrates WordNet with a machine learning model to improve WSD performance.

----- × -----

6E1552

Total No. of Questions : 22 Total No. of Pages : 04

Roll No. : .....

6E1552

B.Tech. VI-Sem. (Back) Exam. - 2025

COMPUTER SCIENCE AND ENGINEERING

6CS4-02 / Machine Learning

CS, IT

Time : 3 Hours **Maximum Marks : 120**

Attempt all ten questions from Part A, five question out of seven from Part B and four questions out of five from Part C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used/ calculated must be stated clearly.

Use of following supporting material is permitted during examination. (Mentioned in form No.205)

1. ..... 2. .....

PART-A [10x2=20]

(Answer should be given up to 25 words only)
All questions are compulsory

Note : All questions are **compulsory** :

1. State the assumption behind Linear regression and its mathematical form.
2. What is the Gini index in a decision tree ?
3. What is the role of the Sigmoid function in Logistic regression ?

4. What is the role of the support, confidence and life metrics in Association Rule Mining ?

5. Differentiate between feature extraction and feature selection.

6. What is a Markov Decision Process (MDP) ?

7. Differentiate between Q-learning and SARSA algorithms.

8. What is the purpose of Policy Iteration ?

9. How does the Apriori algorithm generate frequent item sets ?

10. What is the FP-Growth algorithm ?

PART-B [5x8=40]

(Analytical/Problem solving questions)

Attempt any Five questions

Q.1. In a Spam detection model, the probability of an email being spam is 30%, and the likelihood of a keyword appearing in spam emails is 0.8, but 0.1 in non-spam. Use Bayes' theorem to calculate the probability that an email is spam if the keyword is present.

Q.2. Give these three data points : (2, 2), (4, 4), (6, 6) and initial centroids (2, 2) and (6, 6), perform one iteration of the K-means algorithm.

Q.3. KNN classifier (k = 3) predicts the class of a point (2, 2) using the following labeled data :

X	Y	Class
1	1	A
2	3	B
3	3	A
3	2	B

What is the predicted class ?

Q.4. Give the data set :

X	Y
1	2
2	3
3	4

Compute the first principal component.

Q.5. A model gives the following confusion matrix :

	Predicted Positive	Predicted Negative
Actual Positive	40	10
Actual Negative	20	30

Calculate precision, recall and F1-score.

Q.6. How SARSA differs from Q-learning in terms of exploration and exploitation ? Explain.

Q.7. Explain the steps of backpropagation and compute the gradient for a single hidden layer neural network.

PART-C [4x15=60]

(Descriptive/Analytical/Problem Solving/Design questions)
Attempt any four questions

Q.1. Given a logistic regression model with the decision boundary :

$$P\left(Y=\frac{1}{X}\right)=\frac{1}{1+e^{-(2X_1+3X_2-4)}}$$

Determine whether the point $(X_1=1, X_2=0)$ will be classified as 1 or 0.

Q.2. Given the following dataset, calculate the information gain if the root node splits on "Weather" :

Weather	Temperature	Play ?
Sunny	Hot	No
Overcast	Mild	Yes
Rainy	Cool	Yes
Sunny	Cool	No

Q.3. Solve the Bellman equation for the following policy :

$$V(s) - R(s) + \gamma \max_{\alpha} Q(s,a)$$

Given :

(i) $R(s) = 5$

(ii) $\gamma = 0.9$

(iii) $Q(s,a) = [3,4]$

Q.4. Given the following transaction :

Transaction ID	Item
1	{A, B, C}
2	{A, B}
3	{B, C}
4	{A, C}

Find frequent itemsets using Apriori algorithm with min-supoort = 2.

Q.5. Design a Convolutional Neural Network (CNN) architecture for image classication with at least two convolutional layers and a softmax output.

-------X-------

6E1555

Total No. of Questions : 14 Total No. of Pages : 04

Roll No. : .....

6E1555

B.Tech. VI-Sem. (Back) Exam. - 2025

COMPUTER SCIENCE AND ENGINEERING

6CS4-05 Artificial Intelligence

CS, IT

Time : 2 Hours **Maximum Marks : 80**

Instructions to Candidates :

Attempt all five questions from Part-A, four questions out of six questions from Part-B and two questions out of three questions from Part-C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used / calculated must be stated clearly.

Use of following supporting material is permitted during examination.

(Mentioned in Form No: 205)

1. ..... 2. .....

PART-A [5x2=10]

(Answer should be given up to 25 words only)

All questions are compulsory

Q.1. If a multi-agent system has 'n' agents, each of which has 'm' possible moves, the search space increases to how many possible states?

Q.2. Define the terms 'syntax' and semantics' for First Order Logic.

Q.3. What is heuristic function?

Q.4. What is an expert system?

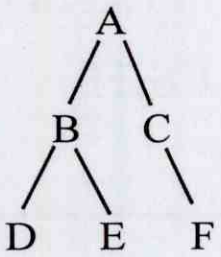
Q.5. What are membership functions with respect to neural networks?

PART-B [4x10=40]

(Analytical/Problem Solving Questions)

Attempt any four questions

Q.1. Explain the working of DFS and BFS algorithms. Apply DFS and BFS to the following graph from node A and show the traversal order for both algorithms.



Q.2 Enumerate classical 'Water jug problem". Describe the state space for this problem. Solve this problem by giving its operation sequence.

Q.3. A hospital uses a Bayesian network to diagnose whether a patient has lung cancer (C) based on two factors

Smoking (S) - whether the patient smokes

Coughing (X) - whether the patient has a cough.

The contidional probabilities are given as follows :

P(S) = 0.3 (30% of people smoke)

P(C/S) = 0.05, P(C/-s) = 0.01 (Lung cancer is more likely in smokers)

P(X/C) = 0.9. (PX | -C) =0.2 (Coughing is more likely in people with lung cancer)

Given that a patient is coughing (X = True). Calculate the probability that the patient has lung cancer (C = True) using Baye's Theorem.

Q.4. (a) Define supervised and unsupervised learning and explain this key differences with appropriate real world examples.

(b) Consider a dataset countaining information about customer purchases.Explain how would you apply supervised learning and unsupervised learning to analyse customer behaviour.

Q 5. (a) Describe the main steps involved in Natural Language Processing (NLP)

(b) Discuss at least four key challenges or issues faced in NLP.

Q.6. Explain how probabilistic reasoning and fuzzy logic approaches can handle uncertainty in AI.

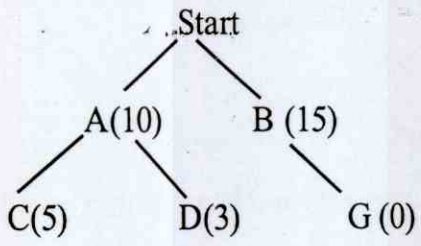
PART-C [2x15=30]

(Descriptive/Analytical/Problem Solving/Design Question)

Attempt any two questions

Q1. (a) Compare and contrast A* and AO* search algorithms based on evaluation functions, heuristic functions, optimality, comleteness, advantages and disadvantages.

(b) Consider the following road network with distances(g) and estimated time (h) :



the heuristic values (h) are given in brackets

The edge costs(g) are :

Start → A = 5

Start → B = 6

$A \rightarrow C = 4$

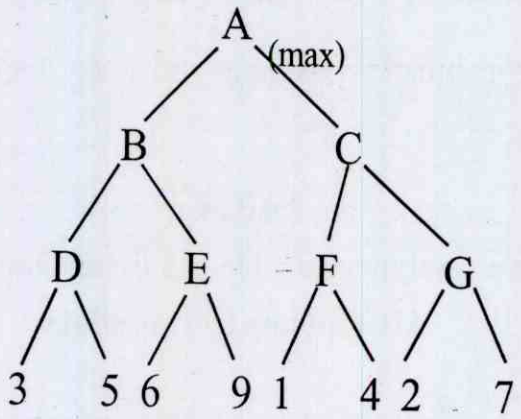
$A \rightarrow D = 2$

$B \rightarrow G = 7$

Using A* search, final the optimal route to G

Q.2 (a) Describe the Mini-Max and Alpha-Beta pruning techniques in detail.

(b) Consider the following game tree with given evaluation values at leaf nodes:



Use the Minimax algorithm to determine the optimal more for the maximizing player at the root A. Apply Alpha-Beta pruning to show which nodes can be pruned and reduce the number of evaluations.

Q.3 Consider the foliowing sentences :

(a) John likes all kind of food.

(b) Apples are food

(c) Chicken is food

(d) Anything anyone eats and is not killed by is food.

(e) Bill eats peanuts and is still alive

(f) She eats everything Bill eats.

Convert the following English statements into First-order Logic expressions and prove that 'John like peanuts by resolution.

----- × -----