13/7/25

Course Title: Artificial Intelligence and Neural Networks
Department of Computer Science and Telecommunication Engineering
Year 3, Term 2, Final Examination (May 2025), Session 2020-21
Course Code: **CSTE 3211**, Full Marks: 70, Time: Four hours
*Answer any **SEVEN** of the following questions.*

1.a) Match each of the following AI problems to one or more suitable AI techniques or paradigms from the list below for solving them. A technique may be used for more than one problem. Justify each of your matches in 1–2 lines. 5

Problems:	Techniques/Paradigms:
Language translation	Search algorithms
Robot path planning in a maze	Reinforcement learning
Diagnosing a disease from symptoms	Neural networks
Playing a board game (e.g., Go)	Expert systems
Identifying objects in an image	Probabilistic reasoning.

b) Define Heuristic. Explain how heuristics improve search efficiency. 2
c) Describe the A* search algorithm. Prove that it is optimal when the heuristic is admissible. 3

2.a) Explain how Genetic Algorithm works using the MAXONE problem. 4
b) Explain how GA can be adapted for function optimization problems. 2
c) We have some probabilities of cancer tests results from empirical data. The test procedure is not 100% perfect, thus we get some incorrect results. 4
Suppose the following probabilities are given:

- P(Test+|Cancer) = 0.9
- P(Test-|Cancer) = 0.1 [False negative]
- P(Test+|No cancer) = 0.001 [False positive]
- P(Test-|No cancer) = 0.999
- P(Cancer) = 0.0001 [Cancer happens in 1 in 10000 people]

Apply Bayes' Rule and answer the following questions:

I. If a patient is tested positive for cancer, what is the actual probability that they have cancer?
II. If the patient now goes through a second test and tests positive again, what is probability that they really have cancer?

3.a) What is the "Naive" assumption in the Naive Bayes classifier? Why is this assumption often unrealistic, and yet why does the algorithm still work well in practice? 2
b) Explain Gaussian Naive Bayes. What type of data does it handle, and how is the class-conditional probability computed for continuous features? 3
c) Explain how a Decision Tree classifier selects the best attribute to split the data at each node. In your answer: 5

- Describe the role of a splitting criterion such as Information Gain or Gini Index
- Discuss why pure splits are desirable
- Mention one potential issue that can arise when growing a very deep tree and how it can be mitigated

4.a) Distinguish between the breath-first search and uniform-cost search. 3
b) Describe the Iterative deepening depth-first search strategies on a binary tree with example. Consider a tree, where b=10 and depth d=5, calculate the total number of nodes generated by a breadth-first search and an iterative deepening depth-first search. 4+3

b^d b^d b^d

13/7/25

5.a) What is Gradient Descent? Use a linear regression example to explain how the algorithm updates weights to minimize the loss. 5

b) Define the following loss functions and explain where each is appropriate: 3
   i. Hinge Loss
   ii. Huber Loss
   iii. Binary Cross Entropy

c) A binary classifier has the following confusion matrix: 2
   TP = 90, FP = 10, FN = 20, TN = 80
   Calculate Precision, Recall, Accuracy, and F1-score.

6.a) Why should we use Batch Normalization? Explain. 3

b) How does L1/L2 regularization affect a neural network? 3

c) What might happen if you set the momentum hyperparameter too close to 1 (e.g., 0.9999) when using an SGD optimizer? Explain. 4

7.a) When it comes to training an artificial neural network, what could be the reason why the loss doesn't decrease in a few epochs? 5

b) Assume that you are using a deep neural network for a prediction task. After training your model, you notice that it is strongly overfitting the training set and that the performance on the test isn't good. What can you do to reduce overfitting? 5

8.a) What is a Multi-Layer Perceptron (MLP)? How is it different from a simple perceptron? Draw the structure of an MLP. 3

b) Explain the architecture of a Convolutional Neural Network (CNN). Describe the function of each layer: 5
   i. Convolution
   ii. Pooling
   iii. Dropout
   iv. Fully Connected

c) What are Hyperparameters? List three examples and describe how tuning them affects model performance. 2

9.a) What is a Recurrent Neural Network (RNN)? What challenges does it face in practice? 3

b) Describe the structure and working of an LSTM cell. How is it different from a GRU? 4

c) What is the Attention Mechanism in Transformers? Explain one key advantage Transformers have over RNNs. 3