Sri Lanka Institute of Information Technology

B.Sc. Honours Degree in Information Technology

Specialized in Data Science

Final Examination
Year 3, Semester 1/2 (2024)

IT3071 – Machine Learning and Optimization Methods

Duration: 2 Hours

October 2024

Instructions to Candidates:

- This paper has 4 questions.
- Answer all questions in the booklet given.
- The total marks for the paper is 100.
- This paper contains 7 pages, including the cover page.
- This paper is preceded by a 10-minute reading period. The supervisor will indicate when answering may commence.
- Calculators are allowed.
- Electronic devices capable of storing and retrieving text are not allowed.

Question 1   **(25 Marks)**

a) Briefly explain the parts of the biological neuron that were considered in developing an artificial neuron.   **(2 marks)**

b) A tech startup is developing a simple binary classification system to determine whether an email is spam or not spam using a **perceptron** model. The features extracted from the email are:

- Number of words (X_1)
- Number of links (X_2)
- Presence of certain keywords (X_3)
- Length of the subject line (X_4)

Given the following:

- Input feature vector: $X = [15,3,1,10]^T$
- Initial weight vector: $W = [0.5,-0.2,0.8,-0.5]^T$
- Bias term: $b = 0.1$
- Activation function: $\phi(x)=\text{sgn}(x)$, where $\text{sgn}(x)$ is the sign function that returns $+1$ for $x \geq 0$ and -1 for $x<0$

  i. Calculate the net input (weighted sum plus bias) to the perceptron. Show your calculations step-by-step.   **(4 marks)**

  ii. Determine the output of the perceptron using the activation function. Show your calculations.   **(3 marks)**

  iii. If the actual label of the email is "spam" (which we consider as $+1$), and the predicted output is different, update the weights using a learning rate of 0.01. Show your calculations step-by-step.   **(3 marks)**

c) An agent is using Q-learning to navigate a simple environment with three states: **A**, **B**, and **C**. The agent starts at **A** and can move between states by taking actions. The Q-values for the state-action pairs are initialized as shown below:

State	Action (Move)	Q-Value
A	B	5.0
A	C	3.0
B	A	2.0
B	C	4.0
C	A	1.0
C	B	0.0

The agent starts in **State A** and takes the following steps:

1. The agent moves from **A** to **B** and receives a reward of **2**.

2. The agent moves from **B** to **C** and receives a reward of **4**.

3. The agent moves from **C** to **A** and receives a reward of **1**.

Using a learning rate $\alpha=0.1$ and discount factor $\gamma=0.9$, calculate the updated Q-values for the state-action pairs **Q(A, B)**, **Q(B, C)**, and **Q(C, A)** after these steps.

(6 marks)

d) State 3 differences between Reinforcement learning and Unsupervised learning.

(3 marks)

e) What are 'support vectors' in the SVM classifier? Why are they used in the classification process?
(2 marks)

f) What is the pre-processing technique used in SVM to make non-linearly separable data linearly separable?

(2 marks)

Question 2 (25 Marks)

You are a machine learning engineer at a tech company that is developing an autonomous driving system. Your task is to build a Convolutional Neural Network (CNN) to detect and classify traffic signs from camera images captured by the car. The goal is to classify traffic signs into one of the 43 categories as defined by the German Traffic Sign Recognition Benchmark (GTSRB).

Scenario:

Data Description:
The dataset contains 50,000 images of traffic signs, each labeled with one of the 43 categories. The images are of various sizes and need preprocessing.
Image Preprocessing:
Each image must be resized to 32x32 pixels and normalized.
Model Architecture:
You propose the following CNN architecture:
- Input layer
- First convolutional layer with 32 filters, kernel size 5x5, ReLU activation, followed by a max pooling layer with pool size 2x2
- Second convolutional layer with 64 filters, kernel size 5x5, ReLU activation, followed by a max pooling layer with pool size 2x2
- Third convolutional layer with 128 filters, kernel size 3x3, ReLU activation, followed by a max pooling layer with pool size 2x2
- Flatten layer
- Fully connected layer with 512 units, ReLU activation
- Dropout layer with a dropout rate of 0.5
- Output layer with 43 units, softmax activation

Training: The dataset is split into 70% training, 20% validation, and 10% test sets. The model is trained using categorical cross-entropy loss and the Adam optimizer over 20 epochs with a batch size of 64.

a) Explain the significance of each layer in the proposed CNN architecture and how they contribute to traffic sign classification. **(4 marks)**

b) Discuss the importance of preprocessing steps such as resizing and normalization of images in the context of training the CNN. **(2 marks)**

c) How would you evaluate the performance of your CNN model on the test set? Include metrics you would use and why they are important. **(3 marks)**

d) Describe potential strategies to further improve the model's performance if the initial results are not satisfactory. **(3 marks)**

e) Consider a convolutional layer in a CNN with the following parameters:

Input volume size: 32x32x3 (height x width x depth)
Number of filters: 16
Filter size: 3x3
Stride: 1
Padding: 'same'

    i. Calculate the dimensions of the output volume produced by this convolutional layer (height x weight x depth).

(3 marks)

    ii. How many parameters (weights and biases) does this convolutional layer have? Show your calculations

(3 marks)

f) A pharmaceutical company is developing an image recognition system to categorize various drug types using images.

    i. If the development team is planning to use a traditional (non-Deep Learning) algorithm such as SVM to enhance the accuracy and to limit the size of the dataset necessary, state whether manual feature engineering is necessary or not for training this model.

(1 mark)

    ii. Briefly justify your answer using 1-2 sentences.

(2 marks)

    iii. Considering the same scenario, explain one advantage and one disadvantage of using an SVM for this image recognition task.   **(2 marks)**

    iv. Suppose the company decides to use a Deep Learning approach with a Convolutional Neural Network (CNN) instead. State whether manual feature engineering is necessary or not for training this model and justify your answer in 1-2 sentences.   **(2 marks)**

Question 3 (25 Marks)

a) Name and describe the main components of Genetic Algorithm.

(5 Marks)

b) Explain adequately the two requirements a problem should satisfy in order to qualify for solving it by a GA

(4 Marks)

c) A foreign traveller has requested a Sri Lankan local tour operator to propose a tour plan satisfying his requirements. During his stay, he wants to visit the following cities: Anuradhapura, Sigiriya, Dambulla, Polonnaruwa, Kandy. The task is to find the **shortest** route through several cities such that each city is visited only once and in the end return to the starting city. The task is to use GA to solve this problem.

    i.   What is the size of the population if the starting city is fixed?   (3 Marks)

    ii.   Randomly select the initial population of four tour plans assuming Badulla as the starting city   (2 Marks)

    iv.   Describe a fitness function suitable for measuring the goodness of every tour plan of the population and find out the fitness values for the four tours selected above.

You may assume that the distances between cities in kilometers as given below:

(4 Marks)

	Anuradhapura	Badulla	Galle	Trincomalee	Kandy
Anuradhapura	0	163	253	95	115
Badulla	163	0	140	176	57
Galle	253	140	0	302	147
Trincomalee	95	176	302	0	156
Kandy	115	57	147	156	0

    v.   Explain why one-point crossover operation cannot be performed to generate the new population

(3 Marks)

    vi.   Suppose that you have applied a crossover operation and generate a new population chromosome. Suggest and elaborate a suitable mutation operation to create the second generation of chromosome.

(2 Marks)

    vii.   Explain and describe the termination of this GA   (2 Marks)

Question 4 (25 Marks)

a) A company produces two products, A and B. Each product requires the use of two resources, labor and capital. The company has 100 hours of labor and 50 units of capital available. Product A requires 2 hours of labor and 1 unit of capital, while product B requires 1 hour of labor and 2 units of capital. The company wants to maximize the profit from the sale of these two products.

    i. What are the decision variables in this problem? **(1 marks)**
   ii. What is the objective function which is to maximize profit? **(3 marks)**
  iii. Write two constraints in this problem. **(4 marks)**

b) A company operates two factories (Factory A and Factory B) that produce two products: P1 and P2. Each product requires processing at two stages: machining and assembly. The company wants to maximize profit while optimizing the use of labor, machinery, and transportation between factories.

Key Details:

- **Factory A** can produce a maximum of 60 units of either P1 or P2 combined, while **Factory B** can produce a maximum of 80 units of either product combined.
- Both factories share the same resource limitations: 150 labor hours and 100 machine hours at each factory.
- **P1 at Factory A** requires 3 labor hours and 2 machine hours per unit.
- **P2 at Factory A** requires 2 labor hours and 4 machine hours per unit.
- **P1 at Factory B** requires 4 labor hours and 3 machine hours per unit.
- **P2 at Factory B** requires 3 labor hours and 2 machine hours per unit.

Additionally, the products must be transported to a central distribution warehouse after production, and the transportation cost per unit is:

- From Factory A to the warehouse: $5 for P1 and $7 for P2.
- From Factory B to the warehouse: $6 for P1 and $4 for P2.

Profits per unit (after production and transportation costs):

- P1 yields a profit of $50 per unit.
- P2 yields a profit of $60 per unit.

The company also has the following demand constraints:

- The demand for P1 is at least 40 units.
- The demand for P2 cannot exceed 70 units.

    i. What are the decision variables in this problem? **(2 marks)**
   ii. Formulate the objective function for this problem. **(6 marks)**
 iii. Write the constraints for this problem. **(9 marks)**

End of Paper