Nova Southeastern University
College of Computing and Engineering

Assignment 2

CISC 670 Artificial Intelligence

Fall 2022

Due date: 11/16/2022 11:59PM ET

Total points: 100

Note: Please include your name and the "Certification of Authorship" (located on Canvas) form in EVERY document you submit. Thanks.

Part 1.1. Text Reading:

Genetic Algorithm (Chap. 4 Sec 4.1)
Neural Networks (Chap. 21, for introduction of deep learning)
Support Vector Machine (Chap. 19, Sec 19.7.5, 19.7.6)

Part 1.2. Additional Reading (optional):

Neural Networks:

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, 29(3), 31-44.
(downloadable version available at http://metalab.uniten.edu.my/~abdrahim/mitm613/Jain1996_ANN%20-%20A%20Tutorial.pdf)
(*Note: To better understand the back-propagation algorithm, this is a better reference than the textbook.*)

Aggarwal, C. C. (2018). An introduction to neural networks. In *Neural Networks and Deep Learning* (pp. 1-52). Springer, Cham.
(downloadable version available at https://link.springer.com/chapter/10.1007/978-3-319-94463-0_1)

Genetic Algorithm:

Deb, K. (1999). An introduction to genetic algorithms. *Sadhana*, 24(4-5), 293-315.
(downloadable version available at https://link.springer.com/content/pdf/10.1007/BF02823145.pdf)
https://en.wikipedia.org/wiki/Genetic_algorithm

Part 2. Problems:

(**Note**: Please include any external materials other than the textbook. Use the APA format where appropriate.)

2.1 Genetic Algorithm [15 points]

The following algorithm is used to implement crossover in a genetic algorithm:
Input: Two strings of n bits x and y
Output: Two strings of n bits x' and y'
The crossover operator is applied as follows:
A crossover site is selected at random (with equal probability) that divides each string into two sub-strings of *non-zero* length. That is x = [x_1 x_2] y = [y_1 y_2], with length of x_1 = length of y_1.
The outputs are generated as x' = [x_1 y_2] and y' = [y_1 x_2]
Given that you start with (x1, y1) = ((1 0 1 0 1 0) (1 1 1 1 1 1)), specify which 6-bit strings are possible values obtained through crossover alone. Justify your answer.

2.2 Genetic Algorithm [15 points]

A genetic algorithm uses the following mutation operator: the bits in the input string are considered one by one independently, with probability 0.01 that each bit is inverted. Given that you apply the mutate operator to the string (1 1 1 1), what is the probability that the output is: (0 0 0 0)? (0 1 0 0)? (1 0 1 0)? (1 1 1 1)? Show the process of your computation.

2.3 Neural Network [50 points]

The data set in the file "data.txt" contains 300 observations for 4 input variables (Temp, Pres, Flow, and Process) and an output variable (Rejects). The first column "No" is simply an identifier. The table below reproduces the first 4 observations:

No	Temp	Pres	Flow	Process	Rejects
1	53.39	10.52	4.82	0	1.88
2	46.23	15.13	5.31	0	2.13
3	42.85	18.79	3.59	0	2.66
4	53.09	18.33	3.67	0	2.03

Train a back-propagation neural network on approximately 80% of the observations, randomly selected. Test the trained network using the remaining 20% observations.

Please write a detailed report that includes the following.

1) A detailed discussion how you set up the tool (if used), tuning of the parameters, and performed the training and testing. If you implemented your own neural network, please also report the hardware/software/programming language/libraries you have used.

2) Answer: Will different parameters yield the same solutions based on your experiments? Please justify your choice on these parameters.

3) Present:
(i) A figure that plots the actual and predicted values of the output "Rejects" for the training and test data sets.
(ii) Sum of squared errors for the training and test data sets.

You should also show important intermediate results, and important steps of your experiments.

Note: The easiest way to solve this problem is to use a Neural Network tool. Do some research and find a tool that you feel comfortable with. If you wish to implement your own neural networks, that is also fine.

Problem 2.4 Support Vector Machine [20 points]

Consider the following set of training data.

x_1	x_2	class
1	1	+
2	2	+
3	0	+
3	1	+
0	0	−
1	0	−
0	1	−
-1	1	−

(i) Plot these eight training points in a two-dimensional space (with x_1 and x_2).
Are the classes {+, −} linearly separable? Why?

(ii) Construct the weight vector of the maximum margin hyperplane by inspection and identify the support vectors.

(iii) If you remove ***one*** of the support vectors, does the size of the optimal margin decrease, stay the same, or increase? Justify your answer.

(iv) Is your answer to (iii) still hold for any dataset in a 2-dimentioanl space? Provide an example if it is not true, or give a short proof if it is true. What if we have datasets in a space with more than two dimensions, do you have the same answer? Please justify.

Note: For additional reading on SVM, you may refer to the following URLs.
http://nlp.stanford.edu/IR-book/html/htmledition/support-vector-machines-the-linearly-separable-case-1.html
http://web.mit.edu/6.034/wwwbob/svm-notes-long-08.pdf