NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER 1 EXAMINATION 2018-2019

CE4042/CZ4042 – NEURAL NETWORKS AND DEEP LEARNING

Nov/Dec 2018 Time Allowed: 2hours

INSTRUCTIONS

1. This paper contains 4 questions and comprises 5 pages.

2. Answer **ALL** questions.

3. This is an open-book examination.

4. All questions carry equal marks.

---

1. (a) Given two inputs x and y, you are to train a neuron to approximate the following function ϕ when $0 \leq x, y \leq 1.0$.

$$\phi(x, y) = x + 2y^3 + xy - 0.5$$

(i) Briefly state how you generate training data.

(3 marks)

(ii) State how you design the inputs to a linear neuron.

(3 marks)

(iii) Write the activation function if a perceptron is used.

(4 marks)

Note: Question No. 1 continues on Page 2

(b) A softmax layer of three neurons receives 2-dimensional inputs $(x_1, x_2) \in \boldsymbol{R}^2$. The weight matrix $\boldsymbol{W}$ and bias vector $\boldsymbol{b}$ of the layer are given by

$$\boldsymbol{W} = \begin{pmatrix} 2.0 & 0.0 & 1.0 \\ -1.0 & 1.0 & -2.0 \end{pmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} 0.5 \\ 1.0 \\ -0.5 \end{pmatrix}.$$

(i) Find the decision boundaries separating each pair of classes.

(7 marks)

(ii) Plot the decision boundaries separating the three classes, clearly indicating the regions belonging to the classes.

(5 marks)

(iii) Find the output class label for an input pattern $\boldsymbol{x} = \begin{pmatrix} -0.5 \\ 1.0 \end{pmatrix}$.

(3 marks)

2. The three-layer feedforward network shown in Figure Q2 receives 2-dimensional inputs $(x_1, x_2) \in \boldsymbol{R}^2$ and produces an output $y \in \boldsymbol{R}$. Each hidden-layer has two perceptrons and the output neuron is a linear neuron. The weights and biases of the network are initialized as indicated in the figure.

The network is trained to produce a desired output $d = 1.0$ for an input $\boldsymbol{x} = \begin{pmatrix} -1.0 \\ 0.5 \end{pmatrix}$ by using gradient descent learning. The learning factor $\alpha = 0.8$.

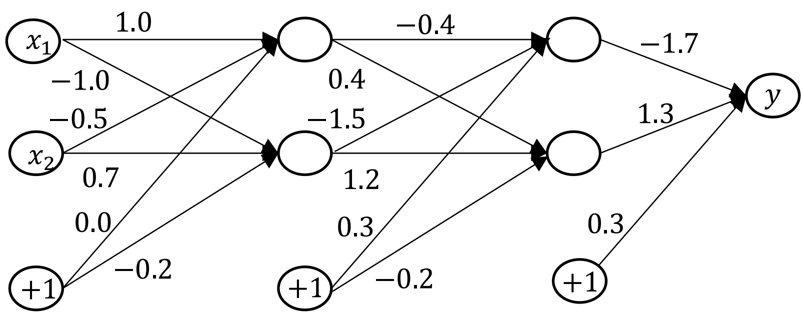


Figure Q2

Note: Question No. 2 continues on Page 3

For one iteration of stochastic gradient descent learning:

(a) Write the initial weight matrix $\boldsymbol{W}_1$ and bias vector $\boldsymbol{b}_1$ of the first hidden layer, the initial weight matrix $\boldsymbol{W}_2$ and bias vector $\boldsymbol{b}_2$ of the second hidden layer, and the initial weight vector $\boldsymbol{w}$ and bias b of the output neuron.

(3 marks)

(b) Find the synaptic input $\boldsymbol{u}_1$ and the activation $\boldsymbol{h}_1$ of the first hidden layer, the synaptic input $\boldsymbol{u}_2$ and the activation $\boldsymbol{h}_2$ of the second hidden layer, and the activation y of the output neuron.

(5 marks)

(c) Find the square error cost $J = \frac{1}{2}(d - y)^2$.

(1 mark)

(d) Find the gradients $\nabla_{\boldsymbol{u}_1} J$, $\nabla_{\boldsymbol{u}_2} J$, and $\nabla_y J$ of cost J with respect to $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, and y, respectively.

(8 marks)

(e) Find the gradients $\nabla_{\boldsymbol{W}_2} J$ and $\nabla_{\boldsymbol{b}_2} J$ of cost J with respect to the weight matrix $\boldsymbol{W}_2$ and the bias vector $\boldsymbol{b}_2$ of the second hidden layer, respectively.

(4 marks)

(f) Find the updated weight matrix $\boldsymbol{W}_2$ and bias vector $\boldsymbol{b}_2$ of the second hidden layer.

(4 marks)

3. (a) An input image $\boldsymbol{X}$ is processed by a convolution layer and thereafter by a pooling layer. The convolution layer has filters with weights $\boldsymbol{w} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and a bias $b = 0.2$, and consists of neurons with *sigmoid* activation function. The convolution is performed at strides = $[1, 1]$ and with 'VALID' padding. The pooling layer performs max pooling and uses a 2×2 pooling window at strides = $[2, 2]$ and with 'SAME' padding.

Given an input image $\boldsymbol{X} = \begin{pmatrix} 0.4 & -0.1 & 0.2 & -0.3 \\ 0.7 & 0.1 & -0.3 & 0.4 \\ -1.5 & 0.2 & 0.0 & -0.3 \end{pmatrix}$, find the feature maps at

(i) the convolution layer and

(7 marks)

(ii) the pooling layer.

(4 marks)

(b) A recurrent neural network (RNN) with top-down recurrence receives 2-dimensional inputs and produces 2-dimensional hidden layer activations and 1-dimensional outputs. The hidden layer neurons have *tanh* activation functions and the output layer neurons have *sigmoid* activation functions.

The weight matrix $\boldsymbol{U}$ from the input layer to the hidden layer, the weight matrix $\boldsymbol{V}$ to the output layer, and the top-down recurrence weight matrix $\boldsymbol{W}$ are given by

$$\boldsymbol{U} = \begin{pmatrix} -1.0 & 0.5 \\ 0.5 & 0.3 \end{pmatrix}, \boldsymbol{V} = \begin{pmatrix} 2.0 \\ -1.5 \end{pmatrix} \text{ and } \boldsymbol{W} = (-2.0 \quad 1.5).$$

The hidden layer bias vector $\boldsymbol{b}$ and the output layer bias c are given by

$$\boldsymbol{b} = \begin{pmatrix} 2.0 \\ 0.3 \end{pmatrix} \text{ and } c = 0.4.$$

The output layer is initialized to an output of 1.0.

Determine the output sequence of the RNN for an input sequence of $\big(\boldsymbol{x}(1), \boldsymbol{x}(2), \boldsymbol{x}(3)\big)$ when

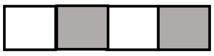
$$\boldsymbol{x}(1) = \begin{pmatrix} -1.0 \\ 2.0 \end{pmatrix}, \boldsymbol{x}(2) = \begin{pmatrix} 1.0 \\ -1.0 \end{pmatrix} \text{ and } \boldsymbol{x}(3) = \begin{pmatrix} 0.0 \\ 3.0 \end{pmatrix}.$$

(14 marks)

4. (a) An autoencoder has four neurons at the input layer and two neurons at the hidden layer. All the neurons have *sigmoid* activation functions. The weight matrix $\boldsymbol{W}$ of the hidden layer, the bias vector $\boldsymbol{b}$ of the hidden layer and the bias vector $\boldsymbol{c}$ of the output layer are given by

$$\boldsymbol{W} = \begin{pmatrix} 0.8 & 0.4 \\ -0.4 & -0.8 \\ 0.2 & 0.4 \\ -0.7 & 0.2 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} 0.0 \\ 0.2 \end{pmatrix} \text{ and } \boldsymbol{c} = \begin{pmatrix} 0.0 \\ -0.6 \\ 0.8 \\ 0.1 \end{pmatrix}.$$

Consider the following two input patterns applied to the autoencoder:

$\boldsymbol{s}_1$  $\boldsymbol{s}_2$ 

(i) Convert each input pattern to their respective vector representations by using the following notation: shaded box = 0 and white box = 1.

(2 marks)

(ii) Find the hidden layer activations and the outputs of the autoencoder.

(7 marks)

(iii) Find the entropy at the output layer.

(4 marks)

(iv) Find the Kullback-Leibler (KL) divergence of hidden layer activations with respect to a constant neuron activation $\rho = 0.1$.

(4 marks)

(b) Describe how the adversarial process between the generator and discriminator networks is implemented in the training of Generative Adversarial Networks (GAN).

(8 marks)

END OF PAPER