T6

Wednesday, June 16, 2021 4:22 PM

Tutorial6

STATS5099: Data Mining

Tutorial Sheet 6

Conceptual

1. What does an artificial neuron compute?
   (a) A neuron computes an activation function followed by a linear function ($\alpha + \sum_j w_j x_{ij}$).
   (b) A neuron computes a linear function ($\alpha + \sum_j w_j x_{ij}$) followed by an activation function.
   (c) A neuron computes a function g that scales the input x linearly ($\alpha + w_i x_i j$).
   (d) A neuron computes the mean of all features before applying the output to an activation function.

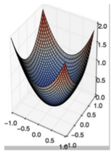
2. Linear and logistic regression models can be seen as neural networks. Which of the following statements are correct? Select all that apply.
   (a) The difference between the two models is the activation function of the output node and nothing else.
   (b) The difference between the two models is the activation function of the output node and the loss function.
   (c) The difference between the two models is the activation functions of the hidden layer and the output node, and the loss function.
   (d) The difference between the two models is the loss function and nothing else.

3. Which of the following statements are correct? Select all that apply.
   (a) A neural network will always outperform a regression model.
   (b) A neural network with multiple hidden layers will always outperform a neural network with one hidden layer.
   (c) As the number of hidden layers increase, model capacity increases.
   (d) As the number of hidden layers increase, model capacity increases.

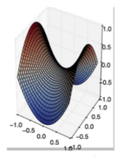
4. What are the steps for using a gradient descent algorithm?
   3 Calculate error between the actual value and the predicted value
   4 Reiterate until you find the best weights of network
   2 Pass an input through the network and get values from output layer
   1 Initialise random weight and bias
   5 Go to each neuron which contributes to the error and change its respective values to reduce the error

   (a) 2, 3, 5, 4, 1
   (b) 1, 3, 2, 5, 4
   (c) 3, 2, 1, 5, 4
   (d) 4, 3, 1, 5, 2

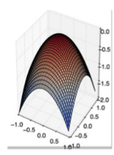
5. First Order Gradient descent would not work correctly (i.e. may get stuck) in which of the following graphs?



(a)



(b)



(c)

6. Which of the following statements about backpropagation are correct? Select all that apply.
   (a) Backpropagation is the transmission of error back through the network to adjust the inputs.
   (b) Backpropagation is the transmission of error back through the network to allow weights to be adjusted so that the network can learn.
   (c) Suppose you are training a neural network using backpropagation. Depending on your random initialisation, your algorithm may converge to different local optima.
   (d) When training a neural network using backpropagation, it is preferable to initialise all the weights to be zero.

7. Below is a diagram if a single artificial neuron (unit):

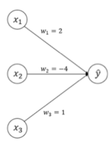


Figure 2: Single unit with three inputs.

The node has three inputs $\boldsymbol{x} = (x_1, x_2, x_3)$ and the weights corresponding to each input are given by w_1, w_2, w_3 respectively. The activation of the output unit (y) is given by the step-function:

$$\phi(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Calculate what will be the output value $\hat{y}$ for each of the following input patterns:

Pattern	P_1	P_2	P_3	P_4
x_1	1	0	1	1
x_2	0	1	0	1
x_3	0	1	1	1

P_1: $t = 2\times1 - 4\times0 + 1\times0 = 2 > 0$, $y = \phi(t) = 1$

P_2: $t = 2\times0 - 4\times1 + 1\times1 = -3 < 0$, $y = \phi(t) = 0$

P_3: $t = 2\times1 - 4\times0 + 1\times1 = 3 > 0$, $y = \phi(t) = 1$

P_4: $t = 2\times1 - 4\times1 + 1\times1 = -1 < 0$, $y = \phi(t) = 0$