CSEN240 Machine Learning
Homework #8

(Level 0: Please watch the guest lecture and answer the following questions)

1. (4 points) What are the common data quality issues?
    a. Duplication records
    b. Inconsistent data formatting
    c. Missing or incomplete data
    d. Reliable data sources
2. (3 points) What best defines data quality?
    a. The speed at which data is processed
    b. The amount of data stored
    c. The accuracy, completeness, and consistency of data
    d. The security measures in place to protect data
3. (3 points) What is the purpose of data profiling in data quality assessment?
    a. To identify missing or incomplete data
    b. To encrypt sensitive data for security purposes
    c. To determine the historical trends of data usage
    d. To automate data entry processes of efficiency
4. (3 points) What are examples of bad data governance practices?
    a. Allowing unrestricted access to data for all employees
    b. Storing data in multiple unconnected databases
    c. Updating data quality policies only once a year

(Level 1: high-level understanding)

5. (6 points) Which of the following statements are true about artificial neural networks and backpropagation? (Select all that apply)
    a. Artificial neural networks are inspired by the structure and function of the human brain.
    b. Activation functions introduce non-linearity to the network and help in modeling complex relationships.
    c. MLP is a universal classifier. Any input-to-output mapping can be modeled by a network of sufficient size and appropriate weights
    d. Backpropagation is a common algorithm used to adjust the weights and biases in a neural network during training.
    e. Backpropagation calculates the gradient of the loss function with respect to the network's parameters.
    f. The chain rule is a fundamental concept used in backpropagation to compute gradients layer by layer.

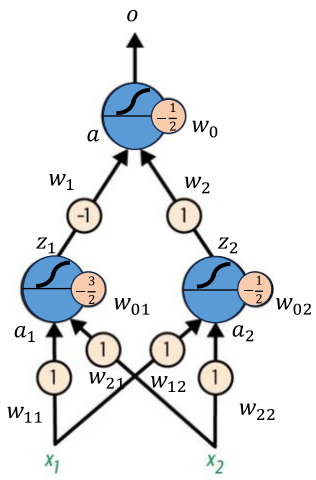
6. (5 points) Which of the following statements are true about convolutional neural networks? (Select all that apply)
    a. CNNs typically consist of convolutional layers, pooling layers, and fully connected layers.
    b. Convolution layers are responsible for extracting local features from the input data.
    c. A convolutional layer simultaneously applies multiple filters to its inputs, making it capable of detecting multiple features in its inputs
    d. Pooling is typically applied after convolutional layers to progressively reduce the spatial size.
    e. Skip connections are a common architectural element in CNNs, enabling better gradient flow and aiding in training deep networks.

(Level 2: manual exercise)

7. (10 points) CNN
    a. 1D conv
        - Given 1D input data: {1, 2, 1, 1, 3, 0}
        - Given 1D 3-point filter: {-1, 2, -1}
        - Zero-pad the input data by 2 points on either side
        - Use stride 2 and compute the filter output
    b. 2-point max pooling
        - Use the result from 1D conv
        - Perform max pooling

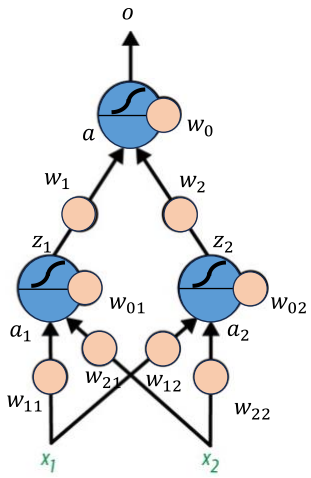
8. (26 points) MLP
    a. A MLP is shown on the right with the current weight

    b. Forward pass
        - Please use (x_1=1, x_2=1) as the input to compute the output
        - Using sigmoid function as the activation function
    c. Backward pass
        - The target output is 0
        - Please
            - Calculate the error using MSE
            - Backpropagate the error
            - Update the weights (using learning rate = 0.1)



(Level 3: computer-based exercise)

9. (20 points) MLP (Using HW8-9.ipynb)
    a. Implement an XOR function using MLP
        - XOR(0,0)=0, XOR(0,1)=1, XOR(1,0)=1, XOR(1,1)=0
    b. Use the MLP architecture on the right



    c. Use the sigmoid function as the activation function
    d. Initialize the weights $\in (-0.1,0.1)$
    e. Use mean squared error as the objective function
    f. Use learning rate = 0.1
    g. Use stochastic gradient descent.
        - For each new epoch, permutate the training examples
    h. Run for 100,000 epochs or until error is less than 0.25
    i. Report your learned function and the final weights

10. (20 points) CNN (Using HW8-10.ipynb)
    a. Implement a CNN for MNIST (hand-written digit recognition)
    b. Using keras build-in library to construct a CNN model, which consists of the following (in the order of the forward pass)
        - 8 3x3 kernel conv layer
        - 2x2 Max pooling layer
        - 16 3x3 kernel conv layer
        - 2x2 Max pooling layer
        - Flatten 2D features maps into a 1D vector
        - A dropout layer with 0.5 dropout probability
        - A fully connected layer with softmax activation for classification
    c. Report the test accuracy