ME 6103 Engineering Optimization
Spring 2025, Dr. Carolyn Conner Seepersad
Assignment 5: Neural Networks
Due: On Canvas by Wed, April 16, 2025
NOTE: This assignment is optional. If you complete it, the grade will be averaged into your overall assignment grade. If not, we will base your assignment grade on the 4 assignments already completed.

1. Consider the neural network below. Assume a linear activation function for all of the nodes in the hidden layer and the output layer.

Input Layer   Hidden Layer   Output Layer
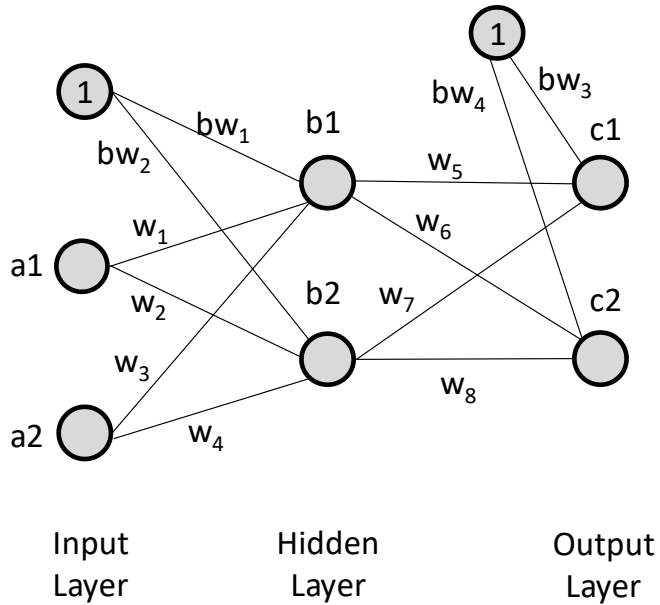
(a) Write an expression for the output of nodes b_1, b_2, and c_2, (i.e., $b_{1\text{-}out}$, $b_{2\text{-}out}$, $c_{2\text{-}out}$) as a function of the other relevant components of the network.

$$b_{1\,out} = w_1 a_{1\,out} + w_3 a_{2\,out} + 1 \cdot b w_1$$

$$b_{2\,oot} = w_2 a_{1\,out} + w_4 a_{2\,out} + 1 \cdot b w_2$$

$$c_{2\,ovt} = 1 \cdot b w_4 + b_{1\,out} w_6 + b_{2\,out} w_8$$

(b) Assume a MSE error function, and write an expression for the partial derivative of the MSE of the output of c_2 with respect to w_1.

$$\frac{\partial E}{\partial z} = 2(z-t) = \frac{\partial E}{\partial c_{2\,out}}$$

$$\frac{\partial E}{\partial w_1} = \frac{\partial E}{\partial z} \cdot \frac{\partial z}{\partial net\, c_2} \underbrace{\frac{\partial net\, c_2}{\partial out\, b_1}}_{} \underbrace{\frac{\partial out\, b_1}{\partial net\, b_1}}_{} \underbrace{\frac{\partial net\, b_1}{\partial w_1}}_{}$$

$$2(z-t)(1) \cdot w_b \quad \cdot (1) \qquad out\, a_1$$

$$= 2(z-t) w_b \cdot out\, a_1$$

2. In class, we are discussing the basics of neural networks. However, there are many advanced categories of neural networks that have been developed for specific applications. Popular examples include convolutional neural networks and generative adversarial networks. Pick one of these networks (or an alternative neural network formulation of your choice) and describe it as if you were explaining it to your classmates. Answer the following questions:
   a. What types of applications are well-suited to the network?
   b. What features make it unique?
   c. Generally, how does it work?
   d. What are the pros and cons of using it?
Include at least 2 references / citations to back up your discussion.

    One of the advanced categories of neural networks convolutional neural networks (CNNs). CNNs are used extensively in grid-structured data such as images, videos, facial recognition, time series analysis. It can also be applied to dynamic PID tuning based on sensor data for better system responsiveness rather than static PID gain parameters (Lecture Slides, Xie).

    What makes a CNN distinct as a neural network implementation is it containing a convolution layer along with a pooling layer which allows for feature extraction and makes high dimensional data more extensible for a fully connected neural network which otherwise would be unable to process high dimensional data. A convolution layer is a mapping of data to a subset of that data by using a filter or a kernel. This allows for the capture of spatial information which is essentially the useful most critical information of the dataset rather than the full dataset. By adding in a pooling layer the dimensionality of the data can further be decreased.

    The basic method of using a convolutional neural network is first starting with some grid structured input data. From this via convolution and pooling layers the critical features are extracted from the input while decreasing its dimensionality in each step. Finally these subsets are flattened and serve as the input to a fully connected neural network. In a fully connected neural network the flattened subsets act as the inputs in layer 1. Then in the hidden layers the inputs are connected to a certain number of bias nodes, weights, and internal nodes. The outputs from the hidden layer have additional weights and bias nodes that they are connected in the output layer before reaching the output node (Lecture Notes, CNN Explainer, Lecture Slides).

    Some pros of using a CNN is their high accuracy and robustness to noise especially in image and video processing. Hence they can be very predictive and applicable for those applications compared to traditional machine learning methods. Some disadvantages of CNNs are that for things like image or video processing they can still be computationally expensive and require advanced hardware to effectively be trained and be deployed. Additionally, they only can be effective if they receive a high amount of good quality training data where despite robustness to noise this is subject to their initial training data's clarity (Tamanna).

Citations:

Handwritten Lecture Notes from NN_Applications

Lecture Slides on Neural Networks in Applications

"CNN Explainer." *Poloclub.github.io*, poloclub.github.io/cnn-explainer/.

Tamanna. "Exploring Convolutional Neural Networks: Architecture, Steps, Use Cases, and Pros and Cons." *Medium*, 24 Apr. 2023, medium.com/@tam.tamanna18/exploring-convolutional-neural-networks-architecture-steps-use-cases-and-pros-and-cons-b0d3b7d46c71.

Xie, Jiahang, et al. "PID-Based CNN-LSTM for Accuracy-Boosted Virtual Sensor in Battery Thermal Management System." *Applied Energy*, vol. 331, 1 Feb. 2023, pp. 120424–120424, https://doi.org/10.1016/j.apenergy.2022.120424.