

# Problem of the Week 8

Name: Ryan Rajaie

Student ID: 016105024

Date: 10/12/2025

## Motivation

This week's Problem of the Week is focused on comparing and contrasting the problems we addressed with our multi-layer perceptrons and our convolutional neural networks.

## Completing this assignment

Fill in the cell as indicated below. When you are happy with your response, produce a clean rendition of your notebook. Then, create a PDF of your notebook and upload both documents (PDF and ipynb file) to Canvas.

## Problem 8.1 - Model Similarities

Write 3-5 sentences explaining how the purpose of each model and the architecture of the neural networks are similar in the multilayer perceptron and CNN examples examined in class. Your response should pertain to the specific models we worked with in class, not generic statements about perceptrons and CNNs.

Both our CNN and the Perceptron models are supervised training neural networks designed to take in images and output labels, using a gradient algorithm to better label the given images. Both of our models feature layers that data passes through so that the model can better understand the features of the data. Both also use the same computational basics of matrix/tensor multiplications and gradient updates (even if our CNN model hides this behind a more simple PyTorch implementation).

## Problem 8.2 - Model Differences

Write 3-5 sentences explaining how the multilayer perceptron and CNN examples examined in class differ. Place specific emphasis on what makes the CNN approach more ideal for this application than a multi-layer perceptron. Your response should pertain to the specific models we worked with in class, not generic statements about perceptrons and CNNs.

Our models of course first differ in the dataset (the Perceptron model uses the MNIST number dataset while the CNN uses a dataset of unusual sea creatures...). They also differ in how they input the data. The perceptron model inputs our data with a 1 dimensional vector (by calling the reshape function on our datasets) while the CNN inputs our data with a 2D structure where we are even able to customize the kernel size and how large each of the layers are. However, in our MLP function (afaik) we can only customize starting parameters. This ability to understand the data in 2D makes the CNN far more effective at the job of learning images. Additionally, our ability to artificially "expand" our dataset with drop-outs and image transformations allows us to

further leverage the CNN's abilities to create a more accurate model without risking overfitting (which we do in part 2 to much success). Our MLP model also uses the MSE loss function while the CNN uses a Cross-Entropy function.