

## Section 9: Deep Learning

---

### 1. Run a Tensor Flow Algorithm

In section train a deep learning algorithm which can detect whether or not a handwritten digit is a 1 or a 0, using the MNIST handwritten digit dataset. Design a simple deep learning model. If you do this on your own outside of section, we suggest referencing the Python notebooks published on the class website.



Some questions that to think about:

- a. What is the difference between a test and a train set?
- b. How can you use log-likelihood to understand training?

### 2. Deep Dream

Deep dream is an algorithm that can imagine what an image would look like if it had certain objects in it (eg take Van Gough's Starry Night and imagine it with fish). It works by using a trained deep learning model to modify image pixels.



Assume that you have a fully trained network similar to the one from the previous question, except that it is trained to detect if 8x8 images ( $\mathbf{x}$ ) are pictures of cats ( $Y = 1$ ) or not-cats ( $Y = 0$ ). Your deep learning network has one hidden layer with 20 hidden neurons.

$$h_j = \sigma\left(\sum_{i=0}^{64} \theta_{i,j}^{(h)} \mathbf{x}_i\right) \qquad \hat{y} = \sigma\left(\sum_{i=0}^{20} \theta_i^{(y)} \mathbf{h}_i\right)$$

- a. How can you choose the image from your training dataset which maximally activates the final neuron ( $\hat{y}$ , the one that indicates the model believes it is looking at a cat)?
- b. How can you find the pixel values for a picture (likely not in your dataset) which maximally activates the second hidden neuron ( $h_2$ )? This is an involved question. Hint: You should not be changing the weights of your model, assume that they are fixed. Solve for any and all partial derivatives required by your answer.
- c. Given an image, how can you alter it slightly so that it is more catlike?