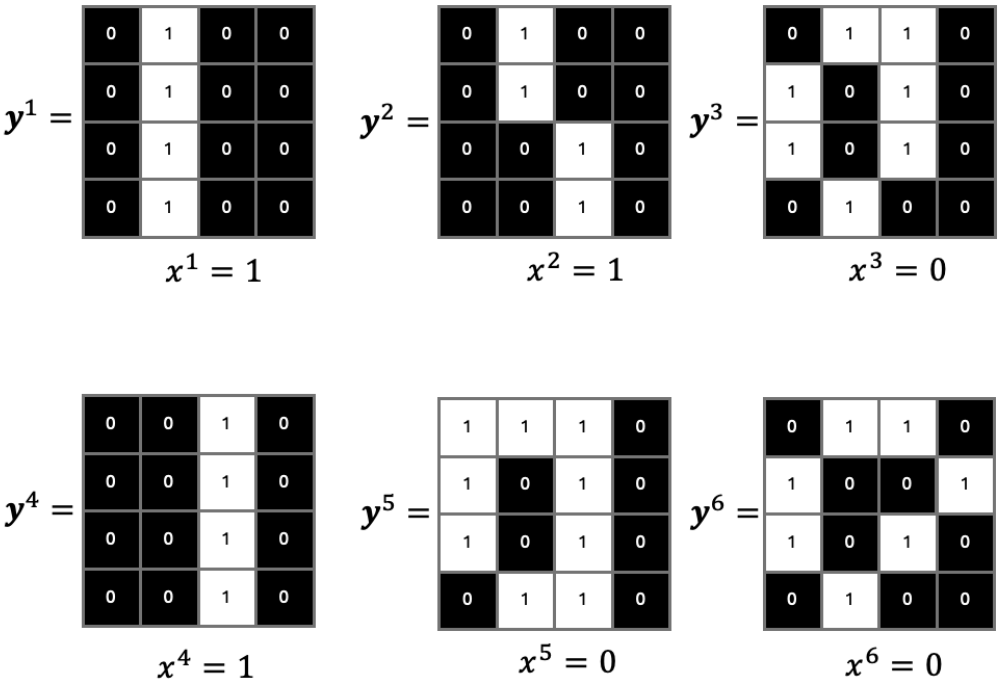
Exercise Sheet 7

Neural Computation
University of Birmingham

Week 7 2025

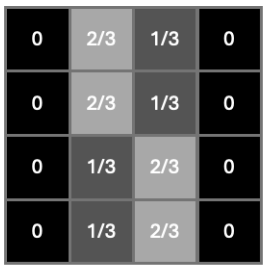
Exercise 1. You want to train a regression network to take a single digit number x as input and the produce a 4×4 pixel hand drawn image $\hat{\mathbf{y}}$ of that digit. You are using gradient descent and a mean square error loss function to achieve this; averaging the quadratic error between all pixels in the predicted image $\hat{\mathbf{y}}$ and the corresponding target image $\mathbf{y}$. Consider the following set of training pairs.



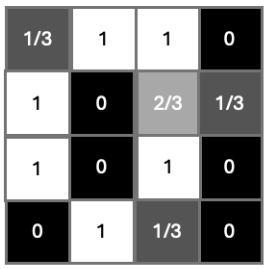
After training the network for a time, it is able to produce the optimal output (minimising the loss function). What are the outputs of the network for inputs $x = 1$ and $x = 0$ after the network has been trained. Draw the pixel grid and fill in numbers for each of the pixels. Do you see any problems with this approach?

Solution:

The best prediction for each digit x will be the average of the training images with with that digit, because the average minimises sum of squared errors. Note, this is the same property we looked at in Week 2, Exercise 1. In this case for each pixel, the network will find the average. The approach does not provide realistic outputs, because the averaged image looks very different from each of the training examples, including for example intermediate values like 1/3 and 2/3, which are not present in the dataset. A conditional generative model could solve this problem and produce more realistic results.



Output for $x = 1$



Output for $x = 0$

Exercise 2. You are training a GAN on a dataset to generate 256×256 pixel colour images. Assume you are using a batch size of 4. What would be the size of the input tensor and output tensor of the discriminator network?

Solution:

The discriminator takes the generated (and real) images as input. The tensor size can thus be derived from the image size. Each image in the dataset has dimensions 256×256 pixels and 3 colour channels (RGB), so a single image has shape $(3, 256, 256)$, i.e. 3 channels, 256 width, 256 height.

Given a batch size of 4, the input tensor to the discriminator will have shape:

$$\text{Input tensor shape} = (4, 3, 256, 256)$$

The discriminator ends with a single scalar output per image, representing the probability that the input is real. The output tensor would have the shape:

$$\text{Output tensor shape} = (4, 1)$$

This means the discriminator outputs one probability for each image in the batch.

Exercise 3. Which of the following statements about GANs are true? Explain your answers.

1. The discriminator is only required for training the generator.
2. The generator in a GAN is trained to maximise the discriminator's ability to distinguish real from fake data.
3. The discriminator in a GAN is trained to minimise the probability of correctly classifying real data as real.
4. When using back propagation to train the GAN, to compute the derivatives for the discriminator's parameters we first have to compute the derivatives for the generator's parameters.

Solution:

1. **True.** The discriminator is only used during training to provide feedback to the generator. Once training is complete, the discriminator is typically discarded, and only the generator is used to produce new samples.
2. **False.** The generator is trained to *fool* the discriminator, i.e., to generate data that the discriminator classifies as real. Therefore, it is trained to *minimise* the discriminator's ability to distinguish real from fake, not maximise it.
3. **False.** The discriminator is trained to *maximise* the probability of correctly classifying real data as real and fake data as fake. It is a binary classifier that aims to distinguish between real and generated samples.

4. **False.** When training the discriminator, we compute gradients with respect to its own parameters only, using both real and fake data. The generator's parameters are not involved in this step. The generator is updated separately, using gradients that flow through the discriminator.