# Handwriting Assignment HW 2

**110024516** 統研所 邱繼賢

## Exercise 5.3

*Discuss why an RBF network is a supervised variant of a nearest-neighbor classifier.*

An RBF (Radial Basis Function) network is a type of neural network that uses prototypes, or points from the training data, as the basis for classification. The network works by assigning weights to each prototype based on its similarity to the input data and the connection strength between the prototype and the output node, which are learned through a supervised training process. As such, an RBF network can be considered a supervised variant of a nearest-neighbor classifier, as it uses a similar approach of identifying the most similar prototype to the input data and using its label to make a prediction, but with additional capabilities for improving accuracy and making nuanced predictions.

## Exercise 7.10

*Consider a setting in which you have a large database of pairs of sentences in different languages. Although you have sufficient representation of each language, some pairs might not be well represented in the database. Show how you can use this training data to (i) create the same universal code for a particular sentence across all languages, and (ii) have the ability to translate even between pairs of languages not well represented in the database.*

To create a universal code for a particular sentence across all languages, we can train encoder-decoder pairs for each language using a large database of sentence pairs. The encoders will map the input sentences to a common latent space, while the decoders will read from this space and generate the translated output. By updating the weights of the encoder-decoder pairs during training, we can learn to create the same code for a particular sentence across all languages. To translate between pairs of languages not well represented in the database, we can use transfer learning techniques to fine-tune the

pre-trained encoder-decoder pairs on a smaller dataset of the specific language pairs we are interested in. This can help the model to adapt to the specific characteristics of the target language pairs and improve translation performance.

# Exercise 8.7

*Compute the convolution of the input volume in the upper-left corner of Figure 8.2 with the horizontal edge detection filter of Figure 8.1(b). Use a stride of 1 without padding.*

Input volume with size (7,7) :

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | 3 | 4 | 4 | 5 | 0 | 3 |
| 4 | 7 | 4 | 0 | 4 | 0 | 4 |
| 7 | 0 | 2 | 3 | 4 | 5 | 2 |
| 3 | 7 | 5 | 0 | 3 | 0 | 7 |
| 5 | 8 | 1 | 2 | 5 | 4 | 2 |
| 8 | 0 | 1 | 0 | 6 | 0 | 0 |
| 6 | 4 | 1 | 3 | 0 | 4 | 5 |

Filter with size (3,3) :

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 0 | 0 | 0 |
| -1 | -1 | -1 |

Because stride = 1, padding = 0, the size of the output obtained through convolution is (5,5), the calculation method is to use the numbers in the filter as weights, and do the weighted sum of the corresponding 9 numbers in the input volume, and then move the filter one unit and recalculate until the (3,3) filter is fully scanned through the (7,7) input volume. The result is as follow :

| | | | | |
|---|---|---|---|---|
| 4 | 6 | 4 | -3 | -3 |
| 0 | -1 | 0 | 1 | -2 |
| -5 | -6 | 1 | 1 | 0 |
| 6 | 11 | 1 | -3 | 4 |
| 3 | 3 | 4 | 4 | 2 |

# Exercise 9.2

*Throughout this chapter, a neural network, referred to as the policy network, has been used in order to implement the policy gradient. Discuss the importance of the choice of network architecture in different settings.*

A softmax regression is a simple machine learning model that is used to predict probabilities for a multiclass classification problem. It can be viewed as a special type of neural network with one layer, where the output layer consists of a set of sigmoid units that predict the probabilities of each class. In contrast, a neural network allows for greater flexibility and more complex representation learning through the use of additional layers and hidden units. This can be especially useful in situations where the inputs are not directly related to the outputs, as the neural network can learn more complex patterns and relationships between the input and output data. Overall, the advantage of using a neural network over softmax regression is that it provides a more powerful and flexible tool for learning complex relationships in the data, which can lead to improved prediction accuracy in different settings.