

Exercise 5

a) Looking @ the error surface graph, the descending gradient causes a backprop to find the global minimum.

b)

Exercise 6

→ As the number of training examples increases, the generalisation comes closer to the training set

→ From a different perspective, overfitting should be taken into consideration, hence the network should be tested.

Exercise 7

→ 10 networks

→ pros → that on unseen data → estimate

→ limited sample order to estimate how the model should perform.

Exercise 8

→ It can be seen that there are missing data, hence preprocessing is an important step.

→ stopping → reduce overfitting → improve generalisation