EFFECTS OF DIFFERENT NUMBER OF HIDDEN LAYERS

Network with one hidden layer

Network with two hidden layer

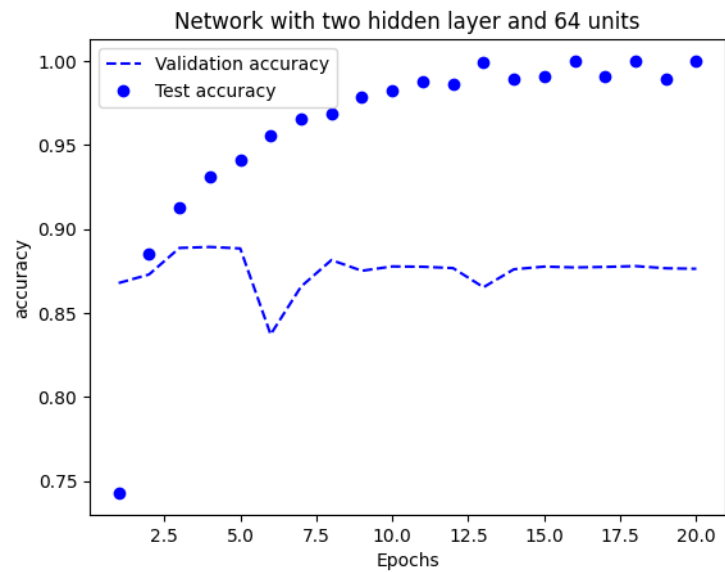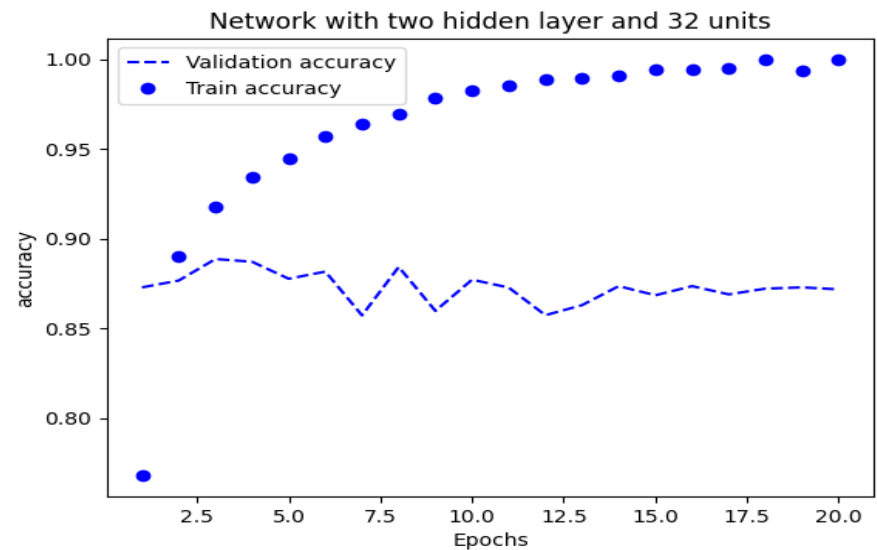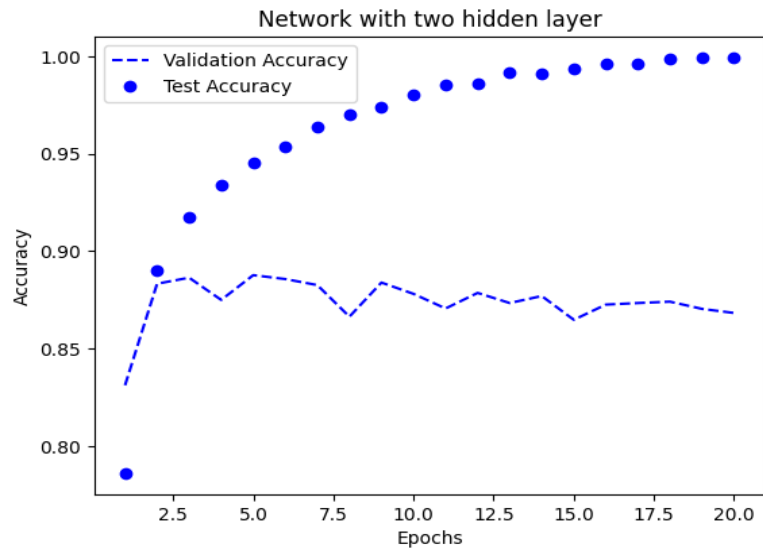Network with three hidden layer
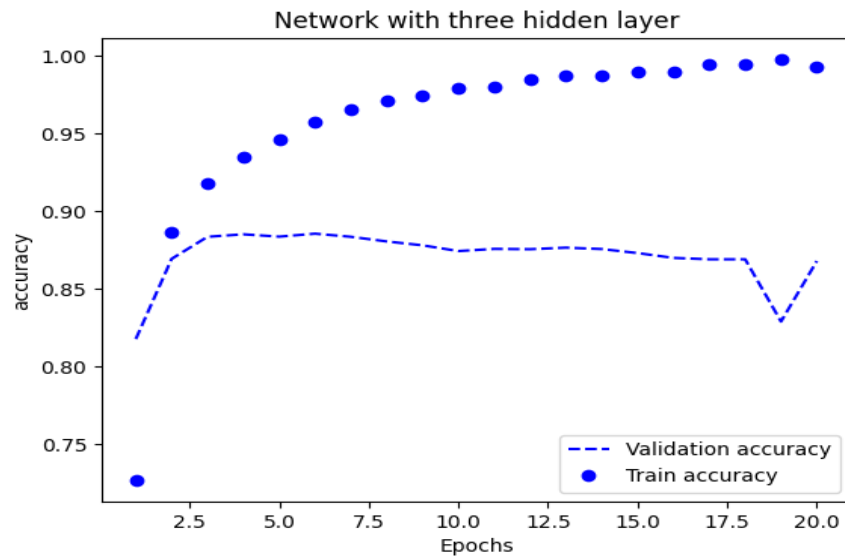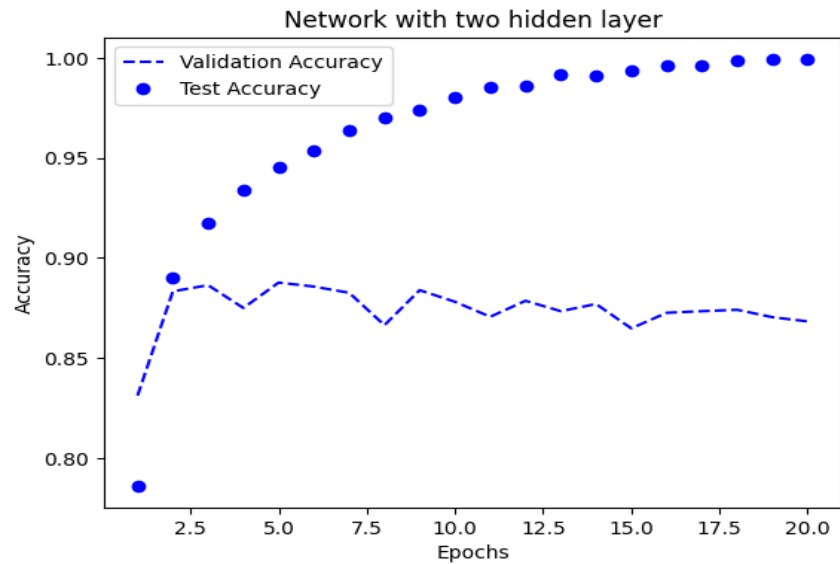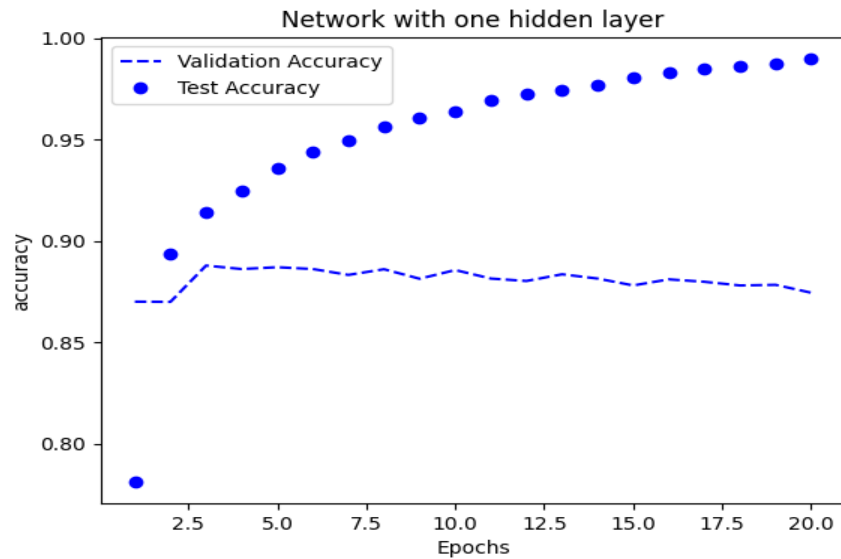
Models with more layers overfit faster than models with fewer ones.

- The training accuracy increases as the number of epochs increases, irrespective of the number of layers we use.

- Conversely, the validation accuracy first increases with training accuracy, then decreases. The decrease happens faster as we increase the number of layers.
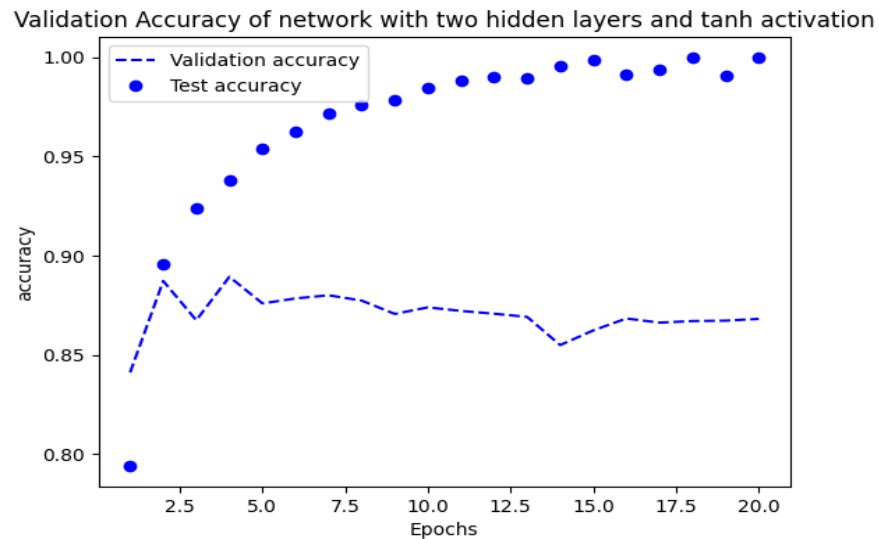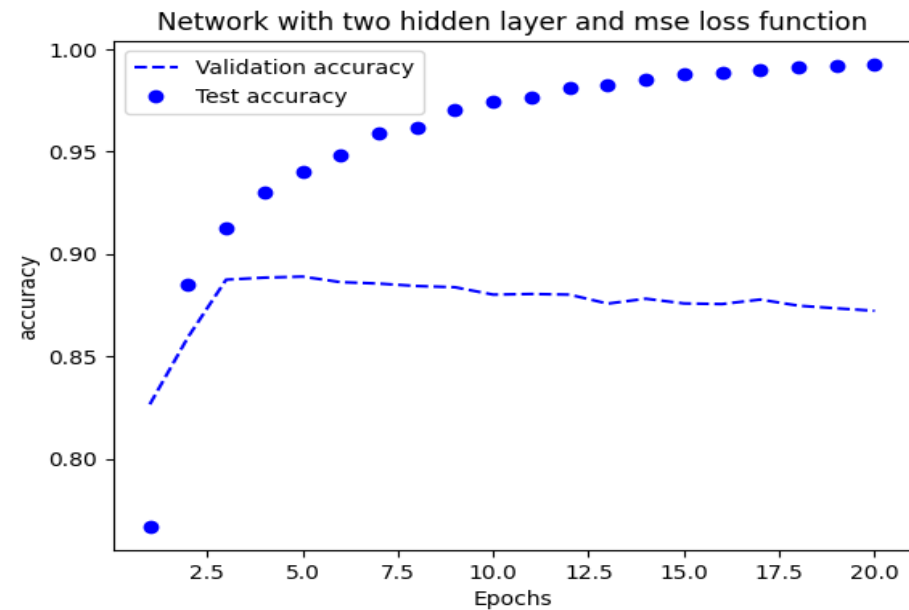
Effect of different number of units

Network with two hidden layer

Network with two hidden layer and 32 units

Network with two hidden layer and 64 units

Models with greater capacities overfit faster than models with lower ones.

- The training accuracy increases as the number of epochs increases, irrespective of the number of layers we use.

- Conversely, the validation accuracy first increases with training accuracy, then decreases. The decrease happens faster as we increase the capacity.
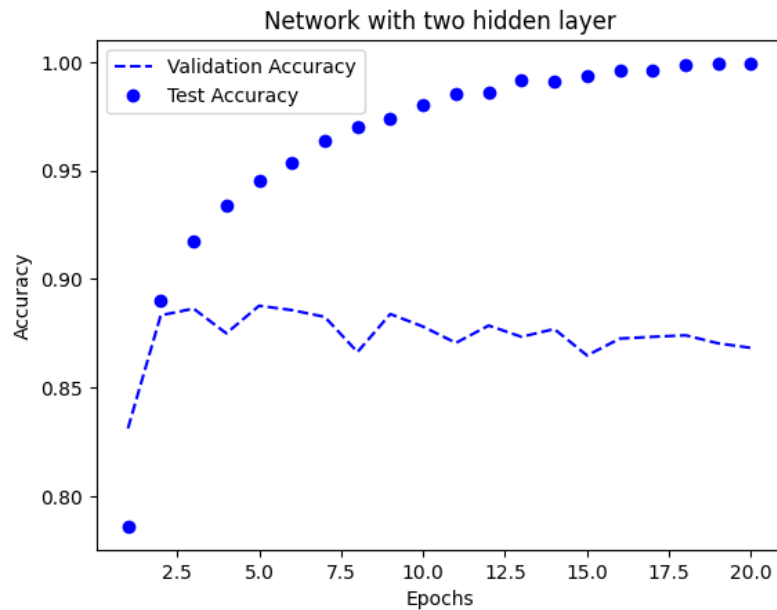
Effects of different loss and activation functions

Network with two hidden layer

Network with two hidden layer and mse loss function

Validation Accuracy of network with two hidden layers and tanh activation

- The training accuracy increases as the number of epochs increases, irrespective of the loss or activation function.
- **The relu activation function attains peak accuracy faster than the tanh activation function**
- **The mse loss function attains peak accuracy earlier than binary cross entropy function.**
This could be because mse does not penalize misclassification enough

The Effect of Dropout on Validation Loss
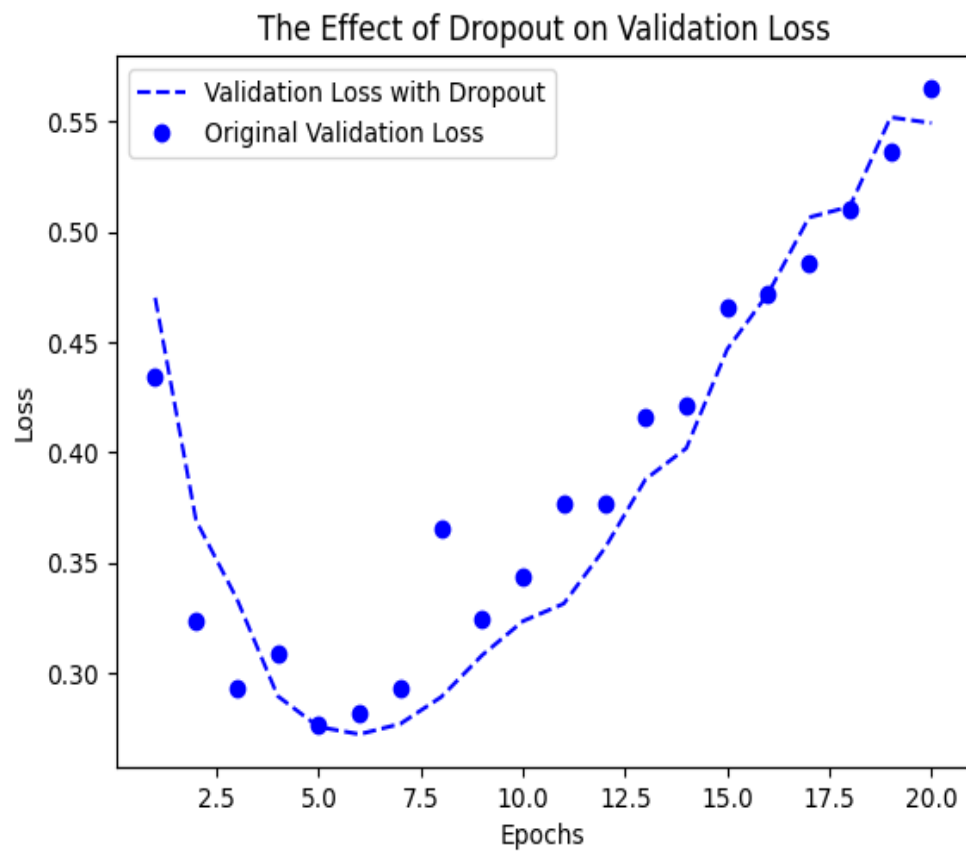
- The model with dropout seems to be an improvement since the lowest validation loss reached has improved when we perform the dropout and occurs around 6 epochs.