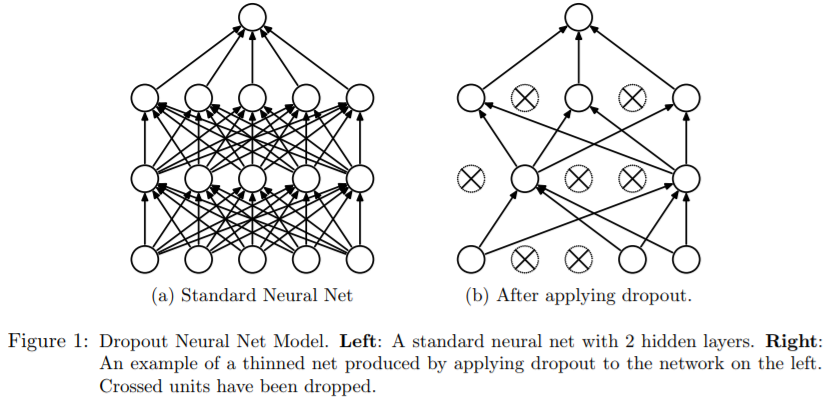
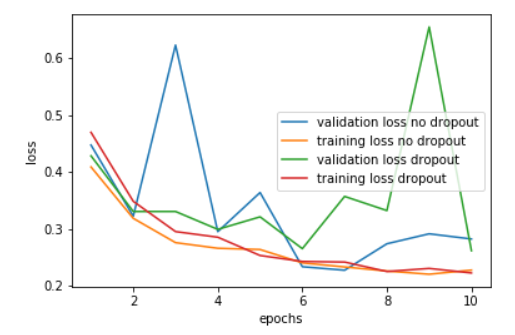
Exercise 4.3

*The model in transfer.py uses a dropout layer. How does dropout work and what is the effect of adding dropout layers the the network architecture? What is the observed effect when removing the dropout layer from this model?*

‘Dropout’ means that units are dropped out of the neural network. This means that this unit and all the connections are temporarily removed from the network. The network architecture is showed in figure 1. Which unit drops out is a random choice. What will be left is a thinned network. A neural network of n units equals a collection of 2n possible thinned networks. During training, dropout samples from the possible thinned network. At test time, a neural network is used without dropout with smaller weights. With this network the effect of averaging the predictions of all these thinned networks can be approximated. Dropout layers prevent overfitting and they provide an efficient way of connecting exponentially many different neural networks. It results in lower generalization errors of classification problems. [1]



Running the model with dropout gives a loss of 0.2224 and an accuracy of 0.9124. Runnig the model without dropout is done by removing the line *output = Dropout(0.5)(output).* This gives a loss of 0.2274 and an accuracy of 0.9129. This gives almost no difference. To check is there is overfitting, the validation losses are plotted. No direct conclusion can be drawn from this. It looks like there is overfitting when there is dropout

[1] Srivastava, N., Hinton, G., Krizhevsky, A., & Salakhutdinov, R. (2014). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*. *Journal of Machine Learning Research* (Vol. 15). Retrieved from <http://www.jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf>