13.3

(a)

We assume there are V_j units in each hidden layer from $j=1$ to $j=L$ and $V_0=N$, $V_{L+1}=1$, we consider a fully-connected feedforward neural network with L hidden layers.

The total number of trainable parameters Q in this NN can be calculated by:

$$Q = \sum_{j=0}^{L} (1+V_j) V_{j+1}$$

(2) if we rewrite it to

$$Q = NV_1 + (V_1 + \sum_{j=1}^{L} (1+V_j) V_{j+1})$$

This means the complexity of this model is

independent of the data size but depends on the architecture of VM