Figure 1: Representation of a multilayer perceptron for binary classification of L+1 layers with an input vector of dimension d and two output units where the hidden L-th layer contains $m^{(l)}$ neurons