Accuracy for the test and train datasets vs number of hidden layers