Accuracy during training process for different data sizes, activation functions, and number of layers

Experiment 4:
number of layers, data size
- 1, 20000
- 2, 20000
- 4, 20000
- 1, 50000
- 2, 50000
- 4, 50000
- 1, 100000
- 2, 100000
- 4, 100000

activation function
— RReLU
-- softmax