

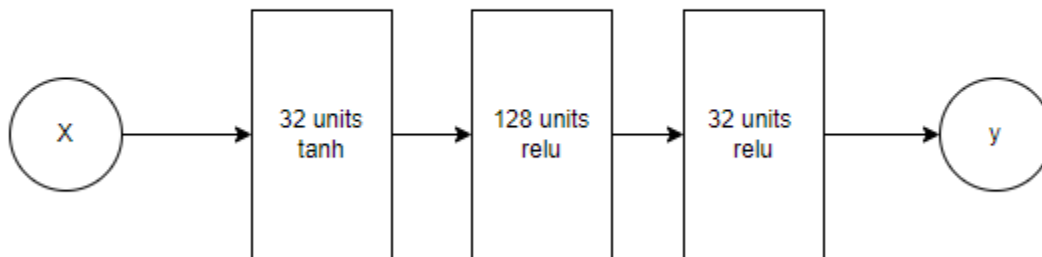
The best 3 architectures using adagrad are as follows

1. Test accuracy = 95.85

num_of_layers: 3

- Layer 0: 32
- Activation function 0: tanh
- Layer 1: 128
- Activation function 1: relu
- Layer 2: 32
- Activation function 2: relu
- Score: 0.9585416515668234

1. Test accuracy = 95.85 %

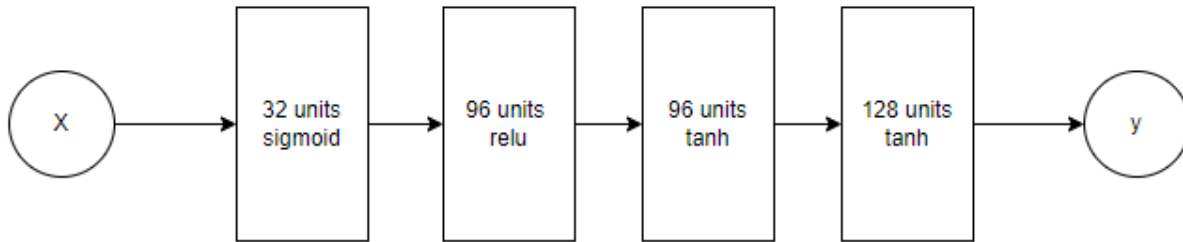


2. Test accuracy = 95 %

num_of_layers: 4

- Layer 0: 32
- Activation function 0: sigmoid
- Layer 1: 96
- Activation function 1: relu
- Layer 2: 96
- Activation function 2: tanh
- Layer 3: 128
- Activation function 3: tanh

2. Test accuracy = 95 %



3. Test Accuracy = 94.89 %

num_of_layers: 5

- Layer 0: 128
- Activation function 0: sigmoid
- Layer 1: 128
- Activation function 1: tanh
- Layer 2: 96
- Activation function 2: tanh
- Layer 3: 32
- Activation function 3: relu
- Layer 4: 32
- Activation function 4: relu
- Score: 0.9489236076672872

3. Test accuracy = 94.89 %

